Repeated Communication and Ramsey Graphs*

Noga Alon[†] Alon Orlitsky[‡]

Abstract

We study the savings afforded by repeated use in two zero-error communication problems. We show that for some random sources, communicating one instance requires arbitrarily-many bits, but communicating multiple instances requires roughly one bit per instance. We also exhibit sources where the number of bits required for a single instance is comparable to the source's size, but two instances require only a logarithmic number of additional bits. We relate this problem to that of communicating information over a channel. Known results imply that some channels can communicate exponentially more bits in two uses than they can in one use.

1 Introduction

Starting with graph definitions below, this section introduces the two coding problems, describes the results obtained, and relates them to known ones. The proofs are given in Sections 2 and 3. Section 4 outlines possible extensions.

A *graph* G consists of a set V of *vertices* and a collection E of *edges*, unordered pairs of distinct vertices. If $\{x, x'\} \in E$, we say that x and x' are *connected* in G. When E needs not be mentioned explicitly, we write $\{x, x'\} \in G$. An *independent* set in G is a collection of its vertices, no two connected. G's *independence number*, $\alpha(G)$, is the size of its largest independent set. A *coloring* of G is an assignment of colors to its vertices such that connected vertices are assigned different colors. G's *chromatic number*, $\chi(G)$, is the minimum number of colors in any of its colorings. The n-th *AND* (or *normal*) *power* of G is the graph $G^{\wedge n}$ whose vertex set is V^n and where distinct vertices $(x_1, \ldots, x_n)$ and $(x'_1, \ldots, x'_n)$ are connected if $\{x_i, x'_i\} \in G$ for all $i \in \{1, \ldots, n\}$ such that $x_i \neq x'_i$.

1.1 Channel coding

A *channel* consists of a finite *input set* $\mathcal{X}$, a (possibly infinite) *output set* $\mathcal{Y}$, and a nonempty *fan-out set* $S_x \subseteq \mathcal{Y}$ for every $x \in \mathcal{X}$. In each channel *use*, a *sender* transmits an *input* $x \in \mathcal{X}$ and a *receiver*

*Appeared, IEEE Transactions on Information Theory, 41:5 (September 1995), pp. 1276-1289. Printed on March 21, 1996.

[†]AT&T Bell Laboratories, 600 Mountain Ave., Murray Hill, NJ 07974, and Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, Israel.

[‡]AT&T Bell Laboratories, 600 Mountain Ave., Murray Hill, NJ 07974.

receives an arbitrary *output* in S_x. Following Shannon [19], we study the amount of information a channel can communicate without error.

Associated with a channel $\mathcal{C}$ is a *characteristic graph* $\mathcal{G}$. Its vertex set is $\mathcal{X}$ and two (distinct) vertices are connected if their fan-out sets intersect, namely, both can result in the same output. Note that every graph (V, E) is the characteristic graph of some channel: its input set is V, its output set is E, and S_v consists of all edges containing v.

The largest number of inputs $\mathcal{C}$ can communicate without error in a single use is $\alpha(\mathcal{G})$, the independence number of its characteristic graph. Intuitively, the sender and the receiver agree in advance on an independent set I. The sender transmits only inputs in I. Every received output belongs to the fan-out set of exactly one input in I, hence the receiver can correctly determine the transmitted input. Conversely, it is easy to see that a set containing two connected vertices cannot be communicated without error. The largest number of bits $\mathcal{C}$ can communicate without error in a single use is therefore

$$\gamma^{(1)} \stackrel{\text{def}}{=} \log \alpha(\mathcal{G}).$$

Note that this definition allows for a non-integral number of bits.

Example 1(a) In a *completely-noisy channel* every two fan-out sets intersect. $\mathcal{G}$ is the *complete graph* over $\mathcal{X}$ where every two vertices are connected, $\alpha(\mathcal{G}) = 1$, and $\gamma^{(1)} = 0$ indicating that in a single use no information can be communicated without error.

In a *noiseless channel* no two fan-out sets intersect. $\mathcal{G}$ is the *empty graph* over $\mathcal{X}$ where no two vertices are connected, $\alpha(\mathcal{G}) = |\mathcal{X}|$, and $\gamma^{(1)} = \log |\mathcal{X}|$.

In the *Pentagon channel* $\mathcal{X} = \mathcal{Y} = \mathcal{Z}_5 \stackrel{\text{def}}{=} \{0,\ldots,4\}$ and $S_i = \{i, i+1 \pmod 5\}$ for all $i \in \mathcal{Z}_5$. $\mathcal{G}$ is the *Pentagon graph* whose vertex set is $\mathcal{Z}_5$ and where vertex i is connected to vertices $i-1$ and $i+1 \pmod 5$. Clearly, $\{0, 2\}$ is a largest-size independent set, hence $\alpha(\mathcal{G}) = 2$ and $\gamma^{(1)} = 1$. □

When the channel $\mathcal{C}$ is used $n \in \mathcal{N}_2$ times[1], the sender transmits a sequence $x_1,\ldots,x_n$ of inputs and the receiver receives a sequence $y_1,\ldots,y_n$ of outputs where each $y_i \in S_{x_i}$. Conceptually, n uses of $\mathcal{C}$ can be viewed as a single use of a larger channel $\mathcal{C}^{(n)}$. Its input set is $\mathcal{X}^n$, its output set is $\mathcal{Y}^n$, and the fan-out set of $\overline{x} = (x_1,\ldots,x_n) \in \mathcal{X}^n$ is the Cartesian product $S_{\overline{x}} \stackrel{\text{def}}{=} S_{x_1} \times \ldots \times S_{x_n}$.

Let $\mathcal{G}^{(n)}$ denote the characteristic graph of $\mathcal{C}^{(n)}$. Its vertex set is $\mathcal{X}^n$, and if $\overline{x} = (x_1,\ldots,x_n)$ and $\overline{x}' \stackrel{\text{def}}{=} (x'_1,\ldots,x'_n)$ are distinct vertices then $\{\overline{x}, \overline{x}'\} \in \mathcal{G}^{(n)}$ iff $S_{\overline{x}}$ and $S_{\overline{x}'}$ intersect iff S_{x_i} intersects $S_{x'_i}$ for all $i \in \{1,\ldots,n\}$ iff $\{x_i, x'_i\} \in \mathcal{G}$ for all $i \in \{1,\ldots,n\}$ such that $x_i \neq x'_i$ iff $\{\overline{x}, \overline{x}'\} \in \mathcal{G}^{\wedge n}$, the n-th AND power of $\mathcal{G}$. Therefore,

$$\mathcal{G}^{(n)} = \mathcal{G}^{\wedge n}$$

It follows that the largest number of bits $\mathcal{C}$ can communicate without error in n uses is

$$\gamma^{(n)} \stackrel{\text{def}}{=} \log \alpha(\mathcal{G}^{\wedge n}).$$

If a set I is independent in G then $I \times I$ is independent in $G^{\wedge 2}$. Therefore, $\alpha(G^{\wedge 2}) \geq (\alpha(G))^2$ for every graph G and $\gamma^{(2)} \geq 2\gamma^{(1)}$ for every channel. Shannon showed that for some channels two uses can result in further gains.

[1] $\mathcal{N}_k$ for a nonnegative integer k is the set of integers $\geq k$.

Example 1(b) For a completely-noisy channel we saw that $\mathcal{G}$ is the complete graph over $\mathcal{X}$. Hence $\mathcal{G}^{\wedge n}$ is the complete graph on $\mathcal{X}^n$, $\alpha(\mathcal{G}^{\wedge n}) = 1$, and $\gamma^{(n)} = 0$ for all $n \in \mathcal{N}_1$.

For the noiseless channel we saw that $\mathcal{G}$ is the empty graph over $\mathcal{X}$. Hence $\mathcal{G}^{\wedge n}$ is the empty graph over $\mathcal{X}^n$, $\alpha(\mathcal{G}^{\wedge n}) = |\mathcal{X}|^n$, and $\gamma^{(n)} = n \log |\mathcal{X}|$.

The Pentagon channel is more revealing. We saw that $\mathcal{G}$ is the Pentagon graph, whose independence number is 2. Shannon showed that $\{(0,0),(1,2),(2,4),(3,1),(4,3)\}$ is a largest-size independent set in $\mathcal{G}^{\wedge t}$, hence $\alpha(\mathcal{G}^{\wedge 2}) = 5$, implying that $\gamma^{(2)} > 2\gamma^{(1)}$. Lovász [14] showed that $\alpha(\mathcal{G}^{\wedge n}) = 5^{n/2}$ for every even n (see also Haemers [10]). □

Consider first the largest possible increase from $\alpha(G)$ to $\alpha(G^{\wedge 2})$ and from $\gamma^{(1)}$ to $\gamma^{(2)}$. For $l \in \mathcal{N}_0$ define

$$\rho_2(l) \stackrel{\text{def}}{=} \max\{\alpha(G^{\wedge 2}) : \alpha(G) \le l\}$$

where $\rho_2(0) = 0$. The subscript 2 refers to the two channel uses. Simple extensions of the Pentagon graph show that for every even l, $\rho_2(l) \ge \frac{5}{4}l^2$. Hence, for arbitrarily large values of $\gamma^{(1)}$ there are channels with $\gamma^{(2)} \ge 2\gamma^{(1)} + \log\frac{5}{4}$. However, more can be gained.

K_r is the complete graph over $\{1,\ldots,r\}$. A *two-coloration* of K_r is an assignment of one of two given colors to every edge (there are no further restrictions on the color assignments). A set S of vertices of K_r is *monochromatic* if all edges connecting vertices in S are assigned the same color. The *Ramsey number*[2] $r_2(l)$ of $l \in \mathcal{N}_0$ is the largest integer r for which there is a two-coloration of K_r where all monochromatic sets have size $\le l$. For example, $r_2(0) = 0$, $r_2(1) = 1$, $r_2(2) = 5$, and $r_2(3) = 17$. For higher values of α, only bounds are known; essentially:

$$2^{\frac{1}{2}l} \le r_2(l) < 2^{2l}. \tag{1}$$

For these and many other results concerning Ramsey numbers, see Graham, Rothschild, and Spencer [9]. In Subsection 2.1 we recount results of Erdős, McEliece, and Taylor [6] proving a general correspondence between Ramsey numbers and independence numbers of AND graph powers[3], showing in particular that for every $l \in \mathcal{N}_0$,

$$\rho_2(l) = r_2(l).$$

It follows that some channels can convey exponentially more bits in two uses than they can in one: for arbitrarily-large values of $\gamma^{(1)}$, there is a channel such that

$$\gamma^{(2)} \ge 2^{\gamma^{(1)}-1}, \tag{2}$$

and that this discrepancy is almost the highest possible: for every channel,

$$\gamma^{(2)} < 2^{\gamma^{(1)}+1}. \tag{3}$$

[2]The Ramsey number of $\alpha \in \mathcal{N}_1$ is often defined to be $r_2(l-1)+1$, the smallest integer r such that every two-coloration of K_r has a monochromatic set of size l.

[3]We proved these results independently and found out about [6] only after the paper was submitted. We kept the results partly because they complement the results of Section 3 and partly because the possible extensions described in Section 4 apply to them as well.

For multiple uses, it is instructive to consider the *per-use* number of bits the channel can convey without error. Shannon defined the *(zero-error) n-use capacity* of a channel $\mathcal{C}$ to be

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

By super-additivity, $C^{(n)}$ tends to a limit, $C^{(\infty)}$, known as Shannon's *zero-error capacity* of the channel. It is the highest per-use number of bits the channel can convey without error.

Example 1(c) For a completely-noisy channel $C^{(n)} = 0$ for all $n \in \{1,\ldots,\infty\}$. For a noiseless channel $C^{(n)} = \log|\mathcal{X}|$ for all $n \in \{1,\ldots,\infty\}$. For the pentagon channel $C^{(1)} = 1$ and, using Lovász's result, $C^{(2)} = C^{(4)} = C^{(6)} = \ldots = C^{(\infty)} = \frac{\log 5}{2} \approx 1.16$. □

Since $C^{(\infty)} \geq C^{(2)}$, Inequality (2) shows that for some channels

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

It is natural to ask whether this discrepancy is the largest possible. In the extreme, one wonders if there are channels that in one use convey only a constant number of bits, but in multiple uses convey arbitrarily many bits per use. Namely, whether there is a constant c such that for every c' there is a channel where

$$C^{(1)} \leq c \qquad \text{but} \qquad C^{(\infty)} \geq c'.$$

We could not resolve this question. However, towards the end of Subsection 2.1 we show that it generalizes an open problem proposed by P. Erdős. We note that the corresponding dual-source coding question is resolved affirmatively in Subsection 3.1.

So far, we considered only the number of bits conveyed and ignored the channel size. A channel where $\gamma^{(1)} \approx \log\log|\mathcal{X}|$ and $\gamma^{(2)} \approx \log|\mathcal{X}|$ is more "interesting" than one where $\gamma^{(1)} \approx \log\log\log|\mathcal{X}|$ and $\gamma^{(2)} \approx \log\log|\mathcal{X}|$, even though both display exponential increase in the number of transmissible bits. To relate the number of bits communicated to the channel's size, we define the *(zero-error) normalized n-use capacity* of a channel $\mathcal{C}$ to be

$$\tilde{C}^{(n)} = \frac{C^{(n)}}{\log|\mathcal{X}|} = \frac{\log \alpha(\mathcal{G}^{\wedge n})}{n \log|\mathcal{X}|}.$$

For every $n \in \{1,\ldots,\infty\}$, $\tilde{C}^{(n)}$ ranges from 0 for channels that can convey very little information error free, to 1 for channels where almost all input sequences can be communicated error free, thereby reflecting the channel's "quality."

Example 1(d) For a completely-noisy channel $\tilde{C}^{(n)} = 0$ for all $n \in \{1,\ldots,\infty\}$. For a noiseless channel $\tilde{C}^{(n)} = 1$ for all $n \in \{1,\ldots,\infty\}$. For the pentagon channel $\tilde{C}^{(1)} = \frac{1}{\log 5} \approx .43$ and $\tilde{C}^{(2)} = \tilde{C}^{(4)} = \tilde{C}^{(6)} = \ldots = \tilde{C}^{(\infty)} = \frac{1}{2}$. □

It is easy to see that $\alpha(G^{\wedge 2}) \leq v\alpha(G)$ for every v-vertex graph G. In Subsection 2.2 we use probabilistic constructions of *self-complementary Ramsey graphs* to show that this bound can be almost achieved even when $\alpha(G)$ is much smaller than v. For every v, multiple of 4, we exhibit a v-vertex graph G such that

$$\alpha(G) \leq \lceil 2\log v \rceil \qquad \text{but} \qquad \alpha(G^{\wedge 2}) \geq v.$$

It follows that for every $\epsilon > 0$ there is a channel such that

$$\tilde{C}^{(1)} \leq \epsilon \qquad \text{but} \qquad \tilde{C}^{(\infty)} \geq \tilde{C}^{(2)} \geq \frac{1}{2}.$$

For two uses, this is essentially the largest possible discrepancy: For every channel,

$$\tilde{C}^{(2)} - \tilde{C}^{(1)} \leq \frac{1}{2}.$$

We do not know whether for every $\epsilon > 0$ there is a channel such that $\tilde{C}^{(1)} \leq \epsilon$ but $\tilde{C}^{(\infty)} \geq 1 - \epsilon$. Namely, in one use almost no information can be transmitted, but in multiple use almost all input sequences can be transmitted. In Section 4 we discuss an extension of our proofs that may lead to such a result.

1.2 Dual-source coding

A *dual source* consists of a finite set $\mathcal{X}$, a (possibly infinite) set $\mathcal{Y}$, and a *support set* $S \subseteq \mathcal{X} \times \mathcal{Y}$. In each *dual-source instance*, a *sender* $P_{\mathcal{X}}$ is given an $x \in \mathcal{X}$ and a *receiver* $P_{\mathcal{Y}}$ is given a $y \in \mathcal{Y}$ such that $(x, y) \in S$. Following Witsenhausen [21] and Ferguson and Bailey [8] we study the number of bits that $P_{\mathcal{X}}$ must transmit in the worst case in order for $P_{\mathcal{Y}}$ to learn x without error. (See Orlitsky [17] for the case where $P_{\mathcal{X}}$ and $P_{\mathcal{Y}}$ are allowed to interact.)

The *fan-out* of $x \in \mathcal{X}$ is the set $S_x \stackrel{\text{def}}{=} \{y : (x, y) \in S\}$ of y's that are *jointly possible* with x. Associated with a dual source $\mathcal{S}$ is a *characteristic graph* $\mathcal{G}$. Its vertex set is $\mathcal{X}$, and two (distinct) vertices x, x' are connected if their fan-out sets intersect, namely, there is a y that is jointly possible with both. Note that every graph (V, E) is the characteristic graph of some dual source: $\mathcal{X} = V$, $\mathcal{Y} = E$, and $S = \{(x, y) : x \in y\}$.

The smallest number of possible messages $P_{\mathcal{X}}$ must transmit for a single instance of $\mathcal{S}$ is $\chi(\mathcal{G})$, the chromatic number of $\mathcal{S}$'s characteristic graph. Intuitively, $P_{\mathcal{X}}$ and $P_{\mathcal{Y}}$ agree in advance on a coloring of $\mathcal{G}$. Given x, $P_{\mathcal{X}}$ transmits its color. $P_{\mathcal{Y}}$, having y, can determine x because there is exactly one element of $\mathcal{X}$ 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.

The smallest number of bits $P_{\mathcal{X}}$ must transmit in the worst case for a single instance of $\mathcal{S}$ is

$$\sigma^{(1)} \stackrel{\text{def}}{=} \log \chi(\mathcal{G}).$$

Note that this definition allows for a non-integral number of bits. The actual number of bits that must be transmitted is $\lceil \sigma^{(1)} \rceil$.

Example 2(a) In an *uncorrelated dual source*, every two fan-out sets intersect. $\mathcal{G}$ is the complete graph over $\mathcal{X}$, $\chi(\mathcal{G}) = |\mathcal{X}|$, and $\sigma^{(1)} = \log |\mathcal{X}|$ indicating that x has to be specified completely for the receiver to learn its value.

In a *completely-correlated dual source* no two fan-out sets intersect. The characteristic graph $\mathcal{G}$ is the empty graph over $\mathcal{X}$, $\chi(\mathcal{G}) = 1$, and $\sigma^{(1)} = 0$ indicating that $P_{\mathcal{Y}}$ knows x and therefore no bits need be transmitted.

The *Pentagon dual source* has $\mathcal{X} = \mathcal{Y} = Z_5$ and $S = \{(x, y) : y = x \text{ or } y = x + 1 \pmod 5\}$. Figuratively, five countries are arranged around a circle. Occasionally, border disputes arise between neighboring countries. $P_\mathcal{Y}$ knows two countries involved in a dispute and $P_\mathcal{X}$ knows one of them, say the one who wins the dispute. We are interested in the number of bits that $P_\mathcal{X}$ must transmit in the worst case for $P_\mathcal{Y}$ to know the winning country. $\mathcal{G}$ is the pentagon graph defined in Example 1(a). Clearly, $\chi(\mathcal{G}) = 3$ and $\sigma^{(1)} = \log 3$, indicating that $\lceil \log 3 \rceil = 2$ bits are needed for $P_\mathcal{Y}$ to learn the winner. $\square$

In $n \in \mathcal{N}_2$ instances of the dual source $\mathcal{S}$, $P_\mathcal{X}$ knows $x_1,\ldots,x_n$ while $P_\mathcal{Y}$ knows $y_1,\ldots,y_n$ such that each $(x_i, y_i) \in S$ and wants to learn $x_1,\ldots,x_n$. Conceptually, these n instances can be viewed as a single instance of a larger dual source $\mathcal{S}^{(n)}$ whose support set is $S^n \subseteq \mathcal{X}^n \times \mathcal{Y}^n$. If $\overline{x} = (x_1,\ldots,x_n) \in \mathcal{X}^n$, then $S^n_{\overline{x}} = \{\overline{y} : (\overline{x}, \overline{y}) \in S^n\} = S_{x_1} \times \ldots \times S_{x_n}$. Let $\mathcal{G}^{(n)}$ denote the characteristic graph of $\mathcal{S}^{(n)}$. The argument used in the previous subsection for n channel uses shows that

$$\mathcal{G}^{(n)} = \mathcal{G}^{\wedge n}.$$

It follows that the number of bits that $P_\mathcal{X}$ must transmit in the worst case to convey n instances of $\mathcal{S}$ without error is

$$\sigma^{(n)} \stackrel{\text{def}}{=} \log \chi(\mathcal{G}^{\wedge n}).$$

If G can be colored with χ colors, $G^{\wedge 2}$ can be colored with χ^2 colors. Therefore, $\chi(G^{\wedge 2}) \leq (\chi(G))^2$ for every graph G, and $\sigma^{(2)} \leq 2\sigma^{(1)}$ for every dual source. Witsenhausen [21] showed that for some dual sources, fewer bits suffice.

Example 2(b) For a completely-correlated source we saw that $\mathcal{G}$ is the complete graph over $\mathcal{X}$. Hence $\mathcal{G}^{\wedge n}$ is the complete graph on $\mathcal{X}^n$, $\chi(\mathcal{G}^{\wedge n}) = |\mathcal{X}|^n$, and $\sigma^{(n)} = n \log |\mathcal{X}|$ for all $n \in \mathcal{N}_1$.

For a completely-correlated dual source we saw that $\mathcal{G}$ is the empty graph over $\mathcal{X}$. Hence $\mathcal{G}^{(n)}$ is the empty graph over $\mathcal{X}^n$, $\chi(\mathcal{G}^{\wedge n}) = 1$, and $\sigma^{(n)} = 0$.

The Pentagon dual source is more interesting. We saw that $\mathcal{G}$ is the Pentagon graph, whose chromatic number is 3. Using the corresponding results by Shannon and Lovász (Example 1(b)), Witsenhausen showed that for every even n, $\chi(\mathcal{G}^{\wedge n}) = 5^{n/2}$. In particular, $\chi(\mathcal{G}) = 5$, hence $\sigma^{(2)} < 2\sigma^{(1)}$. $\square$

Consider first the smallest possible increase from $\chi(G)$ to $\chi(G^{\wedge 2})$ and from $\sigma^{(1)}$ to $\sigma^{(2)}$. Simple extensions of the Pentagon graph shows that for every $\chi(G)$ divisible by 3, there is a graph with $\chi(G^{\wedge 2}) \leq \frac{5}{9}(\chi(G))^2$, hence for arbitrarily-large values of $\sigma^{(1)}$ there is a dual source where $\sigma^{(2)} \leq 2\sigma^{(1)} - \log 1.8$. However, more bits can be saved.

Clearly, $\chi(G^{\wedge 2}) \geq \chi(G)$ for every graph G. Linial and Vazirani [12] showed that for arbitrarily-large values of $\chi(G)$ there are graphs such that $\chi(G^{\wedge 2}) = O(\chi(G))$. In Subsection 3.2 we significantly reduce the implied constant, showing that for arbitrarily-large values of $\chi(G)$ there are graphs such that

$$\chi(G^{\wedge 2}) \leq \lceil 34 \cdot \chi(G) \rceil.$$

It follows that for arbitrarily-large values of $\sigma^{(1)}$ there are dual sources where two instances require only few more bits than one instance:

$$\sigma^{(2)} \le \sigma^{(1)} + 6.$$

For multiple uses, it is instructive to consider the *per-instance* number of bits required to convey the x's without error. Witsenhausen defined the *(zero-error) n-instance rate* of a dual source $\mathcal{S}$ to be

$$R^{(n)} = \frac{\sigma^{(n)}}{n} = \frac{\log \chi(\mathcal{G}^{\wedge n})}{n}.$$

By sub additivity, $R^{(n)}$ tends to $R^{(\infty)}$, *Witsenhausen's zero-error rate* of $\mathcal{S}$. It is the lowest per-instance number of bits that must be transmitted in the worst case to convey the x's to $P_{\mathcal{Y}}$.

Example 2(c) For an uncorrelated dual source $R^{(n)} = \log|\mathcal{X}|$ for all $n \in \{1,\ldots,\infty\}$. For a completely-correlated dual source $R^{(n)} = 0$ for all $n \in \{1,\ldots,\infty\}$. For the pentagon dual source $R^{(1)} = \log 3 \approx 1.58$ and $R^{(2)} = R^{(4)} = R^{(6)} = \ldots = R^{(\infty)} = \frac{\log 5}{2} \approx 1.16$. $\square$

Also in Subsection 3.1, we show that for every $\epsilon, t > 0$ there is a graph G such that for every n,

$$\chi(G) \ge \epsilon t \qquad \text{but} \qquad \chi(G^{\wedge n}) \le \left\lceil (2+\epsilon)^{n+1} n t \ln 2 \right\rceil.$$

For large t and small ϵ, $\chi(G^{\wedge n})$ grows significantly slower than $(\chi(G))^n$. It follows that for every $\epsilon > 0$ and arbitrarily-large $\sigma^{(1)}$ there are sources where

$$R^{(1)} \ge \sigma^{(1)} \qquad \text{but} \qquad R^{(\infty)} \le 1 + \epsilon.$$

Namely, one instance requires arbitrarily many bits, but multiple instances require about one bit per instance. Subsection 3.2 shows another family of graphs for which slightly weaker results hold. For interactive communication, where the communicators are allowed to communicate back and forth, similar results were established by Feder, Kushilevitz, and Naor [7], and by Naor, Orlitsky, and Shor [16].

The graphs used to derive the above results have $\chi(G)$ which is merely logarithmic in the graph's size and therefore the implied source has $\sigma^{(1)}$ which are only about $\log\log|\mathcal{X}|$. The same holds for the afore-mentioned interactive-communication results. Dual sources requiring a large number of bits are of more interest. In fact, recent interest in the number of bits required for multiple instances (see also Subsection 1.3) was partly motivated by Karchmer, Raz, and Wigderson [11] who related certain open problems in computational complexity to the number of bits required to communicate multiple instances of problems with high communication complexity.

In Subsection 3.3 we show that $\chi(G^{\wedge 2})$ can be about $\chi(G)$ even when $\chi(G)$ is close to the the graph's size, and therefore that $\sigma^{(2)}$ can be about $\sigma^{(1)}$ even when $\sigma^{(1)}$ is close to $\log|\mathcal{X}|$. Using probabilistic constructions of *self-complementary Ramsey graphs* that are also *Cayley graphs* we show that for every prime power $v \equiv 1 \bmod 4$ there is a v-vertex graph G such that[4]

$$\chi(G) \ge \frac{v}{(1+o(1))16\log^2 v} \qquad \text{but} \qquad \chi(G^{\wedge 2}) \le v. \tag{4}$$

[4]Throughout this paper the $o(1)$ term diminishes to zero as the relevant parameters (here, v) tend to infinity.

It follows that for arbitrarily-high values of $\chi(G)$ and v such that $\chi(G) \geq \frac{v}{(1+o(1))16\log^2 v}$ there is a graph G where:

$$\chi(G^{\wedge 2}) \leq (1+o(1))16\chi(G)\log^2\chi(G). \tag{5}$$

Therefore, for arbitrarily-high values of $\sigma^{(1)}$ and $|\mathcal{X}|$ such that $\sigma^{(1)} \geq \log|\mathcal{X}| - 2\log\log|\mathcal{X}| - 4 - o(1)$, there are dual sources where

$$\sigma^{(2)} \leq \sigma^{(1)} + 2\log\sigma^{(1)} + 4 + o(1).$$

To relate the number of bits transmitted to the source's size and to account for the number of instances, we define the *(zero-error) normalized n-instance rate* of a dual source $\mathcal{S}$ to be

$$\tilde{R}^{(n)} = \frac{R^{(n)}}{\log|\mathcal{X}|} = \frac{\log\chi(\mathcal{G}^{\wedge n})}{n\log|\mathcal{X}|}.$$

For every $n \in \{1,\ldots,\infty\}$, $\tilde{R}^{(n)}$ ranges from 1 for dual sources where x and y are unrelated, to 0 for dual sources where y determines x, thereby reflecting the "difficulty" of conveying x to $P_{\mathcal{Y}}$.

Example 2(d) For an uncorrelated dual source $\tilde{R}^{(n)} = 1$ for all $n \in \{1,\ldots,\infty\}$. For a completely-correlated dual source $\tilde{R}^{(n)} = 0$ for all $n \in \{1,\ldots,\infty\}$. For the pentagon dual source $\tilde{R}^{(1)} = \frac{\log 3}{\log 5} \approx .683$ and $\tilde{R}^{(2)} = \tilde{R}^{(4)} = \tilde{R}^{(6)} = \ldots = \tilde{R}^{(\infty)} = .5$. $\square$

Inequalities (4) imply that for every $\epsilon > 0$ there is a dual source such that

$$\tilde{R}^{(1)} \geq 1-\epsilon \qquad \text{but} \qquad \tilde{R}^{(\infty)} \leq \tilde{R}^{(2)} \leq \frac{1}{2}.$$

For two instances, this difference is essentially the largest possible. Clearly, $\chi(G^{\wedge 2}) \geq \chi(G)$ for every graph G. Hence, for all dual sources,

$$\tilde{R}^{(2)} - \tilde{R}^{(1)} \leq \frac{1}{2}.$$

We do not know whether for every $\epsilon > 0$ there is a dual source such that $\tilde{R}^{(1)} \geq 1-\epsilon$ but $\tilde{R}^{(\infty)} \leq \epsilon$. Namely, a single instance requires almost complete specification of X, while multiple instances require very little information. In Section 4 we discuss an extension of our proofs that may lead to such a result.

1.3 Relations to OR graph powers and to communication complexity

Besides the aforementioned applications, this work was motivated by two problems that, while similar in appearance, yield very different results.

The n-th *OR* (or *co-normal*, or *inclusive*) *power* of a graph G with vertex set V is the graph $G^{\vee n}$ whose vertex set is V^n and where distinct vertices $(x_1,\ldots,x_n)$ and $(x'_1,\ldots,x'_n)$ are connected if distinct x_i and x'_i are connected in G for some $i \in \{1,\ldots,n\}$. Clearly, $\chi(G^{\vee 2}) \leq (\chi(G))^2$ for every graph G. Linial and Vazirani [12] showed that for every v-vertex graph G,

$$\chi(G^{\vee 2}) \geq \frac{\chi(G)(\chi(G)-1)}{\log v} \tag{6}$$

and that graphs related to *directed line graphs* achieve this bound up to a constant factor. Contrasting (5) and (6) for graphs G with relatively high $\chi(G)$, we see that the chromatic number of G's OR square is always around $(\chi(G))^2$ while the chromatic number of G's AND square may be much closer to $\chi(G)$.

Another contrast relates to the communication setup. In dual-source coding we assume that (x,y) is always in S. Karchmer, Raz, and Wigderson [11] considered a similar problem which for identification purposes we call *communication complexity* (see also Yao [22]). (x,y) can attain any value in $\mathcal{X}\times\mathcal{Y}$. However, $P_{\mathcal{Y}}$ needs to learn x only if $(x,y)\in S$. When $(x,y)\notin S$, $P_{\mathcal{Y}}$ can be wrong on the value of x. Let $L^{(n)}$ be the per-instance number of bits required in the worst case (over all inputs in $\mathcal{X}\times\mathcal{Y}$) for n independent instances of the problem.

For a single instance $L^{(1)}=R^{(1)}$. Any protocol for a communication-complexity problem can be used unaltered for the corresponding dual-source coding problem. Conversely, any protocol for a dual-source problem can be easily modified to work for the corresponding communication-complexity problem. Simply stop communicating after $R^{(1)}$ bits have been transmitted.[5]

For n instances $L^{(n)}$ can be interpreted in two ways:

1. $P_{\mathcal{Y}}$ needs to find $x_1,\ldots,x_n$ iff $(x_i,y_i)\in S$ for all i. Under this interpretation, $L^{(n)}=R^{(n)}$. Hence our results show that for some sets S, $L^{(2)}\approx\frac{1}{2}L^{(1)}$, namely two independent instances require about the same transmission as one.

2. $P_{\mathcal{Y}}$ needs to learn x_i for exactly those i's where $(x_i,y_i)\in S$, and can be wrong about x_i for the other i's. This interpretation may require more bits. Feder, Kushilevitz, and Naor [7] showed that in that case, $L^{(n)}\geq L^{(1)}-\log\log|\mathcal{X}|$ for all n. Namely, hardly any savings arises from multiple instances.

To contrast the two interpretations consider two instances (x_1,y_1) and (x_2,y_2). If $P_{\mathcal{Y}}$ always needs to learn the x_i's such that $(x_i,y_i)\in S$ then roughly $2L^{(1)}$ bits are needed. But if $P_{\mathcal{Y}}$ needs to learn the x_i's such that $(x_i,y_i)\in S$ only if both (x_1,y_1) and (x_2,y_2) are in S, then roughly $L^{(1)}$ bits may suffice for both instances.

2 Channel coding

2.1 Ramsey numbers

Erdős, McEliece, and Taylor [6] related Ramsey numbers to independence numbers of AND graph products. The results are more general than claimed in the introduction and require additional definitions.

An *edge coloration* of a graph is an assignment of a color to each edge (no restrictions imposed on the colors). An edge coloration using at most n colors is an *n-coloration.* A set of vertices is *i-monochromatic* for color i in a coloration if all edges connecting vertices in the set are assigned

[5] This argument does not apply to the average number of bits.

color i. The *Ramsey number* $r(l_1,\ldots,l_n)$ of $n \in \mathcal{N}_1$ and $l_1,\ldots,l_n \in \mathcal{N}_0$ is the largest integer r for which there is an n-coloration of the complete graph K_r using colors from $\{1,\ldots,n\}$ where for every color i all i-monochromatic sets have size $\leq l_i$. In particular, $r(l) = l$ for every $l \in \mathcal{N}_0$, and $r(l_1,\ldots,l_n) = 0$ whenever some $l_i = 0$. Also, $r(l,l)$ is simply $r_2(l)$ defined in the introduction. In general, we abbreviate

$$r_n(l) \stackrel{\text{def}}{=} r(\overbrace{l,\ldots,l}^{n}).$$

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 $V_1 \times \cdots \times V_n$ and where distinct vertices $(v_1,\ldots,v_n)$ and $(v'_1,\ldots,v'_n)$ are connected iff $\{v_i,v'_i\} \in E_i$ for all $i \in \{1,\ldots,n\}$ such that $v_i \neq v'_i$. Note that $G^{\wedge n}$ is the n-fold AND product of G with itself. Define:

$$\rho(l_1,\ldots,l_n) \stackrel{\text{def}}{=} \max\{\alpha(G_1 \wedge \cdots \wedge G_n) : \alpha(G_i) \leq l_i \text{ for all } i \in \{1,\ldots,n\}\}.$$

$\rho(\overbrace{l,\ldots,l}^{n})$ is not directly related to n-use capacities. Rather, we are interested in

$$\rho_n(l) \stackrel{\text{def}}{=} \max\{\alpha(G^{\wedge n}) : \alpha(G) \leq l\}.$$

Theorem 1 below shows that $\rho(l_1,\ldots,l_n) = r(l_1,\ldots,l_n)$. Theorem 2 shows that $\rho_n(l) = \rho(l,\ldots,l)$. Therefore, for every $n \in \mathcal{N}_1$ and $l \in \mathcal{N}_0$,

$$\rho_n(l) = r_n(l).$$

The proofs are most succinct when instead of independent sets and AND products, we consider *cliques* and *OR products*.

A *clique* in a graph is a set of vertices, every two connected. The *clique number*, $\omega(G)$, of a graph G is the size of its largest clique. The *complement* of a graph G is the graph $\overline{G}$ which has the same vertex set as G and where distinct vertices are connected iff they are not connected in G. Clearly, $\omega(G) = \alpha(\overline{G})$ for every graph G.

The *OR product* of n graphs $G_1 = (V_1,E_1),\ldots,G_n = (V_n,E_n)$ is the graph $G_1 \vee \cdots \vee G_n$ whose vertex set is the Cartesian product $V_1 \times \cdots \times V_n$ and where distinct vertices $(v_1,\ldots,v_n)$ and $(v'_1,\ldots,v'_n)$ are connected iff $\{v_i,v'_i\} \in E_i$ for some $i \in \{1,\ldots,n\}$ such that $v_i \neq v'_i$. The n-fold OR product of a graph G with itself is its n-th OR power, denoted $G^{\vee n}$. It is easy to see that for every $G_1,\ldots,G_n$,

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

It follows that $\rho(l_1,\ldots,l_n)$ can be interpreted as:

$$\rho(l_1,\ldots,l_n) = \max\{\omega(G_1 \vee \cdots \vee G_n) : \omega(G_i) \leq l_i \text{ for all } 1 \leq i \leq n\}. \tag{7}$$

Theorem 1 ([6]) For every $n \in \mathcal{N}_1$ and $l_1,\ldots,l_n \in \mathcal{N}_0$,

$$\rho(l_1,\ldots,l_n) = r(l_1,\ldots,l_n).$$

Proof: $\rho(l_1,\ldots,l_n) \geq r(l_1,\ldots,l_n)$. Let $r \stackrel{\text{def}}{=} r(l_1,\ldots,l_n)$. By definition, there is a coloration of K_r with colors from $\{1,\ldots,n\}$ where for all $i \in \{1,\ldots,n\}$, all i-monochromatic sets have size $\leq l_i$. Let G_i be the graph defined over $\{1,\ldots,r\}$ and where distinct vertices a and b are connected if the edge $\{a,b\}$ is colored i in the given coloration. Every clique of G_i is i-monochromatic in K_r, hence $\omega(G_i) \leq l_i$. On the other hand, the set $\{(1,\ldots,1),\ldots,(r,\ldots,r)\}$ is a clique in $G_1 \vee\!\!\!\!\vee \cdots \vee\!\!\!\!\vee G_n$.

$r(l_1,\ldots,l_n) \geq \rho(l_1,\ldots,l_n)$. Let $\rho = \rho(l_1,\ldots,l_n)$. Then there are graphs $G_1,\ldots,G_n$ such that $\omega(G_i) \leq l_i$ and $G_1 \vee\!\!\!\!\vee \cdots \vee\!\!\!\!\vee G_n$ contains a size-ρ clique $S = \{(x_1^1,\ldots,x_n^1),\ldots,(x_1^\rho,\ldots,x_n^\rho)\}$. Color K_ρ with colors from $\{1,\ldots,n\}$ by assigning each edge $\{a,b\}$ the first color i such that $x_i^a \neq x_i^b$ and $\{x_i^a, x_i^b\} \in G_i$. Since S is clique, every edge of K_ρ is colored. Every i-monochromatic set has size $\leq l_i$, because if $M \subseteq \{1,\ldots,r\}$ is i-monochromatic then $\{x_i^m : m \in M\}$ is a clique in G_i. □

When $l_1 = \ldots = l_n$, the product can be taken over a single graph:

Theorem 2 ([6]) For every $n \in \mathcal{N}_1$ and $l \in \mathcal{N}_0$,

$$\rho_n(l) = \rho(\overbrace{l,\ldots,l}^{n}).$$

Proof: Let $\rho = \rho(l,\ldots,l)$. As in (7), $\rho_n(l) = \max\{\omega(G^{\vee\!\!\!\vee n}) : \omega(G) \leq l\}$. It therefore suffices to show a graph G such that $\omega(G) \leq l$ and $\omega(G^{\vee\!\!\!\vee n}) \geq \rho$.

By definition, there are graphs $G_1,\ldots,G_n$ each with clique number $\leq l$ such that $G_1 \vee\!\!\!\!\vee \cdots \vee\!\!\!\!\vee G_n$ contains a size-ρ clique $S = \{(x_{1,1},\ldots,x_{1,n}),\ldots,(x_{\rho,1},\ldots,x_{\rho,n})\}$. Let the graph G have the vertex-set $\{1,\ldots,\rho\} \times \{1,\ldots,n\}$. Vertex (a,i) is connected to all vertices (b,i) where $\{x_{a,i}, x_{b,i}\} \in G_i$; it is not connected to any other (b,i) or any (b,j) for $j \neq i$. Clearly, $\omega(G) \leq \max\{\omega(G_i)\} \leq l$. Yet, it is easy to verify that the set $\{((a,1),\ldots,(a,n)) : a \in \{1,\ldots,\rho\}\}$ is a clique in $G^{\vee\!\!\!\vee n}$. □

Note that any graph achieving $\rho_n(l)$ must have at least $(\rho_n(l))^{1/n}$ vertices. The graphs constructed in the proof have size $n\rho_n(l)$.

Theorems 1 and 2 imply that for every $n \in \mathcal{N}_1$ and $l \in \mathcal{N}_0$,

$$\rho_n(l) = \rho(l,\ldots,l) = r(l,\ldots,l) = r_n(l).$$

In particular, $\rho_2(l) = r_2(l)$, hence results on Ramsey numbers can be used to bound the largest increase from $\gamma^{(1)}$ to $\gamma^{(2)}$.

(A special case of) the well known theorem of Ramsey [18] asserts that any v-vertex graph G contains a logarithmic-size clique or independent set:

$$\max\{\alpha(G), \omega(G)\} > \frac{1}{2}\log v. \tag{8}$$

Erdős [5] showed that up to a constant factor this bound is tight. For every $v \geq 2$ there is a v-vertex graph G containing neither a clique nor an independent set of size $> 2\log v$:

$$\max\{\alpha(G), \omega(G)\} \leq 2\log v. \tag{9}$$

A graph whose independence and clique numbers are both polylogarithmic in the number of vertices is called a *Ramsey graph*.

These results imply the bounds in Inequality (1): for every $l \in \mathcal{N}_0$,

$$2^{\frac{1}{2}l} \leq r_2(l) < 2^{2l}.$$

Which in turn imply the bounds in (2) and (3).

We have therefore demonstrated channels whose infinite-use capacity is exponentially larger than their single-use capacity: $C^{(\infty)} \geq 2^{C^{(1)}-2}$. We would like to know whether there are channels where $C^{(\infty)}$ is arbitrarily higher than $C^{(1)}$. In view of the correspondence between $\rho_n(l)$ and $r_n(l)$,

$$\max\{C^{(n)} : C^{(1)} \leq c\} = \max\{\frac{\log \alpha(G^{\wedge n})}{n} : \alpha(G) \leq 2^c\} = \frac{\log \rho_n(2^c)}{n} = \frac{\log r_n(2^c)}{n}.$$

Therefore, there is an arbitrary gap between $C^{(1)}$ and $C^{(\infty)}$ iff for some constant $c'(= 2^c)$ the Ramsey number $r_n(c')$ grows faster than any exponential in n:

$$r_n(c') \geq 2^{ng(n)} \qquad \text{where} \qquad \limsup_{n\to\infty} g(n) = \infty.$$

This generalizes an open problem proposed by Erdős (see, e.g., Graham, Rothschild, and Spencer [9], page 146), asking whether $r_n(2)$ grows faster than any exponential in n.

2.2 Self-complementary Ramsey graphs

We are interested in large increases in the number of transmissible bits for channels conveying relatively many bits. Later in this section, we prove the following discrepancy between the independence number of a graph and its normal square. The ensuing corollary follows immediately.

Theorem 3 For every $v \in 4\mathcal{N}_1$ $(= \{4, 8, 12, \ldots\})$ there is a v-vertex graph G such that

$$\alpha(G) \leq 2\lceil \log v \rceil \qquad \text{and} \qquad \alpha(G^{\wedge 2}) \geq v. \qquad \square$$

Corollary 1 For every $\epsilon > 0$ there is a channel such that

$$\tilde{C}^{(1)} \leq \epsilon \qquad \text{and} \qquad \tilde{C}^{(\infty)} \geq \tilde{C}^{(2)} \geq \frac{1}{2}. \qquad \square$$

For two channel uses, this discrepancy is essentially the largest possible:

Lemma 1 For every channel and integer n,

$$\tilde{C}^{(n)} - \tilde{C}^{(1)} \leq (1 - \frac{1}{n})(1 - \tilde{C}^{(1)}).$$

Proof: For every v-vertex graph G and integer n, $\alpha(G^{\wedge n}) \leq v^{n-1} \cdot \alpha(G)$. Hence $\tilde{C}^{(n)} \leq (1 - \frac{1}{n}) + \frac{1}{n}\tilde{C}^{(1)}$ and the lemma follows. $\square$

Results in the previous subsection imply that for every $\epsilon > 0$ there is a channel such that $\tilde{C}^{(1)} \leq \epsilon$ and $\tilde{C}^{(2)} \geq \frac{1}{2} - \epsilon$. In that sense Theorem 3 represents only a mild improvement. However, the theorem's proof, given in the rest of the section, constructs more symmetric *self-complementary* Ramsey graphs that can be generalized to yield results needed for dual-source coding (Subsection 3.3) where the improvements yielded by this method are substantial.

A graph G is *self complementary* if it is *isomorphic* to its complement. Namely, if there is a permutation π of its vertices such that for every pair x, x' of distinct vertices, $\{x, x'\} \in G$ iff $\{\pi(x), \pi(x')\} \notin G$.

Lemma 2 Every self-complementary graph G on v vertices has $\alpha(G^{\wedge 2}) \geq v$.

Proof: Let $G = (V, E)$ be self complementary with π mapping G onto its complement. Then the set $\{(x, \pi(x)) : x \in V\}$ is independent in $G^{\wedge 2}$ because if $x \neq x'$ then $\{x, x'\} \in G$ implies that $\{\pi(x), \pi(x')\} \notin G$. □

To establish Theorem 3 it therefore suffices to prove:

Lemma 3 For every $v \in 4\mathcal{N}_1$ there is a self-complementary graph G on v vertices satisfying $\alpha(G) < 2\lceil \log v \rceil$.

Proof: Let $v = 4a$ where $a \in \mathcal{N}_1$ and let $\mathcal{Z}_v$, the additive group modulo v, be the graph's vertex-set. All operations involving vertices are performed in $\mathcal{Z}_v$, where $4a = 0$. Define an equivalence relation on "potential edges:" $\{x, y\} \sim \{x', y'\}$ if $\{x', y'\} = \{x + ia, y + ia\}$ for some $i \in \{0, 1, 2, 3\}$. Let

$$E_{\{x,y\}} \stackrel{\text{def}}{=} \{\{x, y\}, \{x + a, y + a\}, \{x + 2a, y + 2a\}, \{x + 3a, y + 3a\}\}$$

denote the equivalence class of $\{x, y\}$. Note that if $y = x + 2a$ then $E_{\{x,y\}}$ consists of only two elements: $\{x, y\}$ and $\{x + a, y + a\}$.

Choose the edge-set E randomly. For each equivalence class $E_{\{x,y\}}$ such that $y = x + 2a$, randomly and independently select precisely one of the two edges $\{x, y\}$ and $\{x + a, y + a\}$ to be in E. For each equivalence class $E_{\{x,y\}}$ such that $y \notin \{x, x + 2a\}$, randomly and independently select precisely two of its four edges to be in E: either the two edges $\{x, y\}$ and $\{x + 2a, y + 2a\}$, or the two edges $\{x + a, y + a\}$ and $\{x + 3a, y + 3a\}$.

The resulting graph is clearly self-complementary under the permutation $\pi(x) \stackrel{\text{def}}{=} x + a$. We prove that with positive probability its independence number is less than $t \stackrel{\text{def}}{=} 2\lceil \log v \rceil$.

Let T be a fixed t-element subset of $\mathcal{Z}_v$ and let $\{x_i, y_i\}$ for $i \in \{1, \ldots, \binom{t}{2}\}$ be the pairs of distinct T-elements in some (arbitrary) order. T is independent if none of these pairs is in E. We calculate the probability of that event. If T contains some vertices of the form $x, y, x + a, y + a$ it is certainly not independent, hence we restrict our attention to sets T that do not contain such vertices. Let i denote the number of pairs of vertices of the form $x, x + 2a$ that belong to T. Then $0 \leq i \leq t/2$, and T contains $t - 2i$ additional vertices. It is not difficult to check that for each such T with $i \geq 2$, there are $\binom{t}{2} - 2\binom{i}{2}$ pairs of vertices of T whose choices as edges in E are mutually independent. For $i \leq 1$, all the $\binom{t}{2}$ choices are independent. It follows that the expected number of independent sets of size t is at most

$$\begin{aligned}
&\binom{v}{t}2^{-\binom{t}{2}}+\sum_{i=2}^{t/2}\binom{v/2}{i}\binom{v}{t-2i}\;2^{-\binom{t}{2}+i(i-1)}\\
&\le \frac{v^t}{t!}2^{-\binom{t}{2}}+\sum_{i=2}^{t/2}\frac{1}{2^i i!(t-2i)!}v^{t-i}\;2^{-\binom{t}{2}+(t/2)(i-1)}\\
&=\frac{2^{t/2}}{t!}(v/2^{t/2})^t+\sum_{i=2}^{t/2}\frac{1}{2^i i!(t-2i)!}(v/2^{t/2})^{t-i}\\
&\le \frac{2^{t/2}}{t!}+\sum_{i=2}^{t/2}\frac{1}{2^i i!(t-2i)!}<1.
\end{aligned}$$

Therefore, with positive probability there is no independent set of size t, as required. $\square$

Remarks:

1. The graphs constructed by the lemma have the same independence and clique numbers, showing that the bound in Inequality (9) is achievable even by self-complementary Ramsey graphs.

2. It can be shown that the random v-vertex graph G constructed in the proof has $\chi(G^{\wedge 2}) = \Theta(v^2/\log^2 v)$ almost surely (i.e., with probability that tends to one as v tends to infinity). Therefore G cannot be used in the next section where we need self-complementary graphs with large chromatic numbers which are close to those of their squares.

3 Dual-source coding

3.1 Kneser graphs

The *Kneser graph* $K = K(u,t)$ consists of all $\binom{u}{t}$ t-element subsets of $\{1,\ldots,u\}$. Two vertices are connected iff they are disjoint. Every vertex can be colored with one of its elements, hence $\chi(K) \le u$. But fewer colors suffice. Among two disjoint t-element subsets of $\{1,\ldots,u\}$, at least one contains an element $\le u-2t+1$. Therefore, the mapping which assigns to every S the smaller of $u-2t+2$ and $\min(S)$ also colors K. Lovász [13] showed that the number of colors cannot be reduced:

$$\chi(K) = u - 2t + 2.$$

Example 3(a) (Neighborhood games) u basketball players, numbered $1,\ldots,u$, meet at a neighborhood court. Two t-player teams ($t \le \lfloor u/2 \rfloor$) soon form and play each other. $P_{\mathcal{Y}}$ knows the two teams (namely, two disjoint sets $\{i_1,\ldots,i_t\}$ and $\{j_1,\ldots,j_t\}$) while $P_{\mathcal{X}}$ knows the winning team (say, $\{j_1,\ldots,j_t\}$) and would like to convey that information to $P_{\mathcal{Y}}$.

The vertices of the characteristic graph $\mathcal{G}$ are all t-element subsets of $\{1,\ldots,u\}$. Two vertices are connected iff they are disjoint. $\mathcal{G}$ is therefore the *Kneser graph* $K(u,t)$, and

$$\sigma^{(1)} = \log(u - 2t + 2). \qquad \square$$

We show that AND and OR powers of Kneser graphs can be colored with relatively few colors.

Theorem 4

$$\chi(K^{\wedge n}) \le \chi(K^{\vee n}) \le \left\lceil \left(\frac{u}{t}\right)^n \cdot n \cdot \ln \binom{u}{t} \right\rceil.$$

Proof: The vertices of $K^{\vee n}$ are all $\binom{u}{t}^n$ *team-sequences* $(S_1,\ldots,S_n)$ where each S_j is a t-element subset of $\{1,\ldots,u\}$. Distinct vertices $(S_1,\ldots,S_n)$ and $(S'_1,\ldots,S'_n)$ are connected iff S_j and S'_j are disjoint for some $j \in \{1,\ldots,n\}$.

A *player-sequence* $\mathbf{z} = (z_1,\ldots,z_n) \in \{1,\ldots,u\}^n$ *represents* a team-sequence $\mathbf{S} \stackrel{\text{def}}{=} S_1,\ldots,S_n$, written $\mathbf{z} \in \mathbf{S}$, if $z_j \in S_j$ for all $j \in \{1,\ldots,n\}$. Any assignment of a representing player-sequence to each vertex colors $K^{\vee n}$. We show that with positive probability, a randomly chosen set of

$$m = \left\lceil \left(\frac{u}{t}\right)^n \cdot n \cdot \ln \binom{u}{t} \right\rceil$$

player-sequences $\mathbf{z}_1,\ldots,\mathbf{z}_m$ represents every vertex, namely, for every team-sequence $\mathbf{S}$ there is an $i \in \{1,\ldots,m\}$ such that $\mathbf{z}_i \in \mathbf{S}$.

Choose the n players in a player-sequence $\mathbf{z} = z_1,\ldots,z_n$ uniformly and independently from $\{1,\ldots,u\}$. Consider any (fixed for now) team-sequence $\mathbf{S} = (S_1,\ldots,S_n)$. For every $j \in \{1,\ldots,n\}$, $\Pr(z_j \in S_j) = \frac{t}{u}$. Hence, $\Pr(\mathbf{z} \in \mathbf{S}) = \left(\frac{t}{u}\right)^n$. Equivalently, $\Pr(\mathbf{z} \notin \mathbf{S}) = 1 - \left(\frac{t}{u}\right)^n$. Let $\mathbf{z}_1,\ldots,\mathbf{z}_m$ be independent copies of $\mathbf{z}$. Then, $\Pr(\forall i \in \{1,\ldots,n\}, \mathbf{z}_i \notin \mathbf{S}) = \left(1 - \left(\frac{t}{u}\right)^n\right)^m$, and, since there are $\binom{u}{t}$ different team sequences,

$$\Pr(\exists \mathbf{S} \text{ such that } \forall i \in \{1,\ldots,n\}, \mathbf{z}_i \notin S) \le \binom{u}{t}^n \left(1 - \left(\frac{t}{u}\right)^n\right)^m < \binom{u}{t}^n e^{-m\left(\frac{t}{u}\right)^n}.$$

By the choice of m, this probability is < 1, hence every team-sequence contains a representative. $\square$

To obtain small differences between the chromatic number of a graph and its AND/OR squares, let $u = 3.25 \cdot t$. Then

$$\chi(K) \ge 1.25 \cdot t \qquad \text{while} \qquad \chi(K^{\vee 2}) \le \lceil 42.4 \cdot t \rceil.$$

Hence for arbitrarily large values of $\chi(K)$,

$$\chi(K^{\vee 2}) \le 34 \chi(K).$$

For asymptotically many instances,

$$\lim_{n\to\infty} \left(\chi(K^{\vee n})\right)^{1/n} \le \frac{u}{t}.$$

McEliece and Posner [15], and Berge and Simonowitz [4] showed that

$$\lim_{n\to\infty} \left(\chi(G^{\vee n})\right)^{1/n} = \chi^*(\mathcal{G}),$$

the *fractional chromatic number* of a graph G. Letting $u = (2+\epsilon)t$, we see that there can be an arbitrarily-high discrepancy between the standard and fractional chromatic numbers of a graph: for every $\chi, \epsilon > 0$ there is a graph such that

$$\chi(G) \ge \chi \qquad \text{but} \qquad \chi^*(G) \le 2+\epsilon.$$

Namely, $\chi(G)$ is arbitrarily high, but $\chi(G^{\vee n})$ grows like $(2+\epsilon)^n$.

Following are some implications of these results on independent instances of Example 3(a).

Example 3(b) In a happening neighborhood, n sports are played. Each sport engages u distinct players, and a game involves two teams of t players each. On a certain day, n games take place, one in each sport. $P_{\mathcal{Y}}$ knows the $2n$ playing teams while $P_{\mathcal{X}}$ knows the n that won. How many bits must $P_{\mathcal{X}}$ transmit now?

Recall that $\sigma^{(1)} = \log(u-2t+2)$. Theorem 4 shows that

$$\sigma^{(n)} \le \lceil n \log \frac{u}{t} + \log n + \log \ln \binom{u}{t} \rceil.$$

Setting $u = 3.25t$, we see that for arbitrarily high $\sigma^{(1)}$ there are sources with

$$\sigma^{(2)} \le \sigma^{(1)} + 5.1,$$

namely, at most six additional bits are needed for two sports over the number needed for one. For many instances, let $u = (2+\epsilon)t$. Then

$$R^{(1)} \ge \log(\epsilon t) \qquad \text{but} \qquad R^{(\infty)} \le \log(2+\epsilon) \le 1+\epsilon.$$

For small, but fixed, ϵ and increasing t, a single instance requires arbitrarily many bits while multiple instances require roughly one bit per instance. □

We conclude this subsection with three observations on the optimality of the results.

1. We showed that $\chi^*(K) \le \frac{u}{t}$. We now prove equality. By the Erdős-Ko-Rado Theorem, $\alpha(K) = \binom{u-1}{t-1}$. Hence,

$$\alpha(K^{\vee n}) = \binom{u-1}{t-1}^n,$$

and therefore,

$$\chi(K^{\vee n}) \ge \left(\frac{u}{t}\right)^n,$$

implying that

$$\chi^*(K) = \frac{u}{t}.$$

In fact, the same limit holds for AND products as well. Lovász [14] showed that for every n,

$$\alpha(K^{\wedge n}) = \binom{u-1}{t-1}^n,$$

hence

$$\chi(K^{\wedge n}) \geq \left(\frac{u}{t}\right)^n.$$

Combined with Theorem 4, we obtain

$$\lim_{n\to\infty} (\chi(K^{\wedge n}))^{1/n} = \frac{u}{t}.$$

2. For $t \geq 2$, $\chi^*(K)$ is not achieved by any finite power n. If it were, we would have a partition of $\binom{\{1,\ldots,u\}}{t}^n$ into maximal independent sets in $K^{\vee n}$. Each such set is a Cartesian product $S_1 \times \ldots \times S_n$ where every S_i is a maximal independent set in K, hence consists of all t-element subsets containing a *representative element* z_i. Let $\mathbf{S} = S_1 \times, \ldots, \times S_n$ and $\mathbf{S}' = S'_1 \times, \ldots, \times S'_n$ be two maximal independent sets and let $z_1,\ldots,z_n$ and $z'_1,\ldots,z'_n$ be the representative sequences. For every $i \in \{1,\ldots,n\}$, let T_i be a t-element set containing z_i and z'_i, then $T_1 \times \ldots \times T_n \in \mathbf{S}, \mathbf{S}'$ and therefore $\mathbf{S}$ and $\mathbf{S}'$ intersect.

3. If $\chi(G) > 2$ then

$$\chi^*(G) \geq \frac{2}{1 - 1/|G|} > 2.$$

To see this, note that if G is a cycle of odd length ℓ then $\alpha(G) = \frac{\ell-1}{2}$. Therefore $\alpha(G^{\vee n}) = \left(\frac{\ell-1}{2}\right)^n$ and

$$\chi^*(G) \geq \frac{2}{1 - 1/\ell}$$

(in fact, equality holds). A general G with $\chi(G) > 2$ contains a cycle C of odd length $\ell \leq |G|$. Therefore,

$$\chi^*(G) \geq \chi^*(C) \geq \frac{2}{1 - 1/|G|}.$$

Similarly, Lovász showed that the Shannon capacity of an odd cycle of length ℓ is at most $\ell\cos(\pi/\ell)/(1+\cos(\pi/\ell)) = l(0.5 - \Omega(1/\ell^2))$. It follows that if $\chi(G) > 2$ then

$$\lim_{n\to\infty} (\chi(G^{\wedge n}))^{1/n} \geq 2 + \Omega(1/|G|^2).$$

3.2 Directed line graphs

Let H be a directed graph. The *directed line graph* of H is the graph G whose vertices are the edges of H and where two vertices (a,b) and (c,d) are connected iff $b = c$ or $a = d$. Linial and Vazirani [12] showed that for directed line graphs, $\chi(G^{\wedge 2}) = O(\chi(G))$. We improve the implied constant and extend the colorings to higher graph powers, showing directed line graphs with arbitrarily-high chromatic numbers such that for all $n \in \mathcal{N}_1$,

$$\chi(G^{\wedge n}) \leq \chi(G^{\vee n}) \leq \lceil 2\ln 2 \cdot n4^n \cdot \chi(G)\rceil.$$

We consider directed line graphs derived when H is a complete graph. As described following example 4(b), the arguments can be easily extended to all directed line graphs.

Example 4(a) (Luggage) $P_{\mathcal{Y}}$ is a passenger who recently completed a flight from a to c with a plane change at b, and $P_{\mathcal{X}}$ is the airline's employee who wants to write a letter informing $P_{\mathcal{Y}}$ which of the two flight segments (a, b) or (b, c) his luggage got lost on. How short can $P_{\mathcal{X}}$'s letter be?

We assume: the airline serves t airports; $P_{\mathcal{X}}$ knows only the segment the luggage was lost on (not the other segment); all segments connect two distinct airports; all connections are possible (including those where $a = c$); communication is only from $P_{\mathcal{X}}$ to $P_{\mathcal{Y}}$; both $P_{\mathcal{X}}$ and $P_{\mathcal{Y}}$ know the "official" protocol for communicating lost-luggage segments.

The vertices of the characteristic graph $\mathcal{G}$ are all $t(t-1)$ ordered pairs (a, b) of distinct elements in $\{1,\ldots,t\}$. Vertex (a, b) is connected to all vertices (b, c) and to all vertices (d, a). $\mathcal{G}$ is therefore the directed line graph of the complete graph K_t. We repeat a known proof (see, e.g., Alon [2]), showing that

$$\chi(\mathcal{G}) = m$$

where

$$m \stackrel{\text{def}}{=} \min\{\mu : \tbinom{\mu}{\lceil \mu/2 \rceil} \geq t\} \approx \log t + .5 \log\log t + .5 \log \pi + .5.$$

It follows that for all $t \in \mathcal{N}_2$, the number of bits $P_{\mathcal{X}}$ must convey in the worst case is

$$\log\log t \leq \sigma^{(1)} \leq \log\log t + 1.$$

$\chi(\mathcal{G}) \leq m$: A length-m binary sequence is *balanced* if it contains $\lceil m/2 \rceil$ ones. By choice of m one can assign to each $f \in \{1,\ldots,t\}$ a unique balanced sequence $f_1,\ldots,f_m$. Color vertex (a, b) with the first coordinate j such that $a_j = 0$ and $b_j = 1$. The condition $a_j = 0$, $b_j = 1$ is mutually exclusive of the condition $b_j = 0$, $c_j = 1$, and of the condition $d_j = 0$, $a_j = 1$ guaranteeing that adjacent vertices are assigned different coordinates.

$\chi(\mathcal{G}) \geq m$: let A_a be the set of colors assigned to the vertices (a, b) for $b \in \{1,\ldots,t\} - \{a\}$. If the number of colors is smaller than m, then by Sperner's Lemma [20] there are a and b such that $A_a \subseteq A_b$. But then there is a c such that (a, b) is assigned the same color as (b, c). $\square$

The next example shows that multiple instances require few additional bits. Note that the coloring described is valid also for OR graph powers.

Example 4(b) Consider n independent instances of the dual source in Example 4(a). $P_{\mathcal{Y}}$ takes n connecting flights. All flight combinations are possible. Due to great misfortune, his luggage gets lost on one segment of each flight. $P_{\mathcal{X}}$ wants to inform $P_{\mathcal{Y}}$ all segments where his luggage was lost. How long should the letter be?

The vertices of $\mathcal{G}^{\vee n}$ are all $(t(t-1))^n$ n-tuples $((a_1, b_1),\ldots,(a_n, b_n))$ of ordered pairs of distinct elements in $\{1,\ldots,t\}$. Two distinct vertices $((a_1, b_1),\ldots,(a_n, b_n))$ and $((a'_1, b'_1),\ldots,(a'_n, b'_n))$ are connected if for some i, (a_i, b_i) and (a'_i, b'_i) are distinct and connected in $\mathcal{G}$, namely $a_i = b'_i$ or $b_i = a'_i$

The coloring in Example 4(a) implies that $\chi(\mathcal{G}^{\vee n}) \leq m^n$ where $m \approx \log t$ was defined therein. But arguments similar to those used in Example 3(b) show that

$$\chi(\mathcal{G}^{\wedge n}) \leq \chi(\mathcal{G}^{\vee n}) \leq \lceil 2n4^n \ln t \rceil. \tag{10}$$

In particular, for $n = 2$,

$$\chi(\mathcal{G}^{\wedge 2}) \leq \lceil 64 \ln t \rceil$$

and

$$\chi^*(\mathcal{G}) = \lim_{n\to\infty} \left(\chi(\mathcal{G}^{\vee n})\right)^{1/n} \leq 4.$$

It follows that while arbitrarily many bits are needed for one flight, at most six additional bits are needed for two flights, and asymptotically, at most two bits are needed per flight.

□

Examples 4(a) and 4(b) can be easily extended to directed line graphs G of arbitrary directed graphs H. $\chi(G) \geq \log \chi(H)$ because[6] any proper coloring of G can be converted to a proper coloring of H with subsets of the colors. Simply assign to each vertex the set of all colors of directed edges that emanate from it. $\chi(G^{\vee n}) \leq \lceil 2n4^n \ln \chi(H) \rceil$ follows directly from the colorings described in the examples. Instead of assigning sequences to the vertices of H, assign them to the colors in an optimal coloring of H.

3.3 Self-complementary Cayley graphs that are also Ramsey graphs

We are interested in *high-rate* dual sources where the second instance requires only few additional bits. We prove the following discrepancy between the chromatic number of a graph and its AND square. The ensuing corollary follows immediately.

Theorem 5 For every prime power $v \equiv 1 \bmod 4$ there is a v-vertex graph G such that

$$\chi(G) \geq \frac{v}{(1+o(1))16\log^2 v} \qquad \text{and} \qquad \chi(G^{\wedge 2}) \leq v.$$ □

Corollary 2 For every $\epsilon > 0$ there is a dual source such that

$$\tilde{R}^{(1)} \geq 1 - \epsilon \qquad \text{and} \qquad \tilde{R}^{(\infty)} \leq \tilde{R}^{(2)} \leq \frac{1}{2}.$$ □

For two instances, this discrepancy is essentially the largest possible:

Lemma 4 For every dual source and integer n,

$$\tilde{R}^{(1)} - \tilde{R}^{(n)} \leq \left(1 - \frac{1}{n}\right)\tilde{R}^{(1)}.$$

Proof: For every graph G and integer n, $\chi(G^{\wedge n}) \geq \chi(G)$. Hence $\tilde{R}^{(n)} \geq \frac{\tilde{R}^{(1)}}{n}$ and the lemma follows. □

[6] Colorings of H ignore the orientation of the edges.

Note that while the theorem and its corollary apply only to one and two instances, the implied decrease in the normalized rate is substantially higher than that described in the previous section. There, both $\tilde{R}^{(1)}$ and $\tilde{R}^{(2)}$ tend to 0 as the source's size grows.

The proof of Theorem 5, given in the rest of the section, is based on the existence of self-complementary *Cayley graphs* with small independence numbers. It is similar to the method used by Agarwal, Alon, Aronov, and Suri [1] but requires some additional ideas.

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.

Lemma 5 Every self-complementary Cayley graph G over A has $\chi(G^{\wedge 2}) \leq |A|$.

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

$$c(x, y) \stackrel{\text{def}}{=} x - \pi(y)$$

has range cardinality $|A|$. We show that it colors $G^{\wedge 2}$.

Suppose that $c(x, y) = c(x', y')$. We prove that either $(x, y) = (x', y')$ or $\{(x, y), (x', y')\} \notin G^{\wedge 2}$. By the definition of c,

$$x - x' = \pi(y) - \pi(y'). \tag{11}$$

There are two possibilities. If $x = x'$ then $\pi(y) = \pi(y')$ and the vertices (x, y) and (x', y') coincide. Otherwise, $x \neq x'$, implying that $y \neq y'$, and by (11), the definition of G, and self-complementarity, $\{x, x'\} \in G$ iff $\{\pi(y), \pi(y')\} \in G$ iff $\{y, y'\} \notin G$, showing that (x, y) and (x', y') are not connected in $G^{\wedge 2}$. □

To prove Theorem 5 it therefore suffices to show that for every prime power $v \equiv 1 \bmod 4$, there is a self-complementary Cayley graph over F_v, the finite field of order v, with chromatic number $\geq \frac{v}{(1+o(1))16 \log^2 v}$. This follows once we prove the subsection's main result:

Theorem 6 For every prime power $v \equiv 1 \bmod 4$ there is a self-complementary Cayley graph over F_v with independence number $\leq (1 + o(1))16 \log^2 v$. □

Remark: It would be interesting to determine if the theorem can be strengthened to show that there are self-complementary Cayley graphs of order v with independence number proportional to $\log v$.

The proof resembles that of Lemma 3, but is more involved. A crucial part of Lemma 3's proof relied on the construction of a random graph where each random choice determined at most four edges. At most half the vertex pairs in any set of vertices were "positively correlated with previous pairs" in the sense that if one was an edge, the other was more likely to be one too. For Cayley graphs, every element of K (or lack thereof) determines $|A|$ different edges. Therefore there are many more correlations. We first prove a result whose proof is fashioned after that of Lemma 3, then extend it to the result we need.

A *difference* of a set $T \subseteq A$ is an element $a-b$ such that $a, b \in T$. We let $T-T$ denote the set $\{a-b : a, b \in T\}$ of differences of T. Clearly, $T-T$ contains at most $|T|(|T|-1)$ nonzero elements. If $T-T$ contains $\geq \frac{1}{2}|T|^2$ distinct non-zero elements, it is said to have *many-differences.*

Lemma 6 For every prime power $v \equiv 1 \bmod 4$ there is a self-complementary Cayley graph over F_v containing no $\lceil 4 \log v \rceil$-element independent set with many differences.

Proof: The (cyclic) multiplicative group F_v^* contains an element a of order 4. Let $\{1, a, a^2, a^3\} = \{1, a, -1, -a\}$ be the subgroup generated by a in F_v^*. We construct a random symmetric set $K \subseteq F_v^*$ as follows. For each coset $x \cdot \{1, a, -1, -a\} = \{x, xa, -x, -xa\}$, randomly and independently choose either the two elements x and $-x$ or the two elements xa and $-xa$ to be elements of K, where these two possible choices are equally likely. Let G be the (random) Cayley graph over F_v with respect to K.

G is self complementary since the bijection $\pi : F_v \mapsto F_v$ defined by $\pi(x) = ax$ is an isomorphism between G and its complement. We show that with positive probability G contains no $\lceil 4 \log v \rceil$-element independent set with many differences.

Fix a t-element independent set with many 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 is certainly not independent. Otherwise, the probability that none of the $\geq \frac{1}{4}t^2$ distinct pairs of non-zero elements $x, -x$ in $T-T$ belongs to K is $\leq 2^{-\frac{1}{4}t^2}$.

It follows that the probability that there is a t-element independent set with many differences is $\leq \binom{v}{t} 2^{-\frac{1}{4}t^2}$. Since $t \stackrel{\text{def}}{=} \lceil 4 \log v \rceil$, this probability is < 1. $\square$

To prove Theorem 6, we show that any large set in an Abelian group contains a large subset with many-differences. Therefore, if a graph over an Abelian group contains a large independent set, it contains a large independent set with many-differences. We first prove a simple lemma that provides a slightly weaker estimate than the one needed to prove the theorem, and then improve it to get the assertion of the theorem.

Lemma 7 Every s-element set in an Abelian group of odd order contains a $\lfloor (2s)^{1/3} \rfloor$-element subset whose non-zero differences are all distinct.

Proof: Let S be an s-element set in an Abelian group of odd order and let $\mathcal{N}_0 \ni t \leq (2s)^{1/3} - 1$. We show that for every t-element subset T of S there exists $x \in S \setminus T$ such that $x - T$ is disjoint from $(T-T) \cup (T-x)$. The lemma follows by sequential construction.

For $T \subseteq S$ with $|T| = t$, define $\mathcal{T} \stackrel{\text{def}}{=} \{a+b-c : a, b, c \in T\} \cup \{(a+b)/2 : a, b \in T\}$ where $g/2$ is the element x such that $x + x = g$ (exists and is unique because the group is of odd order). If $x - T$ intersects $(T-T) \cup (T-x)$ then $x \in \mathcal{T}$. But,

$$|\mathcal{T}| \leq t + \binom{t}{2} + t(t-1) + \binom{t}{2}(t-2) = \frac{t(t^2+1)}{2} < s.$$

Hence there exists $x \in S \setminus \mathcal{T}$. Clearly, $x \in S \setminus T$, and $x - T$ is disjoint from $(T-T) \cup (T-x)$. $\square$

The estimate $(2s)^{1/3}$ can be improved to $\Omega(s^{1/3}(\log s)^{1/3})$ using the methods of Alon, Lefmann, and Rődl [3], but it is not known if it can be improved to $\Omega(s^{1/2})$. However, as long as t is around $\sqrt{s}$, S contains a t-element subset with a relatively large number of differences:

Lemma 8 Let S be an s-element set in an Abelian group of odd order. Then for every $t \leq s$, S contains a t-element subset T such that

$$|T - T| \geq t(t-1)\left(1 - \frac{t-2}{s-2} - \frac{(t-2)(t-3)}{2(s-3)}\right) + 1.$$

In particular, if $t = c\sqrt{s}$ for some constant c, then S contains a t-element subset T such that

$$|T - T| \geq t^2\left(1 - \frac{c^2}{2}\right)(1 + o(1)),$$

where, again, the $o(1)$ term diminishes to 0 as s (and t) tend to infinity.

Proof: For a set T, let $n_4(T)$ denote the number of unordered pairs $\{(t_1, t_2), (t_3, t_4)\}$ of ordered pairs of elements of T so that all four elements t_i are distinct and $t_1 - t_2 = t_3 - t_4$. Similarly, let $n_3(T)$ denote the number of ordered triples (t_1, t_2, t_3) of elements of T so that all three elements t_i are distinct and $t_1 - t_2 = t_2 - t_3$.

There are $|T|(|T| - 1)$ ordered pairs (t_1, t_2) of distinct elements of T, and each of them supplies a nonzero element $t_1 - t_2$ of $T - T$. Moreover, it is not difficult to check that if the same group element $t_1 - t_2$ is obtained $r > 1$ times as a difference of this form, then the contribution of pairs containing (t_1, t_2) and of triples containing t_1 and t_2 to $n_4(T) + n_3(T)$ is at least $r - 1$. It follows that the number of distinct nonzero elements in $T - T$ is at least $\eta(T) \stackrel{\text{def}}{=} |T|(|T| - 1) - n_3(T) - n_4(T)$.

Let T be a random t-element subset of S. Then $\eta(T)$ is a random variable with expectation $t(t-1) - E(n_3(T)) - E(n_4(T))$. By linearity of expectation,

$$E(n_3(T)) = n_3(S)\frac{t(t-1)(t-2)}{s(s-1)(s-2)}$$

and

$$E(n_4(T)) = n_4(S)\frac{t(t-1)(t-2)(t-3)}{s(s-1)(s-2)(s-3)}.$$

However, $n_3(S) \leq s(s-1)$, since for any two distinct elements s_1 and s_2 of S there is at most one ordered triple (s_1, s_2, w) that contributes to $n_3(S)$ (as the group is of odd order). Similarly, for every three distinct elements $s_1, s_2, s_3 \in S$, there is at most one unordered pair of the form $\{(s_1, s_2), (s_3, w)\}$ that contributes to $n_4(S)$, and each such pair is counted twice in this manner, implying that $n_4(S) \leq s(s-1)(s-2)/2$. Therefore, the expectation of $\eta(T)$ is at least

$$t(t-1) - s(s-1)\frac{t(t-1)(t-2)}{s(s-1)(s-2)} - \frac{s(s-1)(s-2)}{2}\frac{t(t-1)(t-2)(t-3)}{s(s-1)(s-2)(s-3)}$$

$$= t(t-1)\left(1 - \frac{t-2}{s-2} - \frac{(t-2)(t-3)}{2(s-3)}\right).$$

It follows that there is a choice of a subset T with $\eta(T)$ having at least the above value. Any such T satisfies the conclusion of the lemma. $\square$

We can now prove Theorem 6. By Lemma 8, every $(1+o(1))t^2$-element set in an Abelian group contains a t-element subset with many differences (the $o(1)$ term is easily derived from that in the lemma and diminishes to zero as t tends to infinity). By Lemma 6, there is a self-complementary Cayley graph G over F_v having no $\lceil 4\log v\rceil$-element independent sets with many differences. G cannot contain *any* independent set of size $(1+o(1))16\log^2 v$. If it did, this set would by the above contain a $\lceil 4\log v\rceil$-element subset T with many differences. T, a subset of an independent set, would be independent itself, contradicting the choice of G.

4 Possible improvements and codes correcting certain errors

We have proved that for every $\epsilon > 0$ there is a channel such that $\tilde{C}^{(\infty)} - \tilde{C}^{(1)} \geq \frac{1}{2} - \epsilon$ and a dual source such that $\tilde{R}^{(1)} - \tilde{R}^{(\infty)} \geq \frac{1}{2} - \epsilon$. In an (unsuccessful) attempt to improve these discrepancies, we consider error-correcting codes that can correct only certain kinds of errors.

Throughout this section, v is a prime power, $F(=F_v)$ is the field of order v, $K(=-K)$ is a symmetric subset of F, G is the Cayley graph of (the additive group) F with respect to K, and $G^{\wedge n}$ is the nth AND power of G. F^n is the vector space of n-tuples of elements of F, each a *length-n word over F*. A word over F is a *K-word* if at least one of its coordinates is not in $K\cup\{0\}$, namely, it is nonzero and has a nonzero coordinate which is not in K.

Lemma 9 Let $\overline{x}, \overline{x}' \in F^n$. If $\overline{x} - \overline{x}'$ is a K-word, then $\{\overline{x}, \overline{x}'\} \notin G^{\wedge n}$.

Proof: Let $\overline{x} \stackrel{\text{def}}{=} (x_1, \ldots, x_n)$ and $\overline{x}' \stackrel{\text{def}}{=} (x'_1, \ldots, x'_n)$. Then, there is an $i \in \{1, \ldots, n\}$ such that $0 \neq (x_i - x'_i) \notin K$. Hence $\{x_i, x'_i\} \notin G$, implying that $\{\overline{x}, \overline{x}'\} \notin G^{\wedge n}$. □

A collection of length-n words over F is a *length-n code over F* and each of the words is a *codeword*. The code is *linear* if it is a vector subspace of F^n. A linear code over F has size v^k for some integer $k \geq 0$ called the code's *dimension*. A linear code is a *K-code* if all its nonzero codewords are K-words[7].

Lemma 10 If there is a length-n linear K-code of dimension k then

$$\alpha(G^{\wedge n}) \geq v^k.$$

Proof: We show that every length-n linear K-code is independent (as a set) in $G^{\wedge n}$. Take two distinct codewords. Their difference is a nonzero codeword, hence a K-word. By Lemma 9, they are not connected in $G^{\wedge n}$. □

Using standard linear-algebra, we can strengthen the lemma. Every k-dimensional linear code over F has an $(n-k)\times n$ *parity-check matrix* H such that $\overline{x}$ is a codeword iff $H\cdot\overline{x} = 0$ ($\overline{x}$ is viewed as a column vector).

[7]The definition can be generalized. The *K-weight* of a word is its number of coordinates that are *not* in $K\cup\{0\}$. The *minimum K-weight* of a linear code is the minimum weight of any nonzero codeword. A K-code is therefore a code of minimum K-weight 1. Other minimum K-weights can be considered as well, but are not of use here. Note also that if K is the empty set or $\{0\}$, the minimum K-weight of a code is its minimum distance.

Lemma 11 If there is a length-n linear K-code of dimension k then

$$\chi(G^{\wedge n}) \leq v^{n-k}.$$

Proof: Let H be an $(n-k)\times n$ parity-check matrix for the claimed code. The mapping $\overline{x} \mapsto H\cdot\overline{x}$ has range of cardinality v^{n-k}. We show that it colors $G^{\wedge n}$. If $H\cdot\overline{x} = H\cdot\overline{x}'$ then

$$H\cdot(\overline{x}-\overline{x}') = \overline{0}$$

and therefore $\overline{x}-\overline{x}'$ is a codeword. Lemma 9 says that if $\overline{x}\neq\overline{x}'$, then $\{\overline{x},\overline{x}'\}\notin G^{\wedge n}$ □

Corollary 3 If there is a length-n linear K-code of dimension k then:

1. The channel whose characteristic graph is G has n-use normalized capacity:

$$\tilde{C}^{(n)} \geq \frac{k}{n}.$$

2. The dual source whose characteristic graph is G has n-instance normalized rate:

$$\tilde{R}^{(n)} \leq 1-\frac{k}{n}.$$ □

To establish $\tilde{C}^{(n)}-\tilde{C}^{(1)}$ and $\tilde{R}^{(1)}-\tilde{R}^{(n)}$ discrepancies of $\alpha-\epsilon$ where $0<\alpha\leq 1$ and ϵ is arbitrarily small, it therefore suffices to find a field F of large order v and a symmetric set $K\subseteq F$ such that:

(1) G's independence number is poly-logarithmic in v.

(2) There is a length-n linear K-code of dimension $\geq \alpha n$.

Condition (1) implies that $\tilde{C}^{(1)}\leq\epsilon$ and that $\tilde{R}^{(1)}\geq 1-\epsilon$. Condition (2) and Corollary 3 imply that $\tilde{C}^{(n)}\geq\alpha$ and $\tilde{R}^{(n)}\leq 1-\alpha$.

Example 5 In Subsection 3.3 we proved that $\tilde{R}^{(1)}-\tilde{R}^{(n)}\geq\frac{1}{2}-\epsilon$. For an element a of order 4 in the multiplicative group of F, we (probabilistically) constructed a symmetric set $K\subseteq F$ such that:

(a) G's independence number was $\leq (1+o(1))16\log^2 v$.

(b) For all nonzero x, $\quad x\in K$ iff $ax\notin K$.

Condition (a) implies (1) above. Condition (b) appears slightly different from (2). We used it to infer that G is self complementary, hence $\chi(G^{\wedge 2})\leq v$ and $\tilde{R}^{(2)}\leq\frac{1}{2}$. However, Condition (b) has a simple K-code interpretation. Since $0\neq x\in K$ implies $ax\notin K$, the set $\{(x,ax): x\in F\}$ is a linear K-code. It has length 2 and dimension 1, hence $\tilde{R}^{(2)}\leq\frac{1}{2}$. □

We have not found a symmetric set K satisfying Conditions (1) and (2) for $\alpha>1/2$, and the problem of deciding if such a set exists remains open.

Acknowledgement

We thank Jim Roche for helpful discussions.

References

[1] P.K. Agarwal, N. Alon, B. Aronov, and S. Suri. Can visibility graphs be represented compactly? *Discrete and Computational Geometry*, 12:347–365, 1994.

[2] N. Alon. Monochromatic directed walks in arc colored directed graphs. *Acta Math. Acad. Sci. Hungar.*, 49:163–167, 1987.

[3] N. Alon, H. Lefmann, and V. Rődl. On an anti-ramsey type result. In *Colloq. Math. Soc. János Bolyai 60; Sets, Graphs and Numbers*, pages 9–22. Budapest, Hungary, 1991.

[4] C. Berge and M. Simonovits. The coloring numbers of direct product of two hypergraphs. In C. Berge and D. Ray-Chaudhuri, editors, *Hypergraph Seminar*, Lecture Notes on Mathematics, # 411. Springer Verlag, 1974.

[5] P. Erdős. Some remarks on the theory of graphs. *Bulletin of the American Mathematical Society*, 53:292–294, 1947.

[6] P. Erdős, R.J. McEliece, and H. Taylor. Ramsey bounds for graph products. *Pacific Journal of Mathematics*, 37(1):45–46, 1971.

[7] T. Feder, E. Kushilevitz, and M. Naor. Amortized communication complexity. In *Proceedings of the 32nd Annual Symposium on Foundations of Computer Science*, pages 239–248, 1991.

[8] M.J. Ferguson and D.W. Bailey. Zero-error coding for correlated sources. Unpublished, 1975.

[9] R.L. Graham, B.L. Rothschild, and J.H. Spencer. *Ramsey Theory (Second Edition)*. Wiley Interscience, 1990.

[10] W. Haemers. On some problems of Lovász concerning the Shannon capacity of a graph. *IEEE Transactions on Information Theory*, 25(2):231–232, March 1979.

[11] M. Karchmer, R. Raz, and A. Wigderson. On proving super-logarithmic depth lower bounds via the direct sum in communication complexity. In *Proceedings of the 6th IEEE Symposium on Structure in Complexity Theory*, pages 299–304, 1991.

[12] N. Linial and U. Vazirani. Graph products and chromatic numbers. In *Proceedings of the 30th Annual Symposium on Foundations of Computer Science*, pages 124–128, 1989.

[13] L. Lovász. Kneser's conjecture, chromatic number and homotopy. *Journal of Combinatorial Theory*, 25:319–324, 1978.

[14] L. Lovász. On the Shannon capacity of a graph. *IEEE Transactions on Information Theory*, 25(1):1–7, January 1979.

[15] R.J. McEliece and E.C. Posner. Hide and seek, data storage, and entropy. *The Annals of Mathematical Statistics*, 42(5):1706–1716, 1971.

[16] M. Naor, A. Orlitsky, and P. Shor. Three results on interactive communication. *IEEE Transactions on Information Theory*, 39(5):1608–1615, September 1993.

[17] A. Orlitsky. Worst-case interactive communication I: Two messages are almost optimal. *IEEE Transactions on Information Theory*, 36(5):1111–1126, September 1990.

[18] F. P. Ramsey. On a problem of formal logic. *Proceedings of the London Mathematical Society*, 30(2):264–286, 1929.

[19] C. E. Shannon. The zero-error capacity of a noisy channel. *IRE Transactions on Information Theory*, 2(3):8–19, 1956.

[20] E. Sperner. Ein Satz über Untermengen einer endlichen Menge. *Math. Z.*, 27:544–548, 1928.

[21] H. Witsenhausen. The zero-error side information problem and chromatic numbers. *IEEE Transactions on Information Theory*, 22(5):592–593, September 1976.

[22] A.C. Yao. Some complexity questions related to distributive computing. In *Proceedings of the 11th Annual ACM Symposium on Theory of Computing*, pages 209–213, 1979.