Zero-error capacities and very different sequences

(Preliminary version)

G. Cohen J. Körner* G. Simonyi*

ENST

46, rue Barrault 75634 Paris Cedex 13

FRANCE

Abstract

Perfect hash functions, superimposed codes as well as some other fashionable questions in computer science and random-access communication are special cases of early-day information theoretic models in the zero-error case.

A new class of problems in asymptotic combinatorics can be formulated as the determination of the zero-error capacity of a class of discrete memoryless channels. (This model is also known as the compound channel). We solve an interesting class of these problems using our recent results in polyhedral combinatorics.

1 Introduction

We should like to argue that zero-error cases of classical information theory problems offer a natural language for many known and new problems in asymptotic combinatorics. While Shannon's theory does not always have the solution to these difficult mathematical questions, the very possibility of treating them in a unified manner is a non-negligible advantage.

To illustrate this, let us start with two examples. Here and in the sequel log's are to the base 2.

Example 1

We shall say that the ternary sequences $\mathbf{x} \in \{0,1,2\}^t$ and $\mathbf{x}' \in \{0,1,2\}^t$ are *symmetrically different* if for any two-element subset of $\{0,1,2\}$ there is a coordinate i for which the set $\{x_i, x'_i\}$ is precisely this subset. Let us denote by $N(t)$ the maximum cardinality of a set $\{0,1,2\}^t$ any two sequences of which are symmetrically different. What is

$$\limsup_{t\to\infty} \frac{1}{t} \log N(t)?$$

*On leave from the Mathematics Institute of HAS, 1364 Budapest, POB 127, HUNGARY. Research partially supported by Hungarian National Foundation for Scientific Research grant n^o 1806

Example 2

Let us denote by Q the quaternary alphabet $Q = \{0, 1, 2, 3\}$. Let us say that two quaternary sequences $\mathbf{y} \in Q^t, \mathbf{y}' \in Q^t$ are *well separated* if both

1. there is a coordinate i for which $|y_i - y_i'|$ is odd

2. there is a coordinate j for which $|y_j - y_j'|$ is even (and non-zero).

Let us denote by $P(t)$ the maximum cardinality of a set $D \subset Q^t$ any two sequences of which are well-separated.

Once again, what is

$$\limsup_{t\to\infty} \frac{1}{t} \log P(t)?$$

A common framework to treat these two problems is available, in information theory. A realization of this will immediately furnish non-trivial upper bounds. The information theory bound will be shown to be tight for Example 2. Further, an understanding of the information-theoretic background helps us to relate our present problems to the more established topic of separating partition systems, including qualitatively independent partitions in the sense of Rényi (cf. N. Alon [1], etc. ...).

A *discrete memoryless channel* (DMC) is characterized by a stochastic matrix $W : X \to Y$. Here X is the input alphabet and Y the output alphabet. The rows of the matrix are probability distributions on Y. The different rows are indexed by the different elements of X. Shannon's basic model of information transmission described one-way communication between two terminals, called sender and receiver, respectively. The sender can select an element $x \in X$ of the input alphabet. Correspondingly, the elements of the output alphabet Y will appear at the output randomly, according to the probability distribution $W(.|x)$, the x-th row of the matrix W. Perfect transmission would mean $X = Y$ and $W(y|x) = 1 - \delta_{xy}$ where δ_{xy} is Kronecker's delta. To counterbalance random errors in transmission, the channel can be used repeatedly, to transmit long sequences from X. A *code* is simply an appropriate subset of X^t, the t-length input sequences. Lack of memory is modeled by setting

$$W(\mathbf{y}|\mathbf{x}) = \Pi_{i=1}^t W(y_i|x_i) \tag{1}$$

for $\mathbf{x} = x_1, x_2, \ldots, x_t$, $\mathbf{y} = y_1, y_2, \ldots, y_t$. Here $W(\mathbf{y}|\mathbf{x})$ represents the probability of appearance of the sequence $\mathbf{y} \in Y^t$ at the output provided that $\mathbf{x} \in X^t$ has been selected at the input. Formula (1) means that the conditional probability of seeing y_i in the i-th position of the output sequence depends solely on x_i and no other position of $\mathbf{x}$. The receiver, upon seeing a particular sequence $\mathbf{y} \in Y^t$ has to decide which $\mathbf{x} \in X^t$ has been selected by the sender.

If the code contains several sequences $\mathbf{x} \in X^t$ for which $\mathbf{y} \in Y^t$ has a positive conditional probability, errors will occur in transmission. We will not discuss the various ways of evaluating the probability of such an error. The combinatorial approach is concerned with error-free transmission. Clearly, a subset $C \subset X^t$ is a *code for error-free transmission* over the channel W iff for no two elements $\mathbf{x} \in C$, $\mathbf{x}' \in C$ is there a $\mathbf{y} \in Y^t$ with

$$W(\mathbf{y}|\mathbf{x}) > 0, \quad W(\mathbf{y}|\mathbf{x}') > 0$$

Let $C(t)$ denote the largest cardinality of any such subset of X^t. The quantity

$$\limsup_{t\to\infty} \frac{1}{t} \log C(t)$$

(which can be proved to be a true limit) is called the *zero-error capacity of the discrete memoryless channel* W. Its numerical determination is a tremendous open problem, cf. Lovàsz [2] and Haemers [3]. It is customary to reformulate the above problem in a purely graph theoretic language as we will do somewhat later.

Shannon's model is based on the assumption that the overall statistical behaviour of the communication channel as described by the matrix W is known to the sender and the receiver. Somewhat later, Blackwell, Breiman and Thomasian [4] and Dobrusin [5] have generalized Shannon's original model to the case when W is an unknown member of a set $\mathcal{W}$ of stochastic matrices, each having the same input and output alphabet. In a way this was a trivial generalization and finding a formula for capacity was routine. Not so for zero-error capacity. In fact, somewhat surprisingly, this question has not been raised before.

Let us be formal. Given a finite family $\mathcal{W}$ of stochastic matrices each having input alphabet X and output alphabet Y we shall say that a set $C \subset X^t$ is an *error-free code for the compound discrete memoryless channel* $\mathcal{W}$ if it is an error-free code for each $W \in \mathcal{W}$. Let $C(\mathcal{W}, t)$ denote the largest cardinality of an error-free code for $\mathcal{W}$ with elements in X^t. The quantity

$$C_0(\mathcal{W}) = \limsup_{t\to\infty} \frac{1}{t} \log C(\mathcal{W}, t) \tag{2}$$

is called the *zero-error capacity of the compound channel* $\mathcal{W}$.
Clearly if $|\mathcal{W}| = 1$, we are back to Shannon's zero-capacity. Hence there is little hope to find a nice formula for (2) in the general case.

Although one should stay away from generalizations of unsolved problems in the above spirit, we feel the present case is different. In fact, we would like to show that even in the case when each channel in the family $\mathcal{W}$ is "trivial" in every sense, the determination of $C_0(\mathcal{W})$ is in general a formidable problem. To explain this in more detail, some more information theory is needed. (Our reference is [6]).

Given a DMC $W : X \to Y$ and a probability distribution (PD) P on X let

$$I(P, W) = \sum_{x\in X} P(x) \sum_{y\in Y} W(y|x) \log \frac{W(y|x)}{\sum_{x\in X} P(x)W(y|x)}$$

denote the *mutual information* between an input of distribution P and the corresponding output over W. It is well-known that the capacity of the DMC is $\max I(P,W)$. It is shown in [4] and [5], (cf. also [6]) that the *capacity of the compound channel* $\mathcal{W}$ is

$$C(\mathcal{W}) = \max_P \min_{W \in \mathcal{W}} I(P,W) \tag{3}$$

Since the zero-error capacity is an analogous maximum under more severe criteria, it is an elementary fact from information theory that

$$C_0(\mathcal{W}) \leq C(\mathcal{W}) \tag{4}$$

In order to understand why the zero-error capacity of the compound channel seems to feature an interesting new mathematical problem the intrinsic difficulty of which is, in a sense, "disjoint" from that of the zero-error capacity of a single channel, we should take a look at (3). Without going into tedious technicalities, we would like to explain why (3) is a trivial result. The uninterested reader might skip the rest of this section, after the next definition.

Definition 1 *Given a sequence* $\mathbf{x} \in X^t$, *its type is the probability distribution* $P_{\mathbf{x}}$ *on* X, *defined for every* $a \in X$ *by*

$$P_{\mathbf{x}}(a) = \frac{1}{t}\left|\{i : x_i = a, i = 1, 2, \ldots, t\}\right|$$

Since the capacity $C(\mathcal{W})$ of the compound channel (which has not been defined here) is the asymptotic exponent of the largest cardinality of the codeword set of a nearly error-free code for the channel, it is quite clear that it does not change if we restrict ourselves to codes in which each codeword has the same type, (cf. [6]). Standard Shannon theory shows that if for a sequence of codes the type of the codewords converges to P, then the asymptotic exponent of the largest cardinality of a nearly error-free code is $I(P,W)$ for the DMC W. Hence

$$C(\mathcal{W}) \leq \max_P \min_{W \in \mathcal{W}} I(P,W)$$

follows immediately. (The analogous bound for the zero-error capacity will be derived in a precise manner).

Our main point here is that the tightness of this simple bound is obvious. In fact, it is well-known (cf. [6]) that for a single DMC the best code and the "average code" with the same fixed type of codewords behave in the same asymptotic manner, i.e. "most codes are good". Hence if $|\mathcal{W}| < \infty$ we immediately see that within a fixed type, "most codes are good for every $W \in \mathcal{W}$ at once". Hence,

$$C(\mathcal{W}) \geq \max_P \min_{W \in \mathcal{W}} I(P,W).$$

In other words, the ease with which we proved result (3) is due to the "banal efficiency" of the method of random choice. Now, random choice does not produce good error-free codes and this seems to be at the core of the difficulty we will encounter with the combinatorial model.

2 The graph theory model

Shannon [7] has observed that the determination of the zero-error capacity of a DMC is a purely graph-theoretic problem. He associated with a stochastic matrix W the following graph: let the vertex set of the graph G be X, the input alphabet of W. Let $E(G)$ the edge set of G, consist of those pairs (x, x') of elements from X for which

$$\sum_{y \in Y} W(y|x)W(y|x') = 0$$

In other words, (x, x') constitutes an edge iff x and x' cannot result in the same output with positive probability. (Actually, Shannon's graph is the complement of ours). The information-theoretic problem leads to the following notion of the t-th *power* of G. The graph G^t has vertex set X^t and $(\mathbf{x}, \mathbf{x}') \in E(G^t)$ iff $(x_i, x_j) \in E(G)$ for at least one coordinate $i \in \{1, 2, \ldots, t\}$. Let $\mathcal{H}(G)$ denote the largest cardinality of a complete subgraph of G. It should be clear that $C \subset X^t$ is an error-free code for the DMC W iff the vertices $C \subset X^t$ form a complete subgraph of G^t. Hence the zero-error capacity of W can be defined equivalently as

$$C(G) = \limsup_{t \to \infty} \frac{1}{t} \log \mathcal{H}(G^t).$$

Let us call $C(G)$ the logarithmic capacity of the graph G. More on it can be found in Lovàsz's brilliant paper [2].

Definition 2 *Let $\mathcal{G}$ be a finite family of graphs, each having the same vertex set X. Let $\mathcal{H}(\mathcal{G}, t)$ be the largest cardinality of a subset $C \subset X^t$ for which the vertices of C are a complete subgraph in each graph of $\mathcal{G}$. Then the logarithmic capacity of the family of graphs $\mathcal{G}$ is*

$$C(\mathcal{G}) = \limsup_{t \to \infty} \frac{1}{t} \log \mathcal{H}(\mathcal{G}, t).$$

It should be clear that if we associate a graph G with each of the matrices W of a compound channel $\mathcal{W}$ in the above manner, then for the family $\mathcal{G}$ of graphs so obtained

$$C(\mathcal{G}) = C_0(\mathcal{W})$$

This means that, by (4),

$$C(\mathcal{G}) \leq C(\mathcal{W}).$$

We will derive a better upper bound on $C(\mathcal{G})$.

Let $X^t(P, \varepsilon)$ denote the set of those $\mathbf{x} \in X^t$ for which

$$|P_{\mathbf{x}} - P| < \varepsilon.$$

Let $G^t(P, \varepsilon)$ be the subgraph induced by G^t on $X^t(P, \varepsilon)$. Write

$$C(G, P) := \lim_{\varepsilon \to 0} \limsup_{t \to \infty} \mathcal{H}(G^t(P, \varepsilon)).$$

This quantity was introduced in Csiszàr-Körner [8] and studied in more detail in Marton [9]. Next we state a technical lemma that will be in a complete analogy to the upper bound

$$C(\mathcal{W}) \leq \max_P \min_{W \in \mathcal{W}} I(P, W)$$

for the capacity of the compound channel.

Lemma 1 *Given a family of graphs $\mathcal{G}$ we have*

$$C(\mathcal{G}) \leq \max_P \min_{G \in \mathcal{G}} C(G, P).$$

Proof:

Clearly, the number of possible types of sequences in X^t is upper bounded by $(t+1)^{|X|}$. Let us denote the family of these types by $\mathcal{P}_t$. Then, for every $\varepsilon > 0$,

$$X^t = \bigcup_{P \in \mathcal{P}_t} X^t(P, \varepsilon).$$

This means that, for every $\varepsilon > 0$,

$$\mathcal{N}(\mathcal{G}, t) \leq |\mathcal{P}^t| \max_{P \in \mathcal{P}_t} \min_{G \in \mathcal{G}} \mathcal{N}(G^t(P, \varepsilon)).$$

Hence, for every $\varepsilon > 0$,

$$\frac{1}{t} \log \mathcal{N}(\mathcal{G}, t) \leq \frac{|X| \log(t+1)}{t} + \max_P \frac{1}{t} \log \min_{G \in \mathcal{G}} \mathcal{N}(G^t(P, \varepsilon)).$$

Since $|\mathcal{G}| < \infty$, the lemma follows. $\square$

We are unable to decide whether Lemma 1 is tight. In fact, this is the main difficulty we would like to address in the rest of the paper. For the other problem of trying to convert the bound into a nice formula is not due to us; it has been dealt with extensively in [9], and is not independent of [7], [2] and [3].

At this point, there is little evidence that the bound should be tight. Yet, surprisingly enough, it will be shown to be precise in a non-trivial manner in some disconnected special cases. Typically, in these examples, determination of $C(G, P)$ will be a trivial matter for every G and P. The intriguing part will be to prove the lower bound, i.e. the converse of Lemma 1. This will be done in several different ways, depending on the special case in question.

3 Two complementary graphs

Let G and $\overline{G}$ be the only elements of $\mathcal{G}$. We will put $C(G, \overline{G}) = C(\mathcal{G})$. We have

$$C(G, \overline{G}) \leq \max_P \min[C(G, P), C(\overline{G}, P)]. \tag{5}$$

This bound is not computable. The following computable bound, however, is not always tight.

Theorem 1

$$C(G,\overline{G}) \le \frac{1}{2}\log|V(G)|.$$

Moreover, if $G \cong \overline{G}$, i.e. G is isophormic to its complement, then

$$C(G,\overline{G}) = \frac{1}{2}\log|V(G)|.$$

Proof:

Let us start by the trivial second part of the statement. Suppose we have $G \cong \overline{G}$, and let $\phi : V(G) \to V(\overline{G})$ be a mapping for which

$$(x,x') \in E(G) \Longleftrightarrow (\phi(x),\phi(x')) \in E(\overline{G}).$$

Then, obviously, the vertices $(x,\phi(x))$ form a complete subgraph both in G^2 and $\overline{G}^2$, and hence

$$C(G,\overline{G}) \ge \frac{1}{2}\log|V(G)|.$$

Let us now turn to the first inequality of the Theorem. We shall apply an interesting inequality due to Marton [9]. (Her result is a consequence of recent works on convex corners including an information-theoretic characterization of perfect graphs, cf. Csiszàr-Kőrner-Lovàsz-Marton-Simonyi [10]). A combination of theorems 1 and 2 of Marton [9] shows that for every distribution P on $V(G)$ we have

$$C(G,P) + C(\overline{G},P) \le H(P)$$

for every graph G. Hence, by (5),

$$C(G,\overline{G}) \le \max_P \min[C(G,P), C(\overline{G},P)] \le \max_P \frac{1}{2}[C(G,P) + C(\overline{G},P)]$$

$$\le \max_P \frac{1}{2}H(P) \le \frac{1}{2}\log|V(G)|. \quad \square$$

It is easy to see that the result fails to be tight if G is not self-complementary. In fact, (5) shows that for the graph I having 3 vertices and a single edge

$$C(I,\overline{I}) = \max\{q : h(q) = q\} < 0.78 < \frac{1}{2}\log 3,$$

where $h(q) = -q\log q - (1-q)\log(1-q)$ is the binary entropy.

Unfortunately, we have been unable to determine $C(I,\overline{I})$. A somewhat complicated construction shows that

$$C(I,\overline{I}) \ge 0.71$$

We have no reason to believe that any of these two is tight.

On the other hand, the statement of our Theorem might be sharp also in the case of non-self-complementary graphs. A case in point is our starting Example 2 to which we shall now return. (In fact, it is easily seen that the problem of the example can be reformulated in the form below).

Proposition 1 *Let C_4 be the cycle of length 4. Then*

$$C(C_4, \overline{C}_4) = 1.$$

Proof:

By our Theorem 1,

$$C(C_4, \overline{C}_4) \leq 1.$$

To prove that this bound is actually achievable, let us represent the vertices of C_4 and $\overline{C}_4$ by the elements of $\{0,1\}^2$. Clearly, C_4 can be represented by setting

$$((x,y),(x',y')) \in E(C_4) \text{ iff } x \neq x'.$$

Consequently, $((x,y),(x',y')) \in E(\overline{C}_4)$ iff $x = x', y \neq y'$.
Consider the set $A_t \subset [\{0,1\} \times \{0,1\}]^t$ defined by the rule $((x_1,y_1),(x_2,y_2),\ldots,(x_t,y_t)) \in A_t$ iff

i) $x_1 = 0$

ii) $y_i = x_{i+1}$

where the indices are understood modulo t, i.e. $t+1$ is replaced by 1. Clearly, the elements of A_t are fully determined by the sequence $(x_1, x_2, \ldots, x_t)$. Hence the different sequences from A_t must differ somewhere in their x-coordinates. This means that

$$|A_t| = 2^{t-1}.$$

Further, for any two elements of A_t there is an i such that $x_i \neq x_i'$. Hence, between any two elements of A_t there is an edge in $(C_4)^t$.

We claim that the same is true in $(\overline{C}_4)^t$. In fact, let us look at two different elements of A_t ,$((x_i,y_i),\ldots,(x_t,y_t))$ and $((x_i',y_i'),\ldots,(x_t',y_t'))$, say. Reading the x-parts of the two sequences from left to right, there must be a first coordinate in which they differ. If this is the j-th coordinate, then $j > 1$, by definition. This means that

$$x_{j-1} = x_{j-1}',\ x_j \neq x_j'.$$

Hence $y_{j-1} \neq y_{j-1}'$, and thus

$$((x_{j-1},y_{j-1}),\ (x_{j-1}',y_{j-1}')) \in E(\overline{C}_4),$$

implying that our two sequences are connected by an edge in $(\overline{C}_4)^t$. Hence

$$C(C_4, \overline{C}_4) \geq \lim_{t\to\infty} \frac{1}{t} \log |A_t| = 1. \quad \square$$

The last Proposition gives a construction that has some easy generalizations. In fact, for every p the complete graph on p vertices gives rise to a graph on p^2 vertices in the way C_4 has been constructed here from the complete graph on 2 vertices. Then a similar argument gives an analogous result. We omit the details.

4 Other cases

Another example where we can actually prove the bound of Theorem 1 to be tight arises when we define $\mathcal{G}$ to consist of all the star subgraphs of a complete graph. More formally we have

Proposition 2 *Let X finite and let G_X be the graph for which $V(G_X) = X$, $E(G_X) = \{(x, y) : y \in X, x \neq y\}$. Then for the family $\mathcal{G} = \{G_X, x \in X\}$ we have*

$$C(\mathcal{G}) = h\left(\frac{1}{|X|}\right).$$

Proof:

The upper bound $C(\mathcal{G}) \leq h(1/|X|)$ is obvious. To prove the lower bound, let us consider those values of t which are integral multiples of $|X|$. Write $t = k|X|$, and consider t-length sequences in which every element of X appears exactly k times. If for two such sequences no two elements of X appear at exactly the same k positions, then those two sequences are joined by an edge of $(G_X)^t$ for every $x \in X$.

By Baranyai's theorem [11], there exist $\binom{t-1}{k-1}$ equipartitions of a t-element set such that the classes of the different partitions are all different k-element subsets of the t-element set. Assigning the elements of X to the classes of a partition in an arbitrary manner, we thus obtain $\binom{t-1}{k-1}$ sequences with the desired properties. Hence

$$C(\mathcal{G}) \geq \lim_{t\to\infty} \frac{1}{t} \log \binom{t-1}{\frac{t}{|X|} - 1} = h\left(\frac{1}{|X|}\right). \quad \square$$

Particularly intriguing special cases of our problem arise when all the graphs of $\mathcal{G}$ are edge-disjoint and each of them contains just one edge. Unfortunately, we do not know the exact answer in any but one of these problems. For such a family the edges of the various graphs in the family form a single graph G and the family itself is fully determined by G. Let us denote such a family by $\mathcal{F}(G)$.

Proposition 3 *Let S_r be a star with $(r-1)$ edges. Then*

$$C(\mathcal{F}(S_R)) \leq \max_{\alpha} (1 - (r-2)\alpha)\, h\left(\frac{\alpha}{1 - (r-2)\alpha}\right).$$

Further, this bound is tight for $r = 3$.

Proof:

The upper bound is clear.

Let us turn to the case $r = 3$. The upper bound becomes

$$C(\mathcal{F}(S_3)) \leq \max_{\alpha} (1 - \alpha)\, h\left(\frac{\alpha}{1 - \alpha}\right).$$

It is not hard to see that this bound equals the logarithm of the reciprocal of $(\sqrt{5}-1)/2$, as it is the asymptotic exponent of the number of t-length binary sequences without consecutive ones.

To prove that

$$C(\mathcal{F}(S_3)) \geq \log(2/(\sqrt{5}-1)),$$

let us represent the vertices of S_3 by $\{0,1,2\}$. Let 0 be at the center of the star. It is then clear that the set C_t of those t-length ternary sequences in which every 1 is followed by a 2 has the property that for any $x \in C_t$ there is a coordinate i in which $\{x_j, x'_j\} = \{0,1\}$ and a coordinate j in which $\{x_j, x'_j\} = \{0,2\}$.

Hence

$$C(\mathcal{F}(S_3)) \geq \limsup_{t\to\infty} \frac{1}{t} \log |C_t|.$$

On the other hand, clearly, C_t has asymptotically as many elements as there are binary sequences of length t without consecutive ones. □

Finally, to reveal the extent of our ignorance, we should add that we do not even know the value of $C(\mathcal{F}(K_n))$ where K_n is the complete graph of n vertices. Our upper bound becomes

$$C(\mathcal{F}(K_n)) \leq \frac{2}{n}.$$

Surprisingly, for the cycle of n vertices, our upper bound gives the same value,

$$C(\mathcal{F}(C_n)) \leq \frac{2}{n}.$$

It is, however, quite unlikely that $C(\mathcal{F}(C_n)) = C(\mathcal{F}(K_n))$, if $n > 3$.

As for $n = 3$, this is our Example 1. We can prove

$$0.61 \leq C(\mathcal{F}(K_3)) \leq \frac{2}{3}.$$

Once again, we omit the details.

5 Conclusions

In a recent paper, [12] J. Körner and K. Marton have shown that the problem of perfect hashing can be interpreted as finding the zero-error capacity of a discrete memoryless channel for list codes of some fixed list size, a classical problem in information theory. In another recent paper [13], the connection of perfect hashing to other problems of separating partition systems is pointed out. Foremost among those is the problem of (i,j)-separating systems. A system of bipartitions $P_1, P_2, \ldots P_t$ of an n-element set is said to be an (i,j)-*separating system* if for any disjoint pairs of subsets (A,B) of the n-set with respective size i and j there is at least one partition among the given ones so that A and B are in the two different classes. The problem then is to determine, for given i and j, asymptotically in n, the minimum number t of partitions needed for said purpose. In [13] we have been dealing with lower bounding techniques for such problems. We have failed to point out that qualitatively k-independent bipartitions [1] can be regarded as a system $P_1', P_2', \dots P_t'$ of bipartitions of the n-set such that for any disjoint pair of subsets (A,B) with $|A|+|B|=k$ there is at least one partition among the given ones such that the two sets fall into its different classes. Hence good lower bounds on the minimum size of (i,j)-separating systems might help in dealing with the latter problem.

In conclusion, we would like to point out that the problem of Example 1 is very similar to that of finding the maximum number $N(t)$ of qualitatively independent partitions of a t-set into 3 classes. In fact, if a 3-partition of the t-set is represented by an element of $\{0,1,2\}^t$ in the obvious manner, then the two 3-partitions $\mathbf{x}$ and $\mathbf{x}'$ are qualitatively independent iff for any ordered pair $(a,b) \in \{0,1,2\}^2$ there is a coordinate i such that for the ordered pairs we have $(a,b) = (x_i, x_j')$. This condition is more demanding than that of Example 1 inasmuch here we have two requirements for every unordered pair $\{a,b\}$ with $a \neq b$, plus a requirement for (a,b) with $a = b$, too. The latter, however does not play any role in the asymptotic problem. For qualitatively independent 3-partitions, cf. PolJak-Tuza [14].

References

[1] N. Alon "Explicit construction of exponential sized families of k-independent sets", Discrete Math 58 (1986), pp. 191-193.

[2] L. Lovàsz "On the Shannon capacity of a graph", IEEE Trans. on Information Theory, IT-25 n° 1, pp.1-7 (1979).

[3] W. Haemers "On some problems of Lovàsz concerning the Shannon capacity of a graph", IEEE Trans. on Information Theory, IT-25 nc2, pp. 231-323 (1979).

[4] D. Blackwell, L. Breiman, A.J. Thomasian "The capacity of a class of channels", Annals of Math. Stat. 30 (1959), pp. 1229-1241.

[5] R.L. Dobrushin "Optimal information transfer over a channel with unknown parameters", (in Russian) Radiotechn i Electron. 4 (1959), pp. 1951-1956.

[6] I. Csiszàr, J. Kõrner "Information Theory. Coding Theorems for Discrete Memoryless Systems", Academic Press, New York (1982).

[7] C.E. Shannon "The zero-error capacity of a noisy channel", IRE Trans. on Information Theory 2, pp. 8-19 (1956).

[8] I. Csiszàr, J. Körner "On the capacity of the arbitrarily varying channel for maximum probability of error", Z. Wahrscheinlichkeitstheorie verw, Geb. 57, pp. 87-101 (1981).

[9] K. Marton " On the Shannon capacity of graphs within a given type", preprint.

[10] I. Csiszàr, J. Körner, L. Lovàsz, K. Marton, G. Simonyi "Entropy splitting for antiblocking pairs and perfect graphs", Combinatorica, to appear.

[11] Zs. Baranyai "On the factorization of the complete uniform hypergraph", in Infinite and finite sets, Ed. A. Hajnal, R. Rado, V.T.-Sòs, NorthHolland, Amsterdam-London (1975).

[12] J. Körner, K. Marton "On the capacity of uniform hypergraphs", to be submitted to the IEEE Trans. on Information Theory.

[13] J. Körner, G. Simonyi "Separating partition systems and locally different sequences", SIAM Journal Disc. Math. 1, pp. 351-359 (1988).

[14] S. Poljak, Zs. Tuza "On the maximum number of qualitatively independent partitions", preprint.