The Open University of Israel

Department of Mathematics and Computer Science

Zero vs. ε error in interference channels

Author:
Ilia Levi

Supervisor:
Prof. Michael Langberg

Thesis submitted as partial fulfillment of the requirements towards an M.Sc. degree in Computer Science

The Open University of Israel
Computer Science Division

September 2014

Acknowledgments

This work would not have come into existence if it wasn't for my supervisor, Prof. Michael Langberg. His skillful guidance, patience, and above all his immense dedication and kindness was all I could hope for and more. I am particularly thankful for his support through all stages of research and writing, and not the least, for motivating me in times of uncertainty.

Highly regarded is the contribution of Dr. Danny Vilenchik to this work and the homonymous publication in ITW2013. Our correspondence and collaboration was an enriching and pleasant experience.

A special thanks is due to all my friends and colleagues, whose concern for my studies has pushed me forward.

None of this would be possible but for my parents, who instilled in me the importance of education, and their unflagging belief in me.

Katia, my wife, is the woman behind the scenes. It is her unconditional support, love, small everyday sacrifices, and simply being there that allowed me to complete this work. I am also greatly indebted to my son, Nimrod, for providing inspiration and perspective.

Sincere apologies to anyone inadvertently omitted.

Contents

List of Figures

Abstract

Traditional studies of multi-source, multi-terminal interference channels typically allow a vanishing probability of error in communication. Motivated by the study of network coding, this thesis addresses the task of quantifying the loss in rate when insisting on *zero* error communication in the context of interference channels.

Parts of this work were published in: I. Levi, D. Vilenchik, M. Langberg and M. Effros. *Zero vs. epsilon error in interference channels.* In proceedings of IEEE Information Theory Workshop (ITW), 2013, p. 1-5.

1 Introduction

In the distributed multi-source/multi-terminal network coding paradigm, independent sources wish to convey their information to a set of terminals over a given network $\mathcal{N}$ via a communication scheme in which internal nodes of the network may mix (i.e., encode) the information content of received packets before forwarding them (see e.g., [1, 22, 17, 15, 13] and references therein). In such a communication scheme each terminal eventually receives a certain function of the source information and is required to decode based on the information received. For example, in the multiple-unicast scenario, there are k source/terminal pairs and terminal i is required to decode information of source i.

One may abstractly model the end-to-end behavior of a given multiple-unicast communication scheme by a corresponding k-source/k-terminal interference channel $W : \mathcal{X}^k \to \hat{\mathcal{X}}^k$. Such a channel receives as input the encoded information of the k independent sources $x = x_1, \ldots, x_k \in \mathcal{X}$ and returns as output a vector $\hat{x} = \hat{x}_1, \ldots, \hat{x}_k \in \hat{\mathcal{X}}$, where $\hat{x}_i$ is the information available at terminal node i. As an example, consider the famous butterfly network in Figure 1. The channel W, corresponding to the well known encoding scheme presented in the figure, sets $W(x_1, x_2) = (\hat{x}_1, \hat{x}_2)$ with $\hat{x}_1 = (x_2, x_1 + x_2)$ and $\hat{x}_2 = (x_1, x_1 + x_2)$.

As in the butterfly example, it is common in the network coding literature to assume that the corresponding channel W is *deterministic* (i.e., it is completely determined by the source information) and that communication is successful if all terminals are able to decode the information they require no matter what source information was transmitted. We refer to the latter requirement as *zero error communication*.

The question whether zero error communication poses a restriction on the achievable rate has seen recent interest [5, 20] and has been found in [21, 7] to be closely related to additional intriguing questions in the context of network communication (such as the *edge-removal* problem [12, 16]). Relaxing the requirement of zero error communication with that of $\varepsilon > 0$ error (in which one allows communication to fail with probability ε over the source messages) the following question remains open [5, 20]. [1]

Question 1.1. *Let $\varepsilon > 0$ be a small constant. In the network coding paradigm, can one obtain a strictly higher rate of communication when allowing ε error in communication as opposed to zero error?*

To better understand the price in rate of the zero-error constraint in the context of network coding, in this thesis we study a relaxed version of Question 1.1. Specifically, we view communication via network coding as communication over deterministic interference channels and study the potential gap in rate when communicating with zero error over deterministic interference channels as opposed to $\varepsilon > 0$ error.

Question 1.2. *Let $\varepsilon > 0$ be a small constant. Do there exist deterministic interference channels for which one can obtain a strictly higher rate of communication with ε error as opposed to zero error?*

[1] We note that several statements below will be made informally. Formal definitions and statements will be made in Section 2.

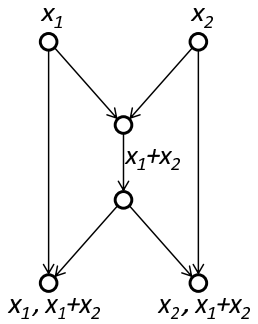


Figure 1: The interference channel corresponding to the butterfly network takes input x_1 and x_2 and returns (x_2, x_1+x_2) to the right terminal (which requires x_1) and (x_1, x_1+x_2) to the left terminal (which requires x_2).

We note that a negative answer to Question 1.2 will imply a negative answer to Question 1.1. However, understanding Q.1.2 will not necessarily resolve Q.1.1. Namely, a positive answer to Q.1.2 will not necessarily imply any consequences on Q.1.1. The latter follows from two main reasons. Primarily, interference channels that result from network communication schemes must take into account the given topology of the network, while the channels that may answer Q.1.2 in the affirmative may not correspond to any given topology. Secondly, to resolve Q.1.1 one must take into account all possible network coding schemes corresponding to a given topology and collection of communication requirements (as the coding scheme that achieves ε error may differ significantly from the best zero error scheme), whereas in Q.1.2 we fix the coding scheme (i.e., interference channel) under study.

1.1 Our contribution

The main focus of the thesis at hand is to better understand Question 1.2 posed above, and in light of its connections with Question 1.1, to gain a better understanding of the tradeoff between $\varepsilon > 0$ and zero error in network coding.

Our work focuses on the 2-source/2-terminal setting. We present and analyze a family of deterministic interference channels $\mathcal{W}$ which we believe can act as witnesses to an affirmative answer of Q.1.2. We do not resolve Q.1.2 in this thesis. Rather, we present what we believe to be evidence in the support of an affirmative answer to Q.1.2 with arbitrarily small values of $\varepsilon > 0$.

In Sections 2 and 3 we present our channel model in detail and define a refined version of Question 1.2 alongside preliminary results and previous work. In Sections 4 and 5 we define the family $\mathcal{W}$ discussed above and show some of its properties. In Section 6 we build upon our results from previous sections to present a positive answer to Q.1.2 assuming a finite communication block length n. In Section 7 we study what we view as a natural approach to refute Q.1.2, and show that it does not necessarily

succeed. In Section 8 we show how our questions and model can be extended to the k-source/k-terminal setting where $k > 2$. Finally we conclude in Section 9.

2 Model

In a multiple unicast communication network the objective is for k source nodes, $s_1, s_2, \ldots, s_k$, to communicate their information to k corresponding terminal nodes, $t_1, t_2, \ldots, t_k$ over a channel W. In this thesis, we will focus on the case of two sources and two terminals (i.e., $k = 2$). A brief discussion regarding our model for larger values of k will appear in Section 8. One can model a deterministic multiple unicast communication network with block length n by the following components (note that the model presented here differs slightly in notation from that presented informally in the Introduction; namely, to simplify notation, encoded source information is denoted by the pair (x, y) and not (x_1, x_2)). Our model is depicted in Figure 2.

Message space: For $i = 1, 2$, source s_i holds a message from a set of size M_i. Without loss of generality, the message space can be considered $[M_i] = \{1, \ldots, M_i\}$.

Encoding: For alphabet $[Q] = [2^q]$ and block length n, each source s_i holds an encoding function $E_i : [M_i] \to [Q]^n$. For binary alphabets ($q = 1$), we will use $\{0, 1\}$ instead of $\{1, 2\} = [2]$. Here, and throughout, we denote the coded information corresponding to source s_1 by $x^{(n)} = (x_1, \ldots, x_n) \in [Q]^n$, and that corresponding to s_2 by $y^{(n)} = (y_1, \ldots, y_n) \in [Q]^n$.

Network W: The network $W : [Q]^2 \to [Q]^2$ is a deterministic function which takes as input elements from $[Q] \times [Q]$ and returns elements from the same alphabet. Namely, denoting W as (W_1, W_2), terminal t_i will receive the evaluation of $W_i : [Q] \times [Q] \to [Q]$ on input $(x, y) \in [Q] \times [Q]$.

Network $W^{(n)}$: Applying the network W multiple times (over block length n) yields the network $W^{(n)} : [Q]^n \times [Q]^n \to [Q]^n \times [Q]^n$ which is a deterministic function that takes as input two n vectors and returns two n vectors. Namely, denoting $W^{(n)}$ as $(W_1^{(n)}, W_2^{(n)})$, the evaluation of $W_1^{(n)} : [Q]^n \times [Q]^n \to [Q]^n$ on input $(x^{(n)}, y^{(n)}) \in [Q]^n \times [Q]^n$, is a vector $\hat{x}^{(n)} \in [Q]^n$ received at terminal t_1, where $\hat{x}_j^{(n)} = W_1(x_j^{(n)}, y_j^{(n)})$. Similarly, $W_2^{(n)}(x^{(n)}, y^{(n)})$ is a vector $\hat{y}^{(n)} \in [Q]^n$ received at terminal t_2, where $\hat{y}_j^{(n)} = W_2(x_j^{(n)}, y_j^{(n)})$.

Decoding: Each terminal t_i holds a decoding function $D_i : [Q]^n \to [M_i]$.

Communication with block length n is successful for terminal t_i and source information (m_1, m_2) if for $i = 1, 2$, $D_i[W_i^{(n)}(E_1(m_1), E_2(m_2))] = m_i$. We say that communication is successful with probability $1 - \varepsilon$ if for source information (m_1, m_2) chosen uniformly at random from $[M_1] \times [M_2]$ it holds with probability $1 - \varepsilon$ that communication is successful for all terminals. Rate (R_1, R_2) is achievable with probability $1 - \varepsilon$ and block length n over network W if for $M_i = 2^{R_i n}$ there exist encoding and decoding functions such that communication is successful with probability $1 - \varepsilon$.

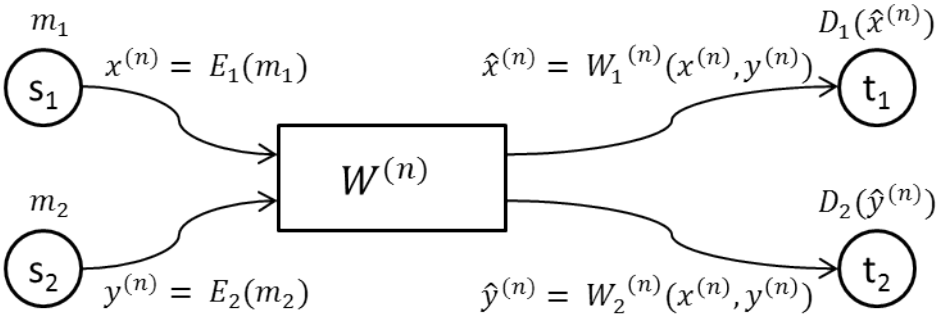


Figure 2: Our model of a two user interference channel. Sources s_1, s_2 communicate messages m_1, m_2 to terminals t_1, t_2 over channel W.

The ε-error sum capacity of network W and block length n is defined to be

$$\mathcal{R}_{W,n}^{(\varepsilon)} = \sup_{(R_1,R_2)\in\Gamma_{n,\varepsilon}} (R_1 + R_2),$$

where the supremum is taken over $\Gamma_{n,\varepsilon}$ which consists of all pairs (R_1, R_2) that are achievable with probability $1-\varepsilon$ and block length n over W. The ε-error sum capacity of network W is defined to be

$$\mathcal{R}_W^{(\varepsilon)} = \sup_n \mathcal{R}_{W,n}^{(\varepsilon)}$$

In particular, for $\varepsilon = 0$, we have $\mathcal{R}_W^{(0)}$. The major question we study in this thesis is the relation between $\mathcal{R}_W^{(\varepsilon)}$ and $\mathcal{R}_W^{(0)}$.

Some remarks are in place. Our model implies independence in encoding (i.e., sources cannot communicate with each other) and independence in decoding (i.e., terminals cannot communicate with each other), which is a commonly used and realistic model. Also notice that W can be defined probabilistically and not deterministically as above. We do not address probabilistic W in this thesis, however one may prove that Question 1.2 has a positive answer in this context: for example, consider the channel W which on input $(x, y) \in [Q] \times [Q]$ returns (x, y) with probability $1-\varepsilon$ and a random pair (x', y') chosen uniformly from $[Q] \times [Q]$ with probability ε. In this case zero-error communication is not possible if we have more than one message per source, and thus $\mathcal{R}_W^{(0)} = 0$. On the other hand, ε-error allows successful communication of any pair $(x, y) \in [Q] \times [Q]$ with probability $1-\varepsilon$, and thus $\mathcal{R}_W^{(\varepsilon)} \geq 2q$.

3 Preliminaries and previous work

Given a channel W, our main interest in this thesis is the relationship between $\mathcal{R}_W^{(0)}$ and $\mathcal{R}_W^{(\varepsilon)}$. In words, $\mathcal{R}_W^{(0)}$ represents the achievable rate when communicating with no error at all, while $\mathcal{R}_W^{(\varepsilon)}$ represents the rate when allowing a small ε probability of error. There are

several examples in communication in which allowing an ε-error significantly increases the rate of communication when compared to zero-error [6]. One specific example is that of Slepian-Wolf [27]. Specifically, we explore the plausibility of the following open statement which claims a large gap between $\mathcal{R}_W^{(\varepsilon)}$ and $\mathcal{R}_W^{(0)}$. The statement below is a refined version of Question 1.2 presented in the Introduction.

Statement 3.1. *Let $\varepsilon > 0$. There exists $\delta = \delta(\varepsilon) > 0$ that tends to 0 when ε tends to 0 such that for every network W it holds that*

$$\frac{\mathcal{R}_W^{(\varepsilon)}}{2} - \delta \leq \mathcal{R}_W^{(0)}.$$

Moreover, for any $\varepsilon > 0$ and δ as above, there exists a network W_ε, such that:

$$\mathcal{R}_{W_\varepsilon}^{(0)} \leq \frac{\mathcal{R}_{W_\varepsilon}^{(\varepsilon)}}{2} + \delta.$$

In other words, for certain networks W, requiring zero-error in communication may reduce the sum capacity by a factor of 2 (or equivalently, allowing an ε error may increase the sum capacity by a factor of 2), and this 2-factor is tight.

We note that it is simple to obtain the first part of Statement 3.1 via a *time sharing* scheme.

Lemma 3.1. *For any $n, \varepsilon > 0$, and $\delta = -\frac{\log(1-\varepsilon)}{n}$, any channel W satisfies*

$$\frac{\mathcal{R}_{W,n}^{(\varepsilon)}}{2} - \delta \leq \mathcal{R}_{W,n}^{(0)}.$$

Here, $\delta > 0$ tends to 0 as ε tends to 0 or n to ∞.

Proof. Let M_1, M_2 be such that $M_i = 2^{R_i n}$, and $\mathcal{R}_{W,n}^{(\varepsilon)} = R_1 + R_2$. W.l.o.g. assume that $R_1 \geq R_2$. Let $E_1,\ E_2,\ D_1,\ D_2$ be the encoders and decoders realizing the sum-mate $\mathcal{R}_{W,n}^{(\varepsilon)}$. By averaging, there exists an $m^* \in M_2$ and a subset $S \subseteq M_1$ of size at least $(1-\varepsilon)|M_1|$, such that for every $m \in S$, $D_1[W_1^{(n)}(E_1(m), E_2(m^*))] = m$ and $D_2[W_2^{(n)}(E_1(m), E_2(m^*))] = m^*$. Otherwise, the overall probability of error would be above ε. Taking $M_1' = S$ and $M_2' = \{m^*\}$, we get a zero-error communication scheme over W_n whose rate is $\frac{1}{n}\left(\log|M_1'| + \log|M_2'|\right) = R_1 + \frac{\log(1-\varepsilon)}{n}$. Since we assumed $R_1 \geq R_2$, we get

$$\mathcal{R}_{W,n}^{(0)} \quad \geq \quad R_1 + \frac{\log(1-\varepsilon)}{n} \geq \frac{\mathcal{R}_{W,n}^{(\varepsilon)}}{2} + \frac{\log(1-\varepsilon)}{n}.$$

□

As a corollary of Lemma 3.1 we get the first part of Statement 3.1: Fix ε, and let $\delta^* = -\log(1-\varepsilon)$. Lemma 3.1 implies that for all n,

$$\sup_n \left(\frac{\mathcal{R}_{W,n}^{(\varepsilon)}}{2} - \delta^* \right) \leq \sup_n \mathcal{R}_{W,n}^{(0)}.$$

Since δ^* does not depend on n, we can take it out of the parenthesis and obtain the first part of Statement 3.1 (recalling the definition $\sup_n \mathcal{R}_{W,n}^{(0)} = \mathcal{R}_W^{(0)}$ and the same for $\mathcal{R}_{W,n}^{(\varepsilon)}$).

3.1 Previous work

The literature on this subject is rather thin. In an excellent survey, Körner and Orlitsky [18] discuss the problem under study, and describe a special case of a 2-user network in which $[Q] = [2] \equiv \{0,1\}$ and $W = (W_1, W_2)$ with

$$W_1(x,y) = \max(x,y), \quad W_2(x,y) = \min(x,y)$$

The problem addressed in [18] is to find $\mathcal{R}_W^{(0)}$. It is not hard to verify that $\mathcal{R}_W^{(1/4)} = 2$. The authors note that this problem has a combinatorial formulation, and that $\mathcal{R}_W^{(0)}$ is conjectured (by [26] and [3]) to be equal to 1 (which matches Statement 3.1 for $\mathcal{R}_W^{(1/4)} = 2$). However, the best upper bound that was proven is $\mathcal{R}_W^{(0)} \leq 1.2118$ ([14]). Note that achieving $\mathcal{R}_W^{(0)} = 1$ is simple by using the network to transmit the information of one user only. E.g., for $n = 1$ define $E_1(0) = 0, E_1(1) = 1, E_2(0) = 0, E_2(1) = 0$ and $D_1(x,y) = x$. Essentially, using a time-sharing scheme, we can convey information to both users, one at a time, with $R_W^{(0)} = 1$. In terms of Statement 3.1, the above shows that:

Theorem 3.1 ([14])**.** *There exists a binary channel W, such that for $\varepsilon = 1/4$:*

$$\mathcal{R}_W^{(0)} \leq 0.6059 \cdot \mathcal{R}_W^{(\varepsilon)}.$$

Our work addresses the potential gap between $\mathcal{R}_W^{(0)}$ and $\mathcal{R}_W^{(\varepsilon)}$ for arbitrary small values of $\varepsilon > 0$.

4 "Erasure/identity" channels

As we have seen, the first part of Statement 3.1 is true. In this thesis, we explore the second part of that statement. We conjecture that it is correct, and provide evidence that supports this conjecture. To this end, we analyze the gap between $\mathcal{R}_W^{(\varepsilon)}$ and $\mathcal{R}_W^{(0)}$ on a family of channels W for which $W : [Q]^2 \to \left([Q]^2 \cup \{(\phi,\phi)\}\right)$ is either the identity function (i.e., $W(x,y) = (x,y)$) or W returns an "erasure value" (i.e., for a new symbol

$\phi \notin [Q]$, $W(x,y) = (\phi,\phi)$). Notice that we change the model slightly by allowing our output alphabet to have an additional symbol. We refer to such channels as *erasure/identity* channels. More specifically, we consider a distribution over erasure/identity channels W, and study the properties of the resulting channels. Our distribution is very natural and is parametrized by ε.

Definition 4.1. *Let $\mathcal{W}_{Q,\varepsilon}$ be the distribution over erasure/identity channels in which for every $(x,y) \in [Q]^2$ we fix $W(x,y) = (\phi,\phi)$ independently with probability ε (otherwise $W(x,y) = (x,y)$).*

Figure 3 presents two possible visualizations for an erasure/identity channel. The first form is a two-dimensional matrix, where a cell $(x,y) \in [Q]^2$ contains ϕ iff $W(x,y) = (\phi,\phi)$ and is empty otherwise (i.e., if $W(x,y) = (x,y)$). The second, more useful form, is given in the following definition.

Definition 4.2. *Let $W : [Q]^2 \to \left([Q]^2 \cup \{(\phi,\phi)\}\right)$ be an erasure/identity channel. We say that a graph $G = (V,E)$ represents channel W if G is a balanced bipartite graph where $V = Q_L \cup Q_R$ and there exist two bijections $f_L : [Q] \to Q_L$, $f_R : [Q] \to Q_R$ such that for any $x,y \in [Q]$, $(f_L(x), f_R(y)) \in E$ if and only if $W(x,y) = (\phi,\phi)$.*

Remark 4.1. *Clearly for any channel $W \in \mathcal{W}_{Q,\varepsilon}$ we can construct a graph G that represents it. In addition, all graphs that represent W are identical up to an isomorphism, thus we can simply denote such a graph by G_W.*

This form of a bipartite graph allows us another useful definition.

Definition 4.3. *Given a bipartite graph $G = (U \cup V, E)$, a bipartite independent set (BPIS) is a pair (A,B), $A \subseteq U$ and $B \subseteq V$, such that for any $(x,y) \in A \times B$, $(x,y) \notin E$. We define the size of the BPIS (A,B) to be $|A||B|$.*

The relationship between the rate $\mathcal{R}_W^{(0)}$ of an erasure/identity channel W and the size of a BPIS in G_W is illustrated by the following proposition.

Proposition 4.1. *Let $W \in \mathcal{W}_{Q,\varepsilon}$, and let (A,B) be a BPIS in G_W. Then:*

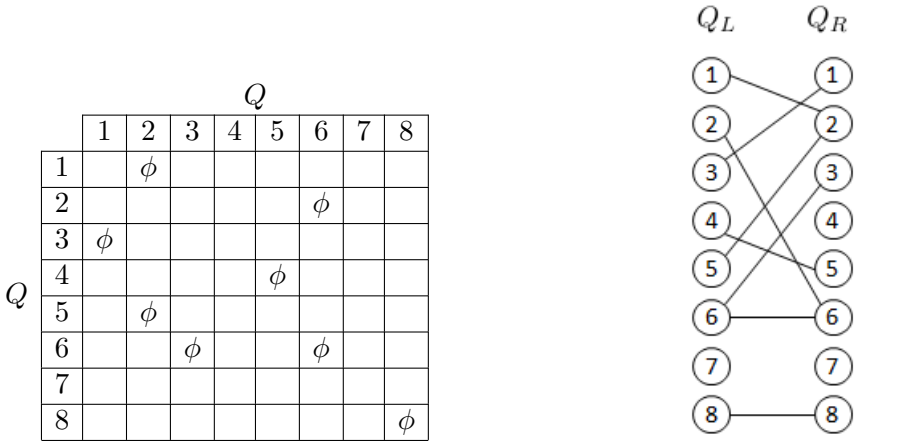
$$\mathcal{R}_W^{(0)} \geq \log\left(|A||B|\right)$$

Proof. Consider

$$Q_A = \{f_L^{-1}(a) \mid a \in A\}$$
$$Q_B = \{f_R^{-1}(b) \mid b \in B\}$$

Where f_L, f_R are the bijections from Definition 4.2. Since f_L, f_R are bijections, $|Q_A| = |A|$ and $|Q_B| = |B|$. Thus setting,

$$M_1 = |A|$$
$$M_2 = |B|$$



(a) matrix form (b) bipartite graph form

Figure 3: Two visualization forms for a channel with $Q = 2^3 = 8$.

and using any bijections

$$E_1 : [M_1] \to Q_A$$
$$E_2 : [M_2] \to Q_B$$

for encoding (and their inverse for decoding), we receive a scheme in which communication is successful for any $(m_1, m_2) \in [M_1] \times [M_2]$ with probability 1. Thus

$$\mathcal{R}^{(0)}_{W,1} \geq \log |A| + \log |B| = \log (|A||B|)$$

and by definition $\mathcal{R}^{(0)}_{W} \geq \mathcal{R}^{(0)}_{W,1}$, concluding the proof. □

We are now ready to explore additional properties of the erasure/identity channels. First, any *typical* channel $W \in \mathcal{W}_{Q,\varepsilon}$ is *almost* the identity function. It only deviates from the identity function on an ε-fraction of input values in expectation, and in such case returns the value (ϕ, ϕ). In addition, the following proposition holds:

Proposition 4.2.

$$\Pr_{W \in \mathcal{W}_{Q,\varepsilon}} \left[\mathcal{R}^{(2\varepsilon)}_{W} \geq 2q \right] \geq \frac{1}{2}$$

Proof. Let us pick a channel W uniformly at random from $\mathcal{W}_{Q,\varepsilon}$. Define a random variable X that counts the number of "erasure" occurrences in channel W. Explicitly,

$$X = \sum_{(i,j) \in [Q] \times [Q]} X_{ij}$$

where $\{X_{ij}\}_{(i,j)\in[Q]\times[Q]}$ are independent indicator variables for the event $W(i,j)=(\phi,\phi)$. Clearly $E[X_{ij}]=\varepsilon$. By linearity of expectation:

$$E[X]=\sum_{(i,j)\in[Q]\times[Q]}E[X_{ij}]=\varepsilon Q^2$$

By Markov's inequality, for any $a>0$:

$$\Pr[X\geq a]\leq\frac{E[X]}{a}$$

In particular, for $a=2\varepsilon Q^2=2E[X]$:

$$\Pr_{W\in\mathcal{W}_{Q,\varepsilon}}\left[X\geq 2\varepsilon Q^2\right]\leq\frac{E[X]}{2\varepsilon Q^2}=\frac{1}{2}$$

Thus:

$$\Pr_{W\in\mathcal{W}_{Q,\varepsilon}}\left[X<2\varepsilon Q^2\right]\geq\frac{1}{2}$$

Note that if, for a given W, the random variable $X<2\varepsilon Q^2$ then for (x,y) chosen uniformly at random from $[Q]\times[Q]$, communication is successful over W with probability at least $1-2\varepsilon$. This means that rate (q,q) is achievable with probability $1-2\varepsilon$, i.e. $\mathcal{R}_{W,1}^{(2\varepsilon)}\geq 2q$. By definition, $\mathcal{R}_{W}^{(2\varepsilon)}\geq\mathcal{R}_{W,1}^{(2\varepsilon)}$ and thus $\mathcal{R}_{W}^{(2\varepsilon)}\geq 2q$, and the proposition follows. □

Proposition 4.2 implies the existence of channels $W\in\mathcal{W}_{Q,\varepsilon}$ for which $\mathcal{R}_{W}^{(2\varepsilon)}\geq 2q$. In light of Statement 3.1, we ask in this case how far $\mathcal{R}_{W}^{(0)}$ is from q. Proving that $\mathcal{R}_{W}^{(0)}\simeq q$ would suffice to prove the gap suggested in Statement 3.1. Thus, the main focus of this thesis is in gaining a better understanding of the channel family $\mathcal{W}_{Q,\varepsilon}$ and specifically studying the question:

Question 4.1. *Is $\mathcal{R}_{W}^{(0)}\simeq q$?*

First of all we note that for parameters Q and ε in which Q is small with respect to ε (e.g., $Q<\frac{1}{\varepsilon}$ or equivalently $\varepsilon<1/Q$) it holds for typical $W\in\mathcal{W}_{Q,\varepsilon}$ that $\mathcal{R}_{W}^{(0)}$ is close to $2q$ (which does not support Statement 3.1). In particular, the following proposition holds:

Proposition 4.3. *Let $\varepsilon\leq\frac{1}{Q}$.*

$$\Pr_{W\in\mathcal{W}_{Q,\varepsilon}}\left[\mathcal{R}_{W}^{(0)}<2q-2\right]\leq e^{-\frac{Q}{12}}$$

Proof. First, let us pick a channel W uniformly at random from $\mathcal{W}_{Q,\varepsilon}$. Define a random variable X that counts the number of edges in G_W. Explicitly,

$$X=\sum_{(i,j)\in[Q]\times[Q]}X_{ij}$$

where $\{X_{ij}\}_{(i,j)\in[Q]\times[Q]}$ are independent indicator variables for the event $(f_L(i), f_R(j)) \in E(G_W)$.

Next, consider the following lemma:

Lemma 4.1. *[9, Lemma 2.2] Let $G = (Q_L \cup Q_R, E)$ be a bipartite graph where $|Q_L| = |Q_R| = Q$. If $|E| \leq \frac{3}{2}Q$ then there exists a BPIS (A, B) such that $|A| = |B| \geq \frac{Q}{2}$.*

Thus, in our case if $X \leq \frac{3}{2}Q$, then there exists a BPIS (A, B) in G such that $|A||B| \geq \frac{Q^2}{4}$. By Proposition 4.1 this means that $\mathcal{R}_W^{(0)} \geq 2q - 2$. It follows that:

$$\Pr_{W\in\mathcal{W}_{Q,\varepsilon}}\left[\mathcal{R}_W^{(0)} < 2q-2\right] \leq \Pr\left[X > \frac{3}{2}Q\right]$$

In order to bound this expression, let us define a random variable Y in the following way:

$$Y = \sum_{(i,j)\in[Q]\times[Q]} Y_{ij}$$

where $\{Y_{ij}\}_{(i,j)\in[Q]\times[Q]}$ are independent indicator variables such that $\Pr[Y_{ij} = 1] = \frac{1}{Q}$.

Note that since $\Pr[X_{ij} = 1] = \varepsilon$ and $\varepsilon \leq \frac{1}{Q}$, for any $a \geq 0$ it holds that $\Pr[X \geq a] \leq \Pr[Y \geq a]$. In particular,

$$\Pr\left[X > \frac{3}{2}Q\right] \leq \Pr\left[Y \geq \frac{3}{2}Q\right]$$

To bound the above we can use the Chernoff bound. Notice that by linearity of expectation:

$$E[Y] = \sum_{(i,j)\in[Q]\times[Q]} E[Y_{ij}] = Q$$

And by the Chernoff bound [23, Theorem 4.4]:

$$\Pr\left[Y \geq \left(1+\frac{1}{2}\right)E[Y]\right] \leq e^{-E[Y]\left(\frac{1}{2}\right)^2/3}$$

That is:

$$\Pr\left[Y \geq \frac{3}{2}Q\right] \leq e^{-\frac{Q}{12}}$$

Putting it all together we get:

$$\begin{aligned}\Pr_{W\in\mathcal{W}_{Q,\varepsilon}}\left[\mathcal{R}_W^{(0)} < 2q-2\right] &\leq \Pr\left[X > \frac{3}{2}Q\right] \\ &\leq \Pr\left[Y \geq \frac{3}{2}Q\right] \\ &\leq e^{-\frac{Q}{12}}\end{aligned}$$

as required. □

4.1 Our results on "erasure/identity" channels

In light of the above properties of our channel model $\mathcal{W}_{Q,\varepsilon}$, for any $\varepsilon > 0$, we focus on values of Q which are large and satisfy $Q \geq \Omega(1/\varepsilon)$. For such values of Q we ask if indeed $\mathcal{R}_W^{(0)}$ is roughly q with high probability (which will imply the existence of channels for the second part of Statement 3.1).

In this thesis we do not prove our conjecture that $\mathcal{R}_W^{(0)} \simeq q$, but rather we present theoretical results in its support. In Section 6 we show (with high probability over $W \in \mathcal{W}_{Q,\varepsilon}$) that for sufficiently large Q, $\mathcal{R}_{W,n}^{(0)} \leq (1 - \frac{1}{n})2q$. This implies the existence of channels W for which on one hand $\mathcal{R}_W^{(2\varepsilon)} \geq 2q$, while on the other $\mathcal{R}_{W,n}^{(0)} \leq (1 - \frac{1}{n})2q$. This is stated formally in Theorem 6.1 and Corollary 6.1. Our results do not have any asymptotic significance since as n grows our upper bound yields the trivial bound of $\mathcal{R}_W^{(0)} \leq 2q$. However for every fixed n we establish a gap between $\mathcal{R}_{W,n}^{(0)}$ and $\mathcal{R}_W^{(2\varepsilon)}$ (which can be viewed as a relaxed version of the second part of Statement 3.1).

In Section 7, we turn to study lower bounds on $\mathcal{R}_W^{(0)}$ obtained by coding schemes which we view to be natural. We study two coding schemes, one scheme based on concatenated coding suggested to us by [24] and another based on certain combinatorial primitives. In both cases, showing that $\mathcal{R}_W^{(0)}$ is larger than q would refute our conjecture that for our channels $\mathcal{R}_W^{(0)} \simeq q$. Indeed, we show that for the specific coding schemes we study, our conjecture still holds. This of course does not rule out the existence of other coding schemes that may refute our conjecture! Nevertheless, we believe our results to be interesting, even if only as a case study that may help in future analysis of "erasure/identity" channels in the context of Statement 3.1. We start by first discussing the graph G_W in greater detail in the upcoming Section 5.

5 BPIS in "erasure/identity" graphs

As we have seen in Proposition 4.1, a BPIS in G_W plays a significant role in establishing a bound on $\mathcal{R}_W^{(0)}$. In this section we introduce the graph $G_{W,n}$ that represents the network $W^{(n)}$, and then analyze the nature of BPIS in G_W and $G_{W,n}$. The results of this analysis will then be used in order to bound $\mathcal{R}_{W,n}^{(0)}$ in Section 6.

We now extend Definition 4.2 for block length n.

Definition 5.1. *Let $W : [Q]^2 \to \left([Q]^2 \cup \{(\phi,\phi)\}\right)$ be an erasure/identity channel, and consider the network $W^{(n)}$. We say that a graph $G = (V, E)$ represents network $W^{(n)}$ if G is a balanced bipartite graph where $V = Q_{L,n} \cup Q_{R,n}$ and there exist two bijections $f_L : [Q]^n \to Q_{L,n}$, $f_R : [Q]^n \to Q_{R,n}$ such that for any $x^{(n)}, y^{(n)} \in [Q]^n$, $(f_L(x^{(n)}), f_R(y^{(n)})) \in E$ if and only if there exists at least one index i such that $W(x_i, y_i) = (\phi, \phi)$.*

Remark 5.1. *Clearly for any network $W^{(n)}$ we can construct a graph G that represents it. In addition, all graphs that represent $W^{(n)}$ are identical up to an isomorphism, thus we can simply denote such a graph by $G_{W,n}$.*

Remark 5.2. *By definition, G_W and $G_{W,1}$ are the same.*

Our interest lies primarily in the maximum size of the BPIS in $G_{W,n}$. The relationship between BPIS in $G_{W,1}$ and $G_{W,n}$ is presented in Section 5.1. In Section 5.2 we establish some upper bounds on the maximum size of various BPIS in $G_{W,n}$, and Section 5.3 discusses the lower bounds in similar cases. Our study addresses two cases. We first study BIPS (A, B) for which $|A| = |B|$, we then address general BPIS and study bounds on the size $|A||B|$. The former case does not directly relate to the rate $\mathcal{R}_W^{(0)}$ (via Proposition 4.1), while the latter study does. Nevertheless, we include its study in this thesis as a detailed analysis of the channel family $\mathcal{W}_{Q,\varepsilon}$.

5.1 Relationship between $G_{W,1}$ and $G_{W,n}$

In this section we establish the following relationship between sizes of BPIS in $G_{W,1}$ and $G_{W,n}$:

Theorem 5.1. *Let $W \in \mathcal{W}_{Q,\varepsilon}$. Then:*

$$\max_{(A',B') \text{ is a BPIS in } G_{W,n}} \left(|A'||B'|\right) = \left(\max_{(A,B) \text{ is a BPIS in } G_{W,1}} (|A||B|) \right)^n$$

Proof. Corollary of Propositions 5.1 and 5.2 given below. □

We note that Theorem 5.1 implies that there is no advantage in studying the BPIS of $G_{W,n}$ over the BPIS of $G_{W,1}$ with respect to the corresponding coding scheme described in Proposition 4.1.

Proposition 5.1. *Let $W \in \mathcal{W}_{Q,\varepsilon}$. If there exists a BPIS (A, B) in G_W such that $|A||B| \geq \alpha$, then there exists a BPIS (A', B') in $G_{W,n}$ such that $|A'||B'| \geq \alpha^n$.*

Proof. Consider $A' = A^n, B' = B^n$. (A', B') is a BPIS because for any $(x^{(n)}, y^{(n)}) \in A' \times B'$ it holds that for any index i, $W(x_i, y_i) \neq (\phi, \phi)$. Thus:

$$|A'||B'| = |A|^n|B|^n = (|A||B|)^n \geq \alpha^n.$$

□

Proposition 5.2. *Let $W \in \mathcal{W}_{Q,\varepsilon}$. If for any BPIS (A, B) in G_W it holds that $|A||B| \leq \alpha$ then for any BPIS (A', B') in $G_{W,n}$ it holds that $|A'||B'| \leq \alpha^n$.*

Proof. By induction on n. For $n = 1$ it holds trivially. Suppose the relationship holds for all $l < n$, and let (A', B') be some BPIS in $G_{W,n}$. Consider some index set I, such that $1 \leq |I| < n$. For a vector $s^{(n)} \in Q^n$, denote by s_I the vector obtained by the projection of $s^{(n)}$ onto I. Similarly, for a set $S \subseteq Q^n$, denote by S_I the set obtained by projection of S onto I. In addition, for a set S and some $p \in S_I$, let

$$S_{I,p} = \{s_{[n]-I} \mid s \in S, s_I = p\}$$

Namely, $S_{I,p}$ is the set obtained by projection of S onto the complement of I and selection of elements that, when merged with p, appear in S.

Next, we claim that for any set $S \subseteq Q^n$ and index set I (where $1 \leq |I| < n$), there exists a $p \in S_I$ such that:

$$|S_{I,p}| \geq \frac{|S|}{|S_I|}.$$

As otherwise, we get

$$\sum_{p \in S_I} |S_{I,p}| < |S_I| \cdot \frac{|S|}{|S_I|} = |S|$$

which is not possible by our definition of $S_{I,p}$. Thus, let $a \in A'_I$ and $b \in B'_I$ be such that:

$$|A'_{I,a}| \geq \frac{|A'|}{|A'_I|}$$
$$|B'_{I,b}| \geq \frac{|B'|}{|B'_I|}$$

Note that (A'_I, B'_I) and $(A'_{I,a}, B'_{I,b})$ are BPIS in $G_{W,|I|}$ and $G_{W,n-|I|}$ respectively. Thus by the induction hypothesis:

$$\begin{aligned} |A'_I||B'_I| &\leq \alpha^{|I|} \\ |A'_{I,a}||B'_{I,b}| &\leq \alpha^{n-|I|} \end{aligned}$$

Finally,

$$\begin{aligned} |A'||B'| &= \left(|A'_I| \cdot \frac{|A'|}{|A'_I|}\right)\left(|B'_I| \cdot \frac{|B'|}{|B'_I|}\right) \\ &= \left(|A'_I||B'_I|\right)\left(\frac{|A'|}{|A'_I|} \cdot \frac{|B'|}{|B'_I|}\right) \\ &\leq \left(|A'_I||B'_I|\right)\left(|A'_{I,a}||B'_{I,b}|\right) \\ &\leq \alpha^{|I|}\alpha^{n-|I|} \\ &= \alpha^n \end{aligned}$$

as required. □

5.2 Probabilistic upper bounds for BPIS in $G_{W,1}$ (and thus also in $G_{W,n}$)

In this section we show probabilistic upper bounds on the maximum size of BPIS in $G_{W,n}$. Proposition 5.3 considers the case of a balanced BPIS in $G_{W,1}$. Proposition 5.4 considers any BPIS in $G_{W,1}$. Finally, Theorem 5.2 considers any BPIS in $G_{W,n}$. In the analysis below, recall that $Q = 2^q$.

Proposition 5.3. *Let $W \in \mathcal{W}_{Q,\varepsilon}$. The probability that there exists a BPIS (A, B) in G_W such that $|A| = |B| = cq$ where $c \geq \frac{3}{\log \frac{1}{1-\varepsilon}}$ is at most $\left(\frac{1}{2}\right)^{cq^2}$.*

Proof. Let τ be the event that there exists a BPIS (A, B) such that $|A| = |B| = cq$. For fixed A, B of size cq each, let $\tau_{A,B}$ be the event that (A, B) is a BPIS. Clearly, $\Pr[\tau_{A,B}] = (1-\varepsilon)^{(cq)^2}$. By the union bound:

$$\begin{aligned}
\Pr[\tau] &= \Pr\left[\bigcup_{\substack{A,B \\ |A|=|B|=cq}} \tau_{A,B}\right] \\
&\leq \sum_{\substack{A,B \\ |A|=|B|=cq}} \Pr[\tau_{A,B}] \\
&= \binom{2^q}{cq}^2 \cdot (1-\varepsilon)^{(cq)^2} \\
&\leq ((2^q)^{cq})^2 \cdot (1-\varepsilon)^{(cq)^2} \\
&= 2^{2cq^2} \cdot (1-\varepsilon)^{(cq)^2} \\
&= 4^{cq^2} \cdot ((1-\varepsilon)^c)^{cq^2} \\
&= (4 \cdot (1-\varepsilon)^c)^{cq^2}
\end{aligned}$$

Note that $c \geq \frac{3}{\log \frac{1}{1-\varepsilon}}$, which means that:

$$\begin{aligned}
c \log(1-\epsilon) &\leq -3 \\
2^{c \log(1-\epsilon)} &\leq 2^{-3} \\
(1-\epsilon)^c &\leq \frac{1}{8}
\end{aligned}$$

Substituting this bound in the probability bound above, we get that

$$(4 \cdot (1-\epsilon)^c)^{cq^2} \leq \left(\frac{1}{2}\right)^{cq^2}$$

as required. □

Proposition 5.4. *Let $W \in \mathcal{W}_{Q,\varepsilon}$. The probability that there exists a BPIS (A, B) in G_W such that $\log|A||B| > q + \log c$ where $c \geq \frac{4}{\log \frac{1}{1-\varepsilon}}$ is at most $\left(\frac{1}{2}\right)^Q$.*

Proof. We use the union bound, much like in our proof for Proposition 5.3. The probability of finding such a BPIS (A, B) is thus bounded by:

$$\sum_{\substack{|A|,|B| \\ \log|A||B| > q + \log c}} \left(\binom{2^q}{|A|}\binom{2^q}{|B|}(1-\epsilon)^{|A||B|}\right)$$

But since $|A|, |B| \leq 2^q$ this is in turn bounded by:

$$\begin{aligned} 2^{2q} \cdot \max_{\substack{|A|,|B| \\ \log|A||B| > q + \log c}} \left(\binom{2^q}{|A|} \binom{2^q}{|B|} (1-\epsilon)^{|A||B|} \right) &\leq 2^{2q} \cdot \left(2^{2^q} \cdot 2^{2^q} (1-\epsilon)^{2^{q+\log c}} \right) \\ &\leq 2^{2^q} \cdot \left(2^{2^q} \cdot 2^{2^q} (1-\epsilon)^{2^{q+\log c}} \right) \\ &\leq (2^3)^{2^q} \cdot (1-\epsilon)^{2^{q+\log c}} \\ &\leq (8 \cdot (1-\epsilon)^c)^{2^q} \end{aligned}$$

Note that $c \geq \frac{4}{\log \frac{1}{1-\varepsilon}}$, which means that:

$$\begin{aligned} c \log(1-\epsilon) &\leq -4 \\ 2^{c \log(1-\epsilon)} &\leq 2^{-4} \\ (1-\epsilon)^c &\leq \frac{1}{16} \end{aligned}$$

Substituting this bound in the probability bound above, we get that

$$(8 \cdot (1-\epsilon)^c)^{2^q} \leq \left(\frac{1}{2} \right)^{2^q} = \left(\frac{1}{2} \right)^Q$$

as required. □

Combining Proposition 5.4 and Proposition 5.2 yields:

Theorem 5.2. *Let $W \in \mathcal{W}_{Q,\varepsilon}$. The probability that there exists a BPIS (A, B) in $G_{W,n}$ such that $\frac{1}{n} \log|A||B| > q + \log c$ where $c \geq \frac{4}{\log \frac{1}{1-\varepsilon}}$ is at most $\left(\frac{1}{2}\right)^Q$.*

Proof. By Proposition 5.2 if there exists a BPIS (A', B') in $G_{W,n}$ such that $\frac{1}{n} \log|A'||B'| > q + \log c$ then there exists a BPIS (A, B) in $G_{W,1}$ such that $\log|A||B| > q + \log c$. By Proposition 5.4 the probability of the latter is at most $\left(\frac{1}{2}\right)^Q$. It follows that the probability in question is also at most $\left(\frac{1}{2}\right)^Q$. □

Theorem 5.2 can also be restated in the following, weaker but simplified form which we use later in our proof:

Theorem 5.3. *Let $W \in \mathcal{W}_{Q,\varepsilon}$. The probability that there exists a BPIS (A, B) in $G_{W,n}$ such that $\frac{1}{n} \log|A||B| > q + \log \frac{4 \ln 2}{\varepsilon}$ is at most $\left(\frac{1}{2}\right)^Q$.*

Proof. This is an immediate consequence of Theorem 5.2 and Lemma A.1 of the Appendix. □

5.3 Probabilistic lower bounds for BPIS in $G_{W,1}$ (and thus also in $G_{W,n}$)

In this section we show probabilistic lower bounds on the maximum size of BPIS in $G_{W,n}$. Proposition 5.5 considers the case of a balanced BPIS in $G_{W,1}$. Proposition 5.6 considers any BPIS in $G_{W,1}$. Finally, Theorem 5.4 considers any BPIS in $G_{W,n}$. We start with a "converse" to Proposition 5.3.

Proposition 5.5. *Let $W \in \mathcal{W}_{Q,\varepsilon}$ where ε is such that $\varepsilon \leq \frac{1}{2(1+\delta)}$ for some constant $0 < \delta < 1$. For a sufficiently large Q (w.r.t. $1/\varepsilon$), the probability that there exists a BPIS (A, B) in G_W such that $|A| = |B| = \lceil \frac{\delta(q+\log\varepsilon+1)}{4\log(e)\varepsilon} \rceil$ is at least $1 - 2e^{-\varepsilon Q/8}$.*

Proof. In this proof we follow the idea of Feige & Kogan [10, Lemma 3.4]. Given the graph $G_W = (Q_L \cup Q_R, E)$, let us pick uniformly at random a set B of k (which will be defined later) vertices from Q_R, and let X be the random variable which counts the number of vertices in Q_L which have no neighbors in B. If $E[X] \geq k$ then there must exist a set $A \subseteq Q_L$ of size k, such that (A, B) is a BPIS (here, the expectation is over the random choices of B). All that is left is to show is that for some $k \geq \lceil \frac{\delta(q+\log\varepsilon+1)}{4\log(e)\varepsilon} \rceil$ and a sufficiently large Q, $\Pr_{W \in \mathcal{W}_{Q,\varepsilon}}[E[X] \geq k] \geq 1 - 2e^{-\varepsilon Q/8}$.

Assume that the vertices in Q_L are $u_1, \ldots, u_Q$ and that their degrees are $d_1, \ldots, d_Q$. Suppose that the average degree d in G_W is such that $\varepsilon Q/2 \leq d \leq 2\varepsilon Q$. We define $k = \lceil \frac{\delta Q \ln d}{2d} \rceil$. Notice that since $d \leq 2\varepsilon Q$:

$$k = \left\lceil \frac{\delta Q \ln d}{2d} \right\rceil \geq \left\lceil \frac{\delta Q \ln(2\varepsilon Q)}{2(2\varepsilon Q)} \right\rceil = \left\lceil \frac{\delta Q \frac{\log(2\varepsilon Q)}{\log(e)}}{2(2\varepsilon Q)} \right\rceil = \left\lceil \frac{\delta\left(q + \log\varepsilon + 1\right)}{4\log(e)\varepsilon} \right\rceil$$

The probability that vertex u_i has no neighbors in B is $\binom{Q-d_i}{k} / \binom{Q}{k}$. Note that the sequence $\left(\binom{Q-x}{y}\right)_{x=0}^{Q-y}$ is convex. To see this, one needs to verify that for any integer $0 \leq x \leq Q - y - 2$:

$$2\binom{Q-(x+1)}{y} \leq \binom{Q-x}{y} + \binom{Q-(x+2)}{y}$$

And indeed this holds if and only if:

$$\begin{aligned}
2\frac{(Q-x-1)!}{(y)!(Q-x-y-1)!} &\leq \frac{(Q-x)!}{(y)!(Q-x-y)!} + \frac{(Q-x-2)!}{(y)!(Q-x-y-2)!} \\
2\frac{(Q-x-1)!}{(Q-x-y-1)} &\leq \frac{(Q-x)!}{(Q-x-y)(Q-x-y-1)} + (Q-x-2)! \\
2(Q-x-y)(Q-x-1)! &\leq (Q-x)! + (Q-x-y)(Q-x-y-1)(Q-x-2)! \\
2(Q-x-y)(Q-x-1) &\leq (Q-x)(Q-x-1) + (Q-x-y)(Q-x-y-1) \\
2((Q-x)(Q-x-1) - y(Q-x) + y) &\leq (Q-x)(Q-x-1) + (Q-x)(Q-x-y-1) - y(Q-x) \\
&\quad +y(y+1) \\
2((Q-x)(Q-x-y-1) + y) &\leq (Q-x)(2Q-2x-2y-2) + y(y+1) \\
2y &\leq y(y+1) \\
y &\leq y^2
\end{aligned}$$

Which holds for all $y \geq 1$, and in our case for all k-s of interest. Therefore:

$$E[X] = \sum_{i=1}^{Q} \frac{\binom{Q-d_i}{k}}{\binom{Q}{k}} \geq Q\frac{\binom{Q-d}{k}}{\binom{Q}{k}}$$

Notice that:

$$\begin{aligned}
\frac{\binom{Q-d}{k}}{\binom{Q}{k}} &= \frac{(Q-d)!}{k!(Q-d-k)!} \Big/ \frac{Q!}{k!(Q-k)!} \\
&= \frac{(Q-d)! \cdot (Q-k)!}{(Q-d-k)! \cdot Q!} \\
&= \frac{\left(\prod_{i=0}^{k-1}(Q-d-i)\right) \cdot (Q-k)!}{Q!} \\
&= \frac{\left(\prod_{i=0}^{k-1}(Q-d-i)\right)}{\left(\prod_{i=0}^{k-1}(Q-i)\right)} \\
&= \prod_{i=0}^{k-1}\left(\frac{Q-d-i}{Q-i}\right)
\end{aligned}$$

Thus:

$$E[X] \geq Q\frac{\binom{Q-d}{k}}{\binom{Q}{k}} = Q\prod_{i=0}^{k-1}\left(\frac{Q-d-i}{Q-i}\right)$$

It is thus enough to prove that for a sufficiently large d:

$$\prod_{i=0}^{k-1}\left(\frac{Q-d-i}{Q-i}\right) \geq \frac{k}{Q}$$

Or alternatively that:

$$\frac{Q}{k} \geq \prod_{i=0}^{k-1} \left(\frac{Q-i}{Q-d-i} \right) \tag{1}$$

Note that a sufficiently large d can be obtained by choosing a sufficiently large Q (w.r.t. $1/\varepsilon$), since by our assumptions $d \geq \varepsilon Q/2$.

First, let us bound the RHS of inequality 1 from above. Notice that:

$$\begin{aligned}
\ln\left(\prod_{i=0}^{k-1}\left(\frac{Q-i}{Q-d-i}\right)\right) &= \sum_{i=0}^{k-1} \ln\left(1+\frac{d}{Q-d-i}\right) \\
&\leq k\ln\left(1+\frac{d}{Q-d-k+1}\right) \\
&\leq \frac{dk}{Q-d-k+1} && \text{as } \ln(1+x)\leq x \text{ for } x\geq 0 \\
&\leq \frac{d\left(\frac{\delta Q\ln d}{2d}+1\right)}{Q-d-\frac{\delta Q\ln d}{2d}} && \text{as } k \leq \frac{\delta Q\ln d}{2d}+1 \\
&= \frac{Q\left(\frac{\delta\ln d}{2}+\frac{d}{Q}\right)}{Q\left(1-\frac{d}{Q}-\frac{\delta\ln d}{2d}\right)} \\
&= \frac{\frac{\delta\ln d}{2}+\frac{d}{Q}}{1-\frac{d}{Q}-\frac{\delta\ln d}{2d}} \\
&\leq \frac{\frac{\delta\ln d}{2}+\frac{1}{1+\delta}}{1-\frac{1}{1+\delta}-\frac{\delta\ln d}{2d}} && \text{as } \frac{d}{Q}\leq\frac{2\varepsilon Q}{Q}=2\varepsilon\leq\frac{1}{1+\delta} \\
&= \frac{\frac{\delta\ln d}{2}+\frac{1}{1+\delta}}{\frac{\delta}{1+\delta}-\frac{\delta\ln d}{2d}} \\
&= \frac{\frac{(1+\delta)\ln d}{2}+\frac{1}{\delta}}{1-\frac{(1+\delta)\ln d}{2d}} && \text{as a result of multiplication by } \frac{1+\delta}{\delta} \\
&\leq (1-\alpha)\ln d + O(1) && \text{for some constant } 0<\alpha<1 \text{ and large enough } d
\end{aligned}$$

Thus we get:

$$\prod_{i=0}^{k-1}\left(\frac{Q-i}{Q-d-i}\right) = O(d^{1-\alpha}) \tag{2}$$

Next, let us bound the LHS of inequality 1 from below.

$$\begin{aligned}
\frac{Q}{k} &\geq \frac{Q}{\frac{\delta Q \ln d}{2d} + 1} && \text{as } k \leq \frac{\delta Q \ln d}{2d} + 1 \\
&= \frac{1}{\frac{\delta \ln d}{2d} + \frac{1}{Q}} \\
&= \frac{\frac{2}{\delta} d}{\ln d + \frac{2d}{\delta Q}} && \text{as a result of multiplication by } \frac{2d}{\delta} \\
&\geq \frac{\frac{2}{\delta} d}{\ln d + \frac{2}{\delta(1+\delta)}} && \text{as } \frac{d}{Q} \leq \frac{2\varepsilon Q}{Q} = 2\varepsilon \leq \frac{1}{1+\delta}
\end{aligned}$$

Thus we get:

$$\frac{Q}{k} = \Omega(\frac{d}{\ln d}) \tag{3}$$

Inequality 1 follows from equations 2 and 3 for large enough d.

Finally, recall that for a vertex $u \in Q_L \cup Q_R$ it holds that $E[d(u)] = \varepsilon Q$. Since $d = \frac{\sum_{u \in Q_L \cup Q_R} d(u)}{2Q}$, by linearity of expectation, $E[d] = \varepsilon Q$ as well. Thus, by the Chernoff bound [23, Theorem 4.4-4.5]:

$$\begin{aligned}
\Pr_{W \in \mathcal{W}_{Q,\varepsilon}} [d \leq \varepsilon Q/2] &\leq e^{-\varepsilon Q/8} \\
\Pr_{W \in \mathcal{W}_{Q,\varepsilon}} [d \geq 2\varepsilon Q] &\leq e^{-\varepsilon Q/3}
\end{aligned}$$

By the union bound:

$$\Pr_{W \in \mathcal{W}_{Q,\varepsilon}} [d \leq \varepsilon Q/2 \vee d \geq 2\varepsilon Q] \leq e^{-\varepsilon Q/8} + e^{-\varepsilon Q/3} \leq 2e^{-\varepsilon Q/8}$$

And thus:

$$\Pr_{W \in \mathcal{W}_{Q,\varepsilon}} [E[X] \geq k] \geq \Pr_{W \in \mathcal{W}_{Q,\varepsilon}} [\varepsilon Q/2 \leq d \leq 2\varepsilon Q] \geq 1 - 2e^{-\varepsilon Q/8}$$

□

With regards to lower bounds on the size of a BPIS in G_W, a simple analysis gives the following proposition.

Proposition 5.6. *Let $W \in \mathcal{W}_{Q,\varepsilon}$. The probability that there exists a BPIS (A, B) in G_W such that $\log |A||B| > q - (1 + \log \frac{1}{1-\varepsilon})$ is at least $1 - e^{-(1-\varepsilon)Q/8}$.*

Proof. Given the graph $G_W = (Q_L \cup Q_R, E)$, set A to have one vertex $v \in Q_L$, and let X be the random variable which counts the number of vertices in $B = \{w \in Q_R \mid$

$(v,w) \notin E\}$. Clearly, (A,B) is a BPIS. Note that $E[X] = (1-\varepsilon)Q = (1-\varepsilon)2^q$. Thus, by Chernoff bound [23, Theorem 4.5]:

$$\Pr_{W \in \mathcal{W}_{Q,\varepsilon}}[X \leq (1-\varepsilon)Q/2] \leq e^{-(1-\varepsilon)Q/8}$$

Therefore:

$$\Pr_{W \in \mathcal{W}_{Q,\varepsilon}}[X > (1-\varepsilon)Q/2] \geq 1 - e^{-(1-\varepsilon)Q/8}$$

Namely that with probability at least $1 - e^{-(1-\varepsilon)Q/8}$ it holds that:

$$\log|A||B| > \log((1-\varepsilon)Q/2) = q - 1 + \log(1-\varepsilon) = q - \left(1 + \log\frac{1}{1-\varepsilon}\right).$$

□

Combining Propositions 5.6 and 5.1 yields:

Theorem 5.4. *Let $W \in \mathcal{W}_{Q,\varepsilon}$. The probability that there exists a BPIS (A,B) in $G_{W,n}$ such that $\frac{1}{n}\log|A||B| > q - (1 + \log\frac{1}{1-\varepsilon})$ is at least $1 - e^{-(1-\varepsilon)Q/8}$.*

Remark 5.3. *All in all, we note that the results of the last two subsections together with Proposition 4.1 imply that a zero error code for (most) channels $W \in \mathcal{W}_{Q,\varepsilon}$ based on BPIS will have rate $\mathcal{R}_W^{(0)} \simeq q$.*

6 Upper bounds on $\mathcal{R}_{W,n}^{(0)}$ (i.e., for fixed block length)

In this section, we present an upper bound on the rate $\mathcal{R}_{W,n}^{(0)}$ for channels W chosen from the aforementioned distribution $\mathcal{W}_{Q,\varepsilon}$. As discussed, for any error value $\varepsilon > 0$, we study the distribution $\mathcal{W}_{Q,\varepsilon}$ for values of Q which are sufficiently large and satisfy $Q = \Omega(1/\varepsilon)$. Specifically, we prove:

Theorem 6.1. *For every integer $n \geq 2$, $\varepsilon \in [0,1]$ and $\gamma > 0$, let $Q = 2^q$ with $q \geq \max\{\log n, \frac{4}{\gamma}\log\frac{4\ln 2}{\varepsilon}\}$. Then with probability at least $3/4$, a random channel $W \in \mathcal{W}_{Q,\varepsilon}$ satisfies*

$$\mathcal{R}_{W,n}^{(0)} \leq 2q\left(1 - \frac{1}{n}\right)(1+\gamma). \tag{4}$$

Specifically, for $n = 2$

$$\mathcal{R}_{W,2}^{(0)} \leq q(1+\gamma).$$

In words, Theorem 6.1 states that the gap suggested in Statement 3.1 indeed holds if we restrict ourselves to block length $n = 2$, and a smaller gap holds for increasing block length. For asymptotically large block length n the bound of Theorem 6.1 approaches $2q$ which is not enough to support Statement 3.1.

Our proof of Theorem 6.1 will follow the statement and proof of a number of technical claims given below. Recall that Proposition 4.1 derived a lower bound on the rate of the channel $W \in \mathcal{W}_{Q,\varepsilon}$, given a lower bound on the size of some BPIS. We start by proving the following lemma that demonstrates the *inverse* claim.

Lemma 6.1. *Let $W \in \mathcal{W}_{Q,\varepsilon}$. If $\mathcal{R}^{(0)}_{W,n} = r$, then there exists a BPIS (A, B) in $G_{W,n}$ such that:*

$$|A||B| \geq 2^{rn} - 2^{rn/2}((Q+1)^n - Q^n)$$

Proof. Note that $\mathcal{R}^{(0)}_{W,n} = r$ implies that there exists $(A', B') \subseteq [Q]^n \times [Q]^n$ such that for any $x^{(n)}, x'^{(n)} \in A'$ and $y^{(n)}, y'^{(n)} \in B'$:

(a) $x^{(n)} \neq x'^{(n)} \Rightarrow W_1^{(n)}(x^{(n)}, y^{(n)}) \neq W_1^{(n)}(x'^{(n)}, y'^{(n)})$

(b) $y^{(n)} \neq y'^{(n)} \Rightarrow W_2^{(n)}(x^{(n)}, y^{(n)}) \neq W_2^{(n)}(x'^{(n)}, y'^{(n)})$

and, in addition:

(c) $R_1 + R_2 = r$ where $R_1 = \frac{\log|A'|}{n}$, $R_2 = \frac{\log|B'|}{n}$.

Otherwise, if condition (a) or (b) does not hold, it is not possible to devise a proper decoding function D_1 or D_2 as it receives the same channel output and is required to produce two different decoded values. Condition (c) holds by our definitions, i.e., there must exist an achievable rate (R_1, R_2) such that $R_1 + R_2 = r$.

Now, assume without loss of generality, that $|A'| \geq |B'|$. Consider the pair (A, B) where:

$$\begin{aligned} S &= \{x^{(n)} \mid x^{(n)} \in A' \text{ s.t. } \exists y^{(n)} \in B' \text{ for which } \exists 1 \leq i \leq n : W(x_i, y_i) = (\phi, \phi)\} \\ A &= A' - S \\ B &= B' \end{aligned}$$

Namely A is obtained from A' by removing all $x^{(n)} \in A'$ that cause at least one erasure in the output of $W^{(n)}(x^{(n)}, y^{(n)})$ for some $y^{(n)} \in B'$. Clearly (A, B) induces a BPIS in $G_{W,n}$ of size $|A||B|$, and

$$|A||B| = (|A'| - |S|)|B'| \tag{5}$$

Now let us compute an upper bound on $|S|$. For this purpose, given $I \subseteq [n]$ let us define the subset $S_I \subseteq A'$, where S_I contains all $x^{(n)} \in A'$ for which there exists $y^{(n)} \in B'$ such that:

(i) for any $i \in I$, $W(x_i, y_i) = (\phi, \phi)$.

(ii) for any $j \in [n] - I$, $W(x_j, y_j) \neq (\phi, \phi)$.

Namely S_I is the set of all elements from A' which cause erasures in W exactly at indices of I. Note that if $x^{(n)}, x'^{(n)} \in S_I$ and $x^{(n)} \neq x'^{(n)}$ then $x_{[n]-I} \neq x'_{[n]-I}$ (the vectors obtained by projection of $x^{(n)}$ and $x'^{(n)}$ onto $[n] - I$), as otherwise condition (a) would not hold. This means that:

$$|S_I| \leq Q^{n-|I|}$$

And finally:

$$\begin{aligned} |S| &\leq \sum_{\substack{I\subseteq[n]\\ I\neq\emptyset}} |S_I| \\ &\leq \sum_{|I|=1}^{n} \left(\binom{n}{|I|} Q^{n-|I|} \right) \\ &= (Q+1)^n - Q^n \end{aligned}$$

By plugging this upper bound in equation 5, we get:

$$\begin{aligned} |A||B| &= (|A'| - |S|)|B'| \\ &\geq (2^{R_1 n} - ((Q+1)^n - Q^n))2^{R_2 n} \\ &= 2^{(R_1+R_2)n} - 2^{R_2 n}\left((Q+1)^n - Q^n\right) \end{aligned}$$

By our assumption, $|A'| \geq |B'|$, which means that $R_1 \geq R_2$, and since $R_1 + R_2 = r$, it follows that $R_2 \leq \frac{r}{2}$. Thus:

$$\begin{aligned} |A||B| &\geq 2^{(R_1+R_2)n} - 2^{R_2 n}\left((Q+1)^n - Q^n\right) \\ &\geq 2^{rn} - 2^{rn/2}\left((Q+1)^n - Q^n\right) \end{aligned}$$

□

The following lemma is a refinement of Lemma 6.1 for certain values of r and Q.

Lemma 6.2. *Let $Q = 2^q$ and let $W \in \mathcal{W}_{Q,\varepsilon}$. If $\mathcal{R}^{(0)}_{W,n} \geq r$ where $r \geq 2q\left(1 - \frac{1}{n}\right) + 4\frac{\log n}{n}$ and $q \geq \log n$, then there exists a BPIS (A, B) in $G_{W,n}$ such that:*

$$|A||B| \geq 2^{rn} - 2^{rn/2 + 2\log n + q(n-1)}$$

Proof. Let $\mathcal{R}^{(0)}_{W,n} = r_0$. Then, by Lemma 6.1 there exists a BPIS (A, B) in $G_{W,n}$ such that:

$$|A||B| \geq 2^{r_0 n} - 2^{r_0 n/2}\left((Q+1)^n - Q^n\right)$$

Next, notice that

$$\begin{aligned} (Q+1)^n - Q^n &= \sum_{k=1}^{n} \left(\binom{n}{k} Q^{n-k} \right) \\ &\leq \sum_{k=1}^{n} \left(n^k Q^{n-k} \right) \\ &= \sum_{k=1}^{n} \left(n^{k-1+1} \cdot Q^{n-1-(k-1)} \right) \\ &= n \cdot Q^{n-1} \cdot \sum_{k=1}^{n} \left(\left(\frac{n}{Q} \right)^{k-1} \right) \end{aligned}$$

Since $q \geq \log n$ implies $Q \geq n$, we conclude:

$$\begin{aligned}(Q+1)^n - Q^n &\leq n^2 \cdot Q^{n-1} \\ &= 2^{2\log n + q(n-1)}\end{aligned}$$

Thus there exists a BPIS (A, B) in $G_{W,n}$ such that:

$$|A||B| \geq 2^{r_0 n} - 2^{r_0 n/2 + 2\log n + q(n-1)} \tag{6}$$

Now, by our assumption $r \geq 2q\left(1 - \frac{1}{n}\right) + 4\frac{\log n}{n}$, which means that:

$$\begin{aligned}rn &\geq 2qn\left(1 - \frac{1}{n}\right) + 4\log n \\ \frac{rn}{2} &\geq qn\left(\frac{n-1}{n}\right) + 2\log n \\ \frac{rn}{2} &\geq q(n-1) + 2\log n\end{aligned}$$

Since $r_0 \geq r$, clearly $2\log n + q(n-1) \leq r_0 n/2$ and thus

$$\begin{aligned}|A||B| &\geq 2^{r_0 n} - 2^{r_0 n/2 + 2\log n + q(n-1)} \\ &= 2^{r_0 n/2}\left(2^{r_0 n/2} - 2^{2\log n + q(n-1)}\right) \\ &\geq 2^{rn/2}\left(2^{rn/2} - 2^{2\log n + q(n-1)}\right) \\ &= 2^{rn} - 2^{rn/2 + 2\log n + q(n-1)}\end{aligned}$$

□

We are now ready to prove Theorem 6.1:

Proof. (Theorem 6.1) First, suppose that $\mathcal{R}^{(0)}_{W,n} > r = 2q\left(1 - \frac{1}{n}\right)(1+\gamma)$. The conditions of Lemma 6.2 are satisfied iff:

$$2q\left(1 - \frac{1}{n}\right)(1+\gamma) \geq 2q\left(1 - \frac{1}{n}\right) + 4\frac{\log n}{n}$$

Namely, this relationship holds iff:

$$\begin{aligned}2q\gamma\left(1 - \frac{1}{n}\right) &\geq 4\frac{\log n}{n} \\ q\gamma\,(n-1) &\geq 2\log n \\ q\gamma &\geq 2\frac{\log n}{n-1}\end{aligned}$$

And indeed, since $\frac{\log n}{n-1} \leq 1$ for $n \geq 2$ it suffices to show that:

$$q\gamma \geq 2$$

and per our selection of q the relationship holds. Thus, by Lemma 6.2, there exists a BPIS (A, B) in $G_{W,n}$ such that:

$$\begin{aligned}|A||B| &\geq 2^{2q\left(1-\frac{1}{n}\right)(1+\gamma)n} - 2^{2q\left(1-\frac{1}{n}\right)(1+\gamma)n/2+2\log n+q(n-1)} \\ &= 2^{2qn\left(1-\frac{1}{n}\right)(1+\gamma)} - 2^{2qn\left(1-\frac{1}{n}\right)\left(\frac{1+\gamma}{2}+\frac{\log n}{q(n-1)}+\frac{1}{2}\right)} \\ &= 2^{2qn\left(1-\frac{1}{n}\right)(1+\gamma)} - 2^{2qn\left(1-\frac{1}{n}\right)\left(1+\frac{\gamma}{2}+\frac{\log n}{q(n-1)}\right)}\end{aligned}$$

Again, per our selection of q, $q \geq \frac{4}{\gamma}$ and since $n \geq 2$ it holds that:

$$\frac{\log n}{q(n-1)} \leq \frac{1}{q} \leq \frac{\gamma}{4}$$

Thus:

$$\begin{aligned}|A||B| &\geq 2^{2qn\left(1-\frac{1}{n}\right)(1+\gamma)} - 2^{2qn\left(1-\frac{1}{n}\right)\left(1+\frac{\gamma}{2}+\frac{\log n}{q(n-1)}\right)} \\ &\geq 2^{2qn\left(1-\frac{1}{n}\right)(1+\gamma)} - 2^{2qn\left(1-\frac{1}{n}\right)\left(1+\frac{\gamma}{2}+\frac{\gamma}{4}\right)} \\ &= 2^{2qn\left(1-\frac{1}{n}\right)(1+\gamma)} - 2^{2qn\left(1-\frac{1}{n}\right)\left(1+\frac{3\gamma}{4}\right)} \\ &= 2^{2qn\left(1-\frac{1}{n}\right)\left(1+\frac{3\gamma}{4}\right)}\left(2^{2qn\left(1-\frac{1}{n}\right)\frac{\gamma}{4}} - 1\right) \\ &\geq 2^{2qn\left(1-\frac{1}{n}\right)\left(1+\frac{3\gamma}{4}\right)}\end{aligned}$$

where the last inequality is due to our selection of $q \geq \frac{4}{\gamma}$ and $n \geq 2$. Now notice that this implies (for $n \geq 2$):

$$\begin{aligned}\frac{\log|A||B|}{n} &\geq 2q\left(1-\frac{1}{n}\right)\left(1+\frac{3\gamma}{4}\right) \\ &\geq q\left(1+\frac{3\gamma}{4}\right) \\ &= q+\frac{3q\gamma}{4}\end{aligned}$$

By our choice of $q \geq \frac{4}{\gamma}\log\frac{4\ln 2}{\varepsilon}$, we get:

$$\frac{\log|A||B|}{n} \geq q + 3\log\left(\frac{4\ln 2}{\varepsilon}\right) > q + \log\left(\frac{4\ln 2}{\varepsilon}\right)$$

However, recall that by Theorem 5.3 the probability that this relationship holds is at most $\left(\frac{1}{2}\right)^Q$ which for $Q \geq n \geq 2$ is at most $\frac{1}{4}$. Thus, all in all the probability that $\mathcal{R}_{W,n}^{(0)} > r = 2q\left(1-\frac{1}{n}\right)(1+\gamma)$ is at most $1/4$. This concludes the proof of our assertion. □

We thus conclude (based on Proposition 4.2) that:

Corollary 6.1. *For every integer $n \geq 2$, $\varepsilon \in [0,1]$ and $\gamma > 0$, let $Q = 2^q$ with $q \geq \max\{\log n, \frac{4}{\gamma}\log\frac{4\ln 2}{\varepsilon}\}$. Then there exist channels $W \in \mathcal{W}_{Q,\varepsilon}$ such that $\mathcal{R}_W^{(2\varepsilon)} \geq 2q$ and $\mathcal{R}_{W,n}^{(0)} \leq 2q\left(1-\frac{1}{n}\right)(1+\gamma)$.*

7 Constructing coding schemes

In this section we study two coding schemes that partially support the conjecture that $\mathcal{R}_W^{(0)} \simeq q$. In the first coding scheme we suggest and analyze, we tie the existence of a certain natural combinatorial structure to zero error communication. Namely, in Section 7.1 we define a combinatorial criteria (we refer to as the γ-uniform criteria) on subsets of $[Q]^n \times [Q]^n$, and show that subsets satisfying this criteria yield good zero error encoding schemes for the typical deterministic interference channels $W \in \mathcal{W}_{Q,\varepsilon}$. We then study upper bounds and lower bounds on the sizes of γ-uniform sets in Section 7.2. On one hand, the lower bounds we obtain, imply zero error coding schemes for which $\mathcal{R}_W^{(0)} \simeq q$ (and in such support our conjecture). On the other hand, appropriate upper bounds have the potential to prove that the suggested coding scheme (via γ-uniform sets) cannot exceed rate q. Unfortunately, we were not able to present tight upper bounds, and those we present do not rule out the possibility that the coding scheme suggested allows $\mathcal{R}_W^{(0)}$ to exceed q. Thus, although we believe that $\mathcal{R}_W^{(0)} \simeq q$, our results in this section only support this conjecture in a fairly weak manner.

The second scheme we study is a natural concatenated scheme mentioned in Section 3 and suggested to us by [24]. Specifically, in Section B of the Appendix, we define and analyze the suggested scheme and show that our analysis does not allow $\mathcal{R}_W^{(0)} > q$ for large values of Q. The ideas and proof of the second scheme are based on those in [24] and appear here only for completeness.

7.1 γ-uniform set systems and their connection to $\mathcal{R}_W^{(0)}$

Definition 7.1. *Given $x^{(n)} \in [Q]^n$ and $i \in [n]$, denote by x_i the i-th coordinate of $x^{(n)}$. A pair $(x^{(n)}, y^{(n)}) \in [Q]^n \times [Q]^n$ is called γ-uniform if for each pair $(\alpha, \beta) \in [Q]^2$ it holds that*

$$(1-\gamma)\frac{n}{Q^2} \leq |\{i \in [n] \mid (x_i, y_i) = (\alpha, \beta)\}| \leq (1+\gamma)\frac{n}{Q^2}.$$

In other words, the number of appearances of any pair $(\alpha, \beta) \in [Q]^2$ in $(x^{(n)}, y^{(n)})$ is bounded by $(1 \pm \gamma)\frac{n}{Q^2}$; i.e., the type *of $(x^{(n)}, y^{(n)})$ is γ-far from being uniform (under the $\|\cdot\|_\infty$ norm). Similarly, the subsets $A \subseteq [Q]^n, B \subseteq [Q]^n$ are called γ-uniform if for any $x^{(n)} \in A, y^{(n)} \in B$, $(x^{(n)}, y^{(n)})$ are γ-uniform.*

The following theorem ties the existence of γ-uniform set systems to good zero error codes for typical channels W in $\mathcal{W}_{Q,\varepsilon}$. Roughly speaking, given a γ-uniform pair A and B one can construct a zero error code for W by taking large subsets A' of A and B' of B with *large* minimum distance. Here the term *large* depends on ε and γ.

Theorem 7.1. *Let $Q = 2^q$. Let $A \subseteq [Q]^n, B \subseteq [Q]^n$ be γ-uniform with $|A| \geq Q^{n(1-\delta_1)}$ and $|B| \geq Q^{n(1-\delta_2)}$. Let $\delta > 0$ be arbitrarily small. Consider a channel W chosen from the distribution $\mathcal{W}_{Q,\varepsilon}$. With probability at least $1/2$ it holds that:*

$$\mathcal{R}_{W,n}^{(0)} \geq 2q\left(1 - \left(\frac{\delta_1+\delta_2}{2} + 2(1+\gamma)\varepsilon + \delta\right)\right) - 2. \tag{7}$$

To prove Theorem 7.1 we will introduce an additional combinatorial criteria on sets. We refer to the additional criteria as the (d, ε)-diversity criteria.

Definition 7.2. *A pair $(x^{(n)}, y^{(n)}) \in [Q]^n \times [Q]^n$ is called (d, ε)-diverse if for each index set $I \subseteq [n]$ of size dn it holds that $|\{(x_j, y_j) \mid j \in I\}| > \varepsilon Q^2$. Similarly, the subsets $A \subseteq [Q]^n, B \subseteq [Q]^n$ are called (d, ε)-diverse if for any $x^{(n)} \in A, y^{(n)} \in B$, $(x^{(n)}, y^{(n)})$ are (d, ε)-diverse.*

In words, a pair $(x^{(n)}, y^{(n)})$ is (d, ε)-diverse if every projection to a *sufficiently large* subset of coordinates contains a *large* number of distinct pairs (x_j, y_j). We first connect (d, ε)-diverse set systems to good zero error codes for channels in $\mathcal{W}_{Q,\varepsilon}$ (via Theorem 7.2 below). We then turn to prove Theorem 7.1.

Theorem 7.2. *Let $Q = 2^q$. Let $A \subseteq [Q]^n, B \subseteq [Q]^n$ be $(d, 2\varepsilon)$-diverse with $|A| \geq Q^{n(1-\delta_1)}$ and $|B| \geq Q^{n(1-\delta_2)}$. Consider a channel W chosen from the distribution $\mathcal{W}_{Q,\varepsilon}$. With probability at least $1/2$ it holds that*

$$\mathcal{R}^{(0)}_{W,n} \geq 2q \left(1 - \left(\frac{\delta_1 + \delta_2}{2} + d\right)\right) - 2.$$

To prove Theorem 7.2, we first require the following lemma:

Lemma 7.1. *Let $W \in \mathcal{W}_{Q,\varepsilon}$. Let $(A, B) \subseteq [Q]^n \times [Q]^n$ be such that for any $x^{(n)}, x'^{(n)} \in A$ and $y^{(n)}, y'^{(n)} \in B$:*

(a) $x^{(n)} \neq x'^{(n)} \Rightarrow W_1^{(n)}(x^{(n)}, y^{(n)}) \neq W_1^{(n)}(x'^{(n)}, y'^{(n)})$

(b) $y^{(n)} \neq y'^{(n)} \Rightarrow W_2^{(n)}(x^{(n)}, y^{(n)}) \neq W_2^{(n)}(x'^{(n)}, y'^{(n)})$

Then it holds that:

$$\mathcal{R}^{(0)}_{W,n} \geq \frac{1}{n} \log(|A||B|)$$

Proof. Note that conditions (a) and (b) in the proposition imply that any pair $(x^{(n)}, y^{(n)}) \in (A, B)$ passed as input to $W^{(n)}$ has a unique output $(\hat{x}^{(n)}, \hat{y}^{(n)})$. Thus, if we define the decoding function $D_1(\hat{x}^{(n)}) = x^{(n)}$ for any $x^{(n)} \in A$ and $D_2(\hat{y}^{(n)}) = y^{(n)}$ for any $y^{(n)} \in B$ then communication is successful with probability 1 for any $(x^{(n)}, y^{(n)}) \in (A, B)$. Let $M_1 = 2^{R_1 n} = |A|$, $M_2 = 2^{R_2 n} = |B|$. Then $R_1 = \frac{\log |A|}{n}$, $R_2 = \frac{\log |B|}{n}$ and

$$R_1 + R_2 = \frac{\log |A| + \log |B|}{n} = \frac{1}{n} \log(|A||B|)$$

and the proposition follows. □

Next, we require the following lemma that follows from a standard packing argument.

Lemma 7.2. *Let $A \subseteq [Q]^n$. Then for any $d \in [0, 1]$ there exists $A' \subseteq A$ such that $|A'| > \frac{|A|}{2^n Q^{dn}}$ and for any $x^{(n)}, x'^{(n)} \in A'$: $h(x^{(n)}, x'^{(n)}) > dn$, where $h : [Q]^n \times [Q]^n \to \{0, 1, \ldots, n\}$ is the Hamming distance function.*

Proof. Consider a graph $G = (V, E)$ where the vertices are elements of A, and there is an edge between two vertices $x^{(n)}, x'^{(n)}$ if and only if $h(x^{(n)}, x'^{(n)}) \leq dn$. The maximal degree of a vertex in this graph is $\binom{n}{dn}Q^{dn}$. Thus, the size of the independent set in G is at least $\frac{|A|}{\binom{n}{dn}Q^{dn}} > \frac{|A|}{2^n Q^{dn}}$, and the vertices of this independent set satisfy the conditions on A' in the lemma. □

Now we can prove Theorem 7.2:

Proof. (Theorem 7.2) Let W be chosen from the distribution $\mathcal{W}_{Q,\varepsilon}$ and let $A' \subseteq A, B' \subseteq B$ be the subsets whose existence is guaranteed by Lemma 7.2. Suppose $W, (A', B')$ satisfy the conditions in Lemma 7.1. Then Lemma 7.1 implies:

$$\mathcal{R}^{(0)}_{W,n} \geq \frac{1}{n}\log(|A'||B'|)$$

By Lemma 7.2:

$$|A'| > \frac{|A|}{2^n Q^{dn}}$$
$$|B'| > \frac{|B|}{2^n Q^{dn}}$$

And thus:

$$\begin{aligned}
\mathcal{R}^{(0)}_{W,n} &\geq \frac{1}{n}\log\left(|A'||B'|\right) \\
&> \frac{1}{n}\log\left(\frac{|A|}{2^n Q^{dn}} \cdot \frac{|B|}{2^n Q^{dn}}\right) \\
&\geq \frac{1}{n}\log\left(\frac{Q^{n(1-\delta_1)}}{2^n Q^{dn}} \cdot \frac{Q^{n(1-\delta_2)}}{2^n Q^{dn}}\right) \\
&= \frac{1}{n}\left((qn(1-\delta_1) - n - qdn) + (qn(1-\delta_2) - n - qdn)\right) \\
&= (q(1-\delta_1) - 1 - qd) + (q(1-\delta_2) - 1 - qd) \\
&= q\left(2 - \delta_1 - \delta_2 - 2d\right) - 2 \\
&= 2q\left(1 - \left(\frac{\delta_1 + \delta_2}{2} + d\right)\right) - 2
\end{aligned}$$

as Theorem 7.2 states.

Thus, it is left is to prove that $W, (A', B')$ satisfy the conditions in Lemma 7.1 with probability at least $1/2$ (over $\mathcal{W}_{Q,\varepsilon}$). Recall that by Markov inequality (see proof of Proposition 4.2) if X is a random variable that represents the number of "erasure" occurrences in W then:

$$\Pr\left[X < 2\varepsilon Q^2\right] \geq \frac{1}{2}.$$

In other words, with probability at least $1/2$ there are less than $2\varepsilon Q^2$ distinct values $(x, y) \in [Q] \times [Q]$ such that $W(x, y) = (\phi, \phi)$. Assume to the contrary that W is

indeed such a channel and that $W, (A', B')$ do not satisfy the conditions in Lemma 7.1. Without loss of generality, assume that condition (a) is not satisfied. Namely, that there exist $x^{(n)}, x'^{(n)} \in A'$, $y^{(n)}, y'^{(n)} \in B'$ such that $x^{(n)} \neq x'^{(n)}$ and $W_1^{(n)}(x^{(n)}, y^{(n)}) = W_1^{(n)}(x'^{(n)}, y'^{(n)})$. By our definition of A' via Lemma 7.2, $h(x^{(n)}, x'^{(n)}) > dn$. Namely $x^{(n)}, x'^{(n)}$ differ in at least dn coordinates. Formally, this means that there exists an index set $I \subset [n]$ of size dn such that for any $i \in I$, $x_i \neq x'_i$. By our assumption, $W_1^{(n)}(x^{(n)}, y^{(n)}) = W_1^{(n)}(x'^{(n)}, y'^{(n)})$ and in particular for $i \in I$ it must be the case that

$$W_1(x_i, y_i) = W_1(x'_i, y'_i) = \phi \tag{8}$$

as otherwise the channel returns the identity value, which would imply $x_i = W_1(x_i, y_i) = W_1(x'_i, y'_i) = x'_i$. Now, as A, B are $(d, 2\varepsilon)$-diverse, so are A', B'. Thus by definition,

$$|\{(x_i, y_i) \mid i \in I\}| > 2\varepsilon Q^2 \tag{9}$$

By combining equations 8 and 9 we get that there are more than $2\varepsilon Q^2$ distinct values (x, y) such that $W(x, y) = (\phi, \phi)$ in contradiction to our assumption on W above.

All in all, we conclude that with probability at least $1/2$ there are less than $2\varepsilon Q^2$ distinct values $(x, y) \in [Q] \times [Q]$ such that $W(x, y) = (\phi, \phi)$, which implies the conditions in Lemma 7.1, which in turn implies our assertion. □

We now tie γ-uniform set systems to (d, ε)-diverse systems.

Lemma 7.3. *If $d > (1 + \gamma)\varepsilon$ then: (A, B) is γ-uniform $\Rightarrow$ (A, B) is (d, ε)-diverse.*

Proof. Let $(x^{(n)}, y^{(n)}) \in [Q]^n \times [Q]^n$ be γ-uniform and $I \subseteq [n]$ some index set of size dn. A pair $(\alpha, \beta) \in [Q]^2$ can appear at most $(1 + \gamma)\frac{n}{Q^2}$ times in $(x^{(n)}, y^{(n)})$. Particularly, $|\{i \mid (x_i, y_i) = (\alpha, \beta), i \in I\}| \leq (1 + \gamma)\frac{n}{Q^2}$. This means that $|\{(x_i, y_i) \mid i \in I\}| \geq \frac{dn}{(1+\gamma)\frac{n}{Q^2}} = \frac{d}{1+\gamma}Q^2$. But if $d > (1 + \gamma)\varepsilon$ then $\frac{d}{1+\gamma}Q^2 > \varepsilon Q^2$, and thus $(x^{(n)}, y^{(n)})$ is (d, ε)-diverse. □

Finally, we conclude with the proof of Theorem 7.1:

Proof. (Theorem 7.1) The proof of Theorem 7.1 follows directly by combining Lemma 7.3 with Theorem 7.2. □

7.2 Upper and lower bounds on γ-uniform set systems

The previous section presented a scheme to construct codes for channels W chosen at random from the distribution $\mathcal{W}_{Q,\varepsilon}$ based on γ-uniform set systems. We now attempt to better understand the parameters for which such set systems exist (and as a consequence what the implications are on $\mathcal{R}_W^{(0)}$). Our analysis is divided into several case studies that include binary alphabets $Q = 2$, general alphabets, the case in which $\gamma = 0$, and the case in which $\gamma > 0$. In our statements that follow, upper bounds will hold for every pair of sets A and B, and lower bounds will hold for certain pairs we construct.

7.2.1 Bounds on zero-uniform set systems, ($\gamma = 0$)

In this section we will show the following bounds on sets (A, B) which are zero-uniform.

When $Q = 2$:

$$\left(\frac{2}{n}\right) 2^n \leq |A||B| \leq 2^n$$

For other values of Q (explicit details appear in the claims below):

$$\left(\frac{Q}{e\sqrt{n}}\right)^{Q(Q-1)} Q^n \leq |A||B| \leq Q^n$$

First, we study the case $Q = 2$. Below we list some basic properties of zero-uniform set systems in this case.

Proposition 7.1. *Let $A \subseteq \{0,1\}^n, B \subseteq \{0,1\}^n$ be a zero-uniform pair. Then:*

(i) n is divisible by 4.

(ii) If $a \in A$ then a contains $n/2$ zeroes and $n/2$ ones. Same for $b \in B$.

If $|A||B|$ is maximal then it also holds that:

(iii) $|A|$ and $|B|$ are even.

(iv) $|A||B| \geq 2 \cdot \binom{n/2}{n/4}^2$.

Proof. Note that a pair $(a, b) \in \{0,1\}^n \times \{0,1\}^n$ is zero-uniform if and only if each of the pairs $(0,0), (0,1), (1,0), (1,1)$ appears exactly $n/4$ times in it. Propositions (i) and (ii) easily follow. Now consider a bipartite balanced graph $G_n = (V_L \cup V_R, E)$ where V_L and V_R are vertices that represent elements in the set

$$\{u \mid u \in \{0,1\}^n, u \text{ contains } n/2 \text{ zeroes and } n/2 \text{ ones}\}$$

and $(u, v) \in E$ if and only if (u, v) is zero-uniform. Note that $|V_L| = |V_R| = \binom{n}{n/2}$, as it is the number of ways to distribute $n/2$ ones in n places. Also every vertex in G has the same degree, equal to $\binom{n/2}{n/4}^2$, as it is the number of ways to distribute $n/4$ ones to match $n/2$ ones $\times$ the number of ways to distribute $n/4$ ones to match $n/2$ zeroes. Finally note that a vertex v and the vertex representing its bitwise inverse have the same neighbors. Thus given that $|A||B|$ is maximal, if $a \in A$ then its bitwise inverse is also in A, and the same for $b \in B$. Proposition (iii) now follows. For proposition (iv), take A to be any u in V_L and its bitwise inverse, and take B to be their neighbors in V_R. □

Note that the lower bound in Proposition 7.1 (iv) is tight for $n = 4$, but no longer tight for $n = 8$. To see that it is tight consider the graph G_4 (described in the proof of Proposition 7.1 above) and its complement in Figure 4 and Figure 5 respectively. As we can see from Figure 5, the maximal BPIS (which corresponds to a zero-uniform

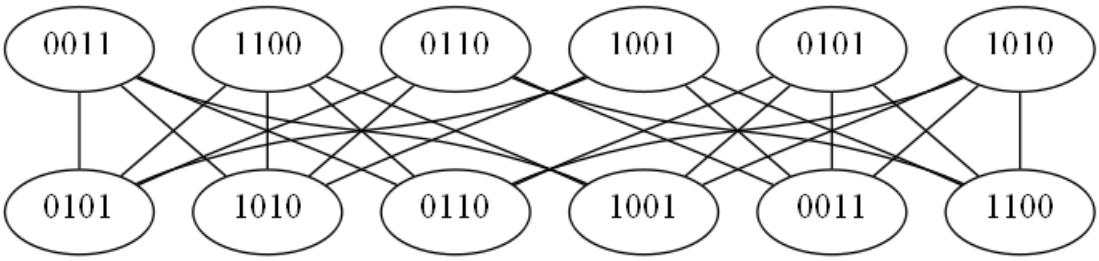


Figure 4: G_4

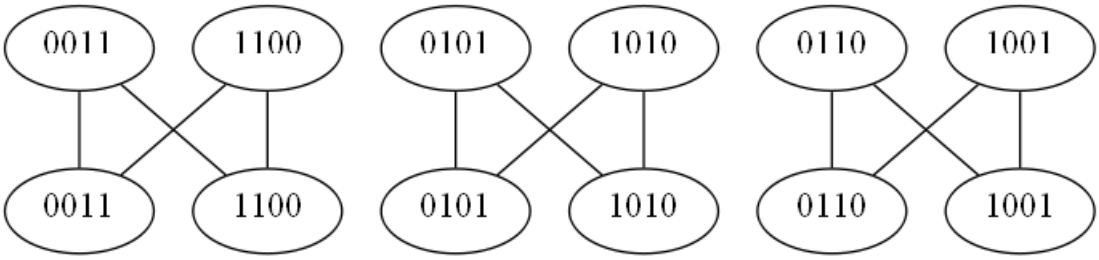


Figure 5: Complement graph of G_4

set system) contains at most 2 vertices from one side and 4 vertices from the other, matching the bound of Proposition 7.1 (iv). For $n = 8$ however, consider the BPIS:

$$\begin{aligned} A &= \{00001111, 11110000, 00110011, 11001100, 00111100, 11000011\} \\ B &= \{01010101, 10101010, 01010110, 10101001, 01011001, 10100110, 01011010, 10100101, \\ &\quad 01100101, 10011010, 01100110, 10011001, 01101001, 10010110, 01101010, 10010101\} \end{aligned}$$

Its size is $6 \times 16 = 96$, which is higher than the given bound $2 \cdot \binom{4}{2}^2 = 72$.

Proposition 7.2. *Let $n \geq 4$ be divisible by 4. There exists a zero-uniform pair (A, B), $A \subseteq \{0,1\}^n, B \subseteq \{0,1\}^n$ such that $|A||B| \geq \binom{2}{n} 2^n$.*

Proof. By Proposition 7.1 (iv), there exists a zero-uniform pair (A, B), $A \subseteq \{0,1\}^n, B \subseteq \{0,1\}^n$ such that:

$$|A||B| \geq 2 \cdot \binom{n/2}{n/4}^2.$$

By Lemma A.3, we get:

$$|A||B| \geq 2 \cdot \binom{n/2}{n/4}^2 \geq 2 \cdot \frac{2^n}{n} = \binom{2}{n} 2^n.$$

□

Proposition 7.3. *Let $Q \geq 2$ such that n is divisible by Q^2. There exists a zero-uniform pair $A \subseteq [Q]^n, B \subseteq [Q]^n$ such that:*

$$|A||B| \geq \left(\frac{Q}{e\sqrt{n}}\right)^{Q(Q-1)} Q^n.$$

Proof. Define $a^{(n)} \in [Q]^n$ to be:

$$a_i = \left\lceil \frac{Qi}{n} \right\rceil \text{ for all } 1 \leq i \leq n$$

Namely $a^{(n)}$ has the form $1^{\frac{n}{Q}} 2^{\frac{n}{Q}} \dots Q^{\frac{n}{Q}}$. Consider the sets $A \subseteq [Q]^n, B \subseteq [Q]^n$ where $A = \{a^{(n)}\}$ and B is the maximal set such that (A, B) is zero-uniform. Such B can be obtained by distributing $\frac{n}{Q^2}$ j-s ($1 \leq j \leq Q$) in each block of size $\frac{n}{Q}$. In each block we have $\binom{\frac{n}{Q}}{\frac{n}{Q^2}}$ possibilities for distributing 1-s, $\binom{\frac{n}{Q}-\frac{n}{Q^2}}{\frac{n}{Q^2}}$ possibilities for distributing 2-s, etc. Namely there are $\prod_{i=1}^{Q} \binom{i\cdot\frac{n}{Q^2}}{\frac{n}{Q^2}}$ possibilities for each block. It holds that

$$|A||B| = |B| = \left(\prod_{i=1}^{Q} \binom{i \cdot \frac{n}{Q^2}}{\frac{n}{Q^2}}\right)^Q$$

Notice that by Lemma A.4:

$$\prod_{i=1}^{Q} \binom{i \cdot \frac{n}{Q^2}}{\frac{n}{Q^2}} > \left(\frac{Q}{e\sqrt{n}}\right)^{(Q-1)} \frac{\sqrt{2\pi Q}}{e} \cdot Q^{\frac{n}{Q}}$$

And thus:

$$\left(\prod_{i=1}^{Q} \binom{i \cdot \frac{n}{Q^2}}{\frac{n}{Q^2}}\right)^Q > \left(\frac{Q}{e\sqrt{n}}\right)^{Q(Q-1)} \left(\frac{\sqrt{2\pi Q}}{e}\right)^Q \cdot Q^n$$

For any $Q \geq 2$, $\sqrt{2\pi Q} > e$. Therefore, we get:

$$|A||B| = \left(\prod_{i=1}^{Q} \binom{i \cdot \frac{n}{Q^2}}{\frac{n}{Q^2}}\right)^Q > \left(\frac{Q}{e\sqrt{n}}\right)^{Q(Q-1)} \cdot Q^n$$

as claimed. □

Remark 7.1. *We could improve the bound in Proposition 7.3 by a factor of $Q!$ by setting A to contain all elements "isomorphic" to $a^{(n)}$ as well. Those elements are obtained by rearranging the $\frac{n}{Q}$-sized blocks of $a^{(n)}$. We omit the improved bound for the sake of clarity.*

Theorem 7.3. *If $A \subseteq \{0,1\}^n, B \subseteq \{0,1\}^n$ is a zero-uniform pair then $|A||B| \leq 2^n$.*

Below we describe two proofs, marked A and B. Proof A is our own, and Proof B is a corollary of a theorem we discovered later.

Proof A. First, notice that instead of considering the alphabet $\{0,1\}$ we can consider the alphabet $\{-1,1\}$. This does not affect the definition of zero-uniformity, or the sizes of A and B, since there is a one-to-one mapping between elements over those alphabets, explicitly $x \to 1-2x$. This transition allows us to state the following: if (x,y) is zero-uniform then $x \cdot y = 0$, where the dot stands for the dot product. This follows from simple and standard calculations. Now, consider the subspaces $A_1 = span(A), B_1 = span(B)$ in $\mathbb{R}^n$. Those subspaces are orthogonal, and hence $dim(A_1) + dim(B_1) \leq n$. Now, let $A_2 = A_1 \cap \{-1,1\}^n$ and $B_2 = B_1 \cap \{-1,1\}^n$. Clearly $|A_2| \geq |A|$ and $|B_2| \geq |B|$. If $|A_2| \leq 2^{dim(A_1)}$ and $|B_2| \leq 2^{dim(B_1)}$ as we prove below then $|A||B| \leq |A_2||B_2| \leq 2^{dim(A_1)} \cdot 2^{dim(B_1)} = 2^{dim(A_1)+dim(B_1)} \leq 2^n$, as required.

All is left, is to prove that for a subspace α of dimension l in $\mathbb{R}^n$, the intersection of α with $\{-1,1\}^n$ is of size at most 2^l. Then, setting α to be A_1 or B_1 and $l = dim(A_1)$ or $l = dim(B_1)$ accordingly, will conclude the proof of the lemma. Assume $S = \alpha \cap \{-1,1\}^n$ is of size $\geq 2^l + 1$. In that case, we claim that α has at least $l+1$ linearly independent vectors of the form $u - v$ for some $u, v \in S$, in contradiction to the fact that $dim(\alpha) = l$. We prove this by induction on l.

Basis: $l = 1$. If $|S| \geq 3$, then there exist $u, v, w \in S$, such that $u_i = v_i \neq w_i$ at some coordinate i (this follows from the fact that there are only two possible values at each coordinate, and the vectors should differ in at least one coordinate). Thus, $u - v$ and $w - v$ are linearly independent vectors in α, and the claim holds.

Inductive step: Assume the claim holds for $l = t$, and consider $l = t+1$, i.e. $|S| \geq 2^{t+1} + 1$. Let:

$$\begin{aligned} S_i^{-1} &= \{u \mid u \in S, u_i = -1\} \\ S_i^{1} &= \{u \mid u \in S, u_i = 1\} \end{aligned}$$

As before, there is a coordinate i for which S_i^{-1} and S_i^1 are both not empty. Denote those sets by S_1 and S_2 such that $|S_1| \geq |S_2| > 0$. Note that $|S_1| \geq \lceil \frac{2^{t+1}+1}{2} \rceil = 2^t + 1$, since $S_1 \cup S_2 = S$ and $|S| \geq 2^{t+1} + 1$. Thus, by the induction hypothesis, there are at least $t+1$ linearly independent vectors of the form $u - v$ where $u, v \in S_1$. On the other hand, the vector $w - v$ for $v \in S_1, w \in S_2$ is linearly independent of all vectors $u - v$ where $u, v \in S_1$, thus there are at least $t+2$ linearly independent vectors in S.

□

Proof B. Consider the following theorem:

Theorem 7.4. *[8, Theorem 2.5.3] Let $A \subseteq \{0,1\}^n, B \subseteq \{0,1\}^n$ and let "$\cdot$" be the dot product operation. If for any $x \in A$, $y \in B$ it holds that $x \cdot y$ is even, then $|A||B| \leq 2^n$. If for any $x \in A$, $y \in B$ it holds that $x \cdot y$ is odd, then $|A||B| \leq 2^{n-1}$.*

Let $A \subseteq \{0,1\}^n, B \subseteq \{0,1\}^n$ be zero-uniform, and consider some zero-uniform pair $(x,y) \in A \times B$. In (x,y) there are exactly $\frac{n}{4}$ $(1,1)$ pairs, thus $x \cdot y = \frac{n}{4}$ which is

even when n is divisible by 8 and odd otherwise. In particular, by the theorem above: $|A||B| \leq 2^n$. □

It turns out that Theorem 7.3 can be generalized, as demonstrated by the following theorem.

Theorem 7.5. *Let Q be a prime. If $A \subseteq [Q]^n, B \subseteq [Q]^n$ is a zero-uniform pair then $|A||B| \leq Q^n$.*

Proof. For $Q = 2$ this is exactly Theorem 7.3 above. For $Q \geq 3$, we can use, without loss of generality, $GF(Q)$ instead of $[Q]$. So let $A \subseteq GF(Q)^n, B \subseteq GF(Q)^n$ be a zero-uniform pair. The structure of the proof is similar to Proof A for Theorem 7.3. First we claim that if $(x, y) \in A \times B$ is zero uniform, then $x \cdot y \equiv 0 \pmod{Q}$, where $\cdot$ is the dot product. This is because every pair $(i, j) \in GF(Q) \times GF(Q)$ must appear exactly $\frac{n}{Q^2}$ times in (x, y), and thus:

$$\begin{aligned} x \cdot y &= \frac{n}{Q^2} \cdot \sum_{i=0}^{Q-1} \sum_{j=0}^{Q-1} i \cdot j \\ &= \frac{n}{Q^2} \cdot \sum_{i=1}^{Q-1} \sum_{j=1}^{Q-1} i \cdot j \\ &= \frac{n}{Q^2} \cdot \sum_{i=1}^{Q-1} i \cdot \sum_{j=1}^{Q-1} j \\ &= \frac{n}{Q^2} \cdot \frac{Q(Q-1)}{2} \cdot \frac{Q(Q-1)}{2} \\ &= n \cdot \left(\frac{Q-1}{2}\right)^2 \end{aligned}$$

For $Q \geq 3$ prime, the term $\frac{Q-1}{2}$ is an integer, and since n is divisible by Q (by our definition of zero-uniform) it follows that $x \cdot y$ is divisible by Q, i.e., $x \cdot y \equiv 0 \pmod{Q}$.

Now, consider the subspaces $A_1 = span(A), B_1 = span(B)$ in $GF(Q)^n$. Those subspaces are orthogonal, and hence $dim(A_1) + dim(B_1) \leq n$. Clearly $|A_1| \geq |A|$ and $|B_1| \geq |B|$. If $|A_1| \leq Q^{dim(A_1)}$ and $|B_1| \leq Q^{dim(B_1)}$ then $|A||B| \leq |A_1||B_1| \leq Q^{dim(A_1)} \cdot Q^{dim(B_1)} = Q^{dim(A_1)+dim(B_1)} \leq Q^n$, as required. And indeed, given a base $v_1, v_2, ..., v_{dim(A_1)}$ for A_1, A_1 contains elements of the form $\sum_{i=1}^{dim(A_1)} \alpha_i v_i$. Since $\alpha_i \in \{0, 1, \ldots, Q-1\}$ this means that there are at most $Q^{dim(A_1)}$ different elements in A_1, as needed. The proof for B_1 is identical, and thus we conclude the proof of the theorem. □

7.2.2 Bounds on γ-uniform set systems ($\gamma > 0$)

The following lemmas present both upper and lower bounds on the size of γ-uniform set systems for $\gamma > 0$. After we present our bounds we will elaborate on their implication

on Theorem 7.1. As before we study the case of $Q = 2$ and more general values of Q. In this section we will presents the following bounds on sets (A, B) which are γ-uniform.

When $Q = 2$:

$$\frac{2}{n(n+1)} \cdot 2^{\left(1+H\left(\frac{\gamma}{4}\right)\right)n} \leq |A||B| \leq \min\left(2^{\left(1+\frac{\gamma}{2}+H\left(\frac{\gamma}{2}\right)\right)n}, 2^{2H\left(\frac{1+\gamma}{4}\right)n}\right)$$

where H is the binary entropy function. Note that the upper bound of $2^{\left(1+\frac{\gamma}{2}+H\left(\frac{\gamma}{2}\right)\right)n}$ is tighter for small values of γ (approximately $\gamma < 0.36$), while the upper bound of $2^{2H\left(\frac{1+\gamma}{4}\right)n}$ is tighter for larger values of γ, see Figure 6 for comparison of the factor in the exponent.

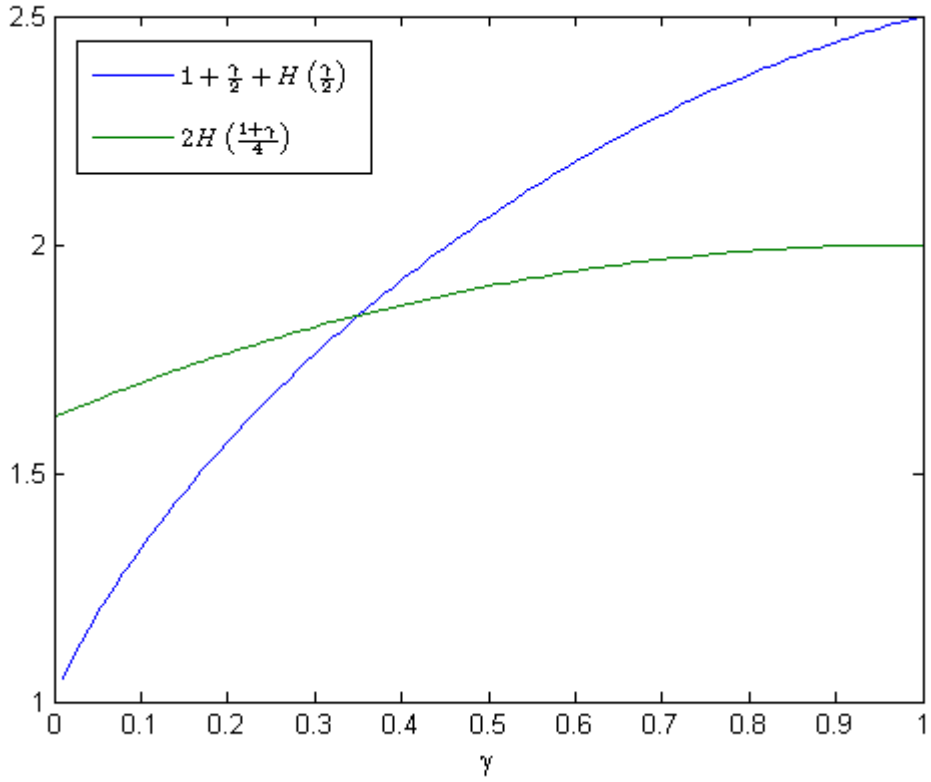


Figure 6: Factor of n in our upper bounds for $Q = 2$ case.

For other values of Q (explicit details appear in the claims below):

$$\left(\frac{1}{n+1}\right)\left(\frac{Q}{e\sqrt{n}}\right)^{Q(Q-1)} Q^{\left(1+\frac{\log(Q-1)}{\log Q}\cdot\frac{\gamma}{Q^2}+\frac{1}{\log Q}\cdot H\left(\frac{\gamma}{Q^2}\right)\right)n} \leq |A||B| \leq \min\left((2Q)^{\left(1+\frac{\gamma}{2}+H\left(\frac{\gamma}{2}\right)\right)n}, (2Q)^{2H\left(\frac{1+\gamma}{4}\right)n}\right)$$

We start with the case $Q = 2$.

Proposition 7.4. *Let $0 < \gamma < 2$. Let n be divisible by 4. There exists a γ-uniform pair $A \subseteq \{0,1\}^n, B \subseteq \{0,1\}^n$ such that:*

$$|A||B| \geq \frac{2}{n(n+1)} \cdot 2^{\left(1+H\left(\frac{\gamma}{4}\right)\right)n}.$$

where H is the binary entropy function.

Proof. Define $a^{(n)}, \bar{a}^{(n)} \in \{0,1\}^n$ to be:

$$a_i = \begin{cases} 0 & \text{if } i \leq \frac{n}{2} \\ 1 & \text{otherwise} \end{cases}$$

$$\bar{a}_i = \begin{cases} 0 & \text{if } i > \frac{n}{2} \\ 1 & \text{otherwise} \end{cases}$$

Namely $a^{(n)}$ has the form $0^{\frac{n}{2}}1^{\frac{n}{2}}$ and $\bar{a}^{(n)}$ is its bitwise inverse and has the form $1^{\frac{n}{2}}0^{\frac{n}{2}}$. Consider the sets $A' \subseteq \{0,1\}^n, B \subseteq \{0,1\}^n$ where $A' = \{a^{(n)}, \bar{a}^{(n)}\}$ and B is the maximal set such that (A', B) is zero-uniform. As we have seen in our proof of Proposition 7.1 (iv):

$$|A'||B| = 2 \cdot \binom{\frac{n}{2}}{\frac{n}{4}}^2$$

since such B can be obtained by selecting $y^{(n)}$-s that have exactly $\frac{n}{4}$ zeroes in the range $y_1 \dots y_{\frac{n}{2}}$ and exactly $\frac{n}{4}$ zeroes in the range $y_{\frac{n}{2}+1} \dots y_n$.

Let $A = \{x^{(n)} \mid \min(d(x^{(n)}, a^{(n)}), d(x^{(n)}, \bar{a}^{(n)})) \leq \frac{\gamma n}{4}\}$, where d is the Hamming distance. It holds that (A, B) is γ-uniform (as each $x^{(n)}$ in A is *close* to either $a^{(n)}$ or $\bar{a}^{(n)}$). Since $\gamma < 2$:

$$|A||B| = \left(2 \sum_{i \leq \frac{\gamma n}{4}} \binom{n}{i}\right) \cdot \binom{\frac{n}{2}}{\frac{n}{4}}^2$$

Using the lower bounds (see, e.g., Lemma A.3 for the first one and [23, Lemma 9.2] for the second inequality):

$$\binom{\frac{n}{2}}{\frac{n}{4}}^2 \geq \frac{2^n}{n}$$

$$\sum_{i \leq \frac{\gamma n}{4}} \binom{n}{i} > \binom{n}{\frac{\gamma n}{4}} \geq \frac{2^{H(\frac{\gamma}{4})n}}{n+1}$$

We obtain the bound of Proposition 7.4:

$$|A||B| > \frac{2}{n(n+1)} \cdot 2^{(1+H(\frac{\gamma}{4}))n}.$$

□

Proposition 7.5. *Let $Q \geq 3$ such that n is divisible by Q^2. For $\gamma \leq Q^2/2$ there exists a γ-uniform pair $A \subseteq [Q]^n, B \subseteq [Q]^n$ such that:*

$$|A||B| \geq \left(\frac{1}{n+1}\right)\left(\frac{Q}{e\sqrt{n}}\right)^{Q(Q-1)} Q^{n\left(1+\frac{\log(Q-1)}{\log Q}\cdot\frac{\gamma}{Q^2}+\frac{1}{\log Q}\cdot H\left(\frac{\gamma}{Q^2}\right)\right)}.$$

Proof. We use the same techniques used in Propositions 7.3 and 7.4. As in Proposition 7.3, define $a^{(n)} \in [Q]^n$ to have the form $1^{\frac{n}{Q}}2^{\frac{n}{Q}}\ldots Q^{\frac{n}{Q}}$. Similar to Proposition 7.4, define $A = \{x^{(n)} \mid d(x^{(n)}, a^{(n)}) \leq \frac{\gamma n}{Q^2}\}$, where d is the Hamming distance, and define B to be the maximal set such that (A, B) is γ-uniform.

Notice that:

$$|A| = \sum_{i \leq \frac{\gamma n}{Q^2}} \left(\binom{n}{i}(Q-1)^i\right) \geq \binom{n}{\frac{\gamma n}{Q^2}}(Q-1)^{\frac{\gamma n}{Q^2}} \geq \left(\frac{2^{H\left(\frac{\gamma}{Q^2}\right)n}}{n+1}\right)(Q-1)^{\frac{\gamma n}{Q^2}}$$

Also recall that:

$$|B| = \left(\prod_{i=1}^{Q}\binom{i\cdot\frac{n}{Q^2}}{\frac{n}{Q^2}}\right)^Q \geq \left(\frac{Q}{e\sqrt{n}}\right)^{Q(Q-1)} Q^n.$$

Thus:

$$\begin{aligned}|A||B| &\geq \left(\frac{2^{H\left(\frac{\gamma}{Q^2}\right)n}}{n+1}\right)(Q-1)^{\frac{\gamma n}{Q^2}}\left(\frac{Q}{e\sqrt{n}}\right)^{Q(Q-1)} Q^n \\ &= \left(\frac{1}{n+1}\right)\left(\frac{Q}{e\sqrt{n}}\right)^{Q(Q-1)} Q^{n\left(1+\frac{\log(Q-1)}{\log Q}\cdot\frac{\gamma}{Q^2}+\frac{1}{\log Q}\cdot H\left(\frac{\gamma}{Q^2}\right)\right)}\end{aligned}$$

as claimed.

□

We now address upper bounds on the size of $|A||B|$. We start with the binary case and present two upper bounds.

Theorem 7.6. *Let $A \subseteq \{0,1\}^n, B \subseteq \{0,1\}^n$ be γ-uniform ($\gamma \leq 1$). Then:*

$$|A||B| \leq 2^{2H\left(\frac{1+\gamma}{4}\right)n}.$$

In order to prove Theorem 7.6 we need an additional lemma.

Lemma 7.4. *For any $A \subseteq \{0,1\}^n, B \subseteq \{0,1\}^n$ it holds that:*

$$|A||B| \leq |A \vee B||A \wedge B|$$

where $\vee$ and $\wedge$ represent bitwise OR *and* AND *respectively and*

$$A \vee B = \{x^{(n)} \vee y^{(n)} \mid x^{(n)} \in A, y^{(n)} \in B\}$$
$$A \wedge B = \{x^{(n)} \wedge y^{(n)} \mid x^{(n)} \in A, y^{(n)} \in B\}$$

Proof. Consider the four functions theorem of Ahlswede and Daykin [2] below.

Theorem 7.7. *[2][4, Theorem 6.1.1] Suppose $n \geq 1$ and put $N = \{1, 2, \ldots, n\}$. Let $P(N)$ denote the set of all subsets of N, and let $\mathbb{R}^+$ denote the set of nonnegative real numbers. For a function $f : P(N) \to \mathbb{R}^+$ and for a family $\mathcal{A}$ of subsets of N denote $f(\mathcal{A}) = \sum_{X \in \mathcal{A}} f(X)$. For two families $\mathcal{A}$ and $\mathcal{B}$ of subsets of N define*

$$\mathcal{A} \cup \mathcal{B} = \{X \cup Y \mid X \in \mathcal{A}, Y \in \mathcal{B}\}$$
$$\mathcal{A} \cap \mathcal{B} = \{X \cap Y \mid X \in \mathcal{A}, Y \in \mathcal{B}\}$$

Let $\{f_i\}_{i=1}^4 : P(N) \to \mathbb{R}^+$ be four functions from the set of all subsets of N to the nonnegative reals. If, for every two subsets $X, Y \subseteq N$ the inequality

$$f_1(X) f_2(Y) \leq f_3(X \cup Y) f_4(X \cap Y)$$

holds, then, for every two families of subsets $\mathcal{A}, \mathcal{B} \subseteq P(N)$,

$$f_1(\mathcal{A}) f_2(\mathcal{B}) \leq f_3(\mathcal{A} \cup \mathcal{B}) f_4(\mathcal{A} \cap \mathcal{B}).$$

Note that $x^{(n)} \in A$ can be regarded as a characteristic vector of a set $X \in P(N)$ (and similarly for $y^{(n)} \in B$). Thus, our sets $A, B \subseteq \{0,1\}^n$ correspond to the families $\mathcal{A}, \mathcal{B} \subseteq P(N)$. Now, let us set for any $X \in P(N)$ and all i, $f_i(X) = 1$. Then, for any family $\mathcal{A} \subseteq P(N)$ and all i, $f_i(\mathcal{A}) = |\mathcal{A}|$. Clearly, the condition of the theorem holds, since we get 1 in both sides of the inequality. Thus, we get that for any $\mathcal{A}, \mathcal{B} \subseteq P(N)$, $|\mathcal{A}||\mathcal{B}| \leq |\mathcal{A} \cup \mathcal{B}||\mathcal{A} \cap \mathcal{B}|$, which is equivalent to $|A||B| \leq |A \vee B||A \wedge B|$ for any $A, B \subseteq \{0,1\}^n$ in our notation (since union and intersection of sets of $P(N)$ correspond to bitwise OR and bitwise AND of their characteristic vectors). □

Now we can prove Theorem 7.6.

Proof. (Theorem 7.6)

Note that since A, B are γ-uniform the number of $(1,1)$ pairs that appear in any $(x^{(n)}, y^{(n)}) \in A \times B$ is between $\lceil (1-\gamma)\frac{n}{4} \rceil$ and $\lfloor (1+\gamma)\frac{n}{4} \rfloor$. This means that $\lceil (1-\gamma)\frac{n}{4} \rceil \leq |x^{(n)} \wedge y^{(n)}| \leq \lfloor (1+\gamma)\frac{n}{4} \rfloor$. By the definitions in Lemma 7.4, $|A \wedge B|$ contains all distinct intersections $x^{(n)} \wedge y^{(n)}$ (when $x^{(n)}$ and $y^{(n)}$ are ranged in A and B respectively). Thus,

$$|A \wedge B| \leq \sum_{i=\lceil (1-\gamma)\frac{n}{4} \rceil}^{\lfloor (1+\gamma)\frac{n}{4} \rfloor} \binom{n}{i}.$$

In a complement manner, the number of non-$(0,0)$ pairs that appear in any $(x^{(n)}, y^{(n)}) \in A \times B$ is between $n - \lceil (1-\gamma)\frac{n}{4} \rceil$ and $n - \lfloor (1+\gamma)\frac{n}{4} \rfloor$. This means that $n - \lceil (1-\gamma)\frac{n}{4} \rceil \leq$

$|x^{(n)} \vee y^{(n)}| \leq n - \lfloor (1+\gamma)\frac{n}{4} \rfloor$. Similarly, by definitions in Lemma 7.4,

$$\begin{aligned} |A \vee B| &\leq \sum_{i=n-\lceil (1-\gamma)\frac{n}{4} \rceil}^{n-\lfloor (1+\gamma)\frac{n}{4} \rfloor} \binom{n}{i} \\ &= \sum_{i=n-\lceil (1-\gamma)\frac{n}{4} \rceil}^{n-\lfloor (1+\gamma)\frac{n}{4} \rfloor} \binom{n}{n-i} \\ &= \sum_{i=\lceil (1-\gamma)\frac{n}{4} \rceil}^{\lfloor (1+\gamma)\frac{n}{4} \rfloor} \binom{n}{i}. \end{aligned}$$

As $\gamma \leq 1$, it holds that $\lceil (1-\gamma)\frac{n}{4} \rceil \geq 0$ and $\lfloor (1+\gamma)\frac{n}{4} \rfloor \leq \lfloor \frac{1}{2}n \rfloor$. By Lemma A.2 we get:

$$\sum_{i=\lceil (1-\gamma)\frac{n}{4} \rceil}^{\lfloor (1+\gamma)\frac{n}{4} \rfloor} \binom{n}{i} \leq \sum_{i=0}^{\lfloor \frac{1+\gamma}{4} \cdot n \rfloor} \binom{n}{i} \leq 2^{H(\frac{1+\gamma}{4})n}.$$

Putting it all together, by Lemma 7.4 and our bounds above:

$$\begin{aligned} |A||B| &\leq |A \vee B||A \wedge B| \\ &\leq \left(\sum_{i=\lceil (1-\gamma)\frac{n}{4} \rceil}^{\lfloor (1+\gamma)\frac{n}{4} \rfloor} \binom{n}{i} \right)^2 \\ &\leq \left(2^{H(\frac{1+\gamma}{4})n} \right)^2 = 2^{2H(\frac{1+\gamma}{4})n} \end{aligned}$$

□

Our second upper bound for the case $Q = 2$ follows from restating the results from [25, Corollary 3.5].

Theorem 7.8. *[25, Corollary 3.5] Let $A \subseteq [2]^n, B \subseteq [2]^n$ be γ-uniform ($\gamma \leq 2$). Then:*

$$|A||B| \leq \binom{n}{\frac{\gamma n}{2}} 2^{n(1+\frac{\gamma}{2})} \leq 2^{n(1+\frac{\gamma}{2}+H(\frac{\gamma}{2}))}$$

We now address the case of general Q. Our next theorem provides the upper bound for this case, given the upper bound for $Q = 2$. Its corollary below follows from plugging in the bounds of Theorems 7.6 and 7.8.

Theorem 7.9. *Let $A \subseteq [Q]^n, B \subseteq [Q]^n$ be γ-uniform with $Q = 2^q$. In addition, let α be such that for any γ-uniform pair $A' \subseteq [2]^n, B' \subseteq [2]^n$ it holds that $|A'||B'| \leq 2^{\alpha n}$. Then:*

$$|A||B| \leq \begin{cases} \left(2^{\lceil q \rceil}\right)^{\alpha n} & Q \text{ is even} \\ \left(2^{4\lceil \frac{q}{4} \rceil}\right)^{\alpha n} & Q \text{ is odd} \end{cases}$$

.

Corollary 7.1. *Let $A \subseteq [Q]^n, B \subseteq [Q]^n$ be γ-uniform with $Q = 2^q$, $\gamma \leq 1$. Then:*

$$|A||B| \leq \min\left((2Q)^{\left(1+\frac{\gamma}{2}+H\left(\frac{\gamma}{2}\right)\right)n}, (2Q)^{2H\left(\frac{1+\gamma}{4}\right)n}\right)$$

In order to prove Theorem 7.9 we need a few other lemmas. The next two lemmas provide a framework for creating a reduction (in our upper bound) from any Q to $Q = 2$.

Lemma 7.5. *Let $Q \geq 2$ be even and let $f, g : [Q] \to \{0,1\}^{\lceil \log Q \rceil}$ be any functions such that for any $1 \leq i \leq \lceil \log Q \rceil$ it holds that*

$$|\{x \in [Q] \mid f(x)_i = 0\}| = \frac{Q}{2}$$
$$|\{y \in [Q] \mid g(y)_i = 0\}| = \frac{Q}{2}$$

Define functions $f^{(n)}, g^{(n)} : [Q]^n \to \{0,1\}^{\lceil \log Q \rceil n}$ to be

$$f^{(n)}(x^{(n)}) = f(x_1^{(n)}) f(x_2^{(n)}) \dots f(x_n^{(n)})$$
$$g^{(n)}(y^{(n)}) = g(y_1^{(n)}) g(y_2^{(n)}) \dots g(y_n^{(n)})$$

Let $A \subseteq [Q]^n, B \subseteq [Q]^n$ and define subsets $A' \subseteq \{0,1\}^{\lceil \log Q \rceil n}, B' \subseteq \{0,1\}^{\lceil \log Q \rceil n}$ to be

$$A' = \{f^{(n)}(x^{(n)}) \mid x^{(n)} \in A\}$$
$$B' = \{g^{(n)}(y^{(n)}) \mid y^{(n)} \in B\}$$

If A, B are γ-uniform then A', B' are γ-uniform.

Proof. Let f, g be functions as in the Lemma. For any $1 \leq i \leq \lceil \log Q \rceil$ and some $(\alpha, \beta) \in \{0,1\}^2$ let us define $S_{f,g,i}(\alpha, \beta) = \{(x, y) \in [Q]^2 \mid (f(x)_i, g(y)_i) = (\alpha, \beta)\}$. Namely, $S_{f,g,i}(\alpha, \beta)$ is the set of pairs in $[Q]^2$ that "generate" the pair (α, β) at position i, with respect to functions f and g. Note that the restrictions on f, g in the Lemma imply that $|S_{f,g,i}(\alpha, \beta)| = \frac{Q^2}{4}$.

Let A, B be γ-uniform, $(x^{(n)}, y^{(n)}) \in A \times B$ and consider the number of times a pair $(\alpha, \beta) \in \{0,1\}^2$ appears in $\left(f^{(n)}(x^{(n)}), g^{(n)}(y^{(n)})\right)$. It is equal to

$$\sum_{i=1}^{\lceil \log Q \rceil} \sum_{(x,y) \in S_{f,g,i}(\alpha,\beta)} |\{k \in [n] \mid (x_k^{(n)}, y_k^{(n)}) = (x, y)\}| \tag{10}$$

Since $(x^{(n)}, y^{(n)})$ is γ-uniform, $|\{k \in [n] \mid (x_k^{(n)}, y_k^{(n)}) = (x, y)\}|$ is bounded by $(1 \pm \gamma)\frac{n}{Q^2}$. Thus, Expression 10 is bounded by $\lceil \log Q \rceil \cdot \frac{Q^2}{4} \cdot (1 \pm \gamma)\frac{n}{Q^2} = (1 \pm \gamma)\frac{\lceil \log Q \rceil n}{4}$. Hence $\left(f^{(n)}(x^{(n)}), g^{(n)}(y^{(n)})\right)$ is γ-uniform. □

Lemma 7.6. *Let $Q \geq 3$ be odd and let $f, g : [Q] \to \{0,1\}^{4\lceil \frac{\log Q}{4} \rceil}$ be any functions where for any $0 \leq i < \lceil \frac{\log Q}{4} \rceil$ there exist distinct $j_1, j_2, j_3, j_4 \in [4]$ such that it holds that*

$$|\{x \in [Q] \mid f(x)_{4i+j_1} = 0\}| = \lfloor \frac{Q}{2} \rfloor \text{ and } |\{y \in [Q] \mid g(y)_{4i+j_1} = 0\}| = \lfloor \frac{Q}{2} \rfloor$$
$$|\{x \in [Q] \mid f(x)_{4i+j_2} = 0\}| = \lceil \frac{Q}{2} \rceil \text{ and } |\{y \in [Q] \mid g(y)_{4i+j_2} = 0\}| = \lceil \frac{Q}{2} \rceil$$
$$|\{x \in [Q] \mid f(x)_{4i+j_3} = 0\}| = \lceil \frac{Q}{2} \rceil \text{ and } |\{y \in [Q] \mid g(y)_{4i+j_3} = 0\}| = \lfloor \frac{Q}{2} \rfloor$$
$$|\{x \in [Q] \mid f(x)_{4i+j_4} = 0\}| = \lfloor \frac{Q}{2} \rfloor \text{ and } |\{y \in [Q] \mid g(y)_{4i+j_4} = 0\}| = \lceil \frac{Q}{2} \rceil$$

Define functions $f^{(n)}, g^{(n)} : [Q]^n \to \{0,1\}^{4\lceil \frac{\log Q}{4} \rceil n}$ to be

$$f^{(n)}(x^{(n)}) = f(x_1^{(n)}) f(x_2^{(n)}) \dots f(x_n^{(n)})$$
$$g^{(n)}(y^{(n)}) = g(y_1^{(n)}) g(y_2^{(n)}) \dots g(y_n^{(n)})$$

Let $A \subseteq [Q]^n, B \subseteq [Q]^n$ and define subsets $A' \subseteq \{0,1\}^{4\lceil \frac{\log Q}{4} \rceil n}, B' \subseteq \{0,1\}^{4\lceil \frac{\log Q}{4} \rceil n}$ to be

$$A' = \{f^{(n)}(x^{(n)}) \mid x^{(n)} \in A\}$$
$$B' = \{g^{(n)}(y^{(n)}) \mid y^{(n)} \in B\}$$

If A, B are γ-uniform then A', B' are γ-uniform.

Proof. Let f, g be functions as in the Lemma. For any $0 \leq i < \lceil \frac{\log Q}{4} \rceil$, $j \in [4]$ and some $(\alpha, \beta) \in \{0,1\}^2$ let us define $S_{f,g,i,j}(\alpha, \beta) = \{(x, y) \in [Q]^2 \mid (f(x)_{4i+j}, g(y)_{4i+j}) = (\alpha, \beta)\}$. Namely, $S_{f,g,i,j}(\alpha, \beta)$ is the set of pairs in $[Q]^2$ that "generate" the pair (α, β) at position $4i+j$, with respect to functions f and g. Note that the restrictions on f, g in the Lemma imply that $\sum_{j=1}^{4} |S_{f,g,i,j}(\alpha, \beta)| = Q^2$.

Let A, B be γ-uniform, $(x^{(n)}, y^{(n)}) \in A \times B$ and consider the number of times a pair $(\alpha, \beta) \in \{0,1\}^2$ appears in $\left(f^{(n)}(x^{(n)}), g^{(n)}(y^{(n)})\right)$. It is equal to

$$\sum_{0 \leq i < \lceil \frac{\log Q}{4} \rceil} \sum_{j=1}^{4} \sum_{(x,y) \in S_{f,g,i,j}(\alpha,\beta)} |\{k \in [n] \mid (x_k^{(n)}, y_k^{(n)}) = (x, y)\}| \tag{11}$$

Since $(x^{(n)}, y^{(n)})$ is γ-uniform, $|\{k \in [n] \mid (x_k^{(n)}, y_k^{(n)}) = (x, y)\}|$ is bounded by $(1 \pm \gamma)\frac{n}{Q^2}$. Thus, Expression 11 is bounded by $\lceil \frac{\log Q}{4} \rceil \cdot Q^2 \cdot (1 \pm \gamma)\frac{n}{Q^2} = (1 \pm \gamma)\lceil \frac{\log Q}{4} \rceil n = (1 \pm \gamma)\frac{4\lceil \frac{\log Q}{4} \rceil n}{4}$. Hence $\left(f^{(n)}(x^{(n)}), g^{(n)}(y^{(n)})\right)$ is γ-uniform. □

Now we can prove Theorem 7.9.

Proof. (Theorem 7.9) Let $A \subseteq [Q]^n, B \subseteq [Q]^n$ be γ-uniform. For $Q = 2$, the upper bound holds trivially by the conditions of the theorem. The upper bounds for other Q-s, even or odd, are obtained by constructing A', B' via Lemma 7.5 or Lemma 7.6 respectively, and then by using the upper bound of the theorem's condition on $|A'||B'|$. All that remains is to prove the existence of f, g that satisfy the conditions of Lemma 7.5 or Lemma 7.6 and are injective.

Indeed, for the even $Q \geq 4$ define f greedily in the following way: let $P = \{P_1, P_2, \ldots, P_{\frac{Q}{2}}\}$ be any partition of $[Q]$ such that $|P_i| = 2$ (for any i). For each $P_i = \{v_1, v_2\}$ set $f(v_1)$ to any previously unused value $w \in \{0,1\}^{\lceil q \rceil}$, and set $f(v_2)$ to the bitwise inverse of w (0-s replaced with 1-s and 1-s replaced with 0-s). Also set $g = f$. It is easy to see that f, g satisfy the conditions of Lemma 7.5 and are injective.

For the odd $Q \geq 3$, define f, g in the following manner: let

$$P = \{P_1, P_2, \ldots, P_{\lfloor \frac{Q}{2} \rfloor}, P_{\lceil \frac{Q}{2} \rceil}\}$$
$$P' = \{P'_1, P'_2, \ldots, P'_{\lfloor \frac{Q}{2} \rfloor}, P'_{\lceil \frac{Q}{2} \rceil}\}$$

be two partitions of $[Q]$ such that

$$|P_i| = |P'_i| = \begin{cases} 1 & \text{if } i = \lceil \frac{Q}{2} \rceil \\ 2 & \text{otherwise} \end{cases}$$

For each $P_i = \{v_1, v_2\}$ $(1 \leq i \leq \lfloor \frac{Q}{2} \rfloor)$, set $f(v_1)$ to any previously unused value $w \in \{0,1\}^{4\lceil \frac{q}{4} \rceil}$, such that w is not one of the two *alternating* strings in the set $\{0101\ldots0101, 1010\ldots1010\}$, and set $f(v_2)$ to the bitwise inverse of w. For $P_{\lceil \frac{Q}{2} \rceil} = \{v_1\}$, set $f(v_1) = 0101\ldots0101$.

Similarly, for each $P'_i = \{v_1, v_2\}$ $(1 \leq i \leq \lfloor \frac{Q}{2} \rfloor)$, set $g(v_1)$ to any previously unused value $w \in \{0,1\}^{4\lceil \frac{q}{4} \rceil}$, such that $w \notin \{0110\ldots0110, 1001\ldots1001\}$, and set $g(v_2)$ to the bitwise inverse of w. For $P'_{\lceil \frac{Q}{2} \rceil} = \{v_1\}$, set $g(v_1) = 0110\ldots0110$.

To see that the conditions are satisfied, observe that for function f we have

$$|\{x \in [Q] \mid f(x)_j = 0\}| = \begin{cases} \lceil \frac{Q}{2} \rceil & \text{if } j \text{ is odd} \\ \lfloor \frac{Q}{2} \rfloor & \text{otherwise} \end{cases}$$

and for function g we have

$$|\{y \in [Q] \mid g(y)_j = 0\}| = \begin{cases} \lceil \frac{Q}{2} \rceil & \text{if } j = 4i \text{ or } j = 4i+1 \text{ for any } i \geq 0 \\ \lfloor \frac{Q}{2} \rfloor & \text{otherwise} \end{cases}$$

It is now easy to see that the conditions of Lemma 7.6 are satisfied for $j_1 = 2$, $j_2 = 1$, $j_3 = 3$, $j_4 = 4$, and clearly f, g are injective. □

7.3 Implications on $\mathcal{R}_W^{(0)}$

In this section we study the implications of our analysis of γ-uniform and (d, ε)-diverse set systems on the rate $\mathcal{R}_W^{(0)}$ as expressed in Theorem 7.1 and Theorem 7.2. Specifically,

we "plug in" the upper and lower bounds obtained in the previous sections to obtain upper and lower bounds on $\mathcal{R}_W^{(0)}$ based on our suggested encoding scheme. As we will see, the upper and lower bounds presented above combined with Theorem 7.1 do not resolve the question whether $\mathcal{R}_W^{(0)}$ is greater than q. Namely, we show that the lower bound on γ-uniform set systems do not imply that $\mathcal{R}_W^{(0)} > q$. In addition, to put our results in context, we also show that an optimistic assumption that there exist γ-uniform set systems that match the upper bound of the previous section will indeed imply that $\mathcal{R}_W^{(0)} > q$, however our upper bound may be loose and such set systems are not known to exist. All in all, even though we cannot conclude any bounds on the value of $\mathcal{R}_W^{(0)}$ for our channels $W \in \mathcal{W}_{Q,\varepsilon}$, we believe that the concept of γ-uniform set systems is an interesting one and that a better understanding of bounds for such systems may give a better understanding on the value of $\mathcal{R}_W^{(0)}$.

Claim 7.1. *Let $\varepsilon \geq \frac{1}{4}$. In this case, Theorem 7.2 will not imply $\mathcal{R}_W^{(0)} > q$.*

In order to prove this, we first need the following lemma:

Lemma 7.7. *Let $(x, y) \in [Q]^n \times [Q]^n$ be a (d, ε)-diverse pair. Then:*

$$d \geq \varepsilon - \frac{1}{Q^2} + \frac{1}{n}$$

Proof. Consider the pairs $(\alpha, \beta) \in [Q] \times [Q]$ that appear in (x, y) sorted in descending order by the number of their appearances. Let us take the first t such (α, β) pairs. The sum of all appearances of these pairs should be greater or equal than t multiplied by the average number of appearances of all (α, β) pairs, namely:

$$\sum_{i=1}^{t} (\#(\alpha, \beta)_i) \geq t \frac{n}{Q^2}$$

Here, $(\alpha, \beta)_i$ is the i'th pair according to the ordering mentioned above, and $\#(\alpha, \beta)_i$ is the number of appearences of the pair $(\alpha, \beta)_i$. In particular this means that if we take $t = \varepsilon Q^2 - 1$, then:

$$\sum_{i=1}^{\varepsilon Q^2 - 1} (\#(\alpha, \beta)_i) \geq \left(\varepsilon Q^2 - 1\right) \frac{n}{Q^2}$$

Now, consider the coordinate set J consisting of locations j for which (x_j, y_j) is in the set $\{(\alpha, \beta)_i\}_{i=1}^{\varepsilon Q^2 - 1}$. Since (x, y) is (d, ε)-diverse, every dn coordinates of (x, y) must contain at least εQ^2 distinct pairs, thus $dn > |J|$. Namely,

$$\begin{aligned} dn &\geq \left(\varepsilon Q^2 - 1\right) \frac{n}{Q^2} + 1 \\ dn &\geq \varepsilon n - \frac{n}{Q^2} + 1 \\ d &\geq \varepsilon - \frac{1}{Q^2} + \frac{1}{n} \end{aligned}$$

□

Proof. (Claim 7.1) Let $\varepsilon \geq \frac{1}{4}$. Consider the setting of Theorem 7.2. Any $(x, y) \in A \times B$ is $(d, 2\varepsilon)$-diverse. Thus by Lemma 7.7:

$$d \geq 2\varepsilon - \frac{1}{Q^2} + \frac{1}{n} \geq \frac{1}{2} - \frac{1}{Q^2} + \frac{1}{n}$$

By substituting this lower bound for d in the rate of Theorem 7.2, we get:

$$\begin{aligned}
2q\left(1 - \left(\frac{\delta_1 + \delta_2}{2} + d\right)\right) - 2 &\leq 2q\left(1 - \left(\frac{\delta_1 + \delta_2}{2} + \frac{1}{2} - \frac{1}{Q^2} + \frac{1}{n}\right)\right) - 2 \\
&\leq 2q\left(1 - \left(\frac{1}{2} - \frac{1}{Q^2} + \frac{1}{n}\right)\right) - 2 \\
&= 2q\left(\frac{1}{2} + \frac{1}{Q^2} - \frac{1}{n}\right) - 2 \\
&= q + \frac{2q}{Q^2} - \frac{2q}{n} - 2 \\
&= q + \frac{2q}{2^{2q}} - \frac{2q}{n} - 2 \\
&< q
\end{aligned}$$

□

Claim 7.2. *Let $A \subseteq [Q]^n, B \subseteq [Q]^n$ be γ-uniform such that $|A||B|$ is equal to the lower bound in Proposition 7.5. If $\varepsilon > \frac{1}{4Q^2}$ then the RHS of Equation 7 in Theorem 7.1 is no larger than $q - 1$.*

Proof. We will prove a stronger statement by using a lower bound of $2^{nq(1+\gamma/Q^2+1/q)}$ which is larger than the lower bound in Proposition 7.5.

To see that it is indeed larger, recall that the lower bound in Proposition 7.5 is:

$$\begin{aligned}
\left(\frac{1}{n+1}\right)\left(\frac{Q}{e\sqrt{n}}\right)^{Q(Q-1)} Q^{n\left(1+\frac{\log(Q-1)}{\log Q}\cdot\frac{\gamma}{Q^2}+\frac{1}{\log Q}\cdot H\left(\frac{\gamma}{Q^2}\right)\right)} &< Q^{n\left(1+\frac{\log(Q-1)}{\log Q}\cdot\frac{\gamma}{Q^2}+\frac{1}{\log Q}\cdot H\left(\frac{\gamma}{Q^2}\right)\right)} \\
&= 2^{nq\left(1+\frac{\log(Q-1)}{q}\cdot\frac{\gamma}{Q^2}+\frac{1}{q}\cdot H\left(\frac{\gamma}{Q^2}\right)\right)} \\
&< 2^{nq(1+\gamma/Q^2+1/q)}.
\end{aligned}$$

Let $|A| = Q^{n(1-\delta_1)}$ and $|B| = Q^{n(1-\delta_2)}$. By our definitions

$$2^{nq(2-\delta_1-\delta_2)} = 2^{nq(1-\delta_1)+nq(1-\delta_2)} = |A||B| = 2^{nq(1+\gamma/Q^2+1/q)}$$

Thus:

$$1 + \gamma/Q^2 + 1/q = 2 - \delta_1 - \delta_2$$

Which means that:

$$\frac{\delta_1 + \delta_2}{2} = \frac{1}{2} - \frac{\gamma}{2Q^2} - \frac{1}{2q}$$

Thus we can compute the RHS of Equation 7 as:

$$2q\left(1-\left(\frac{\delta_1+\delta_2}{2}+2(1+\gamma)\varepsilon+\delta\right)\right)-2=$$
$$2q\left(1-\left(\frac{1}{2}-\frac{\gamma}{2Q^2}-\frac{1}{2q}+2(1+\gamma)\varepsilon+\delta\right)\right)-2=$$
$$2q\left(\frac{1}{2}+\frac{\gamma}{2Q^2}-2(1+\gamma)\varepsilon-\delta\right)-1$$

And since $\varepsilon > \frac{1}{4Q^2}$ the whole expression is less than $q-1$. □

In the following claim, we assume for simplicity that $Q = 2^q$ for an integer q.

Claim 7.3. *Let $A \subseteq [Q]^n, B \subseteq [Q]^n$ be γ-uniform such that $|A||B|$ is equal to the upper bound in Corollary 7.1. Then the RHS of Equation 7 in Theorem 7.1 is larger than q for*

$$\varepsilon < \frac{\max\left(\gamma + 2H(\gamma/2), 4H((1+\gamma)/4) - 2\right)}{8(1+\gamma)},$$

sufficiently large Q, and sufficiently small δ.

Proof. We will prove a stronger statement by using an upper bound of $\min\left(2^{nq\left(1+\frac{\gamma}{2}+H\left(\frac{\gamma}{2}\right)\right)}, 2^{nq\left(2H\left(\frac{1+\gamma}{4}\right)\right)}\right)$ which is clearly smaller than the upper bound in Corollary 7.1.

Let $|A| = Q^{n(1-\delta_1)}$ and $|B| = Q^{n(1-\delta_2)}$. By our definitions

$$2^{nq(2-\delta_1-\delta_2)} = 2^{nq(1-\delta_1)+nq(1-\delta_2)} = |A||B| = \min\left(2^{nq\left(1+\frac{\gamma}{2}+H\left(\frac{\gamma}{2}\right)\right)}, 2^{nq\left(2H\left(\frac{1+\gamma}{4}\right)\right)}\right)$$

Thus:

$$2-\delta_1-\delta_2 = \min\left(1+\frac{\gamma}{2}+H\left(\frac{\gamma}{2}\right), 2H\left(\frac{1+\gamma}{4}\right)\right)$$

Which means that:

$$\frac{\delta_1+\delta_2}{2} = \min\left(\frac{1}{2}-\frac{\gamma}{4}-\frac{H\left(\frac{\gamma}{2}\right)}{2}, 1-H\left(\frac{1+\gamma}{4}\right)\right)$$

Thus the RHS of Equation 7 is:

$$2q\left(1-\left(\frac{\delta_1+\delta_2}{2}+2(1+\gamma)\varepsilon+\delta\right)\right)-2=$$
$$2q\left(1-\left(\min\left(\frac{1}{2}-\frac{\gamma}{4}-\frac{H\left(\frac{\gamma}{2}\right)}{2}, 1-H\left(\frac{1+\gamma}{4}\right)\right)+2(1+\gamma)\varepsilon+\delta\right)\right)-2=$$
$$q\max\left(1+\frac{\gamma}{2}+H\left(\frac{\gamma}{2}\right)-4\varepsilon(1+\gamma)-2\delta, 2H\left(\frac{1+\gamma}{4}\right)-4\varepsilon(1+\gamma)-2\delta\right)-2$$

Consider the coefficient of q in both cases of the maximum function. In the first case, for sufficiently small δ, the coefficient of q is greater than 1 if

$$\begin{aligned} 1+\frac{\gamma}{2}+H\left(\frac{\gamma}{2}\right)-4\varepsilon(1+\gamma) &> 1 \\ 4\varepsilon(1+\gamma) &< \frac{\gamma}{2}+H\left(\frac{\gamma}{2}\right) \\ \varepsilon &< \frac{\gamma+2H\left(\frac{\gamma}{2}\right)}{8(1+\gamma)} \end{aligned}$$

In the second case, for sufficiently small δ, the coefficient of q is greater than 1 if

$$\begin{aligned} 2H\left(\frac{1+\gamma}{4}\right)-4\varepsilon(1+\gamma) &> 1 \\ 4\varepsilon(1+\gamma) &< 2H\left(\frac{1+\gamma}{4}\right)-1 \\ \varepsilon &< \frac{4H\left(\frac{1+\gamma}{4}\right)-2}{8(1+\gamma)} \end{aligned}$$

Putting it all together, it means that for

$$\varepsilon < \frac{\max\left(\gamma+2H(\gamma/2), 4H((1+\gamma)/4)-2\right)}{8(1+\gamma)},$$

a sufficiently small δ and a sufficiently large Q, the RHS of Equation 7 is greater than q, as required. □

We finally, present a general claim that computes the implications (in terms of $\mathcal{R}_W^{(0)}$) of γ-uniform sets $A \subseteq [Q]^n, B \subseteq [Q]^n$ such that $|A||B| \geq Q^{n(1+f(\gamma))}$ for some function f.

Claim 7.4. *Assume the existence of γ-uniform sets $A \subseteq [Q]^n, B \subseteq [Q]^n$ such that $|A||B| = Q^{n(1+f(\gamma))}$. Then if for some constant $c > 0$ it holds that $f(\gamma) > 4\varepsilon(1+\gamma)+\frac{2}{q}+c$, we have by Theorem 7.1 that $\mathcal{R}_W^{(0)} > q$.*

Proof. Let $|A| = Q^{n(1-\delta_1)}$ and $|B| = Q^{n(1-\delta_2)}$.

$$Q^{n(1-\delta_1)}Q^{n(1-\delta_2)} = |A||B| = Q^{n(1+f(\gamma))}$$

Thus:

$$1+f(\gamma) = 2-\delta_1-\delta_2$$

Which means that:

$$\frac{\delta_1+\delta_2}{2} = \frac{1}{2}-\frac{f(\gamma)}{2}$$

Thus the RHS of Equation 7 is:

$$2q\left(1-\left(\frac{\delta_1+\delta_2}{2}+2(1+\gamma)\varepsilon+\delta\right)\right)-2=$$
$$2q\left(1-\left(\frac{1}{2}-\frac{f(\gamma)}{2}+2(1+\gamma)\varepsilon+\delta\right)\right)-2=$$
$$q\left(1+f(\gamma)-4\varepsilon(1+\gamma)-2\delta\right)-2$$

If $f(\gamma) > 4\varepsilon(1+\gamma)+\frac{2}{q}+c$ then the RHS of equation 7 is bounded from below by $q(1+c-2\delta)$. Since we can select any $\delta > 0$, we have that $q(1+c-2\delta) > q$. □

7.4 δ-additively-uniform set systems

We conclude this section with a study of γ-uniform sets that are extremely close to being 0-uniform. Namely, we address sets that have an *additive* gap from being 0-uniform (rather than a *multiplicative* gap in context of γ-uniform sets). Our precise definitions are given below.

Definition 7.3. *A pair $(x^{(n)}, y^{(n)}) \in [Q]^n \times [Q]^n$ is called δ-additively-uniform if for each pair $(\alpha, \beta) \in [Q]^2$ it holds that*

$$\frac{n}{Q^2}-\delta \leq |\{i \in [n] \mid (x_i, y_i) = (\alpha, \beta)\}| \leq \frac{n}{Q^2}+\delta$$

In other words, the number of appearances of any pair $(\alpha, \beta) \in [Q]^2$ in $(x^{(n)}, y^{(n)})$ is bounded by $\frac{n}{Q^2} \pm \delta$. Here, δ can be considered an integer. Similarly, the subsets $A \subseteq [Q]^n, B \subseteq [Q]^n$ are called δ-additively-uniform if for any $x^{(n)} \in A, y^{(n)} \in B$, $(x^{(n)}, y^{(n)})$ are δ-additively-uniform.

Remark 7.2. *A γ-uniform set system is thus a special case of a δ-additively-uniform set system where $\delta = \gamma\frac{n}{Q^2}$.*

The study of δ-additively-uniform sets does not have direct implications on $\mathcal{R}_W^{(0)}$. Nevertheless, we were intrigued by such sets and during our studies we tried to bound the size $|A||B|$ of δ-additively-uniform set systems from above. We believe that such a bound should be closely related to the upper bounds on 0-uniform sets systems we exhibit in Theorem 7.5 (which is $|A||B| \leq Q^n$). We note that our upper bounds presented earlier on γ-uniform sets imply bounds for δ-additively-uniform sets (by setting $\gamma = \delta\frac{Q^2}{n}$). However, these bounds (e.g., Corollary 7.1) do not approach $O(Q^n)$ even for $\delta = 1$ (an additive gap of 1!). In our studies, we were not able to obtain improved upper bounds for δ-additively-uniform sets. In what follows we present a claim we found interesting. Our claim addresses only the case of $Q = 2$, i.e., where the alphabet is $\{0, 1\}$.

Let $c_0(x), c_1(x)$ denote the number of zeroes and ones in $x \in \{0,1\}^n$ respectively. Also let $c_{0,0}(x,y), c_{0,1}(x,y), c_{1,0}(x,y), c_{1,1}(x,y)$ denote the number of appearances of $(0,0), (0,1), (1,0), (1,1)$ in $(x,y) \in \{0,1\}^n \times \{0,1\}^n$ respectively. We start with a small technical lemma.

Lemma 7.8. *If $(x, y) \in \{0,1\}^n \times \{0,1\}^n$ are δ-additively-uniform then*

$$\frac{n}{2} - 2\delta \leq c_0(x) \leq \frac{n}{2} + 2\delta$$
$$\frac{n}{2} - 2\delta \leq c_1(x) \leq \frac{n}{2} + 2\delta$$
$$\frac{n}{2} - 2\delta \leq c_0(y) \leq \frac{n}{2} + 2\delta$$
$$\frac{n}{2} - 2\delta \leq c_1(y) \leq \frac{n}{2} + 2\delta$$

Proof. This follows from the fact that

$$c_0(x) = c_{0,0}(x, y) + c_{0,1}(x, y)$$
$$c_1(x) = c_{1,0}(x, y) + c_{1,1}(x, y)$$

and similarly

$$c_0(y) = c_{0,0}(x, y) + c_{1,0}(x, y)$$
$$c_1(y) = c_{0,1}(x, y) + c_{1,1}(x, y)$$

□

Definition 7.4. *Let $x', x \in \{0,1\}^n$. We say that $x' \succeq x$ if x' can be obtained from x by "balancing", namely flipping t zeroes to ones where $0 \leq t \leq \left(c_0(x) - \frac{n}{2}\right)$ if $c_0(x) > c_1(x)$, or flipping t ones to zeroes where $0 \leq t \leq \left(c_1(x) - \frac{n}{2}\right)$ otherwise.*

Claim 7.5. *If $A \subseteq \{0,1\}^n, B \subseteq \{0,1\}^n$ are δ-additively-uniform then $\cup_{x \in A} S(x), \cup_{y \in B} S(y)$ are also δ-additively-uniform, where $S(x) = \{x' \mid x' \succeq x\}$.*

Proof. Let $x \in A$. Assume, w.l.o.g. that $c_0(x) > c_1(x)$, and let $x' \succeq x$. We can write

$$c_0(x) = \frac{n}{2} + t'$$
$$c_0(x') = \frac{n}{2} + t' - t$$

where by Lemma 7.8, $0 < t' \leq 2\delta$ and $t \leq t'$.

Now, for any $y \in B$ we can express:

$$c_{0,0}(x, y) = \frac{n}{4} + l_{t'}(y)$$
$$c_{0,1}(x, y) = \frac{n}{4} + r_{t'}(y)$$

Such that

$$l_{t'}(y) + r_{t'}(y) = t'$$
$$|l_{t'}(y)| \leq \delta$$
$$|r_{t'}(y)| \leq \delta$$

Note that $l_{t'}(y)$ or $r_{t'}(y)$ can also be negative. Thus,

$$c_{0,0}(x', y) = \frac{n}{4} + l_{t'}(y) - l_t(y)$$
$$c_{0,1}(x', y) = \frac{n}{4} + r_{t'}(y) - r_t(y)$$

Such that $l_t(y) + r_t(y) = t$, and $l_t(y), r_t(y) \geq 0$. Clearly,

$$c_{0,0}(x', y) \leq \frac{n}{4} + \delta$$
$$c_{0,1}(x', y) \leq \frac{n}{4} + \delta$$

Note that $l_t(y)$ is maximized if $l_t(y) = t = t'$. But even in this case:

$$\begin{aligned} c_{0,0}(x', y) &= \frac{n}{4} + l_{t'}(y) - l_t(y) \\ &= \frac{n}{4} + l_{t'}(y) - t' \\ &= \frac{n}{4} + l_{t'}(y) - (l_{t'}(y) + r_{t'}(y)) \\ &= \frac{n}{4} - r_{t'}(y) \\ &\geq \frac{n}{4} - \delta \end{aligned}$$

Similarly, $r_t(y)$ is maximized if $r_t(y) = t = t'$. And in this case:

$$\begin{aligned} c_{0,1}(x', y) &= \frac{n}{4} + r_{t'}(y) - r_t(y) \\ &= \frac{n}{4} + r_{t'}(y) - t' \\ &= \frac{n}{4} + r_{t'}(y) - (l_{t'}(y) + r_{t'}(y)) \\ &= \frac{n}{4} - l_{t'}(y) \\ &\geq \frac{n}{4} - \delta \end{aligned}$$

Similarly, let us write

$$c_1(x) = \frac{n}{2} - t'$$
$$c_1(x') = \frac{n}{2} - t' + t$$

where $0 < t' \leq 2\delta$ and $t \leq t'$.

And for any $y \in B$ we can express:

$$c_{1,0}(x, y) = \frac{n}{4} - l_{t'}(y)$$
$$c_{1,1}(x, y) = \frac{n}{4} - r_{t'}(y)$$

Such that

$$\begin{aligned} l_{t'}(y) + r_{t'}(y) &= t' \\ |l_{t'}(y)| &\leq \delta \\ |r_{t'}(y)| &\leq \delta \end{aligned}$$

Note that $l_{t'}(y)$ or $r_{t'}(y)$ can also be negative. Thus,

$$\begin{aligned} c_{1,0}(x', y) &= \frac{n}{4} - l_{t'}(y) + l_t(y) \\ c_{1,1}(x', y) &= \frac{n}{4} - r_{t'}(y) + r_t(y) \end{aligned}$$

Such that $l_t(y) + r_t(y) = t$, and $l_t(y), r_t(y) \geq 0$. Clearly,

$$\begin{aligned} c_{1,0}(x', y) &\geq \frac{n}{4} - \delta \\ c_{1,1}(x', y) &\geq \frac{n}{4} - \delta \end{aligned}$$

Note that $l_t(y)$ is maximized if $l_t(y) = t = t'$. But even in this case:

$$\begin{aligned} c_{1,0}(x', y) &= \frac{n}{4} - l_{t'}(y) + l_t(y) \\ &= \frac{n}{4} - l_{t'}(y) + t' \\ &= \frac{n}{4} - l_{t'}(y) + (l_{t'}(y) + r_{t'}(y)) \\ &= \frac{n}{4} + r_{t'}(y) \\ &\leq \frac{n}{4} + \delta \end{aligned}$$

Similarly, $r_t(y)$ is maximized if $r_t(y) = t = t'$. And in this case:

$$\begin{aligned} c_{1,1}(x', y) &= \frac{n}{4} - r_{t'}(y) + r_t(y) \\ &= \frac{n}{4} - r_{t'}(y) + t' \\ &= \frac{n}{4} - r_{t'}(y) + (l_{t'}(y) + r_{t'}(y)) \\ &= \frac{n}{4} + l_{t'}(y) \\ &\leq \frac{n}{4} + \delta \end{aligned}$$

Thus, (x', y) are δ-additively-uniform. The proof for (x, y') is similar, and the claim follows. □

8 k-user networks

In the previous sections we focused on the two user erasure/identity networks. Our model and definitions can be extended quite naturally for multiple users. Instead of representing channels with graphs, we now represent them with hypergraphs. In what follows we briefly reiterate on the model and definitions, extending them to k-users. We do not elaborate on any technical claims regarding the extended model.

8.1 Our model revisited

Recall that in a multiple unicast communication network the objective is for k source nodes, $s_1, s_2, \ldots, s_k$, to communicate their information to k corresponding terminal nodes, $t_1, t_2, \ldots, t_k$ over a channel W. One can model a deterministic multiple unicast communication network with block length n by the following components.

Message space: For $i = 1 \ldots k$, source s_i holds a message from a set of size M_i. Without loss of generality, the message space can be considered $[M_i] = \{1, \ldots, M_i\}$.

Encoding: For alphabet $[Q] = [2^q]$ and block length n, each source s_i holds an encoding function $E_i : [M_i] \to [Q]^n$.

Network W: The network $W : [Q]^k \to [Q]^k$ is a deterministic function which takes as input elements from $[Q]^k$ and returns elements from the same alphabet. Namely, denoting W as $(W_1, \ldots, W_k)$, terminal t_i will receive the evaluation of $W_i : [Q]^k \to [Q]$ on input $(x_1, \ldots, x_k) \in [Q]^k$.

Network $W^{(n)}$: Applying the network W multiple times (over block length n) yields the network $W^{(n)} : ([Q]^n)^k \to ([Q]^n)^k$ which is a deterministic function that takes as input k n-vectors and returns k n-vectors. Namely, denoting $W^{(n)}$ as $(W_1^{(n)}, \ldots, W_k^{(n)})$, the evaluation of $W_i^{(n)} : ([Q]^n)^k \to [Q]^n$ on input $({x_1}^{(n)}, \ldots, {x_k}^{(n)}) \in ([Q]^n)^k$, is a vector $\widehat{x_i}^{(n)} \in [Q]^n$ received at terminal t_i, where $(\widehat{x_i}^{(n)})_j = W_i(({x_1}^{(n)})_j, \ldots, ({x_k}^{(n)})_j)$.

Decoding: Each terminal t_i holds a decoding function $D_i : [Q]^n \to [M_i]$.

Communication with block length n is successful for terminal t_i and source information $(m_1, \ldots, m_k)$ if for $i = 1 \ldots k$, $D_i[W_i^{(n)}(E_1(m_1), \ldots, E_k(m_k))] = m_i$. We say that communication is successful with probability $1-\varepsilon$ if for source information $(m_1, \ldots, m_k)$ chosen uniformly at random from $[M_1] \times \ldots \times [M_k]$ it holds with probability $1 - \varepsilon$ that communication is successful for all terminals. Rate $(R_1, \ldots, R_k)$ is achievable with probability $1 - \varepsilon$ and block length n over network W if for $M_i = 2^{R_i n}$ there exist encoding and decoding functions such that communication is successful with probability $1 - \varepsilon$.

The ε-error sum capacity of network W and block length n is defined to be

$$\mathcal{R}_{W,n}^{(\varepsilon)} = \sup_{(R_1,\ldots,R_k)\in\Gamma_{n,\varepsilon}} \left(\sum_{i=1}^{k} R_i \right),$$

where the supremum is taken over $\Gamma_{n,\varepsilon}$ which consists of all tuples $(R_1, \ldots, R_k)$ that are achievable with probability $1 - \varepsilon$ and block length n over W. The ε-error sum capacity of network W is defined to be

$$\mathcal{R}_W^{(\varepsilon)} = \sup_n \mathcal{R}_{W,n}^{(\varepsilon)}$$

8.2 Statement 3.1 revisited

The equivalent of Statement 3.1 is therefore:

Statement 8.1. *Let $\varepsilon > 0$. There exists $\delta = \delta(\varepsilon) > 0$ that tends to 0 when ε tends to 0 such that for every k-user network W it holds that*

$$\frac{\mathcal{R}_W^{(\varepsilon)}}{k} - \delta \leq \mathcal{R}_W^{(0)}.$$

Moreover, for any $\varepsilon > 0$ and δ as above, there exists a k-user network W_ε, such that:

$$\mathcal{R}_{W_\varepsilon}^{(0)} \leq \frac{\mathcal{R}_{W_\varepsilon}^{(\varepsilon)}}{k} + \delta.$$

In other words, for certain networks W, requiring zero-error in communication may reduce the sum capacity by a factor of k (or equivalently, allowing an ε error may increase the sum capacity by a factor of k), and this k-factor is tight.

8.3 Our definitions revisited

We further show how the definitions used in our work can be extended in a natural way to k-users.

Definition 8.1. *Let $\mathcal{W}_{Q,\varepsilon,k}$ be the distribution over erasure/identity channels in which for every $(x_1, \ldots, x_k) \in [Q]^k$ we fix $W(x_1, \ldots, x_k) = (\phi, \ldots, \phi)$ independently with probability ε (otherwise $W(x_1, \ldots, x_k) = (x_1, \ldots, x_k)$).*

Definition 8.2. *Let $W \in \mathcal{W}_{Q,\varepsilon,k}$. We say that a hypergraph $H = (V, E)$ represents channel W if H is a k-uniform k-partite hypergraph where $V = Q_1 \cup \ldots \cup Q_k$ and there exist bijections $f_i : [Q] \to Q_i$ for $i = 1 \ldots k$ such that for any $x_i \in [Q]$, $(f_1(x_1), \ldots, f_k(x_k)) \in E$ if and only if $W(x_1, \ldots, x_k) = (\phi, \ldots, \phi)$.*

Definition 8.3. *Given a k-partite hypergraph $H = (V_1 \cup \ldots \cup V_k, E)$, a k-partite independent set (KPIS) is a sequence $(A_i)_{i=1}^k$ such that $A_i \subseteq V_i$ and for any $(x_1, \ldots, x_k) \in A_1 \times \ldots \times A_k$, $(x_1, \ldots, x_k) \notin E$. We define the size of the KPIS $(A_i)_{i=1}^k$ to be $\Pi_{i=1}^k |A_i|$.*

Definition 8.4. *Let $W \in \mathcal{W}_{Q,\varepsilon,k}$, and consider the network $W^{(n)}$. We say that a hypergraph $H = (V, E)$ represents network $W^{(n)}$ if H is a k-uniform k-partite hypergraph where $V = Q_{1,n} \cup \ldots \cup Q_{k,n}$ and there exist bijections $f_i : [Q]^n \to Q_{i,n}$ such that for any $x_i^{(n)} \in [Q]^n$, $(f_1(x_1^{(n)}), \ldots, f_k(x_k^{(n)})) \in E$ if and only if there exists at least one index j such that $W({x_1^{(n)}}_j, \ldots, {x_k^{(n)}}_j) = (\phi, \ldots, \phi)$.*

Remark 8.1. *Clearly for any network $W^{(n)}$ where $W \in \mathcal{W}_{Q,\varepsilon,k}$, we can construct a hypergraph that represents it. In addition, all hypergraphs that represent $W^{(n)}$ are identical up to an isomorphism.*

9 Conclusion and open problems

In this thesis we address the potential gap between $\mathcal{R}_W^{(0)}$ and $\mathcal{R}_W^{(\varepsilon)}$ in the context of 2-source/2-terminal deterministic interference channels. Our work is motivated by similar questions in the context of network coding. In Statement 3.1 we conjecture that there exist channels W for which $\mathcal{R}_W^{(0)} \leq \mathcal{R}_W^{(\varepsilon)}/2 + \delta$ (and more generally for the k-source/k-terminal case that $\mathcal{R}_W^{(0)} \leq \mathcal{R}_W^{(\varepsilon)}/k + \delta$). Studying the channels that result from the distribution $\mathcal{W}_{Q,\varepsilon}$, we support Statement 3.1 by presenting upper bounds on $\mathcal{R}_{W,n}^{(0)}$ (which take into account the block length n) and by studying the limitations of a natural encoding scheme based on γ-uniform set systems. We view our posing of Statement 3.1, our upper bounds, and the study of γ-uniform set systems as the main contributions of this thesis. Whether Statement 3.1 is true or not remains open and will be subject to our future research.

Appendix

A Useful inequalities

Lemma A.1. *Let $\alpha > 0$. For any $0 < \varepsilon < 1$:*

$$\frac{\alpha}{\log \frac{1}{1-\varepsilon}} < \frac{\alpha \ln 2}{\varepsilon}$$

Proof. The inequality above is identical to:

$$\begin{aligned}
\frac{1}{\log \frac{1}{1-\varepsilon}} &< \frac{\ln 2}{\varepsilon} \\
\frac{1}{\ln 2} &< \frac{\log \frac{1}{1-\varepsilon}}{\varepsilon} \\
\frac{1}{\ln 2} &< \frac{-\log(1-\varepsilon)}{\varepsilon} \\
\frac{1}{\ln 2} &< \frac{-\ln(1-\varepsilon)}{\varepsilon \ln 2} \\
1 &< \frac{-\ln(1-\varepsilon)}{\varepsilon}
\end{aligned}$$

And indeed, by using the following Maclaurin series expansion for $|\varepsilon| < 1$:

$$\ln(1-\varepsilon) = -\sum_{i=1}^{\infty} \frac{\varepsilon^i}{i}$$

We conclude that:

$$\begin{aligned}\frac{-\ln(1-\varepsilon)}{\varepsilon} &= \frac{\varepsilon+\frac{\varepsilon^2}{2}+\frac{\varepsilon^3}{3}+\dots}{\varepsilon}\\ &= 1+\frac{\varepsilon}{2}+\frac{\varepsilon^2}{3}+\dots\\ &> 1\end{aligned}$$

□

Lemma A.2. *[11, Lemma 16.19] Let $n \geq 1$ and $0 < \delta \leq \frac{1}{2}$. Then*

$$\sum_{i=0}^{\lfloor \delta n \rfloor} \binom{n}{i} \leq 2^{H(\delta)n}$$

where $H(\delta) := -\delta \log \delta - (1-\delta)\log(1-\delta)$ is the binary entropy of δ.

Lemma A.3. *[19] Let m and n be integers with $m \geq 2$ and $n \geq 1$. Then*

$$\binom{mn}{n} \geq \frac{m^{m(n-1)+1}}{(m-1)^{(m-1)(n-1)}} n^{-\frac{1}{2}}$$

Lemma A.4. *Let Q and n be integers with $Q, n \geq 1$. Then*

$$\prod_{i=1}^{Q} \binom{in}{n} > \left(\frac{1}{e\sqrt{n}}\right)^{(Q-1)} \frac{\sqrt{2\pi Q}}{e} \cdot Q^{nQ}$$

Proof. The proof is basically a corollary of Lemma A.3, since:

$$\begin{aligned}\prod_{i=1}^{Q} \binom{in}{n} &= \binom{2n}{n}\binom{3n}{n}\cdots\binom{Qn}{n}\\ &\geq \left(2\cdot 2^{2(n-1)}\cdot\frac{1}{\sqrt{n}}\right)\left(3\cdot\frac{3^{3(n-1)}}{2^{2(n-1)}}\cdot\frac{1}{\sqrt{n}}\right)\cdots\left(Q\cdot\frac{Q^{Q(n-1)}}{(Q-1)^{(Q-1)(n-1)}}\cdot\frac{1}{\sqrt{n}}\right)\\ &= Q!\cdot Q^{Q(n-1)}\cdot\left(\frac{1}{\sqrt{n}}\right)^{(Q-1)}\end{aligned}$$

By Stirling's formula:

$$Q! > \sqrt{2\pi Q}\left(\frac{Q}{e}\right)^{Q}$$

Thus:

$$\begin{aligned}\prod_{i=1}^{Q} \binom{in}{n} &\geq Q!\cdot Q^{Q(n-1)}\cdot\left(\frac{1}{\sqrt{n}}\right)^{(Q-1)}\\ &> \sqrt{2\pi Q}\left(\frac{Q^n}{e}\right)^{Q}\cdot\left(\frac{1}{\sqrt{n}}\right)^{(Q-1)}\\ &= \left(\frac{1}{e\sqrt{n}}\right)^{(Q-1)}\frac{\sqrt{2\pi Q}}{e}\cdot Q^{nQ}\end{aligned}$$

□

B Concatenated coding scheme

We now consider a simple and natural concatenated coding scheme based on erasure codes suggested by [24]. Let $R < 1$ be a given value that will be specified later. Let $M_1 = M_2 = Q^{Rn}$. Let $E : [Q^{Rn}] \to [Q]^n$ be an erasure code that enables to reconstruct a codeword under $\alpha n = (1 - R - \delta)n$ erasures for some small $\delta > 0$ (that depends on Q and tends to 0 as Q tends to ∞). That is, for any $m \in M_i$, if αn locations in $E(m) = E(m)_1, \ldots, E(m)_n$ are erased then one can still reconstruct m. Such codes exist, e.g., in [6]. Let $\{\sigma_{i,j}\}_{i=1,j=1}^{i=2,j=n}$ be independent random permutations on $[Q]$ and define $E_i : [M_i] \to [Q]^n$ to satisfy $E_i(m) = \sigma_{i,1}(E(m)_1), \ldots, \sigma_{i,n}(E(m)_n)$. It holds that E_i are also α-erasure codes with rate $\log(Q^{Rn})/n = Rq$. The difference between E_i and the original E lies in the fact that now entry j of $E_i(m)$ is uniformly distributed due to the permutation $\sigma_{i,j}$.

A natural concatenation scheme uses E_1, E_2 to communicate over W. Namely, for a pair of messages $m_1 \in M_1, m_2 \in M_2$, we send $E_1(m_1), E_2(m_2)$ over network $W^{(n)}$ as depicted in Figure 2. The scheme will work if for every pair $m_1 \in M_1, m_2 \in M_2$ it holds that $W_1(E_1(m_1), E_2(m_2))$ (and consequently $W_2(E_1(m_1), E_2(m_2))$) have at most αn erasure values ϕ. We analyze the probability over $\{\sigma_{i,j}\}$ that indeed this is the case. Let A be the required event, namely that for every pair $m_1 \in M_1, m_2 \in M_2$ it holds that $W_1(E_1(m_1), E_2(m_2))$ has at most αn erasure values ϕ. Let B_{m_1,m_2} be the "bad" event, that for a specific $m_1 \in M_1, m_2 \in M_2$, the number of erasure values ϕ in $W_1(E_1(m_1), E_2(m_2))$ is more than αn. By the union bound:

$$\Pr[A] = 1 - \Pr\left[\bar{A}\right] \geq 1 - \sum_{m_1,m_2} \Pr[B_{m_1,m_2}]$$

Clearly,

$$\Pr[B_{m_1,m_2}] > \varepsilon^{\alpha n}$$

since $\Pr[B_{m_1,m_2}]$ is greater than the probability to get αn erasure symbols in the first αn coordinates (and this probability is equal to $\varepsilon^{\alpha n}$). Even if we were optimistic enough to assume $\Pr[B_{m_1,m_2}] = \varepsilon^{\alpha n}$, we would receive:

$$\Pr[A] \geq 1 - Q^{2nR}\varepsilon^{\alpha n}$$

If $1 - Q^{2nR}\varepsilon^{\alpha n} > 0$ then we could conclude the existence of permutations $\{\sigma_{i,j}\}$ as desired

(namely that will yield zero-error communication). But $1 - Q^{2nR}\varepsilon^{\alpha n} > 0$ only if:

$$\begin{aligned} Q^{2nR}\varepsilon^{\alpha n} &< 1 \\ \varepsilon^{\alpha n} &< \frac{1}{Q^{2nR}} \\ \varepsilon &< \left(\frac{1}{Q^{2nR}}\right)^{\frac{1}{\alpha n}} \\ \varepsilon &< \left(\frac{1}{Q^{2R/\alpha}}\right) \end{aligned}$$

Recall that $\alpha = 1 - R - \delta$. So we have for any $R \geq \frac{1}{2}$ that $1 - Q^{2nR}\varepsilon^{\alpha n} > 0$ implies $\varepsilon < \frac{1}{Q^{2+4\delta}}$ (since in this case $2R/\alpha \geq 2R/(1/2 - \delta) \geq 2/(1 - 2\delta)$; and for $\delta > 0$ it holds that $2 + 4\delta < \frac{2}{1-2\delta}$). Thus, to prove rate greater than q using the suggested concatenation scheme (with our analysis via the union bound), we need Q to be *small* compared to $1/\varepsilon$. Notice that for such values of Q and ε there are very few inputs mapped to erasure symbols in a typical $W \in \mathcal{W}_{Q,\varepsilon}$, and thus one would indeed expect a large zero-error rate. Namely, this analysis of the natural concatenated scheme will not succeed in obtaining a sum rate larger than q when $Q \geq \Omega(1/\varepsilon)$, which we view as some support in favor of Statement 3.1.

References

[1] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, *Network Information Flow*, IEEE Transactions on Information Theory **46** (2000), no. 4, 1204–1216.

[2] R. Ahlswede and D. E. Daykin, *An inequality for the weights of two families of sets, their unions and intersections*, Probability Theory and Related Fields **43** (1978), 183–185.

[3] R. Ahlswede and G. Simonyi, *On the optimal structure of recovering set pairs and the sandglass conjecture*, Discrete Mathematics **128** (1994), 389–394.

[4] N. Alon and J. H. Spencer, *The probabilistic method*, third ed., Wiley-Interscience, 2008.

[5] T. Chan and A. Grant, *On capacity regions of non-multicast networks*, In proceedings of IEEE International Symposium on Information Theory, 2010, pp. 2378–2382.

[6] T. M. Cover and J. A. Thomas, *Elements of information theory*, second ed., ch. 15.4, Wiley-Interscience, 2006.

[7] M. Effros, S. El Rouayheb, and M. Langberg, *An equivalence between network coding and index coding*, In proceedings of IEEE International Symposium on Information Theory (ISIT), 2013, pp. 967–971.

[8] K. Engel, *Sperner theory*, Encyclopedia of mathematics and its applications, Cambridge University Press, 1997.

[9] O. Favaron, P. Mago, and O. Ordaz, *On the bipartite independence number of a balanced bipartite graph*, Discrete Mathematics **121** (1993), no. 13, 55–63.

[10] U. Feige and S. Kogan, *Balanced coloring of bipartite graphs*, Journal of Graph Theory **64** (2010), no. 4, 277–291.

[11] J. Flum and M. Grohe, *Parameterized complexity theory*, Springer-Verlag Berlin / Heidelberg, 2006.

[12] T. Ho, M. Effros, and S. Jalali, *On equivalences between network topologies*, In proceedings of the 48th Annual Allerton Conference on Communication, Control, and Computing (Allerton), 2010, pp. 391–398.

[13] T. Ho, M. Médard, R. Koetter, D. R. Karger, M. Effros, J. Shi, and B. Leong, *A Random Linear Network Coding Approach to Multicast*, IEEE Transactions on Information Theory **52** (2006), no. 10, 4413–4430.

[14] R. Holzman and J. Korner, *Cancellative pairs of families of sets*, European Journal of Combinatorics **16** (1995), 263–266.

[15] S. Jaggi, P. Sanders, P. A. Chou, M. Effros, S. Egner, K. Jain, and L. Tolhuizen, *Polynomial Time Algorithms for Multicast Network Code Construction*, IEEE Transactions on Information Theory **51** (2005), no. 6, 1973–1982.

[16] S. Jalali, M. Effros, and T. Ho, *On the impact of a single edge on the network coding capacity*, In proceedings of Information Theory and Applications Workshop (ITA), 2011, pp. 1–5.

[17] R. Koetter and M. Medard, *An Algebraic Approach to Network Coding*, IEEE/ACM Transactions on Networking **11** (2003), no. 5, 782–795.

[18] J. Körner and A. Orlitsky, *Zero-error information theory*, IEEE Transactions on Information Theory **44** (1998), no. 6, 2207–2229.

[19] O. Krafft, *Problem 10819*, The American Mathematical Monthly **107** (2000), no. 7, 652.

[20] M. Langberg and M. Effros, *Network coding: Is zero error always possible?*, In proceedings of the 49th Annual Allerton Conference on Communication, Control, and Computing (Allerton), 2011, pp. 1478–1485.

[21] M. Langberg and M. Effros, *Source coding for dependent sources*, In proceedings of IEEE Information Theory Workshop (ITW), 2012, pp. 70–74.

[22] S.-Y. R. Li, R. W. Yeung, and N. Cai, *Linear Network Coding*, IEEE Transactions on Information Theory **49** (2003), no. 2, 371–381.

[23] M. Mitzenmacher and E. Upfal, *Probability and computing: Randomized algorithms and probabilistic analysis*, Cambridge University Press, 2005.

[24] C. Nair, private communication.

[25] J. Sgall, *Bounds on pairs of families with restricted intersections*, Combinatorica **19** (1999), 555–566.

[26] G. Simonyi, *On write uni-directional memory codes*, IEEE Transactions on Information Theory **35** (1989), 663–669.

[27] D. Slepian and J. K. Wolf, *Noiseless coding of correlated information sources*, IEEE Transactions on Information Theory **19** (1973), 471–480.

תקציר

מחקרים סטנדרטיים של ערוצי הפרעה מרובי מקורות ומרובי יעדים מאפשרים בדרך כלל הסתברות הולכת ונעלמת של שגיאה בתקשורת. תזה זו עוסקת בכימות אובדן קצב התקשורת כאשר נדרשים לתקשורת עם אפס שגיאה בהקשר של ערוצי הפרעה. המוטיבציה לעבודה זו נובעת בין היתר מבעיה פתוחה דומה בקידוד רשתות.

חלקים מעבודה זו התפרסמו ב:

I. Levi, D. Vilenchik, M. Langberg and M. Effros.
Zero vs. epsilon error in interference channels. In proceedings of IEEE Information Theory Workshop (ITW), 2013, p. 1-5.

תוכן עניינים

האוניברסיטה הפתוחה
המחלקה למתמטיקה ולמדעי המחשב

אפס מול ε שגיאה בערוצי הפרעה

עבודת תזה זו הוגשה כחלק מהדרישות לקבלת תואר
"מוסמך למדעים" .M.Sc במדעי המחשב
באוניברסיטה הפתוחה
החטיבה למדעי המחשב

על-ידי
איליה לוי

העבודה הוכנה בהדרכתו של פרופ' מיכאל לנגברג

ספטמבר 2014