A VC-dimension-based Outer Bound on the Zero-Error Capacity of the Binary Adder Channel

Or Ordentlich
Tel Aviv University
ordent@eng.tau.ac.il

Ofer Shayevitz
Tel Aviv University
ofersha@eng.tau.ac.il

***Abstract*—The binary adder is a two-user multiple access channel whose inputs are binary and whose output is the real sum of the inputs. While the Shannon capacity region of this channel is well known, little is known regarding its zero-error capacity region, and a large gap remains between the best inner and outer bounds. In this paper, we provide an improved outer bound for this problem. To that end, we introduce a soft variation of the Saur-Perles-Shelah Lemma, that is then used in conjunction with an outer bound for the Shannon capacity region with an additional common message.**

I. Introduction

The binary adder is a multiple access channel with two binary inputs X_1 and X_2 and output $Y = X_1 + X_2 \in \{0, 1, 2\}$. The capacity region of this channel is well known and consists of all rate-pairs (R_1, R_2) satisfying

$$\begin{aligned} R_1 &\leq 1, \\ R_2 &\leq 1, \\ R_1 + R_2 &\leq \tfrac{3}{2}. \end{aligned} \tag{1}$$

The zero-error capacity region of the binary adder channel is the closure of the set of all rate-pairs (R_1, R_2) such that for n large enough there exist two codebooks $\mathcal{C}_1, \mathcal{C}_2 \subseteq \{0, 1\}^n$ with cardinalities $|\mathcal{C}_i| = 2^{nR_i}$, $i = 1, 2$, such that all elements in the *sumset*

$$\mathcal{C}_1 + \mathcal{C}_2 \triangleq \{\mathbf{a} + \mathbf{b} : \mathbf{a} \in \mathcal{C}_1, \mathbf{b} \in \mathcal{C}_2\} \quad \text{with multiplicities} \tag{2}$$

appear with multiplicity exactly one, where addition is taken over the reals. We say that the pair (R_1, R_2) is *admissible* if it belongs to the zero-error capacity region, and we call the codebooks $(\mathcal{C}_1, \mathcal{C}_2)$ a *zero-error codebook pair* if all elements in their sumset $\mathcal{C}_1 + \mathcal{C}_2$ appear with multiplicity exactly one.

Despite its apparent simplicity, the problem of characterizing the zero-error capacity region of this channel is wide open. Many inner bounds have been established over the last four decades, see, e.g., [1]–[10]. However, to date, the best known lower bound on the zero error sum-capacity is $\log(240/6) \approx 1.3178$ [10], where logarithms are taken in base 2. To put this result in perspective, note that a sum-rate of $R_1 + R_2 = \frac{1}{2}\log(6) \approx 1.2924$ can be attained by the two-dimensional construction $\mathcal{C}_1 = \{00, 11\}$, $\mathcal{C}_2 = \{00, 01, 10\}$. In terms of outer bounds, the current state of knowledge is even less satisfying. Clearly, any admissible pair must be inside the Shannon capacity region and must therefore satisfy (1). However, to date the only improvement upon the trivial outer bound (1) was obtained by Urbanke and Li [8] who showed that near the corner points $(1, \frac{1}{2})$ and $(\frac{1}{2}, 1)$ the zero-error capacity region is strictly contained in (1). Specifically, for $R_1 = 1$ it was shown that the maximal admissible R_2 must satisfy $R_2 < 0.49216$. Our main result is a new outer bound on the zero-error capacity region that strictly improves upon the bound from [8], using different techniques.

Write $h(p) = -p \log p - (1-p) \log(1-p)$ for the binary entropy function, and $h^{-1}(x)$ for its inverse restricted to $[0, \frac{1}{2}]$. For $0 \leq p, q \leq 1$, write $p \star q \triangleq p(1-q) + q(1-p)$. Let

$$L(\eta) \triangleq h(\eta) + 1 - \eta \tag{3}$$

and

$$J(p, \eta) \triangleq \begin{cases} 2h\left(\frac{1}{2}\left(1 - \sqrt{1 - 2\eta}\right)\right) - \eta & \eta \geq p \star p \\ 2h\left(\frac{1}{2}\left(1 - \frac{1-\eta-p\star p}{\sqrt{1-2(p\star p)}}\right)\right) & \\ \quad -\frac{1}{2}\left(1 - \frac{(1-\eta-p\star p)^2}{1-2(p\star p)}\right) & \eta < p \star p \end{cases} \tag{4}$$

and

$$R_\Sigma(r_0, r_1) \triangleq \max_{h^{-1}(r_1) \leq \eta \leq \frac{1}{2}} \min\{L(\eta),\ J(h^{-1}(r_1), \eta) + r_0\} \tag{5}$$

Our main result is the following.

Theorem 1: Any admissible (R_1, R_2) satisfies

$$R_2 < \min_{0 \leq \alpha \leq h^{-1}(R_1)} (1 - \alpha) \left(R_\Sigma\left(\frac{\alpha}{1-\alpha}, \Gamma\right) - \Gamma\right)$$

where

$$\Gamma = \Gamma(R_1, \alpha) \triangleq h\left(\frac{h^{-1}(R_1) - \alpha}{1 - \alpha}\right)$$

The work of O. Ordentlich was supported by the Admas Fellowship Program of the Israel Academy of Science and Humanities, a fellowship from The Yitzhak and Chaya Weinstein Research Institute for Signal Processing at Tel Aviv University, and the Feder Family Award. The work of O. Shayevitz has received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement no. 639573, the Marie Curie Career Integration Grant (CIG) under grant agreement no. 631983, and the Israel Science Foundation under grant agreement no. 1367/14.

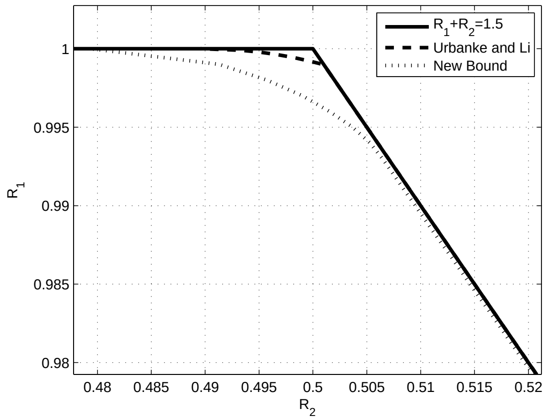

Fig. 1. Illustration of the three outer bounds.

For the maximal value of $R_1 = 1$, this bound yields $R_2 < 0.4794$. Figure 1 depicts the three outer bounds for values of R_1 close to 1. The question of whether $R_1 + R_2 = \frac{3}{2}$ is admissible for some (R_1, R_2) remains open.

II. Proof of Theorem 1

We first note that it suffices to prove inadmissibility in the limit of large n, by the simple fact that if $(\mathcal{C}_1, \mathcal{C}_2)$ is a zero-error codebook pair, so is the concatenation $(\mathcal{C}_1 \times \mathcal{C}_1, \mathcal{C}_2 \times \mathcal{C}_2)$. To avoid cumbersome notations, we can therefore assume without loss of generality that nR_1 and nR_2 (and all similar quantities) are integers.

A. Motivation

Let $\mathcal{C} \subseteq \{0,1\}^n$ be a codebook and let $S \subseteq [n]$ be a subset of coordinates, where $[n] \triangleq \{1, \ldots, n\}$. The projection $\mathbf{a}(S)$ maps the vector $\mathbf{a} \in \{0,1\}^n$ to a vector in $\{0,1\}^{|S|}$ by taking only the values of $\mathbf{a}$ on the coordinates in S. We say that S is *shattered* by $\mathcal{C}$ [11], if the projection multiset

$$P_S^+(\mathcal{C}) \triangleq \{\mathbf{c}(S) : \mathbf{c} \in \mathcal{C}\} \quad \text{with multiplicities}$$

of $\mathcal{C}$ on S contains all $2^{|S|}$ binary vectors of length $|S|$.[1] A codebook $\mathcal{C}$ is said to be *systematic* if it shatters some $S \subseteq [n]$ of cardinality $\log |\mathcal{C}|$. Weldon proved the following.

Theorem 2 (Weldon [4]): If $\mathcal{C}_1$ is systematic and $(\mathcal{C}_1, \mathcal{C}_2)$ form a zero-error codebook pair, then $R_2 \leq (1 - R_1) \log 3$.

Proof. Let S be a set of cardinality nR_1 that is shattered by $\mathcal{C}_1$. For every $\mathbf{c}_2 \in \mathcal{C}_2$, there exists a $\mathbf{c}_1 \in \mathcal{C}_1$ such that $\mathbf{c}_1$ and $\mathbf{c}_2$ are an S*-complement* pair, i.e.,

$$\mathbf{c}_1(S) + \mathbf{c}_2(S) = \mathbf{1}_{|S|}, \tag{6}$$

where $\mathbf{1}_m$ denotes a vector of 1s of length m. Hence, there are at least 2^{nR_2} such S-complement pairs. By the assumption that $(\mathcal{C}_1, \mathcal{C}_2)$ form a zero-error codebook pair, $\mathbf{c}_1(\overline{S}) + \mathbf{c}_2(\overline{S})$ must be distinct for all S-complement pairs. Therefore, the number of such pairs cannot be larger than $3^{|\overline{S}|} = 3^{n(1-R_1)}$, and the theorem follows. ■

For example, if $\mathcal{C}_1$ is systematic and $R_2 = 1$, then the theorem implies that $R_1 \leq 0.37$. This strong bound is a consequence of the restriction to a systematic codebook. However, we note that the only property used in the proof is the existence of a large shattered set. Hence, any lower bound on the size of a maximal shattered set in a general codebook $\mathcal{C}_1$ would lead to a similar result. The cardinality of the maximal set shattered by a code $\mathcal{C} \subseteq \{0,1\}^n$ is referred to in the machine-learning literature as its *Vapnik-Chervonenkis dimension*, or *VC-dimension*. The Sauer-Perles-Shelah lemma provides a lower bound on the VC-dimension of a code.

Lemma 1 (Sauer-Perles-Shelah Lemma [11]): If the cardinality of the maximal subset shattered by the codebook $\mathcal{C} \subseteq \{0,1\}^n$ is d, then

$$|\mathcal{C}| \leq \sum_{k=0}^{d} \binom{n}{k}.$$

Remark 1: It is easy to see that this bound is attained with equality if $\mathcal{C}$ is a n-Hamming ball of radius d.

Corollary 1: Let $\varepsilon > 0$. If $|\mathcal{C}| = 2^{n(R+\varepsilon)}$ then for any n large enough, $\mathcal{C}$ shatters a set $S \subseteq [n]$ with $|S| \geq nh^{-1}(R)$.

Plugging the above into Weldon's argument yields:

Proposition 1: If $(\mathcal{C}_1, \mathcal{C}_2)$ form a zero-error codebook pair, then $R_2 \leq (1 - h^{-1}(R_1)) \log 3$.

Unfortunately, this bound is trivial since for any R_1, we have that $R_1 + (1 - h^{-1}(R_1)) \log 3 > \frac{3}{2}$. This stems from two main weaknesses. First, we have taken the worst case assumption that each codeword $\mathbf{c}_2 \in \mathcal{C}_2$ has only one codeword $\mathbf{c}_1 \in \mathcal{C}_1$ such that $\mathbf{c}_1$ and $\mathbf{c}_2$ are S-complement, where S is a shattered set in $\mathcal{C}_1$. Second, bounding the number of S-complement pairs by $3^{|\overline{S}|}$ may be loose, as it ignores the sumset structure. In the next two subsections, we provide the technical tools to handle each of these weaknesses, and apply them to prove the theorem in the subsection that follows.

B. A Soft Sauer-Perles-Shelah Lemma

Let $\mathcal{C} \subseteq \{0,1\}^n$ be a codebook and let $S \subseteq [n]$ be a subset of coordinates. We say that S is k*-shattered* by $\mathcal{C}$, if the projection multiset $P_S^+(\mathcal{C})$ of $\mathcal{C}$ on S contains all binary vectors in $\{0,1\}^{|S|}$ each with multiplicity of at least k. For $k = 1$, this definition reduces to the regular definition of a shattered set.

The proof of the following lemma is given in Section III.

Lemma 2: If the cardinality of the maximal subset that is k-shattered by the codebook $\mathcal{C} \subseteq \{0,1\}^n$ is $d - 1$, then

$$|\mathcal{C}| \leq \sum_{t=1}^{t^*} \binom{n}{t} + \binom{n}{t^*} \sum_{t=t^*+1}^{n} \frac{\binom{t^*}{d}}{\binom{t}{d}}$$

[1] Taking the multiplicities into account in the definition of the projection multiset is not necessary here, but will become important in the sequel.

where t^* is the smallest integer t satisfying $\binom{n-d}{t-d} \geq k$ if such an integer exists, and $t^* = n$ otherwise.

Remark 2: Note that if $k = \binom{n-d}{t^*-d}$ for some t^*, then our bound is tight for a n-Hamming ball of radius t^*, up to a multiplicative gap of $O(n/d)$. This coincides with the Sauer-Perles-Shelah Lemma for $k = 1$ (and $t^* = d$), up to the aforementioned multiplicative factor. Since we are only interested in exponential behavior, no attempt has been made to reduce this gap.

Corollary 2: Let $\varepsilon > 0$. If $|\mathcal{C}| = 2^{n(R+\varepsilon)}$ then for any $0 \leq \alpha \leq h^{-1}(R)$ and any n large enough, there exists a set $S \subseteq [n]$ with $|S| \geq n\alpha$ that is $2^{n\beta}$-shattered by $\mathcal{C}$, where

$$\beta = (1-\alpha) \cdot h\left(\frac{h^{-1}(R) - \alpha}{1-\alpha}\right) \tag{7}$$

Proof. Let $0 \leq \alpha \leq h^{-1}(R)$ and assume to the contrary that no subset of size $d = n\alpha$ is $2^{n\beta}$-shattered by $\mathcal{C}$. Denote $t^* = \gamma_n n$, and write

$$\begin{aligned}\frac{1}{n}\log\binom{n-d}{t^*-d} &= \frac{n-d}{n}\left(h\left(\frac{t^*-d}{n-d}\right) + o(1)\right)\\ &= (1-\alpha+o(1))h\left(\frac{\gamma_n - \alpha}{1-\alpha}\right)\end{aligned}$$

We can set γ_n to the minimal value guaranteeing that the above is at least β, which is $\gamma_n = \alpha + (1-\alpha)h^{-1}\left(\frac{\beta}{1-\alpha}\right) + o(1)$. Invoking Lemma 2, it must then be that $|\mathcal{C}| \leq 2^{n(h(\gamma_n)+o(1))} = 2^{n(R+o(1))}$, contradicting the assumption. ■

C. The Binary Adder Channel with an Additional Common Message

In the Weldon-type arguments mentioned above, the number of S-complement pairs was bounded by $3^{|\overline{S}|}$, thereby ignoring the sumset structure. As we shall see in the next subsection, this structure can be accounted for by partitioning each codebook according to its projection on S, which naturally gives rise to a zero-error communication problem with an additional common message of rate at most $|S|/|\overline{S}|$. Upper bounding the corresponding admissible sum-rate in this new setup can in turn be translated into an upper bound on the number of S-complement pairs in our original setup.

More precisely, assume that there are three messages $W_i \in [2^{nr_i}]$, $i = 0, 1, 2$, to be conveyed to the receiver over the binary adder channel, where the first user has access to the messages (W_0, W_1) and the second user has access to the messages (W_0, W_2). The Shannon capacity region for this problem was found by Slepian and Wolf [12] to be the set of all rate triplets satisfying

$$\begin{aligned} r_1 &\leq H(X_1|U),\\ r_2 &\leq H(X_2|U),\\ r_1 + r_2 &\leq H(X_1+X_2|U),\\ r_0 + r_1 + r_2 &\leq H(X_1+X_2)\end{aligned} \tag{8}$$

for some $P_{U,X_1,X_2} = P_U P_{X_1|U} P_{X_2|U}$, where X_1 and X_2 are binary random variables and the random variable U has a finite support.

A coding scheme for this problem consists of a *system* $\mathcal{V}$, which is a set of codebook pairs $\{\mathcal{C}_{1,i}, \mathcal{C}_{2,i}\}_{i=1}^{M_0}$, where each $\mathcal{C}_{1,i}$ (resp. $\mathcal{C}_{2,i}$) is a codebook in $\{0,1\}^n$ with fixed cardinality $|\mathcal{C}_{1,i}| = M_1$ (resp. $|\mathcal{C}_{2,i}| = M_2$). We say that $\mathcal{V}$ is a *zero-error system* if each pair $(\mathcal{C}_{1,i}, \mathcal{C}_{2,i})$ is a zero-error codebook pair, and the sumsets $\mathcal{C}_{1,i} + \mathcal{C}_{2,i}$ are mutually disjoint. A triplet (r_0, r_1, r_2) is called admissible if there exists a zero-error system $\mathcal{V}$ with $M_\ell = 2^{n(r_\ell + o(1))}$ for $\ell \in \{0, 1, 2\}$.

Clearly, any admissible triplet must satisfy (8). The bounds we obtain in this subsection are based on outer bounding this latter region. More specifically, as will become clear in the next subsection, our goal is to upper bound the maximal sum of admissible rates $r_0 + r_1 + r_2$ as a function of r_0 and r_1. Although the bounds in (8) are given in a single-letter form, in order to guarantee the inadmissibility of a rate triplet, one must go over all valid distributions P_{U,X_1,X_2}. While it is not difficult to show that for our needs there is no loss of generality in considering only random variables U with cardinality no greater than 3, the number of remaining parameters makes the evaluation of (8) within a satisfactory resolution infeasible for a brute-force grid search. Instead, the following lemma provides an analytic upper bound on the sum-capacity as a function of r_0 and r_1, in terms of the solution to a single-parameter optimization problem. The proof is omitted due to space limitations, but can be found in the full version of this paper [13].

Lemma 3: Let $L(\eta)$ and $J(p, \eta)$ be as defined in (3) and (4). If (r_0, r_1, r_2) is admissible, then

$$r_0 + r_1 + r_2 \leq \max_{h^{-1}(r_1) \leq \eta \leq \frac{1}{2}} \min\{L(\eta),\ J(h^{-1}(r_1), \eta) + r_0\}$$

Remark 3: Note that it can be shown that the maximization can be further restricted to $h^{-1}(r_1) \star h^{-1}(r_2) \leq \eta \leq \frac{1}{2}$. This however is not useful for our purposes.

The following lemma is not necessary for the proof of Theorem 1, but may be of independent interest.

Lemma 4: The maximal sum of achievable rates (for a vanishing error probability) over the binary adder channel with an additional common message, as a function of the rate of the common message rate r_0, is

$$\begin{aligned} r_0 + r_1 + r_2 = \max_{0 \leq \eta \leq \frac{1}{2}} \min\{&h(\eta) + 1 - \eta,\\ &2h\left(\tfrac{1}{2}\left(1 - \sqrt{1-2\eta}\right)\right) - \eta + r_0\}\end{aligned} \tag{9}$$

Proof sketch. The upper bound on $r_0 + r_1 + r_2$ follows as a corollary of Lemma 3, by noting that for any $0 \leq r_1 \leq 1$ we have $J(h^{-1}(r_1), \eta) \leq 2h\left(\frac{1}{2}\left(1 - \sqrt{1-2\eta}\right)\right) - \eta$. To see that the right hand side of (9) is achievable, let η^* be the maximizer of (9) and evaluate the entropies in (8) with the following distribution:

$$\begin{aligned} &X_1 = U \oplus Z_1,\ X_2 = U \oplus Z_2 \\ &U \sim \text{Bern}\left(\frac{1}{2}\right),\ Z_1 \sim \text{Bern}(p^*),\ Z_2 \sim \text{Bern}(p^*) \end{aligned} \tag{10}$$

where U, Z_1, Z_2 are mutually independent, and $p^* \le \frac{1}{2}$ satisfies $p^* \star p^* = \eta^*$, i.e., $p^* = \frac{1}{2}(1 - \sqrt{1 - 2\eta^*})$. ■

D. Putting it Together

We are now in a position to prove Theorem 1. Let $(\mathcal{C}_1, \mathcal{C}_2)$ be a zero-error codebook pair of cardinalities 2^{nR_1} and 2^{nR_2} respectively. Given this pair, we use Corollary 2 to construct a zero-error system with certain cardinalities, and then apply Lemma 3 to obtain constraints on that system.

By Corollary 2, for any $\alpha < h^{-1}(R_1)$ there exists a subset of coordinates $S \subset [n]$ of cardinality $n\alpha$ that is $2^{n\beta}$-shattered by $\mathcal{C}_1$, where β is given in (7), all up to an $o(1)$ term. Let $\mathcal{C}_0$ be the family of all binary vectors of length $|S|$, and for any $\mathbf{g} \in \mathcal{C}_0$ let $\mathcal{C}_{1,\mathbf{g}} = \{\mathbf{c} \in \mathcal{C}_1 : \mathbf{c}(S) = \mathbf{g}\}$. Define $\mathcal{C}_{2,\mathbf{g}}$ similarly, and note that $\{\mathcal{C}_{j,\mathbf{g}}\}_{\mathbf{g}\in\mathcal{C}_0}$ is a partition of $\mathcal{C}_j$ for each $j \in \{1, 2\}$.

By construction, $|\mathcal{C}_{1,\mathbf{g}}| \ge 2^{n\beta}$. We can therefore arbitrarily choose $\widetilde{\mathcal{C}}_{1,\mathbf{g}} \subseteq \mathcal{C}_{1,\mathbf{g}}$ such that $|\widetilde{\mathcal{C}}_{1,\mathbf{g}}| = 2^{n\beta}$. For each $\mathbf{g}$ with $|\mathcal{C}_{2,\mathbf{g}}| > 0$, arbitrarily choose $\widetilde{\mathcal{C}}_{2,\mathbf{g}} \subseteq \mathcal{C}_{2,\mathbf{g}}$ such that $\log|\widetilde{\mathcal{C}}_{2,\mathbf{g}}| = \lfloor \log|\mathcal{C}_{2,\mathbf{g}}| \rfloor$. Note that this guarantees that $|\widetilde{\mathcal{C}}_{2,\mathbf{g}}| = 2^k$ for some integer $0 \le k \le nR_2$, and that $|\widetilde{\mathcal{C}}_{2,\mathbf{g}}| \ge |\mathcal{C}_{2,\mathbf{g}}|/2$. Moreover, there must exist an integer k' with the property that the union of all $\widetilde{\mathcal{C}}_{2,\mathbf{g}}$ of cardinality $2^{k'}$ contains at least $\frac{1}{2(nR_2+1)}2^{nR_2}$ vectors. Let $\mathcal{G}$ be the set of all $\mathbf{g} \in \mathcal{C}_0$ that correspond to this k', and note that by construction $|\mathcal{G}| = 2^{n\alpha'}$ for some $\alpha' \le \alpha$. Moreover,

$$|\widetilde{\mathcal{C}}_{2,\mathbf{g}}| = 2^{k'} \ge \tfrac{1}{2(nR_2+1)} 2^{n(R_2 - \alpha')}$$

for all $\mathbf{g} \in \mathcal{G}$.

Let $\overline{\mathbf{g}} = \mathbf{g} \oplus \mathbf{1}_{|S|}$ be the binary complement of $\mathbf{g}$, and define the system $\mathcal{V} = \{(\widetilde{\mathcal{C}}_{1,\overline{\mathbf{g}}}, \widetilde{\mathcal{C}}_{2,\mathbf{g}})\}_{\mathbf{g}\in\mathcal{G}}$. Since the original $\mathcal{C}_1$ and $\mathcal{C}_2$ form a zero-error codebook pair, then $\mathcal{V}$ is trivially a zero-error system. Moreover, since any $\mathbf{c}_1 \in \widetilde{\mathcal{C}}_{1,\overline{\mathbf{g}}}$ and $\mathbf{c}_2 \in \widetilde{\mathcal{C}}_{2,\mathbf{g}}$ are an S-complement pair (6), the projection

$$\mathcal{V}_{\overline{S}} \triangleq \{(P^+_{\overline{S}}(\widetilde{\mathcal{C}}_{1,\overline{\mathbf{g}}}), P^+_{\overline{S}}(\widetilde{\mathcal{C}}_{2,\mathbf{g}}))\}_{\mathbf{g}\in\mathcal{G}}$$

of $\mathcal{V}$ onto $\overline{S}$ is also a zero-error system, over $|\overline{S}| = n(1-\alpha)$ coordinates.

We have thus shown that given a zero-error codebook pair over n coordinates with cardinalities 2^{nR_1} and 2^{nR_2}, we can construct a zero-error system $\mathcal{V}_{\overline{S}}$ over $m = n(1-\alpha)$ coordinates with cardinalities $M_0 = 2^{mr_0}$, $M_1 = 2^{mr_1}$ and $M_2 = 2^{m(r_2 + o(1))}$, where

$$r_0 = \frac{\alpha'}{1-\alpha}, \quad r_1 = \frac{\beta}{1-\alpha}, \quad r_2 = \frac{R_2 - \alpha'}{1-\alpha}$$

Thus for this system $r_0 + r_1 + r_2 = \frac{R_2+\beta}{1-\alpha}$, and by Lemma 3, recalling that $\alpha' \le \alpha$, we have that

$$\frac{R_2+\beta}{1-\alpha} \le \max_{h^{-1}\left(\frac{\beta}{1-\alpha}\right) \le \eta \le \frac{1}{2}} \min\left\{ L(\eta),\ J\left(h^{-1}\left(\frac{\beta}{1-\alpha}\right), \eta\right) + \frac{\alpha}{1-\alpha} \right\}$$

The theorem now follows by substituting β from Corollary 2, and noting that the above inequality holds for any $0 \le \alpha \le h^{-1}(R_1)$.

III. Proof of Lemma 2

For the purpose of the proof, it will be convenient to represent any binary vector $\mathbf{c} \in \{0,1\}^n$ by a subset of $F \subseteq [n]$ that contains the indices of the coordinates where $\mathbf{c}$ equals 1. Accordingly, any codebook $\mathcal{C} \subseteq \{0,1\}^n$ can be represented by the corresponding family $\mathcal{F}$ of subsets of $[n]$. Similarly, the multiset projection $P^+_S(\mathcal{C})$ of $\mathcal{C}$ on S corresponds to

$$P^+_S(\mathcal{F}) \triangleq \{F \cap S : F \in \mathcal{F}\} \qquad \text{with multiplicities}$$

and S is k-shattered by $\mathcal{C}$ (equivalently by $\mathcal{F}$) means that $P^+_S(\mathcal{F})$ contains each subset of S with multiplicity at least k.

Let $\mathcal{C}$ be a codebook and let $\mathcal{F}$ be the corresponding family of subsets on $[n]$. We start by applying the shifting argument introduced in [14] on $\mathcal{F}$, to construct another family $\mathcal{G}$ of the same cardinality, such that if S is k-shattered by $\mathcal{G}$ then it is also k-shattered by $\mathcal{F}$. Furthermore, $\mathcal{G}$ will be *monotone*, i.e., will have the property that if $G \in \mathcal{G}$ then all subsets of G are in $\mathcal{G}$.

Set $\mathcal{G} = \mathcal{F}$. If $\mathcal{G}$ is already monotone, we are done. Otherwise there exists some $i \in [n]$ such that the set

$$\widetilde{\mathcal{G}}_i \triangleq \{G \in \mathcal{G} : i \in G,\ G \setminus \{i\} \notin \mathcal{G}\}$$

is not empty. Update $\mathcal{G}$ according to the rule:

$$\mathcal{G} \leftarrow \left(\mathcal{G} \setminus \widetilde{\mathcal{G}}_i\right) \cup \left(\widetilde{\mathcal{G}}_i - i\right) \tag{11}$$

where $\widetilde{\mathcal{G}}_i - i$ is the family of subsets obtained from $\widetilde{\mathcal{G}}_i$ by removing the element i from each subset. The process continues until $\mathcal{G}$ is monotone, and is clearly guaranteed to terminate in finite time. By construction, $|\mathcal{G}| = |\mathcal{F}|$.

We now show that if S is k-shattered by $\mathcal{G}$ then it is also k-shattered by $\mathcal{F}$. Let $\mathcal{G}'$ be the family of subsets before the operation (11) on some element i, and let $\mathcal{G}$ be the family obtained after that operation. Suppose S is k-shattered by $\mathcal{G}$. It now suffices to show that S is also k-shattered by $\mathcal{G}'$. If $i \notin S$ then clearly $P^+_S(\mathcal{G}) = P^+_S(\mathcal{G}')$, hence this does not affect the k-shatterdness of S. Suppose $i \in S$, and let

$$\mathcal{G}_i \triangleq \{G \in \mathcal{G} : i \in G\}.$$

Then $\mathcal{G}_i \subseteq \mathcal{G}'$ since the update rule (11) does not add elements to subsets. Since $\mathcal{G}$ k-shatters S, then every subset of S that contains i has multiplicity at least k in $P^+_S(\mathcal{G}_i) \subseteq P^+_S(\mathcal{G}')$.

Recalling that $\mathcal{G}_i \subseteq \mathcal{G} \cap \mathcal{G}'$, we have that $\mathcal{G}_i - i \subseteq \mathcal{G}'$ since otherwise some replacement would have occurred in (11). Since $\mathcal{G}$ k-shatters S, then every subset of S that does not contain i has multiplicity at least k in $P_S^+(\mathcal{G}_i - i) \subseteq P_S^+(\mathcal{G}')$.

The Lemma now follows directly from the next proposition.

Proposition 2: If $\mathcal{G}$ is a monotone family of subsets of $[n]$ with the property that no subset of cardinality d is k-shattered by $\mathcal{G}$, then

$$|\mathcal{G}| \leq \sum_{t=1}^{t^*} \binom{n}{t} + \binom{n}{t^*} \sum_{t=t^*+1}^{n} \frac{\binom{t^*}{d}}{\binom{t}{d}}$$

where t^* is the smallest integer t satisfying $\binom{n-d}{t-d} \geq k$ if such an integer exists, and $t^* = n$ otherwise.

Proof. Let $\mathcal{G}_t$ denote the family of all subsets in $\mathcal{G}$ with cardinality t. For $t \geq d$, every $G \in \mathcal{G}_t$ has exactly $\binom{t}{d}$ subsets of cardinality d. There is a total of $\binom{n}{d}$ subsets of cardinality d. Hence by a simple counting argument there must exist at least one subset S of cardinality d, that is a subset of no less than $|\mathcal{G}_t|\binom{t}{d}/\binom{n}{d}$ subsets in $\mathcal{G}_t$. Recalling that $\mathcal{G}$ is monotone, this implies that S is $|\mathcal{G}_t|\binom{t}{d}/\binom{n}{d}$-shattered by $\mathcal{G}$. By our assumption, it must be that

$$\frac{\binom{t}{d}|\mathcal{G}_t|}{\binom{n}{d}} < k, \quad t = d, \ldots, n$$

On the other hand, $|\mathcal{G}_t| \leq \binom{n}{t}$, and therefore

$$|\mathcal{G}_t| \leq \min\left\{\binom{n}{t}, \frac{\binom{n}{d}k}{\binom{t}{d}}\right\}, \quad t = d, \ldots, n$$

Summing over t we get

$$|\mathcal{G}| = \sum_{t=1}^{n} |\mathcal{G}_t| \leq \sum_{t=1}^{d-1} \binom{n}{t} + \sum_{t=d}^{n} \min\left\{\binom{n}{t}, \frac{\binom{n}{d}k}{\binom{t}{d}}\right\} \tag{12}$$

Let t^* be the smallest integer t such that $\binom{n}{t} \geq \frac{\binom{n}{d}k}{\binom{t}{d}}$ if such an integer exists. If no such integer t exists, set $t^* = n$. Then

$$\begin{aligned}|\mathcal{G}| &\leq \sum_{t=1}^{t^*} \binom{n}{t} + \sum_{t=t^*+1}^{n} \frac{\binom{n}{d}k}{\binom{t^*}{d}} \cdot \frac{\binom{t^*}{d}}{\binom{t}{d}} \\ &\leq \sum_{t=1}^{t^*} \binom{n}{t} + \binom{n}{t^*} \sum_{t=t^*+1}^{n} \frac{\binom{t^*}{d}}{\binom{t}{d}}\end{aligned}$$

To complete the proof, note that for any $d \leq t \leq n$ we have $\binom{n}{t}\binom{t}{d} = \binom{n}{d}\binom{n-d}{t-d}$, hence t^* is the smallest integer t satisfying $\binom{n-d}{t-d} \geq k$ if such an integer exists, and otherwise $t^* = n$. ■

IV. Discussion

Given a zero-error codebook pair $\mathcal{C}_1, \mathcal{C}_2 \subseteq \{0,1\}^n$ with cardinalities 2^{nR_1} and 2^{nR_2} respectively, our bounding technique was based on a procedure for constructing a zero-error system $\mathcal{V}$ with dimension $(1-\alpha)n$. This was achieved by proving the existence of a subset $S \subset [n]$ of cardinality αn, such that the sumset of the projection multisets of each codebook on S, i.e., $P_S^+(\mathcal{C}_1) + P_S^+(\mathcal{C}_2)$ has a member $\mathbf{v} \in \{0,1,2\}^{|S|}$ with a large number of occurrences, say $2^{n\rho}$. This in turn implied that $r_0 + r_1 + r_2$ for the system is at least $\rho/(1-\alpha)$. To lower bound ρ as a function of α and the cardinalities of the original codebooks, we introduced the soft Sauer-Perles-Shelah Lemma, which enabled us to bound the number of occurrences of the vector $\mathbf{v} = \mathbf{1}_{|S|}$. This lemma offered the additional benefit of a lower bound on r_1. We note in passing that the bound obtained on R_2 as a function of R_1 outperforms previous results even without incorporating the constraint on r_1. We suspect that better bounds on ρ can be obtained, possibly for $\mathbf{v}$ other than $\mathbf{1}_{|S|}$.

References

[1] B. Lindström, “Determination of two vectors from the sum,” *Journal of Combinatorial Theory*, vol. 6, no. 4, pp. 402–407, 1969.

[2] H. van Tilborg, “An upper bound for codes in a two-access binary erasure channel (corresp.),” *IEEE Transactions on Information Theory*, vol. 24, no. 1, pp. 112–116, 1978.

[3] T. Kasami and S. Lin, “Bounds on the achievable rates of block coding for a memoryless multiple-access channel,” *IEEE Transactions on Information Theory*, vol. 24, no. 2, pp. 187–197, 1978.

[4] E. Weldon, “Coding for a multiple-access channel,” *Information and Control*, vol. 36, no. 3, pp. 256–274, 1978.

[5] T. Kasami, S. Lin, V. Wei, and S. Yamamura, “Graph theoretic approaches to the code construction for the two-user multiple-access binary adder channel,” *IEEE Transactions on Information Theory*, vol. 29, no. 1, pp. 114–130, 1983.

[6] P. van den Braak and H. van Tilborg, “A family of good uniquely decodable code pairs for the two-access binary adder channel,” *IEEE Transactions on Information Theory*, vol. 31, no. 1, pp. 3–9, 1985.

[7] S. Bross and I. Blake, “Upper bound for uniquely decodable codes in a binary input N-user adder channel,” *IEEE Transactions on Information Theory*, vol. 44, no. 1, pp. 334–340, Jan 1998.

[8] R. Urbanke and Q. Li, “The zero-error capacity region of the 2-user synchronous bac is strictly smaller than its shannon capacity region,” in *Information Theory Workshop*, Jun 1998, p. 61.

[9] R. Ahlswede and V. Balakirsky, “Construction of uniquely decodable codes for the two-user binary adder channel,” *IEEE Transactions on Information Theory*, vol. 45, no. 1, pp. 326–330, 1999.

[10] M. Mattas and P. Östergård, “A new bound for the zero-error capacity region of the two-user binary adder channel,” *IEEE Transactions on Information Theory*, vol. 51, no. 9, pp. 3289–3291, 2005.

[11] N. Alon and J. H. Spencer, *The probabilistic method*. John Wiley & Sons, 2004.

[12] D. Slepian and J. K. Wolf, “A coding theorem for multiple access channels with correlated sources,” *Bell System Technical Journal*, vol. 52, no. 7, pp. 1037–1076, 1973.

[13] O. Ordentlich and O. Shayevitz, “An upper bound on the sizes of multiset-union-free families,” 2014, available online http://arxiv.org/abs/1412.8415.

[14] N. Alon, “On the density of sets of vectors,” *Discrete Mathematics*, vol. 46, no. 2, pp. 199–202, 1983.