The rate of Index Coding

Noga Alon* Eyal Lubetzky† Uri Stav ‡

Abstract

The problem of Informed Source Coding on Demand, introduced by Birk and Kol [4], describes a setting where a server wishes to transmit an n-bit input word via broadcast to a n receivers; each receiver is interested in a specific block of the input data and may have side-information on other blocks. The goal of the sender is to use a code of minimal word-length, while allowing every receiver to recover his desired bit. This problem can be formulated as a graph parameter: let G be a directed graph on the vertex set $[n]$, where ij is an edge iff the i-th receiver knows the j-th bit, and let $\ell(G)$ denote the length of a minimal index code for G.

In this work, we show that an index code for a disjoint union of graphs can be strictly better than a concatenation of the optimal codes for the individual graphs. This justifies a definition of a "rate" of an index code of a graph, as the (normalized) length of an index code when performing multiple transmissions with a constant side-information graph. The proof is based on a relation between index codes for a disjoint union of graphs and the OR product of certain graphs, and relies on some results in the study of Witsenhausen's rate and colorings of graph powers.

1 Introduction

Source coding deals with a scenario in which a *sender* has some data string x he wishes to transmit through a broadcast channel to *receivers*. In this paper we consider a variant of source coding which was first proposed by Birk and Kol [4]. In this variant, called Informed Source Coding On Demand (ISCOD), each receiver has prior side information, comprising some subset of the input word x. The sender is aware of the portion of x known to each receiver. Moreover, each receiver is interested in just part of the data. Following Bar-Yossef, Birk, Jayram and Kol [2], we restrict ourselves to the problem which is formalized as follows.

Definition 1 (Index code). *A sender wishes to send a word $x \in \{0,1\}^n$ to n receivers $R_1, \ldots, R_n$. Each R_i knows some of the bits of x and is interested solely in the bit x_i. An* index code *of length ℓ for this setting is a binary code of word-length ℓ, which enables R_i to recover x_i for any x and i.*

*Schools of Mathematics and Computer Science, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, 69978, Israel. Email: nogaa@tau.ac.il. Research supported in part by a USA-Israeli BSF grant, by the Israel Science Foundation and by the Hermann Minkowski Minerva Center for Geometry at Tel Aviv University.

†School of Mathematics, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, 69978, Israel. Email: lubetzky@tau.ac.il. Research partially supported by a Charles Clore Foundation Fellowship.

‡School of Computer Science, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, 69978, Israel. Email: uristav@tau.ac.il.

Using a graph model for the side-information, this problem can be restated as a graph parameter. For a directed graph G and a vertex v, let $N_G^+(v)$ be the set of out-neighbors of v in G, and for $x \in \{0,1\}^n$ and $S \subset [n] = \{1,\dots,n\}$, let $x|_S$ be the restriction of x to the coordinates of S.

Definition 2 ($\ell(\mathbf{G})$)**.** *The setting of Definition 1 is characterized by the directed* side information graph *G on the vertex set $[n]$, where (i,j) is an edge iff R_i knows the value of x_j. An* index code *of length ℓ for G is a function $E : \{0,1\}^n \to \{0,1\}^\ell$ and functions $D_1,\dots,D_n$, so that for all $i \in [n]$ and $x \in \{0,1\}^n$, $D_i(E(x), x|_{N_G^+(i)}) = x_i$. Denote the minimal length of an index code for G by $\ell(G)$.*

For instance, one can verify that if G is edgeless then $\ell(G) = n$. More generally, it can be seen that $\ell(G) \geq \alpha(G)$, where $\alpha(G)$ is the cardinality of a maximum *independent set* (a set of vertices which are pairwise disjoint)[1]. On the other hand, if G is the complete graph then $\ell(G) = 1$ (this is achieved by transmitting the Xor of all the input word bits). Another easy property of $\ell(G)$ is monotonicity with respect to removal of edges, namely if G_2 is obtained from G_1 by removing some of the edges then $\ell(G_1) \leq \ell(G_2)$. See [4], [2] and [11] for several other examples of index coding schemes and details on the distributed application which motivates this problem.

In several of our results, it will be convenient to address the more precise notion of the *number of codewords* in an index code. We say that $\mathcal{C}$, an index code for G, is *optimal*, if it contains the minimum possible number of codewords (in which case, $\ell(G) = \lceil \log_2 |\mathcal{C}| \rceil$).

Before reviewing the main previous results on index coding, we define the following graph theoretic parameters. Let $G = (V,E)$ be a directed graph on the vertex set $V = [n]$. The adjacency matrix of G, denoted by $A_G = (a_{ij})$, is the $n \times n$ binary matrix where $a_{ij} = 1$ iff $(i,j) \in E$. The *chromatic number* of G, $\chi(G)$, is the minimum number of independent sets whose union is all of V (each such set is referred to as a *color group*). Let $\overline{G}$ denote the *graph complement* of G. We denote by $G + H$ the disjoint union of the graphs G and H, and denote by $k \cdot G$ the disjoint union of k copies of the graph G.

Let $A = (a_{ij})$ be an $n \times n$ matrix over some field $\mathbb{F}$. We say that A *represents* the graph G over $\mathbb{F}$ if $a_{ii} \neq 0$ for all i, and $a_{ij} = 0$ whenever $i \neq j$ and $(i,j) \notin E$. The *minrank* of a directed graph G with respect to the field $\mathbb{F}$ is defined by

$$\mathrm{minrk}_{\mathbb{F}}(G) = \min\{\mathrm{rank}_{\mathbb{F}}(A) : A \text{ represents } G \text{ over } \mathbb{F}\} . \tag{1}$$

For the common case where $\mathbb{F}$ is a finite field, we abbreviate: $\mathrm{minrk}_{p^k}(G) = \mathrm{minrk}_{GF(p^k)}(G)$. The notion of $\mathrm{minrk}(G)$ for an undirected graph G was first considered in the context of graph capacities by Haemers [6],[7]. Bar-Yossef, Birk, Jayram and Kol [2] showed that for any graph G, $\ell(G) \leq \mathrm{minrk}_2(G)$, and that $\mathrm{minrk}_2(G)$ is in fact the length of the optimal *linear* index code for G. The authors of [2] proved that in various cases, linear codes are in fact optimal, and their main conjecture was that linear index coding is *always* optimal, that is $\ell(G) = \mathrm{minrk}_2(G)$ for any graph G. In [11], This conjecture was disproved in, essentially, the strongest possible way, showing that for any positive $\varepsilon > 0$ there exist graphs where every linear index code requires $n^{1-\varepsilon}$ bits (barely improving the n bits required by the naïve protocol), whereas a given non-linear index code utilizes a word-length of only n^ε bits.

[1] A slightly more sophisticated argument is used in [2] to show that for a directed graph G: $\ell(G) \geq \mathrm{MAIS}(G)$ where $\mathrm{MAIS}(G)$ denotes the maximum number of vertices in an induced acyclic subgraph of G.

In this paper, we investigate an interesting relation between optimal index codes of a disjoint union of graphs, and vertex-colorings of OR graph products. Using this connection, we obtain several rather surprising results, both on index coding and on generalized index coding problems considered in [4] and [11], where there is a larger-alphabet or multiple-round transmission.

So far, every upper bound on the size of the optimal index code was based on linear codes. Indeed, the non-linear index codes constructed by [11], which drastically outperform all linear index codes, are in fact hybrids of linear index codes over higher-order finite fields. In this work, we present a new probabilistic and inherently non-linear upper bound on the size of the optimal index code for a disjoint union of graphs. This is incorporated in the next theorem, where here and in what follows, all logarithms are in the natural basis unless stated otherwise.

Theorem 1.1. *Let G be a directed side-information graph on n vertices, and let γ denote the maximal number of input-strings in $\{0,1\}^n$, which can be encoded by the same codeword in an index code for G (without danger of confusion). The following holds for any integer k:*

$$1 \leq |\mathcal{C}| \left(\frac{\gamma}{2^n}\right)^k \leq \lceil kn \log 2 \rceil \tag{2}$$

where $\mathcal{C}$ is an optimal index code for $k \cdot G$. In particular, $\lim_{k\to\infty} \frac{\ell(k \cdot G)}{k} = n - \log_2 \gamma$.

As a corollary of the above theorem, we deduce that an index code for the disjoint union of graphs can be strictly better than the concatenation of the individual optimal index codes (see Corollary 3.1). To illustrate this surprising result, suppose that an optimal index code for G consists of, say, 10 distinct codewords. In this case, it seems unlikely that fewer than 10^k words can suffice in an index code for $k \cdot G$, since each of the k copies of G has no side-information on any of the bits corresponding to the remaining copies, and seeks to recover its own independently chosen input-word.

It is not difficult to show that a *linear* index code for $k \cdot G$ can never outperform the k-fold concatenation of optimal linear index codes for G. Hence, as a by-product, we obtain an additional proof that linear index coding is suboptimal (this time, by a multiplicative constant, see Corollary 3.3).

A final corollary of Theorem 1.1 is in the context of multiple-round index coding. We begin by defining a natural extension of index coding, pointed out in [4] and [11]:

Definition 3 (**Generalized index coding**). *Consider the following equivalent formulations*[2]*:*

1. *(Larger alphabet) Each input symbol of the data string x comprises a* block *of $b \geq 1$ bits. Every receiver is interested in a single block, and knows a subset of the other blocks.*
2. *(Multiple-rounds) The sender accepts $b \geq 1$ input-words in $\{0,1\}^n$, all with respect to the same side-information graph G. Each R_i is interested in recovering the i-th bit of all b input-words.*

It was observed in [11] that this generalized problem can be reduced to the original binary setting by considering the *b-blow-up* of G (with independent sets). This graph, denoted by $G[b]$,

[2] These two problems are not equivalent in their general setting. Indeed, the alphabet need not be a power of 2, and furthermore, the multiple-round transmission need not be over the same fixed side-information graph.

has the vertex set $V(G) \times [b]$, and an edge from (u,i) to (v,j) iff $uv \in E(G)$. By definition, an index-code for the generalized problem in Definition 3 is precisely an index-code for $G[b]$.

Clearly, one can always encode each round independently, and hence $\ell(G[b]) \leq b{\cdot}\ell(G)$. In various cases this is optimal (e.g., whenever G is a perfect graphs, cf. [11]), and the natural question arising is whether one can benefit from using a unified scheme that encodes all the words at once. As we later show, the answer for this question is positive, justifying the following definition:

Definition 4 (**Index-code rate**). *The* index-code rate *of a graph G is the asymptotic average communication that is required for a single round when allowing a unified transmission for multiple rounds:*

$$\bar{\ell}(G) = \lim_{b\to\infty} \frac{\ell(G[b])}{b} = \inf_b \frac{\ell(G[b])}{b} \ .$$

The above limit exists and equals the infimum by sub-additivity, and as we will later see (Corollary 3.5), it is nontrivial: there are graphs for which the index-code rate is strictly smaller than $\ell(G)$. In these cases, there is benefit in performing multiple-rounds, and a unified transmission significantly improves the efficiency of the coding scheme.

The rest of this paper is organized as follows. In Section 2, we prove Theorem 1.1, using a connection between index codes for disjoint unions of graphs and the chromatic number of OR graph products. Section 3 contains applications of this theorem, which provide new non-linear index codes for disjoint unions of graphs and for multiple-round transmission over the same graph. Section 4 is devoted to concluding remarks and open problems.

2 Optimal index codes for a disjoint union of graphs

Proof of Theorem 1.1. The OR graph product is equivalent to the complement of the *strong* product[3], which was thoroughly studied in the aspect of the Shannon capacity of a graph, a notoriously challenging graph parameter introduced by Shannon [13].

Definition 5 (**OR graph product**). *The* OR graph product *of G and H, denoted by $G \curlyvee H$, is the graph on the vertex set $V(G) \times V(H)$, where (u,v) and (u',v') are adjacent iff either $uu' \in E(G)$ or $vv' \in E(H)$ (or both). Let $G^{\curlyvee k}$ denote the k-fold OR product of a graph G.*

The size of an optimal index code for a given directed graph may be restated as a problem of determining a chromatic number of a graph, as observed by Bar-Yossef et al. [2]. We need the following definition:

Definition 6 (**Confusion graph**). *Let $G = ([n], E)$ be a directed side-information graph. The* confusion graph *of G, $\mathfrak{C}(G)$, is the undirected graph whose vertex set is $\{0,1\}^n$, and two vertices $x, y \in \{0,1\}^n$ are adjacent iff for some $i \in [n]$, $x_i \neq y_i$ and yet $x|_{N_G^+(i)} = y|_{N_G^+(i)}$.*

In other words, $\mathfrak{C}(G)$ is the graph whose vertex set is all possible input-words, and two vertices are adjacent iff they cannot be encoded by the same codeword in an index code for G (otherwise, the decoding of one of the receivers would be ambiguous). Hence, every index code for G is equivalent

[3]Namely, the OR product of G and H is the complement of the strong product of $\overline{G}$ and $\overline{H}$.

to a legal vertex coloring of $\mathfrak{C}(G)$, where each color-group corresponds to a distinct codeword. Altogether, if $\mathcal{C}$ is an optimal index code for G, then $|\mathcal{C}| = \chi(\mathfrak{C}(G))$.

Let G and H denote directed graphs on the vertex-sets $[m]$ and $[n]$ respectively, and consider an index code for their disjoint union, $G + H$. As there are no edges between G and H, such an index code cannot encode two input-words $x, y \in \{0,1\}^{m+n}$ by the same codeword iff this forms an ambiguity either with respect to G or with respect to H (or both). Hence:

Observation 2.1. *For any two directed graphs G and H, the graphs $\mathfrak{C}(G+H)$ and $\mathfrak{C}(G) \curlyvee \mathfrak{C}(H)$ are isomorphic.*

Altogether, the number of codewords in an optimal index code for $k \cdot G$ is equal to $\chi(\mathfrak{C}(G)^{\curlyvee k})$. The chromatic numbers of strong powers of a graph, as well as those of OR graph powers, have been thoroughly studied. In the former case, they correspond to the Witsenhausen rate of a graph (see [14]). In the latter case, the following was proved by McEliece and Posner [12], and also by Berge and Simonovits [3]:

$$\lim_{k\to\infty} \left(\chi(G^{\curlyvee k})\right)^{1/k} = \chi_f(G) \ , \tag{3}$$

where χ_f is the *fractional chromatic number* of a graph, defined as follows. A legal vertex coloring corresponds to an assignment of $\{0,1\}$-weights to independent-sets, such that every vertex will be "covered" by a total weight of at least 1. A fractional coloring is the relaxation of this problem where the weights belong to $[0,1]$, and χ_f is the minimal sum of weights in such a coloring.

To obtain an estimate on the rate of the convergence in (3), we will use the following well-known properties of the fractional chromatic number and OR graph products (cf. [1],[10],[9] and also [8]):

(i) For any graph H, $\chi_f(H^{\curlyvee k}) = \chi_f(H)^k$.

(ii) For any graph H, $\chi_f(H) \leq \chi(H) \leq \lceil\ \chi_f(H) \log |V(H)|\ \rceil$ (to see this, collect independent sets, chosen randomly and independently according to the weight distribution, dictated by the optimal weight-function achieving χ_f).

(iii) For any vertex transitive graph H (that is, a graph whose automorphism group is closed under all possible vertex substitutions), $\chi_f(H) = |V(H)|/\alpha(H)$ (cf., e.g., [5]).

In order to translate (ii) to the statement of (2), notice that γ, as defined in Theorem 1.1 is precisely $\alpha(\mathfrak{C}(G))$. In addition, the graph $\mathfrak{C}(G)$ is indeed vertex transitive (in fact, it is a Cayley graph over $GF(2)^n$), and combining the above facts we obtain that:

$$\chi_f\left(\mathfrak{C}(G)^{\curlyvee k}\right)^{1/k} = \frac{2^n}{\alpha(\mathfrak{C}(G))} = \frac{2^n}{\gamma} \ .$$

This completes the proof of the theorem. ■

Remark 2.2: The right-hand-side of (2) can be replaced by $\lceil 1 + k \log \gamma \rceil$. To see this, combine the simple fact that $\alpha(G^{\curlyvee k}) = \alpha(G)^k$ with the bound $\chi(G) \leq \lceil \chi_f(G)(1 + \log \alpha(G)) \rceil$ given in [10] (choose $\chi_f(G) \log \alpha(G)$ independent sets as before, leaving at most $\chi_f(G)$ uncovered vertices).

3 Applications of non-linear coding schemes

Recall that, as shown by Haemers [6], [7], for any field $\mathbb{F}$ we have

$$\alpha(G) \le c(G) \le \mathrm{minrk}_{\mathbb{F}}(G) \le \chi(\overline{G}) \ ,$$

where $c(G)$ is the Shannon capacity of G. The authors of [2] showed that

$$\alpha(G) \le \ell(G) \le \mathrm{minrk}_2(G) \ ,$$

thus whenever $\alpha(G) = \chi(\overline{G})$ we have $\ell(G) = \mathrm{minrk}_{\mathbb{F}}(G) = \alpha(G)$. In particular, this holds for all perfect graphs. It is not difficult to verify that, by definition

$$\mathrm{minrk}_{\mathbb{F}}(G+H) = \mathrm{minrk}_{\mathbb{F}}(G) + \mathrm{minrk}_{\mathbb{F}}(H) \ \text{ for all } G, H \text{ and field } \mathbb{F}. \tag{4}$$

As the disjoint union of perfect graphs is perfect as well, we conclude that for any perfect graph G and any integer k, $\ell(k \cdot G) = k \cdot \ell(G)$. Therefore, the smallest example where $\ell(k \cdot G)$ might be nontrivial is C_5, the smallest non-perfect graph. Indeed, one can do better than $k \cdot \ell(C_5)$:

Corollary 3.1. *The following holds:* $\lim_{k\to\infty} \dfrac{\ell(k \cdot C_5)}{k} = 5 - \log_2 5 \approx 2.678$*, whereas* $\ell(C_5) = 3$.

Proof. One can verify that the following is a maximum independent set of size 5 in $\mathfrak{C}(C_5)$:

$$\{00000, 01100, 00011, 11011, 11101\}$$

Theorem 1.1 now implies that $\ell(k \cdot C_5)/k$ tends to $5 - \log_2 5$ as $k \to \infty$. On the other hand, one can verify[4] that $\chi(\mathfrak{C}(C_5)) = 8$, hence $\ell(C_5) = 3$. ∎

Remark 3.2: Using the upper bound of (2) in its alternate form, as stated in Remark 2.2, we obtain that $\ell(k \cdot C_5) < k \cdot \ell(C_5)$ already for $k = 15$.

It is well known that the Shannon capacity of C_5 is $\sqrt{5}$, and as the chromatic number of its complement (which is isomorphic) is 3, we have $\mathrm{minrk}_{\mathbb{F}}(C_5) = 3$ for any field $\mathbb{F}$. Therefore, the above discussion provides an alternative proof that linear index coding is suboptimal, this time by a multiplicative constant:

Corollary 3.3. *For any field $\mathbb{F}$, linear index coding over $\mathbb{F}$ for the side-information graph C_5 is suboptimal:*

$$\frac{\ell(k \cdot C_5)}{\mathrm{minrk}_{\mathbb{F}}(k \cdot C_5)} = \frac{5 - \log_2 5}{3} + o(1) \approx 0.89 + o(1) \ ,$$

where the $o(1)$-term tends to 0 as $k \to \infty$.

Remark 3.4: It is possible to show that not only is any linear index coding scheme for the above graph suboptimal, but so is any hybrid of linear index codes over different fields (as in the approach by the authors of [11]). To see this, notice that it suffices to examine a single copy of C_5 (by the behavior of minrk with respect to disjoint unions), where this statement is immediate.

[4] This fact can be verified using computer simulations, as stated in [2].

Recall that a unified index code for $b \geq 1$ rounds over the same side-information graph G is equivalent to an index code for $G[b]$, the b-blow-up of G. In what follows, the lower bound $\alpha(G)$ can be refined in directed graphs to $\mathrm{MAIS}(G)$ (the size of a maximum acyclic induced subgraph).

Corollary 3.5. *For any undirected side-information G, $\alpha(G) \leq \overline{\ell}(G) \leq n - \log_2 \alpha(\mathfrak{C}(G))$.*

Proof. For the lower bound, note that $\alpha(G[b]) = b \cdot \alpha(G)$, hence $\ell(G[b]) \geq b \cdot \alpha(G)$ and $\overline{\ell}(G) \geq \alpha(G)$.

For the upper bound, notice that $G[b]$ contains the subgraph $b \cdot G$ (for each $i \in [b]$, the graph $V(G) \times \{i\}$ is isomorphic to G). Thus, by the monotonicity of index coding with respect to addition of edges, $\ell(G[b]) \leq \ell(b \cdot G)$, and Theorem 1.1 completes the proof. ■

Whenever G satisfies $\alpha(G) = \chi(\overline{G})$, we have

$$\alpha(G) \leq \ell(G[b]) \leq b \cdot \ell(G) = b \cdot \alpha(G) \ ,$$

and hence in this case $\ell(G[b]) = b \cdot \alpha(G)$ for any integer b, and $\overline{\ell}(G) = \ell(G) = \alpha(G)$. Once again, the smallest example where multiple-round index coding can improve the performance of a single round is C_5 (the smallest non-perfect graph). Indeed, combining the above with Corollary 3.1 gives:

Corollary 3.6. *The following holds:* $2 \leq \overline{\ell}(C_5) \leq 5 - \log_2 5 \approx 2.678$*, whereas* $\ell(C_5) = 3$.

4 Conclusion and open problems

- In this work, we have shown that for sufficiently large values of k and for every graph G, $\ell(k \cdot G) = (n - \log_2 \alpha(\mathfrak{C}(G)) + o(1))\, k$, where the $o(1)$-term tends to 0 as $k \to \infty$.
- As a corollary, we demonstrated that $\ell(k \cdot C_5)/k \approx 2.678$ for large values of k, whereas $\ell(C_5) = 3$. Hence, the optimal index code for a disjoint union of k copies of a graph G can be strictly better than the concatenation of k optimal index codes for G (benefit is gained already for $k = 15$). This is surprising, considering the lack of mutual information between receivers which correspond to distinct copies.
- Furthermore, as $\mathrm{minrk}_2(k \cdot C_5) = 3k$, this provides another proof for the fact that linear index coding is suboptimal (by a multiplicative constant).
- Our results also imply that multiple-round index coding over the same given graph G can be strictly better on average than the optimal index code for G. In the above case of C_5, $2 \leq \overline{\ell}(C_5) \leq 2.678$, and it would be interesting to determine the index code rate of C_5 precisely.

Acknowledgement: We would like to thank Amit Weinstein for helpful and efficient computer simulations.

References

[1] N. Alon and A. Orlitsky, Repeated communication and Ramsey graphs, IEEE Transactions on Information Theory 41 (1995), 1276-1289.

[2] Z. Bar-Yossef, Y. Birk, T.S. Jayram and T. Kol, Index coding with side information, Proc. of the 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2006), pp. 197-206.

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

[4] Y. Birk and T. Kol, Coding-on-demand by an informed source (ISCOD) for efficient broadcast of different supplemental data to caching clients, IEEE Transactions on Information Theory 52 (2006), 2825-2830. An earlier version appeared in INFOCOM '98.

[5] C. Godsil and G. Royle, Algebraic Graph Theory, volume 207 of Graduate Text in Mathematics, Springer, New York, 2001.

[6] W. Haemers, An upper bound for the Shannon capacity of a graph, Colloq. Math. Soc. János Bolyai 25, Algebraic Methods in Graph Theory, Szeged, Hungary (1978), 267-272.

[7] W. Haemers, On some problems of Lovász concerning the Shannon capacity of a graph, IEEE Transactions on Information Theory 25 (1979), 231-232.

[8] U. Feige, Randomized graph products, chromatic numbers, and the Lovász ϑ-function, Proc. of the 27th ACM Symposium on Theory of Computing (STOC 1995), pp. 635-640.

[9] N. Linial and U. Vazirani, Graph products and chromatic numbers, Proc. of the 30th Annual IEEE Symposium on Foundations of Computer Science (FOCS 1989), pp. 124-128.

[10] L. Lovász, On the ratio of optimal integral and fractional covers, Discrete Math. 13 (1975), no. 4, 383-390.

[11] E. Lubetzky and U. Stav, Non-linear index coding outperforming the linear optimum, to appear.

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

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

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