Error Correction for Index Coding With Side Information

Son Hoang Dau, Vitaly Skachek, and Yeow Meng Chee, *Senior Member, IEEE*

***Abstract*—A problem of index coding with side information was first considered by Birk and Kol in 1998. In this study, a generalization of index coding scheme, where transmitted symbols are subject to errors, is studied. Error-correcting methods for such a scheme, and their parameters, are investigated. In particular, the following question is discussed: given the side information hypergraph of index coding scheme and the maximal number of erroneous symbols δ, what is the shortest length of a linear index code, such that every receiver is able to recover the required information? This question turns out to be a generalization of the problem of finding a shortest length error-correcting code with a prescribed error-correcting capability in the classical coding theory. The Singleton bound and two other bounds, referred to as the α-bound and the κ-bound, for the optimal length of a linear error-correcting index code (ECIC) are established. For large alphabets, a construction based on concatenation of an optimal index code with a maximum distance separable classical code is shown to attain the Singleton bound. For smaller alphabets, however, this construction may not be optimal. A random construction is also analyzed. It yields another inexplicit bound on the length of an optimal linear ECIC. Further, the problem of error-correcting decoding by a linear ECIC is studied. It is shown that in order to decode correctly the desired symbol, the decoder is required to find one of the vectors, belonging to an affine space containing the actual error vector. The syndrome decoding is shown to produce the correct output if the weight of the error pattern is less or equal to the error-correcting capability of the corresponding ECIC. Finally, the notion of static ECIC, which is suitable for use with a family of instances of an index coding problem, is introduced. Several bounds on the length of static ECICs are derived, and constructions for static ECICs are discussed. Connections of these codes to weakly resilient Boolean functions are established.**

***Index Terms*—Broadcast, error correction, index coding, minimum distance, network coding, side information.**

Manuscript received May 10, 2011; revised March 05, 2012; accepted September 04, 2012. Date of publication November 15, 2012; date of current version February 12, 2013. This work was supported by the National Research Foundation of Singapore under Research Grant NRF-CRP2-2007-03. This paper was presented in part at the 2011 IEEE International Symposium on Information Theory.

S. H. Dau was with the Division of Mathematical Sciences, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 637371. He is now with SUTD-MIT International Design Centre, Singapore University of Technology and Design, 20 Dover Drive, Singapore 138682 (e-mail: dausonhoang84@gmail.com).

V. Skachek was with the Division of Mathematical Sciences, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 637371. He is now with the Institute of Computer Science, Faculty of Mathematics and Computer Science, University of Tartu, Tartu 50409, Estonia (e-mail: vitaly.skachek@gmail.com).

Y. M. Chee is with the Division of Mathematical Sciences, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore 637371 (e-mail: ymchee@ntu.edu.sg).

Communicated by A. Ashikhmin, Associate Editor for Coding Techniques.

Digital Object Identifier 10.1109/TIT.2012.2227674

I. Introduction

A. Background

The problem of index coding with side information (ICSI) was introduced by Birk and Kol [1], [2]. During the transmission, each client might miss a certain part of the data, due to intermittent reception, limited storage capacity, or any other reasons. Via a slow backward channel, the clients let the server know which messages they already have in their possession, and which messages they are interested to receive. The server has to find a way to deliver to each client all the messages he requested, yet spending a minimum number of transmissions. As it was shown in [1], the server can significantly reduce the number of transmissions by coding the messages.

The toy example in Fig. 1 presents a scenario with one broadcast transmitter and four receivers. Each receiver requires a different information packet (we sometimes simply call it message). The naive approach requires four separate transmissions, one transmission per an information packet. However, by exploiting the knowledge on the subsets of messages that clients already have, and by using coding of the transmitted data, the server can just broadcast one coded packet.

Possible applications of index coding include communications scenarios, in which a satellite or a server broadcasts a set of messages to a set of clients, such as daily newspaper delivery or video-on-demand. ICSI can also be used in opportunistic wireless networks. These are the networks in which a wireless node can opportunistically listen to the wireless channel. The client may obtain packets that are not designated to it (see [3]–[5]). As a result, a node obtains some side information about the transmitted data. Exploiting this additional knowledge may help to increase the throughput of the system.

The ICSI problem has been a subject of several recent studies [3], [6]–[13]. This problem can be viewed as a special case of the network coding (NC) problem [14], [15]. On the other hand, it was shown that every instance of the NC problem can be reduced to an instance of the ICSI problem in the following sense. For each NC instance, we can construct an ICSI instance such that there exists a scalar linear network code for the NC instance if and only if there exists a perfect scalar linear index code for the corresponding ICSI instance (see [3] and [11] for more details).

B. Our Contribution

The preceding works on the ICSI problem consider scenario where the transmissions are error-free. In practice, of course, this might not be the case. In this study, we assume that the transmitted symbols are subject to errors. We extend some known

0018-9448/$31.00 © 2012 IEEE

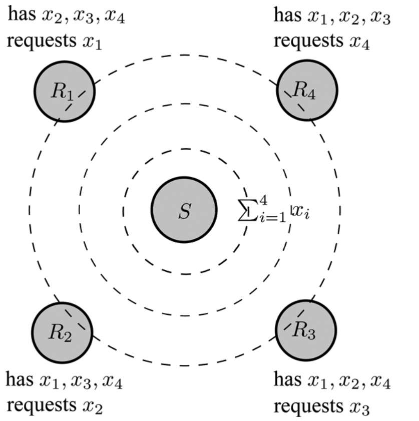


Fig. 1. Example of the ICSI problem.

results on index coding to a case where any receiver can correct up to a certain number of errors. It turns out that the problem of designing such error-correcting index codes (ECICs) naturally generalizes the problem of constructing classical error-correcting codes.

More specifically, assume that the number of messages that the server possesses is n, and that the designed maximal number of errors is δ. We show that the problem of constructing ECIC of minimal possible length is equivalent to the problem of constructing a matrix $\boldsymbol{L}$ which has n rows and the minimal possible number of columns, such that

$$\mathsf{wt}\,(\boldsymbol{z}\boldsymbol{L}) \geqslant 2\delta + 1 \quad \text{for all } \boldsymbol{z} \in \mathcal{I}$$

where $\mathcal{I}$ is a certain subset of $\mathbb{F}_q^n \backslash \{\mathbf{0}\}$. Here, $\mathsf{wt}(\boldsymbol{x})$ denotes the Hamming weight of the vector $\boldsymbol{x}$, $\mathbb{F}_q$ stands for a finite field with q elements, and $\mathbf{0}$ is the all-zeros vector. If $\mathcal{I} = \mathbb{F}_q^n \backslash \{\mathbf{0}\}$, this problem becomes equivalent to the problem of designing a shortest length linear code of given dimension and minimum distance.

In this study, we establish an upper bound (the κ-bound) and a lower bound (the α-bound) on the shortest length of a linear ECIC, which is able to correct any error pattern of size up to δ. More specifically, let $\mathcal{H}$ be the side information hypergraph that describes the instance of the ICSI problem. Let $\mathcal{N}_q[\mathcal{H}, \delta]$ denote the length of a shortest length linear ECIC over $\mathbb{F}_q$, such that every R_i can recover the desired message, if the number of errors is at most δ. We use notation $N_q[k, d]$ for the length of an optimal linear error-correcting code of dimension k and minimum distance d over $\mathbb{F}_q$. We obtain

$$N_q[\alpha(\mathcal{H}), 2\delta + 1] \leqslant \mathcal{N}_q[\mathcal{H}, \delta] \leqslant N_q[\kappa_q(\mathcal{H}), 2\delta + 1] \qquad (1)$$

where $\alpha(\mathcal{H})$ is the generalized independence number and $\kappa_q(\mathcal{H})$ is the min-rank (over $\mathbb{F}_q$) of $\mathcal{H}$.

For linear index codes, we also derive an analog of the Singleton bound. This result implies that (over sufficiently large alphabet) the concatenation of a standard maximum distance separable (MDS) error-correcting code with an optimal linear index code yields an optimal linear ECIC. Finally, we consider random ECICs. By analyzing its parameters, we obtain an upper bound on its length.

When the side information hypergraph is a pentagon, $\delta = 2$ and $q = 2$, the inequalities in (1) are shown to be strict. This implies that a concatenated scheme based on a classical error-correcting code and on a linear non-error-correcting index code does not necessarily yield an optimal linear error-correcting index code. Since ICSI problem can also be viewed as a source coding problem [6], [13], this example demonstrates that sometimes designing a single code for both source and channel coding can result in a smaller number of transmissions.

The decoding of a linear ECIC is somewhat different from that of a classical error-correcting code. There is no longer a need for a complete recovery of the whole information vector. We analyze the decoding criteria for the ECICs and show that the syndrome decoding, which might be different for each receiver, results in a correct result, provided that the number of errors does not exceed the error-correcting capability of the code.

An ECIC is called static under a family of instances of the ICSI problem if it works for all of these instances. Such an ECIC is interesting since it remains useful as long as the parameters of the problem vary within a particular range. Bounds and constructions for static ECICs are studied in Section VIII. Connections between static ECICs and weakly resilient vectorial Boolean functions are also discussed.

The problem of error correction for NC was studied in several previous works. However, these results are not directly applicable to the ICSI problem. First, there is only a very limited variety of results for nonmulticast networks in the existing literature. The ICSI problem, however, is a special case of the nonmulticast NC problem. Second, the ICSI problem can be modeled by the NC scenario [3], yet, this requires that there are directed edges from particular sources to each sink, which provide the side information. The symbols transmitted on these special edges are not allowed to be corrupted. By contrast, for error-correcting NC, symbols transmitted on all edges can be corrupted.

C. Organization

This paper is organized as follows. Basic notations and definitions, used throughout this paper, are provided in Section II. The problem of index coding with and without error correction is introduced in Section III. Some basic results are presented in that section. The α-bound and the κ-bound are derived in Section IV. The Singleton bound is presented in Section V. Random codes are discussed in Section VI. Syndrome decoding is studied in Section VII. A notion of static ECICs is presented in Section VIII. Several bounds on the length of such codes are derived, and connections to resilient function are shown in that section. Finally, the results are summarized in Section IX, and some open questions are proposed therein.

II. Preliminaries

In this section, we introduce some useful notation. Here, $\mathbb{F}_q$ is the finite field of q elements, where q is a power of prime, and $\mathbb{F}_q^*$ is the set of all nonzero elements of $\mathbb{F}_q$.

Let $[n]=\{1,2,\dots,n\}$. For the vectors $\boldsymbol{u}=(u_1,u_2,\dots,u_n)\in\mathbb{F}_q^n$ and $\boldsymbol{v}=(v_1,v_2,\dots,v_n)\in\mathbb{F}_q^n$, the (Hamming) distance between $\boldsymbol{u}$ and $\boldsymbol{v}$ is defined to be the number of coordinates where $\boldsymbol{u}$ and $\boldsymbol{v}$ differ, namely:

$$\mathsf{d}(\boldsymbol{u},\boldsymbol{v})=|\{i\in[n]\ :\ u_i\neq v_i\}|.$$

If $\boldsymbol{u}\in\mathbb{F}_q^n$ and $\boldsymbol{M}\subseteq\mathbb{F}_q^n$ is a set of vectors (or a vector subspace), then the last definition can be extended to

$$\mathsf{d}(\boldsymbol{u},\boldsymbol{M})=\min_{\boldsymbol{v}\in\boldsymbol{M}}\mathsf{d}(\boldsymbol{u},\boldsymbol{v}).$$

The *support* of a vector $\boldsymbol{u}\in\mathbb{F}_q^n$ is defined to be the set $\operatorname{supp}(\boldsymbol{u})=\{i\in[n]\ :\ u_i\neq 0\}$. The (Hamming) weight of a vector $\boldsymbol{u}$, denoted $\mathsf{wt}(\boldsymbol{u})$, is defined to be $|\operatorname{supp}(\boldsymbol{u})|$, the number of nonzero coordinates of $\boldsymbol{u}$. Suppose $E\subseteq[n]$. We write $\boldsymbol{u}\triangleleft E$ whenever $\mathsf{supp}(\boldsymbol{u})\subseteq E$.

A k-dimensional subspace $\mathcal{C}$ of $\mathbb{F}_q^n$ is called a linear $[n,k,d]_q$ code over $\mathbb{F}_q$ if the minimum distance of $\mathcal{C}$, i.e.,

$$\mathsf{d}(\mathcal{C})\overset{\triangle}{=}\min_{\boldsymbol{u}\in\mathcal{C},\ \boldsymbol{v}\in\mathcal{C},\ \boldsymbol{u}\neq\boldsymbol{v}}\mathsf{d}(\boldsymbol{u},\boldsymbol{v})$$

is equal to d. Sometimes, we may use the notation $[n,k]_q$ for the sake of simplicity. The vectors in $\mathcal{C}$ are called codewords. It is easy to see that the minimum weight of a nonzero codeword in a linear code $\mathcal{C}$ is equal to its minimum distance $\mathsf{d}(\mathcal{C})$. A *generator matrix* $\boldsymbol{G}$ of an $[n,k]_q$ code $\mathcal{C}$ is a $k\times n$ matrix whose rows are linearly independent codewords of $\mathcal{C}$. Then, $\mathcal{C}=\{\boldsymbol{y}\boldsymbol{G}\ :\ \boldsymbol{y}\in\mathbb{F}_q^k\}$. The *parity-check matrix* of $\mathcal{C}$ is an $(n-k)\times n$ matrix $\boldsymbol{H}$ over $\mathbb{F}_q$ such that $\boldsymbol{c}\in\mathcal{C}\iff\boldsymbol{H}\boldsymbol{c}^T=\boldsymbol{0}^T$. Given q, k, and d, let $N_q[k,d]$ denote the length of the shortest linear code over $\mathbb{F}_q$ which has dimension k and minimum distance d.

We use $\boldsymbol{e}_i=(\underbrace{0,\dots,0}_{i-1},1,\underbrace{0,\dots,0}_{n-i})\in\mathbb{F}_q^n$ to denote the unit vector, which has a one at the ith position, and zeros elsewhere. For a vector $\boldsymbol{y}=(y_1,y_2,\dots,y_n)$ and a subset $B=\{i_1,i_2,\dots,i_b\}$ of $[n]$, where $i_1<i_2<\cdots<i_b$, let $\boldsymbol{y}_B$ denote the vector $(y_{i_1},y_{i_2},\dots,y_{i_b})$.

For an $n\times N$ matrix $\boldsymbol{L}$, let $\boldsymbol{L}_i$ denote its ith row. For a set $E\subseteq[n]$, let $\boldsymbol{L}_E$ denote the $|E|\times N$ matrix obtained from $\boldsymbol{L}$ by deleting all the rows of $\boldsymbol{L}$ which are not indexed by the elements of E. For a set of vectors $\boldsymbol{M}$, we use notation $\mathsf{span}(\boldsymbol{M})$ to denote the linear space spanned by the vectors in $\boldsymbol{M}$. We also use notation $\mathsf{colspan}(\boldsymbol{L})$ for the linear space spanned by the columns of the matrix $\boldsymbol{L}$.

Let $\mathcal{G}=(\mathcal{V},\mathcal{E})$ be a graph with a vertex set $\mathcal{V}$ and an edge set $\mathcal{E}$. The graph is called *undirected* if every edge $e\in\mathcal{E}$, $e=\{u,v\}$, and $u,v\in\mathcal{V}$. A graph $\mathcal{G}$ is *directed* if every edge $e\in\mathcal{E}$ is an ordered pair $e=(u,v)$, $u,v\in\mathcal{V}$. A directed graph $\mathcal{G}$ is called *symmetric* if

$$(u,v)\in\mathcal{E}\qquad\iff\qquad(v,u)\in\mathcal{E}.$$

There is a natural correspondence between undirected graph $\mathcal{G}=(\mathcal{V},\mathcal{E})$ and directed symmetric graph $\mathcal{G}'=(\mathcal{V},\mathcal{E}')$ defined as

$$\mathcal{E}=\{\{u,v\}\ :\ (u,v)\in\mathcal{E}'\}\,.\qquad(2)$$

Let $\mathcal{G}$ be an undirected graph. A subset of vertices $\mathcal{S}\subseteq\mathcal{V}$ is called an *independent set* if $\forall u,v\in\mathcal{S}$, $\{u,v\}\notin\mathcal{E}$. The size of the largest independent set in $\mathcal{G}$ is called the *independence number* of $\mathcal{G}$ and is denoted by $\alpha(\mathcal{G})$. The graph $\bar{\mathcal{G}}=(\mathcal{V},\bar{\mathcal{E}})$ is called the *complement* of $\mathcal{G}=(\mathcal{V},\mathcal{E})$ if

$$\bar{\mathcal{E}}=\{\{u,v\}\ :\ u\in\mathcal{V},\ v\in\mathcal{V},\ \{u,v\}\notin\mathcal{E}\}\,.$$

A *coloring* of $\mathcal{G}$ using χ colors is a function $\psi\ :\ \mathcal{V}\to[\chi]$, such that

$$\forall e=\{u,v\}\in\mathcal{E}\ :\ \psi(u)\neq\psi(v).$$

The *chromatic number* of $\mathcal{G}$ is the smallest number χ such that there exists a coloring of $\mathcal{G}$ using χ colors, and it is denoted by $\chi(\mathcal{G})$. By using the correspondence (2), the definitions of independence number, graph complement, and chromatic number are trivially extended to directed symmetric graphs.

III. Index Coding and Error Correction

A. Index Coding With Side Information

ICSI problem considers the following communications scenario. There is a unique sender (or source) S, who has a vector of messages $\boldsymbol{x}=(x_1,x_2,\dots,x_n)$ in his possession. There are also m receivers $R_1,R_2,\dots,R_m$, receiving information from S via a broadcast channel. For each $i\in[m]$, R_i has side information, i.e., R_i owns a subset of messages $\{x_j\}_{j\in\mathcal{X}_i}$, where $\mathcal{X}_i\subseteq[n]$. Each R_i, $i\in[m]$, is interested in receiving the message $x_{f(i)}$ (we say that R_i requires $x_{f(i)}$), where the mapping $f\ :\ [m]\to[n]$ satisfies $f(i)\notin\mathcal{X}_i$ for all $i\in[m]$. Hereafter, we use the notation $\mathcal{X}=(\mathcal{X}_1,\mathcal{X}_2,\dots,\mathcal{X}_m)$. An instance of the ICSI problem is given by a quadruple $(m,n,\mathcal{X},f)$. It can also be conveniently described by a directed hypergraph [13].

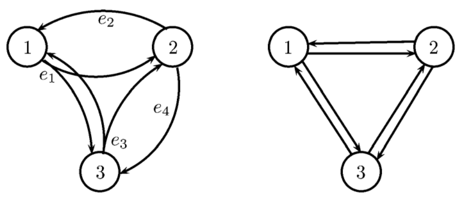
Definition 3.1: Let $(m,n,\mathcal{X},f)$ be an instance of the ICSI problem. The corresponding *side information (directed) hypergraph* $\mathcal{H}=\mathcal{H}(m,n,\mathcal{X},f)$ is defined by the vertex set $\mathcal{V}=[n]$ and the (directed) hyperedge set $\mathcal{E}_{\mathcal{H}}$, where

$$\mathcal{E}_{\mathcal{H}}=\{(f(i),\mathcal{X}_i)\ :\ i\in[m]\}.$$

Each hyperedge represents the side information and the demand from one receiver. We often refer to $(m,n,\mathcal{X},f)$ as an instance of the ICSI problem described by the hypergraph $\mathcal{H}$.

For instance, consider an ICSI instance where $n=3$ (three messages), $m=4$ (four receivers), $f(1)=1$, $f(2)=2$, $f(3)=3$, $f(4)=2$, $\mathcal{X}_1=\{2,3\}$, $\mathcal{X}_2=\{1\}$, $\mathcal{X}_3=\{1,2\}$, and $\mathcal{X}_4=\{3\}$. The hypergraph $\mathcal{H}_1$ that describes this instance has three vertices 1, 2, 3, and has four directed hyperedges. These are $e_1=(1,\{2,3\})$, $e_2=(2,1)$, $e_3=(3,\{1,2\})$, and $e_4=(2,3)$. This hypergraph is depicted in Fig. 2(a).

Each side information hypergraph $\mathcal{H}=(\mathcal{V},\mathcal{E}_{\mathcal{H}})$ can be associated with the directed graph $\mathcal{G}_{\mathcal{H}}=(\mathcal{V},\mathcal{E})$ in the following way. For each directed edge $(f(i),\mathcal{X}_i)\in\mathcal{E}_{\mathcal{H}}$, there will be $|\mathcal{X}_i|$ directed edges $(f(i),v)\in\mathcal{E}$, for $v\in\mathcal{X}_i$. For instance, $\mathcal{G}_{\mathcal{H}_1}$ is depicted in Fig. 2(b). When $m=n$ and $f(i)=i$ for all $i\in[m]$, the graph $\mathcal{G}_{\mathcal{H}}$ is, in fact, the *side information graph*, defined in [6].



(a) (b)

Fig. 2. (a) Hypergraph $\mathcal{H}_1$ and (b) its corresponding directed graph $\mathcal{G}_{\mathcal{H}_1}$.

The goal of the ICSI problem is to design a coding scheme that allows S to satisfy the requests of all receivers R_i in the least number of transmissions. More formally, we have the following definition.

Definition 3.2: An *index code* over $\mathbb{F}_q$ for an instance of the ICSI problem described by $\mathcal{H} = \mathcal{H}(m, n, \mathcal{X}, f)$ (or just an $\mathcal{H}$-IC over $\mathbb{F}_q$) is an encoding function

$$\mathfrak{E} : \mathbb{F}_q^n \to \mathbb{F}_q^N$$

such that for each receiver $R_i,\ i \in [m]$, there exists a decoding function

$$\mathfrak{D}_i : \mathbb{F}_q^N \times \mathbb{F}_q^{|\mathcal{X}_i|} \to \mathbb{F}_q$$

satisfying

$$\forall \boldsymbol{x} \in \mathbb{F}_q^n \ : \ \mathfrak{D}_i(\mathfrak{E}(\boldsymbol{x}), \boldsymbol{x}_{\mathcal{X}_i}) = x_{f(i)}.$$

Sometimes we refer to such $\mathfrak{E}$ as a *non-error-correcting index code*. The parameter N is called the *length* of the index code. In the scheme corresponding to this code, S broadcasts a vector $\mathfrak{E}(\boldsymbol{x})$ of length N over $\mathbb{F}_q$.

Definition 3.3: A *linear index code* is an index code, for which the encoding function $\mathfrak{E}$ is a linear transformation over $\mathbb{F}_q$. Such a code can be described as

$$\forall \boldsymbol{x} \in \mathbb{F}_q^n \ : \ \mathfrak{E}(\boldsymbol{x}) = \boldsymbol{x}\boldsymbol{L}$$

where $\boldsymbol{L}$ is an $n \times N$ matrix over $\mathbb{F}_q$. The matrix $\boldsymbol{L}$ is called the *matrix corresponding to the index code* $\mathfrak{E}$. The code $\mathfrak{E}$ is also referred to as the *linear index code based on* $\boldsymbol{L}$.

Hereafter, we assume that $\mathcal{X} = (\mathcal{X}_i)_{i \in [m]}$ is known to S. Moreover, we also assume that the code $\mathfrak{E}$ is known to each receiver $R_i,\ i \in [m]$. In practice, this can be achieved by a preliminary communication session, when the knowledge of the sets $\mathcal{X}_i$ for $i \in [m]$ and of the code $\mathfrak{E}$ is disseminated between the participants of the scheme.

Definition 3.4: Suppose $\mathcal{H} = \mathcal{H}(m, n, \mathcal{X}, f)$ corresponds to an instance of the ICSI problem. Then, the *min-rank* of $\mathcal{H}$ over $\mathbb{F}_q$ is defined as

$$\kappa_q(\mathcal{H}) \overset{\triangle}{=} \min\left\{\mathsf{rank}_{\mathbb{F}_q}(\{\boldsymbol{v}_i + \boldsymbol{e}_{f(i)}\}_{i \in [m]}) \ : \right. \\ \left. \boldsymbol{v}_i \in \mathbb{F}_q^n,\ \boldsymbol{v}_i \lhd \mathcal{X}_i\right\}.$$

For example, it is straightforward to verify that $\kappa_2(\mathcal{H}_1) = 2$, where $\mathcal{H}_1$ is depicted in Fig. 2(a). We may select $\boldsymbol{v}_1 = (0, 1, 1)$, $\boldsymbol{v}_2 = (0, 0, 0)$, $\boldsymbol{v}_3 = (1, 1, 0)$, and $\boldsymbol{v}_4 = (0, 0, 0)$. Then

$$\begin{aligned}&\mathsf{rank}_{\mathbb{F}_2}(\{\boldsymbol{v}_1 + \boldsymbol{e}_1, \boldsymbol{v}_2 + \boldsymbol{e}_2, \boldsymbol{v}_3 + \boldsymbol{e}_3, \boldsymbol{v}_4 + \boldsymbol{e}_2\}) \\ &\quad = \mathsf{rank}_{\mathbb{F}_2}\{(1, 1, 1), (0, 1, 0), (1, 1, 1), (0, 1, 0))\} \\ &\quad = 2.\end{aligned}$$

Observe that $\kappa_q(\mathcal{H})$ generalizes the $\mathrm{min\text{-}rank}$ over $\mathbb{F}_q$ of the side information graph, which was defined in [6]. More specifically, when $m = n$ and $f(i) = i$ for all $i \in [m]$, $\mathcal{G}_{\mathcal{H}}$ becomes the side information graph, and $\kappa_q(\mathcal{H}) = \mathrm{min\text{-}rank}_q(\mathcal{G}_{\mathcal{H}})$. The $\mathrm{min\text{-}rank}$ of an undirected graph was first introduced by Haemers [16] to bound the Shannon capacity of a graph, and was later proved in [6] and [7] to be the smallest number of transmissions in a linear index code. It was shown by Peeters [17] that finding the min-rank of a general graph is an NP-hard problem. Studies on graph parameters related to min-ranks can also be found in the works of Peeters [18], [19].

The intuition behind the concept of min-rank is explained as follows. For each $i \in [m]$, if R_i obtains $\boldsymbol{x}(\boldsymbol{v}_i + \boldsymbol{e}_{f(i)})^T$ where $\boldsymbol{v}_i \lhd \mathcal{X}_i$, then R_i is able to determine $x_{f(i)}$. Indeed, as R_i possesses x_j for every $j \in \mathcal{X}_i$, he can calculate $\boldsymbol{x}\boldsymbol{v}_i^T$ if $\boldsymbol{v}_i \lhd \mathcal{X}_i$. Hence, R_i can retrieve $x_{f(i)}$ as follows:

$$x_{f(i)} = \boldsymbol{x}\boldsymbol{e}_{f(i)}^T = \boldsymbol{x}(\boldsymbol{v}_i + \boldsymbol{e}_{f(i)})^T - \boldsymbol{x}\boldsymbol{v}_i^T.$$

Thus, in order to satisfy all the demands, the sender S may broadcast m packets $\boldsymbol{x}(\boldsymbol{v}_i + \boldsymbol{e}_{f(i)})^T$, where $\boldsymbol{v}_i \lhd \mathcal{X}_i,\ i \in [m]$. In fact, it is sufficient to broadcast only

$$\mathsf{rank}_{\mathbb{F}_q}\left(\{\boldsymbol{v}_i + \boldsymbol{e}_{f(i)}\}_{i \in [m]}\right)$$

packets. Therefore, the minimum number of packets (transmissions) required in this way is $\kappa_q(\mathcal{H})$. It turns out that $\kappa_q(\mathcal{H})$ is the smallest possible number of transmissions required if scalar linear index codes are used, according to Lemma 3.5. This lemma was implicitly formulated in [6] for the case where $m = n$, $q = 2$, $f(i) = i$ for all $i \in [n]$, and generalized to its current form in [20].

Lemma 3.5: Consider an instance of the ICSI problem described by $\mathcal{H} = \mathcal{H}(m, n, \mathcal{X}, f)$.

1) The matrix $\boldsymbol{L}$ corresponds to a linear $\mathcal{H}$-IC over $\mathbb{F}_q$ if and only if for each $i \in [m]$ there exists $\boldsymbol{v}_i \in \mathbb{F}_q^n$ such that
   - $\boldsymbol{v}_i \lhd \mathcal{X}_i$;
   - $\boldsymbol{v}_i + \boldsymbol{e}_{f(i)} \in \mathsf{colspan}(\boldsymbol{L})$.
2) The smallest possible length of a linear $\mathcal{H}$-IC over $\mathbb{F}_q$ is $\kappa_q(\mathcal{H})$.

B. Error-Correcting Index Code With Side Information

Due to noise, the symbols received by $R_i,\ i \in [m]$, may be subject to errors. Consider an ICSI instance $(m, n, \mathcal{X}, f)$, and assume that S broadcasts a vector $\mathfrak{E}(\boldsymbol{x}) \in \mathbb{F}_q^N$. Let $\boldsymbol{\epsilon}_i \in \mathbb{F}_q^N$ be the error affecting the information received by $R_i,\ i \in [m]$. Then, R_i actually receives the vector

$$\boldsymbol{y}_i = \mathfrak{E}(\boldsymbol{x}) + \boldsymbol{\epsilon}_i \in \mathbb{F}_q^N$$

instead of $\mathfrak{E}(\boldsymbol{x})$. The following definition is a generalization of Definition 3.2.

Definition 3.6: Consider an instance of the ICSI problem described by $\mathcal{H} = \mathcal{H}(m, n, \mathcal{X}, f)$. A *$\delta$-error-correcting index code* ($(\delta, \mathcal{H})$-ECIC) over $\mathbb{F}_q$ for this instance is an encoding function

$$\mathfrak{E} : \mathbb{F}_q^n \rightarrow \mathbb{F}_q^N$$

such that for each receiver $R_i,\ i \in [m]$, there exists a decoding function

$$\mathfrak{D}_i : \mathbb{F}_q^N \times \mathbb{F}_q^{|\mathcal{X}_i|} \rightarrow \mathbb{F}_q$$

satisfying

$$\forall \boldsymbol{x}, \boldsymbol{\epsilon}_i \in \mathbb{F}_q^n,\ \mathsf{wt}(\boldsymbol{\epsilon}_i) \leqslant \delta\ :\ \mathfrak{D}_i(\mathfrak{E}(\boldsymbol{x}) + \boldsymbol{\epsilon}_i, \boldsymbol{x}_{\mathcal{X}_i}) = x_{f(i)}.$$

The definitions of the length, of a linear index code, and of the matrix corresponding to an index code are naturally extended to an ECIC. Note that if $\mathfrak{E}$ is an $\mathcal{H}$-IC, then it is a $(0, \mathcal{H})$-ECIC, and vice versa.

Definition 3.7: An *optimal* linear $(\delta, \mathcal{H})$-ECIC over $\mathbb{F}_q$ is a linear $(\delta, \mathcal{H})$-ECIC over $\mathbb{F}_q$ of the smallest possible length $\mathcal{N}_q[\mathcal{H}, \delta]$.

Consider an instance of the ICSI problem described by $\mathcal{H} = \mathcal{H}(m, n, \mathcal{X}, f)$. We define the set of vectors

$$\mathcal{I}(q, \mathcal{H}) \triangleq \left\{\boldsymbol{z} \in \mathbb{F}_q^n\ :\ \exists i \in [m] \text{ such that } \boldsymbol{z}_{\mathcal{X}_i} = \boldsymbol{0} \text{ and } z_{f(i)} \neq 0\right\}.$$

For all $i \in [m]$, we also define

$$\mathcal{Y}_i \triangleq [n] \backslash \Big(\{f(i)\} \cup \mathcal{X}_i\Big).$$

Then, the collection of supports of all vectors in $\mathcal{I}(q, \mathcal{H})$ is given by

$$\mathcal{J}(\mathcal{H}) \triangleq \bigcup_{i \in [m]} \Big\{\{f(i)\} \cup Y_i\ :\ Y_i \subseteq \mathcal{Y}_i\Big\}. \qquad (3)$$

The necessary and sufficient condition for a matrix $\boldsymbol{L}$ to be the matrix corresponding to some $(\delta, \mathcal{H})$-ECIC is given in the following lemma.

Lemma 3.8: The matrix $\boldsymbol{L}$ corresponds to a $(\delta, \mathcal{H})$-ECIC over $\mathbb{F}_q$ if and only if

$$\mathsf{wt}\Big(\boldsymbol{z}\boldsymbol{L}\Big) \geqslant 2\delta + 1 \quad \text{for all } \boldsymbol{z} \in \mathcal{I}(q, \mathcal{H}). \qquad (4)$$

Equivalently, $\boldsymbol{L}$ corresponds to a $(\delta, \mathcal{H})$-ECIC over $\mathbb{F}_q$ if and only if

$$\mathsf{wt}\left(\sum_{i \in K} z_i \boldsymbol{L}_i\right) \geqslant 2\delta + 1 \qquad (5)$$

for all $K \in \mathcal{J}(\mathcal{H})$ and for all choices of $z_i \in \mathbb{F}_q^*,\ i \in K$.

Proof: For each $\boldsymbol{x} \in \mathbb{F}_q^n$, we define

$$B(\boldsymbol{x}, \delta) = \{\boldsymbol{y} \in \mathbb{F}_q^N\ :\ \boldsymbol{y} = \boldsymbol{x}\boldsymbol{L} + \boldsymbol{\epsilon},\ \boldsymbol{\epsilon} \in \mathbb{F}_q^N,\ \mathsf{wt}(\boldsymbol{\epsilon}) \leqslant \delta\}$$

the set of all vectors resulting from at most δ errors in the transmitted vector associated with the information vector $\boldsymbol{x}$. Then, the receiver R_i can recover $x_{f(i)}$ correctly if and only if

$$B(\boldsymbol{x}, \delta) \cap B(\boldsymbol{x}', \delta) = \varnothing$$

for every pair $\boldsymbol{x}, \boldsymbol{x}' \in \mathbb{F}_q^n$ satisfying

$$\boldsymbol{x}_{\mathcal{X}_i} = \boldsymbol{x}'_{\mathcal{X}_i} \text{ and } x_{f(i)} \neq x'_{f(i)}.$$

(Observe that R_i is interested only in the bit $x_{f(i)}$, not in the whole vector $\boldsymbol{x}$.)

Therefore, $\boldsymbol{L}$ corresponds to a $(\delta, \mathcal{H})$-ECIC if and only if the following condition is satisfied: for all $i \in [m]$ and for all $\boldsymbol{x}, \boldsymbol{x}' \in \mathbb{F}_q^n$ such that $\boldsymbol{x}_{\mathcal{X}_i} = \boldsymbol{x}'_{\mathcal{X}_i}$ and $x_{f(i)} \neq x'_{f(i)}$, it holds

$$\forall \boldsymbol{\epsilon}, \boldsymbol{\epsilon}' \in \mathbb{F}_q^N,\ \mathsf{wt}(\boldsymbol{\epsilon}) \leqslant \delta,\ \mathsf{wt}(\boldsymbol{\epsilon}') \leqslant \delta\ :\ \boldsymbol{x}\boldsymbol{L} + \boldsymbol{\epsilon} \neq \boldsymbol{x}'\boldsymbol{L} + \boldsymbol{\epsilon}'. \qquad (6)$$

Denote $\boldsymbol{z} = \boldsymbol{x}' - \boldsymbol{x}$. Then, the condition in (6) can be reformulated as follows: for all $i \in [n]$ and for all $\boldsymbol{z} \in \mathbb{F}_q^n$ such that $\boldsymbol{z}_{\mathcal{X}_i} = \boldsymbol{0}$ and $z_{f(i)} \neq 0$, it holds

$$\forall \boldsymbol{\epsilon}, \boldsymbol{\epsilon}' \in \mathbb{F}_q^N,\ \mathsf{wt}(\boldsymbol{\epsilon}) \leqslant \delta,\ \mathsf{wt}(\boldsymbol{\epsilon}') \leqslant \delta\ :\ \boldsymbol{z}\boldsymbol{L} \neq \boldsymbol{\epsilon} - \boldsymbol{\epsilon}'. \qquad (7)$$

Note that the two sets

$$\{\boldsymbol{\epsilon} - \boldsymbol{\epsilon}' : \boldsymbol{\epsilon}, \boldsymbol{\epsilon}' \in \mathbb{F}_q^N,\ \mathsf{wt}(\boldsymbol{\epsilon}) \leqslant \delta,\ \mathsf{wt}(\boldsymbol{\epsilon}') \leqslant \delta\}$$

and

$$\{\boldsymbol{\epsilon}'' \in \mathbb{F}_q^N : \mathsf{wt}(\boldsymbol{\epsilon}'') \leqslant 2\delta\}$$

coincide. Therefore, an equivalent condition to (7) is that for all $\boldsymbol{z} \in \mathcal{I}(q, \mathcal{H})$,

$$\mathsf{wt}(\boldsymbol{z}\boldsymbol{L}) \geqslant 2\delta + 1.$$

Since for $\boldsymbol{z} \in \mathcal{I}(q, \mathcal{H})$, we have

$$\boldsymbol{z}\boldsymbol{L} = \sum_{i \in \mathsf{supp}(\boldsymbol{z})} z_i \boldsymbol{L}_i$$

the condition (4) can be restated as

$$\mathsf{wt}\left(\sum_{i \in K} z_i \boldsymbol{L}_i\right) \geqslant 2\delta + 1$$

for all $K \in \mathcal{J}(\mathcal{H})$ and for all choices of nonzero $z_i \in \mathbb{F}_q, i \in K$. ■

Corollary 3.9: For all $i \in [m]$, let

$$\boldsymbol{M}_i \triangleq \mathrm{span}\left(\{\boldsymbol{L}_j\ :\ j \in \mathcal{Y}_i\}\right).$$

Then, the matrix $\boldsymbol{L}$ corresponds to a $(\delta, \mathcal{H})$-ECIC over $\mathbb{F}_q$ if and only if

$$\forall i \in [m] \ : \ \mathsf{d}(\boldsymbol{L}_{f(i)}, \boldsymbol{M}_i) \geqslant 2\delta + 1. \tag{8}$$

Proof: It suffices to show that the conditions (5) and (8) are equivalent. First, to avoid confusion, we rewrite (5) as follows:

$$\mathsf{wt}\left(\sum_{j \in K} z_j \boldsymbol{L}_j\right) \geqslant 2\delta + 1 \tag{9}$$

for all $K \in \mathcal{J}(\mathcal{H})$ and for all choices of $z_j \in \mathbb{F}_q^*$, $j \in K$. By definition of $\mathcal{J}(\mathcal{H})$ (see (3)), the condition (9) is equivalent to the condition that

$$\mathsf{wt}\left(z_{f(i)} \boldsymbol{L}_{f(i)} + \sum_{j \in Y_i} z_j \boldsymbol{L}_j\right) \geqslant 2\delta + 1$$

for all $Y_i \subseteq \mathcal{Y}_i$ and for all choices of $z_{f(i)} \in \mathbb{F}_q^*$ and $z_j \in \mathbb{F}_q^*$, $j \in Y_i$. We can also rewrite this condition as

$$\mathsf{d}\left(\boldsymbol{L}_{f(i)}, \sum_{j \in Y_i} z_j \boldsymbol{L}_j\right) \geqslant 2\delta + 1 \tag{10}$$

for all $Y_i \subseteq \mathcal{Y}_i$ and for all choices of $z_j \in \mathbb{F}_q^*$, $j \in Y_i$. The conditions (10) and (8) are obviously equivalent. ■

The next corollary also follows directly from Lemma 3.8 by considering an error-free setup, i.e., $\delta = 0$. It is easy to verify that the conditions stated in this corollary and in Lemma 3.5 are equivalent, as expected.

Corollary 3.10: The matrix $\boldsymbol{L}$ corresponds to an $\mathcal{H}$-IC over $\mathbb{F}_q$ if and only if

$$\mathsf{wt}\left(\sum_{i \in K} z_i \boldsymbol{L}_i\right) \geqslant 1$$

for all $K \in \mathcal{J}(\mathcal{H})$ and for all choices of $z_i \in \mathbb{F}_q^*$, $i \in K$, or, equivalently,

$$\forall i \in [m] \ : \ \boldsymbol{L}_{f(i)} \notin \mathsf{span}(\{\boldsymbol{L}_j\}_{j \in \mathcal{Y}_i}).$$

Example 3.11: Let $q = 2$, $m = n = 3$, and $f(i) = i$ for $i \in [3]$. Suppose $\mathcal{X}_1 = \{2, 3\}$, $\mathcal{X}_2 = \{1, 3\}$, and $\mathcal{X}_3 = \{1, 2\}$. Let

$$\boldsymbol{L} = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \end{pmatrix}.$$

Note that $\boldsymbol{L}$ generates a $[4, 3, 1]_2$ code, which has minimum distance one. However, the index code based on $\boldsymbol{L}$ can still correct one error. Indeed, let $\mathcal{H} = \mathcal{H}(3, 3, \mathcal{X}, f)$; we have

$$\mathcal{I}(2, \mathcal{H}) = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}.$$

Since each row of $\boldsymbol{L}$ has weight at least three, it follows that $\mathsf{wt}(\boldsymbol{z}\boldsymbol{L}) \geqslant 3$ for all $\boldsymbol{z} \in \mathcal{I}(2, \mathcal{H})$. By Lemma 3.8, $\boldsymbol{L}$ corresponds to a $(1, \mathcal{H})$-ECIC over $\mathbb{F}_2$.

In fact, for this instance, even a simpler index code of length three, based on

$$\boldsymbol{L}' = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix},$$

is a $(1, \mathcal{H})$-ECIC over $\mathbb{F}_2$.

Example 3.12: Assume that $m = n$ and $f(i) = i$ for all $i \in [m]$. Furthermore, suppose that $\mathcal{X}_i = \varnothing$ for all $i \in [m]$ (i.e., there is no side information available to the receivers). Let $\mathcal{H} = \mathcal{H}(m, n, \mathcal{X}, f)$. Then, $\mathcal{I}(q, \mathcal{H}) = \mathbb{F}_q^n \backslash \{\boldsymbol{0}\}$. Hence, by Lemma 3.8, the $n \times N$ matrix $\boldsymbol{L}$ corresponding to a $(\delta, \mathcal{H})$-ECIC over $\mathbb{F}_q$ (for some integer $\delta \geqslant 0$) is a generating matrix of an $[N, n, \geqslant 2\delta + 1]_q$ linear code. Thus, under these settings, the problem of designing an optimal ECIC is reduced to the problem of constructing an optimal classical linear error-correcting code.

Observe, however, that for general $\mathcal{X}$, changing the order of rows in $\boldsymbol{L}$ can lead to ECICs with different error-correcting capabilities, according to Corollary 3.9. Therefore, the problem of designing an optimal linear ECIC is essentially the problem of finding the matrix $\boldsymbol{L}$ corresponding to that code. However, the minimum distance of the code generated by the rows of $\boldsymbol{L}$ is not necessarily a valid indicator for goodness of an ECIC. Sometimes, as Example 3.11 shows, matrix $\boldsymbol{L}$ with redundant rows yields a good ECIC.

IV. α-Bound and the κ-Bound

Let $(m, n, \mathcal{X}, f)$ be an instance of the ICSI problem, and let $\mathcal{H}$ be the corresponding side information hypergraph. Next, we introduce the following definitions for the hypergraph $\mathcal{H}$.

Definition 4.1: A subset H of $[n]$ is called a *generalized independent set* in $\mathcal{H}$ if every nonempty subset K of H belongs to $\mathcal{J}(\mathcal{H})$.

Definition 4.2: A generalized independent set of the largest size in $\mathcal{H}$ is called a *maximum generalized independent set*. The size of a maximum generalized independent set in $\mathcal{H}$ is called the *generalized independence number*, and denoted by $\alpha(\mathcal{H})$.

When $m = n$ and $f(i) = i$ for all $i \in [n]$, the generalized independence number of $\mathcal{H}$ is equal to the maximum size of an acyclic induced subgraph of $\mathcal{G}_{\mathcal{H}}$, which was introduced in [6]. In particular, when $\mathcal{G}_{\mathcal{H}}$ is symmetric, $\alpha(\mathcal{H})$ is the independence number of $\mathcal{G}_{\mathcal{H}}$. We prove the latter statement in the Appendix.

Next, we present a lower bound on the length of a $(\delta, \mathcal{H})$-ECIC. We call this bound *α-bound*.

Theorem 4.3 (α-Bound): The length of an optimal linear $(\delta, \mathcal{H})$-ECIC over $\mathbb{F}_q$ satisfies

$$\mathcal{N}_q[\mathcal{H}, \delta] \geqslant N_q[\alpha(\mathcal{H}), 2\delta + 1].$$

Moreover, the equality is attained if there exists an $n \times \alpha(\mathcal{H})$ matrix $\boldsymbol{B} = (b_{i,j})$ over $\mathbb{F}_q$ satisfying the following condition: for all $K \in \mathcal{J}(\mathcal{H})$ and for all choices of $z_i \in \mathbb{F}_q^*$, $i \in K$, there always exists some j such that

$$\sum_{i \in K} z_i \, b_{i,j} \neq 0.$$

Proof: Consider an $n \times N$ matrix $\boldsymbol{L}$, which corresponds to a $(\delta, \mathcal{H})$-ECIC. Let $H = \{i_1, i_2, \dots, i_{\alpha(\mathcal{H})}\}$ be a maximum generalized independent set in $\mathcal{H}$. Then, every subset $K \subseteq H$ satisfies $K \in \mathcal{J}(\mathcal{H})$. Therefore

$$\mathsf{wt}\left(\sum_{i \in K} z_i \boldsymbol{L}_i\right) \geqslant 2\delta + 1$$

for all $K \subseteq H$, $K \neq \varnothing$, and for all choices of $z_i \in \mathbb{F}_q^*$, $i \in K$. Hence, the $\alpha(\mathcal{H})$ rows of $\boldsymbol{L}$, namely $\boldsymbol{L}_{i_1}, \boldsymbol{L}_{i_2}, \dots, \boldsymbol{L}_{i_{\alpha(\mathcal{H})}}$, form a generator matrix of an $[N, \alpha(\mathcal{H}), 2\delta + 1]_q$ code. Therefore

$$N \geqslant N_q[\alpha(\mathcal{H}), 2\delta + 1].$$

Next, we assume the existence of a matrix $\boldsymbol{B}$ satisfying the properties stated in the theorem. Let $\boldsymbol{L}'$ be a generator matrix of some $[N', \alpha(\mathcal{H}), 2\delta+1]_q$ code, where $N' = N_q[\alpha(\mathcal{H}), 2\delta + 1]$. We construct the $n \times N'$ matrix $\boldsymbol{L}$ as follows. For $i \in [n]$, let

$$\boldsymbol{L}_i = \sum_{j=1}^{\alpha(\mathcal{H})} b_{i,j} \boldsymbol{L}'_j.$$

For every $K \in \mathcal{J}(\mathcal{H})$ and for all choices of $z_i \in \mathbb{F}_q^*$, $i \in K$, we have

$$\begin{aligned}\mathsf{wt}\left(\sum_{i \in K} z_i \boldsymbol{L}_i\right) &= \mathsf{wt}\left(\sum_{i \in K} z_i \sum_{j=1}^{\alpha(\mathcal{H})} b_{i,j} \boldsymbol{L}'_j\right)\\ &= \mathsf{wt}\left(\sum_{j=1}^{\alpha(\mathcal{H})} \left(\sum_{i \in K} z_i\, b_{i,j}\right) \boldsymbol{L}'_j\right)\\ &\geqslant 2\delta + 1\end{aligned}$$

where the last transition is due to the existence of $j \in [\alpha(\mathcal{H})]$ such that

$$\sum_{i \in K} z_i\, b_{i,j} \neq 0$$

and the fact that $\boldsymbol{L}'_j$'s are linearly independent nonzero codewords of a code of minimum distance $2\delta + 1$.

We conclude that the index code based on $\boldsymbol{L}$ is capable of correcting δ errors. Therefore, $\mathcal{N}_q[\mathcal{H}, \delta] = N_q[\alpha(\mathcal{H}), 2\delta + 1]$. ■

Example 4.4: Let $q = 2$, $m = n = 5$, $f(i) = i$ for all $i \in [m]$, and $\delta = 2$. Assume

$$\begin{aligned}\mathcal{X}_1 = \{2,3,4\}, \quad \mathcal{X}_2 = \{3,4,5\}, \quad \mathcal{X}_3 = \{4,5,1\},\\ \mathcal{X}_4 = \{5,1,2\}, \quad \mathcal{X}_5 = \{1,2,3\}.\end{aligned}$$

Let $\mathcal{H} = \mathcal{H}(5, 5, \mathcal{X}, f)$. Then

$$\begin{aligned}\mathcal{J}(\mathcal{H}) = \Big\{&\{1\}, \{1,5\}, \{2\}, \{2,1\}, \{3\},\\ &\{3,2\}, \{4\}, \{4,3\}, \{5\}, \{5,4\}\Big\}.\end{aligned}$$

It is easy to check that $\alpha(\mathcal{H}) = 2$. Therefore, Theorem 4.3 implies that

$$\mathcal{N}_2[\mathcal{H}, 2] \geqslant N_2[2, 5] = 8.$$

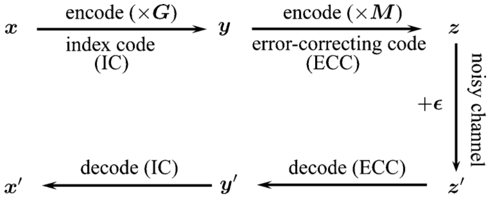


Fig. 3. Concatenation of an error-correcting code and an index code.

The last equality can be verified by [21].

On the other hand, take the matrix

$$\boldsymbol{B} \triangleq \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}.$$

The matrix $\boldsymbol{B}$ satisfies the property that for all $K \in \mathcal{J}(\mathcal{H})$, $K \neq \varnothing$, there exists $j \in [2]$ such that

$$\sum_{i \in K} b_{i,j} \neq 0.$$

From Theorem 4.3, we have $\mathcal{N}_2[\mathcal{H}, 2] = N_2[2, 5] = 8$.

Remark 4.5: In [6], when $m = n$ and $f(i) = i$ for all $i \in [n]$, $\alpha(\mathcal{H})$ was shown to be a lower bound on the length of a (non-error-correcting) linear index code. However, the α-bound in Theorem 4.3 does not follow from the results in [6]. The reason is that a concatenation of an optimal linear error-correcting code with an optimal non-error-correcting index code might fail to produce an optimal linear error-correcting index code. This is illustrated later in Example 4.8.

The following proposition is based on the fact that concatenation of a δ-error-correcting code with an optimal (non-error-correcting) $\mathcal{H}$-IC yields a $(\delta, \mathcal{H})$-ECIC (see Fig. 3).

Proposition 4.6 (κ-Bound): The length of an optimal $(\delta, \mathcal{H})$-ECIC over $\mathbb{F}_q$ satisfies

$$\mathcal{N}_q[\mathcal{H}, \delta] \leqslant N_q[\kappa_q(\mathcal{H}), 2\delta + 1].$$

Proof: Let $\boldsymbol{G}$, which is an $n \times \kappa_q(\mathcal{H})$ matrix, correspond to an optimal $\mathcal{H}$-IC over $\mathbb{F}_q$. Denote

$$\boldsymbol{y} = \boldsymbol{x}\boldsymbol{G} \in \mathbb{F}_q^{\kappa_q(\mathcal{H})}.$$

Let $\boldsymbol{M}$ be a generator matrix of an optimal $[N, \kappa_q(\mathcal{H}), 2\delta + 1]_q$ code $\mathcal{C}'$, where

$$N = N_q[\kappa_q(\mathcal{H}), 2\delta + 1].$$

Consider a scheme where S broadcasts the vector $\boldsymbol{y}\boldsymbol{M} \in \mathbb{F}_q^N$. If less than δ errors occur, then each receiver R_i is able to recover $\boldsymbol{y}$ by using $\mathcal{C}'$. Hence, each R_i is able to recover $x_{f(i)}$. Therefore, for the index code based on $\boldsymbol{L}$,

$$\boldsymbol{L} = \boldsymbol{G}\boldsymbol{M},$$

each receiver R_i is capable to recover $x_{f(i)}$ if the number of errors is less or equal to δ. The length of the corresponding ECIC is $N = N_q[\kappa_q(\mathcal{H}), 2\delta + 1]$. Therefore

$$\mathcal{N}_q[\mathcal{H}, \delta] \leqslant N_q[\kappa_q(\mathcal{H}), 2\delta + 1]. \qquad \blacksquare$$

By combining the results in Theorem 4.3 and in Proposition 4.6, we obtain the following corollary.

Corollary 4.7: The length of an optimal linear $(\delta, \mathcal{H})$-ECIC over $\mathbb{F}_q$ satisfies

$$N_q[\alpha(\mathcal{H}), 2\delta + 1] \leqslant \mathcal{N}_q[\mathcal{H}, \delta] \leqslant N_q[\kappa_q(\mathcal{H}), 2\delta + 1].$$

It is shown in the example below that the inequalities in Corollary 4.7 can be strict. In particular, it follows that mere application of an error-correcting code on top of an index code may fail to provide us with an optimal linear ECIC. This fact motivates the study of ECICs in Sections III–VII.

Example 4.8: Let $q = 2$, $m = n = 5$, $\delta = 2$, and $f(i) = i$ for all $i \in [m]$. Assume

$$\mathcal{X}_1 = \{2, 5\}, \quad \mathcal{X}_2 = \{1, 3\}, \quad \mathcal{X}_3 = \{2, 4\},$$
$$\mathcal{X}_4 = \{3, 5\}, \quad \mathcal{X}_5 = \{1, 4\}.$$

Let $\mathcal{H} = \mathcal{H}(5, 5, \mathcal{X}, f)$. Then, we have

$$\begin{aligned}\mathcal{J}(\mathcal{H}) = \Big\{&\{1\}, \{1,3\}, \{1,4\}, \{1,3,4\},\\ &\{2\}, \{2,4\}, \{2,5\}, \{2,4,5\},\\ &\{3\}, \{1,3\}, \{3,5\}, \{1,3,5\},\\ &\{4\}, \{1,4\}, \{2,4\}, \{1,2,4\},\\ &\{5\}, \{2,5\}, \{3,5\}, \{2,3,5\}\Big\}.\end{aligned}$$

The side information graph $\mathcal{G}_\mathcal{H}$ of this instance is symmetric and can be regarded as a (undirected) pentagon (Fig. 4). It is easy to verify that $\alpha(\mathcal{H}) = \alpha(\mathcal{G}_\mathcal{H}) = 2$. It follows from [7, Th. 9] that $\kappa_2(\mathcal{H}) = \text{min-rank}_2(\mathcal{G}_\mathcal{H}) = 3$. Thus, from [21], we have

$$N_2[2, 5] = 8 \quad \text{and} \quad N_2[3, 5] = 10.$$

Due to Corollary 4.7, we have

$$8 \leqslant \mathcal{N}_2[\mathcal{H}, 2] \leqslant 10.$$

Using a computer search, we obtain that $\mathcal{N}_2[\mathcal{H}, 2] = 9$, and the corresponding optimal scheme is based on

$$\boldsymbol{L} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 \end{pmatrix}.$$

It is technical to verify that for all $K \in \mathcal{J}(\mathcal{H})$,

$$\text{wt}\left(\sum_{i \in K} \boldsymbol{L}_i\right) \geqslant 5.$$

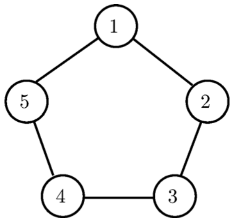


Fig. 4. Side information graph $\mathcal{G}_\mathcal{H}$.

Therefore, by Lemma 3.8, for the index code based on $\boldsymbol{L}$, each receiver R_i is able to recover x_i, if the number of errors is less than or equal to 2. Observe that the length of the ECIC corresponding to $\boldsymbol{L}$ lies strictly between the α-bound and the κ-bound.

When the graph $\mathcal{G}$ is undirected (or symmetric), the following theorem holds (see, for instance, [16]).

Theorem 4.9: Let $\chi(\bar{\mathcal{G}})$ denote the chromatic number of the complement of the graph $\mathcal{G}$. Then

$$\alpha(\mathcal{G}) \leqslant \text{min-rank}_q(\mathcal{G}) \leqslant \chi(\bar{\mathcal{G}}).$$

When $m = n$ and $f(i) = i$ for all $i \in [m]$, we have that $\alpha(\mathcal{H}) = \alpha(\mathcal{G}_\mathcal{H})$ and $\kappa_q(\mathcal{H}) = \text{min-rank}_q(\mathcal{G}_\mathcal{H})$. Moreover, if the graph $\mathcal{G}_\mathcal{H}$ is symmetric and satisfies $\alpha(\mathcal{G}_\mathcal{H}) = \chi(\bar{\mathcal{G}}_\mathcal{H})$, then from Corollary 4.7, we have

$$\mathcal{N}_q[\mathcal{H}, \delta] = N_q[\alpha(\mathcal{H}), 2\delta + 1] = N_q[\kappa_q(\mathcal{H}), 2\delta + 1]$$

for all q, and the corresponding bounds in Corollary 4.7 are tight.

Definition 4.10: An undirected (or symmetric) graph $\mathcal{G}$ is called perfect if for every induced subgraph $\mathcal{G}'$ of $\mathcal{G}$, $\alpha(\mathcal{G}') = \chi(\bar{\mathcal{G}}')$.

Perfect graphs include families of graphs such as trees, bipartite graphs, interval graphs, and chordal graphs. If $m = n$, $f(i) = i$ for all $i \in [m]$, and $\mathcal{G}_\mathcal{H}$ is perfect, then the bounds in Corollary 4.7 are tight. For the full characterization of perfect graphs, the reader can refer to [22].

V. Singleton Bound

The following bound is analogous to the Singleton bound for classical linear error-correcting codes.

Theorem 5.1 (Singleton Bound): The length of an optimal linear $(\delta, \mathcal{H})$-ECIC over $\mathbb{F}_q$ satisfies

$$\mathcal{N}_q[\mathcal{H}, \delta] \geqslant \kappa_q(\mathcal{H}) + 2\delta.$$

Proof: Let $\boldsymbol{L}$ be the $n \times \mathcal{N}_q[\mathcal{H}, \delta]$ matrix corresponding to some optimal $(\delta, \mathcal{H})$-ECIC. Let $\boldsymbol{L}'$ be the matrix obtained by deleting any 2δ columns from $\boldsymbol{L}$.

By Lemma 3.8, $\boldsymbol{L}$ satisfies

$$\text{wt}\left(\sum_{i \in K} z_i \boldsymbol{L}_i\right) \geqslant 2\delta + 1$$

for all $K \in \mathcal{J}(\mathcal{H})$ and all choices of $z_i \in \mathbb{F}_q^*,\ i \in K$. We deduce that the rows of $\boldsymbol{L}'$ also satisfy

$$\mathsf{wt}\left(\sum_{i\in K} z_i \boldsymbol{L}'_i\right) \geqslant 1.$$

By Corollary 3.10, $\boldsymbol{L}'$ corresponds to a linear $\mathcal{H}$-IC. Therefore, by Lemma 3.5, part 2, $\boldsymbol{L}'$ has at least $\kappa_q(\mathcal{H})$ columns. We deduce that

$$\mathcal{N}_q[\mathcal{H},\delta] - 2\delta \geqslant \kappa_q(\mathcal{H}),$$

which concludes the proof. ■

The following corollary from Proposition 4.6 and Theorem 5.1 demonstrates that, for sufficiently large alphabets, a concatenation of a classical MDS error-correcting code with an optimal (non-error-correcting) index code yields an optimal ECIC. However, as it was illustrated in Example 4.8, this does not hold for the index coding schemes over small alphabets.

Corollary 5.2 (MDS Error-Correcting Index Code): For $q \geqslant \kappa_q(\mathcal{H}) + 2\delta - 1$,

$$\mathcal{N}_q[\mathcal{H},\delta] = \kappa_q(\mathcal{H}) + 2\delta. \tag{11}$$

Proof: From Theorem 5.1, we have

$$\mathcal{N}_q[\mathcal{H},\delta] \geqslant \kappa_q(\mathcal{H}) + 2\delta.$$

On the other hand, from Proposition 4.6

$$\mathcal{N}_q[\mathcal{H},\delta] \leqslant N_q[\kappa_q(\mathcal{H}), 2\delta+1] = \kappa_q(\mathcal{H}) + 2\delta,$$

for $q \geqslant \kappa_q(\mathcal{H}) + 2\delta - 1$ (by taking doubly extended Reed–Solomon (RS) codes). Therefore, for these q, (11) holds. ■

Remark 5.3: Let $q = 2$, $m = n = 2\ell + 1$ ($\ell \geqslant 2$), and $f(i) = i$ for all $i \in [m]$. Let $\mathcal{X}_1 = \{2, n\}$ and $\mathcal{X}_n = \{1, n-1\}$. For $2 \leqslant i \leqslant n$, let $\mathcal{X}_i = \{i-1, i+1\}$. Let $\mathcal{H} = \mathcal{H}(n, n, \mathcal{X}, f)$. Then, $\mathcal{G}_\mathcal{H}$ is the (symmetric) odd cycle of length n. Therefore, $\alpha(\mathcal{H}) = \alpha(\mathcal{G}_\mathcal{H}) = \ell$. From [7], $\kappa_2(\mathcal{H}) = \text{min-rank}_2(\mathcal{G}_\mathcal{H}) = \ell + 1$. From α-bound

$$\mathcal{N}_2[\mathcal{H},\delta] \geqslant N_2[\ell, 2\delta+1].$$

By contrast, from Theorem 5.1

$$\mathcal{N}_2[\mathcal{H},\delta] \geqslant (\ell+1) + 2\delta.$$

As there are no nontrivial binary MDS codes, we have

$$N_2[\ell, 2\delta+1] \geqslant \ell + 2\delta + 1$$

for all choices of $\delta > 0$. Therefore, for these choices, the α-bound is at least as good as the Singleton bound.

VI. Random Codes

In this section, we prove an inexplicit upper bound on the optimal length of the ECICs. The proof is based on constructing a random ECIC and analyzing its parameters.

Theorem 6.1: Let $\mathcal{H} = \mathcal{H}(m, n, \mathcal{X}, f)$ describe an instance of the ICSI problem. Then, there exists a $(\delta,\mathcal{H})$-ECIC over $\mathbb{F}_q$ of length N if

$$\sum_{i\in[m]} q^{n-|\mathcal{X}_i|-1} < \frac{q^N}{V_q(N, 2\delta)}$$

where

$$V_q(N, 2\delta) = \sum_{\ell=0}^{2\delta} \binom{N}{\ell} (q-1)^\ell \tag{12}$$

is the volume of the q-ary sphere in $\mathbb{F}_q^N$.

Proof: We construct a random $n \times N$ matrix $\boldsymbol{L}$ over $\mathbb{F}_q$, row by row. Each row is selected independently of other rows, uniformly over $\mathbb{F}_q^N$. Define vector spaces

$$\boldsymbol{M}_i \stackrel{\triangle}{=} \text{span}\left(\{\boldsymbol{L}_j \ : \ j \in \mathcal{Y}_i\}\right)$$

for all $i \in [m]$. We also define the following events:

$$\forall i \in [m] \ : \ \text{Event } E_i \ \stackrel{\triangle}{=} \ \left\{\mathsf{d}(\boldsymbol{L}_{f(i)}, \boldsymbol{M}_i) < 2\delta + 1\right\}$$

and

$$\text{Event } E_{\text{Fail}} \stackrel{\triangle}{=} \\ \{\boldsymbol{L} \text{ does not correspond to a } (\delta,\ \mathcal{H})\text{-ECIC}\}.$$

The event E_i represents the situation when the receiver R_i cannot recover $x_{f(i)}$. Then, by Corollary 3.9, the event E_{Fail} is equivalent to $\bigcup_{i\in[m]} E_i$. Therefore

$$\Pr\left(E_{\text{Fail}}\right) = \Pr\left(\bigcup_{i\in[m]} E_i\right) \leqslant \sum_{i\in[m]} \Pr\left(E_i\right). \tag{13}$$

For a particular event $E_i,\ i \in [m]$,

$$\Pr\left(E_i\right) \leqslant \frac{q^{|\mathcal{Y}_i|}\, V_q(N, 2\delta)}{q^N}. \tag{14}$$

There exists a matrix $\boldsymbol{L}$ that corresponds to a $(\delta,\mathcal{H})$-ECIC if $\Pr\left(E_{\text{Fail}}\right) < 1$. It is enough to require that the right-hand side of (13) is smaller than 1. By plugging in the expression in (14), we obtain a sufficient condition on the existence of a $(\delta,\mathcal{H})$-ECIC over $\mathbb{F}_q$:

$$\frac{V_q(N, 2\delta)}{q^N} \sum_{i\in[m]} q^{|\mathcal{Y}_i|} < 1.$$

■

Remark 6.2: The bound in Theorem 6.1 does not take into account the structure of the sets $\mathcal{X}_i$'s, other than their cardinalities. Therefore, this bound generally is weaker than the κ-bound. On the other hand, for a particular instance of the ICSI problem, it is easier to compute this bound, while calculating the κ-bound in general is an NP-hard problem.

Remark 6.3: The bound in Theorem 6.1 implies a bound on $\kappa_q(\mathcal{H})$, which is tight for some $\mathcal{H}$. Indeed, fix $\delta = 0$. The bound implies that there exists a linear index code of length N whenever

$$\sum_{i \in [m]} q^{n-|\mathcal{X}_i|-1} < q^N. \tag{15}$$

Let $m = n = 2\ell + 1$ ($\ell \geqslant 2$), and $f(i) = i$ for all $i \in [n]$. Let $\mathcal{X}_1 = [n] \backslash \{1, 2, n\}$ and $\mathcal{X}_n = [n] \backslash \{1, n-1, n\}$. For $2 \leqslant i \leqslant n-1$, let $\mathcal{X}_i = [n] \backslash \{i-1, i, i+1\}$. Let $\mathcal{H} = \mathcal{H}(n, n, \mathcal{X}, f)$ be the corresponding side information hypergraph. Then, $\mathcal{G}_{\mathcal{H}}$ is the complement of the (symmetric directed) odd cycle of length n. We have $|\mathcal{X}_i| = 2\ell - 2$ for all $i \in [n]$. Then, (15) becomes

$$N > 2 + \log_q(2\ell + 1).$$

If $q > 2\ell + 1$, then we obtain $N \geqslant 3$. Observe that in this case, $\kappa_q(\mathcal{H}) = \text{min-rank}_q(\mathcal{G}_{\mathcal{H}}) = 3$ (see [13, Claim A.1]), and thus, the bound is tight.

VII. Syndrome Decoding

Consider the $(\delta, \mathcal{H})$-ECIC based on a matrix $\boldsymbol{L}$. Suppose that the receiver R_i, $i \in [m]$, receives the vector

$$\boldsymbol{y}_i = \boldsymbol{x}\boldsymbol{L} + \boldsymbol{\epsilon}_i \tag{16}$$

where $\boldsymbol{x}\boldsymbol{L}$ is the codeword transmitted by S, and $\boldsymbol{\epsilon}_i$ is the error pattern affecting this codeword.

In the classical coding theory, the transmitted vector $\boldsymbol{c}$, the received vector $\boldsymbol{y}$, and the error pattern $\boldsymbol{e}$ are related by $\boldsymbol{y} = \boldsymbol{c} + \boldsymbol{e}$. Therefore, if $\boldsymbol{y}$ is known to the receiver, then there is a one-to-one correspondence between the values of unknown vectors $\boldsymbol{c}$ and $\boldsymbol{e}$. For index coding, however, this is no longer the case. The following theorem shows that, in order to recover the message $x_{f(i)}$ from $\boldsymbol{y}_i$ using (16), it is sufficient to find just one vector from a set of possible error patterns. This set is defined as follows:

$$\mathcal{L}_i(\boldsymbol{\epsilon}_i) = \{\boldsymbol{\epsilon}_i + \boldsymbol{z} \; : \; \boldsymbol{z} \in \mathsf{span}(\{\boldsymbol{L}_j\}_{j \in \mathcal{Y}_i})\}.$$

We henceforth refer to the set $\mathcal{L}_i(\boldsymbol{\epsilon}_i)$ as the *set of relevant error patterns*.

Lemma 7.1: Assume that the receiver R_i receives $\boldsymbol{y}_i$.

1) If R_i knows the message $x_{f(i)}$, then it is able to determine the set $\mathcal{L}_i(\boldsymbol{\epsilon}_i)$.
2) If R_i knows some vector $\hat{\boldsymbol{\epsilon}} \in \mathcal{L}_i(\boldsymbol{\epsilon}_i)$, then it is able to determine $x_{f(i)}$.

Proof:

1) From (16), we have

$$\boldsymbol{y}_i = x_{f(i)}\boldsymbol{L}_{f(i)} + \boldsymbol{x}_{\mathcal{X}_i}\boldsymbol{L}_{\mathcal{X}_i} + \boldsymbol{x}_{\mathcal{Y}_i}\boldsymbol{L}_{\mathcal{Y}_i} + \boldsymbol{\epsilon}_i. \tag{17}$$

If R_i knows $x_{f(i)}$, then it is also able to determine

$$\boldsymbol{\epsilon}_i + \boldsymbol{x}_{\mathcal{Y}_i}\boldsymbol{L}_{\mathcal{Y}_i} = \boldsymbol{y}_i - x_{f(i)}\boldsymbol{L}_{f(i)} - \boldsymbol{x}_{\mathcal{X}_i}\boldsymbol{L}_{\mathcal{X}_i} \in \mathcal{L}_i(\boldsymbol{\epsilon}_i).$$

Since R_i has a knowledge of $\boldsymbol{L}$, it is also able to determine the whole $\mathcal{L}_i(\boldsymbol{\epsilon}_i)$.

2) Suppose that R_i knows a vector

$$\hat{\boldsymbol{\epsilon}} = \boldsymbol{\epsilon}_i + \sum_{j \in \mathcal{Y}_i} z_j \boldsymbol{L}_j \in \mathcal{L}_i(\boldsymbol{\epsilon}_i)$$

for some $\boldsymbol{z} = (z_j)_{\mathcal{Y}_i} \in \mathbb{F}_q^{|\mathcal{Y}_i|}$. We show that R_i is able then to determine $x_{f(i)}$. Indeed, we rewrite (17) as

$$\boldsymbol{y}_i = x_{f(i)}\boldsymbol{L}_{f(i)} + \boldsymbol{x}_{\mathcal{X}_i}\boldsymbol{L}_{\mathcal{X}_i} + (\boldsymbol{x}_{\mathcal{Y}_i} - \boldsymbol{z})\boldsymbol{L}_{\mathcal{Y}_i} + \hat{\boldsymbol{\epsilon}}. \tag{18}$$

The receiver R_i can find some solution of the equation

$$\boldsymbol{y}_i = \hat{x}_{f(i)}\boldsymbol{L}_{f(i)} + \boldsymbol{x}_{\mathcal{X}_i}\boldsymbol{L}_{\mathcal{X}_i} + \hat{\boldsymbol{x}}_{\mathcal{Y}_i}\boldsymbol{L}_{\mathcal{Y}_i} + \hat{\boldsymbol{\epsilon}} \tag{19}$$

with respect to the unknowns $\hat{x}_{f(i)}$ and $\hat{\boldsymbol{x}}_{\mathcal{Y}_i}$. Observe that (19) has at least one solution due to (18).

From (18) and (19), we deduce that

$$\boldsymbol{0} = (\hat{x}_{f(i)} - x_{f(i)})\boldsymbol{L}_{f(i)} + (\hat{\boldsymbol{x}}_{\mathcal{Y}_i} - \boldsymbol{x}_{\mathcal{Y}_i} + \boldsymbol{z})\boldsymbol{L}_{\mathcal{Y}_i}.$$

This equality implies that $\hat{x}_{f(i)} = x_{f(i)}$ (otherwise, by Corollary 3.9, the sum in the right-hand side will have nonzero weight). Hence, R_i is able to determine $x_{f(i)}$, as claimed.

■

We now describe a syndrome decoding algorithm for linear ECICs. From (17), we have

$$\boldsymbol{y}_i - \boldsymbol{x}_{\mathcal{X}_i}\boldsymbol{L}_{\mathcal{X}_i} - \boldsymbol{\epsilon}_i \in \mathsf{span}\big(\{\boldsymbol{L}_{f(i)}\} \cup \{\boldsymbol{L}_j\}_{j \in \mathcal{Y}_i}\big).$$

Let $\mathcal{C}_i = \mathsf{span}(\{\boldsymbol{L}_{f(i)}\} \cup \{\boldsymbol{L}_j\}_{j \in \mathcal{Y}_i})$, and let $\boldsymbol{H}^{(i)}$ be a parity check matrix of $\mathcal{C}_i$. We obtain that

$$\boldsymbol{H}^{(i)}\boldsymbol{\epsilon}_i^T = \boldsymbol{H}^{(i)}(\boldsymbol{y}_i - \boldsymbol{x}_{\mathcal{X}_i}\boldsymbol{L}_{\mathcal{X}_i})^T. \tag{20}$$

Let $\boldsymbol{\beta}_i$ be a column vector defined by

$$\boldsymbol{\beta}_i = \boldsymbol{H}^{(i)}(\boldsymbol{y}_i - \boldsymbol{x}_{\mathcal{X}_i}\boldsymbol{L}_{\mathcal{X}_i})^T. \tag{21}$$

Observe that each R_i is capable of determining $\boldsymbol{\beta}_i$. Then, we can rewrite (20) as

$$\boldsymbol{H}^{(i)}\boldsymbol{\epsilon}_i^T = \boldsymbol{\beta}_i.$$

This leads us to the formulation of the decoding procedure for R_i, which is presented in Fig. 5.

Remark 7.2: The solution $\hat{\boldsymbol{\epsilon}}$ in (22) might not be unique. Nevertheless, any such solution of the lowest Hamming weight yields the correct answer in the algorithm.

Remark 7.3: Gaussian elimination can be used to solve (23) for $\hat{x}_{f(i)}$. However, since $\boldsymbol{L}$ also corresponds to an $\mathcal{H}$-IC, there is more efficient way to do so. From Lemma 3.5, there exists a vector $\boldsymbol{v}_i \lhd \mathcal{X}_i$ satisfying $\boldsymbol{v}_i + \boldsymbol{e}_{f(i)} \in \mathsf{colspan}(\boldsymbol{L})$. Hence, $\boldsymbol{v}_i + \boldsymbol{e}_{f(i)} = \boldsymbol{u}\boldsymbol{L}^T$ for some $\boldsymbol{u} \in \mathbb{F}_q^N$. Therefore

$$\begin{aligned} \hat{x}_{f(i)} &= \hat{\boldsymbol{x}}(\boldsymbol{v}_i + \boldsymbol{e}_{f(i)})^T - \hat{\boldsymbol{x}}\boldsymbol{v}_i^T \\ &= \hat{\boldsymbol{x}}\boldsymbol{L}\boldsymbol{u}^T - \hat{\boldsymbol{x}}\boldsymbol{v}_i^T \\ &= (\boldsymbol{y}_i - \hat{\boldsymbol{\epsilon}})\boldsymbol{u}^T - \hat{\boldsymbol{x}}\boldsymbol{v}_i^T. \end{aligned}$$

With the knowledge of $\boldsymbol{L}$ and $\boldsymbol{x}_{\mathcal{X}_i}$, R_i can determine $\boldsymbol{u}$ and $\hat{\boldsymbol{x}}\boldsymbol{v}_i^T$. Therefore, it can also determine $\hat{x}_{f(i)}$. Note that (23) may have more than one solution $\hat{\boldsymbol{x}}$ with $\hat{\boldsymbol{x}}_{\mathcal{X}_i} = \boldsymbol{x}_{\mathcal{X}_i}$. However, as shown in the next theorem, if at most δ errors occur in $\boldsymbol{y}_i$, then it always holds that $\hat{x}_{f(i)} = x_{f(i)}$.

Example 7.4: Consider the ICSI instance presented in Example 4.8. Suppose the binary ECIC based on the following matrix is used

$$\boldsymbol{L} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 \end{pmatrix}.$$

Let $\boldsymbol{x} = (1,0,1,1,1)$. The sender broadcasts $\boldsymbol{xL}$. As $\mathcal{X}_1 = \{2,5\}$, the receiver R_1 knows in advance $x_2 = 0$ and $x_5 = 1$. Since $f(1) = 1$, R_1 requests x_1. Suppose two errors occur and R_1 receives the erroneous vector $\boldsymbol{y}_1 = \boldsymbol{xL} + \boldsymbol{\epsilon}_1$, where

$$\boldsymbol{\epsilon}_1 = (0,0,1,0,1,0,0,0,0).$$

Then

$$\boldsymbol{y}_1 = (1,1,0,1,1,0,1,1,0).$$

As $f(1) = 1$ and $\mathcal{Y}_1 = \{3,4\}$,

$$\mathcal{C}_1 = \mathsf{span}(\{\boldsymbol{L}_1, \boldsymbol{L}_3, \boldsymbol{L}_4\}).$$

A parity check matrix of $\mathcal{C}_1$ is

$$\boldsymbol{H}^{(1)} = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \end{pmatrix}.$$

Recall that $\mathcal{X}_1 = \{2,5\}$. We have

$$\begin{aligned} \boldsymbol{x}_{\mathcal{X}_1} \boldsymbol{L}_{\mathcal{X}_1} &= (0\ 1) \begin{pmatrix} 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 \end{pmatrix} \\ &= (1,0,1,0,1,0,0,1,1) \end{aligned}$$

and

$$\boldsymbol{y}_1 - \boldsymbol{x}_{\mathcal{X}_1} \boldsymbol{L}_{\mathcal{X}_1} = (0,1,1,1,0,0,1,0,1).$$

Therefore, the syndrome is

$$\beta_1 = \boldsymbol{H}^{(1)} (\boldsymbol{y}_1 - \boldsymbol{x}_{\mathcal{X}_1} \boldsymbol{L}_{\mathcal{X}_1})^T = (0,1,0,0,0,1)^T.$$

In this case, $\hat{\boldsymbol{\epsilon}} = \boldsymbol{\epsilon}_1$ is the unique lowest Hamming weight vector that has syndrome β_1. Choosing $\boldsymbol{v}_1 = (0,1,0,0,1) \triangleleft \mathcal{X}_1$, $\boldsymbol{u} = (0,0,0,0,1,0,0,0,0)$, we have

$$\boldsymbol{e}_1 + \boldsymbol{v}_1 = \boldsymbol{u} \boldsymbol{L}^T.$$

Since R_1 knows x_2 and x_5, R_1 can compute

$$\hat{\boldsymbol{x}} \boldsymbol{v}_1^T = \hat{x}_2 + \hat{x}_5 = x_2 + x_5 = 0 + 1 = 1.$$

Therefore, as discussed in Remark 7.3, R_1 obtains

$$\hat{x}_1 = (\boldsymbol{y}_1 - \hat{\boldsymbol{\epsilon}}) \boldsymbol{u}^T - \hat{\boldsymbol{x}} \boldsymbol{v}_1^T = 0 - 1 = 1,$$

which is equal to x_1—the message that R_1 requests.

- *Input:* $\boldsymbol{y}_i$, $\boldsymbol{x}_{\mathcal{X}_i}$, $\boldsymbol{L}$.
- *Step 1*: Compute the syndrome
$$\boldsymbol{\beta}_i = \boldsymbol{H}^{(i)} (\boldsymbol{y}_i - \boldsymbol{x}_{\mathcal{X}_i} \boldsymbol{L}_{\mathcal{X}_i})^T .$$
- *Step 2*: Find the lowest Hamming weight solution $\hat{\boldsymbol{\epsilon}}$ of the system
$$\boldsymbol{H}^{(i)} \hat{\boldsymbol{\epsilon}}^T = \boldsymbol{\beta}_i . \qquad (22)$$
- *Step 3*: Given that $\hat{\boldsymbol{x}}_{\mathcal{X}_i} = \boldsymbol{x}_{\mathcal{X}_i}$, solve the system for $\hat{x}_{f(i)}$:
$$\boldsymbol{y}_i = \hat{\boldsymbol{x}} \boldsymbol{L} + \hat{\boldsymbol{\epsilon}}. \qquad (23)$$
- *Output:* $\hat{x}_{f(i)}$.

Fig. 5. Syndrome decoding procedure.

Theorem 7.5: Let $\boldsymbol{y}_i = \boldsymbol{xL} + \boldsymbol{\epsilon}_i$ be the vector received by R_i, and let $\mathsf{wt}(\boldsymbol{\epsilon}_i) \leqslant \delta$. Assume that the procedure in Fig. 5 is applied to $(\boldsymbol{y}_i, \boldsymbol{x}_{\mathcal{X}_i}, \boldsymbol{L})$. Then, its output satisfies $\hat{x}_{f(i)} = x_{f(i)}$.

Proof: By Lemma 7.1, it is sufficient to prove that $\hat{\boldsymbol{\epsilon}} \in \mathcal{L}_i(\boldsymbol{\epsilon}_i)$. Indeed, since

$$\boldsymbol{H}^{(i)} \boldsymbol{\epsilon}_i^T = \boldsymbol{H}^{(i)} \hat{\boldsymbol{\epsilon}}^T = \boldsymbol{\beta}_i$$

we have

$$\boldsymbol{H}^{(i)} (\hat{\boldsymbol{\epsilon}} - \boldsymbol{\epsilon}_i)^T = \boldsymbol{0}.$$

Hence, $\hat{\boldsymbol{\epsilon}} - \boldsymbol{\epsilon}_i \in \mathcal{C}_i$, and therefore,

$$\hat{\boldsymbol{\epsilon}} - \boldsymbol{\epsilon}_i = z_{f(i)} \boldsymbol{L}_{f(i)} + \sum_{j \in \mathcal{Y}_i} z_j \boldsymbol{L}_j \qquad (24)$$

for some $z_{f(i)} \in \mathbb{F}_q$ and $z_j \in \mathbb{F}_q$, $j \in \mathcal{Y}_i$.

Since $\boldsymbol{\epsilon}_i$ is a solution of (22), and $\mathsf{wt}(\boldsymbol{\epsilon}_i) \leqslant \delta$, we deduce that $\mathsf{wt}(\hat{\boldsymbol{\epsilon}}) \leqslant \delta$ as well. Hence

$$\mathsf{wt} \left(z_{f(i)} \boldsymbol{L}_{f(i)} + \sum_{j \in \mathcal{Y}_i} z_j \boldsymbol{L}_j \right) = \mathsf{wt} \left(\hat{\boldsymbol{\epsilon}} - \boldsymbol{\epsilon}_i \right) \leqslant 2\delta.$$

Therefore, by Corollary 3.9, $z_{f(i)} = 0$. Hence, $\hat{\boldsymbol{\epsilon}} \in \mathcal{L}_i(\boldsymbol{\epsilon}_i)$, as desired, and therefore, $\hat{x}_{f(i)} = x_{f(i)}$. ■

Remark 7.6: We anticipate Step 2 in Fig. 5 to be computationally hard. Indeed, the problem of finding $\hat{\boldsymbol{\epsilon}}$ over $\mathbb{F}_2$ of the lowest weight satisfying

$$\boldsymbol{H}^{(i)} \hat{\boldsymbol{\epsilon}}^T = \boldsymbol{\beta}_i \qquad (25)$$

for a given binary vector $\boldsymbol{\beta}_i$ is at least as hard as a decision problem *coset weights* that was shown in [23] to be NP-complete.

VIII. Static Codes and Related Problems

A. Static Error-Correcting Index Codes

In the previous sections, we focused on linear δ-error-correcting index codes for a *particular* instance of the ICSI problem. When some of the parameters m, n, $\mathcal{X}$, and f are variable or not known, it is very likely that an ECIC for the instance with particular values of these parameters cannot be used for the instances with different values of some of these parameters. Therefore, it is interesting to design an ECIC which will be suitable for a *family* of instances of the ICSI problem.

Definition 8.1: Let $\Gamma = \{(m, n, \mathcal{X}, f)\}$ be a set of instances for an ICSI problem. A δ-error-correcting index code over $\mathbb{F}_q$ is said to be *static* under the set Γ if it is a δ-error-correcting $(m, n, \mathcal{X}, f)$-IC over $\mathbb{F}_q$ for all instances $(m, n, \mathcal{X}, f) \in \Gamma$.

Recall that an instance $(m, n, \mathcal{X}, f)$ can be described by the side information hypergraph $\mathcal{H}(m, n, \mathcal{X}, f)$. For a set Γ of instances $(m, n, \mathcal{X}, f)$, let

$$\mathfrak{J}(\Gamma) \triangleq \bigcup_{(m,n,\mathcal{X},f)\in\Gamma} \mathcal{J}(\mathcal{H}(m, n, \mathcal{X}, f)) \tag{26}$$

where $\mathcal{J}(\mathcal{H}(m, n, \mathcal{X}, f))$ is defined as in (3). We also define

$$n(\Gamma) \triangleq \max\{n : (m, n, \mathcal{X}, f) \in \Gamma\}.$$

Lemma 8.2: The $n(\Gamma) \times N$ matrix $\boldsymbol{L}$ corresponds to a δ-error-correcting index code which is static under Γ if and only if

$$\mathsf{wt}\left(\sum_{i\in K} z_i \boldsymbol{L}_i\right) \geqslant 2\delta + 1$$

for all $K \in \mathfrak{J}(\Gamma)$ and for all choices of $z_i \in \mathbb{F}_q^*,\ i \in K$.

Proof: The proof follows from Definition 8.1 and Lemma 3.8. ■

Notice that when $\boldsymbol{L}$ is used for an instance $(m, n, \mathcal{X}, f) \in \Gamma$ with $n < n(\Gamma)$, then the last $n(\Gamma) - n$ rows of $\boldsymbol{L}$ are simply discarded.

One particular family of interest is $\Gamma(n, \rho)$, the family that contains all instances where each receiver owns at least $n - \rho$ messages as its side information. More formally

$$\Gamma(n, \rho) = \big\{(m, n', \mathcal{X}, f)\ :\ n' \leqslant n \text{ and } \forall i \in [m],\ |\mathcal{X}_i| \geqslant n - \rho\big\}.$$

A δ-error-correcting index code which is static under $\Gamma(n, \rho)$ will provide successful communication between the sender and the receivers under the presence of at most δ errors, despite a possible change of the collection of the side information sets $\mathcal{X}$, a change of the set of receivers, and a change of the demand function, as long as each receiver still possesses at least $n - \rho$ messages.

In the rest of this section, we assume that $N \geqslant 1$, $n \geqslant \rho \geqslant 1$ and $\delta \geqslant 0$.

Definition 8.3: An $n \times N$ matrix $\boldsymbol{L}$ is said to satisfy the (ρ, δ)-Property if any nontrivial linear combination of at most ρ rows of $\boldsymbol{L}$ has weight at least $2\delta + 1$.

Proposition 8.4: The $n \times N$ matrix $\boldsymbol{L}$ corresponds to a δ-error-correcting linear index code, which is static under $\Gamma(n, \rho)$, if and only if $\boldsymbol{L}$ satisfies the (ρ, δ)-Property.

Proof: Let $\boldsymbol{L}$ be an $n \times N$ matrix that satisfies the (ρ, δ)-Property. We show that this is equivalent to the condition that $\boldsymbol{L}$ corresponds to a δ-error-correcting linear index code, which is static under $\Gamma(n, \rho)$. By Lemma 8.2, it suffices to show that $\mathfrak{J}(\Gamma(n, \rho))$ is the collection of all nonempty subsets of $[n]$, whose cardinalities are not greater than ρ.

Consider an instance $(m, n', \mathcal{X}, f) \in \Gamma(n, \rho)$. For all $i \in [m]$, we have $|\mathcal{X}_i| \geqslant n - \rho$ and $\mathcal{Y}_i = [n'] \backslash (f(i) \cup \mathcal{X}_i)$, and thus we deduce that

$$|\mathcal{Y}_i| \leqslant n' - 1 - (n - \rho) \leqslant n' - 1 - (n' - \rho) = \rho - 1.$$

Hence by (3), the cardinality of each set in $\mathcal{J}(\mathcal{H}(m, n', \mathcal{X}, f))$ is at most

$$1 + (\rho - 1) = \rho.$$

Therefore, due to (26), every set in $\mathfrak{J}(\Gamma(n, \rho))$ has at most ρ elements.

It remains to show that every nonempty subset of $[n]$ whose cardinality is at most ρ belongs to $\mathfrak{J}(\Gamma(n, \rho))$. Consider an arbitrary ρ'-subset $K = \{i_1, i_2, \ldots, i_{\rho'}\}$ of $[n]$, with $1 \leqslant \rho' \leqslant \rho$. Consider an instance $(m = 1, n, \mathcal{X}, f) \in \Gamma(n, \rho)$ with $\mathcal{X}_1 = [n] \backslash K$ and $f(1) = i_1$. Since

$$\mathcal{Y}_1 = K \backslash \{i_1\}$$

we have

$$K = \{i_1\} \cup \mathcal{Y}_1 \in \mathcal{J}(\mathcal{H}(m, n, \mathcal{X}, f)) \subseteq \mathfrak{J}(\Gamma(n, \rho)).$$

The proof follows. ■

B. Application: Weakly Resilient Functions

In this section, we introduce the notion of weakly resilient functions. Hereafter, we restrict the discussion to the binary alphabet.

The concept of *binary resilient functions* was first introduced by Chor *et al.* [24] and independently by Bennet *et al.* [25].

Definition 8.5: A function $\boldsymbol{f} : \mathbb{F}_2^N \to \mathbb{F}_2^n$ is called *t-resilient* if $\boldsymbol{f}$ satisfies the following property: when t arbitrary inputs of $\boldsymbol{f}$ are fixed and the remaining $N - t$ inputs run through all the 2^{N-t}-tuples exactly once, the value of $\boldsymbol{f}$ runs through every possible output n-tuple an equal number of times. Moreover, if $\boldsymbol{f}$ is a linear transformation, then it is called a *linear t-resilient function*. We refer to the parameter t as the *resiliency* of $\boldsymbol{f}$.

The applications of resilient functions can be found in fault-tolerant distributed computing, quantum cryptographic key distribution [24], privacy amplification [25], and random sequence generation for stream ciphers [26]. Connections between linear error-correcting codes and resilient functions were established in [24].

Theorem 8.6 [24]: Let $\boldsymbol{L}$ be an $n \times N$ binary matrix. Then, $\boldsymbol{L}$ is a generator matrix of a linear error-correcting code with minimum distance $d = t + 1$ if and only if $\boldsymbol{f}(\boldsymbol{z}) = \boldsymbol{L}\boldsymbol{z}^T$ is t-resilient.

Remark 8.7: Vectorial Boolean functions with certain properties are useful for design of stream ciphers. These properties include high resiliency and high nonlinearity (see, for instance, [26]). However, linear resilient functions are still particularly interesting, since they can be transformed into highly nonlinear resilient functions with the same parameters. This can be achieved by a composition of the linear function with a highly nonlinear permutation (see [27] and [28] for more details).

Below we introduce a definition of a ρ-weakly t-resilient function, which is a weaker version of a t-resilient function.

Definition 8.8: A function $\boldsymbol{f} : \mathbb{F}_2^N \to \mathbb{F}_2^n$ is called *ρ-weakly t-resilient* if $\boldsymbol{f}$ satisfies the property that every set of ρ coordinates in the image of $\boldsymbol{f}$ runs through every possible output ρ-tuple an equal number of times, when t arbitrary inputs of $\boldsymbol{f}$ are fixed and the remaining $N-t$ inputs run through all the 2^{N-t}-tuples exactly once.

Remark 8.9: A ρ-weakly t-resilient function $\boldsymbol{f} : \mathbb{F}_2^N \to \mathbb{F}_2^n$ can be viewed as a collection of $\binom{n}{\rho}$ different t-resilient functions $\mathbb{F}_2^N \to \mathbb{F}_2^\rho$, each such function is obtained by taking some ρ coordinates in the image of $\boldsymbol{f}$. Similarly to [24], consider a scenario, in which two parties are sharing a secret key, which consists of N randomly selected bits. Suppose that at some moment t out of the N bits of the key are leaked to an adversary. By applying a t-resilient function to the current N-bit key, two parties are able to obtain a completely new and secret key of n bits, without requiring any communication or randomness generation. However, if the parties use various parts of the key for various purposes, they may only require one of the ρ-bit secret keys (instead of the larger n-bit key). In that case, a ρ-weakly t-resilient function can be used. By applying a ρ-weakly t-resilient function to the current N-bit key, the parties obtain a set of $\binom{n}{\rho}$ different ρ-bit keys, each key is new and secret (however, these keys might not be independent of each other).

Theorem 8.10: Let $\boldsymbol{L}$ be an $n \times N$ binary matrix. Then, $\boldsymbol{L}$ satisfies the (ρ, δ)-Property if and only if the function $\boldsymbol{f} : \mathbb{F}_2^N \to \mathbb{F}_2^n$ defined by $\boldsymbol{f}(\boldsymbol{z}) = \boldsymbol{L}\boldsymbol{z}^T$ is ρ-weakly 2δ-resilient.

Proof:

1) Suppose that $\boldsymbol{L}$ satisfies the (ρ, δ)-Property. Take any ρ-subset $K \subseteq [n]$. By Definition 8.3, the $\rho \times N$ submatrix $\boldsymbol{L}_K$ of $\boldsymbol{L}$ is a generating matrix of the error-correcting code with the minimum distance $\geqslant 2\delta + 1$. By Theorem 8.6, the function $\boldsymbol{f}_K : \mathbb{F}_2^N \to \mathbb{F}_2^\rho$ defined by $\boldsymbol{f}_K(\boldsymbol{z}) = \boldsymbol{L}_K\boldsymbol{z}^T$ is 2δ-resilient. Since K is an arbitrary ρ-subset of $[n]$, the function $\boldsymbol{f}$ is ρ-weakly 2δ-resilient.
2) Conversely, assume that the function $\boldsymbol{f}$ is ρ-weakly 2δ-resilient. Take any subset $K \subseteq [n]$, $|K| = \rho$. Then, the function $\boldsymbol{f}_K : \mathbb{F}_2^N \to \mathbb{F}_2^\rho$ defined by $\boldsymbol{f}_K(\boldsymbol{z}) = \boldsymbol{L}_K\boldsymbol{z}^T$ is 2δ-resilient. Therefore, by Theorem 8.6, $\boldsymbol{L}_K$ is a generating matrix of a linear code with minimum distance $2\delta + 1$. Since K is an arbitrary ρ-subset of $[n]$, by Proposition 8.4, $\boldsymbol{L}$ satisfies the (ρ, δ)-Property.

■

C. Bounds and Constructions

In this section, we study the problem of constructing a matrix $\boldsymbol{L}$ satisfying the (ρ, δ)-Property. Such $\boldsymbol{L}$ with the minimal possible number of columns is called *optimal*. First, observe that from Proposition 8.4, we have

$$\mathfrak{J}(\Gamma(n,\rho)) = \bigcup_{i=1}^{\rho} \binom{[n]}{i}$$

which is the set of all nonempty subsets of $[n]$ of cardinality at most ρ. Next, consider an instance $(m^*, n, \mathcal{X}^*, f^*)$ satisfying

$$\mathcal{J}(\mathcal{H}^*) = \mathfrak{J}(\Gamma(n,\rho)) \tag{27}$$

where $\mathcal{H}^* = \mathcal{H}(m^*, n, \mathcal{X}^*, f^*)$ is the side information hypergraph corresponding to that instance. Such an instance can be constructed as follows. For each subset $K = \{i_1, i_2, \ldots, i_{\rho'}\} \subseteq [n]$ $(1 \leqslant \rho' \leqslant \rho)$, we introduce a receiver which requests the message x_{i_1} and has a set $\{x_j : j \in [n] \backslash K\}$ as its side information. It is straightforward to verify that indeed we obtain an instance $(m^*, n, \mathcal{X}^*, f^*)$ satisfying (27). The problem of designing an optimal matrix $\boldsymbol{L}$ satisfying the (ρ, δ)-Property then becomes equivalent to the problem of finding an optimal $(\delta, \mathcal{H}^*)$-ECIC. Thus, $\mathcal{N}_q[\mathcal{H}^*, \delta]$ is equal to the number of columns in an optimal matrix which satisfies the (ρ, δ)-Property.

The corresponding α-bound and κ-bound for $\mathcal{N}_q[\mathcal{H}^*, \delta]$ can be stated as follows.

Theorem 8.11: Let ρ^* be the smallest number such that a linear $[n, n-\rho^*, \geqslant \rho + 1]_q$ code exists. Then, we have

$$N_q[\rho, 2\delta + 1] \leqslant \mathcal{N}_q[\mathcal{H}^*, \delta] \leqslant N_q[\rho^*, 2\delta + 1].$$

Proof: The first inequality follows from the α-bound and from the fact that $\alpha(\mathcal{H}^*) = \rho$, which is due to (27).

For the second inequality, it suffices to show that $\kappa_q(\mathcal{H}^*) = \rho^*$. By Corollary 3.10, an $n \times N$ matrix $\boldsymbol{L}$ corresponds to an $\mathcal{H}^*$-IC if and only if $\{\boldsymbol{L}_i : i \in K\}$ is linearly independent for every $K \in \mathcal{J}(\mathcal{H}^*)$. Since $\mathcal{J}(\mathcal{H}^*)$ is the set of all nonempty subsets of cardinality at most ρ, this is equivalent to saying that every set of at most ρ rows of $\boldsymbol{L}$ is linearly independent. This condition is equivalent to the condition that $\boldsymbol{L}^T$ is a parity check matrix of a linear code with the minimum distance at least $\rho + 1$ [29, Ch. 1]. Therefore, a linear $\mathcal{H}^*$-IC of length N exists if and only if an $[n, n-N, \geqslant \rho+1]_q$ linear code exists. Since ρ^* is the smallest number such that an $[n, n-\rho^*, \geqslant \rho + 1]_q$ code exists, we conclude that $\kappa_q(\mathcal{H}^*) = \rho^*$. ■

Corollary 8.12: The length of an optimal δ-error-correcting linear index code over $\mathbb{F}_q$ which is static under $\Gamma(n, \rho)$ satisfies

$$\mathcal{N}_q[\delta, \mathcal{H}^*] \geqslant \rho^* + 2\delta$$

where ρ^* is the smallest number such that an $[n, n-\rho^*, \geqslant \rho+1]_q$ code exists.

Proof: This is a straightforward corollary of Theorem 5.1 (the Singleton bound) and Theorem 8.11. ■

Corollary 8.13: For $q \geqslant \max\{n-1, \rho+2\delta-1\}$, the length of an optimal δ-error-correcting linear index code over $\mathbb{F}_q$ which is static under $\Gamma(n, \rho)$ is $\rho + 2\delta$.

Proof: For $q \geqslant n - 1$, there exists an $[n, n-\rho^*, \rho+1]_q$ linear code with $\rho^* = \rho$ (for example, one can take an extended RS code [29, Ch. 11]). Due to the Singleton bound, we conclude that $\rho^* = \rho$ is the smallest value such that $[n, n-\rho^*, \rho+1]_q$ linear code exists. Following the lines of the proof of Theorem 8.11, there exists a δ-error-correcting index code of length $N_q[\rho, 2\delta+1]$, which is static under $\Gamma(n, \rho)$. As $q \geqslant \rho + 2\delta - 1$, we have

$$N_q[\rho, 2\delta + 1] = \rho + 2\delta$$

(for example, by taking an extended RS code). Due to Corollary 8.12, this static ECIC is optimal. ■

Remark 8.14: We observe from the proof of Theorem 8.11 that the problem of constructing an optimal linear (non-error-correcting) index code, which is static under $\Gamma(n,\rho)$, is, in fact, equivalent to the problem of constructing a parity check matrix of a classical linear error-correcting code.

Example 8.15: Let $n = 20$, $\rho = 10$, $\delta = 1$, and $q = 2$. From [21], the smallest possible dimension of a binary linear code of length 20 and minimum distance 11 is 3. We obtain that $\rho^* = 17$. We also have $N_2[17,3] = 22$. Theorem 8.11 implies the existence of a one-error-correcting binary index code of length 22 which can be used for any instance of IC problem, in which each receiver owns at least 10 out of (at most) 20 messages, as side information. It also implies that the length of any such static ECIC is at least $N_2[10,3] = 14$. Corollary 8.12 provides a better lower bound on the minimum length, which is $17 + 2 = 19$.

Example 8.16: Below we show that with the same number of inputs N and outputs n, a weakly resilient function may have strictly higher resiliency t.

From Example 8.15, there exists a linear vectorial Boolean function $\boldsymbol{f} : (\mathbb{F}_2)^{22} \to (\mathbb{F}_2)^{20}$ which is 10-weakly 2-resilient. According to [21], an optimal linear $[22,20]_2$ code has minimum distance $d = 2$. Hence, due to Theorem 8.6, the resiliency of any linear vectorial Boolean function $\boldsymbol{g} : (\mathbb{F}_2)^{22} \to (\mathbb{F}_2)^{20}$ cannot exceed one.

The problem of constructing an $n \times N$ matrix $\boldsymbol{L}$ that satisfies the (ρ,δ)-Property is a natural generalization of the problem of constructing the parity check matrix $\boldsymbol{H}$ of a linear $[n,k,d \geqslant \rho+1]_q$ code. Indeed, $\boldsymbol{H}$ is a parity check matrix of an $[n,k,d \geqslant \rho+1]_q$ code if and only if every set of ρ columns of $\boldsymbol{H}$ is linearly independent. Equivalently, any nontrivial linear combination of at most ρ columns of $\boldsymbol{H}$ has weight at least one. For comparison, $\boldsymbol{L}$ satisfies the (ρ,δ)-Property if and only if any nontrivial linear combination of at most ρ columns of $\boldsymbol{L}^T$ has weight at least $2\delta + 1$.

Some classical methods for deriving bounds on the parameters of error-correcting codes can be generalized to the case of linear static ECICs. Below we present a Gilbert–Varshamov-like bound.

Theorem 8.17: Let $V_q(N,2\delta)$ denotes the volume of q-ary sphere of radius 2δ in $\mathbb{F}_q^N$ given by (12). If

$$\sum_{i=0}^{\rho-1}\binom{n-1}{i}(q-1)^i < \frac{q^N}{V_q(N,2\delta)}$$

then there exists an $n \times N$ matrix $\boldsymbol{L}$ which satisfies the (ρ,δ)-Property.

Proof: We build up the set $\mathcal{R}$ of rows of $\boldsymbol{L}$ one by one. The first row can be any vector in $\mathbb{F}_q^N$ of weight at least $2\delta+1$. Now suppose we have chosen r rows so that no nontrivial linear combination of at most ρ among these r rows have weight less than $2\delta + 1$. There are at most

$$V_q(N,2\delta)\sum_{i=0}^{\rho-1}\binom{r}{i}(q-1)^i$$

vectors which are at distance less than $2\delta + 1$ from any linear combination of at most $\rho - 1$ among r chosen rows (this includes vectors at distance less than $2\delta + 1$ from $\mathbf{0}$). If this quantity is smaller than q^N, then we can add another row to the set $\mathcal{R}$ so that no nontrivial linear combination of at most ρ rows in $\mathcal{R}$ has weight less than $2\delta + 1$. The claim follows if we replace r by $n - 1$. ■

Remark 8.18: If we apply Theorem 6.1 to the instance $(m^*, n, \mathcal{X}^*, f^*)$ defined in the beginning of this section, then we obtain a bound, which is somewhat weaker then its counterpart in Theorem 8.17, namely the $n \times N$ matrix $\boldsymbol{L}$ as above exists if

$$\sum_{i=1}^{\rho}\binom{n}{i}q^{i-1} < \frac{q^N}{V_q(N,2\delta)}.$$

IX. Conclusion

In this work, we generalize the Index Coding with Side Information problem toward a setup with errors. Under this setup, each receiver should be able to recover its desired message even if a certain amount of errors happen in the transmitted data. This is the first work that considers such a problem.

A number of bounds on the length of an optimal error-correcting index code are constructed. As it is shown in Example 4.8, a separation of error-correcting code and index code sometimes leads to a nonoptimal scheme. This raises a question of designing coding schemes in which the two layers are treated as a whole. Therefore, the question of constructing error-correcting index codes with good parameters is still open.

A general decoding procedure for linear error-correcting index codes is discussed. The difference between decoding of a classical error-correcting code and decoding of an error-correcting index code is that in the latter case, each receiver does not require a complete knowledge of the error vector. This difference may help to ease the decoding process. Finding an efficient decoding method for error-correcting index codes (together with their corresponding constructions) is also still an open problem.

The notion of error-correcting index code is further generalized to static index code. The latter is designed to serve a family of instances of error-correcting index coding problem. The problem of designing an optimal static error-correcting index code is studied, and several bounds on the length of such codes are presented.

Appendix

Lemma A.1: If $\mathcal{G}_{\mathcal{H}}$ is symmetric, then the generalized independence number of $\mathcal{H}$ is the independence number of $\mathcal{G}_{\mathcal{H}}$.

Proof: It suffices to show that if $\mathcal{G}_{\mathcal{H}}$ is symmetric, then the set of generalized independent sets of $\mathcal{H}$ and the set of independent sets of $\mathcal{G}_{\mathcal{H}}$ coincide.

Let H be a generalized independent set in $\mathcal{H}$. If $|H| = 1$, then obviously H is an independent set in $\mathcal{G}_{\mathcal{H}}$. Assume that $|H| \geqslant 2$. For any pair of vertices $i, j \in H$, the set $\{i,j\}$ belongs to $\mathcal{J}(\mathcal{H})$. By definition of $\mathcal{J}(\mathcal{H})$, either there is no edge from i to j, or there is no edge from j to i, in $\mathcal{G}_{\mathcal{H}}$. Since $\mathcal{G}_{\mathcal{H}}$ is symmetric, there are no edges between i and j, in neither directions. Therefore, H is an independent set in $\mathcal{G}_{\mathcal{H}}$.

Conversely, let H be an independent set in $\mathcal{G}_{\mathcal{H}}$. For each $i \in H$, since there are no edges from i to all other vertices in H, we deduce that $H \backslash \{i\} \subseteq \mathcal{Y}_i$. Due to (3), every subset of H which contains i belongs to $\mathcal{J}(\mathcal{H})$. This holds for an arbitrary $i \in H$. Therefore, every nonempty subset of H belong to $\mathcal{J}(\mathcal{H})$. We obtain that H is a generalized independent set of $\mathcal{H}$. ■

Acknowledgment

The authors would like to thank the authors of [7] for providing a preprint of their paper.

References

[1] Y. Birk and T. Kol, "Informed-source coding-on-demand (ISCOD) over broadcast channels," in *Proc. IEEE Conf. Comput. Commun.*, San Francisco, CA, 1998, pp. 1257–1264.

[2] Y. Birk and T. Kol, "Coding-on-demand by an informed source (ISCOD) for efficient broadcast of different supplemental data to caching clients," *IEEE Trans. Inf. Theory*, vol. 52, no. 6, pp. 2825–2830, Jun. 2006.

[3] S. E. Rouayheb, A. Sprintson, and C. Georghiades, "On the index coding problem and its relation to network coding and matroid theory," *IEEE Trans. Inf. Theory*, vol. 56, no. 7, pp. 3187–3195, Jul. 2010.

[4] S. Katti, H. Rahul, W. Hu, D. Katabi, M. Médard, and J. Crowcroft, "XORs in the air: Practical wireless network coding," in *Proc. ACM SIGCOMM*, 2006, pp. 243–254.

[5] S. Katti, D. Katabi, H. Balakrishnan, and M. Médard, "Symbol-level network coding for wireless mesh networks," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 38, no. 4, pp. 401–412, 2008.

[6] Z. Bar-Yossef, Z. Birk, T. S. Jayram, and T. Kol, "Index coding with side information," in *Proc. 47th Annu. IEEE Symp. Found. Comput. Sci.*, 2006, pp. 197–206.

[7] Z. Bar-Yossef, Z. Birk, T. S. Jayram, and T. Kol, "Index coding with side information," *IEEE Trans. Inf. Theory*, vol. 57, no. 3, pp. 1479–1494, Mar. 2011.

[8] E. Lubetzky and U. Stav, "Non-linear index coding outperforming the linear optimum," in *Proc. 48th Annu. IEEE Symp. Found. Comput. Sci.*, 2007, pp. 161–168.

[9] Y. Wu, J. Padhye, R. Chandra, V. Padmanabhan, and P. A. Chou, "The local mixing problem," presented at the presented at the Inf. Theory Appl. Workshop, San Diego, CA, 2006.

[10] S. E. Rouayheb, M. A. R. Chaudhry, and A. Sprintson, "On the minimum number of transmissions in single-hop wireless coding networks," in *Proc. IEEE Inf. Theory Workshop*, 2007, pp. 120–125.

[11] S. E. Rouayheb, A. Sprintson, and C. Georghiades, "On the relation between the index coding and the network coding problems," in *Proc. IEEE Symp. Inf. Theory*, Toronto, ON, Canada, 2008, pp. 1823–1827.

[12] M. A. R. Chaudhry and A. Sprintson, "Efficient algorithms for index coding," in *Proc. IEEE Conf. Comput. Commun.*, 2008, pp. 1–4.

[13] N. Alon, A. Hassidim, E. Lubetzky, U. Stav, and A. Weinstein, "Broadcasting with side information," in *Proc. 49th Annu. IEEE Symp. Found. Comput. Sci.*, 2008, pp. 823–832.

[14] R. Ahlswede, N. Cai, S. Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inf. Theory*, vol. 46, no. 4, pp. 1204–1216, Jul. 2000.

[15] R. Koetter and M. Médard, "An algebraic approach to network coding," *IEEE/ACM Trans. Netw.*, vol. 11, no. 5, pp. 782–795, Oct. 2003.

[16] W. Haemers, "An upper bound for the Shannon capacity of a graph," *Algebr. Methods Graph Theory*, vol. 25, pp. 267–272, 1978.

[17] R. Peeters, "Orthogonal representations over finite fields and the chromatic number of graphs," *Combinatorica*, vol. 16, no. 3, pp. 417–431, 1996.

[18] R. Peeters, "Uniqueness of strongly regular graphs having minimal p -rank," *Linear Algebra Applicat.*, vol. 226–228, pp. 9–31, 1995.

[19] R. Peeters, "On the p-ranks of net graphs," *Designs, Codes Cryptogr.*, vol. 5, pp. 139–153, 1995.

[20] S. H. Dau, V. Skachek, and Y. M. Chee, "On the security of index coding with side information," *IEEE Trans. Inf. Theory*, vol. 58, no. 6, pp. 3975–3988, Jun. 2012.

[21] M. Grassl, Bounds on the Minimum Distance of Linear Codes and Quantum Codes [Online]. Available: http://www.codetables.de

[22] M. Chudnovsky, N. Robertson, P. Seymour, and R. Thomas, "The strong perfect graph theorem," *Ann. Math.*, vol. 164, pp. 51–229, 2006.

[23] E. R. Berlekamp, R. J. McEliece, and H. C. A. van Tilborg, "On the inherent intractability of certain coding problems," *IEEE Trans. Inf. Theory*, vol. IT-24, no. 3, pp. 384–386, May 1978.

[24] B. Chor, O. Goldreich, J. Håstad, J. Freidmann, S. Rudich, and R. Smolensky, "The bit extraction problem or t-resilient functions," in *Proc. 26th Annu. IEEE Symp. Found. Comput. Sci.*, 1985, pp. 396–407.

[25] C. H. Bennet, G. Brassard, and J. M. Robert, "Privacy amplification by public discussion," *SIAM J. Comput.*, vol. 17, pp. 210–229, 1988.

[26] C. Carlet*, Vectorial Boolean Functions for Cryptography*, ser. (ser. Boolean Models and Methods in Mathematics, Computer Science and Engineering). Cambridge, U.K.: Cambridge Univ. Press, 2010, ch. 9.

[27] X.-M. Zhang and Y. Zheng, "On nonlinear resilient functions," in *Proc. 14th Annu. Int. Conf. Theory Appl. Cryptogr. Technol.*, 1995, pp. 274–288.

[28] K. Gupta and P. Sarkar, "Improved construction of nonlinear resilient s-boxes," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 339–348, Jan. 2005.

[29] F. J. MacWilliams and N. J. A. Sloane*, The Theory of Error-Correcting Codes*. Amsterdam, The Netherlands: North Holland, 1977.

Son Hoang Dau received the B.S. degree in applied mathematics and informatics from the College of Science, Vietnam National University, Hanoi, Vietnam, in 2006. He received the M.S. degree (2009) and Ph.D. degree (2012) in mathematical sciences from the Division of Mathematical Sciences, Nanyang Technological University, Singapore.

During 2011-2012, he held research positions with Nanyang Technological University. He is currently a research fellow at SUTD-MIT International Design Centre, Singapore University of Technology and Design, Singapore.

His research interests are coding theory, network coding, and combinatorics.

Vitaly Skachek received the B.A. (Cum Laude), M.Sc. and Ph.D. degrees in computer science from the Technion—Israel Institute of Technology, in 1994, 1998 and 2007, respectively.

In the summer of 2004, he visited the Mathematics of Communications Department at Bell Laboratories under the DIMACS Special Focus Program in Computational Information Theory and Coding. During 2007-2012, he held research positions with the Claude Shannon Institute, University College Dublin, with the School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore, with the Coordinated Science Laboratory, University of Illinois at Urbana-Champaign, Urbana, and with the Department of Electrical and Computer Engineering, McGill University, Montreal. He is now a Senior Lecturer with the Institute of Computer Science, University of Tartu.

Dr. Skachek is a recipient of the Permanent Excellent Faculty Instructor award, given by Technion.

Yeow Meng Chee (SM'08) received the B.Math. degree in computer science and combinatorics and optimization and the M.Math. and Ph.D. degrees in computer science from the University of Waterloo, Waterloo, ON, Canada, in 1988, 1989, and 1996, respectively.

Currently, he is an Associate Professor at the Division of Mathematical Sciences, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore. Prior to this, he was Program Director of Interactive Digital Media R&D in the Media Development Authority of Singapore, Postdoctoral Fellow at the University of Waterloo and IBM's Zürich Research Laboratory, General Manager of the Singapore Computer Emergency Response Team, and Deputy Director of Strategic Programs at the Infocomm Development Authority, Singapore.

His research interest lies in the interplay between combinatorics and computer science/engineering, particularly combinatorial design theory, coding theory, extremal set systems, and electronic design automation.