Locality in Index Coding for Large Min-Rank

by

Lakshmi Prasad Natrajan, Hoang Dau, Prasad Krishnan, Lalitha Vadlamani

in

The 2019 IEEE International Symposium on Information Theory
(*ISIT*)

Paris, France

Report No: IIIT/TR/2019/-1

Centre for Communications
International Institute of Information Technology
Hyderabad - 500 032, INDIA
July 2019

Locality in Index Coding for Large Min-Rank

Lakshmi Natarajan, Hoang Dau, Prasad Krishnan and V. Lalitha

***Abstract*—An index code is said to be *locally decodable* if each receiver can decode its demand using its side information and by querying only a subset of the transmitted codeword symbols instead of observing the entire codeword. Local decodability can be a beneficial feature in some communication scenarios, such as when the receivers can afford to listen to only a part of the transmissions because of limited availability of power. The *locality* of an index code is the ratio of the maximum number of codeword symbols queried by a receiver to the message length. In this paper we analyze the optimum locality of linear codes for the family of index coding problems whose min-rank is one less than the number of receivers in the network. We first derive the optimal trade-off between the index coding rate and locality with vector linear coding when the side information graph is a directed cycle. We then provide the optimal trade-off achieved by scalar linear coding for a larger family of problems, viz., problems where the min-rank is only one less than the number of receivers. While the arguments used for achievability are based on known coding techniques, the converse arguments rely on new results on the structure of locally decodable index codes.**

I. Introduction

Index coding [1] is a central problem in network coding theory, because of its applications, such as in video-on-demand and daily newspaper delivery [2], and because of its strong relation to other coding theoretic problems, such as network coding [3], [4], coded caching [5], codes for distributed data storage [6], [7] etc. The index coding problem is to design a code for a broadcast channel where each receiver has prior side information of a subset of messages being transmitted. The objective is to minimize the number of uses of the broadcast channel, or equivalently, the *broadcast rate*.

Conventional index coding solutions in the literature require each receiver to observe the entire transmitted codeword in order to decode its demand. If the network involves a large number of receivers the number of transmissions that each receiver has to observe could be significantly larger than the size of the message itself. Thus conventional index coding solutions could be unfavorable in certain applications, such as when the power available at the wireless receivers is limited and they can not afford to listen to radio transmissions for an extended period of time. In such scenarios, it is desirable to use *locally decodable index codes* [8], which provide reductions in broadcast rate while requiring the receivers to query only a part of the transmitted codeword. The *locality* of an index code is the ratio of the number of codeword symbols queried by a receiver to the number of message symbols it demands [9]. The objective of locally decodable index coding is to minimize both broadcast rate and locality simultaneously, and achieve the optimal trade-off between these two parameters.

To the best of our knowledge, the idea of local decodability in index coding was introduced in [8] where the broadcast rates of random index coding problems, modeled as random graphs, were analyzed under a locality requirement. The results in [8] and [9] characterize the optimal broadcast rate of an arbitrary index coding problem when locality is set to the minimum possible value, which is unity. Constructions of index codes with locality greater than one were given in [9]. Locally decodable index codes were shown to be related to privacy in index coding in [10] and studied under the terminology 'k-limited access schemes'. The authors of [10] provide constructions that modify any given binary scalar linear index code into a locally decodable scalar linear code at the cost of increased broadcast rate.

Determining the optimal trade-off between broadcast rate and locality of a given index coding problem is yet to be addressed in the literature. This will not only involve designing good achievability schemes but also formulating tight lower bounds on locality and rate.

In this paper we consider linear index codes for the family of index coding problems whose *min-rank* is one less than the number of receivers in the problem. We consider the problems where the side information graph is a directed cycle, and derive the optimal trade-off between rate and locality when vector linear codes are used. We also analyse the dependence of locality on the length of the vector messages in vector linear index codes for directed cycles (Section IV). We then consider the larger class of index coding problems, viz., problems whose min-rank is one less than the number of receivers, and provide the exact trade-off between rate and locality when scalar linear codes are used (Section V). Note that directed cycles are a subset of this larger class of problems. The derivation of the optimal trade-off provided in Sections IV and V rely on new results on the structural properties of locally decodable index codes given in Section III.

The achievability schemes used in this paper are based on known index coding techniques for directed cycles. However, one of the main contributions of this paper is the development of new tools to derive good lower bounds on rate and locality that are vital in proving the optimality of these schemes.

Notation: For any positive integer N, we will denote the set $\{1, \ldots, N\}$ by $[N]$. Matrices and column vectors are denoted by bold upper and lower case letters, respectively, such as $\boldsymbol{A}$ and $\boldsymbol{x}$. The finite field of size q is denoted as $\mathbb{F}_q$.

Dr. Natarajan is with the Department of Electrical Engineering, Indian Institute of Technology Hyderabad, email: lakshminatarajan@iith.ac.in.

Dr. Dau is with the Department of Electrical and Computer Systems Engineering, Monash University, Melbourne, Australia, email: hoang.dau@monash.edu.

Dr. Krishnan and Dr. Lalitha are with the Signal Processing & Communications Research Center, International Institute of Information Technology Hyderabad, India, email:{prasad.krishnan, lalitha.v}@iiit.ac.in.

The subspace spanned by vectors $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_N$ is denoted by $\mathsf{span}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_N)$. The column space of a matrix $\boldsymbol{A}$ is denoted as $\mathcal{C}(A)$ and the null space of $\boldsymbol{A}$ is $\mathcal{N}(\boldsymbol{A}) = \{\boldsymbol{x} | \boldsymbol{A}\boldsymbol{x} = \boldsymbol{0}\}$. The support set of a vector $\boldsymbol{x}$ is denoted as $\mathsf{supp}(\boldsymbol{x})$.

II. System Model & Preliminaries

We consider index coding for a broadcast channel consisting of N receivers $\mathsf{Rx}_1, \ldots, \mathsf{Rx}_N$. The transmitter holds N messages $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N$ where the i^{th} message is demanded by Rx_i, and the messages $\boldsymbol{x}_j$, $j \in K_i$, are known at this receiver as side information, where $K_i \subset [N]$. The side information graph $G = (\mathcal{V}, \mathcal{E})$ is the directed graph with vertex set $\mathcal{V} = [N]$ and edge set $\mathcal{E} = \{(i, j) \,|\, i \in [N], j \in K_i\}$, and it completely specifies the index coding problem.

We are interested in linear index codes, and hence, we will assume that each message $\boldsymbol{x}_i$ is a vector of length M over a finite field $\mathbb{F}_q$. Note that for scalar linear index coding problems the message length $M = 1$. The M components of the i^{th} message vector $\boldsymbol{x}_i$ are denoted as $x_{i,1}, \ldots, x_{i,M}$. Encoding is performed by first concatenating the N messages into $\boldsymbol{x} = (\boldsymbol{x}_1^{\mathsf{T}}, \ldots, \boldsymbol{x}_N^{\mathsf{T}})^{\mathsf{T}} \in \mathbb{F}_q^{MN}$ and multiplying this vector with an encoding matrix $\boldsymbol{L} \in \mathbb{F}_q^{MN \times \ell}$ to generate a length ℓ the codeword $\boldsymbol{c}^{\mathsf{T}} = \boldsymbol{x}^{\mathsf{T}}\boldsymbol{L}$. Note that the MN components of the concatenated vector $\boldsymbol{x} = (x_1, \ldots, x_{MN})^{\mathsf{T}}$ and the components of the individual message vectors are related as $x_{(i-1)M+m} = x_{i,m}$ for $i \in [N]$ and $m \in [M]$. The code length corresponding to the encoding matrix $\boldsymbol{L}$ is ℓ and the broadcast rate is $\beta = \frac{\ell}{M}$.

Unlike the conventional index coding scenario where each receiver is required to query or download the entire codeword $\boldsymbol{c}$, we allow the receivers to query only a part of the transmitted codeword in order to decode their demands. Index codes that satisfy this property are called *locally decodable* [8]. We will assume that Rx_i queries the subvector $\boldsymbol{c}_{R_i} = (c_j, j \in R_i)$, where $R_i \subseteq [\ell]$ is chosen in such a way that Rx_i can decode $\boldsymbol{x}_i$ using $\boldsymbol{c}_{R_i}$ and the available side information $\boldsymbol{x}_j$, $j \in K_i$. The *locality* of Rx_i is $r_i = \frac{|R_i|}{M}$. Since Rx_i demands message symbols $x_{i,1}, \ldots, x_{i,M}$, it needs to query at least M components of the codeword to be able to decode them, and hence $r_i \geq 1$. The *overall locality* or simply the *locality* of the index code is $r = \max_{i \in [N]} r_i$, and the *average locality* is

$$r_{\text{avg}} = \sum_{i \in [N]} \frac{r_i}{N} = \sum_{i \in [N]} \frac{|R_i|}{MN}. \tag{1}$$

Observe that the average locality is upper bounded by overall locality $r_{\text{avg}} \leq r$.

Example 1 (*A simple scalar linear code for directed cycles*)**.** Let the message length $M = 1$ and let G be a directed cycle of length N, i.e., $K_i = \{i+1\}$ for $i \in [N-1]$ and $K_N = \{1\}$. Consider the $N \times (N-1)$ encoder matrix

$$\boldsymbol{L} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix},$$

that generates the codeword

$$\boldsymbol{c} = (c_1, \ldots, c_{N-1}) = \boldsymbol{x}^{\mathsf{T}}\boldsymbol{L} = (x_1 + x_2, x_1 + x_3, \ldots, x_1 + x_N).$$

Rx_1 and Rx_N can decode their demands by querying c_1 and c_{N-1}, respectively, hence, $|R_1| = |R_N| = 1$. For $1 < i < N$, receiver Rx_i queries the symbols $c_{i-1} = x_1 + x_i$ and $c_i = x_1 + x_{i+1}$, and uses its side information x_{i+1} to compute $c_{i-1} - c_i - x_{i+1}$, which equals its demand x_i. Hence, $|R_i| = 2$ for $1 < i < N$. Since $M = 1$, the locality of each receiver $r_i = 1$ if $i = 1$ or N, and $r_i = 2$ otherwise. The overall locality $r = 2$ and the average locality $r_{\text{avg}} = 2(N-1)/N$.

Since the graph G is symmetric, for any choice of $i \in [N-1]$, the above coding scheme can be modified by an appropriate permutation of the rows of $\boldsymbol{L}$ to allow receiver localities $r_i = 1$ and $r_{i+1} = 1$ and locality $r_j = 2$ at all other receivers $j \neq i, i+1$. ■

We would like to characterize the trade-off between the broadcast rate β and the locality r of linear index codes over $\mathbb{F}_q$ for a given index coding problem G. We define the optimum locality-rate trade-off among all linear index codes for G over $\mathbb{F}_q$ as

$$\beta^*_{G,q}(r) = \inf \{\beta \,|\, \exists \text{ a linear code of rate } \beta, \text{ locality} \leq r\},$$

where the infimum considers linear index codes over all possible message lengths $M \geq 1$ for the index coding problem G.

We would like to view the vector linear index coding problem involving N vector messages $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N \in \mathbb{F}_q^M$ as a scalar linear problem defined over MN scalar messages $x_1, \ldots, x_{MN} \in \mathbb{F}_q$. In this case the i^{th} receiver demands M scalar messages $x_{(i-1)M+1}, x_{(i-1)M+2}, \ldots, x_{iM}$, which correspond to the M components of the vector $\boldsymbol{x}_i$. This set of demands of Rx_i is represented by the index set $\mathcal{D}_i = \{(i-1)M + m \,|\, m \in [M]\}$. The scalar symbols available as side information at Rx_i correspond to the index set

$$\mathcal{K}_i = \{(j-1)M + m \,|\, j \in K_i, m \in [M]\}.$$

Thus the vector linear problem corresponding to the side information graph G with message length M, is equivalent to a scalar linear problem with N receivers and MN messages, where the index set of demands of Rx_i is $\mathcal{D}_i$ and the set corresponding to the side information at Rx_i is $\mathcal{K}_i$. Note that $\mathcal{D}_i$ and $\mathcal{K}_i$ are subsets of $[MN]$.

III. Structure of Locally Decodable Index Codes

We will first derive some properties of locally decodable (vector) linear index codes (codes for any message length $M \geq 1$) in Section III-A and then analyze scalar linear codes

(message length $M = 1$) specifically in Section III-B. These properties will be useful in deriving the rate-locality trade-off results presented in later sections.

Following the notation from [11], for a vector $\boldsymbol{u} \in \mathbb{F}_q^{MN}$ and set $E \subset [MN]$, we write $\boldsymbol{u} \lhd E$ to denote $\mathsf{supp}(\boldsymbol{u}) \subseteq E$. Let us denote the columns of the encoder matrix $\boldsymbol{L}$ as $\boldsymbol{L}_1, \ldots, \boldsymbol{L}_\ell \in \mathbb{F}_q^{MN}$. Then the k^{th} symbol of the codeword is $c_k = \boldsymbol{x}^\mathsf{T}\boldsymbol{L}_k$. Note that the i^{th} receiver queries the subvector $\boldsymbol{c}_{R_i} = (c_k, k \in R_i)$, and utilizes the side information $\boldsymbol{x}_{\mathcal{K}_i} = (x_j, j \in \mathcal{K}_i)$, to decode the demand $\boldsymbol{x}_{\mathcal{D}_i} = (x_j, j \in \mathcal{D}_i)$.

A. Locally decodable linear index codes

We observe that the proofs of Lemmas 3.1 and 4.3 and Corollary 4.4 of [11] can be directly adapted to the scenario of locally decodable index codes, immediately yielding the following constraints on the encoder matrix. Let $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_{MN}$ be the standard basis of $\mathbb{F}_q^{MN}$.

Theorem 1. *For each $i \in [N]$, let Rx_i query the subvector $\boldsymbol{c}_{R_i}$ of the codeword $\boldsymbol{c} = \boldsymbol{x}^\mathsf{T}\boldsymbol{L}$ and have the side information $\boldsymbol{x}_{\mathcal{K}_i}$. Then Rx_i can decode its demand $\boldsymbol{x}_{\mathcal{D}_i}$ if and only if for each $j \in \mathcal{D}_i$ there exists a $\boldsymbol{u}_j \in \mathbb{F}_q^{MN}$ that $\boldsymbol{u}_j \lhd \mathcal{K}_i$ and $\boldsymbol{u}_j + \boldsymbol{e}_j \in \mathsf{span}(\boldsymbol{L}_k, k \in R_i)$.*

We will say that $\boldsymbol{L} \in \mathbb{F}_q^{MN\times\ell}$ is a valid encoder matrix corresponding to the queries $R_1, \ldots, R_N \subseteq [\ell]$ if it satisfies the criterion stated in Theorem 1 for decodability at all the receivers.

Observe that among the component symbols in the codeword $\boldsymbol{c} = (c_1, \ldots, c_\ell)^\mathsf{T}$, some are queried exactly once, i.e., queried by a single receiver, and the other symbols are queried by multiple receivers in the network. Let $\mathcal{S}_i = R_i \setminus (R_1 \cup \cdots \cup R_{i-1} \cup R_{i+1} \cdots \cup R_N)$ denote the index set of coded symbols that are queried only by Rx_i. Also, let $\mathcal{M}_i = R_i \cap (R_1 \cup \cdots \cup R_{i-1} \cup R_{i+1} \cdots \cup R_N)$ denote the index set of coded symbols that are queried by Rx_i and at least one other receiver. Note that $R_i = \mathcal{S}_i \cup \mathcal{M}_i$ for each $i \in [N]$.

The following result shows that certain entries of the encoder matrix can be set to be equal to zero without affecting the locality or rate of the index code.

Theorem 2. *Let $\boldsymbol{L} \in \mathbb{F}_q^{MN\times\ell}$ be a valid encoding matrix corresponding to the receivers' queries $R_1, \ldots, R_N$. Then there exists a valid encoding matrix $\boldsymbol{L}' \in \mathbb{F}_q^{MN\times\ell}$ for the queries $R_1, \ldots, R_N$ such that for each $i \in [N]$*

$$\boldsymbol{L}'_k \lhd \mathcal{D}_i \text{ for all } k \in \mathcal{S}_i.$$

Proof: See Appendix A. ∎

The new index code guaranteed by Theorem 2 employs the same code length and the same set of queries as the given index code. Hence, the broadcast rate β, overall locality r and the average locality r_{avg} of the new code are identical to those of the given index code. Additionally, the new code guarantees that for any $i \in [N]$ any codeword symbol c_k, $k \in \mathcal{S}_i$, queried only by Rx_i, can be expressed as a linear combination of the demands of Rx_i. Since any valid encoder matrix can be modified to satisfy this property using Theorem 2, in the sequel, without loss of generality, we will only consider encoder matrices $\boldsymbol{L}$ that satisfy

$$\boldsymbol{L}_k \lhd \mathcal{D}_i \text{ for all } k \in \mathcal{S}_i \text{ and } i \in [N]. \tag{2}$$

Further, without loss of generality, we will assume that for each $i \in [N]$, the vectors $\boldsymbol{L}_k$, $k \in R_i$, are linearly independent. If this is not the case, then at least one of the codeword symbols c_k queried by Rx_i is a linear combination of the other queried symbols $\boldsymbol{c}_{R_i\setminus\{k\}}$. Reducing the index set of the queries of Rx_i from R_i to $R_i\setminus\{k\}$ does not affect decodability at Rx_i since c_k can be reconstructed from $\boldsymbol{c}_{R_i\setminus\{k\}}$. Note that this reduction in the queries does not increase the value of either the overall locality r or the average locality r_{avg} of the index code. This process can be repeated till the columns of $\boldsymbol{L}$ corresponding to the queries of each of the receivers are linearly independent. Finally, any codeword symbol that is not queried by any of the receivers can be removed from the transmission since this symbol will not be used for decoding.

Let us denote the index set of codeword symbols that are queried exactly once by

$$\mathcal{S} = \mathcal{S}_1 \cup \cdots \cup \mathcal{S}_N. \tag{3}$$

Note that $\mathcal{S}_i \cap \mathcal{S}_j = \phi$ for any $i \neq j$. Hence, $|\mathcal{S}| = \sum_{i=1}^N |\mathcal{S}_i|$. Let

$$\mathcal{M} = \mathcal{M}_1 \cup \cdots \cup \mathcal{M}_N \tag{4}$$

denote the index set corresponding to the codeword symbols that have been queried by more than one receiver. Observe that $\mathcal{S} \cap \mathcal{M} = \phi$ and $\mathcal{S} \cup \mathcal{M} = [\ell]$.

Lemma 1. *For any valid index code for message length M, number of receivers N, rate β and average locality r_{avg},*

$$|\mathcal{S}| \geq M(2\beta - Nr_{\text{avg}}),$$

where $|\mathcal{S}|$ is the number of codeword symbols that have been queried exactly once.

Proof: We will count the total number of queries made by all the receivers in two different ways and relate these expressions to arrive at the statement of this lemma.

The number of queries made by Rx_i is $|R_i|$. Hence, the total number of queries made by all the receivers is $\sum_{i\in[N]} |R_i|$. From (1) this is equal to MNr_{avg}. The number of times a codeword symbol c_k is queried is equal to 1 if $k \in \mathcal{S}$, and is at least 2 if $k \in \mathcal{M}$. Thus the total number of queries satisfies

$$\begin{aligned} MNr_{\text{avg}} &\geq \sum_{k\in\mathcal{S}} 1 + \sum_{k\in\mathcal{M}} 2 = |\mathcal{S}| + 2|\mathcal{M}| \\ &= |\mathcal{S}| + 2(\ell - |\mathcal{S}|) = 2\ell - |\mathcal{S}|, \end{aligned}$$

where we have used the fact $|\mathcal{S}| + |\mathcal{M}| = \ell$. Substituting $\ell = M\beta$ in the above inequality, we arrive at $MNr_{\text{avg}} \geq 2M\beta - |\mathcal{S}|$ thereby proving the lemma. ∎

B. Scalar linear index codes with local decodability

We will now derive a few results that hold for the case $M = 1$. Note that, in this case $\mathcal{D}_i = \{i\}$ and $\mathcal{K}_i = K_i$ for all $i \in [N]$. The locality of each receiver $r_i = |R_i|$ is an integer, and so is the overall locality $r = \max_i r_i$.

A matrix $\boldsymbol{A} \in \mathbb{F}_q^{N\times N}$ *fits* $G = (\mathcal{V}, \mathcal{E})$ if the diagonal elements of $\boldsymbol{A}$ are all equal to 1, and the $(j,i)^{\text{th}}$ entry of $\boldsymbol{A}$ is zero if $j \notin K_i$, i.e., if $(i,j) \notin \mathcal{E}$. The *minrank* of G over $\mathbb{F}_q$ is the minimum among the ranks of all possible matrices $\boldsymbol{A} \in \mathbb{F}_q^{N\times N}$ that fit G, and is denoted as $\text{minrk}_q(G)$. It is known that the smallest possible scalar linear index coding rate is equal to $\text{minrk}_q(G)$ [1], [11]. We also know from [1], [11] that a matrix $\boldsymbol{L}$ is a valid encoder matrix for G if and only if for each receiver $i \in [N]$, there exists a vector $\boldsymbol{u}_i \in \mathbb{F}_q^N$ such that $\boldsymbol{u}_i \lhd K_i$ and $\boldsymbol{u}_i + \boldsymbol{e}_i \in \mathcal{C}(\boldsymbol{L})$, where $\mathcal{C}$ denotes the column span of a matrix. If $\boldsymbol{L}$ is a valid encoder matrix, stacking these vectors we obtain the $N \times N$ matrix

$$\boldsymbol{A} = [\boldsymbol{u}_1 + \boldsymbol{e}_1 \;\; \boldsymbol{u}_2 + \boldsymbol{e}_2 \;\; \cdots \;\; \boldsymbol{u}_N + \boldsymbol{e}_N].$$

Notice that $\boldsymbol{A}$ fits G and $\mathcal{C}(\boldsymbol{A}) \subseteq \mathcal{C}(\boldsymbol{L})$. We will say that $\boldsymbol{A}$ is a *fitting matrix corresponding to the encoder matrix* $\boldsymbol{L}$.

Suppose $\boldsymbol{L} \in \mathbb{F}_q^{N\times\ell}$ is a valid scalar linear encoder and the queries of the N receivers are $R_1, \ldots, R_N \subseteq [\ell]$. From Theorem 1, for each $i \in [N]$, there exists a vector $\boldsymbol{u}_i \lhd K_i$ such that $\boldsymbol{u}_i + \boldsymbol{e}_i \in \mathsf{span}(\boldsymbol{L}_k, k \in R_i)$. Thus, there exist scalars $\alpha_{i,k}$, $k \in R_i$, such that $\boldsymbol{u}_i + \boldsymbol{e}_i = \sum_{k\in R_i} \alpha_{i,k}\boldsymbol{L}_k$. The i^{th} receiver decodes its demand by computing $\sum_{k\in R_i} \alpha_{i,k} c_k - \boldsymbol{x}^{\mathsf{T}}\boldsymbol{u}_i$, which is equal to

$$\sum_{k\in R_i} \alpha_{i,k}\boldsymbol{x}^{\mathsf{T}}\boldsymbol{L}_k - \boldsymbol{x}^{\mathsf{T}}\boldsymbol{u}_i = \boldsymbol{x}^{\mathsf{T}}(\boldsymbol{u}_i + \boldsymbol{e}_i) - \boldsymbol{x}^{\mathsf{T}}\boldsymbol{u}_i = x_i.$$

Notice that the receiver can compute $\boldsymbol{x}^{\mathsf{T}}\boldsymbol{u}_i$ using its side information since $\mathsf{supp}(\boldsymbol{u}_i) \subseteq K_i$. We will assume that each scalar $\alpha_{i,k}$ is non-zero since if $\alpha_{i,k} = 0$ the receiver does not need to query the coded symbol c_k. Finally, notice that stacking the vectors $\boldsymbol{u}_i + \boldsymbol{e}_i$, $i \in [N]$, we obtain a fitting matrix $\boldsymbol{A}$ corresponding to $\boldsymbol{L}$.

For certain choices of $S \subseteq [N]$, we will now relate the sizes of R_i, $i \in S$, and their union $\cup_{i\in S} R_i$. Let $\mathcal{N}(\boldsymbol{A})$ denote the null space of $\boldsymbol{A}$.

Lemma 2. *Let $\boldsymbol{A}$ be any fitting matrix corresponding to a valid scalar linear encoder $\boldsymbol{L}$, and $S \subseteq [N]$ be such that S is the support of a non-zero vector in $\mathcal{N}(\boldsymbol{A})$ and the vectors $\boldsymbol{L}_k$, $k \in \cup_{i\in S} R_i$, are linearly independent. Then*

$$\sum_{i\in S} |R_i| \;\geq\; 2\left|\bigcup_{i\in S} R_i\right|.$$

Proof: Denote the columns of $\boldsymbol{A}$ by $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_N$. Let $\boldsymbol{z} \in \mathcal{N}(\boldsymbol{A}) \setminus \{\boldsymbol{0}\}$ be such that $S = \mathsf{supp}(\boldsymbol{z})$. Since $\boldsymbol{A}\boldsymbol{z} = \boldsymbol{0}$, we have $\sum_{i\in S} z_i \boldsymbol{A}_i = \boldsymbol{0}$, where the components z_i, $i \in S$, of the vector $\boldsymbol{z}$ are non-zero. Notice that there exist non-zero scalars $\alpha_{i,k}$ such that $\boldsymbol{A}_i = \sum_{k\in R_i} \alpha_{i,k}\boldsymbol{L}_k$. Hence, we have

$$\boldsymbol{0} = \sum_{i\in S} z_i \boldsymbol{A}_i = \sum_{i\in S}\sum_{k\in R_i} z_i \alpha_{i,k}\boldsymbol{L}_k.$$

All the scalars $z_i\alpha_{i,k}$ in the above linear combination are non-zero, and the set of vectors $\boldsymbol{L}_k$, $k \in \cup_{i\in S} R_i$, appearing in this linear combination are linearly independent. Hence, this linear combination is zero only if each $\boldsymbol{L}_k$, where $k \in \cup_{i\in S} R_i$, appears at least twice in the expansion $\sum_{i\in S}\sum_{k\in R_i} z_i\alpha_{i,k}\boldsymbol{L}_k$, i.e., only if each $k \in \cup_{i\in S} R_i$ is contained in at least two distinct sets R_i and R_j, $i \neq j$ and $i, j \in S$. Then a simple counting argument leads to the statement of this lemma. ∎

The following result can be used to manipulate the bound in Lemma 2 to derive explicit lower bounds on locality.

For any $S \subseteq [N]$, let G_S denote the subgraph of G induced by the vertices in S, i.e., the vertex set of G is S and edge set is $\{(i,j) \in \mathcal{E} | i, j \in S\}$. The subgraph G_S is the side information graph of the index coding problem obtained by restricting the index coding problem G to the messages x_i, $i \in S$.

Lemma 3. *Let $\boldsymbol{L}$ be a valid scalar linear index code for G with receiver queries $R_1, \ldots, R_N$. For any $S \subseteq [N]$, we have*

$$\left|\bigcup_{i\in S} R_i\right| \geq \text{minrk}_q(G_S).$$

Proof: The submatrix $\boldsymbol{L}_S$ of $\boldsymbol{L}$ consisting of the rows indexed by S is a valid encoder matrix for the index coding problem G_S. Since, the receivers $i \in S$ query only the coded symbols with indices $k \in \cup_{i\in S} R_i$, the submatrix of $\boldsymbol{L}_S$ consisting of the columns with indices in $\cup_{i\in S} R_i$ is also a valid scalar linear encoder for G_S. Hence, the codelength $|\cup_{i\in S} R_i|$ of this index code is lower bounded by $\text{minrk}_q(G_S)$. ∎

The next result follows immediately from Lemmas 2 and 3.

Corollary 1. *If $\boldsymbol{L}$ is an optimal scalar linear encoder for G, i.e., has codelength equal to $\text{minrk}_q(G)$, $\boldsymbol{A}$ is a fitting matrix corresponding to $\boldsymbol{L}$ and $\boldsymbol{z} \in \mathcal{N}(\boldsymbol{A}) \setminus \{\boldsymbol{0}\}$, then*

$$\sum_{i\in S} r_i \;\geq\; 2\,\text{minrk}_q(G_S),$$

where $S = \mathsf{supp}(\boldsymbol{z})$.

Proof: The matrix $\boldsymbol{L}$ has linearly independent columns since the number of columns ℓ of $\boldsymbol{L}$ satisfies

$$\ell = \text{minrk}_q(G) \leq \text{rank}(\boldsymbol{A}) \leq \text{rank}(\boldsymbol{L}) \leq \ell.$$

The corollary holds since $r_i = |R_i|$ for scalar linear codes and the vectors $\boldsymbol{L}_k$, $k \in \cup_{i\in S} R_i$ satisfy the conditions of Lemma 2. ∎

IV. Vector Linear Coding for Directed Cycles

In this section, we will consider the index coding problem where G is a directed N-cycle, i.e., for $i = 1, \ldots, N-1$, $K_i = \{i+1\}$ and $K_N = \{1\}$. For $N \geq 3$, and over any finite field $\mathbb{F}_q$, we will show that

$$\beta^*_{G,q}(r) = \max\left\{N - 1, \frac{N(N-1-r)}{N-2}\right\}, \quad r \geq 1. \qquad (5)$$

Note that locality $r \geq 1$ for any valid index coding scheme, and hence, $\beta^*_{G,q}(r)$ is defined for $r \geq 1$ only. The trade-off

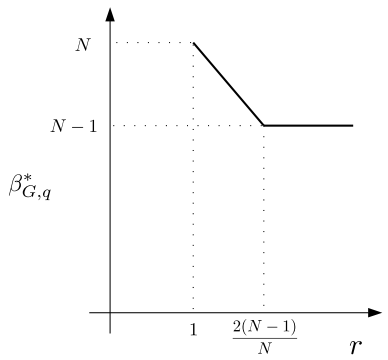

Fig. 1. The locality-rate trade-off of linear index codes for the directed N-cycle index coding problem.

between rate and locality is shown in Fig. 1. Sections IV-A and IV-B provide the proofs for the converse and achievability, respectively, of this rate-locality trade-off. The smallest locality at which the rate $N-1$, which is the minimum possible rate, is achievable is $r = 2(N-1)/N$. This locality is achievable if the message length M of the vector linear code is chosen carefully. In Section IV-C, we provide a detailed analysis of the effect of the message length M on the locality r when the broadcast rate is $N-1$.

A. Converse

In this subsection we will show that $\beta^*_{G,q}(r)$ is lower bounded by both $N-1$ and $N(N-1-r)/(N-2)$. It is clear that $\beta^*_{G,q}(r) \geq N-1$, since even without any locality constraints the smallest possible broadcast rate for the directed N-cycle is $N-1$. To complete the converse, we only need to show that $\beta^*_{G,q}(r) \geq N(N-1-r)/(N-2)$.

Suppose that the encoding matrix $\boldsymbol{L}$ is valid with respect to a set of queries $R_1, \ldots, R_N$. From Lemma 1, the number of codeword symbols queried exactly once

$$|\mathcal{S}| = \sum_{i \in [N]} |\mathcal{S}_i| \geq M(2\beta - Nr_{\text{avg}}).$$

Hence there exists an $i \in [N]$ such that $|\mathcal{S}_i| \geq M(2\beta - Nr_{\text{avg}})/N$. Since G is a directed cycle, without loss of generality, let us assume that

$$|\mathcal{S}_N| \geq M(2\beta - Nr_{\text{avg}})/N. \tag{6}$$

We now relate this lower bound on $|\mathcal{S}_N|$ to the rank of $\boldsymbol{L}$ to complete the converse.

For $i = 1, \ldots, N-1$, we have $\mathcal{K}_i = \{iM+1, iM+2, \ldots, (i+1)M\}$ and $\mathcal{D}_i = \{(i-1)M+1, \ldots, iM\}$. From Theorem 1, for each $j \in \mathcal{D}_i$ there exists a $\boldsymbol{u}_j \lhd \mathcal{K}_i$ such that $\boldsymbol{e}_j + \boldsymbol{u}_j \in \mathcal{C}(\boldsymbol{L})$, where $\mathcal{C}(\boldsymbol{L})$ denotes the column span of the matrix $\boldsymbol{L}$. Considering the first $N-1$ receivers $i = 1, \ldots, N-1$ and each of their demands $j \in \mathcal{D}_i$, we obtain $M(N-1)$ such vectors $\boldsymbol{e}_j + \boldsymbol{u}_j$, all which lie in $\mathcal{C}(\boldsymbol{L})$. Arranging these vectors into a matrix of size $MN \times M(N-1)$ we arrive at

$$\begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{C}_1 & \boldsymbol{I} & \cdots & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{C}_2 & \cdots & \boldsymbol{0} & \boldsymbol{0} \\ \vdots & \vdots & & \vdots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{C}_{N-2} & \boldsymbol{I} \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{C}_{N-1} \end{bmatrix}, \tag{7}$$

where each of the submatrices is of size $M \times M$. Note that the columns of this matrix are linearly independent.

Now considering Rx_N, we note that $\mathcal{D}_N = \{(N-1)M+1, \ldots, MN\}$. The coded symbols $\boldsymbol{x}^{\mathsf{T}}\boldsymbol{L}_k$, $k \in \mathcal{S}_N$, are queried only by Rx_N. From (2), we deduce that $\boldsymbol{L}_k \lhd \mathcal{D}_N$ for all $k \in \mathcal{S}_N$, i.e., $\mathsf{supp}(\boldsymbol{L}_k) \subseteq \mathcal{D}_N$. Since the set of vectors $\{\boldsymbol{L}_k \,|\, k \in R_N\}$ is linearly independent and $\mathcal{S}_N \subseteq R_N$, we observe that the vectors $\boldsymbol{L}_k$, $k \in \mathcal{S}_N$, are linearly independent as well. Note that each of these vectors is a column of $\boldsymbol{L}$ and hence lies in $\mathcal{C}(\boldsymbol{L})$. Appending these $|\mathcal{S}_N|$ vectors as columns to the matrix in (7), we arrive at the block matrix

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{C}_1 & \boldsymbol{I} & \cdots & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{C}_2 & \cdots & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \vdots & \vdots & & \vdots & \vdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{I} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{C}_{N-2} & \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{0} & \boldsymbol{C}_{N-1} & \boldsymbol{B}_N \end{bmatrix},$$

where $\boldsymbol{B}_N$ is an $M \times |\mathcal{S}_N|$ matrix with linearly independent columns. Note that each column of $\boldsymbol{A}$ lies in $\mathcal{C}(\boldsymbol{L})$, i.e., $\mathcal{C}(\boldsymbol{A}) \subseteq \mathcal{C}(\boldsymbol{L})$, and the columns of $\boldsymbol{A}$ are linearly independent, i.e., $\text{rank}(\boldsymbol{A}) = M(N-1) + |\mathcal{S}_N|$. Thus, we have

$$\begin{aligned} \ell &\geq \text{rank}(\boldsymbol{L}) \geq \text{rank}(\boldsymbol{A}) && \text{(since } \mathcal{C}(\boldsymbol{L}) \supseteq \mathcal{C}(\boldsymbol{A})\text{)} \\ &= M(N-1) + |\mathcal{S}_N| \\ &\geq M(N-1) + M(2\beta - Nr_{\text{avg}})/N && \text{(using (6))} \end{aligned}$$

Using the fact that the broadcast rate $\beta = \ell/M$, the above inequality yields $\beta \geq (N-1) + (2\beta - Nr_{\text{avg}})/N$, which upon manipulation results in

$$\beta \geq N(N-1-r_{\text{avg}})/(N-2). \tag{8}$$

Since $r_{\text{avg}} \leq r$, we arrive at $\beta \geq \frac{N(N-1-r)}{N-2}$.

B. Achievability

In this subsection we show that the trade-off in (5) is achievable using linear index codes. We will show that the points $(r, \beta) = (1, N)$ and $(2(N-1)/N, N-1)$ are achievable. Then any point on the line segment $\beta = N(N-1-r)/(N-2)$, $1 \leq r \leq 2(N-1)/N$ can be achieved using time sharing between these two schemes. The achievability of the points $\beta = N-1$ and $r > 2(N-1)/N$ will follow immediately since the rate $N-1$ is already achievable with $r = 2(N-1)/N$.

1) Achieving $r = 1$, $\beta = N$: The point $(r, \beta) = (1, N)$ can be achieved trivially using the uncoded scheme, i.e., the transmitted codeword equals the message vector $\boldsymbol{c} = \boldsymbol{x} \in \mathbb{F}_q^{MN}$. Each receiver Rx_i queries $\boldsymbol{c}_{\mathcal{D}_i} = \boldsymbol{x}_{\mathcal{D}_i}$ to meet its demand. The side information available at the receivers is not utilized by this scheme. Since the code length $\ell = MN$, we have $\beta = N$ and since $|R_i| = |\mathcal{D}_i| = M$, we have $r = 1$. Also note that this is a linear index code corresponding to the encoding matrix $\boldsymbol{L} = \boldsymbol{I}$.

2) Achieving $r = 2(N-1)/N$, $\beta = N-1$: Example 1 provides a family of N scalar linear codes for G, one for each choice of $i \in [N]$, with rate $\beta = N-1$. The i^{th} code provides localities $r_i = r_{i+1} = 1$ and $r_j = 2$ for all $j \neq i, i+1$, where we interpret $i+1$ as 1 if $i = N$. Using $\boldsymbol{r} = (r_1, r_2, \ldots, r_N)$ to represent the tuple of receiver localities, we observe that rate $N-1$ can be achieved with the following values of $\boldsymbol{r}$

$$\begin{aligned} \boldsymbol{r}_1 = (1,1,2,2,\ldots,2),\ \boldsymbol{r}_2 = (2,1,1,2,\ldots,2),\ldots, \\ \boldsymbol{r}_{N-1} = (2,2,\ldots,2,1,1),\ \boldsymbol{r}_N = (1,2,\ldots,2,1). \end{aligned} \tag{9}$$

If N is an odd integer, we time share the N scalar linear codes corresponding to $\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N$. Observe that the overall scheme is a vector linear code for message length $M = N$, rate $N-1$ and locality $r = r_{\text{avg}} = 2(N-1)/N$.

If N is an even integer, we time share $N/2$ scalar linear codes corresponding to $\boldsymbol{r}_1, \boldsymbol{r}_3, \ldots, \boldsymbol{r}_{N-1}$, that yields a vector linear code with $M = N/2$, rate $N-1$ and $r = r_{\text{avg}} = 2(N-1)/N$.

C. Dependence of locality on message length

From the achievability scheme in Section IV-B we observed that for $\beta = N-1$, the locality $r = r_{\text{avg}} = 2(N-1)/N$ can be achieved using message length $M = N$ if N is odd, and $M = N/2$ if N is even.

Lemma 4. *Let $N \geq 3$. The message length M of any index code that achieves locality $2(N-1)/N$ for the directed N-cycle satisfies $M \geq N$ if N is odd, and $M \geq N/2$ if N is even.*

Proof: Consider any valid coding scheme with locality $r = 2(N-1)/N$. There exists an $i \in [N]$ such that $\frac{2(N-1)}{N} = r = r_i = \frac{|R_i|}{M}$, that is

$$M = \frac{|R_i|\, N}{2(N-1)}.$$

If N is odd, N and $2(N-1)$ have no common factors, and since M is an integer, we deduce that M must be a multiple of N, i.e., $M \geq N$. If N is even, using the fact $N/2$ and $N-1$ have no common factors we arrive at $M \geq N/2$. ∎

From Lemma 4, it is clear that the minimum M required to attain $r = 2(N-1)/N$ at rate $N-1$ is $M = N$ if N is odd and $M = N/2$ if N is even. We will now derive the optimal locality when the message length is smaller than this quantity, i.e., $M < N$. We do so by analysing the two cases, $M < N/2$ and $N/2 \leq M < N$.

1) Locality when $M < N/2$: From (8) we deduce that for any vector linear scheme of rate $\beta = N-1$, we have

$$r_{\text{avg}} \geq 2(N-1)/N.$$

Thus, $\sum_{i=1}^{N} |R_i| = MNr_{\text{avg}} \geq 2M(N-1)$. It follows that there exists an $i \in [N]$ such that

$$|R_i| \geq \frac{2M(N-1)}{N} = 2M - \frac{M}{N/2}. \tag{10}$$

If $M < N/2$, considering the fact that $|R_i|$ is an integer, we deduce that $|R_i| \geq 2M$. Hence, $r_i = |R_i|/M \geq 2$, and thus, $r \geq 2$. This lower bound on r can be achieved by simply using the scalar linear code of Example 1 M times, leading to a vector linear code for message length M, rate $N-1$ and $r = 2$. Note that this code still achieves the optimal value of average locality $r_{\text{avg}} = 2(N-1)/N$.

2) Locality when $N/2 \leq M < N$: If N is even, the message length $M = N/2$ is sufficient to attain $r = 2(N-1)/N$. Thus it is enough to consider larger values of M, i.e., $N/2 \leq M < N$ only for N odd. From (10) and using the fact that $|R_i|$ is an integer, we arrive at $|R_i| \geq 2M - 1$. Thus,

$$r \geq r_i \geq 2 - \frac{1}{M}.$$

Assuming N is odd, this lower bound on r is achieved by time sharing the M scalar linear codes from Section IV-B corresponding to the tuples of localities

$$\boldsymbol{r}_1, \boldsymbol{r}_3, \ldots, \boldsymbol{r}_{N-3}, \boldsymbol{r}_N, \boldsymbol{r}_2, \boldsymbol{r}_4, \ldots, \boldsymbol{r}_{2M-(N+1)},$$

see (9). It is straightforward to show that this scheme has rate $N-1$, $r = 2 - 1/M$ and $r_{\text{avg}} = 2(N-1)/N$. Note that in the interval $N/2 \leq M < N$, the value of the optimal locality increases with M. Hence, the choice $M = (N+1)/2$ yields the smallest locality in this interval.

V. Scalar Linear Coding when Minrank is $N-1$

We now characterize the optimal localities r and r_{avg} among scalar linear index codes for index coding problems G with $\text{minrk}_q(G) = N-1$. The message length $M = 1$ for scalar codes, and hence, the receiver localities r_i and the rate β are integers. Since the minimum scalar coding rate is equal to minrk_q, we are interested in the operating points corresponding to $\beta = N$ and $\beta = N-1$. In the rest of this section we will assume that $\text{minrk}_q(G) = N-1$.

The side information graph G contains at least one directed cycle. Otherwise, G is a directed acyclic graph and its minrank is equal to N [1], a contradiction. Let N_c denote the length of the smallest directed cycle contained in G.

If $N_c = 2$, there exist $i, j \in [N]$ such that $(i,j), (j,i) \in \mathcal{E}$, i.e., $i \in K_j$ and $j \in K_i$. The following scalar linear code attains the minimum possible locality $r = r_{\text{avg}} = 1$ and the minimum possible rate $\beta = N-1$ simultaneously. Transmit $x_i + x_j$ followed by transmitting the remaining $N-2$ information symbols uncoded. Rx_i and Rx_j can decode using $x_i + x_j$, and the remaining receivers query their demands directly from the codeword.

In the rest of this section we will assume that $N_c \geq 3$. Observe that rate $\beta = N$ can be achieved with smallest possible localities $r = r_{\text{avg}} = 1$ using uncoded transmission.

We will now consider the case $\beta = \text{minrk}_q(G) = N-1$. Let $\boldsymbol{L}$ be any scalar encoder matrix with codelength $\ell = N-1$, and $\boldsymbol{A}$ be a corresponding fitting matrix. Since $\ell = \text{minrk}_q(G)$, we have $\ell \leq \text{rank}(\boldsymbol{A}) \leq \text{rank}(\boldsymbol{L}) \leq \ell$, and hence, $\text{rank}(\boldsymbol{A}) = \text{rank}(\boldsymbol{L}) = N-1 = \ell$. Thus, the nullspace of $\boldsymbol{A}$ contains a non-zero vector.

Lemma 5. *If $\boldsymbol{z} \in \mathcal{N}(\boldsymbol{A}) \setminus \{\boldsymbol{0}\}$ and $S = \mathsf{supp}(\boldsymbol{z})$, the subgraph G_S of G induced by the vertices S contains at least one directed cycle.*

Proof: Observe that the $|S| \times |S|$ submatrix $\boldsymbol{A}'$ of $\boldsymbol{A}$ composed of the rows and columns of $\boldsymbol{A}$ with indices in S fits G_S. Since $\boldsymbol{A}\boldsymbol{z} = \boldsymbol{0}$, the columns of $\boldsymbol{A}$ indexed by S are linearly dependent. This implies that the columns of the submatrix $\boldsymbol{A}'$ are linearly dependent as well, and hence, $\text{rank}(\boldsymbol{A}') \leq |S|-1$. It follows that $\text{minrk}_q(G_S) \leq \text{rank}(\boldsymbol{A}') \leq |S|-1$. Clearly G_S is not a directed acyclic graph since otherwise $\text{minrk}_q(G_S) = |S|$. ∎

In order to use Corollary 1, we now derive a lower bound on $\text{minrk}_q(G_S)$.

Lemma 6. *If $\text{minrk}_q(G) = N-1$, then for any $S \subseteq [N]$, $\text{minrk}_q(G_S) \geq |S| - 1$.*

Proof: Consider the following valid scalar linear code for G. Encode the information symbols x_i, $i \in S$, using the optimal scalar linear code for G_S, and use uncoded transmission for the remaining symbols. The length of this code is lower bounded by $\text{minrk}_q(G) = N-1$, hence we obtain $\text{minrk}_q(G_S) + N - |S| \geq N - 1$. ∎

We now prove the main result of this section.

Theorem 3. *If $\text{minrk}_q(G) = N-1$ and the smallest directed cycle in G is of length $N_c \geq 3$, the optimal locality for scalar linear coding for G with rate $N-1$ is*

$$r = 2 \text{ and } r_{\text{avg}} = \frac{N + N_c - 2}{N}.$$

Proof: Converse: Let $\boldsymbol{A}$ be a fitting matrix corresponding to any valid scalar linear code for G with rate $N-1$. Let $\boldsymbol{z} \in \mathcal{N}(\boldsymbol{A}) \setminus \{\boldsymbol{0}\}$ and $S = \mathsf{supp}(\boldsymbol{z})$. From Corollary 1 and Lemma 6, $\sum_{i \in S} r_i \geq 2\,\text{minrk}_q(G_S) \geq 2(|S|-1)$. Using the trivial bound $r_i \geq 1$ for $i \notin S$, we have

$$\begin{aligned}\sum_{i \in [N]} r_i &= \sum_{i \in S} r_i + \sum_{i \notin S} r_i \\ &\geq 2(|S|-1) + N - |S| = N + |S| - 2.\end{aligned}$$

From Lemma 5, we know that G_S contains a cycle, and hence, the number of vertices $|S|$ in G_S is at least N_c. Thus,

$$r_{\text{avg}} = \frac{\sum_{i \in [N]} r_i}{N} \geq \frac{N + |S| - 2}{N} \geq \frac{N + N_c - 2}{N}.$$

Since $N_c \geq 3$, we have $r \geq r_{\text{avg}} \geq (N+1)/N$, and since r is an integer we conclude that $r \geq 2$.

Achievability: Let $C \subseteq [N]$ be the set of vertices that form the smallest directed cycle in G. Note that the subgraph G_C is a directed cycle of length $|C| = N_c$. We encode the symbols x_i, $i \in C$, using the scalar linear code given in Example 1 and send the remaining $N - N_c$ symbols uncoded. This achieves the codelength $N-1$. From Example 1, the sum locality within the cycle $\sum_{i \in C} r_i = 2(N_c - 1)$ and the maximum locality within the cycle $\max_{i \in C} r_i = 2$. The locality of the remaining receivers is $r_i = 1$, $i \notin C$. This scheme achieves the optimal values of r and r_{avg} for rate $N-1$. ∎

Remark 1. Corollary V.2 of [10] shows that, over the binary field $q = 2$, if $\text{minrk}_q(G) = N-1$, then a scalar linear coding rate of $N-1$ is achievable for any choice of locality $r \geq 2$. In contrast, our results show that $r = 2$ is optimal (if $N_c \geq 3$) and also provide the optimal value of the average locality r_{avg} for scalar linear codes over an arbitrary finite field $\mathbb{F}_q$.

Appendix A
Proof of Theorem 2

We will design a new encoding matrix $\boldsymbol{L}'$ by modifying the subset of the columns of the given matrix $\boldsymbol{L}$ corresponding to the column indices $\mathcal{S}_1 \cup \cdots \cup \mathcal{S}_N$. For the remaining indices $k \in \mathcal{M}_1 \cup \cdots \cup \mathcal{M}_N$, the k^{th} columns of $\boldsymbol{L}$ and $\boldsymbol{L}'$ are equal, i.e., $\boldsymbol{L}_k = \boldsymbol{L}'_k$. For an arbitrary $i \in [N]$, we will now explain the construction of the column vectors $\boldsymbol{L}'_k$, $k \in \mathcal{S}_i$. Since the symbols $\boldsymbol{x}^{\mathsf{T}}\boldsymbol{L}'_k$, $k \in \mathcal{S}_i$, are queried only by Rx_i and are unused by other receivers, we only need to consider the constraints that are imposed by the demands of Rx_i while designing the column vectors $\boldsymbol{L}'_k$, $k \in \mathcal{S}_i$.

We will introduce the notation which will be used in the rest of the proof. For any $E \subseteq [MN]$, let $U_E = \mathsf{span}(\boldsymbol{e}_k, k \in E)$, i.e., U_E is the subspace of all vectors whose support is a subset of E. For any $F \subseteq [\ell]$, let $V_F = \mathsf{span}(\boldsymbol{L}_k, k \in F)$ and $V'_F = \mathsf{span}(\boldsymbol{L}'_k, k \in F)$. Since $R_i = \mathcal{S}_i \cup \mathcal{M}_i$, we have $V_{R_i} = V_{\mathcal{S}_i} + V_{\mathcal{M}_i}$ and $V'_{R_i} = V'_{\mathcal{S}_i} + V'_{\mathcal{M}_i}$, where the addition corresponds to sum of subspaces. From Theorem 1 and using the fact that $\boldsymbol{L}$ is a valid encoder matrix, we have $\boldsymbol{e}_j \in V_{R_i} + U_{\mathcal{K}_i}$, for all $j \in \mathcal{D}_i$, i.e., we have

$$U_{\mathcal{D}_i} = \mathsf{span}(\boldsymbol{e}_j, j \in \mathcal{D}_i) \subseteq V_{R_i} + U_{\mathcal{K}_i} = V_{\mathcal{S}_i} + V_{\mathcal{M}_i} + U_{\mathcal{K}_i}. \tag{11}$$

Again using Theorem 1, we observe that $\boldsymbol{L}'$ allows Rx_i to decode its demand if and only if

$$U_{\mathcal{D}_i} \subseteq V'_{R_i} + U_{\mathcal{K}_i} = V'_{\mathcal{S}_i} + V'_{\mathcal{M}_i} + U_{\mathcal{K}_i}. \tag{12}$$

Using the validity of the encoder matrix $\boldsymbol{L}$, we will first lower bound $|\mathcal{S}_i|$ which is the number of coded symbols queried uniquely by Rx_i. From (11), we obtain

$$U_{\mathcal{D}_i} = (V_{\mathcal{S}_i} + V_{\mathcal{M}_i} + U_{\mathcal{K}_i}) \cap U_{\mathcal{D}_i}. \tag{13}$$

Let $V_{\mathsf{in}} = (V_{\mathcal{M}_i} + U_{\mathcal{K}_i}) \cap U_{\mathcal{D}_i}$ denote the subspace of $V_{\mathcal{M}_i} + U_{\mathcal{K}_i}$ contained in $U_{\mathcal{D}_i}$, and let V_{out} be any subspace such that $V_{\mathsf{out}} \cap U_{\mathcal{D}_i} = \{\boldsymbol{0}\}$ and $V_{\mathsf{in}} + V_{\mathsf{out}} = V_{\mathcal{M}_i} + U_{\mathcal{K}_i}$. Continuing from (13), we claim that

$$\begin{aligned}U_{\mathcal{D}_i} &= (V_{\mathcal{S}_i} + V_{\mathsf{in}} + V_{\mathsf{out}}) \cap U_{\mathcal{D}_i} &(14)\\ &= ((V_{\mathcal{S}_i} + V_{\mathsf{out}}) \cap U_{\mathcal{D}_i}) + V_{\mathsf{in}}. &(15)\end{aligned}$$

It is clear that the subspace in (14) contains the subspace in (15) since $V_{\mathsf{in}} \subseteq U_{\mathcal{D}_i}$. To prove that (14) is contained in (15), assume that $\boldsymbol{y} \in V_{\mathcal{S}_i} + V_{\mathsf{out}}$ and $\boldsymbol{z} \in V_{\mathsf{in}}$ are such that $\boldsymbol{y} + \boldsymbol{z} \in U_{\mathcal{D}_i}$. Since $\boldsymbol{z} \in V_{\mathsf{in}} \subseteq U_{\mathcal{D}_i}$ and $\boldsymbol{y} + \boldsymbol{z} \in U_{\mathcal{D}_i}$, we conclude that $\boldsymbol{y} \in U_{\mathcal{D}_i}$ as well. Thus $\boldsymbol{y} \in (V_{\mathcal{S}_i} + V_{\mathsf{out}}) \cap U_{\mathcal{D}_i}$, and hence, $\boldsymbol{y} + \boldsymbol{z} \in ((V_{\mathcal{S}_i} + V_{\mathsf{out}}) \cap U_{\mathcal{D}_i}) + V_{\mathsf{in}}$.

Considering the dimensions of the subspaces in (15), we have

$$\dim(U_{\mathcal{D}_i}) \leq \dim((V_{\mathcal{S}_i} + V_{\mathsf{out}}) \cap U_{\mathcal{D}_i}) + \dim(V_{\mathsf{in}}). \quad (16)$$

In order to proceed with the proof of the theorem, we will now show that $\dim((V_{\mathcal{S}_i} + V_{\mathsf{out}}) \cap U_{\mathcal{D}_i}) \leq |\mathcal{S}_i|$. To do so, assume that $\boldsymbol{y}_j + \boldsymbol{z}_j$, $j = 1, \ldots, n$, form a basis for $(V_{\mathcal{S}_i} + V_{\mathsf{out}}) \cap U_{\mathcal{D}_i}$ where $\boldsymbol{y}_j \in V_{\mathcal{S}_i}$ and $\boldsymbol{z}_j \in V_{\mathsf{out}}$ and $n = \dim((V_{\mathcal{S}_i} + V_{\mathsf{out}}) \cap U_{\mathcal{D}_i})$. If the scalars $\alpha_1, \ldots, \alpha_n$ are such that $\sum_j \alpha_j \boldsymbol{y}_j = \boldsymbol{0}$, then

$$\sum_{j=1}^{n} \alpha_j (\boldsymbol{y}_j + \boldsymbol{z}_j) = \sum_{j=1}^{n} \alpha_j \boldsymbol{z}_j \in V_{\mathsf{out}}.$$

Since $\boldsymbol{y}_j + \boldsymbol{z}_j \in U_{\mathcal{D}_i}$, we also observe that $\sum_j \alpha_j (\boldsymbol{y}_j + \boldsymbol{z}_j) \in U_{\mathcal{D}_i}$. Using the fact $V_{\mathsf{out}} \cap U_{\mathcal{D}_i} = \{\boldsymbol{0}\}$, we deduce that $\sum_j \alpha_j (\boldsymbol{y}_j + \boldsymbol{z}_j) = \boldsymbol{0}$, and hence, $\alpha_1 = \cdots = \alpha_n = 0$. We conclude that $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n \in V_{\mathcal{S}_i}$ are linearly independent, and therefore

$$\dim((V_{\mathcal{S}_i} + V_{\mathsf{out}}) \cap U_{\mathcal{D}_i}) = n \leq \dim(V_{\mathcal{S}_i}) \leq |\mathcal{S}_i|. \quad (17)$$

Note that the columns of $\boldsymbol{L}$ and $\boldsymbol{L}'$ corresponding to the column indices $\mathcal{M}_i$ are equal, and hence $V'_{\mathcal{M}_i} = V_{\mathcal{M}_i}$. Using this fact together with (16) and (17), we have

$$\begin{aligned} |\mathcal{S}_i| &\geq \dim((V_{\mathcal{S}_i} + V_{\mathsf{out}}) \cap U_{\mathcal{D}_i}) \\ &\geq \dim(U_{\mathcal{D}_i}) - \dim(V_{\mathsf{in}}) \\ &= \dim(U_{\mathcal{D}_i}) - \dim((V_{\mathcal{M}_i} + U_{\mathcal{K}_i}) \cap U_{\mathcal{D}_i}) \\ &= \dim(U_{\mathcal{D}_i}) - \dim((V'_{\mathcal{M}_i} + U_{\mathcal{K}_i}) \cap U_{\mathcal{D}_i}). \end{aligned}$$

From (12), in order to satisfy the claim of this theorem, it is sufficient to chose the $|\mathcal{S}_i|$ vectors $\boldsymbol{L}'_k$, $k \in \mathcal{S}_i$, such that $\boldsymbol{L}'_k \lhd \mathcal{D}_i$, i.e., $\boldsymbol{L}'_k \in U_{\mathcal{D}_i}$ and

$$V'_{\mathcal{S}_i} + \left((V'_{\mathcal{M}_i} + U_{\mathcal{K}_i}) \cap U_{\mathcal{D}_i}\right) \supseteq U_{\mathcal{D}_i}.$$

This is always possible since the difference in the dimensions of $U_{\mathcal{D}_i}$ and $(V'_{\mathcal{M}_i} + U_{\mathcal{K}_i}) \cap U_{\mathcal{D}_i}$ is at the most $|\mathcal{S}_i|$. One way to construct $\boldsymbol{L}'_k$, $k \in \mathcal{S}_i$, is as follows. We begin with a basis for $(V'_{\mathcal{M}_i} + U_{\mathcal{K}_i}) \cap U_{\mathcal{D}_i}$. These vectors form a linearly independent set in $U_{\mathcal{D}_i}$. We extend this set to a basis for $U_{\mathcal{D}_i}$. The number of additional vectors in this basis is $\dim(U_{\mathcal{D}_i}) - \dim((V'_{\mathcal{M}_i} + U_{\mathcal{K}_i}) \cap U_{\mathcal{D}_i})$. These additional vectors together with $|\mathcal{S}_i| - \dim(U_{\mathcal{D}_i}) + \dim((V'_{\mathcal{M}_i} + U_{\mathcal{K}_i}) \cap U_{\mathcal{D}_i})$ all-zero vectors are chosen as the columns $\boldsymbol{L}'_k$, $k \in \mathcal{S}_i$.

REFERENCES

[1] Z. Bar-Yossef, Y. 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.

[2] Y. Birk and T. Kol, "Informed-source coding-on-demand (ISCOD) over broadcast channels," in *Proc. 17th Annu. Joint Conf. IEEE Computer and Communications Societies (INFOCOM)*, vol. 3, Mar. 1998, pp. 1257–1264.

[3] S. El 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] M. Effros, S. E. Rouayheb, and M. Langberg, "An equivalence between network coding and index coding," *IEEE Transactions on Information Theory*, vol. 61, no. 5, pp. 2478–2487, May 2015.

[5] M. A. Maddah-Ali and U. Niesen, "Fundamental limits of caching," *IEEE Transactions on Information Theory*, vol. 60, no. 5, pp. 2856–2867, May 2014.

[6] A. Mazumdar, "On a duality between recoverable distributed storage and index coding," in *2014 IEEE International Symposium on Information Theory*, June 2014, pp. 1977–1981.

[7] K. Shanmugam and A. G. Dimakis, "Bounding multiple unicasts through index coding and locally repairable codes," in *2014 IEEE International Symposium on Information Theory*, June 2014, pp. 296–300.

[8] I. Haviv and M. Langberg, "On linear index coding for random graphs," in *2012 IEEE International Symposium on Information Theory Proceedings*, July 2012, pp. 2231–2235.

[9] L. Natarajan, P. Krishnan, and V. Lalitha, "On locally decodable index codes," in *2018 IEEE International Symposium on Information Theory (ISIT)*, June 2018, pp. 446–450.

[10] M. Karmoose, L. Song, M. Cardone, and C. Fragouli, "Privacy in index coding: k-limited-access schemes," *CoRR*, vol. abs/1809.08263, 2018. [Online]. Available: http://arxiv.org/abs/1809.08263

[11] S. H. Dau, V. Skachek, and Y. M. Chee, "On the security of index coding with side information," *IEEE Transactions on Information Theory*, vol. 58, no. 6, pp. 3975–3988, June 2012.