Zero-error Function Computation on a Directed Acyclic Network

Ardhendu Tripathy and Aditya Ramamoorthy
Department of Electrical and Computer Engineering
Iowa State University
Ames, IA 50011, USA
Email: {ardhendu, adityar}@iastate.edu

***Abstract*—We study the rate region of variable-length source-network codes that are used to compute a function of messages observed over a network. The particular network considered here is the simplest instance of a directed acyclic graph (DAG) that is not a tree. Existing work on zero-error function computation in DAG networks provides bounds on the *computation capacity*, which is a measure of the amount of communication required per edge in the worst case. This work focuses on the average case: an achievable rate tuple describes the expected amount of communication required on each edge, where the expectation is over the probability mass function of the source messages.**

We describe a systematic procedure to obtain outer bounds to the rate region for computing an arbitrary demand function at the terminal. Our bounding technique works by lower bounding the entropy of the descriptions observed by the terminal conditioned on the function value and by utilizing the Schur-concave property of the entropy function.

I. Introduction

Zero-error function computation over a graphical network using network coding was studied in [1]. There they considered two variants of the communication load on the network, called *worst-case* and *average-case* complexity, depending on whether the probability information of the source messages was used or not. They characterized the rate region of achievable rate tuples that allowed zero-error function computation in tree-networks, each entry in a rate tuple was the rate of a code employed on the corresponding edge in the tree-network. They also made the observation that finding the rate region of a directed acyclic graph (DAG) network is challenging because of multiple paths between a source node and the terminal, which allows for different ways of combining information at the intermediate nodes.

Worst-case zero-error function computation was also studied in [2], where the authors defined the *computation capacity* of a function computation problem instance. This is a generalization of the coding capacity of a network, which is the supremum of the ratio $\frac{k}{n}$ over all achievable (k, n) fractional coding solutions for that communication network. A (k, n) fractional network code is one in which k source messages are block encoded at each encoder and every edge in the network transmits n symbols from the alphabet in one channel use. The authors in [2] characterized the computation capacity of multi-stage tree-networks by finding the necessary and sufficient amount of information that must be transmitted across all graph cuts that separate one or more source nodes from the terminal. Upper bounds on the computation capacity of DAG networks are more complicated and have been obtained in [3], [4]. These upper bounds were shown to be unachievable for a function computation problem on a particular DAG in recent work [5, Sec. V]. A qualitatively different line of work considers the computation of simple functions such as finite-field sum [11]–[13] over arbitrary acyclic networks with multiple terminals. Furthermore, [14]–[16] discuss the multiple unicast problem which is an important special case of function computation.

In this paper, we focus on average-case complexity of computing a general demand function over the DAG network shown in figure 1. We summarize our contributions below.

- For an arbitrary demand function, we give a procedure to obtain lower bounds on the rates that must be used on the edges of the DAG in figure 1 in order to compute the function with zero-error at the terminal. The key fact used in this is the Schur-concavity of entropy and the equivalence relations given in [3]–[5].
- Applying this procedure for the arithmetic sum demand function gives us a tighter lower bound to the sum rate than that implied by the work in [6].

The paper is organized as follows. Section II formulates the problem and defines its rate region. Section III gives a procedure to find an outer bound to the rate region; we evaluate this outer bound for arithmetic sum computation in Section IV. Due to space constraints, some proofs and calculations are omitted and can be found in the full paper [10].

II. Problem formulation

The edges in figure 1 model error-free communication links and are later denoted by an ordered pair of vertices. In what follows, logarithms are to the base 2 unless specified otherwise. Suppose that $\mathcal{Z}$ is the alphabet used for communication, and $|\mathcal{Z}| > 1$. Vertices s_1, s_2, s_3 are the three source nodes that observe random variables X_1, X_2, X_3 respectively, each from a discrete alphabet $\mathcal{A}$ with size $|\mathcal{A}| > 1$. The source r.v.s are assumed to be i.i.d. uniformly distributed over $\mathcal{A}$. Terminal node t wants to compute $B \triangleq f(X_1, X_2, X_3)$ with zero error and zero distortion for a known demand function:

$$f : \mathcal{A} \times \mathcal{A} \times \mathcal{A} \to \mathcal{B}.$$

This work was supported in part by the National Science Foundation (NSF) under grants CCF-1149860, CCF-1320416, CCF-1718470 and DMS-1120597.

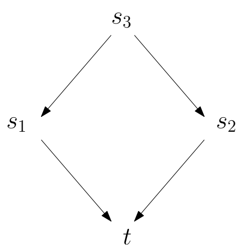

Fig. 1. A directed acyclic network with three sources and one terminal.

An example f is shown in table I. We use variable-length network codes to compute it. To do that, we use the *Source-Network Code* as described in [7] and use it in the function computation setting described above. The quantity of interest here is the rate region $\mathcal{R}$, which is a region containing all achievable rate tuples $\boldsymbol{R}$. Each rate tuple has four components, one for each edge in the network. We define the source-network code and the admissible rate tuples below. We denote a k-length vector r.v. as $B^k \triangleq (B^{(1)}, B^{(2)}, \ldots, B^{(k)})$, where $B^{(i)}$ indicates the i-th r.v. in B^k. Lowercase boldface are used to denote vector realizations, i.e., $\boldsymbol{b} \triangleq (b^{(1)}, b^{(2)}, \ldots, b^{(k)})$.

Definition 1: Let $\mathcal{Z}^*$ denote the set of all finite-length sequences with alphabet $\mathcal{Z}$. A source-network code $\mathcal{C}_{f,k}$ for computing B^k in the network of Figure 1 consists of:

1) Encoding functions for edges in Figure 1:

$$\begin{aligned}
\phi_{(s_3,s_1)}(X_3^k) &: \mathcal{A}^k \to \mathcal{Z}^* \\
\phi_{(s_3,s_2)}(X_3^k) &: \mathcal{A}^k \to \mathcal{Z}^* \\
\phi_{(s_1,t)}\left(X_1^k, \phi_{(s_3,s_1)}(X_3^k)\right) &: \mathcal{A}^k \times \mathcal{Z}^* \to \mathcal{Z}^* \\
\phi_{(s_2,t)}\left(X_2^k, \phi_{(s_3,s_2)}(X_3^k)\right) &: \mathcal{A}^k \times \mathcal{Z}^* \to \mathcal{Z}^*
\end{aligned}$$

For brevity, we denote $\phi_{(s_1,t)}\left(X_1^k, \phi_{(s_3,s_1)}(X_3^k)\right)$ by the random variable $\boldsymbol{Z}_1$, and similarly define the r.v.s $\boldsymbol{Z}_2, \boldsymbol{Z}_{31}, \boldsymbol{Z}_{32}$.

2) Decoding function for terminal t: $\psi_t : \mathcal{Z}^* \times \mathcal{Z}^* \to \mathcal{B}^k$ is such that $\Pr\{\psi_t(\boldsymbol{Z}_1, \boldsymbol{Z}_2) \neq B^k\} = 0$, where the sample space of the probability consists of all realizations of the i.i.d. messages X_1^k, X_2^k and X_3^k.

Thus the outputs of the encoders are variable length, and the terminal has a decoder that takes in a pair of variable length inputs and returns the block of k function computations on the message tuple. The rate tuple of a source-network code is defined below, taking into account the different alphabets in which the messages and the codewords reside.

Definition 2: $\boldsymbol{R} = (R_{31}, R_{32}, R_1, R_2)$ is an admissible rate tuple for the code $\mathcal{C}_{f,k}$ if for any $\epsilon > 0$ there exists a sufficiently large k such that

$$\mathbb{E}\,\ell(\boldsymbol{Z}_1) \log|\mathcal{Z}| \leq k(R_1 + \epsilon) \log|\mathcal{A}|,$$

where $\mathbb{E}\,\ell(\boldsymbol{Z}_1)$ is the expected length (in symbols from $\mathcal{Z}$, over the probability mass function of the source messages) of the codeword $\boldsymbol{Z}_1$. A similar definition is used for the rates R_{31}, R_{32} and R_2.

III. Outer bound to the rate region

We use $\mathcal{C}^\star_{\text{NS}} : \mathcal{S}_{\boldsymbol{Z}} \to \mathcal{Z}^\star$ to denote a non-singular code for a r.v. $\boldsymbol{Z}$ supported on $\mathcal{S}_{\boldsymbol{Z}}$ having minimum expected length.

Lemma 1 (Adapted from Theorem 3 in [8]): Let $H_{|\mathcal{Z}|}(\cdot)$ denote entropy in base $|\mathcal{Z}|$. The expected length of the best non-singular code $\mathcal{C}^\star_{\text{NS}}(\boldsymbol{Z})$ for a r.v. $\boldsymbol{Z}$ has the lower bound:

$$\mathbb{E}\,\ell\left(\mathcal{C}^\star_{\text{NS}}(\boldsymbol{Z})\right) \geq H_{|\mathcal{Z}|}(\boldsymbol{Z}) - 2\log_{|\mathcal{Z}|}\left(H_{|\mathcal{Z}|}(\boldsymbol{Z}) + |\mathcal{Z}|\right).$$

Since the identity mapping is also a non-singular code for $\boldsymbol{Z}$,

$$\mathbb{E}\,\ell(\boldsymbol{Z}) = \sum_{\boldsymbol{z}} \Pr\{\boldsymbol{Z} = \boldsymbol{z}\}\ell(\boldsymbol{z}) \geq \mathbb{E}\,\ell\left(\mathcal{C}^\star_{\text{NS}}(\boldsymbol{Z})\right). \tag{1}$$

Lemma 2: Consider an equivalence relation on $\mathcal{A}^k$ for which $\boldsymbol{x}_3 \equiv \boldsymbol{x}_3'$ if and only if for all $(\boldsymbol{x}_1, \boldsymbol{x}_2) \in \mathcal{A}^k \times \mathcal{A}^k$, we have that $f(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3) = f(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3')$. Define the function $g(X_3^k)$ which returns the equivalence class that X_3^k belongs to under the above relation. Then the range of $g(X_3^k)$ is a subset of $\{1, 2, \ldots, |\mathcal{A}|^k\}$ and we have that $R_{31} + R_{32} \geq H(g(X_3^k))/k \log|\mathcal{A}|$.

Proof: Suppose that $H_{|\mathcal{Z}|}(g(X_3^k)|\boldsymbol{Z}_{31}, \boldsymbol{Z}_{32}) > 0$, then one cannot obtain $g(X_3^k)$ from the pair $(\boldsymbol{Z}_{31}, \boldsymbol{Z}_{32})$, i.e., there exist $\boldsymbol{x}_3 \not\equiv \boldsymbol{x}_3'$ but their associated codewords satisfy $\boldsymbol{z}_{31} = \boldsymbol{z}_{31}'$ and $\boldsymbol{z}_{32} = \boldsymbol{z}_{32}'$. There exists a pair $(\boldsymbol{x}_1, \boldsymbol{x}_2) \in \mathcal{A}^k \times \mathcal{A}^k$ such that $f(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3) \neq f(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3')$. However, since $\boldsymbol{z}_{31} = \boldsymbol{z}_{31}'$ and $\boldsymbol{z}_{32} = \boldsymbol{z}_{32}'$, the codewords transmitted on the edges $(s_1, t), (s_2, t)$ in the two cases satisfy $\boldsymbol{z}_1 = \boldsymbol{z}_1'$ and $\boldsymbol{z}_2 = \boldsymbol{z}_2'$. Thus the decoder receives the same input arguments in both the cases and consequently causes an error.

Thus we have that $H_{|\mathcal{Z}|}(g(X_3^k)|\boldsymbol{Z}_{31}, \boldsymbol{Z}_{32}) = 0$. That gives us $H_{|\mathcal{Z}|}(g(X_3^k)) \leq H_{|\mathcal{Z}|}(\boldsymbol{Z}_{31}) + H_{|\mathcal{Z}|}(\boldsymbol{Z}_{32})$, and using the upper bound to the entropy in terms of the expected codeword length (*cf.* equation 1 and lemma 1), we have the following.

$$\begin{aligned}
H_{|\mathcal{Z}|}(g(X_3^k)) \leq\, &\mathbb{E}\,\ell(\boldsymbol{Z}_{31}) + 2\log_{|\mathcal{Z}|}(H_{|\mathcal{Z}|}(\boldsymbol{Z}_{31}) + |\mathcal{Z}|) + \\
&\mathbb{E}\,\ell(\boldsymbol{Z}_{32}) + 2\log_{|\mathcal{Z}|}(H_{|\mathcal{Z}|}(\boldsymbol{Z}_{32}) + |\mathcal{Z}|), \\
\implies H(g(X_3^k)) \leq\, &k(R_{31} + R_{32} + 2\epsilon + \delta)\log|\mathcal{A}|,
\end{aligned}$$

the second inequality uses the same ϵ for both rates, and δ can be made small enough for large k as $H_{|\mathcal{Z}|}(\boldsymbol{Z}_{31}) \leq H_{|\mathcal{Z}|}(X_3^k) = k\log_{|\mathcal{Z}|}|\mathcal{A}|$ and similarly for $H_{|\mathcal{Z}|}(\boldsymbol{Z}_{32})$. ■

In the rest of the paper, we outline a systematic procedure to get a lower bound for the sum rate $R_1 + R_2$. We can also use a similar process to get lower bounds for R_1 and R_2 individually; the details can be found in [10]. There are two key parts in the proof of lemma 2: (i) an equivalence relation for the message values, which gives (ii) a lower bound to the entropy of the codewords transmitted. The same two ideas will also be used in bounding the sum rate. The equivalence relation used for X_1^k and X_2^k is originally given in [3], [4] and we adapt them to our particular network instance. This equivalence relation allows us to give a lower bound to the relevant entropy as follows.

$$\begin{aligned}
H_{|\mathcal{Z}|}(\boldsymbol{Z}_1) + H_{|\mathcal{Z}|}(\boldsymbol{Z}_2) &\geq H_{|\mathcal{Z}|}(\boldsymbol{Z}_1, \boldsymbol{Z}_2|B^k, X_3^k) + H_{|\mathcal{Z}|}(B^k) \\
&\geq \alpha k + H_{|\mathcal{Z}|}(B^k),
\end{aligned}$$

where the first inequality is true by the zero-error criterion and α is a quantity that can be evaluated for a general demand function as described in the sequel (*cf.* (6)). The lower bound to the sum rate can then be given in terms of α as

$$\begin{aligned}\alpha k + H_{|\mathcal{Z}|}(B^k) \le& \,\mathbb{E}\,\ell(\boldsymbol{Z}_1) + 2\log_{|\mathcal{Z}|}(H_{|\mathcal{Z}|}(\boldsymbol{Z}_1)+|\mathcal{Z}|)+ \\ & \mathbb{E}\,\ell(\boldsymbol{Z}_2) + 2\log_{|\mathcal{Z}|}(H_{|\mathcal{Z}|}(\boldsymbol{Z}_2)+|\mathcal{Z}|), \\ \le& \, k(R_1+R_2+2\epsilon+\delta)\log_{|\mathcal{Z}|}|\mathcal{A}|. \qquad (2)\end{aligned}$$

A. Equivalence relation and its associated classes

Definition 3: For all possible realizations (x_1,x_2,x_3), $(y_1,y_2,y_3) \in \mathcal{A}^3$ such that $x_3 = y_3 \triangleq a_3$, we say[1] that $x_1 \overset{a_3}{\equiv} y_1|_1$ if and only if $f(x_1,x_2,a_3) = f(y_1,y_2,a_3)$ for all $x_2 = y_2 \in \mathcal{A}$. Similarly, for all possible realizations $(x_1,x_2,x_3),(y_1,y_2,y_3) \in \mathcal{A}^3$ such that $x_3 = y_3 \triangleq a_3$, we say that $x_2 \overset{a_3}{\equiv} y_2|_2$ if and only if $f(x_1,x_2,a_3) = f(y_1,y_2,a_3)$ for all $x_1 = y_1 \in \mathcal{A}$.
Note that for both $u = 1,2$ and any value of $a_3 \triangleq x_3 = y_3$,

- $x_u = y_u$ implies $x_u \overset{a_3}{\equiv} y_u|_u$,
- $x_u \overset{a_3}{\equiv} y_u|_u$ implies $y_u \overset{a_3}{\equiv} x_u|_u$, and
- $x_u \overset{a_3}{\equiv} w_u|_u$ and $w_u \overset{a_3}{\equiv} y_u|_u$ implies $x_u \overset{a_3}{\equiv} y_u|_u$.

Thus $\overset{a_3}{\equiv}|_u$ is an equivalence relation on $\mathcal{A}$ for any demand function $f(X_1,X_2,X_3)$, choice of $u \in \{1,2\}$ and $a_3 \in \mathcal{A}$. The number of equivalence classes of $\mathcal{A}$ induced by $\overset{a_3}{\equiv}|_u$ is denoted as $V_u(a_3)$ for both $u=1$ and 2.

The above equivalence relation gives the minimum number of $\boldsymbol{Z}_u$ codewords that must be transmitted on the edges, because of the following lemma, which can also be obtained by adapting lemma 3 in [4] to the network here.

Lemma 3: Consider a block of k independent realizations of X_1, X_2 and X_3 and let $\boldsymbol{a}_3 \in \mathcal{A}^k$ be the realization for X_3^k. Then for $u \in \{1,2\}$, the number of distinct $\boldsymbol{Z}_u$-labels that must be transmitted on the edge (s_u,t) to allow the terminal to recover B^k with zero error is at least $V_u(\boldsymbol{a}_3) \triangleq \prod_{i=1}^k V_u(a_3^{(i)})$.
Let $\overset{\boldsymbol{a}_3}{\equiv}|_u$ denote the collection of equivalence relations of definition 3 for each component of $\boldsymbol{a}_3$. In this notation,

$$\boldsymbol{x}_u \overset{\boldsymbol{a}_3}{\equiv} \boldsymbol{y}_u \Leftrightarrow x_u^{(j)} \overset{a_3^{(j)}}{\equiv} y_u^{(j)}, \quad \forall j \in \{1,2,\ldots,k\}.$$

If $\boldsymbol{x}_1 \overset{\boldsymbol{a}_3}{\not\equiv} \boldsymbol{y}_1$, then $\phi_{(s_1,t)}(\boldsymbol{x}_1,\boldsymbol{a}_3) \neq \phi_{(s_1,t)}(\boldsymbol{y}_1,\boldsymbol{a}_3)$, i.e., their $\boldsymbol{Z}_1$ labels must be different. An analogous statement is true for the $\boldsymbol{Z}_2$ label as well.

For either $u=1$ or 2 and each $i \in \{1,2,\ldots,k\}$, the equivalence classes under $\overset{a_3^{(i)}}{\equiv}|_u$ are denoted as $\mathrm{Cl}_u^{(j)}(a_3^{(i)})$, where the superscript $j \in \{1,2,\ldots,V_u(a_3^{(i)})\}$ indexes the classes such that

$$|\mathrm{Cl}_u^{(1)}(a_3^{(i)})| \ge |\mathrm{Cl}_u^{(2)}(a_3^{(i)})| \ge \cdots \ge |\mathrm{Cl}_u^{(V_u(a_3^{(i)}))}(a_3^{(i)})|.$$

Based on the value of $\boldsymbol{a}_3$ at each component, the set $\mathcal{A}^k$ can be partitioned into $V_u(\boldsymbol{a}_3) = \prod_{i=1}^k V_u(a_3^{(i)})$ partitions. These partitions of $\mathcal{A}^k$ under $\overset{\boldsymbol{a}_3}{\equiv}|_u$ can be represented using a index vector $\boldsymbol{v}$ having k components, each of which satisfies $v^{(i)} \in \{1,2,\ldots,V_u(a_3^{(i)})\}$ and

$$\boldsymbol{x}_u \in \mathrm{Cl}_u^{(\boldsymbol{v})}(\boldsymbol{a}_3) \Leftrightarrow x_u^{(i)} \in \mathrm{Cl}_u^{(v^{(i)})}(a_3^{(i)}), \quad \forall i \in \{1,2,\ldots,k\}.$$

Like in the scalar case, we add a subscript $t \in \{1,2,\ldots,V_u(\boldsymbol{a}_3)\}$ to get the index vector $\boldsymbol{v}_t$ such that the equivalence classes under $\overset{\boldsymbol{a}_3}{\equiv}|_u$ satisfy the ordering

$$|\mathrm{Cl}_u^{(\boldsymbol{v}_1)}(\boldsymbol{a}_3)| \ge |\mathrm{Cl}_u^{(\boldsymbol{v}_2)}(\boldsymbol{a}_3)| \ge \cdots \ge |\mathrm{Cl}_u^{(\boldsymbol{v}_{V_u(\boldsymbol{a}_3)})}(\boldsymbol{a}_3)|.$$

From the definition, for every $t \in \{1,2,\ldots,V_u(\boldsymbol{a}_3)\}$, we have that $\mathrm{Cl}_u^{(\boldsymbol{v}_t)}(\boldsymbol{a}_3) = \bigtimes_{i=1}^k \mathrm{Cl}_u^{(v_t^{(i)})}(a_3^{(i)})$, i.e., a cartesian product of k scalar equivalence classes and accordingly $|\mathrm{Cl}_u^{(\boldsymbol{v}_t)}(\boldsymbol{a}_3)| = \prod_{i=1}^k |\mathrm{Cl}_u^{(v_t^{(i)})}(a_3^{(i)})|$.

We characterize the family of valid conditional p.m.f.s for the pair $(\boldsymbol{Z}_1,\boldsymbol{Z}_2)$ given the values of the demand function $f(X_1^k,X_2^k,X_3^k)$ and the realization of the message X_3^k. Towards this end we first find the number of distinct $(\boldsymbol{Z}_1,\boldsymbol{Z}_2)$-labels that must be assigned by the network code to message tuples that result in a particular value, say, $\boldsymbol{b} \in \mathcal{B}^k$ of the demand function. The set $A_3(\boldsymbol{b})$ has all possible realizations $\boldsymbol{a}_3$ of X_3^k that can result in the value of $\boldsymbol{b}$ for the demand function, i.e.,

$$A_3(\boldsymbol{b}) \triangleq \{\boldsymbol{a}_3 \in \mathcal{A}^k : \exists\, \boldsymbol{x}_1,\boldsymbol{x}_2 \in \mathcal{A}^k \text{ s.t. } f(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{a}_3) = \boldsymbol{b}\}.$$

Let $M(\boldsymbol{a}_3,\boldsymbol{b})$ be the number of distinct $(\boldsymbol{Z}_1,\boldsymbol{Z}_2)$ pair labels used for message tuples that have $X_3^k = \boldsymbol{a}_3$ and $B^k = \boldsymbol{b}$. Consider two message tuples $(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{a}_3)$ and $(\boldsymbol{y}_1,\boldsymbol{y}_2,\boldsymbol{a}_3)$ which satisfy $f(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{a}_3) = f(\boldsymbol{y}_1,\boldsymbol{y}_2,\boldsymbol{a}_3) = \boldsymbol{b}$. If either $\boldsymbol{x}_1 \overset{\boldsymbol{a}_3}{\not\equiv} \boldsymbol{y}_1|_1$ or $\boldsymbol{x}_2 \overset{\boldsymbol{a}_3}{\not\equiv} \boldsymbol{y}_2|_2$, then the pair of labels $(\boldsymbol{Z}_1,\boldsymbol{Z}_2)$ assigned to the two message tuples must be different. This motivates us to define the pair index set:

$$\mathcal{V}_{12}(\boldsymbol{a}_3,\boldsymbol{b}) \triangleq \left\{ (\mathrm{Cl}_1^{(\boldsymbol{v}_j)}, \mathrm{Cl}_2^{(\boldsymbol{w}_t)}) : \begin{array}{r} \exists\, \boldsymbol{x}_1 \in \mathrm{Cl}_1^{(\boldsymbol{v}_j)}(\boldsymbol{a}_3), \boldsymbol{x}_2 \in \\ \mathrm{Cl}_2^{(\boldsymbol{w}_t)}(\boldsymbol{a}_3) \text{ s.t. } f(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{a}_3) = \boldsymbol{b}, \\ 1 \le j \le V_1(\boldsymbol{a}_3), 1 \le t \le V_2(\boldsymbol{a}_3). \end{array} \right\}.$$

Then for any $\boldsymbol{b} \in \mathcal{B}^k, \boldsymbol{a}_3 \in A_3(\boldsymbol{b})$, $M(\boldsymbol{a}_3,\boldsymbol{b}) \ge |\mathcal{V}_{12}(\boldsymbol{a}_3,\boldsymbol{b})|$.

Example 1: Consider block size $k=1$ and a realization $b=0$ of the demand function of table I. We can evaluate that

$$\mathrm{Cl}_1^{(1)}(1) = \{0,1\}, \mathrm{Cl}_1^{(2)}(1) = \{2\}, \text{ and}$$
$$\mathrm{Cl}_2^{(1)}(1) = \{1,2\}, \mathrm{Cl}_2^{(2)}(1) = \{0\}.$$

TABLE I
FUNCTION TABLE FOR A DEMAND FUNCTION TO BE COMPUTED OVER THE NETWORK IN FIGURE 1. THE MESSAGE ALPHABET IS $\mathcal{A} = GF(3)$. TABLE IA SHOWS THE FUNCTION VALUES FOR ALL (X_1,X_2) PAIRS WHEN $X_3 = 0$, TABLE IB SHOWS THE FUNCTION VALUES WHEN $X_3 = 1$ AND TABLE IC SHOWS THE FUNCTION VALUES WHEN $X_3 = 2$.

$X_3=0$		X_2 = 0	1	2
	0	0	2	1
X_1	1	1	0	2
	2	2	1	0

(a)

$X_3=1$		X_2 = 0	1	2
	0	0	0	0
X_1	1	0	0	0
	2	1	0	0

(b)

$X_3=2$		X_2 = 0	1	2
	0	1	1	0
X_1	1	1	1	1
	2	1	1	1

(c)

[1]Read as 'x_1 is a_3-equivalent to y_1 for the message at s_1'.

The pair index set $\mathcal{V}_{12}(1,0) = \{(\mathrm{Cl}_1^{(1)}, \mathrm{Cl}_2^{(1)}), (\mathrm{Cl}_1^{(1)}, \mathrm{Cl}_2^{(2)}), (\mathrm{Cl}_1^{(2)}, \mathrm{Cl}_2^{(1)})\}$. Note that $(\mathrm{Cl}_1^{(2)}, \mathrm{Cl}_2^{(2)}) \notin \mathcal{V}_{12}(1,0)$ as the elements of that pair of equivalence classes do not result in the demand function value of 0, i.e.,

$$\mathrm{Cl}_1^{(2)}(1) = \{2\},\ \mathrm{Cl}_2^{(2)}(1) = \{0\},$$
$$\text{but for } x_1 = 2, x_2 = 0, x_3 = 1,\ f(x_1, x_2, x_3) = 1 \neq 0.$$

Thus in this case we have that $|\mathcal{V}_{12}(1,0)| = 3$. The other pair index sets for this demand function are given in table II.

B. Lower bound for conditional entropy

We now explicitly derive a p.m.f. whose entropy is a lower bound to $H_{|\mathcal{Z}|}(\boldsymbol{Z}_1, \boldsymbol{Z}_2 | B^k = \boldsymbol{b}, X_3^k = \boldsymbol{a}_3)$. Let $A_{123}(\boldsymbol{b}) \subseteq \mathcal{A}^{3k}$ contain all message tuples that are present in the pre-image of the demand function value of $\boldsymbol{b}$. For a $\boldsymbol{a}_3 \in A_3(\boldsymbol{b})$ we define an associated subset of the pre-image set as follows.

$$A_{123}(\boldsymbol{b}, \boldsymbol{a}_3) \triangleq \{(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{a}_3) : f(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{a}_3) = \boldsymbol{b}\}.$$

Let $(\mathrm{Cl}_1^{(\boldsymbol{v})}, \mathrm{Cl}_2^{(\boldsymbol{w})}) \in \mathcal{V}_{12}(\boldsymbol{a}_3, \boldsymbol{b})$. The number of different message tuples that cause the membership of the equivalence class pair $(\mathrm{Cl}_1^{(\boldsymbol{v})}, \mathrm{Cl}_2^{(\boldsymbol{w})}) \in \mathcal{V}_{12}(\boldsymbol{a}_3, \boldsymbol{b})$ is denoted as follows.

$$h_{\boldsymbol{a}_3}(\boldsymbol{v}, \boldsymbol{w}) \triangleq \left| \left\{ (\boldsymbol{x}_1, \boldsymbol{x}_2) : \begin{array}{l} \boldsymbol{x}_1 \in \mathrm{Cl}_1^{(\boldsymbol{v})}(\boldsymbol{a}_3),\ \boldsymbol{x}_2 \in \mathrm{Cl}_2^{(\boldsymbol{w})}(\boldsymbol{a}_3), \\ \text{and } f(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{a}_3) = \boldsymbol{b}. \end{array} \right\} \right|$$
$$= |\mathrm{Cl}_1^{(\boldsymbol{v})}(\boldsymbol{a}_3)| \cdot |\mathrm{Cl}_2^{(\boldsymbol{w})}(\boldsymbol{a}_3)|. \quad (3)$$

Equality (3) is true above as by definition 3 every element of an equivalence class under $\overset{\boldsymbol{a}_3}{\equiv} |_1$ results in the same demand function value (while the other message $\boldsymbol{x}_2$ is held constant), and since $(\mathrm{Cl}_1^{(\boldsymbol{v})}, \mathrm{Cl}_2^{(\boldsymbol{w})}) \in \mathcal{V}_{12}(\boldsymbol{a}_3, \boldsymbol{b})$, there is at least one $\boldsymbol{x}_1 \in \mathrm{Cl}_1^{(\boldsymbol{v})}(\boldsymbol{a}_3)$ and one $\boldsymbol{x}_2 \in \mathrm{Cl}_2^{(\boldsymbol{w})}(\boldsymbol{a}_3)$ such that $f(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{a}_3) = \boldsymbol{b}$. Hence every other pair of elements in $\mathrm{Cl}_1^{(\boldsymbol{v})}(\boldsymbol{a}_3) \times \mathrm{Cl}_2^{(\boldsymbol{w})}(\boldsymbol{a}_3)$ would also result in the same demand function value with $X_3^k = \boldsymbol{a}_3$.

Example 2: For block size $k = 1$ and demand function realization $b = 0$, from table I, $A_3(0) = \{0, 1, 2\}$. Following the indexing of the equivalence partitions and Table II we have that $h_1(1,1) = |\mathrm{Cl}_1^{(1)}(1)| \cdot |\mathrm{Cl}_2^{(1)}(1)| = 4$ as $\mathrm{Cl}_1^{(1)}(1) = \{0, 1\}$ and $\mathrm{Cl}_2^{(1)}(1) = \{1, 2\}$. One can similarly check that $h_1(1,2) = h_1(2,1) = 2$ and for other values of a_3, that $h_0(1,1) = h_0(2,2) = h_0(3,3) = 1$ and $h_2(2,2) = 1$.

The following notation is useful in stating a necessary condition for any valid p.m.f. for the $(\boldsymbol{Z}_1, \boldsymbol{Z}_2)$ pair label.

Definition 4: For any index i of a vector $\boldsymbol{p}$, we use $p^{[i]}$ to denote the ith component of $\boldsymbol{p}$ when it is arranged in non-increasing order. For two vectors $\boldsymbol{p}, \boldsymbol{q}$ of the same length l, the vector $\boldsymbol{p}$ is majorized by $\boldsymbol{q}$, denoted as $\boldsymbol{p} \prec \boldsymbol{q}$, if

$$\sum_{i=1}^{t} p^{[i]} \leq \sum_{i=1}^{t} q^{[i]} \quad \forall t \leq l-1, \text{ and } \sum_{i=1}^{l} p^{[i]} = \sum_{i=1}^{l} q^{[i]}.$$

As an example, the vector $[0.5\ \ 0.5]$ is majorized by $[0.25\ \ 0.75]$. Note that any vector $\boldsymbol{p}$ is majorized by itself. Interpreting p.m.f.s as vectors over non-negative real numbers (denoted as $\mathbb{R}_{\geq 0}$), we have the following.

Lemma 4: For any $\boldsymbol{a}_3 \in A_3(\boldsymbol{b})$, define a vector $\boldsymbol{h}_{\boldsymbol{b}, \boldsymbol{a}_3}$ as

$$\boldsymbol{h}_{\boldsymbol{b}, \boldsymbol{a}_3} \triangleq$$
$$\begin{bmatrix} h_{\boldsymbol{a}_3}((\boldsymbol{v}, \boldsymbol{w})_1) & h_{\boldsymbol{a}_3}((\boldsymbol{v}, \boldsymbol{w})_2) & \cdots & h_{\boldsymbol{a}_3}((\boldsymbol{v}, \boldsymbol{w})_{|\mathcal{V}_{12}(\boldsymbol{a}_3, \boldsymbol{b})|}) & \boldsymbol{0}_t \end{bmatrix}^\top$$

where $\boldsymbol{0}_t$ indicates a vector of zeros of length $t \triangleq M(\boldsymbol{a}_3, \boldsymbol{b}) - |\mathcal{V}_{12}(\boldsymbol{a}_3, \boldsymbol{b})|$ and subscript j in $(\boldsymbol{v}, \boldsymbol{w})_j$ indexes all the equivalence class pairs in $\mathcal{V}_{12}(\boldsymbol{a}_3, \boldsymbol{b})$ such that

$$h_{\boldsymbol{a}_3}((\boldsymbol{v}, \boldsymbol{w})_1) \geq h_{\boldsymbol{a}_3}((\boldsymbol{v}, \boldsymbol{w})_2) \geq \cdots \geq h_{\boldsymbol{a}_3}((\boldsymbol{v}, \boldsymbol{w})_{|\mathcal{V}_{12}(\boldsymbol{a}_3, \boldsymbol{b})|}).$$

Then all conditional p.m.f.s $\boldsymbol{p} \in \mathbb{R}_{\geq 0}^{M(\boldsymbol{a}_3, \boldsymbol{b})}$ on $M(\boldsymbol{a}_3, \boldsymbol{b})$ valid $(\boldsymbol{z}_1, \boldsymbol{z}_2)$-labels given the value $\boldsymbol{b}$ of the demand function and the realization $\boldsymbol{a}_3$ of X_3^k satisfy $\boldsymbol{p} \prec \boldsymbol{h}_{\boldsymbol{b}, \boldsymbol{a}_3} / |A_{123}(\boldsymbol{b}, \boldsymbol{a}_3)|$.

Proof: We first note that $\boldsymbol{h}_{\boldsymbol{b}, \boldsymbol{a}_3} / |A_{123}(\boldsymbol{b}, \boldsymbol{a}_3)|$ is a valid p.m.f. as its components are non-negative and sum up to 1. Suppose that there is an encoding scheme for $\boldsymbol{Z}_1$ and $\boldsymbol{Z}_2$ such that $\Pr\{\boldsymbol{Z}_1, \boldsymbol{Z}_2 | X_3^k = \boldsymbol{a}_3, B^k = \boldsymbol{b}\} \triangleq \boldsymbol{p} \not\prec \boldsymbol{h}_{\boldsymbol{b}, \boldsymbol{a}_3} / |A_{123}(\boldsymbol{b}, \boldsymbol{a}_3)|$. Furthermore let $\boldsymbol{p}$ be supported on $L_{12}(\boldsymbol{a}_3)$ components. Then the assumption implies that there is a $m < L_{12}(\boldsymbol{a}_3)$ such that

$$\sum_{j=1}^{m} p^{[j]} > \frac{1}{|A_{123}(\boldsymbol{b}, \boldsymbol{a}_3)|} \sum_{j=1}^{m} h_{\boldsymbol{a}_3}((\boldsymbol{v}, \boldsymbol{w})_j). \quad (4)$$

Since each realization of the pair (X_1^k, X_2^k) is equally likely, the RHS in the above equation is the conditional probability given the value of $X_3^k = \boldsymbol{a}_3$ and $B^k = \boldsymbol{b}$ of the event that (X_1^k, X_2^k) belongs to one of the m largest equivalence class pairs under the pair of relations $(\overset{\boldsymbol{a}_3}{\equiv} |_1, \overset{\boldsymbol{a}_3}{\equiv} |_2)$. The LHS is the conditional probability of observing any of the m most probable $(\boldsymbol{Z}_1, \boldsymbol{Z}_2)$ pair labels. Thus equation (4) implies such an encoding scheme gives a total of m distinct $(\boldsymbol{Z}_1, \boldsymbol{Z}_2)$ pair labels to as many (X_1^k, X_2^k) pairs for which the pair belongs to at least $m+1$ different equivalence class pairs in $\mathcal{V}_{12}(\boldsymbol{a}_3, \boldsymbol{b})$. By the pigeonhole principle, this contradicts lemma 3. ■

To find a lower bound to the conditional entropy, we use the order-preserving property of the entropy function with respect to the majorization relation. The entropy function $H : \mathbb{R}^{L_{12}(\boldsymbol{a}_3)} \to \mathbb{R}$ is a strictly Schur-concave function [9, Chap. 3], i.e., for two p.m.f.s $\boldsymbol{p}, \boldsymbol{q} \in \mathbb{R}^{L_{12}(\boldsymbol{a}_3)}$ that are not equal to each other under any permutation of their components,

$$\boldsymbol{p} \prec \boldsymbol{q} \implies H(\boldsymbol{p}) > H(\boldsymbol{q}).$$

Using this and lemma 4, we obtain

$$H_{|\mathcal{Z}|}(\boldsymbol{Z}_1, \boldsymbol{Z}_2 | X_3^k = \boldsymbol{a}_3, B^k = \boldsymbol{b}) \geq H_{|\mathcal{Z}|}\Big(\frac{\boldsymbol{h}_{\boldsymbol{b}, \boldsymbol{a}_3}}{|A_{123}(\boldsymbol{b}, \boldsymbol{a}_3)|} \Big). \quad (5)$$

Having evaluated $H_{|\mathcal{Z}|}(\boldsymbol{h}_{\boldsymbol{b}, \boldsymbol{a}_3} / |A_{123}(\boldsymbol{b}, \boldsymbol{a}_3)|)$, we can find the value of α as follows.

$$H_{|\mathcal{Z}|}(\boldsymbol{Z}_1, \boldsymbol{Z}_2 | B^k, X_3^k)$$
$$= \sum_{\boldsymbol{b}, \boldsymbol{x}_3} \Pr\{X_3^k = \boldsymbol{x}_3, B^k = \boldsymbol{b}\} H_{|\mathcal{Z}|}(\boldsymbol{Z}_1, \boldsymbol{Z}_2 | B^k = \boldsymbol{b}, X_3^k = \boldsymbol{x}_3)$$
$$\geq \sum_{\boldsymbol{b}, \boldsymbol{x}_3} \Pr\{X_3^k = \boldsymbol{x}_3, B^k = \boldsymbol{b}\} H_{|\mathcal{Z}|}\Big(\frac{\boldsymbol{h}_{\boldsymbol{b}, \boldsymbol{a}_3}}{|A_{123}(\boldsymbol{b}, \boldsymbol{a}_3)|} \Big) \triangleq \alpha k, \quad (6)$$

TABLE II

THE SETS $\mathcal{V}_{12}(a_3, b)$ FOR DIFFERENT VALUES OF THE DEMAND FUNCTION REALIZATION b AND DIFFERENT VALUES OF a_3 IN DIFFERENT ROWS.

a_3	$b=0$	$b=1$	$b=2$
0	$\{(\mathrm{Cl}_1^{(1)},\mathrm{Cl}_2^{(1)}),(\mathrm{Cl}_1^{(2)},\mathrm{Cl}_2^{(2)}),(\mathrm{Cl}_1^{(3)},\mathrm{Cl}_2^{(3)})\}$	$\{(\mathrm{Cl}_1^{(1)},\mathrm{Cl}_2^{(3)}),(\mathrm{Cl}_1^{(2)},\mathrm{Cl}_2^{(1)}),(\mathrm{Cl}_1^{(3)},\mathrm{Cl}_2^{(2)})\}$	$\{(\mathrm{Cl}_1^{(1)},\mathrm{Cl}_2^{(2)}),(\mathrm{Cl}_1^{(2)},\mathrm{Cl}_2^{(3)}),(\mathrm{Cl}_1^{(3)},\mathrm{Cl}_2^{(1)})\}$
1	$\{(\mathrm{Cl}_1^{(1)},\mathrm{Cl}_2^{(1)}),(\mathrm{Cl}_1^{(1)},\mathrm{Cl}_2^{(2)}),(\mathrm{Cl}_1^{(2)},\mathrm{Cl}_2^{(1)})\}$	$\{(\mathrm{Cl}_1^{(2)},\mathrm{Cl}_2^{(2)})\}$	$\emptyset$
2	$\{(\mathrm{Cl}_1^{(2)},\mathrm{Cl}_2^{(2)})\}$	$\{(\mathrm{Cl}_1^{(1)},\mathrm{Cl}_1^{(1)}),(\mathrm{Cl}_1^{(1)},\mathrm{Cl}_2^{(2)}),(\mathrm{Cl}_1^{(2)},\mathrm{Cl}_2^{(1)})\}$	$\emptyset$

where the above inequality is true by the lower bound in (5). By the i.i.d. uniform assumption on the message tuples, the value of $\Pr\{X_3^k = \boldsymbol{x}_3, B^k = \boldsymbol{b}\}$ is given by $\frac{|A_{123}(\boldsymbol{b},\boldsymbol{x}_3)|}{|\mathcal{A}|^{3k}}$. We calculate α for the example demand function in [10, Sec. IIIB].

IV. Arithmetic sum demand function

Suppose the message alphabet is $\mathcal{A} = \{0,1\}$, such that the messages X_1, X_2, X_3 are independent bits each equally likely to be 0 or 1. The demand function $B = f(X_1, X_2, X_3) = X_1 + X_2 + X_3$ is the sum of the messages over the integers, such that $\mathcal{B} = \{0,1,2,3\}$. We use the codeword alphabet $\mathcal{Z} = \{0,1\}$. This case of arithmetic sum computation in the variable-length network code framework was considered in [6] and we recover the results there in our general framework.

For any value of X_1^k, X_3^k and B^k the value of X_2^k is fixed by $X_2^k = B^k - X_1^k - X_3^k$, where the subtraction over the integers operates componentwise. Hence, for every $(\mathrm{Cl}_1^{(\boldsymbol{v})}, \mathrm{Cl}_2^{(\boldsymbol{w})}) \in \mathcal{V}_{12}(\boldsymbol{a}_3, \boldsymbol{b})$ there is exactly one message tuple whose X_1^k and X_2^k belong to the equivalence classes $\mathrm{Cl}_1^{(\boldsymbol{v})}(\boldsymbol{a}_3)$ and $\mathrm{Cl}_2^{(\boldsymbol{w})}(\boldsymbol{a}_3)$ respectively. Thus we have that

$$h_{\boldsymbol{a}_3}(\boldsymbol{v},\boldsymbol{w}) = 1 \ \forall \boldsymbol{a}_3 \in A_3(\boldsymbol{b}), (\mathrm{Cl}_1^{(\boldsymbol{v})}, \mathrm{Cl}_2^{(\boldsymbol{w})}) \in \mathcal{V}_{12}(\boldsymbol{a}_3, \boldsymbol{b}). \quad (7)$$

Consider an arithmetic sum realization $\boldsymbol{b}$ with t_0 zeros, t_1 ones, t_2 twos and $k - t_0 - t_1 - t_2$ threes. For every $p \in \{0,1,2,3\}$ and $q \in \{0,1\}$, define $t_{p,q} \triangleq |\{i : b^{(i)} = p, X_3^{(i)} = q, 1 \le i \le k\}|$. Then for any choice of $\boldsymbol{b}$ and $\boldsymbol{a}_3 \in A_3(\boldsymbol{b})$, $t_{0,1} = t_{3,0} = 0$. The cardinality of the pre-image set $|A_{123}(\boldsymbol{b}, \boldsymbol{a}_3)| = 2^{t_{1,0}+t_{2,1}}$. Using (7) the entropy $H(\boldsymbol{h}_{\boldsymbol{b},\boldsymbol{a}_3}/|A_{123}(\boldsymbol{b},\boldsymbol{a}_3)|) = t_{1,0} + t_{2,1}$. From the function definition, we can check that $|A_{123}(\boldsymbol{b})| = 3^{t_1+t_2}$. Hence,

$$\Pr\{X_3^k = \boldsymbol{a}_3 | B^k = \boldsymbol{b}\} = \frac{|A_{123}(\boldsymbol{b},\boldsymbol{a}_3)|}{|A_{123}(\boldsymbol{b})|} = \frac{2^{t_{1,0}+t_{2,1}}}{3^{t_1+t_2}}.$$

For a given $\boldsymbol{b}$, the number of different $\boldsymbol{a}_3 \in A_3(\boldsymbol{b})$ that have the same value for $t_{1,0}$ and $t_{2,1}$ are $\binom{t_1}{t_{1,0}} \cdot \binom{t_2}{t_{2,1}}$. Using these in (6), the value of α can be found as follows.

$$\alpha k = \sum_{\boldsymbol{b}} \Pr\{B^k = \boldsymbol{b}\} \sum_{\boldsymbol{a}_3 \in A_3(\boldsymbol{b})} \Pr\{X_3^k = \boldsymbol{a}_3 | B^k = \boldsymbol{b}\} H\Big(\frac{\boldsymbol{h}_{\boldsymbol{b},\boldsymbol{a}_3}}{|A_{123}(\boldsymbol{b},\boldsymbol{a}_3)|}\Big)$$

$$= \sum_{\boldsymbol{b}} \Pr\{B^k = \boldsymbol{b}\} \sum_{t_{1,0}=0}^{t_1} \sum_{t_{2,1}=0}^{t_2} \binom{t_1}{t_{1,0}} \left(\frac{2}{3}\right)^{t_{1,0}} \left(\frac{1}{3}\right)^{t_1 - t_{1,0}} \cdot \binom{t_2}{t_{2,1}} \left(\frac{2}{3}\right)^{t_{2,1}} \left(\frac{1}{3}\right)^{t_2 - t_{2,1}} (t_{1,0} + t_{2,1})$$

$$= \sum_{t_1=0}^{k} \sum_{t_2=0}^{k-t_1} \frac{k! 2^{k-t_1-t_2} 3^{t_1+t_2}}{t_1! t_2! (k - t_1 - t_2)! 8^k} \frac{2(t_1 + t_2)}{3}$$

$$= \frac{2/3}{4^k} \sum_{t_1=0}^{k} \sum_{t_2=0}^{k-t_1} \frac{k!(t_1+t_2)}{t_1! t_2! (k - t_1 - t_2)!} \left(\frac{3}{2}\right)^{t_1+t_2} = \frac{k}{2}.$$

Using $\alpha = 0.5$ and $H(B^k)/k = 3 - 0.75 \log 3$ in (2), we get

$$R_1 + R_2 + \epsilon \ge 0.5 + 3 - 0.75 \log 3 \approx 2.31128.$$

Remark 1: We note that the lower bound for the sum rate shown above is tighter than the bound $R_1 + R_2 > 2.25$ obtained in [6] for the same problem.

Remark 2: For computing arithmetic sum, a source-network code having sum rate $R_1 + R_2 = 2.5$ was given in [6, Sec. IV]. Thus there is currently a gap between the lower bound and the achieved sum rate for this function computation problem.

References

[1] H. Kowshik and P. R. Kumar, "Optimal function computation in directed and undirected graphs," in *IEEE Trans. on Info. Th.*, vol. 58, no. 6, pp. 3407–3418, June 2012.

[2] R. Appuswamy, M. Franceschetti, N. Karamchandani, and K. Zeger, "Network coding for computing: Cut-set bounds," *IEEE Trans. on Info. Th.*, vol. 57, no. 2, pp. 1015–1030, Feb. 2011.

[3] C. Huang, Z. Tan, and S. Yang, "Upper bound on function computation in directed acyclic networks," in *IEEE Info. Th. Workshop*, April 2015, pp. 1–5.

[4] X. Guang, S. Yang, and C. Li, "An improved upper bound on network function computation using cut-set partition," in *IEEE Info. Th. Workshop*, Sept. 2016, pp. 11–15.

[5] X. Guang, R. W. Yeung, S. Yang and C. Li, "Improved upper bound on the network function computing capacity," online at http://arxiv.org/abs/1710.02252.

[6] A. Tripathy and A. Ramamoorthy, "On computation rates for arithmetic sum," in *IEEE Intl. Symposium on Info. Th.*, July 2016, pp. 2354–2358.

[7] L. Song, R. W. Yeung and N. Cai, "Zero-error network coding for acyclic networks," in *IEEE Trans. on Info. Th.*, vol. 49, no. 12, pp. 3129–3139, Dec. 2003.

[8] S. Leung-Yan-Cheong and T. Cover, "Some equivalences between Shannon entropy and Kolmogorov complexity," in *IEEE Trans. on Info. Th.*, vol. 24, no. 3, pp. 331–338, May 1978.

[9] A. W. Marshall, I. Olkin, and B. C. Arnold, *Inequalities: Theory of Majorization and Its Applications*, ser. Springer Series in Statistics. Springer New York, 2011.

[10] A. Tripathy and A. Ramamoorthy, "Zero-error function computation on a directed acyclic network," online at https://arxiv.org/abs/1805.03730.

[11] A. Ramamoorthy and M. Langberg, "Communicating the sum of sources over a network," *IEEE J. Select. Areas Comm.*, vol. 31, no. 4, pp. 655–665, April 2013.

[12] B. K. Rai and B. K. Dey, "On network coding for sum-networks," *IEEE Trans. on Info. Th.*, vol. 58, no. 1, pp. 50–63, Jan. 2012.

[13] A. Tripathy and A. Ramamoorthy, "Sum-networks from incidence structures: construction and capacity analysis," *IEEE Trans. on Info. Th.*, vol. 64, no. 5, pp. 3461–3480, May 2018.

[14] R. Dougherty, C. Freiling, and K. Zeger, "Insufficiency of linear coding in network information flow," *IEEE Trans. on Info. Th.*, vol. 51, no. 8, pp. 2745–2759, Aug. 2005.

[15] S. Huang and A. Ramamoorthy, "An achievable region for the double unicast problem based on a minimum cut analysis," *IEEE Trans. on Comm.*, vol. 61, no. 7, pp. 2890–2899, Jul. 2013.

[16] ——, "On the multiple unicast capacity of 3-source, 3-terminal directed acyclic networks," *IEEE/ACM Trans. on Networking*, vol. 22, no. 1, pp. 285–299, Feb. 2014.