Upper Bound on Function Computation in Directed Acyclic Networks

Cupjin Huang*, Zihan Tan* and Shenghao Yang†

*The Institute for Theoretical Computer Science (ITCS), Institute for Interdisciplinary Information Sciences, Tsinghua University, Beijing, China
†Institute of Network Coding, The Chinese University of Hong Kong, Hong Kong, China

Abstract—Function computation in directed acyclic networks is considered, where a sink node wants to compute a target function with the inputs generated at multiple source nodes. The network links are error-free but capacity-limited, and the intermediate network nodes perform network coding. The target function is required to be computed with zero error. The computing rate of a network code is measured by the average number of times that the target function can be computed for one use of the network. We propose a cut-set bound on the computing rate using an equivalence relation associated with the inputs of the target function. Our bound holds for general target functions and network topologies. We also show that our bound is tight for some special cases where the computing capacity can be characterized.

I. INTRODUCTION

We consider function computation in a directed acyclic network, where a target function \( f \) is intended to be calculated at a sink node, and the input symbols of the target function are generated at multiple source nodes. As a special case, network communication is just the computation of the identity function. Network function computation naturally arises in sensor networks and Internet of Things, and may find applications in big data processing.

Various models and special cases of this problem have been studied in literature (see the summarizations in [2]–[4]). We are interested in the following network coding model for function computation. Specifically, we assume that the network links have limited (unit) capacity and are error-free. Each source node generates multiple input symbols, and the network codes perform vector network coding by using the network multiple times\(^1\). An intermediate network node can transmit the output of a certain fixed function of the symbols it receives. Here all the intermediated nodes are considered with unbounded computing ability. The target function is required to be computed correctly for all possible inputs. We are interested in the computing rate of a network code that computes the target function, i.e., the average number of times that the target function can be computed for one use of the network. The maximum achievable computing rate is called the computing capacity.

When computing the identity function, the problem becomes the extensively studied network coding. For linear target functions over a finite field, a complete characterization of the computing capacity is not available for networks with one sink node. Certain necessary and sufficient conditions have been obtained such that linear network codes are sufficient to calculate a linear target function. But in general, linear network codes are not sufficient to achieve the computing capacity of linear target functions.

Networks with a single sink node are discussed in this paper, while both the target function and the network code can be non-linear. In this scenario, the computing capacity is known when the network is a multi-edge tree or when the target function is the identity function. For the general case, various bounds on the computing capacity based on cut sets have been studied. But we find that the upper bounds claimed in [2], [3] are not valid. Specifically, the proof of [2, Theorem 1] has an error\(^2\). The condition provided in the beginning of the second paragraph is not always necessary, which is illustrated by an example given in this paper. We show that the computing capacity of our example is strictly larger than the two upper bounds claimed in [2], [3].

Towards a general upper bound, we define an equivalence relation associated with the inputs of the target function (but does not depend on the network topology) and propose a cut-set bound on the computing capacity using this equivalence relation. Our bound holds for general target functions and general network topologies in the network coding model. We also show that our bound is tight when the network is a multi-edge tree or when the target function is the identity function.

In the remainder of this paper, Section II formally introduces the network computing model. The upper bound of the computing rate is given in Theorem 3 and is proved in Section IV. Section III compares with the previous results and discusses the tightness of our upper bound.

---

1. A function \( f : A \to A \) is identity if \( f(x) = x \) for all \( x \in A \).
2. One use of a network means the use of each link in the network at most once.

---

This work was supported in part by the National Basic Research Program of China Grant 2011CBA00300, 2011CBA00301, the National Natural Science Foundation of China Grant 61033001, 61361136003, 61471215. The work described in this paper was partially supported by a grant from University Grants Committee of the Hong Kong Special Administrative Region, China (Project No. AoE/E-02/08).

\( ^2 \) No proof is provided for the upper bound (Lemma 3) in [3].
II. Main Results

In this section, we will first introduce the network computing model. Then we will define cut sets and discuss some special cases of the function computation problem. Last we head to the main theorem about the cut-set bound for function computation.

A. Function-Computing Network Codes

Let $G = (V, E)$ be a directed acyclic graph (DAG) with a finite vertex set $V$ and an edge set $E$, where multi-edges between a certain pair of nodes are allowed. A network over $G$ is denoted as $\mathcal{N} = (G, S, \rho)$, where $S \subseteq V$ is called the source nodes and $\rho \in V \setminus S$ is called the sink node $\rho$. Let $s = |S|$, and without loss of generality (WLOG), let $S = \{1, 2, \ldots, s\}$. For an edge $e = (u, v)$, we call $u$ the tail of $e$ (denoted by tail($e$)) and $v$ the head of $e$ (denoted by head($e$)). Moreover, for each node $u \in V$, let $\mathcal{E}_i(u) = \{e \in E : \text{head}(e) = u\}$ and $\mathcal{E}_o(u) = \{e \in E : \text{tail}(e) = u\}$ be the set of incoming edges and the set of outgoing edges of $u$, respectively. Fix an order of the vertex set $V$ that is consistent with the partial order induced by the directed graph $G$. This order naturally induces an order of the edge set $E$, where edges $e > e'$ if either tail($e$) $> \text{tail}(e')$ or tail($e$) $= \text{tail}(e')$ and head($e$) $> \text{head}(e')$. WLOG, we assume that $\mathcal{E}_i(j) = \emptyset$ for all source nodes $j \in S$, and $\mathcal{E}_o(\rho) = \emptyset$. We will illustrate in Section III-C how to apply our results on a network with $\mathcal{E}_i(j) \neq \emptyset$ for certain $j \in S$.

The network defined above is used to compute a function, where multiple inputs are generated at the source nodes and the output of the function is demanded by the sink node. The computation units with unbounded computing ability are allocated at all the network nodes. However, the computing capability of the network will be bounded by the network transmission capability. Denote by $B$ a finite alphabet. We assume that each edge can transmit a symbol in $B$ reliably for each use.

Denote by $A$ and $O$ two finite alphabets. Let $f : A^s \to O$ be the target function, which is the function to be computed via the network and whose $i$th input is generated at the $i$th source node. We may use the network to compute the function multiple times. Suppose that the $j$th source node consecutively generates $k$ symbols in $A$ denoted by $x_{1j}, x_{2j}, \ldots, x_{kj}$, and the symbols generated by all the source nodes can be given as a matrix $x = (x_{ij})_{k \times s}$. We denote by $x_j$ the $j$th column of $x$, and denote by $x^i$ the $i$th row of $x$. In other words, $x_j$ is the vector of the symbols generated at the $j$th source node, and $x^i$ is the input vector of the $i$th computation of the function $f$. Define for $x \in A^{k \times s}$

$$f^{(k)}(x) = (f(x^1), f(x^2), \ldots, f(x^k))^T.$$ 

For convenience, we denote by $x_j$ the submatrix of $x$ formed by the columns indexed by $J \subseteq S$, and denote by $x^i$ the submatrix of $x$ formed by the rows indexed by $I \subseteq \{1, 2, \ldots, k\}$. We equate $A^{1 \times s}$ with $A^s$ in this paper.

For two positive integers $n$ and $k$, a $(n, k)$ (function-computing) network code over network $\mathcal{N}$ with target function $f$ is defined as follows. Let $x \in A^{k \times s}$ be the matrix formed by symbols generated at the source nodes. The purpose of the code is to compute $f^{(k)}(x)$ by transmitting at most $n$ symbols in $A$ on each edge in $E$. Denote the symbols transmitted on edge $e$ by $g^e(x) \in B^n$. For a set of edges $E \subseteq E$ we define

$$g^E(x) = (g^e(x) | e \in E)$$

where $g^{e_1}(x)$ comes before $g^{e_2}(x)$ whenever $e_1 < e_2$. The $(n, k)$ network code contains the encoding function for each edge $e$, define:

$$h^e : \begin{cases} A^{k} \to B^n, & \text{if } u \in S; \\
\prod_{e' \in E_{\text{tail}}(e)} B^n \to B^n, & \text{otherwise.}
\end{cases}$$

Functions $h^e, e \in E$ determine the symbols transmitted on the edges. Specifically, if $e$ is an outgoing edge of the $i$th source node, then

$$g^e(x) = h^e \left( g^{E_u}(x) \right).$$

The $(n, k)$ network code also contains a decoding function

$$\phi : \prod_{e' \in E_{\text{tail}}(\rho)} B^n \to O^k.$$ 

Define

$$\psi(x) = \phi \left( g^{E_u}(x) \right).$$

If the network code computes $f$, i.e., $\psi(x) = f^{(k)}(x)$ for all $x \in A^{k \times s}$, we then call $\frac{1}{n} \log_2 |B|^s |A|$ an achievable computing rate, where we multiply $\frac{1}{n}$ by $\log_2 |B|^s |A|$ in order to normalize the computing rate for target functions with different input alphabets. The computing capacity of network $\mathcal{N}$ with respect to a target function $f$ is defined as

$$C(\mathcal{N}, f) = \sup \left\{ \frac{k}{n} \log_2 |B|^s |A| : \frac{k}{n} \log_2 |B|^s |A| \text{ is achievable} \right\}.$$

B. Cut Sets and Special Cases

For two nodes $u$ and $v$ in $V$, denote the relation $u \to v$ if there exists a directed path from $u$ to $v$ in $G$. If there is no directed path from $u$ to $v$, we say $u$ is separated from $v$. Given a set of edges $C \subseteq E$, $I_C$ is defined to be the set of source nodes which are separated from the sink node $\rho$ if $C$ is deleted from $E$. Set $C$ is called a cut set if $I_C \neq \emptyset$, and the family of all cut sets in network $\mathcal{N}$ is denoted as $\Lambda(\mathcal{N})$. Additionally, we define the set $K_C$ as

$$K_C = \{ i \in S | \exists (u, v) \in V, i \to v, (v, t) \in C \}.$$ 

It is reasonable to assume that $u \to \rho$ for all $u \in V$. Then one can easily see that $K_C$ is the set of source nodes from which there exists a path to the sink node through $C$. Define $J_C = K_C \setminus I_C$.

The problem also becomes simple when $s = 1$. 
Proposition 1. For a network $\mathcal{N}$ with a single source node and any target function $f : \mathcal{A} \rightarrow \mathcal{O}$, 
\[ C(\mathcal{N}, f) = \min_{C \in \Lambda(\mathcal{N})} \frac{|C|}{\log |\mathcal{A}| |f(\mathcal{A})|} \]
where $f[\mathcal{A}]$ is the image of $f$ on $\mathcal{O}$.

Proof: We first show that the right hand side in the above equation could be achieved. Let 
\[ M = \min_{C \in \Lambda(\mathcal{N})} |C|. \]
Fix integers $k$ and $n$ such that 
\[ k/n \log |\mathcal{B}| |A| \leq C(\mathcal{N}, f), \]
which implies 
\[ |f[\mathcal{A}]|^k \leq |\mathcal{B}|^{nM}. \tag{1} \]
By the max-flow min-cut theorem of directed acyclic graphs, there must exist $M$ edge-disjoint paths from the source node to the sink node, and each of them can transmit one symbol in $\mathcal{B}$. We can apply the following network computing code to compute $f^{(k)}$. The source node first computes $f^{(k)}(x)$, and then encodes $f^{(k)}(x)$ into a unique sequence $y$ in $\mathcal{B}^{nM}$ using a one-to-one mapping $m$, whose existence is guaranteed by (1). The network then forwards $y$ from the source node to the sink node by $n$ uses of the $M$ paths. The sink node decodes by $m^{-1}(y)$.

We then prove the converse. Suppose that we have a $(n, k)$ code with 
\[ k/n \log |\mathcal{B}| |\mathcal{A}| > C(\mathcal{N}, f), \]
which implies 
\[ |f[\mathcal{A}]|^k > |\mathcal{B}|^{nM}. \tag{2} \]
Fix a cut set $C$ with $|C| = M$. It can be shown that 
\[ \psi(x) = \psi^C (g^C(x)), \]
for certain function $\psi^C$ (see Lemma 5 in Section V). By (2) and the pigeonhole principle, there must exist $x, x' \in \mathcal{A}^{k \times s}$ such that 
\[ f^{(k)}(x) \neq f^{(k)}(x'), \]
\[ g^C(x) = g^C(x'), \]
where the second equality implies $\psi(x) = \psi(x')$. Thus this network code cannot compute both $f^{(k)}(x)$ and $f^{(k)}(x')$ correctly. The proof is completed.

C. Upper Bounds

In this paper, we are interested in the general upper bound on $C(\mathcal{N}, f)$. The first upper bound is induced by Proposition 1.

Proposition 2. For a network $\mathcal{N}$ with target function $f$, 
\[ C(\mathcal{N}, f) \leq \min_{C \in \Lambda(\mathcal{N}) : I_C = S} \frac{|C|}{\log |\mathcal{A}| |f(\mathcal{A}^*)|}. \]

Proof: Build a network $\mathcal{N}'$ by joining all the source nodes of $\mathcal{N}$ into a single “super” source node. Since a code for network $\mathcal{N}$ can be naturally converted to a code for network $\mathcal{N}'$ (where the super source node performs the operations of all the source nodes in $\mathcal{N}$), we have 
\[ C(\mathcal{N}, f) \leq C(\mathcal{N}', f). \]

The proof is completed by applying Proposition 1 on $\mathcal{N}'$.

The above upper bound only uses the image of function $f$. We propose an enhanced upper bound by investigating an equivalence relation on the input vectors of $f$. We will compare this equivalence relation with similar definitions proposed in [2], [3] in the next section.

Definition 1 (Equivalence Class). For any function $f : \mathcal{A}^s \rightarrow \mathcal{O}$, any two disjoint index sets $I, J \subseteq \mathcal{S}$, and any $a, b \in \mathcal{A}^{|I|}, c \in \mathcal{A}^{|J|}$, we say $a \equiv b|_{I,J}$ if for every $x, y \in \mathcal{A}^{1 \times s}$, we have $f(x) = f(y)$ whenever $x_I = a, y_I = b, x_J = y_J = c$ and $x_S(I \cup J) = y_S(I \cup J)$. Two vectors $a$ and $b$ satisfying $a \equiv b|_{I,J}$ are said to be $(I, J, c)$-equivalent. When $J = \emptyset$ in the above definition, we use the convention that $c$ is an empty matrix.

Note that the equivalence as defined above does not depend on the structure of the network. However, it will soon be clear that with a network, the division of equivalence classes naturally leads to an upper bound of the network function-computing capacity based on cut sets.

For every $f, I, J$ and $c \in \mathcal{A}^{|I|}$, let $W_{I,J,f}^{(c)}$ denote the total number of equivalence classes induced by $\equiv|_{I,J}$. Given a network $\mathcal{N}$ and a cut set $C$, let $W_{C,f} = \max_{c \in \mathcal{A}^{|C|}} W_{I,J,C,f}^{(c)}$. Our main result is stated as following. The proof of the theorem is presented in Section V.

Theorem 3. If $\mathcal{N}$ is a network and $f$ is a target function, then 
\[ C(\mathcal{N}, f) \leq \min_{C \in \Lambda(\mathcal{N}) : I_C = S} \frac{|C|}{\log |\mathcal{A}| W_{C,f}} := \text{min-cut}(\mathcal{N}, f). \]

III. DISCUSSION OF UPPER BOUND

In this section, we first give an example to illustrate the upper bound. We compare our result with the existing ones, and proceed by a discussion about the tightness of the bound.

A. An Illustration of the Bound

First we give an example to illustrate our result. Consider the network $\mathcal{N}_1$ in Fig. 1 with the object function 
\[ f(x_1, x_2, x_3) = x_1 x_2 + x_3, \]
where $\mathcal{A} = E = \mathcal{O} = \{0, 1\}$.

Let us first compare the upper bounds in Theorem 3 and Proposition 2. Let $\mathcal{C}_0 = \{e_0, e_1\}$. Here we have

- $|\mathcal{C}_0| = 2, I_{\mathcal{C}_0} = \{3\}, J_{\mathcal{C}_0} = \{1, 2\}$;
- For any given inputs of nodes 1 and 2, different inputs from node 3 generate different outputs of $f$. Therefore $W_{I_{\mathcal{C}_0}, J_{\mathcal{C}_0}, f} = 2$ for any $c \in \mathcal{A}^2$ and hence $W_{\mathcal{C}_0, f} = 2$.

By Theorem 3, we have
\[ C(\mathcal{N}_1, f) \leq \text{min-cut}(\mathcal{N}_1, f) \leq \frac{|\mathcal{C}_0|}{\log |\mathcal{A}| W_{\mathcal{C}_0, f}} = 2. \]
While Proposition 2 induces that
\[
\mathcal{C}(N_1, f) \leq \min_{c \in \Lambda(N_1)} |C| = \min_{c \in \Lambda(N_1)} |C| = 4,
\]
where the first equality follows from \( f[A^s] = \{0, 1\} \), and the second equality follows from
\[
\min_{c \in \Lambda(N_1)} |C| = \{ \{e_4, e_5, e_6, e_7\} | = 4.
\]
Therefore, Theorem 3 gives a strictly better upper bound than Proposition 2.

The upper bound in Theorem 3 is actually tight in this case. We claim that there exists a \((1, 2)\) network code that computes \( f \) in \( N_1 \). Consider an input matrix \( x = (x_{ij})_{2 \times 3} \). Node \( i \) sends \( x_{1i} \) to node \( v \) and sends \( x_{2i} \) to node \( \rho \) for \( i = 1, 2, 3 \), respectively, i.e., for \( i = 1, 2, 3 \)
\[
g_{2i} = x_{1i}, \quad g_{4i+3} = x_{2i}.
\]
Node \( v \) then computes \( f(x^1) = x_{11}x_{12} + x_{13} \) and sends it to node \( \rho \) via edge \( e_7 \). Node \( \rho \) receives \( f(x^1) \) from \( e_7 \) and computes \( f(x^2) = x_{21}x_{22} + x_{23} \) using the symbols received from edges \( e_4, e_5 \) and \( e_6 \).

B. Comparison with Previous Works

Upper bounds on the computing capacity have been studied in [2], [3] based on a special case of the equivalence class defined in Definition 1. However, we will demonstrate that the bounds therein do not hold for the example we studied in the last subsection.

In Definition 1 when \( J = \emptyset \), we will say \( a \equiv b \) if \( a \) and \( b \) are \( I \)-equivalent. That is \( a \equiv b \) if for every \( x, y \in A^{1 \times 5} \) with \( x_1 = a, y_1 = b \) and \( x_{S \setminus I} = y_{S \setminus I} \), we have \( f(x) = f(y) \). For target function \( f \) and \( I \subseteq S \), denote by \( R_{I,f} \) the total number of equivalence classes induced by \( f \). For a cut \( C \in \Lambda(N) \), let \( R_{C,f} = R_{I_C,f} \). For \( J \subseteq S \), let \( i_J : J \to \{1, \ldots, |J|\} \) be the one-to-one mapping preserving the order on \( J \), i.e., \( i_J(i) < i_J(j) \) if and only if \( i < j \). Then we have the following lemma:

**Lemma 1.** Let \( I, J \) be disjoint subsets of \( S \) and \( J' \subseteq J \). Then for \( a, b \in A^{|I|} \) and \( c \in A^{|J|} \), we have that \( a \equiv b \) implies \( a \equiv b \) where \( c' = e_{i_J(J')} \). In particular, \( a \equiv b \) implies \( a \equiv b \) for all \( J \subseteq S \setminus I \) and \( c \in A^{|J|} \).

**Proof:** Assume that \( a \equiv b \) where \( I \subseteq S, J \subseteq S \setminus I, J' \subseteq J \), \( a, b \in A^{|I|} \), \( c \in A^{|J|} \) and \( c' = e_{i_J(J')} \). We want to prove that \( a \equiv b \).

It suffices to show that \( f(x) = f(y) \) for all \( x, y \in A^s \) satisfying \( x_I = a, y_I = b, x_{S \setminus I} = y_{S \setminus I} \). We know that \( x_{J'} = (x_{i_J(J')})_{J'} = c_{i_J(J')} \). Define \( f(x) = f(y) \). The proof is finished.

**Lemma 2.** Let \( I \subseteq I' \subseteq S \) and \( J = J' \setminus I \). For all \( c \in A^{|J|} \), \( a \equiv b \) implies \( a' \equiv b' \), where \( a'_{i_J(I)} = a, b'_{i_J(I)} = b \) and \( a'_{i_J(J)} = b'_{i_J(J)} = c \).

**Proof:** We have \( a \equiv b \) implies \( a \equiv b \) by Lemma 1 and \( a' \equiv b' \) is equivalent to \( a \equiv b \) by definition.

**Lemma 3.** Fix network \( N \) and function \( f \). Then, i) for any \( C \in \Lambda(N) \), we have \( R_{C,f} \geq W_{C,f} \); ii) for any \( C, C' \in \Lambda(N) \) with \( C' \subseteq C \) and \( I_{C'} = I_{C} \), we have \( W_{C',f} \geq W_{C,f} \).

**Proof:** Let \( I = I_C, J = J_C \) and \( J' = J_C \). By Lemma 3 for \( c \in A^{|J_C|} \), we have
\[
W_{I,J',f}^c \geq W_{I',J,f}^c
\]
where \( c' = e_{i_J(J')} \). In particular,
\[
R_{I,f} \geq W_{I,J',f}^c.
\]
Then \( R_{I,f} \geq \max_{c \in A^{|J_C|}} W_{I,J',f}^c = W_{C,f} \).

Fix \( c^* \in A^{|J_C|} \) such that \( W_{I,J',f}^c = W_{C,f} \). Then, \( W_{C',f} \geq W_{I,J',f}^c \) for \( c^* \) where \( c^* = e_{i_J(J')} \). Define
\[
\min\text{-cut}_A(N, f) = \min_{C \in \Lambda(N)} \frac{|C|}{\log_2 |A||C|}.
\]
By Lemma 3 we have \( \min\text{-cut}_A(N, f) \geq \min\text{-cut}_A(N, f) \). It is claimed in [2] that \( \min\text{-cut}_A(N, f) \) is an upper bound on \( C(N, f) \). We find, however, \( \min\text{-cut}_A(N, f) \) is not universally an upper bound for the computing capacity. Consider the example in Figure 1 For cut set \( C = \{e_4, e_5, e_7\} \), we have \( I_{C_1} = \{1, 3\} \). On the other hand, it can be proved that \( R_{C_1,f} = 4 \) since \( f \) is an affine function of \( x_2 \) given that \( x_1 \) and \( x_3 \) are fixed, and ii) it takes 2 bits to represent this affine function over the binary field. Hence
\[
\min\text{-cut}_A(N, f) = \min_{C \in \Lambda(N)} \frac{|C|}{\log_2 |A||C|} = \frac{3}{2} < 2 = C(N_1, f).
\]

For a network \( N \) as defined in Section II-A we say a subset of nodes \( U \subset V \) is a cut if \( |U \cap S| > 0 \) and \( \rho \notin U \). For a cut \( U \), denote by \( \mathcal{E}(U) \) the cut set determined by \( U \), i.e.,
\[
\mathcal{E}(U) = \{e \in \mathcal{E} : \text{tail}(e) \in V, \text{head}(e) \in V \setminus U\}.
\]
Let
\[ \Lambda^*(N) = \{ E(U) : U \text{ is a cut in } N \} .\]

Define
\[ \text{min-cut}_k(N, f) = \min \frac{|C|}{\log_{|A|} R_{C,f}}. \]

Since \( \Lambda^*(N) \subset \Lambda(N) \), \( \text{min-cut}_k(N, f) \geq \text{min-cut}_A(N, f) \).

It is implied by [3] Lemma 3 that \( \text{min-cut}_k(N, f) \) is an upper bound on \( C(N, f) \). However, \( \text{min-cut}_k(N, f) \) is also not universally an upper bound for the computing capacity. Consider the example in Fig. 1.

For the cut \( U_1 = \{1, 3, v\} \), the corresponding cut set \( E(U_1) = C_1 = \{ e_4, e_6, e_7 \} \). Hence,
\[ \text{min-cut}_k(N_1, f) \leq \frac{|C_1|}{\log_{|A|} R_{C_1,f}} = \frac{3}{2} < 2 = \mathcal{C}(N_1, f). \]

Though in general the upper bounds in [2], [3] are not valid, for various special cases discussed in [2], e.g., multi-edge tree networks, these bounds still hold.

C. Tightness

The upper bound in Theorem 5 is tight when the network is a multi-edge tree. We may alternatively prove the same result using [2] Theorem III.3, together with the facts that
\[ \text{min-cut}(N, f) = \min \frac{|C|}{\log_{|A|} W_{C,f}} \]

and that for a multi-edge tree network, \( W_{C,f} = R_{C,f} \) for \( C \in \Lambda^*(N) \).

Theorem 4. If \( G \) is a multi-edge tree, for network \( N = (G, S, \rho) \) and any target function \( f \),
\[ \mathcal{C}(N, f) = \text{min-cut}(N, f). \]

Proof: Fix a pair of \( (n, k) \) such that \( \frac{k}{n} \log_{|B|} |A| \leq \text{min-cut}(N, f) \). It suffices to show that there exists an \( (n, k) \) code computing \( f \) on \( N \). We have
\[ W_{C_i,f}^k \leq |B|^{|C_i|} \]

for all \( C_i \in \Lambda(N) \). For node \( u \), define
\[ P(u) = \{ v \in S \mid v \rightarrow u \} , \]
\[ \text{prec}(u) = \{ v \in V \mid (v, u) \in E \} . \]

For all \( u \in G \), \( E_o(u) \) is a cut set, and
\[ I_{E_o(u)} = P(u) , \quad J_{E_o(u)} = \emptyset , \]
\[ R_{E_o(u)}^i \leq |B|^{|E_o(u)|} \]

We assign each \( u \) a function
\[ \gamma_u : A^{P(u)} \rightarrow \{1, 2, \ldots, R_{P(u)}, f\} \]
such that
\[ \gamma_u(x) = \gamma_u(y) \iff x = y|P(u). \]

For any \( s \in A^{P(u)} \), the value \( \gamma_u(s) \) determines the equivalence class \( s \) lies in. For \( x \in A^{k+1} \), let
\[ \gamma_k^u(x) = (\gamma_u(x^1), \gamma_u(x^2), \ldots, \gamma_u(x^k)) \]

Then consider the following \( (n, k) \) code, where we claim that \( \gamma_k^u(x_{P(u)}) \) can be computed by all nodes \( u \), given the initial input \( x \in A^{k+1} \). This claim is proved inductively with the outline of the code.

Each source node \( i \in S \) computes \( \gamma_k^i(x_i) \). Note that there are only \( R_k^{i,j} \leq |B|^{|E_o(i)|} \) possible outputs of \( \gamma_k^i \). So we can encode each output into a distinct string of \( B^{n|E_o(i)|} \), and send the \( n|E_o(i)| \) entries of the string in \( n \) uses of the edges in \( E_o(i) \). Thus the claim holds for all source nodes \( i \).

For an intermediate node \( u \), let \( m = \text{prec}(u) \) and denote \( \text{prec}(u) = \{ v_1, \ldots, v_m \} \). Assume that the claim holds for all \( v \in \text{prec}(u) \). Node \( u \) first recovers \( \gamma_k^u(x_{P(u)}) \) for all \( v \in \text{prec}(u) \) using the symbols received from \( E_o(u) \). This is possible by the induction hypothesis and \( E_o(v) \subseteq E_o(u) \) for all \( v \in \text{prec}(u) \). Then node \( u \) fixes \( s \in A^{k \times |P(u)|} \) such that
\[ \gamma_k^i (s_{P(u)}(P(v))) = \gamma_k^u (x_{P(v)}) \]

for all \( 1 \leq j \leq m \), i.e.,
\[ s^i_{P(u)}(P(v)) = x_{P(v)}, \forall 1 \leq j \leq m, 1 \leq i \leq k. \]

Such an \( s \) can be found by enumerating the matrices in \( A^{k \times |P(u)|} \).

By [4], we can encode each output of \( \gamma_k^u \) into a distinct string of \( B^{n|E_o(u)|} \). In this \( (n, k) \) code, node \( u \) encodes \( \gamma_k^u(s) \) and sends the \( n|E_o(u)| \) entries of the string in \( n \) uses of the edges in \( E_o(u) \).

We claim that \( \gamma_k^u(s) = \gamma_k^u(x_{P(u)}) \). As \( G \) is a tree, we have
\[ P(u) = \cup_1 \leq j \leq m P(v_j) \]

where \( P(v) \cap P(v') = \emptyset \) if \( v \neq v' \). Then define a sequence of strings \( \{a_j\}_{j=0}, a_j \in A^{1 \times |P(u)|} \) as follows:
\[ a_0 = s^i, \quad a_m = x_{P(u)} \]
\[ (a_j)_{P(u)}(P(v)) = \begin{cases} s^j_{P(u)}(P(v)) & l > j; \\ x_{P(v)} & l \leq j. \end{cases} \]

Consider strings \( a_{j-1} \) and \( a_j \), for all \( 1 \leq j \leq m \). In Lemma 2 let
\[ \begin{cases} a' = a_{j-1} , \\ b = a_j , \\ a = (a')_{P(u)}(P(v)) = (s^i)_{P(u)}(P(v)), \\ b = (b')_{P(u)}(P(v)) = (x^j)_{P(v)}, \\ c = (a')_{P(u)}(P(u) \setminus P(v)) = (b')_{P(u)}(P(u) \setminus P(v)) \end{cases} \]

we have \( a' \equiv b'|P(u) \), i.e. \( a_{j-1} \equiv a_j |P(u) \) for all \( 1 \leq j \leq m \). The equivalence then extends to
\[ a_0 = s^i \equiv a_m = x_{P(u)} |P(u) \]

for all \( 1 \leq i \leq k \). Therefore \( s^i \equiv x_{P(u)} |P(u) \) always holds, and we have
\[ \gamma_k^u(s) = \gamma_k^u(x_{P(u)}) \]

finishing the induction.

The sink node can compute \( \gamma_k^u(x) \), identifying \( f^{(k)}(x) \).
that in $N_2$, source nodes 1 and 2 have incoming edges. To
match our model described in Section [17A] we can modify
$N_2$ to $N_2'$ shown in Fig. 2(b), where the number of edges
from node $i$ to node $i'$ is infinity, $i = 1, 2$. Every network
code in $N_2'$ naturally induces a network code in $N_2$ and vice versa. Hence, we have

$$C(N_2, f) = C(N_2', f).$$

We then evaluate min-cut($N_2', f$). Note that

$$\frac{|C|}{\log_{|\mathcal{A}|} W_{C, f}} < \infty$$

holds only if $|C| < \infty$, and we can thus consider only the
finite cut sets. For a finite cut set $C$, we denote by $C' = C \cap \{e_1, \ldots, e_4\}$. We have $|C'| \leq |C|$ and $I(C') \subseteq I(C)$, and we claim $I(C') = I(C)$. Note that $I(C') \subseteq I(C)$. Suppose that there exists $i \in I(C) \setminus I(C')$, then there exists a path from $i$ to $\rho$ which is disjoint with $C'$, but shares a subset $D$ of edges with $C$. Then $D \subset E_i(i)$ and hence $|D| = 1$. We simply replace the edge in $D$ by an arbitrary edge in $E_i(i) \setminus C$ and form a new path from $i$ to $\rho$. This is always possible, since $C \cap E_i(i)$ is finite while $E_i(i)$ is not. The newly formed path is disjoint with $C$, and then we have $i \notin I(C)$, a contradiction.

According to Lemma 3 we have $W_{C', f} \geq W_{C, f}$ and hence

$$\frac{\log_{|\mathcal{A}|} W_{C, f}}{\log_{|\mathcal{A}|} W_{C', f}} \leq \frac{|C'|}{|C|}.$$ 

Therefore we can consider only cut sets $C' \subseteq \{e_1, e_2, e_3, e_4\}$. We then have min-cut($N_2', f$) = 1, which is obtained by the cut set $\{e_2, e_4\}$. While for network $N_2$, it has been proved in [2] that $C(N_2, f) = \log_6 4 < 1$. Hence min-cut($N_2', f$) = 1 > $C(N_2, f)$.

IV. PROOF OF MAIN THEOREM

To prove Theorem 3 we first give the definition of F-
extension and two lemmas.

**Definition 2.** [F-Extension] Given a network $N$ and a cut set $C \in \Lambda(N)$, define $D(C) \subseteq \mathcal{E}$ as

$$D(C) = \bigcup_{i \notin I_C} E_i(i).$$

Then the F-extension of $C$ is defined as

$$F(C) = C \cup D(C).$$

**Lemma 4.** For every cut set $C$, $F(C)$ is a global cut set, i.e.

$$\forall C \in \Lambda(N), I_{F(C)} = S.$$

**Proof:** Clearly, $I_C \subseteq I_{F(C)}$, then it suffices to show that for all $i \notin I_C$, we have $i \in I_{F(C)}$. This is true, since $E_i(i) \subseteq F(C)$ and $i \in E_i(i)$ imply $i \in I_{F(C)}$.

**Lemma 5.** Consider a $(n, k)$ network code in $N = (G, S, \rho)$. For any global cut set $C$, $\psi(x)$ is a function of $g^C(x)$, i.e., $\psi(x) = \psi^C(g^C(x))$ for certain function $\psi^C$.

**Proof:** For a global cut set $C$ of $N$. Let $G_C$ be the
subgraph of $G$ formed by the (largest) connected component
of $G$ including $\rho$ after removing $C$ from $E$. Let $S_C$ be the set
of nodes in $G_C$ that do not have incoming edges. Since $G_C$
also is a DAG, $S_C$ is not empty. For each node $u \in S_C$, we have $i$ is not a source node in $N$ since otherwise $C$ would not be a global cut set, and ii) all the incoming edges of $u$ in $G$ are in $C$ since otherwise $G_C$ can be larger. For each node $u$ in $G_C$ but not in $S_C$, the incoming edges of $u$ are either in $G_C$ or in $C$, since otherwise the cut set $C$ would not be global. If we can show that for any edge $e$ in $G_C$, $g^C(x)$ is a function of $g^C(x)$, then $\psi(x) = \varphi(E_i(\rho))$ is a function of $g^C(x)$.

Suppose that $G_C$ has $K$ nodes. Fix an order on the set of
nodes in $G_C$ that is consistent with the partial order induced
by $G_C$, and number these nodes as $u_1 < \ldots < u_K$, where
$u_K = \rho$. Denote by $E_o(u_C)$ the set of outgoing edges of
$u$ in $G_C$. We claim that $g^{E_o(u_C)}(x)$ is a function of $g^C(x)$
for $i = 1, \ldots, K$, which implies that for any edge $e$ in $G_C$, $g^C(x)$ is a function of $g^C(x)$. We prove this inductively. First $g^{E_o(u_C)}(x)$ is a function of $g^C(x)$ since $u_1 \in S_C$ and hence all the incoming edges of $u_1$ in $G$ are in $C$. Assume that the claim holds for the first $k$ nodes in $G_C, k \geq 1$. For $u_{k+1}$ we have two cases: If $u_{k+1} \in S_C$, the claim holds since all the incoming edges of $u_{k+1}$ in $G$ are in $C$. If $u_{k+1} \notin S_C$, we know that $E_i(u_{k+1}) \subseteq \bigcup_{u_i \in G_C} E_o(u_i) \cup C$. By the induction hypothesis, we have that $E_o(u_{k+1}G_C)$ is a function of $g^C(x)$. The proof is completed.

In the following proof of Theorem 3 it will be handy to extend the equivalence relation for a block of function inputs. For disjoint sets $I, J \in S$ and $c \in \mathcal{A}^{1 \times |I|}$ we say $a, b \in \mathcal{A}^{k \times |I|}$ are $(I, J, c)$-equivalent if for any $x, y \in \mathcal{A}^{k \times s}$, with

$$x_I = a, y_I = b, x_J = y_J = (e^T, e^T, \ldots, e^T)$$

and $x_S \setminus I, J = y_S \setminus I, J$, we have $f(b)(x) = f(b)(y)$. Then for the set $\mathcal{A}^{k \times |I|}$, the number of equivalence classes induced by the equivalence relation is $|W_{c, 1, I, J, f}|^k$.

**Proof of Theorem 3** Suppose that we have a $(n, k)$ code with

$$k / n \log_{|\mathcal{A}|} |A| > \text{min-cut}(N, f).$$
We show that this code cannot compute \( f(x) \) correctly for all \( x \in A^k \times s \). Denote
\[
C^* = \arg \min_{C \in \Lambda(N)} \frac{|C|}{\log |A| W_{C,f}}
\]
(8) and
\[
c^* = \arg \max_{c \in A^{|C^*|}} W_{I_{C^*}, J_{C^*}, f}^c.
\]
(9)
By (7)-(9), we have
\[
k n \log |B| |A| > \frac{|C^*|}{\log |A| W_{I_{C^*}, J_{C^*}, f}}
\]
which leads to
\[
|B|^{|C^*|n} < \left( W_{C^*, f} \right)^k.
\]
(10)

Note that \( g_{C^*}(x) \) only depends on \( x_{K_{C^*}} \). By (10) and the pigeonhole principle, there exist \( a, b \in A^k \times |I_{C^*}| \) such that i) \( a \) and \( b \) are not \( (I_{C^*}, J_{C^*}, c^*) \)-equivalent and ii) \( g_{C^*}(x) = g_{C^*}(y) \) for any \( x, y \in A^k \times s \) with
\[
\begin{align*}
x_{I_{C^*}} &= a, & y_{I_{C^*}} &= b, \\
x_{J_{C^*}} &= y_{J_{C^*}} = (c^*^T, c^*^T, \ldots, c^*^T)^T, \\
x_{S \setminus K_{C^*}} &= y_{S \setminus K_{C^*}}.
\end{align*}
\]
(11)
Fix \( x, y \in A^k \times s \) satisfying (11) and \( f^{(k)}(x) \neq f^{(k)}(y) \). The existence of such \( x \) and \( y \) is due to i). Since \( C^* \) and \( D(C^*) \) are disjoint (see Definition 2) and for any \( i \notin I_{C^*}, x_i = y_i \), together with ii), we have
\[
g_{F(C^*)}(x) = g_{F(C^*)}(y).
\]
Thus, applying Lemma 5 we have \( \psi(x) = \psi(y) \). Therefore, the code cannot compute both \( f^{(k)}(x) \) and \( f^{(k)}(y) \) correctly. The proof is completed.

References

[1] A. Giridhar and P. Kumar, “Computing and communicating functions over sensor networks,” Selected Areas in Communications, IEEE Journal on, vol. 23, no. 4, pp. 755–764, April 2005.

[2] R. Appuswamy, M. Franceschetti, N. Karamchandani, and K. Zeger, “Network coding for computing: Cut-set bounds,” Information Theory, IEEE Transactions on, vol. 57, no. 2, pp. 1015–1030, Feb 2011.

[3] H. Kowshik and P. Kumar, “Optimal function computation in directed and undirected graphs,” Information Theory, IEEE Transactions on, vol. 58, no. 6, pp. 3407–3418, June 2012.

[4] A. Ramamoorthy and M. Langberg, “Communicating the sum of sources over a network,” Selected Areas in Communications, IEEE Journal on, vol. 31, no. 4, pp. 655–665, April 2013.

[5] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, “Network information flow,” IEEE Trans. Inform. Theory, vol. 46, no. 4, pp. 1204–1216, Jul. 2000.

[6] S.-Y. R. Li, R. W. Yeung, and N. Cai, “Linear network coding,” IEEE Trans. Inform. Theory, vol. 49, no. 2, pp. 371–381, Feb. 2003.

[7] R. Koetter and M. Medard, “An algebraic approach to network coding,” IEEE/ACM Trans. Networking, vol. 11, no. 5, pp. 782–795, Oct. 2003.

[8] R. Appuswamy and M. Franceschetti, “Computing linear functions by linear coding over networks,” Information Theory, IEEE Transactions on, vol. 60, no. 1, pp. 422–431, Jan 2014.

[9] B. Rai and B. Dey, “On network coding for sum-networks,” Information Theory, IEEE Transactions on, vol. 58, no. 1, pp. 50–63, Jan 2012.