The Encoding Complexity of Network Coding for Two Simple Multicast Sessions

Wentu Song, Kai Cai, Rongquan Feng, and Wang Rui

Abstract

The intersession network coding problem, which is also known as the multiple source network coding problem is a challenging topic, and has attracted significant attention from the network coding community. In this paper, we study the encoding complexity for intersession network coding with two simple multicast sessions. The encoding complexity is characterized from two aspects: the number of encoding links and the finite field size for achieving a solution. We prove that (1) the number of encoding links required to achieve a solution is upper-bounded by $2N - 1$; and (2) the size of the field required to achieve a linear solution is upper-bounded by $\lceil \sqrt{4N - 31/4 + 3/2} \rceil$, where $|E|$ is the number of edges and $N$ is the number of sinks of the network.

Index Terms

Network coding, encoding edge, information flow, region decomposition.

I. INTRODUCTION

We investigate the multiple source network coding problem, of which the underlying network is assumed to be a finite, directed, acyclic multigraph. A number of messages are generated at some nodes, named sources and demanded by some other nodes, named sinks. The messages are assumed to be elements of a fixed finite alphabet, usually a finite field. A directed edge represents a communication link with unit capacity, i.e., 1 symbol per transmission. A multiple source network coding problem is called solvable if all the source messages can be successfully recovered at the corresponded sinks; otherwise, it is called unsolvable. It is well known that the multiple source network coding problem is in general very challenging [8]. Koetter and Mard [3] showed that to find a solution of a multiple source network coding problem is NP-complete. The recent work of [9] showed that to characterize the achievable rate region is very hard even for the simplest case, i.e., the two simple multicast sessions.

The multicast network coding problem have been extensively studies after the seminal work of Ahlswede, Cai, Li, and Yeung [1] and Li et. al [2]. The encoding complexity is obviously an important issue for network coding, which attracted many researchers [3], [4], [5], [6], [7]. In previous works, the encoding complexity is generally studied from two aspects: the number of encoding links, and the encoding field size for achieving a network solution.

In [3], the authors first categorized the network links into two classes, i.e., the forwarding links and the encoding links. The forwarding links only forward the data of its incoming links. While, the encoding links transmit new packets by combining date from its incoming links, and hence more expensive due to the computing process and the equipping of encoding capabilities. It is shown that, in an acyclic multicast network, the number of encoding links required to achieve the capacity of the network is independent of the size of the underlying network and bounded by $h^3 N^2$, where $N$ is the number of the sinks and $h$ is maximal flow from the source to each of the sinks.

The second issue of encoding complexity is the size of the encoding fields, and this issue has been addressed in many literatures [3], [6], [4], [7]. As stated in [6], large encoding field size may cause difficulties, i.e., either large delays or large bandwidth for implementation of network coding, hence a small alphabet is more preferred. For a multicast network, the required alphabet size to achieve a solution is upper bound by $\sqrt{2N + 1/4 - 1/2}$ (see [7]). In [4], the authors showed that an finite field with size $\sqrt{2N - 7/4 + 1/2}$ is sufficient for achieving a solution of a multicast network.

The intersession network coding problem with two simple multicast sessions was first investigated by C. C. Wang and N. B. Shroff in [9], where they characterized the solvability by using controlled path overlaps conditions and under the assumption of sufficient large encoding fields. They derive a polynomial time algorithm using pebbling games [10] to determine the solvability of such networks. However, they did not consider the issues of the encoding complexity.

In this study, we determine the encoding complexity for intersession network coding with two simple multicast sessions. We will prove that the number of encoding links required to achieve a solution is upper-bounded by $2N - 1$; and (2) the size of the field required to achieve a linear solution is upper-bounded by $\lceil \sqrt{4N - 31/4 + 3/2} \rceil$, where $|E|$ is the number of the edges and $N$ is the number of sinks of the network.

We obtain the encoding complexity by using a region decomposition method, which is an extension of the information flow decomposition approach of C. Fragouli and E. Soljanin [4]. Unlike their method, which decomposes a subgraph of the
underlying network and does not yield unique decomposition in general, the region decomposition method decomposes the whole network uniquely into mutually disjoint regions and can be applied to arbitrary multiple multicast sessions. When the network is divided into particular regions, the messages in each region can then be totally decided by one link, namely, the head of the region. Then, the solution of the network problem can be obtained by a simple labelling operation. To obtain the number of the encoding links and the finite field size for a network solution, we define the region graph, the labelled region graph and obtain the minimal feasible region graph by contraction operations over the feasible region graph. It can be illustrated that the information of the encoding links and the encoding field can be derived by the minimal feasible region graph of the original network.

The rest of the paper is organized as follows. In Section II, we introduce the network models, the notations and the methodology. In Section III, we discuss the encoding complexity from the aspect of encoding links and encoding field respectively. Finally, the paper is concluded in Section IV.

II. NETWORK MODEL, NOTATIONS, AND METHODOLOGY

A. Network Model

In this paper, a network is a finite, directed, acyclic multi-graph \( G = (V, E) \) with node set \( V \) and edge set \( E \). Two mutually independent messages \( X_1, X_2 \) are generated at source nodes \( s_1 \) and \( s_2 \) and are demanded by sink node sets \( T_1 = \{ t_{1,1}, \ldots, t_{1,N_1} \} \) and \( T_2 = \{ t_{2,1}, t_{2,2}, \ldots, t_{2,N_2} \} \) respectively, where \( s_i \notin T_i (i = 1, 2) \). We denote such network as a 2-URS network. The messages are assumed to be elements of a fixed finite field \( F \). A directed edge \((u, v) \in E \) can transmit one symbol from \( u \) to \( v \) per transmission. We add one imaginary incoming edge to each \( s_i \), called the \( X_i \)-source link, and one imaginary outgoing edge to each of the sinks in \( T_i \), called the \( X_i \)-demand link. \( X_i \)-source link and \( X_i \)-demand link are called \( X_i \)-link. And when we say that \( e \) is a source (demand) link we mean that \( e \) is an \( X_1 \)-source (demand) link or an \( X_2 \)-source (demand) link. We use \( \varepsilon \) to denote the union of \( E \) and the imaginary links.

III. MODEL AND BASIC NOTIONS

We consider a communication network modeled as a finite, directed, acyclic multi-graph \( G = (V, E) \), where \( V \) is the set of nodes and \( E \) is the set of edges. There are two source nodes \( s_1 \) and \( s_2 \) and two groups of receiver nodes \( T_1 = \{ t_{1,1}, \ldots, t_{1,N_1} \} \) and \( T_2 = \{ t_{2,1}, t_{2,2}, \ldots, t_{2,N_2} \} \) such that \( s_i \notin T_i, i = 1, 2 \). The message \( X_i \) is generated at source node \( s_i \) and are demanded by receiver nodes in \( T_i \). The messages are assumed to be elements of a fixed finite field \( F \). A directed edge \((u, v) \in E \) can transmit one symbol from \( u \) to \( v \) per transmission. We denote such network as a TSMS network. If \( T_1 = T_2 \), such networks are known as a multicast network, which have been well studied by the research community up to now. If \( T_1 = \{ t_{1,1} \} \) and \( T_2 = \{ t_{2,1} \} \), the network is called a 2-pair network.

We add one imaginary incoming edge to each \( s_i \), called the \( X_i \) source edge, and one imaginary outgoing edge to each of the receiver nodes in \( T_i \), called \( X_i \) receiver edge. We use \( \varepsilon \) to denote the union of \( E \) and the imaginary edges.

Remark 3.1: Throughout this paper, we mean that the number of receiver nodes is \( N_1 + N_2 \), i.e., if a receiver node receives two messages, it is counted twice. Thus, the number of receiver edges is equal to the number of receiver nodes.

For each edge \((u, v) \in \varepsilon \), \( u \) and \( v \) are termed as the tail and the head of \( e \) and are denoted by \( u = \text{head}(e) \) and \( v = \text{tail}(e) \) respectively. Note that the source edges have no tail and the receiver edges have no head. But it doesn’t matter in our discussion. For \( e_1, e_2 \in \varepsilon \), we call \( e_1 \) an incoming edge of \( e_2 \) if \( \text{head}(e_1) = \text{tail}(e_2) \). Denote by \( \text{In}(e) \) the set of incoming edges of \( e \). We assume that \( \text{In}(e) \neq \emptyset \) if \( e \in \varepsilon \) is not a source edge. Otherwise \( e \) has no impact on the network and can be removed from \( G \). A path is an ordered set of edges \( \{ e_1, \ldots, e_m \} \) such that the head of the previous edge is the tail of the next one. It is well known that \( \varepsilon \) can be totally ordered such that \( e_i < e_j \) if \( e_i \) is an incoming edge of \( e_j \).

Definition 3.2: A valid network code of \( G \) over the field \( F \) is a collection of non-zero vectors \( C = \{ d_e \in F^2 : e \in \varepsilon \} \) such that

1. If \( e \) is an \( X_1 \) source edge or an \( X_1 \) receiver edge, then \( d_e = \alpha_e \), where \( \alpha_1 = (1, 0) \) and \( \alpha_2 = (0, 1) \);
2. If \( e \) is not a source edge, then \( d_e \) is an \( F \)-linear combination of \( \{ d_u : u \in \text{In}(e) \} \).

The vector \( d_e \) is called the coding vector of edge \( e \). \( G \) is said to be solvable if \( G \) has a valid network code.

A valid network code of \( G \) specifies a mechanism for message transmission over the network as follows. Let \( x_1, x_2 \in F \) be any instances of \( X_1 \) and \( X_2 \). If \( d_e = (c_1, c_2) \), then the message transmitted on \( e \) is \( c_1 x_1 + c_2 x_2 \), which is a linear combination of the messages transmitted on edges in \( \text{In}(e) \) if \( e \) is not a source edge. The message transmitted on the \( X_1 \)-receiver edge(s) is \( x_1 \). By this mechanism, each receiver node can get its demanded message through the network. We require that the coding vector of each edge \( e \) is non-zero. Otherwise, the message transmitted on \( e \) is zero. Thus, \( e \) has no impact on the network and can be removed from \( G \).

Definition 3.3: Let \( C = \{ d_e \in F^2 : e \in \varepsilon \} \) be a valid network code of \( G \). \( e \in \varepsilon \) is said to be an encoding edge of \( C \) if \( d_e \) is an \( F \)-linear combination of \( \{ d_u : u \in \text{In}(e) \} \) with at least two coefficients not being zero. Else, \( e \) is said to be a forwarding edge of \( C \).

If \( e \) is a forwarding edge of \( C \), then \( d_e = \lambda d_u \) for some \( \lambda \in F \) and \( u \in \text{In}(e) \). Without loss of generality, we can let \( d_e = d_u \).
IV. REGION DECOMPOSITION

In this section, we develop a region decomposition approach for network coding of TSMS networks. Our approach can be viewed as a generalization of the information flow decomposition method [4] for multicast networks and is an efficient method for designing the optimal network codes for TSMS networks.

Definition 4.1: A region is a subset of $\varepsilon$ with an $e_0 \in R$ such that any $e \in R \setminus \{e_0\}$ has an incoming edge in $R$. Or, equivalently, there is a path $\{e_0, e_{i_1}, \ldots, e_{i_k}, e\}$ such that $\{e_0, e_{i_1}, \ldots, e_{i_k}, e\} \subseteq R$.

The edge $e_0$ is called the head of $R$ and is denoted by $e_0 = \text{head}(R)$. A region decomposition of $G$ is a partition of $\varepsilon$ into mutually disjoint regions $D = \{R_1, R_2, \ldots, R_K\}$. $R \in D$ is called an $X_i$ source region ($X_i$ receiver region) if $R$ contains the $X_i$ source edge ($X_i$ receiver edge). It is possible that a region contains both $X_i$ source edge and $X_i$ receiver edge. Hence, an $X_i$ source region may be an $X_i$ receiver region simultaneously. For the sake of convenience, we call $R$ a non-source region if $R$ is not a source region.

Definition 4.2: Let $D$ and $D'$ be two region decompositions of $G$. $D'$ is said to be a contraction of $D$ if any region in $D'$ is either a region in $D$ or the union of some regions in $D$.

For each $e \in \varepsilon$, let $R_e = \{e\}$. Then $R_e$ is a region of $G$ with head($R_e$) = $e$ and $D_0 = \{R_e : e \in \varepsilon\}$ is a region decomposition of $G$. We call $D_0$ the trivial region decomposition of $G$. Any region decomposition of $G$ is a contraction of $D_0$.

It is more transparent to work with the so-called region graph of $G$. Let $D$ be a region decomposition of $G$. A region graph of $G$ with respect to $D$, denoted by $RG(D)$, is defined as a directed graph with node set $D$ and edge set $\{(R_i, R_j) \in D^2 : \text{head}(R_j) \text{ has an incoming edge in } R_i\}$. $R_i$ is said to be a parent of $R_j$ ($R_j$ is a child of $R_i$) if $R_i, R_j$ is an edge of $RG(D)$. Denoted by $\text{Par}(R_i)$ the set of all parents of $R_i$. Clearly, $RG(D)$ is acyclic and simple. So $D$ can be totally ordered such that $R_i < R_j$ if $R_i$ is a parent of $R_j$. $RG(D_0)$ coincides with the so-called line graph of $G$, where $D_0$ is the trivial region decomposition of $G$. If $D'$ is a contraction of $D$, we say that $RG(D')$ is a contraction of $RG(D)$.

Definition 4.3: Let $D$ be a region decomposition of $G$. $RG(D)$ is said to be feasible if there is a collection of non-zero vectors $C = \{d_R \in F^2 : R \in D\}$ such that

1. If $R$ is an $X_i$ source region or an $X_i$ receiver region, then $d_R = a_1$, where $a_1 = (1, 0)$ and $a_2 = (0, 1)$;
2. If $R$ is not a source region, then $d_R$ is an $F$-linear combination of $\{d_P : P \in \text{Par}(R)\}$.

The vector $d_R$ is called the coding vector of region $R$ and $C$ is called a valid network code of $RG(D)$.

Clearly, a valid network code of $RG(D_0)$ is exactly a valid network code of $G$, where $D_0$ is the trivial region decomposition of $G$. In general, if $C = \{d_c : e \in \varepsilon\}$ is a valid network code of $G$ and $D$ is a region decomposition of $G$ such that $d_c = d_{\text{head}(R)}$ for all $R \in D$ and $e \in R$, then $\tilde{C} = \{d_R : d_R = d_{\text{head}(R)}, R \in D\}$ is a valid network code of $RG(D)$. Conversely, suppose $D$ is a region decomposition of $G$ and $C = \{d_R : R \in D\}$ is a valid network code of $RG(D)$. Let $C = \{d_c : e \in \varepsilon\}$ be such that $d_c = d_R$ if $e \in R$ and $R \in D$. Then $C$ is a valid network code of $G$.

Example 4.4: Let $G$ be a TSMS network shown in Fig.1 (a). The message $X_1$ is generated at source node $s_1$ and is demanded by receiver nodes $t_1$ and $t_3$. The message $X_2$ is generated at $s_2$ and is demanded by $t_2$ and $t_3$. The imaginary edges $e_1$ and $e_2$ are the $X_1$ source edge and $X_2$ source edge respectively. The imaginary edges $e_{21}$ and $e_{23}$ are the $X_1$ receiver edges. The imaginary edges $e_{20}$ and $e_{22}$ are the $X_2$ receiver edges. A valid network code $C = \{d_c : e \in \varepsilon\}$ of $G$ is shown in Fig.1 (b). Let $R_1 = \{e_1, e_3, e_5, e_8, e_9, e_{11}\}$, $R_2 = \{e_2, e_4, e_7, e_{10}\}$, $R_3 = \{e_6, e_{13}, e_{14}\}$, $R_4 = \{e_{12}, e_{16}, e_{21}\}$, $R_5 = \{e_{15}, e_{17}, e_{18}, e_{23}\}$, $R_6 = \{e_{20}\}$, $R_7 = \{e_{19}\}$, $R_8 = \{e_{22}\}$. Then $D = \{R_1, R_2, R_3, R_4, R_5, R_6, R_7, R_8\}$ is a region decomposition of $G$. The region graph $RG(D)$ is shown in Fig.2 (a) and the valid network code $C = \{d_R : R \in D\}$ of $RG(D)$ derived from $C$ is shown in Fig.2 (b). Conversely, $C$ can be derived from $C$ by letting $d_c = d_R$ if $e \in R$ and $R \in D$.
unchanged results in a valid network code of $RG$ if $RG$ is a valid network code of $G$.

**Lemma 4.5:** Let $D'$ be a contraction of $D$. If $RG(D')$ is feasible, then $RG(D)$ is feasible.

**Proof:** Suppose $\tilde{C}' = \{d_{R'} : R' \in D'\}$ is a valid network code of $RG(D')$. Let $\tilde{C} = \{d_R : R \in D\}$ be such that $d_R = d_{R'}$ if $R \subseteq R'$ and $R' \in D'$. $\tilde{C}$ is well defined because $D'$ is a contraction of $D$. It is easy to see that $\tilde{C}$ is a valid network code of $RG(D)$. Hence, $RG(D)$ is feasible.

**Lemma 4.6:** Let $D$ be a region decomposition of $G$ and $P, Q \in D$ be adjacent in $RG(D)$. Suppose $\tilde{C} = \{d_R : R \in D\}$ is a valid network code of $RG(D)$ such that $d_P = d_Q$. Let $P' = P \cup Q$ and $D' = \{P'\} \cup D \setminus \{P, Q\}$. then $D'$ is a contraction of $D$ and $RG(D')$ is feasible.

**Proof:** Without loss of generality, we can assume that $P$ is a parent of $Q$. Clearly, $P \cup Q$ is a region of $G$ with $head(P \cup Q) = head(P)$ (definition 4.1) and $D' = \{P \cup Q\} \cup D \setminus \{P, Q\}$ is a region decomposition of $G$. By definition 4.2, $D'$ is a contraction of $D$. Note that $d_P = d_Q$. Letting $d_{P \cup Q} = d_P$ and keeping the coding vectors of regions in $D \setminus \{P, Q\}$ unchanged results in a valid network code of $RG(D')$. Hence $RG(D')$ is still feasible.

**Example 4.7:** We consider again the network $G$ in example 4.4. Note that $d_{R_1} = d_{R_4}$ and $R_1, R_4$ are adjacent; $d_{R_3} = d_{R_7}$ and $R_3, R_7$ are adjacent. Let $D' = \{R_1 \cup R_4, R_2, R_3 \cup R_7, R_5, R_6, R_8\}$. Then $D'$ is a contraction of $D$ and $D'$ is still feasible. The region graph $RG(D')$ and the corresponding valid network code of $RG(D')$ are shown in Fig.3 (a) and Fig. (b) respectively. Note that $R_1 \cup R_4$ is both $X_1$ source region and $X_1$ receiver region of $RG(D')$. 

![Fig.1: (a) is a feasible TSMS network $G$; (b) shows a valid network code of $G$.](image1)

![Fig.2: (a) is the region graph $RG(D)$ of $G$; (b) shows a valid network of $RG(D)$.](image2)
By definition \( G \) is solvable if and only if \( G \) has a feasible region graph. We remark that this result is useful because there is a simple method to determine whether a region graph is feasible (theorem 4.11). However, we are more interested in the problem that if we can find the "simplest" feasible region graphs of \( G \) and what properties the "simplest" feasible region graphs may have. This is an issue we shall discuss in section IV. We now give a characterization of feasible region graph of \( G \). The following definition is needed.

Definition 4.8: Let \( D \) be a region decomposition of \( G \) and \( i \in \{1, 2\} \). The \( X_i \)-type region of \( RG(D) \) is defined recursively as follows.

1. An \( X_i \) source region or an \( X_i \) receiver region is an \( X_i \)-type region;
2. A region whose parents are all \( X_i \)-type region is an \( X_i \)-type region.

A region which is neither \( X_1 \)-type region nor \( X_2 \)-type region is called a coding region. A region which is both \( X_1 \)-type region and \( X_2 \)-type region is called a singular region. The following lemma is obvious.

Lemma 4.9: Suppose \( D \) is a region decomposition of \( G \) and \( RG(D) \) is feasible. Let \( \tilde{C} = \{d_R : R \in D\} \) be a valid network code of \( RG(D) \). Then \( d_R = \alpha_i \) for all \( X_i \)-type region \( R \), \( i \in \{1, 2\} \). In particular, \( RG(D) \) has no singular region.

Lemma 4.10: Let \( D \) be a region decomposition of \( G \). Then \( RG(D) \) is feasible if the following two conditions hold.

1. \( RG(D) \) has no singular region;
2. Any non-source region in \( D \) has at least two parents in \( RG(D) \).

Proof: We are to construct a valid network code of \( RG(D) \) as follows. First, let \( d_R = \alpha_1 \) for all \( X_1 \)-type region \( R \), \( i \in \{1, 2\} \). This is reasonable because \( RG(D) \) has no singular region. Next, we select a coding vector for each coding region of \( RG(D) \). Assume that \( \{R_1, R_2, \ldots, R_K\} \) is the set of coding regions of \( RG(D) \).

Let \( d_{R_k} \) be any fixed vector of \( F^2\setminus\{(\alpha_1)\cup(\alpha_2)\} \). For \( 2 \leq k \leq K \), suppose \( d_{R_j} \) has been determined for all \( j \in \{1, \ldots, k-1\} \), let \( d_{R_k} \in F^2\setminus\{(\alpha_1)\cup(\alpha_2)\} \) be such that for any \( j \in \{1, \ldots, k-1\} \), \( d_{R_k} \) and \( d_{R_j} \) are linearly independent if \( R_k \) and \( R_j \) have a common child. We can select \( d_{R_k} \) correctly if the size of field \( F \) is sufficiently large. We now prove that \( \tilde{C} = \{d_R : R \in D\} \) is a valid network code of \( RG(D) \).

Clearly, by the construction of \( C \), \( d_R = \alpha_i \) if \( R \in D \) is an \( X_i \) source region or an \( X_i \) receiver region. If \( R \) is not a source region, by condition (2), \( R \) has at least two parents in \( RG(D) \). We have the following two cases.

Case 1: The parents of \( R \) are all \( X_i \)-type region for a fixed \( i \in \{1, 2\} \). By definition 4.8, \( R \) is an \( X_i \)-type region. By the construction of \( C \), \( d_R = d_{P_1} = \alpha_i \) for all \( P \in Par(R) \).

Case 2: There are two parents of \( R \), say \( P_1 \) and \( P_2 \), such that \( P_1 \) and \( P_2 \) are both coding regions or \( P_1 \) and \( P_2 \) are of different types. By the construction, \( d_{P_1} \) and \( d_{P_2} \) are linearly independent, hence span \( F^2 \). So \( d_R \) is a linear combination of \( d_{P_1} \) and \( d_{P_2} \).

By definition 4.8, \( \tilde{C} \) is a valid network code of \( RG(D) \) and \( RG(D) \) is feasible.

Theorem 4.11: Let \( D \) be a region decomposition of \( G \). Then \( RG(D) \) is feasible if and only if \( D \) has a contraction \( D' \) such that

1. \( D' \) contains no singular region;
2. Any non-source region in \( D' \) has at least two parents in \( RG(D') \).

In particular, if the two conditions hold, \( RG(D') \) is feasible.

Proof: Suppose \( RG(D) \) is feasible. Let \( \tilde{C} = \{d_R : R \in D\} \) be a valid network code of \( RG(D) \). If \( Q \in D \) has only one parent \( P \) in \( RG(D) \), then by definition 4.3, \( d_Q \) is a linear combination of \( d_P \). Without loss of generality, we can assume that \( d_Q = d_P \). By lemma 4.6, \( D' = \{P'\} \cup D \setminus \{P, Q\} \) is a contraction of \( D \) and \( RG(D') \) is feasible, where \( P' = Q \cup P \). So \( D' \) contains no singular region (lemma 4.9). This operator can be performed continuously until any non-source region in \( D' \) has at least two parents in \( RG(D') \), which completes one direction of the proof.

Conversely, suppose \( D' \) is a contraction of \( D \) satisfying conditions (1) and (2). By lemma 4.10, \( RG(D') \) is feasible. Hence, by lemma 4.5, \( RG(D) \) is feasible.

Corollary 4.12: Let \( D \) be a region decomposition of \( G \) such that \( RG(D) \) is feasible. Suppose \( R_0 \in D \) is a coding region and \( i \in \{1, 2\} \) is a fixed number such that any child of \( R_0 \) not being an \( X_i \)-type region has a parent not being an \( X_i \)-type region. Then \( RG(D) \) has a valid network code \( \tilde{C}' = \{d_R \in F^2: R \in D\} \) such that \( d_{R_0} = \alpha_i \), where \( \alpha_i \) is the vector of \( F^2 \) with the \( i \)th component being one and the other component being zero.
Proof: Consider a valid network code $\tilde{C} = \{d_R \in F^2 : R \in D\}$ of $RG(D)$ constructed as in lemma 4.10. We alter $\tilde{C}$ by letting $d_{R_0} = \alpha_1$ and keeping the rest of coding vectors unchanged. We show that the resulted code $\tilde{C}'$ is still a valid network code of $RG(D)$. Clearly, condition (1) of definition 4.3 still holds. For any non-source region $R \in D$, we have the following four cases.

Case 1: $R = R_0$. Since $R$ is a coding region, by the construction $\tilde{C}$, $R$ has two parents, say $P_1$ and $P_2$, such that $d_{P_1}$ and $d_{P_2}$ are linearly independent and form a basis of $F^2$. Hence, $d_{R_0} = \alpha_1$ is a linear combination of $d_{P_1}$ and $d_{P_2}$.

Case 2: $R$ is a child of $R_0$ and $R$ is an $X_i$-type region. By the construction of $\tilde{C}$, $d_R = \alpha_i$. Thus $d_R$ is a linear combination of $d_{R_0} = \alpha_1$.

Case 3: $R$ is a child of $R_0$ and $R$ is not an $X_i$-type region. By the assumption of this lemma, $R$ has a parent $P$ not being an $X_i$-type region. By the construction of $\tilde{C}$, $d_P$ and $\alpha_i$ are linearly independent and form a basis of $F^2$. Hence, $d_R$ is a linear combination of $d_P$ and $d_{R_0} = \alpha_1$.

Case 4: $R \neq R_0$ and $R$ is not a child of $R$. Since $\tilde{C}$ is a valid network code of $RG(D)$ and the coding vector(s) of $R$ and $R$’s parent(s) keep unchanged, $d_R$ is a linear combination of the coding vector(s) of $R$’s parent(s).

Thus condition (2) of definition 4.3 also holds. So $\tilde{C}'$ is a valid network code of $RG(D)$ and $d_{R_0} = \alpha_1$.

Corollary 4.13: Let $D$ be a region decomposition of $G$ such that $RG(D)$ is feasible. Suppose $P, Q \in D$ are two coding regions such that $P$ and $Q$ have no common child. Then $RG(D)$ has a valid network code $\tilde{C}' = \{d_R \in F^2 : R \in D\}$ such that $d_P = d_Q$.

Proof: Consider a valid network code $\hat{C} = \{d_R \in F^2 : R \in D\}$ of $RG(D)$ constructed as in lemma 4.10. Let $\alpha \in F^2 \setminus ((\alpha_1) \cup (\alpha_2))$ be such that for any coding region $R \in D$, $\alpha$ and $d_R$ are linearly independent if $P$ and $R$ have a common child or $Q$ and $R$ have a common child. We alter $\hat{C}$ by letting $d_P = d_Q = \alpha$ and keeping the rest of coding vectors unchanged. Since $P$ and $Q$ have no common child, it is easy to see that the resulted code $\tilde{C}'$ is a valid network code of $RG(D)$.

V. Encoding Complexity

In this section, we investigate the encoding complexity of network coding for TSMS networks. We shall prove an upper bound on the number of encoding edges as well as an upper bound on the field size needed to achieve a valid network code of a TSMS network. The following definition is needed.

Definition 5.1: Let $D$ be a region decomposition of $G$ such that $RG(D)$ is feasible. $RG(D)$ is said to be minimal feasible if the following two conditions hold.

1) Deleting any edge of $RG(D)$ results in a subgraph of $RG(D)$ which is not feasible.

2) For any proper contraction $D'$ of $D$ (i.e., $D'$ is a contraction of $D$ and $D' \neq D$), $RG(D')$ is not feasible.

Lemma 5.2: Suppose $RG(D)$ is minimal feasible and $\hat{C} = \{d_R \in F^2 : R \in D\}$ is a valid network code of $RG(D)$. If $P, Q \in D$ are adjacent or have a common child, then $d_P$ and $d_Q$ are linearly independent.

Proof: Suppose $d_P$ and $d_Q$ are linearly dependent. Without loss of generality, we can assume that $d_P = d_Q$.

If $P$ is a parent of $Q$, by lemma 5.2 $D$ has a proper contraction $D'$ and $RG(D')$ is feasible, which contradicts to definition 5.1.

If $P$ and $Q$ have a common child $R_0$, then the subgraph obtained by deleting the edge between $R_0$ and $Q$ has $\hat{C}$ as a valid network code, hence is still feasible, which contradicts to definition 5.1.

For the sake of convenience, we say that a region $Q$ is an $X_i$-parent (or an $X_i$-child) of a region $R$ if $Q$ is a parent (or a child) of $R$ and $Q$ is an $X_i$-type region.

Theorem 5.3: Let $D$ be a region decomposition of $G$ such that $RG(D)$ is minimal feasible. The following items hold:

1) Any non-source region $R$ in $D$ has exactly two parents in $RG(D)$.

2) Two regions which are adjacent or have a common child can’t be both $X_i$-type region (or both $X_2$-type region).

3) Two adjacent coding regions have a common child.

4) If a coding region has an $X_i$-parent or an $X_i$-child, then it has a child not being an $X_i$-type region and having an $X_i$-parent, where $i \in \{1, 2\}$ is a fixed number.

Proof: 1) Since $RG(D)$ is minimal feasible, by theorem 4.1, any non-source region $Q$ has at least two parent in $RG(D)$. Let $\hat{C} = \{d_R \in F^2 : R \in D\}$ be a valid network code of $RG(D)$. Since the dimension of $F^2$ is 2, there are two parents of $Q$, say $P_1$ and $P_2$, such that $d_{P_1} = d_{P_2} = \{(d_P : P \in Par(Q))\}$. If $Q$ has more than two parents, then deleting the edge(s) between $Q$ and all of its parents but $P_1$ and $P_2$ results in a subgraph which has $\hat{C}$ as a valid network code and is still feasible, which contradicts to definition 5.1.

2) This claim is a direct consequence of lemma 4.9 and lemma 5.2.

3) Suppose the converse is true, i.e., there are two adjacent coding regions $P, Q \in D$ such that $P$ and $Q$ have no common child. By corollary 4.13 $RG(D)$ has a valid network code $\tilde{C} = \{d_R : R \in D\}$ such that $d_P = d_Q$, which contradicts to lemma 5.2.

4) Suppose the converse is true, i.e., there is a coding region $R_0 \in D$ such that $R_0$ has an $X_i$-parent (or an $X_i$-child) $P$ and any child $Q$ of $R_0$ which is not an $X_i$-type region has no $X_i$-parent. By claim 1), $Q$ has two parents in $RG(D)$. So $Q$
has a parent not being \( X_i \)-type region. By corollary 4.12, \( RG(D) \) has a valid network code \( \tilde{C} = \{ d_R : R \in D \} \) such that \( d_{R_0} = \alpha_1 \). But \( d_R = \alpha_1 \) (lemma 4.9) and \( P, R_0 \) are adjacent, which contradicts to lemma 5.2.

**Corollary 5.4:** Let \( D \) be a region decomposition of \( G \) such that \( RG(D) \) is minimal feasible. The following items hold.

1. An \( X_i \)-type region is either an \( X_i \) source region or an \( X_i \) receiver region \((i \in \{1, 2\})\).
2. A coding region has at least two children being receiver regions.
3. There exists a coding region which has two children, say \( R_1 \) and \( R_2 \), such that for each \( i \in \{1, 2\}, R_i \) is an \( X_i \) receiver region and \( R_i \) has an \( X_j \)-parent, \( j \in \{1, 2\} \) and \( j \neq i \).

**Proof:** 
1) This claim is a direct consequence of Definition 4.8 and 2) of Theorem 5.3.

2) let \( R \) be a coding region of \( RG(D) \). By Theorem 5.3, \( R \) has two parents, say \( P_1 \) and \( P_2 \), such that \( P_1 \) and \( P_2 \) can’t be both \( X_i \)-type region or both \( X_j \)-type region. We have the following three cases.

- Case 1: \( P_1 \) is an \( X_i \)-type region for some \( i \in \{1, 2\} \) and \( P_2 \) is a coding region. First, consider \( P_1 \) and \( R \). By 4) of theorem 5.3, \( R \) has a child \( Q_1 \) such that \( Q_1 \) is not an \( X_i \)-type region and \( Q_1 \) has an \( X_j \)-parent. If \( Q_1 \) is an \( X_j \)-type region for some \( j \in \{1, 2\} \), by claim 1), \( Q_1 \) is an \( X_j \)-receiver region. If \( Q_1 \) is a coding region, then by 3) of theorem 5.3, \( R \) and \( Q_1 \) have a common child, say \( Q_2 \). Similarly, either \( Q_2 \) is an \( X_j \)-receiver region or \( R \) and \( Q_2 \) have a common child \( Q_3 \). Since \( RG(D) \) is a finite graph, we can finally find a \( Q_m \) such that \( Q_m \) is an \( X_j \)-receiver region. Next, consider \( P_2 \) and \( R \). By 3) of theorem 5.3, \( P_2 \) and \( R \) have a common child \( W_1 \). Also, either \( W_1 \) is an \( X_l \)-receiver region for some \( l \in \{1, 2\} \) or \( R \) and \( W_1 \) have a common child \( W_2 \). Similarly, we can finally find a \( W_n \) such that \( W_n \) is an \( X_l \)-receiver region. Note that \( Par(Q_1) \) contains an \( X_i \)-type region but \( Par(W_1) = \{P_2, R\} \). So \( Q_1 \neq W_1 \). We can prove inductively that \( Q_m \neq W_n \). Hence \( Q_m \) and \( W_n \) are two children of \( R \) being receiver regions.

- Case 2: Both \( P_1 \) and \( P_2 \) are coding regions.

- Case 3: \( P_1 \) is an \( X_j \)-type region and \( P_2 \) is an \( X_j \)-type region.

In the latter two cases, by the similar analysis, we can find two children of \( R \) being receiver regions.

3) Without loss of generality, we can assume that each coding region \( R \) has at least one child. Otherwise, \( R \) has no impact on the network and can be removed from \( RG(D) \). Since \( RG(D) \) is a finite graph, we can find a coding region \( R_0 \) such that no child of \( R_0 \) is coding region. Let \( Q \) be a child of \( R_0 \). Then \( Q \) is an \( X_i \)-type region for some \( i \in \{1, 2\} \). Without loss of generality, we can assume that \( Q \) is an \( X_i \)-type region. By 4) of theorem 5.3, \( R_0 \) has a child \( R_2 \), such that \( R_2 \) is not an \( X_i \)-type region and \( R_2 \) has an \( X_j \)-parent. Similarly, since \( R_0 \) is a coding region, \( R_2 \) is an \( X_j \)-type region. By claim 1), \( R_2 \) is an \( X_j \)-receiver region. Again, by 4) of theorem 5.3, \( R_0 \) has a child \( R_1 \) such that \( R_1 \) is an \( X_i \) receiver region and \( R_1 \) has an \( X_j \)-parent.

**Theorem 5.5:** Let \( D \) be a region decomposition of \( G \) such that \( RG(D) \) is minimal feasible. Then the number of coding regions is smaller than the number of receiver regions.

**Proof:** Let \( K \) be the number of coding regions and \( N \) be the number of receiver regions. Denoted by \( J \) the number of edges in \( RG(D) \) from a coding region to a receiver region. We count \( J \) in two different ways. By 1) of theorem 5.3, \( N \) receiver regions have \( 2N \) incoming edges. By 3) of corollary 5.4, there are two receiver regions each of which have an incoming edge not being from a coding region. So \( J \leq 2N - 2 \). On the other hand, by 2) of corollary 5.4, \( K \) coding regions have at least \( 2K \) outgoing edges pointing to receiver regions. So \( 2K \leq 2N \). Thus, \( 2K \leq 2N - 2 \) and \( K \) is smaller than \( N - 1 \). Hence \( K \) is smaller than \( N - 1 \).

**Theorem 5.6:** Let \( G \) be a solvable TSMS network, then \( G \) has a valid network with at most \( 2N - 1 \) encoding edges, where \( N \) is the number of receiver nodes.

**Proof:** By remark 5.1, we only need to prove that \( G \) has a valid network with at most \( 2N - 1 \) encoding edges, where \( N \) is the number of receiver nodes. Let \( D \) be a region decomposition of \( G \) such that \( RG(D) \) is minimal feasible. Let \( C = \{ d_R : R \in D \} \) be a valid network code of \( RG(D) \) and \( C = \{ d_e : e \in e \} \) be the valid network code of \( G \) derived from \( C \) by letting \( d_e = d_R \) if \( e \in R \) and \( R \in D \). Clearly, an edge \( e \) is an encoding edge of \( C \) if and only if \( e = head(R) \) for some non-source region \( R \). By 1) of corollary 5.4, the number of non-source regions is the total number of coding regions and receiver regions. Note that the number of coding regions is smaller than or equal to \( N - 1 \) (theorem 5.5) and the number of receiver regions is smaller than or equal to \( N \). The conclusion follows.

We finally investigate the problem of the field size of a valid network code of a solvable TSMS network. Let \( G \) be a solvable TSMS network and \( D \) be a region decomposition of \( G \) such that \( RG(D) \) is minimal feasible and has \( K \) coding regions, say \( R_1, R_2, \ldots, R_K \). As in the proof of lemma 4.10, constructing a valid network code of \( RG(D) \) is equivalent to find \( K \) vectors \( \{d_{R_1}, d_{R_2}, \ldots, d_{R_K}\} \subseteq F^2\setminus(\langle\alpha_1\rangle \cup \langle\alpha_2\rangle) \) such that \( d_{R_i} \) and \( d_{R_j} \) are linearly independent if \( R_i \) and \( R_j \) have a common child, where \( \alpha_1 = (1, 0) \) and \( \alpha_2 = (0, 1) \). By lemma 5.2, \( d_{R_i} \) and \( d_{R_j} \) are linearly independent if \( R_i \) and \( R_j \) are adjacent. However, we need not worry about this case separately since by theorem 5.5, two adjacent coding regions have a common child.

Let \( \Omega_D \) be an undirected graph with \( \{R_1, R_2, \ldots, R_K\} \) being vertex set. We connect two vertices in \( \Omega_D \) with an edge if the corresponding coding regions in \( RG(D) \) have a common child. Then constructing a valid network code of \( RG(D) \) is equivalent to vertex coloring of \( \Omega_D \) with vectors in \( F^2 \setminus (\langle\alpha_1\rangle \cup \langle\alpha_2\rangle) \). It is known that the number of mutually linearly independent coding vectors is the chromatic number \( \chi(\Omega_D) \).

**Lemma 5.7:** Let \( F \) be a field of size \( q \), then there are \( q - 1 \) vectors in \( F^2 \setminus (\langle\alpha_1\rangle \cup \langle\alpha_2\rangle) \) which are mutually linearly independent.
Proof: Suppose $F = \{0, x_1 = 1, x_2, \cdots , x_{q-1}\}$. Let \( \beta_k = (1, x_k), k = 1, 2, \cdots , q - 1 \). Then \( \{\beta_1, \beta_2, \cdots , \beta_{q-1}\} \subseteq F^2 \setminus (\{\alpha_1\} \cup \{\alpha_2\}) \) are mutually linearly independent.

By lemma 5.7, a field of size $q = k + 1$ is sufficient to achieve a valid network code of $RG(D)$, where $k$ is the chromatic number $\chi(\Omega_D)$.

Lemma 5.8: [17, Ch. 9] Every $k$-chromatic graph has at least $k$ vertices of degree at least $k - 1$.

Lemma 5.9: Let $G$ be a TSMS network and $D$ be a region decomposition of $G$. There are two coding regions of $RG(D)$ which are not the common child of any two coding regions.

Proof: Let $R_1, R_2, \cdots , R_K$ be the set of coding regions of $RG(D)$. Without loss of generality, we can assume that $i < j$ if $R_i$ is a parent of $R_j$. Clearly, $R_1$ and $R_2$ can't be the common child of any two coding regions.

Theorem 5.10: For any solvable TSMS network $G$ with $N$ receiver edges, there exists a valid network code of $G$ over the field of size $\sqrt{4N - 39/4 + 3/2}$.

By lemma 5.7, a field of size $q = k + 1$ is sufficient to achieve a valid network code of $RG(D)$, where $k$ is the chromatic number $\chi(\Omega_D)$.

Lemma 5.11: [17, Ch. 9] Every $k$-chromatic graph has at least $k$ vertices of degree at least $k - 1$.

Lemma 5.12: Let $G$ be a TSMS network and $D$ be a region decomposition of $G$. There are two coding regions of $RG(D)$ which are not the common child of any two coding regions.

Proof: Let $R_1, R_2, \cdots , R_K$ be the set of coding regions of $RG(D)$. Without loss of generality, we can assume that $i < j$ if $R_i$ is a parent of $R_j$. Clearly, $R_1$ and $R_2$ can't be the common child of any two coding regions.

Theorem 5.13: For any solvable TSMS network $G$ with $N$ receiver nodes, there exists a valid network code of $G$ over the field of size $\sqrt{4N - 39/4 + 3/2}$.

Proof: By remark 3.1, $N$ is the number of receiver edges. Let $D$ be a region decomposition of $G$ such that $RG(D)$ is minimal feasible with $K$ coding regions. Let $k$ be the chromatic number $\chi(\Omega_D)$ and $J$ be the number of edges of $\Omega_D$. We count $J$ in two different ways. By lemma 5.11 we have

\[ k(k - 1)/2 \leq J. \] (1)

On the other hand, a region is a common child of two coding only if it is a non-source region, i.e., it is a coding region or a receiver region (corollary 5.4). By the assumption, there are $K$ coding regions and at most $N$ receiver regions of which two coding regions and two receiver regions are not the common child of any two coding regions (lemma 5.12 and corollary 5.4). So there are at most $K + N - 4$ regions which are the common child of two coding regions. Thus

\[ J \leq K + N - 4. \] (2)

From (1) and (2), we obtain

\[ k(k - 1)/2 \leq K + N - 4. \] (3)

But $K \leq N - 1$ (theorem 5.5). So (3) implies

\[ k(k - 1)/2 \leq 2N - 5. \] (4)

Solving the inequality (4) for $k$ we have

\[ k \leq \sqrt{4N - 39/4 + 1/2} \]

and

\[ q = k + 1 \leq \sqrt{4N - 39/4 + 3/2}. \]

VI. CONCLUSION

In this paper, we consider the encoding complexity of network coding for solvable TSMS networks. We proved that the number of encoding edges is upper bounded by $2N - 1$ and a field of size $\sqrt{4N - 39/4 + 3/2}$ is sufficient for a valid network code, where $N$ is the number of the receiver nodes and a receiver node is counted two times if it receives both messages.

REFERENCES