Exponentially Simpler Network Rate Regions

Yirui Liu and John MacLaren Walsh
Department of Electrical and Computer Engineering
Drexel University
Philadelphia, PA 19104, USA
Email: {yl926, jmw96}@drexel.edu

***Abstract*—Determining the rate region of a network is of great importance in the research area of network coding. Lots of attempts have been made and significant progress has been achieved over the last decade on this topic. Although these researches provide us with multiple ways of calculating the outer or inner bounds of rate region, the sheer complexity of the problem, which involves expressing and projecting a very high dimensional polyhedra, makes it computationally infeasible beyond networks with 10s of edges.**

Aimed at reducing the complexity of the rate region calculation, in this paper a new theorem that implicitly determines the rate region of a network is proved and a corresponding systematic way of applying the theorem to calculate explicitly the outer bounds to a rate region is proposed. Compared with the traditional method, the proposed method has the potential to calculate the true rate region via the projection of simpler polyhedra that has exponentially less dimensions and is characterized by exponentially less facets.

***Index Terms*—network coding cuts, rate regions, complexity reduction**

I. Introduction

Determining the rate region of a network is of great importance in the research area of network coding. Lots of attempts have been made and significant progress has been achieved over the last decade on this topic [1]–[4]. From an applied perspective, understanding network coding problems will help us to determining the fundamental limits as well as to constructing practical designs that approach them, for engineering problems range form index coding [5], [6], coded caching [7] to distributed storage system [8]–[11], to delay mitigating codes and delay tradeoffs for streaming information [12]–[14]. Although these researches provide us with multiple ways of calculating the outer or inner bounds of rate region, the sheer complexity of the problem [15], which involves expressing and projecting a very high dimensional polyhedra, makes it computationally infeasible beyond networks with 10s of edges.

Aimed at reducing the complexity of the rate region calculation, in this paper a new theorem that implicitly determines the rate region of a network is proved and a corresponding systematic way of applying the theorem to calculate explicitly the outer bounds to a rate region is proposed. Compared with the traditional method, the proposed method has the potential to calculate the true rate region via the projection of simpler polyhedra that has exponentially less dimensions and is characterized by exponentially less facets. This enables our new method to finish way sooner than the traditional one when tested on rate region calculation software. Besides, an inductive argument is given at the end of the paper to show that there are infinitely many networks that our new method can be applied upon.

II. Background

This paper focuses on multi-source multi-sink network coding problems on acyclic networks with directed hyperedges, which we hereafter refer to as the MSNC problems.

A. MSNC Network

The purpose of a network is to capture a structure of information exchange whose goal is to enable some sink demands to be met. Thus, the MSNC network studied in this paper is defined primarily to indicate what must be encoded from what, and what must be capable of being decoded from what, in a series of message exchanges.

Definition 1. *(MSNC network) A MSNC hyperedge network* A *is a directed hyperedge graph defined as a tuple* $\mathsf{A} = (\mathcal{S}, \mathcal{E}, \mathcal{T}, \mathcal{Q}, \mathcal{W})$ *with* $\mathcal{S}$, $\mathcal{E}$ *and* $\mathcal{T}$ *indicating three different types of hyperedges,* $\mathcal{Q}$ *and* $\mathcal{W}$ *indicating two different types of nodes. In particular,* $\mathcal{S} = \{1, ..., K\}$ *is defined as labels of source messages,* $\mathcal{E} = \{K+1, ..., K+L\}$ *as labels of edge messages,* $\mathcal{T} \subseteq 2^{\mathcal{S}}$ *as labels of sink demands. Each node* g *in the network is labeled as an order pair of incoming messages* $in(g) \subseteq \mathcal{S} \cup \mathcal{E}$ *and outgoing messages* $out(g) \subseteq (\mathcal{S} \cup \mathcal{E}) \backslash in(g)$ *as* $g = (out(g), in(g))$*. Let* $\mathcal{G}$ *be the collection of all the node* g *in network* A*, then the labels of sink nodes (decoders)* $\mathcal{W}$ *is defined as* $\mathcal{W} = \{g \in \mathcal{G} | out(g) \in \mathcal{T}\}$*, while the labels of intermediate nodes (encoders)* $\mathcal{Q}$ *is defined as* $\mathcal{Q} = \mathcal{G} \backslash \mathcal{W}$*.*

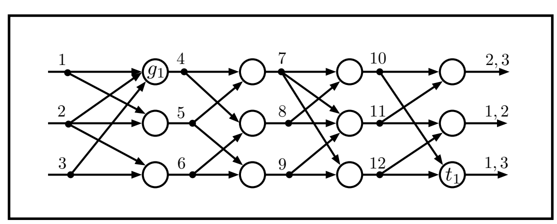

Figure 1: Network A that contains 3 sources and 9 edges

Take the network A in Fig.1 as an example to help readers understand the above defined notations. A is a 3-source 9-edge network with source messages labelled as $\mathcal{S} = \{1, 2, 3\}$, edge messages labelled as $\mathcal{E} = \{4, ..., 12\}$, sink demands are labelled as $\mathcal{T} = \{\{2, 3\}, \{1, 2\}, \{1, 3\}\}$. An intermediate node

$g_1 \in \mathcal{Q}$ is denoted as $g_1 = (\{1,2,3\},\{4\})$, a sink node $t_1 \in \mathcal{W}$ is denoted as $t_1 = (\{10,12\},\{1,3\})$.

B. Valid Cut of A MSNC Network

The **key** definition introduced in this paper is so called the *valid cut*. Unlike the traditional definition of cut [16], [17] that partitions the nodes of a graph into two subsets such that all source nodes are in one and all sink nodes are in the other, the valid cut introduced in this paper is defined as follows,

Definition 2. *(Valid cut) Let $E = \mathcal{S} \cup \mathcal{E}$ be the collection of all sources and edges of a MSNC network, a partition $E = A_1 \cup A_2 \cup A_o$ is called a valid cut if the followings are satisfied:*
1) $A_1 \neq \emptyset$, $A_2 \neq \emptyset$ and $A_o \neq \emptyset$.
2) $\{out(g), in(g)\} \subseteq 2^{E_1}$ or $\{out(g), in(g)\} \subseteq 2^{E_2}$, $\forall g \in \mathcal{Q} \cup \mathcal{W}$, where $E_1 = A_o \cup A_1$, $E_2 = A_o \cup A_2$.
3) $\mathcal{S} \subseteq E_1$ or $\mathcal{S} \subseteq E_2$, where $E_1 = A_o \cup A_1$, $E_2 = A_o \cup A_2$.

As shown in Fig.2 is a valid cut of the network in Fig.1, where $A_1 = \{1,2,3\}$, $A_2 = \{7,8,9\}$ and $A_o = \{4,5,6,10,11,12\}$.

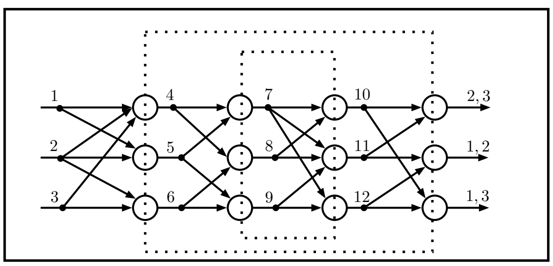

Figure 2: Network A and a partition of $E = \mathcal{S} \cup \mathcal{E}$ that forms a valid cut, where $A_1 = \{1,2,3\}$, $A_2 = \{7,8,9\}$ and $A_o = \{4,5,6,10,11,12\}$.

C. Region of Entropic Vectors Γ_E^* and the $\boxtimes$ operator

Definition 3. *(PMF, marginalization operator and region of entropic vectors Γ_E^*) For two sets A and B, we will denote by $A^B = \{\boldsymbol{x} \,|\, \boldsymbol{x} : B \to A\}$. Let the ground set $E \subseteq \mathbb{N}$ index a set of $|E|$ finite discrete random variables (RVs), and let*

$$\Delta_E = \left\{ p_E \middle| \begin{array}{l} \mathcal{X} \subseteq \mathbb{N}, p_E : \mathcal{X}^E \to [0,1] \\ \sum_{\boldsymbol{x}_E \in \mathcal{X}^E} p_E(\boldsymbol{x}_E) = 1 \end{array} \right\} \quad (1)$$

be a set of valid joint probability mass functions (PMFs) for these RVs paired with their ranges. For any such joint PMF $p_E \in \Delta_E$ and any subset $A \subset E$ define the marginalization operator $M : \Delta_E \times 2^E \to \bigcup_{A \subseteq E} \Delta_A$ with

$$M(p_E, A) = (p_A, \mathcal{X}^A), \quad (2)$$

with,

$$p_A(\boldsymbol{x}_A) = \sum_{\substack{\boldsymbol{x}_E \in \mathcal{X}^E \text{ s.t.} \\ \boldsymbol{x}_E(e) = \boldsymbol{x}_A(e)\ \forall e \in A}} p_E(\boldsymbol{x}_E) \quad (3)$$

define the function $\hat{\boldsymbol{h}}_E : \Delta_E \to \mathbb{R}^{2^E}$ such that for any $A \subseteq E$,

$$\hat{\boldsymbol{h}}_E(p_E) : A \mapsto -\sum_{\boldsymbol{x}_A \in \mathcal{X}^A} p_A(\boldsymbol{x}_A) \log_2 (p_A(\boldsymbol{x}_A)) \quad (4)$$

The entropy region for the ground set E will be the image of the set of all joint PMFs Δ_E under this this map $\hat{\boldsymbol{h}}_E$

$$\Gamma_E^* = \hat{\boldsymbol{h}}_E(\Delta_E) \quad (5)$$

Denoted as $\boldsymbol{h} = [\boldsymbol{h}(A) | A \subseteq E] \in \Gamma_E^*$ an entropic vector in Γ_E^*.

Definition 4. *(The $\boxtimes$ operator) For two subsets $\mathscr{P}_1 \subseteq \mathbb{R}^{A_1}$ and $\mathscr{P}_2 \subseteq \mathbb{R}^{A_2}$ of Euclidean space, the pasting operator $\boxtimes$ is defined as,*

$$\mathscr{P}_1 \boxtimes \mathscr{P}_2 = \left\{ \boldsymbol{f} \in \mathbb{R}^{A_1 \cup A_2} \middle| \begin{array}{l} \exists\, \boldsymbol{g}_1 \in \mathscr{P}_1, \boldsymbol{g}_2 \in \mathscr{P}_2 \text{ s.t} : \\ \boldsymbol{f}(a) = \boldsymbol{g}_1(a), \forall a \in A_1 \\ \boldsymbol{f}(b) = \boldsymbol{g}_2(b), \forall b \in A_2 \end{array} \right\} \quad (6)$$

D. Rate Region of A MSNC Network

Given a MSNC network $\mathsf{A} = (\mathcal{S}, \mathcal{E}, \mathcal{T}, \mathcal{Q}, \mathcal{W})$, denoted as $Y_s, s \in \mathcal{S}$ and $U_e, e \in \mathcal{E}$ the discrete random variables representing source messages and edge messages respectively. Let $\boldsymbol{\omega} = [\omega_s | s \in \mathcal{S}]$ and $\boldsymbol{r} = [R_e | e \in \mathcal{E}]$ be the *source rates* and *edge rates* respectively, where $\omega_s \leqslant H(Y_s)$ denotes the slack variables on source entropy and $R_e \geqslant H(U_e)$ the edge capacity. The rate region $\mathcal{R}_c(\mathsf{A})$ of a network A is defined as follow,

Definition 5. *(rate region of a network) The rate region $\mathcal{R}_c(\mathsf{A})$ of a network A is the collection of all achievable rate vectors $\boldsymbol{r}$.*

As was originally proved in [2] and later extended in [18], $\mathcal{R}_c(\mathsf{A})$ can be determined implicitly from the following equation,

$$\mathcal{R}_c(\mathsf{A}) = \text{proj}_{\boldsymbol{r},\boldsymbol{\omega}}(\overline{\text{con}(\Gamma_E^* \cap \mathcal{L}_{12})} \cap \mathcal{L}_{34}) \quad (7)$$

where $\text{con}(\mathcal{B})$ is the conic hull of $\mathcal{B}$, and $\text{proj}_{\boldsymbol{r},\boldsymbol{\omega}}(\mathcal{B})$ is the projection of $\mathcal{B}$ onto the coordinates $\boldsymbol{r}$, $\mathcal{L}_{ij}$ indicates the intersection of spaces $\mathcal{L}_i$ and $\mathcal{L}_j$ with each of which denoted as,

$$\mathcal{L}_1 = \{[\boldsymbol{r}, \boldsymbol{\omega}, \boldsymbol{h}] |\ \boldsymbol{h}(\mathcal{S}) = \sum_{s \in \mathcal{S}} \boldsymbol{h}(\{s\})\} \quad (8)$$

$$\mathcal{L}_2 = \{[\boldsymbol{r}, \boldsymbol{\omega}, \boldsymbol{h}] | \boldsymbol{h}(\text{out}(g)|\text{in}(g)) = 0,\ \forall g \in \mathcal{Q}\} \quad (9)$$

$$\mathcal{L}_3 = \{[\boldsymbol{r}, \boldsymbol{\omega}, \boldsymbol{h}] | \boldsymbol{h}(\text{out}(g)|\text{in}(g)) = 0,\ \forall g \in \mathcal{W}\} \quad (10)$$

$$\mathcal{L}_4 = \left\{ [\boldsymbol{r}, \boldsymbol{\omega}, \boldsymbol{h}] \middle| \begin{array}{l} \boldsymbol{h}(\{s\}) \geqslant \omega_s,\ \forall s \in \mathcal{S} \\ \boldsymbol{h}(\{e\}) \leqslant R_e,\ \forall e \in \mathcal{E} \end{array} \right\} \quad (11)$$

This paper studies a closely related outer bound $\mathcal{R}_c^o(\mathsf{A})$ of $\mathcal{R}_c(\mathsf{A})$,

$$\mathcal{R}_c(\mathsf{A}) \subseteq \mathcal{R}_c^o(\mathsf{A}) = \text{proj}_{\boldsymbol{r},\boldsymbol{\omega}}(\overline{\Gamma_E^*} \cap \mathcal{L}_{\mathsf{A}}) \quad (12)$$

where $\mathcal{L}_{\mathsf{A}} = \mathcal{L}_{12} \cap \mathcal{L}_{34}$ denoting the collection of all network constraints of A. . We can infer from corollary 1 and footnote 1 in [19] that $\mathcal{R}_c(\mathsf{A})$ is strictly equal to $\mathcal{R}_c^o(\mathsf{A})$ if $\overline{\Gamma_E^*} = \text{con}(\Gamma_E^*)$.

As explained in [20], an outer bound of $\mathcal{R}_c(\mathsf{A})$ can be obtained by replacing $\overline{\Gamma_E^*}$ in (7) with its outer bound. For example, let Γ_E be the Shannon outer bound of $\overline{\Gamma_E^*}$, then the corresponding outer bound $\mathcal{R}(\mathsf{A})$ to the rate region is given by

$$\mathcal{R}_c(\mathsf{A}) \subseteq \mathcal{R}_c^o(\mathsf{A}) \subseteq \mathcal{R}(\mathsf{A}) = \text{proj}_{\boldsymbol{r},\boldsymbol{\omega}}(\Gamma_E \cap \mathcal{L}_{\mathsf{A}}) \tag{13}$$

The problem with calculating $\mathcal{R}(\mathsf{A})$ is that is involves projecting a very high dimensional polyhedral cone, which makes it computationally infeasible beyond networks with 10s of sources and edges.

E. Main Results

There are three main results in this paper. Firstly a **new theorem** that implicitly determine the rate region is proposed in section III and is proved in detail in section IV. Then a systematic way of applying the theorem to calculate explicitly the outer bound to a rate region is proposed in section V. For some MSNC networks (for example the network in Fig.1), our method involves the projection of a **exponentially simpler** polyhedra than the traditional method of calculating $\mathcal{R}(\mathsf{A})$ in (13). Finally in section V-C, we show that there are **infinitely many** networks where our new rate region calculation method can be applied upon.

III. Definitions and Problem Formulation

Consider a MSNC network $\mathsf{A} = (\mathcal{S}, \mathcal{E}, \mathcal{T}, \mathcal{Q}, \mathcal{W})$ with $E = \mathcal{S} \cup \mathcal{E}$ denoted as the set of source and edge labels. Define the following,

Definition 6. *Let $E = E_1 \cup E_2$ and $E_1 \cap E_2 \neq \varnothing$ and denoted as $2^{E_1} \cup 2^{E_2}$ the union of two power sets, define.*
1)

$$\mathcal{G}^e_{E_1,E_2} = \overline{\Gamma^*_{E_1}} \bigcap\!\!\!\!\times \overline{\Gamma^*_{E_2}} \tag{14}$$

so that

$$\mathcal{G}^e_{E_1,E_2} = \left\{ \boldsymbol{h} \in \mathbb{R}^{2^{E_1} \cup 2^{E_2}} \;\middle|\; \begin{array}{l} \exists \boldsymbol{h}_1 \in \overline{\Gamma^*_{E_1}},\ \boldsymbol{h}_2 \in \overline{\Gamma^*_{E_2}}: \\ \boldsymbol{h}(A) = \boldsymbol{h}_1(A),\ \forall A \subseteq E_1, \\ \boldsymbol{h}(A) = \boldsymbol{h}_2(A),\ \forall A \subseteq E_2 \end{array} \right\} \tag{15}$$

2) Define the following region

$$\mathcal{R}'_{one}(\mathsf{A}) = \textit{proj}_{\boldsymbol{r},\boldsymbol{\omega}}(\mathcal{G}^e_{E_1,E_2} \cap \mathcal{L}_{\mathsf{A}}) \tag{16}$$

where,

$$\mathcal{L}_{\mathsf{A}} = \mathcal{L}_1 \cap \mathcal{L}_2 \cap \mathcal{L}_3 \cap \mathcal{L}_4$$

Condition 1. *Let $E = A_1 \cup A_2 \cup A_o$ be a set partition, and $E_1 = A_1 \cup A_o$ $E_2 = A_2 \cup A_o$. For any two entropic vectors $\boldsymbol{h}_1 \in \Gamma^*_{E_1}$ and $\boldsymbol{h}_2 \in \Gamma^*_{E_2}$, if*

$$\boldsymbol{h}_1(A) = \boldsymbol{h}_2(A),\ \forall A \subseteq E_1 \cap E_2$$

then we can find a joint PMF p_E over the ground set E with $\boldsymbol{h} = \hat{\boldsymbol{h}}_E(p_E)$ such that,

$$\begin{cases} \boldsymbol{h}_1(A) = \boldsymbol{h}(A),\ \forall A \subseteq E_1 \\ \boldsymbol{h}_2(B) = \boldsymbol{h}(B),\ \forall B \subseteq E_2 \end{cases}$$

Theorem 1. *For a network A with the set of source and edge variables denoted as E, if there exists a non-empty partition $E = A_1 \cup A_2 \cup A_o$, with $E_1 = A_1 \cup A_o$ and $E_2 = A_2 \cup A_o$ such that: 1) A_1, A_2 and A_o forms a valid cut and 2) **Condition** 1 are satisfied. Then we must have,*

$$\mathcal{R}_c^o(\mathsf{A}) = \mathcal{R}'_{one}(\mathsf{A}) \tag{17}$$

This theorem will be proved in Section IV. Notice that although in this paper it served as a premise of **Theorem** 1, **Condition** 1 itself is of important theoretical value. Readers who are interested may refer to (the other ISIT 2021 draft of ours) for more details where we proved **Condition** 1 is true for some special cases.

IV. Proof of Theorem 1

Lemma 1. *For a network A with the set of source and edge variables labeled as E and a non-empty partition $E = A_1 \cup A_2 \cup A_o$, let $E_1 = A_1 \cup A_o$ and $E_2 = A_2 \cup A_o$ we have*

$$\textit{proj}_{2^{E_1} \cup 2^{E_2}}(\overline{\Gamma^*_E}) \subseteq \mathcal{G}^e_{E_1,E_2} \tag{18}$$

Proof. (Proof of **Lemma** 1) Pick an arbitrary vector $\overline{\boldsymbol{v}} \in \text{proj}_{2^{E_1} \cup 2^{E_2}}(\overline{\Gamma^*_E})$, then we can find a vector $\overline{\boldsymbol{h}} \in \overline{\Gamma^*_E}$ such that $\overline{\boldsymbol{v}} = \text{proj}_{2^{E_1} \cup 2^{E_2}}(\overline{\boldsymbol{h}})$, which means,

$$\overline{\boldsymbol{v}}(A) = \overline{\boldsymbol{h}}(A), \forall A \subseteq E_1 \text{ or } A \subseteq E_2 \tag{19}$$

By definition of $\overline{\Gamma^*_E}$, $\overline{\boldsymbol{h}}$ is a limit point of Γ^*_E, so we can find a sequence of entropic vectors $(\boldsymbol{h}_n)_{n \in \mathbb{N}}$ in Γ^*_E that $\lim_{n\to\infty} \boldsymbol{h}_n = \overline{\boldsymbol{h}}$. Create from $(\boldsymbol{h}_n)_{n\in\mathbb{N}}$ and $\overline{\boldsymbol{h}}$ the following terms,

$$\overline{\boldsymbol{k}} = [\overline{k}(A) | \overline{k}(A) = \overline{\boldsymbol{h}}(A), \forall A \subseteq E_1] \tag{20}$$

$$\overline{\boldsymbol{l}} = [\overline{l}(A) | \overline{l}(A) = \overline{\boldsymbol{h}}(A), \forall A \subseteq E_2] \tag{21}$$

$$\boldsymbol{k}_n = [k_n(A) | k_n(A) = \boldsymbol{h}_n(A), \forall A \subseteq E_1],\ \forall n \in \mathbb{N} \tag{22}$$

$$\boldsymbol{l}_n = [l_n(A) | l_n(A) = \boldsymbol{h}_n(A), \forall A \subseteq E_2],\ \forall n \in \mathbb{N} \tag{23}$$

It is not hard to see that by the above construction, $(\boldsymbol{k}_n)_{n\in\mathbb{N}}$ and $(\boldsymbol{l}_n)_{n\in\mathbb{N}}$ are sequence of entropic vectors and that,

$$\lim_{n\to\infty} \boldsymbol{k}_n = \overline{\boldsymbol{k}} \tag{24}$$

$$\lim_{n\to\infty} \boldsymbol{l}_n = \overline{\boldsymbol{l}} \tag{25}$$

which implies that $\overline{\boldsymbol{k}} \in \overline{\Gamma^*_{E_1}}$ and $\overline{\boldsymbol{l}} \in \overline{\Gamma^*_{E_2}}$. Now combining (19), (20) and (21) we have,

$$\overline{\boldsymbol{v}}(A) = \overline{\boldsymbol{k}}(A), \forall A \subseteq E_1 \tag{26}$$

$$\overline{\boldsymbol{v}}(A) = \overline{\boldsymbol{l}}(A), \forall A \subseteq E_2 \tag{27}$$

which then implies that

$$\overline{\boldsymbol{v}} \in \overline{\Gamma^*_{E_1}} \bigcap\!\!\!\!\times \overline{\Gamma^*_{E_2}} = \mathcal{G}^e_{E_1,E_2}$$

So we have $\boldsymbol{v} \in \text{proj}_{\boldsymbol{h}_p}(\bar{\Gamma}^*_E)$ implies that $\boldsymbol{v} \in \mathcal{G}^e_{E_1,E_2}$, which ends the proof. ■

Lemma 2. *For a network A with the set of source and edge variables labeled as E and a non-empty partition $E = A_1 \cup A_2 \cup A_o$ such that **condition** 1 is satisfied, let $E_1 = A_1 \cup A_o$ and $E_2 = A_2 \cup A_o$ we have*

$$\textit{proj}_{2^{E_1} \cup 2^{E_2}}(\Gamma^*_E) \supseteq \Gamma^*_{E_1} \bigcap\!\!\!\!\times \Gamma^*_{E_2} \tag{28}$$

Proof. (Proof of **Lemma** 2) Pick an arbitrary vector $\boldsymbol{v} \in \Gamma^*_{E_1} \bigcap\!\!\!\!\times \Gamma^*_{E_2}$, then we can find a pair of vectors $\boldsymbol{h}_1 \in \Gamma^*_{E_1}$ and $\boldsymbol{h}_2 \in \Gamma^*_{E_2}$ such that

$$\boldsymbol{v}(A) = \boldsymbol{h}_1(A), \forall A \subseteq E_1 \tag{29}$$

$$\boldsymbol{v}(A) = \boldsymbol{h}_2(A), \forall A \subseteq E_2 \tag{30}$$

which implies that,

$$\boldsymbol{h}_1(A) = \boldsymbol{h}_2(A), \forall A \subseteq E_1 \cap E_2 \tag{31}$$

By **condition** 1 we can find a joint PMF p_E over the ground set E with $\boldsymbol{h} = \hat{\boldsymbol{h}}_E(p_E)$such that,

$$\begin{cases} \boldsymbol{h}_1(A) = \boldsymbol{h}(A), \ \forall A \subseteq E_1 \\ \boldsymbol{h}_2(B) = \boldsymbol{h}(B), \ \forall B \subseteq E_2 \end{cases} \tag{32}$$

Now, combining (29), (30) with (32) we have

$$\boldsymbol{v} = \text{proj}_{2^{E_1} \cup 2^{E_2}}(\boldsymbol{h}) \in \text{proj}_{2^{E_1} \cup 2^{E_2}}(\Gamma_E^*) \tag{33}$$

So we have $\boldsymbol{v} \in \Gamma_{E_1}^* \boxtimes \Gamma_{E_2}^*$ implies that $\boldsymbol{v} \in \text{proj}_{2^{E_1} \cup 2^{E_2}}(\Gamma_E^*)$, which ends the proof. ∎

Lemma 3. *For a network* A *with the set of source and edge variables labeled as E and a non-empty partition $E = A_1 \cup A_2 \cup A_o$ such that* ***condition*** *1 is satisfied, let $E_1 = A_1 \cup A_o$ and $E_2 = A_2 \cup A_o$ we have*

$$\textit{proj}_{2^{E_1} \cup 2^{E_2}}(\overline{\Gamma_E^*}) \supseteq \mathcal{G}^e_{E_1,E_2} \tag{34}$$

Proof. (Proof of **Lemma** 3) Apply **Lemma** 2 and take closure on both sides of (28), which gives

$$\overline{\text{proj}_{2^{E_1} \cup 2^{E_2}}(\Gamma_E^*)} \supseteq \overline{\Gamma_{E_1}^* \boxtimes \Gamma_{E_2}^*} \tag{35}$$

Notice that,

$$\overline{\Gamma_{E_1}^* \boxtimes \Gamma_{E_2}^*} \supseteq \mathcal{G}^e_{E_1,E_2} \tag{36}$$

$$\text{proj}_{2^{E_1} \cup 2^{E_2}}(\overline{\Gamma_E^*}) \supseteq \overline{\text{proj}_{2^{E_1} \cup 2^{E_2}}(\Gamma_E^*)} \tag{37}$$

So we have,

$$\begin{aligned} \text{proj}_{2^{E_1} \cup 2^{E_2}}(\overline{\Gamma_E^*}) &\supseteq \overline{\text{proj}_{2^{E_1} \cup 2^{E_2}}(\Gamma_E^*)} \\ &\supseteq \overline{\Gamma_{E_1}^* \boxtimes \Gamma_{E_2}^*} \\ &\supseteq \mathcal{G}^e_{E_1,E_2} \end{aligned}$$ ∎

Proof. (Proof of **Theorem** 1) Let $E_1 = A_1 \cup A_o$, $E_2 = A_2 \cup A_o$

$$\begin{aligned} \mathcal{R}_c^o(\mathsf{A}) &= \text{proj}_{\boldsymbol{r},\boldsymbol{\omega}}(\overline{\Gamma_E^*} \cap \mathcal{L}_\mathsf{A}) \\ &\stackrel{T_1}{=} \text{proj}_{\boldsymbol{r},\boldsymbol{\omega}}(\text{proj}_{2^{E_1} \cup 2^{E_2}}(\overline{\Gamma_E^*} \cap \mathcal{L}_\mathsf{A})) \\ &\stackrel{T_2}{=} \text{proj}_{\boldsymbol{r},\boldsymbol{\omega}}(\text{proj}_{2^{E_1} \cup 2^{E_2}}(\overline{\Gamma_E^*}) \cap \mathcal{L}_\mathsf{A}) \\ &\stackrel{T_3}{=} \text{proj}_{\boldsymbol{\omega},\boldsymbol{r}}(\mathcal{G}^e_{E_1,E_2} \cap \mathcal{L}_\mathsf{A}) \\ &= \mathcal{R}'_{one}(\mathsf{A}) \end{aligned} \tag{38}$$

where T_1 holds because projections can be done gradually by projecting out first the dimensions not in $2^{E_1} \cup 2^{E_2}$ and then the dimensions not in $\boldsymbol{r}$; T_2 holds because it can be implied from the definition of valid cut that each inequality or equality that defines $\mathcal{L}_\mathsf{A}$ will act only on dimensions in $2^{E_1} \cup 2^{E_2}$; T_3 holds because combining **Lemma** 1 and **Lemma** 3 we have

$$\text{proj}_{2^{E_1} \cup 2^{E_2}}(\overline{\Gamma_E^*}) = \mathcal{G}^e_{E_1,E_2}$$ ∎

V. Application of Theorem 1

The following 4 questions will be addressed in this section: 1) How to apply **Theorem** 1 to rate region calculation? 2) What is the advantage of applying **Theorem** 1 compared with the traditional way of calculating $\mathcal{R}(\mathsf{A})$ to determine the rate region? 3) How many networks are there such that a valid cut exists? 4) What if **Condition** 1 is really hard to verify and even to be false in some cases?

A. How to Apply *Theorem 1*

In general, $\overline{\mathcal{R}'_{one}(\mathsf{A})}$ can not be calculated directly due to the fact that $\overline{\Gamma_{E_1}^*}$ (or $\overline{\Gamma_{E_2}^*}$) is non-polyhedral when $|E_1| \geqslant 4$ (or $|E_2| \geqslant 4$). So to apply **Theorem** 1, we followed the same trick as (13) to substitute Shannon outer bound of $\overline{\Gamma_{E_1}^*}$ (and $\overline{\Gamma_{E_2}^*}$) into (16) and obtained an outer bound $\mathcal{R}_{one}(\mathsf{A})$ of $\mathcal{R}'_{one}(\mathsf{A})$ as,

$$\begin{aligned} \mathcal{R}'_{one}(\mathsf{A}) &\subseteq \mathcal{R}_{one}(\mathsf{A}) \\ &= \text{proj}_{\boldsymbol{r},\boldsymbol{\omega}}\left(\text{proj}_{2^{E_1} \cup 2^{E_2}}\left(\Gamma_{E_1} \boxtimes \Gamma_{E_2}\right) \cap \mathcal{L}_\mathsf{A}\right) \end{aligned} \tag{39}$$

After calculating $\mathcal{R}_{one}(\mathsf{A})$, we then test to see if $\mathcal{R}_{one}(\mathsf{A})$ is the true rate region (achievability proof). Notice that for the achievability proof it is suffice to show that all the extreme rays of $\mathcal{R}_{one}(\mathsf{A})$ are achievable, the rest of non-extremal rays can be achieved by time-sharing between the codes that achieve the extreme rays.

B. The Advantage of Applying *Theorem 1*

The basic idea is that computing $\mathcal{R}_{one}(\mathsf{A})$ is in general faster than computing $\mathcal{R}(\mathsf{A})$, which is due to the fact that $\Gamma_{E_1} \boxtimes \Gamma_{E_2}$ may have **exponentially less** dimensions and inequalities than Γ_E. For example for the network A shown in Fig.2, $\Gamma_{E_1} \boxtimes \Gamma_{E_2}$ is a $2^9 + 2^9 - 2^6 - 1 = 959$ dimensional polyhedra cone that is characterized by $2(9 + \binom{9}{2}2^7) = 9234$ inequalities while Γ_E is $2^{12} - 1 = 4095$ dimensional and contains $12 + \binom{12}{2}2^{10} = 67586$ inequalities. To directly compare the difference in terms of computational complexity, we tested our ideas on ITCP [21], a rate region computation software, and found that for the network A in Fig.2, while $\mathcal{R}_{one}(\mathsf{A})$ turned out to be **equal to** $\mathcal{R}(\mathsf{A})$, it took less than 10 minutes to calculate $\mathcal{R}_{one}(\mathsf{A})$ but more than 24 hours to calculate $\mathcal{R}(\mathsf{A})$.

C. Infinitely Many Networks Where a Valid Cut Exists

When moving from theory to application, one of the natural questions to ask is how many instances are there satisfying the premise of a theorem. To show that the power of **Theorem** 1 is not limited by the requirement of valid cuts, in this subsection we offer an inductive argument to show that there are infinitely many networks where valid cut exists.

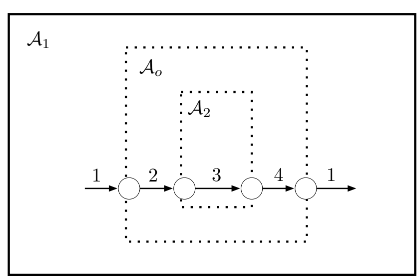

Figure 3: Network A_1 and a valid cut with $A_1 = \{1\}$, $A_2 = \{3\}$ and $A_o = \{2, 4\}$

Take the network A_1 shown in fig.3, which is arguably the simplest network such that a valid cut exists, as the base case in our inductive argument. Two dotted rectangle boundaries are introduced as helper lines to separate the figure into 3 disjoint areas $\mathcal{A}_1$, $\mathcal{A}_2$ and $\mathcal{A}_o$. Collecting the network variables within each area into a set, then we have three disjoint sets $A_1 = \{1\}$, $A_2 = \{3\}$ and $A_o = \{2, 4\}$ obviously forms a valid cut of the network.

The key of our inductive argument is to define proper network extensions such that any network extended from the base network A_1 is also a premise-satisfied network. To do this let's first define the notion of proper sources, proper edges and proper nodes.

Definition 7. *A source s is proper source if $s \in \mathcal{A}_1$. An edge e is proper if e satisfies EXACTLY one of the followings: $e \in \mathcal{A}_1$ or $e \in \mathcal{A}_2$ or $e \in \mathcal{A}_o$. A node g is a proper node if its outgoing and incoming messages satisfy AT LEAST one of the followings: $(in(g) \cup out(g)) \cap \mathcal{A}_2 = \emptyset$ or $(in(g) \cup out(g)) \cap \mathcal{A}_1 = \emptyset$.*

Then we argue that the premise-satisfied extension can be achieved by adding proper sources, proper edges and nodes. As shown in fig.4 is such an example where the network A_2 is extended from A_1 by adding one proper source 5, proper edges $6, 7$ and proper nodes g_1, g_2 and t_1. Notice that now the set A_i associated with area $\mathcal{A}_i$ becomes $A_1 = \{1, 5\}$, $A_2 = \{3\}$ and $A_o = \{2, 4, 6, 7\}$ and one can easily verify that they also form a valid cut.

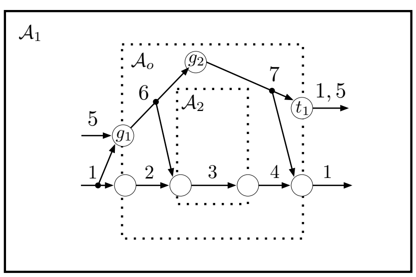

Figure 4: Network A_2 and a valid cut with $A_1 = \{1, 5\}$, $A_2 = \{3\}$ and $A_o = \{2, 4, 6, 7\}$

Now, to finish our inductive argument, let's assume that A_i is a premise-satisfied network extended from A_1 by adding proper sources, nodes and edges, and let $\mathcal{A}_1$, $\mathcal{A}_2$ and $\mathcal{A}_o$ be those three distinct areas associated with $A_{i,1}$, $A_{i,2}$ and $A_{i,o}$ respectively. Let $\mathsf{A}_j = (\mathcal{S}_j, \mathcal{E}_j, \mathcal{T}_j, \mathcal{Q}_j, \mathcal{W}_j)$ be the network extended from A_i by adding finite number of proper sources, proper edges, and denoted as $A_{j,1}$, $A_{j,2}$ and $A_{j,o}$ the set of network variables associated with $\mathcal{A}_1$, $\mathcal{A}_2$ and $\mathcal{A}_o$ respectively after extension and let $E_{j,1} = A_{j,1} \cup A_{j,o}$, $E_{j,2} = A_{j,2} \cup A_{j,o}$. Firstly, It is not hard to see that $A_{j,1}$, $A_{j,2}$ and $A_{j,o}$ forms a partition and each of them is not an empty set. What's left is to show that:

B.1 $\{\text{out}(g), \text{in}(g)\} \subseteq 2^{E_{j,1}}$ or $\{\text{out}(g), \text{in}(g)\} \subseteq 2^{E_{j,2}}$, $\forall g \in \mathcal{Q}_j \cup \mathcal{W}_j$

B.2 $\mathcal{S}_j \subseteq E_{j,1}$ or $\mathcal{S}_j \subseteq E_{j,2}$

To show **B.1**, consider that any node $g \in \mathcal{Q}_j \cup \mathcal{W}_j$ is a proper node so we have $(\text{in}(g) \cup \text{out}(g)) \cap \mathcal{A}_2 = \emptyset$ or $(\text{in}(g) \cup \text{out}(g)) \cap \mathcal{A}_1 = \emptyset$, which implies that $\{\text{out}(g), \text{in}(g)\} \subseteq 2^{E_{j,1}}$ or $\{\text{out}(g), \text{in}(g)\} \subseteq 2^{E_{j,2}}$. To show **B.2** consider that any proper source $s \in \mathcal{S}_j$ is a proper source, so we have $s \in \mathcal{A}_1$, which implies that $\mathcal{S}_j \subseteq A_{j,1} \subseteq E_{j,1}$.

A generate network of such is shown in Fig.5 to help readers better visualize the above proved extension process.

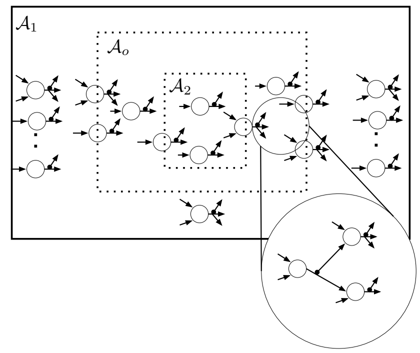

Figure 5: A general network extended from A_1 that a valid cut exists

D. The Uncertainty of *Condition 1*

In a companion manuscript [22] we prove an infinite series of situations in which **Condition** 1 holds. More broadly, even when condition 1 is not satisfied for a valid cut, an outer bound is still obtained. Indeed if we assume only the existence of a valid cut but not **Condition** 1, then based on **Lemma** 1, T_3 in (38) changes from "is equal to" to "is a subset of", which gives,

$$\mathcal{R}_c^o(\mathsf{A}) \subseteq \mathcal{R}'_{one}(\mathsf{A}) \tag{40}$$

Now combining (39) with (40) we have,

$$\mathcal{R}_c^o(\mathsf{A}) \subseteq \mathcal{R}'_{one}(\mathsf{A}) \subseteq \mathcal{R}_{one}(\mathsf{A}) \tag{41}$$

So one can see that no matter **Condition** 1 is assumed or not, $\mathcal{R}_{one}(\mathsf{A})$ is an outer bound of $\mathcal{R}_c^o(\mathsf{A})$, the systematic way of first computing $\mathcal{R}_{one}(\mathsf{A})$ and then determining if it the true rate region is still valid.

VI. Conclusion

In this paper, a new theorem that implicitly determines the rate region of a network is proved and a corresponding systematic way of applying the theorem to calculate explicitly the outer bounds to a rate region is proposed. Compared with the traditional method, the major advantage of our method is that for some networks it is suffice to project an exponentially simpler polyhedra to obtain their true rate region. An inductive construction shows that there are infinitely many networks where the proposed method can be applied upon.

References

[1] J. Cannons and K. Zeger, "Network coding capacity with a constrained number of coding nodes," *IEEE Transactions on Information Theory*, vol. 54, pp. 1287–1291, March 2008.

[2] X. Yan, R. W. Yeung, and Z. Zhang, "An implicit characterization of the achievable rate region for acyclic multisource multisink network coding," *IEEE Transactions on Information Theory*, vol. 58, pp. 5625–5639, Sept 2012.

[3] T. H. Chan and A. Grant, "Network coding capacity regions via entropy functions," *IEEE Transactions on Information Theory*, vol. 60, no. 9, pp. 5347 – 5374, 2014. Achievable rate region;Acyclic networks;Capacity regions;Decoding errors;Entropy function;General networks;Linearity constraints;Secure Network Coding;.

[4] R. Dougherty, C. Freiling, and K. Zeger, “Achievable rate regions for network coding,” *IEEE Transactions on Information Theory*, vol. 61, pp. 2488–2509, May 2015.
[5] M. Effros, S. El Rouayheb, and M. Langberg, “An equivalence between network coding and index coding,” *IEEE Transactions on Information Theory*, vol. 61, no. 5, pp. 2478–2487, 2015.
[6] H. Sun and S. A. Jafar, “Index coding capacity: How far can one go with only shannon inequalities?,” *IEEE Transactions on Information Theory*, vol. 61, no. 6, pp. 3041–3055, 2015.
[7] C. Tian, “Symmetry, outer bounds, and code constructions: A computer-aided investigation on the fundamental limits of caching,” *Entropy*, vol. 20, no. 8, p. 603, 2018.
[8] C. Tian, “Rate region of the (4, 3, 3) exact-repair regenerating codes,” in *2013 IEEE International Symposium on Information Theory*, pp. 1426–1430, 2013.
[9] C. Tian, “Characterizing the rate region of the (4,3,3) exact-repair regenerating codes,” *IEEE Journal on Selected Areas in Communications*, vol. 32, pp. 967–975, May 2014.
[10] A. G. Dimakis, P. B. Godfrey, Y. Wu, M. J. Wainwright, and K. Ramchandran, “Network coding for distributed storage systems,” *IEEE transactions on information theory*, vol. 56, no. 9, pp. 4539–4551, 2010.
[11] M. Asteris, D. Papailiopoulos, A. G. Dimakis, R. Vadali, S. Chen, and D. Borthakur, “Xoring elephants: Novel erasure codes for big data,” *Proceedings of the VLDB Endowment*, vol. 6, no. 5, 2013.
[12] S. Weber, C. Li, and J. M. Walsh, “Rate region for a class of delay mitigating codes and p2p networks,” in *2012 46th Annual Conference on Information Sciences and Systems (CISS)*, pp. 1–6, 2012.
[13] J. M. Walsh, S. Weber, and C. w. Maina, “Optimal rate delay tradeoffs for multipath routed and network coded networks,” in *2008 IEEE International Symposium on Information Theory*, pp. 682–686, 2008.
[14] J. Maclaren Walsh, S. Weber, J. Cavalcante De Oliveira, A. Eryilmaz, and M. Medard, “Guest editorial trading rate for delay at the application and transport layers,” *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 5, pp. 913–915, 2011.
[15] Y. Liu and J. M. Walsh, “On the complexity of computing the shannon outer bound to a network coding capacity region,” in *2019 IEEE International Symposium on Information Theory (ISIT)*, pp. 206–210, 2019.
[16] G. Dantzig and D. R. Fulkerson, “On the max flow min cut theorem of networks,” *Linear inequalities and related systems*, vol. 38, pp. 225–231, 2003.
[17] A. El Gamal and Y.-H. Kim, *Network information theory*. Cambridge university press, 2011.
[18] C. Li, S. Weber, and J. M. Walsh, “Network combination operations preserving the sufficiency of linear network codes,” in *2015 IEEE Information Theory Workshop - Fall (ITW)*, pp. 44–48, Oct 2015.
[19] C. Li, S. Weber, and J. M. Walsh, “On multi-source networks: Enumeration, rate region computation, and hierarchy,” *IEEE Transactions on Information Theory*, vol. 63, pp. 7283–7303, Nov 2017.
[20] C. Li, S. Weber, and J. M. Walsh, “Multilevel diversity coding systems: Rate regions, codes, computation, amp; forbidden minors,” *IEEE Transactions on Information Theory*, vol. 63, pp. 230–251, Jan 2017.
[21] J. Apte and J. M. Walsh, “Explicit polyhedral bounds on network coding rate regions via entropy function region: Algorithms, symmetry, and computation,” *CoRR*, vol. abs/1607.06833, 2016.
[22] Y. Liu and J. M. Walsh, “Linear complexity entropy regions,” in *submitted to 2021 IEEE International Symposium on Information Theory (ISIT)*, 2021.