The Capacity Region for Multi-source Multi-sink Network Coding

Xijin Yan
Dept. of Electrical Eng. - Systems
University of Southern California
Los Angeles, CA, U.S.A.
xyan@usc.edu

Raymond W. Yeung
Dept. of Information Eng.
The Chinese University of Hong Kong
N.T., Hong Kong
whyeung@ie.cuhk.edu.hk

Zhen Zhang
Dept. of Electrical Eng. - Systems
University of Southern California
Los Angeles, CA, U.S.A.
zzhang@commsci1.usc.edu

***Abstract*— The capacity problem for general acyclic multi-source multi-sink networks with arbitrary transmission requirements has been studied in [1], [2] and [3]. Specifically, inner and outer bounds of the capacity region were derived respectively in terms of Γ_n^* and $\bar{\Gamma}_n^*$, the fundamental regions of the entropy function. In this paper, we show that by carefully bounding the constrained regions in the entropy space, we obtain the exact characterization of the capacity region, thus closing the existing gap between the above inner and outer bounds.**

I. Introduction

Consider a multi-source multi-sink network in which more than one mutually independent information sources are generated at possibly different nodes and each of the information sources is multicast to a specific set of sink nodes. We assume the network is acyclic and the channels are free of error.

Unlike the single-source multicast network coding problem, where the capacity region has an explicit Max-flow Min-cut representation [4], the counterpart problem for multi-source multi-sink network coding with arbitrary transmission requirements is considerably more complex. Except for a few explicit outer bounds that have been discovered recently in [2], [5], [6] and [7], the tightest theoretical characterization which has been obtained so far makes use of the tools developed in the theory of information inequalities [1]. Specifically, an inner bound and an outer bound were derived in terms of Γ_n^* and $\bar{\Gamma}_n^*$ respectively, which are fundamental regions of the entropy function.

In this paper, we determine the exact capacity region for general acyclic multi-source multi-sink networks using an entropy function characterization. In particular, we show that by carefully bounding the constrained regions in the entropy space, we obtain the exact characterization of the capacity region, thus closing the existing gap between the above inner and outer bounds.

The rest of the paper is organized as follows. In section II, we present a formal problem formulation and introduce some preliminaries on strongly typical sequences and entropy functions. In section III, we present the main result, i.e., an exact characterization of the capacity region. Proofs of the main result are given in section IV and final conclusions are drawn in section V.

II. Preliminaries

A. Network Model

Let $G = (V, E)$ denote an acyclic multi-source multi-sink communication network, where V and E are the set of all nodes and the set of all channels. We assume each channel $e \in E$ is error-free with a positive capacity constraint R_e and all nodes $i \in V$ are ordered in a way such that if there exists a channel from a node i to a node j, then the node i precedes the node j. We further define $In(i) = \{(j,i) \in E : j \in V\}$ to be the set of channels directed into node i and $Out(i) = \{(i,j) \in E : j \in V\}$ to be the set of channels directed from i. Let $S \subset V$ be the set of all source nodes and $T \subset V$ be the set of all sink nodes. Without loss of generality, we assume G has the structure such that each source node has no input channels and each sink node has no outgoing channels.

We assume all sources are uniformly distributed and mutually independent with finite alphabet $\mathcal{X}_s = \{1, 2, \cdots, \lceil 2^{n\tau_s} \rceil\}$ for all $s \in S$, where τ_s is the source information rate at $s \in S$. A sink node $t \in T$ requires the data from a set of sources $\beta(t) \subset S$ to be decoded. In the case when $\beta(t) = S$ for all $t \in T$, the given network can be simply treated as a single source multicast network [4]. To allow for a general treatment of networks with arbitrary transmission requirements, we assume that $\beta(t)$ can be any subset of S for all $t \in T$.

For clarity of notation, we sometimes use a superscript on a vector (e.g. $\mathbf{x}^n$) to specify the dimension of the vector, which will be distinguished from a superscript in parentheses (e.g. $\mathbf{x}^{(k)}$) used to specify the index of a vector in a sequence (e.g. $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \cdots$). The complement of a set A is represented by A^c. For consistency, we further assume all the logarithms are in the base 2.

B. Capacity Region

Consider a block code of length n.

Definition 1: An $(n, (\eta_e, e \in E), (\tau_s, s \in S), (\Delta_t, t \in T))$ block code of length n on a given communication network is defined by

1) for all source node $s \in S$ and all channel $e \in Out(s)$, a local encoding mapping

$$k_e : \mathcal{X}_s \to \{0, 1, \cdots, \eta_e\}; \tag{1}$$

2) for all node $i \in V \setminus (S \cup T)$ and all channel $e \in Out(i)$, a local encoding mapping

$$k_e : \prod_{d \in In(i)} \{0, 1, \cdots, \eta_d\} \to \{0, 1, \cdots, \eta_e\}; \quad (2)$$

3) for all sink node $t \in T$, a decoding mapping

$$g_t : \prod_{d \in In(t)} \{0, 1, \cdots, \eta_d\} \to \prod_{s \in \beta(t)} \mathcal{X}_s; \quad (3)$$

4) for all sink node $t \in T$, a decoding error probability

$$\Delta_t = \Pr\left\{\tilde{g}_t(X_S) \neq X_{\beta(t)}\right\}, \quad (4)$$

where $\tilde{g}_t(X_S)$ is the value of g_t as a function of X_S.

Definition 2: An information rate tuple $\boldsymbol{\omega} = (\omega_s : s \in S)$, where $\boldsymbol{\omega} \geq 0$ (componentwise) is *achievable* if for any $\epsilon > 0$, there exists for sufficient large n an $(n, (\eta_e, e \in E), (\tau_s, s \in S), (\Delta_t, t \in T))$ code such that

$$n^{-1} \log \eta_e \leq R_e + \epsilon \quad (5)$$

for all $e \in E$,

$$\tau_s \geq \omega_s - \epsilon \quad (6)$$

for all $s \in S$, and

$$\Delta_t \leq \epsilon \quad (7)$$

for all $t \in T$.

Definition 3: The *capacity region* denoted by $\mathcal{R}$ is the set of all achievable information rate tuple $\boldsymbol{\omega}$.

C. Strongly Typical Sequences

Consider an information source $\{X_k, k \geq 1\}$ where X_k are i.i.d. with probability distribution $p(x)$. Let X denote the generic random variable, where $H(X) < \infty$ and $\mathcal{S}_X$ be the support of X.

Definition 4: The *strong δ-typical set* $T^n_{[X]\delta}$ with respect to $p(x)$ is the set of sequences $\boldsymbol{x}^n = (x_1, x_2, \cdots, x_n) \in \mathcal{X}^n$ such that $N(x; \boldsymbol{x}^n) = 0$ for $x \notin \mathcal{S}_X$, and

$$\sum_x \left| \frac{1}{n} N(x; \boldsymbol{x}^n) - p(x) \right| \leq \delta,$$

where $N(x; \boldsymbol{x}^n)$ is the number of occurrences of x in $\boldsymbol{x}^n$, and δ is an arbitrarily small positive real number.

Lemma 1: (*Strong AEP*) Let η be a small positive quantity such that $\eta \to 0$ as $\delta \to 0$.

1) If $\boldsymbol{x}^n \in T^n_{[X]\delta}$, then

$$2^{-n(H(X)+\eta)} \leq p(\boldsymbol{x}^n) \leq 2^{-n(H(X)-\eta)}. \quad (8)$$

2) For n sufficiently large,

$$Pr\left\{\boldsymbol{X}^n \in T^n_{[X]\delta}\right\} > 1 - \delta. \quad (9)$$

3) For n sufficiently large,

$$(1-\delta) 2^{n(H(X)-\eta)} \leq \left| T^n_{[X]\delta} \right| \leq 2^{n(H(X)+\eta)}. \quad (10)$$

For an i.i.d. bivariate information source $\{(X_k, Y_k) : k \geq 1\}$ with probability distribution $p(x, y)$. Let (X, Y) denote the pair of generic random variables, where $H(X, Y) < \infty$.

Definition 5: The *strongly jointly δ-typical* set $T^n_{[XY]\delta}$ with respect to $p(x, y)$ is the set of $(\boldsymbol{x}^n, \boldsymbol{y}^n) \in \mathcal{X}^n \times \mathcal{Y}^n$ such that $N(x, y; \boldsymbol{x}^n, \boldsymbol{y}^n) = 0$ for $(x, y) \notin \mathcal{S}_{XY}$, and

$$\sum_x \sum_y \left| \frac{1}{n} N(x, y; \boldsymbol{x}^n, \boldsymbol{y}^n) - p(x, y) \right| \leq \delta,$$

where $N(x, y; \boldsymbol{x}^n, \boldsymbol{y}^n)$ is the number of occurrences of (x, y) in $(\boldsymbol{x}^n, \boldsymbol{y}^n)$ and δ is an arbitrarily small positive real number.

Strong typicality satisfies the following properties.

Lemma 2: (*Consistency*) If $(\boldsymbol{x}^n, \boldsymbol{y}^n) \in T^n_{[XY]\delta}$, then $\boldsymbol{x}^n \in T^n_{[X]\delta}$ and $\boldsymbol{y}^n \in T^n_{[Y]\delta}$.

Lemma 3: (*Preservation*) Let $Y = f(X)$. If $\boldsymbol{x}^n = (x_1, x_2, \cdots, x_n) \in T^n_{[X]\delta}$, then $f(\boldsymbol{x}^n) = (y_1, y_2, \cdots, y_n) \in T^n_{[Y]\delta}$, where $y_i = f(x_i)$ for $1 \leq i \leq n$.

Lemma 4: (*Strong JAEP*) Let

$$(\boldsymbol{X}^n, \boldsymbol{Y}^n) = ((X_1, Y_1), (X_2, Y_2), \cdots, (X_n, Y_n)),$$

where (X_i, Y_i) are i.i.d. with generic pair of random variables (X, Y). Let λ be a small positive quantity such that $\lambda \to 0$ as $\delta \to 0$.

1) If $(\boldsymbol{x}^n, \boldsymbol{y}^n) \in T^n_{[XY]\delta}$, then

$$2^{-n(H(X,Y)+\lambda)} \leq p(\boldsymbol{x}^n, \boldsymbol{y}^n) \leq 2^{-n(H(X,Y)-\lambda)}. \quad (11)$$

2) For n sufficiently large,

$$\Pr\left\{(\boldsymbol{X}^n, \boldsymbol{Y}^n) \in T^n_{[XY]\delta}\right\} > 1 - \delta. \quad (12)$$

3) For n sufficiently large,

$$(1-\delta) 2^{n(H(X,Y)-\lambda)} \leq \left| T^n_{[XY]\delta} \right| \leq 2^{n(H(X,Y)+\lambda)}. \quad (13)$$

Lemma 5: For any $\boldsymbol{x}^n \in T^n_{[X]\delta}$, define

$$T^n_{[Y|X]\delta}(\boldsymbol{x}^n) = \left\{\boldsymbol{y}^n \in T^n_{[Y]\delta} : (\boldsymbol{x}^n, \boldsymbol{y}^n) \in T^n_{[XY]\delta}\right\}. \quad (14)$$

If $\left| T^n_{[Y|X]\delta}(\boldsymbol{x}^n) \right| \geq 1$, then

$$2^{n(H(Y|X)-\gamma)} \leq \left| T^n_{[Y|X]\delta} \right| \leq 2^{n(H(Y|X)+\gamma)}, \quad (15)$$

where $\gamma \to 0$ as $n \to \infty$ and $\delta \to 0$.

The generalization to a multivariate distribution is straightforward. A more thorough introduction to strongly δ-typical sequences can be found in [2] Chapter 5.

D. The Region $\Gamma^*_{\mathcal{N}}$

Let $\mathcal{N}$ be a nonempty set of random variables and $\mathcal{Q}_{\mathcal{N}} = 2^{\mathcal{N}} \setminus \{\phi\}$ with cardinality $|\mathcal{Q}_{\mathcal{N}}| = 2^{|\mathcal{N}|} - 1$. Let $\mathcal{H}_{\mathcal{N}}$ be the $|\mathcal{Q}_{\mathcal{N}}|$-dimensional Euclidean space with the coordinates labeled by $h_A, A \in \mathcal{Q}_{\mathcal{N}}$. A vector $\boldsymbol{h} = (h_A : A \in \mathcal{Q}_{\mathcal{N}})$ in $\mathcal{H}_{\mathcal{N}}$ is said to be an *entropy function* if there exists a joint distribution for all $X \in \mathcal{N}$ such that $h_A = H(X : X \in A)$ for all $A \in \mathcal{Q}_{\mathcal{N}}$. We then define the region

$$\Gamma^*_{\mathcal{N}} = \{\boldsymbol{h} \in \mathcal{H}_{\mathcal{N}} : \boldsymbol{h} \text{ is an entropy function}\}. \quad (16)$$

Therefore, by the above definition, there exists an one-to-one mapping between each vector $\boldsymbol{h}$ in $\Gamma^*_{\mathcal{N}}$ and some set of random variables whose joint entropies correspond to the elements in $\boldsymbol{h}$. Since an arbitrary information inequality (equality) can be regarded as a half-space (hyperplane) in $\mathcal{H}_{\mathcal{N}}$, it cuts $\Gamma^*_{\mathcal{N}}$ into a subregion that maps to all sets of random variables that possess this property.

Lemma 6: Basic properties of $\Gamma^*_{\mathcal{N}}$:

1) $\Gamma^*_{\mathcal{N}}$ contains the origin.
2) $\bar{\Gamma}^*_{\mathcal{N}}$, the closure of $\Gamma^*_{\mathcal{N}}$, is convex.
3) $\Gamma^*_{\mathcal{N}}$ is in the nonnegative orthant of the space $\mathcal{H}_{\mathcal{N}}$, i.e.,

$$\Gamma^*_{\mathcal{N}} \subseteq \{\boldsymbol{h} \in \mathcal{H}_{\mathcal{N}} : h_A \geq 0 \text{ for all } A \in \mathcal{Q}_{\mathcal{N}}\}.$$

III. Main Result

Consider the set of all information rate tuples $\boldsymbol{\omega}$ such that there exist auxiliary random variables $\{Y_s, s \in S\}$ and $\{U_e, e \in E\}$ which satisfy the following conditions:

$$\begin{aligned} H(Y_s) &\geq \omega_s, s \in S & (17)\\ H(Y_S) &= \sum_{s\in S} H(Y_s) & (18)\\ H(U_{Out(s)}|Y_s) &= 0, s \in S & (19)\\ H(U_{Out(i)}|U_{In(i)}) &= 0, i \in V \setminus (S \cup T) & (20)\\ H(U_e) &\leq R_e, e \in E & (21)\\ H(Y_{\beta(t)}|U_{In(t)}) &= 0, t \in T, & (22)\end{aligned}$$

where U_e is an auxiliary random variable associated with the codeword sent on channel e, and Y_S, $U_{In(i)}$ denote respectively the sets $\{Y_s : s \in S\}$, $\{U_e : e \in In(i)\}$, etc.

For a given acyclic multi-source multi-sink network G, let $\mathcal{N} = \{Y_s : s \in S; U_e : e \in E\}$ and define the following constrained regions in $\mathcal{H}_{\mathcal{N}}$:

$$\mathcal{C}_1 = \left\{\boldsymbol{h} \in \mathcal{H}_{\mathcal{N}} : h_{Y_S} = \sum_{s\in S} h_{Y_s}\right\} \quad (23)$$

$$\mathcal{C}_2 = \left\{\boldsymbol{h} \in \mathcal{H}_{\mathcal{N}} : h_{U_{Out(s)}|Y_s} = 0, s \in S\right\} \quad (24)$$

$$\mathcal{C}_3 = \left\{\boldsymbol{h} \in \mathcal{H}_{\mathcal{N}} : h_{U_{Out(i)}|U_{In(i)}} = 0, i \in V \setminus (S \cup T)\right\} \quad (25)$$

$$\mathcal{C}_4 = \{\boldsymbol{h} \in \mathcal{H}_{\mathcal{N}} : h_{U_e} \leq R_e, e \in E\} \quad (26)$$

$$\mathcal{C}_5 = \left\{\boldsymbol{h} \in \mathcal{H}_{\mathcal{N}} : h_{Y_{\beta(t)}|U_{In(t)}} = 0, t \in T\right\}, \quad (27)$$

where we used the notations $h_{A|A'}$, $h_{AA'}$ and $h_{A'}$ for $H(A|A')$, $H(AA')$ and $H(A')$ for brevity. Clearly, (23) to (27) are the corresponding forms of (18) to (22) in $\mathcal{H}_{\mathcal{N}}$.

Let $con(\Gamma^*_{\mathcal{N}})$ be the convex hull of $\Gamma^*_{\mathcal{N}}$. By time sharing, for any $\boldsymbol{h} \in con(\Lambda)$, $\Lambda \subseteq \Gamma^*_{\mathcal{N}}$, there exists a random variable W and a set of jointly distributed random variables $\{X : X \in \mathcal{N}\}$ such that for any $A \in \mathcal{Q}_{\mathcal{N}}$, $h_{A|W} = H(X : X \in A|W)$. Let $\mathcal{C}_\alpha = \bigcap_{i\in\alpha} \mathcal{C}_i, \alpha \subseteq \{1, 2, 3, 4, 5\}$, we have the following theorem.

Theorem 1: The capacity region for an arbitrary acyclic multi-source multi-sink network is characterized by

$$\mathcal{R} = \Lambda\left(\text{proj}_{Y_S}\left(\overline{con(\Gamma^*_{\mathcal{N}} \cap \mathcal{C}_{123})} \cap \mathcal{C}_4 \cap \mathcal{C}_5\right)\right), \quad (28)$$

where for any $A \subset \mathcal{H}_{\mathcal{N}}$, $\text{proj}_{Y_S}(A) = \{\boldsymbol{h}_{Y_S} : \boldsymbol{h} \in A\}$ is the projection of A on the coordinates $h_{Y_s}, s \in S$, $\Lambda(A) = \{\boldsymbol{h} \in \mathcal{H}_{\mathcal{N}} : \mathbf{0} \leq \boldsymbol{h} \leq \boldsymbol{h}', \boldsymbol{h}' \in A\}$ and $\bar{A}$ is the closure of region A.

IV. Proof of Theorem

A. Proof of Achievability

Let $\boldsymbol{\omega} \in \text{proj}_{Y_S}(\overline{con(\Gamma^*_{\mathcal{N}} \cap \mathcal{C}_{123})} \cap \mathcal{C}_4 \cap \mathcal{C}_5)$. Then there exists an $\boldsymbol{h} \in \overline{con(\Gamma^*_{\mathcal{N}} \cap \mathcal{C}_{123})} \cap \mathcal{C}_4 \cap \mathcal{C}_5$ such that $\boldsymbol{\omega} = \text{proj}_{Y_S}(\boldsymbol{h})$. This implies that there exists a sequence $\boldsymbol{h}^{(k)} \in con(\Gamma^*_{\mathcal{N}} \cap \mathcal{C}_{123})$ such that $\boldsymbol{h} = \lim_{k\to\infty} \boldsymbol{h}^{(k)}$. Note that $\boldsymbol{h}^{(k)}$ might not be in $\Gamma^*_{\mathcal{N}}$, thus is not necessarily an entropy function. However, this can be resolved by using time-sharing. Let $W^{(k)}$ be a time-sharing variable, thus by the definition of $\Gamma^*_{\mathcal{N}}$, there exists a set of random variables

$$\mathcal{N}^{(k)} = \left\{W^{(k)}; Y_s^{(k)} : s \in S; U_e^{(k)} : e \in E\right\}$$

whose entropy function corresponds to $\boldsymbol{h}^{(k)}$, i.e.,

$$\begin{aligned} H\left(Y_s^{(k)}\middle| W^{(k)}\right) &= \omega_s^{(k)}, s \in S & (29)\\ H\left(Y_S^{(k)}\middle| W^{(k)}\right) &= \sum_{s\in S} H\left(Y_s^{(k)}\middle| W^{(k)}\right) & (30)\\ H\left(U_{Out(s)}^{(k)}\middle| Y_s^{(k)}, W^{(k)}\right) &= 0, s \in S & (31)\\ H\left(U_{Out(i)}^{(k)}\middle| U_{In(i)}^{(k)}, W^{(k)}\right) &= 0, i \in V \setminus (S \cup T), & (32)\end{aligned}$$

where $\lim_{k\to\infty} \omega_s^{(k)} = \omega_s$ for all $s \in S$. Since $\boldsymbol{h} \in \mathcal{C}_4 \cap \mathcal{C}_5$ and $\boldsymbol{h} = \lim_{k\to\infty} \boldsymbol{h}^{(k)}$, it is implied that $\mathcal{N}^{(k)}$ must also satisfy

$$\begin{aligned} H\left(U_e^{(k)}\middle| W^{(k)}\right) &\leq R_e + \epsilon_k, e \in E & (33)\\ H\left(Y_{\beta(t)}^{(k)}\middle| U_{In(t)}^{(k)}, W^{(k)}\right) &= \delta_k, t \in T, & (34)\end{aligned}$$

where $\epsilon_k \to 0, \delta_k \to 0$ as $k \to \infty$. From the Caratheodory's Theorem, we may assume that the support $\mathcal{S}_W$ of $W^{(k)}$ has size at most $2^{|\mathcal{N}|}$.

$Y_s^{(k)}$ and $U_e^{(k)}$ are random variables representing the information source X_s and the codeword sent on the channel e. View $W^{(k)}$ as a time-sharing random variable. To prove the result, we need to construct a code for each value w of $W^{(k)}$. The time sharing of the codes gives the overall performance of the code. By (30), Y_s are conditionally independent given $W^{(k)} = w$, and we consider n i.i.d. copies of them having generic distributions $p(y_s|w) : s \in S$. For sequences $\boldsymbol{Y}_s^n$ with block length n, the corresponding joint distribution is $p(\boldsymbol{y}_s^n|w) = p^n(y_s|w)$ where $\boldsymbol{y}_s^n = (y_s^{(1)}, \cdots, y_s^{(n)})$.

Since we next consider only the random variables in $\mathcal{N}^{(k)}$, we temporarily drop the index k for convenience. Similarly, being aware of the conditioning on $W^{(k)} = w$, we drop the conditioning index w unless otherwise specified.

Now we construct a random $(n, (\eta_e, e \in E), (\omega_s - \epsilon, s \in S), (\Delta_t, t \in T))$ code by the following procedure:

1. **Codebook**

a) For each source $s \in S$, generate $2^{n\tau_s}$ sequences of length n randomly and independently according to $p^n(y_s|w)$. We call this codebook $\mathcal{C}_s$ with indices $\{0, 1, \cdots, 2^{n\tau_s} - 1\}$ and cardinality $M_s = 2^{n\tau_s}$, where we assume $2^{n\tau_s}$ is an integer for convenience.

b) The construction of the random codes $\mathcal{C}_s$ for the sources $s \in S$ are done independently.

2. **Encoding**

a) If the message is j at source node s, map it to the jth codeword in $\mathcal{C}_s$ and call this sequence $\boldsymbol{y}_{s_j}^n$.

b) By (31) and (32), for each channel $e \in Out(i), i \in V \setminus (S \cup T)$, there exists a deterministic function u_e such that $U_e = u_e(U_d : d \in In(i))$. Since the network is acyclic, we can show inductively that there exists another deterministic function $\tilde{u}_e$ such that $U_e = \tilde{u}_e(Y_s, s \in S)$. Let $\zeta_e = |T^n_{[U_e|W=w]\delta}|$, by Lemma 5, we have $\zeta_e \leq 2^{n(H(U_e|W=w)+\eta)}$, where $\eta \to 0$ as $\delta \to 0$. Then from (33), by time sharing the code with respect to the probability distribution of W and letting δ be sufficiently small, we have the average rate over channel e being

$$H(U_e|W) + \eta \leq R_e + \epsilon_k + \eta. \qquad (35)$$

Thus for the fixed value $W = w$, by choosing an integer η_e such that

$$2^{n(H(U_e|W=w)+\eta)} \leq \eta_e \leq 2^{n(R_e+\epsilon_k+\eta)}, \qquad (36)$$

we can define a local encoding function k_e such that if $\boldsymbol{U}^n_{In(i)} \in T^n_{[U_{In(i)}|W=w]\delta}$, $\boldsymbol{U}^n_e \in T^n_{[U_e|W=w]\delta}$. Thus we transmit the index of $\boldsymbol{U}^n_e$ in $T^n_{[U_e|W=w]\delta}$ as the codeword on channel e. Otherwise, we transmit a zero codeword on e.

3. **Decoding** For $t \in T$, define the decoding function

$$g_t : \prod_{d \in In(t)} \{0, 1, \cdots, \eta_d\} \to \prod_{s \in \beta(t)} \mathcal{C}_s. \qquad (37)$$

Let $\mathcal{C}_{\beta(t)} = \prod_{s \in \beta(t)} \mathcal{C}_s$ be the joint codebook for all the sources in $\beta(t)$ with cardinality $M_{\beta(t)} = \prod_{s \in \beta(t)} M_s$. We use the following strong typicality decoding. If the received codeword is nonzero for all $d \in In(t)$ and there exists a unique codeword $\boldsymbol{y}^n_{\beta(t)}$ in $\mathcal{C}_{\beta(t)}$ such that $(\boldsymbol{U}^n_{In(t)}, \boldsymbol{y}^n_{\beta(t)}) \in T^n_{[U_{In(t)} Y_{\beta(t)}|W=w]\delta}$, then let $g_t(\boldsymbol{U}^n_{In(t)}) = \boldsymbol{y}^n_{\beta(t)}$. Otherwise, declare a decoding error.

For the above coding procedure, assuming that $\boldsymbol{y}^n_S$ is the source message from all sources and the codewords of $\mathcal{C}_{\beta(t)}$ are ordered as $\{\boldsymbol{c}_i^{n|\beta(t)|} : i = 1, \cdots, M_{\beta(t)}\}$, where the joint codeword $\boldsymbol{c}_i^{n|\beta(t)|}$ is concatenated by $|\beta(t)|$ codewords from $\mathcal{C}_s, s \in \beta(t)$. For notational convenience, we omit the dimension index on $\boldsymbol{c}_i^{n|\beta(t)|}$ and use $\mathbf{c}_i$ instead. Thus

$$\begin{aligned}
&\Pr\{\text{error}|W=w\} \\
=&\Pr\left\{\text{error}\,\middle|\,\boldsymbol{y}^n_S \notin T^n_{[Y_S|W=w]\delta}\right\}\Pr\left\{\boldsymbol{y}^n_S \notin T^n_{[Y_S|W=w]\delta}\right\} \\
&+\Pr\left\{\text{error}\,\middle|\,\boldsymbol{y}^n_S \in T^n_{[Y_S|W=w]\delta}\right\}\Pr\left\{\boldsymbol{y}^n_S \in T^n_{[Y_S|W=w]\delta}\right\} \\
=&\Pr\left\{\boldsymbol{y}^n_S \notin T^n_{[Y_S|W=w]\delta}\right\} + \Pr\left\{\text{error}\,\middle|\,\boldsymbol{y}^n_S \in T^n_{[Y_S|W=w]\delta}\right\} \\
\leq&\lambda_n + \Pr\left\{\text{error}\,\middle|\,\boldsymbol{y}^n_S \in T^n_{[Y_S|W=w]\delta}\right\} \qquad (38)
\end{aligned}$$

where $\lambda_n \to 0$ as $n \to \infty$.

Since $\boldsymbol{y}^n_S \in T^n_{[Y_S|W=w]\delta}$ implies $\boldsymbol{U}^n_{In(t)} \in T^n_{[U_{In(t)}|W=w]\delta}$, a decoding error occurs if and only if there are more than one codeword in $\mathcal{C}_{\beta(t)}$ that satisfy the strong joint typicality condition. Let $\mathbf{c}_i \in \mathcal{C}_{\beta(t)}$ be such a codeword, that is $(\mathbf{c}_i, \boldsymbol{U}^n_{In(t)}) \in T^n_{[Y_{\beta(t)} U_{In(t)}|W=w]\delta}$ and $\mathbf{c}_i \neq \boldsymbol{y}^n_{\beta(t)}$. Call this event E_i, then we have

$$\begin{aligned}
&\Pr\left\{\text{error}\,\middle|\,\boldsymbol{y}^n_S \in T^n_{[Y_S|W=w]\delta}\right\} \\
=\;&\Pr\left\{\cup_{\mathbf{c}_i \neq \boldsymbol{y}^n_{\beta(t)}} \mathrm{E}_i\,\middle|\,\boldsymbol{y}^n_S \in T^n_{[Y_S|W=w]\delta}\right\} \\
=\;&1 - \Pr\left\{\cap_{\mathbf{c}_i \neq \boldsymbol{y}^n_{\beta(t)}} \mathrm{E}^c_i\,\middle|\,\boldsymbol{y}^n_S \in T^n_{[Y_S|W=w]\delta}\right\} \\
\overset{(a)}{=}\;&1 - \Pr\left\{\mathrm{E}^c_1\,\middle|\,\boldsymbol{y}^n_S \in T^n_{[Y_S|W=w]\delta}\right\}^{(M_{\beta(t)}-1)} \\
\leq\;&1 - \left(1 - \Pr\left\{\mathrm{E}_1\,\middle|\,\boldsymbol{y}^n_S \in T^n_{[Y_S|W=w]\delta}\right\}\right)^{M_{\beta(t)}} \\
\overset{(b)}{\leq}\;&1 - \left(1 - \delta' 2^{-n\left(H(Y_{\beta(t)}|W=w)-\delta_k-3\eta\right)}\right)^{M_{\beta(t)}} \\
\overset{(c)}{\leq}\;&1 - \left(1 - M_{\beta(t)}\delta' 2^{-n\left(H(Y_{\beta(t)}|W=w)-\delta_k-3\eta\right)}\right) \\
=\;&M_{\beta(t)}\delta' 2^{-n\left(H(Y_{\beta(t)}|W=w)-\delta_k-3\eta\right)} \\
\overset{(d)}{=}\;&M_{\beta(t)}\delta' 2^{-n\left(\sum_{s\in\beta(t)} H(Y_s|W=w)-\delta_k-3\eta\right)}, \qquad (39)
\end{aligned}$$

where $\delta' = \frac{1}{1-\delta} \to 1$ as $\delta \to 0$. The noted inequalities are explained as follows:

(a) follows since $\mathbf{c}_i \in \mathcal{C}_{\beta(t)}$ are i.i.d. uniformly distributed (we abuse the notation here by assuming $\mathbf{c}_1 \neq \boldsymbol{y}^n_{\beta(t)}$), and the total number of such codewords is $M_{\beta(t)} - 1$.

(b) follows from Lemma 4, 5 and the conditioning on $W = w$. We omit the details for brevity.

(c) follows from the fact that $(1 + a)^n \geq 1 + na$.

(d) follows by applying (30).

Now combine (38) and (39), we have

$$\begin{aligned}
&\Pr\{\text{error}|W=w\} \\
\leq\;&\lambda_n + M_{\beta(t)}\delta' 2^{-n\left(\sum_{s\in\beta(t)} H(Y_s|W=w)-\delta_k-3\eta\right)} \\
=\;&\lambda_n + \delta' 2^{-n(\mu-\delta_k-3\eta)}
\end{aligned}$$

by choosing the codebook cardinalities $M_s = 2^{n\tau_s} = 2^{n(H(Y_s|W=w)-\mu)}$ and $\mu > \delta_k + 3\eta$ for all $s \in S$. Since the support size $|\mathcal{S}_W|$ of W is uniformly bounded above, the time shared code has a vanishing error probability, i.e.,

$$\begin{aligned}
\Pr\{\text{error}\} &= \mathbb{E}[\Pr\{\text{error}|W=w\}] \\
&\leq |\mathcal{S}_W| \cdot \left(\lambda_n + \delta' 2^{-n(\mu-\delta_k-3\eta)}\right) \\
&\to 0, \text{ as } n \to \infty.
\end{aligned}$$

Thus we proved the existence of $\mathcal{N}^{(k)}$, which asserts that $\mathrm{proj}_{Y_S}(\boldsymbol{h}^{(k)}) \in \mathcal{R}$. Letting $k \to \infty$, we obtain $\boldsymbol{\omega} = \mathrm{proj}_{Y_S}(\boldsymbol{h}) \in \mathcal{R}$. Invoking the definition of achievability, we obtain $\Lambda(\boldsymbol{\omega}) \subset \mathcal{R}$, completing the proof of achievability. □

B. Proof of Converse

Let $\boldsymbol{\omega} \in \mathcal{R}$, then for $0 < \epsilon_k \to 0$, we have a sequence of

$$\left(n_k, (\eta_e^{(k)}, e \in E), (\tau_s^{(k)}, s \in S), (\Delta_t^{(k)}, t \in T)\right)$$

codes satisfying

$$\begin{aligned}
n_k^{-1} \log \eta_e^{(k)} &\leq R_e + \epsilon_k, e \in E \qquad (40)\\
\tau_s^{(k)} &\geq \omega_s - \epsilon_k, s \in S \qquad (41)\\
\Delta_t^{(k)} &\leq \epsilon_k, t \in T. \qquad (42)
\end{aligned}$$

In the following discussion, we fix the value of k, and drop k from all notations. Let X_s be uniformly distributed in the codebook $\mathcal{X}_s$ and $\{X_s : s \in S\}$ be independent, then

$$\begin{aligned} H(X_S) &= \sum_{s\in S} H(X_s) & (43)\\ H(U_{Out(s)}|X_s) &= 0, s\in S & (44)\\ H(U_{Out(i)}|U_{In(i)}) &= 0, i\in V\setminus(S\cup T) & (45)\\ H(U_e) &\le n(R_e+2\epsilon_k), e\in E & (46)\\ H(X_{\beta(t)}|U_{In(t)}) &\le n\phi_t(n,\epsilon_k), t\in T, & (47)\\ H(X_s) &\ge n(\omega_s-\epsilon_k), s\in S & (48)\end{aligned}$$

where

- (43) follows from the independence of $X_s, s \in S$.
- (44), (45) follows from the definition of the code.
- (46) follows from (40).
- (47) can be proved similarly as in [3] section 6.3 where

$$\phi_t(n,\epsilon_k) = \frac{1}{n} + 2\epsilon_k \sum_{e\in In(t)} (R_e + \epsilon_k). \qquad (49)$$

- (48) follows from (41) and the fact that X_s is uniformly distributed, i.e., for all $s \in S$

$$H(X_s) = \log|\mathcal{X}_s| = \log\lceil 2^{n\tau_s}\rceil \ge n\tau_s \ge n(\omega_s - \epsilon_k).$$

By letting $Y_s = X_s$ for all $s \in S$, we see there exists a sequence $\boldsymbol{h}^{(k)}$ such that

$$h^{(k)}_{Y_s} \ge n(\omega_s - \epsilon_k), \qquad (50)$$

for all $s \in S$, and

$$\boldsymbol{h}^{(k)} \in \Gamma^*_{\mathcal{N}} \cap \mathcal{C}_{123} \cap \mathcal{C}^n_{4\epsilon_k} \cap \mathcal{C}^n_{5\epsilon_k}, \qquad (51)$$

where

$$\mathcal{C}^n_{4\epsilon_k} = \{\boldsymbol{h} \in \mathcal{H}_{\mathcal{N}} : h_{U_e} \le n(R_e + 2\epsilon_k), e \in E\}$$
$$\mathcal{C}^n_{5\epsilon_k} = \{\boldsymbol{h} \in \mathcal{H}_{\mathcal{N}} : h_{Y_{\beta(t)}|U_{In(t)}} \le n\phi_t(n,\epsilon_k), t \in T\}.$$

Dividing (50) by n, we obtain

$$n^{-1}h^{(k)}_{Y_s} \ge \omega_s - \epsilon_k \qquad (52)$$

for all $s \in S$. Since $\overline{con(\Gamma^*_{\mathcal{N}} \cap \mathcal{C}_{123})}$ is convex and contains the zero vector, we have

$$n^{-1}\boldsymbol{h}^{(k)} \in \overline{con(\Gamma^*_{\mathcal{N}} \cap \mathcal{C}_{123})} \cap \mathcal{C}_{4\epsilon_k} \cap \mathcal{C}_{5\epsilon_k}, \qquad (53)$$

where

$$\mathcal{C}_{4\epsilon_k} = \{\boldsymbol{h} \in \mathcal{H}_{\mathcal{N}} : h_{U_e} \le R_e + 2\epsilon_k, e \in E\}$$
$$\mathcal{C}_{5\epsilon_k} = \{\boldsymbol{h} \in \mathcal{H}_{\mathcal{N}} : h_{Y_{\beta(t)}|U_{In(t)}} \le \phi_t(n,\epsilon_k), t \in T\}.$$

Now define the set

$$\mathcal{B}^{(n,k)} = \Big\{\boldsymbol{h} \in \overline{con(\Gamma^*_{\mathcal{N}} \cap \mathcal{C}_{123})} \cap \mathcal{C}_{4\epsilon_k} \cap \mathcal{C}_{5\epsilon_k} : h_{Y_s} \ge \omega_s - \epsilon_k, \text{ for all } s \in S\Big\}. \qquad (54)$$

Without loss of generality, we let $\epsilon_k \to 0$ monotonically. Note from (49) that $\phi_t(n,\epsilon_k)$ is monotonic with respect to n and k, thus

$$\mathcal{B}^{(n+1,k)} \subset \mathcal{B}^{(n,k)} \text{ and } \mathcal{B}^{(n,k+1)} \subset \mathcal{B}^{(n,k)}.$$

For any fixed k, for sufficiently large n, from (52) and (53), we see that $\mathcal{B}^{(n,k)}$ is nonempty. Since each $\mathcal{B}^{(n,k)}$ is compact (closeness by the definition, boundedness by the constraint in $\mathcal{C}_{4\epsilon_k}$), we see that

$$\lim_{n\to\infty} \mathcal{B}^{(n,k)}$$

is both compact and nonempty. By the same argument,

$$\lim_{k\to\infty}\lim_{n\to\infty} \mathcal{B}^{(n,k)}$$

is also nonempty. Thus there exists some $\boldsymbol{h}'$ such that

$$\boldsymbol{h}' \in \overline{con(\Gamma^*_{\mathcal{N}} \cap \mathcal{C}_{123})} \cap \mathcal{C}_4 \cap \mathcal{C}_5, \qquad (55)$$
$$h'_{Y_s} \ge \omega_s, \text{ for all } s \in S. \qquad (56)$$

Let $\boldsymbol{r} = \text{proj}_{Y_S}(\boldsymbol{h}')$, then we have

$$\boldsymbol{r} \in \text{proj}_{Y_S}\left(\overline{con(\Gamma^*_{\mathcal{N}} \cap \mathcal{C}_{123})} \cap \mathcal{C}_4 \cap \mathcal{C}_5\right), \qquad (57)$$
$$\boldsymbol{r} \ge \boldsymbol{\omega} \text{ (componentwise)} . \qquad (58)$$

By (57) and (58), we finally obtain that

$$\boldsymbol{\omega} \in \Lambda\left(\text{proj}_{Y_S}\left(\overline{con(\Gamma^*_{\mathcal{N}} \cap \mathcal{C}_{123})} \cap \mathcal{C}_4 \cap \mathcal{C}_5\right)\right). \qquad (59)$$

This completes the proof. □

V. Conclusion

In this work, we extended the previous work in [1], [2] and [3] on the capacity region for general acyclic multi-source multi-sink networks. Specifically, we closed the gap between the existing inner and outer bounds by refining the constrained regions in the entropy space. This leads to an exact characterization of the capacity region for general acyclic multi-source multi-sink networks with arbitrary transmission requirements and thus completes the work along this line of research. However, how to explicitly evaluate the obtained capacity region remains an open problem in general.

Acknowledgment

The authors would like to give special thanks to Prof. Ning Cai and Dr. Lihua Song for their valuable comments.

References

[1] L. Song and R. W. Yeung, "Zero-error network coding for acyclic network," *IEEE Trans. Inform. Theory*, vol. 49, no. 12, pp. 3129–3139, Dec. 2003.

[2] R. W. Yeung, *A First Course in Information Theory*. New York: Kluwer/Plenum, 2002.

[3] R. W. Yeung, N. Cai, S.-Y. R. Li, and Z. Zhang, "Network coding theory," *Foundations and Trends in Communications and Information Theory*, vol. 2, no. 4 and 5, pp. 241–381, 2005.

[4] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inform. Theory*, vol. 46, no. 7, pp. 1204–1216, July 2000.

[5] X. Yan, J. Yang, and Z. Zhang, "An outer bound for multisource multisink network coding with minimum cost consideration," *IEEE Trans. Inform. Theory & IEEE/ACM Trans. Networking (joint issue)*, vol. 52, no. 6, pp. 2373–2385, June 2006.

[6] N. Harvey, R. Kleinberg, and A. Lehman, "On the capacity of information networks," *IEEE Trans. Inform. Theory*, vol. 52, no. 6, pp. 2345–2364, June 2006.

[7] G. Kramer and S. A. Savari, "Edge-cut bounds on network coding rates," *Journal of Network and Systems Management*, vol. 14, no. 1, pp. 49–67, Mar. 2006.