

Polymatroids with Network Coding

Te Sun Han
Waseda University
Tokyo, Japan
Email: han@aoni.waseda.jp

Abstract—The problem of network coding for multicasting a single source to multiple sinks has first been studied by Ahlswede, Cai, Li and Yeung in 2000, in which they have established the celebrated max-flow mini-cut theorem on non-physical information flow over a network of independent channels. On the other hand, in 1980, Han has studied the case with correlated multiple sources and a single sink from the viewpoint of polymatroidal functions in which a necessary and sufficient condition has been demonstrated for reliable transmission over the network. This paper presents an attempt to unify both cases, which leads to establish a necessary and sufficient condition for reliable transmission over a noisy network for multicasting all the correlated multiple sources to all the multiple sinks. Furthermore, we address also the problem of transmitting “independent” sources over a multiple-access-type of network as well as over a broadcast-type of network, which reveals that the (co-) polymatroidal structures are intrinsically involved in these types of network coding.

1. INTRODUCTION

On the other hand, in 1980, Han [3] had studied the case with *correlated* multiple sources and a single sink from the viewpoint of polymatroidal functions in which a necessary and sufficient condition has been demonstrated for reliable transmission over a network.

This paper presents an attempt to unify both cases and to generalize it to quite a general case with stationary ergodic correlated sources and *noisy* channels (with *arbitrary* nonnegative real values of capacity that are *not* necessarily *integers*) satisfying the strong converse property (cf. Verdú and Han [7], Han [4]), which leads to establish a necessary and sufficient condition for reliable transmission over a *noisy* network for multicasting correlated multiple sources altogether to every multiple sinks.

It should be noted here that in such a situation with *correlated* multiple sources, the central issue turns out to be how to construct the *matching condition* between source and channel (i.e., joint source-channel coding), instead of of the traditional concept of *capacity region* (i.e., channel coding), although in the special case with *non-correlated* independent multiple sources the problem reduces again to how to describe the capacity region.

Several network models with correlated multiple sources have been studied by some people, e.g., by Barros and Servetto [10], Ho, Médard, Effros and Koetter [14], Ho, Médard, Koetter, Karger, Effros, Shi and Leong [15], Ramamoorthy, Jain, Chou and Effros [16]. Among others, [14], [15] and [16] consider (without attention to the *converse* part) *error-free* network coding for two (or possibly *multiple*) stationary memoryless correlated sources with a single (or possibly *multiple*) sink(s) to study the error exponent problem, where we notice that all the arguments in [14], [15] and [16] can be validated only within the class of stationary memoryless sources of *integer* bit rates and *error-free* channels (i.e., the *identity* mappings) all with *one bit* capacity (or *integer bits* capacity, which is allowed by introducing *multiple* edges); these restrictions are needed solely to invoke “Menger’s theorem” in graph theory. The main result in the present paper is quite free from such seemingly severe restrictions, because we can dispense with the use of Menger’s theorem.

On the other hand, [10] revisits the same model as in Han [3], while [16] mainly focuses on the network with two correlated sources and two sinks to discuss the separation problem of distributed source coding and network coding, where, in addition, cases of two-sources three-sinks, and three-

sources two-sinks are also studied. It should be noted that, in the case of networks with correlated multiple sources, such a *separation* problem is another central issue. In this paper, we describe a necessary and sufficient condition for separability in the case with multiple sources and multiple sinks. (cf. Remark 5.2).

On the other hand, we may consider another network model with *independent* multiple sources but with multiple sinks each of which is required to reliably reproduce a prescribed subset of the multiple sources that depends on each sink. However, the problem with this general model looks quite hard, although, e.g., Yan, Yeung and Zhang [12] and Song, Yeung and Cai [13] have demonstrated the entropy characterizations of the capacity region, which still contain limiting operations and are not computable. Incidentally, Yan, Yang and Zhang [23] have considered, as a computable special case, degree-2 three-layer networks with K -pairs transmission requirements to derive the explicit capacity region. In this paper, for the same reason, we focus on the case in which all the correlated multiple sources is to be multicast to all the multiple sinks and derive a simple necessary and sufficient matching condition in terms of conditional entropy rates and capacity functions. This case can be regarded as the network counterpart of the non-network compound Slepian-Wolf system [22].

We notice here the following; although throughout in the paper we are encountered with the subtleties coming from the general channel and source characteristics assumed, the main logical stream remains essentially unchanged if we consider simpler models, e.g., such as stationary correlated Markov sources together with stationary memoryless noisy channels. This means that considering only simple cases does not help so much at both of the conceptual and notational levels of the arguments. For this reason, we preferred here the compact general settings.

The present paper consists of six sections. In Section 2 notations and preliminaries are described, and in Section 3 we state the main result as

well as its proof. In Section 4 two examples are shown. Section 5 provides another type of necessary and sufficient condition for transmissibility. Also, some detailed comments on the previous papers are given. Finally, in Section 6 we address the routing capacity problems with polymatroids and co-polymatroids.

2. PRELIMINARIES AND NOTATIONS

A. Communication networks

Let us consider an acyclic directed graph $G = (V, E)$ where $V = \{1, 2, \dots, |V|\}$ ($|V| < +\infty$), $E \subset V \times V$, but $(i, i) \notin E$ for all $i \in V$. Here, elements of V are called *nodes*, and elements (i, j) of E are called *edges* or *channels* from i to j . Each edge (i, j) is assigned the *capacity* $c_{ij} \geq 0$, which specifies the maximum amount of information flow passing through the channel (i, j) . If we want to emphasize the graph thus capacitated, we write it as $G = (V, E, C)$ where $C = (c_{ij})_{(i,j) \in E}$. A graph $G = (V, E, C)$ is sometimes called a (communication) network, and indicated also by $\mathcal{N} = (V, E, C)$. We consider two fixed subsets Φ, Ψ of V such that $\Phi \cap \Psi = \emptyset$ (the empty set) with

$$\Phi = \{s_1, s_2, \dots, s_p\},$$

$$\Psi = \{t_1, t_2, \dots, t_q\},$$

where elements of Φ are called *source nodes*, while elements of Ψ are called *sink nodes*. Here, to avoid subtle irregularities, we assume that there are no edges (i, s) such that $s \in \Phi$.

Informally, our problem is how to simultaneously transmit the information generated at the source nodes in Φ altogether to all the sink nodes in Ψ . More formally, this problem is described as in the following subsection.

B. Sources and channels

Each source node $s \in \Phi$ generates a stationary and ergodic source process

$$X_s = (X_s^{(1)}, X_s^{(2)}, \dots), \quad (2.1)$$

where $X_s^{(i)}$ ($i = 1, 2, \dots$) takes values in finite source alphabet \mathcal{X}_s . Throughout in this paper we consider the case in which the whole joint process $X_\Phi \equiv (X_s)_{s \in \Phi}$ is stationary and ergodic. It is then evident that the joint process $X_T \equiv (X_s)_{s \in T}$ is also stationary and ergodic for any T such that $\emptyset \neq T \subset \Phi$. The component processes X_s ($s \in \Phi$) may be correlated. We write X_T as

$$X_T = (X_T^{(1)}, X_T^{(2)}, \dots) \quad (2.2)$$

and put

$$X_T^n = (X_T^{(1)}, X_T^{(2)}, \dots, X_T^{(n)}), \quad (2.3)$$

where $X_T^{(i)}$ ($i = 1, 2, \dots$) takes values in $\mathcal{X}_T \equiv \prod_{s \in T} \mathcal{X}_s$.

On the other hand, it is assumed that all the channels $(i, j) \in E$, specified by the transition probabilities $w_{ij} : A_{ij}^n \rightarrow B_{ij}^n$ with finite input alphabet A_{ij} and finite output alphabet B_{ij} , are *statistically independent* and satisfy the *strong converse* property (see Verdú and Han [7]). It should be noted here that stationarity and memoryless (*noisy* or *noiseless*) channels with finite input/output alphabets satisfy, as very special cases, this property (cf. Gallager [8], Han [4]). Barros and Servetto [10] have considered the case of stationary and memoryless sources/channels with finite alphabets. The following lemma plays a crucial role in establishing the relevant converse of the main result:

Lemma 2.1: (Verdú and Han [7]) The channel capacity c_{ij} of a channel w_{ij} satisfying the strong converse property with finite input/output alphabets is given by

$$c_{ij} = \lim_{n \rightarrow \infty} \frac{1}{n} \max_{X^n} I(X^n; Y^n),$$

where X^n, Y^n are the input and the output of the channel w_{ij} , respectively, and $I(X^n; Y^n)$ is the mutual information (cf. Cover and Thomas [9]). ■

C. Encoding and decoding

In this section let us state the necessary operation of encoding and decoding for network coding with

correlated multiple sources to be multicast to multiple sinks.

With arbitrarily small $\delta > 0$ and $\varepsilon > 0$, we introduce an $(n, (R_{ij})_{(i,j) \in E}, \delta, \varepsilon)$ code as the one as specified by (2.4) ~ (2.9) below, where we use the notation $[1, M]$ to indicate $\{1, 2, \dots, M\}$. How to construct a “good” $(n, (R_{ij})_{(i,j) \in E}, \delta, \varepsilon)$ code will be shown in Direct part of the proof of Theorem 3.1.

1) For all (s, j) ($s \in \Phi$), the encoding function is

$$f_{sj} : \mathcal{X}_s^n \rightarrow [1, 2^{n(R_{sj}-\delta)}], \quad (2.4)$$

where the output of f_{sj} is carried over to the encoder φ_{sj} of channel w_{sj} , while the decoder ψ_{sj} of w_{sj} outputs an estimate of the output of f_{sj} , which is specified by the stochastic composite function:

$$h_{sj} \equiv \psi_{sj} \circ w_{sj} \circ \varphi_{sj} \circ f_{sj} : \mathcal{X}_s^n \rightarrow [1, 2^{n(R_{sj}-\delta)}]; \quad (2.5)$$

2) For all (i, j) ($i \notin \Phi$), the encoding function is

$$f_{ij} : \prod_{k:(k,i) \in E} [1, 2^{n(R_{ki}-\delta)}] \rightarrow [1, 2^{n(R_{ij}-\delta)}], \quad (2.6)$$

where the output of f_{ij} is carried over to the encoder φ_{ij} of channel w_{ij} , while the decoder ψ_{ij} of w_{ij} outputs an estimate of the output of f_{ij} , which is specified by the stochastic composite function:

$$h_{ij} \equiv \psi_{ij} \circ w_{ij} \circ \varphi_{ij} \circ f_{ij} : \prod_{k:(k,i) \in E} [1, 2^{n(R_{ki}-\delta)}] \rightarrow [1, 2^{n(R_{ij}-\delta)}]. \quad (2.7)$$

Here, if $\{k : (k, i) \in E\}$ is empty, we use the convention that f_{ij} is an arbitrary constant function taking a value in $[1, 2^{n(R_{ij}-\delta)}]$;

3) For all $t \in \Psi$, the decoding function is

$$g_t : \prod_{k:(k,t) \in E} [1, 2^{n(R_{kt}-\delta)}] \rightarrow \mathcal{X}_\Phi^n. \quad (2.8)$$

4) *Error probability*

All sink nodes $t \in \Psi$ are required to reproduce a “good” estimate $\hat{X}_{\Phi,t}^n$ (\equiv the output of the decoder g_t) of X_{Φ}^n , through the network $\mathcal{N} = (V, E, C)$, so that the error probability $\Pr\{\hat{X}_{\Phi,t}^n \neq X_{\Phi}^n\}$ be as small as possible. Formally, for all $t \in \Psi$, the probability $\lambda_{n,t}$ of decoding error committed at sink t is required to satisfy

$$\lambda_{n,t} \equiv \Pr\{\hat{X}_{\Phi,t}^n \neq X_{\Phi}^n\} \leq \varepsilon \quad (2.9)$$

for all sufficiently large n . Clearly, $\hat{X}_{\Phi,t}^n$ are the random variables induced by X_{Φ}^n that were generated at all source nodes $s \in \Phi$.

We now need the following definitions.

Definition 2.1 (rate achievability): If there exists an $(n, (R_{ij})_{(i,j) \in E}, \delta, \varepsilon)$ code for any arbitrarily small $\varepsilon > 0$ as well as any sufficiently small $\delta > 0$, and for all sufficiently large n , then we say that the rate $(R_{ij})_{(i,j) \in E}$ is *achievable* for the network $G = (V, E)$. ■

Definition 2.2 (transmissibility): If, for any small $\tau > 0$, the augmented capacity rate $(R_{ij} = c_{ij} + \tau)_{(i,j) \in E}$ is achievable, then we say that the source X_{Φ} is *transmissible* over the network $\mathcal{N} = (V, E, C)$, where $c_{ij} + \tau$ is called the τ -capacity of channel (i, j) . ■

The proof of Theorem 3.1 (both of the converse part and the direct part) are based on these definitions.

D. Capacity functions

Let $\mathcal{N} = (V, E, C)$ be a network. For any subset $M \subset V$ we say that (M, \bar{M}) (or simply, M) is a *cut* and

$$E_M \equiv \{(i, j) \in E | i \in M, j \in \bar{M}\}$$

the *cutset* of (M, \bar{M}) (or simply, of M), where \bar{M} denotes the complement of M in V . Also, we call

$$c(M, \bar{M}) \equiv \sum_{(i,j) \in E, i \in M, j \in \bar{M}} c_{ij} \quad (2.10)$$

the value of the cut (M, \bar{M}) . Moreover, for any subset S such that $\emptyset \neq S \subset \Phi$ (the source node

set) and for any $t \in \Psi$ (the sink node sets), define

$$\rho_t(S) = \min_{M: S \subset M, t \in \bar{M}} c(M, \bar{M}); \quad (2.11)$$

$$\rho_{\mathcal{N}}(S) = \min_{t \in \Psi} \rho_t(S). \quad (2.12)$$

We call this $\rho_{\mathcal{N}}(S)$ the capacity function of $S \subset V$ for the network $\mathcal{N} = (V, E, C)$.

Remark 2.1: A set function $\sigma(S)$ on Φ is called a co-polymatroid * (function) if it holds that

- 1) $\sigma(\emptyset) = 0$,
- 2) $\sigma(S) \leq \sigma(T) \quad (S \subset T)$,
- 3) $\sigma(S \cap T) + \sigma(S \cup T) \geq \sigma(S) + \sigma(T)$.

It is not difficult to check that $\sigma(S) = H(X_S | X_{\bar{S}})$ is a co-polymatroid (see, Han [3]). On the other hand, a set function $\rho(S)$ on Φ is called a polymatroid if it holds that

- 1') $\rho(\emptyset) = 0$,
- 2') $\rho(S) \leq \rho(T) \quad (S \subset T)$,
- 3') $\rho(S \cap T) + \rho(S \cup T) \leq \rho(S) + \rho(T)$.

It is also not difficult to check that for each $t \in \Psi$ the function $\rho_t(S)$ in (2.11) is a polymatroid (cf. Han [3], Meggido [24]), but $\rho_{\mathcal{N}}(S)$ in (2.12) is not necessarily a polymatroid. These properties have been fully invoked in establishing the *matching condition* between source and channel for the special case of $|\Psi| = 1$ (cf. Han [3]). ■

With these preparations we will demonstrate the main result in the next section.

3. MAIN RESULT

The problem that we deal with here is not that of establishing the “capacity region” as usual, because the concept of “capacity region” does not make sense for the general network with correlated sources. Instead, we are interested in the *matching* problem between the correlated source X_{Φ} and the network $\mathcal{N} = (V, E, C)$ (transmissibility: cf.

*In Zhang, Chen, Wicker and Berger [19], the co-polymatroid here is called the *contra*-polymatroid.

Definition 2.2). Under what condition is such a matching possible? This is the key problem here. An answer to this question is just our main result to be stated here.

Theorem 3.1: The source X_Φ is transmissible over the network $\mathcal{N} = (V, E, C)$ if and only if

$$H(X_S|X_{\bar{S}}) \leq \rho_{\mathcal{N}}(S) \quad (\emptyset \neq \forall S \subset \Phi) \quad (3.13)$$

holds. ■

Remark 3.1: The case of $|\Psi| = 1$ was investigated by Han [3], and subsequently revisited by Barros and Servetto [10], while the case of $|\Phi| = 1$ was investigated by Ahlswede, Cai, Li and Yeung [1]. ■

Remark 3.2: If the sources are *mutually independent*, (3.13) reduces to

$$\sum_{i \in S} H(X_i) \leq \rho_{\mathcal{N}}(S) \quad (\emptyset \neq \forall S \subset \Phi).$$

Then, setting the rates as $R_i = H(X_i)$ we have another equivalent form:

$$\sum_{i \in S} R_i \leq \rho_{\mathcal{N}}(S) \quad (\emptyset \neq \forall S \subset \Phi). \quad (3.14)$$

This specifies the *capacity region* of independent message rates in the traditional sense. In other words, in case the sources are independent, the concept of capacity region makes sense. In this case too, channel coding looks like for *non-physical* flows (as for the case of $|\Phi| = 1$, see Ahlswede, Cai, Li and Yeung [1]; and as for the case of $|\Phi| > 1$ see, e.g., Koetter and Médard [17], Li and Yeung [18]). It should be noted that formula (3.14) is *not* derivable by a naive extension of the arguments as used in the case of single-source ($|\Phi| = 1$), irrespective of the comment in [1]. ■

Proof of Theorem 3.1: The proof is based on joint typicality, strong converse property of the channel, acyclicity of the network, ergodicity of the source, random coding arguments, Fano's inequality, subtle classification of the error patterns, and so on. As for the details, see Han [5]).

4. EXAMPLES

In this section we show two examples of Theorem 3.1 with $\Phi = \{s_1, s_2\}$ and $\Psi = \{t_1, t_2\}$.

Example 1. Consider the network as in Fig.1 (called the *butterfly*) where all the solid edges have capacity 1 and the independent sources X_1, X_2 are binary and uniformly distributed. The capacity function of this network is computed as follows:

$$\begin{aligned} \rho_{t_1}(\{s_2\}) &= \rho_{t_2}(\{s_1\}) = 1, \\ \rho_{t_1}(\{s_1\}) &= \rho_{t_2}(\{s_2\}) = 2, \\ \rho_{t_1}(\{s_1, s_2\}) &= \rho_{t_2}(\{s_1, s_2\}) = 2; \end{aligned}$$

$$\begin{aligned} \rho_{\mathcal{N}}(\{s_1\}) &= \min(\rho_{t_1}(\{s_1\}), \rho_{t_2}(\{s_1\})) = 1, \\ \rho_{\mathcal{N}}(\{s_2\}) &= \min(\rho_{t_1}(\{s_2\}), \rho_{t_2}(\{s_2\})) = 1, \\ \rho_{\mathcal{N}}(\{s_1, s_2\}) &= \min(\rho_{t_1}(\{s_1, s_2\}), \rho_{t_2}(\{s_1, s_2\})) \\ &= 2. \end{aligned}$$

On the other hand,

$$\begin{aligned} H(X_1|X_2) &= H(X_1) = 1, \\ H(X_2|X_1) &= H(X_2) = 1, \\ H(X_1 X_2) &= H(X_1) + H(X_2) = 2. \end{aligned}$$

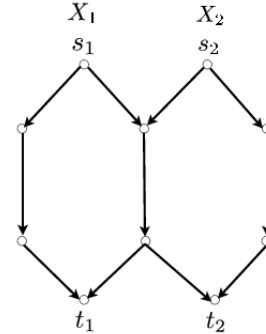


Fig. 1. Example 1

Therefore, condition (3.13) in Theorem 3.1 is satisfied with equality, so that the source is transmissible over the network. Then, how to attain this transmissibility? That is depicted in Fig.2 where \oplus denotes the exclusive OR. Fig. 3 depicts the corresponding capacity region, which is within the

framework of the previous work (e.g., see Ahlswede *et al.* [1]).

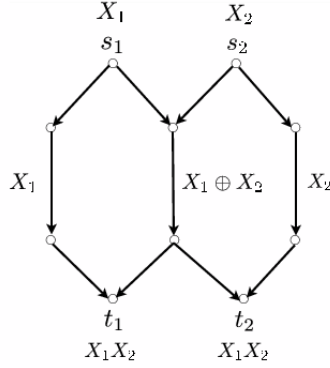


Fig. 2. Coding for Example 1

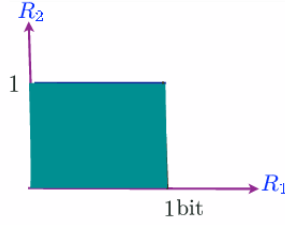


Fig. 3. Capacity region for Example 1

Example 2. Consider the network with *noisy* channels as in Fig.4 where the solid edges have capacity 1 and the broken edges have capacity $h(p) < 1$. Here, $h(p)$ ($0 < p < \frac{1}{2}$) is the binary entropy defined by $h(p) = -p \log_2 p - (1-p) \log_2 (1-p)$. The source (X_1, X_2) generated at the nodes s_1, s_2 is the binary symmetric source with crossover probability p , i.e.,

$$\begin{aligned} \Pr\{X_1 = 1\} &= \Pr\{X_1 = 0\} = \Pr\{X_2 = 1\} \\ &= \Pr\{X_2 = 0\} = \frac{1}{2}, \end{aligned}$$

$$\Pr\{X_2 = 1|X_1 = 0\} = \Pr\{X_2 = 0|X_1 = 1\} = p.$$

Notice that X_1, X_2 are not independent. The capacity function of this network is computed as

follows:

$$\begin{aligned} \rho_{t_1}(\{s_2\}) &= \rho_{t_2}(\{s_1\}) = h(p), \\ \rho_{t_1}(\{s_1\}) &= \rho_{t_2}(\{s_2\}) = 1 + h(p), \\ \rho_{t_1}(\{s_1, s_2\}) &= \rho_{t_2}(\{s_1, s_2\}) = \min(1 + 2h(p), 2); \\ \rho_{\mathcal{N}}(\{s_2\}) &= \min(\rho_{t_1}(\{s_2\}), \rho_{t_2}(\{s_2\})) = h(p), \\ \rho_{\mathcal{N}}(\{s_1\}) &= \min(\rho_{t_1}(\{s_1\}), \rho_{t_2}(\{s_1\})) = h(p), \\ \rho_{\mathcal{N}}(\{s_1, s_2\}) &= \min(\rho_{t_1}(\{s_1, s_2\}), \rho_{t_2}(\{s_1, s_2\})) \\ &= 2. \end{aligned}$$

On the other hand,

$$\begin{aligned} H(X_1|X_2) &= h(p), \\ H(X_2|X_1) &= h(p), \\ H(X_1X_2) &= 1 + h(p). \end{aligned}$$

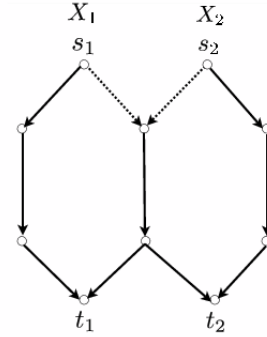


Fig. 4. Example 2

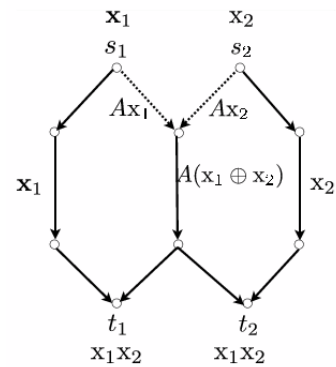


Fig. 5. Coding for Example 2

Therefore, condition (3.13) in Theorem 3.1 is satisfied with strict inequality, so that the source is transmissible over the network. Then, how to attain this transmissibility? That is depicted in Fig.5 where $\mathbf{x}_1, \mathbf{x}_2$ are n independent copies of X_1, X_2 , respectively, and A is an $m \times n$ matrix ($m = nh(p) < n$). Notice that the entropy of $\mathbf{x}_1 \oplus \mathbf{x}_2$ (component-wise exclusive OR) is $nh(p)$ bits and hence it is possible to recover $\mathbf{x}_1 \oplus \mathbf{x}_2$ from $A(\mathbf{x}_1 \oplus \mathbf{x}_2)$ (of length $m = nh(p)$) with asymptotically negligible probability of decoding error, provided that A is appropriately chosen (see Körner and Marton [21]). It should be remarked that this example cannot be justified by the previous works such as Ho *et al.* [14], Ho *et al.* [15], and Ramamoorthy *et al.* [16], because all of them assume *noiseless* channels with capacity of *one bit*, i.e., this example is outside the previous framework.

5. ALTERNATIVE TRANSMISSIBILITY CONDITION

In this section we demonstrate an alternative transmissibility condition equivalent to the necessary and sufficient condition (3.13) given in Theorem 3.1.

To do so, for each $t \in \Psi$ we define the polyhedron \mathcal{C}_t as the set of all nonnegative rates $(R_s; s \in \Phi)$ such that

$$\sum_{i \in S} R_i \leq \rho_t(S) \quad (\emptyset \neq \forall S \subset \Phi), \quad (5.15)$$

where $\rho_t(S)$ is the capacity function as defined in (2.11) of Section 2. Moreover, define the polyhedron \mathcal{R}_{SW} as the set of all nonnegative rates $(R_s; s \in \Phi)$ such that

$$H(X_S|X_{\bar{S}}) \leq \sum_{i \in S} R_i \quad (\emptyset \neq \forall S \subset \Phi), \quad (5.16)$$

where $H(X_S|X_{\bar{S}})$ is the conditional entropy rate as defined in Section 2. Then, we have the following theorem on the transmissibility over the network $\mathcal{N} = (V, E, C)$.

Theorem 5.1: The following two statements are equivalent:

$$1) \quad H(X_S|X_{\bar{S}}) \leq \rho_{\mathcal{N}}(S) \quad (\emptyset \neq \forall S \subset \Phi), \quad (5.17)$$

$$2) \quad \mathcal{R}_{\text{SW}} \cap \mathcal{C}_t \neq \emptyset \quad (\forall t \in \Psi). \quad (5.18)$$

In order to prove Theorem 5.1 we need the following lemma:

Lemma 5.1 (Han [3]): Let $\sigma(S)$, $\rho(S)$ be a copolymatroid and a polymatroid, respectively, as defined in Remark 2.1. Then, the necessary and sufficient condition for the existence of some non-negative rates $(R_s; s \in \Phi)$ such that

$$\sigma(S) \leq \sum_{i \in S} R_i \leq \rho(S) \quad (\emptyset \neq \forall S \subset \Phi) \quad (5.19)$$

is that

$$\sigma(S) \leq \rho(S) \quad (\emptyset \neq \forall S \subset \Phi). \quad (5.20)$$

Proof of Theorem 5.1: It is easy to see this.

Remark 5.1: The necessary and sufficient condition of the form (5.18) appears (*without* the proof) in Ramamoorthy, Jain, Chou and Effros [16], which they call the *feasibility*. They attribute the sufficiency part simply to Ho, Médard, Effros and Koetter [14] with $|\Phi| = 2, |\Psi| = 1$ (also, cf. Ho, Médard, Koetter, Karger, Effros, Shi, and Leong [15] with $|\Phi| = 2, |\Psi| = 1$), while attributing the necessity part to Han [3], Barros and Servetto [10]. However, notice that all the arguments in [14], [15] ([14] is included in [15]) can be validated only within the class of stationary memoryless sources of *integer* bit rates and *error-free* channels (i.e., the *identity* mappings) all with *one bit* capacity (this restriction is needed to invoke “Menger’s theorem” in graph theory); while the present paper, without such “seemingly” severe restrictions, treats “general” acyclic networks, allowing for general correlated stationary ergodic sources as well as general statistically independent channels with each satisfying the strong converse property (cf. Lemma 2.1). Moreover, as long as we are concerned also with *noisy* channels, the way of approaching the

problem as in [14], [15] does *not* work as well, because in this noisy case we have to cope with two kinds of error probabilities, one due to error probabilities for source coding and the other due to error probabilities for network coding (i.e., channel coding); thus in the noisy channel case or in the noiseless channel case with *non-integer* capacities and/or i.i.d. sources of *non-integer* bit rates, [16] cannot attribute the sufficiency part of (5.18) to [14], [15].

It should be noted here also that [14] and [15], though demonstrating relevant error exponents (the *direct* part), do not have the *converse* part. ■

Remark 5.2 (Separation): Here, the term of *separation* is used to mean “two step” separation of distributed source coding (*Step 1*: based on Slepian-Wolf theorem) and network coding with *independent* sources (*Step 2*: based on Theorem 3.1 with Remark 3.2). Now suppose that there exist some nonnegative rates R_i ($i \in \Phi$) such that

$$H(X_S|X_{\bar{S}}) \leq \sum_{i \in S} R_i \leq \rho_{\mathcal{N}}(S) \quad (\emptyset \neq \forall S \subset \Phi). \quad (5.21)$$

Then, the first inequality ensures *reliable* distributed source coding by virtue of the theorem of Slepian and Wolf (*Step 1*), while the second inequality ensures *reliable* network coding by virtue of Theorem 3.1 with Remark 3.2 for independent sources (*Step 2*), which looks like for *non-physical* flows, with *independent* distributed sources of rates R_i ($i \in \Phi$; see Remark 3.2). Thus, (5.21) is *sufficient* for separability. Condition (5.21) is equivalently written as

$$\mathcal{R}_{\text{SW}} \cap \left(\bigcap_{t \in \Psi} \mathcal{C}_t \right) \neq \emptyset \quad (5.22)$$

for any general network \mathcal{N} . Moreover, in view of Remark 3.2, it is not difficult to check that (5.22) is also necessary. Thus, our conclusion is that, in general, condition (5.22) is not only *sufficient* but also *necessary* for separability in the sense above stated. ■

Remark 5.3: In this connection, [16] claims that, in the case with $|\Phi| = |\Psi| = 2$ and with *rational*

capacities as well as sources of *integer* bit rates, “*separation*” always holds, while in the case of $|\Phi| > 2$ or $|\Psi| > 2$ no conclusive claim is still not made. Although this seems to contradict Remark 5.2, it should be merely due to the different definitions for separability.

Remark 5.4: It is possible also to consider network coding with *cost*. In this regard the reader may refer to, e.g., Han [3], Ramamoorthy [28], Lee *et al.* [29]. ■

Remark 5.5: So far we have focused on the case where the channels of a network are quite general but are statistically independent. On the other hand, we may think of the case where the channels are not necessarily statistically independent. This problem is quite hard in general. A typical tractable example of such networks would be a class of acyclic deterministic relay networks with no interference (called the Aref network) in which the concept of “channel capacity” is irrelevant. In this connection, Ratnakar and Kramer [25] have studied Aref networks with a single source and multiple sinks, while Korada and Vasudevan [26] have studied Aref networks with multiple correlated sources and multiple sinks. The network capacity formula as well as the network matching formula obtained by them are in nice correspondence with the formula obtained by Ahlswede *et al.* [1] as well as Theorem 3.1 established in this paper, respectively. ■

6. ROUTING CAPACITY REGIONS

So far we have considered the problem of multicasting multiple sources to multiple sinks over a *noisy* network in which all the sources may be mutually *correlated*. The fundamental tools for this kind of reliable transmission are mainly *routing* and *coding* at each node of the network. Along this line we have established Theorem 3.1 to give a necessary and sufficient condition for reliable transmission. On the other hand, several class of network coding may not need the operation of *coding* but only that of *routing*. However, Theorem 3.1 does not provide any explicit suggestions or answers in this respect.

In this section, we address this problem with mutually *independent* sources specified only by their *rates* R_i 's. The set of all such achievable rates will be called the *routing capacity* of the network. In what follows, as illustrative cases, we take three types of network routings, i.e., multiple-access-type of routing, broadcast-type of routing, and interference-type of routing. In doing so, for each edge $(i, j) \in E$ we impose not only *upper* capacity c_{ij} restriction but also *lower* capacity d_{ij} restriction ($0 \leq d_{ij} \leq c_{ij}$), which means that information flows g_{ij} passing the channel $(i, j) \in E$ are restricted so that $d_{ij} \leq g_{ij} \leq c_{ij}$ for all $(i, j) \in E$. A motivation for the introduction of such lower capacities d_{ij} is that in some situations “informational outage” over a network is to be avoided, for example.

Let us first state “Hoffman’s theorem” in a graph theory which is needed to discuss the routing capacities. For simplicity we put $D = (d_{ij})_{(i,j) \in E}$ like $C = (c_{ij})_{(i,j) \in E}$, and also write as $[d_{ij}, c_{ij}]$.

Definition 6.1: Given a graph $G = (V, E, C, D)$, a flow g_{ij} is said to be *circular* if $d_{ij} \leq g_{ij} \leq c_{ij}$ for all $(i, j) \in E$, and $\sum_{j \in V} g_{ij} = \sum_{j \in V} g_{ji} = 0$ for all $i \in V$ (the conservation law). ■

Theorem 6.1: [Hoffman’s theorem] There exists a circular flow $(g_{ij} : (i, j) \in E)$ if and only if

$$c(M, \bar{M}) \geq d(\bar{M}, M) \quad \text{for all subset } M \subset V, \quad (6.23)$$

where $c(M, \bar{M})$ was specified in (2.10) and $d(\bar{M}, M)$ is similarly defined by replacing c, M, \bar{M} by d, \bar{M}, M , respectively, where \bar{M} denotes the complement of M in V . ■

E. Multiple-access-type of network routing

Suppose that we are given a network $\mathcal{N} = (V, E, C, D)$ as in Fig.6:

We modify this network by adding p fictitious edges as in Fig.7:

to obtain the modified network $\mathcal{N}^* = (V, E^*, C, D, [R_1, R_1], \dots, [R_p, R_p])$. We notice here that a rate

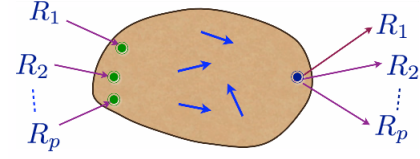


Fig. 6. Network \mathcal{N}

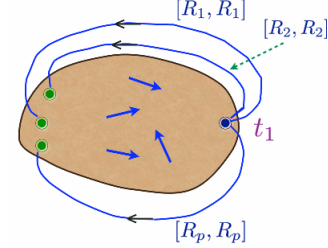


Fig. 7. Modified network \mathcal{N}^*

(R_1, \dots, R_p) is achievable over the original network \mathcal{N} if and only if the network \mathcal{N} has a circular flow. Thus, by writing down all the conditions in Hoffman’s theorem for \mathcal{N}^* , we obtain the following two kinds of inequalities:

1. $M \rightarrow \bar{M}$:

$$\sum_{i \in A} R_i \geq d(\bar{M}, M) - c(M, \bar{M}) \equiv d^*(\bar{M}, M)$$

$$(A \subset \Phi, M \ni t_1, \bar{M} \supset A; M \supset \bar{A})$$
2. $\bar{M} \rightarrow M$:

$$\sum_{i \in A} R_i \leq c(\bar{M}, M) - d(M, \bar{M}) \equiv c^*(\bar{M}, M)$$

$$(A \subset \Phi, M \ni t_1, \bar{M} \supset A; M \supset \bar{A})$$

can be omitted

can be omitted

Fig. 8. Conditions for achievability

In order to express these conditions in a compact form, define the “capacity function” as follows:

Definition 6.2: For each subset A of Φ , let

$$\begin{aligned} \rho_m(A) &\equiv \min\{c^*(\bar{M}, M) | M \ni t_1, \bar{M} \supset A\}, \\ \sigma_m(A) &\equiv \max\{d^*(\bar{M}, M) | M \ni t_1, M \supset \bar{A}\}. \end{aligned}$$

Theorem 6.2: $\rho_m(A)$, $\sigma_m(A)$ are a polymatroid and a co-polymatroid, respectively.

Theorem 6.3: There exists a circular flow if and only if

$$\sigma_m(A) \leq \rho_m(A) \quad \text{for all } A \subset \Phi, \quad (6.24)$$

and the routing capacity region of multiple-access-type of network is given by

$$\sigma_m(A) \leq \sum_{i \in A} R_i \leq \rho_m(A) \quad \text{for all } A \subset \Phi. \quad (6.25)$$

Proof: It suffices only to use Lemma 5.1. ■

Remark 6.1: It is easy to see that $\sigma_m(A) = 0$ for all $A \subset \Phi$ if $d_{ij} = 0$ (for all $(i, j) \in E$), and hence Theorem 6.3 turns out to be a special case of Theorem 3.1 with independent sources. This implies that in this case the capacity region given by Theorem 3.1 can actually be attained only with network routing but without network coding.

F. Broadcast-type of network routing

Suppose that we are given a network $\mathcal{N} = (V, E, C, D)$ as in Fig.9:

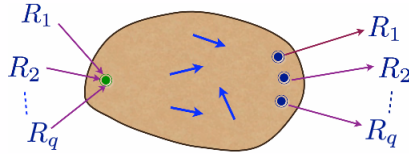


Fig. 9. Broadcast-type of network

We modify this network by adding q fictitious edges as in Fig.10:

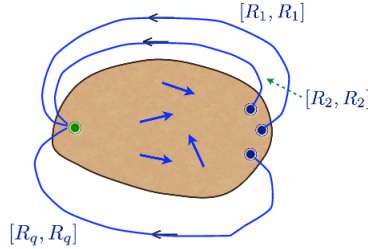


Fig. 10. Modified network \mathcal{N}^*

to obtain the modified network $\mathcal{N}^* = (V, E^*, C, D, [R_1, R_1], \dots, [R_q, R_q])$. Thus, in a similar manner, we have the achievability conditions:

1. $M \rightarrow \bar{M}$:

$$\sum_{i \in A} R_i \leq c(M, \bar{M}) - d(\bar{M}, M) \equiv c^*(M, \bar{M})$$

$$(A \subset \Psi, M \ni s_1, \bar{M} \supset A; M \supset \bar{A})$$

2. $\bar{M} \rightarrow M$:

$$\sum_{i \in A} R_i \geq d(M, \bar{M}) - c(\bar{M}, M) \equiv d^*(M, \bar{M})$$

$$(A \subset \Psi, M \ni s_1, \bar{M} \supset A; M \supset \bar{A})$$

can be omitted
can be omitted

Fig. 11. Conditions for achievability

Accordingly, the “capacity function” is defined as:

Definition 6.3: For each subset A of Ψ , let

$$\rho_b(A) \equiv \min\{c^*(M, \bar{M}) | M \ni s_1, \bar{M} \supset A\},$$

$$\sigma_b(A) \equiv \max\{d^*(M, \bar{M}) | M \ni s_1, M \supset \bar{A}\}.$$

Theorem 6.4: $\rho_b(A)$, $\sigma_b(A)$ are a polymatroid and a co-polymatroid, respectively.

Theorem 6.5: There exists a circular flow if and only if

$$\sigma_b(A) \leq \rho_b(A) \quad \text{for all } A \subset \Psi, \quad (6.26)$$

and the routing capacity region of broadcast-type of network is given by

$$\sigma_b(A) \leq \sum_{i \in A} R_i \leq \rho_b(A) \quad \text{for all } A \subset \Psi. \quad (6.27)$$

Proof: It suffices only to use Lemma 5.1. ■

Remark 6.2: The capacity region (6.27) remains unchanged if network coding is allowed in addition to network routing. ■

G. Interference-type of network routing

Suppose that we are given a network $\mathcal{N} = (V, E, C, D)$ as in Fig.12:

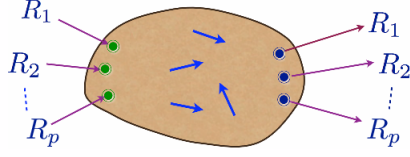


Fig. 12. Interference-type of network

We modify this network by adding p fictitious edges as in Fig.13 below.

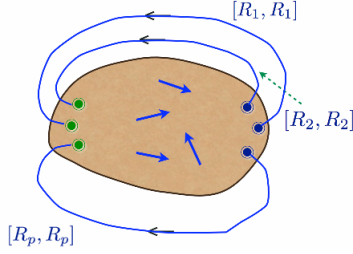


Fig. 13. Modified network \mathcal{N}^*

In this case too, again by virtue of Hoffman's theorem, the achievability condition is summarized as

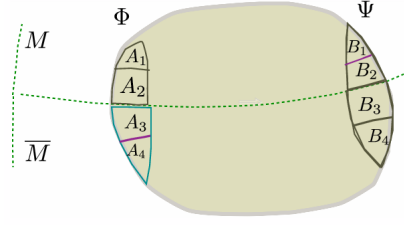
$$\begin{aligned} \max_{(A,B)} d^*(\overline{M}, M) &\leq \sum_{i \in A_2} R_i - \sum_{i \in A_3} R_i \\ &\leq \min_{(A,B)} c^*(M, \overline{M}), \end{aligned} \quad (6.28)$$

where $\max_{(A,B)}$ (resp. $\min_{(A,B)}$) denotes \max (resp. \min) over all possible cuts (M, \overline{M}) and partitions (with A_2, A_3 fixed) as was shown in Fig.14:

For simplicity, let us now consider the case with $|\Phi| = |\Psi| = 2$:

Then, the achivability condition reduces to that as in Fig.16 below, where, on the contrary to the case of multiple-access-type of network routing as well as to the case of broadcast-type of network routing, the term $R_1 - R_2$ appears in addition to the term $R_1 + R_2$, which suggests that the interference-type of network routing may be more complicated with a rather pathological behavior.

The following Fig.17 shows a typical region of



$$\Phi = A_1 \cup A_2 \cup A_3 \cup A_4$$

$$\Psi = B_1 \cup B_2 \cup B_3 \cup B_4$$

$$A_1 \Leftarrow B_1, A_2 \Leftarrow B_3, A_3 \Leftarrow B_2, A_4$$

$$|A_1| = |B_1|, |A_2| = |B_3|,$$

$$|A_3| = |B_2|, |A_4| = |B_4|$$

Fig. 14. Conditions for achievability

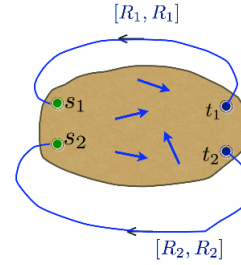


Fig. 15. Modified network \mathcal{N}^*

Case 1: $0 \leq R_1 \leq \min_M c(M, \overline{M})$ **where**

$$[s_1 \in M, t_1 \in \overline{M}] \text{ and } [\{s_2 \in M, t_2 \in M\} \text{ or } \{s_2 \in \overline{M}, t_2 \in \overline{M}\}]$$

Case 2: $0 \leq R_2 \leq \min_M c(M, \overline{M})$ **where**

$$[s_2 \in M, t_2 \in \overline{M}] \text{ and } [\{s_1 \in M, t_1 \in M\} \text{ or } \{s_1 \in \overline{M}, t_1 \in \overline{M}\}]$$

Case 3: $0 \leq R_1 + R_2 \leq \min_M c(M, \overline{M})$ **where**

$$[s_1 \in M, t_1 \in \overline{M}] \text{ and } [s_2 \in M, t_2 \in \overline{M}]$$

Case 4: $0 \leq R_1 \leq \min_M c(M, \overline{M}) + R_2,$

$$0 \leq R_2 \leq \min_M c(\overline{M}, M) + R_1 \text{ **where**}$$

$$[s_1 \in M, t_1 \in \overline{M}] \text{ and } [s_2 \in \overline{M}, t_2 \in M]$$

Fig. 16. Conditions for achievability with $d_{ij} = 0$

Fig.16, which looks like a rather untraditional shape.

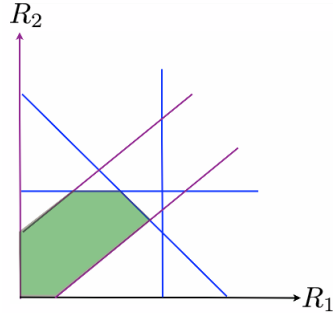


Fig. 17. “achievable” region

For example, let us consider the network \mathcal{N} as in Fig.18. It is easy to check that the achievability condition in Fig. 16 for this network reduces to $0 \leq R_1 \leq R_2 \leq 1$. Fig.19 shows two elementary directed cycles: one with $R_1 = 0, R_2 = 1$; the other with $R_1 = 1, R_2 = 1$. The first flow actually specifies a flow in \mathcal{N} , while the latter flow contains two fictitious edges and hence does not specify any flow in \mathcal{N} . Thus, in the case of the interference-type of network, we have to take another approach to establish the routing capacity region. An approach, though brute, would be to list up all the elementary paths in \mathcal{N} connecting source s_i and sink t_i for every $i = 1, 2, \dots, p$; and, for every edge e of \mathcal{N} , to check whether the total flow passing through e satisfies the upper and lower capacity constraints. Incidentally, it should also be remarked that any polymatroidal structure does not appear here. Anyway, Hoffman’s theorem does not work as well for interference-type of networks. Thus, the routing capacity problem for interference-type of networks would be rather intractable and elusive, compared to multiple-access-type of networks and broadcast-type of networks.

ACKNOWLEDGMENTS

@ The author is very grateful to Prof. Shin’ichi Oishi for providing him with pleasant research facilities during this work. Thanks are also due to

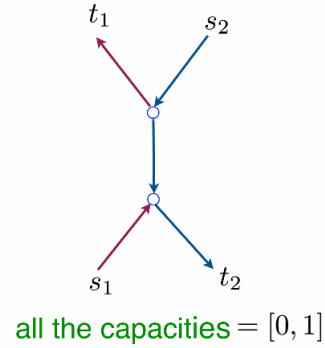


Fig. 18. a simple graph

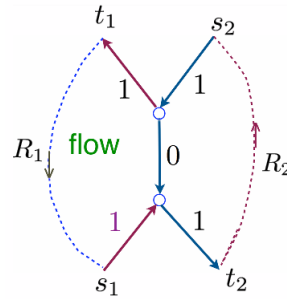


Fig. 19. non-operational cycle with $R_1 = R_2 = 1$

Dinkar Vasudevan for bringing reference [26] to the author’s attention.

REFERENCES

- [1] R. Ahlswede, N.Cai, S.Y. R. Li and R.W.Yeung, “Network information flow,” *IEEE Transactions on Information Theory*, vol.IT-46, no.4, pp. 1204-1216, 2000
- [2] R.W. Yeung, *A First Course in Information Theory*, Kluwer, 2002
- [3] T.S. Han, “Slepian-Wolf-Cover theorem for a network of channels,” *Information and Control*, vol.47, no.1, pp. 67-83, 1980
- [4] T.S. Han, *Information-Spectrum Methods in Information Theory*, Springer-verlag, Berlin, 2003

- [5] T.S. Han, "Multicasting correlated multi-source to multi-sink over a network," arXiv:0901.0608, submitted to *IEEE Transactions on Information Theory*
- [6] T. M. Cover, "A simple proof of the data compression theorem of Slepian and Wolf for ergodic sources," *IEEE Transactions on Information Theory*, vol.IT-21, pp. 226-228, 1975
- [7] S. Verdú and T.S.Han, "A general formula for channel capacity," *IEEE Transactions on Information Theory*, vol.IT-40, no.4, pp.1147-1157, 1994
- [8] R. G. Gallager, *Information Theory and Reliable Communication*, Wiley, New York, 1968
- [9] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, Wiley, New York, 1991
- [10] J. Barros and S. D. Servetto, "Network information flow with correlated sources," *IEEE Transactions on Information Theory*, vol.IT-52, no.1, pp.155-170, 2006
- [11] L. Song, R.W.Yeung and N.Cai, "A separation theorem for single-source network coding," *IEEE Transactions on Information Theory*, vol.IT-52, no.5, pp.1861-1871, 2006
- [12] X. Yan, R.W.Yeung and Z. Zhang, "The capacity region for multiple-source multiple-sink network coding," *Proc. IEEE International Symposium on Information Theory*, June 2007
- [13] L. Song, R.W. Yeung and N.Cai, "Zero-error network coding for acyclic networks," *IEEE Transactions on Information Theory*, vol.IT-49, no.12, pp.3129-3139, 2003
- [14] T. Ho, M. Médard, M.Effros and R. Koetter, "Network coding for correlated sources," *Proc. Conference on Information Science and Systems*, 2004
- [15] T. Ho, M. Médard, and R. Koetter, D.R.Karger, M.Effros, Jun Shi, and Ben Leong, "A random linear network coding approach to multicast," *IEEE Transactions on Information Theory*, vol.IT-52, no.10, pp.4413-4430, 2006
- [16] A. Ramamoorthy, K. Jain, A. Chou and M.Effros, "Separating distributed source coding from network coding," *IEEE Transactions on Information Theory*, vol.IT-52, no.6, pp.2785-2795, 2006
- [17] R. Koetter and M. Médard, "An algebraic approach to network coding," *IEEE Transactions on Information Theory*, vol.IT-49, no.5, pp.782-795, 2003
- [18] A. Li and R.W. Yeung, "Network information flow - multiple source," *IEEE International Symposium on Information Theory*, 2001
- [19] X. Zhang, Jun Chen, S. B. Wicker and T. Berger, "Successive coding in multiuser information theory," *IEEE Transactions on Information Theory*, vol.IT-53, no.6, pp.2246-2254, 2007
- [20] I. Csiszár, "Linear codes for sources and source networks: error exponents, universal coding," *IEEE Transactions on Information Theory*, vol.IT-28, no.4, pp.585-592, 1982
- [21] J. Körner and Marton, "How to encode the modulo-two sum of binary sources," *IEEE Transactions on Information Theory*, vol.IT-25, pp. 219-221, 1979
- [22] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, Academic Press, New York, 1981
- [23] X. Yan, J. Yang and Z. Zhang, "An outer bound for multiple source multiple sink coding with minimum cost consideration," *IEEE Transactions on Information Theory*, vol.IT-52, no.6, pp. 2373-2385, 2006
- [24] N. Meggido, "Optimal flows in networks with multiple sources and multiple sinks," *Mathematical Programming*, vol.7, pp.97-107, 1974
- [25] N. Ratnakar and G. Kramer, "The multicast capacity of deterministic relay networks with no interference," *IEEE Trans. Inform. Theory*, vol. 52, no. 6, pp.2425-2432, Jun. 2006.
- [26] S. B. Korada and D. Vasudevan, "Broadcast and Slepian-Wolf multicast over Aref networks," *Proc. IEEE International Symposium on Information Theory*, pp. 1656-1660, Toronto, Canada, July 6-11, 2008
- [27] S. Avestimehr *Wireless network information flow*, PhD thesis, University of California, Berkeley, 2008
- [28] A. Ramamoorthy, "Minimum cost distributed source coding over a network," *Proc. IEEE International Symposium on Information Theory*, Nice, France, 2007; accepted by *IEEE Trans. Inform. Theory*; <http://arxiv.org/abs/0704.2808>
- [29] A. Lee, M.Médard, K.Z. Haigh, S. Gowan, and P. Ruppel, "Minimum-cost subgraphs for joint distributed source and network coding," *Thurs Workshop on Network Coding, Theory, and Applications*, January 2007
- [30] S. Avestimehr, "Minimum cost distributed source coding over a network," *Wireless network information flow: a deterministic approach*, PhD thesis, University of California, Berkeley, 2008; <http://www.eecs.berkeley.edu/Pubs/TechRpts/2008/EECS-2008-128.html>