# UC Irvine

**Title**

On Routing-Optimal Networks for Multiple Unicasts

**Permalink**

**Authors**

Meng, Chun
Markopoulou, Athina

**Publication Date**

**DOI**

Peer reviewed

# On Routing-Optimal Networks for Multiple Unicasts

Chun Meng, Athina Markopoulou

EECS Department, University of California, Irvine

email: {cmeng1,athina}@uci.edu

*Abstract*—In this paper, we consider the problem of multiple unicast sessions over a directed acyclic graph. It is well known that linear network coding is insufficient for achieving the capacity region, in the general case. However, there exist networks for which routing is sufficient to achieve the whole rate region, and we refer to them as *routing-optimal networks*. We identify a class of routing-optimal networks, which we refer to as *information-distributive networks*, defined by three topological features. Due to these features, for each rate vector achieved by network coding, there is always a routing scheme such that it achieves the same rate vector, and the traffic transmitted through the network is exactly the information transmitted over the cut-sets between the sources and the sinks in the corresponding network coding scheme. We present examples of information-distributive networks, including some examples from (1) index coding and (2) from a single unicast session with hard deadline constraint.

## I. Introduction

In this paper, we consider network coding for multiple unicast sessions over directed acyclic graphs. In general, non-linear network coding should be considered in order to achieve the whole rate region of network coding [1]. Yet, there exist networks, for which routing is sufficient to achieve the whole rate region. We refer to these networks as *routing-optimal* networks. We attempt to answer the following questions: 1) What are the distinct topological features of these networks? 2) Why do these features make a network routing-optimal? The answers to these questions will not only explain which kind of networks can or cannot benefit from network coding, but will also deepen our understanding on how network topologies affect the rate region of network coding.

A major challenge is that there is currently no effective method to calculate the rate region of network coding. Some researchers proposed to use information inequalities to approximate the rate region [2]. However, except for very simple networks, it is very difficult to use this approach since there is potentially an exponential number of inequalities that need to be considered. [3] provides a formula to calculate the rate region by finding all possible entropy functions, which are vectors of an exponential number of dimensions, thus very difficult to solve even for simple networks.

In this paper, we employ a graph theoretical approach in conjunction with information inequalities to identify topological features of routing-optimal networks. Our high-level idea is as follows. Consider a network code. For each unicast session, we choose a cut-set $C$ between source and sink, and a set $\mathcal{P}$ of paths from source to sink such that each path in $\mathcal{P}$ passes through an edge in $C$. Since the information transmitted from the source is totally contained in the information transmitted along the edges in $C$, we can think of distributing the source

information along the edges in $C$ (details will be explained later). Moreover, we consider a routing scheme in which the traffic transmitted along each path $P \in \mathcal{P}$ is exactly the source information distributed over the edge in $C$ that is traversed by $P$. Such a routing scheme achieves the same rate vector as the network code. However, since the edges might be shared among multiple unicast sessions, such a routing scheme might not satisfy the edge capacity constraints. This suggests that the cut-sets and path-sets we choose for the unicast sessions should have special features. These are essentially the features we are looking for to describe routing-optimal networks.

We make the following contributions:

- We identify a class of networks, called *information-distributive* networks, which are defined by three topological features. The first two features capture how the edges in the cut-sets are connected to the sources and the sinks, and the third feature captures how the paths in the path-sets overlap with each other. Due to these features, given a network code, there is always a routing scheme such that it achieves the same rate vector as the network code, and the traffic transmitted through the network is exactly the source information distributed over the cut-sets between the sources and the sinks.

- We prove that if a network is information-distributive, it is routing-optimal. We also show that the converse is not true. This indicates that the three features might be too restrictive in describing routing-optimal networks.

- We present examples of information-distributive networks taken from the index coding problem [4] and single unicast with hard deadline constraint.

We expect that our work will provide helpful insights towards characterizing all possible routing-optimal networks.

## II. Preliminaries

### A. Network Model

The network is represented by an acyclic directed multigraph $G = (V, E)$, where $V$ and $E$ are the set of nodes and the set of edges in the network respectively. Edges are denoted by $e = (u, v, i) \in V \times V \times \mathbb{Z}_{\geq 0}$, or simply by $(u, v)$, where $v = \text{head}(e)$ and $u = \text{tail}(e)$. Each edge represents an error-free and delay-free channel with capacity rate of one. Let $\text{In}(v)$ and $\text{Out}(v)$ denote the set of incoming edges and the set of outgoing edges at node $v$.

There are $K \geq 1$ unicast sessions in the network. The $i$th unicast session is denoted by a tuple $\omega_i = (s_i, d_i)$, where $s_i$ and $d_i$ are the source and the sink of $\omega_i$ respectively. The message sent from $s_i$ to $d_i$ is assumed to be a uniformly distributed random variable $Y_i$ with finite alphabet $\mathcal{Y}_i = \{1, \cdots, \lceil 2^{nR_i} \rceil\}$,

where $R_i$ is the source information rate at $s_i$. All $Y_i$'s are mutually independent. Given $1 \leq i \leq j \leq K$, denote $Y_{i:j} = \{Y_m : i \leq m \leq j\}$. We assume $\text{In}(s_i) = \text{Out}(d_i) = \emptyset$ for all $1 \leq i \leq K$.

Let $\text{mincut}(u,v,G)$ denote the minimum capacity of all cut-sets between two nodes $u$ and $v$. Given two nodes $u, v$, let $\mathcal{P}_{uv}$ denote the set of directed paths from $u$ to $v$. The *routing domain* of $\omega_i$, denoted by $G_i$, is the sub-graph induced by the edges of the paths in $\mathcal{P}_{s_i d_i}$.

### B. Routing Scheme

A *routing scheme* is a transmission scheme where each node only replicates and forwards the received messages onto its outgoing edges. Define the following linear constraints:

$$\sum_{P \in \mathcal{P}_{s_i d_i}} f_i(P) \geq R_i' \qquad \forall 1 \leq i \leq K \qquad (1)$$

$$\sum_{i=1}^{K} \sum_{P \in \mathcal{P}_{s_i d_i}, e \in P} f_i(P) \leq 1 \qquad \forall e \in E \qquad (2)$$

where $f_i(P) \in \mathbb{R}_{\geq 0}$ represents the amount of traffic routed through path $P$ for $\omega_i$. A rate vector $\mathbf{R} = (R_i' : 1 \leq i \leq K) \in \mathbb{R}_{\geq 0}^K$ is achievable by routing scheme if there exist $f_i(P)$'s such that (1) and (2) are satisfied. The rate region of routing scheme, denoted by $\mathcal{R}_r$, is the set of all rate vectors achievable by routing scheme.

### C. Network Coding Scheme

A network coding scheme is defined as follows: [3]

**Definition 1.** An $(n, (\eta_e : e \in E), (R_i : 1 \leq i \leq K), (\delta_i : 1 \leq i \leq K))$ *network code* with block length $n$ is defined by:

1) for each $1 \leq i \leq K$ and $e \in \text{Out}(s_i)$, a local encoding function: $\phi_e : \mathcal{Y}_i \to \{1, \cdots, \eta_e\}$;
2) for each $v \in V - \{s_i, d_i : 1 \leq i \leq K\}$ and $e \in \text{Out}(v)$, a local encoding function: $\phi_e : \prod_{e' \in \text{In}(v)} \{1, \cdots, \eta_{e'}\} \to \{1, \cdots, \eta_e\}$;
3) for each $1 \leq i \leq K$, a decoding function: $\psi_i : \prod_{e' \in \text{In}(d_i)} \{1, \cdots, \eta_{e'}\} \to \mathcal{Y}_i$;
4) for each $1 \leq i \leq K$, the decoding error for $\omega_i$ is $\delta_i = Pr(\tilde{\psi}_i(Y_{1:K}) \neq Y_i)$, where $\tilde{\psi}_i(Y_{1:K})$ is the value of $\psi_i$ as a function of $Y_{1:K}$.

Given $e \in E$, let $U_e = \tilde{\phi}_e(Y_{1:K})$, where $\tilde{\phi}_e(Y_{1:K})$ is the value of $\phi_e$ as a function of $Y_{1:K}$, denote the random variable transmitted along $e$ in a network code. For a subset $C \subseteq E$, denote $U_C = \{U_e : e \in C\}$.

**Definition 2.** A rate vector $\mathbf{R} = (R_i' : 1 \leq i \leq K) \in \mathbb{R}_{\geq 0}^K$ is *achievable* by network coding if for any $\epsilon > 0$, there exists for sufficiently large $n$, an $(n, (\eta_e : e \in E), (R_i : 1 \leq i \leq K), (\delta_i : 1 \leq i \leq K))$ network code such that the following conditions are satisfied:

$$\frac{1}{n} \log \eta_e \leq 1 + \epsilon \qquad \forall e \in E \qquad (3)$$

$$R_i \geq R_i' - \epsilon \qquad \forall 1 \leq i \leq K \qquad (4)$$

$$\delta_i \leq \epsilon \qquad \forall 1 \leq i \leq K \qquad (5)$$

The *capacity region* achieved by network coding, denoted by $\mathcal{R}_{nc}$, is the set of all rate vectors $\mathbf{R}$ achievable by network coding.

Given a network code that satisfies (3)-(5), the following inequalities must hold:

$$\frac{1}{n} H(U_e) \leq \frac{1}{n} \log(\eta_e) \leq 1 + \epsilon \qquad \forall e \in E \qquad (6)$$

$$\frac{1}{n} H(Y_i) = \frac{1}{n} \log(\lceil 2^{nR_i} \rceil) \geq R_i \geq R_i' - \epsilon \quad \forall 1 \leq i \leq K \quad (7)$$

$$\frac{1}{n} I(Y_i; U_{\text{In}(d_i)}) \geq (1 - \epsilon)(R_i' - \epsilon) - \frac{1}{n} \qquad \forall 1 \leq i \leq K \quad (8)$$

where (8) is due to Fano's Inequality:

$$\frac{1}{n} I(Y_i; U_{\text{In}(d_i)}) \geq \frac{1}{n}(H(Y_i) - \delta_i \log |\mathcal{Y}_i| - 1)$$

$$= \frac{1}{n}(1 - \delta_i) H(Y_i) - \frac{1}{n} \geq (1 - \epsilon)(R_i' - \epsilon) - \frac{1}{n}$$

### D. Routing-Optimal Networks

Since routing scheme is a special case of network coding, $\mathcal{R}_r \subseteq \mathcal{R}_{nc}$.

**Definition 3.** A network is said to be *routing-optimal*, if $\mathcal{R}_{nc} = \mathcal{R}_r$, *i.e.*, for such network, routing is sufficient to achieve the whole rate region of network coding.

## III. A CLASS OF ROUTING-OPTIMAL NETWORKS

In this section, we present a class of routing-optimal networks, called *information-distributive* networks. We first use examples to illustrate the topological features of these networks, and show why they make the networks routing-optimal. Then, we define these networks more rigorously.

### A. Illustrative Examples

**Example 1.** We start with the simplest case of single unicast. It is well known that for this case, a network is always routing-optimal [5]. In this example, we re-investigate this case from a new perspective in order to highlight some of the important features that make it routing optimal. Let $m = \text{mincut}(s_1, d_1, G)$, and $C = \{e_1, \cdots, e_m\}$ is a cut-set between $s_1$ and $d_1$. Assume $R_1' \in \mathcal{R}_{nc}$. Therefore, for $\epsilon = \frac{1}{k} > 0$ ($k \in \mathbb{Z}_{>0}$), there exists a network code such that (3)-(5) are satisfied. In the followings, all the random variables are defined in this network code.

One important feature of this network is that each path from $s_1$ to $d_1$ must pass through at least an edge in $C$. Thus, $U_{\text{In}(d_1)}$ is a function of $U_C$. The following inequality holds:

$$I(Y_1; U_{\text{In}(d_1)}) \leq I(Y_1; U_C) \qquad (9)$$

The following equation holds:

$$I(Y_1; U_C) = \sum_{j=1}^{m} I(Y_1; U_{e_j} | U_{\{e_1, \cdots, e_{j-1}\}}) \qquad (10)$$

Intuitively, we can interpret (10) as follows: $I(Y_1; U_{e_1})$ is the amount of information about $Y_1$ that can be obtained from $U_{e_1}$, $I(Y_1; U_{e_2} | U_{e_1})$ the amount of information about $Y_1$ that can be obtained from $U_{e_2}$, excluding those already obtained from $U_{e_1}$, and so on. Hence, (10) can be seen as a "distribution" of the source information over the edges in $C$. Moreover, for each $1 \leq j \leq m$, we have:

$$I(Y_1; U_{e_j} | U_{\{e_1, \cdots, e_{j-1}\}}) \leq H(U_{e_j}) \qquad (11)$$

Another important feature is that due to Menger's Theorem, there exist $m$ edge-disjoint paths, $P_1, \cdots, P_m$, from $s_1$ to $d_1$ such that $e_j \in P_j$ for $1 \leq j \leq m$. Due to this feature, we can

(a) 2 unicast sessions      (b) 3 unicast sessions
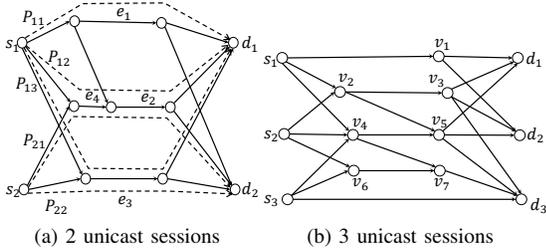
Fig. 1. Examples of information-distributive networks, where $s_i, d_i$ ($1 \leq i \leq 3$) are the source and the sink of the $i$th unicast session respectively.

construct a routing scheme by simply letting each $P_j$ transmit the information distributed on $e_j$:

$$f^{n,k}(P) = \begin{cases} \frac{1}{n}I(Y_1; U_{e_j}|U_{\{e_1,\cdots,e_{j-1}\}}) & \text{if } P = P_j, 1 \leq j \leq m \\ 0 & \text{otherwise.} \end{cases} \tag{12}$$

Clearly, due to (6) and (11), the above routing scheme satisfies the following inequalities:

$$f^{n,k}(P_j) \leq \frac{1}{n}H(U_{e_j}) \leq 1 + \frac{1}{k} \tag{13}$$

Moreover, due to (8)-(10), we have:

$$\sum_{P \in \mathcal{P}_{s_1 d_1}} f^{n,k}(P) = \sum_{j=1}^{m} f^{n,k}(P_j) = \frac{1}{n}I(Y_1; U_C)$$
$$\geq \frac{1}{n}I(Y_1; U_{\text{In}(d_1)}) \geq \left(1 - \frac{1}{k}\right)\left(R_i' - \frac{1}{k}\right) - \frac{1}{n} \tag{14}$$

Since $f^{n,k}(P_j)$ have an upper bound (see (13)), there exists a sub-sequence $(n_l, k_l)_{l=1}^{\infty}$ such that each sequence $(f^{n_l,k_l}(P_j))_{l=1}^{\infty}$ approaches a finite limit. Define the following routing scheme:

$$f_1(P) = \begin{cases} \lim_{l\to\infty} f^{n_l,k_l}(P) & \text{if } P = P_j (1 \leq j \leq m); \\ 0 & \text{otherwise.} \end{cases}$$

Due to (13) and (14), the above routing scheme satisfies (1) and (2). Hence, $R_1' \in \mathcal{R}_r$, which implies $\mathcal{R}_{nc} \subseteq \mathcal{R}_r$. Therefore, the network is routing-optimal. ∎

As shown above, two features are essential in making a network with single-unicast routing-optimal. The first feature is the existence of a cut-set such that each path from the source to the sink must pass through an edge in the cut-set. Due to this feature, the source information contained in $U_{\text{In}(d_1)}$ can be completely obtained from the messages transmitted through the cut-set $C$ (see (9)). The second feature is the existence of edge-disjoint paths $P_1, \cdots, P_m$, each of which passes through exactly one edge in $C$. Due to this feature, a routing scheme can be constructed such that the traffic transmitted along the paths $P_1, \cdots, P_m$ is exactly the information distributed on the edges in $C$ (see (12)). These two features together guarantee that the routing scheme achieves the same rate as network coding (see (13), (14)).

However, extending these features to multiple unicast sessions is not straightforward. One difference from single unicast is that $U_{\text{In}(d_i)}$ may not be a function of $U_C$, where $C$ is a cut-set between $s_i$ and $d_i$, and thus (9) might not hold. Another difference is that the information from multiple unicast sessions might be distributed on an edge, and thus (11) might not hold. Moreover, the paths for multiple unicast sesssions might overlap with each other, and thus (13) might not hold. These differences suggest that the cut-sets and the paths, over which

a routing scheme is to be constructed, should have additional features in order for the resulting routing scheme to achieve the same rate vector as network coding. We use an example to illustrate some of these features.

**Example 2.** Consider the network shown in Fig. 1aa. Consider an arbitrary rate vector $\mathbf{R} = (R_1', R_2') \in \mathcal{R}_{nc}$. Therefore, for $\epsilon = \frac{1}{k}$ ($k \in \mathbb{Z}_{>0}$), there exists a network code that satisfies (3)-(5). In the sequel, all the random variables are defined in this network code.

For $\omega_1$, we choose a cut-set $C_1 = \{e_1, e_2, e_3\}$ between $s_1$ and $d_1$, and a set of paths $\mathcal{P}_1 = \{P_{11}, P_{12}, P_{13}\}$ that pass through $e_1, e_2, e_3$ respectively; for $\omega_2$, we choose a cut-set $C_2 = \{e_2, e_3\}$ between $s_2$ and $d_2$, and a set of paths $\mathcal{P}_2 = \{P_{21}, P_{22}\}$ that pass through $e_2, e_3$ respectively.

We first investigate $C_1, C_2$. One important feature is that each path from $s_2$ to $d_1$ passes through at least an edge in $C_1$. Thus, $C_1$ is also a cut-set between $\{s_1, s_2\}$ and $d_1$, and $U_{\text{In}(d_1)}$ is a function of $U_{C_1}$. Hence, we have:

$$I(Y_1; U_{\text{In}(d_1)}) \leq I(Y_1; U_{C_1}) \tag{15}$$

Moreover, $\text{Out}(s_1) \cup C_2$ is a cut-set between $\{s_1, s_2\}$ and $d_2$, and $U_{\text{Out}(s_1)}$ is a function of $Y_1$. Hence $U_{\text{In}(d_2)}$ is a function of $Y_1, U_{C_2}$, which implies:

$$I(Y_2; U_{\text{In}(d_2)}|Y_1) \leq I(Y_2; U_{C_2}|Y_1) \tag{16}$$

We distribute the source information over $C_1, C_2$ as follows:

$$I(Y_1; U_{C_1}) = I(Y_1; U_{e_1}) + I(Y_1; U_{e_2}|U_{e_1}) + I(Y_1; U_{e_3}|U_{\{e_1,e_2\}}) \tag{17}$$

$$I(Y_2; U_{C_2}|Y_1) = I(Y_2; U_{e_2}|Y_1) + I(Y_2; U_{e_3}|Y_1, U_{e_2})$$

Another feature about $C_1, C_2$ is that edge $e_1$ is connected to only one source $s_1$, and thus $U_{e_1}$ is a function of $Y_1$. As shown below, this feature guarantees that the information distributed on an edge $e \in C_1 \cup C_2$ is completely contained in $U_e$. First, for $e_1$, it can be easily seen that:

$$I(Y_1; U_{e_1}) \leq H(U_{e_1}) \tag{18}$$

For $e_2$, we have:

$$I(Y_1; U_{e_2}|U_{e_1}) + I(Y_2; U_{e_2}|Y_1)$$
$$\stackrel{(b)}{=} I(Y_1; U_{e_2}|U_{e_1}) + I(Y_2; U_{e_2}|Y_1, U_{e_1}) \tag{19}$$
$$= I(Y_1, Y_2; U_{e_2}|U_{e_1}) \leq H(U_{e_2})$$

where $(b)$ is due to the fact that $U_{e_1}$ is a function of $Y_1$, and thus, $I(Y_2; U_{e_2}|Y_1) = I(Y_2; U_{e_2}|Y_1, U_{e_1})$. Similarly, for $e_3$, we have:

$$I(Y_1; U_{e_3}|U_{\{e_1,e_2\}}) + I(Y_2; U_{e_3}|Y_1, U_{e_2})$$
$$\stackrel{(c)}{=} I(Y_1; U_{e_3}|U_{\{e_1,e_2\}}) + I(Y_2; U_{e_3}|Y_1, U_{\{e_1,e_2\}}) \tag{20}$$
$$= I(Y_1, Y_2; U_{e_3}|U_{\{e_1,e_2\}}) \leq H(U_{e_3})$$

where $(c)$ is again due to the fact that $U_{e_1}$ is a function of $Y_1$.

Next, we investigate $\mathcal{P}_1, \mathcal{P}_2$. One important feature is that if $P \in \mathcal{P}_1$ overlaps with $P' \in \mathcal{P}_2$, $P \cap C_1 = P' \cap C_2$. For example, $P_{12}$ overlaps with $P_{21}$, and $P_{12} \cap C_1 = P_{21} \cap C_2 = \{e_2\}$. This feature ensures that the information distributed over $C_1, C_2$ can be further distributed over the paths in $\mathcal{P}_1, \mathcal{P}_2$. To see this, we construct the following routing scheme:

$$f_1^{n,k}(P) = \begin{cases} \frac{1}{n}I(Y_1; U_{e_j}|U_{\{e_1,\cdots,e_{j-1}\}}) & \text{if } P = P_{1j}, 1 \leq j \leq 3 \\ 0 & \text{otherwise.} \end{cases}$$

$$f_2^{n,k}(P) = \begin{cases} \frac{1}{n}I(Y_2;U_{e_2}|Y_1) & \text{if } P = P_{21}; \\ \frac{1}{n}I(Y_2;U_{e_3}|Y_1,U_{e_2}) & \text{if } P = P_{22}; \\ 0 & \text{otherwise.} \end{cases}$$

Due to (18)-(20), we can derive that for each $e \in C_1 \cup C_2$,

$$\sum_{i=1}^{2}\sum_{P \in \mathcal{P}_{s_i d_i}, e \in P} f_i^{n,k}(P) \leq \frac{1}{n}H(U_e) \leq 1 + \frac{1}{k} \qquad (21)$$

For $e_4$, we have:

$$\sum_{i=1}^{2}\sum_{P \in \mathcal{P}_{s_i d_i}, e_4 \in P} f_i^{n,k}(P)$$

$$= f_1^{n,k}(P_{12}) + f_2^{n,k}(P_{21}) \leq \frac{1}{n}H(U_{e_2}) \leq 1 + \frac{1}{k}$$

Likewise, we can prove that (21) holds for all the other edges of the paths in $\mathcal{P}_1 \cup \mathcal{P}_2$. Due to (15)-(17), the following inequalities hold for $i = 1, 2$

$$\sum_{P \in \mathcal{P}_{s_i d_i}} f_i^{n,k}(P) \geq \left(1 - \frac{1}{k}\right)\left(R_i' - \frac{1}{n}\right) + \frac{1}{n} \qquad (22)$$

By (21), there exists a sub-sequence $(n_l, k_l)_{l=1}^{\infty}$ such that for all $P \in \mathcal{P}_1 \cup \mathcal{P}_2$ and $i = 1, 2$, the sub-sequence $(f_i^{n_l, k_l}(P))_{l=1}^{\infty}$ approaches a finite limit. Define a routing scheme:

$$f_i(P) = \begin{cases} \lim_{l \to \infty} f_i^{n_l, k_l}(P) & \text{if } P \in \mathcal{P}_i, i = 1, 2; \\ 0 & \text{otherwise.} \end{cases} \qquad (23)$$

Due to (21) and (22), $f_i(P)$ satisfies (1) and (2). Hence, $\mathbf{R} \in \mathcal{R}_r$, and $\mathcal{R}_{nc} \subseteq \mathcal{R}_r$. The network is routing-optimal. ∎

### B. Information Distributive Networks

In this subsection, we present the definition of information-distributive networks. Similarly to single unicast, for each unicast session $\omega_i$ ($1 \leq i \leq K$), we choose a cut-set $C_i$ between $s_i$ and $d_i$ such that $|C_i| = \text{mincut}(s_i, d_i, G_i)$, and a set of paths $\mathcal{P}_i$ from $s_i$ to $d_i$. The collection of these cut-sets, denoted by $\mathcal{W} = (C_i)_{i=1}^{K}$, is called a *cut-set sequence*, and the collection of these path-sets, denoted by $\mathcal{K} = (\mathcal{P}_i)_{i=1}^{K}$, is called a *path-set sequence*. For instance, in Example 2, we choose a cut-set sequence $\mathcal{W} = (C_i)_{i=1}^{2}$, where $C_1 = \{e_1, e_2, e_3\}$ is a cut-set between $s_1$ and $d_1$, and $C_2 = \{e_2, e_3\}$ is a cut-set between $s_2$ and $d_2$, and a path-set sequence $\mathcal{K} = (\mathcal{P}_i)_{i=1}^{2}$, where $\mathcal{P}_1$ is a path-set from $s_1$ to $d_1$, and $\mathcal{P}_2$ a path-set from $s_2$ to $d_2$. Moreover, we arrange the edges in each cut-set in $\mathcal{W}$ in some ordering. For instance, in Example 2, we arrange the edges in $C_1$ in the ordering $T_1 = (e_1, e_2, e_3)$, and the edges in $C_2$ in the ordering $T_2 = (e_2, e_3)$. Each such ordering is called a permutation of the edges in the corresponding cut-set. The collection of these permutations, denoted $\mathcal{T} = (T_i)_{i=1}^{K}$, is called a *permutation sequence*. For $e \in C_i$, let $T_i(e)$ denote the subset of edges before $e$ in $T_i$. For $e \in E$, define $\mathcal{W}(e) = \{C_i \in \mathcal{W} : e \in C_i\}$, and $\alpha(e)$ the largest index of the source to which $\text{tail}(e)$ is connected. The first feature is described below.

Next, we formalize the three features we have shown in Example 2. The first feature is described below.

**Definition 4.** Given a cut-set sequence $\mathcal{W}$, if for all $1 \leq i < j \leq K$, each path from $s_j$ to $d_i$ must pass through an edge in $C_i$, we say that $\mathcal{W}$ is *cumulative*.

This feature guarantees that the source information con-

tained in the incoming messages at each sink $d_i$ can be completely obtained from $Y_{1:i-1}, U_{C_i}$.

**Lemma 1.** Consider a network code as defined in Definition 1. If $\mathcal{W}$ is a cumulative cut-set sequence, then for each $1 \leq i \leq K$, $Y_i$ is a function of $Y_{1:i-1}, U_{C_i}$, and the following inequality holds:

$$I(Y_i; U_{\text{In}(d_i)}|Y_{1:i-1}) \leq I(Y_i; U_{C_i}|Y_{1:i-1}) \qquad (24)$$

*Proof:* See Appendix B. ∎

Given a cumulative cut-set sequence $\mathcal{W}$ and a permutation sequence $\mathcal{T}$ for $\mathcal{W}$, we can distribute the source information $Y_i$ over the edges in $C_i$ as follows:

$$I(Y_i; U_{C_i}|Y_{1:i-1}) = \sum_{e \in C_i} I(Y_i; U_e|Y_{1:i-1}, U_{T_i(e)}) \qquad (25)$$

The second feature is presented below. Without loss of generality, let $\mathcal{W}(e) = \{C_{n_1}, \cdots, C_{n_k}\}$, where $1 \leq n_1 < \cdots < n_k \leq K$.

**Definition 5.** Given a cut-set sequence $\mathcal{W}$, we say that it is *distributive* if there exists a permutation sequence $\mathcal{T}$ for $\mathcal{W}$ such that for each $e \in \bigcup_{i=1}^{K} C_i$, the following conditions are satisfied: for all $1 \leq j < k$,

$$\alpha(e') \leq n_k \qquad \forall e' \in T_{n_{j+1}}(e) - T_{n_j}(e) \qquad (26)$$

$$\alpha(e') \leq n_{j+1} - 1 \qquad \forall e' \in T_{n_j}(e) - T_{n_{j+1}}(e) \qquad (27)$$

As shown in Example 2, let $T_1 = (e_1, e_2, e_3)$, and $T_2 = (e_2, e_3)$. For $e_3$, $\mathcal{W}(e_3) = \{C_1, C_2\}$, $T_2(e_3) - T_1(e_3) = \emptyset$, and thus, (26) is trivially satisfied; $T_1(e_3) - T_2(e_3) = \{e_1\}$, $\alpha(e_1) = 1$, and (27) is satisfied. Similarly, we can verify other edges. Hence, $\mathcal{W}$ is distributive.

The above two features ensure that the information from multiple unicast sessions that is distributed on an edge $e \in \bigcup_{i=1}^{K} C_i$ can be completely obtained from $U_e$.

**Lemma 2.** Consider a network code as defined in Definition 1. Given a cumulative cut-set sequence $\mathcal{W}$, if $\mathcal{W}$ is distributive, for each $e \in \bigcup_{i=1}^{K} C_i$, the following inequality holds:

$$\sum_{1 \leq i \leq K, e \in C_i} I(Y_i; U_e|Y_{1:i-1}, U_{T_i(e)}) \leq H(U_e) \qquad (28)$$

*Proof:* See Appendix B. ∎

The third feature is presented below.

**Definition 6.** Given a path-set sequence $\mathcal{K}$ for $\mathcal{W}$, we say that $\mathcal{K}$ is *extendable*, if for all $1 \leq i < j \leq K$, $P_1 \in \mathcal{P}_i$ and $P_2 \in \mathcal{P}_j$ such that $P_1$ overlaps with $P_2$, $P_1 \cap C_i = P_2 \cap C_j$.

As shown in Example 2, let $\mathcal{K} = \{\mathcal{P}_1, \mathcal{P}_2\}$. Clearly, we have $P_{12} \cap P_{21} = \{e_2, e_4\}$, $P_{13} \cap C_1 = P_{21} \cap C_2 = \{e_2\}$, and $P_{13} \cap P_{22} = \{e_3\}$, $P_{13} \cap C_1 = P_{22} \cap C_2 = \{e_3\}$. Thus, $\mathcal{K}$ is extendable.

**Definition 7.** A network with multiple unicast sessions is said to be *information-distributive*, if there exist a cumulative and distributive cut-set sequence $\mathcal{W}$, and an extendable path-set sequence $\mathcal{K}$ for $\mathcal{W}$ in the network.

As shown in the next theorem, the three features together guarantee that the network is routing-optimal.

**Theorem 1.** If a network is information-distributive, it is routing-optimal.

*Proof:* See Appendix B. ∎

**Example 3.** Consider the network shown in Fig. 1b. Define the following cut-sets:

$$C_1 = \{(s, v_1), (v_2, v_3), (v_4, v_5)\}$$
$$C_2 = \{(v_2, v_3), (v_4, v_5)\}$$

$$C_3 = \{(v_6, v_7), (s_3, d_3)\}$$

Define $\mathcal{W} = (C_i)_{i=1}^3$. Define the following paths:

$$P_{11} = \{(s_1, v_1), (v_1, d_1)\}$$
$$P_{12} = \{(s_1, v_2), (v_2, v_3), (v_3, d_1)\}$$
$$P_{13} = \{(s_1, v_4), (v_4, v_5), (v_5, d_1)\}$$
$$P_{21} = \{(s_2, v_2), (v_2, v_3), (v_3, d_2)\}$$
$$P_{31} = \{(s_3, v_6), (v_6, v_7), (v_7, d_3)\}$$
$$P_{33} = \{(s_3, d_3)\}$$

Define $\mathcal{K} = \{\{P_{11}, P_{12}, P_{13}\}, \{P_{21}, P_{22}\}, \{P_{31}, P_{32}\}\}$. It can be verified that $\mathcal{W}$ is cumulative and distributive, and $\mathcal{K}$ is extendable. The network is information-distributive. ∎

## IV. More Examples

### A. Index Coding

We consider a multiple-unicast version of index coding problem [4]. In this problem, there are $K$ terminals $t_1, \cdots, t_K$, a broadcast station $s$, and $K$ source messages $X_1, \cdots, X_K$, all available at $s$. All $X_i$'s are mutually independent random variables uniformly distributed over alphabet $\mathcal{X}_i = \{1, \cdots, 2^m\}$. Each terminal requires $X_i$, and has acquired a subset of source messages $\mathcal{H}_i$ such that $X_i \notin \mathcal{H}_i$. $s$ uses an encoding function $\phi : \prod_{i=1}^K \mathcal{X}_i \to \{1, \cdots, 2^l\}$ to encode the source messages, and broadcasts the encoded message to the terminals through an error-free broadcast channel. Each $t_i$ uses a decoding function $\psi_i$ to decode $X_i$ by using the received message and the messages in $\mathcal{H}_i$. The encoding function $\phi$ and the decoding functions $\psi_i$'s are collectively called an index code, and $l$ is the length of this index code. The minimum length of an index code is denoted by $l_{min}$.

This index coding problem can be cast to a multiple-unicast network coding problem over a network $G_1 = (V_1, E_1)$, where $V_1 = \{s_i, d_i : 1 \leq i \leq K\} \cup \{u, v\}$, $E_1 = \{(s_i, u), (v, d_i) : 1 \leq i \leq K\} \cup \{(u, v)\} \cup \{(s_j, d_i) : X_j \in \mathcal{H}_i\}$. The $K$ unicast sessions are $(s_1, d_1), \cdots, (s_K, d_K)$. It can be verified that there exists an index code of length $l$, if and only if $\mathbf{R} = (\frac{l}{m}, \cdots, \frac{l}{m})$ is achievable by network coding in $G_1$.

Let $C_i = \{(u, v)\}$, $P_i = \{(s_i, u), (u, v), (v, d_i)\}$. Define $\mathcal{W} = (C_i)_{i=1}^K$ and $\mathcal{K} = (\mathcal{P}_i)_{i=1}^K$, where $\mathcal{P}_i = \{P_i\}$. Since each $C_i$ contains only one edge, $\mathcal{W}$ is distributive. Meanwhile, since all $P_i$'s overlap at $(u, v)$, $\mathcal{K}$ is extendable.

The following theorem states that if the optimal solution to the index coding problem is to let the broadcast station transmit raw packet, *i.e.*, no coding is needed, then the corresponding multiple-unicast network is information-distributive, and the converse is also true.

**Theorem 2.** $l_{min} = mK$ if and only if $\mathcal{W}$ is cumulative, *i.e.*, $G_1$ is information-distributive.

*Proof:* See Appendix C. ∎

**Example 4.** In Fig. 2, we show an example of $G_1$, which corresponds to an index coding problem defined by: $\mathcal{H}_1 = \emptyset$, $\mathcal{H}_2 = \{X_1\}$, $\mathcal{H}_3 = \{X_1, X_2\}$, and $\mathcal{H}_4 = \{X_2, X_3\}$. Clearly, $\mathcal{W}$ is cumulative, and thus $l_{min} = mK$. ∎

### B. Single Unicast with Hard Deadline Constraint

In this example, we consider the network coding problem for a single-unicast session $(s, d)$ over a network $G = (V, E)$,
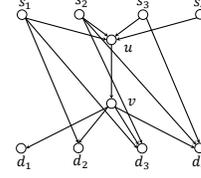


Fig. 2. The equivalent network coding problem for an index coding problem. The network is information-distributive, and thus no coding is needed in the index coding problem.

where each edge $e$ is associated with a delay $d_e \in \mathbb{Z}_{>0}$, and each node has a memory to hold received data. Given a directed path $P$, let $d(P) = \sum_{e \in P} d_e$ denote its delay. For $e \in E$, let $\delta(e)$ denote the minimum delay of directed paths from $s$ to $\text{tail}(e)$. The data transmission in the network proceeds in time slots. The messages transmitted from $s$ is represented by a sequence $(Y[t])_{t=0}^K$, where $Y[t]$ is a uniformly distributed random variable, and represents the message transmitted from $s$ at time slot $t$. All $Y[t]$'s are mutually independent. We require that each $Y[t]$ must be received by $d$ within $\tau$ time slots. Otherwise, it is regarded as useless, and is discarded. This problem was first proposed by [6] [7]. Recently, it has been shown that network coding can improve throughput by utilizing over-delayed information [8].

This problem can be cast to an equivalent network coding problem for multiple unicast sessions. We construct a time-extended graph $\tilde{G} = (\tilde{V}, \tilde{E})$ as follows: the node set is $\tilde{V} = \{s_t, d_t : 0 \leq t \leq K\} \cup \{v[t] : 0 \leq t \leq K + \tau\}$; for each $e = (u, v) \in E$ and $0 \leq t \leq K + \tau - d_e$, we add an edge $e[t] = (u[t], v[t + d_e])$ to $\tilde{E}$; for $u \in V$, and $0 \leq t \leq K + \tau - 1$, we add $M$ edges from $u[t]$ to $u[t + 1]$, where $M$ is the amount of memory available at $u$; for each $0 \leq t \leq K$, we add $J$ edges from $s_t$ to $s[t]$ and $J$ edges from $d[t + \tau]$ to $d_t$, where $J$ is a sufficiently large integer. Thus, the original single unicast session $(s, d)$ is cast to $K + 1$ unicast sessions $(s_0, d_0), \cdots, (s_K, d_K)$ over $\tilde{G}$.

Let $\tilde{G}[t]$ denote the routing domain for $(s_t, d_t)$, and $m = \text{mincut}(s_0, d_0, \tilde{G}[0])$. It can be seen that each $\tilde{G}[t]$ is simply a time-shifted version of $\tilde{G}[0]$. Given a subset of edges $U \subseteq \tilde{E}$, define $U[t] = \{(u[k + t], v[l + t]) : (u[k], v[l]) \in U\}$. Let $C = \{e_j[t_j] : 1 \leq j \leq m\}$ be a cut-set between $s_0$ and $d_0$ such that $e_j \in E$ for $1 \leq j \leq m$, and $\mathcal{P} = \{P_j, \cdots, P_m\}$ a set of edge disjoint paths from $s_0$ to $d_0$ such that $e_j[t_j] \in P_j$ for $1 \leq j \leq m$. Let $\mathcal{P}[t] = \{P[t] : P \in \mathcal{P}\}$. We consider the cut-set sequence $\mathcal{W} = (C[t])_{t=0}^K$, and the path-set sequence $\mathcal{K} = (\mathcal{P}[t])_{t=0}^K$.

**Lemma 3.** $\mathcal{W}$ is cumulative.

*Proof:* See Appendix C. ∎

Given $U \subseteq \tilde{E}$, a *recurrent* sequence of $U$ is a sequence consisting of all the edges in $U$ that are time-shifted versions of the same edge. $C[0]$ is said to be *distributive* if there is a re-indexing of the edges in $C[0]$ such that for each recurrent sequence $(e_p[t_{n_j}])_{j=1}^k$ of $C[0]$, the following conditions are satisfied:

1) for each $1 < j \leq k$, if $e_q[t_q] \in C[0]$ lies before $e_p[t_{n_j}]$, and $e_q[t_q - t_{n_j} + t_{n_{j-1}}] \notin C[0]$, then $t_q - \delta(e_q) \leq t_{n_j} - t_{n_{j-1}} - 1$;

2) for each $1 \leq j < k$, if $e_q[t_q] \in C[0]$ lies before $e_p[t_{n_j}]$,
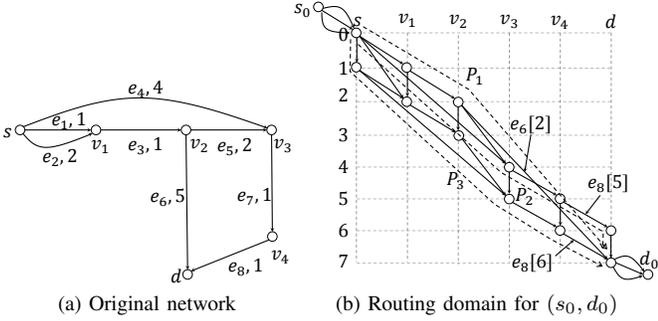
Fig. 3. An example of single unicast with deadline constraint $\tau = 7$. (a) shows an network with a single unicast $(s, d)$, where $e_k, i$ denotes the alias of an edge and its corresponding delay respectively. (b) shows the routing-domain between $s_0$ and $d_0$ over the corresponding time-extended graph $\tilde{G}$, where the node at coordinate $(v, t)$ is $v[t]$. In this routing-domain, $C[0] = \{e_8[5], e_6[2], e_8[6]\}$ is distributive, and $\mathcal{P} = \{P_1, P_2, P_3\}$ is extendable. Hence, $\tilde{G}$ is information-distributive, and therefore, routing-optimal.

and $e_q[t_q + t_{n_{j+1}} - t_{n_j}] \notin C[0]$, then $t_q - \delta(e_q) \leq t_{n_j} - t_{n_1}$.

$\mathcal{P}$ is said to be *extendable* if for all $P_i, P_j \in \mathcal{P}$ and $e[k], e[l] \in \tilde{E}$ such that $e[k] \in P_i$ and $e[l] \in P_j$, $e_i = e_j$ and $t_i - t_j = k - l$.

**Theorem 3.** If $C[0]$ is distributive, and $\mathcal{P}$ is extendable, $\tilde{G}$ is information-distributive, and thus is routing-optimal.

*Proof:* See Appendix C. ∎

**Example 5.** In Fig. 3a, we show an example of single unicast with delay constraint $\tau = 7$. In Fig. 3b, we show the routing domain $\tilde{G}[0]$ for $(s_0, d_0)$. Let $C[0] = \{e_8[5], e_6[2], e_8[6]\}$, and $\mathcal{P} = \{P_1, P_2, P_3\}$, where $P_1, P_2, P_3$ are marked as black dashed lines in Fig. 3b. It can be verified that $C[0]$ is distributive, and $\mathcal{P}$ is extendable. Thus, according to Theorem 3, $\tilde{G}$ is information-distributive. ∎

## V. THE CONVERSE IS NOT TRUE

Note that information-distributive networks don't subsume all possible routing-optimal networks. In the following, we show an example of such a network.

**Example 6.** Consider the network as shown in Fig. 4. We first show that it is not information-distributive. Define the following paths:

$$P_{11} = \{a_1, e_1, b_1\}, P_{12} = \{a_2, e_3, b_2\}$$
$$P_{21} = \{a_3, e_3, b_3\}, P_{22} = \{a_4, e_5, b_4\}$$
$$P_{31} = \{a_5, e_5, b_5\}, P_{32} = \{a_6, e_6, b_6\}$$

For $1 \leq i \leq 3$, let $\mathcal{P}_i = \{P_{i1}, P_{i2}\}$, and $\mathcal{K} = (\mathcal{P}_i)_{i=1}^3$. Since each source has only two outgoing edges, $\mathcal{K}$ is the only-possible path-set sequence. It can be verified that for all cumulative and distributive cut-set sequences, $\mathcal{K}$ is not extendable. For instance, let $C_1 = \{a_1, e_3\}$, $C_2 = \{e_3, b_4\}$, and $C_3 = \{e_5, b_6\}$. Clearly, the cut-set sequence $\mathcal{W} = (C_i)_{i=1}^3$ is cumulative and distributive. However, it can be seen that $P_{22}$ overlaps with $P_{31}$, but $P_{22} \cap C_2 = \{b_4\}$, and $P_{31} \cap C_3 = \{e_5\}$. Hence, $\mathcal{K}$ doesn't satisfy the condition of Definition 6. Similarly, we can verify other cases. Thus, the network is not information-distributive.

Nevertheless, we can show that the network is routing-optimal. Consider an arbitrary rate vector $\mathbf{R} = (R_1', R_2', R_3') \in \mathcal{R}_{nc}$. For $\epsilon = \frac{1}{k}$ ($k \geq 2$), there exists a network code of length $n$ such that (6)-(8) are satisfied.
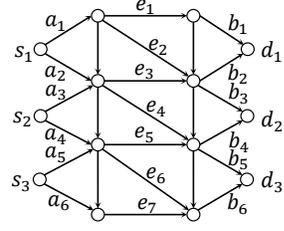
Fig. 4. A routing-optimal network that is not information-distributive.

Define the following cut-sets, and permutations of edges:

$$C_1 = \{e_1, e_2, e_3\} \; C_2 = \{e_3, e_4, e_5\} \; C_3 = \{e_5, e_6, e_7\}$$
$$T_1 = (e_1, e_2, e_3) \; T_2 = (e_3, e_4, e_5) \; T_3 = (e_5, e_6, e_7)$$

Define the following permutations:

$$T_1' = (b_1, b_2) \; T_2' = (b_3, b_4) \; T_3' = (b_5, b_6)$$

Let $\mathcal{W} = (C_i)_{i=1}^3$, and $\mathcal{T} = (T_i)_{i=1}^3$. Clearly, $\mathcal{W}$ satisfies the condition of Definition 4. Thus, according to Lemma 1, for $i = 1, 2, 3$, the following inequality holds:

$$I(Y_i; U_{\text{In}(d_i)} | Y_{1:i-1}) \leq I(U_{C_i}; Y_i | Y_{1:i-1}) \quad (29)$$

Moreover, since $\mathcal{T}$ satisfies the conditions of Definition 5. By Lemma 2, for $e \in \bigcup_{i=1}^3 C_i$, the following inequality holds:

$$\sum_{i=1}^3 \sum_{e \in C_i} I(Y_i; U_e | Y_{1:i-1}, U_{T_i(e)}) \leq H(U_e) \quad (30)$$

Define the following paths:

$$P_{11} = \{a_1, e_1, b_1\}, P_{12} = \{a_1, e_2, b_2\}, P_{13} = \{a_2, e_3, b_2\}$$
$$P_{21} = \{a_3, e_3, b_3\}, P_{22} = \{a_3, e_4, b_4\}, P_{23} = \{a_4, e_5, b_4\}$$
$$P_{31} = \{a_5, e_5, b_5\}, P_{32} = \{a_5, e_6, b_6\}, P_{33} = \{a_6, e_7, b_6\}$$

Let $\mathcal{P}_i = \{P_{i1}, P_{i2}, P_{i3}\}$. Define a routing scheme as follows:

$$f_i^{n,k}(P) = \begin{cases} \frac{1}{n} I(Y_i; U_{P \cap C_i} | Y_{1:i-1}, U_{T_i(e)}) & \text{if } P \in \mathcal{P}_i \\ 0 & \text{otherwise.} \end{cases}$$

Note the following inequalities hold for $i = 1, 2, 3$:

$$\frac{1}{n} H(Y_i) \geq \sum_{P \in \mathcal{P}_i} f_i^{n,k}(P)$$
$$= \frac{1}{n} I(Y_i; U_{C_i} | Y_{1:i-1}) \overset{(a)}{\geq} \frac{1}{n} I(Y_i; U_{\text{In}(d_i)} | Y_{1:i-1})$$
$$= \frac{1}{n} \sum_{e \in \text{In}(d_i)} I(Y_i; U_e | Y_{1:i-1}, U_{T_i'(e)}) \quad (31)$$
$$\overset{(b)}{\geq} \frac{1}{n} (1 - \frac{1}{k}) H(Y_i) - \frac{1}{n} \geq (1 - \frac{1}{k})(R_2' - \frac{1}{k}) - \frac{1}{n}$$

where $(a)$ holds because $U_{\text{In}(d_i)}$ is a function of $U_{C_2}, Y_{1:i-1}$; $(b)$ is due to Fano's Inequality. For $i = 1, 2, 3$, $e' \in \bigcup_{i=1}^3 C_i$, and $e \in \text{In}(d_i)$, define the following notations:

$$y_i^{n,k} = \frac{1}{n} H(Y_i) \quad u_{e'}^{n,k} = \frac{1}{n} H(U_{e'})$$
$$g_{i,e}^{n,k} = \frac{1}{n} I(Y_i; U_e | Y_{1:i-1}, U_{T_i'(e)})$$

Thus, (31) can be rewritten in a concise form as:

$$y_i^{n,k} \geq \sum_{P \in \mathcal{P}_i} f_i^{n,k}(P) \geq \sum_{e \in \text{In}(d_i)} g_{i,e}^{n,k}$$
$$\geq (1 - \frac{1}{k}) y_i^{n,k} - \frac{1}{n} \geq (1 - \frac{1}{k})(R_2' - \frac{1}{k}) - \frac{1}{n} \quad (32)$$

Due to (31), it can be seen that:

$$\frac{1}{2} y_i^{n,k} - 1 \leq (1 - \frac{1}{k}) y_i^{n,k} - \frac{1}{n}$$

$$\leq \frac{1}{n} I(Y_i; U_{\text{In}(d_i)} | Y_{1:i-1})$$

$$\leq \frac{1}{n} \sum_{e \in \text{In}(d_i)} H(U_e) \leq 2(1 + \frac{1}{k}) \leq 3$$

This means that $y_i^{n,k} \leq 8$. Clearly, all $y_i^{n,k}$'s, $u_{e'}^{n,k}$'s, $g_{i,e}^{n,k}$'s and $f_i^{n,k}(P)$'s have upper bounds. Thus, there exists a sub-sequence $(n_l, k_l)_{l=1}^{\infty}$ such that $y_i^{n_l,k_l}$, $u_{e'}^{n_l,k_l}$, $g_{i,e}^{n_l,k_l}$ and $f_i^{n_l,k_l}(P)$ approach finite limits when $l \to \infty$. Define the following notations:

$$y_i = \lim_{l \to \infty} y_i^{n_l,k_l} \quad u_{e'} = \lim_{l \to \infty} u_{e'}^{n_l,k_l} \quad g_{i,e} = \lim_{l \to \infty} g_{i,e}^{n_l,k_l}$$

Clearly, the following inequalities holds:

$$u_{e'} \leq 1 \quad g_{i,e} \leq u_e \leq 1$$

Define the following routing scheme:

$$f_i(P) = \begin{cases} \lim_{l \to \infty} f_i^{n_l,k_l}(P) & \text{if } P \in \mathcal{P}_i \\ 0 & \text{otherwise.} \end{cases}$$

We will prove that this routing scheme satisfies (1) and (2). According to (31), we see that $\sum_{P \in \mathcal{P}_i} f_i(P) \geq R'_2$, and thus, (1) is satisfied. Moreover, due to (30), (2) is satisfied for $e \in \bigcup_{i=1}^{3} C_i$. For $a_3$, we have:

$$f_2^{n,k}(P_{21}) + f_2^{n,k}(P_{22})$$
$$= \frac{1}{n} I(Y_2; U_{\{e_3,e_4\}} | Y_1) \overset{(c)}{\leq} \frac{1}{n} I(Y_2; U_{a_3} | Y_1)$$
$$\leq H(U_{a_3}) \leq 1 + \frac{1}{k}$$

where $(c)$ is due to the fact that $U_{\{e_3,e_4\}}$ is a function of $U_{a_3}, Y_1$. Thus, $f_2(P_{21}) + f_2(P_{22}) \leq 1$, and (2) is satisfied for $a_3$. Using similar arguments, we can prove that (2) is satisfied for $a_1, a_5$. Now consider $b_4$. Due to (32), the following equations hold:

$$y_2 = g_{2,b_3} + g_{2,b_4} = f_2(P_{21}) + f_2(P_{22}) + f_2(P_{23}) \quad (33)$$

Meanwhile, since $U_{b_3}$ is a function of $U_{e_3}, Y_1$, the following equations hold:

$$f_2^{n,k}(P_{21}) = \frac{1}{n} I(Y_2; U_{e_3} | Y_1) \geq \frac{1}{n} I(Y_2; U_{b_3} | Y_1) = g_{2,b_3}^{n,k}$$

Hence, $f_2(P_{21}) \geq g_{2,b_3}$. Combining with (33), we have:

$$f_2(P_{22}) + f_2(P_{23}) \leq g_{2,b_3} \leq 1$$

Hence, (2) holds for $b_3$. Similarly, we can prove that (2) holds for $b_2, b_6$. It can be easily seen that for all the other edges, (2) also holds. Therefore, we have proved that $\mathbf{R} \in \mathcal{R}_r$. This means that $\mathcal{R}_{nc} \subseteq \mathcal{R}_r$, and the network is routing-optimal. ∎

## VI. Conclusion

In this paper, we present a class of routing-optimal networks, called information-distributive networks, defined by three topological features. Due to these features, there is always a routing scheme that achieves the same rate vector as network coding such that the traffic transmitted through the network is the information distributed over the cut-sets between the sources and the sinks in the corresponding network coding scheme. We then present some examples of information-distributive networks related to index coding and single unicast with hard deadline constraint.

## Acknowledgment

## Appendix A
### Useful Tools

In this section, we present some useful tools to be used in the sequel.

**Proposition 1.** The following equations hold:

1) $H(X|Y) = H(X|Y, f(Y))$.
2) $I(X; Y|Z) = I(X; Y|Z, f(Z))$.
3) $H(X|f(Y)) \geq H(X|Y)$.
4) $I(X; Y|Z, W) \geq I(X; f(Y, Z)|Z, W)$.

*Proof:* 1) The following equation holds:

$$H(X, Y, f(Y)) = H(Y) + H(X|Y) + H(f(Y)|X, Y)$$
$$= H(Y) + H(X|Y) \quad (34)$$

Meanwhile, we have:

$$H(X, Y, f(Y)) = H(Y) + H(f(Y)|Y) + H(X|Y, f(Y))$$
$$= H(Y) + H(X|Y, f(Y)) \quad (35)$$

Combining Eq. (34) and Eq. (35), we have $H(X|Y) = H(X|Y, f(Y))$.

2) Due to 1), we can derive:

$$I(X; Y|Z, f(Z)) = H(X|Z, f(Z)) - H(X|Y, Z, f(Z))$$
$$= H(X|Z) - H(X|Y, Z)$$
$$= I(X; Y|Z)$$

3) First, the following equalities hold:

$$H(X, Y, f(Y)) = H(f(Y)) + H(X|f(Y)) + H(Y|X, f(Y)) \quad (36)$$

Combining Eq. (34) and Eq. (36), we then have:

$$H(X|f(Y)) = H(X|Y) + H(Y) - H(f(Y)) - H(Y|X, f(Y))$$
$$\overset{(a)}{=} H(X|Y) + H(Y|f(Y)) - H(Y|X, f(Y))$$
$$= H(X|Y) + I(X; Y|f(Y))$$
$$\geq H(X|Y)$$

where $(a)$ follows from the equation: $H(Y) = H(Y, f(Y)) = H(f(Y)) + H(Y|f(Y))$.

4) We have the following equations:

$$I(X; Y|Z, W) - I(X; f(Y, Z)|Z, W)$$
$$= H(X|Z, W) - H(X|Y, Z, W) -$$
$$[H(X|Z, W) - H(X|f(Y, Z), Z, W)]$$
$$= H(X|f(Y, Z), Z, W) - H(X|Y, Z, W) \geq 0$$

where the last inequality is due to 3) and the fact that $(f(Y, Z), Z, W)$ is a function of $(Y, Z, W)$. ∎

**Proposition 2.** If $Y \to (X, W) \to Z$, then $I(X; Y|W) \geq I(X; Y|W, Z)$ and $I(X; Y|W) \geq I(Z; Y|W)$. As a special case, we have $I(X; Y|W) \geq I(X; Y|W, f(X, W))$ and $I(X; Y|W) \geq I(f(X, W); Y|W)$.

*Proof:* We have the following equations:

$$I(X, Z; Y|W) = I(Z; Y|W) + I(X; Y|W, Z)$$
$$= I(X; Y|W) + I(Z; Y|X, W) = I(X; Y|W)$$

Thus, it must be that $I(X;Y|W) \geq I(X;Y|W,Z)$ and $I(X;Y|W) \geq I(Z;Y|W)$. Since the following chain: $Y \to (X,W) \to f(X,W)$ holds, we must have $I(X;Y|W) \geq I(X;Y|W, f(X,W))$ and $I(X;Y|W) \geq I(f(X,W);Y|W)$. ∎

## APPENDIX B
### PROOFS FOR INFORMATION-DISTRIBUTIVE NETWORKS

*Proof of Lemma 1:* Let $S'_i$ denote the set consisting of the outgoing edges of $s_1, \cdots, s_{i-1}$. Since each path from $s_j$ ($i \leq j < K$) to $d_i$ must pass through an edge in $C_i$, $S'_i \cup C_i$ forms a cut-set between $\{s_1, \cdots, s_K\}$ and $d_i$. Thus $U_{\text{In}(d_i)}$ is a function of $U_{S'_i}, U_{C_i}$. Meanwhile, $Y_{S'_i}$ is a function of $Y_{1:i-1}$. Thus, $U_{\text{In}(d_i)}$ is a function of $Y_{1:i-1}, U_{C_i}$. According to Proposition 2, (24) holds. ∎

*Proof of Lemma 2:* Let $\mathcal{T}$ be the permutation sequence as defined in Definition 5. Consider an arbitrary edge $e \in \bigcup_{i=1}^{K} C_i$. Without loss of generality, let $\mathcal{W}(e) = \{C_{n_1}, \cdots, C_{n_k}\}$, where $1 \leq n_1 < \cdots < n_k \leq K$. Then we have:

$$\sum_{1 \leq i \leq K, e \in C_i} I(Y_i; U_e | Y_{1:i-1}, U_{T_i(e)})$$
$$= \sum_{i=1}^{k} I(Y_i; U_e | Y_{1:n_i-1}, U_{T_{n_i}(e)})$$

For $k = 1$, the following equation holds:

$$\sum_{i=1}^{k} I(Y_i; U_e | Y_{1:n_i-1}, U_{T_{n_i}(e)})$$
$$= I(Y_i : U_e | Y_{1:n_1-1}, U_{T_{n_1}(e)}) \leq H(U_e)$$

Hence, (28) holds for $k = 1$. We now consider the case $k > 1$. We will prove the following inequality holds for $1 \leq p \leq k$:

$$\sum_{i=p}^{k} I(Y_i; U_e | Y_{1:n_i-1}, U_{T_{n_i}(e)}) \tag{37}$$
$$\leq I(Y_{n_p:n_k}; U_e | Y_{1:n_p-1}, U_{T_{n_p}(e)})$$

Clearly, (37) holds trivially for $p = k$. Assume it holds for $p > 1$. We will prove it also holds for $p - 1$.

$$\sum_{i=p-1}^{k} I(Y_i; U_e | Y_{1:n_i-1}, U_{T_{n_i}(e)})$$
$$\overset{(a)}{\leq} I(Y_{n_p:n_k}; U_e | Y_{1:n_p-1}, U_{T_{n_p}(e)}) +$$
$$I(Y_{n_{p-1}}; U_e | Y_{1:n_{p-1}-1}, U_{T_{n_{p-1}}(e)})$$
$$\overset{(b)}{=} I(Y_{n_p:n_k}; U_e | Y_{1:n_p-1}, U_{T_{n_p}(e) \cap T_{n_{p-1}}(e)}, U_{T_{n_p}(e) - T_{n_{p-1}}(e)})$$
$$+ I(Y_{n_{p-1}}; U_e | Y_{1:n_{p-1}-1}, U_{T_{n_p}(e) \cap T_{n_{p-1}}(e)}, U_{T_{n_{p-1}}(e) - T_{n_p}(e)})$$
$$\overset{(c)}{\leq} I(Y_{n_p:n_k}; U_e | Y_{1:n_p-1}, U_{T_{n_p}(e) \cap T_{n_{p-1}}(e)}) +$$
$$I(Y_{n_{p-1}}; U_e | Y_{1:n_{p-1}-1}, U_{T_{n_p}(e) \cap T_{n_{p-1}}(e)}, U_{T_{n_{p-1}}(e) - T_{n_p}(e)})$$
$$\overset{(d)}{=} I(Y_{n_p:n_k}; U_e | Y_{1:n_p-1}, U_{T_{n_p}(e) \cap T_{n_{p-1}}(e)}, U_{T_{n_p}(e) - T_{n_p}(e)})$$
$$+ I(Y_{n_{p-1}}; U_e | Y_{1:n_{p-1}-1}, U_{T_{n_p}(e) \cap T_{n_{p-1}}(e)}, U_{T_{n_{p-1}}(e) - T_{n_p}(e)})$$
$$\overset{(e)}{=} I(Y_{n_p:n_k}; U_e | Y_{1:n_p-1}, U_{T_{n_{p-1}}(e)}) +$$
$$I(Y_{n_{p-1}}; U_e | Y_{1:n_{p-1}-1}, U_{T_{n_{p-1}}(e)})$$
$$\leq I(Y_{n_p:n_k}; U_e | Y_{1:n_p-1}, U_{T_{n_{p-1}}(e)}) +$$

$$I(Y_{n_{p-1}:n_p-1}; U_e | Y_{1:n_{p-1}-1}, U_{T_{n_{p-1}}(e)})$$
$$\overset{(f)}{=} I(Y_{n_{p-1}:n_k}; U_e | Y_{1:n_{p-1}-1}, U_{T_{n_{p-1}}(e)})$$

where $(a)$ is due to our assumption that (37) holds for $p$; $(b)$ is due to the equalities, $T_{n_p}(e) = (T_{n_p}(e) \cap T_{n_{p-1}}(e)) \cup (T_{n_p}(e) - T_{n_{p-1}}(e))$ and $T_{n_{p-1}}(e) = (T_{n_p}(e) \cap T_{n_{p-1}}(e)) \cup (T_{n_{p-1}}(e) - T_{n_p}(e))$; $(c)$ is due to our premise that $\mathcal{W}$ is distributive: for each $e' \in T_{n_p}(e) - T_{n_{p-1}}(e)$, $\alpha(e') \leq n_k$, and thus $U_{e'}$ is a function of $Y_{1:n_k}$; therefore, according to Proposition 2, we have:

$$I(Y_{n_p:n_k}; U_e | Y_{1:n_p-1}, U_{T_{n_p}(e) \cap T_{n_{p-1}}(e)}, U_{T_{n_p}(e) - T_{n_{p-1}}(e)})$$
$$\leq I(Y_{n_p:n_k}; U_e | Y_{1:n_p-1}, U_{T_{n_p}(e) - T_{n_{p-1}}(e)})$$

$(d)$ is also due to our premise that $\mathcal{W}$ is distributive: for each $e' \in T_{n_{p-1}}(e) - T_{n_p}(e)$, $\alpha(e') \leq n_p - 1$, and thus $U_{e'}$ is a function of $Y_{1:n_p-1}$; therefore, the following equality holds according to Proposition 1:

$$I(Y_{n_{p-1}}; U_e | Y_{1:n_{p-1}-1}, U_{T_{n_p}(e) \cap T_{n_{p-1}}(e)}, U_{T_{n_{p-1}}(e) - T_{n_p}(e)})$$
$$= I(Y_{n_{p-1}}; U_e | Y_{1:n_{p-1}-1}, U_{T_{n_p}(e) \cap T_{n_{p-1}}(e)})$$

$(e)$ is again due to $T_{n_{p-1}}(e) = (T_{n_p}(e) \cap T_{n_{p-1}}(e)) \cup (T_{n_{p-1}}(e) - T_{n_p}(e))$; $(f)$ is due to chain rule of mutual information. Thus, (37) holds for $p - 1$. This means that (37) must hold for all $1 \leq p \leq k$. Letting $p = 1$ in (37), we have:

$$\sum_{i=1}^{k} I(Y_i; U_e | Y_{1:n_i-1}, U_{T_{n_i}(e)})$$
$$\leq I(Y_{n_1:n_k}; U_e | Y_{1:n_1-1}, U_{T_{n_1}(e)}) \leq H(U_e)$$

Thus, the lemma holds. ∎

Let $e$ be an edge that is passed through by at least one path in an extendable path-set sequence $\mathcal{K}$. According to the above definition, all the paths in $\mathcal{K}$ that pass through $e$ must pass through a single edge in $\mathcal{W}$. We use $\mu_e$ to denote this edge, and refer to it as the *representative* of $e$ in $\mathcal{W}$.

*Proof of Theorem 1:* Let $\mathcal{W} = \{C_i : 1 \leq i \leq K\}$ be a cumulative and distributive cut-set sequence, $\mathcal{T}$ a permutation sequence for $\mathcal{W}$ that satisfies the conditions of Definition 5, and $\mathcal{K} = \{\mathcal{P}_i : 1 \leq i \leq K\}$ an extendable path-set sequence for $\mathcal{W}$. Let $\mathbf{R} = (R'_i : 1 \leq i \leq K)$ be an arbitrary rate vector in $\mathcal{R}_{nc}$. Therefore, for $\epsilon = \frac{1}{k} > 0$ ($k \in \mathbb{Z}_{>0}$), there exists a network code which satisfies (3)-(5). In the rest of this proof, all the random variables are defined in this network code.

We then define the following routing scheme: for $1 \leq i \leq K$,

$$f_i^{n,k}(P) = \begin{cases} \frac{1}{n} I(Y_i; U_e | Y_{1:i-1}, U_{T_i(e)}) & \text{if } P \in \mathcal{P}_i, e \in P \cap C_i; \\ 0 & \text{otherwise.} \end{cases}$$

Since $\mathcal{W}$ is cumulative, the following equation holds:

$$\sum_{P \in \mathcal{P}_{s_i d_i}} f_i^{n,k}(P) = \sum_{P \in \mathcal{P}_i} f_i^{n,k}(P)$$
$$= \frac{1}{n} \sum_{e \in C_i} I(Y_i; U_e | Y_{1:i-1}, U_{T_i(e)}) \tag{38}$$
$$\overset{(a)}{=} \frac{1}{n} I(Y_i; U_{C_i} | Y_{1:i-1}) \overset{(b)}{\geq} \frac{1}{n} I(Y_i; U_{\text{In}(d_i)} | Y_{1:i-1})$$

where $(a)$ is due to (25), and $(b)$ is due to (24). Define $\delta'_i = Pr(Y_i$ cannot be decoded from $U_{\text{In}(d_i)}, Y_{1:i-1})$. Clearly, $\delta'_i \leq$

$\delta_i \leq \frac{1}{k}$. Then, we can derive the following equation:

$$\frac{1}{n} I(Y_i; U_{\text{In}(d_i)} | Y_{1:i-1})$$

$$= \frac{1}{n} (H(Y_i | Y_{1:i-1}) - H(Y_i | U_{\text{In}(d_i)}, Y_{1:i-1}))$$

$$\overset{(c)}{=} \frac{1}{n} (H(Y_i) - H(Y_i | U_{\text{In}(d_i)}, Y_{1:i-1}))$$

$$\overset{(d)}{\geq} \frac{1}{n} (H(Y_i) - 1 - \delta_i' \log |\mathcal{Y}_i|)$$

$$= (1 - \delta_i') \frac{1}{n} H(Y_i) - \frac{1}{n}$$

$$\overset{(e)}{\geq} \left(1 - \frac{1}{k}\right) \left(R_i' - \frac{1}{k}\right) - \frac{1}{n}$$

where $(c)$ is due to the fact that $Y_i$ is independent from $Y_{1:i-1}$; $(d)$ is due to Fano Inequality; $(e)$ is due to (7). Combining the above equation with (38), the following inequality holds:

$$\sum_{P \in \mathcal{P}_{s_i d_i}} f_i^{n,k}(P) \geq \left(1 - \frac{1}{k}\right) \left(R_i' - \frac{1}{k}\right) - \frac{1}{n} \quad (39)$$

Let $e$ be an edge that is passed through by at least one path in $\mathcal{K}$. Since $\mathcal{K}$ is extendable, the paths in $\mathcal{K}$ that pass through $e$ must pass through $e$'s representative $\mu_e$ in $\mathcal{W}$. Hence, the following equation holds:

$$\sum_{i=1}^{K} \sum_{P \in \mathcal{P}_{s_i d_i}, e \in P} f_i^{n,k}(P)$$

$$= \sum_{i=1}^{K} \sum_{P \in \mathcal{P}_i, e \in P} f_i^{n,k}(P)$$

$$\leq \sum_{i=1}^{K} \sum_{P \in \mathcal{P}_i, \mu_e \in P} f_i^{n,k}(P) \quad (40)$$

$$= \frac{1}{n} \sum_{1 \leq i \leq K, \mu_e \in C_i} I(Y_i; U_{\mu_e} | Y_{1:i-1}, U_{T_i(\mu_e)})$$

$$\overset{(f)}{\leq} \frac{1}{n} H(U_{\mu_e}) \overset{(g)}{\leq} 1 + \frac{1}{k}$$

where $(f)$ is due to (28); $(g)$ is due to (6).

Since each $f_i^{n,k}(P)$ has an upper bound, there exists a sequence $(n_l, k_l)_{l=1}^{\infty}$ such that for $1 \leq i \leq K$, the sequence $(f_i^{n_l, k_l}(P))_{l=1}^{\infty}$ approaches a finite limit. Define the following routing scheme:

$$f_i(P) = \begin{cases} \lim_{l \to \infty} f_i^{n_l, k_l}(P) & \text{if } P \in \mathcal{P}_i \\ 0 & \text{otherwise.} \end{cases}$$

Due to (39) and (40), $f_i(P)$ satisfies (1) and (2). Hence, $\mathbf{R} \in \mathcal{R}_r$. This implies that $\mathcal{R}_{nc} \subseteq \mathcal{R}_r$, and the network is routing-optimal. $\blacksquare$

## APPENDIX C
## PROOFS FOR EXAMPLES

*Proof of Theorem 2:* Assume $\mathcal{W}$ is cumulative. Hence, $G_1$ is information-distributive According to Theorem 1, $G_1$ is routing-optimal. Since routing can achieve a common rate of at most $\frac{1}{K}$, $l_{min} = mK$.

Now assume $l_{min} = mK$. We consider a side-information graph $G' = (V', E')$ [4], where $V' = \{1, \cdots, K\}$, and $E' = \{(j,i) : X_i \in \mathcal{H}_j, 1 \leq i, j \leq K\}$. It has been shown that if $l_{min} = mK$, then $G'$ is acyclic [4]. We will show that

$\mathcal{W}$ is information-distributive. Since $G'$ is acyclic, we can re-index the nodes in $G'$, such that if $(j,i) \in E'$, $j < i$. Let $1 \leq i < j \leq K$. Consider a path $P$ from $s_j$ to $d_i$. Since $(j,i) \notin E'$, $X_j \notin \mathcal{H}_i$. Thus, there is no directed edge from $s_j$ to $d_i$ in $G_1$, and $P$ must pass through $(u,v) \in C_i$. Hence, $\mathcal{W}$ is cumulative, and $G_1$ is information-distributive. $\blacksquare$

*Proof of Lemma 3:* Let $0 \leq i < j \leq K$. Assume there is a directed path $P$ from $s_j$ to $d_i$. Let $P_1$ be the part of $P$ after $s[j]$. Clearly, $P' = \{(s_i, s[i]), (s[i], s[i+1]), \cdots, (s[j-1], s[j])\} \cup P_1$ is a directed path from $s_i$ to $d_i$. Since $C[i]$ is a cut-set between $s[i]$ and $d[i]$, $P'$ must pass through an edge $e[k] \in C[i]$. Thus, $e[k] \in P$. This means that $\mathcal{W}$ is cumulative. $\blacksquare$

Since the duration between $e[t]$ and $s[\alpha(e[t])]$ is $\delta(e)$, we have:

$$\alpha(e[t]) = t - \delta(e) \quad (41)$$

**Lemma 4.** If $C[0]$ is distributive, $\mathcal{W}$ is distributive.

*Proof:* Let $T[t] = (e_i[t_i + t])_{i=1}^{k}$, and define a permutation sequence $\mathcal{T} = (T[t])_{t=0}^{K}$ for $\mathcal{W}$. We will prove that if $C[0]$ is distributive, $\mathcal{T}$ satisfies (26) and (27).

Consider an edge $e_p[t_p] \in C[0]$. Let $(e_p[t_{n_i}])_{i=1}^{k}$ be the recurrent sequence in $C[0]$, in which all the edges are time-shifted versions of $e_p$. Without loss of generality, let $n_j = p$. Next, consider $e_p[t_p + k] \in C[k]$. Let $\mathcal{W}(e_p[t_p + k]) = \{C[t] : e_p[t_p + k] \in C[t], 0 \leq t \leq K\}$ denote the subset of cut-sets which contain $e_p[t_p + k]$. Clearly, $C[k - t_{n_{j+1}} + t_{n_j}]$ and $C[k + t_{n_j} - t_{n_{j-1}}]$ are the cut-sets in $\mathcal{W}(e_p[t_p + k])$ that lies immediately before and after $C[k]$ respectively, and $C[k + t_{n_j} - t_{n_1}]$ is the last cut-set in $\mathcal{W}(e_p[t_p + k])$.

Consider an edge $e_q[t_q + k] \in C[k]$ be an edge that lies before $e_p[t_p + k]$ in $T[k]$, but doesn't appear before $e_p[t_p + k]$ in $T[k - t_{n_{j+1}} + t_{n_j}]$. This means that $e_q[t_q + t_{n_{j+1}} - t_{n_j}] \notin C[0]$. Thus, the following equation holds:

$$\alpha(e_q[t_q + k]) = k + t_q - \delta(e_q) \overset{(a)}{\leq} k + t_{n_j} - t_{n_1}.$$

where $(a)$ is due to the premise that $C[0]$ is distributive. Hence, (26) is satisfied.

Now assume that $e_q[t_q + k] \in C[k]$ lies before $e_p[t_p + k]$ in $T[k]$, but doesn't appear before $e_p[t_p + k]$ in $T[k - t_{n_j} + t_{n_{j-1}}]$. This implies that $e_q[t_q - t_{n_j} + t_{n_{j-1}}] \notin C[0]$. Thus, the following equation holds:

$$\alpha(e_q[t_q + k]) = k + t_q - \delta(e_q) \overset{(b)}{\leq} k + t_{n_j} - t_{n_{j-1}} - 1$$

where $(b)$ is again due to the premise that $C[0]$ is distributive. Hence, (27) is satisfied. $\mathcal{W}$ is distributive. $\blacksquare$

**Lemma 5.** If $\mathcal{P}$ is extendable, $\mathcal{K}$ is extendable.

*Proof:* Consider two paths $P_i, P_j \in \mathcal{P}$. Assume $P_i[k_1]$ overlaps with $P_j[k_2]$ at $e[t]$. Thus, $e[t - k_1] \in P_i$ and $e[t - k_2] \in P_j$. Since $\mathcal{P}$ is extendable, this means that $e_i = e_j$ and

$$t_i - t_j = t - k_1 - (t - k_2) = k_2 - k_1$$

Note that $e_i[t_i + k_1]$ is the edge in $\mathcal{W}$ that is passed through by $P_i[k_1]$. We have:

$$e_i[t_i + k_1] = e_j[t_j + k_2] \in P_j[k_2] \cap C[k_2].$$

Thus, $P_i[k_1]$ and $P_j[k_2]$ pass through the same edge $e_i[t_i + k_1]$ in $\mathcal{W}$. Hence, $\mathcal{K}$ is extendable. $\blacksquare$

*Proof of Theorem 3:* Due to Lemmas 5, 3 and 4, the theorem holds. $\blacksquare$

## REFERENCES

[1] R. Dougherty, C. Freiling, and K. Zeger, "Insufficiency of linear coding in network information flow," *IEEE Transactions on Information Theory*, vol. 51, no. 8, pp. 2745–2759, 2005.

[2] N. J. Harvey, R. Kleinberg, and A. R. Lehman, "On the capacity of information networks," *IEEE/ACM Transactions on Networking (TON)*, vol. 14, no. SI, pp. 2345–2364, 2006.

[3] X. Yan, R. W. Yeung, and Z. Zhang, "The capacity region for multi-source multi-sink network coding," in *the Proceedings of IEEE International Symposium on Information Theory*, Nice, France, June 2007, pp. 116–120.

[4] Z. Bar-Yossef, Y. Birk, T. Jayram, and T. Kol, "Index coding with side information," *IEEE Transactions on Information Theory*, vol. 57, no. 3, pp. 1479–1494, 2011.

[5] R. Ahlswede, N. Cai, S.-Y. Li, and R. W. Yeung, "Network information flow," *IEEE Transactions on Information Theory*, vol. 46, no. 4, pp. 1204–1216, 2000.

[6] M. Kodialam and T. Lakshman, "On allocating capacity in networks with path length constrained routing," in *the Proceedings of Allerton Conference on Communication, Control, and Computing*, Monticello, IL, U.S.A., Sept. 2002.

[7] M. Chen, personal communication, Feb. 2013.

[8] C. Wang and M. Chen, "Sending perishable information: Coding improves delay-constrained throughput even for single unicast," in *the Proceedings of IEEE International Symposium on Information Theory*, Hawaii, U.S.A., June 2014.