# Distributed Decode–Forward for Broadcast

Sung Hoon Lim
EPFL
1015 Lausanne, Switzerland
sung.lim@epfl.ch

Kwang Taik Kim
Samsung Electronics
Suwon, Gyeonggi-do 443-742, Korea
kwangtaik.kim@samsung.com

Young-Han Kim
University of California, San Diego
La Jolla, CA 92093, USA
yhk@ucsd.edu

***Abstract*—A new coding scheme for broadcasting multiple messages over a general relay network is presented. The proposed distributed decode–forward scheme combines Marton coding for single-hop broadcast channels and partial decode–forward for relay channels by Cover and El Gamal. For the $N$-node Gaussian broadcast relay network, the scheme achieves within $0.5N$ bits from the capacity region, extending and refining a recent result by Kannan, Raja, and Viswanath. The main idea of the scheme is to precode all the codewords initially at the source and to decode and forward parts of them on the fly at the relays.**

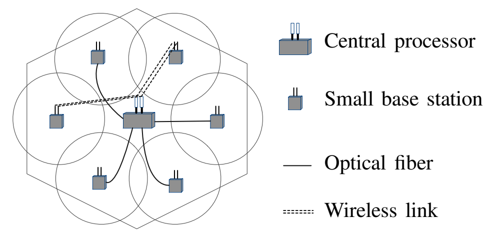


Fig. 1. A typical dense heterogeneous network with centralized multicell joint processing. The central processor encodes and decodes messages over multiple cells jointly for multihop transmission and reception, respectively.

## I. Introduction

Motivated by ever-increasing demands for higher data rates and broader coverage, future cellular systems are expected to deploy many small base stations. While such dense deployment provides the benefit of bringing the radio closer to end users, it also increases the amount of interference from neighboring cells. Consequently, several system architectures have been proposed [1], [2] to mitigate multicell interference, whereby multiple base stations are connected via optical or wireless links to a central processor that encodes and decodes messages over multiple cells jointly; see Figure 1 for a typical deployment scenario of a coordinated cellular network.

One of the main challenges in designing coordinated cellular networks is to develop optimal transmission schemes for *networked multiplexing*, namely, simultaneous transmission of multiple messages for both multihop uplink and downlink communications, which are multihop extensions of the traditional multiple access and broadcast channels in a single cell. What is the fundamental limit on the performance of uplink (multiple access) and downlink (broadcast) relay networks? Which coding scheme achieves this limit?

For multiple access relay networks, this problem is relatively well studied as a special instance of general multimessage multicast. In their award-winning paper [3], Avestimehr, Diggavi, and Tse developed the quantize–map–forward coding scheme that uniformly achieves the capacity for single-message multicast Gaussian networks within a constant gap. This coding scheme was further streamlined and extended by noisy network coding [4], [5], [6], the main features of which include a single-letter performance bound that can be applied to arbitrary multimessage multicast network models. In particular, noisy network coding achieves the capacity region for $N$-node Gaussian multiple access networks within $0.63N$ bits per dimension.

For broadcast relay networks, however, much less is known. The main difficulty for broadcast lies with the need for careful coordination among codewords for multiple messages. This is evident even with single-hop broadcast channels, for which the capacity region is not known in general. Recently, Kannan, Raja, and Viswanath [7] made a seminal contribution to the problem by showing that the capacity region for Gaussian broadcast relay networks can be achieved within a uniform constant number of bits from the cutset outer bound. The main idea behind their coding scheme is the "pruning" technique for coordinating superletter codewords, which was originally introduced by Anand and Kumar [8] for multicast.

This paper presents a single-letter coding scheme that extends the multiletter coding scheme by Kannan et al. [7] to arbitrary noisy networks, with the ultimate goal of establishing general design principles for broadcast relay networks. The development of an efficient and scalable coding scheme entails the following two considerations. First, among the three canonical relaying schemes, decode–forward [9], compress–forward [9], and amplify–forward [10], since the relay processing of compress–forward and amplify–forward inherit random noise components, coordination for broadcasting multiple codewords is not possible as the exact status of the relays is not available at the source. This leaves (some version of) decode–forward as the only building block that is feasible for coordination. Second, since the network model simplifies to the classical broadcast channel when the network is single-hop, the coding scheme must include Marton coding [11] for broadcast channels.

Guided by these two criteria, our coding scheme carefully combines Marton's multicoding scheme [12, Section 8.3.2] with a variant of partial decode–forward [13]. In particular, by jointly encoding all the messages with compatible codewords via multicoding, the source node coordinates and controls the transmission over the entire network, and the relays simply recover their desired messages as well as parts of other messages in a block-by-block manner. The details will be explained in Section III.

When applied to deterministic networks, the achievable rate region of this distributed decode–forward coding scheme simplifies to the same rate expression as the cutset outer bound [7, Eq. (9)] except for a difference in the form of input pmfs. For Gaussian networks, we show that our scheme achieves the capacity region within $0.5N$ bits per dimension, which provides the best known gap result in general. For the standard single-hop broadcast channel, the achievable rate region coincides with Marton's inner bound. In this sense, the proposed coding scheme is an extension of Marton coding to multihop networks.

Throughout the paper, we use the notation in [12]. In particular, a sequence of random variables with node index $k$ and time index $i \in [1:n] := \{1, \ldots, n\}$ is denoted as $X_k^n := (X_{k1}, \ldots, X_{kn})$. A tuple of random variables is denoted as $X(\mathcal{A}) := (X_k : k \in \mathcal{A})$.

## II. Problem Setup and Main Results

Consider the $N$-node discrete memoryless broadcast relay network (BRN) $p(y_1, \ldots, y_N | x_1, \ldots, x_N)$. The noise, interference, and broadcast effects in the network as well as the topology of the network (which nodes can communicate directly to which other nodes) are defined through the structure of the conditional pmf $p(y^N|x^N)$, namely, the probability that the output symbols $y_1, \ldots, y_N$ are received at nodes $1, \ldots, N$, respectively, when the input symbols $x_1, \ldots, x_N$ are transmitted from nodes $1, \ldots, N$, respectively.

Suppose that source node 1 wishes to communicate messages $M_2, \ldots, M_N$ to their respective destination nodes $2, \ldots, N$. The $((2^{nR_2}, \ldots, 2^{nR_N}), n)$ code for the BRN consists of

- message sets $[1:2^{nR_2}]$, …, $[1:2^{nR_N}]$,
- a source encoder that assigns a symbol $x_{1i}(m_2, \ldots, m_N, y_1^{i-1})$ to each message tuple $(m_2, \ldots, m_N) \in [1:2^{nR_2}] \times \cdots \times [1:2^{nR_N}]$ and received sequence $y_1^{i-1} \in \mathcal{Y}_1^{i-1}$ for $i \in [1:n]$,
- a set of relay encoders, where encoder $k \in [2:N]$ assigns $x_{ki}(y_k^{i-1})$ to each $y_k^{i-1}$ for $i \in [1:n]$, and
- a set of decoders, where decoder $k \in [2:N]$ assigns an estimate $\hat{m}_k$ or an error message $\mathrm{e}$ to each $y_k^n$.

The performance of the code is measured by the *average probability of error* $P_e^{(n)} = \mathsf{P}\{\hat{M}_k \neq M_k$ for some $k \in [2:N]\}$, where $(M_2, \ldots, M_N)$ be uniformly distributed over $[1:2^{nR_2}] \times \cdots \times [1:2^{nR_N}]$. A rate tuple $(R_2, \ldots, R_N)$ is said to be *achievable* if there exists a sequence of $((2^{nR_2}, \ldots, 2^{nR_N}), n)$ codes such that $\lim_{n\to\infty} P_e^{(n)} = 0$. The *capacity region* of the BRN is the closure of the set of achievable rate tuples $(R_2, \ldots, R_N)$.

We are ready to state the main result of the paper.

**Theorem 1.** *For the discrete memoryless broadcast relay network $p(y^N|x^N)$, a rate tuple $(R_2, \ldots, R_N)$ is achievable if*

$$R(\mathcal{S}^c) < I(X(\mathcal{S}); U(\mathcal{S}^c) | X(\mathcal{S}^c)) - \sum_{k \in \mathcal{S}^c} I(U_k; U(\mathcal{S}_k^c), X^N | X_k, Y_k) \quad (1)$$

*for all $\mathcal{S} \subseteq [1:N]$ such that $1 \in \mathcal{S}$ and $\mathcal{S}^c \neq \emptyset$ for some $(\prod_{k=2}^n p(x_k)) p(x_1, u_2^N | x_2^N)$, where $\mathcal{S}_k^c = \mathcal{S}^c \cap [2:k-1]$ and $R(\mathcal{S}^c) = \sum_{k \in \mathcal{S}^c} R_k$.*

The proof of Theorem 1 along with the description and analysis of the associated distributed decode–forward coding scheme is deferred to Section III. By setting some rates to zero (say, $R_k = 0$ if $k \notin \mathcal{D}$) in Theorem 1 and removing inactive inequalities, we can also establish the following inner bound on the capacity region when the message tuple $(M_k : k \in \mathcal{D})$ is broadcast to a set of destination nodes, $\mathcal{D} \subseteq [2:N]$.

**Theorem 2.** *For the discrete memoryless broadcast relay network $p(y^N|x^N)$ with destination set $\mathcal{D} \subseteq [2:N]$, a rate tuple $(R_k : k \in \mathcal{D})$ is achievable if*

$$R(\mathcal{T}) < \min_{\mathcal{S}: \mathcal{T} \subseteq \mathcal{S}^c} I(X(\mathcal{S}); U(\mathcal{S}^c) | X(\mathcal{S}^c)) - \sum_{k \in \mathcal{S}^c} I(U_k; U(\mathcal{S}_k^c), X^N | X_k, Y_k) \quad (2)$$

*for all $\mathcal{T} \subseteq \mathcal{D}$ for some $(\prod_{k=2}^n p(x_k)) p(x_1, u_2^N | x_2^N)$, where $\mathcal{S}_k^c = \mathcal{S}^c \cap [2:k-1]$ and $R(\mathcal{T}) = \sum_{k \in \mathcal{T}} R_k$.*

The capacity inner bound in Theorem 2 has a similar structure to the cutset bound [14, Theorem 15.10.1],

$$R(\mathcal{T}) \leq \min_{\mathcal{S}: \mathcal{T} \subseteq \mathcal{S}^c} I(X(\mathcal{S}); Y(\mathcal{S}^c) | X(\mathcal{S}^c)) \quad (3)$$

for all $\mathcal{T} \subseteq \mathcal{D}$ for some $p(x^N)$. The first term of (2), however, has the auxiliary random variables $U_j$ instead of $Y_j$, and there is a negative term that quantifies the cost of multicoding. In addition, the pmfs for (2) are of the form $(\prod_{k=2}^{n} p(x_k))p(x_1, u_2^N|x_2^N)$ rather than the full joint form.

We illustrate the utility of Theorem 1 via three canonical examples.

**Example 1** (Deterministic networks)**.** Suppose $Y_k = g_k(X_1, \ldots, X_N)$, $k \in [1:N]$. Then, by setting $U_k = Y_k$, $k \in [2:N]$ in (2), Theorem 2 simplifies as:

$$R(\mathcal{T}) < \min_{\mathcal{S}:\mathcal{T}\subseteq\mathcal{S}^c} H(Y(\mathcal{S}^c)|X(\mathcal{S}^c)), \quad \mathcal{T} \subseteq \mathcal{D}, \tag{4}$$

for some pmf $(\prod_{k=2}^{N} p(x_k))p(x_1|x_2^N)$. This region refines the results by Kannan, Raja, and Viswanath [7, Theorem 2] for deterministic broadcast networks by considering a slightly more general form of input pmfs. If the cutset bound in (3) is attained by pmfs of the same form, then the lower bound in (4) is tight. In particular, for the graphical network, whereby each link is noise-free and orthogonal, Theorem 2 recovers the result by Federgruen and Groenevelt [15]:

$$R(\mathcal{T}) < \min_{\mathcal{S}:\mathcal{T}\subseteq\mathcal{S}^c} C(\mathcal{S}),$$

where $C(\mathcal{S})$ denotes the capacity of the cut $(\mathcal{S}, \mathcal{S}^c)$.

**Example 2** (Gaussian networks)**.** Consider the additive white Gaussian noise network, in which the channel outputs are $Y_k = g_{k1}X_1 + \cdots + g_{kN}X_N + Z_k$, $k \in [1:N]$. Here $g_{kj}$ is the channel gain from node $j$ to node $k$ and $Z_1, \ldots, Z_N$ are independent Gaussian noise components with zero mean and unit variance. We assume average power constraint $P$ on each $X_k$ [12, Section 19.1]. On the one hand, by similar steps as in [13], the cutset bound in (3) can be relaxed to the set of rate tuples that satisfy,

$$R(\mathcal{S}^c) \le \frac{1}{2}\log\left|I + PG(\mathcal{S})G^T(\mathcal{S})\right| + \frac{|\mathcal{S}|}{2}.$$

On the other hand, in the inner bound (2) we set $X_k$, $k \in [1:N]$, i.i.d. $\mathrm{N}(0, P)$, and $U_k = g_{k1}X_1 + \cdots + g_{kN}X_N + \hat{Z}_k$, $k \in [2:N]$, where $\hat{Z}_k \sim \mathrm{N}(0,1)$ are independent of each other and of $(X^N, Y^N)$. Again, similar to the steps in [13], Theorem 1 simplifies as

$$R(\mathcal{S}^c) < \frac{1}{2}\log\left|I + PG(\mathcal{S})G^T(\mathcal{S})\right| - \frac{|\mathcal{S}^c|}{2}.$$

Comparing the inner and outer bounds, we can conclude that distributed decode–forward achieves within $0.5N$ bits per dimension from the cutset bound and thus from the capacity region. This improves upon the existing gap result of $O(N\log(N))$ by Kannan et al. [7, Theorem 1], establishing the tightest known gap for Gaussian broadcast relay networks. A similar gap result can be established for Gaussian vector (MIMO) networks. In this case, distributed decode–forward achieves within $0.5T$ bits per dimension from the cutset bound, where $T$ is the total number of antennas in the network.

**Example 3** (Broadcast channels)**.** Consider the single-hop discrete memoryless broadcast channel $p(y_2, \ldots, y_N|x_1)$, which corresponds to setting $Y_1 = X_2 = \cdots = X_N = \emptyset$ in our broadcast relay network model. For this case, Theorem 1 simplifies to Marton's inner bound [11], namely, a rate tuple $(R_2, \ldots, R_N)$ is achievable if

$$R(\mathcal{S}^c) < \sum_{k\in\mathcal{S}^c} I(U_k; Y_k) - \sum_{k\in\mathcal{S}^c} I(U_k; U(\mathcal{S}_k^c)), \tag{5}$$

for all $\mathcal{S} \subseteq [1:N]$ such that $1 \in \mathcal{S}$ and $\mathcal{S}^c \neq \emptyset$ for some pmfs $p(u_2^N)$ and functions $x_1(u_2^N)$. Equation (5) follows from

$$\begin{aligned}
&I(X_1; U(\mathcal{S}^c)) - \sum_{k\in\mathcal{S}^c} I(U_k; U(\mathcal{S}_k^c), X_1|Y_k)\\
&\quad= \sum_{k\in\mathcal{S}^c} I(U_k; X_1|U(\mathcal{S}_k^c)) - \sum_{k\in\mathcal{S}^c} I(U_k; U(\mathcal{S}_k^c), X_1|Y_k)\\
&\quad\overset{(a)}{=} \sum_{k\in\mathcal{S}^c} I(U_k; Y_k) + \sum_{k\in\mathcal{S}^c} I(U_k; X_1|U(\mathcal{S}_k^c))\\
&\qquad - \sum_{k\in\mathcal{S}^c} I(U_k; U(\mathcal{S}_k^c), X_1)\\
&\quad= \sum_{k\in\mathcal{S}^c} I(U_k; Y_k) - \sum_{k\in\mathcal{S}^c} I(U_k; U(\mathcal{S}_k^c))
\end{aligned}$$

where equality $(a)$ is due to the Markovity $Y^N \to X_1 \to U^N$. This result verifies that the distributed decode–forward coding scheme that achieves the inner bound in Theorem 1 is a natural generalization of the single-hop Marton coding scheme to multihop relaying.

## III. Proof of Theorem 1

We use a block Markov coding scheme in which a sequence of $b$ i.i.d. message tuples $\mathbf{M}_j = (M_{2j}, \ldots, M_{Nj})$, $j \in [1:b]$, is sent over $b$ blocks each consisting of $n$ transmissions. For each block, we generate codewords $U_k$, $k \in [2:N]$, to be recovered at node $k$. Using multicoding [12, Sections 7.8 and 8.3], we design these codewords to be dependent among themselves and on the transmitted codewords $X_1, \ldots, X_N$. The key difference from multicoding for single-hop networks is that here multicoding is performed using *backward encoding* over all blocks and the dependency among the codewords are satisfied simultaneously among them. Node $k$ recovers its intended message explicitly as well as some part of other messages implicitly by recovering $U_k$. The recovered part of the unintended messages, captured by an auxiliary index, is then forwarded to the destination nodes in the next block. The encoding and decoding operations are summarized in Table I. We now describe the full detail of the coding scheme.

*Codebook generation.* Fix $\prod_{k=2}^{N} p(x_k)p(x_1, u^N|x_2^N)$. For each block $j \in [1:b]$, randomly and independently generate $2^{n\hat{R}_k}$ sequences $x_{kj}^n(l_{k,j-1})$, $l_{k,j-1} \in$

| Block | 1 | 2 | $\cdots$ | $b-1$ | $b$ |
|---|---|---|---|---|---|
| Multicoding | $\mathbf{l}_0$ | $\leftarrow \mathbf{l}_1$ | $\ldots$ | $\leftarrow \mathbf{l}_{b-2}$ | $\leftarrow \mathbf{l}_{b-1}$ |
| $X_1$ | $x_1^n(\mathbf{m}_1, \mathbf{l}_1 \vert \mathbf{l}_0)$ | $x_1^n(\mathbf{m}_2, \mathbf{l}_2 \vert \mathbf{l}_1)$ | $\ldots$ | $x_1^n(\mathbf{m}_{b-1}, \mathbf{l}_{b-1} \vert \mathbf{l}_{b-2})$ | $x_1^n(\mathbf{m}_b, \mathbf{l}_b \vert \mathbf{l}_{b-1})$ |
| $X_k$ | $x_k^n(l_{k0})$ | $x_k^n(l_{k1})$ | $\ldots$ | $x_k^n(l_{k,b-2})$ | $x_k^n(l_{k,b-1})$ |
| $Y_k$ | $(\hat{m}_{k1}, l_1) \rightarrow$ | $(\hat{m}_{k2}, l_{k2}) \rightarrow$ | $\ldots$ | $(\hat{m}_{k,b-1}, l_{k,b-1}) \rightarrow$ | $(\hat{m}_{kb}, l_{kb}) \rightarrow$ |

TABLE I
ENCODING AND DECODING OF THE DISTRIBUTED DECODE–FORWARD CODING SCHEME.

$[1 : 2^{n\hat{R}_k}]$, each according to $\prod_{i=1}^n p_{X_k}(x_{ki})$, $k \in [2 : N]$. For each $l_{k,j-1}$, randomly and independently generate $2^{n(R_k+\hat{R}_k)}$ sequences $u_k^n(m_{kj}, l_{kj}|l_{k,j-1})$, $(m_{kj}, l_{kj}) \in [1 : 2^{nR_k}] \times [1 : 2^{n\hat{R}_k}]$, each according to $\prod_{i=1}^n p_{U_k|X_k}(u_{ki}|x_{ki}(l_{k,j-1}))$. For each $\mathbf{m}_j = (m_{2j}, \ldots, m_{Nj})$, $\mathbf{l}_j = (l_{2j}, \ldots, l_{Nj})$, and $\mathbf{l}_{j-1} = (l_{2,j-1}, \ldots, l_{N,j-1})$, randomly and independently generate sequences $x_1^n(\mathbf{m}_j, \mathbf{l}_j|\mathbf{l}_{j-1})$, each according to

$$\prod_{i=1}^n p_{X_1|U_2^N, X_2^N}(x_{1i}|u_{2i}(l_{2j}|l_{2,j-1}), \ldots, \\ u_{Ni}(l_{Nj}|l_{N,j-1}), x_{2i}(l_{2j}), \ldots, \\ x_{Ni}(l_{Nj})).$$

*Encoding.* For $j = b, b-1, \ldots, 1$, given $\mathbf{m}_j$, find an index tuple $\mathbf{l}_{j-1}$ such that

$$(u_2^n(m_{2j}, l_{2j}|l_{2,j-1}), \ldots, u_{Nj}^n(m_{Nj}, l_{Nj}|l_{N,j-1}), \\ x_{2j}^n(l_{2,j-1}), \ldots, x_{Nj}^n(l_{N,j-1})) \in \mathcal{T}_{\epsilon'}^{(n)},$$

successively with the initial condition $l_{2b} = \cdots = l_{Nb} = 1$. If there is more than one such index tuple, select one of them arbitrarily. If there is none, select one arbitrarily from $[1 : 2^{n\hat{R}_2}] \times \cdots \times [1 : 2^{n\hat{R}_N}]$. By a direct application of the properties of multivariate typicality [12, Section 2.5], induction on backward encoding, and steps similar to those of the multivariate covering lemma [12, Lemma 8.2], it can be shown that encoding is successfully with high probability if

$$\hat{R}(\mathcal{T}) > \sum_{k\in\mathcal{T}} I(U_k; U(\mathcal{T}_k), X(\mathcal{T})|X_k) + \delta(\epsilon') \qquad (6)$$

for all $\mathcal{T} \subseteq [2 : N]$, where $\mathcal{T}_k = \mathcal{T} \cap [2 : k-1]$.

Before the actual transmission of the messages, we use additional $(N-1)^2$ blocks to transmit each $l_{k0}$ to node $k \in [2 : N]$ using multihop coding, as in the initialization phase for short-message noisy network coding in [6] and distributed decode–forward for multicast in [13]. The additional transmission needed for this phase is in the order of $O(nN^2)$, independent of $b$. Thus, the realized transmission rate converges to $R$ as $b \to \infty$. In the following, we assume that all $l_{k0}$ indices are known prior to transmission.

To send message tuple $\mathbf{m}_j$ in block $j$, the source node transmits $x_1^n(\mathbf{m}_j, \mathbf{l}_j|\mathbf{l}_{j-1})$, where $(\mathbf{l}_j, \mathbf{l}_{j-1})$ is the chosen index tuple.

*Relay encoding and decoding.* Let $\epsilon > \epsilon'$. At the end of block $j$, node $k$ finds a *unique* pair $(\tilde{m}_{kj}, \tilde{l}_{kj}) \in [1 : 2^{nR_k}] \times [1 : 2^{n\hat{R}_k}]$ such that

$$(u_k^n(\tilde{m}_{kj}, \tilde{l}_{kj}|\tilde{l}_{k,j-1}), x_{kj}^n(\tilde{l}_{k,j-1}), y_{kj}^n) \in \mathcal{T}_\epsilon^{(n)}$$

and declares $\tilde{m}_{kj}$ as its message estimate. (If there is none or more than one pair, declare $\tilde{m}_{kj} = 1$.) By the packing lemma [12, Lemma 3.1], this is successful with high probability if $\tilde{l}_{k,j-1}$ was recovered correctly at the end of the previous block and

$$R_k + \hat{R}_k < I(U_k; Y_k|X_k) - \delta(\epsilon). \qquad (7)$$

In the next block $j+1$, node $k$ transmits $x_{k,j+1}^n(\tilde{l}_{kj})$. By identifying $\mathcal{S} = [1 : N] \setminus \mathcal{T}$ and eliminating the auxiliary rates $\hat{R}_2, \ldots, \hat{R}_N$ from (6) and (7), we have

$$R(\mathcal{S}^c) < I(X(\mathcal{S}); U(\mathcal{S}^c)|X(\mathcal{S}^c)) \\ - \sum_{k\in\mathcal{S}^c} I(U_k; U(\mathcal{S}_k^c), X^N|X_k, Y_k) - \delta'(\epsilon)$$

for some $\mathcal{S} \subseteq [1 : N]$ such that $1 \in \mathcal{S}$ and $\mathcal{S}^c \neq \emptyset$.

## IV. DISCUSSION

As a dual setting to the broadcast relay network, consider the *multiple access relay network* $p(y^N|x^N)$, in which source nodes $k \in [2 : N]$ communicate independent messages to the common destination node 1. This is a special case of the multimessage multicast network [12, Section 18.4] and the noisy network coding inner bound [4], [5], [6] simplifies as follows.

**Theorem 3.** *For the discrete memoryless multiple access relay network, a rate tuple* $(R_2, \ldots, R_N)$ *is achievable if*

$$R(\mathcal{S}) < I(X(\mathcal{S}); \hat{Y}(\mathcal{S}^c), Y_1|X(\mathcal{S}^c)) \\ - I(Y(\mathcal{S}); \hat{Y}(\mathcal{S})|X^N, \hat{Y}(\mathcal{S}^c), Y_1) \qquad (8)$$

*for all* $\mathcal{S} \subseteq [1 : N]$ *such that* $1 \in \mathcal{S}^c$ *and* $\mathcal{S} \neq \emptyset$ *for some joint pmf* $\prod_{k=1}^N p(x_k)p(\hat{y}_k|y_k, x_k)$.

This inner bound has a form that is, in a sense, dual to Theorem 1. The first term in (8) has auxiliary random variables $\hat{Y}_j$, which is to be encoded at node $j$ and to be decoded at node 1. In comparison, the first term in (1) has auxiliary random variables $U_j$, which is to be encoded at node 1 and to be decoded at node $j$. In addition, the second term in (8) quantifies the cost of

| | Distributed decode–forward | Noisy network coding |
|---|---|---|
| Network model | Broadcast relay networks | Multiple access relay networks |
| Key component | Backward encoding at the source | Backward decoding at the destination |
| Single–hop | Marton inner bound | MAC capacity region |
| Relay channel | Partial decode–forward lower bound | Compress–forward lower bound |
| Gaussian capacity gap | $0.5N$ bits | $0.63N$ bits |

TABLE II
COMPARISON BETWEEN DISTRIBUTED DECODE–FORWARD AND NOISY NETWORK CODING.

decoding $\hat{Y}_j$ at node 1, while the second term in (1) quantifies the cost of encoding $U_j$ at node 1.

As with distributed decode–forward, for deterministic networks, the noisy network coding scheme of Theorem 3 achieves

$$R(\mathcal{S}) < H(Y(\mathcal{S}^c)|X(\mathcal{S}^c)),$$

which is of the same form as the cutset bound except for the set of input pmfs. For Gaussian networks, noisy network coding uniformly achieves within $0.63N$ bits per dimension from the capacity region. Finally, for the single-hop multiple access channel, Theorem 3 simplifies to the multiple access channel capacity region (before time sharing), namely,

$$R(\mathcal{S}) < I(X(\mathcal{S}); Y_1|X(\mathcal{S}^c))$$

for all $\mathcal{S}$ for some $\prod_{k=1}^{n} p(x_k)$. Thus, both distributed decode–forward (DDF) and noisy network coding (NNC) naturally extend the standard broadcast channel and multiple access channel coding schemes by combining them with (partial) decode–forward and compress–forward, respectively.

The duality between Theorems 1 and 3 is also reflected by the operations of the DDF and NNC schemes. In destination-centric NNC, the source and the relays are relatively simple, but the major burden is on the destination so as to recover the messages and the compression indices from the entire network over multiple blocks. This scheme fits well with (and currently is the only reasonable solution to) general *multiple access* relay networks. In source-centric DDF, the relays and the destinations are relatively simple, but the source needs to precode dependent codewords for the entire network over multiple blocks. The scheme fits well with (and currently is the only reasonable solution to) general *broadcast* relay networks.

This operational reciprocity in the roles of source and destination for multiple access and broadcast has been well noted by Kannan, Raja, and Viswanath [7], which was the key intuition for their coding scheme that parallels the quantize–map–forward scheme by Avestimehr, Diggavi, and Tse [3]. Compared to these nested multiletter schemes [3], [7], however, NNC and DDF are more general (designed for arbitrary wired or wireless or hybrid networks) and provide easy-to-evaluate single-letter performance bounds.

We finally note that just as NNC can deal with multiple destination nodes easily (the feature of which goes back to the original network coding scheme by Ahlswede, Cai, Li, and Yeung [16]), DDF can be also adapted to the groupcast scenario, in which a single-source broadcasts multiple messages to nonoverlapping groups of destination nodes.

## REFERENCES


[1] 3GPP, "Coordinated multi-point operation for LTE physical layer aspects," TR 36.819 v11.2.0., 3GPP, Tech. Rep., 2013.

[2] China Mobile, "C-RAN: The road towards green RAN," White paper, ver. 2.5, China Mobile Research Institute, 2011.

[3] A. S. Avestimehr, S. N. Diggavi, and D. N. C. Tse, "Wireless network information flow: A deterministic approach," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 1872–1905, Apr. 2011.

[4] S. H. Lim, Y.-H. Kim, A. El Gamal, and S.-Y. Chung, "Noisy network coding," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 3132–3152, May 2011.

[5] M. H. Yassaee and M. R. Aref, "Slepian-Wolf coding over cooperative relay networks," *IEEE Trans. Inf. Theory*, vol. 57, no. 6, pp. 3462–3482, 2011.

[6] J. Hou and G. Kramer, "Short message noisy network coding with a decode–forward option," 2013. [Online]. Available: http://arxiv.org/abs/1304.1692/

[7] S. Kannan, A. Raja, and P. Viswanath, "Approximately optimal wireless broadcasting," *IEEE Trans. Inf. Theory*, vol. 58, no. 12, pp. 7154–7167, 2012.

[8] M. Anand and P. R. Kumar, "A digital interface for Gaussian relay and interference networks: Lifting codes from the discrete superposition model," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 2548–2564, 2011.

[9] T. M. Cover and A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 5, pp. 572–584, Sep. 1979.

[10] B. Schein and R. G. Gallager, "The Gaussian parallel relay channel," in *Proc. IEEE Int. Symp. Inf. Theory*, Sorrento, Italy, Jun. 2000, p. 22.

[11] K. Marton, "A coding theorem for the discrete memoryless broadcast channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 3, pp. 306–311, 1979.

[12] A. El Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge: Cambridge University Press, 2011.

[13] S. H. Lim, K. T. Kim, and Y.-H. Kim, "Distributed decode–forward for multicast," in *Proc. IEEE Int. Symp. Inf. Theory*, Honolulu, HI, June/July 2014.

[14] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. New York: Wiley, 2006.

[15] A. Federgruen and H. Groenevelt, "Polymatroidal flow network models with multiple sinks," *Networks*, vol. 18, no. 4.

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