Coding Schemes for Discrete Memoryless Multicast Networks With and Without Feedback

Youlong Wu
Lehrstuhl für Nachrichtentechnik
Technische Universität München, Germany
youlong.wu@tum.de

***Abstract*—Coding scheme for discrete memoryless multicast networks with rate-limited feedback from the receivers and relays to the transmitter is proposed. The coding scheme is based on block-Markov coding, joint backward decoding and hybrid relaying strategy. In each block, the relays use partial decode-forward strategy to decode part of the source message. Meanwhile, the receivers and relays use compress-forward strategy to compress their channel outputs and send the compression indices to the transmitter through the feedback links. In the next block, after obtaining the compression indices, the transmitter sends them together with the source message. Each receiver uses backward decoding to jointly decode the source message and all compression indices. It is shown that our coding scheme generalizes Gabbai and Bross's results for the single relay channel with partial feedback, where they proposed coding schemes based on restricted decoding and deterministic partitioning. For the single relay channel with relay-transmitter feedback, our coding scheme can strictly improve on noisy network coding, distributed decode-forward coding and all known lower bounds on the achievable rate in the absence of feedback. Furthermore, motivated by the feedback coding scheme, we propose a new coding scheme for discrete memoryless multicast networks without feedback, which also improves noisy network coding and distributed decode-forward coding.**

I. Introduction

The relay channel [1] describes a 3-node communication channel where the transmitter sends a message to the receiver with the assistance of relay. Cover and El Gamal [2] proposed two basic coding strategies: compress-forward and decode-forward that are based on block-Markov coding. The compress-forward strategy has the relay compress its outputs and send the compression index to the receiver. The decode-forward strategy has the relay first decode all or part of the message and then send the decoded message to the receiver. Both strategies have been generalized to multiple-relay channels in [3]. The compress-forward strategy was later extended to multi-message multicast and multi-messages networks, called *noisy network coding* (NNC) [4], [5], [6]. Recently, a *distributed decode-forward coding* (DDF) scheme was proposed for multicast [7] and broadcast relay networks [8], which uses the partial decode-forward strategy at the relays and backward coding at the transmitter.

Both decode-forward and compress-forward require sophisticated operations. A much simpler strategy, called amplify-forward, was introduced by Schein and Gallager [9] for the 4-node Gaussian diamond network. In amplify-forward, the relay scales its received signal and forwards it to the receiver. A hybrid coding scheme for the general noisy-relay networks was proposed in [10] that unifies both amplify-forward and NNC.

Perfect feedback from the receiver to the relay makes the relay channel *physically degraded* [2], and therefore decode-forward achieves the capacity. If there is feedback from the receiver or relay to the transmitter, the capacity is unknown in general. In [11] Gabbai and Bross studied this problem and proposed inner bounds by using restricted decoding and deterministic partitioning [12].

In this paper, we consider the general discrete memoryless multicast network with *rate-limited* feedback. This network consists of $N \geq 3$ nodes where the transmitter sends a source message to different receivers with the assistance of multiple relays and in the presence of rate-limited feedback from the receivers and relays to the transmitter. We propose a new coding scheme based on block-Markov coding, joint backward decoding and hybrid relaying. In our scheme, In each block, the relays use partial decode-forward strategy to decode part of the source message. Meanwhile, the receivers and relays compress their channel outputs and send the compression indices to the transmitter through the feedback links. In the next block, after obtaining the compression indices, the transmitter sends them together with the source message. Each receiver uses backward decoding to jointly decode the source message and all compression indices.

Our coding scheme is reminiscent of the noisy network coding for general networks [5], [6] in the sense the relays and receivers compress their channel outputs and send these compression indices over the feedback links. However, we introduce combined compress-forward and partial decode-forward strategy into the relays. And our scheme has the transmitter *forward* the receivers and relays' compression messages, instead of creating a new compression message. This is similar to the schemes in [13] for the broadcast channel, where the transmitter forwards the receivers' compression messages. It is shown that our coding scheme generalizes Gabbai and Bross's results [11] for the relay channel with relay-transmitter feedback. For some channels, such as the Gaussian relay channel and Z relay channels, our coding scheme improve over the NNC scheme [5], the DDF coding scheme [7], [8] and all known lower bounds on the achievable rate in the absence of feedback.

Motivated by our feedback coding scheme, we propose a new coding scheme for multicast networks *without* feedback, which can strictly improve NNC and DDF for some channels.

Notation: We use capital letters to denote random variables and small letters for their realizations, e.g. X and x. For $k, j \in \mathbb{Z}^+$, let $X_k^j := (X_k, \ldots, X_j)$ and $x_k^j := (x_{k,1}, \ldots, x_{k,j})$. Given a set of integers $\mathcal{A} \subseteq [2:N]$ and $k \in \mathcal{A}$, we denote by $|\mathcal{A}|$ its cardinality and define $\mathcal{A}^c := [2:N]\backslash\mathcal{A}$. A tuple of random variables is denoted as $X(\mathcal{A}) := [X_k : k \in \mathcal{A}]$. Given a positive integer n, let $\mathbf{1}_{[n]}$ denote the all-one tuple of length n, e.g., $\mathbf{1}_{[3]} = (1,1,1)$. Define a function $\mathcal{C}(x) := \frac{1}{2}\log_2(1+x)$.

II. System model

Consider an N-node discrete memoryless (DM) multicast networks with feedback from the receivers and relays to the transmitter, see Figure 1. Let $\mathcal{R}$ and $\mathcal{D}$ denote the set of relays and receivers, respectively, where $\mathcal{R} \subset [2:N]$ and $\mathcal{R} = [2:N]\backslash\mathcal{D}$. This setup is characterized by $2N$ finite alphabets $\mathcal{X}_1, \ldots, \mathcal{X}_N, \mathcal{Y}_1, \ldots, \mathcal{Y}_N$, a channel law $P_{Y_1 \cdots Y_N | X_1, \ldots, X_N}$ and nonnegative feedback rates $R_{\text{Fb},k}$, for $k \in [2:N]$. Specifically, at discrete-time i, node $j \in [1:N]$ sends input $x_{j,i} \in \mathcal{X}_j$ and then observes output $y_{j,i} \in \mathcal{Y}_j$. After observing $y_{k,i}$, for $k \in [2:N]$, node k sends a feedback signal $f_{k,i} \in \mathcal{F}_{k,i}$ to the transmitter, where $\mathcal{F}_{k,i}$ denotes the finite alphabet of $f_{k,i}$. The feedback link between the transmitter and node k is noiseless and *rate-limited* to $R_{\text{Fb},k}$ bits per channel use. In other words, if the transmission takes place over a total blocklength n, then

$$|\mathcal{F}_{k,1}| \times \cdots \times |\mathcal{F}_{k,n}| \leq 2^{nR_{\text{Fb},k}}, \quad k \in [2:N]. \tag{1}$$

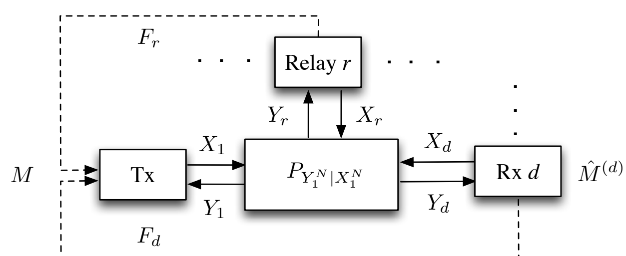

Fig. 1. N-node multicast network with partial feedback

The transmitter communicates a message $m \in [1:2^{nR}]$ to the set of receivers $\mathcal{D}$ with the assistance of the relays $\mathcal{R}$. A $(2^{nR}, n)$ code for this channel has

- a message set $[1:2^{nR}]$,
- a source encoder that maps $(m, y_1^{i-1}, f_2^{i-1}, \ldots, f_N^{i-1})$ to the channel input $x_{1,i}(m, y_1^{i-1}, f_2^{i-1}, \ldots, f_N^{i-1})$, for each time $i \in [1:n]$,
- relay and receiver encoders that maps y_k^{i-1} to a sequence $x_{k,i}(y_k^{i-1})$, for each $k \in [2:N]$ and $i \in [1:n]$,
- feedback-encoders that produce feedback symbols $f_{k,i}(y_k^i)$, for each $k \in [2:N]$ and $i \in [1:n]$,
- decoders that estimates $\hat{m}^{(d)}$ based on y_d^n, for $d \in \mathcal{D}$.

Suppose m is uniformly distributed over the message set. A rate R with average feedback rates $R_{\text{Fb},k}$, for $k \in [2:N]$, is called achievable if for every blocklength n, there exists a $(2^{nR}, n)$ code such that the average probability of error $P_e^{(n)} = \Pr[\hat{m}^{(d)} \neq m, \text{ for some } d \in \mathcal{D}]$ tends to 0 as the n tends to infinity. The capacity is the supremum of the set of achievable rates R such that $\lim_{n\to\infty} P_e^{(n)} = 0$.

III. Main Results

This section presents our main results. The proofs are given in Sections V and VI.

Theorem 1. *For DM multicast networks with feedback from the receivers and relays to the transmitter, the rate R is achievable if*

$$\begin{aligned} R &\leq I(X_1; \hat{Y}_2^N, Y_d | U_2^N, X_2^N) + \min_{r \in \mathcal{R}} I(U_r; Y_r | X_r) \\ R &\leq I(X_1, X(\mathcal{T}), U(\mathcal{T}); \hat{Y}(\mathcal{T}^c), Y_d | X(\mathcal{T}^c), U(\mathcal{T}^c)) \\ &\quad - I(\hat{Y}(\mathcal{T}); Y(\mathcal{T}) | U_2^N, X_1^N, \hat{Y}(\mathcal{T}^c), Y_d) \end{aligned} \tag{2}$$

for all $d \in \mathcal{D}$, $\mathcal{T} \subset [2:N]$ with $\mathcal{T}^c \cap \mathcal{D} \neq \emptyset$, and for some pmf

$$\begin{aligned} &\left[\prod_{k=2}^N P_{X_k U_k}\right] P_{X_1|X_2^N U_2^N} P_{Y_1^N|X_1^N} \\ &\quad \times \left[\prod_{r \in \mathcal{R}} P_{\hat{Y}_r|U_r X_r Y_r}\right] \left[\prod_{d \in \mathcal{D}} P_{\hat{Y}_d|X_d Y_d}\right] \end{aligned} \tag{3}$$

such that

$$R_{Fb,r} \geq I(\hat{Y}_r; Y_r | X_r, U_r), \quad \textit{for } r \in \mathcal{R} \tag{4a}$$

$$R_{Fb,d} \geq I(\hat{Y}_d; Y_d | X_d), \quad \textit{for } d \in \mathcal{D}. \tag{4b}$$

Proof: See Section V-A ■

By setting $U_k = \emptyset$, for all $k \in [2:N]$, we obtain the following corollary.

Corollary 1. *For DM multicast networks with feedback from the receivers and relays to the transmitter, the rate R is achievable if*

$$\begin{aligned} R &\leq I(X_1, X(\mathcal{T}); \hat{Y}(\mathcal{T}^c), Y_d | X(\mathcal{T}^c)) \\ &\quad - I(\hat{Y}(\mathcal{T}); Y(\mathcal{T}) | X_1^N, Y_d, \hat{Y}(\mathcal{T}^c)) \end{aligned} \tag{5}$$

for all $d \in \mathcal{D}$, $\mathcal{T} \subset [2:N]$ with $\mathcal{T}^c \cap \mathcal{D} \neq \emptyset$ and for some pmf

$$\left[\prod_{k=2}^N P_{X_k}\right] P_{X_1|X_2^N} P_{Y_1^N|X_1^N} \left[\prod_{k=2}^N P_{\hat{Y}_k|X_k Y_k}\right] \tag{6}$$

such that

$$R_{Fb,k} \geq I(\hat{Y}_k; Y_k | X_k), \quad \textit{for } k \in [2:N]. \tag{7}$$

Remark 1. *Comparing the lower bound in Corollary 1 with the NNC lower bound [5, Theorem 1], our rates includes NNC if the feedback rates are sufficient large, i.e., if* (7) *holds for all pmfs* (6)*, since in* (6) *we allow the joint distribution $\prod_{k=2}^N P_{X_k} P_{X_1|X_2^N}$ instead of $\prod_{k=1}^N P_{X_k}$.*

Based on the coding scheme for Theorems 1, we propose another coding scheme for DM multicast networks *without* feedback. The new achievable rate is shown below.

$$R \leq I(X_1;\hat{Y}_2^N, Y_d|U_2^N, V_2^N, X_2^N) + \min_{r\in\mathcal{R}} I(U_r;Y_r|V_r,X_r) \tag{8a}$$

$$R \leq I(X_1, X(\mathcal{T}), U(\mathcal{T}), V(\mathcal{T}); \hat{Y}(\mathcal{T}^c), Y_d|U(\mathcal{T}^c), V(\mathcal{T}^c), X(\mathcal{T}^c)) - I(\hat{Y}(\mathcal{T}); Y(\mathcal{T})|U_2^N, V_2^N, X_1^N, \hat{Y}(\mathcal{T}^c), Y_d) \tag{8b}$$

Theorem 2. *For DM multicast networks without feedback, the rate* R *is achievable if* (8) *holds for all* $d \in \mathcal{D}$, $\mathcal{T} \subset [2:N]$ *with* $\mathcal{T}^c \cap \mathcal{D} \neq \emptyset$, *and for some pmf*

$$\left[\prod_{k=2}^{N} P_{V_k}P_{X_k|V_k}P_{U_k|V_k}\right] P_{X_1|V_2^N U_2^N} \times P_{Y_1^N|X_1^N}\left[\prod_{r\in\mathcal{R}} P_{\hat{Y}_r|U_rV_rX_rY_r}\right]\left[\prod_{d\in\mathcal{D}} P_{\hat{Y}_d|V_dX_dY_d}\right] \tag{9}$$

such that

$$\sum_{r\in\mathcal{T}\cap\mathcal{R}} I(\hat{Y}_r;Y_r|U_r,V_r,X_r) + \sum_{d\in\mathcal{T}\cap\mathcal{D}} I(\hat{Y}_d;Y_d|V_d,X_d) \leq I(X(\mathcal{T});Y_1|U_2^N,V_2^N,X(T^c),X_1). \tag{10}$$

Remark 2. *Theorem 2 requires the transmitter to decode the compression messages generated by all receivers and relays, which may limit the performance if their are weak links from the receivers or relays to the transmitter. One easy way to improve the scheme is to allow the transmitter to decode some nodes' compression messages. Suppose the transmitter decodes only the compression messages generated by the set of nodes* $\mathcal{A} \subseteq [2:N]$. *Then by a scheme similar to that for Theorem 2, we obtain a new lower bound satisfying* (11) *for all* $d \in \mathcal{D}$, $\mathcal{T} \subset [2:N]$ *with* $\mathcal{T}^c \cap \mathcal{D} \neq \emptyset$, *and for some pmf*

$$\left[\prod_{k=2}^{N} P_{V_k}P_{X_k|V_k}P_{U_k|V_k}\right] P_{X_1|V(\mathcal{A})U(\mathcal{A})} \times P_{Y_1^N|X_1^N}\left[\prod_{r\in\mathcal{R}} P_{\hat{Y}_r|U_rV_rX_rY_r}\right]\left[\prod_{d\in\mathcal{D}} P_{\hat{Y}_d|V_dX_dY_d}\right] \tag{12}$$

such that

$$\sum_{r\in\mathcal{T}_\mathcal{A}\cap\mathcal{R}} I(\hat{Y}_r;Y_r|U_r,V_r,X_r) + \sum_{d\in\mathcal{T}_\mathcal{A}\cap\mathcal{D}} I(\hat{Y}_d;Y_d|V_d,X_d) \leq I(X(T_\mathcal{A});Y_1|U(\mathcal{A}),V(\mathcal{A}),X(T_\mathcal{A}^c),X_1) \tag{13}$$

where $\mathcal{T}_\mathcal{A} = \mathcal{T}\cap\mathcal{A}$, $\mathcal{T}_\mathcal{A}^c$ *is the complement of* $\mathcal{T}_\mathcal{A}$ *in* $\mathcal{A}$.

This lower bound reduces to the lower bound in Theorem 2 when $\mathcal{A} = [2:N]$, *and to the NNC lower bound when* $\mathcal{A} = \emptyset$ *and* $V_k = U_k = \emptyset$, *for all* $k \in [2:N]$.

IV. Examples

A. The relay channel with relay-transmitter feedback

Consider the relay channel with perfect feedback from the relay to the transmitter, see Figure 2.

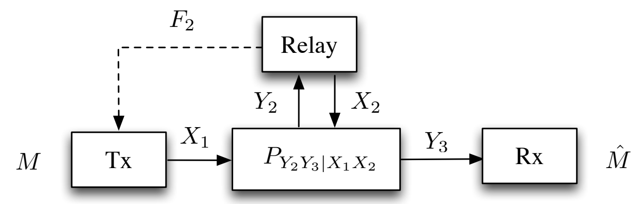

Fig. 2. Relay channel with relay-transmitter feedback

Let $U_3 = \hat{Y}_3 = \emptyset$, then Theorem 1 specializes to

$$\left.\begin{aligned} R &\leq I(X_1;\hat{Y}_2,Y_3|U_2,X_2) + I(U_2;Y_2|X_2)\\ R &\leq I(X_1,X_2;Y_3) - I(\hat{Y}_2;Y_2|U_2,X_1,X_2,Y_3)\end{aligned}\right. \tag{14}$$

for some pmf $P_{X_1X_2U_2}P_{\hat{Y}_2|X_2U_2Y_2}$.

Let $\hat{Y}_3 = \emptyset$, then Corollary 1 specializes to

$$\left.\begin{aligned} R &\leq I(X_1;\hat{Y}_2,Y_3|X_2)\\ R &\leq I(X_1,X_2;Y_3) - I(\hat{Y}_2;Y_2|X_1,X_2,Y_3)\end{aligned}\right\} \tag{15}$$

for some pmf $P_{X_1X_2}P_{\hat{Y}_2|X_2Y_2}$.

In [11] Gabbai and Bross studied this channel and proposed coding schemes based on restricted decoding and deterministic partitioning. The rates (15) and (14) recover Gabbai and Bross's rates of Theorems 2 and 3 in [11], respectively.

By using NNC [5], the rate R satisfying

$$\left.\begin{aligned} R &\leq I(X_1;\hat{Y}_2,Y_3|X_2)\\ R &\leq I(X_1,X_2;Y_3) - I(\hat{Y}_2;Y_2|X_1,X_2,Y_3)\end{aligned}\right\} \tag{16}$$

is achievable for any pmf $P_{X_1}P_{X_2}P_{\hat{Y}_2|X_2Y_2}$, which coincides with the compress-forward lower bound [2, Theorem 6].

By using DDF [7], [8], the rate R satisfying

$$R \leq I(X_1,X_2;Y_3) \tag{17a}$$

$$R \leq I(U_2;Y_2|X_2) + I(X_1;Y_3|X_2,U_2) \tag{17b}$$

is achievable for any pmf $P_{X_1X_2U_2}$, which coincides with the partial decode-forward lower bound [2, Theorem 7].

The lower bound (14) includes (16) and (17). In [11] Gabbai and Bross showed that for the Gaussian and Z relay channels, the lower bound (14) improves on the known lower bounds on the achievable rate in the absence of feedback, including the compress-forward lower bound in (16), and the partial decode-forward lower bound in (17). In view of this fact, we have the following corollary:

Corollary 2. *For the DM single-relay channel with relay-transmitter feedback, our coding scheme recovers Gabbai and Bross's results, and can strictly improve on NNC [5], DDF [7] and all known lower bounds on the achievable rate in the absence of feedback.*

B. Enhanced Gaussian relay channel

Consider an enhanced Gaussian relay channel where the transmitter can access the output Y_1, see Figure 3. The channel

$$R \leq \max_{\mathcal{A}\subseteq[2:N]} \left\{ I(X_1;\hat{Y}_2^N, Y_d|U_2^N, V_2^N, X_2^N) + \min_{r\in\mathcal{R}} I(U_r;Y_r|V_r,X_r) \right\} \tag{11a}$$

$$R \leq \max_{\mathcal{A}\subseteq[2:N]} \left\{ I(X_1,X(\mathcal{T}),U(\mathcal{T}),V(\mathcal{T});\hat{Y}(\mathcal{T}^c),Y_d|U(\mathcal{T}^c),V(\mathcal{T}^c),X(\mathcal{T}^c)) - I(\hat{Y}(\mathcal{T});Y(\mathcal{T})|U_2^N,V_2^N,X_1^N,\hat{Y}(\mathcal{T}^c),Y_d) \right\} \tag{11b}$$

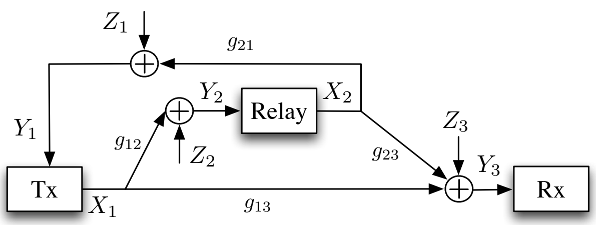

Fig. 3. The enhanced Gaussian relay channel

outputs are: $Y_1 = g_{21}X_2 + Z_1$, $Y_2 = g_{12}X_1 + Z_2$ and $Y_3 = g_{13}X_1 + g_{23}X_2 + Z_3$, where g_{21}, g_{23}, g_{12} and g_{13} are channel gains and $Z_1 \sim \mathcal{N}(0,1)$, $Z_2 \sim \mathcal{N}(0,1)$ and $Z_3 \sim \mathcal{N}(0,1)$ are independent Gaussian noise variables. The input power constraints are $\mathbb{E}|X_1^2| \leq P_1$ and $\mathbb{E}|X_2^2| \leq P_2$. Let $s_{12} = g_{12}^2 P_1$, $s_{13} = g_{13}^2 P_1$, $s_{23} = g_{23}^2 P_2$ and $s_{21} = g_{21}^2 P_2$.

We compare the lower bound in Theorem 2 with the cut-set outer bound and the previous known lower bounds, such as amplify-forward, NNC, DDF and Cover-El Gama's general lower bound [2, Theorem 7].

Achievable rate in Theorem 2: Let $U_3 = V_3 = \hat{Y}_3 = \emptyset$, then Theorem 2 reduces to

$$\begin{aligned} &R \leq I(X_1;\hat{Y}_2, Y_3|U_2, V_2, X_2) + I(U_2;Y_2|V_2,X_2) \\ &R \leq I(X_1,X_2;Y_3) - I(\hat{Y}_2;Y_2|U_2,V_2,X_1,X_2,Y_3) \end{aligned} \tag{18}$$

for some pmf $P_{V_2}P_{X_2|V_2}P_{U_2|V_2}P_{X_1|V_2U_2}P_{\hat{Y}_2|X_2V_2U_2Y_2}$ such that $I(\hat{Y}_2;Y_2|U_2,V_2,X_2) \leq I(X_2;Y_1|U_2,X_1,V_2)$. To compute (26), we choose distributions as in [14]:

$$\begin{aligned} &U_2 = aV_2 + W_0, X_2 = cV_2 + W_2 \\ &X_1 = bU_2 + W_1, \hat{Y}_2 = Y_2 + Z' \end{aligned} \tag{19}$$

where $V_2 \sim \mathcal{N}(0,P_1), W_0 \sim \mathcal{N}(0,\frac{\bar{\alpha}\beta P_1}{b^2}), W_1 \sim \mathcal{N}(0,\alpha P_1), W_2 \sim \mathcal{N}(0,\gamma P_2)$ and $Z' \sim \mathcal{N}(0,N')$ are independent, for $\alpha,\beta,\gamma \in [0,1]$. For this choice, we have,

$$\begin{aligned} I(X_1;\hat{Y}_2,Y_3|X_2,V_2,U_2) &= \mathcal{C}\Big(\alpha s_{13} + \frac{\alpha s_{12}}{1+N'}\Big) \\ I(U_2;Y_2|V_2,X_2) &= \mathcal{C}\Big(\frac{s_{12}\beta\bar{\alpha}}{\alpha s_{12}+1}\Big) \\ I(X_1,X_2;Y_3) &= \mathcal{C}\Big(2\sqrt{\bar{\alpha}\bar{\beta}\bar{\gamma}s_{13}s_{23}}+s_{13}+s_{23}\Big) \\ I(\hat{Y}_2;Y_2|U_2,V_2,X_1,X_2,Y_3) &= \mathcal{C}\Big(\frac{1}{N'}\Big), \end{aligned} \tag{20}$$

and

$$\begin{aligned} I(\hat{Y}_2;Y_2|U_2,V_2,X_2) &= \mathcal{C}\Big(\frac{1+\alpha s_{12}}{N'}\Big) \\ I(X_2;Y_1|U_2,X_1,V_2) &= \mathcal{C}(\gamma s_{21}). \end{aligned} \tag{21}$$

Thus we obtain the lower bound

$$\begin{aligned} R \leq \min\Big\{&\mathcal{C}\Big(\alpha s_{13} + \frac{\alpha s_{12}}{1+N'}\Big) + \mathcal{C}\Big(\frac{s_{12}\beta\bar{\alpha}}{\alpha s_{12}+1}\Big), \\ &\mathcal{C}\Big(2\sqrt{\bar{\alpha}\bar{\beta}\bar{\gamma}s_{13}s_{23}} + s_{13} + s_{23}\Big) - \mathcal{C}\Big(\frac{1}{N'}\Big)\Big\} \end{aligned} \tag{22}$$

subject to the constraint

$$N' \geq \frac{1+\alpha s_{12}}{\gamma s_{21}}. \tag{23}$$

Amplify-forward: For the general Gaussian relay channel with linear relaying functions, finding the channel capacity is a non-convex optimization problem for blocklength $k \geq 2$, which is almost intractable. The paper [15] proposed an achievable rate:

$$R \leq \max_{0<\alpha\leq 1} \frac{1}{2}\mathcal{C}\bigg(2\alpha P\Big(1 + \frac{\big(\sqrt{(1-\alpha)/\alpha} + g_{12}g_{23}d\big)^2}{1+g_{23}^2 d^2}\Big)\bigg)$$

where $d = \sqrt{2P_2/(2\alpha s_{13}^2+1)}$.

NNC: When using NNC [5], the achievable rate is:

$$\begin{aligned} &R \leq I(X_1;\hat{Y}_2,Y_3|X_2) - I(\hat{Y}_1;Y_1|X_1,X_2,\hat{Y}_2,Y_3), \\ &R \leq I(X_1,X_2;Y_3) - I(\hat{Y}_2;Y_2|X_2,Y_3) - I(\hat{Y}_1;Y_1|X_1,X_2,Y_3) \end{aligned}$$

for some pmf $P_{X_1}P_{X_2}P_{\hat{Y}_2|X_2Y_2}P_{\hat{Y}_1|X_1Y_1}$. It's easy to check that the optimal choice of Y_1 is $\hat{Y}_1 = \emptyset$, which leads to the compress-forward lower bound (16). The optimal distribution of $\hat{Y}_2$ is generally unknown. Choose $\hat{Y}_2 = Y_2 + Z'$ where $Z' \sim \mathcal{N}(0,\sigma^2)$ and optimise over σ^2. We obtain the achievable rate

$$R \leq \mathcal{C}\Big(s_{13} + \frac{s_{12}s_{23}}{s_{13}+s_{12}+s_{23}+1}\Big). \tag{24}$$

DDF: When using DDF [7], the achievable rate is same as the partial decode-forward lower bound (17). For the Gaussian relay channels, partial decode-forward coding doesn't improve the decode-forward lower bound [15], thus we obtain the achievable rate

$$R \leq \min\left\{\mathcal{C}(s_{13}+s_{23}+2\rho\sqrt{s_{13}s_{23}}), \mathcal{C}\big(s_{12}(1-\rho^2)\big)\right\} \tag{25}$$

for $0 \leq \rho \leq 1$.

Cover–El Gamal's general lower bound [2, Theorem 7]: In [2] Cover and El Gamal proposed a general lower bound for the relay channel by combining compress-forward and decode-forward, which can be written as:

$$\begin{aligned} &R \leq I(X_1;\hat{Y}_2,Y_3|X_2,U_2) + I(U_2;Y_2|V_2,X_2), \\ &R \leq I(X_1,X_2;Y_3) - I(\hat{Y}_2;Y_2|U_2,X_1,X_2,Y_3) \end{aligned} \tag{26}$$

for some pmf $P_{V_2}P_{X_2|V_2}P_{U_2|V_2}P_{X_1|U_2}P_{\hat{Y}_2|X_2U_2Y_2}$ such that $I(\hat{Y}_2;Y_2|U_2,X_1,X_2,Y_3) \leq I(X_2;Y_3|V_2)$.

Choosing the distributions as in [14], we obtain the lower bound with same expression as (22) but subject to the constraint

$$N' \geq (\alpha(s_{13}+s_{23})+1)\frac{(\beta-\alpha\beta+\alpha)s_{13}+1}{\gamma s_{23}(\alpha s_{13}+1)}. \quad (27)$$

Comparing (23) with (27), if

$$\frac{1+\alpha s_{12}}{s_{21}(\alpha(s_{13}+s_{23})+1)} < \frac{(\beta-\alpha\beta+\alpha)s_{13}+1}{s_{23}(\alpha s_{13}+1)} \quad (28)$$

for all $\alpha, \beta \in [0,1]$ (e.g. $s_{21} > s_{23}, s_{12} < s_{13}$), our coding scheme always improves Cover–El Gama's general lower bound [2, Theorem 7]. This general lower bound includes both the partial decode-forward and compress-forward lower bounds [2], thus we have the following corollary:

Corollary 3. *For the enhanced Gaussian relay channel which satisfies* (28)*, our coding scheme improves the known inner bounds, including the NNC and DDF lower bounds and Cover–El Gama's general lower bound [2, Theorem 7].*

Note that both NNC and DDF fail to use Y_1. (In NNC , the optimal choice of $\hat{Y}_1$ is $\hat{Y}_1 = \emptyset$, which means that the transmitter doesn't compress Y_1). As we will see in Theorem 2 (see Section VI), instead of compressing or ignoring Y_1, the transmitter decodes the compression messages sent by the receivers and relays based on Y_1. This is particularly useful when the link from the relay to the transmitter is stronger than the link from the relay to the receiver.

Based on (22–27), the achievable rates for $g_{12} = g_{13} = g_{21} = 1$, $g_{23} = 0.7$, and $P_1 = P_2 = P$ are shown in figure 4. Note that Cover-El Gama's general lower bound is better than NNC but worse than our proposed lower bound (not shown).

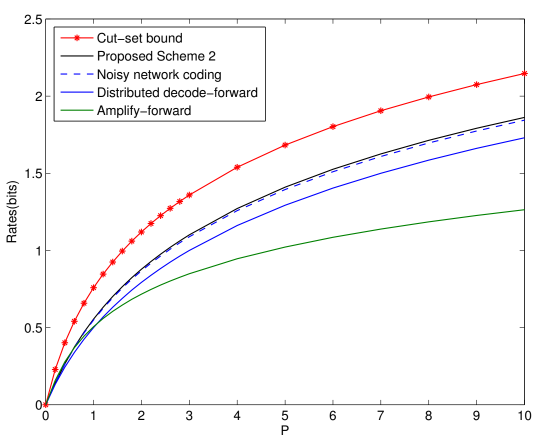

Fig. 4. Achievable rates for the enhanced Gaussian relay channel

Table I compares achievable rates for this enhanced Gaussian relay channel for $g_{12} = 1/d, g_{13} = 1$, $g_{23} = g_{21} = 1/|1-d|$, and with $P_1 = 5, P_2 = 1$. Here R_{NNC}, R_{DDF}, R_{CE}, R_{Pro1} and R_{Pro2} denote rates achieved by NNC, DDF, rates from [2, Theorem 7] and rates from our proposed Theorem 1 and 2, respectively. The feedback scheme (R_{Pro1}) obtains the best performance, and our non-feedback scheme for Theorem 2 (R_{Pro2}) strictly improves the known lower bounds in the absence of feedback.

TABLE I
ACHIEVABLE RATES FOR THE ENHANCED GAUSSIAN RELAY WITH AND WITHOUT FEEDBACK

d	R_{NNC}	R_{DDF}	R_{CE}	R_{Pro1}	R_{Pro2}
0.73	1.6908	1.6881	1.6927	1.7069	1.6996
0.74	1.6971	1.6703	1.6971	1.7111	1.7032
0.75	1.7033	1.6529	1.7033	1.7153	1.7077
0.76	1.7094	1.6358	1.7094	1.7195	1.7129

V. ACHIEVABLE RATES FOR DM MULTIPLE-RELAY CHANNELS WITH PARTIAL FEEDBACK

A. Scheme 1B

Note that in Scheme 1A above, the relays and receivers use only compress-forward. In this subsection we present a scheme where relays perform mixed compress-forward and partial decode-forward.

1) Codebook: Fix pmf in (3). Transmission takes place in $B+1$ blocks each consisting of n transmissions. For block $b \in [1:B]$, split the message m_b into (m'_b, m''_b), where m'_b and m''_b are independently and uniformly distributed over the sets $[1:2^{nR'}]$ and $[1:2^{nR''}]$, respectively, where $R', R'' \geq 0$ and so that $R = R' + R''$. Let $m''_{B+1} = m'_{B+1} = 1$.

For each $r \in \mathcal{R}$ and block $b \in [1:B+1]$, randomly and independently generate $2^{n(R'+\hat{R}_r)}$ sequences $x^n_{r,b}(m'_{b-1}, l_{r,b-1}) \sim \prod_{i=1}^n P_{X_r}(x_{r,b,i})$, with $m'_{b-1} \in [1:2^{nR'}]$ and $l_{r,b-1} \in [1:2^{n\hat{R}_r}]$. For each $(m'_{b-1}, l_{r,b-1})$, randomly and independently generate $2^{nR'}$ sequences $u^n_{r,b}(m'_b|m'_{b-1}, l_{r,b-1}) \sim \prod_{i=1}^n P_{U_r|X_r}(u_{r,b,i}|x_{r,b,i})$. For each $(m'_b, m'_{b-1}, l_{r,b-1})$, randomly and independently generate $2^{n\hat{R}_r}$ sequences $\hat{y}^n_{r,b}(l_{r,b}|m'_b, m'_{b-1}, l_{r,b-1}) \sim \prod_{i=1}^n P_{\hat{Y}_r|U_rX_r}(\hat{y}_{r,b,i}|u_{r,b,i}, x_{r,b,i})$.

For each $d \in \mathcal{D}$ and block $b \in [1:B+1]$, randomly and independently generate $2^{n\hat{R}_d}$ sequences $x^n_{d,b}(l_{d,b-1}) \sim \prod_{i=1}^n P_{X_d}(x_{d,b,i})$, $l_{d,b-1} \in [1:2^{n\hat{R}_d}]$. For each $l_{d,b-1}$, randomly and independently generate $2^{nR'}$ sequences $u^n_{d,b}(m'_b|l_{d,b-1}) \sim \prod_{i=1}^n P_{U_d|X_d}(u_{d,b,i}|x_{d,b,i})$. Similarly, for each $l_{d,b-1}$, randomly and independently generate $2^{n\hat{R}_d}$ sequences $\hat{y}^n_{d,b}(l_{d,b}|l_{d,b-1}) \sim \prod_{i=1}^n P_{\hat{Y}_d|X_d}(\hat{y}_{d,b,i}|x_{d,b,i})$.

For each $(m'_b, m'_{b-1}, \mathbf{l}_{b-1})$, randomly and independently generate $2^{nR''}$ sequences $x^n_{1,b}(m''_b|m'_b, m'_{b-1}, \mathbf{l}_{b-1}) \sim \prod_{i=1}^n P_{X_1|U_2^N X_2^N}(x_{1,b,i}|x_{2,b,i}, u_{2,b,i}, \ldots, x_{N,b,i}, u_{N,b,i})$.

Encoding and decoding are explained with the help of Table II.

2) Source encoding: In each block $b \in [1:B+1]$, assume that the transmitter already knows $\mathbf{l}_{b-1}$ through the feedback links. It sends $x^n_{1,b}(m''_b|m'_b, m'_{b-1}, \mathbf{l}_{b-1})$.

To ensure that the transmitter perfectly knows $\mathbf{l}_{b-1}$, we have

$$\hat{R}_k \leq R_{\text{Fb},k}, \quad \text{for } k \in [2:N]. \quad (29)$$

TABLE II
CODING SCHEME 1B FOR MULTICAST NETWORK WITH PARTIAL FEEDBACK

Block	1	2	...	B	$B+1$
X_1	$x_{1,1}(m_1''\|m_1',1,\mathbf{1})$	$x_{1,2}(m_2''\|m_2',m_1',\mathbf{l}_1)$	...	$x_{1,B}(m_B''\|m_B',m_{B-1}',\mathbf{l}_{B-1})$	$x_{1,B+1}(1\|1,m_B',\mathbf{l}_B)$
X_r	$x_{r,1}(1,1)$	$x_{r,2}(m_1,l_{r,1})$	...	$x_{r,B}(m_{B-1}',l_{r,B-1})$	$x_{r,B+1}(m_B',l_{r,B})$
U_r	$u_{r,1}(m_1'\|1,1)$	$u_{r,2}(m_2'\|m_1,l_{r,1})$	...	$u_{r,B}(m_B'\|m_{B-1}',l_{r,B-1})$	$u_{r,B+1}(1\|m_B',l_{r,B})$
$\hat{Y}_r$	$\hat{y}_{r,1}(l_{r,1}\|1,1)$	$\hat{y}_{r,2}(l_{r,2}\|m_1,l_{r,1})$	...	$\hat{y}_{r,B}(l_{r,B}\|m_{B-1}',l_{r,B-1})$	$\hat{y}_{r,B+1}(1\|m_B',l_{r,B})$
X_d	$x_{d,1}(1)$	$x_{d,2}(l_{d,1})$	...	$x_{d,B}(l_{d,B-1})$	$x_{d,B+1}(l_{d,B})$
U_d	$u_{d,1}(m_1'\|1)$	$u_{d,2}(m_2'\|l_{d,1})$	...	$u_{d,B}(m_B'\|l_{d,B-1})$	$u_{d,B+1}(1\|l_{d,B})$
$\hat{Y}_d$	$\hat{y}_{d,1}(l_{d,1}\|1)$	$\hat{y}_{d,2}(l_{d,2}\|l_{d,1})$	...	$\hat{y}_{d,B}(l_{d,B}\|l_{d,B-1})$	$\hat{y}_{d,B+1}(1\|l_{d,B})$
Y_d	$\hat{m}_1''$	$\leftarrow(\hat{m}_2'',\hat{m}_1',\hat{\mathbf{l}}_1)$	...	$\leftarrow(\hat{m}_B'',\hat{m}_{B-1}',\hat{\mathbf{l}}_{B-1})$	$\leftarrow(\hat{m}_B',\hat{\mathbf{l}}_B)$

3) Relay encoding: Relay nodes perform hybrid compress-forward and decode-forward. For each block $b\in[1:B+1]$, assume that Relay $r\in\mathcal{R}$ already knows $\hat{m}_{b-1}'$ from block $b-1$. It looks for a unique index $\hat{m}_b'$ such that[1]

$$\left(x_{r,b}^n(\hat{m}_{b-1}',l_{r,b-1}),u_{r,b}^n(\hat{m}_b'|\hat{m}_{b-1}',l_{r,b-1}),y_{r,b}^n\right)\in\mathcal{T}_{\epsilon/4}^n(P_{X_rY_rU_r}).$$

then it compresses $y_{r,b}^n$ by finding a unique index $l_{r,b}$ such that

$$\begin{aligned}&\big(u_{r,b}^n(\hat{m}_b'|\hat{m}_{b-1}',l_{r,b-1}),x_{r,b}^n(\hat{m}_{b-1}',l_{r,b-1}),\\&\quad\hat{y}_{r,b}^n(l_{r,b}|\hat{m}_b',\hat{m}_{b-1}',l_{r,b-1}),y_{r,b}^n\big)\in\mathcal{T}_{\epsilon/2}^n(P_{U_rX_rY_r\hat{Y}_r}).\end{aligned}$$

Then, it sends $l_{r,b}$ through the feedback link at rate $\hat{R}_r\le R_{\mathrm{Fb},r}$ and in block $b+1$ sends $x_{r,b+1}^n(\hat{m}_b',l_{r,b})$.

By the covering and packing lemmas, this is successful with high probability if

$$\begin{aligned}&R'<I(U_r;Y_r|X_r)-\delta(\epsilon/4)\\&\hat{R}_r>I(\hat{Y}_r;Y_r|X_r,U_r)+\delta(\epsilon/2),\quad\text{for } r\in\mathcal{R}.\end{aligned}\tag{30}$$

4) Receiver encoding: Receiver $d\in\mathcal{D}$ compresses $y_{d,b}^n$ by finding a unique index $l_{d,b}$ such that

$$\left(x_{d,b}^n(l_{d,b-1}),\hat{y}_{d,b}(l_{d,b}|l_{d,b-1}),y_{d,b}^n\right)\in\mathcal{T}_{\epsilon/2}^n(P_{X_dY_d\hat{Y}_d}).$$

Then, it sends $l_{d,b}$ through the feedback link at rate $\hat{R}_d\le R_{\mathrm{Fb},d}$ and in block $b+1$ sends $x_{d,b+1}^n(l_{d,b})$.

By the covering lemma, this is successful with high probability if

$$\hat{R}_d>I(\hat{Y}_d;Y_d|X_d)+\delta(\epsilon/2),\quad\text{for } d\in\mathcal{D}.\tag{31}$$

5) Decoding: Receiver $d\in\mathcal{D}$ performs backward decoding. For each block $b\in[B+1,\dots,1]$, it looks for $(\hat{m}_b'',\hat{m}_{b-1}',\hat{\mathbf{l}}_{b-1})$ such that [2]

$$\begin{aligned}&\big(x_{1,b}^n(\hat{m}_b''|\hat{m}_b',\hat{m}_b',\hat{\mathbf{l}}_{b-1}),\mathbf{x}_b^n(\mathcal{R}),\mathbf{x}_b^n(\mathcal{D}),\mathbf{u}_b^n(\mathcal{R}),\\&\quad\mathbf{u}_b^n(\mathcal{D}),\hat{\mathbf{y}}_b^n(\mathcal{R}),\hat{\mathbf{y}}_b^n(\mathcal{D}),y_{d,b}^n\big)\in\mathcal{T}_\epsilon^n(P_{X_1^NU_2^N\hat{Y}_2^NY_d})\end{aligned}$$

where $\mathbf{x}_b^n(\mathcal{R}):=[x_{r,b}^n(\hat{m}_{b-1}',\hat{l}_{r,b-1}):r\in\mathcal{R}]$, $\mathbf{x}_b^n(\mathcal{D}):=[x_{d,b}^n(\hat{l}_{d,b-1}):d\in\mathcal{D}]$, $\mathbf{u}_b^n(\mathcal{R}):=[u_{r,b}^n(\hat{m}_b'|\hat{m}_{b-1}',\hat{l}_{r,b-1}):r\in\mathcal{R}]$, $\mathbf{u}_b^n(\mathcal{D}):=[u_{d,b}^n(\hat{m}_b'|\hat{l}_{d,b-1}):d\in\mathcal{D}]$ and $\hat{\mathbf{y}}_b^n(\mathcal{R}):=[\hat{y}_{r,b}^n(\hat{l}_{r,b}|\hat{m}_b',\hat{m}_{b-1}',\hat{l}_{r,b-1}):r\in\mathcal{R}]$, $\hat{\mathbf{y}}_b^n(\mathcal{D}):=[\hat{y}_{d,b}^n(\hat{l}_{d,b}|\hat{l}_{d,b-1}):d\in\mathcal{D}]$.

By the independence of the codebooks, the Markov lemma, packing lemma and the induction on backward decoding, the decoding is successful with high probability if

$$R''<I(X_1;\hat{Y}_2^N,Y_d|U_2^N,X_2^N)-\delta(\epsilon),\tag{32}$$

$$\begin{aligned}R+\hat{R}(\mathcal{T})<&\\&I(X_1,X(\mathcal{T}),U(\mathcal{T});\hat{Y}(\mathcal{T}^c),Y_d|X(\mathcal{T}^c)))\\&+\sum_{k\in\mathcal{R}\cap\mathcal{T}}H(\hat{Y}_k|U_k,X_k)+\sum_{j\in\mathcal{D}\cap\mathcal{T}}H(\hat{Y}_j|X_j)\\&-H(\hat{Y}(\mathcal{T})|X_2^N,U_2^N,\hat{Y}(\mathcal{T}^c),Y_d)-\delta(\epsilon)\end{aligned}\tag{33}$$

satisfying $\mathcal{T}\supseteq\mathcal{R}$, and

$$\begin{aligned}R''+\hat{R}(\mathcal{T})<&\\&I(X_1,X(\mathcal{T}),U(\mathcal{T});\hat{Y}(\mathcal{T}^c),Y_d|X(\mathcal{T}^c)))\\&+\sum_{k\in\mathcal{R}\cap\mathcal{T}}H(\hat{Y}_k|U_k,X_k)+\sum_{j\in\mathcal{D}\cap\mathcal{T}}H(\hat{Y}_j|X_j)\\&-H(\hat{Y}(\mathcal{T})|X_2^N,U_2^N,\hat{Y}(\mathcal{T}^c),Y_d)-\delta(\epsilon)\end{aligned}\tag{34}$$

for all $\mathcal{T}\subset[2:N]$ with $\mathcal{T}^c\cap\mathcal{D}\neq\emptyset$.

For convenience, we ignore the constraints (33) and (34) by introducing the following constraint which is stricter than both (33) and (34):

$$\begin{aligned}R+\hat{R}(\mathcal{T})<&\\&I(X_1,X(\mathcal{T}),U(\mathcal{T});\hat{Y}(\mathcal{T}^c),Y_d|X(\mathcal{T}^c),U(\mathcal{T}^c))\\&+\sum_{k\in\mathcal{R}\cap\mathcal{T}}H(\hat{Y}_k|U_k,X_k)+\sum_{j\in\mathcal{D}\cap\mathcal{T}}H(\hat{Y}_j|X_j)\\&-H(\hat{Y}(\mathcal{T})|X_2^N,U_2^N,\hat{Y}(\mathcal{T}^c),Y_d)-\delta(\epsilon)\end{aligned}\tag{35}$$

for all $\mathcal{T}\subset[2:N]$ with $\mathcal{T}^c\cap\mathcal{D}\neq\emptyset$.

Combining (29–32) and (35), and using Fourier-Motzkin elimination to eliminate $R',R'',\hat{R}_2,\dots,\hat{R}_N$, we obtain Theorem 1.

VI. DISCRETE MEMORYLESS MULTICAST NETWORK

In Section V we proposed a block-Markov coding scheme for DM multicast networks in the presence of instantaneous, rate-limited and noisy-free feedback. Recall the NNC scheme

[1] Since each Relay $r\in\mathcal{R}$ makes its own estimate of m_b', the precise notation $\hat{m}_b'^{(r)}$. For simplicity, we omit the superscript (r).

[2] Receiver $d\in\mathcal{D}$ knows $l_{d,b-1}$ since it generated this index. Since each Receiver d makes its own estimate of $(m_b'',m_{b-1}',\mathbf{l}_{b-1})$, the precise notation is $(\hat{m}_b''^{(d)},\hat{m}_{b-1}'^{(d)},\hat{\mathbf{l}}_{b-1}^{(d)})$. For simplicity, we omit the superscript (d).

[4], [5], [6] for DM multicast networks without feedback, where each node (including the transmitter) compresses its observation and sends the new compression index in the next block. Comparing our coding scheme with NNC, we observe that both schemes involve block-Markov coding, compressing channel outputs and sending compression messages. However, our scheme allows hybrid relaying strategies at relay nodes, and in each block, instead of creating a new compression index, the transmitter forwards all compression indices sent by receivers and relays from the previous block. In our scheme, different nodes operate differently according to the features of the network, which leads to a larger achievable rate than NNC, as shown by examples in Section IV.

Motivated by our feedback coding scheme, we propose another scheme for the N-node DM multicast networks *without* feedback. The main idea is as follows: in each block b, each node $k \in [2:N]$ creates a compression index $l_{k,b-1}$ and sends $(l_{k,b-1}, l_{k,b-2})$. The transmitter, after observing $Y_{1,b}^n$, first decodes compression indices $\mathbf{l}_{b-1}$, which is in essence a coding problem on a multiple access channel $P_{Y_1|X_2,\ldots,X_N}$ with side information X_1. Then in block $b+1$, the transmitter sends compression messages $\mathbf{l}_{b-1}$ with source message m_{b+1}.

1) Codebook: Fix the pmf in (9). Transmission takes place in $B+2$ blocks each consisting of n transmissions. For block $b \in [1:B]$, split the message m_b into (m'_b, m''_b), where m'_b and m''_b are independently and uniformly distributed over the sets $\in [1:2^{nR'}]$ and $[1:2^{nR''}]$, respectively, where $R', R'' \geq 0$ and so that $R = R' + R''$. Let $\mathbf{l}_{-1} = \mathbf{l}_0 = \mathbf{1}_{[N-1]}$ and $m''_{B+1} = m'_{B+1} = m''_{B+2} = m'_{B+2} = 1$.

For each $r \in \mathcal{R}$ and block $b \in [1:B+2]$, randomly and independently generate $2^{n(R'+\hat{R}_r)}$ sequences $v_{r,b}^n(m'_{b-1}, l_{r,b-2}) \sim \prod_{i=1}^n P_{V_r}(v_{r,b,i})$, with $m'_{b-1} \in [1:2^{nR'}]$ and $l_{r,b-2} \in [1:2^{n\hat{R}_r}]$. For each $(m'_{b-1}, l_{r,b-2})$, randomly and independently generate $2^{n\hat{R}_r}$ sequences $x_{r,b}^n(l_{r,b-1}|m'_{b-1}, l_{r,b-2}) \sim \prod_{i=1}^n P_{X_r|V_r}(x_{r,b,i}|v_{r,b,i})$. For each pair $(m'_{b-1}, l_{r,b-2})$, randomly and independently generate $2^{nR'}$ sequences $u_{r,b}^n(m'_b|m'_{b-1}, l_{r,b-2}) \sim \prod_{i=1}^n P_{U_r|V_r}(u_{r,b,i}|v_{r,b,i})$. For each $(m'_b, m'_{b-1}, l_{r,b-2}, l_{r,b-1})$, randomly and independently generate $2^{n\hat{R}_r}$ sequences $\hat{y}_{r,b}^n(l_{r,b}|m'_b, m'_{b-1}, l_{r,b-2}, l_{r,b-1}) \sim \prod_{i=1}^n P_{\hat{Y}_r|U_r X_r V_r}(\hat{y}_{r,b,i}|u_{r,b,i}, x_{r,b,i}, v_{r,b,i})$.

For each $d \in \mathcal{D}$ and block $b \in [1:B+2]$, randomly and independently generate $2^{n\hat{R}_d}$ sequences $v_{d,b}^n(l_{d,b-2}) \sim \prod_{i=1}^n P_{V_d}(v_{d,b,i})$, with $l_{d,b-2} \in [1:2^{n\hat{R}_d}]$. For each $l_{d,b-2}$, randomly and independently generate $2^{n\hat{R}_d}$ sequences $x_{d,b}^n(l_{d,b-1}|l_{d,b-2}) \sim \prod_{i=1}^n P_{X_d|V_d}(x_{d,b,i}|v_{d,b,i})$. For each $l_{d,b-2}$, randomly and independently generate $2^{nR'}$ sequences $u_{d,b}^n(m'_b|l_{d,b-2}) \sim \prod_{i=1}^n P_{U_d|V_d}(u_{d,b,i}|v_{d,b,i})$. For each $(l_{d,b-2}, l_{d,b-1})$, randomly and independently generate $2^{n\hat{R}_d}$ sequences $\hat{y}_{d,b}^n(l_{d,b}|l_{d,b-2}, l_{d,b-1}) \sim \prod_{i=1}^n P_{\hat{Y}_d|X_d V_d}(\hat{y}_{d,b,i}|x_{d,b,i}, v_{d,b,i})$.

For each $(m'_b, m'_{b-1}, \mathbf{l}_{b-2})$, randomly and independently generate $2^{nR''}$ sequences $x_{1,b}^n(m''_b|m'_b, m'_{b-1}, \mathbf{l}_{b-2}) \sim \prod_{i=1}^n P_{X_1|U_2^N V_2^N}(x_{1,b,i}|v_{2,b,i}, u_{2,b,i}, \ldots, v_{N,b,i}, u_{N,b,i})$.

Let

$$\begin{aligned}
\mathbf{v}'^n_b(\mathcal{R}) &:= [v_{r,b}^n(\hat{m}'_{b-1}, \hat{l}_{r,b-2}), r \in \mathcal{R}] \\
\mathbf{v}'^n_b(\mathcal{D}) &:= [v_{d,b}^n(\hat{l}_{d,b-2}), d \in \mathcal{D}] \\
\mathbf{x}'^n_b(\mathcal{R}) &:= [x_{r,b}^n(\hat{l}_{r,b-1}|\hat{m}'_{b-1}, \hat{l}_{r,b-2}) : r \in \mathcal{R}] \\
\mathbf{x}'^n_b(\mathcal{D}) &:= [x_{d,b}^n(\hat{l}_{d,b-1}|\hat{l}_{d,b-2}) : d \in \mathcal{D}] \\
\mathbf{u}'^n_b(\mathcal{R}) &:= [u_{r,b}^n(\hat{m}'_b|\hat{m}'_{b-1}, \hat{l}_{r,b-2}) : r \in \mathcal{R}] \\
\mathbf{u}'^n_b(\mathcal{D}) &:= [u_{d,b}^n(\hat{m}'_b|\hat{l}_{d,b-2}) : d \in \mathcal{D}] \\
\hat{\mathbf{y}}'^n_b(\mathcal{R}) &:= [\hat{y}_{r,b}^n(\hat{l}_{r,b}|\hat{m}'_b, \hat{m}'_{b-1}, \hat{l}_{r,b-2}, \hat{l}_{r,b-1}) : r \in \mathcal{R}] \\
\hat{\mathbf{y}}'^n_b(\mathcal{D}) &:= [\hat{y}_{d,b}^n(\hat{l}_{d,b}|\hat{l}_{d,b-2}, \hat{l}_{d,b-1}) : d \in \mathcal{D}].
\end{aligned}$$

2) Source encoding: At each block $b \in [1:B+1]$, after observing $Y_{1,b}^n$, it looks for $\hat{\mathbf{l}}_{b-1}$ such that

$$\begin{aligned}
&\big(x_{1,b}^n(m''_b|m'_b, m'_{b-1}, \hat{\mathbf{l}}_{b-2}), \mathbf{v}'^n_b(\mathcal{R}), \mathbf{v}'^n_b(\mathcal{D}), \mathbf{x}'^n_b(\mathcal{R}), \\
&\quad \mathbf{x}'^n_b(\mathcal{D}), \mathbf{u}'^n_b(\mathcal{R}), \mathbf{u}'^n_b(\mathcal{D}), y_{1,b}^n\big) \in \mathcal{T}_{\epsilon/8}^n(P_{V_2^N X_1^N U_2^N Y_1})
\end{aligned}$$

where $\hat{m}''_b = m''_b$, $\hat{m}'_b = m'_b$ and $\hat{m}'_{b-1} = m'_{b-1}$ in (36) since the transmitter knows the source messages it sent.

After finding compression indices $\hat{\mathbf{l}}_{b-1}$, in block $b+1$ the transmitter sends $x_{1,b+1}^n(m''_{b+1}|m'_{b+1}, m'_b, \hat{\mathbf{l}}_{b-1})$.

By the packing lemma, this step is successful with high probability if for $\mathcal{T} \subseteq [2:N]$, we have

$$\hat{R}(\mathcal{T}) < I(X(\mathcal{T}); Y_1|X(\mathcal{T}^c), V_2^N, U_2^N, X_1) - \delta(\epsilon/8). \quad (36)$$

3) Relay encoding: Relay nodes perform mixed compress-forward and partial decode-forward. In each block $b \in [1:B+1]$, Relay $r \in \mathcal{R}$ looks for a unique index $\hat{m}'_b$ such that[3]

$$\begin{aligned}
&\big(v_{r,b}^n(\hat{m}'_{b-1}, l_{r,b-2}), x_{r,b}^n(l_{r,b-1}|\hat{m}'_{b-1}, l_{r,b-2}), \\
&\quad u_{r,b}^n(\hat{m}'_b|\hat{m}'_{b-1}, l_{r,b-2}), y_{r,b}^n\big) \in \mathcal{T}_{\epsilon/6}^n(P_{X_r Y_r U_r V_r}),
\end{aligned}$$

then it compresses $y_{r,b}^n$ by finding a unique index $l_{r,b}$ such that

$$\begin{aligned}
&\big(v_{r,b}^n, u_{r,b}^n, x_{r,b}^n, y_{r,b}^n, \\
&\quad \hat{y}_{r,b}^n(l_{r,b}|\hat{m}'_b, \hat{m}'_{b-1}, l_{r,b-2}, l_{r,b-1})\big) \in \mathcal{T}_{\epsilon/4}^n(P_{V_r U_r X_r Y_r \hat{Y}_r}).
\end{aligned}$$

Then, in block $b+1$ it sends $x_{r,b+1}^n(l_{r,b}|\hat{m}'_b, l_{r,b-1})$.

By the covering and packing lemma, this step is successful with high probability if

$$\begin{aligned}
R' &< I(U_r; Y_r|V_r, X_r) - \delta(\epsilon/6) \\
\hat{R}_r &> I(\hat{Y}_r; Y_r|V_r, X_r, U_r) + \delta(\epsilon/4), \quad \text{for } r \in \mathcal{R}. \quad (37)
\end{aligned}$$

4) Receiver encoding: Receiver $d \in \mathcal{D}$ compresses $y_{d,b}^n$ by finding a unique index $l_{d,b}$ such that

$$\begin{aligned}
&\big(v_{d,b}^n(l_{d,b-2}), x_{d,b}^n(l_{d,b-1}|l_{d,b-2}), \\
&\quad \hat{y}_{d,b}(l_{d,b}|l_{d,b-2}, l_{d,b-1}), y_{d,b}^n\big) \in \mathcal{T}_{\epsilon/4}^n(P_{V_d X_d Y_d \hat{Y}_d}).
\end{aligned}$$

Then, in block $b+1$ it sends $x_{d,b+1}^n(l_{d,b}|l_{d,b-1})$.

By the covering and packing lemmas, this step is successful with high probability if

$$\hat{R}_d > I(\hat{Y}_d; Y_d|V_d, X_d) + \delta(\epsilon/4), \quad \text{for } d \in \mathcal{D}. \quad (38)$$

[3]Since each Relay $r \in \mathcal{R}$ makes its own estimate of m'_b, thus the precise notation should be $\hat{m}'^{(r)}_b$. For simplicity, we omit the superscript (r).

TABLE III

CODING SCHEME FOR MULTICAST NETWORK WITHOUT FEEDBACK

Block	1	...	B	$B+1$	$B+2$
X_1	$x_{1,1}(m''_1\|m'_1,1,\mathbf{1})$	...	$x_{1,B}(m''_B\|m'_B,m'_{B-1},\mathbf{l}_{B-2})$	$x_{1,B+1}(1\|1,m'_B,\mathbf{l}_{B-1})$	$x_{1,B+2}(1\|1,1,\mathbf{l}_B)$
V_r	$v_{r,1}(1,1)$	...	$v_{r,B}(m'_{B-1},l_{r,B-2})$	$v_{r,B+1}(m'_B,l_{r,B-1})$	$v_{r,B+2}(1,l_{r,B})$
X_r	$x_{r,1}(1\|1,1)$	...	$x_{r,B}(l_{r,B-1}\|m'_{B-1},l_{r,B-2})$	$x_{r,B+1}(l_{r,B}\|\hat{m}'_B,l_{r,B-1})$	$x_{r,B+2}(1\|1,l_{r,B})$
U_r	$u_{r,1}(m'_1\|1,1)$	...	$u_{r,B}(m'_B\|m'_{B-1},l_{r,B-2})$	$u_{r,B+1}(1\|m'_B,l_{r,B-1})$	$u_{r,B+2}(1\|1,l_{r,B})$
$\hat{Y}_r$	$\hat{y}_{r,1}(l_{r,1}\|m'_1,1,1,1)$	...	$\hat{y}_{r,B}(l_{r,B}\|m'_{B-1},m'_B,l_{r,B-2},l_{r,B-1})$	$\hat{y}_{r,B+1}(1\|m'_B,1,l_{r,B-1},l_{r,B})$	$\hat{Y}_{r,B+2}(1\|1,1,l_{r,B},1)$
V_d	$v_{d,1}(1)$	...	$v_{d,B}(l_{d,B-2})$	$v_{d,B+1}(l_{d,B-1})$	$v_{d,B+2}(l_{d,B})$
U_d	$u_{d,1}(m'_1\|1)$	...	$u_{d,B}(m'_B\|l_{d,B-2})$	$u_{d,B+1}(1\|l_{d,B-1})$	$u_{d,B+2}(1\|l_{d,B})$
X_d	$x_{d,1}(1\|1)$	...	$x_{d,B}(l_{d,B-1}\|l_{d,B-2})$	$x_{d,B+1}(l_{d,B}\|l_{d,B-1})$	$x_{d,B+2}(1\|l_{d,B})$
$\hat{Y}_d$	$\hat{y}_{d,1}(l_{d,1}\|1,1)$	...	$\hat{y}_{d,B}(l_{d,B}\|l_{d,B-2},l_{d,B-1})$	$\hat{y}_{d,B+1}(1\|l_{d,B-1},l_{d,B})$	$\hat{y}_{d,B+2}(1\|l_{d,B},1)$
Y_d	$\hat{m}''_1$	...	$\leftarrow(\hat{m}''_B,\hat{m}'_{B-1},\hat{\mathbf{l}}_{B-2})$	$\leftarrow(\hat{m}'_B,\hat{\mathbf{l}}_{B-1})$	$\leftarrow\hat{\mathbf{l}}_B$

5) Decoding: Receiver $d \in \mathcal{D}$ performs backward decoding. For each block $b \in [B+2,\ldots,1]$, it looks for $(\hat{m}''_b, \hat{m}'_{b-1}, \hat{\mathbf{l}}_{b-2})$ such that[4]

$$\big(x^n_{1,b}(\hat{m}''_b|\hat{m}'_b,\hat{m}'_{b-1},\hat{\mathbf{l}}_{b-2}), \mathbf{v}'^n_b(\mathcal{R}), \mathbf{v}'^n_b(\mathcal{D}), \mathbf{x}'^n_b(\mathcal{R}), \mathbf{x}'^n_b(\mathcal{D}), \mathbf{u}'^n_b(\mathcal{R}), \mathbf{u}'^n_b(\mathcal{D}), \hat{\mathbf{y}}'^n_b(\mathcal{R}), \hat{\mathbf{y}}'^n_b(\mathcal{D}), y^n_{d,b}\big) \in \mathcal{T}^n_\epsilon(P_{V^N_2 X^N_1 U^N_2 \hat{Y}^N_2 Y_d}).$$

By the independence of the codebooks, the Markov lemma, packing lemma and induction on backward decoding, the decoding is successful with high probability if

$$R'' < I(X_1;\hat{Y}^N_2, Y_d|X^N_2, U^N_2, V^N_2) - \delta(\epsilon) \tag{39}$$

and

$$\begin{aligned} &R + \hat{R}(\mathcal{T}) < \\ &I(X_1,V(\mathcal{T}),U(\mathcal{T}),X(\mathcal{T});\hat{Y}(\mathcal{T}^c),Y_d|V(\mathcal{T}^c),X(\mathcal{T}^c),U(\mathcal{T}^c)) \\ &\quad + \sum_{k\in\mathcal{R}\cap\mathcal{T}} H(\hat{Y}_k|X_k,U_k,V_k) + \sum_{j\in\mathcal{D}\cap\mathcal{T}} H(\hat{Y}_j|X_j,V_j) \\ &\quad -H(\hat{Y}(\mathcal{T})|V^N_2,X^N_1,U^N_2,\hat{Y}(\mathcal{T}^c),Y_d) - \delta(\epsilon) \end{aligned} \tag{40}$$

such that $\mathcal{T} \supseteq \mathcal{R}$, and

$$\begin{aligned} &R'' + \hat{R}(\mathcal{T}) < \\ &I(X_1,V(\mathcal{T}),U(\mathcal{T}),X(\mathcal{T});\hat{Y}(\mathcal{T}^c),Y_d|V(\mathcal{T}^c),X(\mathcal{T}^c),U(\mathcal{T}^c)) \\ &\quad + \sum_{k\in\mathcal{R}\cap\mathcal{T}} H(\hat{Y}_k|X_k,U_k,V_k) + \sum_{j\in\mathcal{D}\cap\mathcal{T}} H(\hat{Y}_j|X_j,V_j) \\ &\quad -H(\hat{Y}(\mathcal{T})|V^N_2,X^N_1,U^N_2,\hat{Y}(\mathcal{T}^c),Y_d) - \delta(\epsilon) \end{aligned} \tag{41}$$

for all $\mathcal{T} \subset [2:N]$ with $\mathcal{T}^c \cap \mathcal{D} \neq \emptyset$.

For convenience, we ignore the constraints (40) and (41) by introducing the following constraint which is stricter than both (40) and (41):

$$\begin{aligned} &R + \hat{R}(\mathcal{T}) < \\ &I(X_1,V(\mathcal{T}),U(\mathcal{T}),X(\mathcal{T});\hat{Y}(\mathcal{T}^c),Y_d|V(\mathcal{T}^c),X(\mathcal{T}^c),U(\mathcal{T}^c)) \\ &\quad + \sum_{k\in\mathcal{R}\cap\mathcal{T}} H(\hat{Y}_k|X_k,U_k,V_k) + \sum_{j\in\mathcal{D}\cap\mathcal{T}} H(\hat{Y}_j|X_j,V_j) \\ &\quad -H(\hat{Y}(\mathcal{T})|V^N_2,X^N_1,U^N_2,\hat{Y}(\mathcal{T}^c),Y_d) - \delta(\epsilon) \end{aligned} \tag{42}$$

for all $\mathcal{T} \subset [2:N]$ with $\mathcal{T}^c \cap \mathcal{D} \neq \emptyset$. .

Combining (36–39) and (42), and using Fourier-Motzkin elimination to eliminate $R', R'', \hat{R}_2, \ldots, \hat{R}_N$, we obtain Theorem 2.

[4]Receiver $d \in \mathcal{D}$ knows $l_{d,b-2}$ since it generated itself. Since each Receiver $d \in$ makes its own estimate of $(m''_b, m'_{b-1}, \mathbf{l}_{b-2})$, the precise notation is $(\hat{m}''^{(d)}_b, \hat{m}'^{(d)}_{b-1}, \hat{\mathbf{l}}^{(d)}_{b-2})$. For simplicity, we omit the superscript (d).

REFERENCES

[1] E. C. van der Meulen, “Three terminal communication channels,” *Adv. Appl. Probab.*, vol. 3, pp. 120–154, 1971.

[2] 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.

[3] G. Kramer, M. Gastpar and P. Gupta, ”Cooperative strategies and capacity theorems for relay networks,” *IEEE Trans. Inf. Theory*, vol. 51, no. 9, pp. 3037–3063, Sep. 2005.

[4] M. Yassaee and M. Aref, “Generalized compress-and-forward strategy for relay networks,” *Information Theory (ISIT), 2008 IEEE International Symposium on*, Toronto, Canada, pp. 2683–2687, July 2008.

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

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

[7] S. Lim, K. T. Kim and Y-Han Kim, “Distributed decode-forward for multicast,” *Information Theory Workshop (ITW), 2014 IEEE*, pp. 556–560, Nov. 2014

[8] S. Lim, K. T. Kim and Y-Han Kim, “Distributed decode-forward for broadcast,” *Information Theory (ISIT), 2014 IEEE International Symposium on*, pp. 636–640, June, 2014

[9] B. Schein and R. Gallager, “The Gaussian parallel relay network,” *Information Theory (ISIT), 2000 IEEE International Symposium on*, pp. 22, June, 2000

[10] Y-Han Kim, S. Lim and P. Minero, “Relaying via hybrid coding,” *Information Theory (ISIT), 2011 IEEE International Symposium on*, pp. 1881–1885, July, 2011

[11] Y. Gabbai and S. I. Bross, “Achievable rates for the discrete memoryless relay channel with partial feedback configurations,” *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 4989–5007, Nov. 2006

[12] F. M. J. Willems and E. C. van der Meulen, “Partial feedback for the discrete memoryless multiple-access channel,” *IEEE Trans. Inf. Theory*, vol. 29, no. 2, pp. 287–290, Mar. 1983

[13] Y. Wu, M. Wigger, “Coding schemes for discrete memoryless broadcast channels with rate–limited feedback,” *Information Theory (ISIT), 2014 IEEE International Symposium on*, pp.2127–2131, June 2014.

[14] H. Chong, M. Motani and H. Garg, “New coding strategies for the relay channel,” *Information Theory (ISIT), 2005 IEEE International Symposium on*, pp. 1086–1090, Sept. 2005

[15] S. Zahedi, “On reliable communication over relay channels,” Ph.D thesis. Stanford: Stanford University, 2005.

[16] A. El Gamal and Y-Han Kim, *Network information theory.* Cambridge, U.K.: Cambridge Univ. Press, 2011.