The Three-Node Wireless Network: Achievable Rates and Cooperation Strategies

Lifeng Lai, Ke Liu, *Member, IEEE*, and Hesham El Gamal, *Senior Member, IEEE*

Abstract—We consider a wireless network composed of three nodes and limited by the half-duplex and total power constraints. This formulation encompasses many of the special cases studied in the literature and allows for capturing the common features shared by them. Here, we focus on three special cases, namely, 1) relay channel, 2) multicast channel, and 3) three-way channel. These special cases are judicially chosen to reflect varying degrees of complexity while highlighting the common ground shared by the different variants of the three-node wireless network. For the relay channel, we propose a new cooperation scheme that exploits the wireless feedback gain. This scheme combines the benefits of the decode-and-forward (DF) and compress-and-forward (CF) strategies and avoids the noiseless feedback assumption adopted in earlier works. Our analysis of the achievable rate of this scheme reveals the diminishing feedback gain in both the low and high signal-to-noise ratio (SNR) regimes. Inspired by the proposed feedback strategy, we identify a greedy cooperation framework applicable to both the multicast and three-way channels. Our performance analysis reveals the asymptotic optimality of the proposed greedy approach and the central role of list source–channel decoding in exploiting the receiver side information in the wireless network setting.

Index Terms—List source–channel decoding, multicast channel, noisy feedback, relay channel, three-way channel.

WE are in the midst of a new wireless revolution, brought on by the adoption of wireless networks for consumer, military, scientific, and wireless applications. For example, the consumer potential is clearly evident in the exploding popularity of wireless local-area networks (LANs) and Bluetooth-protocol devices. The military potential is also clear: wireless networks can be rapidly deployed, and the failure of individual nodes does not imply the failure of the network. Scientific data-collection applications using wireless sensor networks are also gaining in numbers. These applications have sparked a renewed interest in network information theory. Despite the recent progress (see [1]–[9] and references therein), developing a unified theory for network information flow remains an elusive task.

In our work, we consider, perhaps, the most simplified scenario of wireless networks. Our network is composed of only three nodes and limited by the half-duplex and total power constraints. Despite this simplicity, this model encompasses many of the special cases that have been extensively studied in the literature. These special *channels*[1] are induced by the traffic generated at the nodes and the requirements imposed on the network.[2] More importantly, this model exposes the common features shared by these special cases and allows for constructing universal cooperation strategies that yield significant performance gains. In particular, we focus here on three special cases, namely 1) relay channel, 2) multicast channel, and 3) three-way channel. These channels are defined rigorously in Section I. We adopt a greedy framework for designing cooperation strategies and characterize the achievable rates of the proposed schemes. Our analysis reveals the structural similarities of the proposed strategies, in the three special cases, and establishes the asymptotic optimality of such strategies in several cases. More specifically, our contributions can be summarized as follows.

1) We propose a novel cooperation strategy for the relay channel with feedback. Our scheme combines the benefits of both the decode-and-forward (DF) and compress-and-forward (CF) strategies and avoids the noiseless feedback assumption adopted in earlier works. Our analysis of the achievable rate of the proposed strategy reveals the diminishing gain of feedback in the asymptotic scenarios of low and large signal-to-noise ratio (SNR).

2) Inspired by the feedback strategy for the relay channel, we construct a new cooperation strategy for the multicast scenario. Motivated by a greedy approach, we show that the *weak* receiver is led to help the *strong* receiver first.[3] Based on the same greedy motivation, the strong user starts to assist the weak receiver after successfully decoding the transmitted codeword. We compute the corresponding achievable rate and use it to establish the significant gains offered by this strategy, as compared with the noncooperative scenario.

3) Motivated by the sensor networks application, we identify the three-way channel model as a special case of our

[1]With a slight abuse of notation, we interchange "channel" and "network" in different places of the sequel for maximal consistency with the literature.

[2]For example, the relay channel corresponds to the special case where the traffic is generated at one node and is required to be transmitted to only one of the remaining two nodes.

[3]The notions of weak and strong receivers will be defined rigorously in the sequel.

general formulation. In this model, the three nodes observe correlated date streams and every node wishes to communicate its observations to the other two nodes. Our proposed cooperation strategy in this scenario consists of three stages of *multicast with side information*, where the multicasting order is determined by a low-complexity greedy scheduler. In every stage, we use a cooperation strategy obtained as a generalization of the greedy multicast approach. This strategy highlights the central role of list source–channel decoding in exploiting the side information available at the receivers. By contrasting the minimum energy required by the proposed strategy with the genie-aided and noncooperative schemes, we establish its superior performance.

4) In summary, we identify the greedy principle as the basis for constructing efficient cooperation strategies in the three considered scenarios. Careful consideration of other variants of the three-node network reveals the fact that such principle carries over with slight modifications.

The rest of the paper is organized as follows. Section I introduces our modeling assumptions and notation. In Section II, we present the new cooperation strategy for the wireless relay channel with *noisy* feedback and analyze its performance. Building on the relay channel strategy, Section III develops the greedy cooperation framework for the multicast channel. We devote Section IV to the three-way channel. Finally, we offer some concluding remarks in Section V. To enhance the flow of the paper, all the proofs are collected in the Appendices.

I. THE THREE NODE WIRELESS NETWORK

Fig. 1 illustrates a network consisting of three nodes each observing a different source. In the general case, the three sources can be correlated. Nodes are interested in obtaining a subset or all the source variables at the other nodes. To achieve this goal, nodes are allowed to coordinate and exchange information over the wireless channel. Different variants of this problem have been previously considered in the literature. For example, [12], [13] investigate the capacity region for general multiterminal networks, [14] studies the multiterminal source coding problem, whereas [15] discusses the cost–distortion tradeoff in a similar setup. In this paper, we focus on the half-duplex wireless setting and assume that each node generates only one source sequence. This source sequence must be reconstructed losslessly at one of the two other nodes, or both of them, with (or without) receiver side information. Although this model is not the most general, it encompasses many important wireless communication scenarios as argued in the sequel. Mathematically, the three node wireless network studied in this paper consists of following elements.

1) The three sources $S_i, i = 1, 2, 3$, drawn independent and identically from certain known joint distribution $p(s_1, s_2, s_3)$ over a finite set $\mathcal{S}_1 \times \mathcal{S}_2 \times \mathcal{S}_3$. We denote by S_i^K the length-K discrete source sequence $S_i(1), \ldots, S_i(K)$ at the ith node. Throughout the sequel, we use capital letters to refer to random variables and lower case letters for realizations.
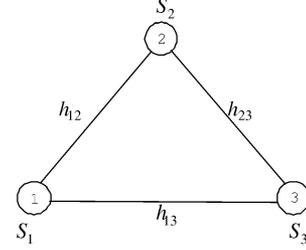
Fig. 1. An illustration of the three-node (half-duplex) wireless network. Each node may be interested in a subset or all the observation variables distributed across the network.

2) We consider the discrete-time *additive white Gaussian noise* (AWGN) channel. At time instant n, node j receives

$$Y_j(n) = \sum_{i \neq j} h_{ij} X_i(n) + Z_j(n) \qquad (1)$$

where $X_i(n)$ is the transmitted signal by node-i and h_{ij} is the channel coefficient from node i to j. To simplify the discussion, we assume the channel coefficients are symmetric, i.e., $h_{ij} = h_{ji}$. These channel gains are assumed to be known *a priori* at the three nodes. We also assume that the additive zero-mean Gaussian noise is spatially and temporally white and has the same unit variance ($\sigma^2 = 1$).

3) We consider half-duplex nodes that cannot transmit and receive *simultaneously* using the same degree of freedom. Without loss of generality, we split the degrees of freedom available to each node in the temporal domain, so that, at each time instant n, a node-i can either transmit (T-mode, $Y_i(n) = 0$) or receive (R-mode, $X_i(n) = 0$), but never both. Due to the half-duplex constraint, at any time instant, the network nodes are divided into two groups: the T-mode nodes (denoted by $\mathcal{T}$) and the R-mode nodes ($\mathcal{R}$). A partition ($\mathcal{T}, \mathcal{R}$) is called a network state.

4) Let $P_i^{(l)}$ denote the average transmit power at the ith node during the m_l network state. We adopt a short-term power constraint such that the total power of all the T-mode nodes at any network state is limited to P, that is,

$$\sum_{i \in \mathcal{T}_l} P_i^{(l)} \leq P, \qquad \forall m_l. \qquad (2)$$

This short-term constraint avoids large peak powers and simplifies the power allocation algorithm (at the expense of a possible performance loss).

5) We associate with node-i an index set I_i, such that $j \in I_i$ indicates that node-i is interested in obtaining S_j from node-j ($j \neq i$).

6) At node-i, a causal joint source–channel encoder converts a length-K block of source sequence into a length-N codeword. The encoder output at time n is allowed to depend on the received signal in the previous $n - 1$ instants, i.e.,

$$X_i(n) = f_i\left(n, S_i^K, Y_i^{n-1}\right). \qquad (3)$$

In the special case of a separate source–channel coding approach, the encoder decomposes into the following.

- A source encoder f_{si} maps S_i^K into a node message W_i, i.e., $W_i = f_{si}(S_i^K)$, $W_i \in [1, M_i]$.
- A channel encoder $f_{ci}(n)$ encodes the node message into a channel input sequence $X_i(n) = f_{ci}(n, W_i, Y_i^{n-1})$.

7) At node-i, decoder d_i estimates the source variables indexed by I_i

$$\{\hat{S}_{ij}^K\} = d_i\left(Y_i^N, S_i^K\right), \qquad \forall j \in I_i \tag{4}$$

where $\hat{S}_{ij}^K$ denotes the estimation of S_j^K at node i. In the case of a separate source–channel coding scheme, decoder d_i consists of the following.

- A channel decoder d_{ci}, $\hat{W}_{ij} = d_{ci}(Y_i^N)$.
- A source decoder d_{si}, $\hat{S}_{ij}^K = d_{si}(\hat{W}_{ij}, S_i^K)$.

8) A decoding error is declared if any node fails to reconstruct its intended source variables correctly. Thus, the joint error probability can be expressed as

$$P_e^{N,K} = \mathrm{Prob}\left\{ \bigcup_{j \in I_i, i=1,2,3} \left\{\hat{S}_{ij}^K \neq S_j^K\right\} \right\}. \tag{5}$$

In the case of a separate coding scheme, the error probability of the channel coding scheme is

$$P_e^N = \mathrm{Prob}\{ \bigcup_{j \in I_i, i=1,2,3} \{\hat{W}_{ij} \neq W_j\} \}. \tag{6}$$

9) An efficient cooperation strategy should strive to maximize the achievable rate given by $\frac{KH(S_1, S_2, S_3)}{N}$, where N is the minimum number of channel uses necessary to satisfy the network requirements. For a fixed $H(S_1, S_2, S_3)$, this optimization is equivalent to minimizing the bandwidth expansion factor $\tau = \frac{N}{K}$.[4] Due to a certain additive property, using the bandwidth expansion factor will be more convenient in the three-way channel scenario. A bandwidth expansion factor τ is said to be achievable if there exists a series of source–channel codes with $N, K \to \infty$ but $\frac{N}{K} \to \tau$, such that $P_e^{N,K} \to 0$. In the feedback-relay and multicast channel, minimizing the bandwidth expansion factor reduces to the more conventional concept of maximizing the rate given by $R = \frac{\log_2(M)}{N}$, where M is the size of message set at the source node.

10) Throughout the sequel we will use the shorthand notation

$$C(x) = \frac{1}{2}\log(1 + x). \tag{7}$$

The model described here encompasses many important network communication scenarios with a wide range of complexity, controlled by various configurations of the index sets and the sources. From this perspective, the relay channel represents the simplest situation where one node serves as the relay for the other source–destination pair, e.g., $\mathcal{S}_2 = \mathcal{S}_3 = \phi$,

[4]The bandwidth expansion factor terminology is motivated by the real-time application where the bandwidth of the channel must be N/K times the bandwidth of the source process.

$I_1 = I_2 = \phi$, and $I_3 = \{1\}$. If we enlarge the index set $I_2 = \{1\}$, meaning node-2 now is also interested in obtaining the source message, then the problem becomes the multicast channel. Furthermore, if the two receivers (nodes 2 and 3) in the multicast case have additional observations, i.e., S_2 and S_3, which are correlated with the source variable S_1, then the problem generalizes to the so-called multicast with side information. We refer to the most complex scenario as the three-way channel. In this scenario, the three sources are correlated and every node attempts to reconstruct the other two sources, i.e., $I_i = \{1, 2, 3\}\backslash\{i\}$. While it is easy to envision other variants of the three-node network, we limit ourselves to these special cases. This choice stems from our belief that other scenarios do not add further insights to our framework. For example, another variant of the feedback-relay channel would allow the relay to observe its own side information. Careful consideration of this case, however, shows that our analysis in Section II extends to this case with only slight modifications. Similarly, inspired by our modular approach for the three-way channel, one can decompose the multiple-access channel with correlated sources into two stages of a feedback-relay channel with side information.

II. THE FEEDBACK-RELAY CHANNEL

Our formulation for the three-node network allows for a more realistic investigation of the relay channel with feedback. In this scenario, node-1 is designated as the source node, node-3 the destination, and node-2 the relay. Since there is only one source in this case, one can easily see that maximizing the achievable rate R is equivalent to minimizing the bandwidth expansion factor. Contrary to previous works on the relay channel, we allow the destination to transmit over the noisy wireless channel and investigate the achievable rates in this context.

Before proceeding to our scenario of interest, we review briefly the available results on the AWGN relay channel. In a recent work [8], Kramer *et al.* present a comprehensive overview of existing cooperation strategies, and the corresponding achievable rates, for full- and half-duplex relay channels. In our work, we focus on two classes of cooperation strategies, namely 1) the DF and 2) the CF strategies.

In DF cooperation, the relay node first decodes the source message and then starts aiding the destination node in decoding. More specifically, the transmission cycle is divided into two stages. In the first stage, which occupies a fraction t of the total time, the source node sends a common messages to both the relay and the destination node. Typically, more information is sent in this stage than can be decoded by the destination node. Having successfully decoded the source message in this stage, the relay node uses the second stage to help the destination resolve its uncertainty about the transmitted codeword. During the second stage, a new message is also sent to the destination node from the source node, along with the information from the relay. When the source–relay link is very noisy, one can argue that requiring the relay node to decode the message before starting to help the destination may, in fact, adversely affect performance. The CF strategy avoids this drawback by asking the relay to

	m_1	m_3	m_2
	at	(1-a)t	1-t
source	transmit	transmit	transmit
relay	receive	receive	transmit
destination	receive	transmit	receive
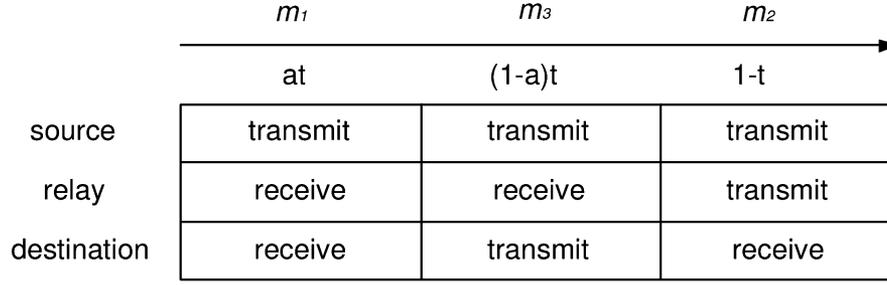
Fig. 2.   The operation sequence of the half-duplex relay channel with noisy feedback.

"compress" its observations and send it to the destination. In this approach, Wyner–Ziv source compression is employed by the relay to allow the destination node to obtain a (noisy) copy of the relay observations. Similar to the DF strategy, the transmission cycle is divided into two stages. During the first stage, both the relay and the destination listen to the source node. The relay then quantizes its observations and sends the quantized data to the destination node during the second stage. In general, the correlation between the relay and destination observations can be exploited by the Wyner–Ziv coding to reduce the data rate at the relay. During the second stage, new information is also sent by the source to further boost the total throughput.

Lemma 1: The achievable rate of the DF and CF strategies are given by

$$
\begin{aligned}
R_{\mathrm{DF}} = \max_{t,r_{12},P_i^{(j)}} \min \Big\{ & tC\left(h_{12}^2 P\right) \\
& + (1-t)C\left(\left(1-r_{12}^2\right)h_{13}^2 P_1^{(2)}\right); \\
& tC\left(h_{13}^2 P\right) + (1-t)C \\
& \left(h_{13}^2 P_1^{(2)} + 2r_{12}h_{13}h_{23}\sqrt{P_1^{(2)}P_2^{(2)}} + h_{23}^2 P_2^{(2)}\right) \Big\}
\end{aligned}
\tag{8}
$$

$$
R_{\mathrm{CF}} = \max_{t,P_i^{(j)}} tC\left(\left(h_{13}^2 + \frac{h_{12}^2}{1+\sigma_2^2}\right)P\right) \\
+ (1-t)C\left(h_{13}^2 P_1^{(2)}\right).
\tag{9}
$$

where

$$
\sigma_2^2 = \frac{\left(h_{12}^2 + h_{13}^2\right)P + 1}{\left(h_{13}^2 P + 1\right)\left(\left(1 + \frac{h_{23}^2 P_2^{(2)}}{h_{13}^2 P_1^{(2)}+1}\right)^{\frac{1-t}{t}} - 1\right)}
\tag{10}
$$

and

$$
P_1^{(2)} + P_2^{(2)} = P.
\tag{11}
$$

Here we omit the detailed proofs and refer the interested readers to the relevant works ([8], [10], [16]–[22]). We note, however, that the statement of the results allows for employing optimal power allocation policies to maximize the throughput.

One can leverage the feedback, from the destination to the relay, to further increase the achievable rate [10]. The capacity of the full-duplex relay channel with *noiseless* feedback, where the exact received signal at the destination is available to the relay, is known [10]. Applying the coding scheme of [10] one can get the following expression for capacity of the half-duplex

relay channel with noiseless feedback [11], which also serves as the upper bound of the achievable rate:

$$
\begin{aligned}
R_{\mathrm{UB}} = \max_{t,r_{12},P_i^{(j)}} \min \Big\{ & tC\left((h_{12}^2 + h_{13}^2)P_1^{(1)}\right) \\
& + (1-t)C\left(\left(1-r_{12}^2\right)h_{13}^2 P_1^{(2)}\right); \\
& tC\left(h_{13}^2 P_1^{(1)}\right) + (1-t)C \\
& \left(h_{13}^2 P_1^{(2)} + 2r_{12}h_{13}h_{23}\sqrt{P_1^{(2)}P_2^{(2)}} + h_{23}^2 P_2^{(2)}\right) \Big\}.
\end{aligned}
\tag{12}
$$

In the following, we present a cooperation strategy for the relay channel with *noisy* feedback. Our model for the noisy feedback represents a more faithful model for the wireless environments. In a nutshell, the proposed strategy combines the DF and CF strategies to overcome the bottleneck of a noisy source–relay channel. In this FeedBack (FB) approach, the destination first assists the relay in decoding via CF cooperation. After decoding, the relay starts helping the destination via a DF configuration. Due to the half-duplex constraint, every cycle of transmission is divided into the following three stages (as shown in Fig. 2).

- The first state lasts for a fraction αt of the cycle ($0 \le t, \alpha \le 1$). In this stage, both the relay and the destination listen to the source, thus we set $X_2 = X_3 = 0$. We refer to the network state in this stage as m_1.
- The feedback stage lasts for a fraction $(1-\alpha)t$ of the cycle. In this stage, the relay listens to both the destination and the source, so $X_2 = 0$. Since the destination is not yet able to completely decode the source message, it sends to the relay node a Wyner–Ziv compressed version of its observations. We refer to the network state in this stage as m_3.
- The final stage lasts for a fraction $(1-t)$ of the cycle. Having obtained the source information, the relay is now able to help the destination node in decoding the source message. In this stage $X_3 = 0$. We refer to the network state in this stage as m_2.

The time-division parameters t and α control the relative duration of each network state. In particular, t represents the total time when the relay node is in the receive mode. The feedback parameter α controls the amount of feedback, i.e., a $(1-\alpha)$ fraction of the total relay listening time is dedicated to feedback. Here, we stress that this formulation for a relay channel with feedback represents a "realistic" view that attempts to capture

the constraints imposed by the wireless scenario (as opposed to the *noiseless* feedback mentioned above). The feedback considered here simply refers to transmission from the destination to the relay over the same (noisy) wireless channel. Using random coding arguments we obtain the following achievable rate for the proposed feedback scheme.

Theorem 1: The achievable rate of the noisy feedback scheme for discrete memoryless channel (DMC) is

$$
R_{\text{FB}} = \sup_{\alpha, t, p(x_1 \mid m_1), p(x_1, x_3 \mid m_3), p(x_1, x_2 \mid m_2), p(\hat{y}_3)}
$$
$$
\min\{\alpha t I(X_1; Y_2, \hat{Y}_3 \mid m_1)
$$
$$
+ (1-\alpha) t I(X_1; Y_2 \mid X_3, m_3)
$$
$$
+ (1-t) I(X_1; Y_3 \mid X_2, m_2);
$$
$$
\alpha t I(X_1; Y_3 \mid m_1) + (1-t) I(X_1, X_2; Y_3 \mid m_2)\}
$$

(13)

subject to

$$
(1-\alpha) I(X_3; Y_2 \mid m_3) \geq \alpha I(Y_3; \hat{Y}_3 \mid Y_2, m_1). \tag{14}
$$

Proof: Please refer to Appendix I. □

The gain leveraged from the noisy feedback can be seen in the increased rate $\alpha t I(X_1; Y_2, \hat{Y}_3 \mid m_1)$ that the relay can decode after combining the signal from the source and the feedback signal from the destination, where (14) captures the constraint on the amount of information the destination can send to the relay. The result for the Gaussian channel now follows.

Lemma 2: The achievable rate of the noisy feedback scheme in the Gaussian channel is given by

$$
R_{\text{FB}} = \max_{\alpha, t, r_{12}, P_i^{(j)}} \min \left\{ \alpha t C \left(\left(\frac{h_{13}^2}{1+\sigma_3^2} + h_{12}^2 \right) P_1^{(1)} \right) \right.
$$
$$
+ (1-\alpha) t C \left(h_{12}^2 P_1^{(3)} \right)
$$
$$
+ (1-t) C \left(\left(1 - r_{12}^2\right) h_{13}^2 P_1^{(2)} \right);
$$
$$
\alpha t C \left(h_{13}^2 P_1^{(1)} \right) + (1-t) C
$$
$$
\left. \left(h_{13}^2 P_1^{(2)} + 2 r_{12} h_{13} h_{23} \sqrt{P_1^{(2)} P_2^{(2)}} + h_{23}^2 P_2^{(2)} \right) \right\}
$$

(15)

where

$$
\sigma_3^2 = \frac{\left(h_{12}^2 + h_{13}^2\right) P_1^{(1)} + 1}{\left(h_{12}^2 P_1^{(1)} + 1\right) \left(\left(1 + \frac{h_{23}^2 P_3^{(3)}}{h_{12}^2 P_1^{(3)} + 1} \right)^{\frac{1-\alpha}{\alpha}} - 1 \right)} \tag{16}
$$

and r_{12} is the correlation between X_1, X_2 during state m_2. In the proposed strategy, the total power constraint specializes to

$$
P_1^{(1)} = P, \quad P_1^{(2)} + P_2^{(2)} = P, \quad P_1^{(3)} + P_3^{(3)} = P. \tag{17}
$$

Proof: Please refer to Appendix II. □

Armed with Lemmas 1 and 2, we can now contrast the performance of the DF, CF, and FB strategies. Our emphasis is to characterize the fundamental properties of the feedback scheme and quantify the gain offered by it under different assumptions on the channel gains and total power. The *relay-off* performance,
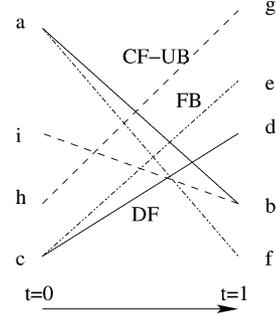
Fig. 3. A geometric representation of FB-, DF- and CF-relay schemes. The solid lines are for R_{DF} in (8), the dash–dotted for R_{FB} in (15), and the dashed for the upper bound of R_{CF} in (19). The various endpoints in the figure are (a) $C(h_{13}^2 P_1^{(2)} + 2 r_{12} h_{13} h_{23} \sqrt{P_1^{(2)} P_2^{(2)}} + h_{23}^2 P_2^{(2)})$, (b) $C(h_{13}^2 P_1^{(1)})$, (c) $C((1 - r_{12}^2) h_{13}^2 P_1^{(2)})$, (d) $C(h_{12}^2 P_1^{(1)})$, (e) $\alpha C((\frac{h_{13}^2}{1+\sigma_3^2} + h_{12}^2) P_1^{(1)}) + (1 - \alpha) C(h_{12}^2 P_1^{(3)})$, (f) $\alpha C(h_{13}^2 P_1^{(1)})$, (g) $C((h_{13}^2 + h_{12}^2) P_1^{(1)})$, (h) $C(h_{13}^2 P_1^{(2)})$, and (i) $C(h_{13}^2 P_1^{(2)} + h_{23}^2 P_2^{(2)})$.

i.e., $R_{\text{ro}} = C(h_{13}^2 P)$, serves as a lower bound on the achievable rate. In fact, the relay-off benchmark can be viewed as a special case of the three cooperative schemes. For example, setting $P_2^{(2)} = 0$ and $t = 0$ effectively reduces both DF and CF strategies to the relay-off case. Therefore, one can conceptually describe the order of containment of various schemes as "relay-off $\subset$ DF $\subset$ FB" and "relay-off $\subset$ CF." As for the performance upper bounds, the cut-set bounds [10] give rise to 1) a multi-transmitter rate $R_{(1,2)-3} = C((h_{13}^2 + h_{23}^2)P)$ corresponding to perfect cooperation between the source and relay nodes (notice that with the total power constraint $P_1 + P_2 = P$, the optimal power for node-1 is $P_1 = \frac{h_{13}^2 P}{h_{13}^2 + h_{23}^2}$); and 2) a multireceiver rate $R_{1-(2,3)} = C((h_{13}^2 + h_{12}^2)P)$ corresponding to perfect cooperation between the relay and destination nodes.

The achievable rate of the DF strategy, i.e., R_{DF}, enjoys an intuitive geometric interpretation: each expression within the min operator is a linear segment in the parameter $t \in [0, 1]$ (see (8)). Hence, the optimal time t, assuming the other variables remain fixed, can be simply determined by the intersection point of the two associated line segments, as illustrated in Fig. 3. On the other hand, R_{FB} and R_{CF} are characterized by more complicated expressions due to the dependency of σ_3^2 and σ_2^2 upon the time-division parameters. Our next result finds upper bounds on R_{FB} and R_{CF} which allow for the same simple line-crossing interpretation as R_{DF}.

Lemma 3: The achievable rate of the feedback scheme is upper bounded by

$$
R_{\text{FB}} \leq \max_{\alpha, t, r_{12}, P_i^{(j)}} \min \left\{ \alpha t C \left(h_{12}^2 P_1^{(1)} \right) \right.
$$
$$
+ (1-\alpha) t C \left(h_{12}^2 P_1^{(3)} + h_{23}^2 P_3^{(3)} \right)
$$
$$
+ (1-t) C \left(\left(1 - r_{12}^2\right) h_{13}^2 P_1^{(2)} \right);
$$
$$
\alpha t C \left(h_{13}^2 P_1^{(1)} \right) + (1-t) C
$$
$$
\left. \left(h_{13}^2 P_1^{(2)} + 2 r_{12} h_{13} h_{23} \sqrt{P_1^{(2)} P_2^{(2)}} + h_{23}^2 P_2^{(2)} \right) \right\}.
$$

(18)

The achievable rate of CF is bounded by

$$R_{\mathrm{CF}} \leq \max_{t, P_i^{(j)}} \min \left\{ tC\left(h_{13}^2 P_1^{(1)}\right) \right.$$
$$+ (1-t)C\left(h_{13}^2 P_1^{(2)} + h_{23}^2 P_2^{(2)}\right);$$
$$\left. tC\left(\left(h_{13}^2 + h_{12}^2\right)P_1^{(1)}\right) + (1-t)C\left(h_{13}^2 P_1^{(2)}\right) \right\}. \tag{19}$$

Proof: Please refer to Appendix III. □

Comparing (15) and (18) sheds more light on the intuition behind the upper bound on R_{FB}. The upper bound replaces the actual rate received at the relay node with the maximum rate corresponding to the following two stages. In the first stage, only the source is transmitting and the second stage is a multiple-access channel where the source and the destination play the roles of transmitting users. Fig. 3 compares the performance of the three schemes. For example, when $h_{12}^2 \leq h_{13}^2$, the intersection point corresponding to decode–forward would fall below the flat line $C(h_{13}^2 P)$ associated with the relay-off rate. More rigorously, we have the following statement.

Theorem 2:
1) If $h_{12}^2 \leq h_{13}^2$ then $R_{\mathrm{DF}} \leq R_{\mathrm{ro}}$.
2) If $h_{23}^2 \leq h_{13}^2$ then $R_{\mathrm{CF}} \leq R_{\mathrm{ro}}$.
3) If $h_{23}^2 \leq h_{12}^2$ then $R_{\mathrm{FB}} \leq R_{\mathrm{DF}}$.
   Proof: Please refer to Appendix IV. □

Theorem 2 reveals the fundamental impact of channel coefficients on the performance of the different cooperation strategies. In particular, the DF strategy is seen to work well with a "strong" source–relay link. If, at the same time, the relay–destination link is stronger, then one may exploit feedback, i.e., $\alpha \neq 1$, to improve performance. The next result shows that the feedback gain is diminishing in certain asymptotic scenarios.

Theorem 3:
1) As h_{12} increases, both DF and FB schemes approach the optimal beam-forming benchmark, while the CF scheme is limited by a suboptimal rate $C(\max\{h_{13}^2, h_{23}^2\}P)$.
2) As h_{23} increases, both CF and FB schemes approach the optimal multireceiver benchmark, while the DF scheme only approaches a suboptimal rate $C(\max\{h_{12}^2, h_{13}^2\}P)$.

The proof of Theorem 3 is a straightforward limit computation, and hence, is omitted for brevity. So far, we have kept the total power P constant. But in fact, the achievable rate as a function of P offers another important dimension to the problem. First, we investigate the low-power regime. In this case, we study the slope S of the achievable rate with respect to P (i.e., $R \sim \frac{1}{2}(\log e)SP$). This slope determines the minimum energy per bit [23] according to the relationship

$$\left(\frac{E_b}{N_0}\right)_{\min} = 1.44 - 10\log_{10}(S). \tag{20}$$

We observe that (20) suffers from a 3-dB loss due to our *real* channel model (as opposed to the complex channel model [23]).

Theorem 4: Let

$$f_1(\theta, r_{12}, h_{13}, h_{23}) = h_{13}^2 \cos^2\theta + 2r_{12}h_{13}h_{23}\cos\theta\sin\theta + h_{23}^2\sin^2\theta$$

and

$$f_2(\theta, r_{12}, h_{13}) = (1 - r_{12}^2)h_{13}^2\cos^2\theta$$

be a shorthand notation, then we have the following.
1) When $h_{12}^2 \geq h_{13}^2$

$$S_{\mathrm{DF}}$$
$$= \max_{\theta, r_{12}} \frac{f_1(\theta, r_{12}, h_{13}, h_{23})h_{12}^2 - f_2(\theta, r_{12}, h_{13})h_{13}^2}{f_1(\theta, r_{12}, h_{13}, h_{23}) + h_{12}^2 - f_2(\theta, r_{12}, h_{13}) - h_{13}^2} \tag{21}$$

and

$$\frac{\left(h_{13}^2 + h_{23}^2\right)h_{12}^2}{h_{23}^2 + h_{12}^2} \leq S_{\mathrm{DF}} \leq \frac{\left(h_{13}^2 + h_{23}^2\right)h_{12}^2 - h_{13}^4}{h_{23}^2 + h_{12}^2 - h_{13}^2}. \tag{22}$$

2) $S_{\mathrm{CF}} = h_{13}^2$ with $t_{\mathrm{opt}} \to 1$.
3) $S_{\mathrm{FB}} = S_{\mathrm{DF}}$ with $\alpha_{\mathrm{opt}} \to 1$.
   Proof: Please refer to Appendix V. □

It follows from Theorem 4 that given $h_{12}^2 > h_{13}^2$, DF cooperation delivers a larger slope than the relay-off, i.e.,

$$S_{\mathrm{DF}} - h_{13}^2 \geq \frac{h_{23}^2\left(h_{12}^2 - h_{13}^2\right)}{h_{23}^2 + h_{12}^2} > 0. \tag{23}$$

In this case, the signal the relay receives is better than the signal the destination receives. If the relay is off, the transmission rate is dictated by the smaller channel gain h_{13}^2. In DF cooperation, the source can send information at the rate that only the relay can decode $(h_{12}^2 > h_{13}^2)$. After decoding, the relay can help the source to send information to the destination. However, CF cooperation does not yield any gain in the low power regime. Similarly, we see that the CF stage of the proposed FB becomes useless, and hence, the scheme reduces to the DF approach in the low-power regime. The reason lies in the fact that for small P, the channel output is dominated by the noise, and hence, the compression algorithm inevitably operates on the noise, resulting in diminishing gains.

We next quantify the power offset of the three schemes in the high-power regime, that is, to characterize $R \sim \frac{1}{2}\log P + \frac{1}{2}G$ as $P \to \infty$ [24].

Theorem 5: Following the same shorthand notation as in Theorem 4, we obtain
1) Given $h_{12}^2 \geq h_{13}^2$ we get (24) at the bottom of the page.
2)

$$G_{\mathrm{CF}} = \max_{t, \theta} t\log\left(h_{13}^2 + \frac{h_{12}^2}{1 + \sigma_2^2(\infty)}\right) + (1-t)\log\left(h_{13}^2\cos^2\theta\right) \tag{25}$$

$$G_{\mathrm{DF}} = \max_{\theta, r_{12}} \frac{\log f_1(\theta, r_{12}, h_{13}, h_{23}) \cdot \log h_{12}^2 - \log f_2(\theta, r_{12}, h_{13}) \cdot \log h_{13}^2}{\log\left[f_1(\theta, r_{12}, h_{13}, h_{23})h_{12}^2\right] - \log\left[f_2(\theta, r_{12}, h_{13})h_{13}^2\right]}. \tag{24}$$
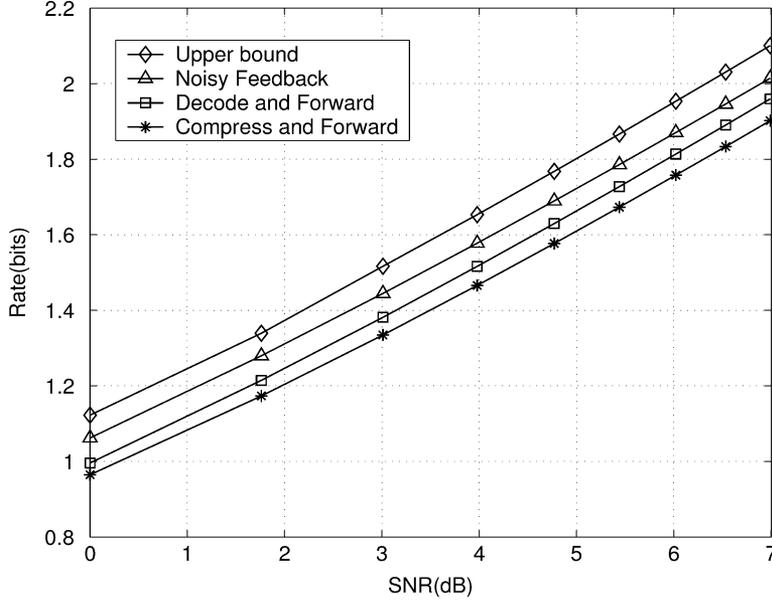
Fig. 4. The achievable rate of various schemes in the half-duplex relay channel, $h_{12} = 2.55$dB, $h_{13} = 0$ dB, $h_{23} = 23$ dB.

where

$$\sigma_2^2(\infty) = \frac{h_{12}^2 + h_{13}^2}{h_{13}^2 \left(\left(1 + \frac{h_{23}^2}{h_{13}^2} \tan^2 \theta \right)^{\frac{1-t}{t}} - 1 \right)}. \qquad (26)$$

3) $G_{\mathrm{FB}} = G_{\mathrm{DF}}$ with $\alpha_{\mathrm{opt}} \to 1$.

Proof: Please refer to Appendix VI. $\qquad \square$

Theorem 5 reveals the fact that strict feedback ($\alpha \neq 1$) does not yield a gain in high-power regime. The reason for this behavior can be traced back to the half-duplex constraint. When $\alpha \neq 1$, the destination spends a fraction $(1 - \alpha)t$ of time transmitting to the relay, which cuts off the time in which it would have been listening to the source in nonfeedback schemes. Such a time loss reduces the pre-log constant, which cannot be compensated by the cooperative gain when P becomes large. The quantity $G_{\mathrm{DF}} - G_{\mathrm{CF}}$ determines the relative order of the DF scheme and CF scheme in the high-SNR region. In general, this quantity depends on the channel gains and can be computed using numerical methods. For example, when $h_{12} = 3$ dB, $h_{13} = 0$ dB, $h_{23} = 10$ dB, we find that $G_{\mathrm{DF}} - G_{\mathrm{CF}} = 1.24$ dB.

We conclude this section with simulation results that validate our theoretical analysis. Fig. 4 reports the achievable rate of various schemes, when $h_{12} = 2.55$ dB, $h_{13} = 0$ dB, and $h_{23} = 23$ dB. This corresponds to the case when the source–relay channel is a little better than the source–destination channel, and the relay–destination channel is quite good. This is the typical scenario when feedback results in a significant gain, as demonstrated in the figure. Fig. 5 reports the ratios $r_{\mathrm{FB}} = R_{\mathrm{FB}}/R_{\mathrm{UB}}$ and $r_{\mathrm{DF}} = R_{\mathrm{DF}}/R_{\mathrm{UB}}$. We can see that, as SNR increases, the achievable rate of the FB scheme converges to that of the DF scheme, which confirms our claim in Theorem 5. Fig. 6 reports the achievable rates of various schemes, as we vary the relay–destination channel gain h_{23}. It is clear that as the relay–destination channel becomes better, the advantage of feedback increases. Fig. 7 shows the position of the relay where

each relay scheme achieves the highest achievable rate. Here we put a source node at $(0, 0)$ and a destination node at $(1, 0)$ and change the position of the relay node. The channel condition between nodes is given by $h_{ij} = \frac{1}{d_{ij}^\beta}$, where d_{ij} is the distance between node i and j. Overall, we can see that the proposed FB cooperation scheme combines the benefits of both the DF and CF cooperation strategies, and hence, attains the union of the asymptotic optimality properties of the two strategies. On the other hand, the gain offered by feedback seems to be limited to certain operating regions, as defined by the channel gains, and diminishes in either the low- or high-power regime.

III. THE MULTICAST CHANNEL

The relay channel, considered in the previous section, represents the simplest example of a three-node wireless network. Another example can be obtained by requiring node-2 to decode the message generated at node-1. This corresponds to the multicast scenario. Similar to the relay scenario, we focus on maximizing the achievable rate from node-1 to both nodes 2 and 3, without any loss of generality. The half-duplex and total power constraints, adopted here, introduce an interesting design challenge. To illustrate the idea, suppose that node-2 decides to help node-3 in decoding. In this case, not only does node-2 compete with the source node for transmit power, but it also sacrifices its listening time for the sake of helping node-3. It is, therefore, not clear *a priori* if the network would benefit from this cooperation.

In a recent work [25], the authors considered another variant of the multicast channel and established the benefits of receiver cooperation in this setup. The fundamental difference between the two scenarios is that, in [25], the authors assumed the existence of a dedicated link between the two receivers. This dedicated link was used by the *strong* receiver to help the *weak* receiver in decoding through a DF strategy. As expected, such a cooperation strategy was shown to strictly enlarge the achievable rate region
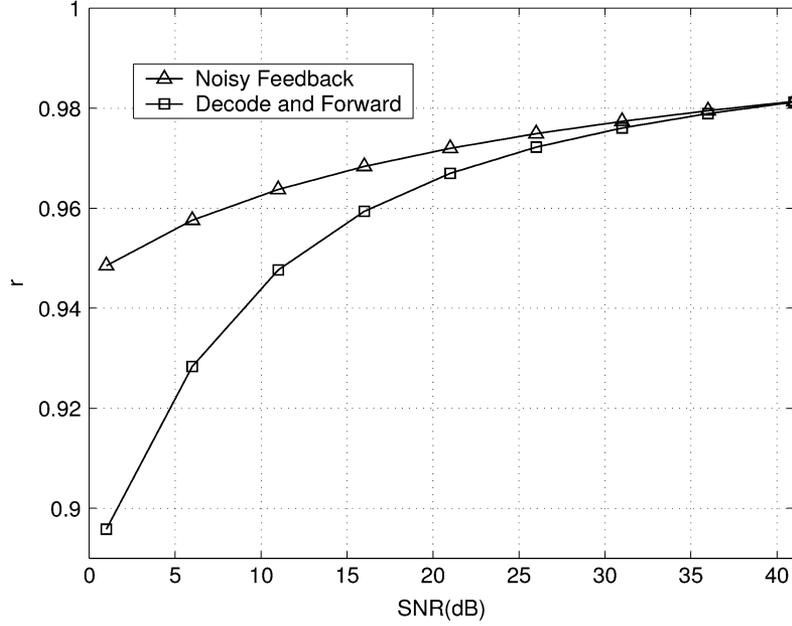
Fig. 5.   The ratio of achievable rate of various schemes to the cut-set upper bound, $h_{12} = 2.55$ dB, $h_{13} = 0$ dB, $h_{23} = 23$ dB.

[25]. In our work, we consider a more representative model of the wireless network in which all communications take place over the same channel, subject to the half-duplex and total power constraints. Despite these constraining assumptions, we still demonstrate the significant gains offered by receiver cooperation. Inspired by the feedback–relay channel, we further construct a greedy cooperation strategy that significantly outperforms the DF scheme [25] in many relevant scenarios.

In the noncooperative scenario, both node-2 and node-3 will listen all the time, and hence, the achievable rate is given by

$$C_{\text{noncoop}} = C(\min\{h_{12}^2, h_{13}^2\}P). \qquad (27)$$

Due to the half-duplex constraint, time is valuable to both nodes, which makes them selfish and unwilling to help each other at first thought. A more careful consideration, however, reveals that such a *greediness* will lead the nodes to cooperate. The enabling observation stems from the feedback strategy proposed for the relay channel in which the destination was found to get a higher achievable rate if it sacrifices some of its receiving time to help the relay. Motivated by this observation, our strategy decomposes into three stages, without loss of generality we assume $h_{12}^2 > h_{23}^2$, 1) m_1 lasting for a fraction αt of the frame during which both receivers listen to node-1; 2) m_3 occupying $(1 - \alpha)t$ fraction of the frame during which node-3 sends its compressed signal to node-2; and 3) m_2 (the rest $1 - t$ fraction) during which node-2 helps node-3 finish decoding. One major difference between the multicast and relay scenarios is that in the m_2 stage, the source cannot send additional (new) information to node-3, for it would not be decoded by node-2, thus violating the multicast requirement that both receivers obtain the same source information. Here, we observe that the last stage of cooperation, in which node-2 is helping node-3, is still motivated by the greedy approach. The idea is that node-1 will continue transmitting the same codeword until both receivers can successfully decode. It is, therefore, beneficial for node-2

to help node-3 in decoding faster to allow the source to move on to the next packet in the queue.

Lemma 4: The achievable rate of the greedy multicast strategy strategy is given by

$$R_g = \max_{\alpha,t,P_i^{(j)}} \min \left\{ \alpha t C \left(\left(\frac{h_{13}^2}{1 + \sigma_4^2} + h_{12}^2 \right) P \right) \right. $$
$$+ (1 - \alpha) t C \left(h_{12}^2 P_1^{(3)} \right); $$
$$\left. \alpha t C \left(h_{13}^2 P \right) + (1 - t) C \left((h_{13}^2 + h_{23}^2) P \right) \right\} \qquad (28)$$

where

$$\sigma_4^2 = \frac{(h_{12}^2 + h_{13}^2) P + 1}{(h_{12}^2 P + 1) \left(\left(1 + \frac{h_{23}^2 P_3^{(3)}}{h_{12}^2 P_1^{(3)} + 1} \right)^{\frac{1-\alpha}{\alpha}} - 1 \right)}. \qquad (29)$$

The proof follows in the footsteps of the proof of Lemma 2, the only difference is that now the source cannot send new information to the relay in the m_2 state.

We observe that the DF multicast scheme corresponds to the special case of $\alpha = 1$, which has a rate

$$R_{\text{DF}} = \max_t \min \left\{ t C \left(h_{12}^2 P \right); t C \left(h_{13}^2 P \right) \right. $$
$$\left. + (1 - t) C \left((h_{13}^2 + h_{23}^2) P \right) \right\}. \qquad (30)$$

The cut-set upper bounds give rise to the two following benchmarks: beam-forming $R_{(1,2)-3} = C((h_{13}^2 + h_{23}^2)P)$ and multireceiver $R_{1-(2,3)} = C((h_{13}^2 + h_{12}^2)P)$. Similar to the relay channel scenario, we examine in the following the asymptotic behavior of the greedy strategy as a function of the channel coefficients and available power.

Theorem 6:

1) The greedy cooperative multicast scheme strictly increases the achievable rate (as compared to the noncooperative scenario).
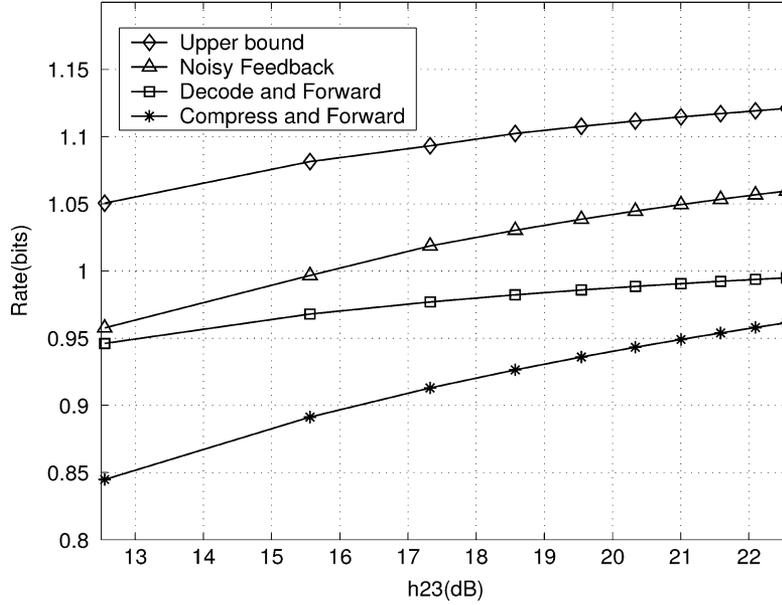
Fig. 6.   The achievable rate of various schemes in the half-duplex relay channel, $h_{12} = 2.55$ dB, $h_{13} = 0$ dB, SNR $= 0$ dB.

2) The greedy strategy approaches the beam-forming benchmark as h_{12} increases, i.e.,

$$\lim_{h_{12} \to \infty} R_g = C\left(\left(h_{13}^2 + h_{23}^2\right) P\right). \tag{31}$$

3) The greedy strategy approaches the multireceiver benchmark as h_{23} increases, i.e.,

$$\lim_{h_{23} \to \infty} R_g = C\left(\left(h_{12}^2 + h_{13}^2\right) P\right). \tag{32}$$

4) As $P \to 0$, the slope of the greedy strategy achievable rate is given by

$$S_g = \frac{h_{12}^2 \left(h_{23}^2 + h_{13}^2\right)}{h_{12}^2 + h_{23}^2}. \tag{33}$$

5) As $P \to \infty$, the SNR gain $G_g = G_{\text{non-coop}} = \log h_{13}^2$ with $t_{\text{opt}} \to 1$.

Proof: Please refer to Appendix VII.   □

Parts 2) and 3) demonstrate the asymptotic optimality of the greedy multicast as the channel gains increase (the proof follows the same line as that of Theorem 3). On the other hand, we see that the large-power asymptotic of the multicast channel differs significantly from that of the relay channel. In the relay case (Theorem 5), the contribution of feedback diminishes $(G_{\text{FB}} = G_{\text{DF}})$ in this asymptotic scenario, but cooperation was found to be still beneficial, that is, $G_{\text{DF}} > \log h_{13}^2$. On the contrary, the gain of receiver cooperation in the multicast channel disappears as P increases. This is because, unlike the relay scenario, at least one receiver must cut its listening time in any cooperative multicast scheme due to the half-duplex constraint. Such a reduction induces a pre-log penalty in the rate, which results in substantial loss that cannot be compensated by cooperation as $P \to \infty$, and hence, the greedy strategy reduces to the noncooperative mode automatically.

Fig. 8 compares the achievable rate of the various multicast schemes where the DF cooperation strategy is shown

to outperform the noncooperation scheme. It is also shown that optimizing the parameter α provides an additional gain. (Note that R_{DF} in the figure corresponds to $\alpha = 1$.) Fig. 9 reports the ratio $r = R_i/C_{\text{noncoop}}$ for the greedy and DF cooperation scheme, where we observe that the gain offered by cooperation decreases as the SNR increases. Fig. 10 reports the achievable rate of the three schemes when $h_{12} = h_{13}$. In this case, it is easy to see that DF strategy yields **exactly** the same performance as the noncooperative strategy. On the other hand, as illustrated in the figure, the proposed greedy strategy is still able to offer a sizable gain. Fig. 11 illustrates the fact that the gain of greedy strategy increases as h_{23} increases. The noncooperation scheme is not able to exploit the inter-receiver channel, and hence, its achievable rate corresponds to a flat line. The DF scheme can benefit from the inter-receiver channel, but its maximum rate is limited by $C(h_{12}^2 P)$, whereas the greedy strategy approaches a rate $R_g = C((h_{12}^2 + h_{13}^2)P)$ as $h_{23} \to \infty$.

IV. THE THREE-WAY CHANNEL

Arguably, the most demanding instantiation of the three-node network is the three-way channel [12], [26], [27]. To satisfy the three-way channel requirements, every node needs to transmit its message to the other two nodes and receive their messages from them. Due to the half-duplex constraint, these two tasks cannot be completed simultaneously. Take node-1 as an example (see Fig. 12) and consider the transmission of a block of observations S_1^K to the other two nodes using N_t channel uses. To obtain a lower bound on the bandwidth expansion factor, we assume that node-2 and node-3 can fully cooperate, from a joint source–channel coding perspective, which converts the problem into a point-to-point situation. Then node-1 only needs to randomly divide its source sequences into $2^{KH(S_1 \mid S_2, S_3)}$ bins and transmit the corresponding bin index [14], [28], [29]. With N_t channel uses, the information rate is $\frac{KH(S_1 \mid S_2, S_3)}{N_t}$. The channel capacity between node-1 and the multiple-antenna node-2, 3 is
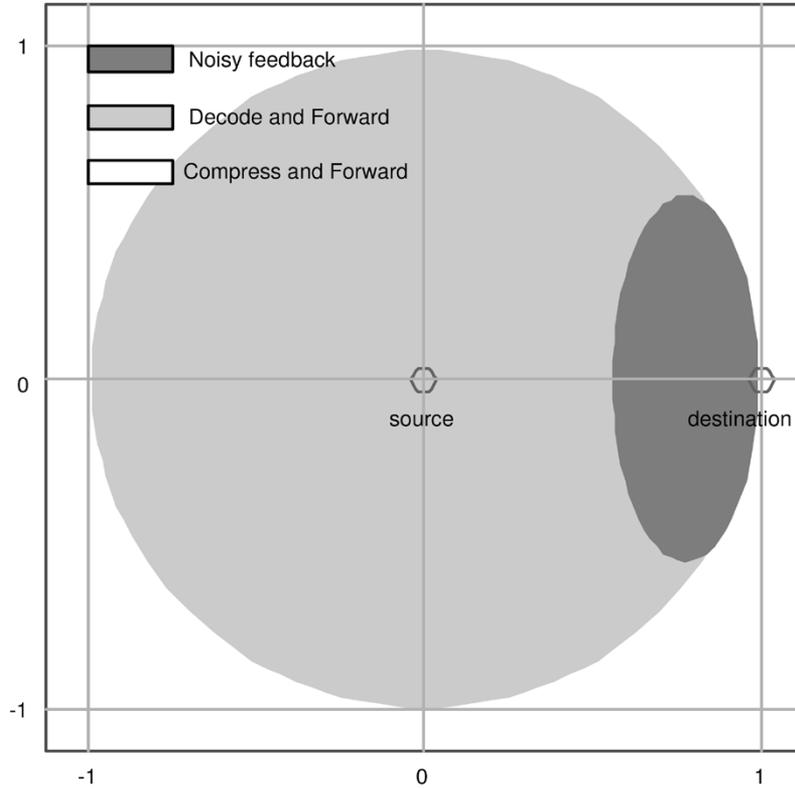
Fig. 7.   The position of the relay where different scheme achieves the largest achievable rate, $P = 3$ dB, $\beta = 4$.
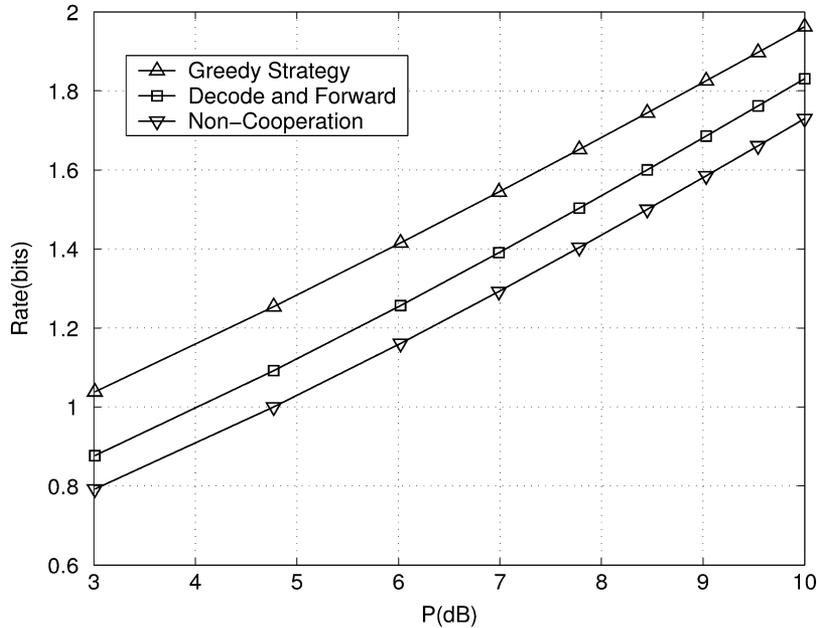
Fig. 8.   The achievable rate of various schemes in the multicast channel, $h_{12} = 0.4$ dB, $h_{13} = 0$ dB, and $h_{23} = 23$ dB.

$C((h_{12}^2 + h_{13}^2)P)$. In order to decode S_1^K at node-2, 3 with a vanishingly small error probability, the following condition must be satisfied [30], [31]:

$$\frac{KH(S_1 \mid S_2, S_3)}{N_t} \leq C\left(\left(h_{12}^2 + h_{13}^2\right) P\right).$$

Similarly, with full cooperation between node-2 and node-3, the following condition is needed to ensure the decoding of the se-

quence S_2^K, S_3^K at node-1 with a vanishingly small error probability:

$$\frac{KH(S_2, S_3 \mid S_1)}{N_r} \leq C((h_{12}^2 + h_{13}^2)P).$$

These two genie-aided bounds at node-1 imply that the minimum bandwidth expansion factor required for node-1 is

$$\tau_{1,\text{gen}} = \frac{H(S_1 \mid S_2, S_3) + H(S_2, S_3 \mid S_1)}{C((h_{12}^2 + h_{13}^2)P)}.$$
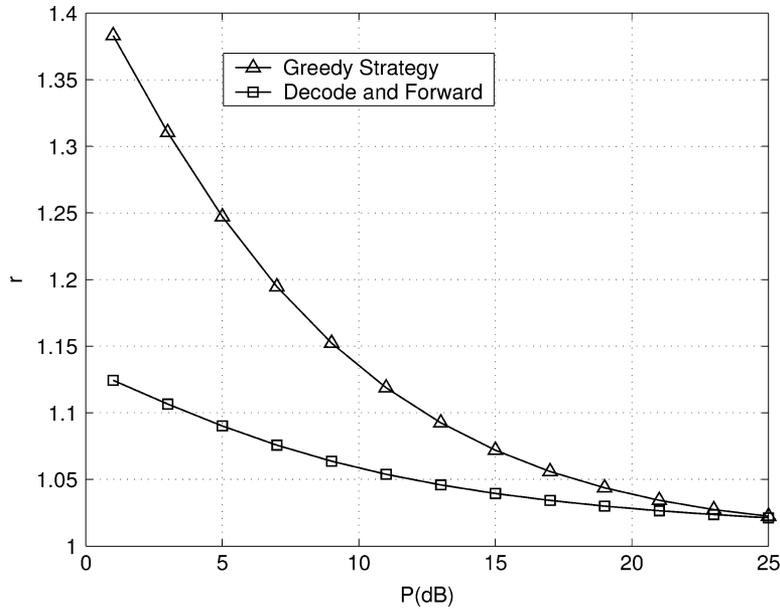
Fig. 9.   The ratio of the achievable rates of various schemes to the noncooperative scheme, $h_{12} = 0.4$ dB, $h_{13} = 0$ dB, and $h_{23} = 23$ dB.
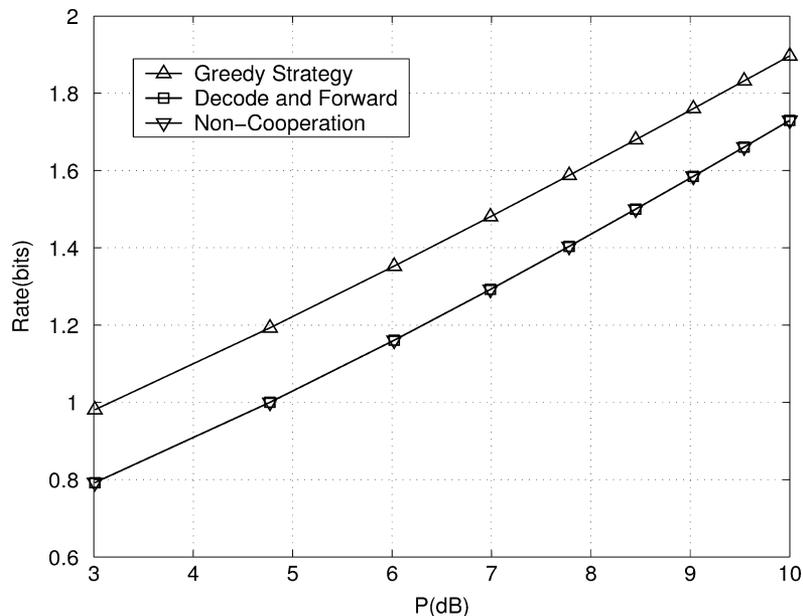
Fig. 10.   The achievable rate of various schemes in the multicast channel, $h_{12} = 0$ dB, $h_{13} = 0$ dB, and $h_{23} = 23$ dB.

Similarly, we can obtain the corresponding genie-aided bounds for node-2 and node-3. To satisfy the requirement for all these three nodes, the minimum bandwidth expansion factor for this half-duplex three-way channel is therefore

$$\tau_{\text{gen}} \geq \max_{i=1,2,3} \tau_{i,\text{gen}}. \tag{34}$$

At this point, it is not clear whether the genie-aided bound in (34) is achievable or not. Moreover, finding the optimal cooperation strategy for the three-way channel remains an elusive task. However, inspired by our greedy multicast strategy, we propose in the following a modular cooperation approach composed of three *cooperative multicast with side information* stages, which achieves the optimal performance in certain asymptotic scenarios.

A. Multicast With Side Information

To simplify the presentation, without sacrificing any generality, we assume that node-1 is the source and nodes 2 and 3 are provided with the side information S_2 and S_3, respectively. A related work appears in the paper [32], where the authors consider the broadcast channel with arbitrarily correlated sources. In that paper, the sender has two correlated messages to send to the two users.

Before presenting our greedy cooperation strategy, we briefly discuss the noncooperative scenario to establish a performance benchmark where the two receive nodes are not allowed to communicate. For the convenience of exposition, we assume that $H(S_1 \mid S_2) > H(S_1 \mid S_3)$. In the nested binning approach of [33], a source sequence s_1^K is randomly assigned to one
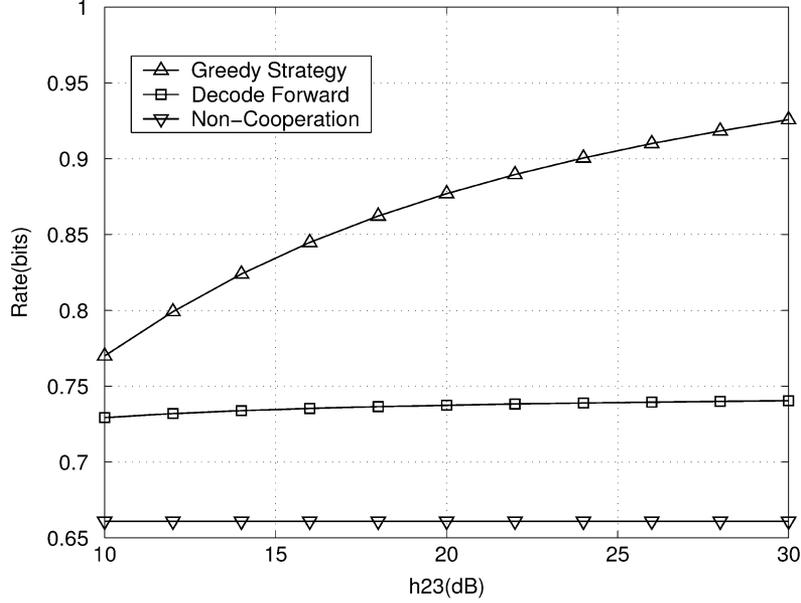
Fig. 11.   The achievable rate of various schemes in the multicast channel, $h_{12} = 0.4$ dB, $h_{13} = 0$ dB, and SNR $= 1.8$ dB.
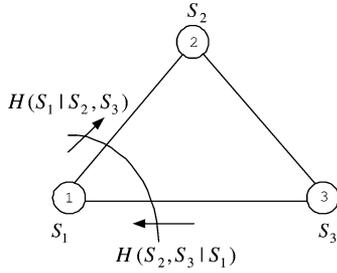
Fig. 12.   The genie-aided bound in node-1, in which node-2 and node-3 can fully cooperate with each other.

of $2^{KH(S_1 \mid S_2)}$ bins. This is the low-level indexing sufficient for node-2 to decode with side information S_2. These indices are then (randomly) divided into $2^{KH(S_1 \mid S_3)}$ equal-sized groups, which corresponds to the random binning approach for node-3. Therefore, a source sequence s_1^K is associated with an index-pair (b, c), where $b \in [1, 2^{KH(S_1 \mid S_3)}]$ is the group index and $c \in [1, 2^{K(H(S_1 \mid S_2) - H(S_1 \mid S_3))}]$ identifies the bin index within a group. Given side information S_3 (more correlated with the source), node-3 needs only the group index b to recover the source sequence. But the low-level bin index is necessary for node-2 to decode. In summary, the above nested binning scheme permits the source node to send (b, c) to node-2 while only b to node-3. Such a structured message is called the *degraded information set* in [34] where b is the "common" information for both receivers and c the "private" information required by only one of the two receivers. Using the capacity region given by [34], we get the following benchmark.

Lemma 5: For noncooperative multicast with side-information, the achievable bandwidth expansion factor $\tau = N/K$ based on nested binning source coding and degraded information set broadcasting is given by

1) if $h_{12}^2 < h_{13}^2$

$$H(S_1 \mid S_2) \le \tau C \left(h_{12}^2 P \right); \tag{35}$$

2) if $h_{12}^2 > h_{13}^2$

$$H(S_1 \mid S_2) - H(S_1 \mid S_3) \le \tau C \left(\gamma h_{12}^2 P \right),$$
$$H(S_1 \mid S_3) \le \tau C \left(\frac{(1 - \gamma) h_{13}^2 P}{1 + \gamma h_{13}^2 P} \right). \tag{36}$$

for some γ.

Now, we are ready to describe our greedy cooperation approach (which is not necessarily optimal). In our scheme, each node calculates the expected bandwidth expansion factor assuming no receiver cooperation $\tau_{\mathrm{ex},i} = H(S_1 \mid S_i)/C_i$, where C_i denotes the link capacity $C(h_{1i}^2 P)$. The receive node with the smaller τ_{ex} is deemed as the *strong* node, and hence, will decode first. Without loss of generality, we assume that node-2 is the strong node. To better describe the proposed approach, we consider first the simple case where node-3 does not help node-2. We randomly bin the sequences S_1^K into $2^{KH(S_1 \mid S_2)}$ bins and denote the bin index by $w \in [1, 2^{KH(S_1 \mid S_2)}]$. We further denote by f_{s1} the mapping function $w = f_{s1}(s_1^K)$. We then independently generate another bin index b for every sequence S_1^K by picking b uniformly from $\{1, 2, \ldots, 2^{KR}\}$, where R is to be determined later. Let $B(b)$ be the set of all sequences S_1^K allocated to bin b. Thus, every source sequence has two bin indices $\{w, b\}$ associated with it. A full cooperation cycle is divided into two stages, where we refer to the network state in these two stages as m_1 and m_2, respectively. In the first stage using for N_1 channel uses, node-1 sends the message w to node-2 using a capacity-achieving code. This stage is assumed to last for N_1 channel uses. At the end of this state, node-2 can get a reliable estimate $\hat{w} = w$ if the condition $KH(S_1 \mid S_2) \le N_1 C(h_{12}^2 P)$ is satisfied. Next, node-2 searches in the bin specified by $\hat{w}$ for the one and only one $\hat{s}_{21}^K$ that is typical with s_2^K. If none exists, decoding error is declared, otherwise, $\hat{s}_{21}^K$ is the decoding sequence. During this stage, node-3 computes a list $\ell(\boldsymbol{y}_{3,m_1})$ such that if $w' \in \ell(\boldsymbol{y}_{3,m_1})$ then $\{\boldsymbol{x}_{1,m_1}(w'), \boldsymbol{y}_{3,m_1}\}$ are jointly typical. A key point of our scheme is that node-3 does not attempt to decode w, but rather proceeds to decoding s_1^K directly.

After node-2 decodes s_1^K correctly, it knows the pair $\{w, b\}$, and hence, in the second stage node-2 and node-1 cooperate to send the message b to node-3. At the end of this stage, if the parameters are appropriately chosen, node-3 can decode b correctly. Node-3 then searches in the bin $B(b)$ for the one and only one $\hat{s}_{31}^K$ that is jointly typical with s_3^K and that $f_{s1}(\hat{s}_{31}^K) \in \ell(\boldsymbol{y}_{3,m_1})$.

Lemma 6: With the proposed scheme, both nodes 2 and 3 can decode S_1^K with a vanishingly small probability of error if $\tau_0 = N_0/K, \tau_1 = N_1/K$ satisfy the following conditions :

$$H(S_1 \mid S_2) \leq \tau_1 C\left(h_{12}^2 P\right)$$

$$H(S_1 \mid S_3) - \frac{\min\left\{C\left(h_{13}^2 P\right), C\left(h_{12}^2 P\right)\right\} H(S_1 \mid S_2)}{C\left(h_{12}^2 P\right)}$$
$$\leq \tau_0 C\left(\left(h_{13}^2 + h_{23}^2\right)P\right). \quad (37)$$

Proof: Please refer to Appendix VIII. □

Next, we allow for the weak node-3 to assist the strong node-2 in decoding. The original N_1 channel uses now split into two parts: 1) state m_1 occupying αN_1 channel uses during which both receiver nodes listen to the source node; and 2) state m_3 for the remaining $(1-\alpha)N_1$ channel uses during which node-3 sends a compressed version of its received signal to node-2. At the end of the N_1 network uses, node-2 decodes the source sequence and then proceeds to facilitate the same list-decoding at the other receiver as described above. The simple case where node-3 does not assist node-2 can be regarded as a special case of the greedy scheme when $\alpha = 1$. Slightly modifying the proof of Lemma 6, we obtain the following.

Lemma 7: If τ_0, τ_1 satisfy the following conditions, both node-2, 3 will decode S_1^K with vanishingly small probability of error (see (38) at the bottom of the page), where $R_{\text{CF2}}(\alpha)$ is the achievable rate of the CF scheme for the following relay channel: node-1 acts as the source, node-3 the relay that spends $1 - \alpha$ part of the time in helping the destination using the CF scheme, and node-2 the destination. The symbol $\hat{Y}_3$ stands for the compressed version of the received signal at node-3 (Y_3).

Unfortunately, the expressions for the achievable bandwidth expansion factors do not seem to allow for further analytical manipulation. In order to shed more light on the relative performance of the different schemes, we introduce the *minimum energy per source observation* metric. Given the total transmission power P, the bandwidth expansion factor τ translates to the energy requirement per source observation as

$$E(P) = \tau(P)P = \frac{N(P)P}{K}. \quad (39)$$

Let $E_1(P)$ denotes the energy per source symbol for the benchmark based on broadcast with a degraded information set and $E_2(P)$ for the proposed cooperative multicast scheme. It is easy to see that both $E_1(P)$ and $E_2(P)$ are nondecreasing function

of P, and hence, approach their minimal values as $P \to 0$, that is,

$$E_{i,m} = \lim_{P \to 0} E_i(P) \quad \text{for } i \in \{1, 2\}. \quad (40)$$

Under the assumption that $\tau_{\text{ex},2} < \tau_{\text{ex},3}$ and using Lemmas 5 and 7, we obtain the following.

Theorem 7:
1) Broadcast with degraded information set:
   When $H(S_1 \mid S_2) > H(S_1 \mid S_3)$,

$$E_{1,m} = \frac{2}{\log e}\left(\frac{H(S_1 \mid S_2)}{h_{12}^2} + \left(\frac{1}{h_{13}^2} - \frac{1}{h_{12}^2}\right)^+ H(S_1 \mid S_3)\right). \quad (41)$$

When $H(S_1 \mid S_2) < H(S_1 \mid S_3)$,

$$E_{1,m} = \frac{2}{\log e}\left(\frac{H(S_1 \mid S_3)}{h_{13}^2} + \left(\frac{1}{h_{12}^2} - \frac{1}{h_{13}^2}\right)^+ H(S_1 \mid S_2)\right). \quad (42)$$

Here $x^+ = \max\{x, 0\}$.
2) Greedy strategy:

$$E_{2,m} = \frac{2}{\log e}\left(\frac{H(S_1 \mid S_2)}{h_{12}^2}\right.$$
$$\left. + \frac{h_{12}^2 H(S_1 \mid S_3) - \min\{h_{13}^2, h_{12}^2\} H(S_1 \mid S_2)}{(h_{13}^2 + h_{23}^2)h_{12}^2}\right). \quad (43)$$

3) $E_{2,m} < E_{1,m}$.

Proof: Please refer to Appendix IX. □

Combined with the results in Section III, this result argues strongly for receiver cooperation in the multicast scenario even under the stringent half-duplex and total power constraints. Finally, Figs. 13 and 14 validate our theoretical claims.

B. Multicast Scheduler

The second step in the proposed solution for the three-way channel is the design of the scheduler. The *optimal* scheduler will choose the multicast order corresponding to the minimum bandwidth expansion factor among all possible permutations. The following result argues for the efficiency of our proposed cooperation scheme for the three-way channel.

Theorem 8: The proposed multicast with side information scheme with the optimal scheduler has the following properties (in the three-way channel).
1) It is asymptotically optimal, i.e., approaches the genie-aided bound, when any one of the channel coefficients is sufficiently large.
2) It always outperforms the broadcast with a degraded set based multicast scheme with the optimal scheduler.
   Proof: Please refer to Appendix X. □

$$H(S_1 \mid S_2) \leq \tau_1 R_{\text{CF2}}(\alpha)$$

$$H(S_1 \mid S_3) - \frac{\alpha \min\{I(X_1; Y_3 \mid m_1); I(X_1; \hat{Y}_3, Y_2 \mid m_1)\} H(S_1 \mid S_2)}{R_{\text{CF2}}(\alpha)} \leq \tau_0 C\left(\left(h_{13}^2 + h_{23}^2\right)P\right) \quad (38)$$
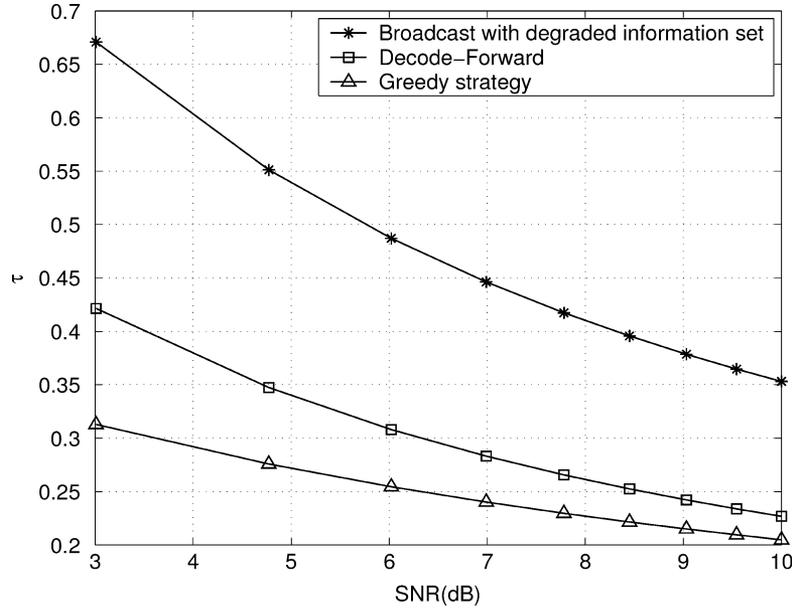
Fig. 13. The bandwidth expansion factor of various schemes in the multicast channel with side information, $h_{12} = 3$ dB, $h_{13} = 0$ dB, $h_{23} = 19.5$ dB, $H(S_1 \mid S_2) = 0.9, H(S_1 \mid S_3) = 0.3$.

To reduce the computational complexity of the scheduler, one can adopt the following greedy strategy. At the beginning of every multicast stage, every node that has not finished multicasting yet will calculate its expected bandwidth expansion factor based on the cooperative scheme for multicast with side information. The greedy scheduler chooses the node with the least expected bandwidth expansion factor to transmit at this stage. After this node finishes and the side information is updated, the scheduler computes the expected bandwidth expansion factor for the rest of the nodes and selects the one with the least bandwidth expansion factor to multicast next. In general, this greedy scheduler constitutes a potential source for further suboptimality. However, it approaches the genie-aided bound in the asymptotic limit when one of the channel gains is sufficient-lylarge. Take $h_{23} \to \infty$ as an example. In this case, one can easily verify that, if one of the following conditions is satisfied, then the greedy scheduler will approach the genie-aided bound:

1) $H(S_2 \mid S_1) < \min\{H(S_1 \mid S_2), H(S_1 \mid S_3), H(S_3 \mid S_1)\}$ and $H(S_3 \mid S_1, S_2) < H(S_1 \mid S_2, S_3)$,
2) $H(S_3 \mid S_1) < \min\{H(S_1 \mid S_2), H(S_1 \mid S_3), H(S_2 \mid S_1)\}$ and $H(S_2 \mid S_1, S_3) < H(S_1 \mid S_2, S_3)$.

The numerical results in Figs. 15 and 16 validate our claims on the efficiency of the proposed cooperation strategy. These figures compare the minimum energy required per source observation by each scheme. In our experiments, we randomly generate the channel coefficients according to a normalized zero-mean Gaussian distribution. For each source realization, we randomly generate $H(S_1), H(S_2), H(S_3)$ with uniform distribution in $[0, 1]$ (corresponds to binary source). To ensure $0 \leq H(S_1 \mid S_2) \leq H(S_1)$, we generate $H(S_1, S_2)$ according to uniform distribution in

$$[\max\{H(S_1), H(S_2)\}, H(S_1) + H(S_2)].$$

Similarly, we generate $H(S_1, S_3)$, $H(S_2, S_3)$, $H(S_1, S_2, S_3)$ to ensure that all the entropy inequalities are satisfied. Then,

we use the entropy equalities to compute all other conditional entropy [35]. For example, $H(S_1 \mid S_2, S_3)$ is given

$$H(S_1 \mid S_2, S_3) = H(S_1, S_2, S_3) - H(S_2, S_3).$$

In such way, for each realization, we get different source entropies and correlation patterns that satisfy the appropriate entropy inequalities. For each (source and channel) realization, we then use numerical methods to find the optimal order and greedy order for cooperative scheme, and the corresponding minimum energy required per source observation, namely, $E_{\mathrm{oc}}, E_{\mathrm{gc}}$. We also find the optimal order for the noncooperative scheme and the corresponding minimum energy required per source observation E_{nc}. The minimum energy required per source observation by the genie-aided bound E_{gen} is used as a benchmark. In particular, for each realization, we calculate the ratio of the minimum energy required by the three schemes to the genie aided bound. We repeat the experiment 100 000 times and report the histogram of the ratios in the figures. In Fig. 15, we see that 94% of the time, the proposed cooperative scheme with the greedy scheduler operates within 3 dB of the genie-aided bound. We also see that the performance of greedy scheduler is almost identical to the optimal scheduler. Fig. 16 shows that the noncooperative scheme operates more than 3 dB away from the genie-aided bound for 90% of the time. Moreover, there is a nonnegligible probability, i.e., 8%, that this scheme operates 100 dB away from genie-aided bound. It is clear that receiver cooperation reduces this probability significantly.

V. CONCLUSION

We have adopted a formulation of the three-node wireless network based on the half-duplex and total power constraints. In this setup, we have proposed a greedy cooperation strategy in which the *weak* receiver first helps the *strong* receiver to decode in a CF configuration. After successfully decoding, the strong user starts assisting the weak user in a DF configuration.
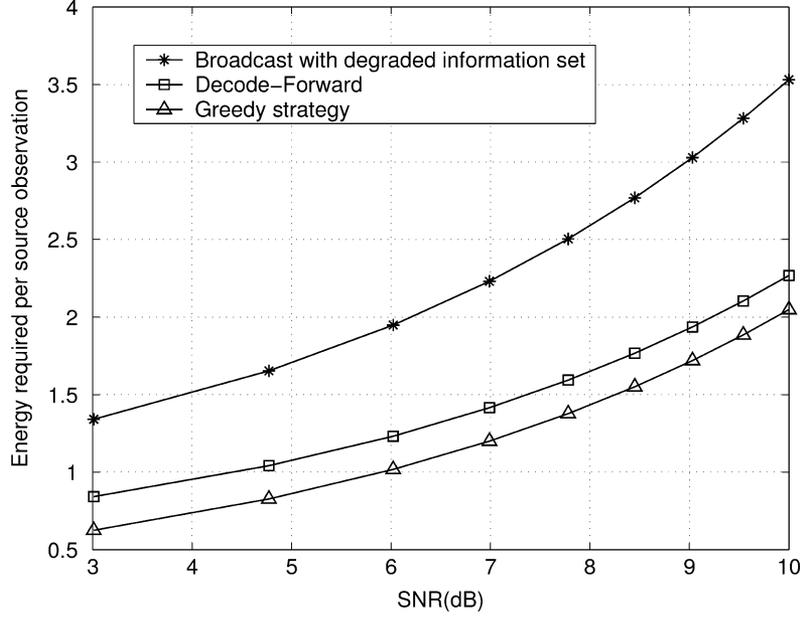
Fig. 14. The energy required per source observation of various schemes in the multicast channel with side information, $h_{12} = 3$ dB, $h_{13} = 0$ dB, $h_{23} = 19.5$ dB, $H(S_1 \mid S_2) = 0.9, H(S_1 \mid S_3) = 0.3$.

We have shown that different instantiations of this strategy yield excellent performance in the relay channel with noisy feedback, multicast channel, and three-way channel. Our analysis for the achievable rates in such special cases elucidates the value of noisy feedback in relay channel and the need for a list source–channel decoding approach to efficiently exploit receiver side information in the wireless setting.

Extending our work to networks with an arbitrary number of nodes appears to be a natural next step. In particular, the generalization of the greedy cooperation strategy is an interesting avenue worthy of further research. Our preliminary investigations reveal that such a strategy can get sizable performance gains, over the traditional multihop routing approach, in certain network configurations.

APPENDIX I
PROOF OF THEOREM 1

In this paper, we refer to typical sequences as strong typical sequences (see [8], [10], [29] for details of strong typical sequences).

A. Discrete Memoryless Channel

1) Outline: Suppose we want to send an independent and identically distributed (i.i.d.) source $w(i), w(i) \in [1, M]$, in which $M = 2^{NR}$ to the destination. Equally divide these 2^{NR} messages into $M_1 = 2^{N\alpha t R_1}$ cells, index the cell number as $b(i)$. Index the element in every cell as $d(i), d(i) \in [1, M_2], M_2 = 2^{N(1-t) \cdot R_2}$. Thus,

$$2^{NR} = 2^{N\alpha t R_1} 2^{N(1-t)R_2} \qquad (44)$$

that is,

$$R = \alpha t R_1 + (1 - t)R_2. \qquad (45)$$

The main idea is that the relay and the destination help each other to decode $b(i)$.

- In the first state m_1, the source sends the cell index $b(i)$ to both the relay and the destination. At this time, neither the relay nor the destination can decode this information.
- In the feedback state m_3, the destination sends the compressed version of the received noisy signal to the destination. At the same time, the source sends additional information to the relay.
- At the end of the relay receive mode, the relay gets an estimation of $b(i)$, namely, $\hat{b}(i)$. Thus, in m_2, the relay sends its knowledge of $\hat{b}(i)$ to the destination to help it decode $b(i)$. At the same time, the source sends $d(i)$ to the destination.

2) Random Code Generation: Fix

$$p(x_1 \mid m_1), p(x_1, x_3 \mid m_3), p(x_1, x_2 \mid m_2), p(\hat{y}_3).$$

- state m_1:
  At the source, generate $2^{\alpha t N R_1}$ i.i.d. length-$\alpha t N$ sequence $\boldsymbol{x}_{1,m_1}$ each with probability

$$p(\boldsymbol{x}_{1,m_1}) = \prod_{j=1}^{\alpha t N} p(x_{1j} \mid m_1).$$

  Label these sequences as $\boldsymbol{x}_{1,m_1}(b)$, where $b \in [1, 2^{\alpha t N R_1}]$ is called the cell index.
- state m_3:
  — source node:
    Generate $2^{(1-\alpha)t N R_4}$ i.i.d. length-$(1 - \alpha)t N$ codewords $\boldsymbol{x}_{1,m_3}$ with

$$p(\boldsymbol{x}_{1,m_3}) = \prod_{j=1}^{(1-\alpha)t N} p(x_{1j} \mid m_3).$$

  Label these sequences as

$$\boldsymbol{x}_{1,m_3}(q), \quad q \in [1, 2^{(1-\alpha)t N R_4}].$$
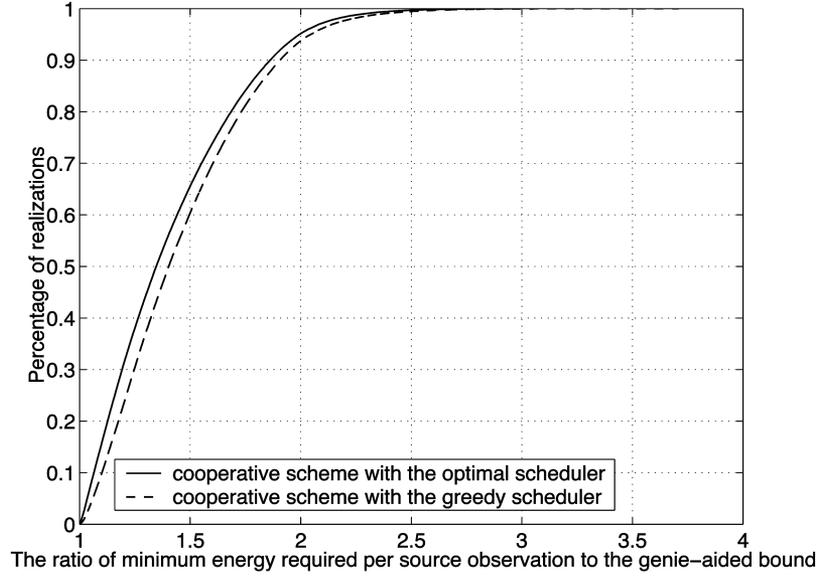
Fig. 15.   The ratio of the energy required per source observation of the proposed schemes to the genie-aided bound.
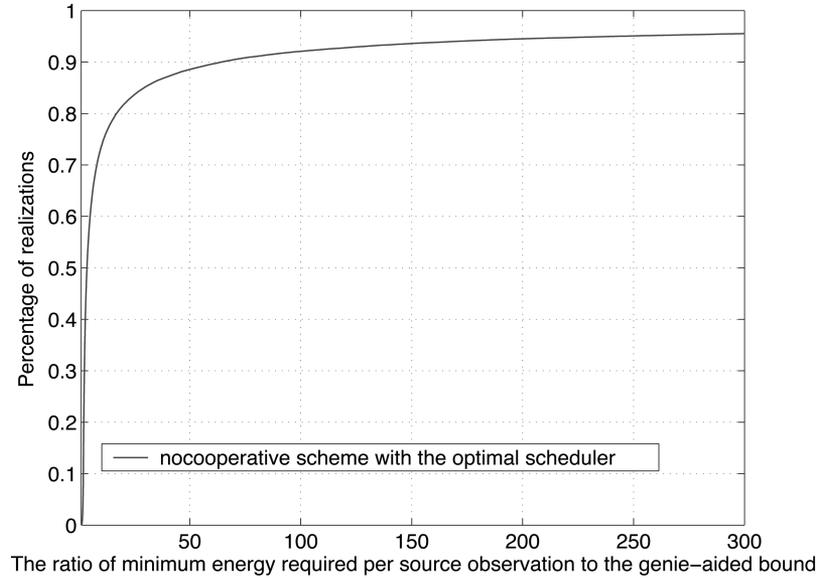
Fig. 16.   The ratio of the energy required per source observation of the noncooperative scheme to the genie-aided bound.

Randomly partition the $2^{\alpha tNR_1}$ cell indices $\{b\}$ into $2^{(1-\alpha)tNR_4}$ bins Q_q with $q \in [1, 2^{(1-\alpha)tNR_4}]$.

— destination node:
Generate $2^{(1-\alpha)tNR_3}$ i.i.d. length-$(1-\alpha)tN$ codewords $\boldsymbol{x}_{3,m_3}$ with

$$p(\boldsymbol{x}_{3,m_3}) = \prod_{j=1}^{(1-\alpha)tN} p(x_{3j} \,|\, m_3).$$

Index them as $\boldsymbol{x}_{3,m_3}(u)$. Generate $2^{\alpha tN\hat{R}}$ i.i.d. length-αtN sequences $\hat{\boldsymbol{y}}_3(z)$ with

$$p(\hat{\boldsymbol{y}}_3) = \prod_{j=1}^{\alpha tN} p(\hat{y}_{3j}).$$

Randomly partition the set $z \in [1, 2^{\alpha tN\hat{R}}]$ into bins $U_u, u \in [1, 2^{(1-\alpha)tNR_3}]$.

• state m_2:
— relay node:
Randomly generate $M_0 = 2^{(1-t)NR_0}$ i.i.d. length-$(1-t)N$ sequences $\boldsymbol{x}_{2,m_2}$ with

$$p(\boldsymbol{x}_{2,m_2}) = \prod_{j=1}^{(1-t)N} p(x_{2j} \,|\, m_2).$$

Index them as

$$\boldsymbol{x}_{2,m_2}(c), c \in [1, 2^{(1-t)NR_0}].$$

Randomly partition the $2^{\alpha t N R_1}$ cell indices into $2^{(1-t)NR_0}$ bins C_c.
— source node:
Generate $M_2 = 2^{(1-t)NR_2}$ i.i.d. length-$(1-t)N$ sequences $\boldsymbol{x}_{1,m_2}$ with

$$p(\boldsymbol{x}_{1,m_2}) = \prod_{j=1}^{(1-t)N} p(x_{1j}\,|\,x_{2j}, m_2)$$

for every $\boldsymbol{x}_{2,m_2}$ sequence. Index them as

$$\boldsymbol{x}_{1,m_2}(d\,|\,c), d \in [1, 2^{(1-t)NR_2}].$$

3) Encoding: Partition the source message set into $2^{\alpha t N R_1}$ equal-sized cells. Let $w(i)$ be the message to be sent in block i. Suppose $w(i)$ is the $d(i)$th message in cell-$b(i)$ and the cell index $b(i)$ is in bin-$q(i)$ and bin-$c(i)$, respectively. For brevity we drop the block index i in the following.

- state m_1:
  The source sends $\boldsymbol{x}_{1,m_1}(b)$.
- state m_3:
  — The source node knows that the cell index b is in bin-q, so it sends $\boldsymbol{x}_{1,m_3}(q)$.
  — The destination first selects $\hat{\boldsymbol{y}}_3(z)$ that is jointly typical with $\boldsymbol{y}_{3,m_1}$. It then sends $\boldsymbol{x}_{3,m_3}(u)$ where z is in the bin U_u.
- state m_2:
  — Knowing the cell index b is in bin-c, the source node sends the corresponding $\boldsymbol{x}_{1,m_3}(d\,|\,c)$.
  — Using the information received in state m_1 and m_3, the relay gets an estimation of the cell index $\hat{\hat{b}}$. Suppose $\hat{\hat{b}}$ is in bin-$\hat{\hat{c}}$. Then it sends $\boldsymbol{x}_{2,m_3}(\hat{\hat{c}})$.

4) Decoding: In the following, code length N is chosen sufficiently large.

- at the end of m_1:
  The destination has received $\boldsymbol{y}_{3,m_1}$ and it decides a sequence $\hat{\boldsymbol{y}}_3(z)$ if $(\hat{\boldsymbol{y}}_3(z), \boldsymbol{y}_{3,m_1})$ are jointly typical. There exists such a z with high probability if

$$\hat{R} \geq I(\hat{Y}_3; Y_3). \tag{46}$$

- at the end of m_3:
  At this stage, only the relay decodes the message.
  — The relay estimates u by looking for the unique $\hat{u}$ such that $(\boldsymbol{x}_{3,m_3}(\hat{u}), \boldsymbol{y}_{2,m_3})$ are jointly typical. $\hat{u} = u$ with high probability if

$$R_3 \leq I(X_3; Y_2). \tag{47}$$

  — Knowing $\hat{u}$, the relay tries to decode q by selecting the unique $\hat{q}$ such that $(\boldsymbol{x}_{1,m_3}(\hat{q}), \boldsymbol{x}_{3,m_3}(\hat{u}), \boldsymbol{y}_{2,m_3})$ are jointly typical. $\hat{q} = q$ with high probability if

$$R_4 \leq I(X_1; Y_2\,|\,X_3). \tag{48}$$

  — The relay calculates a list $\ell(\boldsymbol{y}_{2,m_1})$ such that $z \in \ell(\boldsymbol{y}_{2,m_1})$ if $(\hat{\boldsymbol{y}}_{3,m_1}(z), \boldsymbol{y}_{2,m_1})$ are jointly typical.

Assuming u decoded successfully at the relay, $\hat{z}$ is selected if it is the unique $\hat{z} \in U_u \bigcap \ell(\boldsymbol{y}_{2,m_1})$. Using the same argument as in [10], it can be shown that $\hat{z} = z$ occurs with high probability if

$$\alpha t \hat{R} \leq \alpha t I(\hat{Y}_3; Y_2\,|\,m_1) + (1-\alpha)t R_3. \tag{49}$$

  — The relay computes another list $\ell(\boldsymbol{y}_{2,m_1}, \hat{\boldsymbol{y}}_{3,m_1})$ such that $b \in \ell(\boldsymbol{y}_{2,m_1}, \hat{\boldsymbol{y}}_{3,m_1})$ if $(\boldsymbol{x}_{1,m_1}(b), \boldsymbol{y}_{2,m_1}, \hat{\boldsymbol{y}}_{3,m_1})$ are jointly typical.
  — Finally, the relay declares $\hat{\hat{b}}$ is received if it is the unique $\hat{\hat{b}} \in Q_q \bigcap \ell(\boldsymbol{y}_{2,m_1}, \hat{\boldsymbol{y}}_{3,m_1})$. Using the same arguement as in [10], one can show $\hat{\hat{b}} = b$ with high probability if

$$\begin{aligned} \alpha t R_1 \leq {} & \alpha t I(X_1; \hat{Y}_3, Y_2\,|\,m_1) \\ & + (1-\alpha)t I(X_1; Y_2\,|\,X_3, m_3). \end{aligned} \tag{50}$$

- at the end of m_2:
  — The destination declares that $\hat{c}$ was sent from the relay if there exists one and only one $\hat{c}$ such that $(\boldsymbol{x}_{2,m_2}(\hat{c}), \boldsymbol{y}_{3,m_2})$ are jointly typical. Then $\hat{c} = c$ with high probability if

$$R_0 \leq I(X_2; Y_3\,|\,m_2). \tag{51}$$

  — After decoding $\hat{c}$, the destination further declares that $\hat{d}$ was sent from the source if it is the unique $\hat{d}$ such that $(\boldsymbol{x}_{1,m_2}(\hat{d}), \boldsymbol{x}_{2,m_2}(\hat{c}), \boldsymbol{y}_{3,m_2},)$ are joint typical. Assuming c decoded correctly, the probability of error of $\hat{d}$ is small if

$$R_2 \leq I(X_1; Y_3\,|\,X_2, m_2). \tag{52}$$

  — At first, the destination calculates a list $\ell(\boldsymbol{y}_{3,m1})$, such that $b \in \ell(\boldsymbol{y}_{3,m1})$ if $(\boldsymbol{x}_{1,m_1}(b), \boldsymbol{y}_{3,m1})$ are jointly typical. Assuming c decoded successfully at the destination, $\hat{b}$ is declared to be the cell index if there is a unique $\hat{b} \in C_c \bigcap \ell(\boldsymbol{y}_{3,m_1})$. As in [10], the decoding error is small if

$$\begin{aligned} \alpha t R_1 & \leq \alpha t I(X_1; Y_3\,|\,m_1) + (1-t)R_0 \\ & \leq \alpha t I(X_1; Y_3\,|\,m_1) + (1-t)I(X_2; Y_3\,|\,m_2). \end{aligned} \tag{53}$$

  From the cell index $\hat{b}$ and the message index $\hat{d}$ within the cell, the destination can recover the source message.

Combining (50) and (52), we have

$$\begin{aligned} R < {} & \alpha t I(X_1; \hat{Y}_3, Y_2\,|\,m_1) + (1-\alpha)t I(X_1; Y_2\,|\,X_3, m_3) \\ & + (1-t)I(X_1; Y_3\,|\,X_2, m_2). \end{aligned} \tag{54}$$

It follows from (53) and (52) that

$$R < \alpha t I(X_1; Y_3\,|\,m_1) + (1-t)I(X_1, X_2; Y_3\,|\,m_2). \tag{55}$$

From (46) and (49), we have the constraint

$$(1-\alpha)I(X_3; Y_2) > \alpha I(\hat{Y}_3; Y_3\,|\,Y_2, m_1). \tag{56}$$

Thus, if (54), (55), and (56) are satisfied, there exist a channel code that makes the decoding error at destination less than ϵ.

As mentioned in [8], strong typicality does not apply to continuous random variables in general, but for Gaussian input distributions, one can generalize the Markov lemma along the lines of [36], [37] and thereby the DMC result derived above applies to the Gaussian inputs $(X_1, X_2, X_3, \hat{Y})$. Since $\hat{Y}_3$ is a degraded version of Y_3, we write $\hat{Y}_3 = Y_3 + Z'$ where Z' is Gaussian noise with variance σ_3^2 (see [9], [21] for a similar analysis).

First, we examine the constraint (56) under the Gaussian inputs

$$
\begin{aligned}
I(X_3; Y_2 \,|\, m_3) &= h(Y_2 \,|\, m_3) - h(Y_2 \,|\, X_3, m_3) \\
&= \frac{1}{2} \log \left(\frac{1}{1 - \rho_{Y_2, X_3}^2} \right)
\end{aligned}
\tag{57}
$$

and

$$
\begin{aligned}
\rho_{Y_2, X_3}^2 &= \frac{E^2\{(h_{12}X_1 + h_{23}X_3 + Z_2)X_3\}}{\mathrm{Var}(Y_2)\mathrm{Var}(X_3)} \\
&= \frac{\left(h_{12}r_{13}\sqrt{P_1^{(3)}} + h_{23}\sqrt{P_3^{(3)}} \right)^2}{h_{12}^2 P_1^{(3)} + h_{23}^2 P_3^{(3)} + \sigma^2 + 2h_{12}h_{23}r_{13}\sqrt{P_1^{(3)}P_3^{(3)}}}.
\end{aligned}
\tag{58}
$$

Hence we get (59) at the bottom of the page.

We observe that the correlation coefficient $r_{13} = 0$ because neither the source nor the destination knows the codeword sent by the other duing the feedback state. Thus, one has

$$
\begin{aligned}
I(X_3; Y_2 \,|\, m_3) &= \frac{1}{2} \log \left(1 + \frac{h_{23}^2 P_3^{(3)}}{h_{12}^2 P_1^{(3)} + \sigma^2} \right) \\
&= C \left(\frac{h_{23}^2 P_3^{(3)}}{h_{12}^2 P_1^{(3)} + \sigma^2} \right).
\end{aligned}
\tag{60}
$$

Similarly, one has

$$
\begin{aligned}
&I(\hat{Y}_3; Y_3 \,|\, Y_2, m_1) \\
&= h(\hat{Y}_3 \,|\, m_1) - \frac{1}{2} \log \left(\frac{1}{1 - \rho_{\hat{Y}_3 Y_2}^2} \right) \\
&\quad - h(\hat{Y}_3 \,|\, Y_3, m_1) \\
&= h(h_{13}X_1 + Z_3 + Z') - h(h_{13}X_1 + Z_3 + Z' \,|\, h_{13}X_1 + Z_3) \\
&\quad - \frac{1}{2} \log \left(\frac{1}{1 - \rho_{\hat{Y}_3 Y_2}^2} \right) \\
&= \frac{1}{2} \log \left(\frac{h_{13}^2 P_1^{(1)} + \sigma^2 + \sigma_3^2}{\sigma_3^2} \right) - \frac{1}{2} \log \left(\frac{1}{1 - \rho_{\hat{Y}_3 Y_2}^2} \right)
\end{aligned}
\tag{61}
$$

where

$$
\begin{aligned}
\rho_{\hat{Y}_3 Y_2}^2 &= \frac{E^2(\hat{Y}_3 Y_2)}{\mathrm{Var}(\hat{Y}_3)\mathrm{Var}(Y_2)} \\
&= \frac{E^2\{(h_{13}X_1 + Z_3 + Z')(h_{12}X_1 + Z_2)\}}{\mathrm{Var}(\hat{Y}_3)\mathrm{Var}(Y_2)} \\
&= \frac{\left(h_{12}h_{13}P_1^{(1)} \right)^2}{\left(h_{13}^2 P_1^{(1)} + \sigma^2 + \sigma_3^2 \right)\left(h_{12}^2 P_1^{(1)} + \sigma^2 \right)}.
\end{aligned}
\tag{62}
$$

So

$$
\begin{aligned}
&I(\hat{Y}_3; Y_3 \,|\, Y_2, m_1) \\
&= \frac{1}{2} \log \left(1 + \frac{\sigma^2}{\sigma_3^2} + \frac{1}{\sigma_3^2} \left(\frac{\sigma^2 h_{13}^2 P_1^{(1)}}{h_{12}^2 P_1^{(1)} + \sigma^2} \right) \right).
\end{aligned}
\tag{63}
$$

Setting

$$
(1 - \alpha)tI(X_3; Y_2) = \alpha t I(\hat{Y}_3; Y_3 \,|\, m_1) - \alpha t I(\hat{Y}_3; Y_2 \,|\, m_1)
\tag{64}
$$

to solve for σ_3^2

$$
\sigma_3^2 = \frac{\sigma^2 + \frac{\sigma^2 h_{13}^2 P_1^{(1)}}{h_{12}^2 P_1^{(1)} + \sigma^2}}{\left(1 + \frac{h_{23}^2 P_3^{(3)}}{h_{12}^2 P_1^{(3)} + \sigma^2} \right)^{\frac{1-\alpha}{\alpha}} - 1}.
\tag{65}
$$

Next, we examine the achievable rate expression (54).

$$
\begin{aligned}
I(X_1; \hat{Y}_3, Y_2 \,|\, m_1) &= \frac{1}{2} \log \left(1 + \frac{h_{12}^2 P_1^{(1)}}{\sigma^2} + \frac{h_{13}^2 P_1^{(1)}}{\sigma^2 + \sigma_3^2} \right) \\
&= C \left(\frac{h_{12}^2 P_1^{(1)}}{\sigma^2} + \frac{h_{13}^2 P_1^{(1)}}{\sigma^2 + \sigma_3^2} \right) \\
I(X_1; Y_2 | X_3, m_3) &= h(Y_2 | X_3, m_3) - h(Y_2 | X_1, X_3, m_3) \\
&= \frac{1}{2} \log \left(1 + \frac{h_{12}^2 P_1^{(3)}}{\sigma^2} \right) = C \left(\frac{h_{12}^2 P_1^{(3)}}{\sigma^2} \right) \\
I(X_1; Y_3 | X_2, m_2) &= h(Y_3 | X_2, m_2) - h(Y_3 | X_1, X_2, m_2) \\
&= \frac{1}{2} \log \left(1 + \frac{\left(1 - r_{12}^2\right) h_{13}^2 P_1^{(2)}}{\sigma^2} \right) \\
&= C \left(\frac{\left(1 - r_{12}^2\right) h_{13}^2 P_1^{(2)}}{\sigma^2} \right).
\end{aligned}
\tag{66}
$$

Combining them together, we get

$$
\begin{aligned}
&\alpha t I(X_1; \hat{Y}_3, Y_2 \,|\, m_1) + (1 - \alpha)t I(X_1; Y_2 \,|\, X_3, m_3) \\
&\quad + (1 - t)I(X_1; Y_3 \,|\, X_2, m_2) \\
&= \alpha t C \left(\frac{h_{12}^2 P_1^{(1)}}{\sigma^2} + \frac{h_{13}^2 P_1^{(1)}}{\sigma^2 + \sigma_3^2} \right) \\
&\quad + (1 - \alpha)t C \left(\frac{h_{12}^2 P_1^{(3)}}{\sigma^2} \right) \\
&\quad + (1 - t)C \left(\frac{\left(1 - r_{12}^2\right) h_{13}^2 P_1^{(2)}}{\sigma^2} \right).
\end{aligned}
\tag{67}
$$

$$
I(X_3; Y_2 \,|\, m_3) = \frac{1}{2} \log \left(\frac{h_{12}^2 P_1^{(3)} + h_{23}^2 P_3^{(3)} + \sigma^2 + 2h_{12}h_{23}r_{13}\sqrt{P_1^{(3)}P_3^{(3)}}}{h_{12}^2(1 - r_{13}^2)P_1^{(3)} + \sigma^2} \right).
\tag{59}
$$

Similarly for (55), one has

$$
\begin{aligned}
&I(X_1; Y_3 \mid m_1) \\
&= h(h_{13}X_1 + Z_3 \mid m_1) - h(h_{13}X_1 + Z_3 \mid X_1, m_1) \\
&= \frac{1}{2}\log\left(1 + \frac{h_{13}^2 P_1^{(1)}}{\sigma^2}\right) = C\left(\frac{h_{13}^2 P_1^{(1)}}{\sigma^2}\right)
\end{aligned}
$$

$$
\begin{aligned}
&I(X_1, X_2; Y_3 \mid m_2) \\
&\quad = h(h_{13}X_1 + h_{23}X_2 + Z_3 \mid m_2) \\
&\qquad - h(h_{12}X_1 + h_{23}X_2 + Z_3 \mid X_1, X_2, m_2) \\
&\quad = \frac{1}{2}\log\left(1 + \frac{h_{13}^2 P_1^{(2)} + h_{23}^2 P_2^{(2)} + 2h_{13}h_{23}r_{12}\sqrt{P_1^{(2)}P_2^{(2)}}}{\sigma^2}\right) \\
&\quad = C\left(\frac{h_{13}^2 P_1^{(2)} + h_{23}^2 P_2^{(2)} + 2h_{13}h_{23}r_{12}\sqrt{P_1^{(2)}P_2^{(2)}}}{\sigma^2}\right) \qquad (68)
\end{aligned}
$$

which gives rise to

$$
\begin{aligned}
&\alpha t I(X_1; Y_3 | m_1) + (1-t)I(X_1, X_2; Y_3 | m_2) \\
&= \frac{\alpha t}{2}\log\left(1 + \frac{h_{13}^2 P_1^{(1)}}{\sigma^2}\right) \\
&\quad + \frac{1-t}{2}\log\left(1 + \frac{h_{13}^2 P_1^{(2)} + h_{23}^2 P_2^{(2)} + +2h_{13}h_{23}r_{12}\sqrt{P_1^{(2)}P_2^{(2)}}}{\sigma^2}\right) \\
&= \frac{\alpha t}{2}C\left(\frac{h_{13}^2 P_1^{(1)}}{\sigma^2}\right) \\
&\quad + (1-t)C\left(\frac{h_{13}^2 P_1^{(2)} + h_{23}^2 P_2^{(2)} + 2h_{13}h_{23}r_{12}\sqrt{P_1^{(2)}P_2^{(2)}}}{\sigma^2}\right).
\end{aligned}
$$

$$(69)$$

Setting the noise variance $\sigma^2 = 1$, the proof is complete.

APPENDIX III
PROOF OF LEMMA 3

We only show the upper bound of R_{FB}. The proof for R_{CF} is similar and thus omitted. Setting shorthand notation $\Delta = 1 + \frac{h_{23}^2 P_3^{(3)}}{h_{12}^2 P_1^{(3)} + 1}$, one has from (15) that

$$
\begin{aligned}
&1 + \frac{h_{13}^2 P_1^{(1)}}{1 + \sigma_3^2} + h_{12}^2 P_1^{(1)} \\
&= \frac{h_{13}^2 P_1^{(1)}}{1 + \frac{(h_{12}^2 + h_{13}^2)P_1^{(1)} + 1}{(h_{12}^2 P_1^{(1)} + 1)\left(\Delta^{\frac{1-\alpha}{\alpha}} - 1\right)}} + \left(1 + h_{12}^2 P_1^{(1)}\right) \\
&= \frac{h_{13}^2 P_1^{(1)}\left(h_{12}^2 P_1^{(1)} + 1\right)\left(\Delta^{\frac{1-\alpha}{\alpha}} - 1\right)}{\Delta^{\frac{1-\alpha}{\alpha}}\left(h_{12}^2 P_1^{(1)} + 1\right) + h_{13}^2 P_1^{(1)}} + \left(1 + h_{12}^2 P_1^{(1)}\right) \\
&\leq \left(1 + h_{12}^2 P_1^{(1)}\right)\Delta^{\frac{1-\alpha}{\alpha}}. \qquad (70)
\end{aligned}
$$

Hence,

$$
\begin{aligned}
&\alpha t C\left(\left(\frac{h_{13}^2}{1 + \sigma_3^2} + h_{12}^2\right)P_1^{(1)}\right) + (1-\alpha)t C\left(h_{12}^2 P_1^{(3)}\right) \\
&\leq \alpha t C\left(h_{12}^2 P_1^{(1)}\right) + (1-\alpha)t C\left(\frac{h_{23}^2 P_3^{(3)}}{h_{12}^2 P_1^{(3)} + 1}\right) \\
&\quad + (1-\alpha)t C\left(h_{12}^2 P_1^{(3)}\right) \\
&= \alpha t C\left(h_{12}^2 P_1^{(1)}\right) + (1-\alpha)t C\left(h_{12}^2 P_1^{(3)} + h_{23}^2 P_3^{(3)}\right)
\end{aligned}
$$

$$(71)$$

which proves (18).

APPENDIX IV
PROOF OF THEOREM 2

In view of R_{DF} in (8), $h_{12}^2 \leq h_{13}^2$ implies that

$$
\begin{aligned}
R_{\mathrm{DF}} &\leq t C\left(h_{12}^2 P_1^{(1)}\right) + (1-t)C\left(\left(1 - r_{12}^2\right)h_{13}^2 P_1^{(2)}\right) \\
&\leq C\left(h_{13}^2 P\right) = R_{\mathrm{ro}} \qquad (72)
\end{aligned}
$$

where we have used the total power constraint (17). To prove 2), consider the upper bound for R_{CF} in (19). Given the total power constraint $P_1^{(2)} + P_2^{(2)} \leq P$, it is easy to verify that

$$
h_{13}^2 P_1^{(2)} + h_{23}^2 P_2^{(2)} \leq \max\{h_{13}^2, h_{23}^2\}P.
$$

Therefore, the condition $h_{23}^2 \leq h_{13}^2$ implies that

$$
\begin{aligned}
R_{\mathrm{CF}} &\leq t C\left(h_{13}^2 P_1^{(1)}\right) + (1-t)C\left(h_{13}^2 P_1^{(2)} + h_{23}^2 P_2^{(2)}\right) \\
&\leq C\left(h_{13}^2 P\right) = R_{\mathrm{ro}}. \qquad (73)
\end{aligned}
$$

The last statement of the theorem can be shown in a similar fashion using the R_{FB} upper bound in (18).

APPENDIX V
PROOF OF THEOREM 4

Since $C(h_{12}^2 P) \geq C(h_{13}^2 P)$ by $h_{12}^2 \geq h_{13}^2$, the two line segments in the R_{DF} expression intersect at some optimal $t^* \in (0, 1)$ (see Fig. 3). The corresponding rate is given by (74) at the bottom of the page, where we have set $P_1^{(2)} = P\cos^2\theta$ and $P_2^{(2)} = P\sin^2\theta$ according to the total power constraint. Taking $P \to 0$, the Taylor expansion is sufficient to establish (21). To prove the lower bound in (21), note that $f_1(\theta, r_{12}, h_{13}, h_{23}) \leq (h_{13}^2 + h_{23}^2)$ with equality when $r_{12} = 1$ and $\tan(\theta) = \frac{h_{23}}{h_{13}}$, which, together with $f_2(\theta, r_{12}, h_{13}) \leq h_{13}^2$ also proves the upper bound of S_{DF} in (21).

On the other hand, as $P \to 0$, it is seen from (9) that $\sigma_2^2 \to \infty$, thus showing $S_{\mathrm{CF}} \leq h_{13}^2$. Similar behavior holds for the feedback scheme, that is, $\sigma_3^2 \to \infty$ as $P \to 0$, in which case $S_{\mathrm{FB}} \leq S_{\mathrm{DF}}$ with the optimal α approaches 1.

APPENDIX VI
PROOF OF THEOREM 5

The results for G_{DF} and G_{CF} follow from direct computation of large P limit. We only show the last statement concerning

$$
R_{\mathrm{DF}} = \frac{C(f_1(\theta, r_{12}, h_{13}, h_{23})P)C\left(h_{12}^2 P\right) - C(f_2(\theta, r_{12}, h_{13})P)C\left(h_{13}^2 P\right)}{C(f_1(\theta, r_{12}, h_{13}, h_{23})P) + C\left(h_{12}^2 P\right) - C(f_2(\theta, r_{12}, h_{13})P) - C\left(h_{13}^2 P\right)} \qquad (74)
$$

the feedback scheme. As in the case of decode–forward, the line-crossing point gives the optimal t and the associated rate R_{FB} is given by (75) at the bottom of the page, in which

$$A = \alpha C\left(\left(\frac{h_{13}^2}{1+\sigma_3^2} + h_{12}^2\right)P\right) + (1-\alpha)C\left(h_{12}^2 P \cos^2\psi\right) \tag{76}$$

where we set $P_1^{(3)} = P\cos^2\psi$ and $P_3^{(3)} = P\sin^2\psi$. Taking $P \to \infty$

$$\sigma_3^2 \to \frac{h_{12}^2 + h_{13}^2}{h_{12}^2\left[\left(1+\frac{h_{23}^2}{h_{12}^2}\tan^2\psi\right)^{\frac{1-\alpha}{\alpha}} - 1\right]} \quad \left(=\sigma_3^2(\infty)\right). \tag{77}$$

Denoting

$$f_3(\psi, \alpha, h_{13}, h_{12}, h_{23})$$
$$= \alpha \log\left(\frac{h_{13}^2}{1+\sigma_3^2(\infty)} + h_{12}^2\right) + (1-\alpha)\log h_{12}^2\cos^2\psi$$

one gets (78) also at the bottom of the page. It follows that if $\alpha < 1$

$$G_{\mathrm{FB}} \le \log f_2 \le \log h_{13}^2 \quad \text{(relay-off)} \tag{79}$$

which forces $\alpha = 1$, that is, $G_{\mathrm{FB}} = G_{\mathrm{DF}}$.

APPENDIX VII
PROOF OF THEOREM 6

Here we only prove the Part 1) of this theorem. Parts 2)—5) follow the same lines as the corresponding results in the relay case.

To prove Part 1), it suffices to show the statement for $\alpha = 1$. The capacity of the multicast channel without cooperation is given by $R_{\mathrm{noncoop}} = C(\min\{h_{13}^2, h_{12}^2\}P)$. With the assumption that $h_{12}^2 > h_{13}^2$, we have $R_{\mathrm{noncoop}} = C(h_{13}^2 P)$.

Note that the rate expression of (30) admits the same line-crossing interpretation as in the relay case. Thus, the intersection determines the optimal rate point. Equate the two terms

$$tC\left(h_{12}^2 P_1^{(1)}\right) = tC\left(h_{13}^2 P_1^{(1)}\right) + (1-t)C\left(\left(h_{13}^2 + h_{23}^2\right)P\right) \tag{80}$$

to solve

$$t^* = \frac{C\left(\left(h_{13}^2 + h_{23}^2\right)P\right)}{C\left(\left(h_{13}^2 + h_{23}^2\right)P\right) + C\left(h_{12}^2 P\right) - C\left(h_{13}^2 P\right)} \tag{81}$$

which gives the corresponding rate

$$R_{\mathrm{DF}} = \frac{C\left(\left(h_{13}^2 + h_{23}^2\right)P\right)C\left(h_{12}^2 P\right)}{C\left(\left(h_{13}^2 + h_{23}^2\right)P\right) + C\left(h_{12}^2 P\right) - C\left(h_{13}^2 P\right)}. \tag{82}$$

Therefore, using $h_{12}^2 > h_{13}^2$, one has

$$R_{\mathrm{DF}} - R_{\mathrm{noncoop}}$$
$$= (1 - t^*)\left(C\left(\left(h_{13}^2 + h_{23}^2\right)P\right) - C\left(h_{13}^2 P\right)\right) > 0 \tag{83}$$

which proves the theorem.

APPENDIX VIII
PROOF OF LEMMA 6

Here we first prove the result for the DMC case, then apply the result to the Gaussian channel.

A. Source Coding

Randomly bin all the sequence S_1^K into $2^{K(H(S_1|S_2)+\epsilon)}$ bins by independently generating an index w uniformly distributed on $\{1, 2, \ldots, 2^{K(H(S_1|S_2)+\epsilon)}\}$. Let f_{s1} be the mapping function, such that $w = f_{s1}(s_1^K)$. Independently generate another bin index b for every sequence S_1^K by picking b uniformly from $\{1, 2, \ldots, 2^{KR}\}$. Let $B(b)$ be the set of all sequences S_1^K allocated to bin b. Thus, every source sequence is associated with two bin indexes $\{w, b\}$.

B. Channel Coding

1) Random Code Generation:
- At state m_1, generate $2^{K(H(S_1|S_2)+\epsilon)}$ i.i.d. length-N_1 sequence $\boldsymbol{x}_{1,m_1}$, each with probability

$$p(\boldsymbol{x}_{1,m_1}) = \prod_{j=1}^{N_1} p(x_{1j}|m_1)$$

in which $p(x_1|m_1)$ is the input distribution that maximizes $I(X_1; Y_2)$. Assign every bin index w to one sequence $\boldsymbol{x}_{1,m_1}(w), w \in [1, 2^{K(H(S_1|S_2)+\epsilon)}]$.
- At state m_2, randomly generate 2^{KR} i.i.d. length-N_0 $\boldsymbol{x}_{2,m_2}$ at node-2, each with probability

$$p(\boldsymbol{x}_{2,m_2}) = \prod_{j=1}^{N_0} p(x_{2j}|m_2).$$

Generate 2^{KR} i.i.d. length-N_0 $\boldsymbol{x}_{1,m_2}$ at node-1, each with probability

$$p(\boldsymbol{x}_{1,m_2}) = \prod_{j=1}^{N_0} p(x_{1j}|m_2)$$

in which

$$p(x_1|m_2) = \sum_{x_2} p(x_1, x_2|m_2)$$

and $p(x_1, x_2|m_2)$ is the input distribution that maximizes $I(X_1, X_2; Y_3)$. Associate every bin index b to one sequence pair $\{\boldsymbol{x}_{1,m_2}(b), \boldsymbol{x}_{2,m_2}(b)\}$.

$$\frac{C(f_1(\theta, r_{12}, h_{13}, h_{23})P)A - C(f_2(\theta, r_{12}, h_{13})P)\alpha C\left(h_{13}^2 P\right)}{C(f_1(\theta, r_{12}, h_{13}, h_{23})P) + A - C(f_2(\theta, r_{12}, h_{13})P) - \alpha C\left(h_{13}^2 P\right)}, \tag{75}$$

$$R_{\mathrm{FB}} \sim \frac{1}{2}\log P + \frac{\frac{1}{4}(1-\alpha)\log f_2 \cdot \log P + \frac{1}{4}\left[\log f_1 \cdot \log f_3 - \alpha \log f_2 \cdot \log h_{13}^2\right]}{\frac{1-\alpha}{2}\log P + \frac{1}{2}\left[\log f_1 + \log f_3 - \log f_2 - \alpha \log h_{13}^2\right]}. \tag{78}$$

2) Coding: Suppose we want to send source sequence $s_1^K(i)$ at block i, and $w(i) = f_{s1}(s_1^K(i))$, $s_1^K(i) \in B(b(i))$. For brevity of notation, we drop block index i in the following.

- State m_1:
  Node-1 sends $\boldsymbol{x}_{1,m_1}(w)$.
- State m_2:
  — Node-1 knows s_1^K is in b, so it sends $\boldsymbol{x}_{1,m_2}(b)$.
  — At the end of state m_1, node-2 gets an estimation of $\hat{s}_{21}^K$ (details will be given in the following), and suppose $\hat{s}_{21}^K$ is in bin $\hat{b}$. Then in state m_2 nodes-2 sends the corresponding $\boldsymbol{x}_{2,m_2}(\hat{b})$.

C. Decoding

At the end of state m_1:

- At node-2:
  At first, node-2 looks for the one and only one $\hat{w}$ such that $\{\boldsymbol{x}_{1,m_1}(\hat{w}), \boldsymbol{y}_{2,m_1}\}$ are jointly typical. Then node-2 searches in the bin indexed by $\hat{w}$ for source sequence $\hat{s}_{21}^K$ such that $\{\hat{s}_{21}^K, s_2^K\}$ are jointly typical. If it finds only one such sequence, it declares it has received $\hat{s}_{21}^K$, otherwise, it declares an error.
- At node-3:
  Node-3 calculates a list $\ell(\boldsymbol{y}_{3,m_1})$, such that $w' \in \ell(\boldsymbol{y}_{3,m_1})$ if $\{\boldsymbol{x}_{1,m_1}(w'), \boldsymbol{y}_{3,m_1}\}$ are jointly typical.

At the end of state m_2, only node-3 needs to decode:

- Step 1:
  Node-3 declares it receives $\hat{\hat{b}}$, if $\hat{\hat{b}}$ is the one and only one index such that $\{\boldsymbol{x}_{1,m_2}(\hat{\hat{b}}), \boldsymbol{x}_{2,m_2}(\hat{\hat{b}}), \boldsymbol{y}_{3,m_2}\}$ are jointly typical.
- Step 2:
  Node-3 searches in the bin $B(\hat{\hat{b}})$ for the one and only one source sequence $\hat{s}_{31}^K$, such that $\{\hat{s}_{31}^K, s_3^K\}$ are jointly typical and $f_{s1}(\hat{s}_{31}^K) \in \ell(\boldsymbol{y}_{3,m_1})$. If it finds such a unique one, it declares that $\hat{s}_{31}^K$ is the source sequence. Otherwise, it declares an error.

D. Calculation of Probability of Error

1) Node-2: For node-2 there are following error events:

$$E_0 = \left\{ (s_1^K, s_2^K) \notin A_\epsilon^K \right\} \tag{84}$$

$$E_1 = \{\hat{w} \neq w\} \tag{85}$$

$$E_2 = \Big\{ \exists s_1'^K : s_1'^K \neq s_1^K, f_{s1}\left(s_1'^K\right)$$
$$= w \text{ and } \left(s_1'^K, s_2^K\right) \in A_\epsilon^K \Big\}. \tag{86}$$

And

$$P_e^{N_1,K} = P(E_0 \cup E_1 \cup E_2)$$
$$\leq P(E_0) + P(E_1 \,|\, E_0^c) + P(E_2 \,|\, E_0^c, E_1^c). \tag{87}$$

When K is sufficiently large, using the asymptotic equipartition property (AEP), $P(E_0) \to 0$. Now consider $P(E_1 \,|\, E_0^c)$, if channel code rate is less than the capacity, the receiver will decode channel code with error probability less than ϵ. Here, there are $2^{K(H(S_1 \,|\, S_2)+\epsilon)}$ codewords, and the channel code length is N_1, then the rate of channel code is $\frac{K(H(S_1 \,|\, S_2)+\epsilon)}{N_1}$. Thus, for sufficently large N_1 and K, $P(E_1 \,|\, E_0^c) \leq \epsilon$ if

$$\frac{K(H(S_1 \,|\, S_2) + \epsilon)}{N_1} < \max_{p(x_1)} I(X_1; Y_2 \,|\, m_1) = C_2 \tag{88}$$

which is the same as

$$H(S_1 \,|\, S_2) + \epsilon < \tau_1 C_2. \tag{89}$$

Because the source code rate is $H(S_1 \,|\, S_2) + \epsilon$, using the same argument as in [29], one can get $P(E_2 \,|\, E_0^c, E_1^c) < \epsilon$, if K is sufficiently large. So if (89) is satisfied, and N_1, K are sufficiently large, there exists a source–channel code that make the error probability at node-2

$$P_e^{N_1,K} = P\left(\hat{s}_{21}^K \neq s_1^K\right) \leq 3\epsilon. \tag{90}$$

2) Node-3: For node-3, there are the following error events:

$$E_0 = \left\{ (s_1^K, s_3^K) \notin A_\epsilon^K \right\} \tag{91}$$

$$E_1 = \{\text{node-2 cannot decode successfully}\} \tag{92}$$

$$E_2 = \{\hat{\hat{b}} \neq b\} \tag{93}$$

$$E_3 = \Big\{ \exists s_1'^K : s_1'^K \neq s_1^K, f_{s1}\left(s_1'^K\right) \in \ell(\boldsymbol{y}_3 \,|\, m_1), s_1'^K$$
$$\in B(\hat{\hat{b}}), \left(s_1'^K, s_3^K\right) \in A_\epsilon^K \Big\} \tag{94}$$

$$P_e^{N,K} = P(E_0 \cup E_1 \cup E_2 \cup E_3) \leq P(E_0) + P(E_1)$$
$$+ P\left(E_2 \,|\, E_0^c, E_1^c\right) + P\left(E_3 \,|\, E_0^c, E_1^c, E_2^c\right). \tag{95}$$

When K is sufficiently large, $P(E_0) \to 0$. And if (89) is satisfied, $P(E_1) \leq 3\epsilon$. Now consider $P(E_2 \,|\, E_0^c, E_1^c)$, the channel code rate is $\frac{KR}{N_0}$. So, $P(E_2 \,|\, E_0^c, E_1^c) \leq \epsilon$ for sufficiently large N_0, if

$$\frac{KR}{N_0} \leq \max_{p(x_1,x_2)} I(X_1, X_2; Y_3) = C_{(1,2)-3} \tag{96}$$

that is,

$$R \leq \tau_0 C_{(1,2)-3}. \tag{97}$$

Now consider $P(E_3 \,|\, E_0^c, E_1^c, E_2^c)$:

$$P\left(E_3 \,|\, E_0^c, E_1^c, E_2^c\right)$$
$$= P\left(\exists s_1'^K : s_1'^K \neq S_1^K, f_{s1}\left(s_1'^K\right)\right.$$
$$\left. \in \ell\left(\boldsymbol{Y}_3 \,|\, m_1\right), s_1'^K \in B(b), \left(s_1'^K, S_3^K\right) \in A_\epsilon^K\right)$$
$$= \sum_{(s_1^K, s_3^K)} p\left(s_1^K, s_3^K\right) P\left(\exists s_1'^K \neq s_1^K, f_{s1}\left(s_1'^K\right)\right.$$
$$\left. \in \ell\left(\boldsymbol{y}_3 \,|\, m_1\right), s_1'^K \in B(b), \left(s_1'^K, s_3^K\right) \in A_\epsilon^K\right)$$
$$\leq \sum_{(s_1^K, s_3^K)} p\left(s_1^K, s_3^K\right) \sum_{s_1'^K \neq s_1^K \text{ and } (s_1'^K, s_3^K) \in A_\epsilon^K} P\left(f_{s1}\left(s_1'^K\right)\right.$$
$$\left. \in \ell(\boldsymbol{y}_3 \,|\, m_1), s_1'^K \in B(b)\right)$$
$$= \sum_{(s_1^K, s_3^K)} p\left(s_1^K, s_3^K\right) \sum_{s_1'^K \neq s_1^K \text{ and } (s_1'^K, s_3^K) \in A_\epsilon^K} P\left(f_{s1}\left(s_1'^K\right)\right.$$
$$\left. \in \ell(\boldsymbol{y}_3 \,|\, m_1)\right) P\left(s_1'^K \in B(b)\right)$$
$$\leq \sum_{(s_1^K, s_3^K)} p\left(s_1^K, s_3^K\right) 2^{-K(H(S_1 \,|\, S_2)+\epsilon)}$$
$$\times \|\ell(\boldsymbol{y}_3 \,|\, m_1)\| 2^{-KR} \|A_\epsilon\left(S_1^K \,|\, s_3^K\right)\|$$
$$\leq 2^{-K(H(S_1 \,|\, S_2)+\epsilon)} E\{\|\ell(\boldsymbol{y}_3 \,|\, m_1)\|\} 2^{-KR} 2^{K(H(S_1 \,|\, S_3)+\epsilon)}. \tag{98}$$

Follow the same steps in the [10], one has

$$E\{\|\ell(\boldsymbol{y}_3 \,|\, m_1)\|\} \leq 1 + 2^{K(H(S_1 \,|\, S_2)+\epsilon)} 2^{-N_1(I(X_1; Y_3 \,|\, m_1)-7\epsilon)}.$$

So

$$P(E_3 \mid E_0^c, E_1^c, E_2^c)$$
$$\leq 2^{-K(H(S_1 \mid S_2)+\epsilon)}$$
$$\times \left(1 + 2^{K(H(S_1 \mid S_2)+\epsilon)} 2^{-N_1(I(X_1;Y_3 \mid m_1)-7\epsilon)}\right)$$
$$\times 2^{-KR} 2^{K(H(S_1 \mid S_3)+\epsilon)}$$
$$= 2^{-K\{R-(H(S_1 \mid S_3)+\epsilon)+(H(S_1 \mid S_2)+\epsilon)\}}$$
$$+ 2^{-K\{R+\frac{N_1}{K}(I(X_1;Y_3 \mid m_1)-7\epsilon)-(H(S_1 \mid S_3)+\epsilon)\}}. \quad (99)$$

So if

$$R > H(S_1 \mid S_3) + \epsilon - (H(S_1 \mid S_2) + \epsilon), \quad (100)$$

and

$$R > H(S_1 \mid S_3) + \epsilon - \frac{N_1}{K}(I(X_1;Y_3 \mid m_1) - 7\epsilon)$$
$$> H(S_1 \mid S_3) + \epsilon - \tau_1 I(X_1;Y_3 \mid m_1) \quad (101)$$

and K is sufficiently large, $P(E_3 \mid E_0^c, E_1^c, E_2^c) \leq \epsilon$. Together with (89) and (97), one can get

$$H(S_1 \mid S_3) + \epsilon$$
$$- \frac{\min\{I(X_1;Y_3 \mid m_1) - 7\epsilon, C_2\}(H(S_1 \mid S_2) + \epsilon)}{C_2}$$
$$\leq \tau_0 C_{(1,2)-3}. \quad (102)$$

Thus, if both (89) and (102) are satisfied, there exists a source–channel code that makes the error probability at node-3 $P_e^{N,K} < 6\epsilon$.

Next step is to apply the result to the Gaussian channel. In this case, we have

$$C_2 = C\left(h_{12}^2 P\right)$$
$$I(X_1;Y_3 \mid m_1) = C\left(h_{13}^2 P\right)$$
$$C_{(1,2)-3} = C\left(\left(h_{13}^2 + h_{23}^2\right)P\right). \quad (103)$$

Inserting (89) to (102) and (103) completes the proof.

APPENDIX IX
PROOF OF THEOREM 7

Parts 1) and 2) of this theorem follow straightforward limit calculation, we only prove Part 3).

The assumption $\tau_{\text{ex},2} < \tau_{\text{ex},3}$ becomes

$$\frac{H(S_1 \mid S_2)}{h_{12}^2} < \frac{H(S_1 \mid S_3)}{h_{13}^2}$$

when $P \to 0$. Under this assumption, there are two different cases corresponding to different cost functions for the benchmark scheme

$$H(S_1 \mid S_2) > H(S_1 \mid S_3)$$

and

$$H(S_1 \mid S_2) < H(S_1 \mid S_3).$$

When $H(S_1 \mid S_2) > H(S_1 \mid S_3)$, in which case $h_{12}^2 > h_{13}^2$ and

$$E_{1,m} = \frac{2}{\log e}\left(\frac{H(S_1 \mid S_2)}{h_{12}^2}\right.$$
$$\left. + \left(\frac{1}{h_{13}^2} - \frac{1}{h_{12}^2}\right)^+ H(S_1 \mid S_3)\right) \quad (104)$$

$$E_{2,m} = \frac{2}{\log e}\left(\frac{H(S_1 \mid S_2)}{h_{12}^2}\right.$$
$$\left. + \frac{h_{12}^2 H(S_1 \mid S_3) - h_{13}^2 H(S_1 \mid S_2)}{(h_{13}^2 + h_{23}^2) h_{12}^2}\right)$$
$$< \frac{2}{\log e}\left(\frac{H(S_1 \mid S_2)}{h_{12}^2}\right.$$
$$\left. + \frac{h_{12}^2 H(S_1 \mid S_3) - h_{13}^2 H(S_1 \mid S_3)}{(h_{13}^2 + h_{23}^2) h_{12}^2}\right)$$
$$< \frac{2}{\log e}\left(\frac{H(S_1 \mid S_2)}{h_{12}^2}\right.$$
$$\left. + \frac{h_{12}^2 H(S_1 \mid S_3) - h_{13}^2 H(S_1 \mid S_3)}{h_{13}^2 h_{12}^2}\right)$$
$$\leq \frac{2}{\log e}\left(\frac{H(S_1 \mid S_2)}{h_{12}^2}\right.$$
$$\left. + \left(\frac{1}{h_{13}^2} - \frac{1}{h_{12}^2}\right)^+ H(S_1 \mid S_3)\right)$$
$$= E_{1,m}. \quad (105)$$

When $H(S_1 \mid S_2) < H(S_1 \mid S_3)$

$$E_{1,m} = \frac{2}{\log e}\left(\frac{H(S_1 \mid S_3)}{h_{13}^2} + \left(\frac{1}{h_{12}^2} - \frac{1}{h_{13}^2}\right)^+ H(S_1 \mid S_2)\right) \quad (106)$$

so

$$E_{2,m} = \frac{2}{\log e}\left(\frac{H(S_1 \mid S_2)}{h_{12}^2}\right.$$
$$\left. + \frac{h_{12}^2 H(S_1 \mid S_3) - \min\{h_{13}^2, h_{12}^2\} H(S_1 \mid S_2)}{(h_{13}^2 + h_{23}^2) h_{12}^2}\right)$$
$$< \frac{2}{\log e}\left(\frac{H(S_1 \mid S_2)}{h_{12}^2}\right.$$
$$\left. + \frac{h_{12}^2 H(S_1 \mid S_3) - \min\{h_{13}^2, h_{12}^2\} H(S_1 \mid S_2)}{h_{13}^2 h_{12}^2}\right)$$
$$= \frac{2}{\log e}\left(\frac{H(S_1 \mid S_2)}{h_{12}^2}\right.$$
$$\left. + \frac{H(S_1 \mid S_3)}{h_{13}^2} - \min\left\{\frac{1}{h_{13}^2}, \frac{1}{h_{12}^2}\right\} H(S_1 \mid S_2)\right)$$
$$= \frac{2}{\log e}\left(\frac{H(S_1 \mid S_3)}{h_{13}^2}\right.$$
$$\left. + \left(\frac{1}{h_{12}^2} - \frac{1}{h_{13}^2}\right)^+ H(S_1 \mid S_2)\right)$$
$$= E_{1,m}. \quad (107)$$

APPENDIX X
PROOF OF THEOREM 8

For Part 1) of this theorem, without loss of generality, we only prove the case when $h_{23}^2 \to \infty$. In this case, $\lim_{h_{23}^2 \to \infty} \tau_{2,\text{gen}} = \lim_{h_{23}^2 \to \infty} \tau_{3,\text{gen}} = 0$, $\tau_{\text{gen}} = \tau_{1,\text{gen}}$. In the following, we will show that the genie-aided bound could be approached using the following multicast order $2 \to 3 \to 1$.

When node-2 multicasts S_2^K to both node-3 and node-1 using the proposed cooperative multicast with side-information scheme, from Lemma 7 we know it requires

$$\tau_{2-(3,1)} = \frac{H(S_2 \mid S_3)}{R_{\mathrm{CF}r1d3}(\alpha)}$$
$$+ \frac{H(S_2 \mid S_1) - \alpha H(S_2 \mid S_3) \frac{\min\{I(X_2;Y_1), I(X_2;\hat{Y}_1,Y_3)\}}{R_{\mathrm{CF}r1d3}(\alpha)}}{C\left((h_{12}^2 + h_{13}^2)P\right)}.$$

$R_{\mathrm{CF}r1d3}$ means the achievable rate of the following relay channel using the CF scheme: node-2 is the source, node-1 acts as relay that spends $1 - \alpha$ part of the time in helping destination using CF scheme, and node-3 acts as the destination.

Next consider node-3 multicasts S_3^K to both node-1 and node-2. At this time, node-1 already has S_1, S_2, thus this step requires

$$\tau_{3-(2,1)} = \frac{H(S_3 \mid S_2)}{R_{\mathrm{CF}r1d2}(\alpha)}$$
$$+ \frac{H(S_3 \mid S_1, S_2) - \alpha H(S_2 \mid S_3) \frac{\min\{I(X_3;Y_1), I(X_3;\hat{Y}_1,Y_2)\}}{R_{\mathrm{CF}r1d2}(\alpha)}}{C\left((h_{12}^2 + h_{13}^2)P\right)}.$$

The final step: node-1 multicasts $H(S_1 \mid S_2, S_3)$ to both node-2 and node-3 using the greedy multicast scheme developed in the multicast section, this step requires

$$\tau_{1-(2,3)} = \frac{H(S_1 \mid S_2, S_3)}{R_g}.$$

Thus, the total bandwidth expansion factor of this scheme is

$$\tau = \tau_{2-(3,1)} + \tau_{3-(1,2)} + \tau_{1-(2,3)}. \tag{108}$$

Based on the results on the relay channel and multicast channel $h_{23}^2 \to \infty$, $R_{\mathrm{CF}r1d3} \to \infty$, $R_{\mathrm{CF}r1d2} \to \infty$, and $\lim_{h_{23}^2 \to \infty} R_g = C((h_{12}^2 + h_{13}^2)P)$. Then

$$\lim_{h_{23} \to \infty} \tau = \frac{H(S_2 \mid S_1) + H(S_3 \mid S_1, S_2) + H(S_1 \mid S_2, S_3)}{C\left((h_{12}^2 + h_{13}^2)P\right)}$$
$$= \frac{H(S_2, S_3 \mid S_1) + H(S_1 \mid S_2, S_3)}{C\left((h_{12}^2 + h_{13}^2)P\right)} = \tau_{1,\mathrm{gen}} = \tau_{\mathrm{gen}}. \tag{109}$$

To prove the second part of this theorem, without loss of generality, suppose $1 \to 2 \to 3$ is the optimal multicast order for the scheme that uses broadcast with degraded information set. Then, just use the same order for the list source–channel decoding scheme based multicast with side information. Theorem 7 shows that at every multicast step, the list source–channel decoding scheme outperforms the broadcast with degraded information set. Thus, even with this not necessarily optimal order, the list source–channel decoding scheme outperforms the scheme that uses broadcast with degraded information set with optimal order.

REFERENCES

[1] L. L. Xie and P. R. Kumar, "A network information theory for wireless communication: Scaling laws and optimal operation," *IEEE Trans. Inf. Theory*, vol. 50, no. 5, pp. 748–767, May 2004.

[2] S. Toumpis and A. J. Goldsmith, "Capacity regions for wireless ad hoc networks," *IEEE Trans. Wireless Commun.*, vol. 2, no. 4, pp. 736–748, Jul. 2003.

[3] P. Gupta and P. R. Kumar, "The capacity of wireless networks," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 388–404, Mar. 2000.

[4] M. Grossglauser and D. N. C. Tse, "Mobility increases the capacity of ad hoc wireless networks," *IEEE/ACM Trans. Netw.*, vol. 10, no. 4, pp. 477–486, Aug. 2002.

[5] A. Høst-Madsen, "On the achievable rate for receiver cooperation in ad-hoc networks," in *Proc. 2004 IEEE Int. Symp. Information Theory*, Chicago, IL, Jun./Jul. 2004, p. 272.

[6] A. Høst-Madsen, "Capacity bounds for cooperative diversity," *IEEE Trans. Inf. Theory*, submitted for publication.

[7] M. A. Khojastepour, A. Sabharwal, and B. Aazhang, "Improved achievable rates for user cooperation and relay channels," in *Proc. 2004 IEEE Intl. Symp. Information Theory*, Chicago, IL, Jun./Jul. 2004, p. 4.

[8] 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.

[9] M. Gastpar, G. Kramer, and P. Gupta, "The multiple-relay channel: Coding and antenna-clustering capacity," in *Proc. IEEE Int. Symp. Information Theory*, Lausanne, Switzerland, Jun./Jul. 2002, p. 136.

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

[11] M. A. Khojastepour, A. Sabharwal, and B. Aazhang, "Cut-set theorems for multi-state networks," *IEEE Trans. Inf. Theory*, submitted for publication.

[12] E. van der Meulen, "A survey of multi-way channels in information theory: 1961–1976," *IEEE Trans. Inf. Theory*, vol. IT-23, no. 1, pp. 1–37, Jan. 1977.

[13] G. Kramer, "Capacity results for the discrete memoryless network," *IEEE Trans. Inf. Theory*, vol. 49, no. 1, pp. 4–21, Jan. 2003.

[14] I. Csiszár and J. Körner, "Toward a general theory of source networks," *IEEE Trans. Inform. Theory*, vol. IT-26, no. 2, pp. 155–165, Mar. 1980.

[15] M. Gastpar, "On source-channel communication in networks," in *Proc. DIMACS Workshop on Network Information Theory*, Piscataway, NJ, Mar. 2003.

[16] A. El Gamal, M. Mohseni, and S. Zahedi, "On reliable communication over additive white Gaussian noise relay channels," *IEEE Trans. Inf. Theory*, submitted for publication.

[17] M. A. Khojastepour, A. Sabharwal, and B. Aazhang, "On the capacity of 'cheap' relay networks," in *Proc. 37th Annu. Conf. Information Sciences and Systems*, Baltimore, MD, Mar. 2003.

[18] ——, "On capacity of Gaussian 'cheap' relay channel," in *Proc. 2003 IEEE Global Telecommunications Conf.*, San Francisco, CA, Dec. 2003, pp. 1776–1780.

[19] A. Høst-Madsen, "On the capacity of wireless relaying," in *Proc. 56th IEEE Vehicular Technology Conf.*, Vancouver, BC, Canada, Sep. 2002, pp. 1333–1337.

[20] G. Kramer, "Models and theory for relay channels with receive constraints," in *Proc. 42nd Annu. Allerton Conf. Communication, Control, and Computing*, Monticello, IL, Sep./Oct. 2004.

[21] A. Høst-Madsen and J. Zhang, "Capacity bounds and power allocation for the wireless relay channel," *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 2020–2040, Jun. 2005.

[22] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Trans. Inf. Theory*, vol. 50, no. 12, pp. 3062–3080, Dec. 2004.

[23] S. Verdú, "Spectral efficiency in the wideband regime," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1319–1343, Jun. 2002.

[24] S. Shamai (Shitz) and S. Verdú, "The impact of frequency-flat fading on the spectral efficiency of CDMA," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1302–1327, May 2001.

[25] R. Dabora and S. Servetto, "Broadcast channels with cooperating receivers: A downlink for the sensor reachback problem," in *Proc. 2004 IEEE Int. Symp. Information Theory*, Chicago, IL, Jun./Jul. 2004, p. 176.

[26] E. van der Meulen, "Transmission of Information in a T-Terminal Discrete Memoryless Channel," Ph.D. dissertation, Univ. California, Berkeley, CA, Jun. 1968.

[27] ——, "Three-terminal communication channels," *Adv. Appl. Probab.*, vol. 3, pp. 120–154, 1971.

[28] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inf. Theory*, vol. IT-19, no. 4, pp. 471–480, Jul. 1973.

[29] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991, p. 410.

[30] S. Shamai (Shitz) and S. Verdú, "Capacity of channels with side information," *Europ. Trans. Telecommun.*, vol. 6, no. 5, pp. 587–600, Sep.–Oct. 1995.

[31] J. Nayak, E. Tuncel, and K. Rose, "Zero-error source-channel coding with side-information," *IEEE Trans. Inf. Theory*, submitted for publication.

[32] T. Han and M. Costa, "Broadcast channels with arbitrarily correlated sources," *IEEE Trans. Inform. Theory*, vol. 33, pp. 641–650, Sept. 1987.

[33] A. D. Wyner, J. K. Wolf, and F. M. J. Willems, "Communicating via a processing broadcast satellite," *IEEE Trans. Inf. Theory*, vol. 48, no. 6, pp. 1243–1249, Jun. 2002.

[34] J. Körner and K. Marton, "General broadcast channels with degraded message sets," *IEEE Trans. Inf. Theory*, vol. IT-23, no. 1, pp. 60–64, Jan. 1977.

[35] R. W. Yeung, *A First Course in Information Theory*. New York: Kluwer, 2002, ch. 13.

[36] A. D. Wyner, "The rate-distortion function for source coding with side information at the decoder II: General sources," *Inf. Contr.*, vol. 38, pp. 60–80, July 1978.

[37] Y. Oohama, "Gaussian multiterminal source coding," *IEEE Trans. Inf. Theory*, vol. 43, no. 6, pp. 1912–1923, Nov. 1997.