KTH Electrical Engineering

Capacity Bounds for Backhaul-Supported Wireless Multicast Relay Networks with Cross-Links

in Proc. of IEEE ICC, Jun. 2011.

JINFENG DU, MING XIAO, MIKAEL SKOGLUND

Stockholm June 2011

School of Electrical Engineering and the ACCESS Linnaeus Center,
Royal Institute of Technology (KTH), SE-100 44 Stockholm, Sweden

IR-EE-KT 2011:004

Capacity Bounds for Backhaul-Supported Wireless Multicast Relay Networks with Cross-Links

Jinfeng Du, Ming Xiao, and Mikael Skoglund
ACCESS Linnaeus Center, Royal Institute of Technology, Stockholm, Sweden
Email: jinfeng@kth.se, ming.xiao@ee.kth.se, mikael.skoglund@ee.kth.se

***Abstract*—We investigate the capacity bounds for a wireless multicast relay network where two sources simultaneously multicast to two destinations through Gaussian channels with the help of a full-duplex relay node. All the individual channel gains are assumed to be time-invariant and known to every nodes in the network. The transmissions from two sources and from the relay use the same channel resource (i.e. co-channel transmission) and the two source nodes are connected with an orthogonal error-free backhaul. This multicast relay network is generic in the sense that it can be extended to more general networks by tuning the channel gains within the range $[0, \infty)$. By extending the proof of the converse developed by Cover and El Gamal for the Gaussian relay channel, we characterize the cut-set bound for this multicast relay network. We also present a lower bound by using decoding-and-forward relaying combined with network beam-forming.**

I. Introduction

As shown in Fig. 1, we consider a relay-aided two-source two-destination multiple multicast network where two source nodes $\mathcal{S}_1$ and $\mathcal{S}_2$ multicast their individual message W_1 at rate R_1 and W_2 at rate R_2, respectively, to both destinations $\mathcal{D}_1$ and $\mathcal{D}_2$, with the help of a relay $\mathcal{R}$. Two source nodes can cooperate with each other through an orthogonal error-free backhaul with capacity C_b. The relay forwards the information it receives in previous time slot to both destinations. The transmissions from $\mathcal{S}_1$, $\mathcal{S}_2$, and $\mathcal{R}$ use the same channel resource (i.e. co-channel transmission) and mix up at all the receiving terminals with independent Gaussian noise. The individual channel gains $g_{ij}\geq 0, i, j = 1, 2, r$ are assumed to be time-invariant and known to every nodes in this network. The model in Fig. 1 is generic since it covers a class of networks by tuning the channel gains g_{ij} within the range $[0, \infty)$. Such a system is interesting since it arises, for example, in a wireless cellular downlink where two base stations multicast multimedia information to two mobile terminals, one in each cell, with the help of a dedicated relay deployed at the common cell boundary. Since the base stations are connected through the (fiber or microwave) backhaul, various cooperative strategies can be used to boost the throughput.

The full understanding of such systems, even for the original three-node relay network, is not ready yet. Since 1970s, numerous research efforts have been casted on the relay networks. Capacity bounds for three-node relaying networks (source-relay-sink, or two cooperative sources and one sink) have been studied in [1], [2], where the capacity regions for degraded or reversely degraded scenarios are characterized. Coding schemes have been investigated for multiple-access relay channels (MARC) [3], [4] involving multiple sources and a single destination, and for broadcast relay channels (BRC) [4], [5] where a single source transmits to multiple destinations. Recent results on capacity bounds for multiple-source multiple-destination relay networks, [6]–[9] and references therein, have provided valuable insights. Apart from introducing a dedicated relay, the rate region can also be enlarged by cooperation among conferencing/cognitive encoders [10]–[12] and/or among conferencing decoders [12] .

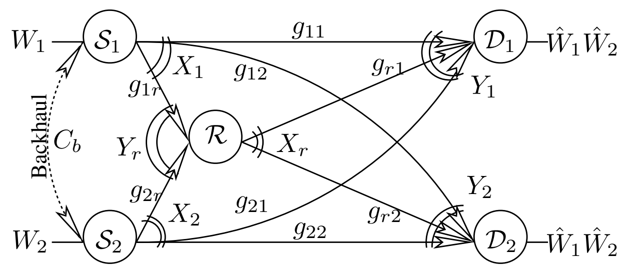

Fig. 1. Two source nodes $\mathcal{S}_1$ and $\mathcal{S}_2$, connected with backhaul, multicast information W_1 at rate R_1 and W_2 at rate R_2 respectively to both destinations $\mathcal{D}_1$ and $\mathcal{D}_2$ through Gaussian channels, with aid from a full-duplex relay $\mathcal{R}$.

In this paper, we aim at characterizing the capacity upper and lower bounds for the system shown in Fig. 1, where relaying is combined with source cooperation. To simplify our analysis, we will restrict our analysis with a full-duplex relay and an error-free backhaul with sufficiently high capacity (larger than the sum-rate, i.e. $C_b{>}R_1{+}R_2$). Extensions to a half-duplex relay or low-rate backhaul ($C_b{<}R_1{+}R_2$) will be left to future work. Note that, the high-rate backhaul makes the capacity region (R_1, R_2) only bounded by $R_1 + R_2$, which makes our system closely related to the MIMO relay channels [13], [14], where $\mathcal{S}_1$ and $\mathcal{S}_2$ can be grouped together as a virtual transmitter equipped with two antennas. As discussed in our previous work [15] for the case without cross-link (i.e. $g_{12}{=}g_{21}{=}0$), the capacity upper bounds derived based on the MIMO relay channel [13], [14] are in general larger than the cut-set bound: (i) the equality in (4) of [13] is achieved only if when the relay has at least the same number of transmit antennas compared with the virtual transmitter, which is not the case here; (ii) the bound (9) of [14] is derived based on the sum-power constraint, while in our case only per-antenna/node power constraint is applied. Moreover, for the cases of low-rate backhaul, the MIMO upper bounds developed in [13], [14] seem irrelevant, but our proof developed here can still be helpful by combining with the strategies in [10]–[12].

In our previous work [16], we have successfully characterized the exact cut-set bound for the case without cross-link by extending the proof of the converse developed by Cover and El Gamal for the Gaussian relay channel, and verified that this cut-set bound is tight (coincide with a lower bound) when the transmitting power is lower than a sphere defined by the individual channel gains. We extend this method to the system in Fig. 1 and present capacity upper and lower bounds.

The rest of this paper is organized as follows. The system model is introduced in Sec. II. Our main results are summarized in Sec. III with the detailed proof of the cut-set bound presented in Sec. IV. Numerical results are presented in Sec. V with the concluding remarks presented in Sec. VI.

Notations: Capital letter X indicates a real valued random variable and $p(X)$ indicates its probability density/mass function. $X^{(n)}$ denotes a vector of random variables of length n and $I(X;Y)$ denotes the mutual information between X and Y. $C(x)=\frac{1}{2}\log_2(1+x)$ is the Gaussian capacity function.

II. System model

Given an average transmit power constraint P, fixed channel gain g, and noise power σ^2, the signal-to-noise ratio (SNR) of that link can be written as $\gamma = g^2P/\sigma^2$. We can therefore characterize the transmission links only by their individual SNR γ, without distinguishing the SNR contribution from the transmit power or the channel gain. Therefore we can model the system shown in Fig. 1 as follows

$$\begin{aligned} Y_1^{(n)} &= \sqrt{\gamma_{11}}X_1^{(n)} + \sqrt{\gamma_{21}}X_2^{(n)} + \sqrt{\gamma_{r1}}X_r^{(n)} + Z_1^{(n)}, \\ Y_2^{(n)} &= \sqrt{\gamma_{21}}X_1^{(n)} + \sqrt{\gamma_{22}}X_2^{(n)} + \sqrt{\gamma_{r2}}X_r^{(n)} + Z_2^{(n)}, \\ Y_r^{(n)} &= \sqrt{\gamma_{1r}}X_1^{(n)} + \sqrt{\gamma_{2r}}X_2^{(n)} + Z_r^{(n)}, \end{aligned} \quad (1)$$

where $\gamma_{ij} \geq 0, i,j=1,2,r$ are the effective SNR for the corresponding links. $X_i^{(n)}$, $Y_i^{(n)}$, $Z_i^{(n)}$, $i=1,2,r$ are n-dimensional transmitted signals, received signals, and noise at nodes $\mathcal{S}_1$, $\mathcal{S}_2$ and $\mathcal{R}$, respectively. An average unit-power constraints are applied to all the transmitted signals, i.e.,

$$\frac{1}{n}\sum_{k=1}^{n} X_i^2[k] \leq 1. \quad (2)$$

The noise components $Z_i[k]$, $i=1,2,r$ and $k=1,...,n$ are i.i.d. zero-mean unit-variance Gaussian random variables.

III. Main Results

Theorem 1: Define a rate region as

$$\begin{aligned} &R_1+R_2 \leq C_0 = \sup_{0\leq\alpha_1,\alpha_2,\rho\leq 1} \min\{ \\ &C\left((\gamma_{11}+\gamma_{1r})\alpha_1+(1-\rho^2)\alpha_1\alpha_2(\sqrt{\gamma_{11}\gamma_{2r}}-\sqrt{\gamma_{21}\gamma_{1r}})^2 \right.\\ &\quad \left.+(\gamma_{21}+\gamma_{2r})\alpha_2+2\rho\sqrt{\alpha_1\alpha_2}(\sqrt{\gamma_{11}\gamma_{21}}+\sqrt{\gamma_{1r}\gamma_{2r}})\right), \\ &C\left((\gamma_{12}+\gamma_{1r})\alpha_1+(1-\rho^2)\alpha_1\alpha_2(\sqrt{\gamma_{12}\gamma_{2r}}-\sqrt{\gamma_{22}\gamma_{1r}})^2 \right.\\ &\quad \left.+(\gamma_{22}+\gamma_{2r})\alpha_2+2\rho\sqrt{\alpha_1\alpha_2}(\sqrt{\gamma_{12}\gamma_{22}}+\sqrt{\gamma_{1r}\gamma_{2r}})\right), \\ &C\left(\gamma_{11}+\gamma_{21}+\gamma_{r1}+2\sqrt{\bar\alpha_1\gamma_{11}\gamma_{r1}}+2\sqrt{\bar\alpha_2\gamma_{21}\gamma_{r1}}\right.\\ &\quad \left.+2(\rho\sqrt{\alpha_1\alpha_2}+\sqrt{\bar\alpha_1\bar\alpha_2})\sqrt{\gamma_{11}\gamma_{21}}\right), \\ &C\left(\gamma_{12}+\gamma_{22}+\gamma_{r2}+2\sqrt{\bar\alpha_1\gamma_{12}\gamma_{r2}}+2\sqrt{\bar\alpha_2\gamma_{22}\gamma_{r2}}\right.\\ &\quad \left.+2(\rho\sqrt{\alpha_1\alpha_2}+\sqrt{\bar\alpha_1\bar\alpha_2})\sqrt{\gamma_{12}\gamma_{22}}\right)\}, \end{aligned} \quad (3)$$

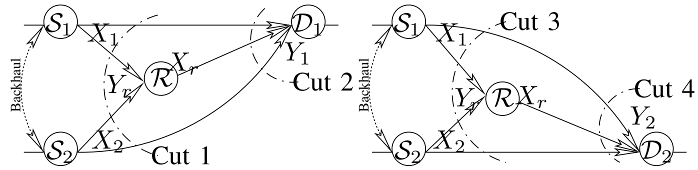

Fig. 2. The sum multicast capacity is bounded by the cut-set bound based on the four cuts shown in the figure.

where $\bar\alpha_1=1-\alpha_1$ and $\bar\alpha_2=1-\alpha_2$. For the system in Fig. 1 with power constraints (2), its cut-set bound $C_{\text{cut-set}}$ equals to C_0.

Proof: The proof can be found in Section IV. ■

We also present a lower bound given by the cooperative strategy network beam-forming (NBF) [16], which essentially first generates the network coded message W_b at source nodes, and then cooperates its transmission at $\mathcal{S}_1$ and $\mathcal{S}_2$ with the relaying signal in a beam-forming fashion to take advantage of the coherent combining gain. Since $\mathcal{S}_1$ and $\mathcal{S}_2$ transmit the same network coded message W_b, the achievable sum-rate can be split arbitrarily between them.

Proposition 1: The capacity region is lower bounded by the achievable rate region of NBF, which is the union of all rate pairs (R_1,R_2) which satisfy $R_1\geq 0$, $R_2 \geq 0$, and

$$\begin{aligned} &R_1+R_2 < \min\left\{C\left((\sqrt{\alpha_1\gamma_{1r}}+\sqrt{\alpha_2\gamma_{2r}})^2\right),\right. \\ &C\left((\sqrt{\alpha_1\gamma_{11}}+\sqrt{\alpha_2\gamma_{21}})^2+(\sqrt{\bar\alpha_1\gamma_{11}}+\sqrt{\bar\alpha_2\gamma_{21}}+\sqrt{\gamma_{r1}})^2\right), \\ &\left.C((\sqrt{\alpha_1\gamma_{12}}+\sqrt{\alpha_2\gamma_{22}})^2+(\sqrt{\bar\alpha_1\gamma_{12}}+\sqrt{\bar\alpha_2\gamma_{22}}+\sqrt{\gamma_{r2}})^2)\right\}, \end{aligned} \quad (4)$$

with the union taken over all $0\leq\alpha_1,\alpha_2\leq 1$.

Proof: This result is a straightforward extension of *Proposition 4* in [16] by replacing the system model with (1) and formulating the transmitted signals as

$$\begin{aligned} X_{r,b}^{(n)} &= U^{(n)}(W_{b-1}), \\ X_{1,b}^{(n)} &= \sqrt{\alpha_1}V^{(n)}(W_b,W_{b-1})+\sqrt{\bar\alpha_1}U^{(n)}(W_{b-1}), \\ X_{2,b}^{(n)} &= \sqrt{\alpha_2}V^{(n)}(W_b,W_{b-1})+\sqrt{\bar\alpha_2}U^{(n)}(W_{b-1}), \end{aligned} \quad (5)$$

where $0\leq\alpha_1,\alpha_2\leq 1$ are power allocation parameters, $U^{(n)}$, $V^{(n)}$ are random codewords generated in the same way as in [16]. Detailed description on encoding/decoding is skipped due to space limitation. The constraints in (4) correspond to successful decoding at $\mathcal{R}$, $\mathcal{D}_1$, and $\mathcal{D}_2$, respectively. ■

IV. Proof of the Cut-Set Bound

By the maximum-flow min-cut theorem [17], the maximum achievable sum-rate from the source nodes to any of the destinations can be no larger than the minimum of the mutual information flows across all possible cuts, maximized over a joint distribution for the transmitted signals. The cut-set bound between the two sources and each of the sink for the network in Fig. 1 can be derived based on four cuts shown in Fig. 2 as follows (the dimension super script $^{(n)}$ is suppressed hereafter to simplify the notation)

$$\begin{aligned} &R_1+R_2 \leq C_{\text{cut-set}} = \sup_{p(X_1,X_2,X_r)} \min\{ \\ &\quad \frac{1}{n}I(X_1,X_2;Y_1,Y_r|X_r), \frac{1}{n}I(X_1,X_2,X_r;Y_1), \\ &\frac{1}{n}I(X_1,X_2;Y_2,Y_r|X_r), \frac{1}{n}I(X_1,X_2,X_r;Y_2)\}+\epsilon_n, \end{aligned} \quad (6)$$

where X_1, X_2 and X_r are potentially correlated, and $\epsilon_n\to 0$ as $n\to\infty$. We will first find an upper bound $C_{\text{upp}} \geq C_{\text{cut-set}}$ by using the new technique, and then find a lower bound $C_{\text{cut-set, G}} \leq C_{\text{cut-set}}$ by restricting the source distribution to Gaussian. Finally by showing that $C_{\text{cut-set, G}} = C_{\text{upp}}$ we can find the exact cut-set bound $C_{\text{cut-set}} = C_{\text{upp}} = C_{\text{cut-set, G}}$.

A. The upper bound C_{upp}

Following the conventional notation for the differential entropy $h(X)$ of a continuous valued random variable X, the mutual information corresponding to cut 2 can be written as

$$\begin{aligned} I(X_1, X_2, X_r; Y_1) &= h(Y_1) - h(Y_1|X_1, X_2, X_r) \\ &= h(Y_1) - h(Z_1) = h(Y_1) - \frac{n}{2}\log(2\pi e). \end{aligned} \tag{7}$$

From the maximum entropy lemma [17], we get

$$h(Y_1) \leq \sum_{i=1}^{n} h(Y_{1,i}) \leq \sum_{i=1}^{n} \frac{1}{2}\log(2\pi e \text{Var}[Y_{1,i}]), \tag{8}$$

where the second equality is achieved when $Y_{1,i}$ is Gaussian distributed. Hence

$$\begin{aligned} \frac{1}{n} I(X_1, X_2, X_r; Y_1) &\leq \frac{1}{n}\sum_{i=1}^{n}\frac{1}{2}\log(\text{Var}[Y_{1,i}]) \\ &\leq \frac{1}{2}\log(\frac{1}{n}\sum_{i=1}^{n}\text{Var}[Y_{1,i}]), \end{aligned} \tag{9}$$

where the last steps follow from Jensen's inequality, with

$$\text{Var}[Y_{1,i}] = 1 + \text{Var}[\sqrt{\gamma_{11}}X_{1,i} + \sqrt{\gamma_{21}}X_{2,i} + \sqrt{\gamma_{r1}}X_{r,i}]. \tag{10}$$

According to the *law of total variance*, for two random variables X and Y on the same probability space, and the variance of X is finite, then

$$\text{Var}[X] = E(\text{Var}[X|Y]) + \text{Var}[E(X|Y)].$$

We can therefore rewrite (10) as

$$\begin{aligned} \text{Var}[Y_{1,i}] &= 1 + \text{Var}[E(\sqrt{\gamma_{11}}X_{1,i} + \sqrt{\gamma_{21}}X_{2,i} + \sqrt{\gamma_{r1}}X_{r,i}|X_{r,i})] \\ &\quad + E(\text{Var}[\sqrt{\gamma_{11}}X_{1,i} + \sqrt{\gamma_{21}}X_{2,i} + \sqrt{\gamma_{r1}}X_{r,i}|X_{r,i}]) \\ &= 1 + E(\text{Var}[\sqrt{\gamma_{11}}X_{1,i} + \sqrt{\gamma_{21}}X_{2,i}|X_{r,i}]) \\ &\quad + \text{Var}[\sqrt{\gamma_{11}}E(X_{1,i}|X_{r,i}) + \sqrt{\gamma_{21}}E(X_{2,i}|X_{r,i}) + \sqrt{\gamma_{r1}}X_{r,i}], \end{aligned} \tag{11}$$

where

$$\begin{aligned} E(\text{Var}[\sqrt{\gamma_{11}}X_{1,i} + \sqrt{\gamma_{21}}X_{2,i}|X_{r,i}]) &= E(\gamma_{11}\text{Var}[X_{1,i}|X_{r,i}] \\ &\quad + \gamma_{21}\text{Var}[X_{2,i}|X_{r,i}] + 2\sqrt{\gamma_{11}\gamma_{21}}\text{Cov}(X_{1,i}, X_{2,i}|X_{r,i})) \\ &= \gamma_{11}E(\text{Var}[X_{1,i}|X_{r,i}]) + \gamma_{21}E(\text{Var}[X_{2,i}|X_{r,i}]) \\ &\quad + 2\sqrt{\gamma_{11}\gamma_{21}}E(\text{Cov}(X_{1,i}, X_{2,i}|X_{r,i})), \end{aligned} \tag{12}$$

and

$$\begin{aligned} &\text{Var}[\sqrt{\gamma_{11}}E(X_{1,i}|X_{r,i}) + \sqrt{\gamma_{21}}E(X_{2,i}|X_{r,i}) + \sqrt{\gamma_{r1}}X_{r,i}] \\ &\leq E[(\sqrt{\gamma_{11}}E[X_{1,i}|X_{r,i}] + \sqrt{\gamma_{21}}E[X_{2,i}|X_{r,i}] + \sqrt{\gamma_{r1}}X_{r,i})^2] \\ &= \gamma_{11}E(E^2[X_{1,i}|X_{r,i}]) + 2\sqrt{\gamma_{11}\gamma_{r1}}E(X_{r,i}E[X_{1,i}|X_{r,i}]) \\ &\quad + \gamma_{21}E(E^2[X_{2,i}|X_{r,i}]) + 2\sqrt{\gamma_{21}\gamma_{r1}}E(X_{r,i}E[X_{2,i}|X_{r,i}]) \\ &\quad + \gamma_{r1}E(X_{r,i}^2) + 2\sqrt{\gamma_{11}\gamma_{21}}E(E[X_{1,i}|X_{r,i}]E[X_{2,i}|X_{r,i}]). \end{aligned} \tag{13}$$

As in [1], define

$$\bar{\alpha}_1 = \frac{1}{n}\sum_{i=1}^{n} E[E^2(X_{1,i}|X_{r,i})], \ \alpha_1 \in [0, 1], \tag{14}$$

then we have

$$\begin{aligned} \frac{1}{n}\sum_{i=1}^{n} E[\text{Var}(X_{1,i}|X_{r,i})] &= \frac{1}{n}\sum_{i=1}^{n}(E[X_{1,i}^2] - E[E^2(X_{1,i}|X_{r,i})]) \\ &= \frac{1}{n}\sum_{i=1}^{n} E[X_{1,i}^2] - \frac{1}{n}\sum_{i=1}^{n} E[E^2(X_{1,i}|X_{r,i})] \leq \alpha_1, \end{aligned} \tag{15}$$

where the inequality comes from (2). Similarly we have

$$\bar{\alpha}_2 = \frac{1}{n}\sum_{i=1}^{n} E[E^2(X_{2,i}|X_{r,i})], \frac{1}{n}\sum_{i=1}^{n} E[\text{Var}(X_{2,i}|X_{r,i})] \leq \alpha_2, \tag{16}$$

where $\alpha_2 \in [0, 1]$. On the other hand, as

$$\text{Cov}(X_{1,i}, X_{2,i}|X_{r,i}) = \phi_i\sqrt{\text{Var}(X_{1,i}|X_{r,i})\text{Var}(X_{2,i}|X_{r,i})},$$

where $|\phi_i| \leq 1$ is the correlation coefficient, we have

$$\begin{aligned} &\frac{1}{n}\sum_{i=1}^{n} E[\text{Cov}(X_{1,i}, X_{2,i}|X_{r,i})] \\ &\leq \sqrt{\frac{1}{n}\sum_{i=1}^{n}\phi_i E[\text{Var}(X_{1,i}|X_{r,i})]\frac{1}{n}\sum_{i=1}^{n}\phi_i E[\text{Var}(X_{2,i}|X_{r,i})]} \\ &\leq \sqrt{\alpha_1\alpha_2}, \end{aligned} \tag{17}$$

where the first inequality is due to the Cauchy–Schwarz inequality and the last step is given by (15) and (16). Given that $|\phi_i| \leq 1$, we can introduce an auxiliary variable $0 \leq \rho \leq 1$ such that

$$\frac{1}{n}\sum_{i=1}^{n} E[\text{Cov}(X_{1,i}, X_{2,i}|X_{r,i})] = \rho\sqrt{\alpha_1\alpha_2}. \tag{18}$$

Also, using the Cauchy–Schwarz inequality we get

$$\begin{aligned} &\frac{1}{n}\sum_{i=1}^{n} E(X_{r,i}E[X_{1,i}|X_{r,i}]) \\ &\leq \sqrt{\frac{1}{n}\sum_{i=1}^{n} E[X_{r,i}^2]\frac{1}{n}\sum_{i=1}^{n} E(E^2[X_{1,i}|X_{r,i}])} \leq \sqrt{\bar{\alpha}_1}, \end{aligned} \tag{19}$$

$$\frac{1}{n}\sum_{i=1}^{n} E(X_{r,i}E[X_{2,i}|X_{r,i}]) \leq \sqrt{\bar{\alpha}_2}, \tag{20}$$

$$\frac{1}{n}\sum_{i=1}^{n} E(E[X_{1,i}|X_{r,i}]E[X_{2,i}|X_{r,i}]) \leq \sqrt{\bar{\alpha}_1\bar{\alpha}_2}. \tag{21}$$

Now, substituting (11)–(21) into (9), and apply the same approach also to cut 4, we get

$$\begin{aligned} \frac{1}{n} I(X_1, X_2, X_r; Y_1) &\leq C(\gamma_{11} + \gamma_{21} + \gamma_{r1} + 2\sqrt{\bar{\alpha}_1\gamma_{11}\gamma_{r1}} \\ &\quad + 2\sqrt{\bar{\alpha}_2\gamma_{21}\gamma_{r1}} + 2(\rho\sqrt{\alpha_1\alpha_2} + \sqrt{\bar{\alpha}_1\bar{\alpha}_2})\sqrt{\gamma_{11}\gamma_{21}}), \end{aligned} \tag{22}$$

$$\begin{aligned} \frac{1}{n} I(X_1, X_2, X_r; Y_2) &\leq C(\gamma_{12} + \gamma_{22} + \gamma_{r2} + 2\sqrt{\bar{\alpha}_1\gamma_{12}\gamma_{r2}} \\ &\quad + 2\sqrt{\bar{\alpha}_2\gamma_{22}\gamma_{r2}} + 2(\rho\sqrt{\alpha_1\alpha_2} + \sqrt{\bar{\alpha}_1\bar{\alpha}_2})\sqrt{\gamma_{12}\gamma_{22}}). \end{aligned} \tag{23}$$

For cut 1 we have

$$
\begin{aligned}
&I(X_1,X_2;Y_1,Y_r|X_r) = h(Y_1,Y_r|X_r){-}h(Y_1,Y_r|X_1,X_2,X_r)\\
&{=}h(Y_1,Y_r|X_r){-}h(Y_1|X_1,X_2,X_r){-}h(Y_r|X_1,X_2,X_r) \quad (24)\\
&= h(Y_1,Y_r|X_r){-}h(Z_1){-}h(Z_r) = h(Y_1,Y_r|X_r){-}n\log(2\pi e)\\
&\leq \tfrac{1}{2}\textstyle\sum_{i=1}^{n}\log\left((2\pi e)^2|\boldsymbol{K}_i|\right) - n\log(2\pi e) = \tfrac{1}{2}\sum_{i=1}^{n}\log\left(|\boldsymbol{K}_i|\right),
\end{aligned}
$$

where the second equality in (24) comes from the fact that Y_1 and Y_r are independent given (X_1, X_2, X_r) and the inequality is due to the maximum entropy lemma [17], with equality achieved by joint Gaussian distributed $(Y_{1,i}, Y_{r,i})$ with conditional covariance matrices $\boldsymbol{K}_i$, which is defined by

$$
\boldsymbol{K}_i = \begin{bmatrix} E(\text{Var}[Y_{1,i}|X_{r,i}]) & E[\text{Cov}(Y_{1,i},Y_{r,i}|X_{r,i})] \\ E[\text{Cov}(Y_{1,i},Y_{r,i}|X_{r,i})] & E(\text{Var}[Y_{r,i}|X_{r,i}]) \end{bmatrix}.
$$

Obviously, the covariance matrices $\boldsymbol{K}_i$ are positive semi-definite. Since the function $\log|\boldsymbol{K}|$ is concave [18], we can thus bound the throughput of cut 1 from (24) as follows

$$
\begin{aligned}
&\frac{1}{n}I(X_1,X_2;Y_1,Y_r|X_r) \leq \frac{1}{2}\log\left(\left|\tfrac{1}{n}\textstyle\sum_{i=1}^{n}\boldsymbol{K}_i\right|\right)\\
&= \frac{1}{2}\log\Big(\tfrac{1}{n}\textstyle\sum_{i=1}^{n}E(\text{Var}[Y_{1,i}|X_{r,i}])\tfrac{1}{n}\sum_{j=1}^{n}E(\text{Var}[Y_{r,j}|X_{r,j}])\\
&\qquad - \left(\tfrac{1}{n}\textstyle\sum_{i=1}^{n}E[\text{Cov}(Y_{1,i},Y_{r,i}|X_{r,i})]\right)^2\Big). \qquad (25)
\end{aligned}
$$

Furthermore, since

$$
\begin{aligned}
&E(\text{Var}[Y_{1,i}|X_{r,i}]) = 1 + E(\text{Var}[\sqrt{\gamma_{11}}X_{1,i}{+}\sqrt{\gamma_{21}}X_{2,i}|X_{r,i}]),\\
&E(\text{Var}[Y_{r,i}|X_{r,i}]) = 1 + E(\text{Var}[\sqrt{\gamma_{1r}}X_{1,i}{+}\sqrt{\gamma_{2r}}X_{2,i}|X_{r,i}]),\\
&E[\text{Cov}(Y_{1,i},Y_{r,i}|X_{r,i})] =\\
&\quad\sqrt{\gamma_{11}\gamma_{1r}}E(\text{Var}[X_{1,i}|X_{r,i}]) + \sqrt{\gamma_{21}\gamma_{2r}}E(\text{Var}[X_{2,i}|X_{r,i}])\\
&\qquad +(\sqrt{\gamma_{11}\gamma_{2r}}+\sqrt{\gamma_{21}\gamma_{1r}})E[\text{Cov}(X_{1,i},X_{2,i}|X_{r,i})],
\end{aligned}
$$

by combining with (12) and (15)–(18), we can conclude that

$$
\begin{aligned}
&\frac{1}{n}I(X_1,X_2;Y_1,Y_r|X_r) \leq\\
&C\left((\gamma_{11}{+}\gamma_{1r})\alpha_1{+}(1{-}\rho^2)\alpha_1\alpha_2(\sqrt{\gamma_{11}\gamma_{2r}}{-}\sqrt{\gamma_{21}\gamma_{1r}})^2\right.\\
&\quad\left.+(\gamma_{21}{+}\gamma_{2r})\alpha_2 + 2\rho\sqrt{\alpha_1\alpha_2}(\sqrt{\gamma_{11}\gamma_{21}}{+}\sqrt{\gamma_{1r}\gamma_{2r}})\right). \quad (26)
\end{aligned}
$$

Similarly, we can bound the throughput of cut 3 as follows

$$
\begin{aligned}
&\frac{1}{n}I(X_1,X_2;Y_2,Y_r|X_r) \leq\\
&C\left((\gamma_{12}{+}\gamma_{1r})\alpha_1{+}(1{-}\rho^2)\alpha_1\alpha_2(\sqrt{\gamma_{12}\gamma_{2r}}{-}\sqrt{\gamma_{22}\gamma_{1r}})^2\right.\\
&\quad\left.+(\gamma_{22}{+}\gamma_{2r})\alpha_2 + 2\rho\sqrt{\alpha_1\alpha_2}(\sqrt{\gamma_{12}\gamma_{22}}{+}\sqrt{\gamma_{1r}\gamma_{2r}})\right). \quad (27)
\end{aligned}
$$

By substituting (22) (23) (26) (27) into (6) and $n{\to}\infty$, comparing the resulting region to (3) we can conclude that

$$
C_{\text{cut-set}} \leq C_{\text{upp}} = C_0.
$$

B. The lower bound $C_{\text{cut-set, G}}$

By restricting $p(X_1, X_2, X_r)$ in (6) to be Gaussian, we can partition X_1, X_2 and X_r as follows

$$X_r = U, \qquad (28a)$$

$$X_1 = \sqrt{(1-\rho)\alpha_1}S_1 + \sqrt{\rho\alpha_1}V + \sqrt{(1-\alpha_1)}U, \qquad (28b)$$

$$X_2 = \sqrt{(1-\rho)\alpha_2}S_2 + \sqrt{\rho\alpha_2}V + \sqrt{(1-\alpha_2)}U, \qquad (28c)$$

where S_1, S_2, V, U are n-dimensional independent Gaussian random vectors with zero-mean and unit-variance. Auxiliary variables $0{\leq}\alpha_1, \alpha_2, \rho{\leq}1$ are introduced to represent the potential correlation among X_1, X_2 and X_r due to cooperation. by substituting (28) into (1), we can derive from (6) that

$$
\begin{aligned}
&C_{\text{cut-set}} \geq C_{\text{cut-set, G}} = \sup_{0\leq\alpha_1,\alpha_2,\rho\leq 1}\min\frac{1}{2n}\sum_{i=1}^{n}\{ \qquad (29)\\
&\log(\text{Var}[Y_{1,i}]), \log(\text{Var}[Y_{2,i}]), \log\left(|\boldsymbol{K}_{1,i}|\right), \log\left(|\boldsymbol{K}_{2,i}|\right)\}{+}\epsilon_n,
\end{aligned}
$$

where $\epsilon_n{\to}0$ as $n{\to}\infty$, and for $i = 1, ..., n$ we have

$$
\begin{aligned}
\text{Var}[Y_{1,i}] &= 1{+}\bar{\rho}\alpha_1\gamma_{11}{+}\bar{\rho}\alpha_2\gamma_{21}{+}\rho(\sqrt{\alpha_1\gamma_{11}}{+}\sqrt{\alpha_2\gamma_{21}})^2\\
&\quad+(\sqrt{\bar{\alpha}_1\gamma_{11}} + \sqrt{\bar{\alpha}_2\gamma_{21}} + \sqrt{\gamma_{r1}})^2,\\
\text{Var}[Y_{2,i}] &= 1{+}\bar{\rho}\alpha_1\gamma_{12}{+}\bar{\rho}\alpha_2\gamma_{22}{+}\rho(\sqrt{\alpha_1\gamma_{12}}{+}\sqrt{\alpha_2\gamma_{22}})^2\\
&\quad+(\sqrt{\bar{\alpha}_1\gamma_{12}} + \sqrt{\bar{\alpha}_2\gamma_{22}} + \sqrt{\gamma_{r2}})^2, \qquad (30)\\
|\boldsymbol{K}_{1,i}| &= (1{+}\bar{\rho}\alpha_1\gamma_{11}{+}\bar{\rho}\alpha_2\gamma_{21}{+}\rho(\sqrt{\alpha_1\gamma_{11}}{+}\sqrt{\alpha_2\gamma_{21}})^2)\\
&\quad*(1{+}\bar{\rho}\alpha_1\gamma_{1r}{+}\bar{\rho}\alpha_2\gamma_{2r}{+}\rho(\sqrt{\alpha_1\gamma_{1r}}{+}\sqrt{\alpha_2\gamma_{2r}})^2)\\
&-(\sqrt{\gamma_{11}\gamma_{1r}}\alpha_1{+}\sqrt{\gamma_{21}\gamma_{2r}}\alpha_2{+}(\sqrt{\gamma_{11}\gamma_{2r}}{+}\sqrt{\gamma_{21}\gamma_{1r}})\rho\sqrt{\alpha_1\alpha_2})^2,\\
|\boldsymbol{K}_{2,i}| &= (1{+}\bar{\rho}\alpha_1\gamma_{12}{+}\bar{\rho}\alpha_2\gamma_{22}{+}\rho(\sqrt{\alpha_1\gamma_{12}}{+}\sqrt{\alpha_2\gamma_{22}})^2)\\
&\quad*(1{+}\bar{\rho}\alpha_1\gamma_{1r}{+}\bar{\rho}\alpha_2\gamma_{2r}{+}\rho(\sqrt{\alpha_1\gamma_{1r}}{+}\sqrt{\alpha_2\gamma_{2r}})^2)\\
&-(\sqrt{\gamma_{12}\gamma_{1r}}\alpha_1{+}\sqrt{\gamma_{22}\gamma_{2r}}\alpha_2{+}(\sqrt{\gamma_{12}\gamma_{2r}}{+}\sqrt{\gamma_{22}\gamma_{1r}})\rho\sqrt{\alpha_1\alpha_2})^2.
\end{aligned}
$$

By substituting (30) into (29) and letting $n{\to}\infty$, comparing the resulting region to (3) we can conclude that

$$
C_{\text{cut-set}} \geq C_{\text{cut-set, G}} = C_0.
$$

Recall that $C_{\text{cut-set, G}} \leq C_{\text{cut-set}} \leq C_{\text{upp}}$, we can finally conclude that $C_{\text{cut-set}} = C_0$, i.e., Theorem 1 holds.

V. Numerical Results

We present the numerical results of the cut-set bound and the NBF lower bound on the capacity regions with different link quality. Unless stated otherwise, the following heuristic parameters will be used: The $\mathcal{S}_1$–$\mathcal{D}_1$ link SNR γ_{11}=5dB, the source-destination link SNR γ_{22}=10dB, the source-relay link SNR γ_{1r}=γ_{2r}=10dB, the relay-destination link SNR γ_{r1}=γ_{r2}=10dB, and the cross-link SNR γ_{12}=γ_{21}=0dB.

In Fig. 3 we investigate the impact of the relay-destination link quality on the capacity region. When the $\mathcal{R}$–$\mathcal{D}_1$ link is not strong, the $\mathcal{R}$–$\mathcal{D}_2$ link SNR γ_{r2} is not a limiting factor and therefore the capacity will monotonically increase with γ_{r1} until it is large enough to reach the bottleneck set by γ_{r2}, as demonstrated in Fig. 3. The gap between the cut-set bound the NBF lower bound is always less than 0.1 bits/channel use.

In Fig. 4 we fix the $\mathcal{S}_2$–$\mathcal{R}$ link SNR γ_{2r} and vary the SNR of the $\mathcal{S}_1$–$\mathcal{R}$ link. When the $\mathcal{S}_2$–$\mathcal{R}$ link is poor (γ_{2r} = 0dB), the $\mathcal{S}_1$–$\mathcal{R}$ link is the only reliable channel between $\mathcal{S}_1$, $\mathcal{S}_2$ and $\mathcal{R}$ and therefore its quality has a large impact on the capacity. When the $\mathcal{S}_2$–$\mathcal{R}$ link is strong, however, $\mathcal{S}_1$ can communicate with $\mathcal{R}$ reliably via $\mathcal{S}_2$ thanks to the backhual, even for a poor $\mathcal{S}_1$–$\mathcal{R}$ link. The impact of γ_{2r} will diminish when γ_{1r} becomes very large, as shown in the zoom-in plot. Note that the decoding requirement at relay for NBF introduces a large gap when both γ_{1r} and γ_{2r} are small.

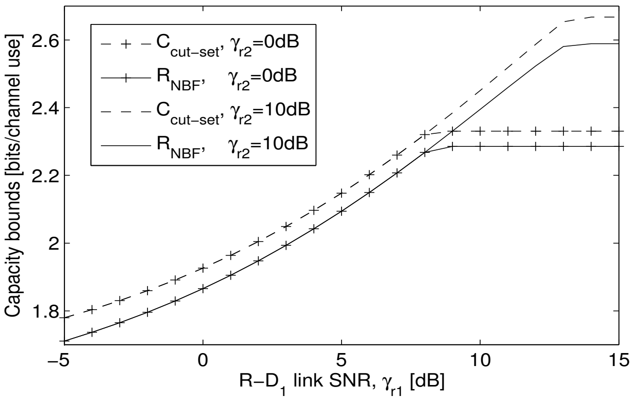

Fig. 3. Capacity bounds for varying $\mathcal{R} - \mathcal{D}_1$ link SNR γ_{r1}.

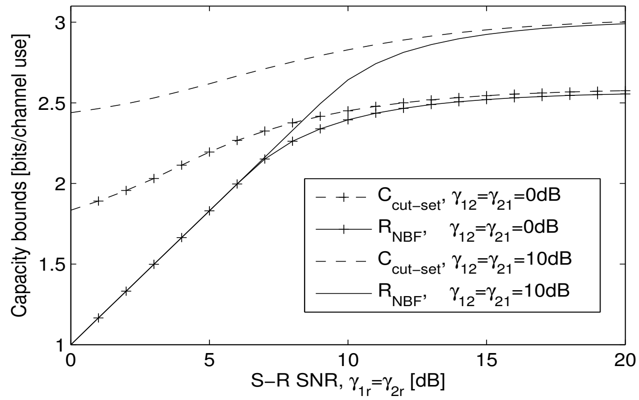

Fig. 5. Capacity bounds for varying symmetric source-relay SNR $\gamma_{1r} = \gamma_{2r}$.

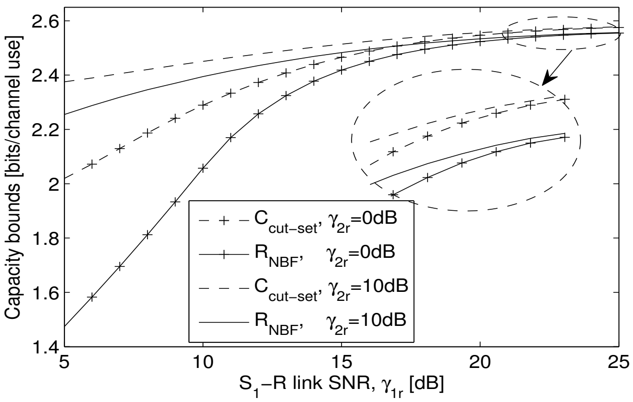

Fig. 4. Capacity bounds for varying $\mathcal{S}_1 - \mathcal{R}$ link SNR γ_{1r}.

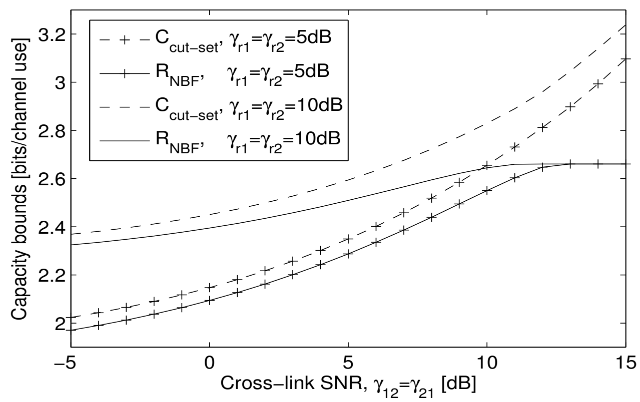

Fig. 6. Capacity bounds for varying symmetric cross-link SNR $\gamma_{12}{=}\gamma_{21}$.

In Fig. 5 we show the impact of the source-relay link for weak and strong cross-link quality. Obviously the NBF strategy does not utilize the benefit from the cross-link when the source-relay connection is not good enough. When the source-relay link is good, as shown in Fig. 6, the achievable rate of NBF increases with improved cross-link quality until the decoding at the relay becomes the bottleneck. The cut-set bound has no restriction on decoding at $\mathcal{R}$ and therefore will increase with the improved cross-link quality.

VI. Conclusion

We have derived the cut-set bound for backhaul-supported wireless multicast relay networks with cross-link and provided achievable rate based on network beam-forming strategy. The gap between the cut-set bound and the achievable rate is not large in general, as illustrated in the numerical results. Since the NBF strategy cannot take full advantage of the cross-link, a better scheme has to be sought in further work.

Acknowledgments

This work is funded by VR and VINNOVA.

References

[1] T. M. Cover and A. El Gamal, "Capacity theorems for the relay channel," *IEEE Trans. Inf. Theory*, vol. 25, pp. 572–584, Sep. 1979.

[2] A. Høst-Madsen and J. Zhang, "Capacity bounds and power allocation for wireless relay channels," *IEEE Trans. Inf. Theory*, vol. 51, pp. 2020–2040, Jun. 2005.

[3] G. Kramer and A. J. van Wijngaarden, "On the white Gaussian multiple-access relay channel", in *Proc. of IEEE ISIT*, Jan. 2000.

[4] G. Kramer, M. Gastpar and P. Gupta, "Cooperative strategies and capacity theorems for relay networks," *IEEE Trans. Inf. Theory*, vol. 51, pp. 3037–3063, Sep. 2005.

[5] Y. Liang and G. Kramer, "Rate regions for relay broadcast channels," *IEEE Trans. Inf. Theory*, vol. 53, pp. 3517–3535, Oct. 2007.

[6] O. Sahin and E. Erkip, "Achievable rates for the Gaussian interference relay channel", in *Proc. of IEEE GLOBECOM*, Nov. 2007.

[7] A. Avestimehr and T. Ho, "Approximate capacity of the symmetric half-duplex Gaussian butterfly network," in *Proc. of IEEE ITW*, Jun. 2009.

[8] D. Gündüz, O. Simeone, A. J. Goldsmith, H. V. Poor, and S. Shamai, "Relaying simultaneous multicast messages," in *Proc. of IEEE ITW*, Jun. 2009.

[9] ——, "Multiple multicasts with the help of a relay," *IEEE Trans. Inf. Theory*, vol. 56, pp. 6142–6158, Dec. 2010.

[10] N. Devroye, P. Mitran, and V. Tarokh, "Achievable rates in cognitive radio channels," *IEEE Trans. Inf. Theory*, vol. 52, pp. 1813–1827, May 2006.

[11] S. I. Bross, A. Lapidoth, and M. A. Wigger, "The Gaussian MAC with conferencing encoders," in *Proc. of IEEE ISIT*, Jul. 2008.

[12] O. Simeone, D. Gündüz, H. V. Poor, A. J. Goldsmith, and S. Shamai, "Compound multiple-access channels with partial cooperation," *IEEE Trans. Inf. Theory*, vol. 55, pp. 2425–2441, Jun. 2009.

[13] B. Wang, J. Zhang, and A. Høst-Madsen, "On the capacity of MIMO relay channels," *IEEE Trans. Inf. Theory*, vol. 51, pp. 29–43, Jan. 2005.

[14] S. Simoens, O. Muñoz-Medina, J. Vidal, and A. del Coso, "On the Gaussian MIMO relay channel with full channel state information," *IEEE Trans. Signal Proc.*, vol. 57, pp. 3588–3599, Sep. 2009.

[15] J. Du, M. Xiao, and M. Skoglund, "Capacity bounds for relay-aided wireless multiple multicast with backhaul," in proc. of *WCSP*, Oct. 2010.

[16] J. Du, M. Xiao, and M. Skoglund, "Cooperative network coding strategies for wireless relay networks with backhaul," accepted for publication in *IEEE Trans. Commu.*, to appear in Sep. 2011.

[17] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, New York, Wiley, 2006.

[18] ——, "Determinant inequalities via information theory," *SIAM journal on matrix analysis and applications*, vol. 9, pp. 384–392, Jul. 1998.