2017 IEEE International Symposium on Information Theory (ISIT)

On the Capacity of Cloud Radio Access Networks

Shouvik Ganguly and Young-Han Kim
Department of Electrical and Computer Engineering
University of California, San Diego
Email: shgangul@eng.ucsd.edu, yhk@ucsd.edu

***Abstract*—Uplink and downlink cloud radio access networks are modeled as two-hop K-user L-relay networks, whereby small base-stations act as relays and are connected to a central processor via orthogonal links of finite capacity. Simplified versions of noisy network coding and distributed decode–forward are used to establish inner bounds on the capacity region for uplink and downlink communications, respectively. Through a careful analysis, the uplink inner bound is shown to achieve the cutset bound on the capacity region universally within $O(\log L)$ bits per user. The downlink inner bound achieves the cutset bound with a slightly looser gap of $O(\log(KL))$. These tight per-user gap results are extended to the situations in which the nodes have multiple antennas.**

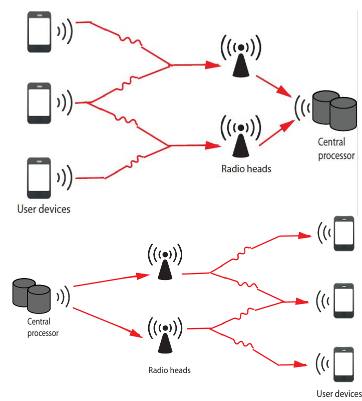

Fig. 1: (a) Uplink and (b) downlink cloud radio access network

I. Introduction

With ever-increasing demands for higher data rates, better coverage, and reliability of communication for a large number of devices, novel network protocols and architectures are expected to play an important role in future communication systems. The cloud radio access network (C-RAN) architecture [1] is one of the promising candidates, in which communication over a group of cells is coordinated by a cloud-based central processor. Fig. 1 depicts C-RAN uplink and downlink systems schematically. Base-stations in the C-RAN architecture, unlike traditional wireless systems, do not perform the complete processing locally, but are instead connected to the central processor through wired or wireless fronthaul links.

If these links have unbounded capacities, the C-RAN can be viewed as a "distributed" multiple input multiple output (MIMO) system. Base-stations act as remote radio heads that use beamforming to coordinate transmission and mitigate interference among multiple cells. For the more realistic situation of limited capacities, the optimal beamforming solution is typically computed, assuming infinite fronthaul capacities, and then compressed individually, which is then applied at the base-stations.

As an alternative to this greedy beamforming-compression approach, this paper investigates the optimal coding scheme for the entire system by modeling the C-RAN as a two-hop relay network. In this model, the base-stations act as relays that summarize the received signals to the central processor (uplink) and transmit the prescribed signals from the central processor (downlink) [1].

The two-hop relay network model for the uplink C-RAN was studied by Zhou et al. [2], who applied network compress–forward [3] to this case and showed, by optimizing over quantizers, that under some symmetry assumptions, it is possible to achieve a sum rate gap from the cutset bound that is linear in the number of base-stations. Sanderovich et al. [4] used the same scheme without optimization and analyzed the large-user asymptotics of achievable rates under the scenario where each fronthaul link has a fixed capacity. Zhou et al. [5] subsequently showed that under a sum-capacity constraint on the fronthaul links, the coding scheme in [4] and [2] can be simplified through successive cancellation decoding, generalizing an earlier single-user multiple relay scheme [6].

In this paper, we apply noisy network coding [7] to the uplink two-hop relay network model and the scheme achieves within $O(\log L)$ bits per user from the capacity region, where L is the number of relays (base-stations), regardless of the channel gain matrix, power constraint, and the number of users. Compared to network compress–forward, noisy network coding is simpler in that relays send the compression indices of the received signals directly without binning (hashing). This simpler operation inter alia achieves higher rates than network compress–forward for general networks, but for the uplink C-RAN, the achievable rate regions coincide. Hence, our main contribution for the uplink C-RAN can be viewed as a refinement of the capacity analyses in [4] and [2].

For the downlink, a variety of coding schemes have been proposed. Hong and Caire [8] studied a low-complexity reverse compute–forward scheme for symmetric rates. Liu et al. [9] applied network coding and beamforming to the downlink model with a noiseless multi-hop fronthaul. Motivated by the MAC–BC duality, Liu et al. [10] proposed suboptimal compression-based schemes and established a duality between achievable rate regions for the uplink and downlink C-RANs.

In this paper, as an alternative approach, we specialize

978-1-5090-4096-4/17/$31.00 ©2017 IEEE

and simplify the distributed decode–forward coding scheme [11] to the downlink C-RAN with capacity-limited single-hop fronthaul. In this scheme, multicoding at the encoder (as in Marton coding for broadcast channels [12]) is coupled with coding for fronthaul links, which allows more efficient coordination among the transmitted signals at the base-stations. We show that our rate region achieves a per-user gap of $\frac{1}{2}(1.45+\log(LK))$ from the cutset bound, where L and K are the number of relays and users. This refines the best-known linear gap from capacity for this model (implicit in [10]).

The rest of the paper is organized as follows. Section II studies the uplink model. Section II-A describes the general inner and outer bounds on the capacity region; Section II-B specializes the noisy network coding inner bound to the Gaussian network model and estblishes the capacity gap; and Section II-C generalizes the gap result to the MIMO case. Section III parallels the same flow for the downlink C-RAN. Throughout the paper, we follow the notation in [12]. In addition, $||A||_F := \sqrt{\mathbf{tr}(AA^T)} = \sqrt{\mathbf{tr}(A^TA)}$ denotes the Frobenius norm of a matrix A. All logarithms are to base 2 and all information measures are in bits.

II. Uplink Communication

A. General Model

We model the uplink C-RAN as a two-hop relay network in Fig. 2, where the first hop, namely, the (wireless) channel from the user devices to the radio heads, is modeled as a discrete memoryless network $p(y^L|x^K)$, and the second hop, namely, the channel from the radio heads to the central processor, consists of orthogonal links of capacities $C_1, \ldots, C_L$ bits per transmission, decoupled from the first hop. To be more precise, the channel output at the central processor (receiver) is $(W_1, \ldots, W_L)$, where $W_l \in [1 : 2^{nC_l}]$ is a reliable estimate of what relay l communicates to the receiver over n transmissions. We assume without loss of generality that these communication links are noiseless.

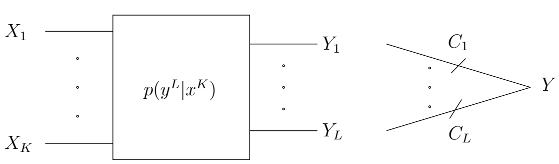

Fig. 2: Uplink network model

A $(2^{nR_1}, \ldots, 2^{nR_K}, n)$ code for this network consists of K message sets $[1 : 2^{nR_1}], \ldots, [1 : 2^{nR_K}]$; K encoders, where encoder $k \in [1 : K]$ assigns a codeword x_k^n to each $m_k \in [1 : 2^{nR_k}]$; L relay encoders, where relay encoder $l \in [1 : L]$ assigns an index $w_l \in [1 : 2^{nC_l}]$ to each received sequence y_l^n; and a decoder that assigns message estimates $(\hat{m}_1, \ldots, \hat{m}_K)$ to each index tuple w^L. We assume that the messages $M_1, \ldots, M_K$ are uniformly distributed and independent of each other. The average probability of error is defined as $P_e^{(n)} = P(\cup_{k=1}^K \{\hat{M}_k \neq M_k\})$. A rate tuple $(R_1, \ldots, R_K)$ is achievable if there is a sequence of $(2^{nR_1}, \ldots, 2^{nR_K}, n)$ codes with $\lim_{n\to\infty} P_e^{(n)} = 0$. The capacity region is defined as the closure of the set of all achievable rate tuples.

We have the following inner bound [4] on the capacity region of this network.

Proposition 1: A rate tuple $(R_1, \ldots, R_K)$ is achievable if

$$\sum_{k\in S_1} R_k < I(X(S_1); \hat{Y}(S_2^c)|X(S_1^c)) + \sum_{l\in S_2} C_l - \sum_{l\in S_2} I(Y_l; \hat{Y}_l|X^K) \tag{1}$$

for all $S_1 \subseteq [1 : K]$ and $S_2 \subseteq [1 : L]$ for some pmf $\prod_{k=1}^K p(x_k) \prod_{l=1}^L p(\hat{y}_l|y_l)$.

This inner bound was established by specializing the network compress–forward scheme in [6]. Roughly speaking, each relay compresses its received sequence and sends the bin index of the compression index. A more straightforward scheme can be developed by specializing noisy network coding [7] and simplifying it to our network model. This simplified coding scheme and its analysis will be presented in a longer version of this manuscript and is omitted here. The cutset bound [13] for this network can be characterized by the set of $(R_1, \ldots, R_K)$ such that

$$\sum_{k\in S_1} R_k < I(X(S_1); Y(S_2^c)|X(S_1^c)) + \sum_{l\in S_2} C_l \tag{2}$$

for all $S_1 \subseteq [1 : K]$ and $S_2 \subseteq [1 : L]$ for some pmf $p(x^K)$.

B. Gaussian Model

We now assume that

$$Y^L = GX^K + Z^L,$$

where $G \in \mathbb{R}^{L\times K}$ is a (deterministic) channel gain matrix and Z^L is a vector of independent $\mathrm{N}(0, 1)$ noise components. Assume the average power constraint P on each sender, i.e.,

$$\sum_{i=1}^n x_{ki}^2(m_k) \le nP, \quad m_k \in [1 : 2^{nR_k}], \quad k \in [1 : K].$$

Our main goal of this section is to quantify how well network compress–forward (or noisy network coding), with achievable rates in (1), performs for this Gaussian network. In particular, we bound the per-user rate gap Δ such that if $(R_1, \ldots, R_K)$ lies in the cutset bound, the rate tuple $((R_1 - \Delta), \ldots, (R_K - \Delta))$ will lie in the inner bound in (1), regardless of G and P.

It can be shown that Proposition 1 can be simplified to establish the achievability of all $(R_1, \ldots, R_K)$ such that

$$\begin{aligned}\sum_{k\in S_1} R_k &< \frac{1}{2}\log\left|\frac{P}{\sigma^2+1} G_{S_2^c,S_1} G_{S_2^c,S_1}^T + I\right| \\ &\quad + \sum_{l\in S_2} C_l - \frac{|S_2|}{2}\log\left(1+\frac{1}{\sigma^2}\right) \\ &=: f_{\mathrm{in}}(S_1, S_2),\end{aligned} \tag{3}$$

where $G_{S_2^c,S_1}$ is the submatrix of G formed by the rows with indices in S_2^c and the columns with indices in S_1. This follows by considering X^K to be a vector of i.i.d. $\mathrm{N}(0, P)$ random variables, and setting $\hat{Y}_l = Y_l + \hat{Z}_l$, $l \in [1 : L]$, where $\hat{Z}_l \sim \mathrm{N}(0,\sigma^2)$. Note that for any $S_2 \subseteq [1:L]$ and $S_1' \subseteq S_1$, $G_{S_2^c,S_1}G_{S_2^c,S_1}^T \succeq G_{S_2^c,S_1'}G_{S_2^c,S_1'}^T$ and thus, $f_{\text{in}}(S_1',S_2) \le f_{\text{in}}(S_1,S_2)$, which implies that

$$\min_{S_2} f_{\text{in}}(S_1',S_2) \le \min_{S_2} f_{\text{in}}(S_1,S_2). \tag{4}$$

The cutset bound in (2) can also be simplified and relaxed as

$$\begin{aligned}\sum_{k\in S_1} R_k &< \frac{1}{2}\log\left|G_{S_2^c,S_1}\Sigma_{S_1|S_1^c}G_{S_2^c,S_1}^T + I\right| + \sum_{l\in S_2} C_l \\ &\overset{(a)}{\le} \frac{1}{2}\log\left|PG_{S_2^c,S_1}G_{S_2^c,S_1}^T + I\right| + \frac{|S_1|}{2} + \sum_{l\in S_2} C_l \\ &=: f_{\text{out}}(S_1,S_2),\end{aligned} \tag{5}$$

where (a) follows in a similar manner as equation (34) in Section V of [11], and $\Sigma_{S_1|S_1^c}$ is the conditional covariance matrix of $X(S_1)$ given $X(S_1^c)$. For this Gaussian network model, we have the following upper bound on the gap from capacity.

Theorem 1: For every $G \in \mathbb{R}^{L\times K}$ and every $P \in \mathbb{R}^+$, if a rate tuple $(R_1,\ldots,R_K)$ is in the cutset bound (5), then the rate tuple $((R_1-\Delta)^+,\ldots,(R_K-\Delta)^+)$ is achievable, where

$$\Delta \le \frac{1}{2}(2.45 + \log L).$$

Proof: Let

$$\Delta := \max_{\substack{S_1\subseteq[1:K]\\ S_1\neq\emptyset}} \frac{\min_{S_2} f_{\text{out}}(S_1,S_2) - \min_{S_2} f_{\text{in}}(S_1,S_2)}{|S_1|}. \tag{6}$$

Suppose that $(R_1,\ldots,R_K)$ lies in the cutset bound, and let $\mathcal{A} = \{k : R_k > \Delta\}$. Then, for every nonempty $S_1 \subseteq [1:K]$,

$$\begin{aligned}&\sum_{k\in S_1}(R_k-\Delta)^+ \\ &= \sum_{k\in S_1\cap\mathcal{A}}(R_k-\Delta) \\ &= \sum_{k\in S_1\cap\mathcal{A}} R_k - |S_1\cap\mathcal{A}|\Delta \\ &\overset{(a)}{\le} \min_{S_2}\Big[f_{\text{out}}(S_1\cap\mathcal{A},S_2)\Big] \\ &\qquad - \Big(\min_{S_2} f_{\text{out}}(S_1\cap\mathcal{A},S_2) - \min_{S_2} f_{\text{in}}(S_1\cap\mathcal{A},S_2)\Big) \\ &= \min_{S_2} f_{\text{in}}(S_1\cap\mathcal{A},S_2) \\ &\overset{(b)}{\le} \min_{S_2} f_{\text{in}}(S_1,S_2),\end{aligned}$$

where (a) follows from the cutset bound (5), and the fact that

$$\begin{aligned}\Delta &= \max_{S_1} \frac{\min_{S_2} f_{\text{out}}(S_1,S_2) - \min_{S_2} f_{\text{in}}(S_1,S_2)}{|S_1|} \\ &\ge \frac{\min_{S_2} f_{\text{out}}(S_1\cap\mathcal{A},S_2) - \min_{S_2} f_{\text{in}}(S_1\cap\mathcal{A},S_2)}{|S_1\cap\mathcal{A}|},\end{aligned}$$

and (b) follows from (4). Hence, Δ, as defined in (6) satisfies the requirements of Theorem 1. Now, for every $\sigma^2 > 0$,

$$\begin{aligned}\Delta &= \max_{S_1} \frac{\min_{S_2} f_{\text{out}}(S_1,S_2) - \min_{S_2} f_{\text{in}}(S_1,S_2)}{|S_1|} \\ &\overset{(a)}{\le} \max_{S_1,S_2} \frac{f_{\text{out}}(S_1,S_2) - f_{\text{in}}(S_1,S_2)}{|S_1|} \\ &\overset{(b)}{=} \max_{S_1,S_2} \frac{1}{2|S_1|}\left[\log\frac{\left|PG_{S_2^c,S_1}G_{S_2^c,S_1}^T + I\right|}{\left|\frac{P}{\sigma^2+1}G_{S_2^c,S_1}G_{S_2^c,S_1}^T + I\right|} + |S_1|\right. \\ &\overset{(c)}{=} \max_{S_1,S_2} \frac{1}{2|S_1|}\left[\sum_{i=1}^{\text{rank}(G_{S_2^c,S_1})}\log\frac{P\beta_i+1}{\frac{P}{\sigma^2+1}\beta_i+1} + |S_1|\right. \\ &\qquad\qquad \left. + |S_2|\log\Big(1+\frac{1}{\sigma^2}\Big)\right] \\ &\overset{(d)}{\le} \max_{\substack{k\in[1:K]\\ l\in[0:L]}} \Big[\frac{\min\{L-l,k\}}{2k}\log(1+\sigma^2) \\ &\qquad\qquad + \frac{l}{2k}\log\Big(1+\frac{1}{\sigma^2}\Big) + \frac{1}{2}\Big].\end{aligned} \tag{7}$$

Here, (a) follows from the fact that for functions f and g defined over a finite set $\mathcal{X}$, such that $g \ge f$ everywhere on $\mathcal{X}$, $\min_{x\in\mathcal{X}} g(x) - \min_{x\in\mathcal{X}} f(x) \le \max_{x\in\mathcal{X}}[g(x)-f(x)]$, (b) follows from (5) and (3), and in (c), $\beta_1,\beta_2,\ldots$ are the (non-negative) eigen-values of $G_{S_2^c,S_1}G_{S_2^c,S_1}^T$. Finally, in (d), we take $|S_1| = k$, $|S_2| = l$, and upper-bound $\text{rank}(G_{S_2^c,S_1})$ by $\min\{L-l,k\}$. The maximization in (7) yields

$$\begin{cases}\frac{1}{2}\log(\sigma^2+1) + \frac{L-1}{2}\log(1+\frac{1}{\sigma^2}) + \frac{1}{2}, & \sigma^2 \ge 1,\\ \frac{L}{2}\log(1+\frac{1}{\sigma^2}) + \frac{1}{2}, & \sigma^2 \le 1.\end{cases}$$

Since this holds for every $\sigma^2 > 0$, we set $\sigma^2 = L-1$ for $L \ge 2$ to obtain

$$\begin{aligned}\Delta &\le \frac{1}{2}\log L + \frac{L-1}{2}\log\Big(1+\frac{1}{L-1}\Big) + \frac{1}{2} \\ &\le \frac{1}{2}\log L + \frac{L-1}{2}\cdot\frac{1}{(L-1)\ln 2} + \frac{1}{2} \\ &\le \frac{1}{2}(2.45 + \log L).\end{aligned} \tag{8}$$

For $L = 1$, we take $\sigma^2 = 1$ to obtain $\Delta \le 1$ bit. This, together with (8), establishes Theorem 1. ■

C. MIMO Model

We now generalize Theorem 1 to the situation in which the senders and relays have multiple antennas. For simplicity, we assume that every sender has t antennas and every relay has r antennas. We also assume the average power constraint P at each transmit antenna.

Proposition 2: If $(R_1,\ldots,R_K)$ lies in the cutset bound, then $((R_1-\Delta)^+,\ldots,(R_K-\Delta)^+)$ is achievable, where

$$\Delta \le \begin{cases}\frac{t}{2}\Big(2.45 + \log\Big(\frac{Lr}{t}\Big)\Big), & Lr > 2t\\ \frac{Lr+t}{2}, & Lr \le 2t.\end{cases} \tag{9}$$

Proof sketch: First assume that $t = 1$. In this case, the sequence of steps leading to (3) and (5) go through almost unchanged, except for a slight change of notation in that $G_{S_2^c,S_1}$ is now the $r|S_2^c| \times |S_1|$ channel gain matrix from the senders in S_1 to the relays in S_2^c. Also, the last term in (3) becomes $\frac{r|S_2|}{2} \log\left(1+\frac{1}{\sigma^2}\right)$. The relation (7) now reads:

$$\Delta \leq \max_{k,l} \Big[\frac{\min\{(L-l)r, k\}}{2k} \log(1+\sigma^2) + \frac{lr}{2k} \log\left(1+\frac{1}{\sigma^2}\right) + \frac{1}{2}\Big].$$

Manipulating this expression as before, we show that $\Delta \leq \frac{1}{2}(2.45 + \log(Lr))$.

For general r and t, (7) becomes

$$\Delta \leq \max_{k,l} \Big[\frac{\min\{(L-l)r, kt\}}{2k} \log(1+\sigma^2) + \frac{lr}{2k} \log\left(1+\frac{1}{\sigma^2}\right) + \frac{t}{2}\Big].$$

For $Lr \leq 2t$, the maximization, followed by substituting $\sigma^2 = 1$, yields $\Delta \leq \frac{Lr+t}{2}$. For $Lr > 2t$ and $\sigma^2 \geq 1$, the maximization yields $\Delta \leq \frac{t}{2}\log(1+\sigma^2) + \frac{Lr-t}{2}\log(1+\frac{1}{\sigma^2}) + \frac{t}{2}$; by setting $\sigma^2 = \frac{Lr-t}{t}$, we have

$$\begin{aligned}\Delta &\leq \frac{Lr}{2}\Big(\frac{t}{rL}\log(Lr/t) + \frac{Lr-t}{Lr}\log\Big(1+\frac{t}{Lr-t}\Big)\Big) + \frac{t}{2} \\ &\leq \frac{t}{2}\Big(1+\log(Lr/t)\Big) + \frac{Lr-t}{2}\cdot\frac{t}{(Lr-t)\ln 2} \\ &= \frac{t}{2}\Big(2.45+\log(Lr/t)\Big),\end{aligned}$$

which establishes (9). ■

III. Downlink Communication

A. General Model

Similar to the uplink, we model the downlink C-RAN as a two-hop relay network in Fig. 3, where the first hop (central processor to radio heads) consists of orthogonal noiseless links of capacities $C_1, \ldots, C_L$ bits per transmission and the second hop (radio heads to user devices) is modeled as a discrete memoryless network $p(y^K|x^L)$.

A $(2^{nR_1}, \ldots, 2^{nR_K}, n)$ code for this network consists of K message sets $[1:2^{nR_1}], \ldots, [1:2^{nR_K}]$; an encoder $w^L(m_1, \ldots, m_K) \in \prod_{l=1}^L [1:2^{nC_l}]$; relay encoders $x_l^n(w_l)$, $l \in [1:L]$; and decoders $\hat{m}_k(y_k^n) \in [1:2^{nR_k}]$, $k \in [1:K]$. The average probability of error, achievability of a rate tuple, and the capacity region are defined as before.

The following inner bound can be established by specializing the distributed decode–forward scheme [11]. We describe a simplified version of the coding scheme in the Appendix.

Proposition 3: A rate tuple $(R_1, \ldots, R_K)$ is achievable, if

$$\sum_{k\in S_2^c} R_k < I(X(S_1); U(S_2^c)|X(S_1^c)) + \sum_{l\in S_1^c} C_l - \sum_{k\in S_2^c} I(U_k; X^L|Y_k)$$

for all $S_1 \subseteq [1:L]$ and $S_2 \subseteq [1:K]$ for some pmf $\prod_{l=1}^L p(x_l)\prod_{k=1}^K p(u_k|x^L)$.

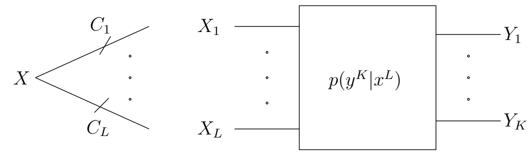

Fig. 3: Downlink network model

The cutset bound for this network is characterized by

$$\sum_{k\in S_2^c} R_k < I(X(S_1); Y(S_2^c)|X(S_1^c)) + \sum_{l\in S_1^c} C_l \tag{10}$$

for all $S_1 \subseteq [1:L]$ and $S_2 \subseteq [1:K]$ for some pmf $p(x^L)$.

B. Gaussian Model

Similar to Section II-B, we now assume that $Y^K = GX^L + Z^K$, where $G \in \mathbb{R}^{K\times L}$ is a channel gain matrix and Z^K consists of i.i.d. $\mathrm{N}(0,1)$ noise components, and assume the average power constraint P at each relay. The rest of this section is devoted to bounding the achievable per-user rate gap Δ from the cutset bound. First, Proposition 3 can be simplified to establish the achievability of all $(R_1, \ldots, R_K)$ such that

$$\begin{aligned}\sum_{k\in S_2^c} R_k &< \frac{1}{2}\log\Big|\frac{P}{\sigma^2}G_{S_2^c,S_1}G_{S_2^c,S_1}^T + I\Big| + \sum_{l\in S_1^c} C_l - \frac{|S_2^c|}{2}\log\Big(1+\frac{1}{\sigma^2}\Big) \\ &=: F_{\text{in}}(S_1, S_2)\end{aligned} \tag{11}$$

for $S_1 \subseteq [1:L]$ and $S_2 \subseteq [1:K]$. This follows by setting X^L to be a vector of i.i.d. $\mathrm{N}(0,P)$ random variables and defining $U^K = GX^L + \hat{Z}^K$, where $\hat{Z}^K \sim \mathrm{N}(0, \sigma^2 I)$ is independent of Z^K. The cutset bound (10) simplifies to

$$\begin{aligned}\sum_{k\in S_2^c} R_k &< \frac{1}{2}\log\Big|G_{S_2^c,S_1}\Sigma_{S_1|S_1^c}G_{S_2^c,S_1}^T + I\Big| + \sum_{l\in S_1^c} C_l \\ &=: F_{\text{out}}(S_1, S_2)\end{aligned} \tag{12}$$

for all $S_1 \subseteq [1:L]$ and $S_2 \subseteq [1:K]$.

We now have the following.

Theorem 2: For every G and P, if $(R_1, \ldots, R_K)$ is in the cutset bound (12), then $((R_1-\Delta)^+, \ldots, (R_K-\Delta)^+)$ is achievable, where

$$\Delta \leq \frac{1}{2}(1.45 + \log(KL)).$$

Proof: Note that unlike (4), F_{in} is not necessarily monotonic. We overcome this difficulty by rephrasing the inner bound (11) as

$$\sum_{k\in S_2^c} R_k < \min_{T_2\subseteq S_2} F_{\text{in}}(S_1, T_2). \tag{13}$$

We observe that the right-hand side of (13) is increasing with S_2^c for a fixed S_1, so we can apply the technique developed in Section II-B to compute an upper bound on Δ. We thus write

$$\Delta = \max_{S_2 \subsetneq [1:K]} \Big[\frac{\min_{S_1} F_{\text{out}}(S_1, S_2)}{|S_2^c|}$$

$$- \frac{\min_{S_1} \min_{T_2 \subseteq S_2} F_{\rm in}(S_1, T_2)}{|S_2^c|}\Bigg]$$

$$\leq \max_{\substack{S_1 \subseteq [1:L] \\ S_2 \subsetneq [1:K] \\ T_2 \subseteq S_2}} \frac{F_{\rm out}(S_1, S_2) - F_{\rm in}(S_1, T_2)}{|S_2^c|}$$

$$= \max_{\substack{S_1 \subseteq [1:L] \\ S_2 \subsetneq [1:K] \\ T_2 \subseteq S_2}} \frac{1}{2|S_2^c|} \Bigg[\log \frac{\left|G_{S_2^c,S_1} \Sigma_{S_1|S_1^c} G_{S_2^c,S_1}^T + I\right|}{\left|\frac{P}{\sigma^2} G_{T_2^c,S_1} G_{T_2^c,S_1}^T + I\right|} + |T_2^c| \log\left(1 + \frac{1}{\sigma^2}\right)\Bigg]$$

$$\overset{(a)}{\leq} \max_{\substack{S_1 \subseteq [1:L] \\ S_2 \subsetneq [1:K] \\ T_2 \subseteq S_2}} \frac{1}{2|S_2^c|} \Bigg[\log \frac{\left|G_{S_2^c,S_1} \Sigma_{S_1} G_{S_2^c,S_1}^T + I\right|}{\left|\frac{P}{\sigma^2} G_{S_2^c,S_1} G_{S_2^c,S_1}^T + I\right|} + |T_2^c| \log\left(1 + \frac{1}{\sigma^2}\right)\Bigg], \quad (14)$$

where (a) follows since $\Sigma_{S_1} \succeq \Sigma_{S_1|S_1^c}$ and for any matrix A and $\alpha > 0$, $|I + \alpha AA^T|$ increases when we add more rows to A. Writing $\Sigma_{S_1} = U\Lambda U^T$, where U is orthogonal and Λ is diagonal, and letting $G_{S_2^c,S_1}U = \begin{bmatrix} b_1 & b_2 & \cdots & b_{|S_1|}\end{bmatrix}$, where $b_1, \ldots, b_{|S_1|}$ are $|S_2^c| \times 1$ vectors constrained by $\sum_{l=1}^{|S_1|} ||b_l||^2 = ||G_{S_2^c,S_1}||_F^2$, we have

$$log \frac{|G_{S_2^c,S_1} \Sigma_{S_1} G_{S_2^c,S_1}^T + I|}{|\frac{P}{\sigma^2} G_{S_2^c,S_1} G_{S_2^c,S_1}^T + I|} = \log \frac{\left|I + \sum_{l=1}^{|S_1|} \lambda_l b_l b_l^T\right|}{\left|I + \frac{P}{\sigma^2} \sum_{l=1}^{|S_1|} b_l b_l^T\right|}$$

$$\overset{(a)}{\leq} \log \frac{\left|I + P|S_1| \sum_{l=1}^{|S_1|} b_l b_l^T\right|}{\left|I + \frac{P}{\sigma^2} \sum_{l=1}^{|S_1|} b_l b_l^T\right|}$$

$$\overset{(b)}{=} \sum_{k=1}^{|S_2^c|} \log \frac{1 + P|S_1|\mu_k}{1 + \frac{P}{\sigma^2}\mu_k}$$

$$\leq |S_2^c| \log\left(\sigma^2 |S_1|\right),$$

provided that $\sigma^2 \geq \frac{1}{|S_1|}$. Here, (a) follows since the trace of Σ_{S_1} is upper bounded by $P|S_1|$ and in (b), $\mu_1, \ldots, \mu_{|S_2^c|}$ are the (nonnegative) eigenvalues of $\sum_{l=1}^{|S_1|} b_l b_l^T$. Continuing from (14), we thus have

$$\Delta \leq \max_{\substack{S_1 \subseteq [1:L] \\ S_2 \subsetneq [1:K] \\ T_2 \subseteq S_2}} \left[\frac{|T_2^c| \log\left(1 + \frac{1}{\sigma^2}\right)}{2|S_2^c|} + \frac{1}{2} \log\left(\sigma^2 |S_1|\right)\right]$$

$$= \frac{K}{2} \log\left(1 + \frac{1}{\sigma^2}\right) + \frac{1}{2} \log(\sigma^2 L).$$

This holds for every $\sigma^2 \geq 1$, so we set $\sigma^2 = K - 1$ to obtain

$$\Delta \leq \frac{1}{2} \log L + \frac{1}{2}\Big(K \log K - (K-1)\log(K-1)\Big)$$

$$\overset{(a)}{\leq} \frac{1}{2}\left(\log L + \log K + \frac{1}{\ln 2}\right).$$

(a) follows since $(d/dx)(x \log x) = \log x + (1/\ln 2)$ and the latter is an increasing function of x. ∎

C. MIMO Model

As before, assume that every relay has t antennas and every receiver has r antennas with average power constraint P at each antenna. By slightly modifying the arguments of the previous section, it can be shown that the following holds.

Proposition 4: If $(R_1, \ldots, R_K)$ lies in the cutset bound, then $((R_1 - \Delta)^+, \ldots, (R_K - \Delta)^+)$ is achievable, where

$$\Delta \leq \begin{cases} \frac{r}{2}(\log(Lt) + \log K + 1.45), & Lt > r \\ \frac{Lt}{2}(\log r + \log K + 1.45), & Lt \leq r. \end{cases}$$

Appendix

The following coding scheme is a specialization of distributed decode–forward to our general downlink C-RAN model and achieves the inner bound in Proposition 3.

Codebook generation: Fix a pmf $p(x^L)\prod_{k=1}^K p(u_k|x^L)$. For each w_l, $l \in [1:L]$, generate $x_l^n(w_l) \sim \prod_{i=1}^n p_{X_l}(x_{li})$. Define auxiliary indices $s_k \in [1:2^{n\tilde{R}_k}]$, $k \in [1:K]$. Here, each $\tilde{R}_k$ is some non-negative auxiliary rate. For each $(m_k, s_k) \in [1:2^{nR_k}] \times [1:2^{n\tilde{R}_k}]$ and $k \in [1:K]$, generate $u_k^n(m_k, s_k) \sim \prod_{i=1}^n p_{U_k}(u_{ki})$.

Encoding: The encoder sends w^L such that $(x_1^n(w_1), \ldots, x_L^n(w_L), u_1^n(m_1, s_1), \ldots, u_K^n(m_K, s_K)) \in \mathcal{T}_{\epsilon'}^{(n)}$.

Relay encoding: Relay l transmits $x_l^n(w_l)$.

Decoding: Let $\epsilon > \epsilon'$. Upon receiving y_k^n, receiver k finds $\hat{m}_k$ such that $(u_k^n(\hat{m}_k, s_k), y_k^n) \in \mathcal{T}_\epsilon^{(n)}$ for some s_k.

References

[1] T. Q. S. Quek, M. Peng, O. Simone, and W. Yu, Eds., *Cloud Radio Access Networks: Principles, Technologies, and Applications*. Cambridge Univ. Press, March 2017.

[2] Y. Zhou and W. Yu, "Optimized backhaul compression for uplink cloud radio access network," *IEEE Journ. Sel. Are. Comm.*, vol. 32, no. 6, pp. 1295–1307, June 2014.

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

[4] A. Sanderovich, O. Somekh, H. V. Poor, and S. Shamai, "Uplink macro diversity of limited backhaul cellular network," *IEEE Trans. Inf. Theory*, vol. 55, no. 8, pp. 3457–3478, Aug 2009.

[5] Y. Zhou, Y. Xu, W. Yu, and J. Chen, "On the optimal fronthaul compression and decoding strategies for uplink cloud radio access networks," *IEEE Trans. Inf. Theory*, vol. 62, no. 12, pp. 7402–7418, Dec 2016.

[6] A. Sanderovich, S. Shamai, Y. Steinberg, and G. Kramer, "Communication via decentralized processing," *IEEE Trans. Inf. Theory*, vol. 54, no. 7, pp. 3008–3023, July 2008.

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

[8] S. N. Hong and G. Caire, "Compute-and-forward strategies for cooperative distributed antenna systems," *IEEE Trans. Inf. Theory*, vol. 59, no. 9, pp. 5227–5243, Sept 2013.

[9] L. Liu and W. Yu, "Joint sparse beamforming and network coding for downlink multi-hop cloud radio access networks," December 2016.

[10] L. Liu, P. Patil, and W. Yu, "An uplink-downlink duality for cloud radio access network," in *Proc. IEEE Int. Symp. Inf. Theory*, July 2016, pp. 1606–1610.

[11] S. Lim, K. Kim, and Y.-H. Kim, "Distributed decode-forward for relay networks," *arXiv:1510.00832 [cs.IT]*, October 2015.

[12] A. E. Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2011.

[13] A. E. Gamal, "On information flow in relay networks," in *IEEE Nat. Telecom Conf.*, vol. 2, November 1981, pp. D4.1.1–D4.1.4.