Approximate Capacity of Gaussian Relay Networks: Is a Sublinear Gap to the Cutset Bound Plausible?

Thomas A. Courtade
Department of Electrical Engineering and Computer Sciences
University of California, Berkeley
Email: courtade@eecs.berkeley.edu

Ayfer Özgür
Department of Electrical Engineering
Stanford University
Email: aozgur@stanford.edu

***Abstract*—Beginning with work by Avestimehr, Diggavi and Tse, there have been a series of papers showing that the capacity of Gaussian relay networks can be closely approximated by the cutset bound. More precisely, it is known that the gap between the cutset bound and capacity in these networks can be bounded by a function that grows linearly with the number of nodes in the network and is otherwise independent of network topology and channel configurations. We argue that this linear gap is fundamental to such approximations, and prove that improvement to a sublinear function is possible if, and only if, capacity is equal to the cutset bound for all Gaussian relay networks.**

I. Introduction

Characterizing the capacity of relay networks has been a long-standing open question in network information theory. The seminal work of Cover and El-Gamal has established several basic achievability schemes and upper bounds for the single relay channel [1]. More recently, motivated by the rapid proliferation of wireless devices, there has been significant interest in extending these techniques to understand the capacity of larger networks. With limited scope for an exact capacity characterization, an approach initiated by Avestimehr, Diggavi and Tse [2] and widely adapted by follow-up work [3]–[6] has been to approximate the capacity by providing a universal guarantee on the gap to optimality. In particular, this line of work explicitly bounds the gap between the capacity of the network and its information-theoretic cutset bound by a term that is independent of the channel coefficients and SNRs and that depends on the actual network only through the total number of nodes N. In particular, the sharpest currently known approximation is within $0.5N$ bits/s/Hz [6]. This approach has been useful in designing new achievable schemes that can perform well across different network configurations and topologies, and has also identified the cutset bound as a proxy for the true capacity.

However, an approximation gap that increases linearly in the total number of nodes quickly becomes too large even for networks of moderate size. An interesting question is whether this linear gap to the cutset bound can be substantially improved, for example, to scale sublinearly in the total number of nodes. [7]–[10] show that this can be done if additional restrictions are imposed on the network. For example, [7] shows that in the diamond topology with N relays the cutset bound can be achieved within a gap that is only logarithmic in N. By further restricting the channel gains of the diamond network to be identical, [8] shows that the cutset bound can be achieved within a constant gap independent of N. These improvements may lead one to hope for tighter approximations, where the gap between capacity and the cutset bound scales sublinearly in the network size, also in the general case.

In this paper, we present a somewhat negative result. We show that an improvement of the cutset-to-capacity approximation gap beyond linear in the number of nodes for general Gaussian relay networks would imply that the cutset bound is tight for *all* Gaussian relay networks. This strongly suggests that improving the gap to a sublinear function of N presents a formidable (if not impossible) challenge. In this sense, the result suggests a converse for the currently available approximation results and emphasizes the importance of continued investigation into upper bounds on capacity for specific small networks with few nodes.

Organization

This paper is organized as follows. Section II describes the Gaussian relay networks we consider and briefly surveys known approximations to capacity. Section III formalizes the notion of a capacity approximation theorem and presents our main results; proofs follow in Section IV. We conclude with a short discussion of our results and various generalizations in Section V.

II. System Model and Capacity Approximations

Throughout this paper, we consider a discrete memoryless Gaussian relay network of N nodes, in which a source node s aims to reliably communicate a message to a destination node d. For each node $i \in \{1, 2, \dots, N\} \triangleq \mathcal{N}$, we let R_i denote the number of node i's receive-antennas, and let T_i denote the number of node i's transmit-antennas. We adopt the usual Gaussian relay network setting, wherein if $\mathbf{x}_j[t] \in \mathbb{R}^{T_j}$ is the signal transmitted by node j at time instant t, the signal received at node i is given by

$$\mathbf{y}_i[t] = \sum_{j \in \mathcal{N}} G_{ij}\mathbf{x}_j[t] + \mathbf{z}_i[t], \tag{1}$$

This work was supported by the NSF Center for Science of Information under grant agreement CCF-0939370 and NSF CAREER award 1254786.

where $G_{ij} \in \mathbb{R}^{R_i \times T_j}$ is a known $R_i \times T_j$ matrix describing the channel gain from node j to i, $\mathbf{z}_i[t] \sim \mathcal{N}(0, I_{R_i \times R_i})$ is additive Gaussian noise with $\{\mathbf{z}_1[t], \mathbf{z}_2[t], \ldots, \mathbf{z}_N[t]\}_{t=1,2,\ldots}$ being mutually independent. In this manner, a Gaussian relay network is completely characterized by the triple (G, s, d), where G denotes the collection of channel gain matrices $\{G_{ij} : i, j \in \mathcal{N}\}$. For a network (G, s, d), it will be convenient to define the quantity

$$\kappa(G, s, d) \triangleq \sum_{i \in \mathcal{N}} \max\{T_i, R_i\} \tag{2}$$

since it will be referred to frequently. When the network (G, s, d) under consideration is clear from context, we will abbreviate $\kappa \equiv \kappa(G, s, d)$.

For a positive rate R and a non-negative integer n, a $(2^{nR}, n)$-code consists of a set of encoders

$$\mathsf{x}_i^{(k)} : \mathbf{y}_i[1:k-1] \mapsto \mathbf{x}_i[k] \quad i \in \mathcal{N} \setminus \{s\} \tag{3}$$

$$\mathsf{x}_s^{(k)} : (\mathsf{m}, \mathbf{y}_s[1:k-1]) \mapsto \mathbf{x}_s[k] \quad \mathsf{m} \in \{1, 2, \ldots, 2^{nR}\} \tag{4}$$

which satisfy the power constraints

$$\frac{1}{n} \sum_{k=1}^{n} \mathbb{E}\left[\|\mathsf{x}_i^{(k)}(\mathbf{Y}_i[1:k-1])\|^2\right] \le T_i P \quad i \in \mathcal{N} \setminus \{s\}$$

$$\frac{1}{n} \sum_{k=1}^{n} \mathbb{E}\left[\|\mathsf{x}_s^{(k)}(M, \mathbf{Y}_s[1:k-1])\|^2\right] \le T_s P, \tag{5}$$

where expectation is taken with respect to a uniformly random message $M \sim \mathrm{Unif}(\{1, 2, \ldots, 2^{nR}\})$ and the channel noise.

Given a $(2^{nR}, n)$-code, the corresponding average probability of error is defined to be

$$P_e^{(n)} = \min_{\phi} \mathbb{P}\{\phi(\mathbf{Y}_d[1:n]) \neq M\}, \tag{6}$$

where the minimum is taken over all decoders that reproduce a message based on the decoder's observation $\mathbf{y}_d[1:n]$

$$\phi : \mathbf{y}_d[1:n] \mapsto \hat{M} \in \{1, 2, \ldots, 2^{nR}\}. \tag{7}$$

For the network (G, s, d), a rate R is said to be *achievable* if there exists a sequence of $(2^{nR}, n)$-codes for which $P_e^{(n)} \to 0$ as $n \to \infty$. For a Gaussian relay network (G, s, d), the *capacity* $\mathcal{C}(G, s, d)$ is defined to be the supremum of achievable rates.

For a network (G, s, d), the cutset bound [11] is given by:

$$\bar{\mathcal{C}}(G, s, d) \triangleq \sup_{p(\mathbf{x}_1, \ldots, \mathbf{x}_N)} \min_{S : s \in S, d \in S^c} I(\mathbf{X}_S; \mathbf{Y}_{S^c} | \mathbf{X}_{S^c}), \tag{8}$$

where the supremum is over all joint distributions $p(\mathbf{x}_1, \ldots, \mathbf{x}_N)$ on $\prod_{i=1}^{N} \mathbb{R}^{T_i}$ satisfying the power constraints $\mathbb{E}\left[\|\mathbf{X}_i\|^2\right] \le T_i P$ for $i \in \mathcal{N}$, the minimum is over all subsets $S \subset \mathcal{N}$ that separate s from d, and the conditional distribution of $\mathbf{y}_1, \ldots, \mathbf{y}_N$ given $\mathbf{x}_1, \ldots, \mathbf{x}_N$ is induced by the channel model (1).

Following a series of other works (e.g., [2]–[4]), Lim et al. have proved the following approximation result:

Theorem 1. *[6] For any Gaussian relay network (G, s, d),*

$$\mathcal{C}(G, s, d) \ge \bar{\mathcal{C}}(G, s, d) - 0.5\,\kappa(G, s, d). \tag{9}$$

Since $\mathcal{C}(G, s, d) \le \bar{\mathcal{C}}(G, s, d)$ always, Theorem 1 establishes that the cutset bound approximates the capacity $\mathcal{C}(G, s, d)$ within a factor that is linear in the parameter κ. An interesting question is whether (9) can be substantially improved. For example, is it possible to replace the slack term 0.5κ with 0.1κ, or with a sublinear term such as $\frac{\kappa}{\log \log \kappa}$?

When additional restrictions are imposed on the network topology, it is already known that Theorem 1 can be improved. As a specific example, it has been shown in [7] that, for the diamond network with $N-2$ relays,

$$\mathcal{C}(G, s, d) \ge \bar{\mathcal{C}}(G, s, d) - 2 \log(\kappa - 2) \tag{10}$$

when all nodes have one antenna (i.e., $\kappa = N$). Niesen and Diggavi [8] have sharpened the approximation to

$$\mathcal{C}(G, s, d) \ge \bar{\mathcal{C}}(G, s, d) - 1.8 \tag{11}$$

for the N-relay diamond network in the special case where all channel gains are identical. For certain layered networks in which intermediate relays have one antenna, Kolte et al. [10] have shown that

$$\mathcal{C}(G, s, d) \ge \bar{\mathcal{C}}(G, s, d) - O(K^2 \log D + K \log K), \tag{12}$$

where D is the number of layers, and K is the number of relays in each layer. In this setting, if D grows sufficiently fast relative to K, the $O(K^2 \log D + K \log K)$ term scales as $o(\kappa)$.

III. Main Results

In light of the empirical evidence described in the previous section, it is natural to conjecture that the inequality in Theorem 1 can be improved to a bound of the form[1]

$$\mathcal{C}(G, s, d) \ge \bar{\mathcal{C}}(G, s, d) - o(\kappa(G, s, d)). \tag{13}$$

Indeed, this was posed as an open question by Niesen and Diggavi in [8]. In the present paper, we essentially put this question to rest by observing that such an improvement is impossible, unless the cutset bound is tight for *all* Gaussian relay networks. Toward doing so, we now define a general template for approximating capacity via the cutset bound.

In the spirit of the approximation results proved in [2]–[10], we define a *Gaussian Relay Network Approximation Theorem* with parameter γ (abbreviated as γ-GRNAT) to be a claim of the following form:

There exists a constant $\gamma \ge 0$ and a function $f(n) = o(n)$ such that, for any Gaussian relay network (G, s, d),

$$\mathcal{C}(G, s, d) \ge \bar{\mathcal{C}}(G, s, d) - (\gamma\kappa + f(\kappa)). \tag{14}$$

It should be emphasized that a γ-GRNAT makes an assertion that is independent of network topology, channel SNRs, and so forth. In particular, Theorem 1 provides a concrete example of a 0.5-GRNAT, with the $f(\kappa)$ term being zero.

[1] Recall that the notation $f(n) = o(n)$ signifies that $\lim_{n \to \infty} \frac{f(n)}{n} = 0$.

Our first main result shows that improving the linear term 0.5κ in (9) to a sublinear term $o(\kappa)$ as suggested by (13) is equivalent to proving the cutset bound is tight for *all* Gaussian relay networks. Formally,

Theorem 2. *A* 0*-GRNAT exists iff* $\mathcal{C}(G,s,d) = \bar{\mathcal{C}}(G,s,d)$ *for* all *Gaussian relay networks* (G,s,d).

Hence, improving the slack in (9) to a sublinear function of κ presents a formidable (if not impossible) challenge. Since Theorem 2 asserts that the $\Theta(\kappa)$ term in approximations of the form (14) is fundamental, the following definition is well-motivated:

$$\gamma^\star = \inf\{\gamma : \text{a } \gamma\text{-GRNAT holds}\}. \tag{15}$$

In words, $\gamma^\star$ characterizes the best possible linear factor in (14). Clearly, Theorems 1 and 2 imply

$$0 \le \gamma^\star \le 0.5, \tag{16}$$

with $\gamma^\star = 0$ if and only if the cutset bound is tight for all Gaussian relay networks.

Much progress has already been made in proving upper bounds on $\gamma^\star$. In particular, Avestimehr, Diggavi and Tse pioneered the study of GRNATs, establishing[2] $\gamma^\star \le 7.5$ [2]. Subsequently, this was improved to $\gamma^\star \le 1.5$ by Ozgur and Diggavi in [3] using structured lattice codes. Lim et al. proved that $\gamma^\star \le 0.63$ in [4] using a noisy network coding scheme. Finally, as already noted in Theorem 1, Lim et al. have recently established $\gamma^\star \le 0.5$ in [6], which is presently the best known upper bound. Given the substantial increments in the improvements, it is conceivable that the upper bound can be further improved as new achievability schemes are proposed and analyzed. Hence, we turn our attention to the problem of establishing meaningful lower bounds on $\gamma^\star$. To this end, we note the following simple observation:

Theorem 3. *If* (G,s,d) *is a Gaussian relay network and* $\mathcal{C}(G,s,d) \le \bar{\mathcal{C}}(G,s,d) - \beta$*, then*

$$\gamma^\star \ge \frac{\beta}{\kappa(G,s,d)}. \tag{17}$$

Proof. The theorem is a consequence of the following proposition, which is proved in Section IV.

Proposition 1. *A* γ*-GRNAT implies that*

$$\mathcal{C}(G,s,d) \ge \bar{\mathcal{C}}(G,s,d) - \gamma\,\kappa(G,s,d) \tag{18}$$

for all Gaussian relay networks (G,s,d).

Now, (17) easily follows since Proposition 1 and the hypothesis of Theorem 3 imply that there is a Gaussian relay network (G,s,d) satisfying

$$\bar{\mathcal{C}}(G,s,d) - \beta \ge \mathcal{C}(G,s,d) \ge \bar{\mathcal{C}}(G,s,d) - \gamma^\star\kappa(G,s,d).$$

□

[2]The $1/2$-factor discrepancy is due to the consideration of complex channels in [2].

While the recent research trend has shifted toward studying general classes of Gaussian relay networks, Theorem 3 underscores the importance of continued investigation into upper bounds on capacity for specific 'small' networks with few nodes. For example, the capacity of the non-degraded Gaussian single-relay channel remains unknown, and the cutset bound is widely believed to not be tight, although counterexamples are only known for non-Gaussian networks (cf. [12], [13]). As a second example, the diamond network with two relays and one transmit/receive-antenna at each node introduced by Schein and Gallager [14], [15] remains unknown. Hypothetically speaking, if the lower bound $\mathcal{C}(G,s,d) \ge \bar{\mathcal{C}}(G,s,d) - 2$ for this network implied by (10) were tight for some particular choice of channel configurations, it would condense the interval (16) to the single point $\gamma^\star = 0.5$. This would be an exciting result since it would characterize a fundamental limit on how closely capacity can be approximated by the cutset bound in general Gaussian relay networks.

IV. Proofs

The claims in this paper boil down to a simple observation: we can exploit the fact that a γ-GRNAT is independent of network topology to show that the $o(\kappa)$ term can always be eliminated from (14). More formally, we show that the following restatement of Proposition 1 holds, from which Theorem 2 follows immediately:

Lemma 1. *Let* $f(n)$ *be a nonnegative function satisfying* $f(n) = o(n)$*, and let* $\gamma \ge 0$ *be a given constant. If, for any Gaussian relay network* (G,s,d)*, the capacity approximation*

$$\mathcal{C}(G,s,d) \ge \bar{\mathcal{C}}(G,s,d) - (\gamma\kappa + f(\kappa)) \tag{19}$$

holds, then we must also have

$$\mathcal{C}(G,s,d) \ge \bar{\mathcal{C}}(G,s,d) - \gamma\kappa. \tag{20}$$

Proof. Fix a network (G,s,d) and an integer $m \ge 1$. The idea behind the proof relies on the construction of a new network (G_m, s_0, d_0) that roughly consists of m parallel copies of (G,s,d), with the corresponding source nodes connected to a super-source s_0, and the corresponding destination nodes connected to a super-destination d_0. An example of this construction is illustrated in Figure 1 for the example of $m = 3$ and (G,s,d) being a three-terminal MIMO relay channel.

Formally, construct (G_m, s_0, d_0) from (G,s,d) as follows: For the network (G,s,d) with node set $\mathcal{N} = \{1,\dots,N\}$, recall that the channel gains are described by the collection of matrices $\{G_{i,j} : i,j \in \mathcal{N}\}$. The new network (G_m, s_0, d_0) will consist of mN relay nodes indexed by elements of the set $\{1,2,\dots,m\} \times \mathcal{N}$, plus a source node labeled s_0, and a destination node labeled d_0. We equip s_0 with one transmit antenna, and d_0 with one receive antenna.

For $i,j \in \mathcal{N}$ and $1 \le k,k' \le m$, the channel gain matrices for the network (G_m, s_0, d_0) are given by

$$G_{(k,i),(k',j)} = \begin{cases} G_{i,j} & \text{if } k = k' \\ 0 & \text{otherwise.} \end{cases} \tag{21}$$

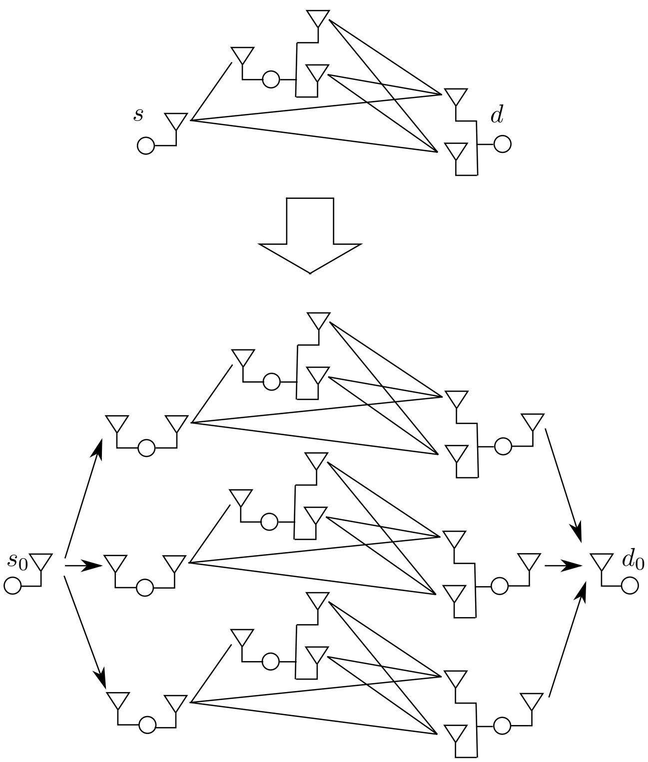

Fig. 1: Replication of a three-terminal MIMO relay channel for $m = 3$. Transmit and receive antennas are drawn to the right and left of nodes, respectively.

Recall that R_s and T_d denote the number of receive and transmit antennas at source s and destination d in (G, s, d), respectively. If $R_s = 0$, then we add a receive antenna to each node (k, s) for $1 \le k \le m$. Similarly, if $T_d = 0$, then we add a transmit antenna to each node (k, d) for $1 \le k \le m$. Since T_s and R_d must both be at least 1, our construction implies the relationship $\kappa(G_m, s_0, d_0) = m\,\kappa(G, s, d) + 2$.

Now, for $j \in \mathcal{N}$ and $1 \le k \le m$, we define the channel gain matrices:

$$G_{(k,j),s_0} = \begin{cases} \infty & \text{if } j = s \\ 0 & \text{otherwise,} \end{cases} \tag{22}$$

and

$$G_{d_0,(k,j)} = \begin{cases} \infty & \text{if } j = d \\ 0 & \text{otherwise,} \end{cases} \tag{23}$$

Finally, defining $G_{d_0,s_0} = 0$ completes the construction of (G_m, s_0, d_0).

Based on this construction, we claim that the following are true:

$$\kappa(G_m, s_0, d_0) = m\,\kappa(G, s, d) + 2 \tag{24}$$

$$\mathcal{C}(G_m, s_0, d_0) \le m\,\mathcal{C}(G, s, d) \tag{25}$$

$$\bar{\mathcal{C}}(G_m, s_0, d_0) \ge m\,\bar{\mathcal{C}}(G, s, d). \tag{26}$$

Indeed, (24) holds by definition of (G_m, s_0, d_0), so only (25) and (26) need further explanation. These inequalities are rather straightforward, but we argue them formally below for completeness. We remark that (25) and (26) can actually be shown to be equalities, but this is not needed in the sequel.

Claim 1: $\mathcal{C}(G_m, s_0, d_0) \le m\,\mathcal{C}(G, s, d)$.

Note that a $(2^{nR}, n)$-code for (G_m, s_0, d_0) with probability of error $P_e^{(n)}$ satisfies

$$n(R - P_e^{(n)}) \le H(M) - H(M|\hat{M}) = I(M; \hat{M}) \tag{27}$$

$$\le I(M; \mathbf{Y}^n_{(1,d)}, \dots \mathbf{Y}^n_{(m,d)}) \tag{28}$$

$$\le \sum_{j=1}^{m} I(M; \mathbf{Y}^n_{(j,d)}), \tag{29}$$

where (27) is Fano's inequality, (28) is the data processing inequality with $\mathbf{Y}^n_{(j,d)}$ abbreviating $\mathbf{Y}_{(j,d)}[1\colon n]$, and (29) follows by the chain rule and the fact that the $\mathbf{Y}^n_{(j,d)}$'s are conditionally independent given M by construction of the network. Now, we must have

$$\limsup_{n\to\infty} \frac{1}{n} I(M; \mathbf{Y}^n_{(j,d)}) \le \mathcal{C}(G, s, d), \tag{30}$$

else a rate exceeding $\mathcal{C}(G, s, d)$ would be possible in the original network (G, s, d) by designing a code for the channel $P_{\mathbf{Y}^n_{(j,d)}|M} : M \mapsto \mathbf{Y}^n_{(j,d)}$. Thus, by recalling (29), we have shown $\mathcal{C}(G_m, s_0, d_0) \le m\,\mathcal{C}(G, s, d)$ as desired. ♢

Claim 2: $\bar{\mathcal{C}}(G_m, s_0, d_0) \ge m\,\bar{\mathcal{C}}(G, s, d)$.

Let $\mathcal{N}_m$ denote the set of nodes for the network (G_m, s_0, d_0). In evaluating the cutset bound (8) for the network (G_m, s_0, d_0), any candidate cut (S, S^c) separating s_0 from d_0 should also satisfy $(k, s) \in S$ and $(k, d) \in S^c$ for all $1 \le k \le m$. If this is not the case, such a cut leads to the degenerate bound $\bar{\mathcal{C}}(G_m, s_0, d_0) \le \infty$.

Now, let the distribution $p^\star(\mathbf{x}_1, \dots, \mathbf{x}_N)$ achieve $\bar{\mathcal{C}}(G, s, d)$ in (8) for the network (G, s, d). Define the candidate distribution

$$\tilde{p}(\mathbf{x}_{s_0}, \mathbf{x}_{(1,1)}, \dots, \mathbf{x}_{(m,N)}) = q(\mathbf{x}_{s_0}) \prod_{k=1}^{m} p^\star(\mathbf{x}_{(k,1)}, \dots, \mathbf{x}_{(k,N)}),$$

where $q(\cdot)$ is an arbitrary probability measure on $\mathbb{R}$. From the definition of $\tilde{p}$, we have the inequality

$$\begin{aligned} \sup_{p(\mathbf{x}_{s_0}, \mathbf{x}_{(1,1)}, \dots, \mathbf{x}_{(m,N)})} & \min_{S \subseteq \mathcal{N}_m : s_0 \in S, d_0 \in S^c} I(\mathbf{X}_S; \mathbf{Y}_{S^c} | \mathbf{X}_{S^c}) \\ &\ge \min_{S \subseteq \mathcal{N}_m : s_0 \in S, d_0 \in S^c} I_{\tilde{p}}(\mathbf{X}_S; \mathbf{Y}_{S^c} | \mathbf{X}_{S^c}) \\ &= \sum_{k=1}^{m} \min_{\substack{S_k \subseteq \{k\} \times \mathcal{N}: \\ (k,s) \in S_k, (k,d) \in S_k^c}} I_{\tilde{p}}(\mathbf{X}_{S_k}; \mathbf{Y}_{S_k^c} | \mathbf{X}_{S_k^c}) \\ &= m\,\bar{\mathcal{C}}(G, s, d). \end{aligned}$$

Hence, the claim is proved. ♢

Now, with (24)-(26) in hand, suppose the approximation (19) holds for any Gaussian relay network. Then it must also hold for (G_m, s_0, d_0) implying

$$m\,\mathcal{C}(G,s,d) \geq \mathcal{C}(G_m, s_0, d_0) \tag{31}$$
$$\geq \bar{\mathcal{C}}(G_m, s_0, d_0) - (\gamma\kappa_m + f(\kappa_m)) \tag{32}$$
$$\geq m\,\bar{\mathcal{C}}(G,s,d) - \Big(\gamma\,(m\,\kappa+2) + f\big(m\,\kappa+2\big)\Big), \tag{33}$$

where $\kappa \equiv \kappa(G,s,d)$ and $\kappa_m \equiv \kappa(G_m, s_0, d_0)$ were defined for convenience. Dividing both sides by m, we have

$$\mathcal{C}(G,s,d) \geq \bar{\mathcal{C}}(G,s,d) - \Big(\gamma\,\kappa\left(1+\frac{2}{\kappa\,m}\right) + \frac{f(m\,\kappa+2)}{m}\Big).$$

Since m was arbitrary and $f(n) = o(n)$, we can conclude that

$$\mathcal{C}(G,s,d) \geq \bar{\mathcal{C}}(G,s,d) - \gamma\,\kappa \tag{34}$$

as desired. □

Remark: The described transformation from (G, s, d) to (G_m, s_0, d_0) does not rely on the Gaussian nature of (G, s, d) in any particular way. The crucial point needed here is that a γ-GRNAT provides an approximation that holds universally across the class of Gaussian relay networks, independent of network topology and channel configurations. By exploiting this generality, it is possible to obtain an apparently tighter capacity approximation for the network (G, s, d) by first transforming to another related Gaussian relay network (G_m, s_0, d_0), and then applying the γ-GRNAT.

V. Concluding Remarks

For the class of Gaussian relay networks, recent papers [2]–[6] have shown that the gap between the cutset bound and capacity can be bounded by a function that grows linearly in $\kappa(G, s, d)$, and is otherwise independent of network topology and channel configurations. Motivated by specific examples where this can be improved to a sublinear or constant function, Niesen and Diggavi [8] have suggested the problem of tightening the approximation gap beyond linear in the general case. In this paper, we argue that a linear gap is fundamental to such approximation results which apply universally to the class of Gaussian relay networks. Specifically, if the gap between the cutset bound and capacity of a Gaussian relay network is uniformly bounded by a function that grows sublinearly in $\kappa(G, s, d)$, then the cutset bound must be tight for all Gaussian relay networks.

We have focused our treatment on approximations with respect to the parameter $\kappa(G, s, d)$ since this is the prevailing quantity in published capacity approximation results (e.g., [2]–[6]). However, it is straightforward to see that our reduction described in Section IV applies more generally. For instance, if $\eta(G, s, d)$ is a functional defined on the set of Gaussian networks which satisfies

$$\lim_{m\to\infty} \frac{m\,\eta(G,s,d)}{\eta(G_m,s_0,d_0)} = 1 \tag{35}$$

for any Gaussian relay network (G, s, d), then Lemma 1 generalizes as follows.

Lemma 1'. *Let $f(n)$ be a nonnegative function satisfying $f(n) = o(n)$, and let $\gamma \geq 0$ be a given constant. If, for any Gaussian relay network (G, s, d), the capacity approximation*

$$\mathcal{C}(G,s,d) \geq \bar{\mathcal{C}}(G,s,d) - (\gamma\,\eta(G,s,d) + f(\eta(G,s,d)))$$

holds, then we must also have

$$\mathcal{C}(G,s,d) \geq \bar{\mathcal{C}}(G,s,d) - \gamma\,\eta(G,s,d).$$

To give a more concrete example along these lines, Kolte, Özgür and El Gamal have shown in [10] that

$$\mathcal{C}(G,s,d) \geq \bar{\mathcal{C}}(G,s,d) - d_0^*\log\left(1+\frac{\kappa(G,s,d)}{d_0^*}\right) - \frac{\kappa(G,s,d)}{Q} - d_Q^*\log(1+Q), \tag{36}$$

for all $Q \geq 0$, where $d_Q^* \equiv d_Q^*(G, s, d)$ denotes the degrees of freedom of the MIMO channel corresponding to the min-cut of the network (G, s, d) when transmit power is scaled as $P \mapsto \frac{P}{1+Q}$ (see [10] for a precise definition and details).

By appealing to the definition of d_Q^*, it can be shown that $d_Q^*(G_m, s_0, d_0) = m\,d_Q^*(G, s, d)$. Thus, we can conclude that the gap in (36) cannot be narrowed to a function that grows sublinearly in $d_Q^*(G, s, d)$ for all $Q \geq 0$ unless $\mathcal{C}(G, s, d) = \bar{\mathcal{C}}(G, s, d)$ for all Gaussian relay networks.

References

[1] T. Cover and A. El Gamal, "Capacity theorems for the relay channel," vol. 25, no. 5, pp. 572–584, Sep 1979.

[2] A. Avestimehr, S. Diggavi, and D. Tse, "Wireless network information flow: A deterministic approach," *Information Theory, IEEE Transactions on*, vol. 57, no. 4, pp. 1872–1905, April 2011.

[3] A. Ozgur and S. Diggavi, "Approximately achieving Gaussian relay network capacity with lattice-based QMF codes," *Information Theory, IEEE Transactions on*, vol. 59, no. 12, pp. 8275–8294, Dec 2013.

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

[5] A. Raja and P. Viswanath, "Compress-and-forward scheme for a relay network: Approximate optimality and connection to algebraic flows," in *Information Theory Proceedings (ISIT), 2011 IEEE International Symposium on*, July 2011, pp. 1698–1702.

[6] S. H. Lim, K. T. Kim, and Y.-H. Kim, "Distributed decode-forward for multicast," in *IEEE Int. Symp. Inf. Theory*, 2014, pp. 636–640.

[7] B. Chern and A. Ozgur, "Achieving the capacity of the N-relay Gaussian diamond network within $\log n$ bits," in *Information Theory Workshop (ITW), 2012 IEEE*, Sept 2012, pp. 377–380.

[8] U. Niesen and S. Diggavi, "The approximate capacity of the Gaussian N-relay diamond network," *Information Theory, IEEE Transactions on*, vol. 59, no. 2, pp. 845–859, Feb 2013.

[9] U. Niesen, B. Nazer, and P. Whiting, "Computation alignment: Capacity approximation without noise accumulation," vol. 59, no. 6, pp. 3811–3832, 2013.

[10] R. Kolte, A. Özgür, and A. E. Gamal, "Capacity approximations for Gaussian relay networks," *CoRR*, vol. abs/1407.3841, 2014. [Online]. Available: http://arxiv.org/abs/1407.3841

[11] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. John Wiley & sons, 1991.

[12] Z. Zhang, "Partial converse for a relay channel," *Information Theory, IEEE Transactions on*, vol. 34, no. 5, pp. 1106–1110, 1988.

[13] M. Aleksic, P. Razaghi, and W. Yu, "Capacity of a class of modulo-sum relay channels," *Information Theory, IEEE Transactions on*, vol. 55, no. 3, pp. 921–930, 2009.

[14] B. Schein and R. Gallager, "The Gaussian parallel relay network," in *Information Theory, 2000. Proceedings. IEEE International Symposium on*, 2000, p. 22.

[15] B. E. Schein, "Distributed coordination in network information theory," Ph.D. dissertation, Massachusetts Institute of Technology, 2001.