Capacity Bounds for Relay Networks

Gerhard Kramer
Communications and Statistical Sciences Dept.
Bell Labs, Lucent Technologies
600 Mountain Ave., Murray Hill, NJ 07974 USA
Email: gkr@bell-labs.com

Serap A. Savari
Dept. of Electrical Engineering and Computer Science
University of Michigan
Ann Arbor, MI 48109 USA
Email: savari@eecs.umich.edu

***Abstract*— Recently developed PdE outer bounds on the capacity region of classical networks are extended to networks with broadcasting, interference, and noise. Examples demonstrate that the new bounds improve on cut-set bounds.**

I. Introduction

Consider a directed graph $\mathcal{N} = (\mathcal{V}, \mathcal{E})$ with vertex set $\mathcal{V}$ and edge set $\mathcal{E}$. We interpret this graph in progressively more general ways.

- Classical networks: Every vertex u represents a processor that transmits a different $x_{u,v}$ into each of its outgoing edges (u, v). The edges represent channels, and the channel (u, v) has capacity $C_{u,v}$ and is usually taken to be noise-free. The channels are "orthogonal" in the sense that there is no interference between them.
- Aref networks: Every vertex u broadcasts a common x_u into its outgoing edges. The output $y_{u,v}$ of edge (u, v) is a deterministic function $h_{u,v}(x_u)$ of x_u. This type of network is also known as a *deterministic relay network with no interference* [2], [5].
- Relay networks with no interference and with independent edge noise: Similar to Aref networks except that the output $y_{u,v}$ is some noisy function $h_{u,v}(x_u, z_{u,v})$ where $z_{u,v}$ is a realization of the noise random variable $Z_{u,v}$. The $Z_{u,v}$ are independent across edges.
- Relay networks with independent vertex noise: Every vertex u broadcasts a common x_u into its outgoing edges. Let $\mathcal{E}_u$ be the set of edges terminated by vertex u and let $\mathcal{V}(\mathcal{E}_u)$ be the set of vertices from which these edges emanate. Vertex u receives a common output y_u that is some function $h_u(x_{\mathcal{V}(\mathcal{E}_u)}, z_u)$ where z_u is a realization of the noise random variable Z_u. The Z_u are independent across vertices.

The purpose of this paper is to generalize recently developed *progressive d-separating edge* set bounds (or PdE bounds [4]) for classical networks to all of the above networks. This paper is organized as follows. In Sec. II, we review existing PdE bounds. In Sec. III, we describe how these bounds can generalize to other networks. Sec. IV concludes the paper.

II. Preliminaries

We use the network model described in [4, Sec. 3] where the network is governed by a clock that ticks N times. We write $X^N = X^{(1)}, X^{(2)}, \ldots, X^{(N)}$. The noisy classical network has K messages W_k, $k = 1, 2, \ldots, K$, $|\mathcal{E}|$ channel input sequences $X_{u,v}^N$, $|\mathcal{E}|$ channel output sequences $Y_{u,v}^N$, $|\mathcal{E}|$ noise sequences $Z_{u,v}^N$, and D_k channel estimates $\hat{W}_k^{(i)}$, $i = 1, 2, \ldots, D_k$, for $k = 1, 2, \ldots, K$. The message W_k has rate R_k bits per clock tick, it originates at source vertex s_k, and it is destined for the sink vertices $\{t_k(1), t_k(2), \ldots, t_k(D_k)\}$ where $t_k(i) \neq s_k$ for all i. The message estimate $\hat{W}_k^{(i)}$ is generated at vertex $t_k(i)$.

We review the PdE bound for classical networks [4, Sec. 3]. This bound begins with:

- a set of edges $\mathcal{E}_d$
- a set of sources $\mathcal{S}_d$
- an ordering of the source indices in $\mathcal{S}_d$ via a one-to-one mapping $\pi(\cdot)$ from $\{1, 2, \ldots, |\mathcal{S}_d|\}$ to $\mathcal{S}_d$, where $|\mathcal{S}_d|$ is the cardinality of $\mathcal{S}_d$.

We use the notation $X_{\mathcal{E}_d} = \{X_{u,v} : (u, v) \in \mathcal{E}_d\}$, and similarly for $Y_{\mathcal{E}_d}$ and $Z_{\mathcal{E}_d}$. The PdE bound was developed in [4] by using a functional dependence graph (FDG) $\mathcal{G}$ that is a modified version of $\mathcal{N}$. Let $\mathcal{S}_d^C$ be the complement of $\mathcal{S}_d$ in $\{1, 2, \ldots, K\}$ and let $\mathcal{E}_d^C$ be the complement of $\mathcal{E}_d$ in $\mathcal{E}$. The following describes the PdE bound.

1) (Initialization) Consider the FDG $\mathcal{G}$ for $\mathcal{N}$ that has vertices and edges representing the channel input, noise, and output sequences, the messages and their estimates.
   - Remove all vertices and edges in $\mathcal{G}$ except those encountered when moving backward one or more edges starting from any of the vertices representing: (1) $Y_{\mathcal{E}_d}^N$ and (2) some choice of non-empty subset of $\{\hat{W}_k^{(i)} : i = 1, 2, \ldots, D_k\}$ for all $k \in \mathcal{S}_d$.
   - Further remove the edges coming out of the vertices representing $Z_{\mathcal{E}_d^C}^N$ and $W_{\mathcal{S}_d^C}$, and successively remove edges coming out of vertices and on cycles that have no incoming edges, excepting source vertices. Call the resulting graph $\mathcal{G}_{\mathcal{E}_d}$. Set $k = 1$.
2) (Iterations) If $W_{\pi(k)}$ is not disconnected (in an undirected sense) from one of its estimates $\hat{W}_{\pi(k)}^{(i)}$, $i = 1, 2, \ldots, D_k$, then stop (one has no bound). If $W_{\pi(k)}$ is disconnected (in an undirected sense) from all of its estimates then:
   - Remove the edges coming out of the vertex representing $W_{\pi(k)}$.

Presented at the Workshop on Information Theory & Applications, UCSD Campus, La Holla, CA, USA, Feb. 6-10, 2006

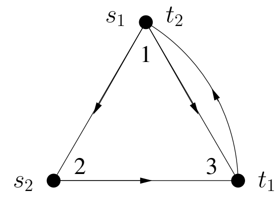

Fig. 1. A network graph $\mathcal{N}$.

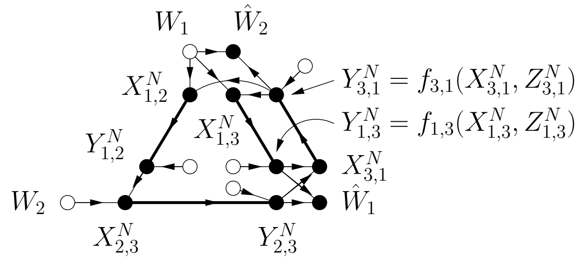

Fig. 2. An FDG for a classical network with the graph in Fig. 1.

- Successively remove edges coming out of vertices and on cycles that have no incoming edges, excepting source vertices. Call the resulting graph $\mathcal{G}_{\mathcal{E}_d W_\pi^k}$.

3) (Termination and Bound) Increment k. If $k \leq K$ go to the previous step. If $k = K + 1$, then we have

$$\sum_{k \in \mathcal{S}_d} R_k \leq \sum_{e \in \mathcal{E}_d} C_e. \tag{1}$$

For example, consider the network depicted in Fig. 1 that has two source-destination pairs. The FDG $\mathcal{G}$ for this network with noisy channels is shown in Fig. 2 (we have not labeled the four noise vertices). The cut-set bound gives

$$R_1 \leq \min\{C_{1,2} + C_{1,3},\, C_{1,3} + C_{2,3}\} \tag{2}$$

$$R_2 \leq \min\{C_{2,3},\, C_{3,1}\}. \tag{3}$$

For the PdE bound, consider the choice $\mathcal{E}_d = \{(1,3),(2,3)\}$, $\mathcal{S}_d = \{1,2\}$, and $[\pi(1), \pi(2)] = [1,2]$. We find that

$$R_1 + R_2 \leq C_{1,3} + C_{2,3} \tag{4}$$

which can be stronger than (2) and (3). It was known that (4) applies to routing, but it applies to network coding [1] as well (see [4]). We remark that the cut-set bounds (2) and (3) are also PdE bounds with different choices of $\mathcal{S}_d$, $\mathcal{E}_d$, and $\pi(\cdot)$.

III. Application to Other Networks

A. Aref Networks

Our first extension of the PdE bound is for Aref networks. Let $\mathcal{V}(\mathcal{E}_d)$ and $\bar{\mathcal{V}}(\mathcal{E}_d)$ be the respective sets of vertices from which the edges $\mathcal{E}_d$ emanate and terminate in $\mathcal{N}$ (note that $\mathcal{V}(\mathcal{E}_d)$ and $\bar{\mathcal{V}}(\mathcal{E}_d)$ are not necessarily disjoint). Let $\bar{\mathcal{V}}_u(\mathcal{E}_d)$ be the set of vertices that terminate the edges in $\mathcal{E}_d$ that emanate from u. The procedure described above remains the same, but rather than (1) we now have the bound

$$\sum_{k \in \mathcal{S}_d} R_k \leq \sum_{u \in \mathcal{V}(\mathcal{E}_d)} H(Y_{u,\bar{\mathcal{V}}_u(\mathcal{E}_d)}). \tag{5}$$

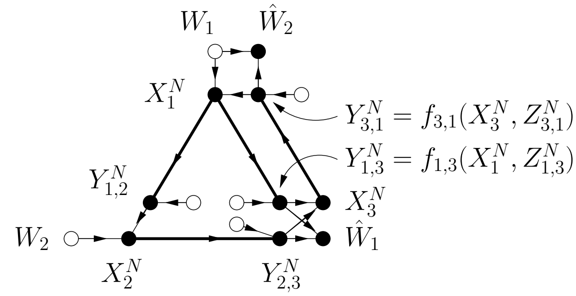

Fig. 3. An FDG for an Aref network with the graph in Fig. 1.

for some choice of independent X_u, $u \in \mathcal{V}$. For example, the cut-set bound for the graph in Fig. 1 gives

$$R_1 \leq \min\{H(Y_{1,2}Y_{1,3}),\, H(Y_{1,3}) + H(Y_{2,3})\} \tag{6}$$

$$R_2 \leq \min\{H(Y_{2,3}),\, H(Y_{3,1})\}. \tag{7}$$

For the PdE bound, consider the FDG in Fig. 3 when the network is noise-free. For $\mathcal{E}_d = \{(1,3),(2,3)\}$, we have $\mathcal{V}(\mathcal{E}_d) = \{1,2\}$ and $\bar{\mathcal{V}}(\mathcal{E}_d) = \{3\}$. The bound (5) thus gives

$$R_1 + R_2 \leq H(Y_{1,3}) + H(Y_{2,3}). \tag{8}$$

We remark that (8) includes (4) as a special case if the Aref network happens to be a noise free classical network [5].

B. Relay Networks with No Interference and with Independent Edge Noise

Consider next relay networks with no interference, like Aref networks, but where every edge has independent noise. These networks model problems where only one transmitter uses a time and/or frequency slot at once. The procedure described above remains the same, but rather than (1) we now have

$$\sum_{k \in \mathcal{S}_d} R_k \leq \sum_{u \in \mathcal{V}(\mathcal{E}_d)} I(X_u; Y_{u,\bar{\mathcal{V}}_u(\mathcal{E}_d)}). \tag{9}$$

for some choice of independent X_u, $u \in \mathcal{V}$. For example, we can replace (8) with

$$R_1 + R_2 \leq I(X_1; Y_{1,3}) + I(X_2; Y_{2,3}). \tag{10}$$

C. Relay Networks with Independent Vertex Noise

Consider next relay networks *with* interference but where every *vertex* has independent noise. Let $s(\mathcal{S}_d) = \{s_k : k \in \mathcal{S}_d\}$ and let $s(\mathcal{S}_d)^C$ be the complement of $s(\mathcal{S}_d)^C$ in $\mathcal{V}$. The procedure described above changes because we require that $\mathcal{E}_d$ have the property that if $(u,v) \in \mathcal{E}_d$ then $(w,v) \in \mathcal{E}_d$ for all w with $(w,v) \in \mathcal{E}$. We must further replace $Z_{\mathcal{E}_d^C}$ with $Z_{\bar{\mathcal{V}}(\mathcal{E}_d^C)}$. With these changes, we have the bound

$$\sum_{k \in \mathcal{S}_d} R_k \leq I(X_{\mathcal{V}(\mathcal{E}_d)}; Y_{\bar{\mathcal{V}}(\mathcal{E}_d)} \,|\, X_{\bar{\mathcal{V}}(\mathcal{E}_d) \cap s(\mathcal{S}_d)^C}) \tag{11}$$

for some choice of jointly distributed X_u, $u \in \mathcal{V}$. For example, the cut-set bound for the graph in Fig. 1 gives

$$R_1 \leq \min\{I(X_1; Y_2Y_3|X_2X_3),\, I(X_1X_2; Y_3|X_3)\} \tag{12}$$

$$R_2 \leq \min\{I(X_2; Y_3|X_1X_3),\, I(X_3; Y_1|X_1)\}. \tag{13}$$

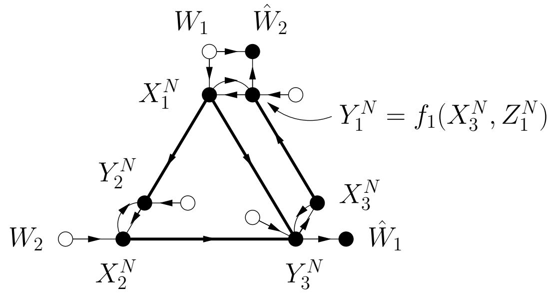

Fig. 4. An FDG for a relay network with the graph in Fig. 1 when there is independent vertex noise.

For the PdE bound, consider the FDG in Fig. 4 and note that $\mathcal{E}_d = \{(1,3),(2,3)\}$ is permitted. Note further that vertex v's transmit symbols X_v^N can interfere with its own received symbols Y_v^N, as depicted by the functional dependence of Y_v^N on X_v^N for all v. We thus have $\mathcal{V}(\mathcal{E}_d) = \{1,2,3\}$ and $\bar{\mathcal{V}}(\mathcal{E}_d) = \{3\}$. We further have $s(\mathcal{S}_d) = \{1,2\}$ and $s(\mathcal{S}_d)^C = \{3\}$ so that (11) gives

$$R_1 + R_2 \le I(X_1X_2; Y_3 \,|\, X_3). \tag{14}$$

IV. Concluding Remarks

We have generalized the applicability of PdE bounds to networks other than classical networks. We remark that extensions to general relay networks are possible but then seem to give only cut-set bounds. It would be useful to have other approaches for applying the PdE concept to general memoryless relay networks.

Acknowledgment

The work of G. Kramer was partially supported by the Board of Trustees of the University of Illinois Subaward No. 04-217 under NSF Grant No. CCR-0325673. The work of S. A. Savari was supported by NSF Grant No. CCF-0430201.

Appendix

We prove the validity of the PdE bound for relay networks with independent vertex noise. The proofs for the other two networks are similar. Recall that W_k is associated with the vertices $(s_k, t_k(1), t_k(2), \ldots, t_k(D_k))$, and that we are considering some subset $\hat{\mathcal{V}}_k$ of the vertices $t_k(i)$, $i = 1,2,\ldots,D_k$. We write the corresponding subset of estimates as $\hat{W}_k(\hat{\mathcal{V}}_k)$. Similarly, we denote the incoming edges of the vertices in $\hat{\mathcal{V}}_k$ by $\hat{\mathcal{E}}_k$.

The first steps are the same as in [4, Appendix]. For reliable communication, Fano's inequality [3, p. 39] requires that

$$\sum_{k\in\mathcal{S}_d} R_k \le \sum_{k\in\mathcal{S}_d} \frac{1}{N}\, I(W_k; \hat{W}_k(\hat{\mathcal{V}}_k)) = \sum_{k=1}^{|\mathcal{S}_d|} \frac{1}{N}\, I(\,W_{\pi(k)}; \hat{W}_{\pi(k)}(\hat{\mathcal{V}}_{\pi(k)})\,). \tag{15}$$

We define $W_\pi^{k-1} = [W_{\pi(1)}, W_{\pi(2)}, \ldots, W_{\pi(k-1)}]$ and bound

$$\begin{aligned} I\left(W_{\pi(k)}; \hat{W}_{\pi(k)}(\hat{\mathcal{V}}_{\pi(k)})\right) &\le I(W_{\pi(k)}; \hat{W}_{\pi(k)}(\hat{\mathcal{V}}_{\pi(k)})\; Y_{\mathcal{E}_d}^N Z_{\mathcal{E}_d^C}^N\; W_{\mathcal{S}_d^C} W_\pi^{k-1}) \\ &= I(W_{\pi(k)}; \hat{W}_{\pi(k)}(\hat{\mathcal{V}}_{\pi(k)})\; Y_{\mathcal{E}_d}^N \,|\, Z_{\mathcal{E}_d^C}^N W_{\mathcal{S}_d^C} W_\pi^{k-1}) \\ &= I(W_{\pi(k)}; Y_{\mathcal{E}_d}^N \,|\, Z_{\mathcal{E}_d^C}^N W_{\mathcal{S}_d^C} W_\pi^{k-1}) \end{aligned} \tag{16}$$

where the last step follows because success in step 2) of the PdE procedure implies that

$$I(W_{\pi(k)}; \hat{W}_{\pi(k)}(\hat{\mathcal{V}}_{\pi(k)}) \,|\, Y_{\mathcal{E}_d}^N Z_{\mathcal{E}_d^C}^N\; W_{\mathcal{S}_d^C} W_\pi^{k-1}) = 0 \tag{17}$$

via fd-separation. Inserting (16) into (15), and applying the chain rule for mutual information, we have

$$\sum_{k\in\mathcal{S}_d} R_k \le \frac{1}{N}\, I(W_{\mathcal{S}_d}; Y_{\mathcal{E}_d}^N \,|\, Z_{\mathcal{E}_d^C}^N W_{\mathcal{S}_d^C}). \tag{18}$$

We continue as follows:

$$\begin{aligned} &I(W_{\mathcal{S}_d}; Y_{\mathcal{E}_d}^N \,|\, Z_{\mathcal{E}_d^C}^N W_{\mathcal{S}_d^C}) \\ &= \sum_{n=1}^N I(W_{\mathcal{S}_d}; Y_{\mathcal{E}_d}^{(n)} \,|\, Y_{\mathcal{E}_d}^{n-1} Z_{\mathcal{E}_d^C}^N W_{\mathcal{S}_d^C}) \\ &\overset{(a)}{=} \sum_{n=1}^N I(W_{\mathcal{S}_d}; Y_{\mathcal{E}_d}^{(n)} \,|\, Y_{\mathcal{E}_d}^{n-1} Z_{\mathcal{E}_d^C}^N W_{\mathcal{S}_d^C}\; X^n_{\bar{\mathcal{V}}(\mathcal{E}_d)\cap s(\mathcal{S}_d)^C}) \\ &\le \sum_{n=1}^N I(W_{\mathcal{S}_d} X^{(n)}_{\mathcal{V}(\mathcal{E}_d)}; Y_{\mathcal{E}_d}^{(n)} \,|\, Y_{\mathcal{E}_d}^{n-1} Z_{\mathcal{E}_d^C}^N W_{\mathcal{S}_d^C}\; X^n_{\bar{\mathcal{V}}(\mathcal{E}_d)\cap s(\mathcal{S}_d)^C}) \\ &\overset{(b)}{\le} \sum_{n=1}^N \left[H(Y_{\mathcal{E}_d}^{(n)} \,|\, X^{(n)}_{\bar{\mathcal{V}}(\mathcal{E}_d)\cap s(\mathcal{S}_d)^C}) - H(Y_{\mathcal{E}_d}^{(n)} \,|\, X^{(n)}_{\mathcal{V}(\mathcal{E}_d)}) \right] \\ &= \sum_{n=1}^N I(X^{(n)}_{\mathcal{V}(\mathcal{E}_d)}; Y_{\mathcal{E}_d}^{(n)} \,|\, X^{(n)}_{\bar{\mathcal{V}}(\mathcal{E}_d)\cap s(\mathcal{S}_d)^C}) \\ &\overset{(c)}{=} N \cdot I(X^{(Q)}_{\mathcal{V}(\mathcal{E}_d)}; Y_{\mathcal{E}_d}^{(Q)} \,|\, X^{(Q)}_{\bar{\mathcal{V}}(\mathcal{E}_d)\cap s(\mathcal{S}_d)^C}\; Q) \\ &\le N \cdot I(X^{(Q)}_{\mathcal{V}(\mathcal{E}_d)}; Y_{\mathcal{E}_d}^{(Q)} \,|\, X^{(Q)}_{\bar{\mathcal{V}}(\mathcal{E}_d)\cap s(\mathcal{S}_d)^C}) \end{aligned} \tag{19}$$

where (a) follows because $X^n_{\bar{\mathcal{V}}(\mathcal{E}_d)\cap s(\mathcal{S}_d)^C}$ is a function of $Y_{\mathcal{E}_d}^{n-1}$ and $W_{\mathcal{S}_d^C}$, (b) follows because $Y_{\mathcal{E}_d}^{(n)}$ is a function of $X^{(n)}_{\mathcal{V}(\mathcal{E}_d)}$ and noise, and (c) follows by choosing Q to be uniform over $\{1,2,\ldots,n\}$. We can now use standard arguments (see [3, Sec. 14.10]) to show that there is a joint distribution on X_u, $u \in \mathcal{V}$, such that (11) is satisfied for all choices of $\mathcal{E}_d$, $\mathcal{S}_d$, and $\pi(\cdot)$ for which the PdE procedure gives a rate bound.

References

[1] R. Ahlswede, N. Cai, S.-Y .R. Li, and R. W. Yeung, "Network information flow," *IEEE Trans. Inf. Theory*, vol. 46, pp. 1204–1216, July 2000.
[2] M. R. Aref, "Information flow in relay networks," Ph.D. dissertation, Stanford Univ., Stanford, CA, Oct. 1980.
[3] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, New York: Wiley, 1991.
[4] G. Kramer and S. A. Savari, "Edge-cut bounds on network coding rates," *J. Network and Systems Management*, to appear March 2006.
[5] N. Ratnakar and G. Kramer, "The multicast capacity of deterministic relay networks with no interference," *IEEE Trans. Inf. Theory*, to appear June 2006.