The Common Randomness Capacity of a Finite Network of Channels

Sivarama Venkatesan
School of Electrical Engineering
Cornell University
Ithaca, NY 14853
svenkat@eclair.eecs.berkeley.edu

Venkat Anantharam
EECS Department
University of California
Berkeley, CA 94720
ananth@vyasa.eecs.berkeley.edu

Abstract

Consider a finite number of agents interconnected by an arbitrary network of independent, point-to-point, discrete memoryless channels. The agents wish to generate common randomness by interactive communication over the network. Our main result is an exact characterization of the common randomness capacity of such a network, i.e. the maximum number of bits of randomness that all the agents can agree on, per step of communication. As a by-product, we also obtain a description by linear inequalities of the blocking-type polyhedron whose extreme points are precisely the incidence vectors of all arborescences in a digraph, with a prescribed root of out-degree 1.

1. Introduction

Two or more communicating agents are said to have *common randomness* if there is a random variable, e.g., a random bit string, whose value is known to all of them. This notion of shared randomness turns out to be of significance in many problems of information theory. For example, common randomness available to a transmitter and receiver allows them to use random codes for data transmission, which can outperform deterministic codes in certain situations, e.g., with arbitrarily varying channels [1], [6]. In the theory of identification over noisy channels [4], [3], [5], the maximum achievable identification rate is essentially determined by the amount of common randomness that the transmitter and receiver can set up. Common randomness can also significantly reduce the communication complexity of certain distributed computations [7], [9]. Finally, *secret* common randomness available to a transmitter and receiver allows them to communicate securely over a channel with eavesdroppers [2], [8].

An important observation is the possibility of exploiting *channel noise* to generate common randomness. To see how channel noise could actually be useful in this context, consider a situation where there are two agents A and B connected to each other in both directions by a pair of channels. As an extreme case, assume that neither agent has access to any external sources of randomness (such as a random bit generator). Even then, A and B may be able to generate common randomness! The intuition is this: suppose A transmits some known input sequence to B. If the channel from A to B is noisy, then the resulting output sequence seen by B will be random. Since the input sequence is known, B could somehow "cancel" out its effect on the output sequence, and extract the randomness due to noise. Now, if the channel from B to A has positive Shannon capacity, then B could reliably convey the randomness thus obtained to A, using suitable encoding techniques; A and B would then have *common* randomness.

A natural question that arises now is: what is the maximum *rate*, in bits per step of communication, at which the two agents can generate common randomness this way from the noise on the two channels, i.e., what is the common randomness *capacity* of the given pair of channels? This question was posed and answered in [11], under the assumption that the two channels are independently operating discrete memoryless channels (DMC's). The main result of [11] is that A and B can generate

$$\max_{X_A, X_B} \quad \{\min\left[H(Y_B|X_B), I(X_A;Y_A)\right] + \min\left[H(Y_A|X_A), I(X_B;Y_B)\right]\} \quad (1)$$

bits of common randomness per step of communication over such a pair of channels. In (1), (X_A, Y_A) and (X_B, Y_B) are random input-output pairs for the A-to-B and B-to-A DMC's, respectively, and the maximum is over all possible distributions for the inputs X_A and X_B. The rate in (1) was shown to be optimal with a "strong" converse, which was proved by developing a novel typical sequence machinery for interactive communication. Further, it was shown that the common randomness capacity in the presence of independent, discrete memoryless sources of external randomness at the two terminals could also be derived from

(1).

In the special case where both channels are binary symmetric with crossover probabilities p and q, respectively, the expression in (1) reduces to $\min\{h(p) + h(q), 2 - h(p) - h(q)\}$, where $h(\cdot)$ is the binary entropy function. Observe that this is 0 if and only if either $h(p) = h(q) = 0$ (no randomness on either channel) or $h(p) = h(q) = 1$ (plenty of randomness but no ability to agree); moreover the capacity equals its maximum of 1 bit per step when $h(p) = h(q) = 1/2$. For details and other examples, see [11].

In this document we consider a network of agents, represented by a digraph $G = (V, E)$. The vertex set V of this digraph is just the set of agents, and the edge set $E \subseteq V \times V$ describes their interconnections — $(u, v) \in E$ means there is a channel whose input is controlled by u, and whose output is seen by v. As in [11], we assume that these channels are all DMC's, and that they operate independently. Communication occurs simultaneously on all the channels, and in synchronism (i.e., there is a common clock).

As in the two-agent case, there are basically two steps here in the process of generating common randomness. The first step is for the agents to bring channel noise into play by communicating over the channels, so that each agent can then extract randomness from the observed channel outputs. The second step is for each agent to convey reliably the randomness thus obtained to each of the other agents, using suitable encoding techniques. However, important new features show up in the general problem that are not present in the problem for two agents. In the two-agent case, there is only one path along which an agent can deliver randomness to the other, and this path consists of a single channel. Also, the flows of randomness originating from the two agents do not interact — each is confined to a different channel. In contrast, in an arbitrary network of channels, there could be several paths from one agent to each of the others, many of these consisting of more than one channel, and all these paths could be used simultaneously to deliver randomness. Moreover, a given channel could be on several different paths, which means that the flows of randomness from different agents must interact. For these reasons, the problem of optimally disseminating randomness from each agent to all the others is quite non-trivial in the general situation. In fact, the solution to this problem leads to some purely combinatorial results about blocking polyhedra, which are of independent interest.

2. Problem Formulation

The following conventions will be in effect throughout the remainder of this document : all logarithms and exponentials will be to the base two. If N is a positive integer, then $[N] \stackrel{\text{def}}{=} \{1, 2, \ldots, N\}$. $\lfloor z \rfloor$ will denote the largest integer not exceeding z. The standard sequence notation $\mathbf{z}^k = (z_1, z_2, \ldots, z_k)$ will be employed. If S is a finite set, then $(z_s : s \in S)$ will mean a vector whose components are indexed by the elements of S. $\mathbf{R}_+$ will denote the set of all non-negative real numbers, and $\mathbf{R}_+^S$ the set of all vectors $(z_s \in \mathbf{R}_+ : s \in S)$.

If $\mathbf{Q} = \big(Q(y|x) : (x, y) \in \mathcal{X} \times \mathcal{Y}\big)$ is the matrix of transition probabilities of a discrete memoryless channel (DMC) with input alphabet $\mathcal{X}$ and output alphabet $\mathcal{Y}$, P is a probability distribution on $\mathcal{X}$, and X and Y are random variables with the joint distribution $Pr\{X = x, Y = y\} = P(x)Q(y|x)$, then we will also write $H(\mathbf{Q}|P)$ and $I(P; \mathbf{Q})$ for $H(Y|X)$ and $I(X; Y)$, respectively.

As mentioned earlier, the network of DMC's connecting the agents will be represented by the digraph $G = (V, E)$. The DMC corresponding to the edge $e \in E$ will be assumed to have finite input alphabet $\mathcal{X}_e$, finite output alphabet $\mathcal{Y}_e$, and transition probabilities $\mathbf{Q}_e = \big(Q_e(y|x) : (x, y) \in \mathcal{X}_e \times \mathcal{Y}_e\big)$.

We will say that the edge (u, v) *exits* u and *enters* v. If $W \subseteq V$, then

$$\delta^-(W) \stackrel{\text{def}}{=} \{(u, v) \in E : u \notin W, v \in W\}; \quad (2)$$

$$\delta^+(W) \stackrel{\text{def}}{=} \{(u, v) \in E : u \in W, v \notin W\}; \quad (3)$$

$$\sigma(W) \stackrel{\text{def}}{=} \{(u, v) \in E : u \in W, v \in W\}. \quad (4)$$

Thus $\delta^-(W)$ (resp. $\delta^+(W)$) is the set of edges that exit (resp. enter) a vertex not in W and enter (resp. exit) one in W, while $\sigma(W)$ is the set of edges that exit *and* enter vertices in W. To simplify notation, we will write $\delta^-(v)$, $\delta^+(v)$, and $\sigma(v)$ for $\delta^-(\{v\})$, $\delta^+(\{v\})$, and $\sigma(\{v\})$, respectively. Note that $\sigma(v)$, if non-empty, contains exactly one edge, viz. a self-loop on v. It will also be convenient further to define

$$\delta_{in}(W) \stackrel{\text{def}}{=} \delta^-(W) \cup \sigma(W); \quad (5)$$

$$\delta_{out}(W) \stackrel{\text{def}}{=} \delta^+(W) \cup \sigma(W). \quad (6)$$

Thus $\delta_{in}(W)$ (resp. $\delta_{out}(W)$) is the set of edges that enter (resp. exit) a vertex in W. In particular, $\delta_{in}(v)$ (resp. $\delta_{out}(v)$) is just $\delta^-(v)$ (resp. $\delta^+(v)$) with the self-loop on v thrown in, if it exists.

To generate common randomness, the agents communicate interactively over the network for a certain number, say n, of steps. This communication proceeds according to an agreed-upon set of rules that specifies each agent's channel inputs in each step, based on the channel outputs available to him from all previous steps. More elaborately, in step k $(1 \leq k \leq n)$, agent v does the following in sequence and in synchronism with all the other agents:

- He determines the input symbol to be transmitted in step k on each exiting channel $e \in \delta_{out}(v)$, as a function $X_{e,k} = f_{e,k}(\mathbf{Y}_{e'}^{k-1} : e' \in \delta_{in}(v))$ of the sequences of outputs $\mathbf{Y}_{e'}^{k-1}$ received in the previous $k-1$ steps on all entering channels $e' \in \delta_{in}(v)$.

- He then transmits the symbols $X_{e,k}$ on their respective exiting channels.
- Finally, he receives the outputs $Y_{e',k}$ corresponding to the symbols transmitted in step k on all his entering channels.

After n steps, each agent either computes a random output taking values in a common finite set of size, say, K — without loss of generality, we will take this set to be $[K] \stackrel{\text{def}}{=} \{1, 2, \ldots, K\}$ — or decides that the attempt to generate common randomness failed. Each agent's decision is based solely on the output sequences available to him. Formally, agent v computes a *decision random variable* S_v that is a function of all the output sequences $\mathbf{Y}_{e'}^n$, $e' \in \delta_{in}(v)$, and takes values in the set $\{*\} \cup [K]$. Here, $S_v = *$ is supposed to indicate that v declared failure to generate common randomness.

Let $f_e = (f_{e,1}, \ldots, f_{e,n})$, and $\mathbf{f} = (f_e : e \in E)$. Let $\mathbf{S} = (S_v : v \in V)$. Then the pair $(\mathbf{f}, \mathbf{S})$, which all the agents agree on before communication begins, sums up the set of rules according to which the agents communicate over the network and make their final decisions. We will refer to $(\mathbf{f}, \mathbf{S})$ as an (n, K) *protocol* for generating common randomness. Of course, the "amount" of randomness generated by this protocol, and the extent to which it is truly "common," are determined by the joint distribution of the decision random variables S_v, $v \in V$. Ideally we would like to have

$$Pr\{S_v = l \text{ for all } v \in V\} = \frac{1}{K} \quad \text{for each } l \in [K], \quad (7)$$

with K as large as possible. If (7) were true, then all the S_v's would be equal with probability 1, and uniformly distributed over $[K]$. (There would be no "failure" events of positive probability.) Such a protocol could reasonably be said to generate $\log K$ bits of common randomness in n steps of communication.

In general, however, it is not possible to satisfy (7) except in the trivial case $K = 1$. Therefore, we will have to settle for *approximate* equality and uniformity of the S_v's. To this end, we make the following definition: $(\mathbf{f}, \mathbf{S})$ *is an* (n, K, λ) *protocol if*

$$\frac{1-\lambda}{K} \leq Pr\{S_v = l \text{ for all } v \in V\} \leq \frac{1+\lambda}{K} \quad \text{for each } l \in [K]. \quad (8)$$

To motivate this definition, suppose that $K = \exp\{nR - o(n)\}$ for some $R > 0$, and $\lambda = o(1)$ — what this means is that $(\mathbf{f}, \mathbf{S})$ is the n^{th} term in a *sequence* of (n, K_n, λ_n) protocols, with $\liminf_{n\to\infty}(1/n)\log K_n = R$ and $\lim_{n\to\infty} \lambda_n = 0$. Then, by (8),

$$Pr\{\exists\, l \in [K] \text{ such that } S_v = l \text{ for all } v \in V\} \geq 1 - \lambda. \quad (9)$$

This means that all the agents compute the *same* random output with high probability. In particular, the probability that some agent declares failure to generate common randomness is small.

The above considerations motivate the following definition of the common randomness capacity of the given network:

Definition 2.1 *R is an achievable rate of generating common randomness over the given network if there exists a sequence of (n, K_n, λ_n) protocols such that*

$$\lim_{n\to\infty} \lambda_n = 0 \quad \textit{and} \quad \liminf_{n\to\infty} \frac{\log K_n}{n} = R. \quad (10)$$

The common randomness capacity of the network is the supremum of all achievable rates.

3. Results

Our main result is a "single-letter" characterization of the common randomness capacity, in terms of the topology of the network and the characteristics of the channels constituting it.

For ease of reference, we first record the definitions of some standard graph-theoretic concepts. All definitions are with respect to the given digraph $G = (V, E)$.

A *path* from vertex u to vertex w is a set of $k \geq 1$ edges

$$\{(v_0, v_1), (v_1, v_2), (v_2, v_3), \ldots, (v_{k-1}, v_k)\},$$

with $v_0 = u$ and $v_k = w$. A *circuit* is a path from some vertex to itself. Note that a self-loop (v, v) constitutes a circuit by itself.

An *arborescence rooted at* v is a set T of edges, with the following properties: (i) no edge in T enters v; (ii) for each $u \neq v$, exactly one edge in T enters u; and (iii) T does not contain any circuits.

It can be verified easily that the above properties imply that (iv) $|T| = |V| - 1$; (v) for each $u \neq v$, there is a *unique* path in T from v to u; and (vi) the edges of T form a spanning tree in the undirected graph underlying G. We will denote the set of all arborescences rooted at v by $\mathcal{T}(v)$, and let $\mathcal{T} = \bigcup_{v\in V} \mathcal{T}(v)$.

If $(u, v) \in T$, then we will say that u is the *parent of* v *in* T, and v is a *child of* u *in* T. Note that a vertex can have more than one child in T, but no more than one parent — in fact, every vertex other than the root has exactly one parent in T (the root has none). A vertex with no children in T will be called a *leaf of* T (every arborescence has at least one leaf).

For each $e \in E$, let P_e be a probability distribution on the input alphabet $\mathcal{X}_e$ of channel e, and let $\mathbf{P} = (P_e : e \in E)$. Let the vectors $\mathbf{a}(\mathbf{P}) \in \mathbf{R}_+^V$ and $\mathbf{b}(\mathbf{P}) \in \mathbf{R}_+^E$ be given by

$$a_v(\mathbf{P}) \stackrel{\text{def}}{=} \sum_{e \in \delta_{in}(v)} H\left(\mathbf{Q}_e | P_e\right), \tag{11}$$

$$b_e(\mathbf{P}) \stackrel{\text{def}}{=} I\left(P_e; \mathbf{Q}_e\right), \tag{12}$$

and let $\mathcal{R}(\mathbf{P})$ be the polyhedron of all vectors $\mathbf{r} \in \mathbf{R}_+^{\mathcal{T}}$ satisfying the following constraints:

$$\sum_{T \in \mathcal{T}(v)} r_T \leq a_v(\mathbf{P}), \quad \text{for each } v \in V; \tag{13}$$

$$\sum_{T : e \in T} r_T \leq b_e(\mathbf{P}), \quad \text{for each } e \in E. \tag{14}$$

Note that the constraint in (14) can be ignored if e is a self-loop, because the LHS is then a summation over an empty set (no arborescence contains a self-loop), and is therefore equal to zero.

The "achievability" part of our main result essentially states that for any $\mathbf{P}$ and any $\mathbf{r} \in \mathcal{R}(\mathbf{P})$, the rate $\sum_T r_T$ is achievable.

Theorem 3.1 (Achievability result) *Let*

$$C_*(\mathbf{P}) \stackrel{\text{def}}{=} \max_{\mathbf{r} \in \mathcal{R}(\mathbf{P})} \sum_{T \in \mathcal{T}} r_T, \tag{15}$$

$$C_* \stackrel{\text{def}}{=} \max_{\mathbf{P}} C_*(\mathbf{P}). \tag{16}$$

Then the common randomness capacity of the network is bounded from below by C_*.

To establish that the common randomness capacity is actually *equal* to C_*, we must also prove an appropriate "converse" result. For this purpose, it will be convenient first to derive different, more explicit, expressions for $C_*(\mathbf{P})$ and C_*.

Note that $C_*(\mathbf{P})$ is defined in (15) to be the optimal value of a certain linear program (LP). We will now write down the *dual* to this LP. From now on, we will refer to the LP in (15) as the *primal*. The dual LP has a variable $x_v \in \mathbf{R}_+$ for each $v \in V$, and a variable $y_e \in \mathbf{R}_+$ for each $e \in E$. The dual constraints are

$$x_v + \sum_{e \in T} y_e \geq 1, \quad \text{for each } v \in V \text{ and } T \in \mathcal{T}(v). \tag{17}$$

Let $\mathcal{D}$ denote the dual feasible region (which does not depend on $\mathbf{P}$). Thus $\mathcal{D}$ is the polyhedron of all vectors $(\mathbf{x}, \mathbf{y}) \in \mathbf{R}_+^V \times \mathbf{R}_+^E$ satisfying (17).

The dual objective is to minimize $\sum_{v \in V} a_v(\mathbf{P}) x_v + \sum_{e \in E} b_e(\mathbf{P}) y_e$. By linear programming duality, the optimal values of the primal and dual problems are equal, i.e.,

$$C_*(\mathbf{P}) = \min_{(\mathbf{x},\mathbf{y}) \in \mathcal{D}} \left[\sum_{v \in V} a_v(\mathbf{P}) x_v + \sum_{e \in E} b_e(\mathbf{P}) y_e \right]. \tag{18}$$

The key step in obtaining more convenient expressions for $C_*(\mathbf{P})$ and C_* is to decompose the polyhedron $\mathcal{D}$ as the vector sum of the convex hull of its extreme points and the cone generated by its extreme directions. (Every polyhedron of non-negative vectors can be so decomposed.) Now, the cone generated by the extreme directions of $\mathcal{D}$ equals all of $\mathbf{R}_+^V \times \mathbf{R}_+^E$ because $\mathcal{D}$ has the following property: if $(\mathbf{x}, \mathbf{y}) \in \mathcal{D}$, and $\mathbf{x}' \geq \mathbf{x}$, $\mathbf{y}' \geq \mathbf{y}$, then $(\mathbf{x}', \mathbf{y}') \in \mathcal{D}$. As for the extreme points of $\mathcal{D}$, the following result identifies a finite set $D_0 \subseteq \mathcal{D}$ that contains all of them.

Theorem 3.2 (Combinatorial result) *For each non-empty subset W of V, let $\mathbf{x}(W) \in \mathbf{R}_+^V$ and $\mathbf{y}(W) \in \mathbf{R}_+^E$ be given by*

$$x_v(W) \stackrel{\text{def}}{=} \begin{cases} 1 & \text{if } v \in W; \\ 0 & \text{otherwise;} \end{cases} \quad \text{and}$$

$$y_e(W) \stackrel{\text{def}}{=} \begin{cases} 1 & \text{if } e \in \delta^-(W); \\ 0 & \text{otherwise.} \end{cases} \tag{19}$$

For each $1 \leq m < |V|$, and each collection of non-empty and pairwise disjoint subsets $V_1, \ldots, V_{m+1}$ of V, let $\mathbf{y}(V_1, \ldots, V_{m+1}) \in \mathbf{R}_+^E$ be given by

$$y_e\left(V_1, \ldots, V_{m+1}\right) \stackrel{\text{def}}{=} \begin{cases} 1/m & \text{if } e \in \bigcup_{i=1}^{m+1} \delta^-(V_i); \\ 0 & \text{otherwise.} \end{cases} \tag{20}$$

Let D_0 be the set consisting of the vectors $(\mathbf{x}(W), \mathbf{y}(W))$ and $(\mathbf{0}, \mathbf{y}(V_1, \ldots, V_{m+1}))$ defined above. Then $D_0 \subseteq \mathcal{D}$; in fact,

$$\mathcal{D} = \mathit{conv}(D_0) + \mathbf{R}_+^V \times \mathbf{R}_+^E. \tag{21}$$

Here, conv (D_0) denotes the convex hull of all the vectors in D_0.

The desired decomposition of $\mathcal{D}$ is given by (21).

We are now in a position to state the "converse" part of the main result.

Theorem 3.3 (Converse result) *Let*

$$C^*(\mathbf{P}) \stackrel{\text{def}}{=} \min_{(\mathbf{x},\mathbf{y}) \in D_0} \left[\sum_{v \in V} a_v(\mathbf{P}) x_v + \sum_{e \in E} b_e(\mathbf{P}) y_e \right] \tag{22}$$

$$C^* \stackrel{\text{def}}{=} \max_{\mathbf{P}} C^*(\mathbf{P}). \tag{23}$$

Then the common randomness capacity of the network is bounded from above by C^*.

We claim now that $C_*(\mathbf{P}) = C^*(\mathbf{P})$ for every $\mathbf{P}$, and hence $C_* = C^*$. To see this, note that by (21), and the non-negativity of the vectors $\mathbf{a}(\mathbf{P})$ and $\mathbf{b}(\mathbf{P})$, the optimal value of the dual LP equals the minimum of the dual objective function over $\text{conv}(D_0)$, which in turn equals the minimum over D_0, by linearity. But the dual optimal value also equals $C_*(\mathbf{P})$ by LP duality (see (18)), and the minimum of the dual objective over D_0 equals $C^*(\mathbf{P})$ by definition (see (22)). Hence $C_*(\mathbf{P}) = C^*(\mathbf{P})$. It follows now that both C_* and C^* equal the common randomness capacity of the given network.

Theorem 3.3 essentially states that if $\lim_{n\to\infty} \lambda_n = 0$, then there does not exist a sequence of (n, K_n, λ_n) protocols for generating common randomness with $\liminf_{n\to\infty}(1/n)\log K_n > C^*$. (We can actually prove a slightly stronger result, with liminf replaced by limsup.) In the usual terminology, this is a "weak" converse to Theorem 3.1. A "strong" converse would state that even if only $\limsup_{n\to\infty} \lambda_n < 1$ is assumed (instead of $\lim_{n\to\infty} \lambda_n = 0$), there does not exist a sequence of (n, K_n, λ_n) protocols with $\limsup_{n\to\infty}(1/n)\log K_n > C^*$. Such a result was proved in [11] for the two-agent case. By similar methods, it is indeed possible to prove a "strong" converse in the general case, too.

Proofs of the theorems stated in this document are available in [12], and a paper based on these results is currently being prepared for the archival literature.

4. Acknowledgements

The research reported in this document was supported by the National Science Foundation through grants IRI 9005849, IRI 9310670, and NCR 9422513, and by the AT&T Foundation. The first author was resident at the University of California at Berkeley, as an exchange scholar, while the research was carried out.

References

[1] R. Ahlswede. "Elimination of Correlation in Random Codes for Arbitrarily Varying Channels", *Z. Wahrsch. Verw. Gebiete*, Vol. 33, 1978, pp. 159-175.

[2] R. Ahlswede and I. Csisz"à"r. "Common Randomness in Information Theory and Cryptography, Part 1: Secret Sharing", *IEEE-IT*, Vol. 39, No. 4, July 1993, pp. 1121 -1132.

[3] R. Ahlswede and G. Dueck. "Identification via Channels", *IEEE-IT*, Vol. 35, No. 1, January 1989, pp. 15 -29.

[4] R. Ahlswede and G. Dueck. "Identification in the Presence of Feedback - A Discovery of New Capacity Formulas", *IEEE-IT*, Vol. 35, No. 1, January 1989, pp. 30 -36.

[5] R. Ahlswede and B. Verboven. "On Identification via Multiway Channels with Feedback", *IEEE-IT*, Vol. 37, No. 6, November 1991, pp. 1519 -1526.

[6] I. Csiszar and P. Narayan. "The Capacity of the Arbitrarily Varying Channel Revisited: Positivity Constraints", *IEEE-IT*, Vol. 34, No. 2, March 1988, pp. 181-193.

[7] L. Lovász. "Communication Complexity: A Survey", In *Paths, Flows and VLSI layout*, B.H. Korte et. al, Editors, Springer-Verlag, 1990.

[8] U.M. Maurer". "Secret Key Agreement by Public Discussion from Common Information", *IEEE-IT*, Vol. 39, No. 3, May 1993, pp. 733 -742.

[9] A. Orlitsky and A. El Gamal. "Communication Complexity", In *Complexity in Information Theory*, Y. Abu-Mostafa, Editor, Springer-Verlag, 1988.

[10] A. Orlitsky and A. El Gamal, "Average and Randomized Communication Complexity", *IEEE-IT*, Vol. 36, No. 1,January 1990, pp. 3 -16.

[11] S. Venkatesan and V. Anantharam. "The Common Randomness Capacity of a Pair of Independent Discrete Memoryless Channels", *Electronics Research Laboratory*, University of California, Berkeley, UCB/ERL M95/85, September 1995.

[12] S. Venkatesan and V. Anantharam. "Generating Common Randomness in an Arbitrary Network of Channels : Capacity Formulas and some Combinatorial Results", *Electronics Research Laboratory*, University of California, Berkeley, UCB/ERL M97/9, January 1997.