数理解析研究所講究録
第 1906 巻 2014 年 103-110

Inequalities for Sums of Joint Entropy Based on the Strong Subadditivity

Yuta Kishi
Department of Mathematical Information Science, Graduate School of Science
Tokyo University of Science

Nozomu Ochiumi
Faculty of Engineering, Shonan Institute of Technology

Masahiro Yanagida
Department of Mathematical Information Science, Faculty of Science
Tokyo University of Science

1 Introduction

In what follows, $V=\{1,\ldots,n\}$ is the set of indices of given random variables $X_1,\ldots,X_n$, and $\mathcal{B}=\{B_1,\ldots,B_m\}$ is a set of subsets (possibly with repeat) of V. Furthermore, for $S=\{i_1,\ldots,i_\ell\}\subseteq V$, X_S and $H(X_S)$ denote the random vector $(X_{i_1},\ldots,X_{i_\ell})$ and its Shannon entropy $H(X_{i_1},\ldots,X_{i_\ell})$ $(H(X_\emptyset)=0)$. The power set (the set of all subsets) and the set of all ℓ-subsets of V are written as 2^V and $\binom{V}{\ell}$, respectively. For simplicity, we state results only for discrete random variables with finite alphabets for which the entropy functions are always well-defined.

The following entropy inequality, which is called Shearer's inequality, is given in [1] as a key lemma used in certain combinatorial argument.

Theorem A (Shearer's inequality [1]). *If every element of V appears in at least λ members of $\mathcal{B}$, i.e., $|\{j\mid i\in B_j\}|\geqslant\lambda$ for each $i\in V$, then*

$$\lambda H(X_V)\leqslant\sum_{B\in\mathcal{B}}H(X_B).$$

Theorem A yields as a special case the subadditivity of joint entropy $H(X_V)\leqslant\sum_{i\in V}H(X_i)$, which as well as other basic properties of entropy has played important roles in deriving a number of combinatorial results (see for example [2]–[6]). A simple and intuitively clear proof of Theorem A is given in [7] by proposing the "dropping method" explained in the following paragraph.

Joint entropy has the strong subadditivity

$$H(X_{S\cap T})+H(X_{S\cup T})\leqslant H(X_S)+H(X_T)\tag{1}$$

for $S,T\subseteq V$ since $H(X_{S\cup T})-H(X_T)=H(X_{S-S\cap T}\mid X_T)\leqslant H(X_{S-S\cap T}\mid X_{S\cap T})=H(X_S)-H(X_{S\cap T})$. In (1), $S\cap T$ and $S\cup T$ result from arranging S and T in upper and lower rows and then "dropping from S to T" all the elements $i \in S$ and $i \notin T$ as the following example:

$$\begin{array}{ccccc} S = \{\ \mathbf{1},\ 2,\quad \mathbf{4}\ \} & & S \cap T = \{\quad 2 \quad\quad \} \\ \downarrow \quad \downarrow & \longrightarrow & \downarrow \quad \downarrow \\ T = \{\quad 2,\ 3 \quad \} & & S \cup T = \{\ \mathbf{1},\ 2,\ 3,\ \mathbf{4}\ \} \end{array}$$

Consider the following simple algorithm D for $\beta_1, \ldots, \beta_m \subseteq V$.

1. $r \leftarrow m$.
2. Compute $\beta_r \cap \beta_{r-1}$ and $\beta_r \cup \beta_{r-1}$ by the dropping from β_r to β_{r-1}, and let these be new β_r and β_{r-1}, respectively.
3. $r \leftarrow r - 1$ if $r > 2$, and go to 2. Stop if $r = 2$.

Run D with the initialization $\beta_j \leftarrow B_j$ for $1 \leqslant j \leqslant m$. In step 2, first β_m and β_{m-1} change from B_m and B_{m-1} to $B_m \cap B_{m-1}$ and $B_m \cup B_{m-1}$ when $r = m$, and next β_{m-1} and β_{m-2} change from $B_m \cup B_{m-1}$ and B_{m-2} to $(B_m \cup B_{m-1}) \cap B_{m-2}$ and $B_m \cup B_{m-1} \cup B_{m-2}$ when $r = m - 1$. Thus β_r and β_{r-1} change from $B_m \cup B_{m-1} \cup \cdots \cup B_r$ and B_{r-1} to $(B_m \cup B_{m-1} \cup \cdots \cup B_r) \cap B_{r-1}$ and $B_m \cup B_{m-1} \cup \cdots \cup B_r \cup B_{r-1}$ for each $r = m, m-1, \ldots, 2$. Hence by $(m - 1)$ times applications of the strong subadditivity, we have

$$H(X_{B_1^{(1)}}) + H(X_{B_2^{(1)}}) + \cdots + H(X_{B_m^{(1)}}) \leqslant H(X_{B_1}) + H(X_{B_2}) + \cdots + H(X_{B_m}),$$

where $B_j^{(1)} = (B_m \cup B_{m-1} \cup \cdots \cup B_j) \cap B_{j-1}$ for $2 \leqslant j \leqslant m$ and $B_1^{(1)} = B_m \cup B_{m-1} \cup \cdots \cup B_2 \cup B_1$ because D finishes with $\beta_j = B_j^{(1)}$. For each $i \in V$, let λ_i be the number of members of $\mathcal{B}$ containing i, that is,

$$\lambda_i = |\{j \mid i \in B_j\}|,$$

then $i \in B_1^{(1)}$ and there are $(\lambda_i - 1)$ sets containing i among $B_2^{(1)}, \ldots, B_m^{(1)}$ if $\lambda_i \geqslant 1$. Let $B_1^{(2)}, \ldots, B_m^{(2)}$ be the result of running D again with the initialization $\beta_j \leftarrow B_j^{(1)}$ for $1 \leqslant j \leqslant m$, then

$$H(X_{B_1^{(2)}}) + H(X_{B_2^{(2)}}) + \cdots + H(X_{B_m^{(2)}}) \leqslant H(X_{B_1^{(1)}}) + H(X_{B_2^{(1)}}) + \cdots + H(X_{B_m^{(1)}}),$$

$i \in B_1^{(2)}$, $i \in B_2^{(2)}$ and there are $(\lambda_i - 2)$ sets containing i among $B_3^{(2)}, \ldots, B_m^{(2)}$ if $\lambda_i \geqslant 2$ for each $i \in V$. Therefore at most $(m - 1)$ times applications of D to the list obtained thus far yield $\beta_j = A_j$ for $1 \leqslant j \leqslant m$, and we have

$$\sum_{j=1}^{m} H(X_{A_j}) \leqslant \sum_{j=1}^{m} H(X_{B_j}), \tag{2}$$

where $A_1, \ldots, A_m \subseteq V$ are defined by $i \in A_1, \ldots, i \in A_{\lambda_i}$, $i \notin A_{\lambda_i+1}, \ldots, i \notin A_m$, i.e.,

$$A_j = \{i \in V \mid j \leqslant \lambda_i\} \text{ for } 1 \leqslant j \leqslant m.$$

The assumption of Theorem A is equivalent to $\lambda \leqslant \lambda_i$ for all $i \in V$, hence $A_1 = \cdots = A_\lambda = V$ holds and we obtain Theorem A by (2).

Han's inequality [8] is a classic result in information theory. It essentially states that

$$h_\ell^{(n)} = \frac{1}{\binom{n}{\ell}} \sum_{S \in \binom{V}{\ell}} \frac{H(X_S)}{\ell} \text{ for } 1 \leqslant \ell \leqslant n, \tag{3}$$

which means the average entropy per symbol of randomly drawn ℓ-subset of $\{X_1, \ldots, X_n\}$, decrease as the size of subset increases.

Theorem B (Han's inequality [8]). *Let $h_\ell^{(n)}$ be defined as* (3). *Then $h_\ell^{(n)} \leqslant h_{\ell-1}^{(n)}$ holds for $2 \leqslant \ell \leqslant n$.*

This inequality implies the subadditivity of joint entropy since $nh_n^{(n)} = H(X_V)$ and $nh_1^{(n)} = \sum_{i \in V} H(X_i)$. Han's inequality was first shown in [8] and another proof was given in [7] by using Theorem A (see also [9, 10]). This inequality has found applications in multi-user information theoretic problems; e.g. [11]–[13]. Furthermore, a generalization of Theorem B to allow common components among the random variables is given in [14].

In this paper, we give a generalization of Shearer's inequality in case that every t-subset of V is contained in at least λ members of $\mathcal{B}$. We also give a refinement of Han's inequality on monotonicity of the average entropy by applying the new inequality. We hope that our inequalities may find their applications in the future, just as Han's inequality finds applications in [11]–[13] some 20 or 30 years after its discovery.

2 A Generalization of Shearer's Inequality

In this section, for each $S \subseteq V$, let λ_S be the number of members of $\mathcal{B}$ containing S, i.e.,

$$\lambda_S = |\{j \mid S \subseteq B_j\}|,$$

and Ω_S the set in the right-hand side.

The following result is a generalization of Shearer's inequality. In fact, Theorem 1 coincides with Theorem A in case $t = 1$.

Theorem 1. *Let $X_1, \ldots, X_n$ be discrete random variables with finite alphabets. If every t-subset of $V = \{1, \ldots, n\}$ is contained in at least λ members of $\mathcal{B} = \{B_1, \ldots, B_m\} \subseteq 2^V$, i.e., $\lambda_T = |\{j \mid T \subseteq B_j\}| \geqslant \lambda$ for each $T \in \binom{V}{t}$, then*

$$\lambda \binom{n}{t-1} H(X_V) + \left\lceil \frac{\lambda(n-k)}{k-t+1} \right\rceil \sum_{S \in \binom{V}{t-1}} H(X_S) \leqslant \binom{k}{t-1} \sum_{B \in \mathcal{B}} H(X_B),$$

where k is an upper bound for the sizes of members of $\mathcal{B}$, i.e., $|B_j| \leqslant k$ for $1 \leqslant j \leqslant m$.

We prepare the following lemma to prove Theorem 1.

Lemma 2. *Let $|B_j| \leqslant k$ for $1 \leqslant j \leqslant m$. If $\lambda_T \geqslant \lambda$ for each $T \in \binom{V}{t}$, then*

$$\lambda_S \geqslant \left\lceil \frac{\lambda(n-t+1)}{k-t+1} \right\rceil$$

holds for each $S \in \binom{V}{t-1}$.

Proof. Let $\mathbb{I}_{B_j}$ be the indicator function of B_j. Since

$$|B_j \setminus S| = |B_j| - |S| \leqslant k - t + 1$$

when $j \in \Omega_S$, we have

$$\sum_{j\in\Omega_S}\sum_{i\in V\setminus S} \mathbb{I}_{B_j}(i) = \sum_{j\in\Omega_S} |B_j \setminus S| \leqslant \lambda_S(k - t + 1).$$

On the other hand,

$$|\{j \in \Omega_S \mid i \in B_j\}| = \lambda_{S\cup\{i\}} \tag{4}$$

for each $S \subseteq V$ and $i \in V$, hence

$$\sum_{i\in V\setminus S}\sum_{j\in\Omega_S} \mathbb{I}_{B_j}(i) = \sum_{i\in V\setminus S} \lambda_{S\cup\{i\}} \geqslant \lambda(n - t + 1).$$

because $|S \cup \{i\}| \leqslant t$ and thus $\lambda_{S\cup\{i\}} \geqslant \lambda$. □

Proof of Theorem 1. For each $S \in \binom{V}{t-1}$, by applying the dropping method to $\{B_j \mid j \in \Omega_S\} = \{B_{j_1}, \dots, B_{j_{\lambda_S}}\}$ as mentioned in the previous section, we have

$$\sum_{\ell=1}^{\lambda_S} H(X_{A_{j_\ell}}) \leqslant \sum_{\ell=1}^{\lambda_S} H(X_{B_{j_\ell}})$$

by strong subadditivity where

$$A_{j_\ell} = \{i \in V \mid \ell \leqslant \lambda_{S\cup\{i\}}\} \text{ for } 1 \leqslant \ell \leqslant \lambda_S$$

because it follows from (4) that each $i \in V$ belongs to $\lambda_{S\cup\{i\}}$ members of $\{B_{j_1}, \dots, B_{j_{\lambda_S}}\}$. If $i \notin S$, then $|S \cup \{i\}| = t$, so that $\lambda_{S\cup\{i\}} \geqslant \lambda$ and $i \in A_{j_1}, \dots, A_{j_\lambda}$. If $i \in S$, then $\lambda_{S\cup\{i\}} = \lambda_S$ $(\geqslant \lambda)$ and $i \in A_{j_1}, \dots, A_{j_{\lambda_S}}$. Therefore $A_{j_1} = \cdots = A_{j_\lambda} = V$ and $A_{j_{\lambda+1}}, \dots, A_{j_{\lambda_S}} \supseteq S$. Thus we have

$$\begin{aligned}\sum_{j\in\Omega_S} H(X_{B_j}) &\geqslant \sum_{j\in\Omega_S} H(X_{A_j})\\ &= \sum_{\ell=1}^{\lambda} H(X_{A_{j_\ell}}) + \sum_{\ell=\lambda+1}^{\lambda_S} H(X_{A_{j_\ell}})\\ &\geqslant \lambda H(X_V) + (\lambda_S - \lambda)\, H(X_S).\end{aligned} \tag{5}$$

by monotonicity of entropy functions. Summing up both sides of (5) over all $S \in \binom{V}{t-1}$, by Lemma 2 and nonnegativity of entropy functions, we have

$$\sum_{S\in\binom{V}{t-1}}\sum_{j\in\Omega_S} H(X_{B_j}) \geqslant \lambda\binom{n}{t-1} H(X_V) + \left\lceil \frac{\lambda(n-k)}{k-t+1} \right\rceil \sum_{S\in\binom{V}{t-1}} H(X_S). \tag{6}$$

Let $\mathcal{M} = (\mathcal{M}_{S,j})$ be a $\binom{n}{t-1} \times m$ matrix defined by $\mathcal{M}_{S,j} = H(X_{B_j})$ if $j \in \Omega_S$ and $\mathcal{M}_{S,j} = 0$ otherwise for each $S \in \binom{V}{t-1}$ and $1 \leqslant j \leqslant m$. Then the sum of all row-sums is

$$\sum_{S\in\binom{V}{t-1}} \sum_{j=1}^{m} \mathcal{M}_{S,j} = \sum_{S\in\binom{V}{t-1}} \sum_{j\in\Omega_S} H(X_{B_j}) \tag{7}$$

and the sum of all column-sums is

$$\begin{aligned} \sum_{j=1}^{m} \sum_{S\in\binom{V}{t-1}} \mathcal{M}_{S,j} &= \sum_{j=1}^{m} \sum_{S\in\binom{B_j}{t-1}} H(X_{B_j}) \\ &\leqslant \binom{k}{t-1} \sum_{j=1}^{m} H(X_{B_j}) \end{aligned} \tag{8}$$

because $j \in \Omega_S$ is equivalent to $S \subseteq B_j$, and also $|B_j| \leqslant k$ and nonnegativity of entropy functions. The desired inequality follows from (6), (7) and (8). $\square$

A special case of V and $\mathcal{B}$ satisfying the conditions required in Theorem 1 is the case that they form a t-(n, k, λ) design. The pair $(V, \mathcal{B})$ is said to be a t-(n, k, λ) design if

$$\lambda_T = |\{j \mid T \subseteq B_j\}| = \lambda \quad \text{and} \quad |B_j| = k$$

for each $T \in \binom{V}{t}$ and $1 \leqslant j \leqslant m$. Moreover

$$\lambda_S = |\{j \mid S \subseteq B_j\}| = \frac{\lambda(n-t+1)}{k-t+1}$$

holds for each $S \in \binom{V}{t-1}$ by the property of t-design. Combinatorial design theory is a fundamental branch of combinatorics connecting coding theory and other applications in computer science (see for example [15], [16]).

Theorem 3. *If $(V, \mathcal{B})$ is a t-(n, k, λ) design, then*

$$\lambda\binom{n}{t-1} H(X_V) + \frac{\lambda(n-k)}{k-t+1} \sum_{S\in\binom{V}{t-1}} H(X_S) \leqslant \binom{k}{t-1} \sum_{B\in\mathcal{B}} H(X_B).$$

3 A Refinement of Han's Inequality

As an application of the results in the previous section, we obtain a refinement of Han's inequality on monotonicity of the average entropy. The following theorem states that differences between consecutive terms of the sequence $h_1^{(n)}, \ldots, h_n^{(n)}$ are monotone in a sense, and thus they turn out to be nonnegative. Therefore this result is seen to be a refinement of Han's inequality.

Theorem 4. *Let $h_\ell^{(n)}$ be defined as*

$$h_\ell^{(n)} = \frac{1}{\binom{n}{\ell}} \sum_{S \in \binom{V}{\ell}} \frac{H(X_S)}{\ell} \tag{3}$$

for $1 \leqslant \ell \leqslant n$. Then

$$0 \leqslant (\ell-2)(\ell-1)\left(h_{\ell-2}^{(n)} - h_{\ell-1}^{(n)}\right) \leqslant (\ell-1)\ell\left(h_{\ell-1}^{(n)} - h_\ell^{(n)}\right)$$

holds for $3 \leqslant \ell \leqslant n$.

Proof. Let $2 \leqslant \ell \leqslant n$. For each ℓ-subset U of V,

$$\binom{\ell}{\ell-2} H(X_U) + \sum_{S \in \binom{U}{\ell-2}} H(X_S) \leqslant (\ell-1) \sum_{T \in \binom{U}{\ell-1}} H(X_T) \tag{9}$$

holds by Theorem 3 because $\left(U, \binom{U}{\ell-1}\right)$ is a $(\ell-1)$-$(\ell, \ell-1, 1)$ design, that is, every $(\ell-1)$-subset of U is contained in exactly one member of $\binom{U}{\ell-1}$. Summing up both sides of (9) over all $U \in \binom{V}{\ell}$, we have

$$\binom{\ell}{\ell-2} \sum_{U \in \binom{V}{\ell}} H(X_U) + \sum_{U \in \binom{V}{\ell}} \sum_{S \in \binom{U}{\ell-2}} H(X_S) \leqslant (\ell-1) \sum_{U \in \binom{V}{\ell}} \sum_{T \in \binom{U}{\ell-1}} H(X_T).$$

The right hand side is equal to $(\ell-1)(n-\ell+1)\sum_{T \in \binom{V}{\ell-1}} H(X_T)$ by the double-counting on the $\binom{n}{\ell} \times \binom{n}{\ell-1}$ matrix whose rows are labeled by U's $\in \binom{V}{\ell}$, columns are labeled by T's $\in \binom{V}{\ell-1}$, and (U,T)-element is given by $H(X_T)$ if $T \subseteq U$ and by 0 otherwise, because we have that $\sum_{U \in \binom{V}{\ell}} \sum_{T \in \binom{U}{\ell-1}} H(X_T)$ = the sum of all row-sums = the sum of all column-sums = $(n-\ell+1)\sum_{T \in \binom{V}{\ell-1}} H(X_T)$. Similarly, the second term in the left hand side is equal to $\binom{n-\ell+2}{2} \sum_{S \in \binom{V}{\ell-2}} H(X_S)$. Thus, we have

$$\binom{\ell}{\ell-2} \sum_{U \in \binom{V}{\ell}} H(X_U) + \binom{n-\ell+2}{2} \sum_{S \in \binom{V}{\ell-2}} H(X_S) \leqslant (\ell-1)(n-\ell+1) \sum_{T \in \binom{V}{\ell-1}} H(X_T). \tag{10}$$

Dividing both sides of (10) by $\binom{\ell}{2}\binom{n}{\ell}$ finishes the proof. □

Moreover we obtain a generalization of Theorem 4 for differences between two terms which are not necessarily consecutive which holds if one pair (i, j) exists on the left side of another pair (k, ℓ) in some sense.

Theorem 5. *Let $h_\ell^{(n)}$ be defined as*

$$h_\ell^{(n)} = \frac{1}{\binom{n}{\ell}} \sum_{S \in \binom{V}{\ell}} \frac{H(X_S)}{\ell} \tag{3}$$

for $1 \leqslant \ell \leqslant n$. Then

$$0 \leqslant \frac{ij}{j-i}\left(h_i^{(n)} - h_j^{(n)}\right) \leqslant \frac{k\ell}{\ell-k}\left(h_k^{(n)} - h_\ell^{(n)}\right) \tag{11}$$

holds for $1 \leqslant i, j, k, \ell \leqslant n$ such that $i < j$, $k < \ell$, $i \leqslant k$, $j \leqslant \ell$.

Proof. By Theorem 4,

$$\begin{aligned} h_i^{(n)} - h_j^{(n)} &= \left(h_i^{(n)} - h_{i+1}^{(n)}\right) + \cdots + \left(h_{j-1}^{(n)} - h_j^{(n)}\right) \\ &\leqslant \left(\frac{1}{i(i+1)} + \cdots + \frac{1}{(j-1)j}\right) i(i+1)\left(h_i^{(n)} - h_{i+1}^{(n)}\right) \\ &= \frac{j-i}{ij} \cdot i(i+1)\left(h_i^{(n)} - h_{i+1}^{(n)}\right) \end{aligned}$$

and

$$0 \leqslant i(i+1)\left(h_i^{(n)} - h_{i+1}^{(n)}\right) \leqslant \frac{ij}{j-i}(h_i^{(n)} - h_j^{(n)}) \tag{12}$$

holds, and thus we have the first inequality in (11). In the same way as (12), we obtain

$$\frac{ij}{j-i}\left(h_i^{(n)} - h_j^{(n)}\right) \leqslant (j-1)j\left(h_{j-1}^{(n)} - h_j^{(n)}\right),$$

$$k(k+1)\left(h_k^{(n)} - h_{k+1}^{(n)}\right) \leqslant \frac{k\ell}{\ell-k}\left(h_k^{(n)} - h_\ell^{(n)}\right),$$

hence (11) holds in case $j \leqslant k$ by Theorem 4. In case $j > k$,

$$\begin{aligned} \frac{ik}{k-i}\left(h_i^{(n)} - h_j^{(n)}\right) &= \frac{ik}{k-i}\left(h_i^{(n)} - h_k^{(n)} + h_k^{(n)} - h_j^{(n)}\right) \\ &\leqslant \left(\frac{kj}{j-k} + \frac{ik}{k-i}\right)\left(h_k^{(n)} - h_j^{(n)}\right) \\ &= \frac{k(j-i)}{j(k-i)} \cdot \frac{kj}{j-k}\left(h_k^{(n)} - h_j^{(n)}\right) \end{aligned}$$

holds by (11) for the previous case, then we have

$$\frac{ij}{j-i}\left(h_i^{(n)} - h_j^{(n)}\right) \leqslant \frac{kj}{j-k}\left(h_k^{(n)} - h_j^{(n)}\right).$$

We also obtain the following in the same way;

$$\frac{kj}{j-k}\left(h_k^{(n)} - h_j^{(n)}\right) \leqslant \frac{k\ell}{\ell-k}\left(h_k^{(n)} - h_\ell^{(n)}\right).$$

Combining them finishes the proof. □

References

[1] F. R. K. Chung, R. L. Graham, P. Frankl and J. B. Shearer, "Some intersection theorems for ordered sets and graphs," *J. Comb. Theory Ser. A,* vol. 43, pp. 23–37, 1986.

[2] N. Pippenger, "An information-theoretic method in combinatorial theory," *J. Comb. Theory Ser. A,* vol. 23, pp. 99–104, 1977.

[3] N. Pippenger, "Entropy and enumeration of Boolean functions," *IEEE Trans. Inf. Theory,* vol. 45, pp. 2096–2100, 1999.

[4] J. Kahn, "An entropy approach to the hard-core model on bipartite graphs," *Comb. Probabil. Comput.,* vol. 10, pp. 219–237, 2001.

[5] J. Radhakrishnan. *Entropy and Counting* [Online]. Available: http://www.tcs.tifr.res.in/~jaikumar/Papers/EntropyAndCounting.pdf

[6] M. Madiman and P. Tetali, "Information inequalities for joint distributions, with interpretations and applications," *IEEE Trans. Inf. Theory,* vol. 56, pp. 2699–2713, 2010.

[7] M. Yanagida and Y. Horibe, "A dropping proof of an entropy inequality," *Appl. Math. Lett.,* vol. 21, pp. 840–842, 2008.

[8] T. S. Han, "Nonnegative entropy measures of multivariate symmetric correlations," *Inform. Control,* vol. 36, pp. 133–156, 1978.

[9] T. M. Cover and J. A. Thomas, *Elements of Information Theory,* 2nd ed. Hoboken, NJ: Wiley-Interscience (John Wiley & Sons), 2006.

[10] Y. Horibe, *Information Entropy Theory,* 2nd ed. Tokyo: Morikita Publ., 1997.

[11] A. Albanese, J. Blomer, J. Edmonds, M. Luby and M. Sudan, "Priority encoding transmission," *IEEE Trans. Inf. Theory,* vol. 42, pp. 1737–1744, 1996.

[12] H. Wang and P. Viswanath, "Vector Gaussian multiple description with two levels of receivers," *IEEE Trans. Inf. Theory,* vol. 55, pp. 401–410, 2009.

[13] C. Tian, S. Mohajer and S. Diggavi, "Approximating the Gaussian multiple description rate region under symmetric distortion constraints," *IEEE Trans. Inf. Theory,* vol. 55, pp. 3869–3891, 2009.

[14] C. Tian, "Inequalities for entropies of sets of subsets of random variables," in 2011 *IEEE International Symposium on Information Theory,* St. Petersburg, 2011, pp. 1950–1954.

[15] R. P. Stanley, *Enumerative Combinatorics,* vol. 2 (Cambridge Studies in Advanced Mathematics 62). Cambridge: Cambridge University Press, 1999.

[16] J. H. Dinitz and D. R. Stinson, *Contemporary Design Theory: A Collection of Surveys* (Wiley Series in Discrete Mathematics and Optimization). New York: Wiley-Interscience (John Wiley & Sons), 1992.