

Inequalities for Entropies of Sets of Subsets of Random Variables

Chao Tian
AT&T Labs-Research
Florham Park, NJ 07932, USA.
tian@research.att.com

Abstract—Han’s inequality on the entropy rates of subsets of random variables is a classic result in information theory, which often finds its application in multiuser information theoretic problems. In this note, we generalize Han’s inequality to allow common components among the random variables, or, in an equivalent manner, to replace the simple random variables in Han’s inequality by subsets of random variables. This additional ingredient significantly complicates the matter and the form of the resultant inequalities are rather different from the original Han’s inequality. Our proof only relies on the sub-modularity property of the entropy function and the super-modularity property of the conditional entropy function. This new set of inequalities also provides a new link between Han’s inequality and the n -way sub-modularity inequality.

I. INTRODUCTION

Han’s inequality [1] on the entropy rates of subsets of random variables is a classic result in information theory. It essentially states that the average entropy rates of subsets decrease as the size of subset increases [2] (p. 490). This inequality has found applications in multi-user information theoretic problems; e.g. [3]–[5]. Han’s inequality, however, does not naturally deal with common components among random variables efficiently. For example, consider two random variables¹ $V_1 = \{W_1, W_{12}\}$ and $V_2 = \{W_2, W_{12}\}$, where W_{12} is the common component between them. Then a simple application of Han’s inequality to the random variables V_1 and V_2 , which in this case is equivalent to the non-negativity of mutual information, gives

$$H(V_1) + H(V_2) \geq H(V_1, V_2). \quad (1)$$

This is not a good lower bound for $H(V_1) + H(V_2)$ for the random variable under consideration, as can be seen easily

$$H(V_1) + H(V_2) \geq H(V_1, V_2) + H(W_{12}). \quad (2)$$

In fact, (2) is simply the sub-modularity property of the entropy function [6]. The inequality (2) can indeed be reached by applying Han’s inequality in its conditional form, yet a natural question arises from this simple example: what is the form of the generalized Han’s inequality which takes into account of the common components? In this note, we provide exactly such a new set of inequalities.

¹Throughout this note, $\{\cdot\}$ is used to denote the collections of random variables, and the resulting random variable, such as V_1 and V_2 , is defined on the product of the alphabet of each of its elements.

The same question can be considered from another perspective, that is: what is Han’s inequality when the random variables are in fact subsets of random variables, instead of simple random variables? In the above example, V_1 and V_2 are themselves subsets of the set $\{W_1, W_2, W_{12}\}$. Since both perspectives may be of interest, we will present the main theorem in two equivalent forms.

The proof of this new set of inequalities hinges only on the sub-modular property of the entropy function, and it relies on another inequality previously discovered in [7], namely, the n -way sub-modularity inequality. As a consequence, this new set of inequalities provides a connection between Han’s inequality and the n -way sub-modularity inequality, as they are special cases of a wider spectrum of inequalities regarding subsets of random variables.

In addition to the well-known Han’s inequalities, inequalities on the entropies of subsets of random variables have been considered in other works, e.g. [8], [9]. The new inequalities we present in this work do not appear to be included in these existing inequalities, though the intersection-cover problem considered in those work is indeed closely related to the proof idea of our main result. This connection is hardly surprising, since all these results are directly related to sub-modularity functions, which naturally appear in the intersection-cover and related counting problems.

The rest of this note is organized as follows. In Section II, we first present the main results on the new set of inequalities, and then discuss its relationship with Han’s inequality and the n -way sub-modular inequality. The proofs are given in Section III, and Section IV gives a few concluding remarks.

II. MAIN RESULT

The set $\{1, 2, \dots, n\}$ is written as \mathcal{I}_n , and its power set (i.e., the collection of all its subsets) is written as $\mathcal{P}(\mathcal{I}_n)$. Define $\hat{\mathcal{P}}(\mathcal{I}_n) = \mathcal{P}(\mathcal{I}_n) \setminus \{\emptyset\}$, i.e., $\hat{\mathcal{P}}(\mathcal{I}_n)$ is the collection of all subsets of \mathcal{I}_n excluding the empty set. For a set \mathcal{A} , we write its cardinality as $|\mathcal{A}|$. For simplicity, we state the results only for discrete random variables with finite alphabets for which the entropy functions are always well-defined, however, they can be extended to continuous random variables (or mixed random variables) with well-defined differential entropy functions.

Theorem 1: Let $\{W_{\mathcal{A}}, \mathcal{A} \in \hat{\mathcal{P}}(\mathcal{I}_n)\}$ be a set of $2^n - 1$ discrete random variables jointly distributed according to some

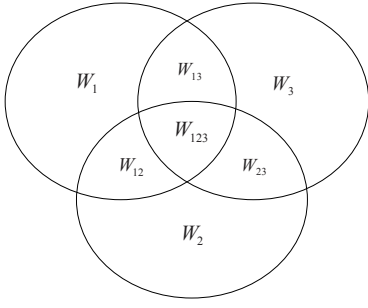


Fig. 1. Subsets and their relation.

probability mass function on some finite alphabets. Define

$$\bar{h}_k^{(n)} = \frac{1}{k \binom{n}{k}} \sum_{\mathcal{A} \in \hat{\mathcal{P}}(\mathcal{I}_n): |\mathcal{A}|=k} \sum_{j=1}^k H(W_{\mathcal{A}}(j)), \quad (3)$$

where

$$W_{\mathcal{A}}(j) \triangleq \{W_{\mathcal{B}} : |\mathcal{B} \cap \mathcal{A}| \geq j\} \quad (4)$$

Then

$$\bar{h}_1^{(n)} \geq \bar{h}_2^{(n)} \geq \dots \geq \bar{h}_n^{(n)}. \quad (5)$$

We can understand the set of random variables

$$\{W_{\mathcal{B}} : i \in \mathcal{B}\}, \quad i = 1, 2, \dots, n, \quad (6)$$

as n random variables with certain common components, or as subsets of random variables instead of simple random variables; see. Fig. 1 for an illustration when $n = 3$. To make this more explicit, it is beneficial to present an alternative statement of Theorem 1.

Theorem 1': Let $\{W_i, i = 1, 2, \dots, N\}$ be a set of random variables jointly distributed according to some probability mass function on some finite alphabets. Let $\mathcal{C}_i, i = 1, 2, \dots, n$ be subsets of \mathcal{I}_N . Define

$$\bar{h}_k^{(n)} = \frac{1}{k \binom{n}{k}} \sum_{\mathcal{A} \in \hat{\mathcal{P}}(\mathcal{I}_n): |\mathcal{A}|=k} \sum_{j=1}^k H(W'_{\mathcal{A}}(j)) \quad (7)$$

where

$$W'_{\mathcal{A}}(j) \triangleq \left\{ W_k, k \in \bigcup_{\{i_1, i_2, \dots, i_j\} \subseteq \mathcal{A}} (\mathcal{C}_{i_1} \cap \mathcal{C}_{i_2} \cap \dots \cap \mathcal{C}_{i_j}) \right\}, \quad (8)$$

then

$$\bar{h}_1^{(n)} \geq \bar{h}_2^{(n)} \geq \dots \geq \bar{h}_n^{(n)}. \quad (9)$$

It is straightforward to verify that Theorem 1 and Theorem 1' are equivalent. This result essentially provides the relation among the average entropy rates of the fixed cardinality intersections of subsets of random variables, however with special consideration of the specified common parts. It should be noted that there is a layered structure in $W'_{\mathcal{A}}(j)$, which is generated by taking intersection of different number of subsets

of a set.

The following results are complementary to Theorem 1, in a similar manner as those to Han's inequality [2] (pp. 492-493).

Theorem 2: Let $\{W_{\mathcal{A}}, \mathcal{A} \in \hat{\mathcal{P}}(\mathcal{I}_n)\}$ be a set of $2^n - 1$ discrete random variables jointly distributed according to some probability mass function. Define

$$\bar{g}_k^{(n)} = \frac{1}{k \binom{n}{k}} \sum_{\mathcal{A} \in \hat{\mathcal{P}}(\mathcal{I}_n): |\mathcal{A}|=k} \sum_{j=1}^k H(W_{\mathcal{A}}(j) | W_{\mathcal{A}^c}^c(j)), \quad (10)$$

where $W_{\mathcal{A}^c}^c(j) = W_{\mathcal{I}_n}(1) \setminus W_{\mathcal{A}}(j)$. Then

$$\bar{g}_1^{(n)} \leq \bar{g}_2^{(n)} \leq \dots \leq \bar{g}_n^{(n)}. \quad (11)$$

Combining Theorem 1 and Theorem 2, we have the following theorem.

Theorem 3: Let $\{W_{\mathcal{A}}, \mathcal{A} \in \hat{\mathcal{P}}(\mathcal{I}_n)\}$ be a set of $2^n - 1$ discrete random variables jointly distributed according to some probability mass function. Define

$$\bar{f}_k^{(n)} = \frac{1}{k \binom{n}{k}} \sum_{\mathcal{A} \in \hat{\mathcal{P}}(\mathcal{I}_n): |\mathcal{A}|=k} \sum_{j=1}^k I(W_{\mathcal{A}}(j); W_{\mathcal{A}^c}^c(j)) \quad (12)$$

Then

$$\bar{f}_1^{(n)} \geq \bar{f}_2^{(n)} \geq \dots \geq \bar{f}_n^{(n)}. \quad (13)$$

As a special case to Theorem 1, we can recover Han's inequality on subsets, by letting $W_{\mathcal{A}}$ be a constant for all $|\mathcal{A}| > 1$, for which $W_{\{i\}}$ can be conveniently written as W_i .

Proposition 4: Let $\{W_i\}$ be a set of n discrete random variables jointly distributed according to some probability mass function. Define

$$h_k^{(n)} = \frac{1}{k \binom{n}{k}} \sum_{\mathcal{A} \in \hat{\mathcal{P}}(\mathcal{I}_n): |\mathcal{A}|=k} H(\{W_i : i \in \mathcal{A}\}). \quad (14)$$

Then

$$h_1^{(n)} \geq h_2^{(n)} \geq \dots \geq h_n^{(n)}. \quad (15)$$

As another special case to Theorem 1, we can recover the n -way sub-modularity inequality [7], which is the following proposition.

Proposition 5: Let $\{W_{\mathcal{A}}, \mathcal{A} \in \hat{\mathcal{P}}(\mathcal{I}_n)\}$ be a set of $2^n - 1$ discrete random variables jointly distributed according to some probability mass function. Then

$$\sum_{i=1}^n H(\{W_{\mathcal{B}} : i \in \mathcal{B}\}) \geq \sum_{j=1}^n H(\{W_{\mathcal{B}} : |\mathcal{B}| \geq j\}). \quad (16)$$

The left hand side of (16) is simply $\bar{h}_1^{(n)}$ and the right hand side is simply $\bar{h}_n^{(n)}$. For this reason, it is tempting to state Proposition 5 as a corollary of Theorem 1. However, we refrain from doing so because Proposition 5 in fact plays an instrumental role in our proof of Theorem 1. The proof of Proposition 5 can be found in [7].

It should be noted that though Han's inequality and the n -way sub-modularity inequality can be recovered this way, this does not mean Theorem 1 is stronger. Indeed the new

$$\begin{aligned}
& \frac{1}{3} \left[H(W_1, W_{12}, W_{13}, W_{123}) + H(W_2, W_{12}, W_{23}, W_{123}) + H(W_3, W_{13}, W_{23}, W_{123}) \right] \\
& \geq \frac{1}{6} \left[H(W_1, W_2, W_{12}, W_{13}, W_{23}, W_{123}) + H(W_{12}, W_{123}) + H(W_1, W_3, W_{12}, W_{13}, W_{23}, W_{123}) + H(W_{13}, W_{123}) \right. \\
& \quad \left. + H(W_2, W_3, W_{12}, W_{13}, W_{23}, W_{123}) + H(W_{23}, W_{123}) \right] \\
& \geq \frac{1}{3} \left[H(W_1, W_2, W_3, W_{12}, W_{13}, W_{23}, W_{123}) + H(W_{12}, W_{13}, W_{23}, W_{123}) + H(W_{123}) \right] \tag{17}
\end{aligned}$$

$$\begin{aligned}
& \frac{1}{3} \left[H(W_1, W_{12}, W_{13}, W_{123}) + H(W_2, W_{12}, W_{23}, W_{123}) + H(W_3, W_{13}, W_{23}, W_{123}) \right] \\
& \geq \frac{1}{6} \left[H(W_1, W_2, W_{12}, W_{13}, W_{23}, W_{123}) + H(W_1, W_3, W_{12}, W_{13}, W_{23}, W_{123}) + H(W_2, W_3, W_{12}, W_{13}, W_{23}, W_{123}) \right] \\
& \geq \frac{1}{3} H(W_1, W_2, W_3, W_{12}, W_{13}, W_{23}, W_{123}) \tag{18}
\end{aligned}$$

set of inequalities, or rather its general form on sub-modular functions (given in the next section) can be specialized to the n -way sub-modularity inequality, however, since the n -way sub-modularity inequality is sufficient to prove the new set of inequalities, they are completely equivalent for general sub-modular functions. The new set of inequalities on entropy function can be specialized to Han's equality, yet the conditional version of Han's inequality implies the sub-modularity of the entropy function, and thus the new set of inequalities in its entropy form and the conditional version of Han's inequality are equivalent. In the language of classification of information inequalities, these inequalities are all Shannon type inequalities [10]. Despite this relationship, it is our hope that this new set of inequalities may find its applications in the future, just as Han's inequality finds applications in [3]–[5] some 20 or 30 years after its discovery.

Example: Consider $n = 3$, and the random variables in Theorem 1 are

$$\{W_1, W_2, W_3, W_{12}, W_{13}, W_{23}, W_{123}\},$$

where for simplicity we have written $W_{\{1,2,3\}}$ as W_{123} , and similarly for $W_{\{i,j\}}$ and $W_{\{i\}}$; see Fig. 1. Theorem 1 then gives (17) on the top of this page. We can also naively apply Han's inequality on the three sets of random variables, and treat each of them as a random variable,

$$\begin{aligned}
T_1 &= \{W_1, W_{12}, W_{13}, W_{123}\} \\
T_2 &= \{W_2, W_{12}, W_{23}, W_{123}\} \\
T_3 &= \{W_3, W_{13}, W_{23}, W_{123}\},
\end{aligned}$$

which gives (18) on the top of this page. Note that when all the random variables are independent, (17) becomes equalities, however, (18) will not give equalities in general. The form of (17) is not obvious, particularly the coefficients of the entropies are not easy to determine a priori.

III. PROOFS OF THE THEOREMS

It is well known that the entropy function is sub-modular, thus the following theorem regarding sub-modular functions suffices for the proof of Theorem 1. For notational simplicity, we reserve the calligraphic letters \mathcal{A} , \mathcal{A}' and \mathcal{B} for non-empty subset of \mathcal{I}_n , i.e., $\mathcal{A} \in \hat{\mathcal{P}}(\mathcal{I}_n)$ (and similarly for \mathcal{A}' and \mathcal{B}), and will not explicitly include this condition from here on.

Theorem 6: Let θ be a submodular function on sets, i.e., for any two sets \mathcal{C}_1 and \mathcal{C}_2 ,

$$\theta(\mathcal{C}_1) + \theta(\mathcal{C}_2) \geq \theta(\mathcal{C}_1 \cup \mathcal{C}_2) + \theta(\mathcal{C}_1 \cap \mathcal{C}_2). \tag{19}$$

Define

$$\theta_k^{(n)} \triangleq \frac{1}{k \binom{n}{k}} \sum_{\mathcal{A}: |\mathcal{A}|=k} \sum_{j=1}^k \theta(\{\mathcal{B} : |\mathcal{B} \cap \mathcal{A}| \geq j\}), \tag{20}$$

then

$$\theta_1^{(n)} \geq \theta_2^{(n)} \geq \dots \geq \theta_n^{(n)}, \tag{21}$$

Note here \mathcal{C}_1 and \mathcal{C}_2 should be understood as sets of subsets in the context of Theorem 1.

We will need the following lemma from [7] to prove Theorem 6. It is clear that Proposition 5 follows directly from this lemma.

Lemma 1 ([7] n -way sub-modularity): Let θ be a sub-modular function, then

$$\sum_{i=1}^n \theta(\mathcal{C}_i) \geq \sum_{k=1}^n \theta(\mathcal{C}_{k,n}), \tag{22}$$

where

$$\mathcal{C}_{k,n} \triangleq \bigcup_{\{j_1, j_2, \dots, j_k\} \subseteq \mathcal{I}_n} (\mathcal{C}_{j_1} \cap \mathcal{C}_{j_2} \cap \dots \cap \mathcal{C}_{j_k}). \tag{23}$$

Proof of Theorem 6:

We first prove that $\theta_n^{(n)} \leq \theta_{n-1}^{(n)}$. For notational simplicity, \mathcal{A}_i is used to denote $\mathcal{I}_n \setminus \{i\}$ for $i = 1, 2, \dots, n$. By Lemma 1 we can write for a fixed j that

$$\sum_{i=1}^n \theta(\{\mathcal{B} : |\mathcal{B} \cap \mathcal{A}_i| \geq j\}) \geq \sum_{i=1}^n \theta(\mathcal{C}_i^{(j)}), \quad (24)$$

where

$$\mathcal{C}_i^{(j)} \triangleq \bigcup_{|\mathcal{A}'|=i} \bigcap_{m \in \mathcal{A}'} \{\mathcal{B} : |\mathcal{B} \cap \mathcal{A}_m| \geq j\}. \quad (25)$$

We claim that

$$\mathcal{C}_i^{(j)} = \begin{cases} \{\mathcal{B} : |\mathcal{B}| \geq j\} & i+j \leq n \\ \{\mathcal{B} : |\mathcal{B}| \geq j+1\} & i+j > n \end{cases} \quad (26)$$

To see this, observe that due to its complete symmetry, $\mathcal{C}_i^{(j)}$ must be in the form $\{\mathcal{B} : |\mathcal{B}| \geq v\}$ for some integer v . Clearly $v \geq j$ because of the special form of $\mathcal{C}_i^{(j)}$. Furthermore, since any \mathcal{B} that has at least $j+1$ elements in \mathcal{I}_n , it must have at least j elements in \mathcal{A}_i for $i = 1, 2, \dots, n$, i.e.,

$$\{\mathcal{B} : |\mathcal{B} \cap \mathcal{A}_i| \geq j\} \supseteq \{\mathcal{B} : |\mathcal{B}| \geq j+1\}, \quad (27)$$

it is clear that $v \leq j+1$. Next we show that $v = j$ if and only if $i+j \leq n$.

If $i+j \leq n$, then let us consider a special choice of \mathcal{A}'

$$\mathcal{A}' = \{n, n-1, \dots, n-i+1\}, \quad (28)$$

and thus for this choice

$$\mathcal{I}_j \in \bigcap_{m \in \mathcal{A}'} \{\mathcal{B} : |\mathcal{B} \cap \mathcal{A}_m| \geq j\} \quad (29)$$

because

$$\mathcal{I}_j \in \{\mathcal{B} : |\mathcal{B} \cap \mathcal{A}_m| \geq j\}, \quad m = n, n-1, \dots, n-i+1. \quad (30)$$

Thus $v = j$ for this case.

Next we show that if $i+j > n$, then there does not exist a set \mathcal{A}' such that (29) can hold, and it will follow that $\mathcal{I}_j \notin \mathcal{C}_i^{(j)}$, and thus $v = j+1$ for this case. We prove this by contradiction. For (29) to hold, it must be true that

$$\mathcal{I}_j \in \{\mathcal{B} : |\mathcal{B} \cap \mathcal{A}_m| \geq j\}, \quad m \in \mathcal{A}'. \quad (31)$$

This implies that

$$\mathcal{I}_j \subseteq \mathcal{A}_m, \quad m \in \mathcal{A}'. \quad (32)$$

However there are only a total of $n-j$ possible choices of \mathcal{A}_m such that $\mathcal{I}_j \subseteq \mathcal{A}_m$, but at the same time $|\mathcal{A}'| = i > n-j$ by assumption. Thus a contradiction arises by the pigeonhole principle. Thus (26) is indeed true.

We can now write

$$\begin{aligned} n(n-1)\theta_{n-1}^{(n)} &= \sum_{i=1}^n \sum_{j=1}^{n-1} \theta(\{\mathcal{B} : |\mathcal{B} \cap \mathcal{A}_i| \geq j\}) \\ &= \sum_{j=1}^{n-1} \sum_{i=1}^n \theta(\{\mathcal{B} : |\mathcal{B} \cap \mathcal{A}_i| \geq j\}) \\ &\geq \sum_{j=1}^{n-1} (n-j)\theta(\{\mathcal{B} : |\mathcal{B}| \geq j\}) + j\theta(\{\mathcal{B} : |\mathcal{B}| \geq j+1\}) \\ &= (n-1) \sum_{j=1}^n \theta(\{\mathcal{B} : |\mathcal{B}| \geq j\}) = n(n-1)\theta_n^{(n)}, \end{aligned} \quad (33)$$

which completes the proof for $\theta_n^{(n)} \leq \theta_{n-1}^{(n)}$.

To prove $\theta_k^{(n)} \leq \theta_{k-1}^{(n)}$ for $k \leq n$, we take a similar approach as used in [2] (p. 491). We first limit the choice of \mathcal{A} and \mathcal{B} to a k -element subset of \mathcal{I}_n , and then taking a uniform choice over its $(k-1)$ -element subsets. For each such k -element subset, $\theta_k^{(k)} \leq \theta_{k-1}^{(k)}$, and the inequality remains true after taking the expectation over all k -element subset chosen uniformly from \mathcal{I}_n . This completes the proof. ■

Theorem 1 now directly follows from Theorem 6. Clearly, if the sub-modular functions in Theorem 6 are replaced by super-modular functions, the direction of the inequalities are simply reversed. With this observation, Theorem 2 follows directly from the super-modularity of the conditional entropy function $H(\mathcal{C}|\mathcal{C}^c)$, where $\mathcal{C}^c = \Omega \setminus \mathcal{C}$ is the complement of set \mathcal{C} . This can be seen as follows.

$$\begin{aligned} &H(\mathcal{C}_1|\mathcal{C}_1^c) + H(\mathcal{C}_2|\mathcal{C}_2^c) \\ &= 2H(\Omega) - H(\mathcal{C}_1^c) - H(\mathcal{C}_2^c) \\ &\leq 2H(\Omega) - H((\mathcal{C}_1 \cap \mathcal{C}_2)^c) - H((\mathcal{C}_1 \cup \mathcal{C}_2)^c) \\ &= H(\mathcal{C}_1 \cap \mathcal{C}_2 | (\mathcal{C}_1 \cap \mathcal{C}_2)^c) + H(\mathcal{C}_1 \cup \mathcal{C}_2 | (\mathcal{C}_1 \cup \mathcal{C}_2)^c), \end{aligned} \quad (34)$$

where the inequality is by the sub-modularity of the entropy function and the fact

$$\mathcal{C}_1^c \cap \mathcal{C}_2^c = (\mathcal{C}_1 \cup \mathcal{C}_2)^c \quad \mathcal{C}_1^c \cup \mathcal{C}_2^c = (\mathcal{C}_1 \cap \mathcal{C}_2)^c.$$

The proofs of Theorem 3, Proposition 4 and Proposition 5 are trivial, and thus omitted here.

IV. CONCLUDING REMARKS

We provided a new set of inequalities, which generalizes the classic Han's inequality on entropy rates of subsets of random variables to include common components. These inequalities provide a link between Han's inequality and the n -way sub-modularity inequality, both of which belong to this new set of inequalities as special cases.

Admittedly, this new set of inequality has yet to find applications in information theoretical problems. These inequalities, however, are indeed motivated by one of author's recent works [11], and the consideration to include common components comes from the common message requirement in broadcast channel. It appears possible to simplify the rather lengthy proof given in [11] with the help of this new set of information

inequalities. It remains to be seen whether this new set of inequalities can find its application elsewhere.

REFERENCES

- [1] T. S. Han, "Nonnegative entropy measures of multivariate symmetric correlations," *Inform. Contr.*, vol. 36, no. 2, pp. 133–156, 1978.
- [2] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, New York: Wiley, 1991.
- [3] A. Albanese, J. Blomer, J. Edmonds, M. Luby, and M. Sudan, "Priority encoding transmission," *IEEE Trans. Information Theory*, vol. 42, no. 6, pp. 1737–1744, Nov. 1996.
- [4] H. Wang and P. Viswanath, "Vector Gaussian multiple description with two levels of receivers," *IEEE Trans. Information Theory*, vol. 55, no. 1, pp. 401–410, Jan. 2009.
- [5] C. Tian, S. Mohajer, and S. Diggavi, "Approximating the Gaussian multiple description rate region under symmetric distortion constraints," *IEEE Trans. Information Theory*, vol. 55, no. 8, pp. 3869–3891, Aug. 2009.
- [6] S. Fujishige, "Polymatroidal dependence structure of a set of random variables," *Inform. Contr.*, vol. 39, no. 1, pp. 55–72, Oct. 1978.
- [7] H. J. A. Harvey, R. Kleinberg, and A. R. Lehman, "On the capacity of information networks," *IEEE Trans. Information Theory*, vol. 52, no. 6, pp. 2345–2364, Jun. 2006.
- [8] F. R. K. Chung, R. L. Graham, P. Frankl, and J. B. Shearer, "Some intersection theorems for ordered sets and graphs," *J. Combinatorial Theory, Ser. A*, vol. 52, pp. 2337, 1986.
- [9] M. Madiman and P. Tetali, "Information inequalities for joint distributions, with interpretations and applications," *IEEE Trans. Information Theory*, vol. 56, no. 6, pp. 2699–2713, Jun. 2010.
- [10] R. Yeung, *A First Course in Information Theory*, Kluwer Academic/Plenum Publishers, New York, 2002.
- [11] C. Tian, "Latent capacity region: a case study on symmetric broadcast with common messages," *IEEE Trans. Information Theory*, vol. 57, no. 6, pp. 3273–3285, Jun. 2011.