Characteristic-Dependent Linear Rank Inequalities and Network Coding Applications

Randall Dougherty, Eric Freiling, and Kenneth Zeger

***Abstract*—Two characteristic-dependent linear rank inequalities are given for eight variables. Specifically, the first inequality holds for all finite fields whose characteristic is not three and does not in general hold over characteristic three. The second inequality holds for all finite fields whose characteristic is three and does not in general hold over characteristics other than three. Applications of these inequalities to the computation of capacity upper bounds in network coding are demonstrated.**

I. Introduction

The study of information inequalities is a subfield of information theory that describes linear constraints on the entropies of finite collections of jointly distributed discrete random variables. Historically, the known information inequalities were orignally all special cases of Shannon's conditional mutual information inequality $I(X;Y|Z) \geq 0$, but later were generalized to other types of inequalities, called non-Shannon inequalities. Information inequalities have been shown to be useful for computing upper bounds on the network coding capacities of certain networks.

Analagously, the study of linear rank inequalities is a topic of linear algebra, which describes linear constraints on the dimensions of collections of subspaces of finite dimensional vector spaces. In fact, the set of all information inequalities can be viewed as subclass of the set of all linear rank inequalities.

Information inequalities hold over all collections of a certain number of random variables. In constrast, linear rank inequalities may hold over only certain vector spaces, such as those whose scalars have particular field characteristics.

This work was supported by the Institute for Defense Analyses and the National Science Foundation.

R. Dougherty is with the Center for Communications Research, 4320 Westerra Court, San Diego, CA 92121-1969 (rdough@ccrwest.org).

E. Freiling and K. Zeger are with the Department of Electrical and Computer Engineering, University of California, San Diego, La Jolla, CA 92093-0407 (efreilin@ucsd.edu, zeger@ucsd.edu).

This paper appeared in the proceedings of the IEEE International Symposium on Information Theory (ISIT), held in Honolulu, Hawaii June-July 2014.

In this paper, we present two new linear rank inequalities over finite fields, which are not information inequalities, and with the peculiar property that they only hold for certain fields, depending on the associated vector space. The first inequality is shown to hold over all vector spaces when the field characteristic is anything but three (Theorem II.1), but does not always hold when the field characteristic is three (Theorem II.2). In contrast, the second inequality is shown to hold over all vector spaces when the field characteristic is three (Theorem III.1), but does not always hold when the field characteristic is not three (Theorem III.2). We also show how these inequalities can be used to obtain bounds on the capacities of certain networks (Corollaries II.3 and III.3).

Due to space limitations, we omit the proofs of results stated in this paper. The full proofs of all such results can be found in [11].

A. Background

In 2000, Ahlswede, Cai, Li, and Yeung introduced the field of Network Coding [1] and showed that coding can outperform routing in directed acyclic networks.[1] There are presently no known algorithms to determine the capacity or the linear capacity of a given network. In fact, it is not even known if such algorithms exist.

Information inequalities are linear inequalities that hold for all jointly distributed random variables, and Shannon inequalities are information inequalities of a certain form [19]. Both are defined in Section I-B. It is known [22] that all information inequalities containing three or fewer variables are Shannon inequalities. The first "non-Shannon" information inequality was of four variables and was published in 1998 by Zhang and Yeung [25]. Since 1998, various other non-Shannon inequalities have been found, for example, by Lněnička [14], Makarychev, Makarychev, Romashchenko, and Vereshchagin [15], Zhang [23], Zhang and Yeung [24], Dougherty, Freiling, and Zeger [5], and

[1] In what follows, by "network" we shall always mean a directed acyclic network.

Matúš [16]. Additionally, in 2007, Matúš demonstrated an infinite collection of independent non-Shannon information inequalities [16] and there were necessarily an infinite number of such inequalities. In 2008, Xu, Wang, and Sun [20] also gave an infinite list of inequalities but did not establish their necessity.

There is a close connection between information inequalities and network coding [4]. Capacities of some networks have been computed by finding matching lower and upper bounds [6]. Lower bounds have been found by deriving coding solutions. Upper bounds have been found by using information inequalities and treating the sources as independent random variables that are uniformly distributed over the alphabet. One "holy grail" problem of network coding is to develop an algorithm to compute the coding capacity of an arbitrary network. If such an algorithm exists, information inequalities may potentially play a role in the solution.

It has been shown that linear codes are insufficient for network coding in general [7]. However, linear codes may be desirable to use in practice due to ease of analysis and implementation. It has been shown that the coding capacity is independent of the alphabet size [3]. However, the linear coding capacity is dependent on alphabet size, or more specifically the field characteristic. In other words, one can potentially achieve a higher rate of linear communication by choosing one characteristic over another. To provide upper bounds for the linear coding capacity for a particular field one can look at linear rank inequalities [10]. Linear rank inequalities are linear inequalities that are always satisfied by ranks[2] of subspaces of a vector space. All information inequalities are linear rank inequalities but not all linear rank inequalities are information inequalities. The first example of a linear rank inequality that is not an information inequality was found by Ingleton [13]. Information inequalities can provide an upper bound for the capacity of a network, but this upper bound would hold for all alphabets. Therefore, to determine the linear coding capacity over a certain characteristic one would have to consider linear rank inequalities.

All linear rank inequalities up to and including five variables are known and none of these depend on the vector spaces' field characteristics [8]. The set of all linear rank inequalities for six variables has not yet been determined. Characteristic-dependent linear rank inequalities are given, for example, in [2] and [10].

An inequality is given in [10] which is valid for characteristic two and another inequality is given which is valid for every characteristic except for two. These inequalities are then used to provide upper bounds for the linear coding capacity of two networks.

In the present paper, we give two characteristic-dependent linear rank inequalities on eight variables. One is valid for characteristic three and the other is valid for every characteristic except for three. These inequalities are then used to provide upper bounds for the linear coding capacity of two networks.

It is our intention that the techniques presented here may prove useful or otherwise motivate further progress in determining network capacities.

B. Information Theory and Linear rank Inequalities

Let A, B, C be collections of discrete random variables over a finite alphabet $\mathcal{X}$, and let p be the probability mass function of A. The *entropy* of A is defined by

$$H(A) = -\sum_u p(u) \log_{|\mathcal{X}|} p(u).$$

The *conditional entropy* of A given B is

$$H(A|B) = H(A, B) - H(B), \tag{1}$$

the *mutual information* between A and B is

$$I(A; B) = H(A) - H(A|B), \tag{2}$$

and the *conditional mutual information* between A and B given C is

$$I(A; B|C) = H(A|C) - H(A|B, C). \tag{3}$$

Definition I.1. Let q be a positive integer, and let $S_1, \ldots, S_k$ be subsets of $\{1, \ldots, q\}$. Let $\alpha_i \in \mathbb{R}$ for $1 \leq i \leq k$. A linear inequality of the form

$$\alpha_1 H(\{A_i : i \in S_1\}) + \cdots + \alpha_k H(\{A_i : i \in S_k\}) \geq 0 \tag{4}$$

is called an *information inequality* if it holds for all jointly distributed random variables $A_1, \ldots, A_q$.

A *Shannon information inequality* is any information inequality that can be expressed as a finite sum of the form

$$\sum_i \alpha_i I(A_i; B_i|C_i) \geq 0$$

where each α_i is a nonnegative real number. Any information inequality that cannot be expressed in the form above will be called a *non-Shannon information inequality*.

Linear rank inequalities are closely related to information inequalities. In fact, in order to describe linear rank

[2]Throughout this paper, we will use the terminology "rank" of a subspace to mean the dimension of the subspace (i.e. the rank of a matrix whose columns are a basis for the subspace), in order to parallel the terminology of matroid theory.

inequalities we will borrow notation from information theory to use in the context of linear algebra in the following manner.

Suppose A and B are subspaces of a given vector space V, and let $\langle A, B\rangle$ denote the span of $A \cup B$. We will let $H(A)$ denote the rank of A, and let $H(A, B)$ denote the rank of $\langle A, B\rangle$. The meanings of some other information theoretic notation in the context of linear algebra then follows from (1)-(3). Specifically, note that the conditional entropy notation $H(A|B)$ denotes the excess rank of subspace A over that of subspace $A \cap B$, or equivalently, the codimension of $A \cap B$ in A; and the mutual information notation $I(A; B)$ denotes the rank of $A \cap B$.

A *linear rank inequality* over a vector space V is a linear inequality of the form in (4), that is satisfied by every assignment of subspaces of V to the variables $A_1, \ldots, A_q$.

All information inequalities are linear rank inequalities over all finite vector spaces, but not all linear rank inequalities are information inequalities. For background material on these concepts, the reader is referred to Hammer, Romashchenko, Shen, and Vereshchagin [12].

The first known example of a linear rank inequality over all finite vector spaces that is not an information inequality is the *Ingleton inequality* [13]:

$$I(A; B) \leq I(A; B|C) + I(A; B|D) + I(C; D).$$

C. Network Coding

A *network* is a finite, directed, acyclic multigraph with messages and demands. Network *messages* are arbitrary vectors of k symbols over a finite alphabet $\mathcal{A}$. Each network edge carries a vector of n symbols from $\mathcal{A}$. Each message originates at a particular node called the *source node* for that message and is required by one or more *demand nodes*. When we draw a network, a message variable appearing above a node indicates the message is generated by such node[3], and a message variable appearing below a node indicates the message is demanded by such node, For a given network, the values of k and n can be chosen in order to implement certain codes and to obtain certain throughput k/n.

The inputs to a network node are the vectors carried on its in-edges as well as the messages, if any, generated at the node. The outputs of a network node are the packets carried on its out-edges as well as any demanded messages at the node. Each output of a node must be a function only of its inputs. A *coding solution* for the network is an assignment of such functions to the network edges. When the values of k and n need to be emphasized, the coding solution will be called a (k, n)-coding solution. The *capacity* of a network is defined as:

$$\mathcal{C} = \sup\{k/n : \exists \text{ a } (k, n)\text{-coding solution}\}.$$

A solution is called a *linear solution*, if the alphabet $\mathcal{A}$ is a finite field and the edge functions are linear (i.e. linear combinations of their input vectors where the coefficients are matrices over the field).

The *linear capacity* is defined the same as the capacity but restricting solutions to be linear. It is also easily verified that if x is a message, then $H(x) = k$, and if x is a vector carried by an edge, then $H(x) \leq n$.

Information inequalities can be used to obtain capacity bounds for networks. In the proofs of Corollaries II.3 and III.3, we obtain bounds on the capacities of networks by using linear rank inequalities, instead of information inequalities. In those cases, certain vector subspaces will be used instead of random variables.

II. A Linear Rank Inequality for fields of characteristic other than 3

In this section, we use the known T8 matroid to construct a "T8 network", and then in turn we use the T8 network to guide a construction of a "T8 linear rank inequality" that is shown to hold for all vector spaces having finite scalar fields of characteristic not equal to 3. Then we show that the T8 inequality does not necessarily hold when such scalar fields have characteristic 3. Finally, we determine the exact coding capacity of the T8 network and its linear coding capacity over finite field alphabets of characteristic 3, as well as a linear capacity upper bound for finite field alphabets whose characteristic is not 3.

The T8 matroid [18] is a vector matroid which is represented by the following matrix, where column dependencies are over characteristic 3:

$$\begin{array}{c} A \quad B \quad C \quad D \quad W \quad X \quad Y \quad Z \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \end{pmatrix}. \end{array}$$

The T8 matroid is representable over a field if and only if the field is of characteristic 3. Figure 1 is a network whose dependencies and independencies are consistent with the T8 matroid. It was designed by the construction

[3]We note that in Figures 1 and 2, for convenience, we label source messages above nodes lying in both the top and bottom layers in each diagram. This is meant to indicate that there is, in fact, a separate (but hidden) distinct node for each such source message, whose out-edges go directly to the nodes labeled by the source message in the top and bottem layers.

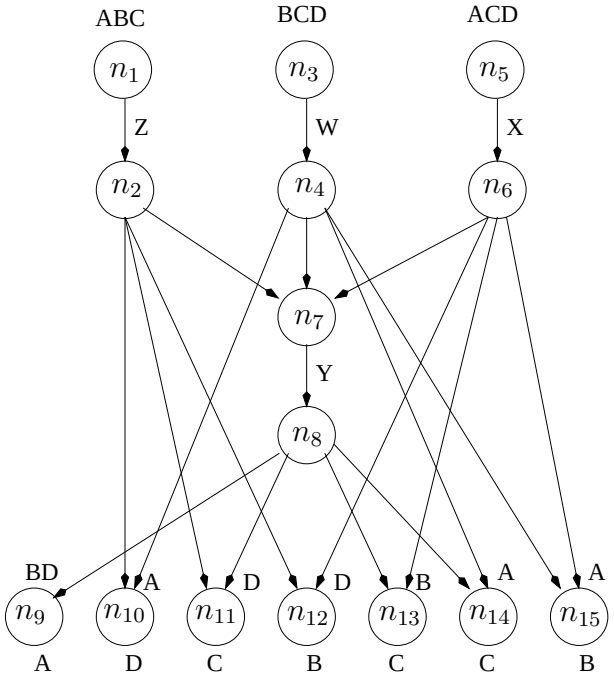

Fig. 1. The T8 network has source messages A, B, C, and D generated at hidden source nodes with certain hidden out-edges pointing to corresponding displayed nodes n_1, n_3, n_5, and n_9–n_{15} (which are labeled by incoming messages above such nodes). The nodes n_9–n_{15} each demand one message, as labeled below such nodes.

process described in [6], and we will refer to it as the T8 network. Theorem II.1 uses the T8 network as a guide to derive a linear rank inequality valid for every characteristic except for 3. We refer to the inequality in the following theorem as the *T8 linear rank inequality*.

Theorem II.1. *Let A, B, C, D, W, X, Y, and Z be subspaces of a vector space V whose scalar field is finite and of characteristic other than 3. Then the following is a linear rank inequality over V:*

$$\begin{aligned}
&H(A) \leq 8H(Z)+29H(Y)+3H(X)+8H(W)-6H(D)\\
&-17H(C)-8H(B)-17H(A)+55H(Z|A,B,C)\\
&+35H(Y|W,X,Z)+50H(X|A,C,D)\\
&+49H(W|B,C,D)+18H(A|B,D,Y)+7H(B|D,X,Z)\\
&+H(B|A,W,X)+7H(C|D,Y,Z)+7H(C|B,X,Y)\\
&+3H(C|A,W,Y)+6H(D|A,W,Z)\\
&+49(H(A)+H(B)+H(C)+H(D)-H(A,B,C,D)).
\end{aligned}$$

The next theorem demonstates that the inequality in Theorem II.1 does not in general hold for vector spaces with finite fields of characteristic 3.

Theorem II.2. *There exists a vector space V with a finite scalar field of characteristic 3 such that the T8 inequality in Theorem II.1 is not a linear rank inequality over V.*

The following corollary uses the T8 linear rank inequality to derive capacities and a capacity bound on the T8 network. Note that although the T8 network itself was used as a guide in obtaining the T8 linear rank inequality, subsequently using the inequality to bound the network capacity is not circular reasoning.

The proof of Corollary II.3 makes use of the T8 linear rank inequality.

Corollary II.3. *For the T8 network, the linear coding capacity is at most $48/49$ over any finite field alphabet of characteristic not equal to 3. The linear coding capacity over finite field alphabets of characteristic 3 and the coding capacity are both equal to 1.*

III. A Linear Rank Inequality for Fields of Characteristic 3

In the T8 matroid, $W + X + Y + Z = (3, 3, 3, 3)$, which equals $(0, 0, 0, 0)$ in characteristic 3. We define the *non-T8 matroid* to be the T8 matroid except that we force the T8's characteristic 3 circuit $\{W, X, Y, Z\}$ to be a base in the non-T8 matroid. Figure 2 is a network that we call the *non-T8 network*, whose dependencies and independencies are consistent with the non-T8 matroid. The non-T8 network was designed by the construction process described in [6]. Theorem III.1 uses the non-T8 network as a guide to derive a linear rank inequality valid for characteristic 3. The new linear rank inequality can then be used to prove the non-T8 network has linear capacity less than 1 if the field characteristic is 3.

Theorem III.1. *Let A, B, C, D, W, X, Y, and Z be subspaces of a vector space V whose scalar field is finite and of characteristic 3. Then the following is a linear rank inequality over V:*

$$\begin{aligned}
&H(A) \leq 9H(Z)+8H(Y)+5H(X)+6H(W)-4H(D)\\
&-12H(C)-11H(B)-H(A)+19H(Z|A,B,C)\\
&+17H(Y|A,B,D)+13H(X|A,C,D)\\
&+11H(W|B,C,D)+H(A|W,X,Y,Z)+H(A|B,W,X)\\
&+7H(B|D,X,Z)+4H(B|C,X,Y)+7H(C|D,Y,Z)\\
&+5H(C|A,W,Y)+4H(D|A,W,Z)\\
&+29(H(A)+H(B)+H(C)+H(D)-H(A,B,C,D)).
\end{aligned}$$

The next theorem demonstates that the inequality in Theorem III.1 does not in general hold for vector spaces with finite fields of characteristic other than 3.

Theorem III.2. *For each prime number $p \neq 3$ there exists a vector space V with a finite scalar field of*

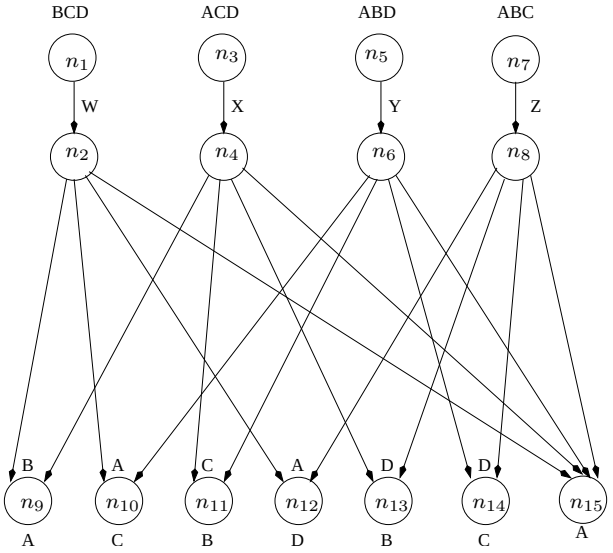

Fig. 2. The Non-T8 Network has source messages A, B, C, and D generated at hidden source nodes with certain hidden out-edges pointing to corresponding displayed nodes n_1, n_3, n_5, n_7, and n_9–n_{14} (which are labeled by incoming messages above such nodes). The nodes n_9–n_{15} each demand one message, as labeled below them.

characteristic p such that the non-T8 inequality in Theorem III.1 is not a linear rank inequality over V.

Corollary III.3. *For the non-T8 network, the linear coding capacity is at most $28/29$ over any finite field alphabet of characteristic equal to 3. The linear coding capacity over finite field alphabets of characteristic not 3 and the coding capacity are all equal to 1.*

REFERENCES

[1] R. Ahlswede, N. Cai, S.-Y.R. Li, and R.W. Yeung, "Network information flow," *IEEE Transactions on Information Theory*, vol. 46, no. 4, pp. 1204–1216, 2000.

[2] A. Blasiak, R. Kleinberg, E. Lubetzky, "Lexicographic products and the power of non-linear network coding", *arXiv* 1108.2489.

[3] J. Cannons, R. Dougherty, C. Freiling, and K. Zeger, "Network routing capacity", *IEEE Transactions on Information Theory*, vol. 52, no. 3, pp. 777-788, March 2006.

[4] T. Chan and A. Grant, "Entropy vectors and network codes," *IEEE International Symposium on Information Theory*, pp. 1586–1590, 2007.

[5] R. Dougherty, C. Freiling, and K. Zeger, "Six new non-Shannon information inequalities," *IEEE International Symposium on Information Theory*, pp. 233–236, 2006.

[6] R. Dougherty, C. Freiling, and K. Zeger, "Networks, matroids, and non-Shannon information inequalities," *IEEE Transactions on Information Theory*, vol. 53, no. 6, pp. 1949–1969, 2007.

[7] R. Dougherty, C. Freiling, and K. Zeger, "Insufficiency of linear coding in network information flow," *IEEE Transactions on Information Theory*, vol. 51, no. 8, pp. 2745–2759, 2005.

[8] R. Dougherty, C. Freiling, and K. Zeger, "Linear rank inequalities on five or more variables," *arXiv 0910.0284*, 2012.

[9] R. Dougherty, C. Freiling, and K. Zeger, "Linear network codes and systems of polynomial equations," *IEEE Transactions on Information Theory*, vol. 54, no. 5, pp. 2303–2316, 2008.

[10] R. Dougherty, C. Freiling, and K. Zeger, "Achievable Rate Regions for Network Coding" *IEEE Transactions on Information Theory*, submitted November 18, 2013.
Also *arXiv 1311.4601*, 2013.

[11] R. Dougherty, E. Freiling, and K. Zeger, "Characteristic-Dependent Linear Rank Inequalities with Applications to Network Coding," *IEEE Transactions on Information Theory*, submitted November 19, 2013.
Also *arXiv 1401.2507*, 2014.

[12] D. Hammer, A.E. Romashchenko, A. Shen, and N.K. Vereshchagin, "Inequalities for Shannon entropy and Kolmogorov complexity," *Journal of Computer and Systems Sciences*, vol. 60, pp. 442–464, 2000.

[13] A.W. Ingleton, "Representation of matroids," *Combinatorial Mathematics and its Applications*, pp. 149–167, 1971.

[14] R. Lněnička, "On the tightness of the Zhang-Yeung inequality for Gaussian vectors," *Communications in Information and Systems*, vol. 3, no. 1, pp. 41–46, 2003.

[15] K. Makarychev, Y. Makarychev, A. Romashchenko, and N. Vereshchagin, "A new class of non-Shannon-type inequalities for entropies," *Communications in Information and Systems*, vol. 2, no. 2, pp. 147–166, 2002.

[16] F. Matúš, "Infinitely many information inequalities," *IEEE International Symposium on Information Theory*, pp. 2101–2105, 2007.

[17] C. Ngai and R. Yeung, "Network coding gain of combination networks," *IEEE Information Theory Workshop*, pp. 283–287, 2004.

[18] J. Oxley, *Matroid Theory*, Oxford, New York, 1992.

[19] C. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. Journal*, vol. 27, pp. 379–423 and 623–656, 1948.

[20] W. Xu, J. Wang, and J. Sun, "A projection method for derivation of non-Shannon-type information inequalities," *IEEE International Symposium on Information Theory*, pp. 2116–2120, 2008.

[21] R. Yeung, *Information Theory and Network Coding*, Springer, 2008.

[22] R. Yeung, *A First Course in Information Theory*, Kluwer, Norwell, MA, 2002.

[23] Z. Zhang, "On a new non-Shannon type information inequality," *Communications in Information and Systems*, vol. 3, no. 1, pp. 47–60, 2003.

[24] Z. Zhang and R. Yeung, "A non-Shannon-type conditional inequality of information quantities," *IEEE Transactions on Information Theory*, vol. 43, pp. 1982–1985, 1997.

[25] Z. Zhang and R. Yeung, "On characterization of entropy function via information inequalities," *IEEE Transactions on Information Theory*, vol. 44, no. 4, pp. 1440–1452, 1998.