Characteristic-Dependent Linear Rank Inequalities With Applications to Network Coding

Randall Dougherty, Eric Freiling, and Kenneth Zeger

***Abstract*— Two characteristic-dependent linear rank inequalities are given for eight variables. In particular, 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.**

***Index Terms*— Shannon entropy, network coding, vector spaces, capacity.**

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.

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 3.1), but does not always hold when the field characteristic is three (Theorem 3.2). In contrast, the second inequality is shown to hold over all vector spaces when the field characteristic is three (Theorem 4.1), but does not always hold when the field characteristic is not three (Theorem 4.2). We also show how these inequalities can be used to obtain bounds on the capacities of certain networks (Corollaries 3.4 and 4.3).

It will be assumed that the reader has familiarity with linear algebra, finite fields, information theory, and network coding. Nevertheless, we will give some brief tutorial descriptions of these topics for completeness.

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 [18]. Both are defined in Section C. It is known [21] 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 [24]. Since 1998, various other non-Shannon inequalities have been found, for example, by Lněnička [13], Makarychev, Makarychev, Romashchenko, and Vereshchagin [14], Zhang [22], Zhang and Yeung [23], Dougherty, Freiling, and Zeger [5], and Matúš [15]. Additionally, in 2007, Matúš demonstrated an infinite collection of independent non-Shannon information inequalities [15] and there were necessarily an infinite number of such inequalities. In 2008, Xu, Wang, and Sun [19] 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

Manuscript received November 20, 2013; revised October 5, 2014; accepted January 12, 2015. Date of publication February 13, 2015; date of current version April 17, 2015. This work was supported in part by the U.S. National Science Foundation and in part by the Institute for Defense Analyses.

R. Dougherty is with the Center for Communications Research, San Diego, CA 92121 USA (e-mail: rdough@ccrwest.org).

E. Freiling and K. Zeger are with the Department of Electrical and Computer Engineering, University of California at San Diego, La Jolla, CA 92093 USA (e-mail: efreiling@gmail.com; zeger@ucsd.edu).

Communicated by Y. Liang, Associate Editor for Shannon Theory.

Digital Object Identifier 10.1109/TIT.2015.2403361

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

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 [12]. 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. Matroids

In this section a very brief review of matroids is given which will enable discussion in subsequent sections of a matroid-based method for constructing a particular network that helps in the derivation of the linear rank inequalities presented in this paper.

A matroid is an abstract structure that captures a notion of "independence" that is found in finite dimensional vector spaces, graphs, and various other mathematical topics. We will follow the notation and results of [17].

Definition 1.1: A *matroid*, M, is a pair (E, I), where E is a finite set and I is a set of subsets of E that satisfies the following properties:

(I1) $\varnothing \in I$.

(I2) $\forall A, B \subseteq E$, if $A \subseteq B \in I$, then $A \in I$.

(I3) $\forall A, B \subseteq E$, if $A, B \in I$ and $|A| > |B|$, then $\exists u \in A \setminus B$ such that $B \cup \{u\} \in I$.

[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.

The sets in I are called *independent sets*. If a subset of E is not in I, then it is called *dependent*.

An example of a matroid is obtained from linear algebra. Let F be a finite field and let $V(m, F)$ be the vector space of all m-dimensional vectors whose components are elements of F. Suppose A is an $m \times n$ matrix over F. Let $E = \{1, \ldots, n\}$ and I be the set of all $X \subseteq E$ such that the multiset of columns of A indexed by the elements of X is linearly independent in the vector space $V(m, F)$. Then $M = (E, I)$ is a matroid called the *vector matroid* of A.

A matroid is said to be *representable* over the field F if it is isomorphic to some vector matroid over $V(m, F)$.

For example, if F is the binary field and

$$A = \begin{array}{c} \begin{matrix} a & b & c & d & e \end{matrix} \\ \begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

where a, b, c, d, e denote the columns of A from left to right, then $M = (E, I)$ is a vector matroid of A, where $E = \{a, b, c, d, e\}$ and

$$I = \{\varnothing, \{a\}, \{b\}, \{d\}, \{e\}, \{a, b\}, \{a, e\}, \{b, d\}, \{b, e\}, \{d, e\}\}.$$

A *base* is a maximal independent set. Let $B(M)$ denote the set of all bases of a matroid M. In our example,

$$B(M) = \{\{a, b\}, \{a, e\}, \{b, d\}, \{b, e\}, \{d, e\}\}.$$

It is well known that all the bases of a matroid are of the same cardinality.

If we let $X \subseteq E$ and $I|X = \{i \subseteq X : i \in I\}$, then it is easy to see that $(X, I|X)$ is a matroid. The *rank* of X, denoted by $r(X)$, is defined to be the cardinality of a base in $M|X$. In our example, $r(M) = 2$. A *circuit* is a minimal dependent set. The circuits in our example are $\{\{c\}, \{a, d\}, \{a, b, e\}, \{b, d, e\}\}$.

C. Information Theory and Linear Rank Inequalities

In this section we will use the information theoretic concepts of entropy and mutual information to define and use the linear algebraic concept of linear rank inequalities. Connections between information inequalities and linear rank inequalities is also discussed.

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

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

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

$$\begin{aligned} I(A; B|C) &= H(A|C) - H(A|B, C) && (3) \\ &= H(A, C) + H(B, C) - H(C) - H(A, B, C). && (4) \end{aligned}$$

We will make use of the following basic information-theoretic facts [21]:

$$0 = H(\varnothing) \tag{5}$$
$$0 \leq H(A) = H(A|\varnothing) \tag{6}$$
$$0 \leq H(A|B) \tag{7}$$
$$0 \leq I(A;B) \tag{8}$$
$$H(A,B|C) \leq H(A|C) + H(B|C) \tag{9}$$
$$H(A|B,C) \leq H(A|B) \leq H(A,C|B) \tag{10}$$
$$I(A;B,C) = I(A;B|C) + I(A;C). \tag{11}$$

The equations (6)-(10) were originally given by Shannon in 1948 [18], and can all be obtained from the single inequality $I(A;B|C) \geq 0$.

Definition 1.2: 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

$$\alpha_1 H(\{A_i : i \in S_1\}) + \cdots + \alpha_k H(\{A_i : i \in S_k\}) \geq 0 \tag{12}$$

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

As an example, taking $q = 2$, $S_1 = \{1\}$, $S_2 = \{2\}$, $S_3 = \varnothing$, $S_4 = \{1,2\}$, $\alpha_1 = \alpha_2 = 1$, $\alpha_4 = -1$, and using (9) shows that $H(A_1) + H(A_2) - H(A_1, A_2) \geq 0$ is an information inequality.

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 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)-(4). 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 (12), 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 [11].

The first known example of a linear rank inequality over all finite vector spaces that is not an information inequality is the *Ingleton inequality* [12]:

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

To see that the Ingleton inequality is not an information inequality, let A, B, C, D be binary random variables, and let $X = (A, B, C, D)$ with probabilities:

$$P(X = 0000) = 1/4$$
$$P(X = 1111) = 1/4$$
$$P(X = 0101) = 1/4$$
$$P(X = 0110) = 1/4.$$

Then the Ingleton inequality fails since:

$$\underbrace{I(A;B)}_{(5-\log_2 27)/2} - \underbrace{I(A;B|C)}_{0} - \underbrace{I(A;B|D)}_{0} - \underbrace{I(C;D)}_{0} > 0.$$

D. Network Coding

In this section, we will briefly review some concepts of network coding. This will enable the discussion later in this paper of our construction of linear rank inequalities using networks constructed from two particular matroids (T8 and non-T8). For more details on network coding, see [20].

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:

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

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).

[3]We note that in Figures 2 and 3, 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.

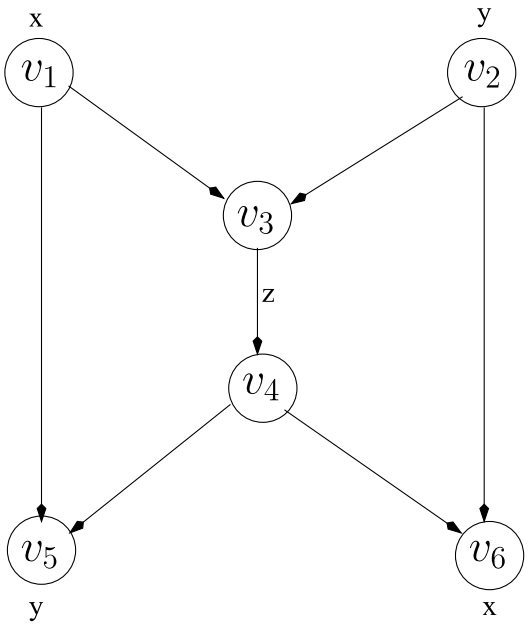

Fig. 1. The Butterfly network with source messages x and y, generated by source nodes v_1 and v_2, respectively. Demand nodes v_5 and v_6 demand messages y and x, respectively.

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

Let us illustrate a method for finding capacity bounds by examining the well-known Butterfly network, depicted in Figure 1. We assume the network messages x and y are independent, k-dimensional, random vectors with uniformly distributed components. Then in any solution it must be the case that

$$H(y|x, z) = 0 \tag{13}$$

since y is a function of x and z, and also that

$$\begin{aligned} 2k &= H(x) + H(y) && (14)\\ &= H(x, y) && \text{[from independence of x and y]}\\ &\leq H(x, y, z) && \text{[from (10)]}\\ &= H(x, z) + H(y|x, z) && \text{[from (1)]}\\ &= H(x, z) && \text{[from (13)]}\\ &\leq H(x) + H(z) && \text{[from (9)]}\\ &\leq k + n. && (15) \end{aligned}$$

This implies $2k \leq k + n$, or equivalently $k/n \leq 1$. Since this bound holds for all choices of k and n, the coding capacity must be at most 1. On the other hand, a solution with $k = n = 1$ is obtained by taking $z = x + y$ over any finite field alphabet, so the coding capacity is at least 1. Thus the coding capacity for the Butterfly network is the same as the linear coding capacity which is exactly equal to 1.

The inequalities in (15) were based on random variables x, y, z. Later, in the proofs of Corollaries 3.4 and 4.3, we will 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, but the procedure will appear similar.

II. Preliminaries

In this section, we given some technical lemmas which will be useful for proving the main results of the paper. Although some of them have appeared in the literature before, we present them here for completeness.

If A is a subspace of vector space V, and $\overline{A}$ is a subspace of A, then we will use the notation

$$\text{codim}_A(\overline{A}) = \dim(A) - \dim(\overline{A})$$

to represent the codimension of $\overline{A}$ in A. We will omit the subscript when it is obvious from the context which space the codimension is with respect to.

The proofs of all lemmas in this section are given in the Appendix.

Lemma 2.1: Let V be a finite dimensional vector space with subspaces A and B. Then the subspace $A \cap B$ has codimension at most $codim(A) + codim(B)$ in V.

Lemma 2.2: Let A and B be vector spaces over the same finite scalar field and with subspaces $\overline{A}$ and $\overline{B}$, respectively. Let $f : A \to B$ be a linear function such that $f(A \backslash \overline{A}) \subseteq B \backslash \overline{B}$. Then the codimension of $\overline{A}$ in A is at most the codimension of $\overline{B}$ in B.

Lemma 2.3: Let A and B be vector spaces over the same finite scalar field, let $\overline{B}$ be a subspace of B, and let $f : A \to B$ be a linear function. Then $f \in \overline{B}$ on a subspace of A of codimension at most the codimension of $\overline{B}$.

Lemma 2.4: Let V be a finite dimensional vector space and let $A_1, \ldots, A_k, B$ be subspaces of V. Then for $i = 1, \ldots, k$, there exist linear functions $f_i : B \to A_i$ such that $f_1 + \cdots + f_k = I$ on a subspace of B of codimension $H(B|A_1, \ldots, A_k)$.

Lemma 2.5: Let V be a finite-dimensional vector space and let A, B, and C be subspaces of V. Let $f : A \to B$ and $g : A \to C$ be linear functions such that $f + g = 0$ on A. Then $f = g = 0$ on a subspace of A of codimension at most $I(B; C)$.

Lemma 2.6: Let V be a finite dimensional vector space and let $A, B_1, \ldots, B_k$ be subspaces of V. For each $i = 1, \ldots, k$ let $f_i : A \to B_i$ be a linear function such that $f_1 + \cdots + f_k = 0$ on A. Then $f_1 = \cdots = f_k = 0$ on a subspace of A of codimension at most $H(B_1) + \cdots + H(B_k) - H(B_1, \ldots, B_k)$.

Lemma 2.7: Let A, B, C, D, E be subspaces of a vector space V and let f_R, f_L, g_R, and g_L be functions such that $f_R : A \to C$, $f_L : C \to A$, $g_R : B \to D$, and $g_L : D \to E$. If $f_L f_R = I$ on A and $g_L g_R$ is injective on B, then $g_L f_R$ is injective on $f_L(f_R A \cap g_R B)$.

III. 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

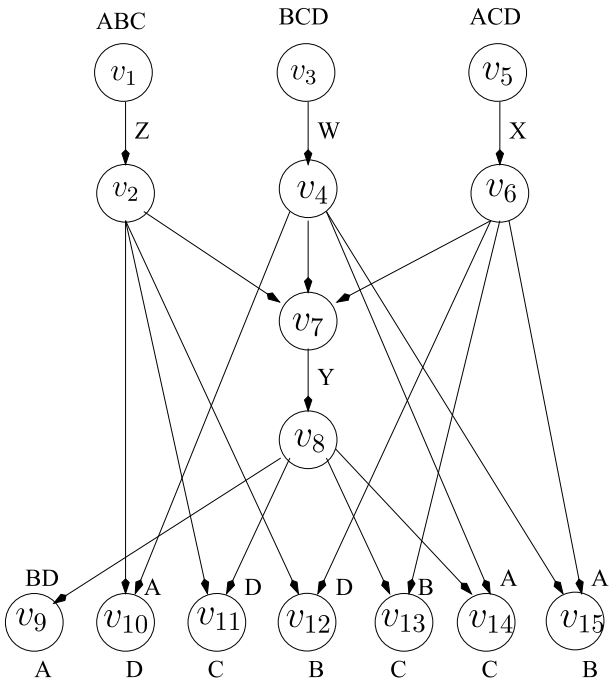

Fig. 2. The T8 network has source messages A, B, C, and D generated at hidden source nodes with certain hidden out-edges pointing to corresponding displayed nodes v_1, v_3, v_5, and v_9–v_{15} (which are labeled by incoming messages above such nodes). The nodes v_9–v_{15} each demand one message, as labeled below such nodes. The values W, X, Y, Z are labels assigned to certain edges to facilitate proofs.

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

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

The T8 matroid is representable over a field if and only if the field is of characteristic 3. Figure 2 is a network whose dependencies and independencies are consistent with the T8 matroid. It was designed by the construction process described in [6], and we will refer to it as the T8 network. Theorem 3.1 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 3.1: 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:*

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

Proof: See the Appendix. ∎

The next theorem demonstates that the inequality in Theorem 3.1 does not in general hold for vector spaces with finite fields of characteristic 3.

Theorem 3.2: There exists a vector space V with a finite scalar field of characteristic 3 *such that the T8 inequality in Theorem 3.1 is not a linear rank inequality over V.*

Proof: Let V be the vector space of 4-dimensional vectors whose components are from the field $GF(3)$, and define the following subspaces of V:

$$\begin{aligned} A &= \langle (1,0,0,0) \rangle \\ B &= \langle (0,1,0,0) \rangle \\ C &= \langle (0,0,1,0) \rangle \\ D &= \langle (0,0,0,1) \rangle \\ W &= \langle (0,1,1,1) \rangle \\ X &= \langle (1,0,1,1) \rangle \\ Y &= \langle (1,1,0,1) \rangle \\ Z &= \langle (1,1,1,0) \rangle. \end{aligned}$$

We have:

$$\begin{aligned} 0 &= H(Z|A, B, C) \\ &\quad [\text{from } (1,1,1,0) = (1,0,0,0) + (0,1,0,0) + (0,0,1,0)] \\ &= H(W|B, C, D) \\ &\quad [\text{from } (0,1,1,1) = (0,1,0,0) + (0,0,1,0) + (0,0,0,1)] \\ &= H(X|A, C, D) \\ &\quad [\text{from } (1,0,1,1) = (1,0,0,0) + (0,0,1,0) + (0,0,0,1)] \\ &= H(Y|W, X, Z) \\ &\quad [\text{from } (1,1,0,1) = 2^{-1}((0,1,1,1) + (1,0,1,1) \\ &\qquad + (1,1,1,0))] \\ &= H(A|B, D, Y) \\ &\quad [\text{from } (1,0,0,0) = (1,1,0,1) - (0,1,0,0) - (0,0,0,1)] \\ &= H(D|A, W, Z) \\ &\quad [\text{from } (0,0,0,1) = (0,1,1,1) + (1,0,0,0) - (1,1,1,0)] \\ &= H(C|D, Y, Z) \\ &\quad [\text{from } (0,0,1,0) = (1,1,1,0) + (0,0,0,1) - (1,1,0,1)] \\ &= H(B|D, X, Z) \\ &\quad [\text{from } (0,1,0,0) = (1,1,1,0) + (0,0,0,1) - (1,0,1,1)] \\ &= H(C|B, X, Y) \\ &\quad [\text{from } (0,0,1,0) = (1,0,1,1) + (0,1,0,0) - (1,1,0,1)] \\ &= H(C|A, W, Y) \\ &\quad [\text{from } (0,0,1,0) = (0,1,1,1) + (1,0,0,0) - (1,1,0,1)] \\ &= H(B|A, W, X) \\ &\quad [\text{from } (0,1,0,0) = (0,1,1,1) + (1,0,0,0) - (1,0,1,1)]. \end{aligned} \tag{16}$$

Note that the characteristic 3 assumption is used above in showing $H(Y|W, X, Z) = 0$, by using the fact that the ranks of Y and $Y \cap \langle W, X, Z\rangle$ are both 1, since

$$(1, 1, 0, 1) = 2^{-1} \cdot ((0, 1, 1, 1) + (1, 0, 1, 1) + (1, 1, 1, 0))$$

which holds for scalar fields of characteristic 3 (in fact, for all characteristics except 2).

We know $H(A) = H(B) = H(C) = H(D) = H(W) = H(X) = H(Y) = H(Z) = 1$. Also, we have

$$H(A) + H(B) + H(C) + H(D) = H(A, B, C, D).$$

So, if the inequality in Theorem 3.1 were to hold over V, then we would have

$$\begin{aligned} 1 &= H(A) \\ &\le 8H(Z) + 29H(Y) + 3H(X) + 8H(W) - 6H(D) \\ &\quad -17H(C) - 8H(B) - 17H(A) \\ &= 8 + 29 + 3 + 8 - 6 - 17 - 8 - 17 \\ &= 0 \end{aligned}$$

which is impossible. ∎

Consider a network over finite field F with a (k, n) linear code. The *vector space associated with any message* is defined to be F^k. The *vector space associated with any edge* is defined to be the set of all possible vectors from F^n that can be carried on that edge (i.e. taking into account the linear code).

Since each output of a network node is a function of the node's inputs, the conditional entropy of the vector carried by a node's out-edge, given the entropies of the vectors carried by the node's in-edges, is zero, assuming the network messages are uniform random vectors. The following lemma extends this idea from random variables to vector spaces and will be useful for the proof of Corollary 3.4.

Lemma 3.3: Suppose a network has a node with an out-edge (or demand) x and in-edges and messages (in some order) $y_1, \ldots, y_m$. Suppose the network has a finite field alphabet and a linear code. Let us view $X, Y_1, \ldots, Y_m$ as the vector spaces associated with $x, y_1, \ldots, y_m$, respectively. Then we have $H(X|Y_1, \ldots, Y_m) = 0$.

Proof: The vector carried on the node's out-edge (or demand) x is a linear combination of the vectors carried on the node's in-edges and the node's messages $y_1, \ldots, y_m$. Thus, every vector appearing on the node's out-edge (or demand) lies in the span of the subspaces $Y_1, \ldots, Y_m$. This implies

$$\dim(X) = \dim(X \cap \langle Y_1, \ldots, Y_m\rangle)$$

or equivalently,

$$H(X|Y_1, \ldots, Y_m) = 0.$$

∎

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 3.4 below makes use of the T8 linear rank inequality, and resembles the example shown earlier in (15) for computing the capacity of the Butterfly network using information inequalities and random variables.

Corollary 3.4: 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.

Proof: Let F be a finite field alphabet. Consider a (k, n) linear solution of the T8 network over F, such that the characteristic of F is not 3. Let A, B, C, D be message random variables in the T8 network, that are uniformly distributed over vectors in F^k. Let W, X, Y, Z be the resulting random variables associated with the corresponding labeled edges of T8 in Figure 2.

Equations (16) now hold with random variables A, B, C, D, W, X, Y, Z (i.e., not as subspaces as in Theorem 3.2) by Lemma 3.3:

$$\begin{aligned} 0 &= H(Z|A, B, C) && [\text{from } (v_1, v_2)] \\ &= H(W|B, C, D) && [\text{from } (v_3, v_4)] \\ &= H(X|A, C, D) && [\text{from } (v_5, v_6)] \\ &= H(Y|W, X, Z) && [\text{from } (v_4, v_7)] \\ &= H(A|B, D, Y) && [\text{from } v_9] \\ &= H(D|A, W, Z) && [\text{from } v_{10}] \\ &= H(C|D, Y, Z) && [\text{from } v_{11}] \\ &= H(B|D, X, Z) && [\text{from } v_{12}] \\ &= H(C|B, X, Y) && [\text{from } v_{13}] \\ &= H(C|A, W, Y) && [\text{from } v_{14}] \\ &= H(B|A, W, X) && [\text{from } v_{15}] \end{aligned}$$

and since the vector spaces A, B, C, D are associated with independent random variables, we have

$$H(A) + H(B) + H(C) + H(D) = H(A, B, C, D)$$

so the T8 inequality in Theorem 3.1 reduces to

$$\begin{aligned} H(A) \le\ & 8H(Z) + 29H(Y) + 3H(X) + 8H(W) - 6H(D) \\ & - 17H(C) - 8H(B) - 17H(A). \end{aligned}$$

Now since $H(A) = H(B) = H(C) = H(D) = k$ and $H(W) = H(X) = H(Y) = H(Z) \le n$, we have

$$\begin{aligned} k &\le 8n + 29n + 3n + 8n - 6k - 17k - 8k - 17k \\ k/n &\le 48/49. \end{aligned}$$

So, the linear coding capacity over every characteristic except for 3 is at most $48/49 < 1$.

The T8 network has a scalar linear solution over characteristic 3 by using the following edge functions (here we are using the notations A, B, C, D, W, X, Y, Z to denote edge variables rather than vector spaces):

$$\begin{aligned} Z &= A + B + C \\ W &= B + C + D \\ X &= A + C + D \\ Y &= W + X + Z. \end{aligned}$$

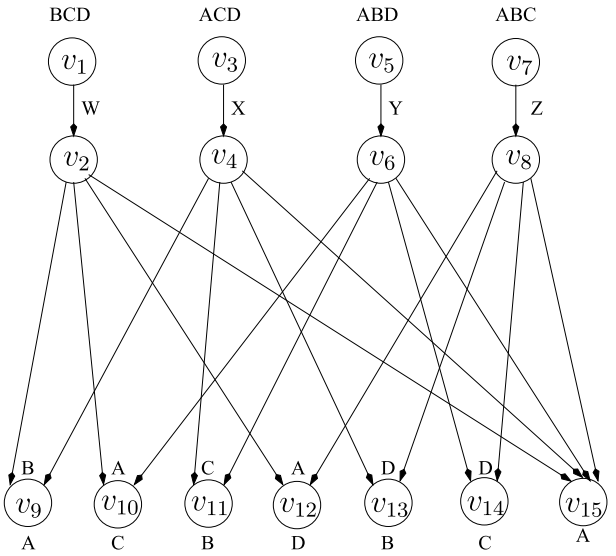

Fig. 3. The Non-T8 Network has source messages A, B, C, and D generated at hidden source nodes with certain hidden out-edges pointing to corresponding displayed nodes v_1, v_3, v_5, v_7, and v_9–v_{14} (which are labeled by incoming messages above such nodes). The nodes v_9–v_{15} each demand one message, as labeled below such nodes.

and decoding functions:

$$\begin{aligned} v_9 &: A = (2^{-1} \cdot Y) - B - D \\ v_{10} &: D = W - Z + A \\ v_{11} &: C = Z - (2^{-1} \cdot Y) + D \\ v_{12} &: B = Z - X + D \\ v_{13} &: C = X - (2^{-1} \cdot Y) + B \\ v_{14} &: C = W - (2^{-1} \cdot Y) + A \\ v_{15} &: B = W - X + A. \end{aligned}$$

Thus the linear coding capacity for characteristic 3 is at least 1.

We know the coding capacity is at most 1 because every path from source A to node v_9 passes through the single edge (v_7, v_8). Since the coding capacity is at least as large as the linear coding capacity for characteristic 3, we conclude that the coding capacity is exactly equal to 1. ■

IV. 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 3 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 4.1 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 4.1: 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:

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

Proof: See the Appendix. ■

The next theorem demonstates that the inequality in Theorem 4.1 does not in general hold for vector spaces with finite fields of characteristic other than 3.

Theorem 4.2: For each prime number $p \neq 3$ there exists a vector space V with a finite scalar field of characteristic p such that the non-T8 inequality in Theorem 4.1 is not a linear rank inequality over V.

Proof: Let V be the vector space of 4-dimensional vectors whose components are from $GF(p)$, and define the following subspaces of V:

$$\begin{aligned} A &= \langle (1, 0, 0, 0) \rangle \\ B &= \langle (0, 1, 0, 0) \rangle \\ C &= \langle (0, 0, 1, 0) \rangle \\ D &= \langle (0, 0, 0, 1) \rangle \\ W &= \langle (0, 1, 1, 1) \rangle \\ X &= \langle (1, 0, 1, 1) \rangle \\ Y &= \langle (1, 1, 0, 1) \rangle \\ Z &= \langle (1, 1, 1, 0) \rangle. \end{aligned}$$

We have:

$$\begin{aligned} 0 &= H(W|B, C, D) \\ &\quad [\text{from } (0, 1, 1, 1) = (0, 1, 0, 0) + (0, 0, 1, 0) + (0, 0, 0, 1)] \\ &= H(X|A, C, D) \\ &\quad [\text{from } (1, 0, 1, 1) = (1, 0, 0, 0) + (0, 0, 1, 0) + (0, 0, 0, 1)] \\ &= H(Y|A, B, D) \\ &\quad [\text{from } (1, 1, 0, 1) = (1, 0, 0, 0) + (0, 1, 0, 0) + (0, 0, 0, 1)] \\ &= H(Z|A, B, C) \\ &\quad [\text{from } (1, 1, 1, 0) = (1, 0, 0, 0) + (0, 1, 0, 0) + (0, 0, 1, 0)] \\ &= H(A|B, W, X) \\ &\quad [\text{from } (1, 0, 0, 0) = (1, 0, 1, 1) + (0, 1, 0, 0) - (0, 1, 1, 1)] \\ &= H(C|A, W, Y) \\ &\quad [\text{from } (0, 0, 1, 0) = (0, 1, 1, 1) + (1, 0, 0, 0) - (1, 1, 0, 1)] \\ &= H(B|C, X, Y) \\ &\quad [\text{from } (0, 1, 0, 0) = (1, 1, 0, 1) + (0, 0, 1, 0) - (1, 0, 1, 1)] \\ &= H(D|A, W, Z) \\ &\quad [\text{from } (0, 0, 0, 1) = (0, 1, 1, 1) + (1, 0, 0, 0) - (1, 1, 1, 0)] \\ &= H(B|D, X, Z) \\ &\quad [\text{from } (0, 1, 0, 0) = (1, 1, 1, 0) + (0, 0, 0, 1) - (1, 0, 1, 1)] \\ &= H(C|D, Y, Z) \\ &\quad [\text{from } (0, 0, 1, 0) = (1, 1, 1, 0) + (0, 0, 0, 1) - (1, 1, 0, 1)] \\ &= H(A|W, X, Y, Z) \\ &\quad [\text{from } (1, 0, 0, 0) = 3^{-1}((1, 0, 1, 1) + (1, 1, 0, 1) \\ &\qquad + (1, 1, 1, 0) - 2(0, 1, 1, 1))]. \end{aligned} \quad (17)$$

We know $H(A) = H(B) = H(C) = H(D) = H(W) = H(X) = H(Y) = H(Z) = 1$, Also, we have

$$H(A) + H(B) + H(C) + H(D) = H(A, B, C, D).$$

So, if the inequality in Theorem 4.1 were to hold over V, then we would have

$$\begin{aligned} 1 &= H(A) \\ &\le 9H(Z) + 8H(Y) + 5H(X) + 6H(W) - 4H(D) \\ &\quad - 12H(C) - 11H(B) - H(A) \\ &= 9 + 8 + 5 + 6 - 4 - 12 - 11 - 1 \\ &= 0 \end{aligned}$$

which is impossible. ∎

Corollary 4.3: 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.

Proof: Let F be a finite field alphabet. Consider a (k, n) linear solution of the non-T8 network over F, such that the characteristic of F is 3. Let A, B, C, D be message random variables in the T8 network, that are uniformly distributed over vectors in F^k. Let W, X, Y, Z be the resulting random variables associated with the corresponding labeled edges of T8 in Figure 3.

Equations (17) now hold when A, B, C, D, W, X, Y, Z are taken as random variables (i.e. not as subspaces as in Theorem 4.2) by Lemma 3.3:

$$\begin{aligned} 0 &= H(W|B, C, D) && [\text{from } (v_1, v_2)] \\ &= H(X|A, C, D) && [\text{from } (v_3, v_4)] \\ &= H(Y|A, B, D) && [\text{from } (v_5, v_6)] \\ &= H(Z|A, B, C) && [\text{from } (v_7, v_8)] \\ &= H(A|B, W, X) && [\text{from } v_9] \\ &= H(C|A, W, Y) && [\text{from } v_{10}] \\ &= H(B|C, X, Y) && [\text{from } v_{11}] \\ &= H(D|A, W, Z) && [\text{from } v_{12}] \\ &= H(B|D, X, Z) && [\text{from } v_{13}] \\ &= H(C|D, Y, Z) && [\text{from } v_{14}] \\ &= H(A|W, X, Y, Z) && [\text{from } v_{15}] \end{aligned}$$

and since the source messages A, B, C, D are independent random variables, we have

$$H(A) + H(B) + H(C) + H(D) = H(A, B, C, D)$$

so the non-T8 inequality in Theorem 4.1 reduces to

$$\begin{aligned} H(A) \le\ & 9H(Z) + 8H(Y) + 5H(X) + 6H(W) - 4H(D) \\ & - 12H(C) - 11H(B) - H(A). \end{aligned}$$

Now, since

$$H(A) = H(B) = H(C) = H(D) = k$$

and

$$H(W) = H(X) = H(Y) = H(Z) \le n$$

we have

$$\begin{aligned} k &\le 9n + 8n + 5n + 6n - 4k - 12k - 11k - k \\ k/n &\le 28/29. \end{aligned}$$

So, the linear coding capacity over characteristic 3 is at most $28/29 < 1$.

The non-T8 network has a scalar linear solution over every characteristic except for 3 by using the following edge functions (here we are using the notations A, B, C, D, W, X, Y, Z to denote edge variables rather than vector spaces):

$$\begin{aligned} W &= B + C + D \\ X &= A + C + D \\ Y &= A + B + D \\ Z &= A + B + C \end{aligned}$$

and decoding functions:

$$\begin{aligned} v_9 &: A = X - W + B \\ v_{10} &: C = W - Y + A \\ v_{11} &: B = Y - X + C \\ v_{12} &: D = W - Z + A \\ v_{13} &: B = Z - X + D \\ v_{14} &: C = Z - Y + D \\ v_{15} &: A = 3^{-1} \cdot (X + Y + Z - 2W). \end{aligned}$$

We know the coding capacity is at most 1 because there is a unique path from source A to node v_9 (through node v_4). Since the coding capacity is at least as large as the linear coding capacity for characteristics other than 3, we conclude that the coding capacity is exactly equal to 1. ∎

V. Conclusion

We have demonstrated a linear rank inequality which holds over all vector spaces when the scalar field characteristic is anything but three, and have shown that this inequality does not generally hold over characteristic three. Similarly, we have demonstrated a linear rank inequality which holds over all vector spaces with scalar field of characteristic three, and have shown that this inequality does not generally hold over characteristics other than three. We have applied these inequalities to the problem of bounding the network coding capacity of certain directed acyclic networks. An open problem is how to use these ideas to bound the capacities of more general networks using linear rank inequalities.

Appendix

Proof of Lemma 2.1: We know $H(A) + H(B) - I(A; B) = H(A, B) \le H(V)$. Then adding $H(V)$ to both sides of the inequality gives

$$H(V) - I(A; B) \le H(V) - H(A) + H(V) - H(B).$$

Thus, $\operatorname{codim}(A \cap B) \le \operatorname{codim}(A) + \operatorname{codim}(B)$. ∎

Proof of Lemma 2.2: Suppose a base for A consists of a base for $\overline{A}$ together with the vectors $a_1, \ldots, a_n$. Let $\gamma_1, \ldots, \gamma_n$ be field elements which are not all zero. Then

$$\gamma_1 a_1 + \cdots + \gamma_n a_n \in A \setminus \overline{A}$$

so

$$\gamma_1 f(a_1) + \cdots + \gamma_n f(a_n) = f(\gamma_1 a_1 + \cdots + \gamma_n a_n) \in B \setminus \overline{B}.$$

Thus, the vectors $f(a_1), \ldots, f(a_n)$ are linearly independent over the subspace $\overline{B}$, and therefore

$$\operatorname{codim}_A(\overline{A}) = n \le \operatorname{codim}_B(\overline{B}).$$ ∎

Proof of Lemma 2.3: Let $\overline{A} = \{t \in A : f(t) \in \overline{B}\}$. Then $f(A\backslash\overline{A}) \subseteq B\backslash\overline{B}$ and the result follows from Lemma 2.2. ∎

Proof of Lemma 2.4: Let W be a subspace of B defined by

$$W = \langle A_1, \ldots, A_k \rangle \cap B.$$

The subspace on which this lemma holds is W. If $H(W) = 0$, then the lemma would be trivially true. So, assume that $H(W) > 0$, and let $\{w_1, \ldots, w_n\}$ be a basis for W. For each $j = 1, \ldots, n$, choose $x_{i,j} \in A_i$ for $i = 1, \ldots, k$ such that

$$w_j = x_{1,j} + \cdots + x_{k,j}.$$

For each $i = 1, \ldots, k$, define a linear mapping $g_i : W \to A_i$ so that $g_i(w_j) = x_{i,j}$ for all i and j. Then extend g_i arbitrarily to $f_i : B \to A_i$. Now we have linear functions $f_1, \ldots, f_k$ such that

$$f_1 + \cdots + f_k = I$$

on W. The dimension of W is

$$H(W) = I(A_1, \ldots, A_k; B)$$

so the codimension of W is

$$H(B) - I(A_1, \ldots, A_k; B) = H(B|A_1, \ldots, A_k). \quad ∎$$

Proof of Lemma 2.5: Let K be the kernel of f. Clearly, f maps A into $B \cap C$ and since f is linear the rank of its domain is at most the sum of the ranks of its kernel and range, so

$$\operatorname{codim}(K) = H(A) - H(K) \leq I(B; C). \quad ∎$$

Proof of Lemma 2.6: First we apply Lemma 2.5 to f_1 and $(f_2 + \cdots + f_k)$ to get

$$f_1 = (f_2 + \cdots + f_k) = 0$$

on a subspace A_1 of A of codimension at most

$$I(B_1; B_2, \ldots, B_k) = H(B_1) + H(B_2, \ldots, B_k) - H(B_1, B_2, \ldots, B_k).$$

Then apply Lemma 2.5 to f_2 and $(f_3 + \cdots + f_k)$ to get

$$f_2 = (f_3 + \cdots + f_k) = 0$$

on a subspace A_2 of A_1 of codimension at most

$$I(B_2; B_3, \ldots, B_k) = H(B_2) + H(B_3, \ldots, B_k) - H(B_2, B_3, \ldots, B_k).$$

Continue on until we apply Lemma 2.5 to f_{k-1} and f_k to get

$$f_{k-1} = f_k = 0$$

on a subspace A_{k-1} of A_{k-2} of codimension at most

$$I(B_{k-1}; B_k) = H(B_{k-1}) + H(B_k) - H(B_{k-1}, B_k).$$

Now A_{k-1} is a subspace of A of codimension at most

$$H(B_1) + \cdots + H(B_k) - H(B_1, \ldots, B_k),$$

on which $f_1 = f_2 = \cdots = f_k = 0$. ∎

Proof of Lemma 2.7: Let $x, y \in f_L(f_R A \cap g_R B)$. We know $f_R f_L = I$ on $f_R A$ because

$$f_R f_L(f_R(w)) = f_R(f_L f_R(w)) = f_R(w)$$

for all $w \in A$. Since $x \in f_L(f_R A \cap g_R B)$, we know

$$f_R(x) \in f_R f_L(f_R A \cap g_R B) = f_R A \cap g_R B$$

which implies $f_R(x) = g_R(b_x)$ for some $b_x \in B$. Similarly, we know $f_R(y) = g_R(b_y)$ for some $b_y \in B$. So, we have

$$g_L g_R(b_x) = g_L f_R(x)$$

and

$$g_L g_R(b_y) = g_L f_R(y).$$

If we assume $g_L f_R(x) = g_L f_R(y)$, then we have

$$g_L g_R(b_x) = g_L g_R(b_y).$$

Since $g_L g_R$ is injective on B, we know $b_x = b_y$. Thus

$$f_R(x) = g_R(b_x) = g_R(b_y) = f_R(y)$$

which implies

$$f_L f_R(x) = f_L f_R(y).$$

Since $f_L f_R = I$ on A, we know $x = y$. Thus $g_L f_R$ is injective on $f_L(f_R A \cap g_R B)$. ∎

Proof of Theorem 3.1: The main idea is to establish the existence of certain linear functions, some of which are injective on particular subspaces of the original vector space V. Inequalities relating the dimensions (or co-dimensions) of various subspaces ultimately use the assumption that the field is of characteristic other than 3, and then the final linear rank inequality is obtained. Many of the subspace co-dimension computations and manipulations are fairly tedious, although they can be readily followed and verified in a line-by-line manner.

By Lemma 2.4 we get linear functions:

$$\begin{array}{lll}
f_1 : Z \to A, & f_2 : Z \to B, & f_3 : Z \to C, \\
f_4 : W \to B, & f_5 : W \to C, & f_6 : W \to D, \\
f_7 : X \to A, & f_8 : X \to C, & f_9 : X \to D, \\
f_{10} : Y \to Z, & f_{11} : Y \to W, & f_{12} : Y \to X, \\
f_{13} : A \to B, & f_{14} : A \to D, & f_{15} : A \to Y, \\
f_{16} : D \to Z, & f_{17} : D \to W, & f_{18} : D \to A, \\
f_{19} : C \to Z, & f_{20} : C \to Y, & f_{21} : C \to D, \\
f_{22} : B \to Z, & f_{23} : B \to X, & f_{24} : B \to D, \\
f_{25} : C \to Y, & f_{26} : C \to X, & f_{27} : C \to B, \\
f_{28} : C \to Y, & f_{29} : C \to W, & f_{30} : C \to A, \\
f_{31} : B \to W, & f_{32} : B \to X, & f_{33} : B \to A
\end{array}$$

such that

$$f_1 + f_2 + f_3 = I$$
on a subspace of Z of codimension $H(Z|A, B, C)$ (A.1)

$$f_4 + f_5 + f_6 = I$$
on a subspace of W of codimension $H(W|B, C, D)$ (A.2)

$$f_7 + f_8 + f_9 = I$$
on a subspace of X of codimension $H(X|A, C, D)$ (A.3)

$$f_{10} + f_{11} + f_{12} = I$$
on a subspace of Y of codimension $H(Y|W, X, Z)$ (A.4)

$$f_{13} + f_{14} + f_{15} = I$$
on a subspace of A of codimension $H(A|B, D, Y)$ (A.5)

$$f_{16} + f_{17} + f_{18} = I$$
on a subspace of D of codimension $H(D|A, W, Z)$ (A.6)
$$f_{19} + f_{20} + f_{21} = I$$
on a subspace of C of codimension $H(C|D, Y, Z)$ (A.7)
$$f_{22} + f_{23} + f_{24} = I$$
on a subspace of B of codimension $H(B|D, X, Z)$ (A.8)
$$f_{25} + f_{26} + f_{27} = I$$
on a subspace of C of codimension $H(C|B, X, Y)$ (A.9)
$$f_{28} + f_{29} + f_{30} = I$$
on a subspace of C of codimension $H(C|A, W, Y)$ (A.10)
$$f_{31} + f_{32} + f_{33} = I$$
on a subspace of B of codimension $H(B|A, W, X)$. (A.11)

Now let

$$f_A \triangleq f_7 f_{12} + f_1 f_{10}$$
$$f_B \triangleq f_4 f_{11} + f_2 f_{10}$$
$$f_C \triangleq f_8 f_{12} + f_5 f_{11} + f_3 f_{10}$$
$$f_D \triangleq f_9 f_{12} + f_6 f_{11}.$$

Combining the functions we obtained from Lemma 2.4 gives new functions:

$$f_A f_{15} : A \to A$$
$$f_B f_{15} + f_{13} : A \to B$$
$$f_C f_{15} : A \to C$$
$$f_D f_{15} + f_{14} : A \to D.$$

Using (A.1) - (A.5), Lemma 2.1, and Lemma 2.3 we know the sum of these functions is equal to I on a subspace of A of codimension at most

$$H(Z|A, B, C) + H(W|B, C, D) + H(X|A, C, D) + H(Y|W, X, Z) + H(A|B, D, Y)$$

since, on that subspace,

$$\begin{aligned}
&(f_A f_{15}) + (f_B f_{15} + f_{13}) + (f_C f_{15}) + (f_D f_{15} + f_{14}) \\
&= f_7 f_{12} f_{15} + f_1 f_{10} f_{15} + f_4 f_{11} f_{15} + f_2 f_{10} f_{15} + f_{13} \\
&\quad + f_8 f_{12} f_{15} + f_5 f_{11} f_{15} + f_3 f_{10} f_{15} + f_9 f_{12} f_{15} \\
&\quad + f_6 f_{11} f_{15} + f_{14} \\
&= f_{13} + f_{14} + (f_1 + f_2 + f_3) f_{10} f_{15} + (f_4 + f_5 + f_6) f_{11} f_{15} \\
&\quad + (f_7 + f_8 + f_9) f_{12} f_{15} \\
&= f_{13} + f_{14} + (f_{10} + f_{11} + f_{12}) f_{15} \\
&= f_{13} + f_{14} + f_{15} \\
&= I.
\end{aligned}$$

Applying Lemma 2.6 and Lemma 2.1 to

$$f_A f_{15} - I$$
$$f_B f_{15} + f_{13}$$
$$f_C f_{15}$$
$$f_D f_{15} + f_{14}$$

we get a subspace $\overline{A}$ of A of codimension at most

$$\begin{aligned}
\Delta_{\overline{A}} = {} & H(Z|A, B, C) + H(W|B, C, D) + H(X|A, C, D) \\
& + H(Y|W, X, Z) + H(A|B, D, Y) \\
& + H(A) + H(B) + H(C) + H(D) - H(A, B, C, D)
\end{aligned}$$

on which

$$f_A f_{15} = I \tag{A.12}$$
$$f_B f_{15} + f_{13} = 0 \tag{A.13}$$
$$f_C f_{15} = 0 \tag{A.14}$$
$$f_D f_{15} + f_{14} = 0. \tag{A.15}$$

To see how the T8 network is used as a guide, consider receiver node v_9, which demands A. Let M_1, M_7, M_{10}, M_{12}, M_{15} be matrices corresponding to the transformations along the edges (A, Z), (A, X), (Z, Y), (X, Y), (Y, A), respectively. Using algebra to solve for A one deduces that

$$M_{15} M_{10} M_1 + M_{15} M_{12} M_7 = I.$$

Equation (A.12) was designed to model this property.

Similarly, we get a subspace $\overline{B}$ of B of codimension at most

$$\begin{aligned}
\Delta_{\overline{B}} = {} & H(Z|A, B, C) + H(X|A, C, D) + H(B|D, X, Z) \\
& + H(A) + H(B) + H(C) + H(D) - H(A, B, C, D)
\end{aligned}$$

on which

$$f_7 f_{23} + f_1 f_{22} = 0 \tag{A.16}$$
$$f_2 f_{22} = I \tag{A.17}$$
$$f_8 f_{23} + f_3 f_{22} = 0 \tag{A.18}$$
$$f_{24} + f_9 f_{23} = 0. \tag{A.19}$$

We get a subspace $\widehat{B}$ of B of codimension at most

$$\begin{aligned}
\Delta_{\widehat{B}} = {} & H(W|B, C, D) + H(X|A, C, D) + H(B|A, W, X) \\
& + H(A) + H(B) + H(C) + H(D) - H(A, B, C, D)
\end{aligned}$$

on which

$$f_{33} + f_7 f_{32} = 0 \tag{A.20}$$
$$f_4 f_{31} = I \tag{A.21}$$
$$f_8 f_{32} + f_5 f_{31} = 0 \tag{A.22}$$
$$f_9 f_{32} + f_6 f_{31} = 0. \tag{A.23}$$

We get a subspace $\overline{C}$ of C of codimension at most

$$\begin{aligned}
\Delta_{\overline{C}} = {} & 2H(Z|A, B, C) + H(W|B, C, D) + H(X|A, C, D) \\
& + H(Y|W, X, Z) + H(C|D, Y, Z) \\
& + H(A) + H(B) + H(C) + H(D) - H(A, B, C, D)
\end{aligned}$$

on which

$$f_A f_{20} + f_1 f_{19} = 0 \tag{A.24}$$
$$f_B f_{20} + f_2 f_{19} = 0 \tag{A.25}$$
$$f_C f_{20} + f_3 f_{19} = I \tag{A.26}$$
$$f_D f_{20} + f_{21} = 0. \tag{A.27}$$

We get a subspace $\widehat{C}$ of C of codimension at most

$$\begin{aligned}\Delta_{\widehat{C}} = {} & H(Z|A,B,C)+H(W|B,C,D)+2H(X|A,C,D)\\ & +H(Y|W,X,Z)+H(C|B,X,Y)\\ & +H(A)+H(B)+H(C)+H(D)-H(A,B,C,D)\end{aligned}$$

on which

$$f_A f_{25}+f_7 f_{26}=0 \tag{A.28}$$
$$f_B f_{25}+f_{27}=0 \tag{A.29}$$
$$f_C f_{25}+f_8 f_{26}=I \tag{A.30}$$
$$f_D f_{25}+f_9 f_{26}=0. \tag{A.31}$$

We get a subspace $\widetilde{C}$ of C of codimension at most

$$\begin{aligned}\Delta_{\widetilde{C}} = {} & H(Z|A,B,C)+2H(W|B,C,D)+H(X|A,C,D)\\ & +H(Y|W,X,Z)+H(C|A,W,Y)\\ & +H(A)+H(B)+H(C)+H(D)-H(A,B,C,D)\end{aligned}$$

on which

$$f_A f_{28}+f_{30}=0 \tag{A.32}$$
$$f_B f_{28}+f_4 f_{29}=0 \tag{A.33}$$
$$f_C f_{28}+f_5 f_{29}=I \tag{A.34}$$
$$f_D f_{28}+f_6 f_{29}=0. \tag{A.35}$$

We get a subspace $\overline{D}$ of D of codimension at most

$$\begin{aligned}\Delta_{\overline{D}} = {} & H(Z|A,B,C)+H(W|B,C,D)+H(D|A,W,Z)\\ & +H(A)+H(B)+H(C)+H(D)-H(A,B,C,D)\end{aligned}$$

on which

$$f_{18}+f_1 f_{16}=0 \tag{A.36}$$
$$f_4 f_{17}+f_2 f_{16}=0 \tag{A.37}$$
$$f_5 f_{17}+f_3 f_{16}=0 \tag{A.38}$$
$$f_6 f_{17}=I. \tag{A.39}$$

First notice that (A.12) implies

$$f_{15} \text{ is injective on } \overline{A}. \tag{A.40}$$

We need to define a subspace of $\overline{A}$ on which f_{13} and f_{14} are injective. The justifications can be found on (A.44) and (A.45). Let

$$\begin{aligned}\overline{C}^* &\triangleq f_3(f_{19}(\overline{C}\cap f_{20}^{-1}f_{15}\overline{A})\cap f_{22}\overline{B})\subseteq\overline{C}\\ \widetilde{C}^* &\triangleq f_5(f_{29}(\widetilde{C}\cap f_{28}^{-1}f_{15}\overline{A})\cap f_{17}\overline{D})\subseteq\widetilde{C}\\ \overline{A}^* &\triangleq f_A(f_{15}\overline{A}\cap f_{20}\overline{C}^*\cap f_{28}\widetilde{C}^*)\subseteq\overline{A}.\end{aligned}$$

To justify why $\overline{C}^*\subseteq\overline{C}$, by (A.14) we know $f_C f_{15}=0$ on $\overline{A}$ and by (A.26) we know

$$f_C f_{20}+f_3 f_{19}=I.$$

Thus for each $\overline{c}\in\overline{C}\cap f_{20}^{-1}f_{15}\overline{A}$, we have $f_C f_{20}=0$ on $\overline{C}$ which gives

$$f_3 f_{19}=I \text{ on } \overline{C}\cap f_{20}^{-1}f_{15}\overline{A}. \tag{A.41}$$

Using (A.14) and (A.34) we have

$$f_5 f_{29}=I \text{ on } \widetilde{C}\cap f_{28}^{-1}f_{15}\overline{A}. \tag{A.42}$$

Using (A.14) and (A.30) we have

$$f_8 f_{26}=I \text{ on } \widehat{C}\cap f_{25}^{-1}f_{15}\overline{A}. \tag{A.43}$$

We are now going to show f_{13} is injective on $\overline{A}^*$. First we need to apply Lemma 2.7 to show $f_2 f_{19}$ is injective on $\overline{C}^*$ and then again to show $f_B f_{15}$ is injective on $\overline{A}^*$. By (A.17) and (A.41), we know $f_2 f_{22}$ is injective on $\overline{B}$ and $f_3 f_{19}=I$ on $\overline{C}\cap f_{20}^{-1}f_{15}\overline{A}$. So, we can apply Lemma 2.7 by letting $g_L=f_2$, $g_R=f_{22}$, $f_L=f_3$, and $f_R=f_{19}$ to get that $f_2 f_{19}$ is injective on $\overline{C}^*$. Then using (A.25), we know $f_B f_{20}$ is injective on $\overline{C}^*$. Now we can apply Lemma 2.7 again by using the fact that $f_A f_{15}=I$ on $\overline{A}$ and by letting $g_L=f_B$, $g_R=f_{20}$, $f_L=f_A$, and $f_R=f_{15}$ to get $f_B f_{15}$ is injective on $\overline{A}^*$. Thus by (A.13),

$$f_{13} \text{ is injective on } \overline{A}^*. \tag{A.44}$$

Similarly, we are going to show f_{14} is injective on $\overline{A}^*$. We will first apply Lemma 2.7 to show $f_6 f_{29}$ is injective on $\widetilde{C}^*$ and then again to show $f_D f_{15}$ is injective on $\overline{A}^*$. By (A.39) and (A.42), we know $f_6 f_{17}$ is injective on $\overline{D}$ and $f_5 f_{29}=I$ on $\widetilde{C}\cap f_{28}^{-1}f_{15}\overline{A}$. So, we can apply Lemma 2.7 by letting $g_L=f_6$, $g_R=f_{17}$, $f_L=f_5$, and $f_R=f_{29}$ to get that $f_6 f_{29}$ is injective on $\widetilde{C}^*$. Then using (A.35), we know $f_D f_{28}$ is injective on $\widetilde{C}^*$. Now we can apply Lemma 2.7 again by using the fact that $f_A f_{15}=I$ on $\overline{A}$ and by letting $g_L=f_D$, $g_R=f_{28}$, $f_L=f_A$, and $f_R=f_{15}$ to get $f_D f_{15}$ is injective on $\overline{A}^*$. Thus by (A.15),

$$f_{14} \text{ is injective on } \overline{A}^*. \tag{A.45}$$

Now we are going to find an upper bound for $\mathrm{codim}_A(\overline{A}^*)$. First we need to find upper bounds for $\mathrm{codim}_C(\overline{C}^*)$ and $\mathrm{codim}_C(\widetilde{C}^*)$. Using (A.40) to show $\dim(f_{15}\overline{A})=\dim(\overline{A})$, and again using Lemma 2.1 and Lemma 2.3, we have

$$\begin{aligned}&\mathrm{codim}_C(\overline{C}^*)\\ &=H(C)-\dim(\overline{C}^*)\\ &=H(C)-\dim(f_3(f_{19}(\overline{C}\cap f_{20}^{-1}f_{15}\overline{A})\cap f_{22}\overline{B}))\\ &=H(C)-\dim(f_{19}(\overline{C}\cap f_{20}^{-1}f_{15}\overline{A})\cap f_{22}\overline{B})\\ &=H(C)-H(Z)+\mathrm{codim}_Z(f_{19}(\overline{C}\cap f_{20}^{-1}f_{15}\overline{A})\cap f_{22}\overline{B})\\ &\le H(C)-H(Z)+\mathrm{codim}_Z(f_{19}(\overline{C}\cap f_{20}^{-1}f_{15}\overline{A}))\\ &\quad+\mathrm{codim}_Z(f_{22}\overline{B})\\ &=H(C)-H(Z)+H(Z)-\dim(f_{19}(\overline{C}\cap f_{20}^{-1}f_{15}\overline{A}))\\ &\quad+H(Z)-\dim(f_{22}\overline{B})\\ &=H(C)+H(Z)-\dim(\overline{C}\cap f_{20}^{-1}f_{15}\overline{A})-\dim(\overline{B})\\ &=H(C)+H(Z)-H(C)+\mathrm{codim}_C(\overline{C}\cap f_{20}^{-1}f_{15}\overline{A})\\ &\quad-H(B)+\mathrm{codim}_B(\overline{B})\\ &=H(Z)-H(B)+\mathrm{codim}_C(\overline{C}\cap f_{20}^{-1}f_{15}\overline{A})+\mathrm{codim}_B(\overline{B})\\ &\le H(Z)-H(B)+\Delta_{\overline{C}}+\mathrm{codim}_C(f_{20}^{-1}f_{15}\overline{A})+\Delta_{\overline{B}}\\ &\le H(Z)-H(B)+\Delta_{\overline{C}}+\mathrm{codim}_Y(f_{15}\overline{A})+\Delta_{\overline{B}}\\ &\le H(Z)-H(B)+\Delta_{\overline{C}}+H(Y)-\dim(f_{15}\overline{A})+\Delta_{\overline{B}}\\ &=H(Z)-H(B)+\Delta_{\overline{C}}+H(Y)-\dim(\overline{A})+\Delta_{\overline{B}}\\ &=H(Z)-H(B)+\Delta_{\overline{C}}+H(Y)-H(A)\\ &\quad+\mathrm{codim}_A(\overline{A})+\Delta_{\overline{B}}\\ &\le H(Z)-H(B)+H(Y)-H(A)+\Delta_{\overline{C}}+\Delta_{\overline{A}}+\Delta_{\overline{B}}\end{aligned} \tag{A.46}$$

$$
\begin{aligned}
&\operatorname{codim}_C(\widetilde{C}^*)\\
&= H(C) - \dim(\widetilde{C}^*)\\
&= H(C) - \dim(f_5(f_{29}(\widetilde{C} \cap f_{28}^{-1} f_{15}\overline{A}) \cap f_{17}\overline{D}))\\
&= H(C) - \dim(f_{29}(\widetilde{C} \cap f_{28}^{-1} f_{15}\overline{A}) \cap f_{17}\overline{D})\\
&= H(C) - H(W) + \operatorname{codim}_W(f_{29}(\widetilde{C} \cap f_{28}^{-1} f_{15}\overline{A}) \cap f_{17}\overline{D})\\
&\le H(C) - H(W) + \operatorname{codim}_W(f_{29}(\widetilde{C} \cap f_{28}^{-1} f_{15}\overline{A}))\\
&\quad + \operatorname{codim}_W(f_{17}\overline{D})\\
&= H(C) - H(W) + H(W) - \dim(f_{29}(\widetilde{C} \cap f_{28}^{-1} f_{15}\overline{A}))\\
&\quad + H(W) - \dim(f_{17}\overline{D})\\
&= H(C) + H(W) - \dim(\widetilde{C} \cap f_{28}^{-1} f_{15}\overline{A}) - \dim(\overline{D})\\
&= H(C) + H(W) - H(C) + \operatorname{codim}_C(\widetilde{C} \cap f_{28}^{-1} f_{15}\overline{A})\\
&\quad - H(D) + \operatorname{codim}_D(\overline{D})\\
&= H(W) - H(D) + \operatorname{codim}_C(\widetilde{C} \cap f_{28}^{-1} f_{15}\overline{A}) + \operatorname{codim}_D(\overline{D})\\
&\le H(W) - H(D) + \Delta_{\widetilde{C}} + \operatorname{codim}_C(f_{28}^{-1} f_{15}\overline{A}) + \Delta_{\overline{D}}\\
&\le H(W) - H(D) + \Delta_{\widetilde{C}} + \operatorname{codim}_Y(f_{15}\overline{A}) + \Delta_{\overline{D}}\\
&= H(W) - H(D) + \Delta_{\widetilde{C}} + H(Y) - \dim(f_{15}\overline{A}) + \Delta_{\overline{D}}\\
&= H(W) - H(D) + \Delta_{\widetilde{C}} + H(Y) - \dim(\overline{A}) + \Delta_{\overline{D}}\\
&= H(W) - H(D) + \Delta_{\widetilde{C}} + H(Y) - H(A)\\
&\quad + \operatorname{codim}_A(\overline{A}) + \Delta_{\overline{D}}\\
&\le H(W) - H(D) + H(Y) - H(A) + \Delta_{\widetilde{C}} + \Delta_{\overline{A}} + \Delta_{\overline{D}}.
\end{aligned}
\tag{A.47}
$$

In the justification for (A.44), we concluded that $f_B f_{20}$ is injective on $\overline{C}^*$, which implies f_{20} is injective on $\overline{C}^*$. In the justification for (A.45), we concluded that $f_D f_{28}$ is injective on $\widetilde{C}^*$, which implies f_{28} is injective on $\widetilde{C}^*$. These facts combined with (A.40) will be used to arrive on line (A.48).

$$
\begin{aligned}
&\operatorname{codim}_A(\overline{A}^*)\\
&= H(A) - \dim(f_A(f_{15}\overline{A} \cap f_{20}\overline{C}^* \cap f_{28}\widetilde{C}^*))\\
&= H(A) - \dim(f_{15}\overline{A} \cap f_{20}\overline{C}^* \cap f_{28}\widetilde{C}^*)\\
&= H(A) - H(Y) + \operatorname{codim}_Y(f_{15}\overline{A} \cap f_{20}\overline{C}^* \cap f_{28}\widetilde{C}^*)\\
&\le H(A) - H(Y) + \operatorname{codim}_Y(f_{15}\overline{A}) + \operatorname{codim}_Y(f_{20}\overline{C}^*)\\
&\quad + \operatorname{codim}_Y(f_{28}\widetilde{C}^*)\\
&= H(A) - H(Y) + H(Y) - \dim(f_{15}\overline{A}) + H(Y)\\
&\quad - \dim(f_{20}\overline{C}^*) + H(Y) - \dim(f_{28}\widetilde{C}^*)\\
&= H(A) + 2H(Y) - \dim(\overline{A}) - \dim(\overline{C}^*) - \dim(\widetilde{C}^*)
\end{aligned}
\tag{A.48}
$$

$$
\begin{aligned}
&= H(A) + 2H(Y) - H(A) + \operatorname{codim}_A(\overline{A}) - H(C)\\
&\quad + \operatorname{codim}_C(\overline{C}^*) - H(C) + \operatorname{codim}_C(\widetilde{C}^*)\\
&= 2H(Y) - 2H(C) + \operatorname{codim}_A(\overline{A}) + \operatorname{codim}_C(\overline{C}^*)\\
&\quad + \operatorname{codim}_C(\widetilde{C}^*)\\
&\le 2H(Y) - 2H(C) + \Delta_{\overline{A}}\\
&\quad + H(Z) - H(B) + H(Y) - H(A) + \Delta_{\overline{C}} + \Delta_{\overline{A}} + \Delta_{\overline{B}}\\
&\quad + H(W) - H(D) + H(Y) - H(A) + \Delta_{\widetilde{C}} + \Delta_{\overline{A}} + \Delta_{\overline{D}}\\
&= H(W) + 4H(Y) + H(Z) - 2H(A) - H(B)\\
&\quad - 2H(C) - H(D) + 3\Delta_{\overline{A}} + \Delta_{\overline{B}} + \Delta_{\overline{C}} + \Delta_{\widetilde{C}} + \Delta_{\overline{D}}\\
&\triangleq \Delta_{\overline{A}^*}.
\end{aligned}
\tag{A.49}
$$

Let $t \in A$. We will next make a collection of assumptions on t in (A.50)–(A.55). Each such assumption gives rise to an upper bound on the codimension of a particular subspace of A. The justification of these upper bounds will be given in what follows. Ultimately, we will show that these assumptions imply that $3t = 0$ and thus for field characteristics other than 3, no nonzero t can satisfy this condition. This in turn implies that the codimension of the intersection of the subspaces of A in the upper bounds of (A.50)–(A.55) must be at least as big as the dimension of A, which then yields the desired inequality.

We will assume $t \in \overline{A}^*$.

This is true on a subspace of A of codimension at most $\Delta_{\overline{A}^*}$. (A.50)

We will assume $f_{10}f_{15}t \in f_{19}(\overline{C} \cap f_{20}^{-1} f_{15}\overline{A}^*)$.

This is true on a subspace of A of codimension at most

$$H(Z) - H(C) + H(Y) - H(A) + \Delta_{\overline{C}} + \Delta_{\overline{A}^*}. \tag{A.51}$$

We will assume $f_{11}f_{15}t \in f_{29}(\widetilde{C} \cap f_{28}^{-1} f_{15}\overline{A}^*)$.

This is true on a subspace of A of codimension at most

$$H(W) - H(C) + H(Y) - H(A) + \Delta_{\widetilde{C}} + \Delta_{\overline{A}^*}. \tag{A.52}$$

We will assume $f_{12}f_{15}t \in f_{26}(\widehat{C} \cap f_{25}^{-1} f_{15}\overline{A}^*)$.

This is true on a subspace of A of codimension at most

$$H(X) - H(C) + H(Y) - H(A) + \Delta_{\widehat{C}} + \Delta_{\overline{A}^*}. \tag{A.53}$$

We will assume $f_{10}f_{15}t \in f_{22}(\overline{B} \cap f_{23}^{-1} f_{26}(\widehat{C} \cap f_{25}^{-1} f_{15}\overline{A}^*))$.

This is true on a subspace of A of codimension at most

$$
\begin{aligned}
&H(Z) - H(B) + H(X) - H(C) + H(Y) - H(A)\\
&\qquad + \Delta_{\overline{A}^*} + \Delta_{\overline{B}} + \Delta_{\widehat{C}}.
\end{aligned}
\tag{A.54}
$$

We will assume $f_{11}f_{15}t \in f_{31}(\widehat{B} \cap f_{32}^{-1} f_{26}(\widehat{C} \cap f_{25}^{-1} f_{15}\overline{A}^*))$.

This is true on a subspace of A of codimension at most

$$
\begin{aligned}
&H(W) - H(B) + H(X) - H(C) + H(Y) - H(A)\\
&\qquad + \Delta_{\overline{A}^*} + \Delta_{\widehat{B}} + \Delta_{\widehat{C}}.
\end{aligned}
\tag{A.55}
$$

To justify (A.51), first we know f_{19} is injective on $\overline{C} \cap f_{20}^{-1} f_{15}\overline{A}^*$ by (A.41). Then by Lemma 2.3, we know

$$f_{10}f_{15}t \in f_{19}(\overline{C} \cap f_{20}^{-1} f_{15}\overline{A}^*)$$

on a subspace of A of codimension at most

$$H(Z) - H(C) + \operatorname{codim}_C(\overline{C} \cap f_{20}^{-1} f_{15}\overline{A}^*).$$

By Lemma 2.1, we know

$$\operatorname{codim}_C(\overline{C} \cap f_{20}^{-1} f_{15}\overline{A}^*) \le \Delta_{\overline{C}} + \operatorname{codim}_C(f_{20}^{-1} f_{15}\overline{A}^*).$$

Then using Lemma 2.3 and (A.40), we know

$$
\begin{aligned}
\operatorname{codim}_C(\overline{C} \cap f_{20}^{-1} f_{15}\overline{A}^*) &\le \Delta_{\overline{C}} + \operatorname{codim}_Y(f_{15}\overline{A}^*)\\
&= \Delta_{\overline{C}} + H(Y) - \dim(f_{15}\overline{A}^*)\\
&= \Delta_{\overline{C}} + H(Y) - \dim(\overline{A}^*)\\
&\le \Delta_{\overline{C}} + H(Y) - H(A) + \Delta_{\overline{A}^*}.
\end{aligned}
\tag{A.56}
$$

So, we have

$$f_{10}f_{15}t \in f_{19}(\overline{C} \cap f_{20}^{-1} f_{15}\overline{A}^*)$$

on a subspace of A of codimension at most

$$H(Z) - H(C) + H(Y) - H(A) + \Delta_{\overline{C}} + \Delta_{\overline{A}^*}.$$

To justify (A.52), first we know f_{29} is injective on $\widetilde{C} \cap f_{28}^{-1} f_{15} \overline{A}^*$ by (A.42). Then by Lemma 2.3, we know

$$f_{11} f_{15} t \in f_{29}(\widetilde{C} \cap f_{28}^{-1} f_{15} \overline{A}^*)$$

on a subspace of A of codimension at most

$$H(Z) - H(C) + \mathrm{codim}_C(\widetilde{C} \cap f_{28}^{-1} f_{15} \overline{A}^*).$$

By Lemma 2.1, we know

$$\mathrm{codim}_C(\widetilde{C} \cap f_{28}^{-1} f_{15} \overline{A}^*) \le \Delta_{\widetilde{C}} + \mathrm{codim}_C(f_{28}^{-1} f_{15} \overline{A}^*).$$

Then using Lemma 2.3 and (A.40), we know

$$\begin{aligned} \mathrm{codim}_C(\widetilde{C} \cap f_{28}^{-1} f_{15} \overline{A}^*) &\le \Delta_{\widetilde{C}} + \mathrm{codim}_Y(f_{15} \overline{A}^*) \\ &= \Delta_{\widetilde{C}} + H(Y) - \dim(f_{15} \overline{A}^*) \\ &= \Delta_{\widetilde{C}} + H(Y) - \dim(\overline{A}^*) \\ &\le \Delta_{\widetilde{C}} + H(Y) - H(A) + \Delta_{\overline{A}^*}. \end{aligned} \tag{A.57}$$

So, we have

$$f_{11} f_{15} t \in f_{29}(\widetilde{C} \cap f_{28}^{-1} f_{15} \overline{A}^*)$$

on a subspace of A of codimension at most

$$H(Z) - H(C) + H(Y) - H(A) + \Delta_{\widetilde{C}} + \Delta_{\overline{A}^*}.$$

To justify (A.53), first we know f_{26} is injective on $\widehat{C} \cap f_{25}^{-1} f_{15} \overline{A}^*$ by (A.43). Then by Lemma 2.3, we know

$$f_{12} f_{15} t \in f_{26}(\widehat{C} \cap f_{25}^{-1} f_{15} \overline{A}^*)$$

on a subspace of A of codimension at most

$$H(Z) - H(C) + \mathrm{codim}_C(\widehat{C} \cap f_{25}^{-1} f_{15} \overline{A}^*).$$

By Lemma 2.1, we know

$$\mathrm{codim}_C(\widehat{C} \cap f_{25}^{-1} f_{15} \overline{A}^*) \le \Delta_{\widehat{C}} + \mathrm{codim}_C(f_{25}^{-1} f_{15} \overline{A}^*).$$

Then using Lemma 2.3 and (A.40), we know

$$\begin{aligned} \mathrm{codim}_C(\widehat{C} \cap f_{25}^{-1} f_{15} \overline{A}^*) &\le \Delta_{\widehat{C}} + \mathrm{codim}_Y(f_{15} \overline{A}^*) \\ &= \Delta_{\widehat{C}} + H(Y) - \dim(f_{15} \overline{A}^*) \\ &= \Delta_{\widehat{C}} + H(Y) - \dim(\overline{A}^*) \\ &\le \Delta_{\widehat{C}} + H(Y) - H(A) + \Delta_{\overline{A}^*}. \end{aligned} \tag{A.58}$$

So, we have

$$f_{12} f_{15} t \in f_{26}(\widehat{C} \cap f_{25}^{-1} f_{15} \overline{A}^*)$$

on a subspace of A of codimension at most

$$H(Z) - H(C) + H(Y) - H(A) + \Delta_{\widehat{C}} + \Delta_{\overline{A}^*}.$$

To justify (A.54), we first know f_{22} is injective on $\overline{B} \cap f_{23}^{-1} f_{26}(\widehat{C} \cap f_{25}^{-1} f_{15} \overline{A}^*)$ by (A.17). Then by Lemma 2.3, we know

$$f_{10} f_{15} t \in f_{22}(\overline{B} \cap f_{23}^{-1} f_{26}(\widehat{C} \cap f_{25}^{-1} f_{15} \overline{A}^*))$$

on a subspace of A of codimension at most

$$H(Z) - H(B) + \mathrm{codim}_B(\overline{B} \cap f_{23}^{-1} f_{26}(\widehat{C} \cap f_{25}^{-1} f_{15} \overline{A}^*)).$$

Now again we are going to use Lemma 2.1, Lemma 2.3, and (A.40). Also on line (A.59) we will use the fact that f_{26} is injective on $\widehat{C} \cap f_{25}^{-1} f_{15} \overline{A}^*$ from (A.43).

$$\begin{aligned} &\mathrm{codim}_B(\overline{B} \cap f_{23}^{-1} f_{26}(\widehat{C} \cap f_{25}^{-1} f_{15} \overline{A}^*)) \\ &\le \Delta_{\overline{B}} + \mathrm{codim}_B(f_{23}^{-1} f_{26}(\widehat{C} \cap f_{25}^{-1} f_{15} \overline{A}^*)) \\ &\le \Delta_{\overline{B}} + \mathrm{codim}_X(f_{26}(\widehat{C} \cap f_{25}^{-1} f_{15} \overline{A}^*)) \\ &= \Delta_{\overline{B}} + H(X) - \dim(f_{26}(\widehat{C} \cap f_{25}^{-1} f_{15} \overline{A}^*)) \\ &= \Delta_{\overline{B}} + H(X) - \dim(\widehat{C} \cap f_{25}^{-1} f_{15} \overline{A}^*) \qquad \text{(A.59)} \\ &\le \Delta_{\overline{B}} + H(X) - H(C) + \mathrm{codim}_C(\widehat{C}) + \mathrm{codim}_C(f_{25}^{-1} f_{15} \overline{A}^*) \\ &\le \Delta_{\overline{B}} + H(X) - H(C) + \Delta_{\widehat{C}} + \mathrm{codim}_Y(f_{15} \overline{A}^*) \\ &= \Delta_{\overline{B}} + H(X) - H(C) + \Delta_{\widehat{C}} + H(Y) - \dim(f_{15} \overline{A}^*) \\ &= \Delta_{\overline{B}} + H(X) - H(C) + H(Y) + \Delta_{\widehat{C}} - \dim(\overline{A}^*) \\ &= \Delta_{\overline{B}} + H(X) - H(C) + H(Y) + \Delta_{\widehat{C}} - H(A) \\ &\quad + \mathrm{codim}_A(\overline{A}^*) \\ &\le \Delta_{\overline{B}} + H(X) - H(C) + H(Y) - H(A) + \Delta_{\widehat{C}} + \Delta_{\overline{A}^*}. \end{aligned} \tag{A.60}$$

So, we have

$$f_{10} f_{15} t \in f_{22}(\overline{B} \cap f_{23}^{-1} f_{26}(\widehat{C} \cap f_{25}^{-1} f_{15} \overline{A}^*))$$

on a subspace of A of codimension at most

$$H(Z) - H(B) + H(X) - H(C) + H(Y) - H(A) + \Delta_{\overline{A}^*} + \Delta_{\overline{B}} + \Delta_{\widehat{C}}.$$

To justify (A.55), we first know f_{31} is injective on $\widehat{B} \cap f_{32}^{-1} f_{26}(\widehat{C} \cap f_{25}^{-1} f_{15} \overline{A}^*)$ by (A.21). Then by Lemma 2.3, we know

$$f_{11} f_{15} t \in f_{31}(\widehat{B} \cap f_{32}^{-1} f_{26}(\widehat{C} \cap f_{25}^{-1} f_{15} \overline{A}^*))$$

on a subspace of A of codimension at most

$$H(W) - H(B) + \mathrm{codim}_B(\widehat{B} \cap f_{32}^{-1} f_{26}(\widehat{C} \cap f_{25}^{-1} f_{15} \overline{A}^*)).$$

Now again we are going to use Lemma 2.1 and Lemma 2.3,

$$\begin{aligned} &\mathrm{codim}_B(\widehat{B} \cap f_{32}^{-1} f_{26}(\widehat{C} \cap f_{25}^{-1} f_{15} \overline{A}^*)) \\ &\le \Delta_{\widehat{B}} + \mathrm{codim}_B(f_{32}^{-1} f_{26}(\widehat{C} \cap f_{25}^{-1} f_{15} \overline{A}^*)) \\ &\le \Delta_{\widehat{B}} + \mathrm{codim}_X(f_{26}(\widehat{C} \cap f_{25}^{-1} f_{15} \overline{A}^*)) \\ &\le \Delta_{\widehat{B}} + H(X) - H(C) + H(Y) - H(A) + \Delta_{\widehat{C}} + \Delta_{\overline{A}^*}. \end{aligned}$$

The last line was derived by copying the argument from (A.60). So, we have

$$f_{11} f_{15} t \in f_{31}(\widehat{B} \cap f_{32}^{-1} f_{26}(\widehat{C} \cap f_{25}^{-1} f_{15} \overline{A}^*))$$

on a subspace of A of codimension at most

$$H(W) - H(B) + H(X) - H(C) + H(Y) - H(A) + \Delta_{\overline{A}^*} + \Delta_{\widehat{B}} + \Delta_{\widehat{C}}.$$

From (A.51) and (A.54) we know $\exists \overline{c} \in \overline{C}, \overline{b} \in \overline{B}$ such that

$$\begin{aligned} f_{10} f_{15} t &= f_{19} \overline{c} = f_{22} \overline{b} \\ \text{where} \quad f_{20} \overline{c} &\in f_{15} \overline{A}^* \text{ and} \\ f_{23} \overline{b} &\in f_{26}(\widehat{C} \cap f_{25}^{-1} f_{15} \overline{A}^*). \end{aligned} \tag{A.61}$$

From (A.52) and (A.55) we know $\exists \widetilde{c} \in \widetilde{C}, \widehat{b} \in \widehat{B}$ such that

$$\begin{aligned} & f_{11} f_{15} t = f_{29} \widetilde{c} = f_{31} \widehat{b} \\ \text{where} \quad & f_{28} \widetilde{c} \in f_{15} \overline{A}^* \text{ and} \\ & f_{32} \widehat{b} \in f_{26}(\widehat{C} \cap f_{25}^{-1} f_{15} \overline{A}^*). \end{aligned} \tag{A.62}$$

From (A.53) we know $\exists \widehat{c} \in \widehat{C}$ such that

$$\begin{aligned} f_{12} f_{15} t &= f_{26} \widehat{c} \\ \text{where } f_{25} \widehat{c} &\in f_{15} \overline{A}^*. \end{aligned} \tag{A.63}$$

From (A.12) and (A.13), we know

$$\begin{aligned} f_B f_{15} &= -f_{13} \text{ on } \overline{A} \\ f_B &= -f_{13} f_A \text{ on } f_{15} \overline{A}. \end{aligned} \tag{A.64}$$

From (A.12) and (A.15), we know

$$\begin{aligned} f_D f_{15} &= -f_{14} \text{ on } \overline{A} \\ f_D &= -f_{14} f_A \text{ on } f_{15} \overline{A}. \end{aligned} \tag{A.65}$$

From (A.12) we have

$$f_7 f_{12} f_{15} t + f_1 f_{10} f_{15} t = t.$$

Then (A.63), (A.61), (A.28), and (A.24) give

$$\begin{aligned} f_7 f_{12} f_{15} t + f_1 f_{10} f_{15} t &= t \\ f_7 f_{26} \widehat{c} + f_1 f_{19} \overline{c} &= t \\ -f_A f_{25} \widehat{c} - f_A f_{20} \overline{c} &= t \\ f_A f_{25} \widehat{c} + f_A f_{20} \overline{c} &= -t. \end{aligned} \tag{A.66}$$

From (A.13) we have

$$f_4 f_{11} f_{15} t + f_2 f_{10} f_{15} t = -f_{13} t.$$

Then (A.62), (A.61), (A.33), and (A.25) give

$$\begin{aligned} f_4 f_{11} f_{15} t + f_2 f_{10} f_{15} t &= -f_{13} t \\ f_4 f_{29} \widetilde{c} + f_2 f_{19} \overline{c} &= -f_{13} t \\ -f_B f_{28} \widetilde{c} - f_B f_{20} \overline{c} &= -f_{13} t. \end{aligned}$$

By (A.62) and (A.61), we know $f_{28} \widetilde{c} \in f_{15} \overline{A}^*$ and $f_{20} \overline{c} \in f_{15} \overline{A}^*$. Now by (A.64), we have

$$\begin{aligned} -f_B f_{28} \widetilde{c} - f_B f_{20} \overline{c} &= -f_{13} t \\ f_{13} f_A f_{28} \widetilde{c} + f_{13} f_A f_{20} \overline{c} &= -f_{13} t. \end{aligned}$$

Then using (A.12), we know $f_A f_{28} \widetilde{c} \in \overline{A}^*$ and $f_A f_{20} \overline{c} \in \overline{A}^*$. By (A.44), we have

$$\begin{aligned} f_{13} f_A f_{28} \widetilde{c} + f_{13} f_A f_{20} \overline{c} &= -f_{13} t \\ f_A f_{28} \widetilde{c} + f_A f_{20} \overline{c} &= -t. \end{aligned} \tag{A.67}$$

From (A.15) we have

$$f_9 f_{12} f_{15} t + f_6 f_{11} f_{15} t = -f_{14} t.$$

Then (A.63), (A.62), (A.35), and (A.31) give

$$\begin{aligned} f_9 f_{12} f_{15} t + f_6 f_{11} f_{15} t &= -f_{14} t \\ f_9 f_{26} \widehat{c} + f_6 f_{29} \widetilde{c} &= -f_{14} t \\ -f_D f_{25} \widehat{c} + -f_D f_{28} \widetilde{c} &= -f_{14} t. \end{aligned}$$

By (A.63) and (A.62), we know $f_{25} \widehat{c} \in f_{15} \overline{A}^*$ and $f_{28} \widetilde{c} \in f_{15} \overline{A}^*$. Now by (A.65), we have

$$\begin{aligned} -f_D f_{25} \widehat{c} + -f_D f_{28} \widetilde{c} &= -f_{14} t \\ f_{14} f_A f_{25} \widehat{c} + f_{14} f_A f_{28} \widetilde{c} &= -f_{14} t. \end{aligned}$$

Then using (A.12), we know $f_A f_{25} \widehat{c} \in \overline{A}^*$ and $f_A f_{28} \widetilde{c} \in \overline{A}^*$. By (A.45), we have

$$\begin{aligned} f_{14} f_A f_{25} \widehat{c} + f_{14} f_A f_{28} \widetilde{c} &= -f_{14} t \\ f_A f_{25} \widehat{c} + f_A f_{28} \widetilde{c} &= -t. \end{aligned} \tag{A.68}$$

From (A.24) and (A.41), we know

$$\begin{aligned} f_1 f_{19} &= -f_A f_{20} \text{ on } \overline{C} \\ f_1 &= -f_A f_{20} f_3 \text{ on } f_{19}(\overline{C} \cap f_{20}^{-1} f_{15} \overline{A}^*). \end{aligned} \tag{A.69}$$

From (A.28) and (A.43), we know

$$\begin{aligned} f_7 f_{26} &= -f_A f_{25} \text{ on } \widehat{C} \\ f_7 &= -f_A f_{25} f_8 \text{ on } f_{26}(\widehat{C} \cap f_{25}^{-1} f_{15} \overline{A}^*). \end{aligned} \tag{A.70}$$

From (A.16), we have

$$f_7 f_{23} \overline{b} + f_1 f_{22} \overline{b} = 0.$$

By (A.61), we know

$$f_{23} \overline{b} \in f_{26}(\widehat{C} \cap f_{25}^{-1} f_{15} \overline{A}^*).$$

By (A.61), we also know

$$f_{22} \overline{b} = f_{19} \overline{c}$$

which implies

$$f_{22} \overline{b} \in f_{19}(\overline{C} \cap f_{20}^{-1} f_{15} \overline{A}^*).$$

Now we can apply (A.69) and (A.70) to give us

$$\begin{aligned} f_7 f_{23} \overline{b} + f_1 f_{22} \overline{b} &= 0 \\ -f_A f_{25} f_8 f_{23} \overline{b} - f_A f_{20} f_3 f_{22} \overline{b} &= 0. \end{aligned}$$

Now using (A.18), (A.61), and (A.41), we have

$$\begin{aligned} -f_A f_{25} f_8 f_{23} \overline{b} - f_A f_{20} f_3 f_{22} \overline{b} &= 0 \\ f_A f_{25} f_3 f_{22} \overline{b} - f_A f_{20} f_3 f_{22} \overline{b} &= 0 \\ f_A f_{25} f_3 f_{22} \overline{b} &= f_A f_{20} f_3 f_{22} \overline{b} \\ f_A f_{25} f_3 f_{19} \overline{c} &= f_A f_{20} f_3 f_{19} \overline{c} \\ f_A f_{25} \overline{c} &= f_A f_{20} \overline{c}. \end{aligned} \tag{A.71}$$

From (A.31) and (A.43), we know

$$\begin{aligned} f_9 f_{26} &= -f_D f_{25} \text{ on } \widehat{C} \\ f_9 &= -f_D f_{25} f_8 \text{ on } f_{26}(\widehat{C} \cap f_{25}^{-1} f_{15} \overline{A}^*). \end{aligned} \tag{A.72}$$

From (A.35) and (A.42), we know

$$\begin{aligned} f_6 f_{29} &= -f_D f_{28} \text{ on } \widetilde{C} \\ f_6 &= -f_D f_{28} f_5 \text{ on } f_{29}(\widetilde{C} \cap f_{28}^{-1} f_{15} \overline{A}^*). \end{aligned} \tag{A.73}$$

From (A.23), we have

$$f_9 f_{32} \widehat{b} + f_6 f_{31} \widehat{b} = 0.$$

From (A.62) we know $f_{31}\widehat{b} = f_{29}\widetilde{c}$ so $f_{31}\widehat{b} \in f_{29}(\widetilde{C} \cap f_{28}^{-1} f_{15}\overline{A}^*)$. From (A.62) we also know that

$$f_{32}\widehat{b} \in f_{26}(\widehat{C} \cap f_{25}^{-1} f_{15}\overline{A}^*)$$

so (A.72) and (A.73) give us

$$\begin{aligned} f_9 f_{32}\widehat{b} + f_6 f_{31}\widehat{b} &= 0 \\ -f_D f_{25} f_8 f_{32}\widehat{b} - f_D f_{28} f_5 f_{31}\widehat{b} &= 0. \end{aligned}$$

From (A.62), we know

$$f_{32}\widehat{b} \in f_{26}(\widehat{C} \cap f_{25}^{-1} f_{15}\overline{A}^*).$$

From (A.43), we know $f_8 f_{26} = I$ on $\widehat{C} \cap f_{25}^{-1} f_{15}\overline{A}^*$. So $f_8 f_{32}\widehat{b} \in f_{25}^{-1} f_{15}\overline{A}^*$, which implies $f_{25} f_8 f_{32}\widehat{b} \in f_{15}\overline{A}^*$. By (A.62) and (A.42), we know

$$f_{28} f_5 f_{31}\widehat{b} = f_{28} f_5 f_{29}\widetilde{c} = f_{28}\widetilde{c} \in f_{15}\overline{A}^*.$$

Now we can apply (A.65) to give us

$$\begin{aligned} -f_D f_{25} f_8 f_{32}\widehat{b} - f_D f_{28} f_5 f_{31}\widehat{b} &= 0 \\ f_{14} f_A f_{25} f_8 f_{32}\widehat{b} + f_{14} f_A f_{28} f_5 f_{31}\widehat{b} &= 0. \end{aligned}$$

Since we already established that $f_{25} f_8 f_{32}\widehat{b} \in f_{15}\overline{A}^*$ and $f_{28} f_5 f_{31}\widehat{b} \in f_{15}\overline{A}^*$, by (A.12) and (A.45) we know

$$\begin{aligned} f_{14} f_A f_{25} f_8 f_{32}\widehat{b} + f_{14} f_A f_{28} f_5 f_{31}\widehat{b} &= 0 \\ f_A f_{25} f_8 f_{32}\widehat{b} + f_A f_{28} f_5 f_{31}\widehat{b} &= 0. \end{aligned}$$

Now by (A.22)

$$\begin{aligned} f_A f_{25} f_8 f_{32}\widehat{b} + f_A f_{28} f_5 f_{31}\widehat{b} &= 0 \\ -f_A f_{25} f_5 f_{31}\widehat{b} + f_A f_{28} f_5 f_{31}\widehat{b} &= 0 \\ f_A f_{25} f_5 f_{31}\widehat{b} &= f_A f_{28} f_5 f_{31}\widehat{b}. \end{aligned}$$

By (A.62) and (A.42), we have

$$\begin{aligned} f_A f_{25} f_5 f_{31}\widehat{b} &= f_A f_{28} f_5 f_{31}\widehat{b} \\ f_A f_{25} f_5 f_{29}\widetilde{c} &= f_A f_{28} f_5 f_{29}\widetilde{c} \\ f_A f_{25}\widetilde{c} &= f_A f_{28}\widetilde{c}. \end{aligned} \tag{A.74}$$

Now adding (A.66), (A.67), and (A.68), we have

$$-3t = 2(f_A f_{20}\overline{c} + f_A f_{25}\widehat{c} + f_A f_{28}\widetilde{c}).$$

Now using (A.71) and (A.74) we have

$$\begin{aligned} -3t &= 2(f_A f_{25}\overline{c} + f_A f_{25}\widehat{c} + f_A f_{25}\widetilde{c}) \\ -3t &= 2 f_A f_{25}(\overline{c} + \widehat{c} + \widetilde{c}). \end{aligned}$$

By (A.41), (A.42), and (A.43) we know

$$-3t = 2 f_A f_{25}(f_3 f_{19}\overline{c} + f_8 f_{26}\widehat{c} + f_5 f_{29}\widetilde{c}).$$

By (A.61), (A.62), (A.63), and (A.14), we have

$$\begin{aligned} -3t &= 2 f_A f_{25}(f_3 f_{10} f_{15} t + f_8 f_{12} f_{15} t + f_5 f_{11} f_{15} t) \\ -3t &= 2 f_A f_{25}(0) \\ 3t &= 0. \end{aligned} \tag{A.75}$$

Thus if the field is of characteristic other than 3, then no nonzero t can satisfy conditions (A.50)–(A.55). Therefore the sum of the codimensions given in the assumptions must be at least the dimension of A. So we have a linear rank inequality for fields of characteristic other than 3:

$$\begin{aligned} &H(A) \\ &\le \Delta_{\overline{A}^*} + H(Z) - H(C) + H(Y) - H(A) + \Delta_{\overline{C}} + \Delta_{\overline{A}^*} \\ &\quad + H(W) - H(C) + H(Y) - H(A) + \Delta_{\widetilde{C}} + \Delta_{\overline{A}^*} \\ &\quad + H(X) - H(C) + H(Y) - H(A) + \Delta_{\widehat{C}} + \Delta_{\overline{A}^*} \\ &\quad + H(Z) - H(B) + H(X) - H(C) + H(Y) - H(A) \\ &\quad + \Delta_{\overline{A}^*} + \Delta_{\overline{B}} + \Delta_{\widehat{C}} \\ &\quad + H(W) - H(B) + H(X) - H(C) + H(Y) - H(A) \\ &\quad + \Delta_{\overline{A}^*} + \Delta_{\widehat{B}} + \Delta_{\widehat{C}} \\ &= 2H(Z) + 5H(Y) + 3H(X) + 2H(W) - 5H(A) \\ &\quad - 2H(B) - 5H(C) \\ &\quad + 6\Delta_{\overline{A}^*} + \Delta_{\overline{B}} + \Delta_{\widehat{B}} + \Delta_{\overline{C}} + \Delta_{\widetilde{C}} + 3\Delta_{\widehat{C}} \\ &= 2H(Z) + 5H(Y) + 3H(X) + 2H(W) - 5H(A) \\ &\quad - 2H(B) - 5H(C) \\ &\quad + 6\,(H(W) + 4H(Y) + H(Z) \\ &\qquad - 2H(A) - H(B) - 2H(C) - H(D)) \\ &\quad + 6(3\Delta_{\overline{A}} + \Delta_{\overline{B}} + \Delta_{\overline{C}} + \Delta_{\widetilde{C}} + \Delta_{\overline{D}}) \\ &\quad + \Delta_{\overline{B}} + \Delta_{\widehat{B}} + \Delta_{\overline{C}} + \Delta_{\widetilde{C}} + 3\Delta_{\widehat{C}} \\ &= 8H(Z) + 29H(Y) + 3H(X) + 8H(W) \\ &\quad - 6H(D) - 17H(C) - 8H(B) - 17H(A) \\ &\quad + 18\Delta_{\overline{A}} + 7\Delta_{\overline{B}} + \Delta_{\widehat{B}} + 7\Delta_{\overline{C}} + 7\Delta_{\widetilde{C}} + 3\Delta_{\widehat{C}} + 6\Delta_{\overline{D}} \\ &= 8H(Z) + 29H(Y) + 3H(X) + 8H(W) - 6H(D) \\ &\quad - 17H(C) - 8H(B) - 17H(A) \\ &\quad + 55H(Z|A, B, C) + 35H(Y|W, X, Z) \\ &\quad + 50H(X|A, C, D) + 49H(W|B, C, D) \\ &\quad + 18H(A|B, D, Y) + 7H(B|D, X, Z) + H(B|A, W, X) \\ &\quad + 7H(C|D, Y, Z) \\ &\quad + 7H(C|B, X, Y) + 3H(C|A, W, Y) + 6H(D|A, W, Z) \\ &\quad + 49(H(A) + H(B) + H(C) + H(D) - H(A, B, C, D)). \end{aligned}$$

■

Proof of Theorem 4.1: The proof of this theorem follows a similar strategy as discussed at the beginning of the proof of Theorem 3.1, and again, is rather tedious.

By Lemma 2.4 we get linear functions:

$$\begin{aligned} &f_1 : W \to B, \quad f_2 : W \to C, \quad f_3 : W \to D, \\ &f_4 : X \to A, \quad f_5 : X \to C, \quad f_6 : X \to D, \\ &f_7 : Y \to A, \quad f_8 : Y \to B, \quad f_9 : Y \to D, \\ &f_{10} : Z \to A, \quad f_{11} : Z \to B, \quad f_{12} : Z \to C, \\ &f_{13} : A \to B, \quad f_{14} : A \to W, \quad f_{15} : A \to X, \\ &f_{16} : C \to A, \quad f_{17} : C \to W, \quad f_{18} : C \to Y, \\ &f_{19} : B \to C, \quad f_{20} : B \to X, \quad f_{21} : B \to Y, \\ &f_{22} : D \to W, \quad f_{23} : D \to A, \quad f_{24} : D \to Z, \\ &f_{25} : B \to X, \quad f_{26} : B \to D, \quad f_{27} : B \to Z, \\ &f_{28} : C \to Y, \quad f_{29} : C \to Z, \quad f_{30} : C \to D, \\ &\qquad f_{31} : A \to W, \quad f_{32} : A \to X \\ &\qquad f_{33} : A \to Y, \quad f_{34} : A \to Z \end{aligned}$$

such that

$$f_1 + f_2 + f_3 = I \text{ on a subspace of } W \text{ of codimension } H(W|B, C, D) \tag{A.76}$$

$$f_4 + f_5 + f_6 = I \text{ on a subspace of } X \text{ of codimension } H(X|A, C, D) \tag{A.77}$$

$$f_7 + f_8 + f_9 = I \text{ on a subspace of } Y \text{ of codimension } H(Y|A, B, D) \tag{A.78}$$

$$f_{10} + f_{11} + f_{12} = I \text{ on a subspace of } Z \text{ of codimension } H(Z|A, B, C) \tag{A.79}$$

$$f_{13} + f_{14} + f_{15} = I \text{ on a subspace of } A \text{ of codimension } H(A|B, W, X) \tag{A.80}$$

$$f_{16} + f_{17} + f_{18} = I \text{ on a subspace of } C \text{ of codimension } H(C|A, W, Y) \tag{A.81}$$

$$f_{19} + f_{20} + f_{21} = I \text{ on a subspace of } B \text{ of codimension } H(B|C, X, Y) \tag{A.82}$$

$$f_{22} + f_{23} + f_{24} = I \text{ on a subspace of } D \text{ of codimension } H(D|A, W, Z) \tag{A.83}$$

$$f_{25} + f_{26} + f_{27} = I \text{ on a subspace of } B \text{ of codimension } H(B|D, X, Z) \tag{A.84}$$

$$f_{28} + f_{29} + f_{30} = I \text{ on a subspace of } C \text{ of codimension } H(C|D, Y, Z) \tag{A.85}$$

$$f_{31} + f_{32} + f_{33} + f_{34} = I \text{ on a subspace of } A \text{ of codimension } H(A|W, X, Y, Z). \tag{A.86}$$

Now combining some functions we obtained from Lemma 2.4 gives four new functions:

$$\begin{aligned} f_4 f_{32} + f_7 f_{33} + f_{10} f_{34} &: A \to A \\ f_1 f_{31} + f_8 f_{33} + f_{11} f_{34} &: A \to B \\ f_2 f_{31} + f_5 f_{32} + f_{12} f_{34} &: A \to C \\ f_3 f_{31} + f_6 f_{32} + f_9 f_{33} &: A \to D. \end{aligned}$$

Using (A.76)–(A.79), (A.86), Lemma 2.1, and Lemma 2.3 we know the sum of these four functions is equal to I on a subspace of A of codimension at most

$$\begin{aligned} H(W|B, C, D) + H(X|A, C, D) + H(Y|A, B, D) \\ + H(Z|A, B, C) + H(A|W, X, Y, Z) \end{aligned}$$

since, on that subspace,

$$\begin{aligned} &(f_4 f_{32} + f_7 f_{33} + f_{10} f_{34}) + (f_1 f_{31} + f_8 f_{33} + f_{11} f_{34}) \\ &+ (f_2 f_{31} + f_5 f_{32} + f_{12} f_{34}) + (f_3 f_{31} + f_6 f_{32} + f_9 f_{33}) \\ &\quad = (f_1 + f_2 + f_3) f_{31} + (f_4 + f_5 + f_6) f_{32} \\ &\qquad + (f_7 + f_8 + f_9) f_{33} + (f_{10} + f_{11} + f_{12}) f_{34} \\ &\quad = f_{31} + f_{32} + f_{33} + f_{34} \\ &\quad = I. \end{aligned}$$

Now applying Lemma 2.6 and Lemma 2.1 to the functions

$$\begin{aligned} &f_4 f_{32} + f_7 f_{33} + f_{10} f_{34} - I \\ &f_1 f_{31} + f_8 f_{33} + f_{11} f_{34} \\ &f_2 f_{31} + f_5 f_{32} + f_{12} f_{34} \\ &f_3 f_{31} + f_6 f_{32} + f_9 f_{33} \end{aligned}$$

we get a subspace $\widehat{A}$ of A of codimension at most

$$\begin{aligned} \Delta_{\widehat{A}} = {} & H(W|B, C, D) + H(X|A, C, D) + H(Y|A, B, D) \\ & + H(Z|A, B, C) + H(A|W, X, Y, Z) \\ & + H(A) + H(B) + H(C) + H(D) - H(A, B, C, D) \end{aligned}$$

on which

$$f_4 f_{32} + f_7 f_{33} + f_{10} f_{34} = I \tag{A.87}$$

$$f_1 f_{31} + f_8 f_{33} + f_{11} f_{34} = 0 \tag{A.88}$$

$$f_2 f_{31} + f_5 f_{32} + f_{12} f_{34} = 0 \tag{A.89}$$

$$f_3 f_{31} + f_6 f_{32} + f_9 f_{33} = 0. \tag{A.90}$$

Similarly, we get a subspace $\overline{A}$ of A of codimension at most

$$\begin{aligned} \Delta_{\overline{A}} = {} & H(W|B, C, D) + H(X|A, C, D) + H(A|B, W, X) \\ & + H(A) + H(B) + H(C) + H(D) - H(A, B, C, D) \end{aligned}$$

on which

$$f_4 f_{15} = I \tag{A.91}$$

$$f_{13} + f_1 f_{14} = 0 \tag{A.92}$$

$$f_2 f_{14} + f_5 f_{15} = 0 \tag{A.93}$$

$$f_3 f_{14} + f_6 f_{15} = 0. \tag{A.94}$$

We get a subspace $\overline{B}$ of B of codimension at most

$$\begin{aligned} \Delta_{\overline{B}} = {} & H(X|A, C, D) + H(Y|A, B, D) + H(B|C, X, Y) \\ & + H(A) + H(B) + H(C) + H(D) - H(A, B, C, D) \end{aligned}$$

on which

$$f_4 f_{20} + f_7 f_{21} = 0 \tag{A.95}$$

$$f_8 f_{21} = I \tag{A.96}$$

$$f_{19} + f_5 f_{20} = 0 \tag{A.97}$$

$$f_6 f_{20} + f_9 f_{21} = 0. \tag{A.98}$$

We get a subspace $\widehat{B}$ of B of codimension at most

$$\begin{aligned} \Delta_{\widehat{B}} = {} & H(X|A, C, D) + H(Z|A, B, C) + H(B|D, X, Z) \\ & + H(A) + H(B) + H(C) + H(D) - H(A, B, C, D) \end{aligned}$$

on which

$$f_4 f_{25} + f_{10} f_{27} = 0 \tag{A.99}$$
$$f_{11} f_{27} = I \tag{A.100}$$
$$f_5 f_{25} + f_{12} f_{27} = 0 \tag{A.101}$$
$$f_6 f_{25} + f_{26} = 0. \tag{A.102}$$

We get a subspace $\overline{C}$ of C of codimension at most

$$\Delta_{\overline{C}} = H(W|B,C,D) + H(Y|A,B,D) + H(C|A,W,Y) + H(A) + H(B) + H(C) + H(D) - H(A,B,C,D)$$

on which

$$f_{16} + f_7 f_{18} = 0 \tag{A.103}$$
$$f_1 f_{17} + f_8 f_{18} = 0 \tag{A.104}$$
$$f_2 f_{17} = I \tag{A.105}$$
$$f_3 f_{17} + f_9 f_{18} = 0. \tag{A.106}$$

We get a subspace $\widehat{C}$ of C of codimension at most

$$\Delta_{\widehat{C}} = H(Y|A,B,D) + H(Z|A,B,C) + H(C|D,Y,Z) + H(A) + H(B) + H(C) + H(D) - H(A,B,C,D)$$

on which

$$f_7 f_{28} + f_{10} f_{29} = 0 \tag{A.107}$$
$$f_8 f_{28} + f_{11} f_{29} = 0 \tag{A.108}$$
$$f_{12} f_{29} = I \tag{A.109}$$
$$f_9 f_{28} + f_{30} = 0. \tag{A.110}$$

We get a subspace $\overline{D}$ of D of codimension at most

$$\Delta_{\overline{D}} = H(W|B,C,D) + H(Z|A,B,C) + H(D|A,W,Z) + H(A) + H(B) + H(C) + H(D) - H(A,B,C,D)$$

on which

$$f_{23} + f_{10} f_{24} = 0 \tag{A.111}$$
$$f_1 f_{22} + f_{11} f_{24} = 0 \tag{A.112}$$
$$f_2 f_{22} + f_{12} f_{24} = 0 \tag{A.113}$$
$$f_3 f_{22} = I. \tag{A.114}$$

Let

$$\widehat{B}^* = f_{11}(f_{27}\widehat{B} \cap f_{29}\widehat{C}) \subseteq \widehat{B}.$$

Considering (A.100) and (A.109), we can apply Lemma 2.7 to show that $f_{12} f_{27}$ is injective on $\widehat{B}^*$. By (A.101), we know

$$f_5 f_{25} \text{ is injective on } \widehat{B}^*. \tag{A.115}$$

Let

$$\widehat{C}^* = f_{12}(f_{29}\widehat{C} \cap f_{27}\widehat{B}) \subseteq \widehat{C}.$$

Considering again (A.100) and (A.109), we can apply Lemma 2.7 to show that $f_{11} f_{29}$ is injective on $\widehat{C}^*$. By (A.108), we know

$$f_8 f_{28} \text{ is injective on } \widehat{C}^*. \tag{A.116}$$

Let

$$\overline{A}^* = f_4(f_{15}\overline{A} \cap f_{25}\widehat{B}^*) \subseteq \overline{A}.$$

Considering (A.91) and (A.115), we can apply Lemma 2.7 to show that $f_5 f_{15}$ is injective on $\overline{A}^*$. By (A.93), we know $f_2 f_{14}$ is injective on $\overline{A}^*$ which implies

$$f_{14} \text{ is injective on } \overline{A}^*. \tag{A.117}$$

Let

$$\overline{C}^* = f_2(f_{17}\overline{C} \cap f_{22}\overline{D}) \subseteq \overline{C}.$$

Considering (A.105) and (A.114),we can apply Lemma 2.7 to show that $f_3 f_{17}$ is injective on $\overline{C}^*$. Then by (A.106), we know

$$f_9 f_{18} \text{ is injective on } \overline{C}^*. \tag{A.118}$$

Let

$$\overline{B}^* = f_8(f_{21}\overline{B} \cap f_{18}\overline{C}^*) \subseteq \overline{B}.$$

Considering (A.96) and (A.118), we can apply Lemma 2.7 to show that

$$f_9 f_{21} \text{ is injective on } \overline{B}^*. \tag{A.119}$$

By (A.98), we know

$$f_6 f_{20} \text{ is injective on } \overline{B}^* \tag{A.120}$$

which implies

$$f_{20} \text{ is injective on } \overline{B}^*. \tag{A.121}$$

Let us define the functions

$$g_{14} = (f_{14}|\overline{A}^*)^{-1}$$
$$g_{20} = (f_{20}|\overline{B}^*)^{-1}$$

where $f_{14}|\overline{A}^*$ and $f_{20}|\overline{B}^*$ are the restrictions of the functions f_{14} and f_{20} to the sets $\overline{A}^*$ and $\overline{B}^*$, respectively. Now, considering (A.96), (A.100), (A.105), and (A.109) we have

$$f_1 = -f_8 f_{18} f_2 \text{ on } f_{17}\overline{C} \quad \text{[from (A.104)]} \tag{A.122}$$
$$f_2 = -f_5 f_{15} g_{14} \text{ on } f_{14}\overline{A}^* \quad \text{[from (A.93)]} \tag{A.123}$$
$$f_3 = -f_6 f_{15} g_{14} \text{ on } f_{14}\overline{A}^* \text{ and}$$
$$f_3 = -f_9 f_{18} f_2 \text{ on } f_{17}\overline{C} \quad \text{[from (A.94), (A.106)]} \tag{A.124}$$
$$f_4 = -f_7 f_{21} g_{20} \text{ on } f_{20}\overline{B}^* \quad \text{[from (A.95)]} \tag{A.125}$$
$$f_6 = -f_9 f_{21} g_{20} \text{ on } f_{20}\overline{B}^* \quad \text{[from (A.98)]} \tag{A.126}$$
$$f_7 = -f_4 f_{20} f_8 \text{ on } f_{21}\overline{B} \quad \text{[from (A.95)]} \tag{A.127}$$
$$f_9 = -f_6 f_{20} f_8 \text{ on } f_{21}\overline{B} \quad \text{[from (A.98)]} \tag{A.128}$$
$$f_{10} = -f_4 f_{25} f_{11} \text{ on } f_{27}\widehat{B} \text{ and}$$
$$f_{10} = -f_7 f_{28} f_{12} \text{ on } f_{29}\widehat{C} \quad \text{[from (A.99), (A.107)]} \tag{A.129}$$
$$f_{11} = -f_8 f_{28} f_{12} \text{ on } f_{29}\widehat{C} \quad \text{[from (A.108)]} \tag{A.130}$$
$$f_{12} = -f_5 f_{25} f_{11} \text{ on } f_{27}\widehat{B}. \quad \text{[from (A.101)]} \tag{A.131}$$

Next, we provide upper bounds for the codimensions of $\overline{A}^*$, $\widehat{B}^*$, $\overline{B}^*$, $\widehat{C}^*$, and $\overline{C}^*$. From (A.100), we know f_{11} is injective on $f_{27}\widehat{B}$ and f_{27} is injective on $\widehat{B}$. These facts will be used to arrive on lines (A.132) and (A.134). From (A.109), we know

f_{29} is injective on $\widehat{C}$, which will also be used to arrive on line (A.134). Lemma 2.1 will be used to arrive on (A.133).

$$\begin{aligned}
&\operatorname{codim}_B \widehat{B}^* \\
&\quad = H(B) - \dim(\widehat{B}^*) \\
&\quad = H(B) - \dim(f_{11}(f_{27}\widehat{B} \cap f_{29}\widehat{C})) \\
&\quad = H(B) - \dim(f_{27}\widehat{B} \cap f_{29}\widehat{C}) && \text{(A.132)} \\
&\quad = H(B) - H(Z) + \operatorname{codim}_Z(f_{27}\widehat{B} \cap f_{29}\widehat{C}) \\
&\quad \le H(B) - H(Z) + \operatorname{codim}_Z(f_{27}\widehat{B}) + \operatorname{codim}_Z(f_{29}\widehat{C}) && \text{(A.133)} \\
&\quad = H(B) - H(Z) + H(Z) - \dim(f_{27}\widehat{B}) \\
&\qquad + H(Z) - \dim(f_{29}\widehat{C}) \\
&\quad = H(B) + H(Z) - \dim(\widehat{B}) - \dim(\widehat{C}) && \text{(A.134)} \\
&\quad \le H(B) + H(Z) - H(B) + \Delta_{\widehat{B}} - H(C) + \Delta_{\widehat{C}} && \text{(A.135)} \\
&\quad \le H(Z) - H(C) + \Delta_{\widehat{B}} + \Delta_{\widehat{C}} && \text{(A.136)} \\
&\quad \triangleq \Delta_{\widehat{B}^*}.
\end{aligned}$$

From (A.91), we know f_4 is injective on $f_{15}\overline{A}$ and f_{15} is injective on $\overline{A}$. These facts will be used on lines (A.137) and (A.139). From (A.115), we know f_{25} is injective on $\widehat{B}^*$, which will also be used to arrive on line (A.139). Lemma 2.1 will be used to arrive on (A.138).

$$\begin{aligned}
&\operatorname{codim}_A \overline{A}^* \\
&\quad = H(A) - \dim(\overline{A}^*) \\
&\quad = H(A) - \dim(f_4(f_{25}\widehat{B}^* \cap f_{15}\overline{A})) \\
&\quad = H(A) - \dim(f_{25}\widehat{B}^* \cap f_{15}\overline{A}) && \text{(A.137)} \\
&\quad = H(A) - H(X) + \operatorname{codim}_X(f_{25}\widehat{B}^* \cap f_{15}\overline{A}) \\
&\quad \le H(A) - H(X) + \operatorname{codim}_X(f_{25}\widehat{B}^*) + \operatorname{codim}_X(f_{15}\overline{A}) && \text{(A.138)} \\
&\quad = H(A) + H(X) - \dim(f_{25}\widehat{B}^*) - \dim(f_{15}\overline{A}) \\
&\quad = H(A) + H(X) - \dim(\widehat{B}^*) - \dim(\overline{A}) && \text{(A.139)} \\
&\quad \le H(A) + H(X) - H(B) + \Delta_{\widehat{B}^*} - H(A) + \Delta_{\overline{A}} \\
&\quad = H(X) - H(B) + H(Z) - H(C) + \Delta_{\widehat{B}} + \Delta_{\widehat{C}} + \Delta_{\overline{A}} \\
&\quad \triangleq \Delta_{\overline{A}^*}.
\end{aligned}$$

From (A.105), we know f_2 is injective on $f_{17}\overline{C}$ and f_{17} is injective on $\overline{C}$. These facts will be used to arrive on lines (A.140) and (A.142). From (A.114), we know f_{22} is injective on $\overline{D}$, which will also be used on line (A.142). Lemma 2.1 will be used to arrive on (A.141).

$$\begin{aligned}
&\operatorname{codim}_C \overline{C}^* \\
&\quad = H(C) - \dim(\overline{C}^*) \\
&\quad = H(C) - \dim(f_2(f_{17}\overline{C} \cap f_{22}\overline{D})) \\
&\quad = H(C) - \dim(f_{17}\overline{C} \cap f_{22}\overline{D}) && \text{(A.140)} \\
&\quad = H(C) - H(W) + \operatorname{codim}_W(f_{17}\overline{C} \cap f_{22}\overline{D}) \\
&\quad \le H(C) - H(W) + \operatorname{codim}_W(f_{17}\overline{C}) + \operatorname{codim}_W(f_{22}\overline{D}) && \text{(A.141)} \\
&\quad = H(C) - H(W) + H(W) - \dim(f_{17}\overline{C}) \\
&\qquad + H(W) - \dim(f_{22}\overline{D}) \\
&\quad = H(C) + H(W) - \dim(\overline{C}) - \dim(\overline{D}) && \text{(A.142)} \\
&\quad \le H(C) + H(W) - H(C) + \Delta_{\overline{C}} - H(D) + \Delta_{\overline{D}} \\
&\quad = H(W) - H(D) + \Delta_{\overline{C}} + \Delta_{\overline{D}} \\
&\quad \triangleq \Delta_{\overline{C}^*}.
\end{aligned}$$

From (A.96), we know f_8 is injective on $f_{21}\overline{B}$ and f_{21} is injective on $\overline{B}$. These facts will be used to arrive on lines (A.143) and (A.145). From (A.118), we know f_{18} is injective on $\overline{C}^*$, which will also be used on line (A.145). Lemma 2.1 will be used to arrive on (A.144).

$$\begin{aligned}
&\operatorname{codim}_B \overline{B}^* \\
&\quad = H(B) - \dim(\overline{B}^*) \\
&\quad = H(B) - \dim(f_8(f_{21}\overline{B} \cap f_{18}\overline{C}^*)) \\
&\quad = H(B) - \dim(f_{21}\overline{B} \cap f_{18}\overline{C}^*) && \text{(A.143)} \\
&\quad = H(B) - H(Y) + \operatorname{codim}_Y(f_{21}\overline{B} \cap f_{18}\overline{C}^*) \\
&\quad \le H(B) - H(Y) + \operatorname{codim}_Y(f_{21}\overline{B}) + \operatorname{codim}_Y(f_{18}\overline{C}^*) && \text{(A.144)} \\
&\quad = H(B) - H(Y) + H(Y) - \dim(f_{21}\overline{B}) \\
&\qquad + H(Y) - \dim(f_{18}\overline{C}^*) \\
&\quad = H(B) + H(Y) - \dim(\overline{B}) - \dim(\overline{C}^*) && \text{(A.145)} \\
&\quad \le H(B) + H(Y) - H(B) + \Delta_{\overline{B}} - H(C) + \Delta_{\overline{C}^*} \\
&\quad = H(Y) - H(C) + \Delta_{\overline{B}} + \Delta_{\overline{C}^*} \\
&\quad = H(Y) - H(C) + H(W) - H(D) + \Delta_{\overline{C}} + \Delta_{\overline{D}} + \Delta_{\overline{B}} \\
&\quad \triangleq \Delta_{\overline{B}^*}.
\end{aligned}$$

From (A.109), we know f_{12} is injective on $f_{29}\widehat{C}$ and f_{29} is injective on $\widehat{C}$. These facts will be used to arrive on lines (A.146) and (A.148). From (A.100), we know f_{27} is injective on $\widehat{B}$, which will also be used on line (A.148). Lemma 2.1 will be used to arrive on (A.147).

$$\begin{aligned}
&\operatorname{codim}_C \widehat{C}^* \\
&\quad = H(C) - \dim(\widehat{C}^*) \\
&\quad = H(C) - \dim(f_{12}(f_{27}\widehat{B} \cap f_{29}\widehat{C})) \\
&\quad = H(C) - \dim(f_{27}\widehat{B} \cap f_{29}\widehat{C}) && \text{(A.146)} \\
&\quad = H(C) - H(Z) + \operatorname{codim}_Z(f_{27}\widehat{B} \cap f_{29}\widehat{C}) \\
&\quad \le H(C) - H(Z) + \operatorname{codim}_Z(f_{27}\widehat{B}) + \operatorname{codim}_Z(f_{29}\widehat{C}) && \text{(A.147)} \\
&\quad = H(C) - H(Z) + H(Z) - \dim(f_{27}\widehat{B}) \\
&\qquad + H(Z) - \dim(f_{29}\widehat{C}) \\
&\quad = H(C) + H(Z) - \dim(\widehat{B}) - \dim(\widehat{C}) && \text{(A.148)} \\
&\quad \le H(C) + H(Z) - H(B) + \Delta_{\widehat{B}} - H(C) + \Delta_{\widehat{C}} \\
&\quad = H(Z) - H(B) + \Delta_{\widehat{B}} + \Delta_{\widehat{C}} \\
&\quad \triangleq \Delta_{\widehat{C}^*}.
\end{aligned}$$

Let $t \in A$. Now, we will assume t satisfies conditions (A.149)–(A.154). The justifications can be found below.

$t \in \widehat{A}$:

this is true on a subspace of A of codimension at most $\Delta_{\widehat{A}}$ (A.149)

$f_{32}t \in f_{20}\overline{B}^* \cap f_{25}\widehat{B}^*$:

this is true on a subspace of A of codimension at most

$$2H(X) - 2H(B) + \Delta_{\overline{B}^*} + \Delta_{\widehat{B}^*} \qquad \text{(A.150)}$$

$f_{33}t \in f_{28}\widehat{C}^* \cap f_{21}\overline{B}^*$:
this is true on a subspace of A of codimension at most
$$2H(Y) - H(B) - H(C) + \Delta_{\overline{B}^*} + \Delta_{\widehat{C}^*} \tag{A.151}$$
$f_{34}t \in f_{29}\widehat{C}^* \cap f_{27}\widehat{B}^*$:
this is true on a subspace of A of codimension at most
$$2H(Z) - H(C) - H(B) + \Delta_{\widehat{C}^*} + \Delta_{\widehat{B}^*} \tag{A.152}$$
$f_{18}f_2f_{31}t \in f_{21}\overline{B}^* \cap f_{28}\widehat{C}^*$:
this is true on a subspace of A of codimension at most
$$2H(Y) - H(B) - H(C) + \Delta_{\overline{B}^*} + \Delta_{\widehat{C}^*}. \tag{A.153}$$

Now, we need to make two assumptions on t simultaneously:

$f_{31}t \in f_{17}\overline{C} \cap f_{14}\overline{A}^*$ and $f_{15}g_{14}f_{31}t \in f_{20}\overline{B}^* \cap f_{25}\widehat{B}^*$;
this is true on a subspace of A of codimension at most
$$2H(X) - 2H(B) + 2H(W) - H(C) - H(A) + \Delta_{\overline{C}} + \Delta_{\overline{A}^*} + \Delta_{\overline{B}^*} + \Delta_{\widehat{B}^*}. \tag{A.154}$$

To justify (A.150), first we know f_{20} is injective on $\overline{B}^*$ by (A.120). Then by Lemma 2.3, we know $f_{32}t \in f_{20}\overline{B}^*$ on a subspace of A of codimension at most

$$H(X) - H(B) + \operatorname{codim}_B(\overline{B}^*) \le H(X) - H(B) + \Delta_{\overline{B}^*}.$$

By (A.115), we also know f_{25} is injective on $\widehat{B}^*$. Then by Lemma 2.3, we know $f_{32}t \in f_{25}\widehat{B}^*$ on a subspace of A of codimension at most

$$H(X) - H(B) + \operatorname{codim}_B(\widehat{B}^*) \le H(X) - H(B) + \Delta_{\widehat{B}^*}.$$

Then using Lemma 2.1, we have $f_{32}t \in f_{20}\overline{B}^* \cap f_{25}\widehat{B}^*$ on a subspace of A of codimension at most

$$2H(X) - 2H(B)\Delta_{\overline{B}^*} + \Delta_{\widehat{B}^*}.$$

Conditions (A.151)–(A.153) can be justified similarly.

To justify (A.154), first we know f_{17} is injective on $\overline{C}$ by (A.105). Then by Lemma 2.3, we know $f_{31}t \in f_{17}\overline{C}$ on a subspace of A of codimension at most

$$H(W) - H(C) + \operatorname{codim}_C(\overline{C}) \le H(W) - H(C) + \Delta_{\overline{C}}.$$

By (A.117), we also know f_{14} is injective on $\overline{A}^*$. Then by Lemma 2.3, we know $f_{31}t \in f_{14}\overline{A}^*$ on a subspace of A of codimension at most

$$H(W) - H(A) + \operatorname{codim}_A(\overline{A}^*) \le H(W) - H(A) + \Delta_{\overline{A}^*}.$$

Then using Lemma 2.1, we have

$$f_{31}t \in f_{17}\overline{C} \cap f_{14}\overline{A}^*$$

on a subspace, S, of A of codimension at most

$$2H(W) - H(C) - H(A) + \Delta_{\overline{C}} + \Delta_{\overline{A}^*}.$$

Since f_{14} is injective on $\overline{A}^*$, the function $f_{15}g_{14}f_{31}$ is defined on S. Using the same technique as before we can show that

$$f_{15}g_{14}f_{31}t \in f_{20}\overline{B}^* \cap f_{25}\widehat{B}^*$$

on a subspace, $\overline{S}$, of codimension with respect to S at most

$$2H(X) - 2H(B) + \Delta_{\overline{B}^*} + \Delta_{\widehat{B}^*}.$$

Thus both conditions are true on $\overline{S}$, which has codimension with respect to A at most

$$\operatorname{codim}_S\overline{S} + \operatorname{codim}_A S \le 2H(X) - 2H(B) + 2H(W) - H(C) - H(A) + \Delta_{\overline{C}} + \Delta_{\overline{A}^*} + \Delta_{\overline{B}^*} + \Delta_{\widehat{B}^*}.$$

Our final goal is to show that $t = 3x$ for some x so that we may conclude that $t = 0$ if the characteristic is 3. We will accomplish this by using (A.87) and by proving that $f_4f_{32}t = f_7f_{33}t = f_{10}f_{34}t$.

Claim 1: $f_4f_{32}t = f_{10}f_{34}t$

Proof: First we must show that $f_{28}f_{12}f_{34}t = f_{21}g_{20}f_{32}t$. By (A.88), we know

$$f_8f_{33}t = -f_{11}f_{34}t - f_1f_{31}t.$$

Then by using (A.130) and condition (A.152), we have

$$f_8f_{33}t = f_8f_{28}f_{12}f_{34}t - f_1f_{31}t.$$

Now, by using (A.122) and condition (A.154), we have

$$f_8f_{33}t = f_8f_{28}f_{12}f_{34}t + f_8f_{18}f_2f_{31}t.$$

By (A.116), we know f_8 is injective on $f_{28}\widehat{C}^*$. By condition (A.151), we know

$$f_{33}t \in f_{28}\widehat{C}^*.$$

By condition (A.153), we know

$$f_{18}f_2f_{31}t \in f_{28}\widehat{C}^*.$$

By condition (A.152), we know

$$f_{34}t \in f_{29}\widehat{C}^*.$$

Using (A.109), we know

$$f_{12}f_{34}t \in \widehat{C}^*.$$

Thus, we have

$$f_{33}t = f_{28}f_{12}f_{34}t + f_{18}f_2f_{31}t. \tag{A.155}$$

By (A.90), we have

$$f_9f_{33}t = -f_6f_{32}t - f_3f_{31}t.$$

Then by using (A.126) and condition (A.150), we have

$$f_9f_{33}t = f_9f_{21}g_{20}f_{32}t - f_3f_{31}t.$$

Now, by using (A.124) and condition (A.154), we have

$$f_9f_{33}t = f_9f_{21}g_{20}f_{32}t + f_9f_{18}f_2f_{31}t.$$

By (A.119), we know f_9 is injective on $f_{21}\overline{B}^*$. By condition (A.151), we know

$$f_{33}t \in f_{21}\overline{B}^*.$$

By condition (A.150), we know

$$f_{32}t \in f_{20}\overline{B}^*$$

so

$$f_{21}g_{20}f_{32}t \in f_{21}\overline{B}^*.$$

By condition (A.153), we know

$$f_{18}f_2f_{31}t \in f_{21}\overline{B}^*.$$

Thus, we have

$$f_{33}t = f_{21}g_{20}f_{32}t + f_{18}f_2f_{31}t. \tag{A.156}$$

Now, setting (A.155) and (A.156) equal to each other, we have

$$f_{21}g_{20}f_{32}t = f_{28}f_{12}f_{34}t. \tag{A.157}$$

By (A.125) and condition (A.150), we know

$$f_4f_{32}t = -f_7f_{21}g_{20}f_{32}t.$$

Using (A.157), we have

$$f_4f_{32}t = -f_7f_{28}f_{12}f_{34}t.$$

Then using (A.129) and condition (A.152), we know

$$f_4f_{32}t = f_{10}f_{34}t. \qquad \blacksquare$$

Claim 2: $f_7f_{33}t = f_{10}f_{34}t$.

Proof: First we must show that $f_{25}f_{11}f_{34}t = f_{20}f_8f_{33}t$. By (A.89), we know

$$f_5f_{32}t = -f_{12}f_{34}t - f_2f_{31}t.$$

Then by using (A.131) and condition (A.152), we have

$$f_5f_{32}t = f_5f_{25}f_{11}f_{34}t - f_2f_{31}t.$$

Now, by using (A.123) and condition (A.154), we have

$$f_5f_{32}t = f_5f_{25}f_{11}f_{34}t + f_5f_{15}g_{14}f_{31}t.$$

By (A.115), we know f_5 is injective on $f_{25}\widehat{B}^*$. By condition (A.150), we know

$$f_{32}t \in f_{25}\widehat{B}^*.$$

By condition (A.152), we know

$$f_{34}t \in f_{27}\widehat{B}^*.$$

Now, using (A.100), we know

$$f_{11}f_{34}t \in \widehat{B}^*.$$

By condition (A.154), we know

$$f_{15}g_{14}f_{31}t \in f_{25}\widehat{B}^*.$$

Thus, we have

$$f_{32}t = f_{25}f_{11}f_{34}t + f_{15}g_{14}f_{31}t. \tag{A.158}$$

By (A.90), we have

$$f_6f_{32}t = -f_9f_{33}t - f_3f_{31}t.$$

Then using (A.128) and condition (A.151), we have

$$f_6f_{32}t = f_6f_{20}f_8f_{33}t - f_3f_{31}t.$$

Now, by using (A.124) and condition (A.154), we have

$$f_6f_{32}t = f_6f_{20}f_8f_{33}t + f_6f_{15}g_{14}f_{31}t.$$

By (A.120), we know that f_6 is injective on $f_{20}\overline{B}^*$. By condition (A.150), we know

$$f_{32}t \in f_{20}\overline{B}^*.$$

By condition (A.151), we know

$$f_{33}t \in f_{21}\overline{B}^*.$$

Now, using (A.96), we know

$$f_8f_{33}t \in \overline{B}^*.$$

By condition (A.154), we know

$$f_{15}g_{14}f_{31}t \in f_{20}\overline{B}^*.$$

Thus, we have

$$f_{32}t = f_{20}f_8f_{33}t + f_{15}g_{14}f_{31}t. \tag{A.159}$$

Now, setting (A.158) and (A.159) equal to each other, we have

$$f_{25}f_{11}f_{34}t = f_{20}f_8f_{33}t. \tag{A.160}$$

By (A.127) and condition (A.151), we know

$$f_7f_{33}t = -f_4f_{20}f_8f_{33}t.$$

Using (A.160), we have

$$f_7f_{33}t = -f_4f_{25}f_{11}f_{34}t.$$

Then using (A.129) and condition (A.152), we know

$$f_7f_{33}t = f_{10}f_{34}t. \qquad \blacksquare$$

Now, by (A.87) and the two claims, we have

$$\begin{aligned} t &= f_4f_{32}t + f_7f_{33}t + f_{10}f_{34}t \\ &= f_{10}f_{34}t + f_{10}f_{34}t + f_{10}f_{34}t \\ &= 3f_{10}f_{34}t. \end{aligned}$$

Thus if the field has characteristic 3, then

$$t = 0. \tag{A.161}$$

No nonzero t can satisfy all of the conditions (A.149)–(A.154), so we must have

$$\begin{aligned} &H(A) \\ &\le \Delta_{\widehat{A}} + 2H(W) - H(C) - H(A) + \Delta_{\overline{C}} + \Delta_{\overline{A}^*} \\ &\quad + 2H(X) - 2H(B) + \Delta_{\overline{B}^*} + \Delta_{\widehat{B}^*} \\ &\quad + 2H(Y) - H(B) - H(C) + \Delta_{\overline{B}^*} + \Delta_{\widehat{C}^*} \\ &\quad + 2H(Z) - H(C) - H(B) + \Delta_{\widehat{C}^*} + \Delta_{\widehat{B}^*} \\ &\quad + 2H(Y) - H(B) - H(C) + \Delta_{\overline{B}^*} + \Delta_{\widehat{C}^*} \\ &\quad + 2H(X) - 2H(B) + \Delta_{\overline{B}^*} + \Delta_{\widehat{B}^*} \\ &= 2H(Z) + 4H(Y) + 4H(X) + 2H(W) - 4H(C) \\ &\quad - 7H(B) - H(A) \\ &\quad + \Delta_{\overline{A}^*} + 4\Delta_{\overline{B}^*} + 3\Delta_{\widehat{B}^*} + 3\Delta_{\widehat{C}^*} + \Delta_{\widehat{A}} + \Delta_{\overline{C}} \\ &= 2H(Z) + 4H(Y) + 4H(X) + 2H(W) - 4H(C) \\ &\quad - 7H(B) - H(A) \\ &\quad + H(X) - H(B) + H(Z) - H(C) + \Delta_{\widehat{B}} + \Delta_{\widehat{C}} + \Delta_{\overline{A}} \\ &\quad + 4(H(Y) - H(C) + H(W) - H(D) + \Delta_{\overline{C}} + \Delta_{\overline{D}} + \Delta_{\overline{B}}) \end{aligned}$$

$$
\begin{aligned}
&+3(H(Z)-H(C)+\Delta_{\widehat{B}}+\Delta_{\widehat{C}})\\
&+3(H(Z)-H(B)+\Delta_{\widehat{B}}+\Delta_{\widehat{C}})\\
&+\Delta_{\widehat{A}}+\Delta_{\overline{C}}\\
&=9H(Z)+8H(Y)+5H(X)+6H(W)-4H(D)\\
&-12H(C)-11H(B)-H(A)\\
&+\Delta_{\widehat{A}}+\Delta_{\overline{A}}+7\Delta_{\widehat{B}}+4\Delta_{\overline{B}}+7\Delta_{\widehat{C}}+5\Delta_{\overline{C}}+4\Delta_{\overline{D}}\\
&=9H(Z)+8H(Y)+5H(X)+6H(W)-4H(D)\\
&-12H(C)-11H(B)-H(A)\\
&+H(W|B,C,D)+H(X|A,C,D)+H(Y|A,B,D)\\
&+H(Z|A,B,C)+H(A|W,X,Y,Z)+H(W|B,C,D)\\
&+H(X|A,C,D)+H(A|B,W,X)\\
&+7(H(X|A,C,D)+H(Z|A,B,C)+H(B|D,X,Z))\\
&+4(H(X|A,C,D)+H(Y|A,B,D)+H(B|C,X,Y))\\
&+7(H(Y|A,B,D)+H(Z|A,B,C)+H(C|D,Y,Z))\\
&+5(H(W|B,C,D)+H(Y|A,B,D)+H(C|A,W,Y))\\
&+4(H(W|B,C,D)+H(Z|A,B,C)+H(D|A,W,Z))\\
&+29(H(A)+H(B)+H(C)+H(D)-H(A,B,C,D))\\
&=9H(Z)+8H(Y)+5H(X)+6H(W)-4H(D)\\
&-12H(C)-11H(B)-H(A)\\
&+19H(Z|A,B,C)+17H(Y|A,B,D)\\
&+13H(X|A,C,D)+11H(W|B,C,D)\\
&+H(A|W,X,Y,Z)+H(A|B,W,X)+7H(B|D,X,Z)\\
&+4H(B|C,X,Y)\\
&+7H(C|D,Y,Z)+5H(C|A,W,Y)+4H(D|A,W,Z)\\
&+29(H(A)+H(B)+H(C)+H(D)-H(A,B,C,D)).
\end{aligned}
$$

■

REFERENCES

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

[2] A. Blasiak, R. Kleinberg, and E. Lubetzky. (2011). "Lexicographic products and the power of non-linear network coding." [Online]. Available: http://arxiv.org/abs/1108.2489

[3] J. Cannons, R. Dougherty, C. Freiling, and K. Zeger, "Network routing capacity," *IEEE Trans. Inf. Theory*, vol. 52, no. 3, pp. 777–788, Mar. 2006.

[4] T. Chan and A. Grant, "Entropy vectors and network codes," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2007, pp. 1586–1590.

[5] R. Dougherty, C. Freiling, and K. Zeger, "Six new non-Shannon information inequalities," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2006, pp. 233–236.

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

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

[8] R. Dougherty, C. Freiling, and K. Zeger. (2012). "Linear rank inequalities on five or more variables." [Online]. Available: http://arxiv.org/abs/0910.0284

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

[10] R. Dougherty, C. Freiling, and K. Zeger. (2013). "Achievable rate regions for network coding." [Online]. Available: http://arxiv.org/abs/1311.4601

[11] D. Hammer, A. Romashchenko, A. Shen, and N. Vereshchagin, "Inequalities for Shannon entropy and Kolmogorov complexity," *J. Comput. Syst. Sci.*, vol. 60, no. 2, pp. 442–464, Apr. 2000.

[12] A. W. Ingleton, "Representation of matroids," in *Proc. Conf. Combinat. Math. Appl.*, 1971, pp. 149–167.

[13] R. Lněnička, "On the tightness of the Zhang–Yeung inequality for Gaussian vectors," *Commun. Inf. Syst.*, vol. 3, no. 1, pp. 41–46, 2003.

[14] K. Makarychev, Y. Makarychev, A. Romashchenko, and N. Vereshchagin, "A new class of non-Shannon-type inequalities for entropies," *Commun. Inf. Syst.*, vol. 2, no. 2, pp. 147–166, 2002.

[15] F. Matúš, "Infinitely many information inequalities," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2007, pp. 41–44.

[16] C. K. Ngai and R. W. Yeung, "Network coding gain of combination networks," in *Proc. IEEE Inf. Theory Workshop*, Oct. 2004, pp. 283–287.

[17] J. G. Oxley, *Matroid Theory*. New York, NY, USA: Oxford Univ. Press, 1992.

[18] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, Jul. 1948.

[19] W. Xu, J. Wang, and J. Sun, "A projection method for derivation of non-Shannon-type information inequalities," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2008, pp. 2116–2120.

[20] R. W. Yeung, *Information Theory and Network Coding*. New York, NY, USA: Springer-Verlag, 2008.

[21] R. W. Yeung, *A First Course in Information Theory*. Norwell, MA, USA: Kluwer, 2002.

[22] Z. Zhang, "On a new non-Shannon type information inequality," *Commun. Inf. Syst.*, vol. 3, no. 1, pp. 47–60, Jun. 2003.

[23] Z. Zhang and R. W. Yeung, "A non-Shannon-type conditional inequality of information quantities," *IEEE Trans. Inf. Theory*, vol. 43, no. 6, pp. 1982–1986, Nov. 1997.

[24] Z. Zhang and R. Yeung, "On characterization of entropy function via information inequalities," *IEEE Trans. Inf. Theory*, vol. 44, no. 4, pp. 1440–1452, Jul. 1998.

Randall Dougherty, photograph and biography not available at the time of publication.

Eric Freiling, photograph and biography not available at the time of publication.

Kenneth Zeger, photograph and biography not available at the time of publication.