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. Specifically, the first inequality holds for all finite fields whose characteristic is not three and does not in general hold over characteristic three. The second inequality holds for all finite fields whose characteristic is three and does not in general hold over characteristics other than three. Applications of these inequalities to the computation of capacity upper bounds in network coding are demonstrated.

I. INTRODUCTION

The study of information inequalities is a subfield of information theory that describes linear constraints on the entropies of finite collections of jointly distributed discrete random variables. Historically, the known information inequalities were originally 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.

Analogously, 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 contrast, 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 III.1), but does not always hold when the field characteristic is three (Theorem III.2). In contrast, the second inequality is shown to hold over all vector spaces when the field characteristic is three (Theorem IV.1), but does not always hold when the field characteristic is not three (Theorem IV.2). We also show how these inequalities can be used to obtain bounds on the capacities of certain networks (Corollaries III.4 and IV.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 Matuš [15]. Additionally, in 2007, Matuš 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.

¹In what follows, by “network” we shall always mean a directed acyclic network.
It has been shown that linear codes are insufficient for network coding in general [7]. However, linear codes may be desirable to use in practice due to ease of analysis and implementation. It has been shown that the coding capacity is independent of the alphabet size [3]. However, the linear coding capacity is dependent on alphabet size, or more specifically the field characteristic. In other words, one can potentially achieve a higher rate of linear communication by choosing one characteristic over another. To provide upper bounds for the linear coding capacity for a particular field one can look at linear rank inequalities [10]. Linear rank inequalities are linear inequalities that are always satisfied by ranks 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 characteristic. In other words, one can potentially achieve an upper bound for the linear coding capacity of two networks.

In the present paper, we give two characteristic-dependent linear rank inequalities in the derivation of the linear rank inequalities presented 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].

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.

**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:

1. $\emptyset \in I$.
2. $\forall A, B \subseteq E$, if $A \subseteq B \in I$, then $A \in I$.
3. $\forall A, B \subseteq E$, if $A, B \in I$ and $|A| > |B|$, then $\exists u \in A \setminus B$ such that $B \cup \{u\} \in I$.

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

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

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

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

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
$$H(A|B) = H(A, B) - H(B), \quad (1)$$
the mutual information between $A$ and $B$ is
$$I(A; B) = H(A) - H(A|B) = H(A) + H(B) - H(A, B), \quad (2)$$
and the conditional mutual information between $A$ and $B$ given $C$ is
$$I(A; B|C) = H(A|C) - H(A|B, C) = H(A, C) + H(B, C) - H(A, B, C) \quad (3)$$
$$= H(A) + H(B, C) - H(C) - H(A, B, C). \quad (4)$$

We will make use of the following basic information-theoretic facts [21]:
$$0 = H(\emptyset) \quad (5)$$
$$0 \leq H(A) = H(A|\emptyset) \quad (6)$$
$$0 \leq H(A|B) \quad (7)$$
$$0 \leq I(A; B) \quad (8)$$
$$H(A, B|C) \leq H(A|C) + H(B|C) \quad (9)$$
$$H(A|B, C) \leq H(A|B) \leq H(A, C|B) \quad (10)$$
$$I(A; B|C) = I(A; B|C) + I(A; C). \quad (11)$$

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 L2.** 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
$$\alpha_1 H(\{A_i : i \in S_1\}) + \cdots + \alpha_k H(\{A_i : i \in S_k\}) \geq 0 \quad (12)$$
is called an information inequality if it holds for all jointly distributed random variables $A_1, \ldots, A_q$.

As an example, taking $q = 2$, $S_1 = \{1\}$, $S_2 = \{2\}$, $S_3 = \emptyset$, $S_4 = \{1, 2\}$, $\alpha_1 = \alpha_2 = 1$, $\alpha_4 = -1$, 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 $\alpha_i$ 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 $(A, B)$ 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 $(A, B)$. 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]:
$$I(A; B) \leq I(A; B|C) + I(A; B|D) + I(C; D). \quad (13)$$

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

**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, 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:

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

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

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

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

\[
2k = H(x) + H(y) \tag{14}
\]

\[
= H(x, y) \quad \text{[from indep. of } x \text{ and } y]\]

\[
\leq H(x, y, z) \quad \text{[from (10)]}
\]

\[
= H(x, z) + H(y|x, z) \quad \text{[from (1)]}
\]

\[
= H(x, z) \quad \text{[from (13)]}
\]

\[
\leq H(x) + H(z) \quad \text{[from (9)]}
\]

\[
\leq k + n. \tag{15}
\]

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 III.4 and IV.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 II.1. Let \( V \) be a finite dimensional vector space with subspaces \( A \) and \( B \). Then the subspace \( A \cap B \) has codimension at most \( \text{codim}(A) + \text{codim}(B) \) in \( V \).

Lemma II.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 \( f(A) \subseteq B \). Then the codimension of \( A \) in \( A \) is at most the codimension of \( \overline{B} \) in \( B \).

Lemma II.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 II.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 \( f_1 + \cdots + f_k = 1 \) on a subspace of \( B \) of codimension \( H(B|A_1, \ldots, A_k) \).
Lemma II.5. Let \( V \) be a finite-dimensional vector space and let \( A, B, \) and \( C \) be subspaces of \( V \). Let \( f : A \rightarrow B \) and \( g : A \rightarrow C \) 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 II.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 \( f_i : A \rightarrow B_i \) 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 \( H(B_1) + \cdots + H(B_k) - H(A) \).

Lemma II.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 \( f_R : A \rightarrow C, f_L : C \rightarrow A, g_R : B \rightarrow D, \) and \( g_L : D \rightarrow E \). If \( f_L f_R = 1 \) 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 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:

\[
\begin{pmatrix}
A & B & C & D & W & X & Y & Z \\
1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\
0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \\
0 & 0 & 0 & 1 & 1 & 1 & 1 & 0
\end{pmatrix}
\]

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

\[
H(A) \leq 8H(Z) + 29H(Y) + 3H(X) + 8H(W) - 6H(D) - 17H(C) - 8H(B) - 17H(A) + 55H(Z|A, B, C) + 35H(Y|W, X, Z) + 50H(X|A, C, D) + 49H(W|B, C, D) + 18H(A|B, D, Y) + 7H(B|D, X, Z) + H(B|A, W, X) + 7H(C|D, Y, Z) + 7H(C|B, X, Y) + 3H(C|A, W, Y) + 6H(D|A, W, Z) + 49(H(A) + H(B) + H(C) + H(D) - H(A, B, C, D)).
\]

Proof. See the Appendix.

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

Theorem III.2. There exists a vector space \( V \) with a finite scalar field of characteristic 3 such that the T8 inequality in Theorem III.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 $V$:

$$A = \langle (1, 0, 0, 0) \rangle$$

$$B = \langle (0, 1, 0, 0) \rangle$$

$$C = \langle (0, 0, 1, 0) \rangle$$

$$D = \langle (0, 0, 0, 1) \rangle$$

$$W = \langle (0, 1, 1, 1) \rangle$$

$$X = \langle (1, 0, 1, 1) \rangle$$

$$Y = \langle (1, 1, 0, 1) \rangle$$

$$Z = \langle (1, 1, 1, 0) \rangle.$$ 

We have:

$$0 = H(Z|A, B, C)$$

$$= H(W|B, C, D)$$

$$= H(X|A, C, D)$$

$$= H(Y|W, X, Z)$$

$$= H(Y|W, X, Z) = 2^{-1}(0, 1, 1, 1) + (0, 1, 1, 1) + (1, 1, 1, 0) + (0, 1, 1, 1) + (1, 1, 1, 0)$$

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

$$= H(A)|B, D, Y)$$

$$= H(D|A, W, Z)$$

$$= H(C|D, Y, Z)$$

$$= H(C|D, Y, Z)$$

$$= H(C|D, Y, Z)$$

$$= H(C|D, Y, Z)$$

$$= H(C|D, Y, Z)$$

So, if the inequality in Theorem III.1 were to hold over $V$, then we would have

$$1 = H(A)$$

$$\leq 8H(Z) + 39H(Y) + 8H(X) + 8H(W) - 6H(D) - 17H(C) - 8H(B) - 17H(A)$$

$$= 8 + 29 + 3 + 8 - 6 - 17 - 8 - 17$$

$$= 0$$

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 III.4.

**Lemma III.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 III.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 III.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 III.2) by Lemma III.3:

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

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 III.1 reduces to

$$
H(A) \leq 8H(Z) + 29H(Y) + 3H(X) + 8H(W) - 6H(D)
- 17H(C) - 8H(B) - 17H(A).
$$

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

$$
k \leq 8n + 29n + 3n + 8n - 6k - 17k - 8k - 17k
$$

$$
k/n \leq 48/49.
$$

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):

$$
Z = A + B + C
$$
$$
W = B + C + D
$$
$$
X = A + C + D
$$
$$
Y = W + X + Z.
$$

and decoding functions:

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

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 IV.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 IV.1. Let $A, B, C, D, W, X, Y, Z$ 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 $V$:

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

Proof. See the Appendix.

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

Theorem IV.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 IV.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 $V$:

$$A = \langle (1, 0, 0, 0) \rangle$$
$$B = \langle (0, 1, 0, 0) \rangle$$
$$C = \langle (0, 0, 1, 0) \rangle$$
$$D = \langle (0, 0, 0, 1) \rangle$$
$$W = \langle (0, 1, 1, 1) \rangle$$
$$X = \langle (1, 0, 1, 1) \rangle$$
$$Y = \langle (1, 1, 0, 1) \rangle$$
$$Z = \langle (1, 1, 1, 1) \rangle$$

We have:

$$0 = H(W|B, C, D)$$
[from $(0, 1, 1, 1) = (0, 1, 0, 0) + (0, 0, 1, 0) + (0, 0, 0, 1)$]

\[ = H(X|A, C, D) \]
[from $(1, 0, 1, 1) = (1, 0, 0, 0) + (0, 0, 1, 0) + (0, 0, 0, 1)$]

\[ = H(Y|A, B, D) \]
[from $(1, 1, 0, 1) = (1, 0, 0, 0) + (0, 1, 0, 0) + (0, 0, 0, 1)$]

\[ = H(Z|A, B, C) \]
[from $(1, 1, 1, 0) = (1, 0, 0, 0) + (0, 1, 0, 0) + (0, 0, 1, 0)$]

\[ = H(A|B, W, X) \]
[from $(1, 0, 0, 0) = (1, 0, 1, 1) + (0, 1, 0, 0) - (0, 1, 1, 1)$]

\[ = H(C|A, W, Y) \]
[from $(0, 0, 1, 0) = (1, 0, 1, 1) + (1, 0, 0, 0) - (1, 1, 0, 1)$]

\[ = H(B|C, X, Y) \]
[from $(0, 1, 0, 0) = (1, 1, 0, 1) + (0, 0, 0, 1) - (1, 0, 1, 1)$]

\[ = H(D|A, W, Z) \]
[from $(0, 0, 0, 1) = (0, 1, 1, 1) + (1, 0, 0, 0) - (1, 1, 0, 0)$]

\[ = H(B|D, X, Z) \]
[from $(0, 1, 0, 0) = (1, 1, 0, 1) + (0, 0, 0, 1) - (1, 0, 1, 1)$]

\[ = H(C|D, Y, Z) \]
[from $(0, 0, 1, 0) = (1, 1, 1, 1) + (0, 0, 0, 1) - (1, 1, 0, 1)$]

\[ = H(A|W, X, Y, Z) \]
[from $(1, 0, 0, 0) = 3^{-1} \left( (1, 0, 1, 1)+(1, 1, 0, 1)+(1, 1, 1, 0) -2(0,1,1,1) \right). \] (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 IV.1 were to hold over $V$, then we would have

\[ 1 < H(A) \]
\[ \leq 9H(Z) + 8H(Y) + 5H(X) + 6H(W) - 4H(D) \]
\[ - 12H(C) - 11H(B) - H(A) \]
\[ = 9 + 8 + 5 + 6 - 4 - 12 - 11 - 1 \]
\[ = 0 \]

which is impossible. $\blacksquare$

**Corollary IV.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 with random variables $A, B, C, D, W, X, Y, Z$ are taken as random variables (i.e., not as subspaces as in Theorem IV.2) by Lemma III.3:

\begin{align*}
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{align*}

and since the source message $A, B, C, D$ 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 IV.1 reduces to

\[ H(A) \leq 9H(Z) + 8H(Y) + 5H(X) + 6H(W) - 4H(D) - 12H(C) - 11H(B) - H(A). \]

Now, since

\[ H(A) = H(B) = H(C) = H(D) = k \]

and

\[ H(W) = H(X) = H(Y) = H(Z) \leq n \]

we have

\[ k \leq 9n + 8n + 5n + 6n - 4k - 12k - 11k - k \]

\[ k/n \leq 28/29. \]

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):

\[ W = B + C + D \]
\[ X = A + C + D \]
\[ Y = A + B + D \]
\[ Z = A + B + C \]

and decoding functions:

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

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. ■

\section*{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.

\section*{APPENDIX}

\subsection*{Proof of Lemma II.1}

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

\[ H(V) - I(A; B) \leq H(V) - H(A) + H(V) - H(B). \]

Thus, $\text{codim}(A \cap B) \leq \text{codim}(A) + \text{codim}(B)$. ■

\subsection*{Proof of Lemma II.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

\[ \text{codim}_A(\overline{A}) = n \leq \text{codim}_B(\overline{B}). \]

\[ \text{codim}_A(\overline{A}) = n \leq \text{codim}_B(\overline{B}). \]

\subsection*{Proof of Lemma II.3}

Let $\overline{A} = \{ t \in A : f(t) \in \overline{B} \}$. Then $f(A \setminus \overline{A}) \subseteq B \setminus \overline{B}$ and the result follows from Lemma II.2. ■
Proof of Lemma II.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

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

Proof of Lemma II.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

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

Proof of Lemma II.6.
First we apply Lemma II.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 II.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 II.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 II.7. Let $x, y \in f_L(f_R A \cap g_R B)$. We know $f_R f_L = f$ 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

$$f_R(x) = f_R f_L(f_R A \cap g_R B) = f_R A \cap g_R B$$
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 III.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 II.4 we get linear functions:

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

\[ f_1 + f_2 + f_3 = I \]

on a subspace of \( Z \) of codimension \( H(Z|A, B, C) \)  \hspace{1cm} (A.1)

\[ f_4 + f_5 + f_6 = I \]

on a subspace of \( W \) of codimension \( H(W|B, C, D) \)  \hspace{1cm} (A.2)

\[ f_7 + f_8 + f_9 = I \]

on a subspace of \( X \) of codimension \( H(X|A, C, D) \)  \hspace{1cm} (A.3)

\[ f_{10} + f_{11} + f_{12} = I \]

on a subspace of \( Y \) of codimension \( H(Y|W, X, Z) \)  \hspace{1cm} (A.4)

\[ f_{13} + f_{14} + f_{15} = I \]

on a subspace of \( A \) of codimension \( H(A|B, D, Y) \)  \hspace{1cm} (A.5)

\[ f_{16} + f_{17} + f_{18} = I \]

on a subspace of \( D \) of codimension \( H(D|A, W, Z) \)  \hspace{1cm} (A.6)

\[ f_{19} + f_{20} + f_{21} = I \]

on a subspace of \( C \) of codimension \( H(C|D, Y, Z) \)  \hspace{1cm} (A.7)

\[ f_{22} + f_{23} + f_{24} = I \]

on a subspace of \( B \) of codimension \( H(B|D, X, Z) \)  \hspace{1cm} (A.8)

\[ f_{25} + f_{26} + f_{27} = I \]

on a subspace of \( C \) of codimension \( H(C|B, X, Y) \)  \hspace{1cm} (A.9)

\[ f_{28} + f_{29} + f_{30} = I \]

on a subspace of \( C \) of codimension \( H(C|A, W, Y) \)  \hspace{1cm} (A.10)

\[ f_{31} + f_{32} + f_{33} = I \]

on a subspace of \( B \) of codimension \( H(B|A, W, X) \)  \hspace{1cm} (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 II.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 II.1, and Lemma II.3 we know the sum of these functions is equal to \( I \) on a subspace of \( A \) of codimension at most


since, on that subspace,

\[ (f_A f_{15}) + (f_B f_{15} + f_{13}) + (f_C f_{15}) + (f_D f_{15} + f_{14}) = f_7 f_{12} f_{15} + f_1 f_{10} f_{15} + f_4 f_{11} f_{15} + f_2 f_{10} f_{15} + f_{13} + f_8 f_{12} f_{15} + f_5 f_{11} f_{15} + f_3 f_{10} f_{15} + f_6 f_{12} f_{15} + f_6 f_{11} f_{15} + f_{14} = f_{13} + f_{14} + (f_1 + f_2 + f_3) f_{10} f_{15} + (f_4 + f_5 + f_6) f_{11} f_{15} + (f_7 + f_8 + f_9) f_{12} f_{15} = f_{13} + f_{14} + (f_{10} + f_{11} + f_{12}) f_{15} = f_{13} + f_{14} + f_{15} = I. \]

Applying Lemma II.6 and Lemma II.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 \( \mathcal{A} \) of \( A \) of codimension at most


on which

\[ f_A f_{15} = I \]  \hspace{1cm} (A.12)

\[ f_B f_{15} + f_{13} = 0 \]  \hspace{1cm} (A.13)

\[ f_C f_{15} = 0 \]  \hspace{1cm} (A.14)

\[ f_D f_{15} + f_{14} = 0. \]  \hspace{1cm} (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

\[ \Delta_{\overline{B}} = H(Z|A, B, C) + H(X|A, C, D) + H(B|D, X, Z) + H(A) + H(B) + H(C) + H(D) - H(A, B, C, D) \]

on which

\[ f_7 f_{23} + f_1 f_{22} = 0 \]  \hspace{1cm} (A.16)

\[ f_2 f_{22} = I \]  \hspace{1cm} (A.17)

\[ f_8 f_{23} + f_3 f_{22} = 0 \]  \hspace{1cm} (A.18)

\[ f_{24} + f_9 f_{23} = 0. \]  \hspace{1cm} (A.19)
We get a subspace $\hat{B}$ of $B$ of codimension at most
\[
\Delta_{\hat{B}} = H(W|B, C, D) + H(X|A, C, D) \\
+ H(B|A, W, X) \\
+ H(A) + H(B) + H(C) + H(D) - H(A, B, C, D)
\]
on which
\[
f_{33} + f_{7}f_{32} = 0 \quad (A.20) \\
f_{4}f_{31} = I \quad (A.21) \\
f_{6}f_{32} + f_{5}f_{31} = 0 \quad (A.22) \\
f_{9}f_{32} + f_{6}f_{31} = 0. \quad (A.23)
\]
We get a subspace $\overline{C}$ of $C$ of codimension at most
\[
\Delta_C = 2H(Z|A, B, C) + H(W|B, C, D) + H(X|A, C, D) \\
+ H(Y|W, X, Z) + H(C|B, X, Y) \\
+ H(A) + H(B) + H(C) + H(D) - H(A, B, C, D)
\]
on which
\[
f_{A}f_{20} + f_{1}f_{19} = 0 \quad (A.24) \\
f_{B}f_{20} + f_{2}f_{19} = 0 \quad (A.25) \\
f_{C}f_{20} + f_{3}f_{19} = I \quad (A.26) \\
f_{D}f_{20} + f_{1}f_{21} = 0. \quad (A.27)
\]
We get a subspace $\hat{C}$ of $C$ of codimension at most
\[
\Delta_{\hat{C}} = H(Z|A, B, C) + H(W|B, C, D) + 2H(X|A, C, D) \\
+ H(Y|W, X, Z) + H(C|B, X, Y) \\
+ H(A) + H(B) + H(C) + H(D) - H(A, B, C, D)
\]
on which
\[
f_{A}f_{25} + f_{1}f_{26} = 0 \quad (A.28) \\
f_{B}f_{25} + f_{2}f_{27} = 0 \quad (A.29) \\
f_{C}f_{25} + f_{3}f_{26} = I \quad (A.30) \\
f_{D}f_{25} + f_{5}f_{26} = 0. \quad (A.31)
\]
We get a subspace $\overline{C}$ of $C$ of codimension at most
\[
\Delta_{\overline{C}} = H(Z|A, B, C) + H(W|B, C, D) + H(D|A, W, Z) \\
+ H(A) + H(B) + H(C) + H(D) - H(A, B, C, D)
\]
on which
\[
f_{A}f_{28} + f_{30} = 0 \quad (A.32) \\
f_{B}f_{28} + f_{1}f_{29} = 0 \quad (A.33) \\
f_{C}f_{28} + f_{3}f_{29} = I \quad (A.34) \\
f_{D}f_{28} + f_{6}f_{29} = 0. \quad (A.35)
\]
We get a subspace $\overline{D}$ of $D$ of codimension at most
\[
\Delta_{\overline{D}} = H(Z|A, B, C) + H(W|B, C, D) + H(D|A, W, Z) \\
+ H(A) + H(B) + H(C) + H(D) - H(A, B, C, D)
\]
on which
\[
f_{18} + f_{1}f_{16} = 0 \quad (A.36) \\
f_{4}f_{17} + f_{2}f_{16} = 0 \quad (A.37) \\
f_{5}f_{17} + f_{3}f_{16} = 0 \quad (A.38) \\
f_{6}f_{17} = I. \quad (A.39)
\]
First notice that (A.12) implies
\[
f_{15} \text{ is injective on } \overline{A}. \quad (A.40)
\]
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
\[
\overline{C}' \triangleq f_{3}(f_{19}(\overline{C} \cap f_{20}^{-1}f_{15}\overline{A}) \cap f_{22}\overline{B}) \subseteq \overline{C} \\
\overline{C}'' \triangleq f_{3}(f_{29}(\overline{C} \cap f_{28}^{-1}f_{15}\overline{A}) \cap f_{17}\overline{D}) \subseteq \overline{C} \\
\overline{A} \triangleq f_{A}(f_{15}\overline{A} \cap f_{20}\overline{C}' \cap f_{28}\overline{C}'') \subseteq \overline{A}.
\]
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
\[
f_{3}f_{19} = I \quad \text{on } \overline{C} \cap f_{20}^{-1}f_{15}\overline{A}. \quad (A.41)
\]
Using (A.14) and (A.34) we have
\[
f_{5}f_{29} = I \quad \text{on } \overline{C} \cap f_{28}^{-1}f_{15}\overline{A}. \quad (A.42)
\]
Using (A.14) and (A.30) we have
\[
f_{8}f_{26} = I \quad \text{on } \overline{C} \cap f_{25}^{-1}f_{15}\overline{A}. \quad (A.43)
\]
We are now going to show $f_{13}$ is injective on $\overline{A}$. First we need to apply Lemma II.7 to show $f_{2}f_{19}$ is injective on $\overline{C}'$ and then again to show $f_{2}f_{15}$ 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 II.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 II.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),
\[
f_{13} \text{ is injective on } \overline{A}. \quad (A.44)
\]
Similarly, we are going to show $f_{14}$ is injective on $\overline{A}$. We will first apply Lemma II.7 to show $f_{6}f_{29}$ is injective on $\overline{C}^*$ 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 $\overline{C} \cap f_{28}^{-1}f_{15}\overline{A}$. So, we can apply Lemma II.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 $\overline{C}^*$. Then using (A.35), we know $f_{D}f_{28}$ is injective on $\overline{C}^*$. Now we can apply Lemma II.7 again by using the fact that $f_{A}f_{15} = I$ on $\overline{A}$ and by letting $g_{L} = f_{D}$,
\[ g_R = f_{28}, \quad f_L = f_A, \quad \text{and} \quad f_R = f_{15} \text{ to get } f_D f_{15} \text{ is injective on } \overline{A}. \text{ Thus by (A.15),} \]

\[ f_{14} \text{ is injective on } \overline{A}. \quad (A.45) \]

Now we are going to find an upper bound for \( \text{codim}_A(\overline{A}) \). First we need to find upper bounds for \( \text{codim}_C(\overline{C}^*) \) and \( \text{codim}_C(\tilde{C}^*) \). Using (A.40) to show \( \dim(f_{15}\overline{A}) = \dim(\overline{A}) \), and again using Lemma II.1 and Lemma II.3, we have

\[ \text{codim}_C(\overline{C}^*) \]
\[ = H(C) - \dim(\overline{C}^*) \]
\[ = H(C) - \dim(f_5(f_{29}(\overline{C} \cap f_{28}^{-1}f_{15}\overline{A}) \cap f_{17}\overline{D}))) \]
\[ = H(C) - \dim(f_5(\overline{C} \cap f_{28}^{-1}f_{15}\overline{A}) \cap f_{17}\overline{D})) \]
\[ = H(C) - H(W) + \text{codim}_W(f_29(\overline{C} \cap f_{28}^{-1}f_{15}\overline{A}) \cap f_{17}\overline{D}) \]
\[ \leq H(C) - H(W) + \text{codim}_W(f_29(\overline{C} \cap f_{28}^{-1}f_{15}\overline{A})) \]
\[ + \text{codim}_W(f_{17}\overline{D}) \]
\[ = H(C) - H(W) + H(W) - \dim(f_{29}(\overline{C} \cap f_{28}^{-1}f_{15}\overline{A})) \]
\[ + H(W) - \dim(f_{17}\overline{D}) \]
\[ = H(C) + H(W) - \dim(\overline{C} \cap f_{28}^{-1}f_{15}\overline{A}) - \dim(\overline{D}) \]
\[ = H(C) + H(W) - H(C) + \text{codim}_C(\overline{C} \cap f_{28}^{-1}f_{15}\overline{A}) \]
\[ - H(D) + \text{codim}_D(\overline{D}) \]
\[ = H(W) - H(C) + \text{codim}_C(f_{15}\overline{A} \cap f_{28}^{-1}f_{15}\overline{A}) + \text{codim}_D(\overline{D}) \]
\[ \leq H(W) - H(D) + \Delta_{\overline{C}} + \text{codim}_C(f_{28}^{-1}f_{15}\overline{A}) + \Delta_{\overline{D}} \]
\[ \leq H(W) - H(D) + \Delta_{\overline{C}} + \text{codim}_V(f_{15}\overline{A}) + \Delta_{\overline{D}} \]
\[ = H(W) - H(D) + \Delta_{\overline{C}} + H(Y) - \dim(f_{15}\overline{A}) + \Delta_{\overline{D}} \]
\[ = H(W) - H(D) + \Delta_{\overline{C}} + H(Y) - \dim(\overline{A}) + \Delta_{\overline{D}} \]
\[ = H(W) - H(D) + \Delta_{\overline{C}} + H(Y) - H(A) \]
\[ + \text{codim}_A(\overline{A}) + \Delta_{\overline{D}} \]
\[ \leq H(W) - H(D) + H(Y) - H(A) + \Delta_{\overline{C}} + \Delta_{\overline{D}} + \Delta_{\overline{D}} \]

(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 \( C^* \), which implies \( f_{28} \) is injective on \( C^* \). These facts
combined with (A.40) will be used to arrive on line (A.48).

\[
\text{codim}_A(\overline{A}) = H(A) - \dim(f_A(f_{15}\overline{A} \cap f_{28}\overline{C}^* \cap f_{28}\tilde{C}^*)) = H(A) - \dim(f_{15}\overline{A} \cap f_{20}\overline{C}^* \cap f_{28}\tilde{C}^*) = H(A) - H(Y) + \text{codim}_Y(f_{15}\overline{A} \cap f_{20}\overline{C}^* \cap f_{28}\tilde{C}^*) \\
\leq H(A) - H(Y) + \text{codim}_Y(f_{15}\overline{A}) + \text{codim}_Y(f_{20}\overline{C}^*) + \text{codim}_Y(f_{28}\tilde{C}^*).
\]

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

This is true on a subspace of \( A \) of codimension at most

\[
H(X) - H(C) + H(Y) - H(A) + \Delta_{\overline{C}} + \Delta_{\overline{C}^*}.
\]

(A.53)

We will assume \( f_{10}f_{15}t \in f_{22}(\overline{B} \cap f_{20}\overline{C} \cap f_{28}(\tilde{C} \cap f_{25}^{-1}f_{15}\overline{A}^*)) \).

This is true on a subspace of \( A \) of codimension at most

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

(A.54)

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

\[
H(Z) - H(C) + H(Y) - H(A) + \Delta_{\overline{C}} + \Delta_{\overline{A}}.
\]

(A.51)

We will assume \( f_{11}f_{15}t \in f_{20}(\tilde{C} \cap f_{28}^{-1}f_{15}\overline{A}^*) \).

This is true on a subspace of \( A \) of codimension at most

\[
H(W) - H(C) - H(\tilde{C} \cap f_{28}^{-1}f_{15}\overline{A}^*) \leq \Delta_{\overline{A}}.
\]

(A.49)

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 \( \overline{C} \cap f_{28}^{-1}f_{15}\overline{A}^* \) by (A.42). Then by Lemma II.3, we know

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

on a subspace of \( A \) of codimension at most

\[
H(Z) - H(C) + \text{codim}_C(\tilde{C} \cap f_{28}^{-1}f_{15}\overline{A}^*).
\]

By Lemma II.1, we know

\[
\text{codim}_C(\overline{C} \cap f_{28}^{-1}f_{15}\overline{A}^*) \leq \Delta_{\overline{C}} + \text{codim}_C(f_{15}\overline{A}^*).
\]

Then using Lemma II.3 and (A.40), we know

\[
\text{codim}_C(\overline{C} \cap f_{28}^{-1}f_{15}\overline{A}^*) \leq \Delta_{\overline{C}} + \text{codim}_Y(f_{15}\overline{A}^*) = \Delta_{\overline{C}} + H(Y) - \dim(f_{15}\overline{A}^*) = \Delta_{\overline{C}} + H(Y) - \dim(\overline{A}^*) \leq \Delta_{\overline{C}} + H(Y) - H(A) + \Delta_{\overline{A}}.
\]

(A.56)
Then using Lemma II.3 and (A.40), we know
\[
\text{codim}_C(\tilde{C} \cap f_{25}^{-1}f_{15}\mathcal{A}) \leq \Delta_{\tilde{C}} + \text{codim}_Y(f_{15}\mathcal{A})
\]
\[
= \Delta_{\tilde{C}} + H(Y) - \dim(f_{15}\mathcal{A})
\]
\[
= \Delta_{\tilde{C}} + H(Y) - \dim(\mathcal{A}^*)
\]
\[
\leq \Delta_{\tilde{C}} + H(Y) - H(A) + \Delta_T.
\]
(A.57)

So, we have
\[
f_{12}f_{15}t \in f_{26}(\tilde{C} \cap f_{25}^{-1}f_{15}\mathcal{A})
\]
on a subspace of \(A\) of codimension at most
\[
H(Z) - H(C) + H(Y) - H(A) + \Delta_{\tilde{C}} + \Delta_T.
\]

To justify (A.53), first we know \(f_{26}\) is injective on \(\tilde{C} \cap f_{25}^{-1}f_{15}\mathcal{A}\) by (A.43). Then by Lemma II.3, we know
\[
f_{12}f_{15}t \in f_{26}(\tilde{C} \cap f_{25}^{-1}f_{15}\mathcal{A})
\]
on a subspace of \(A\) of codimension at most
\[
H(Z) - H(C) + \text{codim}_C(\tilde{C} \cap f_{25}^{-1}f_{15}\mathcal{A}).
\]

By Lemma II.1, we know
\[
\text{codim}_C(\tilde{C} \cap f_{25}^{-1}f_{15}\mathcal{A}) \leq \Delta_{\tilde{C}} + \text{codim}_C(f_{25}^{-1}f_{15}\mathcal{A})
\]
Then using Lemma II.3 and (A.40), we know
\[
\text{codim}_C(\tilde{C} \cap f_{25}^{-1}f_{15}\mathcal{A}) \leq \Delta_{\tilde{C}} + \text{codim}_Y(f_{15}\mathcal{A})
\]
\[
= \Delta_{\tilde{C}} + H(Y) - \dim(f_{15}\mathcal{A})
\]
\[
= \Delta_{\tilde{C}} + H(Y) - \dim(\mathcal{A}^*)
\]
\[
\leq \Delta_{\tilde{C}} + H(Y) - H(A) + \Delta_T.
\]
(A.58)

So, we have
\[
f_{12}f_{15}t \in f_{26}(\tilde{C} \cap f_{25}^{-1}f_{15}\mathcal{A})
\]
on a subspace of \(A\) of codimension at most
\[
H(Z) - H(C) + H(Y) - H(A) + \Delta_{\tilde{C}} + \Delta_T.
\]

To justify (A.54), we first know \(f_{22}\) is injective on \(\tilde{B} \cap f_{23}^{-1}f_{26}(\tilde{C} \cap f_{25}^{-1}f_{15}\mathcal{A})\) by (A.17). Then by Lemma II.3, we know
\[
f_{10}f_{15}t \in f_{22}(\tilde{B} \cap f_{23}^{-1}f_{26}(\tilde{C} \cap f_{25}^{-1}f_{15}\mathcal{A}))
\]
on a subspace of \(A\) of codimension at most
\[
H(Z) - H(B) + \text{codim}_B(\tilde{B} \cap f_{23}^{-1}f_{26}(\tilde{C} \cap f_{25}^{-1}f_{15}\mathcal{A})).
\]

Now again we are going to use Lemma II.1, Lemma II.3, and (A.40). Also on line (A.59) we will use the fact that \(f_{26}\) is injective on \(\tilde{C} \cap f_{25}^{-1}f_{15}\mathcal{A}\) from (A.43).
\[
\text{codim}_B(\tilde{B} \cap f_{23}^{-1}f_{26}(\tilde{C} \cap f_{25}^{-1}f_{15}\mathcal{A}))
\]
\[
\leq \Delta_{\tilde{B}} + \text{codim}_B(f_{23}^{-1}f_{26}(\tilde{C} \cap f_{25}^{-1}f_{15}\mathcal{A}))
\]
\[
\leq \Delta_{\tilde{B}} + \text{codim}_X(f_{26}(\tilde{C} \cap f_{25}^{-1}f_{15}\mathcal{A}))
\]
\[
= \Delta_{\tilde{B}} + H(X) - \dim(f_{26}(\tilde{C} \cap f_{25}^{-1}f_{15}\mathcal{A}))
\]
\[
= \Delta_{\tilde{B}} + H(X) - \dim(\tilde{C} \cap f_{25}^{-1}f_{15}\mathcal{A}).
\]
(A.59)

So, we have
\[
f_{10}f_{15}t \in f_{22}(\tilde{B} \cap f_{23}^{-1}f_{26}(\tilde{C} \cap f_{25}^{-1}f_{15}\mathcal{A}))
\]
on a subspace of \(A\) of codimension at most
\[
H(Z) - H(B) + H(X) - (H(C) + H(Y) - H(A)) + \Delta_T + \Delta_{\tilde{B}} + \Delta_{\tilde{C}}.
\]

To justify (A.55), we first know \(f_{31}\) is injective on \(\tilde{B} \cap f_{32}^{-1}f_{26}(\tilde{C} \cap f_{25}^{-1}f_{15}\mathcal{A})\) by (A.21). Then by Lemma II.3, we know
\[
f_{11}f_{15}t \in f_{31}(\tilde{B} \cap f_{32}^{-1}f_{26}(\tilde{C} \cap f_{25}^{-1}f_{15}\mathcal{A}))
\]
on a subspace of \(A\) of codimension at most
\[
H(W) - H(B) + \text{codim}_B(\tilde{B} \cap f_{32}^{-1}f_{26}(\tilde{C} \cap f_{25}^{-1}f_{15}\mathcal{A})).
\]

Now again we are going to use Lemma II.1 and Lemma II.3,
\[
\text{codim}_B(\tilde{B} \cap f_{32}^{-1}f_{26}(\tilde{C} \cap f_{25}^{-1}f_{15}\mathcal{A}))
\]
\[
\leq \Delta_{\tilde{B}} + \text{codim}_B(f_{32}^{-1}f_{26}(\tilde{C} \cap f_{25}^{-1}f_{15}\mathcal{A}))
\]
\[
\leq \Delta_{\tilde{B}} + \text{codim}_X(f_{26}(\tilde{C} \cap f_{25}^{-1}f_{15}\mathcal{A}))
\]
\[
= \Delta_{\tilde{B}} + H(X) - H(C) + H(Y) - H(A) + \Delta_T.
\]
(A.60)

The last line was derived by copying the argument from (A.60). So, we have
\[
f_{11}f_{15}t \in f_{31}(\tilde{B} \cap f_{32}^{-1}f_{26}(\tilde{C} \cap f_{25}^{-1}f_{15}\mathcal{A}))
\]
on a subspace of \(A\) of codimension at most
\[
H(W) - H(B) + H(X) - (H(C) + H(Y) - H(A)) + \Delta_T + \Delta_{\tilde{B}} + \Delta_{\tilde{C}}.
\]

From (A.51) and (A.54) we know \(\exists \tilde{a} \in \tilde{C}, \tilde{b} \in \tilde{B}\) such that
\[
f_{10}f_{15}t = f_{19}\tilde{a} = f_{22}\tilde{b}
\]
where \(f_{20}\tilde{a} \in f_{15}\mathcal{A}\) and \(f_{23}\tilde{b} \in f_{26}(\tilde{C} \cap f_{25}^{-1}f_{15}\mathcal{A})\).
(A.61)

From (A.52) and (A.55) we know \(\exists \tilde{c} \in \tilde{C}, \tilde{b} \in \tilde{B}\) such that
\[
f_{11}f_{15}t = f_{29}\tilde{c} = f_{31}\tilde{b}
\]
where \(f_{28}\tilde{c} \in f_{15}\mathcal{A}\) and \(f_{32}\tilde{b} \in f_{26}(\tilde{C} \cap f_{25}^{-1}f_{15}\mathcal{A})\).
(A.62)
From (A.53) we know $\exists \tilde{c} \in \tilde{C}$ such that
\[ f_{12} f_{15} t = f_{26} \tilde{c}, \]
where $f_{25} \tilde{c} \in f_{15} \tilde{A}$.

From (A.12) and (A.13), we know
\[ f_B f_{15} = -f_{13} \text{ on } \tilde{A} \]
\[ f_B = f_{13} f_A \text{ on } f_{15} \tilde{A}. \]

From (A.12) and (A.15), we know
\[ f_D f_{15} = -f_{14} \text{ on } \tilde{A} \]
\[ f_D = -f_{14} f_A \text{ on } f_{15} \tilde{A}. \]

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
\[ f_{7} f_{12} f_{15} t + f_{1} f_{10} f_{15} t = t \]
\[ f_{7} f_{20} \tilde{c} + f_{1} f_{19} \tilde{c} = t \]
\[ -f_{A} f_{25} \tilde{c} - f_{A} f_{20} \tilde{c} = t \]
\[ f_{A} f_{25} \tilde{c} + f_{A} f_{20} \tilde{c} = -t. \]

From (A.13) we have
\[ f_{4} f_{11} f_{15} t + f_{4} f_{10} f_{15} t = -f_{13} t. \]

Then (A.62), (A.61), (A.33), and (A.25) give
\[ f_{4} f_{11} f_{15} t + f_{4} f_{10} f_{15} t = -f_{13} t \]
\[ f_{4} f_{29} \tilde{c} + f_{4} f_{19} \tilde{c} = -f_{13} t \]
\[ -f_{B} f_{28} \tilde{c} - f_{B} f_{20} \tilde{c} = -f_{13} t. \]

By (A.62) and (A.61), we know $f_{28} \tilde{c} \in f_{15} \tilde{A}$ and $f_{20} \tilde{c} \in f_{15} \tilde{A}$. Now by (A.64), we have
\[ -f_{B} f_{28} \tilde{c} - f_{B} f_{20} \tilde{c} = -f_{13} t \]
\[ f_{13} f_{A} f_{28} \tilde{c} + f_{13} f_{A} f_{20} \tilde{c} = -f_{13} t. \]

Then using (A.12), we know $f_{A} f_{28} \tilde{c} \in \tilde{A}$ and $f_{A} f_{20} \tilde{c} \in \tilde{A}$.

By (A.44), we have
\[ f_{13} f_{A} f_{28} \tilde{c} + f_{13} f_{A} f_{20} \tilde{c} = -f_{13} t \]
\[ f_{A} f_{28} \tilde{c} + f_{A} f_{20} \tilde{c} = -t. \]  \hspace{1cm} (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
\[ f_{9} f_{12} f_{15} t + f_{6} f_{11} f_{15} t = -f_{14} t \]
\[ f_{9} f_{25} \tilde{c} + f_{6} f_{29} \tilde{c} = -f_{14} t \]
\[ -f_{D} f_{25} \tilde{c} - f_{D} f_{28} \tilde{c} = -f_{14} t. \]

By (A.63) and (A.62), we know $f_{25} \tilde{c} \in f_{15} \tilde{A}$ and $f_{28} \tilde{c} \in f_{15} \tilde{A}$. Now by (A.65), we have
\[ -f_{D} f_{25} \tilde{c} - f_{D} f_{28} \tilde{c} = -f_{14} t \]
\[ f_{14} f_{A} f_{25} \tilde{c} + f_{14} f_{A} f_{28} \tilde{c} = -f_{14} t. \]

Then using (A.12), we know $f_{A} f_{25} \tilde{c} \in \tilde{A}$ and $f_{A} f_{28} \tilde{c} \in \tilde{A}$.

By (A.45), we have
\[ f_{14} f_{A} f_{25} \tilde{c} + f_{14} f_{A} f_{28} \tilde{c} = -f_{14} t \]
\[ f_{A} f_{25} \tilde{c} + f_{A} f_{28} \tilde{c} = -t. \]  \hspace{1cm} (A.68)

From (A.24) and (A.41), we know
\[ f_{1} f_{19} = -f_{A} f_{20} \text{ on } \tilde{C} \]
\[ f_{1} = -f_{A} f_{20} f_{3} \text{ on } f_{19} (\tilde{C} \cap f_{20} f_{15} \tilde{A}). \]  \hspace{1cm} (A.69)

From (A.28) and (A.43), we know
\[ f_{7} f_{26} = -f_{A} f_{25} \text{ on } \tilde{C} \]
\[ f_{7} = -f_{A} f_{25} f_{8} \text{ on } f_{26} (\tilde{C} \cap f_{25} f_{15} \tilde{A}). \]  \hspace{1cm} (A.70)

From (A.16), we have
\[ f_{7} f_{23} = f_{1} f_{22} = 0. \]

By (A.61), we know
\[ f_{23} = f_{26} (\tilde{C} \cap f_{25} f_{15} \tilde{A}). \]

By (A.61), we also know
\[ f_{22} = f_{19} \]

which implies
\[ f_{22} = f_{19} (\tilde{C} \cap f_{20} f_{15} \tilde{A}). \]

Now we can apply (A.69) and (A.70) to give us
\[ f_{7} f_{23} = f_{1} f_{22} = 0 \]
\[ f_{A} f_{25} f_{8} f_{23} = f_{A} f_{20} f_{3} f_{22} = 0. \]

Now using (A.18), (A.61), and (A.41), we have
\[ f_{A} f_{25} f_{8} f_{23} = f_{A} f_{20} f_{3} f_{22} = 0 \]
\[ f_{A} f_{25} f_{3} f_{22} = f_{A} f_{20} f_{3} f_{22} = 0. \]

From (A.31) and (A.43), we know
\[ f_{9} f_{26} = -f_{D} f_{25} \text{ on } \tilde{C} \]
\[ f_{9} = -f_{D} f_{25} f_{8} \text{ on } f_{26} (\tilde{C} \cap f_{25} f_{15} \tilde{A}). \]  \hspace{1cm} (A.72)

From (A.35) and (A.42), we know
\[ f_{6} f_{29} = -f_{D} f_{28} \text{ on } \tilde{C} \]
\[ f_{6} = -f_{D} f_{28} f_{5} \text{ on } f_{29} (\tilde{C} \cap f_{28} f_{15} \tilde{A}). \]  \hspace{1cm} (A.73)

From (A.23), we have
\[ f_{9} f_{32} + f_{6} f_{31} = 0. \]

From (A.62) we know $f_{31} = f_{29}$ so $f_{31} = f_{29} (\tilde{C} \cap f_{28} f_{15} \tilde{A})$. From (A.62) we also know that
\[ f_{32} = f_{26} (\tilde{C} \cap f_{25} f_{15} \tilde{A}) \]
so (A.72) and (A.73) give us

\[ f_9 f_{32} \hat{b} + f_6 f_{31} \hat{b} = 0 \]
\[ -f_D f_2 s f_{32} \hat{b} - f_D f_8 f_{32} \hat{b} = 0. \]

From (A.62), we know

\[ f_{32} \hat{b} \in f_{26}(\bar{C} \cap f_{25}^{-1} f_{15} \bar{A}^T). \]

From (A.43), we know \( f_8 f_{26} = I \) on \( \bar{C} \cap f_{25}^{-1} f_{15} \bar{A}^T \). So

\[ f_{8} f_{32} \hat{b} \in f_{25}^{-1} f_{15} \bar{A}, \]

which implies \( f_{25} f_8 f_{32} \hat{b} \in f_{15} \bar{A} \). By (A.62) and (A.42), we know

\[ f_{28} f_5 f_{31} \hat{b} = f_{28} f_5 f_{29} \hat{c} = f_{28} \hat{c} \in f_{15} \bar{A}. \]

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

\[ -f_D f_2 s f_{32} \hat{b} - f_D f_8 f_{32} \hat{b} = 0 \]
\[ f_{14} f_A f_2 s f_{32} \hat{b} + f_{14} f_A f_8 f_{32} \hat{b} = 0. \]

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

\[ f_{14} f_A f_2 s f_{32} \hat{b} + f_{14} f_A f_8 f_{32} \hat{b} = 0 \]
\[ f_{A} f_2 s f_{32} \hat{b} + f_{A} f_8 f_{32} \hat{b} = 0. \]

Now by (A.22)

\[ f_{A} f_2 s f_{32} \hat{b} + f_{A} f_8 f_{32} \hat{b} = 0 \]
\[ f_{A} f_2 s f_{32} \hat{b} + f_{A} f_8 f_{32} \hat{b} = 0 \]
\[ f_{A} f_2 s f_{32} \hat{b} = f_{A} f_2 s f_{32} \hat{b}. \]

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

\[ f_{A} f_2 s f_{32} \hat{b} = f_{A} f_2 s f_{32} \hat{b} \]
\[ f_{A} f_2 s f_{32} \hat{c} = f_{A} f_2 s f_{32} \hat{c} \]
\[ f_{A} f_2 s f_{32} \hat{c} = f_{A} f_2 s f_{32} \hat{c}. \]

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

\[ -3t = 2( f_{A} f_2 s f_{32} \hat{c} + f_{A} f_2 s f_{32} \hat{c}). \]

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

\[ -3t = 2( f_{A} f_2 s f_{32} \hat{c} + f_{A} f_2 s f_{32} \hat{c}) \]
\[ -3t = 2 f_{A} f_2 s f_{32} \hat{c} + f_{A} f_2 s f_{32} \hat{c}. \]

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

\[ -3t = 2 f_{A} f_2 s f_{32} \hat{c} + f_{A} f_2 s f_{32} \hat{c}. \]

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

\[ -3t = 2 f_{A} f_2 s f_{10} f_{15} t + f_{A} f_8 f_{12} f_{15} t + f_{A} f_{11} f_{15} t \]
\[ -3t = 2 f_{A} f_2 s f_{15} t \]
\[ -3t = 0. \]

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:

\[
H(A) \leq \\
\Delta_{\bar{A}} + H(Z) - H(C) + H(Y) - H(A) + \Delta_{\bar{C}} + \Delta_{\bar{A}} \\
+ H(W) - H(C) + H(Y) - H(A) + \Delta_{\bar{C}} + \Delta_{\bar{A}} \\
+ H(X) - H(C) + H(Y) - H(A) + \Delta_{\bar{C}} + \Delta_{\bar{A}} \\
+ H(Z) - H(B) + H(X) - H(C) + H(Y) - H(A) \\
+ \Delta_{\bar{A}} + \Delta_{\bar{C}} + \Delta_{\bar{A}} \\
+ H(W) - H(B) + H(X) - H(C) + H(Y) - H(A) \\
+ \Delta_{\bar{A}} + \Delta_{\bar{C}} + \Delta_{\bar{A}} \\
= 2H(Z) + 5H(Y) + 3H(X) + 2H(W) - 5H(A) \\
- 2H(B) + 5H(C) \\
+ 6\Delta_{\bar{A}} + \Delta_{\bar{C}} + \Delta_{\bar{A}} + \Delta_{\bar{C}} + 3\Delta_{\bar{A}} \\
= 2H(Z) + 5H(Y) + 3H(X) + 2H(W) - 5H(A) \\
- 2H(B) + 5H(C) \\
+ 6( H(W) + 4H(Y) + H(Z) \\
- 2H(A) - H(B) - 2H(C) - H(D)) \\
+ 6( 3\Delta_{\bar{A}} + \Delta_{\bar{C}} + \Delta_{\bar{A}} + \Delta_{\bar{C}} + 3\Delta_{\bar{A}} \\
+ \Delta_{\bar{C}} + \Delta_{\bar{A}} + \Delta_{\bar{C}} + 3\Delta_{\bar{A}} \\
= 8H(Z) + 29H(Y) + 3H(X) + 8H(W) \\
- 6H(D) - 17H(C) - 8H(B) - 17H(A) \\
+ 18\Delta_{\bar{A}} + 7\Delta_{\bar{C}} + 7\Delta_{\bar{A}} + 7\Delta_{\bar{C}} + 3\Delta_{\bar{A}} + 6\Delta_{\bar{C}} \\
= 8H(Z) + 29H(Y) + 3H(X) + 8H(W) - 6H(D) \\
- 17H(C) - 8H(B) - 17H(A) \\
+ 55H(Z|A,B,C) + 35H(Y|W,X,Z) \\
+ 50H(X|A,C,D) + 49H(W|B,C,D) \\
+ 18H(A|B,D,Y) + 7H(B|D,X,Z) + H(B|A,W,X) \\
+ 7H(C|D,Y,Z) \\
+ 7H(C|B,X,Y) + 3H(C|A,W,Y) + 6H(D|A,W,Z) \\
+ 49H(A) + H(B) + H(C) + H(D) - H(A,B,C,D)).
\]

Proof of Theorem IV.1.

The proof of this theorem follows a similar strategy as discussed at the beginning of the proof of Theorem III.1, and again, is rather tedious.

By Lemma II.4 we get linear functions:
\[ f_1 : W \to B, \quad f_2 : W \to C, \quad f_3 : W \to D, \]
\[ f_4 : X \to A, \quad f_5 : X \to C, \quad f_6 : X \to D, \]
\[ f_7 : Y \to A, \quad f_8 : Y \to B, \quad f_9 : Y \to D, \]
\[ f_{10} : Z \to A, \quad f_{11} : Z \to B, \quad f_{12} : Z \to C, \]
\[ f_{13} : A \to B, \quad f_{14} : A \to W, \quad f_{15} : A \to X, \]
\[ f_{16} : C \to A, \quad f_{17} : C \to W, \quad f_{18} : C \to Y, \]
\[ f_{19} : B \to C, \quad f_{20} : B \to X, \quad f_{21} : B \to Y, \]
\[ f_{22} : D \to W, \quad f_{23} : D \to A, \quad f_{24} : D \to Z, \]
\[ f_{25} : B \to X, \quad f_{26} : B \to D, \quad f_{27} : B \to Z, \]
\[ f_{28} : C \to Y, \quad f_{29} : C \to Z, \quad f_{30} : C \to D, \]
\[ f_{31} : A \to W, \quad f_{32} : A \to X, \quad f_{33} : A \to Y, \quad f_{34} : A \to Z. \]

such that

\[ f_1 + f_2 + f_3 = I \]

on a subspace of \( W \) of codimension \( H(W|B, C, D) \)

\[ (A.76) \]

\[ f_4 + f_5 + f_6 = I \]

on a subspace of \( X \) of codimension \( H(X|A, C, D) \)

\[ (A.77) \]

\[ f_7 + f_8 + f_9 = I \]

on a subspace of \( Y \) of codimension \( H(Y|A, B, D) \)

\[ (A.78) \]

\[ f_{10} + f_{11} + f_{12} = I \]

on a subspace of \( Z \) of codimension \( H(Z|A, B, C) \)

\[ (A.79) \]

\[ f_{13} + f_{14} + f_{15} = I \]

on a subspace of \( A \) of codimension \( H(A|B, W, X) \)

\[ (A.80) \]

\[ f_{16} + f_{17} + f_{18} = I \]

on a subspace of \( C \) of codimension \( H(C|A, W, Y) \)

\[ (A.81) \]

\[ f_{19} + f_{20} + f_{21} = I \]

on a subspace of \( B \) of codimension \( H(B|C, X, Y) \)

\[ (A.82) \]

\[ f_{22} + f_{23} + f_{24} = I \]

on a subspace of \( D \) of codimension \( H(D|A, W, Z) \)

\[ (A.83) \]

\[ f_{25} + f_{26} + f_{27} = I \]

on a subspace of \( B \) of codimension \( H(B|D, X, Z) \)

\[ (A.84) \]

\[ f_{28} + f_{29} + f_{30} = I \]

on a subspace of \( C \) of codimension \( H(C|D, Y, Z) \)

\[ (A.85) \]

\[ f_{31} + f_{32} + f_{33} + f_{34} = I \]

on a subspace of \( A \) of codimension \( H(A|W, X, Y, Z) \).

\[ (A.86) \]

Now combining some functions we obtained from Lemma II.4 gives four new functions:

\[ f_{4} f_{32} + f_{7} f_{33} + f_{10} f_{34} : A \to A, \]
\[ f_{1} f_{31} + f_{8} f_{33} + f_{11} f_{34} : A \to B, \]
\[ f_{2} f_{31} + f_{5} f_{32} + f_{12} f_{34} : A \to C, \]
\[ f_{3} f_{31} + f_{6} f_{32} + f_{9} f_{33} : A \to D. \]

Using \((A.76)-(A.79), (A.86), \) Lemma II.1, and Lemma II.3 we know the sum of these four functions is equal to \( I \) on a subspace of \( A \) of codimension at most

\[ H(W|B, C, D) + H(X|A, C, D) + H(Y|A, B, D) \]
\[ + H(Z|A, B, C) + H(A|W, X, Y, Z) \]

since, on that subspace,

\[ (f_{4} f_{32} + f_{7} f_{33} + f_{10} f_{34}) + (f_{1} f_{31} + f_{8} f_{33} + f_{11} f_{34}) + (f_{2} f_{31} + f_{5} f_{32} + f_{12} f_{34}) + (f_{3} f_{31} + f_{6} f_{32} + f_{9} f_{33}) \]
\[ = (f_{1} + f_{2} + f_{3}) f_{31} + (f_{4} + f_{5} + f_{6}) f_{32} + (f_{7} + f_{8} + f_{9}) f_{33} \]
\[ + (f_{10} + f_{11} + f_{12}) f_{34} \]
\[ = f_{31} + f_{32} + f_{33} + f_{34} \]
\[ = I. \]

Now applying Lemma II.6 and Lemma II.1 to the functions

\[ f_{4} f_{32} + f_{7} f_{33} + f_{10} f_{34} - I \]
\[ f_{1} f_{31} + f_{8} f_{33} + f_{11} f_{34} \]
\[ f_{2} f_{31} + f_{5} f_{32} + f_{12} f_{34} \]
\[ f_{3} f_{31} + f_{6} f_{32} + f_{9} f_{33} \]

we get a subspace \( \tilde{A} \) of \( A \) of codimension at most

\[ \Delta_{\tilde{A}} = H(W|B, C, D) + H(X|A, C, D) + H(Y|A, B, D) \]
\[ + H(Z|A, B, C) + H(A|W, X, Y, Z) \]
\[ + H(A) + H(B) + H(C) + H(D) - H(A, B, C, D) \]

on which

\[ f_{4} f_{32} + f_{7} f_{33} + f_{10} f_{34} = I \]
\[ f_{1} f_{31} + f_{8} f_{33} + f_{11} f_{34} = 0 \]
\[ f_{2} f_{31} + f_{5} f_{32} + f_{12} f_{34} = 0 \]
\[ f_{3} f_{31} + f_{6} f_{32} + f_{9} f_{33} = 0. \]

Similarly, we get a subspace \( \tilde{T} \) of \( A \) of codimension at most

\[ \Delta_{\tilde{T}} = H(W|B, C, D) + H(X|A, C, D) + H(A|B, W, X) \]
\[ + H(A) + H(B) + H(C) + H(D) - H(A, B, C, D) \]

on which

\[ f_{4} f_{15} = I \]
\[ f_{13} + f_{14} = 0 \]
\[ f_{2} f_{14} + f_{5} f_{15} = 0 \]
\[ f_{3} f_{14} + f_{6} f_{15} = 0. \]

We get a subspace \( \widetilde{B} \) of \( B \) of codimension at most

\[ \Delta_{\widetilde{B}} = H(X|A, C, D) + H(Y|A, B, D) + H(B|C, X, Y) \]
\[ + H(A) + H(B) + H(C) + H(D) - H(A, B, C, D) \]

on which

\[ f_{4} f_{20} + f_{7} f_{21} = 0 \]
\[ f_{8} f_{21} = I \]
\[ f_{19} + f_{5} f_{20} = 0 \]
\[ f_{6} f_{20} + f_{9} f_{21} = 0. \]
We get a subspace $\widehat{B}$ of $B$ of codimension at most
\[ \Delta_{\widehat{B}} = H(X|A, C, D) + H(Z|A, B, C) + H(B|D, X, Z) + H(A) + H(B) + H(C) + H(D) - H(A, B, C, D) \]
on which
\[
\begin{align*}
&f_4 f_{25} + f_{10} f_{27} = 0 \\
&f_{11} f_{27} = 1 \\
&f_5 f_{25} + f_{12} f_{27} = 0 \\
&f_6 f_{25} + f_{26} = 0.
\end{align*}
\] (A.99) (A.100) (A.101) (A.102)

Considering again (A.100) and (A.109), we can apply Lemma II.7 to show that $f_{11} f_{29}$ is injective on $\widehat{C}^*$. By (A.108), we know
\[ f_{8} f_{28} \text{ is injective on } \widehat{C}^*. \] (A.116)

Let $\widehat{A}^* = f_4 (f_{15} A \cap f_{25} \widehat{B}^*) \subseteq \widehat{A}$. Considering (A.91) and (A.115), we can apply Lemma II.7 to show that $f_{5} f_{15}$ is injective on $\widehat{A}^*$. By (A.93), we know $f_{2} f_{14}$ is injective on $\widehat{A}^*$ which implies
\[ f_{14} \text{ is injective on } \widehat{A}. \] (A.117)

Let $\widehat{C}^* = f_2 (f_{17} C \cap f_{22} D) \subseteq \widehat{C}$. Considering (A.105) and (A.114), we can apply Lemma II.7 to show that $f_{3} f_{17}$ is injective on $\widehat{C}^*$. Then by (A.106), we know
\[ f_{9} f_{18} \text{ is injective on } \widehat{C}^*. \] (A.118)

Let $\widehat{B}^* = f_8 (f_{21} B \cap f_{18} \widehat{C}^*) \subseteq \widehat{B}$. Considering (A.96) and (A.118), we can apply Lemma II.7 to show that
\[ f_{9} f_{21} \text{ is injective on } \widehat{B}^*. \] (A.119)

By (A.98), we know
\[ f_{6} f_{20} \text{ is injective on } \widehat{B}^*. \] (A.120)

which implies
\[ f_{20} \text{ is injective on } \widehat{B}^*. \] (A.121)

Let us define the functions
\[ g_{14} = (f_{14} | \widehat{A}^*)^{-1}, \quad g_{20} = (f_{20} | \widehat{B}^*)^{-1} \]
where $f_{14} | \widehat{A}^*$ and $f_{20} | \widehat{B}^*$ are the restrictions of the functions $f_{14}$ and $f_{20}$ to the sets $\widehat{A}$ and $\widehat{B}$, respectively. Now, considering (A.96), (A.100), (A.105), and (A.109) we have
\[ f_1 = -f_8 f_{18} f_2 \text{ on } f_{17} C \]  

from (A.104)  

(A.122)

\[ f_2 = -f_5 f_{15} g_{14} \text{ on } f_{14} \overline{A} \]  

from (A.93)  

(A.123)

\[ f_3 = -f_6 f_5 g_{14} \text{ on } f_{14} \overline{A} \]  

and \[ f_3 = -f_6 f_{18} f_2 \text{ on } f_{17} C \]  

from (A.94), (A.106)  

(A.124)

\[ f_4 = -f_7 f_{21} g_{20} \text{ on } f_{20} \overline{B} \]  

from (A.95)  

(A.125)

\[ f_6 = -f_3 f_{21} g_{20} \text{ on } f_{20} \overline{B} \]  

from (A.98)  

(A.126)

\[ f_7 = -f_4 f_{20} f_8 \text{ on } f_{21} \overline{B} \]  

from (A.95)  

(A.127)

\[ f_9 = -f_6 f_{20} f_8 \text{ on } f_{21} \overline{B} \]  

from (A.98)  

(A.128)

\[ f_{10} = -f_4 f_{25} f_{11} \text{ on } f_{27} \overline{B} \]  

and \[ f_{10} = -f_7 f_{28} f_{12} \text{ on } f_{29} \overline{C} \]  

from (A.99), (A.107)  

(A.129)

\[ f_{11} = -f_8 f_{28} f_{12} \text{ on } f_{29} \overline{C} \]  

from (A.108)  

(A.130)

\[ f_{12} = -f_5 f_{25} f_{11} \text{ on } f_{27} \overline{B}. \]  

from (A.101)  

(A.131)

Next, we provide upper bounds for the codimensions of \( \overline{A} \), \( \overline{B} \), \( \overline{B}^* \), \( \overline{C} \), and \( \overline{C}^* \). From (A.100), we know \( f_{11} \) is injective on \( f_{27} \overline{B} \) and \( f_{27} \overline{B} \) is injective on \( \overline{B} \). These facts will be used to arrive on lines (A.132) and (A.134). From (A.109), we know \( f_{29} \) is injective on \( \overline{C} \), which will also be used to arrive on line (A.134). Lemma II.1 will be used to arrive on (A.133).

\[ \text{codim}_{\overline{A}} \overline{B}^* \]  

\[ = H(B) - \text{dim}(\overline{B}^*) \]  

\[ = H(B) - \text{dim}(f_{11}(f_{27} \overline{B} \cap f_{29} \overline{C})) \]  

\[ = H(B) - \text{dim}(f_{27} \overline{B} \cap f_{29} \overline{C}) \]  

(A.132)

\[ = H(B) - H(Z) + \text{codim}_{\overline{Z}}(f_{27} \overline{B} \cap f_{29} \overline{C}) \]  

\[ \leq H(B) - H(Z) + \text{codim}_{\overline{Z}}(f_{27} \overline{B}) + \text{codim}_{\overline{Z}}(f_{29} \overline{C}) \]  

(A.133)

\[ = H(B) - H(Z) + \text{dim}(f_{27} \overline{B}) \]  

\[ + H(Z) - \text{dim}(f_{29} \overline{C}) \]  

\[ = H(B) + H(Z) - \text{dim}(\overline{B}) - \text{dim}(\overline{C}) \]  

(A.134)

\[ \leq H(B) + H(Z) - H(B) + \Delta_B + H(C) + \Delta_C \]  

(A.135)

\[ \leq H(Z) - H(C) + \Delta_B + \Delta_C \]  

(A.136)

\[ \triangleq \Delta_{\overline{B}}. \]

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 \( \overline{B}^* \), which will also be used to arrive on line (A.139). Lemma II.1 will be used to arrive on (A.138).

\[ \text{codim}_{\overline{A}} \overline{A} \]  

\[ = H(A) - \text{dim}(\overline{A}) \]  

\[ = H(A) - \text{dim}(f_{15} \overline{A} \cap f_{15} \overline{A}) \]  

\[ = H(A) - \text{dim}(\overline{B}^* \cap f_{15} \overline{A}) \]  

(A.137)

\[ = H(A) - H(X) + \text{codim}_X(f_{25} \overline{B}^* \cap f_{15} \overline{A}) \]  

(\[ \leq H(A) - H(X) + \text{codim}_X(f_{25} \overline{B}^*) + \text{codim}_X(f_{15} \overline{A}) \]  

(A.138)

\[ = H(A) - H(X) - \text{dim}(f_{25} \overline{B}^*) - \text{dim}(f_{15} \overline{A}) \]  

\[ = H(A) - H(X) - \text{dim}(\overline{B}^*) - \text{dim}(\overline{A}) \]  

(A.139)

\[ \leq H(A) - H(X) - H(B) + \Delta_B - H(A) + \Delta_A \]  

\[ = H(X) - H(B) + H(Z) - H(C) + \Delta_B + \Delta_C + \Delta_A \]  

\[ \triangleq \Delta_{\overline{B}}. \]

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 II.1 will be used to arrive on (A.141).

\[ \text{codim}_{\overline{A}} \overline{C} \]  

\[ = H(C) - \text{dim}(\overline{C}) \]  

\[ = H(C) - \text{dim}(f_{17} \overline{C} \cap f_{22} \overline{D}) \]  

\[ = H(C) - \text{dim}(f_{17} \overline{C} \cap f_{27} \overline{D}) \]  

(A.140)

\[ = H(C) - H(W) + \text{codim}_W(f_{17} \overline{C} \cap f_{27} \overline{D}) \]  

\[ \leq H(C) - H(W) + \text{codim}_W(f_{17} \overline{C}) + \text{codim}_W(f_{27} \overline{D}) \]  

(A.141)

\[ = H(C) - H(W) + H(W) - \text{dim}(f_{17} \overline{C}) \]  

\[ + H(W) - \text{dim}(f_{27} \overline{D}) \]  

\[ = H(C) + H(W) - \text{dim}(\overline{C}) - \text{dim}(\overline{D}) \]  

(A.142)

\[ \leq H(C) + H(W) - H(C) + \Delta_{\overline{C}} - H(D) + \Delta_{\overline{D}} \]  

\[ = H(W) - H(D) + \Delta_{\overline{C}} + \Delta_{\overline{D}} \]  

\[ \triangleq \Delta_{\overline{C}}. \]

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 II.1
will be used to arrive on (A.144).

\[
\text{codim}_{H}B^* = H(B) - \dim(B^*) \\
= H(B) - \dim(f_{18}(f_{21}B \cap f_{18}C^*)) \\
= H(B) - \dim(f_{21}B \cap f_{18}C^*) \\
= H(B) - H(Y) + \text{codim}_Y(f_{21}B \cap f_{18}C^*) \\
\leq H(B) - H(Y) + \text{codim}_Y(f_{21}B) + \text{codim}_Y(f_{18}C^*) \\
= H(B) - H(Y) + (H(Y) - \dim(f_{21}B)) \\
+ H(Y) - \dim(f_{18}C^*) \\
= H(B) + H(Y) - \dim(B) - \dim(C^*) \\
\leq H(B) + H(Y) - H(B) + \Delta_{\overline{\text{t}}} - H(C) + \Delta_{\overline{\text{c}}} \\
= H(Y) - H(C) + \Delta_{\overline{\text{t}}} + \Delta_{\overline{\text{c}}} \\
= H(Y) - H(C) + H(W) - H(D) + \Delta_{\overline{\text{c}}} + \Delta_{\overline{\text{t}}} + \Delta_{\overline{\text{b}}} \\
\triangleq \Delta_{\overline{\text{t}}}.
\]

From (A.109), we know \( f_{12} \) is injective on \( f_{29}\overline{C} \) and \( f_{29} \) is injective on \( \overline{C} \). These facts will be used to arrive on lines (A.146) and (A.148). From (A.100), we know \( f_{27} \) is injective on \( \overline{B} \), which will also be used on line (A.148). Lemma II.1 will be used to arrive on (A.147).

\[
\text{codim}_{C\overline{C}^*} = H(C) - \dim(C^*) \\
= H(C) - \dim(f_{12}(f_{27}\overline{B} \cap f_{29}\overline{C})) \\
= H(C) - \dim(f_{27}\overline{B} \cap f_{29}\overline{C}) \\
= H(C) - H(Z) + \text{codim}_Z(f_{27}\overline{B} \cap f_{29}\overline{C}) \\
\leq H(C) - H(Z) + \text{codim}_Z(f_{27}\overline{B}) + \text{codim}_Z(f_{29}\overline{C}) \\
= H(C) - H(Z) + H(Z) - \dim(f_{27}\overline{B}) \\
+ H(Z) - \dim(f_{29}\overline{C}) \\
= H(C) + H(Z) - \dim(\overline{B}) - \dim(\overline{C}) \\
\leq H(C) + H(Z) - H(B) + \Delta_{\overline{B}} - H(C) + \Delta_{\overline{c}} \\
= H(Z) - H(B) + \Delta_{\overline{B}} + \Delta_{\overline{c}} \\
\triangleq \Delta_{\overline{c}}.
\]

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

\[
f_{32}t \in f_{20}\overline{B} \cap f_{25}\overline{B}^* : \\
\text{this is true on a subspace of } A \text{ of codimension at most } \Delta_{\overline{c}}. \\
\text{(A.149)}
\]

\[
f_{33}t \in f_{28}\overline{C}^* \cap f_{21}\overline{B}^* : \\
\text{this is true on a subspace of } A \text{ of codimension at most } 2H(X) - 2H(B) + \Delta_{\overline{t}} + \Delta_{\overline{c}}. \\
\text{(A.150)}
\]

\[
f_{34}t \in f_{29}\overline{C}^* \cap f_{27}\overline{B}^* : \\
\text{this is true on a subspace of } A \text{ of codimension at most } 2H(Y) - H(B) - H(C) + \Delta_{\overline{t}} + \Delta_{\overline{c}}. \\
\text{(A.151)}
\]

Now, we need to make two assumptions on \( t \) simultaneously.

\[
f_{31}t \in f_{17}\overline{C} \cap f_{14}\overline{A}^* \text{ and } f_{15}g_{14}f_{31}t \in f_{20}\overline{B} \cap f_{25}\overline{B}^* : \\
\text{this is true on a subspace of } A \text{ of codimension at most } 2H(X) - 2H(B) + 2H(W) - H(C) - H(A) \\
+ \Delta_{\overline{t}} + \Delta_{\overline{c}} + \Delta_{\overline{t}} + \Delta_{\overline{b}}. \\
\text{(A.152)}
\]

To justify (A.150), first we know \( f_{20} \) is injective on \( \overline{B}^* \) by (A.120). Then by Lemma II.3, we know \( f_{32}t \in f_{20}\overline{B}^* \) on a subspace of \( A \) of codimension at most

\[
H(X) - H(B) + \text{codim}_{B}(\overline{B}^*) \leq H(X) - H(B) + \Delta_{\overline{t}}. \\
\text{(A.153)}
\]

By (A.115), we also know \( f_{25} \) is injective on \( \overline{B}^* \). Then by Lemma II.3, we know \( f_{32}t \in f_{25}\overline{B}^* \) on a subspace of \( A \) of codimension at most

\[
H(X) - H(B) + \text{codim}_{\overline{B}^*}(\overline{B}^*) \leq H(X) - H(B) + \Delta_{\overline{b}}. \\
\text{(A.154)}
\]

Then using Lemma II.1, we have \( f_{32}t \in f_{20}\overline{B} \cap f_{25}\overline{B}^* \) on a subspace of \( A \) of codimension at most

\[
2H(X) - 2H(B) + \Delta_{\overline{t}} + \Delta_{\overline{b}}. \\
\text{(A.155)}
\]

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

To justify (A.154), first we know \( f_{17} \) is injective on \( \overline{B} \) by (A.105). Then by Lemma II.3, we know \( f_{31}t \in f_{17}\overline{C} \) on a subspace of \( A \) of codimension at most

\[
H(W) - H(C) + \text{codim}_{\overline{C}}(\overline{C}) \leq H(W) - H(C) + \Delta_{\overline{t}}. \\
\text{(A.156)}
\]

By (A.117), we also know \( f_{14} \) is injective on \( \overline{A} \). Then by Lemma II.3, we know \( f_{31}t \in f_{14}\overline{A}^* \) on a subspace of \( A \) of codimension at most

\[
H(W) - H(A) + \text{codim}_{\overline{A}}(\overline{A}) \leq H(W) - H(A) + \Delta_{\overline{t}}. \\
\text{(A.157)}
\]

Then using Lemma II.1, we have

\[
f_{31}t \in f_{17}\overline{C} \cap f_{14}\overline{A}^*. \\
\text{(A.158)}
\]
on a subspace, $S$, of $A$ of codimension at most
\[ 2H(W) - H(C) - H(A) + \Delta_{\mathcal{C}} + \Delta_{\mathcal{B}}. \]
Since $f_{14}$ is injective on $\mathcal{A}$, the function $f_{15}g_{14}f_{31}$ is defined on $S$. Using the same technique as before we can show that
\[ f_{15}g_{14}f_{31}t \in f_{20}\hat{\mathcal{B}} \cap f_{25}\hat{\mathcal{B}}^* \]
on a subspace, $\mathcal{S}$, of codimension with respect to $S$ at most
\[ 2H(X) - 2H(B) + \Delta_{\mathcal{B}}^* + \Delta_{\mathcal{B}}^*. \]
Thus both conditions are true on $\mathcal{S}$, which has codimension with respect to $A$ at most
\[ \text{codim}_A \mathcal{S} + \text{codim}_A S \]
\[ \leq 2H(X) - 2H(B) + 2H(W) - H(C) - H(A)
\[ + \Delta_{\mathcal{C}} + \Delta_{\mathcal{B}}^* + \Delta_{\mathcal{B}}^*. \]
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 $f_{4}f_{32}t = f_{10}f_{34}t$.

**Claim.** $f_{4}f_{32}t = f_{10}f_{34}t$

**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
\[ f_{8}f_{33}t = -f_{11}f_{34}t - f_{1}f_{31}t. \]
Then by using (A.130) and condition (A.152), we have
\[ f_{8}f_{33}t = f_{8}f_{28}f_{12}f_{34}t - f_{1}f_{31}t. \]
Now, by using (A.122) and condition (A.154), we have
\[ f_{8}f_{33}t = f_{8}f_{28}f_{12}f_{34}t + f_{8}f_{18}f_{2}f_{31}t. \]
By (A.116), we know $f_{8}$ is injective on $f_{28}\hat{\mathcal{C}}^*$. By condition (A.151), we know
\[ f_{33}t \in f_{28}\hat{\mathcal{C}}^*. \]
By condition (A.153), we know
\[ f_{18}f_{2}f_{31}t \in f_{28}\hat{\mathcal{C}}^*. \]
By condition (A.152), we know
\[ f_{34}t \in f_{25}\hat{\mathcal{C}}^*. \]
Using (A.109), we know
\[ f_{12}f_{34}t \in \hat{\mathcal{C}}^*. \]
Thus, we have
\[ f_{33}t = f_{28}f_{12}f_{34}t + f_{18}f_{2}f_{31}t. \]
(A.155)
By (A.90), we have
\[ f_{9}f_{33}t = -f_{6}f_{34}t - f_{3}f_{31}t. \]
Then by using (A.126) and condition (A.150), we have
\[ f_{9}f_{33}t = f_{9}f_{21}g_{20}f_{32}t - f_{3}f_{31}t. \]
Now, by using (A.124) and condition (A.154), we have
\[ f_{9}f_{33}t = f_{9}f_{21}g_{20}f_{32}t + f_{9}f_{18}f_{2}f_{31}t. \]
By (A.119), we know $f_{9}$ is injective on $f_{21}\hat{\mathcal{B}}^*$. By condition (A.151), we know
\[ f_{33}t \in f_{21}\hat{\mathcal{B}}^*. \]
By condition (A.150), we know
\[ f_{32}t \in f_{20}\hat{\mathcal{B}}^*. \]
so
\[ f_{21}g_{20}f_{32}t \in f_{21}\hat{\mathcal{B}}^*. \]
By condition (A.153), we know
\[ f_{18}f_{2}f_{31}t \in f_{21}\hat{\mathcal{B}}^*. \]
Thus, we have
\[ f_{33}t = f_{21}g_{20}f_{32}t + f_{18}f_{2}f_{31}t. \]
(A.156)
Now, setting (A.155) and (A.156) equal to each other, we have
\[ f_{21}g_{20}f_{32}t = f_{28}f_{12}f_{34}t. \]
(A.157)
By (A.125) and condition (A.150), we know
\[ f_{4}f_{32}t = -f_{7}f_{21}g_{20}f_{32}t. \]
Using (A.157), we have
\[ f_{4}f_{32}t = -f_{7}f_{28}f_{12}f_{34}t. \]
Then using (A.129) and condition (A.152), we know
\[ f_{4}f_{32}t = f_{10}f_{34}t. \]

**Claim.** $f_{7}f_{33}t = f_{10}f_{34}t$.

**Proof.** First we must show that $f_{25}f_{11}f_{34}t = f_{20}f_{6}f_{33}t$. By (A.89), we know
\[ f_{5}f_{32}t = -f_{12}f_{34}t - f_{2}f_{31}t. \]
Then by using (A.131) and condition (A.152), we have
\[ f_{5}f_{32}t = f_{5}f_{25}f_{11}f_{34}t - f_{2}f_{31}t. \]
Now, by using (A.123) and condition (A.154), we have
\[ f_{5}f_{32}t = f_{5}f_{25}f_{11}f_{34}t + f_{5}f_{15}g_{14}f_{31}t. \]
By (A.115), we know $f_{5}$ is injective on $f_{25}\hat{\mathcal{B}}^*$. By condition (A.150), we know
\[ f_{32}t \in f_{25}\hat{\mathcal{B}}^*. \]
By condition (A.152), we know
\[ f_{34}t \in f_{27}\hat{\mathcal{B}}^*. \]
Now, using (A.100), we know
\[ f_{11}f_{34}t \in \hat{\mathcal{B}}^*. \]
By condition (A.154), we know
\[ f_{15}g_{14}f_{31}t \in f_{25}\hat{\mathcal{B}}^*. \]
Thus, we have

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

$$f_{6f_{32}t} = f_{6f_{20}f_{8}f_{33}t} - f_{3f_{31}t}.$$  

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

$$f_{6f_{32}t} = f_{6f_{20}f_{8}f_{33}t} + f_{6f_{15g_{14}f_{31}t}}.$$  

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

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

By condition (A.151), we know

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

Now, using (A.96), we know

$$f_{8f_{33}t} \in B^*.$$  

By condition (A.154), we know

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

Thus, we have

$$f_{32t} = f_{20f_{8}f_{33}t} + f_{15g_{14}f_{31}t}.$$  \hfill (A.159)

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

$$f_{25f_{11}f_{34}t} = f_{20f_{8}f_{33}t}.$$  \hfill (A.160)

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

$$f_{7f_{33}t} = -f_{4f_{20}f_{8}f_{33}t}.$$  

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

$$f_{7f_{33}t} = f_{10f_{34}t}.$$  

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

$$t = f_{4f_{32}t} + f_{7f_{33}t} + f_{10f_{34}t}$$

$$= f_{10f_{34}t} + f_{10f_{34}t} + f_{10f_{34}t}$$

$$= 3f_{10f_{34}t}.$$  

Thus if the field has characteristic 3, then

$$t = 0.$$  \hfill (A.161)

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

$$H(A) \leq$$

$$\Delta_{\tilde{A}} + 2H(W) - H(C) - H(A) + \Delta_{\tilde{X}} + \Delta_{\tilde{T}}$$

$$+ 2H(X) - 2H(B) + \Delta_{\tilde{T}} + \Delta_{\tilde{B}}$$

$$+ 2H(Y) - H(B) - H(C) + \Delta_{\tilde{T}} + \Delta_{\tilde{C}}$$

$$+ 2H(Z) - H(B) - H(C) + \Delta_{\tilde{T}} + \Delta_{\tilde{B}}$$

$$+ 2H(X) - 2H(B) + \Delta_{\tilde{T}} + \Delta_{\tilde{B}}$$

$$= 2H(Z) + 4H(Y) + 4H(X) + 2H(W) - 4H(C)$$

$$- 7H(B) - H(A)$$

$$+ \Delta_{\tilde{T}} + 4\Delta_{\tilde{T}} + 3\Delta_{\tilde{B}} + 3\Delta_{\tilde{C}} + \Delta_{\tilde{A}} + \Delta_{\tilde{T}}$$

$$= 2H(Z) + 4H(Y) + 4H(X) + 2H(W) - 4H(C)$$

$$- 7H(B) - H(A)$$

$$+ H(X) - H(B) + H(Z) - H(C) + \Delta_{\tilde{B}} + \Delta_{\tilde{C}} + \Delta_{\tilde{T}}$$

$$+ H(Y) - H(C) + H(W) - H(D) + \Delta_{\tilde{T}} + \Delta_{\tilde{T}} + \Delta_{\tilde{T}}$$

$$+ 3H(Z) - H(C) + \Delta_{\tilde{B}} + \Delta_{\tilde{C}}$$

$$+ 3H(Z) - H(B) + \Delta_{\tilde{B}} + \Delta_{\tilde{C}}$$

$$+ \Delta_{\tilde{A}} + \Delta_{\tilde{T}}$$

$$= 9H(Z) + 8H(Y) + 5H(X) + 6H(W) - 4H(D)$$

$$- 12H(C) - 11H(B) - H(A)$$

$$+ \Delta_{\tilde{A}} + \Delta_{\tilde{T}} + 7\Delta_{\tilde{B}} + 4\Delta_{\tilde{T}} + 7\Delta_{\tilde{C}} + 5\Delta_{\tilde{B}} + 4\Delta_{\tilde{T}}$$

$$= 9H(Z) + 8H(Y) + 5H(X) + 6H(W) - 4H(D)$$

$$- 12H(C) - 11H(B) - H(A)$$

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


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

$$+ 7H(X|A, C, D) + H(Z|A, B, C) + H(B|D, X, Z))$$

$$+ 4H(X|A, C, D) + H(Y|A, B, D) + H(B|C, X, Y))$$

$$+ 7H(Y|A, B, D) + H(Z|A, B, C) + H(C|D, Y, Z))$$

$$+ 5H(W|B, C, D) + H(Y|A, B, D) + H(C|A, W, Y))$$


$$+ 29H(A) + H(B) + H(C) + H(D) - H(A, B, C, D))$$

$$= 9H(Z) + 8H(Y) + 5H(X) + 6H(W) - 4H(D)$$

$$- 12H(C) - 11H(B) - H(A)$$

$$+ 19H(Z|A, B, C) + 17H(Y|A, B, D)$$

$$+ 13H(X|A, C, D) + 11H(W|B, C, D)$$

$$+ H(A|W, X, Y, Z) + H(A|B, W, X) + 7H(B|D, X, Z)$$

$$+ 4H(B|C, X, Y)$$

$$+ 7H(C|D, Y, Z) + 5H(C|A, W, Y) + 4H(D|A, W, Z)$$

$$+ 29H(A) + H(B) + H(C) + H(D) - H(A, B, C, D)).$$

References


