# Worst-Case Interactive Communication II: Two Messages are Not Optimal*

Alon Orlitsky

AT&T Bell Laboratories

**Abstract**

$X$ and $Y$ are random variables. Person $P_{\mathcal{X}}$ knows $X$, Person $P_{\mathcal{Y}}$ knows $Y$, and both know the joint probability distribution of the pair $(X,Y)$. Using a predetermined protocol, they communicate over a binary, error-free, channel in order for $P_{\mathcal{Y}}$ to learn $X$. $P_{\mathcal{X}}$ may or may not learn $Y$. How many information bits must be transmitted (by both persons) in the worst case if only $m$ messages are allowed?

$\hat{C}_1(X|Y)$ is the number of bits required when at most one message is allowed, necessarily from $P_{\mathcal{X}}$ to $P_{\mathcal{Y}}$. $\hat{C}_2(X|Y)$ is the number of bits required when at most two messages are permitted: $P_{\mathcal{Y}}$ transmits a message to $P_{\mathcal{X}}$, then $P_{\mathcal{X}}$ responds with a message to $P_{\mathcal{Y}}$. $\hat{C}_\infty(X|Y)$ is the number of bits required when communication is unrestricted: $P_{\mathcal{X}}$ and $P_{\mathcal{Y}}$ can communicate back and forth.

It is known that one-message communication may require exponentially more bits than the minimum necessary: for some $(X,Y)$ pairs, $\hat{C}_1(X|Y) = 2^{\hat{C}_\infty(X|Y)-1}$. Yet just two messages suffice to reduce communication to almost the minimum: for all $(X,Y)$ pairs $\hat{C}_2(X|Y) \le 4\hat{C}_\infty(X|Y)+3$. It was further shown that for a large class of $(X,Y)$ pairs, two messages are optimal: $\hat{C}_2(X|Y) = \hat{C}_\infty(X|Y)$. It remained uncertain whether two message are optimal for all $(X,Y)$ pairs.

In this paper we define the chromatic-decomposition number of a hypergraph and show that under general conditions on $(X,Y)$ it can be used to determine $\hat{C}_2(X|Y)$. We use this result to prove that for some $(X,Y)$ pairs, two-message communication requires twice the minimum number of bits: for all positive $\epsilon$ and $c$ we present an $(X,Y)$ pair for which $\hat{C}_2(X|Y) \ge (2-\epsilon)\hat{C}_\infty(X|Y) \ge c$.

*Appeared, IEEE Transactions on Information Theory, July 1991, pp. 995–1005. This version was printed on March 22, 1996.

# 1 Introduction

Consider two *communicators*: an *informant* $P_{\mathcal{X}}$ having a random variable $X$ and a *recipient* $P_{\mathcal{Y}}$ having a, possibly dependent, random variable $Y$. Both communicators want the recipient, $P_{\mathcal{Y}}$, to learn $X$ with no probability of error, whereas the informant, $P_{\mathcal{X}}$, may or may not learn $Y$. To that end they alternate in transmitting *messages*: finite sequences of bits. Messages are transmitted over an error-free channel and are determined by an agreed-upon, deterministic, protocol.

A formal definition of protocols is given in [1]. It follows the definition of [2] for a slightly different model. Essentially, a *protocol for* $(X, Y)$ (i.e., a protocol for transmitting $X$ to a person who knows $Y$) guarantees that the following properties hold. (1) Separate transmissions: each message is based on the random variable known to its transmitter and on previous messages. (2) Implicit termination: when one communicator transmits a message, the other knows when it ends, and when the last message ends, both communicators know that communication has ended. (3) Correct decision: when communication ends, the recipient, $P_{\mathcal{Y}}$, knows $X$.

For every *input* — a possible value assignment for $X$ and $Y$ — the protocol determines a finite sequence of transmitted messages. The protocol is *m-message* if, for all inputs, the number of messages transmitted is at most $m$. The *worst-case complexity* of the protocol is the number of bits it requires both communicators to transmit, maximized over all inputs. $\hat{C}_m(X|Y)$, the *m-message complexity* of a random pair $(X, Y)$, is the minimum complexity of an $m$-message protocol for $(X, Y)$. It is the minimum worst-case number of bits transmitted by both communicators using a protocol that never exchanges more than $m$ messages. A precise definition of complexities is given in Section 2. Here we illustrate them in the next paragraph and the ensuing example.

$\hat{C}_1(X|Y)$, the *one-way complexity* of $(X, Y)$, is the number of bits required in the worst-case when $P_{\mathcal{Y}}$ cannot transmit to $P_{\mathcal{X}}$, that is, when no *interaction* is allowed. $\hat{C}_2(X|Y)$ is the number of bits required in the worst-case when at most two messages are permitted: $P_{\mathcal{Y}}$ transmits a message reflecting $Y$, then $P_{\mathcal{X}}$ responds with a message from which $P_{\mathcal{Y}}$ must infer $X$. Since empty messages are allowed, $\hat{C}_m(X|Y)$ is a decreasing function of $m$ bounded below by 0. We can therefore define $\hat{C}_\infty(X|Y)$, the *unbounded-message complexity of* $(X, Y)$, to be the limit of $\hat{C}_m(X|Y)$ as $m \to \infty$. It is the *minimum* number of bits that must be transmitted for $P_{\mathcal{Y}}$ to know $X$, even if no restrictions are placed on the number of messages exchanged. A protocol for $(X, Y)$ whose worst-case complexity is $\hat{C}_\infty(X|Y)$ is an *optimal* protocol for $(X, Y)$. In summary, for all random pairs $(X, Y)$:

$$\hat{C}_1(X|Y) \geq \hat{C}_2(X|Y) \geq \hat{C}_3(X|Y) \geq \cdots \geq \hat{C}_\infty(X|Y) .$$

The following example illustrates some of these concepts.

**Example 1** [1] A league has $t$ teams named $1, \ldots, t$. Every week two random teams play each other. The outcome of the match is announced over the radio. All announcements have the same format: "The match between team $I$ and team $J$ was won by team $K$" where $1 \leq I < J \leq t$ and $K$ is either $I$ or $J$.

One day, while $P_\mathcal{Y}$ listens to a match announcement, $P_\mathcal{X}$ grabs his radio. Consequently, $P_\mathcal{Y}$ hears the first part of the announcement: "The match between team $I$ and team $J$" and $P_\mathcal{X}$ hears the second part: "was won by team $K$." $P_\mathcal{X}$ and $P_\mathcal{Y}$ agree that $P_\mathcal{Y}$ should know the winner (for it is his radio). They are looking for the most efficient way to do it.

If no feedback is allowed, $P_\mathcal{X}$ has to send a single message enabling $P_\mathcal{Y}$ to uniquely determine the winner. This message is based solely on the winner (for that is all $P_\mathcal{X}$ knows). Let $\sigma(i)$ be the message sent by $P_\mathcal{X}$ when he hears "was won by team $i$." If the messages $\sigma(i)$ and $\sigma(j)$ are the same for $i \neq j$, then in the event of a match between teams $i$ and $j$, $P_\mathcal{Y}$ cannot tell who the winner is. Also, if $\sigma(i)$ is a prefix of $\sigma(j)$ for $i \neq j$ then in the event of a match between $i$ and $j$ $P_\mathcal{Y}$ does not know when the message ends. Therefore, the messages $\sigma(1), \ldots, \sigma(n)$ must all be different and none can be a proper prefix of another. Hence the worst-case one-way communication is at least $\lceil \log t \rceil$ bits. This lower bound is clearly achievable as the one-way protocol instructing $P_\mathcal{X}$ to transmit the binary representation of $K$ requires $\lceil \log t \rceil$ bits in the worst case. In complexity terms, we say that the one-message complexity of this problem is $\lceil \log t \rceil$.

Communication is very different once feedback is allowed. We describe a two-message protocol that requires only $\lceil \log \log t \rceil + 1$ bits in the worst case. $P_\mathcal{Y}$ considers the binary representations $I_1, \ldots, I_{\lceil \log t \rceil}$ and $J_1, \ldots, J_{\lceil \log t \rceil}$ of $I$ and $J$. Since $I \neq J$, there must be a first bit location $L$ where the binary representations differ: $I_L \neq J_L$. Using $\lceil \log \log t \rceil$ bits, $P_\mathcal{Y}$ transmits $L$. $P_\mathcal{X}$ responds by transmitting $K_L$ – the $L$'th bit in the binary representation of the winning team $K$. The total number of bits exchanged is $\lceil \log \log t \rceil + 1$ bits.

It was shown in [1] that $\lceil \log \log t \rceil + 1$ bits are necessary with two or any number of messages. Therefore, for all $m \in \{2, \ldots, \infty\}$, the $m$-message complexity of this problem is $\lceil \log \log t \rceil + 1$. □

The example shows that for some $(X, Y)$ pairs, one-message complexity is exponentially higher than minimum:

$$\hat{C}_1(X|Y) = 2^{\hat{C}_\infty(X|Y)-1} ,$$

yet, at least in the example, just two messages suffice to achieve minimum complexity:

$$\hat{C}_2(X|Y) = \hat{C}_\infty(X|Y) . \tag{1}$$

These two equations were shown in [1] to hold for a large class of $(X, Y)$ pairs. The same paper also showed that two messages always suffice to reduce communication to almost the minimum: for all $(X, Y)$ pairs,

$$\hat{C}_2(X|Y) \leq 4\hat{C}_\infty(X|Y) + 3 . \tag{2}$$

This contrasts with communication complexity where Papadimitriou and Sipser [3] and Duris, Galil, and Schnitger [4] showed that for every $m$ there is a function whose $(m-1)$-message complexity is almost exponentially higher than its $m$-message complexity.

However, the above does not indicate whether two messages are always optimal in conveying information. Namely, whether Equation (1) holds for all $(X, Y)$ pairs, or, Equation (2) is more representative of some $(X, Y)$ pairs for which three or more messages can reduce worst-case communication over two messages.

The purpose of this paper is twofold. We define the *chromatic-decomposition number* of a hypergraph and show that, under general conditions, it determines the two message complexity. We then use this result to prove that two messages are not optimal. For some $(X, Y)$ pairs $\hat{C}_2(X|Y)$ is nearly twice larger than $\hat{C}_\infty(X|Y)$.

In the next section we define protocols, complexities, and $G(X|Y)$, the *characteristic hypergraph* of $(X, Y)$. We mention several previously known results that are needed here including the relation between the one-way complexity and the chromatic number, $\mathcal{G}$, of the characteristic hypergraph: for all $(X, Y)$ pairs,

$$\hat{C}_1(X|Y) = \lceil \log \mathcal{G} \rceil . \tag{3}$$

In Section 3 we describe the *playoffs problem.* Although similar in appearance to the league problem of Example 1, we show in subsequent sections that its two-message complexity is about twice as high as its three-message complexity.

To do so, we need to characterize two-message complexity. In Section 4 we define $\mathcal{D}(G)$, the *chromatic-decomposition number* of a hypergraph $G$. We show that under general conditions,

$$\lceil \log \mathcal{D}(G(X|Y)) \rceil \leq \hat{C}_2(X|Y) \leq \lceil \log \mathcal{D}(G(X|Y)) \rceil + 1 \ . \tag{4}$$

This inequality is the analog of Equation (3) for two-message complexities. It can be used to determine the two-message complexity of certain $(X, Y)$ pairs. Here we use it to lower bound two-message complexities.

In Section 5, we prove a high lower bound on the chromatic-decomposition number of the playoffs problem's characteristic hypergraph. Inequality (4) then shows that the problem has a high two-message complexity. We thereby obtain a family of random pairs (one for every choice of parameters in the playoffs problem) such that for every positive $\epsilon$ and $c$, there is a pair $(X, Y)$ in the family for which

$$\begin{aligned} \hat{C}_2(X|Y) &\geq (2-\epsilon)\hat{C}_3(X|Y) \\ &\geq (2-\epsilon)\hat{C}_\infty(X|Y) \\ &\geq (2-\epsilon)c \ . \end{aligned} \tag{5}$$

Therefore, $\hat{C}_2(X|Y)$ and $\hat{C}_\infty(X|Y)$ can differ: for some $(X, Y)$ pairs allowing more than two messages may decrease the number of transmitted bits by a factor of two. In Section 6, we apply Inequality (4) to show that $\hat{C}_2(X|Y)$ is about twice $\hat{C}_\infty(X|Y)$ for yet another problem, named *duelists and ghouls.*

Having shown that the chromatic-decomposition number is related to the two-message complexity, in Section 7 we describe a technique that improves the lower bound for the chromatic-decomposition number of the playoffs problem. However, this improved lower bound does not suffice to increase the provable gap between two and three message complexities.

In view of Inequalities (2) and (5), the largest possible ratio between $\hat{C}_2(X|Y)$ and $\hat{C}_\infty(X|Y)$ is between 2 and 4. Can this gap be reduced? Also, we see that two messages are not necessarily optimal. But, are three? is there an $M$ such that for all $(X, Y)$ pairs $\hat{C}_M(X|Y) \leq \hat{C}_\infty(X|Y) + o(\hat{C}_\infty(X|Y))$? In Section 8 we discuss these and other open problems related to the gap between the different complexity measures.

## 2 Definitions and Previous Results

First, we formally define the *complexity measures.* The *support set* of a random pair $(X,Y)$ with underlying probability distribution $p(x,y)$ is the set

$$S_{X,Y} \stackrel{\text{def}}{=} \{(x,y) : p(x,y) > 0\}$$

of ordered pairs occurring with positive probability. An *input* is an element of $S_{X,Y}$ viewed as value assignments to $X$ and $Y$. As mentioned in the introduction, for every input $(x,y)$, the protocol used, $\phi$, determines a finite sequence $\sigma_1(x,y),\ldots,\sigma_{m_\phi(x,y)}(x,y)$ of transmitted messages.[1] Since $P_{\mathcal{X}}$ and $P_{\mathcal{Y}}$ alternate in transmitting these messages, all even numbered messages in this sequence are transmitted by one communicator and all odd numbered messages are transmitted by the other. $m_\phi(x,y)$ is the number of messages exchanged for $(x,y)$. The protocol $\phi$ is $m$*-message* if $m_\phi(x,y) \le m$ for all $(x,y) \in S_{X,Y}$. The *length*, $|\sigma|$, of a message $\sigma$ is the number of bits it contains. The *transmission length* of the input $(x,y)$ is

$$l_\phi(x,y) \stackrel{\text{def}}{=} \sum_{i=1}^{m_\phi(x,y)} |\sigma_i(x,y)| \ ,$$

the number of bits transmitted by both communicators when $X = x$, $Y = y$, and the protocol used is $\phi$.

The *worst-case complexity of* $\phi$ is the supremum, over all inputs, of the number of bits transmitted during the execution of $\phi$:

$$\hat{L}_\phi \stackrel{\text{def}}{=} \sup\{l_\phi(x,y) : (x,y) \in S_{X,Y}\} \ .$$

The $m$*-message complexity* of $(X,Y)$ (i.e., of transmitting $X$ to a person who knows $Y$) is

$$\hat{C}_m(X|Y) \stackrel{\text{def}}{=} \inf\{\hat{L}_\phi : \phi \text{ is an } m\text{-message protocol for } (X,Y)\} \ ,$$

the number of bits that the two communicators must transmit in the worst case if restricted to protocols that never require more than $m$ messages. Since empty messages are allowed, $\hat{C}_m(X|Y)$ is a decreasing function of $m$ bounded below by 0. We can therefore define $\hat{C}_\infty(X|Y)$, the *unbounded-message complexity of* $(X,Y)$, to be the limit of $\hat{C}_m(X|Y)$ as $m \to \infty$. It is the *minimum* number of bits that must be transmitted in the worst case for

[1] For notational simplicity, the protocol $\phi$ is implicit in $\sigma_i(x,y)$. Also, we do not assume that the same number of messages is transmitted for all inputs or that the same communicator transmits the first message for all inputs.

$P_{\mathcal{Y}}$ to know $X$. A protocol for $(X,Y)$ whose worst-case complexity is $\hat{C}_\infty(X|Y)$ is an *optimal* protocol for $(X,Y)$. In summary, for all random pairs $(X,Y)$:

$$\hat{C}_1(X|Y) \geq \hat{C}_2(X|Y) \geq \hat{C}_3(X|Y) \geq \cdots \geq \hat{C}_\infty(X|Y) .$$

Several results concerning the possible discrepancy between these quantities were proven in [1]. Some are used in this paper and we repeat them here. The first result shows that since we are interested in worst-case complexities, and do not allow any errors, the precise values of the probability distribution underlying $(X,Y)$ are irrelevant to $\hat{C}_m(X|Y)$.

**Result 1** (Corollary 1 in [1]) For all $(X,Y)$ pairs and all $m \in \{2,\ldots,\infty\}$, the $m$-message complexity, $\hat{C}_m(X|Y)$, is determined by $S_{X,Y}$, the support set of $(X,Y)$. □

This explains why in Example 1 we did not specify the precise probabilities of games and winners. Let $Y_t$ denote the two teams that played and let $X_t$ denote the winner. We (implicitly) described only the support set of $(X_t,Y_t)$; since all games are possible and the winner is one of the playing teams, $S_{X_t,Y_t} = \{(k,(i,j)) : 1 \leq i < j \leq t \text{ and } k \in \{i,j\}\}$.

The simplest complexity measure to determine is $\hat{C}_1(X|Y)$, the one-way complexity of $(X,Y)$. The *support set* of $X$ is the set

$$S_{\mathcal{X}} \stackrel{\text{def}}{=} \{x : (x,y) \in S_{X,Y} \text{ for some } y\}$$

of possible values of $X$. The *support set* $S_{\mathcal{Y}}$ of $Y$ is similarly defined. Instrumental in determining $\hat{C}_1(X|Y)$ is $G(X|Y)$, the *characteristic hypergraph* of $(X,Y)$. Its vertex set is $S_{\mathcal{X}}$ and for every $y \in S_{\mathcal{Y}}$, it contains the edge: $\{x : (x,y) \in S_{X,Y}\}$. The characteristic hypergraph is equivalent to a graph defined by Witsenhausen [5] who considered the one-message version of this problem. For the league problem with $t$ teams, $G(X_t|Y_t)$ has $t$ vertices, one corresponding to each value of $X_t$. It has $\binom{t}{2}$ edges, one corresponding to each possible value of $Y_t$. Each edge contains two vertices (the two possible winning teams in this game). In other words, $G(X_t|Y_t)$ is $K_t$, the complete graph on $t$ vertices.

A convenient way of representing both $S_{X,Y}$ and $G(X|Y)$ is the *characteristic table*. We view $S_{\mathcal{X}} \times S_{\mathcal{Y}}$ as table with $|S_{\mathcal{X}}| \cdot |S_{\mathcal{Y}}|$ entries. Each column corresponds to an element of $S_{\mathcal{X}}$ and each row to an element of $S_{\mathcal{Y}}$. We mark the entry in the column corresponding to $x$ and the row corresponding to $y$ if $(x,y) \in S_{X,Y}$. Figure 1 depicts the characteristic table for the league problem with four teams.

A coloring of a hypergraph $G$ is an assignment of a color to every vertex of $G$ such that no two vertices sharing an edge are assigned the same color. The chromatic number of $G$ is the

| | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $(1,2)$ | ● | ● | | |
| $(1,3)$ | ● | | ● | |
| $(1,4)$ | ● | | | ● |
| $(2,3)$ | | ● | ● | |
| $(2,4)$ | | ● | | ● |
| $(3,4)$ | | | ● | ● |

$S_{Y_4}$ (games)

$S_{X_4}$ (teams)

Figure 1: Characteristic table of the league problem with four teams.

minimum number of colors required to color $G$. Witsenhausen [5] showed that in one-way communication, the minimum number of different messages sent by $P_{\mathcal{X}}$ is $\mathcal{G}$. Although the messages transmitted do not have to be prefix-free (by the implicit-termination property they have to be prefix free only given the value of $X$), it can be shown that:

**Result 2** (Lemma 2 in [1]) For all $(X, Y)$ pairs,

$$\hat{C}_1(X|Y) = \lceil \log \mathcal{G} \rceil \ . \qquad \square$$

This proves, again, that for the league problem with $t$ teams $\hat{C}_1(X_t|Y_t) = \lceil \log \chi(K_t) \rceil = \lceil \log t \rceil$ as was shown in Example 1.

We saw that the two-message complexity of the league problem with $t$ teams satisfies: $\hat{C}_2(X_t|Y_t) \leq \lceil \log \log t \rceil + 1$. Therefore,

$$\hat{C}_2(X_t|Y_t) \leq \lceil \log \hat{C}_1(X_t|Y_t) \rceil + 1 \ .$$

This "almost-logarithmic" reduction in communication is the largest achievable via interaction:

**Result 3** (Theorem 2 in [1]) For all $(X, Y)$ pairs,

$$\hat{C}_\infty(X|Y) \geq \lceil \log \hat{C}_1(X|Y) \rceil + 1 \ . \qquad \square$$

Therefore, the two-message protocol described in Example 1 is optimal for the league problem; no other protocol requires fewer bits in the worst case. The "almost-logarithmic" reduction is typical of a large class of random pairs. $P_{\mathcal{Y}}$'s *ambiguity* when he knows the value $y$ is

$$\mu_{X|Y}(y) \stackrel{\text{def}}{=} |\{x : (x,y) \in S_{X,Y}\}| \ ,$$

the number of possible $X$ values when $Y = y$. The *maximum ambiguity* of the pair $(X, Y)$ is

$$\hat{\mu}_{X|Y} \stackrel{\text{def}}{=} \sup\{\mu_{X|Y}(y) : y \in S_{\mathcal{Y}}\},$$

the maximum ambiguity $P_{\mathcal{Y}}$ can ever have over the value of $X$. In the league problem, the maximum ambiguity is two: for every value known to $P_{\mathcal{Y}}$ (a game), $P_{\mathcal{X}}$ can have two possible values (winning teams). When $\hat{\mu}_{X|Y} = 2$, the characteristic hypergraph is a graph. For all random pairs $(X, Y)$ with maximum ambiguity two, the bound of Result 3 is tight:

**Result 4** (Theorem 3 in [1]) For all $(X, Y)$ pairs with $\hat{\mu}_{X|Y} = 2$,

$$\hat{C}_\infty(X|Y) = \lceil \log \hat{C}_1(X|Y) \rceil + 1 \ . \qquad \Box$$

Stated differently, for all random pairs whose maximum ambiguity is 2,

$$\hat{C}_1(X|Y) \geq 2^{\hat{C}_\infty(X|Y)-2}$$

and a similar inequality holds for many other $(X, Y)$ pairs. Therefore, one-message complexity can be exponentially higher than the minimum. Perhaps the most interesting result in [1] is that just two messages suffice to reduce communication to almost the minimum[2]:

**Result 5** (Corollary 5 in [1]) For all $(X, Y)$ pairs,

$$\hat{C}_2(X|Y) \leq 4\hat{C}_\infty(X|Y) + 3 \ . \qquad \Box$$

If the maximum ambiguity is two, a stronger result holds. Two messages suffice to achieve minimum communication:

**Result 6** (Corollary 4 in [1]) For all random pairs $(X, Y)$ with $\hat{\mu}_{X|Y} = 2$,

$$\hat{C}_2(X|Y) = \ldots = \hat{C}_\infty(X|Y) = \lceil \log \hat{C}_1(X|Y) \rceil + 1 \ . \qquad \Box$$

[2] This contrasts with communication complexity where for every $m$ there is a function with low $m$-message complexity but almost exponentially higher $(m-1)$-message complexity.

This was the case for the league problem where the two-message protocol described was optimal.

The most obvious question left open in [1] was whether two messages are always optimal. Namely, does Result 6 holds in general and $\hat{C}_2(X|Y) = \hat{C}_\infty(X|Y)$ for all $(X, Y)$ pairs? Or, are there $(X, Y)$ pairs for which three or more messages can reduce worst-case communication over two messages?

# 3 Playoffs

Consider a sport played by $\ell t$ teams. The teams are divided among $\ell$ leagues consisting of $t$ teams each. The playoffs for this sport consist of $2\ell$ teams, two from each league. Normally, the playoffs end with one of the $2\ell$ teams declared as a winner. However, at times, the playoffs are canceled (due to bad weather, say) and then there is "no winner." (The reason for including the "no winner" possibility will become clear later.)

Similar to the league problem, a person $P_{\mathcal{Y}}$ knows the set of $2\ell$ teams scheduled to play in the playoffs. Another person, $P_{\mathcal{X}}$, knows the outcome of the playoffs: whether they took place, and if so, who the winner was. This knowledge has to be conveyed to $P_{\mathcal{Y}}$.

We now cast the problem in terms of random variables and graphs. $Y_{\ell,t}$ is a playoffs: a set of $2\ell$ teams, two from each league. $X_{\ell,t}$ is either a winner – a team in the playoffs $Y_{\ell,t}$ – or "no winner." $P_{\mathcal{X}}$ knows $X_{\ell,t}$, $P_{\mathcal{Y}}$ knows $Y_{\ell,t}$ and they both want $P_{\mathcal{Y}}$ to know $X_{\ell,t}$. The characteristic hypergraph for this problem, $G(X_{\ell,t}|Y_{\ell,t})$, consists of $\ell t + 1$ vertices, one corresponding to each possible value of $X_{\ell,t}$. $\ell t$ vertices correspond to the possible winning teams, and one corresponds to "no winner." There are $\binom{t}{2}^\ell$ edges in $G(X_{\ell,t}|Y_{\ell,t})$, each corresponds to a possible value of $Y_{\ell,t}$, hence contains $2\ell + 1$ vertices: the $2\ell$ teams in the playoffs, and "no winner." Figure 2 depicts $G(X_{\ell,t}|Y_{\ell,t})$ for two leagues ($\ell = 2$) each containing three teams ($t = 3$). Note that the last column is full as the "no winner" vertex belongs to all edges.

In a coloring of $G(X_{\ell,t}|Y_{\ell,t})$, no two vertices can be assigned the same color: any two teams can participate in a playoffs, thus any two "team vertices" belong to a common edge and the "no winner" vertex belongs to all edges. Hence

$$\chi(G(X_{\ell,t}|Y_{\ell,t})) = \ell t + 1$$

and, from Result 2,

$$\hat{C}_1(X_{\ell,t}|Y_{\ell,t}) = \lceil \log(\ell t + 1) \rceil.$$

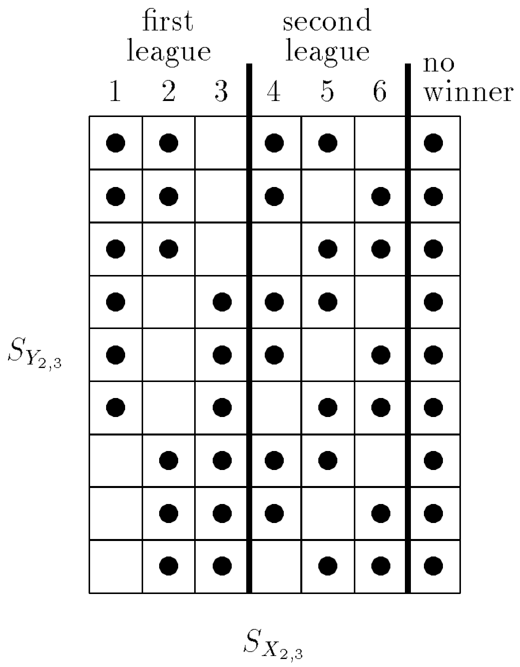


Figure 2: Characteristic table for two leagues, each consisting of three teams.

The problem lends itself to a simple three-message protocol whose worst-case complexity is $\lceil \log(\ell+1) \rceil + \lceil \log\log t \rceil + 1$. First, $P_{\mathcal{X}}$ transmits $\lceil \log(\ell+1) \rceil$ bits notifying $P_{\mathcal{Y}}$ whether the playoffs took place and, if so, which was the winner's league. If there was "no winner," communication stops. Otherwise, $P_{\mathcal{Y}}$ knows the winner's league and two teams in this league, and $P_{\mathcal{X}}$ knows one of them (the winner). They proceed according to the two-message protocol described in Example 1, transmitting at most $\lceil \log\log t \rceil + 1$ additional bits. Hence,

$$\hat{C}_3(X_{\ell,t}|Y_{\ell,t}) \le \lceil \log(\ell+1) \rceil + \lceil \log\log t \rceil + 1.$$

Result 3 can be used to prove a lower bound on $\hat{C}_\infty(X_{\ell,t}|Y_{\ell,t})$ and therefore to lower bound $\hat{C}_3(X_{\ell,t}|Y_{\ell,t})$. However this is not our goal here. In the rest of the paper we show that

$$\hat{C}_2(X_{\ell,t}|Y_{\ell,t}) \ge \log \min\{t, (\ell+1)(2\ell+1)\}. \tag{6}$$

Relating the number of leagues to the number of teams in each league by:

$$t = (\ell+1)(2\ell+1),$$

we derive

$$\hat{C}_3(X_{\ell,(\ell+1)(2\ell+1)}|Y_{\ell,(\ell+1)(2\ell+1)}) \le \log \ell + \log\log \ell + 4$$

while

$$\hat{C}_2(X_{\ell,(\ell+1)(2\ell+1)}|Y_{\ell,(\ell+1)(2\ell+1)}) \geq 2\log\ell + 1.$$

For large $\ell$, the two-message complexity is almost twice the tree-message complexity. Specifically, two messages are not optimal for conveying information: for all $c \in \mathcal{R}$ and all $\epsilon > 0$ there is a random pair $(X,Y)$ such that

$$\hat{C}_2(X|Y) \geq (2-\epsilon)\hat{C}_3(X|Y) \geq c. \tag{7}$$

To prove this discrepancy, we need to lower-bound two-message complexity. Result 2 stated that the one-message complexity of $(X,Y)$ is determined by the chromatic number of its characteristic hypergraph:

$$\hat{C}_1(X|Y) = \lceil \log \mathcal{G} \rceil.$$

In the next section we show that a related measure can help determine two-message complexity.

# 4 Chromatic-decomposition number and 2-message complexity

Let $G = (V,E)$ be a hypergraph. The hypergraphs $G_1 = (V,E_1),\ldots,G_k = (V,E_k)$ *edge cover* $G$ if every edge in $G$ belongs to at least one $G_i$. Namely,

$$E \subseteq \bigcup_{i=1}^{k} E_i\ .$$

The *chromatic-decomposition number* of $G$ is

$$\mathcal{D}(G) \stackrel{\text{def}}{=} \min\left\{\sum_{i=1}^{k}\chi(G_i) : G_1,\ldots,G_k \text{ edge cover } G\right\},$$

the minimum sum of the individual chromatic numbers in any edge cover of $G$. In this section we show that under some general conditions,

$$\lceil \log \mathcal{D}(G(X|Y)) \rceil \leq \hat{C}_2(X|Y) \leq \lceil \log \mathcal{D}(G(X|Y)) \rceil + 1. \tag{8}$$

But, first, we illustrate chromatic-decomposition numbers by proving an upper bound for graphs (not general hypergraphs).

**Example 2** If $G$ is a graph, then

$$\mathcal{D}(G) \leq 3\lceil \log_3 \chi(G) \rceil \ .$$

**Proof:** Let $V$ and $E$ denote the vertex and edge sets of $G$ and let $d \stackrel{\text{def}}{=} \lceil \log_3 \chi(G) \rceil$. Fix a coloring $\psi : V \to \{0,1,2\}^d$ of $G$. Edge cover $G$ with $d$ graphs $G_1, \ldots, G_d$ where, for $i \in \{1, \ldots, d\}$, the graph $G_i$ has $V$ as its vertex set and it contains the edge $\{v', v''\}$ if and only if:

1. $\{v', v''\}$ is an edge of $G$
2. the $i$th digit of $\psi(v')$ differs from the $i$th digit of $\psi(v'')$.

Since $\psi$ is a coloring of $G$, every edge of $G$ belongs to some $G_i$ so $G_1, \ldots, G_d$ edge cover $G$. Also, for $i \in \{1, \ldots, d\}$, the $i$th digit of $\psi$ is a coloring of $G_i$ so $\chi(G_i) \leq 3$. Therefore,

$$\mathcal{D}(G) \leq \sum_{i=1}^{d} \chi(G_i) \leq 3d = 3\lceil \log_3 \chi(G) \rceil \ .$$

This edge cover is illustrated in Figure 3 for a graph with chromatic number 5. □

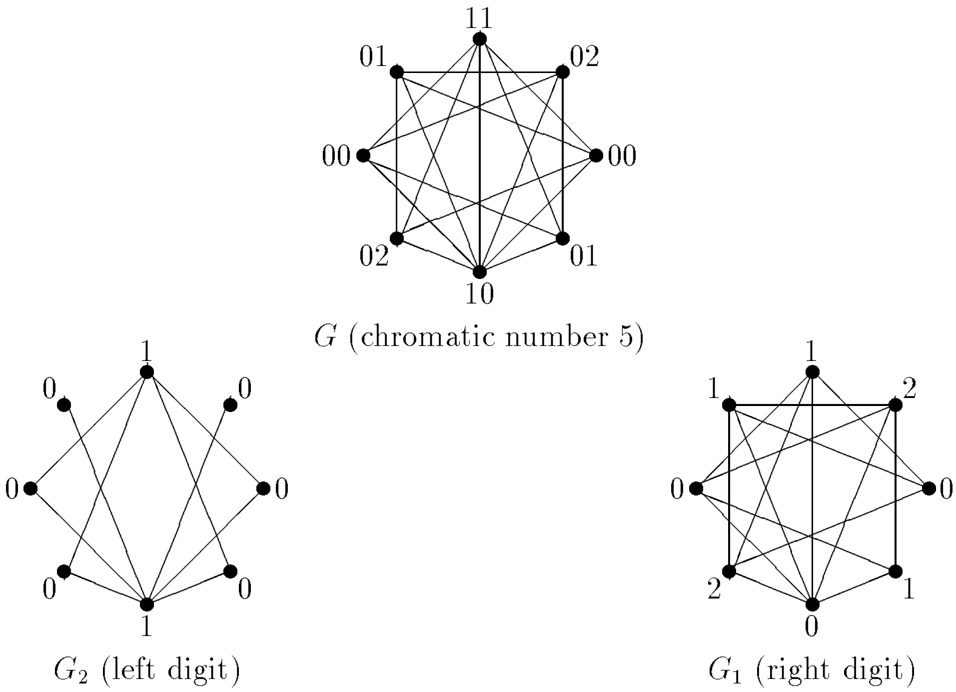


Figure 3: Edge cover derived in Example 2.

It can be shown that when the chromatic number of a graph $G$ is a power of 3, the upper bound just derived is tight:

$$\mathcal{D}(G) = 3 \log_3 \chi(G) \tag{9}$$

However, this digression on graphs is meant only to illustrate chromatic-decomposition numbers, hence we postpone the proof of Equation (9) till the next section (Corollary 1). Our main goal in this section is to prove Inequality (8), relating two-message complexity and chromatic-decomposition numbers. The upper bound always holds:

**Lemma 1** For all $(X, Y)$ pairs,

$$\hat{C}_2(X|Y) \le \lceil \log \mathcal{D}(G(X|Y)) \rceil + 1. \tag{10}$$

**Proof:** Given an edge cover $G_1, \ldots, G_k$ of $G(X|Y)$ with chromatic numbers $\chi_1, \ldots, \chi_k$, we construct a two-message protocol whose worst-case complexity is at most $\lceil \log \sum_{j=1}^k \chi_j \rceil + 1$. Since

$$\sum_{i=1}^k 2^{-\lceil \log(\sum_{j=1}^k \chi_j) - \log \chi_i \rceil} \le \frac{\sum_{i=1}^k \chi_i}{\sum_{j=1}^k \chi_j} = 1,$$

we can, by Kraft's inequality, assign to each $G_i$ a message of length $\lceil \log(\sum_{j=1}^k \chi_j) - \log \chi_i \rceil$ such that no message prefixes another. The protocol then starts with $P_\mathcal{Y}$ transmitting the message associated with any[3] subgraph $G_I$ containing the edge corresponding to $Y$. Following this transmission, $P_\mathcal{X}$ and $P_\mathcal{Y}$ confine their consideration to $G_I$ and use the one-message protocol of Result 2 in which $P_\mathcal{X}$ transmits $\lceil \log \chi_I \rceil$ bits to describe $X$. The total number of transmitted bits is:

$$\left\lceil \log \left( \sum_{j=1}^k \chi_j \right) - \log \chi_I \right\rceil + \lceil \log \chi_I \rceil \le \left\lceil \log \sum_{j=1}^k \chi_j \right\rceil + 1 \ . \tag{11}$$

Taking an edge cover achieving $\mathcal{D}(G(X|Y))$, we prove the lemma. □

It is interesting to compare Example 2 and Lemma 1 with the communication required for the league problem with $t$ teams. Example 1 showed that

$$\hat{C}_2(X_t|Y_t) = \lceil \log \log t \rceil + 1 \ .$$

The characteristic hypergraph for this problem is $K_t$, the complete graph on $t$ vertices. Hence, from Example 2,

$$\mathcal{D}(G(X_t|Y_t)) \le 3\lceil \log_3 t \rceil$$

[3]The subgraphs $G_1, \ldots, G_k$ form an edge cover of $G$ hence do not have to be disjoint.

and, from the left-hand side of Inequality (11),

$$\hat{C}_2(X_t|Y_t) \leq \lceil \log \log_3 t \rceil + 2 \ .$$

If $t = 3^{2^s}$ for some integer $s$, then

$$\lceil \log \log_3 t \rceil + 2 = \lceil \log \log t \rceil + 1 \ .$$

Hence for these values of $t$, one can use either the two-message protocol of Example 1, or the similar two-message protocol suggested by Example 2 and Lemma 1 which is based on the ternary representation of the teams. The two protocols have the same worst-case complexity.

We now prove the lower bound of Inequality (8) for a restricted class of $(X, Y)$ pairs.

**Lemma 2** If for all $y_1, y_2 \in S_{\mathcal{Y}}$ there exists an $x \in S_{\mathcal{X}}$ such that $(x, y_1), (x, y_2) \in S_{X,Y}$ then

$$\hat{C}_2(X|Y) \geq \lceil \log \mathcal{D}(G(X|Y)) \rceil.$$

**Proof:** Let $\phi$ be a two-message protocol for $(X, Y)$. Consider the set $\{m_1, \ldots, m_k\}$ of all possible messages transmitted by $P_{\mathcal{Y}}$ before any bits are transmitted by $P_{\mathcal{X}}$ (if for some input $P_{\mathcal{X}}$ transmits first, one of these messages is empty). From the separate transmissions property, the first message transmitted by $P_{\mathcal{Y}}$ is a function of $Y$ alone, hence the set of messages induces an edge partition $G_1, \ldots, G_k$ of $G(X|Y)$, where the edges of $G_i$ correspond to all values of $Y$ for which $P_{\mathcal{Y}}$ transmit $m_i$. If $i$ and $j$ are distinct indices in $\{1, \ldots, k\}$ then there are distinct $y_i$ and $y_j$ for which $P_{\mathcal{Y}}$ transmits $m_i$ and $m_j$ respectively. But then, by the assumption of the lemma, there is an $x$ such that both $(x, y_i)$ and $(x, y_j)$ are in $S_{X,Y}$. From the implicit termination property, neither $m_i$ nor $m_j$ can be a prefix of the other. Hence the set $\{m_1, \ldots, m_k\}$ is prefix free. Letting $l_i$ denote the length in bits of $m_i$ and using Kraft's inequality, we obtain:

$$\sum_{i=1}^{k} 2^{-l_i} \leq 1. \tag{12}$$

We assumed that $\phi$ is a two-message protocol, hence the next message, transmitted by $P_{\mathcal{X}}$, should inform $P_{\mathcal{Y}}$ of the value of $X$. From Result 2, if $P_{\mathcal{Y}}$ transmits message $m_i$, $P_{\mathcal{X}}$ has to transmit $\lceil \log \chi(G_i) \rceil$ bits in the worst case. Hence, for all $i \in \{1, \ldots, k\}$,

$$\hat{L}_\phi \geq l_i + \lceil \log \chi(G_i) \rceil \geq l_i + \log \chi(G_i).$$

It follows that

$$1 \geq \sum_{i=1}^{k} 2^{-l_i} \geq \sum_{i=1}^{k} 2^{-(\hat{L}_\phi - \log \chi(G_i))} \geq \frac{\sum_{i=1}^{k} \chi(G_i)}{2^{\hat{L}_\phi}}$$

Hence,

$$\hat{L}_\phi \geq \lceil \log \mathcal{D}(G(X|Y)) \rceil.$$

This holds for every two-message protocol, hence

$$\hat{C}_2(X|Y) \geq \lceil \log \mathcal{D}(G(X|Y)) \rceil.$$

The lemma has an alternative proof that does not use Kraft's inequality. For every input, concatenate $P_\mathcal{Y}$'s message with the message transmitted by $P_\mathcal{X}$. The resulting set of strings is not necessarily prefix free. However, as in the proof of Lemma 2 in [1], it can be shown to contain a prefix-free subset of size $\mathcal{D}(G(X|Y))$. Hence at least one string is of length $\geq \lceil \log \mathcal{D}(G(X|Y)) \rceil$. Though more direct, this proof considers individual messages to demonstrate the alleged subset. The proof presented here uses only "top-level" results, hence was preferred. □

We proved the lemma for the restricted class of $(X, Y)$ pairs where every two $Y$ values, share a possible $X$ value. In the characteristic hypergraph of such a pair, every two edges share a vertex. Such hypergraphs are known as *intersecting hypergraphs*. The converse does not hold for all $(X, Y)$ pairs: there are random variables $X$ and $Y$ such that $\hat{C}_2(X|Y) \not\geq \log \mathcal{D}(G(X|Y))$. This is because, from the implicit termination property, the messages transmitted by $P_\mathcal{Y}$ must be prefix-free only given the value of $X$ and previous transmissions. They do not have to be prefix free in general, hence Inequality (12) does not necessarily hold for general $(X, Y)$ pairs. For one-way complexity, this problem can be overcome, leading to the equality of Result 2. We do not know whether a similar argument can help here (see Section 8).

# 5 Lower bound on $\mathcal{D}(G(X_{\ell,t}|Y_{\ell,t}))$

We want to show that for the playoffs problem,

$$\hat{C}_2(X_{\ell,(\ell+1)(2\ell+1)}|Y_{\ell,(\ell+1)(2\ell+1)}) \geq (2-\epsilon)\hat{C}_3(X_{\ell,(\ell+1)(2\ell+1)}|Y_{\ell,(\ell+1)(2\ell+1)})$$

hence we need to lower bound the two-message complexity of this problem. By construction, there is one $X$ value ("no games") for which all $Y$ values are possible. Hence the condition of Lemma 2 is satisfied and

$$\lceil \log \mathcal{D}(G(X_{\ell,t}|Y_{\ell,t})) \rceil + 1 \geq \hat{C}_2(X_{\ell,t}|Y_{\ell,t}) \geq \lceil \log \mathcal{D}(G(X_{\ell,t}|Y_{\ell,t})) \rceil.$$

We therefore show that, $\mathcal{D}(G(X_{\ell,t}|Y_{\ell,t}))$, the chromatic-decomposition number of this problem, is high.

**Lemma 3** For all $\ell$ and $t$,

$$\mathcal{D}(G(X_{\ell,t}|Y_{\ell,t})) \geq \min\{t, (\ell+1)(2\ell+1)\} \ .$$

**Proof:** We show that $\mathcal{D}(G(X_{\ell,t}|Y_{\ell,t})) < t$ implies that $\mathcal{D}(G(X_{\ell,t}|Y_{\ell,t})) \geq (\ell+1)(2\ell+1)$. Let $G_1,\ldots,G_k$ be an edge partition of $G(X_{\ell,t}|Y_{\ell,t})$ with chromatic numbers $\chi_1,\ldots,\chi_k$ achieving $\mathcal{D}(G(X_{\ell,t}|Y_{\ell,t}))$. That is,

$$\sum_{i=1}^{k} \chi_i = \mathcal{D}(G(X_{\ell,t}|Y_{\ell,t})).$$

Without loss of generality, assume that no $G_i$ is empty. Then, as each edge of $G(X_{\ell,t}|Y_{\ell,t})$ contains $2\ell+1$ vertices, $\chi_i \geq 2\ell+1$ for all $i \in \{1,\ldots,k\}$. We prove that $\mathcal{D}(G(X_{\ell,t}|Y_{\ell,t})) < t$ implies that $k \geq \ell+1$, and therefore

$$\mathcal{D}(G(X_{\ell,t}|Y_{\ell,t})) = \sum_{i=1}^{k} \chi_i \geq (\ell+1)(2\ell+1).$$

If $\mathcal{D}(G(X_{\ell,t}|Y_{\ell,t})) < t$, then $\chi_i < t$ for all $i \in \{1,\ldots,k\}$. In particular $\chi_1 < t$ so there must be two vertices, $v_1'$ and $v_1''$, in the first league that do not belong together to any edge of $G_1$. Therefore, all edges of $G$ containing both $v_1'$ and $v_1''$ belong to $\bigcup_{i=2}^{k} G_i$. Similarly, $\chi_2 < t$, hence there must be two vertices $v_2'$ and $v_2''$ in the second league that do not belong together to any edge in $G_2$. Therefore, all edges containing the four vertices $v_1', v_1'', v_2'$, and $v_2''$ cannot belong to either $G_1$ or $G_2$ hence must belong to $\bigcup_{i=3}^{k} G_i$.

Continuing this way, we can find vertices $v_1',\ldots,v_\ell'$ and $v_1'',\ldots,v_\ell''$ such that all edges containing these $2\ell$ vertices belong to $\bigcup_{i=\ell+1}^{k} G_i$. But such an edge exists, namely, the one corresponding to the playoffs containing teams $v_i'$ and $v_i''$ from league $i$ for $i = 1,\ldots,\ell$. Hence $\bigcup_{i=\ell+1}^{k} G_i$ is not empty, so $k \geq \ell+1$. □

Combining Lemmas 2 and 3 we obtain

**Theorem 1** For the league problem with $\ell$ leagues and $t$ teams,

$$\hat{C}_2(X_{\ell,t}|Y_{\ell,t}) \geq \log\min\{t, (\ell+1)(2\ell+1)\} \ . \qquad \square$$

This proves Inequality (6) and therefore the factor of two discrepancy between $\hat{C}_2(X_{\ell,t}|Y_{\ell,t})$ and $\hat{C}_3(X_{\ell,t}|Y_{\ell,t})$ asserted by Inequality (7).

# 6 Duelists and ghouls

We use Lemma 2 to prove that $\hat{C}_2(X|Y)$ is almost twice $\hat{C}_\infty(X|Y)$ for yet another family of random pairs. This construction was derived from an idea suggested by Noga Alon in a discussion aimed at proving that $\hat{C}_3(X|Y) > \hat{C}_\infty(X|Y)$.

Let $n$ be an integer. A town has $n$ duelists who frequently engage in pistol duels. There are also $2\log n$ ghouls[4] in town, each attending every duel. Unfortunately, the duelists are not trained in marksmenship so their shots often hit a ghoul. Each duel therefore has a *victim*: one of the two participating duelists, or one of the $2\log n$ ghouls. $P_\mathcal{Y}$ knows the the two duelists in a duel and $P_\mathcal{X}$ knows the victim. Both $P_\mathcal{X}$ and $P_\mathcal{Y}$ want $P_\mathcal{Y}$ to learn who the victim is.

Formally, the set $S_{X_n}$ consists of $n$ duelists and $2\log n$ ghouls. $Y_n$ is one of the $\binom{n}{2}$ subsets of $S_{X_n}$ containing two duelists and all $2\log n$ ghouls. $X_n$, the victim, is one of the $2\log n + 2$ elements of $Y_n$. Figure 4 depicts the characteristic table of the problem for $n = 4$.

| | 1 | 2 | 3 | 4 | A | B | C | D |
|---|---|---|---|---|---|---|---|---|
| (1, 2) | ● | ● | | | ● | ● | ● | ● |
| (1, 3) | ● | | ● | | ● | ● | ● | ● |
| (1, 4) | ● | | | ● | ● | ● | ● | ● |
| (2, 3) | | ● | ● | | ● | ● | ● | ● |
| (2, 4) | | ● | | ● | ● | ● | ● | ● |
| (3, 4) | | | ● | ● | ● | ● | ● | ● |

$S_{Y_4}$ (duels)

$S_{X_4}$ (duelists and ghouls)

Figure 4: Characteristic table for four duelists: 1, 2, 3, 4 and four ghouls: A, B, C, D.

Every two vertices in the characteristic hypergraph, $G(X_n|Y_n)$, belong to a common edge, hence

$$\chi(G(X_n|Y_n)) = n + 2\log n \ . \tag{13}$$

Result 2 therefore says that,

$$\hat{C}_1(X_n|Y_n) = \log n + 1. \tag{14}$$

[4]For simplicity assume that $n \in 2^{2^{\mathbb{Z}}}$, hence $\log n$ and $\log\log n$ are integers.

The problem has a simple three message protocol showing that for $n \geq 2$

$$\hat{C}_3(X_n|Y_n) \leq \log\log n + 2.$$

$P_{\mathcal{X}}$ transmits a single bit identifying the victim as either a duelist or a ghoul. If the victim is a ghoul, $P_{\mathcal{X}}$ transmits $\log\log n + 1$ additional bits, completely identifying him. If $X$ is a duelist, then $P_{\mathcal{Y}}$ knows two possible victims (out of $n$) and $P_{\mathcal{X}}$ knows one of them. They use the two-message league protocol of Example 1 to convey the victim to $P_{\mathcal{Y}}$ while transmitting $\log\log n + 1$ bits.

Note that Result 3 and Inequality (14) imply that the above bound is tight:

$$\hat{C}_\infty(X_n|Y_n) = \ldots = \hat{C}_3(X_n|Y_n) = \log\log n + 2.$$

We now show that

**Theorem 2** For all $n \geq 4$,

$$\hat{C}_2(X_n|Y_n) \geq 2\log\log n - \log\log\log n.$$

The intuition is simple. Suppose that we are trying to devise a two-message protocol requiring few bits in the worst case, say $\log\log n + 2$ bits. If the victim is a ghoul, the informant, $P_{\mathcal{X}}$, must transmit $\log\log n + 1$ bits. If the victim is a duelist, it can be shown that the recipient, $P_{\mathcal{Y}}$, must transmit at least $\log\log n$ bits to achieve a total of $\log\log n + 2$. However, $P_{\mathcal{Y}}$ knows only who played hence cannot tell whether the victim is a duelist or a ghoul. If he transmits few bits, then if the victim is a duelist many more bits must be transmitted. If $P_{\mathcal{Y}}$ transmits many bits, then if the victim is a ghoul, these bits are wasted as $P_{\mathcal{X}}$ must still transmit $\log\log n + 1$ bits.

Formally, let $G_1, \ldots, G_k$ be an edge cover of $G(X_n|Y_n)$. Without loss of generality, none of theses subgraphs is empty. Each edge in $G(X_n|Y_n)$ contains $2\log n + 2 > \log n$ vertices, hence

$$\sum_{i=1}^{k} \chi(G_i) \geq k \log n.$$

On the other hand, it is easy to show, by induction on $k$, that if $G_1, \ldots, G_k$ edge cover $G$ then

$$\prod_{i=1}^{k} \chi(G_i) \geq \chi(G). \tag{15}$$

Combining this and Inequality (13) with the standard comparison between the geometric and arithmetic mean, we obtain:

$$\frac{\sum_{i=1}^{k}\chi(G_i)}{k} \geq \left(\prod_{i=1}^{k}\chi(G_i)\right)^{\frac{1}{k}} \geq (\chi(G(X_n|Y_n)))^{\frac{1}{k}} > n^{\frac{1}{k}}.$$

Hence,

$$\mathcal{D}(G(X_n|Y_n)) > \min_k \max\{k\log n, kn^{\frac{1}{k}}\}.$$

Every two edges of $G(X_n|Y_n)$ clearly intersect. Hence the lower bound of Lemma 2 applies and we have:

$$\hat{C}_2(X_n|Y_n) \geq \log\mathcal{D}(G(X_n|Y_n)) > \min_k \max\{\log k + \log\log n,\ \log k + \frac{1}{k}\log n\}\ .$$

For $k > \frac{\log n}{\log\log n}$,

$$\log k + \log\log n \geq 2\log\log n - \log\log\log n\ ,$$

while, for $k \leq \frac{\log n}{\log\log n}$,

$$\log k + \frac{\log n}{k} \geq 2\log\log n - \log\log\log n\ ,$$

proving the theorem. □

Note that the theorem is almost tight. The trivial protocol where $P_{\mathcal{Y}}$ transmits $\log\log n$ bits describing a bit location where the two duelists differ and $P_{\mathcal{X}}$ responds by transmitting $\log\log n + 2$ bits identifying the victim, shows that

$$\hat{C}_2(X_n|Y_n) \leq 2\log\log n + 2.$$

# 7 Improved bound on $\mathcal{D}(G(X_{\ell,t}|Y_{\ell,t}))$

In Section 4, we related the the two-message complexity of a random pair $(X, Y)$ to the chromatic-decomposition number of its characteristic hypergraph. We now refine the technique for lower bounding the chromatic-decomposition number.

A hypergraph is *r-regular* for an integer $r$ if each of its edges has size $r$. A 2-regular hypergraph is therefore a graph with no self edges. The following lemma provides a simple lower bound for the chromatic-decomposition number of $r$-regular hypergraphs when $r \neq 2$. Lemma 5 provides the corresponding bound for graphs.

**Lemma 4** If a hypergraph $G$ is $r$-regular ($r \neq 2$) then

$$\mathcal{D}(G) \geq \frac{r}{\log r} \log \chi(G) \ .$$

**Proof:** If $r = 0$ or $r = 1$ then $\mathcal{D}(G) = \chi(G) = 1$ and the lemma holds as, by convention, $\frac{0 \cdot \log 1}{\log 1} = 0$ and $\frac{1 \cdot \log 1}{\log 1} = 1$. For $r \geq 3$, let $G_1, \ldots, G_k$ be an edge cover of $G$ achieving $\mathcal{D}(G)$, namely:

$$\mathcal{D}(G) = \sum_{i=1}^{k} \chi(G_i).$$

Each $G_i$ is $r$ regular so $\chi(G_i) \geq r \geq 3$ and, trivially,

$$\chi(G_i) \geq \frac{r}{\log r} \log \chi(G_i).$$

Inequality (15) noted that if $G_1, \ldots, G_k$ edge cover $G$ then

$$\prod_{i=1}^{k} \chi(G_i) \geq \chi(G).$$

Hence,

$$\begin{aligned}
\mathcal{D}(G) &= \sum_{i=1}^{k} \chi(G_i) \\
&\geq \sum_{i=1}^{k} \frac{r}{\log r} \log \chi(G_i) \\
&= \frac{r}{\log r} \log \prod_{i=1}^{k} \chi(G_i) \\
&\geq \frac{r}{\log r} \log \chi(G).
\end{aligned}$$

$\square$

Using the same method for $r = 2$ and noting that self edges do not effect $\chi(G)$ or $\mathcal{D}(G)$we can prove:

**Lemma 5** For all graphs $G$,

$$\mathcal{D}(G) \geq \frac{3}{\log 3} \log \chi(G) \ . \qquad \square$$

For some 2-regular hypergraphs, this lower bound is tight. Example 2 showed that for all graphs $G$,

$$\mathcal{D}(G) \leq 3 \lceil \log_3 \chi(G) \rceil .$$

Therefore:

**Corollary 1** If $G$ is a graph and $\chi(G)$ is a power of 3, then

$$\mathcal{D}(G) = 3 \log_3 \chi(G) \ . \qquad \Box$$

The characteristic hypergraph of the playoffs problem is $(2\ell+1)$-regular hence the lower bound of Lemma 4 applies. However, this bound is weak yielding:

$$\mathcal{D}(G(X_{\ell,t}|Y_{\ell,t})) \geq \frac{2\ell+1}{\log(2\ell+1)} \log(\ell t + 1) \tag{16}$$

whereas Lemma 3 showed

$$\mathcal{D}(G(X_{\ell,t}|Y_{\ell,t})) \geq \min\{t, (\ell+1)(2\ell+1)\} \ , \tag{17}$$

a bound stronger than (16) for $4.1\ell \leq t \leq (2\ell)^\ell/\ell$. We now use Lemma 4 to show that for all $\alpha \leq t$,

$$\mathcal{D}(G(X_{\ell,t}|Y_{\ell,t})) \geq \min\left\{\alpha, \frac{(\ell+1)(2\ell+1)}{\log(2\ell+1)} \log \frac{t}{\alpha}\right\} \ .$$

Let $\alpha \stackrel{\text{def}}{=} \frac{(\ell+1)(2\ell+1)\log t}{\log(2\ell+1)}$. For large $\ell$ and $t \geq 4.1\ell^2$ we have $\alpha \leq t$, yielding:

$$\mathcal{D}(G(X_{\ell,t}|Y_{\ell,t})) \geq \frac{(\ell+1)(2\ell+1)}{\log(2\ell+1)} \log \frac{t\log(2\ell+1)}{(\ell+1)(2\ell+1)} \ .$$

This bound is stronger than (17) when $\ell$ is large and $t \geq 4.1\ell^3$. For example, if $t = \ell^\delta$ for $\delta > 3$ then, for large $\ell$, this bound is roughly $2(\delta-2)\ell^2$, a factor of $(\delta-2)$ larger than the bound in (17).

Note, however, that this bound is weaker than Inequality (17) when $t = (\ell+1)(2\ell+1)$, the value for which we obtained the largest discrepancy between $\hat{C}_2(X_{\ell,t}|Y_{\ell,t})$ and $\hat{C}_\infty(X_{\ell,t}|Y_{\ell,t})$. Hence the bound we derive now improves the lower bound on $\hat{C}_2(X_{\ell,t}|Y_{\ell,t})$ and the technique for proving it; but it does not change the provable discrepancy between $\hat{C}_2(X|Y)$ and $\hat{C}_\infty(X|Y)$.

**Theorem 3** For the playoffs problem with $\ell$ leagues and $t$ teams,

$$\mathcal{D}(G(X_{\ell,t}|Y_{\ell,t})) \geq \min\left\{\alpha, \frac{(\ell+1)(2\ell+1)}{\log(2\ell+1)} \log \frac{t}{\alpha}\right\}$$

for all $\alpha \leq t$.

**Proof:** Fix $\alpha \leq t$ and let $G_1,\ldots,G_k$ be an edge cover of $G(X_{\ell,t}|Y_{\ell,t})$ with chromatic numbers $\chi_1,\ldots,\chi_k$ achieving $\mathcal{D}(G(X_{\ell,t}|Y_{\ell,t}))$. That is,

$$\sum_{i=1}^{k} \chi_i = \mathcal{D}(G(X_{\ell,t}|Y_{\ell,t})).$$

We will show that

$$\sum_{i=1}^{k}\chi_i \geq \min\left\{\alpha, \frac{(\ell+1)(2\ell+1)}{\log(2\ell+1)}\log\frac{t}{\alpha}\right\} .$$

If $\chi_i \geq \alpha$ for some $i \in \{1,\ldots,k\}$, we are done. We therefore assume that $\chi_i < \alpha$ (hence $\chi_i < t$) for all $i \in \{1,\ldots,k\}$. Sequentially, over $I$, consider the edge union $\bigcup_{i=1}^{I} G_i$. Let $I_1$ be the largest $I$ such that $\chi(\bigcup_{i=1}^{I_1} G_i) < t$ and define

$$H_1 \stackrel{\text{def}}{=} \bigcup_{i=1}^{I_1} G_i.$$

Then,

$$\begin{aligned} t &\leq \chi\left(\bigcup_{i=1}^{I_1+1} G_i\right) \\ &\leq \chi(H_1)\cdot\chi(G_{I_1+1}) \\ &< \chi(H_1)\cdot\alpha \end{aligned}$$

and so

$$\frac{t}{\alpha} < \chi(H_1) < t.$$

Next let $I_2$ be the largest $I$ such that $\chi(\bigcup_{i=I_1+1}^{I_2} G_i) < t$ and define

$$H_2 \stackrel{\text{def}}{=} \bigcup_{i=I_1+1}^{I_2} G_i.$$

It can be similarly shown that

$$\frac{t}{\alpha} < \chi(H_2) < t.$$

Proceed similarly to define $I_3,\ldots,I_r$ and an edge cover $H_1,\ldots,H_r$ of $G(X_{\ell,t}|Y_{\ell,t})$ such that for $j \in \{1,\ldots,r\}$,

$$H_j \stackrel{\text{def}}{=} \bigcup_{i=I_{j-1}+1}^{I_j} G_i.$$

(where we define $I_0 \stackrel{\text{def}}{=} 0$) and

$$\frac{t}{\alpha} < \chi(H_j) < t.$$

We now repeat the argument used to prove Lemma 3. There must be two vertices, $v_1'$ and $v_1''$, in the first league that do not belong together to any edge of $H_1$. Therefore, all edges of $G(X_{\ell,t}|Y_{\ell,t})$ containing both $v_1'$ and $v_1''$ belong to $\bigcup_{j=2}^{r} H_j$. Similarly, $\chi_2 < t$, hence there must be two vertices $v_2'$ and $v_2''$ in the second league that do not belong together to any edge in $H_2$. Therefore, all edges containing the four vertices $v_1', v_1'', v_2'$, and $v_2''$ cannot belong to either $H_1$ or $H_2$ hence must belong to $\bigcup_{j=3}^r H_j$.

We can similarly find vertices $v_1', \ldots, v_\ell'$ and $v_1'', \ldots, v_\ell''$ such that all edges containing these $2\ell$ vertices belong to $\bigcup_{j=\ell+1}^r H_j$. But such an edge exists, namely, the one corresponding to the playoffs containing teams $v_i'$ and $v_i''$ from league $i$ for $i = 1, \ldots, \ell$. Hence $\bigcup_{j=\ell+1}^r H_j$ is not empty, so

$$r \geq \ell + 1.$$

For all $j \in \{1, \ldots, r\}$, $H_j$ is $2\ell + 1$ regular hence, from Lemma 4,

$$\mathcal{D}(H_j) > \frac{2\ell + 1}{\log(2\ell + 1)} \log \frac{t}{\alpha}$$

and therefore

$$\sum_{j=1}^r \mathcal{D}(H_j) > \frac{(\ell + 1)(2\ell + 1)}{\log(2\ell + 1)} \log \frac{t}{\alpha} .$$

In the original cover, the graphs $G_1, \ldots, G_k$ achieved $\mathcal{D}(G(X_{\ell,t}|Y_{\ell,t}))$, hence, for all $j \in \{1, \ldots, r\}$,

$$\mathcal{D}(H_j) = \sum_{i=I_{j-1}+1}^{I_j} \chi(G_i)$$

(or else there would have been an edge cover with a lower chromatic sum). Therefore,

$$\begin{aligned}
\mathcal{D}(G(X_{\ell,t}|Y_{\ell,t})) &= \sum_{i=1}^k \chi_i \\
&= \sum_{j=1}^r \sum_{i=I_{j-1}+1}^{I_j} \chi_i \\
&= \sum_{j=1}^r \mathcal{D}(H_j) \\
&> \frac{(\ell + 1)(2\ell + 1)}{\log(2\ell + 1)} \log \frac{t}{\alpha} .
\end{aligned}$$

We showed that if $\alpha \leq t$ and

$$\sum_{i=1}^k \chi_i = \mathcal{D}(G(X_{\ell,t}|Y_{\ell,t}))$$

then either $\chi_i \geq \alpha$ for some $i \in \{1, \ldots, k\}$, or

$$\sum_{i=1}^k \chi_i > \frac{(\ell + 1)(2\ell + 1)}{\log(2\ell + 1)} \log \frac{t}{\alpha}$$

and the theorem is proven. $\square$

# 8 Open problems

Several problems related to the potential gap between $m$-message complexities for different $m$'s remain unsettled.

1. Lemma 2 showed that whenever $G(X|Y)$ is an intersecting hypergraph,

   $$\hat{C}_2(X|Y) \geq \log \mathcal{D}(G(X|Y))$$

   We noted that the inequality does not hold in general. Does a weaker inequality hold for all $(X, Y)$ pairs? For example, is $\hat{C}_2(X|Y) \geq \log \mathcal{D}(G(X|Y)) - c$ always true for some constant $c$?

2. For the league problem, we showed that for every $\epsilon > 0$, there are $\ell$ and $t$ such that

   $$\hat{C}_2(X_{\ell,t}|Y_{\ell,t}) > (2-\epsilon)\hat{C}_\infty(X_{\ell,t}|Y_{\ell,t}) \ .$$

   A two-message protocol similar to the "hashing protocols" used in [1] shows that for all $\ell$ and $t$,

   $$\hat{C}_2(X_{\ell,t}|Y_{\ell,t}) < 3\hat{C}_\infty(X_{\ell,t}|Y_{\ell,t})$$

   Therefore, the largest ratio between $\hat{C}_2(X_{\ell,t}|Y_{\ell,t})$ and $\hat{C}_\infty(X_{\ell,t}|Y_{\ell,t})$ is between 2 and 3. Can this maximum ratio be determined?

3. More generally, Result 5 stated that for all $(X, Y)$ pairs,

   $$\hat{C}_2(X|Y) \leq 4\hat{C}_\infty(X|Y) + 2$$

   while Inequality (5) stated that for every $\epsilon > 0$ there is an $(X, Y)$ pair for which

   $$\hat{C}_2(X|Y) \geq (2-\epsilon)\hat{C}_\infty(X|Y).$$

   Therefore, the largest ratio between $\hat{C}_2(X|Y)$ and $\hat{C}_\infty(X|Y)$ is between 2 and 4. Can this maximum ratio be determined?

4. Most importantly, what is the largest ratio between $\hat{C}_M(X|Y)$ and $\hat{C}_\infty(X|Y)$ for $M \geq 3$? It would be interesting if an $M$ could be exhibited such that for all $(X, Y)$ pairs,

   $$\hat{C}_M(X|Y) \leq \hat{C}_\infty(X|Y) + o(\hat{C}_\infty(X|Y)) \ .$$

   Thereby showing that $M$ messages are asymptotically optimal.

## Acknowledgements

I am grateful to Thomas Cover for encouraging me to prove that two messages are not optimal; to János Körner and Gábor Simonyi for helpful discussions and for disproving one of my earlier conjectures; and to Noga Alon for a discussion, following a presentation of this paper, in which he proposed an idea that led to the duelists and ghouls example.


## References


[1] A. Orlitsky. Worst-case interactive communication I: Two messages are almost optimal. *IEEE Transactions on Information Theory*, 36(5):1111–1126, September 1990.

[2] A. El Gamal and A. Orlitsky. Interactive data compression. In *Proceedings of the 25th Annual Symposium on Foundations of Computer Science*, pages 100–108, 1984.

[3] C.H. Papadimitriou and M. Sipser. Communication complexity. In *Proceedings of the 14th Annual ACM Symposium on Theory of Computing*, pages 196–200, 1982.

[4] P. Duris, Z. Galil, and G. Schnitger. Lower bounds on communication complexity. In *Proceedings of the 16th Annual ACM Symposium on Theory of Computing*, pages 81–91, 1984.

[5] H. Witsenhausen. The zero-error side information problem and chromatic numbers. *IEEE Transactions on Information Theory*, 22(5):592–593, September 1976.