

NOF-Multiparty Information Complexity Bounds for Pointer Jumping

Andre Gronemeier*

FB Informatik, LS2, Univ. Dortmund, 44221 Dortmund, Germany
andre.gronemeier@cs.uni-dortmund.de

Abstract. We prove a lower bound on the communication complexity of pointer jumping for multiparty one-way protocols in the number on the forehead model that satisfy a certain information theoretical restriction: We consider protocols for which the i th player may only reveal information about the first $i + 1$ inputs. To this end we extend the information complexity approach of Chakrabarti, Shi, Wirth, and Yao (2001) and Bar-Yossef, Jayram, Kumar, and Sivakumar (2004) to our restricted version of the multiparty number on the forehead model. The best currently known multiparty protocol for pointer jumping by Damm, Jukna, and Sgall (1998) works in this model.

1 Introduction

1.1 Multiparty Communication Complexity

In the multiparty communication game by Chandra, Furst, and Lipton [8] k players jointly compute a function $f(x_1, \dots, x_k)$ on k variables such that in the end each of the players knows the result. The players have unlimited computational power, but the i th player does not know the input variable x_i . Thus the players need to communicate to fulfill their task. This is usually called *number on the forehead model* (briefly NOF-model), since we can imagine the input x_i being written on the i th player's forehead. The players exchange messages according to a fixed protocol by writing to a shared blackboard seen by all players. Inscriptions on the blackboard are never deleted, each player appends his message to the previously written messages on the board. The current inscription on the blackboard determines unambiguously whose turn it is to write the next message and when to stop the protocol. The only important computational resource in this model is communication: The cost of the protocol is the worst case length of the inscription on the blackboard. The communication complexity of a function is the minimum cost of a protocol for the function.

Communication complexity for two players has been investigated independently [18] and is well understood [13, 12]. Less is known about general multiparty protocols with more than two players. At the time of writing, only a single general proof method for proving lower bounds on the multiparty communication complexity of functions for more than two players is known: The

* Supported by DFG grant SA 1053/1-1

discrepancy method by Babai, Nisan, and Szegedy [3] (see also [9, 17]). Consequently, restricted versions of the multiparty model received some attention, most notably the simultaneous message model and the one-way model (see [1, 13]).

In the simultaneous message model the players do not interact. Each player sends a single message, depending only on the inputs seen by the player, to a referee, who does not see the input. The referee has to announce the function value. Babai, Gál, Kimmel, and Lokam [1] introduced a proof method for proving lower bounds in the simultaneous message model which has found many applications by now, for example [16, 6, 4].

In the one-way model the interaction of the players is restricted such that the first player sends the first message, then the second player sends a message depending on the inputs seen by him and the first message, and so on. The last player has to announce the function value. Although this model is still severely restricted compared to the general model in which the players can exchange messages in an arbitrary order, for more than three players currently no proof methods are known which make use of this restriction alone. The only known lower bound for a variable number of players in the one-way model by Damm, Jukna, and Sgall [11] uses an additional restriction that is described in Sect. 1.4.

1.2 Information Complexity

Information theory (see [10] for an introduction) has been used before to obtain results on communication complexity, but it has been used mainly as a tool in small parts of the proofs. Some references to these results are contained in [4]. Recently several publications emerged in which information theory is the main ingredient of the proof [7, 4, 5]. Bar-Yossef, Jayram, Kumar, and Sivakumar [4] reduced communication complexity problems to information theory problems and solved these problems in the information theory domain. Chakrabarti, Shi, Wirth, and Yao [7] introduced the concept of information complexity for the two party model. The information complexity of a function is the amount of information about the input that the messages of any protocol for the function must reveal. In the language of information theory, the information complexity of a function is the minimum mutual information between the messages of any protocol for the function and the inputs (see Sect. 2.2 and 2.3). In [5] this concept was further refined by Bar-Yossef, Jayram, Kumar, and Sivakumar.

1.3 Our Result

We consider the NOF-multiparty one-way model with an additional information theoretical restriction. In this model we prove a lower bound on the information complexity of the pointer jumping function.

Definition 1. *Let f_1, \dots, f_k be functions with domain and range $\{1, \dots, n\}$. Then the k -player pointer jumping function Jump_n^k is defined as follows:*

$$\text{Jump}_n^k(f_1, \dots, f_k) := (f_k \circ f_{k-1} \circ \dots \circ f_1)(1) .$$

Note that for the first input f_1 only $f_1(1)$ influences the result. The inputs $f_1(2), \dots, f_1(n)$ are redundant.

Our lower bound on the communication complexity of Jump_n^k holds for all one-way protocols for which the messages M_1, \dots, M_{i-2} of the players $1, \dots, i-2$ do not reveal any information about the input f_i , in information theoretical terms, one-way protocols for which the mutual information between the input f_i and the messages M_1, \dots, M_{i-2} is 0. We call protocols which obey this restriction *myopic* (see Sect. 2.3). The currently best one-way multiparty protocol for Jump_n^k by Damm, Jukna, and Sgall [11] is myopic.

We will prove the following lower bound on the multiparty communication complexity of Jump_n^k in the number on the forehead model.

Theorem 1. *Every myopic ε -error k -party one-way protocol in the number on the forehead model for Jump_n^k has cost $\Omega(n^{(1-\varepsilon)/k} \log n)$.*

This result is based on an information complexity bound and the proof relies solely on information theoretical arguments. Our result is not stronger than bounds which can be proved in the simultaneous message model and we need our additional information theoretical restriction, but the result extends the information complexity approach beyond the number in the hand model. To the best of our knowledge, this is the first extension of the information complexity results of Chakrabarti, Shi, Wirth, and Yao [7] and Bar-Yossef, Jayram, Kumar, and Sivakumar [5] to the one-way multiparty model.

1.4 Related Work

The communication complexity of pointer jumping has been investigated mainly for a two-player version that differs slightly from Def. 1. In this model upper and lower bounds have been proved by Nisan and Wigderson [14] and Ponzio, Radhakrishnan, and Venkatesh [15].

Much less is known about the communication complexity of pointer jumping, as defined in Def. 1, in the multiparty NOF-model. Wigderson proved an $\Omega(n^{1/2})$ bound for the complexity of pointer jumping for three players in the fully general one-way model (The result is contained in the appendix of [2]). The proof method of Babai *et al.* [1] yields an $\Omega(n^{1/k})$ lower bound on the communication complexity of Jump_n^k in the simultaneous message model [16]. The currently best one-way protocol for Jump_n by Damm, Jukna, and Sgall [11] has cost $O(n)$ for $k \geq \log^* n$ players and cost $n \log^{(k-1)} n + O(n)$ for $k < \log^* n$ players. In addition, Damm, Jukna, and Sgall proved an $\Omega(n/k^2)$ lower bound for up to $O(n^{1/3-\varepsilon})$ players in a restricted one-way model which they call *conservative one-way multiparty complexity*. In this model the i th player knows the inputs f_{i+1}, \dots, f_k , but unlike the usual number on the forehead model, instead of the inputs f_1, \dots, f_{i-1} , he does only know the partial result $(f_{i-1} \circ f_{i-2} \circ \dots \circ f_1)(1)$. Since this result can take only n different values whereas the inputs f_1, \dots, f_{i-1} can take $n^{(i-1)n}$ different values, this is a potentially severe restriction.

Note that the above result is complementary to our result in the following sense: In a myopic protocol for Jump_n^k the i th player must not reveal information

about f_{i+2}, \dots, f_k . This is obviously the case if his message does not depend functionally on these inputs. Thus myopic protocols include protocols in which the i th player may only access the inputs f_j with $j < i$ and, in addition, the input f_{i+1} . For conservative protocols, the i th player has unrestricted access to the inputs f_j with $j > i$, while the access to the inputs f_j with $j < i$ is severely restricted. Note that a protocol can be conservative and myopic at the same time. Although this is a very severe restriction, the currently best one-way protocol for Jump_n^k of Damm, Jukna, and Sgall [11] is both conservative and myopic.

2 Preliminaries

2.1 Notation

We use $[n]$ as an abbreviation for the set $\{1, \dots, n\}$. Let P be a k -party one-way protocol for Jump_n^k . If the inputs of P are drawn randomly with respect to some probability distribution, then these inputs and the messages are random variables F_1, \dots, F_k and M_1, \dots, M_{k-1} , respectively. In this case let \mathcal{F} denote the set of the random variables $\{F_1, \dots, F_k\}$ and let \mathcal{M}_i denote the set of the random variables $\{M_1, \dots, M_i\}$. Furthermore let $\tilde{F}_i := F_i \circ F_{i-1} \circ \dots \circ F_1$, hence $\text{Jump}_n^k(F_1, \dots, F_k) = \tilde{F}_k(1)$.

2.2 Tools from Information Theory

In this section some basic facts from information theory are summarized. Results that are needed for an arbitrary number of random variables are only stated for two variables. In most cases the extension to an arbitrary number of variables follows immediately by induction. Most information theoretical facts that we use are elementary. Nevertheless, we see this section merely as an agreement on the notation of information theoretical results. For a proper introduction to information theory we refer the reader to the book by Cover and Thomas [10].

Let X, Y, Z , and W be random variables with a finite range R . Then $H(X) := \sum_{x \in R} \text{Prob}(X = x) \log(1/\text{Prob}(X = x))$ is called the *entropy* of X , $H(X, Y)$ is the entropy of the joint distribution of X and Y , and $H(X | Y) := H(X, Y) - H(Y)$ is called the conditional entropy of X given Y . The *mutual information* between X and Y is defined as $I(X; Y) := H(X) - H(X | Y)$ and the *conditional mutual information* between X and Y given Z is $I(X; Y | Z) := H(X | Z) - H(X | Y, Z)$.

Let E denote an event, for example $W = w$. Then $H(X | E)$ denotes the entropy of X with respect to the conditional distribution of X given the event E occurred. Conditioning on an event for conditional entropy, mutual information and conditional mutual information is defined analogously. For example, $I(X; Y | Z, W=w)$ is the mutual information between X and Y given Z with respect to the conditional distribution given the event $W=w$ occurred.

Proofs of the following elementary properties of entropy and mutual information can be found in most textbooks about information theory, for example [10].

Theorem 2. Let $X, Y, Z,$ and W be random variables with finite range R . Then

1. $0 \leq H(X) \leq \log |R|$ with $H(X) = \log |R|$ iff X is uniformly distributed.
2. $H(X | Y) = \sum_{y \in R} \text{Prob}(Y=y) H(X | Y=y)$.
3. $I(X; Y | Z) = \sum_{z \in R} \text{Prob}(Z=z) I(X; Y | Z=z)$.
4. $I(X; Y | Z, W) = \sum_{w \in R} \text{Prob}(W=w) I(X; Y | Z, W=w)$.
5. $H(X, Y) \leq H(X) + H(Y)$ with equality iff X and Y are independent.
6. If X and Y are jointly independent of Z then $H(X | Y, Z) = H(X | Y)$.
7. If X and Y are jointly independent of Z then $I(X; Y | Z) = I(X; Y)$.
8. If f is a function with domain R then $H(X, f(X)) = H(X)$.

The following useful inequality can be proved easily using Jensen's inequality.

Lemma 1. Let $a_1, \dots, a_n \in \mathbb{R}$ and $\mu : [n] \rightarrow [0, 1]$ be a probability distribution on $[n]$. Then

$$\sum_{i=1}^n \mu(i) \log a_i \leq \log \left(\sum_{i=1}^n \mu(i) a_i \right) .$$

Fano's inequality uses information theory to give bounds on the error of a predictor.

Theorem 3 (Fano's inequality). Let X and Y be random variables with range R_X and R_Y , let $P : R_Y \rightarrow R_X$ be a function that predicts the value of X from an observed value of Y , and let $\varepsilon = \text{Prob}(P(Y) \neq X)$ be the prediction error. Then

$$H_2(\varepsilon) + \varepsilon \log(|R_X| - 1) \geq H(X | Y)$$

where $H_2(\varepsilon) = \varepsilon \log(1/\varepsilon) + (1 - \varepsilon) \log(1/(1 - \varepsilon))$ denotes the binary entropy function.

Note that Fano's inequality implies $1 + \varepsilon \log(|R_X| - 1) \geq H(X | Y)$ since the binary entropy function is bounded from above by 1.

2.3 Communication Complexity and Information Complexity

The multiparty communication game by Chandra, Furst, and Lipton [8] and multiparty one-way protocols [1, 13] have been described already in the introduction. The following definition summarizes the introductory discussion.

Definition 2. In a k -party one-way protocol P with input variables x_1, \dots, x_k the i th player sees all input variables except x_i . Each player $i \in \{1, \dots, k - 1\}$ sends a single message m_i which may depend on the inputs seen by player i and the messages m_1, \dots, m_{i-1} of the previous players. The k th player announces the output $P(x_1, \dots, x_k)$ of the protocol depending on the inputs seen by the k th player and the messages m_1, \dots, m_{k-1} .

Let $f(x_1, \dots, x_k)$ be a function on k variables and let μ be a distribution on the domain of f . The protocol P is called an ε -error protocol for f with respect to μ , if $\text{Prob}_\mu(P(x_1, \dots, x_k) \neq f(x_1, \dots, x_k)) \leq \varepsilon$.

The cost $c(P)$ of the one-way k -party protocol P for f is the length of the longest message sent by any of the players. The ε -error one-way multiparty communication complexity $C_{\mu,\varepsilon}^1(f)$ of the function f with respect to distribution μ is the minimum cost of an ε -error one-way k -party protocol for f . We omit μ , if μ is the uniform distribution.

Note that our definition of the cost of a one-way protocol differs from the usual definition from [8]. We define the cost of a protocol as the length of the longest message sent by any of the players, while usually the worst case length of the whole transcript of the communication is used. For k players these two cost measures can differ at most by a factor of k .

We will impose an additional information theoretical restriction on one-way protocols. In a *myopic* protocol with random inputs X_1, \dots, X_k the messages of the players $1, \dots, i-2$ must not reveal any information about X_i .

Definition 3. Let X_1, \dots, X_k be the inputs and M_1, \dots, M_{k-1} be the messages of a k -party one-way protocol P for f . Let $\mathcal{X}_i = \{X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_k\}$. Then P is called *myopic*, if

$$I(X_i; M_1, \dots, M_{i-2} \mid \mathcal{X}_i) = 0 \quad \text{for all } 1 \leq i \leq k .$$

Let $C_{\mu,\varepsilon}^m(f)$ denote the minimum cost of a myopic ε -error one-way k -party protocol for f w.r.t. μ . We omit μ , if μ is the uniform distribution.

Note that for myopic protocols with independent inputs X_1, \dots, X_k the input X_i is independent of the messages M_1, \dots, M_{i-2} , since

$$I(X_i; M_1, \dots, M_{i-2} \mid \mathcal{X}_i) = H(X_i \mid \mathcal{X}_i) - H(X_i \mid M_1, \dots, M_{i-2}, \mathcal{X}_i) = 0$$

and therefore

$$H(X_i) = H(X_i \mid \mathcal{X}_i) = H(X_i \mid M_1, \dots, M_{i-2}, \mathcal{X}_i) .$$

The following lemma generalizes the information complexity approach of [7] and [5] to multiparty protocols. There are several sensible definitions of information complexity in the multiparty model. Therefore, instead of defining information complexity explicitly, we only state a lower bound on communication complexity in terms of mutual information that is meaningful for our application to myopic protocols.

Lemma 2. Let f be a function on k random variables X_1, \dots, X_k that are jointly distributed with respect to the distribution μ and let M_1, \dots, M_{k-1} be the messages of a k -party one-way protocol P that computes f with error ε with respect to distribution μ . Then the cost of P is bounded from below by

$$\max_{i \in [k-1]} I(M_i; X_{i+1} \mid \mathcal{X}_{i+1}, \mathcal{M}_{i-1}) .$$

Proof. Let $|M_i|$ denote the number of different values that the random variable M_i can take. Clearly, the cost $c(P)$ of P is bounded by $\max_{i \in [k-1]} \log |M_i|$ and, since conditioning reduces entropy, the definition of conditional mutual information implies

$$c(P) \geq \max_{i \in [k-1]} \log |M_i| \geq \max_{i \in [k-1]} \mathbb{H}(M_i) \geq \max_{i \in [k-1]} \mathbb{I}(M_i; X_{i+1} | \mathcal{X}_{i+1}, \mathcal{M}_{i-1}) .$$

□

3 Main Result

3.1 Outline of the Proof

Consider the situation of the i th player in a myopic k -party protocol P for Jump_n^k with uniformly distributed inputs: The i th player knows the inputs $\mathcal{F} \setminus \{F_i\}$, and when it is his turn to send a message, he additionally knows the messages $\mathcal{M}_{i-1} = \{M_1, \dots, M_{i-1}\}$ of the previous players. Since P is myopic, F_{i+1} is independent of the messages \mathcal{M}_{i-1} and, in particular, $F_{i+1}(1), \dots, F_{i+1}(n)$ are independent with respect to the conditional distribution given the first $i - 1$ messages \mathcal{M}_{i-1} .

We will show in Lemma 4 that under these circumstances the message M_i can not reveal much information about $\tilde{F}_{i+1}(1)$, if the conditional entropy of $\tilde{F}_i(1)$ given $\mathcal{F} \setminus \{F_i\}$ and M_1, \dots, M_{i-1} is large and the conditional mutual information between M_i and F_{i+1} given $\mathcal{F} \setminus \{F_{i+1}\}$ and M_1, \dots, M_{i-1} is small. This claim can be used inductively to prove a lower bound on the conditional entropy of $\tilde{F}_k(1)$ given $\mathcal{F} \setminus \{F_k\}$ and M_1, \dots, M_{k-1} , if the conditional mutual information between M_i and F_{i+1} is bounded appropriately from above for all $i \in [k - 1]$.

Intuitively, the last claim holds because the i th player needs to allocate the information about F_{i+1} to $F_{i+1}(1), \dots, F_{i+1}(n)$, since these variables are independent, whereas only one of the variables, namely $F_{i+1}(\tilde{F}_i(1))$, contains information about $\tilde{F}_{i+1}(1)$. It is difficult for the i th player to predict the value of $\tilde{F}_i(1)$, if the conditional entropy of $\tilde{F}_i(1)$ given $\mathcal{F} \setminus \{F_i\}$ and M_1, \dots, M_{i-1} is large. Therefore he has to send a lot of useless information, if he wants to reveal some information about $F_{i+1}(\tilde{F}_i(1))$. The details of this argument are contained in Lemma 3.

Finally, in Theorem 4 we will use Fano's inequality and Lemma 4 to prove a lower bound the cost of myopic protocols for Jump_n^k .

3.2 Proof of the Main Result

First we will show that a random variable M must contain much information about the n independent random variables $X = (X_1, \dots, X_n)$, if it contains much information about a randomly chosen variable from this collection which is chosen independently of X and M with respect to a distribution with large entropy.

Lemma 3. Let $X = (X_1, \dots, X_n)$ be n independent random variables with $H(X_p) \leq \log n$ for all p , let Y and M be random variables such that X and M are jointly independent of Y and let P be a function that maps Y to $[n]$. If $A := \lceil I(X_1, \dots, X_n; M | Y) / \log n \rceil < n/2$ then

$$I(X_{P(Y)}; M | Y) \leq \frac{\log(n - A) + 1 - H(P(Y))}{\log(n - A) - \log A} \log n .$$

Proof. Clearly $I(X; M | Y) = I(X; M)$ since X and M are jointly independent of Y . Since the variables X_1, \dots, X_n are independent, $H(X_1, \dots, X_n) = \sum_{p=1}^n H(X_p)$, and obviously $H(X_1, \dots, X_n | M) \leq \sum_{p=1}^n H(X_p | M)$. Therefore, by using the definition of mutual information, it follows that

$$I(X_1, \dots, X_n; M | Y) = I(X_1, \dots, X_n; M) \geq \sum_{p=1}^n I(X_p; M) .$$

Furthermore $I(X_{P(Y)}; M | Y) = I(X_{P(Y)}; M | Y, P(Y))$ since $P(Y)$ is a function of Y and $I(X_{P(Y)}; M | Y, P(Y) = p) = I(X_p; M)$ since X and M are jointly independent of Y . Therefore

$$\begin{aligned} I(X_{P(Y)}; M | Y) &= I(X_{P(Y)}; M | Y, P(Y)) \\ &= \sum_{p=1}^n \text{Prob}(P(Y) = p) \cdot I(X_{P(Y)}; M | Y, P(Y) = p) \\ &= \sum_{p=1}^n \text{Prob}(P(Y) = p) \cdot I(X_p; M) . \end{aligned}$$

Assume w.l.o.g. that $\text{Prob}(P(Y)=1) \geq \text{Prob}(P(Y)=2) \geq \dots \geq \text{Prob}(P(Y)=n)$. Then the last sum is maximized, if $I(X_p; M)$ is large for small values of p . Since $I(X_p; M) \leq H(X_p) \leq \log n$ and $\sum_{p=1}^n I(X_p; M) \leq I(X_1, \dots, X_n; M) \leq A \cdot \log n$, we get an upper bound for the sum, if we assume that $I(X_p, M) = \log n$ for $p \leq A$ and $I(X_p, M) = 0$ for $p > A$. Let Z be a random variable such that $Z = 1$ if $P(Y) \leq A$ and $Z = 0$ if $P(Y) > A$. Then, using the upper bound described above, we get

$$I(X_{P(Y)}; M | Y) \leq \text{Prob}(Z = 1) \cdot \log n .$$

The value of Z is a function of $P(Y)$. Hence $H(P(Y)) = H(P(Y), Z) = H(Z) + H(P(Y) | Z)$ and

$$H(P(Y) | Z) = H(P(Y)) - H(Z) \geq H(P(Y)) - 1 .$$

On the other hand

$$\begin{aligned} H(P(Y) | Z) &= \text{Prob}(Z = 1) \cdot H(P(Y) | Z = 1) \\ &\quad + (1 - \text{Prob}(Z = 1)) \cdot H(P(Y) | Z = 0) \\ &\leq \text{Prob}(Z = 1) \cdot \log A + (1 - \text{Prob}(Z = 1)) \cdot \log(n - A) \end{aligned}$$

where the last inequality is due to the fact, that under the condition $Z = 1$ the values of $P(Y)$ are from $\{1, \dots, A\}$ while under the condition $Z = 0$ the values of $P(Y)$ are from $\{A + 1, \dots, n\}$. By combining the two inequalities we get

$$\text{Prob}(Z = 1)[\log(n - A) - \log A] \leq \log(n - A) + 1 - H(P(Y))$$

and the using the premise $A < n/2 \Leftrightarrow \log(n - A) - \log A > 0$ we get

$$\text{Prob}(Z = 1) \leq \frac{\log(n - A) + 1 - H(P(Y))}{\log(n - A) - \log A} .$$

Finally, by substituting this bound into our estimate of $I(X_{P(Y)}; M | Y)$, we get the claimed result

$$\begin{aligned} I(X_{P(Y)}; M | Y) &\leq \text{Prob}(Z = 1) \cdot \log n \\ &\leq \frac{\log(n - A) + 1 - H(P(Y))}{\log(n - A) - \log A} \log n . \end{aligned}$$

□

Now we will apply the last lemma to myopic one-way protocols for pointer jumping.

Lemma 4. *Let M_1, \dots, M_{k-1} be the messages of a myopic k -player one-way protocol P for Jump_n^k with uniformly distributed inputs F_1, \dots, F_k such that the cost of P satisfies $\lceil c(P)/\log n \rceil < n/2$ and the messages of P satisfy*

$$I(F_{i+1}; M_i | \mathcal{F} \setminus \{F_{i+1}\}, \mathcal{M}_{i-1}) / \log n \leq C$$

for all $i < k$. Then

$$H(\tilde{F}_i(1) | \mathcal{F} \setminus \{F_i\}, \mathcal{M}_{i-1}) \geq \log n - i - i \log(C + 1)$$

for all $i \leq k$.

Proof. Clearly, by the definition of mutual information,

$$\begin{aligned} I(\tilde{F}_{i+1}(1); M_i | \mathcal{F} \setminus \{F_{i+1}\}, \mathcal{M}_{i-1}) = \\ H(\tilde{F}_{i+1}(1) | \mathcal{F} \setminus \{F_{i+1}\}, \mathcal{M}_{i-1}) - H(\tilde{F}_{i+1}(1) | \mathcal{F} \setminus \{F_{i+1}\}, \mathcal{M}_i) . \end{aligned}$$

Since P is myopic, the first term on the right side of the last equation is equal to $\log n$. Let $B_i := H(\tilde{F}_i(1) | \mathcal{F} \setminus \{F_i\}, \mathcal{M}_{i-1})$. Then we get

$$B_{i+1} = \log n - I(\tilde{F}_{i+1}(1); M_i | \mathcal{F} \setminus \{F_{i+1}\}, \mathcal{M}_{i-1}) .$$

The message M_i does only depend on $\mathcal{F} \setminus \{F_i\}$ and \mathcal{M}_{i-1} . For fixed values of n and k one can easily show that there is only a finite number of different messages. Thus we can assume that $M_i \in \{0, 1\}^m$ for some fixed m and use the messages of the protocol as an index of summation without worrying about convergence. For

fixed f and m let $E_i(f, m)$ denote the event that $(F_1, \dots, F_{i-1}, F_{i+2}, \dots, F_k) = f$ and $(M_1, \dots, M_{i-1}) = m$ (note that both F_i and F_{i+1} are not fixed). Then, by using $\tilde{F}_{i+1}(1) = F_{i+1}(\tilde{F}_i(1))$ and expanding the conditional mutual information, we get

$$B_{i+1} = \sum_{f,m} \text{Prob}(E_i(f, m)) \cdot \left[\log n - \text{I}(F_{i+1}(\tilde{F}_i(1)); M_i \mid F_i, E_i(f, m)) \right] .$$

Under the condition $E_i(f, m)$ the messages M_1, \dots, M_{i-1} and all inputs except F_i and F_{i+1} are fixed to constants. In this case $\tilde{F}_i(1)$ is a function of F_i , since $\tilde{F}_i(1) = F_i(\tilde{F}_{i-1}(1))$ and $\tilde{F}_{i-1}(1)$ is constant under the condition $E_i(f, m)$. Similarly M_i is a function of F_{i+1} , the only variable seen by player i that is not fixed by the conditioning event. Furthermore, since P is myopic, F_{i+1} is independent of M_1, \dots, M_{i-1} and $\mathcal{F} \setminus \{F_{i+1}\}$. Thus $\text{H}(F_{i+1} \mid F_i, E_i(f, m)) = \text{H}(F_{i+1} \mid E_i(f, m))$ implying that F_{i+1} is independent of F_i under the condition $E_i(f, m)$. Hence under the condition $E_i(f, m)$ the random variables F_{i+1} and M_i are jointly independent of F_i and the random variables $F_{i+1}(p)$ for $p = 1, \dots, n$ are independent and satisfy $\text{H}(F_{i+1}(p) \mid E_i(f, m)) \leq \log n$. Clearly even under the condition $E_i(f, m)$ the entropy of M_i is bounded from above by the cost of P , thus $\lceil \text{I}(F_{i+1}; M_i \mid F_i, E_i(f, m)) / \log n \rceil \leq \lceil c(P) / \log n \rceil < n/2$. Therefore we can estimate $S_i(f, m) := \log n - \text{I}(F_{i+1}(\tilde{F}_i(1)); M_i \mid F_i, E_i(f, m))$ using Lemma 3 with $X_p = F_{i+1}(p)$, $Y = F_i$, $P(F_i) = F_i(\tilde{F}_{i-1}(1)) = \tilde{F}_i(1)$, $M = M_i$, and $A_{i+1}(f, m) := \lceil \text{I}(F_{i+1}; M_i \mid F_i, E_i(f, m)) / \log n \rceil$ to get

$$\begin{aligned} S_i(f, m) &= \log n - \text{I}(F_{i+1}(\tilde{F}_i(1)); M_i \mid F_i, E_i(f, m)) \\ &\geq \log n - \frac{\log(n - A_{i+1}(f, m)) + 1 - \text{H}(\tilde{F}_i(1) \mid E_i(f, m))}{\log(n - A_{i+1}(f, m)) - \log A_{i+1}(f, m)} \cdot \log n \\ &= \frac{\text{H}(\tilde{F}_i(1) \mid E_i(f, m)) - \log A_{i+1}(f, m) - 1}{\log(n - A_{i+1}(f, m)) - \log A_{i+1}(f, m)} \cdot \log n \\ &\geq \text{H}(\tilde{F}_i(1) \mid E_i(f, m)) - \log A_{i+1}(f, m) - 1 . \end{aligned}$$

For the last inequality we use that $\log(n - A_{i+1}(f, m)) - \log A_{i+1}(f, m) \leq \log n$. From this estimate of $S_i(f, m)$ we get

$$\begin{aligned} B_{i+1} &= \sum_{f,m} \text{Prob}(E_i(f, m)) \cdot S_i(f, m) \\ &\geq \sum_{f,m} \text{Prob}(E_i(f, m)) \cdot \text{H}(\tilde{F}_i(1) \mid E_i(f, m)) \\ &\quad - \sum_{f,m} \text{Prob}(E_i(f, m)) \cdot \log A_{i+1}(f, m) - 1 . \end{aligned}$$

Let T_1 and T_2 denote the first and second term of the right hand side in the last inequality, respectively. Then

$$T_1 = \text{H}(\tilde{F}_i(1) \mid \mathcal{F} \setminus \{F_i, F_{i+1}\}, \mathcal{M}_{i-1}) \geq B_i$$

where the last inequality holds, because conditioning reduces entropy. We apply Lemma 1 to the second term and get

$$\begin{aligned}
T_2 &= \sum_{f,m} \text{Prob}(E_i(f, m)) \cdot \log A_{i+1}(f, m) \\
&\leq \log \left(\sum_{f,m} \text{Prob}(E_i(f, m)) \cdot A_{i+1}(f, m) \right) \\
&= \log \left(\sum_{f,m} \text{Prob}(E_i(f, m)) \cdot \lceil \text{I}(F_{i+1}; M_i \mid F_i, E_i(f, m)) / \log n \rceil \right) \\
&\leq \log \left(\sum_{f,m} \text{Prob}(E_i(f, m)) \cdot (\text{I}(F_{i+1}; M_i \mid F_i, E_i(f, m)) / \log n + 1) \right) \\
&= \log (\text{I}(F_{i+1}; M_i \mid \mathcal{F} \setminus \{F_{i+1}\}, \mathcal{M}_{i-1}) / \log n + 1) \\
&\leq \log(C + 1) .
\end{aligned}$$

Thus $B_{i+1} \geq T_1 - T_2 - 1 \geq B_i - \log(C + 1) - 1$ and the claim of the Theorem is implied by this recurrence relation and the base case $B_1 = \log n$. \square

The main result follows from the last lemma by a simple application of Fano's inequality.

Theorem 4. $C_\varepsilon^m(\text{Jump}_n^k) \geq (2^{-(1+1/k)} n^{(1-\varepsilon)/k} - 2) \log n$.

Proof. Let P be a myopic ε -error protocol for Jump_n^k , let F_1, \dots, F_k be the random inputs of Jump_n^k , and let M_1, \dots, M_{k-1} be the messages of P for this input. The k th player uses F_1, \dots, F_{k-1} and M_1, \dots, M_{k-1} to predict the value of $\text{Jump}_n^k(F_1, \dots, F_k) = \tilde{F}_k(1)$ with error ε , thus, by Fano's inequality,

$$1 + \varepsilon \log n \geq \text{H}(\tilde{F}_k(1) \mid \mathcal{F} \setminus \{F_k\}, \mathcal{M}_{k-1}) .$$

Recall that $c(P)$ denotes the cost of P and let $C := \lceil c(P) / \log n \rceil$. If $C \geq n/2$ then the claim of the Theorem holds for $k \geq 2$ and sufficiently large n . If $C < n/2$ then, by Lemma 2, $\lceil \text{I}(F_{i+1}; M_i \mid \mathcal{F} \setminus \{F_{i+1}\}, \mathcal{M}_{i-1}) / \log n \rceil \leq C$ for all $i \in [k-1]$, and consequently, by Lemma 4,

$$\text{H}(\tilde{F}_k(1) \mid \mathcal{F} \setminus \{F_k\}, \mathcal{M}_{k-1}) \geq \log n - k - k \log(C + 1) .$$

Combining these inequalities yields $1 + \varepsilon \log n \geq \log n - k - k \log(C + 1)$ which implies $C \geq n^{(1-\varepsilon)/k} / 2^{1+1/k} - 1$. The claim of the theorem follows immediately from the last inequality. \square

Acknowledgment

Thanks to Martin Sauerhoff for proofreading and advice.

References

1. Babai, L., Gál, A., Kimmel, P.G., Lokam, S.V.: Communication complexity of simultaneous messages. *SIAM J. Comput.* **33**(1) (2004) 137–166
2. Babai, L., Hayes, T.P., Kimmel, P.G.: The cost of the missing bit: Communication complexity with help. *Combinatorica* **21**(4) (2001) 455–488
3. Babai, L., Nisan, N., Szegedy, M.: Multiparty protocols, pseudorandom generators for logspace, and time-space trade-offs. *J. Comput. Syst. Sci.* **45**(2) (1992) 204–232
4. Bar-Yossef, Z., Jayram, T.S., Kumar, R., Sivakumar, D.: Information theory methods in communication complexity. In: *Proc. of 17th CCC.* (2002) 93–102
5. Bar-Yossef, Z., Jayram, T.S., Kumar, R., Sivakumar, D.: An information statistics approach to data stream and communication complexity. *J. Comput. Syst. Sci.* **68**(4) (2004) 702–732
6. Beame, P., Pitassi, T., Segerlind, N., Wigderson, A.: A direct sum theorem for corruption and the multiparty NOF communication complexity of set disjointness. In: *Proc. of 20th CCC.* (2005) 52–66
7. Chakrabarti, A., Shi, Y., Wirth, A., Yao, A.C.: Informational complexity and the direct sum problem for simultaneous message complexity. In: *Proc. of 42nd FOCS.* (2001) 270–278
8. Chandra, A.K., Furst, M.L., Lipton, R.J.: Multi-party protocols. In: *Proc. of 15th STOC.* (1983) 94–99
9. Chung, F.R.K., Tetali, P.: Communication complexity and quasi randomness. *SIAM J. Discret. Math.* **6**(1) (1993) 110–125
10. Cover, T.M., Thomas, J.A.: *Elements of Information Theory.* Wiley-Interscience (1991)
11. Damm, C., Jukna, S., Sgall, J.: Some bounds on multiparty communication complexity of pointer jumping. *Comput. Complex.* **7**(2) (1998) 109–127
12. Hromkovič, J.: *Communication Complexity and Parallel Computing.* Springer (2002)
13. Kushilevitz, E., Nisan, N.: *Communication Complexity.* Cambridge University Press (1997)
14. Nisan, N., Wigderson, A.: Rounds in communication complexity revisited. *SIAM J. Comput.* **22**(1) (1993) 211–219
15. Ponzio, S., Radhakrishnan, J., Venkatesh, S.: The communication complexity of pointer chasing. *J. Comput. Syst. Sci.* **62**(2) (2001) 323–355
16. Pudlák, P., Rödl, V., Sgall, J.: Boolean circuits, tensor ranks, and communication complexity. *SIAM J. Comput.* **26**(3) (1997) 605–633
17. Raz, R.: The BNS-Chung criterion for multi-party communication complexity. *Comput. Complex.* **9**(2) (2000) 113–122
18. Yao, A.C.: Some complexity questions related to distributive computing (preliminary report). In: *Proc. of 11th STOC.* (1979) 209–213