Worst-Case Nonzero-Error Interactive Communication

Hugues Mercier, *Student Member, IEEE*, Pierre McKenzie, Stefan Wolf, and Vijay K. Bhargava, *Fellow, IEEE*

Abstract

In the interactive communication model, two distant parties $P_{\mathcal{X}}$ and $P_{\mathcal{Y}}$ respectively possess private but correlated inputs x and y, and $P_{\mathcal{Y}}$ wants to learn x from $P_{\mathcal{X}}$ while minimizing the communication for the worst possible input pair (x, y). Our main contribution is the analysis of nonzero-error models in this correlated data setting. In the private coin randomized model, both players are allowed to toss coins and P_Y must learn x with high probability for every input pair. The public coin randomized model is similar to the first model, but instead of private coins, both players have access to a common source of randomness. The private coin randomized amortized model is also similar to the first model, with the addition that the players are also allowed to solve several independent instances of the same problem simultaneously instead of sequentially. The last model, called the distributional model, is deterministic, but P_Y is allowed to answer incorrectly for a small fraction of the inputs weighted by their probability distribution.

We show that the public coin randomized, private coin randomized amortized and distributional models are equivalent and can reduce the communication compared to the original worst-case deterministic model. Moreover, when the players are not allowed to interact, the difference between the best deterministic and public coin randomized protocols can be arbitrarily large. We prove that one round of communication is almost optimal for the private coin randomized model. We also show that the deterministic model and all the nonzero-error models are equivalent for a large class of symmetric

The material in this paper was presented in part at the IEEE International Symposium on Information Theory, Adelaide, Australia, September 2005.

H. Mercier and V. K. Bhargava are with the Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, BC V6T 1Z4, Canada (email:huguesm@ece.ubc.ca; vijayk@ece.ubc.ca).

P. Mckenzie and S. Wolf are with the Département d'informatique et de recherche opérationnelle, Université de Montréal, Montréal, PQ H3C 3J7, Canada (email: mckenzie@iro.umontreal.ca; wolf@iro.umontreal.ca).

problems arising from several practical applications, although nonzero-error and randomization allows efficient one-way protocols.

Index Terms

Interactive communication, nonzero-error, randomization, worst-case protocols, communication complexity

I. INTRODUCTION

Interactive communication was introduced by Orlitsky [1] and lies at the intersection of information theory and communication complexity. It studies the amount of communication needed for one party to convey information to a second party who has correlated information. Two players, an *informant* $P_{\mathcal{X}}$ and a *recipient* $P_{\mathcal{Y}}$, respectively possess private but correlated inputs x and y. $P_{\mathcal{Y}}$ wants to learn his interlocutor's input without error while minimizing the number of bits that need to be transmitted in the worst case. To do so, they alternately exchange data on a noise-free channel following a deterministic protocol they have agreed upon initially. Unlike the original communication complexity model [2] (see [3] for an exhaustive survey), the function to be computed is trivial ($f(x, y) = x$), but the problem is to exploit the correlation between the parties' knowledge for reducing the required amount of communication. The following example illustrates the model.

Example 1: The League Problem [1]

A sport league has 2^n teams, and the name of each team is a binary string of n bits. $P_{\mathcal{Y}}$ knows the two teams playing in the championship match, but a blackout during the game restricts him to learn who wins. $P_{\mathcal{X}}$, on the other hand, hears the name of the champion team on the radio but has no idea who is the runner-up. $P_{\mathcal{Y}}$ wants to learn the identity of the champion team from $P_{\mathcal{X}}$ with certainty while minimizing the communication.

If only one round of communication from $P_{\mathcal{X}}$ to $P_{\mathcal{Y}}$ is allowed, then $P_{\mathcal{X}}$ has to transmit the n bits of the winning team. If less than n bits are transmitted, there are two teams for which $P_{\mathcal{X}}$ sends the same message; if those two teams happen to play in the championship match, the $P_{\mathcal{Y}}$ is not able to learn the winner with certainty. However, a substantial gain can be achieved when interaction is allowed. $P_{\mathcal{Y}}$ sends the position of one of the bits where the names of the two finalists differ, which requires $\lceil \log n \rceil$ bits. It is then sufficient for $P_{\mathcal{X}}$ to send the bit of the

winning team at the required position. This protocol requires $\lceil \log n \rceil + 1$ communication bits, an exponential gain compared to the one-way protocol. Moreover, Orlitsky has shown that even if more than two rounds of communication are allowed, no protocol can solve the problem by exchanging a smaller number of bits in the worst case.

Although the example above might seem artificial, interactive communication includes a large class of symmetric problems [4] inherent to several practical applications including synchronization of mobile data [5], reconciliation of sequences of symbols such as nucleotides sequences in DNA molecules [6], remote data storage [7] and quantum key distribution [8].

In this paper, we study worst-case nonzero-error interactive communication and compare the results with the original worst-case deterministic model. We allow P_Y to learn P_X's input with a probability of error at most ϵ and study how it can improve the communication, either by reducing the number of bits that need to be exchanged or by reducing the number of rounds of communication. Four nonzero-error models are presented. The first model, worst-case private coin randomized interactive communication, allows $P_{\mathcal{Y}}$ to learn x with probability at least $1-\epsilon$ for all the possible input pairs (x, y). The players can also use randomized protocols: each player has a private, independent source of randomness whose output can be used to decide which bits should be transmitted. The second model, worst-case public coin randomized interactive communication, also allows $P_{\mathcal{Y}}$ to learn $P_{\mathcal{X}}$'s input with probability at least $1-\epsilon$ for all the possible input pairs. It also uses randomized protocols, but instead of private coins, both players can use a public (common) random generator. The third model, worst-case private coin randomized amortized interactive communication, allows the players to solve several independent instances of the same problem simultaneously instead of sequentially. The players are again permitted to use private coins, and $P_{\mathcal{Y}}$ can fail to learn x with probability at most ϵ for every input pair. The fourth model, worst-case distributional interactive communication, permits only deterministic protocols, but $P_{\mathcal{Y}}$ can learn x incorrectly for a fraction at most ϵ of all the inputs weighted by their probability distribution.

We prove that the worst-case public coin randomized, private coin randomized amortized and distributional models are equivalent and that optimal protocols for the three models do not require interaction between the players. The models can be arbitrarily better than the worst-case deterministic model when a single round of communication from $P_{\mathcal{X}}$ to $P_{\mathcal{Y}}$ is allowed.

We show that for Cartesian-product pairs, the deterministic and all the nonzero-error models are equivalent and inefficient. The models are equivalent and efficient for symmetric problems including all the applications previously mentioned, although nonzero error and randomization allows efficient protocols using only one round of communication.

The most challenging model is the worst-case private coin randomized model. We show that the best one-round protocols for this model are at most three times more expensive than the best randomized or deterministic protocols using an unbounded number of rounds of communication. This is a striking difference from the deterministic model, for which Orlitsky [1] has shown that the best one-round protocols can require to the transmission of exponentially more bits than the optimal protocols. It is also different from randomized communication complexity of boolean functions, which exhibits the same phenomenon [9]. We also prove that the worst-case randomized and deterministic models are equivalent for several classes of problems and conjecture that both models are equivalent for all the interactive communication problems.

The outline of the paper is as follows. In Section II, we describe the complexity models and present the existing work. The private coin and public coin randomized models are treated in Sections III and IV, respectively. Randomized amortized interactive communication is studied in Section V, and the distributional model in Section VI. The results for balanced and symmetric problems are presented in Section VII. Finally, in Section VIII, we discuss open problems related to the private coin randomized model.

II. Complexity Models and Known Results

A. Preliminaries

The framework for studying interactive communication was introduced by Orlitsky in his seminal paper [1]. Let X and Y be finite sets, and let $S \subseteq X \times Y$, the *support set* of (X, Y), define an interactive communication problem. Two players, $P_{\mathcal{X}}$ and $P_{\mathcal{Y}}$, possess respectively inputs $x \in X$ and $y \in Y$ such that $(x, y) \in S$, and they want $P_{\mathcal{Y}}$ to learn x while minimizing the communication between them (it is not necessary for $P_{\mathcal{X}}$ to learn y). It is assumed that the communication between the players is binary.

A *k-round protocol* is a protocol such that for every input, there are at most $k-1$ alternations between the data sent by $P_{\mathcal{X}}$ and the data sent by $P_{\mathcal{Y}}$. Due to the asymmetric nature of the interactive communication model, it is assumed that the last round of communication is always

from $P_{\mathcal{X}}$ to $P_{\mathcal{Y}}$. A 1-round protocol is also called *one-way*, and a protocol requiring more that one round of communication is called *two-way*.

A hypergraph is an ordered pair $G = (V, E)$, where V is the set of vertices and E the set of hyperedges. Each hyperedge is a subset of V. Two distinct vertices v_1 and v_2 of a hypergraph are adjacent if there is an edge $e \in E$ such that $v_1 \in e$ and $v_2 \in e$. A *proper coloring* of a hypergraph is a partition of V in colors such that no adjacent vertices have the same color. The *chromatic number* χ of a hypergraph is the smallest number of colors for which there exists a proper coloring of G. A convenient way to analyze an interactive communication problem $S \subseteq X \times Y$ is to use its *characteristic hypergraph* G_S. The vertices of G_S are the elements of X, and for every $y \in Y$ there is a hyperedge $E(y) \triangleq \{x \mid (x, y) \in S\}$. The number of different hyperedges of G_S is denoted σ. It should be noted that all the asymptotic bounds presented in this paper only make sense for the support sets for which χ is a function of the size of the inputs.

The *ambiguity set* of an input $x \in X$, defined as $a(x) \triangleq \{y \in Y \mid (x, y) \in S\}$, is the set of all possible inputs for $P_{\mathcal{Y}}$ given that $P_{\mathcal{X}}$'s input is x, and the *ambiguity* of x is $|a(x)|$. The *maximum ambiguity* of $P_{\mathcal{X}}$, $\widehat{a_{\mathcal{X}}} \triangleq \max_{x \in X}\{|a(x)|\}$, is the maximum number of possible elements of Y for any element in X. Note that $a(y)$, $|a(y)|$ and $\widehat{a_{\mathcal{Y}}}$ are defined similarly with the assumption that $\widehat{a_{\mathcal{Y}}} > 1$, since a support set S with $\widehat{a_{\mathcal{Y}}} = 1$ is trivial and does not require communication.

Example 2: As an illustration for the league problem presented in Example 1, the support set L is defined as $L = \{(t_1, \{t_1, t_2\}) \mid t_1 \neq t_2\}$, where $t_1, t_2 \in \{0, 1\}^n$ are teams. The vertices of G_L are the teams of the league and its hyperedges are the possible match-ups for the championship game. It follows that the maximum ambiguity of $P_{\mathcal{X}}$ is the the number of possible runner-ups given the champion team ($\widehat{a_{\mathcal{X}}} = 2^n - 1$), and the maximum ambiguity of $P_{\mathcal{Y}}$ is the number of possible champion teams given the two finalists ($\widehat{a_{\mathcal{Y}}} = 2$).

A support set $S \subseteq X \times Y$ is a *Cartesian-product* support set if there exists $X' \subseteq X$ and $Y' \subseteq Y$ such that $S = X' \times Y'$. A support set S is *balanced* if $\widehat{a_{\mathcal{X}}} = \widehat{a_{\mathcal{Y}}}$. A *symmetric* support set is a support such that $(x, y) \in S$ if and only if $(y, x) \in S$ (it is clear that symmetric support sets are also balanced). Symmetric support sets arise naturally in all the problems for which the parties' inputs are bounded by a certain "distance", including all the applications mentioned in Section I.

B. Worst-Case Deterministic Model

The worst-case deterministic model was introduced by Orlitsky [1]. It has the following characteristics:

1) Both players send information according to a deterministic protocol. Each player sends messages based on his input and the messages previously received.
2) When a player sends a message, his interlocutor knows when it ends, and both players know that the transmission ends when the protocol halts. The *codeword* of $(x, y) \in S$ is the concatenation of the messages sent by the players on the input pair (x, y). It can be shown that the set of possible codewords is prefix-free.
3) $P_{\mathcal{Y}}$ has to learn x without error for every pair $(x, y) \in S$.

The *worst-case deterministic complexity* of a support set S, written $\hat{C}_\infty(S)$, is the minimum number of bits the players have to exchange in order for $P_{\mathcal{Y}}$ to learn x without error for every pair $(x, y) \in S$. We write $\hat{C}_k(S)$ when the number of rounds of communication is bounded by k; obviously $\hat{C}_k(S)$ decreases with k and $\hat{C}_\infty(S) = \lim_{k\to\infty} \hat{C}_k(S)$. We write $\hat{C}^*(S)$ for the number of bits that need to be transmitted if $P_{\mathcal{X}}$ knows y in advance.

The following results have been shown by Orlitsky [1]. A deterministic protocol requires at least $\lceil \log \widehat{a_{\mathcal{Y}}} \rceil$ bits of communication, otherwise if $|a(y)| = a_{\mathcal{Y}}$, then there are different input pairs (x_1, y) and (x_2, y) for which the communication between the players is the same. Clearly, the bound is tight if $P_{\mathcal{X}}$ knows y in advance and a single round of communication is sufficient.

Result 3:

$$\hat{C}_\infty(S) \geq \lceil \log \widehat{a_{\mathcal{Y}}} \rceil = \hat{C}^*(S).$$

The one-way deterministic complexity is the logarithm of the chromatic number of the underlying hypergraph of the problem, and one-way protocols require at most exponentially more bits than protocols allowing interaction.

Result 4:

$$\hat{C}_\infty(S) \geq \lceil \log \hat{C}_1(S) \rceil + 1 = \lceil \log \lceil \log \chi \rceil \rceil + 1.$$

A remarkable result from Orlitsky is that two rounds of communication are almost optimal for every problem, i.e.,

$$\hat{C}_2(S) \leq 4\hat{C}_\infty(S) + 3.$$

This is quite different from the original communication complexity model, where for every $k > 0$, there is a function whose best k-round protocol requires exponentially more bits than its best $(k+1)$-round protocol [10].

Orlitsky [11] has shown in a subsequent paper that two rounds of communication are not optimal for worst-case deterministic interactive communication, and Zhang and Xia [12] have proved that three rounds are not optimal either. Ahlswede, Cai and Zhang [13] have conjectured that four rounds are optimal, but the problem remains open, i.e., whether there is a k such that

$$\hat{C}_k(S) \leq \hat{C}_\infty(S) + o(\hat{C}_\infty(S)).$$

Naor, Orlitsky and Shor [14] have proved an upper bound on the 4-round deterministic complexity.

Result 5:

$$\begin{aligned}\hat{C}_4(S) &\leq \log\log\sigma + \log\widehat{a_{\mathcal{Y}}} + 3\log\log\widehat{a_{\mathcal{Y}}} + 7\\ &\leq \log\log\chi + 2\log\widehat{a_{\mathcal{Y}}} + 3\log\log\widehat{a_{\mathcal{Y}}} + 7\\ &\leq 3\hat{C}_\infty(S) + o(\hat{C}_\infty(S)).\end{aligned}$$

Balanced and symmetric support sets have been studied by Orlitsky [4], who has proved that the best one-way protocols require at most two times the amount of communication required by optimal protocols, i.e,

$$\hat{C}_1(S) \leq 2\hat{C}_\infty(S) + 1,$$

and that three rounds of communication are optimal.

Result 6: Let S be a balanced support set. Then,

$$\hat{C}_\infty(S) \leq \hat{C}_3(S) \leq \log\widehat{a_{\mathcal{Y}}} + 3\log\log\widehat{a_{\mathcal{Y}}} + 11 \leq \hat{C}_\infty(S) + o(\hat{C}_\infty(S)).$$

C. Worst-Case Private Coin Randomized Model

In our first nonzero-error model, $P_{\mathcal{Y}}$ is allowed to learn x with probability of error ϵ. The players are also allowed to toss coins; $P_{\mathcal{X}}$ and $P_{\mathcal{Y}}$ possess respectively independent finite random strings $r_{\mathcal{X}}$ and $r_{\mathcal{Y}}$ of arbitrary length. The communication bits become random variables: bits sent by $P_{\mathcal{X}}$ depend on x and $r_{\mathcal{X}}$, and bits sent by $P_{\mathcal{Y}}$ depend on y and $r_{\mathcal{Y}}$. It is therefore possible that for a fixed input pair (x, y), a protocol outputs different results for different values of $r_{\mathcal{X}}$ and $r_{\mathcal{Y}}$.

Let S be a support set and let $\mathcal{P}$ be a randomized protocol. $\mathcal{P}$ computes S with error ϵ if, for every pair $(x, y) \in S$, the probability that $P_{\mathcal{Y}}$ answers x on input (x, y) is at least $1 - \epsilon$. The *worst-case communication* of a protocol $\mathcal{P}$ on input (x, y) is the maximum number of bits communicated for any choice of the random strings $r_{\mathcal{X}}$ et $r_{\mathcal{Y}}$. The *worst-case cost* of $\mathcal{P}$ is the maximum, for all the inputs (x, y), of the worst-case communication of $\mathcal{P}$ on (x, y).

The *ϵ-error worst-case randomized complexity* of S, written $\hat{R}^{\epsilon}_{\infty}(S)$, is the minimum worst-case cost of a randomized protocol computing S with error ϵ, for $0 < \epsilon < \frac{1}{2}$. In other words, $\hat{R}^{\epsilon}_{\infty}(S)$ is the number of bits transmitted in the worst-case by the best protocol which, for every pair $(x, y) \in S$, allows $P_{\mathcal{Y}}$ to learn x with probability at least $1 - \epsilon$. We write $\hat{R}^{\epsilon}_{k}(S)$ when the number of rounds is bounded by k. Also, for the rest of this paper unless specified otherwise, ϵ is constant and $c(\epsilon)$ is a function of ϵ.

In his original paper, Orlitsky [1] has briefly studied a weaker randomized model considering the average communication over the choices of $r_{\mathcal{X}}$ and $r_{\mathcal{Y}}$ for the worst input pair. Using this model, he has shown that the one-way randomized complexity is at more four times the worst-case deterministic complexity, i.e.,

$$\mathfrak{R}^{\epsilon}_{1}(S) \leq 4\hat{C}_{\infty}(S) + 2\log\frac{1}{\epsilon}.$$

D. Worst-Case Public Coin Randomized Model

In the randomized model previously defined, each player has his own random generator. $P_{\mathcal{X}}$ cannot see $r_{\mathcal{Y}}$ and vice-versa. In the public coin randomized model, both players can access a common "public" random coin. Formally, both players have a common random string r following a probability distribution Π. Communication bits sent by $P_{\mathcal{X}}$ depend on x and r, and those sent by $P_{\mathcal{Y}}$ depend on y and r. A public coin randomized protocol can also be viewed as a probability distribution over a family of worst-case deterministic protocols.

The *ϵ-error worst-case public coin randomized complexity* of a support set S, written $\hat{R}^{\epsilon,pub}_{\infty}(S)$, is the number of bits transmitted in the worst case by the best public coin protocol which allows $P_{\mathcal{Y}}$ to learn x with an error probability bounded by ϵ for every pair $(x, y) \in S$. We write $\hat{R}^{\epsilon,pub}_{k}(S)$ when the number of rounds is bounded by k.

E. Worst-Case Amortized Models

For several models of computation including interactive communication, the simultaneous resolution of several independent instances of a problem can be more efficient than the sequential resolution of the instances. This phenomenon is named the *direct-sum problem* and was introduced by Karchmer, Raz, and Wigderson [15] for communication complexity of relations as a promising approach to separate the complexity classes $\mathcal{NC}^1$ and $\mathcal{NC}^2$ [16].

Let $S \subseteq X \times Y$ be an interactive communication problem, and let $(x_1, y_1), (x_2, y_2), \ldots, (x_l, y_l)$ be l independent instances of S. P_X knows $(x_1, x_2, \ldots, x_l)$, P_Y knows $(y_1, y_2, \ldots, y_l)$, and the goal is for $P_\mathcal{Y}$ to learn all the x_i from $P_\mathcal{X}$ while minimizing the worst-case communication. We write $\hat{C}_\infty(S^l)$ for the *simultaneous worst-case deterministic complexity* of l instances of a support set S, and the *worst-case deterministic amortized complexity* of S, written $\hat{A}_\infty(S)$, is a complexity measure representing the average communication per instance and given by the expression

$$\hat{A}_\infty(S) \triangleq \lim_{l\to\infty} \frac{1}{l}\hat{C}_\infty(S^l).$$

We write $\hat{C}_k(S^l)$ and $\hat{A}_k(S)$ when the number of rounds is bounded by k. Clearly, $\hat{C}_\infty(S^l) \leq l \cdot \hat{C}_\infty(S)$ and $\hat{A}_\infty(S) \leq \hat{C}_\infty(S)$. Deterministic amortized complexity for interactive communication has been studied by Naor, Orlitsky and Shor [14] and Alon and Orlitsky [17]. In the former paper, it is proven that the deterministic amortized complexity is equal to the complexity when $P_\mathcal{X}$ knows y in advance and that at most four rounds of communication are required to achieve the optimal solution. Ahlswede, Cai and Zhang [13] have subsequently reduced the number of rounds to three.

Result 7:

$$\hat{A}_3(S) = \hat{A}_4(S) = \cdots = \hat{A}_\infty(S) = \hat{C}^*(S) = \log \widehat{a_\mathcal{Y}}.$$

Example 8: We want to solve the league problem for two seasons, assuming that the results are independent. $P_\mathcal{Y}$ wants to learn the identity of the two champion teams from $P_\mathcal{X}$, who knows the two pairs of finalists. Obviously, if both seasons are solved independently, $2(\lceil \log n \rceil + 1)$ communication bits are required. However, by treating the two seasons as one larger problem, it can be shown [18] that $\hat{C}_2(S^2) \leq \lceil \log n \rceil + 6$. When the number of teams in the league is large, solving one or two instances requires roughly the same number of communication bits. Moreover, Result 7 implies that the deterministic amortized complexity of the league problem

is 1 bit per instance.

This paper examines the simultaneous resolution of several instances of interactive communication problems using nonzero-error randomized protocols: $P_{\mathcal{X}}$ and $P_{\mathcal{Y}}$ are allowed to toss private coins, and $P_{\mathcal{X}}$ must learn $(x_1, x_2, \ldots, x_l)$ correctly with probability at least $1 - \epsilon$. We write $\hat{R}^{\epsilon}_{\infty}(S^l)$ for the *simultaneous worst-case private coin randomized complexity* of l instances of a support set S, and the *worst-case private coin randomized amortized complexity* of S is given by the expression

$$\hat{A}^{\epsilon}_{\infty}(S) \triangleq \lim_{l\to\infty} \frac{1}{l} \hat{R}^{\epsilon}_{\infty}(S^l). \tag{1}$$

Again, we write $\hat{R}^{\epsilon}_{k}(S^l)$ and $\hat{A}^{\epsilon}_{k}(S)$ when the number of rounds is bounded by k.

F. Worst-Case Distributional Model

In all the complexity models defined so far, any interactive communication problem is completely described by its support set S. In effect, since we consider the communication in the worst case for all the possible input pairs (x, y), the probability distribution over the inputs is irrelevant.

In our last nonzero-error model, we allow deterministic protocols to fail with probability 1 for some pairs $(x, y) \in S$ as long as they are correct for most of the inputs. Let μ be a probability distribution over S. The *worst-case* (μ, ϵ)*-distributional complexity* of S, written $\hat{D}^{\mu,\epsilon}_{\infty}(S)$, is the number of bits transmitted in the worst case by the best deterministic protocol that allows $P_{\mathcal{Y}}$ to learn x for a fraction at least $1 - \epsilon$ of the inputs $(x, y) \in S$, weighted by μ. It is assumed that the players use a agreed-upon protocol based on μ.

The only model previously studied in the context of interactive communication and considering a probability distribution over the inputs it the zero-error average-case deterministic model. It has been studied by Orlitsky [19] and Alon and Orlitsky [20]. $P_{\mathcal{Y}}$ wants to learn x with certainty using a deterministic protocol and the goal is to minimize the expected number of bits that need to be transmitted.

III. Private Coin Randomized Interactive Communication: One Round of Communication is Almost Optimal

In this section, we study worst-case private coin randomized interactive communication. We show that a single round of communication is almost optimal, more precisely that

$$\hat{R}_1^\epsilon(S) \leq 4\hat{R}_\infty^\epsilon(S) + o(\hat{R}_\infty^\epsilon(S)).$$

Again, as mentioned in the introduction, this is quite different from worst-case deterministic interactive communication and even from worst-case randomized communication complexity of boolean functions.

Notice first that a randomized protocol can simulate a deterministic protocol by ignoring the output of the random generators, thus

$$\hat{R}_\infty^\epsilon(S) \leq \hat{C}_\infty(S). \tag{2}$$

A first lower bound, Lemma 9, shows that for any support set, the difference between private coin randomized complexity and one-way deterministic complexity is at most exponential.

Lemma 9:

$$\hat{R}_\infty^\epsilon(S) \in \Omega(\log \hat{C}_1(S)).$$

Proof:

The proof is similar to that of Lemma 3.8 by Kushilevitz and Nisan [3], but for interactive communication problems instead of boolean functions. Before proving the result, a few preliminary notions are needed. A randomized protocol $\mathcal{P}$ for a support set S can be represented by a binary tree. All the internal nodes of the tree are labeled by functions

$$f_{\mathcal{X},N} : X \times \{0,1\}^* \to \{0,1\}$$

or

$$f_{\mathcal{Y},N} : Y \times \{0,1\}^* \to \{0,1\},$$

and all its leaves are labeled by functions

$$f_{\mathcal{Y},L} : Y \times \{0,1\}^* \to X.$$

During the execution of $\mathcal{P}$ on input (x,y) with random strings $r_\mathcal{X}$ and $r_\mathcal{Y}$, the players travel down the tree following the values of $f_{\mathcal{X},N}$ and $f_{\mathcal{Y},N}$. When a node is labeled by a function

$f_{\mathcal{X},N}$, $P_{\mathcal{X}}$ computes $f_{\mathcal{X},N}(x, r_{\mathcal{X}})$; if $f_{\mathcal{X},N}(x, r_{\mathcal{X}}) = 0$, the next node is the left sibling of node N, and if $f_{\mathcal{X},N}(x, r_{\mathcal{X}}) = 1$, the protocol continues with the right sibling. Since $P_{\mathcal{Y}}$ does not know x nor $r_{\mathcal{X}}$, $P_{\mathcal{X}}$ must tell him which sibling to choose by transmitting him a bit. When a node is labeled by a function $f_{\mathcal{Y},N}(y, r_{\mathcal{Y}})$, $P_{\mathcal{Y}}$ executes the same process. The output of the protocol is the value $f_{\mathcal{Y},L}(y, r_{\mathcal{Y}}) \in X$ of the leaf reached at the end of the execution, and the worst-case cost of $\mathcal{P}$ is the height of the tree.

We now prove the lemma. Given a randomized protocol for S, we transform it into a one-way deterministic protocol. For each leaf L of the protocol tree, $P_{\mathcal{X}}$ sends $p_{\mathcal{X},L}$ to $P_{\mathcal{Y}}$, the probability over the choices of $r_{\mathcal{X}}$ to reach leaf L on input x, skipping all the nodes labeled by the functions $f_{\mathcal{Y},N}$. Each real number is sent using $p = -\log\left(\frac{1}{2} - \epsilon\right) + \hat{R}^{\epsilon}_{\infty}(S)$ bits, which means that the difference between the exact probability and its rounded value is at most 2^{-p}.

$P_{\mathcal{Y}}$ then computes $p_L = p_{\mathcal{X},L} \cdot p_{\mathcal{Y},L}$, the probability over the choices of $r_{\mathcal{X}}$ and $r_{\mathcal{Y}}$ to reach leaf L. Furthermore, using the functions $f_{\mathcal{Y},L} : Y \times \{0,1\}^* \to X$, $P_{\mathcal{Y}}$ computes for each leaf the probability to output x_i, for every $x_i \in a(y)$[1]. Adding the results for all the leaves, he also computes the probability over the choices of $r_{\mathcal{X}}$ and $r_{\mathcal{Y}}$ that $\mathcal{P}$ outputs x_i, for every $x_i \in a(y)$. Since $\mathcal{P}$ computes S with error at most ϵ, there is a single $x_i = x$ which is output with probability at least $1 - \epsilon$. The total rounding error is smaller than

$$\begin{aligned} 2^{\hat{R}^{\epsilon}_{\infty}(S)} \cdot 2^{-p} &= 2^{\hat{R}^{\epsilon}_{\infty}(S)} \cdot 2^{-\log\left(\frac{1}{2}-\epsilon\right)+\hat{R}^{\epsilon}_{\infty}(S)} \\ &= \frac{1}{2} - \epsilon, \end{aligned}$$

because the tree has at most $2^{\hat{R}^{\epsilon}_{\infty}(S)}$ leaves and because only the information sent by $P_{\mathcal{X}}$ has to be rounded. Even with the rounding error, the probability computed by $P_{\mathcal{Y}}$ for $x_i = x$ is more than $\frac{1}{2}$, hence $P_{\mathcal{Y}}$ concludes with certainty that the only $x_i \in a(y)$ with probability greater than $\frac{1}{2}$ is x.

The communication complexity of the protocol is

$$\hat{C}_1(S) \leq 2^{\hat{R}^{\epsilon}_{\infty}(S)} \cdot \left(\log \left(\frac{1}{2} - \epsilon \right)^{-1} + \hat{R}^{\epsilon}_{\infty}(S) \right),$$

thus

$$\hat{R}^{\epsilon}_{\infty}(S) \geq \log \hat{C}_1(S) - \log \hat{R}^{\epsilon}_{\infty}(S) - c(\epsilon). \tag{3}$$

[1]Recall that $a(y)$ is the ambiguity set of y.

The result follows.

■

Lemma 9 can be used to show that when the difference between one-way and two-way deterministic complexity is exponential, the deterministic and private coin randomized models are equivalent.

Corollary 10: If $\hat{C}_\infty(S) \sim {}^2 \log \hat{C}_1(S)$, then

$$\hat{C}_\infty(S) \sim \hat{R}^\epsilon_\infty(S).$$

If $\hat{C}_\infty(S) \in \Theta(\log \hat{C}_1(S))$, then

$$\hat{C}_\infty(S) \in \Theta(\hat{R}^\epsilon_\infty(S)).$$

Proof: Combining (2) and (3), we get

$$\log \hat{C}_1(S) - \log \hat{C}_\infty(S) - c(\epsilon) \leq \hat{R}^\epsilon_\infty(S) \leq \hat{C}_\infty(S).$$

If $\hat{C}_\infty(S) \sim \log \hat{C}_1(S)$, then asymptotically,

$$\hat{C}_\infty(S) - \log \hat{C}_\infty(S) - c(\epsilon) \leq \hat{R}^\epsilon_\infty(S) \leq \hat{C}_\infty(S),$$

thus $\hat{C}_\infty(S) \sim \hat{R}^\epsilon_\infty(S)$. The second case is proved in a similar way. ■

Example 11: Recall that for the league problem presented in Example 1, $\hat{C}_\infty(L) = \lceil \log n \rceil + 1$ and $\hat{C}_1(L) = n$. Using Corollary 10, it follows that $\hat{R}^\epsilon_\infty(L) \sim \log n$; the worst-case deterministic and private coin randomized models are thus equivalent for this problem.

Lemma 9 can also be used to show that the worst-case deterministic and private coin randomized are equivalent when P_y's ambiguity is constant.

Corollary 12: If $\widehat{a_y} \in O(1)$, then

$$\hat{C}_\infty(S) \sim \hat{R}^\epsilon_\infty(S).$$

Proof: From (2) and (3), it follows that

$$\log \hat{C}_1(S) - \log \hat{R}^\epsilon_\infty(S) - c(\epsilon) \leq \hat{R}^\epsilon_\infty(S) \leq \hat{C}_\infty(S).$$

Moreover,

$$\hat{C}_\infty(S) \leq \log \log \chi + O(1) \leq \log \hat{C}_1(S) + O(1)$$

[2] $f(n) \sim g(n)$ if and only if $\lim_{n\to\infty} \frac{f(n)}{g(n)} = 1$

from the assumption and Results 4 and 5, thus

$$\hat{C}_\infty(S) - \log \hat{R}^\epsilon_\infty(S) - O(1) \leq \hat{R}^\epsilon_\infty(S) \leq \hat{C}_\infty(S).$$

■

The next lemma tightens the bound from Lemma 9 for problems with small discrepancy between $\hat{C}_1(S)$ and $\hat{C}_\infty(S)$.

Lemma 13:

$$\hat{R}^\epsilon_\infty(S) \geq \log \widehat{a_\mathcal{Y}} - \log \frac{1}{1-\epsilon}.$$

Proof: Suppose that $\hat{R}^\epsilon_\infty(S) < \log \widehat{a_\mathcal{Y}} - \log \frac{1}{1-\epsilon}$. By definition, there exists a protocol $\mathcal{P}$ for S requiring less than $\log \widehat{a_\mathcal{Y}} - \log \frac{1}{1-\epsilon}$ bits of communication and such that for every pair $(x, y) \in S$, the probability that $P_\mathcal{Y}$ does not learn x correctly is at most ϵ.

Let $y \in Y$ be an input such that $|a(y)| = \widehat{a_\mathcal{Y}}$. For every choice of the random strings $r_\mathcal{X}$ and $r_\mathcal{Y}$, less than $(1-\epsilon) \cdot \widehat{a_\mathcal{Y}}$ distinct messages can be transmitted and it follows that the protocol makes an error for more than $\epsilon \cdot \widehat{a_\mathcal{Y}}$ of the $x_i \in a(y)$. By a simple counting argument, there is at least one element $x' \in a(y)$ such that the error probability of $\mathcal{P}$ on the input pair (x', y) is more than ϵ, which is a contradiction. ■

The previous bound is very weak for problems for which the difference between one-way and two-way worst-case deterministic complexity is large. For the league problem, $\widehat{a_\mathcal{Y}} = 2$ and Lemma 13 gives $\hat{R}^\epsilon_\infty(L) \geq 1$. On the other hand, for balanced and symmetric pairs, it will be shown in Section VII that the bound is tight. In fact, the two lower bounds presented in this section are asymptotically equal to Results 3 and 4 for deterministic protocols. The next lemma gives a strong upper bound on the private coin randomized complexity.

Lemma 14:

$$\hat{R}^\epsilon_1(S) \leq 2 \left\lceil \log\log \chi + \log \widehat{a_\mathcal{Y}} + \log\left(\frac{1}{\epsilon} - 1\right) + 1 \right\rceil.$$

Proof: Let G_S be the characteristic hypergraph of S, and let x be $P_\mathcal{X}$'s input. Recall that $P_\mathcal{Y}$ has an input y defining the ambiguity set (hyperedge) $a(y) = \{x_1, \ldots, x_l\}$ and wants to learn which of the x_i is x. Let $k \triangleq \lceil \log \chi \rceil - 1$. Players agree on a proper coloring Ψ of G_S with χ colors. We write the color of the vertices s of G_S in binary: $\Psi_s \triangleq s_0 s_1 \ldots s_k$, where $s_i \in \{0, 1\}$. We consider these strings as polynomials in $\mathbb{Z}_p$, where p is a prime number such that

$$\left\lfloor \frac{1-\epsilon}{\epsilon} \cdot k \cdot (\widehat{a_\mathcal{Y}} - 1) \right\rfloor < p \leq 2 \left\lfloor \frac{1-\epsilon}{\epsilon} \cdot k \cdot (\widehat{a_\mathcal{Y}} - 1) \right\rfloor.$$

Bertrand's postulate [21] guarantees the existence of such a p. In other words, for a vertex s of G_S whose color is Ψ_s, $\Psi_s(t) \equiv (s_0 + s_1 t + \cdots + s_k t^k) (\text{mod } p)$.

$P_{\mathcal{X}}$ randomly chooses $m \in \mathbb{Z}_p$ and sends m and $\Psi_x(m)$ to $P_{\mathcal{Y}}$. $P_{\mathcal{Y}}$ then constructs the set

$$E_m = \{x_i \mid x_i \in a(y) \wedge \Psi_{x_i}(m) = \Psi_x(m)\},$$

randomly chooses an element of E_m and concludes that it is $P_{\mathcal{X}}$'s input.

In order to show that the protocol works, we have to prove that the probability that $P_{\mathcal{Y}}$ answers x is at least $1 - \epsilon$. First, remark that $x \in E_m$, therefore $|E_m| \geq 1$ and $P_{\mathcal{Y}}$ can only answer incorrectly when $|E_m| \geq 2$. Thus, we get

$$\begin{aligned}
\Pr[P_{\mathcal{Y}} \text{ answers } x] &= \sum_{j \in \mathbb{Z}_p} \Pr[P_{\mathcal{X}} \text{ randomly chooses } j \in \mathbb{Z}_p \wedge P_{\mathcal{Y}} \text{ guesses correctly}] \\
&= \frac{1}{p} \sum_{j=0}^{p-1} \Pr[P_{\mathcal{Y}} \text{ answers } x \mid j \text{ is chosen}] \\
&= \frac{1}{p} \sum_{j=0}^{p-1} \frac{1}{|E_j|}.
\end{aligned}$$

Furthermore, since Ψ is a proper coloring of G_S, all the vertices of the hyperedge $a(y)$ have a different color, thus the corresponding polynomials $\Psi_s(t)$ are different. It implies that two such polynomials can be equal for at most k field elements because their difference, a nonzero polynomial of degree at most k, has at most k roots. Hence, we can deduce that

$$\sum_{j=0}^{p-1} |E_j| \leq k \cdot (\widehat{a_{\mathcal{Y}}} - 1) + p,$$

which implies that

$$\sum_{j=0}^{p-1} \frac{1}{|E_j|} \geq \frac{p}{\left(\frac{k \cdot (\widehat{a_{\mathcal{Y}}} - 1) + p}{p}\right)},$$

and thus

$$\begin{aligned}
\Pr[P_{\mathcal{Y}} \text{ answers } x] &\geq \frac{1}{\frac{k \cdot (\widehat{a_{\mathcal{Y}}} - 1)}{p} + 1} \\
&\geq \frac{1}{\frac{k \cdot (\widehat{a_{\mathcal{Y}}} - 1)}{k \cdot (\widehat{a_{\mathcal{Y}}} - 1) \cdot \frac{1-\epsilon}{\epsilon}} + 1} \\
&= \frac{1}{\frac{\epsilon}{1-\epsilon} + 1} = \frac{1}{\left(\frac{1}{1-\epsilon}\right)} \\
&= 1 - \epsilon.
\end{aligned}$$

Finally, the complexity of the protocol is

$$\begin{aligned}\hat{R}_1^\epsilon(S) &\leq 2\lceil \log p \rceil \\ &\leq 2\left\lceil 1 + \log\left(\frac{1}{\epsilon} - 1\right) + \log\log\chi + \log\widehat{a_{\mathcal{Y}}}\right\rceil .\end{aligned}$$

■

Using Lemma 14 with the league problem, we get $\hat{R}_1^\epsilon(L) \leq 2\lceil \log n \rceil + c(\epsilon)$. More generally, the lemma can be used to prove that the one-way private coin randomized complexity is at most four times larger than the worst-case deterministic complexity.

Corollary 15:

$$\hat{R}_1^\epsilon(S) \leq 4\hat{C}_\infty(S) + \left\lceil 2\log\frac{1}{\epsilon}\right\rceil .$$

Proof: Using Lemma 14 and Results 3 and 4, we get

$$\begin{aligned}\hat{R}_1^\epsilon(S) &\leq 2\left\lceil 1 + \log\left(\frac{1}{\epsilon} - 1\right) + \log\hat{C}_1(S) + \hat{C}_\infty(S)\right\rceil \\ &\leq 2\left\lceil 1 + \log\left(\frac{1}{\epsilon} - 1\right) + \hat{C}_\infty(S) - 1 + \hat{C}_\infty(S)\right\rceil \\ &\leq 4\hat{C}_\infty(S) + 2\left\lceil \log\frac{1}{\epsilon}\right\rceil .\end{aligned}$$

■

Combining Lemma 14 with the two lower bounds proved in this section, we can conclude that one round of communication is almost optimal for worst-case private coin randomized protocols.

Corollary 16:

$$\hat{R}_1^\epsilon(S) \leq 4\hat{R}_\infty^\epsilon(S) + o(\hat{R}_\infty^\epsilon(S)).$$

Proof: From Result 4 and Lemmas 9, 13 and 14 it follows that

$$\begin{aligned}\hat{R}_1^\epsilon(S) &\leq 2\log\hat{C}_1(S) + 2\log\widehat{a_{\mathcal{Y}}} + c_1(\epsilon) \\ &\leq 4\hat{R}_\infty^\epsilon(S) + 2\log\hat{R}_\infty^\epsilon(S) + c_2(\epsilon) \\ &\leq 4\hat{R}_\infty^\epsilon(S) + o(\hat{R}_\infty^\epsilon(S)).\end{aligned}$$

■

IV. Public Coin Randomized Interactive Communication

In this section, we characterize worst-case public coin randomized interactive communication. Firstly, we prove that one round of communication is optimal for this model; the complexity corresponds to the amortized complexity and to the communication required when $P_{\mathcal{Y}}$ knows y in advance. Secondly, we give an upper bound on the difference between the private coin and public coin randomized models and use it to improve the upper bounds on the one-way private coin complexity derived in the previous section. Notice first that the players can use the concatenation of the private strings $r_{\mathcal{X}}$ and $r_{\mathcal{Y}}$ as a public random string, thus the public coin randomized model is at least as powerful as the private coin randomized model, i.e.,

$$\hat{R}_k^{\epsilon,pub}(S) \leq \hat{R}_k^{\epsilon}(S).$$

A. One Round of Communication is Optimal

Lemma 17:

$$\log \widehat{a_{\mathcal{Y}}} - \log \frac{1}{1-\epsilon} \leq \hat{R}_{\infty}^{\epsilon,pub}(S) \leq \hat{R}_1^{\epsilon,pub}(S) \leq \lceil \log(\widehat{a_{\mathcal{Y}}} - 1) \rceil + \left\lceil \log \frac{1-\epsilon}{\epsilon} \right\rceil .$$

Proof:

Lemma 13 can be applied with a public random generator for the lower bound. For the upper bound, recall that $P_{\mathcal{X}}$ has an input x and $P_{\mathcal{Y}}$ has an input y defining the ambiguity set (hyperedge of the characteristic hypergraph) $a(y) = \{x_1, \ldots, x_l\} = \{x \mid (x, y) \in S\}$. It is assumed that the elements of X are coded using $\lceil \log |X| \rceil$ bits. $P_{\mathcal{X}}$ chooses k random subsets of the bits of his input x and, for each of these subsets, sends the parity of the bits to $P_{\mathcal{Y}}$. $P_{\mathcal{Y}}$ then computes the k corresponding parities for each of the inputs $x_i \in a(y)$. Each time a parity differs from the corresponding result sent by $P_{\mathcal{X}}$, $P_{\mathcal{Y}}$ deduces that $x_i \neq x$ and discards it. When $P_{\mathcal{Y}}$ has done all the comparisons, he randomly chooses an input among those which have not been discarded and concludes that it is $P_{\mathcal{X}}$'s input.

The probability of not discarding a vertex x_i is 1 if $x_1 = x$ and $\frac{1}{2^k}$ otherwise. Let Z be a random variable representing the number of inputs not discarded after k iterations. Since there are $a(y) - 1$ inputs to discard, it follows that

$$E(Z) = 1 + \frac{1}{2^k}(a(y) - 1) \leq 1 + \frac{1}{2^k}(\widehat{a_{\mathcal{Y}}} - 1).$$

The probability that $P_{\mathcal{Y}}$ learns $P_{\mathcal{X}}$'s input correctly is $E\left[\frac{1}{Z}\right]$, and Jensen's inequality gives

$$\Pr[P_{\mathcal{Y}} \text{ answers } x] \geq \frac{1}{E[Z]} = \frac{1}{1 + \frac{1}{2^k}(\widehat{a_{\mathcal{Y}}} - 1)}.$$

By letting $k = \lceil \log(\widehat{a_{\mathcal{Y}}} - 1) \rceil + \left\lceil \log \frac{1-\epsilon}{\epsilon} \right\rceil$, it follows that

$$\begin{aligned} \Pr[\text{success}] &\geq \frac{1}{1 + \frac{1}{2^{\lceil \log(\widehat{a_{\mathcal{Y}}} - 1) \rceil + \left\lceil \log \frac{1-\epsilon}{\epsilon} \right\rceil}} \cdot (\widehat{a_{\mathcal{Y}}} - 1)} \\ &\geq \frac{1}{1 + \frac{1}{\frac{1-\epsilon}{\epsilon}}} \\ &= 1 - \epsilon, \end{aligned}$$

from which we conclude that

$$R^1_{\epsilon,pub}(S) \leq k = \lceil \log(\widehat{a_{\mathcal{Y}}} - 1) \rceil + \left\lceil \log \frac{1-\epsilon}{\epsilon} \right\rceil.$$

■

Example 18: If we apply Lemma 17 to the league problem, we get $\hat{R}_1^{\epsilon,pub}(L) \leq \left\lceil \log \frac{1-\epsilon}{\epsilon} \right\rceil \in \Theta(1)$. It shows that one-way public coin randomized complexity can be arbitrarily better than one-way deterministic complexity.

B. Difference Between the Private Coin and Public Coin Models

Private coin randomized complexity cannot be much worse than public coin randomized complexity: every public randomized protocol can be transformed into a private randomized protocol whose error probability is slightly larger, and which uses a few more communication bits. It is inspired by a similar result for boolean functions shown by Newman [22].

Theorem 19: For all $\delta > 0$ and for all $\epsilon > 0$ such that $\epsilon + \delta < 1$,

$$\hat{R}_1^{\epsilon+\delta}(S) \leq \hat{R}_\infty^{\epsilon,pub}(S) + \log\log|S| + \log\frac{1}{\delta^2} + 1.$$

Proof: Let $\mathcal{P}$ be a public coin randomized protocol for S whose error is bounded by ϵ and requiring $\hat{R}_\infty^{\epsilon,pub}(S)$ communication bits. We suppose the random generator r follows a probability distribution μ. Let $Z(x,y,r)$ be a random variable equal to 1 if the answer given by $P_{\mathcal{Y}}$ following the execution of $\mathcal{P}$ on input (x,y) is incorrect (different from x), and equal to 0 if correct. Since $\mathcal{P}$ solves S with error at most ϵ, it follows that $\underset{r\in\mu}{E}[Z(x,y,r)] \leq \epsilon$ for every pair $(x,y) \in S$.

A new protocol for S using less random bits is designed. Let t be a parameter to be fixed later, and let $r_1, r_2, \ldots, r_t$ be binary strings. The protocol $\mathcal{P}_{r_1,r_2,\ldots,r_t}$ is defined as follows: $P_{\mathcal{X}}$ and $P_{\mathcal{Y}}$ randomly choose i between 1 and t and run protocol $\mathcal{P}$ with the common random string r_i.

We now show that there exists strings $r_1, r_2, \ldots, r_t$ such that $\underset{i}{E}[Z(x,y,r_i)] \leq \epsilon + \delta$ for every pair $(x,y) \in S$. We choose the t strings $r_1, r_2, \ldots, r_t$ randomly following the probability distribution μ, consider an arbitrary pair $(x,y) \in S$ and compute the probability that $\underset{i}{E}[Z(x,y,r_i)] > \epsilon + \delta$ (where i is uniformly distributed). This is equivalent to the probability that $\frac{1}{t}\sum_{i=1}^{t} Z(x,y,r_i) > \epsilon + \delta$. Since $\underset{r\in\mu}{E}[Z(x,y,r)] \leq \epsilon$, Chernoff bound yields

$$\Pr_{r_1,\ldots,r_t}\left[\frac{1}{t}\sum_{i=1}^{t} Z(x,y,r_i) - \epsilon > \delta\right] \leq 2e^{-2\delta^2 t}.$$

By choosing $t = \left\lceil \frac{\log|S|}{\delta^2} \right\rceil$, it follows that

$$\begin{aligned} 2e^{-2\delta^2 t} &= 2e^{-2\delta^2\left\lceil \frac{\log|S|}{\delta^2}\right\rceil} \\ &\leq 2e^{-2\log|S|} \\ &= 2\cdot 2^{-2\log e \log|S|} \\ &= 2|S|^{-2\log e} \\ &< \frac{1}{|S|} \text{ when } |S| > 1. \end{aligned}$$

Thus, for a random choice of $r_1, \ldots, r_t$, the probability that there exists at least a pair $(x,y) \in S$ (there are $|S|$ such pairs) such that $\underset{i}{E}[Z(x,y,r_i)] > \epsilon+\delta$ is smaller than $|S|\cdot\frac{1}{|S|} = 1$. Consequently, there exists a choice of $r_1, \ldots, r_t$ such that for every pair $(x,y) \in S$, the error of protocol $\mathcal{P}_{r_1,r_2,\ldots,r_t}$ is at most $\epsilon+\delta$. The number of random bits used by $\mathcal{P}_{r_1,r_2,\ldots,r_t}$ is $\lceil \log t \rceil$, and in order to transform the public protocol into a private protocol, $P_{\mathcal{X}}$ has to randomly choose i between 1 and t and to send it to $P_{\mathcal{Y}}$. Moreover, from Lemma 17, there exists an optimal one-round public coin randomized protocol for S ensuring that $\mathcal{P}_{r_1,r_2,\ldots,r_t}$ is also a one-way protocol. Hence,

$$\begin{aligned} \hat{R}_1^{\epsilon+\delta}(S) &\leq \hat{R}_1^{\epsilon,pub}(S) + \lceil \log t \rceil \\ &\leq \hat{R}_1^{\epsilon,pub}(S) + \left\lceil \log \left\lceil \frac{\log|S|}{\delta^2} \right\rceil \right\rceil \\ &\leq \hat{R}_1^{\epsilon,pub}(S) + \log\log|S| + \log\frac{1}{\delta^2} + 1. \end{aligned}$$

■

Example 20: The previous theorem is applied to the league problem, for which $|S| = (2^n) \cdot (2^n - 1)$. This gives

$$\begin{aligned} \hat{R}_1^{\epsilon'+\delta}(L) &\leq \hat{R}_\infty^{\epsilon,pub}(S) + \log\log|S| + \log\frac{1}{\delta^2} + 1 \\ &\leq \log\log|2^n(2^n-1)| + c(\epsilon',\delta), \end{aligned}$$

thus $\hat{R}_1^{\epsilon}(L) \leq \log n + c(\epsilon)$. There exists an optimal one-way private coin randomized protocol for this problem; it improves the bound given by Lemma 14, which is not optimal, and the result from Example 11, which does not limit interaction between the players. This example also proves that the bound given by Theorem 19 can be reached.

Theorem 19 can also be used to tighten the upper bounds on the one-way private coin randomized complexity presented in Section III.

Corollary 21: For all $\delta > 0$ and for all $\epsilon > 0$ such that $\epsilon + \delta < 1$,

$$\hat{R}_1^{\epsilon+\delta}(S) \leq \lceil \log(\widehat{a_{\mathcal{Y}}} - 1) \rceil + \log\log|S| + \left\lceil \log\frac{1-\epsilon}{\epsilon} \right\rceil + \log\frac{1}{\delta^2} + 1.$$

Corollary 22:

$$\hat{R}_1^{\epsilon}(S) \leq \log\log\sigma + \log\widehat{a_{\mathcal{Y}}} + \log\log\widehat{a_{\mathcal{Y}}} + c(\epsilon).$$

Proof: Recall that σ is the number of hyperedges in the characteristic hypergraph G_S. Suppose that for each $y \in Y$ corresponds a different hyperedge; if is not the case, the players agree on an equivalent support set $S' \subseteq X' \times Y'$, $S' \subset S$, such that for each $y \in Y'$ corresponds a distinct hyperedge. It follows that $|S'| \leq |\mathcal{Y}'| \cdot \widehat{a_{\mathcal{Y}}}' \leq \sigma \cdot \widehat{a_{\mathcal{Y}}}$, and the result follows from Theorem 19 and Lemma 17. ■

If $X = Y = \{0,1\}^n$, the difference between the private and public coin randomized models is at most an additive term of $\log n + O(1)$ bits. Using Result 5, the same thing can be said about the difference between the deterministic and public coin randomized models. In fact, the similarity between Result 5 and Corollary 22 is striking considering how different the models and the proofs are.

Corollary 23:

$$\hat{R}_1^{\epsilon}(S) \leq 3\hat{R}_\infty^{\epsilon}(S) + o(\hat{R}_\infty^{\epsilon}(S)).$$

Proof: Since the hypergraph G_S can be colored with χ colors, there exists a support set S' equivalent to S having at most $\sum_{i=1}^{\widehat{a_{\mathcal{Y}}}} \binom{\chi}{\widehat{a_{\mathcal{Y}}}} \leq \chi^{\widehat{a_{\mathcal{Y}}}}$ distinct hyperedges. The result follows from Corollary 22 and Lemmas 9 and 13. ■

Corollary 24:

$$\hat{R}_1^\epsilon(S) \leq 3\hat{C}_\infty(S) + o(\hat{C}_\infty(S)).$$

V. Private Coin Randomized Amortized Interactive Communication: One Round of Communication is Optimal

In this section, we show that one round of communication is optimal for the private coin randomized amortized model; the complexity is equal to the number of bits that need to be transmitted when $P_{\mathcal{X}}$ knows y in advance. This is an improvement over the deterministic amortized model, for which interaction is required in order to minimize the communication.

Lemma 25:

$$\hat{A}_1^\epsilon(S) = \hat{A}_2^\epsilon(S) = \cdots = \hat{A}_\infty^\epsilon(S) = \log \widehat{a_{\mathcal{Y}}}.$$

Proof:

A mentioned in Section II, l independent instances of a support set $S \subseteq X \times Y$ can be treated as a larger support set S^l. Let G_S^l be the characteristic hypergraph of S^l. It is not hard to show that the vertices of G_S^l are the elements of X^l, and that for each l-tuple $(e_1, e_2, \ldots, e_l)$ of hyperedges of G_S, $e_1 \times e_2 \times \cdots \times e_l$ is a hyperedge of G_S^l. Clearly, the maximum ambiguity of $P_{\mathcal{Y}}$ for the support set S^l is $\widehat{a_{\mathcal{Y}}}^l$, and the number of different hyperedges of G_S^l is σ^l. It follows from Lemma 13 that

$$\hat{R}_\infty^\epsilon(S^l) \geq l \cdot \log \widehat{a_{\mathcal{Y}}} - \log \frac{1}{1-\epsilon},$$

and Corollary 22 gives

$$\hat{R}_1^\epsilon(S^l) \leq \log\log\sigma + l \cdot \log \widehat{a_{\mathcal{Y}}} + \log\log \widehat{a_{\mathcal{Y}}} + 2\log l + c(\epsilon).$$

The result follows from the definition of $\hat{A}_1^\epsilon(S)$ and $\hat{A}_\infty^\epsilon(S)$ given by Eq. (1). ■

Lemma 25 shows that when several instances are solved simultaneously instead of sequentially, there is no advantage to use a public coin over private coins since no interaction is required and the amount communication is the same. Also, comparing Lemma 25 and Corollary 21 when $|S| \in O(2^{cn})$ (this includes the support sets $S \subseteq \{0,1\}^n \times \{0,1\}^n$), the difference between the private coin randomized complexity and the private coin randomized amortized complexity is at most an additive term of $\log n + O(1)$ bits. The bound is tight for the league problem. The same discrepancy between the deterministic complexity and the deterministic amortized complexity is implicit in the work of Naor, Orlitsky and Shor [14] using Results 5 and 7.

VI. Distributional Interactive Communication: One Round of Communication is Optimal

In this section, we study worst-case distributional complexity, starting with yet another example using the league problem.

Example 26: Suppose that the support set L follows a uniform distribution, and without loss of generality, that n, the length of the team names, is an even integer. $P_{\mathcal{X}}$ sends two bits to P_Y: the parity of the first $\frac{n}{2}$ bits of the champion's name, followed by the parity of the last $\frac{n}{2}$ bits. $P_{\mathcal{Y}}$ computes the equivalent parities for the two teams playing for the championship and compares the results with the bits received from P_X. The protocol will err for the pairs of finalist teams with identical parity bits, and there are $|S| \cdot \left(\frac{1}{4} - \frac{1}{2^n}\right)$ such pairs. Hence, $\hat{D}_1^{uniform,\frac{1}{4}}(S) \leq 2$.

It should be noted that for every interactive communication problem and for all $\epsilon > 0$, there exists a probability distribution over the inputs for which the distributional complexity is 0, for example if there is an input pair occurring with probability at least $1 - \epsilon$. Consequently, it is interesting to consider probability distributions maximizing the number of bits that need to be transmitted, and even in this case communication can be more efficient than when no error is allowed. The main result of this section is that the distributional complexity with the worst possible probability distribution over the inputs is equal to the public coin randomized complexity. Again, it is equivalent to the complexity when $P_{\mathcal{X}}$ knows y in advance. This is hardly a surprise, since the equivalence between the randomized and distributional models was first established by Yao [23] for computational complexity using von-Neumann's Minimax Theorem of game theory [24]. The equality is also verified for communication complexity of boolean functions (see, for example, [3]), and the proof can be applied without modification to partial domains $S \subseteq X \times Y$. We present a simpler proof tailored for interactive communication problems and improving the original proof in two ways. Firstly, a family of asymptotically "worst" probability distributions over the inputs is described. The distributions can be used with any interactive communication problem and probability of error. Secondly, an optimal protocol requiring a single round of communication is constructed.

Lemma 27:

$$\log \widehat{a_{\mathcal{Y}}} - \log \frac{1}{1-2\epsilon} \leq \max_{\mu} \hat{D}_{\infty}^{\mu,\epsilon}(S) \leq \max_{\mu} \hat{D}_{1}^{\mu,\epsilon}(S) \leq \lceil \log(\widehat{a_{\mathcal{Y}}} - 1) \rceil + \left\lceil \log \frac{1-\epsilon}{\epsilon} \right\rceil .$$

Proof: Let $S \subseteq X \times Y$ be an interactive communication problem. To prove the lower bound, the following probability distribution μ' over S is used: there is a $y' \in Y$ with $|a(y')| = \widehat{a_{\mathcal{Y}}}$[3], and $\Pr[(x,y) = (x_i, y')] = \frac{1}{2\widehat{a_{\mathcal{Y}}}}$ for every $x_i \in a(y)$. The probability of all the other input pairs $(x,y) \in S$ can be any nonzero value.

Suppose that $\hat{D}_{\infty}^{\mu',\epsilon}(S) < \log \widehat{a_{\mathcal{Y}}} - \log \frac{1}{1-2\epsilon}$, and let $\mathcal{P}$ be an optimal protocol for S. For any pair $(x,y) \in S$, less than $\widehat{a_{\mathcal{Y}}} \cdot (1-2\epsilon)$ distinct messages can be exchanged between $P_{\mathcal{X}}$ and $P_{\mathcal{Y}}$. It follows that when $P_{\mathcal{Y}}$ has the input y', the protocol fails on more than $2\epsilon \cdot \widehat{a_{\mathcal{Y}}}$ input pairs (x_i, y'), each pair occurring with probability $\frac{1}{2\widehat{a_{\mathcal{Y}}}}$. Hence, $\mathcal{P}$ fails with probability more than ϵ over S, which is a contradiction, and

$$\log \widehat{a_{\mathcal{Y}}} - \log \frac{1}{1-2\epsilon} \leq \hat{D}_{\infty}^{\mu',\epsilon}(S) \leq \max_{\mu} \hat{D}_{\infty}^{\mu,\epsilon}(S).$$

To prove the upper bound, the public coin protocol presented in Lemma 17 is used. Recall that by receiving the parity of $k = \lceil \log(\widehat{a_{\mathcal{Y}}} - 1) \rceil + \left\lceil \log \frac{1-\epsilon}{\epsilon} \right\rceil$ random subsets of the bits of x from $P_{\mathcal{X}}$, $P_{\mathcal{Y}}$ can learn x with probability at least $1-\epsilon$ for every valid input pair. Since the protocol works for any probability distribution over the inputs, it succeeds with probability at least $1-\epsilon$ when the probability distribution is taken over the inputs and the choice of the subsets. By a simple counting argument, it follows that there exists k subsets of the bits of x for which the protocol succeeds for at least a fraction $1-\epsilon$ of the inputs weighted by their probability distribution.

The protocol from Lemma 17 is derandomized as follows: for a fixed μ and a fixed ϵ, $P_{\mathcal{X}}$ and $P_{\mathcal{Y}}$ agree on k subsets for which the protocol fails on at most a fraction ϵ of S weighted by μ. $P_{\mathcal{X}}$ sends the parity of the k subsets of x to $P_{\mathcal{Y}}$, who compares the received bits with the k corresponding parities for each of the inputs $x_i \in a(y)$. If more than one of the x_i is not discarded, $P_{\mathcal{Y}}$ chooses the one whose input probability given y is the highest. The complexity of the protocol is

$$\max_{\mu} \hat{D}_1^{\mu,\epsilon}(S) \leq \hat{R}_1^{\epsilon,pub}(S) \leq k \leq \lceil \log(\widehat{a_{\mathcal{Y}}} - 1) \rceil + \left\lceil \log \frac{1-\epsilon}{\epsilon} \right\rceil.$$

■

The family of "worst" probability distributions presented in the previous proof is not unique. For example, when the ambiguity of every input $y \in Y$ is maximal, i.e., $|a(y)| = \widehat{a_{\mathcal{Y}}}$, it is not

[3]Such a y' always exists.

hard to show that the uniform distribution also maximizes the required number of communication bits. The league problem is an example of such a support set: $P_{\mathcal{Y}}$ always picks the winner among two teams.

VII. Equivalence of All the Models for Balanced and Symmetric Pairs

All the worst-case complexity models introduced in this paper require asymptotically at least $\log \widehat{a_{\mathcal{Y}}}$ bits of communication, and with the exception of the deterministic and private coin randomized models, asymptotically $\log \widehat{a_{\mathcal{Y}}}$ bits suffice. For some support sets like the league problem, the difference between $\log \widehat{a_{\mathcal{Y}}}$ and $\hat{C}_{\infty}(S)$ is large, but for other problems, the amortized or nonzero-error models cannot be used to reduce the number of communication bits. Consider the following modification of the league problem: either $P_{\mathcal{Y}}$ knows the two teams playing for the league championship, or he doesn't know anything. It is obvious that in the worst case, whether it is the worst input of the worst probability distribution over the inputs, P_X has to send asymptotically all the bits of the champion's name to $P_{\mathcal{Y}}$ and he can do so in a single round of communication. For the problems mentioned in the next lemma, all the worst-case models presented in this paper are asymptotically equivalent and inefficient: randomization, nonzero-error protocols, solving several instances simultaneously, knowing y in advance and even interaction cannot significantly reduce the communication between $P_{\mathcal{X}}$ and $P_{\mathcal{Y}}$.

Lemma 28: If there is a $y \in Y$ with $|a(y)| = \widehat{a_{\mathcal{Y}}} = |X|$, or if S is a Cartesian-product support set, then

$$\hat{C}_1(S) = \lceil \log \widehat{a_{\mathcal{Y}}} \rceil \, .$$

Proof: Result 3 gives a lower bound of $\lceil \log \widehat{a_{\mathcal{Y}}} \rceil$. If $\widehat{a_{\mathcal{Y}}} = |X|$, then P_X has to describe x completely; he can do so using $\lceil \log |X| \rceil = \lceil \log \widehat{a_{\mathcal{Y}}} \rceil$ bits. If S is a Cartesian-product support set, then there exists $X' \in X$ and $Y' \in Y$ such that $S = X' \times Y'$. Again, $\widehat{a_{\mathcal{Y}}} = |X'|$ and $P_{\mathcal{X}}$ can describe x with $\lceil \log |X'| \rceil = \lceil \log \widehat{a_{\mathcal{Y}}} \rceil$ bits of communication. ■

It is also possible to prove that all the worst-case complexity models are equivalent for balanced and symmetric pairs. Recall that a support set S is balanced if $\widehat{a_{\mathcal{X}}} = \widehat{a_{\mathcal{Y}}}$ and symmetric if $(x, y) \in S$ if and only if $(y, x) \in S$.

Theorem 29: Let S be a balanced support set, and let $0 < \epsilon < 1$. The deterministic, amortized deterministic, deterministic when $P_{\mathcal{X}}$ knows P_Y's input, private coin randomized, public coin randomized, distributional and private coin randomized amortized models are equivalent. More

precisely,

$$\hat{C}^*(S) \leq \quad \hat{C}_3(S) \quad \leq \hat{C}^*(S) + o(\hat{C}^*(S)); \tag{4}$$

$$\hat{C}^*(S) - 1 \leq \quad \hat{A}_3(S) \quad \leq \hat{C}^*(S); \tag{5}$$

$$\hat{C}^*(S) - O(1) \leq \quad \hat{R}_3^{\epsilon}(S) \quad \leq \hat{C}^*(S) + o(\hat{C}^*(S)); \tag{6}$$

$$\hat{C}^*(S) - O(1) \leq \quad \hat{R}_1^{\epsilon,pub}(S) \quad \leq \hat{C}^*(S) + O(1); \tag{7}$$

$$\hat{C}^*(S) - O(1) \leq \quad \hat{D}_1^{\mu,\epsilon}(S) \quad \leq \hat{C}^*(S) + O(1); \tag{8}$$

$$\hat{C}^*(S) - 1 \leq \quad \hat{A}_1^{\epsilon}(S) \quad \leq \hat{C}^*(S). \tag{9}$$

Proof: Inequality (4) can be deduced from Results 3 and 6; Inequality (5) from Results 3 and 7; Inequality (6) from Results 3 and 6, Inequality (2) and Lemma 13; Inequality (7) from Result 3 and Lemma 17; Inequality (8) from Result 3 and Lemma 27; Inequality (9) from Result 3 and Lemma 25. ■

Unfortunately, Theorem 29 means that nonzero-error algorithms cannot significantly reduce the communication for all the practical applications mentioned in the introduction. This is a somewhat "negative" result, but only because deterministic protocols are already very efficient. Nevertheless, nonzero-error models, even the private coin randomized model, allow efficient one-way protocols. For most practical applications of symmetric pairs, the ambiguity of the players increases at least polynomially with the size of the inputs; in this case there exists a private coin randomized protocol using a single round of communication and whose communication complexity is almost optimal. If the ambiguity of the players increases superpolynomially with the size of the inputs, then the best one-round private coin randomized protocol is optimal.

Lemma 30: Let S be a balanced support set with $X = \{0,1\}^n$, and let $0 < \epsilon < 1$ and $k \geq 1$. If $\widehat{a_{\mathcal{Y}}} \in \Theta(n^k)$, then

$$\hat{C}^*(S) - O(1) \leq \hat{R}_1^{\epsilon}(S) \leq \left(1 + \frac{1}{k}\right) \cdot \hat{C}^*(S) + o(\hat{C}^*(S)).$$

If $k \in \omega(n^k)$ for all $k \geq 1$, then

$$\hat{C}^*(S) - O(1) \leq \hat{R}_1^{\epsilon}(S) \leq \hat{C}^*(S) + o(\hat{C}^*(S)).$$

Proof: The lower bound comes from Lemma 13. For the upper bound, Corollary 21 gives

$$\begin{aligned} \hat{R}_1^{\epsilon}(S) &\leq \log \widehat{a_{\mathcal{Y}}} + \log\log|S| + c(\epsilon) \\ &\leq \log \widehat{a_{\mathcal{Y}}} + \log\log(|X| \cdot \widehat{a_{\mathcal{Y}}}) + O(1), \end{aligned}$$

and if $\widehat{a_{\mathcal{Y}}} \in \Theta(n^k)$ it follows that

$$\begin{aligned} \hat{R}_1^{\epsilon}(S) &\leq (k+1)\log n + o(\log n) \\ &\leq \frac{k+1}{k} \cdot \log n^k + o(\log n). \end{aligned}$$

The case $k \in \omega(n^k)$ for all $k \geq 1$ is proved similarly. ■

Example 31: A large database D on a PC has to be synchronized with an updated version D' of the database stored on a PDA. The problem can be viewed as a *set reconciliation* problem [25], and for this example we assume that D and D' are sets of integers in $[1, 10^6]$, that D' has to be conveyed to the PC, and that D and D' differ for at most 1000 entries, i.e., $|D \setminus D'| + |D' \setminus D| \leq 1000$. Translated in the interactive communication framework, D and D' can be expressed as binary strings x and y such that $x_i = 1$ ($y_i = 1$) if and only if $i \in D$ ($i \in D'$). The correlation between D and D' gives an upper bound on the Hamming distance between x and y, thus the support set is symmetric and $\widehat{a_{\mathcal{X}}} = \widehat{a_{\mathcal{Y}}} = \sum_{i=0}^{1000} \binom{1000000}{i}$.

From Result 6, we know there exists a 3-round deterministic algorithm requiring 11452 bits of communication. Nonzero-error algorithms cannot do much better: from Lemma 13, even with a public random string, at least 11401 bits have to be exchanged in order for $P_{\mathcal{Y}}$ to learn x with error at most ϵ, for $0 < \epsilon < \frac{1}{2}$. From Lemma 17, there exists a one-way public coin randomized protocol transmitting 11452 bits with a 2^{-50} probability of error, and a one-way private coin randomized protocol transmitting 11576 bits with a 2^{-50} probability of error from Corollary 21.

The results just stated assume that $P_{\mathcal{X}}$ and $P_{\mathcal{Y}}$ have unbounded memory and computing power. For several classes of symmetric support sets, a lot research has been done to design algorithms with reasonable tradeoffs between the number of rounds, communication complexity, computational complexity and storage space. Of course, the requirements vary widely depending on the application and the size of the maximum ambiguity of the players.

VIII. Open Problems

In this paper, we have studied worst-case nonzero-error interactive communication. We have shown that if the players are allowed to use public coins, to answer correctly on a fraction of the inputs or to solve a large number of instances simultaneously, then interaction is not necessary and in some cases the communication can be significantly reduced. Although we have proved that

one round of communication is almost optimal, we have not been able to completely characterize the private coin randomized model, and in this section we discuss two open problems.

A. One-Way Private Coin Randomized Complexity

Is one round of communication optimal for private coin randomized complexity? If not, is there a k such that an optimal k-round protocol always exists, i.e., a k such that

$$\hat{R}^{\epsilon}_{k}(S) \leq \hat{R}^{\epsilon}_{\infty}(S) + o(\hat{R}^{\epsilon}_{\infty}(S))?$$

The lower bounds for private coin randomized complexity presented earlier do not restrict the number of rounds between $P_{\mathcal{X}}$ and $P_{\mathcal{Y}}$. Moreover, it is not clear if existing techniques for proving lower bounds for other models of computation like communication complexity of non-boolean functions can be used to derive stronger lower bounds in the interactive communication setting. As for upper bounds, although it is not obvious how randomized protocols can use interaction to reduce the the number of bits that need to be transmitted, our best upper bound, Theorem 19, is not always tight.

B. Deterministic Model Versus Private Coin Randomized Model

The private coin randomized model allows almost optimal one-way protocols, but it does not seem that it can be more efficient than determinism when interaction is allowed. In fact, we conjecture that the worst-case deterministic and worst-case private coin randomized models are equivalent, i.e.,

$$\hat{R}^{\epsilon}_{\infty}(S) \sim \hat{C}_{\infty}(S).$$

In Section III, we have shown that the models are equivalent when $P_{\mathcal{Y}}$'s maximum ambiguity is constant. This includes all the maximally unbalanced pairs like the league problem, i.e, pairs with $\widehat{a_{\mathcal{Y}}} \in O(1)$ and $\widehat{a_{\mathcal{X}}} \in \Omega(|X|)$. In Section VII, it has been proved that the models are equivalent for balanced and symmetric support sets. Thus, in order to solve the conjecture, one needs to study problems for which the private coin randomized complexity is not tightly bounded by Lemmas 9 and 13, i.e., problems that are neither balanced nor too unbalanced.

ACKNOWLEDGMENT

The first author would like to thank Hervé Caussinus for suggesting the problem, and Alon Orlitsky and Krishnamurthy Viswanathan for their comments regarding Lemma 14.

REFERENCES

[1] A. Orlitsky, "Worst-case interactive communication I: Two messages are almost optimal," *IEEE Trans. Inform. Theory*, vol. 36, no. 5, pp. 1111–1126, 1990.

[2] A. C.-C. Yao, "Some complexity questions related to distributed computing," in *Proceedings of the 11th ACM Symposium on Theory of Computing (STOC)*, 1979, pp. 209–213.

[3] E. Kushilevitz and N. Nisan, *Communication Complexity*. Cambridge University Press, 1997.

[4] A. Orlitsky, "Interactive communication of balanced distributions and of correlated files," *SIAM Journal on Discrete Mathematics*, vol. 6, no. 4, pp. 548–564, 1993.

[5] D. Starobinski, A. Trachtenberg, and S. Agarwal, "Efficient PDA synchronization," *IEEE Trans. Mobile Comput.*, vol. 2, no. 1, pp. 40–51, 2003.

[6] R. Durbin, S. Eddy, A. Krogh, and G. Mitchison, *Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids*. Cambridge University Press, 1998.

[7] G. Cormode, M. Paterson, S. Sahinalp, and U. Vishkin, "Communication complexity of document exchange," in *Proceedings of the eleventh ACM-SIAM Symposium on Discrete Algorithms (SODA)*, 2000, pp. 197–206.

[8] G. Brassard and L. Salvail, "Secret-key reconciliation by public discussion," in *Advances in Cryptology: Proceedings of Eurocrypt '93*, ser. Lectures Notes in Computer Science, vol. 765. Springer-Verlag, 1994, pp. 410–423.

[9] A. C.-C. Yao, "Lower bounds by probabilistic arguments," in *Proceedings of the 24th IEEE Symposium on Foundations of Computer Science (FOCS)*, 1983, pp. 420–428.

[10] P. Duris, Z. Galil, and G. Schnitger, "Lower bounds on communication complexity," in *Proceedings of the 16th annual ACM Symposium on Theory of Computing (STOC)*, 1984, pp. 81–91.

[11] A. Orlitsky, "Worst-case interactive communication II: Two messages are not optimal," *IEEE Trans. Inform. Theory*, vol. 37, no. 4, pp. 995–1005, 1991.

[12] Z. Zhang and X.-G. Xia, "Three messages are not optimal in worst case interactive communication," *IEEE Trans. Inform. Theory*, vol. 40, no. 1, pp. 3–10, 1994.

[13] R. Ahlswede, N. Cai, and Z. Zhang, "On interactive communication," *IEEE Trans. Inform. Theory*, vol. 43, no. 1, pp. 22–37, 1997.

[14] M. Naor, A. Orlitsky, and P. Shor, "Three results on interactive communication," *IEEE Trans. Inform. Theory*, vol. 39, no. 5, pp. 1608–1615, 1993.

[15] M. Karchmer, R. Raz, and A. Wigderson, "Super-logarithmic depth lower bounds via the direct sum in communication complexity," *Computational Complexity*, vol. 5, no. 3/4, pp. 191–204, 1995.

[16] C. H. Papadimitriou, *Computational Complexity*. Addison-Wesley, 1994.

[17] N. Alon and A. Orlitsky, "Repeated communication and Ramsey graphs," *IEEE Trans. Inform. Theory*, vol. 41, no. 5, pp. 1276–1289, 1995.

[18] T. Feder, E. Kushilevitz, M. Naor, and N. Nisan, "Amortized communication complexity," *SIAM Journal on Computing*, vol. 24, no. 4, pp. 736–750, 1995.

[19] A. Orlitsky, "Average-case interactive communication," *IEEE Trans. Inform. Theory*, vol. 38, pp. 1534–1547, 1992.

[20] N. Alon and A. Orlitsky, "Source coding and graph entropies," *IEEE Trans. Inform. Theory*, vol. 42, no. 5, pp. 1329–1339, 1996.

[21] G. H. Hardy and E. M. Wright, *An introduction to the theory of numbers*. Oxford University Press, 1979.

[22] I. Newman, “Private vs. common random bits in communication complexity,” *Information Processing Letters*, pp. 67–71, 1991.

[23] A. C.-C. Yao, “Probabilistic computations: toward a unified measure of complexity,” in *Proceedings of the 18th IEEE Symposium on Foundations of Computer Science (FOCS)*, 1977, pp. 222–227.

[24] M. Willem, *Minimax Theorems*, ser. Progress in Nonlinear Differential Equations and Their Applications. Birkhäuser, 1996.

[25] Y. Minsky, A. Trachtenberg, and R. Zippel, “Set reconciliation with nearly optimal communication complexity,” *IEEE Trans. Inform. Theory*, vol. 49, no. 9, pp. 2213–2218, 2003.