New strong direct product results in communication complexity

Rahul Jain
Centre for Quantum Technologies and Department of Computer Science
National University of Singapore.
rahul@comp.nus.edu.sg

March 10, 2011

Abstract

We show two new direct product results in two different models of communication complexity. Our first result is in the model of one-way public-coin model. Let $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ be a relation and $\varepsilon > 0$ be a constant. Let $\mathsf{R}_{\varepsilon}^{1,\mathsf{pub}}(f)$ represent the communication complexity of f, with worst case error ε in this model. We show that if for computing f^k (k independent copies of f) in this model, $o(k \cdot \mathsf{R}_{1/3}^{1,\mathsf{pub}}(f))$ communication is provided, then the success is exponentially small in k. To our knowledge this is the first time a strong direct product result holding for all relations has been shown in any model of communication complexity. We show a new tight characterization of communication complexity in this model which strengthens on the tight characterization shown in J., Klauck, Nayak [JKN08]. We use this new characterization to show our direct product result and this characterization may also be of independent interest.

Our second direct product result is in the model of two-way public-coin communication complexity. We show a direct product result for all relations in this model in terms of a new complexity measure that we define. Our new measure is a generalization to non-product distributions, of the two-way product subdistribution bound of J., Klauck and Nayak [JKN08]. Our direct product result therefore generalizes to non-product distributions, their direct product result in terms of the two-way product subdistribution bound. As an application of our new direct product result, we reproduce (via completely different arguments) strong direct product result for the set-disjointness problem which was previously shown by Klauck [Kla10]. We show this by showing that our new complexity measure gives tight lower bound of $\Omega(n)$ for the set-disjointness problem on n-bit inputs (this strengthens on the linear lower bound on the rectangle/corruption bound for set-disjointness shown by Razborov [Raz92]). In addition we show that many previously known direct product results in this model are uniformly implied and often strengthened by our result.

Keywords: Communication complexity, strong direct product, information theory

1 Introduction

Can computing k simultaneous instances of a problem be done more efficiently than just computing them in parallel? This question has been very well studied in many models of computation, first for being a natural, fundamental and interesting question in itself and second for the many implications it has for some other important questions. One way to pose this question for bounded error computation models is as follows. Let the resource required for solving a single instance with constant success be c; then if $o(kc)$ resource is provided for solving k instances together, is the overall success exponentially small in k? This is referred to as the direct product question. We consider this question in two models of communication complexity: the one-way public-coin model and the two-way public-coin model.

The one-way public-coin model

Let $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ be a relation and $\varepsilon > 0$. Let Alice with input $x \in \mathcal{X}$ and Bob with input $y \in \mathcal{Y}$ wish to compute a $z \in \mathcal{Z}$ such that $(x, y, z) \in f$. In this model Alice sends a single message to Bob who outputs z and Alice and Bob use pubic coins. Let $\mathsf{R}^{1,\mathsf{pub}}_{\varepsilon}(f)$ denote the communication of the best protocol $\mathcal{P}$ which achieves this with error at most ε (over the public-coins) for any input (x, y). We answer the direct product question for all relations f in this model in the following manner.

Theorem 1.1 *Let $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ be a relation, $\varepsilon > 0$ be a constant and k be a natural number. Then,*

$$\mathsf{R}^{1,\mathsf{pub}}_{1-2^{-\Omega(\varepsilon^3 k)}}(f^k) = \Omega\left(k \cdot \left(\varepsilon^2 \cdot \mathsf{R}^{1,\mathsf{pub}}_{\varepsilon}(f) - O(1)\right)\right) \quad .$$

To our knowledge this is the first time a direct product statement has been made for all relations in any model of communication complexity[1]. This question can be considered open for more than thirty years since 1979 when the model of communication complexity was first defined by Yao [Yao79]. We present below some previous results which are now implied and strengthened by our result above.

1. J., Klauck, Nayak [JKN08] introduced the so called one-way subdistribution bound and showed a direct product result in terms of the one-way subdistribution bound under product distributions. The one-way subdistribution bound forms a lower bound on $\mathsf{R}^{1,\mathsf{pub}}_{\varepsilon}(f)$ and hence our result implies their's.
2. Gavinsky [Gav08] proves direct product for one-way distributional communication complexity of a certain class of relational problems under the uniform distribution. Gavinsky used his result to show communication v/s entanglement trade-off for communication protocols. Since our result holds for all relations, it implies (in fact with stronger parameters) the result of Gavinsky.
3. De Wolf [dW05] proves a strong direct product theorem for the one-way public-coin randomized communication complexity of the Index function. Ben-Aroya, Regev, and de Wolf [BARdW08] derive a similar direct product theorem for the one-way quantum communication complexity of Index. Since Index captures the notion of VC-dimension, similar results follow for the one-way distributional (classical and quantum) communication complexity of any Boolean function under the worst case product distribution. These results for classical communication complexity are implied by our result above.

[1] Nsote that in case $\mathsf{R}^{1,\mathsf{pub}}_{1/3}(f) \geq 1$ is a constant, then the corresponding direct product statement $\mathsf{R}^{1,\mathsf{pub}}_{1-2^{-\Omega(k)}}(f^k) = \Omega(k)$ is easily argued. Also a direct product result for the one-way private-coin communication complexity follows immediately from our result and the well known relationship $\mathsf{R}^{1,\mathsf{pub}}_{\varepsilon}(f) \leq \mathsf{R}^{1,\mathsf{pvt}}_{\varepsilon}(f) \leq \mathsf{R}^{1,\mathsf{pub}}_{\varepsilon}(f) + O(\log|\mathcal{X}| + \log|\mathcal{Y}|)$, which holds for all relations $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ and constant $\varepsilon > 0$ [New91].

4. J., Radhakrishnan, and Sen [JRS05] show optimal direct sum for all relations in the one-way public coin communication complexity. A direct sum is a weaker question to direct product question and asks the following. Let the resource required for solving a single instance with constant success be c; then if $o(kc)$ resource is provided for solving k instances together, is the overall success is at most a constant? (In direct product the overall success is required to be exponentially small in k). [JRS05] also show similar optimal direct sum result for one-way entanglement assisted quantum communication complexity of all relations; however quantum communication complexity is beyond the scope of this work.

Our techniques

We follow a natural argument for showing the direct product result. Let us say there are totally k coordinates (instances) and we condition on success on $l = d \cdot k$ ($d < 1$ is a small constant) coordinates. If the overall success in these l coordinates is already as small as we want then we are done and stop. Otherwise we try to exhibit another coordinate j outside of these l coordinates such that the success in the j-th coordinate, even conditioned on the success in the l coordinates, is bounded away from 1. This way the overall success keeps going down and becomes exponentially small eventually.

We do this argument in the distributional setting where one is concerned with average error over the inputs coming from a specified distribution rather than the worst case error over all inputs. The distributional setting can then be related to the worst case setting by the well known Yao's principal. Let μ be a hard distribution on $\mathcal{X} \times \mathcal{Y}$, possibly non-product across $\mathcal{X}$ and $\mathcal{Y}$. Let us consider the inputs for f^k drawn from the distribution μ^k (k independent copies of μ). Now consider a one-way protocol for f^k with communication $o(kc)$ and condition on a typical message string m from Alice. Conditioned on this message m and also on success in l coordinates, we analyze how the distribution of Alice and Bob's inputs on a typical coordinate j (outside the l coordinates) looks like. We argue that this distribution is still hard enough, that is Bob will make constant error on this coordinate whichever way he tries to give an answer. We are able to identify some key properties in such a distribution, concerning its relationship to μ, and argue that any distribution with these properties must be a hard distribution, given that μ is a hard distribution.

We do this last argument by showing a new tight characterization of one-way public-coin communication complexity for all relations. We introduce a new measure of complexity which we call the robust conditional relative min-entropy bound (rcment). We show that this bound is equivalent, up to constants, to $\mathsf{R}^{1,\mathsf{pub}}_{\varepsilon}(f)$ (for any constant $\varepsilon > 0$). This bound forms lower bound on the one-way subdistribution bound of J., Klauck, Nayak [JKN08] where they also show that their bound is equivalent, up to constants, to $\mathsf{R}^{1,\mathsf{pub}}_{\varepsilon}(f)$.

One key difficulty that is faced in the argument outlined above is that while Bob is making a decision on the jth coordinate, he can use his inputs in other coordinates while making this decision. Since μ could be a non-product distribution, Bob's inputs in other coordinates potentially provide him information about Alice's inputs, in addition to the information obtained from the message from Alice. This difficulty is overcome by splitting the distribution μ into a convex combination of several product distributions. The particular way in which we split distributions leads us to consider the conditional distributions (conditioned on Bob's inputs) in the definition of rcment. This idea of splitting a non-product distribution into convex combination of product distributions has been used in several previous works to handle non-product distributions in different settings [Raz92, Raz98, BJKS02, Hol07, BBR10].

The two-way public-coin model

In this model Alice on input x and Bob on input y exchange messages using pubic coins and at the end agree on a common output z. Let $\mathsf{R}^{2,\mathsf{pub}}_{\varepsilon}(f)$ denote the communication of the best protocol

$\mathcal{P}$ which achieves this with error at most ε (over the public-coins) for any input (x, y). We show a direct product result in terms of a new complexity measure that we introduce: the ε-error two-way conditional relative entropy bound of f with respect to distribution μ, denoted $\mathsf{crent}^{2,\mu}_{\varepsilon}(f)$. The measure $\mathsf{crent}^{2,\mu}_{\varepsilon}(f)$ forms a lower bound (upto constants) on $\mathsf{R}^{2,\mathsf{pub}}_{\varepsilon}(f)$. Although this result is not an optimal direct product result that one may desire, we show how many previously known direct product results in the two-way model follow as a consequence of our result.

1. Recently Klauck [Kla10] showed a direct product result for the set disjointness problem. In the set disjointness problem, Alice with input $x \in \{0,1\}^n$ and Bob with input $y \in \{0,1\}^n$ are supposed to determine if x and y intersect when viewed as characteristic vectors of subsets of $[n]$. This is arguably one of the most well studied problems in communication complexity. We show (in section 5) that our new complexity measure crent gives tight lower bound for the set-disjointness problem. This combined with the direct product in terms of crent, implies strong direct product result for the set disjointness problem for its two-way public-coin communication complexity. We point here that the arguments used in [Kla10] are arguably specifically geared to handle the set disjointness problem. In contrast our result is much more general as we further argue below.
2. When μ is a product distribution, $\mathsf{crent}^{2,\mu}_{\varepsilon}(f)$ forms an upper bound (upto constants) on the two-way subdistribution bound of J., Klauck, Nayak [JKN08]. Hence our direct product result generalizes to non-product distributions, the direct product result of [JKN08] for their two-way product subdistribution bound and in particular implies their result. It was pointed in [JKN08] that their result provides a unified view of several recent works on the topic, simultaneously generalizing and strengthening them. These works include the strong direct product property for the rectangle/corruption bound for Boolean functions due to Beame et al. [BPSW07].
3. Shaltiel [Sha03] gave strong direct product theorem for the discrepancy bound for communication complexity under the uniform distribution. The discrepancy bound under product distributions (in particular under the uniform distribution) is upper bounded by the rectangle bound which in turn is upper bounded (upto constants) by the crent. Therefore our result implies and strengthens on Shaltiel's result and in particular implies strong direct product for the Inner Product function ($\mathsf{IP}_n(x, y) = \sum_i x_i \cdot y_i \mod 2$), since for this function the discrepancy bound under the uniform distribution is $\Omega(n)$.

Our techniques to show the direct product in the two-way model are quite similar to the techniques in the one-way model. However in the two-way model we do not present an upper bound on the public-coin communication complexity in terms of the new measure crent.

Other related work in communication complexity

Parnafes, Raz, and Wigderson [PRW97] prove a direct product result when a different algorithm works for each of the different instances and each algorithm is only provided communication at most the communication complexity of single instance (with constant error). In their result the bound on the success probability is shown to behave like $2^{k/c}$ for the communication complexity c of the problem at hand. Lee, Shraibman and Spalek [LSS08] have shown strong direct product for the discrepancy bound under arbitrary distributions and Viola and Wigderson [VW08] have extended it to the multiparty case. Recently Sherstov [She11] showed strong direct product for the generalized-discrepancy bound. For deterministic protocols it is known that k times the square root of the deterministic communication complexity of a function f is needed to compute k instances of f (see, e.g., [[KN97], Exercise 4.11, page 46]). It is also straightforward to show that the deterministic one-way communication complexity of every function f has the direct sum property. In a sequence of results [JRS03, HJMR09, BBR10, BR10] the direct sum property for all relations in the two-way public-coin model has been *almost* shown, however the optimal

direct sum result holding for all relations is still open. Direct sum result for all relations in the public-coin SMP model has been shown in [JRS05] both for classical and quantum protocols. In the SMP (simultaneous message passing) model, Alice with input x and Bob with input y, each send a message to a Referee who outputs z. Direct sum results for all relations in the private-coin SMP model has been shown in [JK09] both for classical and quantum protocols. In a weak direct product theorem, one shows that the success probability of solving k instances of a problem with the resources needed to solve one instance (with constant success) goes down exponentially with k. Klauck [Kla04] shows such a result for the rectangle/corruption bound for all functions and all distributions in the two-way model and [JKN08] extend this result for all relations and for all distributions in the same model.

Indeed above is only an incomplete list of the many interesting results concerning direct product and related questions in communication complexity.

Organization

In section 2 we provide some information theory and communication complexity preliminaries that we need. We refer the reader to the texts [CT91, KN97] for good introductions to these topics respectively. In section 3 we introduce our new bound for one-way communication, show that it tightly characterizes one-way public-coin communication complexity and show our direct product result in the one-way model. In section 4 we introduce our new bound for two-way communication and show a direct product result in its terms. In section 5 we present the strong direct product for set disjointness as an application of our two-way direct product result.

2 Preliminaries

Information theory

Let $\mathcal{X}, \mathcal{Y}$ be sets and k be a natural number. Let $\mathcal{X}^k$ represent $\mathcal{X} \times \cdots \times \mathcal{X}$, k times. Let μ be a distribution over $\mathcal{X}$ which we denote by $\mu \in \mathcal{X}$. We let $\mu(x)$ represent the probability of x under μ. The entropy of μ is defined as $S(\mu) = -\sum_{x \in \mathcal{X}} \mu(x) \log \mu(x)$. Let X be a random variable distributed according to μ which we denote by $X \sim \mu$. We use the same symbol to represent a random variable and its distribution whenever it is clear from the context. For distributions $\mu, \mu_1 \in \mathcal{X}$, $\mu \otimes \mu_1$ represents the product distribution $(\mu \otimes \mu_1)(x) = \mu(x) \cdot \mu_1(x)$ and μ^k represents $\mu \otimes \cdots \otimes \mu$, k times. The ℓ_1 distance between distributions μ, μ_1 is defined as $||\mu - \mu_1||_1 = \frac{1}{2} \sum_{x \in \mathcal{X}} |\mu(x) - \mu_1(x)|$. The relative entropy between $\lambda \in \mathcal{X}$ and μ is defined as $S(\lambda||\mu) = \sum_{x \in \mathcal{X}} \lambda(x) \log \frac{\lambda(x)}{\mu(x)}$. The relative min-entropy between λ and μ is defined as $S_\infty(\lambda||\mu) = \max_{x \in \mathcal{X}} \log \frac{\lambda(x)}{\mu(x)}$. It is easily seen that $S(\lambda||\mu) \leq S_\infty(\lambda||\mu)$. Let $\lambda, \mu \in \mathcal{X} \times \mathcal{Y}$. We use $\mu(x|y)$ to represent $\mu(x, y)/\mu(y)$. Let $XY \sim \mu$, here we assume that $X \in \mathcal{X}$ and $Y \in \mathcal{Y}$. We use μ_x and Y_x to represent $Y|\ X = x$. The conditional entropy of Y given X, is defined as $S(Y|X) = \mathbb{E}_{x \leftarrow X} S(Y_x)$. We use the following properties of relative entropy.

Fact 2.1 *1. Relative entropy is jointly convex in its arguments, that is for distributions $\lambda_1, \lambda_2, \mu_1, \mu_2$,*

$$S(p\lambda_1 + (1-p)\lambda_2 \ ||\ p\mu_1 + (1-p)\mu_2) \leq p \cdot S(\lambda_1||\mu_1) + (1-p) \cdot S(\lambda_2||\mu_2) \ .$$

2. Let $XY, X^1Y^1 \in \mathcal{X} \times \mathcal{Y}$. Relative entropy satisfies the following chain rule,

$$S(XY||X^1Y^1) = S(X||X^1) + \mathbb{E}_{x \leftarrow X} S(Y_x||Y^1_x) \ .$$

This in-particular implies, using joint convexity of relative entropy,

$$S(XY||X^1 \otimes Y^1) = S(X||X^1) + \mathbb{E}_{x \leftarrow X} S(Y_x||Y^1) \geq S(X||X^1) + S(Y||Y^1) \quad .$$

3. *For distributions* λ, μ *:* $||\lambda - \mu||_1 \leq \sqrt{S(\lambda||\mu)}$ *and* $S(\lambda||\mu) \geq 0$.
4. *Substate theorem [JRS02]: Let* λ, μ *be distributions. For every* $\delta > 0$*, there exists a distribution* λ_δ *such that* $S_\infty(\lambda_\delta||\mu) \leq O(\frac{1}{\delta}(S(\lambda||\mu) + 1))$ *and* $||\lambda_\delta - \lambda||_1 \leq \delta$.

Let X, Y, Z be random variables. The mutual information between X and Y is defined as

$$I(X:Y) = S(X) + S(Y) - S(XY) = \mathbb{E}_{x \leftarrow X} S(Y_x||Y) = \mathbb{E}_{y \leftarrow Y} S(X_y||X).$$

The conditional mutual information is defined as $I(X:Y|\ Z) = \mathbb{E}_{z \leftarrow Z} I(X:Y|\ Z=z)$. Random variables XYZ form a Markov chain $Z \leftrightarrow X \leftrightarrow Y$ iff $I(Y:Z|\ X=x) = 0$ for each x in the support of X. The following fact is often used in our proofs.

Fact 2.2 *Let* $\lambda, \mu \in \mathcal{X}$ *be distributions. Then* $\sum_{x \in \mathcal{X}: \lambda(x) < \mu(x)} \lambda(x) \log \frac{\lambda(x)}{\mu(x)} \geq -1$.

One-way communication complexity

Let $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ be a relation. We only consider complete relations that is for each $(x, y) \in \mathcal{X} \times \mathcal{Y}$, there exists at least one $z \in \mathcal{Z}$ such that $(x, y, z) \in f$. In the one-way model of communication there is a single message, from Alice with input $x \in \mathcal{X}$ to Bob with input $y \in \mathcal{Y}$, at the end of which Bob is supposed to determine an answer z such that $(x, y, z) \in f$. Let $\varepsilon > 0$ and let $\mu \in \mathcal{X} \times \mathcal{Y}$ be a distribution. We let $\mathsf{D}^{1,\mu}_{\varepsilon}(f)$ represent the distributional one-way communication complexity of f under μ with expected error ϵ, i.e., the communication of the best deterministic one-way protocol for f, with distributional error (average error over the inputs) at most ε under μ. Let $\mathsf{R}^{1,\mathsf{pub}}_{\epsilon}(f)$ represent the one-way public-coin communication complexity of f with worst case error ε, i.e., the communication of the best one-way public-coin protocol for f with error for each input (x, y) being at most ε. The following is a consequence of the min-max theorem in game theory [KN97, Theorem 3.20, page 36].

Lemma 2.3 (Yao principle) $\mathsf{R}^{1,\mathsf{pub}}_{\epsilon}(f) = \max_\mu \mathsf{D}^{1,\mu}_{\epsilon}(f)$.

The following result follows quite directly from the arguments in Braverman and Rao [BR10]. We provide its proof in Appendix A for completeness.

Lemma 2.4 (Braverman and Rao [BR10]) *Let* $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ *be a relation and* $\varepsilon > 0, \delta \geq 0$*. Let* $XY \sim \mu$ *be inputs to a one-way private-coin communication protocol* $\mathcal{P}$ *with distributional error at most* ε*. Let* M *represent the message of* $\mathcal{P}$*. Let* θ *be the distribution of* XYM *and let*

$$\Pr_{(x,y,i) \leftarrow \theta} \left[\log \frac{\theta(i|x)}{\theta(i|y)} > c \right] \leq \delta. \tag{2.1}$$

There exists a deterministic one-way protocol $\mathcal{P}_1$ *for* f *with inputs distributed according to* μ*, such that the communication of* $\mathcal{P}_1$ *is* $c + O(\log(1/\delta))$*, and distributional error of* $\mathcal{P}_1$ *is at most* $\varepsilon + 2\delta$.

Two-way communication complexity

Let $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ be a relation. As mentioned before we only consider complete relations. In the two-way model of communication, Alice with input $x \in \mathcal{X}$ and Bob with input $y \in \mathcal{Y}$, communicate at the end of which they are supposed to determine a common answer z (as a function of the message transcript) such that $(x, y, z) \in f$. Let $\varepsilon > 0$ and let $\mu \in \mathcal{X} \times \mathcal{Y}$ be a distribution. We let $\mathsf{D}^{2,\mu}_{\varepsilon}(f)$ represent the two-way distributional communication complexity of f under μ with expected error ϵ, i.e., the communication of the best deterministic two-way protocol for f, with distributional error (average error over the inputs) at most ε under μ. Let $\mathsf{R}^{2,\mathsf{pub}}_{\epsilon}(f)$ represent the two-way public-coin communication complexity of f with worst case error ε, i.e., the communication of the best two-way public-coin protocol for f with error for each input (x, y) being at most ε. The following is a consequence of the min-max theorem in game theory [KN97, Theorem 3.20, page 36].

Lemma 2.5 (Yao principle) $\mathsf{R}^{2,\mathsf{pub}}_{\epsilon}(f) = \max_{\mu} \mathsf{D}^{2,\mu}_{\epsilon}(f)$.

3 One-way communication

Definitions

We make here the necessary definitions for this section. Let $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ be a relation, $\mu, \lambda \in \mathcal{X} \times \mathcal{Y}$ be distributions and $\varepsilon, \delta > 0$.

Definition 3.1 (One-way distributions) *Distribution λ is called one-way for distribution μ if for all (x, y) in the support of λ we have $\mu(y|x) = \lambda(y|x)$.*

Note that in a one-way protocol if the inputs are drawn from μ and we condition on a message transcript from Alice, then the resulting distribution would be one-way for μ.

Definition 3.2 (Error of a distribution) *Error of distribution μ with respect to f, denoted $\mathsf{err}_f(\mu)$, is defined as*

$$\mathsf{err}_f(\mu) \stackrel{\text{def}}{=} \min \left\{ \Pr_{(x,y)\leftarrow\mu} [(x, y, g(y)) \notin f] \mid g : \mathcal{Y} \to \mathcal{Z} \right\} \ .$$

Let μ be the distribution of inputs for Alice and Bob. Let Bob make an output depending on his input, without any communication from Alice. Then $\mathsf{err}_f(\mu)$ represents the least error that Bob must make.

Definition 3.3 (Robust conditional relative min-entropy) *The δ-robust conditional relative min-entropy of λ with respect to μ, denoted $\mathsf{rcment}^{\mu}_{\delta}(\lambda)$, is defined to be the minimum number c such that*

$$\Pr_{(x,y)\leftarrow\lambda} \left[\log \frac{\lambda(x|y)}{\mu(x|y)} > c \right] \leq \delta.$$

Definition 3.4 (Robust conditional relative min-entropy bound) *The ε-error δ-robust conditional relative min-entropy bound of f with respect to distribution μ, denoted $\mathsf{rcment}^{\mu}_{\varepsilon,\delta}(f)$, is defined as*

$$\mathsf{rcment}^{\mu}_{\varepsilon,\delta}(f) \stackrel{\text{def}}{=} \min \{ \mathsf{rcment}^{\mu}_{\delta}(\lambda) |\ \lambda \textit{ is one-way for } \mu \textit{ and } \mathsf{err}_f(\lambda) \leq \varepsilon \} \ .$$

The ε-error δ-robust conditional relative min-entropy bound of f, denoted $\mathsf{rcment}_{\epsilon,\delta}(f)$, is defined as

$$\mathsf{rcment}_{\varepsilon,\delta}(f) \stackrel{\text{def}}{=} \max \left\{ \mathsf{rcment}^{\mu}_{\varepsilon,\delta}(f) |\ \mu \textit{ is a distribution over } \mathcal{X} \times \mathcal{Y} \right\} \ .$$

We often use the above definition in the following way. Let λ be a distribution which is one-way for μ and with $\mathsf{rcment}^{\mu}_{\delta}(\lambda) < \mathsf{rcment}^{\mu}_{\varepsilon,\delta}(f)$. Then $\mathsf{err}_f(\lambda) > \varepsilon$.

The following bound was defined in [JKN08] where it was referred to as the one-way sub-distribution bound. We call it differently here for consistency of nomenclature with the other bound.

Definition 3.5 (Relative min-entropy bound) *The ε-error relative min-entropy bound of f with respect to distribution μ, denoted $\mathsf{ment}^{\mu}_{\varepsilon}(f)$, is defined as*

$$\mathsf{ment}^{\mu}_{\varepsilon}(f) \stackrel{\text{def}}{=} \min \{ S_{\infty}(\lambda||\mu) |\ \lambda \textit{ is one-way for } \mu \textit{ and } \mathsf{err}_f(\lambda) \leq \varepsilon \} \ .$$

The ε-error relative min-entropy bound of f, denoted $\mathsf{ment}(f)$, is defined as

$$\mathsf{ment}_{\varepsilon}(f) \stackrel{\text{def}}{=} \max \{ \mathsf{ment}^{\mu}_{\varepsilon}(f) |\ \mu \textit{ is a distribution over } \mathcal{X} \times \mathcal{Y} \} \ .$$

Lemma 3.1 *For every λ, μ and $\delta > 0$ we have* $\mathsf{rcment}^{\mu}_{\delta}(\lambda) \leq (S_{\infty}(\lambda||\mu)+1)/\delta$, *hence* $\mathsf{rcment}^{\mu}_{\varepsilon,\delta}(f) = O(\mathsf{ment}^{\mu}_{\varepsilon}(f))$ *and* $\mathsf{rcment}_{\varepsilon,\delta}(f) = O(\mathsf{ment}_{\varepsilon}(f))$.

Proof: Let $XY \sim \lambda$ and $X'Y' \sim \mu$. Then using Part 2. of Fact 2.1 we have,

$$S_{\infty}(\lambda||\mu) \geq S(\lambda||\mu) \geq \mathbb{E}_{y \leftarrow Y} S(X_y||X'_y) = \mathbb{E}_{(x,y)\leftarrow\lambda} \log \frac{\lambda(x|y)}{\mu(x|y)}.$$

Therefore using Fact 2.2 and Markov's inequality we get $\Pr_{(x,y)\leftarrow\lambda}[\log \frac{\lambda(x|y)}{\mu(x|y)} > (S_{\infty}(\lambda||\mu) + 1)/\delta] < \delta$. Hence $\mathsf{rcment}^{\mu}_{\delta}(\lambda) \leq (S_{\infty}(\lambda||\mu) + 1)/\delta$. The other relationships follow from definitions. ■

New characterization

In this section we show a new characterization of one-way public-coin communication complexity in terms of rcment. The following lemma which lower bounds distributional communication complexity using ment appears in [JKN08].

Lemma 3.2 *Let $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ be a relation and $\mu \in \mathcal{X} \times \mathcal{Y}$ be a distribution and $\varepsilon, k > 0$. Then,*

$$\mathsf{D}^{1,\mu}_{\epsilon(1-2^{-k})}(f) \quad \geq \quad \mathsf{ment}^{\mu}_{\varepsilon}(f) - k.$$

We show the following key lemma which upper bounds distributional communication complexity using rcment.

Lemma 3.3 *Let $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ be a relation and $\mu \in \mathcal{X} \times \mathcal{Y}$ be a distribution and $\varepsilon, \delta > 0$. Then,*

$$\mathsf{D}^{1,\mu}_{\varepsilon+4\delta}(f) \leq \mathsf{rcment}_{\varepsilon,\delta}(f) + O(\log \frac{1}{\delta}) \ .$$

Proof: We start with the following key claim where we produce a desired split of μ into several distributions which are one-way for μ. This will enable us to obtain a one-way protocol with small communication as we show later.

Claim 3.4 *There exists a natural number k and a Markov chain $M \leftrightarrow X \leftrightarrow Y$, where $M \in [k]$ and $XY \sim \mu$, such that*

1. *for each $i \in [k]$: $\mathsf{err}_f(P_i) \leq \varepsilon$, where $P_i = (XY|\ M = i)$ and*
2. $\Pr_{(x,y,i)\leftarrow\theta}\left[\log \frac{\theta(i|x)}{\theta(i|y)} > \mathsf{rcment}_{\varepsilon,\delta}(f) + \log \frac{1}{\delta}\right] \leq 2\delta$, *where θ is the distribution of XYM.*

Proof: Let $c = \mathsf{rcment}_{\varepsilon,\delta}(f)$. Let us perform a procedure as follows. Start with $i = 1$.

1. Let us say we have collected distributions $P_1, \ldots, P_{i-1}$, each one-way for μ, and positive numbers $p_1, \ldots, p_{i-1}$ such that $\mu \geq \sum_{j=1}^{i-1} p_j P_j$. If $\mu = \sum_{j=1}^{i-1} p_j P_j$ then set $k = i - 1$ and stop.
2. Otherwise let us express $\mu = \sum_{j=1}^{i-1} p_j P_j + q_i Q_i$, where Q_i is a distribution, one-way for μ. Since $\mathsf{rcment}^{Q_i}_{\varepsilon,\delta}(f) \leq c$, we know that there is a distribution R, one-way for Q_i (hence also one-way for μ), such that $\mathsf{rcment}^{Q_i}_{\delta}(R) \leq c$ and $\mathsf{err}_f(R) \leq \varepsilon$. Let $r = \max\{q|\ Q_i \geq qR\}$. Let $P_i = R, p_i = q_i * r, i = i + 1$ and go back to step 1.

It can be observed that for each new i, there is a new $x \in \mathcal{X}$ such that $Q_i(x) = 0$. Hence the above process converges after at most $|\mathcal{X}|$ iterations. At the end we have $\mu = \sum_{i=1}^{k} p_i P_i$.

Let us define $M \in [k]$ such that $\Pr[M = i] = p_i$. Let us define $XY \in \mathcal{X} \times \mathcal{Y}$ correlated with M such that $(XY|\ M = i) \sim P_i$. It is easily checked that $XY \sim \mu$. Also since each P_i is one-way for μ, XYM form a Markov chain $M \leftrightarrow X \leftrightarrow Y$. This shows Part 1. and it remains to show Part 2. Let θ be the distribution of XYM. Let us define

1. $B = \{(x,y,i) |\ \log \frac{P_i(x|y)}{\mu(x|y)} > c + \log \frac{1}{\delta}\}$,
2. $B_1 = \{(x,y,i) |\ \log \frac{P_i(x|y)}{Q_i(x|y)} > c\}$,
3. $B_2 = \{(x,y,i) |\ \frac{\mu(y)}{q_i Q_i(y)} > \frac{1}{\delta}\}$.

Since $q_i Q(x,y) \leq \mu(x,y)$,

$$\frac{P_i(x|y)}{\mu(x|y)} = \frac{P_i(x|y)}{Q_i(x|y)} \cdot \frac{Q_i(x|y)}{\mu(x|y)} = \frac{P_i(x|y)}{Q_i(x|y)} \cdot \frac{Q_i(x,y)\mu(y)}{Q_i(y)\mu(x,y)} \leq \frac{P_i(x|y)}{Q_i(x|y)} \cdot \frac{\mu(y)}{q_i Q_i(y)} \ .$$

Therefore $B \subseteq B_1 \cup B_2$. Since for each i, $\mathsf{rcment}_\delta^{Q_i}(P_i) \leq c$, we have

$$\Pr_{(x,y,i)\leftarrow\theta}[(x,y,i) \in B_1] \leq \delta.$$

For a given y, let i_y be the smallest i such that $\frac{\mu(y)}{q_i Q_i(y)} > \frac{1}{\delta}$. Then,

$$\Pr_{(x,y,i)\leftarrow\theta}[(x,y,i) \in B_2] = \sum_y q_{i_y} Q_{i_y}(y) < \sum_y \delta\mu(y) = \delta.$$

Hence, $\Pr_{(x,y,i)\leftarrow\theta}[(x,y,i) \in B] < 2\delta$. Finally we note that,

$$\frac{P_i(x|y)}{\mu(x|y)} = \frac{\theta(x|(y,i))}{\theta(x|y)} = \frac{\theta(x|y)\theta(i|(x,y))}{\theta(i|y)\theta(x|y)} = \frac{\theta(i|x)}{\theta(i|y)} \ .$$

■

Now consider the following one-way private-coin protocol $\mathcal{P}_1$ for f with inputs drawn from distribution μ. In $\mathcal{P}_1$ Alice on input x generates i from the distribution $(M|\ X = x)$ and sends i to Bob. Note that from Part 1. of the above claim, conditioned on Alice's message being i, the joint distribution of the inputs of Alice and Bob is P_i. Bob on input y and receiving message i, gives the best possible output assuming distribution of Alice's inputs being $(X|\ M = i, Y = y)$. Since for all i, $\mathsf{err}_f(P_i) \leq \varepsilon$, the distributional error of $\mathcal{P}_1$ is at most ε. Now using Lemma 2.4 we get a deterministic protocol $\mathcal{P}_2$ for f, with distributional error at most $\varepsilon + 4\delta$ and communication at most $d = \mathsf{rcment}_{\varepsilon,\delta}(f) + O(\log \frac{1}{\delta})$. ■

We can now conclude our characterization.

Theorem 3.5 (New characterization) *Let $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ be a relation and $\varepsilon > 0$. Then,*

$$\mathsf{ment}_{2\varepsilon}(f) - 1 \leq \mathsf{R}_\varepsilon^{1,\mathsf{pub}}(f) \leq \mathsf{rcment}_{\varepsilon/5,\varepsilon/5}(f) + O(\log \frac{1}{\varepsilon}) \ .$$

Hence

$$\mathsf{R}_\varepsilon^{1,\mathsf{pub}}(f) = \Theta(\mathsf{ment}_\varepsilon(f)) = \Theta(\mathsf{rcment}_{\varepsilon,\varepsilon}(f)) \ .$$

Proof: The first inequality follows from Lemma 3.2 (set $k = 1$) and maximizing both sides over all distributions μ and using Lemma 2.3 (Yao principal). The second inequality follows from Lemma 3.3 (set $\varepsilon = \varepsilon, \delta = \varepsilon$) and maximizing both sides over all distributions μ and using Lemma 2.3. The other relations now follow from Lemma 3.1 and from the fact that the error in public-coin randomized one-way communication complexity can be made a constant factor down by increasing the communication by a constant factor. ■

Strong direct product

In this section we show strong direct product theorem for one-way public-coin communication complexity. We start with the following key theorem.

Theorem 3.6 (Direct product in terms of ment and rcment) *Let $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ be a relation and $\mu \in \mathcal{X} \times \mathcal{Y}$ be a distribution. Let $0 < 4\sqrt{80\delta} < \varepsilon < 0.5$ be constants, k be a natural number and $\mathsf{rcment}^{\mu}_{\varepsilon,\varepsilon}(f) \geq 4/\delta$. Then*

$$\mathsf{ment}^{\mu^k}_{1-(1-\varepsilon/2)^{\lfloor \delta k \rfloor}}(f^k) \geq \delta \cdot k \cdot \mathsf{rcment}^{\mu}_{\varepsilon,\varepsilon}(f) \ .$$

Proof: Let $c = \mathsf{rcment}^{\mu}_{\varepsilon,\varepsilon}(f)$. Let $\lambda \in \mathcal{X}^k \times \mathcal{Y}^k$ be a distribution which is one-way for μ^k and with $S_\infty(\lambda||\mu^k) < \delta ck$. We show that $\mathsf{err}_{f^k}(\lambda) \geq 1 - (1-\varepsilon/2)^{\lfloor \delta k \rfloor}$. This shows the desired.

Let B be a set. For a random variable distributed in B^k, or a string in B^k, the portion corresponding to the ith coordinate is represented with subscript i. Also the portion except the ith coordinate is represented with subscript $-i$. Similarly portion corresponding to a subset $C \subseteq [k]$ is represented with subscript C. For joint random variables MN, we let M_n to represent $M|\ (N = n)$ and also $MN|\ (N = n)$ and is clear from the context.

Let $XY \sim \lambda$. Let us fix $g : \mathcal{Y}^k \to \mathcal{Z}^k$. For a coordinate i, let the binary random variable $T_i \in \{0, 1\}$, correlated with XY, denote success in the ith coordinate. That is $T_i = 1$ iff $XY = (x, y)$ such that $(x_i, y_i, g(y)_i) \in f$. Using Claim 3.7 below we get that the overall success is upper bounded as desired:

$$\Pr[T_1 \times T_2 \times \cdots \times T_k = 1] \leq \Pr[T_{i_1} \times T_{i_2} \times \cdots \times T_{i_{k'}} = 1] \leq (1 - \varepsilon/2)^{k'} \ .$$

Claim 3.7 *Let $k' = \lfloor \delta k \rfloor$. There exists k' distinct coordinates $i_1, \ldots, i_{k'}$ such that $\Pr[T_{i_1} = 1] \leq 1 - \varepsilon/2$ and for each $r < k'$,*

1. *either $\Pr[T_{i_1} \times T_{i_2} \times \cdots \times T_{i_r} = 1] \leq (1 - \varepsilon/2)^{k'}$,*
2. *or $\Pr[T_{i_{r+1}} = 1|\ (T_{i_1} \times T_{i_2} \times \cdots \times T_{i_r} = 1)] \leq 1 - \varepsilon/2$.*

Proof: Let us say we have identified $r < k'$ coordinates $i_1, \ldots i_r$. Let $C = \{i_1, i_2, \ldots, i_r\}$. Let $T = T_{i_1} \times T_{i_2} \times \cdots \times T_{i_r}$. If $\Pr[T = 1] \leq (1 - \varepsilon/2)^{k'}$ then we will be done. So assume that $\Pr[T = 1] > (1 - \varepsilon/2)^{k'} \geq 2^{-\delta k}$.

Let $X'Y' \sim \mu$. Let $X^1Y^1 = (XY|\ T = 1)$. Let D be uniformly distributed in $\{0, 1\}^k$ and independent of X^1Y^1. Let $U_i = X^1_i$ if $D_i = 0$ and $U_i = Y^1_i$ if $D_i = 1$. Let $U = U_1 \ldots U_k$. Below for any random variable $\tilde{X}\tilde{Y}$, we let $\tilde{X}\tilde{Y}_{d,u}$, represent the random variable obtained by appropriate conditioning on $\tilde{X}\tilde{Y}$: for all i, $\tilde{X}_i = u_i$ if $d_i = 0$ otherwise $\tilde{Y}_i = u_i$ if $d = 1$. Consider,

$$\begin{aligned}
&\delta k + \delta ck \\
&> S_\infty(X^1Y^1||XY) + S_\infty(XY||(X'Y')^{\otimes k}) \\
&\geq S_\infty(X^1Y^1||(X'Y')^{\otimes k}) \geq S(X^1Y^1||(X'Y')^{\otimes k}) \\
&\geq \mathbb{E}_{(d,u,x_C,y_C)\leftarrow(DUX^1_CY^1_C)} S((X^1Y^1)_{d,u,x_C,y_C}||((X'Y')^{\otimes k})_{d,u,x_C,y_C}) \quad \text{(from Part 2. of Fact 2.1)} \\
&\geq \mathbb{E}_{(d,u,x_C,y_C)\leftarrow(DUX^1_CY^1_C)} S(X^1_{d,u,x_C,y_C}||X'_{d_1,u_1,x_C,y_C} \otimes \ldots \otimes X'_{d_k,u_k,x_C,y_C}) \quad \text{(from Part 2. of Fact 2.1)} \\
&\geq \mathbb{E}_{(d,u,x_C,y_C)\leftarrow(DUX^1_CY^1_C)} \sum_{i \notin C} S((X^1_{d,u,x_C,y_C})_i||X'_{d_i,u_i}) \quad \text{(from Part 2. of Fact 2.1)} \\
&= \sum_{i \notin C} \mathbb{E}_{(d,u,x_C,y_C)\leftarrow(DUX^1_CY^1_C)} S((X^1_{d,u,x_C,y_C})_i||X'_{d_i,u_i}) \ . \qquad (3.1)
\end{aligned}$$

Also

$$\begin{aligned}
\delta k &> S_\infty(X^1Y^1||XY) \geq S(X^1Y^1||XY) = \mathbb{E}_{d\leftarrow D} S(X^1Y^1||XY) \\
&\geq \mathbb{E}_{(d,u,x_C,y_C)\leftarrow(DUX^1_CY^1_C)} S(Y^1_{d,u,x_C,y_C} \mid\mid Y_{d_1,u_1,x_C,y_C} \otimes \ldots \otimes Y_{d_k,u_k,x_C,y_C}) \quad \text{(from Part 2. of Fact 2.1)} \\
&\geq \mathbb{E}_{(d,u,x_C,y_C)\leftarrow(DUX^1_CY^1_C)} \sum_{i\notin C} S((Y^1_{d,u,x_C,y_C})_i||Y_{d_i,u_i}) \quad \text{(from Part 2. of Fact 2.1)} \\
&= \sum_{i\notin C} \mathbb{E}_{(d,u,x_C,y_C)\leftarrow(DUX^1_CY^1_C)} S((Y^1_{d,u,x_C,y_C})_i||Y'_{d_i,u_i}) \ . \qquad (3.2)
\end{aligned}$$

From Eq. 3.1 and Eq. 3.2 and using Markov's inequality we get a coordinate j outside of C such that

1. $\mathbb{E}_{(d,u,x_C,y_C)\leftarrow(DUX^1_CY^1_C)} S((X^1_{d,u,x_C,y_C})_j||X'_{d_j,u_j}) \leq \frac{2\delta(c+1)}{(1-\delta)} \leq 4\delta c$, and
2. $\mathbb{E}_{(d,u,x_C,y_C)\leftarrow(DUX^1_CY^1_C)} S((Y^1_{d,u,x_C,y_C})_j||Y'_{d_j,u_j}) \leq \frac{2\delta}{(1-\delta)} \leq 4\delta$.

Therefore,

$$\begin{aligned}
4\delta c &\geq \mathbb{E}_{(d,u,x_C,y_C)\leftarrow(DUX^1_CY^1_C)} S((X^1_{d,u,x_C,y_C})_j||X'_{d_j,u_j}) \\
&= \mathbb{E}_{(d_{-j},u_{-j},x_C,y_C)\leftarrow(D_{-j}U_{-j}X^1_CY^1_C)} \mathbb{E}_{(d_j,u_j)\leftarrow(D_jU_j)|\ (D_{-j}U_{-j}X^1_CY^1_C)=(d_{-j},u_{-j},x_C,y_C)} S((X^1_{d,u,x_C,y_C})_j||X'_{d_j,u_j}).
\end{aligned}$$

And,

$$\begin{aligned}
4\delta &\geq \mathbb{E}_{(d,u,x_C,y_C)\leftarrow(DUX^1_CY^1_C)} S((Y^1_{d,u,x_C,y_C})_j||Y'_{d_j,u_j}) \\
&= \mathbb{E}_{(d_{-j},u_{-j},x_C,y_C)\leftarrow(D_{-j}U_{-j}X^1_CY^1_C)} \mathbb{E}_{(d_j,u_j)\leftarrow(D_jU_j)|\ (D_{-j}U_{-j}X^1_CY^1_C)=(d_{-j},u_{-j},x_C,y_C)} S((Y^1_{d,u,x_C,y_C})_j||Y'_{d_j,u_j}).
\end{aligned}$$

Now using Markov's inequality, there exists set G_1 with $\Pr[D_{-j}U_{-j}X^1_CY^1_C \in G_1] \geq 1-0.2$, such that for all $(d_{-j},u_{-j},x_C,y_C) \in G_1$,

1. $\mathbb{E}_{(d_j,u_j)\leftarrow(D_jU_j)|\ (D_{-j}U_{-j}X^1_CY^1_C)=(d_{-j},u_{-j},x_C,y_C)} S((X^1_{d,u,x_C,y_C})_j||X'_{d_j,u_j}) \leq 40\delta c$, and
2. $\mathbb{E}_{(d_j,u_j)\leftarrow(D_jU_j)|\ (D_{-j}U_{-j}X^1_CY^1_C)=(d_{-j},u_{-j},x_C,y_C)} S((Y^1_{d,u,x_C,y_C})_j||Y'_{d_j,u_j}) \leq 40\delta$.

Fix $(d_{-j},u_{-j},x_C,y_C) \in G_1$. Conditioning on $D_j = 1$ (which happens with probability 1/2) in inequality 1. above we get,

$$\mathbb{E}_{y_j\leftarrow Y^1_j|(D_{-j}U_{-j}X^1_CY^1_C)=(d_{-j},u_{-j},x_C,y_C)} S((X^1_{d_{-j},u_{-j},y_j,x_C,y_C})_j||X'_{y_j}) \leq 80\delta c. \qquad (3.3)$$

Conditioning on $D_j = 0$ (which happens with probability 1/2) in inequality 2. above we get,

$$\mathbb{E}_{x_j\leftarrow X^1_j|(D_{-j}U_{-j}X^1_CY^1_C)=(d_{-j},u_{-j},x_C,y_C)} S((Y^1_{d_{-j},u_{-j},x_j,x_C,y_C})_j||Y'_{x_j}) \leq 80\delta.$$

Using Part 3. of Fact 2.1 and concavity of square root we get,

$$\mathbb{E}_{x_j\leftarrow X^1_j|(D_{-j}U_{-j}X^1_CY^1_C)=(d_{-j},u_{-j},x_C,y_C)} ||(Y^1_{d_{-j},u_{-j},x_j,x_C,y_C})_j - Y'_{x_j}||_1 \leq \sqrt{80\delta}. \qquad (3.4)$$

Let X^2Y^2 be such that $X^2 \sim (X^1_{d_{-j},u_{-j},x_C,y_C})_j$ and $(Y^2|\ X^2 = x_j) \sim Y'_{x_j}$. From Eq. 3.4 we get,

$$||X^2Y^2 - ((X^1Y^1)_{d_{-j},u_{-j},x_C,y_C})_j||_1 \leq \sqrt{80\delta}. \qquad (3.5)$$

We claim the following which we prove in a bit.

Claim 3.8

$$\Pr_{(x,y)\leftarrow X^2Y^2}\left[\log\frac{X^2Y^2(x|y)}{\mu(x|y)} > c\right] \leq 200\delta + \sqrt{80\delta} < \varepsilon.$$

Using above claim and the fact that from construction X^2Y^2 is one-way for μ we get $\mathsf{rcment}^{\mu}_{\varepsilon}(X^2Y^2) < c$. Hence, $\mathsf{err}_f(X^2Y^2) \geq \varepsilon$ and therefore using Eq. 3.5,

$$\mathsf{err}_f(((X^1Y^1)_{d_{-j},u_{-j},x_C,y_C})_j) \geq \varepsilon - \sqrt{80\delta} \geq \frac{3\varepsilon}{4}.$$

Since conditioned on $(Y^1_{d_{-j},u_{-j},x_C,y_C})_j$, the distribution $(X^1Y^1)_{d_{-j},u_{-j},x_C,y_C}$ is product across the $\mathcal{X}^k$ and $\mathcal{Y}^k$ parts, we have,

$$\Pr[T_j = 1|\ (1, d_{-j}, u_{-j}, x_C, y_C) = (TD_{-j}U_{-j}X_CY_C)] \leq 1 - \mathsf{err}_f(((X^1Y^1)_{d_{-j},u_{-j},x_C,y_C})_j).$$

Therefore overall

$$\Pr[T_j = 1|\ (T = 1)] \leq 0.8(1 - \frac{3\varepsilon}{4}) + 0.2 \leq 1 - \varepsilon/2.$$

■

We return to the proof of Claim 3.8.

Proof of Claim 3.8: Let $X^3Y^3 = ((X^1Y^1)_{d_{-j},u_{-j},x_C,y_C})_j$. From Eq. 3.3 we have,

$$80\delta c \geq \mathbb{E}_{y \leftarrow Y^3} S(X^3_y || X'_y) = \mathbb{E}_{(x,y)\leftarrow X^3Y^3} \log \frac{X^3_y(x)}{\mu(x|y)}.$$

Using Fact 2.2 and Markov's inequality on above we get,

$$161\delta \geq \frac{80\delta c + 1}{c/2 + 1/\delta} \geq \Pr_{(x,y)\leftarrow X^3Y^3}\left[\log \frac{X^3_y(x)}{\mu(x|y)} > \frac{c}{2} + \frac{1}{\delta}\right]. \tag{3.6}$$

Now assume for contradiction that,

$$200\delta + \sqrt{80\delta} < \Pr_{(x,y)\leftarrow X^2Y^2}\left[\log \frac{X^2_y(x)}{\mu(x|y)} > c\right].$$

Using Eq. 3.5 we get,

$$200\delta < \Pr_{(x,y)\leftarrow X^3Y^3}\left[\log \frac{X^2_y(x)}{\mu(x|y)} > c\right]. \tag{3.7}$$

Using Eq. 3.6, Eq. 3.7 and since $c \geq 4/\delta$ we get,

$$39\delta < \Pr_{(x,y)\leftarrow X^3Y^3}\left[\log \frac{X^2_y(x)}{X^3_y(x)} > \frac{c}{2} - \frac{1}{\delta} \geq \frac{1}{\delta}\right].$$

Therefore there exists a y such that,

$$39\delta < \Pr_{x\leftarrow X^3_y}\left[\log \frac{X^2_y(x)}{X^3_y(x)} > \frac{1}{\delta}\right].$$

But this is not possible since both X^2_y and X^3_y are probability distributions. ■

■

We now ready to state and prove the main result of this section.

Theorem 3.9 (Strong direct product for one-way public-coin communication complexity) *Let $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ be a relation. Let $0 < 4\sqrt{80\delta} < \varepsilon < 0.5$ be constants, k be a natural number. Let $\delta' = (1 - \varepsilon/10)^{\lfloor \delta k \rfloor} + 2^{-k}$. Then,*

$$\mathsf{R}^{1,\mathsf{pub}}_{1-\delta'}(f^k) = \Omega\left(k \cdot (\delta \cdot \mathsf{R}^{1,\mathsf{pub}}_{\varepsilon}(f) - O(1))\right) \ .$$

In other words,

$$\mathsf{R}^{1,\mathsf{pub}}_{1-2^{-\Omega(\varepsilon^3 k)}}(f^k) = \Omega\left(k \cdot \left(\varepsilon^2 \cdot \mathsf{R}^{1,\mathsf{pub}}_{\varepsilon}(f) - O(1)\right)\right) \ .$$

Proof: Let μ_1 be a distribution such that $\mathsf{D}^{1,\mu_1}_{\varepsilon}(f) = \mathsf{R}^{1,\mathsf{pub}}_{\varepsilon}(f)$. Let μ be a distribution such that $\mathsf{rcment}^{\mu}_{\varepsilon/5,\varepsilon/5}(f) = \mathsf{rcment}_{\varepsilon/5,\varepsilon/5}(f)$. Then,

$$\begin{aligned}
\delta \cdot k \cdot \mathsf{R}^{1,\mathsf{pub}}_{\varepsilon}(f) &= \delta \cdot k \cdot \mathsf{D}^{1,\mu_1}_{\varepsilon}(f) \\
&= O(\delta \cdot k \cdot \mathsf{rcment}_{\varepsilon/5,\varepsilon/5}(f)) \quad \text{(from Lemma 3.3)} \\
&= O(\delta \cdot k \cdot \mathsf{rcment}^{\mu}_{\varepsilon/5,\varepsilon/5}(f)) \\
&= O(\mathsf{ment}^{\mu^k}_{1-(1-\varepsilon/10)^{\lfloor \delta k \rfloor}}(f^k) + k) \quad \text{(from Theorem 3.6)} \\
&= O(\mathsf{D}^{1,\mu^k}_{1-(1-\varepsilon/10)^{\lfloor \delta k \rfloor}-2^{-k}}(f^k) + k) \quad \text{(from Lemma 3.2)} \\
&= O(\mathsf{R}^{1,\mathsf{pub}}_{1-\delta'}(f^k) + k) \ .
\end{aligned}$$

Hence $\mathsf{R}^{1,\mathsf{pub}}_{1-\delta'}(f^k) = \Omega\left(k \cdot (\delta \cdot \mathsf{R}^{1,\mathsf{pub}}_{\varepsilon}(f) - O(1))\right)$. ■

4 Two-way communication

In this section we discuss the two-way public-coin model. We begin with the necessary definitions.

Definitions

Let $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ be a relation, $\mu, \lambda \in \mathcal{X} \times \mathcal{Y}$ be distributions and $\varepsilon > 0$. Let $XY \sim \mu$ and $X_1Y_1 \sim \lambda$ be random variables. Let $S \subseteq \mathcal{Z}$.

Definition 4.1 (Error of a distribution) *Error of distribution μ with respect to f and answer in S, denoted $\mathsf{err}_{f,S}(\mu)$, is defined as*

$$\mathsf{err}_{f,S}(\mu) \stackrel{\text{def}}{=} \min\left\{ \Pr_{(x,y)\leftarrow\mu}[(x,y,z) \notin f] \mid z \in S \right\} \ .$$

Let μ the distribution of inputs of Alice and Bob conditioned on a message transcript in a two-way deterministic protocol. Then if Alice and Bob give an answer in S, they make error on at least $\mathsf{err}_{f,S}(\mu)$ fraction of the inputs.

Definition 4.2 (Essentialness of an answer subset) *Essentialness of answer in S for f with respect to distribution μ, denoted $\mathsf{ess}^{\mu}(f,S)$, is defined as*

$$\mathsf{ess}^{\mu}(f,S) \stackrel{\text{def}}{=} 1 - \Pr_{(x,y)\leftarrow\mu}[\textit{there exists } z \notin S \textit{ such that } (x,y,z) \in f].$$

Above $\mathsf{ess}^{\mu}(f,S)$ represents the fraction of inputs according to μ for which any correct answer must lie in S. For example $\mathsf{ess}^{\mu}(f,\mathcal{Z}) = 1$.

Definition 4.3 (One-way distributions) *λ is called one-way for μ with respect to $\mathcal{X}$, if for all (x, y) in the support of λ we have $\mu(y|x) = \lambda(y|x)$. Similarly λ is called one-way for μ with respect to $\mathcal{Y}$, if for all (x, y) in the support of λ we have $\mu(x|y) = \lambda(x|y)$.*

Definition 4.4 (SM-like) *λ is called SM-like (simultaneous-message-like) for μ, if there is a distribution θ on $\mathcal{X} \times \mathcal{Y}$ such that θ is one-way for μ with respect to $\mathcal{X}$ and λ is one-way for θ with respect to $\mathcal{Y}$.*

Let the inputs of Alice and Bob be distributed according to μ. Then note that conditioned on any message transcript in a two-way deterministic protocol, the resulting distribution on the inputs will be SM-like for μ.

Definition 4.5 (Conditional relative entropy) *The $\mathcal{Y}$-conditional relative entropy of λ with respect to μ, denoted $\mathsf{crent}^{\mu}_{\mathcal{Y}}(\lambda)$, is defined as*

$$\mathsf{crent}^{\mu}_{\mathcal{Y}}(\lambda) \stackrel{\text{def}}{=} \mathbb{E}_{y \leftarrow Y_1} S((X_1)_y || X_y).$$

Similarly the $\mathcal{X}$-conditional relative entropy of λ with respect to μ, denoted $\mathsf{crent}^{\mu}_{\mathcal{X}}(\lambda)$, is defined as

$$\mathsf{crent}^{\mu}_{\mathcal{X}}(\lambda) \stackrel{\text{def}}{=} \mathbb{E}_{x \leftarrow X_1} S((Y_1)_x || Y_x).$$

Definition 4.6 (Conditional relative entropy bound) *The two-way ε-error conditional relative entropy bound of f with answer in S with respect to distribution μ, denoted $\mathsf{crent}^{2,\mu}_{\varepsilon}(f, S)$, is defined as*

$$\mathsf{crent}^{2,\mu}_{\varepsilon}(f, S) \stackrel{\text{def}}{=} \min \left\{ \mathsf{crent}^{\mu}_{\mathcal{X}}(\lambda) + \mathsf{crent}^{\mu}_{\mathcal{Y}}(\lambda) \mid \lambda \text{ is SM-like for } \mu \text{ and } \mathsf{err}_{f,S}(\lambda) \leq \varepsilon \right\} .$$

We use the above definition as follows. Let λ be SM-like for μ and $\mathsf{crent}^{\mu}_{\mathcal{X}}(\lambda) + \mathsf{crent}^{\mu}_{\mathcal{Y}}(\lambda) < c$. Then $\mathsf{err}_{f,S}(\lambda) > \varepsilon$.

The following bound is analogous to a bound defined in [JKN08] where it was referred to as the two-way subdistribution bound. We call it differently here for consistency of nomenclature with the other bounds. [JKN08] typically considered the cases where $S = \mathcal{Z}$ or S is a singleton set.

Definition 4.7 (Relative min entropy bound) *The two-way ε-error relative min entropy bound of f with answer in S with respect to distribution μ, denoted $\mathsf{ment}^{2,\mu}_{\varepsilon}(f, S)$, is defined as*

$$\mathsf{ment}^{2,\mu}_{\varepsilon}(f, S) \stackrel{\text{def}}{=} \min \left\{ S_{\infty}(\lambda || \mu) | \ \lambda \text{ is SM-like for } \mu \text{ and } \mathsf{err}_{f,S}(\lambda) \leq \varepsilon \right\} .$$

The following is easily seen from definitions and Part 2. of Fact 2.1.

Lemma 4.1

$$\mathsf{crent}^{\mu}_{\mathcal{X}}(\lambda) + \mathsf{crent}^{\mu}_{\mathcal{Y}}(\lambda) \leq 2 \cdot S_{\infty}(\lambda || \mu) \quad \text{and} \quad \mathsf{crent}^{2,\mu}_{\varepsilon}(f, S) \leq 2 \cdot \mathsf{ment}^{2,\mu}_{\varepsilon}(f, S).$$

The following lemma states that when μ is a product distribution then ment is upper bounded by crent.

Lemma 4.2 *Let μ be a product distribution across $\mathcal{X}$ and $\mathcal{Y}$. Then,*

$$\mathsf{ment}^{2,\mu}_{\varepsilon}(f, S) = O\left(\frac{1}{\varepsilon} (\mathsf{crent}^{2,\mu}_{\varepsilon/2}(f, S) + 1) \right).$$

Proof: Let $\mu = \mu_1 \otimes \mu_2$, where $\mu_1 \in \mathcal{X}$ and $\mu_2 \in \mathcal{Y}$. Let λ be a distribution such that $\mathsf{crent}^{\mu}_{\mathcal{X}}(\lambda) + \mathsf{crent}^{\mu}_{\mathcal{Y}}(\lambda) = \mathsf{crent}^{2,\mu}_{\varepsilon/2}(f,S)$; λ is SM-like for μ and $\mathsf{err}_{f,S}(\lambda) \leq \varepsilon/2$. Since λ is SM-like for μ, it is easily verified that λ is also a product distribution across $\mathcal{X}$ and $\mathcal{Y}$. Let $\lambda = \lambda_1 \otimes \lambda_2$, where $\lambda_1 \in \mathcal{X}$ and $\lambda_2 \in \mathcal{Y}$. Using Part 4. of Fact 2.1 we get that there exists distributions λ'_1 and λ'_2 such that $S_\infty(\lambda'_1||\mu_1) \leq O(\frac{1}{\varepsilon}(S(\lambda_1||\mu_1)+1))$, $S_\infty(\lambda'_2||\mu_2) \leq O(\frac{1}{\varepsilon}(S(\lambda_2||\mu_2)+1))$, $||\lambda'_1 - \lambda_1||_1 \leq \frac{\varepsilon}{4}$ and $||\lambda'_2 - \lambda_2||_1 \leq \frac{\varepsilon}{4}$. Let $\lambda' = \lambda'_1 \otimes \lambda'_2$. Note that λ' is SM-like for μ. Since $||\lambda' - \lambda||_1 \leq \frac{\varepsilon}{2}$ and $\mathsf{err}_{f,S}(\lambda) \leq \varepsilon/2$, we have $\mathsf{err}_{f,S}(\lambda') \leq \varepsilon$. Now,

$$S_\infty(\lambda'||\mu) = S_\infty(\lambda'_1||\mu_1) + S_\infty(\lambda'_2||\mu_2) = O\left(\frac{1}{\varepsilon}(S(\lambda_1||\mu_1) + S(\lambda_2||\mu_2) + 1)\right)$$
$$= O\left(\frac{1}{\varepsilon}(\mathsf{crent}^{\mu}_{\mathcal{X}}(\lambda) + \mathsf{crent}^{\mu}_{\mathcal{Y}}(\lambda) + 1)\right) \quad = O\left(\frac{1}{\varepsilon}(\mathsf{crent}^{2,\mu}_{\varepsilon/2}(f,S) + 1)\right)$$

Therefore $\mathsf{ment}^{2,\mu}_{\varepsilon}(f,S) = O\left(\frac{1}{\varepsilon}(\mathsf{crent}^{2,\mu}_{\varepsilon/2}(f,S)+1)\right)$. ∎

4.1 Strong direct product

We start with the following theorem.

Theorem 4.3 (Direct product in terms of ment and crent) *Let $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ be a relation, $\mu \in \mathcal{X} \times \mathcal{Y}$ be a distribution and $S \subseteq \mathcal{Z}$. Let $0 < \varepsilon < 1/3$, $0 < 200\delta < 1$ and k be a natural number. Fix $z \in \mathcal{Z}^k$. Let the number of indices $i \in [k]$ with $z_i \in S$ be at least $\delta_1 k$. Then*

$$\mathsf{ment}^{2,\mu^k}_{1-(1-\varepsilon/2)^{\lfloor \delta\delta_1 k \rfloor}}(f^k, \{z\}) \geq \delta \cdot \delta_1 \cdot k \cdot \mathsf{crent}^{2,\mu}_{\varepsilon}(f,S) \ .$$

Proof: Let $c = \mathsf{crent}^{2,\mu}_{\varepsilon}(f,S)$. Let $\lambda \in \mathcal{X}^k \times \mathcal{Y}^k$ be a distribution which is SM-like for μ^k and with $S_\infty(\lambda||\mu^k) < \delta\delta_1 ck$. We show that $\mathsf{err}_{f^k,\{z\}}(\lambda) \geq 1 - (1-\varepsilon/2)^{\lfloor \delta\delta_1 k \rfloor}$. This shows the desired.

We use similar notations as in the previous section. Let $XY \sim \lambda$. For a coordinate i, let the binary random variable $T_i \in \{0,1\}$, correlated with XY, denote success in the ith coordinate. That is $T_i = 1$ iff $XY = (x,y)$ such that $(x_i, y_i, z_i) \in f$. Using Claim 4.4 we conclude the desired upper bound on the overall success:

$$\Pr[T_1 \times T_2 \times \cdots \times T_k = 1] \leq \Pr[T_{i_1} \times T_{i_2} \times \cdots \times T_{i_{k'}} = 1] \leq (1-\varepsilon/2)^{k'} \ .$$

Claim 4.4 *Let $k' = \lfloor \delta\delta_1 k \rfloor$. There exists k' distinct coordinates $i_1, \ldots, i_{k'}$ such that $\Pr[T_{i_1} = 1] \leq 1 - \varepsilon/2$ and for each $r < k'$,*

1. *either $\Pr[T_{i_1} \times T_{i_2} \times \cdots \times T_{i_r} = 1] \leq (1-\varepsilon/2)^{k'}$,*
2. *or $\Pr[T_{i_{r+1}} = 1 |\ (T_{i_1} \times T_{i_2} \times \cdots \times T_{i_r} = 1)] \leq 1 - \varepsilon/2$.*

Proof: Let us say we have identified $r < k'$ coordinates $i_1, \ldots i_r$. Let $C = \{i_1, i_2, \ldots, i_r\}$. Let $T = T_1 \times T_2 \times \cdots \times T_r$. If $\Pr[T=1] \leq (1-\varepsilon/2)^{k'}$ then we will be done. So assume that $\Pr[T=1] > (1-\varepsilon/2)^{k'} \geq 2^{-\delta\delta_1 k}$. Let $X'Y' \sim \mu$. Let $X^1Y^1 = (XY |\ T = 1)$. Let D be uniformly distributed in $\{0,1\}^k$ and independent of X^1Y^1. Let $U_i = X^1_i$ if $D_i = 0$ and $U_i = Y^1_i$ if $D_i = 1$. Let $U = U_1 \ldots U_k$. Below for any random variable $\tilde{X}\tilde{Y}$, we let $\tilde{X}\tilde{Y}_{d,u}$, represent the random variable obtained by appropriate conditioning on $\tilde{X}\tilde{Y}$: for all i, $\tilde{X}_i = u_i$ if $d_i = 0$

otherwise $\tilde{Y}_i = u_i$ if $d = 1$. Let I be the set of indices i such that $z_i \in S$. Consider,

$$
\begin{aligned}
&\delta\delta_1 k + \delta\delta_1 ck \\
&> S_\infty(X^1Y^1||XY) + S_\infty(XY||(X'Y')^{\otimes k}) \\
&\geq S_\infty(X^1Y^1||(X'Y')^{\otimes k}) \geq S(X^1Y^1||(X'Y')^{\otimes k}) \\
&\geq \mathbb{E}_{(d,u,x_C,y_C)\leftarrow(DUX_C^1Y_C^1)} S((X^1Y^1)_{d,u,x_C,y_C}||((X'Y')^{\otimes k})_{d,u,x_C,y_C}) \quad \text{(from Part 2. of Fact 2.1)} \\
&\geq \mathbb{E}_{(d,u,x_C,y_C)\leftarrow(DUX_C^1Y_C^1)} S(X^1_{d,u,x_C,y_C}||X'_{d_1,u_1,x_C,y_C} \otimes \ldots \otimes X'_{d_k,u_k,x_C,y_C}) \quad \text{(from Part 2. of Fact 2.1)} \\
&\geq \mathbb{E}_{(d,u,x_C,y_C)\leftarrow(DUX_C^1Y_C^1)} \sum_{i\notin C, i\in I} S((X^1_{d,u,x_C,y_C})_i||X'_{d_i,u_i}) \quad \text{(from Part 2. of Fact 2.1)} \\
&= \sum_{i\notin C, i\in I} \mathbb{E}_{(d,u,x_C,y_C)\leftarrow(DUX_C^1Y_C^1)} S((X^1_{d,u,x_C,y_C})_i||X'_{d_i,u_i}) \ . \qquad (4.1)
\end{aligned}
$$

Similarly,

$$
\delta\delta_1 k + \delta\delta_1 ck > \sum_{i\notin C, i\in I} \mathbb{E}_{(d,u,x_C,y_C)\leftarrow(DUX_C^1Y_C^1)} S((Y^1_{d,u,x_C,y_C})_i||Y'_{d_i,u_i}) \ . \qquad (4.2)
$$

From Eq. 4.1 and Eq. 4.2 and using Markov's inequality we get a coordinate j outside of C but in I such that

1. $\mathbb{E}_{(d,u,x_C,y_C)\leftarrow(DUX_C^1Y_C^1)} S((X^1_{d,u,x_C,y_C})_j||X'_{d_j,u_j}) \leq \frac{2\delta(c+1)}{(1-\delta)} \leq 4\delta c$, and
2. $\mathbb{E}_{(d,u,x_C,y_C)\leftarrow(DUX_C^1Y_C^1)} S((Y^1_{d,u,x_C,y_C})_j||Y'_{d_j,u_j}) \leq \frac{2\delta(c+1)}{(1-\delta)} \leq 4\delta c$.

Therefore,

$$
\begin{aligned}
4\delta c &\geq \mathbb{E}_{(d,u,x_C,y_C)\leftarrow(DUX_C^1Y_C^1)} S((X^1_{d,u,x_C,y_C})_j||X'_{d_j,u_j}) \\
&= \mathbb{E}_{(d_{-j},u_{-j},x_C,y_C)\leftarrow(D_{-j}U_{-j}X_C^1Y_C^1)} \mathbb{E}_{(d_j,u_j)\leftarrow(D_jU_j)|\ (D_{-j}U_{-j}X_C^1Y_C^1)=(d_{-j},u_{-j},x_C,y_C)} S((X^1_{d,u,x_C,y_C})_j||X'_{d_j,u_j}).
\end{aligned}
$$

And,

$$
\begin{aligned}
4\delta c &\geq \mathbb{E}_{(d,u,x_C,y_C)\leftarrow(DUX_C^1Y_C^1)} S((Y^1_{d,u,x_C,y_C})_j||Y'_{d_j,u_j}) \\
&= \mathbb{E}_{(d_{-j},u_{-j},x_C,y_C)\leftarrow(D_{-j}U_{-j}X_C^1Y_C^1)} \mathbb{E}_{(d_j,u_j)\leftarrow(D_jU_j)|\ (D_{-j}U_{-j}X_C^1Y_C^1)=(d_{-j},u_{-j},x_C,y_C)} S((Y^1_{d,u,x_C,y_C})_j||Y'_{d_j,u_j}).
\end{aligned}
$$

Now using Markov's inequality, there exists set G_1 with $\Pr[D_{-j}U_{-j}X_C^1Y_C^1 \in G_1] \geq 1 - 0.2$, such that for all $(d_{-j}, u_{-j}, x_C, y_C) \in G_1$,

1. $\mathbb{E}_{(d_j,u_j)\leftarrow(D_jU_j)|\ (D_{-j}U_{-j}X_C^1Y_C^1)=(d_{-j},u_{-j},x_C,y_C)} S((X^1_{d,u,x_C,y_C})_j||X'_{d_j,u_j}) \leq 40\delta c$, and
2. $\mathbb{E}_{(d_j,u_j)\leftarrow(D_jU_j)|\ (D_{-j}U_{-j}X_C^1Y_C^1)=(d_{-j},u_{-j},x_C,y_C)} S((Y^1_{d,u,x_C,y_C})_j||Y'_{d_j,u_j}) \leq 40\delta c$.

Fix $(d_{-j}, u_{-j}, x_C, y_C) \in G_1$. Conditioning on $D_j = 1$ (which happens with probability $1/2$) in inequality 1. above we get,

$$
\mathbb{E}_{y_j\leftarrow Y_j^1|(D_{-j}U_{-j}X_C^1Y_C^1)=(d_{-j},u_{-j},x_C,y_C)} S((X^1_{d_{-j},u_{-j},y_j,x_C,y_C})_j||X'_{y_j}) \leq 80\delta c. \qquad (4.3)
$$

Conditioning on $D_j = 0$ (which happens with probability $1/2$) in inequality 2. above we get,

$$
\mathbb{E}_{x_j\leftarrow X_j^1|(D_{-j}U_{-j}X_C^1Y_C^1)=(d_{-j},u_{-j},x_C,y_C)} S((Y^1_{d_{-j},u_{-j},x_j,x_C,y_C})_j||Y'_{x_j}) \leq 80\delta c. \qquad (4.4)
$$

Let $X^2Y^2 = ((X^1Y^1)_{d_{-j},u_{-j},x_C,y_C})_j$. Note that X^2Y^2 is SM-like for μ. From Eq. 4.3 and Eq. 4.4 we get that

$$
\mathsf{crent}^\mu_{\mathcal{X}}(X^2Y^2) + \mathsf{crent}^\mu_{\mathcal{Y}}(X^2Y^2) \leq c.
$$

Hence, $\mathsf{err}_f(((X^1Y^1)_{d_{-j},u_{-j},x_C,y_C})_j) \geq \varepsilon$. This implies,

$$\Pr[T_j = 1|\ (1, d_{-j}, u_{-j}, x_C, y_C) = (TD_{-j}U_{-j}X_CY_C)] \leq 1 - \varepsilon.$$

Therefore overall

$$\Pr[T_j = 1|\ (T = 1)] \leq 0.8(1 - \varepsilon) + 0.2 \leq 1 - \varepsilon/2.$$

■

■

We can now state and prove the main result of this section.

Theorem 4.5 (Direct product in terms of D and crent) *Let $f \subseteq \mathcal{X} \times \mathcal{Y} \times \mathcal{Z}$ be a relation, $\mu \in \mathcal{X} \times \mathcal{Y}$ be a distribution and $S \subseteq \mathcal{Z}$. Let $0 < \varepsilon < 1/3$ and k be a natural number. Let $\delta_2 = \mathsf{ess}^{\mu}(f, S)$. Let $0 < 200\delta < \delta_2$. Let $\delta' = 3(1 - \varepsilon/2)^{\lfloor \delta\delta_2 k/2 \rfloor}$. Then,*

$$\mathsf{D}^{2,\mu^k}_{1-\delta'}(f^k) \geq \delta \cdot \delta_2 \cdot k \cdot \mathsf{crent}^{2,\mu}_{\varepsilon}(f, S) - k\ .$$

Proof: Let $\mathsf{crent}^{2,\mu}_{\varepsilon}(f, S) = c$. For input $(x, y) \in \mathcal{X}^k \times \mathcal{Y}^k$, let $b(x, y)$ be the number of indices i in $[k]$ for which there exists $z_i \notin S$ such that $(x_i, y_i, z_i) \in f$. Let

$$B = \left\{(x, y) \in \mathcal{X}^k \times \mathcal{Y}^k |\ b(x, y) \geq (1 - \delta_2/2)k\right\}.$$

By Chernoff's inequality we get,

$$\Pr_{(x,y)\leftarrow\mu^k}[(x, y) \in B] \leq \exp(-\delta_2^2 k/2).$$

Let $\mathcal{P}$ be a protocol for f^k with inputs $XY \sim \mu^k$ with communication at most $d = (kc\delta\delta_2/2) - k$ bits. Let $M \in \mathcal{M}$ represent the message transcript of $\mathcal{P}$. Let

$$B_M = \{m \in \mathcal{M}|\ \Pr[(XY)_m \in B] \geq \exp(-\delta_2^2 k/4)\}.$$

Then $\Pr[M \in B_M] \leq \exp(-\delta_2^2 k/4)$. Let

$$B^1_M = \{m \in \mathcal{M}|\ \Pr[M = m] \leq 2^{-d-k}\}.$$

Then $\Pr[M \in B^1_M] \leq 2^{-k}$. Fix $m \notin B_M \cup B^1_M$. Let z_m be the output of $\mathcal{P}$ when $M = m$. Let $b(z_m)$ be the number of indices i such that $z_{m,i} \notin S$. If $b(z_m) \geq 1 - \delta_2 k/2$ then success of $\mathcal{P}$ when $M = m$ is at most $\exp(-\delta_2^2 k/4) \leq (1 - \varepsilon/2)^{\lfloor \delta\delta_2 k/2 \rfloor}$. If $b(z_m) < 1 - \delta_2 k/2$ then from Theorem 4.3 (by setting $z = z_m$ and $\delta_1 = \delta_2/2$), success of $\mathcal{P}$ when $M = m$ is at most $(1 - \varepsilon/2)^{\lfloor \delta\delta_2 k/2 \rfloor}$. Therefore overall success of $\mathcal{P}$ is at most

$$\delta' = 2^{-k} + \exp(-\delta_2^2 k/4) + (1 - 2^{-k} - \exp(-\delta_2^2 k/4)(1 - \varepsilon/2)^{\lfloor \delta\delta_2 k/2 \rfloor} \leq 3(1 - \varepsilon/2)^{\lfloor \delta\delta_2 k/2 \rfloor}.$$

■

We point that when μ is a product distribution, the result above and Lemma 4.2 imply the direct product result of [JKN08] in terms of $\mathsf{ment}^{2,\mu}_{\varepsilon}(f, S)$.

5 Strong direct product for set disjointness

In this section we present an application of Theorem 4.5 to show a strong direct product result for two-way public-coin communication complexity of the set-disjointness function. For a string $x \in \{0, 1\}^n$ we let x also represent the subset of $[n]$ for which x is the characteristic vector. The set disjointness function $\mathsf{disj}_n : \{0, 1\}^n \times \{0, 1\}^n \to \{0, 1\}$ is defined as $\mathsf{disj}_n(x, y) = 1$ iff the subsets x and y do not intersect.

Theorem 5.1 (Strong Direct product for set disjointness) *Let k be a natural number. Then* $\mathsf{R}^{2,\mathsf{pub}}_{1-2^{-\Omega(k)}}(\mathsf{disj}_n^k) = \Omega(k \cdot n)$.

Proof: Let $n = 4l - 1$ (for some integer l). Let $T = (T_1, T_2, I)$ be a uniformly random partition of $[n]$ into three disjoint sets such that $|T_1| = |T_2| = 2l - 1$ and $|I| = 1$. Conditioned on $T = t = (t_1, t_2, \{i\})$, let X be a uniformly random subset of $t_1 \cup \{i\}$ and Y be a uniformly random subset of $t_2 \cup \{i\}$. Note that $X \leftrightarrow T \leftrightarrow Y$ is a Markov chain. It is easily seen that $\mathsf{ess}^{XY}(\mathsf{disj}_n, \{1\}) = 0.75$. Therefore using Theorem 4.5, Lemma 5.2 (below) and Lemma 2.5 (Yao principal) we conclude the desired,

$$\mathsf{R}^{2,\mathsf{pub}}_{1-2^{-\Omega(k)}}(\mathsf{disj}_n^k) = \Omega(k \cdot n).$$

Lemma 5.2 $\mathsf{crent}^{2,XY}_{1/70}(\mathsf{disj}_n, \{1\}) = \Omega(n)$.

Proof: Our proof follows on similar lines as the proof of Razborov [Raz92] showing linear lower bound on the rectangle bound for set-disjointness (see e.g. [KN97], Lemma 4.49). However there are important differences since we are lower bounding a weaker quantity.

Let $\delta = 1/(200)^2$. Let $X'Y'$ be such that $\mathsf{crent}^{XY}_{\mathcal{X}}(X'Y') + \mathsf{crent}^{XY}_{\mathcal{Y}}(X'Y') \leq \delta n$ and $X'Y'$ is SM-like for XY. We will show that $\mathsf{err}_{\mathsf{disj}_n,\{1\}}(X'Y') = \Pr[\mathsf{disj}_n(X'Y') = 0] \geq 1/70$. This will show the desired. We assume that $\Pr[\mathsf{disj}_n(X'Y') = 1] \geq 0.5$ otherwise we are done already. Let $A, B \in \{0,1\}$ be binary random variables such that $A \leftrightarrow X \leftrightarrow Y \leftrightarrow B$ and $X'Y' = (XY|\ A = B = 1)$ (it is easily verified that such A, B must exist since $X'Y'$ is SM-like for XY).

Claim 5.3

1. $\Pr[A = B = 1, \mathsf{disj}_n(XY) = 0]$
   $= \frac{1}{4}\mathbb{E}_{t=(t_1,t_2,\{i\})\leftarrow T} \Pr[A = 1|\ T = t, X_i = 1] \Pr[B = 1|\ T = t, Y_i = 1].$

2. $\Pr[A = B = 1, \mathsf{disj}_n(XY) = 1]$
   $= \frac{3}{4}\mathbb{E}_{t=(t_1,t_2,\{i\})\leftarrow T} \Pr[A = 1|\ T = t, X_i = 0] \Pr[B = 1|\ T = t, Y_i = 0].$

Proof: We first show Part 1.

$$\begin{aligned}
&\Pr[A = B = 1, \mathsf{disj}_n(XY) = 0] = \Pr[A = B = 1, X_I = Y_I = 1] \\
&= \mathbb{E}_{t=(t_1,t_2,\{i\})\leftarrow T} \Pr[A = B = 1, X_i = Y_i = 1|\ T = t] \\
&= \mathbb{E}_{t=(t_1,t_2,\{i\})\leftarrow T} \Pr[X_i = Y_i = 1|\ T = t] \Pr[A = B = 1|\ T = t, X_i = Y_i = 1] \\
&= \frac{1}{4}\mathbb{E}_{t=(t_1,t_2,\{i\})\leftarrow T} \Pr[A = B = 1|\ T = t, X_i = Y_i = 1] \\
&= \frac{1}{4}\mathbb{E}_{t=(t_1,t_2,\{i\})\leftarrow T} \Pr[A = 1|\ T = t, X_i = 1] \Pr[B = 1|\ T = t, Y_i = 1].
\end{aligned}$$

Now we show Part 2. Note that the distribution of $(XY|\ \mathsf{disj}_n(X, Y) = 1)$ is identical to the distribution of $(XY|\ X_I = Y_I = 0)$ (both being uniform distribution on disjoint x, y such that

$|x| = |y| = l$). Also $\Pr[\mathsf{disj}_n(XY) = 1] = 3\Pr[X_I = Y_I = 0]$. Therefore,

$$\begin{aligned}
&\Pr[A = B = 1, \mathsf{disj}_n(XY) = 1] = \Pr[\mathsf{disj}_n(XY) = 1]\Pr[A = B = 1|\ \mathsf{disj}_n(XY) = 1] \\
&= 3\Pr[X_I = Y_I = 0]\Pr[A = B = 1|\ X_I = Y_I = 0] = 3\Pr[A = B = 1, X_I = Y_I = 0] \\
&= 3\mathbb{E}_{t=(t_1,t_2,\{i\})\leftarrow T}\Pr[A = B = 1, X_i = 0, Y_i = 0|\ T = t] \\
&= 3\mathbb{E}_{t=(t_1,t_2,\{i\})\leftarrow T}\Pr[X_i = 0, Y_i = 0|\ T = t]\Pr[A = B = 1|\ T = t, X_i = 0, Y_i = 0] \\
&= \frac{3}{4}\mathbb{E}_{t=(t_1,t_2,\{i\})\leftarrow T}\Pr[A = B = 1|\ T = t, X_i = 0, Y_i = 0] \\
&= \frac{3}{4}\mathbb{E}_{t=(t_1,t_2,\{i\})\leftarrow T}\Pr[A = 1|\ T = t, X_i = 0]\Pr[B = 1|\ T = t, Y_i = 0].
\end{aligned}$$

■

Recall that in our notation $X_{t_2} = (X|\ T_2 = t_2)$ and so on.

Claim 5.4 *Let* $B_x^1 = \{t_2|\ S(X'_{t_2}||X_{t_2}) > 100\delta n\}, \quad B_y^1 = \{t_1|\ S(Y'_{t_1}||Y_{t_1}) > 100\delta n\}$.

$$B_x^2 = \{t|\ \Pr[A = 1|\ X_i = 1, T = t] < \frac{1}{3}\Pr[A = 1|\ X_i = 0, T = t]\}.$$

$$B_y^2 = \{t|\ \Pr[B = 1|\ Y_i = 1, T = t] < \frac{1}{3}\Pr[B = 1|\ Y_i = 0, T = t]\}.$$

1. $\Pr[A = B = 1, T_2 \in B_x^1] < \frac{1}{100}\Pr[A = B = 1]$.
2. $\Pr[A = B = 1, T_1 \in B_y^1] < \frac{1}{100}\Pr[A = B = 1]$.
3. *Let* $t_2 \notin B_x^1$, *then* $\Pr[T \in B_x^2|\ T_2 = t_2] < \frac{1}{100}$.
4. *Let* $t_1 \notin B_y^1$, *then* $\Pr[T \in B_y^2|\ T_1 = t_1] < \frac{1}{100}$.

Proof: We show the proof of Part 1. and Part 2. follows similarly. Let $T' = (T|\ A = B = 1)$. Consider (below the second equality follows since $\forall(x, y) :\ (T|\ XY = (x, y))$ is identically distributed as $(T'|\ X'Y' = (x, y))$) and using Part 2. of Fact 2.1),

$$\begin{aligned}
\delta n &\geq \mathsf{crent}_{\mathcal{Y}}^{XY}(X'Y') = \mathbb{E}_{y\leftarrow Y'}S(X'_y||X_y) \\
&= \mathbb{E}_{y\leftarrow Y'}S((X'T')_y||(XT)_y) \\
&\geq \mathbb{E}_{(y,t)\leftarrow(Y'T')}S(X'_{y,t}||X_{y,t}) \quad \text{(from Part 2. of Fact 2.1)} \\
&= \mathbb{E}_{t\leftarrow T'}S(X'_t||X_t) \quad \text{(since } X \leftrightarrow T \leftrightarrow Y \text{ and } X' \leftrightarrow T' \leftrightarrow Y' \text{ are Markov chains)} \\
&= \mathbb{E}_{t_2\leftarrow T'_2}S(X'_{t_2}||X_{t_2}).
\end{aligned}$$

Above the last equality follows since for all $t = (t_1, t_2, \{i\})$, X_{t_2} is identically distributed as X_t and similarly X'_{t_2} is identically distributed as X'_t . Therefore using Markov's inequality,

$$\frac{1}{100} > \Pr[T'_2 \in B_x^1] = \Pr[T_2 \in B_x^1|\ A = B = 1] = \frac{\Pr[T_2 \in B_x^1, A = B = 1]}{\Pr[A = B = 1]}.$$

We show the proof of Part 3. and Part 4. follows similarly. Fix $t_2 \notin B_x^1$. Then using Part 2. of Fact 2.1 (recall that in our notation X_i represents the i-th bit of X and so on),

$$100\delta n \geq S(X'_{t_2}||X_{t_2}) \geq \sum_{i\notin t_2} S((X'_{t_2})_i||(X_{t_2})_i).$$

Let $R = \{i \notin t_2|\ S((X'_{t_2})_i||(X_{t_2})_i) > 0.01\}$. From above $\frac{|R|}{2l} < \frac{1}{100}$. For $i \notin R \cup t_2$ (using Part 3. of Fact 2.1),

$$S((X'_{t_2})_i||(X_{t_2})_i) \leq 0.01 \quad \Rightarrow \quad ||(X'_{t_2})_i - (X_{t_2})_i||_1 \leq \sqrt{0.01} = 0.1$$
$$\Rightarrow \Pr[(X'_{t_2})_i = 1] \geq 0.4 \geq \frac{1}{3}\Pr[(X'_{t_2})_i = 0] \quad (\text{since } \Pr[(X_{t_2})_i = 1] = 0.5)$$
$$\Rightarrow \Pr[X_i = 1|\ T_2 = t_2, A = 1] \geq \frac{1}{3}\Pr[X_i = 0|\ T_2 = t_2, A = 1]$$
$$\Rightarrow \frac{\Pr[A = 1|\ T_2 = t_2]}{\Pr[X_i = 1|\ T_2 = t_2]}\Pr[X_i = 1|\ T_2 = t_2, A = 1] \geq \frac{1}{3}\frac{\Pr[A = 1|\ T_2 = t_2]}{\Pr[X_i = 0|\ T_2 = t_2]}\Pr[X_i = 0|\ T_2 = t_2, A = 1]$$
$$\Rightarrow \Pr[A = 1|\ X_i = 1, T_2 = t_2] \geq \frac{1}{3}\Pr[A = 1|\ X_i = 0, T_2 = t_2].$$

Therefore $i \notin R \cup t_2$ implies $t = (t_1, t_2, \{i\}) \notin B^2_x$. Hence,

$$\Pr[T \in B^2_x|\ T_2 = t_2] \leq \Pr[i \in R|\ T_2 = t_2] = \frac{|R|}{2l} < \frac{1}{100}.$$

■

Claim 5.5 *1. Let $Bad^1_x = 1$ iff $T_2 \in B^1_x$ otherwise 0. Then*

$$\mathbb{E}_{t=(t_1,t_2,\{i\})\leftarrow T}\Pr[A = 1|\ X_i = 0, T = t]\Pr[B = 1|\ Y_i = 0, T = t]Bad^1_x$$
$$\leq \frac{6}{100}\mathbb{E}_{t=(t_1,t_2,\{i\})\leftarrow T}\Pr[A = 1|\ X_i = 0, T = t]\Pr[B = 1|\ Y_i = 0, T = t].$$

2. Let $Bad^1_y = 1$ iff $T_1 \in B^1_y$ otherwise 0. Then

$$\mathbb{E}_{t=(t_1,t_2,\{i\})\leftarrow T}\Pr[A = 1|\ X_i = 0, T = t]\Pr[B = 1|\ Y_i = 0, T = t]Bad^1_y$$
$$\leq \frac{6}{100}\mathbb{E}_{t=(t_1,t_2,\{i\})\leftarrow T}\Pr[A = 1|\ X_i = 0, T = t]\Pr[B = 1|\ Y_i = 0, T = t].$$

3. Fix $t_2 \notin B^1_x$. Let $T_{t_2} = (T|\ T_2 = t_2)$. Let $Bad^2_x = 1$ iff $T \in B^2_x$ otherwise 0. Then

$$\mathbb{E}_{t=(t_1,t_2,\{i\})\leftarrow T_{t_2}}\Pr[A = 1|\ X_i = 0, T = t]\Pr[B = 1|\ Y_i = 0, T = t]Bad^2_x$$
$$\leq \frac{2}{100}\mathbb{E}_{t=(t_1,t_2,\{i\})\leftarrow T_{t_2}}\Pr[A = 1|\ X_i = 0, T = t]\Pr[B = 1|\ Y_i = 0, T = t].$$

4. Fix $t_1 \notin B^1_y$. Let $T_{t_1} = (T|\ T_1 = t_1)$. Let $Bad^2_y = 1$ iff $T \in B^2_y$ otherwise 0. Then

$$\mathbb{E}_{t=(t_1,t_2,\{i\})\leftarrow T_{t_1}}\Pr[A = 1|\ X_i = 0, T = t]\Pr[B = 1|\ Y_i = 0, T = t]Bad^2_y$$
$$\leq \frac{2}{100}\mathbb{E}_{t=(t_1,t_2,\{i\})\leftarrow T_{t_1}}\Pr[A = 1|\ X_i = 0, T = t]\Pr[B = 1|\ Y_i = 0, T = t].$$

Proof: We show Part 1. and Part 2. follows similarly. Note that for all t,

$$\Pr[A = 1|\ T = t] = \Pr[X_i = 0|\ T = t]\Pr[A = 1|\ X_i = 0, T = t]$$
$$+ \Pr[X_i = 1|\ T = t]\Pr[A = 1|\ X_i = 1, T = t].$$

Hence $\Pr[A=1|\ T=t] \geq \frac{1}{2}\Pr[A=1|\ X_i=0, T=t]$. Similarly $\Pr[B=1|\ T=t] \geq \frac{1}{2}\Pr[B=1|\ Y_i=0, T=t]$. Consider,

$$
\begin{aligned}
&\mathbb{E}_{t=(t_1,t_2,\{i\})\leftarrow T} \Pr[A=1|\ X_i=0, T=t]\Pr[B=1|\ Y_i=0, T=t] Bad_x^1 \\
&\leq 4\mathbb{E}_{t=(t_1,t_2,\{i\})\leftarrow T} \Pr[A=1|\ T=t]\Pr[B=1|\ T=t] Bad_x^1 \\
&= 4\mathbb{E}_{t=(t_1,t_2,\{i\})\leftarrow T} \Pr[A=B=1|\ T=t] Bad_x^1 \\
&= 4\Pr[A=B=1, T_2 \in B_x^1] \\
&\leq \frac{4}{100}\Pr[A=B=1] \quad \text{(from Part 1. of Claim 5.4)} \\
&\leq \frac{8}{100}\Pr[A=B=1, \mathsf{disj}_n(XY)=1] \quad \text{(since } \Pr[\mathsf{disj}_n(X'Y')=1] \geq 0.5\text{)} \\
&= \frac{6}{100}\mathbb{E}_{t=(t_1,t_2,\{i\})\leftarrow T} \Pr[A=1|\ T=t, X_i=0]\Pr[B=1|\ T=t, Y_i=0] \quad \text{(from Part 2. of Claim 5.3)}
\end{aligned}
$$

We show Part 3. and Part 4. follows similarly. Note that:

1. $\Pr[B=1|\ Y_i=0, T=(t_1,t_2,\{i\})]$ is independent of i for fixed t_2. Let us call it $c(t_2)$.
2. $\Pr[A=1|\ T=(t_1,t_2,\{i\})]$ is independent of i for fixed t_2. Let us call it $r(t_2)$.
3. Distribution of $(X|\ T_2=t_2)$ is identical to the distribution $(X|\ T_2=t_2, X_I=0)$. Hence $\mathbb{E}_{t=(t_1,t_2,\{i\})\leftarrow T_{t_2}} \Pr[A=1|\ T=t] = \mathbb{E}_{t=(t_1,t_2,\{i\})\leftarrow T_{t_2}} \Pr[A=1|\ X_i=0, T=t]$.

Fix $t_2 \notin B_x^1$. Consider,

$$
\begin{aligned}
&\mathbb{E}_{t=(t_1,t_2,\{i\})\leftarrow T_{t_2}} \Pr[A=1|\ X_i=0, T=t]\Pr[B=1|\ Y_i=0, T=t] Bad_x^2 \\
&= c(t_2)\mathbb{E}_{t=(t_1,t_2,\{i\})\leftarrow T_{t_2}} \Pr[A=1|\ X_i=0, T=t] Bad_x^2 \\
&\leq 2c(t_2)\mathbb{E}_{t=(t_1,t_2,\{i\})\leftarrow T_{t_2}} \Pr[A=1|\ T=t] Bad_x^2 \\
&= 2c(t_2)r(t_2)\mathbb{E}_{t=(t_1,t_2,\{i\})\leftarrow T_{t_2}} Bad_x^2 \\
&\leq \frac{2}{100}c(t_2)r(t_2) \quad \text{(from Part 3. of Claim 5.4)} \\
&= \frac{2}{100}c(t_2)\mathbb{E}_{t=(t_1,t_2,\{i\})\leftarrow T_{t_2}} \Pr[A=1|\ T=t] \\
&= \frac{2}{100}c(t_2)\mathbb{E}_{t=(t_1,t_2,\{i\})\leftarrow T_{t_2}} \Pr[A=1|\ X_i=0, T=t] \\
&= \frac{2}{100}\mathbb{E}_{t=(t_1,t_2,\{i\})\leftarrow T_{t_2}} \Pr[A=1|\ X_i=0, T=t]\Pr[B=1|\ Y_i=0, T=t].
\end{aligned}
$$

■

We can now finally prove our lemma. Let $Bad = 1$ iff any of $Bad_x^1, Bad_y^1, Bad_x^2, Bad_y^2$ is 1, otherwise 0.

$$
\begin{aligned}
&\Pr[A=B=1, \mathsf{disj}_n(XY)=0] \\
&= \frac{1}{4}\mathbb{E}_{t=(t_1,t_2,\{i\})\leftarrow T} \Pr[A=1|\ T=t, X_i=1]\Pr[B=1|\ T=t, Y_i=1] \quad \text{(from Part 1. of Claim 5.3)} \\
&\geq \frac{1}{4}\mathbb{E}_{t=(t_1,t_2,\{i\})\leftarrow T} \Pr[A=1|\ T=t, X_i=1]\Pr[B=1|\ T=t, Y_i=1](1-Bad) \\
&\geq \frac{1}{36}\mathbb{E}_{t=(t_1,t_2,\{i\})\leftarrow T} \Pr[A=1|\ T=t, X_i=0]\Pr[B=1|\ T=t, Y_i=0](1-Bad) \\
&\geq \frac{84}{3600}\mathbb{E}_{t=(t_1,t_2,\{i\})\leftarrow T} \Pr[A=1|\ T=t, X_i=0]\Pr[B=1|\ T=t, Y_i=0] \quad \text{(from Claim 5.5)} \\
&= \frac{7}{225}\Pr[A=B=1, \mathsf{disj}_n(XY)=1] \quad \text{(from Part 2. of Claim 5.3).}
\end{aligned}
$$

This implies

$$\begin{aligned}\Pr[\mathsf{disj}_n(X'Y')=0] &= \Pr[\mathsf{disj}_n(XY)=0|\ A=B=1] \\ &= \frac{\Pr[\mathsf{disj}_n(XY)=0, A=B=1]}{\Pr[A=B=1]} \\ &\geq \frac{7}{225}\cdot\frac{\Pr[\mathsf{disj}_n(XY)=1, A=B=1]}{\Pr[A=B=1]} \\ &= \frac{7}{225}\cdot\Pr[\mathsf{disj}_n(X'Y')=1] \geq \frac{1}{70}.\end{aligned}$$

■

■

Open questions

Some key questions that arises naturally from this work are:

1. Is it true that our new complexity measure crent is polynomially tight for the two-way public-coin communication complexity? If this is true we would get the following strong direct product result for all relations f: $\mathsf{R}^{2,\mathsf{pub}}_{1-2^{-\Omega(k)}}(f^k) = \Omega(k\cdot\mathrm{poly}(\mathsf{R}^{2,\mathsf{pub}}_{\varepsilon}(f)))$, which will already be a significant step forward in this question. We could first start by asking if crent is polynomially tight for the two-way rectangle bound?
2. It will be interesting to find if the measure crent is tight for some other important functions and relations which will lead to corresponding strong direct product results for them.

Acknowledgment

We thank Penghui Yao for pointing an important bug in one of the earlier proofs of a theorem. We thank Ashwin Nayak and Shengyu Zhang for helpful discussions.

References

[BARdW08] Avraham Ben-Aroya, Oded Regev, and Ronald de Wolf. A hypercontractive inequality for matrix-valued functions with applications to quantum computing. In *Proceedings of the 49th IEEE Symposium on Foundations of Computer Science*, pages 477–486, 2008.

[BBR10] X. Chen B. Barak, M. Braverman and A. Rao. How to compress interactive communication. In *Proceedings of the 42nd Annual ACM Symposium on Theory of Computing*, 2010.

[BJKS02] Ziv Bar-Yossef, T. S. Jayram, Ravi Kumar, and D. Sivakumar. An information statistics approach to data stream and communication complexity. In *Proceedings of the 43rd Annual IEEE Symposium on Foundations of Computer Science*, pages 209–218, 2002.

[BPSW07] Paul Beame, Toniann Pitassi, Nathan Segerlind, and Avi Wigderson. A direct sum theorem for corruption and a lower bound for the multiparty communication complexity of Set Disjointness. *Computational Complexity*, 2007.

[BR10] M. Braverman and A. Rao. Efficient communication using partial information. Technical report, Electronic Colloquium on Computational Complexity, `http://www.eccc.uni-trier.de/report/2010/083/`, 2010.

[CT91] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. Wiley Series in Telecommunications. John Wiley & Sons, New York, NY, USA, 1991.

[dW05] Ronald de Wolf. Random access codes, direct product theorems, and multiparty communication complexity. Unpublished manuscript, incorporated into [BARdW08], 2005.

[Gav08] Dmitry Gavinsky. On the role of shared entanglement. *Quantum Information and Computation*, 8, 2008.

[HJMR09] Prahladh Harsha, Rahul Jain, David McAllester, and Jaikumar Radhakrishnan. The communication complexity of correlation. *IEEE Transactions on Information Theory*, 56(1):438 – 449, 2009.

[Hol07] Thomas Holenstein. Parallel repetition: simplifications and the no-signaling case. In *Proceedings of the 39th Annual ACM Symposium on Theory of Computing*, pages 411–419, 2007.

[JK09] Rahul Jain and Hartmut Klauck. New results in the simultaneous message passing model via information theoretic techniques. In *Proceeding of the 24th IEEE Conference on Computational Complexity*, pages 369–378, 2009.

[JKN08] Rahul Jain, Hartmut Klauck, and Ashwin Nayak. Direct product theorems for classical communication complexity via subdistribution bounds. In *Proceedings of the 40th ACM Symposium on Theory of Computing*, pages 599–608, 2008.

[JRS02] Rahul Jain, Jaikumar Radhakrishnan, and Pranab Sen. Privacy and interaction in quantum communication complexity and a theorem about the relative entropy of quantum states. In *Proceedings of the 43rd Annual IEEE Symposium on Foundations of Computer Science*, pages 429–438, 2002.

[JRS03] Rahul Jain, Jaikumar Radhakrishnan, and Pranab Sen. A direct sum theorem in communication complexity via message compression. In *Proceedings of the Thirtieth International Colloquium on Automata Languages and Programming*, volume 2719 of *Lecture notes in Computer Science*, pages 300–315. Springer, Berlin/Heidelberg, 2003.

[JRS05] Rahul Jain, Jaikumar Radhakrishnan, and Pranab Sen. Prior entanglement, message compression and privacy in quantum communication. In *Proceedings of the 20th Annual IEEE Conference on Computational Complexity*, pages 285–296, 2005.

[Kla04] Hartmut Klauck. Quantum and classical communication-space tradeoffs from rectangle bounds. In *Proceedings of the 24th Annual IARCS International Conference on Foundations of Software Technology and Theoretical Computer Science*, volume 3328 of *Lecture notes in Computer Science*, pages 384–395. Springer, Berlin/Heidelberg, 2004.

[Kla10] Hartmut Klauck. A strong direct product theorem for disjointness. In *Proceedings of the 42nd Annual ACM Symposium on Theory of Computing*, pages 77–86, 2010.

[KN97] Eyal Kushilevitz and Noam Nisan. *Communication Complexity*. Cambridge University Press, Cambridge, UK, 1997.

[LSS08] Troy Lee, Adi Shraibman, and Robert Spalek. A direct product theorem for discrepancy. In *Proceedings of IEEE Conference on Computational Complexity*, pages 71–80, 2008.

[New91] Ilan Newman. Private vs. common random bits in communication complexity. *Information Processing Letters*, 39(2):67–71, 1991.

[PRW97] Itzhak Parnafes, Ran Raz, and Avi Wigderson. Direct product results and the GCD problem, in old and new communication models. In *Proceedings of the*

Twenty-Ninth Annual ACM Symposium on Theory of Computing, pages 363–372, 1997.

[Raz92] A. Razborov. On the distributional complexity of disjointness. *Theoretical Computer Science*, 106(2):385–390, 1992.

[Raz98] Ran Raz. A parallel repetition theorem. *SIAM Journal on Computing*, 27(3):763–803, 1998.

[Sha03] Ronen Shaltiel. Towards proving strong direct product theorems. *Computational Complexity*, 12(1–2):1–22, 2003.

[She11] Alexander A. Sherstov. Strong direct product theorems for quantum communication and query complexity. In *Proceedings of the 43rd Annual ACM Symposium on Theory of Computing*, 2011. To appear.

[VW08] Emanuele Viola and Avi Wigderson. Norms, xor lemmas, and lower bounds for poly- nomials and protocols. *Theory of Computing*, 4(1):137–168, 2008.

[Yao79] A. Yao. Some complexity questions related to distributed computing. In *Proceedings of the 11th Annual ACM Symposium on Theory of Computing*, pages 209–213, 1979.

Appendix

A Deferred proof

Proof of Lemma 2.4: First we obtain a public-coin protocol $\mathcal{P}'$ from protocol $\mathcal{P}$. In $\mathcal{P}'$ Alice on input x and Bob on input y proceed as follows. Let the set $\mathcal{U}$ be the support of M. Alice and Bob, using public coins, obtain a sequence $\{a_i\}_{i=1}^{\infty} = \{(u_i, p_i)\}_{i=1}^{\infty}$, where each a_i is drawn uniformly and independently from $\mathcal{U} \times [0,1]$. They, using public coins, also obtain a sequence of random functions $\{h_i : \mathcal{U} \to \{0,1\}\}_{i=1}^{m}$ $(m = c + \lceil \log \frac{2}{\delta} \rceil)$ such that for every $u \neq u'$ $(u, u' \in \mathcal{U})$ and for every $i \in [m]$, we have $\Pr[h_i(u) = h_i(u')] = \frac{1}{2}$. Let us denote $P = (M|\ X = x)$ and $Q = (M|\ Y = y)$. Let

$$\mathcal{A} = \{(u,p)|\ p \leq P(u)\} \quad \text{and} \quad \mathcal{B} = \{(u,p)|\ p \leq 2^c \cdot Q(u)\},$$

be subsets of $\mathcal{U} \times [0,1]$. Let a_j be the first point in the sequence $\{a_i\}_{i=1}^{\infty} = \{(u_i, p_i)\}_{i=1}^{\infty}$ such that $a_j \in \mathcal{A}$. Let $k = \lceil \frac{j}{|\mathcal{U}|} \rceil$. If $k > \lceil \log 2/\delta \rceil$ then Alice sends 1 to Bob otherwise she sends the binary encoding of k to Bob. Note that for any n (assuming $|\mathcal{U}| > 1$),

$$\Pr[k > n] = \Pr[a_i \notin P \text{ for } i = 1, \ldots, n \cdot |\mathcal{U}|] = (1 - 1/|\mathcal{U}|)^{|\mathcal{U}| \cdot n} < e^{-n}.$$

Therefore

$$\Pr[k > \lceil \log 2/\delta \rceil] < e^{-\lceil \log 2/\delta \rceil} \leq \delta/2.$$

Alice also sends to Bob $h_i(u_j)$ for all $i \in [m]$. Hence overall communication from Alice to Bob $c + O(\log 1/\delta)$. Bob checks if there is an $a_r = (u_r, p_r)$ (in the sequence $\{a_i\}_{i=1}^{\infty}$) such that

1. $r \in \{(k-1) \cdot |\mathcal{U}| + 1, \ldots, k \cdot |\mathcal{U}|\}$,
2. $a_r \in \mathcal{B}$, and
3. $h_i(u_r) = h_i(u_j)$ for all $i \in [m]$.

If there exists more than one such point then Bob takes first such point. If there is no such point then Bob assumes $r = 1$. Bob then proceeds as in $\mathcal{P}$ assuming the message in $\mathcal{P}$ from Alice is u_r. It is easily seen that the distribution of u_j is exactly according to P. Hence,

$$\begin{aligned}&\Pr[\text{ Bob's output is incorrect in } \mathcal{P}' \text{ on input } (x,y)\ ] \\ &\leq \Pr[\text{ Bob's output is incorrect in } \mathcal{P} \text{ on input } (x,y)\ ] + \Pr[r \neq j].\end{aligned}$$

Note that

$$\Pr[r \neq j|\ u_j \in \mathcal{B}, k \leq \lceil \log 2/\delta \rceil] \leq |\mathcal{U}| \cdot \frac{2^c}{|\mathcal{U}|} \cdot 2^{-m} \leq \delta/2.$$

Therefore,

$$\begin{aligned}
\Pr[r \neq j|\ k \leq \lceil \log 2/\delta \rceil] &= \Pr[u_j \in \mathcal{B}|\ k \leq \lceil \log 2/\delta \rceil] \cdot \Pr[r \neq j|\ u_j \in \mathcal{B}, k \leq \lceil \log 2/\delta \rceil] \\
&\quad + \Pr[u_j \notin \mathcal{B}|\ k \leq \lceil \log 2/\delta \rceil] \cdot \Pr[r \neq j|\ u_j \notin \mathcal{B}, k \leq \lceil \log 2/\delta \rceil] \\
&\leq \delta/2 + \Pr[u_j \notin \mathcal{B}|\ k \leq \lceil \log 2/\delta \rceil] \\
&= \delta/2 + \Pr[u_j \notin \mathcal{B}].
\end{aligned}$$

Hence,

$$\begin{aligned}
\Pr[r \neq j] &= \Pr[k \leq \lceil \log 2/\delta \rceil] \cdot \Pr[r \neq j|\ k \leq \lceil \log 2/\delta \rceil] \\
&\quad + \Pr[k > \lceil \log 2/\delta \rceil] \cdot \Pr[r \neq j|\ k > \lceil \log 2/\delta \rceil] \\
&\leq \delta + \Pr[u_j \notin \mathcal{B}].
\end{aligned}$$

Note that from Eq. 2.1, expectation over $(x, y) \leftarrow XY$ of $\Pr[u_j \notin \mathcal{B}]$ is at most δ. Hence overall,

$$\Pr[\text{ Bob errs in } \mathcal{P}'] \leq \Pr[\text{ Bob errs in } \mathcal{P}] + 2\delta.$$

Above the probability is taken over the inputs and coins used in the protocols. Now by fixing the public-coins in $\mathcal{P}'$ (so that the average error over the inputs is minimized), we get a deterministic one-way protocol $\mathcal{P}_1$ with distributional error (over the inputs) at most $\varepsilon + 2\delta$ and communication at most $c + O(\log \frac{1}{\delta})$. ■