Lower bounds on information complexity via zero-communication protocols and applications

Iordanis Kerenidis*, Sophie Laplante†, Virginie Lerays‡, Jérémie Roland§ and David Xiao¶
*CNRS, LIAFA, Université Paris 7 and CQT, NUS Singapore. Email: jkeren@liafa.univ-paris-diderot.fr
†LRI, Université Paris-Sud 11. Email: laplante@lri.fr
‡LRI, Université Paris-Sud 11. Email: virginie.lerays@lri.fr
§Université Libre de Bruxelles, QuIC, Ecole Polytechnique de Bruxelles. Email: jroland@ulb.ac.be
¶CNRS, LIAFA, Université Paris 7. Email: dxiao@liafa.univ-paris-diderot.fr

***Abstract*—We show that almost all known lower bound methods for communication complexity are also lower bounds for the information complexity. In particular, we define a relaxed version of the *partition bound* of Jain and Klauck [1] and prove that it lower bounds the information complexity of any function. Our relaxed partition bound subsumes all norm based methods (e.g. the γ_2 method) and rectangle-based methods (e.g. the rectangle/corruption bound, the smooth rectangle bound, and the discrepancy bound), except the partition bound.**

Our result uses a new connection between rectangles and *zero-communication* protocols where the players can either output a value or abort. We prove the following compression lemma: given a protocol for a function f with information complexity I, one can construct a zero-communication protocol that has non-abort probability at least $2^{-O(I)}$ and that computes f correctly with high probability conditioned on not aborting. Then, we show how such a zero-communication protocol relates to the relaxed partition bound.

We use our main theorem to resolve three of the open questions raised by Braverman [2]. First, we show that the information complexity of the Vector in Subspace Problem [3] is $\Omega(n^{1/3})$, which, in turn, implies that there exists an exponential separation between quantum communication complexity and classical information complexity. Moreover, we provide an $\Omega(n)$ lower bound on the information complexity of the Gap Hamming Distance Problem.

***Keywords*-communication complexity, information complexity, information theory**

I. Introduction

Information complexity is a way of measuring the amount of information Alice and Bob must reveal to each other in order to solve a distributed problem. The importance of this notion has been made apparent in recent years through a flurry of results that relate the information complexity of a function and its communication complexity. One of the main applications of information complexity is to prove direct sum theorems in communication complexity, namely to show that computing k copies of a function costs k times the communication of computing a single copy. Chakrabarti, Shi, Wirth and Yao [4] used information complexity to prove a direct sum theorem for simultaneous messages protocols (their notion is now usually called the *external* information complexity, whereas in this paper we work exclusively with what is often called the *internal* information complexity). Bar-Yossef et al. [5], used the information cost in order to prove a linear lower bound on the two-way randomized communication complexity of Disjointness. More recently, information-theoretic techniques enabled the proof of the first non-trivial direct sum result for general two-way randomized communication complexity: the randomized communication complexity of k copies of a function f is at least $\sqrt{k}$ times the randomized communication complexity of f [6]. Then, Braverman and Rao [7], showed a tight relation between the amortized distributional communication complexity of a function and its internal information cost. Braverman [2], defined interactive information complexity, a notion which is independent of the prior distribution of the inputs and proved that it is equal to the amortized communication complexity of the function. Braverman and Weinstein [8] showed that the information complexity is lower bounded by discrepancy.

The main question pertaining to information complexity is its relation to communication complexity. On the one hand, the information complexity provides a lower bound on the communication complexity of the function, since there cannot be more information leaked than the length of the messages exchanged. However, it is still open, whether the information complexity of a function can be much smaller than its communication complexity or whether the two notions are basically equivalent. In order to make progress towards this question, it is imperative to provide strong lower bounds

for information complexity, and more specifically to see whether the lower bound methods for communication complexity can be compared to the model of information complexity.

Lower bound methods in communication complexity can be seen to fall into three main categories: the norm based methods, such as the γ_2 method of Linial and Shraibman [9] (see Lee and Shraibman's survey for an overview [10]); the rectangle based methods, such as discrepancy and the rectangle bound; and, of course, the information theoretic methods, among which, information complexity. Recently, Jain and Klauck [1] introduced the smooth rectangle bound, as well as the stronger partition bound, and showed that they subsume both γ_2 and the rectangle bound [1].

The first lower bound on information complexity was proved by Braverman [2], who showed that it is lower bounded by the logarithm of the communication complexity. Recently, Braverman and Weinstein showed that the discrepancy method lower bounds the information complexity [8]. Their result follows from a compression lemma for protocols: a protocol for a function f that leaks I bits of information implies the existence of a protocol with communication complexity $O(I)$ and advantage on computing f (over a random guess) of $2^{-O(I)}$.

A. Our results

In this paper, we show that all known lower bound methods for communication complexity, with the notable exception of the partition bound, generalize to information complexity. More precisely, we introduce the *relaxed partition bound* (in Definition III.2) denoted by $\bar{\mathsf{prt}}^{\mu}_{\epsilon}(f)$, which depends on the function to be computed f, the input distribution μ, and the error parameter ϵ, and such that the distributional communication complexity $D^{\mu}_{\epsilon}(f) \geq \log(\bar{\mathsf{prt}}^{\mu}_{\epsilon}(f))$ for any f. We prove that the information complexity of a function f is bounded below by the relaxed partition bound:

Theorem I.1. *There is a positive constant C such that for all functions $f : \mathcal{I} \to \mathcal{Z}$, all $\epsilon, \delta \in (0, \frac{1}{2}]$, and all distributions μ, it holds that:*

$$\mathsf{IC}_{\mu}(f, \epsilon) \geq \tfrac{\delta^2}{C} \cdot \left(\log \bar{\mathsf{prt}}^{\mu}_{\epsilon+3\delta}(f) - \log|\mathcal{Z}|\right) - \delta$$

Since we show in Lemma III.3 that the relaxed partition bound subsumes the norm based methods (e.g. the γ_2 method) and the rectangle-based methods (e.g. the rectangle/corruption bound, the smooth rectangle bound, and the discrepancy bound), all of these bounds are also lower bounds on the information complexity. Moreover, together with the direct sum theorem for information complexity, our main result implies a direct sum theorem on communication complexity for many notable functions (see Corollary I.4).

Technique: The key idea of our result is a new connection between communication rectangles and zero-communication protocols, where the players can either output a value or abort but *without communicating*. A priori, it is surprising that protocols with no communication can actually provide some insight on the communication or information complexity of a function. However, this model, which has been extensively used in quantum information for the study of non-local games and Bell inequalities, turns out to be a very powerful tool for the study of classical communication and information complexity. The communication complexity of simulating distributions is known to be related to the probability of not aborting in zero-communication protocols that can abort [11, 12, 13, 14]. More recently connections have been shown for specific lower bound methods. It has been shown that zero-communication protocols with error give rise to the factorization norm method [15], and the connection between the partition bound and zero-communication protocols with abort was studied in [16].

In a deterministic zero-communication protocol with abort, each of the two players looks at their input and decides either to abort the protocol or to output some value z. The output of the protocol is z if both players agree on z, or it aborts otherwise. It is easy to see that for any deterministic zero-communication protocol with abort, the set of inputs where both players choose to output z forms a rectangle, and so the protocol is characterized by a set of rectangles each labeled by an output. In a randomized protocol, we have instead a distribution over labeled rectangles.

This connection between rectangles and zero-communication protocols with abort allows us to obtain our lower bound for information complexity from a new compression lemma for protocols (Lemma III.4): a protocol for a function f that leaks I bits of information implies the existence of a *zero*-communication protocol that has non-abort probability at least $2^{-O(I)}$ and that computes f correctly with high probability when not aborting. Our main theorem follows from this new compression.

The technical tools we use are drawn from Braverman [2] and in particular Braverman and Weinstein [8]. We describe the difference between our compression and that of [8]. There, they take a protocol for computing a function f that has information cost I and compress it to a protocol with communication $O(I)$ and advantage

of computing f of $2^{-O(I)}$ (*i.e.* the error increases considerably). Then, they apply the discrepancy method, which can handle such small advantage.

In our compression, we suppress the communication entirely, and, moreover, we only introduce an arbitrarily small error since the compressed protocol aborts when it does not believe it can correctly compute the output. This compression enables us to provide much sharper lower bounds on the information complexity and in particular, the lower bound in terms of the relaxed partition bound.

Applications: Our lower bound implies that for most functions for which there exists a lower bound on their communication complexity, the same bound extends to their information complexity. Specifically, we can apply our lower bound in order to resolve three of the open questions in [2].

First, we show that there exists a function f, such that the quantum communication complexity of f is exponentially smaller than the information complexity of f (Open Problem 3 in [2]).

Theorem I.2. *There exists a function f, s.t. for all $\epsilon \in (0, \frac{1}{2}), Q(f, \epsilon) = O(\log(\mathsf{IC}(f, \epsilon)))$.*

In order to prove the above separation, we show that the proof of the lower bound on the randomized communication complexity of the Vector in Subspace Problem ($\widetilde{\mathrm{VSP}}$) [3] provides, in fact, a lower bound on the relaxed partition bound. By our lower bound, this implies that $\mathsf{IC}(\widetilde{\mathrm{VSP}}_{\theta,n}, 1/3) = n^{1/3}$ (Open Problem 7 in [2]). Since the quantum communication complexity of $\widetilde{\mathrm{VSP}}_{\theta,n}$ is $O(\log n)$, we have the above theorem. Moreover, this implies an exponential separation between classical and quantum information complexity. We refrain from defining quantum information cost in this paper (see [17] for a definition), but since the quantum information cost is always smaller than the quantum communication complexity, the separation follows trivially from the above theorem.

In addition, we resolve the question of the information complexity of the Gap Hamming Distance Problem (GHD) (Open Problem 6 in [2]), since the lower bounds on the randomized communication complexity of this problem go through the rectangle/corruption bound [18] or smooth rectangle bound [19, 20].

Theorem I.3. $\mathsf{IC}(\mathrm{GHD}_n, 1/3) = \Omega(n)$.

Regarding direct sum theorems, it was shown [2] that the information complexity satisfies a direct sum theorem, namely $\mathsf{IC}_{\mu^k}(f^k, \epsilon) \geq k \cdot \mathsf{IC}_\mu(f, \epsilon)$. If in addition it holds that $D^\mu_{\epsilon'}(f) = O(\mathsf{IC}_\mu(f, \epsilon))$, then we can immediately deduce that $D^{\mu^k}_\epsilon(f) \geq \mathsf{IC}_{\mu^k}(f^k, \epsilon) \geq k \cdot \mathsf{IC}_\mu(f, \epsilon) \geq \Omega(k \cdot D^\mu_{\epsilon'}(f))$, *i.e.* the direct sum theorem holds for f. Therefore our main result also gives the following corollary:

Corollary I.4. *For any ϵ, μ and any $f : \mathcal{I} \to \mathcal{Z}$, if $D^\mu_\epsilon(f) = O(\log \bar{\mathsf{prt}}^\mu_\epsilon(f))$, then for all $\delta > 0$ and integers k, it holds that $D^{\mu^k}_\epsilon(f) \geq \Omega\left(k \cdot \delta^2 (D^\mu_{\epsilon+3\delta}(f) - \log|\mathcal{Z}|) - k\delta\right)$.*

For example, since $D^\mu_\epsilon(\mathrm{GHD}) \leq n$ holds trivially, this corollary along with the fact that $\log \bar{\mathsf{prt}}^\mu_\epsilon(\mathrm{GHD}) = \Omega(n)$ ([18, 19, 20], see Section V-B) immediately implies a direct sum theorem for GHD.

Finally, regarding the central open question of whether or not it is possible to compress communication down to the information complexity for any function, we note that our result says that if one hopes to prove a negative result and separate information complexity from communication complexity, then one must use a lower bound technique that is stronger than the relaxed partition bound. To the best of our knowledge, the only such technique in the literature is the (standard) partition bound. We note, however, that to the best of our knowledge there are no known problems whose communication complexity can be lower-bounded by the partition bound but not by the relaxed partition bound.

Many proofs are omitted from this extended abstract, and can be found in the full version of the paper.

B. Related work

Definitions of information complexity with some variations extend back to the work on privacy in interactive protocols [21], and related definitions in the privacy literature appear [22, 23, 24]. Information complexity as a tool in communication complexity was first used to prove direct sum theorems in the simultaneous message model [4], and subsequently to prove direct sum theorems and to study amortized communication complexity as stated in the first paragraph of this paper [5, 6, 7, 2, 8]. There are many other works using information complexity to prove lower bounds for specific functions or to prove direct sum theorems in restricted models of communication complexity, for example [25, 26, 27, 28].

In independent and concurrent work, Chakrabarti et al. proved that information complexity is lower bounded by the smooth rectangle bound under product distributions [29]. While our result implies the result of [29] as a special case, we note that their proof uses entirely different techniques and may be of independent interest.

II. Preliminaries

A. Notation and information theory facts

Let μ be a probability distribution over a (finite) universe $\mathcal{U}$. We will often treat μ as a function $\mu : 2^{\mathcal{U}} \to [0,1]$. For $T, S \subseteq \mathcal{U}$, we let $\mu(T \mid S) = \Pr_{u \leftarrow \mu}[u \in T \mid S]$. For singletons $u \in U$, we write interchangeably $\mu_u = \mu(u) = \mu(\{u\})$. Random variables are written in uppercase and fixed values in lowercase. We sometimes abuse notation and write a random variable in place of the distribution of that random variable.

For two distributions μ, ν, we let $|\mu - \nu|$ denote their statistical distance, *i.e.* $|\mu - \nu| = \max_{T \subseteq \mathcal{U}}(\mu(T) - \nu(T))$. We let $D(\mu \parallel \nu) = \mathbb{E}_{U \sim \mu}[\log \frac{\mu(U)}{\nu(U)}]$ be the relative entropy (*i.e.* KL-divergence). For two random variables X, Y, the mutual information is defined as $I(X:Y) = H(X) - H(X \mid Y) = H(Y) - H(Y \mid X)$, where $H(\cdot)$ is the Shannon entropy.

A *rectangle* of $\mathcal{X} \times \mathcal{Y}$ is a product set $A \times B$ where $A \subseteq \mathcal{X}$ and $B \subseteq \mathcal{Y}$. We let R denote a rectangle in $\mathcal{X} \times \mathcal{Y}$. We let $(x, y) \in \mathcal{X} \times \mathcal{Y}$ denote a fixed input, and (X, Y) be random inputs sampled according to some distribution (specified from context and usually denoted by μ).

B. Information complexity

We study 2-player communication protocols for calculating a function $f : \mathcal{I} \to \mathcal{Z}$, where $\mathcal{I} \subseteq \mathcal{X} \times \mathcal{Y}$. Let π be a randomized protocol (allowing both public and private coins, unless otherwise specified). We denote the randomness used by the protocol π by r_π. Let $\pi(x, y)$ denote its output, *i.e.* the value in $\mathcal{Z}$ the two parties wish to compute.

The transcript of a protocol includes all messages exchanged, the output of the protocol (in fact we just need that both players can compute the output of the protocol from the transcript), as well as any public coins (but no private coins). The complexity of π is the maximum (over all random coins) of the number of bits exchanged.

Let μ be a distribution over $\mathcal{X} \times \mathcal{Y}$. Define $\mathsf{err}_f(\pi; x, y) = \Pr_{r_\pi}[f(x, y) \neq \pi(x, y)]$ if $(x, y) \in \mathcal{I}$ and 0 otherwise and $\mathsf{err}_f(\pi; \mu) = \mathbb{E}_{(X,Y) \sim \mu} \mathsf{err}_f(\pi; X, Y) = \Pr_{r_\pi, (X,Y) \sim \mu}[(X, Y) \in \mathcal{I} \wedge f(X, Y) \neq \pi(X, Y)]$.

Definition II.1. Fix f, μ, ϵ. Let (X, Y, Π) be the tuple distributed according to (X, Y) sampled from μ and then Π being the transcript of the protocol π applied to X, Y. Then define:

1) $\mathsf{IC}_\mu(\pi) = I(X; \Pi \mid Y) + I(Y; \Pi \mid X)$
2) $\mathsf{IC}_\mu(f, \epsilon) = \inf_{\pi : \mathsf{err}_f(\pi; \mu) \leq \epsilon} \mathsf{IC}_\mu(\pi)$
3) $\mathsf{IC}_D(f, \epsilon) = \max_\mu \mathsf{IC}_\mu(f, \epsilon)$

Braverman [2] also defined the non-distributional information cost IC, and all of our results extend to it trivially by the inequality $\mathsf{IC}_D \leq \mathsf{IC}$. (We do not require the reverse inequality $\mathsf{IC} \leq O(\mathsf{IC}_D)$, whose proof is non-trivial and was given in [2]).

III. Zero-communication protocols and the relaxed partition bound

A. The zero-communication model and rectangles

Let us consider a (possibly partial) function f. We say that (x, y) is a valid input if $(x, y) \in \mathcal{I}$, that is, (x, y) satisfies the promise. In the zero-communication model with abort, the players either output a value $z \in \mathcal{Z}$ (they *accept* the run) or output $\perp$ (they *abort*).

Definition III.1. The *zero-communication* model with abort is defined as follows:

- **Inputs.** Alice and Bob receive inputs x and y respectively.
- **Output.** Alice outputs $a \in \mathcal{Z} \cup \{\perp\}$ and Bob outputs $b \in \mathcal{Z} \cup \{\perp\}$. If both Alice and Bob output the same $z \in \mathcal{Z}$, then the output is z. Otherwise, the output is $\perp$.

We will study (public-coin) randomized *zero-communication* protocols for computing functions in this model.

B. Relaxed partition bound

The relaxed partition bound with error ϵ and input distribution μ, denoted by $\bar{\mathsf{prt}}^\mu_\epsilon(f)$, is defined as follows.

Definition III.2. The distributional relaxed partition bound $\bar{\mathsf{prt}}^\mu_\epsilon(f)$ is the value of the following linear program. (The value of z ranges over $\mathcal{Z}$ and R over all rectangles, including the empty rectangle.)

$$\bar{\mathsf{prt}}^\mu_\epsilon(f) = \min_{\eta, p_{R,z} \geq 0} \frac{1}{\eta} \quad \text{subject to:}$$

$$\sum_{(x,y) \in \mathcal{I}} \mu_{x,y} \sum_{R : (x,y) \in R} p_{R, f(x,y)} + \sum_{(x,y) \notin \mathcal{I}} \mu_{x,y} \sum_{z, R : (x,y) \in R} p_{R,z} \geq (1 - \epsilon)\eta \quad (1)$$

$$\forall (x, y) \in \mathcal{X} \times \mathcal{Y}, \quad \sum_{z, R : (x,y) \in R} p_{R,z} \leq \eta \quad (2)$$

$$\sum_{R,z} p_{R,z} = 1. \quad (3)$$

The relaxed partition bound is defined as $\bar{\mathsf{prt}}_\epsilon(f) = \max_\mu \bar{\mathsf{prt}}^\mu_\epsilon(f)$.

We can identify feasible solutions to the program in Definition III.2 as a particular type of randomized zero-communication protocol: Alice and Bob sample (R, z) according to the distribution given by the $p_{R,z}$, and each individually sees if their inputs are in R and if so they output z, otherwise they abort. The parameter η is the *efficiency* of the protocol [16], that is, the probability that the protocol does not abort, and ideally we want it to be as large as possible.

There is also a natural way to convert any zero-communication protocol π into a distribution over (R, z): sample z uniformly from $\mathcal{Z}$, sample random coins r_π for π, and let $R = A \times B$ be such that A is the set of inputs on which Alice outputs z in the protocol π using random coins r_π, and similarly for B. (The sampling of a random z incurs a loss of $|\mathcal{Z}|$ in the efficiency, which is why our bounds have a loss depending on $|\mathcal{Z}|$. See Section III-D for details.)

Relation to other bounds: The relaxed partition bound is, as its name implies, a relaxation of the partition bound $\mathsf{prt}_\epsilon(f)$ [1]. It can also be shown that the relaxed partition bound is stronger than the smooth rectangle bound $\mathsf{srec}^z_\epsilon(f)$.

Lemma III.3. *For all f, ϵ and $z \in \mathcal{Z}$, we have $\mathsf{srec}^z_\epsilon(f) \le \overline{\mathsf{prt}}_\epsilon(f) \le \mathsf{prt}_\epsilon(f)$.*

Since Jain and Klauck have shown in [1] that the smooth rectangle bound is stronger than the rectangle/corruption bound, the γ_2 method and the discrepancy method, this implies that the relaxed partition bound subsumes all these bounds as well. Therefore, our result implies that all these bounds are also lower bounds for information complexity.

We briefly explain the difference between the relaxed partition bound and the partition bound. The partition bound includes two types of constraints. The first is a correctness constraint: on every input, the output of the protocol should be correct with probability at least $(1 - \epsilon)\eta$. The second is a completeness constraint: on every input, the efficiency of the protocol (*i.e.* the probability it does not abort) should be *exactly* η. In the relaxed partition bound, we keep the same correctness constraint. Since in certain applications the function is partial (such as the Vector in Subspace Problem [3]), one also has to handle the inputs where the function is not defined. We make this explicit in our correctness constraint. On the other hand, we relax the completeness constraint so that the efficiency may lie anywhere between $(1 - \epsilon)\eta$ and η. This relaxation seems to be crucial for our proof of the lower bound on information complexity, since we are unable to achieve efficiency exactly η.

C. Compression lemma

Lemma III.4 (Main compression lemma)**.** *There exists a universal constant C such that for all distributions μ, communication protocols π and $\delta \in (0, 1)$, there exists a zero-communication protocol π' and a real number $\lambda \ge 2^{-C(\mathsf{IC}_\mu(\pi)/\delta^2 + 1/\delta)}$ such that*

$$|(X, Y, \pi(X, Y)) - (X, Y, \pi'(X, Y) | \pi'(X, Y) \neq \perp)| \le \delta \quad (4)$$

(in statistical distance) and

$$\forall (x, y) \quad \Pr_{r_{\pi'}}[\pi'(x, y) \neq \perp] \le (1 + \delta)\lambda \quad (5)$$

$$\Pr_{r_{\pi'}, (X,Y)\sim\mu}[\pi'(X, Y) \neq \perp] \ge (1 - \delta)\lambda. \quad (6)$$

Our compression π' extends the strategy outlined by [8]. At a high level, the protocol π' does the following:

- **Sample transcripts.** Alice and Bob use their shared randomness to repeat T independent executions of an experiment to sample transcripts (Protocol IV.1). Alice and Bob each decide whether the experiment is accepted (they may differ in their opinions).
- **Find common transcript.** Let $\mathcal{A}$ be the set of accepted experiments for Alice, and $\mathcal{B}$ the set of accepted experiments for Bob. They try to guess an element of $\mathcal{A} \cap \mathcal{B}$. If they find one, they output according to the transcript from this experiment.

We prove our compression lemma in Section IV.

D. Information cost is lower bounded by the relaxed partition bound

We show how our compression lemma implies the main theorem.

Proof of Theorem I.1: Let π be a randomized communication protocol achieving $\mathsf{IC}_\mu(f, \epsilon)$ and let $\mathcal{R}$ be the following relation that naturally arises from the function f

$$\mathcal{R} = \{(x, y, f(x, y)) : (x, y) \in \mathcal{I}\} \cup \{(x, y, z) : (x, y) \notin \mathcal{I}, z \in \mathcal{Z}\}.$$

Let us now consider the zero-communication protocol π' from Lemma III.4. As mentioned in Section III-B, there is a natural way to identify π' with a distribution over labeled rectangles (R, z): sample z uniformly from $\mathcal{Z}$, sample r_π and let $R = A \times B$ where A is the set of inputs on which Alice outputs z, and similarly for B. The sampling of z incurs a loss of $|\mathcal{Z}|$ in the efficiency.

We make this formal: for any fixed randomness r occurring with probability p_r, we define the rectangle

$R(z,r)$ as the set of (x,y) such that the protocol outputs z, and we let $p_{R,z} = \sum_{r:R=R(z,r)} p_r/|\mathcal{Z}|$.

We check the normalization constraint

$$\sum_{R,z} p_{R,z} = \frac{1}{|\mathcal{Z}|} \sum_{R,z} \sum_{r:R=R(z,r)} p_r$$
$$= \frac{1}{|\mathcal{Z}|} \sum_{r} p_r \sum_{R,z:R=R(z,r)} 1 = \sum_r p_r = 1.$$

To see that Equation 2 is satisfied, we have by definition of $p_{R,z}$ that for any (x,y):

$$\sum_{z,R:(x,y)\in R} p_{R,z} = \frac{1}{|\mathcal{Z}|} \Pr_{r_{\pi'}}[\pi'(x,y) \neq \perp] \leq \frac{(1+\delta)\lambda}{|\mathcal{Z}|}.$$

Finally, to see that Equation 1 is satisfied, we have

$$\begin{aligned}
&\sum_{(x,y)\in\mathcal{I}} \mu_{x,y} \sum_{R:(x,y)\in R} p_{R,f(x,y)} \\
&\quad + \sum_{(x,y)\notin\mathcal{I}} \mu_{x,y} \sum_{z,R:(x,y)\in R} p_{R,z} \\
&= \frac{1}{|\mathcal{Z}|} \Pr_{r_{\pi'},(X,Y)\sim\mu}[(X,Y,\pi'(X,Y)) \in \mathcal{R}] \\
&= \frac{1}{|\mathcal{Z}|} \Pr_{r_{\pi'},(X,Y)\sim\mu}[\pi'(X,Y) \neq \perp] \\
&\quad \cdot \Pr_{r_{\pi'},(X,Y)\sim\mu}[(X,Y,\pi'(X,Y)) \in \mathcal{R} \mid \pi'(X,Y) \neq \perp] \\
&\geq \frac{1}{|\mathcal{Z}|}(1-\delta)\,\lambda \\
&\quad \cdot \left(\Pr_{r_{\pi'},(X,Y)\sim\mu}[(X,Y,\pi(X,Y)) \in \mathcal{R}] - \delta \right) \\
&\geq \frac{1}{|\mathcal{Z}|}(1-\delta)\,\lambda\,(1-\epsilon-\delta) \\
&\geq \frac{1}{|\mathcal{Z}|}\,\lambda\,(1-\epsilon-2\delta) \quad \geq \quad \frac{\lambda(1+\delta)}{|\mathcal{Z}|}(1-\epsilon-3\delta)
\end{aligned}$$

where for the last line we used the fact that π has error ϵ, and so $\Pr_{r_\pi,(X,Y)\sim\mu}[(X,Y,\pi(X,Y)) \in \mathcal{R}] \geq 1-\epsilon$. This satisfies the constraints in the linear program (Definition III.2) for $\bar{\mathsf{prt}}^{\mu}_{\epsilon+3\delta}(f)$ with objective value $\eta = (1+\delta)\lambda/|\mathcal{Z}| \geq 2^{-C(\mathsf{IC}_\mu(\pi)/\delta^2+1/\delta)}/|\mathcal{Z}|$. ∎

By the definitions of the information complexity and the relaxed partition bound, we have immediately

Corollary III.5. *There exists a universal constant C such that for all functions $f : \mathcal{I} \to \mathcal{Z}$, all $\epsilon,\delta \in (0,1/2)$, we have $\mathsf{IC}_D(f,\epsilon) \geq \frac{\delta^2}{C}[\log \bar{\mathsf{prt}}_{\epsilon+3\delta}(f) - \log|\mathcal{Z}|] - \delta$.*

IV. The zero-communication protocol

The zero-communication protocol consists of two stages. First, Alice and Bob use their shared randomness to come up with candidate transcripts, based on the a priori information they have on the distribution of the transcripts given by the information cost of the protocol. To do this, they run some sampling experiments and decide which ones to accept. Second, they use their shared randomness in order to choose an experiment that they have both accepted. If anything fails in the course of the protocol, they abort by outputting $\perp$.

A. Single sampling experiment

The single sampling experiment is described in Protocol IV.1 and appeared first in [8] (variants also appeared in [2] and [7]). Roughly, Protocol IV.1 takes a distribution τ and two distributions $\nu_{\mathsf{A}}, \nu_{\mathsf{B}}$ over a universe $\mathcal{U}$ such that $\nu_{\mathsf{A}}, \nu_{\mathsf{B}}$ are not too far from τ and tries to sample an element of $\mathcal{U}$ that is close to being distributed according to τ.

Let us informally describe the goal of this sampling experiment in our context. Alice knowing x and Bob y want to sample transcripts according to $\Pi_{x,y}$ which is the distribution over the transcripts of the protocol π applied to (x,y). When inputs x,y are fixed, the probability of a transcript u occurring is the product of the probabilities of each bit in the transcript. The product of the probabilities for Alice's bits is some function $p_{\mathsf{A}}(u)$ which depends on x and the product of the probabilities for Bob's bits is some function $p_{\mathsf{B}}(u)$ which depends on y and $\Pi_{x,y}(u) = p_{\mathsf{A}}(u)p_{\mathsf{B}}(u)$. Alice can also estimate $p_{\mathsf{B}}(u)$ by taking the average over y of $\Pi_y(u)$. Call this estimate $q_{\mathsf{A}}(u)$; similarly for Bob's estimate $q_{\mathsf{B}}(u)$. Set $\nu_{\mathsf{A}} = p_{\mathsf{A}}q_{\mathsf{A}}$ and $\nu_{\mathsf{B}} = q_{\mathsf{B}}p_{\mathsf{B}}$.

The challenge is that Alice and Bob know only $(p_{\mathsf{A}}, q_{\mathsf{A}})$ and $(p_{\mathsf{B}}, q_{\mathsf{B}})$ respectively and do not know τ (in our setting, $\tau = \Pi_{x,y}$). They use a variant of rejection sampling, in which Alice will overestimate q_{A} by a factor 2^{Δ}; likewise for Bob. Let us define the set of Δ-bad elements with respect to τ,ν as follows:

$$B_\Delta(\tau,\nu) = \{u \in \mathcal{U} \mid 2^{\Delta}\nu(u) < \tau(u)\}.$$

Intuitively, u is bad if τ gives much more weight to it than ν. Observe that if $\tau = p_{\mathsf{A}}p_{\mathsf{B}}, \nu_{\mathsf{A}} = p_{\mathsf{A}}q_{\mathsf{A}}$, then $u \notin B_\Delta(\tau,\nu_{\mathsf{A}})$ implies that $2^{\Delta}q_{\mathsf{A}}(u) \geq p_{\mathsf{B}}(u)$.

To prove our compression lemma, we use the following claim about the single sampling experiment.

Claim IV.2. *Let $B = B_\Delta(\tau,\nu_{\mathsf{A}}) \cup B_\Delta(\tau,\nu_{\mathsf{B}})$. Let $\gamma = \tau(B)$. Then the following holds about Protocol IV.1:*

1) *The probability that Alice accepts equals $\frac{1}{|\mathcal{U}|2^{\Delta}}$ and the same for Bob.*
2) *The probability that the experiment is accepted is at most $\frac{1}{|\mathcal{U}|2^{2\Delta}}$ and at least $\frac{1-\gamma}{|\mathcal{U}|2^{2\Delta}}$.*
3) *Let τ' denote the distribution of the output of the experiment, conditioned on it being accepted. Then $|\tau - \tau'| \leq \gamma$.*

Fix a finite universe $\mathcal{U}$. Let $p_{\mathsf{A}}, q_{\mathsf{A}}, p_{\mathsf{B}}, q_{\mathsf{B}} : \mathcal{U} \to [0,1]$ such that $\tau = p_{\mathsf{A}} p_{\mathsf{B}}$, $\nu_{\mathsf{A}} = p_{\mathsf{A}} q_{\mathsf{A}}$, $\nu_{\mathsf{B}} = p_{\mathsf{B}} q_{\mathsf{B}}$ are all probability distributions.
Alice's input: $p_{\mathsf{A}}, q_{\mathsf{A}}$. Bob's input: $p_{\mathsf{B}}, q_{\mathsf{B}}$. Common input: parameter $\Delta > 0$.

1) Using public coins, sample $u \leftarrow \mathcal{U}$, and $\alpha, \beta \leftarrow [0, 2^{\Delta}]$.
2) Alice accepts the run if $\alpha \le p_{\mathsf{A}}(u)$ and $\beta \le 2^{\Delta} q_{\mathsf{A}}(u)$.
3) Bob accepts the run if $\alpha \le 2^{\Delta} q_{\mathsf{B}}(u)$ and $\beta \le p_{\mathsf{B}}(u)$.
4) If both Alice and Bob accept, then we say that the experiment is accepted and the output is u. Otherwise, the output is $\perp$.

Protocol IV.1. Single sampling experiment

Intuitively, this claim says that Alice accepts each single experiment with probability $\frac{1}{|\mathcal{U}|2^{\Delta}}$, and also implies that conditioned on Alice accepting the i'th experiment, it is relatively likely that Bob accepts it. Therefore, by repeating this experiment enough times, there is reasonable probability of Alice and Bob both accepting the same execution of the experiment. Conditioned on the experiment accepting, the output of the experiment is distributed close to the original distribution τ. In the next section, we show how to use a hash function to select a common accepting execution of the experiment out of many executions.

B. Description and analysis of the zero-communication protocol

Let μ be any distribution on inputs and π be any protocol with information complexity $I = \mathsf{IC}_{\mu}(\pi)$. Let (X, Y, Π) be the joint random variables where X, Y are distributed according to μ and Π is the distribution of the transcript of the protocol π applied to X, Y (by slight abuse of notation we use the letter Π for both the transcript and its distribution). Let $\Pi_{x,y}$ be Π conditioned on $X = x, Y = y$, Π_x be Π conditioned $X = x$, and Π_y likewise.

Let $\mathcal{U}$ be the space of all possible transcripts. We assume that each transcript contains the output of the protocol. As shown in [2] and described above, Alice can construct functions $p_{\mathsf{A}}, q_{\mathsf{A}} : \mathcal{U} \to [0,1]$ and Bob can construct functions $p_{\mathsf{B}}, q_{\mathsf{B}} : \mathcal{U} \to [0,1]$, such that for all $u \in \mathcal{U}$, $\Pi_{x,y}(u) = p_{\mathsf{A}}(u) p_{\mathsf{B}}(u)$, $\Pi_x(u) = p_{\mathsf{A}}(u) q_{\mathsf{A}}(u)$, and $\Pi_y(u) = p_{\mathsf{B}}(u) q_{\mathsf{B}}(u)$.

The zero-communication protocol π' is described in Protocol IV.3. This protocol is an extension of the one in [8], where here Alice uses public coins to guess the hash function value instead of calculating and transmitting it to Bob and both players are allowed to abort when they do not believe they can output the correct value.

Alice's input: x. Bob's input: y. Common inputs: $\delta > 0, I > 0$.
Set parameters: $\Delta = \frac{4}{\delta} \cdot (\frac{8 \cdot I}{\delta} + 1)$ and $T = |\mathcal{U}| 2^{\Delta} \ln(8/\delta)$ and $k = \Delta + \log(\frac{64}{\delta} \ln(8/\delta)^2)$.

1) Alice constructs functions $p_{\mathsf{A}}, q_{\mathsf{A}} : \mathcal{U} \to [0,1]$ and Bob constructs functions $p_{\mathsf{B}}, q_{\mathsf{B}} : \mathcal{U} \to [0,1]$, such that for all transcripts $u \in \mathcal{U}$, $\Pi_{x,y}(u) = p_{\mathsf{A}}(u) p_{\mathsf{B}}(u)$, $\Pi_x(u) = p_{\mathsf{A}}(u) q_{\mathsf{A}}(u)$, and $\Pi_y(u) = p_{\mathsf{B}}(u) q_{\mathsf{B}}(u)$.
2) (**Run experiments.**) Using public coins, Alice and Bob run Protocol IV.1 T independent times with inputs $p_{\mathsf{A}}, q_{\mathsf{A}}, p_{\mathsf{B}}, q_{\mathsf{B}}$ and Δ.
3) Let $\mathcal{A} = \{i \in [T] : \text{Alice accepts experiment } i\}$ and similarly $\mathcal{B}$ for Bob. If either set is empty, that party outputs the abort symbol $\perp$.
4) (**Find intersection.**) Using public coins, Alice and Bob choose a random function $h : [T] \to \{0,1\}^k$ and a random string $r \in \{0,1\}^k$.
   a) Alice finds the smallest $i \in \mathcal{A}$. If $h(i) \neq r$ then Alice outputs $\perp$. Otherwise, Alice outputs in accordance with the transcript of experiment i.
   b) Bob finds the smallest $j \in \mathcal{B}$ such that $h(j) = r$. If no such j exists, he outputs $\perp$. Otherwise, Bob outputs in accordance with the transcript of experiment j.

Protocol IV.3. Zero-communication protocol π' derived from π

In order to analyze our protocol, we first define some events and give bounds on their probabilities.

Definition IV.4. We define the following events over the probability space of sampling (X, Y) according to μ and running π' on (X, Y) to produce a transcript Π:

1) **Large divergence.** B_D occurs if $(X, Y) = (x, y)$ such that $D(\Pi_{x,y} \parallel \Pi_x) > \frac{8\mathsf{IC}_{\mu}(\pi)}{\delta}$ or $D(\Pi_{x,y} \parallel \Pi_y) > \frac{8\mathsf{IC}_{\mu}(\pi)}{\delta}$. We will also let B_D denote the set of such (x, y).
2) **Collision.** B_C occurs if there exist distinct $i, j \in \mathcal{A} \cup \mathcal{B}$ such that $h(i) = h(j) = r$.
3) **Protocol outputs something.** H occurs if $\pi'(X, Y) \neq \perp$.

The proof of the main compression lemma (Lemma III.4) uses the following claim.

Claim IV.5. *The probability of the above events are bounded as follows:*

1) *The inputs rarely have large divergence:*

$\Pr_{(X,Y)\sim\mu}[B_D] \le \delta/4$.

2) *For all (x,y), the hash function rarely has a collision:* $\Pr_{r_{\pi'}}[B_C] \le \frac{\delta}{16}\cdot 2^{-(k+\Delta)}$.
3) *For all $(x,y) \notin B_D$, the probability of outputting something is not too small:* $\Pr_{r_{\pi'}}[H] \ge (1 - \frac{11\delta}{16})2^{-(k+\Delta)}$.
4) *For all (x,y) the probability of outputting something is not too large:* $\Pr_{r_{\pi'}}[H] \le (1 + \frac{\delta}{16})2^{-(k+\Delta)}$.
5) *For all protocols π, input distributions μ and $\delta > 0$, the protocol π' in Protocol IV.3 satisfies: For all $(x,y) \notin B_D$, let $\Pi'_{x,y,H}$ be the distribution of $\pi'(x,y)$ conditioned on H (namely, on $\pi'(x,y) \neq \perp$). Then $|\Pi_{x,y} - \Pi'_{x,y,H}| \le 3\delta/4$.*

C. Proof of the Compression Lemma

Proof of Lemma III.4: Set $\lambda = 2^{-(k+\Delta)}$. It holds that $\lambda \ge 2^{-C(\mathsf{IC}_\mu(\pi)/\delta^2+1/\delta)}$ for $C = 64$. Let $\mathcal{R}$ be any subset of the support of $(X,Y,\pi(X,Y))$. Then

$$\begin{aligned}&\Pr_{r_\pi,(X,Y)\sim\mu}[(X,Y,\pi(X,Y)) \in \mathcal{R}]\\ &\le \Pr_{(X,Y)\sim\mu}[B_D] + \Pr_{(X,Y)\sim\mu}[\neg B_D]\\ &\qquad\cdot \Pr_{r_\pi,(X,Y)\sim\mu}[(X,Y,\pi(X,Y)) \in \mathcal{R} \mid \neg B_D].\end{aligned}$$

Applying Item 5 of Claim IV.5 and the fact that $\mathcal{R}$ is simply an event, it follows that for all $(x,y) \notin B_D$

$$\begin{aligned}&\Pr_{r_\pi}[(x,y,\pi(x,y)) \in \mathcal{R}]\\ &\quad\le \Pr_{r_{\pi'}}[(x,y,\pi'(x,y)) \in \mathcal{R} \mid \pi'(x,y) \neq \perp] + \tfrac{3\delta}{4}.\end{aligned}$$

Since $\Pr[B_D] \le \delta/4$ (Item 1 of Claim IV.5),

$$\begin{aligned}&\Pr_{r_\pi,(X,Y)\sim\mu}[(X,Y,\pi(X,Y)) \in \mathcal{R}]\\ &\le \Pr_{r_{\pi'},(X,Y)\sim\mu}[(X,Y,\pi'(X,Y)) \in \mathcal{R} \mid \pi'(x,y) \neq \perp]\\ &\qquad+\delta.\end{aligned}$$

This proves one direction of Equation 4. For the other direction, we have that

$$\begin{aligned}&\Pr[(X,Y,\pi'(X,Y)) \in \mathcal{R} \mid \pi'(X,Y) \neq \perp] &(7)\\ &\le \Pr[B_D] + \Pr[\neg B_D]\\ &\qquad\cdot \Pr[(X,Y,\pi'(X,Y))) \in \mathcal{R} \mid \neg B_D, \pi'(X,Y) \neq \perp]\\ &\le \tfrac{\delta}{4} + \Pr[\neg B_D]\\ &\qquad\cdot \left(\Pr[(X,Y,\pi(X,Y)) \in \mathcal{R} \mid \neg B_D] + \tfrac{3\delta}{4}\right) &(8)\\ &\le \Pr[(X,Y,\pi(X,Y)) \in \mathcal{R} \wedge \neg B_D] + \delta\\ &\le \Pr[(X,Y,\pi(X,Y)) \in \mathcal{R}] + \delta\end{aligned}$$

where in Equation 8 we applied Item 5 of Claim IV.5. This proves Equation 4 of Lemma III.4. ∎

Equation 5 follows immediately from Item 4 of Claim IV.5.

Finally, for Equation 6, we may write:

$$\begin{aligned}&\Pr_{r_{\pi'},(X,Y)\sim\mu}[\pi'(X,Y) \neq \perp]\\ &\ge \Pr_{(X,Y)\sim\mu}[\neg B_D] \Pr_{r_{\pi'},(X,Y)\sim\mu}[\pi'(X,Y) \neq \perp | \neg B_D]\\ &\ge (1-\tfrac{\delta}{4})(\Pr_{r_{\pi'},(X,Y)\sim\mu}[H \mid \neg B_D])\\ &\ge (1-\tfrac{\delta}{4})(1-\tfrac{11\delta}{16})\,\lambda\\ &> (1-\delta)\,\lambda\end{aligned}$$

where we used Item 3 of Claim IV.5. ∎

V. Applications

We can prove lower bounds on the information complexity of specific problems, by checking that their communication lower bounds were obtained by one of the methods subsumed by the relaxed partition bound, including the factorization norm, smooth rectangle, rectangle, or discrepancy. However, a bit of care is required to ascertain this. For example, while a paper may say it uses the "rectangle bound", we must still verify that the value of the linear program for $\bar{\mathsf{prt}}$ (or one of the subsumed programs such as srec or rec) is at least the claimed bound, since different authors may use the term "rectangle bound" to mean different things. In particular what they call "rectangle bound" may not satisfy the constraints of the rectangle/smooth rectangle linear programs given by Jain and Klauck [1]. After we have verified that $\bar{\mathsf{prt}}$ is appropriately bounded, then we can apply our main theorem (Theorem I.1). We do this for the problems below.

A. Exponential separation of quantum communication and classical information complexity

We prove that the quantum communication complexity of the Vector in Subspace Problem is exponentially smaller than its classical information complexity (Theorem I.2). In the Vector in Subspace Problem $\mathrm{VSP}_{0,n}$, Alice is given an $n/2$ dimensional subspace of an n dimensional space over $\mathbb{R}$, and Bob is given a vector. This is a partial function, and the promise is that either Bob's vector lies in the subspace, in which case the function evaluates to 1, or it lies in the orthogonal subspace, in which case the function evaluates to 0. Note that the input set of $\mathrm{VSP}_{0,n}$ is continuous, but it can be discretized by rounding, which leads to the problem $\widetilde{\mathrm{VSP}}_{\theta,n}$ (see [3] for details).

Klartag and Regev [3] show that the Vector in Subspace Problem can be solved with an $O(\log n)$ quantum protocol, but the randomized communication

complexity of this problem is $\Omega(n^{1/3})$. Their lower bound uses a modified version of the rectangle bound, which can be shown to be still weaker than our relaxed partition bound.

Lemma V.1. *There exist universal constants C and γ such that for any ϵ,*

$$\bar{\mathsf{prt}}_\epsilon(\widetilde{\mathrm{VSP}}_{\theta,n}) \geq \frac{1}{C}(0.8 - 2.8\epsilon)\exp(\gamma n^{1/3}).$$

This allows us to conclude that the information complexity of this function is at least $\Omega(n^{1/3})$. This solves Braverman's Open Problem 3 (Are there problems for which $Q(f,\epsilon) = O(\mathrm{polylog}(\mathsf{IC}(f,\epsilon)))$?) and Open Problem 7 (is it true that $\mathsf{IC}(\widetilde{\mathrm{VSP}}_{\theta,n}, 1/3) = n^{\Omega(1)}$?)

Moreover, our result implies an exponential separation between classical and quantum information complexity. We refrain from defining quantum information cost and complexity in this paper (see [17] for a definition), but since the quantum information complexity is always smaller than the quantum communication complexity, the separation follows trivially from Theorem I.2.

B. Information complexity of the Gap Hamming Distance Problem

We prove that the information complexity of Gap Hamming Distance is $\Omega(n)$ (Theorem I.3; Open Problem 6 in [2]). In the Gap Hamming Distance Problem (GHD_n), Alice and Bob each receive a string of length n and they need to determine whether their Hamming distance is at least $n/2 + \sqrt{n}$ or less than $n/2 - \sqrt{n}$. We prove that the information complexity of Gap Hamming Distance is $\Omega(n)$ (Theorem I.3; Open Problem 6 in [2]). The communication complexity of Gap Hamming Distance was shown to be $\Omega(n)$ by Chakrabarti and Regev [19]. The proof was subsequently simplified by Vidick [20] and Sherstov [18]. The first two proofs use the smooth rectangle bound, while Sherstov uses the rectangle/corruption bound.

The corruption bound used by Sherstov is a slight refinement of the rectangle bound as defined by Jain and Klauck [1], since it can handle distributions that put small weight on the set of inputs that map to some function value z. It can be shown that this bound is weaker than our relaxed partition bound, which implies Theorem I.3.

Lemma V.2. *There exist universal constants C and δ such that for any small enough ϵ, $\bar{\mathsf{prt}}_\epsilon(\mathrm{GHD}) \geq C2^{\delta n}$.*

VI. Conclusions and open problems

We have shown that the information complexity is lower bounded by a relaxed version of the partition bound. This subsumes all known algebraic and rectangle based methods, except the partition bound. It remains to see if the partition bound also provides a lower bound on the information complexity. Alternatively, if we would like to separate the communication and information complexities, then possible candidates could be functions whose partition bound is strictly larger than their relaxed partition bound.

Moreover, we have seen how the relaxed partition bound naturally relates to zero-communication protocols with abort. Actually, we can relate all other lower bound methods to different variants of zero-communication protocols [15, 16]. This provides new insight on the inherent differences between these bounds and may lead to new lower bound methods, coming from different versions of zero-communication protocols. Moreover, since these protocols have been extensively studied in the field of quantum information, it is intriguing to see what other powerful tools can be transferred to the model of classical communication complexity.

Acknowledgment

We would like to thank Amit Chakrabarti and Oded Regev for helpful comments. J.R. acknowledges support from the action *Mandats de Retour 2010* of the *Politique Scientifique Fédérale Belge.* This research was funded in part by the EU grant QCS, ANR Jeune Chercheur CRYQ, ANR Blanc QRAC (ANR-08-EMER-012), and EU ANR Chist-ERA DIQIP.

References

[1] R. Jain and H. Klauck, "The partition bound for classical complexity and query complexity," in *Proc. 25th CCC*, 2010, pp. 247–258.

[2] M. Braverman, "Interactive information complexity," in *Proc. 44th STOC*, 2012, pp. 505–524.

[3] B. Klartag and O. Regev, "Quantum one-way communication can be exponentially stronger than classical communication," in *Proc. 43rd STOC*, 2011, pp. 31–40.

[4] A. Chakrabarti, Y. Shi, A. Wirth, and A. Yao, "Informational complexity and the direct sum problem for simultaneous message complexity," in *Proc. 42nd FOCS*, 2001, pp. 270–278.

[5] Z. Bar-Yossef, T. S. Jayram, R. Kumar, and D. Sivakumar, "An information statistics approach to data stream and communication complexity," *Journal of Computer and System Sciences*, vol. 68, no. 4, pp. 702–732, 2004.

[6] B. Barak, M. Braverman, X. Chen, and A. Rao, "How to compress interactive communication," in *Proc. 42nd STOC*, 2010, pp. 67–76.

[7] M. Braverman and A. Rao, "Information equals amortized communication," in *Proc. 52nd FOCS*, 2011, pp. 748–757.

[8] M. Braverman and O. Weinstein, "A discrepancy lower bound for information complexity," in *Proc. 16th RANDOM*, 2012, to appear.

[9] N. Linial and A. Shraibman, "Lower bounds in communication complexity based on factorization norms," *Random Structures and Algorithms*, vol. 34, no. 3, pp. 368–394, 2009.

[10] T. Lee and A. Shraibman, "Lower bounds in communication complexity," *Foundations and Trends in Theoretical Computer Science*, vol. 3, no. 4, pp. 263–399, 2009.

[11] B. Gisin and N. Gisin, "A local hidden variable model of quantum correlation exploiting the detection loophole," *Phys. Lett. A*, vol. 260, pp. 323–327, 1999.

[12] S. Massar, "Non locality, closing the detection loophole and communication complexity," *Phys. Rev. A*, vol. 65, 2002.

[13] H. Buhrman, P. Høyer, S. Massar, and H. Röhrig, "Combinatorics and quantum nonlocality," *Phys. Rev. Lett.*, vol. 91, 2003.

[14] ——, "Multipartite nonlocal quantum correlations resistant to imperfections," *Phys. Rev. A*, vol. 73, 2006.

[15] J. Degorre, M. Kaplan, S. Laplante, and J. Roland, "The communication complexity of non-signaling distributions," *Quantum Information and Computation*, vol. 11, no. 7–8, pp. 649–676, 2011.

[16] S. Laplante, V. Lerays, and J. Roland, "Classical and quantum partition bound and detector inefficiency," in *Proc. 39th ICALP*, 2012, pp. 617–628.

[17] R. Jain and A. Nayak, "The space complexity of recognizing well-parenthesized expressions in the streaming model: the Index function revisited," ECCC, Tech. Rep. TR10-071, 2010.

[18] A. Sherstov, "The communication complexity of Gap Hamming Distance," *Theory of Computing*, vol. 8, no. 8, pp. 197–208, 2012.

[19] A. Chakrabarti and O. Regev, "An optimal lower bound on the communication complexity of gap-hamming-distance," in *Proc. 43rd STOC*, 2011, pp. 51–60.

[20] T. Vidick, "A concentration inequality for the overlap of a vector on a large set with application to the communication complexity of the Gap-Hamming-Distance problem," *Chicago Journal of Theoretical Computer Science*, vol. 2012, pp. 1–12, 2012.

[21] R. Bar-Yehuda, B. Chor, E. Kushilevitz, and A. Orlitsky, "Privacy, additional information and communication," *IEEE Transactions on Information Theory*, vol. 39, no. 6, pp. 1930–1943, 1993.

[22] H. Klauck, "On quantum and approximate privacy," in *Proc. 19th STACS*, vol. 2285, 2002, pp. 735–735.

[23] J. Feigenbaum, A. D. Jaggard, and M. Schapira, "Approximate privacy: foundations and quantification (extended abstract)," in *Proc. 11th ACM EC'10*, 2010, pp. 167–178.

[24] A. Ada, A. Chattopadhyay, S. Cook, L. Fontes, M. Koucky, and T. Pitassi, "The hardness of being private," in *Proc. 27th CCC*, 2012, pp. 192–202.

[25] T. S. Jayram, R. Kumar, and D. Sivakumar, "Two applications of information complexity," in *Proc. 35th STOC*, 2003, pp. 673–682.

[26] R. Jain, J. Radhakrishnan, and P. Sen, "A direct sum theorem in communication complexity via message compression," in *Proc. 30th ICALP*, 2003, pp. 300–315.

[27] ——, "Prior entanglement, message compression and privacy in quantum communication," in *Proc. 20th CCC*, 2005, pp. 285–296.

[28] P. Harsha, R. Jain, D. McAllester, and J. Radhakrishnan, "The communication complexity of correlation," in *Proc. 22nd CCC*, 2007, pp. 10–23.

[29] A. Chakrabarti, R. Kondapally, and Z. Wang, "Information Complexity versus Corruption and Applications to Orthogonality and Gap-Hamming," in *Proc. 16th RANDOM*, 2012, to appear.