Equality Alone Does not Simulate Randomness

Arkadev Chattopadhyay
School of Technology and Computer Science, Tata Institute of Fundamental Research, Mumbai, India
arkadev.c@tifr.res.in

Shachar Lovett
Department of Computer Science and Engineering, University of California, San Diego, USA
slovett@ucsd.edu

Marc Vinyals
School of Technology and Computer Science, Tata Institute of Fundamental Research, Mumbai, India
marc.vinyals@tifr.res.in

Abstract

The canonical problem that gives an exponential separation between deterministic and randomized communication complexity in the classical two-party communication model is "Equality". In this work we show that even allowing access to an "Equality" oracle, deterministic protocols remain exponentially weaker than randomized ones. More precisely, we exhibit a total function on n bits with randomized one-sided communication complexity $O(\log n)$, but such that every deterministic protocol with access to "Equality" oracle needs $\Omega(n)$ cost to compute it.

Additionally we exhibit a natural and strict infinite hierarchy within BPP, starting with the class P^{EQ} at its bottom.

2012 ACM Subject Classification Theory of computation → Communication complexity

Keywords and phrases Communication lower bound, derandomization

Digital Object Identifier 10.4230/LIPIcs.CCC.2019.14

Funding *Arkadev Chattopadhyay*: Part of the work done while visiting the Simons Institute for the Theory of Computing.
Shachar Lovett: Supported by NSF award 1614023.
Marc Vinyals: Supported by the Prof. R Narasimhan Foundation.

Acknowledgements We are grateful to Bruno Loff, Jaikumar Radhakrishnan and especially Suhail Sherif for helpful discussions in the early stages of this work. We thank Sagnik Mukhopadhyay and anonymous reviewers for providing useful feedback to an earlier version of this manuscript. We also thank the Simons Institute for the Theory of Computing at Berkeley where some of this work took place.

1 Introduction

A deterministic communication protocol in Yao's two-party model is a strategy for a collaborative game between two parties, Alice and Bob, each of whom receives an input and whose task is to compute a function while communicating as little as possible.

It has been known since the origins of communication complexity that randomized protocols, where the parties are given access to a source of randomness and are allowed to make errors with small probability, are strictly more powerful than deterministic protocols. The classic example is the Equality function over n-bit strings, which has a randomized protocol with $\mathrm{O}(\log n)$ bits of communication, while every deterministic protocol requires at least $n + 1$ bits [15].

34th Computational Complexity Conference (CCC 2019).
Editor: Amir Shpilka; Article No. 14; pp. 14:1–14:11

Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

An efficient protocol for Equality is obtained by using a fingerprinting technique: use the randomness source to obtain a fingerprint of the strings to be compared of length $O(\log n)$, exchange the fingerprints, and answer whether the fingerprints are equal.

A few more examples of functions where randomness is helpful are the 'Greater-Than' function [13], the sparse set disjointness problem [8], and the Hamming distance problem with a small threshold [16]. In all cases the fingerprinting technique is enough to efficiently solve the problems. Is fingerprinting all there is to randomized protocols?

To state this question in a formal way we consider a model of communication where the parties are given access to an oracle that solves the Equality problem and are charged a cost of one bit each time the parties call the oracle. The set of functions that can be computed by some protocol in this model with cost $\operatorname{polylog} n$ bits is called P^{EQ}. The set of functions that have randomized protocols of cost $\operatorname{polylog} n$ is called BPP. We overload notation and use P^{EQ} and BPP to refer to both the class of functions and the corresponding communication models respectively. The question then is whether every function that has a randomized protocol with c bits of communication, also has a P^{EQ} protocol with $\operatorname{poly}(c, \log n)$ bits of communication and oracle calls. In other words, is $\mathsf{P}^{\mathrm{EQ}} = \mathsf{BPP}$?

The P^{EQ} model was first considered in [3]. The knowledge about it until our work (for total functions, see discussion below) can be summarized as follows:

$$\mathsf{P} \subsetneq \mathsf{P}^{\mathrm{EQ}} \subseteq \mathsf{BPP}.$$

P^{EQ} is also strictly weaker than the P^{NP} model, since EQ calls can be simulated with an NP oracle but P^{EQ} cannot efficiently solve the coNP-complete set disjointness problem. It also is worth mentioning that giving access to an Equality oracle is equivalent to giving access to a Greater-Than oracle up to a logarithmic factor. The latter model was introduced as real communication by Krajíček [10], with a connection to proof complexity in mind, and later found further applications in the same area [5, 4].

Partial functions

There are many examples in the literature of *partial functions* that separate P^{EQ} from BPP. One such example is the gap Hamming distance problem with a large gap. Concretely, the problem is to distinguish between pairs of input strings whose Hamming distance is less than a 1/3-fraction and more than a 2/3-fraction. This can be solved with a randomized protocol with $O(1)$ bits that samples a position in the strings uniformly at random and answers whether the strings are the same at that position. On the other hand, this problem has cost $\Omega(n)$ in the P^{NP} model [14], and hence in the P^{EQ} model too.

A different example follows from the simulation theorem of [5], made explicit in [6], and it is to lift a (partial) function that exhibits an exponential gap between deterministic and randomized query complexity, say promised majority. To be more precise, we consider the majority function of n bits with the promise that the fraction of zeros is either less than 1/3 or more than 2/3, which can be computed with a randomized decision tree by querying the input at a constant number of randomly sampled points, but requires linearly many queries to be solved by a deterministic decision tree. If we compose this function with the indexing gadget with pointers of size $O(\log n)$ then we have a randomized protocol of cost $O(\log n)$ that evaluates a constant number of instances of the gadget, while the simulation theorem tells us that it requires real communication $\Omega(n \log n)$.

Total functions

The question about a separation between P^{EQ} and BPP for *total functions* requires a different approach. If one uses the same means as before, namely lifting theorems, then a quadratic separation follows for example from the pointer chasing function [2] composed with indexing. However, this is where the lifting from query complexity approach seems to end, since deterministic and randomized query complexity are known to be polynomially related for total functions [12]. Our main result is a non-lifted total function, which exhibits an exponential separation between P^{EQ} and randomized communication.

▶ **Definition 1.1.** *The integer inner product problem* $\mathrm{IIP}_{m,t}(x,y)$ *is defined as follows. The inputs are integer vectors* $x, y \in [-M, M]^t$ *where* $M = 2^m$*. The output is* 1 *if* $\langle x, y\rangle = 0$*, where the inner product is computed over the integers.*

We denote by IIP_t the family of functions $\mathrm{IIP}_{m,t}$ with fixed $t = \mathrm{O}(1)$ and growing m. Note that the input size of $\mathrm{IIP}_{m,t}$ is $n = (m+1)t$.

▶ **Theorem 1.2** (Main theorem, informal)**.** *For any* $t \geq 6$*, the total function* IIP_t *on* n *bits can be computed with* $\mathrm{O}(\log n)$ *bits of randomized communication but requires* $\Omega(n)$ *cost to be solved by* P^{EQ} *protocols.*

Once we settled that EQ is not enough to simulate BPP because P^{EQ} cannot efficiently solve IIP, the next natural candidate for an oracle A such that $\mathsf{P}^A = \mathsf{BPP}$ becomes IIP itself. However, we also show that for any fixed t, IIP_t is not enough to simulate BPP, and in fact the complexity classes defined by IIP oracles form a strict infinite hierarchy.

▶ **Theorem 1.3.** *There is an infinite sequence* $(t_i)_{i\in\mathbb{N}}$ *such that*

$$\mathsf{P} \subsetneq \mathsf{P}^{\mathrm{EQ}} \subsetneq \mathsf{P}^{\mathrm{IIP}_{t_1}} \subsetneq \cdots \subsetneq \mathsf{P}^{\mathrm{IIP}_{t_i}} \subsetneq \mathsf{P}^{\mathrm{IIP}_{t_{i+1}}} \subsetneq \cdots \subset \mathsf{BPP} \ .$$

2 Preliminaries

We assume familiarity with standard definitions in communication complexity, such as in [11]. The only somewhat non-standard definition we need is that of protocols with access to an oracle.

If A is a family of communication problems $A_N \colon \{0,1\}^N \times \{0,1\}^N \to \{0,1\}$ for $N \in \mathbb{N}$, then the parties involved in a P^A protocol communicate via an oracle for A. Informally, if the players hold inputs $(x,y) \in \{0,1\}^n \times \{0,1\}^n$, every message is a pair of inputs $(g_1(x), g_2(y)) \in \{0,1\}^N \times \{0,1\}^N$ for one of the functions A_N, where g_1 and g_2 have been agreed beforehand, and the output $A_N(x,y)$ is visible to both parties. We assume that A is nontrivial in the sense that it can simulate sending one-bit messages from each party to the other one. The cost of such a protocol is the number of bits the oracle outputs, and $\mathsf{P}^A(f)$ is the minimum over all protocols. In particular, P^{EQ} is a protocol with oracle access to the Equality oracle, and P^{GT} is a protocol with oracle access to the Greater-Than oracle, both of which are nontrivial.

Usually the cost of an oracle call is defined with an additional term logarithmic in the size of its inputs, since otherwise we could solve any function with a single call to a strong oracle such as set disjointness. The kind of oracles we consider are weak enough that we do not need any limits on the input size to prove lower bounds, hence we omit the additional term for simplicity.

In fact, in our analysis, after a call to the oracle we immediately partition the set of inputs compatible with the answer into a set of rectangles. This makes it convenient to work

with a stronger model where all the possible sets of answers are partitioned beforehand, and the oracle tells the players not only the answer to their query, but also which rectangle in the partition their input belongs to, at no extra cost.

Observe that calling a function A_N with inputs transformed by g_1 and g_2 is equivalent to calling a function $B = A_N \circ (g_1, g_2)$ and that the matrix of B can be obtained from the matrix of A_N by removing, duplicating, and permuting some rows or columns. Therefore we identify an oracle A with the smallest family of matrices $\mathcal{M}_A$ that contains all the communication matrices of the functions A_N, and is closed under removing, duplicating, and permuting rows or columns. To each matrix $M \in \mathcal{M}_A$ we associate a monochromatic rectangle partition $\mathcal{R}(M)$, i.e., a set of rectangles such that for each rectangle $R \in \mathcal{R}(M)$ the submatrix of M defined by R is an all-zeros or all-ones matrix. In general, there may be many such choices; a good choice will be crucial for our lower bound technique. The only requirement is that this partition is to monochromatic rectangles, and hence a refinement of the answer given by the oracle.

A P^A protocol to compute a function $f\colon \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ is a tree where each node corresponds to a rectangle $R \subseteq \{0,1\}^n \times \{0,1\}^n$ of compatible inputs. Each internal node is associated with a matrix $M \in \mathcal{M}_A$ of the same dimensions as R, and has one child for each rectangle $R' \in \mathcal{R}(M)$. Upon reaching a node labelled by R the players move to the child R' that contains their input. Each leaf is labelled 0 or 1, and the label of a leaf R equals $f(x, y)$ for each $(x, y) \in R$.

Analogous to how one bit of deterministic communication induces a refined partition of the input space where each rectangle is split into two, one call to an oracle induces a refined partition where each rectangle R is replaced by the partition $\mathcal{R}(M(R))$ associated to a matrix $M(R) \in \mathcal{M}_A$ of the same size. This is, we start with a single rectangle $\mathcal{R}_0 = \{\{0,1\}^n \times \{0,1\}^n\}$, and after i calls to the oracle we have the partition $\mathcal{R}_i = \bigcup_{R \in \mathcal{R}_{i-1}} \mathcal{R}(M(R))$. If a protocol computes a function f after c calls, then the partition $\mathcal{R}_c$ applied to M_f yields a set of monochromatic rectangles.

3 A Lower Bound Technique for P with Oracle Access

The goal of this section is to develop a lower bound technique for P^{EQ}, and more generally for P with oracle access. The key property of EQ that we exploit is that, no matter how it is transformed by an oracle call, we can always partition the matrix of EQ into few rectangles so that a large area is monochromatic. More generally, if we denote the number of elements in a matrix M by $|M|$, we define the property as follows.

▶ **Definition 3.1.** *A family of Boolean matrices $\mathcal{M}$ has* ϵ-monochromatic rectangles *if every matrix $M \in \mathcal{M}$ contains a monochromatic rectangle – i.e., an all-zeros or all-ones submatrix – of size at least $\epsilon|M|$.*

We obtain our lower bounds by estimating the following complexity measure.

▶ **Definition 3.2.** *If $\mathcal{R}$ is a set of rectangles and $\eta \in (1/2, 1)$ is a real number, we denote the* η-area *of $\mathcal{R}$ by $p_\eta(\mathcal{R}) = \sum_{R_i \in \mathcal{R}} |R_i|^\eta$. The η-area of a matrix M is the minimum of $p_\eta(\mathcal{R})$ over all monochromatic partitions $\mathcal{R}$ of M.*

Observe that the η-area of a matrix M is bounded below by $|M|^\eta$, which is attained if and only if the matrix is monochromatic, and above by $|M|$, which corresponds to partitioning the matrix into singletons. In fact, partitioning into either rows or columns gives a better upper bound of $2|M|^{(1+\eta)/2}$ for any matrix, and it can be shown that the matrix of inner product modulo 2 attains this bound up to a constant factor.

The *relative η-area* of a matrix M is $q_\eta(M) = p_\eta(M)/|M|^\eta$. Note that $q_\eta(M) \geq 1$ with equality attained if and only if M is monochromatic. The relative η-area of a family of matrices is the maximum relative η-area over all matrices in the family.

▶ **Lemma 3.3.** *For any η such that $1/(1-\log_2(1-\epsilon)) < \eta < 1$ there exists a constant $\xi = \xi(\epsilon,\eta)$ such that every ϵ-monochromatic family of matrices $\mathcal{M}$ that is closed under taking submatrices satisfies $q_\eta(\mathcal{M}) \leq \xi$.*

Proof. We prove the lemma by induction over the size of the matrices in the family. This is clearly true for 1×1 matrices; otherwise consider a matrix $M \in \mathcal{M}$ of size $r = |M|$. By assumption M contains a monochromatic rectangle R_1 of size $r_1 \geq \epsilon r$, so we can partition M into R_1 and two non-monochromatic rectangles R_2 and R_3 of respective sizes r_2 and r_3. We then apply the induction hypothesis to each non-monochromatic rectangle, while noting that the η-area of R_1 is r_1^η:

$$\begin{aligned} p_\eta(M) &\leq r_1^\eta + p_\eta(R_2) + p_\eta(R_3) \\ &\leq r_1^\eta + \xi r_2^\eta + \xi r_3^\eta \\ &\leq (1+2\xi)\left(\frac{r_1 + \xi r_2 + \xi r_3}{1+2\xi}\right)^\eta \\ &= (1+2\xi)^{1-\eta}\left(r_1 + \xi r_2 + \xi r_3\right)^\eta \\ &\leq (1+2\xi)^{1-\eta}\big(\xi + (1-\xi)\epsilon\big)^\eta r^\eta \ . \end{aligned}$$

We can write $(1-\epsilon) = (2+\delta)^{1-1/\eta}$ with $\delta > 0$ by the assumption on η. Set $\alpha = (2+\delta/2)^{1-1/\eta}$ so that $\alpha > (1-\epsilon)$ and set $\xi = \max\{2/\delta, \epsilon/(\alpha-(1-\epsilon))\}$. Then we can bound

$$1 + 2\xi = \xi(2 + 1/\xi) \leq \xi(2+\delta/2) = \xi\alpha^{1/(1-1/\eta)}$$

and

$$\xi + (1-\xi)\epsilon = \xi(1 - \epsilon + \epsilon/\xi) \leq \xi\alpha$$

so that

$$p_\eta(M) \leq (1+2\xi)^{1-\eta}\big(\xi + (1-\xi)\epsilon\big)^\eta r^\eta \leq \xi r^\eta\big(\alpha^{-\eta}\alpha^\eta\big) = \xi r^\eta \ . \qquad \blacktriangleleft$$

For simplicity we can take $\eta = 1 - \epsilon > 1/(1-\log_2(1-\epsilon))$ whenever $0 < \epsilon < 1/2$.

▶ **Lemma 3.4.** *Assume that f is a function which has a P^A protocol with cost c. For any $\eta \in (0,1)$ the communication matrix of f has relative η-area $q_\eta(f) \leq (q_\eta(\mathcal{M}_A))^c$.*

Proof. First we associate to each matrix $M \in \mathcal{M}_A$ a partition $\mathcal{R}(M)$ with relative η-area at most $q = q_\eta(\mathcal{M}_A)$. Next, assume that we have a partition of the input space into rectangles $\mathcal{R}$ with η-area $p_\eta(\mathcal{R})$. For each rectangle $R_i \in \mathcal{R}$ choose a matrix $M_i \in \mathcal{M}_A$ of the same dimensions. We obtain a refined partition $\mathcal{R}'$ by replacing each rectangle R_i by $\mathcal{R}(M_i)$. We can bound the total η-area of $\mathcal{R}'$ by

$$p_\eta(\mathcal{R}') = \sum_{R_i \in \mathcal{R}} p_\eta(\mathcal{R}(M_i)) \leq \sum_{R_i \in \mathcal{R}} q \cdot |R_i|^\eta = q \cdot p_\eta(\mathcal{R}) \ .$$

As $\mathcal{R}, \mathcal{R}'$ are partitions of the same dimensions, their relative η-areas satisfy

$$q_\eta(\mathcal{R}') \leq q \cdot q_\eta(\mathcal{R}) \ .$$

To conclude the proof, let $\mathcal{R}_0, \ldots, \mathcal{R}_c$ denote the intermediate partitions induced by the protocol, where $\mathcal{R}_i$ is the partition obtained after the first i calls. Then $\mathcal{R}_0$ is the singleton partition, $\mathcal{R}_c$ is a monochromatic partition of M_f, and all partitions have the same dimensions. Thus $q_\eta(\mathcal{R}_0) = 1$ and $q_\eta(\mathcal{R}_i) \leq q \cdot q_\eta(\mathcal{R}_{i-1})$ for $i = 1, \ldots, c$. We conclude that $q_\eta(M_f) \leq q_\eta(\mathcal{R}_c) \leq q^c$ as claimed. ◀

The next lemma gives an easy to verify condition under which Lemma 3.4 can be applied.

▶ **Lemma 3.5.** *Fix $0 < \eta < 1$. Let A be an oracle with constant relative η-area and let f be an n-bit function with a corresponding $2^n \times 2^n$ communication matrix M. Assume that:*
1. *The number of entries i, j with $M_{i,j} = 1$ is $\alpha 2^{2n}$.*
2. *For any 1-monochromatic rectangle R in M it holds that $|R| \leq \beta 2^{2n}$.*

Then the communication complexity of f in P^A is $\Omega(\log(\alpha\beta^{\eta-1}))$.

Proof. Let $\mathcal{R}$ be a partition of $f^{-1}(1)$ with minimum η-area. Let $x_i = |R_i|/2^{2n}$ denote the density of each rectangle R_i. Then the following minimization problem lower bounds the η-area of $\mathcal{R}$:

$$p_\eta(\mathcal{R}) \geq 2^{2\eta n} \cdot \min_{\sum_i x_i = \alpha, 0 \leq x_i \leq \beta} \sum_i x_i^\eta \ .$$

The minimum of a concave function over a convex polytope is attained at a vertex, in this case any point with $\lfloor \alpha/\beta \rfloor$ coordinates equal to β, one coordinate equal to $\alpha - \lfloor \alpha/\beta \rfloor \beta$, and the rest equal to 0. Hence

$$p_\eta(\mathcal{R}) \geq 2^{2\eta n} \lfloor \alpha/\beta \rfloor \beta^\eta \ .$$

If f has a P^A protocol with cost c, then by Lemma 3.4

$$p_\eta(\mathcal{R}) \leq 2^{2\eta n} (q_\eta(\mathcal{M}_A))^c = 2^{2\eta n + \mathrm{O}(c)} \ .$$

Rearranging these gives $c \geq \Omega(\log(\alpha\beta^{\eta-1}))$ as claimed. ◀

3.1 An Improved Bound for Equality

Coming back to the particular case of P^{EQ}, it is not hard to prove that the $\mathcal{M}_{\mathrm{EQ}}$ family of matrices has 1/9-monochromatic rectangles, and hence Lemma 3.5 applies to EQ with $\eta = 8/9$. While this is already enough to separate P^{EQ} and BPP, some of our applications require a tighter bound on η.

To obtain a better bound it is convenient to consider instead the model of P^{GT}, where the players have oracle access to a Greater-Than oracle. Note that as an EQ oracle can be simulated by two calls to a GT oracle, the latter model is stronger.

We show that $\mathcal{M}_{\mathrm{GT}}$ has constant η-area for any $\eta > 1/2$. The matrix GT_N is *monotone*, in the sense that it satisfies $M_{i_1,j_1} \leq M_{i_2,j_2}$ for all pairs of entries such that $i_1 \leq i_2$ and $j_1 \leq j_2$, and duplicating or removing rows and columns preserves monotonicity. Therefore every matrix $M \in \mathcal{M}_{\mathrm{GT}}$ is (a permutation of) a monotone matrix.

▶ **Lemma 3.6.** *A monotone matrix M can be partitioned into four rectangles R_1, R_2, R_3, R_4, such that R_1, R_2 are monochromatic and $|R_1| + |R_2| \geq |R_3| + |R_4|$.*

Proof. Let a and b be the dimensions of the matrix M and assume without loss of generality that $a \geq b$. Let a_1 be the maximal number such that $M_{a_1,b_1} = 0$, with $b_1 = \lceil a_1 b/a \rfloor$. Then the rectangle $R_1 = [1, a_1] \times [1, b_1]$ is 0-monochromatic, while the rectangle $R_2 = [a_1 + 1, a] \times$

$[b_1+1, b]$ is 1-monochromatic. We define $R_3 = [1, a_1] \times [b_1+1, b]$ and $R_4 = [a_1+1, a] \times [1, b_1]$. To complete the proof let $a_2 = a - a_1$ and $b_2 = b - b_2$, and observe that if $a_1 > a_2$ then $b_1 \geq b_2$, while if $a_1 < a_2$ then $b_1 \leq b_2$. Therefore by the rearrangement inequality

$$|R_1| + |R_2| = a_1 b_1 + a_2 b_2 \geq a_1 b_2 + a_2 b_1 = |R_3| + |R_4| \ . \qquad \blacktriangleleft$$

We use this partition to prove a more refined version of Lemma 3.3.

▶ **Lemma 3.7.** *For any $1/2 < \eta < 1$ there exists a constant $\xi = \xi(\eta)$ such that $q_\eta(\mathcal{M}_{\mathrm{GT}}) \leq \xi$.*

Proof. The proof is analogous to that of Lemma 3.3, except that we use Lemma 3.6 to partition each matrix into two monochromatic rectangles R_1 and R_2, and two non-monochromatic rectangles R_3 and R_4. We then get a bound

$$\begin{aligned} p_\eta(M) &\leq r_1^\eta + r_2^\eta + p_\eta(R_3) + p_\eta(R_4) \\ &\leq r_1^\eta + r_2^\eta + \xi r_3^\eta + \xi r_4^\eta \\ &\leq (2+2\xi)\left(\frac{r_1 + r_2 + \xi r_3 + \xi r_4}{2+2\xi}\right)^\eta \\ &\leq (2+2\xi)\left(\frac{r_1 + r_2 + r_3 + r_4}{4}\right)^\eta \\ &= \xi r^\eta \end{aligned}$$

for $\xi = 1/(2^{2\eta-1} - 1)$. $\blacktriangleleft$

It follows that Lemma 3.5 holds for both EQ and GT with $1/2 < \eta < 1$.

4 Separation

We demonstrate the separation by considering the inner product function over the integers. We recall the definition from the introduction.

▶ **Definition 4.1.** *The integer inner product problem $\mathrm{IIP}_{m,t}(x, y)$ is defined as follows. The inputs are integer vectors $x, y \in [-M, M]^t$ where $M = 2^m$. The output is 1 if $\langle x, y \rangle = 0$, where the inner product is computed over the integers.*

We use n to denote the input length, where $n = (m+1)t$. We recall that we consider $t = \mathrm{O}(1)$ and growing m.

▶ **Lemma 4.2.** *There is a* coRP *protocol for $\mathrm{IIP}_{m,t}$ of cost $\mathrm{O}(t \log m)$.*

Proof. Consider the following protocol: sample a uniformly random prime q among the first $4m + 2\log t$ primes, compute $\langle x, y \rangle \pmod q$ by having Alice send t integers $x_i \pmod q$ to Bob, and accept if and only if $\langle x, y \rangle = 0 \pmod q$. The protocol uses $\mathrm{O}(t \log q) = \mathrm{O}(t \log m)$ bits of communication.

The protocol is always correct on 1-inputs. To see that it is correct on 0-inputs with probability at least $1/2$ we observe that the probability of failure is the probability of picking a prime q that divides $\langle x, y \rangle$. Since the number $\langle x, y \rangle$ is bounded by tM^2 in absolute value, it is divisible by at most $\log(tM^2) = 2m + \log t$ primes, and since we have $4m + 2\log t$ primes to choose from, the probability of failure is at most $1/2$. $\blacktriangleleft$

▶ **Lemma 4.3.** *If t is even then $\Pr_{x,y}[\mathrm{IIP}_{m,t}(x, y) = 1] = \Omega(1/tM^2)$.*

Proof. Write $x = (x', -x'')$ and $y = (y', y'')$ where $x', y', x'', y'' \in [-M, M]^{t/2}$, so that $\langle x, y\rangle = \langle x', y'\rangle - \langle x'', y''\rangle$. The distribution of $\langle x', y'\rangle$ and $\langle x'', y''\rangle$ are i.i.d. and take at most $\mathrm{O}(tM^2)$ possible values. So the collision probability is $\Omega(1/tM^2)$. ◀

▶ **Lemma 4.4.** *For any rectangle $R \subseteq \mathrm{IIP}^{-1}_{m,t}(1)$ we have $|R| \leq (4M)^t$.*

Proof. Let $A, B \subset [-M, M]^t$ such that $\langle x, y\rangle = 0$ for all $x \in A, y \in B$. Let p be a prime between $2M + 1$ and $4M$, and consider the problem modulo p. Note that we can injectively identify A, B with subsets of $\mathbb{F}_p^t$. Let V, W denote the linear subspaces of $\mathbb{F}_p^t$ spanned by A, B, respectively. Then $V \perp W$ and hence $|V||W| \leq p^t$. This implies that $|A||B| \leq p^t \leq (4M)^t$. ◀

▶ **Lemma 4.5.** *Any P^{EQ} protocol for $\mathrm{IIP}_{m,t}$ with even $t \geq 6$ has cost $\Omega(n)$.*

Proof. Apply Lemma 3.5 with $\eta = \frac{1}{2} + \frac{1}{100}$, $\alpha = \Omega(1/tM^2)$ as given by Lemma 4.3, and $\beta = (4M)^t/(2M+1)^{2t} \leq 1/M^t$ as given by Lemma 4.4. We obtain

$$\mathsf{P}^{\mathrm{EQ}}(\mathrm{IIP}_{m,t}) = \Omega(\log(\alpha\beta^{\eta-1})) = \Omega(\log(M^{t(1-\eta)-2}/t)) = \Omega(tm) = \Omega(n) \ . \qquad ◀$$

Theorem 1.2 follows immediately from Lemma 4.2 and Lemma 4.5.

A related example

We give a similar separation by the inner product function over polynomials. Let $\mathbb{F}_2[z]$ denote the ring of univariate polynomials over $\mathbb{F}_2$.

▶ **Definition 4.6.** *The polynomial inner product problem $\mathrm{PIP}_{m,t}(x, y)$ is defined as follows. The inputs x, y are t-tuples of polynomials in $\mathbb{F}_2[z]$, each of degree at most m. The output is 1 if $\langle x, y\rangle = 0$, where the inner product is computed over $\mathbb{F}_2[z]$.*

Note that also here the input size is $n = (m+1)t$. Again we consider large m and $t = \mathrm{O}(1)$.

▶ **Lemma 4.7.** *There is a coRP protocol for $\mathrm{PIP}_{m,t}$ of cost $\mathrm{O}(t \log m)$.*

Proof. Consider the following protocol. Alice and Bob interpret their polynomials as polynomials in $\mathbb{F}_q[z]$ with $q = 2^{\lceil \log m \rceil + 2}$. They sample a uniformly random point $z \in \mathbb{F}_q$ and compute $\langle x, y\rangle(z)$ by having Alice send the result of evaluating each of her polynomials at z. The protocol uses $\mathrm{O}(t \log q) = \mathrm{O}(t \log m)$ bits of communication.

The protocol is always correct on 1-inputs. To see that it is correct on 0-inputs with probability at least $1/2$ we observe that the probability of failure is the probability of picking a root of $\langle x, y\rangle$. Since the number of roots is at most $2m$ and we have $q \geq 4m$ points in $\mathbb{F}_q$ to choose from, the probability of failure is at most $1/2$. ◀

▶ **Lemma 4.8.** *Any P^{EQ} protocol for $\mathrm{PIP}_{m,t}$ with even $t \geq 6$ has cost $\Omega(n)$.*

The proof is analogous to that of Lemma 4.5. We can use Lemma 4.3 unchanged, and we adapt Lemma 4.4 by considering the inner product function over $\mathbb{F}_q$ with $q = 2^m$.

Set disjointness

Babai et al. [3] were the first who attempted to prove a strong lower bound on the cost of any P^{EQ} protocol solving DISJ, however their method only yielded lower bounds for one-way protocols. The subsequent breakthrough tight bound of $\Omega(n)$ by Kalyanasundaram and Schnitger [9] on the randomized complexity of DISJ yields an $\Omega(n/\log n)$ bound on the P^{EQ} cost of DISJ. Using the techniques developed here, we prove a simple tight lower bound of $\Omega(n)$ on the cost of P^{EQ} protocols for set disjointness that does not rely on lower bounds for BPP.

▶ **Lemma 4.9.** *Any* P^{EQ} *protocol for* DISJ *has cost* $\Omega(n)$.

Proof. Apply Lemma 3.5 with $\eta = \frac{1}{2} + \frac{1}{100}$, $\alpha = (3/4)^n$, and $\beta = 1/2^n$. We obtain

$$\mathsf{P}^{\mathrm{EQ}}(\mathrm{DISJ}) = \Omega(\log(\alpha\beta^{\eta-1})) = \Omega(\log(2^{(\log 3 - 2 + 0.49)n}) = \Omega(\log(2^{0.07n})) = \Omega(n) \ . \qquad \blacktriangleleft$$

5 Hierarchy

A generic way to obtain ϵ-monochromatic families is by extracting large rectangles from matrices of small *sign-rank*. A real matrix M, each of whose entries are non-zero, is said to be sign represented by another matrix A if each entry of A and M agree in sign. The sign-rank of M is the minimum r such that there exists an A of rank r that sign represents it. A corollary of the following theorem allows us to extract large rectangles from matrices of small sign-rank.

▶ **Theorem 5.1** ([1])**.** *Let* U *and* V *be finite multisets of vectors in* $\mathbb{R}^d$ *and let* $\delta = 1/2^{d+1}$. *Then there are subsets* $U' \subset U$ *and* $V' \subset V$ *such that* $|U'| \geq \delta|U|$, $|V'| \geq \delta|V|$, *and either* $\langle u, v\rangle \geq 0$ *for all* $u, v \in U' \times V'$ *or* $\langle u, v\rangle < 0$ *for all* $u, v \in U' \times V'$.

▶ **Corollary 5.2.** *A Boolean matrix of sign-rank* d *and size* r *contains a monochromatic rectangle of size at least* $1/2^{2(d+1)}r$.

Proof. Let M be a matrix of size $n \times m$ and sign rank d, and let A and B be matrices of size $n \times d$ and $d \times m$ such that $M = \mathrm{sign}(AB)$. Apply Theorem 5.1 to the set of rows of A and the set of columns of B. $\blacktriangleleft$

Since sign-rank does not increase with respect to removing, duplicating, or permuting rows or columns, in order to establish that IIP_t is ϵ-monochromatic, it is sufficient to look at the sign-rank of IIP_t.

▶ **Lemma 5.3.** *The sign-rank of* $\mathrm{IIP}_{m,t}$ *is at most* $t^2 + 1$.

Proof. $\mathrm{IIP}_{m,t}(x, y) = \mathrm{sign}(\langle x, y\rangle^2 - 1/2)$, which can be decomposed into a linear combination of t^2 rank-one matrices of the form $M_{x,y} = \langle x_i x_j, y_i y_j\rangle$ and the all-ones matrix. $\blacktriangleleft$

We can now put all the pieces together and prove a lower bound for $\mathsf{P}^{\mathrm{IIP}_t}$.

▶ **Lemma 5.4.** *Any* $\mathsf{P}^{\mathrm{IIP}_t}$ *protocol for* $\mathrm{IIP}_{m,t'}$ *with even* $t' \geq 2^{3t^2}$ *has cost* $\Omega(n)$.

Proof. Let $\epsilon = 1/2^{2(t^2+1)}$ given by Corollary 5.2. Then $\mathcal{M}_{\mathrm{IIP}_t}$ is an ϵ-monochromatic family, therefore we can apply Lemma 3.5 with $\eta = 1 - \epsilon$. Choose t' to be the smallest even integer such that $(2\log t')/t' < (1 - \eta)$. We can bound t' by

$$t' \leq \frac{4}{1-\eta}\log\left(\frac{1}{1-\eta}\right) = \frac{4}{\epsilon}\log\left(\frac{1}{\epsilon}\right) \leq 2^{3t^2} \ .$$

Apply Lemma 3.5 with $\alpha = \Omega(1/t'M^2)$ as given by Lemma 4.3, and $\beta \leq 1/M^{t'}$ as given by Lemma 4.4. We obtain

$$\mathsf{P}^{\mathrm{IIP}_t}(\mathrm{IIP}_{m,t'}) = \Omega(\log(\alpha\beta^{\eta-1})) = \Omega\left(t'm\left(-\frac{2\log t'}{t'} + 1 - \eta\right)\right) = \Omega(t'm) = \Omega(n) \ . \qquad \blacktriangleleft$$

To prove Theorem 1.3 we consider the sequence of classes $\mathsf{P}^{\mathrm{IIP}_{t_i}}$ where $t_1 = 6$ and $t_{i+1} = 2^{3t_i^2}$. The inclusion $\mathsf{P}^{\mathrm{EQ}} \subseteq \mathsf{P}^{\mathrm{IIP}_{t_1}}$ follows from the observation that $\mathrm{EQ}(x, y) = \mathrm{IIP}_2((x, 1), (-1, y))$, and Lemma 4.5 shows that it is strict. The inclusions $\mathsf{P}^{\mathrm{IIP}_{t_i}} \subseteq \mathsf{P}^{\mathrm{IIP}_{t_{i+1}}}$ are immediate since we can solve $\mathrm{IIP}_{m,t}$ with a single call to $\mathrm{IIP}_{m,t'}$ padding the additional coordinates with zeros, and we just proved the non-inclusions in Lemma 5.4.

6 Concluding Remarks

This work belongs to the general area of understanding the power of randomness in communication complexity. We use this opportunity to remind the readers of a fascinating open problem, posed explicitly by Göös, Pitassi and Watson [7], which is whether $\mathsf{BPP} \subset \mathsf{P}^{\mathsf{NP}}$ for *total functions*. It is known that this containment is not true for partial functions. Göös et al. suggested, as a first step, separating the class of total functions in BPP from an interesting subclass of P^{NP}. In this work, we took this step by providing the first (exponential) separation between BPP and P^{EQ}, the latter being one of the most natural subclasses of P^{NP}. However, the original problem of separating BPP from P^{NP} remains open.

To state this in combinatorial terms, a function f has a P^{NP} protocol of cost c if the following holds. There exists a list of 2^c rectangles R_i and values $z_i \in \{0,1\}$, such that $f(x,y) = z_i$ for the *first* rectangle R_i in the list for which $(x,y) \in R_i$ (We may assume that the last rectangle contains all possible inputs, to make this model well defined). In particular, if $\mathsf{BPP} \subset \mathsf{P}^{\mathsf{NP}}$ then there must exist a monochromatic rectangle in f of density $2^{-\mathrm{O}(c)}$ for $c = \operatorname{polylog} n$. Understanding this question seems to be pivotal towards understanding the relation between BPP and P^{NP}.

▶ **Problem 6.1.** *Let f be an n-bit total Boolean function with a randomized protocol of cost c. Is it true that f must contain a monochromatic rectangle R of size $|R| \geq 2^{-\mathrm{O}(c)}2^{2n}$?*

References

1. Noga Alon, János Pach, Rom Pinchasi, Radoš Radoičić, and Micha Sharir. Crossing patterns of semi-algebraic sets. *Journal of Combinatorial Theory, Series A*, 111(2):310–326, 2005. doi:10.1016/j.jcta.2004.12.008.
2. Andris Ambainis, Kaspars Balodis, Aleksandrs Belovs, Troy Lee, Miklos Santha, and Juris Smotrovs. Separations in Query Complexity Based on Pointer Functions. *Journal of the ACM*, 64(5):32:1–32:24, 2017. Preliminary version in *STOC '16*. doi:10.1145/3106234.
3. László Babai, Péter Frankl, and János Simon. Complexity Classes in Communication Complexity Theory. In *Proceedings of the 27th Annual Symposium on Foundations of Computer Science (FOCS '86)*, pages 337–347, October 1986. doi:10.1109/SFCS.1986.15.
4. Paul Beame, Noah Fleming, Russell Impagliazzo, Antonina Kolokolova, Denis Pankratov, Toniann Pitassi, and Robert Robere. Stabbing Planes. In *Proceedings of the 9th Innovations in Theoretical Computer Science Conference (ITCS '18)*, volume 94 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 10:1–10:20, January 2018. doi:10.4230/LIPIcs.ITCS.2018.10.
5. María Luisa Bonet, Juan Luis Esteban, Nicola Galesi, and Jan Johannsen. On the Relative Complexity of Resolution Refinements and Cutting Planes Proof Systems. *SIAM Journal on Computing*, 30(5):1462–1484, 2000. Preliminary version in *FOCS '98*. doi:10.1137/S0097539799352474.
6. Susanna F. de Rezende, Jakob Nordström, and Marc Vinyals. How Limited Interaction Hinders Real Communication (and What It Means for Proof and Circuit Complexity). In *Proceedings of the 57th Annual IEEE Symposium on Foundations of Computer Science (FOCS '16)*, pages 295–304, October 2016. doi:10.1109/FOCS.2016.40.
7. Mika Göös, Toniann Pitassi, and Thomas Watson. The Landscape of Communication Complexity Classes. *Computational Complexity*, 27(2):245–304, June 2018. Preliminary version in *ICALP '16*. doi:10.1007/s00037-018-0166-6.
8. Johan Håstad and Avi Wigderson. The Randomized Communication Complexity of Set Disjointness. *Theory of Computing*, 3(1):211–219, 2007. doi:10.4086/toc.2007.v003a011.
9. Bala Kalyanasundaram and Georg Schnitger. The Probabilistic Communication Complexity of Set Intersection. *SIAM J. Discrete Math.*, 5(4):545–557, 1992. doi:10.1137/0405044.

10 Jan Krajíček. Interpolation by a Game. *Mathematical Logic Quarterly*, 44(4):450–458, 1998. doi:10.1002/malq.19980440403.

11 Eyal Kushilevitz and Noam Nisan. *Communication Complexity*. Cambridge University Press, New York, NY, USA, 1997.

12 Noam Nisan. CREW PRAMs and decision trees. *SIAM Journal on Computing*, 20(6):999–1007, December 1991. doi:10.1137/0220062.

13 Noam Nisan. The communication complexity of threshold gates. In *Proceedings of Combinatorics, Paul Erdős is Eighty*, volume 1, pages 301–315, 1993.

14 Periklis A. Papakonstantinou, Dominik Scheder, and Hao Song. Overlays and Limited Memory Communication. In *Proceedings of the 29th IEEE Conference on Computational Complexity (CCC '14)*, pages 298–308, June 2014. doi:10.1109/CCC.2014.37.

15 Andrew Chi-Chih Yao. Some Complexity Questions Related to Distributive Computing (Preliminary Report). In *Proceedings of the 11th Annual ACM Symposium on Theory of Computing (STOC '79)*, pages 209–213, April 1979. doi:10.1145/800135.804414.

16 Andrew Chi-Chih Yao. On the power of quantum fingerprinting. In *Proceedings of the 35th Annual ACM Symposium on Theory of Computing (STOC '03)*, pages 77–81, June 2003. doi:10.1145/780542.780554.