Applications of Monotone Rank to Complexity Theory

Yang D. Li [*]

November 15, 2011

Abstract

Raz's recent result [Raz10] has rekindled people's interest in the study of *tensor rank*, the generalization of matrix rank to high dimensions, by showing its connections to arithmetic formulas. In this paper, we follow Raz's work and show that *monotone rank*, the monotone variant of tensor rank and matrix rank, has applications in algebraic complexity, quantum computing and communication complexity. A common point of tensor rank and monotone rank is that they are both NP-hard to compute [Has90, Vav09], and are also hard to bound. This paper differs from Raz's paper in that it leverages existing results to show unconditional bounds while Raz's result relies on some assumptions.

More concretely, we show the following things.

- We show a super-exponential separation between monotone and non-monotone computation in the non-commutative model, and thus provide a strong solution to Nisan's question [Nis91] in algebraic complexity. More specifically, we exhibit that there exists a homogeneous algebraic function f of degree d (d even) on n variables with the monotone algebraic branching program (ABP) complexity $\Omega(d^2 \log n)$ and the non-monotone ABP complexity $O(d^2)$.

- In Bell's theorem [Bel64, CHSH69], a basic assumption is that players have free will, and under such an assumption, local hidden variable theory still cannot predict the correlations produced by quantum mechanics. Using tools from monotone rank, we show that even if we disallow the players to have free will, local hidden variable theory still cannot predict the correlations produced by quantum mechanics.

- We generalize the log-rank conjecture [LS88] in communication complexity to the multiparty case, and prove that for super-polynomial parties, there is a super-polynomial separation between the deterministic communication complexity and the logarithm of the rank of the communication tensor. This means that the log-rank conjecture does not hold in high dimensions.

---

[*]Email: danielliy@gmail.com. University of Illinois at Urbana-Champaign.

1 Introduction

Computational complexity focuses on studying the minimum amount of resources required for carrying out computational tasks. The resources may be time, space, randomness (public or private), communication, quantum entanglement and so on. Readers can refer to the textbook by Arora and Barak [AB09] for more information on this subject.

Matrix rank plays a key role for proving lower bounds, and sometimes upper bounds, in many models of computation with an algebraic or combinatorial flavor, like algebraic branching programs, span programs, and communication complexity. Another important notion is *tensor rank*, which is shown to be crucial in arithmetic formulas [Raz10]. Tensor rank is the generalization of matrix rank. The monotone variant of tensor rank and matrix rank is called *monotone rank* and is defined as follows, where $[n]$ denotes the set $\{1, 2, \ldots, n\}$, $\otimes$ means tensor product and $\mathbb{R}^+$ represents the set of nonnegative real numbers.

Definition 1 (monotone rank) *The* monotone rank *of a tensor* $M : \Pi_{j=1}^d [n_j] \to \mathbb{R}^+$ *(d ≥ 2) is the minimum r such that $M = \sum_{i=1}^r v_{1,i} \otimes v_{2,i} \otimes \cdots \otimes v_{d,i}$, where $v_{j,i} \in (\mathbb{R}^+)^{n_j}, j \in [d], i \in [r]$. It is denoted as $mr(M)$.*

When $d \geq 3$, monotone rank is the *monotone tensor rank*, as discussed in [AFT11]; when $d = 2$, monotone rank specializes to *monotone matrix rank* (or positive rank/nonnegative rank), first mentioned in [Yan91]. According to a recent report by Lee and Shraibman [LS09], there are no lower bounds which actually use monotone rank in practice. The main problem with monotone rank as a complexity measure is that it is extremely difficult to bound for explicit functions. Consequently, to the best of our knowledge, this paper is the first to connect monotone rank to real applications in complexity theory.

A brief preview of our results is that we want to quantify the power of negation. Just like Valiant's famous result [Val79] on the complexities for computing the permanent and determinant of a matrix, the usual and monotone ranks of a nonnegative tensor can also differ dramatically. Morally speaking, the reason for these differences is the obvious possibility of cancelation in the non-monotone case. In algebraic complexity, monotone computation does not allow substraction and the coefficients of the monomials are all positive; while the non-monotone model does not have such restrictions. In quantum computing, Feynman [Fey82] points out that the only difference between the probabilistic world and the quantum world is that it happens as if the "probabilities" would have to go negative in the quantum world. Moreover, the separation of communication complexity and the log-rank of the communication tensor is partly due to whether or not we allow negative decomposition.

1.1 Algebraic Complexity

Valiant [Val80] shows an exponential separation between monotone and non-monotone computation in terms of algebraic circuit complexity. Nisan [Nis91] asks if the same difference between monotone and non-monotone can be achieved in a restricted model called non-commutative model, which prohibits the commutativity of multiplication, for some complexity measure.

We answer this question by showing the following theorem. The gap is more than exponential in terms of algebraic branching program (ABP) complexity, where ABP can be regarded as the analog of branching program in algebraic computation.

Theorem 1 *There exists an homogeneous algebraic function f of degree d (d even) on n variables with the monotone ABP complexity $\Omega(d^2 \log n)$ and the non-monotone ABP complexity $O(d^2)$.*

1.2 Quantum Computing

Bell's theorem [Bel64,CHSH69] basically states that local hidden variable theory cannot predict the correlations produced by quantum mechanics. In Bell's theorem, the measurements are
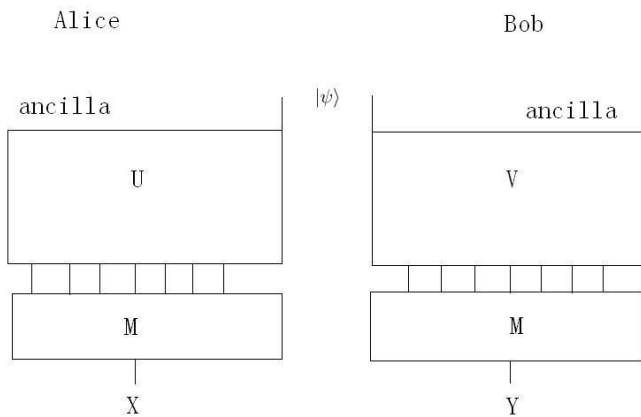
Figure 1: Bell's Theorem with Fixed Measurement

versatile and can be chosen be with respect to various bases; namely the players have the free will to have their own choices and make their own decisions. In a game-theoretic term, they are selfish players. This paper departs from Bell's paradigm by assuming that there are no choices for Alice and Bob and that the measurements Alice and Bob will make are fixed from the start. The model is roughly illustrated by Figure 1.

In Figure 1, $|\psi\rangle$ represents an entangled state of size $2Q$ ($Q \leq n$) qubits, where the first Q qubits are owned by Alice and the second Q qubits belong to Bob. We may also add some ancilla qubits to make sure that both Alice and Bob have n qubits. The state in the beginning is ϕ_0. Alice applies unitary operation U and Bob applies unitary operation V. The state becomes $\phi_1 = (U \otimes V)\phi_0$. U and V are fixed, so that Alice and Bob have no freedom at all. The role of U and V here is to amplify the hardness of simulating quantum correlations.

Then Alice and Bob both apply the measurement M, which is fixed to be with respect to the standard basis. M is assumed to be fixed from the start. In the end, they output a correlation (X, Y) according to results of the measurement, where X and Y are random variables taking values in $\{0,1\}^n$. X and Y are correlated in the sense that their distributions are not independent. Suppose the distribution of (X, Y) is P_r, and we want to reproduce P_r using local hidden variables. Note that P_r can be treated as a matrix, namely

$$(P_r)_{xy} = Prob\{X = x, Y = y\}, \tag{1}$$

for all $x, y \in \{0,1\}^n$.

For the classical simulation, Alice and Bob initially share some random bits (shared randomness, public coins, or local hidden variables). We denote the shared random variable as Z. Alice and Bob also have some private random bits, denoted as r_A and r_B respectively. They use Z, r_A and r_B to generate a correlation (X', Y'), such that $X' = f_A(Z, r_A)$ and $Y' = f_B(Z, r_B)$. X' and Y' are random variables taking values in $\{0,1\}^n$. Suppose that the probability distribution of (X', Y') is P_c, we want to make sure that P_c is exactly the same as P_r. If P_c and P_r are the same, then we succeed in simulating quantum correlations using local hidden variables.

Based on the model above, we have the following result.

Theorem 2 *There exists a 2-qubit ($Q = 1$) quantum state $|\psi\rangle$, some proper U and V, such that at least $\log n$ shared random bits are needed to simulate the quantum correlations.*

3

Theorem 2 is very interesting and subtle: local hidden variables cannot account for quantum correlations even when the measurements Alice and Bob will make are prescribed from the start. That is, Alice and Bob share $Q = 1$ pair and produce $2n$ classical correlated shared bits (X, Y). Local hidden variables fail, not at any particular value of n, but in the limit as n tends to infinity, because the number of local hidden variables needed to account for the results grows in a super-exponential speed ($\Omega(\log n)$ vs. $O(1)$) to infinity. Maybe we can say that the problem is that there is no thermodynamic limit. That is to say, the number of hidden variables per pair shared by Alice and Bob is not an intensive quantity. Our argument is not as convincing as Bell's, but it goes beyond Bell's theorem in the sense that it shows that if hidden variables explain quantum correlations, then there is some pathology in the explanation.

1.3 Communication Complexity

Arguably the most well-known open problem in communication complexity is the log-rank conjecture [LS88], which is stated as follows:

Conjecture 1 *There exists a constant c, such that for every function $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$,*
$$D(f) = O(\log^c rk(M(f))),$$
where $D(f)$ is the deterministic communication complexity of f and $M(f)$ is the communication matrix of f.

Although a number of people aspire to resolve this conjecture, very little progress has been made in the last two decades [NW95, RS95]. In fact, the conjecture can be easily generalized to the number-in-hand multiparty communication complexity model. Suppose there are d parties in the communication.

Conjecture 2 *There exists a constant c, such that for every function $f : (\{0,1\}^n)^d \to \{0,1\}$,*
$$D(f) = O(\log^c rk(M(f))),$$
where $M(f)$ is the communication tensor of f.

We have the following theorem for the generalized log-rank conjecture.

Theorem 3 *For every $d = \omega(n^{c'}), \forall c' > 0$, there exists a function $f : (\{0,1\}^n)^d \to \{0,1\}$, such that for every constant $c > 0$,*
$$D(f) = \omega(\log^c rk(M(f))).$$

Thus, we provide a super-polynomial separation between the deterministic communication complexity and the logarithm of the rank of the communication tensor when there are super-polynomial parties. This means that the log-rank conjecture does not hold in high dimensions.

1.4 Related Work

Independent from our paper, in a completely different model and using another complexity measure, [HY11] shows a much weaker (super-polynomial) separation between the monotone and non-monotone computation in the non-commutative model. We note that [DKW09] also involves the generalized log-rank conjecture, but it focuses on the case when the number of players is small. There are lots of follow-ups of Bell's theorem. The idea of a great deal of papers (such as [BCT99, BCvD01, BT03, TB03, RT09]) is that local hidden variables augmented by communication could reproduce the results of quantum entanglement. Quantum entanglement has plenty of applications in areas such as quantum teleportation [BBC+93], superdense coding [BW92] and quantum cryptography [BB84]. [Zha12] and [Win05] are also related in simulating quantum correlations, but they are in a game-theoretic and/or an information-theoretic setting and the selfish players there have the free will. [Hru11] builds up a relation between monotone rank and the complexity of boolean formula.

2 Preliminaries

2.1 ABP Complexity

We briefly recall some necessary definitions, notations, and results from Nisan's paper [Nis91]. For more details, please refer to [Nis91]. All the logarithms in this paper have base 2.

Definition 2 *[Nis91] An* algebraic branching program (ABP) *is a directed acyclic graph with one source and one sink. The vertices of the graph are partitioned into levels numbered from 0 to d, where edges may only go from level i to level $i+1$. d is called the degree of the ABP. The source is the only vertex at level 0 and the sink is the only vertex at level d. Each edge is labeled with a homogeneous linear function of $x_1, x_2, \ldots, x_n$ (namely a function of the form $\sum_i c_i x_i$). The size of an ABP is the number of vertices, which is denoted by $B(f)$.*

The ABP model is non-commutative (see [Nis91]). An ABP computes a function in the following sense: the sum over all paths from the source to the sink, of the product of the linear functions by which the edges of the path are labeled. An ABP of degree d computes a homogeneous polynomial of degree d.

Definition 3 *[Nis91] An ABP is called* monotone *if all constants used as coefficients in the linear forms are positive. The monotone ABP complexity of f are denoted as $B^+(f)$.*

Definition 4 *[Nis91] For a function f of degree d, and $0 \leq k \leq d$, the k-monotone-ABP complexity of f, $B_k^+(f)$ is the minimum, over all monotone ABPs that compute f, of the size of the k'th level of the ABP.*

We use $rk(M)$ to denote the rank of a matrix M. Let f be a homogeneous function of degree d on n variables. For each $0 \leq k \leq d$, we define a real matrix $M_k(f)$ of dimensions n^k by n^{d-k} as follows: there is a row for each sequence of k variables (k-term), and a column for each $d-k$-term. The entry at $(x_{i_1} \cdots x_{i_k}, x_{j_1} \cdots x_{j_{d-k}})$ is defined to be the real coefficient of the monomial $x_{i_1} \cdots x_{i_k} x_{j_1} \cdots x_{j_{d-k}}$ in f.

Lemma 4 *[Nis91] For any homogeneous function f of degree d, $B(f) = \sum_{k=0}^d rk(M_k(f))$.*

Lemma 5 *[Nis91] For every homogeneous function f of degree d and all $0 \leq k \leq d$, $B_k^+(f) = mr(M_k(f))$. Also, $B^+(f) \geq \sum_{k=0}^d B_k^+(f)$.*

2.2 Monotone Rank

We review some of the known results on monotone rank.

Given n distinct real numbers $a_1, a_2, \ldots, a_n$, a $n \times n$ matrix M can be defined by $M_{ij} = (a_j - a_i)^2$, $i, j \in [n]$. Such a matrix is called a Euclidean distance matrix and has the following properties.

Lemma 6 *[BL09] $rk(M) = 3$.*

Lemma 7 *[BL09] $mr(M) \geq \log n$.*

[BL09] conjectures that all Euclidean distance matrices satisfy $mr(M) = n$ and [LC10] claims that they prove the conjecture. But Hrubes [Hru11] refutes the conjecture in [BL09] and "theorem" in [LC10] by showing the following counter-example.

Lemma 8 *[Hru11] Let $a_i = i$ for all $i \in [n]$. Then, $mr(M) \leq 2\log n + 2$.*

So we put forward a new conjecture based on the new result of Hrubes.

Conjecture 3 *For all possible distinct real numbers $a_1, a_2, \ldots, a_n$, $mr(M) = \Theta(\log n)$.*

A tensor $M : [n]^d \to \{0, 1\}$, which satisfies $M(i_1, i_2, i_3, \dots, i_d) = 1$ if and only if $\sum_{j=1}^{d} i_j$ is divisible by n, has the following properties.

Lemma 9 *[AFT11]* $rk(M) \leq dn$.

Lemma 10 *[AFT11]* $mr(M) = n^{d-1}$.

2.3 Hadamard Product

Definition 5 *Let A and B be $m \times n$ matrices with entries in $\mathbb{R}$. The* Hadamard product *of A and B is defined by $[A \circ B]_{ij} = [A]_{ij}[B]_{ij}$, for all $1 \leq i \leq m, 1 \leq j \leq n$.*

We need a folklore property of Hadamard product.

Proposition 11 *Let A and B be $m \times n$ real matrices, then $rk(A \circ B) \leq rk(A)rk(B)$.*

Proof: Just note that $A \circ B$ is a submatrix of $A \otimes B$. $\qquad\square$

3 Monotone vs. Non-monotone Computation

3.1 Remarks

The proof is under the non-commutative model. The method and result do not hold in a commutative setting. In the commutative model the separation between monotone and non-monotone is exponential as shown by Valiant [Val80] while in our case the separation is super-exponential.

3.2 Proof of Theorem 1

Now we prove Theorem 1 by defining a function f with ABP complexity $O(d^2)$ and monotone ABP complexity $\Omega(d^2 \log n)$. A homogeneous function f of degree d (d even) on n variables is in the form of

$$f(x_1, x_2, \dots, x_n) = \sum_{i_1, i_2, \dots, i_d \in [n]} c_{i_1 i_2 \cdots i_d} x_{i_1} x_{i_2} \cdots x_{i_d}.$$

There are totally n^d monomials and n^d corresponding coefficients. We define a function $g : \{i_1, i_2, \dots, i_{d/2} : i_1, i_2, \dots, i_{d/2} \in [n]\} \to [n^{d/2}]$ as follows.

$$g(i_1, i_2, \dots, i_{d/2}) = \sum_{k=1}^{d/2} (i_k - 1) n^{d/2 - k} + 1.$$

It is easy to verify that g is a bijective function. For instances, $g(1, 1, \dots, 1, 1) = 1$, $g(1, 1, \dots, 1, 2) = 2$, $g(1, 1, \dots, 1, 3) = 3$, ..., $g(n, n, \dots, n, n - 1) = n^{d/2} - 1$, $g(n, n, \dots, n, n) = n^{d/2}$.

Then we define f by specifying all its coefficients.

$$c_{i_1 i_2 \cdots i_d} = (g(i_1, i_2, \dots, i_{d/2}) - g(i_{d/2+1}, i_{d/2+2}, \dots, i_d))^2.$$

It is clear that all the coefficients are nonnegative. In a word, f is the following.

$$f(x_1, x_2, \dots, x_n) = \sum_{i_1, i_2, \dots, i_d \in [n]} (g(i_1, i_2, \dots, i_{d/2}) - g(i_{d/2+1}, i_{d/2+2}, \dots, i_d))^2 x_{i_1} x_{i_2} \cdots x_{i_d}.$$

6

We arrange $M_{d/2}(f)$ in a way such that k-term is in the ascending order of its corresponding $g(i_1, i_2, \ldots, i_{d/2})$, and $d-k$-term is in the ascending order of its corresponding $g(i_{d/2+1}, i_{d/2+2}, \ldots, i_d)$. Then for $M_{d/2}(f)$, it is not hard to verify that $[M_{d/2}(f)]_{ij} = (j-i)^2, \forall i, j \in [n^{d/2}]$. More explicitly,

$$M_{d/2}(f) = \begin{bmatrix} 0^2 & 1^2 & 2^2 & \cdots & (n^{d/2}-1)^2 \\ 1^2 & 0^2 & 1^2 & \cdots & (n^{d/2}-2)^2 \\ 2^2 & 1^2 & 0^2 & \cdots & (n^{d/2}-3)^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ (n^{d/2}-1)^2 & (n^{d/2}-2)^2 & (n^{d/2}-3)^2 & \cdots & 0 \end{bmatrix}.$$

If we use R_i to represent the i-th row of $M_{d/2}(f)$, $M_{d/2}(f)$ could be written into another way.

$$M_{d/2}(f) = \begin{bmatrix} R_1 \\ R_2 \\ R_3 \\ \vdots \\ R_{n^{d/2}} \end{bmatrix}.$$

More generally, $\forall k \in [d/2]$,

$$M_{d/2-k}(f) = \begin{bmatrix} R_1 & R_2 & R_3 & \cdots & R_{n^k} \\ R_{n^k+1} & R_{n^k+2} & R_{n^k+3} & \cdots & R_{2n^k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ R_{n^{d/2}-n^k+1} & R_{n^{d/2}-n^k+2} & R_{n^{d/2}-n^k+3} & \cdots & R_{n^{d/2}} \end{bmatrix}.$$

This can be easily verified by definition. For example, when $k = 1$,

$$[M_{d/2-1}(f)]_{(x_{i_1}\cdots x_{i_{d/2-1}}, x_{i_{d/2}}x_{j_1}\cdots x_{j_{d/2}})} = [M_{d/2}(f)]_{(x_{i_1}\cdots x_{i_{d/2}}, x_{j_1}\cdots x_{j_{d/2}})}, \tag{2}$$

and thus,

$$M_{d/2-1}(f) = \begin{bmatrix} R_1 & R_2 & R_3 & \cdots & R_n \\ R_{n+1} & R_{n+2} & R_{n+3} & \cdots & R_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ R_{n^{d/2}-n+1} & R_{n^{d/2}-n+2} & R_{n^{d/2}-n+3} & \cdots & R_{n^{d/2}} \end{bmatrix}.$$

Next we will show the following two lemmas, the combination of which will immediately yield the result in Theorem 1.

Lemma 12 $B(f) = O(d^2)$.

Proof: By Lemma 6, $rk(M_{d/2}(f)) = 3$. We define the following subsidiary matrix S of dimension $n^{d/2-1} \times (n-1)$.

$$S = \begin{bmatrix} 1^2 & 2^2 & 3^2 & \cdots & (n-1)^2 \\ (1+n)^2 & (2+n)^2 & (3+n)^2 & \cdots & (2n-1)^2 \\ (1+2n)^2 & (2+2n)^2 & (3+2n)^2 & \cdots & (3n-1)^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ (n^{d/2}-n+1)^2 & (n^{d/2}-n+2)^2 & (n^{d/2}-n+3)^2 & \cdots & (n^{d/2}-1)^2 \end{bmatrix}$$

A careful comparison of $M_{d/2}(f)$ and $M_{d/2-1}(f)$ would reveal the following observation,

Observation 1 $rk(M_{d/2-1}(f)) \leq rk(M_{d/2}(f)) + rk(S)$.

The correctness of this observation can be very easily verified. Here is a very simple example. Let $d = 4$ and $n = 2$, then

$$M_{d/2}(f) = \begin{bmatrix} 0^2 & 1^2 & 2^2 & 3^2 \\ 1^2 & 0^2 & 1^2 & 2^2 \\ 2^2 & 1^2 & 0^2 & 1^2 \\ 3^2 & 2^2 & 1^2 & 0^2 \end{bmatrix},$$

$$M_{d/2-1}(f) = \begin{bmatrix} 0^2 & 1^2 & 2^2 & 3^2 & 1^2 & 0^2 & 1^2 & 2^2 \\ 2^2 & 1^2 & 0^2 & 1^2 & 3^2 & 2^2 & 1^2 & 0^2 \end{bmatrix},$$

and

$$S = \begin{bmatrix} 1^2 \\ 3^2 \end{bmatrix}.$$

It is not hard to see that the first four columns of $M_{d/2-1}(f)$ is a submatrix of $M_{d/2}(f)$. The last four columns of $M_{d/2-1}(f)$ differ from the first four columns of $M_{d/2-1}(f)$ just by one column, which is S. Thus, we know that $rk(M_{d/2-1}(f)) \leq rk(M_{d/2}(f)) + rk(S)$ for this example.

We define another auxiliary matrix S_1.

$$S_1 = \begin{bmatrix} 1 & 2 & 3 & \cdots & (n-1) \\ (1+n) & (2+n) & (3+n) & \cdots & (2n-1) \\ (1+2n) & (2+2n) & (3+2n) & \cdots & (3n-1) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ (n^{d/2}-n+1) & (n^{d/2}-n+2) & (n^{d/2}-n+3) & \cdots & (n^{d/2}-1) \end{bmatrix}$$

It is easy to see that $rk(S_1) = 2$ (Gaussian elimination), and that $S = S_1 \circ S_1$. According to Proposition 11, $rk(S) \leq (rk(S_1))^2 = 4$. Now we have

$$rk(M_{d/2-1})(f) \leq rk(M_{d/2}(f)) + 4.$$

In a similar way and by a straightforward generalization, we know that

$$\forall k \in [d/2], rk(M_{d/2-k}(f)) \leq rk(M_{d/2-k+1}(f)) + 4.$$

So we know that $\forall k \in [d/2]$,

$$rk(M_{d/2-k}(f)) \leq 3 + 4k.$$

Noting that $M_{d/2+k}(f)$ and $M_{d/2-k}(f)$ are symmetric,

$$\forall k \in [d/2], rk(M_{d/2+k}(f)) = rk(M_{d/2-k}(f)) \leq 3 + 4k.$$

Therefore, according to Lemma 4,

$$B(f) = \sum_{k=0}^{d} rk(M_k(f)) = O(d^2).$$

$\square$

Lemma 13 $B^+(f) = \Omega(d^2 \log n)$.

Proof: By Lemma 7, we know that $mr(M_{d/2}(f)) \geq \frac{d}{2} \log n$. For $M_{d/2-1}(f)$, it is not hard to see that after some permutation of the columns, we can obtain a sub-matrix with dimension $n^{d/2-1} \times n^{d/2-1}$ from $M_{d/2-1}(f)$ as follows:

$$
\begin{bmatrix}
0^2 & n^2 & (2n)^2 & \cdots & (n^{d/2}-n)^2 \\
n^2 & 0^2 & n^2 & \cdots & (n^{d/2}-2n)^2 \\
(2n)^2 & n^2 & 0^2 & \cdots & (n^{d/2}-3n)^2 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
(n^{d/2}-n)^2 & (n^{d/2}-2n)^2 & (n^{d/2}-3n)^2 & \cdots & 0^2
\end{bmatrix}.
$$

So $mr(M_{d/2-1}(f)) \geq (d/2-1)(\log n)$ by Lemma 7. Similarly, we can show that $\forall k \in [d/2]$,

$$mr(M_{d/2-k}(f)) \geq (d/2-k)(\log n).$$

Noting the fact that $M_{d/2-k}(f)$ and $M_{d/2+k}(f)$ are symmetric, we know that $\forall k \in [d/2]$,

$$mr(M_{d/2+k}(f)) = mr(M_{d/2-k}(f)) \geq (d/2-k)(\log n).$$

By Lemma 5,

$$B^+(f) = \Omega(\sum_{k=0}^{d} B_k^+(f)) = \Omega(d^2 \log n).$$

$\square$

4   Shared Randomness vs. Quantum Entanglement

4.1   Proof of Theorem 2

In this part, we shall prove Theorem 2. Remember that U and V are unitary matrices of dimension $2^n \times 2^n$, and that M is a measurement with respect to the standard basis. We set $|\psi\rangle$ to be $\frac{|00\rangle+|11\rangle}{\sqrt{2}}$.

First we calculate that $\phi_0 = \frac{1}{\sqrt{2}}(|0^{2n}\rangle + |0^{n-1}\rangle|11\rangle|0^{n-1}\rangle)$ and $\phi_1 = \frac{1}{\sqrt{2}}(u_0 \otimes v_0 + u_1 \otimes v_1)$, where u_0 is the first column of U, v_0 is the first column of V, u_1 is the second column of U, and v_1 is the $(2^{n-1}+1)$-th column of V. After the measurement M, there are 4^n possibilities. And $Prob\{X = x, Y = y\} = \frac{1}{2}|u_0(x)v_0(y) + u_1(x)v_1(y)|^2$, for all $x, y \in \{0,1\}^n$. Here we use x and y as the index for vector or matrix. It is clear that the unitary operation U and V can decide the distribution of the measurement outcome. Suppose $N = 2^n$. We use a nonnegative matrix P of dimension $N \times N$ to demonstrate the distribution of (X,Y), $P = [Prob\{X = x, Y = y\}]_{xy}$. Suppose we want to use shared randomness to simulate this distribution generated from quantum entanglement. In the beginning Alice and Bob share a random variable Z, whose sample space is Ω. We would show two lemmas.

Lemma 14 $|\Omega| \geq mr(P)$.

Lemma 15 *There exists U and V such that $mr(P) \geq \log N$.*

From these lemmas, it is clear that $\log |\Omega| \geq \log n$, implying that we need at least $\log n$ bits of shared randomness to simulate the correlation.

$\square$

4.2   Proof of Lemma 14

We observe that conditional on Z, X and Y are independent. That is to say,

$$Prob\{X = x, Y = y\} = \sum_{z \in \Omega} Prob\{Z = z\} \times Prob\{X = x, Y = y | Z = z\}$$

$$= \sum_{z \in \Omega} Prob\{Z = z\} \times Prob\{X = x | Z = z\} \times Prob\{Y = y | Z = z\}.$$

For a fixed z, let v_z be the vector of size 2^n, such that $v_z(x) = Prob\{X = x | Z = z\}$, and v'_z be the vector of size 2^n, such that $v'_z(y) = Prob\{Y = y | Z = z\}$. So,

$$P = \sum_{z \in \Omega} Prob\{Z = z\}(v_z)(v'_z)^T,$$

which means that P can be decomposed into $|\Omega|$ nonnegative rank-1 matrices. By the definition of $mr(P)$, $|\Omega| \geq mr(P)$.

$\square$

4.3   Proof of Lemma 15

The aim of this lemma is to find a hard distance (the distribution of a quantum correlation) P, based on some plausible U and V, such that the monotone rank of P is high. Next we will show an explicit P and show why its monotone rank is high.

Let $\{c_x : x \in \{0,1\}^n\}$ be a set of $N = 2^n$ distinct elements of $\mathbb{R}^+$. and define matrix C to be $C_{xy} = c_y - c_x, x, y \in \{0,1\}^n$. Thus the Hadamard product of C and its conjugate matrix is $C \circ \bar{C} = [(c_y - c_x)^2]_{xy}$. Using Guassian elimination, we know that $rk(C) = 2$. Since C is an antisymmetric matrix, the eigenvalues of C are λ, $-\lambda$ and $N - 2$ 0's. The characteristic polynomial of C is

$$\sum_{k \in \{0,1...,N\}} e_k \lambda^k,$$

where e_k is the coefficient of λ^k. It is easy to see that $e_N = 1, e_{N-1} = 0$ and

$$e_{N-2} = \sum_{1 \leq x < y \leq N} (c_y - c_x)^2.$$

For $k \leq N - 3$,

$$e_k = \sum_{I \subseteq \{0,1\}^n : |I| = N-k} |C_I|,$$

where C_I is the submatrix obtained by restricting C on those rows and columns in I, and $|C_I|$ stands for the determinant of C_I. Since $rk(C) = 2$, $rk(C_I) \leq rk(C) = 2$, so $|C_I| = 0$. Consequently, $e_k = 0, \forall k \leq N - 3$. Hence, the characteristic polynomial of C is

$$\lambda^N + \sum_{1 \leq x < y \leq N} (c_y - c_x)^2 \lambda^{N-2},$$

and

$$\lambda = i \sqrt{\sum_{1 \leq x < y \leq N} (c_y - c_x)^2}.$$

10

Since C is antisymmetric, C is normal. Using spectral decomposition, we know that

$$C = \lambda |u_0\rangle\langle u_0| - \lambda |u_1\rangle\langle u_1|,$$

where $|u_0\rangle$ is the eigenvector of λ and $|u_1\rangle$ is the eigenvector of $-\lambda$. It is easy to take proper distinct values of $\{c_x : x \in \{0,1\}^n\}$ to satisfy $\sqrt{\sum_{1 \leq x < y \leq n}(c_y - c_x)^2} = \sqrt{1/2}$, which implies $\lambda = i\sqrt{1/2}$. Let $v_0 = \bar{u}_0$ and $v_1 = -\bar{u}_1$ and we get

$$
\begin{aligned}
P_{xy} &= \frac{1}{2}|u_0(x)v_0(y) + u_1(x)v_1(y)|^2 \\
&= \frac{1}{2}|u_0(x)\bar{u}_0(y) - u_1(x)\bar{u}_1(y)|^2.
\end{aligned}
$$

Also, we have

$$
\begin{aligned}
(C \circ \bar{C})_{xy} &= (\lambda u_0(x)\overline{u_0(y)} - \lambda u_1(x)\overline{u_1(y)})\overline{(\lambda u_0(x)\overline{u_0(y)} - \lambda u_1(x)\overline{u_1(y)})} \\
&= \frac{1}{2}|u_0(x)\bar{u}_0(y) - u_1(x)\bar{u}_1(y)|^2.
\end{aligned}
$$

Therefore, $P = C \circ \bar{C}$, which means that we are able to construct P by selecting proper values for entries of C, where C also determines some columns of U and V. By Lemma 7, $mr(P) = mr(C \circ \bar{C}) \geq \log_2(N)$. $\qquad\square$

5 Generalized Log-Rank Conjecture

5.1 Discussions

Our communication complexity separation does not hold for randomized communication.

5.2 Proof of Theorem 3

A function $f : \{0,1\}^n \to \{0,1\}$ could be written into an equivalent form $f : [N] \to \{0,1\}$ with $N = 2^n$. For convenience, we will use the latter representation. We define a function $f : [N]^d \to \{0,1\}$ by requiring $f(i_1, i_2, \ldots, i_d) = 1$ if and only if $\sum_{j=1}^d i_j$ is divisible by N. By Lemma 9, $rk(M(f)) \leq dN$. By Lemma 10, $mr(M(f)) = N^{d-1}$. Therefore, $\log rk(M(f)) \leq \log d + n$, and $\log mr(M(f)) = (d-1)n$. Suppose $d = \omega(n^{c'}), \forall c' > 0$. Now it is easy to see that for any constant $c > 0$, $\log mr(M(f)) = \omega(\log^c rk(M(f)))$, and because of an obvious relation $\log(rk(M(f))) \leq \log(mr(M(f))) \leq D(f)$, we know that for any constant $c > 0$, $D(f) = \omega(\log^c rk(M(f)))$.

$\qquad\square$

6 Acknowledgment

References

[AB09]   Sanjeev Arora and Boaz Barak. *Computational Complexity: A Modern Approach.* Cambridge University Press, 2009.

[AFT11]   Boris Alexeev, Michael Forbes, and Jacob Tsimerman. Tensor rank: Some lower and upper bounds. *Proceedings of the 26th Annual Conference on Computational Complexity*, pages 283 – 291, 2011.

[BB84]   Charles Bennett and Gilles Brassard. Quantum cryptography: Public key distribution and coin tossing. *Proceedings of the IEEE International Conference on Computers, Systems, and Signal Processing*, pages 175–179, 1984.

[BBC$^+$93] Charles Bennett, Gilles Brassard, Claude Crepeau, Richard Jozsa, Ashes Peres, and William Wootters. Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels. *Physical Review Letters*, 70:1895–1899, 1993.

[BCT99]   Gilles Brassard, Richard Cleve, and Alain Tapp. The cost of exactly simulating quantum entanglement with classical communication. *Physical Review Letters*, 83(9):1874–1877, 1999.

[BCvD01]   Harry Buhrman, Richard Cleve, and Wim van Dam. Quantum engtanglement and communication complexity. *SIAM Journal on Computing*, 30(6):1829–1841, 2001.

[Bel64]   John Bell. On the Einstein-Podolsky-Rosen paradox. *Physics*, 1:195–200, 1964.

[BL09]   Leroy Beasley and Thomas Laffey. Real rank versus nonnegative rank. *Linear Algebra and its Applications*, 431(12):2330–2335, 2009.

[BT03]   Dave Bacon and Ben Tonor. Bell inequalities with auxiliary communication. *Physical Review Letters*, 90:157904, 2003.

[BW92]   Charles Bennett and Stephen Wiesner. Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states. *Physical Review Letters*, 69(20):2881–2884, 1992.

[CHSH69]   John Clauser, Michael Horne, Abner Shimony, and Richard Holt. Proposed experiment to test local hidden-variable theories. *Physical Review Letters*, 23(15):880–884, 1969.

[DKW09]   Jan Draisma, Eyal Kushilevitz, and Enav Weinreb. Partition arguments in multiparty communication complexity. *Proceedings of the 36th International Colloquium on Automata, Languages and Programming*, pages 390–402, 2009.

[Fey82]   Richard Feynman. Simulating physics with computers. *International Journal of Theoretical Physics*, 21:467–488, 1982.

[Has90]   Johan Hastad. Tensor rank is NP-complete. *Journal of Algorithms*, 11(4):644–654, 1990.

[Hru11]   Pavel Hrubes. On the nonnegative rank of distance matrices. *submitted to Information Processing Letters*, 2011.

[HY11]   Pavel Hrubes and Amir Yehudayoff. Homogeneous formulas and symmetric polynomials. *Computational Complexity*, pages 559–578, 2011.

[LC10]   Matthew Lin and Moody Chu. On the nonnegative rank of Euclidean distance matrices. *Linear Algebra and its Applications*, 433(3):681–689, 2010.

[Li11]   Yang Li. Monotone rank and separations in computational complexity. *ECCC, TR-025*, 2011.

[LS88]   Laszlo Lovasz and Michael Saks. Lattices, mobius functions and communications complexity. *Proceedings of the 29th Annual IEEE Symposium on Foundations of Computer Science*, pages 81–90, 1988.

[LS09]   Troy Lee and Adi Shraibman. Lower bounds in communication complexity. *Foundations and Trends in Theoretical Computer Science*, 3(4):263–399, 2009.

[Nis91]   Noam Nisan. Lower bounds for non-commutative computation. *Proceedings of the 23rd Annual ACM Symposium on Theory of Computing*, pages 410–418, 1991.

[NW95]   Noam Nisan and Avi Wigderson. On rank vs. communication complexity. *Combinatorica*, 15(4):557–565, 1995.

[Raz10]   Ran Raz. Tensor-rank and lower bounds for arithmetic formulas. *Proceedings of the 42th Annual ACM Symposium on Theory of Computing*, pages 659–666, 2010.

[RS95]   Ran Raz and Boris Spieker. On the 'log-rank' conjecture in communication complexity. *Combinatorica*, 15(4):567–588, 1995.

[RT09]   Oded Regev and Ben Tonor. Simulating quantum correlations with finite communication. *SIAM Journal on Computing*, 39(4):1562–1580, 2009.

[TB03]   Ben Tonor and Dave Bacon. Communication cost of simulating Bell correlations. *Physical Review Letters*, 91:187904, 2003.

[Val79]   Leslie Valiant. The complexity of computing the permanent. *Theoretical Computer Science*, 8:189–201, 1979.

[Val80]   Leslie Valiant. Negation can be exponentially powerful. *Theoretical Computer Science*, 12:303–314, 1980.

[Vav09]   Stephen Vavasis. On the complexity of nonnegative matrix factorization. *SIAM Journal on Optimization*, 20(3):1364–1377, 2009.

[Win05]   Andreas Winter. Secret, public and quantum correlation cost of triples of random variables. *Proceeding of ISIT*, pages 2270–2274, 2005.

[Yan91]   Mihalis Yannakakis. Expressing combinatorial optimization problems by linear programs. *Journal of Computer and System Sciences*, 43(3):441–466, 1991.

[Zha12]   Shengyu Zhang. Quantum strategic game theory. *The 3rd Innovations in Theoretical Computer Science*, page to appear, 2012.