Electronic Colloquium on Computational Complexity, Report No. 181 (2016)

The Bipartite Formula Complexity of Inner-Product is Quadratic

Avishay Tal *

Abstract

A bipartite formula on binary variables $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$ is a binary tree whose internal nodes are marked with AND or OR gates and whose leaves can compute any function of either the x or y variables. We show that any bipartite formula for the Inner-Product modulo 2 function, namely $\mathrm{IP}_n(x,y) = \sum_{i=1}^n x_i y_i \mod 2$, must be of size $\widetilde{\Omega}(n^2)$. The result is tight up to logarithmic factors. To the best of our knowledge, this is the first super-linear lower bound on the bipartite formula complexity of any explicit function.

We give two simple proofs that rely on the deep results of Reichardt [Rei11] and Forster [For02]. Moreover, the second proof establishes an average-case lower bound for the Inner-Product function. Namely, we show that any bipartite formula that agrees with IP_n on at least $\frac{1}{2} + \frac{1}{n^{\log n}}$ of the inputs must be of size $\widetilde{\Omega}(n^2)$.

1 Introduction

A bipartite de Morgan formula on variables $x_1, \ldots, x_n, y_1, \ldots, y_n$ is a binary tree whose internal nodes are marked with AND or OR gates, and such that each leaf computes an arbitrary function in either $(x_1, \ldots, x_n)$ or $(y_1, \ldots, y_n)$.

This is a generalization of the well-know *standard* de Morgan formula, defined next. A de Morgan formula on variables $x_1, \ldots, x_n, y_1, \ldots, y_n$ is a binary tree whose internal nodes are marked with AND or OR gates, and such that each leaf is marked with either a variable or its negation.

In both cases, the formula size is the number of leaves in the tree. For a given function $f(x,y)$ on $2n$ variables we define the bipartite formula complexity of f, denoted $L_{\mathrm{bip}}(f)$, to be the size of the smallest bipartite de Morgan formula on $x_1, \ldots, x_n, y_1, \ldots, y_n$ that computes f. Similarly, the formula complexity of f, denoted $L(f)$, is the size of the smallest de Morgan formula on $x_1, \ldots, x_n, y_1, \ldots, y_n$ that computes f. Clearly for any $f(x,y)$ we have

$$L_{\mathrm{bip}}(f) \leq L(f),$$

however for some functions $L_{\mathrm{bip}}(f) \ll L(f)$. For example, for the parity function given by $f(x,y) = \sum_{i=1}^n x_i + \sum_{i=1}^n y_i \mod 2$, we have $L(f) = \Theta(n^2)$ and $L_{\mathrm{bip}}(f) \leq 4$.

Bipartite formula complexity is sometimes referred to as graph complexity. See [Lok09, Chapter 6] and [Juk12, Chapters 1,6] for more on bipartite formula complexity and graph complexity.

We show that the Inner-Product function, $\mathrm{IP}_n(x,y) = \sum_{i=1}^n x_i y_i \mod 2$, have $L_{\mathrm{bip}}(\mathrm{IP}_n) \geq \Omega(n^2/\log^2 n)$. This is tight up to the logarithmic factors, as $L_{\mathrm{bip}}(\mathrm{IP}_n) \leq L(\mathrm{IP}_n) \leq O(n^2)$ by a straight forward de Morgan formula for IP_n. To the best of our knowledge, our lower bound is the first super-linear lower bound on the bipartite formula complexity of any explicit function. This (essentially) solves an open problem raised in [Juk12, Chapter 6] (more precisely, Research Problem

*Institute for Advanced Study, Princeton, NJ. Email: avishay.tal@gmail.com. Supported by the Simons Collaboration on Algorithms and Geometry, and by the National Science Foundation grant No. CCF-1412958.

ISSN 1433-8092

6.57 in [Juk12] asks to give better than linear lower bounds on a measure $\chi(f)$, that lower bounds $L_{\text{bip}}(f)$. We lower bound $L_{\text{bip}}(f)$ directly.)

We remark that it is a long-standing open question to find better than n^3 lower bounds for $L(f)$ for any explicit function f (the state of the art is a $\widetilde{\Omega}(n^3)$ lower bound for Andreev's function [And87], given in [Hås98, Tal14]). Thus, any better than n^3 lower bounds for the bipartite formula complexity of explicit functions would be extremely interesting.

Organization. In Section 2 we prove that any matrix associated with small bipartite formula has small sign-rank. In section 3 we deduce that $L_{\text{bip}}(\text{IP}_n) = \widetilde{\Omega}(n^2)$. In section 4, we prove that even if a function is ε-correlated with a small bipartite formula, then the matrix associated with the *function* has high discrepancy. This allows us to get $\widetilde{\Omega}(n^2)$ lower bounds for IP_n also in the average-case, as done in Section 5.

2 Sign Rank of Matrices Associated with Small Bipartite Formulas

Throughout the paper, we will treat the Boolean domain as $\{-1, 1\}$. Our result heavily relies on the following result from quantum query complexity by Reichardt [Rei11] (completing a long line of work [BBC+01, LLS06, HLS07, FGG08, Rei09, ACR+10, RS12]). The fact that the proof of this result involves quantum algorithms seems quite surprising, as the statement is completely classical. Indeed, this can be seen as an example of the quantum method, giving a quantum based proof for a classical theorem.

Theorem 2.1 ([Rei11])**.** *Let F be a de Morgan formula of size s, computing a Boolean function f. Then, $\widetilde{\deg}(f) \leq O(\sqrt{s})$. That is, there is a multi-linear polynomial p of degree $O(\sqrt{s})$ over the reals, such that for any $x \in \{-1, 1\}^n$, the value $p(x)$ is in the range $[F(x) - 1/3, F(x) + 1/3]$.*

Next, we define the sign-rank of a matrix.

Definition 2.2 (Sign-Rank)**.** *Let A by an n-by-m matrix with ± 1 entries. The* sign-rank *of A, denoted $\text{rank}_{\pm}(A)$, is the least rank of a matrix $B \in \mathbb{R}^{n \times m}$ such that $B_{i,j} \neq 0$ and $\text{sgn}(B_{i,j}) = A_{i,j}$ for all $(i, j) \in [n] \times [m]$.*

We state the main lemma of this section.

Lemma 2.3 (The Sign-Rank of matrices associated with Bipartite de Morgan formula)**.** *Let F be a bipartite de Morgan formula of size s. Then F has sign-rank at most $s^{O(\sqrt{s})}$.*

Proof. F is a formula where each input gate is either a function on $x_1, \ldots, x_n$ or a function on $y_1, \ldots, y_n$. In particular, for $i \in [s]$, we can write the function computed by the i-th input gate as the product of some function $f_i(x)$ and a function $g_i(y)$ (where one of (f_i, g_i) is the constant 1 function). Denote by $h_i(x, y) := f_i(x) \cdot g_i(y)$.

Let F' be a read-once de Morgan formula obtained from F by replacing the i-th input gate with a formal variable z_i, for all $i \in [s]$. Note that F' is a *standard* de Morgan formula on $z_1, \ldots, z_s$. Apply Theorem 2.1 to F' to get a degree $d = O(\sqrt{s})$ polynomial $p(z)$ such that for any $z \in \{-1, 1\}^s$, $|p(z) - F'(z)| \leq 1/3$. In particular, $\text{sgn}(p(z)) = F'(z)$ for any $z \in \{-1, 1\}^s$. We write $p(z)$ as a multi-linear polynomial:

$$p(z) = \sum_{S \subseteq [s] : |S| \leq d} \hat{p}(S) \cdot \prod_{i \in S} z_i$$

Note that there are at most $s^{O(\sqrt{s})}$ monomials in p. We replace each z_i with $g_i(x) \cdot f_i(y)$ and get that for any $x \in \{-1,1\}^n$ and $y \in \{-1,1\}^n$

$$F(x,y) = \operatorname{sgn}\left(p(f_1(x)\cdot g_1(y), \ldots, f_s(x)\cdot g_s(y))\right) = \operatorname{sgn}\left(\sum_{S\subseteq[s]:|S|\leq d} \hat{p}(S)\cdot\left(\prod_{i\in S} f_i(x)\right)\cdot\left(\prod_{i\in S} g_i(y)\right)\right).$$

Observe that each summand in the RHS, $\hat{p}(S)\cdot\left(\prod_{i\in S} f_i(x)\right)\cdot\left(\prod_{i\in S} g_i(y)\right)$, corresponds to a matrix of rank-1 over the reals, hence their sum corresponds to a matrix of rank at most at most $s^{O(\sqrt{s})}$ over the reals. Overall, we got that the matrix M_F has sign-rank at most $s^{O(\sqrt{s})}$. □

3 Lower Bounds for The Inner-Product Function

The Inner-Product modulo 2 function is usually defined with the $\{0,1\}$ realization of the Boolean domain as $\mathrm{IP_n}(x,y) = \sum_{i=1}^n x_i y_i \mod 2$. In this paper however, we realize the Boolean domain as $\{-1,1\}$, and in this notation, $\mathrm{IP_n} : \{-1,1\}^n \times \{-1,1\}^n \to \{-1,1\}$ is defined as $\mathrm{IP_n}(x,y) = (-1)^{\sum_{i=1}^n (1-x_i)\cdot(1-y_i)/4}$. The 2^n-by-2^n matrix associated with the Inner-Product function, $M_{\mathrm{IP_n}}$, is an Hadamard matrix (which means that $M_{\mathrm{IP_n}} \cdot M_{\mathrm{IP_n}} = 2^n \cdot I$). In his celebrated work, Forster [For02] gave exponential lower bounds on the sign-rank of Hadamard matrices and $M_{\mathrm{IP_n}}$ in particular.

Theorem 3.1 ([For02])**.** *The sign-rank of $M_{\mathrm{IP_n}}$ is at least $2^{n/2}$.*

From Forster's Theorem and Lemma 2.3, we deduce our main theorem.

Theorem 3.2 (Main Theorem)**.** *The Bipartite Formula Complexity of* $\mathrm{IP_n}$ *is* $\Omega(n^2/\log^2 n)$.

Proof. Let F be the smallest bipartite formula computing $\mathrm{IP_n}$, and denote its size by s. By Forster's Theorem (Thm. 3.1) and Lemma 2.3, we get

$$2^{n/2} \leq \mathrm{rank}_{\pm}(M_{\mathrm{IP_n}}) \leq s^{O(\sqrt{s})},$$

hence $s = \Omega(n^2/\log^2 n)$. □

4 The Discrepancy of Matrices Associated with Small Bipartite Formulas

The discrepancy of a matrix (defined next) is a well-studied complexity measure in the field of communication complexity, where upper bounds on the discrepancy of a matrix yield lower bounds on the randomized communication complexity of the two-party communication problem associated with this matrix.

Definition 4.1. *Let A by an n-by-m matrix with ± 1 entries. The discrepancy of A is defined by*

$$\mathrm{disc}(A) = \frac{1}{nm} \cdot \max_{I\subseteq[n], J\subseteq[m]} \left|\sum_{i\in I}\sum_{j\in J} A_{i,j}\right|$$

Next, we state the main lemma of this section.

Lemma 4.2. *Let $f(x_1,\ldots,x_n,y_1,\ldots,y_n)$ be a Boolean function. Let $F(x_1,\ldots,x_n,y_1,\ldots,y_n)$ be a bipartite de Morgan formula of size s such that*

$$\mathop{\mathbf{Pr}}_{x\in_R\{-1,1\}^n,y\in_R\{-1,1\}^n}[F(x,y)=f(x,y)]\geq 1/2+\varepsilon.$$

Let M_f be the $2^n\times 2^n$ matrix defined by $(M_f)_{x,y}=f(x,y)$. Then, $\mathrm{disc}(M_f)\geq 1/s^{O(\sqrt{s}\cdot\log(1/\varepsilon))}$.

Before proving the main lemma of this section, we define and discuss the approximate degree of Boolean functions.

Definition 4.3 (Approximate Degree)**.** *Let $f:\{-1,1\}^n\to\{-1,1\}$. Given $\varepsilon\geq 0$, we define the ε-approximate degree, denoted by $\widetilde{\deg}_\varepsilon(f)$, as the minimal degree of a multi-linear polynomial p such that for all $x\in\{-1,1\}^n$, $|f(x)-p(x)|\leq\varepsilon$. We denote $\widetilde{\deg}_{1/3}(f)$ by $\widetilde{\deg}(f)$.*

When defining the approximate degree the choice of $1/3$ may seem arbitrary. The next fact ([BNRdW07, Lemma 1], see also [Tal14, Appendix B]) shows how approximate degrees for different error parameters relate. We would like to point out that we are not sure if Fact 4.4 is tight.

Fact 4.4. *Let f be a Boolean function and $0<\epsilon<1$. Then, $\widetilde{\deg}_\epsilon(f)\leq O(\widetilde{\deg}(f)\cdot\log(1/\varepsilon))$.*

In this notation, Reichardt's result [Rei11], Theorem 2.1, states that $\widetilde{\deg}(f)\leq O(\sqrt{L(f)})$. Combining this with Fact 4.4 gives $\widetilde{\deg}_\varepsilon(f)\leq O(\sqrt{L(f)}\cdot\log(1/\varepsilon))$. Equipped with this result, we are ready to prove Lemma 4.2.

Proof of lemma 4.2. Without loss of generality we assume $\varepsilon<1/3$. F is a formula where each input gate is either a function on $x_1,\ldots,x_n$ or a function on $y_1,\ldots,y_n$. In particular, for $i\in[s]$, we can write the function computed by the i-th input gate as the product of some function $f_i(x)$ and a function $g_i(y)$ (where one of f_i,g_i is the constant 1 function). Denote by $h_i(x,y):=f_i(x)\cdot g_i(y)$.

Let F' be a read-once de Morgan formula obtained from F by replacing the i-th input gate with a formal variable z_i, for all $i\in[s]$. Apply Theorem 2.1 and Fact 4.4 on F' to get a degree $d=O(\sqrt{s}\cdot\log(1/\varepsilon))$ polynomial $p(z)$ such that for any $z\in\{-1,1\}^s$, $|p(z)-F'(z)|\leq\varepsilon$. In particular, $p(z)\cdot F'(z)\in[1-\varepsilon,1+\varepsilon]$ for any $z\in\{-1,1\}^s$.

Next, we show that under the uniform distribution, the function $\tilde{p}(x,y)\triangleq p(h_1(x,y),\ldots,h_s(x,y))$ correlates with the function $f(x,y)$ that $F(x,y)$ approximates. Let $q:=\mathbf{Pr}[f(x,y)=F(x,y)]\geq 1/2+\varepsilon$. We break the expectation $\mathbf{E}[\tilde{p}(x,y)\cdot f(x,y)]$ according to whether or not $f(x,y)=F(x,y)$:

$$\begin{aligned}\mathop{\mathbf{E}}_{x,y}[\tilde{p}(x,y)\cdot f(x,y)] &= \mathop{\mathbf{E}}_{x,y}[\tilde{p}(x,y)\cdot f(x,y)|f(x,y)=F(x,y)]\cdot\mathbf{Pr}[f(x,y)=F(x,y)]\\ &+\mathop{\mathbf{E}}_{x,y}[\tilde{p}(x,y)\cdot f(x,y)|f(x,y)\neq F(x,y)]\cdot\mathbf{Pr}[f(x,y)\neq F(x,y)]\\ &\geq(1-\varepsilon)\cdot q+(-1-\varepsilon)\cdot(1-q)\\ &=2q-1-\varepsilon\geq 2\cdot(1/2+\varepsilon)-1-\varepsilon=\varepsilon\,.\end{aligned}$$

Next, we write $p(z)$ as a multi-linear polynomial:

$$p(z)=\sum_{S\subseteq[s]:|S|\leq d}\hat{p}(S)\cdot\prod_{i\in S}z_i\,.$$

Since $\hat{p}(S) = \mathbf{E}_{z\in\{-1,1\}^s}[p(z)\cdot\prod_{i\in S} z_i]$ we have that $|\hat{p}(S)| \le 1+\varepsilon$. Note that there are at most $s^{O(d)} = s^{O(\sqrt{s}\cdot\log(1/\varepsilon))}$ monomials in p. We have

$$\begin{aligned}
\varepsilon &\le \mathop{\mathbf{E}}_{x,y\in\{-1,1\}^n}[p(h_1(x,y),\ldots,h_s(x,y))\cdot f(x,y)] \\
&= \mathop{\mathbf{E}}_{x,y\in\{-1,1\}^n}\left[\sum_{S\subseteq[s]:|S|\le d}\hat{p}(S)\cdot\prod_{i\in S}h_i(x,y)\cdot f(x,y)\right] \\
&= \sum_{S\subseteq[s]:|S|\le d}\hat{p}(S)\cdot\mathop{\mathbf{E}}_{x,y\in\{-1,1\}^n}\left[\prod_{i\in S}h_i(x,y)\cdot f(x,y)\right] \\
&\le \sum_{S\subseteq[s]:|S|\le d}(1+\varepsilon)\cdot\left|\mathop{\mathbf{E}}_{x,y\in\{-1,1\}^n}\left[\prod_{i\in S}h_i(x,y)\cdot f(x,y)\right]\right|
\end{aligned}$$

Hence there must exists a set $S\subseteq[s]$ with size at most d such that

$$\left|\mathop{\mathbf{E}}_{x,y\in\{-1,1\}^n}\left[\prod_{i\in S}h_i(x,y)\cdot F(x,y)\right]\right| \ge \frac{\varepsilon}{(1+\varepsilon)\cdot s^{O(d)}} \ge \frac{1}{s^{O(\sqrt{s}\cdot\log(1/\varepsilon))}}.$$

We get that $\prod_{i\in S}h_i(x,y) = (\prod_{i\in S}f_i(x))\cdot(\prod_{i\in S}g_i(y))$ is of the form $f_S(x)\cdot g_S(y)$ and is $1/s^{O(\sqrt{s}\cdot\log(1/\varepsilon))}$ correlated with F. In addition, observe that $f_S(x)\cdot g_S(y)\in\{-1,1\}$ for any $x\in\{-1,1\}^n$ and $y\in\{-1,1\}^n$. For any $a\in\{-1,1\}$ and $b\in\{-1,1\}$, let $I_a = \{x\in\{-1,1\}^n : f_S(x) = a\}$ and $J_b = \{y\in\{-1,1\}^n : g_S(y) = b\}$. We have

$$\begin{aligned}
\frac{1}{s^{O(\sqrt{s}\cdot\log(1/\varepsilon))}} &\le \left|\mathop{\mathbf{E}}_{x,y\in\{-1,1\}^n}[f_S(x)\cdot g_S(y)\cdot f(x,y)]\right| \\
&= \frac{1}{2^n\cdot 2^n}\cdot\left|\sum_{x\in\{-1,1\}^n}\sum_{y\in\{-1,1\}^n} f_S(x)\cdot g_S(y)\cdot f(x,y)\right| \\
&\le \frac{1}{2^n\cdot 2^n}\cdot\sum_{(a,b)\in\{-1,1\}^2}\left|\sum_{x\in I_a}\sum_{y\in J_b} a\cdot b\cdot f(x,y)\right| \\
&= \frac{1}{2^n\cdot 2^n}\cdot\sum_{(a,b)\in\{-1,1\}^2}\left|\sum_{x\in I_a}\sum_{y\in J_b} f(x,y)\right| \\
&\le 4\cdot\mathrm{disc}(M_f),
\end{aligned}$$

which completes the proof. □

5 Average-Case Lower Bounds for The Inner-Product Function

We strengthen Theorem 3.2 to the average-case. Namely, we show that any bipartite formula that even slightly correlates with the Inner-Product function must be of size $\widetilde{\Omega}(n^2)$. Our result will rely on Lemma 4.2 and the famous Lindsey Lemma (cf. [Juk11, Lemma 14.5]) that shows that the discrepancy of $\mathrm{IP_n}$ is exponentially small.

Lemma 5.1 (Lindsey's Lemma)**.** $\mathrm{disc}(M_{\mathrm{IP_n}}) \le 2^{-n/2}$.

Theorem 5.2 (Main Theorem - Average Case)**.** *Let* $\mathrm{IP_n}(x_1,\ldots,x_n,y_1,\ldots,y_n)$ *be the Inner-Product function. Let* F *be a bipartite Formula on* $x_1,\ldots,x_n,y_1,\ldots,y_n$ *with* $\mathbf{Pr}[\mathrm{IP_n}(x,y)=F(x,y)]\geq\frac{1}{2}+\varepsilon$*. Then,* F *is of size at least* $\Omega\left(\frac{n^2}{\log^2(n)\cdot\log^2(1/\varepsilon)}\right)$.

In particular, if F is a bipartite formula that has agreement $1/2+1/\mathrm{poly}(n)$ with $\mathrm{IP_n}$ (or even $1/2+1/n^{\mathrm{polylog}(n)}$), then F must be of size at least $\widetilde{\Omega}(n^2)$.

Proof. By Lindsey's Lemma (Lemma 5.1) and Lemma 4.2, we get

$$2^{-n/2}\geq\mathrm{disc}(M_{\mathrm{IP_n}})\geq 1/s^{O(\sqrt{s}\cdot\log(1/\varepsilon))},$$

hence $s=\Omega\left(\frac{n^2}{\log^2(n)\cdot\log^2(1/\varepsilon)}\right)$. □

Acknowledgements

I wish to thank Dima Gavinsky, Or Meir and Avi Wigderson for introducing me to this problem, for helpful discussions, and for their encouragement.

References

[ACR+10] A. Ambainis, A. M. Childs, B. Reichardt, R. Spalek, and S. Zhang. Any and-or formula of size n can be evaluated in time $\mathrm{n}^{1/2+\mathrm{o}(1)}$ on a quantum computer. *SIAM J. Comput.*, 39(6):2513–2530, 2010.

[And87] A. E. Andreev. On a method for obtaining more than quadratic effective lower bounds for the complexity of π-schemes. *Moscow Univ. Math. Bull.*, 42:63–66, 1987. In Russian.

[BBC+01] R. Beals, H. Buhrman, R. Cleve, M. Mosca, and R. de Wolf. Quantum lower bounds by polynomials. *J. ACM*, 48(4):778–797, 2001.

[BNRdW07] H. Buhrman, I. Newman, H. Röhrig, and R. de Wolf. Robust polynomials and quantum algorithms. *Theory Comput. Syst.*, 40(4):379–395, 2007.

[FGG08] E. Farhi, J. Goldstone, and S. Gutmann. A quantum algorithm for the hamiltonian nand tree. *Theory of Computing*, 4(1):169–190, 2008.

[For02] J. Forster. A linear lower bound on the unbounded error probabilistic communication complexity. *J. Comput. Syst. Sci.*, 65(4):612–625, 2002.

[Hås98] J. Håstad. The shrinkage exponent is 2. *SIAM J. on Computing*, 27:48–64, 1998.

[HLS07] P. Høyer, T. Lee, and R. Spalek. Negative weights make adversaries stronger. In *STOC*, pages 526–535, 2007.

[Juk11] S. Jukna. *Extremal combinatorics: with applications in computer science*. Springer Science & Business Media, 2011.

[Juk12] S. Jukna. *Boolean Function Complexity: Advances and Frontiers*. Springer Berlin Heidelberg, 2012.

[LLS06] S. Laplante, T. Lee, and M. Szegedy. The quantum adversary method and classical formula size lower bounds. *Computational Complexity*, 15(2):163–196, 2006.

[Lok09] S. V. Lokam. Complexity lower bounds using linear algebra. *Foundations and Trends in Theoretical Computer Science*, 4(1-2):1–155, 2009.

[Rei09] B. Reichardt. Span programs and quantum query complexity: The general adversary bound is nearly tight for every boolean function. In *FOCS*, pages 544–551, 2009.

[Rei11] B. Reichardt. Reflections for quantum query algorithms. In *SODA*, pages 560–569, 2011.

[RS12] B. Reichardt and R. Spalek. Span-program-based quantum algorithm for evaluating formulas. *Theory of Computing*, 8(1):291–319, 2012.

[Tal14] A. Tal. Shrinkage of de Morgan formulae from quantum query complexity. In *FOCS*, pages 551–560, 2014.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de