Electronic Colloquium on Computational Complexity, Report No. 32 (2019)

Strongly Exponential Separation Between Monotone VP and Monotone VNP

Srikanth Srinivasan*
Department of Mathematics
IIT Bombay

March 4, 2019

Abstract

We show that there is a sequence of explicit multilinear polynomials $P_n(x_1, \ldots, x_n) \in \mathbb{R}[x_1, \ldots, x_n]$ with non-negative coefficients that lies in monotone VNP such that any monotone algebraic circuit for P_n must have size $\exp(\Omega(n))$. This builds on (and strengthens) a result of Yehudayoff (2018) who showed a lower bound of $\exp(\tilde{\Omega}(\sqrt{n}))$.

1 Introduction

This paper deals with a problem in *Algebraic Complexity*, which is the study of the complexity of computing multivariate polynomials over some underlying field $\mathbb{F}$. The model of computation is the *Algebraic circuit* model, which computes polynomials from $\mathbb{F}[x_1, \ldots, x_n]$ using the basic sum and product operations in this ring. This model and its variants have been studied by a large body of work (see, e.g. the surveys [12, 10]).

The central question in the area is Valiant's [13] VP vs. VNP question. The set VP contains sequences $(P_n(x_1, \ldots, x_n))_{n\geq 1}$ of polynomials of polynomially bounded degree[1] that can be computed by polynomial-sized algebraic circuits. The class VNP contains sequences $(Q_n(x_1, \ldots, x_n))_{n\geq 1}$ where

$$Q_n(x_1, \ldots, x_n) = \sum_{b_1,\ldots,b_m \in \{0,1\}} P_{n+m}(x_1, \ldots, x_n, b_1, \ldots, b_m)$$

where m is polynomially bounded by n and $(P_r(x_1, \ldots, x_r))_{r\geq 1}$ is in VP.

Like its Boolean analogue, the VP vs. VNP question has proved stubbornly hard to resolve, the principal bottleneck being our inability to prove explicit algebraic circuit lower bounds. Given this, it is natural to look at variants of this question.

In a recent paper [15], Yehudayoff considered the *monotone* version of the VP vs. VNP question, which is defined as follows. The underlying field is $\mathbb{R}$ and the polynomials being computed have non-negative coefficients. A *monotone* algebraic circuit is one where all the constants appearing in the circuit are non-negative. The monotone versions of VP and VNP, denoted MVP and MVNP respectively, are defined analogously: MVP contains (sequences of) polynomials that have small

*Email: `srikanth@math.iitb.ac.in`
[1] i.e. $\deg(P_n) \leq n^{O(1)}$

ISSN 1433-8092

monotone algebraic circuits; MVNP contains (sequences of) polynomials that can be written as exponential Boolean sums over polynomials in MVP.

Monotone algebraic circuits have been studied since the 80s, and explicit exponential lower bounds are known for this model via the work of Jerrum and Snir [7] (see also [14, 11, 8]). However, as Yehudayoff [15] pointed out, these results do not imply a separation between MVP and MVNP. In fact, most[2] of the monotone circuit lower bounds proved in earlier work also imply that the same polynomials do not belong to MVNP, and hence do not imply a separation between these two classes.

The main result of [15] was the resolution of the MVP vs. MVNP question. More precisely, Yehudayoff showed that there is an explicit sequences of multilinear polynomials $(P_n(x_1,\ldots,x_n))_{n\geq 1}$ in MVNP such that any monotone algebraic circuit for P_n must have size $\exp(\tilde{\Omega}(\sqrt{n}))$.

In this paper, we strengthen this result to a strongly exponential lower bound.

Theorem 1. *There is an explicit sequence of multilinear polynomials* $(P_n(x_1,\ldots,x_n))_{n\geq 1}$ *in* MVNP *such that any monotone algebraic circuit for* P_n *must have size* $2^{\Omega(n)}$.

This theorem bears a similar relation to Yehudayoff's result as a paper of Raz and Yehudayoff [8] bears to the result of Jerrum and Snir [7]. Jerrum and Snir [7] proved a lower bound of $\exp(\Omega(\sqrt{n}))$ for an explicit family of polynomials.[3] This bound was strengthened to a strongly exponential lower bound by Raz and Yehudayoff [8].

1.1 Some Motivation for the proof

We rely on a connection between monotone algebraic circuit lower bounds and communication complexity that was made explicit by Raz and Yehudayoff [8]. As shown in [8], if a multilinear polynomial $P\in\mathbb{R}[x_1,\ldots,x_n]$ has a monotone algebraic circuit of size s, then we get a decomposition

$$P=\sum_{i=1}^{s} g_i h_i \tag{1}$$

where each summand g_ih_i satisfies the property that g_i and h_i are *non-negative* multilinear polynomials that depend on disjoint sets of at least $n/3$ variables each. We call each such term a *non-negative product polynomial.* Thus, to prove a lower bound on the circuit complexity of P, it suffices to lower bound the number of terms in any decomposition as in (1).

As noted by Jerrum and Snir [7], one way to do this is via the *support of the polynomial* P, by which we mean the set of monomials that have non-zero coefficients in P. We think of this set, denoted $\mathrm{Supp}(P)$, as a subset of $2^{[n]}$ by identifying each multilinear monomial on $x_1,\ldots,x_n$ with a subset of $[n]$ in the natural way. Given a decomposition of P into non-negative product polynomials as in (1), we immediately get $\mathrm{Supp}(P)=\bigcup_{i\in[s]}\mathrm{Supp}(g_i\cdot h_i)$. And so it suffices to obtain a P such that any such decomposition of $\mathrm{Supp}(P)$ must have large size.

Such decompositions are closely related to a model of communication complexity known as *Multipartition Communication Complexity*, introduced by Ďuris, Hromkovič, Jukna, Sauerhoff and Schnitger [4] (see also the earlier result of Borodin, Razborov and Smolensky [2]). The multipartition communication complexity of a subset $\mathcal{S}\subseteq 2^{[n]}$ (or equivalently a Boolean function

[2]The one exception to this seems to be a lower bound of Raz and Yehudayoff [8]. Here, it is unclear whether the hard polynomials lie in MVNP but we are unable to rule it out.

[3]The family is just the Permanent of a $\sqrt{n}\times\sqrt{n}$ matrix with distinct variable entries.

$f : \{0,1\}^n \to \{0,1\}$)) is defined as follows. We define a *rectangle* $\mathcal{R} \subseteq 2^{[n]}$ to be any set of the form $\{A \cup B \mid A \in \mathcal{A}, B \in \mathcal{B}\}$, where $\mathcal{A} \subseteq 2^X$ and $\mathcal{B} \subseteq 2^Y$ and (X, Y) is a partition of $[n]$. Further, we say that $\mathcal{R}$ is *balanced* if $|X|, |Y| \geq n/3$. Finally, the multipartition communication complexity of $\mathcal{S}$ is defined to be $\lceil \log_2 k \rceil$ where k is the smallest integer such that $\mathcal{S}$ can be decomposed as the union of k many balanced rectangles.

To see the connection to algebraic complexity, note that if $P \in \mathbb{R}[x_1, \ldots, x_n]$ has monotone algebraic circuits of size s, then $\mathrm{Supp}(P)$ has multipartition communication complexity at most $\lceil \log_2 s \rceil$. In particular, linear lower bounds in this model for some explicit $\mathcal{S}$ implies that any non-negative polynomial P with support exactly $\mathcal{S}$ cannot be computed by monotone algebraic circuits of subexponential size.

Polynomial (but sublinear) lower bounds for multipartition communication complexity were implicit in the work of Borodin et al. [2] and were extended to linear (but somewhat non-explicit) lower bounds in the work of Ďuris et al. [4]. An explicit linear lower bound for this model is implicit in a result of Bova, Capelli, Mengel and Slivovsky [3]. (See also the related work of Hayes [5]. A similar construction is attributed to Wigderson in [8].) The hard problem of [3] is quite easy to describe. Fix a regular expander graph[4] G on vertex set $[n]$ with constant-degree d. The associated hard problem is given by taking $\mathcal{S}$ to be the set of all vertex covers in G. (Alternately, we consider the Boolean function $f(x_1, \ldots, x_n) = \bigwedge_{\{i,j\} \in E(G)} (x_i \vee x_j)$.)

As mentioned above, the communication complexity lower bound on $\mathcal{S}$ immediately yields a strongly exponential lower bound on the monotone algebraic complexity of some explicitly defined polynomial. Unfortunately, as observed by Yehudayoff [15], this does not yield a separation between MVNP and MVP. This is because the above argument implies that *any* polynomial P_0 that has support $\mathcal{S}$ requires monotone algebraic circuits of exponential size. Yehudayoff showed that for any polynomial P in MVNP, there is a polynomial-sized monotone algebraic circuit that computes a polynomial Q with the same support. In particular, the polynomial P_0 cannot be in MVNP as that would contradict our lower bound above. Thus, to obtain a separation between MVNP and MVP along these lines, some new idea is necessary.

We take our cue from the multipartition communication complexity lower bound above, but modify it suitably to obtain a somewhat different lower bound candidate polynomial P. Our proof method for the lower bound, as in [15], is not just based on the support of P, but rather on the sizes of the coefficients of P. We define a probability distribution μ on the monomials of P and show that for any non-negative product polynomial $g_i h_i$ in a decomposition as in (1), a random monomial (chosen according to μ) has much smaller coefficient in the product polynomial than in P. As the product polynomials sum to P, there must be many of them. This yields the lower bound.

2 Defining the hard polynomial

Notation. Throughout, let $n \geq 1$ be a growing integer parameter. Let $X = \{x_1, \ldots, x_n\}$ be a set of indeterminates. We use x^S to denote the monomial $\prod_{i \in S} x_i$. Given a polynomial $P \in \mathbb{R}[x_1, \ldots, x_n]$ and $S \subseteq [n]$, we use $\mathrm{Coeff}(x^S, P)$ to denote the coefficient of the monomial x^S in the polynomial P.

[4]Recall that we call a d-regular graph G an expander if the second largest (in absolute value) eigenvalue of its adjacency matrix A is at most $d(1 - \Omega(1))$.

Let $(G_n)_{n>d}$ be an explicit sequence of d-regular expander graphs on n vertices with second largest eigenvalue at most $d^{0.75}$. Here, d is a large enough constant as specified below. Such an explicit sequence of expander graphs can be constructed using, say, [9]. The only fact we will use about expanders is the following, which is an easy consequence of the Expander Mixing Lemma [1] (see also [6, Lemma 2.5]).

For any pair of *disjoint* sets $U, V \subseteq V(G_n)$, we use $E(U,V)$ to denote the set of edges $\{u,v\} \in E(G_n)$ such that $u \in U$ and $v \in V$. Also, let $E(U)$ denote the set of edges $e = \{u,v\} \in E(G_n)$ such that $u, v \in U$.

Lemma 2 (Corollary to Expander Mixing Lemma)**.** *Let G_n be as above. Then, for any disjoint sets $U, V \subseteq [n]$ such that $|U|, |V| \in [n/3, 2n/3]$, we have*

$$|E(U,V)| \geq \frac{|E(G_n)|}{10}.$$

as long as d is a large enough constant.

From now on, d will be fixed to be a large enough constant so that the inequality in Lemma 2 holds.

We define the polynomial $P_n(x_1, \ldots, x_n)$ as follows. We assume that $V(G_n) = [n]$. For each edge $e \in E(G)$ introduce a variable x'_e and let $X' = \{x'_e \mid e \in E(G)\}$. Notice that for each Boolean assignment to the variables in X', we obtain a subgraph H of G. In particular, if the variables in X' are set *randomly* to Boolean values, we get a random subgraph H of G with the same vertex set $[n]$. We use $\deg_H(i)$ to denote the degree of the vertex i in the graph H.

We now define

$$P_n(x_1, \ldots, x_n) = \mathop{\mathbf{E}}_{\substack{x'_e \in \{0,1\} \\ \forall e \in E(G)}} \left[\prod_{i \in [n]} \left(1 + x_i \cdot 2^{\deg_H(i)} \right) \right] \tag{2}$$

$$= \mathop{\mathbf{E}}_{\substack{x'_e \in \{0,1\} \\ \forall e \in E(G)}} \left[\sum_{S \subseteq [n]} x^S \cdot 2^{\sum_{i \in S} \deg_H(i)} \right] \tag{3}$$

where the variables x'_e are set to one of $\{0,1\}$ independently and uniformly at random.

Lemma 3. *The sequence of polynomials P_n as defined above is in* mVNP*.*

Proof. Using (2), we see that

$$P_n(x_1, \ldots, x_n) = \frac{1}{2^{|E(G_n)|}} \sum_{x'_e \in \{0,1\} : e \in E(G)} \prod_{i \in [n]} \left(1 + x_i \cdot 2^{\sum_{e \ni i} x'_e} \right).$$

Since G is d-regular, it suffices to show that each function $f : \{0,1\}^d \to \mathbb{R}$ defined by $f(x'_1, \ldots, x'_d) = 2^{\sum_{j \in [d]} x'_j}$ can be represented by a constant-sized polynomial over $x'_1, \ldots, x'_d$ with *non-negative* coefficients.

But this is clear since $f(x'_1, \ldots, x'_d) = \sum_{S \subseteq [d]} \prod_{i \in S} x'_i$. □

Remark 4. *The above lemma also holds if we change the definition of P_n in (2) with the constant* 2 *replaced by any fixed $c > 1$.*

3 The lower bound

The main theorem of this section is the following.

Theorem 5. *Any monotone circuit computing P_n has size $2^{\Omega(n)}$.*

We need the following lemma from [8]. We say that a pair of multilinear polynomials $(g, h) \in \mathbb{R}[X]$ form a *non-negative product pair* if g, h are polynomials with non-negative coefficients, and there is a partition of $X = Y \cup Z$ where $n/3 \le |Y|, |Z| \le 2n/3$ and $g \in \mathbb{R}[Y], h \in \mathbb{R}[Z]$.

Lemma 6 ([8], Lemma 3.3)**.** *Assume that P_n has a monotone circuit of size s. Then*

$$P_n(X) = \sum_{i=1}^{s+1} g_i h_i$$

where for each $i \in [s]$, (g_i, h_i) forms a non-negative product pair.

Corollary 7. *Assume that P_n has a monotone circuit of size s. Let μ be any probability distribution on subsets $S \subseteq [n]$. Then, there is a non-negative product pair (g, h) such that*

- *$gh \le P$, i.e.,* $\mathrm{Coeff}(x^S, gh) \le \mathrm{Coeff}(x^S, P_n)$ *for each $S \subseteq [n]$,*
- $\boldsymbol{E}_{S\sim\mu}\left[\mathrm{Coeff}(x^S, gh)/\mathrm{Coeff}(x^S, P_n)\right] \ge 1/(s+1)$. *(The quantity* $\mathrm{Coeff}(x^S, gh)/\mathrm{Coeff}(x^S, P_n)$ *is well defined since by (3), the denominator is non-zero for all $S \subseteq [n]$.)*

Proof. Write $P_n = \sum_{i \le s+1} g_i h_i$ as in Lemma 6. For any fixed $S \subseteq [n]$ and a uniformly random $i \in [s+1]$, we have

$$\mathop{\mathbf{E}}_{i\in[s+1]}\left[\frac{\mathrm{Coeff}(x^S, g_i h_i)}{\mathrm{Coeff}(x^S, P_n)}\right] = \frac{1}{s+1}\sum_{i\in[s]}\frac{\mathrm{Coeff}(x^S, g_i h_i)}{\mathrm{Coeff}(x^S, P_n)} = \frac{1}{s+1}.$$

In particular, the above also holds when S is chosen according to μ. The result now follows by averaging over $i \in [s+1]$. $\square$

Given Corollary 7, to prove Theorem 5, it suffices to show the following.

Lemma 8. *There is a probability distribution μ on subsets $S \subseteq [n]$ such that for any non-negative product pair (g, h) with $gh \le P_n$, we have*

$$\mathop{\boldsymbol{E}}_{S\sim\mu}\left[\mathrm{Coeff}(x^S, gh)/\mathrm{Coeff}(x^S, P_n)\right] \le \exp(-\Omega(n)). \tag{4}$$

We need some preparatory work before proving Lemma 8.

Lemma 9. *There exist constants $A, B > 1$ such that*

$$P_n(X) = \sum_{S\subseteq[n]} x^S B^{|S|} A^{|E(S)|}.$$

Proof. Using (3), we obtain

$$P_n(x_1, \ldots, x_n) = \mathop{\mathbf{E}}_{x'_e : e \in E(G)} \left[\sum_{S \subseteq [n]} x^S \cdot 2^{\sum_{i \in S} \deg_H(i)} \right] = \sum_{S \subseteq [n]} x^S \cdot \mathop{\mathbf{E}}_{x'_e : e \in E(G)} \left[2^{\sum_{i \in S} \deg_H(i)} \right]$$

where H is the random subgraph of G defined by a uniformly random Boolean assignment to the variables in X'. Note that

$$\sum_{i \in S} \deg_H(i) = \sum_{i \in S} \sum_{e \ni i} x'_e = \sum_{e \in E(S, \bar{S})} x'_e + \sum_{e \in E(S)} 2x'_e.$$

Hence, we get for any $S \subseteq [n]$,

$$\begin{aligned} \mathop{\mathbf{E}}_{x'_e : e \in E(G)} \left[2^{\sum_{i \in S} \deg_H(i)} \right] &= \prod_{e \in E(S, \bar{S})} \mathop{\mathbf{E}}_{x'_e} \left[2^{x'_e} \right] \cdot \prod_{e \in E(S)} \mathop{\mathbf{E}}_{x'_e} \left[4^{x'_e} \right] = (3/2)^{|E(S, \bar{S})|} \cdot (5/2)^{|E(S)|} \\ &= (3/2)^{|S|d} \cdot \frac{(5/2)^{|E(S)|}}{(3/2)^{2|E(S)|}} \end{aligned}$$

where for the last equality, we have used the fact that $2|E(S)| + |E(S, \bar{S})| = |S|d$. Note that this proves the lemma with $B = (3/2)^d$ and $A = (10/9)$. □

We now define the probability distribution μ that will be shown to have the property in (4). The distribution is defined by the following sampling process. Let $m = \alpha n$ where $\alpha \in (0, 1)$ is a small constant specified below.

Sampling Algorithm $\mathcal{S}$:

1. Set $M = \emptyset$. (Eventually, M will be a matching in G.)
2. For $i = 1$ to $(m/2)$, do the following.
   (a) Remove all vertices from G_n that are at distance at most 2 from any vertex in the matching M. Let $G_n^{(i)}$ be the resulting graph.
   (b) Choose a uniformly random edge e_i from $E(G_n^{(i)})$ and add it to M.
3. Output M.

The above algorithm defines a distribution ν over matchings M in G_n of size $m/2$. We define $S = V(M)$ to be the set of vertices sampled by the algorithm. This defines a probability distribution μ over subsets of $[n]$.

We will need the following properties of the above algorithm.

Lemma 10 (Properties of $\mathcal{S}$.)**.** *Let M be sampled as in $\mathcal{S}$ above and let $S = V(M)$. Then we have*

1. $|M| = (m/2)$, $|S| = m$ *and* $E(S) = M$ *with probability* 1.

2. *Let (U,V) be any partition of $V(G_n)$ such that $n/3 \leq |U|,|V| \leq 2n/3$. Then, as long as $\alpha \leq 1/(100 \cdot d^2)$, for some absolute constant $\gamma > 0$, we have*

$$\Pr_M[|M \cap E(U,W)| \leq \gamma m] \leq \exp(-\gamma m).$$

3. *Let M_1 and M_2 be two independent samples obtained by running $\mathcal{S}$ twice, and let $S_i = V(M_i)$ ($i \in [2]$). Let (U,V) be a partition of $V(G_n)$ as above. Define $R_i = S_i \cap U$ and $T_i = S_i \cap V$. Then, for $\alpha \leq \gamma \ln A/(100 \cdot A^4 d^2)$ we have*

$$\mathop{\boldsymbol{E}}_{M_1,M_2}\left[A^{|E(R_1,T_2)|+|E(R_2,T_1)|}\right] \leq A^{\gamma m/4}.$$

Proof. Item 1 is immediate from the definition of the Sampling algorithm $\mathcal{S}$.

For Item 2, we proceed as follows. For $i \in \{1,\ldots,m/2\}$, let e_i be the edge chosen by the sampling algorithm $\mathcal{S}$ in the ith iteration of Step 2. Fix any choice of e_j for $j < i$ and consider the ith iteration of Step 2. The probability that e_i lies in $E(U,V)$ is $|E_i(U,V)|/|E(G_n^{(i)})|$ where $E_i(U,V)$ is the set of edges in $G_n^{(i)}$ with one endpoint each in U and V. Note that

$$|E_i(U,V)| \geq |E(U,V)| - |E(G_n) \setminus E(G_n^{(i)})| \geq \frac{nd}{10} - 2(i-1)\cdot(d^3+d^2+d) \geq \frac{nd}{10} - \alpha n(3d^3) \geq \frac{nd}{20}$$

where the second inequality follows from Lemma 2 and the fact that for each vertex incident to one of $e_1,\ldots,e_{i-1}$, we remove at most d^2+d+1 vertices (and hence at most d^3+d^2+d edges) from G_n to obtain $G_n^{(i)}$; and the last two inequalities follow from the fact that $2(i-1) < m = \alpha n \leq n/(100\cdot d^2)$. Hence, we have shown that for each i,

$$\Pr[e_i \in E(U,V) \mid e_1,\ldots,e_{i-1}] = \frac{|E_i(U,V)|}{|E(G_n^{(i)})|} \geq \frac{(nd)/20}{(nd)/2} = \frac{1}{10}.$$

In particular, for any $T \subseteq [m]$, the probability that for every $i \in T$, $e_i \notin E(U,V)$ can be upper bounded by $(9/10)^{|T|}$.

Thus, the probability that $|M \cap E(U,V)| \leq \ell = \gamma m$ can be bounded by

$$\begin{aligned}\Pr_M[\exists T \in \binom{[m]}{m-\ell} \ s.t.\ \forall i \in T, e_i \notin E(U,V)] &\leq \sum_T \Pr_M[\forall i \in T, e_i \notin E(U,V)] \\ &\leq \binom{m}{\ell}\left(\frac{9}{10}\right)^{m-\ell} \leq \left(\frac{em}{\ell}\right)^{\ell}\cdot\left(\frac{9}{10}\right)^{m-\ell} \\ &= \left(\frac{e}{\gamma}\cdot\left(\frac{9}{10}\right)^{(1/\gamma)-1}\right)^{\gamma m} \leq \exp(-\gamma m)\end{aligned}$$

as long as γ is bounded by a small enough absolute constant. This finishes the proof of Item 2.

We now prove Item 3. Fix any possible value of M_1 as sampled by the algorithm $\mathcal{S}$. It suffices to bound $\mathbf{E}_{M_2}\left[A^{|E(R_2,T_1)|+|E(R_1,T_2)|}\right]$ for each such M_1. Let $\tilde{S}_1$ denote the set of vertices that are at distance at most 1 from S_1 and let E_1 denote the set of edges that have at least one endpoint in $\tilde{S}_1$. Note that $|E_1| \leq |\tilde{S}_1|d \leq |S_1|d^2 = md^2$.

We claim that $|E(R_2,T_1)| + |E(R_1,T_2)| \leq 4|E_1 \cap M_2|$. The reason for this is that if a vertex $i \in S_2$ is incident to an edge e in $E(R_1,T_2) \cup E(R_2,T_1)$ then $i \in \tilde{S}_1$ and hence the edge $e' \in M_2$

involving i is an edge in $E_1 \cap M_2$. In particular, the number of such vertices $i \in S_2$ is at most $2|E_1 \cap M_2|$. Further, each such vertex i is adjacent to at most 2 vertices in S_1 since vertices in S_1 that are not adjacent via an edge in M_1 are at distance at least 3 from each other. Thus each such vertex i contributes at most 2 to $|E(R_2, T_1)| + |E(R_1, T_2)|$. This yields the claimed inequality. Thus it suffices to bound $\mathbf{E}_{M_2}\left[A^{4|E_1 \cap M_2|}\right]$.

We start with a tail bound for $|E_1 \cap M_2|$. Let $M_2 = \{e'_1, \ldots, e'_{m/2}\}$ where e'_j is the jth edge added to M_2 by the algorithm $\mathcal{S}$. Conditioned on $e'_1, \ldots, e'_{j-1}$, the probability that $e'_j \in E_1$ is at most

$$\frac{|E_1|}{|E(G_n^{(j)})|} = \frac{|E_1|}{|E(G_n)| - |E(G_n) \setminus E(G_n^{(j)})|} \le \frac{\alpha n d}{(nd/2) - 3\alpha n d^3} \le \frac{\alpha n d}{(nd)/4} = 4\alpha$$

where for the first inequality we have bounded $|E(G_n) \setminus E(G_n^{(j)})|$ as above and for the second inequality we have used the bound on α. Hence, we have

$$\Pr_{M_2}[|E_1 \cap M_2| \ge i] \le \sum_{T \in \binom{m/2}{i}} \Pr_{M_2}[\forall j \in T, e'_j \in E_1] \le \binom{m/2}{i} (4\alpha)^i.$$

This allows us to bound $\mathbf{E}_{M_2}\left[A^{4|E_1 \cap M_2|}\right]$ for any fixed M_1 output by $\mathcal{S}$.

$$\begin{aligned}
\mathop{\mathbf{E}}_{M_2}\left[A^{4|E_1 \cap M_2|}\right] &\le \sum_{i=0}^{m/2} A^{4i} \Pr_{M_2}[|E_1 \cap M_2| \ge i] \\
&\le \sum_{i=0}^{m/2} A^{4i} \cdot \binom{m/2}{i} (4\alpha)^i = \left(1 + 4\alpha A^4\right)^{m/2} \\
&\le (1 + (\gamma \ln A)/2)^{m/2} \le \exp((m\gamma \ln A)/4) = A^{\gamma m/4}
\end{aligned}$$

where the third inequality follows from the bound on α. $\square$

We are now ready to prove Lemma 8, which will complete the proof of Theorem 5.

Proof of Lemma 8. We set $m = \alpha n$ so that α is a positive constant upper bounded by $\gamma \ln A/(100 \cdot A^4 \cdot d^2)$ and m is even. Assume that M is as sampled above by sampling algorithm $\mathcal{S}$ and $S = V(M)$. This defines the distribution μ on subsets of $[n]$.

Let (g, h) be any non-negative product pair such $gh \le P_n$. Consequently, there exists a partition (U, V) of $V(G_n) = [n]$ such that $n/3 \le |U|, |V| \le 2n/3$ and $g \in \mathbb{R}[x_i : i \in U], h \in \mathbb{R}[x_j : j \in V]$.

Let $\mathcal{E} = \mathcal{E}(M)$ denote the event that $|M \cap E(U, V)| \le \gamma m$. By Lemma 10 item 2, we know that $\Pr_M[\mathcal{E}] \le \exp(-\Omega(n))$ and hence we have

$$\begin{aligned}
\mathop{\mathbf{E}}_{M}\left[\frac{\mathrm{Coeff}(x^S, gh)}{\mathrm{Coeff}(x^S, P_n)}\right] &\le \mathop{\mathbf{E}}_{M}\left[\frac{\mathrm{Coeff}(x^S, gh)}{\mathrm{Coeff}(x^S, P_n)} \mid \mathcal{E}\right] \Pr_M[\mathcal{E}] + \mathop{\mathbf{E}}_{M}\left[\frac{\mathrm{Coeff}(x^S, gh)}{\mathrm{Coeff}(x^S, P_n)} \mid \overline{\mathcal{E}}\right] \Pr_M[\bar{\mathcal{E}}] \\
&\le \Pr_M[\mathcal{E}] + \mathop{\mathbf{E}}_{M}\left[\frac{\mathrm{Coeff}(x^S, gh)}{\mathrm{Coeff}(x^S, P_n)} \mid \overline{\mathcal{E}}\right] \\
&\le \exp(-\Omega(n)) + \frac{1}{B^m \cdot A^{m/2}} \mathop{\mathbf{E}}_{M}\left[\mathrm{Coeff}(x^S, gh) \mid \overline{\mathcal{E}}\right] \qquad (5)
\end{aligned}$$

where for the second inequality we have used that $gh \le P_n$, and for the final inequality we have used our bound on $\Pr_M[\mathcal{E}]$ along with Lemma 9 and Lemma 10 item 1.

We now bound the latter term in (5). For any i, j, k, let $\mathcal{E}_{i,j,k} = \mathcal{E}_{i,j,k}(M)$ denote the event that $|M \cap E(U, V)| = i, |M \cap E(U)| = j$, and $|M \cap E(V)| = k$. The event $\bar{\mathcal{E}}$ is partitioned into $\mathcal{E}_{i,j,k}$ where $i + j + k = (m/2)$ and $i \geq \gamma m$. Let $\mathcal{T}$ denote the set of such triples (i, j, k). We have

$$\mathop{\mathbf{E}}_{M}\left[\mathrm{Coeff}(x^S, gh) \mid \overline{\mathcal{E}}\right] = \sum_{(i,j,k)\in\mathcal{T}} \mathop{\mathbf{E}}_{M}\left[\mathrm{Coeff}(x^S, gh) \mid \mathcal{E}_{i,j,k}\right] \cdot \Pr_{M}[\mathcal{E}_{i,j,k} \mid \overline{\mathcal{E}}]$$

Call a triple $(i, j, k) \in \mathcal{T}$ *heavy* if $\Pr_M[\mathcal{E}_{i,j,k}] \geq A^{-\gamma m/4}$ and *light* otherwise. Note that as $\Pr_M[\overline{\mathcal{E}}] = 1 - \exp(-\Omega(n)) \geq 1/2$, we have $\Pr_M[\mathcal{E}_{i,j,k} \mid \overline{\mathcal{E}}] \leq 2\Pr_M[\mathcal{E}_{i,j,k}]$. In particular, if (i, j, k) is light, we have $\Pr_M[\mathcal{E}_{i,j,k} \mid \overline{\mathcal{E}}] \leq 2A^{-\gamma m/4} = \exp(-\Omega(n))$. Plugging this into the above, we get

$$\begin{aligned}
&\mathop{\mathbf{E}}_{M}\left[\mathrm{Coeff}(x^S, gh) \mid \overline{\mathcal{E}}\right] = \sum_{(i,j,k)\in\mathcal{T}} \mathop{\mathbf{E}}_{M}\left[\mathrm{Coeff}(x^S, gh) \mid \mathcal{E}_{i,j,k}\right] \cdot \Pr_{M}[\mathcal{E}_{i,j,k} \mid \overline{\mathcal{E}}] \\
&\leq |\{(i,j,k) \mid (i,j,k) \text{ light}\}| \cdot \mathrm{Coeff}(x^S, P_n) \cdot \exp(-\Omega(n)) + \max_{(i,j,k) \text{ heavy}} \mathop{\mathbf{E}}_{M}\left[\mathrm{Coeff}(x^S, gh) \mid \mathcal{E}_{i,j,k}\right] \\
&\leq \exp(-\Omega(n)) B^m A^{m/2} + \max_{(i,j,k) \text{ heavy}} \mathop{\mathbf{E}}_{M}\left[\mathrm{Coeff}(x^S, gh) \mid \mathcal{E}_{i,j,k}\right]. \qquad (6)
\end{aligned}$$

It suffices therefore to bound $\mathbf{E}_M\left[\mathrm{Coeff}(x^S, gh) \mid \mathcal{E}_{i,j,k}\right]$ for any heavy (i, j, k). This is the main part of the proof.

Fix some $(i, j, k) \in \mathcal{T}$ that is heavy. Let $C = \mathbf{E}_M\left[\mathrm{Coeff}(x^S, gh) \mid \mathcal{E}_{i,j,k}\right]$. Thus, we get

$$C^2 = \mathop{\mathbf{E}}_{M_1,M_2}\left[\mathrm{Coeff}(x^{S_1}, gh)\mathrm{Coeff}(x^{S_2}, gh)\right]$$

where M_1 and M_2 are independent samples of M conditioned on the event $\mathcal{E}_{i,j,k}(M)$, and $S_\ell = V(M_\ell)$ for $\ell \in \{1, 2\}$. Define $R_\ell = S_\ell \cap U$ and $T_\ell = S_\ell \cap V$. We make some simple observations. For each $\ell \in [2]$

1. $|R_\ell| = i + 2j$ and $|T_\ell| = i + 2k$,
2. $|E(R_\ell)| = j, |E(T_\ell)| = k$,
3. $\mathrm{Coeff}(x^{S_\ell}, gh) = \mathrm{Coeff}(x^{R_\ell}, g) \cdot \mathrm{Coeff}(x^{T_\ell}, h)$.

Thus, we have

$$\begin{aligned}
C^2 &= \mathop{\mathbf{E}}_{M_1,M_2}\left[\mathrm{Coeff}(x^{R_1}, g)\mathrm{Coeff}(x^{T_1}, h)\mathrm{Coeff}(x^{R_2}, g)\mathrm{Coeff}(x^{T_2}, h)\right] \\
&= \mathop{\mathbf{E}}_{M_1,M_2}\left[\mathrm{Coeff}(x^{R_1\cup T_2}, gh)\mathrm{Coeff}(x^{R_2\cup T_1}, gh)\right] \\
&\leq \mathop{\mathbf{E}}_{M_1,M_2}\left[\mathrm{Coeff}(x^{R_1\cup T_2}, P_n)\mathrm{Coeff}(x^{R_2\cup T_1}, P_n)\right] \\
&= \mathop{\mathbf{E}}_{M_1,M_2}\left[B^{|R_1|+|T_2|} \cdot A^{|E(R_1\cup T_2)|} \cdot B^{|R_2|+|T_1|} \cdot A^{|E(R_2\cup T_1)|}\right] \\
&= \mathop{\mathbf{E}}_{M_1,M_2}\left[B^{4(i+j+k)} \cdot A^{|E(R_1)|+|E(T_1)|+|E(R_2)|+|E(T_2)|+|E(R_1,T_2)|+|E(R_2,T_1)|}\right] \\
&= B^{2m}A^{m-2i} \cdot \mathop{\mathbf{E}}_{M_1,M_2}\left[A^{|E(R_1,T_2)|+|E(R_2,T_1)|}\right] \\
&\leq B^{2m}A^{m(1-2\gamma)} \cdot \mathop{\mathbf{E}}_{M_1,M_2}\left[A^{|E(R_1,T_2)|+|E(R_2,T_1)|}\right] \qquad (7)
\end{aligned}$$

where we used the observations above for the equalities and for the final inequality, we used the fact that $i \geq \gamma m$ for all $(i, j, k) \in \mathcal{T}$.

To bound the latter term in (7), we consider a similar expression where M_1 and M_2 are replaced by M_1' and M_2' which are independent random outputs of the algorithm $\mathcal{S}$ (without any conditioning). In this case, by Lemma 10 item 3, we have

$$\mathop{\mathbf{E}}_{M_1', M_2'}\left[A^{|E(R_1', T_2')| + |E(R_2', T_1')|}\right] \leq A^{\gamma m/4},$$

where R_ℓ', T_ℓ' are defined analogously for $\ell \in [2]$. Thus, using Bayes' rule we have

$$\mathop{\mathbf{E}}_{M_1, M_2}\left[A^{|E(R_1, T_2)| + |E(R_2, T_1)|}\right] \leq \frac{\mathbf{E}_{M_1', M_2'}\left[A^{|E(R_1', T_2')| + |E(R_2', T_1')|}\right]}{\Pr_{M_1', M_2'}[\mathcal{E}_{i,j,k}(M_1') \wedge \mathcal{E}_{i,j,k}(M_2')]} \leq A^{3\gamma m/4}$$

where the last inequality uses the fact that (i, j, k) is heavy. Plugging the above into (7), we have

$$C \leq B^m A^{m/2 - 5\gamma m/8} = B^m A^{m/2} \exp(-\Omega(n)).$$

As this holds for any heavy (i, j, k), using (6) and (5), we obtain the statement of the lemma. $\square$

Acknowledgements. The author is grateful to Mrinal Kumar and Amir Yehudayoff for very helpful discussions and encouragement. This work was done during a visit to the "Lower Bounds Program in Computational Complexity" program at the Simons Institute for the Theory of Computing. The author is grateful to the organizers of this program and the Simons Institute for their hospitality.

References

[1] N. Alon and F. Chung. Explicit construction of linear sized tolerant networks. *Discrete Mathematics*, 72(1):15 – 19, 1988.

[2] A. Borodin, A. A. Razborov, and R. Smolensky. On lower bounds for read-k-times branching programs. *Computational Complexity*, 3:1–18, 1993.

[3] S. Bova, F. Capelli, S. Mengel, and F. Slivovsky. Expander cnfs have exponential DNNF size. *CoRR*, abs/1411.1995, 2014.

[4] P. Duris, J. Hromkovic, S. Jukna, M. Sauerhoff, and G. Schnitger. On multi-partition communication complexity. *Inf. Comput.*, 194(1):49–75, 2004.

[5] T. P. Hayes. Separating the k-party communication complexity hierarchy: an application of the zarankiewicz problem. *Discrete Mathematics & Theoretical Computer Science*, 13(4):15–22, 2011.

[6] S. Hoory, N. Linial, and A. Wigderson. Expander graphs and their applications. *Bulletin of the American Mathematical Society*, 43(4):439–561, 2006.

[7] M. Jerrum and M. Snir. Some exact complexity results for straight-line computations over semirings. *J. ACM*, 29(3):874–897, 1982.

[8] R. Raz and A. Yehudayoff. Multilinear formulas, maximal-partition discrepancy and mixed-sources extractors. *J. Comput. Syst. Sci.*, 77(1):167–190, 2011.

[9] O. Reingold, S. Vadhan, and A. Wigderson. Entropy waves, the zig-zag graph product, and new constant-degree expanders. *Annals of mathematics*, pages 157–187, 2002.

[10] R. Saptharishi. A survey of lower bounds in arithmetic circuit complexity. *Github survey*, 2015.

[11] E. Shamir and M. Snir. *Lower bounds on the number of multiplications and the number of additions in monotone computations*. IBM Thomas J. Watson Research Division, 1977.

[12] A. Shpilka and A. Yehudayoff. Arithmetic circuits: A survey of recent results and open questions. *Foundations and Trends in Theoretical Computer Science*, 5(3-4):207–388, 2010.

[13] L. G. Valiant. Completeness classes in algebra. In *Proceedings of the 11h Annual ACM Symposium on Theory of Computing, April 30 - May 2, 1979, Atlanta, Georgia, USA*, pages 249–261, 1979.

[14] L. G. Valiant. Negation can be exponentially powerful. *Theor. Comput. Sci.*, 12:303–314, 1980.

[15] A. Yehudayoff. Separating monotone VP and VNP. *Electronic Colloquium on Computational Complexity (ECCC)*, 25:124, 2018.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il