Log-concavity and lower bounds for arithmetic circuits

Ignacio García-Marco[1⋆], Pascal Koiran[1⋆], and Sébastien Tavenas[2⋆]

[1] LIP, ENS Lyon, France ⋆⋆
ignacio.garcia-marco@ens-lyon.fr, pascal.koiran@ens-lyon.fr
[2] Max-Planck-Insitut für Informatik, Saarbrücken, Germany
stavenas@mpi-inf.mpg.de

Abstract. One question that we investigate in this paper is, how can we build log-concave polynomials using sparse polynomials as building blocks? More precisely, let $f = \sum_{i=0}^{d} a_i X^i \in \mathbb{R}^+[X]$ be a polynomial satisfying the log-concavity condition $a_i^2 > \tau a_{i-1}a_{i+1}$ for every $i \in \{1, \ldots, d-1\}$, where $\tau > 0$. Whenever f can be written under the form $f = \sum_{i=1}^{k} \prod_{j=1}^{m} f_{i,j}$ where the polynomials $f_{i,j}$ have at most t monomials, it is clear that $d \leqslant kt^m$. Assuming that the $f_{i,j}$ have only non-negative coefficients, we improve this degree bound to $d = \mathcal{O}(km^{2/3}t^{2m/3}\log^{2/3}(kt))$ if $\tau > 1$, and to $d \leqslant kmt$ if $\tau = d^{2d}$.
This investigation has a complexity-theoretic motivation: we show that a suitable strengthening of the above results would imply a separation of the algebraic complexity classes VP and VNP. As they currently stand, these results are strong enough to provide a new example of a family of polynomials in VNP which cannot be computed by monotone arithmetic circuits of polynomial size.

1 Introduction

Let $f = \sum_{j=0}^{d} a_j X^j \in \mathbb{R}[X]$ be a univariate polynomial of degree $d \in \mathbb{Z}^+$. It is a classical result due to Newton (see [4], §2.22 and §4.3 for two proofs) that whenever all the roots of f are real, then the coefficients of f satisfy the following log-concavity condition:

$$a_i^2 \geqslant \frac{d-i+1}{d-i}\frac{i+1}{i}\, a_{i-1}a_{i+1} \text{ for all } i \in \{1, \ldots, d-1\}. \tag{1}$$

Moreover, if the roots of f are not all equal, these inequalities are strict. When $d = 2$, condition (1) becomes $a_1 \geqslant 4a_0a_2$, which is well known to be a necessary and sufficient condition for all the roots of f to be real. Nevertheless, for $d \geqslant 3$, the converse of Newton's result does not hold any more [13].

When $f \in \mathbb{R}^+[X]$, i.e., when $f = \sum_{j=0}^{d} a_j X^j$ with $a_j \geqslant 0$ for all $j \in \{0, \ldots, d\}$, a weak converse of Newton's result holds true. Namely, a sufficient condition for f to only have real (and distinct) roots is that

$$a_i^2 > 4a_{i-1}a_{i+1} \text{ for all } i \in \{1, \ldots, d-1\}.$$

⋆ This work was supported by ANR project CompA (project number: ANR-13-BS02-0001-01).
⋆⋆ UMR 5668 ENS Lyon - CNRS - UCBL - INRIA, Université de Lyon

Whenever a polynomial fulfills this condition, we say that it satisfies the *Kurtz condition* since this converse result is often attributed to Kurtz [13]. Note however that it was obtained some 70 years earlier by Hutchinson [6].

If f satisfies the Kurtz condition, all of its $d+1$ coefficients are nonzero except possibly the constant term. Such a polynomial is therefore very far from being sparse (recall that a polynomial is informally called *sparse* if the number of its nonzero coefficients is small compared to its degree). One question that we investigate in this paper is: how can we construct polynomials satisfying the Kurtz condition using sparse polynomials as building blocks? More precisely, consider f a polynomial of the form

$$f = \sum_{i=1}^{k} \prod_{j=1}^{m} f_{i,j} \tag{2}$$

where $f_{i,j}$ are polynomials with at most t monomials each. By expanding the products in (2) we see that f has at most kt^m monomials. As a result, $d \leqslant kt^m$ if f satisfies the Kurtz condition. Our goal is to improve this very coarse bound. For the case of polynomials $f_{i,j}$ with nonnegative coefficients, we obtain the following result.

Theorem 1. *Consider a polynomial $f \in \mathbb{R}^+[X]$ of degree d of the form*

$$f = \sum_{i=1}^{k} \prod_{j=1}^{m} f_{i,j},$$

where $m \geqslant 2$ and the $f_{i,j} \in \mathbb{R}^+[X]$ have at most t monomials. If f satisfies the Kurtz condition, then $d = \mathcal{O}(km^{2/3}t^{2m/3}\log^{2/3}(kt))$.

We prove this result in Section 2. After that, in Section 3, we study the following stronger log-concavity condition

$$a_i^2 > d^{2d} a_{i-1} a_{i+1} \text{ for all } i \in \{1, \ldots, d-1\}. \tag{3}$$

In this setting we prove the following improved analogue of Theorem 1.

Theorem 2. *Consider a polynomial $f \in \mathbb{R}^+[X]$ of degree d of the form*

$$f = \sum_{i=1}^{k} \prod_{j=1}^{m} f_{i,j},$$

where $m \geqslant 2$ and the $f_{i,j} \in \mathbb{R}^+[X]$ have at most t monomials. If f satisfies (3), then $d \leqslant kmt$.

This investigation has a complexity-theoretic motivation: we show in Section 4 that a suitable extension of Theorem 2 (allowing negative coefficients for the polynomials f_{ij}) would imply a separation of the algebraic complexity classes VP and VNP. The classes VP of "easily computable polynomial families" and VNP of "easily definable polynomial families" were proposed by Valiant [15] as algebraic analogues of P and NP. As shown in Theorem 7, Theorem 2 as it now stands is strong enough to provide a new example of a family of polynomials in VNP which cannot be computed by monotone arithmetic circuits of polynomial size.

2 The Kurtz log-concavity condition

Our main tool in this section is a result of convex geometry [3]. To state this result, we need to introduce some definitions and notations. For a pair of planar finite sets $R, S \subset \mathbb{R}^2$, the *Minkowski sum* of R and S is the set $R + S := \{y + z \,|\, y \in R, z \in S\} \subset \mathbb{R}^2$. A finite set $C \subset \mathbb{R}^2$ is *convexly independent* if and only if its elements are vertices of a convex polygon. The following result provides an upper bound for the number of elements of a convexly independent set contained in the Minkowski sum of two other sets.

Theorem 3. *[3, Theorem 1] Let R and S be two planar point sets with $|R| = r$ and $|S| = s$. Let C be a subset of the Minkowski sum $R + S$. If C is convexly independent we have that $|C| = \mathcal{O}(r^{2/3}s^{2/3} + r + s)$.*

From this result the following corollary follows easily.

Corollary 1. *Let $R_1, \ldots, R_k, S_1, \ldots, S_k, Q_1, Q_2$ be planar point sets with $|R_i| = r$, $|S_i| = s$ for all $i \in \{1, \ldots, k\}$, $|Q_1| = q_1$ and $|Q_2| = q_2$. Let C be a subset of $\cup_{i=1}^{k}(R_i+S_i)+Q_1+Q_2$. If C is convexly independent, then $|C| = \mathcal{O}(kr^{2/3}s^{2/3}q_1^{2/3}q_2^{2/3} + krq_1 + ksq_2)$.*

Proof. We observe that $\cup_{i=1}^{k}(R_i + S_i) + Q_1 + Q_2 = \cup_{i=1}^{k}((R_i + Q_1) + (S_i + Q_2))$. Therefore, we partition C into k convexly independent disjoint sets $C_1, \ldots, C_k$ such that $C_i \subset (R_i + Q_1) + (S_i + Q_2)$ for all $i \in \{1, \ldots, k\}$. Since $|R_i + Q_1| = rq_1$ and $|S_i + Q_2| \leqslant sq_2$, by Theorem 3, we get that $|C_i| = \mathcal{O}(r^{2/3}s^{2/3}q_1^{2/3}q_2^{2/3} + rq_1 + sq_2)$ and the result follows. □

Theorem 4. *Consider a polynomial $f \in \mathbb{R}^+[X]$ of degree d of the form*

$$f = \sum_{i=1}^{k} g_i h_i,$$

where $g_i, h_i \in \mathbb{R}^+[X]$, the g_i have at most r monomials and the h_i have at most s monomials. If f satisfies the Kurtz condition, then $d = \mathcal{O}(kr^{2/3}s^{2/3}\log^{2/3}(kr) + k(r+s)\log^{1/2}(kr))$.

Proof. We write $f = \sum_{i=0}^{d} c_i X^i$, where $c_i > 0$ for all $i \in \{1, \ldots, d\}$ and $c_0 \geqslant 0$. Since f satisfies the Kurtz condition, setting $\epsilon := \log(4)/2$ we get that

$$2\log(c_i) > \log(c_{i-1}) + \log(c_{i+1}) + 2\epsilon. \tag{4}$$

for every $i \geqslant 2$. For every $\delta_1, \ldots, \delta_d \in \mathbb{R}$, we set $C_{(\delta_1,\ldots,\delta_d)} := \{(i, \log(c_i) + \delta_i) \,|\, 1 \leqslant i \leqslant d\}$. We observe that (4) implies that $C_{(\delta_1,\ldots,\delta_d)}$ is convexly independent whenever $0 \leqslant \delta_i < \epsilon$ for all $i \in \{1, \ldots, d\}$.

We write $g_i = \sum_{j=1}^{r_i} a_{i,j} X^{\alpha_{i,j}}$ and $h_i = \sum_{j=1}^{s_i} b_{i,j} X^{\beta_{i,j}}$, with $r_i \leqslant r$, $s_i \leqslant s$ and $a_{i,j}, b_{i,j} > 0$ for all i, j. Then, $c_l = \sum_{i=1}^{k} (\sum_{\alpha_{i,j_1}+\beta_{i,j_2}=l} a_{i,j_1} b_{i,j_2})$. So, setting $M_l := \max\{a_{i,j_1} b_{i,j_2} \,|\, i \in \{1,\dots,k\}, \alpha_{i,j_1} + \beta_{i,j_2} = l\}$ for all $l \in \{1,\dots,d\}$, we have that $M_l \leqslant c_l \leqslant k r M_l$, so $\log(M_l) \leqslant \log(c_l) \leqslant \log(M_l) + \log(kr)$.

For every $l \in \{1,\dots,d\}$, we set

$$\lambda_l := \left\lceil \frac{\log(c_l) - \log(M_l)}{\epsilon} \right\rceil \text{ and } \delta_l := \log(M_l) + \lambda_l \epsilon - \log(c_l), \tag{5}$$

and have that $0 \leqslant \lambda_l \leqslant \lceil (\log(kr))/\epsilon \rceil$ and that $0 \leqslant \delta_l < \epsilon$.

Now, we consider the sets

- $R_i := \{(\alpha_{i,j}, \log(a_{i,j})) \,|\, 1 \leqslant j \leqslant r_i\}$ for $i = 1,\dots,k$,
- $S_i := \{(\beta_{i,j}, \log(b_{i,j})) \,|\, 1 \leqslant j \leqslant s_i\}$ for $i = 1,\dots,k$,
- $Q := \{(0, \lambda\epsilon) \,|\, 0 \leqslant \lambda \leqslant \lceil \log(kr)/\epsilon \rceil\}$,
- $Q_1 := \{(0, \mu\epsilon) \,|\, 0 \leqslant \mu \leqslant \lceil \sqrt{\log(kr)/\epsilon} \rceil\}$, and
- $Q_2 := \{(0, \nu\lceil \sqrt{\log(kr)/\epsilon} \rceil \epsilon) \,|\, 0 \leqslant \nu \leqslant \lceil \sqrt{\log(kr)/\epsilon} \rceil\}$.

If $(0, \lambda\epsilon) \in Q$, then there exist μ and ν such that $\lambda = \nu\lceil \sqrt{\log(kr)/\epsilon} \rceil + \mu$ where $\mu, \nu \leqslant \lceil \sqrt{\log(kr)/\epsilon} \rceil$. We have,

$$(0, \lambda\epsilon) = (0, \nu\lceil \sqrt{\log(kr)/\epsilon} \rceil \epsilon) + (0, \mu\epsilon) \in Q_1 + Q_2,$$

so $Q \subset Q_1 + Q_2$. Then, we claim that $C_{(\delta_1,\dots,\delta_d)} \subset \cup_{i=1}^{k} (R_i + S_i) + Q$. Indeed, for all $l \in \{1,\dots,d\}$, by (5),

$$\log(c_l) + \delta_l = \log(M_l) + \lambda_l \epsilon = \log(a_{i,j_1}) + \log(b_{i,j_2}) + \lambda_l \epsilon$$

for some $i \in \{1,\dots,k\}$ and some j_1, j_2 such that $\alpha_{i,j_1} + \beta_{i,j_2} = l$; thus

$$(l, \log(c_l) + \delta_l) = (\alpha_{i,j_1}, \log(a_{i,j_1})) + (\beta_{i,j_2}, \log(b_{i,j_1})) + (0, \lambda_l \epsilon) \in \cup_{i=1}^{k} (R_i + S_i) + Q.$$

Since $C_{(\delta_1,\dots,\delta_d)}$ is a convexly independent set of d elements contained in $\cup_{i=1}^{k} (R_i + S_i) + Q_1 + Q_2$, a direct application of Corollary 1 yields the result. □

From this result it is easy to derive an upper bound for the general case, where we have the products of $m \geqslant 2$ polynomials. If suffices to divide the m factors into two groups of approximately $m/2$ factors, and in each group we expand the product by brute force.

Proof of Theorem 1. We write each of the k products as a product of two polynomials $G_i := \prod_{j=1}^{\lfloor m/2 \rfloor} f_{i,j}$ and $H_i := \prod_{j=\lfloor m/2 \rfloor + 1}^{m} f_{i,j}$. We can now apply Theorem 4 to $f = \sum_{i=1}^{k} G_i H_i$ with $r = t^{\lfloor m/2 \rfloor}$ and $s = t^{m - \lfloor m/2 \rfloor}$ and we get the result. □

Remark 1. We observe that the role of the constant 4 in the Kurtz condition can be played by any other constant $\tau > 1$ in order to obtain the conclusion of Theorem 1, i.e., we obtain the same result for $f = \sum_{i=0}^{d} a_i X^i$ satisfying that $a_i^2 > \tau a_{i-1} a_{i+1}$ for all $i \in \{1, \ldots, d-1\}$. For proving this it suffices to replace the value $\epsilon = \log(4)/2$ by $\epsilon = \log(\tau)/2$ in the proof of Theorem 4 to conclude this more general result.

For $f = gh$ with $g, h \in \mathbb{R}^+[X]$ with at most t monomials, whenever f satisfies the Kurtz condition, then f has only real (and distinct) roots and so do g and h. As a consequence, both g and h satisfy (1) with strict inequalities and we derive that $d \leqslant 2t$. Nevertheless, in the similar setting where $f = gh + x^i$ for some $i > 0$, the same argument does not apply and a direct application of Theorem 1 yields $d = \mathcal{O}(t^{4/3} \log^{2/3}(t))$, a bound which seems to be very far from optimal.

Comparison with the setting of Newton polygons

A result similar to Theorem 1 was obtained in [12] for the Newton polygons of bivariate polynomials. Recall that the Newton polygon of a polynomial $f(X, Y)$ is the convex hull of the points (i, j) such that the monomial $X^i Y^j$ appears in f with a nonzero coefficient.

Theorem 5 (Koiran-Portier-Tavenas-Thomassé). *Consider a bivariate polynomial of the form*

$$f(X, Y) = \sum_{i=1}^{k} \prod_{j=1}^{m} f_{i,j}(X, Y) \tag{6}$$

where $m \geqslant 2$ and the $f_{i,j}$ have at most t monomials. The Newton polygon of f has $O(kt^{2m/3})$ edges.

In the setting of Newton polygons, the main issue is how to deal with the cancellations arising from the addition of the k products in (6). Two monomials of the form cX^iY^j with the same pair (i, j) of exponents but oppositive values of the coefficient c will cancel, thereby deleting the point (i, j) from the Newton polygon.

In the present paper we associate to the monomial cX^i with $c > 0$ the point $(i, \log c)$. There are no cancellations since we only consider polynomials $f_{i,j}$ with nonnegative coefficients in Theorems 1 and 4. However, the addition of two monomials $cX^i, c'X^i$ with the same exponent will "move" the corresponding point along the coefficient axis. By contrast, in the setting of Newton polygons points can be deleted but cannot move. In the proof of Theorem 4 we deal with the issue of "movable points" by an approximation argument, using the fact that the constant $\epsilon = \log(4)/2 > 0$ gives us a little bit of slack.

3 A stronger log-concavity condition

The objective of this section is to improve the bound provided in Theorem 1 when $f = \sum_{i=0}^{d} a_i X^i \in \mathbb{R}^+[x]$ satisfies a stronger log-concavity condition, namely, when $a_i^2 > d^{2d} a_{i-1} a_{i+1}$ for all $i \in \{1, \ldots, d-1\}$.

To prove this bound, we make use of the following well-known lemma (a reference and similar results for polytopes in higher dimension can be found in [8]). For completeness, we provide a short proof.

Lemma 1. *If $R_1, \ldots, R_s$ are planar sets and $|R_i| = r_i$ for all $i \in \{1, \ldots, s\}$, then the convex hull of $R_1 + \cdots + R_s$ has at most $r_1 + \cdots + r_s$ vertices.*

Proof. We denote by k_i the number of vertices of the convex hull of R_i. Clearly $k_i \leqslant r_i$. Let us prove that the convex hull of $R_1 + \cdots + R_s$ has at most $k_1 + \cdots + k_s$ vertices. Assume that $s = 2$. We write $R_1 = \{a_1, \ldots, a_{r_1}\}$, then $a_i \in R_1$ is a vertex of the convex hull of R_1 if and only if there exists $w \in S^1$ (the unit Euclidean sphere) such that $w \cdot a_i > w \cdot a_j$ for all $j \in \{1, \ldots, r_1\} \setminus \{i\}$. Thus, R_1 induces a partition of S^1 into k_1 half-closed intervals. Similarly, R_2 induces a partition of S^1 into k_2 half-closed intervals. Moreover, these two partitions induce a new one on S^1 with at most $k_1 + k_2$ half-closed intervals; these intervals correspond to the vertices of $R_1 + R_2$ and; thus, there are at most $k_1 + k_2$. By induction we get the result for any value of s. □

Proposition 1. *Consider a polynomial $f = \sum_{i=0}^{d} a_i X^i \in \mathbb{R}^+[X]$ of the form*

$$f = \sum_{i=1}^{k} \prod_{j=1}^{m} f_{i,j}$$

where the $f_{i,j} \in \mathbb{R}^+[x]$. If f satisfies the condition

$$a_i^2 > k^2 d^{2m} a_{i-1} a_{i+1},$$

then there exists a polynomial $f_{i,j}$ with at least d/km monomials.

Proof. Every polynomial $f_{i,j} := \sum_{l=0}^{d_{i,j}} c_{i,j,l}\, X^l$, where $d_{i,j}$ is the degree of $f_{i,j}$, corresponds to a planar set

$$R_{i,j} := \{(l, \log(c_{i,j,l})) \,|\, c_{i,j,l} > 0\} \subset \mathbb{R}^2.$$

We set, $C_{i,l} := \max\{0, \prod_{r=1}^{m} c_{i,r,l_r} \,|\, l_1 + \cdots + l_m = l\}$, for all $i \in \{1, \ldots, k\}$, $l \in \{0, \ldots, d\}$, and $C_l := \max\{C_{i,l} \,|\, 1 \leqslant i \leqslant k\}$ for all $l \in \{0, \ldots, d\}$. Since the polynomials $f_{i,j} \in \mathbb{R}^+[X]$ and

$$a_l = \sum_{i=1}^{k} \left(\sum_{l_1 + \cdots + l_m = l} \prod_{r=1}^{m} c_{i,r,l_r} \right)$$

for all $l \in \{0, \ldots, d\}$, we derive the following two properties:

- $C_l \leqslant a_l \leqslant k d^m C_l$ for all $l \in \{0, \ldots, d\}$,
- either $C_{i,l} = 0$ or $(l, \log(C_{i,l})) \in R_{i,1} + \cdots + R_{i,m}$ for all $i \in \{1, \ldots, k\}$, $l \in \{0, \ldots, d\}$. Since $a_l > 0$ for all $l \in \{1, \ldots, d\}$, we have that $C_l > 0$ and $(l, \log(C_l)) \in \bigcup_{i=1}^{k} (R_{i,1} + \cdots + R_{i,m})$

We claim that the points in the set $\{(l, \log(C_l)) \mid 1 \leqslant l \leqslant d\}$ belong to the upper convex envelope of $\bigcup_{i=1}^{k}(R_{i,1} + \cdots + R_{i,m})$. Indeed, if $(a, \log(b)) \in \bigcup_{i=1}^{k}(R_{i,1} + \cdots + R_{i,m})$, then $a \in \{0, \ldots, d\}$ and $b \leqslant C_a$; moreover, for all $l \in \{1, \ldots, d-1\}$, we have that

$$C_l^2 \geqslant a_l^2/(k^2 d^{2m}) > a_{l-1}\, a_{l+1} \geqslant C_{l-1} C_{l+1}.$$

Hence, there exist $i_0 \in \{1, \ldots, k\}$ and $L \subset \{1, \ldots, d\}$ such that $|L| \geqslant d/k$ and $C_l = C_{i_0,l}$ for all $l \in L$. Since the points in $\{(l, \log(C_l)) \mid 1 \leqslant l \leqslant d\}$ belong to the upper convex envelope of $\bigcup_{i=1}^{k}(R_{i,1} + \cdots + R_{i,m})$ we easily get that the set $\{(l, \log(C_{i_0,l})) \mid l \in L\}$ is a subset of the vertices in the convex hull of $R_{i_0,1} + \cdots + R_{i_0,m}$. By Lemma 1, we get that there exists j_0 such that $|R_{i_0,j_0}| \geqslant |L|/m \geqslant d/km$ points. Finally, we conclude that f_{i_0,j_0} involves at least d/km monomials. □

Proof of Theorem 2. If $d \leqslant k$ or $d \leqslant m$, then $d \leqslant kmt$. Otherwise, $d^{2d} > k^2 d^{2(d-1)} \geqslant k^2 d^{2m}$ and, thus, f satisfies (3). A direct application of Proposition 1 yields the result. □

4 Applications to Complexity Theory

We first recall some standard definitions from algebraic complexity theory (see e.g. [2] or [15] for more details). Fix a field K. The elements of the complexity class VP are sequences (f_n) of multivariate polynomials with coefficients from K. By definition, such a sequence belongs to VP if the degree of f_n is bounded by a polynomial function of n and if f_n can be evaluated in a polynomial number of arithmetic operations (additions and multiplications) starting from variables and from constants in K. This can be formalized with the familiar model of *arithmetic circuits*. In such a circuit, input gates are labeled by a constant or a variable and the other gates are labeled by an arithmetic operation (addition or multiplication). In this paper we take $K = \mathbb{R}$ since there is a focus on polynomials with nonnegative coefficients. An arithmetic circuit is *monotone* if input gates are labeled by nonnegative constants only.

A family of polynomials belongs to the complexity class VNP if it can be obtained by summation from a family in VP. More precisely, $f_n(\overline{x})$ belongs to VNP if there exists a family $(g_n(\overline{x}, \overline{y}))$ in VP and a polynomial p such that the tuple of variables $\overline{y}$ is of length $l(n) \leqslant p(n)$ and

$$f_n(\overline{x}) = \sum_{\overline{y} \in \{0,1\}^{l(n)}} g_n(\overline{x}, \overline{y}).$$

Note that this summation over all boolean values of $\overline{y}$ may be of exponential size. Whether the inclusion $\mathsf{VP} \subseteq \mathsf{VNP}$ is strict is a major open problem in algebraic complexity.

Valiant's criterion [2, 15] shows that "explicit" polynomial families belong to VNP. One version of it is as follows.

Lemma 2. *Suppose that the function $\phi : \{0,1\}^* \to \{0,1\}$ is computable in polynomial time. Then the family (f_n) of multilinear polynomials defined by*

$$f_n = \sum_{e \in \{0,1\}^n} \phi(e) x_1^{e_1} \cdots x_n^{e_n}$$

belongs to VNP.

Note that more general versions of Valiant's criterion are know. One may allow polynomials with integer rather than 0/1 coefficients [2], but in Theorem 7 below we will only have to deal with 0/1 coefficients. Also, one may allow f_n to depend on any (polynomially bounded) number of variables rather than exactly n variables and in this case, one may allow the algorithm for computing the coefficients of f_n to take as input the index n in addition to the tuple e of exponents (see [9], Theorem 2.3).

Reduction of arithmetic circuits to depth 4 is an important ingredient in the proof of the forthcoming results. This phenomenon was discovered by Agrawal and Vinay [1]. Here we will use it under the form of [14], which is an improvement of [11]. We will also need the fact that if the original circuit is monotone, then the resulting depth 4 circuit is also monotone (this is clear by inspection of the proof in [14]). Recall that a depth 4 circuit is a sum of products of sums of products of inputs; sum gates appear on layers 2 and 4 and product gates on layers 1 and 3. All gates may have arbitrary fan-in.

Lemma 3. *Let C be an arithmetic circuit of size $s > 1$ computing a v-variate polynomial of degree d. Then, there is an equivalent depth 4 circuit Γ of size $2^{\mathcal{O}\left(\sqrt{d \log(ds) \log(v)}\right)}$ with multiplication gates at layer 3 of fan-in $\mathcal{O}(\sqrt{d})$. Moreover, if C is monotone, then Γ can also be chosen to be monotone.*

We will use this result under the additional hypothesis that d is polynomially bounded by the number of variables v. In this setting, since $v \leqslant s$, we get that the resulting depth 4 circuit Γ provided by Lemma 3 has size $s^{\mathcal{O}(\sqrt{d})}$.

Before stating the main results of this section, we construct an explicit family of log-concave polynomials.

Lemma 4. *Let $n, s \in \mathbb{Z}^+$ and consider $g_{n,s}(X) := \sum_{i=0}^{2^n-1} a_i X^i$, with*

$$a_i := 2^{si(2^n-i-1)} \text{ for all } i \in \{0, \ldots, 2^n - 1\}.$$

Then, $a_i^2 > 2^s \, a_{i-1} \, a_{i+1}$.

Proof. Take $i \in \{1, \ldots, 2^n - 2\}$, we have that

$$\begin{aligned}
\log\left(2^s a_{i-1} a_{i+1}\right) &= s + s2^n(i-1) - s(i-1)i + s2^n(i+1) - s(i+1)(i+2) \\
&= 2s2^n i - 2si(i+1) - s \\
&< 2s2^n i - 2si(i+1) \\
&= \log(a_i^2).
\end{aligned}$$

□

In the next theorem we start from the family $g_{n,s}$ of Lemma 4 and we set $s = n2^{n+1}$.

Theorem 6. *Let $(f_n) \in \mathbb{N}[X]$ be the family of polynomials $f_n(x) = g_{n,n2^{n+1}}(x)$.*

(i) f_n has degree $2^n - 1$ and satisfies the log-concavity condition (3).
(ii) If $\mathsf{VP} = \mathsf{VNP}$, f_n can be written under form (2) *with $k = n^{O(\sqrt{n})}$, $m = O(\sqrt{n})$ and $t = n^{O(\sqrt{n})}$.*

Proof. It is clear that $f_n \in \mathbb{N}[X]$ has degree $2^n - 1$ and, by Lemma 4, f_n satisfies (3).

Consider now the related family of bivariate polynomials $g_n(X, Y) = \sum_{i=0}^{2^n-1} X^i Y^{e(n,i)}$, where $e(n, i) = si(2^n - i - 1)$. One can check in time polynomial in n whether a given monomial $X^i Y^j$ occurs in g_n: we just need to check that $i < 2^n$ and that $j = e(n, i)$. By mimicking the proof of Theorem 1 in [12] and taking into account Lemma 3 we get that, if $\mathsf{VP} = \mathsf{VNP}$, one can write

$$g_n(X, Y) = \sum_{i=1}^{k} \prod_{j=1}^{m} g_{i,j,n}(X, Y) \tag{7}$$

where the bivariate polynomials $g_{i,j,n}$ have $n^{O(\sqrt{n})}$ monomials, $k = n^{O(\sqrt{n})}$ and $m = O(\sqrt{n})$. Performing the substitution $Y = 2$ in (7) yields the required expression for f_n. □

We believe that there is in fact no way to write f_n under form (2) so that the parameters k, m, t satisfy the constraints $k = n^{O(\sqrt{n})}$, $m = O(\sqrt{n})$ and $t = n^{O(\sqrt{n})}$. By part (ii) of Theorem 6, a proof of this would separate VP from VNP. The proof of Theorem 7 below shows that our belief is actually correct in the special case where the polynomials $f_{i,j}$ in (2) have nonnegative coefficients.

The main point of Theorem 7 is to present an unconditional lower bound for a polynomial family (h_n) in VNP derived from (f_n). Note that (f_n) itself is not in VNP since its degree is too high. Recall that

$$f_n(X) := \sum_{i=0}^{2^n-1} 2^{2n2^n i(2^n-i-1)} X^i. \tag{8}$$

To construct h_n we write down in base 2 the exponents of "2" and "X" in (8). More precisely, we take h_n of the form:

$$h_n := \sum_{\substack{\alpha \in \{0,1\}^n \\ \beta \in \{0,1\}^{4n}}} \lambda(n, \alpha, \beta)\, X_0^{\alpha_0} \cdots X_{n-1}^{\alpha_{n-1}} Y_0^{\beta_0} \cdots Y_{4n-1}^{\beta_{4n-1}}, \tag{9}$$

where $\alpha = (\alpha_0, \ldots, \alpha_{n-1})$, $\beta = (\beta_0, \ldots, \beta_{4n-1})$ and $\lambda(n, \alpha, \beta) \in \{0, 1\}$; we set $\lambda(n, \alpha, \beta) = 1$ if and only if $\sum_{j=0}^{4n-1} \beta_j 2^j = 2n2^n i(2^n - i - 1) < 2^{4n}$, where $i := \sum_{k=0}^{n-1} \alpha_{i,k} 2^k$. By construction, we have:

$$f_n(X) = h_n(X^{2^0}, X^{2^1}, \ldots, X^{2^{n-1}}, 2^{2^0}, 2^{2^1}, \ldots, 2^{2^{4n-1}}). \tag{10}$$

This relation will be useful in the proof of the following lower bound theorem.

Theorem 7. *The family (h_n) in* (9) *is in* VNP. *If (h_n) is computed by depth 4 monotone arithmetic circuits of size $s(n)$, then $s(n) = 2^{\Omega(n)}$. If (h_n) is computed by monotone arithmetic circuits of size $s(n)$, then $s(n) = 2^{\Omega(\sqrt{n})}$. In particular, (h_n) cannot be computed by monotone arithmetic circuits of polynomial size.*

Proof. Note that h_n is a polynomial in $5n$ variables, of degree at most $5n$, and its coefficients $\lambda(n, \alpha, \beta)$ can be computed in polynomial time. Thus, by Valiant's criterion we conclude that $(h_n) \in$ VNP.

Assume that (h_n) can be computed by depth 4 monotone arithmetic circuits of size $s(n)$. Using (10), we get that $f_n = \sum_{i=1}^{k} \prod_{j=1}^{m} f_{i,j}$ where $f_{i,j} \in \mathbb{R}^+[X]$ have at most t monomials and k, m, t are $\mathcal{O}(s(n))$. Since the degree of f_n is $2^n - 1$, by Theorem 2, we get that $2^n - 1 \leqslant kmt$. We conclude that $s(n) = 2^{\Omega(n)}$.

To complete the proof of the theorem, assume that (h_n) can be computed by monotone arithmetic circuits of size $s(n)$. By Lemma 3, it follows that the polynomials h_n are computable by depth 4 monotone circuits of size $s'(n) := s(n)^{\mathcal{O}(\sqrt{n})}$. Therefore $s'(n) = 2^{\Omega(n)}$ and we finally get that $s(n) = 2^{\Omega(\sqrt{n})}$. □

Lower bounds for monotone arithmetic circuits have been known for a long time (see for instance [7, 16]). Theorem 7 provides yet another example of a polynomial family which is hard for monotone arithmetic circuits, with an apparently new proof method.

5 Discussion

As explained in the introduction, log-concavity plays a role in the study of real roots of polynomials. In [10] bounding the number of real roots of sums of products of sparse polynomials was suggested as an approach for separating VP from VNP. Hrubeš [5] suggested to bound the multiplicities of roots, and [12] to bound the number of edges of Newton polygons of bivariate polynomials.

Theorem 6 provides another plausible approach to VP $\neq$ VNP: it suffices to show that if a polynomial $f \in \mathbb{R}^+[X]$ under form (2) satisfies the Kurtz condition or the stronger log-concavity condition (3) then its degree is bounded by a "small" function of the parameters k, m, t. A degree bound which is polynomial bound in k, t and 2^m would be good enough to separate VP from VNP. Theorem 1 improves on the trivial kt^m upper bound when f satisfies the Kurtz condition, but certainly falls short of this goal: not only is the bound on $\deg(f)$ too coarse, but we would also need to allow negative coefficients in the polynomials $f_{i,j}$. Theorem 2 provides a polynomial bound on k, m and t under a stronger log-concavity condition, but still needs the extra assumption that the coefficients in the polynomials $f_{i,j}$ are nonnegative. The unconditional lower bound in Theorem 7 provides a "proof of concept" of this approach for the easier setting of monotone arithmetic circuits.

References

1. Manindra Agrawal and V. Vinay. Arithmetic circuits: A chasm at depth four. In *49th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2008, October 25-28, 2008, Philadelphia, PA, USA*, pages 67–75, 2008.
2. Peter Bürgisser. *Completeness and reduction in algebraic complexity theory*, volume 7 of *Algorithms and Computation in Mathematics*. Springer-Verlag, Berlin, 2000.
3. Friedrich Eisenbrand, János Pach, Thomas Rothvoß, and Nir B. Sopher. Convexly independent subsets of the Minkowski sum of planar point sets. *Electron. J. Combin.*, 15(1):Note 8, 4, 2008.
4. G. H. Hardy, J. E. Littlewood, and G. Pólya. *Inequalities*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, 1988. Reprint of the 1952 edition.
5. P. Hrubes. A note on the real τ-conjecture and the distribution of complex roots. *Theory of Computing*, 9(10):403–411, 2013. eccc.hpi-web.de/report/2012/121/.
6. J. I. Hutchinson. On a remarkable class of entire functions. *Trans. Amer. Math. Soc.*, 25(3):325–332, 1923.
7. Mark Jerrum and Marc Snir. Some exact complexity results for straight-line computations over semirings. *Journal of the ACM (JACM)*, 29(3):874–897, 1982.
8. Menelaos I Karavelas and Eleni Tzanaki. The maximum number of faces of the Minkowski sum of two convex polytopes. In *Proceedings of the twenty-third annual ACM-SIAM Symposium on Discrete Algorithms*, pages 11–28, 2012.
9. P. Koiran. Valiant's model and the cost of computing integers. *Computational Complexity*, 13:131–146, 2004.
10. P. Koiran. Shallow circuits with high-powered inputs. In *Proc. Second Symposium on Innovations in Computer Science (ICS 2011)*, 2011. arxiv.org/abs/1004.4960.
11. Pascal Koiran. Arithmetic circuits: The chasm at depth four gets wider. *Theoretical Computer Science*, 448(0):56 – 65, 2012.
12. Pascal Koiran, Natacha Portier, Sébastien Tavenas, and Stéphan Thomassé. A τ-conjecture for Newton polygons. *Foundations of Computational Mathematics*, pages 1–13, 2014.
13. David C. Kurtz. A sufficient condition for all the roots of a polynomial to be real. *Amer. Math. Monthly*, 99(3):259–263, 1992.
14. S. Tavenas. Improved bounds for reduction to depth 4 and depth 3. In *Proc. 38th International Symposium on Mathematical Foundations of Computer Science (MFCS)*, 2013.
15. L. G. Valiant. Completeness classes in algebra. In *Proceedings of the Eleventh Annual ACM Symposium on Theory of Computing*, STOC '79, pages 249–261, New York, NY, USA, 1979. ACM.
16. Leslie G Valiant. Negation can be exponentially powerful. In *Proceedings of the eleventh annual ACM symposium on Theory of computing*, pages 189–196. ACM, 1979.