Electronic Colloquium on Computational Complexity, Report No. 53 (2025)

On Approximate Symmetric Polynomials and Tightness of Homogenization Results

Amir Shpilka*

Abstract

Motivated by questions concerning the multilinear and homogeneous complexity of the elementary symmetric polynomials, we prove the following results:

We first show that by making small modifications to the nonzero coefficients of the degree-K, N-variate elementary symmetric polynomial $\sigma_{N,K}$, one obtains a polynomial that can be computed by a monotone formula of size $K^{O(\log K)} \cdot N$.

As a corollary, we show that a result of Raz [Raz13] concerning the homogenization of algebraic multilinear or monotone formulas is tight.

Another corollary is that the monotone bounded rigidity of the inclusion matrix between K-subsets and $N - K$ subsets of a universe of size N is small.

*This research was co-funded by the European Union by the European Union (ERC, EACTP, 101142020), the Israel Science Foundation (grant number 514/20) and the Len Blavatnik and the Blavatnik Family Foundation. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

ISSN 1433-8092

1 Introduction

This work is motivated by questions concerning the homogenization of algebraic formulas and the multilinear homogeneous formula complexity of elementary symmetric polynomials. Results related to the homogenization and multilinearization of these polynomials, as well as generalized symmetric polynomials, have played a crucial role in recent lower bound proofs for small-depth circuits [LST24]. Understanding their complexity is therefore of significant importance.

The first nontrivial result on the homogenization of algebraic formulas was established by Raz [Raz13].

Theorem 1.1 ([Raz13]). *Let Φ be a formula of size s over a field $\mathbb{F}$, computing an N-variate homogeneous polynomial $f \in \mathbb{F}[\mathbf{x}]$ of degree K. Then there exists a homogeneous formula Φ' over $\mathbb{F}$ computing f, of size* poly $\left(s, \binom{K+\log s}{\log s}\right)$.

An analysis of the proof of Theorem 1.1 reveals that if the original formula is multilinear (and/or monotone) then the resulting homogeneous formula is multilinear (and/or monotone) as well.

Theorem 1.2 (Implicit in [Raz13]). *Let $\mathbb{F}$ be a field.*[1] *Let Φ be a multilinear (monotone) formula of size s computing an N-variate homogeneous multilinear (monotone) polynomial f of degree K over $\mathbb{F}$. Then, there is a homogeneous multilinear (monotone) formula Φ' computing f of size* poly $\left(s, \binom{K+\log s}{\log s}\right)$ *over $\mathbb{F}$.*

In [FLST24] Fournier, Limaye, Srinivasan and Tavenas obtained the following improvement to Theorem 1.1, for a certain parameter's regime.

Theorem 1.3 (Theorem 5 in [FLST24]). *Let $\mathbb{F}$ be a field of characteristic zero or of large enough characteristic. Assume that Φ is an algebraic formula (with unbounded fan-in) of size s and depth Δ computing a homogeneous polynomial f of degree K over $\mathbb{F}$. Then f is also computed by a homogeneous formula Φ' of size $s \cdot K^{O(\Delta+\log K)}$.*

This improves upon Theorem 1.1 when $\Delta = o(\log s)$ and $K = s^{o(1)}$. However, we note that the proof technique of [FLST24] is inherently non-multilinear and non-monotone, even when starting from a monotone multilinear formula.

The following questions regarding the tightness of Theorem 1.1, 1.2 and 1.3 are still open.

Question 1.4. *Can the homogenization bounds of Raz or Fournier et al. be improved when the original formula is multilinear or monotone?*

Question 1.5. *Are there nontrivial cases where the bounds in Theorem 1.1, Theorem 1.2, or Theorem 1.3 are tight?*

Another motivation comes from trying to understand the complexity of multilinear homogeneous formulas for the elementary symmetric polynomials. The breakthrough result of Limaye, Srinivasan and Tavenas [LST24] relies on efficient set-multilinearization and homogenization, in small depth, of (extended) elementary symmetric polynomials. Hrubeš and Yehudayoff obtained the following results.

Theorem 1.6 ([HY11]). *Let $\sigma_{N,K}(\mathbf{x}) = \sum_{T \in \binom{[N]}{K}} \prod_{i \in T} x_i$.*

1. *Over any field $\mathbb{F}$, such that char$(\mathbb{F}) > K$, $\sigma_{N,K}$ has a* homogeneous *formula of size $K^{O(\log K)} \cdot N$.*
2. *Over any field $\mathbb{F}$, every* homogeneous multilinear *formula computing $\sigma_{N,K}$ has size at least $K^{\Omega(\log K)} \cdot N$.*
3. *Over $\mathbb{F} = \mathbb{Q}$, , for $K = O(\log N)$, $\sigma_{N,K}$ can be computed by a* monotone *formula of size* poly(N).

Note that the lower bound on the multilinear homogeneous size is essentially the same as the upper bound on the homogeneous formula size. This motivated Hrubeš and Yehudayoff to ask whether $\sigma_{N,K}$ has a homogeneous multilinear formula of size poly$(N)K^{O(\log K)}$. An even stricter restriction than homogeneity and multilinearity is monotonicity.

Question 1.7. *Can the monotone formula complexity of the elementary symmetric polynomial $\sigma_{N,K}$ be upper bounded by* poly$(N)K^{O(\log K)}$*?*

[1] When speaking of monotone formulae we only consider subfields of $\mathbb{R}$.

Hrubeš and Yehudayoff say that two polynomials are *weakly equivalent* if they have the same support. They mention that the proof of the lower bound in Theorem 1.6(2) actually holds for every polynomial which is weakly equivalent to $\sigma_{N,K}$. In other words, every polynomial that contains exactly all the multilinear monomials of degree K, cannot be computed by a homogeneous multilinear formula of size smaller than $K^{O(\log K)} \cdot N$.

Theorem 1.8 ([HY11]). *Let $f = \sum_{T \in \binom{[N]}{K}} c_T \cdot \prod_{i \in T} x_i$, where $c_T \neq 0$ for all T, be a polynomial in $\mathbb{F}[\mathbf{x}]$. Then, every homogeneous multilinear formula computing f has size at least $K^{\Omega(\log K)} \cdot N$.*

1.1 Results

We construct a polynomial $f_{N,K}$ with the same support as $\sigma_{N,K}$, whose nonzero coefficients lie in the interval $[0.99, 1]$. We prove that $f_{N,K}$ can be computed by a multilinear formula of size $\mathrm{poly}(N, K)$. Since the lower bound of [HY11] applies to $f_{N,K}$, it follows that the homogenization technique of Raz [Raz13] is tight for multilinear formula homogenization.

This result provides a *negative answer* to Question 1.4 and a *positive answer* to Question 1.5. Moreover, the monotone formula complexity of $f_{N,K}$ matches (up to polynomial blowup) the homogeneous formula complexity of $\sigma_{N,K}$. This implies that in order to disprove Question 1.7, one would need to prove lower bounds that fundamentally rely on all coefficients being 1.

In [Yeh19], Yehudayoff proved a monotone lower bound that depends on the actual coefficients, rather than just the support of the polynomial. This was later improved by Srinivasan [Sri20]. As both results are in the monotone setting, they offer hope for resolving Question 1.7, either negatively (via a lower bound) or positively (via an explicit construction). We are unaware of any other lower bound techniques that depend on the actual coefficients rather than the support.

We begin by showing the existence of such a polynomial $f_{N,K}$ that can be computed as the degree-K homogeneous component of a small depth-3 formula.

Theorem 1.9. *There is a polynomial $f_{N,K}$ with the same support as $\sigma_{N,K}$, and whose nonzero coefficients are in $[1 - 1/K, 1]$, that can be computed as the degree-K homogeneous component of a depth-3 formula, using only 0 and c as coefficients (for some $0 < c < 1$), with top fan-in $O(K^7 \log^2 N)$ and total degree $O(K^5 \log N)$.*

Applying Raz's result (Theorem 1.1) to each multiplication gate yields, for $K = \Omega(\log N)$, a homogeneous (and in fact monotone) formula of size $N \cdot K^{O(\log K)}$. This matches the lower bound of Hrubeš and Yehudayoff (Theorem 1.8) up to constants in the exponent, thereby demonstrating the tightness of Theorem 1.2. Furthermore, this implies that Theorem 1.3 cannot be improved under the additional requirement of multilinearity, even if the original formula is multilinear.

Theorem 1.10. *The polynomial $f_{N,K}$ from Theorem 1.9 can be computed by a monotone formula of size:*

- $N \cdot K^{O(\log K)}$, *for* $K \geqslant \log N$,
- $N \cdot \left(1 + \frac{K}{\log \log N}\right)^{O(\log \log N)}$, *for* $\log \log N < K < \log N$,
- $N \cdot \mathrm{poly}(\log N)$, *for* $K < \log \log N$.

Corollary 1.11. *Theorem 1.2 is tight when $K = s^{\Omega(1)}$. Moreover, Theorem 1.3 cannot be improved for multilinear formulas: for $K = \Omega(\log s)$ and $\Delta = 3$, there exists a multilinear formula Φ such that any homogeneous multilinear formula Φ' computing the same polynomial must have size $s \cdot K^{\Omega(\log K)}$.*

Recall that Hrubeš and Yehudayoff asked whether $\sigma_{N,K}$ can be computed by a homogeneous multilinear formula of size $\mathrm{poly}(N) \cdot K^{O(\log K)}$. Since any monotone formula for this polynomial is also homogeneous and multilinear, Theorem 1.10 shows that resolving this question negatively would require a lower bound that depends on the actual coefficients, not merely on the support.

By expanding each multiplication gate of the formula in Theorem 1.9 as a sum over $\binom{O(K^5 \log N)}{K}$ (monotone) multiplication gates of exact degree K, we obtain a monotone depth-3 formula for $f_{N,K}$ with $(K \log N)^{O(K)}$ multiplication gates. We note that an even more efficient construction is possible if we restrict to depth-3 formulas for a slightly modified polynomial:

Theorem 1.12. *There exists a polynomial $g_{N,K}$ with the same support as $\sigma_{N,K}$, such that each of its nonzero coefficients is in $[0.99, 1]$, and that can be computed by a monotone depth-3 formula with $2^{O(K)} \log N$ multiplication gates.*

1.1.1 Bounded monotone rigidity

Another outcome of our results is that while the partial derivative matrix of $\sigma_{N,K}$ has full rank, this is not the case for $f_{N,K}$ or $g_{N,K}$. As a consequence, since the coefficients of these polynomials are close to each other, we conclude that the partial derivative matrix of $\sigma_{N,K}$ has low bounded monotone rigidity. In other words, we can change each of its entries by a small amount and reduce its rank considerably.

Definition 1.13 (Rigidity)**.** The rigidity $R(r)(A)$ of a matrix A is defined as

$$R(r)(A) = \min_B \{|B|_0 \mid \operatorname{rank}(A - B) \leqslant r\}.$$

In words, $R(r)(A)$ equals the minimal number of entries of A that must be changed in order to reduce the rank of A to r.

Definition 1.14 (Bounded Rigidity)**.** The bounded rigidity $R(\theta, r)(A)$ of a matrix A is defined as

$$R(\theta, r)(A) = \min_B \{|B|_0 \mid \operatorname{rank}(A - B) \leqslant r,\ \|B\|_\infty \leqslant \theta\}.$$

That is, $R(\theta, r)(A)$ is the minimal number of entries of A that must be changed, by at most θ each, so that the rank of A drops to r.

Definition 1.15 (Bounded Monotone Rigidity)**.** The bounded monotone rigidity $R^{(M)}(\theta, r)(A)$ of a non-negative matrix A is defined as

$$R^{(M)}(\theta, r)(A) = \min_B \{|B|_0 \mid \operatorname{rank}(A - B) \leqslant r,\ \|B\|_\infty \leqslant \theta,\ \operatorname{support}(B) \subseteq \operatorname{support}(A)\}.$$

In other words, $R^{(M)}(\theta, r)(A)$ is the minimal number of nonzero entries of A that must be changed, by at most θ each, so that the rank of A drops to r.

It is clear from the above definitions that

$$R(r)(A) \leqslant R(\theta, r)(A) \leqslant R^{(M)}(\theta, r)(A).$$

Bounded rigidity has been studied in [KR98, Lok01, dW06, Ras16], primarily focusing on the bounded rigidity of the Walsh-Hadamard matrix.[2] Recall that for N a power of 2, $N = 2^n$, the Walsh-Hadamard matrix of order N is defined by $H_{S,T} = (-1)^{|S \cap T|}$ for rows and columns indexed by subsets $S, T \subset [N]$.

Theorem 1.16 ([dW06])**.** *The bounded rigidity of the Walsh-Hadamard matrix H satisfies*

$$R(\theta, r)(H) \geqslant \frac{N^2(N - r)}{2\theta N + r(\theta^2 + 2\theta)}.$$

For example, for $r = N/2$ and $\theta = 1/10$, Theorem 1.16 implies that more than N^2 entries must be changed. In other words, it is not possible to reduce the rank of H to $N/2$ by perturbing each entry by at most 0.1. However, Alman and Williams [AW17] showed that without the restriction on bounded changes, the rank can drop significantly, demonstrating that H does not exhibit full rigidity.

Theorem 1.17 ([AW17])**.** *For every $\varepsilon \in (0, 1/2)$,*

$$R\left(N^{1-\Theta(\varepsilon^2 \log(1/\varepsilon))}\right)(H) \leqslant N^{(1+\varepsilon)}.$$

[2] We can also discuss the monotone bounded rigidity of $H + J$, where J is the all-1 matrix, but the lower bounds in the bounded case are already very strong.

For a survey of recent results on rigidity, see [Ram20, Xie22].

Next, we consider the bounded monotone rigidity of the partial derivative matrix of $\sigma_{N,2K}$. Observe that it is actually the inclusion matrix $A_{N,K}$, defined as follows: The rows and columns of $A_{N,K}$ are indexed by all subsets of $[N]$ of size K, $\binom{[N]}{K}$. For two such subsets S and T, the (S,T) entry of $A_{N,K}$ is 1 if and only if $S \cap T = \emptyset$ (equivalently, if $S \subset [N] \setminus T$). It is known that this matrix has full rank [Got66]. We show that by making small changes to each nonzero entry of $A_{N,K}$, we can reduce its rank to $2^{O(K)} \cdot \log N$.

Theorem 1.18. *The partial derivative matrix of the polynomial $g_{N,2K}$ from Theorem 1.12, denoted $B_{N,K}$, satisfies the following two properties:*

1. $\operatorname{rank}(B_{N,K}) = 2^{O(K)} \cdot \log N$.
2. *For every two K-sets S, T, it holds that*

$$0 \leqslant (A_{N,K} - B_{N,K})_{S,T} < \frac{1}{100}.$$

Thus, for $K = o(N)$, we can reduce the rank of $A_{N,K}$ to $|A_{N,K}|^{o(1)}$ by changing each nonzero coordinate by at most $1/100$.

1.2 Proof idea

Our proof is based on an idea similar to color coding. We begin by hashing the variables into M buckets using a condenser (though random functions also work). For a fixed hash function, we sum the variables assigned to each bucket, yielding M linear forms $\ell_1, \ldots, \ell_M$, each with 0–1 coefficients and disjoint variable supports. Consider now the polynomial $\prod_{i=1}^{M}(\ell_i + 1)$. Its degree-$K$ homogeneous component collects all multilinear monomials corresponding to K-tuples of variables that are hashed injectively into distinct buckets.

Repeating this process across all seeds of the condenser ensures that each K-subset is mapped injectively by approximately the same number of seeds. This results in a monotone depth-3 $\Sigma\Pi\Sigma$ formula whose degree-K homogeneous component coefficient-wise approximates $\sigma_{N,K}$.

Finally, we apply Raz's homogenization theorem (Theorem 1.2) to each multiplication gate in the formula. This yields a monotone formula for a multilinear polynomial that approximates $\sigma_{N,K}$.

1.3 Related work

In addition to the work of Raz [Raz13] and Fournier et al. [FLST24] discussed earlier, a related line of research concerns the minimal Waring rank of polynomials that have the same support as $\sigma_{N,K}$ [Pra19].

Recall that a degree-K homogeneous polynomial $f \in \mathbb{F}[x_1, \ldots, x_n]$ is said to have Waring rank at most r if it can be written as a sum of K-th powers of r linear forms

$$f = \ell_1^K + \ldots + \ell_r^K.$$

In [Pra19], Pratt demonstrated that several techniques from parameterized algorithms—including color-coding [AYZ95, AG10], the group-algebra/determinant-sum method [Kou08, Wil09, Bjö14], and inclusion-exclusion—can be leveraged to reduce certain algorithmic problems to the task of upper bounding the Waring rank of a specific family of polynomials. As a result, obtaining tighter explicit upper bounds on the Waring rank of these polynomials directly translates into faster algorithms.

Pratt focused in particular on polynomials over the reals that share the same support as $\sigma_{N,K}$. He considered three natural settings: (1) polynomials with arbitrary coefficients, (2) polynomials with positive coefficients (over fields of characteristic zero), and (3) polynomials whose nonzero coefficients lie in the interval $[1 \pm \varepsilon]$.

Pratt's construction closely resembles our approach. He begins by mapping the variables into buckets using a family of functions that constitute a δ-*balanced* (N, K, M)-*splitter*.[3] Informally, such a splitter is a collection of hash functions from $[N]$ to $[M]$ with the guarantee that, for any K-subset $T \subseteq [N]$, the number of functions in the family that are injective on T lies within the range $[\delta \cdot c, c/\delta]$ (with δ close to 1).

For each hash function in the family, Pratt constructs the same linear forms that we do—namely, summing all variables mapped to the same bucket. He then applies a Waring rank decomposition to the polynomial $\sigma_{M,K}(\ell_1, \ldots, \ell_M)$ formed from these linear functions. Averaging the resulting polynomials (one from each hash function) and normalizing by $1/c$, he obtains a Waring rank decomposition of a polynomial that has the same support as $\sigma_{N,K}$, with all nonzero coefficients lying in the interval $[1 - \varepsilon, 1]$.

The key difference between Pratt's construction and ours lies in the choice of the universe size M into which the variables are hashed. Since Pratt's goal was to upper bound the Waring rank—and given that the Waring rank of a degree-K multilinear monomial is already 2^{K-1}—he hashed the variables into a relatively small number of buckets, roughly $M = 1.55 \cdot K$. Approximating $\sigma_{N,K}$ in this way requires $\Omega(\exp(K))$ hash functions, corresponding to the size of the balanced splitter family. In contrast, our primary concern was to keep the dependence on K subexponential—ideally polynomial. To achieve this, before applying Theorem 1.1, we hashed the variables into significantly more buckets ($M = 4K^5 \log N$), which allowed us to substantially reduce the number of required hash functions.

In the non-monotone setting Pratt obtained the following variant of Theorem 1.12.

Theorem 1.19 (Theorem 58 in [Pra19])**.** *There exists a polynomial with the same support as $\sigma_{N,K}$ and nonzero coefficients in the range $[1 \pm \varepsilon]$, which has a Waring rank at most $4.075^K/\varepsilon^2 \cdot \log N$.*

In comparison to Theorem 1.12, this result provides a Waring rank decomposition rather than a $\Sigma\Pi\Sigma$ formula. However, a Waring rank decomposition of a multilinear polynomial cannot be monotone, which means this approach does not yield a monotone depth-3 formula. Theorem 1.19 gives a polynomial that slightly improves upon the guarantee in Theorem 1.18, though we did not attempt to optimize the $2^{O(K)}$ term.

Finally, we observe that Pratt's upper bound on the Waring rank of a polynomial with non-negative coefficients and the same support as $\sigma_{N,K}$ depends solely on K.

Theorem 1.20 (Theorem 41 in [Pra19])**.** *There exists a polynomial with non-negative coefficients and the same support as $\sigma_{N,K}$, whose Waring rank is at most 6.75^K.*

As a corollary, we obtain the following version of Theorem 1.18:

Theorem 1.21. *The partial derivative matrix of the degree $2K$ polynomial given in Theorem 1.20, denoted $P_{N,K}$, satisfies:*

1. $\operatorname{rank}(P_{N,K}) = 2^{O(K)}$.
2. *For every two K-sets S, T, it holds that*

$$0 \leqslant (A_{N,K} - P_{N,K})_{S,T} < 1.$$

2 Preliminaries

We denote $[N] \triangleq 1, \ldots, N$. For a polynomial f, $\operatorname{Supp}(f)$ denotes the set of monomials in f with nonzero coefficients, and $|f|_0$ represents the size of $\operatorname{Supp}(f)$.

For a distribution X on $0,1^N$, we denote by $H_\infty(X)$ the min-entropy of X, i.e., $H_\infty(X) = -\log(\max_x \Pr[X = x])$. We say that X is a k-source if $H_\infty(X) \geqslant k$.

For an N-variate homogeneous polynomial f of degree K, its partial derivative matrix for order d derivatives is an $\binom{N+d-1}{d} \times \binom{N+K-d-1}{K-d}$ matrix. The rows correspond to N-variate monomials of degree d, and the columns correspond to N-variate monomials of degree $K - d$. The (m_1, m_2) entry in the matrix is the coefficient of the monomial m_2 in $\frac{\partial^d f}{\partial m_1}$, where m_1 and m_2 are monomials of appropriate degrees.

[3] For our parameter regime, a condenser also satisfies the properties of a δ-balanced (N, K, M)-splitter.

3 Proofs

The construction of $f_{N,K}$ relies on finding a family of hash functions such that every set of size K is hashed without collisions for most functions of the family. As in [RS22], we construct such hash family using condensers.

Definition 3.1. A function $C : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ is called a (k, ε, t) condenser if, for every distributions X on $\{0,1\}^n$ such that $H_\infty(X) \geqslant k$, the distribution $C(X, U_d)$ is ε-close (in L_1 norm) to some distribution Z on $\{0,1\}^m$ that satisfies $H_\infty(Z) \geqslant t$.

A simple counting argument gives the following result.

Fact 3.2. *For any $k \leqslant n$ there exists a (k, ε, t) condenser $C : \{0,1\}^n \times \{0,1\}^d \to \{0,1\}^m$ with parameters:*

$$d = \lceil \log(n) + \log(1/\varepsilon) \rceil, \quad m = \lceil t + \log(1/\varepsilon) \rceil, \quad \textit{and} \quad t = k + d.$$

Let C be the condenser from Fact 3.2 with parameters $n = \log N$, $k = \log K$, and $\varepsilon = 1/2K^2$.[4] With this choice of parameters we have that

$$2^d = n/\varepsilon = 2K^2 \log N \quad \text{and} \quad 2^m = 2^{k+d}/\varepsilon = 4K^5 \log N.$$

Claim 3.3 ([RS22, Claim 2.4])**.** *Let $X \subseteq \{0,1\}^n$ be a k-source. Then, for all but $\sqrt{\varepsilon}$ of the seeds $y \in \{0,1\}^d$, it holds that $C_y(X)$ is $\sqrt{\varepsilon}$-close to a k-source.*

Proof. The proof is an easy application of Markov's inequality. □

Corollary 3.4 ([RS22, Corollary 2.5])**.** *Let $I \subseteq [N] = \{0,1\}^n$ be a set of size $|I| \leqslant 2^k = K$. Then, with probability at least $1 - \sqrt{\varepsilon}$ over $y \in \{0,1\}^d$, the map C_y is injective on I.*

Proof. Let X be a random variable uniformly distributed over a set of size exactly 2^k that contains I. Let y be such that $C_y(X)$ is $\sqrt{\varepsilon}$-close to a k-source Z. Then, for all $z \in \{0,1\}^m$ we have

$$\Pr[C_y(X) = z] \leqslant \Pr[Z = z] + \sqrt{\varepsilon} \leqslant 2^{-k} + \sqrt{\varepsilon} < 2 \cdot 2^{-k},$$

where the third inequality follows from the choice of ε. In particular, this implies that no two elements of I were mapped to the same element z. □

Construction of $f_{N,K}$. In what follows, for $N = 2^n$, we identify the set $[N]$ with $\{0,1\}^n$.

Let C be the condenser guaranteed by Fact 3.2. Fix a seed $a \in \{0,1\}^d$. For every $z \in \{0,1\}^m$, define the set

$$S_{a,z} = \{i \in [N] = \{0,1\}^n \mid C(i, a) = z\}$$

and the linear function

$$\ell_{a,z} = \sum_{i \in S_{a,z}} x_i.$$

We then define $f_{N,K}$ as

$$f_{N,K}(x) = \frac{1}{2^d} \sum_{a \in \{0,1\}^d} \sigma_{2^m,K}\left(\ell_{a,z}(x) \mid z \in \{0,1\}^m\right).$$

The next claim summarizes the properties of our construction.

Claim 3.5 (Properties of $f_{N,K}$)**.** 1. *$S_{a,z} \cap S_{a,z'} = \emptyset$ for $z \neq z'$.*

2. *$f_{N,K}$ is a homogeneous multilinear polynomial of degree K.*

3. *The support of $f_{N,K}$ contains all the multilinear monomials of degree K, and only these. Moreover, each such monomial has a coefficient in the interval $[1 - 1/K, 1]$.*

[4] To avoid the use of ceilings and floors we assume for simplicity that both N and K are powers of 2.

4. $f_{N,K}$ *can be computed by a depth-3 formula with* $2^d \cdot (2^m + 1) < 9K^7 \log^2 N$ *multiplication gates, each of degree* $2^m = 4K^5 \log N$.

5. $f_{N,K}$ *can be computed by a monotone depth-3 formula with* $2^d \cdot \binom{2^m}{K} = (K \log N)^{O(K)}$ *multiplication gates.*

6. *For* $K \geqslant \log N$, $f_{N,K}$ *can be computed by a monotone formula of size*

$$K^2 \log N \cdot N \cdot \text{poly}\left(4K^5 \log N, \binom{K + \log(4K^5 \log N)}{\log(4K^5 \log N)}\right) = \begin{cases} N \cdot K^{O(\log K)} & K > \log N \\ N \cdot \text{poly}(\log N) & K \leqslant \log \log N \\ N \cdot \left(1 + \frac{K}{\log \log N}\right)^{O(\log \log N)} & \textit{otherwise} \end{cases}$$

Proof. 1. The first claim follows directly from the fact that for every fixed a, $C(\cdot, a)$ is a map.

2. Since the sets $S_{a,z}$ are disjoint, the support of $\sigma_{2^m,K}(\ell_{a,z}(\mathbf{x}) \mid z \in \{0,1\}^m)$ contains only multilinear monomials of degree K. Therefore, $f_{N,K}$ is multilinear and homogeneous of degree K.

3. Let $I \subset N = \{0,1\}^n$ be of size K. Think of I as a k-source for $k = \log K$. By Claim 3.3, for $1 - \sqrt{\varepsilon}$ of $a \in \{0,1\}^d$, the map $C(\cdot, a)$ is injective on I. For such a, the monomial $\prod_{i \in I} x_i$ appears in $\sigma_{2^m,K}(\ell_{a,z}(\mathbf{x}) \mid z \in \{0,1\}^m)$ with coefficient 1. Thus, the coefficient of each multilinear monomial is at most 1 (since we divide by 2^d) and at least $(1 - \sqrt{\varepsilon}) > 1 - 1/K$.

4. Ben-Or's trick (see e.g. [SW01]) gives a depth-3 multilinear formula with $(2^m + 1)$ multiplication gates computing $\sigma_{2^m,K}(y_1, \ldots, y_{2^m})$. Substituting $\ell_{a,z}$ for y_z, we obtain depth-3 multilinear formula with $(2^m + 1)$ multiplication gates computing $\sigma_{2^m,K}(\ell_{a,z}(\mathbf{x}) \mid z \in \{0,1\}^m)$, proving the claim.

5. Since each $\sigma_{2^m,K}(\ell_{a,z}(\mathbf{x}) \mid z \in \{0,1\}^m)$ has a trivial monotone depth-3 formula with $\binom{2^m}{K}$ multiplication gates, the claim follows.

6. Note that $\sigma_{2^m,K}(\ell_{a,z}(\mathbf{x}) \mid z \in \{0,1\}^m)$ is the degree-K homogeneous component of $\prod_z (1 + \ell_{a,z}(\mathbf{x}))$. By Theorem 1.2, the degree-K homogeneous component of $\prod_z (1 + y_z)$ has a monotone formula of size $\text{poly}\left(2^m, \binom{K+m}{m}\right)$. A simple calculation gives

$$\binom{K+m}{m} = \binom{K + \log(4K^5 \log N)}{\log(4K^5 \log N)} = \begin{cases} K^{O(\log K)} & \log N \leqslant K \\ \left(1 + \frac{K}{\log \log N}\right)^{O(\log \log N)} & \log \log N < K < \log N \\ \text{poly}(\log N) & K \leqslant \log \log N \end{cases}.$$

Substituting $\ell_{a,z}$ for y_z, we obtain a monotone formula for $\sigma_{2^m,K}(\ell_{a,z}(\mathbf{x}) \mid z \in \{0,1\}^m)$ of size

$$N \cdot \text{poly}\left(K, \log N, \binom{K+m}{m}\right) = \begin{cases} N \cdot K^{O(\log K)} & \log N \leqslant K \\ N \cdot \left(1 + \frac{K}{\log \log N}\right)^{O(\log \log N)} & \log \log N < K < \log N \\ N \cdot \text{poly}(\log N) & K \leqslant \log \log N \end{cases}. \quad (1)$$

The result follows as there are $2^d = O(K^2 \log N)$ different seeds a, each contributes according to (1). □

Remark 3.6. A tighter construction could be obtained by recursively approximating each $\sigma_{2^m,K}(\ell_{a,z}(\mathbf{x}) \mid z \in \{0,1\}^m)$, but as this would not alter the main message of the result, have not attempted to get the optimal construction.

Remark 3.7. Instead of using a condenser we could have employed a random map. However, since this would not affect the overall result, we chose to use condensers to emphasize the explicitness of the construction.

The proofs of Theorems 1.9, 1.10 and 1.18 follow directly from Claim 3.5.

Proof of Theorem 1.9. This is exactly Claim 3.5(4). □

Proof of Theorem 1.10. This is a consequence of Claim 3.5(3) and Claim 3.5(6). □

Proof of Corollary 1.11. The proof of Claim 3.5(6) relies on applying Theorem 1.2 to each of the polynomials $\sigma_{2^m,K}(\ell_{a,z}(\mathbf{x}) \mid z \in \{0,1\}^m)$. Each such polynomial is derived through a linear substitution to $\sigma_{2^m,K}(y_1,\ldots,y_{2^m})$, which can be computed by a multilinear depth-3 formula of degree 2^m and top fan-in $2^m+1 = \text{poly}(K, \log N)$. Therefore, if the result in Theorem 1.2 could be improved when applied to the multilinear depth-3 formula computing $\sigma_{2^m,K}(\mathbf{y})$, it would yield a homogeneous multilinear formula for $f_{N,K}$ of size smaller than $N \cdot K^{O(\log K)}$, contradicting the lower bound given in Theorem 1.8.

To demonstrate that the statement of Theorem 1.3 cannot be strengthened if we insist on obtaining a homogeneous and multilinear formula, we again consider $f_{N,K}$. Suppose that Theorem 1.3 also holds in the multilinear case. Applying it to the multilinear depth-3 formula derived in Claim 3.5(4) would produce a homogeneous multilinear formula for $f_{N,K}$ of size $N \cdot K^{O(3+\log K)}$. Up to the constant hidden in the big-O notation, this size is asymptotically optimal, as shown by the lower bound in Theorem 1.8. □

To prove Theorem 1.12, we construct a slightly different polynomial $g_{N,K}$, though the approach follows the same general framework as that for $f_{N,K}$. The construction is probabilistic in nature: we sample $2^{O(K)} \cdot \log N$ random maps forming a family $\mathcal{F} \subset \{f : [N] \to [K]\}$, and show that there exists a choice of such a family $\mathcal{F}$ of size $|\mathcal{F}| = 2^{O(K)} \cdot \log N$ with the property that every K-set is mapped injectively by at least 0.99 of the functions in $\mathcal{F}$.We then define $g_{N,K}$ in a way analogous to the definition of $f_{N,K}$.

Claim 3.8. *Let $I \subset [N]$ be of size $|I| = K$. Let $f : [N] \to [K]$ be chosen uniformly at random among all such maps. Then*

$$\Pr[|f(I)| = K] = \frac{K!}{K^K} = 2^{-O(K)}.$$

Chernoff's inequality (see e.g., [AS16, Appendix A]) implies that if we choose many such maps f independently at random, then for any fixed K-set I, the probability that I is not mapped injectively by $(1 \pm 0.001)\frac{K!}{K^K}$ of them, is exponentially small.

Claim 3.9. *Let $I \subset [N]$ be of size $|I| = K$. Let $f_1,\ldots,f_t : [N] \to [K]$ be chosen uniformly and independently at random among all maps from $[N]$ to $[K]$. Denote the event that f_i mapped I injectively by $E_i(I)$. Then*

$$\Pr\left[\left|\sum_{i=1}^{t} \mathbb{1}_{E_i}(I) - t\frac{K!}{K^K}\right| > 0.001t \cdot \frac{K!}{K^K}\right] < e^{-2t\frac{K!}{K^K}/10^6}.$$

Corollary 3.10. *For $t = O\left(\frac{K^K}{K!} \cdot K \log N\right) = 2^{O(K)} \cdot \log N$ there exist $f_1,\ldots,f_t$ such that every K-set I is mapped injectively by $(1 \pm 0.001)t$ of them.*

Let t be as in Claim 3.9 and $f_1,\ldots,f_t$ be as in Corollary 3.10. For each f_i let $T_i = \prod_{j=1}^{K} \sum_{\ell : f_i(\ell)=j} x_\ell$. Set $g_{N,K} = \frac{K^K}{1.001t \cdot K!} \sum_{i=1}^{t} T_i$. By definition, $g_{N,K}$ has a monotone depth-3 formula with $t = O\left(\frac{K^K}{K!} \cdot K \log N\right) = 2^{O(K)} \cdot \log N$ multiplication gates.

Claim 3.11. *The polynomial $g_{N,K}$ has the same support as $\sigma_{N,K}$ and each of its coefficients lies in the interval $[0.99, 1]$.*

Proof. This follows directly from the selection of $f_1,\ldots,f_t$ as guaranteed by Corollary 3.10, and from the fact that if f_i mapped a K-set I injectively, then the coefficient of the multilinear monomial $\prod_{j \in I} x_j$ in T_i, is 1. □

Proof of Theorem 1.12. The result follows immediately from the construction of of $g_{N,K}$ and Claim 3.11. □

Proof of Theorem 1.18. By Theorem 1.12, the partial derivative space of order K derivatives of $g_{N,2K}$ has dimension $2^{O(K)} \log N \cdot \binom{2K}{K} = 2^{O(K)} \log N$. □

Proof of Theorem 1.21. Theorem 1.20 implies that the partial derivative space of order K derivatives of the polynomial given in the theorem has dimension $\exp(K)$. □

4 Discussion

This paper demonstrates that the analysis of Raz's homogenization result (Theorem 1.1) is tight. It further shows that the the statement of Theorem 1.3 cannot be strengthened for small depth multilinear formulas, assuming we require the resulting formula to be both multilinear and homogeneous. Additionally, we obtain that the proof method of Theorem 1.6 yields asymptotically tight bounds.

A somewhat surprising result is that a slight perturbation of the coefficients of $\sigma_{N,K}$ produces a polynomial, whose partial derivative space has significantly smaller dimension (Theorem 1.18). Moreover, this polynomial has monotone formula complexity roughly equivalent to its homogeneous formula complexity.

In his seminal work [Kal89], Kaltofen showed that the factors of polynomial size algebraic circuit can themselves be computed by polynomial sized circuits. Similarly, Sinhababu and Thierauf [ST21] obtained an analogous result for algebraic branching programs. However, whether an equivalent result holds for algebraic formulas remains an open question (see the survey[FS15] for further discussion on related questions and results in factorization). Another related question that also seems to be open is the following.

Question 4.1. *What is the homogeneous formula complexity of factors of homogeneous formulas?*

We note that if the lower bound in Theorem 1.6(2) could be extended to general homogeneous formulas (without the requirement of multilinearity), it would imply that the polynomial $z^{N-K} \cdot \sigma_{N,K}$ admits a homogeneous depth-3 formula, of size $O(N^2)$, while its factor $\sigma_{N,K}$ requires super-polynomial size homogeneous formulas (for $K = 2^{\omega(\sqrt{\log N})}$). Alternatively, if the answer to Question 4.1 is affirmative, i.e., homogeneous formulas are closed under factorization, then this would immediately yield a polynomial-size homogeneous formula for $\sigma_{N,K}$.

Another interesting question is to establish lower bounds on the Waring rank of an approximation to $\sigma_{N,K}$. Pratt showed that the Waring rank of an ε approximation is at most $4.075^K \varepsilon^{-2} \log n$ (Theorem 1.19), while the trivial lower bound is 2^{K-1}.

Bibliography

[AG10] Noga Alon and Shai Gutner. Balanced families of perfect hash functions and their applications. *ACM Trans. Algorithms*, 6(3):Art. 54, 12, 2010. 5

[AS16] Noga Alon and Joel H. Spencer. *The probabilistic method*. Wiley Series in Discrete Mathematics and Optimization. John Wiley & Sons, Inc., Hoboken, NJ, fourth edition, 2016. 9

[AW17] Josh Alman and Ryan Williams. Probabilistic rank and matrix rigidity. In *STOC'17—Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pages 641–652. ACM, New York, 2017. 4

[AYZ95] Noga Alon, Raphael Yuster, and Uri Zwick. Color-coding. *J. Assoc. Comput. Mach.*, 42(4):844–856, 1995. 5

[Bjö14] Andreas Björklund. Determinant sums for undirected Hamiltonicity. *SIAM J. Comput.*, 43(1):280–299, 2014. 5

[dW06] Ronald de Wolf. Lower bounds on matrix rigidity via a quantum argument. In *Automata, languages and programming. Part I*, volume 4051 of *Lecture Notes in Comput. Sci.*, pages 62–71. Springer, Berlin, 2006. 4

[FLST24] Hervé Fournier, Nutan Limaye, Srikanth Srinivasan, and Sébastien Tavenas. On the power of homogeneous algebraic formulas. In Bojan Mohar, Igor Shinkar, and Ryan O'Donnell, editors, *Proceedings of the 56th Annual ACM Symposium on Theory of Computing, STOC 2024, Vancouver, BC, Canada, June 24-28, 2024*, pages 141–151. ACM, 2024. 2, 5

[FS15] Michael A Forbes and Amir Shpilka. Complexity theory column 88: Challenges in polynomial factorization. *ACM SIGACT News*, 46(4):32–49, 2015. 10

[Got66] Daniel H. Gottlieb. A certain class of incidence matrices. *Proc. Amer. Math. Soc.*, 17:1233–1237, 1966. 5

[HY11] Pavel Hrubes and Amir Yehudayoff. Homogeneous formulas and symmetric polynomials. *Comput. Complex.*, 20(3):559–578, 2011. 2, 3

[Kal89] Erich Kaltofen. Factorization of polynomials given by straight-line programs. *Adv. Comput. Res.*, 5:375–412, 1989. 10

[Kou08] Ioannis Koutis. Faster algebraic algorithms for path and packing problems. In *Automata, languages and programming. Part I*, volume 5125 of *Lecture Notes in Comput. Sci.*, pages 575–586. Springer, Berlin, 2008. 5

[KR98] Boris S Kashin and Alexander A Razborov. Improved lower bounds on the rigidity of hadamard matrices. *Mathematical Notes*, 63(4):471–475, 1998. 4

[Lok01] Satyanarayana V Lokam. Spectral methods for matrix rigidity with applications to size–depth trade-offs and communication complexity. *Journal of Computer and System Sciences*, 63(3):449–473, 2001. 4

[LST24] Nutan Limaye, Srikanth Srinivasan, and Sébastien Tavenas. Superpolynomial lower bounds against low-depth algebraic circuits. *Communications of the ACM*, 67(2):101–108, 2024. 2

[Pra19] Kevin Pratt. Waring rank, parameterized and exact algorithms. In David Zuckerman, editor, *60th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2019, Baltimore, Maryland, USA, November 9-12, 2019*, pages 806–823. IEEE Computer Society, 2019. 5, 6

[Ram20] C Ramya. Recent progress on matrix rigidity–a survey. *arXiv preprint arXiv:2009.09460*, 2020. 5

[Ras16] Cyrus Rashtchian. Bounded matrix rigidity and john's theorem. In *Electronic Colloquium on Computational Complexity (ECCC)*, volume 23, page 93, 2016. 4

[Raz13] Ran Raz. Tensor-rank and lower bounds for arithmetic formulas. *J. ACM*, 60(6):40:1–40:15, 2013. 2, 3, 5

[RS22] Yuval Rabani and Amir Shpilka. Corrigendum: Explicit construction of a small epsilon-net for linear threshold functions. *SIAM Journal on Computing*, 51(5):1692–1702, 2022. 7

[Sri20] Srikanth Srinivasan. Strongly exponential separation between monotone VP and monotone VNP. *ACM Trans. Comput. Theory*, 12(4):Art. 23, 12, 2020. 3

[ST21] Amit Sinhababu and Thomas Thierauf. Factorization of polynomials given by arithmetic branching programs. *Comput. Complex.*, 30(2):15, 2021. 10

[SW01] Amir Shpilka and Avi Wigderson. Depth-3 arithmetic circuits over fields of characteristic zero. *Computational Complexity*, 10:1–27, 2001. 8

[Wil09] Ryan Williams. Finding paths of length k in $O^*(2^k)$ time. *Inform. Process. Lett.*, 109(6):315–318, 2009. 5

[Xie22] Yangxinyu Xie. Matrix rigidity : a survey. Master's thesis, The University of Texas at Austin, 2022. 5

[Yeh19] Amir Yehudayoff. Separating monotone VP and VNP. In Moses Charikar and Edith Cohen, editors, *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing, STOC 2019, Phoenix, AZ, USA, June 23-26, 2019*, pages 425–429. ACM, 2019. 3

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il