Barriers for Rank Methods in Arithmetic Complexity*†

Klim Efremenko[1], Ankit Garg[2], Rafael Oliveira[3], and Avi Wigderson[4]

1 **Ben Gurion University of the Negev, Beer-Sheva, Israel**
klimefrem@gmail.com
2 **Microsoft Research New England, 1 Memorial Dr, Cambridge, MA, USA**
garga@microsoft.com
3 **Department of Computer Science, University of Toronto, 10 King's College Road, Toronto, Canada**
rafael@cs.toronto.edu
4 **Institute for Advanced Study, 1 Einstein Dr, Princeton, USA**
avi@math.ias.edu

Abstract

Arithmetic complexity, the study of the cost of computing polynomials via additions and multiplications, is considered (for many good reasons) simpler to understand than *Boolean complexity*, namely computing Boolean functions via logical gates. And indeed, we seem to have significantly more lower bound techniques and results in arithmetic complexity than in Boolean complexity. Despite many successes and rapid progress, however, foundational challenges, like proving super-polynomial lower bounds on circuit or formula size for explicit polynomials, or super-linear lower bounds on explicit 3-dimensional tensors, remain elusive.

At the same time (and possibly for similar reasons), we have plenty more excuses, in the form of "barrier results" for failing to prove basic lower bounds in Boolean complexity than in arithmetic complexity. Efforts to find barriers to arithmetic lower bound techniques seem harder, and despite some attempts we have no excuses of similar quality for these failures in arithmetic complexity. This paper aims to add to this study.

In this paper we address *rank methods*, which were long recognized as encompassing and abstracting almost all known arithmetic lower bounds to-date, including the most recent impressive successes. Rank methods (under the name of *flattenings*) are also in wide use in algebraic geometry for proving tensor rank and symmetric tensor rank lower bounds. Our main results are barriers to these methods. In particular,

- Rank methods *cannot* prove better than $\Omega_d(n^{\lfloor d/2 \rfloor})$ lower bound on the tensor rank of *any* d-dimensional tensor of side n. (In particular, they cannot prove super-linear, indeed even $> 8n$ tensor rank lower bounds for *any* 3-dimensional tensors.)
- Rank methods *cannot* prove $\Omega_d(n^{\lfloor d/2 \rfloor})$ on the *Waring rank*[1] of any n-variate polynomial of degree d. (In particular, they cannot prove such lower bounds on stronger models, including depth-3 circuits.)

The proofs of these bounds use simple linear-algebraic arguments, leveraging connections between the *symbolic* rank of matrix polynomials and the usual rank of their evaluations. These

* This work was partially supported by NSF grants CCF-1149888, CCF-1523816, CCF-1412958, CAREER award DMS-1451191, European Community's Seventh Framework Programme (FP7/2007- 2013) under grant agreement number 257575, Simons Collaboration on Algorithms and Geometry, Simons Fellowship in Theoretical Computer Science and Siebel Scholarship.

† A full version of the paper is available at [14], https://arxiv.org/abs/1710.09502

[1] A very restricted form of depth-3 circuits

9th Innovations in Theoretical Computer Science Conference (ITCS 2018).
Editor: Anna R. Karlin; Article No. 1; pp. 1:1–1:19

Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

techniques can perhaps be extended to barriers for other arithmetic models on which progress has halted.

To see how these barrier results directly inform the state-of-art in arithmetic complexity we note the following. First, the bounds above nearly match the best explicit bounds we know for these models, hence offer an explanations why the rank methods got stuck there. Second, the bounds above are a far cry (quadratically away) from the true complexity (e.g. of random polynomials) in these models, which *if* achieved (by any methods), are known to imply super-polynomial formula lower bounds.

We also explain the relation of our barrier results to other attempts, and in particular how they significantly differ from the recent attempts to find analogues of "natural proofs" for arithmetic complexity. Finally, we discuss the few arithmetic lower bound approaches which fall outside rank methods, and some natural directions our barriers suggest.

1998 ACM Subject Classification F. Theory of Computation – Algebraic complexity theory

Keywords and phrases Lower Bounds, Barriers, Partial Derivatives and Flattenings

Digital Object Identifier 10.4230/LIPIcs.ITCS.2018.1

1 Introduction

Arithmetic complexity theory (often also called algebraic complexity theory) addresses the computation of algebraic objects (like polynomials, matrices, tensors) using the arithmetic field operations (and sometimes other operations like taking roots). Within computational complexity this field is nearly as old as Boolean complexity theory, which addresses the computation of discrete functions via logical operations, but of course mathematicians were interested in arithmetic computation for centuries before computer science was born. Indeed, Euclid's algorithm for computing GCD, Gauss' discovery of the FFT, and Abel's impossibility result for solving quintic equations by radicals are all precursors of arithmetic complexity theory. Today algebraic algorithms pervade mathematics! Extensive surveys of this field are presented in the books [8, 52], and, more focused on the present material are the recent monographs [48, 12], as well as the book [33] which offers an algebro-geometric perspective.

Structurally, the Boolean and arithmetic theories, and especially the quest for lower bounds which we will focus on, progressed almost hand in hand. Shortly after the important discoveries of reductions and completeness leading to the definitions of P, NP, and complete problems for them, Valiant [51] developed the arithmetic analog notions of VP, VNP and complete problems for them. Separating these pairs of classes stand as the long-term challenges of these fields, and their difficulty has led to the study of a large variety of restricted models in both. Definitions, techniques and results have propagated back and forth and inspired progress, but, all in all, we understand the arithmetic models much better. This of course comes as no surprise. In the arithmetic setting (especially over fields that are large, of characteristic zero, or are algebraically closed) the diverse tools of algebra are available, but have no analogs in the Boolean setting. Moreover, as arithmetic computation is mostly *symbolic* it is (essentially) more stringent than the Boolean computation of functions[2]; indeed, it is known that proving (a non-uniform version of) $\mathsf{P} \neq \mathsf{NP}$ implies $\mathsf{VP} \neq \mathsf{VNP}$ when the underlying field is $\mathbb{C}$ [7]. and thus arithmetic lower bounds are also formally easier to prove!

[2] For example, the *polynomial* $x^p - x$ over $\mathbb{F}_p$ is nontrivial to compute, while the (identically zero) function it represents is trivial.

Despite exciting and impressive progress on arithmetic lower bounds (we will detail many later), some of the most basic questions remain open, and this seeming weakness of current techniques begs explanation, which will hopefully lead to new ones. In Boolean complexity there is a rich interplay between the discovery of the power of new techniques, and then their limitations, in the form of *barrier results*. Such results formally encapsulate a set of lower bound methods, and then prove (unconditional, or sometimes conditionally on natural assumptions) that these cannot solve basic questions. Well known barriers to large classes of techniques include the *relativization* barrier of Baker, Gill and Solovay [6], the *natural proof* barrier of Razborov and Rudich [45] and the *algebrization* barrier of Aaronson and Wigderson [2]. But there are many other important barriers, to more concrete lower bound methods, including [42, 43, 38]. Finding analogous barriers for arithmetic complexity has been much harder; while encapsulation of general lower bound techniques exists, e.g. in [20, 16, 21], there are really no proofs of their limitations (we will discuss these in the related works subsection below).

This paper provides, to the best of our knowledge, the first unconditional barrier results on a very general class of methods, capturing many of the known lower bounds, including the very exciting recent ones. We now begin to describe, through examples, the techniques we encompass under *rank methods* and then explain their limitations.

1.1 Sub-Additive Measures, Rank Bounds and Barriers

Throughout, we will discuss the computation of multivariate polynomials over any field, by arithmetic circuits of various forms, in a way that will not necessitate too many specific details; we will give these as needed, and give formal details in the technical sections. The examples we start with below will demonstrate many "cheap" computations may be encompassed by writing the output polynomial as a "short" sum of *simpler* ones. Thus lower bounds on the number of summands can yield (important) complexity lower bounds. We continue with discussing classes of such lower bound techniques, and then barrier results that put a limit on how large lower bounds such classes of techniques can prove.

Sub-additive measures

Let us start with some examples and then generalize them.

- One of the earliest basic results in arithmetic complexity, due to Hyafil [26] states the following: if a homogeneous circuit of size s computes an n-variate polynomial f of degree d, then

  $$f = g_1 + g_2 + \cdots + g_s$$

  where each g_i is *simple*, which here means *highly reducible*: $g_i = p_i \cdot q_i$, where the degrees of p_i, q_i do not exceed $2d/3$. This result was developed towards parallelizing arithmetic computation, but can also be used for lower bounds: if we could find any sub-additive measure μ on polynomials, which is small on all possible g_i but is large on f, we would have a lower bound on the minimum circuit size s of f! In particular, Hyafil's theorem implies that if the ratio of "large" and "small" values of μ is super-polynomial in n, d, this would imply[3] $\mathsf{VP} \neq \mathsf{VNP}$! We note that Hyafil's theorem is today only one example of numerous other decomposition theorems of similar nature used in lower bounds, e.g. [37, 36, 41, 24] to mention a few.

[3] Since homogenous computation can efficiently simulate non-homogeneous one.

- An even simpler example, where a similar decomposition follows directly from the definition, is tensor rank. Assume that a d-dimensional tensor (with n variables in each dimension) has rank s. This means[4] that

  $$f = g_1 + g_2 + \cdots + g_s$$

  where each g_i is *simple*, which here means *of rank 1*: $g_i = \ell_i^{(1)} \otimes \ell_i^{(2)} \cdots \otimes \ell_i^{(d)}$, where $\ell_i^{(j)}$ is a linear form in the variables of dimension j. Again, any sub-additive measure μ on tensors which is small on all possible rank 1 tensors g_i, but is large on f would yield a lower bound on its tensor rank. This question is no less important than the previous one even though tensor rank seems like a more restricted complexity measure: Raz [40] proved that presenting an explicit tensor f of super-constant dimension $d \leq \log n / \log\log n$, with a nearly-tight tensor rank lower bound of $n^{d(1-o(1))}$ (which holds for most tensors) will imply $\mathsf{VP}_e \neq \mathsf{VNP}$ (namely, explicit super-polynomial lower bounds on formulas)! We note that a similar example as tensor rank, where a decomposition suggests itself by definition, is Waring rank, where each g_i is a d-power of a linear form.
- A third set of examples which directly gives such decompositions of computations is when considering bounded-depth circuits. In almost all computations one can assume without loss of generality that the top (output) gate is a plus gate, and so if a polynomial f is computed by a depth-h circuit of size s, then

  $$f = g_1 + g_2 + \cdots + g_s$$

  where each g_i is *simple* in being of depth $h-1$ (and moreover, with a top product gate). Sub-additive measures small on such simple polynomials and large on f were the key to the many successes on remarkably tight lower bounds for depth-3 and then depth-4 circuits [36, 28, 22, 27, 17, 30, 29]. These include the breakthrough of $(nd)^{\sqrt{d}}$ explicit lower bounds [22] on the size of homogeneous depth-4 circuits, which again seem much more restricted than it is: any super-constant improvement of the exponent will imply $\mathsf{VP} \neq \mathsf{VNP}$!

There are many other examples in which obtaining such decompositions as above uses extra tools like approximations, random restrictions, or iterations. Abstracting all these examples and indeed most known lower bounds in arithmetic complexity[5], can be done in a simple way. Let S be a set of *simple* polynomials, and let $\hat{S}$ be their linear span. The S-complexity $c_S(f)$ of a polynomial $f \in \hat{S}$ is simply the smallest number s such that $f = g_1 + g_2 + \cdots + g_s$ and each $g_i \in S$. A *sub-additive* measure μ is a function $\mu : \hat{S} \to \mathbb{R}^+$ such that

$$\mu(g+h) \leq \mu(g) + \mu(h)$$

for any $g, h \in \hat{S}$. Extending μ to sets, denoting $\mu(T) = \max\{\mu(g) : g \in T\}$, we can immediately derive a lower bound on $c_S(f)$ for any polynomial f by

$$c_S(f) \geq \mu(f)/\mu(S).$$

Let Δ_S denote all possible sub-additive measures on $\hat{S}$. It is a triviality that c_S itself is a sub-additive measure in Δ_S, and hence this method can in principle provide tight lower

[4] Directly generalizing matrix rank, which is the case $d = 2$.
[5] The discussion below is quite general and indeed applies to lower bounds and barriers that use sub-additive measures in practically any computational model.

bound on the complexity $c_S(f)$ for every f. However, the difficulty of proving lower bounds precisely means that c_S is hard to understand, and so we try to "approximate it" with simpler measures $\mu \in \Delta$ for some family $\Delta \subseteq \Delta_S$ of sub-additive measures which are hopefully simpler to understand, compute and reason about.

Barriers for sub-additive measures

This brings us to the topic of this paper: barriers, or limits to the power of such class of lower bound methods. A *barrier* result for any such class of sub-additive measures $\Delta \subseteq \Delta_S$ simply asserts that $\mu(f)$ is *small* for *every* $\mu \in \Delta$ and any $f \in \hat{S}$ (whenever $\mu(S)$ is small). The quantity

$$c(\Delta) = \mu(\hat{S})/\mu(S)$$

upper bounds the best lower bound which can be proven using *any* $\mu \in \Delta$ on *any* polynomial $f \in \hat{S}$, simply as $\mu(f) \leq c(\Delta) \cdot \mu(S)$ for all of them.

Of course, concrete lower bounds are obtained using specific measures μ, and there is always hope that a clever variant of such a choice will give even better bounds; indeed, much of the progress in lower bounds is of this nature. The quality of barrier result is in classifying as large as possible a class of measures Δ, which captures many complexity measures, such that either $c(\Delta)$ is close to the best known lower bounds, or it is well separated with a "desired" lower bound (e.g. one that would approach the complexity of a random polynomial, or that would significantly improve the state of art). In this paper we focus on *rank methods*, which we turn to describe now.

Rank methods

The rank function of matrices, is at once extremely well studied and understood in linear algebra, and is sub-additive. This has made numerous (implicit and explicit) choices of sub-additive measures, for a variety of computational models, to be defined via matrix rank, as follows. Fix a field $\mathbb{F}$, and let $\mathrm{Mat}_m(\mathbb{F})$ denote the set of all $m \times m$ matrices over $\mathbb{F}$. Fix the set of (simple) polynomials S, (and thereby also their span $\hat{S}$) as before. Define the class $\Delta_0^S \subseteq \Delta_S$ to be the set of sub-additive measures μ which arise in the following way. Let $L : S \to \mathrm{Mat}_m(\mathbb{F})$ be any *linear* map for some integer m. Namely, for all $g, h \in S$ (and hence also in $\hat{S}$) we have $L(g+h) = L(g) + L(h)$, and that $L(bg) = L(g)$ for any non-zero constant $b \in \mathbb{F}$. Define

$$\mu_L(f) = rank_{\mathbb{F}}(L(f)).$$

Clearly, all these $\mu_L \in \Delta_0^S$ are sub-additive measures on S. We call the elements of Δ_0^S as *rank methods* for S.

As mentioned, rank methods abound in arithmetic (and other) lower bounds. The possibly familiar names including *partial derivatives, shifted partial derivatives, evaluation dimension, coefficient dimension* which are used e.g. in these lower bounds for monotone, non-commutative, homogeneous, multilinear, bounded-depth and other models [37, 49, 42, 36, 28, 22, 27, 15, 17, 30, 29] are all rank methods, and in many of these papers are explicitly stated as such. Moreover, in algebraic geometry, rank methods (usually called *flattenings*) are responsible for almost all tensor rank and symmetric tensor rank lower bounds (see e.g. [33]).

What should be stressed is that rank methods are extremely general. We do not restrict the size m of matrices used in any way (and indeed in some applications, like shifted partial derivatives [22], m grows super exponentially in the basic size parameters n, d). Moreover,

we demand no explicitness in the specification of the linear map L (and indeed, in some applications, like the multilinear formula lower bounds in [39, 41] the map is chosen at random). The barrier results hold for all.

We prove barrier results for two classes of very weak computational models, *tensor rank* and *Waring rank*, which are very special cases of (respectively) multilinear and homogeneous depth-3 circuits (which themselves are the weakest class of circuits studied[6]. As with all barrier results, the weaker the model for which they are proved, the better, as they scale up for stronger models automatically! As discussed above, we will compare our barriers both to the state-of-art lower bounds in these models, as well to the best one can hope for, namely the complexity of random polynomials.

1.2 Main results

Our results below work for infinite fields $\mathbb{F}$.[7] We start with tensor rank, and proceed with Waring rank, which may be viewed as a symmetric version of tensor rank. In both cases, our barrier results nearly match (up to a function of d, the degree[8]) the best explicit lower bounds (obtained by rank methods), and are roughly quadratically away from the (desired) lower bounds that hold for random polynomials.

Tensor rank

Tensors abound in mathematics and physics, and have been studied for centuries. We refer the reader to the book [32] for one good survey. From a computational perspective tensors have been extremely interesting as well, as many problems naturally present themselves in tensor form. In arithmetic complexity they are often called *set-multilinear* polynomials. While 2-dimensional tensors, namely matrices, are very well understood, d-dimensional tensors possess far less structure, and one way this is manifested is that the problem of computing tensor rank of 3-dimensional tensors is already NP-complete [23]. Many special cases, approximations and related decompositions of tensors were studied, especially recently with machine learning applications [11, 35, 5, 25, 18]. Let us define the model and problem formally.

Fix n, d. The family of polynomials of interest here is $\hat{S} = \mathrm{Ten}_{n,d}(\mathbb{F})$, namely degree d polynomials in d sets of n variables (so, total of nd variables), in which each monomial has precisely one variable from each set. The coefficients of a tensor are naturally described by an $[n]^d$ box with entries from $\mathbb{F}$. The simple polynomials S are *rank-1* tensors, namely those which are products of d linear forms, one in each set of variables (equivalently, the coefficients are described by the tensor product of d vectors). The tensor rank of a tensor f is the smallest number of rank-1 tensors which add up to it.

Most tensors have rank about n^{d-1}/d. Explicit lower bounds are way worse. It is trivial to construct an explicit d-dimensional tensor of rank $n^{\lfloor d/2 \rfloor}$, and the best known lower bound is only a factor of 2 larger. Specifically, [4] give an explicit tensor with 0,1 coefficients of tensor rank at least $2n^{\lfloor d/2 \rfloor} + n - d \log n$. Note in particular that the best lower bound for $d = 3$ is about $3n$. Although the lower bounds of [4] are not attained via a rank method, many other lower bounds for tensor rank are attained via a rank method in Δ_0^T (T for

[6] As depth-2 circuits simply represent polynomials trivially, as sums of monomials.

[7] Our results below hold for all large enough fields $\mathbb{F}$ (polynomial in n, m, d), however, in most cases the dimension m of the matrices is exponentially large in the parameters of interest – that is, n, d.

[8] Which is a constant in the very interesting cases where the degree d is a constant!

Tensor), namely using a sub-additive measure in the class of rank methods [34, 31]. Our barrier result proves that no bound better $2^d \cdot n^{\lfloor d/2 \rfloor}$ can be proven by rank methods, and in particular for $d = 3$, they cannot beat $8n$ (a factor $8/3$ away from the best explicit lower bound!).

▶ **Theorem 1** (Statement of Theorem 3). $c(\Delta_0^T) \leq 2^d \cdot n^{\lfloor d/2 \rfloor}$.

Waring rank

The Waring problem has a long history in mathematics, first in its number theoretic form initiated by Waring [53] in 1770 (writing integers as short sums of d-powers of other integers), and then in its algebraic form we care about, initiated by Sylvester [50] in 1851 (writing polynomials as short sums of powers of linear forms). Some of the basic questions (computing this minimum for monomials and for random polynomials) were only very recently resolved, using algebraic geometric techniques [10, 3]. In arithmetic complexity this model is often referred to as *depth-3 powering circuits*. Let us formalize the problem.

Fix n, d. The family of polynomials of interest here is $\hat{S} = \text{poly}_{n,d}$, all n-variate polynomials of total degree d. The simple generating set S we care about here is the set of all d-powers, namely all polynomials of the form ℓ^d, where ℓ is an affine function in the n given variables. So, $c_S(f)$ is the smallest number s such that f can be written as a sum of such d powers.

For most polynomials, the Waring rank was settled by [3], and is about $(n-1)^d$ for d much smaller than n, and is precisely

$$\left\lceil \frac{1}{n} \cdot \binom{n+d-1}{n-1} \right\rceil.$$

It is trivial to find an explicit $f \in \text{poly}_{n,d}$ whose Waring rank is $\Omega(n^{\lfloor d/2 \rfloor})$, and the best known lower bound, due to [19] (again via rank method in Δ_0^W), is only a little better,

$$\binom{n+\lfloor d/2 \rfloor - 1}{\lfloor d/2 \rfloor} + \lfloor n/2 \rfloor - 1.$$

Our barrier result proves that rank methods cannot improve this lower bound even by a factor of roughly d.

▶ **Theorem 2** (Barriers for Waring Rank[9]). $c(\Delta_0^W) \leq (d+1) \cdot \binom{n+\lfloor d/2 \rfloor}{n}$.

1.3 High-level ideas of the proof

As mentioned, the proofs of our barrier results use only simple tools of linear algebra (although their use and combination is a bit subtle). Here are the key ideas of the proof, written abstractly in the general notation established above (again, we believe that they can be applied in other settings beyond the two we consider in this paper).

Consider any simple set S of polynomials, and rank methods Δ_0^S for it. Thus, we need to provide an upper bound on the quantity $c(\Delta_0^S)$, namely on the ratio $\mu_L(f)/\mu_L(S)$ for every $f \in \hat{S}$, and every linear map $L : S \to \text{Mat}_m(\mathbb{F})$. Set $r = \mu_L(S)$.

[9] A proof of this theorem appears in the full version of the paper [14]

- We view linear map L, which gives rise to a sub-additive measure in Δ_0^S, as a matrix polynomial, namely as a polynomial with matrix coefficients, or equivalently as a symbolic matrix whose entries are polynomials. The variables of these polynomials will be the *parameters* of the family of *simple* polynomials S (these parameters are the coefficients of the linear forms appearing in the decompositions in both the tensor rank and Waring rank settings). Call this symbolic matrix $L(S)$.
- Next, the *symbolic* rank of $L(S)$ (over the field of rational functions in these variables) is bounded by the maximum rank of any *evaluation* of this matrix polynomial (this is the only place we use the fact that the field is large enough). By assumption, as these evaluations are all in the image of L on the simple polynomials S, this maximum rank is at most r, and so is the symbolic rank.
- The symbolic rank gives rise to a decomposition $L(S) = KM$ with M, K having dimensions $m \times r$ and $r \times m$ respectively, and their entries are *rational functions* in the variables appearing in $L(S)$. We show that with a small loss in the dimension r, this affords a much nicer decomposition $L(S) = K'M'$, with dimensions $m \times r'$ and $r' \times m$ respectively, but now the entries of K', M' are *polynomial* functions of the variables. Moreover, the polynomials in every column of K' and every row of M' are homogeneous of the same degree. For tensor rank we obtain $r' = r2^d$, and for Waring rank we have $r' = r(d+1)$.
- As all entries in matrix $L(S)$ are polynomials of degree d, we must have for every $i \in [r']$, that either the i'th column of K' or the i'th row of M' have degree at most $\lfloor d/2 \rfloor$. The dimension of the space of (vector) coefficients of these vectors of polynomials is an appropriate function D of n, d (which in both cases we care about is about $n^{\lfloor d/2 \rfloor}$). Each such vector of polynomials generates at most D *constant* vectors of their coefficients.
- Combining what we have, we see that for every $g \in S$, we have a decomposition $L(g) = C(g) + R(g)$, where the columns of $C(g)$ are spanned by at most $r'D$ vectors *independent of* g, and the rows of R are spanned by at most $r'D$ vectors *independent of* g (indeed the total number of these vectors is $r'D$). This gives an upper bound of $r'D$ on the rank of each $L(g)$, which of course is not interesting as we already have an upper bound of r on each.
- The punchline is obtained by using the linearity of L, and the fact that $\hat{S}$ is the linear span of S. Together, these imply that *every* matrix $L(f)$ with $f \in \hat{S}$ is also in the linear span of the matrices $\{L(g) : g \in S\}$, and so the same decomposition holds for them. Thus, the rank of each $L(f)$ is at most $r'D$, which is a bound on $\mu_L(\hat{S})$. Thus, $c(\Delta_0^S) \leq r'D/r$. In the two settings we consider, D is roughly the best known explicit lower bound, and r'/r is a function of d (namely, $d+1$ for Waring rank, and 2^d for tensor rank).

1.4 Related Work

We now mention other attempts to provide barriers to arithmetic circuit lower bounds. We also mention rank lower bounds in Boolean complexity, and barriers for them. As will be evident, our work is very different than both sets.

All barrier results we are aware of in arithmetic complexity theory attempt to find analogs of the *natural proof* barrier in Boolean circuit complexity of Razborov and Rudich [45]. Roughly, a lower bound technique is *natural* if it satisfies three properties: usefulness, constructively, largeness which we will not need to define. They show how many Boolean circuit lower bound techniques satisfy these properties. Now crucially, the barrier results for natural proofs in the Boolean setting are *conditional*: they hold under a computational assumption on the existence of efficient pseudorandom generators. In this setting, this assumption is widely believed, and is known to follow from e.g. the existence of exponentially hard one-way functions (one which the world relies for cryptography and e-commerce).

In several works, starting with [1, 20], and following with the recent [16, 21], it was understood that an analogous framework with the same three properties is simple to describe (replacing the representation of Boolean functions by their truth tables by the representation of low-degree multivariate polynomials by their list of coefficients). And indeed, it captures essentially all arithmetic lower bounds known. Unfortunately, the main difference from the Boolean setting is the non-existence of an analogous pseudo-randomness theory, and a believable complexity assumption. Several suggestions for such an assumption were made in the works above, and as articulated in [16, 21], they all take the form of the existence of *succinct* hitting sets for small arithmetic circuits (indeed, such existence is *equivalent* to a barrier result). This assumption is related to PIT (polynomial identity testing) and GCT (geometric complexity theory), but the confidence in it is still shaky (initial work in [16] shows succinct hitting sets against extremely weak models of arithmetic circuits). But regardless how believable this assumption is, note that this barrier is again, conditional!

As mentioned earlier, our barrier results are completely unconditional, and moreover require no constructivity from the lower bound proof (thus capturing methods which are not strictly natural in the sense above). On the other hand, our framework of rank methods capture only a large subset, but certainly not all of the known lower bound techniques.

It is interesting that rank methods were used not only in arithmetic complexity, but also in Boolean complexity. While not directly related to our arithmetic setting, we mention where it was used, and which barriers were studied. First, Razborov has used the rank of matrices in an essential way for his lower bound on $AC^0[2]$ (although an elegant route around it was soon after devised by Smolensky [49]). In another work, Razborov [44] has shown how rank methods can be used to prove superpolynomial lower bounds on *monotone* Boolean formulas. His methods were recently beautifully extended to other monotone variants of other models including span programs and comparator circuits in [46]. The potential of such methods to proving *non-monotone* lower bounds for Boolean formulas was considered by Razborov [42], where he proves a strong barrier result in this Boolean setting. Observing that rank is a *submodular* function, he presents a barrier for any submodular progress measure on Boolean formulae: *no such method can prove a super-linear lower bound!*. His barrier was recently made more explicit in [38].

1.5 Organization

In Section A we establish the notation that will be used throughout the paper and provide some lemmas which we will need in the later sections. In Section B, we establish the main technical content of our paper: we define three notions of matrix decomposition and relate these new definitions to commutative rank. In Section 2, we apply the new decompostions from Section B to obtain the main results of the paper, which are the limitations of the rank techniques. Finally, in Section 3 we conclude the paper and present some open questions and future directions of this work.

2 Rank Bounds

In this section, we show how the matrix decomposition techniques developed in Section B can be used to establish barriers to rank-based methods used to prove lower bounds for tensor rank.[10]

[10] For our results on barriers for Waring rank and constant depth circuits, please see the full version of the paper [14].

We show that any linear map, denoted here by $L : \mathrm{Ten}_{n,d}(\mathbb{F}) \to \mathrm{Mat}_m(\mathbb{F})$, for which $\mathrm{rank}(L(\mathbf{u}_1 \otimes \cdots \otimes \mathbf{u}_d)) \leq r$ for all rank one tensors has the property that $\mathrm{rank}(L(T)) \leq r \cdot 2^d \cdot n^{\lfloor d/2 \rfloor}$ for any tensor $T \in \mathrm{Ten}_{n,d}(\mathbb{F})$. This in turn, implies that such a technique cannot yield better lower bounds than

$$\mathrm{rank}(T) > 2^d \cdot n^{\lfloor d/2 \rfloor}$$

for any explicit tensor $T \in \mathrm{Ten}_{n,d}(\mathbb{F})$.

To put this matter into perspective, it is very easy to obtain explicit tensors $T \in \mathrm{Ten}_{n,d}(\mathbb{F})$ whose tensor rank is lower bounded by $\mathrm{rank}(T) \geq n^{\lfloor d/2 \rfloor}$. For instance, one can just take a full-rank matrix in $\mathrm{Mat}_{n^{\lfloor d/2 \rfloor}}(\mathbb{F})$. Nevertheless, despite much work on tensor rank lower bounds, the best lower bounds for the rank of explicit tensors are still of the form $\Omega(n^{\lfloor d/2 \rfloor})$, as seen in the works [9, 4, 32].

On the other hand, it is well-known, see for instance [33], that a random tensor has rank on the order of $\frac{n^{d-1}}{d}$. Thus, our paper shows that rank-based methods for proving tensor rank lower bounds will not suffice to prove strong tensor lower bounds. We now state the main theorem of this section.

▶ **Theorem 3** (Tensor Rank Upper Bounds). *Let $m, n \in \mathbb{N}$ be positive integers and $L : \mathrm{Ten}_{n,d}(\mathbb{F}) \to \mathrm{Mat}_m(\mathbb{F})$ be a linear map such that each rank one tensor $\mathbf{u}_1 \otimes \cdots \otimes \mathbf{u}_d$ is mapped into a matrix $L(\mathbf{u}_1 \otimes \cdots \otimes \mathbf{u}_d)$ such that*

$$\mathrm{rank}(L(\mathbf{u}_1 \otimes \cdots \otimes \mathbf{u}_d)) \leq r.$$

Then it holds that

$$\mathrm{rank}(L(f)) \leq r \cdot 2^d \cdot n^{\lfloor d/2 \rfloor}$$

for any tensor $f \in \mathrm{Ten}_{n,d}(\mathbb{F})$.

Proof. Let $\mathbf{x}_1 \otimes \cdots \otimes \mathbf{x}_d$ be a generic rank one tensor, where $\mathbf{x}_i = (x_{i1}, \ldots, x_{in})$, with x_{ij} being variables which take values from $\mathbb{F}$, for all $i \in [d]$. Additionally, let $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_d)$, that is, $\mathbf{x}$ is the set of all variables involved, taking into account the partitions of the variables. As the map $L : \mathrm{Ten}_{n,d}(\mathbb{F}) \to \mathrm{Mat}_m(\mathbb{F})$ is a linear map, we must have that

$$L(\mathbf{x}_1 \otimes \cdots \otimes \mathbf{x}_d) = \sum_{i_1, i_2, \ldots, i_d = 1}^{n} A_{i_1, i_2, \ldots, i_d} \prod_{j=1}^{d} x_{j i_j}$$

where each $A_{i_1, i_2, \ldots, i_d} \in \mathrm{Mat}_m(\mathbb{F})$ is a complex $m \times m$ matrix.[11] Hence, $M(\mathbf{x}) = L(\mathbf{x}_1 \otimes \cdots \otimes \mathbf{x}_d)$ is a matrix with set-multilinear polynomial entries, where each polynomial is set-multilinear over the sets of variables $\mathbf{x}_1, \ldots, \mathbf{x}_d$.

By Lemma 10 and the assumption that $\mathrm{rank}(L(\mathbf{u}_1 \otimes \cdots \otimes \mathbf{u}_d)) \leq r$ for any multiset of vectors $\mathbf{u}_i \in \mathbb{F}^n$, we have that

$$\mathrm{rank}_{\mathbb{F}(\mathbf{x})}(L(\mathbf{x}_1 \otimes \cdots \otimes \mathbf{x}_d)) \leq r.$$

In this case, the conditions of Lemma 17 apply and therefore there exist $R \leq r \cdot 2^d$ vectors of homogeneous set-multilinear polynomials $\mathbf{f}_i(\mathbf{x}), \mathbf{g}_i(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ for which

$$M(\mathbf{x}) = \sum_{i=1}^{R} \mathbf{f}_i(\mathbf{x}) \otimes \mathbf{g}_i(\mathbf{x}).$$

[11] One can see this by looking at the standard basis of the space $\mathrm{Ten}_{n,d}(\mathbb{F})$ given by tensoring the standard basis vectors $\mathbf{e}_{i_1} \otimes \cdots \otimes \mathbf{e}_{i_d}$.

Moreover, for all $i \in [R]$, there exists a set S_i such that $\mathbf{f}_i(\mathbf{x})$ is set-multilinear with respect to the partition $(\mathbf{x}_j)_{j \in S_i}$ and $\mathbf{g}_i(\mathbf{x})$ is set-multilinear with respect to the partition $(\mathbf{x}_j)_{j \in [d] \setminus S_i}$. Thus, $\deg(\mathbf{f}_i) + \deg(\mathbf{g}_i) \leq d$, which implies that $\min(\deg(\mathbf{f}_i), \deg(\mathbf{g}_i)) \leq \lfloor d/2 \rfloor$, for each $i \in [R]$. This bound on the minimum degree, combined with Corollary 14 and the fact that $\mathbf{f}_i(\mathbf{x})$ and $\mathbf{g}_i(\mathbf{x})$ are set-multilinear, yield

$$\operatorname{rank}(\mathcal{C}(\mathbf{f}_i(\mathbf{x}) \otimes \mathbf{g}_i(\mathbf{x}))) \leq n^{\lfloor d/2 \rfloor}.$$

As $\operatorname{rank}(\mathcal{C}(M(\mathbf{x}))) \leq \sum_{i=1}^{R} \operatorname{rank}(\mathcal{C}(\mathbf{f}_i(\mathbf{x}) \otimes \mathbf{g}_i(\mathbf{x})))$, we have that

$$\operatorname{rank}(\mathcal{C}(M(\mathbf{x}))) \leq R \cdot n^{\lfloor d/2 \rfloor}.$$

To finish the proof, it is enough to show that $L(f) \in \mathcal{C}(M(\mathbf{x}))$, for any $f \in \operatorname{Ten}_{n,d}(\mathbb{F})$.

For any rank one tensor $\mathbf{u}_1 \otimes \cdots \otimes \mathbf{u}_d$, we have that $L(\mathbf{u}_1 \otimes \cdots \otimes \mathbf{u}_d) \in \mathcal{C}(M(\mathbf{x}))$, as $L(\mathbf{u}_1 \otimes \cdots \otimes \mathbf{u}_d) = M(\mathbf{u})$. As any element $f \in \operatorname{Ten}_{n,d}(\mathbb{F})$ can be written as a linear combination of rank one tensors and by linearity of L, we have that

$$L(f) \in \operatorname{span}\{L(\mathbf{u}_1 \otimes \cdots \mathbf{u}_d) \mid \mathbf{u}_1, \ldots, \mathbf{u}_d \in \mathbb{F}^n\} \subseteq \mathcal{C}(M(\mathbf{x})).$$

Thus, $L(f) \in \mathcal{C}(M(\mathbf{x}))$ and we have that

$$\operatorname{rank}(L(f)) \leq \operatorname{rank}(\mathcal{C}(M(\mathbf{x}))) \leq R \cdot n^{\lfloor d/2 \rfloor},$$

as we wanted. ◀

The theorem above implies the following barrier on rank-based techniques.

▶ **Corollary 4.** *Let $m, n \in \mathbb{N}$ be positive integers and $L : \operatorname{Ten}_{n,d}(\mathbb{F}) \to \operatorname{Mat}_m(\mathbb{F})$ be a linear map (i.e., a flattening). Then, any rank methods which use this linear map cannot prove lower bounds better than*

$$\operatorname{rank}(f) > 2^d \cdot n^{\lfloor d/2 \rfloor}$$

for any tensor $f \in \operatorname{Ten}_{n,d}(\mathbb{F})$.

3 Conclusion and Open Problems

In this paper, we prove the first unconditional barrier for a wide class of lower bound techniques for tensor rank as well as the Waring rank of a polynomial. In particular, for 3-dimensional tensor rank, we show for the first time that a wide class of techniques cannot improve a known linear lower bound (of $2n$) even beyond $8n$. Additionally, we provide an explicit instantiation of the rank method for depth-3 circuits, suggesting it will either help prove better lower bounds, or help develop a barrier for this model that explains the difficulty of proving better lower bounds.

We now provide a list of interesting directions for further research, both on the computational side as well as on the mathematical side.

1. Expand the set of methods for which *unconditional* barrier results be proven in arithmetic complexity theory, beyond the rank methods we study in this paper. In particular, can they be expanded to the use of *non-linear* mappings L, possibly of low degree?
2. Expand the set of arithmetic models for which barriers can be established for rank methods, beyond the two models studied here.
3. In some sense, rank methods "flatten" polynomials of degree $d > 2$ into matrices (in 2 dimensions), in a similar fashion flattening methods in algebraic geometry are used (for very similar purposes). Can this connection be further formalized and used?

References

1 Scott Aaronson and Andrew Drucker. Arithmetic natural proofs theory is sought. *Blog post*, 2008.

2 Scott Aaronson and Avi Wigderson. Algebrization: A new barrier in complexity theory. *ACM Transactions on Computation Theory (TOCT)*, 1(1):2, 2009.

3 James Alexander and André Hirschowitz. Polynomial interpolation in several variables. *Journal of Algebraic Geometry*, 4(2):201–222, 1995.

4 Boris Alexeev, Michael A Forbes, and Jacob Tsimerman. Tensor rank: Some lower and upper bounds. In *Computational Complexity (CCC), 2011 IEEE 26th Annual Conference on*, pages 283–291. IEEE, 2011.

5 Anima Anandkumar, Dean P Foster, Daniel J Hsu, Sham M Kakade, and Yi-Kai Liu. A spectral algorithm for latent dirichlet allocation. In *Advances in Neural Information Processing Systems*, pages 917–925, 2012.

6 Theodore Baker, John Gill, and Robert Solovay. Relativizations of the p=?np question. *SIAM Journal on computing*, 4(4):431–442, 1975.

7 Peter Bürgisser. *Completeness and reduction in algebraic complexity theory*, volume 7. Springer Science & Business Media, 2013.

8 Peter Bürgisser, Michael Clausen, and Amin Shokrollahi. *Algebraic complexity theory*, volume 315. Springer Science & Business Media, 2013.

9 Peter Bürgisser and Christian Ikenmeyer. Geometric complexity theory and tensor rank. In *Proceedings of the forty-third annual ACM symposium on Theory of computing*, pages 509–518. ACM, 2011.

10 Enrico Carlini, Maria Virginia Catalisano, and Anthony V Geramita. The solution to the waring problem for monomials and the sum of coprime monomials. *Journal of algebra*, 370:5–14, 2012.

11 Joseph T Chang. Full reconstruction of markov models on evolutionary trees: identifiability and consistency. *Mathematical biosciences*, 137(1):51–73, 1996.

12 Xi Chen, Neeraj Kayal, and Avi Wigderson. *Partial derivatives in arithmetic complexity and beyond*. Now Publishers Inc, 2011.

13 Richard DeMillo and Richard Lipton. A probabilistic remark on algebraic program testing. *Information Processing Letters*, 7(4):193–195, 1978.

14 Klim Efremenko, Ankit Garg, Rafael Oliveira, and Avi Wigderson. Barriers for rank methods in arithmetic complexity. *arXiv preprint arXiv:1710.09502*, 2017.

15 Michael A Forbes, Ramprasad Saptharishi, and Amir Shpilka. Hitting sets for multilinear read-once algebraic branching programs, in any order. In *Proceedings of the forty-sixth annual ACM symposium on Theory of computing*, pages 867–875. ACM, 2014.

16 Michael A Forbes, Amir Shpilka, and Ben Lee Volk. Succinct hitting sets and barriers to proving algebraic circuits lower bounds. *arXiv preprint arXiv:1701.05328*, 2017.

17 Hervé Fournier, Nutan Limaye, Guillaume Malod, and Srikanth Srinivasan. Lower bounds for depth-4 formulas computing iterated matrix multiplication. *SIAM Journal on Computing*, 44(5):1173–1201, 2015.

18 Rong Ge and Tengyu Ma. On the optimization landscape of tensor decompositions. *arXiv preprint arXiv:1706.05598*, 2017.

19 Fulvio Gesmundo and JM Landsberg. Explicit polynomial sequences with maximal spaces of partial derivatives and a question of k. mulmuley. *arXiv preprint arXiv:1705.03866*, 2017.

20 Joshua A Grochow. Unifying known lower bounds via geometric complexity theory. *computational complexity*, 24(2):393–475, 2015.

21 Joshua A Grochow, Mrinal Kumar, Michael Saks, and Shubhangi Saraf. Towards an algebraic natural proofs barrier via polynomial identity testing. *arXiv preprint arXiv:1701.01717*, 2017.

22 Ankit Gupta, Pritish Kamath, Neeraj Kayal, and Ramprasad Saptharishi. Approaching the chasm at depth four. *J. ACM*, 61(6):33:1–33:16, 2014. doi:10.1145/2629541.

23 Johan Håstad. Tensor rank is np-complete. *Journal of Algorithms*, 11(4):644–654, 1990.

24 Pavel Hrubes, Avi Wigderson, and Amir Yehudayoff. Non-commutative circuits and the sum-of-squares problem. *Journal of the American Mathematical Society*, 24(3):871–898, 2011.

25 Daniel Hsu and Sham M Kakade. Learning mixtures of spherical gaussians: moment methods and spectral decompositions. In *Proceedings of the 4th conference on Innovations in Theoretical Computer Science*, pages 11–20. ACM, 2013.

26 Laurent Hyafil. On the parallel evaluation of multivariate polynomials. *SIAM Journal on Computing*, 8(2):120–123, 1979.

27 N. Kayal, N. Limaye, C. Saha, and S. Srinivasan. Super-polynomial lower bounds for depth-4 homogeneous arithmetic formulas. In *Symposium on Theory of Computing, STOC 2014*, pages 119–127, 2014.

28 Neeraj Kayal. An exponential lower bound for the sum of powers of bounded degree polynomials. *Electronic Colloquium on Computational Complexity (ECCC)*, 19:81, 2012. URL: http://eccc.hpi-web.de/report/2012/081.

29 Mrinal Kumar and Ramprasad Saptharishi. An exponential lower bound for homogeneous depth-5 circuits over finite fields. *arXiv preprint arXiv:1507.00177*, 2015.

30 Mrinal Kumar and Shubhangi Saraf. Superpolynomial lower bounds for general homogeneous depth 4 arithmetic circuits. In *International Colloquium on Automata, Languages, and Programming*, pages 751–762. Springer, 2014.

31 JM Landsberg. Nontriviality of equations and explicit tensors in cm cm cm of border rank at least 2m- 2. *Journal of Pure and Applied Algebra*, 219(8):3677–3684, 2015.

32 Joseph M Landsberg. *Tensors: geometry and applications*, volume 128. American Mathematical Society Providence, RI, 2012.

33 Joseph M Landsberg. *Geometry and Complexity Theory*. Cambridge, 2017.

34 Joseph M Landsberg and Giorgio Ottaviani. New lower bounds for the border rank of matrix multiplication. *Theory of Computing*, 11(11):285–298, 2015.

35 Elchanan Mossel and Sébastien Roch. Learning nonsingular phylogenies and hidden markov models. In *Proceedings of the thirty-seventh annual ACM symposium on Theory of computing*, pages 366–375. ACM, 2005.

36 N. Nisan and A. Wigderson. Lower bound on arithmetic circuits via partial derivatives. *Computational Complexity*, 6:217–234, 1996.

37 Noam Nisan. Lower bounds for non-commutative computation. *STOC*, pages 410–418, 1991.

38 Aaron Potechin. A note on amortized space complexity. *arXiv preprint arXiv:1611.06632*, 2016.

39 Ran Raz. Multi-linear formulas for permanent and determinant are of super-polynomial size. *Journal of the ACM (JACM)*, 56(2):8, 2009.

40 Ran Raz. Tensor-rank and lower bounds for arithmetic formulas. In Leonard J. Schulman, editor, *Proceedings of the 42nd ACM Symposium on Theory of Computing, STOC 2010, Cambridge, Massachusetts, USA, 5-8 June 2010*, pages 659–666. ACM, 2010. doi:10.1145/1806689.1806780.

41 Ran Raz and Amir Yehudayoff. Lower bounds and separations for constant depth multilinear circuits. *Computational Complexity*, 18(2):171–207, 2009. doi:10.1007/s00037-009-0270-8.

42 Alexander Razborov. On submodular complexity measures. URL: http://people.cs.uchicago.edu/~razborov/files/sub.pdf.

43 Alexander A Razborov. On the method of approximations. In *Proceedings of the twenty-first annual ACM symposium on Theory of computing*, pages 167–176. ACM, 1989.

44 Alexander A Razborov. Applications of matrix methods to the theory of lower bounds in computational complexity. *Combinatorica*, 10(1):81–93, 1990.

45 Alexander A Razborov and Steven Rudich. Natural proofs. In *Proceedings of the twenty-sixth annual ACM symposium on Theory of computing*, pages 204–213. ACM, 1994.

46 Robert Robere, Toniann Pitassi, Benjamin Rossman, and Stephen A Cook. Exponential lower bounds for monotone span programs. In *Foundations of Computer Science (FOCS), 2016 IEEE 57th Annual Symposium on*, pages 406–415. IEEE, 2016.

47 Jack Schwartz. Fast probabilistic algorithms for verification of polynomial identities. *Journal of the ACM*, 27:701–717, 1980.

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

49 Roman Smolensky. On representations by low-degree polynomials. In *Foundations of Computer Science, 1993. Proceedings., 34th Annual Symposium on*, pages 130–138. IEEE, 1993.

50 James Joseph Sylvester. Lx. on a remarkable discovery in the theory of canonical forms and of hyperdeterminants. *Philosophical Magazine Series 4*, 2(12):391–410, 1851.

51 Leslie G Valiant. Completeness classes in algebra. In *Proceedings of the eleventh annual ACM symposium on Theory of computing*, pages 249–261. ACM, 1979.

52 Joachim Von Zur Gathen and Jürgen Gerhard. *Modern computer algebra*. Cambridge university press, 2013.

53 E. Waring. Meditationes algebraicæ. In *Archdeacon*, Cambridge, 1770.

54 Richard Zippel. Probabilistic algorithms for sparse polynomials. *EUROSAM*, pages 216–226, 1979.

A Preliminaries

In this section, we establish the notation which will be used throughout the paper and some important background which we shall need to prove our claims in the next sections.

A.1 General Facts and Notations

For simplicity of exposition, we will work over a field $\mathbb{F}$ which is algebraically closed and of characteristic zero, even though our results also hold over infinite fields which need not be algebraically closed.[12] From now on we will use boldface to denote a vector of variables or of field elements. For instance, $\mathbf{x} = (x_1, \ldots, x_n)$ is the vector of variables $x_1, \ldots, x_n$ and $\mathbf{a} = (a_1, \ldots, a_n) \in \mathbb{F}^n$ is a vector of elements $a_1, \ldots, a_n$ from the field $\mathbb{F}$.

For any vector of non-negative integers $\mathbf{a} \in \mathbb{N}^n$ and a vector of n variables $\mathbf{x}$, we define $\mathbf{a}! = \prod_{i=1}^{n} a_i!$ and $\mathbf{x}^{\mathbf{a}} = \frac{1}{\mathbf{a}!} \cdot \prod_{i=1}^{n} x_i^{a_i}$. Since the monomials $\mathbf{x}^{\mathbf{a}}$, $\mathbf{a} \in \mathbb{N}^n$, form a linear basis for the ring of polynomials $\mathbb{F}[\mathbf{x}]$, we can write any polynomial $f(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ as

$$f(\mathbf{x}) = \sum_{\mathbf{a} \in \mathbb{N}^n} \alpha_{\mathbf{a}} \mathbf{x}^{\mathbf{a}}.$$

[12] In general, we only need a field with characteristic polynomial in the number of variables, the degree of the polynomials and the dimension of matrices being studied. We cannot work over field extensions, as we need to use Lemma 10 over the base field.

We will denote the coefficients of the polynomial $f(\mathbf{x})$ by $\text{coeff}_{\mathbf{a}}(f(\mathbf{x})) = \alpha_{\mathbf{a}}$.

The degree of a polynomial $f(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ with respect to a variable x_i, denoted by $\deg_i(f(\mathbf{x}))$ is the maximum degree of x_i in a nonzero monomial of $f(\mathbf{x})$. If $\deg_i(f(\mathbf{x})) \leq 1$ for every variable x_i, we say that the polynomial $f(\mathbf{x})$ is a *multilinear* polynomial. Moreover, if $f(\mathbf{x})$ is multilinear and the variables in $\mathbf{x}$ can be partitioned into sets $\mathbf{x}_1, \ldots, \mathbf{x}_d$ such that each monomial from $f(\mathbf{x})$ has at most one variable from each of the sets $\mathbf{x}_i$, we say that $f(\mathbf{x})$ is a *set-multilinear* polynomial.

▶ **Definition 5** (Homogeneous Components)**.** For a polynomial $f(\mathbf{x})$, denote its homogeneous part of degree t by $H_t[f(\mathbf{x})]$. Additionally, define

$$H_{\leq t}[f(\mathbf{x})] = \sum_{i=0}^{t} H_i[f(\mathbf{x})],$$

that is, $H_{\leq t}[f]$ is the sum of the homogeneous components of $f(\mathbf{x})$ up to degree t. We can extend this definition to matrices of polynomials in the natural way. Namely, if $\mathbf{f}(\mathbf{x})$ is a matrix of polynomials of the form $(f_{ij}(\mathbf{x}))_{i,j}$, we define $H_t[\mathbf{f}(\mathbf{x})] = (H_t[f_{ij}(\mathbf{x})])_{i,j}$, that is, $H_t[\mathbf{f}(\mathbf{x})]$ is the matrix given by the homogeneous components of degree t of each entry of $\mathbf{f}(\mathbf{x})$.

▶ **Definition 6** (Homogeneous Set Multilinear Components)**.** Let $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_d)$ be a set of variables, partitioned into d sets of variables $\mathbf{x}_1, \ldots, \mathbf{x}_d$. For a polynomial $f(\mathbf{x})$ of degree d, let $H_S^{SM}[f(\mathbf{x})]$ denote its homogeneous set-multilinear part corresponding to subpartition $S \subseteq [d]$. That is, $H_S^{SM}[f(\mathbf{x})]$ consists of the sum of all monomials (with the appropriate coefficients) of $f(\mathbf{x})$ of degree exactly $|S|$ which are set-multilinear with respect to the partition $(\mathbf{x}_i)_{i \in S}$.

The following lemma tells us that any nonzero polynomial cannot vanish on a large portion of any sufficiently large grid.

▶ **Lemma 7** (Schwartz-Zippel-DeMillo-Lipton [47, 54, 13])**.** *Let $\mathbb{F}$ be any field such that $|\mathbb{F}| > d$ and let $S \subseteq \mathbb{F}$ be such that $|S| > d$. If $p(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ is a nonzero polynomial of degree d, then*

$$\Pr_{\mathbf{a} \in S^n}[p(\mathbf{a}) = 0] \leq \frac{d}{|S|}.$$

A.2 Matrix Spaces

In this section, we introduce the concept of matrix spaces and establish some of their important properties which we will use in the next sections. We begin by establishing some notations for matrices and tensors.

If V is a vector space of dimension n over a field $\mathbb{F}$, we can identify $V = \mathbb{F}^n$. In this case, we denote the d^{th} tensor power of V by $\text{Ten}_{n,d}(\mathbb{F}) = V^{\otimes d}$. We denote the space of $n \times n$ matrices $V^{\otimes 2}$ by $\text{Mat}_n(\mathbb{F}) = \text{Ten}_{n,2}(\mathbb{F})$. Sometimes we will abuse notation and write $\text{Mat}_n(R)$ for the ring of matrices whose entries take value over a ring R.

A tensor $T \in \text{Ten}_{n,d}(\mathbb{F})$ is a rank-1 tensor if it can be written in the form $T = \mathbf{v}_1 \otimes \cdots \otimes \mathbf{v}_d$, where each $\mathbf{v}_i \in \mathbb{F}^n$. Given any tensor $T \in \text{Ten}_{n,d}(\mathbb{F})$, its rank over $\mathbb{F}$ (denoted by $\text{rank}_{\mathbb{F}}(T)$) is the minimum number r of rank-1 tensors $T_1, \ldots, T_r$ such that $T = T_1 + \cdots + T_r$. Whenever the base field is clear from context, we will denote $\text{rank}_{\mathbb{F}}(T)$ simply by $\text{rank}(T)$.

If $M_1, \ldots, M_k$ are matrices in $\text{Mat}_m(\mathbb{F})$ and $x_1, \ldots, x_k$ are commuting variables, we denote $\text{rank}_{\mathbb{F}(x_1,\ldots,x_k)}(\sum_{i=1}^k x_i M_i)$ the *symbolic rank* of the matrix $\sum_{i=1}^k x_i M_i$.

▶ **Definition 8** (Rank of a Set of Matrices)**.** If $\mathcal{M} \subset \mathrm{Mat}_m(\mathbb{F})$ is a set of $m \times m$ matrices over the field $\mathbb{F}$, define

$$\mathrm{rank}(\mathcal{M}) = \max_{M \in \mathcal{M}} \mathrm{rank}(M).$$

That is, the rank of the set $\mathcal{M}$ is given by the maximum rank (over $\mathbb{F}$) among its elements.

The symbolic rank is important as it characterizes the rank of a linear space of matrices, as seen in the following proposition.

▶ **Proposition 9.** *Let $\mathcal{M} \subseteq \mathrm{Mat}_m(\mathbb{F})$ be a space of matrices. If $M_1, \ldots M_m$ is a basis for $\mathcal{M}$ and $x_1, x_2 \ldots x_m$ are variables then*

$$\mathrm{rank}(\mathcal{M}) = \mathrm{rank}_{\mathbb{F}(x_1,\ldots x_m)} \left(\sum_{i=1}^{m} x_i M_i \right).$$

The propostion above, together with Lemma 7, imply the following lemma:

▶ **Lemma 10** (Rank Upper Bound on Polynomial Matrices)**.** *Let $\mathbf{x} = (x_1, \ldots, x_n)$. If $M(\mathbf{x}) \in \mathrm{Mat}_m(\mathbb{F}[\mathbf{x}])$ is a matrix such that $\mathrm{rank}_{\mathbb{F}}(M(\mathbf{a})) \leq r$ for all $\mathbf{a} \in \mathbb{F}^n$, then $\mathrm{rank}_{\mathbb{F}(\mathbf{x})}(M(\mathbf{x})) \leq r$.*

The following proposition shows one way in which a linear space of matrices is of low rank. This decomposition and its variants will be very useful to us throughout the paper.

▶ **Proposition 11.** *Let $\mathcal{M} \subset \mathrm{Mat}_m(\mathbb{F})$ be a vector space of matrices such that $\mathcal{M} = \mathsf{span}(U \otimes V)$, where $U, V \subset \mathbb{F}^m$ are vector spaces of dimensions r and s, respectively. Then,*

$$\mathrm{rank}(\mathcal{M}) = \min(r, s).$$

A.3 Coefficient Spaces and Their Properties

As we saw in Section 1.3, linear spaces of matrices may possess special structure if they are generated by the coefficients of a matrix of polynomials. This observation, together with the definition below, are crucial in obtaining upper bounds for the rank techniques which we study.

▶ **Definition 12** (Coefficient Space)**.** Let $M(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]^{m \times k}$ be a symbolic matrix of polynomials. Considering the monomial basis $\{\mathbf{x}^{\mathbf{e}}\}_{\mathbf{e} \in \mathbb{N}^n}$ for the space $\mathbb{F}[\mathbf{x}]$, we can write $M(\mathbf{x}) = \sum_{\mathbf{e} \in \mathbb{N}^n} M_{\mathbf{e}} \cdot \mathbf{x}^{\mathbf{e}}$, where each $M_{\mathbf{e}} \in \mathbb{F}^{m \times k}$ is a matrix of field elements. We define the *coefficient space* of $M(\mathbf{x})$, denoted by $\mathcal{C}(M(\mathbf{x}))$, as the vector space spanned by the vectors $M_{\mathbf{e}}$. That is,

$$\mathcal{C}(M(\mathbf{x})) = \mathsf{span}\{M_{\mathbf{e}} \mid \mathbf{e} \in \mathbb{N}^n\}.$$

Note that $\mathcal{C}(M(\mathbf{x})) \subseteq \mathbb{F}^{m \times k}$.

Having the definition above, we proceed to show some nice properties of the coefficient space of a matrix of polynomials.

▶ **Proposition 13.** *Let $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_d)$ be a set of nd variables, partitioned into d sets of n variables each, denoted by $\mathbf{x}_i$. If $\mathbf{f}(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]^m$ is a vector of homogeneous and set-multilinear polynomials of degree d, with respect to the partition $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_d)$, then*

$$\dim(\mathcal{C}(\mathbf{f}(\mathbf{x}))) \leq n^d.$$

By using this new proposition and Propostion 11, we have the following corollary:

▶ **Corollary 14.** *Let $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_d)$ be a set of nd variables, partitioned into d sets of n variables each, denoted by $\mathbf{x}_i$. Additionally, let $S_f \sqcup S_g = [d]$ be a partition of the set $[d]$ such that $|S_f| = d_f$ and $|S_g| = d_g$. If $\mathbf{f}(\mathbf{x}), \mathbf{g}(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]^m$ are vectors of homogeneous set-multilinear polynomials, where $\mathbf{f}(\mathbf{x})$ is partitioned with respect to the variables $(\mathbf{x}_i)_{i \in S_f}$ and $\mathbf{g}(\mathbf{x})$ is partitioned with respect to the variables $(\mathbf{x}_i)_{i \in S_g}$, then we have:*

$$\operatorname{rank}(\mathcal{C}(\mathbf{f}(\mathbf{x}) \otimes \mathbf{g}(\mathbf{x}))) \leq \min\left\{n^{d_f}, n^{d_g}\right\}.$$

B Restricted Forms of Symbolic Matrix Rank Decompositions

If some matrix M over a field $\mathbb{F}$ has rank r, then we can write M as sum of r matrices $M = M_1 + \ldots + M_r$, where each M_i is a rank one matrix over $\mathbb{F}$, and thus can be written as $M_i = \mathbf{u}_i \otimes \mathbf{v}_i$, where $\mathbf{u}_i, \mathbf{v}_i$ are vectors over $\mathbb{F}$. In this section we would like to discus what happens when we impose additional conditions on the matrix M and on the rank one matrices M_i.

For instance, let $M(\mathbf{x}) \in \mathrm{Mat}_m(\mathbb{F}[\mathbf{x}])$ be a matrix of homogeneous polynomials of degree d such that $\operatorname{rank}_{\mathbb{F}(\mathbf{x})}(M) = r$. We want to know the minimal r' such that $M(\mathbf{x})$ can be written as sum of r' matrices $M_i(\mathbf{x})$ of rank one, where each $M_i(\mathbf{x})$ decomposes as $\mathbf{u}_i(\mathbf{x}) \otimes \mathbf{v}_i(\mathbf{x})$ for $\mathbf{u}_i(\mathbf{x}), \mathbf{v}_i(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]^m$ being vectors of homogeneous polynomials. Notice that this decomposition imposes the condition that the vectors $\mathbf{u}_i(\mathbf{x}), \mathbf{v}_i(\mathbf{x})$ be vectors of polynomials, whereas in the general rank decomposition these vectors could be vectors of rational functions, that is, elements of $\mathbb{F}(\mathbf{x})^m$.

In this section, we define one non-standard notion of rank, along with some properties which will be useful to us in the main sections of the paper. We begin with the definition of set-multilinear rank.

▶ **Definition 15** (Set-Multilinear Rank)**.** Let $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_d)$ be a set of variables, partitioned into sets of variables $\mathbf{x}_i$, and $M(\mathbf{x}) \in \mathrm{Mat}_m(\mathbb{F}[\mathbf{x}])$ be a matrix with polynomial entries such that each entry $M_{ij}(\mathbf{x})$ is a homogeneous set-multilinear polynomial of degree d, where the partition is given by $\mathbf{x}$.
The *set-multilinear rank* of $M(\mathbf{x})$, denoted by $\text{sm-rank}(M(\mathbf{x}))$, is the smallest integer r for which there exist r pairs of vectors $\mathbf{f}_i(\mathbf{x}), \mathbf{g}_i(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]^m$ such that

$$M(\mathbf{x}) = \sum_{i=1}^{r} \mathbf{f}_i(\mathbf{x}) \otimes \mathbf{g}_i(\mathbf{x}), \tag{1}$$

where:

- $\mathbf{f}_i(\mathbf{x})$ and $\mathbf{g}_i(\mathbf{x})$ are homogeneous vectors of set-multilinear polynomials,
- for each $i \in [r]$, there exists a partition $S_f^i \sqcup S_g^i = [d]$ of the set $[d]$ such that $\mathbf{f}_i(\mathbf{x})$ is set-multilinear with respect to the variables $(\mathbf{x}_j)_{j \in S_f^i}$ and $\mathbf{g}_i(\mathbf{x})$ is set-multilinear with respect to the variables $(\mathbf{x}_j)_{j \in S_g^i}$.

In particular, $\deg(\mathbf{f}_i(\mathbf{x})) + \deg(\mathbf{g}_i(\mathbf{x})) = d$.

Now that we have a notion of rank, we will need the following decomposition lemma to prove that low rank matrices must also have low set-multilinear rank. Let $M(\mathbf{x}) \in \mathrm{Mat}_m(\mathbb{F}[\mathbf{x}])$ be a matrix whose entries are homogeneous polynomials of degree d. The following lemma shows that if $\operatorname{rank}(M(\mathbf{x})) = r$, then it can be written as the homogeneous component of degree d of a sum of r rank one matrices with polynomial entries.

▶ **Lemma 16** (Symbolic Matrix Decomposition Lemma). *Let $M(\mathbf{x}) \in \mathrm{Mat}_m(\mathbb{F}[\mathbf{x}])$ be a matrix of homogeneous polynomials of degree d. If $\mathrm{rank}_{\mathbb{F}(\mathbf{x})}(M(\mathbf{x})) = r$ then there are vectors $\mathbf{f}_1(\mathbf{x}), \ldots, \mathbf{f}_r(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]^m$ and $\mathbf{g}_1(\mathbf{x}), \ldots, \mathbf{g}_r(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]^m$ such that*

$$M(\mathbf{x}) = \sum_{i=1}^{r} H_d[\mathbf{f}_i(\mathbf{x}) \otimes \mathbf{g}_i(\mathbf{x})].$$

Proof. Since $\mathrm{rank}_{\mathbb{F}(\mathbf{x})}(M(\mathbf{x})) = r$, there exist r pairs of vectors of polynomials $\mathbf{p}_i(\mathbf{x}), \mathbf{q}_i(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]^m$ and nonzero polynomials $t_i(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ such that

$$M(\mathbf{x}) = \sum_{i=1}^{r} \frac{1}{t_i(\mathbf{x})} \mathbf{p}_i(\mathbf{x}) \otimes \mathbf{q}_i(\mathbf{x}).$$

Since $t_i(\mathbf{x})$ are nonzero polynomials for all $i \in [r]$, the polynomial given by $Q(\mathbf{x}) = \prod_{i=1}^{r} t_i(\mathbf{x})$ is a nonzero polynomial. By $\mathrm{char}(\mathbb{F}) = 0$ and Lemma 7, there exists $\mathbf{a} \in \mathbb{F}^n$ such that $Q(\mathbf{a}) \neq 0$. In particular, this implies that we can write $t_i(\mathbf{x} + \mathbf{a}) = b_i \cdot (1 - \hat{t}_i(\mathbf{x}))$, where $b_i \in \mathbb{F}$ are nonzero field elements and $\hat{t}_i(\mathbf{x})$ are polynomials such that $\hat{t}_i(\mathbf{0}) = 0$. Namely, the constant terms of $\hat{t}_i(\mathbf{x})$ are zero, for all $i \in [r]$.

Writing $\widehat{\mathbf{p}}_i(\mathbf{x}) = \mathbf{p}_i(\mathbf{x} + \mathbf{a})$, $\widehat{\mathbf{q}}_i(\mathbf{x}) = \mathbf{q}_i(\mathbf{x} + \mathbf{a})$, and from the power series expansion of $1/(1 - x)$, it follows that

$$\begin{aligned} M(\mathbf{x} + \mathbf{a}) &= \sum_{i=1}^{r} \frac{1}{t_i(\mathbf{x} + \mathbf{a})} \widehat{\mathbf{p}}_i(\mathbf{x}) \otimes \widehat{\mathbf{q}}_i(\mathbf{x}) \\ &= \sum_{i=1}^{r} \frac{1}{b_i \cdot (1 - \hat{t}_i(\mathbf{x}))} \widehat{\mathbf{p}}_i(\mathbf{x}) \otimes \widehat{\mathbf{q}}_i(\mathbf{x}) \\ &= \sum_{i=1}^{r} \frac{1}{b_i} [\widehat{\mathbf{p}}_i(\mathbf{x}) \otimes \widehat{\mathbf{q}}_i(\mathbf{x})] \cdot \left(\sum_{j=0}^{\infty} \hat{t}_i(\mathbf{x})^j \right). \end{aligned}$$

As $M(\mathbf{x} + \mathbf{a})$ is a matrix of polynomials of degree no larger than d, the equality above becomes:

$$\begin{aligned} M(\mathbf{x} + \mathbf{a}) &= H_{\leq d}[M(\mathbf{x} + \mathbf{a})] \\ &= H_{\leq d} \left\{ \sum_{i=1}^{r} \frac{1}{b_i} [\widehat{\mathbf{p}}_i(\mathbf{x}) \otimes \widehat{\mathbf{q}}_i(\mathbf{x})] \cdot \left(\sum_{j=0}^{\infty} \hat{t}_i(\mathbf{x})^j \right) \right\} \\ &= H_{\leq d} \left\{ \sum_{i=1}^{r} \frac{1}{b_i} [\widehat{\mathbf{p}}_i(\mathbf{x}) \otimes \widehat{\mathbf{q}}_i(\mathbf{x})] \cdot \left(\sum_{j=0}^{d} \hat{t}_i(\mathbf{x})^j \right) \right\} \\ &= \sum_{i=1}^{r} H_{\leq d}[\widetilde{\mathbf{p}}_i(\mathbf{x}) \otimes \widetilde{\mathbf{q}}_i(\mathbf{x})], \end{aligned}$$

where $\widetilde{\mathbf{p}}_i(\mathbf{x}) = \frac{1}{b_i} \widehat{\mathbf{p}}_i(\mathbf{x})$ and $\widetilde{\mathbf{q}}_i(\mathbf{x}) = \widehat{\mathbf{q}}_i(\mathbf{x}) \cdot \left(\sum_{j=0}^{d} \hat{t}_i(\mathbf{x})^j \right)$.

Moreover, from homogeneity of $M(\mathbf{x})$, we have $M(\mathbf{x}) = H_d[M(\mathbf{x} + \mathbf{a})]$, which implies

$$M(\mathbf{x}) = H_d[M(\mathbf{x} + \mathbf{a})] = \sum_{i=1}^{r} H_d[\widetilde{\mathbf{p}}_i(\mathbf{x}) \otimes \widetilde{\mathbf{q}}_i(\mathbf{x})].$$

Taking $\mathbf{f}_i(\mathbf{x}) = \widetilde{\mathbf{p}}_i(\mathbf{x})$ and $\mathbf{g}_i(\mathbf{x}) = \widetilde{\mathbf{q}}_i(\mathbf{x})$ completes the proof. ◀

With this concept of set-multilinear decomposition and the lemma above, we obtain the following relationship between the symbolic rank and the set-multilinear rank of a set-multilinear polynomial matrix.

▶ **Lemma 17** (Set-Multilinear Rank of Polynomial Matrices). *Let $M(\mathbf{x}) \in \mathrm{Mat}_m(\mathbb{F}[\mathbf{x}])$ be a set-multilinear matrix of degree d, with partition $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_d)$.*

If $\mathrm{rank}_{\mathbb{F}(\mathbf{x})}(M(\mathbf{x})) \leq r$ then $\text{sm-rank}(M(\mathbf{x})) \leq r \cdot 2^d$.

Proof. W.l.o.g., we can assume that $\mathrm{rank}_{\mathbb{F}(\mathbf{x})}(M(\mathbf{x})) = r$. From Lemma 16, there exist vectors of polynomials $\mathbf{p}_1(\mathbf{x}), \mathbf{q}_1(\mathbf{x}), \ldots, \mathbf{p}_r(\mathbf{x}), \mathbf{q}_r(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]^m$ such that

$$M(\mathbf{x}) = \sum_{i=1}^{r} H_d[\mathbf{p}_i(\mathbf{x}) \otimes \mathbf{q}_i(\mathbf{x})]. \tag{2}$$

Decomposing equality (2) into its homogeneous and set multilinear components, according to the partition $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_d)$ we obtain:

$$M(\mathbf{x}) = \sum_{i=1}^{r} H^{SM}_{[d]}[\mathbf{p}_i(\mathbf{x}) \otimes \mathbf{q}_i(\mathbf{x})] = \sum_{i=1}^{r} \sum_{S \subseteq [d]} H^{SM}_{S}[\mathbf{p}_i(\mathbf{x})] \otimes H^{SM}_{[d]\setminus S}[\mathbf{q}_i(\mathbf{x})].$$

The last line of the equality above giving us the decomposition of $M(\mathbf{x})$ into $R \leq r \cdot 2^d$ rank-1 polynomial matrices. ◀