Weighted Sum-of-Squares Lower Bounds for Univariate Polynomials imply VP $\neq$ VNP

Pranjal Dutta * Nitin Saxena † Thomas Thierauf ‡

Abstract

For a polynomial f, a *weighted sum-of-squares representation (SOS)* has the form $f = \sum_{i\in[s]} c_i f_i^2$, where the *weights* c_i are field elements. The size of the representation is the number of monomials that appear across the f_i's. Its minimum across all such decompositions is called the *support-sum* $S(f)$ of f.

For a univariate polynomial f of degree d of full support, a lower bound for the support-sum is $S(f) \geq \sqrt{d}$. We show that the existence of an explicit univariate polynomial f with support-sum just slightly larger than the lower bound, that is, $S(f) \geq d^{0.5+\varepsilon}$, for some $\varepsilon > 0$, implies that VP $\neq$ VNP, *the* major open problem in algebraic complexity. In fact, our proof works for some subconstant functions $\varepsilon(d) > 0$ as well.

We also consider the *sum-of-cubes representation (SOC)* of polynomials. We show that an explicit hard polynomial implies both blackbox-PIT is in P, and VP $\neq$ VNP.

2010 Mathematics Subject Classification. Primary 03D15, 68Q17; Secondary 11E25, 12Y05.

Key words and phrases. Algebraic complexity, sum-of-squares, sum-of-cubes, VP, VNP, PIT.

Contents

*National University of Singapore, pranjal@nus.edu.sg
†CSE, Indian Institute of Technology, Kanpur, nitin@cse.iitk.ac.in
‡Aalen University, Germany, thomas.thierauf@uni-ulm.de

1 Introduction

We consider the *weighted sum-of-squares* (SOS) representation of polynomials over some field,

$$f = \sum_{i \in [s]} c_i f_i^2,$$

where the weights c_i are arbitrary field elements, i.e., they can also be negative in case of reals. We show a connection to central complexity questions with the parameters of the representation.

The *standard* SOS-representation in the literature mostly has no weights, i.e. all weights are 1, or, if there are weights, they are required to be positive (over the reals). The point is, that the SOS-representation can then take only non-negative values. There is a huge literature for this setting in mathematics, going back to Hilbert's *17th problem*, with applications in approximation, optimization and control theory [Rez78, Las07, Lau09, BM16]. Clearly, this non-negativity property is lost when we allow negative weights, and hence, we consider the representation from a somewhat different angle than the classic literature.

1.1 Algebraic circuits and univariate polynomials

Valiant defined the algebraic complexity classes VP and VNP based on algebraic circuits. They are considered as the algebraic analog of boolean classes P and NP. Separating VP from VNP is a long-standing open problem. One of the popular ways has been via depth-reduction results [AV08, Koi12, GKKS13, Tav15]. It seems that showing strong lower bounds require a deeper understanding of the algebraic-combinatorial structure of circuits, which may be easier to unfold for more analytic models that appear in wider mathematics.

Starting from the basics, it is known that most of the degree d univariate polynomials are *hard*, i.e. they require $\Omega(d)$ size circuits (see [CKW11, Theorem 4.2])[1]. For example, for p_i being the i-th prime, $\sum_{i=0}^{d} \sqrt{p_i}\, x^i$ and $\sum_{i=0}^{d} 2^{2^i} x^i$ both require circuits of size $\Omega(d/\log d)$ (see [BCS13, Str74]). Such polynomials can be transformed into multivariate polynomials that require exponential size circuits, i.e. they are *exponentially hard*. Unfortunately, these

[1] The size-bound in the literature usually counted only the number of nodes in the circuit, which gives a $\Omega(\sqrt{d})$ bound. When counting also the edges, the size is $\Omega(d)$.

strong lower bounds are insufficient to separate VP and VNP because the polynomials may not be in VNP (see [HS80b, Bür13] for details). The problem is, that the coefficients of the polynomials seem to be hard to compute, i.e., the polynomials are not *explicit*, a notion defined in the next section.

The interplay between proving lower bounds and derandomization is one of the central themes in complexity theory [NW94]. Blackbox Polynomial Identity Testing (PIT) asks for an algorithm to test the zeroness of a given algebraic circuit via mere query access. It is still an open problem to design an efficient (in circuit size) deterministic PIT algorithm. However, since a non-zero polynomial evaluated at a random point is non-zero with high probability (by the *Polynomial Identity Lemma* [Ore22, DL78, Zip79, Sch80]), one gets a randomized polynomial-time algorithm for PIT.

One important direction, from hardness to derandomization, is to design deterministic PIT algorithms for small circuits assuming access to *explicit multivariate hard polynomials* [NW94, KI04]. Most of the constructions use the concept of a *hitting-set generator* (HSG), see Definition 2.3. PIT is also amenable to the phenomenon of *bootstrapping* (w.r.t. variables) [AGS19, KST19]. This led Guo et al. [GKSS19] to show: ample circuit-hardness of *constant*-variate polynomials implies blackbox-PIT in P.

1.2 Sum-of-squares model (SOS)

We give some background on sum-of-square representation, give some examples, and define our hardness condition. We first define the model and a complexity measure.

Definition 1.1 (Weighted SOS and support-sum size $S_R(f)$)**.** *Let* R *be a ring. An* n*-variate polynomial* $f(\mathbf{x}) \in R[\mathbf{x}]$ *is represented as a* (weighted) sum-of-squares (SOS)*, if*

$$f = \sum_{i=1}^{s} c_i f_i^2 \,, \tag{1}$$

for some top-fanin s*, where* $f_i(\mathbf{x}) \in R[\mathbf{x}]$ *and* $c_i \in R$*.*

The size *of the representation of* f *in* (1) *is the* support-sum*, the sum of the support size (or sparsity) of the polynomials* f_i*. The* support-sum size *of* f*, is defined as the minimum support-sum of* f*, denoted by* $S_R(f)$*, or simply* $S(f)$*, when the ring* R *is clear from the context.*

Remark. In real analysis, the SOS-representation of a polynomial is defined without the coefficients c_i, that is, only for non-negative polynomials f. In these terms, what we define in (1) is a *weighted* SOS. However, we will skip the term "weighted" in the following.

Consider the expression in (1) as a $\sum \bigwedge^2 \sum \prod$-formula, i.e. as a depth-4 layered circuit with a top sum-gate followed by a wedge-2-gate that represents the squaring operation, followed by a depth-2 sum-product circuit that represents the f_i-polynomials as a sum of monomials. Then the support-sum is the number of $\prod$-operations directly above the input level.

For any N-variate polynomial f, let $\mathrm{sp}(f)$ denote the sparsity (i.e., the number of monomials) of f. For any field $R = \mathbb{F}$ of characteristic $\neq 2$, we have

$$\mathrm{sp}(f)^{1/2} \;\leq\; S_{\mathbb{F}}(f) \;\leq\; 2\,\mathrm{sp}(f) + 2\,.$$

The lower bound can be shown by counting monomials. The upper bound is because

$$f \;=\; (f+1)^2/4 - (f-1)^2/4\,. \tag{2}$$

In particular, when f is univariate of degree d and has full sparsity, $\mathrm{sp}(f) = d+1$, we get

$$\sqrt{d} \;\le\; S(f) \;\le\; 2d+2\,. \tag{3}$$

By (2), the SOS-model is *complete* for any field of characteristic $\neq 2$. It can be argued by a dimension-counting argument that for most N-variate (constant $N \geq 1$) polynomials f of degree d, we have $S_{\mathbb{F}}(f(\mathbf{x})) = \Theta(d^N)$, as for random f, we know that $\mathrm{sp}(f) = \Theta(d^N)$. Note that this matches the upper bound given in (3) for univariate f.

We give two examples.

Example 1. Let $f(x) = \sum_{k=0}^{d-1} x^k$. Note that

$$\sum_{k=0}^{d-1} x^k \;=\; \left(\sum_{k=0}^{\sqrt{d}-1} x^k\right)\left(\sum_{k=0}^{\sqrt{d}-1} x^{k\sqrt{d}}\right)\,.$$

Hence, we have a representation of f as $f = gh$, where $\mathrm{sp}(g), \mathrm{sp}(h) \leq \sqrt{d}$. Such a product can be written as a SOS,

$$gh = \frac{(g+h)^2}{4} - \frac{(g-h)^2}{4} \tag{4}$$

Because $\mathrm{sp}(g \pm h) \leq 2\sqrt{d}$ we get that $S(f) \leq 4\sqrt{d}$.

Observe that $S(f)$ essentially hits the lower bound in (3), except for a constant factor.

Example 2. Let $f(x) = (x+1)^d$. This has a trivial SOS-representation with one summand: $(x+1)^d = \left((x+1)^{d/2}\right)^2$, for even d. So we get $S(f) \leq d/2 + 1$.

Note that this bound meets the upper bound in (3), except for a constant factor. We conjecture that it is optimal, i.e. that $S(f) = \Omega(d)$. This is somewhat in contrast to that f has small circuits. By repeated squaring, the circuit size of f_d is $O(\log d)$.

We call a polynomial family SOS-hard, if its support-sum is just *slightly* larger than the trivial lower bound from (3). For our results, it actually suffices to consider *univariate* polynomials. We discuss more on univariate SOS-hardness vs. multivariate SOS-hardness at the end of this section.

Definition 1.2 (SOS-hardness)**.** *Let* $(f_d(x))_d$ *be a polynomial family, where* f_d *has degree* d*. Family* $(f_d(x))_d$ *is* SOS-hard with hardness ε*, if* $S(f_d) = \Omega(d^{0.5+\varepsilon})$.

Main results. Our main results with respect to SOS-representation show that the existence of explicit SOS-hard families of polynomials imply circuit lower bounds. The precise bounds depend on the size of ε:

1. For $\varepsilon = \omega(1/\sqrt{\log d})$, we show that the permanent cannot be computed by small ABPs, i.e., VBP $\neq$ VNP (Corollary 3.6).

2. For $\varepsilon = \omega(\sqrt{\log\log d/\log d})$ we show that the permanent cannot be computed by small circuits, i.e., VP $\neq$ VNP (Theorem 3.2).

3. For $\varepsilon > 0$ constant, we show that the permanent requires exponential size circuits, i.e., we have an exponential separation of VP and VNP (Theorem 4.2).

The technical foundation for these results are SOS-decompositions for circuits (Lemma 3.1 and 3.8) that are based on the known depth-reductions techniques. We show how to express a polynomial $p(\mathbf{x})$ of degree d, given by a circuit of size s, as a sum of squares

- of $\mathtt{quasi\text{-}poly}(d,s)$ many polynomials, i.e. $2^{\mathrm{poly}(\log sd)}$ many, each of degree at most $d/2$, in case of Lemma 3.1, and
- of $\mathrm{poly}(s)$ many polynomials, each of degree close to $d/2$, in case of Lemma 3.8.

Hence, by our results, the major challenge in arithmetic complexity, to separate VP from VNP, can be solved by exhibiting an explicit univariate polynomial family $f_d(x) \in \mathbb{C}[x]$ of degree d with SOS-hardness parameter ε, just slightly above the general lower bound, even for vanishing small $\varepsilon = \varepsilon(d)$.

This would also have consequences for PIT, because Kabanets and Impagliazzo [KI04, Theorem 7.7] showed that VP $\neq$ VNP implies blackbox-PIT $\in$ SUBEXP. In the case of constant ε, where we have an exponential separation of VP and VNP, we get blackbox-PIT $\in$ QP (*quasi*-polynomial time).

Remark. Our results hold similarly when we use a *multivariate* SOS-hard polynomial as starting point instead of a *univariate* polynomial. A polynomial f with n variables and degree d has full sparsity $\binom{n+d}{d}$. Hence, the SOS-hardness condition would be $S(f) = \Omega(\binom{n+d}{d}^{0.5+\varepsilon})$.

We give our results in terms of univariate polynomials because, intuitively, it seems easier to prove the hardness of a univariate polynomial than of a multivariate polynomial. Also, the proofs are technically somewhat easier because we save the parameter for the number of variables.

1.3 Sum-of-cubes model (SOC)

It is not clear whether a strong lower bound in the SOS-model can give a polynomial-time blackbox-PIT. However, a different complexity measure on the sum-of-cube representation of polynomials indeed leads to a complete derandomization of blackbox-PIT. We give a more detailed technical explanation for the reason to use a different measure just before in the outline just before the main SOC-theorem (Theorem 4.2). We start by defining the model and give some background on it.

Definition 1.3 (SOC and support-union size $U_R(f,s)$)**.** *Let* R *be a ring. An* n*-variate polynomial* $f(\mathbf{x}) \in R[\mathbf{x}]$ *is represented as a* sum-of-cubes (SOC)*, if*

$$f = \sum_{i=1}^{s} c_i f_i^3 \,, \tag{5}$$

for some top-fanin s*, where* $f_i(\mathbf{x}) \in R[\mathbf{x}]$ *and* $c_i \in R$.

The size *of the representation of* f *in* (5) *is the size of the* support-union*, namely the number of distinct monomials in the representation,* $\left|\bigcup_{i=1}^{s} \mathrm{supp}(f_i)\right|$*, where* support $\mathrm{supp}(f_i)$ *denotes the set of monomials with a nonzero coefficient in* $f_i(\mathbf{x})$*. The* support-union size *of* f *with respect to* s*, denoted* $U_R(f,s)$*, is defined as the minimum support-union size when* f *is written as in (5).*

If we consider the expression in (5) as a $\sum\bigwedge^3\sum\prod$-circuit, then the support-union size is the number of *distinct* $\prod$-operations directly above the input level.

The support-sum measure is potentially larger than the support-union measure since the same monomial can be counted several times in the support sum, but only once in the support-union. More formally, the two measures are largely incomparable in our case, since $U(\cdot)$ has the extra argument s and is defined for a cubic representation. Still one can show: $S_{\mathbb{F}}(f) \geq \min_s \left(U_{\mathbb{F}}(f,4s) - 1\right)$ (Lemma 6.9).

For any polynomial f of sparsity $\mathrm{sp}(f)$, we have

$$\mathrm{sp}(f)^{1/3} \;\leq\; U_{\mathbb{F}}(f,s) \;\leq\; \mathrm{sp}(f) + 1,$$

where the upper bound is for $s \geq 3$ and for fields $R = \mathbb{F}$ of characteristic $\neq 2,3$. The lower bound can be shown by counting monomials. The upper bound is because

$$f \;=\; (f+2)^3/24 + (f-2)^3/24 - f^3/12\,.$$

Hence, the SOC-model is *complete* for any field of characteristic $\neq 2,3$.

In particular, when f is univariate of degree d and has full sparsity, $\mathrm{sp}(f) = d+1$, we get

$$d^{1/3} \;\leq\; U_{\mathbb{F}}(f,s) \;\leq\; d+1\,. \tag{6}$$

More bounds and examples for the trade-off between s and the measure $U(f,s)$ can be found in Section 6. Here, we summarize:

Example 3.

1. For small $s = \Theta(d^{1/2})$, we have $U(f,s) = O(d^{1/2})$ (Corollary 6.6).
2. For large $s = \Omega(d^{2/3})$, we have $U(f,s) = \Theta(d^{1/3})$ (Theorem 6.8).

However, it is unclear whether it is possible to have a very small fanin s, like $s = o(\sqrt{d})$, and at the same time a support-union of $o(d)$. This motivated us to define the hardness of univariate polynomials in the SOC-model as follows.

Definition 1.4 (SOC-hardness)**.** *A polynomial family* $(f_d(x))_d$ *is* SOC-hard*, if there is a constant* $0 < \varepsilon < 1/2$ *such that* $U_{\mathbb{F}}\left(f_d, d^{\varepsilon}\right) = \Omega(d)$.

Main results. Our main result with respect to SOC-representation shows that the existence of an explicit SOC-hard family of polynomials leads to a *complete* derandomization of blackbox-PIT (Theorem 4.2). In fact, the proof produces an explicit constant-variate hard polynomial family, from which also the separation of VP and VNP follows.

The technical basis for our result is again a decomposition lemma (Lemma 4.1), an extension of Lemma 3.8. It shows how to express a polynomial $p(\mathbf{x})$ of degree d, given by a circuit of size s, as a sum of cubes of $\mathrm{poly}(s)$ many polynomials, each of degree close to $d/3$.

A similar remark with respect to univariate vs. multivariate hardness at the end of Section 1.2 for SOS holds for SOC as well.

1.4 SOS (and SOC)-hardness and Geometric Complexity Theory (GCT)

In computer science, the notion of approximative complexity emerged in the context of tensors for matrix multiplication, i.e. the notion of border rank; see [LL89, BCS13] and references therein. Bürgisser [Bür04] used this concept in the context of arithmetic circuits. Approximative closure of algebraic complexity classes is of great interest in the GCT program [MS01, MS08, GMQ16, Mul17], which aims to study the symmetries of different actions of groups on algebraic varieties, and settle a stronger version of the permanent vs. determinant problem. The SOS- and SOC-hardness, defined above, can also be extended in the border or approximative complexity-theoretic sense, which would eventually strengthen the lower bound and PIT consequences. For formal definitions and related results in the GCT paradigm, we refer to Section 5.

1.5 Small SOS- and SOC-representations

In Section 6, we study small representations of a generic univariate polynomial. Eventually we show that every univariate d-degree polynomial can be *optimally* represented as a sum of powers, where the measure is support-union, see Theorem 6.1; for example any univariate polynomial of degree d has an SOC-representations with top-fanin $O(d)$ and support-union $O(d^{1/3})$, and further a simple counting argument shows a lower bound of $\Omega(d^{1/3})$, for a d-sparse univariate polynomial. However, our construction requires the top-fanin to be *large*. Subsequently, we show a nice trade-off between the top-fanin and support-union size, for both SOS- and SOC-representations; for details see Theorem 6.4 and 6.8. The algorithmic essence of these constructions make the trade-off very interesting.

1.6 Related works

In this particular section, we will use weighted SOS, or SOC, for the clarity of comparison with prior works. We will not use the term weighted later on, but we *only* work with the weighted models throughout the paper.

Real τ-conjecture for weighted SOS. The τ-*conjecture* by Shub and Smale [SS95, Sma98] is about the number of *distinct integer roots* of a polynomial. The *real-τ-conjecture* by Koiran [Koi11] is about the number of *distinct real roots* of certain polynomials: any polynomial computed by a $\sum^k \prod^m \sum^t \wedge$-circuit has most $\mathrm{poly}(kmt)$ distinct real roots.

The first author [Dut21, Dut22] came up with a real τ-conjecture for polynomials with respect to the weighted SOS-representation: a polynomial f has at most $S(f)$ distinct real roots. He showed that if the conjecture is true, we get again $\mathsf{VP} \neq \mathsf{VNP}$.

SOS to non-commutative hardness. Hrubeš, Wigderson, and Yehudayoff [HWY11] considered the sum-of-squares representation in the *non-commutative* setting. They showed that lower bounds for the SOS-representation of a specific multivariate polynomial imply exponential lower bounds on the circuit size of the permanent. Besides the non-commutative algebra, their setting differs in the precise SOS-model and the complexity measure. So, hardly any comparison is possible.

Depth-4 circuits with unbounded powering. Much of the previous works are concerned with multivariate depth-4 circuits that are a sum of unbounded-powers, i.e., $\sum \bigwedge^{\omega(1)} \sum \prod$-circuits, because this is the form one gets after applying the depth-reduction results [AV08, Koi12, GKKS13, AGS19]. The sufficiency of proving lower bound on restricted models of *univariate* polynomials was shown by Koiran [Koi11]. He considered univariate explicit polynomials f_d of degree d over an algebraically closed field $\mathbb{F}$ that are written as

$$f_d(x) = \sum_{i=1}^{s} c_i \, Q_i^{e_i}(x) \,,$$

where $c_i \in \mathbb{F}$, and Q_i are polynomials with sparsity $sp(Q_i) \leq t$ with unbounded exponents $e_i \geq 1$. He showed that when every such presentation of f_d requires $s = \left(\frac{d}{t}\right)^{\Omega(1)}$ summands, then VP $\neq$ VNP.

Some initial lower bounds have been established for this model.

- When $\deg(Q_i) \leq t$, there is a family such that $s \geq \Omega(\sqrt{d/t})$ [KKPS15].
- For $\deg(Q_i) \leq 1$, the bound $s \geq \Omega(d)$ has been established for certain polynomials using the concept of *Birkhoff Interpolation* [GMK17, KPGM18].

Clearly, allowing arbitrary exponents gives much more flexibility than fixed exponents as in weighted SOS and SOC. In that sense, it should be easier to obtain lower bounds in the weighted SOS- or SOC-model. Also the complexity measure is different, as Koiran considers the number of summands, whereas we consider the support-sum.

Existence of $(r, 2)$-elusive function vs. weighted SOS-hardness. Raz [Raz10] formalized a notion of elusive maps and established a connection between the existence of explicit elusive maps and VP vs. VNP. A polynomial map $L : \mathbb{F}^n \to \mathbb{F}^m$ is $(r, 2)$-*elusive*, if for every polynomial of degree 2 that maps $M : \mathbb{F}^r \to \mathbb{F}^m$, we have $\text{Image}(L) \not\subseteq \text{Image}(M)$. Formally, Raz showed that any explicit polynomial map which is $(r, 2)$-elusive, with $m = n^{\omega(1)}$ and $r = n^{0.9}$, implies VP $\neq$ VNP.

Observe that one can reinterpret the coefficients of the f_i^2's in Equation (1) as expressing $\text{coef}(f)$ via quadratic forms, like M. However, the elusiveness notion is *too* general in the following sense: the parameters r and m have a super-polynomial large gap, and still M has to elude all L. On the other hand, weighted SOS-hardness, say for $N = 1$, goes a step further and optimizes the gap to be almost quadratic, i.e, $r = n^{0.5+\epsilon}$ and $m = n$. Further, weighted SOS gives a rather specialized degree-2 polynomial mapping.

From hardness to derandomization. With respect to the derandomization of blackbox-PIT, there are a few conditional results. For example, it has been shown that multivariate hard polynomials lead to blackbox-PIT $\in$ QP (*quasi-poly time*) [KI04, AGS19]. Closer to our result is the work of Guo et al. [GKSS19]. They showed that the circuit hardness of a constant-variate polynomial family yields blackbox-PIT $\in$ P (Theorem 2.4). Still, our hardness assumption is merely in the SOC-model and for univariate polynomials. For now, weighted SOC seems to be the simplest model where hardness implies a complete derandomization.

2 Preliminaries

Basic notation. We work with fields $\mathbb{F} = \mathbb{Q}, \mathbb{Q}_p$, or their fixed extensions. Our results hold also for fields with large enough characteristic.

We denote $[n] = \{1, \ldots, n\}$. For $i \in \mathbb{N}$ and $b \geq 2$, we denote by $\text{base}_b(i)$ the unique k-tuple $(i_1, \ldots, i_k)$ such that $i = \sum_{j=1}^{k} i_j \cdot b^{j-1}$.

For binomial coefficients, we use the following well known bounds. For $1 \leq k \leq n$,

$$\left(\frac{n}{k}\right)^k \leq \binom{n}{k} \leq \left(\frac{en}{k}\right)^k. \tag{7}$$

Polynomials. For $p \in \mathbb{F}[\boldsymbol{x}]$, where $\boldsymbol{x} = (x_1, \ldots, x_m)$, for some $m \geq 1$, the *support of* p, denoted by $\text{supp}(p)$, is the set of nonzero monomials in p. The *sparsity* or *support size of* p is $\text{sp}(p) := |\text{supp}(p)|$. If p is m-variate of degree d, its sparsity is bounded by

$$\text{sp}(p) \leq \binom{m+d}{d}. \tag{8}$$

By $\text{coef}(p)$ we denote the *coefficient vector* of p (in some fixed order).

For an exponent vector $\boldsymbol{e} = (e_1, \ldots, e_m)$, we use $\boldsymbol{x}^{\boldsymbol{e}}$ to denote the monomial $x_1^{e_1} \cdots x_m^{e_m}$. For a polynomial $p(\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{F}[\boldsymbol{x}, \boldsymbol{y}]$, the $\boldsymbol{x}$-*degree of* p, denoted by $\deg_{\boldsymbol{x}}(p)$, is the maximum degree of $\boldsymbol{x}$ in p. That is, for $p(\boldsymbol{x}, \boldsymbol{y}) = \sum_{\boldsymbol{e}} p_{\boldsymbol{e}}(\boldsymbol{x})\boldsymbol{y}^{\boldsymbol{e}}$, we define $\deg_{\boldsymbol{x}}(p) = \max_{\boldsymbol{e}} \deg(p_{\boldsymbol{e}}(\boldsymbol{x}))$.

Kronecker map and its inverse. The *Kronecker substitution* is a bijective map between univariate and multivariate polynomials. We define two variants: The first one is the standard one, the second one is a multilinear version of it. In our application, we consider the sparsity of the polynomials. There it seems as the standard Kronecker substitution does not yield the bounds we need. Let $p(x)$ be a nonzero univariate polynomial of degree d.

1) *Standard Kronecker substitution.* We map p to a k-variate polynomial in variables $\boldsymbol{x} = (x_1, \ldots, x_k)$, for a given k. Think of $k \leq \log d$. We define n such that

$$n^k \leq d < (n+1)^k. \tag{9}$$

The *Kronecker map* $\phi_{k,n}$ is defined as

$$\phi_{k,n} : x^i \mapsto \boldsymbol{x}^{\text{base}_{n+1}(i)}, \tag{10}$$

for all $i \in [d]$. By linear extension, define polynomial $P_{k,n} = \phi_{k,n}(p)$. Note that $\phi_{k,n}$ maps each x^i to a *distinct* k-variate monomial of individual degree $\leq n$, for $0 \leq i \leq d$.

Next, we consider the inverse map. Let $P(x_1, \ldots, x_k)$ be a polynomial, where the variables have individual degree bounded by n. Define $\psi_{k,n}$ by

$$\psi_{k,n} : x_i \mapsto x^{(n+1)^{i-1}},$$

for $0 \leq i \leq k$, and $\psi_{k,n}(P)$ by linear extension. The degree of $\psi_{k,n}(P)$ is bounded by $\sum_{i=1}^{k} n(n+1)^{i-1} = (n+1)^k - 1$ [Kro82].

Note that $\phi_{k,n}$ and $\psi_{k,n}$ are defined in principal for every k ($\leq \log d$). However, the properties we want are only fulfilled for the right n which depends on d as defined in (9). In this case we have $\psi_{k,n} \circ \phi_{k,n}(p) = p$.

2) *Multilinear Kronecker substitution.* Let k again be given. Here, we choose n such that $(k-1)^n \leq d < k^n$. Introduce kn variables $y_{j,\ell}$, where $1 \leq j \leq n$ and $0 \leq \ell \leq k-1$. For every $i = 0, 1, \ldots, d$, write i in base-k representation, $\mathrm{base}_k(i) = (i_1, \ldots, i_n)$. Define the injective map $\phi^{\mathrm{lin}}_{k,n}$ by

$$\phi^{\mathrm{lin}}_{k,n} : x^i \mapsto \prod_{j=1}^{n} y_{j,i_j} . \tag{11}$$

By linear extension, define polynomial $P_{k,n} = \phi^{\mathrm{lin}}_{k,n}(p)$. Note that $P_{k,n}$ is a kn-variate multilinear, homogeneous polynomial of degree n.

Mapping $\phi^{\mathrm{lin}}_{k,n}$ can be inverted by $\psi^{\mathrm{lin}}_{k,n}$,

$$\psi^{\mathrm{lin}}_{k,n} : y_{j,\ell} \mapsto x^{\ell \cdot k^{j-1}} . \tag{12}$$

Again by linear extension and with n as defined above, we have $\psi^{\mathrm{lin}}_{k,n} \circ \phi^{\mathrm{lin}}_{k,n}(p) = p$.

It is also important to note that the sparsity of the polynomials stays the same, for the standard and the multilinear Kronecker map and their inverses.

Algebraic circuits. An *algebraic circuit* over a field $\mathbb{F}$ is a layered directed acyclic graph that uses field operations $\{+, \times\}$ and computes a polynomial. It can be thought of as an algebraic analog of boolean circuits. The leaf nodes are labeled with the input variables $x_1, \ldots, x_n$ and constants from $\mathbb{F}$. Other nodes are labeled as addition and multiplication *gates*. The root node outputs the polynomial computed by the circuit.

Complexity parameters of a circuit are: **1)** *size*, i.e. the number of edges and nodes, **2)** *depth*, i.e. the number of layers, **3)** *fanin* and *fan-out*, i.e. the maximum number of inputs to, respectively, outputs of a node.

When the graph is in fact a tree, i.e., the fan-out is 1, we call the circuit an *algebraic formula*.

For a polynomial f, the size of the smallest circuit that computes f is denoted by $\mathrm{size}(f)$, it is the *algebraic circuit complexity* of f. By $\mathcal{C}(n, d, s)$, we denote the set of circuits C that compute n-variate polynomials of degree d such that $\mathrm{size}(C) \leq s$.

Algebraic complexity classes. Valiant's class VP contains the families of n-variate polynomials of degree $\mathrm{poly}(n)$, i.e. polynomial in n, over $\mathbb{F}$, computed by circuits of size $\mathrm{poly}(n)$. A family of n-variate polynomials $(f_n)_n$ over $\mathbb{F}$ is in VNP, if there exists a family of polynomials $(g_n)_n \in \mathsf{VP}$ such that for every $\boldsymbol{x} = (x_1, \ldots, x_n)$ one can write $f_n(\boldsymbol{x}) = \sum_{w \in \{0,1\}^{t(n)}} g_n(\boldsymbol{x}, w)$, for some polynomial $t(n)$ which is called the *witness size*. It is straightforward to see that $\mathsf{VP} \subseteq \mathsf{VNP}$, and further Valiant's Hypothesis [Val79] conjectures that these two classes are *different*. For more details see [Mah14, SY10, BCS13]. Unless specified otherwise, we consider the field $\mathbb{F} = \mathbb{Q}$ (resp. a finite field with large characteristic).

Valiant [Val79] showed a *sufficient* condition for a polynomial family $(f_n(\boldsymbol{x}))_n$ to be in VNP. We use a slightly modified version of the criterion and formulate it only for multilinear polynomials.

Theorem 2.1 (VNP criterion, [Val79], see also[Bür13]). *Let* $f_n(\mathbf{x}) = \sum_{\mathbf{e}\in\{0,1\}^n} c_n(\mathbf{e})\,\mathbf{x}^{\mathbf{e}}$ *be a polynomial family such that the coefficients* $c_n(\mathbf{e})$ *have length* $\leq n$ *in binary. Then*

$$c_n(\mathbf{e}) \in \#\mathsf{P}/\mathrm{poly} \implies f_n \in \mathsf{VNP}.$$

One can further relax Theorem 2.1 such that the coefficients $c_n(\mathbf{e})$ can actually be 2^n bits long. Koiran and Perifel [KP11, Lem. 3.2] used a similar idea. We also use the fact that VNP is closed under substitution. That is, for a family of polynomials $(f(\mathbf{x},\mathbf{y})) \in \mathsf{VNP}$, it also holds that $(f(\mathbf{x},\mathbf{y}_0)) \in \mathsf{VNP}$, for any value $\mathbf{y}_0 \in \mathbb{F}^n$ assigned to the variables in $\mathbf{y}$.

Theorem 2.2. *Let* $f_n(\mathbf{x}) = \sum_{\mathbf{e}\in\{0,1\}^n} c_n(\mathbf{e})\,\mathbf{x}^{\mathbf{e}}$ *be a polynomial family such that the coefficients* $c_n(\mathbf{e})$ *have length* $\leq 2^n$ *in binary. Let* $c_{n,j}(\mathbf{e})$ *be the* j*-th bit of* $c_n(\mathbf{e})$*. Then*

$$\forall j,\ c_{n,j}(\mathbf{e}) \in \#\mathsf{P}/\mathrm{poly} \implies f_n \in \mathsf{VNP}.$$

Proof. For $j \in \{0,1,\ldots,2^n-1\}$ let $\mathrm{bin}(j) = (j_1,\ldots,j_n)$ denote the n-bit base-2 representation of j such that $j = \sum_{i=1}^{n} j_i\,2^{i-1}$. Introduce new variables $\mathbf{y} = (y_1,\ldots,y_n)$ and define $\tilde{c}_n(\mathbf{e},\mathbf{y}) = \sum_{j=0}^{2^n-1} c_{n,j}(\mathbf{e})\,\mathbf{y}^{\mathrm{bin}(j)}$. Let $\mathbf{y}_0 = (2^{2^0},\ldots,2^{2^{n-1}})$. Then we have $\tilde{c}_n(\mathbf{e},\mathbf{y}_0) = c_n(\mathbf{e})$. Finally, consider the $2n$-variate auxiliary polynomial $h_n(\mathbf{x},\mathbf{y})$.

$$h_n(\mathbf{x},\mathbf{y}) \;=\; \sum_{\mathbf{e}\in\{0,1\}^n} \tilde{c}_n(\mathbf{e},\mathbf{y})\,\mathbf{x}^{\mathbf{e}} \;=\; \sum_{\mathbf{e}\in\{0,1\}^n}\ \sum_{j=0}^{2^n-1} c_{n,j}(\mathbf{e})\,\mathbf{y}^{\mathrm{bin}(j)}\,\mathbf{x}^{\mathbf{e}}\ .$$

Then we have $h_n(\mathbf{x},\mathbf{y}_0) = f_n(\mathbf{x})$. Since $c_{n,j}(\mathbf{e})$ can be computed in $\#\mathsf{P}/\mathrm{poly}$, we have $(h_n(\mathbf{x},\mathbf{y}))_n \in \mathsf{VNP}$. As VNP is closed under substitution, it follows that $(f_n(\mathbf{x}))_n \in \mathsf{VNP}$. □

Explicit univariate polynomials. We will consider univariate polynomials and define associated multivariate polynomials via Kronecker maps. We want all of these polynomials to be in VNP. For this, we use Theorem 2.2.

Let $(f_d)_d$ be a univariate polynomial family, where $f_d(x)$ is of degree d. The family is called *explicit*, if its coefficient-function is computable in $\#\mathsf{P}/\mathrm{poly}$ and each coefficient can be at most $\mathrm{poly}(d)$-bits long. The coefficient-function gets input (j,i,d) and outputs the j-th bit of the coefficient of x^i in f_d.

An *explicit* candidate for the hard family is the *Pochhammer-Wilkinson* polynomial, $f_d(x) := \prod_{i=1}^{d}(x-i)$. Other explicit families are $(x+1)^d$ and the *Chebyshev* polynomial (that writes $\cos d\theta$ as a function of $\cos\theta$) [MH02], and also $\sum_{i\in[d]} 2^{i^2}x^i$.

Hitting-set generators and blackbox-PIT from lower bounds. The technical tool to solve blackbox-PIT is to construct an efficient hitting-set generator.

Definition 2.3 (Hitting-set generator (HSG)). *A polynomial map* $G : \mathbb{F}^k \to \mathbb{F}^n$ *given by* $G(z) = (g_1(z), g_2(z), \ldots, g_n(z))$ *is a* hitting-set generator (HSG) *for a class* $\mathcal{C} \subseteq \mathbb{F}[\mathbf{x}]$ *of polynomials, if for every nonzero* $f \in \mathcal{C}$*, we have that* $f \circ G = f(g_1, g_2, \ldots, g_n)$ *is nonzero.*

We say that G *is* t-time HSG, *if* $\mathrm{coef}(g_i)$ *can be computed in time* t *and the maximum degree of* g_i *is* $\leq t$.

Given a HSG, one can construct a *hitting-set*, a set H such that a non-zero circuit is non-zero at some points in H. Crucial here is the size of H which depends on the parameters of the HSG. A t-time HSG G gives a $(td)^{O(k)}$ time blackbox-PIT algorithm, for circuits that compute polynomials of degree $\leq d$, over popular fields like rationals $\mathbb{Q}$ or their extensions, local fields $\mathbb{Q}_p$ or their extensions, or finite fields $\mathbb{F}_q$. When k is constant, we get a poly-time blackbox-PIT.

Guo et al. [GKSS19] showed how to use the hardness of a *constant* variate explicit polynomial family to derandomize PIT. They need the algebraic circuit hardness to be more than d^3; which requires $k \geq 4$ for the family to exist.

Theorem 2.4. [GKSS19] *Let $P \in \mathbb{F}[\mathbf{x}]$ be a k-variate polynomial of degree d such that $\mathrm{coef}(P)$ can be computed in $\mathrm{poly}(d)$-time. If $\mathrm{size}(P) > s^{10k+2}\, d^3$, then there is a $\mathrm{poly}(s)$-time HSG for $C(s,s,s)$.*

Algebraic branching programs (ABP). An *algebraic branching program (ABP)* in variables $\mathbf{x}$ over a field $\mathbb{F}$ is a directed acyclic graph with a *starting vertex* s with in-degree zero, an *end vertex* t with out-degree zero. The edge between any two vertices is labeled by an affine form $a_1x_1 + \ldots + a_nx_n + c \in \mathbb{F}[\mathbf{x}]$, where $a_i, c \in \mathbb{F}$.

The *weight of a path* in an ABP is the product of labels of the edges in the path. The *polynomial computed at a vertex v* is the sum of weights of all paths from the starting vertex s to v. The *polynomial computed by the ABP* is the polynomial computed at the end vertex t.

An ABP can be seen as a very restricted circuit, but still being able to compute determinants [MV99]. The class VBP contains all families of n-variate polynomials that can be computed by ABPs of size $\mathrm{poly}(n)$. This implies that the degree is $\mathrm{poly}(n)$ too. Clearly, $\mathsf{VBP} \subseteq \mathsf{VP}$.

We say that an ABP is *homogeneous*, if the polynomial computed at every vertex is a homogeneous polynomial. It is known that for an ABP B of size s that computes a homogeneous polynomial $p(\mathbf{x})$, there is an equivalent homogeneous ABP B' of size $\mathrm{poly}(s)$, where each edge-label is a *linear form* $a_1x_1 + \cdots + a_nx_n$. Moreover, when p has degree d, then B' has $d+1$ layers and each vertex in the ℓ-th layer computes a homogeneous polynomial of degree ℓ (see [IL17, Thm. 4.1(5)], [Kum19, Lem.15] or [Sap19]).

We remark that each homogeneous part of a polynomial $p(\mathbf{x})$ of degree d, computed by s-size circuit, can also be computed by a *homogeneous* circuit of size $O(sd^2)$, see [SY10, Sap19].

3 Sum of Squares

In this section, let $\mathbb{F}$ be a field of characteristic $\neq 2, 3$.

3.1 From SOS-hardness to $\mathsf{VP} \neq \mathsf{VNP}$

The connection between the SOS-model and general circuits is mainly established by the next lemma. It shows that a multivariate polynomial $p(\mathbf{x})$ of degree d, computed by a

circuit of size s, has a SOS-representation with $(sd)^{O(\log d)}$ summands, where each summand polynomial has degree precisely $d/2$.

This is achieved by transforming the given circuit for $p(\mathbf{x})$ in several steps into a *homogeneous* ABP. The point here is that degrees of the polynomials computed at the intermediate nodes of the ABP increase gradually, as the labels are linear forms. In particular, there exists a layer of vertices that computes polynomials of degree exactly $d/2$. By cutting the ABP at that layer, we get a representation of p as a sum of products of two polynomials of degree $d/2$ each. This immediately yields the desired SOS-representation.

We present a similar SOS-decomposition in Lemma 3.8 below. It uses the frontiers based depth-reduction technique [VSBR83]. However, this approach yields intermediate polynomials of degree only close to $d/2$, whereas we want degree exactly $d/2$ here.

Lemma 3.1 (SOS Decomposition). *Let $p \in \mathbb{F}[\mathbf{x}]$ be an n-variate polynomial of degree d, with $\text{size}(p) = s$.*

Then there exist $p_i \in \mathbb{F}[\mathbf{x}]$ and $c_i \in \mathbb{F}$ such that

$$p(\mathbf{x}) \;=\; \sum_{i=1}^{s'} c_i p_i(\mathbf{x})^2, \tag{13}$$

for $s' = (sd)^{O(\log d)}$ and $\deg(p_i) \leq \lceil d/2 \rceil$, for all $i \in [s']$.

Proof. Let C be a circuit of size s that computes p. Let us first assume that p is a homogeneous polynomial. We transform C by the following steps.

1. We apply *depth reduction* to C [VSBR83], and get a homogeneous circuit C' of depth $\log d$ and size $\text{poly}(s)$ that computes p.

2. Then we convert C' into a formula F by unfolding the gates with fan-out larger than one. By induction on the depth of the circuit, one can show that F has size $s^{O(\log d)}$.

3. Next, we convert F to an ABP B. It is well known that for any formula of size t, there exists an equivalent ABP of size at most $t + 1$, for details see [Sau12, Lemma 2.14]. Thus, the ABP B that computes p has size at most $s^{O(\log d)}$.

4. Finally, we *homogenize* B to a *layered* ABP B' as explained at the end of the preliminary section. Its size is $|B'| = \text{poly}(s^{O(\log d)}) = s^{O(\log d)}$.

To obtain the representation (13) of p, we cut ABP B' in half. That is, we split B' along the nodes in the $\lceil d/2 \rceil$-th layer. The i-th node v_i in that layer (in some order) defines two ABPs, one between the starting node of B' and v_i as end node, and a second one between v_i as starting node and the end node of B'. Let $p_{i,1}$ and $p_{i,2}$ be the two polynomials computed by these ABPs, respectively. By the definition of how ABPs compute polynomials, we have

$$p = \sum_{i=1}^{|B'|} p_{i,1}\, p_{i,2}\,,$$

where the degree of $p_{i,1}, p_{i,2}$ is at most $\lceil d/2 \rceil$. Now each product can be written as a SOS by (4) as $p_{i,1}\, p_{i,2} = \frac{1}{4}\left((p_{i,1} + p_{i,2})^2 - (p_{i,1} - p_{i,2})^2\right)$ to obtain (13). Hence, we get a SOS-representation of p with top fanin $s' = 2|B'|$.

For a non-homogeneous polynomial p, it is known that the homogeneous parts can be computed by homogeneous circuits of size $O(sd^2)$. Thus, for non-homogeneous polynomials, we can replace the s from above by $O(sd^2)$. Then the top-fanin of the SOS-representation is $(sd^2)^{O(\log d)} = (sd)^{O(\log d)}$. □

Now we come to our main result. We show how to lift the hardness of univariate polynomial f_d of degree d in the SOS-model to a multivariate polynomial that has circuits of super-polynomial size; this lifted polynomial will be in VNP and not in VP.

Our technique is to convert f_d into a multivariate polynomial $P_{k,n}$ via the multilinear Kronecker substitution defined in the preliminary section. Polynomial $P_{k,n}$ will have kn variables and degree n, for carefully chosen parameters k and n that depend on d and the SOS-hardness parameter ε for f_d. Since n is a function in k, it would actually suffice to index the family over k. We will eventually show that $\text{size}(P_{k,n}) = (kn)^{\omega(1)}$.

The proof goes via contradiction. If the size is smaller than claimed, then, by Lemma 3.1, we can write $P_{k,n}$ as the sum of $d^{o(\varepsilon)}$-many Q_i^2's, where the polynomials Q_i have kn variables and degree at most $n/2$. Thus, the support-sum of $P_{k,n}$, and hence of f_d as well, is bounded by $d^{o(\varepsilon)}\binom{kn+n/2}{n/2}$. We show that, for carefully chosen parameters, the latter expression is bounded by $o(d^{1/2+\varepsilon})$. Hence, we get a contradiction to the SOS-hardness of f_d.

Theorem 3.2. *If there exists an SOS-hard explicit family* (f_d) *with hardness* $\varepsilon = \varepsilon(d) = \omega\left(\sqrt{\frac{\log\log d}{\log d}}\right)$, *then* VP $\neq$ VNP.

Proof. Let $f_d(x)$ be an explicit SOS-hard polynomial with hardness ε as in the theorem statement. We define parameters k and n as follows. Choose $k \geq 7$ large enough such that

$$(k-1)^{\varepsilon} \geq 6\,. \tag{14}$$

That is, define $k = \lceil 6^{\frac{1}{\varepsilon}} + 1 \rceil$. Then choose n such that

$$(k-1)^n \leq d < k^n\,.$$

Note that $n = \Theta(\varepsilon \log d) = O(\log d)$.

Now we apply the multilinear Kronecker map $\phi_{k,n}^{\text{lin}}$ from (11) to f_d and define polynomial

$$P_{k,n}(\mathbf{y}) = \phi_{k,n}^{\text{lin}}(f_d(x))\,.$$

Recall that $P_{k,n}$ is multilinear of degree n and has kn variables $y_{j,\ell}$, where $1 \leq j \leq n$ and $0 \leq \ell \leq k-1$. We show that $P_{k,n} \in$ VNP and $\notin$ VP, thereby separating the classes.

Part 1: $P_{k,n} \in$ VNP**.** Let

$$P_{k,n} = \sum_{\boldsymbol{e} \in \{0,1\}^{kn}} c_n(\boldsymbol{e})\, \mathbf{y}^{\boldsymbol{e}}\,.$$

By the inverse multilinear Kronecker map $\psi_{k,n}^{\text{lin}}$ from (12), we get an exponent e such that $x^e = \psi_{k,n}^{\text{lin}}(\mathbf{y}^{\boldsymbol{e}})$. Note that coefficient $c_n(\boldsymbol{e})$ in $P_{k,n}$ is the coefficient of x^e in f_d. We can compute e in time $\text{poly}(n,k)$ and each bit of $c_n(\boldsymbol{e})$ in time $\text{poly}(\log d) = \text{poly}(n \log k)$, by the explicitness of f_d. Hence, $P_{k,n} \in$ VNP by Theorem 2.2.

Part 2: $P_{k,n} \notin \mathsf{VP}$. Define

$$\mu = \frac{1}{\sqrt{\log d \log\log d}}\,.$$

We will show that $\mathrm{size}(P_{k,n}) \geq d^{\mu}$.

Assume to the contrary that $\mathrm{size}(P_{k,n}) \leq d^{\mu}$. By Lemma 3.1, there exist polynomials Q_i such that $P_{k,n} = \sum_{i=1}^{s} c_i\, Q_i^2$, where $s = (d^{\mu} n)^{O(\log n)}$ and $\deg(Q_i) \leq \lceil n/2 \rceil$.

We apply the inverse multilinear Kronecker map $\psi_{k,n}^{\mathrm{lin}}$ to the Q_i's: Define $g_i(x) = \psi_{k,n}^{\mathrm{lin}}(Q_i(\mathbf{y}))$. Note that the Q_i's might no longer be multilinear. Anyway we can apply the ψ^{lin}-transformation. Then we get

$$f_d = \sum_{i=1}^{s} c_i\, g_i^2\,.$$

Recall that sparsity of g_i can be at most that of Q_i. For the sparsity of Q_i, we use the general bound (8). That is, $\mathrm{sp}(Q_i) \leq \binom{kn+\lceil n/2 \rceil}{\lceil n/2 \rceil}$, for all $i \in [s]$. Thus,

$$S(f_d) \leq s \binom{kn + \lceil n/2 \rceil}{\lceil n/2 \rceil}\,. \tag{15}$$

In the following two claims, we give upper bounds for s and the binomial coefficient in (15). Let

$$\delta = \sqrt{\frac{\log\log d}{\log d}}\,.$$

Note that $\delta = \mu \log\log d = o(\varepsilon)$.

Claim 3.3. $s = d^{O(\delta)} = d^{o(\varepsilon)}$.

Proof. Recall that $s = (d^{\mu} n)^{O(\log n)}$. We show that $(d^{\mu} n)^{O(\log n)} = d^{O(\delta)}$. Taking logarithms, we have to show that

$$\log n\,(\mu \log d + \log n) = O(\delta) \log d\,.$$

Recall that $n = O(\log d)$. Hence, we have $\log n = O(\log\log d)$. Now it suffices to show that

$$\mu \log\log d + \frac{(\log\log d)^2}{\log d} = O(\delta)\,.$$

But this holds because by the definitions of μ and δ and some elementary calculation, we have

$$\frac{(\log\log d)^2}{\log d} < \mu \log\log d = \delta\,.$$

This proves the claim. □

Claim 3.4. $\binom{kn+\lceil n/2 \rceil}{\lceil n/2 \rceil} \leq d^{\frac{1+\varepsilon}{2}}$.

Proof. We use (7) to bound the binomial coefficient. We omit the ceiling brackets for better readability.

$$\binom{kn + n/2}{n/2} \;\leq\; \left(\frac{e(kn + \frac{n}{2})}{\frac{n}{2}}\right)^{\frac{n}{2}} \;=\; (2ek + e)^{\frac{n}{2}} \;\leq\; (6(k-1))^{\frac{n}{2}}\,. \tag{16}$$

The last inequality is because $2e < 6$ and $2ek + e \leq 6(k-1)$, by our choice of $k \geq 7$.

By (14), we get that $6(k-1) \leq (k-1)^{1+\varepsilon}$. Hence, we can continue (16) by

$$(6(k-1))^{\frac{n}{2}} \;\leq\; (k-1)^{\frac{n}{2}(1+\varepsilon)} \;\leq\; d^{\frac{1+\varepsilon}{2}}.$$

The last inequality follows by our choice of n such that $(k-1)^n \leq d$. This proves the claim. □

We plug in the bounds from the two claims in (15) and get

$$S(f_d) = d^{o(\varepsilon)}\, d^{\frac{1+\varepsilon}{2}} = o(d^{1/2+\varepsilon}).$$

This is a contradiction to the SOS-hardness of f_d. So $\mathrm{size}(P_{k,n}) \geq d^{\mu}$.

It remains to show that d^{μ} is super-polynomial in parameters k and n.

Claim 3.5. $d^{\mu} = (kn)^{\omega(1)}$.

Proof. Taking logarithms, we have to show that

$$\mu \log d = \omega(\log k + \log n). \tag{17}$$

For the left hand side of (17), we have

$$\mu \log d = \sqrt{\frac{\log d}{\log\log d}} = \omega(1/\varepsilon).$$

For the right hand side of (17), we have

$$\log k = \log\lceil 6^{1/\varepsilon} + 1\rceil = O(1/\varepsilon),$$
$$\log n \leq \log\log d = o(1/\varepsilon).$$

This proves (17), and the claim follows. □

We conclude that $P_{k,n}$ requires super-polynomial size circuits, and therefore, $P_{k,n} \notin \mathrm{VP}$. This proves the theorem. □

Remark. 1. We used the multilinear Kronecker substitution ϕ^{lin} because the standard one ϕ from (10) would *not* give our result. For d, k, n as above, polynomial $\phi_{k,n}(f_d)$ would have only k variables but higher degree, kn, compared to $P_{k,n}$ from above. Then the binomial coefficient in (15) would become $\binom{k+kn/2}{k} > (n+1)^k > d$. Hence, Claim 3.4 would no longer hold.

2. Recall from the proof that $\deg(Q_i) \leq n/2$. Hence, for $g_i(x) = \psi_{k,n}(Q_i(\mathbf{y}))$, by the definition of ψ, we have

$$\deg(g_i) \leq \frac{n}{2}(k-1)\,k^{n-1} < n\,k^n = O(nd) = O(d\log d).$$

Thus, in the SOS-hardness assumption for f_d we could additionally restrict the degree of the polynomials in the SOS-representation to $O(d\log d)$, and still Theorem 3.2 would hold.

3. Similarly, by Claim 3.3, we could additionally restrict the top fanin s in the SOS-representation to $s = d^{\delta}$ and still Theorem 3.2 would hold. Note that this is very small compared to d since $d^{\delta} = d^{o(\varepsilon)}$.

Separating VBP and VNP. Recall that VBP $\subseteq$ VP. If we are interested in the weaker separation of VBP and VNP, then actually a smaller hardness parameter ε suffices in the assumption. The reason comes from Lemma 3.1: When we start with a polynomial p given by an ABP of size s, we can skip transformation steps 1, 2, 3, and only do the homogenization step 4. Then the resulting ABP has size only $\mathrm{poly}(s)$, i.e., we do not have the $\log d$-term in the exponent. So we can modify the proof of Theorem 3.2 and set $\varepsilon = \omega(1/\sqrt{\log d})$ and $\mu = \delta = 1/\sqrt{\log d}$, and still all the calculations go through, in particular Claim 3.3.

Corollary 3.6 (Determinant vs Permanent)**.** *If there exists an SOS-hard explicit family* (f_d) *with hardness parameter* $\varepsilon = \omega(1/\sqrt{\log d})$*, then* VBP $\neq$ VNP*.*

3.2 An exponential separation of VP and VNP

The argument for an exponential separation of VP and VNP follows the proof of Theorem 3.2. However, we use a different decomposition lemma and a different parameter setting. The decomposition lemma is based on the circuit depth-reduction technique of Valiant et al. [VSBR83]. Saptharishi [Sap19] has written a very nice survey on *frontier decomposition* (for e.g., see lemma 5.12 and Theorem 5.15 in the same) the technique to prove the following lemma.

Lemma 3.7 (Sum of product-of-2)**.** *Let* $p \in \mathbb{F}[\mathbf{x}]$ *be an* n*-variate homogeneous polynomial of degree* d*, computed by a homogeneous circuit of size* s*. Then there exist polynomials* $p_{i,j} \in \mathbb{F}[\mathbf{x}]$ *such that*

$$p \;=\; \sum_{i=1}^{s} p_{i,1}\, p_{i,2}\,,$$

and for all $i \in [s]$ *and* $j = 1, 2$*,*

1. $\frac{d}{3} \leq \deg(p_{i,j}) \;\leq\; \frac{2d}{3}$*,*
2. $\deg(p_{i,1}) + \deg(p_{i,2}) = d$*, and*
3. $p_{i,j}$ *has a homogeneous circuit of size* $O(s)$*.*

Remark. For a non-homogeneous polynomial $p(\mathbf{x})$, we can apply Lemma 3.7 for each homogeneous part of $p(\mathbf{x})$. It is well known that each homogeneous part can be computed by a homogeneous circuit of size $O(sd^2)$. Thus, for non-homogeneous polynomials, s can be replaced by $O(sd^2)$ and we get a similar conclusion.

The following lemma iterates the decomposition in Lemma 3.7 to bring the degree of the intermediate polynomials close to $d/2$, while keeping the circuit size polynomial in s. Note that, this is in contrast to Lemma 3.1 where we got intermediate polynomials of degree precisely $d/2$ but paid a super-polynomial blowup in the top fanin.

Lemma 3.8. *Let* $\frac{1}{2} < \gamma < 1$ *be a constant. Then there exists a constant* c*, such that for any* n*-variate homogeneous polynomial* $p \in \mathbb{F}[\mathbf{x}]$ *of degree* d *that can be computed by a homogeneous circuit of size* s*, we have a representation*

$$p \;=\; \sum_{i=1}^{s^c} q_i^2\,, \tag{18}$$

where $q_i \in \mathbb{F}[\mathbf{x}]$, *for all* $i \in [s^c]$, *such that*

1. $\deg(q_i) < \gamma d$,
2. $\text{size}(q_i) = O(s)$.

Proof. By Lemma 3.7, we can write $p(\mathbf{x}) = \sum_{i=1}^{s} \tilde{p}_{i,1}\, \tilde{p}_{i,2}$, where

- $\deg(\tilde{p}_{i,j}) \leq 2d/3$ and $\deg(\tilde{p}_{i,1}) + \deg(\tilde{p}_{i,2}) = d$,
- $\tilde{p}_{i,j}$ has a homogeneous circuit of size $O(s)$.

Let $\delta = \gamma - 1/2$. Choose constant m such that $(2/3)^m < \delta$. That is, let $m = \lceil \log_{3/2}(1/\delta) \rceil$. Now we apply Lemma 3.7 recursively m-times to each $\tilde{p}_{i,j}$. It follows that we can write $p(\mathbf{x})$ as

$$p(\mathbf{x}) \;=\; \sum_{i=1}^{s^{2^m-1}} \widehat{p}_{i,1}\, \widehat{p}_{i,2} \cdots \widehat{p}_{i,2^m}\,, \tag{19}$$

where $\deg(\widehat{p}_{i,j}) \leq (2/3)^m\, d < \delta d$. For all $i \in [s^{2^m-1}]$, we have $\sum_{j=1}^{2^m} \deg(\widehat{p}_{i,j}) = d$ and $\text{size}(\widehat{p}_{i,j}) = O(s)$, for all $j \in [2^m]$.

For each product $\widehat{p}_{i,1} \cdots \widehat{p}_{i,2^m}$, pick the smallest $j_0 \in [2^m]$ such that

$$\frac{d}{2} \;\leq\; \sum_{j=1}^{j_0} \deg(\widehat{p}_{i,j}) \;<\; \gamma d\,.$$

Define $p_{i,1} = \widehat{p}_{i,1} \cdots \widehat{p}_{i,j_0}$ and $p_{i,2} = \widehat{p}_{i,j_0+1} \cdots \widehat{p}_{i,2^m}$. Then we have

$$p \;=\; \sum_{i=1}^{s^{2^m-1}} p_{i,1}\, p_{i,2}\,.$$

By definition, $d/2 \leq \deg(p_{i,1}) < \gamma d$, and therefore, $\deg(p_{i,2}) = d - \deg(p_{i,1}) \leq d/2 < \gamma d$. Because each $\widehat{p}_{i,j}$ has a homogeneous circuit of size $O(s)$, so does $p_{i,j}$. Finally, we use equality (4) as $p_{i,1}\, p_{i,2} = \frac{1}{4}\left((p_{i,1} + p_{i,2})^2 - (p_{i,1} - p_{i,2})^2\right)$ to obtain (18) with $c = 2^m$. □

Remark. Similar as remarked for Lemma 3.7, for a non-homogeneous polynomial $p(\mathbf{x})$, the size s can be replaced by $O(sd^2)$ and we get a similar conclusion.

Lemma 3.8 provides the tool for an exponential separation of VP and VNP. The argument follows the proof of Theorem 3.2. Instead of Lemma 3.1, we use Lemma 3.8. Also we use a different parameter setting.

Theorem 3.9 (Constant ε)**.** *If there exists an SOS-hard explicit family with constant hardness parameter* $\varepsilon > 0$*, then* VNP *is exponentially harder than* VP*, i.e., there exists a polynomial family* $(g_n)_n \in \text{VNP}$ *such that* $\text{size}(g_n) = 2^{\Omega(n)}$.

Proof. Let $f_d(x)$ be an explicit SOS-hard polynomial with constant hardness parameter $\varepsilon < \frac{1}{2}$. First, we define parameters k and n. Let

$$\gamma = \frac{1}{2} + \frac{\varepsilon}{4}\,.$$

Choose constant $k \geq 7$ large enough such that

$$(k-1)^{\frac{\varepsilon}{12}} \geq 6^{\gamma}, \tag{20}$$

and n again such that $(k-1)^n \leq d < k^n$. Note that $n = O(\frac{\log d}{\log k}) = O(\log d)$. Then define $P_{k,n}(\mathbf{y}) = \phi^{\text{lin}}_{k,n}(f_d(x))$. Recall that $P_{k,n}$ is a homogeneous polynomial of degree n with kn variables. Again, we have $P_{k,n} \in \mathsf{VNP}$.

We show that $P_{k,n}$ requires exponential size circuits. We apply Lemma 3.8 to $P_{k,n}$ with parameter γ. Let c be the constant such that s^c bounds the top fanin in (18). That is, when $\text{size}(P_{k,n}) = s$, we get a representation

$$P_{k,n} \;=\; \sum_{i=1}^{s^c} c_i Q_i^2,$$

where $\deg(Q_i) \leq \gamma n$. Define constant

$$\mu \;=\; \frac{\varepsilon}{3c}.$$

Claim 3.10. $\text{size}(P_{k,n}) > d^{\mu}$.

Proof. Assume that $\text{size}(P_{k,n}) \leq d^{\mu}$. Via the inverse Kronecker map applied to the Q_i's, we get a bound similar to (15):

$$S(f_d) \leq d^{c\mu} \binom{kn + \lceil \gamma n \rceil}{\lceil \gamma n \rceil}. \tag{21}$$

We bound the binomial coefficient similar as in Claim 3.4:

$$\binom{kn+\gamma n}{\gamma n} \;\leq\; \left(\frac{e(kn+\gamma n)}{\gamma n}\right)^{\gamma n} \;\leq\; (2ek+e)^{\gamma n} \;\leq\; (6(k-1))^{\gamma n}.$$

By (20) and the definition of γ, we get that

$$(6(k-1))^{\gamma n} \;\leq\; (k-1)^{n(\frac{\varepsilon}{12}+\gamma)} \;\leq\; d^{\frac{1}{2}+\frac{\varepsilon}{3}}.$$

Plugging the bound into (21), we get by definition of μ

$$S(f_d) \;\leq\; d^{c\mu}\, d^{\frac{1}{2}+\frac{\varepsilon}{3}} \;=\; d^{\frac{1}{2}+\frac{2}{3}\varepsilon} \;=\; o(d^{\frac{1}{2}+\varepsilon}).$$

This is a contradiction to the SOS-hardness of f_d. This proves the claim. □

Finally, observe that by the definition of n, and since μ and k are constants, we have $d^{\mu} \geq (k-1)^{n\mu} = 2^{\Omega(kn)}$. Hence, $P_{k,n}$ requires exponential size circuits. This shows an exponential separation between VP and VNP. □

4 Sum of Cubes

In this section, let $\mathbb{F}$ be a field of characteristic $\neq 2, 3$. The following lemma is the crucial ingredient to connect general circuits to a SOC-representation. It is similar to Lemma 3.8. There, we represented a polynomial p as a sum of squares of polynomials with degree close to $1/2$. Now, we write p as a sum of cubes of polynomials with degree close to $1/3$.

Lemma 4.1 (SOC decomposition). *There exists a constant c, such that for any n-variate homogeneous polynomial $p \in \mathbb{F}[\mathbf{x}]$ of degree d that can be computed by a homogeneous circuit of size s, we have a representation*

$$p \;=\; \sum_{i=1}^{s^c} q_i^3\,,$$

where $q_i \in \mathbb{F}[\mathbf{x}]$, for all $i \in [s^c]$, such that

1. $\deg(q_i) < \frac{4}{11}\,d$,
2. q_i *has a circuit of size* $O(s)$.

Proof. We start exactly as in the proof of Lemma 3.8, with parameters $\gamma = 4/11$ and $\delta = \gamma - 1/3 = 1/33$. Then we choose m such that $(2/3)^m < \delta$. Hence, we can set $m = 9$ and we can write p as in (19):

$$p \;=\; \sum_{i=1}^{s^{2^m-1}} \widehat{p}_{i,1}\,\widehat{p}_{i,2}\cdots\widehat{p}_{i,2^m}\,,$$

where $\deg(\widehat{p}_{i,j}) \leq (2/3)^m\,d < \delta d$. For all $i \in [s^{2^m-1}]$, we have $\sum_{j=1}^{2^m}\deg(\widehat{p}_{i,j}) = d$ and $\mathrm{size}(\widehat{p}_{i,j}) = O(s)$, for all $j \in [2^m]$.

In Lemma 3.8, we split each product $\widehat{p}_{i,1}\cdots\widehat{p}_{i,2^m}$ into two parts of degree close to $d/2$. Now, we similarly split it into three parts of degree close to $1/3$. So we first pick the smallest $j_0 \in [2^m]$ such that

$$\frac{d}{3} \;\leq\; \sum_{j=1}^{j_0}\deg(\widehat{p}_{i,j}) \;<\; \gamma d\,,$$

and define $p_{i,1} = \widehat{p}_{i,1}\cdots\widehat{p}_{i,j_0}$. Then we pick the smallest j_1, where $j_0 < j_1 \leq 2^m$, such that

$$\frac{d}{3} \;\leq\; \sum_{j=j_0+1}^{j_1}\deg(\widehat{p}_{i,j}) \;<\; \gamma d\,,$$

and define $p_{i,2} = \widehat{p}_{i,j_0+1}\cdots\widehat{p}_{i,j_1}$ and $p_{i,3} = \widehat{p}_{i,j_1+1}\cdots\widehat{p}_{i,2^m}$. Then we have

$$p \;=\; \sum_{i=1}^{s^{2^m-1}} p_{i,1}\,p_{i,2}\,p_{i,3}\,, \tag{22}$$

where $d/3 \leq \deg(p_{i,j}) < \gamma d$, for all $i \in [s^c]$ and $j = 1, 2$. For, $p_{i,3}$, note that $\deg(p_{i,3}) = d - \deg(p_{i,1}) - \deg(p_{i,2}) \leq d/3 < \gamma d$.

Finally, we write the products in (22) as sums of cubes by the following identity:

$$24abc \;=\; (a+b+c)^3 \;-\; (a-b+c)^3 \;-\; (a+b-c)^3 \;+\; (a-b-c)^3\,. \tag{23}$$

□

Remark. In case of non-homogeneous polynomials, we can consider the homogeneous parts separately. The size s has then again to by replaced by $O(sd^2)$.

We now come to the main result of this section, that the existence of a SOC-hard family implies the derandomization of blackbox-PIT. The proof outline is roughly similar to the proof of Theorem 3.2, but with some crucial modifications.

Given a SOC-hard polynomial $f_d(x)$, we apply the standard Kronecker map to construct a polynomial $P_{k,n}$ that is k-variate, for some constant k, and the variables have individual degree n. We show that $\text{size}(P_{k,n}) = n^{\Omega(1)}$. Note that despite the lower bound, we have $P_{k,n} \in \mathsf{VP}$, because $P_{k,n}$ is constant-variate. However, it turns out that constant-variate hardness is good enough for constructing an efficient HSG.

The proof of the size lower bound goes again by contradiction, and this is where Lemma 4.1 comes into the play. Via the SOC-decomposition of $P_{k,n}$ and the inverse Kronecker map, we get a SOC-representation of f_d that would be smaller than the assumed SOC-hardness of f_d. Thus $P_{k,n}$ fulfills the assumptions made in Theorem 2.4 by Guo et al. [GKSS19], and hence, we can conclude that blackbox-PIT $\in \mathsf{P}$.

We also reiterate that using our techniques and analysis, SOS-decomposition *does not* yield a polynomial-time blackbox PIT. This is mainly because conversions from univariate f_d to a k-variate polynomial $P_{k,n}$ would naively give an upper-bound on (both) the measures

$$\binom{k+kn/2}{k} > (n+1)^k > d\,.$$

Here we use $kn/2$, as the degree of $P_{k,n}$ is kn, whereas the degree of the intermediate polynomial halves. Similarly, for SOC-decomposition, we have to be slightly more restrictive and only work with the support-union, instead of support-sum, as it still becomes big in our analysis.

Theorem 4.2. *If there is an SOC-hard family, then (i) blackbox-PIT* $\in \mathsf{P}$ *and (ii)* VNP *is exponentially harder than* VP.

Proof. Let $f_d(x)$ be an explicit SOC-hard polynomial such that $U(f_d, d^\varepsilon) \geq \delta d$, for constants $0 < \varepsilon < 1/2$ and $\delta > 0$. Let furthermore c be the constant from Lemma 4.1.

We define parameters k and n as follows. Let $\alpha = 1 - \frac{1}{110}$. Choose k large enough such that

$$k > \frac{9c}{\varepsilon} \quad \text{and} \quad \alpha^k \;<\; \delta, \tag{24}$$

and define n such that $n^k \leq d < (n+1)^k$. Now we apply the Kronecker map $\phi_{k,n}$ from (10) to f_d and define polynomial

$$P_{k,n}(y) = \phi_{k,n}(f_d(x))\,.$$

Recall that $P_{k,n}$ has k variables of individual degree n, and therefore total degree kn. Since f_d is explicit, so is $P_{k,n}$.

Define μ as

$$\mu = \frac{1}{2}\left(\frac{\varepsilon}{c} - \frac{1}{k}\right). \tag{25}$$

Note that $\mu > 0$ by our choice of k in (24).

Claim 4.3 (Hardness of $P_{k,n}$)**.** $\mathrm{size}(P_{k,n}) > d^{\mu}$, *for large enough* n.

Proof. Assume to the contrary that $\mathrm{size}(P_{k,n}) \leq d^{\mu}$. By Lemma 4.1, there exist polynomials Q_i such that

$$P_{k,n} = \sum_{i=1}^{s_0} c_i \, Q_i^3,$$

where $s_0 \leq (d^{\mu}\, kn)^c$ and $\deg(Q_i) \leq \frac{4}{11}\, kn$.

We apply the inverse Kronecker map $\psi_{k,n}$ to the Q_i's: Define $g_i(x) = \psi_{k,n}(Q_i(\mathbf{y}))$. Then we get

$$f_d = \sum_{i=1}^{s_0} c_i \, g_i^3 \,.$$

Recall that g_i and Q_i have the same sparsity. Therefore

$$s_1 = \left|\bigcup_i \mathrm{supp}(g_i)\right| \;\leq\; \left|\bigcup_i \mathrm{supp}(Q_i)\right| \;\leq\; \binom{k + \frac{4}{11}kn}{k}.$$

Thus, $U(f_d, s_0) \leq s_1$.

We want to show that $s_0 < d^{\varepsilon'}$ and $s_1 < \delta d$, for large enough n. Then, we have $U(f_d, d^{\varepsilon}) < \delta d$, for large enough d, which contradicts the SOC-hardness of f_d.

Bound on s_0. Recall that $n^k \leq d$. Therefore we get

$$s_0 \;\leq\; (d^{\mu}\, kn)^c \;\leq\; (d^{\mu}\, k d^{\frac{1}{k}})^c \;=\; (k d^{\mu + \frac{1}{k}})^c \;<\; d^{\varepsilon}\,.$$

In the last inequality we used that $\mu + 1/k < \varepsilon/c$, by the definition of μ in (25), and further, since it is a strict inequality and both k and c are constants, the inequality holds for large enough d.

Bound on s_1. By (7), we have

$$s_1 \;=\; \binom{k + \frac{4}{11}kn}{k} \;\leq\; \left(e\left(1 + \frac{4}{11}n\right)\right)^k \;<\; (\alpha n)^k \;<\; \alpha^k d \;<\; \delta d\,.$$

Note that by our choice of the constants, we have $e\left(1 + \frac{4}{11}n\right) < \alpha n$, for large enough n. Also, we used that $n^k < d$ and $\alpha^k < \delta$ by (24). This proves Claim 4.3. □

(i) Blackbox-PIT $\in$ P. We show that from the hardness of $P_{k,n}$, we can fulfill the assumption in Theorem 2.4, that $\text{size}(P_{k,n}) > s^{10k+2}\deg(P_{k,n})^3$, for some growing function $s = s(n)$. Recall that $\deg(P_{k,n}) \leq kn$. We define, $s(n) = n^{\frac{1}{10k+3}}$. Then we have

$$s^{10k+2}(kn)^3 = n^{\frac{10k+2}{10k+3}}(kn)^3 = k^3 n^{4-\frac{1}{10k+3}} < n^4, \tag{26}$$

for large enough n. By the first condition in our choice of k in (24), we have

$$\mu = \frac{1}{2}\left(\frac{\varepsilon}{c} - \frac{1}{k}\right) \geq \frac{1}{2}\left(\frac{9}{k} - \frac{1}{k}\right) = \frac{4}{k},$$

and therefore $k\mu \geq 4$. Recall also that $n^k \leq d$. Hence, we can continue (26) as

$$n^4 \leq n^{k\mu} \leq d^{\mu} < \text{size}(P_{k,n}). \tag{27}$$

Equations (26) and (27) give the desired hardness of $P_{k,n}$. Thus, Theorem 2.4 gives a $\text{poly}(s)$-time HSG for $C(s,s,s)$. Hence, blackbox-PIT $\in$ P.

(ii) Exponential separation of VP and VNP. This part of the proof follows standard arguments in the literature, see for example [Bür13, Koi11, KKPS15].

We extend the Kronecker substitution to map the k-variate $P_{k,n}$ to a polynomial with even more variables, simply by applying the standard Kronecker substitution to each of the k variables $\mathbf{y} = (y_1, y_2, \ldots, y_k)$ of $P_{k,n}$. Define $m = \lceil \log(n+1) \rceil$. For each y_i, we introduce m variables $\mathbf{x}_i = (x_{i,j})_{j\in[m]}$. Let $\widehat{\phi}_{k,m}$ be the map

$$\widehat{\phi}_{k,m} : y_1^{e_1} \cdots y_k^{e_k} \mapsto \prod_{i=1}^{k} \mathbf{x}_i^{\text{base}_2(e_i)}.$$

By linear extension, define $R_{k,m} = \widehat{\phi}_{k,m}(P_{k,n})$. Hence, $R_{k,m}$ is a multilinear polynomial with $km = O(\log d)$ variables. Note that $\widehat{\phi}_{k,m}$ maps each monomial of $P_{k,n}$ to a distinct multilinear monomial in $R_{k,m}$. The inverse map is also obvious: $\widehat{\psi}_{k,m} : x_{i,j} \mapsto y_i^{2^{j-1}}$. Since $P_{k,n}$ is explicit, so is $R_{k,m}$, implying that it is in VNP.

Observe that

$$P_{k,n}(y_1, \ldots, y_k) = R_{k,m}(y_1^{2^0}, y_1^{2^1}, \ldots, y_1^{2^{m-1}}, \ \cdots \ , y_k^{2^0}, y_k^{2^1}, \ldots, y_k^{2^{m-1}}).$$

Since we can compute all the powers of the k variables in circuit size $2km$, we have

$$\text{size}(P_{k,n}) \leq \text{size}(R_{k,m}) + 2km.$$

Since $\text{size}(P_{k,n}) > d^{\mu}$ by Claim 4.3, we get $\text{size}(R_{k,m}) > d^{\mu} - 2km = 2^{\Omega(km)}$ □

Remark. The degree of the Q_i's in the above proof is bounded by $\frac{4}{11}kn$. Hence, the degree of the g_i's obtained via the inverses Kronecker substitution is bounded by

$$(n+1)^{k-1}\frac{4}{11}kn < \frac{4}{11}k(n+1)^k \leq O(d),$$

where in the last equality, we used that $(n+1)^k \leq (2n)^k < 2^k d$, and k is a constant. Thus, it suffices to study the representation of f_d as sum-of-cubes g_i^3, where $\deg(g_i) = O(d)$, and still Theorem 4.2 would hold. From the SOC-hard family in Theorem 4.2 we also get VP $\neq$ VNP.

5 Approximative SOS-hardness and SOC-hardness

In this section, we study the SOS-hardness, respectively, the SOC-hardness, in the *border* or *approximative sense*. Eventually, we show similar consequences as of Theorem 3.2 and Theorem 4.2 in the border algebraic complexity setup.

In a slight abuse of notation, for an arithmetic C with input variables $\mathbf{x}$, the output polynomial computed by $C(\mathbf{x})$ we denote by $C(\mathbf{x})$ as well.

Definition 5.1 (Approximative computation)**.** *A circuit* C *over* $\mathbb{F}(\epsilon)[\mathbf{x}]$ *is said to* approximate a polynomial $P(\mathbf{x})$*, if for some* $M \geq 0$*,*

$$\lim_{\epsilon \to 0} \frac{1}{\epsilon^M} C(\mathbf{x}, \epsilon) \;=\; P(\mathbf{x})\,. \tag{28}$$

The approximative circuit complexity of P*, denote by* $\overline{\mathrm{size}}(P)$*, is the size of the smallest circuit that approximates* P*. The class* $\overline{\mathsf{VP}}$ *contains all families of* n*-variate polynomials of degree* $\mathrm{poly}(n)$ *over* $\mathbb{F}$ *of approximative complexity* $\mathrm{poly}(n)$*.*

Equation (28) can be interpreted as P being approximated by the circuit $C(\mathbf{x}, \epsilon)/\epsilon^M$ over the function field $\mathbb{F}(\epsilon)$. An equivalent way to express the approximation in (28) is that the polynomial computed by circuit C can be written as $C(\mathbf{x}, \epsilon) = \epsilon^M P(\mathbf{x}) + \epsilon^{M+1} Q(\mathbf{x}, \epsilon)$, for some polynomial $Q(\mathbf{x}, \epsilon) \in \mathbb{F}[\mathbf{x}, \epsilon]$.

Note that $\mathsf{VP} \subseteq \overline{\mathsf{VP}}$ because VP is the special case in Definition 5.1 where we fix $M = 0$. However, the definition does not bound M at all. Thus, $\overline{\mathsf{VP}}$ could potentially be much larger than VP. Bürgisser [Bür01] gave a bound on M. He showed that any polynomial in $\overline{\mathsf{VP}}$ can be approximated with $M \leq 2^{\mathrm{poly}(n)}$. It is still an open question whether VP is different from $\overline{\mathsf{VP}}$.

5.1 Approximative SOS-hardness and $\mathsf{VNP} \not\subseteq \overline{\mathsf{VP}}$

Mulmuley and Sohoni [MS01, MS08] proposed the *Geometric Complexity Theory* (GCT) program, which is an approach to the VP vs. VNP problem, via representation theory and algebraic geometry. Eventually, it strengthens Valiant's conjecture and focuses on separating $\overline{\mathsf{VP}}$ (or $\overline{\mathsf{VBP}}$) from VNP. We show that proving a slightly non-trivial lower bound in the SOS-model, in the border sense, is enough to separate these classes.

Definition 5.2 (Approximative SOS and border-support-sum size $\overline{S}_R(f)$)**.** *Let* R *be a ring. An* n*-variate polynomial* $f(\mathbf{x}) \in R[\mathbf{x}]$ *is* approximated as a (weighted) SOS*, if there exists an integer* $M \geq 0$ *such that*

$$f(\mathbf{x}) = \lim_{\epsilon \to 0} \frac{1}{\epsilon^M} \sum_{i=1}^{s} c_i f_i^2(\mathbf{x}, \epsilon)\,, \tag{29}$$

for some top-fanin s*, where* $f_i \in R[\mathbf{x}, \epsilon]$ *and* $c_i \in R[\epsilon]$*.*

The size in the representation of f *in (29) is the* border support-sum*, the sum of the support size (or sparsity) of the polynomials* f_i *over* $R[\epsilon]$*. The* border-support-sum size *of* f*, is defined as the minimum border-support-sum of* f*, denoted by* $\overline{S}_R(f)$*, or simply* $\overline{S}(f)$*, when the ring* R *is clear from the context.*

Note that, by definition, $\overline{S}_R(f) \leq S_R(f)$. In particular, when f is univariate and has sparsity, $sp(f) = d+1$, over any field $R = \mathbb{F}$, of characteristic $\neq 2$, similar bounds hold:

$$\sqrt{d} \;\leq\; \overline{S}(f) \;\leq\; S(f) \;\leq\; 2d+2\,.$$

We call a polynomial family approximative SOS-hard, if its border-support-sum size is just *slightly* larger than the trivial lower bound.

Definition 5.3 (Approximative SOS-hardness). *A polynomial family* $(f_d(x))_d$ *is* approximative SOS-hard *with hardness* ε, *if* $\overline{S}(f_d) = \Omega(d^{0.5+\varepsilon})$.

We point out that the SOS decomposition lemma (Lemma 3.1) works for approximative circuits as well. This lemma plays the pivotal role to establish a connection between approximative SOS-hardness and the general circuit hardness, in the border sense.

Lemma 5.4 (Border SOS Decomposition). *Let* $f(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ *be a polynomial of degree* d *that can be approximated by a circuit* C *of size* s. *Then there exist polynomials* $f_i \in \mathbb{F}[\mathbf{x}, \epsilon]$ *and* $c_i \in \mathbb{F}[\epsilon]$ *such that*

$$C(\mathbf{x}, \epsilon) \;=\; \sum_{i=1}^{s'} c_i\, f_i^2\,,$$

where $s' = (sd)^{O(\log d)}$ *and* $\deg_x(f_i) \leq \lceil d/2 \rceil$, *for all* $i \in [s']$.

Proof. We adapt the proof of Lemma 3.1. Simply consider $C \in \mathbb{F}(\epsilon)[\mathbf{x}]$ and observe that the earlier proof is independent of the underlying field. However, there is one subtlety that we have to take care of: Let C be the circuit of size s that approximates f. That is, for some $M \geq 0$, we have

$$C(\mathbf{x}, \epsilon) = \epsilon^M f + \epsilon^{M+1} g(\mathbf{x}, \epsilon)\,.$$

Now it could happen that $\deg_x(g) > d$. However, using the homogenization technique, we can extract all the terms upto degree d in $\mathbf{x}$, which does not effect the f-part. In particular, there is a circuit $\widehat{C} \in \mathbb{F}(\epsilon)[\mathbf{x}]$ of size $O(sd^2)$, such that $\widehat{C}(\mathbf{x}, \epsilon) = \epsilon^M f + \epsilon^{M+1}\, \widehat{g}(\mathbf{x}, \epsilon)$, where $\deg_x(\widehat{g}) \leq d$. Now, we can work with the circuit $\widehat{C}$ instead C and the proof of Lemma 3.1 goes through over $\mathbb{F}(\epsilon)$. □

We come to our main result in this section. We lift the approximative hardness of a univariate polynomial of degree d in the SOS-model to a multivariate polynomial that has approximative circuits of super-polynomial size, implying it is not in $\overline{\mathsf{VP}}$, but its explicitness ensures it to be in VNP.

Theorem 5.5. *If there exists an approximative SOS-hard explicit family* (f_d) *with hardness parameter* $\varepsilon = \varepsilon(d) = \omega\left(\sqrt{\frac{\log\log d}{\log d}}\right)$, *then* $\mathsf{VNP} \not\subseteq \overline{\mathsf{VP}}$.

Proof. The proof is similar to the proof of Theorem 3.2. We define $P_{k,n}$ with the similar parameters as in that proof. As f_d is explicit, so is $P_{k,n}$. Therefore $P_{k,n} \in \mathsf{VNP}$.

To show that $P_{k,n} \notin \overline{\mathsf{VP}}$, we define μ similarly. We will argue that

$$\overline{\mathsf{size}}(P_{k,n}) \geq d^{1/\mu} = (kn)^{\omega(1)}\,. \tag{30}$$

It follows that $P_{k,n} \notin \overline{\mathsf{VP}}$.

To show (30), assume to the contrary that $\overline{\mathrm{size}}(P_{k,n}) \leq d^{\mu}$. Then there is a circuit $C(\mathbf{y}, \epsilon) \in \mathbb{F}(\epsilon)[\mathbf{x}]$ of size d^{μ} and a $M \geq 0$, such that $C(\mathbf{y}, \epsilon) = \epsilon^M P_{k,n} + \epsilon^{M+1} Q(\mathbf{y}, \epsilon)$. By Lemma 5.4, there exist polynomials $Q_i(\mathbf{y}, \epsilon)$ such that $C(\mathbf{y}, \epsilon) = \sum_{i=1}^{s} c_i Q_i(\mathbf{y}, \epsilon)^2$, where $s = (d^{\mu} n)^{O(\log n)}$ and $\deg_{\mathbf{y}}(Q_i) \leq \lceil n/2 \rceil$.

If we apply the inverse multilinear Kronecker map $\psi_{k,n}^{\mathrm{lin}}$ to the Q_i's, we get

$$\epsilon^M f_d + \epsilon^{M+1} \psi_{k,n}^{\mathrm{lin}}(Q) = \sum_{i=1}^{s} c_i g_i^2 ,$$

where $g_i(x) = \psi_{k,n}^{\mathrm{lin}}(Q_i(\mathbf{y}))$. Note that, $\mathrm{sp}(g_i) \leq \mathrm{sp}(Q_i)$ over $\mathbb{F}(\epsilon)$. For the sparsity of Q_i, we use the general bound (8). That is, $\mathrm{sp}(Q_i) \leq \binom{kn+\lceil n/2 \rceil}{\lceil n/2 \rceil}$, for all $i \in [s]$. Thus, by definition, $\overline{S}(f_d) \leq s\binom{kn+\lceil n/2 \rceil}{\lceil n/2 \rceil}$. The same calculation as in the proof of Theorem 3.2 shows that $\overline{S}(f_d) = o(d^{1/2+\varepsilon})$, a contradiction. □

With a slight modification in the parameters ε, μ and δ, we get a similar consequence as Corollary 3.6 .

Corollary 5.6. *If there exists an approximative SOS-hard explicit family* (f_d) *with hardness parameter* $\varepsilon = \omega(1/\sqrt{\log d})$*, then* $\mathsf{VNP} \not\subseteq \overline{\mathsf{VBP}}$.

When ε is constant, we get an exponential separation between $\overline{\mathsf{VP}}$ and VNP, similar to Theorem 3.9. The basic tool is a border-decomposition version of Lemma 3.8. We omit the proofs as they are similar.

Theorem 5.7 (Constant ε in the border)**.** *If there exists an approximative SOS-hard explicit polynomial family with constant hardness parameter* $\varepsilon > 0$*, then* VNP *is exponentially harder than* $\overline{\mathsf{VP}}$.

5.2 Approximative SOC-hardness and efficient hitting set for $\overline{\mathsf{VP}}$

In this subsection, we introduce approximative SOC-hardness and show its intrinsic connection to construct efficient hitting sets for $\overline{\mathsf{VP}}$. Though the existence of a poly-size hitting set is known due to [HS80a], the best complexity bound known for constructing a hitting set for $\overline{\mathsf{VP}}$ is PSPACE [FS18, GSS19]. The main difficulty comes from certifying that the set that has been constructed is indeed a hitting set. Kumar, Saptharishi, and Solomon [KSS19] showed that the hardness of constant-variate polynomials in the approximative sense suffices to construct an HSG for $\overline{\mathsf{VP}}$.

Theorem 5.8. [KSS19, Thm.1.6] *Let* P *be a* k*-variate polynomial in* $\mathbb{F}[\mathbf{x}]$ *of degree* d *such that* $\mathrm{coef}(P)$ *can be computed in time* $\mathrm{poly}(d)$*. Suppose* $\overline{\mathrm{size}}(P) > s^{10k+2}\, d$*, for some parameter* s*. Then there is a* $\mathrm{poly}(s)$*-size hitting set for* $\overline{C(s, s, s)}$.

Next, we define the approximative SOC-model and its complexity measure.

Definition 5.9 (Approximative SOC and border-support-union size $\overline{U}_R(f,s)$)**.** *Let* R *be a ring. An* n*-variate polynomial* $f(\mathbf{x}) \in R[\mathbf{x}]$ *is* approximated as a SOC*, if there exists an integer* $M \geq 0$ *such that*

$$f(\mathbf{x}) = \lim_{\epsilon \to 0} \frac{1}{\epsilon^M} \sum_{i=1}^{s} c_i f_i^3(\mathbf{x}, \epsilon)\,, \tag{31}$$

for some top-fanin s*, where* $f_i \in R[\mathbf{x}, \epsilon]$ *and* $c_i \in R[\epsilon]$*.*

The size of the representation of f *in* (31) *is the size of the support-union over* $R[\epsilon]$*, i.e.* $\left|\bigcup_{i=1}^{s} \mathrm{supp}(f_i)\right|$*, where* $\mathrm{supp}(f_i)$ *denotes the set of monomials with a nonzero coefficient in* f_i*. The* border support-union size *of* f *with respect to* s*, denoted* $\overline{U}_R(f,s)$*, is defined as the minimum border support-union size when* f *is written as in (31).*

By definition, we have $\overline{U}_R(f,s) \leq U_R(f,s)$. In particular, when f is univariate and has sparsity $\mathrm{sp}(f) = d+1$, over any field $R = \mathbb{F}$, of characteristic $\neq 2, 3$, equation (6) extends to

$$d^{1/3} \;\leq\; \overline{U}(f,s) \;\leq\; U(f,s) \;\leq\; d+1\,.$$

The lower bound is again by a counting argument.

Thus, it follows that for s large enough, $\overline{U}(f,s)$ is small. However, it is unclear whether this is true when $s = o(\sqrt{d})$. We call a polynomial family approximative SOS-hard, if its border support-union size attains the trivial upper bound.

Definition 5.10 (Approximative SOC-hardness)**.** *A polynomial family* $(f_d(x))_d$ *is* approximative SOC-hard*, if there is a constant* $0 < \varepsilon < 1/2$ *such that* $\overline{U}_{\mathbb{F}}(f_d, d^\varepsilon) = \Omega(d)$*.*

One can show that an explicit approximative SOC-hard univariate family can be converted to an explicit hitting set for $\overline{\mathsf{VP}}$. The main ingredient is a SOC-decomposition in the approximative sense. This decomposition is very similar to Lemma 4.1, except that the working field is $\mathbb{F}(\epsilon)$.

Lemma 5.11 (Approximative SOC decomposition)**.** *There exists a constant* c*, such that for any* n*-variate polynomial* $p \in \mathbb{F}[\mathbf{x}]$ *of degree* d *that can be approximated by a circuit of size* s*, we have a representation*

$$\epsilon^M p + \epsilon^{M+1} q(\mathbf{x}, \epsilon) \;=\; \sum_{i=1}^{(sd)^c} q_i^3\,,$$

where $q_i \in \mathbb{F}[\epsilon][\mathbf{x}]$*, for all* $i \in [(sd)^c]$*, such that*

1. $\deg(q_i) < \frac{4}{11}\, d$*,*
2. q_i *has a circuit of size* $\mathrm{poly}(s,d)$ *over* $\mathbb{F}(\epsilon)$*.*

Using the above lemma and Theorem 5.8, it is not hard to construct an explicit and efficient hitting set for $\overline{\mathsf{VP}}$. The proof goes along the lines of Theorem 4.2.

Theorem 5.12. *If there is an approximative SOC-hard family, then we have a* $\mathrm{poly}(s)$*-explicit hitting set for* $\overline{\mathsf{VP}}$*.*

6 Sum of powers of small support-union

In this section, let $\mathbb{F}$ be a field of characteristic 0 or large. We give a way to represent any univariate polynomial as sum of r-th powers of polynomials. Here we use the notion of sumsets. In additive combinatorics, the *sumset*, also called the *Minkowski sum* of two subsets A and B of an abelian group G, is defined to be the set of all sums of an element from A with an element from B,

$$A + B \;=\; \{\, a + b \mid a \in A,\; b \in B \,\}.$$

The n-*fold iterated sumset* of A is $nA = A + \;\cdots\; + A$, where there are n summands.

We want a *small support-union* representation of a polynomial f of degree d as a sum of r-th powers, where r is constant.

Let t be the unique non-negative integer such that $(t-1)^r < d+1 \le t^r$. Define set B as

$$B \;=\; \{\, a\, t^{\ell} \mid 0 \le a \le t-1 \text{ and } 0 \le \ell \le r-1 \,\}.$$

Hence $|B| = rt = O(d^{1/r})$. Let $k \in \{0, 1, \ldots, d\}$. The base-$t$ representation of k is a sum of at most r elements from B. Hence, $\{0, 1, \ldots, d\} \subseteq rB$. The largest element in B is $m = (t-1)t^{r-1} = O(d)$. Since r is a constant, the largest element in rB is $rm = O(d)$.

We show next that any polynomial can be written as a sum of r-th powers of polynomials with support in B.

Theorem 6.1. *For any* $f \in \mathbb{F}[x]$ *of degree* d*, there exist* $\ell_i \in \mathbb{F}[x]$ *with* $\mathrm{supp}(\ell_i) \subseteq B$ *and* $c_i \in \mathbb{F}$*, for* $i = 0, 1, \ldots, mr$*, such that* $f \;=\; \sum_{i=0}^{mr} c_i\, \ell_i^r$.

Proof. Let us set up the polynomials ℓ_i we seek as

$$\ell_i(x) = \sum_{j \in B} a_{i,j} x^j,$$

for unknown coefficients $a_{i,j} \in \mathbb{F}$, for $i = 0, 1, \ldots, rm$ and $j \in B$. We determine the $a_{i,j}$'s via the multivariate polynomial

$$L_i(z_i, x) = \sum_{j \in B} z_{i,j} x^j,$$

where we replaced the coefficients $a_{i,j}$ of ℓ_i by distinct indeterminates $z_{i,j}$.

Note that $\deg_x(L_i) \le m$. Taking the r-th power, we can write

$$L_i^r \;=\; \sum_{j=0}^{mr} Q_j(z_i)\, x^j,$$

for $0 \le i \le rm$, for polynomials Q_j of degree r with $|B| = rt$ many variables, $0 \le j \le rm$.

Let $S = \{\, j \mid Q_j \neq 0 \,\} \subseteq \{0, 1, \ldots, rm\}$. Note that from any monomial in Q_j we can recover j. This follows because $\mathrm{supp}(Q_{j_1}) \cap \mathrm{supp}(Q_{j_2}) = \emptyset$, for any $j_1 \neq j_2$ in S. Therefore, the polynomials $\{\, Q_j \mid j \in S \,\}$ are $\mathbb{F}$-linearly independent.

Note that by the definition of B, we have $\{0, 1, \ldots, d\} \subseteq S$.

We want to find $\mathbf{c} = \begin{pmatrix} c_1 & c_2 & \cdots & c_{|S|} \end{pmatrix} \in \mathbb{F}^{|S|}$ and $\mathbf{a}_i = (a_{i,j})_j$ such that

$$f(x) \;=\; \sum_{i=0}^{mr} c_i \, \ell_i^r(x) = \sum_{i=0}^{mr} c_i \, L_i^r(\mathbf{a}_i, x)\,. \tag{32}$$

Let $f(x) = \sum_{i=0}^{d} f_i \, x^i$. We set up a linear system to determine the unknowns. Define the coefficient vector $\mathbf{f}$ of f over S and a $|S| \times |S|$-matrix A as

$$\mathbf{f} = \begin{pmatrix} f_0 & f_1 & \cdots & f_d & 0 & \cdots & 0 \end{pmatrix},$$

$$A = \begin{pmatrix} Q_{j_1}(z_1) & Q_{j_2}(z_1) & \cdots & Q_{j_s}(z_1) \\ Q_{j_1}(z_2) & Q_{j_2}(z_2) & \cdots & Q_{j_s}(z_2) \\ \vdots & \vdots & \cdots & \vdots \\ Q_{j_1}(z_{|S|}) & Q_{j_2}(z_{|S|}) & \cdots & Q_{j_s}(z_{|S|}) \end{pmatrix}.$$

Then (32) is equivalent to

$$\mathbf{c}\, A(\mathbf{a}) = \mathbf{f}\,.$$

As the z_i's are distinct variables, the first column of A consists of different variables at each coordinate. Moreover, the first row of A contains $\mathbb{F}$-linearly independent Q_j's. Thus, for a random $\mathbf{a} = (a_{i,j})$, matrix $A(\mathbf{a})$ has full rank over $\mathbb{F}$. Fix such an $\mathbf{a}$. This yields $\mathbf{c} = \mathbf{f}\,(A(\mathbf{a}))^{-1}$. For these values $\mathbf{c}$ and $\mathbf{a}$, we get (32) as desired. □

Remark.

1. The above calculation does *not* give small support-sum representation of f, as the top-fanin is already $\Omega(d)$.

2. The above representation crucially requires a *field* $\mathbb{F}$. E.g. it does not exist for f_d over the ring $\mathbb{Z}$.

The number of *distinct* monomials across the ℓ_j's in the above proof is $|B| = O(d^{1/r})$, while the top-fanin is $\leq mr + 1 = \Theta(d)$. Of particular interest for us are the cases $r = 2, 3$.

Corollary 6.2. *Any polynomial $f \in \mathbb{F}[x]$ of degree d has a SOS- and a SOC-representation with top-fanin $O(d)$ and support-union $O(\sqrt{d})$, respectively $O(\sqrt[3]{d})$.*

In the following, we improve Theorem 6.1 for $r = 2, 3$. We show a SOS- and SOC-representation for any polynomial $f(x)$, wherein both the top-fanin *and* the support-union size are small, namely $O(\sqrt{d})$. We assume that characteristic of $\mathbb{F}$ is $\neq 2$ in case of SOS, and $\neq 3$, in case of SOC. The representations are based on discussions with Agrawal [Agr20].

6.1 Small SOS

By Corollary 6.2, any polynomial f of degree d has a SOS-representation with top-fanin $O(d)$ and support-union $O(\sqrt{d})$ We show that also the top-fanin can be reduced to $O(\sqrt{d})$. The technical key for this is the following lemma. It shows how to decrease the top-fanin in a representation without increasing the support-union.

Lemma 6.3. *Let* $f \in \mathbb{F}[x]$ *be written as* $f = \sum_{i=1}^{s} c_i f_{i,1} f_{i,2}$, *with support-union* $t = |\bigcup_{i,j} \mathrm{supp}(f_{i,j})|$. *Then there exists a representation* $f = \sum_{i=1}^{t} c'_i f'_{i,1} f'_{i,2}$ *with support-union* $\leq t$.

Proof. For the given representation of f, we assume w.l.o.g. that $\deg(f_{i,1}) \geq \deg(f_{i,2})$ and that $f_{i,1}, f_{i,2}$ are monic, for $i = 1, 2, \ldots, s$. Let $S = \bigcup_{i,j} \mathrm{supp}(f_{i,j})$.

We construct the representation claimed in the lemma by ensuring the following properties:

1. For every $x^e \in S$, there is exactly one i such that $\deg(f'_{i,1}) = e$,
2. $\bigcup_{i,j} \mathrm{supp}(f'_{i,j}) \subseteq S$,

Since we also maintain that $\deg(f'_{i,1}) \geq \deg(f'_{i,2})$, it follows that the top-fanin is indeed bounded by $t = |S|$ as claimed.

We handle the monomials in S successively according to decreasing degree. Let $x^e \in S$ be the monomial with the largest e that occurs more than once as the degree of a $f_{i,1}$, say $\deg(f_{1,1}) = \deg(f_{2,1}) = e$.

Define $g_1 = f_{2,1} - f_{1,1}$. Then we have $f_{2,1} = f_{1,1} + g_1$ and $\deg(g_1) < e$. Moreover, the support of g_1 is contained in the support of $f_{1,1}$ and $f_{2,1}$ If $\deg(f_{2,2}) = e$, then we define similarly $g_2 = f_{2,2} - f_{1,1}$. Then $f_{2,2} = f_{1,1} + g_2$ and $\deg(g_2) < e$. Now we can write

$$\begin{aligned} c_1 f_{1,1} f_{1,2} + c_2 f_{2,1} f_{2,2} &= c_1 f_{1,1} f_{1,2} + c_2 (f_{1,1} + g_1)(f_{1,1} + g_2) \\ &= f_{1,1} \left(c_1 f_{1,2} + c_2 f_{1,1} + c_2 g_1 + c_2 g_2\right) + c_2 g_1 g_2 \end{aligned}$$

The second line is a new sum of two products, where only the first product has terms of degree e, whereas in the second product, g_1, g_2 have smaller degree. Also, the support-union set has not increased.

In case when $\deg(f_{2,2}) < e$, we can just work with $f_{2,2}$ directly instead of $f_{1,1} + g_2$, and the above equations gets even simpler. □

So when we start with the SOS-representation for polynomial f provided by Theorem 6.1 and apply Lemma 6.3, It follows that f can be re-written as $f(x) = \sum_{i=1}^{O(\sqrt{d})} c'_i f_{i,1} f_{i,2}$, where $|\bigcup_{i,j} \mathrm{supp}(f_{ij})| = O(\sqrt{d})$. This can be turned into a SOS-representation by $f_{i,1} f_{i,2} = (f_{i,1} + f_{i,2})^2/4 - (f_{i,1} - f_{i,2})^2/4$. Note that the last step does not change the support-union, and at most doubles the top-fanin. Hence, we get

Theorem 6.4 (Small SOS-Representation)**.** *Any polynomial* $f \in \mathbb{F}[x]$ *of degree* d *has a SOS-representation such that the top-fanin and the support-union are bounded by* $O(\sqrt{d})$.

6.2 Small SOC

We show two small SOC-representation with different parameters. First, we show a $\sqrt{d}$ SOC-representation that follows essentially from Theorem 6.4. We use the following lemma that a given representation of a polynomial as a sum of m-powers can be rewritten as a sum of r-powers, for any $r \geq m$. In particular, for $m = 2$ and $r = 3$, we see how to rewrite a SOS-representation as a SOC-representation.

Lemma 6.5. *Let $\mathbb{F}$ be a field of characteristic 0 or large. Let $h(x) \in \mathbb{F}[x]$ and $0 \le m \le r$. There exist $c_{m,i} \in \mathbb{F}$ and* distinct *$\lambda_i \in \mathbb{F}$, for $0 \le i \le r$, such that*

$$h(x)^m \;=\; \sum_{i=0}^{r} c_{m,i}\,(h(x)+\lambda_i)^r\,. \tag{33}$$

Proof. Consider the polynomial $(h(x)+t)^r$, where t is a new indeterminate different from x. We have

$$(h(x)+t)^r \;=\; \sum_{i=0}^{r} \binom{r}{i}\, h(x)^i\, t^{r-i}\,.$$

Choose $r+1$ many distinct λ_i's and put $t=\lambda_i$, for $i=0,1,\ldots,r$. We get $r+1$ many linear equations which can be represented in matrix form $A\boldsymbol{\nu}=\mathbf{b}$, for matrix $A=\left(\binom{r}{j}\lambda_i^{r-j}\right)_{0\le i,j\le r}$, and vectors $\boldsymbol{\nu}=\left(h^i\right)_{0\le i\le r}$ and $\mathbf{b}=((h+\lambda_i)^r)_{0\le i\le r}$.

Note that except for the binomial factors, A is a Vandermonde matrix. When computing the determinant, one can pull out the binomial factor $\binom{r}{j}$ from the j-th column, for $j=0,1,\ldots,r$. Then a Vandermonde matrix remains, and hence

$$\det(A)=\prod_{j=0}^{r}\binom{r}{j}\prod_{0\le i<j\le r}(\lambda_j-\lambda_i)\neq 0\,.$$

Therefore, A is invertible and we have $\boldsymbol{\nu}=A^{-1}\mathbf{b}$. Let $\mathbf{c}_m$ be the $(m+1)$-th row of A^{-1}. Then we have $h(x)^m = \mathbf{c}_m\mathbf{b}$ which is exactly (33). □

Observe that the support on both sides of (33) is the same, except maybe for an extra constant term on the right hand side. Hence, for any given polynomial f, we can take the SOS-representation from Theorem 6.4 and rewrite each square as a sum of four cubes by Lemma 6.5. Then we get

Corollary 6.6 ($\sqrt{d}$ SOC-representation)**.** *Any polynomial $f \in \mathbb{F}[x]$ of degree d has a SOC-representation such that the top-fanin and the support-union are bounded by $O(\sqrt{d})$.*

Remark. Recall Definition 1.4 that f_d is SOC-hard if $U_{\mathbb{F}}(f_d, d^\varepsilon) = \Omega(d)$, for some $0<\varepsilon<1/2$. Corollary 6.6 shows, that SOC-hardness is not possible for $\varepsilon = 1/2$.

The second way to get a small SOC-representation technically follows the way we got Theorem 6.4. We first show a reduction similar to Lemma 6.3 for the sum of product-of-3.

Lemma 6.7. *Let $f \in \mathbb{F}[x]$ be written as If $f=\sum_{i=1}^{s} c_i\, f_{i,1}f_{i,2}f_{i,3}$ with support-union t, then there exists a representation $f=\sum_{i=1}^{t^2} c_i'\, f_{i,1}'f_{i,2}'f_{i,3}'$ with support-union $\le t$.*

Proof. The argument is similar to the proof of Lemma 6.3. For the given representation of f, we assume that $\deg(f_{i,1}) \ge \deg(f_{i,2}) \ge \deg(f_{i,3})$ and that $f_{i,1}, f_{i,2}, f_{i,3}$ are monic, for $i=1,2,\ldots,s$. Let $S=\bigcup_{i,j}\operatorname{supp}(f_{i,j})$.

Let $x^e \in S$ be the monomial with the largest e that occurs more than once as the degree of a $f_{i,1}$. W.l.o.g. assume $\deg(f_{1,1})=e$. Write all the other $f_{i,j}$'s where x^e occurs as

$$f_{i,j} = f_{1,1} + g_{i,j}, \tag{34}$$

for $j \in [s]$ and $k \in [3]$. Note that $\deg(g_{i,j}) < e$.

Now we plug in (34) in the representation of f given by assumption and multiply out. This gives

$$f = \sum_{i \in [m]} c_i \, f_{i,1} f_{i,2} f_{i,3} = f_{1,1} \, P + R,$$

where P is a sum of product-of-2 and R is a sum of product-of-3, where each intermediate polynomial has degree $< e$. Note that the last expression still has the same support-union.

Apply Lemma 6.3 on P, to reduce its top-fanin to t. Observe that then $f_{1,1}P$ has a sum of product-of-3 expression with top fanin at most t. Iterating the procedure to R, we finally get a representation of f with top fanin bounded by t^2. □

By Corollary 6.2, any polynomial f of degree d has a SOC-representation with top-fanin $O(d)$ and support-union $O(\sqrt[3]{d})$. By Lemma 6.7, this can be re-written as a sum product-of-3 with top-fanin $O(d^{2/3})$. Finally, any product-of-3 can be written as a sum of four cubes, by (23). Hence, we get

Theorem 6.8 ($d^{2/3}$ SOC-representation)**.** *Any polynomial $f \in \mathbb{F}[x]$ of degree d has a SOC-representation with top-fanin $O(d^{2/3})$ and support-union $O(d^{1/3})$.*

Finally, we observe that Lemma 6.5 also provides a connection between the two complexity measures $S(f)$ from SOS and $U(f, s)$ from SOC.

Lemma 6.9. *For any $f \in \mathbb{F}[x]$, we have $S(f) \geq \min_s (U(f, 4s) - 1)$.*

Proof. Suppose $f = \sum_{i=1}^{s} c_i \, f_i^2$. By Lemma 6.5, each f_i^2 can be written as $f_i^2 = \sum_{j=1}^{4} c_{ij} \, (f_i + \lambda_{ij})^3$, for distinct $\lambda_{ij} \in \mathbb{F}$. Thus, $U(f, 4s) \leq 1 + \sum_{i=1}^{s} sp(f_i)$. Taking minimum over s gives the desired inequality. □

Corollary 6.10. *For $s = \Omega(d^{2/3})$, we have $U(f, s) = \Theta(d^{1/3})$.*

7 Conclusion

This work established that studying the univariate sum-of-squares representation (resp. cubes) is fruitful. Proving a *vanishingly* better lower bound than the trivial one, suffices to both derandomize and prove hardness in algebraic complexity.

Here are some immediate questions which require rigorous investigation.

1. Does existence of a SOS-hard family solve PIT completely? The current proof technique fails to reduce from cubes to squares.

2. Prove existence of a SOS-hard family for the *sum of constantly* many squares.

3. Prove existence of a SOC-hard family for a 'generic' polynomial f with rational coefficients ($\mathbb{Q}$). Does it fail when we move to *complex* coefficients ($\mathbb{C}$)?

4. Can we optimize ε in the SOS-hardness condition (& Corollary 3.6)? In particular, does proving an SOS lower-bound of $\sqrt{d} \cdot \text{poly}(\log d)$, suffice to deduce a separation between determinant and permanent (similarly VP and VNP)?

Acknowledgments. P.D. is supported by the project "Foundation of Lattice-based Cryptography", funded by NUS-NCS Joint Laboratory for Cyber Security, Singapore. This work was mostly carried when P.D. was a research scholar at CMI, and a visiting scholar at CSE, IIT Kanpur funded by Google PhD Fellowship (2018-2022). N.S. thanks the funding support from DST (SJF/MSA-01/2013-14), DST-SERB (CRG/2020/000045), and N.Rama.Rao Chair. Thanks to Manindra Agrawal for many useful discussions to optimize the SOS representations; to J. Maurice Rojas for several comments; to Arkadev Chattopadhyay for organizing a TIFR Seminar on this work. T.T. thanks DFG for the funding (grants TH 472/5-1 and TH 472/5-2), and CSE, IIT Kanpur for the hospitality.

References

[Agr20] Manindra Agrawal. Private Communication, 2020. 29

[AGS19] Manindra Agrawal, Sumanta Ghosh, and Nitin Saxena. Bootstrapping variables in algebraic circuits. *Proceedings of the National Academy of Sciences*, 116(17):8107–8118, 2019. Earlier in Symposium on Theory of Computing, 2018 (STOC'18). 3, 8

[AV08] Manindra Agrawal and V Vinay. Arithmetic Circuits: A Chasm at Depth Four. In *Foundations of Computer Science, 2008. FOCS'08. IEEE 49th Annual IEEE Symposium on*, pages 67–75. IEEE, 2008. 2, 8

[BCS13] Peter Bürgisser, Michael Clausen, and Amin Shokrollahi. *Algebraic Complexity Theory*, volume 315. Springer Science & Business Media, 2013. 2, 7, 10

[BM16] Boaz Barak and Ankur Moitra. Noisy tensor completion via the Sum-of-squares Hierarchy. In *Conference on Learning Theory*, pages 417–445, 2016. 2

[Bür01] Peter Bürgisser. The complexity of factors of multivariate polynomials. In *In Proc. 42th IEEE Symp. on Foundations of Comp. Science*, 2001. 24

[Bür04] Peter Bürgisser. The complexity of factors of multivariate polynomials. *Foundations of Computational Mathematics*, 4(4):369–396, 2004. (Preliminary version in FOCS 2001). 7

[Bür13] Peter Bürgisser. *Completeness and Reduction in Algebraic Complexity Theory*, volume 7. Springer Science & Business Media, 2013. 3, 11, 23

[CKW11] Xi Chen, Neeraj Kayal, and Avi Wigderson. *Partial derivatives in arithmetic complexity and beyond*. Now Publishers Inc, 2011. 2

[DL78] Richard A. Demillo and Richard J. Lipton. A probabilistic remark on algebraic program testing. *Information Processing Letters*, 7(4):193 – 195, 1978. 3

[Dut21] Pranjal Dutta. Real τ-Conjecture for Sum-of-Squares: A Unified Approach to Lower Bound and Derandomization. In *International Computer Science Symposium in Russia*, pages 78–101. Springer, 2021. 7

[Dut22] Pranjal Dutta. *A tale of hardness, de-randomization and de-bordering in complexity theory*. PhD thesis, Chennai Mathematical Institute, 2022. 7

[FS18] Michael A Forbes and Amir Shpilka. A pspace construction of a hitting set for the closure of small algebraic circuits. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, pages 1180–1192, 2018. 26

[GKKS13] Ankit Gupta, Pritish Kamath, Neeraj Kayal, and Ramprasad Saptharishi. Arithmetic circuits: A chasm at depth three. In *2013 IEEE 54th Annual Symposium on Foundations of Computer Science*, pages 578–587. IEEE, 2013. 2, 8

[GKSS19] Zeyu Guo, Mrinal Kumar, Ramprasad Saptharishi, and Noam Solomon. Derandomization from Algebraic Hardness: Treading the Borders. In *60th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2019*, pages 147–157, 2019. Online version: https://mrinalkr. bitbucket. io/papers/newprg. pdf. 3, 8, 12, 21

[GMK17] Ignacio Garcia-Marco and Pascal Koiran. Lower bounds by Birkhoff interpolation. *Journal of Complexity*, 39:38–50, 2017. 8

[GMQ16] Joshua A. Grochow, Ketan D. Mulmuley, and Youming Qiao. Boundaries of VP and VNP. In *43rd International Colloquium on Automata, Languages, and Programming (ICALP 2016)*, volume 55, pages 34:1–34:14, 2016. 7

[GSS19] Zeyu Guo, Nitin Saxena, and Amit Sinhababu. Algebraic dependencies and pspace algorithms in approximative complexity over any field. *Theory of Computing*, 15(1):1–30, 2019. 26

[HS80a] Joos Heintz and Claus-Peter Schnorr. Testing polynomials which are easy to compute. In *Proceedings of the twelfth annual ACM symposium on Theory of computing*, pages 262–272. ACM, 1980. 26

[HS80b] Joos Heintz and Malte Sieveking. Lower bounds for polynomials with algebraic coefficients. *Theoretical Computer Science*, 11(3):321–330, 1980. 3

[HWY11] Pavel Hrubeš, 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. 7

[IL17] Christian Ikenmeyer and JM Landsberg. On the complexity of the permanent in various computational models. *Journal of Pure and Applied Algebra*, 221(12):2911–2927, 2017. 12

[KI04] Valentine Kabanets and Russell Impagliazzo. Derandomizing polynomial identity tests means proving circuit lower bounds. *Computational Complexity*, 13(1-2):1–46, 2004. 3, 5, 8

[KKPS15] Neeraj Kayal, Pascal Koiran, Timothée Pecatte, and Chandan Saha. Lower bounds for sums of powers of low degree univariates. In *International Colloquium on Automata, Languages, and Programming*, pages 810–821. Springer, 2015. 8, 23

[Koi11] Pascal Koiran. Shallow circuits with high-powered inputs. In *Innovations in Computer Science - ICS 2010, Tsinghua University, Beijing, China, January 7-9, 2011. Proceedings*, pages 309–320, 2011. 7, 8, 23

[Koi12] Pascal Koiran. Arithmetic circuits: The chasm at depth four gets wider. *Theoretical Computer Science*, 448:56–65, 2012. 2, 8

[KP11] Pascal Koiran and Sylvain Perifel. Interpolation in Valiants theory. *Computational Complexity*, 20(1):1–20, 2011. 11

[KPGM18] Pascal Koiran, Timothée Pecatte, and Ignacio Garcia-Marco. On the linear independence of shifted powers. *Journal of Complexity*, 45:67–82, 2018. 8

[Kro82] Leopold Kronecker. Grundzüge einer arithmetischen Theorie der algebraischen Grössen.(Abdruck einer Festschrift zu Herrn EE Kummers Doctor-Jubiläum, 10. September 1881.). *Journal für die reine und angewandte Mathematik*, 92:1–122, 1882. 9

[KSS19] Mrinal Kumar, Ramprasad Saptharishi, and Noam Solomon. Derandomization from Algebraic Hardness: Treading the Borders. `https://arxiv.org/pdf/1905.00091v1.pdf`, 2019. 26

[KST19] Mrinal Kumar, Ramprasad Saptharishi, and Anamay Tengse. Near-optimal Bootstrapping of Hitting Sets for Algebraic Circuits. In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 639–646, 2019. 3

[Kum19] Mrinal Kumar. A quadratic lower bound for homogeneous algebraic branching programs. *computational complexity*, 28(3):409–435, 2019. 12

[Las07] Jean B Lasserre. A sum of squares approximation of nonnegative polynomials. *SIAM review*, 49(4):651–669, 2007. 2

[Lau09] Monique Laurent. Sums of squares, moment matrices and optimization over polynomials. In *Emerging applications of algebraic geometry*, pages 157–270. Springer, 2009. 2

[LL89] Thomas Lehmkuhl and Thomas Lickteig. On the order of approximation in approximative triadic decompositions of tensors. *Theoretical computer science*, 66(1):1–14, 1989. 7

[Mah14] Meena Mahajan. Algebraic Complexity Classes. In *Perspectives in Computational Complexity*, pages 51–75. Springer, 2014. 10

[MH02] John C Mason and David C Handscomb. *Chebyshev polynomials*. CRC press, 2002. 11

[MS01] Ketan D Mulmuley and Milind Sohoni. Geometric complexity theory I: An approach to the P vs. NP and related problems. *SIAM Journal on Computing*, 31(2):496–526, 2001. 7, 24

[MS08] Ketan D Mulmuley and Milind Sohoni. Geometric complexity theory II: Towards explicit obstructions for embeddings among class varieties. *SIAM Journal on Computing*, 38(3):1175–1206, 2008. 7, 24

[Mul17] Ketan Mulmuley. Geometric complexity theory V: Efficient algorithms for Noether normalization. *Journal of the American Mathematical Society*, 30(1):225–309, 2017. 7

[MV99] Meena Mahajan and V Vinay. Determinant: Old algorithms, new insights. *SIAM Journal on Discrete Mathematics*, 12(4):474–490, 1999. 12

[NW94] Noam Nisan and Avi Wigderson. Hardness vs randomness. *Journal of computer and System Sciences*, 49(2):149–167, 1994. 3

[Ore22] Øystein Ore. Über höhere kongruenzen. *Norsk Mat. Forenings Skrifter*, 1(7):15, 1922. 3

[Raz10] Ran Raz. Elusive Functions and Lower Bounds for Arithmetic Circuits. *Theory Comput.*, 6(1):135–177, 2010. 8

[Rez78] Bruce Reznick. Extremal PSD forms with few terms. *Duke mathematical journal*, 45(2):363–374, 1978. 2

[Sap19] Ramprasad Saptharishi. A survey of lower bounds in arithmetic circuit complexity. Github survey, 2019. 12, 17

[Sau12] Nitin Saurabh. ALGEBRAIC MODELS OF COMPUTATION. MS Thesis, 2012. 13

[Sch80] J. T. Schwartz. Fast Probabilistic Algorithms for Verification of Polynomial Identities. *J. ACM*, 27(4):701–717, October 1980. 3

[Sma98] Steve Smale. Mathematical problems for the next century. *The mathematical intelligencer*, 20(2):7–15, 1998. 7

[SS95] Michael Shub and Steve Smale. On the intractability of Hilberts Nullstellensatz and an algebraic version of NP$\neq$ P?. *Duke Mathematical Journal*, 81(1):47–54, 1995. 7

[Str74] Volker Strassen. Polynomials with rational coefficients which are hard to compute. *SIAM Journal on Computing*, 3(2):128–149, 1974. 2

[SY10] 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. 10, 12

[Tav15] Sébastien Tavenas. Improved bounds for reduction to depth 4 and depth 3. *Information and Computation*, 240:2–11, 2015. 2

[Val79] Leslie G Valiant. Completeness classes in algebra. In *Proceedings of the 11th Annual ACM symposium on Theory of computing*, pages 249–261. ACM, 1979. 10, 11

[VSBR83] Leslie G. Valiant, Sven Skyum, S. Berkowitz, and Charles Rackoff. Fast Parallel Computation of Polynomials Using Few Processors. *SIAM Journal of Computing*, 12(4):641–644, 1983. 13, 17

[Zip79] Richard Zippel. Probabilistic Algorithms for Sparse Polynomials. In *Proceedings of the International Symposium on Symbolic and Algebraic Computation*, EUROSAM '79, pages 216–226, 1979. 3