Lower-bounding the sum of 4^{th}-powers of univariates leads to derandomization and hardness

Pranjal Dutta * Nitin Saxena †

Abstract

We study the sum of fourth-powers (SO4) model to compute univariate polynomials over $\mathbb{F} = \mathbb{Q}$. We conjecture : the univariate polynomial $(x+1)^d$ when written as a sum of $o(d)$-many fourth-powers requires $\Omega(d)$ distinct monomials. We show that the conjecture puts blackbox-PIT (polynomial identity testing) in P; and proves $\mathsf{VP} \neq \mathsf{VNP}$ (under Generalized Riemann Hypothesis). Thus, studying very simple polynomials, in very simple computational models, suffices to solve the major questions of algebraic complexity theory.

Recently, (Dutta et al. , 2020), demonstrated such a connection for the sum of 25^{th}-powers model. Our work optimizes the exponent (from 25 to 4). We achieve this by employing a new 'CNF' *normal-form*, for general circuits, that is highly specialized towards SO4 expression. Basically, it 'reduces' the intermediate degrees to $1/3^{rd}$ before invoking the conjecture.

2012 ACM CCS concept: Theory of computation - Algebraic complexity theory, Problems, reductions and completeness, Pseudorandomness and derandomization; Computing methodologies - Algebraic algorithms; Mathematics of computing - Combinatoric problems.
Keywords: VP vs VNP, hitting set, circuit, CNF, normal form, univariate polynomial, 4th powers, PIT, lower bound, sparsity, monomials, support, CH, GRH.

1 Introduction

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. Some of the *natural* complexity parameters of a circuit are: **1)** the *size*, i.e. number of edges and nodes, **2)** the *depth*, i.e. number of layers, **3)** the *fanin*, i.e. maximum number of inputs to a node, (resp. the *fan-out*, i.e. maximum number of outputs of a node). In complexity classes, we specify only an upper bound on these parameters.

The class VP contains the families of n-variate polynomials of degree $\mathsf{poly}(n)$ over $\mathbb{F}$, computed by circuits of $\mathsf{poly}(n)$-size. The class VNP can be seen as a non-deterministic analog of the class VP. 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 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 *conjectured* to be different (Valiant's Hypothesis [Val79a]). For more details see, [Mah14, SY10, BCS13]. Unless specified particularly, we consider the field $\mathbb{F} = \mathbb{Q}$ (resp. a finite field with 'large' characteristic).

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]. These results demand a lower bound

*Chennai Mathematical Institute, India (& CSE, IIT Kanpur), pranjal@cmi.ac.in
†CSE, Indian Institute of Technology, Kanpur, nitin@cse.iitk.ac.in

of $n^{\omega(\sqrt{d})}$ on the top-fanin of an explicit n variate, d-degree polynomial when written as *sum of* $O(\sqrt{d})$*-th powers* of polynomials of degree at most $O(\sqrt{d})$. It seems that showing strong lower bounds *require* deeper understanding of algebraic-combinatorial structure of circuits. *Perhaps*, univariate polynomials could be easier to study due to the existing analytic tools in mathematics. Thus, the stimulus for the study of univariate lower bounds comes quite *naturally*. One should first try to develop techniques towards lower bounds for restricted univariate models.

In the world of univariate polynomials, the *Pochhammer-Wilkinson* polynomial, $P_d(x) := \prod_{i=1}^{d}(x-i)$, is conjectured to be hard, i.e. $\text{size}(P_d) \geq \Omega(d)$. It is known that such a hardness proof would imply $\mathsf{VP} \neq \mathsf{VNP}$, assuming GRH (General Riemann Hypothesis) [Bür09, Cor.4.2]. But, sufficiency of proving lower bound on *restricted* models of univariate polynomials came later when Koiran [Koi11] essentially showed that if there exists a univariate polynomial $f(x)$ of degree d such that any representation of the form $f(x) = \sum_{i=1}^{s} c_i \cdot Q_i^{e_i}$, where $\text{sparsity}(Q_i) \leq t$ and arbitrary e_i 's, requires top-fanin $s \geq (d/t)^{\Omega(1)}$, then $\mathsf{VP} \neq \mathsf{VNP}$.

Very recently, in [DST20, Thm.3], it was established that, in fact, showing a lower bound of $\Omega(d)$ on the number of monomials required to write $f_d := (x+1)^d$ as *sum of* 25^{th}*-powers of univariate* polynomials, *suffices* to separate VP from VNP (assuming GRH). In fact, the same hardness can be used to derandomize *blackbox* Polynomial Identity Testing (PIT) which asks for an algorithm to test the zeroness of a given algebraic circuit via mere query access. We *optimize* the recent connection to a surprising extent (see Theorems 2 and 3):

> *If $(x+1)^d$ written as sum of 4^{th}-powers of univariates requires sum of the sparsity of the univariates to be $\Omega(d^{0.98289})$, then assuming GRH, we have* $\mathsf{VP} \neq \mathsf{VNP}$.
> *Strengthening the requirement, to $\Omega(d)$ many* distinct *monomials, puts blackbox-PIT in P.*

In this work, we improve the result of [DST20] in three ways– exponent, measure and requirement-wise. To achieve this, we devise a new 'CNF' representation theorem (Theorem 1), which basically expresses a degree d polynomial of size s as a sum of $\text{poly}(s,d)$ many products of 4 polynomials, each of size $\text{poly}(s,d)$ and degree at most $d/3$. This improves the classic result; which is a sum of product of 5 polynomials, each of degree at most $d/2$ [VSBR83, Sap19]. Our refinement organically comes from the connection to sum of 4^{th}-powers (SO4) representation:

SO4 model and our measure. We say that a polynomial $f(x) \in R[x]$ over a ring R is computed as a *sum of* 4^{th}*-powers* (SO4) if

$$f = \sum_{i=1}^{s} c_i \cdot \ell_i^4 \,, \tag{1}$$

for some top-fanin s, where $c_i \in R$ and $\ell_i(x) \in R[x]$. Interestingly, the sum of 4^{th}-powers is a *complete model* for $R = \mathbb{F}$, a field of characteristic zero (resp. ≥ 5); for details see Lemma 14. A natural complexity measure in Eqn.(1) is the *sparsity-sum*, the sum of the sparsity of ℓ_i's. The sparsity-sum size of f, denoted by $S_{\mathbb{F}}(f)$, is defined as the minimum sparsity-sum size when f is written as in Eqn.(1).

A forthright counting argument shows that $S_R(f) \geq \Omega(|f|_1^{1/4})$ where $|f|_1$ denotes the number of nonzero monomials in f (i.e. *sparsity of* f). We want to inspect how $S_R(f_d)$ *behaves* w.r.t d for the 'simplest' polynomial family $f_d := (x+1)^d$. To the best of our knowledge, here are the estimates on $S_{\mathbb{F}}(f_d)$.

Upper bound on $S_{\mathbb{F}}(f_d)$. For $f_d(x) = (x+1)^d$, it is easy to see that if $4 \mid d$, then $S_R(f_d) \leq d/4+1$ as $(x+1)^d = ((x+1)^{d/4})^4$. In fact, Lemma 15 shows that $S_{\mathbb{F}}(f_d) \leq 5 \cdot (d/4+4)$ for *any* $d \in \mathbb{N}$.

$S_{\mathbb{F}}$ is large for *'almost all'* polynomials f. Wlog, in Eqn. (1), $\deg(\ell_i) \leq \text{poly}(d)$, as $\deg(f) = d$. We claim that the algebraic-circuit complexity $\text{size}(f) \leq O\left(S_{\mathbb{F}}(f) \cdot \log d\right)$. As, $\text{size}(f) \leq \left(\sum_{i\in[s]} \text{size}(\ell_i)\right)$ and by repeated-squaring, $\text{size}(\ell_i) \leq |\ell_i|_1 \cdot O(\log d)$. Moreover, for *random* f (of degree d), $\text{size}(f) \geq \Omega(d)$. Thus, one expects: $S_{\mathbb{F}}(f) \geq \Omega(d/\log d)$.

To play safe, we shall restrict d to the domain $I := \{2^m - 1 \mid m \in \mathbb{N}\}$. Let $\mathbb{F}$ be $\mathbb{Q}$, or a finite field of characteristic ≥ 5. Motivated from the upper bound on $S_{\mathbb{F}}(f_d)$ and the heuristic about the *largeness* of $S_{\mathbb{F}}$ for random f, we conjecture the following.

Conjecture 1 (C1). *For $d \in I$, $S_{\mathbb{F}}(f_d) \geq \Omega(d^{0.98289})$.*

Remarks.
1. We work with this particular domain I mainly because I *suffices* for later implications (Theorem 2). Additionally, it can be shown that the above conjecture is *true* over $\mathbb{Z}$. For $d \in I$, one can show that $\binom{d}{i} = 1 \bmod 2$ for all $0 \leq i \leq d$, implying $|(x+1)^d \bmod 2|_1 = d+1$, whereas $\sum c_i \cdot \ell_i(x)^4 \equiv \sum c_i \cdot \ell_i(x^4) \bmod 2$. Thus, $S_{\mathbb{Z}}(f_d) \geq d+1$.
2. $\Omega(d^{0.98289})$, instead of $\Omega(d)$ [DST20], is a *weaker* requirement; yet, we show that it is strong enough to separate VP and VNP (see Theorem 2).
3. We could even restrict the degrees of ℓ_i, to be $O(d \log d)$, in Eqn.(1), to prove the results in this paper (Remark 3, Section 3.2). This might help in proving the conjecture.
4. We believe the conjecture to hold for *any* $d \in \mathbb{N}$ (i.e. beyond I). We also believe the conjecture to be true for *most* polynomial families, e.g. $f := \sum_{i=0}^{d} 3^{i^2} x^i$ or $f := \prod_{i=1}^{d} (x - i)$.

1.1 Our Results: New circuit normal-form and implications of Conjecture C1

Algebraic circuits are quite well-structured, for instance, there is a famous depth-$O(\log d)$ reduction result [VSBR83, SY10, Sap19]. This is made possible by discovering a specialized *normal-form decomposition* for a circuit. Then, one recurses to reduce the degree with depth.

Connecting algebraic circuits to SO4 representation is a major challenge as the prior best decomposition (with polynomial blow-up in size) required multiplication fanin 5 (*hinting* a sum of 5^{th}-powers, in a way). For details, see [VSBR83, Sap19]. It was established that an n-variate, degree d polynomial $f(\boldsymbol{x})$, computed by a circuit of size s, can be decomposed as

$$f(\boldsymbol{x}) \;=\; \sum_{i=1}^{s'} f_{i1} \cdot f_{i2} \cdot f_{i3} \cdot f_{i4} \cdot f_{i5}\,, \tag{2}$$

for some top-fanin $s' = \mathsf{poly}(s, d)$, where each f_{ij} has circuit size at most s' and $\deg(f_{ij}) \leq d/2$, for all i, j. This circuit normal-form (CNF) has played a key role in all recent depth-reduction results [AV08, Koi12, Tav15, GKKS13]. We *optimize* Eqn.(2) in the following two ways: **(i)** reducing multiplication-fanin from 5 to 4, and **(ii)** further reducing degree $d/2$ to $d/3$. The new CNF is:

Theorem 1 (New CNF). *Let $f(\boldsymbol{x})$ be an n-variate, degree d polynomial computed by a circuit of size s. Then, there exist polynomials $f_{ij} \in \mathbb{F}[\boldsymbol{x}]$ such that*

$$f(\boldsymbol{x}) \;=\; \sum_{i=1}^{s'} f_{i1} \cdot f_{i2} \cdot f_{i3} \cdot f_{i4}\,, \tag{3}$$

for some top-fanin $s' = \mathsf{poly}(s, d)$, where each f_{ij} has circuit size at most $s'' = \mathsf{poly}(s, d)$ and $\deg(f_{ij}) \leq d/3$, for all i, j.

Remarks.
1. Concretely, we bound: $s' \leq O(s^6 \cdot d^{13})$ and $s'' \leq O(s \cdot d^2)$.
2. Notes on *optimality*: **1)** we cannot expect *sum of product of* ≤ 2 decomposition for the degree requirement $d/3$, **2)** *sum of product of* 3 decomposition is also not expected, since it essentially means that each f_{ij} has degree $= d/3$. Such intermediate nodes (with degree $= d/3$) might not exist in the circuit of f. (See the concept of 'frontiers' in Definition 5).

The leitmotif of this paper is the interplay between Conjecture C1 and derandomization/ hardness questions in algebraic complexity. Could the suspected hardness of $(x+1)^d$, in SO4 representation, *settle* the infamous VP vs. VNP? We evince a positive answer.

Theorem 2 (Conditional lower bound)**.** *If GRH and Conjecture C1 hold, then* VP $\neq$ VNP.

Remarks.
1. [DST20] showed that an $\Omega(d)$ lower bound on the sum of 25^{th}-powers of univariates, for f_d, suffices to prove VP $\neq$ VNP (assuming GRH). The nearly *optimal* reduction of the exponent 25, to 4, and a *weaker demand* of $\Omega(d^{0.98289})$, manifest from the new CNF (and a new proof) that overcomes the prior technical bottlenecks; see Section 3.2.
2. Our choice of $f_d = (x+1)^d$ is mostly because it is *simple*, namely, by repeated squaring, it has circuit size $\Theta(\log d)$. Interestingly, if Conjecture C1 holds for more 'intricate' polynomial families, e.g. $f = \sum_{i=0}^{d} 3^{i^2} x^i$, then we *do not* require GRH to conclude VP $\neq$ VNP! (See Remark 1 at the end of Sec.3.2.) GRH is not needed if $\mathbb{F}$ is finite with a large characteristic.
3. Whether 'merely' SOS (sum of squares of univariates) gives strong algebraic lower bounds from our proof technique, is unclear. However, in the *non-commutative* setting, lower bound on sum-of-squares (of multivariates) implies that Permanent is hard [HWY11]. Our theorem can be seen as its natural analog in the commutative setting.

Hardness of general circuits have often lead to efficient *derandomization* [AGS19, GKSS19]. Our methods in Theorem 2 consequently put blackbox-PIT in *quasi*-polynomial time (in a way similar to [KI03]). Recently, [DST20] demonstrated that the hardness of $(x+1)^d$, even for the restricted model of sum of constant (≥ 25) powers, can *completely* derandomize blackbox-PIT. The measure, which was used to establish such a connection was the *support-union*.

Support-union size of f with respect to s, denoted $U_{\mathbb{F}}(f,s)$, is defined to be the *minimum* number of distinct monomials in the representation of Eqn.(1); in other words, $\left|\bigcup_{i=1}^{s} \mathrm{supp}(\ell_i)\right|$ when f is written as Eqn.(1); it is ∞, if no such representation exists. Here, *support* $\mathrm{supp}(\ell)$ denotes the set of nonzero monomials in the polynomial $\ell(x)$. Note that, s is the *top-fanin* when Eqn.(1) is considered as a depth-4 circuit. It is easy to see: $S_{\mathbb{F}}(f) \geq \min_s \max\{s, U_{\mathbb{F}}(f,s)\}$.

A direct counting argument shows: $U_{\mathbb{F}}(f,s) \geq \Omega(|f|_1^{1/4})$. As before, we want to investigate the behaviour of $U_{\mathbb{F}}(f_d,s)$ w.r.t. d for the polynomial family $f_d := (x+1)^d$ and a given top-fanin s. Here are some interesting examples of the behaviour of $U_{\mathbb{F}}(f_d,\cdot)$.

Examples.
1. (Small s) We show that $U_{\mathbb{F}}(f_d,2) \geq d/4+1$ (Theorem 18). Also, we show that $U_{\mathbb{F}}(f_d,5) \leq d/4+4$ (Lemma 15).
2. (Large s) For $s \geq c\cdot(d+1)$ for *any* $c > 4$, we show that $U_{\mathbb{F}}(f_d,s) \leq O(d^{1/4})$ (Lemma 16). Thus, for large s, we get $U_{\mathbb{F}}(f_d,s) = \Theta(d^{1/4})$, which resolves this case.

This enthralling trade-off between the measure U and the top-fanin s in the above examples, motivated us to conjecture the following (same as [DST20, Conj.C1] with $r = 4$).

Conjecture 2 (C2)**.** *There exist positive constants $\delta_1 \leq 1$, $\delta_2 \geq 1$ such that $U_{\mathbb{F}}\left(f_d, d^{\delta_1}\right) \geq d/4^{\delta_2}$, for all large enough $d \in I$.*

Remarks.
1. The example above for large s does not apply for $\delta_1 \in (0,1]$, as $s = d^{\delta_1} \leq d$. On the other hand, by picking a large δ_2, the lower bound on U required is smaller than $d/4$.
2. We believe the conjecture to hold for *any* large $d \in \mathbb{N}$ (i.e. beyond I). We *also* believe the conjecture to be true for *most* polynomial families, e.g. $f := \sum_{i=0}^{d} 3^{i^2} x^i$ or $f := \prod_{i=1}^{d}(x-i)$.
3. The proof of Theorem 2 could be made to work with the 'stronger' Conjecture C2 as well.
4. One can ask for the number of distinct monomials required to *approximate* $f_d(x)$ as a sum of 4^{th}-powers. We believe the above conjecture to hold in the approximative computation model as well. See Conjecture C3 and its consequences in Section E.2.

[DST20] showed that studying representations like $f_d := (x+1)^d = \sum_i \ell_i^{25}$ *solves* PIT; it *requires* proving $U_{\mathbb{F}}(f_d, d^{\delta_1}) \geq \Omega(d)$ for some $\delta_1 \leq 1$. Here, we optimize the exponent (25 to 4).

Theorem 3 (Conditional derandomization)**.** *If Conjecture C2 holds, then blackbox-PIT* $\in$ P.

Remarks.

1. Older results too lead to various conditional derandomizations. E.g. *multi*-variate hard polynomials lead to blackbox-PIT $\in$ QP (*quasipoly-time*) [KI03, AGS19]. Recently, [GKSS19] showed that the circuit hardness of a k-variate polynomial yields blackbox-PIT $\in$ P, where constant $k \geq 4$ (see Theorem 28).
2. Very recent work of [DST20] improved it to $k = 1$ and showed that the hardness of a *simple univariate* polynomial, in a much *weaker* model (sum of 25^{th}-powers of univariates), also translates to complete derandomization. In this paper, we weaken the model substantially.
3. One could also work with more intricate polynomials, e.g. $\prod_{i=1}^{d}(x-i)$ or $\sum_{i=0}^{d} 3^{i^2}x^i$, whose circuit complexity is unclear, but may well be $\Omega(d)$. Showing Conjecture C2 for any of these polynomials would similarly lead us to Theorem 3.
4. For Theorem 3, we could restrict the degrees of ℓ_i, to be $O(d)$. See Section 3.3, Remark 2.
5. One can show that the *approximate* version of the conjecture (see Conjecture C3) implies a poly-time hitting-set for $\overline{\mathsf{VP}}$-circuits (Theorem 32).

1.2 Proof ideas

Proof idea of Theorem 1. The principal notions are those of *gate quotient* and *frontier decomposition*, first developed by [VSBR83], although there are important contrasts requiring non-trivial observations. Wlog, f is homogeneous, computed by a circuit Φ with fanin 2; further, for every node the degree of the right child is at least that of the left child (*right-heavy*). Let $[u]$ denote the polynomial computed at node u in Φ and the quotient $[f : u]$ somewhat behaves like a derivative $\partial_u f$, which has a *small* circuit (by inductive definition). Identify the *frontier* $\mathcal{F}$: nodes v such that $\deg(v) \geq d/3$ with the children having degree $< d/3$ (the usual frontier definition *uses* the threshold $d/2$). Induct on depth to express f as

$$f \;=\; \sum_{v \in \mathcal{F}} [v] \cdot [f : v]\,. \tag{4}$$

By homogeneity $\deg([f : v]) = \deg(f) - \deg([v])$, implying that either $[v]$ or $[f : v]$ has degree in the range $[d/3, d/2]$; and the other is in $[d/2, 2d/3]$. This is a better estimate than the *trivial* upper bound of $2d/3$ on both. See Lemma 9.

Next, in each product, we decompose the *larger* degree factor. The standard way is to prove a sum identity for $[f : u]$, in terms of quotient gates, and recurse. Instead, we will think of $[f : v]$ as new polynomials, relabel, and decompose using Eqn.(4) to get an expression like $f = \sum f_{i1}f_{i2}f_{i3}$, where each f_{ij} has degree at most $d/2$. Careful calculations show that, in each such product, at least one f_{ij} has degree $\leq d/3$ (see Lemma 10). If, *two* of them are so, we further decompose the one with degree $> d/3$ using Eqn.(4); and show that the respective polynomial factors, after the decomposition, do have degree $\leq d/3$. So, we are done in this case.

The remaining case (for a fixed i) is: $\deg(f_{i1}) \leq d/3 < \deg(f_{i2}) \leq \deg(f_{i3}) \leq d/2$. In this case, the idea is to decompose some, and redistribute the factors among the three to finally achieve a product of 4, where each factor has degree $\leq d/3$. To achieve that, we decompose f_{i1} as sum of product of 3, while f_{i2} and f_{i3} as in Eqn.(4). So that, for the time being, it is a sum of product of 7. Next, we redistribute the first 3 in the remaining 4 appropriately; clubbing (or multiply) them to get the desired sum of product of 4 representation. (See Section 3.1.)

Proof idea of Theorem 2. The elemental idea is to use $f_d := (x+1)^d$ to construct $P_{k,n}$ (*non-constant* k), a k-variate polynomial of individual degree at most n (*constant*). We show that, under GRH and assuming VP = VNP, $(P_{k,n})_k$ is an *explicit* family in VP, and is *exponentially hard* assuming Conjecture C1; leading to a contradiction.

$P_{k,n}(\boldsymbol{x})$ is s.t. after a standard Kronecker substitution: $P_{k,n}(x_1,\ldots,x_k) \mapsto P_{k,n}(x^{(n+1)^0},\ldots, x^{(n+1)^{k-1}}) = f_d(x)$, where $d := d(k) \leq (n+1)^k - 1$ is the *largest* possible in I. This map is a bijection between $\mathrm{supp}(P_{k,n})$ and $\mathrm{supp}(f_d)$. Note that, from definition of I and d, $k = \Theta(\log d)$.

We prove that $P_{k,n}$ requires $s := d^{1/c} = 2^{\Omega(k)}$-size circuit, where c is a constant to be fixed later. Assume the contrary, $\mathrm{size}(P_{k,n}) \leq s$. Using Theorem 1 and Fischer's trick, one can write $P_{k,n} = \sum c_i \cdot \tilde{f}_i^4$, with $\deg(\tilde{f}_i) \leq \deg(P_{k,n})/3 \leq kn/3$, and top-fanin $\mathsf{poly}(d^{1/c}, k)$ (as n is constant). A standard combinatorial argument shows: $|\tilde{f}_i|_1 \leq \binom{k+kn/3}{k}$. As Kronecker substitution cannot increase sparsity, $S_{\mathbb{F}}(f_d) \leq \mathsf{poly}(d^{1/c}, k) \cdot \binom{k+kn/3}{k}$.

Stirling approximation (Eqn.5) gives: $\binom{k+kn/3}{k} \leq (e(1+n/3))^k$. We want to find the *minimum* α such that $(e(1+n/3))^k \leq d^\alpha$. As $d = \Omega((n+1)^k)$, we basically want to optimize $\alpha := \min_n (\log(e(1+n/3))/\log(n+1)) \approx 0.98285$ (at $n \approx 113.62$). For easier representation, we choose $n+1 = 2^7$ for which $\alpha \approx 0.982872$; $c := 6 \times 10^5$ so that $\alpha + 6/c < 0.98289$; factor 6 comes because calculation shows that 6 appears in the exponent of s in Theorem 1. This translates to the fact that $S_{\mathbb{F}}(f_d) \leq o(d^{0.98289})$, contradicting Conjecture C1. (See Section 3.2.)

Now assume that GRH is true and VP = VNP. Theorem 7 shows that the counting hierarchy (CH) collapses to P/poly. It is not hard to show that each bit of $\binom{d}{i}$, the coefficients of f_d, is computable in $\mathsf{CH} \subseteq \mathsf{P/poly}$. Thus, Valiant's criterion implies $(P_{k,n})_k \in \mathsf{VNP} = \mathsf{VP}$; contradicting the $2^{\Omega(k)}$-hardness of $P_{k,n}$ proved above from Conjecture C1. Hence, we conclude VP $\neq$ VNP.

Proof idea of Theorem 3. Again, the basic idea is to use $f_d := (x+1)^d$ to construct $P_{k,n}$, but now it is a k-variate (for *constant* k) polynomial of individual degree at most n (for *non-constant* n). It is shown hard assuming Conjecture C2. With appropriate parameters, this hardness will lead us to an efficient hitting-set for VP using the recent result of [GKSS19]; see Theorem 28.

The construction of $P_{k,n}(\boldsymbol{x})$ is such that $P_{k,n}(x_1,\ldots,x_k) \mapsto P_{k,n}(x^{(n+1)^0},\ldots,x^{(n+1)^{k-1}}) = f_d(x)$, where $d := d(n)$ is the *largest* element in I which is $\leq (n+1)^k - 1$, and k depends on δ_1, δ_2.

We prove that $\mathrm{size}(P_{k,n}) > d^{1/\mu} =: s$, where $\mu \geq 1$ is a constant which depends on δ_1, δ_2. Assume the contrary, $\mathrm{size}(P_{k,n}) \leq s$. Using Theorem 1 and Fischer's trick, one can write $P_{k,n} = \sum c_i \cdot \tilde{f}_i^4$, with $\deg(\tilde{f}_i) \leq \deg(P_{k,n})/3 \leq kn/3$, and top-fanin $\mathsf{poly}(d^{1/\mu}, n)$ (as k is constant). Direct combinatorial argument shows: $|\cup_i \mathrm{supp}(\tilde{f}_i)| \leq \binom{k+kn/3}{k}$. Kronecker map yields, $f_d = \sum c_i \cdot g_i^4$; wherein there are at most d^{δ_1} summands and the support-union $|\cup_i \mathrm{supp}(g_i)| < d/4^{\delta_2}$. As Kronecker map does not increase the top-fanin and support, it contradicts Conjecture C2.

The coefficients $\binom{d}{i}$ of $P_{k,n}$ can be computed in poly(d)-time. Hence, $P_{k,n}$ is both explicit and hard! The hardness is $d^{1/\mu} \geq \Omega((n+1)^{k/\mu}) = \Omega(n^{k/\mu})$, and $\deg(P_{k,n}) = O(n)$. Thus, for $k > 3\mu$, we invoke Theorem 28, using $P_{k,n}$ to construct a poly-time hitting-set for VP-circuits.

*N*ote: One can invoke Eqn.(5) to show that $\binom{k+kn/3}{k} \geq (1+n/3)^k \geq \Omega(d)$, thus one *cannot* hope to weaken the lower bound in Conjecture C2 & still get Theorem 2. (See Section 3.3.)

1.3 Previously known results– Circuit complexity, sum-of-powers

It is known that the computation of most of the d-degree polynomials require $\Omega(d)$ many arithmetic operations [Mot55, Bel58]. In fact, $\sum_{i=0}^{d} 2^{2^i} x^i$ requires $\Omega\left(\sqrt{d/\log d}\right)$ size circuits [Str74]. It can be converted to an *exponentially hard* polynomial family (f_n), but unfortunately *cannot* separate VP and VNP; because of the *non-explicitness* of (f_n); for details see [HS80, Bür13].

The classical *Waring problem* is to find the number $g(k)$, for every $k \in \mathbb{N}$, such that every natural number can be written as the sum of $g(k)$-many k^{th}-powers of numbers. Some remarkable examples are $g(3) = 9$ [Kem12] and $g(4) = 19$ [BDD86]. Many variants of Waring's problem for polynomials have been investigated using analytic tools [FOS12, CCG12, BT15].

For representations like $f = \sum c_i \cdot Q_i^{e_i}$, for $\deg(Q_i) \leq t$, a lower bound of $s \geq \Omega(\sqrt{d/t})$ was shown in [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]. Note that we want lower bounds with $|Q_i|_1 \leq t$ (with unrestricted degree) to separate VP and VNP [Koi12]. These lower bound questions innately appear in the algebraic-geometry approaches to the P $\neq$ NP question, e.g. [Muk16, GMQ16, Gro15, Mul17].

'Hardness to derandomization' *intrinsically* demands deterministic PIT algorithms given *explicit hard polynomials*. The celebrated *Polynomial Identity Lemma* [Ore22, DL78, Zip79, Sch80]) and efficient evaluation at *random* points lead to randomized poly-time algorithm for blackbox-PIT. For details, see the surveys [Sax09, Sax14, SY10, KS19] or review articles [Wig17, Mul17].

Most lower bound connections used *hitting-set generator* (HSG) via *combinatorial designs* to give *quasi*-poly-time blackbox-PIT, see [AV08, Bür09, Koi11, Koi12, Tav15]. Recently, it was shown that HSG is compliant to *bootstrapping* (of variables) [AGS19, KST19]. Finally, [GKSS19] came up with an HSG without designs; showing how ample hardness of *constant*-variate ($k \geq 4$) polynomials implies blackbox-PIT $\in$ P. They used *generalized* derivatives of the hard polynomial.

2 Preliminaries

Basic notation. Denote the underlying field as $\mathbb{F}$ and assume that it is $\mathbb{Q}, \mathbb{Q}_p$, or their fixed extensions. Our results hold also for finite fields of large characteristic.

Let $[n] = \{1, \ldots, n\}$. 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 $i = \sum_{j=1}^{k} i_j \, b^{j-1}$. In the special case $b = 2$, we define $\text{bin}(i) = \text{base}_2(i)$.

For binomial coefficients, we use an easy bound based on the e^k-series [Wik], for $1 \leq k \leq n$,

$$\left(\frac{n}{k}\right)^k \leq \binom{n}{k} \leq \left(\frac{en}{k}\right)^k. \tag{5}$$

Polynomials. For a multivariate polynomial $p \in \mathbb{F}[\boldsymbol{x}]$, 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 $|p|_1 := |\text{supp}(p)|$. By $\text{coef}(p)$ we denote the *coefficient vector* of p (in some fixed order). For polynomials $p_1, \ldots, p_s \in \mathbb{F}[\boldsymbol{x}]$, their *span* is the vector space $\text{span}_{\mathbb{F}}(p_1, \ldots, p_s) := \{ \sum_i c_i p_i \mid c_i \in \mathbb{F} \}$.

For an exponent vector $\boldsymbol{e} = (e_1, \ldots, e_k)$, we use $\boldsymbol{x}^{\boldsymbol{e}}$ to denote the monomial $x_1^{e_1} \ldots x_k^{e_k}$.

By $\mathbb{F}[x]^{\leq d}$ we denote the $\mathbb{F}$-vector space of univariate polynomials of degree at most d.

Algebraic circuit complexity. For a polynomial f, the size of the smallest circuit computing f is denoted by $\text{size}(f)$, 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 $\text{size}(C) \leq s$. The *circuit complexity of a family* $(P_n)_n$ is $g(n)$, if $\text{size}(P_n) = \Theta(g(n))$.

Circuit Normal Form (CNF) and Frontier. In an elegant and influential work, [VSBR83] showed that every efficiently computable polynomial family (by algebraic circuits) is also efficiently computable in *parallel*. Work of [AJMV98] proved a similar result with top-down approach. In both the proof techniques, the notion of *gate quotients* was used. A hybrid, detailed discussion can be sought in [Sap19]. For completeness, we define and go through the important details which will be required for proving our new CNF (Theorem 1).

We assume without loss of generality that the given circuit Φ has the following properties: **(i)** Φ is a *homogeneous* circuit, **(ii)** all multiplication gates in Φ have fanin at most *two*, and **(iii)** Φ is a *right heavy* circuit, i.e. the degree of the right child of any multiplication gate is at least as large as the degree of its left child. For any gate u in Φ, we denote by $[u]$ the polynomial computed at gate u. We denote u_L and u_R as left and right child of u respectively.

Definition 4 (Gate quotient). *For gates u, v, the* quotient *polynomial $[u : v]$ is defined as follows:*

1. *If u and v are same nodes, then $[u : v] = 1$*

2. *If u is a leaf, and $u \neq v$, then* $[u : v] = 0$
3. *If $u = u_1 + u_2$, then* $[u : v] = [u_1 : v] + [u_2 : v]$
4. *If $u = u_1 \times u_2$, then* $[u : v] = [u_1] \cdot [u_2 : v]$

Observation. $[u : v]$ is a homogeneous polynomial of degree $\deg(u) - \deg(v)$.

Definition 5 (Frontier)**.** *For any parameter m, define the* frontier at degree m*, denoted $\mathcal{F}_m$, as the deepest nodes in the circuit that have degree at least m. Formally, $\mathcal{F}_m := \{v : deg(v) \geq m,$ and $deg(v_L), deg(v_R) < m\}$.*

Observe that all frontier nodes must be multiplication gates. They give a nice decomposition:

Lemma 6 (CNF by frontier decomposition)**.** *[Sap19, Lem.5.12] Let Φ be a homogeneous, right-heavy circuit. Let u be a node such that $deg(u) \geq m$. Then, $[u] = \sum_{w \in \mathcal{F}_m} [u : w] \cdot [w]$.*

Valiant's hypothesis and GRH. Bürgisser [Bür00, Cor.1.2] showed that if Valiant's hypothesis is false, and GRH holds, then the polynomial hierarchy collapses. Similarly,

Theorem 7 (CH collapse)**.** [DST20, Thm.9] *If GRH is true and* $\mathsf{VP} = \mathsf{VNP}$*, then* $\mathsf{CH} \subseteq \mathsf{P/poly}$.

Over finite fields, GRH is not needed; GRH is required only for $\mathbb{Q}$. See Section C.
A useful *sufficient* condition for a polynomial family $(f_n(\boldsymbol{x}))_n$ to be in VNP is known, due to Valiant [Val79b]. For a proof, see [Bür13]. We use the slightly modified version of the criterion.

Theorem 8 (Valiant's criterion, [Val79b])**.** *Let function $\phi : [0, c]^* \to \mathbb{N}$ be in $\#\mathsf{P/poly}$, for some constant $c \in \mathbb{N}$. Then, the family of polynomials defined by $f_n(\boldsymbol{x}) := \sum_{\boldsymbol{e} \in [0,c]^n} \phi(\boldsymbol{e}) \cdot \boldsymbol{x}^{\boldsymbol{e}}$, is in* VNP.

Kronecker map and its inverse. Let $p(x_1, \ldots, x_k)$ be a polynomial, where the variables have individual degree bounded by n. The *Kronecker map* $\phi_{k,n}(p)(x)$ yields a univariate polynomial by replacing variable x_i in p by $x^{(n+1)^{i-1}}$, for all $i \in [k]$.

The map has the property that any polynomial with individual degree at most n gets *uniquely* mapped to a univariate polynomial of degree at most $d = \sum_{i=1}^{k} n(n+1)^{i-1} = (n+1)^k - 1$ [Kro82].

Next, we consider the inverse map. Let $q(x)$ be a univariate polynomial of degree d. For $k \geq 1$, let $\boldsymbol{x} := (x_1, \ldots, x_k)$ and $n := \lceil (d+1)^{1/k} \rceil - 1 \geq 1$. The *inverse Kronecker map* $\psi_{k,d}(q)(\boldsymbol{x})$ yields a k-variate polynomial by replacing x^i, in q, by the monomial $\boldsymbol{x}^{\mathrm{base}_{n+1}(i)}$, for all $i \in [k]$.

It is easy to see that $\psi_{k,d}$ maps each x^i to a *distinct* k-variate monomial of individual degree $\leq n$, for $0 \leq i \leq d$. Also, we have $\phi_{k,n} \circ \psi_{k,d}(q) = q$. Thus, $\phi_{k,n} \circ \psi_{k,d} = \mathrm{id}$ over $\mathbb{F}[x]^{\leq d}$.

3 Proof of the main results

3.1 The new CNF : Proof of Theorem 1

Before proving the representation theorem for $f(\boldsymbol{x})$ as the sum of product of 4, we study the decomposition of a polynomial as sum of product of 2 (and *subsequently* 3). This will help us to get the desired structure theorem. Here is a very pertinent decomposition lemma:

Lemma 9 (Sum of product of 2)**.** *Let $f(\boldsymbol{x})$ be an n-variate, homogeneous, degree d polynomial computed by a right-heavy homogeneous circuit Φ of size s. Then, there exist polynomials $f_{ij} \in \mathbb{F}[\boldsymbol{x}]$ s.t.*

$$f(\boldsymbol{x}) = \sum_{i=1}^{s} f_{i1} \cdot f_{i2}\,, \quad \textit{with the following properties:} \tag{6}$$

1. *$d/3 \leq deg(f_{i1}) \leq d/2 \leq deg(f_{i2}) \leq 2d/3$, for all $i \in [s]$,*

2. $deg(f_{i1}) + deg(f_{i2}) = d$, *for all* $i \in [s]$, *and*
3. *each* f_{ij} *has a right-heavy homogeneous circuit of size at most* $s_2 := O(s)$.

Proof of Lemma 9. Choose $m := d/3$ in Lemma 6, to conclude that, $f = \sum_{u \in \mathcal{F}_m} [f : u] \cdot [u]$. Note that, $|\mathcal{F}_m| \leq s$. By definition of the frontier, $\deg(u_L), \deg(u_R) < d/3$ (implying $\deg([u]) < 2d/3$). Recall that $\deg([u]) + \deg([f : u]) = d$. Combining together, we get the range $d/3 \leq \deg([u])\ ,\ \deg([f : u]) \leq 2d/3$.

If $\deg([u]) \in [d/3, d/2]$, then $\deg([u]) + \deg([f : u]) = d \implies \deg([f : u]) \in [d/2, 2d/3]$. Otherwise, $\deg([u]) \in [d/2, 2d/3]$. In that case, $\deg([f : u]) \in [d/3, d/2]$. After suitable relabelling, we get Eqn.(6) with the desired degree properties.

Size analysis. We will prove that for *any* node u in Φ, $[f : u]$ can be computed by a right-heavy homogeneous circuit of size $\leq 2s$.

Fix node u. Maintain a 'growing' disjoint *copy* of Φ as circuit Φ', which is intended to inductively compute $[v : u]$, for the 'current' node v. We induct on depth (of v in Φ). The *base case* (namely, a leaf) occupies size ≤ 1 in Φ', but no extra edge/node needs to be added.

Suppose, we have computed Φ' bottom-up, till i-th level ($i \geq 1$). For $i+1$-th level, we need to compute at most two quotients of the form $[v : u]$ (by definition). **Addition:** If $v =: v_L + v_R$, then $[v : u] = [v_L : u] + [v_R : u]$. By induction hypothesis, both the quotients are already computed in Φ'. The two input edges and the $+$ gate are already in Φ'. Thus, we do not need to add any extra edge or node in Φ' (i.e. they are copies of those in Φ). **Multiplication:** If $v = v_L \cdot v_R$, then $[v : u] = [v_L] \cdot [v_R : u]$. By induction hypothesis, both are pre-computed; $[v_L]$ in Φ and $[v_R : u]$ in Φ'. We *delete* the left incoming-edge of $\times$ in Φ', replacing it with an edge from v_L of Φ. So, no extra edge or node is required.

Thus, $[f : u]$ has a circuit Φ' (with Φ) of size at most $s_2 := 2s$. Φ' is *homogeneous* because all the intermediate nodes $[v]$, and $[v : u]$, are homogeneous polynomials. It has fanin 2 again, by the definition. The *right-heaviness* follows by swapping the sub-circuits appropriately without incurring any size blowup. □

The above lemma can be applied many-fold, to decompose $f(\boldsymbol{x})$ as a sum of product of 3 polynomials with better degree restrictions. We claim the following (for proof refer Section B).

Lemma 10 (Sum of product of 3)**.** *Let* $f(\boldsymbol{x})$ *be an* n*-variate, homogeneous, degree* d *polynomial computed by a right-heavy homogeneous circuit of size* s*. Then,* $\exists f_{ij} \in \mathbb{F}[\boldsymbol{x}]$ *such that (for* $s_1 := O(s^2)$*),*

$$f(\boldsymbol{x}) \ = \ \sum_{i=1}^{s_1} f_{i1} \cdot f_{i2} \cdot f_{i3}\ , \quad \textit{with the following properties:} \tag{7}$$

1. $deg(f_{i1}) \leq deg(f_{i2}) \leq deg(f_{i3}) \leq d/2$, *for all* $i \in [s_1]$.
2. $deg(f_{i1}) \leq d/3 \leq deg(f_{i3})$ *and* $deg(f_{i1}) + deg(f_{i2}) + deg(f_{i3}) = d$, *for all* $i \in [s_1]$.
3. *each* f_{ij} *has right-heavy homogeneous circuit of size at most* $s_3 = O(s)$.

Proof of Theorem 1. Now, we are set to prove the main theorem. For the time being, assume that the given $f(\boldsymbol{x})$ is a degree d polynomial which can be computed by a right-heavy homogeneous circuit of size s (as later, we can use this on each *homogeneous part* to finish). Using Lemma 10, we can decompose $f(\boldsymbol{x})$ with summand being $s_1 := O(s^2)$ as follows:

$$f(\boldsymbol{x}) \ = \ \left(\sum_{i \in S_1} f_{i1} \cdot f_{i2} \cdot f_{i3} \right) + \left(\sum_{i \in S_2} f_{i1} \cdot f_{i2} \cdot f_{i3} \right) \ =: \ G \ + \ H\ , \tag{8}$$

where S_1 and S_2 are disjoint subsets of $[s_1]$ such that $S_1 \cup S_2 = [s_1]$ with the following properties:
(i) $\deg(f_{i1}) \leq \deg(f_{i2}) \leq d/3 \leq \deg(f_{i3}) \leq d/2\ ,\ \forall\ i \in S_1$.
(ii) $\deg(f_{i1}) \leq d/3 \leq \deg(f_{i2}) \leq \deg(f_{i3}) \leq d/2\ ,\ \forall\ i \in S_2$.

(iii) $\deg(f_{i1}) + \deg(f_{i2}) + \deg(f_{i3}) = d$, for all $i \in [s_1]$.
(iv) Each f_{ij} can be computed by a right-heavy homogeneous circuit of size at most $s_3 := O(s)$, $\forall\ (i,j) \in [s_1] \times [3]$.

Decomposition of G**.** For $i \in S_1$, we further decompose each f_{i3} (of size s_3) as Eqn.(6), using Lemma 9, to get:

$$G = \sum_{i \in S_1} f_{i1} \cdot f_{i2} \cdot f_{i3} = \sum_{i \in S_1, j \in [s_3]} f_{i1} \cdot f_{i2} \cdot f_{i3,j1} \cdot f_{i3,j2} \,. \tag{9}$$

As, $i \in S_1$, we have $\deg(f_{i1}) \leq \deg(f_{i2}) \leq d/3$. By the degree property in Lemma 9, it follows: $\deg(f_{i3,j1}) \leq \deg(f_{i3,j2}) \leq (2/3) \cdot \deg(f_{i3}) \leq 2/3 \cdot d/2 = d/3$.

Each polynomial $f_{i3,jk}$, for $j \in [s_3], k \in [2]$, can be computed by a right-heavy homogeneous circuit of size at most $O(s_3) = O(s)$. By Lemma 10, each f_{ik}, for $i \in S_1, k \in [2]$, already has $s_2 = O(s)$ size right-heavy homogeneous circuit. The top fanin is at most $|S_1| \cdot s_3 = O(s^3)$. Thus, G admits the desired decomposition.

Decomposition of H**.** For $i \in S_2$, we decompose each f_{i1} (of size s_3) as Eqn.(7) using Lemma 10; and each f_{i2} (resp. f_{i3}), of size s_3, as Eqn.(6) using Lemma 9 to get:

$$f_{i1} \cdot f_{i2} \cdot f_{i3} = \sum_{(j,k,\ell) \in [s'] \times [s_3] \times [s_3]} (f_{i1,j1} \cdot f_{i1,j2} \cdot f_{i1,j3}) \cdot (f_{i2,k1} \cdot f_{i2,k2}) \cdot (f_{i3,\ell 1} \cdot f_{i3,\ell 2}) \,. \tag{10}$$

Note that, $s' := O(s_3^2) = O(s^2)$. Thus, the top fanin is $s'' := s' \cdot s_3^2 = O(s^4)$. Given (j,k,ℓ), we want to club the 7 factors in Eqn.(10) into 4. To attain that, we consider the following two cases.

Case I: ($\deg(f_{i1,j1}) + \deg(f_{i2,k2}) \leq d/3$**).** In this case, we club the 7 product as follows:

$$(f_{i1,j1} \cdot f_{i2,k2}) \cdot (f_{i1,j2} \cdot f_{i2,k1}) \cdot (f_{i1,j3} \cdot f_{i3,\ell 1}) \cdot (f_{i3,\ell 2})$$

We claim that each product polynomial inside the brackets has degree $\leq d/3$. Because,

(i) $\deg(f_{i1,j1} \cdot f_{i2,k2}) = \deg(f_{i1,j1}) + \deg(f_{i2,k2}) \leq d/3$ by assumption.

(ii) $\deg(f_{i1,j2} \cdot f_{i2,k1}) = \deg(f_{i1,j2}) + \deg(f_{i2,k1}) \leq \deg(f_{i1})/2 + \deg(f_{i2})/2 \leq d/3$. The first inequality follows directly from the degree bound on the decomposition of f_{i1} using Lemma 10 and f_{i2} using Lemma 9. The last inequality follows from degree bound on the decomposition of f using Lemma 10, and since $\deg(f_{i3}) \geq d/3 \implies \deg(f_{i1}) + \deg(f_{i2}) \leq 2d/3$.

(iii) $\deg(f_{i1,j3} \cdot f_{i3,\ell 1}) = \deg(f_{i1,j3}) + \deg(f_{i3,\ell 1}) \leq \deg(f_{i1})/2 + \deg(f_{i3})/2 \leq d/3$. The first inequality follows directly from the degree bound on the decomposition of f_{i1} using Lemma 10 and f_{i3} using Lemma 9. The last inequality follows, since $i \in S_2$ means that $\deg(f_{i2}) \geq d/3$; implying $\deg(f_{i1}) + \deg(f_{i3}) \leq 2d/3$.

(iv) $\deg(f_{i3,\ell 2}) \leq (2/3) \cdot (\deg(f_{i3})) \leq (2/3) \cdot (d/2) = d/3$. Both the inequalities follows directly from the degree bounds on the decomposition of f_{i3} and f using Lemma 9 and 10 respectively.

Case II: ($\deg(f_{i1,j1}) + \deg(f_{i2,k2}) > d/3$**).** In this case, we club the 7 factors in Eqn.(10) as follows:

$$(f_{i1,j1} \cdot f_{i3,\ell 1}) \cdot (f_{i1,j2} \cdot f_{i1,j3} \cdot f_{i2,k1}) \cdot (f_{i2,k2}) \cdot (f_{i3,\ell 2})$$

We claim that each factor polynomial inside the bracket has degree $\leq d/3$. Because:

(i) $\deg(f_{i1,j1} \cdot f_{i3,\ell 1}) = \deg(f_{i1,j1}) + \deg(f_{i3,\ell 1}) \leq \deg(f_{i1})/3 + \deg(f_{i3})/2 < 1/2 \cdot (\deg(f_{i1}) + \deg(f_{i3})) \leq d/3$. The first inequality follows directly from the degree bound on the decomposition of f_{i1} using Lemma 10 and f_{i3} using Lemma 9. The last inequality follows, as $i \in S_2$ means $\deg(f_{i2}) \geq d/3$; implying $\deg(f_{i1}) + \deg(f_{i3})) \leq 2d/3$.

(ii) $\deg(f_{i1,j2} \cdot f_{i1,j3} \cdot f_{i2,k1}) = \deg(f_{i1,j2}) + \deg(f_{i1,j3}) + \deg(f_{i2,k1}) = \deg(f_{i1}) - \deg(f_{i1,j1}) + \deg(f_{i2}) - \deg(f_{i2,k2}) = (\deg(f_{i1}) + \deg(f_{i2})) - (\deg(f_{i1,j1}) + \deg(f_{i2,k2})) < 2d/3 - d/3 = d/3$. The second equality follows from homogeneity in the decomposition in Lemmas 9-10: $\deg(f_{i1,j1}) + \deg(f_{i1,j2}) + \deg(f_{i1,j3}) = \deg(f_{i1})$ and $\deg(f_{i2,k1}) + \deg(f_{i2,k2}) = \deg(f_{i2})$. The last inequality follows from the degree bound on the decomposition of f using Lemma 10,

i.e. $\deg(f_{i3}) \geq d/3 \implies \deg(f_{i1}) + \deg(f_{i2}) \leq 2d/3$. The subtracted part follows from the assumption in Case II.

(iii) $\deg(f_{i2,k2}) \leq (2/3) \cdot (\deg(f_{i2})) \leq (2/3) \cdot (d/2) = d/3$. Both the inequalities follow directly from the degree bounds on the decomposition of f_{i2} and f, using Lemmas 9-10 respectively.

(iv) $\deg(f_{i3,\ell 2}) \leq (2/3) \cdot (\deg(f_{i3})) \leq (2/3) \cdot (d/2) = d/3$. Both the inequalities follow directly from the degree bounds on the decomposition of f_{i3} and f, using Lemmas 9-10 respectively.

Thus, in both the cases, we clubbed the product of 7, into that of 4, with the desired degree properties. Thus, Eqn.(10) can be re-written to get:

$$H \;=\; \sum_{i \in S_2} f_{i1} \cdot f_{i2} \cdot f_{i3} \;=\; \sum_{m=1}^{|S_2| \cdot s''} \tilde{f}_{m1} \cdot \tilde{f}_{m2} \cdot \tilde{f}_{m3} \cdot \tilde{f}_{m4}$$

as each $f_{i1} \cdot f_{i2} \cdot f_{i3}$ could be decomposed as Eqn.(3) with the desired degree properties. Adding up G and H in Eqn.(8), it follows that $f(\boldsymbol{x})$ has the decomposition as Eqn.(3), with the desired degree bound; and the top fanin being at most $|S_2| \cdot s'' = O(s^6)$.

Size analysis. As in Lemmas 9-10, the size of the factor polynomials in the product is at most $O(s)$. Also, each factor polynomial in the product in Eqns.(9-10) is computed by a circuit of size at most $O(s)$. While clubbing 7 factors into 4, we multiply at most 3 of them together into one which again is of size at most $O(s)$. As discussed above, the top fanin can be at most $O(s^6)$.

Given an (*in*homogeneous) $f(\boldsymbol{x})$ of degree d, computed by Φ of size s, one can show that there is a right-heavy homogeneous Φ' of size $O(s \cdot d^2)$ computing each homogeneous-part of f (see [SY10]). One can apply the above decomposition on each i-th degree homogeneous-part; to get the desired decomposition with the top fanin being $O(d \cdot (sd^2)^6) = O(s^6 \cdot d^{13})$. The degree of each factor polynomial inside the product is $\leq d/3$ and each can be computed by a circuit of size at most $O(sd^2)$. □

3.2 Conjecture C1 to VP $\neq$ VNP: Proof of Theorem 2

Proof of Theorem 2. Let GRH and Conjecture C1 be true. For a *non-constant* parameter k, let $\boldsymbol{x} := (x_1, \ldots, x_k)$. For all $k \in \mathbb{N}$, take $d := d(k) = 2^{7k} - 1 \in I$, by definition. Thus, $k = \Theta(\log d)$.

Define the polynomial family $P_{k,n}(\boldsymbol{x}) := \psi_{k,d}(f_d)$ via the inverse Kronecker map (with $n := 127$) applied to $f_d := (x+1)^d$. Note that $P_{k,n}$ is a k-variate polynomial with individual degree being at most $\lceil (d+1)^{1/k} \rceil - 1 = 2^7 - 1 = 127 = n$. Hence, $P_{k,n}$ has total degree at most $kn = 127 \cdot k$.

For the sake of contradiction, assume that VP = VNP. Then, we show that $(P_{k,n})_k \in$ VP.

Claim 11. *GRH and* VP = VNP $\implies$ $(P_{k,n})_k \in$ VP.

Proof of Claim 11. By definition of the polynomial family $P_{k,n}$, we have

$$P_{k,n}(\boldsymbol{x}) \;=\; \sum_{e=0}^{128^k - 1} \phi(e) \cdot \boldsymbol{x}^{\mathrm{base}_{128}(e)} \,,$$

where $\mathrm{base}_{128}(e) =: (e_1, \ldots, e_k)$ such that, $e \;=\; \sum_{i=1}^{k} e_i \cdot 128^{i-1}$ and $\phi(e) := \mathrm{coef}(\boldsymbol{x}^{\boldsymbol{e}})(P_{k,n}) = \binom{d}{e}$. Also, the number of monomials in $P_{k,n}$ is $|\mathrm{supp}(P_{k,n})| = |\mathrm{supp}(f_d)| = d + 1$.

Clearly, $\phi(e) < 2^d \leq 2^{128^k - 1} < 2^{128^k} - 1$. Write $\phi(e)$ in binary, i.e. $\phi(e) =: \sum_{j=0}^{128^k - 1} \gamma_{e,j} \, 2^j$, where $\gamma_{e,j} \in \{0, 1\}$. From Theorem 22, we know that the sequence of coefficients $\phi(e) = \binom{d}{e}$ is CH-definable. In particular, this means that $\gamma_{e,j}$'s are computable in CH, and hence in P/poly, by our assumptions together with Theorem 7.

We introduce new variables $\boldsymbol{y} = (y_1, \ldots, y_{7k})$ and consider the auxiliary polynomial $\tilde{\phi}_e(\boldsymbol{y}) := \sum_{j=0}^{128^k-1} \gamma_{e,j}\, \boldsymbol{y}^{\mathrm{bin}(j)}$, where $\mathrm{bin}(j) =: (j_1, \ldots, j_{7k})$ so that $j = \sum_{i=1}^{7k} j_i\, 2^{i-1}$. Let $\boldsymbol{y}_0 := (2^{2^0}, 2^{2^1}, \ldots, 2^{2^{7k-1}})$. Note that $\boldsymbol{y}^{\mathrm{bin}(j)}|_{\boldsymbol{y}=\boldsymbol{y}_0} = 2^j$. Therefore $\tilde{\phi}_e(\boldsymbol{y}_0) = \phi(e)$. Now define

$$\tilde{P}_{k,n}(\boldsymbol{x}, \boldsymbol{y}) \;:=\; \sum_{j=0}^{2^{7k}-1} \sum_{e=0}^{2^{7k}-1} \gamma_{e,j} \cdot \boldsymbol{y}^{\mathrm{bin}(j)} \cdot \boldsymbol{x}^{\mathrm{base}_{128}(e)} \,,$$

It is straightforward to see that $P_{k,n}(\boldsymbol{x}) = \tilde{P}_{k,n}(\boldsymbol{x}, \boldsymbol{y}_0)$. Since $\gamma_{e,j}$ (as a function in $\mathrm{bin}(j), \mathrm{base}_{128}(e)$) is in $\mathsf{P/poly}$, we have $(\tilde{P}_{k,n})_k \in \mathsf{VNP} = \mathsf{VP}$, by Theorem 8 (taking $c = 127$). As VP is closed under substitution, we have $(P_{k,n})_k \in \mathsf{VP}$ as well. □

On the other hand, next we show that Conjecture C1 implies that $(P_{k,n})_k \notin \mathsf{VP}$.

Claim 12. *Conjecture C1* $\implies$ $(P_{k,n})_k \notin \mathsf{VP}$.

Proof of Claim 12. We prove that $\mathrm{size}(P_{k,n}) > d^{10^{-5}/6} = 2^{\Omega(k)}$. Assume the contrary. Then, there exists an infinite subset $J \subset \mathbb{N}$ such that $\mathrm{size}(P_{k,n}) \le d^{10^{-5}/6}$, for all $k \in J$. We will show that Conjecture C1 is false over an infinite subset $J' := \{d(k) \;: k \in J\} \subseteq I$ which is a contradiction.

Let C be a circuit of size $\le d^{10^{-5}/6}$ that computes $P_{k,n}$, for some k. Then, using Theorem 1, we know that there exists $f_{ij} \in \mathbb{F}[\boldsymbol{x}]$ of degree at most $\deg(P_{k,n})/3 \le 127k/3$ such that

$$P_{k,n} \;=\; \sum_{i=1}^{s'} f_{i1} \cdot f_{i2} \cdot f_{i3} \cdot f_{i4} \,,$$

where $s' \le c \cdot d^{10^{-5}} \cdot (127k)^{13}$, for some constant c. We apply Fischer's formula on each product $\prod_j f_{ij}$, to write $P_{k,n}$ as

$$P_{k,n} \;=\; \sum_{i \in [16 \cdot s']} c_i \cdot \tilde{f}_i^4 \,, \tag{11}$$

where $\tilde{f}_i \in \mathrm{span}_{\mathbb{F}} \left(f_{jk} \mid j \in [s'], k \in [4] \right)$, for every i. Thus, $\deg(\tilde{f}_i) \le 127k/3$. Applying the Kronecker map $\phi_{k,n}$ to $P_{k,n}$ yields

$$f_d \;=\; \phi_{k,n}(P_{k,n}) \;=\; \sum_{i=1}^{16s'} c_i \cdot \left(\phi_{k,n}(\tilde{f}_i)\right)^4 \,,$$

Note that, a simple combinatorial argument shows that $|\tilde{f}_i|_1 \le \binom{k+127k/3}{k} = \binom{130k/3}{k}$. As Kronecker map $\phi_{k,n}$ cannot increase the sparsity, it follows that $S_{\mathbb{F}}(f_d) \le 16s' \cdot \binom{130k/3}{k} =: s_1$.

We want to show that $s_1 < o(d^{0.98289})$, for all large enough k. Then, we will have $S_{\mathbb{F}}(f_d) < o(d^{0.98289})$, for all large $d \in J' \subseteq I$ which contradicts Conjecture C1. For all large enough d,

$$\begin{aligned} s_1 \;\le\; 16 \cdot c \cdot d^{10^{-5}} \cdot (127k)^{13} \cdot \binom{130k/3}{k} \;&<\; (16c \cdot 127^{13}) \cdot k^{13} \cdot d^{10^{-5}} \cdot (130e/3)^k \\ &<\; c' \cdot k^{13} \cdot d^{10^{-5}} \cdot (128)^{\alpha \cdot k} \\ &<\; c' \cdot k^{13} \cdot d^{10^{-5}} \cdot 2^{\alpha} \cdot d^{\alpha} \\ &=\; c' \cdot k^{13} \cdot 2^{\alpha} \cdot d^{0.982882} \;<\; o(d^{0.98289}) \,, \end{aligned}$$

where $\alpha := 0.982872 > \log_{128}(130e/3) = 0.982871\ldots$ and $c' := 16c \cdot (127)^3$. We invoked the inequality in Eqn.(5) and also the fact that $d = (2^{7k} - 1) > (1/2) \cdot 128^k$, for $k \ge 1$. This proves Claim 12. □

Since Claim 12 contradicts Claim 11, we conclude that $\mathsf{VP} \neq \mathsf{VNP}$, finishing Theorem 3. □

Remarks. 1. We can consider $\tilde{f}_d(x) := \sum_{i=0}^{d} 3^{i^2} x^i$ and redefine $P_{k,n}$ as above. Consider the polynomial $\tilde{P}_{k,n}(\boldsymbol{x}, \boldsymbol{y})$ defined on $3k$ variables $\boldsymbol{x} = (x_1, \ldots, x_k)$, $\boldsymbol{y} = (y_1, \ldots, y_{2k})$ by $\tilde{P}_{k,n}(\boldsymbol{x}, \boldsymbol{y}) := \sum_{e=0}^{2^{7k}-1} \phi(e) \cdot \boldsymbol{y}^{\mathrm{base}_{128}(e^2)} \cdot \boldsymbol{x}^{\mathrm{base}_{128}(e)}$, where $\phi(e) := 1$ for all $0 \le e \le d$, and 0 otherwise. Note that, substituting $y_j = 3^{128^{j-1}}$ for all $j \in [2k]$ in $\tilde{P}_{k,n}$, we get $P_{k,n}$.

Given (k-tuples) $\mathrm{base}_{128}(i)$ and $\mathrm{base}_{128}(j)$, one can easily calculate whether $j = i^2$ or not, in $O(k^2)$ time. Hence, the function ϕ is in FP. As, FP $\subset$ #P/poly, therefore by Valiant's criterion, we have $(\tilde{P}_{k,n})_k \in$ VNP. As VNP is closed under substitution, we get $(P_{k,n})_k \in$ VNP as well. The *hard*ness part for $(P_{k,n})_k$ follows similarly as in the above proof. Thus, we do *not* need GRH for this particular polynomial!

2. The similar proof works for $\prod_{i\in[d]}(x \pm i)$. The hardness part remains the same. The only non-trivial part is to show that $(P_{k,n})_k \in$ VP, assuming GRH and VP = VNP. The polynomial family $\prod_{i\in[d]}(x \pm i)$ is CH-explicit (see Section C) and the rest follows similarly.

3. Recall the proof notation. As the degree of $\tilde{f}_i$'s is $\le 127k/3$, the degree of $\phi_{k,n}(\tilde{f}_i)$ is $\le 128^{k-1} \cdot 127k/3 = O(d \log d)$ ($\because k = \Theta(\log d)$). Thus, it suffices to study the representation of f_d as sum-of-4^{th}-powers ℓ_i^4, where $\deg(\ell_i) \le O(d \log d)$. Thus, Conjecture C1 even with this restriction, leads to the same conclusion as that in Theorem 1.

4. Note that Eqn.(5) gives a good approximation for the binomial $\binom{k+kn/3}{k}$, and the 1/3-rd degree factor in Theorem 1 is *critical* to make the combinatorial estimate work in the proof. We do not know how to *significantly* reduce the exponent 0.98289.

3.3 Conjecture C2 to blackbox-PIT $\in$ P : Proof of Theorem 3

Proof of Theorem 3. Let Conjecture C2 be true for some $0 < \delta_1 \le 1$, and $\delta_2 \ge 1$. Let k be a *constant* that will be specified later and $\boldsymbol{x} := (x_1, \ldots, x_k)$. For all large $n \in \mathbb{N}$, define $d := d(n)$ to be the *largest* element in I which is $\le (n+1)^k - 1$. Also, $d \ge (1/3) \cdot ((n+1)^k - 1)$ as the ratio of two consecutive elements in I is $(2^{m+1}-1)/(2^m - 1) < 3$, for $m \ge 2$.

Define the polynomial family $P_{k,n}(\boldsymbol{x}) := \psi_{k,d}(f_d)$ via the inverse Kronecker map applied to $f_d = (x+1)^d$. It is clear that $P_{k,n}$ is a k-variate polynomial with individual degree at most n, because the individual degree is bounded by $\lceil (d+1)^{1/k} \rceil - 1 \le n$. Thus, $\deg(P_{k,n}) \le kn$.

Note that $(P_{k,n})_n$ is an *explicit* family of polynomials because its coefficient vector $\mathrm{coef}(P_{k,n})$ can be computed in $\mathrm{poly}(d) = \mathrm{poly}(n)$ time. To see this, observe that for $\boldsymbol{e} = (e_1, \ldots, e_k)$, we have $\mathrm{coef}(\boldsymbol{x}^{\boldsymbol{e}})(P_{k,n}) = \binom{d}{e}$, where $e := \sum_{i=1}^{k} e_i (n+1)^{i-1}$. Also, the number of monomials in $P_{k,n}$ is $|\mathrm{supp}(P_{k,n})| = |\mathrm{supp}(f_d)| = d+1$.

Next we will show the hardness of the polynomial family $(P_{k,n})_n$. Let $\mu := 6/(\delta_1 - 14/k)$. We want $\mu > 0$. This enforces a condition on k, namely $k > 14/\delta_1$.

Claim 13 (Hardness of $P_{k,n}$)**.** $\mathrm{size}(P_{k,n}) > d^{1/\mu}$, *for all large enough* n.

Proof of Claim 13. Assume to the contrary, that there exists an infinite subset $J \subset \mathbb{N}$ such that $\mathrm{size}(P_{k,n}) \le d^{1/\mu}$, for all $n \in J$. We will show that Conjecture C2 is false over an infinite subset $J' := \{d(n) : n \in J\} \subseteq I$; which is a contradiction.

Let C be a circuit of size $\le d^{1/\mu}$ that computes $P_{k,n}$, for some n. Then, using Theorem 1, we know that there exist $f_{ij} \in \mathbb{F}[\boldsymbol{x}]$, of degree at most $\deg(P_{k,n})/3 \le kn/3$, such that

$$P_{k,n} = \sum_{i=1}^{s'} f_{i1} \cdot f_{i2} \cdot f_{i3} \cdot f_{i4} \ ,$$

where $s' \le c \cdot \left(d^{1/\mu}\right)^6 \cdot (kn)^{13}$, for some constant c. We apply Fischer's formula on each $\prod_j f_{ij}$ to write $P_{k,n}$ as

$$P_{k,n} \;=\; \sum_{i\in[s_0]} c_i \cdot \tilde{f}_i^4\,, \tag{12}$$

where $s_0 \;:=\; 16\cdot s'$ and $\tilde{f}_i \in \mathrm{span}_{\mathbb{F}}\left(f_{jk} \mid j\in[s'], k\in[4]\right)$, for every i. Thus, $\deg(\tilde{f}_i) \le kn/3$. Applying the Kronecker map $\phi_{k,n}$ to $P_{k,n}$ yields

$$f_d \;=\; \phi_{k,n}(P_{k,n}) \;=\; \sum_{i=1}^{s_0} c_i \cdot \left(\phi_{k,n}(\tilde{f}_i)\right)^4\,.$$

Note that, a simple combinatorial argument shows that $\left|\bigcup_i \mathrm{supp}(\tilde{f}_i)\right| \le \binom{k+kn/3}{k} =: s_1$. Since Kronecker substitution cannot increase the support size, $\left|\bigcup_i \mathrm{supp}(\phi_{k,n}\left(\tilde{f}_i\right))\right| \le s_1$, and therefore $U_{\mathbb{F}}(f_d, s_0) \le s_1$.

We want to show that $s_0 < d^{\delta_1}$ and $s_1 < d/4^{\delta_2}$, for all large enough n. Then we have $U_{\mathbb{F}}(f_d, d^{\delta_1}) < d/4^{\delta_2}$, for all large $d \in J' \subseteq I$; which contradicts Conjecture C2.

Bound on s_0. We have for large enough n (and thus d),

$$\begin{aligned} s_0 \;=\; 16\cdot s' \;\le\; 16c\cdot\left(d^{1/\mu}\right)^6\cdot(kn)^{13} \;&\le\; (16c\cdot k^{13})\cdot d^{6/\mu}\cdot(3d)^{13/k} \\ &=\; (16c\cdot k^{13}\cdot 3^{13/k})\cdot d^{6/\mu+13/k} = c'\cdot d^{\delta_1-1/k} < d^{\delta_1} \end{aligned}$$

where $c' := 16c\cdot k^{13}\cdot 3^{13/k}$ is a constant. We used that $d \ge ((n+1)^k-1)/3 \ge n^k/3$ for $n\ge 1$, and $\delta_1 = 6/\mu + 14/k$.

Bound on s_1. Finally, we show that $s_1 < d/4^{\delta_2}$. By Eqn.(5), we have

$$s_1 \;=\; \binom{k+k\,n/3}{k} \;\le\; \left(e\,(1+n/3)\right)^k \;<\; (2.8/3)^k\cdot n^k \;\le\; \left(3\cdot(2.8/3)^k\right)\cdot d\,,$$

We used the fact that $e(1+n/3) < (2.8\,n/3) = (14n/15)$, for all large enough n and $d \ge n^k/3$.

Therefore, it suffices to show that $3\cdot(14/15)^k \le 1/4^{\delta_2}$. As, $\delta_2 \ge 1$, it is *enough* to choose $k\cdot\log(15/14) \ge \delta_2\log(12)$ [$\because \delta_2\log(12) \ge \log 3 + \delta_2\log 4$]. The last condition is satisfied if $k \ge 37\cdot\delta_2$ [$\because \log(12)/\log(15/14) \approx 36.02$]. Thus, from the above calculations, it is enough to pick $k > \max(37\delta_2, 14/\delta_1)$. This proves Claim 13. □

From hardness to HSG. We show that from the hardness of $P_{k,n}$ in Claim 13, we can fulfil the assumption in Theorem 28: $\mathrm{size}(P_{k,n}) > s^{10k+2}\deg(P_{k,n})^3$, for some 'growing' function $s = s(n)$. Recall that $\deg(P_{k,n}) \le kn$. We define, $s(n) \;:=\; n^{1/(10k+3)}$. Then we have

$$s^{10k+2}\,(kn)^3 \;=\; n^{(10k+2)/(10k+3)}\,(kn)^3 \;=\; k^3\,n^{4-(1/(10k+3))} \;<\; \frac{n^4}{3^{1/\mu}}\,. \tag{13}$$

For the last inequality note that k, μ are constants. So for large enough n, the inequality will hold. Additionally assume that $4 \le k/\mu$. Recall the fact: $n^k/3 \le d$ for $n \ge 1$. So, we can continue Eqn.(13) as

$$\frac{n^4}{3^{1/\mu}} \;\le\; \frac{n^{k/\mu}}{3^{1/\mu}} \;\le\; d^{1/\mu} \;<\; \mathrm{size}(P_{k,n})\,. \tag{14}$$

Equations (13) and (14) give the desired hardness of $P_{k,n}$. It remains to ensure the last requirement of $4 \le k/\mu$. We show below that $k \ge 38/\delta_1$ suffices:

$$\mu \;=\; 6/(\delta_1 - 14/k) \;\le\; 6/\left(\delta_1 - (14/38)\delta_1\right) \;=\; (6\times 38)/(24\delta_1) \;=\; 19/(2\delta_1) \;\le\; k/4\,.$$

Hence our final choice for k is: $k \ge \max(37\,\delta_2, 38/\delta_1)$.

Thus, Theorem 28 gives a $\mathrm{poly}(s)$-time HSG for $\mathcal{C}(s,s,s)$. Hence, blackbox-PIT $\in$ P. □

Remarks. 1. The same proof works for other polynomials like, $\prod_{i\in[d]}(x\pm i)$ or $\sum_{i=0}^{d} 3^{i^2}x^i$. The hardness-proof part does not change at all (assuming the corresponding Conjecture C2). Their explicitness is also clear as their coefficient vector is computable in $\mathrm{poly}(d)$-time. So, the corresponding $P_{k,n}$ will be k (=constant) variate and $\mathrm{poly}(n)$-time explicit.

2. Recall the proof notation. As the degree of $\tilde{f}_i$'s is $\leq kn/3$, the degree of $\phi_{k,n}(\tilde{f}_i)$ is $\leq (n+1)^{k-1} \cdot kn/3 < k/3 \cdot (n+1)^k \leq k/3 \cdot (3d+1) = O(d)$ ($\because k$ is a constant). Thus, it *suffices* to study the representation of f_d as sum-of-4^{th}-powers ℓ_i^4, where $\deg(\ell_i) \leq O(d)$.

4 Conclusion

This work effectively establishes that studying the univariate sum-of-r^{th}-powers representation, for any $r \geq 4$, suffices to both derandomize and prove hardness in algebraic complexity (see Section F).

Here are some immediate questions which require rigorous investigation (also see [DST20]).

1. Does the exponent 3, or 2, suffice to solve PIT, or VP $\neq$ VNP? In particular, can we cleverly define the frontier to improve Theorem 1?
2. Prove Conjectures C1-C3 for SOS (sum-of-squares) model. They relate to matrix rigidity [DST20, Thm.4].

   Here is a concrete structural conjecture to write f_d via a 'square of vectors': Pick $s, |S| = o(d)$ and $S \subseteq \{0, 1, 2, ..., O(d)\}$. There do *not* exist column-vectors $\mathbf{c}, \mathbf{u}_i \in \mathbb{Q}^s$, s.t. $(x+1)^d = \mathbf{c}^T . \left(\sum_{i \in S} \mathbf{u}_i \cdot x^i\right)^2$.
3. Prove Conjecture C1 for *sum of constant* many 4^{th}-powers.
4. Prove Conjecture C2 for a 'generic' polynomial f with rational coefficients ($\mathbb{Q}$). Does it fail when we move to *complex* coefficients ($\mathbb{C}$)?
5. Prove $S_{\mathbb{Z}}((x+1)^d) \geq \Omega(d)$, for *all* large enough d (i.e. for the ones outside I too).
6. Remove the GRH assumption for the polynomial $(x+1)^d$ (in Theorem 2).
7. Does proving bounds like $\Omega(d^{1/2})$ or $\Omega(d^{1/3})$ on $S_{\mathbb{F}}(f_d)$ *suffice* to conclude Theorem 2? Note that, the trivial lower bound is $\Omega(d^{1/4})$. What's a natural barrier on the lower bound?

Acknowledgments. We thank Meena Mahajan for the useful discussions which motivated us to improve the exponent significantly. P.D. thanks CSE, IIT Kanpur for the hospitality, and acknowledges the support of Google PhD Fellowship. N.S. thanks the funding support from DST (DST/SJF/MSA-01/2013-14) and N. Rama Rao Chair.

References

[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). 4, 5, 7

[AJMV98] Eric Allender, Jia Jiao, Meena Mahajan, and V. Vinay. Non-Commutative Arithmetic Circuits: Depth Reduction and Size Lower Bounds. *Theoretical Computer Science*, 209(1-2):47–86, 1998. 7

[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. 1, 3, 7

[BCS13] Peter Bürgisser, Michael Clausen, and Amin Shokrollahi. *Algebraic Complexity Theory*, volume 315. Springer Science & Business Media, 2013. 1

[BDD86] Ramachandran Balasubramanian, Jean-Marc Deshouillers, and François Dress. Probleme de Waring pour les bicarrés. I. Schéma de la solution. *CR Acad. Sci. Paris Sér. I Math*, 303(4):85–88, 1986. 6

[Bel58] Edouard Grigor'evich Belaga. Some problems involved in the calculation of polynomials. In *Doklady Akademii Nauk*, volume 123, pages 775–777. Russian Academy of Sciences, 1958. 6

[BT15] Grigoriy Blekherman and Zach Teitler. On maximum, typical and generic ranks. *Mathematische Annalen*, 362(3-4):1021–1031, 2015. 6

[Bür00] Peter Bürgisser. Cook's versus Valiant's hypothesis. *Theoretical Computer Science*, 235(1):71–88, 2000. 8

[Bür09] Peter Bürgisser. On Defining Integers and Proving Arithmetic Circuit Lower Bounds. *Computational Complexity*, 18(1):81–103, 2009. 2, 7, 23

[Bür13] Peter Bürgisser. *Completeness and Reduction in Algebraic Complexity Theory*, volume 7. Springer Science & Business Media, 2013. 6, 8, 23

[CCG12] Enrico Carlini, Maria Virginia Catalisano, and Anthony V Geramita. The solution to the Waring problem for monomials and the sum of coprime monomials. *Journal of algebra*, 370:5–14, 2012. 6

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

[DST20] Pranjal Dutta, Nitin Saxena, and Thomas Thierauf. Lower bounds on the sum of 25^{th}-powers of univariates lead to complete derandomization of PIT. *Electronic Colloquium on Computational Complexity (ECCC link)*, 39, 2020. 2, 3, 4, 5, 8, 15, 19, 23, 26, 27, 28

[Fis94] Ismor Fischer. Sums of like Powers of Multivariate Linear Forms. *Mathematics Magazine*, 67(1):59–61, 1994. 23

[FOS12] Ralf Fröberg, Giorgio Ottaviani, and Boris Shapiro. On the Waring problem for polynomial rings. *Proceedings of the National Academy of Sciences*, 109(15):5600–5602, 2012. 6

[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. 1, 3

[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. 4, 5, 6, 7, 24

[GMK17] Ignacio Garcia-Marco and Pascal Koiran. Lower bounds by Birkhoff interpolation. *Journal of Complexity*, 39:38–50, 2017. 7

[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 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 34:1–34:14, 2016. 7

[Gro15] Joshua A Grochow. Unifying known lower bounds via geometric complexity theory. *Computational Complexity*, 24(2):393–475, 2015. 7

[HS80] Joos Heintz and Malte Sieveking. Lower bounds for polynomials with algebraic coefficients. *Theoretical Computer Science*, 11(3):321–330, 1980. 6

[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. 4

[Joh90] David S Johnson. A catalog of complexity classes. In *Algorithms and complexity*, pages 67–161. Elsevier, 1990. 22

[Kem12] Aubrey Kempner. Bemerkungen zum Waringschen Problem. *Mathematische Annalen*, 72(3):387–399, 1912. 6

[KI03] Valentine Kabanets and Russell Impagliazzo. Derandomizing polynomial identity tests means proving circuit lower bounds. In *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*, pages 355–364. ACM, 2003. 4, 5

[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. 6

[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. 2, 7

[Koi12] Pascal Koiran. Arithmetic circuits: The chasm at depth four gets wider. *Theoretical Computer Science*, 448:56–65, 2012. 1, 3, 7

[KPGM18] Pascal Koiran, Timothée Pecatte, and Ignacio Garcia-Marco. On the linear independence of shifted powers. *Journal of Complexity*, 45:67–82, 2018. 7

[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. 8

[KS19] Mrinal Kumar and Ramprasad Saptharishi. Hardness-Randomness Tradeoffs for Algebraic Computation. *Bulletin of EATCS*, 1(129), 2019. 7, 24

[KSS19] Mrinal Kumar, Ramprasad Saptharishi, and Noam Solomon. Derandomization from Algebraic Hardness: Treading the Borders. `https://arxiv.org/pdf/1905.00091v1.pdf`, 2019. 25

[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. 7

[Mah14] Meena Mahajan. Algebraic Complexity Classes. In *Perspectives in Computational Complexity*, pages 51–75. Springer, 2014. 1

[Mot55] T.S. Motzkin. Evaluation of polynomials and evaluation of rational functions. *Bull. Amer. Math. Soc*, 61(163):10, 1955. 6

[Muk16] Partha Mukhopadhyay. Depth-4 identity testing and Noether's normalization lemma. In *International Computer Science Symposium in Russia*, pages 309–323. Springer, 2016. 7

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

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

[Sap19] Ramprasad Saptharishi. A survey of lower bounds in arithmetic circuit complexity. Github survey, 2019. 2, 3, 7, 8

[Sax09] Nitin Saxena. Progress on Polynomial Identity Testing. *Bulletin of the EATCS*, 99:49–79, 2009. 7

[Sax14] Nitin Saxena. Progress on Polynomial Identity Testing - II. *Perspectives in Computational Complexity*, 26:131–146, 2014. 7

[Sch80] J. T. Schwartz. Fast Probabilistic Algorithms for Verification of Polynomial Identities. *J. ACM*, 27(4):701–717, October 1980. 7

[Str74] Volker Strassen. Polynomials with rational coefficients which are hard to compute. *SIAM Journal on Computing*, 3(2):128–149, 1974. 6

[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. 1, 3, 7, 11

[Tav15] Sébastien Tavenas. Improved bounds for reduction to depth 4 and depth 3. *Information and Computation*, 240:2–11, 2015. 1, 3, 7

[Val79a] Leslie G. Valiant. Completeness Classes in Algebra. In *Proceedings of the 11h Annual ACM Symposium on Theory of Computing, 1979*, pages 249–261, 1979. 1

[Val79b] Leslie G Valiant. Completeness classes in algebra. In *Proceedings of the 11th Annual ACM symposium on Theory of computing*, pages 249–261. ACM, 1979. 8, 23

[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. 2, 3, 5, 7

[Wag86] Klaus W Wagner. The complexity of combinatorial problems with succinct input representation. *Acta informatica*, 23(3):325–356, 1986. 22

[Wig17] Avi Wigderson. Low-depth arithmetic circuits: technical perspective. *Communications of the ACM*, 60(6):91–92, 2017. 7

[Wik] Wikipedia. Binomial coefficient– bounds and asymptotic formulas. https://en.wikipedia.org/wiki/Binomial_coefficient#Bounds_and_asymptotic_formulas. 7

[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. 7

A Bounds for the sum of 4^{th}-powers: Details for Section 1

This section reminds some of the motivating observations made in [DST20]. We show some upper/lower bounds for univariate polynomials represented as sum of 4^{th}-powers. We will show *two* different ways to write *any* f as sum of 4^{th}-power of polynomials over $\mathbb{F}$ of characteristic 0 or large; thus, establishing that SO4 is a complete model.

A.1 Sum of $r+1$ many r-th powers of polynomials– Upper bound

The first representation shows that any polynomial can be written as a sum of $(r+1)$-many r^{th}-powers of polynomials. [However, the support-union is as large as the support of the original polynomial.]

Lemma 14. *Let $\mathbb{F}$ be a field of characteristic 0 or $\geq r+1$. Let $\boldsymbol{x} = (x_1, \ldots, x_k)$ for some $k \in \mathbb{N}$ and $h(\boldsymbol{x}) \in \mathbb{F}[\boldsymbol{x}]$ with $0 \leq m \leq r$. There exist $c_{m,i} \in \mathbb{F}$ and* distinct *$\lambda_i \in \mathbb{F}$, for $0 \leq i \leq r$, such that*

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

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

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

Choose $r+1$ distinct λ_j's in $\mathbb{F}$, and put $t = \lambda_j$, for $j = 0, \ldots, r$. We get $r+1$ many linear equations which can be represented in matrix form $Av = b$, where matrix $A := \left(\binom{r}{i}\,\lambda_j^i\right)_{0\leq j,i\leq r}$, and vectors $v = \left(h^{r-i}\right)_{0\leq i\leq r}$ and $b = \left((h+\lambda_j)^r\right)_{0\leq j\leq r}$.

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

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

Therefore, A is invertible and $v = A^{-1}b$. Let the $(m+1)$-th row of A^{-1} be $[c_{m,0}\;\; c_{m,1}\;\; \cdots\;\; c_{m,r}]$. Then we have,

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

□

Using Lemma 14, we show an upper bound on $S_{\mathbb{F}}(f_d)$ and $U_{\mathbb{F}}(f_d, 5)$ for $f_d(x) = (x+1)^d$.

Lemma 15. *For $f_d(x) = (x+1)^d$, we have $S_{\mathbb{F}}(f_d) \leq 5 \cdot (d/4 + 4)$ and $U_{\mathbb{F}}(f_d, 5) \leq d/4 + 4$.*

Proof. Suppose $d =: 4 \cdot k + t$, where $0 \leq t \leq 3$ and $0 \leq k \leq d/4$. Then, from Lemma 14, it follows that there exists $c_i, \lambda_i \in \mathbb{F}$ such that

$$\begin{aligned}
(x+1)^d \;&=\; \left((x+1)^k\right)^4 \cdot (x+1)^t \\
&=\; \left((x+1)^k\right)^4 \cdot \left(\sum_{i=0}^{4} c_i \left((x+1)^t + \lambda_i\right)^4\right) \\
&=\; \sum_{i=0}^{4} c_i \left((x+1)^{t+k} + \lambda_i (x+1)^k\right)^4 \;=:\; \sum_{i=0}^{4} c_i \cdot \ell_i^4
\end{aligned}$$

where $\ell_i := (x+1)^{t+k} + \lambda_i(x+1)^k$. Note that, $|\ell_i|_1 = |\bigcup_{i=0}^{4} \mathrm{supp}(\ell_i)| \leq t+k+1 \leq d/4+4$. Thus, the respective bounds on $S_{\mathbb{F}}(f_d)$ and $U_{\mathbb{F}}(f_d, 5)$ follow. □

A.2 Sum of powers of 'small' support-union– Upper bound

The second representation is a bit more complicated than the first one. Here we use the notion of *sumset*. In additive combinatorics, the sumset (also called the Minkowski sum) of two subsets A and B of an abelian group G (written additively) is defined to be the set of all possible sums of an element from A with an element from B. That is,

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

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 d-degree polynomial f as a sum of 4^{th}-powers. We consider a *small* B such that $4B$ covers $\{0, \ldots, d\}$. In particular, we know that there exists a *unique* non-negative integer t such that $(t-1)^4 < d+1 \leq t^4$. Define the set B as,

$$B \ := \ \{a_i \cdot t^k \mid 0 \leq a_i \leq t-1, 0 \leq k \leq 3\} \ .$$

So, $|B| = 4 \cdot t = O(d^{1/4})$. Clearly, $4B \supseteq \{0, \ldots, d\}$ (using base-t representation). Note that the largest element in B is $m := (t-1) \cdot t^3$. For any $\epsilon > 0$, there exists d_0 such that $t < (1+\epsilon) \cdot (d+1)^{1/4}$, for all $d \geq d_0$. Thus, for *any* constant $c > 1$ and large enough d, we have $m < c \cdot (d+1)$. Therefore, the largest element in $4B$ is at most $4 \cdot m < 4 \cdot c \cdot (d+1) = O(d)$. We claim the following:

Lemma 16. *For any $f(x) \in \mathbb{F}[x]$ of degree d, where $\mathbb{F}$ is a field of characteristic 0 or large, there exists ℓ_i supported on B, and $c_i \in \mathbb{F}$ such that $f(x) \ = \ \sum_{i=0}^{4m} c_i \cdot \ell_i^4$.*

Proof. Consider $\ell_i(z_i, x) = \sum_{j \in B} z_{ij} x^j$, for distinct indeterminates z_{ij}, for all i, j. Surely, $\deg_x(\ell_i) = m$. There exists $4m+1$ many degree-4 polynomials $Q_j(z_i)$, over $|B| = 4t$ many variables, s.t.

$$\ell_i(z_i, x)^4 \ = \ \sum_{j=0}^{4m} Q_j(z_i) \cdot x^j \qquad \forall i \in [4m] \ .$$

Note that from any monomial in Q_j we could recover j uniquely. Thus, we could conclude that $Q_j(z_i)$ $(0 \leq j \leq 4m)$ are $\mathbb{F}$-linearly independent.

Suppose $f(x) =: \sum_{i=0}^{d} f_i x^i$. Define $\tilde{f} \in \mathbb{F}^{1 \times (4m+1)}$ and $A \in \mathbb{F}[z]^{(4m+1) \times (4m+1)}$ as,

$$\tilde{f} := \begin{bmatrix} f_0 & f_1 & \ldots & f_d & 0 & \ldots & 0 \end{bmatrix}, A := \begin{bmatrix} Q_0(z_0) & Q_1(z_0) & \ldots & Q_{4m}(z_0) \\ Q_0(z_1) & Q_1(z_1) & \ldots & Q_{4m}(z_1) \\ \vdots & \vdots & \ldots & \vdots \\ Q_0(z_{4m}) & Q_1(z_{4m}) & \ldots & Q_{4m}(z_{4m}) \end{bmatrix} .$$

We want to find $\bar{c} =: \begin{bmatrix} c_0 & c_1 & \ldots & c_{4m} \end{bmatrix} \in \mathbb{F}^{1 \times (4m+1)}$ and $\boldsymbol{\alpha} = (\alpha_{ij})_{i,j}$ such that

$$\sum_{i=0}^{4m} c_i \cdot \ell_i(\boldsymbol{\alpha}, x)^r \ = \ \sum_{i=0}^{d} f_i \cdot x^i$$

$$\iff \ \bar{c} \cdot A|_{z=\boldsymbol{\alpha}} \cdot \begin{bmatrix} 1 \\ x \\ \vdots \\ x^{4m} \end{bmatrix} \ = \ \tilde{f} \cdot \begin{bmatrix} 1 \\ x \\ \vdots \\ x^{4m} \end{bmatrix}$$

$$\iff \ \bar{c} \cdot A|_{z=\boldsymbol{\alpha}} = \tilde{f} \ .$$

As z_i are distinct variables, first column of A consists of different variables at each coordinate. Moreover, first row of A contains $\mathbb{F}$-linearly independent Q_j's. Thus, for *random* $\alpha_{ij} \in \mathbb{F}$, $A|_{z=\boldsymbol{\alpha}}$ is *full* rank over $\mathbb{F}$. Fix such an $\boldsymbol{\alpha}$. This fixes $\bar{c} = \tilde{f} \cdot (A|_{z=\boldsymbol{\alpha}})^{-1}$.

From the above construction, it follows that $f(x) \ = \ \sum_{i=0}^{4m} c_i \cdot \ell_i(\boldsymbol{\alpha}, x)^4$. □

Remarks. 1. Thus, for any d-degree f, $U_{\mathbb{F}}(f, s := 4m+1) \leq O(d^{1/4})$. As seen before, $4m = \Theta(d)$. When $s \geq c \cdot (d+1)$ for $c > 4$, we have a small base representation for large enough d. It is unclear, though, whether even for $s \leq d$, such a small support-union representation exists.

2. Unfortunately, the above calculation does *not* give 'small' sparsity-sum representation of f: the top fanin is already $\Omega(d)$ and each $|\ell_i|_1 = O(d^{1/4})$; giving $\sum |\ell_i|_1 = O(d^{5/4})$. This is *worse* than the bound on $S_{\mathbb{F}}(f_d)$, derived from Lemma 15.

3. Both the above representations (small top-fanin s resp. small support-union) crucially require a *field* $\mathbb{F}$. E.g. they do not exist for f_d over the ring $\mathbb{Z}$ (hint: go modulo 2).

A.3 $(x+1)^d$ as sum of two 4^{th}-powers– Lower bound

We show a strong lower bound of $\Omega(d)$ for $f_d(x) := (x+1)^d$ when expressed as a sum of *two* 4^{th}-powers. Wlog, we consider $\mathbb{F}$ algebraically closed, as $S_{\mathbb{F}}(\cdot) \leq S_{\mathbb{F}}(\cdot)$ and $U_{\mathbb{F}}(\cdot) \leq U_{\mathbb{F}}(\cdot)$. Note that, $c_1 \cdot \ell_1^4 + c_2 \cdot \ell_2^4 = \tilde{\ell}_1^4 - \tilde{\ell}_2^4$ where $\tilde{\ell}_1 = c_1^{1/4} \cdot \ell_1$ and $\tilde{\ell}_2 = (-c_2)^{1/4} \cdot \ell_2$. Also, $|\bigcup_{i=1}^2 \mathrm{supp}(\ell_i)| = |\bigcup_{i=1}^2 \mathrm{supp}(\tilde{\ell}_i)|$. Thus, it suffices to prove the bounds when f_d is written as $\ell_1^4 - \ell_2^4$. Before that, we prove the following.

Lemma 17. *For a fixed $d \geq 1$, if $(x+1)^d = \ell_1^4 - \ell_2^4$, for some $\ell_i \in \mathbb{F}[x]$, then ℓ_1 and ℓ_2 must share a non-trivial gcd.*

Proof. Suppose, $\gcd(\ell_1, \ell_2) = 1$. Note that, $\ell_1^4 - \ell_2^4$ has the following factorization over $\mathbb{F}[x]$,

$$(x+1)^d \;=\; (\ell_1 - \ell_2)\,(\ell_1 - \zeta_4\,\ell_2)\,(\ell_1 - \zeta_4^2\,\ell_2)\,(\ell_1 - \zeta_4^3\,\ell_2)$$

where $\zeta_4 := \sqrt{-1}$, primitive 4-th root of unity. If $(x+1) \mid (\ell_1 - \zeta_4^i\,\ell_2)$ and $(\ell_1 - \zeta_4^j\,\ell_2)$, for $i \neq j$, then subtraction would imply: $(x+1) \mid \ell_1, \ell_2$. This contradicts our assumption; hence, there must exist i: $\ell_1 - \zeta_4^i\,\ell_2 \;=\; c \cdot (x+1)^d$. In particular, it means: $\ell_1 - \zeta_4^j\,\ell_2$ is constant, for all $j \neq i$. Subtracting two such equations immediately gives us: ℓ_1, ℓ_2 are constants; a contradiction again as $d \geq 1$. □

Theorem 18. *For any $d \geq 1$, we have*

$$U_{\mathbb{F}}(f_d, 2) = \begin{cases} d/4 + 1 & \text{if } 4 \mid d\,, \\ \infty & \text{otherwise.} \end{cases}$$

Proof. We prove the following claim.

Claim 19. *For $1 \leq d$: $(x+1)^d \;=\; \ell_1^4 - \ell_2^4$ iff $4 \mid d$. In that case, $\exists \alpha_1, \alpha_2 \in \mathbb{F}$ such that $\ell_i = \alpha_i \cdot (x+1)^{d/4}$*

Proof. Assume $(x+1)^d = \ell_1^4 - \ell_2^4$. By Lemma 17, $\gcd(\ell_1, \ell_2) =: p(x)$ is non-constant. Therefore, $p^4 \mid (x+1)^d$; implying that $p(x)$ is a power of $x+1$. After dividing out, we can again apply the lemma. Eventually, we deduce: $4 \mid d$, and $\ell_i = \alpha_i \cdot (x+1)^{d/4}$, for some $\alpha_i \in \mathbb{F}$. □

The theorem follows directly from the claim. □

B Decomposition as a sum of product of 3: Details for Section 3.1

Lemma 10 (restated). *Let $f(\boldsymbol{x})$ be an n-variate, homogeneous, degree d polynomial computed by a right-heavy homogeneous circuit of size s. Then, $\exists f_{ij} \in \mathbb{F}[\boldsymbol{x}]$ such that (for $s_1 := O(s^2)$),*

$$f(\boldsymbol{x}) \;=\; \sum_{i=1}^{s_1} f_{i1} \cdot f_{i2} \cdot f_{i3}\,, \quad \textit{with the following properties:} \tag{15}$$

1. $deg(f_{i1}) \le deg(f_{i2}) \le deg(f_{i3}) \le d/2$, *for all* $i \in [s_1]$.
2. $deg(f_{i1}) \le d/3 \le deg(f_{i3})$ *and* $deg(f_{i1}) + deg(f_{i2}) + deg(f_{i3}) = d$, *for all* $i \in [s_1]$.
3. *each* f_{ij} *has right-heavy homogeneous circuit of size at most* $s_3 = O(s)$.

Proof of Lemma 10. Invoke Lemma 9 to conclude that $f(\boldsymbol{x})$ has a decomposition of the form Eqn.(6): Let $d_{i1} := \deg(f_{i1})$ and $d_{i2} := \deg(f_{i2})$, for all $i \in [s]$. From Lemma 9, we have $d/3 \le d_{i1} \le d/2 \le d_{i2} \le 2d/3$. Next, we further decompose each f_{i2} as sum of product of 2 polynomials using Lemma 9, as follows (for $s_2 = O(s)$):

$$f_{i2} \;=\; \sum_{j=1}^{s_2} f_{i2,j1} \cdot f_{i2,j2}\;, \quad \text{with the following properties:} \tag{16}$$

(1) $d_{i2}/3 \;\le\; \deg(f_{i2,j1}) \;\le\; d_{i2}/2 \;\le\; (1/2)\cdot(2d/3) \;=\; d/3$, for all $j \in [s_2]$.
(2) $d_{i2}/2 \;\le\; \deg(f_{i2,j2}) \;\le\; 2d_{i2}/3 \;\le\; (2/3)\cdot(2d/3) \;=\; 4d/9 \;<\; d/2$, for all $j \in [s_2]$.
(3) $\deg(f_{i2,j1}) + \deg(f_{i2,j2}) \;=\; d_{i2}$ and $\deg(f_{i1}) + \deg(f_{i2,j1}) + \deg(f_{i2,j2}) \;=\; d$, for all $j \in [s_2]$.

Plug Eqn.(16) into Eqn.(6) to get a decomposition as a sum of products of 3 (*unordered degree*):

$$f(\boldsymbol{x}) \;=\; \sum_{i\in[s],\, j\in[s_2]} f_{i1} \cdot f_{i2,j1} \cdot f_{i2,j2}\;,$$

where the top fanin is $s_1 := s \cdot s_2 = O(s^2)$. By the above calculation, we have

$$\min\,\big(\deg(f_{i1}), \deg(f_{i2,j1}), \deg(f_{i2,j2})\big) \;\le\; \deg(f_{i2,j1}) \;\le\; d/3\,, \quad \text{and}$$

$$d/3 \;\le\; \deg(f_{i1}) \;\le\; \max\,\big(\deg(f_{i1}), \deg(f_{i2,j1}), \deg(f_{i2,j2})\big) \;\le\; d/2$$

with $\deg(f_{i1}) + \deg(f_{i2,j1}) + \deg(f_{i2,j2}) \;=\; d$. After relabelling the factor polynomials (with degree in increasing order), we get Eqn.(15) with the desired properties.

Size analysis. By Lemma 9, each f_{i1}, f_{i2} has right-heavy homogeneous circuit of size at most $s_2 = O(s)$. Applying the same argument on f_{i2}, we can conclude that each $f_{i2,j1}$, $f_{i2,j2}$, for $j \in [s_2]$ has right-heavy homogeneous circuit of size at most $s_3 = O(s_2) = O(s)$. □

C Primer on complexity classes: Details for Section 3.2

C.1 Complexity classes

The counting hierarchy (CH) was first introduced in [Wag86]. It can be defined by a counting operator **C** that can be applied to complexity classes. We denote by $\{0,1\}^* \times \{0,1\}^* \to \{0,1\}^*, (x,y) \mapsto \langle x,y\rangle$, a pairing function (e.g. by duplicating each bit of x and y and inserting 01 in between).

Definition 20. *If K is a complexity class, then we define an operator* **C** *acting on K. The action, denoted by* $\mathbf{C}\cdot K$*, produces a set of languages A, such that there exists a language $B \in K$ and a polynomial $p(\cdot)$, obeying :*

$$x \in A \iff \#\Big\{y \in \{0,1\}^{p(|x|)} : \langle x,y\rangle \in B\Big\} \;>\; \frac{1}{2}\cdot 2^{p(|x|)}\,.$$

The i-th level $\mathsf{C}_i\mathsf{P}$ of the counting hierarchy is defined recursively as $\mathsf{C}_0\mathsf{P} := \mathsf{P}$ and $\mathsf{C}_i\mathsf{P} = \mathbf{C}\cdot\mathsf{C}_{i-1}\mathsf{P}$. Finally, we define the counting hierarchy*: $\mathsf{CH} := \bigcup_{i\ge 0}\mathsf{C}_i\mathsf{P}$. Often, $\mathsf{PP} := \mathsf{C}_1\mathsf{P}$ is used in the literature. Observe that $\mathsf{C}_2\mathsf{P} = \mathsf{PP}^{\mathsf{PP}}$.*

Let us recall the definition of other complexity classes. For a survey of complexity classes, see [Joh90].

- FP denotes the class of all string *functions* which can be computed by a polynomial time Turing machine.
- A *polynomial advice* is a function $\alpha : \mathbb{N} \to \{0,1\}^*$ such that $n \mapsto \alpha(n)$ is poly-bounded. The (non-uniform) class C/poly for a complexity class C consists of all string functions of the form $\psi(x) =: \phi(\langle x, \alpha(|x|)\rangle)$, where $\phi \in C$ and α is some polynomial advice function.

C.2 GRH, VP vs VNP and the CH collapse

The proof of Theorem 2 requires CH explicitness of certain families. Here is the formal definition.

Definition 21. *A sequence $a = (a(n,k))_{n\in\mathbb{N},\, k\le 2^{p(n)}}$ of integers of exponential-bitsize is said to be* CH-definable *iff Sgn(a) and Bit(|a|) $\in$ CH, where (think of inputs n in unary & k in binary)*

$$Sgn(a) := \{(1^n, k) \mid a(n,k) \geq 0\}$$
$$Bit(|a|) := \{(1^n, k, j, b) \mid \text{the } j\text{-th bit of } a(n,k) \text{ equals } b\}\ .$$

It was established that the family $\prod_{i\in[d]}(x+i)$ is CH-explicit, for details see [Bür09, Cor.3.12]. A similar proof was used in [DST20, Thm. 31] to show that $(x+1)^d$ is CH-explicit as well.

Theorem 22 ([DST20])**.** *The sequence $a = a(n,k) := \binom{d}{k}$, where $d \le 2^{p(n)}$ for some polynomial $p(\cdot)$ (d given in binary) and $k \le d$, is definable in* CH.

A useful *sufficient* condition for a polynomial family $(f_n(\boldsymbol{x}))_n$ to be in VNP is known, due to Valiant [Val79b]. For a proof, see [Bür13].

Theorem 23 (Valiant's criterion, [Val79b])**.** *Suppose $\phi : \{0,1\}^* \to \mathbb{N}$ is a function in the class $\#\mathsf{P}/\mathsf{poly}$. Then, the family $(f_n)_n$ of polynomials defined by $f_n(\boldsymbol{x}) := \sum_{\boldsymbol{e}\in\{0,1\}^n} \phi(\boldsymbol{e}) \cdot \boldsymbol{x}^{\boldsymbol{e}} \in$* VNP.

Theorem 8 (restated). *Let function $\phi : [0,c]^* \to \mathbb{N}$ be in $\#\mathsf{P}/\mathsf{poly}$, for some constant $c \in \mathbb{N}$. Then, the family of polynomials defined by $f_n(\boldsymbol{x}) := \sum_{\boldsymbol{e}\in[0,c]^n} \phi(\boldsymbol{e}) \cdot \boldsymbol{x}^{\boldsymbol{e}}$, is in* VNP.

Proof sketch. The trick is to introduce y_{ij}, for every $i \in [n]$, $j \in [c']$, where $c' := \lceil \log(c+1) \rceil$ is a constant. Define

$$\tilde{f}_n(\boldsymbol{y}) := \sum_{e\in\{0,1\}^{c'\cdot n}} \phi'(\boldsymbol{e}) \cdot \boldsymbol{y}^{e}$$

where $\phi(e_1, \ldots, e_n) =: \phi'(\mathrm{bin}(e_1), \ldots, \mathrm{bin}(e_n))$. Note that under the substitution, $\psi : y_{ij} \mapsto x_i^{2^{j-1}}$, $\tilde{f}_n(\boldsymbol{y})$ becomes $f_n(\boldsymbol{x})$. By the previous theorem, $\tilde{f}_n(\boldsymbol{y}) \in$ VNP. As VNP is closed under polynomial substitution, $f_n(\boldsymbol{x}) \in$ VNP as well. □

VP and VNP have several closure properties. In particular, they are closed under substitution. That is, for a polynomial $f(\boldsymbol{x}, \boldsymbol{y}) \in$ VP (or VNP), also $f(\boldsymbol{x}, \boldsymbol{y}_0) \in$ VP (resp. VNP), for any values $\boldsymbol{y}_0$ from $\mathbb{F}$ assigned to the variables in $\boldsymbol{y}$. In fact, one can assign $y = g(\boldsymbol{x})$ where $g(\boldsymbol{x}) \in$ VP.

Fischer's formula. By a formula due to Fischer [Fis94] one can write any monomial as an exponential sum of powers. It requires $\mathrm{char}\,\mathbb{F} = 0$ or large. Also, it fails over $\mathbb{Z}$.

Lemma 24. [Fis94] *Let $\mathbb{F}$ be a field of characteristic 0 or ≥ 5. Any expression of the form $g = g_1 \cdot g_2 \cdot g_3 \cdot g_4$ can be written as $g = \sum_{j=1}^{16} c_j h_j^4$, where $c_j \in \mathbb{F}$ and $h_j \in \mathrm{span}_{\mathbb{F}}(g_i \mid i \in [4])$, for $j \in [16]$.*

D Hardness to derandomization: Details for Section 3.3

Very recently, Guo et al. in [GKSS19] showed utility of the hardness of *constant* variate polynomials to derandomize PIT. To make this discussion formal, we start with the following definition.

Definition 25 (Hitting-set generator (HSG))**.** *A polynomial map $G : \mathbb{F}^k \longrightarrow \mathbb{F}^n$ given by $G(\boldsymbol{z}) = (g_1(\boldsymbol{z}), g_2(\boldsymbol{z}), \ldots, g_n(\boldsymbol{z}))$ is said to be a* hitting-set generator *(HSG) for a class $\mathcal{C} \subseteq \mathbb{F}[x_1, x_2, \ldots, x_n]$ 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.*

Remark. We say that G is *t-time* HSG if $\mathrm{coef}(g_i)$ can be computed in t-time and maximum degree of g_i is also at most t. This gives $(t \cdot d)^{O(k)}$ time blackbox-PIT algorithm, for circuits computed by 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.

Given an HSG, it can be seen that there is a corresponding *hitting-set* H such that one needs to *only* query the given input circuit at points on H to determine non-zeroness.

Guo et al. in [GKSS19] came up with the following HSG connection.

Definition 26. *For $P(\boldsymbol{z}) \in \mathbb{F}[\boldsymbol{z}]$, define the map $\mathtt{Gen}_P^2 : \mathbb{F}^k \times \mathbb{F}^k \to \mathbb{F}^{n+1}$ such that $\mathtt{Gen}_P^2 =: (\Delta_0(P)(\boldsymbol{z}, \boldsymbol{y}), \ldots, \Delta_n(P)(\boldsymbol{z}, \boldsymbol{y}))$, where $\Delta_i(P)$ is the homogeneous degree i (in $\boldsymbol{y}$) component in the Taylor expansion of $P(\boldsymbol{z} + \boldsymbol{y})$, i.e. $\Delta_i(P)(\boldsymbol{z}, \boldsymbol{y}) = \sum_{\boldsymbol{e} \in S_i} \frac{\boldsymbol{y}^{\boldsymbol{e}}}{\boldsymbol{e}!} \cdot \frac{\partial^i P}{\partial \boldsymbol{z}^{\boldsymbol{e}}}$, where $S_i \subset \mathbb{N}^k$ such that for any $\boldsymbol{e} \in S_i$, $|\boldsymbol{e}|_1 = i$.*

The following theorem says that for a sufficiently hard and explicit P, $\mathtt{Gen}_P^2$ is an efficient HSG. It requires $k \geq 4$.

Theorem 27. [GKSS19, KS19] *Let P be a k-variate polynomial of degree d in $\mathbb{F}[\boldsymbol{x}]$. Suppose P cannot be computed by algebraic circuits of size $\tilde{s} = s \cdot D \cdot d^3 \cdot n^{10k}$ for parameters n, D, s. Then, for any $C \in \mathcal{C}(n+1, D, s)$, we have $C \neq 0 \iff C \circ \mathtt{Gen}_P^2 \neq 0$.*

We use the same theorem with the parameters $n + 1 = D = s$ to get PIT for $\mathcal{C}(s, s, s)$.

Theorem 28. [GKSS19] *Let $P \in \mathbb{F}[\boldsymbol{x}]$ be a k-variate polynomial of degree d such that $\mathrm{coef}(P)$ can be computed in $\mathsf{poly}(d)$-time. If $\mathrm{size}(P) > s^{10k+2} \cdot d^3$, then there is a $\mathsf{poly}(s)$-time HSG for $\mathcal{C}(s, s, s)$.*

E Approximative version of Conjecture C2 to Hitting-set for $\overline{\mathsf{VP}}$

Here we study hitting-set for the approximative class $\overline{\mathsf{VP}}$. Before doing that, it is important to recall the meaning of approximation in the algebraic setting.

Definition 29 (Approximative computation)**.** *A circuit C over $\mathbb{F}(\epsilon)[\boldsymbol{x}]$ is said to* approximate *a polynomial $P(\boldsymbol{x})$ if the output of the circuit C is a polynomial in $\mathbb{F}[\boldsymbol{x}, \epsilon]$ such that $C(\boldsymbol{x}, \epsilon) =: \epsilon^M \cdot P(\boldsymbol{x}) + \epsilon^{M+1} \cdot Q(\boldsymbol{x}, \epsilon)$, for some polynomial $Q(\boldsymbol{x}, \epsilon) \in \mathbb{F}[\boldsymbol{x}, \epsilon]$ and $M \in \mathbb{N}_{\geq 0}$. In other words,*

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

We denote by $\overline{\mathrm{size}}(P)$, the *approximative circuit complexity* of P, to be the size of the smallest circuit approximating P. The class $\overline{\mathsf{VP}}$ contains the families of n-variate polynomials, of degree $n^{O(1)}$, over $\mathbb{F}$, of approximative complexity $n^{O(1)}$.

Note: Equivalently, one could think of P as being 'approximated' by the circuit $C(\boldsymbol{x}, \epsilon)/\epsilon^M$ over the function field $\mathbb{F}(\epsilon)$. $\overline{\mathsf{VP}}$ could potentially be larger than VP, because the degree wrt ϵ (i.e. M above) could be *exponentially*-larger than n.

E.1 Tools for $\overline{\mathsf{VP}}$

We point out that the CNF Theorem (Theorem 1) works for approximative circuits as well.

Theorem 30. *Suppose $f(\boldsymbol{x}) \in \mathbb{F}[\boldsymbol{x}]$ is a polynomial of degree d which can be approximated by a size-s circuit C. Then, there exist polynomials $f_{ij} \in \mathbb{F}(\epsilon)[\boldsymbol{x}]$ such that*

$$C(\boldsymbol{x}) \;=\; \sum_{i=1}^{s'} f_{i1} \cdot f_{i2} \cdot f_{i3} \cdot f_{i4}\;, \tag{17}$$

for $s' = \mathsf{poly}(s,d)$, where each f_{ij} has circuit size (over $\mathbb{F}(\epsilon)$) at most $s'' = \mathsf{poly}(s,d)$ and $deg(f_{ij}) \leq d/3$, for all i,j. (Wherein $deg(\cdot)$ is wrt $\boldsymbol{x}$.)

Proof sketch. Essentially the same proof (of Theorem 1 in Section 3.1) works. One needs to consider $C \in \mathbb{F}(\epsilon)[\boldsymbol{x}]$ and realize that the earlier proof is field *independent*; so run it over $\mathbb{F}(\epsilon)$. □

Kumar et al. in [KSS19] proved that the hardness of constant-variate polynomials in the approximative sense, suffices to construct an HSG for $\overline{\mathsf{VP}}$ using the generator $\mathtt{Gen}_P^2$ (see Definition 26).

Theorem 31. [KSS19, Thm.1.6] *Let P be a k-variate polynomial of degree d in $\mathbb{F}[\boldsymbol{x}]$. Suppose $\tilde{s} := \overline{\mathsf{size}}(P) > s \cdot D \cdot d \cdot n^{10k}$ for parameters n, D, s. Then, for any $(n+1)$-variate polynomial $Q(x_0,\ldots,x_n)$ of degree D such that $\overline{\mathsf{size}}(Q) \leq s$, we have $Q \neq 0 \iff Q \circ \mathtt{Gen}_P^2 \neq 0$.*

E.2 Hitting-set for $\overline{\mathsf{VP}}$: Approximative version of Conjecture C2 and Theorem 1

Let field $\mathbb{F}$ be $\mathbb{Q}, \mathbb{Q}_p$ (or their fixed extensions), or a finite field of large characteristic. Let us first formalize Conjecture C2 in the approximative setting. For a ring R, we define *support-union approximative size* $\overline{U}_R(f,s)$ as the number of distinct monomials (in $\boldsymbol{x}$) required to *approximate* f as sum-of-4^{th}-powers. In particular, define

$$\overline{U}_R(f,s) \;:=\; \min\left(\left| \bigcup_{i=1}^{s} \mathrm{supp}(\ell_i) \right| \;:\; g(x,\epsilon) = \sum_{i=1}^{s} c_i \cdot \ell_i^4 \;\text{ and }\; \lim_{\epsilon\to 0} \frac{1}{\epsilon^M} \cdot g = f \right)\;.$$

Obviously, $\overline{U}_R(\cdot) \leq U_R(\cdot)$. We conjecture that even $\overline{U}_{\mathbb{F}}(f_d,s)$ is large for $f_d := (x+1)^d$.

Conjecture 3 (C3). *There exist positive constants $\delta_1 \leq 1, \delta_2 \geq 1$ such that $\overline{U}_{\mathbb{F}}\left(f_d, d^{\delta_1}\right) \geq d/4^{\delta_2}$, for all large enough $d \in I$.*

Theorem 32. *If Conjecture C3 holds true, then there is a* poly*-time* HSG *for $\overline{\mathsf{VP}}$-circuits.*

Proof sketch. The proof is almost the same as that of Theorem 3. We define $P_{k,n}$ similarly (i.e. inverse Kronecker applied on f_d where d was chosen uniquely from an interval based on n). We claim that $\overline{\mathsf{size}}(P_{k,n}) > d^{1/\mu}$, where $\mu \geq 6/(\delta_1 - 14/k)$ (same as in Section 3.3).

Hardness of $P_{k,n}$: We assume that there is a circuit C of size at most $d^{1/\mu}$ computing a polynomial $C(\boldsymbol{x},\epsilon) \in \mathbb{F}(\epsilon)[\boldsymbol{x}]$, which approximates $P_{k,n}$ over large enough $n \in J$, where $J \subseteq \mathbb{N}$ is an infinite subset. In particular, $C(\boldsymbol{x},\epsilon) =: \epsilon^M \cdot P_{k,n} + \epsilon^{M+1} \cdot Q(\boldsymbol{x},\epsilon)$ for some $M \in \mathbb{N}_{\geq 0}$. Using Theorem 30 and Fischer's trick, one can write

$$C(\boldsymbol{x},\epsilon) \;=\; \sum_{i\in[s_0]} c_i \cdot \tilde{f}_i^4$$

where $s_0 \leq 16c \cdot d^{6/\mu} \cdot (kn)^{13}$, for some constant c and $\deg(\tilde{f}_i) \leq kn/3$. Apply $\phi_{k,n}$ on $C(\boldsymbol{x},\epsilon)$. As, $\phi_{k,n} \circ \psi_{k,d} = \mathrm{id}$, over $\mathbb{F}[x]^{\leq d}$. Thus,

$$\epsilon^M \cdot f_d + \epsilon^{M+1} \cdot \tilde{Q} \;:=\; (\phi_{k,n} \circ \psi_{k,d})(C) \;=\; \sum_{i=1}^{s_0} c_i \cdot \left(\phi_{k,n}\left(\tilde{f}_i\right)\right)^4$$

where we have used that $\phi_{k,n}(P_{k,n}) = f_d$ and $\phi_{k,n}(Q(\boldsymbol{x}, \epsilon)) = \tilde{Q}(x, \epsilon)$, for some $\tilde{Q} \in \mathbb{F}[x, \epsilon]$. It is important to observe that $|\bigcup_i \text{supp}(\tilde{f}_i)| \leq s_1 := \binom{k+kn/3}{k}$. Since Kronecker map can not increase the support size, therefore $|\bigcup_i \text{supp}((\phi_{k,n}(\tilde{f}_i)))| \leq s_1$. Thus, we must have $\overline{U}_{\mathbb{F}}(f_d, s_0) \leq s_1$ from the definition of $\overline{U}_{\mathbb{F}}(\cdot)$.

We can show that $s_0 < d^{\delta_1}$ and $s_1 < d/4^{\delta_2}$, for all large enough n, where $k \geq 37\delta_2$ (as shown in Section 3.3). Therefore, we have $\overline{U}_{\mathbb{F}}(f_d, d^{\delta_1}) < d/4^{\delta_2}$, over all large $d \in J' := \{d(n) \mid n \in J\} \subseteq I$. This contradicts Conjecture C3. Thus, $\overline{\text{size}}(P_{k,n}) > d^{1/\mu}$, for a suitable constant μ and all large enough n.

Like in Section 3.3, $P_{k,n}$ is explicit and hard; thus Theorem 31 gives us a $\text{poly}(s)$-time HSG for $\overline{\text{size}}$-s degree-s polynomials. □

F Generalizing to the sum of r^{th}-powers

We could generalize our results to the sum of r^{th}-powers model, for a *constant prime-power* $r \geq 4$. This was done in [DST20], but only for $r \geq 25$.

We say a polynomial $f(x) \in R[x]$ over a ring R is computed as the sum of r^{th}-powers if

$$f = \sum_{i=1}^{s} c_i \cdot \ell_i^r \,. \tag{18}$$

The sum of r^{th}-powers is a *complete model* for $R = \mathbb{F}$, a field of characteristic zero or large characteristic. Like $S_{\mathbb{F}}(f)$, the *sparsity-sum size of f wrt a fixed r*, denoted by $S_{\mathbb{F}}(f, r)$, is defined as the minimum sparsity-sum size when f is written as in Eqn.(18).

We can restrict the domain to $I_r := \{r^m - 1 \mid m \in \mathbb{N}\}$. Similar heuristic (as given in Section 1) shows that for *random* f, $S_{\mathbb{F}}(f, r) \geq \Omega(d/\log d)$. For $f_d := (x+1)^d$, it is not hard to show that $S_{\mathbb{Z}}(f_d, r) \geq \Omega(d)$. To argue, let $r = r_0^\ell$ for some $\ell \in \mathbb{N}$ and a prime number r_0. Then, $\left|f_d \bmod r_0\right|_1 = d + 1$; as using Lucas's Theorem, $\binom{d}{i} = \pm 1$ for $0 \leq i \leq d$ where $d \in I_r$. Therefore, we could conjecture the following (similar to Conjecture C1):

Conjecture 4 (C1′). *If r is a prime-power and $d \in I_r$, then $S_{\mathbb{F}}(f_d, r) \geq \Omega(d^{0.98289})$.*

Theorem 2 establishes that assuming GRH, Conjecture C1 implies $\mathsf{VP} \neq \mathsf{VNP}$. We could show the same assuming Conjecture C1′. To formally state:

Theorem 33 (Conditional l.b.). *If GRH and Conjecture C1′ for an $r \geq 4$ hold, then $\mathsf{VP} \neq \mathsf{VNP}$.*

Proof sketch. The proof is very similar to that of Theorem 2, see Section 3.2. We define $P_{k,n}$, for a *non*-constant k, to be the k-variate polynomial with constant individual degree n, via the inverse Kronecker map applied on f_d (for $d := d(k)$ picked the *largest* possible element in I_r which is $\leq 2^{7k} - 1$). It is clear that $n := \lceil (d+1)^{1/k} \rceil - 1 = 127$. As the ratio between two consecutive elements in I_r can be at most $(r^{m+1} - 1)/(r^m - 1) < r + 1$, for $m \geq 2$, it is clear that $d \geq (2^{7k} - 1)/(r+1) = \Omega(128^k)$.

The proof strategy remains the same. We could prove that Claim 11 holds i.e. GRH and $\mathsf{VP} = \mathsf{VNP} \implies (P_{k,n})_k \in \mathsf{VP}$. The proof also remains unchanged. Assuming, Conjecture C1′, we next show that $(P_{k,n})_k \notin \mathsf{VP}$ (same as Claim 12).

To prove the hardness part, the only crucial difference is to relate CNF (Theorem 1) to sum of r^{th}-powers. This can be tackled by using Lemma 14. In particular, we prove that $\text{size}(P_{k,n}) > d^{10^{-5}/6} = 2^{\Omega(k)}$ by assuming a contradiction. Suppose there is an infinite $J \subset \mathbb{N}$ such that $\text{size}(P_{k,n}) \leq d^{10^{-5}/6}$, for all $k \in J$. This will prove that Conjecture C1′ is false over infinite subset $J' := \{d(k) : k \in J\} \subseteq I_r$ which is a contradiction.

Assume that C is a circuit of size $\leq d^{10^{-5}/6}$ computing $P_{k,n}$. Then, as done in the proof of Theorem 2, one could conclude that $P_{k,n}$ can be expressed as Eqn. (11):

$$P_{k,n} \;=\; \sum_{i=1}^{16\cdot s'} c_i \cdot \tilde{f}_i^4 \,,$$

where $s' \leq c \cdot d^{10^{-5}} \cdot (127k)^{13}$, for some constant c and $\deg(\tilde{f}_i) \leq 127k/3$. Then, we apply Lemma 14 on each $\tilde{f}_i^4$ yielding the desired sum-of-r^{th}-powers:

$$P_{k,n} \;=\; \sum_{i=1}^{16\cdot s'\cdot(r+1)} c_i' \cdot g_i^r \,,$$

with $\deg(g_i) \leq \max_j \deg(\tilde{f}_j) \leq 127k/3$. Apply Kronecker $\phi_{k,n}$ on both sides; as it cannot increase the sparsity, we get: $S_{\mathbb{F}}(f,r) \leq 16(r+1)s' \cdot \binom{130k/3}{3} = (r+1)\cdot s_1$ (we follow the notation of the proof of Theorem 2). As r is constant, similar calculation establishes that $(r+1)\cdot s_1 \leq o(d^{0.98289})$ contradicting Conjecture C1′. □

To proclaim PIT result, we change the measure to support-union with respect to f and r.

Support-union size of f with respect to s and exponent r, denoted $U_{\mathbb{F}}(f,r,s)$, is defined to be the *minimum* number of distinct monomials in the representation of Eqn.(18); in other words, $\left|\bigcup_{i=1}^{s} \mathrm{supp}(\ell_i)\right|$ when f is written as Eqn.(18); it is ∞, if no such representation exists. Similar to Conjecture C2, one could conjecture the following (same as [DST20, Conj. C1]):

Conjecture 5 (C2′). *Let r be a prime-power. There exist positive constants $\delta_1 \leq 1$, $\delta_2 \geq 1$ such that $U_{\mathbb{F}}\left(f_d, r, d^{\delta_1}\right) \geq d/r^{\delta_2}$, for all large enough $d \in I_r$.*

By [DST20, Thm. 1] the above conjecture implies blackbox-PIT $\in$ P for $r \geq 25$. We could argue the same for a much less $r \geq 4$. Formally,

Theorem 34. *If Conjecture C2′ holds for an $r \geq 4$, then blackbox-PIT $\in$ P.*

Proof sketch. This proof is very similar to the proof of Theorem 3, see Section 3.3. Define, for a constant k, a k-variate, individual degree-n, polynomial $P_{k,n}$ via the inverse Kronecker map applied on f_d. It is an explicit family as the coefficients are poly-time computatble. Similar to Claim 13, we could show that $\mathrm{size}(P_{k,n}) > d^{1/\mu}$, where $\mu := 6/(\delta_1 - 14/k)$.

To show the hardness, assume to the contrary that there is an infinite $J \subset \mathbb{N}$ such that $\mathrm{size}(P_{k,n}) \leq d^{1/\mu}$ for $n \in J$. We will show that Conjecture C2′ is false over an infinite $J' := \{d(n) \mid n \in J\} \subseteq I_r$; which is a contradiction.

Let C be a circuit of size $\leq d^{1/\mu}$ computing $P_{k,n}$. Then, similar to Eqn. (12), we have

$$P_{k,n} \;=\; \sum_{i\in[s_0]} c_i \cdot \tilde{f}_i^4 \,,$$

where $s_0 := 16 \cdot c \cdot (d^{1/\mu})^6 \cdot (kn)^{13}$ for some constant c and $\deg(\tilde{f}_i) \leq kn/3$. Then, we apply Lemma 14 on each $\tilde{f}_i^4$ yielding the desired sum-of-r^{th}-powers:

$$P_{k,n} \;=\; \sum_{i=1}^{s_0\cdot(r+1)} c_i \cdot g_i^r \,,$$

with $\deg(g_i) \leq \max_j \deg(\tilde{f}_j) \leq kn/3$. Apply Kronecker $\phi_{k,n}$ on both sides; as it cannot increase the union-support or the top fan-in, we get $U_{\mathbb{F}}\left(f_d, r, s_0\cdot(r+1)\right) \leq s_1 := \binom{k+kn/3}{k}$.

As, r is just a constant, the bound on $s_0\cdot(r+1)$ holds similar to that in the proof of Theorem 3, establishing $s_0\cdot(r+1) < d^{\delta_1}$.

Bound on s_1 requires a slightly *different* parameter settings as we want to show that $s_1 < d/r^{\delta_2}$ (which is tighter than $d/4^{\delta_2}$). Once we prove this, we are done as, in that case, $U_{\mathbb{F}}(f_d, r, d^{\delta_1}) < d/r^{\delta_2}$, for all large $d \in J'$ contradicting Conjecture C2′.

We already showed in the proof of Theorem 3 (Section 3.3), that $s_1 < \left(3 \cdot (14/15)^k\right) \cdot d$. It suffices to show that $3 \cdot (14/15)^k \leq 1/r^{\delta_2}$. As $r \geq 4$, it is enough to choose $k \cdot \log(15/14) \geq (\delta_2 + 1) \cdot \log r$. It is enough to pick $k > \max((\delta_2 + 1) \cdot \log r / \log(15/14), 14/\delta_1)$.

This hardness result, then, will directly give an efficient HSG using Theorem 28, as shown in the proof of Theorem 3. Hardness to HSG requires $k \geq 38/\delta_1$, same as in the earlier proof. Hence, our final choice for k is $k \geq \max((\delta_2 + 1) \cdot \log r / \log(15/14), 38/\delta_1)$. Thus, we have a $\mathsf{poly}(s)$-time HSG for $\mathcal{C}(s, s, s)$.

□

Remark. It is *natural* to define the approximative measure $\overline{U}_{\mathbb{F}}(f, r, s)$, for any constant prime-power $r \geq 4$. One could prove that similar lower bound on this measure will lead to efficient HSG for $\overline{\mathsf{VP}}$. A similar result was proved in [DST20, Thm. 29], but for $r \geq 25$.