Electronic Colloquium on Computational Complexity, Revision 2 of Report No. 63 (2020)

On the Existence of Algebraic Natural Proofs*

Prerona Chatterjee† Mrinal Kumar‡ C. Ramya§ Ramprasad Saptharishi¶
Anamay Tengse||

Abstract

The framework of *algebraically natural proofs* was independently introduced in the works of Forbes, Shpilka and Volk (2018), and Grochow, Kumar, Saks and Saraf (2017), to study the efficacy of commonly used techniques for proving lower bounds in algebraic complexity. We use the known connections between algebraic hardness and pseudorandomness to shed some more light on the question relating to this framework, as follows.

- The subclass of VP that contains polynomial families with bounded coefficients, *has* efficient equations. Over finite fields, this result holds without any restriction on coefficients. Further, both these results also extend to the class VNP as is.
- Over fields of characteristic zero, VNP *does not have* any efficient equations, if the Permanent is exponentially hard for algebraic circuits.
  This gives the only known barrier to "natural" lower bound techniques (that follows from believable hardness assumptions), and also shows that the restriction on coefficients in the first category of results about VNP is necessary.

The first set of results follows essentially by algebraizing the well-known method of generating hardness from non-trivial hitting sets (e.g. Heintz and Schnorr 1980). The conditional hardness of equations for VNP uses the fact that pseudorandomness against a class can be extracted from a polynomial that is (sufficiently) hard for that class (Kabanets and Impagliazzo, 2004).

*An extended version of a combination of two preliminary works, titled "On the Existence of Algebraically Natural Proofs" (FOCS 2020) and "If VNP is hard, then so are equations for it" (STACS 2022).

†prerona.ch@gmail.com. Tel Aviv University, Israel. Research supported by the Azrieli International Postdoctoral Fellowship, the Israel Science Foundation (grant number 514/20) and the Len Blavatnik and the Blavatnik Family foundation. Most of this work was done while the author was a PhD student at the School of Technology and Computer Science, Tata Institute of Fundamental Research, Mumbai, India.

‡mrinal.kumar@tifr.res.in. School of Technology and Computer Science, Tata Institute of Fundamental Research, Mumbai, India. Most of this work was done while working at the Department of Computer Science & Engineering, IIT Bombay.

§ramyac@imsc.res.in. The Institute of Mathematical Sciences (a CI of Homi Bhabha National Institute), Chennai, India. A part of this work was done while the author was a post-doctoral fellow at TIFR, Mumbai, India.

¶ramprasad@tifr.res.in. School of Technology and Computer Science, Tata Institute of Fundamental Research, Mumbai, India. Research supported by Ramanujan Fellowship of DST, and by DAE, Government of India.

||anamay.tengse@gmail.com. Efi Arazi School of Computer Science, Reichman University, Herzliya, supported by the Israeli Science Foundation grant 843/23. A major part of this work was done when the author was a PhD student at TIFR, Mumbai, India.

ISSN 1433-8092

1 Introduction

The quest for proving strong lower bounds for algebraic circuits is one of the fundamental challenges in algebraic complexity, and maybe the most well-studied one. Yet, progress on this problem has been painfully slow and sporadic. Perhaps the only thing more frustrating than the inability to prove such lower bounds is the difficulty in coming up with plausible approaches towards them. This lack of progress has spurred an interest towards understanding the viability of some commonly used lower bound approaches; the idea being that a good sense of what approaches will *not* work would aid in the search of those that *might* work. Moreover, such meta-studies could help identify the strengths of the current and future approaches that show promise.

In the broader context of lower bounds in computational complexity, there are various results of this flavor which establish that various families of techniques cannot be used for proving very strong lower bounds. For instance, the barrier of *Relativization* due to Baker, Gill and Solovay [BGS75], that of *Algebraization* due to Aaronson and Wigderson [AW09] and that of *Natural Proofs* due to Razborov and Rudich [RR97].[1] While none of these barrier results are directly applicable to the setting of algebraic computation, there have been recent attempts towards generalizing these ideas to the algebraic set up. A key notion in this line of work is the notion of *algebraically natural proofs* alluded to and defined in the works of Aaronson and Drucker [AD08], Forbes, Shpilka and Volk [FSV18], and Grochow, Kumar, Saks and Saraf [GKSS17].

We now discuss this notion, starting with a discussion on *Natural Proofs* which motivated the definition.

1.1 The Natural Proofs framework of Razborov and Rudich

Razborov and Rudich [RR97] noticed that underlying many of the lower bound proofs known in Boolean circuit complexity, there was some common structure. They formalized this common structure via the notion of a *Natural Property*, which we now define.

Definition 1.1. *A subset $\mathcal{P} \subseteq \{f : \{0,1\}^n \to \{0,1\}\}$ of Boolean functions is said to be a natural property useful against a class $\mathscr{C}$ of Boolean circuits if the following are true.*

- ***Usefulness.*** *Any Boolean function $f : \{0,1\}^n \to \{0,1\}$ that can be computed by a Boolean circuit in $\mathscr{C}$ does not have the property $\mathcal{P}$.*
- ***Constructibility.*** *Given the truth table of a Boolean function $f : \{0,1\}^n \to \{0,1\}$, whether it has the property $\mathcal{P}$ can be decided in time polynomial in the length of the input, i.e. in time $2^{O(n)}$.*
- ***Largeness.*** *For all large enough n, at least a $2^{-O(n)}$ fraction of all n variate Boolean functions have the property $\mathcal{P}$.* $\lozenge$

A proof that a certain family of Boolean functions cannot be computed by circuits in $\mathscr{C}$ is said to be a *natural lower bound proof* if it (perhaps implicitly) proceeds via establishing a natu-

[1] Sometimes, these results are conditional, as in [RR97].

ral property useful against $\mathscr{C}$, and showing that the candidate hard function has this property. Razborov and Rudich then showed that most of the Boolean circuit lower bound proofs that we know (for example, lower bounds for AC^0 circuits [FSS84, Hås86] or lower bounds for $\mathsf{AC}^0[\oplus]$ circuits [Raz87, Smo87]) fit into this framework, maybe with some work, and hence are *natural* in this sense. Further they argue that, under standard cryptographic assumptions, the proof of a lower bound against any sufficiently rich circuit class (such as the class $\mathsf{P/poly}$) cannot be natural! Thus, under standard cryptographic assumptions, most of the current lower bound techniques are not strong enough to show super-polynomial lower bounds for general Boolean circuits.

We now move on to discuss a relatively recent analogue of the notion of Natural Proofs, formalized in the context of algebraic computation.

1.2 Algebraically Natural Proofs

Algebraic complexity is the study of computational questions about polynomials as formal objects. The basic model of computation here, an algebraic circuit, is an algebraic analogue of a boolean circuit with the gates of the circuit being labeled by $+$ (sum) and $\times$ (product) gates as opposed to Boolean functions; the size of a circuit is the number of wires (edges) in it.[2] The algebraic analogue of $\mathsf{P/poly}$ is the class VP of polynomial families $\{f_n\}$, where f_n is an n variate polynomial of degree and algebraic circuit size $\mathrm{poly}(n)$. A fundamental question in this setting is to come up with explicit families of polynomials, those in the class VNP (the algebraic analog of $\mathsf{NP/poly}$), which are not in VP. While the state of the art of size lower bounds for algebraic circuits is a bit better than that for Boolean circuits, with slightly super linear lower bounds having been shown by Strassen [Str73] and Baur & Strassen [BS83], this lower bound has seen no improvements for nearly four decades. The recent breakthrough by Limaye, Srinivasan and Tavenas [LST21] does indeed suggest that stronger lower bounds might be within reach in the near future, but it is far from clear how that could be done. This absence of progress has led to some research towards understanding the limitations of the current proof techniques in proving strong lower bounds for algebraic circuits.

Considering that algebraic circuits seem like a fairly general and powerful model of computation, it is tempting to think that the *natural proofs barrier* of Razborov and Rudich [RR97] also extends to this setting. However, this problem turns out to be a non-trivial one, and indeed, it is not known whether their results extend to algebraic circuits. This question is closely related to the existence of cryptographically secure, algebraic pseudorandom functions that can be computed by small and low degree[3] algebraic circuits, and there does not seem to be substantial evidence one way or the other on this. We refer the reader to [AD08] and [FSV18] for a more detailed discussion on this issue.

In the last few years, this question of trying to find an algebraic analogue of the barrier results

[2] See Definition 2.5 for a formal definition.

[3] Throughout this paper, by a *low degree* polynomial family, we mean a polynomial family whose degree is polynomially bounded in its number of variables.

in [RR97] has received substantial attention. It was observed by various authors [AD08, Gro15, FSV18, GKSS17] that most of the currently known proofs of algebraic circuit lower bounds fit into a common unifying framework, not unlike that in [RR97], although of a more algebraic nature. Indeed, these proofs also implicitly go via defining a property for the set of all polynomials and using this property to separate the hard polynomial from the easy ones. Moreover, the notions of *largeness* and *constructibility* in Definition 1.1 also seem to extend to these proofs.

We now discuss this framework in a bit more detail. The key notion here is that of an equation for a set of polynomials.

Definition 1.2 (Equations for a set of polynomials)**.** *For some $n, d \in \mathbb{N}$, let $\mathcal{C}_{n,d}$ be a set of n-variate polynomials of* total *degree at most d; i.e. $\mathcal{C}_{n,d} \subseteq \mathbb{F}[\mathbf{x}]^{\leq d}$.*

Then, for $N = \binom{n+d}{n}$, a nonzero polynomial $P_N(\mathbf{Z})$ is said to be an equation *for $\mathcal{C}_{n,d}$ if for all $f(\mathbf{x}) \in \mathcal{C}_{n,d}$, we have that $P_N(\overline{\text{coeff}}(f)) = 0$, where $\overline{\text{coeff}}(f)$ is the coefficient vector of f.* ◊

The definition naturally extends to a class of polynomial families, as opposed to just a set of polynomials as defined above. In particular, suppose that $\mathscr{C}$ is a class of polynomial families $\{\{f_n\} : f_n \in \mathcal{C}_{n,d_n}\}$, and $\{P_N\}$ is a polynomial family. Then, the family $\{P_N\}$ is said to be a family of equations for $\mathscr{C}$ if $P_{N(n)}$ is an equation for $\mathcal{C}_{n,d_n}$ for all large enough n, for $N(n) := \binom{n+d_n}{n}$. That is, there is some n_0 such that for all $n \geq n_0$ the polynomial $P_{N(n)}$ is an equation for $\mathcal{C}_{n,d_n}$.

Intuitively, non-vanishing of an equation (for a set $\mathcal{C}$) on the coefficient vector of a given polynomial f is a proof that f is not in $\mathcal{C}$. We note that the equations for a set $\mathcal{C}$ evaluate to zero not just on the coefficient vectors of polynomials in $\mathcal{C}$ but also on the coefficient vectors of polynomials in the Zariski closure of $\mathcal{C}$. This framework comes up very naturally in the context of algebraic geometry (and geometric complexity theory), where it is often geometrically nicer to work with the variety obtained by taking the Zariski closure of a complexity class.

Getting our hands on an equation of a variety gives us a plausible way to test and certify non-membership in the variety, in other words, to prove a lower bound for the corresponding complexity class. Thus, families of equations for a class gives an algebraic analogue of the notion of *natural properties useful against a class* in [RR97]. Moreover, since a nonzero polynomial does not vanish very often on a random input from a large enough grid, it follows that a nonzero equation for a set $\mathcal{C}$ will be nonzero on the coefficient vector of a "random polynomial". Here by a random polynomial we mean a polynomial whose coefficients are independent and uniformly random elements from some large enough set in the underlying field. With appropriate quantitative bounds, this observation can be formalized to give an appropriate algebraic analogue of the notion of *largeness*. Lastly, the algebraic circuit complexity of the equation gives a natural algebraic analog of the notion of *constructibility*. Intuitively, any algebraic circuit lower bound which goes via defining a nonzero proof polynomial of polynomially bounded degree that can be efficiently computed by an algebraic circuit is an *Algebraically Natural Proof* of a lower bound.

We now formally define an algebraically natural proof.

Definition 1.3 (Algebraically natural proofs [FSV18, GKSS17])**.** *Let $\mathscr{C}$ be a class of polynomial families $\{\{f_{n,d}\} : f_{n,d} \in \mathcal{C}_{n,d}\}$.*

Then, for a class $\mathscr{D}$ of polynomial families, we say that $\mathscr{C}$ has *$\mathscr{D}$*-natural proofs *if there is a family $\{P_N\} \in \mathscr{D}$ which is a non-trivial family of equations for $\mathscr{C}$.* ◊

In the rest of this paper, whenever we say a natural proof, without specifying the class $\mathscr{D}$, we mean a VP-natural proof.

Analogous to the abstraction of *natural proofs* for Boolean circuit lower bounds, this framework of *algebraically natural proofs* turns out to be rich and general enough that almost all of our current proofs of algebraic circuit lower bounds are in fact algebraically natural, or can be viewed in this framework with a little work [Gro15]. Thus, this definition seems like an important first step towards understanding the strengths and limitations of many of our current lower bound techniques in algebraic complexity.

The immediate next question to ask is whether algebraically natural proofs are rich enough to give strong algebraic circuit lower bounds. This can naturally be worded in terms of the complexity of equations for the class VP as follows.

Question 1.4. *For every constant $c > 0$, does there exist a nonzero polynomial family $\{P_{N,c}\}$ in* VP *such that for all large enough n, the following is true?*

> *For every family of polynomials $\{f_n\}$ in* VP*, such that f_n is an n variate polynomial of degree n^c, $P_N^{(c)}$ vanishes on the coefficient vector of f_n for $N = \binom{n+n^c}{n}$.*

The works [FSV18] and [GKSS17] argue that under an appropriate (but non-standard) pseudorandomness assumption, the answer to the question above is negative, i.e., algebraically natural proof techniques cannot be used to show strong lower bounds for algebraic circuits. To discuss this pseudorandomness assumption formally, we need to define *succinct hitting sets*.

Definition 1.5 (Succinct hitting sets for a set of polynomials)**.** *For some $n, d \in \mathbb{N}$, let $\mathcal{C}_{n,d}$ be a set of n-variate polynomials of* total *degree at most d; that is, $\mathcal{C}_{n,d} \subseteq \mathbb{F}[\mathbf{x}]^{\leq d}$.*

Then for $N = \binom{n+d}{n}$, we say that a set of N variate polynomials $\mathcal{D}_N$ has *$\mathcal{C}_{n,d}$*-succinct hitting sets *if for all nonzero $P(\mathbf{Z}) \in \mathcal{D}_N$, there exists some $f \in \mathcal{C}_{n,d}$ such that $P_N(\overline{\text{coeff}}(f)) \neq 0$.* ◊

As with Definition 1.2, this definition naturally extends to polynomial families (see Definition 2.16).

It immediately follows from these definitions that non-existence of $\mathscr{D}$-natural proofs against a class $\mathscr{C}$ is equivalent to the existence of $\mathscr{C}$-succinct hitting sets for the class $\mathscr{D}$. Forbes, Shpilka and Volk [FSV18] showed that for various restricted circuit classes $\mathscr{C}$ and $\mathscr{D}$, the class $\mathscr{D}$ has $\mathscr{C}$ succinct hitting sets. Or equivalently, lower bounds for $\mathscr{C}$ cannot be proved via proof polynomial families in $\mathscr{D}$. However, we already have super-polynomial lower bounds against these classes $\mathscr{C}$, making the evidence weak. Further, this question has remained unanswered for more general circuit classes $\mathscr{C}$ and $\mathscr{D}$. In particular, if we take both $\mathscr{C}$ and $\mathscr{D}$ to be VP, we do not seem to have significant evidence on the existence of VP-succinct hitting sets[4] for VP.

In [FSV18], the authors observed that showing VP succinct hitting sets for VP would immediately imply non-trivial deterministic algorithms for polynomial identity testing which, via well

[4]The definition of VP-succinct hitting sets (Definition 2.19) is perhaps slightly non-intuitive.

known connections between algebraic hardness and derandomization, will in turn imply new lower bounds [HS80a, KI04]. Thus, the problem of proving an unconditional barrier result for algebraically natural proof techniques via this route seems as hard as proving new circuit lower bounds! It is, however, conceivable that one can show such a barrier conditionally. And in some more structured settings, such as for the case of matrix completion, such results are indeed known [BIJL18]. However, Question 1.4 remains open. In particular, even though many of the structured subclasses of VP have low degree equations which are very efficiently computable, perhaps hoping that this extends to richer and more general circuit classes is too much to ask for?

1.3 Our results

We are now ready to state our results. Our first set of results can be viewed as evidence *towards* the efficacy of natural techniques for proving lower bounds against VP, and possibly even VNP.

Equations for polynomials in VP with coefficients of small complexity

We first show that over the field of complex numbers, there are efficiently computable equations for the set of polynomials in VP that have small coefficients. Here for a field $\mathbb{F}$, $\mathsf{VP}^{\mathbb{F}}$ denotes the class VP where the coefficients are from the field $\mathbb{F}$.

Theorem 1.6. *Let $c > 0$ be any constant. There is a polynomial family $\{P_N^{(c)}\} \in \mathsf{VP}^{\mathbb{Q}}$ such that for $N(n) = \binom{n+n^c}{n}$, the following are true.*

- *For every $t(n) = \mathrm{poly}(n)$, for all large enough n, and every family $\{f_n\} \in \mathsf{VP}^{\mathbb{C}}$, where f_n is an n-variate, degree-n^c, size-$t(n)$ polynomial with coefficients in $\{-1, 0, 1\}$, we have that*

$$P_N^{(c)}(\overline{\mathrm{coeff}}(f_n)) = 0\,,$$

  where $\overline{\mathrm{coeff}}(f_n)$ is the coefficient vector of f_n.
- *There exists a family $\{h_n\}$ of n variate polynomials and degree $\leq n^c$ with coefficients in $\{-1, 0, 1\}$ such that for all large enough n,*

$$P_N^{(c)}(\overline{\mathrm{coeff}}(h_n)) \neq 0\,.$$

We note that even though Theorem 1.6 is stated for polynomials with $\{-1, 0, 1\}$ coefficients, the theorem holds for polynomials with coefficients as large as N.

However, for brevity, we will confine the discussion in this paper to polynomials with coefficients in $\{-1, 0, 1\}$. We also note that the same statement holds over other fields of characteristic zero as well. That is, the $\mathsf{VP}^{\mathbb{C}}$ in the above statement can be replaced with $\mathsf{VP}^{\mathbb{R}}$ or $\mathsf{VP}^{\mathbb{Q}}$.

We also prove an analogous theorem for finite fields.

Theorem 1.7. *Let $\mathbb{F}$ be any finite field, and let $c > 0$ be any constant. There is a polynomial family $\left\{P_N^{(c)}\right\} \in \mathsf{VP}^{\mathbb{F}}$ such that for $N(n) = \binom{n+n^c}{n}$, the following are true.*

- *For every $t(n) = \text{poly}(n)$, for all large enough n, and every family $\{f_n\} \in \mathsf{VP}^{\mathbb{F}}$, where f_n is an n-variate, degree-n^c, size-$t(n)$ polynomial, we have that*

$$P_N^{(c)}(\overline{\text{coeff}}(f_n)) = 0\,.$$

- *There exists a family $\{h_n\}$ of n variate polynomials and degree $\leq n^c$ with coefficients in $\mathbb{F}$ such that for all large enough n,*

$$P_N^{(c)}(\overline{\text{coeff}}(h_n)) \neq 0\,.$$

Equations for polynomials in VNP with coefficients of small complexity

Furthermore, we also prove analogous statements for the seemingly larger class VNP, as follows.

Theorem 1.8. *Let $c > 0$ be any constant. There is a polynomial family $\{Q_N^{(c)}\} \in \mathsf{VP}^{\mathbb{Q}}$ such that for $N(n) = \binom{n+n^c}{n}$, the following are true.*

- *For every $t(n) = \text{poly}(n)$, for all large enough n, and every family $\{f_n\} \in \mathsf{VNP}^{\mathbb{C}}$, where f_n is an n-variate, degree-n^c, size-$t(n)$ polynomial with coefficients in $\{-1, 0, 1\}$, we have that*

$$Q_N^{(c)}(\overline{\text{coeff}}(f_n)) = 0\,.$$

- *There exists a family $\{h_n\}$ of n variate polynomials and $\leq n^c$ with coefficients in $\{-1, 0, 1\}$ such that for all large n,*

$$Q_N^{(c)}(\overline{\text{coeff}}(h_n)) \neq 0\,.$$

As before, we note that the above theorem holds for polynomials with coefficients as large as N and also holds over other fields of characteristic zero, like $\mathbb{R}$ or $\mathbb{Q}$.

Theorem 1.9. *Let $\mathbb{F}$ be any finite field and $c > 0$ be any constant. There is a polynomial family $\left\{Q_N^{(c)}\right\} \in \mathsf{VP}^{\mathbb{F}}$ such that for $N(n) = \binom{n+n^c}{n}$, the following are true.*

- *For every $t(n) = \text{poly}(n)$, for all large enough n, and every family $\{f_n\} \in \mathsf{VNP}^{\mathbb{F}}$, where f_n is an n-variate, degree-n^c, size-$t(n)$ polynomial, we have that*

$$Q_N^{(c)}(\overline{\text{coeff}}(f_n)) = 0\,.$$

- *There exists a family $\{h_n\}$ of n variate polynomials and degree $\leq n^c$ with coefficients in $\mathbb{F}$ such that for all large n,*

$$Q_N^{(c)}(\overline{\text{coeff}}(h_n)) \neq 0\,.$$

In fact, we show that the existence of efficient hitting sets for any class is sufficient to give efficient equations for its subclass that contains polynomials with bounded coefficients. This is formalized in Theorem 4.3 for fields of characteristic zero, and in Theorem 4.1 for finite fields.

Conditional hardness of equations for VNP

Over fields of characteristic zero, we also show that assuming the Permanent is hard enough, the constraint of bounded coefficients in Theorem 1.8 is necessary for efficient equations for VNP. More formally, we show the following.

Theorem 1.10 (Conditional Hardness of Equations for VNP)**.** *Let $\varepsilon > 0$ be a constant. Suppose that the permanent family Perm_m requires circuits of size 2^{m^ε}.*

Then, VP has VNP-succinct hitting sets. Therefore, there are no VP-natural proofs for VNP.

We remark that the above theorem holds over any field of characteristic zero. For our proofs, we will work with complexes for better readability.

Remark 1.11. *Extending the result in Theorem 1.10 to hardness of equations for VP, even under the assumption that Permanent is sufficiently hard, is a fascinating open question. Such an extension would answer the main question investigated in [FSV18, GKSS17] and show a natural-proofs-like barrier for a fairly general family of lower bound proof techniques in algebraic complexity. Our proof of Theorem 1.10 however, crucially relies on some key properties of VNP, and does not appear to extend to VP.* ◊

1.4 Discussion and relations to prior work

As is evident from our results, the main message (in our opinion) is that we do not have compelling evidence to rule out, or accept, the efficacy of algebraically natural proofs towards proving strong lower bounds for rich classes of algebraic circuits. In fact, our results seem to provide *some* evidence for both sides.

We first discuss the results that suggest an affirmative answer to Question 1.4. Many of the families of polynomials commonly studied in algebraic complexity have integer coefficients with absolute values bounded by 1, and fall in the setting of Theorem 1.6. Moreover, the condition of computing polynomials with bounded coefficients is a semantic condition on a model, in the sense that even though the final output of the circuit is required to have bounded coefficients, the circuit is free to use arbitrary constants from $\mathbb{C}$ in the intermediate computation. Thus, it is conceivable that we might be able to prove a super-polynomial lower bound on the algebraic circuit size for the permanent polynomial via an algebraically natural proof constructible in VP, thereby separating VP and VNP. However, since analogues of Theorem 1.6 and Theorem 1.7 are also true for VNP, any such separation will have to rely on more fine-grained information on the equations, and not just their degree and algebraic circuit size. Unfortunately, our proofs are all existential and do not give a sense of what the polynomial families $\{P_N^{(c)}\}$ (or $\{Q_N^{(c)}\}$) might look like.

We also note that in the light of some prior works, these results are perhaps a bit surprising. The classes of polynomials in VP and VNP with small coefficients (or over finite fields), are seemingly rich and complex, and the theorems here show — unconditionally — that they have equations which are also efficiently computable. Furthermore, these equations are a rather straightforward consequence of the existence of efficient hitting sets, as shown in Theorems 4.1 and 4.3. As discussed earlier, the existence of efficient equations is known to be true for many structured subclasses of algebraic circuits (for example, homogeneous circuits of depth 3 and 4, multilinear formulas, polynomials of small Waring rank). However, it is unclear if this property extends to more general circuit classes, like VP or VNP.

On the other hand, Theorem 1.10 shows that under the widely believed assumption that Perm is exponentially hard, the bound on coefficients is crucial, at least for the class VNP. It is unclear if a similar situation is also true for VP. However, even if we were to believe that there are efficiently computable equations for VP, it is unclear if the existence of such equations implies VP $\neq$ VNP.

In the other direction, suppose we were to assume that VP and VNP are indeed different, is it then reasonable to expect an efficiently computable equation exhibiting such a separation? Using Theorem 1.10, we now have that if Perm is exponentially hard, then *any* efficiently computable equation for VP will necessarily *not* be an equation for VNP, thus yielding an "algebraically natural proof" that separates VP and VNP.

We will now briefly go over some works related to Question 1.4. As mentioned earlier, the focus of most of the prior works has been to look for evidence that the answer to Question 1.4 is negative, i.e. VP does not have efficiently computable and low degree equations. We hope that the results in this paper highlight that the answer is not so clear.

Relations to prior work. Following the work of Forbes, Shpilka, Volk [FSV18] and Grochow, Kumar, Saks and Saraf [GKSS17], much of the research on the problem, of whether algebraically natural proofs exist, has focused on proving the *non-existence* of efficiently computable equations for VP, and this line of work has made interesting progress in this direction for many structured and special instances of problems of this nature. Forbes, Shpilka and Volk [FSV18] unconditionally ruled out equations for depth-three multilinear formulas computable by certain structured classes of algebraic circuits using this connection. However, this does not imply anything about complexity of equations for general classes of algebraic circuits such as VP and VNP.

In the context of proving lower bounds against algebraic circuits, Efremenko, Garg, Oliveira and Wigderson [EGOW18] and Garg, Makam, Oliveira and Wigderson [GMOW19] explore limitations of proving algebraic circuit lower bounds via rank based methods. In particular, Efremenko *et al.* [EGOW18] show that some of these rank based methods cannot prove lower bounds better than $\Omega_d(n^{\lfloor d/2 \rfloor})$ on tensor rank (respectively, Waring rank) for a d-dimensional tensor of side n. Building on [EGOW18], in [GMOW19], the authors demonstrate that one *cannot* hope to sig-

nificantly improve the known lower bounds for tensor rank for d dimensional tensors by lifting lower bounds on tensors in fewer dimensions. However, we note that a general algebraically natural proof of a lower bound does not necessarily fit into the framework of [EGOW18, GMOW19], and so these limitations for the so-called *rank methods* do not seem to immediately extend to algebraically natural proofs in general. As discussed earlier, in the light of the results here, it is conceivable that we might be able to improve the state of the art for general algebraic circuit lower bounds, using techniques that are algebraically natural.

Bläser, Ikenmeyer, Jindal and Lysikov [BIJL18] studied the complexity of equations in a slightly different context. They draw connections between the existence of efficiently constructible equations of a variety and the problem of testing (non)membership in it and use the conditional hardness of the (non)membership testing problem for certain varieties to rule out the existence of efficiently computable equations for them. More precisely, they show that if *all* the equations for the variety of matrices with zero permanent are constructible by small constant-free algebraic circuits, then the non-membership problem for this variety can be decided in the class $\exists\mathsf{BPP}$. Thus, unless $\mathsf{P}^{\#\mathsf{P}} \subseteq \exists\mathsf{BPP}$, the equations of this variety do not have small, low degree constant free algebraic circuits. In a subsequent work ([BIL$^+$19]), the results of [BIJL18] are generalized to *min-rank* or *slice-rank* varieties. However, in the bounded coefficient setting (and over finite fields), our results show that the contrary is true, and VP does have efficiently computable low degree equations. We also note that the set-up in these papers differ from that in our paper, and that of [GKSS17, FSV18]. One way to interpret this difference is that [BIJL18] shows that "variety of small completion rank tensors" cannot be "cut out" by efficient equations, whereas [GKSS17, FSV18] and our paper asks if *every* equation for this variety requires large complexity.

A positive result on the complexity of equations of naturally-occurring varieties in algebraic complexity appears in a recent work of Kumar and Volk [KV20] where they show polynomial bounds on the degree of the equations of the Zariski closure of the set of non-rigid matrices and small linear circuits over all large enough fields. However, we do not know if any of these low-degree equations can be *efficiently* computed by an algebraic circuit.

For Boolean circuits, Chow [Cho11] shows a way of circumventing the natural proofs' barrier in [RR97] by providing (under standard cryptographic assumptions) an explicit *almost natural proof* that is useful against P/poly as well as constructive in nearly linear time, but compromises on the largeness condition. Furthermore, Chow [Cho11] shows the unconditional existence of a natural property useful against P/poly (infinitely often) constructive in linear size that has a weakened largeness condition. In some sense, Theorem 1.7 and Theorem 1.6 are analogous to the work of Chow [Cho11], albeit in the algebraic world.

On the largeness criterion. In the definitions of algebraically natural proofs [GKSS17, FSV18], the authors observe that in the algebraic setting, an analogue of the *largeness* criterion in Definition 1.1 is often available for free; the reason being that a nonzero equation for any class of polynomials vanishes on a very small fraction of all polynomials over any sufficiently large field.

However, this tradeoff becomes a bit subtle when considering polynomials over finite fields, or polynomials with bounded integer coefficients. In particular, as we observe in the course of the proofs of our results, we still have a sizeable chunk of polynomials whose coefficients will keep $\{P_N^{(c)}\}$ (and $\{Q_N^{(c)}\}$) nonzero, although this set is no longer a significant fraction of the set of all polynomials.

1.5 Proof ideas

Constructible equations for polynomials with coefficients of small complexity

At a high level, the idea behind our results about constructible equations is to try and come up with a *non-trivial* property of polynomials which every polynomial with a small circuit satisfies. By a non-trivial property, we mean that there should exist (nonzero) polynomials which do not have this property. The hope is that once we have such a property (that is nice enough), one can try to transform this into an equation via an appropriate *algebraization*. The property that we finally end up using is the existence of *hitting sets* for polynomials with small circuits.

A hitting set for a class $\mathscr{C}$ of polynomials over a field $\mathbb{F}$ is a set of points $\mathcal{H}$, such that every nonzero polynomial in $\mathcal{C}$ evaluates to a nonzero value on at least one point in $\mathcal{H}$. We then turn this property of *not-vanishing-everywhere on* $\mathcal{H}$ into an equation in some settings. In order to formalize this, let us consider the map $\Phi_{\mathcal{H}}$ defined on the set of all polynomials, using the hitting set $\mathcal{H}$ of a class $\mathscr{C}$, that maps any given polynomial f to its evaluations over the points in $\mathcal{H}$. It is clear from the above observation that any nonzero polynomial in the kernel of $\Phi_{\mathcal{H}}$ is guaranteed to be outside $\mathscr{C}$. Thus, if there were a nonzero polynomial that vanishes on all polynomials $f \notin \ker(\Phi_{\mathcal{H}})$, then we would have an equation for $\mathscr{C}$.

Moreover, if such a polynomial happened to have its degree and circuit complexity polynomially bounded in its number of variables, we would have our required upper bounds. However, note that *not* being in the kernel of a linear map seems to be a tricky condition to check via a polynomial (as opposed to the complementary property of *being* in the kernel, which can be easily checked via a polynomial). To prove our theorems, we get past this issue in the setting of finite fields, and for polynomials over $\mathbb{C}$ with bounded integer coefficients.

Over a finite field $\mathbb{F}$, a univariate polynomial that maps every nonzero $x \in \mathbb{F}$ to zero and vice versa, already exists in $q(x) = 1 - x^{|\mathbb{F}|-1}$. Therefore, for a given polynomial f, the equation essentially outputs $\prod_{\mathbf{h}\in\mathcal{H}} q(f(\mathbf{h}))$. Clearly, for a polynomial f, $\prod_{\mathbf{h}\in\mathcal{H}} q(f(\mathbf{h}))$ is zero if and only if f evaluates to a nonzero value on at least one point in $\mathcal{H}$. To generalize this to other fields, we wish to find a "low-degree" univariate $q(x)$ that maps nonzero values to 0, and zero to a nonzero value. We observe when the polynomials in $\mathscr{C}$ have integer coefficients of bounded magnitude we can still obtain such a univariate polynomial, and in turn a non-trivial equation. In particular, if $q(\mathbf{x})$ were such a univariate, we essentially output $\prod_{\mathbf{h}\in\mathcal{H}} q(f(\mathbf{h}))$, for a given polynomial f. This step relies on a simple application of the Chinese Remainder Theorem.

In order to show that the equations are non-trivial, in the sense that there exist polynomials

with bounded integer coefficients which do not pass this test, we need to show that there are nonzero polynomials with bounded integer coefficients which vanish everywhere on the hitting set $\mathcal{H}$. We show this via a well-known lemma of Siegel[5], which uses a simple pigeon-hole argument to show that an under-determined system of homogeneous linear equations where the constraint matrix has small integer entries has a nonzero solution with small integer entries.

As it turns out, our proofs do not use much about the class VP except for the existence of small hitting sets for polynomials in the class (Theorem 4.3). In fact, even the existence of these hitting sets essentially follows universal circuits or polynomials, for VP (Theorem 4.2). It is not hard to observe that these properties are also true for the seemingly larger class VNP and hence the results here also extend to VNP. We also note that, given the hitting set $\mathcal{H}$ explicitly, the construction of the equation is completely explicit. In other words, the non-explicitness in our construction comes only from the fact that we do not have explicit constructions of hitting sets for algebraic circuits.

VNP-succinct hitting sets for VP

As was observed in [FSV18, GKSS17], a lower bound for equations for a class of polynomials is equivalent to showing the existence of succinctly describable hitting sets for this class. For our proof of Theorem 5.8 we show that, assuming that the permanent is sufficiently hard, the coefficient vectors of polynomials in VNP form a *hitting set* for the class VP. The connection between hardness and randomness in algebraic complexity is well known via a result of Kabanets and Impagliazzo [KI04], and we use this connection for our proof, along with some additional ideas. It is useful to note that the above-mentioned result of Kabanets and Impagliazzo [KI04] is essentially an algebraic analogue of the *hardness vs randomness* paradigm introduced by Nisan and Wigderson [NW94] in the boolean world. We briefly describe a high level sketch of our proof in a bit more detail now.

Kabanets and Impagliazzo [KI04] showed that using any explicit polynomial family $\{f_n\}$ that is sufficiently hard, one can construct (a family of) hitting set generators for VP. That is, we can construct a polynomial map $\mathsf{Gen}_f : \mathbb{F}^k \to \mathbb{F}^t$ that "fools" any small algebraic circuit C on t variables in the sense that $C(y_1, y_2, \ldots, y_t)$ is nonzero if and only if the k-variate polynomial $C \circ \mathsf{Gen}_f$ is nonzero. In a typical invocation of this result, the parameter k is much smaller than t (typically $k = \text{poly} \log t$). Thus, this gives a reduction from the question of polynomial identity testing for t-variate polynomials to polynomial identity testing for k-variate polynomials. Another related way of interpreting this connection is that if $\{f_n\}$ is sufficiently hard then Gen_f is a polynomial map whose image does not have an equation with small circuit size. Thus, assuming the hardness of the Permanent, this immediately gives us a polynomial map (with appropriate parameters) such that its image does not have an efficiently constructible equation.

For the proof of Theorem 5.8, we show that the points in the image of the map $\mathsf{Gen}_{\text{Perm}}$, can be viewed as the coefficient vectors of polynomials in VNP, or, equivalently in the terminology in [FSV18, GKSS17], that the Kabanets-Impagliazzo hitting set generator is VNP-succinct. To this end,

[5] A statement of the lemma can be found here. Refer to [Sie14] for details.

we work with a specific instantiation of the construction of the Kabanets-Impagliazzo generator where the underlying construction of combinatorial designs is based on Reed-Solomon codes. Although this is perhaps the most well known construction of combinatorial designs, there are other (and in some parameters, better) constructions known. However, our proof relies on the properties of this particular construction to obtain the succinct description. Our final proof is fairly short and elementary, and is based on extremely simple algebraic ideas and making generous use of the fact that we are trying to prove a lower bound for equations for VNP and not VP.

On algebraic "PRFs" against VP: As said above, the main proof can be summarized as saying that the Kabanets-Impagliazzo generator [KI04] applied on the symbolic permanent is VNP-succinct. Informally, this states that if the symbolic permanent is exponentially hard, then the coefficient vectors of polynomials in VNP "look random" to polynomials in VP. If the succinctness of this (or any other) generator can be improved to VP, then this would be a definitive step towards completely ruling out the existence of efficiently computable equations for VP.

1.6 Organization of the paper.

We begin with some notations and preliminaries in Section 2. In Section 3, we outline the existence of efficient hitting sets, over characteristic-zero fields, for any class that has a low-degree, low-variate *universal polynomials*; VP and VNP are examples of such classes. Existence of efficient hitting sets for VP and VNP over finite fields is outlined in Section 3.2. Note that all the arguments in Section 3 are present in previous works, but the results over characteristic zero have not been stated in this generality prior to this work, to the best of our knowledge. In Section 4, we use the above results to first prove Theorems 1.6 to 1.9. We then show the conditional hardness of equations for VNP in Section 5. Finally, we provide some open questions that arise from our work, and the prior literature around algebraic natural proofs in Section 6.

2 Notation and preliminaries

2.1 Notation and basics

- We use $[n]$ to denote the set $\{1, \ldots, n\}$ and $[\![n]\!]$ to denote the set $\{0, 1, \ldots, n\}$. We also use $\mathbb{N}_{\geq 0}$ to denote the set of non-negative integers.
- As usual, we identify the elements of $\mathbb{F}_p$ with $\{0, 1, \ldots, p-1\}$ and think of $[\![n]\!]$ as a subset of $\mathbb{F}_p$ in the natural way for any $n < p$.
- We use boldface letters such as $\mathbf{x}, \mathbf{y}$ to denote tuples, typically of variables. When necessary, we adorn them with a subscript such as $\mathbf{y}_{[n]}$ to denote the length of the tuple. We also use $\mathbf{x}^{\mathbf{e}}$ to denote the monomial $\prod x_i^{e_i}$.

- We use $\{f_n\}_{n\in\mathbb{N}}$ to denote families of polynomials. We drop the index set whenever it is clear from context. For a given polynomial f we denote by $\deg(f)$ its degree. For a polynomial $f(\mathbf{x}, \mathbf{y}, \ldots)$ on multiple sets of variables, we use $\deg_{\mathbf{x}}(f)$, $\deg_{\mathbf{y}}(f)$, etc., to denote the degree in the variables from the respective sets.

- We use $\mathbb{F}[\mathbf{x}]^{\leq d}$ to denote polynomials over the field $\mathbb{F}$ in variables $\mathbf{x}$ of degree at most d, and use $\mathbf{x}^{\leq d}$ to denote the set of all monomials in variables $\mathbf{x}$ of degree at most d.

- For a given polynomial $f \in \mathbb{F}[\mathbf{x}]^{\leq d}$ and a monomial $m \in \mathbf{x}^{\leq d}$, we use $\mathrm{coeff}_m(f)$ to refer to the coefficient of m in f. We further use $\overline{\mathrm{coeff}}(f)$ to denote the vector[6] of coefficients of f.

- We denote sets of polynomials and classes of polynomial families by two sets of calligraphic letters. Sets are denoted by $\mathcal{C}$, $\mathcal{D}$, etc., and classes of families are denoted by $\mathscr{C}$, $\mathscr{D}$, etc.

We will require the notions of hitting sets and hitting set generators (HSGs) given below.

Definition 2.1 (Hitting Set)**.** *A set of points $\mathcal{H}$ is said to be a hitting set for a set of polynomials $\mathcal{T}$, if for each $f \in T$ there exists an $h \in \mathcal{H}$ for which $f(h) \neq 0$.* ◊

Definition 2.2 (Hitting Set Generator (HSG))**.** *A vector of polynomials $(g_1(\mathbf{y}), \ldots, g_n(\mathbf{y}))$ is said to be a* hitting set generator *for a set of n-variate polynomials $\mathcal{T}$, if for each $f \in \mathcal{T}$ we have that the composed polynomial $f(g_1(\mathbf{y}), \ldots, g_n(\mathbf{y}))$ is nonzero.* ◊

We will also be using the well-known Polynomial Identity Lemma.

Lemma 2.3 (Polynomial Identity Lemma [Ore22, DL78, Sch80, Zip79])**.** *Let $f \in \mathbb{F}[x_1, x_2, \ldots, x_n]$ be a nonzero polynomial of degree at most d and let S be a subset of $\mathbb{F}$ (or an extension of $\mathbb{F}$). Then, the number of zeroes of f on the grid S^n is at most $d\,|S|^{n-1}$.*

Corollary 2.4. *For $S = \{0, 1, \ldots, d\}$, the grid S^n is a hitting set for the set of all n-variate polynomials of degree at most d.*

2.2 Algebraic circuits and complexity classes

Let us first formally define algebraic circuits.

Definition 2.5 (Algebraic circuits)**.** *An* algebraic circuit *is specified by a directed acyclic graph, with leaves (in-degree zero; also called* inputs*) labelled by field constants or variables, and internal nodes labelled by $+$ or $\times$. The nodes with out-degree zero are called the* outputs *of the circuit. Computation proceeds in the natural way, where inductively each $+$ gate computes the sum of its children and each $\times$ gate computes the product of its children.*

The size *of the circuit is defined as the number of nodes in the underlying graph.* ◊

[6] We do not explicitly mention the monomial ordering used for this vector representation, since all our statements work for any monomial ordering.

Definition 2.6 (Exponential sums of circuits)**.** *For an algebraic circuit $C(\mathbf{x}, \mathbf{y})$, the* exponential sum of C over $\mathbf{y}$ *is defined as follows.*

$$\widetilde{C}(\mathbf{x}) := \sum_{\alpha \in \{0,1\}^{|\mathbf{y}|}} C(\mathbf{x}, \mathbf{y} = \alpha)$$

The size *of such an exponential sum is said to be the sum of the size of C, and the number of "auxiliary" variables $\mathbf{y}$. Similarly, the* degree *of the exponential sum is simply the total degree (in $\mathbf{x}$ and $\mathbf{y}$) of the polynomial computed by C. For instance, the exponential sum $\widetilde{C}$ has size equal to $\text{size}(C) + |\mathbf{y}|$, and its degree is the total degree of $C(\mathbf{x}, \mathbf{y})$.* ◊

Definition 2.7 (Size of a polynomial)**.** *For a polynomial $f(\mathbf{x})$, the* circuit-size of f *is defined as the size of the smallest circuit that computes it. It is denoted as* $\text{CircuitSize}(f)$.

Similarly, the exponential-sum-size of f *is defined as the size of the smallest exponential sum that computes it. This is denoted by* $\text{ExpSumSize}(f)$. ◊

2.2.1 Polynomial families and their complexity

In order to study the asymptotic cost of computing polynomials, we work with the families of polynomials that they naturally define (e.g. $\{\det_n\}$ is the family of $n \times n$ determinants for all natural n). We formally define polynomial families and their complexity for clarity.

Definition 2.8 (Polynomial families)**.** *For a set of variables $\mathbf{x}$ and a field $\mathbb{F}$, a family of polynomials within $\mathbb{F}[\mathbf{x}]$, denoted by $\{f_n\}_{\mathbb{N}}$ (or simply $\{f_n\}$), is a set of polynomials indexed by $n \in \mathbb{N}$ such that the n-th polynomial f_n depends on at most n variables for all $n \in \mathbb{N}$.*

We shall denote the collection of all such polynomial families using $\mathcal{P}$. ◊

Remark 2.9. *It is more common to work with what are called 'p-bounded families', where the number of variables that the polynomials in the family $\{f_n\}$ depend on, grows polynomially with the index n. We choose the stricter definition above for clarity, and brevity of our statements. More importantly, all the contents of this paper can be translated to the language of p-bounded families.* ◊

The main focus of this work is going to be families of polynomials in which the degree grows polynomially with the number of variables, called *low-degree polynomial families*.

Definition 2.10 (Low-degree polynomial families)**.** *For a function $d : \mathbb{N} \to \mathbb{N}$, we define the class of all degree-d polynomial families as follows.*

$$\mathscr{P}_d := \{\{f_n\} \in \mathscr{P} : \forall n, \deg(f_n) \leq d(n)\}$$

The collection of low-degree polynomial families *is then naturally defined as the union of $\mathscr{P}_d$'s over all polynomial functions d.*

$$\mathscr{P}_{\text{low}-\deg} := \bigcup_{c \in \mathbb{N}} \mathscr{P}_{n^c}.$$ ◊

Note that every family in $\mathscr{P}_{\text{low}-\deg}$ implicitly fixes a constant $c > 0$ so that the family belongs to the class $\mathscr{P}_d$ where $d : n \mapsto n^c$. A helpful feature of $\mathscr{P}_d$ is that for each $\{f_n\}$ in $\mathscr{P}_d$, the length of the *coefficient vector* of the n^{th} polynomial f_n is exactly $\binom{n+d(n)}{n}$ and therefore is *purely a function of* n. This makes it easier to handle equations for sets of polynomials, which is the central subject of our work. Therefore, we will always work with specific subclasses of $\mathscr{P}_d$ for a fixed d while dealing with equations, and this choice of d will be mentioned explicitly whenever it is not completely clear from the context.

Complexity classes of polynomial families

We can now define classes of low-degree families that are efficiently expressible using circuits and exponential sums.

Definition 2.11 (Computable families)**.** *For functions $s : \mathbb{N} \to \mathbb{N}$ and $d : \mathbb{N} \to \mathbb{N}$, we define the class of degree-d families that are computable by size s circuits as follows.*

$$\text{Circuit}_{d,s} := \{\{f_n\} \in \mathscr{P}_d : \exists n_0 \in \mathbb{N}, \forall n > n_0, \text{CircuitSize}(f_n) \leq s(n)\}$$

When d and s are polynomials, we shall alternatively refer to this class as $\mathsf{VP}_{d,s}$. ◊

Definition 2.12 (Definable families)**.** *For functions $s : \mathbb{N} \to \mathbb{N}$ and $d : \mathbb{N} \to \mathbb{N}$, we define the class of degree-d families that are expressible by exponential sums of size and degree s, as follows.*

$$\text{ExpSum}_{d,s} := \{\{f_n\} \in \mathscr{P}_d : \exists n_0 \in \mathbb{N}, \forall n > n_0, \text{ExpSumSize}(f_n) \leq s(n)\}$$ ◊

For polynomials d, s, we shall sometimes refer to this class as $\mathsf{VNP}_{d,s}$.

We now define the familiar classes VP and VNP more formally, specifically to help the forthcoming discussion about families of equations for these classes.

Definition 2.13 (VP from first principles)**.** *For a polynomially bounded $d : \mathbb{N} \to \mathbb{N}$, we denote by VP_d, the class of all degree-d polynomial families that are efficiently computable using algebraic circuits, as follows.*

$$\mathsf{VP}_d := \cup_{e \in \mathbb{N}} \mathsf{VP}_{d,n^e} = \cup_{e \in \mathbb{N}} \text{Circuit}_{d,n^e}$$

The collection of all efficiently computable low-degree polynomial families denoted by VP, is then naturally defined as follows.

$$\mathsf{VP} := \cup_{c \in \mathbb{N}} \mathsf{VP}_{n^c}$$ ◊

Definition 2.14 (VNP from first principles)**.** *For a polynomially bounded $d : \mathbb{N} \to \mathbb{N}$, we denote by VNP_d, the class of all degree-d polynomial families that are efficiently definable using exponential sums, as*

follows.

$$\mathsf{VNP}_d := \cup_{e\in\mathbb{N}}\mathsf{VNP}_{d,n^e} = \cup_{e\in\mathbb{N}}\,\mathrm{ExpSum}_{d,n^e}$$

The collection of all efficiently definable low-degree polynomial families denoted by VNP*, is then naturally defined as follows.*

$$\mathsf{VNP} := \cup_{c\in\mathbb{N}}\mathsf{VNP}_{n^c} \qquad \Diamond$$

Dependence on the field. All the definitions discussed so far can be naturally instantiated for any field $\mathbb{F}$. We give a short summary for clarity.

- A polynomial family $\{f_n\}$ over a field $\mathbb{F}$ is a family in which the coefficients of f_n are elements in $\mathbb{F}$ for all n.
  Similarly, we can then define $\mathscr{P}_d^{\mathbb{F}}$ to be the class of degree-d families over $\mathbb{F}$.
- A circuit is said to be over a field $\mathbb{F}$, if all the constants appearing in the circuit are from $\mathbb{F}$. Exponential sums over $\mathbb{F}$ are defined similarly.
- We then define $\mathrm{Circuit}_{d,s}^{\mathbb{F}}$ and $\mathrm{ExpSum}_{d,s}^{\mathbb{F}}$ to be the analogous subclasses of families over the field $\mathbb{F}$, according to circuits and exponential sums over $\mathbb{F}$.
- Finally, $\mathsf{VP}^{\mathbb{F}}$ is defined as the union over all $c, e \in \mathbb{N}$, of the classes $\mathrm{Circuit}_{d,s}^{\mathbb{F}}$, with $d = n^c$ and $s = n^e$.
  Similarly, $\mathsf{VNP}^{\mathbb{F}}$ is the union of all $\mathrm{ExpSum}_{d,s}^{\mathbb{F}}$.

An important point to note here is that $\mathsf{VP}^{\mathbb{F}}$ and $\mathsf{VNP}^{\mathbb{F}}$ are defined for a *fixed* field $\mathbb{F}$. Particularly, in the case of finite fields, the size of the field is a constant with respect to n. The underlying field will usually be clear from the context, and will therefore not be mentioned unless required.

We will now formally define algebraic natural proofs and succinct hitting sets over characteristic zero. Defining these for finite fields adds one more subtlety of largeness, which we address after that.

2.2.2 Equations, natural proofs and succinct hitting sets over characteristic zero

To start with, we recall the definition of equations for a set of polynomials.

Definition 1.2 (Equations for a set of polynomials)**.** *For some $n, d \in \mathbb{N}$, let $\mathcal{C}_{n,d}$ be a set of n-variate polynomials of* total *degree at most d; i.e. $\mathcal{C}_{n,d} \subseteq \mathbb{F}[\mathbf{x}]^{\leq d}$.*

Then, for $N = \binom{n+d}{n}$, a nonzero polynomial $P_N(\mathbf{Z})$ is said to be an equation *for $\mathcal{C}_{n,d}$ if for all $f(\mathbf{x}) \in \mathcal{C}_{n,d}$, we have that $P_N(\overline{\mathrm{coeff}}(f)) = 0$, where $\overline{\mathrm{coeff}}(f)$ is the coefficient vector of f.* $\Diamond$

Observe that any class $\mathscr{C} \subseteq \mathscr{P}_d$ naturally defines a set of n-variate, degree-$d(n)$ polynomials: $\mathscr{C}(n) = \{f_n : \{f_n\} \in \mathscr{C}\}$. We shall use this piece of notation in order to define the concept of natural proofs for a class of polynomial families.

Definition 2.15 (Natural proofs for a class of families)**.** *For some functions $d, D : \mathbb{N} \to \mathbb{N}$, and classes $\mathscr{C} \subseteq \mathscr{P}_d$, $\mathscr{D} \subseteq \mathscr{P}_D$, we say that a family $\{P_N\}$ is a* $\mathscr{D}$-natural proof for $\mathscr{C}$*, if:*

- $\{P_N\} \in \mathscr{D}$*, and*
- *for all large enough n, P_N is an equation for $\mathscr{C}(n)$ for $N = \binom{n+d(n)}{n}$.* ◊

As first defined in the works of Forbes, Shpilka and Volk [FSV18], and Grochow, Kumar, Saks and Saraf [GKSS17], we then get the following definition for 'succinct hitting sets' by essentially negating the definition of natural proofs.

Definition 2.16 (Succinct hitting sets for a class of families)**.** *For functions $d, D : \mathbb{N} \to \mathbb{N}$, and classes $\mathscr{C} \subseteq \mathscr{P}_d$, $\mathscr{D} \subseteq \mathscr{P}_D$, we say that* $\mathscr{C}$-succinct hitting sets exist for $\mathscr{D}$*, if the following is true.*

> For infinitely many $n \in \mathbb{N}$, *the set of coefficient vectors of $\mathscr{C}(n)$ is a hitting set for the set of polynomials $\mathscr{D}(N)$ where $N = \binom{n+d(n)}{n}$.* ◊

The following statement is therefore an immediate consequence of the definitions above.

Proposition 2.17. *Over any field of characteristic zero, for classes $\mathscr{C} \in \mathscr{P}_d$ and $\mathscr{D} \in \mathscr{P}_D$, $\mathscr{C}$ has $\mathscr{D}$-natural proofs if and only if $\mathscr{D}$ does not have $\mathscr{C}$-succinct hitting sets.* □

Handling VP**,** VNP**.** We now instantiate the definitions of natural proofs and succinct hitting sets for the specific cases of VP and VNP. This needs a bit of care because both VP and VNP are collections of countably many classes of polynomial families, each containing polynomials families of a specific degree. Thus, a formal definition of say, "VP-natural proofs for VP", does not directly follow from Definition 2.15, and one has to rely on the purpose of defining the concept, which is that of analyzing the power of known techniques in the context of proving non-membership in VP. This leads us to the following definitions.

Definition 2.18 (VP-natural proofs for VP)**.** *For polynomially bounded functions $d, s, D, S : \mathbb{N} \to \mathbb{N}$, we say that a family $\{P_N\}$ is a* $\mathsf{VP}_{D,S}$-natural proof for $\mathsf{VP}_{d,s}$*, if:*

- $\{P_N\} \in \mathsf{VP}_{D,S}$*, and*
- *for all large enough n, P_N is an equation for $\mathsf{VP}_{d,s}(n)$ for $N = \binom{n+d(n)}{n}$.*

We say that VP-natural proofs exist for VP *if the following is true.*

> *For any $d(n) \in \mathrm{poly}(n)$, there exist $D(N), S(N) \in \mathrm{poly}(N)$, and a family $\{P_N\}$, such that for every $s(n) \in \mathrm{poly}(n)$, $\{P_N\}$ is a $\mathsf{VP}_{D,S}$-natural proof for $\mathsf{VP}_{d,s}$.*

In other words, for every degree-function $d(n)$, there is a single family $\left\{P_N^{(d)}\right\} \in \mathsf{VP}$, that is a family of equations for all of $\mathsf{VP}_d = \cup_s \mathsf{VP}_{d,s}$. ◊

Two important aspects of this definition must be noted.

- The proof family $\{P_N\}$ *may depend* on the degree-function $d(n)$.

  If a polynomial P_N vanishes on the coefficient vector of an n-variate degree-d polynomial f, then the length of the coefficient vector should match the arity of the equation, and therefore $N = \binom{n+d}{d}$. Therefore, it is inevitable to see a dependence between N and the degree-function $d(n)$.

- The proof family $\{P_N\}$ *must not depend* on the size-function $s(n)$.

  This is crucial for a proof that yields a super-polynomial lower bound. That is, if the proof family were to depend on the size-function $s(n)$, it will only yield size-s lower bounds: arbitrarily large polynomial lower bounds when $s(n)$ is a polynomial.

  That said, the "prover" is allowed to select a suitable "largeness threshold n_0" depending on the size-function $s(n)$. This is also easily justified. For instance, if P_N vanishes on families that have circuits of size $\leq n^{\log n}$, then this n_0 would depend on when $n^{\log n}$ overshoots $s(n)$.

We now define VP-succinct hitting sets for VP. Similar to the general definition, this has been done so that "existence of VP-natural proofs for VP" and "existence of VP-succinct hitting sets for VP" are direct logical negations of each other.

Definition 2.19 (VP-Succinct hitting sets for VP)**.** *For polynomially bounded functions $d, s, D, S : \mathbb{N} \to \mathbb{N}$, we say that* $\mathsf{VP}_{d,s}$-succinct hitting sets exist for $\mathsf{VP}_{D,S}$*, if the following is true.*

> *For* infinitely many $n \in \mathbb{N}$*, if* $N = \binom{n+d(n)}{n}$ *then the coefficient vectors of* $\mathsf{VP}_{d,s}(n)$ *form a hitting set for* $\mathsf{VP}_{D,S}(N)$*.*

We say that VP has VP-succinct hitting sets *if there exists a degree-function $d(n) \in \mathrm{poly}(n)$, such that for all choices of $D(N), S(N)\,\mathrm{poly}(N)$, there is a size-function $s(n) \in \mathrm{poly}(n)$, so that* $\mathsf{VP}_{d,s}$*-succinct hitting sets exist for* $\mathsf{VP}_{D,S}$*.* ◊

Proposition 2.20. VP *has* VP*-natural proofs* $\iff$ VP *does not have* VP*-succinct hitting sets.* □

Remark 2.21. *It can be argued that the term '*VP*-succinct hitting sets' intuitively means that the hitting set can be described efficiently using algebraic circuits. This hints towards a fixed pair of degree and size functions working for all of* VP[7]*. Clearly, such a definition would be stronger than Definition 2.19. So a 'succinct-hitting-sets-based-barrier' towards 'lower bounds via natural proofs' in this intuitive sense is also a barrier as per Definition 2.19.* ◊

Moreover, all the succinct hitting sets described in the work of Forbes, Shpilka and Volk [FSV18] are consistent with the stronger, "intuitive" definition, and therefore also imply succinct hitting sets as per Definition 2.19. More generally, as we shall see in Section 2.3, the notion of universal circuits lets us work with Definition 2.19 with a suitable change in parameters.

[7] "$\exists d(n), s(n) \forall D(N), S(N) \ldots$" as opposed to "$\exists d(n) \forall D(N), S(N) \exists s(n) \ldots$".

2.2.3 Algebraic natural proofs over finite fields

As pointed out before, the classes $\mathsf{VP}^{\mathbb{F}}$ and $\mathsf{VNP}^{\mathbb{F}}$ are defined only for a fixed field $\mathbb{F}$, and hence all of $\mathsf{VP}_d(n)$ is a subset of $\mathbb{F}^N$. We then have to rule out "nonzero" equations that vanish on the entire universe $\mathbb{F}^N$.

Taking a queue from the definition of (boolean) natural proofs (Definition 1.1), we include the largeness criterion. We first define the notion of a *large polynomial family*.

Definition 2.22 (Large polynomial families)**.** *Let $\mathbb{F} = \mathbb{F}_q$ be the finite field of size q. For a constant a, a polynomial family $\{P_N\}$ over $\mathbb{F}$ is said to be a-*large*, if for all large enough N, there is a set $W \subseteq \mathbb{F}^N$ of size at least $N^{-a}q^N$ such that $P_N(\mathbf{w}) \neq 0$ for all $\mathbf{w} \in W$.*

A family $\{P_N\}$ is said to be large, if it is a-large for some a. ◊

Definition 2.23 (VP-natural proofs for VP (Finite fields))**.** *Let $\mathbb{F}$ be a finite field. For polynomially bounded $d, s, D, S : \mathbb{N} \to \mathbb{N}$, we say that a family $\{P_N\}$ is a $\mathsf{VP}^{\mathbb{F}}_{D,S}$-*natural proof for $\mathsf{VP}^{\mathbb{F}}_{d,s}$*, if:*

- $\{P_N\} \in \mathsf{VP}^{\mathbb{F}}_{D,S}$,
- *for all large enough n, P_N is an equation for $\mathsf{VP}^{\mathbb{F}}_{d,s}(n)$ for $N = \binom{n+d(n)}{n}$, and*
- *$\{P_N\}$ is large as per Definition 2.22.*

*We say that $\mathsf{VP}^{\mathbb{F}}$-*natural proofs exist for $\mathsf{VP}^{\mathbb{F}}$ *if the following is true.*

> *For any $d(n) \in \mathrm{poly}(n)$, there exists $D(N), S(N) \in \mathrm{poly}(N)$ and a family $\{P_N\}$, such that for every $s(n) \in \mathrm{poly}(n)$, $\{P_N\}$ is a $\mathsf{VP}^{\mathbb{F}}_{D,S}$-natural proof for $\mathsf{VP}^{\mathbb{F}}_{d,s}$.*

In other words, for each degree-function $d(n)$, there is a single family $\left\{P_N^{(d)}\right\} \in \mathsf{VP}^{\mathbb{F}}$, that is a family of equations for all of $\mathsf{VP}^{\mathbb{F}}_d = \cup_s \mathsf{VP}^{\mathbb{F}}_{d,s}$. ◊

By negating this definition, we get a seemingly weaker version of succinct hitting sets; in particular, the coefficient vectors are now only expected to hit the large families within $\mathsf{VP}^{\mathbb{F}}$. Note that hitting all families in "$\mathsf{VP}^{\mathbb{F}}(N)$" is *impossible* for "$\mathsf{VP}^{\mathbb{F}}(n)$", simply because the constant-degree, univariate polynomial $(Z_1^q - Z_1)$ vanishes over all of $\mathbb{F}_q^N$.

Definition 2.24 (VP-Succinct hitting sets for VP (Finite fields))**.** *Let $\mathbb{F}$ be a finite field. For polynomially bounded $d, s, D, S : \mathbb{N} \to \mathbb{N}$, we say that $\mathsf{VP}^{\mathbb{F}}_{d,s}$-*succinct hitting sets exist for $\mathsf{VP}^{\mathbb{F}}_{D,S}$*, if the following is true.*

> *For any large family $\{P_N\} \in \mathsf{VP}^{\mathbb{F}}_{D,S}$, for* infinitely many *$n \in \mathbb{N}$, if $N = \binom{n+d(n)}{n}$ then the polynomial P_N does not vanish on the coefficient vectors of $\mathsf{VP}_{d,s}(n)$.*

We say that $\mathsf{VP}^{\mathbb{F}}$ has *$\mathsf{VP}^{\mathbb{F}}$-*succinct hitting sets *if there exists a degree-function $d(n) \in \mathrm{poly}(n)$ such that, for all choices of $D(N), S(N) \in \mathrm{poly}(N)$, there is a size-function $s(n) \in \mathrm{poly}(n)$ so that $\mathsf{VP}^{\mathbb{F}}_{d,s}$-succinct hitting sets exist for $\mathsf{VP}^{\mathbb{F}}_{D,S}$.* ◊

2.3 Universal Circuits

A universal circuit is an algebraic circuit with the property that any polynomial which is efficiently computable is a simple projection of it. The following lemma due to Raz shows the existence of such circuits; for the sake of completeness, we also include a proof sketch.

Lemma 2.25 (Universal circuit [Raz10])**.** *Let $\mathbb{F}$ be any field and $n, s \geq 1$ and $d \geq 0$. Then there exists an algebraic circuit U of size $\mathrm{poly}(n, d, s)$ computing a polynomial in $\mathbb{F}[x_1, \ldots, x_n, y_1, \ldots, y_r]$ with $r \leq \mathrm{poly}(n, d, s)$ such that:*

- $\deg_{\mathbf{x}}(U(\mathbf{x}, \mathbf{y})), \deg_{\mathbf{y}}(U(\mathbf{x}, \mathbf{y})) \leq \mathrm{poly}(d)$;
- *for any $f(\mathbf{x}) \in \mathrm{Circuit}_{d,s}(n)$, there exists an $\mathbf{a} \in \mathbb{F}^r$ such that $f(\mathbf{x}) = U(\mathbf{x}, \mathbf{a})$.*

Proof. Let f be an n-variate degree d polynomial computable by a circuit C of size s. Using the classical depth reduction result due to Valiant *et al.* [VSBR83], f has a circuit C' of size $s' = \mathrm{poly}(n, d, s)$ and depth $\ell = O(\log d)$ with the following properties (see, e.g., [Sap15] for a complete proof).

- All the product gates have fan-in at most 5.
- C' is *layered*, with alternating layers of sum and product gates.
- The layer above the leaves is of product gates, and the root is an addition gate.

We can therefore construct a *layered* universal circuit U for the given parameters n, d, s. The circuit will have ℓ layers, with $V_1, V_2, \ldots, V_\ell$ being the layers indexed from leaves to the root. So V_ℓ has a single gate, which is the output gate of the circuit, and V_1 has $n + 1$ gates, labeled with the variables $x_1, \ldots, x_n$ and with the constant 1. All the gates in U are then connected using auxiliary variables $\mathbf{y}$, as follows.

- V_2 has $\leq (n+1)^5$ product gates, with each gate computing a unique monomial of degree at most 5 in the variables $\mathbf{x}$.
- For every odd i with $2 < i < \ell$, the layer V_i has s' addition gates that are all connected to all the gates in the layer V_{i-1}, with each of the wires being labeled by a fresh $\mathbf{y}$-variable.
- For every even i with $2 < i < \ell$, the layer V_i has $\binom{s'}{5}$ product gates, each one multiplying a unique subset of 5 gates from V_{i-1}.

It is now easy to see that U has at most $\ell(ns')^5$ gates, which is $\mathrm{poly}(n, d, s)$. Also, $\deg(U) \leq 5^\ell$, which is $\mathrm{poly}(d)$; and $|\mathbf{y}| = r \leq \ell \cdot (ns')^6$, which is $\mathrm{poly}(n, d, s)$. Further, by the depth reduction result [VSBR83], the circuit C' for f can be obtained by setting the auxiliary variables $\mathbf{y}$ appropriately. Since the choice of f was arbitrary, this finishes the proof. □

Universal circuits let us move naturally from succinct hitting sets to 'succinct hitting set generators', which are defined as follows.

Definition 2.26 (Succinct Hitting Set Generators). *Let $\mathbb{F}$ be any field and $n, d, N \in \mathbb{N}$ be such that $N = \binom{n+d}{n}$. For a set of N-variate polynomials $\mathcal{D}(N)$, a degree-d polynomial $G_n \in \mathbb{F}[\mathbf{y}][x_1, \ldots, x_n]$ (seen as a polynomial only over the $\mathbf{x}$ variables) is said to be a* succinct hitting set generator for $\mathcal{D}(N)$ *if the coefficient vector of G_n in $(\mathbb{F}[\mathbf{y}])^N$ is a hitting set generator for each polynomial $Q \in \mathcal{D}(N)$.*

Naturally, for a class of families $\mathscr{D}$, a polynomial family $\{G_n\}$ is said to be a succinct hitting set generator (family) if, for infinitely many $n \in \mathbb{N}$, we have that G_n is a succinct hitting set generator for $\mathcal{D}(N)$, where n and N are related according to the degree of $\{G_n\}$.

When the family of generators $\{G_n\}$ belongs to a class $\mathscr{C}$, we say that $\mathscr{D}$ has a $\mathscr{C}$-succinct hitting set generator. ◊

Lemma 2.27 ([FSV18, GKSS17]). *For functions $d(n), s(n), D(N), S(N)$, if $\mathsf{VP}_{D,S}$ has $\mathsf{VP}_{d,s}$-succinct hitting sets then there is a family $\{U_n\}$, where U_n is the universal circuit for parameters $n, d(n), s(n)$ guaranteed by Lemma 2.25, that is a succinct hitting set generator for $\mathsf{VP}_{D,S}$.*

Proof. Let $\{Q_N\} \in \mathsf{VP}_{D,S}$ be a family. We will show that for infinitely many $N \in \mathbb{N}$, the composition $Q_N(\overline{\text{coeff}}(U_n))$ gives a nonzero polynomial in $\mathbf{y}$.

Firstly, since $\mathsf{VP}_{D,S}$ has $\mathsf{VP}_{d,s}$-succinct hitting sets, there is some polynomial family $f_n \in \mathsf{VP}_{d,s}$ such that for infinitely many $N \in \mathbb{N}$, $Q_N(\overline{\text{coeff}}(f_n)) \neq 0$. From Lemma 2.25 we get that *for all large enough $n \in \mathbb{N}$*, there is an assignment $\mathbf{a}_n$ for which $U_n(\mathbf{x}, \mathbf{y} = \mathbf{a}_n) = f_n(\mathbf{x})$.

Therefore, in particular, $Q_N(\overline{\text{coeff}}(U_n(\mathbf{y} = \mathbf{a}_n))) = Q_N(\overline{\text{coeff}}(U_n))(\mathbf{y} = \mathbf{a}_n) \neq 0$ and this implies that the composition $Q_N(\overline{\text{coeff}}(U_n))$ is a nonzero polynomial. This finishes the proof. □

Lemma 2.28. *Let $t(n) \in n^{\omega(1)}$ be any function. If VP has VP-succinct hitting sets, then there exists a polynomially bounded $d(n)$ such that the family of n-variate, degree-$d(n)$ universal circuits of size $t(n)$, forms a succinct hitting set generator for VP.*

As a result, VP has succinct hitting set generators of size $\text{poly}(t(n))$.

Proof. Let us assume that VP has VP-succinct hitting sets. Then there exists a $d(n) \in \text{poly}(n)$ such that for any functions $D(N), S(N) \in \text{poly}(N)$, there exists a size-function $s(n) \in \text{poly}(n)$ such that $\mathsf{VP}_{D,S}$ has $\mathsf{VP}_{d,s}$-succinct hitting sets.

Since $t(n) \in n^{\omega(1)}$, there is a finite $n_0 \in \mathbb{N}$, beyond which any polynomial in $\mathsf{VP}_{d,s}(n)$ can be simulated using the "$(d(n), t(n))$-universal circuit" U_n. This means that U_n is a hitting set generator for $\mathsf{VP}_{D,S}(N)$ infinitely often, which implies that the family $\{U_n\}$ is a succinct hitting set generator family for $\mathsf{VP}_{D,S}$ as per Definition 2.26.

As the above argument goes through for any $s(n) \in \text{poly}(n)$, the family of "$(d(n), t(n))$-universal circuits" is a succinct hitting set generator for VP. The complexity of this family follows directly from Lemma 2.25. □

2.4 Hardness-Randomness Connections

We will need the following notion of combinatorial designs (a collection of subsets of a universe with small pairwise intersection).

Definition 2.29 (Combinatorial designs)**.** *A family of sets $\{S_1, \ldots, S_N\} \subseteq [\ell]$ is said to be an (ℓ, m, n)-design if*

- *$|S_i| = m$ for each $i \in [N]$*
- *$|S_i \cap S_j| < n$ for any $i \neq j$.* ◊

Kabanets and Impagliazzo [KI04] obtain hitting set generators from polynomials that are hard to compute for algebraic circuits. The following lemma is crucial to the proof of Theorem 5.8.

Lemma 2.30 (HSG from Hardness [KI04])**.** *Let $\{S_1, \ldots, S_N\}$ be an (ℓ, m, n)-design and $f(\mathbf{x}_m)$ be an m-variate, individual degree d polynomial that requires circuits of size s. Then for fresh variables $\mathbf{y}_\ell$, the polynomial map* KI-gen$_{(N,\ell,m,n)}(f) : \mathbb{F}^\ell \to \mathbb{F}^n$ *given by*

$$(f(\mathbf{y}_{S_1}), \ldots, f(\mathbf{y}_{S_N})) \tag{2.31}$$

is a hitting set generator for all N-variate polynomials with degree and circuit-size at most $\left(\frac{s^{0.1}}{N(d+1)^n}\right)$. □

2.5 Some Algebraic-Geometric Concepts

We will also use some concepts from algebraic geometry. We provide some intuition for these, which should be sufficient to understand the results here. For formal definitions of these concepts, the reader can refer to any algebraic geometry text (e.g. [CLO07]).

- *Closed set* or *Variety*: A set of points S in $\mathbb{C}^n$ is called a closed set if there exists a finite set of n-variate polynomials $\{f_1, \ldots, f_r\}$ such that S is exactly the set of common zeroes (roots) of $f_1, \ldots, f_r$. Such a set is sometimes referred to as a *variety*, and we shall do that for the remainder of this section[8].
- *(Zariski) Closure of a set*: The closure of a set S is the smallest variety that contains it.
- *Dimension of a variety*: Some varieties are clearly a single 'component'; e.g. zeroes of the polynomials $\{z, x - y^2\}$ in $\mathbb{R}^3$. In such a case, their dimension is the dimension of this component (1, in the above example).
  When a variety can be seen as a union of several components, its dimension is that of the component with the largest dimension. For example, the dimension of the zeroes of the set $\{xz, yz\}$ in $\mathbb{R}^3$ is 2.
- *Degree of a variety*: The degree of a variety is the maximal (but finite) number of intersections that it can have with a linear affine subspace (common zeroes of a set of linear polynomials). Here, it might be helpful to think of zeroes of a single bivariate polynomial of the form $y - f(x)$ to understand the nomenclature.

We need the following two inequalities about the degree of intersections of varieties.

[8] Some works reserve the term 'variety' to refer to 'irreducible' closed sets; this distinction will not be important here.

Lemma 2.32 (Bezout's inequality [HS80a, Lemma 2.2]). *Let V, W be varieties. Then the degree of their intersection* $\deg(V \cap W) \leq \deg(V) \cdot \deg(W)$. □

Any finite intersection of varieties is also a variety, and hence Lemma 2.32 can be used to bound the degree of any finite intersection. However, when we additionally have a bound on the dimension of one of the varieties, one can prove a different bound, which is sometimes tighter.

Lemma 2.33 ([HS80a, Proposition 2.3]). *For varieties $V_1, V_2, \ldots, V_k$, we have the following.*

$$\deg(V_1 \cap V_2 \cap \cdots \cap V_k) \leq \deg(V_1) \cdot \left(\max_{i>1} \deg V_i \right)^{\dim V_1}$$

Proof sketch. The proof is by induction. The base case $k = 1$ is trivial, so assume it is true for $k-1$. For simplicity, suppose that V_1 is irreducible. Then either $V_1 \cap V_2 = V_1$, which takes us back to the $(k-1)$ case. Or, $W = V_1 \cap V_2$ has dimension that is at most $(\dim V_1 - 1)$ and degree that is at most $\deg(V_1) \cdot \deg(V_2)$, and again we can apply te induction hypothesis to $W, V_3, \cdots, V_k$. In the general case, we apply the lemma to all the irreducible components of V_1. □

3 Existence of Hitting Sets

3.1 Over complex numbers

We state the results for complex numbers, but they extend as is for rationals and reals.

All the ideas required to prove the following theorem exist in previous works [HS80b, HS80a], but the statements have not been worded in that generality before, so we state it here. We also sketch the relevant proofs for completeness.

Theorem 3.1 (Universal polynomials imply hitting sets). *Let $\mathcal{C} \subseteq \mathbb{C}[x_1, \ldots, x_n]$ be a set of polynomials of degree at most d, and suppose for $N = \binom{n+d}{d}$ there is a* universal *polynomial $U(y_1, \ldots, y_m)(\mathbf{x})$ of total degree D that generates all polynomials from $\mathcal{C}$. That is, for each $f \in \mathcal{C}$, there is an $\alpha \in \mathbb{C}^m$ such that $U(\mathbf{y} = \alpha)(\mathbf{x}) = f(\mathbf{x})$.*

Then there exists a set $H \subset [10m]^n$ of size at most $(D \cdot (d+1)^2)$ which is a hitting set for $\mathcal{C}$.

The key ingredient in proving the above theorem is the following theorem from the work of Heintz and Schnorr [HS80a]. It bounds dimension and degree of the variety that contains the coefficient vectors of a set of polynomials $\mathcal{C}$ in terms of the coefficient-generating-map for $\mathcal{C}$, and we note that even though the theorem is originally stated for polynomials that are computable by small algebraic circuits, the proof only uses the properties of the coefficient generating map. It closely follows the arguments in [HS80b, Lemma 1].

Theorem 3.2 (Rewording of [HS80a, Basic Theorem 3.2]). *Let $\mathcal{C} \subseteq \mathbb{C}[x_1, \ldots, x_n]$ be a set of polynomials of total degree at most d, and let $N = \binom{n+d}{n}$ be the length of the coefficient vectors of $\mathcal{C}$.*

Suppose that there is a polynomial map $\mathcal{Q} : \mathbb{C}^m \to \mathbb{C}^N$ given by polynomials $\{Q_{\mathbf{e}}\}$, such that for each $f(\mathbf{x}) \in \mathcal{C}$, there exists an $\alpha \in \mathbb{C}^m$, so that $f(\mathbf{x}) = \sum_{\mathbf{e}} Q_{\mathbf{e}}(\alpha) \cdot \mathbf{x}^{\mathbf{e}}$. Then for $W \subseteq \mathbb{C}^N$ being the closure of the set of coefficient vectors of $\mathcal{C}$, we have that:

- *The dimension,* $\dim W \leq m$.
- *The degree,* $\deg(W) \leq (\deg \mathcal{Q})^{\dim W}$, *where* $\deg(\mathcal{Q}) := \max_{\mathbf{e}} \deg(Q_{\mathbf{e}})$.

Proof sketch. That $\dim W \leq m$ follows from the fact that it is the image of an m-variate map. Now, let $H_1, H_2, \ldots, H_m$ be hyperplanes in $\mathbb{C}^N$ such that the intersection $W_0 := W \cap H_1 \cap H_2 \cap \cdots \cap H_m$ is a finite set of size equal to $\deg(W)$.

Note that $\mathcal{Q}^{-1}(H_1), \mathcal{Q}^{-1}(H_2), \ldots, \mathcal{Q}^{-1}(H_m)$ are *hypersurfaces* in $\mathbb{C}^m$; that is, they are the zeroes of multivariate polynomials in m variables, and their intersection is a finite set. Further, each of these polynomials, and therefore the hyper-surfaces seen as varieties, have degree at most $\deg(\mathcal{Q})$. We can now bound the size of their intersection, say $V := \mathcal{Q}^{-1}(H_1) \cap \mathcal{Q}^{-1}(H_2) \cap \cdots \cap \mathcal{Q}^{-1}(H_m)$, using Lemma 2.32. Thus, $|V| \leq \prod_{i \in [m]} \deg(\mathcal{Q}^{-1}(H_i)) \leq (\deg \mathcal{Q})^m$. Finally, as $\mathcal{Q}$ is a function (and not a relation) from V to W_0, we get that $\deg(W) \leq |V| \leq \deg(\mathcal{Q})^m$. □

Heintz and Schnorr [HS80a] then use these bounds to derive the existence of hitting sets of small size and bit-complexity, as follows.

Theorem 3.3 (Rewording of [HS80a, Theorem 4.4]). *Let $\mathcal{C} \subseteq \mathbb{C}[x_1, \ldots, x_n]$ be a set of degree-d polynomials, and let $\mathcal{Q} : \mathbb{C}^m \to \mathbb{C}^N$ be a polynomial map such that the set $S = \left\{\overline{\text{coeff}}(f) : f \in \mathcal{C}\right\}$ is contained inside its image: $\mathcal{Q}(\mathbb{C}^m)$.*

Then for $b = \deg(\mathcal{Q}) \cdot (d+1)^2$, and $t = 10 \cdot \dim(W)$, there exists a set of t points $\{\mathbf{a}_1, \ldots, \mathbf{a}_t\} \subseteq [b]^n$ that is a hitting set for $\mathcal{C}$.

Proof sketch. We work with sequences of points from $[b]^n$ of length t, instead of subsets of size t; any "hitting sequence" clearly corresponds to a hitting set.

The key steps in the proof are then as follows.

- The variety of coefficient vectors of $\mathcal{C}$: This is just a direct application of Theorem 3.2. The resulting variety W has $\dim(W) \leq m = t/10$ and $\deg(W) \leq \deg(\mathcal{Q})^m = \deg(\mathcal{Q})^{t/10}$.

- The variety of "bad sequences": We first work with t-length sequences of points from the entire space $\mathbb{C}^n$. So consider the $(nt + N)$-dimensional space, where we identify the first $n \cdot t$ coordinates with sequences of t many, n-dimensional points, and the rest with coefficients of polynomials. Within this space, consider the closure $\mathcal{B}$ of the set of points that contain "bad sequences". A sequence $\mathbf{a}_1, \ldots, \mathbf{a}_t$ is bad, if there is a *nonzero* polynomial $f \in \mathcal{C}$ satisfying $f(\mathbf{a}_1) = \cdots = f(\mathbf{a}_t) = 0$. Let $\mathcal{B} \subset \mathbb{C}^{nt}$ be the Zariski closure of the set of all bad sequences.

  Now, $\dim(\mathcal{B}) \leq t(n-1) + \dim W \leq nt - t + t/10$. To see this, consider a projection from $\mathcal{B}$ to the first nt coordinates. The image of this projection — just the bad sequences — has dimension at most $t(n-1)$, since fixing some $(n-1)$ coordinates in each of the points would

give us a univariate with finitely many choices for the last coordinate. Then, the dimension of the "pre-image" $\mathcal{B}$ can be at most the sum of the dimension of the image, and the dimension of the other coordinates, which are contained in W.

In other (very informal) words, each bad sequence can essentially be "described by at most $(n-1)t + \dim W$ complex numbers".

Further, we can also obtain a simple upper-bound on the degree of $\mathcal{B}$. Consider t different equations, each involving $N+n$ coordinates. Each of these equations asserts that the point given by the n coordinates is a zero of the polynomial specified by the other N coordinates. Then, $\mathcal{B}$ is the intersection of the t hyper-surfaces given by these equations, and W. Since each of these equations have degree $(d+1)$, the degree of W, by Lemma 2.32, is at most $\deg(W) \cdot (d+1)^t$.

- Variety of bad sequences, inside the grid: With these bounds on $\mathcal{B}$, we can now bound the number of bad sequences within $[b]^n$ from above.

  For each of the nt coordinates, consider the polynomial $(z_{i,j}-1)(z_{i,j}-2)\cdots(z_{i,j}-b)$ of degree b with $i \in [t], j \in [n]$, and define its variety — a hyper-surface — $V_{i,j}$. Further, let $\mathcal{B}' := \mathcal{B} \cap V_{1,1} \cap \cdots \cap V_{t,n}$. Using Lemma 2.33, the degree of B' is then at most

$$\begin{aligned}
(\deg \mathcal{B}) \cdot b^{\dim \mathcal{B}} &\leq (\deg W) \cdot (d+1)^t \cdot b^{nt+t/10-t}, \\
&\leq \deg(\mathcal{Q})^{t/10} \cdot (d+1)^t \cdot b^{nt} \cdot b^{-9t/10}, && (\deg W \leq \deg(\mathcal{Q})^{t/10}) \\
&\leq \left(\frac{b}{(d+1)^2}\right)^{t/10} \cdot (d+1)^t \cdot b^{nt} \cdot b^{-9t/10}, && \left(\deg(\mathcal{Q}) = \frac{b}{(d+1)^2}\right) \\
&\leq b^{nt} \cdot b^{-4t/5} \cdot (d+1)^{4t/5}, \\
&\ll b^{nt}. && (b/(d+1) \gg 1)
\end{aligned}$$

  Since $\mathcal{B}'$ is a finite set, this is a bound on its size. Thus, most of the sequences of t points within $[b]^n$ are "hitting sequences", and therefore give valid hitting sets for $\mathcal{C}$. □

We now prove Theorem 3.1 by combining Theorem 3.3 and elementary multivariate interpolation.

Theorem 3.1 (Universal polynomials imply hitting sets)**.** *Let $\mathcal{C} \subseteq \mathbb{C}[x_1, \ldots, x_n]$ be a set of polynomials of degree at most d, and suppose for $N = \binom{n+d}{d}$ there is a* universal *polynomial $U(y_1, \ldots, y_m)(\mathbf{x})$ of total degree D that generates all polynomials from $\mathcal{C}$. That is, for each $f \in \mathcal{C}$, there is an $\alpha \in \mathbb{C}^m$ such that $U(\mathbf{y} = \alpha)(\mathbf{x}) = f(\mathbf{x})$.*

Then there exists a set $H \subset [10m]^n$ of size at most $(D \cdot (d+1)^2)$ which is a hitting set for $\mathcal{C}$.

Proof. Due to Theorem 3.3, all that remains to be done is to obtain a coefficient-generating-map $\mathcal{Q}$ from the given universal map $U(\mathbf{y})(\mathbf{x})$. Such a map is a direct consequence of the following fact, coming from multivariate interpolation.

Claim 3.4. *Given an n-variate, degree-d polynomial $f(\mathbf{x})$, any coefficient $\mathrm{coeff}_{\mathbf{e}}(f)$ can be expressed as a linear combination of its evaluations on the $(d+1)^n$ points in the set $[d+1]^n$.* □

The polynomial map $\mathcal{Q}$ is then just an appropriately ordered sequence of linear combinations of the evaluations of U over $[d+1]^n$. As evaluations and linear combinations do not increase the degree or the number of variables, we can invoke Theorem 3.3 with $m = m$ and $\deg(\mathcal{Q}) = \deg(U) = D$. Therefore, the hitting set uses $r \leq 10m$ points from $[t]^n$ for $t \leq \deg(\mathcal{U}) \cdot (d+1)^2 \leq D \cdot (d+1)^2$, as claimed. □

We now show the existence of a "universal polynomial" for exponential sums. This is a fairly easy extension of Lemma 2.25, as seen below.

Lemma 3.5 (Universal Polynomial for Exponential Sums)**.** *Let $s \geq n \geq 1$ and $d \geq 0$. Then for $N = \binom{n+d}{n}$ there exists a polynomial $V(y_1, \ldots, y_m)(\mathbf{x})$ with $m \leq \text{poly}(n, d, s)$ such that:*

- $\deg(V) \leq \text{poly}(s)$;
- *for any $f \in \mathbb{C}[x_1, \ldots, x_n]$ with $\deg_{\mathbf{x}}(f) \leq d$ that can be written as an exponential sum of size s, there exists a vector $\mathbf{a} \in \mathbb{C}^m$ such that $f(\mathbf{x}) = V(\mathbf{y} = \mathbf{a})(\mathbf{x})$.*

Proof. Suppose $f_n(\mathbf{x})$ is an n-variate, degree-d polynomial that can be expressed as an exponential sum of size at most s. Then by Definition 2.6, there exists an s-variate, degree s polynomial $g_s(\mathbf{x}, \mathbf{z}) \in \text{Circuit}_{s,s}(s)$ such that f_n is obtained as a sum of the polynomials given by all the $\{0,1\}$-assignments to the $\mathbf{z}$ variables, in g_s.

Using Lemma 2.25 for number of variables, degree and size, all bounded by s, we get a universal circuit $U((\mathbf{x}, \mathbf{z}), \mathbf{y})$ for $\text{Circuit}_{s,s}(s)$ with $|\mathbf{y}| = m \leq s^k$ for some constant k. Furthermore, $\deg_{\mathbf{y}}(U) \leq s^k$.

The universal polynomial $V(\mathbf{y})(\mathbf{x})$ is then just the sum of all the 2^{s-n} many $\{0,1\}$-assignments to the $\mathbf{z}$-variables in U, and hence $\deg(V) \leq \deg(U) = \text{poly}(s)$. □

We are now ready to prove the existence of hitting sets for circuits and exponential sums, whose sizes and bit-lengths, grow polynomially in the sizes of the corresponding models.

Lemma 3.6 (Hitting sets for efficiently computable polynomials [HS80a])**.** *There are constants c and e such that, there are (non-explicit) hitting sets $\mathcal{H}$ for $\text{Circuit}_{d,s}(n)$ (the set of all n-variate polynomials with degree at most d that are computable by algebraic circuits of size at most s) with $\mathcal{H} \subset [(nds)^c]^n$ and $|\mathcal{H}| = (nds)^e$.*

Proof. First, Lemma 2.25 provides a universal polynomial $U(\mathbf{x}, \mathbf{y})$ for all the polynomials in the set $\text{Circuit}_{d,s}(n)$ of degree at most $(nds)^{c_1}$, and $m := |\mathbf{y}| \leq (nds)^{c_1}$. Then, invoking Theorem 3.1 for this polynomial map finishes the proof. □

Replacing Lemma 2.25 in the above argument by Lemma 3.5 then gives us an analogous statement for exponential sums.

Lemma 3.7 (Hitting sets for efficiently definable polynomials)**.** *There are constants c' and e' such that, there are (non-explicit) hitting sets $\mathcal{H}$ for $\text{ExpSum}_{d,s}(n)$ with $\mathcal{H} \subset [(nds)^{c'}]^n$ and $|\mathcal{H}| = (nds)^{e'}$.* □

3.2 Finite fields: Hitting Sets for VP and VNP

Lemma 3.8 (Folklore (cf. Forbes [For14, Lemma 3.2.14])). *Let $\mathbb{F}$ be a finite field with $|\mathbb{F}| \geq d^2$. Let $\mathcal{C}(n,d,s)$ be the class of polynomials in $\mathbb{F}[x_1,\ldots,x_n]$ of degree at most d that are computable by fan-in 2 algebraic circuits of size at most s. Then, there is a non-explicit hitting set for $\mathcal{C}$ of size at most $\lceil 2s\cdot(\log n + 2\log s + 4)\rceil$.*

3.3 Finite fields: Hitting Sets for VNP

Lemma 3.9. *Let $\mathbb{F}$ be a finite field with $|\mathbb{F}| \geq d^2$. Let $\mathcal{D}(n,d,s)$ be the class of polynomials in $\mathbb{F}[x_1,\ldots,x_n]$ of degree at most d that are s-definable. Then, there is a non-explicit hitting set $\mathcal{H}$ for $\mathcal{D}(n,d,s)$ of size at most $\lceil 2s\cdot(3\log s + 4)\rceil$.*

Proof. In order to prove the existence of a hitting set for the class $\mathcal{D}(n,d,s)$, we will need a bound on the number of polynomials in the class $\mathcal{D}(n,d,s)$ as well as a bound on the size of an explicit hitting set for the class of n-variate degree at most d polynomials. These two bounds are summarized in the following claims, proofs of which can be found in [For14].

Claim 3.10 (Lemma 3.1.6 in [For14]). *Let $\mathbb{F}$ be a finite field and $n,s \geq 1$. There are at most $(8n\,|\mathbb{F}|\,s^2)^s$ n-variate polynomials in $\mathbb{F}[\mathbf{x}]$ computable by (single-output) algebraic circuits of size $\leq s$ and fan-in ≤ 2.*

Claim 3.11 (Lemma 3.2.13 in [For14]). *Let $\mathbb{F}$ be a finite field with $|\mathbb{F}| \geq (1+\varepsilon)d$. Let $\mathcal{C} \subseteq \mathbb{F}[\mathbf{x}]$ be a finite set of n-variate polynomials of degree $< d$. Then there is a non-explicit hitting set for $\mathcal{C}$ of size $\leq \lceil \log_{1+\varepsilon} |\mathcal{C}| \rceil$.*

Note that by definition, the number of n-variate polynomials that are s-definable is at most the number of polynomials in $\mathcal{C}(s,s,s)$; the class of s-variate polynomials of degree $\leq s$ computable by size s algebraic circuits of fan-in ≤ 2. Thus, by Claim 3.10, $|\mathcal{D}(n,d,s)| \leq (8\,|\mathbb{F}|\,s^3)^s$.

The rest of the proof follows exactly along the lines of the proof of Lemma 3.2.14 in [For14].

As $|\mathbb{F}| \geq d^2$, we have $d \leq |\mathbb{F}|$, and so $|\mathbb{F}| \geq (1+\varepsilon)d$ for $(1+\varepsilon) = \sqrt{|\mathbb{F}|}$. Thus, using $\varepsilon = \sqrt{|\mathbb{F}|} - 1$ in Claim 3.11, we get that there is a non-explicit hitting set $\mathcal{H}$ for $\mathcal{D}(n,d,s)$ of size at most

$$\left\lceil \log_{\sqrt{|\mathbb{F}|}} |\mathcal{D}(n,d,s)| \right\rceil \leq \left\lceil \log_{\sqrt{|\mathbb{F}|}} (8\,|\mathbb{F}|\,s^3)^s \right\rceil = \left\lceil s(2 + 2\log_{|\mathbb{F}|}(8s^3)) \right\rceil = \left\lceil s(2 + 6\log_{|\mathbb{F}|}(2s)) \right\rceil$$

Finally, as $|\mathbb{F}| \geq 2$, we have

$$|\mathcal{H}| \leq \lceil s\cdot(2+6\log(2s))\rceil = \lceil 2s\cdot(1+3\log(2s))\rceil = \lceil 2s\cdot(3\log s + 4)\rceil\,.$$

This completes the proof. □

4 Equations in the Bounded Coefficients setting

We will now prove our results about the existence of efficiently computable families of equations. We begin with the results over finite fields, as their proofs are slightly simpler.

4.1 Over finite fields

Theorem 4.1 (Hitting sets give equations)**.** *Let $\mathbb{F}$ be a finite field of size q, and let n be large enough. For some $d \geq 1$, let $\mathbb{K}$ be an extension of $\mathbb{F}$, of size at least d^2. Let $\mathcal{C} \subseteq \mathbb{F}[x_1, \ldots, x_n]$ be a set of polynomials of degree at most d, and let $\mathcal{H} \subset \mathbb{K}^n$ be a hitting set for $\mathcal{C}$, of size $|\mathcal{H}| = t$.*

Then, for $N = \binom{n+d}{d}$, there is an equation $P_N(Z_1, \ldots, Z_N) \not\equiv 0$ *for $\mathcal{C}'$, with* $\deg(P_N)$ *and* $\text{size}(P_N)$ *at most* $10 \cdot q \cdot N \cdot t \cdot (\log d)$.

Proof. Let $r_d = [\mathbb{K} : \mathbb{F}] = a \cdot \log d$, for some $a \leq 2$. Note that the elements of $\mathbb{K}$ can also be interpreted as vectors over $\mathbb{F}$ via an $\mathbb{F}$-linear map $\Phi : \mathbb{K} \to \mathbb{F}^{r_d}$. We can then define for any $i \in [r_d]$, $\Phi_i : \mathbb{K} \to \mathbb{F}$ to be its projection to the i-th coordinate. That is, $\Phi_i : \alpha \mapsto (\Phi(\alpha))_i$ for every $i \in [r_d]$.

For $N = \binom{n+d}{n}$, let us index the set $[N]$ by the set $\mathbf{x}^{\leq d}$ of n-variate monomials of degree at most d. For a point $\mathbf{a} \in \mathcal{H}$, we define the vector $\text{eval}(\mathbf{a}) \in \mathbb{K}^N$ as $\text{eval}(\mathbf{a})_m = m(\mathbf{a})$ where $m \in \mathbf{x}^{\leq d}$ (that is, the m-th coordinate is the evaluation of the monomial m at $\mathbf{a}$). To get vectors over $\mathbb{F}$ instead, for each $i \in [r_n]$, we shall define $\text{eval}(\mathbf{a})^{(i)} \in \mathbb{F}^N$ as $\text{eval}(\mathbf{a})^{(i)}_m = \Phi_i(m(\mathbf{a}))$.

We are now ready to define the polynomial P_N.

$$P_N(z_m \ : \ m \in \mathbf{x}^{\leq d}) := \text{OR}(\mathbf{z}) \cdot \prod_{\mathbf{a} \in \mathcal{H}} \left(\prod_{i=1}^{r_d} \left(1 - \left(\sum_m z_m \cdot \text{eval}(\mathbf{a})^{(i)}_m \right)^{|\mathbb{F}|-1} \right) \right),$$

$$\text{where } \text{OR}(\mathbf{z}) = \left(1 - \prod_{m \in \mathbf{x}^{\leq d_n}} \left(1 - z_m^{|\mathbb{F}|-1} \right) \right)$$

Constructibility: Note that $\deg(P_N) \leq |\mathbb{F}| \cdot (N + (|\mathcal{H}| \cdot r_d)) \leq q \cdot (N + t \cdot 2 \log d)$ and the above expression immediately yields a circuit for P_N of size that is at most $4 \cdot q \cdot t \cdot r_d \cdot N$ for all large enough N.

Usefulness: Now consider any polynomial $f \in \mathcal{C}$, we will show that $P_N(\overline{\text{coeff}}(f)) = 0$.

For any polynomial $g \in \mathbb{F}[x_1, \ldots, x_n]$ with $\deg(g) \leq d$, we have

$$\begin{aligned} P(\overline{\text{coeff}}(g)) &= \text{OR}(\overline{\text{coeff}}(g)) \cdot \prod_{\mathbf{a} \in \mathcal{H}} \left(\prod_{i=1}^{r_d} \left(1 - \left(\sum_m \overline{\text{coeff}}(g)_m \cdot \text{eval}(\mathbf{a})^{(i)}_m \right)^{|\mathbb{F}|-1} \right) \right), \\ &= \text{OR}(\overline{\text{coeff}}(g)) \cdot \prod_{\mathbf{a} \in \mathcal{H}} \left(\prod_{i=1}^{r_d} \left(1 - (\Phi_i(g(\mathbf{a})))^{|\mathbb{F}|-1} \right) \right), \end{aligned}$$

$$= \begin{cases} 1 & \text{if } g \neq 0 \text{ and } g(\mathbf{a}) = 0 \text{ for all } \mathbf{a} \in \mathcal{H}, \\ 0 & \text{if } g = 0 \text{ or } g(\mathbf{a}) \neq 0 \text{ for some } \mathbf{a} \in \mathcal{H}. \end{cases}$$

If $f = 0$, then $\mathrm{OR}(\overline{\mathrm{coeff}}(f)) = 0$. Else, since $f \in \mathcal{C}$, the set $\mathcal{H}$ is a hitting set for f_n. Therefore, there is some point $\mathbf{a} \in \mathcal{H}_n$ such that $f(\mathbf{a}) \neq 0$. Hence, $\{P_N\}$ vanishes on the coefficient vector of every polynomial in $\mathcal{C}$. Thus, P_N is an equation for $\mathcal{C}$. □

Theorem 1.7. *Let $\mathbb{F}$ be any finite field, and let $c > 0$ be any constant. There is a polynomial family $\left\{P_N^{(c)}\right\} \in \mathsf{VP}^{\mathbb{F}}$ such that for $N(n) = \binom{n+n^c}{n}$, the following are true.*

- *For every $t(n) = \mathrm{poly}(n)$, for all large enough n, and every family $\{f_n\} \in \mathsf{VP}^{\mathbb{F}}$, where f_n is an n-variate, degree-n^c, size-$t(n)$ polynomial, we have that*

$$P_N^{(c)}(\overline{\mathrm{coeff}}(f_n)) = 0\,.$$

- *There exists a family $\{h_n\}$ of n variate polynomials and degree $\leq n^c$ with coefficients in $\mathbb{F}$ such that for all large enough n,*

$$P_N^{(c)}(\overline{\mathrm{coeff}}(h_n)) \neq 0\,.$$

Proof. Fix c to be an arbitrary constant; we refer to the family $\left\{P_N^{(c)}\right\}$ as $\{P_N\}$ for ease of notation. The family $\{P_N\}$ is defined by constructing polynomials P_N for each n and $N(n)$ using Theorem 4.1; so let n be any large enough number. Let $d_n = n^c$, and $s_n = n^{\log n}$ (in fact, s_n can be any function that is barely super-polynomial in n). Since the size of $\mathbb{F}$ is a constant with respect to n, and we need fields of sufficiently large size for invoking Lemma 3.8, we work over an extension $\mathbb{K}_n$ of $\mathbb{F}$ of size at least n^{2c} and at most $|\mathbb{F}| \cdot n^{2c}$. By Lemma 3.8, there are hitting sets in $\mathbb{K}_n^n$ for $\mathrm{Circuit}_{d_n,s_n}(n)$ of size at most s_n^2; let $\mathcal{H}_n$ be such a hitting set.

We can now apply Theorem 4.1 for $\mathcal{C} = \mathrm{Circuit}_{d_n,s_n}(n)$, $\mathbb{K} = \mathbb{K}_n$ and $\mathcal{H} = \mathcal{H}_n$ of size $t = s_n^2$, to obtain P_N that has size and degree that is at most $10\,|\mathbb{F}| \cdot N \cdot s_n^2 \cdot (\log d_n) \leq N^2$. The family $\{P_N\}$ is therefore in $\mathsf{VP}^{\mathbb{F}}$.

Usefulness against VP_{d_n} : Let $\{f_n\} \in \mathsf{VP}_{d_n,t(n)}$ for some $t(n) = \mathrm{poly}(n)$, and $n_0 \in \mathbb{N}$ be large enough, and also be such that $n_0^{\log n_0} > t(n_0)$. Then for all $n \geq n_0$, $\mathcal{H}_n$ contains a non-root of f_n, and hence P_N vanishes on $\overline{\mathrm{coeff}}(f_n)$.

A remark on the largeness: From the definition of P_N in the proof of Theorem 4.1, any nonzero $g \in \mathbb{F}[x_1, \ldots, x_n]^{\leq d_n}$ such that $g(\mathbf{a}) = 0$ for all $\mathbf{a} \in \mathcal{H}_n$ will satisfy $P_N(\overline{\mathrm{coeff}}(g)) \neq 0$. If we interpret the coefficients of g as indeterminates, each equation of the form $g(\mathbf{a}) = 0$ introduces one homogeneous linear constraint in these N indeterminates, over the extension $\mathbb{K}_n$. Each such constraint can be interpreted as $O(\log n)$ homogeneous linear constraints, over $\mathbb{F}$. Since $|\mathcal{H}_n| \ll N$, the set of

g's that are not annihilated by P_N form a subspace of dimension at least $N - O(|\mathcal{H}_n| \log n)$. Thus, there are at least $\left(|\mathbb{F}|^{N-O(|\mathcal{H}_n| \log n)} - 1\right)$ many g's such that $P_N(\overline{\text{coeff}}(g)) \neq 0$. □

Theorem 1.9. *Let $\mathbb{F}$ be any finite field and $c > 0$ be any constant. There is a polynomial family $\left\{Q_N^{(c)}\right\} \in \mathsf{VP}^{\mathbb{F}}$ such that for $N(n) = \binom{n+n^c}{n}$, the following are true.*

- *For every $t(n) = \text{poly}(n)$, for all large enough n, and every family $\{f_n\} \in \mathsf{VNP}^{\mathbb{F}}$, where f_n is an n-variate, degree-n^c, size-$t(n)$ polynomial, we have that*

$$Q_N^{(c)}(\overline{\text{coeff}}(f_n)) = 0\,.$$

- *There exists a family $\{h_n\}$ of n variate polynomials and degree $\leq n^c$ with coefficients in $\mathbb{F}$ such that for all large n,*

$$Q_N^{(c)}(\overline{\text{coeff}}(h_n)) \neq 0\,.$$

Proof. Note that the proof of Theorem 1.7 did not use any properties of VP apart from the existence of hitting sets, which was given by Lemma 3.8. It is therefore easy to see that this theorem also follows by using Lemma 3.9 instead, with basically the same asymptotic behaviors for the size, degree, and even the largeness, of the family $\left\{Q_N^{(c)}\right\}$. □

4.2 Over rationals/complexes

We will prove the following general statement, which essentially says that the existence of "efficient" (low-variate, low-degree) universal polynomials for any class $\mathscr{C}$, yield efficient equations for the subclass of $\mathscr{C}$ containing polynomial families with bounded coefficients.

Theorem 4.2 (Equations from universal polynomials)**.** *Let $\mathcal{C} \subseteq \mathbb{C}[x_1, \ldots, x_n]$ be a set of polynomials of degree at most d, and suppose for $N = \binom{n+d}{d}$ there is a* universal *polynomial $U(y_1, \ldots, y_m)(\mathbf{x})$ of total degree D that generates all polynomials from $\mathcal{C}$. That is, for each $f \in C$, there is an $\alpha \in \mathbb{C}^m$ such that $U(\mathbf{y} = \alpha)(\mathbf{x}) = f(\mathbf{x})$.*

Then for $\mathcal{C}'$ being the set of all polynomials in $\mathcal{C}$ with coefficients in $\{-1, 0, 1\}$, there is an equation *$P_N(Z_1, \ldots, Z_N) \not\equiv 0$ for $\mathcal{C}'$, with $\deg(P_N), \text{size}(P_N) = \text{poly}(N)$.*

As mentioned earlier, the proof would also generalize in a straightforward manner for polynomial families in $\mathcal{C}$ whose coefficients are bounded by $N(n)$. We state this for coefficients in $\{-1, 0, 1\}$ just to avoid cumbersome notation.

Above theorem follows from Theorem 3.1 and the following theorem, which is the technical core of our constructions over fields of characteristic zero: Theorem 1.6 and Theorem 1.8.

Theorem 4.3 (Equations from hitting sets)**.** *Let n be large enough, and let $\mathcal{C} \subseteq \mathbb{C}[x_1, \ldots, x_n]$ be a set of polynomials of degree at most d, and let $\mathcal{H} \subseteq \{1, 2, \ldots, B\}^n$ be a hitting set for $\mathcal{C}$ of size t.*

Then, for $N = \binom{n+d}{d}$, and for $\mathcal{C}'$ being the set of all polynomials in $\mathcal{C}$ with coefficients in $\{-1,0,1\}$, there is an equation $P_N(Z_1,\ldots,Z_N) \not\equiv 0$ *for $\mathcal{C}'$, with* $\deg(P_N)$ *and* $\text{size}(P_N)$ *at most $B^4 \cdot t \cdot N^4$.*

Proof. The proof will proceed similar to the proof of Theorem 1.7, with a careful use of the Chinese Remainder Theorem.

Let $\Delta = \{-1,0,1\}$. For $N = \binom{n+d_n}{n}$, let us index the set $[N]$ by the set $\mathbf{x}^{\le d}$ of n-variate monomials of degree at most d_n. For a point $\mathbf{a} \in \mathbb{Z}^n$, we define the vector $\text{eval}(\mathbf{a}) \in \mathbb{Q}^N$ as $\text{eval}(\mathbf{a})_m = m(\mathbf{a})$ where $m \in \mathbf{x}^{\le d_n}$ (that is, the m-th coordinate is the evaluation of the monomial m at $\mathbf{a}$). Therefore, for any n-variate polynomial f of degree at most d, we have $f(\mathbf{a}) = \left\langle \overline{\text{coeff}}(f), \text{eval}(\mathbf{a}) \right\rangle$, the inner-product.

Note that for any n-variate polynomial f of degree at most d and coefficients in Δ, and any $\mathbf{a} \in \mathcal{H}$, we have $|f(\mathbf{a})| \le N \cdot B^d$, which unfortunately is exponential in the degree d. However, we can work with some "proxy evaluations" by simulating Chinese Remaindering.

For any $\mathbf{a} \in \mathcal{H}$ and a positive integer r, define the vector $\widetilde{\text{eval}}_r(\mathbf{a})$ as follows:

$$\widetilde{\text{eval}}_r(\mathbf{a})_m := (m(\mathbf{a}) \bmod r) \quad \text{for all } m \in \mathbf{x}^{\le d_n}.$$

It is to be stressed that $\widetilde{\text{eval}}_r(\mathbf{a})$ is a vector over $\mathbb{Q}$, whose entries are integers between 0 and $r-1$.

Claim 4.4. *Suppose f is a polynomial with integer coefficients, and $\mathbf{a} \in \mathbb{Z}^n$. If $f(\mathbf{a}) \neq 0$ and $|f(\mathbf{a})| \le M$, then there is some $r \le 2(\log M)^2$ such that*

$$\left\langle \overline{\text{coeff}}(f), \widetilde{\text{eval}}_r(\mathbf{a}) \right\rangle \neq 0 \bmod r.$$

Proof of claim. Let $\ell = \log(M+1)$, note that the LCM of the set $[\ell^2]$ is at least $2^\ell > M$. Since $f(\mathbf{a})$ is a nonzero integer with $|f(\mathbf{a})| \le M$, by the Chinese Remainder Theorem there is some prime $r \le \ell^2$ such that $f(\mathbf{a}) \not\equiv 0 \bmod r$.

$$\begin{aligned}\left\langle \overline{\text{coeff}}(f), \widetilde{\text{eval}}_r(\mathbf{a}) \right\rangle &\equiv \left\langle \overline{\text{coeff}}(f), \text{eval}_r(\mathbf{a}) \right\rangle \bmod r \\ &\equiv f(\mathbf{a}) \bmod r \not\equiv 0 \bmod r\end{aligned}$$

□

Let $M = N \cdot B^d$ and $\ell = \log(M+1)$. For any $r \in [\ell^2]$, any $\mathbf{a} \in \mathcal{H}$, and any n-variate polynomial f of degree at most d and coefficients from Δ, we have

$$\left| \left\langle \overline{\text{coeff}}(f), \widetilde{\text{eval}}_r(\mathbf{a}) \right\rangle \right| \le N \cdot \ell^2 =: R.$$

We are now ready to define the polynomial family $\{P_N\}$.

$$\begin{aligned}P_N(z_m \ : \ m \in \mathbf{x}^{\le n}) &= \text{OR}(\mathbf{z}) \cdot \prod_{\mathbf{a} \in \mathcal{H}} \prod_{r=2}^{\ell^2} Q_r\left(\left\langle \mathbf{z}, \widetilde{\text{eval}}_r(\mathbf{a}) \right\rangle \right), \\ \text{where } Q_r(x) &= \prod_{\substack{i \in [-R,\ldots,R] \\ i \bmod r \neq 0}} (x - i),\end{aligned}$$

$$\mathrm{OR}(\mathbf{z}) = 1 - \prod_{m \in \mathbf{x}^{\leq d_n}} (1 - z_m)$$

Constructibility: For our setting of the underlying parameters, $M \leq N \cdot B^d$ and thus $\ell \leq n \cdot d \cdot B$; and $R \leq N \cdot (ndB)^2 \leq B^2 \cdot N \log^4 N$. Therefore, P_N is a polynomial of degree at most $B^4 \cdot t \cdot N^2 \log^9 N$. Moreover, the above expression also shows that P_N is computable by a circuit of size at most $B^4 \cdot t \cdot N^4$. All these bounds hold for all large enough n.

Usefulness: Fix a polynomial $f_n \in \mathcal{C}'$. We need to show that $P_N(\overline{\mathrm{coeff}}(f_n)) = 0$. Note that we have $\mathrm{OR}(\overline{\mathrm{coeff}}(f_n)) \neq 0$ if f_n is nonzero, and 0 if $f_n = 0$. Hence, it suffices to show that $P_N(\overline{\mathrm{coeff}}(f_n)) = 0$ for nonzero f_n.

Since the set $\mathcal{H}$ is a hitting set for $\mathcal{C}$, we know that $f_n(\mathbf{a}) \neq 0$ for some $\mathbf{a} \in \mathcal{H}$. Therefore, for some $r \in [\ell^2]$, we have that $\left\langle \overline{\mathrm{coeff}}(f), \widetilde{\mathrm{eval}}_r(\mathbf{a}) \right\rangle$ is a nonzero integer in $\{-R, \ldots, R\}$ that is not divisible by r. Hence, we have

$$\begin{aligned} Q_r\left(\left\langle \overline{\mathrm{coeff}}(f), \widetilde{\mathrm{eval}}_r(\mathbf{a}) \right\rangle\right) = 0, \\ \implies P_N(\overline{\mathrm{coeff}}(f)) = 0. \end{aligned}$$

Thus, P_N is an equation for $\mathcal{C}'$. □

Theorem 1.6. *Let $c > 0$ be any constant. There is a polynomial family $\{P_N^{(c)}\} \in \mathsf{VP}^{\mathbb{Q}}$ such that for $N(n) = \binom{n+n^c}{n}$, the following are true.*

- *For every $t(n) = \mathrm{poly}(n)$, for all large enough n, and every family $\{f_n\} \in \mathsf{VP}^{\mathbb{C}}$, where f_n is an n-variate, degree-n^c, size-$t(n)$ polynomial with coefficients in $\{-1, 0, 1\}$, we have that*

  $$P_N^{(c)}(\overline{\mathrm{coeff}}(f_n)) = 0\,,$$

  where $\overline{\mathrm{coeff}}(f_n)$ is the coefficient vector of f_n.

- *There exists a family $\{h_n\}$ of n variate polynomials and degree $\leq n^c$ with coefficients in $\{-1, 0, 1\}$ such that for all large enough n,*

  $$P_N^{(c)}(\overline{\mathrm{coeff}}(h_n)) \neq 0\,.$$

Proof. Fix c to be an arbitrary constant; we refer to the family $\left\{P_N^{(c)}\right\}$ as $\{P_N\}$ for ease of notation. The family $\{P_N\}$ is defined by constructing polynomials P_N for each n and $N(n)$ using Theorem 4.3; so let n be any large enough number. Let $d_n = n^c$, and $s_n = n^{\log n}$ (in fact, s_n can be any function that is barely super-polynomial in n). By Lemma 3.6, there is a hitting set $\mathcal{H}_n \subseteq [B]^n$ for $\mathrm{Circuit}_{d_n, s_n}(n)$ of size at most s_n^{3e}, for $B = s_n^{3e}$ for some constant e.

We can now apply Theorem 4.3 for $\mathcal{C} = \mathrm{Circuit}_{d_n, s_n}(n)$, and $\mathcal{H} = \mathcal{H}_n$ of size $t = s_n^{3e}$, to obtain P_N that has size and degree that is at most $s_n^{15e} \cdot N^4 \leq N^5$ for all large n, for our setting of s_n. The

family $\{P_N\}$ is therefore in $\mathsf{VP}^{\mathbb{Q}}$.

Usefulness against VP_{d_n} : Let $\{f_n\} \in \mathsf{VP}_{d_n,t(n)}$ for some $t(n) = \text{poly}(n)$, and $n_0 \in \mathbb{N}$ be large enough, and also be such that $n_0^{\log n_0} > t(n_0)$. Then for all $n \geq n_0$, $\mathcal{H}_n$ contains a non-root of f_n, and hence P_N vanishes on $\overline{\text{coeff}}(f_n)$.

A remark on the largeness: From the definition of P_N in the proof of Theorem 4.3, any nonzero polynomial $g \in \mathbb{F}[x_1, \ldots, x_n]^{\leq d_n}$ such that $g(\mathbf{a}) = \left\langle \overline{\text{coeff}}(g), \text{eval}(\mathbf{a}) \right\rangle = 0$ for all $\mathbf{a} \in \mathcal{H}_n$, will satisfy $P_N(\overline{\text{coeff}}(g)) \neq 0$. In order to show that there are many such g's with coefficients in $\{-1, 0, 1\}$, we use a pigeon-hole argument, which is essentially an instance of a lemma of Siegel [Sie14]. For completeness, we include a sketch of the argument here.

Consider the map $\Gamma : \mathbb{Z}^N \to \mathbb{Z}^{|\mathcal{H}_n|}$ defined as

$$\Gamma(z_m \ : \ m \in \mathbf{x}^{\leq d_n}) := (\langle \mathbf{z}, \text{eval}(\mathbf{a}) \rangle \ : \ \mathbf{a} \in \mathcal{H}_n)$$

The map Γ is linear in the sense that $\Gamma(\mathbf{z} + \mathbf{z}') = \Gamma(\mathbf{z}) + \Gamma(\mathbf{z}')$. Consider the restriction of Γ on just $\{0,1\}^N$; the range of Γ under this restriction is $\{-M, \ldots, M\}^{|\mathcal{H}_n|}$, where $M = N \cdot B^d$. Hence, by the pigeon-hole-principle there must be some $\mathbf{b} \in \{-M, \ldots, M\}^{|\mathcal{H}_n|}$ with at least $2^N/(2M+1)^{|\mathcal{H}_n|}$ pre-images inside $\{0,1\}^N$. If $\mathbf{h}_0$ is any fixed preimage, then

$$\left\{ \mathbf{h} - \mathbf{h}_0 \in \{-1, 0, 1\}^N \ : \ \mathbf{h} \in \Gamma^{-1}(\mathbf{b}) \cap \{0,1\}^N \right\}$$

are all coefficient vectors of polynomials $g \in \mathbb{Z}[x_1, \ldots, x_n]^{\leq d_n}$ with coefficients in $\{-1, 0, 1\}$ whose coefficient vectors are not zeroes of P_N. □

It is worth mentioning that there are 3^N possible polynomials in $\mathbb{Z}[x_1, \ldots, x_n]^{\leq d_n}$ with coefficients in $\{-1, 0, 1\}$. The above remark on the largeness shows that there are $2^{N-q(n)}$ many polynomials g such that $P_N(\overline{\text{coeff}}(g)) \neq 0$; for some $q(n) = n^{O(\log n)}$.

Theorem 1.8. *Let $c > 0$ be any constant. There is a polynomial family $\{Q_N^{(c)}\} \in \mathsf{VP}^{\mathbb{Q}}$ such that for $N(n) = \binom{n+n^c}{n}$, the following are true.*

- *For every $t(n) = \text{poly}(n)$, for all large enough n, and every family $\{f_n\} \in \mathsf{VNP}^{\mathbb{C}}$, where f_n is an n-variate, degree-n^c, size-$t(n)$ polynomial with coefficients in $\{-1, 0, 1\}$, we have that*

$$Q_N^{(c)}(\overline{\text{coeff}}(f_n)) = 0\,.$$

- *There exists a family $\{h_n\}$ of n variate polynomials and $\leq n^c$ with coefficients in $\{-1, 0, 1\}$ such that for all large n,*

$$Q_N^{(c)}(\overline{\text{coeff}}(h_n)) \neq 0\,.$$

Proof. Again, the proof of Theorem 1.6 did not use any properties of VP apart from the existence of hitting sets, which was given by Lemma 3.6. This theorem then follows by using Lemma 3.7 instead, with essentially the same asymptotic behaviors for the size, degree, and even the largeness, of the family $\left\{Q_N^{(c)}\right\}$. □

5 VNP-succinct hitting sets for VP when Permanent is hard

We will now show that if the Permanent is exponentially hard, then so are all the equations for it. This is shown by constructing VNP-succinct hitting sets for VP. We first lay down the ideas behind our construction in some detail, and then formalize those ideas in the later parts.

Constructing VNP-succinct hitting sets for VP. Let us assume that for some constant $\varepsilon > 0$ and for all[9] $m \in \mathbb{N}$, Perm_m requires circuits of size 2^{m^ε}. Kabanets and Impagliazzo [KI04] showed that for every combinatorial design $\mathbf{D}$ (a collection of subsets of a universe, with small pairwise intersection) of appropriate parameters, the map

$$\mathsf{Gen}_{\mathrm{Perm}}(\mathbf{z}) = (\mathrm{Perm}(\mathbf{z}_S) \ : \ S \in \mathbf{D}),$$

where $\mathbf{z}_S$ denotes the variables in $\mathbf{z}$ restricted to the indices in S, is a hitting set generator for circuits of size $2^{o(m^\varepsilon)}$. Our main goal is to construct a *efficient exponential sum* $F(\mathbf{y}, \mathbf{z})$, such that

$$F(\mathbf{y}, \mathbf{z}) = \sum_{S \in \mathscr{D}} \mathrm{mon}_S(\mathbf{y}) \cdot \mathsf{Perm}(\mathbf{z}_S) \tag{5.1}$$

where $\mathrm{mon}_S(\mathbf{y})$ is a *bijective* map between $\mathbf{D}$ and monomials of total degree $\leq d$ in $\mathbf{y}$ variables.

By choosing parameters carefully, this would immediately imply that any equation on N-variables, for $N = \binom{n+d}{d}$, that vanishes on the coefficient vectors of polynomials in $\mathsf{VNP}_d(n)$ (which is the n^{th} slice of polynomial families in VNP_d) requires size 'super-polynomial in N'.

To show that the polynomial $F(\mathbf{y}, \mathbf{z})$ in Equation 5.1 has an efficient exponential sum, we use a specific combinatorial design. For the design $\mathscr{D}$ obtained via Reed-Solomon codes, every set in the design can be interpreted as a univariate polynomial g of appropriate degree over a finite field. The degree of g (say δ) and size of the finite field (say p) are related to the parameters of the design $\mathscr{D}$. Now,

$$F(\mathbf{y}, \mathbf{z}) = \sum_{\substack{g \in \mathbb{F}_p[v] \\ \deg(g) \leq \delta}} \left(\prod_{i=0}^{\delta} y_i^{g_i} \right) \cdot \mathrm{Perm}(\mathbf{z}_{S(g)}), \tag{5.2}$$

where $(g_0, \ldots, g_\delta)$ is the coefficient vector of the univariate polynomial g. Expressing $F(\mathbf{y}, \mathbf{z})$ in

[9] To be more precise, we should work with this condition for "infinitely often" $m \in \mathbb{N}$ and obtain that VNP does not have efficient equations infinitely often. We avoid this technicality for the sake of simplicity and the proof continues to hold for the more precise version with suitable additional care.

Equation 5.2 as a small exponential sum requires us to implement the product $\left(\prod_{i=0}^{\delta} y_i^{g_i}\right)$ as a polynomial when given the binary representation of coefficients $g_0, \ldots, g_\delta$ via a binary vector $\mathbf{t}$ of appropriate length (say r). This is done via the polynomial $\mathrm{Mon}(\mathbf{t}, \mathbf{y})$ in Section 5.1 in a straightforward manner. Furthermore, we want to algebraically implement the selection $\mathbf{z}_S$ for a set S in the combinatorial design when given the vector $\mathbf{g}$ that represents the polynomial g corresponding to S. This is implemented via the polynomial $\text{RS-Design}(\mathbf{t}, \mathbf{z})$ in Section 5.2. Finally, we have

$$F(\mathbf{y}, \mathbf{z}) = \sum_{\mathbf{t} \in \{0,1\}^r} \mathrm{Mon}(\mathbf{t}, \mathbf{y}) \cdot \mathrm{Perm}(\text{RS-Design}(\mathbf{t}, \mathbf{z}))$$

which is clearly a small exponential sum, as $\{\mathrm{Perm}_p\}$ is in VNP and polynomials $\mathrm{Mon}(\mathbf{t}, \mathbf{y})$ and $\text{RS-Design}(\mathbf{t}, \mathbf{z})$ are efficiently computable. We now provide rest of the details of our proof.

Some notation

We will be using the following additional pieces of notation for this section.

1. For a vector $\mathbf{t} = (t_1, \ldots, t_r)$, we will use the shorthand $t_{i,j}^{(a)}$ to denote the variable $t_{(i \cdot a + j + 1)}$. This would be convenient when we consider the coordinates of $\mathbf{t}$ as blocks of length a.
2. For integers a, p, we shall use $\mathrm{Mod}(a, p)$ to denote the unique integer $a_p \in [0, p-1]$ such that $a_p = a \bmod p$.

As mentioned in the overview, the strategy is to convert the hitting set generator given in (2.31) into a succinct hitting set generator. Therefore, we would like to associate the coordinates of (2.31) into coefficients of a suitable polynomial. That is, using exponential sums, we would like to build a polynomial of the form

$$g(y_1, \ldots, y_\ell, z_1, \ldots, z_t) = \sum_{m \in \mathbf{y}^{\leq d}} m \cdot f(\mathbf{z}_{S_m}),$$

with the monomials $m \in \mathbf{y}^{\leq d}$ suitably indexing into the sets of a combinatorial design. The above expression already resembles an exponential sum, and with a little care this can be made effective. We will first show that the different components of the above expression can be made succinct using the following constructions.

5.1 Building monomials from exponent vectors

For $n, r \in \mathbb{N}$, let $a = \lfloor r/n \rfloor$, and define $\mathrm{Mon}_{r,n}(\mathbf{t}, \mathbf{y})$ as follows.

$$\mathrm{Mon}_{r,n}(t_1, \ldots, t_r, y_1, \ldots, y_n) = \prod_{i=0}^{n-1} \prod_{j=0}^{a-1} \left(t_{i,j}^{(a)} y_{i+1}^{2^j} + (1 - t_{i,j}^{(a)}) \right)$$

The following observation is now immediate from the definition above.

Observation 5.3. *For any* $(e_1, \ldots, e_n) \in [\![d]\!]^n$*, we have*

$$\mathrm{Mon}_{r,n}(\mathrm{Bin}(e_1), \ldots, \mathrm{Bin}(e_n), y_1, \ldots, y_n) = y_1^{e_1} \cdots y_n^{e_n},$$

where $\mathrm{Bin}(e)$ *is the tuple corresponding to the binary representation of* e*, and* $r = n \cdot \lceil \log_2 d \rceil$*. Furthermore, the polynomial* $\mathrm{Mon}_{r,n}$ *is computable by an algebraic circuit of size* $\mathrm{poly}(n, r)$*.*

5.2 Indexing combinatorial designs algebraically

Next, we need to effectively compute the hard polynomial f on sets of variables in a combinatorial design, indexed by the respective monomials. We will need to simulate some computations modulo a fixed prime p. The following claim will be helpful for that purpose.

Claim 5.4. *For any* $i, b, p \in \mathbb{N}_{\geq 0}$*, there exists a unique univariate polynomial* $Q_{i,b,p}(v) \in \mathbb{Q}[v]$ *of degree at most* b *such that*

$$Q_{i,b,p}(a) = \begin{cases} 1 & \text{if } 0 \leq a < b \text{ and } a \equiv i \pmod p, \\ 0 & \text{if } 0 \leq a < b \text{ and } a \not\equiv i \pmod p. \end{cases}$$

Proof. We can define a unique univariate polynomial $Q_{i,b,p}(v)$ satisfying the conditions of the claim via interpolation to make a unique univariate polynomial take a value of 0 or 1 according to the conditions of the claim. Since there are b conditions, there always exists such a polynomial of degree at most b. □

For any $n, b, p \in \mathbb{N}_{\geq 0}$ with $n \geq p$, define

$$\mathrm{Sel}_{n,b,p}(u_1, \ldots, u_n, v) \triangleq \sum_{i=1}^{n} u_i \cdot Q_{i,b,p}(v).$$

Observation 5.5. *For any* $n, b, p \in \mathbb{N}_{\geq 0}$ *with* $n \geq p$*, for any* $0 \leq a < b$*, we have that*

$$\mathrm{Sel}_{n,b,p}(u_1, \ldots, u_n, a) = u_{\mathrm{Mod}(a,p)} = u_{a \bmod p}$$

The degree of $\mathrm{Sel}_{n,b,p}$ *is at most* $(b+1)$ *and can be computed by an algebraic circuit of size* $\mathrm{poly}(b)$*.*

Proof. From the definition of the univariate polynomial $Q_{i,b,p}(v)$ of degree b in Claim 5.4, $Q_{i,b,p}(a)$ outputs 1 if and only if $i = a \bmod p$. Hence, $\mathrm{Sel}_{n,b,p}(u_1, \ldots, u_n, a)$ is $u_{a \bmod p}$ and is of degree at most $(b+1)$. □

And finally, we choose a specific combinatorial design to instantiate Lemma 2.30 with.

5.3 Reed-Solomon based combinatorial designs

For any prime p and any choice of $a \leq p$, the following is an explicit construction of a (p^2, p, a)-combinatorial design of size p^a, defined as follows:

With the universe $U = \mathbb{F}_p \times \mathbb{F}_p$, for every univariate polynomial $g(t) \in \mathbb{F}_p[t]$ of degree less than a, we add the set $S_g = \{(i, g(i)) \ : \ i \in \mathbb{F}_p\}$ to the collection.

Since any two distinct univariate polynomials of degree less than a can agree on at most a points, it follows that the above is indeed a (p^2, p, a)-design.

The advantage of this specific construction is that it can be made succinct as follows. For $r = a \cdot \lfloor \log_2 p \rfloor$, let $t_1, \ldots, t_r$ be variables taking values in $\{0, 1\}$. The values assigned to $\mathbf{t}$-variables can be interpreted as a univariate over $\mathbb{F}_p$ of degree $< a$ by considering $\mathbf{t} \in \{0,1\}^r$ as a matrix with a rows and $\lfloor \log_2 p \rfloor$ columns each [10]. The binary vector in each row represents an element in $\mathbb{F}_p$. We illustrate this with an example.

$$\mathbf{t} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix} \quad \longrightarrow \quad \begin{pmatrix} 7 \\ 2 \\ 1 \\ 4 \\ 2 \end{pmatrix} \cong g(v)$$

$$\text{For } p = 11, a = 5, g(v) = 7 + 2v + v^2 + 4v^3 + 2v^4 \in \mathbb{F}_{11}[v],$$
$$\mathbf{t} \text{ is a } 5 \times 3 \text{ matrix that encodes the coefficients of } g(v).$$

Let $\mathbf{z}$ denote the p^2 variables $\left\{z_1, \ldots, z_{p^2}\right\}$, put in into a $p \times p$ matrix. Let S be a set in the Reed-Solomon based (p^2, p, a)-combinatorial design. We want to implement the selection $\mathbf{z}_S$ algebraically. In the following, we design a vector of polynomials that outputs the vector of variables $\left(z^{(p)}_{0,g(0) \bmod p}, \ldots, z^{(p)}_{p-1,g(p-1) \bmod p}\right)$. Note that as mentioned above the polynomial g can be specified via variables $t_1, \ldots, t_r$. That is,

$$\begin{aligned}
\text{RS-Design}_{p,a}(t_1, \ldots, t_r, z_1, \ldots, z_{p^2}) &\in (\mathbb{F}[\mathbf{t}, \mathbf{z}])^p \quad , \quad \text{for } r = a \cdot \lfloor \log_2 p \rfloor, \\
\text{RS-Design}_{p,a}(t_1, \ldots, t_r, z_1, \ldots, z_{p^2})_{i+1} &= \text{Sel}_{p,p^3,p}\left(z^{(p)}_{i,0}, \ldots, z^{(p)}_{i,p-1}, R_{i,a,p}(\mathbf{t})\right), \quad \text{for each } i \in \mathbb{F}_p, \\
\text{where } R_{i,a,p}(\mathbf{t}) &= \sum_{j=0}^{a-1} \left[\left(\sum_{k=0}^{\ell_p - 1} t^{(\ell_p)}_{j,k} \cdot 2^k \right) \cdot \text{Mod}(i^j, p) \right], \\
\text{with } \ell_p &= \lfloor \log_2 p \rfloor .
\end{aligned}$$

Observation 5.6. *For any prime p, $a \leq p$, and $\mathbf{t} \in \{0,1\}^r$ for $r = a \cdot \lfloor \log_2 p \rfloor$, we have*

$$\text{RS-Design}_{p,a}(\mathbf{t}, \mathbf{z}) = \left(z_{i,g(i)} \ : \ i \in \mathbb{F}_p\right),$$

where $g(v) \in \mathbb{F}_p[v]$ is the univariate whose coefficient vector is represented by the bit-vector $\mathbf{t}$. Furthermore, the polynomial $\text{RS-Design}_{p,a}$ *is computable by an algebraic circuit of size* $\text{poly}(p)$.

[10] Working with $\lfloor \log_2 p \rfloor$ bits (as opposed to $\lceil \log_2 p \rceil$) makes the proofs much simpler, and does not affect the size of the design by much.

Proof. Fix some $\mathbf{t} \in \{0,1\}^r$. From the definition of $R_{i,a,p}(\mathbf{t})$, it is clear that $R_{i,a,p}(\mathbf{t})$ returns an integer α such that $g(i) = \alpha \bmod p$ where $\mathbf{t}$ encodes the coefficients of the polynomial $g(t)$ in binary. Furthermore, since $\mathrm{Mod}(i^j, p)$ is the unique integer $c \in [0, p-1]$ with $c = i^j \bmod p$, it also follows that $R_{i,a,p}(\mathbf{t})$ is an integer in the range $[0, p^3]$. Hence,

$$\mathrm{Sel}_{p,p^3,p}\left(z_{i,0}^{(p)}, \ldots, z_{i,p-1}^{(p)}, R_{i,a,p}(\mathbf{t})\right) = z_{i,g(i)}$$

as claimed. □

5.4 The VNP-succinct KI generator

We are now ready to show the VNP-succinctness of the Kabanets-Impagliazzo hitting set generator family when using a hard polynomial family from VNP and Reed-Solomon based combinatorial designs.

For a prime p and for the largest number m such that $m^2 \leq p$, we will use $\mathrm{Perm}_{[p]} \in \mathbb{F}[\mathbf{y}_{[p]}]$ to denote Perm_m applied to the first m^2 variables of $\mathbf{y}$.

We now define the polynomial $F_{n,a,p}(\mathbf{y}_{[n]}, \mathbf{z}_{[p^2]})$ as follows.

$$F_{n,a,p}(y_1, \ldots, y_n, z_1, \ldots, z_{p^2}) = \sum_{\mathbf{t} \in \{0,1\}^r} \mathrm{Mon}_{r,n}(\mathbf{t}, \mathbf{y}) \cdot \mathrm{Perm}_{[p]}(\text{RS-Design}_{p,a}(\mathbf{t}, \mathbf{z})) \tag{5.7}$$
$$\text{where } r = a \cdot \lfloor \log_2 p \rfloor$$

It is evident from the above definition that the polynomial family $\{F_{n,a,p}(\mathbf{y}, \mathbf{z})\}_n$ is in VNP, for any p that is polynomially related to n, when seen as a polynomial only in the $\mathbf{y}$-variables, with coefficients from $\mathbb{C}[\mathbf{z}]$.

From the construction, we have that

$$F_{n,a,p}(y_1, \ldots, y_n, z_1, \ldots z_{p^2}) = \sum_{\mathbf{e}} \mathbf{y}^{\mathbf{e}} \cdot \mathrm{Perm}_{[p]}(\mathbf{z}_{S_{\mathbf{e}}}),$$

where $\{S_{\mathbf{e}}\}$ is an appropriate ordering of the Reed-Solomon based (p^2, p, a)-combinatorial design of size p^a, described in Section 5.3.

5.5 Putting it all together

We are now ready to show that if the Permanent is exponentially hard, then any polynomial family $\{P_N\}$ that vanishes on the coefficient vectors of all polynomials in the class VNP requires super-polynomial size to compute it.

Theorem 5.8 (Conditional Hardness of Equations for VNP)**.** *Let $\varepsilon > 0$ be a constant. Suppose, for an m large enough, we have that* Perm_m *requires circuits of size 2^{m^ε}.*

Then there is a constant c, such that for $n = m^{\varepsilon/4}$, any $d \leq n$ and $N = \binom{n+d}{n}$, we have that every nonzero polynomial $P(x_1, \ldots, x_N)$ of degree $\mathrm{poly}(N)$ *that is an equation for the set $\mathsf{VNP}_d(n)$, has*

$\text{size}(P) \geq N^{c \cdot m^\varepsilon}$.

Proof. Let p be the smallest prime larger than m^2; we know that $p \leq 2m^2$. We will again use $\text{Perm}_{[p]} \in \mathbb{C}[\mathbf{y}_{[p]}]$ to denote Perm_m acting on the first m^2 variables of $\mathbf{y}$. Therefore, if Perm_m requires size 2^{m^ε} then so does $\text{Perm}_{[p]}$.

Consider the polynomial $F_{n,n,p}(\mathbf{y}_{[n]}, \mathbf{z}_{[p^2]})$ defined in (5.7), which we interpret as a polynomial in $\mathbf{y}$ with coefficients in $\mathbb{C}[\mathbf{z}]$. The individual degree in $\mathbf{y}$ is at least d, and at most p. Let $F^{\leq d}_{n,n,p}(\mathbf{y}_{[n]}, \mathbf{z}_{[p^2]})$ denote the polynomial obtained from $F_{n,n,p}$ by discarding all terms whose *total degree* in $\mathbf{y}$ exceeds d. By standard homogenization arguments, it follows that $F^{\leq d}_{n,n,p} \in \mathsf{VNP}_d(n)$, since $F_{n,n,p}(\mathbf{y}_{[n]}, \mathbf{z}_{[p^2]})$ is efficiently computable by exponential sums. Therefore,

$$F^{\leq d}_{n,n,p}(\mathbf{y}, \mathbf{z}) = \sum_{\deg(\mathbf{y}^\mathbf{e}) \leq d} \mathbf{y}^\mathbf{e} \cdot \text{Perm}_{[p]}(\mathbf{z}_{S_\mathbf{e}}),$$

where $S_\mathbf{e}$, for various $\mathbf{e}$, is an appropriate indexing into a (p^2, p, n)-combinatorial design of size N. Since the individual degree in $\mathbf{y}$ of $F_{n,n,p}$ was at least d, every coefficient of $F^{\leq d}_{n,n,p}$ is $\text{Perm}_{[p]}(\mathbf{z}_S)$ for some S in the combinatorial design. In other words, the coefficient vector of $F^{\leq d}_{n,n,p}$ is precisely $\text{KI-gen}_{N,p^2,p,n}(\text{Perm}_{[p]})$.

Now suppose that $P(x_1, \ldots, x_N)$ is a nonzero equation for $\mathsf{VNP}_d(n)$ of degree at most N^e for some e that is independent of N. Then, in particular, it should be zero on the coefficient vector of $F^{\leq d}_{n,n,p}(\mathbf{y}, \mathbf{a}) \in \mathsf{VNP}_d(n)$ for all $\mathbf{a} \in \mathbb{C}^{p^2}$. By the Polynomial Identity Lemma [Ore22, DL78, Zip79, Sch80], this implies that P must be zero on the coefficient vector of $F^{\leq d}_{n,n,p}(\mathbf{y}, \mathbf{z}) \in (\mathbb{C}[\mathbf{z}])[\mathbf{y}]$, where coefficients are formal polynomials in $\mathbb{C}[\mathbf{z}]$. Since the coefficient vector of $F^{\leq d}_{n,n,p}(\mathbf{y}, \mathbf{z})$ is just $\text{KI-gen}_{N,p^2,p,n}(\text{Perm}_{[p]})$, the contrapositive of Lemma 2.30 gives that either $\text{size}(P)$ or $\deg(P)$ has to be at least,

$$\frac{\text{size}(\text{Perm}_{[p]})^{0.1}}{N \cdot 2^n} > \frac{\text{size}(\text{Perm}_m)^{0.1}}{N \cdot 2^n} > \frac{2^{0.1 m^\varepsilon}}{N \cdot 2^n}$$

Since $N = \binom{n+d}{n} \leq 2^{2n} \ll 2^{m^\varepsilon}$, this is at least N^{cm^ε} for some constant c, for all large enough N. Thus, if $\deg(P)$ was indeed at most N^e, then its size must be at least N^{cm^ε}. □

Concluding that VNP has no efficient equations

Theorem 1.10 (Conditional Hardness of Equations for VNP)**.** *Let $\varepsilon > 0$ be a constant. Suppose that the permanent family Perm_m requires circuits of size 2^{m^ε}.*

Then, VP has VNP-succinct hitting sets. Therefore, there are no VP-natural proofs for VNP.

Proof. We will show that for $d(n) = n$, there is no $D(N) = \text{poly}(N)$ for which there are VP_D-natural proofs for the class VNP_d. So suppose that $\{P_N\}$ is a family of equations for VNP_d, that has degree $\text{poly}(N)$. This means that for *all large enough* n, and $N = \binom{n+d}{n}$, the polynomial P_N vanishes on the coefficient vectors of all polynomials in $\mathsf{VNP}_d(n)$.

However, Theorem 5.8 shows that for m large enough, if there is a constant $\varepsilon > 0$ for which we have that $\text{size}(\text{Perm}_m) \geq 2^{m^\varepsilon}$, then for $n = m^{\varepsilon/4}$ and any $d \leq n$, the coefficient vectors of polynomials in $\mathsf{VNP}_d(n)$ form a hitting set for all N-variate, degree-$\text{poly}(N)$ polynomials that are computable by circuits of size $\text{poly}(N)$. Now suppose the Permanent family is 2^{m^ε}-hard for a constant $\varepsilon > 0$, which means that Perm_m is 2^{m^ε}-hard for *infinitely many* $m \in \mathbb{N}$. Then using Theorem 5.8, we can conclude that for any family $\{P_N\} \in \mathsf{VP}$, we must have for *infinitely many* n that $P_N(\overline{\text{coeff}}(f_n)) \neq 0$ for some $f_n \in \mathsf{VNP}_d(n)$. Since the choice of $\{P_N\} \in \mathsf{VP}$ was arbitrary, this means that there are VNP_d-succinct hitting sets for VP, for $d = n$. □

Remark 5.9. Some recent results on algebraic circuit lower bounds, starting with [KSS14], involves studying families of polynomials whose monomials come from a combinatorial design. A natural question is whether the membership of such polynomial families in VNP (often shown via Valiant's criterion, which in turn relies on the explicitness of the underlying designs) somehow implies the VNP-succinctness of the KI generator in a blackbox manner. We do not know if such a blackbox transformation exists. Nevertheless, our proof of VNP-succinctness of the KI generator proceeds along similar lines but crucially relies on the fact that the underlying combinatorial designs were constructed via Reed-Solomon codes. In contrast to this, the VNP membership of polynomial families based on combinatorial designs via Valiant's criterion, as in [KSS14], only seems to rely on the *explicitness* of the designs, and so, at least on the surface, appears to be less dependent on the precise construction of the underlying combinatorial designs. ◊

6 Open questions

Some key directions that are open for further study can be categorized as follows.

Disproving the existence of natural proofs for VP. This would be equivalent to proving the existence of VP-succinct hitting sets for VP, analogous to Theorem 1.10. The key challenge here is that we do not know any constructions of hitting sets for circuits that follow from polynomial hardness.

This is necessary because coefficient-vectors of VP forming a hitting set for VP is *equivalent* to the "evaluation vectors" of VP forming a hitting set for VP. The recent work of Andrews and Forbes [AF22] talks about some of the challenges in constructing hitting sets with parameters similar to this. The question of "algebraic cryptography" (see e.g. [AD08]) alluded to before, is also along the same lines.

Proving the existence of natural proofs. This would be an interesting development for any circuit class for which strong lower bounds are not known. Of course, such a result — unless it proves a new lower bound — would have to rely on some believable "easiness assumption".

A specific question could be to show that constant-free circuits (or formulas, ABPs) have efficient equations without any restrictions on coefficients.

In this context, it can be shown (as a consequence of Lemma 2.27) that a version of Theorem 1.6 which works for integer coefficients with large magnitudes, say $\exp((\log N)^{\log * N})$, will imply VP-natural proofs for all of VP. The proof strategy for Theorem 1.6 gives equations with degree that is at least linear in the magnitude of the coefficients, and is therefore unlikely to be useful for this purpose. It would therefore be interesting to know if there are constructions that achieve a better dependence between degree and the magnitude of the coefficients.

Designing non-natural lower bound strategies. This is a slightly vague question, in that almost any concrete and general strategy for proving lower bounds that circumvents a possible natural-proofs-barrier would be interesting.

In some sense, the recent breakthrough of Limaye, Srinivasan and Tavenas [LST21] provides one such approach: reduce the lower bound question for $\mathcal{C}$ to that for some $\mathcal{C}'$, such that $\mathcal{C}'$ admits natural proofs. However, it is unclear if this can be pursued as a general strategy, because this additionally requires non-trivial upper bounds (from $\mathcal{C}$ to $\mathcal{C}'$).

Acknowledgements

The authors from TIFR acknowledge support of the Department of Atomic Energy, Government of India, under project no. 12-R&D-TFR-5.01-0500.

Mrinal thanks Rahul Santhanam and Ben Lee Volk for many insightful conversations about algebraic natural proofs and succinct hitting sets.

Anamay thanks Robert Andrews, Amir Shpilka and Ben Lee Volk, and the attendees of the workshop on *Proof Complexity and Meta-mathematics* at the Simons Institute for discussing their insights on algebraic natural proofs.

All the authors are grateful to Ben Lee Volk, Amir Shpilka, and the anonymous reviewers of FOCS 2020, STOC 2021, and STACS 2022, for their valuable suggestions on our preliminary works ([CKR+20, KRST22]) on which this paper rests.

References

[AD08] Scott Aaranson and Andrew Drucker. Arithmetic natural proofs theory is sought. Shtetl Optimized: Scott Aaranson's Blog, 2008.

[AF22] Robert Andrews and Michael A. Forbes. Ideals, determinants, and straightening: proving and using lower bounds for polynomial ideals. In *STOC '22: 54th Annual ACM SIGACT Symposium on Theory of Computing, 2022*, pages 389–402. ACM, 2022.

[AW09] Scott Aaronson and Avi Wigderson. Algebrization: A New Barrier in Complexity Theory. *ACM Trans. Comput. Theory*, 1(1), February 2009.

[BGS75] Theodore Baker, John Gill, and Robert Solovay. Relativizations of the $\mathscr{P} =? \mathscr{NP}$ Question. *SIAM Journal on Computing*, 4(4):431–442, 1975.

[BIJL18] Markus Bläser, Christian Ikenmeyer, Gorav Jindal, and Vladimir Lysikov. Generalized matrix completion and algebraic natural proofs. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2018, Los Angeles, CA, USA, June 25-29, 2018*, pages 1193–1206. ACM, 2018.

[BIL+19] Markus Bläser, Christian Ikenmeyer, Vladimir Lysikov, Anurag Pandey, and Frank-Olaf Schreyer. Variety Membership Testing, Algebraic Natural Proofs, and Geometric Complexity Theory. *CoRR*, abs/1911.02534, 2019.

[BS83] Walter Baur and Volker Strassen. The Complexity of Partial Derivatives. *Theoretical Computer Science*, 22:317–330, 1983.

[Cho11] Timothy Y. Chow. Almost-natural proofs. *J. Comput. Syst. Sci.*, 77(4):728–737, 2011.

[CKR+20] Prerona Chatterjee, Mrinal Kumar, C. Ramya, Ramprasad Saptharishi, and Anamay Tengse. On the Existence of Algebraically Natural Proofs. In *Proceedings of the 61st Annual IEEE Symposium on Foundations of Computer Science (FOCS 2020)*, 2020. eccc:TR20-063.

[CLO07] David A. Cox, John B. Little, and Donal O'Shea. *Ideals, Varieties and Algorithms*. Undergraduate texts in mathematics. Springer, 2007.

[DL78] Richard A. DeMillo and Richard J. Lipton. A Probabilistic Remark on Algebraic Program Testing. *Information Processing Letters*, 7(4):193–195, 1978.

[EGOW18] Klim Efremenko, Ankit Garg, Rafael Oliveira, and Avi Wigderson. Barriers for Rank Methods in Arithmetic Complexity. In *9th Innovations in Theoretical Computer Science Conference, ITCS 2018, January 11-14, 2018, Cambridge, MA, USA*, volume 94 of *LIPIcs*, pages 1:1–1:19. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2018.

[For14] Michael Forbes. *Polynomial Identity Testing of Read-Once Oblivious Algebraic Branching Programs*. PhD thesis, Massachusetts Institute of Technology, 2014.

[FSS84] Merrick Furst, James B. Saxe, and Michael Sipser. Parity, circuits, and the polynomial-time hierarchy. *Mathematical systems theory*, 17(1):13–27, 1984.

[FSV18] Michael A. Forbes, Amir Shpilka, and Ben Lee Volk. Succinct Hitting Sets and Barriers to Proving Lower Bounds for Algebraic Circuits. *Theory of Computing*, 14(1):1–45, 2018.

[GKSS17] Joshua A. Grochow, Mrinal Kumar, Michael E. Saks, and Shubhangi Saraf. Towards an algebraic natural proofs barrier via polynomial identity testing. *CoRR*, abs/1701.01717, 2017. Pre-print available at arXiv:1701.01717.

[GMOW19] Ankit Garg, Visu Makam, Rafael Oliveira, and Avi Wigderson. More Barriers for Rank Methods, via a "numeric to Symbolic" Transfer. In *60th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2019, Baltimore, Maryland, USA, November 9-12, 2019*, pages 824–844. IEEE Computer Society, 2019.

[Gro15] Joshua A. Grochow. Unifying Known Lower Bounds via Geometric Complexity Theory. *Comput. Complex.*, 24(2):393–475, 2015.

[Hås86] Johan Håstad. Almost Optimal Lower Bounds for Small Depth Circuits. In *Proceedings of the Eighteenth Annual ACM Symposium on Theory of Computing*, STOC '86, page 6–20, New York, NY, USA, 1986. Association for Computing Machinery.

[HS80a] Joos Heintz and Claus-Peter Schnorr. Testing Polynomials which Are Easy to Compute (Extended Abstract). In *Proceedings of the 12th Annual ACM Symposium on Theory of Computing (STOC 1980)*, pages 262–272. ACM, 1980.

[HS80b] Joos Heintz and Malte Sieveking. Lower Bounds for Polynomials with Algebraic Coefficients. *Theoretical Computer Science*, 3:321–330, 1980.

[KI04] Valentine Kabanets and Russell Impagliazzo. Derandomizing Polynomial Identity Tests Means Proving Circuit Lower Bounds. *Computational Complexity*, 13(1-2):1–46, 2004. Preliminary version in the *35th Annual ACM Symposium on Theory of Computing (STOC 2003)*.

[KRST22] Mrinal Kumar, C. Ramya, Ramprasad Saptharishi, and Anamay Tengse. If VNP Is Hard, Then so Are Equations for It. In *39th International Symposium on Theoretical Aspects of Computer Science, STACS 2022, March 15-18, 2022, Marseille, France (Virtual Conference)*, volume 219 of *LIPIcs*, pages 44:1–44:13. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2022. arXiv:2012.07056.

[KSS14] Neeraj Kayal, Chandan Saha, and Ramprasad Saptharishi. A super-polynomial lower bound for regular arithmetic formulas. In *Proceedings of the 46th Annual ACM Symposium on Theory of Computing (STOC 2014)*, pages 146–153, 2014. Pre-print available at eccc:TR13-091.

[KV20] Mrinal Kumar and Ben Lee Volk. A Polynomial Degree Bound on Defining Equations of Non-rigid Matrices and Small Linear Circuits. *CoRR*, abs/2003.12938, 2020. Pre-print available at arXiv:2003.12938.

[LST21] Nutan Limaye, Srikanth Srinivasan, and Sébastien Tavenas. Superpolynomial Lower Bounds Against Low-Depth Algebraic Circuits. In *62nd IEEE Annual Symposium on Foundations of Computer Science, FOCS 2021, Denver, CO, USA, February 7-10, 2022*, pages 804–814. IEEE, 2021. eccc:TR21-081.

[NW94] Noam Nisan and Avi Wigderson. Hardness vs Randomness. *Journal of Computer and System Sciences*, 49(2):149–167, 1994. Available on citeseer:10.1.1.83.8416.

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

[Raz87] Alexander A. Razborov. Lower bounds on the size of bounded depth circuits over a complete basis with logical addition. *Mathematical notes of the Academy of Sciences of the USSR*, 41(4):333–338, 1987.

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

[RR97] Alexander A. Razborov and Steven Rudich. Natural Proofs. *J. Comput. Syst. Sci.*, 55(1):24–35, 1997.

[Sap15] Ramprasad Saptharishi. A survey of lower bounds in arithmetic circuit complexity. Github survey, 2015.

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

[Sie14] Carl L Siegel. Über einige anwendungen diophantischer approximationen. In *On Some Applications of Diophantine Approximations*, pages 81–138. Springer, 2014.

[Smo87] Roman Smolensky. Algebraic Methods in the Theory of Lower Bounds for Boolean Circuit Complexity. In *Proceedings of the 19th Annual ACM Symposium on Theory of Computing (STOC 1987)*, pages 77–82, 1987.

[Str73] Volker Strassen. Die Berechnungskomplexität Von Elementarsymmetrischen Funktionen Und Von Interpolationskoeffizienten. *Numerische Mathematik*, 20(3):238–251, 1973.

[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. Preliminary version in the *6th Internationl Symposium on the Mathematical Foundations of Computer Science (MFCS 1981)*.

[Zip79] Richard Zippel. Probabilistic algorithms for sparse polynomials. In *Symbolic and Algebraic Computation, EUROSAM '79, An International Symposiumon Symbolic and Algebraic Computation*, volume 72 of *Lecture Notes in Computer Science*, pages 216–226. Springer, 1979.

ECCC ISSN 1433-8092
https://eccc.weizmann.ac.il