2016 IEEE 57th Annual Symposium on Foundations of Computer Science

The Hilbert Function, Algebraic Extractors, and Recursive Fourier Sampling

Zachary Remscrim
MIT
Email:remscrim@mit.edu

***Abstract*—In this paper, we apply tools from algebraic geometry to prove new results concerning extractors for algebraic sets, the recursive Fourier sampling problem, and VC dimension. We present a new construction of an extractor which works for algebraic sets defined by polynomials over GF(2) of substantially higher degree than the current state-of-the-art construction. We also exactly determine the GF(2)-polynomial degree of the recursive Fourier sampling problem and use this to provide new partial results towards a circuit lower bound for this problem. Finally, we answer a question concerning VC dimension, interpolation degree and the Hilbert function.**

***Keywords*-extractors; polynomial degree; oracle separations**

I. Introduction

A. Extractors

For a finite domain Ω and a collection of distributions $\mathcal{C}$ over Ω, we say that a function $E : \Omega \to \{0,1\}^m$ is an *extractor* (sometimes called a *deterministic extractor*) for $\mathcal{C}$ if, for every random variable X distributed according to any distribution in $\mathcal{C}$, $E(X)$ is close to the uniform distribution. We call each distribution $C \in \mathcal{C}$ a *source*. Of course, in order to have any hope of the collection of distributions $\mathcal{C}$ having an extractor, some sort of condition must be satisfied by the sources. While it is trivial to exhibit simple conditions on $\mathcal{C}$ such that a random function will, with high probability, be an extractor, the problem becomes far more interesting when one requires an *explicit* construction of E (that is to say, a construction realizable by some deterministic polynomial time Turing machine). The natural problem is then to exhibit particular distributions $\mathcal{C}$ for which there exist explicit constructions of extractors.

Numerous versions of this problem have been considered. In this paper, we consider the case, originally introduced in [12], where each source is the uniform distribution over the set of common zeros of a collection of polynomials defined over some field. Such a set is called an *algebraic set* and such a source is called an *algebraic source*. Algebraic sources are a natural generalization of *affine sources* (see, for instance [16] and [8]) and *bit-fixing sources* (see, for instance, [17] and [22]) and build naturally on the earlier work of *efficiently samplable sources* (see, for instance, [33], [21], and [11]).

To be precise, for a finite field $\mathbb{F}$, and a positive integer d, we consider algebraic sets $V \subseteq \mathbb{F}^n$ where V is the set of common zeros of a collection of polynomials $f_1, \ldots, f_t \in \mathbb{F}[x_1, \ldots, x_n]$ such that $\deg(f_i) \leq d$. We say that V has *density* ρ if $|V| \geq \rho|\mathbb{F}^n|$. We say that a function $f : \mathbb{F}^n \to \mathbb{F}$ is an *extractor for algebraic sets* defined by polynomials of degree at most d and density ρ if f is close to uniform on every such algebraic set. A closely related weaker notion is that of a *disperser for algebraic sets*, where we say that a function $f : \mathbb{F}^n \to \mathbb{F}$ is a *disperser for algebraic sets* defined by polynomials of degree at most d and density ρ if, for every such algebraic set V, the image of $f : V \to \mathbb{F}$ (the restriction of f to V) is $\mathbb{F}$. Some authors call this a *zero-error* disperser for algebraic sets, to emphasize that all elements of $\mathbb{F}$ lie in the image of f. Clearly any extractor is also a disperser.

As shown in [12], there exist explicit extractors for polynomials of degree d defined over moderately sized fields, where $|\mathbb{F}| = poly(d)$, and density $\rho = 2^{-\frac{n}{2}}$ as well as over large fields, where $|\mathbb{F}| = d^{\Omega(n^2)}$ and very small density. However, very little is known about the extreme case in which $\mathbb{F} = \mathbb{F}_2$, the two element finite field. To the best of our knowledge, the current state of the art construction for extractors and dispersers is that of [10], in which an explicit construction was exhibited for an extractor for algebraic sets defined by at most $(\log\log n)^{\frac{1}{2e}}$ polynomials each of degree at most 2, as well as for a disperser for algebraic sets defined by at most t polynomials each of degree at most $d = (1 - o(1))\frac{\log(\frac{n}{t})}{\log^{0.9} n}$ (in particular, when $t \leq n^\alpha$ for some $\alpha < 1$, then the requirement on degree is $d < (1 - \alpha - o(1))\log^{0.1} n$).

In this paper, we focus on the case in which $\mathbb{F} = \mathbb{F}_2$, and exhibit explicit extractors (and hence explicit dispersers) for algebraic sets defined by polynomials of substantially higher degree than any previous construction. We now formally state our results. For any set $V \subseteq \mathbb{F}_2^n$, we say that a function $f : \mathbb{F}_2^n \to \mathbb{F}_2$ has *bias* ϵ on V if $\text{bias}(f|_V) := |\mathbb{E}_{x \sim V}[(-1)^{f(x)}]| \leq \epsilon$. A function $f : \mathbb{F}_2^n \to \mathbb{F}_2$ is called an *extractor for algebraic sets* defined by polynomials of degree at most d of density ρ with bias ϵ if $\text{bias}(f|_V) \leq \epsilon$ for every such algebraic set V. We show that any δ-versatile function (this will be defined precisely in §3) is an extractor for algebraic sets.

Theorem 1. *Let $f : \mathbb{F}_2^n \to \mathbb{F}_2$ be δ-versatile (on $\mathbb{F}_2^n$), where $\delta \geq \frac{n}{2} - n^\gamma$ for some $0 \leq \gamma < \frac{1}{2}$. Then, there is a constant $c > 0$ such that, for any constants α, β such that $0 < \alpha, \beta < \frac{1}{2}$, and for any $d \leq n^\alpha$ and $\rho \geq 2^{-n^\beta}$, f is an extractor with*

0272-5428/16 $31.00 © 2016 IEEE
DOI 10.1109/FOCS.2016.29

bias $\frac{c\left(n^{\gamma}+d\log(\frac{\sqrt{n}}{\rho})\right)}{\sqrt{n}}$ *for algebraic sets of density at least* ρ *that are the common zeros of a collection of polynomials each of degree at most* d*.*

As will be shown in §3, the Majority function (the function MAJ : $\mathbb{F}_2^n \rightarrow \mathbb{F}_2$ where MAJ$(x) = 1$ when $wt(x) \geq \frac{n}{2}$ and MAJ$(x) = 0$ when $wt(x) < \frac{n}{2}$, where $wt(x)$ denotes the number of 1s in x) is $\frac{n}{2}$-versatile, and so the following corollary is immediate.

Corollary 1. *There is a constant* $c > 0$ *such that, for any constants* α, β *such that* $0 < \alpha, \beta < \frac{1}{2}$*, and for any* $d \leq n^{\alpha}$ *and* $\rho \geq 2^{-n^{\beta}}$*, MAJ* : $\mathbb{F}_2^n \rightarrow \mathbb{F}_2$ *is an extractor with bias* $\frac{cd\log(\frac{\sqrt{n}}{\rho})}{\sqrt{n}}$ *for algebraic sets of density at least* ρ *that are the common zeros of a collection of polynomials each of degree at most* d*.*

Much as was the case in [12] and [10], our construction relies on statements involving the set of zeros of a single low degree polynomial defined over $\mathbb{F}$. The key distinction between our construction and earlier constructions, which allows our construction to work even for rather high degree polynomials over $\mathbb{F}_2$, is that our construction exploits the structure of this set of zeros, rather than simply bounds on the size of the set of zeros that follow directly from the degree of the polynomial (that is to say, bounds that follow directly from the fundamental theorem of algebra, or, in other words, Schwartz-Zippel type bounds).

B. Recursive Fourier Sampling

The recursive Fourier sampling problem is one of the most well studied problems in quantum complexity theory. This problem was first defined, along with the complexity class BQP (Bounded-Error Quantum Polynomial Time), in [6], a foundational work of quantum complexity theory. In that paper, this problem, whose formal definition we delay for now, was used to exhibit an oracle A relative to which BQP is not contained in NP or even MA, that is to say an A such that BQP$^A \not\subset$ NP^A and BQP$^A \not\subset$ MA^A. Such oracle separations are interesting both because they are, perhaps, suggestive of a unrelativized separation, as well as because they concretely exhibit a measure of complexity in which quantum computers provably outperform classical computers: query complexity, where the resource of interest is the number of queries to the (very long) input string.

For this reason, it is natural to seek oracle separations between BQP and increasingly larger classical complexity classes. However, very little progress in this direction has been made. While some results are known, such as the fact, proven in [2], that there is an oracle A such that BQP$^A \not\subset$ BPP$^A_{\text{path}}$ and BQP$^A \not\subset$ SZK^A, even the question of whether or not there exists an oracle A such that BQP$^A \not\subset$ AM^A remains open, as does, of course, the substantially stronger question of whether or not there exists an oracle A such that BQP$^A \not\subset$ PH^A.

It is this potential oracle separation between BQP and the polynomial hierarchy that we now focus on. The natural approach to this problem, which has been used successfully to show many other similar oracle separations between certain complexity classes and the polynomial hierarchy, is to exploit the connection between relativized separations from the polynomial hierarchy and lower bounds against constant depth circuits [15],[34]. Here, the key idea is to reinterpret the $\exists$ and $\forall$ quantifiers of a PH machine as *OR* and *AND* gates, respectively, to convert a PH machine solving some oracle problem on a 2^n bit long oracle string, into a constant depth, $2^{poly(n)}$ sized circuit, consisting of AND, OR, and NOT gates that solves the same problem. Using this idea, and a $2^{\omega(poly(n))}$ lower bound on the size of a constant depth circuit computing the PARITY function (on an input of size 2^n), one concludes that there is an oracle A relative to which $\oplus$P$^A \not\subset$ PH^A. The same idea can, and has, been used to show other such relativized separations.

Therefore, given this connection between relativized separations from the polynomial hierarchy and lower bounds against constant depth circuits, and the powerful techniques that exist to show lower bounds against constant depth circuits, [15],[3],[18],[29],[31], one might very naturally ask why the question of whether or not there exists an A such that BQP$^A \not\subset$ PH^A remains open. Most fundamentally, the problem is that, in order to show that a particular function f cannot be computed by a small circuit, all of these circuit lower bound techniques either explicitly (in the case of [29] or [31]) or implicitly (in the case of [15],[3],[18] as shown by [24]) argue that f cannot be well approximated by a low-degree polynomial. This is a problem because, as shown in [5], any function that can be computed by an efficient quantum algorithm is well approximated by a low degree polynomial.

More precisely, however, [5] only guarantees the existence of a low-degree polynomial over $\mathbb{R}$, whereas the non-existence of a low-degree polynomial over any field $\mathbb{F}$ would suffice (via the Razborov-Smolensky method) to prove a circuit lower bound, and so this certainly does not completely doom the application of traditional circuit lower bound techniques. Nevertheless, the result of [5] does suggest that a deeper understanding of approximation by low-degree polynomials may be necessary to resolve the question of whether or not there exists an oracle A such that BQP$^A \not\subset$ PH^A. It is this issue that we focus on within this paper.

As has been observed by many authors (for instance [6],[7],[1],[19],[2]) the recursive Fourier sampling problem (or a slight variant) is a prime candidate for exhibiting an oracle A such that BQP$^A \not\subset$ PH^A, as this problem seems to perfectly exploit the advantages of a quantum computer at the expense of a classical one.

The recursive Fourier sampling problem will be formally defined in §5. For the moment, we will simply state that it is

a promise problem (that is to say, a partial Boolean function whose value is only defined on a portion of the input space, called the promise) which is known to have an efficient quantum algorithm. By the result of [5], this immediately implies that there is a low degree real polynomial that well approximates the recursive Fourier sampling problem on the promise. In fact, from the standpoint of proving a circuit lower bound, the situation is even "worse" than this, due to the result of [20], which shows that there is an even lower degree real polynomial than the one guaranteed by [5] which exactly represents the recursive Fourier sampling problem on its promise. Moreover, [20] proves exactly matching upper and lower bounds on any real polynomial that represents the recursive Fourier sampling problem on its promise, thereby completely resolving the question of the polynomial degree of the recursive Fourier sampling problem, with respect to polynomials over $\mathbb{R}$.

In this paper, we consider the question of the polynomial degree of the recursive Fourier sampling problem for polynomials defined over $\mathbb{F}_2$. That is to say, we consider the question of what is the lowest degree polynomial defined over $\mathbb{F}_2$ that represents the recursive Fourier sampling problem on its promise. Before proceeding further, we briefly note that this question is only non-trivial because the recursive Fourier sampling problem is a promise problem. For any total function $g : \mathbb{F}_2^n \to \mathbb{F}_2$, there is a unique multilinear polynomial $f \in \mathbb{F}_2[x_1, \ldots, x_n]$ that agrees everywhere with g; the degree of f is, of course, the minimal degree of any polynomial in $\mathbb{F}[x_1, \ldots, x_n]$ that agrees everywhere with g. For a promise problem, however, there can be many multilinear polynomials, of varying degrees, that all agree on the promise.

Over $\mathbb{F}_2$, there is a simple, though relatively high degree, polynomial that exactly computes the recursive Fourier sampling problem. Our key result, stated in the following theorems, is that, for a certain appropriate settings of the parameters, this simple polynomial is, in fact, the lowest degree polynomial that agrees with recursive Fourier sampling everywhere on its promise. In fact, we show something even stronger: no polynomial of lower degree can even non-trivially one-sided agree with the recursive Fourier sampling problem (that is to say, if a polynomial is zero everywhere (on the promise) that the recursive Fourier sampling problem is zero, then that polynomial must be zero on the entire promise). We then use these results to prove new statements about the ability of constant depth circuits to compute the recursive Fourier sampling problem.

Theorem 2. *For any positive integers k, h, Let $n = 2^k - 1$ and let $RFS_{n,h}^{MAJ}$ denote the recursive Fourier sampling function with majority. Then $\nexists g \in \mathbb{F}_2[x_1, \ldots, x_m]$ such that deg$(g) < \left(\frac{n+1}{2}\right)^h$ and $g(x) = RFS_{n,h}^{MAJ}(x)\ \forall x \in U_{p,h}^{MAJ}$. Moreover, if any $g \in \mathbb{F}_2[x_1, \ldots, x_m]$ such that deg$(g) < \left(\frac{n+1}{2}\right)^h$ vanishes everywhere on $U_{0,h}^{MAJ}$, it vanishes everywhere on $U_{1,h}^{MAJ}$.*

Theorem 3. *For any positive integers d, n, h such that $d|n$, and $n \geq d(2^{d^2} + d - 1)$, Let $RFS_{n,h}^{GIP_{n,d}}$ denote the recursive Fourier sampling function with generalized inner product. Then $\nexists g \in \mathbb{F}_2[x_1, \ldots, x_m]$ such that deg$(g) < d^h$ and $g(x) = RFS_{n,h}^{GIP_{n,d}}(x)\ \forall x \in U_{p,h}^{GIP_{n,d}}$. Moreover, if any $g \in \mathbb{F}_2[x_1, \ldots, x_m]$ such that deg$(g) < d^h$ vanishes everywhere on $U_{0,h}^{GIP_{n,d}}$, it vanishes everywhere on $U_{1,h}^{GIP_{n,d}}$.*

C. VC Dimension

We say that a subset $J \subseteq [n]$ is *shattered* by a family of vectors $C \subseteq \{0,1\}^n$ if, $\forall s : J \to \{0,1\},\ \exists c \in C$ such that $c_j = s(j)\ \forall j \in J$ (in other words, if one considers the set of all substrings of elements of C comprised of the positions indexed by J, this collection of substrings is precisely $\{0,1\}^{|J|}$). We then write str$(C) = \{J \subseteq [n] : J$ is shattered by$C\}$ to denote the sets that are shattered with respect to C. We then define the *VC dimension* of C as VC$(C) = \max\{|J| : J \in \text{str}(C)\}$. For a field $\mathbb{F}$ and a set $C \subseteq \{0,1\}^n$, the interpolation degree of C, denoted by reg(C) is the minimum d such that every function $f : C \to \mathbb{F}$ can be expressed as a multilinear polynomial in $\mathbb{F}[x_1, \ldots, x_n]$ of degree at most d.

Recently, in [26], a very interesting connection between VC dimension and interpolation degree was demonstrated. A simple characterization of sets with interpolation degree 1 was provided. This naturally raised the question of whether a similar characterization exists for sets with interpolation degree r, for arbitrary r. In this paper, we provide such a characterization, in terms of the rank of a certain inclusion matrix, which will be defined precisely in §2. As noted in [26], the primary motivation for such a characterization is to better understand the structure of sets with a given VC dimension (see, for instance, [4] and [27]).

Theorem 4. *A set $C \subseteq \{0,1\}^n$ has reg$(C) = r$ if and only if r is the smallest positive integer such that rank$_{\mathbb{F}_2}\mathcal{M}(C, \binom{[n]}{\leq r}) = |C|$.*

D. Organization of this Paper

We begin, in §2, by reviewing several key definitions and results from algebraic geometry that will be used throughout this paper. In §3, we develop the concept of δ-versatile functions, a natural generalization of the concept of versatile functions defined in [23]. In §4, we exhibit a family of extractors for algebraic sets. In §5, we consider the recursive Fourier sampling problem and present new results concerning its polynomial degree and new partial results towards a circuit lower bound. In §6, we use standard results from algebraic geometry to provide a simple answer to a question raised in [26] concerning interpolation degree and VC-dimension. Due to space limitations, several proofs have been omitted in this extended abstract; see the full paper [30].

II. PRELIMINARIES

We begin by recalling several standard definitions from algebraic geometry. Let $\mathbb{F}$ denote a (not necessarily algebraically closed) field and $\mathbb{F}[x_1, \ldots, x_n]$ denote the ring of polynomials in n indeterminates. An *algebraic set* in $\mathbb{F}^n$ is the set of common zeros of a collection of polynomials in $\mathbb{F}[x_1, \ldots, x_n]$. More precisely, given a set of polynomials $f_1, \ldots, f_k \in \mathbb{F}[x_1, \ldots, x_n]$, we denote their set of common zeros by $V(f_1, \ldots, f_k) = \{(x_1, \ldots, x_n) \in \mathbb{F}^n : f_i(x_1, \ldots, x_n) = 0 \ \forall i\}$.

Rather than working with an arbitrary set of polynomials, it will often be convenient to consider an algebraically nicer object: an ideal. For I an ideal in $\mathbb{F}[x_1, \ldots, x_n]$, let $V(I)$ denote the common zero set of all polynomials in I, that is to say $V(I) = \{(x_1, \ldots, x_n) \in \mathbb{F}^n : f(x) = 0 \ \forall f \in I\}$. Given a set of polynomials $f_1, \ldots, f_k \in \mathbb{F}[x_1, \ldots, x_n]$, let $\langle f_1, \ldots, f_k \rangle$ denote the ideal which they generate in $\mathbb{F}[x_1, \ldots, x_n]$. Clearly, $V(\langle f_1, \ldots, f_k \rangle) = V(f_1, \ldots, f_k)$. For an algebraic set V, let its *vanishing ideal* $I(V)$ be the ideal of $\mathbb{F}[x_1, \ldots, x_n]$ consisting of all polynomials which vanish on V and let $R(V) = \mathbb{F}[x_1, \ldots, x_n]/I(V)$ denote its *coordinate ring*.

For a polynomial $f \in \mathbb{F}[x_1, \ldots, x_n]$, let $\deg(f)$ denote its total degree. Let $\mathbb{F}[x_1, \ldots, x_n]_{\leq d}$ denote the vector space of polynomials over $\mathbb{F}$ with degree at most d. For an ideal I, let $I_{\leq d} = I \cap \mathbb{F}[x_1, \ldots, x_n]_{\leq d}$ denote the subspace consisting of all polynomials in I of degree at most d. For an algebraic set V, with vanishing ideal $I = I(V)$ and coordinate ring $R = R(V)$, let $R_{\leq d} = \mathbb{F}[x_1, \ldots, x_n]_{\leq d}/I_{\leq d}$. The *affine Hilbert function* $h^a(R, d)$ of R is then given by $h^a(R, d) = \dim_{\mathbb{F}}(R_{\leq d})$. By slight abuse of notation, we will use the term *affine Hilbert function of an algebraic set* V, which we will denote $h^a(V, d)$, to simply be the affine Hilbert function of the coordinate ring $R(V)$.

Throughout this paper, we consider only zero-dimensional algebraic sets V (that is to say, V is finite). For such a V, we define its *regularity* $\mathrm{reg}(V)$ to be the minimal value of d such that $h^a(V, d) = |V|$. Equivalently, $\mathrm{reg}(V)$ is the minimal value of d such that every function $V \to \mathbb{F}$ can be realized as a polynomial of degree at most d. This quantity is frequently referred to as *interpolation degree*. In the case of zero-dimensional algebraic sets, this quantity is equivalent to the Castelnuovo-Mumford regularity of $R(V)$ (see, for instance [13] Thm.4.1).

For $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$, we define x^α to be the monomial $x_1^{\alpha_1} \cdots x_n^{\alpha_n} \in \mathbb{F}[x_1, \ldots, x_n]$. For any $J \subseteq [n]$ we define the (multilinear) monomial x_J by $x_J = \prod_{j \in J} x_j$. A *degree compatible term order* $<$ is a total order on the monomials x^α which respects multiplication ($x^\alpha < x^\beta \Rightarrow x^\alpha x^\gamma < x^\beta x^\gamma \ \forall x^\alpha, x^\beta, x^\gamma \in \mathbb{F}[x_1, \ldots, x_n]$) and is degree compatible ($\deg(x^\alpha) < \deg(x^\beta) \Rightarrow x^\alpha < x^\beta \ \forall x^\alpha, x^\beta \in \mathbb{F}[x_1, \ldots, x_n]$). For a degree compatible term order $<$, and polynomial $f \in \mathbb{F}[x_1, \ldots, x_n]$, we define its *leading monomial* $\mathrm{lm}(f)$ to be the largest monomial in f with respect to $<$. Similarly, for an ideal I in $\mathbb{F}[x_1, \ldots, x_n]$, we define its *leading monomials* to be $\mathrm{LM}(I) = \{\mathrm{lm}(f) : f \in I\}$ and its *standard monomials* to be $\mathrm{SM}(I) = \{x^\alpha : \alpha \in \mathbb{N}^n\} \backslash \mathrm{LM}(I)$.

For an algebraic set V, we define $\mathrm{LM}(V) = \mathrm{LM}(I(V))$ and $\mathrm{SM}(V) = \mathrm{SM}(I(V))$. We also define $\mathrm{SM}(V, d) = \{x^\alpha \in \mathrm{SM}(V) : \deg(x^\alpha) = d\}$ and $\mathrm{LM}(V, d) = \{x^\alpha \in \mathrm{LM}(V) : \deg(x^\alpha) = d\}$.

Standard monomials provide an extremely convenient tool for computing both the Hilbert function of an algebraic set and its regularity, as illustrated in the following lemma (these are well known facts in algebraic geometry; see, for instance [14]).

Lemma 1. (a) $h^a(V, d) = \sum_{i=0}^{d} |SM(V, i)|$
(b) $reg(V) = \max_{x^\alpha \in SM(V)} deg(x^\alpha)$
(c) $|SM(V)| = |V|$
(d) $V_1 \subseteq V_2 \Rightarrow SM(V_1) \subseteq SM(V_2)$
(e) $V_1 \subseteq V_2 \Rightarrow LM(V_1) \supseteq LM(V_2)$

Let M_n denote the semigroup of all monomials in n indeterminates. That is to say, as a set $M_n = \{x^\alpha : \alpha \in \mathbb{N}^n\}$ with multiplication between monomials defined in the usual way. An ideal U of M_n is simply an upwardly closed subset of M_n ($x^\alpha \in U \Rightarrow x^\alpha x^\beta \in U \ \forall \alpha, \beta$). For an algebraic set $V \subseteq \mathbb{F}^n$, $\mathrm{LM}(V)$ is an ideal of M_n. Similarly, $\mathrm{SM}(V)$ is a dual ideal. In other words, if $x^\alpha \in \mathrm{LM}(V)$, then $x^\alpha x^\beta \in \mathrm{LM}(V)$ and if $x^\alpha \in \mathrm{SM}(V)$ then $x^\beta \in \mathrm{SM}(V)$ for any divisor x^β of x^α.

For I an ideal of $\mathbb{F}[x_1, \ldots, x_n]$, let $a(I)$ denote the minimal degree of any $g \in I$ such that g consists of only monomials from $\mathrm{SM}(\mathbb{F}^n)$. For an algebraic set $V = V(I)$, let $a(V) = a(I)$. The following lemma, proven independently in [14] and [28], provides an extremely useful relationship between $\mathrm{reg}(V)$ and $a(\overline{V})$, where $\overline{V}$ denotes the complement of V.

Lemma 2. *[14], [28]*
If $V \subseteq \mathbb{F}^n$ is a nonempty zero-dimensional algebraic set, then $a(\overline{V}) + reg(V) = n$.

Lastly, we consider another useful tool for computing the Hilbert function: inclusion matrices. Let $\mathbb{F}_2$ denote the finite field of two elements. Let $2^{[n]}$ denote the collection of all subsets of $[n] = \{1, \ldots, n\}$, and let $\mathcal{F}, \mathcal{G} \subseteq 2^{[n]}$ denote two families of subsets. The *inclusion matrix* $\mathcal{M}(\mathcal{F}, \mathcal{G})$ is a $|\mathcal{F}| \times |\mathcal{G}|$ matrix, with entries in $\mathbb{F}_2$, where for any $F \in \mathcal{F}$ and $G \in \mathcal{G}$ the (F, G) entry is 1 precisely when $G \subseteq F$. Let $\binom{[n]}{\leq k}$ denote the family of all subsets of $[n]$ of size at most k.

Given an algebraic set $V \subseteq \mathbb{F}_2^n$, we associate it with a family of subsets in the natural way: for each $x = (x_1, \ldots, x_n) \in V$ the subset $\{i : x_i = 1\}$ is included in the set family. By a slight abuse of notation, we will also denote this set family by V. The following is immediate from definitions (as a nontrivial linear combination of the

columns corresponds to a polynomial in $I(V)$ and hence a leading monomial).

Lemma 3. *For any algebraic set $V \subseteq \mathbb{F}_2^n$, we have*

$$h^a(V,d) = \textit{rank}_{\mathbb{F}_2}\mathcal{M}\left(V, \binom{[n]}{\leq d}\right).$$

Throughout this paper, our key object of interest will be the affine Hilbert function of an algebraic set. We briefly note that this is a slight departure from the typical situation in algebraic geometry in which one considers the "ordinary" Hilbert function (which is defined similarly to the affine Hilbert function, but in which one considers the space of homogeneous polynomials of a particular degree, rather than arbitrary polynomials of a particular degree) of a variety (which is an algebraic set in which the ground field $\mathbb{F}$ is algebraically closed). Much as was the case in [32], this is done in order to allow a better intuitive connection between the Hilbert function and the questions from complexity theory that we consider. However, it should be noted that it is very straightforward to convert between statements involving the affine Hilbert function of an algebraic set and the Hilbert function of a variety as, firstly, one can harmlessly extend the ground field (and, in particular, extend it to its algebraic closure), and, secondly, one can straightforwardly express the value of the affine Hilbert function at degree d as the sum of values of the Hilbert function of degree at most d. While it is true that certain basic statements that would hold over an algebraically closed ground field do not necessarily hold over arbitrary fields, these statements are either facts that we explicitly exploit in the proof (such as the number of roots a particular degree d polynomial has in a particular algebraic set) or are statements that can easily be modified to analogous statements when the ground field is a finite field (for example, Hilbert's Nullstellensatz, which establishes a bijection between varieties and radical ideals can be modified to a bijection between algebraic sets and radical ideals that contain the field polynomials).

III. Generalization of Versatile Functions

In this section, we consider a certain natural generalization of the concept of versatile functions (as defined in [23], see also [31] for the concept of U_F^n − complete elements) to promise problems. We begin with a definition.

Definition 1. *A function $f : \mathbb{F}_2^n \to \mathbb{F}_2$ is Versatile if, $\forall g : \mathbb{F}_2^n \to \mathbb{F}_2$, $\exists u, v \in \mathbb{F}_2[x_1, \ldots, x_n]$ where $deg(u), deg(v) \leq \frac{n}{2}$ and $g(x) = u(x)f(x) + v(x)\ \forall x \in \mathbb{F}_2^n$.*

Versatile functions admit a particularly simple characterization in terms of regularity (this is essentially the same notion as "degree-m independent sets" as considered in [32]), as shown in the following lemma.

Lemma 4. *For a function $f : \mathbb{F}_2^n \to \mathbb{F}_2$, let $U_0 = f^{-1}(0)$ and $U_1 = f^{-1}(1)$. Then f is versatile if and only if $reg(U_0), reg(U_1) \leq \frac{n}{2}$.*

Proof: If f is versatile, then, by definition, $\forall g : \mathbb{F}_2^n \to \mathbb{F}_2$, $\exists u, v \in \mathbb{F}_2[x_1, \ldots, x_n]$ where $\deg(u), \deg(v) \leq \frac{n}{2}$ and $g(x) = u(x)f(x) + v(x)\ \forall x \in \mathbb{F}_2^n$, and so $g(x) = v(x)\ \forall x \in U_0$ and $g(x) = u(x) + v(x)\ \forall x \in U_1$. Since $\deg(u+v) \leq \max(\deg(u), \deg(v))$, it immediately follows that $\text{reg}(U_0), \text{reg}(U_1) \leq \frac{n}{2}$.

If $\text{reg}(U_0), \text{reg}(U_1) \leq \frac{n}{2}$, then, by definition, $\forall g : \mathbb{F}_2^n \to \mathbb{F}_2$, $\exists u', v' \in \mathbb{F}_2[x_1, \ldots, x_n]$ where $\deg(u'), \deg(v') \leq \frac{n}{2}$ such that $g(x) = u'(x)\ \forall x \in U_0$ and $g(x) = v'(x)\ \forall x \in U_1$. Therefore, $g(x) = u(x)f(x) + v(x)\ \forall x \in \mathbb{F}_2^n$, where $u = u' + v'$ and $v = v'$. Since $\deg(u), \deg(v) \leq \frac{n}{2}$, f is versatile. ∎

As shown in [23], the Majority function (the function $\text{MAJ} : \mathbb{F}_2^n \to \mathbb{F}_2$ where $\text{MAJ}(x) = 1$ when $wt(x) \geq \frac{n}{2}$ and $\text{MAJ}(x) = 0$ when $wt(x) < \frac{n}{2}$, where $wt(x)$ denotes the number of 1s in x) is versatile. As a first illustration of the utility of standard monomials, we present a new short proof of this fact.

Lemma 5. *The function $MAJ : \mathbb{F}_2^n \to \mathbb{F}_2$ is versatile.*

Proof: Let $U_0 = \{x \in \mathbb{F}_2^n : \text{MAJ}(x) = 0\}$. Let $S = \{x^\alpha : \alpha \in \{0,1\}^n, wt(\alpha) < \frac{n}{2}\}$. We will show that $\text{SM}(U_0) = S$. Since $|S| = |U_0| = |\text{SM}(U_0)|$, it suffices to show $\overline{S} \subseteq \text{LM}(U_0)$. To see this, note that for any $J \subseteq [n]$, where $|J| \geq \frac{n}{2}$, we clearly have $x_J \in I(U_0)$ (because, for any $x \in U_0$, a strict majority of the x_j are 0 and so any sufficiently large product $x_J = \prod_{j \in J}$ must vanish on U_0) and so $x_J \in \text{LM}(U_0)$. Trivially, $x_j^2 \in \text{LM}(U_0)\ \forall j$, as, of course, $x_j^2 + x_j \in I(U_0)\ \forall j$. Due to the fact that $\text{LM}(U_0)$ is upwardly closed, the previous two facts immediately imply $\overline{S} \subseteq \text{LM}(U_0)$, as desired.

Similarly, if $U_1 = \{x \in \mathbb{F}_2^n : \text{MAJ}(x) = 1\}$, then, by the same logic as above, $\text{SM}(U_1) = \{x^\alpha : \alpha \in \{0,1\}^n, wt(\alpha) \leq \frac{n}{2}\}$. Therefore, by definition, $\text{reg}(U_0), \text{reg}(U_1) \leq \frac{n}{2}$. ∎

We now generalize the notion of versatility to functions of the form $f : U \to \mathbb{F}_2$, for some $U \subseteq \mathbb{F}_2^n$. As shown above, a versatile function partitions the set $\mathbb{F}_2^n$, which has regularity n, into two pieces, the preimage of 0 and the preimage of 1, which each have regularity at most $\frac{n}{2}$. We will call a function f δ-versatile on U if the function f induces a partitioning of U with a regularity gap of at least δ. This notion is formalized in the following definition.

Definition 2. *For a function $f : U \to \mathbb{F}_2$, let $U_0 = \{x \in U : f(x) = 0\}$ and $U_1 = \{x \in U : f(x) = 1\}$. We say that f is δ-versatile on U if $\delta \leq reg(U) - reg(U_0), reg(U) - reg(U_1)$.*

Clearly, this notion generalizes the concept of versatility as a versatile function is $\frac{n}{2}$-versatile on $\mathbb{F}_2^n$. We now prove several useful properties of δ-versatile functions which will be used throughout the paper.

Lemma 6. *If $f : U \to \mathbb{F}_2$ is δ-versatile on U then, $\nexists g \in \mathbb{F}_2[x_1, \ldots, x_n]$ where $deg(g) < \delta$ and $g(x) = f(x)\ \forall x \in U$.*

Proof: Assume, for contradiction, that such a g exists. By the definition of regularity, there exists at least one function $h : U \to \mathbb{F}_2$ such that, $\forall q \in \mathbb{F}_2[x_1, \ldots, x_n]$ with $\deg(q) < \text{reg}(U)$, $\exists x \in U$ such that $h(x) \neq q(x)$.

Let $U_0 = \{x \in U : f(x) = 0\}$ and $U_1 = \{x \in U : f(x) = 1\}$. Due to the fact that f is δ-versatile on U we have, by definition, $\text{reg}(U_0), \text{reg}(U_1) \leq \text{reg}(U) - \delta$. Therefore, $\exists u, v \in \mathbb{F}_2[x_1, \ldots, x_n]$ where $\deg(u), \deg(v) \leq \text{reg}(U) - \delta$ and $h(x) = u(x)\ \forall x \in U_0$, $h(x) = v(x)\ \forall x \in U_1$. If we then define $q \in \mathbb{F}_2[x_1, \ldots, x_n]$ by $q = u(g+1) + vg$, we clearly have $\deg(q) \leq \max(\deg(u) + \deg(g), \deg(v) + \deg(g)) \leq (\text{reg}(U) - \delta) + \deg(g) < (\text{reg}(U) - \delta) + \delta = \text{reg}(U)$ and $h(x) = u(x)(g(x) + 1) + v(x)g(x) = q(x)\ \forall x \in U$, which is a contradiction. ∎

Next, we consider the behavior of δ-versatile functions $f : U \to \mathbb{F}_2$ where the set U has a certain special property. Given any $U \subseteq \mathbb{F}_2^n$, there is, of course, a unique multilinear polynomial (recall that a polynomial is multilinear if every monomial has degree at most 1 in each variable) $r_U \in \mathbb{F}_2[x_1, \ldots, x_n]$ such that $r_U(x) = 1$ if and only if $x \in U$. Clearly, $r_U \in I(\overline{V})$. Moreover, each monomial of r_U is in $\text{SM}(\mathbb{F}_2^n)$ (due to the fact that the standard monomials of $\mathbb{F}_2^n$ are precisely the multilinear monomials), and so we immediately conclude that $a(\overline{V}) \leq \deg(r_U)$. We call an algebraic set U *critical* if $a(\overline{V}) = \deg(r_U)$.

Lemma 7. *Let $U \subseteq \mathbb{F}_2^n$ be a critical algebraic set, let $f : U \to \mathbb{F}_2$ be δ-versatile on U, and let $U_0 = \{x \in U : f(x) = 0\}$ and $U_1 = \{x \in U : f(x) = 1\}$. Then, $\forall q \in \mathbb{F}_2[x_1, \ldots, x_n]$ such that deg$(q) < \delta$, $q \in I(U_0)$ if and only if $q \in I(U_1)$.*

Proof: We show that,$\forall q \in \mathbb{F}_2[x_1, \ldots, x_n]$, where $\deg(q) < \delta$, $q \in I(U_0) \Rightarrow q \in I(U_1)$; the reverse implication follows by symmetry. Assume, for contradiction that $q \in I(U_0)$ but $q \notin I(U_1)$. Let $Y = \{x \in U : q(x) = 1\}$. Clearly $Y \subseteq U_1$ and Y is nonempty. Let $t \in \mathbb{F}_2[x_1, \ldots, x_n]$ denote the unique multilinear polynomial such that $t(x) = r_U(x)q(x)\ \forall x \in \mathbb{F}_2^n$, then $t \in I(\overline{Y})$ and $\deg(t) \leq \deg(r_U) + \deg(q)$. Using Lemma 2, we have

$$\text{reg}(Y) = n - a(\overline{Y}) \geq n - \deg(t)$$

$$\geq n - \deg(r_U) - \deg(q) = \text{reg}(U) - \deg(q)$$

$$> \text{reg}(U) - \delta \geq \text{reg}(U_1).$$

However, we cannot possibly have $\text{reg}(Y) > \text{reg}(U_1)$ because, as noted above, $U_1 \subseteq U$, and so, by Lemma 1(b,d) we must have $\text{reg}(Y) \leq \text{reg}(U_1)$. ∎

The following lemma provides an extremely useful characterization of the behavior of a δ-versatile f on the intersection of a critical U with a certain simple algebraic set, namely the union of the vanishing sets of a collection of low degree polynomials.

Lemma 8. *Let $U \subseteq \mathbb{F}_2^n$ be a critical algebraic set, let $f : U \to \mathbb{F}_2$ be δ-versatile on U, and let $U_0 = \{x \in U : f(x) = 0\}$ and $U_1 = \{x \in U : f(x) = 1\}$. For any $d < \delta$ and for any $g_1, \ldots, g_k \in \mathbb{F}_2[x_1, \ldots, x_n]$ where deg$(g_i) < d\ \forall i$, let $G = \cup_i V(g_i)$. Then,*

$$\textit{SM}(U \cap G, j) = \textit{SM}(U_0 \cap G, j) = \textit{SM}(U_1 \cap G, j)\ \forall j \leq \delta - d$$

Proof: Clearly, $U_0 \cap G \subseteq U \cap G$, $U_1 \cap G \subseteq U \cap G$ and so by Lemma 1(d), $\text{SM}(U_0 \cap G), \text{SM}(U_1 \cap G) \subseteq \text{SM}(U \cap G)$, from which it immediately follows that $\text{SM}(U_0 \cap G, j), \text{SM}(U_1 \cap G, j) \subseteq \text{SM}(U \cap G, j)$.

We will now show $\text{SM}(U_0 \cap G, j), \text{SM}(U_1 \cap G, j) \supseteq \text{SM}(U \cap G, j) \forall j \leq \delta - d$, which will complete the proof. Consider any $j \leq \delta - d$. Due to the fact that, for any particular algebraic set, every monomial is either a leading monomial or a standard monomial, if suffices to show $\text{LM}(U_0 \cap G, j), \text{LM}(U_1 \cap G, j) \subseteq \text{LM}(U \cap G, j)$.

To see that $\text{LM}(U_0 \cap G, j) \subseteq \text{LM}(U \cap G, j)$, assume, for contradiction, that this is not the case. Then $\exists x^\alpha \in \text{LM}(U_0 \cap G, j) \cap \text{SM}(U \cap G, j)$. Due to the fact that $x^\alpha \in \text{LM}(U_0 \cap G, j)$ we have, by definition, that $\exists q \in \mathbb{F}_2[x_1, \ldots, x_n]$ such that $q \in I(U_0 \cap G)$ and $\text{lm}(q) = x^\alpha$. Clearly, $\deg(q) = j \leq \delta - d$. Due to the fact that $x^\alpha \in \text{SM}(U \cap G, j)$, we have, by definition $q \notin I(U \cap G)$. This immediately implies $q \notin I(U_1 \cap G)$ because $U \cap G = (U_0 \cup U_1) \cap G = (U_0 \cap G) \cup (U_1 \cap G)$, and so if q did vanish on $U_1 \cap G$, then it would vanish on $U \cap G$ (because, by construction, it vanishes on $U_0 \cap G$). Moreover, since $U_1 \cap G = U_1 \cap (\cup_i V(g_i)) = \cup_i (U_1 \cap V(g_i)$ we conclude $\exists i$ such that $q \notin I(U_1 \cap V(g_i))$. Fix such an i and consider the set $Y = \{x \in U : q(x) = 1 \text{ and } g_i(x) = 0\}$. Notice that due to the requirements that $x \in U$ and $g_i(x) = 0$, we immediately have $Y \subseteq U \cap V(g_i)$, and since q vanishes on $U_0 \cap V(g_i)$, we then have $Y \subseteq U_1 \cap V(g_i)$. Let $t \in \mathbb{F}_2[x_1, \ldots, x_n]$ be the (unique) multilinear polynomial equal to $(r_U)(q)(g_i + 1)$. By construction, $t(x) = 1$ if and only if $x \in Y$, and so $t \in I(\overline{Y})$. We then have

$$a(\overline{Y}) \leq \deg(t) \leq \deg(r_U) + \deg(q) + \deg(g_i + 1)$$

$$< a(\overline{U}) + (\delta - d) + d = a(\overline{U}) + \delta.$$

Applying Lemma 2, we then have

$$\text{reg}(Y) = n - a(\overline{Y}) > n - (a(\overline{U}) + \delta)$$

$$= (n - a(\overline{U})) - \delta = \text{reg}(U) - \delta \geq \text{reg}(U_1),$$

where the last inequality holds due to the fact that f is δ-versatile. However, we cannot possibly have $\text{reg}(Y) > \text{reg}(U_1)$ because, as noted above, $U_1 \subseteq U$, and so, by Lemma 1(b,d) we must have $\text{reg}(Y) \leq \text{reg}(U_1)$. This contradiction allows us to conclude $\text{LM}(U_0 \cap G, j) \subseteq \text{LM}(U \cap G, j)$. By a precisely symmetric argument, $\text{LM}(U_1 \cap G, j) \subseteq \text{LM}(U \cap G, j)$, which completes the proof. ∎

IV. Extractors for Algebraic Sets

In this section, we exhibit a new construction for an extractor for algebraic sets with extremely strong parameters. We begin with the following lemma, which provides a useful bound on the Hilbert function.

Lemma 9. *Let $V \subseteq \mathbb{F}_2^n$ satisfy reg$(V) \geq \frac{n}{2} - \sqrt{n}$. Then, there is a constant $c > 0$ such that, for any $\beta > 0$ and any $k \leq n^{\frac{1}{2}-\beta}$, we have*

$$h^a(V, reg(V)) - h^a(V, reg(V) - k) \leq \frac{ck}{\sqrt{n}}|V|.$$

Proof: Omitted. See full paper [30]. ∎

Remark 1. *The above bound can be seen to be essentially optimal, as shown by considering the standard monomials of the function MAJORITY computed in the previous section.*

We now show that any δ-versatile function, for appropriately chosen δ is an extractor.

Theorem 1. *Let $f : \mathbb{F}_2^n \to \mathbb{F}_2$ be δ-versatile (on $\mathbb{F}_2^n$), where $\delta \geq \frac{n}{2} - n^\gamma$ for some $0 \leq \gamma < \frac{1}{2}$. Then, there is a constant $c > 0$ such that, for any constants α, β such that $0 < \alpha, \beta < \frac{1}{2}$, and for any $d \leq n^\alpha$ and $\rho \geq 2^{-n^\beta}$, f is an extractor with bias $\frac{c\left(n^\gamma + d\log(\frac{\sqrt{n}}{\rho})\right)}{\sqrt{n}}$ for algebraic sets of density at least ρ that are the common zeros of a collection of polynomials each of degree at most d.*

Proof: Let $U_0 = f^{-1}(0)$ and $U_1 = f^{-1}(1)$. Due to the fact that f is $(\frac{n}{2} - n^\gamma)$-versatile, we immediately have reg(U_0), reg$(U_1) \leq \frac{n}{2} + n^\gamma$. We also have reg$(U_0)$, reg$(U_1) \geq \frac{n}{2} - n^\gamma$ because $2^n = |U_0| + |U_1| = |\text{SM}(U_0)| + |\text{SM}(U_1)|$, and the regularity of an algebraic set is the size of its largest standard monomial(Lemma 1(b)).

Consider any algebraic set $V = V(g_1, \ldots, g_k)$ where $g_i \in \mathbb{F}_2[x_1, \ldots, x_n]$ and deg$(g_i) \leq d$ $\forall i$. Using the Razborov-Smolensky method [29],[31], we have a collection of polynomials $y_1, \ldots, y_l \in \mathbb{F}_2[x_1, \ldots, x_n]$ such that deg$(y_i) \leq d$, $V(g_1, \ldots, g_k) \subseteq V(y_1, \ldots, y_l)$ and $|V(y_1, \ldots, y_l) \setminus V(g_1, \ldots, g_k)| \leq 2^{n-l}$. Setting $y = 1 + \prod_{i=1}^{l}(1 + y_i)$, we then have deg$(y) \leq dl$ and $V(y) = V(y_1, \ldots, y_l)$.

Consider $U_0 \cap V(y)$ and $U_1 \cap V(y)$. By Lemma 8, we have

$$\text{SM}(U_0 \cap V(y), i) = \text{SM}(U_1 \cap V(y), i) = \text{SM}(V(y), i),$$

for $i \leq \frac{n}{2} - n^\gamma - dl$. From this, and Lemma 1(a), we immediately conclude $h^a(U_0 \cap V(y), \frac{n}{2} - n^\gamma - dl) = h^a(U_1 \cap V(y), \frac{n}{2} - n^\gamma - dl)$. Clearly, $U_0 \cap V(y) \subseteq U_0$ and $U_1 \cap V(y) \subseteq U_1$, and so, by Lemma 1(d,b), we have reg$(U_0 \cap V(y)) \leq$ reg$(U_0) \leq \frac{n}{2} + n^\gamma$, and reg$(U_1 \cap V(y)) \leq$ reg$(U_1) \leq \frac{n}{2} + n^\gamma$. Moreover, reg$(U_0 \cap V(y))$, reg$(U_1 \cap V(y)) \geq \frac{n}{2} - n^\gamma - dl$. To see this, first notice that Lemma 2 allows us to conclude reg$(V(y)) \geq n - dl$ (because $y + 1$ vanishes on the complement of $V(y)$), which immediately implies that SM$(V(y))$ consists of an element x^κ of degree at least $n - dl$. As SM$(V(y))$ is a dual ideal, we then also conclude that it consists of an element of degree precisely $\frac{n}{2} - n^\gamma - dl$ (simply take any divisor of x^κ of the appropriate degree). By the above relationship between SM$(V(y))$, SM$(U_0 \cap V(y))$ and SM$(U_1 \cap V(y))$, we then conclude that both SM$(U_0 \cap V(y))$ and SM$(U_1 \cap V(y))$ contain an element of degree $\frac{n}{2} - n^\gamma - dl$, and so, by Lemma 1(b), the claimed lower bound on regularity follows. ∎

V. Recursive Fourier Sampling

In this section, we consider the recursive Fourier sampling problem. Numerous variants of this problem have been considered by many authors (see, for instance, [6], [7], [1], [2], [19]). The version considered in this paper, and the notation used, follows most closely [19], but essentially the same claims hold for all other standard variants. We begin by precisely defining the problem.

A. Definition of the Problem

First, we define the Fourier sampling function. For every positive integer n, we define the partial Boolean function $FS_n : \{0,1\}^{2^{n+1}} \to \{0, 1, *\}$ as follows. We interpret the 2^{n+1} bit long input to FS_n as a pair of truth tables defining the functions $f, g : \{0,1\}^n \to \{0,1\}$. For $x, s \in \{0,1\}^n$, let x_i and s_i denote the i^{th} bit of x and s, respectively. Let $x \cdot s = \sum_i x_i s_i$ denote the usual Boolean inner product (where of course the sum is evaluated modulo 2). We then define $FS_n(f, g) = g(s)$ if $\exists s \in \{0,1\}^n$ such that $f(x) = x \cdot s$ $\forall x$ and $FS_n(f, g) = *$ otherwise.

This function can very naturally be interpreted as encoding a promise problem, called the Fourier sampling problem, in which the promise is that f is a linear function (that is to say a function of the form $f(x) = x \cdot s$), and the value of $FS_n(f, g)$ (when the promise is satisfied) is simply $g(s)$. We will frequently refer to the value s as the *secret* encoded by f.

Next, we define a slight variant of the above problem where the function g is fixed (that is to say that it is not part of the input to the function). Formally, for any positive integer n and any function $g : \{0,1\}^n \to \{0,1\}$, we define the function $FS_n^g : \{0,1\}^{2^n} \to \{0,1\}$ as follows. We now interpret the input to the function as encoding the truth table of a single function $f : \{0,1\}^n \to \{0,1\}$. We then define $FS_n^g(f) = g(s)$ if $\exists s \in \{0,1\}^n$ such that $f(x) = x \cdot s$ $\forall x$ and $FS_n^g(f) = *$ otherwise.

We now define the recursive Fourier sampling function, which is a variant of the Fourier sampling function in which each bit of f is produced, recursively, by a smaller instance of the recursive Fourier sampling problem.

Formally, let $RFS_{n,1} : \{0,1\}^{n+2^n} \to \{0,1\}$ be the (total) Boolean function where the input is interpreted as a pair (s, g) for a secret $s \in \{0,1\}^n$ and a function $g : \{0,1\}^n \to \{0,1\}$ given as a 2^n bit long truth table, and $RFS_{n,1}(s, g) = g(s)$.

For each $h > 1$, we define $RFS_{n,h}$ recursively in terms of $RFS_{n,h-1}$ as follows. Let $M_{n,h} = n2^{n(h-1)} + \sum_{j=1}^{h-1} 2^{jn}$. Then $RFS_{n,h} : \{0,1\}^{M_{n,h}} \to \{0,1,*\}$ is the partial Boolean function defined as follows. The input is interpreted as being of the form $(R_0, R_1, \ldots, R_{2^n-1}, g)$, where for each $\sigma \in \{0,1\}^n$, R_σ is an instance of $RFS_{n,h-1}$ and g is a function $g : \{0,1\}^n \to \{0,1\}$ given as a 2^n bit long truth table. We then define $RFS_{n,h}(R_0, \ldots, R_{2^n-1}, g) = g(s)$ if $\exists s \in \{0,1\}^n$ such that $\forall \sigma \in \{0,1\}^n$ $RFS_{n,h-1}(R_\sigma) = \sigma \cdot s$ and $RFS_{n,h}(R_0, \ldots, R_{2^n-1}, g) = *$ otherwise.

In a precisely analogous fashion, we define $RFS^g_{n,h}$ where now there is a single fixed g used throughout the problem, rather than a collection of functions provided as part of the input.

We very naturally interpret $RFS_{n,h}$ and $RFS^g_{n,h}$ as encoding a particular promise problem, where the promise is that, at every node in the tree, there exists some $s \in \{0,1\}^n$ such that the function $f : \{0,1\}^n \to \{0,1\}$ defined at this node is of the form $f(x) = x \cdot s$.

Fix the entire input to the recursive Fourier sampling function in any way such that every promise is satisfied. For any node t in the tree, we define the *value* of the node, which we denote by $b(t)$ to be the output of the instance of recursive Fourier sampling corresponding to the subtree rooted at t.

Notice that, due to the structure of the promise, in order to determine the value of node t, it is only necessary to know the values of n linearly independent children of t. That is to say, if the children of t are given by $C(t) = \{t_\sigma : \sigma \in \{0,1\}^n\}$, then $b(t)$ is completely determined by the value of a subset of children C' for any $C' \subseteq C$ such that $C' = \{t_{\sigma_1}, \ldots, t_{\sigma_n}\}$ where $\{\sigma_1, \ldots, \sigma_n\}$ are linearly independent (as vectors in $\{0,1\}^n$, in other words the σ_i form a basis of $\{0,1\}^n$).

For $i \in [n]$, let $\chi_i \in \{0,1\}^n$ denote the i^{th} elementary basis element. That is to say χ_i has value 1 in position i and 0 elsewhere. Clearly, the set of χ_i form a basis of $\{0,1\}^n$, and so, for any node t, the value of node t is completely determined by the values of these children. We call this set of children the *elementary children* of t, which we denote by $C_e(t) = \{t_{\chi_i} : i \in [n]\}$.

Therefore, given an instance (a particular single setting of the input) of $RFS_{n,h}$ or $RFS^g_{n,h}$ that is guaranteed to satisfy the promise, the answer (the value of the root of the tree) can be determined by first determining the value of the n elementary children of t. The value of each of these children can be determined from their n elementary children. This process can be repeated until the leaves of the tree are reached, at which point the value of each node is simply the output of an instance of $RFS_{n,1}$. We refer to this collection of leaves obtained by repeatedly finding elementary children as the *elementary leaves*. For a tree of height h, there are clearly n^{h-1} elementary leaves.

B. Recursive Fourier Sampling is δ-versatile

In this section, we show that for certain natural choices of the function g, such as the majority function or the generalized inner product function, $RFS^g_{n,h}$ is δ-versatile, for suitably chosen δ.

Fix n, and let m denote the total length of the input to $RFS^g_{n,h}$. Clearly $m = n2^{(h-1)n}$. Let $U^g_{p,h} \subseteq \mathbb{F}_2^m$ denote the set of all points at which all promises are satisfied (that is to say, the set of all values of inputs to the recursive Fourier sampling function such that, at every node of the tree, every linearity constraint is satisfied). We frequently refer to $U^g_{p,h}$ as the "promise". On the promise, the recursive Fourier sampling problem is, of course, a total function. By slight abuse of notation, we also denote this induced total function as $RFS^g_{n,h} : U^g_{p,h} \to \mathbb{F}_2$. Similarly, we define $U^g_{0,h} = (RFS^g_{n,h})^{-1}(0)$ and $U^g_{1,h} = (RFS^g_{n,h})^{-1}(1)$ as the points at which the recursive Fourier sampling problem evaluates to 0 and 1, respectively. The superscript g will often be omitted when the function is clear from context.

The first key result of this section, which holds for any g, is the following lower bound on regularity of $U^g_{p,h}$,$U^g_{0,h}$, and $U^g_{1,h}$.

Lemma 10. *For any positive integers n, h and for any $g \in \mathbb{F}_2[x_1, \ldots, x_n]$, let $d = deg(g)$ and let $RFS^g_{n,h} : \mathbb{F}_2^m \to \mathbb{F}_2$ denote the recursive Fourier sampling function. Then*

$$reg(U^g_{p,h}) \geq nd^{h-1} + (n-d)\sum_{j=1}^{h-1} 2^{jn} d^{h-j-1}$$

$$reg(U^g_{0,h}), reg(U^g_{1,h}) \geq (n-d)\sum_{j=0}^{h-1} 2^{jn} d^{h-j-1}.$$

Proof: Omitted. See full paper [30]. ∎

We now exhibit certain functions for which the above lower bounds on regularity are exact. The first such example is the majority function, for certain appropriately chosen input sizes. For a $x \in \{0,1\}^n$, let $x = (x_1, \ldots, x_n)$ and let $wt(x) = |\{i : x_i = 1\}|$ denote the number of 1s in x. Let MAJ : $\mathbb{F}_2^n \to \mathbb{F}_2$ be defined such that MAJ$(x) = 1$ if and only if $wt(x) \geq \frac{n}{2}$. We begin by determining the unique squarefree polynomial in $\mathbb{F}_2[x_1, \ldots, x_n]$ that represents MAJ. Let $e_i(x) = \sum_{J \subseteq [n], |J|=i} \prod_{j \in J} x_j$ denote the i^{th} elementary symmetric polynomial. For $y, z \in \{0,1\}^l$, write $y \geq_b z$ if and only if $y_i \geq z_i$ $\forall i$.

Lemma 11. *For any positive integer n, the unique squarefree polynomial in $\mathbb{F}_2[x_1, \ldots, x_n]$ that is identically equal to MAJ : $\mathbb{F}_2^n \to \mathbb{F}_2$ on $\mathbb{F}_2^n$ is given by*

$$\sum_{l \geq \frac{n}{2}} \sum_{j \geq_b l} e_j(x).$$

Proof: Omitted. See full paper [30]. ∎

We now consider $RFS^{\text{MAJ}}_{n,h}$. We begin by demonstrating a useful symmetry in $U^{\text{MAJ}}_{p,h}$. Define the value $\hat{1}_h \in U^{\text{MAJ}}_{p,h}$ as

follows. Consider the recursive Fourier sampling tree. We define $\hat{1}_h$ by first defining $b(t)$ for every node t in the tree (that is to say, we define the value $b(t)$ that node t has with input $\hat{1}_h$). First, assign the root of the tree the value 1. Then, for each node that has been assigned a value, assign values to the children of that node as follows. If node t has value $b(t)$, then set $b(t_\sigma) = b(t)$ for each $t_\sigma \in C_e(t)$. Assign all other children the value forced by the promise: for each $t_\sigma \in C(t) \setminus C_e(t)$, set $b(t_\sigma) = \sum_{j\in[n],\sigma_j=1} b(t_{\chi_j})$. Equivalently, if a node has value 0, all of its children have value 0; if a node has value 1, then each child t_σ has value given by the parity of the string σ. Once the entire tree has been labeled in such a fashion, define $\hat{1}_h$ by setting the portion of the input corresponding to each leaf (that is to say, the n places of the input representing the secret at that leaf) to the value of that leaf.

It is clear that the value $\hat{1}_h \in U^{\text{MAJ}}_{p,h}$ as claimed, due to the fact that $\hat{1}_h$ was constructed in a way such that the promise is satisfied at every node. Moreover, $\hat{1}_h \in U^{\text{MAJ}}_{1,h}$ as, by construction, the value of the root is 1. For any $x \in U^{\text{MAJ}}_{p,h}$, let $\hat{x} = x \oplus \hat{1}_h$ (where $\oplus$ denotes bitwise parity). We then have the following.

Lemma 12. *For any odd positive integer n and any positive integer h, $x \in U^{MAJ}_{0,h}$ if and only if $\hat{x} \in U^{MAJ}_{1,h}$.*

Proof: Given any $x \in U^{\text{MAJ}}_{0,h}$, the root of the corresponding recursive Fourier sampling tree has value 0. The key observation is that adding $\hat{1}_h$ flips the value at every elementary leaf of the tree. That is to say, if on input x, a particular elementary leaf t has value $b \in \{0,1\}$, then on input $\hat{x}$, that leaf has value $\bar{b}$. This occurs because, by construction, $\hat{1}_h$ is 1 at every position in the elementary leaves. It is then straightforward to see that value of the root of the tree flips and that every promise is preserved, which implies $\hat{x} \in U^{\text{MAJ}}_{1,h}$. The reverse implication follows from the fact that $\hat{\hat{x}} = x$ and symmetry. ∎

We now show that $U^{\text{MAJ}}_{0,h}$ and $U^{\text{MAJ}}_{1,h}$ have identical standard monomials.

Lemma 13. *For any odd positive integer n and any positive integer h, $SM(U^{MAJ}_{0,h}) = SM(U^{MAJ}_{1,h})$.*

Proof: For any algebraic set, every monomial is either a leading monomial or a standard monomial, and so it suffices to show $\text{LM}(U^{\text{MAJ}}_{0,h}) = \text{LM}(U^{\text{MAJ}}_{1,h})$.

We first show $\text{LM}(U^{\text{MAJ}}_{0,h}) \subseteq \text{LM}(U^{\text{MAJ}}_{1,h})$. Consider any $x^\alpha \in \text{LM}(U^{\text{MAJ}}_{0,h})$. By definition, $\exists q_\alpha \in \mathbb{F}_2[x_1,\ldots,x_m]$ such that $q_\alpha \in I(U^{\text{MAJ}}_{0,h})$ and $\text{lm}(q_\alpha) = x^\alpha$. Define $\hat{q}_\alpha \in \mathbb{F}_2[x_1,\ldots,x_m]$ such that $\hat{q}_\alpha(x) = q_\alpha(\hat{x})$. Notice that

$$\text{lm}(\hat{q}_\alpha) = \text{lm}(q_\alpha) = x^\alpha.$$

Moreover, for any $x \in U^{\text{MAJ}}_{1,h}$, Lemma 12 implies that $\hat{x} \in U^{\text{MAJ}}_{0,h}$ and so

$$\hat{q}_\alpha(x) = q_\alpha(\hat{x}) = 0,$$

where the last follows from the fact that q vanishes on $U^{\text{MAJ}}_{0,h}$. This implies that $\hat{q}_\alpha \in I(U^{\text{MAJ}}_{1,h})$, and so $x^\alpha \in \text{LM}(U^{\text{MAJ}}_{1,h})$. Therefore, $\text{LM}(U^{\text{MAJ}}_{0,h}) \subseteq \text{LM}(U^{\text{MAJ}}_{1,h})$.

A precisely symmetric argument implies $\text{LM}(U^{\text{MAJ}}_{0,h}) \supseteq \text{LM}(U^{\text{MAJ}}_{1,h})$. ∎

Next, we provide upper bounds for the regularity of $U^{\text{MAJ}}_{p,h}$, $U^{\text{MAJ}}_{0,h}$, and $U^{\text{MAJ}}_{1,h}$.

Lemma 14. *For any odd positive integer n and any positive integer h, let $RFS^{MAJ}_{n,h} : \mathbb{F}_2^m \to \mathbb{F}_2$ denote the recursive Fourier sampling function with majority. Then*

$$reg(U^{MAJ}_{p,h}) \leq \left(\frac{n-1}{2}\right)\sum_{j=1}^{h-1} 2^{jn}\left(\frac{n+1}{2}\right)^{h-j-1} + n\left(\frac{n+1}{2}\right)^{h-1}$$

$$reg(U^{MAJ}_{0,h}), reg(U^{MAJ}_{1,h}) \leq \left(\frac{n-1}{2}\right)\sum_{j=0}^{h-1} 2^{jn}\left(\frac{n+1}{2}\right)^{h-j-1}.$$

Proof: Omitted. See full paper [30]. ∎

We now conclude that, for appropriately chosen input size, $RFS^{\text{MAJ}}_{n,h}$ is versatile.

Lemma 15. *Let $n = 2^k - 1$ for any positive integer k, then $RFS^{MAJ}_{n,h}$ is $\left(\frac{n+1}{2}\right)^h$-versatile on $U^{MAJ}_{p,h}$. Moreover, $U^{MAJ}_{p,h}$ is a critical algebraic set.*

Proof: Omitted. See full paper [30]. ∎

Next, we exhibit another class of functions such that the lower bound on regularity in Lemma 10 is tight. Consider any $g \in \mathbb{F}_2[x_1,\ldots,x_n]$ and let $d = \deg(g)$. $V_0 = g^{-1}(0)$ and $V_1 = g^{-1}(1)$ denote the preimages of 0 and 1, respectively. For any $k \times n$ matrix A with entries in $\mathbb{F}_2$, let $\phi_A : \mathbb{F}_2^n \to \mathbb{F}_2^k$ denote the linear map defined by A. We say a function g is *well-mixed* if, for every $n-d+1 \times n$ matrix A, $\frac{V_0}{\ker\phi_A} \ncong \mathbb{F}_2^{n-d+1}$ and $\frac{V_1}{\ker\phi_A} \ncong \mathbb{F}_2^{n-d+1}$. We then have the following.

Lemma 16. *For any positive integers n, h, let $g \in \mathbb{F}_2[x_1,\ldots,x_n]$ be well-mixed. Let $d = deg(g)$ and let $RFS^g_{n,h} : \mathbb{F}_2^m \to \mathbb{F}_2$ denote the recursive Fourier sampling function with g. Then*

$$reg(U^g_{p,h}) \leq nd^{h-1} + (n-d)\sum_{j=1}^{h-1} 2^{jn} d^{h-j-1}$$

$$reg(U^g_{0,h}), reg(U^g_{1,h}) \leq (n-d)\sum_{j=0}^{h-1} 2^{jn} d^{h-j-1}.$$

Proof: Omitted. See full paper [30]. ∎

This immediately allows us to conclude that, for any well-mixed g, $RFS^g_{n,h}$ is versatile, as shown in the following lemma.

Lemma 17. *For any positive integers n, h, let $g \in \mathbb{F}_2[x_1, \ldots, x_n]$ be well-mixed. Let $d = deg(g)$ and let $RFS^g_{n,h} : \mathbb{F}_2^m \to \mathbb{F}_2$ denote the recursive Fourier sampling function with g. Then $RFS^g_{n,h}$ is d^h-versatile on $U^g_{p,h}$ and $U^g_{p,h}$ is a critical algebraic set.*

Proof: Omitted. See full paper [30]. ■

We now show that a certain natural function, the generalized inner product function, is well-mixed, and therefore the corresponding version of recursive Fourier sampling is versatile. For any positive integer n and any $d|n$, let $\text{GIP}_{n,d} : \mathbb{F}_2^n \to \mathbb{F}_2$ be defined such that

$$\text{GIP}_{n,d} = x_1 \cdots x_d + x_{d+1} \cdots x_{2d} + \ldots + x_{n-d+1} \cdots x_n.$$

Notice that the ordinary inner product function simply corresponds to the case in which $d = 2$.

Lemma 18. *For any positive integers d, n such that $d|n$, and $n \geq d(2^{d^2} + d - 1)$, the function $GIP_{n,d} : \mathbb{F}_2^n \to \mathbb{F}_2$ is well-mixed. Moreover, the function $RFS^{GIP_{n,d}}_{n,h}$ is d^h-versatile on $U^{GIP_{n,d}}_{p,h}$ and $U^{GIP_{n,d}}_{p,h}$ is a critical algebraic set.*

Proof: Omitted. See full paper [30]. ■

C. Polynomial Degree

Using the results of the previous section, we now prove very strong statements about the degree of any polynomial that computes, or even one-sided agrees with, the recursive Fourier sampling problem.

Theorem 2. *For any positive integers k, h, Let $n = 2^k - 1$ and let $RFS^{MAJ}_{n,h}$ denote the recursive Fourier sampling function with majority. Then $\nexists g \in \mathbb{F}_2[x_1, \ldots, x_m]$ such that $deg(g) < \left(\frac{n+1}{2}\right)^h$ and $g(x) = RFS^{MAJ}_{n,h}(x)$ $\forall x \in U^{MAJ}_{p,h}$. Moreover, if any $g \in \mathbb{F}_2[x_1, \ldots, x_m]$ such that $deg(g) < \left(\frac{n+1}{2}\right)^h$ vanishes everywhere on $U^{MAJ}_{0,h}$, it vanishes everywhere on $U^{MAJ}_{1,h}$.*

Proof: By Lemma 15, $RFS^{\text{MAJ}}_{n,h}$ is $\left(\frac{n+1}{2}\right)^h$-versatile on $U^{\text{MAJ}}_{p,h}$ and $U^{\text{MAJ}}_{p,h}$ is a critical algebraic set. The first claim of the theorem is an immediate consequence of Lemma 6 and the second claim is an immediate consequence of Lemma 7. ■

Theorem 3. *For any positive integers d, n, h such that $d|n$, and $n \geq d(2^{d^2} + d - 1)$, Let $RFS^{GIP_{n,d}}_{n,h}$ denote the recursive Fourier sampling function with generalized inner product. Then $\nexists g \in \mathbb{F}_2[x_1, \ldots, x_m]$ such that $deg(g) < d^h$ and $g(x) = RFS^{GIP_{n,d}}_{n,h}(x)$ $\forall x \in U^{GIP_{n,d}}_{p,h}$. Moreover, if any $g \in \mathbb{F}_2[x_1, \ldots, x_m]$ such that $deg(g) < d^h$ vanishes everywhere on $U^{GIP_{n,d}}_{0,h}$, it vanishes everywhere on $U^{GIP_{n,d}}_{1,h}$.*

Proof: By Lemma 18, $RFS^{\text{GIP}_{n,d}}_{n,h}$ is d^h-versatile on $U^{\text{GIP}_{n,d}}_{p,h}$ and $U^{\text{GIP}_{n,d}}_{p,h}$ is a critical algebraic set. The first claim of the theorem is an immediate consequence of Lemma 6 and the second claim is an immediate consequence of Lemma 7. ■

D. Towards a Circuit Lower Bound

In the previous section, an extremely strong lower bound was given on the lowest degree polynomial over $\mathbb{F}_2$ that computes (or even non-trivially one-sided agrees with) the recursive Fourier sampling function on the promise. In this section, we discuss partial results towards a lower bound on the size of a constant depth circuit that computes the recursive Fourier sampling function, as well as what sort of additional results would allow these partial results to be extended to prove such a lower bound. We begin by defining the circuit class of interest. Let n denote, as before, the size of the secret at each node of the recursive Fourier sampling tree, and h denote the height of the recursive Fourier sampling tree. We consider circuits that consist of *AND*, *OR*, and *NOT* gates, where the fan-in of the *AND* and *OR* gates is unbounded, the size of the circuit (the total number of gates) is at most $2^{O(poly(n))}$, and the depth of the circuit (the number of gates on the longest path from the input to the output) is a constant (independent of n and h). This circuit class is of interest due to the fact that proving a lower bound against it (that is to say, proving that no circuit of this form can compute the recursive Fourier sampling function on its promise), would immediately imply the existence of an oracle A such that $\text{BQP}^A \not\subset \text{PH}^A$. This follows due to the relationship between such circuits and the polynomial hierarchy ([15],[34]) and the fact that there is an efficient quantum algorithm for the recursive Fourier sampling problem ([6],[1],[19]), when $h = O(\log n)$. Such a bound is at least plausible as the trivial circuit (which simply computes the recursive Fourier sampling in the brute force, level-by-level way, in which each subproblem is solved by solving n subproblem one level down) has size $2^{\theta(n^h)}$, which, when $h = \theta(\log n)$ is, of course, not $2^{O(poly(n))}$.

One reasonable approach to proving such a lower bound would be to apply a variant of the Razborov-Smolensky method [29],[31]. We begin by briefly sketching the main idea of the Razborov-Smolensky method, specialized to $\mathbb{F}_2$. We consider a (total) function $g : \mathbb{F}_2^m \to \mathbb{F}_2$, where $m = 2^{O(poly(n))}$. We wish to show that no circuit C of the above form, of size at most $2^{O(poly(\log m))} = 2^{O(poly(n))}$, can compute the function g. The key observation is that there is an $f \in \mathbb{F}_2[x_1, \ldots, x_m]$ where $\deg(f) = O(poly(n))$ such that f agrees with C almost everywhere, and so if it can be shown that g is not well approximated by a low degree polynomial, it immediately follows that g is not actually computed by C. To show that a particular g cannot agree almost everywhere with a low degree polynomial, it suffices to show that g has the property that, on any set $R \subseteq \mathbb{F}_2^n$, if g is represented on R by a polynomial of degree at most d, then every function $q : R \to \mathbb{F}_2$ is represented on R by a polynomial of degree not much higher than d. This

suffices because if g agrees almost everywhere with a low degree polynomial f, then there is a very large set R on which every function $q : R \to \mathbb{F}_2$ is represented by a low-degree polynomial; a straightforward counting of the number of functions of that form and the number of low-degree polynomials shows that this is impossible.

The main idea behind the lower bound on the polynomial degree of recursive Fourier sampling, shown in the previous sections, is that there are functions g such that $RFS^g_{n,h}$ has the property that there is a large gap between the regularity of the promise, reg$(U^g_{p,h})$, and the regularities of the preimages of 0 and 1, reg$(U^g_{0,h})$ and reg$(U^g_{1,h})$. In other words, there are functions on $U^g_{p,h}$ which can only be computed by relatively high degree polynomials, whereas every function on $U^g_{0,h}$ and $U^g_{1,h}$ can be computed by relatively low degree polynomials. It then follows that $RFS^g_{n,h}$ itself cannot be computed on $U^g_{p,h}$ by a low degree polynomial, because if it were, then every function on $U^g_{p,h}$ would be computable by a low degree polynomial.

While this is very similar to the observation made in the Razborov-Smolensky method, there is one crucial difference: due to the fact that the promise $U^g_{p,h}$ is extremely small, one cannot conclude, via a straightforward counting argument, that there is a function on $U^g_{p,h}$ that requires a high degree polynomial; instead, this fact was shown via an analysis of the structure of this algebraic set. It is the very fact that such an analysis is possible that gives hope that this technique could be extended to prove the desired circuit lower bound. To be precise, to prove the desired circuit lower bound, it would suffice to show that, not merely is it the case that $RFS^g_{n,h}$ is $\omega(poly(n))$-versatile on $U^g_{p,h}$, as already shown, but in fact $RFS^g_{n,h}$ is $\omega(poly(n))$-versatile on R for any sufficiently large $R \subseteq U^g_{p,h}$. This would suffice because, if $RFS^g_{n,h}$ had this property, then it could not be the case that $RFS^g_{n,h}$ is well approximated by a low degree polynomial on $U^g_{p,h}$, from which it would then follow that $RFS^g_{n,h}$ is not computed by a small circuit on $U^g_{p,h}$. In fact, something substantially weaker would suffice: one only needs to consider the case in which R is of the form $U^g_{p,h} \cap V(f_1, \ldots, f_k)$ where each $f_i \in \mathbb{F}_2[x_1, \ldots, x_m]$ satisfies $\deg(f_i) = O(poly(n))$. In other words, one only needs to consider the case in which R is a large subset of $U^g_{p,h}$ such that R is the intersection of $U^g_{p,h}$ with an algebraic set that is the set of common zeros of a collection of low degree polynomials. This suffices because, much as was done in Braverman's proof of the Linial-Nisan conjecture [9], one can consider the structure of the set of points on which a small circuit agrees with the low degree polynomial produced by the Razborov-Smolensky method. To be precise, consider applying the Razborov-Smolensky method to a *AND* of a collection of polynomials $p_1, \ldots, p_k \in \mathbb{F}_2[x_1, \ldots, x_m]$ where $\deg(p_i) = O(poly(n))\ \forall i$. This *AND* of low degree polynomials is well approximated by a single $p' \in \mathbb{F}_2[x_1, \ldots, x_m]$, given by the product of a collection of a small number of randomly chosen sums of the p_i. Moreover, the output of the *AND* of $p_1, \ldots, p_k$ agrees with p' precisely on $V(p'(1 + p_1), \ldots, p'(1 + p_k))$. Repeating this process for every gate in the circuit, from the bottom up, yields an algebraic set of the form $V = V(f_1, \ldots, f_k)$ where $\deg(f_i) = O(poly(n))\ \forall i$, where, on V, each gate individually agrees with its approximating polynomial. To be clear, this algebraic set V is a (possibly proper) subset of the set of points on which the circuit agrees with the overall approximating polynomial, due to the fact that a local mistake (that is to say, a point at which an individual gate disagreeing with its approximating polynomial) may not propagate through the entire circuit to yield a global mistake (that is to say, a point at which the circuit disagrees with the approximating polynomial); however, the extremely simple form of V makes it a natural choice for performing the required analysis of regularity.

While the current analysis falls short of being able to prove the type of circuit lower bound needed for the desired relativized separation result, it does produce some interesting partial results. For example, consider any circuit C consisting of an *OR* of a collection $p_1, \ldots, p_k \in \mathbb{F}_2[x_1, \ldots, x_m]$ where $\deg(p_i) \leq d = O(poly(n))\ \forall i$. Circuits of this type are interesting as it can easily be seen that if one can prove that such a circuit cannot be a good approximator with one-sided error of the recursive Fourier sampling problem on its promise (where we say C is a good approximator with one-sided error if C outputs 1 everywhere on $U^g_{1,h}$ and outputs 0 almost everywhere on $U^g_{0,h}$) this would immediately yield the existence of an A such that $\text{BQP}^A \not\subset \text{AM}^A$. The existing analysis does provide some insight into the behavior of any such circuit on the promise, though it, unfortunately, falls short of proving the required lower bound. To be precise, by noting that the set of points on which C outputs one is given by $V = \cup_i V(1+p_i)$, and applying Lemma 8, one can immediately conclude that, for any g such that $RFS^g_{n,h}$ is δ-versatile on $U^g_{p,h}$,

$$\text{SM}(U^g_{p,h} \cap V, j) = \text{SM}(U^g_{0,h} \cap V, j) = \text{SM}(U^g_{1,h} \cap V, j),$$

where $j \leq \delta - d$. This is, by itself, a very strong statement about the structure of the set of points on which any such circuit C evaluates to 1. Moreover, due to the fact that, by Lemma 1(c), the size of any algebraic set is equal to the size of the set of standard monomials of that set, the above claim also yields a (weak) statement about the relationship between the sizes of $U^g_{0,h} \cap V$ and $U^g_{1,h} \cap V$.

VI. VC Dimension

In this section, we answer an open question posed in [26]. We begin with a few definitions. We begin by recalling several key results from that paper.

Lemma 19. *[26](Thm.2.2) For any $C \subseteq \{0,1\}^n$, reg$(C) \leq$ VC(C).*

The following result, expressed in the terminology of this paper, was then shown.

Lemma 20. *[26](Prop.6.1) For any $C \subseteq \{0,1\}^n$, reg$(C) = 1$ precisely when rank$_{\mathbb{F}_2}\mathcal{M}(C, \binom{[n]}{\leq 1}) = |C|$*

They then asked if there was a similar simple characterization of when reg$(C) = r$, for $r > 1$, which would be highly desirable as any such characterization would, by the above lemma, provide a characterization of sets with VC dimension at least r. We show the following.

Theorem 4. *A set $C \subseteq \{0,1\}^n$ has reg$(C) = r$ if and only if r is the smallest positive integer such that rank$_{\mathbb{F}_2}\mathcal{M}(C, \binom{[n]}{\leq r}) = |C|$.*

Proof: By Lemma 3, $h^a(C,d) = \text{rank}_{\mathbb{F}_2}\mathcal{M}(C, \binom{[n]}{\leq d})$. By definition, reg$(C)$ is the minimum r such that $h^a(C,r) = |C|$. ■

ACKNOWLEDGMENTS

I am immensely grateful to my advisor, Michael Sipser, for the assistance, feedback and guidance that he provided throughout this research. I also thank Ravi Boppana for bringing [26] to my attention, and for helpful discussions, as well as Scott Aaronson for bringing the recursive Fourier sampling problem to my attention and providing feedback on an earlier draft of this paper.

REFERENCES

[1] S. Aaronson, *Quantum lower bound for recursive Fourier sampling*, Quantum Information and Computation (2003), 3(2):165-174.

[2] S. Aaronson, *BQP and the polynomial hierarchy*, In Stoc '10: : Proceedings of the forty-second annual ACM symposium of Theory of computing (2010), 141-150.

[3] M. Ajtai, Σ_1^1*-Formulae on finite structure*, APAL (1983).

[4] N. Alon, S. Moran, A. Yehudayoff, *Sign rank, VC dimension and spectral gaps*, ECCC (2014), 21:135.

[5] R. Beals, H. Buhrman, R. Cleve, M. Mosca, and R. de Wolf, *Quantum lower bounds by polynomials*, In IEEE Symposium on Foundations of Computer Science (1998), 352-361.

[6] E. Bernstein and U. Vazirani, *Quantum complexity theory*, In STOC '93: Proceedings of the twenty-fifth annual ACM symposium of Theory of computing (1993), 11-20.

[7] E. Bernstein and U. Vazirani, *Quantum complexity theory*, SIAM J. Comput. (1997), 26(5):1411-1473.

[8] J. Bourgain, *On the construction of affine extractors*, Geometric and Functional Analysis (2007), 17(1):33-57.

[9] M. Braverman, *Poly-logarithmic independence fools AC0 circuits*, IEEE Conference on Computational Complexity (2009), 3-8.

[10] G. Cohen and A. Tal, *Two structural results for low degree polynomials and applications*, ECCC (2013), TR. No. 145.

[11] Z. Dvir, A. Gabizon, and A. Wigderson, *Extractors and rank extractors for polynomial sources*, In FOCS '07 (2007).

[12] Z. Dvir, *Extractors for varieties*, Computational Complexity (2012), 21(4):515-572.

[13] D. Eisenbud, *The Geometry of Syzygies*, (2002).

[14] B. Felszeghy, *Grobner Theory of Zero Dimensional Ideals with a View Towards Combinatorics*, Budapest University (2007), Ph. D. thesis.

[15] M. Furst, J. Saxe, and M. Sipser, *Parity, circuits, and the polynomial time hierarchy*, Mathematical Systems Theory (1984), 17:13-27.

[16] A. Gabizon and R. Raz, *Deterministic extractors for affine sources over large fields*, In Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science (2005), 407-418.

[17] A. Gabizon, R. Raz, and R. Shaltiel, *Deterministic extractors for bit-fixing sources by obtaining an independent seed*, In Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science (2004), 394-403.

[18] J. Hastad, *Computational limitations for small depth circuits*, MIT Press (1986), Ph. D. thesis.

[19] B. Johnson, *Upper and lower bounds for recursive Fourier sampling*, University of California at Berkeley (2008), Ph. D. thesis.

[20] B. Johnson, *The polynomial degree of recursive Fourier sampling*, Theory of Quantum Computation, Communication, and Cryptography in Lecture Notes in Computer Science (2011), 6519:104-112.

[21] J. Kamp, A. Rao, S. Vadhan, and D. Zuckerman, *Deterministic extractors for small-space sources*, In Proceedings of the thirty-eighth annual ACM symposium on Theory of computing (2006), 691-700.

[22] J. Kamp and D. Zuckerman, *Deterministic extractors for bit-fixing sources and exposure-resilient cryptography*, In Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science (2003).

[23] S. Kopparty, *On the complexity of powering in finite fields*, In Stoc '11: : Proceedings of the forty-third annual ACM symposium of Theory of computing (2011), 489-498.

[24] N. Linial, Y. Mansour, and N. Nisan, *Constant depth circuits, Fourier transform, and learnability*, J. ACM (1993), 40(3):607-620.

[25] L. Lovász, *Combinatorial problems and exercises*, North-Holland, Amsterdam (1979), 13.31.

[26] S. Moran and C. Rashtchian, *Shattered sets and the Hilbert function*, ECCC (2015), TR. No. 15-189.

[27] S. Moran, A. Shpilka, A. Wigderson, A. Yehudayoff, *Teaching and compressing for low vc-dimension*, In FOCS (2015).

[28] D. Pintér and L. Rónyai, *On the Hilbert Function of Complementary Set Families*, Annales Univ. Sci. Budapest Sect. Comp. (2008), 29:175-198.

[29] A. A. Razborov, *Lower bounds on the size of bounded depth circuits over a complete basis with logical addition*, Mathematical Notes (1987), 41(4):333-338.

[30] Z. Remscrim, *The Hilbert Function, Algebraic Extractors, and Recursive Fourier Sampling*, ECCC TR16-020 (2016).

[31] R. Smolensky, *Algebraic methods in the theory of lower bounds for Boolean circuit complexity*, Proceedings of the nineteenth annual ACM symposium on Theory of computing (1987), 77-82.

[32] R. Smolensky, *On Representations by Low-degree Polynomials*, FOCS (1993).

[33] L. Trevisan and S. Vadhan, *Extracting randomness from samplable distributions*, In Proceedings of the 41st Annual Symposium of Foundations of Computer Science (2000), 32.

[34] A. Yao, *Separating the polynomial-time hierarchy by oracles (preliminary version)*, In Proc. IEEE FOCS (1985), 1-10.