Electronic Colloquium on Computational Complexity, Report No. 20 (2016)

The Hilbert Function, Algebraic Extractors, and Recursive Fourier Sampling

Zachary Remscrim*
MIT
Cambridge, MA 02139

February 8, 2016

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 $\mathbb{F}_2$ of substantially higher degree than the current state-of-the-art construction. We also exactly determine the $\mathbb{F}_2$-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 posed in [MR15] concerning VC dimension, interpolation degree and the Hilbert function.

1 Introduction

1.1 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 $\mathcal{C}$ to have 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 question is then: for which $\mathcal{C}$ do there exist explicit constructions of extractors?

Numerous versions of this question have been considered. In this paper, we consider the case, originally introduced in [Dvi12], 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 [GR05] and [Bou07]) and *bit-fixing sources* (see, for instance, [GRS04] and [KZ03]) and build naturally on the earlier work of *efficiently samplable sources* (see, for instance, [TV00], [KRVZ06], and [DGW07]).

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

*remscrim@mit.edu

ISSN 1433-8092

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 $\mathbb{F}$. Clearly any extractor is also a disperser.

As shown in [Dvi12], 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 [CT13], 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.

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.*

Much as was the case in [Dvi12] and [CT13], 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).

1.2 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 [BV93], the 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 $\text{BQP}^A \not\subset \text{NP}^A$ and $\text{BQP}^A \not\subset \text{MA}^A$. 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 [Aar10], that there is an oracle A such that $\mathrm{BQP}^A \not\subset \mathrm{BPP}^A_{\mathrm{path}}$ and $\mathrm{BQP}^A \not\subset \mathrm{SZK}^A$, even the question of whether or not there exists an oracle A such that $\mathrm{BQP}^A \not\subset \mathrm{AM}^A$ remains open, as does, of course, the substantially stronger question of whether or not there exists an oracle A such that $\mathrm{BQP}^A \not\subset \mathrm{PH}^A$.

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 [FSS84],[Yao85]. 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 $\oplus\mathrm{P}^A \not\subset \mathrm{PH}^A$. 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, [FSS84],[Ajt83],[Has86],[Raz87],[Smo87], one might very naturally ask why the question of whether or not there exists an A such that $\mathrm{BQP}^A \not\subset \mathrm{PH}^A$ 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 [Raz87] or [Smo87]) or implicitly (in the case of [FSS84],[Ajt83],[Has86] as shown by [LMN93]) argue that f cannot be well approximated by a low-degree polynomial. This is a problem because, as shown in [BBC$^+$98], any function that can be computed by an efficient quantum algorithm is well approximated by a low degree polynomial.

More precisely, however, [BBC$^+$98] 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 [BBC$^+$98] 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 $\mathrm{BQP}^A \not\subset \mathrm{PH}^A$. It is this issue that we focus on within this paper.

As has been observed by many authors (for instance [BV93],[BV97],[Aar03],[Joh08],[Aar10]) the recursive Fourier sampling problem (or a slight variant) is a prime candidate for exhibiting an oracle A such that $\mathrm{BQP}^A \not\subset \mathrm{PH}^A$, 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 [BBC$^+$98], 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 [Joh11], which shows that there is an even lower degree real polynomial than the one guaranteed by [BBC+98] which exactly represents the recursive Fourier sampling problem on its promise. Moreover, [Joh11] 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 $U_{1,h}^{GIP_{n,d}}$.*

1.3 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

$$\mathrm{str}(C) = \{J \subseteq [n] : J \text{ is shattered by} C\}$$

to denote the sets that are shattered with respect to C. We then define the *VC dimension* of C as

$$\mathrm{VC}(C) = \max\{|J| : J \in \mathrm{str}(C)\}.$$

For a field $\mathbb{F}$ and a set $C \subseteq \{0,1\}^n$, the interpolation degree of C, denoted by $\mathrm{reg}(C)$ 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 [MR15], 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.

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 $rank_{\mathbb{F}_2}\mathcal{M}(C, \binom{[n]}{\leq r}) = |C|$.*

1.4 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 [Kop11]. In §4, we exhibit an 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 [MR15] concerning interpolation degree and VC-dimension.

2 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 $V(f_1, \ldots, f_k)$ where

$$V(f_1, \ldots, f_k) = \{(x_1, \ldots, x_n) \in \mathbb{F}^n : f_i(x_1, \ldots, x_n) = 0 \ \forall i\}.$$

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 [Eis02] 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

$$\mathrm{SM}(I) = \{x^\alpha : \alpha \in \mathbb{N}^n\} \setminus \mathrm{LM}(I).$$

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 illustracted in the following lemma (these are well known facts in algebraic geometry; see, for instance [Fel07]).

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 [Fel07] and [PR08], provides an extremely useful relationship between $\mathrm{reg}(V)$ and $a(\overline{V})$, where $\overline{V}$ denotes the complement of V.

Lemma 2. *[Fel07], [PR08]*

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*

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

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 [Smo93], 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 analagous 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).

3 Generalization of Versatile Functions

In this section, we consider a certain natural generalization of the concept of versatile functions (as defined in [Kop11], see also [Smo87] 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 [Smo93]), 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 $\mathrm{reg}(U_0), \mathrm{reg}(U_1) \leq \frac{n}{2}$.

If $\mathrm{reg}(U_0), \mathrm{reg}(U_1) \leq \frac{n}{2}$, 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 [Kop11], the Majority function (the function $\mathrm{MAJ} : \mathbb{F}_2^n \to \mathbb{F}_2$ where $\mathrm{MAJ}(x) = 1$ when $wt(x) \geq \frac{n}{2}$ and $\mathrm{MAJ}(x) = 0$ 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 $U_0 = \{x \in \mathbb{F}_2^n : \mathrm{MAJ}(x) = 0\}$. Let $S = \{x^\alpha : \alpha \in \{0,1\}^n, wt(\alpha) < \frac{n}{2}\}$. We will show that $\mathrm{SM}(U_0) = S$. Since $|S| = |U_0| = |\mathrm{SM}(U_0)|$, it suffices to show $\overline{S} \subseteq \mathrm{LM}(U_0)$. 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 $x_J \in \mathrm{LM}(U_0)$. Trivially, $x_j^2 \in \mathrm{LM}(U_0)\ \forall j$, as, of course, $x_j^2 + x_j \in I(U_0)\ \forall j$. Due to the fact that $\mathrm{LM}(U_0)$ is upwardly closed, the previous two facts immediately imply $\overline{S} \subseteq \mathrm{LM}(U_0)$, as desired.

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

□

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 $\deg(q) < \mathrm{reg}(U)$, $\exists x \in U$ 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, $\mathrm{reg}(U_0), \mathrm{reg}(U_1) \leq \mathrm{reg}(U) - \delta$. Therefore, $\exists u, v \in \mathbb{F}_2[x_1, \ldots, x_n]$ where $\deg(u), \deg(v) \leq \mathrm{reg}(U) - \delta$ 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 $\deg(q) \leq \max(\deg(u) + \deg(g), \deg(v) + \deg(g)) \leq (\mathrm{reg}(U) - \delta) + \deg(g) < (\mathrm{reg}(U) - \delta) + \delta = \mathrm{reg}(U)$ 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 $\mathrm{SM}(\mathbb{F}_2^n)$ (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

$$\begin{aligned}
\mathrm{reg}(Y) &= n - a(\overline{Y}) \\
&\geq n - \deg(t) \\
&\geq n - \deg(r_U) - \deg(q) \\
&= \mathrm{reg}(U) - \deg(q) \\
&> \mathrm{reg}(U) - \delta \\
&\geq \mathrm{reg}(U_1).
\end{aligned}$$

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

□

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,*

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

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

We will now show $\mathrm{SM}(U_0 \cap G, j), \mathrm{SM}(U_1 \cap G, j) \supseteq \mathrm{SM}(U \cap G, j) \forall j \leq \delta - d$, 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 $\mathrm{LM}(U_0 \cap G, j), \mathrm{LM}(U_1 \cap G, j) \subseteq \mathrm{LM}(U \cap G, j)$.

To see that $\mathrm{LM}(U_0 \cap G, j) \subseteq \mathrm{LM}(U \cap G, j)$, assume, for contradiction, that this is not the case. Then $\exists x^\alpha \in \mathrm{LM}(U_0 \cap G, j) \cap \mathrm{SM}(U \cap G, j)$. Due to the fact that $x^\alpha \in \mathrm{LM}(U_0 \cap G, j)$ 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 $\mathrm{lm}(q) = x^\alpha$. Clearly,

$\deg(q) = j \leq \delta - d$. Due to the fact that $x^\alpha \in \mathrm{SM}(U \cap G, j)$, 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

$$\begin{aligned} a(\overline{Y}) &\leq \deg(t) \\ &\leq \deg(r_U) + \deg(q) + \deg(g_i + 1) \\ &< a(\overline{U}) + (\delta - d) + d \\ &= a(\overline{U}) + \delta. \end{aligned}$$

Applying Lemma 2, we then have

$$\begin{aligned} \mathrm{reg}(Y) &= n - a(\overline{Y}) \\ &> n - (a(\overline{U}) + \delta) \\ &= (n - a(\overline{U})) - \delta \\ &= \mathrm{reg}(U) - \delta \\ &\geq \mathrm{reg}(U_1), \end{aligned}$$

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

□

4 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. Let $r = \mathrm{reg}(V)$ and set t to be the unique value $r - k + 1 \leq t \leq n$ such that $\binom{t}{r-k+1} \leq |\mathrm{SM}(V, r - k + 1)| \leq \binom{t+1}{r-k+1}$.

First, notice that $|\mathrm{SM}(V, i)| \geq \binom{t}{i}$, $\forall i \leq r - k + 1$. This follows by a straightforward induction on $j = r - k + 1 - i$. The case in which $j = 0$ follows from the above definition of t. If $|\mathrm{SM}(V, r - k + 1 - j)| \geq \binom{t}{r-k+1-j}$, then we immediately have a set $S \subseteq \mathrm{SM}(V, r - k + 1 - j)$ such

that $|S| = \binom{t}{r-k+1-j}$. Define the set ΔS to consist of all monomials that lie immediately below some monomial in S in the monomial order (this is frequently called the *shadow* of S),

$$\Delta(S) = \{x^\alpha : \deg(x^\alpha) = r-k+1-(j+1) \text{ and } \exists x^\gamma \in S \text{ such that } x^\alpha < x^\gamma\}.$$

Due to the fact that $\mathrm{SM}(V)$ is a dual ideal, we note that $S \subseteq \mathrm{SM}(V) \Rightarrow \Delta(S) \subseteq \mathrm{SM}(V)$, from which we immediately conclude $\Delta(S) \subseteq \mathrm{SM}(V, r-k+1-(j+1))$. We then have

$$|\mathrm{SM}(V, r-k+1-(j+1))| \geq |\Delta(S)| \geq \binom{t}{r-k+1-(j+1)},$$

where the last inequality follows immediately from Lovász's version [Lov79] of the Kruskal-Katona theorem.

By a precisely analogous argument, we also have $|\mathrm{SM}(V,i)| \leq \binom{t+1}{i}\ \forall i \geq r-k+1$. By Lemma 1(a) and the above,

$$h^a(V,r) \geq h^a(V, r-k) = \sum_{i=0}^{r-k} |\mathrm{SM}(V,i)| \geq \sum_{i=0}^{r-k} \binom{t}{i} \geq c_1 2^t,$$

for some constant $c_1 > 0$ (where the last inequality follows from the fact that $r-k > \frac{n}{2} - 2\sqrt{n} \geq \frac{t}{2} - 2\sqrt{t}$ combined with elementary bounds on the sum of binomial coefficients). Similarly, $\forall i \geq r-k+1$, we have, for some constant $c_2 > 0$,

$$|\mathrm{SM}(V,i)| \leq \binom{h+1}{i} \leq \binom{h+1}{\lceil \frac{h+1}{2} \rceil} \leq \frac{c_2 2^h}{\sqrt{h}}.$$

We then have, for some constant $c > 0$,

$$\begin{aligned}
\frac{h^a(V,r) - h^a(V,r-k)}{|V|} &= \frac{h^a(V,r) - h^a(V,r-k)}{h^a(V,r)} \\
&= \frac{\sum_{i=r-k}^{r} |\mathrm{SM}(V,i)|}{h^a(V,r)} \\
&\leq \frac{(k+1)(c_2)\frac{2^t}{\sqrt{t}}}{c_1 2^t} \\
&= \frac{(k+1)\frac{c_2}{c_1}}{\sqrt{t}} \\
&\leq \frac{ck}{\sqrt{n}}.
\end{aligned}$$

□

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 $\mathrm{reg}(U_0), \mathrm{reg}(U_1) \leq \frac{n}{2} + n^\gamma$. We also have $\mathrm{reg}(U_0), \mathrm{reg}(U_1) \geq \frac{n}{2} - n^\gamma$ because $2^n = |U_0| + |U_1| = |\mathrm{SM}(U_0)| + |\mathrm{SM}(U_1)|$, 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 [Raz87],[Smo87], 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

$$\mathrm{SM}(U_0 \cap V(y), i) = \mathrm{SM}(U_1 \cap V(y), i) = \mathrm{SM}(V(y)) \; \forall i \leq \frac{n}{2} - n^\gamma - dl.$$

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 $\mathrm{reg}(U_0 \cap V(y)) \leq \mathrm{reg}(U_0) \leq \frac{n}{2} + n^\gamma$, and $\mathrm{reg}(U_1 \cap V(y)) \leq \mathrm{reg}(U_1) \leq \frac{n}{2} + n^\gamma$. Moreover, $\mathrm{reg}(U_0 \cap V(y)), \mathrm{reg}(U_1 \cap V(y)) \geq \frac{n}{2} - n^\gamma - dl$. To see this, first notice that Lemma 2 allows us to conclude $\mathrm{reg}(V(y)) \geq n - dl$ (because $y+1$ vanishes on the complement of $V(y)$), which immediately implies that $\mathrm{SM}(V(y))$ consists of an element x^κ of degree at least $n - dl$. As $\mathrm{SM}(V(y))$ 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 $\mathrm{SM}(V(y))$, $\mathrm{SM}(U_0 \cap V(y))$ and $\mathrm{SM}(U_1 \cap V(y))$, we then conclude that both $\mathrm{SM}(U_0 \cap V(y))$ and $\mathrm{SM}(U_1 \cap V(y))$ contain an element of degree $\frac{n}{2} - n^\gamma - dl$, and so, by Lemma 1(b), the claimed lower bound on regularity follows. In the following, for brevity, we write $H_i(j) = h^a(U_i \cap V(y), j)$, $d_1 = \frac{n}{2} + n^\gamma$,$d_2 = \frac{n}{2} - n^\gamma - dl$.

We then have

$$\begin{aligned}
\mathrm{bias}(f|_{V(g_1,\ldots,g_k)}) &= |\mathbb{E}_{x \sim V(g_1,\ldots,g_k)}[(-1)^{f(x)}]| \\
&= \frac{||U_0 \cap V(g_1, \ldots, g_k)| - |U_0 \cap V(g_1, \ldots, g_k)||}{|V(g_1, \ldots, g_k)|} \\
&\leq \frac{||U_0 \cap V(y)| - |U_1 \cap V(y)|| + |V(y) \setminus V(g_1, \ldots, g_k)|}{|V(g_1, \ldots, g_k)|} \\
&\leq \frac{|H_0(d_1) - H_1(d_1)| + 2^{n-l}}{|V(g_1, \ldots, g_k)|} \\
&= \frac{|H_0(d_2) - H_1(d_2) + (H_0(d_1) - H_0(d_2)) - (H_1(d_1) - H_1(d_2)) + 2^{n-l}}{|V(g_1, \ldots, g_k)|} \\
&= \frac{|(H_0(d_1) - H_0(d_2) - (H_1(d_1) - H_1(d_2)) + 2^{n-l}}{|V(g_1, \ldots, g_k)|} \\
&\leq \frac{\frac{c'(2n^\gamma + dl)}{\sqrt{n}}|V(g_1, \ldots, g_k)| + 2^{n-l}}{|V(g_1, \ldots, g_k)|} \\
&= \frac{c'(2n^\gamma + dl)}{\sqrt{n}} + \frac{2^{n-l}}{|V(g_1, \ldots, g_k)|} \\
&\leq \frac{c'(2n^\gamma + dl)}{\sqrt{n}} + \frac{2^{n-l}}{\rho 2^n}
\end{aligned}$$

$$= \frac{c^{'}(2n^{\gamma} + dl)}{\sqrt{n}} + \frac{1}{\rho 2^{l}}.$$

Setting $l = \log(\frac{\sqrt{n}}{\rho})$ yields the claimed bound.

□

Next, as in [CT13], we consider a variant of the extractor model in which, rather than explicitly considering algebraic sets which satisfy a certain density bound, we consider algebraic sets defined by a limited number of polynomials. The following is immediate.

Corollary 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 $0 < \alpha \leq \beta < \frac{1}{2}$, and for any $d \leq n^{\alpha}$ and $k \leq n^{\beta-\alpha}$, f is an extractor with bias $\frac{c(n^{\gamma}+d(n^{\beta}+\frac{1}{2}\log(n))}{\sqrt{n}}$ for algebraic sets that are the common zeros of a collection of at most k polynomials each of degree at most d.*

Proof. Consider any algebraic set $V = V(g_1, \ldots, g_k)$ where $g_i \in \mathbb{F}_2[x_1, \ldots, x_n]$ and $\deg(g_i) \leq d \ \forall i$. Let $g = 1 + \prod_{i=1}^{k} g_i$. Then $\deg(g) \leq kd \leq n^{\beta}$ and $V = V(g)$. From Lemma 2 it immediately follows that $\mathrm{reg}(V(g)) \geq n - \deg(g) \geq n - n^{\beta}$, and so, by definition $\exists x^{\kappa} \in \mathrm{SM}(V(g))$ such that $\deg(x^{\kappa}) = n - n^{\beta}$. Due to the fact that $\mathrm{SM}(V(g))$ is a dual ideal, every divisor of x^{κ} is also a member of $\mathrm{SM}(V(g))$. As there are precisely $2^{n-n^{\beta}}$ such divisors we have

$$|V| = |V(g)| = |\mathrm{SM}(V(g))| \geq 2^{n-n^{\beta}},$$

and so V has density $\rho \geq 2^{-n^{\beta}}$. The result then follows immediately from Theorem 1.

□

5 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, [BV93], [BV97], [Aar03], [Aar10], [Joh08]). The version considered in this paper, and the notation used, follows most closely [Joh08], but essentially the same claims hold for all other standard variants. We begin by precisely defining the problem.

5.1 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). Then

$$FS_n(f,g) = \begin{cases} g(s), & \text{if } \exists s \in \{0,1\}^n \text{ such that } f(x) = x \cdot s \ \forall x \\ *, & \text{otherwise} \end{cases}$$

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

$$FS_n^g(f) = \begin{cases} g(s), & \text{if } \exists s \in \{0,1\}^n \text{ such that } f(x) = x \cdot s \ \forall x \\ *, & \text{otherwise} \end{cases}$$

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 funciton $g : \{0,1\}^n \to \{0,1\}$ given as a 2^n bit long truth table. We then define

$$RFS_{n,h}(R_0, \ldots, R_{2^n-1}, g) = \begin{cases} g(s), & \text{if } \exists s \in \{0,1\}^n \text{ such that } \forall \sigma \in \{0,1\}^n \ RFS_{n,h-1}(R_\sigma) = \sigma \cdot s \\ *, & \text{otherwise} \end{cases}$$

In a precisely analogous fashion, we define $RFS_{n,h}^g$ 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 $RFS_{n,h}^g$ 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 $RFS_{n,h}^g$ 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.

5.2 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. Let $r_{U_{p,h}} \in \mathbb{F}_2[x_1, \ldots, x_m]$ denote the unique squarefree polynomial such that $r_{U_{p,h}}(x) = 1$ if and only if $x \in U_{p,h}$. By a straightforward counting of the number of promises of each degree, we have $\deg(r_{U_{p,h}}) \leq (2^n - n)\sum_{j=1}^{h-1} 2^{(j-1)n} d^{h-j}$. By construction $r_{U_{p,h}}$ vanishes on $\overline{U_{p,h}}$ and so

$$\begin{aligned}
a(\overline{U_{p,h}}) &\leq \deg(r_{U_{p,h}}) \\
&\leq (2^n - n)\sum_{j=1}^{h-1} 2^{(j-1)n} d^{h-j} \\
&= 2^n \left(\sum_{j=1}^{h-1} 2^{(j-1)n} d^{h-j}\right) - n\left(\sum_{j=1}^{h-1} 2^{(j-1)n} d^{h-j}\right) \\
&= \left(\sum_{j=1}^{h-1} 2^{jn} d^{h-j}\right) - n\left(\sum_{j=1}^{h-1} 2^{(j-1)n} d^{h-j}\right) \\
&= d\left(\sum_{j=2}^{h} 2^{(j-1)n} d^{h-j}\right) - n\left(\sum_{j=1}^{h-1} 2^{(j-1)n} d^{h-j}\right)
\end{aligned}$$

$$= d2^{(h-1)n} - nd^{h-1} - (n-d)\sum_{j=2}^{h-1} 2^{(j-1)n} d^{h-j}.$$

Applying Lemma 2, we then have

$$\mathrm{reg}(U_{p,h}) = n2^{(h-1)n} - a(\overline{U_{p,h}})$$

$$\geq (n-d)2^{(h-1)n} + nd^{h-1} + (n-d)\sum_{j=2}^{h-1} 2^{(j-1)n} d^{h-j}$$

$$= nd^{h-1} + (n-d)\sum_{j=2}^{h} 2^{(j-1)n} d^{h-j}$$

$$= nd^{h-1} + (n-d)\sum_{j=1}^{h-1} 2^{jn} d^{h-j-1}.$$

Similarly, define $r_{U_{0,h}}, r_{U_{1,h}} \in \mathbb{F}_2[x_1, \ldots, x_m]$ as the unique squarefree polynomials such that $r_{U_{0,h}}(x) = 1$ if and only if $x \in U_{0,h}$ and $r_{U_{1,h}}(x) = 1$ if and only if $x \in U_{1,h}$. We then immediately have $\deg(r_{U_{0,h}}), \deg(r_{U_{1,h}}) \leq \deg(r_{U_{p,h}}) + d^i$, and so, by a precisely analogous argument as above

$$\mathrm{reg}(U_{0,h}), \mathrm{reg}(U_{1,h}) \geq (n-d)d^{h-1} + (n-d)\sum_{j=1}^{h-1} 2^{jn} d^{h-j-1}$$

$$= (n-d)\sum_{j=0}^{h-1} 2^{jn} d^{h-j-1}.$$

□

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 $\mathrm{MAJ} : \mathbb{F}_2^n \to \mathbb{F}_2$ be defined such that $\mathrm{MAJ}(x) = 1$ 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 $e_i(x) = \sum_{J \subseteq [n], |J|=i} \sum_{j \in J} x_j$ 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. Begin by noticing that

$$e_i(x) = \binom{wt(x)}{i} \mod 2.$$

By a straightforward application of Kummer's lemma, we then conclude

$$e_i(x) = \begin{cases} 1, & wt(x) \geq_b i \\ 0, & \text{otherwise} \end{cases}.$$

Next, define functions $E_i : \mathbb{F}_2^n \to \mathbb{F}_2$ and $G_i : \mathbb{F}_2^n \to \mathbb{F}_2$ such that $E_i(x) = 1$ if and only if $wt(x) = i$ and $G_i(x) = 1$ if and only if $wt(x) \geq i$. We then have

$$E_i(x) = \sum_{j \geq_b i} e_j(x).$$

To see this, simply notice that if $E_i(x) = 1$ then $wt(x) = i$ and so $e_i(x) = 1$, but $e_j(x) = 0$ for all other terms in the above sum. If $E_i(x) = 0$, then $wt(x) = t \neq i$. There are then two cases: if $i \leq_b t$, then the only terms in the above sum that evaluate to one are precisely all values j such that $i \leq_b j \leq t$, of which there are an even number; if $i \not\leq_b t$, then $\forall j$ such that $j \geq_b i$, $j \not\leq_b t$, and so every term in the above sum evaluates to zero.

We then have

$$\begin{aligned} G_i(x) &= \sum_{l \geq i} E_l(x) \\ &= \sum_{l \geq i} \sum_{j \geq_b l} e_j(x), \end{aligned}$$

and so

$$\begin{aligned} \mathrm{MAJ}(X) &= G_{\frac{n}{2}}(x) \\ &= \sum_{l \geq \frac{n}{2}} \sum_{j \geq_b l} e_j(x). \end{aligned}$$

□

We now consider $RFS_{n,h}^{\mathrm{MAJ}}$. We begin by demonstrating a useful symmetry in $U_{p,h}^{\mathrm{MAJ}}$. Define the value $\hat{1}_h \in U_{p,h}^{\mathrm{MAJ}}$ 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 $\hat{1}_h \in U_{p,h}^{\mathrm{MAJ}}$ 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, $\hat{1}_h \in U_{1,h}^{\mathrm{MAJ}}$ as, by construction, the value of the root is 1. For any $x \in U_{p,h}^{\mathrm{MAJ}}$, 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, $x \in U_{0,h}^{MAJ}$ if and only if $\hat{x} \in U_{1,h}^{MAJ}$.*

Proof. Given any $x \in U_{0,h}^{\mathrm{MAJ}}$, 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 $\overline{b}$. 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 $\hat{x} \in U^{\mathrm{MAJ}}_{1,h}$. The reverse implication follows from the fact that $\hat{\hat{x}} = x$ and symmetry.

□

We now show that $U^{\mathrm{MAJ}}_{0,h}$ and $U^{\mathrm{MAJ}}_{1,h}$ 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 $\mathrm{LM}(U^{\mathrm{MAJ}}_{0,h}) = \mathrm{LM}(U^{\mathrm{MAJ}}_{1,h})$.

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

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

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

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

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

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

□

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

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*

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

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

Proof. We show this by induction on h. First, consider the case in which $h = 1$. Clearly, $U^{\mathrm{MAJ}}_{p,1} = \mathbb{F}_2^n$ and so $\mathrm{reg}(U^{\mathrm{MAJ}}_{p,1}) = n$. Moreover, $U^{\mathrm{MAJ}}_{0,1} = (\mathrm{MAJ})^{-1}(0)$ and $U^{\mathrm{MAJ}}_{1,1} = (\mathrm{MAJ})^{-1}(1)$, and so, by Lemma 5, we have $\mathrm{reg}(U_{0,1}) = \mathrm{reg}(U_{1,1}) = \frac{n-1}{2}$.

We now consider the case in which $h > 1$. First, consider $U^{\mathrm{MAJ}}_{p,h}$. By the definition of regularity, $\mathrm{reg}(U_{p,h})$ is the minimal value of d such that $h^a(U_{p,h}, d) = |U_{p,h}|$. Therefore, if, for some d, $h^a(U_{p,h}, d) = |U_{p,h}|$, then $\mathrm{reg}(U_{p,h}) \leq d$. In particular, let

$$d(h) = n\left(\frac{n+1}{2}\right)^{h-1} + \left(\frac{n-1}{2}\right)\sum_{j=1}^{h-1} 2^{jn}\left(\frac{n+1}{2}\right)^{h-j-1}.$$

Then, in order to show $\mathrm{reg}(U_{p,h}) \leq d(h)$, it suffices to show $h^a(U_{p,h}, d(h)) = |U_{p,h}|$.

To show this, as before, let $m = n2^{n(h-1)}$ denote the total size of the input to $RFS^{\text{MAJ}}_{n,h}$, and let

$$M_d = \mathcal{M}\left(U_{p,h}, \binom{[m]}{\leq d}\right)$$

denote the inclusion matrix in which the rows are indexed by elements of $U_{p,h}$ and the columns are indexed by all squarefree monomials of degree at most d. By Lemma 3, $h^a(U_{p,h}, d) = \text{rank}_{\mathbb{F}_2}(M_d)$, and so it suffices to show $\text{rank}_{\mathbb{F}_2}(M_{d(h)}) = |U_{p,h}|$. Observe that $|U_{p,h}|$ is precisely the number of rows of $M_{d(h)}$ (and is, of course, substantially smaller than the number of columns), and so this is equivalent to showing that the matrix $M_{d(h)}$ is full rank.

To see that $M_{d(h)}$ is full rank, assume, for contradiction, that it is not. By definition, this means that there exists some non-empty $T \subseteq U_{p,h}$ such that the sum of the rows of $M_{d(h))}$ indexed by T is 0 in every column. We now show that, for any $T \subseteq U_{p,h}$, $\exists \alpha$ such that the rows indexed by T have the sum 1 in the column indexed by the monomial x^α, which is, of course, a contradiction.

Let x_i denote the i^{th} input variable. Let $E \subseteq [m]$ denote the indices of all variables that are inputs to the elementary leaves of the recursive Fourier sampling tree. Clearly, $|E| = n^h$ as there are n^{h-1} elementary leaves, each of which have n input variables. Define $\sigma : U_{p,h} \to \{0,1\}^{n^h}$ such that, for any $x \in U_{p,h}$, $\sigma(x)$ is the portion of x at indices E. We refer to this value as the *signature* of x. Consider a partial ordering on the set of signatures given by the usual bitwise ordering. That is to say, for any $y, z \in \{0,1\}^{n^h}$, let y_i and z_i denote the i^{th} bits of y and z, respectively. Define $y \leq z$ if $y_i \leq z_i \ \forall i$. Similarly, define $y < z$ if $y \leq z$ and $y \neq z$. Let $S_T = \{\sigma(x) : x \in T\}$ and M_T denote an (arbitrary) maximal element of S_T with respect to the partial order on signatures. That is to say, M_T is any single value that satisfies $M_T \in S_T$ and $\nexists y \in S_T$ such that $M_T < y$.

Recall that each column of $M_{d(h)}$ is indexed by a squarefree monomial $x^\alpha = x_1^{\alpha_1} \cdots x_m^{\alpha_m}$. Consider any column of $M_{d(h)}$ that is indexed by some x^α such that α agrees with M_T (that is to say, for each $i \in E$, α_i is equal to the corresponding value of M_T). The key observation is that the only rows $x \in T$ that could possibly have value 1 in column x^α are those such that $\sigma(x) = M_T$. To see this, notice that in order for a particular row $x \in T$ to have entry 1 in column x^α, it must be the case that $x_i = 1$ at every $i \in E$ such that $\alpha_i = 1$, and so, by definition, $\sigma(x) \geq M_T$. If $\sigma(x) \neq M_T$, then $\sigma(X) > M_T$, which contradicts the definition of M_T, and so we must have $\sigma(X) = M_T$, as claimed.

Let $Z \subseteq T$ be defined such that $Z = \{x \in T : \sigma(x) = M_T\}$. Then, for any column indexed by an x^α such that α agrees with M_T, the sum over all $x \in T$ and the sum over only those $x \in Z$ must be equal. Therefore, it suffices to exhibit a column indexed by x^α such that α agrees with M_T and the sum over all rows $x \in Z$ in column x^α is 1.

To do this, notice that the set Z is an algebraic set (as it is simply a set of elements in $\mathbb{F}_2^m$) where every $x \in Z$ lies within a particular subspace, namely the subspace consisting of the set of x that satisfy $\sigma(x) = M_T$. We now consider $\widetilde{Z}$, which is the induced algebraic set living within that subspace. More formally, we partition the collection of variables into two pieces: E and $[m] \setminus E$. For any $x \in \mathbb{F}_2^n$, let x_E and $x_{[m]\setminus E}$ denote the portions of x indexed by E and $[m] \setminus E$ respectively. We define the algebraic set $\widetilde{Z} \subseteq \mathbb{F}_2^{m-n^h}$ where $\widetilde{Z} = \{x_{[m]\setminus E} : x \in Z\}$.

We now consider the inclusion matrix

$$\widetilde{M} = \mathcal{M}\left(\widetilde{Z}, \binom{[m-n^h]}{d(h)-n^h}\right),$$

where the rows are indexed by the $x_{[m]\setminus E} \in \widetilde{Z}$ and the columns are indexed by the monomials $x^\beta_{[m]\setminus E}$. The next key observation is that, in order to prove the existence of a column x^α of the desired form, it suffices to show $\text{rank}_{\mathbb{F}_2}\widetilde{M} = |\widetilde{Z}|$, in other words, that the matrix $\widetilde{M}$ is full rank. To

see this, notice that if $\widetilde{M}$ is full rank then, by definition, for every non-empty set of rows $R \subseteq \widetilde{Z}$, there is some column $x^{\beta}_{[m]\setminus E}$ such that the sum in that column over the rows R is equal to 1. In particular, there is some column $x^{\beta}_{[m]\setminus E}$ such that the sum of every row in the column $x^{\beta}_{[m]\setminus E}$ is equal to 1. Fix any such β, and define α such that $\alpha_E = M_T$ and $\alpha_{[m]\setminus E} = \beta$. By construction, the sum of the entries of $M_{d(h)}$ in column x^{α} and rows Z is equal to the sum of the entries of $\widetilde{M}$ in column $x^{\beta}_{[m]\setminus E}$ and rows $\widetilde{Z}$. Therefore, the sum of the entries of $M_{d(h)}$ in column x^{α} and rows Z is 1, as desired.

All that remains is to show $\text{rank}_{\mathbb{F}_2}\widetilde{M} = |\widetilde{Z}|$. By Lemma 3, this is equivalent to showing $\text{reg}(\widetilde{Z}) \leq d(h) - n^h$. Before providing the details of this regularity bound, we briefly state the main idea which is that $\widetilde{Z} \subseteq V_1 \times \cdots \times V_w$ where each V_i is (isomorphic to) either U^{MAJ}_{0,h_i} or U^{MAJ}_{1,h_i} where each $h_i < h$. The induction hypothesis bounds the regularity of each such V_i, which in turn provides the required bound on the regularity of $\widetilde{Z}$, because, by the definition of regularity, $\widetilde{Z} \subseteq V_1 \times \cdots \times V_w$ immediately implies

$$\text{reg}(\widetilde{Z}) \leq \text{reg}(V_1 \times \cdots \times V_w) = \sum_i \text{reg}(V_i).$$

We now show the required bound on $\text{reg}(\widetilde{Z})$. By construction $\widetilde{Z}$ is the algebraic set consisting of the elements of T which reside in the subspace defined by $\sigma(x) = M_T$. We now consider how the constraint $\sigma(x) = M_T$ interacts with the linearity promise of recursive Fourier sampling. Consider the recursive Fourier sampling tree. The key observation is that the constraint $\sigma(x) = M_T$ fixes the value of all of the elementary children, which in turn fixes the value of every "sibling" of an elementary child. This, essentially, "decouples" the problem into the cartesian product of several independent, smaller instances of the recursive Fourier sampling problem.

To be precise, begin by noting that requiring $\sigma(x) = M_T$ directly forces the value (that is to say, the output) of each of the elementary *leaves* of the recursive Fourier sampling tree. By simply propagating this constraint upward through the tree, the value of all of the elementary *children* is also forced. To see this, simply notice that, by construction, if the value of all elementary children of a particular node t is forced, then the value of t itself is forced. Since each elementary child which is not an elementary leaf has its own collection of elementary children, the result immediately follows.

We therefore conclude that the constraint $\sigma(x) = M_T$ forces the value of all n^{h-1} elementary children. Of course, this is only a tiny portion of the $\theta(n2^{nh})$ nodes of the recursive Fourier sampling tree. However, the linearity constraint imposed by the promise within recursive Fourier sampling causes the constraint $\sigma(x) = M_T$ to constrain other portions of the recursive Fourier sampling tree. In particular, begin by considering the root of the recursive Fourier sampling tree. As noted above, the constraint $\sigma(x) = M_T$ directly forces the value of each of the n elementary children of the root. Moreover, due to the linearity constraint, the value of the other $2^n - n$ children of the root are also forced. In particular, if we let t denote the root of the tree, t_i denote its i^{th} child, $b(t_i)$ denote the value of node t_i, i_j denote the j^{th} bit of i, and χ_j denote the element of $\{0,1\}^n$ which has value 1 in position j and value 0 elsewhere, then

$$b(t_i) = \sum_{j:i_j=1} b(t_{\chi_j}).$$

Therefore, for any x that satisfies $\sigma(x) = M_T$, if we consider the portion of x that lies under the subtree rooted at t_i, for any t_i which is *not* an elementary child of the root node t, then this portion of x must lie within an algebraic set isomorphic to $U^{\text{MAJ}}_{b(t_i),h-1}$.

Precisely the same logic applies if we consider any elementary child that is not an elementary leaf. At l levels down from the root of the tree, there are n^l elementary children, each of which have

n elementary children and $2^n - n$ non-elementary children. For any x that satisfies $\sigma(x) = M_T$, the portion of x that lies under the subtree rooted at each of the non-elementary child t must lie within an algebraic set isomorphic to $U^{\mathrm{MAJ}}_{b(t),h-l-1}$.

Next, notice that this process completely partitions the input of the recursive Fourier sampling tree into a piece that lies beneath the elementary leaves and many other pieces which each lie beneath the subtree rooted at a non-elementary sibling of some elementary child. To see this, consider any particular input variable x_i and consider its highest ancestor (other than the root of the tree) which is not an elementary child. If such an ancestor does not exist, then this variable is an input to an elementary leaf. If such an ancestor does exist, then it must be a non-elementary child of an elementary child (or of the root of the tree), and so this ancestor will have elementary children of its parent as siblings. We therefore conclude that $\widetilde{Z} \subseteq V_1 \times \cdots \times V_w$ where each V_i is (isomorphic to) either U^{MAJ}_{0,h_i} or U^{MAJ}_{1,h_i} where each $h_i < h$, as claimed. Counting the number of copies of each $U^{\mathrm{MAJ}}_{0,j}$ and $U^{\mathrm{MAJ}}_{1,j}$, using Lemma 13 to conclude that $\mathrm{reg}(U^{\mathrm{MAJ}}_{0,j}) = \mathrm{reg}(U^{\mathrm{MAJ}}_{1,j})$, and applying the induction hypothesis to bound the regularity of $U^{\mathrm{MAJ}}_{0,j}$ and $U^{\mathrm{MAJ}}_{1,j}$ yields the following.

$$\mathrm{reg}(\widetilde{Z}) \leq \sum_{i=1}^{h-1} n^{i-1}(2^n - n)\mathrm{reg}(U^{\mathrm{MAJ}}_{0,h-i})$$

$$\leq \sum_{i=1}^{h-1} n^{i-1}(2^n - n)\left[\left(\frac{n-1}{2}\right)\sum_{j=0}^{h-i-1} 2^{jn}\left(\frac{n+1}{2}\right)^{h-i-j-1}\right]$$

$$= \left(\frac{n-1}{2}\right)\left[\left(\sum_{i=1}^{h-1} n^{i-1}\sum_{j=0}^{h-i-1} 2^{(j+1)n}\left(\frac{n+1}{2}\right)^{h-i-j-1}\right) - \left(\sum_{i=1}^{h-1} n^{i}\sum_{j=0}^{h-i-1} 2^{jn}\left(\frac{n+1}{2}\right)^{h-i-j-1}\right)\right]$$

$$= \left(\frac{n-1}{2}\right)\left[\left(\sum_{i=0}^{h-2} n^{i}\sum_{j=0}^{h-i-2} 2^{(j+1)n}\left(\frac{n+1}{2}\right)^{h-i-j-2}\right) - \left(\sum_{i=1}^{h-1} n^{i}\sum_{j=0}^{h-i-1} 2^{jn}\left(\frac{n+1}{2}\right)^{h-i-j-1}\right)\right]$$

$$= \left(\frac{n-1}{2}\right)\left[\left(\sum_{i=0}^{h-2} n^{i}\sum_{j=1}^{h-i-1} 2^{jn}\left(\frac{n+1}{2}\right)^{h-i-j-1}\right) - \left(\sum_{i=1}^{h-1} n^{i}\sum_{j=0}^{h-i-1} 2^{jn}\left(\frac{n+1}{2}\right)^{h-i-j-1}\right)\right]$$

$$= \left(\frac{n-1}{2}\right)\left[\left(\sum_{j=1}^{h-1} 2^{jn}\left(\frac{n+1}{2}\right)^{h-j-1}\right) - n^{h-1} - \left(\sum_{i=1}^{h-2} n^{i}\left(\frac{n+1}{2}\right)^{h-i-1}\right)\right]$$

$$= \left(\frac{n-1}{2}\right)\left[\left(\sum_{j=1}^{h-1} 2^{jn}\left(\frac{n+1}{2}\right)^{h-j-1}\right) - \left(\sum_{i=1}^{h-1} n^{i}\left(\frac{n+1}{2}\right)^{h-i-1}\right)\right]$$

$$= \left(\frac{n-1}{2}\right)\left(\sum_{j=1}^{h-1} 2^{jn}\left(\frac{n+1}{2}\right)^{h-j-1}\right) - \left(\frac{n-1}{2}\right)\left(\sum_{i=1}^{h-1} n^{i}\left(\frac{n+1}{2}\right)^{h-i-1}\right)$$

$$= \left(\frac{n-1}{2}\right)\left(\sum_{j=1}^{h-1} 2^{jn}\left(\frac{n+1}{2}\right)^{h-j-1}\right) - \left(n - \left(\frac{n+1}{2}\right)\right)\left(\sum_{i=1}^{h-1} n^{i}\left(\frac{n+1}{2}\right)^{h-i-1}\right)$$

$$= \left(\frac{n-1}{2}\right)\left(\sum_{j=1}^{h-1} 2^{jn}\left(\frac{n+1}{2}\right)^{h-j-1}\right) - \left(\sum_{i=1}^{h-1} n^{i+1}\left(\frac{n+1}{2}\right)^{h-i-1}\right) + \left(\sum_{i=1}^{h-1} n^{i}\left(\frac{n+1}{2}\right)^{h-i}\right)$$

$$= \left(\frac{n-1}{2}\right)\left(\sum_{j=1}^{h-1} 2^{jn}\left(\frac{n+1}{2}\right)^{h-j-1}\right) - \left(\sum_{i=2}^{h} n^{i+1}\left(\frac{n+1}{2}\right)^{h-i-1}\right) + \left(\sum_{i=1}^{h-1} n^{i}\left(\frac{n+1}{2}\right)^{h-i}\right)$$

$$= \left(\frac{n-1}{2}\right)\left(\sum_{j=1}^{h-1} 2^{jn}\left(\frac{n+1}{2}\right)^{h-j-1}\right) + n\left(\frac{n+1}{2}\right)^{h-1} - n^h$$

$$= d(h) - n^h.$$

This immediately implies

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

Essentially the same argument applies to bound $\mathrm{reg}(U_{0,h}^{\mathrm{MAJ}})$. More precisely, again consider the case in which $h > 1$, we will show

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

$$= \left(\frac{n-1}{2}\right)\left(\frac{n+1}{2}\right)^{h-1} + \left(\frac{n-1}{2}\right)\sum_{j=1}^{h-1} 2^{jn}\left(\frac{n+1}{2}\right)^{h-j-1}$$

$$= \left(n - \left(\frac{n+1}{2}\right)\right)\left(\frac{n+1}{2}\right)^{h-1} + \left(\frac{n-1}{2}\right)\sum_{j=1}^{h-1} 2^{jn}\left(\frac{n+1}{2}\right)^{h-j-1}$$

$$= n\left(\frac{n+1}{2}\right)^{h-1} - \left(\frac{n+1}{2}\right)^{h} + \left(\frac{n-1}{2}\right)\sum_{j=1}^{h-1} 2^{jn}\left(\frac{n+1}{2}\right)^{h-j-1}$$

$$= d(h) - \left(\frac{n+1}{2}\right)^{h}.$$

We perform precisely the same analysis used to bound $\mathrm{reg}(U_{p,h}^{\mathrm{MAJ}})$, with the only change being the fact that when $M_T \in \{0,1\}^{n^h}$ is now constructed, we can now conlude that $wt(M_T) \leq n^h - \left(\frac{n+1}{2}\right)^h$, where $wt(M_T)$ denotes the number of 1s (the weight) of M_T. This follows because, for any $x \in T \subseteq U_{0,h}$ the value of the root node must be 0, by definition. For any node to have value 0, the majority of the elementary children of that node must have value 0 (because the function being evaluated at each node is MAJ). Due to the fact that each node has n elementary children, this requires that any node with value 0 has at least $\frac{n+1}{2}$ (recall that, by assumption, n is odd) elementary children with value 0. In particular, the majority of the elementary children of the root node must have value 0. Moreover, for each elementary child of the root node that has value 0, the majority of its children must have value 0. Continuing in this fashion until we reach the elementary leaves, we conclude that at least $\left(\frac{n+1}{2}\right)^h$ variables that are inputs to the elementary leaves must have value 0, and so at most $n^h - \left(\frac{n+1}{2}\right)^h$ have value 1, which shows the claimed bound on $wt(M_T)$. Therefore, when we construct α by $\alpha_E = M_T$ and $\alpha_{[m]\setminus E} = \beta$, we now have

$$\mathrm{reg}(U_{0,h}^{\mathrm{MAJ}}) \leq wt(\alpha)$$

$$= wt(\alpha_E) + wt(\alpha_{[m]\setminus E})$$
$$= wt(M_T) + wt(\beta)$$
$$\leq \left(n^h - \left(\frac{n+1}{2}\right)^h\right) + \left(d(h) - n^h\right)$$
$$= d(h) - \left(\frac{n+1}{2}\right)^h.$$

Finally, to bound $\mathrm{reg}(U_{1,h}^{\mathrm{MAJ}})$, simply notice that by Lemma 13

$$\mathrm{SM}(U_{0,h}^{\mathrm{MAJ}}) = \mathrm{SM}(U_{1,h}^{\mathrm{MAJ}}) \ \forall h \geq 1.$$

Lemma 1(b) then immediately implies

$$\mathrm{reg}(U_{1,h}^{\mathrm{MAJ}}) = \mathrm{reg}(U_{0,h}^{\mathrm{MAJ}}) \ \forall h \geq 1.$$

□

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

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

Proof. By the assumed form of n, Lemma 11 immediately allows us to conclude

$$\mathrm{MAJ}(x) = e_{\frac{n+1}{2}}(x) \ \forall x \in \mathbb{F}_2^n.$$

Clearly, $\deg(e_{\frac{n+1}{2}}) = \frac{n+1}{2}$, and so Lemma 10 immediately implies

$$\mathrm{reg}(U_{p,h}) \geq n\left(\frac{n+1}{2}\right)^{h-1} + \left(\frac{n-1}{2}\right)\sum_{j=1}^{h-1} 2^{jn}\left(\frac{n+1}{2}\right)^{h-j-1}$$

$$\mathrm{reg}(U_{0,h}), \mathrm{reg}(U_{1,h}) \geq \left(\frac{n-1}{2}\right)\sum_{j=0}^{h-1} 2^{jn}\left(\frac{n+1}{2}\right)^{h-j-1}.$$

By Lemma 14,

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

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

Therefore,

$$\mathrm{reg}(U_{p,h}) = n\left(\frac{n+1}{2}\right)^{h-1} + \left(\frac{n-1}{2}\right)\sum_{j=1}^{h-1} 2^{jn}\left(\frac{n+1}{2}\right)^{h-j-1}$$

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

Finally,

$$\mathrm{reg}(U_{p,h}) - \mathrm{reg}(U_{0,h}) = \mathrm{reg}(U_{p,h}) - \mathrm{reg}(U_{0,h})$$

$$= n\left(\frac{n+1}{2}\right)^{h-1} + \left(\frac{n-1}{2}\right) \sum_{j=1}^{h-1} 2^{jn} \left(\frac{n+1}{2}\right)^{h-j-1} - \left(\frac{n-1}{2}\right) \sum_{j=0}^{h-1} 2^{jn} \left(\frac{n+1}{2}\right)^{h-j-1}$$

$$= n\left(\frac{n+1}{2}\right)^{h-1} - \left(\frac{n-1}{2}\right)\left(\frac{n+1}{2}\right)^{h-1}$$

$$= \left(n - \left(\frac{n-1}{2}\right)\right)\left(\frac{n+1}{2}\right)^{h-1}$$

$$= \left(\frac{n+1}{2}\right)^{h}.$$

Therefore, $RFS_{n,h}^{\mathrm{MAJ}}$ is $\left(\frac{n+1}{2}\right)^h$-versatile on $U_{p,h}^{\mathrm{MAJ}}$. To see that $U_{p,h}^{\mathrm{MAJ}}$ is a critical algebraic set, simply notice that, as shown in the proof of Lemma 10,

$$\deg(r_{U_{p,h}^{\mathrm{MAJ}}}) \leq (2^n - n) \sum_{j=1}^{h-1} 2^{(j-1)n} d^{h-j},$$

where $d = \frac{n+1}{2}$.

By the above,

$$\deg(r_{U_{p,h}^{\mathrm{MAJ}}}) \geq (2^n - n) \sum_{j=1}^{h-1} 2^{(j-1)n} d^{h-j} = a(\overline{U_{p,h}^{\mathrm{MAJ}}}),$$

and so

$$\deg(r_{U_{p,h}^{\mathrm{MAJ}}}) = a(\overline{U_{p,h}^{\mathrm{MAJ}}}),$$

which, by definition implies that $U_{p,h}^{\mathrm{MAJ}}$ is a critical algebraic set.

□

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, $\frac{V_0}{\ker \phi_A} \not\cong \mathbb{F}_2^{n-d+1}$ and $\frac{V_1}{\ker \phi_A} \not\cong \mathbb{F}_2^{n-d+1}$. 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 $RFS_{n,h}^g : \mathbb{F}_2^m \to \mathbb{F}_2$ denote the recursive Fourier sampling function with g. Then*

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

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

Proof. Before proceeding with the proof, we briefly remark that this Lemma could be proven by use of the inclusion matrix, in a similar manner to the proof of Lemma 14, shown above. We provide an different proof to illustrate an alternate method of bounding regularity.

We show this claim by induction on h. First, consider the case in which $h = 1$. Clearly, $U^g_{p,1} = \mathbb{F}_2^n$, and so $\text{reg}(U^g_{p,1}) = n$. We now show $\text{reg}(U^g_{0,1}), \text{reg}(U^g_{1,1}) \leq n-d$. First, consider $\text{reg}(U^g_{1,1})$. Begin by noticing that, by Lemma 1(b), this is equivalent to showing that

$$x^\alpha \in \text{LM}(U^g_{1,1}) \ \forall \alpha \text{ such that } \deg(x^\alpha) > n - d.$$

Due to the fact that, for any algebraic set V, $\text{LM}(V)$ is an ideal (of the semigroup of monomials) and because, for any $V \subseteq \mathbb{F}_2^n$, $x_j^2 \in \text{LM}(V)\ \forall j$, the above is equivalent to showing

$$x^\alpha \in \text{LM}(U^g_{1,1}) \ \forall \alpha \text{ such that } \deg(x^\alpha) = n - d + 1 \text{ and } x^\alpha \text{ is multilinear.}$$

To see this, consider any multilinear monomial x^α where $\deg(x^\alpha) = n-d+1$. Let $J = \{j : \alpha_j = 1\}$. For any $x \in \mathbb{F}_2^n$, let $x_J \in \mathbb{F}_2^{n-d+1}$ denote the substring at positions indexed by J. The key observation is that, because g is well-mixed, there is a $b \in \mathbb{F}_2^{n-d+1}$ such that for every $x \in U^g_{1,1}$, $x_J \neq b$. To see this, let $\phi_A : \mathbb{F}_2^n \to \mathbb{F}_2^{n-d+1}$ denote the unique linear map such that $\phi_A(x) = x_J\ \forall x \in \mathbb{F}_2^n$. Then, because g is well-mixed, we have, by definition, that

$$\frac{U^g_{1,1}}{\ker(\phi_A)} \ncong \mathbb{F}_2^{n-d+1}$$

and so $\exists b \in \mathbb{F}_2^{n-d+1}$ such that, for every $x \in U^g_{1,1}$, $\phi_A(x) \neq b$, as claimed. Fix any such b.

For $k \in [n-d+1]$, let J_k denote the k^{th} element of J (in the natural order), and consider the polynomial $f_\alpha \in \mathbb{F}_2[x_1, \ldots, x_n]$, where $f_\alpha = \prod_{k=1}^{n-d+1}(x_{J_k} + b_k + 1)$. We then have $f_\alpha \in I(U^g_{1,1})$. This holds because, for any $x \in U^g_{1,1}$, $x_J = \phi_A(x) \neq b$, and so $\exists k \in [n-d+1]$ such that $x_{J_k} \neq b_k$. For each k, we have $x_{J_k}, b_k \in \mathbb{F}_2$ and so if $x_{J_k} \neq b_k$, then $x_{J_k} = b_k + 1$. Therefore, for any $x \in U^g_{1,1}$, $\exists k \in [n-d+1]$ such that $x_{J_k} = b_k + 1$, and so f_α vanishes on $U^g_{1,1}$. Clearly, $\text{lm}(f_\alpha) = x^\alpha$, and so

$$x^\alpha \in \text{LM}(U^g_{1,1}) \ \forall \alpha \text{ such that } \deg(x^\alpha) = n - d + 1 \text{ and } x^\alpha \text{ is multilinear,}$$

as desired. Therefore, $\text{reg}(U^g_{1,1}) \leq n - d$. By a precisely symmetric argument, $\text{reg}(U^g_{0,1}) \leq n - d$.

Next, we consider the case in which $h > 1$. Consider $U^g_{1,h}$. Let $r(h) = (n-d)\sum_{j=0}^{h-1} 2^{jn} d^{h-j-1}$. We wish to show

$$\text{reg}(U^g_{1,h}) \leq r(h).$$

For the same reason as above, it is equivalent to show

$$x^\alpha \in \text{LM}(U^g_{1,h}) \ \forall \alpha \text{ such that } \deg(x^\alpha) = r(h) + 1 \text{ and } x^\alpha \text{ is multilinear.}$$

Consider any multilinear monomial x^α, where $\deg(x^\alpha) = r(h) + 1$. Let $J = \{j : \alpha_j = 1\}$. Consider the recursive Fourier sampling tree. For each child t of the root of the tree, say that t is *heavy* if at least $r(h-1) + 1$ of the variables in the subtree rooted at t appear in J (that is to say, there are at least $r(h-1)$ variables x_j, such that $j \in J$ and x_j is a variable that appears at one of the leaves of the subtree rooted at t). Moreover, say that t is *very heavy*, if at least $r(h-1) + d^{h-1} + 1$ of the variables in the subtree rooted at t appear in J. Due to the fact that $\deg(x^\alpha) = r(h) + 1$, it must be the case that at least one of the following two statements is true:

(1):At least one of the children of the root is very-heavy.

(2):At least $n - d + 1$ of the children of the root are heavy.

To see this, assume, for contradiction, that neither of these statements are true. Then at most $n-d$ of the children of the root are heavy, and none of the children of the root are very heavy. We then have

$$\begin{aligned}
\deg(x^\alpha) &\leq (n-d)(r(h-1)+d^{h-1})+(2^n-(n-d))(r(h-1)) \\
&= (n-d)d^{h-1}+2^n r(h-1) \\
&= (n-d)d^{h-1}+2^n(n-d)\sum_{j=0}^{(h-1)-1} 2^{jn}d^{(h-1)-j-1} \\
&= (n-d)\left(d^{h-1}+\sum_{j=0}^{h-2} 2^{(j+1)n}d^{h-(j+1)-1}\right) \\
&= (n-d)\left(d^{h-1}+\sum_{j=1}^{h-1} 2^{jn}d^{h-j-1}\right) \\
&= (n-d)\left(\sum_{j=0}^{h-1} 2^{jn}d^{h-j-1}\right) \\
&= r(h) \\
&< r(h)+1 \\
&= \deg(x^\alpha).
\end{aligned}$$

This contradiction immediately allows us to conclude that at least one of the above statements are true.

We now conclude that $x^\alpha \in \mathrm{LM}(U^g_{1,h})$. We first consider the case in which statement (1) holds. Let t denote an arbitrary very-heavy child of the root of the recursive Fourier sampling tree. Let x^β denote the multilinear monomial consisting of the product of all variables that are in the subtree rooted at t that appear in x^α. Clearly, $x^\beta | x^\alpha$, and so it suffices to show that $x^\beta \in \mathrm{LM}(U^g_{1,h})$. Due to the fact that t is very-heavy, we have,

$$\deg(x^\beta) \geq r(h-1)+d^{h-1}+1 \geq \mathrm{reg}(U_{p,h-1})+1,$$

where the first inequality follows from the definition of a very-heavy child, and the second inequality follows from the induction hypothesis. Let $\widetilde{x}$ denote the portion of the input x within the subtree rooted at t. The key observation is that, since the subtree rooted at t corresponds to an instance of the recursive Fourier sampling problem of height $h-1$, we must have $\widetilde{x} \in \widetilde{V} \cong U_{p,h-1}$ (where $\widetilde{V}$ is simply $U_{p,h-1}$ with variables renamed $\widetilde{x}$). Since $\deg(x^\beta) > \mathrm{reg}(U_{p,h-1}) = \mathrm{reg}(\widetilde{V})$, we have, by the definition of regularity, that $x^\beta \in \mathrm{LM}(\widetilde{V})$, and so $\exists f_\beta$ (which only contains variables in $\widetilde{x}$) such that $x^\beta = \mathrm{lm}(f_\beta)$ and $f_\beta \in I(\widetilde{V})$. Therefore, f_β vanishes on every $\widetilde{x} \in \widetilde{V}$, and so it must also vanish on every $x \in U^g_{1,h}$ because, by construction, if $x \in U^g_{1,h}$, then $\widetilde{x} \in \widetilde{V}$ and f_β only consists of variables in $\widetilde{x}$. Therefore, $x^\beta \in \mathrm{LM}(U^g_{1,h})$, which implies that $x^\alpha \in \mathrm{LM}(U^g_{1,h})$, as desired.

Next, we consider the case in which statement (2) holds. Let $\sigma : U^g_{1,h} \to \mathbb{F}_2^n$ be defined such that, for any $x \in U^g_{1,h}$, $\sigma(x)$ is defined such that the i^{th} position of $\sigma(x)$ is equal to the value of the i^{th} elementary child of the root when the input to the recursive Fourier sampling problem is x. Let

$t_1, \ldots, t_{n-d+1}$ denote an arbitrary collection of (distinct) heavy children of the root of the recursive Fourier sampling tree. Let $\widetilde{\sigma} : U^g_{1,h} \to \mathbb{F}_2^{n-d+1}$ be defined such that, for any $x \in U^g_{1,h}$, $\widetilde{\sigma}(x)$ is defined such that the i^{th} position of $\widetilde{\sigma}(x)$ is equal to the value of t_i. In other words, the function σ simply encodes the values of all elementary children and $\widetilde{\sigma}$ encodes the values of the heavy children of interest. The key observation is that, because g is well-mixed, there is a $b \in \mathbb{F}_2^{n-d+1}$, such that, for every $x \in U^g_{1,h}$, $\widetilde{\sigma}(x) \neq b$. To see this, notice that, by definition, if $x \in U^g_{1,h}$, then $\sigma(x) \in U^g_{1,1}$. Moreover, due to the linear structure of the promise, there is a linear map $\phi : \mathbb{F}_2^n \to \mathbb{F}_2^{n-d+1}$ such that

$$\phi(\sigma(x)) = \widetilde{\sigma}(x) \ \forall x \in U^g_{1,h}.$$

Due to the fact that g is well-mixed,

$$\frac{U^g_{1,1}}{\ker(\phi_A)} \ncong \mathbb{F}_2^{n-d+1},$$

and so the existence of b follows from an identical argument as in the $h = 1$ case above. Fix such a b.

For $i \in [n-d+1]$, let x^{β_i} denote the monomial consisting of all variables in the subtree rooted at t_i that appears in x^α. Let $x^\beta = \prod_i x^{\beta_i}$. Clearly $x^\beta | x^\alpha$ and so it suffices to show $x^\beta \in \text{LM}(U^g_{1,h})$. Notice that, for each i,

$$\deg(x^{\beta_i}) \geq r(h-1) + 1 \geq \text{reg}(U_{1,h-1}) + 1, \text{reg}(U_{0,h-1}) + 1,$$

where the first inequality follows from the fact that t_i is heavy, and the second inequality follows from the induction hypothesis. By the same argument that applied in case (1) above, we conclude that, for each i, there is a polynomial f_{β_i} such that $\text{lm}(f_{\beta_i}) = x^{\beta_i}$ and $f_{\beta_i} \in I(U^g_{b_i+1,h-1})$. To be clear, the bound on the degree of x^{β_i} implies that x^{β_i} is a leading monomial of both the algebraic set isomorphic to $U^g_{1,h-1}$ and the algebraic set isomorphic to $U^g_{0,h-1}$ (where the isomorphism is simply the trivial renaming of variables), we choose $f_{\beta_i} \in I(U^g_{b_i+1,h-1})$ specifically to make the next stage of the construction work.

We now consider the polynomial $f_\beta = \prod_i f_{\beta_i}$. Clearly, $\text{lm}(f_\beta) = x^\beta$. Moreover, we have $f_\beta \in I(U^g_{1,h})$. To see this, notice that, for every $x \in U^g_{1,h}$, $\widetilde{\sigma}(x) \neq b$, and so, for every $x \in U^g_{1,h}$, there must be at least one i such that $\widetilde{\sigma}(x)_i \neq b_i$. Since $\widetilde{\sigma}(x)_i, b_i \in \mathbb{F}_2$, if $\widetilde{\sigma}(x)_i \neq b_i$, then $\widetilde{\sigma}(x)_i = b_i + 1$. Therefore, for every $x \in U^g_{1,h}$, there must be at least one i such that f_{β_i} vanishes at x (because f_{β_i} vanishes whenever the portion of x in the subtree rooted at t_i has value $b_i + 1$ at node t_i). Due to the fact that f_β is the product of the f_{β_i}, if at least one of the f_{β_i} vanish, then f_β vanishes. This implies that $f_\beta \in I(U^g_{1,h})$, which in turn implies that $x^\beta \in \text{LM}(U^g_{1,h})$ which in turn implies that $x^\alpha \in \text{LM}(U^g_{1,h})$.

The above argument shows that $\text{reg}(U^g_{1,h}) \leq r(h)$ for any $h > 1$, given the induction hypothesis. It is easy to see that this argument is precisely symmetric with respect to $U^g_{1,h}$ and $U^g_{0,h}$ and so we immediately also conclude $\text{reg}(U^g_{1,h}) \leq r(h)$. An essentially identical argument shows $\text{reg}(U^g_{p,h}) \leq r(h) + d^h$, with the only changes being the fact that statement (2) now becomes "At Least $n + 1$ children of the root are heavy", and the analysis of the case in which statement (2) holds no longer relies on the fact that g is well-mixed, but instead the fact that, due to the linearity constraint, given any collection of $n + 1$ children of the root, there is at least one tuple of values that violates the promise.

□

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. Combining the bounds from Lemma 10 and Lemma 16, we have

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

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

Therefore,

$$\mathrm{reg}(U^g_{p,h}) - \mathrm{reg}(U^g_{0,h}) = \mathrm{reg}(U^g_{p,h}) - \mathrm{reg}(U^g_{0,h}) = nd^{h-1} + (n-d)\sum_{j=1}^{h-1} 2^{jn} d^{h-j-1} - (n-d)\sum_{j=0}^{h-1} 2^{jn} d^{h-j-1}$$

$$= nd^{h-1} + (n-d)d^{h-1}$$
$$= d^h.$$

To see that $U^g_{p,h}$ is a critical algebraic set, simply notice that, as shown in the proof of Lemma 10,

$$\deg(r_{U^g_{p,h}}) \leq (2^n - n)\sum_{j=1}^{h-1} 2^{(j-1)n} d^{h-j}.$$

By the above,

$$\deg(r_{U^g_{p,h}}) \geq (2^n - n)\sum_{j=1}^{h-1} 2^{(j-1)n} d^{h-j} = a(\overline{U^g_{p,h}}),$$

and so

$$\deg(r_{U^g_{p,h}}) = a(\overline{U^g_{p,h}}),$$

which, by definition implies that $U^g_{p,h}$ is a critical algebraic set.

□

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 $\mathrm{GIP}_{n,d} : \mathbb{F}_2^n \to \mathbb{F}_2$ be defined such that

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

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. We begin by showing that, for any positive integers d, n that satisfy the above requirements, the function $\mathrm{GIP}_{n,d} : \mathbb{F}_2^n \to \mathbb{F}_2$ is well-mixed. To do this, it clearly suffices to show that, for any $(n-d+1) \times n$ matrix A, $\exists t^0, t^1 \in \mathbb{F}_2^{n-d+1}$ such that, for $x \in \mathbb{F}_2^n$, $Ax = t^0 \Rightarrow \mathrm{GIP}_{n,d}(x) = 0$ and $Ax = t^1 \Rightarrow \mathrm{GIP}_{n,d}(x) = 1$. We begin by noting that it suffices to show this claim only for A of a certain very special form. Let $\phi_A : \mathbb{F}_2^n \to \mathbb{F}_2^{n-d+1}$ denote the linear map corresponding to multiplication by the matrix A. Begin by noting that this claim trivially holds when A is not full rank (simply set t^0 and t^1 to be any element not in the image of ϕ_A) and so it suffices to consider only the case in which A is full rank. Next, it suffices to only consider the case in which A is in reduced row echelon form, because, for any invertible $(n-d+1) \times (n-d+1)$ matrix L, $Ax = t$ if and only if $(LA)x = Lt$, and so if the claim holds for every A in reduced row echelon form, then it holds for every A. Divide the n input variables $x_1, \ldots, x_n$ into blocks of size d, where each block consists of the d variables that appear in a single term of the $\mathrm{GIP}_{n,d}$ polynomial. Due to the fact that $\mathrm{rank}(A) = n-d+1$ and that A is in reduced row echelon form, there are precisely $d-1$ columns of A that do not have a leading 1. It suffices to only consider the case in which each of these $d-1$ columns appear as one of the rightmost $d(d-1)$ columns of A, because, due to the symmetry of the generalized inner product function, the variables can be relabled such that these columns always correspond to variables that appear in the rightmost $d-1$ blocks, and hence rightmost $d(d-1)$ columns.

Therefore, in order to show that $\mathrm{GIP}_{n,d}$ is well-mixed, it suffices to show that, for any $(n-d+1) \times n$ matrix A, where $\mathrm{rank}(A) = n-d+1$, A is in reduced row echelon form, and the $d-1$ columns of A that do not contain a leading 1 appear within the rightmost $d(d-1)$ columns, $\exists t^0, t^1 \in \mathbb{F}_2^{n-d+1}$ such that, for $x \in \mathbb{F}_2^n$, $Ax = t^0 \Rightarrow \mathrm{GIP}_{n,d}(x) = 0$ and $Ax = t^1 \Rightarrow \mathrm{GIP}_{n,d}(x) = 1$.

Consider such a matrix A. We now construct t^0 and t^1 with the required properties. Let $y_1, \ldots, y_{d-1}$ denote the x_i that correspond to columns of A that do not have leading 1s, in the natural order. Let $r \leq d$ denote the value such that y_1 is in the r^{th} block from the right; that is to say, r is the minimal value such that all y_m are in the rightmost r blocks. For $i \in [n-d+1]$, and $j \in \{0,1\}$, let t^j_i denote the value of the i^{th} position of t^j.

Begin by noticing that there is a setting of $t^j_{n-dr}, \ldots, t^j_{n-d+1}$ such that, for any such t^j and any $x \in \mathbb{F}_2^n$, $Ax = t^j \Rightarrow x_{n-dr+1} \cdots x_{n-d(r-1)} + \ldots + x_{n-d+1} \cdots x_n = 0$. In other words, there is a way to set the last $dr-d+1$ values of t^j such that, for any x that satisfies $Ax = t^j$, it must be the case that the sum of the rightmost r terms of $\mathrm{GIP}_{n,d}$ is 0. To show this, we will construct the setting of the last $dr-d+1$ values of t^j in a collection of stages, where the values set in the l^{th} stage will force the l^{th} block (from the right) to evaluate to 0. Begin by considering the rightmost block of variables. Let k denote the number of y_m such that y_m correspond to columns in the rightmost block of variables; that is to say, $y_{d-k}, \ldots, y_{d-1}$ are the variables that correspond to the columns within the rightmost block that do not have leading 1s. There is a setting of the last $d-k$ values of t^j such that, for any x that satisfies $Ax = t^j$, we have $x_{n-d+1} \cdots x_n = 0$. To see this, notice that, due to the form of A, the only non-zero entries of A in the last $d-k$ rows are in the last d columns, which correspond precisely to the variables in the rightmost block. Therefore, the last $d-k$ values of Ax are completely determined by the last d values of x. In order to have $x_{n-d+1} \cdots x_n = 1$, it must be the case that $x_{n-d+1} = \ldots = x_n = 1$, and so there is only 1 setting of these rightmost d variables such that $x_{n-d+1} \cdots x_n = 1$. On the other hand, there are $2^{d-k} \geq 2^{d-(d-1)} \geq 2 > 1$ distinct choices of the last $d-k$ values of t^j, from which it immediately follows that there is at least some setting of the last $d-k$ values of t^j such that, for any x that satisfies $Ax = t^j$, we do not have $x_{n-d+1} \cdots x_n = 1$, which then implies $x_{n-d+1} \cdots x_n = 0$. Fix any such setting of the last $d-k$ values of t^j.

In general, in the l^{th} stage, for each l such that $1 < l \leq r$, we consider the l^{th} block of

variables (counting from the right). Within the first $l-1$ stages, we have set every t_i^j such that row i of matrix A has a leading 1 in a column corresponding to a variable in one of the rightmost $l-1$ blocks. This setting forces each of these $l-1$ blocks to evaluate to 0. We now force the l^{th} rightmost block to evaluate to 0 by appropriately setting all t_i^j such that row i of matrix A has a leading 1 in a column corresponding to a variable in block l. To be precise, let k denote the number of y_m that correspond to variables in block l, and let k' denote the number of y_m that appear in the rightmost $l-1$ blocks. Again, due to the form of A, the only non-zero entries in the $d-k$ rows in question are in the last dl columns, and so the corresponding $d-k$ values of Ax are completely determined by the last dl values of x. Again, there is only a single setting of the d values of x in block l such that block l evaluates to 1. Moreover, there are only $2^{k'}$ settings of the $d(l-1)$ values of x in the rightmost $l-1$ blocks which satisfy the constraint imposed by the t_i^j fixed in earlier stages. This follows from the fact that the $(d(l-1)-k')\times(d(l-1))$ submatrix of A corresponding to these constraints has rank $d(l-1)-k'$ and hence nullity k'. Therefore, there are precisely $2^{k'}$ distinct settings of the last dl values of x that both satisfy all earlier constraints and cause block l to evaluate to 1. Moreover, there are 2^{d-k} choices of the portion of t_j currently being set. Due to the fact that $k+k'\leq d-1$ (as there are only a total of $d-1$ variables y_m), we again conclude that there is a setting of the relevant portion of t^j such that, for any x that satisfies $Ax=t^j$, the l^{th} block evaluates to 0, as required.

The above argument shows that all of the rightmost r blocks can be forced to evaluate to 0, by an appropriate setting of a portion of t^j. Next, we show that, similarly, for any $l>r$, the l^{th} rightmost block can be forced to evaluate to 0 by an appropriate setting of another portion of t^j. To be precise, consider the l^{th} rightmost block of variables, for any $l>r$. Due to the fact that every column of A that does not have a leading 1 appears among the rightmost r blocks, we conclude that every column corresponding to the l^{th} block has a leading 1. Consider the submatrix of A consiting of the d for which the leading 1 of that row appears in one of the columns corresponding to block l. The only non-zero entries in this submatrix appear in two parts. First, in the columns corresponding to block l, the submatrix is simply the $d\times d$ identity matrix. Secondly, there are non-zero entries in certain columns indexed by the y_m. In other words, this submatrix expresses the constraint that the values of x in block l are some affine combination of the y_m. To be precise, let $z_1,\ldots,z_d$ denote the d values of x that appear in block l, and let v denote the d values of t^j that correspond to rows in the submatrix in question. Then there is a $d\times(d-1)$ matrix B such that $z=By+v$. Let $\phi_B:\mathbb{F}_2^{d-1}\to\mathbb{F}_2^d$ denote the linear map corresponding to multiplication by B. As before, the key observation is that there is only a single setting of z such that block l evaluates to 0; however, there are 2^d choices of v, and $|\text{Im}(\phi_B)|\leq 2^{d-1}$, from which it immediately follows that there is a choice of v such that, for any z that satisfies $z=By+v$, it must be the case that block l evaluates to 0.

Therefore, to produce t^0, we simply use the first construction above to set the last $dr-d+1$ values of t^0 in such a way as to force the last r blocks to evaluate to 0, and then use the second construction above to set the remaining values of t^0 in such a way as to force all other blocks to evaluate to 0.

To produce t^1, slightly more work is needed. We next show that, given a collection of 2^{d^2} blocks, all of which are not among the rightmost r blocks, it is possible to set the appropriate values of t^1 is such a way as to assure that exactly one of these blocks evaluates to 1, and all other blocks evaluate to 0. To see this, simply consider, as above, the constraint imposed by A on the variables in each block l. To be precise, let $z^l=(z_1^l,\ldots,z_d^l)$ denote the d values of x that appear in block d, $v^l=(v_1^l,\ldots,v_d^l)$ denote the d values of t^1 that correspond to the rows of A that have a leading one in a column corresponding to block l and B^l denote the $d\times(d-1)$ matrix such that $z^l=B^ly+v^l$. As there are 2^{d^2} blocks in question, but only $2^{d(d-1)}$ distinct $d\times(d-1)$ matrices

(with entries in $\mathbb{F}_2$), there must be some particular $d \times (d-1)$ matrix B such that at least 2^d blocks l have $B^l = B$. Fix any such B and let L denote a collection of precisely 2^d blocks l such that $B^l = B$. The key observation is that the portion of t^1 corresponding to the collection of blocks L can be set in such a way so that exactly one block in L evaluates to 1. This can be accomplished by setting the collection of v^l such that $l \in L$ to the 2^d elements of $\mathbb{F}_2^d$. This works because, for any setting of y, the collection of z^l, for $l \in L$ will all be distinct (as each $z^l = By + v^l$ and the v^l are distinct) and exactly one of the z^l will be all 1s (as there are 2^d possible setting of each z^l, so each appears exactly once).

Therefore, to produce t^1, we then simply use the first construction above to set the last $dr - d + 1$ values of t^0 in such a way as to force the last r blocks to evaluate to 0, then the second construction above to set the portion of t^1 that corresponds to every block not in L to force all such blocks to evaluate to 0, and finally the third construction above to set the portion of t^1 that corresponds to the blocks in L to force exactly one block in L to evaluate to 1. Due to the fact that, by assumption, $n \geq d(2^{d^2} + d - 1)$, there are at least $2^{d^2} + d - 1$ blocks, and so this construction is possible.

We have thus shown that, for any positive integers d, n that satisfy the above requirements, the function $\mathrm{GIP}_{n,d} : \mathbb{F}_2^n \to \mathbb{F}_2$ is well-mixed. By Lemma 17, it immediately follows that the function $RFS_{n,h}^{\mathrm{GIP}_{n,d}}$ is d^h-versatile on $U_{p,h}^{\mathrm{GIP}_{n,d}}$ and $U_{p,h}^{\mathrm{GIP}_{n,d}}$ is a critical algebraic set.

□

5.3 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 $RFS_{n,h}^{MAJ}$ 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 $U_{1,h}^{MAJ}$.*

Proof. By Lemma 15, $RFS_{n,h}^{\mathrm{MAJ}}$ is $\left(\frac{n+1}{2}\right)^h$-versatile on $U_{p,h}^{\mathrm{MAJ}}$ and $U_{p,h}^{\mathrm{MAJ}}$ 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 $RFS_{n,h}^{GIP_{n,d}}$ 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 $U_{1,h}^{GIP_{n,d}}$.*

Proof. By Lemma 18, $RFS_{n,h}^{\mathrm{GIP}_{n,d}}$ is d^h-versatile on $U_{p,h}^{\mathrm{GIP}_{n,d}}$ and $U_{p,h}^{\mathrm{GIP}_{n,d}}$ 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.

□

5.4 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 $\mathrm{BQP}^A \not\subset \mathrm{PH}^A$. This follows due to the relationship between such circuits and the polynomial hierarchy ([FSS84],[Yao85]) and the fact that there is an efficient quantum algorithm for the recursive Fourier sampling problem ([BV93],[Aar03],[Joh08]), when $h = O(\log n)$. Such a bound is at least plausable 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 [Raz87],[Smo87]. 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, $\mathrm{reg}(U^g_{p,h})$, and the regularities of the preimages of 0 and 1, $\mathrm{reg}(U^g_{0,h})$ and $\mathrm{reg}(U^g_{1,h})$. 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 [Bra09], 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 relatived 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 existance of an A such that $\mathrm{BQP}^A \not\subset \mathrm{AM}^A$. 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 $U^g_{p,h}$,

$$\mathrm{SM}(U^g_{p,h} \cap V, j) = \mathrm{SM}(U^g_{0,h} \cap V, j) = \mathrm{SM}(U^g_{1,h} \cap V, j) \ \forall j \leq \delta - d.$$

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 $U^g_{1,h} \cap V$.

6 VC Dimension

In this section, we answer an open question posed in [MR15]. We begin with a few definitions. We begin by recalling several key results from that paper.

Lemma 19. *[MR15](Thm.2.2) For any $C \subseteq \{0,1\}^n$, reg(C) $\leq$ VC(C).*

It was shown that, if C_{i_j} denotes the value of the i^{th} element of C in the j^{th} position, then

Lemma 20. *[MR15](Prop.6.1) For any $C \subseteq \{0,1\}^n$, reg(C) = 1 precisely when the matrix*

$$\begin{pmatrix} 1 & C_{1,1} & \cdots & C_{1,n} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ 1 & C_{m,1} & \cdots & C_{m,n} \end{pmatrix}$$

has rank $m = |C|$.

They then asked if there was a similar simple characterization of when $\text{reg}(C) = r$, 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 $\text{rank}_{\mathbb{F}_2}\mathcal{M}(C, \binom{[n]}{\leq r}) = |C|$.*

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 $h^a(C,r) = |C|$.

□

Remark 2. *It is straightforward to see that [MR15](Prop.6.1) is a special case of the above theorem.*

7 Acknowledgements

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 [MR15] to my attention, and for helpful discussions, as well as to Scott Aaronson for bringing the recursive Fourier sampling problem to my attention.

References

[Aar03] S. Aaronson, *Quantum lower bound for recursive Fourier sampling*, Quantum Information and Computation (2003), 3(2):165-174.

[Aar10] 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.

[Ajt83] M. Ajtai, Σ_1^1*-Formulae on finite structure*, APAL (1983).

[BBC+98] 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.

[BV93] 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.

[BV97] E. Bernstein and U. Vazirani, *Quantum complexity theory*, SIAM J. Comput. (1997), 26(5):1411-1473.

[Bou07] J. Bourgain, *On the construction of affine extractors*, Geometric and Functional Analysis (2007), 17(1):33-57.

[Bra09] M. Braverman, *Poly-logarithmic independence fools AC0 circuits*, IEEE Conference on Computational Complexity (2009), 3-8.

[CT13] G. Cohen and A. Tal, *Two structural results for low degree polynomials and applications*, ECCC (2013), TR. No. 145.

[DGW07] Z. Dvir, A. Gabizon, and A. Wigderson, *Extractors and rank extractors for polynomial sources*, In FOCS '07 (2007).

[Dvi12] Z. Dvir, *Extractors for varieties*, Computational Complexity (2012), 21(4):515-572.

[Eis02] D. Eisenbud, *The Geometry of Syzygies*, (2002).

[Fel07] B. Felszeghy, *Grobner Theory of Zero Dimensional Ideals with a View Towards Combinatorics*, Budapest University (2007), Ph. D. thesis.

[FSS84] M. Furst, J. Saxe, and M. Sipser, *Parity, circuits, and the polynomial time hierarchy*, Mathematical Systems Theory (1984), 17:13-27.

[GR05] 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.

[GRS04] 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.

[Has86] J. Hastad, *Computational limitations for small depth circuits*, MIT Press (1986), Ph. D. thesis.

[Joh08] B. Johnson, *Upper and lower bounds for recursive Fourier sampling*, University of California at Berkeley (2008), Ph. D. thesis.

[Joh11] 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.

[KRVZ06] 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.

[KZ03] 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).

[Kop11] 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.

[LMN93] N. Linial, Y. Mansour, and N. Nisan, *Constant depth circuits, Fourier transform, and learnability*, J. ACM (1993), 40(3):607-620.

[Lov79] L. Lovász, *Combinatorial problems and exercises*, North-Holland, Amsterdam (1979), 13.31.

[MR15] S. Moran and C. Rashtchian, *Shattered sets and the Hilbert function*, ECCC (2015), TR. No. 15-189.

[PR08] D. Pintér and L. Rónyai, *On the Hilbert Function of Complementary Set Families*, Annales Univ. Sci. Budapest Sect. Comp. (2008), 29:175-198.

[Raz87] 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.

[Smo87] 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.

[Smo93] R. Smolensky, *On Representations by Low-degree Polynomials*, FOCS (1993).

[TV00] L. Trevisan and S. Vadhan, *Extracting randomness from samplable distributions*, In Proceedings of the 41st Annual Symposium of Foundations of COmputer Science (2000), 32.

[Yao85] A. Yao, *Separating the polynomial-time hierarchy by oracles (preliminary version)*, In Proc. IEEE FOCS (1985), 1-10.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de