Electronic Colloquium on Computational Complexity, Report No. 161 (2010)

A Unified Framework for Testing Linear-Invariant Properties

Arnab Bhattacharyya* Elena Grigorescu† Asaf Shapira‡

October 24, 2010

Abstract

The study of the interplay between the testability of properties of Boolean functions and the invariances acting on their domain which preserve the property was initiated by Kaufman and Sudan (STOC 2008). Invariance with respect to $\mathbb{F}_2$-linear transformations is arguably the most common symmetry exhibited by natural properties of Boolean functions on the hypercube. Hence, an important goal in Property Testing is to describe necessary and sufficient conditions for the testability of linear-invariant properties. This direction was explicitly proposed for investigation in a recent survey of Sudan. We obtain the following results:

1. We show that every linear-invariant property that can be characterized by forbidding induced solutions to a (possibly infinite) set of linear equations can be tested with one-sided error.
2. We show that every linear-invariant property that can be tested with one-sided error can be characterized by forbidding induced solutions to a (possibly infinite) set of *systems* of linear equations.

We conjecture that our result from item (1) can be extended to cover systems of linear equations. We further show that the validity of this conjecture would have the following implications:

1. It would imply that every linear-invariant property that is closed under restrictions to linear subspaces is testable with one-sided error. Such a result would unify several previous results on testing Boolean functions, such as the testability of low-degree polynomials and of Fourier dimensionality.
2. It would imply that a linear-invariant property $\mathcal{P}$ is testable with one-sided error **if and only if** $\mathcal{P}$ is closed under restrictions to linear subspaces, thus resolving Sudan's problem.

*Computer Science and Artificial Intelligence Laboratory, MIT. Email: `abhatt@mit.edu`. Supported in part by a DOE Computational Science Graduate Fellowship and NSF Awards 0514771, 0728645, and 0732334.

†Computer Science and Artificial Intelligence Laboratory, MIT. Email: `elena_g@csail.mit.edu`. Supported by NSF award CCR-0829672.

‡School of Mathematics and School of Computer Science, Georgia Institute of Technology, Atlanta, GA 30332. Email: `asafico@math.gatech.edu`. Supported in part by NSF Grant DMS-0901355.

ISSN 1433-8092

1 Introduction

Let $\mathcal{P}$ be a property of Boolean functions. A *testing* algorithm for $\mathcal{P}$ is a randomized algorithm that can quickly distinguish between the case that f satisfies $\mathcal{P}$ from the case that f is far from satisfying $\mathcal{P}$. The problem of characterizing the properties of Boolean functions for which such an efficient algorithm exists is considered by many to be the most important open problem in this area. Since a complete characterization seems to be out of reach, several researchers have recently considered the problem of characterizing the testable properties $\mathcal{P}$ that belong to certain "natural" subfamilies of properties. One such family that has been extensively studied is the family of so called *linear-invariant* properties. Our main result is two fold. We first show that every property in a large family of linear-invariant properties is indeed testable. Next, we conjecture that an even more general family of properties can be tested and show that such a result would give a *characterization* of the linear-invariant properties that are testable with one-sided error.

1.1 Background on property testing

We start with the formal definitions related to testing Boolean functions. Let $\mathcal{P}$ be a property of Boolean functions over the n-dimensional Boolean hypercube. In other words, $\mathcal{P}$ is simply a subset of the set of functions $f : \{0,1\}^n \to \{0,1\}$. Two functions $f, g : \{0,1\}^n \to \{0,1\}$ are ϵ-far if they differ on at least $\epsilon 2^n$ of the inputs. We say that f is ϵ-far from satisfying a property $\mathcal{P}$ if it is ϵ-far from any function g satisfying $\mathcal{P}$. A *tester* for the property $\mathcal{P}$ is a randomized algorithm which can quickly distinguish between the case that an input function f satisfies $\mathcal{P}$ from the case that it is ϵ-far from satisfying $\mathcal{P}$. Here we assume that the input function f is given to the tester as an oracle, that is, the tester can ask an oracle for the value of the input functions f on a certain $x \in \{0,1\}^n$. We say that $\mathcal{P}$ is *strongly testable* (or simply *testable*) if $\mathcal{P}$ has a tester which makes only a constant number of queries to the oracle, where this constant can depend on ϵ but should be independent[1] of n. Finally, we say that a testing algorithm has *one-sided* error if it always accepts input functions satisfying $\mathcal{P}$. (We always demand that the tester rejects input functions which are ϵ-far from satisfying $\mathcal{P}$ with probability at least, say, 2/3.)

The study of testing of Boolean functions began with the work of Blum, Luby and Rubinfeld [BLR93] on testing linearity of Boolean functions. This work was further extended by Rubinfeld and Sudan [RS96]. Around the same time, Babai, Fortnow and Lund [BFL91] also studied similar problems as part of their work on MIP=NEXP. These works are all related to the PCP Theorem, and an important part of it involves tasks which are similar in nature to testing properties of Boolean functions. The work of Goldreich, Goldwasser and Ron [GGR98] extended these results to more combinatorial settings, and initiated the study of similar problems in various areas. More recently, numerous testing questions in the Boolean functions settings have sparked great interest: testing dictators [PRS02], low-degree polynomials [AKK⁺05, Sam07], juntas [FKR⁺04, Bla09], concise representations [DLM⁺07], halfspaces [MORS09], codes [KL05, KS07, KS09]. These are documented in several surveys [Fis04, Rub06, Ron08, Sud10], and we refer the reader to these surveys for more background and references on property testing.

[1] Observe that since we aim for asymptotic results (that is, we think of $n \to \infty$), our property $\mathcal{P}$ can actually be described as $\mathcal{P} = \bigcup_{i=1}^{\infty} \mathcal{P}_n$, where $\mathcal{P}_n$ is the collection of functions over the n-dimensional Boolean hypercube which satisfy $\mathcal{P}$.

1.2 Invariance in testing Boolean functions

What features of a property make it testable? One area in which this question is relatively well understood is testing properties of dense graphs [AS08a, AFNS06, BCL$^+$06]. In sharp contrast, this question is far from being well understood in the case of testing properties of Boolean functions. In an attempt to remedy this, Sudan and several coauthors [KS08, GKS08, GKS09, BS09] have recently begun to investigate the role of invariance in property testing. The idea is that in order to be able to test if a combinatorial structure satisfies a property using very few queries to its representation, the property we are trying to test must be closed under certain transformations. For example, when testing properties of dense graphs, we are allowed to ask if two vertices i and j are adjacent in the graph, and the assumption is that the property we are testing is invariant under renaming of the vertices. In other words, if we think of the input as an $\binom{n}{2}$ dimensional 0/1 vector encoding the adjacency matrix of the input, then the property should be closed under transformations (of the edges) which result from permuting the vertices of the graph.

A natural notion of invariance that one can consider when studying Boolean functions over the hypercube is linear-invariance, which is in some sense the analogue for graph properties being closed under renaming of the vertices (we further discuss this analogy in Subsection 1.3). Formally, a property of Boolean functions $\mathcal{P}$ is said to be linear-invariant if for every function $f : \mathbb{F}_2^n \to \{0, 1\}$ satisfying $\mathcal{P}$ and for any $\mathbb{F}_2$-linear transformation $L : \mathbb{F}_2^n \to \mathbb{F}_2^n$ the function $f \circ L$ satisfies $\mathcal{P}$ as well, where we define $(f \circ L)(x) = f(L(x))$. Note that here we identify $\{0, 1\}^n$ with $\mathbb{F}_2^n$, and we will use this convention from now on throughout the paper. For a thorough discussion of the importance of linear-invariance, we refer the reader to Sudan's recent survey on the subject [Sud10] and to the paper of Kaufman and Sudan which initiated this line of work [KS08].

1.3 The main result

Our main result in this paper (stated in Theorem 3 below) is that a natural family of linear-invariant properties of Boolean functions can all be tested with one-sided error. The statement requires some preparation.

Definition 1 (**(M, σ)-free**) *Given an $m \times k$ matrix M over $\mathbb{F}_2$ and $\sigma \in \{0, 1\}^k$ for integers $m > 0$ and $k > 2$, we say that a function $f : \mathbb{F}_2^n \to \{0, 1\}$ is* (M, σ)-free *if there is no $x = (x_1, \ldots, x_k) \in (\mathbb{F}_2^n)^k$ such that $Mx = 0$ and for all $1 \le i \le k$ we have $f(x_i) = \sigma_i$.*

Remark: By removing linearly dependent rows, we can ensure that $\mathsf{rank}(M) = m$ without loss of generality. We will assume this fact henceforth.

Let us give some intuition about the above definition. Given a function $f : \mathbb{F}_2^n \to \{0, 1\}$, it is natural to consider the set $S_f = \{x \in \mathbb{F}_2^n : f(x) = 1\}$. Suppose for the rest of this paragraph that in the above definition $\sigma = 1^k$. In this case f is (M, σ)-free if and only if S_f contains no solution to the system of equations $Mx = 0$, that is, if there is no $v \in S_f^k$ satisfying $Mv = 0$. Note that when considering graph properties, the notion of $(M, 1^k)$-freeness is analogous to the graph property of being H-free[2], where H is some fixed graph. Observe that in both cases the property is *monotone* in the sense that if f is $(M, 1^k)$-free, then removing elements from S_f results in a set that contains

[2]If H is a graph on h vertices, then we say that a graph G is H-free if G contains no set of h vertices that contain a copy of H (possibly with some other edges).

no solution to $Mx = 0$. Similarly if G is H-free, then removing edges from G results in an H-free graph.

Let us now go back to considering arbitrary $\sigma \in \{0,1\}^k$ in Definition 1, where again the intuition comes from graph properties. Observe that a natural variant of the monotone graph property of being H-free is the property of being induced H-free[3]. Note that being induced H-free is no longer a monotone property since if G is induced H-free then removing an edge can actually create induced copies of H. Getting back to the property of being (M,σ)-free, observe that we can think of this as requiring S_f to contain no *induced* solution to the system of equations $Mx = 0$. That is, the requirement is that there should be no vector v satisfying $Mv = 0$, where $v_i \in S_f$ if $\sigma_i = 1$ and $v_i \in \mathbb{F}_2^n \setminus S_f$ if $\sigma_i = 0$. So we can think of σ as encoding which elements of a potential solution vector v should belong to S_f and which should belong to its complement. For this reason we will adopt the convention of calling (M,σ) a *forbidden induced system of equations*.

Continuing with the graph analogy, once we have the property of being induced H-free, for some fixed graph H, it is natural to consider the property of being induced $\mathcal{H}$-free where $\mathcal{H}$ is a fixed finite set of graphs. Several natural graph properties can be described as being induced $\mathcal{H}$-free (e.g. being a line-graph), but it is of course natural to further generalize this notion and allow $\mathcal{H}$ to contain an infinite number of forbidden induced graphs. One then gets a very rich family of properties like being Perfect, k-colorable, Interval, Chordal etc. This generalization naturally motivates the following definition which will be key to our main results.

Definition 2 ($\mathcal{F}$-free) *Let $\mathcal{F} = \{(M^1,\sigma^1),(M^2,\sigma^2),\dots\}$ be a (possibly infinite) set of induced systems of linear equations. A function f is said to be $\mathcal{F}$-free if it is (M^i,σ^i)-free[4] for all i.*

Observe that this definition is an OR-AND type restriction, that is, we require that f will not satisfy *any* of the systems (M^i,σ^i), where f satisfies (M^i,σ^i) if it satisfies all the equations of M^i (in the sense of Definition 1). We are now ready to state our main result.

Theorem 3 (Main Result) *Let $\mathcal{F} = \{(M^1,\sigma^1),(M^2,\sigma^2),\dots\}$ be a possibly infinite set of induced equations (that is, all the matrices M^i are of rank one), each on more than two variables. Then the property of being $\mathcal{F}$-free is testable with one-sided error.*

Note that, in the above statement, each M^i contains a single equation, rather than a system of equations as in Definition 2. In fact, though, what we prove is quite a bit stronger: Theorem 3 holds when each M^i is of *complexity* 1, instead of just rank 1. The notion of complexity of a linear system is derived from work by Green and Tao [GT08] (See Section 3.2 for the formal definition.) There, we also show that any matrix of rank at most two is of complexity 1, and, hence, Theorem 3 is obviously a corollary of this stronger result. But for the sake of simplicity, let us restrict ourselves to discussing matrices of rank one in this section.

Let us compare this result to some previous works. One work that initiated some of the recent results on testing Boolean functions was obtained by Green [Gre05]. His result can be formulated as saying that for any rank one matrix M, the property of being $(M,1^k)$-free can be tested with one-sided error. Green conjectured that the same result holds for any *system* of linear equations. This conjecture was recently confirmed by Shapira [Sha09] and Kráľ, Serra and Vena [KSV08].

[3] If H is a graph on h vertices, then we say that a graph G is induced H-free if G contains no set of h vertices that contain a copy of H and no other edges.

[4] In the sense of Definition 1

In our language, the results of [Sha09, KSV08] can be stated as saying that for any matrix M, the property of being $(M, 1^k)$-free is testable with one-sided error. The case of arbitrary σ was first explicitly considered in [BCSX09] where it was shown that if M is a rank one matrix, then (M, σ)-freeness is equivalent to a finite set of properties, all of which were already known to be testable. Tim Austin (see [Sha09]) conjectured that the result of [Sha09] for an arbitrary matrix M can be extended to show testability of (M, σ)-freeness for every vector σ. Shapira [Sha09] further conjectured that his result can be extended to the case when we forbid an infinite set of systems of linear equations as in Definition 2. So Theorem 3 partially resolves the above conjecture, since it can handle an infinite number of induced equations (but not an infinite number of forbidden arbitrary *systems* of equations).

Another way to think of Theorem 3 comes (yet again) from the analogy with graph properties. Alon and Shapira [AS08a] have shown that for every set of graphs $\mathcal{F}$, the property of being induced $\mathcal{F}$-free is testable with one-sided error. Since in many ways[5], copies of a fixed graph H in a graph G correspond to finding solutions of a *single* equation in a set $S \subseteq \mathbb{F}_2^n$, Theorem 3 can be considered to be a Boolean functions analog of the result of [AS08a]. Just like the graph property of being free of a particular subgraph H is analogous to being (M, σ)-free where M has rank 1, the *hypergraph* property of being free of a particular sub-hypergraph $\mathcal{H}$ is analogous to being (M, σ)-free for an arbitrary M. Now, the result of [AS08a] has been later extended to hypergraphs by Austin and Tao [AT08] and Rödl and Schacht [RS09]; so, it is natural to expect that one could also handle an infinite number of forbidden induced systems of equations in the functional case as well. All the above motivates us to raise the following conjecture.

Conjecture 4 *For every (possibly infinite) set of systems of induced equations $\mathcal{F}$, the property of being $\mathcal{F}$-free is testable with one-sided error.*

As the reader can easily convince himself, a graph property $\mathcal{P}$ is equivalent to being induced $\mathcal{H}$-free if and only if $\mathcal{P}$ is closed under vertex removal. Such properties are usually called *hereditary*. This motivates us to define the following analogous notion for properties of Boolean functions.

Definition 5 (Subspace-Hereditary Properties) *A linear-invariant property $\mathcal{P}$ is said to be* subspace-hereditary *if it is closed under restriction to subspaces. That is, if f is in $\mathcal{P}_n$ and H is a m-dimensional linear subspace of $\mathbb{F}_2^n$, then $f|_H \in \mathcal{P}_m$ also, where*[6] *$f|_H : \mathbb{F}_2^m \to \{0,1\}$ is the restriction of f to H.*

When considering linear-invariant properties, one can also obtain the following (slightly cleaner) view of the properties of Definition 2. This equivalence is analogous to the graph properties mentioned above. We stress that this equivalence is a further indication of the "naturalness" of the notion of linear-invariance and its resemblance to the closure of graph properties under vertex renaming. We defer its proof to the appendix.

Proposition 6 *A linear-invariant property $\mathcal{P}$ is subspace-hereditary if and only if there is a (possibly infinite) set of systems of induced equations $\mathcal{F}$ such that $\mathcal{P}$ is equivalent to being $\mathcal{F}$-free.*

[5] This analogy is informal, but see [KSV09] and [Sze10] for some formal connections.

[6] Note that we are implicitly composing $f|_H$ with a linear transformation so that it is now defined on $\mathbb{F}_2^m$. Here, we are using the fact that $\mathcal{F}$ is linear-invariant.

We mention that while the notions of graph properties being hereditary and functions being subspace-hereditary are somewhat more natural than the equivalent notions of being free of induced subgraphs and equations respectively, it is actually easier to think about these properties using the latter notion when proving theorems about them. This was the case in [AS08a], and it will be the case in the present paper as well. Proposition 6 along with Conjecture 4 implies the following:

Corollary 7 *If Conjecture 4 holds, then every linear-invariant subspace-hereditary property is testable with one-sided tester.*

Observe that if Conjecture 4 holds, then Corollary 7 would give yet another surprising similarity between linear-invariant properties of boolean functions and graph properties, since it is known [AS08a] that every hereditary graph property is testable. Actually, as we discuss in the next subsection, if Conjecture 4 holds, then an even stronger similarity would follow.

Many interesting properties of the hypercube that have been studied for testability are linear-invariant. Important examples include linearity [BLR93], being a polynomial of low degree [AKK+05], and low Fourier dimensionality and sparsity [GOS+09]. These properties have all been shown to be testable. Moreover, they all turn out to be subspace-hereditary. Thus, if our Conjecture 4 is true, as we strongly believe, then we could explain the testability of all these properties through a unified perspective that uses no features of these properties other than their linear invariance. Note that our main result, Theorem 3, already shows (yet again!) that linearity is testable but from a completely different viewpoint than used in previous analysis. Furthermore, to show the testability of low degree polynomials (a.k.a., Reed-Muller codes), we would only need to resolve Conjecture 4 for a *finite* [7] family of forbidden induced systems of equations.

1.4 The proposed characterization of testable linear-invariant properties

We now turn to discuss our second result, which based on Conjecture 4 gives a characterization of the linear-invariant properties of Boolean functions that can be tested with one-sided error using "natural" testing algorithms. Let us start with formally defining the types of "natural" testers we consider here.

Definition 8 (Oblivious Tester) *An* oblivious tester *for a property $\mathcal{P} = \{\mathcal{P}_n\}_n$ is a (possibly 2-sided error) non-adaptive, probabilistic algorithm, which, given a distance parameter ϵ, and oracle access to an input function $f : \mathbb{F}_2^n \to \{0,1\}$, performs the following steps:*

1. *Computes an integer $d = d(\epsilon)$. If $d(\epsilon) > n$, let $H = \mathbb{F}_2^n$. Otherwise, let $H \leq \mathbb{F}_2^n$ be a subspace of dimension $d(\epsilon)$ chosen uniformly at random.*
2. *Queries f on all elements $x \in H$.*
3. *Accepts or rejects based only on the outcomes of the received answers, the value of ϵ, and its internal randomness.*

We now discuss the motivation for considering the above type of algorithms. The fact that the tester is non-adaptive and queries a random linear subspace is without loss of generality (see Proposition 33); this is analogous to the fact [AFKS00, GT03] that one can assume a graph property

[7] The characterization of polynomials of degree d using forbidden induced equations is shown in Appendix A.

tester makes its decision only by inspecting a randomly chosen induced subgraph. The only essential restriction we place on oblivious testers is that their behavior cannot depend on the value of n, the domain size of the input function. If we allow the testing algorithm to make its decisions based on n, then it can do very strange and unnatural things. For example, we can now consider properties that depend on the parity of n. As was shown in [AS08b], the algorithm can use the size of the input in order to compute the optimal query complexity. All these abnormalities will not allow us to give any meaningful characterization. As observed in [AS08a] by restricting the algorithm to make its decisions while not considering the size of the input, we can still test any (natural) property while at the same time avoid annoying technicalities. We finally note that all the testing algorithms for testable properties of Boolean functions in prior works were indeed oblivious, and that furthermore many of them implicitly consider only oblivious testers. In particular, these types of testers were considered in [Sud10].

As it turns out, oblivious testers can potentially[8] test properties which are slightly more general than subspace-hereditary properties. These are defined as follows.

Definition 9 (Semi Subspace-Hereditary Property) *A property $\mathcal{P} = \{\mathcal{P}_n\}_n$ is* semi subspace-hereditary *if there exists a subspace-hereditary property $\mathcal{H}$ such that*

1. *Any function f satisfying $\mathcal{P}$ also satisfies $\mathcal{H}$.*
2. *There exists a function $M : (0,1) \to \mathbb{N}$ such that for every $\epsilon \in (0,1)$, if $f : \mathbb{F}_2^n \to \{0,1\}$ is ϵ-far from satisfying $\mathcal{P}$ and $n \geq M(\epsilon)$, then $f|_V$ does not satisfy $\mathcal{H}$.*

The intuition behind the above definition is that a semi subspace-hereditary property can only deviate from being "truly" subspace-hereditary on functions over a finite domain, where the finiteness is controlled by the function M in the definition. Our next theorem connects the notion of oblivious testing and semi subspace-hereditary properties. Assuming Conjecture 4, it essentially characterizes the linear-invariant properties that are testable with one-sided error, thus resolving Sudan's problem raised in [Sud10].

Theorem 10 *If Conjecture 4 holds, then a linear-invariant property $\mathcal{P}$ is testable by a one-sided error oblivious tester* **if and only if** *$\mathcal{P}$ is semi subspace-hereditary.*

Getting back to the similarity to graph properties, we note that [AS08a] obtained a similar characterization for the graph properties that are testable with one-sided error. Let us close by mentioning two points. The first is that most linear-invariant properties are known to be testable with one-sided error, and hence the question of characterizing these properties is well motivated. In fact, for the subclass of linear-invariant properties which also themselves form a linear subspace, [BHR05] showed that the optimal tester is always one-sided and non-adaptive. Our second point is that it is natural to ask if there are linear-invariant properties which are not testable. A linear-invariant property with query complexity $\Omega(2^n)$ arises implicitly from the arguments of [GGR98]; see Section 5 for a brief sketch. A second, more natural, example comes from Reed-Muller codes. [BKS$^+$09] shows that for any $1 \ll q(n) \ll n$ the linear-invariant property of being a $\log_2(q(n))$-Reed-Muller code cannot be tested with $o(q(n))$ queries. We also conjecture that the property of two functions being isomorphic upto linear transformations of the variables is not a testable

[8] The potential relies on the validity of Conjecture 4.

property. Lower bounds for isomorphism testing have been studied both in the Boolean function model [FKR+04, BO10] and in the dense graph model [Fis05], but our problem specifically does not seem to have been examined in a property testing setting.

1.5 Paper overview

The rest of the paper is organized as follows. In Section 2, we discuss the regularity lemma of Green [Gre05]. Just as the graph regularity lemma of Szemerédi [Sze78] guarantees that every graph can be partitioned into a bounded number of pseudorandom graphs, Green's regularity lemma guarantees a similar partition for Boolean functions. This lemma, whose proof relies on Fourier analysis over $\mathbb{F}_2^n$, was used in [Gre05] to show that properties defined by forbidding a single (non-induced) equation are testable. This basic approach falls short of being able to handle an infinite number of forbidden non-induced equations or even a single forbidden induced equation. We thus need to develop a variant of Green's regularity lemma that is strong enough to allow such applications. This new variant is described in Section 2. The overall approach is motivated by that taken by Alon et al. [AFNS06] in their formulation of the functional graph regularity lemma. However, the proof here is somewhat more involved since we need to develop several tools in order to make the approach work. One of them is a certain Ramsey type result for $\mathbb{F}_2^n$ which is key to our proof and that may be useful in other settings (see Theorem 19). The approach of [AFNS06] only allows one to handle a *finite* number of forbidden subgraphs, which translates in our setting to being able to handle a finite number of forbidden equations. So, one last technique we employ is motivated by the ideas from [AS08a] on how to handle an infinite number of forbidden subgraphs. This (somewhat complicated) technique is described in Section 3. We believe that these set of ideas will prove to be instrumental in resolving Conjecture 4. Section 5 is devoted to some concluding remarks and open problems.

2 Pseudorandom Partitions of the Hypercube

The *support* of a Boolean function f refers to the subset of the domain on which f evaluates to 1. If H is a subspace of $\mathbb{F}_2^n$ and given function $f : H \to \{0,1\}$, let $\rho(f)$, the *density* of f, denote $\frac{\sum_{x\in H} f(x)}{|H|}$. Recall that the Fourier coefficients of f, defined for each $\alpha \in H^*$, are:

$$\widehat{f}(\alpha) = \mathop{\mathbb{E}}_{x\in H}\left[f(x)\cdot(-1)^{\langle x,\alpha\rangle}\right]$$

For a parameter $\epsilon \in (0,1)$, we say f is ϵ-*uniform* if $\max_{\alpha\neq 0}|\widehat{f}(\alpha)| < \epsilon$. This definition captures the notion of correlation with a linear function on H, and it will serve as our definition of pseudorandomness.

Given a function $f : \mathbb{F}_2^n \to \{0,1\}$, a subspace $H \leq \mathbb{F}_2^n$ and an element $g \in \mathbb{F}_2^n$, define the function $f_H^{+g} : H \to \{0,1\}$ to be $f_H^{+g}(x) = f(x+g)$ for $x \in H$. The support of f_H^{+g} represents the intersection of the support of f with the coset $g + H$. The following lemma shows that if a uniform function is restricted to a coset of a subspace of low codimension, then the restriction does not become too non-uniform and its density stays roughly the same.

Lemma 11 *Let $f : \mathbb{F}_2^n \to \{0,1\}$ be an ϵ-uniform function of density ρ, and let $H \leq \mathbb{F}_2^n$ be a subspace of codimension k. Then for any $c \in \mathbb{F}_2^n$, the function $f_H^{+c} : H \to \{0,1\}$ is $(2^k\epsilon)$-uniform and of density ρ_c satisfying $|\rho_c - \rho| < 2^k\epsilon$.*

Proof: Let $H^\perp = \{\alpha \in \mathbb{F}_2^n | \ \langle \alpha, h\rangle = 0 \ \forall h \in H\}$ be the dual to the vector space H, and let $H' = \mathbb{F}_2^n/H$ be the quotient of H in $\mathbb{F}_2^n$. We wish to show that, for every $c \in H'$, the Fourier coefficients of f_H^{+c} are small.

For every $\beta \in \mathbb{F}_2^n/H^\perp$ and $\alpha \in H^\perp$:

$$\widehat{f}(\beta+\alpha) = \mathop{\mathbb{E}}_{x\in\mathbb{F}_2^n}[f(x)\chi_{\beta+\alpha}(x)] = \mathop{\mathbb{E}}_{c'\in H'} \mathop{\mathbb{E}}_{h\in H} f_H^{+c'}(h)\chi_{\beta+\alpha}(c'+h) = \mathop{\mathbb{E}}_{c'\in H'} \chi_{\beta+\alpha}(c') \mathop{\mathbb{E}}_{h\in H} f_H^{+c'}(h)\chi_\beta(h)$$
$$= \frac{1}{2^k}\sum_{c'\in H'} \chi_{\beta+\alpha}(c')\widehat{f}_H^{+c'}(\beta)$$

Recall that $\sum_{\alpha\in H^\perp}\chi_\alpha(c') = \begin{cases} 0, & \text{if } c' \neq 0 \\ 1, & \text{if } c' = 0. \end{cases}$ Fixing $\beta \in \mathbb{F}_2^n/H^\perp$ and $c \in H'$ and summing up the quantity computed above over all $\alpha \in H^\perp$, we obtain

$$\begin{aligned}
2^k\left(\sum_{\alpha\in H^\perp}\chi_{\beta+\alpha}(c)\widehat{f}(\beta+\alpha)\right) &= \sum_{c'\in H'}\sum_{\alpha\in H^\perp}\chi_{\beta+\alpha}(c+c')\widehat{f}_H^{+c'}(\beta) \\
&= \sum_{\alpha\in H^\perp}\chi_{\beta+\alpha}(0)\widehat{f}_H^{+c}(\beta) + \sum_{c'\in H'-\{c\}}\sum_{\alpha\in H^\perp}\chi_{\beta+\alpha}(c+c')\widehat{f}_H^{+c'}(\beta) \\
&= 2^k\widehat{f}_H^{+c}(\beta) + \sum_{c'\in H'-\{0\}}\sum_{\alpha\in H^\perp}\chi_{\beta+\alpha}(c')\widehat{f}_H^{+c'+c}(\beta) \\
&= 2^k\widehat{f}_H^{+c}(\beta) + \sum_{c'\in H'-\{0\}}\chi_\beta(c')\left(\sum_{\alpha\in H^\perp}\chi_\alpha(c')\right)\widehat{f}_H^{+c'+c}(\beta) \\
&= 2^k\widehat{f}_H^{+c}(\beta).
\end{aligned}$$

Furthermore,

$$\left|\widehat{f}_H^{+c}(\beta)\right| = \left|\sum_{\alpha\in H^\perp}\chi_{\beta+\alpha}(c)\widehat{f}(\beta+\alpha)\right| \leq \sum_{\alpha\in H^\perp}\left|\chi_{\beta+\alpha}(c)\widehat{f}(\beta+\alpha)\right| = \sum_{\alpha\in H^\perp}\left|\widehat{f}(\beta+\alpha)\right|$$

Since f is ϵ-uniform, setting $\beta = 0$ in the above inequality shows that $|\rho_c - \rho| \leq \sum_{0\neq\alpha\in H^\perp}|\widehat{f}(\alpha)| < 2^k\epsilon$. For nonzero β in $\mathbb{F}_2^n/H^\perp$, it follows again from ϵ-uniformity that $|\widehat{f}_H^{+c}(\beta)| < 2^k\epsilon$. ■

For a subspace $H \leq \mathbb{F}_2^n$, the *H-based partition* refers to the partitioning of $\mathbb{F}_2^n$ into the cosets in $\mathbb{F}_2^n/H$. If $H' \leq H$, then the H'-based partition is called a *refinement* of the H-based partition. The *order* of the H-based partition is defined to be $[G : H]$, i.e., the index of H as a subgroup or the dimension of the quotient space $\mathbb{F}_2^n/H$. Using this notation, Green's regularity lemma can be stated as follows.

Lemma 12 (Green's Regularity Lemma [Gre05]) *For every m and $\epsilon > 0$, there exists $T = T_{12}(m, \epsilon)$ such that the following is true. Given function $f : \mathbb{F}_2^n \to \{0, 1\}$ with $n > T$ and H-based partition of $\mathbb{F}_2^n$ with order at most m, there exists a refined H'-based partition of order k, with $m \leq k \leq T$, for which $f_{H'}^{+g}$ is not ϵ-uniform for at most $\epsilon 2^n$ many $g \in \mathbb{F}_2^n$.*

Our main tool in this work is a functional variant of Green's regularity lemma, in which the uniformity parameter ϵ is not a constant but rather an arbitrary function of the order of the partition. It is quite analogous to a similar lemma, first proved in [AFKS00], in the graph property testing setting. The recent work [GT10] shows a (very strong) functional regularity lemma in the arithmetic setting but it applies over the integers and not $\mathbb{F}_2$.

Lemma 13 (Functional regularity lemma) *For integer m and function $\mathcal{E} : \mathbb{Z}^+ \to (0,1)$, there exists $T = T_{13}(m, \mathcal{E})$ such that the following is true. Given function $f : \mathbb{F}_2^n \to \{0,1\}$ with $n \geq T$, there exist subspaces $H' \leq H \leq \mathbb{F}_2^n$ that satisfy:*

- *Order of H-based partition is $k \geq m$, and order of H'-based partition is $\ell \leq T$.*
- *There are at most $\mathcal{E}(0) \cdot 2^n$ many $g \in \mathbb{F}_2^n$ such that f_H^{+g} is not $\mathcal{E}(0)$-uniform.*
- *For every $g \in \mathbb{F}_2^n$, there are at most $\mathcal{E}(k) \cdot 2^{n-k}$ many $h \in H$ such that $f_{H'}^{+g+h}$ is not $\mathcal{E}(k)$-uniform.*
- *There are at most $\mathcal{E}(0) \cdot 2^n$ many $g \in \mathbb{F}_2^n$ for which there are more than $\mathcal{E}(0) \cdot 2^{n-k}$ many $h \in H$ such that $|\rho(f_H^{+g}) - \rho(f_{H'}^{+g+h})| > \mathcal{E}(0)$.*

Proof: Let us first give an informal overview of the proof. The basic idea is to repeatedly apply Lemma 12, at each step refining the partition obtained in the previous step. At each step, Lemma 12 is applied with a uniformity parameter that depends on the order of the partition obtained in the previous step. We stop when the *index* of the partitions stop increasing substantially. Given a subspace H, the index of the H-based partition is defined to be the variance of the densities in the cosets:

$$\mathsf{ind}(f, H) \stackrel{\text{def}}{=} \frac{1}{2^n} \sum_{g \in \mathbb{F}_2^n} \rho^2(f_H^{+g})$$

We show that when the indexes of two successive partitions are close, then on average, each coset of the finer partitioning has roughly the same density as the coset of the coarser partitioning it is contained in.

To implement the above ideas, we need the following two claims about the index of partitions. Their proofs are essentially identical to those for the corresponding Lemmas 3.6 and 3.7 respectively in [AFKS00], and so we are a bit brief in the following.

Claim 14 *Given subspace $H \leq \mathbb{F}_2^n$ and function $f : \mathbb{F}_2^n \to \{0,1\}$, suppose that there are at least $\epsilon 2^n$ many $g \in \mathbb{F}_2^n$ such that $|\rho(f) - \rho(f_H^{+g})| > \epsilon$. Then:*

$$\mathsf{ind}(f, H) > \rho^2(f) + \frac{\epsilon^3}{2}$$

Proof: Observe that the average of $\rho(f_H^{+g})$ over all $g \in \mathbb{F}_2^n$ equals $\rho(f)$. From our assumptions, either there are $\frac{\epsilon}{2} 2^n$ many $g \in \mathbb{F}_2^n$ such that $\rho(f) - \rho(f_H^{+g}) > \epsilon$ or there are $\frac{\epsilon}{2} 2^n$ many $g \in \mathbb{F}_2^n$ such that $\rho(f) - \rho(f_H^{+g}) < -\epsilon$. For either case, we can use the defect form of the Cauchy-Schwarz inequality to prove our claim. ■

Claim 15 *For function $f : \mathbb{F}_2^n \to \{0,1\}$ and subspaces $H' \leq H \leq \mathbb{F}_2^n$, suppose the H-based partition of order k and its refinement, the H'-based partition, of order ℓ satisfy* $\mathsf{ind}(f, H') - \mathsf{ind}(f, H) \leq \frac{\epsilon^4}{2}$ *for some ϵ. Then, there are at most $\epsilon 2^n$ many $g \in \mathbb{F}_2^n$ for which there are more than $\epsilon 2^{n-k}$ many $h \in H$ satisfying $|\rho(f_H^{+g}) - \rho(f_{H'}^{+g+h})| > \epsilon$.*

Proof: Suppose that there are $> \epsilon 2^n$ many $g \in \mathbb{F}_2^n$ such that there are $> \epsilon 2^{n-k}$ many $h \in H$ satisfying $|\rho(f_H^{+g}) - \rho(f_{H'}^{+g+h})| > \epsilon$. Use Claim 14 to obtain a contradiction:

$$
\begin{aligned}
\mathsf{ind}(f, H') = \frac{1}{2^\ell} \sum_{u \in \mathbb{F}_2^n / H'} \rho^2(f_{H'}^{+u}) &= \frac{1}{2^k} \sum_{v \in \mathbb{F}_2^n / H} \frac{1}{2^{\ell - k}} \sum_{h \in H/H'} \rho^2(f_{H'}^{+v+h}) \\
&= \frac{1}{2^k} \sum_{v \in \mathbb{F}_2^n / H} \mathsf{ind}(f_H^{+v}) \\
&> \frac{1}{2^k} \left(\sum_{v \in \mathbb{F}_2^n / H} \rho^2(f_H^{+v}) + \epsilon \cdot 2^k \frac{\epsilon^3}{2} \right) \\
&= \mathsf{ind}(f, H) + \frac{\epsilon^4}{2}
\end{aligned}
$$

■

Now we have the pieces needed to prove the lemma. We can assume $\mathcal{E}(\cdot)$ is monotone non-increasing. Let $\epsilon = \mathcal{E}(0)$. We define T inductively as follows. Let $T^{(1)} = T_{12}(m, \epsilon)$, and for $i > 1$, let:

$$
T^{(i)} = T_{12} \left(T^{(i-1)}, \mathcal{E} \left(T^{(i-1)} \right) \cdot 2^{-T^{(i-1)}} \right)
$$

Set $T = T_{13}(m, \mathcal{E}) \stackrel{\text{def}}{=} T^{(2\epsilon^{-4}+1)}$.

We now show that this choice of T suffices. Given function $f : \mathbb{F}_2^n \to \{0,1\}$, apply Lemma 12 with m and ϵ to get a subspace H_1, and thereafter repeatedly apply it to get a sequence of finer subspaces $H_2, H_3, H_4, \ldots$, with $H_1 \geq H_2 \geq H_3 \geq H_4 \geq \cdots$, by invoking Lemma 12 at each step $i > 1$ with $T^{(i-1)}$ and $\mathcal{E}\left(T^{(i-1)}\right) \cdot 2^{-T^{(i-1)}}$ as the two input parameters. Stop when $\mathsf{ind}(f, H_{i+1}) - \mathsf{ind}(f, H_i) < \frac{\epsilon^4}{2}$. This happens when i is at most $2\epsilon^{-4} + 1$ because the index of any partition is less than 1. Let $H = H_i$ and $H' = H_{i+1}$. It's clear that the codimension k of H at least m and that the codimension ℓ of H' is at most T. The second item in the lemma follows from the uniformity guarantee of Lemma 12 and from the fact that $\mathcal{E}(T^{(i-1)}) < \mathcal{E}(0)$. For the third, note that Lemma 12 guarantees that there are at most $\mathcal{E}(k) 2^{-k} 2^n = \mathcal{E}(k) 2^{n-k}$ values of $g \in \mathbb{F}_2^n$ such that $f_{H'}^{+g}$ is not $(\mathcal{E}(k)2^{-k})$-uniform and, hence, not $\mathcal{E}(k)$-uniform. So, clearly, there are at most so many g contained in any coset of H. Finally, the fourth item follows from Claim 15. This completes the proof of Lemma 13. ■

We use Lemma 13 in two main ways. For one of them, we use the lemma directly. For the other, we use the following simple but extremely useful corollary which allows us to say that there are many cosets in a partitioning which, on the one hand, are *all* uniform, and on the other hand, are arranged in an algebraically nice structure.

Corollary 16 *For every m and $\mathcal{E} : \mathbb{Z}^+ \to (0,1)$, there exist $T = T_{16}(m, \mathcal{E})$ and $\delta = \delta_{16}(m, \mathcal{E})$ such that the following is true. Given function $f : \mathbb{F}_2^n \to \{0,1\}$ with $n \geq T$, there exist subspaces $H' \leq H \leq \mathbb{F}_2^n$ and an injective linear map $I : \mathbb{F}_2^n / H \to \mathbb{F}_2^n / H'$ such that:*

- *The H-based partition is of order k, where $m \le k \le T$. Additionally, $|H'| \ge \delta 2^n$.*
- *For each $u \in \mathbb{F}_2^n/H$, $I(u) + H'$ lies inside the coset $u + H$. Note that $I(0) = 0$ since I is linear.*
- *For every nonzero $u \in \mathbb{F}_2^n/H$, the set $f_{H'}^{+I(u)}$ is $\mathcal{E}(k)$-uniform.*
- *There are at most $\mathcal{E}(0)2^n$ many $g \in \mathbb{F}_2^n$ for which $|\rho(f_H^{+g}) - \rho(f_{H'}^{+I(u)})| > \mathcal{E}(0)$ where $u = g \pmod{H}$.*

Proof: We can assume $\mathcal{E}$ is a nonincreasing function. Denote $\mathcal{E}(0)$ as ϵ, and set $\mathcal{E}'(r) = \min(\mathcal{E}(r), \frac{\epsilon}{6}, \frac{1}{2^{r+1}})$. We will show that $T = T_{16}(m, \mathcal{E}) \stackrel{\text{def}}{=} T_{13}(m, \mathcal{E}')$ and $\delta = \delta_{16}(m, \mathcal{E}) \stackrel{\text{def}}{=} 1/2^T$ suffice for our proof.

Apply Theorem 13 with m and the function $\mathcal{E}'$ as inputs. Let H and H' be the subspaces obtained there, for the given $f : \mathbb{F}_2^n \to \{0,1\}$. We find I satisfying the conditions of the claim exists using the probabilistic method.

Fix k linearly independent elements $u_1, \dots, u_k \in \mathbb{F}_2^n/H$ (viewing $\mathbb{F}_2^n/H$ as a vector space over $\mathbb{F}_2$). For every $i \in [k]$, choose independently and uniformly at random an element v from H/H' and let $I(u_i)$ equal $u_i + v + H'$. The value of I over the rest of $\mathbb{F}_2^n/H$ is determined by linearity, as the u_i's form a basis for $\mathbb{F}_2^n/H$. It's immediate that $I(u) + H'$ lies inside $u + H$ for every $u \in \mathbb{F}_2^n/H$.

Observe that unless $u = 0$, each $I(u) + H'$ is uniformly distributed among the cosets of H' lying in $u + H$. Hence, for any nonzero u, the probability that $f_{H'}^{+I(u)}$ is not $\mathcal{E}(k)$-uniform is at most $1/2^{k+1}$, by our choice of parameters. Applying the union bound, the probability that there exists nonzero $u \in \mathbb{F}_2^n/H$ such that $f_{H'}^{+I(u)}$ is not $\mathcal{E}(k)$-uniform is at most $1/2$. Also, the expected number of $g \in \mathbb{F}_2^n$, with $u = g \pmod{H}$, for which $|\rho(f_H^{+g}) - \rho(f_{H'}^{+I(u)})| > \epsilon$ is at most $\frac{\epsilon}{6}2^n + \frac{\epsilon}{6}2^n + 1 \le \frac{\epsilon}{2}2^n$, and hence by the Markov inequality, with probability at least $\frac{1}{2}$, the number of $g \in \mathbb{F}_2^n$ satisfying this condition is at most $\epsilon 2^n$. Therefore, there must exist a choice of I making both the third and fourth claims true. ■

The next lemma is in a similar spirit to Corollary 16. It also obtains a set of uniform cosets which are structured algebraically, but in this case, all of them are contained inside the same subspace.

Lemma 17 *For every positive integer d and $\gamma \in (0,1)$, there exists $\delta = \delta_{17}(d, \gamma)$ such that the following is true. Given $f : \mathbb{F}_2^n \to \{0,1\}$, there exists a subspace $H \le \mathbb{F}_2^n$ and a subspace K of dimension d in the quotient space $\mathbb{F}_2^n/H$ with the following properties:*

- *$|H| \ge \delta 2^n$.*
- *For every nonzero $u \in K$, f_H^{+u} is γ-uniform.*
- *Either $\rho(f_H^{+u}) \ge \frac{1}{2}$ for every nonzero $u \in K$ or $\rho(f_H^{+u}) < \frac{1}{2}$ for every nonzero $u \in K$.*

We need a different set of tools to prove this lemma. Specifically, we use linear algebraic variants of the classic theorems of Turán and Ramsey. We note that the (classic) Turán and Ramsey Theorems are key tools in many applications of the graph regularity lemma, for example in the well known bound on the Ramsey numbers of bounded degree graphs [CRSW83]. Hence, the variants that we use of these classic results may be useful in other applications of Greens's regularity lemma.

Proposition 18 (Turán theorem for subspaces) *For positive integers n, if S is a subset of $\mathbb{F}_2^n$ with density greater than $1 - \frac{1}{2^{d-1}}$, then there exists a subspace $H \leq \mathbb{F}_2^n$ of dimension d such that $H - \{0\}$ is contained in S. Moreover, there is a subset of $\mathbb{F}_2^n$ with density $\left(1 - \frac{1}{2^{d-1}}\right)$ which does not contain $H - \{0\}$ for any subspace $H \leq \mathbb{F}_2^n$.*

Proof: Let $S \subseteq \mathbb{F}_2^n$ be a maximal set that does not contain $H - \{0\}$ for any d-dimensional subspace H. Since S is maximal, it must contain $K - \{0\}$ for some $(d-1)$-dimensional subspace K (if not, we can simply add it to S without introducing points of $H - \{0\}$ for any d-dimensional subspace H). Let K' be an $(n-d+1)$-dimension subspace that intersects K only at $\{0\}$.

Now, observe that for any nonzero $\alpha \in K'$, at least one of the elements of $\{\alpha + k : k \in K\}$ must not belong to S. Otherwise, S would contain $(K - \{0\}) \cup \{\alpha + k : k \in K\} = H - \{0\}$ for a d-dimensional subpace $H = \operatorname{span}(K \cup \{\alpha\})$, contradicting our assumption for S. Thus, we can upper-bound the number of points in S by:

$$|S| \leq |K' - \{0\}| \cdot (|K| - 1) + |K - \{0\}| = (2^{n-d+1} - 1) \cdot (2^{d-1} - 1) + (2^{d-1} - 1) = 2^n - 2^{n-d+1}$$

To see that the above bound is tight, let $S = \mathbb{F}_2^n - K'$ for any $(d-1)$-dimensional subspace $K \leq \mathbb{F}_2^n$ and K' as above. It is easy to check that this S does not contain $H - \{0\}$ for any $H \leq \mathbb{F}_2^n$ with $\dim(H) = d$. ■

Theorem 19 (Ramsey theorem for subspaces) [9] *For every positive integer d, there exists $N = N_{19}(d)$ such that for any subset $S \subseteq \mathbb{F}_2^N$, there exists a subspace $H \leq \mathbb{F}_2^N$ of dimension d such that $H - \{0\}$ is contained either in S or in $\bar{S}$.*

Proof: We will show a stronger statement, which we describe in the following lemma.

Lemma 20 *For every positive integer d_1, d_2, there exists $N(d_1, d_2)$ such that for any subset $S \subseteq \mathbb{F}_2^{N(d_1,d_2)}$, either there exists a subspace $H_1 \leq \mathbb{F}_2^{N(d_1,d_2)}$ of dimension d_1 such that $H_1 - \{0\}$ is contained in S or there exists a subspace $H_2 \leq \mathbb{F}_2^{N(d_1,d_2)}$ of dimension d_2 such that $H_2 - \{0\}$ is contained in $\bar{S}$.*

One can immediately deduce the statement of the theorem by taking $d = d_1 = d_2$ in Lemma 20. To prove Lemma 20 we first prove the following helpful result. For a subspace $H \leq \mathbb{F}_2^n$ we say that an affine subspace $a + H$ is *strict* if $a \in \mathbb{F}_2^n / H - \{0\}$.

Lemma 21 *For every positive integer d, there exists $N_a = N_a(d)$ such that for any subset $S \subseteq \mathbb{F}_2^{N_a}$, there exists a strict affine subspace $A \leq \mathbb{F}_2^{N_a}$ of dimension d such that A is contained either in S or in $\bar{S}$.*

Proof: Notice that $N_a(1) = 1$. Assume, by induction that the lemma holds for dimension $d-1$, and let $N_a(d) = 2^{N_a(d-1)+1} + N_a(d-1)$. Let $S \subseteq \mathbb{F}_2^{N_a(d)}$ be an arbitrary set, let $H = \mathbb{F}_2^{N_a(d-1)}$, and $H' = \mathbb{F}_2^{N_a(d)}/H$. Notice that $|H'| = 2^{2^{N_a(d-1)+1}}$. For each $c \in H' - \{0\}$ consider the set $f_H^{+c} \subset H$. Since there are $2^{2^{N_a(d-1)+1}} - 1$ possible such sets, and each set has size at most $2^{N_a(d-1)}$

[9] As pointed to us recently by Noga Alon, this theorem might be implied by the Folkman-Rado-Sanders Theorem, but we include a self-contained proof for the sake of completeness.

it follows that there exists $c_1 \neq c_2 \in H' - \{0\}$ such that $f_H^{+c_1} = f_H^{+c_2}$. By the induction hypothesis, either $f_H^{+c_1}$ or its complement contains a $d-1$ dimensional affine subspace. Assume w.l.o.g. that $f_H^{+c_1}$ contains an affine subspace $\alpha + f_{d-1}$ of dimension $d-1$ (otherwise replace S by $\bar{S}$), for some $\alpha \in H - f_{d-1}$. Then the affine subspaces $\alpha + c_1 + f_{d-1}$ and $\alpha + c_2 + f_{d-1}$ are both contained in S. Let $A_d = (\alpha + c_1 + f_{d-1}) \cup (\alpha + c_2 + f_{d-1}) \subset S$. To conclude the proof, notice that $A_d = \alpha + c_1 + \mathrm{span}(c_2 - c_1, f_{d-1})$ is a strict affine subspace of dimension d, since $\alpha \neq c_1$ and $c_2 - c_1 \notin f_{d-1}$. ■

Proof of Lemma 20: The proof follows by induction on d_1 and d_2, with the base cases $N(0,1) = N(1,0 = 1$. Assume that there exists $N(d_1 - 1, d_2)$ and $N(d_1, d_2 - 1)$ satisfying the conditions of the lemma. Define

$$N(d_1, d_2) = N_a(\max(N(d_1 - 1, d_2), N(d_1, d_2 - 1))),$$

where $N_a(d)$ is the quantity defined in Lemma 21. We show that for any arbitrary set $S \subseteq \mathbb{F}_2^{N(d_1,d_2)}$ either it contains a subspace of dimension d_1 (except 0) or its complement contains a subspace of dimension d_2 (except 0). Suppose $N(d_1 - 1, d_2) \geq N(d_1, d_2 - 1)$. By definition and by Lemma 21, there exists a strict affine subspace $A \subseteq \mathbb{F}_2^{N(d_1,d_2)}$ such that $A = a + H \subseteq S$ or $A \subseteq \bar{S}$ (where H is the subspace underlining A). Assume for now that the former holds. Since $H \cap S \subseteq \mathbb{F}_2^{N(d_1-1,d_2)}$, by the induction hypothesis, either $H \cap S$ contains a subspace of dimension $d_1 - 1$ or $H - S$ contains a subspace of dimension d_2, in which case we are done. If $H \cap S$ contains a subspace $f_{d_1-1} - \{0\}$ of dimension $d_1 - 1$, then define $f_{d_1} = f_{d_1-1} \cup a + f_{d_1-1} = \mathrm{span}(a, f_{d_1-1})$. Clearly $f_{d_1} \in S$ and it has dimension d_1, which completes the proof of this case. It remains to deal with the case when $A \subseteq \bar{S}$. Since $N(d_1 - 1, d_2) \geq N(d_1, d_2 - 1)$, there exists another affine subspace $A' = a' + H' \subset A \subseteq \bar{S}$ of dimension $N(d_1, d_2 - 1)$. Again, by the induction hypothesis, the set $H' \cap S$ either contains a subspace of dimension d_1, in which case we are done, or $H' - S$ contains a subspace f_{d_2-1} of dimension $d_2 - 1$. In the latter case define $f_{d_2} = f_{d_2-1} \cup a' + f_{d_2-1} = \mathrm{span}(a', f_{d_2-1})$. Finally, notice that $f_{d_2} \in \bar{S}$ and it has dimension d_2. ■

This concludes the proof of Theorem 19.■

Given these results, Lemma 17 follows fairly readily.

Proof of Lemma 17: Set $\delta = \delta_{17}(d,\gamma) \stackrel{\mathrm{def}}{=} 2^{-T_{12}(r,\min(2^{-r-2},\gamma))}$ with $r = N_{19}(d)$. Given $f : \mathbb{F}_2^n \to \{0,1\}$, apply Lemma 12 with inputs r and $\min(2^{-r-2}, \gamma)$ to obtain a subspace H such that restrictions of S to at most 2^{-r-2} fraction of the cosets of the H-based partition are not γ-uniform. Using Proposition 18, there exists a subspace $L \leq \mathbb{F}_2^n/H$ of dimension r such that for every nonzero $u \in L$, the set f_H^{+u} is γ-uniform. Furthermore, since L is of dimension $N_{19}(d)$, by Theorem 19, there exists a subspace $K \leq L \leq \mathbb{F}_2^n/H$ satisfying the final condition of the lemma. ■

3 Forbidding Infinitely Many Induced Equations

In this section, we prove our main result (Theorem 3) that properties characterized by infinitely many forbidden induced equations are testable. To begin, let us fix some notation. Given a matrix M over $\mathbb{F}_2$ of size m-by-k, a string $\sigma \in \{0,1\}^k$, and a function $f : \mathbb{F}_2^n \to \{0,1\}$, if there exists

$x = (x_1, \dots, x_k) \in (\mathbb{F}_2^n)^k$ such that $Mx = 0$ and $f(x_i) = \sigma_i$ for all $i \in [k]$, we say that f *induces* (M, σ) *at* x and denote this by $(M, \sigma) \mapsto f$.

The following theorem is the core of the proof of Theorem 3.

Theorem 22 *For every infinite family of equations $\mathcal{F} = \{(E^1, \sigma^1), (E^2, \sigma^2), \dots, (E^i, \sigma^i), \dots\}$ with each E^i being a row vector $[1\ 1\ \cdots\ 1]$ of size k_i and $\sigma^i \in \{0,1\}^{k_i}$ a k_i-tuple, there are functions $N_{\mathcal{F}}(\cdot)$, $k_{\mathcal{F}}(\cdot)$ and $\delta_{\mathcal{F}}(\cdot)$ such that the following is true for any $\epsilon \in (0,1)$. If a function $f : \mathbb{F}_2^n \to \{0,1\}$ with $n > N_{\mathcal{F}}(\epsilon)$ is ϵ-far from being $\mathcal{F}$-free, then f induces $\delta \cdot 2^{n(k_i-1)}$ many copies of some (E^i, σ^i), where $k_i \le k_{\mathcal{F}}(\epsilon)$ and $\delta \ge \delta_{\mathcal{F}}(\epsilon)$.*

Armed with Theorem 22, our main theorem becomes a straightforward consequence. We postpone the proof of this, because we will prove a stronger fact in Section 3.2. To start the proof of Theorem 22, let us relate pseudorandomness (uniformity) of a function to the number of solutions to a single equation induced by it. Similar and more general statements have been shown previously, but we need only the following claim for what follows.

Lemma 23 (Counting Lemma) *For every $\eta \in (0,1)$ and integer $k > 2$, there exist $\gamma = \gamma_{23}(\eta, k)$ and $\delta = \delta_{23}(\eta, k)$ such that the following is true. Suppose E is the row vector $[1\ 1 \cdots 1]$ of size k, $\sigma \in \{0,1\}^k$ is a tuple, H is a subspace of $\mathbb{F}_2^n$, and $f : \mathbb{F}_2^n \to \{0,1\}$ is a function. Furthermore, suppose there are k not necessarily distinct elements $u_1, \dots, u_k \in \mathbb{F}_2^n/H$ such that $Mu = 0$ where $u = (u_1, \dots, u_k)$, $f_H^{+u_i} : H \to \{0,1\}$ is γ-uniform for all $i \in [k]$, and $\rho(f_H^{+u_i})$ is at least η if $\sigma(i) = 1$ and at most $1 - \eta$ if $\sigma(i) = 0$ for all $i \in [k]$. Then, there are at least $\delta |H|^{k-1}$ many k-tuples $x = (x_1, x_2, \dots, x_k)$, with each $x_i \in u_i + H$, such that f induces (E, σ) at x.*

Proof: Fix $v_1 \in u_1 + H$, $v_2 \in u_2 + H, \dots, v_k \in u_k + H$ such that $v_1 + v_2 + \cdots + v_k = 0$; there exist such v_i's because $u_1 + u_2 + \cdots + u_k = 0$ in the quotient space $\mathbb{F}_2^n/H$. Define Boolean functions $f_1, \dots, f_k : H \to \{0,1\}$ so that $f_i(x) = f_H^{+v_i}(x)$ if $\sigma(i) = 1$ and $f_i(x) = 1 - f_H^{+v_i}(x)$ if $\sigma(i) = 0$. By our assumptions, $\widehat{f_i}(0) \ge \eta$ and each $|\widehat{f_i}(\alpha)| < \gamma$ for all $\alpha \neq 0$. Now, observe that, using γ-uniformity and Cauchy-Schwarz, we have:

$$
\begin{aligned}
&\mathop{\mathbb{E}}_{x_1,\dots,x_{k-1} \in H} [f_1(x_1) f_2(x_2) \cdots f_{k-1}(x_{k-1}) f_k(x_1 + x_2 + \cdots + x_{k-1})] \\
&\quad = \sum_{\alpha \in H^*} \widehat{f_1}(\alpha) \widehat{f_2}(\alpha) \cdots \widehat{f_k}(\alpha) \\
&\quad \ge \eta^k - \sum_{\alpha \neq 0} |\widehat{f_1}(\alpha) \widehat{f_2}(\alpha) \cdots \widehat{f_k}(\alpha)| \\
&\quad \ge \eta^k - \gamma^{k-2} \sqrt{\sum_{\alpha} |\widehat{f_1}(\alpha)|^2} \sqrt{\sum_{\alpha} |\widehat{f_2}(\alpha)|^2} \\
&\quad \ge \eta^k - \gamma^{k-2}
\end{aligned}
$$

Setting $\gamma = \gamma_{23}(\eta, k) \stackrel{\text{def}}{=} (\eta^k/2)^{1/(k-2)}$ makes the above expectation at least $\eta^k/2$. Now note that every $x_1, \dots, x_k \in H$ such that $x_1 + \cdots + x_k = 0$ gives $y = (y_1, \dots, y_k)$, where $y_i = v_i + x_i$ for all $i \in [k]$, such that f induces (E, σ) at y. Thus, we have from above that there are at least $\delta |H|^{k-1}$ many such y's, where $\delta = \delta_{23}(\eta, k) \stackrel{\text{def}}{=} \eta^k/2$. ■

3.1 Proof of Theorem 22

Before seeing the full technical details of the proof of Theorem 22 we proceed with a more intuitive overview.

In light of Lemma 23, our strategy will be to partition the domain into uniform cosets, using Green's regularity lemma (Lemma 12) in some fashion, and then to use the above counting lemma to count the number of induced solutions to some equation in $\mathcal{F}$. But one issue that immediately arises is that, because $\mathcal{F}$ is an infinite family of equations, we do not know the size of the equation we would want the input function to induce. Since Lemma 23 needs different uniformity parameters to count equations of different lengths, it is not *a priori* clear how to set the uniformity parameter in applying the regularity lemma. (If $\mathcal{F}$ was finite, one could set the uniformity parameter to correspond to the size of the largest equation in $\mathcal{F}$.)

To handle the infinite case, our basic approach will be to classify the input function into one of a finite set of classes. For each such class c, there will be an associated number k_c such that it is guaranteed that any function classified as c must induce an equation in $\mathcal{F}$ of size at most k_c. If there is such a classification scheme, then we know that *any* input function must induce an equation of size at most $\max_c k_c$. How do we perform this classification? We use the regularity lemma. Consider the following idealized situation. Fix an integer r. Suppose we could modify the input $f : \mathbb{F}_2^n \to \{0, 1\}$ at a small fraction of the domain to get a function $F : \mathbb{F}_2^n \to \{0, 1\}$ and then could apply Lemma 12 to get a partition of order r so that the restrictions of F to each coset was exactly 0-uniform. F is then a constant function (either 0 or 1) on each of the 2^r cosets, and so, we can classify F by a Boolean function $\mu : \mathbb{F}_2^r \to \{0, 1\}$ where $\mu(x)$ is the value of F on the coset corresponding to x. Notice that there are only finitely many such μ's. Since F differs from f at only a small fraction of the domain and since f is far from $\mathcal{F}$-free, F must also induce some equation in $\mathcal{F}$. Then, for every such μ and corresponding F, there is a smallest equation in $\mathcal{F}$ that is induced by F. We can let $\Psi_{\mathcal{F}}(r)$ be the maximum over all such μ of the size of the smallest equation in $\mathcal{F}$ that is induced by the F corresponding to μ. We then might hope that this function $\Psi_{\mathcal{F}}(\cdot)$ can be used to tune the uniformity parameter by using the functional variant of the regularity lemma (Lemma 13).

There are a couple of caveats. First, we will not be able to get the restrictions to every coset to look perfectly uniform. Second, if F induces solutions to an equation, it does not necessarily follow that f also does. To get around the first problem, we use the fact that Lemma 23 is not very restrictive on the density conditions. We think of the uniform cosets which have density neither too close to 0 nor 1 as "wildcard" cosets at which both the restriction of f and its complement behave pseudorandomly and have non-negligible density. Thus, the μ in the above paragraph will map into $\{0, 1, *\}^r$, where a '$*$' denotes a wildcard coset. For the second problem, note that it is not really a problem if $\mathcal{F}$-freeness is known to be monotone. In this case, F inducing an equation automatically means f also induces an equation, if we obtained F by removing elements from the support of f. For induced freeness properties, though, this is not the case. Using ideas from [AFKS00] and the tools from Section 2, we structure the modifications from f to F in such a way so as to force f to induce solutions of an equation if F induces a solution to the same equation. We elaborate much more on this issue during the course of the proof.

The observations described in the proof sketch above motivate the following definitions.

Definition 24 *Given function $\mu : \mathbb{F}_2^r \to \{0, 1, *\}$, a m-by-k matrix M and a k-tuple $\sigma \in \{0, 1\}^k$, suppose there exist $x_1, \ldots, x_k \in \mathbb{F}_2^r$ such that $Mx = 0$ where $x = (x_1, \ldots, x_k)$, and for every $i \in [k]$,*

$\mu(x_i)$ equals either $\sigma(i)$ or $*$. In this case, we say μ partially induces (M,σ) at *x and denote this by $(M,\sigma)\mapsto_* \mu$.*

Definition 25 *Given a positive integer r and an infinite family of systems of equations $\mathcal{F} = \{(M^1,\sigma^1),(M^2,\sigma^2),\dots\}$ with M^i being a m_i-by-k_i matrix of rank m_i and $\sigma^i \in \{0,1\}^{k_i}$ a k_i-tuple, define $\mathcal{F}_r$ to be the set of functions $\mu : \mathbb{F}_2^r \to \{0,1,*\}$ such that there exists some $(M^i,\sigma^i) \in \mathcal{F}$ with $(M^i,\sigma^i)\mapsto_* \mu$. Given $\mathcal{F}$ and integer r for which $\mathcal{F}_r \neq \emptyset$, define the following function:*

$$\Psi_{\mathcal{F}}(r) \stackrel{\text{def}}{=} \max_{\mu\in\mathcal{F}_r} \min_{\{(M^i,\sigma^i):(M^i,\sigma^i)\mapsto_*\mu\}} k_i$$

Proof of Theorem 22: Define the function $\mathcal{E}$ by setting $\mathcal{E}(0) = \epsilon/8$ and for any $r > 0$:

$$\mathcal{E}(r) = \delta_{17}(\Psi_{\mathcal{F}}(r), \gamma_{23}(\epsilon/8, \Psi_{\mathcal{F}}(r))) \cdot \min(\epsilon/8, \gamma_{23}(\epsilon/8, \Psi_{\mathcal{F}}(r)))$$

Additionally, let $T(\epsilon) = T_{16}(8/\epsilon, \mathcal{E})$, and set $N_{\mathcal{F}}(\epsilon) \stackrel{\text{def}}{=} T(\epsilon)$. Also, set $k_{\mathcal{F}}(\epsilon) \stackrel{\text{def}}{=} \Psi_{\mathcal{F}}(T(\epsilon))$ and

$$\delta_{\mathcal{F}}(\epsilon) \stackrel{\text{def}}{=} \left(\delta_{17}(\Psi_{\mathcal{F}}(r), \gamma_{23}(\epsilon/8, \Psi_{\mathcal{F}}(r))) \cdot \delta_{16}(8/\epsilon, \mathcal{E})\right)^{\Psi_{\mathcal{F}}(\epsilon)} \cdot \delta_{23}(\epsilon/8, \Psi_{\mathcal{F}}(T(\epsilon)))$$

We proceed to show that these parameter settings suffice.

Suppose we are given input function $f : \mathbb{F}_2^n \to \{0,1\}$ with $n > N_{\mathcal{F}}(\epsilon) = T_{16}(8/\epsilon, \mathcal{E})$. As mentioned in the paragraphs preceding the proof, our strategy will be to partition the domain in such a way that we can find cosets in the partition satisfying the conditions of Lemma 23. To this end, we apply Corollary 16 with $8/\epsilon$ and the function $\mathcal{E}$ as inputs. This yields subspaces $H' \le H \le \mathbb{F}_2^n$ and linear map $I : \mathbb{F}_2^n/H \to \mathbb{F}_2^n/H'$, where the order of the H-based partition, which we denote ℓ, satisfies $8/\epsilon \le \ell \le T_{16}(8/\epsilon, \mathcal{E})$. Recall that $I(u) + H'$ is contained in $u + H$ for every coset $u \in \mathbb{F}_2^n/H$. Observe that from our setting of parameters, we have that for every *nonzero* $u \in \mathbb{F}_2^n/H$, the restriction $f_{H'}^{+I(u)}$ is $(\delta_{17}(\Psi_{\mathcal{F}}(\ell), \gamma_{23}(\epsilon/8, \Psi_{\mathcal{F}}(\ell))) \cdot \gamma_{23}(\epsilon/8, \Psi_{\mathcal{F}}(\ell)))$-uniform.

But we have no such uniformity guarantee for $f_{H'}^{+0}$. This would not pose an obstacle if $\mathcal{F}$-freeness were a monotone property (i.e., if each σ^i equalled 1^{k_i}). If that were the case, we could simply make f zero on all elements of H. Since H is still only a small fraction of the domain, the modified function would still be far from $\mathcal{F}$-free, and we would be guaranteed that remaining solutions to equations of $\mathcal{F}$ induced by f would only use elements from cosets of H for which we have a guarantee about the corresponding coset of H'. But if $\mathcal{F}$-freeness is not monotone, such a scheme would not work, since it's not clear at all how to change the value of f on H so that any solution to an equation from $\mathcal{F}$ would only involve elements from nonzero shifts of H.

To resolve this issue, we further partition H' to find affine subspaces within H' on which we can guarantee that the restriction of f is uniform. The idea is that once we know that there is a solution involving H, we are going to look not at H' itself but at the smaller affine subspace within H' on which f is known to be uniform. Specifically, apply Lemma 17 to $f_{H'}^{+0}$ with input parameters $\Psi_{\mathcal{F}}(\ell)$ and $\gamma_{23}(\epsilon/8, \Psi_{\mathcal{F}}(\ell))$. This yields subspaces H'' and W, both of which contained in H', such that $|H''| \ge \delta_{17}(\Psi_{\mathcal{F}}(\ell), \gamma_{23}(\epsilon/8, \Psi_{\mathcal{F}}(\ell)))|H'|$ and $\dim(W/H'') = \Psi_{\mathcal{F}}(\ell)$. We further know that for every nonzero $v \in W/H''$, the function $f_{H''}^{+v}$ is $\gamma_{23}(\epsilon/8, \Psi_{\mathcal{F}}(\ell))$-uniform.

Now, let's "copy" W on cosets $I(u)+H'$ for every $u \in \mathbb{F}_2^n/H$. We do this by specifying[10] another linear map $J : \mathbb{F}_2^n/H \to \mathbb{F}_2^n$ so that for any $u \in \mathbb{F}_2^n/H$, the coset[11] $J(u)+W$ lies inside $I(u)+H'$ (which itself lies inside $u+H$). Each coset $J(u)+W$ also has an H''-based partition of order $\Psi_{\mathcal{F}}(\ell)$, just as W itself does. Consider $v \in \mathbb{F}_2^n/H''$ such that $v+H''$ lies inside $J(u)+W$ for some nonzero $u \in \mathbb{F}_2^n/H$. Then, because we know the uniformity of $f_{H'}^{+I(u)}$ and we have a lower bound on the size of H'', it follows from Lemma 11 that $f_{H''}^{+v}$ is $\gamma_{23}(\epsilon/8, \Psi_{\mathcal{F}}(\ell))$-uniform. Thus, for any nonzero $v \in \mathbb{F}_2^n/H''$ such that $v+H''$ lies inside $J(u)+W$ for some $u \in \mathbb{F}_2^n/H$, it is the case that $f_{H''}^{+v}$ is $\gamma_{23}(\epsilon/8, \Psi_{\mathcal{F}}(\ell))$-uniform.

In the following, we will show how to apply Lemma 23 on some of these cosets $f_{H''}^{+v}$. We have already argued their uniformity above. We now need to make sure that the pattern of their densities allow Lemma 23 to infer many induced copies of some equation in $\mathcal{F}$. To this end, we modify f to construct a new function $F : \mathbb{F}_2^n \to \{0,1\}$. F is initially identical to f on the entire domain, but is then modified in the following order:

1. For every nonzero $u \in \mathbb{F}_2^n/H$ such that $|\rho(F_H^{+u}) - \rho(F_{H'}^{+I(u)})| > \epsilon/8$, do the following. If $\rho(F_{H'}^{+I(u)}) \geq \frac{1}{2}$, then make $F(x) = 1$ on all $x \in u+H$. Otherwise, make $F(x) = 0$ on all $x \in u+H$.
2. For every nonzero $u \in \mathbb{F}_2^n/H$ such that $\rho(F_{H'}^{+I(u)}) > 1-\epsilon/4$, make $F(x) = 1$ for all $x \in u+H$. On the other hand, if $u \in \mathbb{F}_2^n/H$ is nonzero and $\rho(F_{H'}^{+I(u)}) < \epsilon/4$, make $F(x) = 0$ for all $x \in u+H$.
3. If for all nonzero $v \in W/H''$, $\rho(F_{H''}^{+v}) \geq \frac{1}{2}$, then make $F(x) = 1$ for all $x \in H$. On the other hand, if for all nonzero $v \in W/H''$, $\rho(F_{H''}^{+v}) < \frac{1}{2}$, them make $F(x) = 0$ for all $x \in H$. (One of these two conditions is true by construction.)

The following observation shows that F also must induce solutions to some equation from $\mathcal{F}$, since F is ϵ-far from being $\mathcal{F}$-free.

Claim 26 *F is ϵ-close to f.*

Proof: We count the number of elements added or removed at each step of the modification. For the first step, Corollary 16 guarantees that at most $\mathcal{E}(0) \leq \epsilon/8$ fraction of cosets $u+H$ have $|\rho(F_H^{+u}) - \rho(F_{H'}^{+I(u)})| > \epsilon/8$. So, F is modified in at most $\frac{\epsilon}{8}2^n$ locations in the first step. In the second step, if $1 > \rho(F_{H'}^{+I(u)}) > 1-\epsilon/4$, then $\rho(F_H^{+u}) > 1-3\epsilon/8$ because the first step has been completed. Similarly, if $0 < \rho(F_{H'}^{+I(u)}) < \epsilon/4$, then $\rho(F_H^{+u}) < 3\epsilon/8$. So, F is modified in at most $\frac{3\epsilon}{4}2^n$ locations in the second step. As for the third step, H contains at most $2^{n-\ell} \leq 2^{n-8/\epsilon} < \frac{\epsilon}{8}2^n$ elements for $\epsilon \in (0,1)$. So, in all, F is ϵ-close to f. ■

Now, we define a function $\mu : \mathbb{F}_2^\ell \to \{0,1,*\}$ based on F and argue that it must partially induce solutions to some equation in $\mathcal{F}$. Since H is of codimension ℓ, $\mathbb{F}_2^n/H \cong \mathbb{F}_2^\ell$ and we identify the two spaces. For $u \in \mathbb{F}_2^n/H$, if $F(x) = 1$ on the entire coset $u+H$, let $\mu(u) = 1$. On the other hand, if $F(x) = 0$ on the entire coset $u+H$, then let $\mu(u) = 0$. In any other case, let $\mu(u) = *$.

[10] One way to accomplish this is to define J appropriately for ℓ linearly independent elements of $\mathbb{F}_2^n/H$ and then use linearity to define it on all of $\mathbb{F}_2^n/H$.

[11] Note that the image of J is to elements of $\mathbb{F}_2^n$ and not $\mathbb{F}_2^n/W$, even though we think of the output as denoting a coset of W. The reason is that we will find it convenient to fix the shift and not make it modulo W.

Claim 27 *There exists* $(E^i, \sigma^i) \in \mathcal{F}$ *such that* $(E^i, \sigma^i) \mapsto_* \mu$.

Proof: As already observed, F is not $\mathcal{F}$-free, and let $(E^i, \sigma^i) \in \mathcal{F}$ be some equation whose solution is induced by F at $(x_1, \ldots, x_{k_i}) \in (\mathbb{F}_2^n)^{k_i}$. Now let $y = (y_1, \ldots, y_{k_i}) \in (\mathbb{F}_2^\ell)^{k_i}$ where for each $j \in [k_i]$, $y_j = x_j \pmod H$. It's clear that $E^i y = 0$. To argue that F partially induces μ at y, suppose for contradiction that for some $j \in [k_i]$, $\mu(y_j) = 0$ but $\sigma^i_j = 1$. But if $\mu(y_j) = 0$, then F is the constant function 0 on all of $y_j + H$, contradicting the existence of $x_j \in y_j + H$ with $F(x) = 1$. We get a similar contradiction if $\mu(y_j) = 1$ but $\sigma^i_j = 0$. ■

Using Definition 25, we immediately get that there is some $(E^i, \sigma^i) \in \mathcal{F}$ of size at most $\Psi_{\mathcal{F}}(\ell)$ such that $(E^i, \sigma^i) \mapsto_* \mu$. Fix $x_1, \ldots, x_{k_i} \in \mathbb{F}_2^n$ where F induces (E^i, σ^i), and as in the above proof, let $y_1, \ldots, y_{k_i} \in \mathbb{F}_2^n/H$ where each $y_j = x_j \pmod H$. Also, pick $k_i - 1$ linearly independent elements $\tilde{v}_1, \ldots, \tilde{v}_{k_i-1}$ from W/H'', which is possible since $\dim(W/H'') = \Psi_{\mathcal{F}}(\ell) > k_i - 1$, and choose $v_1 \in \tilde{v}_1 + H'', \ldots, v_{k_i-1} \in \tilde{v}_{k_i-1} + H''$ such that $v_1, \ldots, v_{k_i}$ are linearly independent. Additionally set $v_{k_i} = \sum_{j=1}^{k_i-1} v_j$. Notice that none of $v_1, \ldots, v_{k_i}$ are in H''. Now, consider the sets $f_{H''}^{+J(y_1)+v_1}, f_{H''}^{+J(y_2)+v_2}, \ldots, f_{H''}^{+J(y_{k_i})+v_{k_i}}$. (Notice these are restrictions of f, not F!) We will show that these sets respect the density and uniformity conditions for Lemma 23 to apply.

As for uniformity, we have already argued that each of these sets is $\gamma_{23}(\epsilon/8, \Psi_{\mathcal{F}}(\ell))$-uniform, since $J(y_j) + v_j$ is not in H'' for every $j \in [k_i]$. For density, we argue as follows. For every $j \in [k_i]$, there are three cases: $\mu(y_j) = 1$, $\mu(y_j) = 0$, and $\mu(y_j) = *$. Consider the first case. If $y_j + H$ was affected by the first modification from f to F, then, $\rho(f_{H'}^{+I(y_j)}) \geq \frac{1}{2}$, and using the $\mathcal{E}(\ell)$-uniformity of $f_{H'}^{+I(y_j)}$ along with Lemma 11, we get that $\rho(f_{H''}^{+J(y_j)+v_j}) \geq \frac{1}{2} - \mathcal{E}(\ell) \cdot \delta_{17}^{-1}(\Psi_{\mathcal{F}}(r), \gamma_{23}(\epsilon/8, \Psi_{\mathcal{F}}(r))) \geq \frac{1}{2} - \frac{\epsilon}{8} \geq \frac{\epsilon}{8}$. If $y_j + H$ was affected by the second modification, then, by the same argument, we get that $\rho(f_{H''}^{+J(y_j)+v_j}) \geq 1 - \frac{\epsilon}{4} - \frac{\epsilon}{8} \geq \frac{\epsilon}{8}$. Else, if $y_j + H$ was affected by the third modification from S to S', we are automatically guaranteed that $\rho(f_{H''}^{+J(y_j)+v_j}) \geq \frac{1}{2}$ since $J(y_j) + v_j \notin H''$. The case $\mu(y_j) = 0$ is similar, and the analysis shows that $\rho(f_{H''}^{+J(y_j)+v_j}) \geq 1 - \frac{\epsilon}{8}$. Finally, consider the "wildcard" case, $\mu(y_j) = *$. This case arises only if $y_j \neq 0$ and $\epsilon/4 \leq \rho(f_{H'}^{+I(y_j)}) \leq 1 - \epsilon/4$. Again using $\mathcal{E}(\ell)$-uniformity of $f_{H'}^{+I(y_j)}$ along with Lemma 11, we get that $\epsilon/8 \leq \rho(f_{H''}^{+J(y_j)+v_j}) \leq 1 - \epsilon/8$.

Thus, we can apply Lemma 23 with $\epsilon/8$ and $\Psi_{\mathcal{F}}(\ell)$ as the parameters to get that there are at least $\delta_{23}(\epsilon/8, \Psi_{\mathcal{F}}(\ell))|H''|^{k_i-1}$ tuples $z = (z_1, \ldots, z_{k_i})$ with each $z_j \in J(y_j) + v_j + H''$ at which (E^i, σ^i) is induced . Finally, each such $z_1, \ldots, z_{k_i}$ leads to a distinct $z' = (z'_1, \ldots, z'_{k_i}) \in (\mathbb{F}_2^n)^{k_i}$ at which (E^i, σ^i) is induced by f, by setting each z'_j to $J(y_j) + v_j + z_j$ and observing that $\sum_{j=1}^{k_i} J(y_j) + v_j = J\left(\sum_{j=1}^{k_i} y_j\right) + \sum_{j=1}^{k_i} v_j = 0$. This completes the proof of Theorem 22. ■

3.2 Extending to Systems of Equations of Complexity 1

As mentioned in the introduction, the result we actually prove is stronger than Theorem 3. To describe the full set of properties for which we can show testability, we first need to make the following definition.

Definition 28 (Complexity of linear system [GT08]) *An* $m \times k$ *matrix* M *over* $\mathbb{F}_2$ *is said to be of* (Cauchy-Schwarz) complexity c*, if* c *is the smallest positive integer for which the following is true. For every* $i \in [k]$*, there exists a partition of* $[k] \backslash \{i\}$ *into* $c+1$ *subsets* $S_1, \cdots, S_{c+1}$ *such that* *for every* $j \in [c+1]$, $\left(\mathbf{e}_i + \sum_{i' \in S_j} \mathbf{e}_{i'}\right) \notin \mathsf{rowspace}(M)$, *where* $\mathsf{rowspace}(M)$ *is the linear subspace of* $\mathbb{F}_2^k$ *spanned by the rows of* M.

In other words, if we view the rowspace of the matrix M as specifying a collection of linear dependencies on k variables $x_1, \ldots, x_k$, then M has complexity c if for every variable x_i, the rest of the variables $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_k$ can be partitioned into $c+1$ sets $S_1, \ldots, S_{c+1}$ such that x_i is not linearly dependent on the variables of any single S_j. Let us make a few remarks to illustrate the definition. Green and Tao show (Lemma 1.6 in [GT08]) that if each of these linear dependencies involves more than two variables, then the complexity of M is at most $\mathsf{rank}(M) = m$. In particular then, if M has one row and is nonzero on more than two coordinates, M has complexity 1. This is the setting we discussed in the introduction. We slightly extend this observation in the claim below. Before we state it, we observe that in the context of property testing, it is only natural to exclude matrices which yield linear dependencies involving less than three variables. If the rowspace of the matrix M contains a vector which is nonzero at only one coordinate i, then for any string σ of length k, the property of (M, σ)-freeness must contain all functions f such that $f(0) = 1 - \sigma_i$, and so *every* function is exponentially close to such a property. Similarly, if $\mathsf{rowspace}(M)$ contains a vector nonzero only at two coordinates i and j, then for any $\sigma \in \{0,1\}^k$, either (M, σ)-freeness is trivial (if $\sigma_i \neq \sigma_j$) or it is equivalent to (M', σ')-freeness where σ' is the string obtained by removing coordinate j and M' is the matrix obtained by removing column j, adding 1 (mod 2) to every element in column i and row-reducing the resulting matrix.

Claim 29 *If* $M \in \mathbb{F}_2^{m \times k}$ *is a matrix with two rows such that every vector in its rowspace has at least three nonzero coordinates, then* M *has complexity* 1.

Proof: Let $R_1 \subseteq [k]$ be the set of coordinates for which the first row is nonzero, and $R_2 \subseteq [k]$ those for which the second row is nonzero. We can assume that $R_1 \not\subseteq R_2$ and $R_2 \not\subseteq R_1$, because if, say, $R_1 \subseteq R_2$, we could replace the second row by the sum of the first and second, making R_1 and R_2 disjoint but preserving the rowspace of the matrix. Also, we we can assume w.l.o.g. that $R_1 \cup R_2 = [k]$.

Fix $i \in [k]$. We want to show a partition of $[k] \backslash \{i\}$ into sets S_1, S_2 such that $\mathbf{e}_i + \sum_{i' \in S_1} \mathbf{e}_{i'} \notin \mathsf{rowspace}(M)$ and similarly for S_2. If $i \in R_1 \backslash R_2$, let S_1 consist of two elements, one from $R_2 \backslash R_1$ and one from $R_1 \backslash \{i\}$, and let S_2 be the rest. If $i \in R_2 \backslash R_1$, let S_1 consist of one element from $R_1 \backslash R_2$ and one from $R_2 \backslash \{i\}$, and let S_2 be the rest. And finally, if $i \in R_1 \cap R_2$, let S_1 consist of one element from $R_1 \backslash R_2$ and one from $R_2 \backslash R_1$, and let S_2 be the rest. It is straightforward to check that the definition of complexity 1 is satisfied by these choices. ■

More generally, an infinitely large class of complexity 1 linear systems is generated by *graphic matroids*. We refer the reader to [BCSX09] for definition and details. That this class contains the class of matrices proved to be of complexity 1 in Claim 29 is easy to show. We proved the claim separately above only to be self-contained without introducing matroid notation. One final remark is that if M is the matrix in the characterization of Reed-Muller codes of order d from Appendix A, then M has complexity exactly d; see Example 3 of [GT08].

Our main result in this section is the extension of Theorem 3 to complexity 1 systems of equations.

Theorem 30 *Let $\mathcal{F} = \{(M^1, \sigma^1), (M^2, \sigma^2), \dots\}$ be a possibly infinite set of induced systems of equations, with each M^i of complexity 1. Then, the property of being $\mathcal{F}$-free is testable with one-sided error.*

We next describe how to modify the previous proof to the new setting. The following analog to Theorem 22 is the core of the proof of Theorem 30.

Theorem 31 *For every infinite family $\mathcal{F} = \{(M^1, \sigma^1), (M^2, \sigma^2), \dots, (M^i, \sigma^i), \dots\}$, where each M^i is a $m_i \times k_i$ matrix over $\mathbb{F}_2$ of complexity 1, there are functions $N_{\mathcal{F}}(\cdot)$, $k_{\mathcal{F}}(\cdot)$ and $\delta_{\mathcal{F}}(\cdot)$ such that the following is true for any $\epsilon \in (0,1)$. If a function $f : \mathbb{F}_2^n \to \{0,1\}$ with $n > N_{\mathcal{F}}(\epsilon)$ is ϵ-far from being $\mathcal{F}$-free, then f induces $\delta \cdot 2^{n(k_i - m_i)}$ many copies of some (M^i, σ^i), where $k_i \leq k_{\mathcal{F}}(\epsilon)$ and $\delta \geq \delta_{\mathcal{F}}(\epsilon)$.*

We show how to deduce Theorem 30 from Theorem 31 next. Note that, as promised earlier, this is also a proof of Theorem 3 assuming Theorem 22.

Proof of Theorem 30: Theorem 31 allows us to devise the following tester T for $\mathcal{F}$-freeness. T, given input $f : \mathbb{F}_2^n \to \{0,1\}$, first checks if $n \leq N_{\mathcal{F}}(\epsilon)$, and in this case, it queries f on the entire domain and decides accordingly. Otherwise, T selects independently and uniformly at random a set D of d elements from $\mathbb{F}_2^n$, where we will specify d at the end of the argument. It then queries all points in the linear subspace spanned by the elements of D and then accepts or rejects based on whether f restricted to this subspace is $\mathcal{F}$-free or not.

Clearly, if f is $\mathcal{F}$-free, then the tester always accepts because the property is subspace-hereditary. Also, if $n \leq N_{\mathcal{F}}(\epsilon)$, then the correctness of the algorithm is trivial. So, suppose f is ϵ-far from $\mathcal{F}$-free and $n > N_{\mathcal{F}}(\epsilon)$. For the M^i guaranteed to exist from Theorem 31, let K be a $k_i \times c$ matrix over $\mathbb{F}_2$, where $c = k_i - m_i \leq k_{\mathcal{F}}(\epsilon)$, such that the columns of K form a basis for the kernel of M^i. Then, every $y = (y_1, \dots, y_c) \in (\mathbb{F}_2^n)^c$ yields a distinct vector $x = (x_1, \dots, x_k) \in (\mathbb{F}_2^n)^k$ formed by letting $x = Ky$ that satisfies $M^i x = M^i K y = 0$. Therefore, because of Theorem 31, the probability that uniformly chosen $y_1, \cdots, y_c \in \mathbb{F}_2^n$ yield $x = (x_1, \dots, x_k)$ such that f induces (M^i, σ^i) at x is at least $\delta_{\mathcal{F}}(\epsilon)$. The probability that D does not contain such $y_1, \dots, y_c$ is at most $(1-\delta)^{d/c} < e^{\delta_{\mathcal{F}}(\epsilon) d/c} < 1/3$ if we choose $d = O(c/\delta_{\mathcal{F}}(\epsilon)) = O(k_{\mathcal{F}}(\epsilon)/\delta_{\mathcal{F}}(\epsilon))$. Thus with probability at least $2/3$, $\mathsf{span}(D)$ contains $x_1, \dots, x_k$ such that f induces (M^i, σ^i) at $x = (x_1, \dots, x_k)$, making the tester reject. ■

To prove Theorem 31, the main ingredient that changes is the counting lemma.

Lemma 32 (Counting Lemma for Complexity 1**)** *For every $\eta \in (0,1)$ and integer $k > 2$, there exist $\gamma = \gamma_{23}(\eta, k)$ and $\delta = \delta_{23}(\eta, k)$ such that the following is true. Suppose M is an $m \times k$ matrix of complexity 1 and rank $m < k$, $\sigma \in \{0,1\}^k$ is a tuple, H is a subspace of $\mathbb{F}_2^n$, and $f : \mathbb{F}_2^n \to \{0,1\}$ is a function. Furthermore, suppose there are k not necessarily distinct elements $u_1, \dots, u_k \in \mathbb{F}_2^n / H$ such that $Mu = 0$ where $u = (u_1, \dots, u_k)$, $f_H^{+u_i} : H \to \{0,1\}$ is γ-uniform for all $i \in [k]$, and $\rho(f_H^{+u_i})$ is at least η if $\sigma(i) = 1$ and at most $1 - \eta$ if $\sigma(i) = 0$ for all $i \in [k]$. Then, there are at least $\delta |H|^{k-m}$ many k-tuples $x = (x_1, x_2, \dots, x_k)$, with each $x_i \in u_i + H$, such that f induces (M, σ) at x.*

Lemma 32 is a special case of the Generalized von Neumann Theorem (Proposition 7.1 in [GT08]). The rest of the proof is a straightforward modification of Section 3.1. Namely, whenever the old proof requires k elements or k cosets $x_1, \dots, x_k$ to satisfy the equation $x_1 + \cdots + x_k = 0$, the new proof would require that they satisfy the equation $Mx = 0$ where $x = (x_1, \dots, x_k)$.

4 Characterization of natural one-sided testable properties

We now turn to showing Theorem 10 which states that for linear-invariant properties, testability with a one-sided error oblivious tester is equivalent to the property being semi subspace-hereditary (recall here Definition 9).

First we formalize the discussion from the introduction regarding the fact that it is always possible to assume that the testing algorithm for a one-sided testable linear-invariant property makes its decision only by querying the input function on a random linear subspace of constant dimension.

Proposition 33 *Let $\mathcal{P}$ be a linear invariant property, and let T be an arbitrary one-sided tester for $\mathcal{P}$ with query complexity $d(\epsilon, n)$. Then, there exists a one-sided tester T' for $\mathcal{P}$ that selects a random subspace H of dimension $d(\epsilon, n)$, queries the input on all points of H, and decides based on the oracle answers, the value of ϵ and n, and internal randomness*[12]*. Note that T' is non-adaptive and has query complexity $2^{d(\epsilon,n)}$.*

Proof: Consider a tester T_2 that acts as follows. If the tester T on the input makes queries $x_1, \ldots, x_d$, then T' queries all points in $\text{span}(x_1, \ldots, x_d)$ but makes its decision based on $x_1, \ldots, x_d$ just as T does. Clearly, T_2 is also a one-sided tester for $\mathcal{P}$ and with query complexity at most $2^{d(\epsilon)}$.

Now, define a tester T' as follows. Given oracle access to a function $f : \mathbb{F}_2^n \to \{0, 1\}$, T' first selects uniformly at random a non-singular linear transformation $L : \mathbb{F}_2^n \to \mathbb{F}_2^n$, and then invokes T_2 providing it with oracle access to the function $f \circ L$. That is, when T_2 makes query x, then algorithm T' makes query $L(x)$. We argue that the sequence of queries made by T' are the elements of a uniformly chosen random subspace of dimension at most $d(\epsilon)$. To see this, fix the input f and the randomness of T_2. Then, for each $i \in [2^{d(\epsilon)}]$ for which the i'th query, x_i, made by T_2 is linearly independent of the previous $i-1$ queries, $x_1, \ldots, x_{i-1}$, it's the case that $L(x_i)$ is a uniformly chosen random element from outside $\text{span}(L(x_1), \ldots, L(x_{i-1}))$. So, for every fixing of the random coins of T_2, the queries made by T' span a uniformly chosen subspace of dimension at most $d(\epsilon)$, and hence, this is also the case when the coins are not fixed. T' is a one-sided tester for $\mathcal{P}$ because if $f \in \mathcal{P}$, then $f \circ L \in \mathcal{P}$ by linear invariance, and if f is ϵ-far from $\mathcal{F}$, then $f \circ L$ is also ϵ-far from $\mathcal{P}$ because L is a permutation on $\mathbb{F}_2^n$. ■

An oblivious tester, as defined in Definition 8, differs from the tester T' of the above proposition in that the dimension of the selected subspace and the decision made by the tester are not allowed to depend on n. As argued there, it is very reasonable to expect natural linear-invariant properties to have such testers, and indeed, prior works have already implicitly restricted themselves in this way.

We can now proceed with the proof of Theorem 10.

Proof of Theorem 10: Let us first prove the forward direction of the theorem. Note that for this direction, we do not need to assume the truth of Conjecture 4. Given a linear-invariant property $\mathcal{P}$ that can be tested with one-sided error by an oblivious tester, we will build a subspace-hereditary property $\mathcal{H}$ containing $\mathcal{P}$, by identifying a (possibly infinite) collection of matrices M^i and binary strings σ^i such that $\mathcal{H}$ is equivalent to the property of being $\{(M^i, \sigma^i)\}_i$- free.

[12] Note here, we leave open the possibility that the decision of the tester may not be based only on properties of the selected subspace. This gap can be resolved using the same techniques as used by [GT03] for the graph case, but this point is not relevant for our purposes and so we do not elaborate more here.

Let $\mathcal{S}$ consist of the pairs (H, S), where H is a subspace of $\mathbb{F}_2^n$ and $S \subseteq H$ is a subset, that satisfy the following two properties: (1) $\dim(H) = d(\epsilon)$ for some ϵ, and (2) if for this ϵ, the tester rejects its input with some positive probability when the evaluation of its input on the sampled subspace is $\mathbf{1}_S$. For $(H, S) \in \mathcal{S}$ let $d = \dim(H)$. Consider the matrix A_H over $\mathbb{F}_2$ with each row representing an element of H in some fixed basis. Notice that A_H is a $(2^\ell \times \ell)$-sized matrix. Define M_H, a matrix over $\mathbb{F}_2$ of size $(2^\ell - \ell) \times 2^\ell$, such that $M_H A_H = 0$. Finally, for each $i \in [2^\ell]$ define $\sigma_S(i) = \mathbf{1}_S(x_i)$, where x_i is the element represented in the i'th row of A_H. Let $\mathcal{M}$ be the set of pairs (M_H, σ_S) obtained in this way from every $(H, S) \in \mathcal{S}$.

We now proceed to verify that $\mathcal{H}$ satisfies the conditions of Definition 9. To show that $\mathcal{P}$ is $\mathcal{M}$-free, let $f \in \mathcal{P}_n$, and suppose that there exists $(M_H, \sigma_S) \in \mathcal{M}$ such that $(M_H, \sigma_S) \mapsto f$, for some ϵ, and for some H with $\dim(H) = d(\epsilon)$ and $S \subseteq H$. We show that f is rejected with some positive probability, a contradiction to the fact that the test is one-sided. If (M_H, σ_S) is induced by f at $(x_1, \ldots, x_{2^{d(\epsilon)}})$, then these elements necessarily span a $d(\epsilon)$-dimensional subspace so that the function restricted to that subspace is $\mathbf{1}_S \circ L$ for some linear transformation $L : \mathbb{F}_2^n \to \mathbb{F}_2^{d(\epsilon)}$ (determined by the choice of basis that was used to represent H). Thus, this immediately implies by the definition of (M_H, σ_S) that the tester rejects f with positive probability.

To verify the second part of the Definition 9, let $M(\epsilon) = d(\epsilon)$. Suppose $f : \mathbb{F}_2^n \to \{0, 1\}$, with $n > M(\epsilon)$ is ϵ-far from satisfying $\mathcal{P}$. In this case, in order for the tester to reject f with positive probability, it must select a $d(\epsilon)$-dimensional subspace H so that the restriction to H equals the indicator function on S (upto a linear transformation), for some $(H, S) \in \mathcal{S}$. Therefore T is not $\mathcal{M}$-free, and thus $T \notin \mathcal{H}$.

It remains to show the opposite direction of Theorem 10. We here assume Conjecture 4 that every subspace-hereditary property $\mathcal{P}$ is testable by a one-sided tester. Our first observation that, in this case, it is actually testable by an *oblivious* one-sided tester. Namely, we show that the clearly oblivious tester, which checks whether the input function restricted to a random linear subspace satisfies $\mathcal{P}$ or not, is a valid tester. We need to argue that if a non-oblivious tester rejects input f that is ϵ-far from $\mathcal{P}$ by querying its values on a random $d(\epsilon)$-dimensional subspace (we already know the tester is of this type from Proposition 33), then with high probability, the input function restricted to a random $3d(\epsilon)$-dimensional subspace does not satisfy the property $\mathcal{P}$. Suppose it did. But then, if the original tester first uniformly selected a $3d(\epsilon)$-dimensional subspace H and then uniformly selected a $d(\epsilon)$-dimension subspace H' inside it, and ran its decision based on $f|_{H'}$, it will accept the input with large probability, which is a contradiction to the soundness of the tester since H'' is a uniformly distributed $d(\epsilon)$-dimensional subspace. Thus, for a testable subspace-hereditary property, we can assume that the tester simply checks for $\mathcal{P}$ on the sampled subspace, and is hence, oblivious to the value of n. This argument is analogous to one of Alon for graph properties, reported in [GT03].

Now, assuming that every subspace-hereditary property is testable by an oblivious one-sided tester (Conjecture 4), we wish to show that every semi subspace-hereditary property is testable by an oblivious one-sided tester. Let $\mathcal{P}$ be a a semi subspace-hereditary property and let $\mathcal{H}$ be the subspace-hereditary property associated to $\mathcal{P}$ in Definition 9. By our assumption, $\mathcal{H}$ has a one-sided tester T', which on input ϵ makes $Q'(\epsilon)$ queries and rejects inputs ϵ-far from $\mathcal{H}$ with probability $2/3$. The tester T for $\mathcal{P}$ makes $Q(\epsilon) = \max(Q'(\epsilon/2), 2^{M(\epsilon/2)})$ queries (where $M(\cdot)$ comes from Definition 9) and proceeds as follows. If the size of the input is at most $Q(\epsilon)$, then by definition, T receives the evaluation of the function all of the input and in this case, it simply checks if the input belongs to $\mathcal{P}$. Otherwise T emulates T' with distance parameter $\epsilon/2$ and accepts if and only if T'

accepts.

Notice that T is one-sided. Indeed, if the input f satisfies $\mathcal{P}$ then $f \in \mathcal{H}$ and thus T' always accepts, causing T to always accept. To prove soundness, we first argue that if f is ϵ-far from $\mathcal{P}$ then it is $\epsilon/2$-far from $\mathcal{H}$. Suppose otherwise, and modify f in at most an $\epsilon/2$ fraction of the domain in order to obtain a function $g \in \mathcal{H}$. Thus g is still $\epsilon/2$-far from $\mathcal{P}$, and by Definition 9 $g \notin \mathcal{H}$, a contradiction. Finally, since f is $\epsilon/2$-far from $\mathcal{H}$ and since T' mistakenly accepts such inputs with probability at most $1/3$ so does T'. ■

5 Concluding Remarks and Open Problems

Obviously, the main open problem we would like to see resolved is Conjecture 4. One appealing way to prove the conjecture would be to proceed as we have but to obtain a stronger notion of pseudorandomness in the regularity lemma. The notion of ϵ-uniformity obtained from Green's regularity lemma corresponds to the Gowers U^2 norm, whereas in order to be able to prove Conjecture 4 in its full generality, we would presumably need a similar regularity lemma with respect to the Gowers U^k norm [Gow01] for any fixed k. Such a higher order regularity lemma has been very recently obtained by Green and Tao [GT10] over the integers and over fields of large characteristic. However, it is not yet available over $\mathbb{F}_2$, as the inverse conjectures for the Gowers norms over $\mathbb{F}_2$ have not yet been completely clarified [Gre10].

Let us mention some other observations and open problems related to this work.

- As we have mentioned in Subsection 1.4, it is not too hard to construct linear-invariant properties which are not testable. Actually, there are properties of this type that cannot be tested with $o(2^n)$ queries. One example can be obtained from a variant of an argument used in [GGR98] as follows; it is shown in [GGR98] (see Proposition 4.1) that for every n there exists a property of Boolean functions that contains $2^{\frac{1}{10}2^n}$ of the Boolean functions over $\mathbb{F}_2^n$ and cannot be tested with less than $\frac{1}{20}2^n$ queries. This family of functions is not necessarily linear invariant, so we just "close" it under linear transformation, by adding to the property all the linear-transformed such functions. Since the number of these linear transformation is bounded by 2^{n^2} (corresponding to all possible $n \times n$ matrices over $\mathbb{F}_2$) we get that the new property contains at most $2^{n^2}2^{\frac{1}{10}2^n} \leq 2^{\frac{1}{5}2^n}$ Boolean functions. One can verify that since this new family contains a small fraction of all possible functions the argument of [GGR98] caries over, and the new property cannot be tested with $o(2^n)$ queries.

- The upper bound one obtains from the general result given in Theorem 3 is terrible in terms of its dependence on $1/\epsilon$. A natural open problem would be to find a characterization of these properties that can be tested with a number of queries that depends polynomially on ϵ. This, however, seems to be a very hard problem. Even if the only forbidden equation is $x + y = z$ it is not known if such an efficient test exists. This question was raised by Green [Gre05]; see [BX10] for current best bounds.

- Our result here gives a (conjectured) characterization of the linear-invariant properties of Boolean functions that can be tested with one-sided error. It is of course natural to try to extend our framework to other families of properties, characterized by other or more general invariances. For instance, can we carry out a full characterization for testable affine invariant properties of Boolean functions on the hypercube?

- It would be valuable to understand formally why the technology developed for handling graph properties can be extended so naturally to linear-invariant properties. This "coincidence" seems part of a larger trend in mathematics where claims about subsets find analogs in claims about vector subspaces. See [Coh04] for an interesting attempt to shed light on this puzzle.

Acknowledgements Arnab would like to thank Eldar Fischer for some initial stimulating discussions during a visit to the Technion and Alex Samorodnitsky for constant encouragement and advice. Many thanks also to Noga Alon and Shachar Lovett for useful suggestions.

References

[AFKS00] Noga Alon, Eldar Fischer, Michael Krivelevich, and Mario Szegedy. Efficient testing of large graphs. *Combinatorica*, 20(4):451–476, 2000.

[AFNS06] Noga Alon, Eldar Fischer, Ilan Newman, and Asaf Shapira. A combinatorial characterization of the testable graph properties: it's all about regularity. In *Proc. 36th Annual ACM Symposium on the Theory of Computing*, pages 251–260, 2006.

[AKK+05] Noga Alon, Tali Kaufman, Michael Krivelevich, Simon Litsyn, and Dana Ron. Testing Reed-Muller codes. *IEEE Transactions on Information Theory*, 51(11):4032–4039, 2005.

[AS08a] Noga Alon and Asaf Shapira. A characterization of the (natural) graph properties testable with one-sided error. *SIAM J. on Comput.*, 37(6):1703–1727, 2008.

[AS08b] Noga Alon and Asaf Shapira. A separation theorem in property testing. *Combinatorica*, 28:261–281, 2008.

[AT08] Tim Austin and Terence Tao. On the testability and repair of hereditary hypergraph properties. *Random Structures and Algorithms (to appear)*, 2008. Preprint available at http://arxiv.org/abs/0801.2179.

[BCL+06] Christian Borgs, Jennifer T. Chayes, László Lovász, Vera T. Sós, Balázs Szegedy, and Katalin Vesztergombi. Graph limits and parameter testing. In *Proc. 36th Annual ACM Symposium on the Theory of Computing*, pages 261–270, 2006.

[BCSX09] Arnab Bhattacharyya, Victor Chen, Madhu Sudan, and Ning Xie. Testing linear-invariant non-linear properties. In *STACS*, pages 135–146, 2009. Full version at http://www.eccc.uni-trier.de/report/2008/088/.

[BFL91] László Babai, Lance Fortnow, and Carsten Lund. Non-deterministic exponential time has two-prover interactive protocols. *Computational Complexity*, 1(1):3–40, 1991.

[BHR05] Eli Ben-Sasson, Prahladh Harsha, and Sofya Raskhodnikova. Some 3cnf properties are hard to test. *SIAM J. on Comput.*, 35(1):1–21, 2005.

[BKS+09] Arnab Bhattacharyya, Swastik Kopparty, Grant Schoenebeck, Madhu Sudan, and David Zuckerman. Optimal testing of Reed-Muller codes. *Electronic Colloquium in Computational Complexity*, TR09-086, October 2009.

[Bla09] Eric Blais. Testing juntas nearly optimally. In *Proc. 41st Annual ACM Symposium on the Theory of Computing*, pages 151–158, 2009.

[BLR93] Manuel Blum, Michael Luby, and Ronitt Rubinfeld. Self-testing/correcting with applications to numerical problems. *J. Comp. Sys. Sci.*, 47:549–595, 1993. Earlier version in STOC'90.

[BO10] Eric Blais and Ryan O'Donnell. Lower bounds for testing function isomorphism. In *Proc. 25th Annual IEEE Conference on Computational Complexity (to appear)*, 2010.

[BS09] Eli Ben-Sasson and Madhu Sudan. Limits on the rate of locally testable affine-invariant codes. November 2009. Manuscript.

[BX10] Arnab Bhattacharyya and Ning Xie. Lower bounds for testing triangle-freeness in boolean functions. In *Proc. 21st ACM-SIAM Symposium on Discrete Algorithms*, pages 87–98, Philadelphia, PA, USA, 2010. Society for Industrial and Applied Mathematics.

[Coh04] Henry Cohn. Projective geometry over $\mathbb{F}_1$ and the Gaussian binomial coefficients. *American Mathematical Monthly*, 111:487–495, 2004.

[CRSW83] C. Chvatál, V. Rödl, E. Szemerédi, and W. T. Trotter Jr. The Ramsey number of a graph with bounded maximum degree. *Journal of Combinatorial Theory, Series B*, 34(3):239–243, 1983.

[DLM+07] Ilias Diakonikolas, Homin K. Lee, Kevin Matulef, Krzysztof Onak, Ronitt Rubinfeld, Rocco A. Servedio, and Andrew Wan. Testing for concise representations. In *Proc. 48th Annual IEEE Symposium on Foundations of Computer Science*, pages 549–558, 2007.

[Fis04] Eldar Fischer. The art of uninformed decisions: A primer to property testing. In G. Paun, G. Rozenberg, and A. Salomaa, editors, *Current Trends in Theoretical Computer Science: The Challenge of the New Century*, volume 1, pages 229–264. World Scientific Publishing, 2004.

[Fis05] Eldar Fischer. The difficulty of testing for isomorphism against a graph that is given in advance. *SIAM J. on Comput.*, 34(5):1147–1158, 2005.

[FKR+04] Eldar Fischer, Guy Kindler, Dana Ron, Shmuel Safra, and Alex Samorodnitsky. Testing juntas. *J. Comp. Sys. Sci.*, 68(4):753–787, 2004.

[GGR98] Oded Goldreich, Shafi Goldwasser, and Dana Ron. Property testing and its connection to learning and approximation. *Journal of the ACM*, 45:653–750, 1998.

[GKS08] Elena Grigorescu, Tali Kaufman, and Madhu Sudan. 2-transitivity is insufficient for local testability. In *IEEE Conference on Computational Complexity*, pages 259–267, 2008.

[GKS09] Elena Grigorescu, Tali Kaufman, and Madhu Sudan. Succinct representation of codes with applications to testing. In *APPROX-RANDOM*, pages 534–547, 2009.

[GOS+09] Parikshit Gopalan, Ryan O'Donnell, Rocco A. Servedio, Amir Shpilka, and Karl Wimmer. Testing Fourier dimensionality and sparsity. In *ICALP (1)*, pages 500–512, 2009.

[Gow01] William T. Gowers. A new proof of Szemḗredi's theorem. *Geometric Functional Analysis*, 11(3):465–588, 2001.

[Gre05] Ben Green. A Szemerédi-type regularity lemma in abelian groups. *Geometric and Functional Analysis*, 15(2):340–376, 2005.

[Gre10] Ben Green. Personal communication, February 2010.

[GT03] Oded Goldreich and Luca Trevisan. Three theorems regarding testing graph properties. *Random Structures and Algorithms*, 23(1):23–57, 2003.

[GT08] Ben Green and Terence Tao. Linear equations in primes. Preprint available at http://arxiv.org/abs/math/0606088v2, April 2008.

[GT10] Ben Green and Terence Tao. An arithmetic regularity lemma, associated counting lemma, and applications. Preprint available at http://arxiv.org/abs/1002.2028, February 2010.

[KL05] Tali Kaufman and Simon Litsyn. Almost orthogonal linear codes are locally testable. *Proc. 46th Annual IEEE Symposium on Foundations of Computer Science*, pages 317–326, 2005.

[KS07] Tali Kaufman and Madhu Sudan. Sparse random linear codes are locally decodable and testable. In *Proc. 48th Annual IEEE Symposium on Foundations of Computer Science*, pages 590–600, 2007.

[KS08] Tali Kaufman and Madhu Sudan. Algebraic property testing: the role of invariance. In *Proc. 40th Annual ACM Symposium on the Theory of Computing*, pages 403–412, New York, NY, USA, 2008. ACM.

[KS09] Swastik Kopparty and Shubhangi Saraf. Tolerant linearity testing and locally testable codes. In *APPROX-RANDOM*, pages 601–614, 2009.

[KSV08] Daniel Král', Oriol Serra, and Lluís Vena. A removal lemma for systems of linear equations over finite fields. *Israel Journal of Mathematics (to appear)*, 2008. Preprint available at http://arxiv.org/abs/0809.1846.

[KSV09] Daniel Král', Oriol Serra, and Lluís Vena. A combinatorial proof of the removal lemma for groups. *Journal of Combinatorial Theory*, 116(4):971–978, May 2009.

[MORS09] Kevin Matulef, Ryan O'Donnell, Ronitt Rubinfeld, and Rocco A. Servedio. Testing halfspaces. In *Proc. 20th ACM-SIAM Symposium on Discrete Algorithms*, pages 256–264, 2009.

[PRS02] Michal Parnas, Dana Ron, and Alex Samorodnitsky. Testing basic boolean formulae. *SIAM J. Discrete Math.*, 16(1):20–46, 2002.

[Ron08] Dana Ron. Property Testing: A Learning Theory Perspective. In *Foundations and Trends in Machine Learning*, volume 1, pages 307–402. 2008.

[RS96] Ronitt Rubinfeld and Madhu Sudan. Robust characterizations of polynomials with applications to program testing. *SIAM J. on Comput.*, 25:252–271, 1996.

[RS09] Vojtěch Rödl and Mathias Schacht. Generalizations of the removal lemma. *Combinatorica*, 29(4):467–502, 2009.

[Rub06] Ronitt Rubinfeld. Sublinear time algorithms. In *Proceedings of International Congress of Mathematicians 2006*, volume 3, pages 1095–1110, 2006.

[Sam07] Alex Samorodnitsky. Low-degree tests at large distances. In *Proc. 37th Annual ACM Symposium on the Theory of Computing*, pages 506–515, 2007.

[Sha09] Asaf Shapira. Green's conjecture and testing linear-invariant properties. In *Proc. 41st Annual ACM Symposium on the Theory of Computing*, pages 159–166, 2009.

[Sud10] Madhu Sudan. Invariance in property testing. *Electronic Colloquium in Computational Complexity*, TR10-051, March 2010.

[Sze78] Endre Szemerédi. Regular partitions of graphs. In J.C. Bremond, J.C. Fournier, M. Las Vergnas, and D. Sotteau, editors, *Proc. Colloque Internationaux CNRS 260 -âĂŤ Problèmes Combinatoires et Théorie des Graphes*, pages 399–401, 1978.

[Sze10] Balázs Szegedy. The symmetry preserving regularity lemma. *Proc. Amer. Math. Soc.*, 138:405–408, 2010.

A Proofs omitted from Section 1

Characterization of Reed Muller codes by forbidding systems of induced equations First recall that Reed Muller codes of order d are defined as

$$\mathcal{RM}(d) = \left\{ f : \mathbb{F}_2^n \to \mathbb{F}_2 : \; f(x) = \sum_{S \subset [n], |S| \le d} \; \prod_{i \in S} x_i . \right\}$$

The most common characterization of $\mathcal{RM}(d)$ (see for example [AKK$^+$05]) is that $f \in \mathcal{RM}(d)$ if and only if f satisfies

$$\sum_{S \subset [n], |S| \le d+1} f\left(\alpha + \sum_{i \in S} \alpha_i\right) = 0, \text{ for all } (\alpha, \alpha_1, \dots, \alpha_{d+1}) \in (\mathbb{F}_2^n)^{d+2}.$$

We use this description to obtain a matrix $M \in \mathbb{F}_2^{(2^{d+1}-d-2)\times(2^{d+1})}$ and a collection of $\sigma^i \in \{0,1\}^{2^{d+1}}$ such that $\mathcal{RM}(d)$ is $\{(M,\sigma^i)\}_i$- free. Intuitively, we want M to encode all the linear relations between the elements of the set $A = \{\alpha + \sum_{i \in S} \alpha_i\}_{0 \le |S| \le d+1}$, and we want to use the σ^i's to enforce the fact f should evaluate to 1 on an even number of elements of A.

More exactly, assume that $B = \{\alpha, \alpha + \alpha_1, \dots, \alpha + \alpha_{d+1}\}$ are linearly independent. For every $\beta \in A - B$, add to M the row which is the vector representing β in the basis B. Further, consider all the $\sigma^i \in \{0,1\}^{2^{d+1}}$ such that $|\{j : \sigma^i_j = 1\}|$ is odd. Clearly the number of such σ^i's is finite, and the patterns allowed by forbidding all (M,σ^i) are only those that satisfy the above characterization.

Finally, notice that setting $d = 1$ the resulting matrix M contains only one row, and thus Theorem 3 applies to testing linearity.

We conclude with the proof of Proposition 6 which was also omitted from the Introduction.

Proof of Proposition 6: In one direction, it is easy to check that $\mathcal{F}$-freeness is a subspace-hereditary linear-invariant property, for any fixed family $\mathcal{F}$.

Now, we show the other direction. For a subspace-hereditary linear-invariant property $\mathcal{P}$, let Obs denote the collection of pairs (d, S), where $d \geq 1$ is an integer and $S \subseteq \mathbb{F}_2^d$ is a subset, such that $\mathbf{1}_S$ does not have property $\mathcal{P}$ and is minimal with respect to restriction to subspaces. In other words, (d, S) is contained in Obs iff $\mathbf{1}_S \notin \mathcal{P}_d$ but for any vector subspace $U \subseteq \mathbb{F}_2^d$ of dimension $d' < d$, $\mathbf{1}_{S|_U} \in \mathcal{P}_{d'}$ where $S|_U \subseteq U$ is the restriction of S to U.

For every $(d, S) \in \mathsf{Obs}$, we construct a matrix M_d and a tuple σ_S such that any f with property $\mathcal{P}$ is (M_d, σ_S)-free. Define A_d to be the 2^d-by-d matrix over $\mathbb{F}_2$, where each of the 2^d rows corresponds to a distinct element of $\mathbb{F}_2^d$ represented using some choice of bases. Now, define M_d to be a $(q^d - d)$-by-q^d matrix over $\mathbb{F}$, such that $M_d A_d = 0$ and $\mathsf{rank}(M_d) = q^d - d$. Define σ_S as $(\sigma(1), \sigma(2), \ldots, \sigma(2^d))$ where $\sigma(i) = \mathbf{1}_S(x_i)$ with x_i being the element of $\mathbb{F}_2^d$ represented in the ith row of A_d. We observe now that any $f : \mathbb{F}_2^n \to \{0, 1\}$ having property $\mathcal{P}$ is (M_d, σ_S)-free. Suppose the opposite, so that there exists $x = (x_1, \ldots, x_{q^d}) \in (\mathbb{F}_2^n)^d$ satisfying $Mx = 0$ and $f(x_i) = \sigma(i)$. Then, by definition of M_d, the $x_1, \ldots, x_{2^d}$ are the elements of a d-dimensional subspace V over $\mathbb{F}_2$, and by definition of σ_S, $S_f|_V = S$ where S_f is the support of f. Thus $f|_V \notin \mathcal{P}$ which is a contradiction to the fact that f has property $\mathcal{P}$ because $\mathcal{P}$ is subspace-hereditary.

Finally, define $\mathcal{F}_{\mathcal{P}} = \{(M_d, \sigma_S)\}$. We have just seen that any f having property $\mathcal{P}$ is $\mathcal{F}_{\mathcal{P}}$-free. On the other hand, suppose f does not have property $\mathcal{P}$. Then, because of heredity, there must be a d-dimensional subspace V such that the support of $f|_V$ is isomorphic to S for some $(d, S) \in \mathsf{Obs}$ under linear transformations, which means by the same argument as above, that f will not be (M_d, σ_S)-free. ■

ECCC ISSN 1433-8092

http://eccc.hpi-web.de