Testing Odd-Cycle-Freeness in Boolean Functions

Arnab Bhattacharyya[*]   Elena Grigorescu[†]   Prasad Raghavendra[‡]   Asaf Shapira[§]

Abstract

Call a function $f : \mathbb{F}_2^n \to \{0,1\}$ *odd-cycle-free* if there are no $x_1, \ldots, x_k \in \mathbb{F}_2^n$ with k an odd integer such that $f(x_1) = \cdots = f(x_k) = 1$ and $x_1 + \cdots + x_k = 0$. We show that one can distinguish odd-cycle-free functions from those ϵ-far from being odd-cycle-free by making $\mathrm{poly}(1/\epsilon)$ queries to an evaluation oracle. To obtain this result, we use connections between basic Fourier analysis and spectral graph theory to show that one can reduce testing odd-cycle-freeness of Boolean functions to testing bipartiteness of dense graphs. Our work forms part of a recent sequence of works that shows connections between testability of properties of Boolean functions and of graph properties.

We also prove that there is a *canonical* tester for odd-cycle-freeness making $\mathrm{poly}(1/\epsilon)$ queries, meaning that the testing algorithm operates by picking a random linear subspace of dimension $O(\log 1/\epsilon)$ and then checking if the restriction of the function to the subspace is odd-cycle-free or not. The test is analyzed by studying the effect of random subspace restriction on the Fourier coefficients of a function. Our work implies that testing odd-cycle-freeness using a canonical tester instead of an arbitrary tester incurs no more than a polynomial blowup in the query complexity. The question of whether a canonical tester with polynomial blowup exists for all linear-invariant properties remains an open problem.

Keywords: property testing, Boolean functions, Fourier analysis, Cayley graphs, eigenvalues

1 Introduction

A property testing algorithm is required to distinguish objects that satisfy a given property from objects that are "far" from satisfying the property. One can trace the origins of property testing as an area of study to two distinct origins: [BLR93] (and subsequently [RS96]) which formally investigated the testability of linearity and other *properties of Boolean functions* and [GGR98] which studied the testability of various *graph properties*. Although [GGR98] was inspired by the preceding work on Boolean functions, the two directions evolved more or less independently in terms of the themes considered and the techniques employed. Recently though, this has dramatically changed, and quite a few surprising connections have emerged. In this work, we draw more connections between these two apparently different areas and show how ideas and tools used in the study of graph properties can be used to test certain properties of Boolean functions.

We start with a few definitional remarks. A property of Boolean functions is a subset $\mathcal{P} \subseteq \{f : \{0,1\}^n \to \{0,1\}\}$. The distance between $f, g : \{0,1\}^n \to \{0,1\}$ is given by the Hamming metric $\delta(f,g) = \Pr_x[f(x) \neq g(x)]$, and the distance from f to $\mathcal{P}$ is $\delta_{\mathcal{P}}(f) = \min_{g \in \mathcal{P}} \delta(f,g)$. A function f is ϵ-far from $\mathcal{P}$ if $\delta_{\mathcal{P}}(f) \geq \epsilon$. These definitions carry over to graph properties[1], where the distance to a graph property $\mathcal{P}$ is said to be ϵ if ϵn^2 edges need to be added to or removed from the given graph on n vertices in order to obtain a graph in $\mathcal{P}$. A *tester* for $\mathcal{P}$ is a randomized algorithm which, given oracle access to the input and a parameter $\epsilon \in (0,1)$, accepts with probability at least $2/3$ when the input is in $\mathcal{P}$ and rejects with probability at least $2/3$ when it is ϵ-far from $\mathcal{P}$. In the case of Boolean functions, the tester can query the value of the function at any element of $\{0,1\}^n$, and in the case of graphs, it can query the adjacency matrix at any location. The complexity of a tester is measured by the number of queries it makes to the oracle, and if this quantity is independent of n, the property is called *(strongly) testable*. A one-sided error tester should accept every object in $\mathcal{P}$ with probability 1 and reject every object that is ϵ-far from $\mathcal{P}$ with probability $2/3$.

Our main focus in this paper is the study of the following property of Boolean functions.

Definition 1 (Odd-Cycle-Freeness) *A function $f : \mathbb{F}_2^n \to \{0,1\}$ is said to be* odd-cycle-free (OCF) *if for all odd $k \geq 1$, there are no $x_1, x_2, \ldots, x_k \in \mathbb{F}_2^n$ satisfying $x_1 + \cdots + x_k = 0$ and $f(x_i) = 1$ for all $i \in [k]$.*

The name "odd-cycle-freeness" arises from the observation that f is OCF if and only if the Cayley graph[2] with the support of f as its generators is free of cycles of odd length, i.e. is bipartite. The property of bipartiteness in general graphs has been extensively studied, and nearly tight upper and lower bounds are known for its query complexity [GGR98, AK02, BT04, KKR04]. In this work, we show that odd-cycle-freeness for Boolean functions is testable with comparable query complexity and moreover, using tests that are very similar to the ones for graph bipartiteness.

Odd-cycle-freeness can also be described in a more algebraic way. As observed in Section 2, given a function $f : \mathbb{F}_2^n \to \{0,1\}$ with density $\rho \overset{\text{def}}{=} \mathbb{E}_x[f(x)]$, the distance of f to OCF is exactly $\frac{1}{2}(\rho + \min_\alpha \hat{f}(\alpha))$. So, a Boolean function's distance to OCF is directly connected to the (signed) value of its smallest Fourier coefficient. This link proves crucial in our analysis of tests for OCF.

Our work is part of a larger program to understand the structure of testable properties of Boolean functions. We explain this perspective next.

Common themes in testing A leading question in the search for common unifying themes in property testing has been that of discovering necessary and sufficient conditions for strong testability. Kaufman and Sudan [KS08] suggest that *linear invariance* is a natural property of boolean functions and play an important role in testing. Formally, a property[3] $\mathcal{P} \subseteq \{f : \mathbb{F}_2^n \to \{0,1\}\}$ is said to be linear-invariant if for any $f \in \mathcal{P}$, it is also the case that $f \circ L \in \mathcal{P}$, for any $\mathbb{F}_2$-linear transformation $L : \mathbb{F}_2^n \to \mathbb{F}_2^n$. Some notable examples of properties that were shown to be testable and which are invariant under linear transformations of the domain include linear functions [BLR93], low degree polynomials [AKK+05], and functions with low Fourier dimensionality or sparsity [GOS+09].

---

[1] In this paper, when we refer to graph properties, we will always mean the dense graph model where the graph is represented by its adjacency matrix.

[2] See Section 3 for the precise definition.

[3] Henceforth, we will identify $\{0,1\}^n$ with the vector space $\mathbb{F}_2^n$.

A general class of linear invariant families can be described in terms of forbidden patterns. The first instance of this perspective appeared in the work of Green [Gre05] in testing if a Boolean property is *triangle-free*. Formally, f is triangle-free if it is free from the pattern $\langle f(x), f(y), f(x+y)\rangle = \langle 1^3\rangle$, for any $x, y \in \mathbb{F}_2^n$. More generally, a function is said to be free from solutions to the linear equation $x_1 + \cdots + x_k = 0$ if there are no $x = (x_1, \ldots, x_k) \in (\mathbb{F}_2^n)^k$ satisfying $x_1 + \cdots + x_k = 0$ and $f(x_i) = 1$ for all $i \in [k]$. Pushing this generalization further, for a matrix $M \in \mathbb{F}_2^{m \times k}$ and $\sigma \in \{0,1\}^k$, we say that f 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 $f(x_i) = \sigma_i$ for all $i \in [n]$. This corresponds to freeness from solutions to a system of linear equations. When $\sigma = 1^k$, notice that if $f \in \mathcal{P}$, then any function obtained from f by removing elements in the support of f also must belong to $\mathcal{P}$, and in this case $\mathcal{P}$ is said to be *monotone*.

Green [Gre05] proved that $(M, 1^k)$-freeness is testable with one sided error when M is a rank 1 matrix. Král', Serra and Vena [KSV11] and Shapira [Sha09] showed that this is true regardless of the matrix M. The setting when $\sigma \neq 1^k$ was introduced in [BCSX11], where it is shown that (M, σ)-freeness is testable for any σ as long as M is of rank one. The case of M being rank 1 was fully resolved by Bhattacharyya, Grigorescu and Shapira [BGS10] who showed that any (possibly infinite) intersection of such properties is also testable. Now, note that odd-cycle-freeness is an example of such a property; it is the intersection of $(C_k, 1^k)$-freeness for all odd $k \geq 1$, where C_k is the row vector $[1\ 1\ \cdots\ 1]$ of size k. We next state this result formally.

Theorem 2 ([BGS10]) *There exists a function $f : (0, 1) \to \mathbb{Z}^+$ such that the following is true. For any $\epsilon > 0$, there is a one-sided tester with query complexity $f(\epsilon)$ that distinguishes OCF functions from functions ϵ-far from OCF.*

In fact, odd-cycle-freeness is not just an "example" of a property shown to be testable by [BGS10]: *any* monotone property that can be expressed as freeness from solutions to an infinite set of linear equations is equivalent to the odd-cycle-freeness property (see Section 5 for the short argument). Thus, we view the study of odd-cycle-freeness as an important step towards understanding the testability of monotone linear-invariant properties.

Surprisingly, a similar picture has been staged in the world of testable properties in the *dense graph model*. Just as for Boolean functions, triangle freeness, which was shown (implicitly) by Ruzsa and Szemerédi [RS78] to be testable, brought up a wealthy perspective in the area. It was followed by exciting results in testing H-freeness [ADL+94] and induced H-freeness [AFKS00] which are somewhat analogous to the results on testing monotone and non-monotone properties of Boolean functions. This direction culminated with a nearly complete characterization of all properties that are testable with one-sided error and constant number of queries [AS08].

The proofs in [BGS10] draw heavily on the ideas used in characterizing general graph properties, such as the Szemerédi regularity lemma and Ramsey theory, and just like in that case, the query complexity bounds obtained are a tower of exponentials. Thus, the bound obtained in [BGS10] for $f(\epsilon)$ in Theorem 2 above is an embarrassingly weak tower of exponentials. This brings up the question of characterizing the class of linear-invariant properties which can be tested with poly$(1/\epsilon)$ queries. In this work, we show that odd-cycle-freeness belongs to this class of properties by the way of two different proofs, each with its own message.

1.1 The Edge-Sampling Test

Our first approach for testing OCF relies on reducing testing OCF in Boolean functions to testing bipartiteness of the Cayley graph associated with the function. More precisely, we will show that

the following algorithm is a tester for OCF.

Edge-sampling test (Input: oracle access to $f : \mathbb{F}_2^n \to \{0, 1\}$)

1. Uniformly pick $\alpha_1, \ldots, \alpha_k \in \mathbb{F}_2^n$, where $k = \tilde{O}(1/\epsilon)$. Let $G = \{\alpha_i - \alpha_j \; : \; i < j\}$.
2. Accept if and only if the restriction of f to G is odd-cycle-free.

Theorem 3 *The edge-sampling test is a one-sided tester for odd-cycle-freeness with query complexity $\tilde{O}(1/\epsilon^2)$.*

For a function $f : \mathbb{F}_2^n \to \{0, 1\}$, define the Cayley graph $\mathcal{G}(f) = (V, E)$ to be the graph with vertex set $V = \mathbb{F}_2^n$ and edge set $E = \{(x, y) \mid f(x - y) = 1\}$. It's not hard to show that f is far from odd-cycle-free if and only if $\mathcal{G}(f)$ is far from any bipartite *Cayley* graph. The crux in the analysis of the above testing algorithm involves showing that if f is ϵ-far from being odd-cycle-free, then its Cayley graph $\mathcal{G}(f)$ is $\epsilon/2$-far from *any* bipartite graph. The proof relies on the well-known fact that the eigenvalues of the adjacency matrix of $\mathcal{G}(f)$ are exactly the Fourier coefficients of f, and uses spectral techniques to analyze the distance to bipartiteness. Finally, the test emulates the test for bipartiteness of [AK02].

In fact, we prove something stronger: the distance of $\mathcal{G}(f)$ from bipartiteness is *exactly* half the distance of f from OCF. Using the fact [GGR98, AdlVKK03] that one can estimate a graph's distance from bipartiteness using $\text{poly}(1/\epsilon)$ queries to within additive error ϵ, it follows that one can estimate the distance of f from OCF.

Theorem 4 *There exists an algorithm that, given oracle access to a function $f : \mathbb{F}_2^n \to \{0, 1\}$ and a parameter $\epsilon \in (0, 1)$, makes $\text{poly}(1/\epsilon)$ queries and returns the distance of f to OCF to within an error of $\pm\epsilon$. The same holds for approximating $\min_\alpha \hat{f}(\alpha)$ to within an error of $\pm\epsilon$.*

The second statement is because of the connection between the distance to OCF and Fourier coefficients mentioned earlier. Using the above, we also obtain a $\text{poly}(1/\epsilon)$-query algorithm to approximate distance to linearity that seems different from previously known ones [BLR93, PRR06].

1.2 The Subspace Restriction Test

Call a tester for a graph property $\mathcal{P}$ *canonical* if it works by picking a set of vertices uniformly at random, querying all the edges spanned by these vertices and checking whether the induced graph satisfies $\mathcal{P}$ or not. [Alo02, GT03] showed that if $\mathcal{P}$ is a *hereditary* graph property (i.e., if a graph G satisfies $\mathcal{P}$, then so does every induced subgraph of G), then $\mathcal{P}$ can be in fact tested using a canonical tester with only a quadratic blowup in the query complexity. Moreover, for many natural hereditary graph properties, and in particular, for the property of graph bipartiteness, there is asymptotically no loss in using a canonical tester. The existence of a canonical tester also makes convenient proving lower bounds for hereditary graph properties. It is thus natural to ask if a similar theorem can be proved in the context of testing properties of Boolean functions.

Suppose $\mathcal{P}$ is a *subspace-hereditary* linear-invariant property of Boolean functions, meaning that if a function $f : \mathbb{F}_2^n \to \{0, 1\}$ satisfies $\mathcal{P}$, then so does f restricted to any linear subspace of the domain. Subspace-hereditary properties are especially interesting because they include most

natural linear-invariant properties and are conjectured in [BGS10] to be exactly the class of one-sided testable linear-invariant properties (modulo some technicalities). Now, just as a canonical tester for a hereditary graph property works by sampling a set of vertices S and querying all edges induced in S, one defines a canonical tester for a subspace-hereditary property $\mathcal{P}$ to be the algorithm that, on input $\epsilon \in (0, 1)$ and oracle access to $f : \mathbb{F}_2^n \to \{0, 1\}$, chooses uniformly at random a linear subspace $H \leq \mathbb{F}_2^n$ of dimension $d(\epsilon, n)$ (for some fixed function $d : [0, 1] \times \mathbb{Z}^+ \to \mathbb{Z}^+$) and accepts if and only if f restricted to H satisfies $\mathcal{P}$. The query complexity of the canonical tester is obviously $2^{d(\epsilon,n)}$, the size of the subspace inspected by the tester. It is shown in [BGS10], using ideas similar to those in [Alo02, GT03], that any tester for a subspace-hereditary linear-invariant property can be converted to be of canonical form, but at the expense of an exponential blowup in the query complexity. The question that arises then is whether, instead of an exponential blowup, only a polynomial blowup in the query complexity is always possible.

Question 5 *Given a subspace-hereditary property $\mathcal{P}$ that can be tested with q queries, is there always a canonical tester of complexity* $\mathrm{poly}(q)$?

This seems to be a hard question in general. However, in this work, we show that for the property of odd-cycle-freeness, the answer to Question 5 is affirmative. (Note that the edge-sampling test is not canonical.)

Subspace restriction test (Input: oracle access to $f : \mathbb{F}_2^n \to \{0, 1\}$)

1. Uniformly pick $\alpha_1, \ldots, \alpha_k \in \mathbb{F}_2^n$, where $k = O(\log \frac{1}{\epsilon})$. Let H be the linear subspace spanned by $\alpha_1, \ldots, \alpha_k$.
2. Accept if and only if the restriction of f to H is odd-cycle-free.

Theorem 6 *The subspace restriction test is a one-sided tester for odd-cycle-freeness with query complexity* $O(1/\epsilon^{20})$.

The analysis of the subspace-restriction test relies on a Fourier analytic argument. One can easily see that the test accepts every OCF function. The main insight is that certain properties of the Fourier spectrum of a function that is ϵ-far from being OCF are preserved under random restrictions to small subspaces.

Note that Theorem 6 implies the combinatorial fact that for any function f that is ϵ-far from OCF, there must exist a short witness to this fact. That is, there must exist $x_1, \ldots, x_k \in \mathrm{supp}(f)$ with $x_1 + \cdots + x_k = 0$ and $k = O(\log 1/\epsilon)$ an odd integer. In fact, Theorem 6 asserts that there must exist many such witnesses, but a priori, it is not clear that even one such witness exists. This is in contrast to properties such as triangle-freeness studied in [Gre05], where witnesses to violations of triangle-freeness are, by definition, short, and for testability, one "only" needs to show that there exist many such witnesses.

1.3 Organization

The rest of the paper is organized as follows. In Section 2 we show that one can relate the distance of f from OCF to the Fourier expansion of f. In Section 3 we use this relation together with

some results from spectral graph theory in order to analyze the edge-sampling test and thus prove Theorem 3. We also show in this section that a strengthening of the analysis for the edge-sampling test yields a distance estimator. In Section 4, we analyze the subspace restriction test and prove Theorem 6. Finally, Section 5 contains some concluding remarks and a discussion of some open problems related to this paper.

2 Odd-Cycle-Freeness and the Fourier Spectrum

Our goal in this section is to give two reformulations of OCF, one of a geometric flavor and one in terms of the coefficients of the Fourier expansion of f. These characterizations of OCF will be useful in the analysis of both the edge-sampling test and the subspace restriction test which will be given in later sections. We begin by recalling some basic facts about Fourier analysis of Boolean functions.

The orthonormal characters $\{\chi_\alpha : \mathbb{F}_2^n \to \mathbb{R}, \ \chi_\alpha(x) = (-1)^{\alpha \cdot x}\}_{\alpha \in \mathbb{F}_2^n}$ form a basis for the set of $\{0,1\}$-valued functions defined over $\mathbb{F}_2^n$, where the inner product is given by $\langle f, g \rangle = \mathbb{E}_x[f(x)g(x)]$. The Fourier coefficient of f at $\alpha \in \mathbb{F}_2^n$ is $\widehat{f}(\alpha) = \mathbb{E}_x[f(x)\chi_\alpha(x)]$. The *density* of f is $\rho = \mathbb{E}_x f(x) = \widehat{f}(0)$, and notice that $\rho = \max_{\alpha \in \mathbb{F}_2^n} |\widehat{f}(\alpha)|$. The support of f is $\mathrm{supp}(f) = \{x \in \mathbb{F}_2^n | f(x) \neq 0\}$. Parsevals's identity states that $\sum_\alpha \widehat{f}(\alpha)^2 = \mathbb{E}_x f(x)^2 = \rho$. Also, for all α, $\widehat{f}(\alpha) \geq \max(-\rho, -1/2)$.

We first notice that the presence of cycles in a function induces a certain distribution of the density of the function on halfspaces.

Claim 7 *Let $f : \mathbb{F}_2^n \to \{0,1\}$. Then:*

(a) *f is OCF if and only if there exists $\alpha \in \mathbb{F}_2^n$ such that for all $x \in \mathrm{supp}(f)$, $\alpha \cdot x = 1$ (i.e., there exists a linear subspace of dimension $n-1$ that does not contain any element of $\mathrm{supp}(f)$).*

(b) *f is ϵ-far from OCF if and only if for every $\alpha \in \mathbb{F}_2^n$, it holds that for at least $\epsilon 2^n$ many $x \in \mathrm{supp}(f)$, $\alpha \cdot x = 0$ (i.e., every linear subspace of dimension $n-1$ must contain at least $\epsilon 2^n$ elements of $\mathrm{supp}(f)$).*

Proof We first prove part (a). To see the "if" direction, suppose f is not odd-cycle-free, but there exists α such that for all $x \in \mathrm{supp}(f)$, $\alpha \cdot x = 1$. Now, let $x_1, \ldots, x_{k-1} \in \mathrm{supp}(f)$ be such that $x_1 + \cdots + x_{k-1} \in \mathrm{supp}(f)$ and k is odd. But then $\alpha \cdot (x_1 + \cdots + x_{k-1}) = \sum_{i=1}^{k-1} \alpha \cdot x_i = 0$, a contradiction. For the opposite direction, suppose f is odd-cycle-free. If f is the zero function, we are vacuously done. Assuming otherwise, let $S = \mathrm{supp}(f)$, and consider the set $H' = \{x_1 + \cdots + x_k \ : \ x_1, \ldots, x_k \in S, \text{ and } k \geq 0 \text{ is even}\}$. Since f is OCF, $H' \cap S = \emptyset$. It is easy to see that H' is a linear subspace of codimension 1 inside $\mathrm{span}(S)$. It follows that H' can be extended to a subspace H of dimension $n-1$ such that $H \cap S = \emptyset$.

For part (b), if f is ϵ-far from being OCF, then by part (a), every linear subspace H of dimension $n-1$ must contain $\epsilon 2^n$ elements of $\mathrm{supp}(f)$ (otherwise removing less than $\epsilon 2^n$ points from $\mathrm{supp}(f)$ would create a function that is OCF.) The converse follows again by part (a). □

Lemma 8 *Let $f : \mathbb{F}_2^n \to \{0,1\}$. Then*

(a) *f is OCF if and only if there exists $\alpha \in \mathbb{F}_2^n$ such that $\widehat{f}(\alpha) = -\rho$.*

6

(b) f is ϵ-far from being OCF if and only if for all $\beta \in \mathbb{F}_2^n$, $\widehat{f}(\beta) \geq -\rho + 2\epsilon$.

(c) The distance of f from OCF is exactly $\frac{1}{2}\left(\rho + \min_\alpha \hat{f}(\alpha)\right)$.

Proof By Claim 7, f is OCF if and only if there exists α such that $\alpha \cdot x = 1$ for all $x \in \mathrm{supp}(f)$, and so it follows that

$$\widehat{f}(\alpha) = \mathbb{E}\, f(x)(-1)^{\alpha \cdot x} = -\rho.$$

This implies item (a) of the lemma. To derive item (b), we get from Claim 7, that f is ϵ-far from OCF if an only if any halfspace contains at least an ϵ fraction of the domain, implying that for each $\beta \in \mathbb{F}_2^n$:

$$
\begin{aligned}
\widehat{f}(\beta) &= \mathop{\mathbb{E}}_{x \in \mathbb{F}_2^n}\, f(x)(-1)^{\beta \cdot x} = \frac{1}{2^n}\left(\sum_{x \in \mathrm{supp}, \beta \cdot x = 0} 1 + \sum_{x \in \mathrm{supp}, \beta \cdot x = 1}(-1)\right) \\
&\geq \epsilon + (-\rho + \epsilon) = -\rho + 2\epsilon.
\end{aligned}
$$

Finally, observe that item (c) is just a restatement of item (b). $\qquad\square$

We see that the minimum Fourier coefficient of f determines its distance from OCF. Since Fourier coefficients also measure correlation to linear functions, it is natural to ask about the relationship between a function's distance to OCF and its distance to linearity[4]. Easy Fourier analysis shows that the distance of a function $f : \mathbb{F}_2^n \to \mathbb{F}_2$ to linearity is exactly $\min(\rho, \frac{1}{2} + \min_\alpha \hat{f}(\alpha))$. So, the distance to linearity, in contrast to OCF, is not always determined by the minimum Fourier coefficient.

3 The Edge-Sampling Test

In this section, we analyze the edge-sampling test and prove Theorem 3. The analysis starts with the characterization of OCF given in the previous section and then proceeds to reduce the problem of testing OCF for Boolean functions to testing bipartiteness in dense graphs. We then show why Theorem 4 follows.

For a function $f : \mathbb{F}_2^n \to \{0, 1\}$, define the Cayley graph $\mathcal{G}(f) = (V, E)$ to be the graph with vertex set $V = \mathbb{F}_2^n$ and edge set $E = \{(x, y) \mid f(x - y) = 1\}$. Let us denote by $N = 2^n$ and let $A_\mathcal{G}$ be the adjacency matrix of $\mathcal{G}$. The next lemma is well-known but we include its proof for the sake of completeness.

Lemma 9 For any $\alpha \in \mathbb{F}_2^n$, the character χ_α is an eigenvector of $A_\mathcal{G}$ of normalized eigenvalue $\widehat{f}(\alpha)$. Moreover, the set $\{2^n \widehat{f}(\alpha)\}_\alpha$ is exactly the set of all the eigenvalues of $A_\mathcal{G}$.

Proof Notice that the entry indexed by x_i in $b = A\chi_\alpha$ is

$$b_i = \sum_{x_j \in \mathbb{F}_2^n} f(x_i - x_j)\chi_\alpha(x_j)$$

---

[4] A function $f : \mathbb{F}_2^n \to \mathbb{F}_2$ is said to be linear if $f(x + y) = f(x) + f(y)$ for all x, y (the range $\{0, 1\}$ has been identified with $\mathbb{F}_2$). Note that linear functions are OCF. However, the converse is certainly false, since the function $f(x) = x_1 x_2$ is OCF but $1/4$-far from linear.

$$= \sum_{x \in \mathbb{F}_2^n} f(x) \chi_\alpha(x_i - x)$$

$$= \chi_\alpha(x_i) \sum_{x \in \mathbb{F}_2^n} f(x) \chi_\alpha(x)$$

$$= \chi_\alpha(x_i)(2^n \widehat{f}(\alpha)).$$

Therefore, $A\chi_\alpha = (2^n \widehat{f}(\alpha))\chi_\alpha$, and since the set of characters contains 2^n orthogonal vectors, the lemma follows. $\square$

We remind the reader that in the context of testing graph properties, a graph is ϵ-far from being bipartite if one needs to remove at least ϵn^2 edges in order to make it odd-cycle-free. In order to be able to apply results concerning testing odd-cycle-freeness in graphs we will have to prove that $\mathcal{G}(f)$ is in fact ϵ-far from *any* bipartite graph. To this end, we show the following lemma that relates the distance to being bipartite to the least eigenvalue of the adjacency matrix. In what follows, we denote by $e(S)$ the number of edges inside a set of vertices S in some graph G.

Lemma 10 *Let λ_{min} be the smallest eigenvalue of the adjacency matrix A of an n-vertex d-regular graph G. Then for every $U \subseteq V(G)$, we have*

$$e(U) \geq \frac{|U|}{2n} \left(|U|d + \lambda_{min}(n - |U|) \right). \tag{1}$$

Proof Let u be the indicator vector of U. We clearly have

$$u^T A u = 2e(U) .$$

Since A is symmetric it has a collection of eigenvectors $v_1, \ldots, v_n$ which form an orthonormal basis for $\mathbb{R}^n$. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues corresponding to these eigenvectors where $\lambda_n = \lambda_{min}$. Suppose we can write $u = \sum_{i=1}^n \alpha_i v_i$ in this basis and note that $\sum_{i=1}^n \alpha_i^2 = |U|$. Since G is d-regular, $(1/\sqrt{n}, \ldots, 1/\sqrt{n})$ is an eigenvector of A. Suppose this is v_1 and note that this means that $\lambda_1 = d$ and $\alpha_1 = |U|/\sqrt{n}$. Combining the above observations we see that

$$
\begin{aligned}
u^T A u &= \sum_{i=1}^n \lambda_i \alpha_i^2 \\
&= d|U|^2/n + \sum_{i=2}^n \lambda_i \alpha_i^2 \\
&\geq d|U|^2/n + \lambda_{min}\left(\sum_{i=2}^n \alpha_i^2 \right) \\
&= d|U|^2/n + \lambda_{min}(|U| - |U|^2/n) .
\end{aligned}
$$

We now get (1) by combining the above two expressions for $u^T A u$. $\square$

Corollary 11 *Let G be an n-vertex d-regular graph with $\lambda_{min} \geq -d + 2\epsilon n$. Then G is $\epsilon/2$-far from being bipartite.*

Proof It is clearly enough to show that in any bipartition of the vertices of G into sets A, B we have $e(A) + e(B) \geq \frac{1}{2}\epsilon n^2$. So let (A, B) be one such bipartition and suppose $|A| = cn$ and $|B| = (1 - c)n$. From Lemma 10 we get that

$$
\begin{aligned}
e(A) &\geq \frac{c}{2}(dcn + (-d + 2\epsilon n)(n - cn)) \\
&= \frac{c}{2}(2\epsilon n^2 - dn) + \frac{c^2}{2}(2dn - 2\epsilon n^2) \,,
\end{aligned}
$$

and similarly

$$
e(B) \geq \frac{1 - c}{2}(2\epsilon n^2 - dn) + \frac{(1 - c)^2}{2}(2dn - 2\epsilon n^2) \,.
$$

Hence

$$
\begin{aligned}
e(A) + e(B) &\geq \frac{1}{2}(2\epsilon n^2 - dn) + \frac{1}{2}(c^2 + (1 - c)^2)(2dn - 2\epsilon n^2) \\
&\geq \frac{1}{2}(2\epsilon n^2 - dn) + \frac{1}{4}(2dn - 2\epsilon n^2) \\
&= \frac{1}{2}\epsilon n^2 \,,
\end{aligned}
$$

where in the second inequality we use the fact that $c^2 + (1 - c)^2$ is minimized when $c = 1/2$. $\square$

We can now derive the following *exact* relation between the OCF property of functions and the bipartiteness of the corresponding Cayley graphs.

Corollary 12 *Let $f : \mathbb{F}_2^n \to \mathbb{F}_2$. If f is ϵ-far from being OCF, then $\mathcal{G}(f)$ is $\epsilon/2$-far from being bipartite. Furthermore, if f is ϵ-close to being OCF, then $\mathcal{G}(f)$ is $\epsilon/2$-close to being bipartite.*

Proof Suppose f is ϵ-far from being OCF. Let us suppose $\text{supp}(f) = \rho N$ where $N = 2^n$. By Lemmas 8 and 9, if f is ϵ-far from being OCF, then the smallest eigenvalue of the adjacency matrix of $\mathcal{G}(f)$ is $\lambda_{min} \geq -(\rho + 2\epsilon)N$. For a function f, recall that $\mathcal{G}(f)$ is a regular graph with degree $d = |\text{supp}(f)| = \rho N$. Hence, we can use Corollary 11 to infer that $\mathcal{G}(f)$ must be $\epsilon/2$-far from being bipartite.

Suppose now that f is ϵ-close to being OCF, and let S be the set of ϵN points in $\mathbb{F}_2^n$ whose removal from the support of f makes it OCF. Call this new function f'. Observe that every $x \in \mathbb{F}_2^n$ for which $f(x) = 1$ accounts for $N/2$ edges in $\mathcal{G}(f)$. Hence, removing from $\mathcal{G}(f)$ all edges corresponding to S results in the removal of at most $\frac{1}{2}\epsilon N^2$ edges. To finish the proof we just need to show that the new graph $\mathcal{G}(f')$ (note that the new graph is indeed the Cayley graph of the new function f') does not contain any odd cycle. Suppose to the contrary that it contains an odd cycle $\alpha_1, \ldots, \alpha_k, \alpha_1$. For $1 \leq i \leq k$ set $x_i = \alpha_{i+1} - \alpha_i$. Then by definition of $\mathcal{G}(f')$ we have $f'(x_1) = \ldots = f'(x_k) = 1$. Furthermore, as $\alpha_1 + \sum_{i=1}^{k} x_i = \alpha_1$ (since we have a cycle in $\mathcal{G}(f')$) we get that $\sum_{i=1}^{k} x_i = 0$ so $x_1, \ldots, x_k$ in an odd-cycle in f' contradicting the assumption that f' is OCF. $\square$

We are now ready to complete the proof of Theorem 3 using the following result of Alon and Krivelevich [AK02].

Theorem 13 ([AK02]) *Suppose a graph G is ϵ-far from being bipartite. Then a random subset of vertices of $V(G)$ of size $\tilde{O}(1/\epsilon)$ spans a non-bipartite graph with probability at least $3/4$.*

Proof of Theorem 3 First, if f is OCF then the test will clearly accept f (with probability 1). Suppose now that f is ϵ-far from being OCF. Then by Corollary 12 we get that $\mathcal{G}(f)$ is $\epsilon/2$-far from being bipartite. Now notice that we can think of the points $\alpha_1, \ldots, \alpha_k \in \mathbb{F}_2^n$ sampled by the edge-sampling test as vertices sampled from $\mathcal{G}(f)$. By Theorem 13, with probability at least $3/4$, the vertices $\alpha_1, \ldots, \alpha_k$ span an odd-cycle of $\mathcal{G}(f)$. We claim that if this event happens, then the edge-sampling test will find an odd-cycle in f. Indeed, if $\alpha_1, \ldots, \alpha_k, \alpha_1$ is an odd-cycle in $\mathcal{G}(f)$, then as in the proof of Corollary 12 this means that $\alpha_2 - \alpha_1, \ldots, \alpha_3 - \alpha_2, \ldots, \alpha_1 - \alpha_k$ is an odd-cycle of f. Finally, the edge-sampling test will find this odd cycle, since it queries f on all points $\alpha_i - \alpha_j$.
$\square$

In order to obtain Theorem 4, observe that by Corollary 12, the distance to OCF for a function f is exactly double the distance to bipartiteness for the graph $\mathcal{G}(f)$. We now invoke the following result of Alon, de la Vega, Kannan and Karpinski [AdlVKK03], which improved upon a previous result of Goldreich, Goldwasser and Ron [GGR98].

Theorem 14 ([AdlVKK03]) *For every $\epsilon > 0$, there exists an algorithm that, given input graph G, inspects a random subgraph of G on $\tilde{O}(1/\epsilon^4)$ vertices and estimates the distance from G to bipartiteness to within an additive error of ϵ.*

Proof of Theorem 4 Combining Theorem 14 with Lemma 12 we immediately obtain a $\text{poly}(1/\epsilon)$ query algorithm that estimates the distance to odd-cycle-freeness with additive error at most ϵ. Since one can use sampling to estimate ρ to within an additive error ϵ using $\text{poly}(1/\epsilon)$ queries, it follows from item (c) of Lemma 8 that one can estimate $\min_\alpha \hat{f}(\alpha)$ to within an additive error of ϵ using $\text{poly}(1/\epsilon)$ queries. $\square$

As we have mentioned earlier, the distance of f from being linear is given by $\min(\rho, \frac{1}{2} + \min_\alpha \hat{f}(\alpha))$, where ρ is the density of f. Therefore, given an estimate of ρ and $\min_\alpha \hat{f}(\alpha)$ for some function $f : \mathbb{F}_2^n \to \mathbb{F}_2$, one can also estimate the distance of f to linearity. Theorem 4 thus gives a new distance estimator for linearity, and hence also a two-sided tester for the property of linearity, both with $\text{poly}(1/\epsilon)$ query complexity.

4 The Subspace Restriction Test

In this section, we analyze the subspace restriction test and prove Theorem 6. We start with a few notational remarks. For $f : \mathbb{F}_2^n \to \{0,1\}$ and subspace $H \le \mathbb{F}_2^n$, let $f_H : H \to \{0,1\}$ be the restriction of f to H, and let ρ_H denote the density of f_H, namely $\rho_H = \Pr_{x \in H}[f_H(x) = 1]$. For $\alpha \in \mathbb{F}_2^n$ and subspace H, define the restriction of the Fourier coefficients of f to a subspace H to be

$$\widehat{f}_H(\alpha) = \mathop{\mathbb{E}}_{x \in H}[f(x)\chi_\alpha(x)].$$

Recall that the character group of H is isomorphic to H itself, and so, $f_H = \sum_{\alpha \in H} \widehat{f}_H(\alpha)\chi_\alpha$. The dual of H is the subspace $H^\perp = \{x \in \mathbb{F}_2^n \mid \langle x, a \rangle = 0 \; \forall a \in H\}$. Note that $\widehat{f}_H(\alpha) = \widehat{f}_H(\beta)$ whenever

$\alpha \in \beta + H^\perp$. The convolution of $f, g : \mathbb{F}_2^n \to \mathbb{F}_2$ is $f * g : \mathbb{F}_2^n \to \mathbb{F}_2$, $(f * g)(c) = \mathbb{E}_{x \in \mathbb{F}_2^n} f(x + c) f(x)$. It is known that $\widehat{f * g} = \widehat{f} \cdot \widehat{g}$. In what follows, we will let h be the size of the the subspace H.

The strategy of the proof is to use Lemma 8 and reduce the analysis to showing that if every nonzero Fourier coefficient of f is at least $-\rho + 2\epsilon$, then for a random linear subspace H, with probability $2/3$, every nonzero Fourier coefficient of f_H is strictly greater than $-\rho_H$ (where, again, $f_H : H \to \{0, 1\}$ is the restriction of f to H and ρ_H is the density of f_H).

A useful insight into why this should be true is that the restricted Fourier coefficients are concentrated around the non-restricted counterparts, deviating from them by an amount essentially inversely proportional with the size of the subspace H. A direct union bound type argument is however too weak to give anything interesting when the size of H is small. The idea of our proof is to separately analyze the restrictions of the large and small coefficients. Understanding the restrictions of the small coefficients is the more difficult part of the argument, and the crux of the proof relies on noticing that the moments of the Fourier coefficients are also preserved under restrictions to subspaces. In particular, an analysis of the deviation of the fourth moment implies that one can balance the parameters involved so that even when H is of size only $\text{poly}(1/\epsilon)$, no restriction of the coefficients of low magnitude can become as small as $-\rho_H$.

We first show that the restriction of f to a random linear subspace does not change an individual Fourier coefficient by more than a small additive term dependent on the size of the subspace. This follows from standard Chebyshev-type concentration bounds.

Lemma 15
$$\Pr_H \left[\left| \widehat{f_H}(\alpha) - \widehat{f}(\alpha) \right| \geq \frac{2}{h} + \eta \right] \leq \frac{14}{h\eta^2}.$$

Proof
Now, consider the deviation of $\mathbb{E}_H[\widehat{f_H}(\alpha)]$ from $\widehat{f}(\alpha)$:

$$\mathbb{E}_H[\widehat{f_H}(\alpha)] = \mathbb{E}_H \mathbb{E}_{x \in H}[f(x)\chi_\alpha(x)]$$
$$\geq \mathbb{E}_H \left(\left(1 - \frac{1}{h}\right) \mathbb{E}_{x \in H - \{0\}}[f(x)\chi_\alpha(x)] \right)$$
$$\geq \mathbb{E}_{x \in \mathbb{F}_2^n - \{0\}}[f(x)\chi_\alpha(x)] - \frac{1}{h}$$
$$\geq \widehat{f}(\alpha) - \frac{1}{2^n} - \frac{1}{h} \geq \widehat{f}(\alpha) - \frac{2}{h}$$

Similarly:

$$\mathbb{E}_H[\widehat{f_H}(\alpha)] \leq \widehat{f}(\alpha) + \frac{2}{h}$$

So, it suffices to show that $\Pr\left[|\widehat{f_H}(\alpha) - \mathbb{E}_H \widehat{f_H}(\alpha)| \geq \eta \right] \leq 10/(h\eta^2)$. We prove this by bounding the variance of $\widehat{f_H}(\alpha)$.

$$\mathbb{E}[\widehat{f_H}(\alpha)^2] = \mathbb{E}_H \left[\left(\mathbb{E}_{x \in H} f(x)\chi_\alpha(x) \right)^2 \right]$$

11

$$= \mathbb{E}_H \left[\mathbb{E}_{x,y \in H} f(x)f(y)\chi_\alpha(x)\chi_\alpha(y) \right]$$

$$\leq \mathbb{E}_H \left[\Pr[\dim(\mathrm{span}(x,y)) < 2] + \mathbb{E}_{\substack{x,y \in H \\ \dim(\mathrm{span}(x,y))=2}} f(x)f(y)\chi_\alpha(x)\chi_\alpha(y) \right]$$

$$\leq \frac{3}{h} + \mathbb{E}_{\substack{x,y \in \mathbb{F}_2^n \\ \dim(\mathrm{span}(x,y))=2}} f(x)f(y)\chi_\alpha(x)\chi_\alpha(y)$$

$$\leq \frac{3}{h} + \frac{3}{2^n} + \widehat{f}^2(\alpha)$$

$$\leq \frac{6}{h} + \left(\mathbb{E}_H \widehat{f}_H(\alpha) + \frac{2}{h} \right)^2 \leq \frac{14}{h} + \left(\mathbb{E}_H \widehat{f}_H(\alpha) \right)^2$$

Hence $\mathrm{Var}[\widehat{f}_H(\alpha)] \leq \frac{14}{h}$, and the lemma now follows by Chebyshev's inequality. $\qquad \square$

As it was the case with the restricted coefficients, it can also be shown using a straightforward variance calculation that the fourth moment is preserved up to small additive error upon restriction to a random H, when h is large enough. For that purpose, define A and A_H as follows:

$$A \stackrel{\mathrm{def}}{=} \sum_{\alpha \in \mathbb{F}_2^n} \widehat{f}^4(\alpha) = \mathbb{E}_{x_1,x_2,x_3 \in \mathbb{F}_2^n} [f(x_1)f(x_2)f(x_3)f(x_1+x_2+x_3)]$$

$$A_H \stackrel{\mathrm{def}}{=} \sum_{\alpha \in H} \widehat{f}_H^4(\alpha) = \mathbb{E}_{x_1,x_2,x_3 \in H} [f(x_1)f(x_2)f(x_3)f(x_1+x_2+x_3)]$$

Then, we have:

Lemma 16

$$\Pr_H \left[|A_H - A| \geq \frac{16}{h} + \eta \right] \leq \frac{500}{h\eta^2}.$$

Proof   As in the proof of Lemma 15, our strategy will be to first show that $\mathbb{E}_H[A_H]$ is likely to be close to A and then to bound the variance of A_H.

Claim 17

$$|A - \mathbb{E}_H[A_H]| \leq \frac{16}{h}.$$

Proof

$$\mathbb{E}_H[A_H] = \mathbb{E}_H \mathbb{E}_{x_1,x_2,x_3 \in H} f(x_1)f(x_2)f(x_3)f(x_1+x_2+x_3)$$

$$\geq \mathbb{E}_H \left[\left(1 - \frac{8}{h} \right) \mathbb{E}_{\substack{x_1,x_2,x_3 \in H \\ \dim(\mathrm{span}(x_1,x_2,x_3))=3}} f(x_1)f(x_2)f(x_3)f(x_1+x_2+x_3) \right]$$

$$\geq \mathbb{E}_H \left[\mathbb{E}_{\substack{x_1,x_2,x_3 \in H \\ \dim(\mathrm{span}(x_1,x_2,x_3))=3}} f(x_1)f(x_2)f(x_3)f(x_1+x_2+x_3) - \frac{8}{h} \right]$$

$$= \mathbb{E}_{\substack{x_1,x_2,x_3 \in \mathbb{F}_2^n \\ \dim(\mathrm{span}(x_1,x_2,x_3))=3}} f(x_1)f(x_2)f(x_3)f(x_1+x_2+x_3) - \frac{8}{h}$$

12

$$\geq \underset{x_1,x_2,x_3\in\mathbb{F}_2^n}{\mathbb{E}} f(x_1)f(x_2)f(x_3)f(x_1+x_2+x_3) - \frac{8}{2^n} - \frac{8}{h} \geq A - \frac{16}{h}$$

and similarly:

$$\underset{H}{\mathbb{E}}[A_H] \leq A + \frac{16}{h}$$

$\square$

Claim 18

$$\mathrm{Var}[A_H] \leq \frac{500}{h}.$$

Proof

$$\underset{H}{\mathbb{E}}[A_H^2] = \underset{H}{\mathbb{E}}\left[\underset{\substack{x_1,x_2,x_3\in H \\ y_1,y_2,y_3\in H}}{\mathbb{E}} f(x_1)f(x_2)f(x_3)f(x_1+x_2+x_3)f(y_1)f(y_2)f(y_3)f(y_1+y_2+y_3) \right]$$

$$\leq \underset{H}{\mathbb{E}}\Big[\Pr[\dim(\mathrm{span}(\{x_1,x_2,x_3,y_1,y_2,y_3\})) < 6]$$

$$+ \underset{\substack{x_1,x_2,x_3,y_1,y_2,y_3\in H \\ \dim(\mathrm{span}(\{x_1,x_2,x_3,y_1,y_2,y_3\}))=6}}{\mathbb{E}} [f(x_1)f(x_2)f(x_3)f(x_1+x_2+x_3)f(y_1)f(y_2)f(y_3)f(y_1+y_2+y_3)]\Big]$$

$$\leq \frac{64}{h} + \underset{\substack{x_1,x_2,x_3,y_1,y_2,y_3\in\mathbb{F}_2^n \\ \dim(\mathrm{span}(\{x_1,x_2,x_3,y_1,y_2,y_3\}))=6}}{\mathbb{E}} [f(x_1)f(x_2)f(x_3)f(x_1+x_2+x_3)f(y_1)f(y_2)f(y_3)f(y_1+y_2+y_3)]$$

$$\leq \frac{64}{h} + \frac{64}{2^n} + A^2$$

$$\leq \frac{128}{h} + \left(\frac{16}{h} + \underset{H}{\mathbb{E}}[A_H]\right)^2 \leq \frac{500}{h} + \underset{H}{\mathbb{E}}[A_H]^2.$$

$\square$

The lemma now follows by Chebyshev's inequality. $\square$

Using Lemma 15 and 16 we can now proceed with the proof of Theorem 6.

Proof of Theorem 6 If f is ϵ-far from odd-cycle-free, then $\rho > \epsilon$, and by Lemma 8, all its Fourier coefficients are $> -\rho + 2\epsilon$. We need to show that with constant probability over random choice of H, each Fourier coefficient of f_H is $> -\rho_H$. We separate these coefficients into the sets of large and small coefficients and analyze them separately. Define

$$L \overset{\mathrm{def}}{=} \{\alpha \mid |\widehat{f}(\alpha)| \geq \gamma\} \subseteq \mathbb{F}_2^n \quad \text{and} \quad S \overset{\mathrm{def}}{=} \mathbb{F}_2^n \backslash L$$

for some $\gamma < \rho$ to be chosen later. Notice that $0 \in L$. Also, by Parseval's identity, $|L| \leq 1/\gamma^2$. Let $L_H \subseteq H$ be the set of elements $\beta \in H$ such that there exists $\alpha \in L$ with $\beta \in \alpha + H^\perp$, that is, β is

the "projection" of some large coefficient. Then $|L_H| \leq |L|$. Let $S_H = H \backslash L_H$ be the complement of L_H in H.

From Lemma 15, for each $\alpha \in L$ and for any $\eta_1 \in (0, 1)$, we have $\Pr_H \left[|\widehat{f}_H(\alpha) - \widehat{f}(\alpha)| \geq \frac{2}{h} + \eta_1 \right] \leq \frac{14}{h\eta_1^2}$. By a union bound, with probability $1 - \frac{1}{\gamma^2} \frac{14}{h\eta_1^2}$, for every $\alpha \in L_H$, it holds $\widehat{f}_H(\alpha) > \widehat{f}(\alpha) - \frac{2}{h} - \eta_1$. Moreover, since $0 \in L$, we know $|\rho_H - \rho_f| \leq \frac{2}{h} + \eta_1$. If $2\eta_1 + \frac{4}{h} < 2\epsilon$, then for any $\alpha \in L_H$, we have

$$\widehat{f}_H(\alpha) > \widehat{f}(\alpha) - \frac{2}{h} - \eta_1 > -\rho + 2\epsilon - \frac{2}{h} - \eta_1 > -\rho_H + 2\epsilon - \frac{4}{h} - 2\eta_1 > -\rho_H$$

with probability at least $1 - \frac{14}{h\gamma^2\eta_1^2}$.

We now analyze the coefficients $\beta \in S_H$ and again show that with constant probability, no $\widehat{f}_H(\beta)$ becomes as small as $-\rho_H$. As we described in the informal proof sketch earlier, for this, we will want to analyze the fourth moment of the Fourier coefficients.

To this end, first observe that for any two Fourier coefficients $\alpha, \alpha' \in L$, their projections are identical if $\alpha - \alpha' \in H^\perp$. Over the random choice of H, this happens with probability at most $\frac{1}{h}$. Therefore, using a union bound, we conclude that with probability at least $1 - |L|^2/h = 1 - \frac{1}{\gamma^4 h}$, all the large Fourier coefficients project to distinct coefficients in H, namely $|L_H| = |L|$. Let us condition on this event that no two large Fourier coefficients in L project to the same restricted coefficient.

Let us also condition on the event that $|A - A_H| < \frac{16}{h} + \eta_2$ for some η_2 to be specified later. Also, condition on the event that for all $\alpha \in L$, $|\widehat{f}_H(\alpha) - \widehat{f}(\alpha)| < \frac{2}{h} + \eta_1$. All of these events occur with probability at least $1 - \frac{500}{h\eta_2^2} - \frac{14}{h\gamma^2\eta_1^2} - \frac{1}{\gamma^4 h}$ by Lemmas 15 and 16.

The following claim shows that the fourth moment of the small Fourier coefficients is also preserved under a random subspace restriction.

Claim 19

$$\left| \sum_{\alpha \in S_H} \widehat{f}_H^4(\alpha) - \sum_{\alpha \in S} \widehat{f}^4(\alpha) \right| \leq \eta_2 + \frac{16}{h} + \frac{4}{\gamma^2} \left(\frac{2}{h} + \eta_1 \right).$$

Proof   We have that

$$\left| \sum_{\alpha \in L_H} \widehat{f}_H^4(\alpha) - \sum_{\alpha \in L} \widehat{f}^4(\alpha) \right| = \left| \sum_{\alpha \in L} \widehat{f}_H^4(\alpha) - \sum_{\alpha \in L} \widehat{f}^4(\alpha) \right|$$

$$\leq \sum_{\alpha \in L} \left| \widehat{f}_H^4(\alpha) - \widehat{f}^4(\alpha) \right|$$

$$\leq \sum_{\alpha \in L} |\widehat{f}_H(\alpha) - \widehat{f}(\alpha)| (\left| \sum_{i=0}^{3} \widehat{f}_H(\alpha)^i \widehat{f}(\alpha)^{3-i} \right|)$$

$$\leq 4 \cdot |L| \cdot \max |\widehat{f}_H(\alpha) - \widehat{f}(\alpha)|$$

$$\leq \frac{4}{\gamma^2} \left(\frac{2}{h} + \eta_1 \right).$$

It follows that

$$\left| \sum_{\alpha \in S_H} \widehat{f}_H^4(\alpha) - \sum_{\alpha \in S} \widehat{f}^4(\alpha) \right| \leq \left| \left(A_H - \sum_{\alpha \in L_H} \widehat{f}_H^4(\alpha) \right) - \left(A - \sum_{\alpha \in L} \widehat{f}^4(\alpha) \right) \right|$$

14

$$\leq \quad |A_H - A| + \left| \sum_{\alpha \in L_H} \widehat{f}_H^4(\alpha) - \sum_{\alpha \in L} \widehat{f}^4(\alpha) \right|$$

$$\leq \quad \eta_2 + \frac{16}{h} + \frac{4}{\gamma^2} \left(\frac{2}{h} + \eta_1 \right).$$

<div align="right">□</div>

Now, on the one hand, we have: $\sum_{\alpha \in S} \widehat{f}^4(\alpha) < \gamma^2 \sum_\alpha \widehat{f}^2(\alpha) \leq \gamma^2$. On the other hand, $\max_{\alpha \in S_H} \widehat{f}_H^4(\alpha) \leq \sum_{\alpha \in S_H} \widehat{f}_H^4(\alpha)$. Therefore, combining and using Claim 19, we have:

$$\max_{\alpha \in S_H} \widehat{f}_H^4(\alpha) < \gamma^2 + \eta_2 + \frac{16}{h} + \frac{4}{\gamma^2} \left(\frac{2}{h} + \eta_1 \right)$$

We need to choose the parameters such that $\max_{\alpha \in S_H} |\widehat{f}_H(\alpha)| < \rho_H$, and so, it is enough to have:

$$\gamma^2 + \eta_2 + \frac{16}{h} + \frac{4}{\gamma^2} \left(\frac{2}{h} + \eta_1 \right) < \left(\epsilon - \frac{2}{h} - \eta_1 \right)^4$$

Additionally, we need to ensure that the events we have conditioned on occur with probability at least $2/3$. So, we want:

$$\frac{500}{h\eta_2^2} + \frac{14}{h\gamma^2\eta_1^2} + \frac{1}{\gamma^4 h} < \frac{1}{3}$$

One can check now that the following setting of parameters satisfies both of the above constraints: $\gamma = \epsilon^2/100$, $h = (10/\epsilon)^{20}$, $\eta_1 = (\epsilon/10)^8$, $\eta_2 = (\epsilon/10)^4$. <div align="right">□</div>

5   Concluding Remarks and Open Problems

- The main open question raised here (Question 5) is whether it is possible in general to obtain canonical testers for subspace-hereditary properties with only a polynomial blow up in the query complexity. Here, we show this to be true for OCF, and [BX10] showed the existence of a canonical tester with quadratic blowup for the triangle-freeness property. On the other hand, there is some evidence to the contrary also. Goldreich and Ron in [GR11] proved a nontrivial gap between canonical and non-canonical testers for graph properties. They showed that there exist hereditary graph properties that can be tested using $\tilde{O}(\epsilon^{-1})$ queries but for which the canonical tester requires $\tilde{\Omega}(\epsilon^{-3/2})$ queries. Perhaps, this indicates that for subspace-hereditary properties also, there is a non-trivial, maybe even super-polynomial in this case, gap between non-canonical and canonical testers.

- As previously mentioned, OCF is in fact the only monotone property characterized by freeness from an infinite number of equations (of rank 1). We briefly comment here on the equivalence between all these properties. It is easy to see that even-length equations can be handled trivially. Suppose now that $\mathcal{P}$ is defined by freeness from all equations of length belonging to the infinite set of odd integers $S = \{k_1, k_2, \ldots\}$. Note that OCF $\subseteq \mathcal{P}$. Now suppose $k \notin S$ and k is odd, and let k' be the smallest element of S such that $k \leq k'$. If $f \in \mathcal{P}$ is not free of solutions to the length k equation, then f is not free of solutions to the equation of length k', since a solution $(x_1, \ldots, x_k)$ to the former induces a solution $(x_1, \ldots, x_k, x_1, x_1 \ldots, x_1)$ to the latter.

- Another open problem that arises is to characterize the class of linear-invariant properties that can be tested using $\text{poly}(1/\epsilon)$ queries. For monotone properties that can be characterized by freeness from solutions to a family $\mathcal{F}$ of equations, we conjecture that there is a sharp dichotomy given by whether $\mathcal{F}$ is infinite or finite. It follows from Theorem 3 and the discussion in the previous item that when $\mathcal{F}$ is infinite, the query complexity is $\text{poly}(1/\epsilon)$. When $\mathcal{F}$ is finite and the property is nontrivial, then the property is equivalent to being free of solutions to a single equation $x_1 + \cdots + x_k = 0$ for some odd integer $k > 1$. In this case, we conjecture that the query complexity is super-polynomial, although the current best lower bound is only slightly non-trivial: $\Omega(1/\epsilon^{2.423})$ for testing triangle-freeness [BX10]. For non-monotone properties characterized by freeness from solutions to a family of equations, [CSX11] showed that $(C_3, 110)$-freeness can be testing using $O(1/\epsilon^2)$ queries (recalling the notation in Section 1), but there is no systematic understanding at present of when $\text{poly}(1/\epsilon)$ query complexity is possible for larger equations or for arbitrary intersections of such non-monotone properties. For properties characterized by freeness from solutions to a system of equations of rank greater than one, even less is known.

- Another open problems left open by our results is whether the $\tilde{O}(1/\epsilon^2)$ bound for odd-cycle-freeness is tight. This is indeed the case for bipartiteness testing in graphs [BT04], but a direct analogue of their hard instances does not seem to work in our case.

- One could also ask Question 5 for linear-invariant properties that are not subspace-hereditary. Given a linear-invariant property $\mathcal{P}$, we say that a tester T is canonical for $\mathcal{P}$ if there exists a fixed linear-invariant property $\mathcal{P}'$ (not necessarily the same as $\mathcal{P}$) such that when T is given oracle access to a function $f : \mathbb{F}_2^n \to \{0, 1\}$, it operates by choosing uniformly at random a subspace $H \leq \mathbb{F}_2^n$ and accepting if and only if f restricted to H satisfies the property $\mathcal{P}'$. Notice that unlike the subspace-hereditary case, the canonical tester now need not be one-sided. The stronger form of Question 5 is whether it is the case that for every linear-invariant property $\mathcal{P}$, there exists a canonical tester for $\mathcal{P}$ with query complexity $\text{poly}(q(n, \epsilon))$ whenever $\mathcal{P}$ is testable with query complexity $q(n, \epsilon)$ by some tester. Goldreich and Trevisan [GT03] showed the existence of such a canonical tester with polynomial blowup for graph properties.

References

[ADL+94]   Noga Alon, Richard A. Duke, Hanno Lefmann, Vojtech Rödl, and Raphael Yuster. The algorithmic aspects of the regularity lemma. *J. Algorithms*, 16(1):80–109, 1994.

[AdlVKK03] Noga Alon, W. Fernandez de la Vega, Ravi Kannan, and Marek Karpinski. Random sampling and approximation of MAX-CSPs. *J. Comp. Sys. Sci.*, 67:212–243, September 2003.

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

[AK02]   Noga Alon and Michael Krivelevich. Testing k-colorability. *SIAM J. Discrete Math.*, 15(2):211–227, 2002.

[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.

[Alo02]   Noga Alon. Testing subgraphs in large graphs. *Random Structures and Algorithms*, 21(3-4):359–370, 2002.

[AS08]   Noga Alon and Asaf Shapira. A characterization of the (natural) graph properties testable with one-sided error. *SIAM J. Comput.*, 37(6):1703–1727, 2008.

[BCSX11]   Arnab Bhattacharyya, Victor Chen, Madhu Sudan, and Ning Xie. Testing linear-invariant non-linear properties. *Theory of Computing*, 7(1):75–99, 2011. Earlier version in STACS '09.

[BGS10]   Arnab Bhattacharyya, Elena Grigorescu, and Asaf Shapira. A unified framework for testing linear-invariant properties. In *Proc. 51st Annual IEEE Symposium on Foundations of Computer Science*, pages 478–487. IEEE Computer Society, 2010.

[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.

[BT04]   Andrej Bogdanov and Luca Trevisan. Lower bounds for testing bipartiteness in dense graphs. In *Proc. 19th Annual IEEE Conference on Computational Complexity*, pages 75–81. IEEE Computer Society, 2004.

[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. SIAM, 2010.

[CSX11]   Victor Chen, Madhu Sudan, and Ning Xie. Property testing via set-theoretic operations. In *Proc. 2nd Innovations in Computer Science*, pages 211–222, 2011.

[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.

[GOS+09]   Parikshit Gopalan, Ryan O'Donnell, Rocco A. Servedio, Amir Shpilka, and Karl Wimmer. Testing Fourier dimensionality and sparsity. In *Proc. 36th Annual International Conference on Automata, Languages, and Programming*, pages 500–512. Springer, 2009.

[GR11]   Oded Goldreich and Dana Ron. Algorithmic aspects of property testing in the dense graphs model. *SIAM J. Comput.*, 40(2):376–445, 2011.

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

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

[KKR04]   Tali Kaufman, Michael Krivelevich, and Dana Ron. Tight bounds for testing bipartiteness in general graphs. *SIAM J. Comput.*, 33(6):1441–1483, 2004.

[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. ACM, 2008.

[KSV11]   Daniel Král', Oriol Serra, and Lluís Vena. A removal lemma for systems of linear equations over finite fields. *Israel Journal of Mathematics*, pages 1–15, 2011.

[PRR06]   Michal Parnas, Dana Ron, and Ronitt Rubinfeld. Tolerant property testing and distance approximation. *J. Comp. Sys. Sci.*, 72(6):1012–1042, 2006.

[RS78]   Imre Ruzsa and Endre Szemerédi. Triple systems with no six points carrying three triangles. *Colloq. Math. Soc. Jnos Bolyai (North-Holland, Amsterdam-New York)*, 18:939–945, 1978.

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

[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. ACM, 2009.