Electronic Colloquium on Computational Complexity, Report No. 90 (2013)

Tight Lower Bounds for Testing Linear Isomorphism

Elena Grigorescu [*] Karl Wimmer [†] Ning Xie [‡]

Abstract

We study lower bounds for testing membership in families of linear/affine-invariant Boolean functions over the hypercube. A family of functions $\mathcal{P} \subseteq \{\{0,1\}^n \to \{0,1\}\}$ is *linear/affine invariant* if for any $f \in \mathcal{P}$, it is the case that $f \circ L \in \mathcal{P}$ for any linear/affine transformation L of the domain. Motivated by the recent resurgence of attention to the permutation isomorphism problem, we first focus on families that are linearly/affinely isomorphic to some fixed function. A function $f : \{0,1\}^n \to \{0,1\}$ is called *linear isomorphic* to a fixed Boolean function g if $f = g \circ A$ for some non-singular transformation A.

Our main result is a tight adaptive, two-sided $\Omega(n^2)$ lower bound for testing linear isomorphism to the inner-product function. This is the first lower bound for testing linear isomorphism to a specific function that matches the trivial upper bound. Our proof exploits the elegant connection between testing and communication complexity discovered by Blais *et al.* (Computational Complexity, 2012.) Our results are also the first instance of this connection that gives better than $\Omega(n)$ lower bound for any property of Boolean functions. These results extend to testing linear isomorphism to any fixed function in the larger class of so-called Maiorana-McFarland bent functions.

Our second result shows an $\Omega(2^{n/4})$ query lower bound for any adaptive, two-sided tester for membership in the Maiorana-McFarland class of bent functions. This class of Boolean functions is also affine-invariant and its rich structure and pseudorandom properties have been well-studied in mathematics, coding theory and cryptography.

[*] Department of Computer Science, Purdue University, West Lafayette, IN. Email: `elena-g@purdue.edu`. Research supported in part by NSF grant 1019343 to the Computing Research Association for the CIFellows Project.

[†] Department of Mathematics, Duquesne University, Pittsburgh, PA. Email: `wimmerk@duq.edu`. Supported by NSF award CCF-1117079.

[‡] SCIS, Florida International University, Miami, FL. Email: `nxie@cs.fiu.edu`. Part of the work was done when the author was at CSAIL, MIT and was supported by NSF awards CCF-1217423 and CCF-1065125.

ISSN 1433-8092

1 Introduction

A *property* $\mathcal{P}$ is a set of objects that share some common features. A local test for a property $\mathcal{P}$ is a randomized algorithm which can distinguish inputs that belong to $\mathcal{P}$ from inputs that are very different from every element in $\mathcal{P}$, by making only a few queries to the input. In this work we focus on families of boolean functions $\mathcal{P} \subseteq \{\mathbb{F}_2^n \to \mathbb{F}_2\}$, where $\mathbb{F}_2 = \{0, 1\}$ is the field on two elements. Formally, a (δ, k)-tester for $\mathcal{P}$ is a randomized algorithm that has oracle access to a function $f : \mathbb{F}_2^n \to \mathbb{F}_2$, makes k queries, and accepts w.p. at least 2/3 if $f \in \mathcal{P}$, while rejecting w.p. at least 2/3 if f is δ-far from $\mathcal{P}$. The notion of distance to a property is given by the relative Hamming distance, namely for $f, g : \mathbb{F}_2^n \to \mathbb{F}_2$, $\mathrm{dist}(f, g) = \frac{1}{2^n}|\{x : f(x) \neq g(x)\}|$ and $\mathrm{dist}(f, \mathcal{P}) = \min_{g\in\mathcal{P}} \mathrm{dist}(f, g)$. If a test always accepts function $f \in \mathcal{P}$ it is called *one-sided*, otherwise it is *two-sided*. If the test must send the queries all at once it is called *non-adaptive*, otherwise, namely when the queries could depend on answers to previous queries, the test is *adaptive*.

The field of property testing was pioneered by Blum, Luby, and Rubinfeld [14], Rubinfeld and Sudan [39] and Goldreich, Goldwasser and Ron [24] who introduced two major directions in property testing: testing algebraic properties and testing combinatorial (e.g. graph) properties. To a large extent, the focus of property testing so far has been to characterize what properties admit testers that make only a constant number of queries (these properties are called *strongly testable*). Alon *et al.* [2] and Borgs *et al.* [15] already showed a complete characterization of strongly testable properties of dense graphs. Very recently, Bhattacharyya *et al.* [8] announced a characterization of one-sided strongly testable boolean families that are invariant under "affine" transformations of the domain.

A systematic study of strongly testable properties that are invariant under natural transformations of the domain was first proposed by Kaufman and Sudan [30]. The most studied and at the same time most natural group of invariances for properties defined over structured, discrete objects such as fields or vector spaces are linear and affine transformations of the domain. A linear transformation $L_C : \mathbb{F}_2^n \to \mathbb{F}_2^n$ is a mapping $L_C(x) = Cx$, where $C \in \mathbb{F}_2^{n\times n}$. An affine transformation $L_{C,b} : \mathbb{F}_2^n \to \mathbb{F}_2^n$ is a mapping $L_{C,b}(x) = Cx + b$, where $C \in \mathbb{F}_2^{n\times n}$ and $b \in \mathbb{F}_2^n$. A property $\mathcal{P} \subset \{\mathbb{F}_2^n \to \mathbb{F}_2\}$ is linear-invariant if $f \in \mathcal{P}$ if and only if $f \circ L_C \in \mathcal{P}$, for any linear transformation L_C, where $f \circ L_C(x) = f(L_C(x))$. Similarly, $\mathcal{P}$ is affine-invariant if $f \in \mathcal{P}$ if and only if the function $f \circ L_{C,b} \in \mathcal{P}$, for any affine $L_{C,b}$, where $f \circ L_{C,b}(x) = f(L_{C,b}(x))$. Following [30] linear/affine invariant families have been intensely studied on two fronts: properties that arise in the setting of linear codes [30, 28, 27, 6, 4, 3, 31, 5, 29], and properties that arise more often in the study of boolean functions [26, 32, 40, 9, 8]. All these works study properties that are testable with a constant number of queries.

Here we work in a somewhat complementary direction: we study linear/affine-invariant properties that are hard to test. Partly motivated by the recent resurgence of interest in the permutation isomorphism problem [11, 1, 20, 19, 12, 10], we focus on testing linear/affine isomorphism to a single function. This study, in a certain sense, combines the directions of testing linear/affine-invariance and testing permutation isomorphism. As isomorphism defines an equivalence relation among functions in the family, we restrict our attention to *non-singular* linear/affine transformations in this paper[1] (non-singular transformations lead to permutations of a function while singular transformations do not). More specifically, the orbit of a function f under the set of non-singular

[1] The reason of such a choice is explained later in Remark 1.2.

linear transformations of $\mathbb{F}_2^n$ is given by $\mathcal{L}(f) = \{f \circ L_C | C \in \mathbb{F}_2^{n\times n}, \det(C) = 1\}$; a function g is said to be *linearly isomorphic* to f if $g \in \mathcal{L}(f)$. Similarly, the orbit of f under affine transformations is the family $\mathcal{A}(f) = \{f \circ L_{C,b} | C \in \mathbb{F}_2^{n\times n}, \det C = 1, b \in \mathbb{F}_2^n\}$ and g is *affinely isomorphic* to f if $g \in \mathcal{A}(f)$. For instance, when $f = x_1$ is a dictator function, $\mathcal{L}(f)$ is just the set of non-constant, linear functions, and $\mathcal{A}(f)$ is the set of non-constant, affine functions.

We exhibit a large family of functions for which testing linear/affine isomorphism to every function in the family requires $\Theta(n^2)$ many queries. Our explicit functions arise from families of Boolean *bent* functions. Bent functions are the functions that are the most 'uncorrelated' with linear functions (see Section 2 for the precise definition that we use). A most common example of bent functions, which is also an object of interest in this work, is the inner-product function $\mathrm{IP}_n : \mathbb{F}_2^n \to \mathbb{F}_2$ defined by $\mathrm{IP}_n(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n/2} x_{2i-1}x_{2i}$ (here and in what follows we will assume that n is a multiple of 4).

Bent functions have been well-investigated in mathematics, coding theory and combinatorial design [35, 38, 18, 21, 22] for their rich structure, and in differential cryptography [16, 17] for their pseudorandom and non-linearity properties that make them applicable to building hash-functions; see e.g. [36] for a comprehensive survey. In property testing they were used before in [7] to show lower bounds for testing triangle freeness.

In this work we also show exponential lower bounds for testing membership in the class of bent functions, which is invariant under non-singular affine transformations of the domain. Hence, our results reveal yet some novel uses of bent functions in property testing, suggesting that they have some inherent feature that make them hard for local testing algorithms.

1.1 Our results

Lower bounds for testing linear/affine isomorphisms. We start with the family $\mathcal{P} = \mathcal{L}(\mathrm{IP}_n)$ that has size $|\mathcal{P}| = 2^{O(n^2)}$. We show that the query complexity of the trivial algorithm (which simply picks $O(n^2)$ random inputs and checks if there is a function in $\mathcal{P}$ that agrees with the answers to the queries) is in fact asymptotically optimal, even for *adaptive, 2-sided tests.* Since the query complexity of the single-sided, non-adaptive tester is no less than that of the 2-sided, adaptive tester, this result shows that $\mathcal{L}(\mathrm{IP})$ is an example of a family that is the hardest to test for linear isomorphism.

Theorem 1.1. *Any 2-sided, adaptive $(1/4, k)$-test for $\mathcal{L}(\mathrm{IP}_n)$ and $\mathcal{A}(\mathrm{IP}_n)$ requires $k = \Omega(n^2)$ queries.*

We remark that a lower bound of $\Omega(n)$ follows from the results of Chakraborty *et al.* [20], since the functions in these families have Fourier dimension n, and can be shown to be far from having (Fourier) dimension $n - 1$.

Remark 1.2. Note that since $\mathcal{P} = \mathcal{L}(\mathrm{IP}_n)$ is a collection of polynomials of degree 2, it is a subset of Reed-Muller codes of order 2, RM(2). However, using Dickson's theorem, one can show that, if $\mathcal{L}$ is not restricted to non-singular linear transformations, then the set of functions $\mathcal{L}(\mathrm{IP}_n)$ is identical to RM(2). The latter is well-known to be testable with 8 queries by a single-sided non-adaptive tester. What we show here is that testing a subset of RM(2), $\mathcal{L}(\mathrm{IP}_n)$, is much harder.

We generalize this result to a broader class of bent functions commonly known as Maiorana-McFarland bent functions, denoted $\mathcal{MM}_n$. Formally, a function in this family is $f : \mathbb{F}_2^n \to \mathbb{F}_2$

defined by $f(x,y) = \langle x,y \rangle + g(y)$, where $x, y \in \mathbb{F}_2^{n/2}$ and $g : \mathbb{F}_2^{n/2} \to \mathbb{F}_2$ is arbitrary, and $\langle x,y \rangle$ denotes the inner product (standard dot product) of x and y.

Notice that these functions are no longer low-degree polynomials, and in fact, since g is arbitrary they could be polynomials of degrees as high as $n/2$. For these families too, we show that testing isomorphism is hard.

Theorem 1.3. *Any 2-sided, adaptive $(1/4, k)$-test for $\mathcal{L}(f)$ and $\mathcal{A}(f)$ where f is a Maiorana-McFarland bent function in n variables requires $k = \Omega(n^2)$ queries.*

We remark that the arguments of Alon and Blais from [1] for testing isomorphism under the symmetric group can be easily adapted to testing isomorphism under the group of non-singular transformations to show that testing linear isomorphism to almost all functions requires $\Omega(n)$ queries.

To the best of our knowledge, no superlinear (in n) 2-sided error lower bounds have been previously established for testing any *explicitly given* function under some class of isomorphisms. More generally, let $f : D \to \mathbb{F}_2$ be a function, and G be a group acting on D, where D is some finite domain; (informally, the elements of the group can be identified with bijections from D to itself). The original question of (permutation) isomorphism considers $D = \mathbb{F}_2^n$ and $G = \mathbb{S}_n$, the symmetric group, acting on D in the natural way. Linear isomorphism, which is the focus of this paper, has $D = \mathbb{F}_2^n$ and $G = GL(n, \mathbb{F}_2)$, the group of invertible matrices over $\mathbb{F}_2$, acting on D in the natural way. The easy upper bound for a testing isomorphism algorithm is $\log |G|$, but for technical reasons, many recent lower bounds and their analyses get bottlenecked by $\log |D|$. For permutation isomorphism, this gap is still open. For linear isomorphism, we close this gap in this paper.

In the case that $D = G$ and the group action is simply the group operation, we remark that the proof of [1] can be easily modified to yield the following:

Theorem 1.4. *Let G be a finite group, and choose a random function $f : G \to \{0,1\}$. Testing isomorphism to f under the group action of multiplying by an element of G requires $\Omega(\log |G|)$ queries with high probability.*

We omit the details of this proof. In this case, almost every function requires as many queries for testing isomorphism as the most simple algorithm.

Lower bounds for testing bentness and Maiorana-McFarland families. We then turn to analyzing bent functions in general. The number of boolean bent functions is known to be at least $2^{2^{n/2+\Omega(\log n)}}$ and the current upper bound is larger than $2^{2^{n-1}}$ [42]. We show a lower bound of $\Omega(2^{n/4})$ queries for testing bentness. A lower bound of significantly more than $2^{n/2+\Omega(\log n)}$ queries for the testing problem would improve the current status regarding the number of boolean bent functions, a long standing open problem in mathematics. We note that no non-trivial test (that makes substantially fewer than 2^n queries) for bentness is currently known.

Theorem 1.5. *Any 2-sided, adaptive $(1/4, k)$-test for the class of boolean bent functions in n variables requires $k = \Omega(2^{n/4})$ queries.*

Theorem 1.5 is in fact an immediate consequence of a same lower bound for testing the Maiorana-McFarland family. Let $\mathcal{L}(\mathcal{MM}_n)$ be the linear closure of the family $\mathcal{MM}_n$, i.e. $\mathcal{L}(\mathcal{MM}_n) = \bigcup_{f \in \mathcal{MM}_n} \mathcal{L}(f)$.

Theorem 1.6. *Any 2-sided, adaptive $(1/4, k)$-test for testing membership in $\mathcal{L}(\mathcal{MM}_n)$ requires $k = \Omega(2^{n/4})$ queries.*

We remark that exponential lower bounds for testing affine-invariant properties were known before. For example, testing the Reed-Muller code of degree $n/2$ requires at least $2^{n/2} - 1$ queries, as $2^{n/2}$ is the minimum distance of its dual code.[2] Hence this lower bound is interesting mainly in the context of testing bent functions and their generalization Maiorana-McFarland families. These results appear in Section 4.

Finally, we remark that all our results generalize to non-boolean functions over prime fields $\mathbb{F}_p$ (See the Appendix.)

1.2 Previous related work

Recently there has been intense interest in testing isomorphism of functions (with respect to the symmetric group). Some initial results of this type can be attributed to [37] who showed the strong testability of dictators and monomials. The problem of testing isomorphism was first explicitly introduced by [23]. In recent years the interest in testing function isomorphism has revived [11, 1, 20, 19, 12, 10] prompted by the works of Blais and O'Donnell [11] and Alon and Blais [1]. The main motivation of many of these works (including the original motivation of [23]) involves testing isomorphism to functions with few relevant variables.

Testing linear isomorphism has been less studied. It has been considered by Chakraborty *et al.* [20] who show a lower bound of $\Omega(k)$ for testing $\mathcal{L}(f)$ for a function f that is far from having (Fourier) dimension $k - 1$. In line with testing juntas, a previous result of [25] implicitly proves an upper bound of $O(k2^k)$ for linear isomorphism to functions that are very close to having dimension k; the "very" here is exponentially small in k. Wimmer and Yoshida [43] give an constant-query algorithm for linear isomorphism to any function close to having dimension k by giving a tolerant tester for functions of dimension k. The technique is an extension of the work of [25], and it applies to functions close to having low spectral norm. They also show lower bounds for testing linear isomorphism, but these lower bounds are no better than $\Omega(n)$ for any fixed function.

1.3 Our techniques

Testing linear/affine isomorphisms to IP. Our lower bounds for testing linear/affine isomorphism are proved using reductions from communication complexity protocols, a powerful technique introduced by Blais *et al.* [13]. In the communication complexity model there are two parties holding inputs x and y, respectively, who are trying to compute a function $f(x, y)$ with as little communication between them as possible. In [13], the authors show an ingenious generic technique to prove lower bounds for property testing, by exploiting the strength of the lower bounds obtained in communication complexity.

The crux of our argument is the observation that one can reduce testing linear isomorphism to IP (and more generally, to any Maiorana-McFarland bent function) from the following natural randomized communication protocol: Alice holds the top half of a matrix C, Bob holds the remaining half of C, and their goal is to determine if C is singular. The main feature of bent functions that we make use of here is the fact that when composed with singular linear transformations they not only become functions that are not bent, but they become functions that are *far* from bent (See

[2] We thank an anonymous referee for pointing this out.

Proposition 2.3). To complete the proof we resort to the recent results of Sun and Wang [41] who show a lower bound of $\Omega(n^2)$ for the randomized communication complexity for this problem.

Testing bentness. To show the lower bound for testing bentness we use Yao's principle, where the Yes distribution is the (linear closure of) Maiorana-McFarland family of functions and the No distribution is supported on random $n/2$ dimensional functions. Our argument that these two distributions are indistinguishable resembles the work of [25]. We show that, for any fixed set Q of $\Omega(2^{n/4})$ queries and for most $(n/2)$-dimensional subspaces H, every vector in Q is in a distinct coset of H. This fact is the statement one would expect given the famous "birthday paradox". Our results shows there are at least $\Omega(2^{n/4})$ "degrees of freedom" in selecting a linear transformation of a Maiorana-McFarland function. This lower bound translates upward to the class of all bent functions, since every bent function is far from the class of $(n/2)$-dimensional functions.

2 Preliminaries

Let $n \geq 1$ be a natural number. We use $[n]$ to denote the set $\{1, \ldots, n\}$. $\mathbb{F}_2 = \{0, 1\}$ is the field with 2 elements, where addition and multiplication are performed mod 2. We view elements in $\mathbb{F}_2^n$ as n-bit binary strings – that is elements of $\{0, 1\}^n$ – alternatively. If x and y are two n-bit strings, then $x + y$ (or $x - y$) denotes bitwise addition (i.e. XOR) of x and y. We view $\mathbb{F}_2^n$ as a vector space equipped with an inner product $\langle x, y\rangle$, which we take to be the standard dot product: $\langle x, y\rangle = \sum_{i=1}^n x_i y_i$, where all operations are performed in $\mathbb{F}_2$.

We start by recalling a basic fact that we have referred to earlier.

Theorem 2.1 (folklore). *Let $\mathcal{P} \subseteq \{f : D \to R\}$ (for some finite domain D and range R) be a property of size $|\mathcal{P}|$. Then there is a one-sided error testing algorithm which tests $\mathcal{P}$ with distance parameter ϵ using $O(\frac{1}{\epsilon}\log|\mathcal{P}|)$ queries.*

Proof. The testing algorithm is: on input function f, pick q points in the domain independently and uniformly at random, query f's values at these points, and check if there exists some function in $\mathcal{P}$ that agrees with f on all these points. If so, the algorithm accepts; otherwise rejects. Clearly if $f \in \mathcal{P}$, the algorithm accepts f with probability 1. On the other hand, if f is ϵ-far from $\mathcal{P}$, then for any fixed member $g \in \mathcal{P}$, $\Pr_{\boldsymbol{x}}[f(\boldsymbol{x}) = g(\boldsymbol{x})] \leq 1 - \epsilon$. Therefore, the probability that f and g agree on all independently and randomly chosen q points is at most $(1-\epsilon)^q < 1/(3|\mathcal{P}|)$, if we set $q \geq \frac{c}{\epsilon}\log|\mathcal{P}|$ for some constant c. Finally, applying a union bound over all members in $\mathcal{P}$ gives that the testing algorithm accepts f with probability at most 1/3. □

Linear/affine isomorphism. We say that two functions $f, g : \mathbb{F}_2^n \to \mathbb{F}_2$ are *linearly isomorphic* (or *linear isomorphic*) if there exists a non-singular linear transformation $C : \mathbb{F}_2^n \to \mathbb{F}_2^n$ such that $g(x) = f(Cx)$ for all $x \in \mathbb{F}_2^n$. Equivalently, g is linearly isomorphic to f if and only if there exist n linearly independent linear functions $\ell_i(x) : \mathbb{F}_2^n \to \mathbb{F}_2$ such that $g(x) = f(\ell_1(x), \ell_2(x), \ldots, \ell_n(x))$ (a linear function, is a function of the form $\ell(x) = \sum_{j\in[n]} l_j x_j$, where $l_j \in \mathbb{F}_2$). Similarly, two functions $f, g : \mathbb{F}_2^n \to \mathbb{F}_2$ are *affinely isomorphic* (or *affine isomorphic*) if there exists a non-singular linear transformation $C : \mathbb{F}_2^n \to \mathbb{F}_2^n$ and a vector $b \in \mathbb{F}_2^n$ such that $g(x) = f(Cx + b)$ for all $x \in \mathbb{F}_2^n$. We define $\mathcal{L}(f)$ to be the set of functions linearly isomorphic to f: $\mathcal{L}(f) = \{f \circ L_C | C \in \mathbb{F}_2^{n\times n}, \det(C) = 1\}$. Similarly, we define $\mathcal{A}$ to be the set of functions affinely isomorphic to f: $\mathcal{A}(f) = \{f \circ L_{C,b} | C \in \mathbb{F}_2^{n\times n}, \det C = 1 \text{ and } b \in \mathbb{F}_2^n\}$.

Bent functions. There are many equivalent definitions of bent functions, for example as functions farthest away from any affine functions, or functions whose Fourier coefficients have the same magnitude. Here we will use another standard definition, due to Rothaus [38].

Definition 2.2. *A Boolean function* $f : \mathbb{F}_2^n \to \mathbb{F}_2$ *is bent if for every nonzero vector* $h \in \mathbb{F}_2^n$*, we have* $\Pr_{\boldsymbol{x}}[f(\boldsymbol{x}) = f(\boldsymbol{x} + h)] = 1/2$.

Given a Boolean function $f : \mathbb{F}_2^n \to \mathbb{F}_2$, we define $\mathrm{Inv}(f) := \{h : f(x) = f(x + h) \text{ for all } x\}$. The set $\mathrm{Inv}(f)$ forms a subspace of $\mathbb{F}_2^n$, and we define the *dimension* of f to be the codimension of $\mathrm{Inv}(f)$. We use $\dim(f)$ to denote the dimension of f. If $\dim(f) \le k$, we say that f is k-dimensional. This notion of dimensionality is equivalent to the notion of Fourier dimension used in [25]. From their definition it immediately follows that the dimension of any bent function $f : \mathbb{F}_2^n \to \mathbb{F}_2$ is n.

The following proposition will be of great importance to us.

Proposition 2.3. *Suppose* $f : \mathbb{F}_2^n \to \mathbb{F}_2$ *is a bent function and* $\dim(g) < n$*. Then* $\Pr_{\boldsymbol{x}}[f(\boldsymbol{x}) \neq g(\boldsymbol{x})] \ge 1/4$.

Proof. Since $\dim(g) < n$, we have $\mathrm{Inv}(g) \neq \{0\}$, so there exists a nonzero vector $h \in \mathbb{F}_2^n$ such that $g(x) = g(x + h)$ for all $x \in \mathbb{F}_2^n$. From Proposition 2.2, we know that $\Pr_{\boldsymbol{x}}[f(\boldsymbol{x}) \neq f(\boldsymbol{x} + h)] = 1/2$. We have

$$\begin{aligned}
\Pr_{\boldsymbol{x}}[f(\boldsymbol{x}) \neq g(\boldsymbol{x})] &= \frac{1}{2}(\Pr_{\boldsymbol{x}}[f(\boldsymbol{x}) \neq g(\boldsymbol{x})] + \Pr_{\boldsymbol{x}}[f(\boldsymbol{x} + h) \neq g(\boldsymbol{x} + h)]) \\
&= \frac{1}{2}(\Pr_{\boldsymbol{x}}[f(\boldsymbol{x}) \neq g(\boldsymbol{x})] + \Pr_{\boldsymbol{x}}[f(\boldsymbol{x} + h) \neq g(\boldsymbol{x})]) \\
&\ge \frac{1}{2}(\Pr_{\boldsymbol{x}}[f(\boldsymbol{x}) \neq f(\boldsymbol{x} + h)] = 1/4.
\end{aligned}$$

□

One well-known class of bent functions are the *Maiorana-McFarland* functions defined as follows. Let $x \in \mathbb{F}_2^{n/2}$, $y \in \mathbb{F}_2^{n/2}$, and let $g : \mathbb{F}_2^{n/2} \to \mathbb{F}_2$ be any Boolean function on $n/2$ variables. Then $\mathrm{MM}_n^g : \mathbb{F}_2^n \to \mathbb{F}_2$ given by $\mathrm{MM}_n^g(x, y) = \langle x, y \rangle + g(y)$ is called a *Maiorana-McFarland function.*

As mentioned before, a formula for the number of bent functions on n variables is unknown. It is worth noting that there are $2^{2^{n/2}}$ distinct bent Maiorana-McFarland functions, so this class accounts for a significant portion of the bent functions known.

We next list a few basic facts about bent functions that will be useful to us later.

Lemma 2.4.

1. *The inner product function* IP_n *is bent.*

2. *Any Maiorana-McFarland function* MM_n^g *is bent.*

3. *If* $f : \mathbb{F}_2^n \to \mathbb{F}_2$ *is bent,* $C \in \mathbb{F}_2^{n \times n}$ *is non-singular and* $b \in \mathbb{F}_2^n$*, then* $f \circ L_C$ *and* $f \circ L_{C,b}$ *are bent.*

4. *If* $f : \mathbb{F}_2^n \to \mathbb{F}_2$ *is bent,* $C \in \mathbb{F}_2^{n \times n}$ *is singular and* $b \in \mathbb{F}_2^n$*, then* $f \circ L_C$ *and* $f \circ L_{C,b}$ *are* $1/4$*-far from bent.*

Proof. Recall that $\mathrm{IP}_n(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n/2} x_{2i-1}x_{2i}$ and $\mathrm{MM}_n^g = \sum_{i=1}^{n/2} x_i x_{i+n/2} + g(x_{n/2+1}, \ldots, x_n)$. First note that $\mathrm{IP}_n = \mathrm{MM}_n^g \circ L_C$, where $g : \mathbb{F}_2^{n/2} \to \mathbb{F}_2$, $g = 0$ and L_C is the full-rank linear transformation that performs the following change of variables: $x_i \mapsto x_{2i-1}$ and $x_{n/2+i} \mapsto x_{2i}$ for $1 \leq i \leq n/2$. Therefore, Item 1 follows from Item 2 and Item 3, and we will show now the latter two. The proof follows from the simple observation that if ℓ is a nonzero linear function $\ell(x) = \langle a, x \rangle$ and $b \in \mathbb{F}_2$ then $\Pr[\ell(x) = b] = 1/2$. We have that for $x, y, h_1, h_2 \in \mathbb{F}_2^{n/2}$

$$\begin{aligned} & \Pr_{x,y}[\langle x, y \rangle + g(y) = \langle x + h_1, y + h_2 \rangle + g(y + h_2)] \\ = \; & \Pr_{x,y}[\langle x, h_2 \rangle = g(y) + g(y + h_2) + \langle h_1, y \rangle + \langle h_1, h_2 \rangle] = 1/2 \end{aligned}$$

Item 3 follows by the observation that if $f : \mathbb{F}_2^n \to \mathbb{F}_2$ is bent and C is non-singular then $Cx = \mathbb{F}_2^n$ and so the functions $f \circ L_C$ and $f \circ L_{C,b}$ are permutations of f. Finally, if C is singular, there exists $h \in \mathbb{F}_2^n$, $h \neq 0$ such that $Ch = 0$ and so $Cx = C(x + h), \forall x \in \mathbb{F}_2^n$ implying that $\dim(f \circ L_C) < n$ and similarly $\dim(f \circ L_{C,b}) < n$. The distance to bentness now follows from Proposition 2.3. □

Basic communication complexity facts. In communication complexity Alice holds an input x and Bob holds an input y and they want to compute a function $f(x, y)$ by exchanging a small number of bit messages. A randomized protocol with ϵ error for computing f is an algorithm whose random bits are known to both the players, and which outputs $f(x, y)$ for any x, y w.p. at least $1 - \epsilon$ over the choice of random bits. In this paper we will fix $\epsilon = 1/3$. The complexity of the protocol is the maximum over all x, y of the number of bits exchanged by Alice and Bob. The number of random bits used in the protocol does not affect the complexity measure of the protocol. For a comprehensive survey on communication complexity, see [34].

3 Lower Bounds for Testing Isomorphism to Inner-Product and Related Functions

We prove Theorem 1.1 and Theorem 1.3 in this section.

As previously mentioned, the main idea of our proofs is an adaption of a technique of proving property testing query lower bounds via communication lower bounds (See Lemma 2.4 of [13]) to the setting of testing linear isomorphisms. Specifically, Alice and Bob are each given half of a linear transformation matrix C, and their goal is to determine if C is singular or not. They can apply their halves, A and B respectively, of matrix C to an arbitrary input x to the inner product function $\mathrm{IP}_n(x)$. Using the fact that $\mathrm{IP}_n(x_1, x_2, \ldots, x_n) = \mathrm{IP}_{n/2}(x_1, \ldots, x_{n/2}) + \mathrm{IP}_{n/2}(x_{n/2+1}, \ldots, x_n)$, Alice computes $\mathrm{IP}_{n/2}(Ax)$, Bob computes $\mathrm{IP}_{n/2}(Bx)$, and both exchange their bits. Now Alice and Bob both know $\mathrm{IP}_n(Cx) = \mathrm{IP}_n(Ax, Bx) = \mathrm{IP}_{n/2}(Ax) + \mathrm{IP}_{n/2}(Bx)$.

Clearly, if the matrix C has full rank, then $\mathrm{IP}_n(Cx) \in \mathcal{L}(\mathrm{IP}_n)$; on the other hand, one can show that if C does not have full rank, then $\mathrm{IP}_n(Cx)$ is far from $\mathcal{L}(\mathrm{IP}_n)$. Therefore, if there is a tester for linear isomorphism of IP_n with q queries, then one can turn such a tester into a communication protocol for Alice and Bob to determine if C has full rank or not, using at most $2q$ bits of communication. The lower bound of Sun and Wang [41] implies that $2q = \Omega(n^2)$, finishing the proof. We formally state their lower bound here.

Theorem 3.1 (Theorem 3, [41])**.** *The randomized communication complexity of computing* $\det(A + B)$*, where Alice holds the matrix* $A \in \mathbb{F}_2^{n \times n}$ *and Bob holds the matrix* $B \in \mathbb{F}_2^{n \times n}$ *is* $\Omega(n^2)$.

We will use two corollaries of this result, the first of which is implicit in [41].

Corollary 3.2. *Let A and B be matrices in $\mathbb{F}_2^{n/2\times n}$ such that the last $n/2$ columns form a basis for $\mathbb{F}_2^{n/2}$. Let C be the matrix $\begin{bmatrix} A \\ B \end{bmatrix}$. The randomized communication complexity of computing* $\det(C)$ *where A is held by Alice and B is held by Bob is $\Omega(n^2)$.*

Proof. By assumption, both Alice and Bob can reduce A and B to $A' = [A'', I_{n/2\times n/2}]$ and $B' = [B'', I_{n/2\times n/2}]$, where $I_{n/2\times n/2}$ is the identity matrix. Then it can be checked that

$$\det(C) = \det\left(\begin{bmatrix} A \\ B \end{bmatrix}\right) = \det\left(\begin{bmatrix} A' \\ B' \end{bmatrix}\right) = \det\left(A'' + B''\right),$$

which needs $\Omega(n^2)$ bits of communication by Theorem 3.1. □

Corollary 3.3. *Let $C \in \mathbb{F}_2^{n\times n}$ be that $C = \begin{bmatrix} A_{n/4\times n/2} & 0_{n/4\times n/2} \\ B_{n/4\times n/2} & 0_{n/4\times n/2} \\ 0_{n/2\times n/2} & I_{n/2\times n/2}. \end{bmatrix}$, where Alice holds A, Bob holds B and matrices A and B are under the same assumptions as in Corollary 3.2. Then the randomized communication complexity of computing* $\det(C)$ *is $\Omega(n^2)$.*

Proof. The statement follows from Corollary 3.2 by noticing that $\det(C) = \det\left(\begin{bmatrix} A \\ B \end{bmatrix}\right)$. □

As mentioned before, it will be convenient to think of the rows of a linear transformation $C \in \mathbb{F}_2^{n\times n}$ as a sequence of linear maps $\ell_1, \ell_2, \ldots, \ell_n : \mathbb{F}_2^n \to \mathbb{F}_2$, where $\ell_i(x) = \sum_{j=1}^{n} a_{ij}x_j$, and so $\mathrm{IP}_n(Cx) = \sum_{i=1}^{n/2} \ell_{2i-1}(x)\ell_{2i}(x)$. Similarly, an affine transformation (C, b) of IP_n can be represented by $\mathrm{IP}_n(Cx + b) = \sum_{i=1}^{n/2}(\ell_{2i-1}(x) + b_{2i-1})(\ell_{2i}(x) + b_{2i})$.

3.1 Proof of Theorem 1.1

We are now ready to complete the proof of Theorem 1.1 by formalizing the reduction to testing $\mathcal{L}(\mathrm{IP}_n)$ ($\mathcal{A}(\mathrm{IP}_n)$, respectively) from computing $\det(A + B)$ as in Theorem 3.1.

Lemma 3.4. *Suppose there exists a 2-sided, adaptive $(1/4, k)$-test for $\mathcal{L}(\mathrm{IP}_n)$ (or $\mathcal{A}(\mathrm{IP}_n)$, respectively), then there exists a randomized communication protocol with public coins, error $1/3$, and communication complexity $O(k)$ for computing* $\det(C)$*, where Alice holds the top half $A \in \mathbb{F}_2^{n/2\times n}$ of C and Bob holds the bottom half $B \in \mathbb{F}_2^{n/2\times n}$ of C.*

Proof. We will show the reduction to $\mathcal{L}(\mathrm{IP}_n)$ only, as the reduction to $\mathcal{A}(\mathrm{IP}_n)$ follows from a very similar argument. Let $C = \begin{bmatrix} A \\ B \end{bmatrix}$ and let $\mathcal{T}$ be a $(1/4, k)$-tester for $\mathcal{L}(\mathrm{IP}_n)$. We will use it to construct a communication protocol for $\det(C)$. Let $\ell_1, \ldots, \ell_{n/2}$ be the linear forms representing the rows of A and $\ell_{n/2+1}, \ldots, \ell_n$ be the linear forms corresponding to the rows of B. Let $f(x) = \mathrm{IP}_n(Cx) = \sum_{i=1}^{n/2} \ell_{2i-1}(x)\ell_{2i}(x)$.

Claim 3.5. *If* $\det(C) = 0$ *then* $f(x) = \mathrm{IP}_n(Cx)$ *is* $1/4$*-far from* $\mathcal{L}(\mathrm{IP}_n)$*.*

Proof. By Items 1 and 3 of Lemma 2.4, every function in $\mathcal{L}(\mathrm{IP}_n)$ is bent. By Item 4 of Lemma 2.4 $f(x) = \mathrm{IP}_n(Cx)$ is 1/4-far from $\mathcal{L}(\mathrm{IP}_n)$. □

In other words, if $\det(C) = 1$ then $f \in \mathcal{L}(\mathrm{IP}_n)$; and if $\det(C) = 0$ then f is 1/4-far from $\mathcal{L}(\mathrm{IP}_n)$.

Let $q_1, q_2, \ldots, q_k$ be the set of (possibly adaptive) queries performed by the tester $\mathcal{T}$ on input f. The protocol for Alice and Bob is to communicate in k rounds. Since Alice and Bob have access to unlimited shared randomness and they exchange bits after generating each query q_i, we can assume that both of them know q_{i+1} given $q_1, q_2, \ldots, q_i$ and $f(q_1), f(q_2), \ldots, f(q_i)$. (Initially, Alice and Bob both know q_1.)

We claim that the following protocol computes $\det(C)$ with probability at least 2/3. Alice computes Aq_1, namely $\ell_1(q_1), \ldots, \ell_{n/2}(q_1)$ and sends to Bob $\sum_{i=1}^{n/4} \ell_{2i-1}(q_1)\ell_{2i}(q_1)$. Bob computes Bq_1 and also $\sum_{i=n/4+1}^{n/2} \ell_{2i-1}(q_1)\ell_{2i}(q_1)$. Using Alice's bit he can now simulate the query $f(q_1)$ by now computing $\mathrm{IP}_n(Cq_1) = \sum_{i=1}^{n/2} \ell_{2i-1}(q_1)\ell_{2i}(q_1)$, and Bob can send $f(q_1)$ to Alice. By repeating this protocol for the remaining queries $q_2, \ldots, q_k$, Bob can finally output the answer that the tester $\mathcal{T}$ would output on f. Since $\mathcal{T}$ succeeds w.p. at least 2/3 on f, it follows that the protocol correctly computes $\det(C)$ w.p. at least 2/3. Notice that the total number of bits exchanged is $O(k)$. □

Proof of Theorem 1.1. Suppose $\mathcal{T}$ is a 2-sided, adaptive $(1/4, k)$-test for $\mathcal{L}(\mathrm{IP}_n)$ (or $\mathcal{A}(\mathrm{IP}_n)$ respectively). Then, in particular, it should distinguish functions $f = \mathrm{IP}_n(Cx)$ (here C is a matrix $C = \begin{bmatrix} A \\ B \end{bmatrix}$ with $A, B \in \mathbb{F}_2^{n/2 \times n}$ such that their last $n/2$ columns form a basis for $\mathbb{F}_2^{n/2}$) from functions that are 1/4-far from $\mathcal{L}(\mathrm{IP}_n)$. By Lemma 3.4, there exists a communication protocol computing $\det(C)$ of complexity $O(k)$. Finally, by Corollary 3.2 it must be that $k = \Omega(n^2)$. □

3.2 Proof of Theorem 1.3

The reduction from Lemma 3.4 can be tweaked to work for the much more general class of Maiorana-McFarland bent functions. Recall that every function in the Maiorana-McFarland $\mathcal{MM}_n$ family of bent functions can be expressed as $\mathrm{MM}_n^g(x) = \sum_{i=1}^{n/2} x_i x_{i+n/2} + g(x_{n/2+1}, \ldots, x_n)$ for some $g : \mathbb{F}_2^{n/2} \to \mathbb{F}_2$, and so $\mathcal{L}(\mathrm{MM}_n^g) = \{\mathrm{MM}_n^g(Ax) |\ A \in \mathbb{F}^{n \times n}, \det(A) = 1\}$.

Our reduction in the previous section can not be directly used, since Alice would have half of the (linear functions acting as) inputs to g and Bob would have the other half. Thus, answering queries might require more than constant communication, degrading the lower bound. In this case, we reduce from a scenario where Alice and Bob will both always know the inputs to g; this preserves the lower bound of $\Omega(n^2)$. The reduction now uses matrices of the special form described in Corollary 3.3.

We note that $\mathrm{MM}_n^0 = \mathrm{IP}_n$, where 0 is the constant 0 function, but our previous reduction to inner product is not equivalent to the following reduction setting $g = 0$.

Lemma 3.6. *Suppose there exists a 2-sided, adaptive $(1/4, k)$-test for $\mathcal{L}(f)$ (or $\mathcal{A}(f)$, respectively), where $f \in \mathcal{MM}_n$. Then there exists a randomized communication protocol with public coins, error at most 1/3 and communication complexity $O(k)$ for computing* $\det\left(\begin{bmatrix} A_{n/4 \times n/2} & 0_{n/4 \times n/2} \\ B_{n/4 \times n/2} & 0_{n/4 \times n/2} \\ 0_{n/2 \times n/2} & I_{n/2 \times n/2}. \end{bmatrix}\right)$, *where Alice holds A and Bob holds B.*

Proof. Once again, we will only show the proof assuming the existence of a tester for $\mathcal{L}$, since the remaining part of the proof follows by similar arguments.

Let $\mathcal{T}$ be a $(1/4, k)$-tester for $\mathcal{L}(\mathrm{MM}_n^g)$, where $\mathrm{MM}_n^g \in \mathcal{M}$. Let f be the function defined by $f(x) = \mathrm{MM}_n^g(Cx)$. Alice and Bob can simulate queries to the function f as before. Let $q_1, \ldots, q_k$ be a set of (possibly adaptive) queries that $\mathcal{T}$ would make on f. As before, the protocol runs in k rounds, Alice and Bob have unlimited shared randomness, and we may assume that both of them know q_{i+1} given $q_1, \ldots, q_i$ and $f(q_1), f(q_2), \ldots, f(q_i)$.

As before, we view the rows of C as linear transformations $\ell_1, \ell_2, \ldots, \ell_n$. The last $n/2$ rows of C are known to both Alice and Bob, so each of them know $\ell_{n/2+1}, \ldots, \ell_n$. Each of these transformation is a projection on to a single coordinate.

Alice computes $[A \;\; 0]q_1$, namely $\ell_1(q_1), \ldots, \ell_{n/4}(q_1)$ and then she sends to Bob $\sum_{i=1}^{n/4} \ell_i(q_1)\ell_{i+n/2}(q_1)$. Bob computes $[B \;\; 0]q_1$, namely $\ell_{n/4+1}(q_1) \ldots, \ell_{n/2}(q_1)$, and uses the result to compute the bit $\sum_{i=n/4+1}^{n/2} \ell_i(q_1)\ell_{i+n/2}(q_1)$. Now Bob can simulate the query $f(q_1)$ by computing

$$\mathrm{MM}_n^g(Cq_1) = \sum_{i=1}^{n/2} \ell_i(q_1)\ell_{i+n/2}(q_1) + g(\ell_{n/2+1}(q_1), \ell_{n/2+2}(q_1), \ldots, \ell_{n-1}(q_1), \ell_n(q_1)).$$

He can then send this bit back to Alice. After simulating all the queries Bob can output the output of $\mathcal{T}$ on f when the test performed these queries. If $\det(C) = 1$ then $f \in \mathcal{L}(\mathrm{MM}_n^g)$ and the test accepts w.p. at least 2/3, and otherwise, by Lemma 2.4 f is 1/4-far from $\mathcal{L}(\mathrm{MM}_n^g)$ and the test rejects w.p. at least 2/3. Therefore, the communication protocol succeeds w.p. at least 2/3. □

The proof of Theorem 1.3 follows by a similar argument as in the proof of Theorem 1.1 but where now we use Lemma 3.6 and Corollary 3.3 instead.

4 Testing Linear Isomorphism to the Class of Maiorana-McFarland Bent Functions

Our lower bounds will be established via Yao's minimax principle. We denote the total variation distance between two distributions D_1 and D_2 as $\|D_1 - D_2\|_{TV} := \frac{1}{2}\sum_x |D_1(x) - D_2(x)|$, and our goal is to show that the query responses over a "yes" distribution and a "no" distribution are close in total variation distance.

We define D_{YES} to be the uniform distribution over $\mathcal{L}(\mathcal{MM}_n)$, and D_{NO} to be the uniform distribution over $(n/2)$-dimensional functions. We remind the reader that every function in the support of D_{YES} is $(1/4)$-far from every function in D_{NO}. In fact, since by Proposition 2.3 every bent function is $(1/4)$-far from every function in D_{NO}, this same argument establishes a lower bound for testing bentness[3], and thus prove Theorem 1.5.

We can simulate random draws from D_{YES} and D_{NO} in the following way. In both experiments, we pick a random function $g : \mathbb{F}_2^n \to \mathbb{F}_2$ and a random full rank $n/2 \times n$ matrix A_b. A draw $f \sim D_{\mathrm{NO}}$ is the function defined by $f(x) = g(A_b x)$, where a draw $f \sim D_{\mathrm{YES}}$ is the function defined by $f(x) = IP_n(\begin{bmatrix} A_t \\ A_b \end{bmatrix} x) + g(A_b x)$, where A_t is a random full rank $n/2 \times n$ matrix such that $\begin{bmatrix} A_t \\ A_b \end{bmatrix}$

[3] The 1/4 can be replaced by any positive constant less than 1/2; we omit the details.

is nonsingular (chosen dependently on A_b, the distribution on A_t given A_b will not matter as long as the nonsingularity condition is satisfied). Our approach here is very reminiscent of the approach in [25] for showing a lower bound for testing functions of Fourier dimension k.

In the following, when we refer to a random matrix in $\mathbb{F}_2^{n/2\times n}$, we mean a matrix whose entries are chosen to be 0 or 1 independently and uniformly at random.

Lemma 4.1. *Let A_b be a random matrix in $\mathbb{F}_2^{n/2\times n}$. Then A_b is full rank (in the $\mathbb{F}_2$ sense) except with probability at most $(n/2)2^{-n/2}$.*

Proof. The probability that A_b is full rank is $\displaystyle\prod_{i=1}^{n/2}(1-2^{i-n}) \leq (1-2^{-n/2})^{n/2} \leq 1-(n/2)2^{-n/2}$. □

Because this probability is subconstant, for the sake of conciseness we will treat this event as always occurring. From now on, we will assume A_b has rank $n/2$ with certainty.

Lemma 4.2. *Let q_1 and q_2 be two distinct vectors in $\mathbb{F}_2^n$, and A_b be a random matrix in $\mathbb{F}_2^{n/2\times n}$. Then $\Pr_{A_b}[A_bq_1 = A_bq_2] = 2^{-n/2}$.*

Proof. The event is equivalent to $A_b(q_1-q_2) = 0$, where q_1-q_2 is a fixed nonzero vector. Since q_1-q_2 is nonzero and the rows of A_b are chosen independently and uniformly from $\mathbb{F}_2^n$, the distribution over $A_b(q_1-q_2)$ is uniform. Thus, $\Pr_{A_b}[A_b(q_1-q_2) = 0] = \Pr_{A_b}[A_bq_1 = A_bq_2] = 2^{-n/2}$. □

Lemma 4.3. *Let Q be a set of $k = 2^{n/4}/10$ vectors in $\mathbb{F}_2^n$. Let A_b be a random $\{0,1\}$ matrix of dimensions $n/2\times n$. Then $\Pr_{A_b}[\exists q_1, q_2 \in Q \text{ such that } A_bq_1 = A_bq_2] \leq 1/100$.*

Proof. We use the previous lemma and the union bound. There are at most $\binom{k}{2} \leq k^2 = 2^{n/2}/100$ pairs of vectors, and

$$\Pr_{A_b}[\exists q_1, q_2 \in Q \text{ such that } A_bq_1 = A_bq_2] \leq \sum_{q_1,q_2\in Q} \Pr_{A_b}[A_bq_1 = A_bq_2] \leq k^2 2^{-n/2} = 1/100.$$

□

Lemma 4.4. *Let f be a random draw from D_{YES}. Let $Q = \{q_1, q_2, \ldots, q_k\}$ be a set of $k = 2^{n/4}/10$ queries. Now, if the conditions in Lemmas 4.1 and 4.3 hold, then the vector $[f(q_1), f(q_2), \ldots, f(q_k)]$ is uniformly distributed.*

Proof. We choose a random matrix $A_b \in \mathbb{F}_2^{n/2\times n}$, and we extend to a full rank matrix A uniformly over all possible choices. Assuming the event from Lemma 4.3 holds, the vectors A_bq_i are all distinct. Since g is a uniformly random function, the values of $g(A_bq_i)$ are all independent and uniformly distributed, and it follows that the values of $f(q_i)$ are all independent and uniformly distributed as well. □

Lemma 4.5. *Let f be a random draw from D_{NO}. Let Q be a set of $2^{n/4}/10$ queries. If the condition in Lemma 4.3 holds, the answers to the queries are uniformly distributed.*

Proof. Essentially the same as the latter portion of Lemma 4.4. □

In order to prove a lower bound for adaptive testers, we can't assume that Q is a fixed query set, since for example q_2 depends on $f(q_1)$. A deterministic adaptive k-query algorithm is equivalent to a decision tree T of depth at most k. The internal nodes of T are labeled by query strings, and the leaves are labeled by "accept" and "reject". However, the best labeling of the leaves is easy to discuss. Given a decision tree T with unlabeled leaves, the best distinguisher one can get by labeling the leaves is exactly $\|L_{\text{YES}} - L_{\text{NO}}\|_{TV}$; this is the result of labeling every leaf v with "accept" if $L_{\text{YES}}(v) > L_{\text{NO}}(v)$ and "reject" otherwise. As in [25], we define L_{YES} and L_{NO} to be distribution on leaves of T induced by a draw from D_{YES} and D_{NO}, respectively.

We fix a deterministic adaptive tester making at most k queries; equivalently, we fix a decision tree T of depth $k \leq 2^{n/4}/10$. Without loss of generality, we can assume that no string appears twice on any root-to-leaf path and the depth of every path is exactly k. It suffices to prove $\|L_{\text{YES}} - L_{\text{NO}}\|_{TV} \leq 1/3$.

Define L_{UNIF} to be the uniform distribution over the leaves of T. Consider a draw $f \sim D_{\text{YES}}$. Drawing A_b is the same as drawing a random $(n/2)$-dimensional subspace of $\mathbb{F}_2^n$. Consider the strings on nodes of a root-to-leaf path in T ending at the leaf v. By Lemma 4.4, all the strings on this path lie in different buckets, except with probability at most $1/100$ over the choice of A_b. Conditioned on this happening, the probability that f is consistent with the root-to-leaf path to v is exactly 2^{-k}, since g is a drawn uniformly at random. Thus, for each leaf v, we have $\Pr_{L_{\text{YES}}}[v \text{ is reached}] \geq (1 - 1/100)2^{-k}$. A similar argument shows that $\Pr_{L_{\text{NO}}}[v \text{ is reached}] \geq (1 - 1/100)2^{-k}$ as well.

The following lemma essentially appears in [25]:

Lemma 4.6. *Let D be a distribution over $\mathbb{F}_2^m$ that becomes the uniform distribution $\mathcal{U}$ conditioned on an event that happens with probability at least* $99/100$. *Then $\|D - \mathcal{U}\|_{TV} \leq 1/100$, where $\mathcal{U}$ is the uniform distribution over $\mathbb{F}_2^m$.*

Proof. Due to the conditioning, each element of $\mathbb{F}_2^m$ has probability mass at least $(99/100)2^{-m}$, so the elements with probability mass less than 2^{-m} contribute at most $1/2(1/100) = 1/200$ in total to the total variation distance. This lower bounding already takes up $99/100$ of the probability mass, so the elements with probability mass at least 2^{-m} contribute at most the remaining $1/2(1/100) = 1/200$ to the total variation distance. Thus $\|D - \mathcal{U}\|_{TV} \leq (1/2)(1/100 + 1/100) = 1/100$. □

Theorem 4.7. *Any $(1/4, k)$-tester for testing membership in $\mathcal{L}(\mathcal{MM}_n)$ requires $2^{n/4}/10$ queries; that is, $k \geq 2^{n/4}/10$. This lower bounds holds for two-sided adaptive testers.*

Proof. The proof is via Yao's minimax principle. Let T be a decision tree of depth $2^{n/4}/10$ representing an adaptive deterministic tester, and let L_{YES} and L_{NO} be the distributions of leaves obtained by taking random draws from D_{YES} and D_{NO} respectively. The distributions L_{YES} and L_{NO} satisfy the conditions of Lemma 4.6 (by Lemmas 4.4 and 4.5), so $\|L_{\text{YES}} - L_{\text{UNIF}}\|_{TV} \leq 1/100$ and $\|L_{\text{NO}} - L_{\text{UNIF}}\|_{TV} \leq 1/100$. By the triangle inequality, $\|L_{\text{YES}} - L_{\text{NO}}\|_{TV} \leq 2/100 < 1/3$. It follows that the two distributions can not be distinguished using the adaptive tester characterized by decision tree T. □

We can now prove Theorem 1.6:

Proof of Theorem 1.6. By Lemma 2.4, every function in $\mathcal{L}(\mathcal{MM}_n)$ is bent. By Proposition 2.3, every bent function is $(1/4)$-far from every $n/2$-dimensional function. Thus, we can use Yao's minimax principle and the same distributions D_{YES} and D_{NO} as before to establish the theorem. The result now follows from mimicking the proof of Theorem 4.7. □

Acknowledgments

We are grateful to Eric Blais, Amit Chakrabarti, Xiaoming Sun, and David Woodruff for helpful discussions. We would like to thank the anonymous referees for their comments and suggestions.

References

[1] N. Alon and E. Blais. Testing boolean function isomorphism. In *Proc. 14th International Workshop on Randomization and Computation*, pages 394–405, 2010.

[2] Noga Alon, Eldar Fischer, Ilan Newman, and Asaf Shapira. A combinatorial characterization of the testable graph properties: It's all about regularity. *SIAM J. Comput.*, 39(1):143–167, 2009.

[3] Eli Ben-Sasson, Elena Grigorescu, Ghid Maatouk, Amir Shpilka, and Madhu Sudan. On sums of locally testable affine invariant properties. In *APPROX-RANDOM*, pages 400–411, 2011.

[4] Eli Ben-Sasson, Ghid Maatouk, Amir Shpilka, and Madhu Sudan. Symmetric LDPC codes are not necessarily locally testable. In *CCC*, pages 55–65, 2011.

[5] Eli Ben-Sasson, Noga Ron-Zewi, and Madhu Sudan. Sparse affine-invariant linear codes are locally testable. In *FOCS*, pages 561–570, 2012.

[6] Eli Ben-Sasson and Madhu Sudan. Limits on the rate of locally testable affine-invariant codes. In *APPROX-RANDOM*, pages 412–423, 2011.

[7] A. Bhattacharyya and N. Xie. Lower bounds on testing triangle-freeness in Boolean functions. In *Proc. 21st ACM-SIAM Symposium on Discrete Algorithms*, pages 87–98, 2010.

[8] Arnab Bhattacharyya, Eldar Fischer, Hamed Hatami, and Pooya Hatami Shachar Lovett. Every locally characterized affine-invariant property is testable. *STOC*, (to appear), 2013.

[9] Arnab Bhattacharyya, Elena Grigorescu, and Asaf Shapira. A unified framework for testing linear-invariant properties. In *FOCS*, pages 478–487, 2010.

[10] E. Blais and D. Kane. Tight bounds for testing k-linearity. In *Proc. 16th International Workshop on Randomization and Computation*, pages 435–446, 2012.

[11] E. Blais and R. O'Donnell. Lower bounds for testing function isomorphism. In *Proc. 25th Annual IEEE Conference on Computational Complexity*, pages 235–246, 2010.

[12] E. Blais, A. Weinstein, and Y. Yoshida. Partially symmetric functions are efficiently isomorphism-testable. In *FOCS*, pages 551–560, 2012.

[13] Eric Blais, Joshua Brody, and Kevin Matulef. Property testing lower bounds via communication complexity. *Computational Complexity*, 21(2):311–358, 2012.

[14] M. Blum, M. Luby, and R. Rubinfeld. Self-testing/correcting with applications to numerical problems. *Journal of Computer and System Sciences*, 47:549–595, 1993.

[15] C. Borgs, J. T. Chayes, L. Lovász, V. Sós, B. Szegedy, and K. Vesztergombi. Graph limits and parameter testing. In *STOC*, pages 261–270, 2006.

[16] Lilya Budaghyan and Claude Carlet. Ccz-equivalence of bent vectorial functions and related constructions. *Des. Codes Cryptography*, 59(1-3):69–87, 2011.

[17] Claude Carlet. Relating three nonlinearity parameters of vectorial functions and building apn functions from bent functions. *Des. Codes Cryptography*, 59(1-3):89–109, 2011.

[18] Claude Carlet and Philippe Guillot. An alternate characterization of the bentness of binary functions, with uniqueness. *Des. Codes Cryptography*, 14(2):133–140, 1998.

[19] S. Chakraborty, E. Fischer, D. García-Soriano, and A. Matsliah. Junto-symmetric functions, hypergraph isomorphism and crunching. In *CCC*, pages 148–158, 2012.

[20] S. Chakraborty, D. García-Soriano, and A. Matsliah. Nearly tight bounds for testing function isomorphism. In *SODA*, pages 1683–1702, 2011.

[21] Chris Charnes, Ulrich Dempwolff, and Josef Pieprzyk. The eight variable homogeneous degree three bent functions. *J. Discrete Algorithms*, 6(1):66–72, 2008.

[22] Ulrich Dempwolff. Geometric and design-theoretic aspects of semibent functions ii. *Des. Codes Cryptography*, 62(2):241–252, 2012.

[23] E. Fischer, G. Kindler, D. Ron, S. Safra, and A. Samorodnitsky. Testing juntas. *Journal of Computer and System Sciences*, 68(4):753–787, 2004.

[24] O. Goldreich, S. Goldwasser, and D. Ron. Property testing and its connection to learning and approximation. *Journal of the ACM*, 45:653–750, 1998.

[25] P. Gopalan, R. O'Donnell, R. Servedio, A. Shpilka, and K. Wimmer. Testing Fourier dimensionality and sparsity. *SIAM Journal on Computing*, 40(4):1075–1100, 2011.

[26] B. Green. A Szemerédi-type regularity lemma in abelian groups, with applications. *Geom. Funct. Anal.*, 15(2):340–376, 2005.

[27] Elena Grigorescu, Tali Kaufman, and Madhu Sudan. Succinct representation of codes with applications to testing. *SIAM J. Discrete Math.*, 26(4):1618–1634, 2012.

[28] Elena Grigorescu, Tali Kaufman, and Madhu Sudan. 2-transitivity is insufficient for local testability. *Computational Complexity*, 22(1), 2013.

[29] Alan Guo, Swastik Kopparty, and Madhu Sudan. New affine-invariant codes from lifting. In *ITCS*, pages 529–540, 2013.

[30] T. Kaufman and M. Sudan. Algebraic property testing: The role of invariance. In *STOC*, pages 403–412, 2008.

[31] Tali Kaufman and Shachar Lovett. New extension of the weil bound for character sums with applications to coding. In *FOCS*, pages 788–796, 2011.

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

[33] P. V. Kumar, R. A. Scholtz, and L. R. Welch. Generalized bent functions and their properties. *Journal of Combinatorial Theory (A)*, 40(1):90–107, 1985.

[34] E. Kushilevitz and N. Nisan. *Communication Complexity.* Cambridge University Press, 1997.

[35] R. McFarland. A family of difference sets in non-cyclic groups. *Journal of Combinatorial Theory (A)*, 15:1–10, 1973.

[36] T. Neumann. Bent functions. Master's thesis, University of Kaiserslautern, 2006.

[37] M. Parnas, D. Ron, and A. Samorodnitsky. Testing basic Boolean formulae. *SIAM Journal on Discrete Mathematics*, 16(1):20–46, 2003.

[38] O. Rothaus. On "bent" functions. *Journal of Combinatorial Theory (A)*, 20:300–305, 1976.

[39] R. Rubinfeld and M. Sudan. Robust characterizations of polynomials with applications to program testing. *SIAM Journal on Computing*, 25:252–271, 1996.

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

[41] X. Sun and C. Wang. Randomized communication complexity for linear algebra problems over finite fields. In *Proc. 29th Annual Symposium on Theoretical Aspects of Computer Science*, pages 477–488, 2012.

[42] Natalia Tokareva. On the number of bent functions: lower bounds and hypotheses. *IACR Cryptology ePrint Archive*, 2011:83, 2011.

[43] K. Wimmer and Y. Yoshida. Testing linear-invariant function isomorphism. In *ICALP*, 2013. To appear.

A Generalization to $\mathbb{F}_p$

In this section we describe how our results generalize to families of functions $f : \mathbb{F}_p^n \to \mathbb{F}_p$. The generalization is straightforward, since the properties of bent functions (and of Maiorana-McFarland functions in particular) generalize to non-boolean fields. Moreover, the reduction of [13] requires $O(\log p)$ bits of communication for each sample, but since the results in [41] hold over $\mathbb{F}_p$ with an additional multiplicative factor of $\Omega(\log p)$, combining them together yields the same query lower bounds as in the $\mathbb{F}_2$ case. The following definition extends Boolean bent functions to general bent functions.

Definition A.1 (Rothaus [38])**.** *A function $f : \mathbb{F}_p^n \to \mathbb{F}_p$ is bent if for every nonzero vector $h \in \mathbb{F}_p^n$ and every $a \in \mathbb{F}_p$, we have*

$$\Pr_{\boldsymbol{x}}[f(\boldsymbol{x}) - f(\boldsymbol{x} + h) = a] = 1/p.$$

It can be easily checked that Proposition 2.3 can be modified to give that if $f, g : \mathbb{F}_p^n \to \mathbb{F}_p$ are bent and $\dim(g) < n$ then $\Pr_x[f(x) \neq g(x)] \geq \frac{1}{2}(1 - 1/p) \geq \frac{1}{4}$. The definition of Maiorana-McFarland remains the same and the statement of Lemma 2.4 follows by essentially the same argument as before. (See also [33]).

In conclusion, our results can be extended to the following formal statements.

Theorem A.2. *Any 2-sided, adaptive $(1/4, k)$-test for $\mathcal{L}(f)$ and $\mathcal{A}(f)$ where f is a Maiorana-McFarland bent function in n variables over $\mathbb{F}_p$ requires $k = \Omega(n^2)$ queries.*

Theorem A.3. *Any 2-sided, adaptive $(1/4, k)$-test for the class of bent functions in n variables over $\mathbb{F}_p$ requires $k = \Omega(p^{n/4})$ queries.*