Electronic Colloquium on Computational Complexity, Report No. 114 (2014)

An Alternative Proof of an $\Omega(k)$ Lower Bound for Testing k-linear Boolean Functions

Roei Tell
Department of Computer Science
Weizmann Institute of Science
roei.tell@weizmann.ac.il

August 27, 2014

Abstract

We provide an alternative proof for a known result stating that $\Omega(k)$ queries are needed to test k-sparse linear Boolean functions. Similar to the approach of Blais and Kane (2012), we reduce the proof to the analysis of Hamming weights of vectors in affine subspaces of the Boolean hypercube. However, we derive our proof from a general result by Linial and Samorodnitsky (2002) that upper bounds the number of vectors with the same Hamming weight in every large affine subspace of the Boolean hypercube. Our line of argument is reminiscent of a technique that is common in communication complexity, and it allows us to derive the lower bound from Linial and Samorodnitsky's result quite easily.

We publish this proof as a self-contained excerpt from a broader work (2014), since it might be of independent interest. In the other work we also extend the result to an $\Omega(s)$ lower bound for testing s-sparse polynomials of degree d, for any $d \in \mathbb{N}$.

Keywords: Property Testing, Affine Subspaces.

1 Introduction

The class of k-linear Boolean functions consists of all linear Boolean functions over $\{0,1\}^n$ that are k-sparse, meaning that exactly k of their coefficients are non-zero. While the class of linear Boolean functions is testable with $O(1)$ queries by the BLR tester [4], testing the subclass of k-linear functions requires $\Omega(k)$ queries if $k \in [0, \frac{n}{2}]$ and $\Omega(n-k)$ queries otherwise. The problem of testing k-linear functions is computationally equivalent to the problem of testing $(n-k)$-linear functions (see [2, Apdx. B]), and hence it suffices to analyze the case of $k \leq \frac{n}{2}$.

For $k \in [0, \frac{n}{2}]$, Blais and O'Donnell [3] proved an $\Omega(\log k)$ lower bound on the query complexity of this property, and Goldreich [6] later proved an $\Omega(k)$ lower bound for non-adaptive testers and an $\Omega(\sqrt{k})$ lower bound for all testers. Blais, Brody, and Matulef [1] proved an $\Omega(k)$ lower bound for all testers, relying on a reduction from communication complexity. Blais and Kane [2] gave an alternative proof for this lower bound by directly analyzing the property testing problem (without a reduction from communication complexity).

In this paper we provide an alternative proof for this lower bound, that also does not rely on a reduction from communication complexity. Specifically, we consider a promise that the input function is linear, which guarantees a relative distance of $\frac{1}{2}$ between every pair of input functions, and prove that the query complexity of testing whether an input function is $\frac{n}{2}$-linear or is l-linear, for any $l \neq \frac{n}{2}$, is $\Omega(n)$. We then use a simple black-box reduction to extend this result to a lower bound of $\Omega(k)$ for testing k-linear functions, for any $k \in [0, \frac{n}{2}]$. Both previous proofs of this lower bound [1, 2] also considered the initial parameter of $k \approx \frac{n}{2}$ and a promise that the input function is linear.

ISSN 1433-8092

High-level overview. We rely on a standard reduction to deterministic testers with an (arbitrary) distribution on the inputs. Identifying linear functions with their coefficient vectors in $\{0,1\}^n$, we follow Blais and Kane [2] and show that any deterministic tester with query complexity q partitions $\{0,1\}^n$ into affine subspaces of co-dimension at most q such that the tester's output on all vectors in each subspace is identical.

We consider testers with query complexity $\frac{1}{3} \cdot n$, and corresponding partitions of $\{0,1\}^n$ to subspaces of co-dimension $\frac{1}{3} \cdot n$. Note that the overall fraction of "yes" instances (i.e., of vectors with Hamming weight $\frac{n}{2}$) in $\{0,1\}^n$ is $O(\frac{1}{\sqrt{n}})$. We rely on a result by Linial and Samorodnitsky [8, Thm 4.4], that upper-bounds the fraction of vectors with the same Hamming weight $k \in [n]$ in every affine subspace of dimension $\lambda \cdot n$ (for $\lambda > \frac{1}{2}$) by $O_\lambda(\frac{1}{\sqrt{n}})$.

We choose a distribution that with probability p is uniform over all vectors with Hamming weight $\frac{n}{2}$, and is otherwise uniform over all other vectors. Relying on Linial and Samorodnitsky's result, and choosing p appropriately, we show that this distribution assigns at most half of the probabilistic mass of each subspace of co-dimension $\frac{1}{3} \cdot n$ to vectors with Hamming weight $\frac{n}{2}$. Thus, if a deterministic tester with query complexity $\frac{1}{3} \cdot n$ accepts a subspace in the corresponding partition of $\{0,1\}^n$, then the tester errs on half of the probabilistic mass of that subspace. On the other hand, if the tester only accepts an overall sub-constant probabilistic mass of subspaces, it incorrectly rejects most of the probabilistic mass assigned to vectors with Hamming weight $\frac{n}{2}$.

Note: This proof is part of a work that was published separately [9], in which we also extend this lower bound to an $\Omega(\min\{s, \binom{n}{d} - s\})$ lower bound for testing s-sparse polynomials of degree d, for any $d \in \mathbb{N}$. We chose to present the current proof independently and in self-contained form, since it might be of independent interest.

2 Preliminaries

For $n \in \mathbb{N}$ and $w \in \{0,1\}^n$, denote the Hamming weight of w by $\|w\|_1 = \sum_{i=1}^n w_i$. For $0 \leq k \leq n$, let $W_k = \{w \in \{0,1\}^n : \|w\|_1 = k\}$. We state a special case of Linial and Samorodnitsky's result [8, Thm 4.4]:

Theorem 1. *For $n \in \mathbb{N}$ and every affine subspace $V \subseteq \{0,1\}^n$ with dimension at least $\frac{2}{3} \cdot n$, it holds that $\frac{|V \cap W_{n/2}|}{|V|} = O(\frac{1}{\sqrt{n}})$.*

For $k \in \mathbb{N}$, we say that a function $f : \{0,1\}^n \to \{0,1\}$ is k-linear if f is a linear function with exactly $k(n)$ non-zero coefficients.

Definition 2. *For $n, k \in \mathbb{N}$, a tester for the problem of* testing k-linear Boolean functions over $\{0,1\}^n$ under the promise that the input function is linear *is a randomized oracle machine T that satisfies the following two conditions:*

1. *If $f : \{0,1\}^n \to \{0,1\}$ is k-linear, then $Pr[T^f(1^n) = 1] \geq \frac{2}{3}$.*
2. *If $f : \{0,1\}^n \to \{0,1\}$ is l-linear, for $l \neq k$, then $Pr[T^f(1^n) = 0] \geq \frac{2}{3}$.*

The query complexity of T is the maximum (over $x \in \{0,1\}^n$ and internal coin tosses of T) number of queries that T makes. The query complexity of the property of k-linear Boolean functions over $\{0,1\}^n$ under the promise that the input function is linear *is the minimum query complexity of all testers for k-linear Boolean functions.*

Note that by standard error reduction, if there exists a tester with query complexity q that satifies both conditions in Definition 2, then for every $\delta > 0$ there exists a tester with query complexity $q \cdot O(\log(1/\delta))$ that satisfies both conditions with probability $1 - \delta$ (instead of $\frac{2}{3}$).

3 The proof

We start by proving the case of $k = \frac{n}{2}$, then we extend the result to all $k \in [0, \frac{n}{2}]$ by a simple black-box reduction to the $k = \frac{n}{2}$ case. Recall that testing k-linear functions is computationally equivalent to testing $(n-k)$-linear functions (see [2, Apdx. B] for a proof), and hence it suffices to focus on the case of $k \in [0, \frac{n}{2}]$.

Theorem 3. *For $n \in \mathbb{N}$, the query complexity of the property of $\frac{n}{2}$-linear Boolean functions over $\{0,1\}^n$ under the promise that the input function is linear is $\Omega(n)$.*

Proof. It is well-known that in order to lower bound the error probability of any probabilistic tester with query complexity $q(n)$, it suffices to lower bound the error probability of all deterministic testers with query complexity $q(n)$ over an (arbitrary) distribution on the inputs. We therefore show a distribution on the inputs such that every deterministic tester with query complexity $\frac{1}{3} \cdot n$ errs on it with probability that is lower bounded by a universal positive constant. According to the discussion after Definition 2 this implies that there does not exists a tester with query complexity $o(n)$ for the property.

Since we consider distributions that are supported only by linear functions, we can represent any distribution over the inputs as a distribution over $\{0,1\}^n$, where each vector $w \in \{0,1\}^n$ represents a $\|w\|_1$-linear function (by representing its coefficients). Correspondingly, $\{0,1\}^n$ is partitioned to $W_{n/2}$, which consists of all "yes" instances, and $\{0,1\}^n \setminus W_{n/2}$, which consists of all "no" instances. A key lemma in our analysis will be the following:

Lemma 4 (following [2, 5]): *Any deterministic tester with query complexity $\frac{1}{3} \cdot n$ for $\frac{n}{2}$-linearity induces a partition of $\{0,1\}^n$ into affine subspaces of dimension at least $\frac{2}{3} \cdot n$ such that its output on all vectors in each subspace is identical.*

The intuition behind this lemma is that a deterministic tester that is given a linear function as input is equivalent to a *parity decision tree* that inputs the coefficients of the linear function. The affine subspaces in the induced partition correspond to the leaves of the tree. For further details see [9, Prop 5.6].

Proof. Let T be a deterministic tester, and we assume without loss of generality that it issues exactly $m \leq \frac{1}{3} \cdot n$ queries on any input. For an arbitrary input $w \in \{0,1\}^n$, denote the corresponding $\|w\|_1$-linear function by f_w, and denote the m queries issued during the execution of $T^{f_w}(1^n)$ by the rows of an m-by-n matrix Q, and the responses received by $r \in \{0,1\}^m$. Then

$$\mathcal{V}_{Q,r} = \{w' \in \{0,1\}^n : Qw' = r\}$$

is an affine subspace of dimension at least $\frac{2}{3} \cdot n$.

Clearly $w \in \mathcal{V}_{Q,r}$. Let $w' \in \mathcal{V}_{Q,r}$. Since T is deterministic, the first query issued by $T^{f_{w'}}(1^n)$ is identical to the first query issued by $T^{f_w}(1^n)$, and since $w' \in \mathcal{V}_{Q,r}$, the first response is also identical in both cases. By induction, all m queries and responses will be identical in both cases, and in particular the final output will also be identical. Note that this is true both for adaptive and for non-adaptive testers.

To see that these subspaces are indeed a partition of $\{0,1\}^n$, consider two subspaces $\mathcal{V}^{(1)}_{Q^{(1)},r^{(1)}}$ and $\mathcal{V}^{(2)}_{Q^{(2)},r^{(2)}}$ such that for $i = 1, 2$, for every input $w \in \mathcal{V}^{(i)}_{Q^{(i)},r^{(i)}}$ it holds that $T^{f_w}(1^n)$ makes queries $Q^{(i)}$ and receives responses $r^{(i)}$. If there exists $w \in \mathcal{V}^{(1)}_{Q^{(1)},r^{(1)}} \cap \mathcal{V}^{(2)}_{Q^{(2)},r^{(2)}}$ then it follows that $Q^{(1)} = Q^{(2)}$ and $r^{(1)} = r^{(2)}$, which implies that $\mathcal{V}^{(1)}_{Q^{(1)},r^{(1)}} = \mathcal{V}^{(2)}_{Q^{(2)},r^{(2)}}$. Also, any $w \in \{0,1\}^n$ belongs to some affine subspace of this form (induced by the queries and responses during the execution of $T^{f_w}(1^n)$). □

For $p \in (0,1)$, let $\mathcal{D}_p$ be a distribution over $\{0,1\}^n$ that with probability p is uniform over $W_{n/2}$ and with probability $1-p$ is uniform over $\{0,1\}^n \setminus W_{n/2}$. Following Lemma 4, it suffices to show that for any partition of $\{0,1\}^n$ to affine subspaces of dimension at least $\frac{2}{3} \cdot n$ and any labeling of these subspaces by "yes" and "no", a constant probabilistic mass of vectors under $\mathcal{D}_p$ is incorrectly labeled (where a vector $u \in W_{n/2}$ is labeled correctly if it is labeled by "yes" and a vector $v \in \{0,1\}^n \setminus W_{n/2}$ is labeled correctly if it is labeled by "no").

Lemma 5. *There exist constants $p \in (0,1)$ and $\mu > 0$ such that for any sufficiently large even integer n and any partition of $\{0,1\}^n$ to affine subspaces of dimension at least $\frac{2}{3} \cdot n$ and any 2-labeling of the subspaces in the partition, the probabilistic mass under $\mathcal{D}_p$ of vectors that are labeled incorrectly is at least μ.*

Proof. We first prove that there exists $p \in (0,1)$ such that $\mathcal{D}_p$ assigns at most half of the probabilistic mass of every subspace $V \subset \{0,1\}^n$ of dimension at least $\frac{2}{3} \cdot n$ to $V \cap W_{n/2}$. To see this, note that for an arbitrary $u \in W_{n/2}$ and $v \in \{0,1\}^n \setminus W_{n/2}$ it holds that $\mathcal{D}_p(u) = \frac{p}{|W_{n/2}|}$ and $\mathcal{D}_p(v) = \frac{1-p}{|\{0,1\}^n \setminus W_{n/2}|}$. Therefore

$$\frac{\mathcal{D}_p(u)}{\mathcal{D}_p(v)} = \frac{p}{1-p} \cdot \frac{|\{0,1\}^n \setminus W_{n/2}|}{|W_{n/2}|} = \frac{p}{1-p} \cdot O(\sqrt{n})$$

From this it follows that

$$\begin{aligned}\frac{\mathcal{D}_p(V \cap W_{n/2})}{\mathcal{D}_p(V)} &= \frac{\mathcal{D}_p(u) \cdot |V \cap W_{n/2}|}{\mathcal{D}_p(v) \cdot |V \setminus W_{n/2}| + \mathcal{D}_p(u) \cdot |V \cap W_{n/2}|} \\ &\leq \frac{\mathcal{D}_p(u)}{\mathcal{D}_p(v)} \cdot \frac{|V \cap W_{n/2}|}{|V|} \\ &= \frac{p}{1-p} \cdot O(\sqrt{n}) \cdot \frac{|V \cap W_{n/2}|}{|V|} \qquad (1)\end{aligned}$$

According to Theorem 1 it holds that (1) is upper bounded by $\frac{p}{1-p} \cdot c$ for some $c > 0$. By setting $p = \frac{1}{1+2\cdot c} \in (0,1)$ we get that $\frac{\mathcal{D}_p(V \cap W_{n/2})}{\mathcal{D}_p(V)} \leq \frac{1}{2}$.

It follows that any labeling of a subspace of dimension at least $\frac{2}{3} \cdot n$ by "yes" incorrectly labels at least half the probabilistic mass assigned by $\mathcal{D}_p$ to that subspace. Now, fix an arbitrary partition of $\{0,1\}^n$ to subspaces of dimension at least $\frac{2}{3} \cdot n$. If the probabilistic mass of subspaces that are labeled by "yes" is at most $\frac{p}{2}$, then there is a probabilistic mass of $\frac{p}{2}$ of vectors in $W_{n/2}$ that are incorrectly labeled by "no". On the other hand, if the probabilistic mass of subspaces that are labeled by "yes" is larger than $\frac{p}{2}$, then vectors in $\{0,1\}^n \setminus W_{n/2}$ with probabilistic mass of at least $\frac{1}{2} \cdot \frac{p}{2} = \frac{p}{4}$ are incorrectly labeled by "yes". The lemma follows with $\mu = \frac{p}{4}$. □

According to Lemmas 4 and 5, the error probability under $\mathcal{D}_p$ of every deterministic tester with query complexity $\frac{1}{3} \cdot n$ is lower bounded by the universal constant $\mu > 0$. The theorem follows. □

For $k < \frac{n}{2}$ we use a simple black-box reduction to the case of Theorem 3 to show that the query complexity of testing k-linear functions is $\Omega(k)$. This black-box reduction is implicit in a padding argument presented by [2] for similar purposes.

Theorem 6. *For $n \in \mathbb{N}$ and $k \in [0, \frac{n}{2}]$, the query complexity of the property of k-linear Boolean functions over $\{0,1\}^n$ under the promise that the input function is linear is $\Omega(k)$.*

Proof. Let $m = 2 \cdot k < n$. Assuming that there exists a tester T' for the property of k-linear functions over $\{0,1\}^n$, we construct a corresponding tester T for $\frac{m}{2}$-linear functions over $\{0,1\}^m$, with the same error probability and query complexity (both problems are under the promise that the input function is linear). Since the query complexity of testing $\frac{m}{2}$-linear functions over $\{0,1\}^m$ is $\Omega(m)$, it follows that the query complexity of testing k-linear functions over $\{0,1\}^n$ is $\Omega(m) = \Omega(k)$.

The construction itself is as follows. The tester T is given oracle access to a function $f : \{0,1\}^m \to \{0,1\}$, and simulates the execution of T' when T' is given access to a function $g : \{0,1\}^n \to \{0,1\}$ that is defined in the following way: For every $z = x \circ \tau \in \{0,1\}^n$, where $x \in \{0,1\}^m$ and $\tau \in \{0,1\}^{n-m}$, let $g(x \circ \tau) = f(x) = \sum_{i=1}^{m} f_i x_i$. Note that T can answer any oracle query that T' makes to g by making a single oracle query to f. Furthermore, f is an $\frac{m}{2}$-linear function if and only if g is a k-linear function. The theorem follows. □

4 Digest

Similar to the approach of Blais and Kane [2], we reduce the proof to a geometric problem regarding the intersection of $W_{n/2}$ with large affine subspaces of $\{0,1\}^n$. However, our line of argument is slightly different: Blais and Kane consider a distribution that with probability $\frac{1}{2}$ is uniform over $W_{\frac{n}{2}-1}$ and is otherwise uniform over $W_{\frac{n}{2}+1}$, whereas we consider a distribution that with probability $p > 0$ is uniform over $W_{n/2}$, and is otherwise uniform over $\{0,1\}^n \setminus W_{n/2}$. This allows us to rely on the general result proved by Linial and Samorodnitsky.

Since we consider a distribution that is supported by all linear functions, our proof does not easily extend to a lower bound on other properties, in the same way that both previous proofs do (see [1, Prop 3.4] for further details). For example, the distribution in the proof of Theorem 3 clearly does not yield a lower bound on testing functions with Fourier degree at most $\frac{n}{2}$ (since it is supported by "yes" instances that have Fourier degree $\frac{n}{2}$ and by "no" instances that have all Fourier degrees in $[n] \setminus \{\frac{n}{2}\}$).

Our line of argument is reminiscent of a technique that is common in communication complexity (see, e.g., [7, Method 1]). Indeed, this proof emerged from an examination of the influential methodology of Blais, Brody, and Matulef [1] for reducing problems in communication complexity to property testing problems (see [9] for further details).

Acknowledgements

The author thanks Tom Gur for suggesting and motivating the search for alternative proofs for the lower bound on testing k-linearity. The author is grateful to Avishay Tal for pointing him to the work of Linial and Samorodnitsky. The author also thanks his advisor, Oded Goldreich, for his guidance and support in the research and writing process. This research was partially supported by the Israel Science Foundation (grant No. 671/13).

References

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

[2] Eric Blais and Daniel M. Kane. Tight bounds for testing k-linearity. In *APPROX-RANDOM*, pages 435–446, 2012.

[3] Eric Blais and Ryan O'Donnell. Lower bounds for testing function isomorphism. In *IEEE Conference on Computational Complexity*, pages 235–246, 2010.

[4] M. Blum, M. Luby, and R. Rubinfeld. Self-testing/correcting with applications to numerical problems. In *Proceedings of the Twenty-second Annual ACM Symposium on Theory of Computing*, STOC '90, pages 73–83, New York, NY, USA, 1990. ACM.

[5] Harry Buhrman, David García-Soriano, Arie Matsliah, and Ronald de Wolf. The non-adaptive query complexity of testing k-parities. *Chicago J. Theor. Comput. Sci.*, 2013, 2013.

[6] Oded Goldreich. On testing computability by small width obdds. In *Proceedings of the 13th International Conference on Approximation, and 14 the International Conference on Randomization, and Combinatorial Optimization: Algorithms and Techniques*, APPROX/RANDOM'10, pages 574–587, Berlin, Heidelberg, 2010. Springer-Verlag.

[7] Hartmut Klauck. Rectangle size bounds and threshold covers in communication complexity. In *IEEE Conference on Computational Complexity*, pages 118–134, 2003.

[8] Nathan Linial and Alex Samorodnitsky. Linear codes and character sums. *Combinatorica*, 22(4):497–522, 2002.

[9] Roei Tell. Deconstructions of reductions from communication complexity to property testing using generalized parity decision trees. *Electronic Colloquium on Computational Complexity (ECCC)*, 21, 2014.

ECCC ISSN 1433-8092

http://eccc.hpi-web.de