Electronic Colloquium on Computational Complexity, Report No. 29 (2013)

A $o(n)$ Monotonicity Tester for Boolean Functions over the Hypercube

D. Chakrabarty and C. Seshadhri

Abstract. Given oracle access to a Boolean function $f : \{0,1\}^n \mapsto \{0,1\}$, we design a randomized tester that takes as input a parameter $\varepsilon > 0$, and outputs Yes if the function is monotone, and outputs No with probability $> 2/3$, if the function is ε-far from monotone. That is, f needs to be modified at ε-fraction of the points to make it monotone. Our non-adaptive, one-sided tester makes $\widetilde{O}(n^{5/6}\varepsilon^{-5/3})$ queries to the oracle.

1. Introduction

Testing monotonicity of Boolean functions is a classical question in property testing. The Boolean hypercube $\{0,1\}^n$ defines a natural partial with $x \prec y$ iff $x_i \leq y_i$ for all $i \in [n]$. A Boolean function $f : \{0,1\}^n \mapsto \{0,1\}$ is *monotone* if $f(x) \leq f(y)$ whenever $x \prec y$.

A Boolean function's *distance* to monotonicity is the minimum fraction of points at which it needs to be modified to make it monotone. In the property testing framework, we are provided oracle access to the function f and given a parameter $\varepsilon > 0$. A *monotonicity tester* is an algorithm that accepts, if the function is monotone and rejects, if the function is ε-far from monotone. The tester is allowed to be randomized, and has to be correct with non-trivial probability (say $> 2/3$). The tester is called *one-sided* if the tester always accepts a monotone function. The tester is *non-adaptive* if the queries made by the algorithm do not depend on the answers given by the oracle.

The quality of a monotonicity tester is governed by the number of oracle queries as well as the running time. Goldreich et al. [8] suggested the following simple tester: query the function value on a pair of points that differ on exactly a single coordinate and reject if monotonicity is violated. In other words, the tester samples a random edge of the hypercube and checks for monotonicity between the two endpoints. This is called the *edge tester* for monotonicity; it is clear the running time is of the same order as the query time.

Goldreich et al. [8] show that $O(n/\varepsilon)$-queries by the edge tester suffice to test monotonicity. They also show that their analysis is tight, so the edge tester can do no better. They explicitly ask whether there exists a tester with an improved query complexity in terms of n? Fischer et al. [7] show that any non-adaptive, one-sided tester[1] for monotonicity must make $\Omega(\sqrt{n})$-queries for constant $\varepsilon > 0$. While monotonicity has been extensively studied in property testing [6, 8, 5, 10, 7, 9, 11, 1, 3, 2, 4], no significant progress had been made on this decade old question of testing monotonicity of Boolean functions.

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

[1][7] also show a $\Omega(\log n)$ lower bound for 2-sided testers.

ISSN 1433-8092

Our main result is an affirmative answer to the above question of [8]. For brevity, we use the tilde notation to hide log factors; $\widetilde{O}(h)$ contains functions dominated by $h \cdot \mathrm{poly}(\log h)$.

Theorem 1. *There exists a one-sided, non-adaptive $\widetilde{O}(n^{5/6}\varepsilon^{-5/3})$-query monotonicity tester for Boolean functions $f : \{0,1\}^n \mapsto \{0,1\}$.*

We get an improved bound for functions with *low average sensitivity.* Given a Boolean function f, the influence of dimension i, denoted as Inf_i, is defined as the fraction of edges on the hypercube crossing the ith dimension whose endpoints have different function values. The average sensitivity, denoted as $\mathbf{I}(f)$, is the sum of all the n influences. The functions defined in [7] to prove lower bounds of $\Omega(\sqrt{n})$ for non-adaptive, one-sided testers have constant average sensitivity, and hence the following is optimal for that setting.

Theorem 2. *There exists a one-sided, non-adaptive $O(\sqrt{n}\mathrm{poly}(\mathbf{I}(f)/\varepsilon))$-query monotonicity tester for Boolean functions of average sensitivity $\mathbf{I}(f)$.*

1.1. **Pair testers.** A *pair tester* [5] describes a fixed distribution (independent of the function) on domain pairs $(x \prec y)$, makes independent queries on pairs drawn from this distribution, and rejects iff some drawn pair violates monotonicity. By definition, pair testers are non-adaptive and one-sided. (Note that the edge tester is a pair tester.) Briët et al. [3] show that any pair tester requires $\Omega(n/(\varepsilon \log n))$ samples and the linear dependence on $O(1/\varepsilon)$ is crucial in their argument. Our tester is also a pair tester. We circumvent the lower bound (on n) of [3] because of the worse dependence on ε.

1.2. **Main Ideas.** Our tester is a combination of the edge tester and what we call the *path* tester. The path tester essentially samples a random point x on the hypercube, performs a sufficiently long random length walk 'up' the directed hypercube to reach y, queries $f(x)$ and $f(y)$ and tests for monotonicity.

Our algorithm is inspired by a recent paper by Ron et al. [12], which shows a $O(\sqrt{n})$-query randomized algorithm to estimate the average sensitivity of a *monotone* function. The algorithm essentially performs the operation above and counts the number of mismatches; Ron et al. [12] explicitly ask whether an algorithm "in the spirit" above can be used for monotonicity. Our answer is yes.

Consider a function f is ε-far from monotone. The aim of any tester is to detect a *violation*, that is, a pair $x \prec y$ such that $1 = f(x) > f(y) = 0$. The success probability of the edge tester is exactly the fraction of violated edges. The intuition is that there are possibly many more violations that are "far away" and the directed random walk will help detect those. Consider the function $f : \{0,1\}^{n+1} \mapsto \{0,1\}$, $f(0,x) = 0$ if $|x| \leq n/2 - 2\sqrt{n}$ and 1 otherwise; $f(1,x) = 0$ if $|x| \leq n/2 + 2\sqrt{n}$ and 1 otherwise. This function has a constant distance to monotonicity, and all the violated edges are of the form $((0,x),(1,x))$ for $n/2 - 2\sqrt{n} \leq |x| \leq n/2 + 2\sqrt{n}$. The edge tester detects a violation with probability only $\Theta(1/n)$. Suppose we pick a uniform random point and perform of length $\sqrt{n}/2 \leq \ell < \sqrt{n}$. If the starting point has 0 in the first coordinate and any of the $\Theta(\sqrt{n})$ steps flip the first coordinate, the walk detects a violation. This happens with probability $\Theta(1/\sqrt{n})$, handily beating the edge tester.

The argument above required the violated edges to be aligned along one dimension. We prove (in §2.2) that directed random walk detects a violation (with sufficiently high probability) when there is a large *matching* of violated edges. One of the ingredients of this proof is the following interesting combinatorial observation. In §2.1, we prove that if an ε-fraction of the hypercube is marked blue, then the probability that the random walk starts and ends at a blue point is $\widetilde{\Omega}(\varepsilon^2)$. It shows that the endpoints of this random walk, which are highly correlated, behave like two independent samples as far as being blue is concerned.

But what if no large matching of violated edges exists? Take the 'anti-majority' function, which is far from monotone, yet only has a matching of violated edges of size $\Theta(2^n/\sqrt{n})$. This is dealt with by our *dichotomy theorem*. In §2.3, we prove that for any $s > 0$, either there exists $\Theta(s\varepsilon 2^n)$ violated edges, or there exists a matching of $\Theta(\varepsilon 2^n/s)$ violated edges. With this we are done; in the former case, the edge tester suffices, in the latter the path tester suffices. The factor $n^{5/6}$ is the optimal tradeoff obtained by our approach.

The proof of our dichotomy theorem combines two ideas discovered earlier in the context of monotonicity testing. The first is a theorem of Lehman and Ron [10] on multiple source-sink routing over the hypercube. The second is an alternating paths machinery developed by the authors in a separate work [4] for general range monotonicity.

2. The Tester and its Analysis

Fix a parameter $\varepsilon > 0$. Our tester accepts if the function is monotone, and rejects with probability $> 2/3$ if f is ε-far from being monotone. We assume without loss of generality that $\varepsilon \leq 1/2$ since any function can be made monotone by changing at most $1/2$ of its values. Set parameter $\ell := 2\lceil C_\varepsilon\sqrt{n}\,\rceil$, where $C_\varepsilon = \sqrt{10\ln(1/\varepsilon)}$. Since $\varepsilon \leq 1/2$, we have $C_\varepsilon > 2$ and $\ell > 4\sqrt{n}$. I_ℓ denotes the index set $[n/2 - \ell/2, n/2 + \ell/2]$.

For a binary vector $x, y \in \{0,1\}^n$, $|x|$ is the number of 1's in x and $\|x - y\|_1$ is the ℓ_1-distance between x and y. The all zeros and all ones vectors are denoted 0^n and 1^n, respectively. The directed hypercube is the directed graph with vertex set $\{0,1\}^n$, and an arc from x to y if $x \prec y$ and $\|y - x\|_1 = 1$. Throughout the paper, u.a.r. stands for 'uniformly at random'.

We describe a random walk based procedure called the *path tester*.

Path Tester.

(1) Let $\mathcal{P}$ be the collection of paths in the directed hypercube from 0^n to 1^n. Pick a path $\mathbf{p} \in \mathcal{P}$ u.a.r. Let $X_\mathbf{p} := \{z \in \mathbf{p} : |z| \in I_\ell\}$.

(2) Sample $x \in X_\mathbf{p}$ u.a.r.

(3) Let $Y_\mathbf{p}(x) := \{z \in X_\mathbf{p} : ||z - x||_1 \geq \frac{\varepsilon\ell}{32C_\varepsilon} - 1\}$. Sample $y \in Y_\mathbf{p}(x)$ u.a.r.

(4) Reject if (x, y) violate monotonicity; i.e. $f(x) < f(y), x \succ y$ or $f(x) > f(y), x \prec y$.

This is clearly a pair tester (and is hence non-adaptive and one-sided). Our final tester runs either the path tester or the edge tester, each with probability $1/2$.

The heart of our work lies in lower bounding the probability of rejection when the function f is ε-far from monotonicity. Henceforth, we assume the function f is ε-far, and we call the rejection event a success. Recall, since f is ε-far from monotonicity, any maximal set M of disjoint, violating pairs satisfies $|M| \geq \varepsilon 2^{n-1}$ [7]. We refer to M as a *matching* of violated pairs,

For $1 \leq i \leq n$, $L_i := \{x \in \{0,1\}^n : |x| = i\}$ denotes the ith layer of the directed hypercube. We refer to $\bigcup_{i \in I_\ell} L_i$ as the *middle layers* of the hypercube.

Proposition 2.1. (a) $|\bigcup_{i \notin I_\ell} L_i| \leq \varepsilon^5 2^n$. (b) *A u.a.r path* $\mathbf{p}$ *contains a u.a.r vertex from* L_i.

Proof. By Chernoff bounds, for a u.a.r $x \in \{0,1\}^n$, $\Pr[\big||x| - n/2\big| > \ell/2] \leq 2e^{-\ell^2/2n}$. Since $\ell = 2\lceil \sqrt{10n \log(1/\varepsilon)}\ \rceil$, this probability is at most ε^5. For the second part, observe the number of paths in $\mathcal{P}$ that pass through a given vertex x depends solely on $|x|$. ☐

2.1. **Going from blue to blue.** Assume ε-fraction of the hypercube is colored blue. Let (x, y) be a random pair sampled by the path tester, and let $\mathcal{E}$ be the event that both x and y are blue. If x and y were chosen *independently* u.a.r., then the probability of both being blue is ε^2. The following lemma shows that this probability does not degrade much even though x and y are correlated (for instance, they form an ancestor-descendant pair).

Lemma 2.1. $\Pr[\mathcal{E}] = \Omega\left(\dfrac{\varepsilon^2}{\ln(1/\varepsilon)}\right)$.

Proof. For notational convenience, set $\mu := \varepsilon/16C_\varepsilon$. This implies[2] $|X_\mathbf{p}| - |Y_\mathbf{p}(x)| \leq \mu\ell$ for any $x \in \mathbf{p}$. Let $b(\mathbf{p})$ be the random variable denoting the number of blue points in $X_\mathbf{p}$ corresponding to a random path $\mathbf{p}$. Let $\mathcal{E}_x$ and $\mathcal{E}_y$ be the probabilities that the first and second points are blue; that is $\mathcal{E} = \mathcal{E}_x \wedge \mathcal{E}_y$. Abusing notation, $\mathbf{p}$ will also denote the event that $\mathbf{p}$ is the sampled path.

Conditioned on a path $\mathbf{p}$ being sampled, the probability of the first point x sampled by the path tester being blue, that is $\Pr[\mathcal{E}_x \mid \mathbf{p}] = b(\mathbf{p})/\ell$.

Conditioned on the path being $\mathbf{p}$ and the first point being x (irrespective of it being blue or not), the probability that the second point y is blue is the number of blue points in $Y_\mathbf{p}(x)$ divided by $|Y_\mathbf{p}(x)|$. The number of blue points is at least $b(\mathbf{p}) - \mu\ell$ since $|X_\mathbf{p}| - |Y_\mathbf{p}(x)| \leq \mu\ell$. Therefore,

$$\Pr[\mathcal{E}_y \mid \mathbf{p}, x \text{ first point}] = \frac{|\text{blue points in } Y_\mathbf{p}(x)|}{|Y_\mathbf{p}(x)|} \geq \frac{b(\mathbf{p}) - \mu\ell}{\ell}.$$

Since the above inequality holds for all x (in particular all the blue x's),

$$\Pr[\mathcal{E}_y \mid \mathbf{p}, \mathcal{E}_x] \geq \frac{b(\mathbf{p}) - \mu\ell}{\ell}.$$

[2]The curious reader may be wonder why we have a "-1" in the distance condition for $Y_\mathbf{p}(x)$ in the description of the path tester. This is a technicality so that we have the bound $|X_\mathbf{p}| - |Y_\mathbf{p}(x)| \leq \mu\ell$. Without the -1, the bound would be $\mu\ell + 2$ and can be made $O(\mu\ell)$ only for large enough ε. So instead of enforcing such a condition or carrying around a $+2$, the "-1" allows for a cleaner presentation.

Together, we get

$$\begin{aligned}\Pr[\mathcal{E}] &= \sum_{\mathbf{p}\in\mathcal{P}} \Pr[\mathcal{E}_y|\mathbf{p},\mathcal{E}_x]\cdot\Pr[\mathcal{E}_x|\mathbf{p}]\cdot\Pr[\mathbf{p}] \\ &\geq \frac{1}{|\mathcal{P}|}\sum_{\mathbf{p}\in\mathcal{P}}\left(\frac{b(\mathbf{p})}{\ell}\cdot\frac{b(\mathbf{p})-\mu\ell}{\ell}\right) \\ &= \frac{1}{|\mathcal{P}|}\sum_{\mathbf{p}\in\mathcal{P}}\left(\frac{b(\mathbf{p})}{\ell}\right)^2 - \frac{\mu}{|\mathcal{P}|}\sum_{\mathbf{p}\in\mathcal{P}}\frac{b(\mathbf{p})}{\ell}. \qquad (1)\end{aligned}$$

In the following, we use $\mathbf{E}[...]$ to denote the expectation over the choice of the path $\mathbf{p}$.

Claim 2.1. $\mathbf{E}[b(\mathbf{p})/\ell] := \frac{1}{|\mathcal{P}|}\sum_{\mathbf{p}\in\mathcal{P}}(b(\mathbf{p})/\ell) \geq \varepsilon/8C_\varepsilon$.

Proof. Note that, for all i, $|L_i| \leq \binom{n}{n/2} \leq \frac{2^n}{\sqrt{n}}$. Let n_i be the number of blue vertices in layer L_i and Z_i be the indicator variable for the ith layer vertex in $\mathbf{p}$ being blue. Hence, $b(\mathbf{p}) = \sum_{i\in I_\ell} Z_i$. For all i, a $\mathbf{p}$ chosen u.a.r from $\mathcal{P}$ contains a uniform random vertex in layer L_i. (Prop. 2.1). Thus,

$$\mathbf{E}[Z_i] = \frac{n_i}{|L_i|} \geq \frac{\sqrt{n}}{2^n}\cdot n_i.$$

By the tail bound of Prop. 2.1, the number of vertices not in the middle layers is at most $\varepsilon^5 2^n$. Since $\varepsilon \leq 1/2$, $\sum_{i\in I_\ell} n_i \geq (\varepsilon - 2\varepsilon^5)2^n \geq \varepsilon 2^{n-1}$. Using linearity of expectation and the boun $\ell < 4C_\varepsilon\sqrt{n}$,

$$\mathbf{E}[b(\mathbf{p})/\ell] \geq \frac{\sqrt{n}}{\ell 2^n}\sum_{i\in I_\ell} n_i \geq \frac{\varepsilon\sqrt{n}}{2\ell} \geq \frac{\varepsilon}{8C_\varepsilon}.$$

□

We express the bound of (1) in terms of expectations and apply Jensen's inequality.

$$\begin{aligned}\Pr[\mathcal{E}] &\geq \mathbf{E}[(b(\mathbf{p})/\ell)^2] - \mu\mathbf{E}[b(\mathbf{p})/\ell] \\ &\geq (\mathbf{E}[b(\mathbf{p})/\ell])^2 - \mu\mathbf{E}[b(\mathbf{p})/\ell] = \mathbf{E}[b(\mathbf{p})/\ell](\mathbf{E}[b(\mathbf{p})/\ell] - \mu)\end{aligned}$$

The function $h(x) = x(x-\mu)$ is increasing when $x \geq \mu/2$. Hence, we can apply the lower bound of Claim 2.1 to get $\Pr[\mathcal{E}] \geq (\varepsilon/8C_\varepsilon)(\varepsilon/8C_\varepsilon - \mu)$. Plugging back $\mu = \varepsilon/16C_\varepsilon$, we complete the proof of Lemma 2.1. □

2.2. **Large violated-edge matchings are good.** We bound the success of the path tester when a large matching of violated edges exists.

Lemma 2.2. *Suppose there exists a matching E of violated edges in the middle layers of the hypercube. Then the path tester succeeds with probability*

$$\widetilde{\Omega}\left(\frac{\varepsilon}{\sqrt{n}}\cdot\left(\frac{|E|}{2^n}\right)^2\right).$$

Proof. We begin with some notation. Denote $B_0 = \{x \in B : f(x) = 0\}$ and $B_1 = \{x \in B : f(x) = 1\}$. Note that $|B_0| = |B_1| = |E|$. For any two points x, y, let $\mathcal{E}_{x,y}$ denote the event that the path tester picks (x, y). For convenience, in what follows, the pairs of E will be ordered according to the directed hypercube (so if $(z, z') \in E$, then $z \prec z'$). We abuse notation to define E as a function. For edge $(z, z') \in E$, we set $E(z) = z'$ and $E(z') = z$.

We define the following sets of pairs of vertices.

$$\begin{aligned} V &= \{(x,y) | x \prec y, \|x - y\|_1 \geq \frac{\varepsilon\ell}{32C_\varepsilon} - 1, x \in B_1, y \in B_1\} \\ V' &= \{(x, E(y)) | (x,y) \in V\} \end{aligned}$$

A few observations. V lies in the support of the pair tester, that is, pairs (x, y) sampled with non-zero probability. Every pair in V' is a violation; for $(x, y) \in V$, we have $x \prec y$, $y \in B_1$ implying $E(y) \succ y$ and $E(y) \in B_0$. Finally, the mapping $(x, y) \in V$ to $(x, E(y)) \in V'$ is one-to-one. This uses the fact that E is a matching (and is a crucial piece of the proof).

Since all pairs in V' are violations,

$$\Pr[\text{ success }] \geq \sum_{(x,y') \in V'} \Pr[\mathcal{E}_{x,y'}].$$

Using the mapping between V' and V and that $\Pr[\mathcal{E}_{x,y}] > 0$ for $(x, y) \in V$,

$$\begin{aligned} \sum_{(x,y') \in V'} \Pr[\mathcal{E}_{x,y'}] &= \sum_{(x,y') \in V'} \Pr[\mathcal{E}_{x,E(y')}] \cdot \frac{\Pr[\mathcal{E}_{x,y'}]}{\Pr[\mathcal{E}_{x,E(y')}]} \\ &= \sum_{(x,y) \in V} \Pr[\mathcal{E}_{x,y}] \cdot \frac{\Pr[\mathcal{E}_{x,E(y)}]}{\Pr[\mathcal{E}_{x,y}]} \end{aligned} \tag{2}$$

We break the remaining proof into simpler claims. For a vertex x, define $s(x) := |Y_{\mathbf{p}}(x)| = |\{z \in X_{\mathbf{p}} : \|z - x\|_1 \geq \varepsilon\ell/32C_\varepsilon - 1\}|$ where $\mathbf{p}$ is some path containing x. This is well-defined since $|Y_{\mathbf{p}}(x)|$ is *independent* of $\mathbf{p}$ for any $\mathbf{p} \ni x$. In fact,

$$s(x) = \left| \left\{ i \in I_\ell : \big| i - |x| \big| \geq \frac{\varepsilon\ell}{32C_\varepsilon} - 1 \right\} \right| \tag{3}$$

The following claim is just a routine calculation.

Claim 2.2. *Suppose x, y are in the middle layers and $\|x - y\|_1 \geq \frac{\varepsilon\ell}{32C_\varepsilon} - 1$. Let $\mathcal{P}_{x,y}$ denote the set of paths containing both x and y. Define*

$$\theta_{x,y} := \frac{1}{\ell}\left(\frac{1}{s(x)} + \frac{1}{s(y)}\right)$$

Then,

$$\Pr[\mathcal{E}_{x,y}] = \theta_{x,y} \frac{|\mathcal{P}_{x,y}|}{|\mathcal{P}|} \tag{4}$$

Proof. Note that

$$\Pr[\mathcal{E}_{x,y}] = \sum_{\mathbf{p}: x,y \in \mathbf{p}} \Pr[\mathbf{p} \text{ sampled }] \cdot \Pr[x, y \text{ sampled } \mid \mathbf{p} \text{ sampled }].$$

Since $\|x-y\|_1 \geq \frac{\varepsilon\ell}{32C_\varepsilon} - 1$, $y \in Y_{\mathbf{p}}(x)$ (and vice versa). Suppose x is the first point to be sampled; this happens with probability $1/\ell$. The probability y is the second point sampled is $\frac{1}{|Y_{\mathbf{p}}(x)|}$. Arguing analogously when y is sampled first, when $x, y \in X_{\mathbf{p}}$,

$$\Pr[x, y \text{ sampled} \mid \mathbf{p} \text{ sampled }] = \frac{1}{\ell}\left(\frac{1}{|Y_{\mathbf{p}}(x)|} + \frac{1}{|Y_{\mathbf{p}}(y)|}\right) = \theta_{x,y}.$$

The proof concludes by noting that $\sum_{\mathbf{p}:x,y\in\mathbf{p}} \Pr[\mathbf{p} \text{ sampled}] = \frac{|\mathcal{P}_{x,y}|}{|\mathcal{P}|}$. $\square$

The next claim shows that for any $(x, y) \in V$, $\theta_{x,E(y)}$ is almost as large as $\theta_{x,y}$.

Claim 2.3. *For* $(x, y) \in V$, $\theta_{x,E(y)} \geq \left(1 - \frac{1}{\sqrt{n}}\right)\theta_{x,y}$

Proof. For convenience, let y' denote $E(y)$; note that $y' \succ y$ and $|y'| = |y| + 1$. There exists some path containing x, y, and y'. From (3), $s(y) \leq s(y') \leq s(y) + 1$.

Putting it all together,

$$\frac{\theta_{x,y'}}{\theta_{x,y}} = \frac{s(x)^{-1} + s(y')^{-1}}{s(x)^{-1} + s(y)^{-1}} \geq \frac{s(y')^{-1}}{s(y)^{-1}} \geq \frac{s(y)}{s(y)+1}.$$

The first inequality follows from the observation $\frac{c+a}{c+b} \geq \frac{a}{b}$ whenever $a \leq b$ and $c \geq 0$. Since $\ell = 2\lceil C_\varepsilon\sqrt{n}\rceil$, $s(y) \geq \ell - \frac{\varepsilon\ell}{16C_\varepsilon} \geq \sqrt{n}$. $\square$

Claim 2.4. *For* $(x, y) \in V$,

$$\frac{\Pr[\mathcal{E}_{x,E(y)}]}{\Pr[\mathcal{E}_{x,y}]} = \Omega\left(\frac{\varepsilon}{\sqrt{n}}\right).$$

Proof. Combining Claim 2.2 and Claim 2.3,

$$\frac{\Pr[\mathcal{E}_{x,E(y)}]}{\Pr[\mathcal{E}_{x,y}]} = \frac{\theta_{x,E(y)}|\mathcal{P}_{x,E(y)}|}{\theta_{x,y}|\mathcal{P}_{x,y}|} \geq \left(1 - \frac{1}{\sqrt{n}}\right)\frac{|\mathcal{P}_{x,E(y)}|}{|\mathcal{P}_{x,y}|} \tag{5}$$

We know exactly what both the numbers in the RHS are. Say $|x| = t$ and $|y| = t + s$. Note $s \geq \varepsilon\ell/32C_\varepsilon - 1$. Also note $|E(y)| = |y| + 1$. Then,

$$|\mathcal{P}_{x,y}| = t!s!(n-s-t)! \quad \text{and} \quad |\mathcal{P}_{x,E(y)}| = t!(s+1)!(n-s-t-1)!$$

Plugging in (5),

$$\frac{\Pr[\mathcal{E}_{x,E(y)}]}{\Pr[\mathcal{E}_{x,y}]} \geq \left(1 - \frac{1}{\sqrt{n}}\right)\frac{s+1}{n-s-t}.$$

The denominator is $\Theta(n)$ since $n/2 - C_\varepsilon\sqrt{n} \leq |y| \leq n/2 + C_\varepsilon\sqrt{n}$. The numerator is at least $\varepsilon\ell/32C_\varepsilon = \Omega(\varepsilon\sqrt{n})$, completing the proof. $\square$

Going back to (2),

$$\Pr[\text{ success }] \geq \sum_{(x,y)\in V} \Pr[\mathcal{E}_{x,y}] \cdot \frac{\Pr[\mathcal{E}_{x,E(y)}]}{\Pr[\mathcal{E}_{x,y}]} = \Omega\Big(\frac{\varepsilon}{\sqrt{n}} \cdot \sum_{(x,y)\in V} \Pr[\mathcal{E}_{x,y}]\Big)$$

Now for the punchline. Color all points in B_1 blue. By Lemma 2.1, the probability that the random path tester samples a pair (x, y) such that both points are blue is $\widetilde{\Omega}((|B_1|/2^n)^2)$.

This is the event that $x \prec y$ (or $y \prec x$), $\|y-x\|_1 \geq \frac{\varepsilon\ell}{32C_\varepsilon} - 1$, and $x, y \in B_1$. The probability of this event is exactly twice $\sum_{(x,y)\in V} \Pr[\mathcal{E}_{x,y}]$. Hence,

$$\Pr[\text{ success }] = \widetilde{\Omega}\left[\frac{\varepsilon}{\sqrt{n}} \cdot \left(\frac{|E|}{2^n}\right)^2\right]$$

□

2.3. **Wrapping it up with the dichotomy.** We state the dichotomy theorem between the total number of violated edges and largest matching of violated edges.

Theorem 3. *For any function f that is ε-far from monotone, and for any $s > 0$, either there is a matching of violated edges completely contained in the middle layers of cardinality at least $\varepsilon 2^{n-5}/s$, or the* total *number of violated edges is at least $\varepsilon s 2^{n-1}$.*

Proof. For any matching M of violated pairs, the *average length of M* is defined to be the quantity

$$|M|^{-1} \sum_{(x,y)\in M} \|y-x\|_1$$

Choose M to be *maximum* cardinality matching of violated pairs with the smallest average length, denoted by r. Since f is ε-far from being monotone, and M is a maximal matching, we know that $|M| \geq \varepsilon 2^{n-1}$. The proof follows from the following two lemmas that we prove in subsequent subsections. □

Lemma 3.1. *If the average length of M is r, then there exists a matching E of violated edges in the middle layers of the hypercube with size at least $\varepsilon 2^{n-5}/r$.*

Proof. Deferred to §2.4. □

Lemma 3.2. *If the average length of M is r, then there are at least $r\varepsilon 2^{n-1}$ violated edges.*

Proof. Deferred to §2.5. □

It is now routine to prove Theorem 1 and Theorem 2.

Proof of Theorem 1: Set $s = n^{1/6}\varepsilon^{2/3}$. Suppose the largest matching in the middle layers has size at least $\varepsilon 2^{n-5}/s$. By Lemma 2.2, we get that the path tester succeeds with probability

$$\widetilde{\Omega}\left[\frac{\varepsilon}{\sqrt{n}} \cdot \left(\frac{\varepsilon}{s}\right)^2\right] = \widetilde{\Omega}\left(\frac{\varepsilon^3}{s^2\sqrt{n}}\right) = \widetilde{\Omega}(n^{-5/6}\varepsilon^{5/3})$$

If the largest matching is at most $\varepsilon 2^{n-5}/s$, then by Theorem 3 there are at least $\varepsilon s 2^{n-1}$ violated edges. The edge tester succeeds with probability $\Omega(s\varepsilon/n) = \Omega(n^{-5/6}\varepsilon^{5/3})$. □

Proof of Theorem 2: The number of violated edges is at most the number of influential edges, $\mathbf{I}(f)2^{n-1}$. From Theorem 3, with $s = \mathbf{I}(f)/\varepsilon$ in Theorem 3, we know there exists a violated edge matching of cardinality $\Omega(\varepsilon^2 2^n/\mathbf{I}(f))$ in the middle layers. By Lemma 2.2, the path tester succeeds with probability $\widetilde{\Omega}(n^{-1/2}\varepsilon^5/\mathbf{I}^2(f)))$. We do not need the edge tester for these functions. □

2.4. Proof of Lemma 3.1.

We first state the routing theorem of Lehman and Ron.

Theorem 4 (Lehman-Ron [10]). *Let $\mathcal{S}, \mathcal{R}$ be two subsets of points of the hypercube such that $|\mathcal{S}| = |\mathcal{R}| = m$ and all points in $\mathcal{S}$ (respectively, $\mathcal{R}$) have j (respectively, i) ones, with $i < j$. Furthermore, suppose there is a mapping $\phi : \mathcal{S} \mapsto \mathcal{R}$ such that $S \succ \phi(S)$, $\forall S \in \mathcal{S}$, that is, $(S, \phi(S))$ are ancestor-descendants. Then there exists m disjoint saturated chains (vertex disjoint paths) that contain all of $\mathcal{S}$ and $\mathcal{R}$.*

For convenience, we represent the matching through ordered pairs, so if $(x, y) \in M$, then $x \prec y$ (and $f(x) = 1$, $f(y) = 0$). Recall M is a maximum cardinality matching in the violated graph with the smallest average length. Among such matchings, let M actually be one maximizing

$$\Phi(M) := \sum_{(x,y)\in M} ||x - y||_1^2$$

We prove a structural claim regarding M. Two pairs (x, y) and (x', y') *cross* if (a) there exists a z such that $x \prec z \prec y$ and $x' \prec z \prec y'$ and (b) the intervals $[|x|, |y|]$ and $[|x'|, |y'|]$ strictly cross, meaning that neither interval contains the other. (By (a), the intervals $[|x|, |y|]$ and $[|x'|, |y'|]$ must intersect.)

Claim 4.1. *There are no crossing pairs in M.*

Proof. Suppose (x, y) and (x', y') cross, then consider M' formed by deleting these pairs from M and adding (x, y') and (x', y). Note that these are valid violations due to the presence of the vertex z. Furthermore, check that $\Phi(M') > \Phi(M)$ since the sum of squares of a pair of numbers having a fixed sum increases as the maximum (of the pair) increases. □

For every two levels $i < j$ of the hypercube, let $M_{i,j} \subseteq M$ be the pairs with endpoints in these level sets. Apply Theorem 4 to get a collection of $|M_{i,j}|$ vertex disjoint paths. Each of these vertex disjoint paths contain at least one violated edge, and let $F_{i,j}$ be the set of these edges. Note that $F_{i,j}$ forms a matching. Consider the multiset F formed by the union of $F_{i,j}$ over the set $\{(i, j) : i, j \in I_\ell,\ i < j,\ j - i \le 2r\}$, *containing* duplicates. Note that all edges of F lie in the middle layers.

Claim 4.2. $|F| \ge |M|/4$.

Proof. Note that $|F| = \sum_{(i,j):i,j\in I_\ell, i<j, j-i\le 2r} |M_{i,j}|$. Since the matching M has average length r, by Markov's inequality, at least $|M|/2$ of these pairs have length at most $2r$. Furthermore, from Prop. 2.1 we get that at most $\varepsilon^5 2^n \le |M|/4$ pairs in M have endpoints not in the middle layer. Looking at the remainder, we get $\sum_{(i,j):i,j\in I_\ell, i<j, j-i\le 2r} |M_{i,j}| \ge |M|/4$. □

Claim 4.3. *No point $z \in \{0, 1\}^n$ has more than $2r$ edges of F incident on it.*

Proof. Pick a vertex z, and pick any two edges f_1 and f_2 of F incident on it. Since each $F_{i,j}$ is a matching, these must lie in different $F_{i,j}$'s. Suppose they are $F_{i,j}$ and $F_{a,b}$, where i could be a and j could be b, but not both together. Note that $i \le |z| \le j$ and $a \le |z| \le b$. We claim that $[i, j]$ and $[a, b]$ cannot cross, and therefore one must strictly lie in the other.

There can be at most $2r$ intervals containing $|z|$ satisfying such containment relationships. Thus, there can be most $2r$ edges of F incident on z.

We now prove that $[i, j]$ and $[a, b]$ cannot cross. Consider the pairs in $M_{i,j}$, and let them be $(x_1, y_1), (x_2, y_2), \ldots, (x_k, y_k)$. Note that Theorem 4 implies k vertex disjoint paths containing all these vertices. Hence, there is some permutation π such that for each $i \in [k]$, there is a path from x_i to $y_{\pi(i)}$. Let $M'_{i,j} = \{(x_i, y_{\pi(i)})\}$. Similarly, get $M'_{a,b}$. Let M' be the matching where $M_{i,j}$ and $M_{a,b}$ are replaced by $M'_{i,j}$ and $M'_{a,b}$ (and all other pairs remain). Note that M' has the same average length, same cardinality *and* $\Phi(M') = \Phi(M)$.

But now, we have a pair $(x, y) \in M'_{i,j}$ such that $x \prec z \prec y$, and a pair $(x', y') \in M'_{a,b}$ such that $x' \prec z \prec y'$. This is because z is incident to an edge in both $F_{i,j}$ and $F_{a,b}$. If $[i, j]$ and $[a, b]$ cross, then (x, y) and (x', y') cross. Since M' maximizes the potential Φ, Claim 4.1 is contradicted. □

Since the multigraph induced by F has maximum degree $2r$, there exists a matching $E \subseteq F$ of size $|E| \geq |F|/4r$. This can be obtained by picking an edge in E arbitrarily and deleting all edges incident to any of its endpoints from F. For every edge added to E there are at most $4r$ edges deleted from F. Therefore, following Claim 4.2, we get a matching E of violated edges of size $\geq |M|/16r \geq \varepsilon 2^{n-5}/r$. Furthermore, they all lie in the middle layers. This completes the proof of Lemma 3.1.

2.5. **Proof of Lemma 3.2.**

Let M_i be the set of pairs in M that cross dimension i, that is, $M_i := \{(x, y) \in M : x_i = 0, y_i = 1\}$. Lemma 3.2 follows from following theorem.

Theorem 5. *The number of violated edges across dimension i is at least $|M_i|$.*

Note that $\sum_i |M_i| = \sum_{(x,y)\in M} \|y - x\|_1 = r|M|$, since a pair (x, y) appears in precisely $|y| - |x|$ different M_i's. Theorem 5 implies that the number of violated edges is at least $\sum_i |M_i|$, and therefore, since $|M| \geq \varepsilon 2^{n-1}$. This completes the proof of Lemma 3.2. We now prove Theorem 5.

Proof. (Proof of Theorem 5.) The proof requires setting up some of the machinery of [4]. Let H be the *perfect* matching of the hypercube formed by the *edges* crossing the dimension i. Let X be the endpoints of M_i. For all $x \in X$, we define a sequence $\mathbf{S}_x$ as follows. The first term $\mathbf{S}_x(0)$ is x. For even i, $\mathbf{S}_x(i + 1) = H(\mathbf{S}_x(i))$. For odd i, if $\mathbf{S}_x(i) \in X$, or is M-unmatched, then $\mathbf{S}_x$ terminates. Otherwise, $\mathbf{S}_x(i + 1) = M(\mathbf{S}_x(i))$. Above, we have used the shorthand $M(v)$ and $H(v)$ to denote the partners of v in the matchings M and H, respectively. (Observe that we use M and not M_i. The matching M_i is only used to define X.) The best way to think about $\mathbf{S}_x$ is via alternating paths and cycles formed by the matchings M and H. We start at x and take the H-edge along the alternating path. We keep on moving till we reach an endpoint or another vertex in X. Thus, each $\mathbf{S}_x$ terminates. It is not hard to see that if $\mathbf{S}_x$ ends at $y \in X$, then $\mathbf{S}_y$ is just $\mathbf{S}_x$ in reverse. Furthermore, $\mathbf{S}_x$ and $\mathbf{S}_y$ are disjoint unless y terminates $\mathbf{S}_x$. Therefore the number of sequences is at least $|X|/2 = |M_i|$. The theorem therefore is a consequence of the following lemma which is the heart of the proof. □

Lemma 5.1. *For all x, $\mathbf{S}_x$ contains a violated edge in H.*

Proof. We will prove this through contradiction, and will henceforth assume that (for some x), $\mathbf{S}_x$ has no violated edge in H. We show that $\mathbf{S}_x$ cannot terminate, completing the contradiction. For brevity, let us use s_i to denote $\mathbf{S}_x(i)$. Let (x, y) be the pair in M_i. We use s_{-1} to denote y. Wlog, assume $x \succ y$, thus $x_i = 1$ and $y_i = 0$. Also $f(y) = 1$ and $f(x) = 0$ since the pair is a violation.

Let B_b $(b = \{0, 1\})$ be the $n - 1$ dimensional hypercube where ith coordinate is b. We will use $d(x, x')$ for the Hamming distance between two points x and x'. We have the following simple claim.

Claim 5.1. *Let $j \geq 0$ be an index and suppose s_j exists. For $j \equiv 0 \pmod 4$, $f(s_j) = 0$ and $s_j \in \mathsf{B}_1$; $j \equiv 1 \pmod 4$, $f(s_j) = 0$ and $s_j \in \mathsf{B}_0$; $j \equiv 2 \pmod 4$, $f(s_j) = 1$ and $s_j \in \mathsf{B}_0$; $j \equiv 3 \pmod 4$, $f(s_j) = 1$ and $s_j \in \mathsf{B}_1$.*

Proof. We prove by induction on j. For the base case, $s_0 = x$, and $f(x) = 0$, $s_0 \in \mathsf{B}_1$. Consider $j \equiv 0 \pmod 4$, $j \geq 1$. By the induction hypothesis, $f(s_{j-1}) = 1$ and $s_{j-1} \in \mathsf{B}_1$. Since $s_j = M(s_{j-1})$ and (s_{j-1}, s_j) is a violation, $f(s_j) = 0$ and $s_j \in \mathsf{B}_1$. Consider $j \equiv 1 \pmod 4$. By the induction hypothesis, $f(s_{j-1}) = 0$ and $s_{j-1} \in \mathsf{B}_1$. Since $s_j = H(s_{j-1})$ and (s_{j-1}, s_j) is not a violation, $f(s_j) = 0$ and $s_j \in \mathsf{B}_0$. The remaining cases are analogous. □

Claim 5.2. *Let $j \geq 0$ be even. Then (s_j, s_{j+3}) is a violation and $d(s_j, s_{j+3}) = d(s_{j+1}, s_{j+2})$. Also, the pair (y, s_1) is a violation and $d(y, s_1) = d(y, s_0) - 1$.*

Proof. Suppose $j \equiv 2 \pmod 4$. Then from the above claim $f(s_j) = 1$ and $f(s_{j+3}) = 0$. The claim also implies $f(s_{j+1}) = 1$ and since $(s_{j+1}, s_{j+2}) \in M$, $s_{j+2} \succ s_{j+1}$. Furthermore, both $s_{j+1}, s_{j+2} \in \mathsf{B}_1$. Since $s_j = H(s_{j+1})$ and $s_{j+3} = H(s_{j+2})$, we get (a) $s_{j+3} \succ s_j$ as well, implying (s_j, s_{j+3}) is a violation, and (b) $d(s_j, s_{j+3}) = d(s_{j+1}, s_{j+2})$. The case $j \equiv 0 \pmod 4$ is analogous.

Observe that (y, s_0) is a violation where $y \in \mathsf{B}_0$ and $s_0 \in \mathsf{B}_1$. Since $s_1 = H(s_0)$, s_1 has the same coordinates as s_0 except for the ith one. Also, $f(s_1) = 0$. Therefore, (y, s_1) is a violation and $d(y, s_1) = d(y, s_0) - 1$. □

We will now argue that $\mathbf{S}_x$ cannot terminate. Consider s_j for $j \equiv 1 \pmod 4$. Because $f(s_j) = 0$ and $s_j \in \mathsf{B}_0$, s_j cannot participate in a pair in M_i. Hence, $s_j \notin X$. Suppose s_j was unmatched. Consider the following set of pairs in M: $A = \{(s_k, s_{k+1}) | k \text{ odd}, -1 \leq k \leq j-2\}$. Suppose we replaced these pairs in M by $B = \{(y, s_1)\} \cup \{(s_k, s_{k+3}) | k \text{ even}, 0 \leq k \leq j - 3\}$. Note that $|A| = \lceil j/2 \rceil = |B|$. Also, by Claim 5.2

$$d(y, s_0) + \sum_{\substack{k=1 \\ k \text{ odd}}}^{j-2} d(s_k, s_{k+1}) = d(y, s_1) + \sum_{\substack{k=0 \\ k \text{ odd}}}^{j-3} d(s_k, s_{k+3}) - 1$$

This means that replacing A and inserting B (in M) leads to a violation matching of the same size with a smaller Hamming distance. This violates the property of M (that of minimum average Hamming distance), and therefore, the sequence $\mathbf{S}_x$ cannot terminate. This cannot occur, and therefore, every $\mathbf{S}_x$ must contain a violated edge in H. This ends the proof of Lemma 5.1. □

3. Conclusion

In this paper, we make progress on the question of testing monotonicity of Boolean functions over the hypercube. Theorem 3, Lemma 2.2 and Lemma 2.1 are tight, and the exponent of $5/6$ is the best we can hope for using our analysis. Our approach in general falls short of the $\sqrt{n}$ bound. Nevertheless, we believe the path tester (alone) is a $O(\sqrt{n})$-query monotonicity tester for Boolean functions. A possible approach is suggested by Theorem 2. Can we perform a different analysis (or even a different algorithm) for high average sensitivity functions?

References

1. N. Ailon, B. Chazelle, S. Comandur, and D. Liu, *Estimating the distance to a monotone function*, Random Structures and Algorithms **31** (2006), no. 3, 1704–1711. 1
2. E. Blais, J. Brody, and K. Matulef, *Property testing lower bounds via communication complexity*, Computational Complexity **21** (2012), no. 2, 311–358. 1
3. J. Briët, S. Chakraborty, D. García-Soriano, and A. Matsliah, *Monotonicity testing and shortest-path routing on the cube*, Combinatorica **32** (2012), no. 1, 35–53. 1, 2
4. D. Chakrabarty and C. Seshadhri, *Optimal bounds for monotonicity and Lipschitz testing over the hypercube*, Tech. Report TR12-030, ECCC, April 2012. 1, 3, 10
5. Y. Dodis, O. Goldreich, E. Lehman, S. Raskhodnikova, D. Ron, and A. Samorodnitsky, *Improved testing algorithms for monotonicity*, Proceedings of the 3rd International Workshop on Randomization and Approximation Techniques in Computer Science (RANDOM) (1999), 97–108. 1, 2
6. F. Ergun, S. Kannan, R. Kumar, R. Rubinfeld, and M. Viswanathan, *Spot-checkers*, Journal of Computer Systems and Sciences (JCSS) **60** (2000), no. 3, 717–751. 1
7. E. Fischer, E. Lehman, I. Newman, S. Raskhodnikova, and R. Rubinfeld, *Monotonicity testing over general poset domains*, Proceedings of the 34th Annual ACM Symposium on the Theory of Computing (STOC) (2002), 474–483. 1, 2, 4
8. O. Goldreich, S. Goldwasser, E. Lehman, D. Ron, and A. Samordinsky, *Testing monotonicity*, Combinatorica **20** (2000), 301–337. 1, 2
9. S. Halevy and E. Kushilevitz, *Testing monotonicity over graph products*, Random Structures and Algorithms **33** (2008), no. 1, 44–67. 1
10. E. Lehman and D. Ron, *On disjoint chains of subsets*, Journal of Combinatorial Theory, Series A **94** (2001), no. 2, 399–404. 1, 3, 9
11. M. Parnas, D. Ron, and R. Rubinfeld, *Tolerant property testing and distance approximation*, Journal of Computer and System Sciences **6** (2006), no. 72, 1012–1042. 1
12. D. Ron, R. Rubinfeld, S. Safra, and O. Weinstein, *Approximating the Influence of Monotone Boolean Functions in $O(\sqrt{n})$ Query Complexity.*, Proceedings of the 15th International Workshop on Randomization and Approximation Techniques in Computer Science (RANDOM), 2011. 2

Microsoft Research India, 9 Lavelle Road, Bangalore, 560001

E-mail address: dechakr@microsoft.com

Sandia National Labs, Livermore

E-mail address: scomand@sandia.gov

ECCC ISSN 1433-8092

http://eccc.hpi-web.de