Electronic Colloquium on Computational Complexity, Report No. 36 (2013)

Lower Bounds for Testing Properties of Functions on Hypergrid Domains

Eric Blais[1], Sofya Raskhodnikova[2], and Grigory Yaroslavtsev[2]

[1] Massachusetts Institute of Technology, eblais@csail.mit.edu
[2] Pennsylvania State University, {sofya,grigory}@cse.psu.edu

Abstract. We introduce strong, and in many cases optimal, lower bounds for the number of queries required to nonadaptively test three fundamental properties of functions $f : [n]^d \to R$ on the hypergrid: monotonicity, convexity, and the Lipschitz property.

Our lower bounds also apply to the more restricted setting of functions $f : [n] \to R$ on the line (i.e., to hypergrids with $d = 1$), where they give optimal lower bounds for all three properties. The lower bound for testing convexity is the first lower bound for that property, and the lower bound for the Lipschitz property is new for tests with 2-sided error.

We obtain our lower bounds via the connection to communication complexity established by Blais, Brody, and Matulef (2012). Our results are the first to apply this method to functions with non-hypercube domains. A key ingredient in this generalization is the set of Walsh functions, an orthonormal basis of the set of functions $f : [n]^d \to \mathbb{R}$.

1 Introduction

We consider the problem of testing properties of functions over the hypergrid[3]: given oracle access to a function $f : [n]^d \to R$, for some (finite or infinite) set $R \subseteq \mathbb{R}$, and given a property $\mathcal{P}$ of functions mapping $[n]^d$ to R, what is the minimum number of queries to f that a randomized algorithm must make to distinguish with high probability between the case where f has the property $\mathcal{P}$ from the case where f is far[4] from having the same property? We focus on *nonadaptive* tests—algorithms that must fix all their queries before observing the value of the function on any of the queried inputs.

The problem of testing properties of functions has been studied extensively (see, for example, the surveys [21,22] and the book [12]), but most of this research has been restricted to functions $f : [n] \to R$ on the line and to functions $f : \{0,1\}^d \to R$ on the hypercube. (These classes of functions correspond to the special cases of the hypergrid where $d = 1$ and $n = 2$, respectively.) The purpose of the current research is to generalize tools developed in these more specialized settings to improve our understanding of property testing of functions with general hypergrid domains. In particular, we show for the first time how the connection with communication complexity established in [3] can be applied to obtain lower bounds on functions with non-hypercube domains. We then use this method to obtain significantly stronger, and in many cases optimal, lower bounds on the number of queries required to nonadaptively test three of the most fundamental properties of functions on the hypergrid: monotonicity, convexity, and the Lipschitz property.

Monotonicity. The function $f : [n]^d \to R$ is *monotone* if for any two inputs $(x_1, \ldots, x_n), (y_1, \ldots, y_n) \in [n]^d$ that satisfy $x_1 \le y_1, \ldots, x_n \le y_n$, the function f satisfies $f(x_1, \ldots, x_n) \le f(y_1, \ldots, y_n)$.

Monotonicity testing is a classic problem in property testing that has been studied extensively for functions on the line [9,10], on the hypercube [13,8,11,3,6,5], on general partially ordered set domains [11], and on hypergrid domains as well: Dodis et al. [8] showed that we can test whether $f : [n]^d \to [r]$ is monotone with $O(d \log n \log r)$ queries. Ailon and Chazelle [1] gave an alternative algorithm with the incomparable

[3] We use $[n]$ to denote the set $\{1, 2, \ldots, n\}$.

[4] See Section 2 for the formal definitions. For the purposes of this introduction, we say that f is far from having the property $\mathcal{P}$ if we need to modify the value of f on a constant fraction of its inputs to turn it into a function that does satisfy $\mathcal{P}$.

ISSN 1433-8092

Functions on the hypergrid

	Our lower bounds	Previous lower bounds	Upper bounds
Monotonicity	$\Omega(d \log n)$	$\Omega(d)$ (adaptive, $n = 2$) [3]	$O(d \log n)$ [6]
Convexity	$\Omega(d \log n)$	—	—
Lipschitz	$\Omega(d \log n)$	$\Omega(d)$ (adaptive, $n = 2$) [15]	$O(d \log n)$ [6]

Functions on the line

	Our lower bounds	Previous lower bounds	Upper bounds
Monotonicity	$\Omega(\min\{\log r, \log n\})$	$\Omega(\min\{\log r, \log n\})$ (1.-s. err.) [9] $\Omega(\log n)$ (adaptive, $r \gg n$) [9,10]	$O(\log n)$ [9]
Convexity	$\Omega(\log n)$ ($r = \Omega(n^2)$)	—	$O(\log n)$ [17]
Lipschitz	$\Omega(\min\{\log r, \log n\})$	$\Omega(\min\{\log r, \log n\})$ (1-s. err.) [15]	$O(\log n)$ [15]

Table 1. Query complexity bounds for testing properties of the function $f : [n]^d \to \mathbb{Z}$ (top) and of the function $f : [n] \to [r]$ (bottom). All the bounds are for nonadaptive tests with two-sided error unless marked otherwise.

query complexity $O(d2^d \log n)$. Very recently, Chakrabarty and Seshadhri [6] improved on both these results by showing that $O(d \log n)$ queries are sufficient for the task.

Prior to this work, however, the only known query complexity lower bounds for the problem of testing whether the function $f : [n]^d \to \mathbb{Z}$ is monotone were for two special cases: When $n = 2$, (i.e., for the hypercube), we know that $\Omega(d)$ queries are required to test monotonicity [3] and that this bound is optimal [5]. And when $d = 1$ and r is large enough, we know that $\Theta(\log n)$ queries are both necessary and sufficient for testing monotonicity [9,10].

We give the first lower bound for testing monotonicity of functions on general hypergrid domains. Furthermore, the bound that we obtain is optimal for nonadaptive tests, since it matches the upper bound of Chakrabarty and Seshadhri [6].

Theorem 1.1. *Fix $\epsilon \in (0, \frac{1}{8}]$; $m, r \in \mathbb{N}$. Let $n = 2^m$. Any nonadaptive ϵ-test for monotonicity of functions $f : [n]^d \to [nd]$ must make $\Omega(d \log n)$ queries.*

The special case $d = 1$ of the theorem also gives the first nontrivial lower bound on the query complexity of two-sided error monotonicity tests for functions $f : [n] \to [r]$ on the line when r is subexponential in n.[5]

Convexity. The function $f : [n]^d \to R$ is *convex* if for all $x, y \in [n]^d$ and all $\rho \in [0, 1]$ such that $\rho x + (1-\rho)y \in [n]^d$, the function f satisfies $f(\rho x + (1-\rho)y) \le \rho f(x) + (1-\rho)f(y)$.

Convexity testing is another classic problem in property testing. This problem was first studied by Parnas, Ron, and Rubinfeld [17], who showed that we can test if $f : [n] \to \mathbb{R}$ is convex with $O(\log n)$ queries. In the same paper, they proposed two open problems: to understand the testing of closely-related properties, and to examine the problem of testing convexity in the setting of functions $f : [n]^d \to \mathbb{R}$ on the hypergrid. While there has been much work on the first open problem—including results on testing submodularity [24,20], convexity of images [19], and convexity of geometric sets in $\mathbb{R}^d$ [18]—our lower bound represents the first progress on the study of testing convexity on the hypergrid.

Theorem 1.2. *Fix $\epsilon \in (0, \frac{1}{8}]$; $m, r \in \mathbb{N}$. Let $n = 2^m$. Any nonadaptive ϵ-test for convexity of functions $f : [n]^d \to \mathbb{R}$ must make $\Omega(d \log n)$ queries.*

[5] Strictly speaking, our result gives the first lower bound in the two-sided error model for *any* finite r, but the Ramsey theory arguments of Fischer [10] can be extended to finite ranges when r is large enough.

We also prove a matching lower bound for *separate convexity*, a closely-related (but weaker) property of functions on hypergrids (see Definition 4.8). Our results also hold for testing concavity.

The special case of our lower bound for $d = 1$ gives the first lower bound for testing convexity on the line.[6] This lower bound matches the query complexity of the nonadaptive test of Parnas, Ron, and Rubinfeld [17], showing that their algorithm and our lower bound are both optimal.

Lipschitz property. The function $f : [n]^d \to R$ is *Lipschitz* if for any two inputs $(x_1, \ldots, x_n), (y_1, \ldots, y_n) \in [n]^d$, the function f satisfies $|f(x_1, \ldots, x_n) - f(y_1, \ldots, y_n)| \le \sum_{i=1}^{n} |x_i - y_i|$.

The problem of testing the Lipschitz property on functions with hypergrid domains has applications to data privacy and program checking [15,7]. Notably, Dixit et al. [7] have used Lipschitz testers to construct privacy testers. Motivated by these applications, Jha and Raskhodnikova [15] initiated the study of testing whether functions are Lipschitz. They showed that testing if a function $f : \{0,1\}^d \to [r]$ is Lipschitz can be done with $O(\min\{d^2, dr\})$ queries, that testing $f : [n] \to [r]$ for the same property can be done with $O(\log \min\{n, r\})$ queries, and that the latter bound is optimal for nonadaptive tests with one-sided error. Awasthi et al. [2] gave the first algorithms for testing the Lipschitz property for functions $f : [n]^d \to [r]$ on the hypergrid, showing that $O(\min\{d^{3/2} n \log n, dr \log r, dr \log n\})$ queries suffice for the task. Finally, Chakrabarty and Seshadhri [6] improved this bound for arbitrary ranges by showing that $O(d \log n)$ queries suffice for testing whether $f : [n]^d \to \mathbb{R}$ is Lipschitz.

We give the first lower bound for testing the Lipschitz property for functions with hypergrid domains.

Theorem 1.3. *Fix $\epsilon \in (0, \frac{1}{8}]$; $m, r \in \mathbb{N}$. Let $n = 2^m$. Any nonadaptive ϵ-test for the Lipschitz property of functions $f : [n]^d \to [R]$, where $R = \Omega(dn)$ must make $\Omega(d \log n)$ queries.*

The lower bound in the theorem is optimal: it shows that no nonadaptive test can improve on the query complexity of Chakrabarty and Seshadhri's test in the hypergrid setting. The special case of the theorem when $d = 1$ is also optimal; showing that the algorithm of Jha and Raskhodnikova for the line is optimal, even if we allow two-sided error.

Our techniques. We obtain our lower bounds by exploiting the connection between property testing and communication complexity discovered in [3]. This connection, which we describe in Section 2, gives a method for establishing reductions between property testing problems and (a special class of) communication problems. These reductions then let us build on the rich body of work in communication complexity to prove lower bounds in property testing. This approach has been particularly successful in establishing lower bounds for testing properties of functions over the hypercube [3,4,15], but until the present work it had yet to be applied to functions over other domains.

One important reason why the previous lower bounds were restricted to properties of functions over the hypercube is that a key ingredient in all these proofs is the use of parity functions—the set of functions $\chi_S : \{0,1\}^d \to \{0,1\}$, $S \subseteq [d]$, defined by $\chi_S(x) = \sum_{i \in S} x_i \pmod 2$. These functions form an orthonormal basis for the set of functions mapping $\{0,1\}^n$ to $\mathbb{R}$, a fact that is exploited in the reductions.

We prove our lower bounds for functions with hypergrid domains by replacing the use of parity functions with *Walsh functions*, a set of functions that forms an orthonormal basis of the functions mapping $[n]^d$ to $\mathbb{R}$. These functions offer different challenges in the construction of reductions and their analysis, but the resulting lower bounds are as clean and natural as the ones obtained for functions on the hypercube.

Remarks on adaptivity. All the lower bounds introduced in this paper are for nonadaptive tests—that is, tests that must fix all their queries in advance, before observing the value of the function on any of the inputs that are queried. Interestingly, *all* the best known upper bounds on the query complexity of testing monotonicity, convexity, or the Lipschitz property (for functions over any domain) are achievable with nonadaptive tests and *nearly all* the lower bounds for testing these properties only apply to nonadaptive

[6] One might ask whether the lower bound for testing monotonicity on the line immediately implies a matching lower bound for testing convexity. It does not, and while it is certainly possible that a direct argument showing this implication exists, this argument would certainly not be trivial (c.f. [24]).

tests. In fact, apart from the result in [3], the only lower bound for monotonicity testing that applies to adaptive tests is obtained via an intricate argument of Fischer [10] that uses deep results in Ramsey theory to show that adaptivity cannot help in this setting.

In light of this general phenomenon, our results suggest two promising directions for future research: if one believes that the upper bounds can be improved in general, then our lower bounds imply that a completely new algorithmic approach which critically makes use of adaptivity will be required; conversely, if one believes that adaptivity does not help for these testing problems, then the proof of this statement will probably lead to a better understanding of the role of adaptivity in property testing and stronger connections between property testing and Ramsey theory or other areas of combinatorics.

Organization. The basic definitions and facts for property testing and communication complexity are introduced in Section 2. In Section 3, we prove our lower bounds for functions on the line; the more general lower bounds for functions with hypergrid domains are presented in Section 4.

2 Preliminaries

Property testing. The basic property testing definitions are as follows. For a more thorough introduction to the area, we recommend [21,22].

Definition 2.1 (Relative distance to a property). *Let $\mathcal{P}$ be a property (i.e., a set) of functions on a domain D, with range R and consider a function $f : D \to R$. The* relative distance *of f to the property is the minimum over all functions $g \in \mathcal{P}$ of the fraction of points in D on which f and g differ. We say f is ϵ-far from $\mathcal{P}$ if its relative distance from $\mathcal{P}$ is at least ϵ.*

Definition 2.2 (Property test [14,23]). *Fix $\epsilon \in (0, 1)$. A* (two-sided *error,* adaptive) *ϵ-test for a property $\mathcal{P}$ is a randomized algorithm which, given oracle access to a function f, accepts with probability at least 2/3 if $f \in \mathcal{P}$, and rejects with probability at least 2/3 if f is ϵ-far from $\mathcal{P}$.*

A test has one-sided error *if it always accepts functions in $\mathcal{P}$. It is* nonadaptive *if the queries to f do not depend on the answers to the previous queries.*

Communication complexity. In a (two-player) *communication game C*, Alice receives some input a, Bob receives some input b, and they must compute the value of some function $f_C(a, b)$ on their joint input. A *protocol* defines how Alice and Bob communicate. The maximum number of bits exchanged by Alice and Bob during the execution of a protocol over the possible inputs a and b is the *complexity* of the protocol. A randomized protocol is valid for f_C if for every input, the protocol computes f_C correctly with probability at least 2/3. The *communication complexity* of f_C is the minimum complexity of any protocol that is valid for f_C.

A number of different communication models have been extensively studied. We focus on the *one-way shared randomness* model. In this model, the only communication allowed is directed from Alice to Bob. Alice and Bob share access to a common source of randomness that can be used to determine the protocol. The communication complexity of f_C in the one-way shared randomness model is denoted $R^{A\to B}(f_C)$.

A fundamental function f_C studied in the one-way shared randomness model is Augmented Index. Alice's input to this function is a set $A \subseteq [t]$ while Bob's input is an index $i \in [t]$ and the set $B = A \cap [i-1]$. The output of Augmented Index is 1 if $i \in A$ and 0 otherwise. No randomized one-way communication protocol for this function does significantly better than the naïve protocol where Alice communicates her whole set to Bob.

Theorem 2.3 ([16]). *The one-way communication complexity of* Augmented Index *in the shared randomness model is $R^{A\to B}(\text{Augmented Index}) = \Theta(t)$.*

The connection between communication complexity and property testing is established via *combining operators*. An operator ψ that takes as input a and b, the inputs of Alice and Bob, respectively, and outputs a function $h(x) = \psi[a, b](x)$ is called a *one-bit one-way combining operator* if for all x in the domain of h, Bob can compute the value $h(x)$ with only one bit of communication from Alice. Next we summarize the conditions on the combining operator which are sufficient for a successful reduction from the AUGMENTED INDEX communication game to the problem of testing a given property.

Definition 2.4 (Reduction operator). *Let $t \in \mathbb{N}$ be a parameter. A combining operator $\psi[A, i, B]$ is called a* reduction operator *for (the* AUGMENTED INDEX *problem and) a property*[7] *$\mathcal{P}_t$ and a value $\epsilon_0 \in (0, 1)$ if it is a one-bit one-way combining operator and the function $h = \psi[A, i, B]$ satisfies the following two conditions for all valid inputs $A \subseteq [t]$, $i \in [t]$ and $B = A \cap [i - 1]$ to* AUGMENTED INDEX:

1. *If* AUGMENTED INDEX$(A, i, B) = 0$ *then $h \in \mathcal{P}_t$.*
2. *If* AUGMENTED INDEX$(A, i, B) = 1$ *then h is ϵ_0-far from $\mathcal{P}_t$.*

The following lemma is implicit in [3].

Lemma 2.5 (Reduction lemma). *If there exists a reduction operator for (the* AUGMENTED INDEX *problem and) a property $\mathcal{P}_t$ and a value $\epsilon_0 \in (0, 1)$ then for all $\epsilon \in (0, \epsilon_0]$, every nonadaptive ϵ-test for $\mathcal{P}_t$ requires $\Omega(t)$ queries.*

Proof. To prove the lemma, we reduce the AUGMENTED INDEX communication game to the problem of ϵ-testing property $\mathcal{P}_t$ and then apply Theorem 2.3.

Let $\psi[A, i, B]$ be a reduction operator. Consider a nonadaptive ϵ-test T for $\mathcal{P}_t$ that makes at most $q(t)$ queries for some $\epsilon \in (0, \epsilon_0]$. Then the following protocol for AUGMENTED INDEX uses $q(t)$ bits of communication from Alice to Bob. In this protocol, both players run the test T using shared randomness to find out the positions $x_1, \ldots, x_q$ queried by the test. Since T is nonadaptive, they can run it on any input of the right size. Then Alice sends to Bob $q \le q(t)$ bits of information that Bob needs to compute $h(x_1), \ldots, h(x_q)$, where $h = \psi[A, i, B]$. One bit per query point is sufficient because ψ is a one-bit one-way combining operator. Bob answers the queries of T with $h(x_1), \ldots, h(x_q)$ and outputs 0 if T accepts and 1 otherwise. The correctness of the protocol for the cases when AUGMENTED INDEX(A, i, B) is 0 and 1, respectively, follows from the conditions 1 and 2 of Definition 2.4.

The reduction establishes that $R^{A \to B}$(AUGMENTED INDEX) $\le q(t)$. Consequently, by Theorem 2.3, the query complexity of the test, $q(t)$, is $\Omega(t)$. □

3 Lower bounds on the line

We use two classes of functions on the domain $[2^m]$ (where[8] $m \in \mathbb{N}$) in the constructions that establish the lower bounds on the query complexity for testing properties of functions on the line: step functions and Walsh functions. Functions in both classes are constant on *blocks* of inputs in $[2^m]$, which we define next.

Definition 3.1 (Blocks). *Let $i \in \{0, \ldots, m\}$. For $k \in [2^{m-i}]$, the kth block of length 2^i is the set of integers $[2^i(k - 1) + 1, \ldots, 2^i k]$. We denote this block B^i_k.*

Observe that blocks of length 2^i partition $[2^m]$.

Definition 3.2 (Step functions). *For $i \in \{0, \ldots, m\}$, the* step function *of block length 2^i is the function $s_i : [2^m] \to [2^{m-i}]$ defined by $s_i(x) = k$, such that $x \in B^i_k$. (Equivalently, $s_i(x) = \lfloor \frac{x-1}{2^i} \rfloor + 1$.)*

[7] In our reductions, the integer t is used to parameterize the size of the domain of functions under consideration. Specifically, for functions on the d-dimensional hypergrid domain $[n]^d$, we set $n = 2^{t/d}$.

[8] The parameter m is set to t in the application of the reduction lemma (Lemma 2.5) to testing functions on the line; in the case of d-dimensional hypergrids, it is set to t/d.

The step functions of block length 2^i are constant on each block B^i_k. Walsh functions indexed by i, which we define next, are equal to 1 on the first half of each block B^i_k and to -1 on the second half. In other words, whether they take the value 1 or -1 on input x is determined by the ith bit of the binary representation of $x-1$, denoted by $\mathsf{bit_i(x-1)}$, where the bits are numbered starting from the least significant.

Definition 3.3 (Walsh functions). *For $i \in [m]$, let $w_i : [2^m] \to \{-1,1\}$ be the function defined by $w_i(x) = (-1)^{\mathsf{bit_i(x-1)}}$. For any $S \subseteq [m]$,* Walsh function $w_S : [2^m] \to \{-1,1\}$ *corresponding to S is $w_S(x) = \prod_{i\in S} w_i(x)$. (If $S = \emptyset$ then $w_S(x) = 1$ for all x.)*

Also, we define $w_{m+1}(x) = 1$. (This is needed only in one of the proofs).

For two functions u, w we denote the the function $v(x) = u(x)w(x)$ by $u \times w$. We use $\|\cdot\|$ to denote the L_1-norm, that is, $\|u\| = \sum_x u(x)$, where the sum is over all x in the domain of u. With this notation, we can express the fundamental properties of Walsh functions that we will use in our proofs.

Proposition 3.4. *1. $\|w_S\| \geq 0$ for all sets $S \subseteq [m]$.*

2. For all sets $A, B \subseteq [m]$, Walsh function $w_{A\triangle B} : [2^m] \to \{-1,1\}$ corresponding to the symmetric difference between A and B satisfies $w_{A\triangle B} = w_A \times w_B$.

3.1 Monotonicity

Theorem 3.5. *Fix $\epsilon \in (0, \frac{1}{4}]$; $m, r \in \mathbb{N}$. Let $n = 2^m$. Any nonadaptive ϵ-test for monotonicity of functions $f : [n] \to [r]$ requires $\Omega(\min(\log n, \log r))$ queries.*

Proof. To prove the lower bound of $\Omega(\log n)$ queries (for $n < r$), we apply the reduction lemma (Lemma 2.5) with the parameter t in the lemma set to m. To get the bound of $\Omega(\log r)$ queries (for $r \leq n$), we use the same proof with t set to $\lfloor \log_2(r-1) \rfloor$, except that the sets given to Alice and Bob reside in $\{m-t+1, \ldots, m\}$ instead of $[m]$. Let ψ be the combining operator that receives Alice's set A, and Bob's index i and set B as input and returns the function $h : [2^m] \to \mathbb{Z}$ defined by

$$h(x) = 2s_i(x) + w_S(x),$$

where $S = A\triangle B = A \cap \{i, \ldots, m\}$. The range of h is $[2 \cdot 2^{t-1} + 1] = [2^t + 1]$, i.e. it is equal to $[n+1]$ when $t = m$ and to $[r]$ when $t = \lfloor \log_2(r-1) \rfloor$.

By Proposition 3.4(2), $w_S = w_A \times w_B$. Bob knows B, so to determine $h(x)$ he only needs Alice to communicate a single bit—namely, the value $w_A(x)$. Thus, ψ is a one-bit one-way combining operator.

To prove that ψ is a reduction operator for monotonicity of functions on the line and $\epsilon_0 = 1/4$, it remains to show that it satisfies Items 1 and 2 of Definition 2.4. To demonstrate this, we prove the following lemma, which in addition to the required statements about h, also contains a statement about a related function h_- needed in Section 4.1.

Lemma 3.6. *Fix $i \in [m]$ and $S \subseteq \{i, \ldots, m\}$. Consider the functions $h = 2s_i + w_S$ and $h_- = 2s_i - w_S$.*

1. *If $i \notin S$, then the functions h and h_- are monotone;*
2. *If $i \in S$, then the function h is $\frac{1}{4}$-far from monotone.*

Proof. Recall the definition of blocks (Definition 3.1). When $i \notin S$, i.e., $S \subseteq \{i+1, \ldots, m\}$, the functions s_i, w_S and $-w_S$ are constant on each block B^i_k (for $k \in [2^{m-i}]$). The value of of functions w_S and $-w_S$ can decrease (from 1 to -1) only between the blocks (i.e., if $w_S(x) > w_S(x+1)$ then $x \in B^i_k$ and $(x+1) \in B^i_{k+1}$ for some $k \in [2^{m-i} - 1]$). But the step function s_i increases by 1 on each subsequent block B^i_k. Thus, h and h_- are monotone (nondecreasing). This completes the proof of Item 1.

When $i \in S$, i.e., i is the smallest element in S, Walsh function w_S changes value in the middle of each block B^i_k. If this change is from 1 to -1, then w_S is 1/2-far from monotone on this block, and so is h because the step function s_i is constant on each B^i_k. Note that this change is from 1 to -1 for all blocks on which $w_{S\setminus\{i\}}$ evaluates to 1. By Proposition 3.4(1), $\|w_{S\setminus\{i\}}\| \geq 0$, so it happens for at least half of the blocks. Thus, h is $\frac{1}{4}$-far from monotone. ∎

By Lemma 3.6, the function h is monotone when $i \notin A$ and it is $\frac{1}{4}$-far from monotone when $i \in A$. That is, ψ is a reduction operator for monotonicity of functions of the form $f : [2^m] \to [t+1]$ and $\epsilon_0 = 1/4$. Then, by Lemma 2.5, any nonadaptive ϵ-test for this property, where $\epsilon \in (0, 1/4]$ requires $\Omega(t)$ queries. That is, when $r > n$, we get a bound of $\Omega(m) = \Omega(\log n)$, and when $r \leq n$, we get a bound of $\Omega(\log r)$. ◻

3.2 Convexity

Recall that the function $f : [n] \to \mathbb{R}$ is convex if for all $x, y \in [n]$ and all $\rho \in [0, 1]$ such that $\rho x + (1-\rho)y$ is also an integer in $[n]$, the function f satisfies $f(\rho x + (1-\rho)y) \leq \rho f(x) + (1-\rho)f(y)$. Equivalently, we can define convexity in terms of the discrete derivative of functions on the line.

Definition 3.7 (Discrete derivative). *The* discrete derivative *of the function $f \colon [n] \to \mathbb{R}$ is the function $f' \colon [n-1] \to \mathbb{R}$ defined by $f'(x) = f(x+1) - f(x)$.*

Definition 3.8 (Convexity, concavity). *The function $f : [n] \to \mathbb{R}$ is* convex *(resp.,* concave*) if its derivative f' is a monotone nondecreasing (resp., nonincreasing) function.*

We present a lower bound for testing the convexity of functions on the line; the same bound also applies for testing concavity.

Theorem 3.9. *Fix $\epsilon \in (0, \frac{1}{8}]$ and $n = 2^m$ for some $m \geq 1$. Any nonadaptive ϵ-test for convexity of functions $f : [n] \to [r]$, where $r = \Omega(n^2)$, requires $\Omega(\log n)$ queries.*

Proof. We apply the reduction lemma (Lemma 2.5) with the parameter t in the lemma set to m. Our construction for the lower bound uses Walsh functions and *rising-step-size functions*, which are built from step functions (see Definition 3.2). The discrete derivatives of these new functions, which we call *double-step functions*, play a crucial role in our construction.

Definition 3.10 (Rising-step-size and double-step functions). *Fix $i \in [m]$. The* rising-step-size function *$r_i : [n] \to [n^2]$ is defined by $r_i(x) = s_i(x) + 2\sum_{y=1}^{x-1} s_i(y)$. Its discrete derivative, $r_i'(x) = s_i(x+1) + s_i(x)$, is called a* double-step function*. Equivalently (by Definitions 3.1 and 3.2), for all $k \in [2^{m-i}]$, function $r_i'(x)$ is equal to $2k$ on all but the last element x of the block B_k^i, and to $2k+1$ on the last element of B_k^i.*

Given Alice's set $A \subseteq [m]$ and Bob's index $i \in [m]$ and the prefix set $B = A \cap [i-1]$ the combining operator $\psi[A, i, B]$ returns the function

$$h(x) = r_i(x) + \frac{1}{2}(w_S(x) + 1),$$

where $S = A \triangle B = A \cap \{i, \ldots, m\}$. Since $w_S = w_A \times w_B$, the operator ψ is a one-bit one-way combining operator. It remains to show that if $i \notin A$, then h is convex and that if $i \in A$, then h is 1/8-far from convex. We do so in the following lemma, which is also used in Section 4.3.

Lemma 3.11. *Fix $i \in [m]$ and $S \subseteq \{i, \ldots, m\}$. The functions $h = r_i + \frac{1}{2}(w_S + 1)$ and $h_- = r_i - \frac{1}{2}(w_S + 1)$ satisfy the following properties.*

1. *If $i \notin S$, then h and h_- are convex;*
2. *If $i \in S$, then h is $\frac{1}{8}$-far from convex.*

Proof. First, consider the case where $i \notin S$. The discrete derivative of h is $h'(x) = r_i'(x) + \frac{1}{2}w_S'(x)$. It is sufficient to prove that h' is nondecreasing. Since $S \subseteq \{i+1, \ldots, m\}$, the function w_S is constant on each block B_k^i (for $k \in [2^{m-i}]$). That is, for all but the last element x of a block B_k^i, the discrete derivative $w'(x) = 0$ and, consequently, $h'(x) = r_i'(x) = 2k$. Now consider $h'(x)$, where x is the last element of a block B_k^i. Recall that $r_i'(x) = 2k+1$. Since Walsh functions are ± 1-valued, the value $\frac{1}{2}w_S'(x)$ is in $\{-1, 0, 1\}$. Thus,

$h'(x) \in [2k, 2k+2]$, i.e., $h'(x-1) \le h'(x) \le h'(x+1)$. Therefore, h' is a nondecreasing function. The same argument shows that when $i \notin S$, the function h_- is also convex.

Now consider the case where $i \in S$. We start the analysis of this case by showing that for at least half of the blocks B_k^i, the derivative $w'_S(x) = -2$ on the 2^{i-1}th element of B_k^i (i.e., on the input $x = 2^i(k-1)+2^{i-1}$.) Note that $w_S = w_i \times w_{S\setminus\{i\}}$. Proposition 3.4(1) gives that $\|w_{S\setminus\{i\}}\| \ge 0$, implying that $\Pr_x[w_{S\setminus\{i\}}(x) = 1] \ge 1/2$. Since $S \cap [i-1] = \emptyset$, the function $w_{S\setminus\{i\}}$ is constant within the blocks B_k^i. Thus, for at least half of these blocks it is a constant 1. For each block B_k^i, the function w_i is 1 on the first half of the block and -1 on the second half. Combining these observations, for half of the blocks B_k^i, the derivative of w_S on the middle point $x = 2^i(k-1) + 2^{i-1}$ of the block satisfies $w'_S(x) = w_S(x+1) - w_S(x) = w_{S\setminus\{i\}}(x+1) \cdot w_i(x+1) - w_{S\setminus\{i\}}(x) \cdot w_i(x) = -2$.

Let B_k^i be a block where $w'_S(x) = -2$ on the 2^{i-1}th element x of B_k^i. Note that $w'_S(x) = 0$ on all other inputs in the block apart from the last one because w_S is constant on all blocks B_j^{i-1}. Consider any three points $x, y, z \in B_k^i$ such that $x \le (k-1)2^i + 2^{i-1} < y < z$, namely, x is in the first half of the block B_k^i while y and z are in the second half. Then $h'(y) = h'(y+1) = \cdots = h'(z-1) = 2k$ so $(h(z) - h(y))/(z-y) = 2k$. However, $h'((k-1)2^i + 2^{i-1}) = 2k-2$ so $(h(y)-h(x))/(y-x) < 2k$, which violates convexity. To fix convexity on all such triples, we must change the value of h on all the points $(k-1)2^i + 1, \dots, (k-1)2^i + 2^{i-1}$ in the first half of the block B_k^i, or on all but one point in the second half of B_k^i. Thus, we need to change at least $1/4$ of the points in B_k^i. Since this is the case for at least half of all blocks, h is $1/8$-far from convex. $\square$

Lemma 3.11 completes the proof that ψ is a reduction operator for convexity and, thus, of the claimed lower bound for convexity. $\square$

3.3 The Lipschitz property

Theorem 3.12. *Fix $\epsilon \in (0, \frac{1}{4}]$; $m, r \in \mathbb{N}$. Let $n = 2^m$. Any nonadaptive ϵ-test for the Lipschitz property of functions $f : [n] \to [r]$ requires $\Omega(\min(\log n, \log r))$ queries.*

Proof. To obtain the $\Omega(\log n)$ bound (for $r > n$), we apply the reduction lemma (Lemma 2.5) with the parameter t in the lemma set to m.

Definition 3.13 (Up-down staircase functions). *For all $i \in \{0, 1, \dots, m\}$, let the* up-down staircase function *of block-length 2^i be the function $u_i : [2^m] \to [2^i]$, such that $u_i(1) = 1$ and the discrete derivative of u_i is*

$$u'_i(x) = \begin{cases} 0 & \text{if } x \text{ is divisible by } 2^i; \\ w_{i+1}(x) & \text{otherwise.} \end{cases}$$

In other words (using Definition 3.1), function u_i takes values $1, \dots, 2^i$ on consecutive inputs from the block B_j^i if j is odd, and values $2^i, \dots, 1$ if j is even.

The combining operator ψ receives Alice's set A, and Bob's index i and set B as input and returns the function $h : [2^m] \to \mathbb{Z}$ defined by

$$h(x) = u_i(x) - \frac{1}{2}(w_S(x) + 1),$$

where $S = A \triangle B = A \cap \{i, \dots, m\}$. Since $w_S = w_A \times w_B$, the operator ψ is a one-bit one-way combining operator. It remains to show that if $i \notin A$, then $h(x)$ is Lipschitz and otherwise it is $1/4$-far from Lipschitz. To demonstrate this, we prove a stronger lemma, which is also used in Section 4.3.

Lemma 3.14. *Fix $i \in [m]$ and $S \subseteq \{i, \dots, m\}$. Consider the functions $h(x) = u_i(x) - \frac{1}{2}(w_S(x) + 1)$ and $h_-(x) = u_i(x) - \frac{1}{2}(-w_S(x) + 1)$.*

1. *If $i \notin S$, then the functions h and h_- are Lipschitz;*
2. *If $i \in S$, then the function h is $\frac{1}{4}$-far from Lipschitz.*

Proof. If $i \notin S$, i.e., $S \subseteq \{i+1, \ldots, m\}$, then the function w_S is constant on each block B_k^i (for $k \in [2^{m-i}]$). Let $w(x) = -\frac{1}{2}(w_S(x) + 1)$. Since Walsh functions are ± 1-valued, the discrete derivative $w'(x)$ is in $\{-1, 0, 1\}$ for all x, and $w'(x) = 0$ for all x *not* divisible by 2^i. By definition of the up-down staircase functions, $u_i'(x) \in \{-1, 0, 1\}$ for all x, and $u_i'(x) = 0$ for all x divisible by 2^i. Thus, $h' = u_i' + w'$ takes values only in $\{-1, 0, 1\}$, implying that h is Lipschitz. The proof that h_- is Lipschitz is analogous.

When $i \in S$, i.e., i is the smallest element in S, the rescaled Walsh function $w(x) = -\frac{1}{2}(w_S(x) + 1)$ changes value in the middle of each block B_k^i. This change is either from -1 to 0 or vice versa. In the former case, the discrete derivative w' is 1 on the 2^{i-1}th element of the block, in the latter, it is -1. In both cases, it is 0 on all other elements of the block besides the last one. Next we show that if the former case occurs on a block with odd i (similarly, if the latter case occurs on a block with even i), then h is 1/2-far from Lipschitz on this block.

Consider the case when i is odd and w' is 1 on the 2^{i-1}th element of a block B_k^i. Since i is odd, u_i' takes value 1 on all but the last element of B_k^i. Then $h' = u_i' + w'$ is 2 on the 2^{i-1}th element of B_k^i, and 1 on all other elements of the block besides the last one. We pair up all elements of B_k^i as follows: each element x in the first half of the block is paired up with the element $x + 2^{i-1}$. The function h is not Lipschitz on each such pair: $h(x + 2^{i-1}) - h(x) = \sum_{y=x}^{x+2^{i-1}-1} h'(y) = 2^{i-1} + 1$. Thus, h is 1/2-far from Lipschitz on each such block. The other case (when i is even and w' is -1 on the 2^{i-1}th element of a block B_k^i) is analogous—the only difference is that h' takes negative values.

We can rephrase what we just proved as follows: the function h is 1/2-far from Lipschitz on all blocks B_k^i with $k \in [2^{m-i}]$, where $w_{S\setminus\{i\}}(x) = w_{i+1}(x)$ for all $x \in B_k^i$. Equivalently, $w_{S\setminus\{i\}}(x) \times w_{i+1}(x) = w_{(S\setminus\{i\})\triangle\{i+1\}}(x) = 1$ for all $x \in B_k^i$. By Proposition 3.4(1), $\|w_{(S\setminus\{i\})\triangle\{i+1\}}\| \geq 0$. Since $w_{(S\setminus\{i\})\triangle\{i+1\}}$ is constant on each block B_k^i, it is 1 on at least half of such blocks. Thus, h is 1/2-far from Lipschitz on at least half of the blocks B_k^i. That is, overall h is 1/4-far from Lipschitz. □

This completes the proof of the $\Omega(\log n)$ lower bound. To get the bound of $\Omega(\log \min\{n, r\})$, we use the same proof with t set to $\min\{m, \lfloor \log_2(r-1) \rfloor\}$. The range of h is $\{0, 1, \ldots, 2^t\}$, i.e., it has size $\min(n+1, r)$. □

4 Lower bounds on the hypergrid

In this section, we generalize the lower bounds for testing functions on the line to the hypergrid setting. To obtain our lower bounds for testing functions on the domain $[2^m]^d$, we give a reduction from the AUGMENTED INDEX problem by applying the reduction lemma (Lemma 2.5) with the parameter t set to md. With this parameter setting, inputs to AUGMENTED INDEX consist of subsets of $[md]$ and an index in $[md]$. We associate each such subset with a d-dimensional vector of subsets of $[m]$ and each such index with a d-dimensional vector of indices in $\{0, 1, \ldots, m\}$.

Definition 4.1 (Vector representation). *Fix $m, d \in \mathbb{N}$. The d-dimensional vector corresponding to the set $S \in [md]$ is $\mathbf{S} = (\mathbf{S}_1, \ldots, \mathbf{S}_d)$, where $\mathbf{S}_j = \{\ell \in [m] : (j-1)m + \ell \in S\}$ for every $j \in [d]$. The d-dimensional vector corresponding to the index $i \in [md]$ is $\boldsymbol{i} = (\boldsymbol{i}_1, \ldots, \boldsymbol{i}_d)$, where $\boldsymbol{i}_j = \max\{0, \min\{m, i - (j-1)m\}\}$ for every $j \in [d]$.*

Equivalently, $\boldsymbol{i} = (m, \ldots, m, \boldsymbol{i}_{j^*}, 0, \ldots, 0)$, where $j^* = \lceil i/m \rceil$ and $\boldsymbol{i}_{j^*} = i - (j^* - 1)m$. Observe that $i \in S$ iff $\boldsymbol{i}_{j^*} \in \mathbf{S}_{j^*}$. Recall that in the AUGMENTED INDEX problem, Bob is given an element i, and he has to find out whether i is in Alice's set. Intuitively, in our reduction from AUGMENTED INDEX to the problem of testing a property $\mathcal{P}$ (such as monotonicity) of d-dimensional functions, the function h returned by the combining operator will satisfy $\mathcal{P}$ on all axis-parallel lines in dimensions other than j^*. Most restrictions of h to lines in the special dimension j^* will behave as in the one-dimensional case: many of them will be far from $\mathcal{P}$ if i is in Alice's set; otherwise, all of them will be in $\mathcal{P}$. This suffices for the proofs for monotonicity, the Lipschitz property and for separate convexity (Definition 4.8), a property closely related (but not equivalent) to convexity. For convexity itself, it doesn't suffice to ensure that restrictions of the function on the axis-parallel lines are convex, so in this case if i is in Alice's set we construct the reduction in such a way that projections on all (not necessarily axis-parallel) lines are convex.

Next we extend the definitions of step functions and Walsh functions (namely, Definitions 3.2 and 3.3, respectively) to multiple dimensions.

Definition 4.2 (Componentwise sum). *For a family of functions $f_i\colon [n] \to \mathbb{R}$, indexed by $i \in \{0, 1, \dots, m\}$, and a vector $\boldsymbol{i} \in \{0, 1, \dots, m\}^d$, the* componentwise sum $\hat{f}_{\boldsymbol{i}}\colon [n]^d \to \mathbb{R}$ *of f is defined by $\hat{f}_{\boldsymbol{i}}(x_1, \dots, x_d) = \sum_{j=1}^{d} f_{\boldsymbol{i}_j}(x_j)$.*

Definition 4.3 (Step functions). *The* step function *indexed by the d-dimensional vector $\boldsymbol{i} \in [m]^d$ is the componentwise sum $\hat{s}_{\boldsymbol{i}} : [2^m]^d \to [d2^m]$ defined by $\hat{s}_{\boldsymbol{i}}(x_1, \dots, x_d) = \sum_{j=1}^{d} s_{\boldsymbol{i}_j}(x_j)$.*

Definition 4.4 (Walsh functions). *The* Walsh function *indexed by the d-dimensional vector $\mathbf{S}$ of subsets of $[m]$ is the function $w_{\mathbf{S}} : [2^m]^d \to \{-1, 1\}$ defined by $w_{\mathbf{S}}(x_1, \dots, x_d) = \prod_{j=1}^{d} w_{\mathbf{S}_j}(x_i)$.*

Next we extend Proposition 3.4 to the hypergrid setting.

Proposition 4.5. *1. $\|w_{\mathbf{S}}\| \geq 0$ for all d-dimensional vectors $\mathbf{S}$ of subsets of $[m]$.*

2. Fix $A, B \subset [md]$, and $S = A \triangle B$. Let $\mathbf{A}, \mathbf{B}, \mathbf{S}$ be the d-dimensional vector representations of sets A, B, S, respectively. Walsh function $w_{\mathbf{S}} : [2^m]^d \to \{-1, 1\}$ satisfies $w_{\mathbf{S}}(x) = w_{\mathbf{A}}(x) \cdot w_{\mathbf{B}}(x)$ for all $x \in [2^m]^d$.

Proof (of Item 1). It is sufficient to prove that if the random variables $X_1, \dots, X_d$ are i.i.d. and uniform over $[2^m]$ then $\Pr[w_{\mathbf{S}}(X_1, \dots, X_d) = 1] \geq 1/2$. If $\mathbf{S}_j = \emptyset$ then $w_{\mathbf{S}_j}(X_j) = 1$. For all $j \in [d]$ such that $\mathbf{S}_j \neq \emptyset$, the random variables $w_{\mathbf{S}_j}(X_j) \in \{-1, 1\}$ are i.i.d. and uniformly distributed over $\{-1, 1\}$. Thus, $\Pr[w_{\mathbf{S}}(X_1, \dots, X_d) = 1] = \Pr[\prod_{j \in [d]} w_{\mathbf{S}_j}(X_j) = 1] \geq 1/2$. ◻

Corollary 4.6. *Let $\mathbf{S}$ be the d-dimensional representation of $S \subseteq [md]$. The product $\prod_{k \in [d] \setminus \{j\}} w_{\mathbf{S}_k}(x_k)$, where $x_k \in [2^m]$ for all $k \in [d] \setminus \{j\}$, evaluates to 1 for at least half of the settings of variables x_k.*

Proof. Let $\mathbf{S}'$ be the $(d-1)$-dimensional vector $(\mathbf{S}_1, \dots, \mathbf{S}_{j-1}, \mathbf{S}_{j+1}, \dots, \mathbf{S}_d)$. Then $\prod_{k \in [d] \setminus \{j\}} w_{\mathbf{S}_k}(x_k) = w_{\mathbf{S}'}(x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_d)$. By Proposition 4.5(1), this expression is 1 for at least half of the settings of x_k. ◻

4.1 Monotonicity

We extend our construction from Section 3.1 to prove Theorem 1.1.

Proof (of Theorem 1.1). We use Lemma 2.5, giving a reduction with parameter $t = md$. Let $A \subseteq [md]$ be Alice's input and $i \in [md]$ and $B = A \cap [i-1]$ be Bob's input.

The combining operator ψ is defined as follows. It receives A, i, B as input. Then it computes $S = A \triangle B = A \cap \{i, \dots, md\}$ and the d-dimensional vectors $\boldsymbol{i}$ and $\mathbf{S}$ corresponding to i and S, respectively. (See Definition 4.1. Also recall definitions of the step functions $\hat{s}_{\boldsymbol{i}}$ and Walsh functions $w_{\mathbf{S}}$ on the hypergrid—specifically, Definitions 3.2, 4.2 and 4.4.) It returns the function $h : [n]^d \to \{d-1, \dots, dn+1\}$ defined by

$$h(x) = 2\hat{s}_{\boldsymbol{i}}(x) + w_{\mathbf{S}}(x).$$

By Proposition 4.5, $w_{\mathbf{S}} = w_{\mathbf{A}} \cdot w_{\mathbf{B}}$, where $\mathbf{A}$ and $\mathbf{B}$ are d-dimensional vector representations of A and B, respectively. Bob knows i and B and can compute their vector representations. To determine $h(x)$, he only needs Alice to communicate the bit $w_{\mathbf{A}}(x)$. Thus, ψ is a one-bit one-way combining operator.

Lemma 4.7 concludes the proof that ψ is a reduction operator for monotonicity and $\epsilon_0 = 1/8$, implying the theorem. ◻

Lemma 4.7. *Fix $i \in [md]$ and $S \subseteq \{i, \dots, dm\}$, and let $\boldsymbol{i}$ and $\mathbf{S}$, respectively, be their d-dimensional vector representations.*

1. *If $i \notin S$, then the function h is monotone;*
2. *If $i \in S$, then the function h is $\frac{1}{8}$-far from monotone.*

Proof. Let $j^* = \lceil i/m \rceil$. We will show that all line restrictions of h to dimensions other than j^* are monotone. If $i \notin S$, we will show that all line restrictions of h to dimension j^* are also monotone, so h itself is monotone. Conversely, if $i \in S$, we will show that at least half of the line restrictions of h to dimension j^* are 1/4-far from monotone, so h itself is 1/8-far from monotone.

Consider the restriction of $h = 2\hat{s}_{\boldsymbol{i}} + w_{\mathbf{S}}$ to a line in dimension $j \in [d]$, i.e., a function $\bar{h} : [2^m] \to \mathbb{N}$ defined by $\bar{h}(x_j) = h(\bar{x}_1, \ldots, \bar{x}_{j-1}, x_j, \bar{x}_{j+1}, \ldots, \bar{x}_d)$, where the values $\bar{x}_k \in [2^m]$ are fixed for all $k \in [d] \setminus \{j\}$. Then

$$\begin{aligned} \bar{h}(x_j) &= 2\sum_{k \neq j} s_{\boldsymbol{i}_k}(\bar{x}_k) + 2s_{\boldsymbol{i}_j}(x_j) + w_{\mathbf{S}_j}(x_j) \cdot \prod_{k \neq j} w_{\mathbf{S}_k}(\bar{x}_k) \\ &= 2s_{\boldsymbol{i}_j}(x_j) \pm w_{\mathbf{S}_j}(x_j) + c, \end{aligned} \tag{1}$$

where $\pm$ means "either + or -" and c is a constant independent of x_j.

If $j < j^*$ then $\mathbf{S}_j = \emptyset$, $\boldsymbol{i}_j = m$ and $\bar{h} = 2s_m \pm w_\emptyset + c = 2 \pm 1 + c$. And if $j > j^*$ then $\boldsymbol{i}_j = 0$, so $\bar{h}(x_j) = 2x_j \pm w_{\mathbf{S}_j}(x_j) + c$. In both cases, the function $\bar{h}$ is monotone.

Finally, if $j = j^*$ then $\boldsymbol{i}_j = i - (j-1)m$. In this case, $i \in S$ iff $\boldsymbol{i}_j \in \mathbf{S}_j$. If $\boldsymbol{i}_j \notin \mathbf{S}_j$ then, by (1) and Lemma 3.6, $\bar{h}(x_j)$ is monotone. Since all line restrictions of $h(x)$ are monotone, the overall function $h(x)$ is monotone. Now suppose $\boldsymbol{i}_j \in \mathbf{S}_j$. Consider the product $\prod_{k \neq j} w_{\mathbf{S}_k}(\bar{x}_k)$ that determines whether the expression $\pm$ in (1) is actually a plus or a minus. By Corollary 4.6, this product evaluates to 1 for at least half of the line restrictions $\bar{h}$ of h in dimension j. For those restrictions, $\bar{h}(x_j) = 2s_{\boldsymbol{i}_j}(x_j) + w_{\mathbf{S}_j}(x_j) + c$ and, since $\boldsymbol{i}_j \in \mathbf{S}_j$, Lemma 3.6 implies that $\bar{h}$ is $\frac{1}{4}$-far from monotone. Thus, at least half of the line restrictions of h in dimension j are 1/4-far from monotone. Since the domains of line restrictions of h in dimension j partition the domain of h, it implies that the overall function $h(x)$ is $\frac{1}{8}$-far from monotone. ◻

4.2 Convexity

In this section, we give lower bounds for testing convexity and a related property called *separate convexity.*

Definition 4.8 (Separate convexity). *The function $f\colon [n]^d \to \mathbb{R}$ is* separately convex *if for all $i \in [d]$ and all sets of values $(\bar{x}_1, \ldots, \bar{x}_{i-1}, \bar{x}_{i+1}, \ldots \bar{x}_d) \in [n]^{d-1}$, the one-dimensional function $\bar{f} : [n] \to \mathbb{R}$ defined by $\bar{f}(x_i) = f(\bar{x}_1, \ldots, \bar{x}_{i-1}, x_i, \bar{x}_{i+1}, \ldots, \bar{x}_d)$ is convex.*

Separate convexity is a weaker property than (standard) convexity: every convex function is also separately convex, but the converse is not true.

Theorem 4.9. *Fix $\epsilon \in (0, \frac{1}{8}]$; $m, r \in \mathbb{N}$. Let $n = 2^m$. Any nonadaptive ϵ-test for separate convexity of functions $f : [n]^d \to [r]$ where $r = \Omega(dn^2)$ must make $\Omega(d \log n)$ queries.*

The proofs of Theorem 4.9 and Theorem 1.2 have some common elements so we present them together.

Proof (of Theorem 1.2 and Theorem 4.9). We apply Lemma 2.5 with parameter $t = md$. Let $A \subseteq [md]$ be the set received by Alice and let $i \in [md]$ and $B = A \cap [i-1]$ be Bob's input. Let $j^* = \lceil i/m \rceil$. Let $\mathbf{A}, \mathbf{B}$ and $\boldsymbol{i}$ be the d-dimensional vectors corresponding to A, B and i respectively. The combining operator ψ receives $\mathbf{A}$ and $\boldsymbol{i}$ as input and returns the function $h : [n]^d \to \mathbb{R}$ defined by

$$h(x_1, \ldots, x_d) = \alpha \left(\frac{1}{2} \left(w_{\mathbf{S}}(x) + 1 \right) + r_{\boldsymbol{i}_{j^*}}(x_{j^*}) \right) + \sum_{j=j^*+1}^{d} {x_j}^2$$

where $\mathbf{S}$ is a d-dimensional vector corresponding to $S = A \triangle B = A \cap \{i, \ldots, md\}$ and r is the family of rising-step-size function (Definition 3.10). The parameter $\alpha > 0$ will be selected later. For any $x \in [n]^d$, Bob only needs the single bit $w_{\mathbf{A}}(x)$ from Alice to compute $h(x)$, so ψ is a one-bit one-way combining operator.

To show that ψ is a reduction operator for convexity (resp., separate convexity) we need to show that if $i \notin S$ (or equivalently $\boldsymbol{i}_{j^*} \notin \mathbf{S}_{j^*}$) then h is convex (resp., separately convex) and otherwise h is $\frac{1}{8}$-far from convex (resp., separately convex).

We first show how to complete the proof using the following two lemmas and then present their proofs below.

Lemma 4.10. *If $\boldsymbol{i}_{j^*} \notin \mathbf{S}_{j^*}$ then :*

1. *For $\alpha = 1$ the function h is separately convex.*
2. *There exists a value of $\alpha > 0$ such that the function h is convex.*

Lemma 4.11. *If $\boldsymbol{i}_{j^*} \in \mathbf{S}_{j^*}$ then the function h is $\frac{1}{8}$-far from separately convex for all $\alpha > 0$.*

The proof of Theorem 4.9 for separate convexity is completed by setting $\alpha = 1$ and noting that the range of h is $[r]$ for $[r] = O(dn^2)$ because for every $k \in [m]$ the range of r_k is $O(n^2)$. In the proof of Theorem 1.2 for convexity we set α to the value from the second part of Lemma 4.10. Then Lemma 4.11 implies that h is $\frac{1}{8}$-far from convex because the distance to convexity is at least the distance to separate convexity. $\square$

Proof (of Lemma 4.10, Part 1). The proof follows by showing that every restriction of h to any dimension $j \in [d]$ is a convex function.

Every one-dimensional restriction $\bar{h}$ of h in dimension j^* can be expressed as $\bar{h}(x_{j^*}) = \alpha(r_{\boldsymbol{i}_{j^*}}(x_i) \pm \frac{1}{2}w_{\mathbf{S}_{j^*}}(x_{j^*})) + c$, where c is some constant independent of x_{j^*}. Because $\boldsymbol{i}_{j^*} \notin \mathbf{S}_{j^*}$ this function is convex by Lemma 3.11. For all $j < j^*$ every one-dimensional restriction $\bar{h}$ of h to dimension j is a constant function. For all $j > j^*$, the restrictions of h to dimension j can be expressed as $\bar{h}(x_j) = \pm\frac{1}{2}\alpha w_{\mathbf{S}_j}(x_j) + {x_j}^2 + c$. The derivative of the first term $w_{\mathbf{S}_j}$ satisfies that $|\frac{1}{2}\alpha w'_{\mathbf{S}_j}(x_j)| \leq \alpha$ and the derivative of the second term is $2x_j$, so for $\alpha \leq 1$ the derivative $\bar{h}'$ is a nondecreasing function and $\bar{h}$ is convex. Hence, the function h is separately convex for all $\alpha \leq 1$. $\square$

Proof (of Lemma 4.10, Part 2). We show how to pick a parameter $0 < \alpha < 1$ such that the function h is convex. By definition, to prove convexity we need to show that for every pair of points $(x, y) \in [n]^d \times [n]^d$ and every $0 < \gamma < 1$ for which $z = \gamma x + (1-\gamma)y \in [n]^d$, we have that $h(z) \leq \gamma h(x) + (1-\gamma)h(y)$.

The function h is independent of the first $j^* - 1$ coordinates, so $h(x) = h(y_1, \ldots, y_{j^*-1}, x_{j^*}, \ldots, x_d)$ and $h(z) = h(y_1, \ldots, y_{j^*-1}, z_{j^*}, \ldots, z_d)$.

First, consider the case when for all $j > j^*$ it holds that $x_j = y_j$ so we have $x = (x_1, \ldots, x_{j^*}, y_{j^*+1}, \ldots, y_d)$. By Lemma 4.10 (Part 1), all the restrictions $\bar{h}$ of h to dimension j^* are convex, so in this case $h(z) \leq \gamma h(x) + (1-\gamma)h(y)$.

Otherwise, fix an index $j > j^*$ such that $x_j \neq y_j$.

Proposition 4.12. *Define $\phi_{j^*}(x) = \sum_{t=j^*+1}^{d} {x_t}^2$. For all $n, d \geq 1$ there exists a value $\delta^*(n, d) > 0$ such that the inequality*

$$\phi_{j^*}(\gamma x + (1-\gamma)y) \leq \gamma\phi_{j^*}(x) + (1-\gamma)\phi_{j^*}(y) - \delta^*(n, d)$$

holds for all pairs (x, y) such that $x_j \neq y_j$ for some $j > j^*$ and all $\gamma \in (0, 1)$ such that $\gamma x + (1-\gamma)y \in [n]^d$.

Proof. We have:

$$\begin{aligned}
&\phi_{j^*}(\gamma x + (1-\gamma)y) - \gamma\phi_{j^*}(x) - (1-\gamma)\phi_{j^*}(y) \\
&\quad = \sum_{t=j^*+1}^{d} (\gamma x_t + (1-\gamma)y_t)^2 - \gamma \sum_{t=j^*+1}^{d} {x_t}^2 - (1-\gamma) \sum_{t=j^*+1}^{d} {y_t}^2 \\
&\quad = \sum_{t=j^*+1}^{d} \left((\gamma x_t + (1-\gamma)y_t)^2 - \gamma {x_t}^2 - (1-\gamma){y_t}^2\right) \\
&\quad \leq \left((\gamma x_j + (1-\gamma)y_j)^2 - \gamma {x_j}^2 - (1-\gamma){y_j}^2\right) < 0.
\end{aligned}$$

The first inequality uses convexity of x^2. The second inequality uses its strict convexity and the fact that $x_j \neq y_j$.
Let $\delta(x, y, j, \gamma, n, d) = -\left((\gamma x_j + (1-\gamma)y_j)^2 - \gamma {x_j}^2 - (1-\gamma){y_j}^2\right) > 0$. Note that j and γ can take at most d and n^d different values respectively for any fixed pair (x, y). Thus there are at most dn^{3d} different valid tuples (x, y, j, γ). The claim follows by letting $\delta^*(n, d) = \min_{x,y,j,\gamma} \delta(x, y, j, \gamma, n, d)$. $\square$

We set $\alpha = \frac{\delta^*(n,d)}{6(2n^2+1)}$. Using the notation introduced above, $h(x) = \alpha\left(\frac{1}{2}\left(w_{\mathbf{S}}(x)+1\right) + r_{\boldsymbol{i}_{j^*}}(x_{j^*})\right) + \sum_{j>j^*} {x_j}^2 = \alpha\left(\frac{1}{2}\left(w_{\mathbf{S}}(x)+1\right) + r_{\boldsymbol{i}_{j^*}}(x_{j^*})\right) + \phi_{j^*}(x)$ Because the range of $c_{\boldsymbol{i}_{j^*}}$ is $[2n^2]$,

$$\begin{aligned} h(z) - \gamma h(x) - (1-\gamma)h(y) &\leq \phi_{j^*}(z) - \gamma\phi_{j^*}(x) - (1-\gamma)\phi_{j^*}(y) + 3\alpha(2n^2+1) \\ &\leq -\delta^*(n,d) + 3\alpha(2n^2+1) = -\delta^*(n,d)/2 < 0, \end{aligned}$$

where the inequalities follows from Proposition 4.12. This concludes the proof of the fact that h is convex. □

Proof (of Lemma 4.11). If $\boldsymbol{i}_{j^*} \in \mathbf{S}_{j^*}$ then by Corollary 4.6 the product $\prod_{k\neq j^*} w_{\mathbf{S}_k}(x_k)$ evaluates to 1 for at least half of the line restrictions $\bar{h}$ of h to dimension j^*. For such restrictions, $\bar{h}(x_{j^*}) = \alpha(\frac{1}{2}w_{\mathbf{S}_{j^*}}(x_{j^*}) + r_{\boldsymbol{i}_{j^*}}(x_{j^*})) + c$, for some constant c. Lemma 3.11 implies that $\bar{h}$ is $\frac{1}{8}$-far from convex. The domains of the restrictions $\bar{h}$ of h in dimension j^* partition the domain of h, so we conclude that the function h is $\frac{1}{8}$-far from separately convex. □

4.3 The Lipschitz property

In this section, we extend our construction from Section 3.3 to prove Theorem 1.3.

Proof (of Theorem 1.3). The starting point of the reduction is the same as in the proof of the lower bound for monotonicity in Section 4.1. We use the same notation for the parameters of the reduction from Augmented Index, Alice's and Bob's inputs, the set $S = A \triangle B = A \cap \{i, \ldots, md\}$ and the vector representation of these objects. Let $\hat{u}_{\boldsymbol{i}}$ be the componentwise sum (see Definition 4.2) of the up-down staircase functions u_i (see Definition 3.13). The combining operator ψ returns the function

$$h(x) = \hat{u}_{\boldsymbol{i}}(x) - \frac{1}{2}(w_{\mathbf{S}}(x)+1).$$

As in the proof of Theorem 1.1, ψ is a one-bit one-way combining operator. The next lemma completes the proof of the theorem. □

Lemma 4.13. *Fix $i \in [md]$ and $S \subseteq \{i, \ldots, dm\}$, and let $\boldsymbol{i}$ and $\mathbf{S}$ be their respective d-dimensional vector representations.*

1. *If $i \notin S$, then the function h is Lipschitz;*
2. *If $i \in S$, then the function h is $\frac{1}{8}$-far from Lipschitz.*

Proof. Consider the restriction of h to a line in dimension $j \in [d]$, i.e., consider the single-variate function $\bar{h}(x_j) = h(\bar{x}_1, \ldots, \bar{x}_{j-1}, x_j, \bar{x}_{j+1}, \ldots, \bar{x}_d)$, where the values $\bar{x}_k \in [2^m]$ are fixed for all $k \in [d] \setminus \{j\}$. Then

$$\begin{aligned} \bar{h}(x_j) &= \sum_{k\neq j} u_{\boldsymbol{i}_k}(\bar{x}_k) + u_{\boldsymbol{i}_j}(x_j) - \frac{1}{2}\Big(w_{\mathbf{S}_j}(x_j)\cdot\prod_{k\neq j} w_{\mathbf{S}_k}(\bar{x}_k) + 1\Big) \\ &= u_{\boldsymbol{i}_j}(x_j) - \frac{1}{2}(\pm w_{\mathbf{S}_j}(x_j)+1) + c, \end{aligned} \tag{2}$$

where $\pm$ means "either + or -" and c is a constant independent of x_j.

Let $j^* = \lceil i/m \rceil$. If $j < j^*$ then $\mathbf{S}_j = \emptyset$, $\boldsymbol{i}_j = m$ and $\bar{h} = u_{\boldsymbol{i}_j} - \frac{1}{2}(\pm 1 + 1) + c$. Since every up-down staircase function u_i is Lipschitz, and since a Lipschitz function plus a constant function is Lipschitz, the resulting function $\bar{h}$ is Lipschitz. If $j > j^*$ then $\boldsymbol{i}_j = 0$, so $\bar{h}(x_j) = 1 - \frac{1}{2}(\pm w_{\mathbf{S}_j}(x_j)+1) + c$,, i.e., $\bar{h}$ is again a Lipschitz function because it is the sum of a Lipschitz function and a constant function.

Finally, if $j = j^*$ then $\boldsymbol{i}_j = i - (j-1)m$. In this case, $i \in S$ iff $\boldsymbol{i}_j \in \mathbf{S}_j$. If $\boldsymbol{i}_j \notin \mathbf{S}_j$ then, by (2) and Lemma 3.14, $\bar{h}$ is Lipschitz. Since all line restrictions of h are Lipschitz, the overall function h is Lipschitz. Now suppose $\boldsymbol{i}_j \in \mathbf{S}_j$. Consider the product $\prod_{k\neq j} w_{\mathbf{S}_k}(\bar{x}_k)$ that determines whether the expression $\pm$ in (1) is actually a plus or a minus. By Corollary 4.6, this product evaluates to 1 for at least half of the line restrictions $\bar{h}(x_j)$ of h in dimension j. For those restrictions, $\bar{h}(x_j) = u_{\boldsymbol{i}_j} + \frac{1}{2}(w_{\mathbf{S}_j}+1)(x_j) + c$ and, since $\boldsymbol{i}_j \in \mathbf{S}_j$, Lemma 3.14 implies that $\bar{h}$ is $\frac{1}{4}$-far from Lipschitz. Thus, at least half of the line restrictions of h in dimension j are 1/4-far from Lipschitz. Since the domains of the line restrictions of h in dimension j partition the domain of h, the overall function h is $\frac{1}{8}$-far from Lipschitz. □

Acknowledgment

The authors would like to thank Madhav Jha for his participation in the project. He declined to be a co-author, but we would like to acknowledge his contributions to the results presented here. We also thank Joshua Brody for insightful conversations about communication complexity.

References

1. Ailon, N., Chazelle, B.: Information theory in property testing and monotonicity testing in higher dimension. Inf. Comput. 204(11), 1704–1717 (2006)
2. Awasthi, P., Jha, M., Molinaro, M., Raskhodnikova, S.: Testing Lipschitz functions on hypergrid domains. In: APPROX-RANDOM. pp. 387–398 (2012)
3. Blais, E., Brody, J., Matulef, K.: Property testing lower bounds via communication complexity. Computational Complexity 21(2), 311–358 (2012)
4. Brody, J., Matulef, K., Wu, C.: Lower bounds for testing computability by small width OBDDs. In: Ogihara, M., Tarui, J. (eds.) Theory and Applications of Models of Computation, LNCS, vol. 6648, pp. 320–331. Springer (2011)
5. Chakrabarty, D., Seshadhri, C.: A $o(n)$ monotonicity tester for boolean functions over the hypercube. In: STOC (2013)
6. Chakrabarty, D., Seshadhri, C.: Optimal bounds for monotonicity and Lipschitz testing over hypercubes and hypergrids. In: STOC (2013)
7. Dixit, K., Jha, M., Raskhodnikova, S., Thakurta, A.: Testing the Lipschitz property over product distributions with applications to data privacy. In: TCC. pp. 418–436 (2013)
8. Dodis, Y., Goldreich, O., Lehman, E., Raskhodnikova, S., Ron, D., Samorodnitsky, A.: Improved testing algorithms for monotonicity. In: RANDOM. pp. 97–108 (1999)
9. Ergun, F., Kannan, S., Kumar, S.R., Rubinfeld, R., Viswanathan, M.: Spot-checkers. J. Comput. Syst. Sci. 60(3), 717–751 (2000)
10. Fischer, E.: On the strength of comparisons in property testing. Information and Computation 189(1), 107–116 (2004)
11. Fischer, E., Lehman, E., Newman, I., Raskhodnikova, S., Rubinfeld, R., Samorodnitsky, A.: Monotonicity testing over general poset domains. In: STOC. pp. 474–483 (2002)
12. Goldreich, O. (ed.): Property Testing: Current Research and Surveys, Lecture Notes in Computer Science, vol. 6390. Springer (2010)
13. Goldreich, O., Goldwasser, S., Lehman, E., Ron, D., Samorodnitsky, A.: Testing monotonicity. Combinatorica 20(3), 301–337 (2000)
14. Goldreich, O., Goldwasser, S., Ron, D.: Property testing and its connection to learning and approximation. J. ACM 45(4), 653–750 (1998)
15. Jha, M., Raskhodnikova, S.: Testing and reconstruction of Lipschitz functions with applications to data privacy. In: Ostrovsky, R. (ed.) FOCS. pp. 433–442. IEEE (2011), full version available at http://eccc.hpi-web.de/report/2011/057/.
16. Miltersen, P.B., Nisan, N., Safra, S., Wigderson, A.: On data structures and asymmetric communication complexity. J. Comput. Syst. Sci. 57(1), 37–49 (1998)
17. Parnas, M., Ron, D., Rubinfeld, R.: On testing convexity and submodularity. SIAM J. Comput. 32(5), 1158–1184 (2003)
18. Rademacher, L., Vempala, S.: Testing geometric convexity. In: FSTTCS. pp. 469–480 (2004)
19. Raskhodnikova, S.: Approximate testing of visual properties. In: RANDOM-APPROX. pp. 370–381 (2003)
20. Raskhodnikova, S., Yaroslavtsev, G.: Learning pseudo-Boolean k-DNF and submodular functions. In: SODA. pp. 1356–1368 (2013)
21. Ron, D.: Algorithmic and analysis techniques in property testing. Foundations and Trends in Theoretical Computer Science 5(2), 73–205 (2009)
22. Rubinfeld, R., Shapira, A.: Sublinear time algorithms. Electronic Colloquium on Computational Complexity (ECCC) 18, 13 (2011)
23. Rubinfeld, R., Sudan, M.: Robust characterization of polynomials with applications to program testing. SIAM J. Comput. 25(2), 252–271 (1996)
24. Seshadhri, C., Vondrák, J.: Is submodularity testable? In: ICS. pp. 195–210 (2011)

ECCC ISSN 1433-8092
http://eccc.hpi-web.de