Testing and Learning Convex Sets in the Ternary Hypercube

Hadley Black
University of California, Los Angeles, CA, USA

Eric Blais
University of Waterloo, Canada

Nathaniel Harms
EPFL, Lausanne, Switzerland

Abstract

We study the problems of testing and learning high-dimensional discrete convex sets. The simplest high-dimensional discrete domain where convexity is a non-trivial property is the *ternary hypercube*, $\{-1, 0, 1\}^n$. The goal of this work is to understand structural combinatorial properties of convex sets in this domain and to determine the complexity of the testing and learning problems. We obtain the following results.

Structural: We prove nearly tight bounds on the edge boundary of convex sets in $\{0, \pm 1\}^n$, showing that the maximum edge boundary of a convex set is $\widetilde{\Theta}(n^{3/4}) \cdot 3^n$, or equivalently that every convex set has influence $\widetilde{O}(n^{3/4})$ and a convex set exists with influence $\Omega(n^{3/4})$.

Learning and sample-based testing: We prove upper and lower bounds of $3^{\widetilde{O}(n^{3/4})}$ and $3^{\Omega(\sqrt{n})}$ for the task of learning convex sets under the uniform distribution from random examples. The analysis of the learning algorithm relies on our upper bound on the influence. Both the upper and lower bound also hold for the problem of *sample-based testing* with two-sided error. For sample-based testing with one-sided error we show that the sample-complexity is $3^{\Theta(n)}$.

Testing with queries: We prove nearly matching upper and lower bounds of $3^{\widetilde{\Theta}(\sqrt{n})}$ for one-sided error testing of convex sets with *non-adaptive queries*.

2012 ACM Subject Classification Theory of computation → Streaming, sublinear and near linear time algorithms; Theory of computation → Randomness, geometry and discrete structures; Theory of computation → Computational geometry

Keywords and phrases Property testing, learning theory, convex sets, testing convexity, fluctuation

Digital Object Identifier 10.4230/LIPIcs.ITCS.2024.15

Related Version *Full Version*: `https://arxiv.org/abs/2305.03194`

Funding *Hadley Black*: Supported by NSF award AF:Small 2007682, NSF Award: Collaborative Research Encore 2217033.
Eric Blais: Supported by an Ontario Early Researcher Award and an NSERC Discovery Grant.
Nathaniel Harms: Some of this work was done while the author was a student at the University of Waterloo. Partly supported by an NSERC Graduate Scholarship, an NSERC Postdoctoral Fellowship, and the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract number MB22.00026.

1 Introduction

A subset $S \subseteq [m]^n$ of the hypergrid is *discrete convex* if it is the intersection of a convex set $C \subseteq \mathbb{R}^n$ with the grid, $S = C \cap [m]^n$, or equivalently if $S = [m]^n \cap \mathsf{Conv}(S)$ where $\mathsf{Conv}(S)$ is the convex hull of S. Discrete convex sets may not even be *connected* (see Figure 1), which, along with some of their other unpleasant features, makes them difficult to handle algorithmically and analytically, the most famous example being the difference between linear programming and integer linear programming.

We are interested in testing and learning discrete convex sets. A learning algorithm should output an approximation of an unknown convex set S by using membership queries to S, while a testing algorithm should decide whether an unknown set S is either convex or ϵ*-far* from convex, meaning that $\mathsf{dist}(S,T) > \epsilon$ for all convex sets T, where $\mathsf{dist}(S,T)$ is the measure of the symmetric difference.

Convexity is particularly interesting for property testing because it can be defined by a local condition: a set $S \subseteq \mathbb{R}^n$ is convex if and only if for every 3 colinear points x, y, z, if $x, z \in S$ then $y \in S$. This means that, to certify the non-convexity of a (continuous) set, it suffices to provide 3 colinear points that violate this condition. Speaking informally, property testing results, especially testing with *one-sided error*, are statements about the difficulty of finding such a certificate of non-membership to the property, when the object S is ϵ-far from satisfying the property. But, the fact that convexity is *defined* by a local condition does not make it easy to find violations of the condition when a set is far from convex. This is particularly evident for *discrete* convex sets where, unlike continuous sets, there may not be *any* lines which witness non-convexity, and one must instead look for up to $n+1$ points that violate Carathéodory's theorem.

We are aware of no non-trivial algorithms for testing or learning discrete convex sets in high dimensional grids $[m]^d$ when m is small. Prior works on testing and learning convex sets include:

1. The analysis of convexity testers, such as the *line tester* and more general *convex hull testers*, which are designed to simply "spot-check" for violations of the local conditions that define convexity in $\mathbb{R}^n$ [23, 7]. These works show that these spot-checkers are not very efficient, requiring $2^{\Omega(n)}$ queries to detect sets that are $\Omega(1)$-far from convex.
2. Testing or learning convex sets in two dimensions, including the continuous square $[0,1]^2$ [25, 4] or the discrete grid $[m]^2$ [24, 5, 6].
3. Testing convexity in high dimensions with *samples*, either in the continuous setting [10, 15] or discrete setting [15], and learning convex sets from random examples of the set [22] or from Gaussian samples [17].

When $m \gg \mathrm{poly}(d)$, a "downsampling" or "gridding" approach can reduce to the case $m = \mathrm{poly}(d)$ [10, 15], but once m is small the only known algorithm for testing or learning is brute-force. So let us see what happens when we make m as small as possible. When $m = 2$, testing and learning convex sets in $[m]^n \equiv \{0,1\}^n$ is trivial, because every subset of $\{0,1\}^d$ is convex and therefore testing is as easy as possible (the tester may simply accept on every input) and learning is as hard as possible (requiring $\Omega(2^n)$ queries).

The story changes significantly when $m = 3$, so that $[m]^n$ is equivalent to the ternary hypercube $\{0,\pm1\}^n$, where the difficulties of handling high-dimensional discrete convex sets suddenly become evident. Although this is the simplest domain where where high-dimensional discrete convex sets are non-trivial, little is known about the structure of discrete convex sets on the ternary hypercube that would help in designing testing and learning algorithms. In this paper we will give the first results towards understanding testing and learning discrete sets in high dimensions by focusing on the ternary hypercube.

1.1 Results

For two sets $S, T \subseteq \{0,\pm1\}^n$, we define $\mathsf{dist}(S,T) := \frac{|S \Delta T|}{3^n}$, where $S \Delta T$ denotes the symmetric difference. A set $S \subseteq \{0,\pm1\}^n$ is ε-far from convex if for every (discrete) convex set $T \subseteq \{0,\pm1\}^n$, $\mathsf{dist}(S,T) \geq \varepsilon$. Given $\varepsilon > 0$, a *convexity tester* is a randomized algorithm which is given membership oracle access to an input $S \subseteq \{0,\pm1\}^n$ and must satisfy

1. If S is convex then the algorithm accepts with probability at least $2/3$.

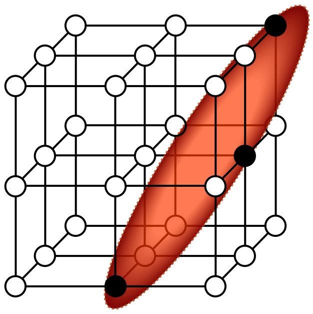

Figure 1 Example of a convex set in $\{0,\pm 1\}^3$. The black dots are the set and the convex red ellipsoid contains them. Note that the set may not be "connected" on the hypergrid.

2. If S is ε-far from convex then the algorithm rejects with probability at least $2/3$.

The tester is *one-sided* if it must accept convex sets S with probability 1 instead of $2/3$. A tester is *non-adaptive* if it chooses its set of queries before receiving the answers to any of the queries and it is *sample-based* if its queries are independently and uniformly random.

A learning algorithm is given membership oracle access to a convex set $S \subseteq \{0,\pm 1\}^n$ and must output (with probability at least $2/3$) a set $T \subseteq \{0,\pm 1\}^n$ with $\mathsf{dist}(S,T) < \varepsilon$; it is *proper* if its output T must be convex.

1.1.1 The Edge Boundary and Influence of Convex Sets

One of the most important things to know about a set is its *edge boundary*. The edge set of the ternary hypercube is defined as

$$E = \left\{ (x,y) \in (\{0,\pm 1\}^n)^2 \colon \sum_{i=1}^{n} |x_i - y_i| = 1 \right\}. \tag{1}$$

Observe that $|E| = 2n \cdot 3^{n-1}$. We will identify a set $S \subseteq \{0,\pm 1\}^n$ with its characteristic function and write $S(x) = 1$ if $x \in S$ and $S(x) = 0$ otherwise. An edge (x,y) is on the boundary of S if $S(x) \neq S(y)$. The influence of a set $S \subseteq \{0,\pm 1\}^n$ is its normalized boundary size:

$$\mathbb{I}(S) := \frac{1}{3^n} \cdot |\{(u,v) \in E \colon S(u) \neq S(v)\}| = \frac{2n}{3} \cdot \mathbb{P}_{(u,v)\sim E}[S(u) \neq S(v)]. \tag{2}$$

Before we state our results, consider some examples. Two important classes of convex sets in $\{0,\pm 1\}^n$ are halfspaces and balls, which often have minimal "boundary size" in various settings.

▶ **Example 1** (Halfspaces). A halfspace is a set $H = \{x \in \{0,\pm 1\}^n \ : \ \langle v, x\rangle < \tau\}$ where $v \in \mathbb{R}^n$ and $\tau \in \mathbb{R}$. To maximize the influence, we want τ to be small, say $\tau = 0$, and we want $v \approx \vec{1}$. The probability that a random edge (x,y) is on the boundary is at most the probability that a uniformly random $x \sim \{0,\pm 1\}^n$ satisfies $|\langle \vec{1}, x\rangle| \leq 1$, and it is not difficult to show that this is at most $O\left(\frac{1}{\sqrt{n}}\right)$, giving an estimate of $O(\sqrt{n})$ for the maximum influence of a halfspace.

▶ **Example 2** (Balls). A ball is a set $B_r = \{x \in \{0,\pm 1\}^n \ : \ \|x\|_2^2 < r\}$ where $r \in \mathbb{R}$ is the radius. The average (squared) norm $\mathbb{E}\left[\|x\|_2^2\right]$ for $x \sim \{0,\pm 1\}^n$ is the same as the expected number of nonzero coordinates of x, which is $\frac{2}{3}n$, so to maximize the edge boundary we think of $r \approx \frac{2}{3}n$. Similar to above, the probability that $x \sim \{0, \pm 1\}^n$ is close enough to this threshold to find a boundary edge is $O\left(\frac{1}{\sqrt{n}}\right)$, again giving an estimate of $O(\sqrt{n})$ for the maximum influence.

Our first result shows that there are convex sets with significantly larger influence, which can be obtained by taking S to be the intersection of roughly $3^{\Theta(\sqrt{n})}$ random halfspaces with thresholds $\tau = \Theta(n^{3/4})$; we think of these sets as interpolating between the halfspaces and the ball. Our construction is inspired by [16], who showed bounds on the influence of intersections random halfspaces on the hypercube $\{0, 1\}^n$, and we note that similar constructions also achieve maximal surface area under the Gaussian distribution on $\mathbb{R}^n$ [20].

▶ **Theorem 3.** *There exists a convex set $S \subseteq \{0, \pm 1\}^n$ with influence $\mathbb{I}(S) = \Omega(n^{3/4})$.*

Our main result on the influence of convex sets is that this construction is essentially optimal: we show a matching upper bound (up to log factors) for any convex set in $\{0, \pm 1\}^n$. Due to the discrete nature of the domain, our proof of this theorem is significantly different from the previous techniques that have been used to bound the surface area of convex sets in continuous domains.

▶ **Theorem 4.** *If $S \subseteq \{0, \pm 1\}^n$ is convex, then $\mathbb{I}(S) = O(n^{3/4} \log^{1/4} n)$.*

1.1.2 Sample-Based Learning and Testing

As an application of our bounds on the influence, we show using standard Fourier analysis that any set $S \subseteq \{0, \pm 1\}^n$ can be approximated with error ε by a polynomial of degree $\mathbb{I}(S)/\varepsilon$. Using Theorem 4 and the "Low-Degree Algorithm" of Linial, Mansour, and Nisan [18] then gives us the following upper bound for learning.

▶ **Theorem 5.** *There is a uniform-distribution learning algorithm for convex sets in $\{0, \pm 1\}^n$ which achieves error at most ε with time and sample complexity $3^{\widetilde{O}(n^{3/4}/\varepsilon)}$. The $\widetilde{O}(\cdot)$ hides a factor of $\log^{1/4} n$.*

A corollary of Theorem 5 is that the same upper bound on the sample complexity holds for sample-based testing, due to the testing-by-learning reduction (which is slightly non-standard because the learner is not *proper*, see the full version of our paper for details).

▶ **Corollary 6.** *There is a sample-based convexity tester for sets in $\{0, \pm 1\}^n$ with sample complexity $3^{\widetilde{O}(n^{3/4}/\varepsilon)}$ where the $\widetilde{O}(\cdot)$ hides a factor of $\log^{1/4} n$.*

To complement our upper bounds, we prove also a lower bound for sample-based testing. Here we remark that one of our motivations for studying convex sets in $\{0, \pm 1\}^n$ is their similarity (in an informal sense) to *monotone* functions on $\{0, 1\}^n$; an analogy between monotone functions on $\{0, 1\}^n$ and convex sets in Gaussian space was proposed in [12] and we are interested in this analogy for *discrete* convex sets. Our lower bound for sample-based testing discrete convex sets uses a version of Talagrand's random DNFs, which were used previously to prove lower bounds for testing monotonicity on $\{0, 1\}^n$ [2, 11].

▶ **Theorem 7.** *For sufficiently small constant $\varepsilon > 0$, every sample-based convexity tester for sets in $\{0, \pm 1\}^n$ has sample complexity $3^{\Omega(\sqrt{n})}$.*

Again, the testing-by-learning reduction implies that this lower bound also holds for learning.

▶ **Corollary 8.** *For sufficiently small constant $\varepsilon > 0$, sample-based learning convex sets in $\{0, \pm 1\}^n$ requires at least $3^{\Omega(\sqrt{n})}$ samples.*

1.1.3 Non-Adaptive One-Sided Testing

A convexity tester with one-sided error is one that finds a *witness* of non-convexity with probability at least 2/3 when the tested set is ε-far from convex. A convexity tester is *non-adaptive* if it must choose its set of membership queries before receiving any of the query results. Bounds on non-adaptive one-sided error testing therefore have a natural combinatorial interpretation as bounds on the likelihood of blindly finding a witness of non-convexity in a random substructure of the domain.

Our first result shows that there is a non-adaptive one-sided error tester with sub-exponential query complexity $3^{o(n)}$. In contrast, a similar bound for the Gaussian setting is not yet known to exist.

▶ **Theorem 9.** *For every $\varepsilon > 0$, there is a non-adaptive convexity tester with one-sided error for sets in $\{0, \pm 1\}^n$ that has query complexity $3^{\widetilde{O}\left(\sqrt{n \ln 1/\varepsilon}\right)}$ where the $\widetilde{O}(\cdot)$ notation is hiding an extra $\ln n$ term.*

Next, we show that Theorem 9 is essentially tight, in that the exponential dependence on $\sqrt{n}$ in its bound is unavoidable.

▶ **Theorem 10.** *For sufficiently small constant $\varepsilon > 0$, every non-adaptive convexity tester with one-sided error for sets in $\{0, \pm 1\}^n$ has query complexity at least $3^{\Omega(\sqrt{n})}$.*

Our Theorem 7 above showed that $3^{\Omega(\sqrt{n})}$ is required for *sample-based* testing. For one-sided error testers, we can improve this lower bound to show that non-adaptive testers are significantly more powerful than sample-based testers for one-sided testing.

▶ **Theorem 11.** *For sufficiently small constant $\varepsilon > 0$, sample-based convexity testing in $\{0, \pm 1\}^n$ with one-sided error requires $3^{\Theta(n)}$ samples.*

This theorem also includes a matching upper bound. The upper bound in Theorem 11 is trivial because a coupon-collector argument shows that one can learn any set $S \subseteq \{0, \pm 1\}^n$ exactly using $O(n3^n)$ samples. A slightly improved bound of $O\left(3^n \cdot \frac{1}{\varepsilon}\log(1/\varepsilon)\right)$ also holds by a general upper bound on one-sided error testing via the VC dimension [8].

1.2 Techniques

The discrete nature of the ternary hypercube, in contrast to the continuity of the domains $\mathbb{R}^n$ or $[0, 1]^n$, provides a new angle in the study of convexity which leads to the development of a new set of *combinatorial* techniques and tools. In this section we give a brief overview of the techniques we use to prove each of our theorems.

1.2.1 The Edge Boundary and Influence of Convex Sets

Influence Upper Bound

Our proof of Theorem 4, which gives an upper bound on the edge boundary of a convex set, is accomplished by relating the number of boundary edges to the expected number of sign-changes of one-dimensional random processes. This is done by constructing a distribution $\mathcal{D}$ over the edge-set E of the ternary hypercube, such that (a) $\mathcal{D}$ is "close" to the uniform distribution over E and (b) the probability that a random edge drawn from $\mathcal{D}$ is influential for our convex set $S \subseteq \{0, \pm 1\}^n$ is equal to the expected number of sign-changes of a certain random process. This process is defined by considering a random walk $\boldsymbol{X}^{(0)}, \ldots, \boldsymbol{X}^{(m)}$ of length $m \approx n^{1/2}$ where $\boldsymbol{X}^{(0)}$ is a random point from the middle layers of $\{0, \pm 1\}^n$ and each $\boldsymbol{X}^{(s)}$ is obtained by flipping a random 0-valued bit of $\boldsymbol{X}^{(s-1)}$ to a uniform random $\{\pm 1\}$-value; the process finally draws $\boldsymbol{s} \sim [m]$ uniformly at random and outputs the edge $(\boldsymbol{X}^{(\boldsymbol{s}-1)}, \boldsymbol{X}^{(\boldsymbol{s})})$.

The crux of the argument is to bound the expected number of times this random walk enters and leaves the set S. Since S is convex, it can be written as an intersection of halfspaces $S = H_1 \cap H_2 \cap \cdots \cap H_k$ of the form $H_i = \{x \in \{0, \pm 1\}^n \colon \langle x, v^{(i)} \rangle < \tau_i\}$ where $v^{(i)} \in \mathbb{R}^n$ and $\tau_i \in \mathbb{R}$. For each halfspace H_i, we define a corresponding one-dimensional random walk $\boldsymbol{W}_i(s) = \langle \boldsymbol{X}^{(s)}, v^{(i)} \rangle - \tau_i$ and observe that the original random walk crosses the boundary of H_i at step s if and only if $\boldsymbol{W}_i$ *changes sign* at step s. Then the number of times the walk $\boldsymbol{X}^{(0)}, \boldsymbol{X}^{(1)}, \ldots$ crosses the boundary of $S = \bigcap_i H_i$ is the number of times the *maximum* of the processes $\boldsymbol{M} = \max_i \boldsymbol{W}_i$ changes sign. Therefore, our goal is to bound the expected number of sign-changes for $\boldsymbol{M}$, which we accomplish by using Sparre Andersen's fluctuation theorem [26] (as stated in [3]) to relate this quantity to the number of sign-changes of a uniform random walk.

High-Influence Set Construction

Our proof of Theorem 3 is inspired by the proof of [16, Theorem 2] which constructs a set in the Boolean hypercube $\{\pm 1\}^n$ with influence $\Omega(\sqrt{n \log k})$ by considering an intersection of k random halfspaces each of which is at distance $\approx \sqrt{n \log k}$ from the origin. In particular, when $k \approx 2^{\sqrt{n}}$ the construction has influence $\approx n^{3/4}$ and when $k \approx 2^n$ the set has influence $\approx n$. On the ternary hypercube $\{0, \pm 1\}^n$, the behaviour is different: here, halfspaces exhibit a "density increment" behaviour as their threshold moves away from the origin, which prevents the influence from increasing as k grows past $2^{\sqrt{n}}$, when $\Omega(\sqrt{n \log k})$ matches our upper bound of $\widetilde{O}(n^{3/4})$.

We can summarize this "density increment" phenomenon as follows. Most of the edges of $\{0, \pm 1\}^n$ occur in the middle layer $\{x \in \{0, \pm 1\}^n : \|x\|_1 = \frac{2}{3}n \pm O(\sqrt{n})\} = \bigcup_{\ell=-O(\sqrt{n})}^{O(\sqrt{n})} \{x : \|x\|_1 = \frac{2}{3}n + \ell\}$. A convex set is an intersection of halfspaces, but for convenience we consider its complement which is a union of halfspaces, and has the same influence. Consider the "density" or measure of the halfspace with normal vector $\vec{1}$ at distance τ from the origin on the points $\{x : \|x\|_1 = \frac{2n}{3} + \ell\}$:

$$\rho(\ell, \tau) := \mathbb{P}_{x \in \{0, \pm 1\}^n \colon \|x\|_1 = \frac{2n}{3} + \ell} \left[\sum_i x_i > \tau \right] .$$

Suppose that there is a fixed value ρ such that $\rho(\ell, \tau) \approx \rho$ up to constant factors for *all* $\ell = \pm O(\sqrt{n})$ simultaneously. Then we can take $k \approx \frac{1}{\rho}$ random halfspaces with threshold τ and combine their boundary edges, since they will be essentially disjoint on the whole middle layer, and it is not hard to show that the influence of the resulting union is roughly τ. It happens that the condition of $\rho(\ell, \tau)$ being approximately equal for all values $\ell = \pm O(\sqrt{n})$ holds for τ up to $\tau \approx n^{3/4}$ but for $\tau \gg n^{3/4}$ the intersection of the halfspace with the set $\{x : \|x\|_1 = \frac{2n}{3} + \ell\}$ grows extremely fast with ℓ making $\rho(-\sqrt{n}, \tau) \ll \rho(\sqrt{n}, \tau)$, and the intersection of halfspaces with threshold τ quickly approaches the ball with influence $O(\sqrt{n})$ (see Example 2).

1.2.2 Sample-Based Learning and Testing

Learning Upper Bound

Our proof of Theorem 5 follows by combining our upper bound on the influence from Theorem 4 with the Low-Degree Algorithm of Linial, Mansour, and Nisan [18]. In particular, using Fourier analysis over $\{0, \pm 1\}^n$ in combination with Theorem 4 we can show that for convex sets, a $(1-\varepsilon)$-fraction of the Fourier mass is on the coefficients with degree at most $\widetilde{O}(n^{3/4})/\varepsilon$. Then we may use the Low-Degree Algorithm for learning the convex sets; see the full version for details. Since the ternary hypercube is a non-standard domain, we state the necessary Fourier analysis for functions over $\{0,\pm1\}^n$ in the full version of our paper, which follows [21, Chapter 8]. One technical difference between Fourier analysis over the Boolean and ternary hypercubes is that the standard Fourier basis over $\{\pm1\}^n$ is given by the parity functions which are bounded in $[0,1]$, whereas any Fourier basis over $\{0,\pm1\}^n$ will have functions taking value $2^{O(n)}$ on some elements $x \in \{0,\pm1\}^n$. Nevertheless, with some care, we show that the Low-Degree Algorithm still works.

Sample-Based Testing Lower Bound

Our proof of Theorem 7 uses a family of functions known as *Talagrand's random DNFs* adapted to the ternary hypercube. As we mentioned, this family of functions has been used to prove lower bounds for monotonicity testing [2, 11]. Our adapted version is described as follows. Each "term" of the DNF is chosen to be a random point $t \in \{0,\pm1\}^n$ with $\|t\|_1 = \sqrt{n}$. We then say that a point $x \in \{0,\pm1\}^n$ "satisfies" t if $x_i = t_i$ for all $i \in [n]$ where $t_i \in \{\pm1\}$. After choosing N random terms $t^{(1)},\ldots,t^{(N)}$ we define the disjoint regions of $\{0,\pm1\}^n$ given by $U_1,\ldots,U_N$ where U_i is the set of points $x \in \{0,\pm1\}^n$ with $\|x\|_1 \in [2n/3 \pm \sqrt{n}]$ which satisfy a unique term. Choosing $N = 3^{\sqrt{n}}$ results in $\bigcup_{i=1}^N U_i$ covering a constant fraction of the domain. We then define two distributions $\mathcal{D}_{\mathsf{yes}}$ and $\mathcal{D}_{\mathsf{no}}$ as follows. Recall that B_r is the radius-r ball in the ternary cube (Example 2) and let D denote the set of points $x \in \{0,\pm1\}^n$ with $\|x\|_1 \in [2n/3 \pm \sqrt{n}]$ that don't satisfy *any* term.

- $S \sim \mathcal{D}_{\mathsf{yes}}$ is drawn by setting $S = B_{\frac{2n}{3}-\sqrt{n}} \cup D \cup \left(\bigcup_{i\in T} U_i\right)$ where T includes each $i \in [N]$ independently with probability $1/2$. Such a set is always convex.
- $S \sim \mathcal{D}_{\mathsf{no}}$ is drawn by setting $S = B_{\frac{2n}{3}-\sqrt{n}} \cup D \cup C$ where C includes each $x \in \bigcup_{i=1}^N U_i$ independently with probability $1/2$. Informally, this set will be $\Omega(1)$-far from convex with constant probability since its intersection with the middle layers is random.

For both distributions, each point $x \in \bigcup_{i=1}^N U_i$ satisfies $\mathbb{P}_S[x \in S] = 1/2$ and if $x \in U_i$ and $y \in U_j$ where $i \neq j$, then the events $x \in S$ and $y \in S$ are independent. Thus, to distinguish $\mathcal{D}_{\mathsf{yes}}$ and $\mathcal{D}_{\mathsf{no}}$ one has to see at least two points from the same U_i and this gives our sample complexity lower bound.

1.2.3 Non-Adaptive One-Sided Testing

The proofs of Theorems 9–11 all rely on a partial order $\preceq$ defined on $\{0,\pm1\}^n$, which we call the *outward-oriented poset*, that has the origin 0^n as the minimum element and the corners of the cube $\{\pm1\}^n$ as the maximum elements. (See Section 2.1 for the formal definition of this poset and a discussion of its properties and history.) For any $y \in \{0,\pm1\}^n$, we define $\mathsf{Up}(y) := \{x \in \{0,\pm1\}^n : y \preceq x\}$ to represent the set of points above y in this poset.

Non-Adaptive One-Sided Upper Bound

An important property of the outward-oriented poset in the context of testing convexity is that any point y in the convex hull of a set of points $X \subseteq \{0,\pm1\}^n$ is also in the convex hull of $X \cap \mathsf{Up}(y)$. Conversely, if a set $S \subseteq \{0,\pm1\}^n$ is *not* convex, then there is a certificate of non-convexity of the form (X, y) where $y \notin S$ is in the convex hull of $X \subseteq S$, and $X \subseteq \mathsf{Up}(y)$. This property implies that a convexity tester can search for certificates of non-convexity by repeatedly choosing a random point y and querying all points in $\mathsf{Up}(y)$. A naïve implementation of this idea leads to a query complexity that is significantly larger than the bound in the theorem. However, the ternary hypercube satisfies a strong concentration of measure property: almost all of the points in the ternary hypercube have $\frac{2}{3}n \pm O(\sqrt{n})$ non-zero coordinates. As a result, we can refine the convexity tester to only query the points in $\mathsf{Up}(y)$ whose number of non-zero coordinates is at most $\frac{2}{3}n + O(\sqrt{n})$ to obtain the desired query complexity. The details of the proof of Theorem 9 are presented in the full version.

Non-Adaptive One-Sided Lower Bound

The lower bound in Theorem 10 is obtained by considering the class of *anti-slabs*, which are defined by choosing a vector $v \in \{0, \pm 1\}^n$ with $n/2$ non-zero coordinates and taking the set of points $\{x \in \{0, \pm 1\}^n : |\langle v, x\rangle| > \tau\}$. It is quite easy to find certificates of non-convexity for anti-slabs – the three points $-x$, 0^n, and x obtained by choosing x uniformly at random in the ternary hypercube forms such a certificate with reasonably large probability whenever τ is small enough. However, we can eliminate these certificates of non-convexity if we "truncate" the anti-slabs by including the set of points whose number of non-zero coordinates is below $\frac{2}{3}n - O(\sqrt{n})$, and excluding the points whose number of non-zero coordinates is above $\frac{2}{3}n + O(\sqrt{n})$. We show that any certificate of non-convexity for these truncated anti-slabs must have two points x, z with a large difference between $\langle v, x\rangle$ and $\langle v, z\rangle$, but on the other hand, any small set of queries has a low probability of including such a pair when v is chosen at random.

Sample-Based One-Sided Lower Bound

Finally, the proof of the lower bound Theorem 11 again uses the outward-oriented poset and the connection between convex hulls and the upwards sets $\mathsf{Up}(y)$ to show that any set of $3^{o(n)}$ samples is unlikely to draw any point y that is contained in the convex hull of the other sampled points and thus to have any possibility of identifying a certificate of non-convexity of any set.

1.3 Discussion and Open Problems

As far as we know, we are the first to study convex sets and their associated algorithmic problems on the ternary hypercube. Thus there are many possible questions one could ask. In this section we discuss a few such questions which we find most interesting.

Learning and sample-based testing

The most obvious question which our work leaves open is that of determining the true sample complexity of learning and sample-based testing of convex sets in the ternary hypercube, where our results leave a gap of $3^{\Omega(\sqrt{n})}$ vs. $3^{\widetilde{O}(n^{3/4})}$. By Theorem 3, our upper bound of $\widetilde{O}(n^{3/4})$ on the influence of convex sets is tight up to a factor of $\log^{1/4} n$, and therefore to improve our learning upper bound would require another method.

▶ **Question 12.** *Can we close the gap of* $3^{\Omega(\sqrt{n})}$ *vs.* $3^{\widetilde{O}(n^{3/4})}$ *for learning convex sets and for sample-based convexity testing in* $\{0, \pm 1\}^n$*?*

Testing with two-sided error

Our results for testing with queries apply only to case of one-sided error. Earlier work on testing convex sets under the Gaussian distribution on $\mathbb{R}^n$ with samples showed that, in that setting, two-sided error was more efficient than one-sided [10].

▶ **Question 13.** *Is there a two-sided error non-adaptive tester for domain* $\{0, \pm 1\}^n$ *with better query complexity than our one-sided error tester?*

Our lower bound technique does not suffice for two-sided error. This is because the class of anti-slabs, which we proved are hard to distinguish from convex sets using a one-sided tester, can be distinguished from convex sets with *two-sided* error using only $O(n)$ samples. To do so, one may use the standard testing-by-learning reduction of [14], together with an $O(n)$ bound on the VC dimension of the anti-slabs (which are essentially the union of two halfspaces).

Testing convexity in other domains

Our results show that queries can be more effective than samples for testing certain discrete convex sets in some high dimensional domains. Is it true for all discrete high-dimensional domains?

▶ **Question 14.** *What are the sample and query complexities for testing discrete convexity over the general hypergrids* $[m]^n$*?*

Note that our techniques do not immediately generalize to larger hypergrids, so answering the last question even for the hypergrid $\{0, \pm 1, \pm 2\}^n$ requires some new ideas.

It would also be interesting to see if the gap between sample and query complexity also holds for continuous sets.

▶ **Question 15.** *Can queries improve upon the bounds of [10, 15] for testing convex sets with samples in* $\mathbb{R}^n$ *under the Gaussian distribution?*

It is not clear if there is a formal connection between testing convex sets on the domain $\{0, \pm 1\}^n$ and on the domain $\mathbb{R}^n$ under the standard Gaussian distribution. One might expect a connection here because the uniform distribution on $\{0, \pm 1\}^n$ acts similarly to the Gaussian in certain ways when $n \to \infty$. But we do not see how to construct direct reductions between these two settings for the problem of convexity testing. Also, there is an intriguing analogy between monotone subsets of $\{\pm 1\}^n$ and convex subsets of $\mathbb{R}^n$ in the Gaussian space [12]. How do convex subsets of $\{0, \pm 1\}^n$ fit into this analogy?

1.4 Organization.

We introduce some preliminaries, including some important structure and concentration of measure results for the ternary hypercube, in Section 2. Section 3 presents the proof of our upper bound on the total influence of convex sets. Due to space constraints, the remainder of our proofs can be found in the full version of the paper.

2 Convexity on the Ternary Hypercube

The main object of study in this paper is the *ternary hypercube*, an analogue of the Boolean hypercube over the ternary set $\{0, \pm 1\}^n$. This set can be viewed as a discrete subset of $\mathbb{R}^n$, as a (hyper)grid graph in which two points $x, y \in \{0, \pm 1\}^n$ are connected by an edge if and only if $\sum_{i=1}^n |x_i - y_i| = 1$, and as a poset that we will describe in more detail in the subsection below.

The study of the ternary hypercube and more general grid graphs goes back at least to Bollobás and Leader [9]. As a poset, its study goes back at least to Metropolis and Rota [19]. The ternary hypercube appears to have some particularly elegant structure that is not necessarily shared by larger hypergrids. We describe some of these fundamental properties in the following subsections.

2.1 The Outward-Oriented Poset

We define a partial order over $\{0, \pm 1\}^n$, which puts the origin 0^n as the minimum element and the corners $\{\pm 1\}^n$ as the maximum elements.

▶ **Definition 16** (Outward-Oriented Poset). *We denote by* $(\{0, \pm 1\}^n, \preceq)$ *the* n*-wise product of the partial order defined by* $0 \prec 1$ *and* $0 \prec -1$*. Equivalently, we write* $y \preceq x$ *when* $\forall i \in [n]\colon (y_i \neq 0 \implies x_i = y_i)$*.*

The outward-oriented poset can easily be extended to a lattice (by adding a global maximum point), though since we do not need this extension we do not pursue it here. The outward-oriented poset appears naturally in many different contexts and, as a result, has received different names. For instance, it arises in the study of the faces of the Boolean hypercube [19], where it is sometimes called the "cubic lattice", and in the study of partial Boolean functions (see, e.g., [13]). We use the name "outward-oriented poset" to emphasize the fact that this poset is distinct from the partial order inherited from $\mathbb{R}^n$.

▶ **Definition 17** (Upper Shadow). *For any point* $y \in \{0, \pm 1\}^n$*, the* upper shadow *of* y *is the set*

$$\mathsf{Up}(y) := \{x \in \{0, \pm 1\}^n : y \preceq x\} \ .$$

2.2 Convexity and Witnesses of Non-Convexity

Given a set of points $X \subseteq \{0, \pm 1\}^n$, we denote the convex hull of X by

$$\mathsf{Conv}(X) := \left\{ \sum_{x \in X} \lambda_x x \colon \sum_{x \in X} \lambda_x = 1 \text{ and } \lambda_x \geq 0, \forall x \in X \right\} .$$

▶ **Definition 18** (Discrete Convexity). *A set* $S \subseteq \{0, \pm 1\}$ *is* convex *if* $S = \mathsf{Conv}(S) \cap \{0, \pm 1\}^n$*.*

Let $\Delta(S, T)$ denote the cardinality of the symmetric difference between S and T. Given $S \subseteq \{0, \pm 1\}^n$, we define $\mathsf{dist}(S, \mathbf{convex})$ as the minimum, over all convex sets $T \subseteq \{0, \pm 1\}^n$, of $\Delta(S, T) \cdot 3^{-n}$. For brevity, we also sometimes use the notation $\varepsilon(S) := \mathsf{dist}(S, \mathbf{convex})$. If $\varepsilon(S) \geq \varepsilon$ for some $\varepsilon \in (0, 1)$, then we say that S is ε-far from convex.

▶ **Definition 19** (Violating Pairs). *Consider* $S \subseteq \{0, \pm 1\}^n$*. If* $X \subseteq S$ *and* $y \in \mathsf{Conv}(X) \cap \{0, \pm 1\}^n$*, but* $y \notin S$*, then we call* (X, y) *a* violating pair *for* S*. The pair is called* minimal *if* $y \notin \mathsf{Conv}(X')$ *for any strict subset* $X' \subset X$*.*

All of our results exploit the following key property of the outward-oriented poset. This fact captures the structure of $\{0, \pm 1\}^n$ which we use throughout the paper.

▶ **Fact 20.** *If a violating pair* (X, y) *is minimal, then* $X \subseteq \mathsf{Up}(y)$*.*

Proof. We have $y = \sum_{x \in X} \lambda_x x$ where $\sum_{x \in X} \lambda_x = 1$. Moreover, the minimality of (X, y) implies that $\lambda_x > 0$ for all $x \in X$. Now, let $i \in [n]$ be some coordinate where $y_i \neq 0$. We need to show that $x_i = y_i$ for all $x \in X$. Without loss of generality, suppose $y_i = 1$. Thus, we have $1 = \sum_{x \in X} \lambda_x x_i$. If $x_i < 1$ for some $x \in X$, then we would have $\sum_{x \in X} \lambda_x x_i < 1$, which is a contradiction. ◀

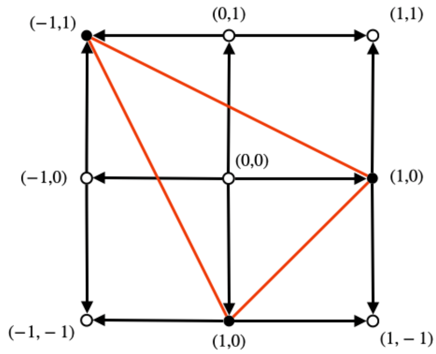

Figure 2 An illustration of $\{0, \pm 1\}^2$. Arrows indicate the direction of the partial order. The red triangle shows the convex hull of $X := \{(-1, 1), (1, 0), (0, 1)\}$, which contains the origin. I.e. $(X, (0, 0))$ is a minimal violating pair for X.

▶ **Fact 21.** *Let $S \subseteq \{0, \pm 1\}^n$. The following two statements are equivalent.*

- *S is not convex.*
- *There exists a minimal violating pair (X, y) for S.*

Proof. Suppose there exists a minimal violating pair (X, y) for S. Since $X \subseteq S$, we have $\mathsf{Conv}(X) \subseteq \mathsf{Conv}(S)$ and so $y \in \mathsf{Conv}(S)$. Thus, $y \notin S$ implies S is not convex. Now suppose S is not convex. Then there exists $y \in (\mathsf{Conv}(S) \cap \{0, \pm 1\}^n) \setminus S$. Let $X \subseteq S$ be a minimal set of points such that $y \in \mathsf{Conv}(X)$. The pair (X, y) is a minimal violating pair for S. ◀

▶ **Fact 22.** *Consider $S, Q \subseteq \{0, \pm 1\}^n$. If Q does not contain any $X \cup \{y\}$ such that (X, y) is a violating pair for S, then there exists a convex set S' such that $S' \cap Q = S \cap Q$.*

Proof. Let $S' = \mathsf{Conv}(S \cap Q)$ and consider an arbitrary $y \in Q$. We need to show that $y \in S$ if and only if $y \in S'$. Clearly, $y \in S$ implies $y \in S'$. Now suppose $y \in S'$ and note this implies $y \in \mathsf{Conv}(S \cap Q) \subseteq \mathsf{Conv}(S)$. Thus, if $y \notin S$, then $(S \cap Q, y)$ is a violating pair for S and this contradicts our assumption about Q. ◀

The following corollary is crucial for proving our lower bounds for testing convexity.

▶ **Corollary 23.** *Let T be a convexity tester for sets $S \subseteq \{0, \pm 1\}^n$ with 1-sided error. Suppose T rejects a set S after querying a set Q. Then Q contains some $X \cup \{y\}$ such that (X, y) is a minimal violating pair for S.*

2.3 Concentration of Mass in the Ternary Hypercube

For $x \in \{0, \pm 1\}^n$, observe that $\|x\|_1 = \|x\|_2^2$ is precisely the number of non-zero coordinates of x. Moreover, each coordinate of a uniformly random x is non-zero with probability $2/3$, and so $\mathop{\mathbb{E}}_{x \in \{0,\pm1\}^n}[\|x\|_1] = \frac{2n}{3}$. Standard concentration inequalities yield the following bound on the number of points $x \in \{0, \pm 1\}^n$ where $\|x\|_1$ is far from this expectation.

▶ **Fact 24.** *For every $\tau \geq 0$,*

$$\mathop{\mathbb{P}}_{x \in \{0,\pm1\}^n}\left[\left|\|x\|_1 - \frac{2n}{3}\right| > \tau\right] \leq 2\exp(-\tau^2/2n).$$

Proof. We have $\|x\|_1 = \sum_{i=1}^n X_i$ where $X_i = 1$ with probability $2/3$ and $X_i = 0$ with probability $1/3$. Thus, the bound follows immediately from Hoeffding's inequality. ◀

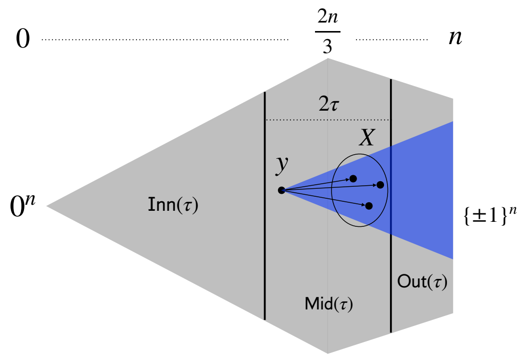

Figure 3 This figure shows a pictorial representation of $\{0, \pm 1\}^n$ as a poset. Any vertical slice represents the set of all points with some fixed number of non-zero coordinates, and this number is increasing from left to right. The left-most point is the origin and the right-most points are the vertices of the hypercube $\{\pm 1\}^n$. The outward-oriented poset goes from left to right. The shaded blue region emanating from y is the set $\mathsf{Up}(y)$ of points above y in the partial order. The set X represents some minimal set of points for which $y \in \mathsf{Conv}(X)$ and thus $y \prec x$ for all $x \in X$, by Fact 20.

Given $\tau \geq 0$, we use the following notation to denote the inner, middle, and outer layers of $\{0, \pm 1\}^n$ with respect to distance τ:

$$\begin{aligned} \mathsf{Inn}(\tau) &:= \left\{ x \colon \|x\|_1 - \frac{2n}{3} < -\tau \right\}, \\ \mathsf{Mid}(\tau) &:= \left\{ x \colon \left| \|x\|_1 - \frac{2n}{3} \right| \leq \tau \right\}, \\ \mathsf{Out}(\tau) &:= \left\{ x \colon \|x\|_1 - \frac{2n}{3} > \tau \right\}. \end{aligned} \tag{3}$$

3 The Influence of Convex Sets

In this section we prove that the maximum edge boundary of convex sets in $\{0, \pm 1\}^n$ is $\widetilde{\Theta}(n^{3/4}) \cdot 3^n$, or equivalently that the influence is $\widetilde{\Theta}(n^{3/4})$.

3.1 Upper Bound

We prove that convex sets in the ternary hypercube have influence $\widetilde{O}(n^{3/4})$. The main idea in the proof is to relate the influence of a convex set S to the number of sign-changes in the maximum of a set of one-dimensional random walks. The proof will consider a random walk $\boldsymbol{X}^{(0)}, \boldsymbol{X}^{(1)}, \ldots, \boldsymbol{X}^{(m)}$ starting from a random position in the middle layer of the ternary hypercube and moving randomly "outward" for $m = O\left(\sqrt{\frac{n}{\log n}}\right)$ steps, and count the number of influential edges crossed near the "middle layers" by relating them to one-dimensional random walks. We begin in Section 3.1.1 with definitions regarding the one-dimensional random walks that we require and then in Section 3.1.2 show how they relate to the number of influential edges of S; finally, in Section 3.1.3 we prove the necessary bound on the number of sign-changes of the one-dimensional random walks.

Notation. In this section it will be convenient to use bold letters like $\boldsymbol{X}$ for random variables, with the non-bold letter X being reserved for a fixed instantiation of $\boldsymbol{X}$.

3.1.1 One-Dimensional Random Walks and the Max-Walk

Let us define the types of one-dimensional random walks that will be necessary for our proof.

▶ **Definition 25** (Random Walks). *Let $x = (x_1, \ldots, x_m) \in \mathbb{R}^m$. Fix any permutation $\sigma : [m] \to [m]$ and sign vector $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_m) \in \{\pm 1\}^m$. For any $a \in \mathbb{R}$, we define the function $W_x^{+a}(t; \sigma, \varepsilon)$ for $t \in \{0\} \cup [m]$ as*

$$W_x^{+a}(t; \sigma, \varepsilon) := \begin{cases} a & \text{if } t = 0 \\ a + \sum_{i=1}^{t} \varepsilon_i x_{\sigma(i)} & \text{if } t > 0 \,. \end{cases}$$

The random walk $\boldsymbol{W}_x^{+a}$ is defined by choosing a uniformly random permutation $\boldsymbol{\sigma}$ and vector $\boldsymbol{\varepsilon} \sim \{\pm 1\}^m$ and setting $\boldsymbol{W}_x^{+a}(t) = W_x^{+a}(t; \boldsymbol{\sigma}, \boldsymbol{\varepsilon})$ for every t. If $a = 0$ we drop the superscript.

The main quantity of interest to us is the number of sign-changes of a random walk, defined as follows.

▶ **Definition 26** (Crossing Number). *Let $W : \{0\} \cup [m] \to \mathbb{R}$ be any sequence. We define the* crossing number *$C(W)$ as the number of sign-changes of W, defined as the number of times $t \in [m]$ such that either $W(t) \geq 0 > W(t-1)$ or $W(t) < 0 \leq W(t-1)$.*

An important feature of our random walks will be that they have the Distinct Subset-Sum (DSS).

▶ **Definition 27** (DSS Random Walk). *We say a sequence $x \in \mathbb{R}^m$ has the Distinct Subset-Sum (DSS) property if for every two disjoint subsets $A, B \subseteq [m]$, it holds that $\sum_{a \in A} x_a \neq \sum_{b \in B} x_b$. In particular, the random walk $\boldsymbol{W}_x$ satisfies*

$$\forall t \in [m] \,, \ \mathop{\mathbb{P}}_{\boldsymbol{\sigma}, \boldsymbol{\varepsilon}} \left[W_x(t; \boldsymbol{\sigma}, \boldsymbol{\varepsilon}) = 0 \right] = 0 \,.$$

Note that, if x has the DSS property, then so does any subsequence of x.

We will require an upper bound on the crossing number of *max-walks*, which are random walks defined as the maximum of a set of *constituent* walks of the type defined above.

▶ **Definition 28** (Max-Walk). *Let X be a set of sequences $x \in \mathbb{R}^m$, and let $a : X \to \mathbb{R}$. For a fixed permutation σ and vector $\varepsilon \in \{\pm 1\}^m$, define*

$$M_X^{+a}(t; \sigma, \varepsilon) := \max_{x \in X} W_x^{+a(x)}(t; \sigma, \varepsilon) \,,$$

and let the random walk $\boldsymbol{M}_X^{+a}$ be defined as

$$\boldsymbol{M}_X^{+a}(t) := M_X^{+a}(t; \boldsymbol{\sigma}, \boldsymbol{\varepsilon})$$

where $\boldsymbol{\sigma}, \boldsymbol{\varepsilon}$ are chosen uniformly at random.

The main fact about max-walks that we require is the following, which we prove in Section 3.1.3.

▶ **Lemma 29** (Max-Walk Crossing Number). *Let X be a set of sequences $x \in \mathbb{R}^m$, each having the DSS property, and let $a : X \to \mathbb{R}$. Then*

$$\mathbb{E}\left[C(\boldsymbol{M}_X^{+a}) \right] = O(\sqrt{m}) \,.$$

3.1.2 Upper Bound on the Number of Influential Edges of a Convex Set

We now prove the following upper bound on the influence of any convex set in the ternary hypercube, restated below for convenience.

▶ **Theorem 4.** *If $S \subseteq \{0, \pm 1\}^n$ is convex, then $\mathbb{I}(S) = O(n^{3/4} \log^{1/4} n)$.*

We require the following basic property of discrete convex sets.

▶ **Proposition 30.** *Let $S \subseteq \{0, \pm 1\}^n$ be any discrete convex set. Then there is a finite set of vectors $V \subseteq \mathbb{R}^n$ and thresholds $\tau : V \to \mathbb{R}$, where each $v \in V$ defines a halfspace $H_v := \{x \in \{0, \pm 1\}^n \; : \; \langle v, x\rangle < \tau(v)\}$, such that $S = \bigcap_{v \in V} H_v$. One may also assume that V satisfies the property that, for every $v \in V$ and every two disjoint subsets $A, B \subseteq [n]$, $\sum_{i \in A} v_i \neq \sum_{j \in B} v_j$.*

Proof. Since S is the intersection of its convex hull $\mathsf{Conv}(S)$ with $\{0, \pm 1\}^n$, it may be written as the intersection of $\{0, \pm 1\}^n$ with a finite set of halfspaces with normal vectors V and thresholds $\tau : V \to \mathbb{R}$, and one may assume that none of the points in $\{0, \pm 1\}^n$ lie on the hyperplane boundary of any of the halfspaces. Then there is some $\delta > 0$ such that the minimum distance between a hyperplane and a point of $\{0, \pm 1\}^n$ is at least $\delta \cdot n$. For each $v \in V$, apply independent random perturbations to each coordinate to obtain $v'_i = v_i + r_i$ where r_i is drawn from $[-\delta, \delta]$ uniformly at random. With probability 1, the resulting set $V' = \{v' \; : \; v \in V\}$ satisfies the required conditions. ◀

Proof of Theorem 4. Recall the definition of the edge-set E of the ternary cube from Equation (1) and the set $\mathsf{Mid}(\ell)$ from Equation (3). Given $\ell > 0$, let

$$E_\ell = \{(u, v) \in E \colon u, v \in \mathsf{Mid}(\ell)\}$$

denote the set of edges lying in the middle ℓ layers of $\{0, \pm 1\}^n$. We consider the following process which samples a random edge in $\{0, \pm 1\}^n$. Define $\ell := \sqrt{2n \log n}$ and $m := \sqrt{\frac{n}{\log n}}$. Let $\mathcal{D}$ denote the distribution over edges defined by the following procedure.

1. Sample $\boldsymbol{X}^{(0)} \sim \mathsf{Mid}(\ell)$.
2. Choose a random subset $\boldsymbol{T} \subseteq \{i \colon \boldsymbol{X}_i^{(0)} = 0\}$ with $|\boldsymbol{T}| = m$ of coordinates where $\boldsymbol{X}^{(0)}$ has a 0.
3. Let $\boldsymbol{\varepsilon} = (\boldsymbol{\varepsilon}_1, \ldots, \boldsymbol{\varepsilon}_m) \in \{\pm 1\}^m$ be independent Rademacher random variables and let $\boldsymbol{\sigma} \colon [m] \to \boldsymbol{T}$ be a random bijection.
4. For each $s \in [m]$, let $\boldsymbol{X}^{(s)} = \boldsymbol{X}^{(s-1)} + \boldsymbol{\varepsilon}_s e_{\boldsymbol{\sigma}(s)} = \boldsymbol{X}^{(0)} + \sum_{i=1}^s \boldsymbol{\varepsilon}_i e_{\boldsymbol{\sigma}(i)}$ where e_j is the unit vector with a 1 in coordinate j.
5. Choose $\boldsymbol{s} \sim [m]$ and return the edge $(\boldsymbol{X}, \boldsymbol{Y}) = (\boldsymbol{X}^{(s-1)}, \boldsymbol{X}^{(s)})$.

Note that the above process can be equivalently defined as obtaining $\boldsymbol{X}^{(s)}$ by selecting a uniform random coordinate $\boldsymbol{i}$ where $\boldsymbol{X}_{\boldsymbol{i}}^{(s-1)} = 0$ and flipping that bit to a random value in $\{\pm 1\}$, with equal probability. This results in a random walk $\boldsymbol{X}^{(0)}, \boldsymbol{X}^{(1)}, \ldots, \boldsymbol{X}^{(m)}$ of length m where each $(\boldsymbol{X}^{(s-1)}, \boldsymbol{X}^{(s)})$ is a random out-going edge from $\boldsymbol{X}^{(s-1)}$. We use two main claims regarding this random walk to complete the proof of the theorem. The first is that choosing an edge $(\boldsymbol{X}, \boldsymbol{Y}) \sim \mathcal{D}$ is approximately the same as choosing a uniformly random edge from the middle layers.

▷ Claim 31. Fix any $z \in \mathsf{Mid}(\ell)$ and $s \in [m]$. Then $\mathbb{P}[\boldsymbol{X}^{(s)} = z] = \Theta(3^{-n})$. I.e., each step of the random walk is approximately uniformly distributed over $\mathsf{Mid}(\ell)$. As a corollary, for any fixed edge $(u, v) \in E_\ell$, we have

$$\mathbb{P}_{(\boldsymbol{X}, \boldsymbol{Y}) \sim \mathcal{D}}[(\boldsymbol{X}, \boldsymbol{Y}) = (u, v)] = \Theta\left(\frac{1}{n \cdot 3^n}\right).$$

The second claim is that the probability of $(\boldsymbol{X}, \boldsymbol{Y}) \sim \mathcal{D}$ being an influential edge is small.

▷ Claim 32. $\mathbb{P}_{(\boldsymbol{X},\boldsymbol{Y})\sim\mathcal{D}}[S(\boldsymbol{X}) \neq S(\boldsymbol{Y})] \leq O\left(\frac{1}{\sqrt{m}}\right)$.

We defer the proof of both claims to the end of the section. We now prove Theorem 4 using Claim 31 and Claim 32 as follows. Let E denote the edges of the ternary hypercube and let $E_\ell = \{(u,v) \in E : u, v \in \mathsf{Mid}(\ell)\}$. By definition,

$$\mathbb{I}(S) = \frac{1}{3^n} \cdot \Big(|\{(u,v) \in E \setminus E_\ell : S(u) \neq S(v)\}| + |\{(u,v) \in E_\ell : S(u) \neq S(v)\}| \Big).$$

The first term is bounded using Fact 24 as

$$\frac{|\{(u,v) \in E \setminus E_\ell : S(u) \neq S(v)\}|}{3^n} \leq \frac{|E \setminus E_\ell|}{3^n} \leq 2n \cdot \frac{|\overline{\mathsf{Mid}(\ell)}|}{3^n} \leq 2n \cdot 2\exp(-\ell^2/2n) = O(1)$$

since every vertex has degree at most $2n$. The second term is bounded as

$$\begin{aligned}\frac{|\{(u,v) \in E_\ell : S(u) \neq S(v)\}|}{3^n} &= \frac{|E_\ell|}{3^n} \cdot \mathbb{P}_{(u,v)\sim E_\ell}[S(u) \neq S(v)] \leq \frac{2n}{3} \cdot \mathbb{P}_{(u,v)\sim E_\ell}[S(u) \neq S(v)] \\ &\leq Ln \cdot \mathbb{P}_{(\boldsymbol{X},\boldsymbol{Y})\sim\mathcal{D}}[S(\boldsymbol{X}) \neq S(\boldsymbol{Y})] \\ &\leq L'n \cdot m^{-1/2} = L' \cdot n^{3/4} \log^{1/4} n\,,\end{aligned}$$

where L, L' are absolute constants. The first inequality follows simply from $E_\ell \subset E$ and $|E| = 2n \cdot 3^{n-1}$. The second inequality follows from Claim 31 and the third inequality follows from Claim 32. This completes the proof of Theorem 4. ◀

Let us now complete the deferred proofs of Claim 31 and Claim 32.

Proof of Claim 31. Let $\|z\|_1 = \frac{2n}{3} + r$ where $|r| = O(\sqrt{n \log n})$. In order for $X^{(s)} = z$ to occur we must have $\left\|X^{(0)}\right\|_1 = \frac{2n}{3} + r - s$. Thus, the probability is

$$\begin{aligned}\mathbb{P}[X^{(s)} = z] &= \frac{1}{3^n}\left(\binom{n}{\frac{2n}{3}+r-s} \cdot 2^{\frac{2n}{3}+r-s}\right) \cdot \left(\binom{n}{\frac{2n}{3}+r} \cdot 2^{\frac{2n}{3}+r}\right)^{-1} \\ &= \frac{1}{3^n} \cdot \frac{1}{2^s} \cdot \binom{n}{\frac{2n}{3}+r-s}\binom{n}{\frac{2n}{3}+r}^{-1} = \Theta(3^{-n})\end{aligned}$$

where the last step is due to the following fact:

If $|r| \leq O(\sqrt{n \log n})$ and $s = O(\sqrt{\frac{n}{\log n}})$, then $\binom{n}{\frac{2n}{3}+r-s}\binom{n}{\frac{2n}{3}+r}^{-1} = \Theta(2^s)$. As a corollary, the number of points in the ternary cube with hamming weight $\frac{2n}{3} + r - s$ and $\frac{2n}{3} + r$ differ by at most a constant factor.

This is proved as follows.

$$\begin{aligned}\frac{\binom{n}{\frac{2n}{3}+r-s}}{\binom{n}{\frac{2n}{3}+r}} &= \frac{\left(\frac{2n}{3}+r\right)!\left(\frac{n}{3}-r\right)!}{\left(\frac{2n}{3}+r-s\right)!\left(\frac{n}{3}-r+s\right)!} = \prod_{p=0}^{s-1} \frac{\frac{2n}{3}+r-p}{\frac{n}{3}-r+s-p} = 2^s \cdot \prod_{p=0}^{s-1} \frac{\frac{n}{3}+\frac{r}{2}-\frac{p}{2}}{\frac{n}{3}-r+s-p} \\ &= 2^s \cdot \prod_{p=0}^{s-1} \frac{\frac{n}{3}-r+s-p+\left(\frac{3r}{2}+\frac{p}{2}-s\right)}{\frac{n}{3}-r+s-p} = 2^s \cdot \prod_{p=0}^{s-1}\left(1 + \frac{\frac{3r}{2}+\frac{p}{2}-s}{\frac{n}{3}-r+s-p}\right)\end{aligned}$$

Observe that the numerator inside the product is $\pm O(\sqrt{n \log n})$ since r is the dominating term and the denominator is $\Omega(n)$ since $n/3$ is the dominating term. Therefore, we have

$$\frac{\binom{n}{\frac{2n}{3}+r-s}}{\binom{n}{\frac{2n}{3}+r}} = 2^s \cdot \left(1 \pm O\left(\sqrt{\frac{\log n}{n}}\right)\right)^s = \Theta(1) \cdot 2^s$$

since $s = O(\sqrt{\frac{n}{\log n}})$. ◁

Proof of Claim 32. Let S be an intersection of halfspaces $S = \bigcap_{v \in V} H_v$, with thresholds $\tau : V \to \mathbb{R}$, in the form promised by Proposition 30. In particular, each vector $v \in V$ has the DSS property (Definition 27). Fix any value of $\boldsymbol{X}^{(0)} = X^{(0)}$ and fix any permutation σ and sign-vector $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_m) \in \{\pm 1\}^m$ in the definition of $\mathcal{D}$, and consider the resulting fixed values of $X^{(0)}, X^{(1)}, \ldots, X^{(m)}$. Define $a : V \to \mathbb{R}$ as $a(v) = \langle v, X^{(0)} \rangle - \tau(v)$.

For each $v \in V$, consider the sequences $W_v^{+a(v)} := W_v^{+a(v)}(\cdot\,; \sigma, \varepsilon)$. For each $X^{(s)}$, observe that $X^{(s)} \in H_v$ if and only if $\langle v, X^{(s)} \rangle < \tau(v)$, which is equivalent to the condition $W_v(s) < 0$, since

$$W_v^{+a(v)}(s) = a(v) + \sum_{j=1}^{s} \varepsilon_j \cdot v_{\sigma(j)} = \left(\sum_{i : X_i^{(0)} \neq 0} v_i X_i^{(0)} \right) - t(v) + \sum_{j=1}^{s} X_{\sigma(j)}^{(s)} v_{\sigma(j)}$$
$$= \left(\sum_{j : X_j^{(s)} \neq 0} X_j^{(s)} v_j \right) - t(v) = \langle X^{(s)}, v \rangle - t(v)\,.$$

Therefore $X^{(s)} \in S$ if and only if $W_v^{+a(v)}(s) < 0$ for all $v \in V$, which is equivalent to $M_V^{+a}(s) < 0$ where M_V^{+a} is the max-walk (recall Definition 28). Then for fixed sequence $X^{(0)}, \ldots, X^{(m)}$ and uniformly random $\boldsymbol{s} \sim [m]$, the probability that $(X^{(\boldsymbol{s}-1)}, X^{(\boldsymbol{s})})$ is an influential edge is equal to $\frac{C(M_V^{+a})}{m}$. Therefore, taking $\boldsymbol{\sigma}$ and $\boldsymbol{\varepsilon}$ to be random, we have

$$\mathop{\mathbb{P}}_{(\boldsymbol{X}, \boldsymbol{Y}) \sim \mathcal{D}} \left[S(\boldsymbol{X}) \neq S(\boldsymbol{Y}) \right] = \frac{1}{m} \cdot \mathbb{E}\left[C(\boldsymbol{M}_V^{+\boldsymbol{a}}) \right] = O(\sqrt{m})\,,$$

where the final bound is due to Lemma 29, since each vector in V was assumed to have the DSS property. This concludes the proof of the claim. $\lhd$

3.1.3 Crossing Bound for the Max-Walk: Proof of Lemma 29

We prove an upper bound on the number of times the maximum of a set of one-dimensional random walks can change sign. Let us define certain special events in a random walk.

Fix any walk time m and let $W : \{0\} \cup [m] \to \mathbb{R}$. We define:

- A *downcrossing* of W is a time $t \in [m]$ such that $W(t) < 0 \leq W(t-1)$. $C_\downarrow(W)$ is the number of downcrossings of W.
- An *upcrossing* of W is a time $t \in [m]$ such that $W(t) \geq 0 > W(t)$. $C_\uparrow(W)$ is the number of upcrossings of W.
- A *downwards level return* of W is any time t such that either:
  - If $W(0) \geq 0$ then the smallest time $t \in [m]$ such that $W(t) < W(0)$ is a downwards level return.
  - For any upcrossing s of W, the first time $t > s$ such that $W(t) < W(s)$ is a downwards level return.

  We write $L_\downarrow(W)$ for the number of downwards level returns of W.
- The *downwards level decrease* times of W is the unique sequence $s_1 < s_2 < \cdots$ defined inductively as follows.
  - If $W(0) \geq 0$ then s_1 is the first time such that $W(s_1) < W(0)$. Otherwise let t be the first upcrossing of W. Then s_1 is the first time such that $W(s_1) < W(t)$.
  - For $i > 1$, if $W(s_{i-1}) \geq 0$ then $s_i \in [m]$ is the smallest time such that $W(s_i) < W(s_{i-1})$. Otherwise, if $W(s_{i-1}) < 0$, then let t be the first upcrossing $t > s_{i-1}$ and define s_i as the first time $s_i > t$ such that $W(s_i) < W(t)$.

  We write $S_\downarrow(W)$ for the number of downwards level decreases of W.

- The *upwards level increase* times of W is the unique sequence $t_1 < t_2 < \cdots$ defined inductively as follows.
  - If $W(0) < 0$ then t_1 is the first time such that $W(t_1) > W(0)$. Otherwise let s be the first downcrossing of W. Then t_1 is the first time such that $W(t_1) > W(t)$.
  - For $i > 1$, if $W(t_{i-1}) < 0$ then t_i is the first time such that $W(t_i) > W(t_{i-1})$. Otherwise if $W(t_{i-1}) \geq 0$, then let s be the first downcrossing $s > t_{i-1}$ and define t_i as the first time $t_i > s$ such that $W(t_i) > W(s)$.

  We write $S_\uparrow(W)$ for the number of upwards level increases of W.

The main technical tool in our analysis is the following version of Sparre Andersen's fluctuation theorem [26], as found in [3, Prop. 4.1]. Recall the definition of $\boldsymbol{W}_x$ from Definition 25 and the DSS property from Definition 27.

▶ **Theorem 33** (Sparre Anderson; see [3], Proposition 4.1)**.** *For every $m \in \mathbb{N}$, if $x \in \mathbb{R}^m$ has the DSS property, then the random walk $\boldsymbol{W}_x$ satisfies*

$$\mathbb{P}\left[\forall t \in [m] : \boldsymbol{W}_x(t) > 0\right] = g(m) := \frac{1}{4^m}\binom{2m}{m}.$$

We define a random variable $\boldsymbol{R}$ on the positive integers with

$$\forall t \in \mathbb{N}\,,\ \mathbb{P}\left[\boldsymbol{R} = t\right] := g(t-1) - g(t) = \frac{1}{4^{t-1}}\binom{2(t-1)}{t-1} - \frac{1}{4^t}\binom{2t}{t},$$

where we define $g(0) := 1$. For each $m \in \mathbb{N}$, we also define a random variable $\boldsymbol{Q}^{(m)}$ by the following process. Set $q = 0$ and $X = 0$; while $X < m$, increment q and set $X \leftarrow X + \boldsymbol{R}$ where $\boldsymbol{R}$ is a new independent copy of the random variable defined above. Then set $\boldsymbol{Q}^{(m)} = q$ once this process terminates; note that $\boldsymbol{Q}^{(0)} = 0$. Observe that for every $k \in \mathbb{N}$,

$$\mathbb{P}\left[\boldsymbol{Q}^{(m)} \geq k\right] = \mathbb{P}\left[\boldsymbol{R}_1 + \boldsymbol{R}_2 + \cdots + \boldsymbol{R}_k \leq m\right]$$

where each $\boldsymbol{R}_i$ is an independent copy of $\boldsymbol{R}$, and

$$\mathbb{E}\left[\boldsymbol{Q}^{(m)}\right] = \sum_{t=1}^{m} \mathbb{P}\left[\boldsymbol{R} = t\right] \cdot \left(1 + \mathbb{E}\left[\boldsymbol{Q}^{(m-t)}\right]\right).$$

The following holds due to Theorem 33.

▶ **Proposition 34.** *Let $x \in \mathbb{R}^m$ have the DSS property, let $a \in \mathbb{R}$, and let $\boldsymbol{s}_1$, $\boldsymbol{t}_1$ denote the first downwards level decrease time and upwards level increase times of $\boldsymbol{W}_x^{+a}$, respectively. Then for all $z \in [m]$,*

1. *If $a \geq 0$ then $\mathbb{P}\left[\boldsymbol{s}_1 = z\right] = \mathbb{P}\left[\boldsymbol{R} = z\right]$; and,*
2. *If $a < 0$ then $\mathbb{P}\left[\boldsymbol{t}_1 = z\right] = \mathbb{P}\left[\boldsymbol{R} = z\right]$.*

▶ **Proposition 35.** *Let $x \in \mathbb{R}^m$ have the DSS property and let $a \in \mathbb{R}$. Then*

$$\mathbb{E}\left[\boldsymbol{Q}^{(m)}\right] = \mathbb{E}\left[S_\downarrow(\boldsymbol{W}_x^{+a}) + S_\uparrow(\boldsymbol{W}_x^{+a})\right].$$

Proof. By induction on m. For $m = 1$ we have $\mathbb{E}\left[\boldsymbol{Q}^{(1)}\right] = \mathbb{P}\left[\boldsymbol{R} = 1\right] = 1/2$ and

$$\mathbb{E}\left[S_\downarrow(\boldsymbol{W}_x^{+a}) + S_\uparrow(\boldsymbol{W}_x^{+a})\right] = 1/2$$

since the random walk has probability $1/2$ of increasing or decreasing in the first step; if $a \geq 0$ then the walk must decrease to create a downwards level decrease, while if $a > 0$ then the walk must increase to create an upwards level increase.

Now let $m > 1$. Suppose $a \geq 0$ without loss of generality. Then the first level increase or decrease is a downwards level decrease. Let $\boldsymbol{s}_1$ be the first downwards level decrease and let $\boldsymbol{y}$ denote the random subsequence of x that remains after removing the first $\boldsymbol{s}_1$ elements according to the random permutation $\boldsymbol{\sigma}$. Then by induction and Proposition 34,

$$\begin{aligned}
&\mathbb{E}\left[S_{\downarrow}(\boldsymbol{W}_x^{+a}) + S_{\uparrow}(\boldsymbol{W}_x^{+a})\right] \\
&= \sum_{t=1}^{m} \mathbb{P}\left[\boldsymbol{s}_1 = t\right] \cdot \left(1 + \mathbb{E}\left[S_{\downarrow}\left(\boldsymbol{W}_{\boldsymbol{y}}^{+\boldsymbol{W}_x(\boldsymbol{s}_1)}\right) + S_{\uparrow}\left(\boldsymbol{W}_{\boldsymbol{y}}^{+\boldsymbol{W}_x(\boldsymbol{s}_1)}\right) \;\middle|\; \boldsymbol{s}_1 = t\right]\right) \\
&= \sum_{t=1}^{m} \mathbb{P}\left[\boldsymbol{R} = t\right] \cdot \left(1 + \mathbb{E}\left[\boldsymbol{Q}^{(m-t)}\right]\right) = \mathbb{E}\left[\boldsymbol{Q}^{(m)}\right].
\end{aligned}$$ ◀

For a sequence $W : \{0\} \cup [m] \to \mathbb{R}$, write $Z(W) = \sum_{t=1}^{m} \mathbb{1}\left[W(t) \in \{0, \pm 1\}\right]$.

▶ **Lemma 36.** *For any m, $\mathbb{E}\left[Z(\boldsymbol{W}_{\vec{1}})\right] = O(\sqrt{m})$.*

Proof. We first bound the number of times t such that $\boldsymbol{W}_{\vec{1}}(t) = 0$. If t is odd then $\mathbb{P}\left[\boldsymbol{W}_{\vec{1}}(t) = 0\right] = 0$. If t is even then there is a universal constant C such that

$$\mathbb{P}\left[\boldsymbol{W}_{\vec{1}}(t) = 0\right] = \frac{1}{2^t}\binom{t}{t/2} \leq C \cdot \frac{1}{\sqrt{t}}.$$

Therefore the expected number of times t with $\boldsymbol{W}_{\vec{1}}(t) = 0$ is at most

$$\sum_{t \text{ even}} \mathbb{P}\left[\boldsymbol{W}_{\vec{1}}(t) = 0\right] \leq C \cdot \sum_{t=1}^{m} \frac{1}{\sqrt{t}} = O(\sqrt{m}).$$

Now observe that the expected number of times t where $\boldsymbol{W}_{\vec{1}}(t) = 1$ is the average of the expected number of times where the shifted walks $\boldsymbol{W}_{\vec{1}}^{+a}$ is 0 on domain $[m-1]$, where $a = \pm 1$, and the same holds for the number of times t where $\boldsymbol{W}_{\vec{1}}(t) = -1$. ◀

▶ **Proposition 37.** *There exists $x \in \mathbb{R}^m$ with the DSS property such that $\mathbb{E}\left[S_{\downarrow}(\boldsymbol{W}_x) + S_{\uparrow}(\boldsymbol{W}_x)\right] \leq \mathbb{E}\left[Z(\boldsymbol{W}_{\vec{1}})\right]$. As a consequence,*

$$\mathbb{E}\left[\boldsymbol{Q}^{(m)}\right] = O(\sqrt{m}).$$

Proof. Let $\delta := \frac{1}{3m}$. Let $\boldsymbol{x} := \vec{1} + \boldsymbol{z}$ where $\boldsymbol{z} \sim [-\delta, \delta]^m$ uniformly at random. Note that $\boldsymbol{x}$ has the DSS property with probability 1. For any fixed $z \in [-\delta, \delta]^m$, any permutation σ, and any $r \in \{\pm 1\}^n$, write $W_x(t) := W_x(t; \sigma, \varepsilon)$ for $x = \vec{1} + z$. Then we have $W_x(t; \sigma, \varepsilon) \in [W_{\vec{1}}(t; \sigma, \varepsilon) - 1/3, W_{\vec{1}}(t; \sigma, \varepsilon) + 1/3]$. Now fix any $z \in [-\delta, \delta]^m$ such that $x = \vec{1} + z$ has the DSS property; we show that it satisfies the required condition.

Let $s_1 < s_2 < \cdots < s_k$ be the downwards level decreasing or upwards level increasing points for W_x, let $s_0 = 0$, and observe that a point cannot be both downwards level decreasing and upwards level increasing. We show by induction on i that $|W_x(s_i)| \leq 1 + 1/3$ and therefore that $W_{\vec{1}}(s_i) \in \{0, \pm 1\}$. Since $W_x(0) = 0$ it must be that s_1 is downwards level decreasing and $W_x(s_1) < 0 = W_x(0) \leq W_x(s_1 - 1)$ and therefore $W_x(s_1) \geq W_x(s_1 - 1) - 1 - 1/3 \geq -4/3$ so it must be that $W_{\vec{1}}(s_1) \in \{-1, 0\}$. For $i > 1$, suppose that s_i is a downwards level decreasing point. If there exists an upcrossing point $a > s_{i-1}$ such that $W_x(s_i) < W_x(a)$, then we observe that $W_x(a-1) < 0 \leq W_x(a)$ and therefore $W_x(a) < 1 + 1/3$ so $W_{\vec{1}}(a) \in \{0, 1\}$. Now $W_x(s_i) < W_x(a) \leq W_x(s_i - 1)$ so it must be that $-1 - 1/3 \leq W_x(s_i) < 1 + 1/3$ so $W_{\vec{1}}(s_i) \in \{0, \pm 1\}$. On the other hand, if $W_x(s_{i-1}) \geq 0$ and $W_x(s_i) < W_x(s_{i-1})$ then by induction we have $W_{\bar{1}}(s_{i-1}) \in \{0,1\}$, and also $W_x(s_i - 1) > W_x(s_{i-1})$, so again we have $-1 - 1/3 \le W_x(s_i) < 1 + 1/3$ and therefore $W_{\bar{1}}(s_i) \in \{0, \pm 1\}$. A similar argument holds for the upwards level increasing points.

The conclusion now follows from Proposition 35 and Lemma 36, since for the $x \in \mathbb{R}^m$ defined in the current proof,

$$\mathbb{E}\left[\boldsymbol{Q}^{(m)}\right] = \mathbb{E}\left[S_\downarrow(\boldsymbol{W}_x) + S_\uparrow(\boldsymbol{W}_x)\right] \le \mathbb{E}\left[Z(\boldsymbol{W}_{\bar{1}})\right] = O(\sqrt{m}) \,. \qquad \blacktriangleleft$$

▶ **Proposition 38.** *Let X be a set of sequences $x \in \mathbb{R}^m$ each having the DSS property, and let $a : X \to \mathbb{R}$ be arbitrary. Then*

$$\mathbb{E}\left[L_\downarrow(\boldsymbol{M}_X^{+a})\right] \le \mathbb{E}\left[\boldsymbol{Q}^{(m)}\right] .$$

Proof. By induction on m. For $m = 1$, the probability that $\boldsymbol{M}_X^{+a}$ has a downwards level return is at most $1/2$, because if $M_X^{+a}(0) \ge 0$, all of the maximizing constituent walks $x \in X$ satisfying $W_x(0)^{+a(x)} = M_X^{+a}(0)$ must decrease. If $M_X^{+a}(0) < 0$ then there is no downwards level return for $m = 1$.

Let $m > 1$ and consider two cases. First assume that $M_X^{+a}(0) \ge 0$ and let $x \in X$ be an arbitrary constituent walk satisfying $W_x^{+a(x)}(0) = M_X^{+a}(0)$. For fixed permutation σ and sign vector ε, let s_1 be the first downwards level return point of $M_X^{+a}(\cdot\,; \sigma, \varepsilon)$ and let s_1' be the first downwards level return point of $W_x^{+a(x)}(\cdot\,; \sigma, \varepsilon)$. Note that $s_1' \le s_1$ since M_X^{+a} is the maximum of its constituents.

Now we may write

$$\mathbb{E}\left[L_\downarrow(\boldsymbol{M}_X^{+a})\right] = \sum_{s=1}^{m} \mathbb{P}\left[\boldsymbol{s}_1 = s\right] \left(1 + \mathbb{E}\left[L_\downarrow\left(\boldsymbol{M}_{\boldsymbol{Y}}^{+\boldsymbol{b}}\right) \;\middle|\; \boldsymbol{s}_1 = s\right]\right) ,$$

where $\boldsymbol{Y}$ denotes the set of vectors X after removing the first t coordinates according to the random permutation $\boldsymbol{\sigma}$ and $\boldsymbol{b} : \boldsymbol{Y} \to \mathbb{R}$ is the starting point $\boldsymbol{b}(\boldsymbol{y}) = \boldsymbol{W}_x^{+a(x)}(t)$ of each walk $\boldsymbol{y} \in \boldsymbol{Y}$ obtained from the original vector $x \in X$ by removing the first t coordinates according to $\boldsymbol{\sigma}$. By induction, this is

$$\mathbb{E}\left[L_\downarrow\left(\boldsymbol{M}_{\boldsymbol{Y}}^{+\boldsymbol{b}}\right) \;\middle|\; \boldsymbol{s}_1 = s\right] \le \mathbb{E}\left[\boldsymbol{Q}^{(m-s)}\right] .$$

Now we write $\boldsymbol{s}_1 = \boldsymbol{s}_1' + (\boldsymbol{s}_1 - \boldsymbol{s}_1')$ where the second term is non-negative. Then

$$\begin{aligned}
\mathbb{E}\left[L_\downarrow(\boldsymbol{M}_X^{+a})\right] &\le \sum_{s=1}^{m} \mathbb{P}\left[\boldsymbol{s}_1' + (\boldsymbol{s}_1 - \boldsymbol{s}_1') = s\right] \cdot \left(1 + \mathbb{E}\left[\boldsymbol{Q}^{(m-s)}\right]\right) \\
&= \sum_{t=1}^{m} \mathbb{P}\left[\boldsymbol{s}_1' = t\right] \sum_{s=0}^{m-t} \mathbb{P}\left[\boldsymbol{s}_1 - \boldsymbol{s}_1' = s \;\middle|\; \boldsymbol{s}_1' = t\right] \left(1 + \mathbb{E}\left[\boldsymbol{Q}^{(m-(s+t))}\right]\right) .
\end{aligned}$$

The inner sum is a convex sum of terms $1 + \mathbb{E}\left[\boldsymbol{Q}^{(m-(s+t))}\right]$ which are each bounded by $1 + \mathbb{E}\left[\boldsymbol{Q}^{(m-t)}\right]$ because $\mathbb{E}\left[\boldsymbol{Q}^{(k)}\right] \ge \mathbb{E}\left[\boldsymbol{Q}^{(k')}\right]$ when $k \ge k'$. We also have $\mathbb{P}\left[\boldsymbol{s}_1' = t\right] = \mathbb{P}\left[\boldsymbol{R} = t\right]$ due to Proposition 34. Therefore,

$$\mathbb{E}\left[L_\downarrow(\boldsymbol{M}_X^{+a})\right] \le \sum_{t=1}^{m} \mathbb{P}\left[\boldsymbol{R} = t\right] \left(1 + \mathbb{E}\left[\boldsymbol{Q}^{(m-t)}\right]\right) = \mathbb{E}\left[\boldsymbol{Q}^{(m)}\right] .$$

We must now handle the case where $M_X^{+a}(0) < 0$. Let $\boldsymbol{t}$ be the smallest time where $\boldsymbol{M}_X^{+a}(\boldsymbol{t}) \ge 0$. Then

$$\mathbb{E}\left[L_\downarrow(\boldsymbol{M}_X^{+a})\right] = \sum_{t=1}^{m} \mathbb{P}\left[\boldsymbol{t} = t\right] \mathbb{E}\left[L_\downarrow\left(\boldsymbol{M}_{\boldsymbol{Y}}^{+\boldsymbol{b}}\right) \;\middle|\; \boldsymbol{t} = t\right] ,$$

where $\boldsymbol{Y}$ and $\boldsymbol{b}$ are defined similarly as before, as the sequences X after removing the first t coordinates according to the random permutation $\boldsymbol{\sigma}$ and $\boldsymbol{b}(\boldsymbol{y}) = \boldsymbol{W}_x^{+a(x)}(t)$ is where the x walk ended up at time t. This new walk starts above 0 so the above argument applies and the conclusion holds. ◀

We can now prove Lemma 29.

Proof of Lemma 29. First observe that $C(\boldsymbol{M}_X^{+a}) \leq 2\,C_\downarrow(\boldsymbol{M}_X^{+a}) + 1$ so it suffices to bound $C_\downarrow(\boldsymbol{M}_X^{+a})$. By definition it holds that $C_\downarrow(\boldsymbol{M}_X^{+a}) \leq L_\downarrow(\boldsymbol{M}_X^{+a})$, so by Proposition 38 and Proposition 37, we have

$$\mathbb{E}\left[C_\downarrow(\boldsymbol{M}_X^{+a})\right] \leq \mathbb{E}\left[\boldsymbol{Q}^{(m)}\right] = O(\sqrt{m})\,.$$ ◀

3.2 Lower Bound

Recall the definition of the *influence* of a set in the ternary hypercube given in Equation (2). In this section, we show that there exists a convex set whose influence is $\Omega(n^{3/4})$, nearly matching the upper bound given in Theorem 4.

The construction of the high-influence set is obtained by considering the intersection of $2^{\sqrt{n}}$ random halfspaces whose distance from the origin is $\Theta(n^{3/4})$. This approach is inspired by [16, Theorem 2], who showed that in the Boolean hypercube, an intersection of k random halfspaces with an appropriately chosen distance from the origin will have expected influence $\Omega(\sqrt{n \log k})$. In the ternary hypercube, this type of argument still works as long as $k \leq 2^{O(\sqrt{n})}$. This type of construction was also used by Nazarov [20] to show the existence of convex sets in $\mathbb{R}^n$ whose Gaussian surface area is $\Omega(n^{1/4})$, which matches the $O(n^{1/4})$ upper bound proven by Ball [1].

We prove the following theorem in the full version of the paper.

▶ **Theorem 3.** *There exists a convex set $S \subseteq \{0, \pm 1\}^n$ with influence $\mathbb{I}(S) = \Omega(n^{3/4})$.*

References

1. Keith Ball. The reverse isoperimetric problem for gaussian measure. *Discrete Comput. Geom.*, 10(4):411–420, December 1993. doi:10.1007/BF02573986.
2. Aleksandrs Belovs and Eric Blais. A polynomial lower bound for testing monotonicity. In *Proceedings, ACM Symposium on Theory of Computing (STOC)*, 2016.
3. Quentin Berger and Loïc Béthencourt. An application of sparre andersen's fluctuation theorem for exchangeable and sign-invariant random variables. *arXiv preprint*, 2023. arXiv:2304.09031.
4. Piotr Berman, Meiram Murzabulatov, and Sofya Raskhodnikova. The power and limitations of uniform samples in testing properties of figures. *Algorithmica*, 81(3):1247–1266, 2019. doi:10.1007/s00453-018-0467-9.
5. Piotr Berman, Meiram Murzabulatov, and Sofya Raskhodnikova. Testing convexity of figures under the uniform distribution. *Random Struct. Algorithms*, 54(3):413–443, 2019. doi:10.1002/rsa.20797.
6. Piotr Berman, Meiram Murzabulatov, and Sofya Raskhodnikova. Tolerant testers of image properties. *ACM Trans. Algorithms*, 18(4):37:1–37:39, 2022. doi:10.1145/3531527.
7. Eric Blais and Abhinav Bommireddi. On testing and robust characterizations of convexity. In *Proceedings, International Workshop on Randomization and Computation (RANDOM)*, 2020. doi:10.4230/LIPIcs.APPROX/RANDOM.2020.18.
8. Eric Blais, Renato Ferreira Pinto Jr, and Nathaniel Harms. VC dimension and distribution-free sample-based testing. In *Proceedings, ACM Symposium on Theory of Computing (STOC)*, 2021.

9 Béla Bollobás and Imre Leader. Edge-isoperimetric inequalities in the grid. *Combinatorica*, 11(4):299–314, 1991.
10 Xi Chen, Adam Freilich, Rocco A. Servedio, and Timothy Sun. Sample-based high-dimensional convexity testing. In *Proceedings, International Workshop on Randomization and Computation (RANDOM)*, 2017.
11 Xi Chen, Erik Waingarten, and Jinyu Xie. Beyond Talagrand: New lower bounds for testing monotonicity and unateness. In *Proceedings, ACM Symposium on Theory of Computing (STOC)*, 2017.
12 Anindya De, Shivam Nadimpalli, and Rocco A Servedio. Convex influences. In *13th Innovations in Theoretical Computer Science Conference (ITCS)*, 2022.
13 Konrad Engel. *Sperner Theory*. Cambridge University Press, 1997.
14 Oded Goldreich, Shafi Goldwasser, and Dana Ron. Property testing and its connection to learning and approximation. *Journal of the ACM*, 45(4):653–750, 1998.
15 Nathaniel Harms and Yuichi Yoshida. Downsampling for testing and learning in product distributions. In *Proceedings, International Colloquium on Automata, Languages and Programming (ICALP)*, 2022.
16 Daniel M. Kane. The average sensitivity of an intersection of half spaces. In David B. Shmoys, editor, *Symposium on Theory of Computing, STOC 2014, New York, NY, USA, May 31 - June 03, 2014*, pages 437–440. ACM, 2014. doi:10.1145/2591796.2591798.
17 Adam R. Klivans, Ryan O'Donnell, and Rocco A. Servedio. Learning geometric concepts via gaussian surface area. In *Proceedings, IEEE Symposium on Foundations of Computer Science (FOCS)*. IEEE Computer Society, 2008. doi:10.1109/FOCS.2008.64.
18 Nathan Linial, Yishay Mansour, and Noam Nisan. Constant depth circuits, fourier transform, and learnability. *J. ACM*, 40(3):607–620, 1993. doi:10.1145/174130.174138.
19 Nicholas Metropolis and Gian-Carlo Rota. Combinatorial structure of the faces of the n-cube. *SIAM J. Appl. Math.*, 35(4):689–694, December 1978.
20 Fedor Nazarov. On the maximal perimeter of a convex set in $\mathbb{R}^n$ with respect to a gaussian measure. In *Geometric Aspects of Functional Analysis: Israel Seminar 2001-2002*, pages 169–187, 2003. URL: https://api.semanticscholar.org/CorpusID:125334165.
21 Ryan O'Donnell. *Analysis of Boolean Functions*. Cambridge University Press, 2014. URL: http://www.cambridge.org/de/academic/subjects/computer-science/algorithmics-complexity-computer-algebra-and-computational-g/analysis-boolean-functions.
22 Luis Rademacher and Navin Goyal. Learning convex bodies is hard. In *Proceedings, Conference on Learning Theory (COLT)*, 2009. URL: http://www.cs.mcgill.ca/%7Ecolt2009/papers/030.pdf#page=1.
23 Luis Rademacher and Santosh S. Vempala. Testing geometric convexity. In *Proceedings, Foundations of Software Technology and Theoretical Computer Science (FSTTCS)*. Springer, 2004. doi:10.1007/978-3-540-30538-5_39.
24 Sofya Raskhodnikova. Approximate testing of visual properties. In *Proceedings, International Workshop on Randomization and Computation (RANDOM)*. Springer, 2003. doi:10.1007/978-3-540-45198-3_31.
25 Bernd Schmeltz. Learning convex sets under uniform distribution. In *Data Structures and Efficient Algorithms, Final Report on the DFG Special Joint Initiative*, Lecture Notes in Computer Science, pages 204–213. Springer, 1992. doi:10.1007/3-540-55488-2_28.
26 Erik Sparre Andersen. On the fluctuations of sums of random variables II. *Math. Scandinavica*, 2(2):195–223, 1954.