Lower Bounds for Convexity Testing

Xi Chen* Anindya De† Shivam Nadimpalli‡ Rocco A. Servedio§ Erik Waingarten¶

Abstract

We consider the problem of testing whether an unknown and arbitrary set $S \subseteq \mathbb{R}^n$ (given as a black-box membership oracle) is convex, versus ε-far from every convex set, under the standard Gaussian distribution. The current state-of-the-art testing algorithms for this problem make $2^{\tilde{O}(\sqrt{n})\cdot \mathrm{poly}(1/\varepsilon)}$ non-adaptive queries, both for the standard testing problem and for tolerant testing.

We give the first lower bounds for convexity testing in the black-box query model:

- We show that any one-sided tester (which may be adaptive) must use at least $n^{\Omega(1)}$ queries in order to test to some constant accuracy $\varepsilon > 0$.
- We show that any non-adaptive *tolerant* tester (which may make two-sided errors) must use at least $2^{\Omega(n^{1/4})}$ queries to distinguish sets that are ε_1-close to convex versus ε_2-far from convex, for some absolute constants $0 < \varepsilon_1 < \varepsilon_2$.

Finally, we also show that for any constant $c > 0$, any non-adaptive tester (which may make two-sided errors) must use at least $n^{1/4-c}$ queries in order to test to some constant accuracy $\varepsilon > 0$.

1 Introduction

High-dimensional convex geometry is a rich topic at the intersection of geometry, probability, and analysis (see [B+97, GW93, LL15, Tro18, Tko18, HW20], among many other works, for general overviews). Apart from its intrinsic interest, a strong motivation for the study of high-dimensional convex sets from the perspective of theoretical computer science is that convexity often translates into a form of mathematical niceness which facilitates efficient computation, as witnessed by the plethora of positive results in algorithms and optimization for convex functions and convex sets. In this work, the object of study is the convex set:

> A set $K \subset \mathbb{R}^n$ is convex if and only if for every two points $x, y \in \mathbb{R}^n$, any point z on the segment between x and y lies in K whenever x and y lie in K.

The above gives a "local" characterization of convex sets, where "local" refers to the fact that (aside from quantifying over *all* co-linear points x, z, y,) an algorithm may make three membership queries to check the condition — in particular, non-convexity can be verified with three queries. Can one relax the "for all" quantification to give a local condition which characterizes *approximately* convex sets? Is there an algorithm which, by making very few queries, can determine whether or not a set is (almost) convex?

A natural vantage point for this broad question is that of *property testing* [BLR93, RS96], which provides an algorithmic framework for studying the above questions. In our setting, we consider property testing of convex sets with respect to the *standard Gaussian distribution*, arguably the most natural distribution over $\mathbb{R}^n$. Indeed, various learning, property testing, and other algorithmic problems in the Gaussian setting have been intensively studied in theoretical computer science [KOS08, Vem10, MORS10, Kan11, Kan12, Kan14, KK14, KNOW14, Kan15, CFSS17, CDS19, DMN19, OSTK21, DMN21, HSSV22, DNS23]. Furthermore, while a large body of mathematical work (e.g. [Bor75, Bal93, LO99, Lat02, Naz03, Bor03, CEFM04, LO05, Bor08, Roy14])

*Columbia University. Email: xc2198@columbia.edu.
†University of Pennsylvania. Email: anindyad@seas.upenn.edu.
‡MIT. Email: shivamn@mit.edu.
§Columbia University. Email: rocco@cs.columbia.edu.
¶University of Pennsylvania. Email: ewaingar@seas.upenn.edu.

investigates the geometry of high-dimensional convex sets against the Gaussian distribution, convexity over Gaussian space arises naturally within theoretical computer science in the context of algorithmic discrepancy theory [Glu89, Ban10, LM15, Rot17, LRR17, Eld22, RR23b] and lattice problems [RR23a, Rot23, RSD24].

We consider the following algorithmic task: A (randomized) testing algorithm has black-box query access to an unknown and arbitrary function $f\colon \mathbb{R}^n \to \{0,1\}$ (the indicator function of a subset of $\mathbb{R}^n$), and its goal is to make as few membership queries on f as possible while deciding whether f is convex or ε-far from convex (meaning f and any indicator of a convex set $g : \mathbb{R}^n \to \{0,1\}$ disagree on $\boldsymbol{x} \sim N(0, I_n)$ with probability at least ε). Thus, a testing algorithm gives an efficiently-checkable (randomized) condition which all convex sets satisfy, and furthermore, any set which satisfies this condition is "almost" convex (with respect to the standard Gaussian distribution). For example, the definition of a convex set naturally leads to the following property testing question, whose positive resolution would directly give a "constant-query" testing algorithm (i.e. an algorithm whose query complexity depends only on ε and not on the ambient dimension n):

> Does there exist a probability distribution over co-linear points $\boldsymbol{x}, \boldsymbol{z}, \boldsymbol{y}$ in $\mathbb{R}^n$ such that the condition $\mathbf{Pr}[\boldsymbol{z} \in K \mid \boldsymbol{x}, \boldsymbol{y} \in K] \geq 1 - \delta(\varepsilon)$ implies that the set K must be ε-close to convex with respect to the standard Gaussian?[1]

In this work, we show the first non-trivial lower bounds for testing convexity under the standard Gaussian distribution. Our lower bounds not only give a negative resolution to the above question, they imply that, in a variety of property testing models (non-adaptive, adaptive, one-sided, two-sided, and tolerant), a dependence on the ambient dimension n is always necessary. Prior to this work, an $O(1/\varepsilon)$-query test was entirely possible for all of those models.[2]

As further discussed in Section 1.3, a number of prior works have studied convexity testing in a range of different settings, yet large gaps remain in our understanding. Most closely related are the works of [KOS08], who study learning convex sets over $N(0, I_n)$, and [CFSS17], who study testing convexity over $N(0, I_n)$ when restricted to sample-based testers (i.e. the algorithm can only query a given number of random points independently drawn from $N(0, I_n)$). On the upper bound side, the best algorithm for convexity testing [CFSS17] is based on [KOS08] and queries $2^{\tilde{O}(\sqrt{n})/\varepsilon^2}$ randomly sampled points from $N(0, I_n)$. Hence, this "sample-based" tester gives a non-adaptive property testing algorithm.[3] Turning to lower bounds, [CFSS17] showed that, when restricted to sample-based testers, (i) algorithms which incur one-sided error must make $2^{\Omega(n)}$ queries,[4] and (ii) algorithms which incur two-sided error must make $2^{\Omega(\sqrt{n})}$ queries. Importantly, lower bounds on sample-based testers do not imply any lower bounds for algorithms which are allowed to make unrestricted queries. There are many prominent property testing problems (e.g., linearity and monotonicity) where the complexity of sample-based testing is significantly higher than the complexity in the (standard) query-based model.[5]

1.1 Our Results and Discussion This work gives the first non-trivial lower bounds for query-based convexity testing. We prove three different lower bounds for three variants of the property testing model, which we now describe. As mentioned, the best known algorithm for convexity testing is the non-adaptive algorithm of [CFSS17], which makes $2^{\tilde{O}(\sqrt{n})/\varepsilon^2}$ non-adaptive queries (and makes two-sided error).

Our first result gives a polynomial lower bound for one-sided adaptive testers:

THEOREM 1.1. (ONE-SIDED ADAPTIVE LOWER BOUND) *For some absolute constant $\varepsilon > 0$, any one-sided (potentially adaptive) ε-tester for convexity over $N(0, I_n)$ must use $n^{\Omega(1)}$ queries.*

[1] Such a distribution would immediately yield a *proximity-oblivious* testing algorithm [GR11], one of the strongest forms of property testing. Prior to this work, the existence of a proximity-oblivious tester for convexity was entirely possible.

[2] An $\Omega(1/\varepsilon)$-query lower bound is easily seen to hold for essentially every non-trivial property, since this many queries are required to distinguish between the empty set (which is convex) and a random set of volume 2ε (which is far from convex and far from having most properties of interest).

[3] Recall that a *non-adaptive* testing algorithm is one in which the choice of its i-th query point does not depend on the responses received to queries $1, \ldots, i-1$.

[4] Recall that a *one-sided* tester for a class of functions is one which must accept (with probability 1) any function f that belongs to the class. This is in contrast to making *two-sided* error, where an algorithm may reject a function in the class with small probability.

[5] For example, linearity testing over $\{0,1\}^n$ admits $O(1/\varepsilon)$-query algorithms [BLR93], but requires $\Omega(n)$ queries for sample-based testers [GR16]. Monotonicity testing over $\{0,1\}^n$ admits poly(n)-query algorithms [GGL+00, CS13, CST14, KMS18], but requires $\Omega(2^{n/2})$ for sample-based testers [GGL+00].

We also consider a challenging and well-studied extension of the standard testing model which is known as *tolerant testing* [PRR06]. Recall that an $(\varepsilon_1, \varepsilon_2)$-tolerant tester for a class of functions is a testing algorithm which must accept with high probability if the input is ε_1-close to some function in the class and reject with high probability if the input is ε_2-far from every function in the class; thus the standard property testing model corresponds to $(0, \varepsilon)$-tolerant testing.

The sample-based algorithm for convexity testing that is given in [CFSS17] is based on agnostic learning results from [KOS08]. It follows easily from the analysis in [CFSS17] and results of [KOS08] that for any $0 \le \varepsilon_1 < \varepsilon_2$ with $\varepsilon_2 - \varepsilon_1 = \varepsilon$, the [CFSS17] approach gives a $2^{\tilde{O}(\sqrt{n})/\varepsilon^4}$-query sample-based algorithm for $(\varepsilon_1, \varepsilon_2)$-tolerant testing of convexity. As our final result, we give a mildly exponential lower bound on the query complexity of two-sided non-adaptive tolerant convexity testing:

THEOREM 1.2. (TWO-SIDED NON-ADAPTIVE TOLERANT TESTING LOWER BOUND) *There exist absolute constants $0 < \varepsilon_1 < \varepsilon_2 < 0.5$ such that any non-adaptive $(\varepsilon_1, \varepsilon_2)$-tolerant tester for convexity over $N(0, I_n)$ (which may make two-sided errors) must use at least $2^{\Omega(n^{1/4})}$ queries.*

Returning to the standard testing model, our final result gives a polynomial lower bound for two-sided non-adaptive testers:

THEOREM 1.3. (TWO-SIDED NON-ADAPTIVE LOWER BOUND) *For any constant $c > 0$, there is a constant $\varepsilon = \varepsilon_c > 0$ such that any non-adaptive ε-tester for convexity over $N(0, I_n)$ (which may make two-sided errors) must use at least $n^{1/4-c}$ queries.*

Since q-query non-adaptive lower bounds imply $(\log q)$-query adaptive lower bounds, Theorem 1.3 implies an $\Omega(\log n)$ two-sided adaptive convexity testing lower bound. (This is in contrast to the $n^{\Omega(1)}$-query lower bound against one-sided adaptive testers given by Theorem 1.1.)

1.2 Techniques Our lower bounds rely on a wide range of techniques and constructions, and draw inspiration from prior work on *monotonicity testing* of Boolean functions $f : \{0,1\}^n \to \{0,1\}$ [CST14, BB16, CWX17, PRW22, CDL+24].[6] Indeed, a conceptual contribution of our work is to highlight a (perhaps unexpected) connection between ideas in monotonicity testing and convexity testing. Our work thus adds to and strengthens a recently emerging analogy between monotone Boolean functions and high-dimensional convex sets [DNS21, DNS22, DNS24]. Establishing this connection requires a number of technical and conceptual innovations for each of our main results; we highlight some of the key ideas below.

1.2.1 The Nazarov Body A central role in our lower bounds in Theorem 1.1 and Theorem 1.2 is played by the so-called "Nazarov body" [Naz03, KOS08, CFSS17]. This is a randomized construction of a convex set $\boldsymbol{B}$ which is a slight variation of a construction originally due to Nazarov [Naz03], which is essentially as follows: we choose $N \approx 2^{\sqrt{n}}$ halfspaces $\boldsymbol{H}_1, \ldots, \boldsymbol{H}_N$ in the space $\mathbb{R}^n$, where each halfspace $\boldsymbol{H}_i$ is a *random halfspace* at a distance roughly $n^{1/4}$ from the origin. In more detail, each halfspace is $\boldsymbol{H}_i(x) := \mathbf{1}\{\boldsymbol{g}^i \cdot x \ge r\}$ where $r \approx n^{3/4}$ and $\boldsymbol{g}^i$ is drawn from $N(0, I_n)$. The convex set $\boldsymbol{B}$ is obtained by taking the intersection of all N halfspaces with $\mathrm{Ball}(\sqrt{n})$, the origin-centered ball of radius $\sqrt{n}$.[7] The exact parameters r and N are set carefully so that with high probability the "Gaussian volume" of $\boldsymbol{B}$, i.e. $\mathbf{Pr}_{\boldsymbol{g} \sim N(0,I_n)}[\boldsymbol{g} \in \boldsymbol{B}]$, is a constant bounded away from 0 and 1.

Note that for the Nazarov body $\boldsymbol{B}$ and any point $x \in \mathrm{Ball}(\sqrt{n}) \setminus \boldsymbol{B}$, there is a non-empty subset $J_x \subseteq [N]$ such that $j \in J_x$ iff $\boldsymbol{H}_j(x) = 0$, i.e., the point x *violates* the halfspace $\boldsymbol{H}_j$ for all $j \in J_x$. Now, define a point $x \in \mathrm{Ball}(\sqrt{n}) \setminus \boldsymbol{B}$ to lie in the set $\boldsymbol{U}_i$ if the set $J_x = \{i\}$, so $x \in \mathrm{Ball}(\sqrt{n})$ lies in $\boldsymbol{U}_i$ if $\boldsymbol{H}_i$ is the unique halfspace violated by x. The set $\boldsymbol{U} := \cup_{i \in [N]} \boldsymbol{U}_i$ is thus the set of "uniquely violated points" in $\mathrm{Ball}(\sqrt{n})$. A crucial feature of the Nazarov construction is that the Gaussian volume of the set of points which are uniquely violated, i.e., Gaussian volume of the set $\boldsymbol{U}$, is "large" compared to the Gaussian volume of the set $\mathrm{Ball}(\sqrt{n}) \setminus \boldsymbol{B}$ (see Lemma 3.5 for the precise statement).

[6] Recall that a Boolean function $f : \{0,1\}^n \to \{0,1\}$ is *monotone* if whenever $x, y \in \{0,1\}^n$ satisfy $x_i \le y_i$ for $i \in [n]$, we have $f(x) \le f(y)$.

[7] We remark that the original construction of [Naz03] differs from our construction in a number of ways: the distribution over random halfspaces is slightly different, and the body is not intersected with $\mathrm{Ball}(\sqrt{n})$. For technical reasons, our specific construction facilitates our lower bound arguments.

The original construction of Nazarov may be viewed as a Gaussian-space analogue of *Talagrand's random CNF formula* [Tal96] (see [DNS24] for a discussion of this connection). Talagrand's random CNF has been very useful in lower bounds for monotonicity testing over the Boolean hypercube, as demonstrated by [BB16, CWX17, CDL+24]. We use our modified Nazarov body to obtain new lower bounds for convexity testing, as described below.

1.2.2 One-Sided Adaptive Lower Bound Recall that a *one-sided* tester always outputs "accept" on convex sets and outputs "reject" on far-from-convex sets with probability at least 2/3 — this requirement implies that the tester rejects only if a certificate of non-convexity is found (i.e. a set of queries $x_1, \dots, x_t$ which lie in the body, and a query y in the convex hull of $x_1, \dots, x_t$ which is not in the body). In order to argue a q-query lower bound, it suffices to (1) design a distribution $\mathcal{D}_{\text{no}}$ over sets which are far-from-convex with high probability, and (2) argue that no q-query deterministic algorithm can find a certificate of non-convexity.

The key will be to "hide" the non-convexity within the uniquely violated sets of the Nazarov body. Consider working in $\mathbb{R}^{2n}$ and first randomly draw an n-dimensional "control subspace" $\boldsymbol{C}$ and the orthogonal n-dimensional "action subspace" $\boldsymbol{A}$; we embed the n-dimensional Nazarov body in the control subspace $\boldsymbol{C}$. A point $x \in \mathbb{R}^{2n}$ lies in our (non-convex) body iff:

- It has Euclidean norm at most $\sqrt{2n}$, and in addition, $x_{\boldsymbol{C}}$ (the projection onto the control subspace) has norm at most $\sqrt{n}$; and
- The point $x_{\boldsymbol{C}}$ lies within an n-dimensional Nazarov body that we randomly sample within the control subspace $\boldsymbol{C}$; *or*, for every $j \in [N]$ where $H_j(x_{\boldsymbol{C}}) = 0$, the projection $x_{\boldsymbol{A}}$ on the action subspace lies outside of a "strip" of width 1 along a randomly sampled direction $\boldsymbol{v}^j$ in the action subspace. (See Section 4.1.2).

Consider a line through a point $x \in \mathbb{R}^{2n}$ in direction $\boldsymbol{v}^j$, for $j \in [N]$ such that $x_{\boldsymbol{C}}$ lies in the uniquely violated region $\boldsymbol{U}_j$ and $x_{\boldsymbol{A}}$ lies inside the strip along $\boldsymbol{v}^j$ (and therefore *outside* our body). Then, as the line proceeds out from x in directions $\boldsymbol{v}^j$ and $-\boldsymbol{v}^j$, it remains in the uniquely violated region $\boldsymbol{U}_j$ (since $\boldsymbol{v}_j$ is orthogonal to $\boldsymbol{C}$) but exits the strip, thereby entering the body and exhibiting non-convexity. By design, the uniquely violated regions and the strips are large enough to constitute a constant fraction of the space, giving the desired distance to convexity (Lemma 4.1). Intuitively, detecting non-convexity is hard because the algorithm does not know $\boldsymbol{C}$, the halfspace $\boldsymbol{H}_j(\cdot)$, or the direction $\boldsymbol{v}^j$. In fact, we show that an algorithm which makes few queries cannot find, with probability at least 2/3, two points x, z outside the same halfspace $\boldsymbol{H}_j(\cdot)$ such that x lies inside and z outside the strip in direction $\boldsymbol{v}^j$.

Roughly speaking, the proof proceeds as follows. First, we show that, except with $o(1)$ probability, any two queries x, z which are far (at distance at least $1000\sqrt{q}n^{1/4}$) cannot lie outside the same halfspace $\boldsymbol{H}_j(\cdot)$ while having projections onto $\boldsymbol{C}$ with norm at most $\sqrt{n}$ (Lemma 4.4), and moreover it is extremely unlikely for a query to be falsified by more than q halfspaces (this follows from a calculation in Lemma 3.2). The argument is geometric in nature and is given in Section 4.5, and essentially argues that it is unlikely, since the algorithm does not know the control subspace $\boldsymbol{C}$ or the vector defining the halfspace, that two far-away queries happen to uniquely falsify the same halfspace.

On the other hand, consider the halfspaces which are falsified by some query (and notice there are most q^2 such halfspaces, since each query is falsified by at most q halfspaces). Since all such queries are within distance $1000\sqrt{q}n^{1/4}$ of each other, the projection of any two such queries onto the direction defining the strip is a segment of length $O(\sqrt{q}/n^{1/4})$ with high probability, and the precise location of this segment is uniform(-like, from Gaussian anti-concentration) (see Section 4.5.1). Therefore, the probability of any particular segment of length $O(\sqrt{q}/n^{1/4})$ which goes from inside to outside the strip of width 1 is roughly $O(\sqrt{q}/n^{1/4})$. We take a union bound over the q^2 possible halfspaces, each containing at most q queries which define segments which may "cross" the strip with probability $O(\sqrt{q}/n^{1/4})$, for a total probability of $O(q^{3.5}/n^{1/4})$. Since this must be at least 2/3 for the algorithm to succeed, this gives the $n^{\Omega(1)}$ lower bound.

1.2.3 Two-Sided Non-Adaptive Tolerant Lower Bound Continuing the analogy with monotonicity testing lower bounds, the proof of Theorem 1.2 is inspired by recent lower bounds on tolerant monotonicity testing, namely [PRW22] and the follow-up work of [CDL+24]. The basic idea of [PRW22] is to construct a family of functions by randomly partitioning the space of variables into *control variables* and *action variables*: if the control variables are

not balanced, i.e. there are more 1s than 0s (or vice-versa), then the function is trivially set to 1 (resp. to 0) both for $\boldsymbol{f} \sim \mathcal{D}_{\text{yes}}$ and for $\boldsymbol{f} \sim \mathcal{D}_{\text{no}}$. If the control variables are balanced, then, at a high level,

1. for $\boldsymbol{f} \sim \mathcal{D}_{\text{yes}}$ the function on the action variables is close to monotone;
2. for $\boldsymbol{f} \sim \mathcal{D}_{\text{no}}$ the function on the action variables is far from monotone.

Roughly speaking, the analysis in [PRW22] shows that unless the algorithm queries two points such that both these points (a) have the same setting of the control variables, and (b) the control variables are balanced, the algorithm cannot distinguish between $\boldsymbol{f} \sim \mathcal{D}_{\text{yes}}$ and $\boldsymbol{f} \sim \mathcal{D}_{\text{no}}$. As the control and action variables are partitioned at random, it turns out that satisfying both (a) and (b) is not possible for a non-adaptive algorithm unless the algorithm makes $2^{\Omega(\sqrt{n})}$ many queries. In particular, [PRW22] shows that distinguishing between functions which are $c_1/\sqrt{n}$-close to monotone versus $c_2/\sqrt{n}$-far from monotone (where $c_2 > c_1 > 0$) cannot be done with $2^{o(\sqrt{n})}$ queries.

The main modification in [CDL+24] vis-a-vis [PRW22] is the following: one can think of the balanced setting of the control variables in the construction described above as the "minimal satisfying assignments" of the Majority function. In [CDL+24], the Majority function is replaced by Talagrand's random monotone DNF [Tal96], a well-studied function in Boolean function analysis and related areas [MO03, OW07]. The specific properties of Talagrand's monotone DNF allows [CDL+24] to obtain a $2^{n^{1/4}}$ query lower bound for non-adaptive testers where the functions in $\mathcal{D}_{\text{yes}}$ are c_1-close to monotone and functions in $\mathcal{D}_{\text{no}}$ are c_2-far from monotone, where $c_2 > c_1$ are positive constants.

For Theorem 1.2, the goal is to obtain lower bounds for tolerant convexity testing rather than monotonicity testing. Towards that goal, let us assume that the ambient space is $\mathbb{R}^{n+1}$. We choose a random n-dimensional subspace $\boldsymbol{C}$ and think of it as the *control subspace*, and we view its one-dimensional orthogonal complement as the *action subspace* $\boldsymbol{A}$ (analogous to the notion of control and action variables in [PRW22, CDL+24]).[8] We embed the Nazarov body $\boldsymbol{B}$ (described earlier) in the control subspace. We define the $\mathcal{D}_{\text{yes}}$ and $\mathcal{D}_{\text{no}}$ distributions in analogy with [CDL+24], roughly as follows: for $x \in \mathbb{R}^{n+1}$,

1. If the projection x_C does not lie in the uniquely violated set of $\boldsymbol{B}$, then $\boldsymbol{f}(x)$ is defined the same way for $\boldsymbol{f} \sim \mathcal{D}_{\text{yes}}$ and $\boldsymbol{f} \sim \mathcal{D}_{\text{no}}$;
2. If the projection x_C lies in the uniquely violated set, then $\boldsymbol{f}(x)$ is set differently for $\boldsymbol{f} \sim \mathcal{D}_{\text{yes}}$ and $\boldsymbol{f} \sim \mathcal{D}_{\text{no}}$ (depending on the projection $x_{\boldsymbol{A}}$ to the action subspace). In particular, for $\boldsymbol{f} \sim \mathcal{D}_{\text{yes}}$, $\boldsymbol{f}$ is defined in such a way that $\boldsymbol{f}^{-1}(1)$ is close to a convex set, and for $\boldsymbol{f} \sim \mathcal{D}_{\text{no}}$, $\boldsymbol{f}$ is defined in such a way that $\boldsymbol{f}^{-1}(1)$ is far from every convex set. This crucially uses the fact that the Gaussian volume of $\boldsymbol{U}$ is "large" compared to the Gaussian volume of the set $\text{Ball}(\sqrt{n}) \setminus \boldsymbol{B}$, as mentioned in our earlier discussion of the Nazarov body.

At a high level, the indistinguishability argument showing that $q = 2^{\Omega(n^{1/4})}$ non-adaptive queries are required to distinguish $\boldsymbol{f} \sim \mathcal{D}_{\text{yes}}$ from $\boldsymbol{f} \sim \mathcal{D}_{\text{no}}$ is a case analysis based on the distance between any given pair of query vectors x and y (see Lemma 5.3 and Lemma 5.4), combined with a union bound over all $\binom{q}{2}$ possible pairs of query vectors. Roughly speaking, if $\|x - y\|$ is small, then the way that $\boldsymbol{f}$ depends on the projection to the action subspace makes it very unlikely to reveal a difference between $\boldsymbol{f} \sim \mathcal{D}_{\text{yes}}$ and $\boldsymbol{f} \sim \mathcal{D}_{\text{no}}$. On the other hand, if $\|x - y\|$ is large, then it is very unlikely for x and y to lie in the same set $\boldsymbol{U}_i$, which must be the case for the pair x, y to reveal a difference between $\boldsymbol{f} \sim \mathcal{D}_{\text{yes}}$ and $\boldsymbol{f} \sim \mathcal{D}_{\text{no}}$. There are many technical issues and geometric arguments required to carry out this rough plan, but when all the dust settles the argument gives a $2^{\Omega(n^{1/4})}$ lower bound for tolerant convexity testing.

1.2.4 Two-Sided Non-Adaptive Bound Our approach to prove Theorem 1.3 is inspired by the lower bounds of [CDST15] on non-adaptive monotonicity testing. As in most property testing lower bounds for non-adaptive algorithms, the high-level approach is to use Yao's principle; we follow [CDST15] in that we use a suitable high-dimensional central limit theorem as the key technical ingredient for establishing indistinguishability between

[8] We remark that in the one-sided adaptive lower bound described above, it would not have been possible to use a one-dimensional action subspace because an adaptive algorithm would be able to detect that "global structure," which is shared across all the $\boldsymbol{U}_i$'s; this is why the dimension of the action subspace $\boldsymbol{A}$ was n in the earlier construction, and there was a different random "action direction" $\boldsymbol{v}^j$ from $\boldsymbol{A}$ for each $j \in [N]$ in the earlier construction.

the yes- and no- distributions. In [CDST15] both the yes- and no- functions are linear threshold functions over $\{-1,+1\}^n$, but since any linear threshold function is trivially a convex set, the [CDST15] construction cannot be directly used to prove a convexity testing lower bound. Instead, in order to ensure that our no- functions are both indistinguishable from the yes- functions and are far from every convex set, we work with *degree-2 polynomial threshold functions (PTFs)* over $\mathbb{R}^n$ rather than linear threshold functions over $\{-1,+1\}^n$. At a high level, degree-2 PTFs of the form $\sum_i \lambda_i x_i^2$ where each λ_i is positive (note that any such PTF is a convex set) play the "yes-function" role that monotone LTFs play in the [CDST15] argument, and degree-2 PTFs of the form $\sum_i \lambda_i' x_i^2$ where a constant fraction of the λ_i''s are negative play the "no-function" role that far-from-monotone LTFs play in the [CDST15] argument. We show that having a constant fraction of the λ_i''s be negative is sufficient, in the context of our construction, to ensure that no-functions are far from convex, and we show that the multi-dimensional central limit theorem used in [CDST15] can be adapted to our context to establish indistinguishability and thereby prove the desired lower bound.

1.3 Related Work A number of earlier papers have considered different aspects of convexity testing. One strand of work deals with testing convexity of (real-valued) functions $f\colon [N] \to \mathbb{R}$, where convexity means the second derivative is positive.[9] This study was initiated by Parnas et al. [PRR03], and extended by Pallavor et al. [PRV18], who gave an improved result parameterized by the image size of the function being tested; by Blais et al. [BRY14b], who gave lower bounds on testing convexity of real-valued functions over the hypergrid $[N]^d$; and by Belovs et al. [BBB20], who gave upper and lower bounds on the number of queries required to test convexity of real-valued functions over various discrete domains including the discrete line, the "stripe" $[3] \times [N]$, and the hypergrid $[N]^d$. (See also the work of Berman et al. [BRY14a], who investigated a notion of "L_1-testing" real-valued functions over $[N]^d$ for convexity.)

A different body of work, which is closer to this paper, deals with testing convexity of *high-dimensional sets* (equivalently, Boolean indicator functions). The earliest work we are aware of along these lines is that of Rademacher and Vempala [RV05].[10] In their formulation, a body $K \subseteq \mathbb{R}^n$ is ε-far from being convex if $\mathrm{Leb}(K \triangle C) \geq \varepsilon \cdot \mathrm{Leb}(K)$ for every convex set C, where $\mathrm{Leb}(\cdot)$ denotes the Lebesgue volume (note that, in contrast, our model uses absolute volume under the Gaussian measure, rather than relative volume under the Lebesgue measure). Moreover, [RV05] allow the testing algorithm access to a black-box membership oracle (as in our model) as well as a "random sample" oracle which can generates a uniform random point from K (for testing with respect to relative measures, such an oracle is necessary). The main positive result of [RV05] is a $(cn/\varepsilon)^n$ sample- and query- algorithm for testing convexity in their model. [RV05] also give an exponential lower bound for a simple "line segment tester," which checks whether a line segment connecting two (uniformly random) points from the body is contained within the body. This lower bound was strengthened and extended to an exponential lower bound for a "convex hull tester" in recent work of Blais and Bommireddi [BB20]. We note that the negative results of [RV05] and [BB20], while they deal with natural and interesting candidate testing algorithms, only rule out very specific kinds of testers and do not provide lower bounds against general testing algorithms in their framework.

The most closely related work for us is the study of sample-based testing algorithms for convexity under the $N(0, I_n)$ distribution [CFSS17]. As was mentioned earlier, [CFSS17] gave a $2^{\tilde{O}(\sqrt{n})/\varepsilon^2}$-sample algorithm for convexity testing and showed that any sample-based tester must use $2^{\Omega(\sqrt{n})}$ samples; we remark that lower bounds for sample-based testers do not have any implications for query-based testing.[11] Finally, another closely related paper is the recent work of Blais et al. [BBH24] which gives nearly matching upper and lower bounds of $3^{\tilde{\Omega}(\sqrt{n})}$ queries for one-sided non-adaptive convexity testing over $\{-1,0,1\}^n$. [BBH24] cites the high-dimensional Gaussian testing problem as motivation for their study of the ternary cube, and asks "Can queries improve upon the bounds of [CFSS17, HY22] for testing convex sets with samples in $\mathbb{R}^n$ under the Gaussian distribution?" (Question 1.15 of [BBH24]). Our work makes progress on this question by establishing the first *lower* bounds for query-based testing under the Gaussian distribution.

[9] These works study discrete domains, where a discrete derivative is used.

[10] The study of convexity testing in two dimensions was initiated in earlier work of Raskhodnikova [Ras03] for the domain $[N]^2$, and has since been extended to sample-based testing [BMR16], testing over the continuous domain $[0,1]^2$ [BMR19], and tolerant testing [BMR22]; see also [BF18].

[11] [CFSS17] also gave a $2^{O(n\log(n/\varepsilon))}$-sample one-sided algorithm, which was generalized to testing under arbitrary product distributions by [HY22].

2 Preliminaries

We use boldfaced letters such as $\boldsymbol{x}, \boldsymbol{X}$, etc. to denote random variables (which may be real- or vector-valued; the intended type will be clear from the context). We write $\boldsymbol{x} \sim \mathcal{D}$ to indicate that the random variable $\boldsymbol{x}$ is distributed according to probability distribution $\mathcal{D}$. We will frequently identify a set $K \subseteq \mathbb{R}^n$ with its 0/1-valued indicator function, i.e., $K(x) = 1$ if $x \in K$ and $K(x) = 0$ otherwise. We write ln to denote natural logarithm and log to denote base-two logarithm.

2.1 Geometry We write $\mathbb{S}^{n-1}$ for the unit sphere in $\mathbb{R}^n$, i.e. $\mathbb{S}^{n-1} = \{x \in \mathbb{R}^n : \|x\| = 1\}$ where $\|x\|$ denotes the ℓ_2-norm of x. We write $\mathrm{Ball}(r) \subseteq \mathbb{R}^n$ to denote the ℓ_2-ball of radius r in $\mathbb{R}^n$, i.e.

$$\mathrm{Ball}(r) := \big\{x \in \mathbb{R}^n : \|x\| \le r\big\}.$$

We will frequently write $\mathrm{Ball} := \mathrm{Ball}(\sqrt{n})$. We recall the following standard bound on the volume of spherical caps (see e.g. Lemma 2.2 of [B+97]):

LEMMA 2.1. *For $0 \le \varepsilon < 1$, we have $\mathbf{Pr}\left[\boldsymbol{u}_1 \ge \varepsilon\right] \le e^{-n\varepsilon^2/2}$, where $\boldsymbol{u} \sim \mathbb{S}^{n-1}$, i.e. $\boldsymbol{u}$ is a Haar random vector drawn uniformly from the unit sphere $\mathbb{S}^{n-1}$.*

2.2 Gaussian and Chi-Squared Random Variables For $\mu \in \mathbb{R}^n$ and $\Sigma \in \mathbb{R}^{n\times n}$, we write $N(\mu, \Sigma)$ to denote the n-dimensional Gaussian distribution centered at μ and with covariance matrix Σ. In particular, identifying $0 \equiv 0^n$ and writing I_n for the $n \times n$ identity matrix, we will denote the n-dimensional standard Gaussian distribution by $N(0, I_n)$. We write $\mathrm{Vol}(K)$ to denote the Gaussian measure of a (Lebesgue measurable) set $K \subseteq \mathbb{R}^n$, i.e.

$$\mathrm{Vol}(K) := \mathop{\mathbf{Pr}}_{\boldsymbol{g}\sim N(0,I_n)}[\boldsymbol{g} \in K].$$

We recall the following standard tail bound on Gaussian random variables:

PROPOSITION 2.1. (THEOREM 1.2.6 OF [DUR19] OR EQUATION 2.58 OF [WAI15]) *Let $\Phi : \mathbb{R} \to (0,1)$ denote the cumulative density function of the (univariate) standard Gaussian distribution, i.e.*

$$\Phi(r) = \mathop{\mathbf{Pr}}_{\boldsymbol{g}\sim N(0,1)}[\boldsymbol{g} \le r].$$

Then for all $r > 0$, we have

$$\varphi(r)\left(\frac{1}{r} - \frac{1}{r^3}\right) \le 1 - \Phi(r) \le \varphi(r)\left(\frac{1}{r} - \frac{1}{r^3} + \frac{3}{r^5}\right)$$

where φ is the one-dimensional standard Gaussian density which is given by

$$\varphi(x) := \frac{1}{\sqrt{2\pi}} e^{-x^2/2}.$$

It is well known that if $\boldsymbol{g} \sim N(0, I_n)$, then $\|\boldsymbol{g}\|$ is distributed according to the chi distribution with n degrees of freedom, i.e. $\|\boldsymbol{g}\| \sim \chi(n)$. It is well known (see e.g. [Wik23]) that the mean of the $\chi^2(n)$ distribution is n, the median is $n(1 - \Theta(1/n))$, and for $n \ge 2$ the probability density function is everywhere at most 1. We note that an easy consequence of these facts is that the origin-centered ball $\mathrm{Ball}(\sqrt{n})$ of radius $\sqrt{n}$ in $\mathbb{R}^n$ has $\mathrm{Vol}(B(\sqrt{n})) = 1/2 + o(1)$.

We will require the following tail bound on $\chi^2(n)$ random variables:

PROPOSITION 2.2. (SECTION 4.1 OF [LM00]) *Suppose $\boldsymbol{y} \sim \chi^2(n)$. Then for any $t > 0$, we have*

$$\mathop{\mathbf{Pr}}_{\boldsymbol{y}\sim\chi^2(n)}\left[\boldsymbol{y} \ge n + 2\sqrt{nt} + 2t\right] \le e^{-t} \quad \text{and} \quad \mathop{\mathbf{Pr}}_{\boldsymbol{y}\sim\chi^2(n)}\left[\boldsymbol{y} \le n - 2\sqrt{nt}\right] \le e^{-t}.$$

2.3 Property Testing and Tolerant Property Testing Let $\mathcal{P}_{\text{conv}} := \mathcal{P}_{\text{conv}}(n)$ denote the class of convex subsets of $\mathbb{R}^n$, i.e.

$$\mathcal{P}_{\text{conv}} = \big\{L \subseteq \mathbb{R}^n : L \text{ is convex}\big\}.$$

Given a set $K \subseteq \mathbb{R}^n$, we define its *distance to convexity* as

$$\text{dist}(K, \mathcal{P}_{\text{conv}}) := \inf_{L \in \mathcal{P}_{\text{conv}}} \text{Vol}(K \triangle L)$$

where $K \triangle L = (K \setminus L) \cup (L \setminus K)$ denotes the symmetric difference of K and L. In particular, we will say that K is ε-close to (resp. ε-far from) a convex set if $\text{dist}(K, \mathcal{P}_{\text{conv}}) \leq \varepsilon$ (resp. $\geq \varepsilon$).

DEFINITION 1. (PROPERTY TESTERS AND TOLERANT PROPERTY TESTERS) *Let $\varepsilon, \varepsilon_1, \varepsilon_2 \in [0, 0.5]$ with $\varepsilon_1 < \varepsilon_2$. An algorithm $\mathcal{A}$ is an* ε-tester *for convexity if, given black-box query access to an unknown set $K \subseteq \mathbb{R}^n$, it has the following performance guarantee:*

- *If K is convex, then $\mathcal{A}$ outputs "accept" with probability at least $2/3$;*
- *If $\text{dist}(K, \mathcal{P}_{\text{conv}}) \geq \varepsilon$, then $\mathcal{A}$ outputs "reject" with probability at least $2/3$.*

An algorithm $\mathcal{A}$ is an $(\varepsilon_1, \varepsilon_2)$-tolerant tester *(or simply an $(\varepsilon_1, \varepsilon_2)$-*tester*) for convexity if it has the following performance guarantee:*

- *If $\text{dist}(K, \mathcal{P}_{\text{conv}}) \leq \varepsilon_1$, then $\mathcal{A}$ outputs "accept" with probability at least $2/3$;*
- *If $\text{dist}(K, \mathcal{P}_{\text{conv}}) \geq \varepsilon_2$, then $\mathcal{A}$ outputs "reject" with probability at least $2/3$.*

In particular, note that every ε-tester is a $(0, \varepsilon)$-tolerant tester.

Our query-complexity lower bounds for non-adaptive property testing algorithms are obtained via Yao's minimax principle [Yao77], which we recall below. (We remind the reader that an algorithm for the problem of $(\varepsilon_1, \varepsilon_2)$-tolerant testing is correct on an input function f provided that it outputs "yes" if f is ε_1-close to the property and outputs "no" if f is ε_2-far from the property; if the distance to the property is between ε_1 and ε_2 then the algorithm is correct regardless of what it outputs.)

THEOREM 2.1. (YAO'S PRINCIPLE) *To prove an $\Omega(q)$-query lower bound on the worst-case query complexity of any non-adaptive randomized testing algorithm, it suffices to give a distribution $\mathcal{D}$ on instances such that for any q-query non-adaptive* deterministic *algorithm $\mathcal{A}$, we have*

$$\Pr_{\boldsymbol{f} \sim \mathcal{D}} \big[\mathcal{A} \text{ is correct on } \boldsymbol{f}\big] \leq c.$$

where $0 \leq c < 1$ is a universal constant.

3 Nazarov's Body

Our constructions in Sections 4 and 5 will employ modifications of a probabilistic construction of a convex body due to Nazarov [Naz03]. Nazarov's randomized construction yields a convex set with asymptotically maximal Gaussian surface area [Bal93, Naz03], and modifications thereof have found applications in learning theory and polyhedral approximation [KOS08, DNS24].

DEFINITION 2. (NAZAROV'S BODY) *For $r, N > 0$, we write $\text{Naz}(r, N)$ to be the distribution over convex subsets of $\mathbb{R}^n$ where a draw $\boldsymbol{B} \sim \text{Naz}(r, N)$ is obtained as follows:*

1. *For $i \in [N]$, draw independent vectors $\boldsymbol{g}^i \sim N(0, I_n)$ and let $\boldsymbol{H}_i \subseteq \mathbb{R}^n$ denote the halfspace*

$$\boldsymbol{H}_i := \{x \in \mathbb{R}^n : x \cdot \boldsymbol{g}^i \leq r\}. \tag{3.1}$$

2. *Output the convex set $\boldsymbol{B} \subseteq \mathbb{R}^n$ where*

$$\boldsymbol{B} := \text{Ball}(\sqrt{n}) \cap \left(\bigcap_{i=1}^{N} \boldsymbol{H}_i\right).$$

Note that for any fixed $x \in \mathbb{R}^n$,

$$\begin{aligned}\mathop{\mathbf{Pr}}_{\boldsymbol{H}_i}[x \in \boldsymbol{H}_i] &= \mathop{\mathbf{Pr}}_{\boldsymbol{g}^i \sim N(0,I_n)}\left[x \cdot \boldsymbol{g}^i \le r\right] \\ &= \mathop{\mathbf{Pr}}_{\boldsymbol{g}^i \sim N(0,I_n)}\left[\sum_{j=1}^{n} x_j \boldsymbol{g}^i_j \le r\right] \\ &= \mathop{\mathbf{Pr}}_{\boldsymbol{g} \sim N(0,1)}\left[\boldsymbol{g} \le \frac{r}{\|x\|}\right] \\ &= \Phi\left(\frac{r}{\|x\|}\right) \end{aligned} \tag{3.2}$$

where $\Phi(\cdot)$ is the univariate Gaussian cumulative density function. Consequently, because of the independence of $\boldsymbol{g}^i$, we have

$$\mathop{\mathbf{Pr}}_{\boldsymbol{B} \sim \mathrm{Naz}(r,N)}[x \in \boldsymbol{B}] = \mathbf{1}\{\|x\| \le \sqrt{n}\} \cdot \Phi\left(\frac{r}{\|x\|}\right)^N. \tag{3.3}$$

Note that $\boldsymbol{B}$ can be also written as

$$\boldsymbol{B} = \mathrm{Ball}(\sqrt{n}) \setminus \bigcup_{i \in [N]} \Big(\mathrm{Ball}(\sqrt{n}) \setminus \boldsymbol{H}_i\Big).$$

For each $i \in [N]$, we define $\boldsymbol{F}_i$ (for "flap") to be points in $\mathrm{Ball}(\sqrt{n})$ which are falsified by $\boldsymbol{H}_i$, i.e.

$$\boldsymbol{F}_i := \mathrm{Ball}(\sqrt{n}) \setminus \boldsymbol{H}_i.$$

Given a non-empty $T \subseteq [N]$, we write $\boldsymbol{F}_T := \bigcap_{i \in T} \boldsymbol{F}_i$. We will be interested in points in $\mathrm{Ball}(\sqrt{n})$ that are falsified by a unique halfspace $\boldsymbol{H}_i$ and denote the set of such points as $\boldsymbol{U}_i$ (for "unique"):

$$\boldsymbol{U}_i := \boldsymbol{F}_i \setminus \bigcup_{j \ne i} \boldsymbol{F}_j.$$

3.1 Useful Estimates Suppose N satisfies $N = n^{\omega_n(1)}$; in both Sections 4 and 5 we will take $N = 2^{\sqrt{n}}$. Let $c_1 > 0$ be a parameter; in Section 4, we will set $c_1 = \ln 2 \pm O(1)/N$, and in Section 5 we will take c_1 to be a suitable small absolute constant.

Throughout this section we will take r to be the unique positive number such that

$$\Phi\left(\frac{r}{\sqrt{n}}\right) = 1 - \frac{c_1}{N}. \tag{3.4}$$

Gaussian tail bounds allow us to relate r and N:

LEMMA 3.1. *We have*

$$r = \sqrt{2n(1-o(1)) \ln\left(\frac{N}{c_1}\sqrt{\frac{n}{2\pi}}\right)}.$$

Proof. Note that because $N = \omega_n(1)$, it follows that $r = \omega(\sqrt{n})$; otherwise, note that $1 - \Phi\left(\frac{r}{\sqrt{n}}\right) = \Omega_n(1)$, contradicting Equation (3.4). Next, it follows from Proposition 2.1 and Equation (3.4) that

$$\left(\frac{\sqrt{n}}{r} - \left(\frac{\sqrt{n}}{r}\right)^3\right) \cdot \varphi\left(\frac{r}{\sqrt{n}}\right) \le \frac{c_1}{N} \le \frac{\sqrt{n}}{r} \cdot \varphi\left(\frac{r}{\sqrt{n}}\right). \tag{3.5}$$

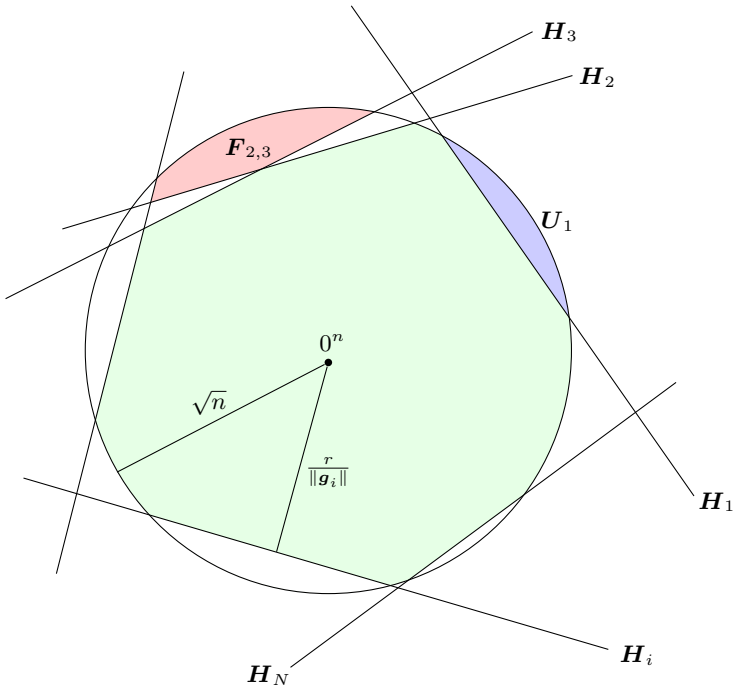

Figure 1: A depiction of $\boldsymbol{B}$ (in green) sampled from $\mathrm{Naz}(r, N)$.

The upper bound implies that

$$r \cdot \exp\left(\frac{r^2}{2n}\right) \le \frac{N}{c_1}\sqrt{\frac{n}{2\pi}}, \qquad \text{and so} \qquad \ln r + \frac{r^2}{2n} \le \ln\left(\frac{N}{c_1}\sqrt{\frac{n}{2\pi}}\right).$$

In particular, this implies that

$$r \le \sqrt{2n \ln\left(\frac{N}{c_1}\sqrt{\frac{n}{2\pi}}\right)}. \tag{3.6}$$

Next, note that the lower bound from Equation (3.5) implies that

$$\frac{N}{c_1}\sqrt{\frac{n}{2\pi}}\left(1 - \frac{n}{r^2}\right) \le r \exp\left(\frac{r^2}{2n}\right), \qquad \text{and so} \qquad \ln\left(\frac{(1-o(1))N}{c_1}\sqrt{\frac{n}{2\pi}}\right) \le \ln r + \frac{r^2}{2n}.$$

This in turn implies that

$$r \ge \sqrt{2n(1-o(1)) \cdot \ln\left(\frac{N}{c_1}\sqrt{\frac{n}{2\pi}}\right)} \tag{3.7}$$

The result follows from Equations (3.6) and (3.7). ∎

We need the following lemma which will be useful in analyzing our construction in Section 4.

LEMMA 3.2. *Let $x \in \mathbb{R}^n$ be a point with $\|x\| \le \sqrt{n}$. Then*

$$\Pr_{\boldsymbol{B} \sim \mathrm{Naz}(r,N)}\left[x \in \bigcup_{|T| \ge q} \boldsymbol{F}_T\right] \le \frac{c_1^q}{q!}.$$

for all $q \in [N]$.

Proof. Note that

$$\begin{aligned}
\Pr_{\boldsymbol{B}\sim\mathrm{Naz}(r,N)}\left[x \in \bigcup_{|T|\geq q} \boldsymbol{F}_T\right] &\leq \binom{N}{q} \Pr_{\boldsymbol{B}\sim\mathrm{Naz}(r,N)}\left[x \in \boldsymbol{F}_1 \cap \ldots \cap \boldsymbol{F}_q\right] \\
&\leq \frac{1}{q!}\left(N\left(1-\Phi\left(\frac{r}{\|x\|}\right)\right)\right)^q \\
&\leq \frac{1}{q!}\left(N\left(1-\Phi\left(\frac{r}{\sqrt{n}}\right)\right)\right)^q \\
&= \frac{c_1^q}{q!}
\end{aligned}$$

where the penultimate equality relies on Equation (3.4). ∎

We will also require a lower bound on the expected volume of $\bigsqcup_{i=1}^N \boldsymbol{U}_i$:

LEMMA 3.3. *For constant* $0 < c_1 \leq 0.9$ *and* $N = 2^{\sqrt{n}}$, *we have*

$$\mathop{\mathbf{E}}_{\boldsymbol{B}\sim\mathrm{Naz}(r,N)}\left[\mathrm{Vol}\left(\bigsqcup_{i=1}^N \boldsymbol{U}_i\right)\right] = \Omega(c_1).$$

Proof. Fix any $x \in \mathbb{R}^n$ and any $i \in [N]$. Note that

$$\Pr_{\boldsymbol{B}\sim\mathrm{Naz}(r,N)}[x \in \boldsymbol{U}_i] = \mathbf{1}\{\|x\| \leq \sqrt{n}\} \cdot \left(1-\Phi\left(\frac{r}{\|\boldsymbol{x}\|}\right)\right)\Phi\left(\frac{r}{\|\boldsymbol{x}\|}\right)^{N-1}. \tag{3.8}$$

It follows that

$$\begin{aligned}
\mathbf{E}\left[\mathrm{Vol}(\boldsymbol{U}_i)\right] &= \mathop{\mathbf{E}}_{\boldsymbol{x}\sim N(0,I_n)}\left[\Pr_{\boldsymbol{B}\sim\mathrm{Naz}(r,N)}[\boldsymbol{x} \in \boldsymbol{U}_i]\right] \\
&= \mathrm{Vol}\big(\mathrm{Ball}(\sqrt{n})\big) \mathop{\mathbf{E}}_{\boldsymbol{x}\sim N(0,I_n)}\left[\left(1-\Phi\left(\frac{r}{\|\boldsymbol{x}\|}\right)\right)\Phi\left(\frac{r}{\|\boldsymbol{x}\|}\right)^{N-1} \,\middle|\, \|\boldsymbol{x}\| \leq \sqrt{n}\right] \\
&\geq \frac{1}{2}\mathop{\mathbf{E}}_{\boldsymbol{x}\sim N(0,I_n)}\left[\left(1-\Phi\left(\frac{r}{\|\boldsymbol{x}\|}\right)\right)\Phi\left(\frac{r}{\sqrt{n}}\right)^{N-1} \,\middle|\, \|\boldsymbol{x}\| \leq \sqrt{n}\right] \\
&\geq \frac{1}{2}\left(1-\frac{c_1}{N}\right)^{N-1} \mathop{\mathbf{E}}_{\boldsymbol{x}\sim N(0,I_n)}\left[1-\Phi\left(\frac{r}{\|\boldsymbol{x}\|}\right) \,\middle|\, \|\boldsymbol{x}\| \leq \sqrt{n}\right] \\
&\geq \frac{1}{2}\left(1-c_1+\frac{c_1}{N}\right) \mathop{\mathbf{E}}_{\boldsymbol{x}\sim N(0,I_n)}\left[1-\Phi\left(\frac{r}{\|\boldsymbol{x}\|}\right) \,\middle|\, \|\boldsymbol{x}\| \leq \sqrt{n}\right] \qquad (3.9)
\end{aligned}$$

where the penultimate inequality follows from Equation (3.4) and the final inequality relies on the fact that $(1-y)^z \geq 1-yz$.

Next, at the cost of 0.01 probability mass (thanks to Proposition 2.2), we can assume that $\|\boldsymbol{x}\| \in [\sqrt{n}-10, \sqrt{n}]$. It follows that

$$\mathop{\mathbf{E}}_{\boldsymbol{x}\sim N(0,I_n)}\left[\left(1-\Phi\left(\frac{r}{\|\boldsymbol{x}\|}\right)\right)\middle|\sqrt{n}-10\le\|\boldsymbol{x}\|\le\sqrt{n}\right]\ge 0.99\cdot\left(1-\Phi\left(\frac{r}{\sqrt{n}-10}\right)\right).$$

Standard Gaussian tail bounds give that

$$\begin{aligned}1\ge\frac{1-\Phi\left(\frac{r}{\sqrt{n}-10}\right)}{1-\Phi\left(\frac{r}{\sqrt{n}}\right)}&\ge\frac{\left(\frac{\sqrt{n}-10}{r}-\frac{(\sqrt{n}-10)^3}{r^3}\right)\exp\left(\frac{-r^2}{2(\sqrt{n}-10)^2}\right)}{\frac{\sqrt{n}}{r}\exp\left(\frac{-r^2}{2n}\right)}\\&\ge(1-o(1))\cdot\exp\left(\frac{r^2(100-20\sqrt{n})}{2n(\sqrt{n}-10)^2}\right)\\&=\Theta(1),\end{aligned}$$

(Proposition 2.1)

where the last line uses our bounds on r from Lemma 3.1 and our bounds on c_1 from the statement of the current lemma. Putting everything together and recalling that $c_1 < 0.9$, we get that for n large enough,

$$\mathbf{E}\left[\mathrm{Vol}(\boldsymbol{U}_i)\right]\ge\Omega\left(1-\Phi\left(\frac{r}{\sqrt{n}}\right)\right)=\Omega\left(\frac{c_1}{N}\right) \tag{3.10}$$

thanks to Equation (3.4). Consequently, we have

$$\mathbf{E}\left[\mathrm{Vol}\left(\bigsqcup_{i\in[N]}\boldsymbol{U}_i\right)\right]\ge\Omega(c_1), \tag{3.11}$$

completing the proof. ∎

Next we show that the volume of $\bigsqcup_i \boldsymbol{U}_i$ is highly concentrated:

LEMMA 3.4. *Suppose $N = 2^{\sqrt{n}}$. With probability at least $1-o(1)$, we have*

$$\mathrm{Vol}\left(\bigsqcup_{i\in[N]}\boldsymbol{U}_i\right)\ge 0.9\cdot\mathop{\mathbf{E}}_{\boldsymbol{B}\sim\mathrm{Naz}(r,N)}\left[\mathrm{Vol}\left(\bigsqcup_{i\in[N]}\boldsymbol{U}_i\right)\right].$$

Proof. Let $\mathrm{Naz}^*(r,N)$ be the same distribution as $\mathrm{Naz}(r,N)$ except that when drawing $\boldsymbol{B}$, each $\boldsymbol{g}^i$ is drawn from $N(0,I_n)$ conditioning on $\|\boldsymbol{g}^i\| = \sqrt{n}\pm 10n^{1/4}$ (instead of just drawing $\boldsymbol{g}^i\sim N(0,I_n)$). Recall $N = 2^{\sqrt{n}}$. By Proposition 2.2, the probability of $\|\boldsymbol{g}^i\|\notin[\sqrt{n}-10n^{1/4},\sqrt{n}+10n^{1/4}]$ for some $i\in[N]$ is at most $o(1)$. As a result, we have

$$\mathop{\mathbf{E}}_{\boldsymbol{B}\sim\mathrm{Naz}^*(r,N)}\left[\mathrm{Vol}\left(\bigsqcup_{i\in[N]}\boldsymbol{U}_i\right)\right]\ge\mathop{\mathbf{E}}_{\boldsymbol{B}\sim\mathrm{Naz}(r,N)}\left[\mathrm{Vol}\left(\bigsqcup_{i\in[N]}\boldsymbol{U}_i\right)\right]-o(1).$$

Moreover, it suffices to show that when $\boldsymbol{B}\sim\mathrm{Naz}^*(r,N)$, we have

$$\mathrm{Vol}\left(\bigsqcup_{i\in[N]}\boldsymbol{U}_i\right)\ge 0.99\cdot\mathop{\mathbf{E}}_{\boldsymbol{B}\sim\mathrm{Naz}^*(r,N)}\left[\mathrm{Vol}\left(\bigsqcup_{i\in[N]}\boldsymbol{U}_i\right)\right]. \tag{3.12}$$

with probability at least $1-o(1)$. To this end, we recall McDiarmid's inequality:

THEOREM 3.1. (MCDIARMID BOUND [McD89]) *Let $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_S$ be independent random variables taking values in a set Ω. Let $G\colon \Omega^S \to \mathbb{R}$ be such that for all $i \in [S]$ we have*

$$\left|G(x_1, \ldots, x_S) - G(x_1, \ldots, x_{i-1}, x_i', x_{i+1}, \ldots, x_S)\right| \le c_i$$

for all $x_1, \ldots, x_S$ and x_i' in Ω. Let $\mu = \mathbf{E}[G(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_S)]$. Then for all $\tau > 0$, we have

$$\mathbf{Pr}\left[G(\boldsymbol{X}_1, \ldots, \boldsymbol{X}_S) < \mu - \tau\right] < \exp\left(-\frac{\tau^2}{\sum_{i\in[S]} c_i^2}\right).$$

We will take $S = N$, $\boldsymbol{X}_i$ to be the halfspaces $\boldsymbol{H}_i$ and $G(\cdot)$ to be the volume of $\sqcup_{i\in[N]}\boldsymbol{U}_i$, as we draw $\boldsymbol{B} \sim \mathrm{Naz}^*(r, N)$. Given the way $\boldsymbol{g}^i$ is drawn in $\boldsymbol{B} \sim \mathrm{Naz}^*(r, N)$, the volume of each $\boldsymbol{H}_i$ is always at least (using $r \ge \sqrt{2n^{3/2}(1 - o(1))}$ by Lemma 3.1)

$$\Phi\left(\frac{r}{\sqrt{n} + 10n^{1/4}}\right) \ge 1 - e^{-(1-o(1))\sqrt{n}},$$

from which we have $c_i \le e^{-(1-o(1))\sqrt{n}}$. As a consequence,

$$\sum_{i\in[N]} c_i^2 \le N \cdot e^{-(1-o(1))\sqrt{n}} = e^{-\Omega(\sqrt{n})}.$$

It follows from McDiarmid that Equation (3.12) holds with probability at least $1 - o(1)$. ∎

Finally, the following lemma will allow us to obtain bounds on the distance to convexity of the "yes"- and "no"-distributions in Section 5:

LEMMA 3.5. *For r satisfying Equation (3.4), we have*

$$\mathop{\mathbf{E}}_{\boldsymbol{B}\sim\mathrm{Naz}(r,N)}\left[\mathrm{Vol}\left(\bigsqcup_{i\in[N]}\boldsymbol{U}_i\right)\right] \ge \left(\frac{2}{c_1} - 2\right)\mathop{\mathbf{E}}_{\boldsymbol{B}\sim\mathrm{Naz}(r,N)}\left[\mathrm{Vol}\left(\bigcup_{|T|\ge 2}\boldsymbol{F}_T\right)\right].$$

Proof. Fix $x \in \mathbb{R}^n$ and $i \in [N]$. Recall Equation (3.8). On the other hand, we have

$$\begin{aligned}
\mathop{\mathbf{Pr}}_{\boldsymbol{B}\sim\mathrm{Naz}(r,N)}\left[x \in \bigcup_{|T|\ge 2}\boldsymbol{F}_T\right] &\le \binom{N}{2}\mathop{\mathbf{Pr}}_{\boldsymbol{B}\sim\mathrm{Naz}(r,N)}\left[x \in \boldsymbol{F}_1 \cap \boldsymbol{F}_2\right] \\
&= \binom{N}{2}\mathop{\mathbf{Pr}}_{\boldsymbol{B}\sim\mathrm{Naz}(r,N)}\left[x \in \mathrm{Ball}(\sqrt{n}) \cap (\mathbb{R}^n \setminus \boldsymbol{H}_1) \cap (\mathbb{R}^n \setminus \boldsymbol{H}_2)\right] \\
&\le \frac{N^2}{2} \cdot \mathbf{1}\{\|x\| \le \sqrt{n}\} \cdot \left(1 - \Phi\left(\frac{r}{\|x\|}\right)\right)^2 \qquad (3.13)
\end{aligned}$$

where we once again used Equation (3.2). It follows from Equations (3.8) and (3.13) that for $x \in \mathrm{Ball}(\sqrt{n})$ (i.e $\|x\| \le \sqrt{n}$), we have

$$\begin{aligned}
\frac{N \cdot \mathbf{Pr}\left[x \in \boldsymbol{U}_i\right]}{\mathbf{Pr}[x \in \bigcup_{|T|\ge 2}\boldsymbol{F}_T]} &\ge \left(\frac{2}{N}\right)\Phi\left(\frac{r}{\|x\|}\right)^{N-1}\left(1 - \Phi\left(\frac{r}{\|x\|}\right)\right)^{-1} \\
&\ge \left(\frac{2}{N}\right)\Phi\left(\frac{r}{\sqrt{n}}\right)^{N-1}\left(1 - \Phi\left(\frac{r}{\sqrt{n}}\right)\right)^{-1} \qquad (3.14) \\
&= \left(\frac{2}{c_1}\right)\left(1 - \frac{c_1}{N}\right)^{N-1} \qquad (3.15) \\
&\ge \left(\frac{2}{c_1}\right)\left(1 - c_1 + \frac{c_1}{N}\right) \qquad (3.16) \\
&> \frac{2}{c_1} - 2
\end{aligned}$$

where Equation (3.14) relies on the fact that $\|x\| \le \sqrt{n}$ and $\Phi(\cdot)$ being increasing, Equation (3.15) relies on our definition of r from Equation (3.4), and Equation (3.16) relies on Bernoulli's inequality: $(1-y)^z \ge 1 - yz$. (Note that for x with $\|x\| > \sqrt{n}$, we have $\mathbf{Pr}[x \in \boldsymbol{U}_i] = \mathbf{Pr}[x \in \bigcup_{|T|\ge 2} \boldsymbol{F}_T] = 0$.)

To conclude, we have

$$
\begin{aligned}
N \cdot \mathop{\mathbf{E}}_{\boldsymbol{B}\sim \mathrm{Naz}(r,N)} \left[\mathrm{Vol}(\boldsymbol{U}_i)\right] &= \mathop{\mathbf{E}}_{\boldsymbol{B}\sim \mathrm{Naz}(r,N)} \left[N \cdot \mathop{\mathbf{Pr}}_{\boldsymbol{x}\sim N(0,I_n)} [\boldsymbol{x} \in \boldsymbol{U}_i] \right] \\
&= \mathop{\mathbf{E}}_{\boldsymbol{x}\sim N(0,I_n)} \left[N \cdot \mathop{\mathbf{Pr}}_{\boldsymbol{B}\sim \mathrm{Naz}(r,N)} [\boldsymbol{x} \in \boldsymbol{U}_i] \right] \\
&\ge \left(\frac{2}{c_1} - 2\right) \mathop{\mathbf{E}}_{\boldsymbol{x}\sim N(0,I_n)} \left[\mathop{\mathbf{Pr}}_{\boldsymbol{B}\sim \mathrm{Naz}(r,N)} \left[\boldsymbol{x} \in \bigcup_{|T|\ge 2} \boldsymbol{F}_T \right] \right] \qquad (3.17) \\
&= \left(\frac{2}{c_1} - 2\right) \mathop{\mathbf{E}}_{\boldsymbol{B}\sim \mathrm{Naz}(r,N)} \left[\mathop{\mathbf{Pr}}_{\boldsymbol{x}\sim N(0,I_n)} \left[\boldsymbol{x} \in \bigcup_{|T|\ge 2} \boldsymbol{F}_T \right] \right] \\
&= \left(\frac{2}{c_1} - 2\right) \mathop{\mathbf{E}}_{\boldsymbol{B}\sim \mathrm{Naz}(r,N)} \left[\mathrm{Vol}\left(\bigcup_{|T|\ge 2} \boldsymbol{F}_T \right) \right]
\end{aligned}
$$

where Equation (3.17) follows from the earlier calculation, completing the proof. ∎

4 One-Sided Adaptive Lower Bound

For this section, it will be most convenient for us to work over $\mathbb{R}^{2n}$. Let us restate Theorem 1.1 in this setting:

THEOREM 4.1. (ONE-SIDED ADAPTIVE LOWER BOUND, RESTATED) *For some absolute constant $\varepsilon > 0$, any one-sided ε-tester for convexity over $N(0, I_{2n})$ (which may be adaptive) must use $n^{\Omega(1)}$ queries.*

At a high level, the proof of Theorem 1.1 works by (1) first defining a distribution $\mathcal{D}_{\mathrm{no}}$ of "no-functions" (Boolean-valued functions over $\mathbb{R}^{2n}$, or equivalently, subsets of $\mathbb{R}^{2n}$), and showing that an $\Omega(1)$ fraction of draws from $\mathcal{D}_{\mathrm{no}}$ yield sets which are $\Omega(1)$-far from convex; and (2) then arguing that for a suitable absolute constant $c > 0$, any n^c-query algorithm (even an adaptive one) has only an $o(1)$ probability of querying a set of points whose labels are inconsistent with every convex set in $\mathbb{R}^{2n}$. In the next subsection we describe the distribution $\mathcal{D}_{\mathrm{no}}$.

4.1 The distribution $\mathcal{D}_{\mathrm{no}}$ of far-from-convex sets

4.1.1 Setup We will see that every function f in the support of $\mathcal{D}_{\mathrm{no}}$ outputs 0 on every input point $x \in \mathbb{R}^{2n}$ with $\|x\| > \sqrt{2n}$. To describe how f behaves within the $\sqrt{2n}$-ball, denoted by

$$\mathrm{Ball}(\sqrt{2n}) := \left\{x \in \mathbb{R}^{2n} : \|x\| \le \sqrt{2n}\right\},$$

we require some more setup.

The "control subspace," the "action subspace," and the Nazarov body. Let $\boldsymbol{C}$ be a Haar random n-dimensional subspace of $\mathbb{R}^{2n}$; we call $\boldsymbol{C}$ the *control subspace*. Let $\boldsymbol{A}$ be the orthogonal complement of $\boldsymbol{C}$ (which is also an n-dimensional subspace); we call $\boldsymbol{A}$ the "action subspace." Given a vector $x \in \mathbb{R}^n$, we write $x_{\boldsymbol{C}}$ to denote the orthogonal projection of x onto $\boldsymbol{C}$ and we write $x_{\boldsymbol{A}}$ to denote the orthogonal projection of x onto $\boldsymbol{A}$, so every vector satisfies $x = x_{\boldsymbol{A}} + x_{\boldsymbol{C}}$.

Fix $N := 2^{\sqrt{n}}$ (we assume without loss of generality that n is a perfect square, so N is an integer). For this choice of N, let $\boldsymbol{B} \sim \mathrm{Naz}(r, N, \boldsymbol{C})$ where $\mathrm{Naz}(r, N, \boldsymbol{C})$ is as defined in Definition 2 but with the n-dimensional control subspace $\boldsymbol{C}$ playing the role of $\mathbb{R}^n$. (We emphasize that $\boldsymbol{B} \sim \mathrm{Naz}(r, N, \boldsymbol{C})$ is a subset of $\mathbb{R}^{2n}$ which is an

"n-subspace junta," meaning that for any $x \in \mathbb{R}^{2n}$, membership of x in $\boldsymbol{B}$ depends only on $x_{\boldsymbol{C}}$.) We take r to be the unique positive number such that

$$\Phi\left(\frac{r}{\sqrt{n}}\right)^N = \frac{1}{2}.$$

In other words, we choose r to be the unique value such that any point x with $\|x_{\boldsymbol{C}}\| = \sqrt{n}$ has probability $1/2$ of being in $\boldsymbol{B} \sim \mathrm{Naz}(r, N, \boldsymbol{C})$ (cf. Equation (3.3)). Note that

$$\Phi\left(\frac{r}{\sqrt{n}}\right) = \left(\frac{1}{2}\right)^{\frac{1}{N}} = 1 - \frac{c_1}{N} \quad \text{for a value } c_1 = \ln 2 \pm \frac{O(1)}{N}$$

by the Taylor expansion of $e^{-\ln(2)/N}$ and setting of r (Lemma 3.1).

The "action directions." For each $i \in [N]$, draw a random vector $\boldsymbol{v}^i$ from the standard Normal distribution $N(0, I_n)$ over the n-dimensional action subspace $\boldsymbol{A}$ (independent of everything else). We say that $\boldsymbol{v}^i$ is the *action direction* for the i-th flap $\boldsymbol{F}_i$ of the Nazarov body $\boldsymbol{B}$. We note that for every pair $i, j \in [N]$, the vector $\boldsymbol{g}^i$ defining the i-th halfspace $\boldsymbol{H}^i$ of the Nazarov body is orthogonal to the vector $\boldsymbol{v}^j$ (because $\boldsymbol{g}^i \in \boldsymbol{C}$ and $\boldsymbol{v}^j \in \boldsymbol{A}$).

4.1.2 The distribution $\mathcal{D}_{\mathrm{no}}$ For a fixed setting of the control subspace C and the (orthogonal) action subspace A, of $\vec{H} := (H_1, \ldots, H_N)$ (which also specifies B and F_i's) and of $\vec{v} := (v^1, \ldots, v^N)$, we define the function $f_{C,A,\vec{H},\vec{v}} : \mathbb{R}^{2n} \to \{0,1\}$ as follows:

$$f_{C,A,\vec{H},\vec{v}}(x) = \begin{cases} 0 & x \notin \mathrm{Ball}(\sqrt{2n}) \text{ or } \|x_C\| > \sqrt{n}; \\ 1 & x \in \mathrm{Ball}(\sqrt{2n}) \text{ and } x_C \in B; \\ \bigwedge_{j \in T} \mathbf{1}\left[\langle v^j, x\rangle \notin [-\frac{\sqrt{n}}{2}, \frac{\sqrt{n}}{2}]\right] & x \in \mathrm{Ball}(\sqrt{2n}) \text{ and } x_C \in F_T \text{ for some } \emptyset \neq T \subseteq [N]. \end{cases}$$

A random function $\boldsymbol{f} \sim \mathcal{D}_{\mathrm{no}}$ is drawn as follows: first we draw a Haar random n-dimensional subspace $\boldsymbol{C}$; then $\boldsymbol{A}$ is taken to be the n-dimensional (Haar random) orthogonal complement of $\boldsymbol{C}$; then we draw $\boldsymbol{B} \sim \mathrm{Naz}(r, N, \boldsymbol{C})$ (which gives a draw of $\vec{\boldsymbol{H}}$ as in Equation (3.1)); then we draw $\vec{\boldsymbol{v}} = (\boldsymbol{v}^1, \ldots, \boldsymbol{v}^N)$ from $\boldsymbol{A}$ as described above; then we set the function $\boldsymbol{f}$ to be $f_{\boldsymbol{C},\boldsymbol{A},\vec{\boldsymbol{H}},\vec{\boldsymbol{v}}}$.

4.2 Sets in $\mathcal{D}_{\mathrm{no}}$ are far from convex We need a constant fraction of the no-functions to be constant-far from convex. This is given by the following lemma:

LEMMA 4.1. *With probability $\Omega(1)$ over a draw of $\boldsymbol{f} \sim \mathcal{D}_{\mathrm{no}}$, we have that $\mathrm{Vol}(\boldsymbol{f} \triangle g) = \Omega(1)$ for every $g : \mathbb{R}^{2n} \to \{0,1\}$ that is the indicator function of a convex set in $\mathbb{R}^{2n}$.*

We require a few definitions. Define $\mathrm{ThinShell} := \{x \in \mathbb{R}^{2n} : \sqrt{2n} - 2 \leq \|x\| \leq \sqrt{2n} - 1\}$. Given an outcome of $\boldsymbol{f} \sim \mathcal{D}_{\mathrm{no}}$ (which determines the $\boldsymbol{g}^i$'s, $\boldsymbol{v}^i$'s, $\boldsymbol{F}^i$'s and $\boldsymbol{U}_i$'s), for $i \in [N]$ define $\boldsymbol{U} := \sqcup_{i \in [N]} \boldsymbol{U}_i$. Define $p := \mathbf{E}_{\boldsymbol{f} \sim \mathcal{D}_{\mathrm{no}}}[\mathrm{Vol}[\boldsymbol{U} \cap \mathrm{ThinShell}]]$.

LEMMA 4.2. $p = \Omega(1) \implies$ *Lemma 4.1.*

Proof. If $p = \Omega(1)$ then $\mathbf{Pr}_{\boldsymbol{f}}[\mathrm{Vol}[\boldsymbol{U} \cap \mathrm{ThinShell}] = \Omega(1)] = \Omega(1)$. We view the draw of $\boldsymbol{f}$ as taking place in two stages: in the first one $\boldsymbol{C}$, $\boldsymbol{A}$, and $\vec{\boldsymbol{g}} = (\boldsymbol{g}^1, \ldots, \boldsymbol{g}^N)$ are drawn, and in the second one $\vec{\boldsymbol{v}} = (\boldsymbol{v}^1, \ldots, \boldsymbol{v}^N)$ is drawn. Observe that the set $\boldsymbol{U}$ depends only on the first stage. Say that any outcome of the first stage for which $\mathrm{Vol}[\boldsymbol{U} \cap \mathrm{ThinShell}] = \Omega(1)$ holds is a *good* outcome of the first stage, so an $\Omega(1)$ fraction of outcomes of the first stage are good.

Fix any good outcome $C, A, \vec{g}$ of the first stage (note that this fixes $U_1, \ldots, U_N$ and hence U), and consider a draw of $\boldsymbol{x} \sim N(0, I_{2n})$. We have the following claim:

CLAIM 3. *For a suitable absolute constant $a > 0$, we have $\mathbf{Pr}_{\boldsymbol{x} \sim N(0,I_{2n})}[\boldsymbol{x} \in U \cap \mathrm{ThinShell}$ and $\|\boldsymbol{x}_C\| \in [\sqrt{n} - a, \sqrt{n}]] = \Omega(1)$.*

Proof. Since we have fixed a good outcome $C, A, \vec{g}$ of the first stage, we have that $\mathbf{Pr}_{\boldsymbol{x}\sim N(0,I_{2n})}[\boldsymbol{x} \in U \cap \text{ThinShell}] \geq c$ for some absolute constant $c > 0$. Moreover, every outcome of $\boldsymbol{x} \in U \cap \text{ThinShell}$ has $\|\boldsymbol{x}_C\| \leq \sqrt{n}$, since U is a subset of B. So to prove the claim we need only show that $\mathbf{Pr}_{\boldsymbol{x}\sim N(0,I_{2n})}[\|\boldsymbol{x}_C\| < \sqrt{n} - a] \leq c/2$.

We first observe that by standard bounds on the chi-square distribution (Proposition 2.2), we have that $\mathbf{Pr}_{\boldsymbol{x}\sim N(0,I_{2n})}[\|\boldsymbol{x}\| \notin [\sqrt{2n}-a', \sqrt{2n}+a']] \leq c/4$ for a suitable constant a'. So fix any length $\ell \in [\sqrt{2n}-a', \sqrt{2n}+a']$. Fix any vector $z \in \mathbb{R}^{2n}$ with $\|z\| = \ell$; by the rotational symmetry of the $N(0, I_{2n})$ distribution and the rotational symmetry of drawing a Haar random n-dimensional subspace $\boldsymbol{C}$ of $\mathbb{R}^{2n}$, the distribution of $\|\boldsymbol{x}_C\|$ conditioned on $\|\boldsymbol{x}\| = \ell$ is the same as the distribution of $\|z_{\boldsymbol{C}}\|$ where $\boldsymbol{C}$ is a Haar random n-dimensional subspace $\boldsymbol{C}$ of $\mathbb{R}^{2n}$. A routine application of the Johnson-Lindenstrauss theorem (see e.g. Theorem 5.3.1 of [Ver18]) gives us that $\mathbf{Pr}_{\boldsymbol{C}}[\|z_{\boldsymbol{C}}\| < \sqrt{n} - a] \leq c/4$, for a suitable choice of the constant a. So $\mathbf{Pr}_{\boldsymbol{x}\sim N(0,I_{2n})}[\|\boldsymbol{x}_C\| < \sqrt{n} - a] \leq c/2$ as required, and the claim is proved. ∎

Now, given an x that lies in $U \cap \text{ThinShell}$ and has $\|x_C\| \in [\sqrt{n} - a, \sqrt{n}]$, consider an outcome of the second stage, i.e. the draw of $\vec{\boldsymbol{v}}$; note that this draw completes the draw of $\boldsymbol{f} \sim \mathcal{D}_{\text{no}}$. Define the vectors

$$x^+ := x + \frac{\boldsymbol{v}^i}{\|\boldsymbol{v}^i\|}, \qquad x^- := x - \frac{\boldsymbol{v}^i}{\|\boldsymbol{v}^i\|}.$$

Let us say that an outcome of $\vec{\boldsymbol{v}}$ for which $\boldsymbol{f}(x) = 0$, $\boldsymbol{f}(x^+) = 1$, $\boldsymbol{f}(x^-) = 1$ is a *fine outcome of* $\vec{\boldsymbol{v}}$ *for* x. We will use the following claim:

CLAIM 4. *For any fixed x that lies in $U \cap \text{ThinShell}$ and has $\|x_C\| \in [\sqrt{n} - a, \sqrt{n}]$, we have $\mathbf{Pr}_{\vec{\boldsymbol{v}}}[\vec{\boldsymbol{v}}$ is fine for $x] = \Omega(1)$.*

Proof. Since $x \in U_i \cap \text{ThinShell}$ for some i, it must be the case that also $x^+, x^- \in U_i$ (because every possible outcome of $\boldsymbol{v}^i$ is orthogonal to every possible outcome of g^j for every $j \in [N]$). So $\vec{\boldsymbol{v}}$ is fine if and only if

$$\left\langle \boldsymbol{v}^i, x^- \right\rangle < -\frac{\sqrt{n}}{2} \leq \left\langle \boldsymbol{v}^i, x \right\rangle \leq \frac{\sqrt{n}}{2} < \left\langle \boldsymbol{v}^i, x^+ \right\rangle, \quad \text{or equivalently,}$$

$$\left\langle \boldsymbol{v}^i, x \right\rangle - \|\boldsymbol{v}^i\| < -\frac{\sqrt{n}}{2} \leq \left\langle \boldsymbol{v}^i, x \right\rangle \leq \frac{\sqrt{n}}{2} < \left\langle \boldsymbol{v}^i, x \right\rangle + \|\boldsymbol{v}^i\|. \tag{4.18}$$

Since $x \in \text{ThinShell}$ we have $\sqrt{2n} - 2 \leq \|x\| \leq \sqrt{2n} - 1$, i.e.

$$2n - 4\sqrt{2n} + 4 \leq \|x\|^2 = \|x_C\|^2 + \|x_A\|^2 \leq 2n - 2\sqrt{2n} + 1,$$

and since $\|x_C\| \in [\sqrt{n} - a, \sqrt{n}]$ we have that $n - 2a\sqrt{n} + a^2 \leq \|x_C\|^2 \leq n$. So

$$n - 4\sqrt{2n} + 4 \leq \|x_A\|^2 \leq n - 2\sqrt{2n} + 2a\sqrt{n} + 1 - a^2. \tag{4.19}$$

Now since $\boldsymbol{v}^i$ is drawn from a standard $N(0, I_n)$ distribution over the subspace A, a routine calculation using (i) Equation (4.19); (ii) the fact that $\|\boldsymbol{v}^i - \langle \boldsymbol{v}^i, x\rangle \frac{x}{\|x\|}\|^2$ and $\langle \boldsymbol{v}^i, x\rangle$ are independent and are distributed as a draw from the $\chi^2(n-1)$ distribution and a draw from $N(0, \|x_A\|^2)$ respectively; and (iii) the fact that a draw from the $\chi^2(n-1)$ distribution takes value $n(1 \pm o(1))$ except with vanishingly small probability, gives that Equation (4.18) holds with $\Omega(1)$ probability. ∎

As an immediate consequence of Claim 4, we get that an $\Omega(1)$ fraction of outcomes of $\vec{v}$ are such that

$$\mathop{\mathbf{Pr}}_{\boldsymbol{x}\sim N(0,I_{2n})}\left[v \text{ is fine for } \boldsymbol{x} \mid \boldsymbol{x} \in U \cap \text{ThinShell} \;\&\; \|\boldsymbol{x}_C\| \in \left[\sqrt{n} - a, \sqrt{n}\right]\right] = \Omega(1). \tag{4.20}$$

Fix any outcome $\vec{v}$ of $\vec{\boldsymbol{v}}$ for which Equation (4.20) holds. To conclude the proof of Lemma 4.1, we observe that since $x \in U_i$ implies that x^+, x^- are also in U_i, it follows that any $z \in \mathbb{R}^n$ can participate in at most three triples of the form (x, x^-, x^+), so the maximum possible degree of overlap across all of the triples is at most a factor of

three. Moreover, for any $x \in \text{ThinShell}$, it holds that $\sqrt{2n} - 3 \le \|x\| - 1 \le \|x^+\|, \|x^-\| \le \|x\| + 1 \le \sqrt{2n}$, and hence the pdf of the $\chi^2(2n)$ distribution is within a constant factor on each of the three inputs $\|x\|, \|x^+\|$ and $\|x^-\|$ (so the $N(0, I_{2n})$ Gaussian's pdf is within a constant factor on each of the three inputs x, x^+, x^-). Combining this with Claim 3, we get that for an $\Omega(1)$ fraction of outcomes of $\boldsymbol{f} \sim \mathcal{D}_{\text{no}}$, the value of $\boldsymbol{f}$ needs to be altered on at least an $\Omega(1)$ fraction of inputs drawn from $N(0, I_n)$ in order to "repair" all of the violating triples (x, x^+, x^-) for which $x \in U \cap \text{ThinShell}$ and $\|x_C\| \in [\sqrt{n} - a, \sqrt{n}]$. This gives Lemma 4.2. ∎

Proof. [Lemma 4.1] To prove Lemma 4.1 it remains only to show that $p = \Omega(1)$, i.e. to show that

$$\mathop{\mathbf{Pr}}_{\boldsymbol{f} \sim \mathcal{D}_{\text{no}}, \boldsymbol{x} \sim N(0, I_{2n})}[x \in (\boldsymbol{U} \cap \text{ThinShell})] = \Omega(1). \tag{4.21}$$

We first observe that we have $\mathbf{Pr}_{\boldsymbol{x} \sim N(0, I_{2n})}[\boldsymbol{x} \in \text{ThinShell}] = \Omega(1)$. Fix any outcome $x \in \text{ThinShell}$. Consider a draw of the Haar random n-dimensional subspace $\boldsymbol{C}$ of $\mathbb{R}^{2n}$ which is part of the draw of $\boldsymbol{f} \sim \mathcal{D}_{\text{no}}$. Similar to the proof of Claim 3, using $\|x\| \in [\sqrt{2n} - 2, \sqrt{2n} - 1]$ the Johnson-Lindenstrauss theorem gives that $\mathbf{Pr}_{\boldsymbol{C}}[\|x_{\boldsymbol{C}}\| \in [\sqrt{n} - 1, \sqrt{n}]] = \Omega(1)$. Finally, fix any outcome C of $\boldsymbol{C}$ such that $\|x_C\| \in [\sqrt{n} - 1, \sqrt{n}]$, and consider the "completion" of the draw of $\boldsymbol{f} \sim \mathcal{D}_{\text{no}}$ (i.e. the draw of $\boldsymbol{B} \sim \text{Naz}(r, N, C)$ which induces an outcome of $\boldsymbol{U}$). We have

$$\mathop{\mathbf{Pr}}_{\boldsymbol{f}}[x \in \boldsymbol{U}] = N \cdot \left(1 - \Phi\left(\frac{r}{\|x_C\|}\right)\right) \Phi\left(\frac{r}{\|x_C\|}\right)^{N-1}, \tag{4.22}$$

so to complete the proof of Lemma 4.1 it suffices to show that (5.29) $= \Omega(1)$. We have

$$\Phi\left(\frac{r}{\|x_C\|}\right)^{N-1} \ge \Phi\left(\frac{r}{\sqrt{n}}\right)^{N-1} = \left(1 - \frac{c_1}{N}\right)^{N-1} = \Omega(1),$$

where the first equality is Equation (3.4) and the second is because $c_1 = \Theta(1)$. Similar to the proof of Lemma 3.3, we have

$$\begin{aligned}
\frac{1 - \Phi\left(\frac{r}{\|x_C\|}\right)}{1 - \Phi\left(\frac{r}{\sqrt{n}}\right)} \ge \frac{1 - \Phi\left(\frac{r}{\sqrt{n}-1}\right)}{1 - \Phi\left(\frac{r}{\sqrt{n}}\right)} &\ge \frac{\left(\frac{\sqrt{n}-1}{r} - \frac{(\sqrt{n}-1)^3}{r^3}\right) \exp\left(\frac{-r^2}{2(\sqrt{n}-1)^2}\right)}{\frac{\sqrt{n}}{r} \exp\left(\frac{-r^2}{2n}\right)} \\
&\ge (1 - o(1)) \exp\left(\frac{r^2(1 - 2\sqrt{n})}{2n(\sqrt{n}-1)^2}\right) \\
&= \Theta(1), \quad \text{using Lemma 3.1.}
\end{aligned}$$

(Proposition 2.1)

So

$$N \cdot \left(1 - \Phi\left(\frac{r}{\|x_C\|}\right)\right) \ge N \cdot \Theta(1) \cdot \left(1 - \Phi\left(\frac{r}{\|x_C\|}\right)\right) = N \cdot \Theta(1) \cdot \frac{c_1}{N} = \Omega(1),$$

where the first equality is by Equation (3.4). This concludes the proof of Lemma 4.1. ∎

4.3 Proof of Theorem 1.1

Definition 5. (One-sided Adaptive Algorithms as Binary Trees) *Fix $n, q \in \mathbb{N}$. A q-query one-sided deterministic algorithm,* Alg, *for testing convexity in $\mathbb{R}^{2n}$ is specified by a rooted binary tree of depth q where each node contains the following information:*

- *Each node v which is a not a leaf contains a query vector $x_v \in \mathbb{R}^{2n}$, as well as two out-going edges, one labeled 0 and one labeled 1, to nodes which we label $v(0)$ and $v(1)$, respectively.*
- *Each leaf node v contains an output o_v which is set to* "accept" *or* "reject." *Let Q_1 (or Q_0) denote the set of points queried along the path that are labelled 1 (or 0, respectively). Then o_v is set to be "reject" if and only if $Q_0 \cap \text{conv}(Q_1) \neq \emptyset$.*

By adding nodes which repeat the queries, we may assume, without loss of generality, that the depth of every leaf of the tree is exactly q.

A q-query deterministic algorithm Alg executes on a function $f\colon \mathbb{R}^{2n} \to \{0,1\}$ by taking the natural root-to-leaf path given by following the function values which the oracle returns at the queries within each of the nodes. In particular, we will make repeated use of the following definitions which capture the execution of the algorithm Alg on a function f:

- The node v^0 is the root of the tree, which is the starting point of the root-to-leaf path. Then, the nodes $v^1, \dots, v^q$ indicate the root-to-leaf path generated by executing the algorithm on the function f. In particular, at time step $t \in \{0, \dots, q-1\}$, we have $v^{t+1} = v^t(f(x_{v^t}))$
- The set Q^0 is defined to be $\emptyset$, and for $t \in \{0, \dots, q-1\}$ the set Q^{t+1} is defined to be $Q^t \cup \{x_{v^t}\} \subset \mathbb{R}^n$. Thus Q^{t+1} is the set of vectors that are queried at time steps prior to $t+1$.

Once the algorithm reaches the leaf node v^q, the algorithm outputs o_{v^q}, and we will refer to $\mathrm{Alg}(f)$ as the output ("accept" or "reject") produced by the algorithm. It is trivial to see that since any q-query deterministic algorithm corresponds to a tree of depth q, the total number of query vectors $x_v \in \mathbb{R}^{2n}$ across all nodes of the tree is at most 2^q. Our goal is to show that, if Alg is a q-query deterministic algorithm which makes *one-sided error*, then

$$\mathop{\mathbf{Pr}}_{\boldsymbol{f} \sim \mathcal{D}_{\mathrm{no}}} \left[\mathrm{Alg}(\boldsymbol{f}) = \text{"reject"}\right] = o(1). \tag{4.23}$$

Recall that implicit in a fixed function f in the support of $\mathcal{D}_{\mathrm{no}}$ are the control and action subspaces $C, A \subset \mathbb{R}^{2n}$, as well as the vectors $g^1, \dots, g^N \in C$ and $v^1, \dots, v^N \in A$, and that $g^1, \dots, g^N$ define B, H_i and F_i regions. In order to simplify our notation, we will often refer to a subset of the queries $\tilde{Q}^k$ for any $k \le q$ whose norm on the control subspace is bounded,

$$\tilde{Q}^k = \left\{ x \in Q^k : \|x_C\| \le \sqrt{n} \right\}.$$

Toward showing the above upper bound, we define two important events (which will depend on the draw $\boldsymbol{f} \sim \mathcal{D}_{\mathrm{no}}$).

Definition 6. *Given* Alg *and a function f from $\mathcal{D}_{\mathrm{no}}$, we consider the following three events:*

- *$\mathcal{E}_1(f)$: This event occurs if at the end of the execution of* Alg *on f, every point $x \in \tilde{Q}^q$ lies in at most q flaps, and for every flap F_i with $\tilde{Q}^q \cap F_i \neq \emptyset$,*

$$\|x - y\| \le 1000\sqrt{q} n^{1/4} \quad \text{for all } x, y \in \tilde{Q}^q \cap F_i. \tag{4.24}$$

- *$\mathcal{E}_2(f)$: This event occurs if at the end of the execution of* Alg *on f, for every flap F_i with $\tilde{Q}^q \cap F_i \neq \emptyset$ and every $x, y \in \tilde{Q}^q \cap F_i$, we have*

$$\mathbf{1}\left[\langle v^i, x\rangle \notin [-\sqrt{n}/2, \sqrt{n}/2]\right] = \mathbf{1}\left[\langle v^i, y\rangle \notin [-\sqrt{n}/2, \sqrt{n}/2]\right].$$

Theorem 1.1 follows immediately from the following three lemmas:

Lemma 4.3. *Let* Alg *be a one-sided, deterministic, q-query algorithm for testing convexity. Then, if* $\mathrm{Alg}(f)$ *outputs "reject," the event $\overline{\mathcal{E}_2(f)}$ occurred.*

Lemma 4.4. *Let* Alg *be a one-sided, deterministic, q-query algorithm. Then,*

$$\mathop{\mathbf{Pr}}_{\boldsymbol{f} \sim \mathcal{D}_{\mathrm{no}}} \left[\mathcal{E}_1(\boldsymbol{f})\right] \ge 1 - o(1).$$

Lemma 4.5. *Let* Alg *be a one-sided, deterministic, q-query algorithm, where $q \le n^{0.05}$. Then,*

$$\mathop{\mathbf{Pr}}_{\boldsymbol{f} \sim \mathcal{D}_{\mathrm{no}}} \left[\overline{\mathcal{E}_2(\boldsymbol{f})} \cap \mathcal{E}_1(\boldsymbol{f})\right] \le o(1).$$

Proof. [Proof of Theorem 1.1 Assuming Lemmas 4.3 to 4.5] We upper bound the expression

$$\mathop{\mathbf{Pr}}_{\boldsymbol{f} \sim \mathcal{D}_{\mathrm{no}}} \left[\mathrm{Alg}(\boldsymbol{f}) = \text{"accept"}\right] \overset{(4.3)}{\ge} \mathop{\mathbf{Pr}}_{\boldsymbol{f} \sim \mathcal{D}_{\mathrm{no}}} \left[\mathcal{E}_2(\boldsymbol{f})\right] \ge \mathop{\mathbf{Pr}}_{\boldsymbol{f} \sim \mathcal{D}_{\mathrm{no}}} \left[\mathcal{E}_1(\boldsymbol{f})\right] - \mathop{\mathbf{Pr}}_{\boldsymbol{f} \sim \mathcal{D}_{\mathrm{no}}} \left[\overline{\mathcal{E}_2(\boldsymbol{f})} \cap \mathcal{E}_1(\boldsymbol{f})\right] \ge 1 - o(1)$$

using Lemmas 4.4 and 4.5. ∎

4.4 Proof of Lemma 4.3 Since Alg is a q-query deterministic algorithm which has one-sided error, in order for the algorithm to output "reject," the set Q^q queried by the root-to-leaf path obtained by executing Alg on f must contain $x_1, \dots, x_\ell, y \in Q^q$ satisfying

$$y \in \mathrm{conv}(x_1, \dots, x_\ell), \quad f(y) = 0, \quad \text{and} \quad f(x_1) = \cdots = f(x_\ell) = 1.$$

In particular, from $y \in \mathrm{conv}(x_1, \dots, x_\ell)$, we must have that, for any vector $u \in \mathbb{R}^{2n}$, there exists a $j \in [\ell]$ such that $\langle x_j, u\rangle \geq \langle y, u\rangle$. This implies that:

- We must have that all $x_1, \dots, x_\ell$ satisfy $\|(x_i)_C\|_2 \leq \sqrt{n}$, and $\|y_C\|_2 \leq \sqrt{n}$, and this means these vectors lie in $\tilde{Q}^q$. The part of $\|(x_i)_C\|_2 \leq \sqrt{n}$ follows trivially from $f(x_1) = \cdots = f(x_\ell) = 1$. On the other hand, if $\|y_C\|_2 > \sqrt{n}$, letting $u \in C$ be the unit vector $u = y_C/\|y_C\|_2$, there exists an x_j with

  $$\|(x_j)_C\|_2 \geq \langle x_j, u\rangle \geq \langle y, u\rangle = \|y_C\|_2 > \sqrt{n},$$

  and hence $f(x_j) = 0$, which would be a contradiction with $f(x_j) = 1$.

- We must have $y \notin B$ since $f(y) = 0$. As a result, there is a nonempty T such that $y \in F_T$. In addition, $f(y) = 0$ implies that there exists an $i \in T$ such that $y \in F_i$ but

  $$\mathbf{1}\Big[\langle v^i, y\rangle \notin [-\sqrt{n}/2, \sqrt{n}/2]\Big] = 0.$$

  Given that $y \in F_i$, setting $u = g^i$, there exists an x_j such that $\langle x_j, g^i\rangle > \langle y, g^i\rangle \geq r$ and thus, $x_j \in F_i$. It follows from $f(x_j) = 1$ and the construction that

  $$\mathbf{1}\Big[\langle v^i, x_j\rangle \notin [-\sqrt{n}/2, \sqrt{n}/2]\Big] = 1.$$

This concludes the proof using i, y and x_j.

4.5 Proof of Lemma 4.4 To prove Lemma 4.4, we introduce five new, easy-to-analyze events $\mathcal{E}_{1,1}, \mathcal{E}_{1,2}, \mathcal{E}_{1,3}$, $\mathcal{E}_{1,4}$ and $\mathcal{E}_{1,5}$, show that each happens with probability at least $1 - o(1)$, and that $\mathcal{E}_{1,1} \cap \mathcal{E}_{1,2} \cap \mathcal{E}_{1,3} \cap \mathcal{E}_{1,4} \cap \mathcal{E}_{1,5}$ implies $\mathcal{E}_1$. For the n-dimensional subspace $C \subset \mathbb{R}^{2n}$ (in particular, the control subspace for f), we denote $\mathrm{Shell}(C) := \{x \in \mathbb{R}^{2n} : \sqrt{n} - 100q \leq \|x_C\|_2 \leq \sqrt{n}\}$, where x_C denotes the orthogonal projection of x onto the subspace C.

- $\mathcal{E}_{1,1}(f)$: This event occurs if no query x in Alg with $\|x_C\|_2 \leq \sqrt{n}$ lies in $\bigcup_{|T|\geq q} F_T$ defined by f;

- $\mathcal{E}_{1,2}(f)$: This event occurs if no query x in Alg satisfies $x \notin \mathrm{Shell}(C)$ and $x \notin B$ (or equivalently, $\|x_C\|_2 < \sqrt{n} - 100q$ and $x \in F_i$ for some $i \in [N]$);

- $\mathcal{E}_{1,3}(f)$: This event occurs if no query x in Alg with $\|x_C\|_2 \leq \sqrt{n}$ has

  $$\langle x, g^i\rangle \geq r + 100qn^{1/4}, \quad \text{for some } i \in [N];$$

- $\mathcal{E}_{1,4}(f)$: This event *does not* occur if there exist $i \in [N]$ and two queries x, z in Alg where (i) x_C and z_C are not scalar multiples of each other, $z_C = (1+a)x_C + by$ denotes the unique decomposition with $x_C \perp y$, $\|y\|_2 = 1$ and $b > 0$, such that $x \in F_i$ and $|\langle y, g^i\rangle| \geq 100\sqrt{q}$.

- $\mathcal{E}_{1,5}(f)$: The event occurs whenever every pair $x, y \in$ Alg satisfy $\|x - y\|_2 \leq 2\|(x-y)_C\|_2$.

We first prove that $\mathcal{E}_1(f)$ is implied by the five events together. Then, we show that each of the events holds individually with probability $1 - o(1)$. By a union bound over the five events, this gives Lemma 4.4.

LEMMA 4.6. $\mathcal{E}_{1,1}(f) \cap \mathcal{E}_{1,2}(f) \cap \mathcal{E}_{1,3}(f) \cap \mathcal{E}_{1,4}(f) \cap \mathcal{E}_{1,5}(f)$ *implies* $\mathcal{E}_1(f)$.

Proof. Recall that $\tilde{Q}^q$ denotes the set of (at most q) queries made by Alg when running on f whose orthogonal projections onto C each have norm at most $\sqrt{n}$. First $\mathcal{E}_{1,1}(f)$ implies that the number of "nonempty" flaps $i \in [N]$, i.e. flaps F_i that have $\tilde{Q}^q \cap F_i \neq \emptyset$, is at most q^2. Fix any nonempty flap F_i and any two points $x, z \in \tilde{Q}^q \cap F_i$. First consider the case that x_C and z_C are scalar multiples of each other. Note that we have $x, z \in \text{Shell}(C)$ by $\mathcal{E}_{1,2}(f)$ and thus, $\|(x-z)_C\|_2 \leq 100q$ (since they are scalar multiples of each other). By $\mathcal{E}_{1,5}(f)$, $\|x-z\|_2 \leq 200q$, which is consistent with the requirement of $\mathcal{E}_1(f)$ since $q = o(\sqrt{q}n^{1/4})$.

So consider the case when x_C, z_C are not scalar multiples of each other, and let $z_C = (1+a)x_C + by$ be the unique decomposition with $x_C \perp y$ and $y \in C$ with $\|y\|_2 = 1$ and $b > 0$. Let $\alpha := \|(x-z)_C\|_2^2 = a^2\|x_C\|_2^2 + b^2$. Our goal is to establish that

$$\alpha \leq 250000q\sqrt{n}, \tag{4.25}$$

so that we may use $\mathcal{E}_{1,5}(f)$ to deduce that (4.24) holds for x and z.

We have $\|z_C\|_2^2 = (1+a)^2\|x_C\|_2^2 + b^2$. Given that $\|z_C\|_2 \leq \sqrt{n}$,

$$(1+2a)\|x_C\|_2^2 + \alpha = \|z_C\|_2^2 \leq n.$$

By $\mathcal{E}_{1,2}(f)$, we have $x \in \text{Shell}(C)$ and thus, $\|x_C\|_2 \geq \sqrt{n} - 100q$. Plugging this in, we have

$$(1+2a)(n - 200q\sqrt{n} + 10000q^2) + \alpha \leq n,$$

or equivalently,

$$\alpha \leq -2an + (1+2a)(200q\sqrt{n} - 10000q^2) \leq 200q\sqrt{n} + a(-2n + 400q\sqrt{n} - 20000q^2). \tag{4.26}$$

Let's consider two cases:

Case 1: $a \geq -200q/\sqrt{n}$. We have from Equation (4.26) (note that the coefficient of a is negative and is a value larger than $-2n$)

$$\|(x-z)_C\|_2^2 = \alpha \leq 200q\sqrt{n} + 2n \cdot \frac{200q}{\sqrt{n}} = 600q\sqrt{n},$$

and we get Equation (4.25).

Case 2: $a < -200q/\sqrt{n}$. In this case, using $r \leq \langle z, g^i\rangle, \langle x, g^i\rangle \leq r + 100qn^{1/4}$ (where the first inequality is because $x, z \in F_i$ and the second is from $\mathcal{E}_{1,3}(f)$) gives

$$r \leq \overbrace{(1+a)\cdot\langle x, g^i\rangle + b\cdot\langle y, g^i\rangle}^{=\langle z, g^i\rangle} \leq a\cdot\langle x, g^i\rangle + \overbrace{r + 100qn^{1/4}}^{\text{b/c } \langle x, g^i\rangle \leq r + 100qn^{1/4}} + b\cdot \overbrace{(100\sqrt{q})}^{\text{b/c } |\langle y, g^i\rangle| \leq 100\sqrt{q}}$$

so (recall that $a < -200q/\sqrt{n}$ is negative and $-a$ is positive)

$$b \geq \frac{-a\cdot\langle x, g^i\rangle - 100qn^{1/4}}{100\sqrt{q}} \geq \frac{-ar}{200\sqrt{q}}.$$

Recalling that $\|z_C\|_2^2 = (1+a)^2\|x_C\|_2^2 + b^2 \leq n$ and that $\|x_C\|_2^2 \geq n - 200q\sqrt{n} + 10000q^2$, we get

$$n \geq (1+2a+a^2)(n - 200q\sqrt{n}) + \frac{a^2r^2}{40000q} \geq (1+2a)(n - 200q\sqrt{n}) + \frac{a^2r^2}{40000q}$$

and hence,

$$a^2 \cdot \frac{r^2}{40000q} \leq 200q\sqrt{n} - 2a(n - 200q\sqrt{n}).$$

Recalling that $a < 0$, dividing through by $-a$ we get

$$(-a)\cdot\frac{r^2}{40000q} \leq \frac{200q\sqrt{n}}{-a} + 2(n - 200q\sqrt{n}) \overbrace{\leq}^{(\text{using } -a \geq 200q/\sqrt{n})} 3n.$$

So we have

$$0 < -a \le \frac{120000qn}{r^2} = \frac{60000q}{\sqrt{n}} \cdot (1 + o(1)),$$

by the setting of r in Lemma 3.1. Recalling Equation (6.54), we get

$$\alpha \le 200q\sqrt{n} - 2an \le 200q\sqrt{n} + 240000q\sqrt{n},$$

as was to be shown. ∎

Event $\mathcal{E}_{1,1}(f)$. We now show that with probability at least $1 - o(1)$ over the draw of $\boldsymbol{f} \sim \mathcal{D}_{\text{no}}$, all 2^q queries specified by Alg avoid the region which is the intersection of at least q flaps. Consider any fixed query x and fix any setting of the control subspace $C \subset \mathbb{R}^{2n}$ with $\|x_C\|_2 \le \sqrt{n}$. Using Lemma 3.2 (and the fact C is isomorphic to $\mathbb{R}^n$),

$$\mathop{\mathbf{Pr}}_{\boldsymbol{B} \sim \text{Naz}(r,N,C)} \left[x \in \bigcup_{|T| \ge q} \boldsymbol{F}_T \right] \le \frac{c_1^q}{q!},$$

so that a union bound over 2^q queries gives $(2c_1)^q/q! = o(1)$ for large q.

Event $\mathcal{E}_{1,2}(f)$. Similarly to above, we proceed by a union bound over all 2^q queries. We consider a fixed control subspace C and we let x be a query with $\|x_C\|_2 < \sqrt{n} - 100q$, so

$$\begin{aligned}
\mathop{\mathbf{Pr}}_{\boldsymbol{B} \sim \text{Naz}(r,N,C)} \left[\exists i \in [N] : x \in \boldsymbol{F}_i \right] &\le N \cdot \left(1 - \Phi\left(\frac{r}{\sqrt{n} - 100q} \right) \right) \\
&\le N \cdot \left(1 - \Phi\left(\frac{r}{\sqrt{n}} \left(1 + \frac{100q}{\sqrt{n}} \right) \right) \right) \\
&\le N \cdot \frac{\sqrt{n}}{r} \cdot \exp\left(-\frac{r^2}{2n} \left(1 + \frac{200q}{\sqrt{n}} \right) \right) \\
&= N \cdot \frac{\sqrt{n}}{r} \cdot \exp\left(-\frac{r^2}{2n} \right) \cdot \exp\left(-\frac{100r^2 q}{n^{3/2}} \right) \le 2c_1 \cdot \exp\left(-10q \right),
\end{aligned}$$

by the setting of r from Lemma 3.1. The desired claim then follows from a union bound over all 2^q queries.

Event $\mathcal{E}_{1,3}(f)$. Consider any query x, and consider a fixed setting of the control subspace C with $\|x_C\|_2 \le \sqrt{n}$. Then,

$$\begin{aligned}
\mathop{\mathbf{Pr}}_{\boldsymbol{B} \sim \text{Naz}(r,N)} \left[\exists i \in [N] : \langle x, \boldsymbol{g}^i \rangle \ge r + 100qn^{1/4} \right] &\le N \left(1 - \Phi\left(\frac{r + 100qn^{1/4}}{\sqrt{n}} \right) \right) \\
&\le N \left(1 - \Phi\left(\frac{r}{\sqrt{n}} \left(1 + \frac{100qn^{1/4}}{r} \right) \right) \right) \le o(2^{-q}),
\end{aligned}$$

where the computation proceeds similarly to $\mathcal{E}_{1,2}(f)$.

Event $\mathcal{E}_{1,4}(f)$. For a fixed control subspace C, we may consider two arbitrary queries x, z among the set of all 2^q queries with $\|x_C\|_2, \|z_C\|_2 \le \sqrt{n}$. This gives 2^{2q} possible settings of the unit vector y which is orthogonal to x_C. In order for the event to fail, there must exists some $i \in [N]$ where $\langle \boldsymbol{g}_i, x_C \rangle \ge r$ and $|\langle \boldsymbol{g}_i, y \rangle| \ge 100\sqrt{q}$. Furthermore, since x_C and y are orthogonal, these two events are independent:

$$\mathop{\mathbf{Pr}}_{\boldsymbol{g}^i} \left[\langle \boldsymbol{g}_i, x_C \rangle \ge r \wedge |\langle \boldsymbol{g}_i, y \rangle| \ge 100\sqrt{q} \right] \le \frac{c_1}{N} \cdot e^{-100^2 q/2},$$

hence, we may take a union bound over all $i \in [N]$ and all 2^{2q} pairs of vectors x and z.

Event $\mathcal{E}_{1,5}(f)$. Finally, consider any two vectors x and y which are queries among the 2^q possible queries in Alg. The Johnson-Lindenstrauss lemma (see Theorem 5.3.1 in [Ver18]) says that a random n-dimensional subspace $\boldsymbol{C}$ of $\mathbb{R}^{2n}$ will satisfy $\|(x-y)_{\boldsymbol{C}}\|_2 \geq (1/\sqrt{2}-\varepsilon)\|x-y\|_2$ except with probability $\exp\left(-\Omega(\varepsilon^2 n)\right)$. Thus, for large enough n, $\|x-y\|_2 \leq 2\|(x-y)_{\boldsymbol{C}}\|_2$ except with probability $\exp\left(-\Omega(n)\right)$, and since $q \ll n$, we may union bound over all 2^{2q} pairs of queries in Alg.

4.5.1 Proof of Lemma 4.5

For $\overline{\mathcal{E}_2(f)} \cap \mathcal{E}_1(f)$ to happen, there must exist a level $k \in [q]$ such that

- After the the first $k-1$ queries $\tilde{Q} = \tilde{Q}^{k-1}$, $\mathcal{E}_1(f)$ holds, i.e., the number of flaps F_i with $\tilde{Q} \cap F_i \neq \emptyset$ is at most q^2. In every such F_i, every two points in $\tilde{Q} \cap F_i$ have distance at most $1000\sqrt{q}n^{1/4}$ and share the same value of
$$\mathbf{1}\Big[\langle v^i, x\rangle \notin [-\sqrt{n}/2, \sqrt{n}/2]\Big],$$
which we denote by $b_i \in \{0,1\}$.
- Let y be the k-th query. There exists an i such that $\tilde{Q} \cap F_i \neq \emptyset$ and $y \in F_i$ such that
$$\|y-x\|_2 \leq 1000\sqrt{q}n^{1/4} \tag{4.27}$$
for all $x \in \tilde{Q} \cap F_i$ (the number of such i is at most q) but
$$\mathbf{1}\Big[\langle v^i, x\rangle \notin [-\sqrt{n}/2, \sqrt{n}/2]\Big] \neq b_i.$$

We prove below that when $q \leq n^{0.05}$, the probability of the event above for a fixed k is $o(1/q)$ and thus, Lemma 4.5 follows by applying a union bound on k. This follows from a union bound on all $i \in [N]$ such that $\tilde{Q} \cap F_i \neq \emptyset$ and every $x \in \tilde{Q} \cap F_i$ satisfies (Equation (4.27)), taking X (or z) below as $\tilde{Q} \cap F_i$ (or y, respectively) projected on the space orthogonal to g^i and b as b_i.

LEMMA 4.7. *Let $b \in \{0,1\}$, and let X be a set of at most q points in $\mathrm{Ball}(\sqrt{2n})$ and $y \in \mathrm{Ball}(\sqrt{2n})$. Suppose that every pair $x \in X$ and y satisfy*
$$\|(x-y)_A\|_2 \leq 1000\sqrt{q}n^{1/4}.$$
Then over the draw of $\boldsymbol{v} \sim N(0, I_n)$ in A, the probability of $\mathbf{1}[\langle \boldsymbol{v}, y\rangle \notin [-\sqrt{n}/2, \sqrt{n}/2]] \neq b$ conditioning on the event that $\mathbf{1}[\langle \boldsymbol{v}, x\rangle \notin [-\sqrt{n}/2, \sqrt{n}/2]] = b$ for all $x \in X$ is at most $O(\sqrt{q}\log n/n^{1/4})$.

Before proving Lemma 4.7, we show how it implies Lemma 4.5 from a union bound. In particular, we have concluded that

$$\Pr_{\boldsymbol{f}\sim\mathcal{D}_{\mathrm{no}}}\left[\overline{\mathcal{E}_2(\boldsymbol{f})} \cap \mathcal{E}_1(\boldsymbol{f})\right] \leq \overbrace{q}^{\text{u.b over } k\in[q]} \times \overbrace{q^2}^{\text{u.b over } i} \times O\left(\sqrt{q}\log n/n^{1/4}\right) = O(q^{3.5}\log n/n^{1/4}).$$

Proof. [Proof of Lemma 4.7] Fix a point $x^* \in X$. We show that (1) the probability of $\mathbf{1}[\langle \boldsymbol{v}, x^*\rangle \notin [-\sqrt{n}/2, \sqrt{n}/2]] = b$ for all x is at least $\Omega(1)$; and (2) the probability of
$$\mathbf{1}[\langle \boldsymbol{v}, y\rangle \notin [-\sqrt{n}/2, \sqrt{n}/2]] \neq \mathbf{1}[\langle \boldsymbol{v}, x^*\rangle \notin [-\sqrt{n}/2, \sqrt{n}/2]]$$
is at most $O(\sqrt{q}\log n/n^{1/4})$. The lemma then follows.

To analyze (2), we consider any $0 < \gamma \leq \sqrt{n}/4$ and any sign $\xi \in \{-1,1\}$. We have that a draw of a Gaussian $\boldsymbol{v}$ lying in the action subspace A satisfies

$$\begin{aligned}\Pr_{\boldsymbol{v}}\left[\langle \boldsymbol{v}, x^*\rangle \in [\xi\sqrt{n}/2-\gamma, \xi\sqrt{n}/2+\gamma]\right] &= \Pr_{\boldsymbol{g}\sim N(0,1)}\left[\boldsymbol{g} \in \left[\frac{\xi\sqrt{n}}{2\|x_A^*\|_2} - \frac{\gamma}{\|x_A^*\|_2}, \frac{\xi\sqrt{n}}{2\|x_A^*\|_2} + \frac{\gamma}{\|x_A^*\|_2}\right]\right] \\ &\leq \min_{\tau>0}\left\{\frac{2\gamma}{\tau}, \frac{4\tau}{n}\cdot e^{-n/8\tau^2}\right\},\end{aligned}$$

where we have used Gaussian anti-concentration (to conclude it does not lie within an interval of width $2\gamma/\|x^*_A\|_2$), as well as Gaussian tail-bounds to say $\boldsymbol{g}$ is is larger than $\sqrt{n}/(4\|x^*_A\|_2)$. The minimum over $\tau > 0$ is meant to quantify over possible values of $\|x^*_A\|_2$. Letting $\gamma = 50000\sqrt{q}n^{1/4}\log n$ allows us to conclude that the probability that $\langle \boldsymbol{v}, x^*\rangle$ lies within distance γ of $-\sqrt{n}/2$ or $\sqrt{n}/2$ is at most $O(\sqrt{q}\log n/n^{1/4})$.

On the other hand, given that $\|(y-x^*)_A\| \le 1000\sqrt{q}n^{1/4}$, we also have that

$$\Pr_{\boldsymbol{v}}\left[\left|\langle \boldsymbol{v}, y-x^*\rangle\right| \ge 1000\sqrt{q}n^{1/4}\log n\right] \le \Pr_{\boldsymbol{g}\sim N(0,1)}\left[|\boldsymbol{g}| \ge \log n\right],$$

which is smaller than any inverse polynomial in n. Therefore, we have that, except with probability at most $O(\sqrt{q}\log n/n^{1/4})$, for both $\xi \in \{-1,1\}$,

- $\left|\langle \boldsymbol{v}, x^*\rangle - \xi\cdot\frac{\sqrt{n}}{2}\right| \le 50000\sqrt{q}n^{1/4}\log n$; and
- $|\langle \boldsymbol{v}, x^* - y\rangle| \le 1000\sqrt{q}n^{1/4}\log n$.

When these two events occur, event (2) cannot occur, which shows that (2) occurs with probability at most $O(\sqrt{q}\log n/n^{1/4})$.

To conclude (1), we note that any x^* with $\|x^*_A\|_2 \le \sqrt{2n}$ satisfies

$$\Pr_{\boldsymbol{v}}\left[\langle \boldsymbol{v}, x^*\rangle \in [-\sqrt{n}/2, \sqrt{n}/2]\right] = \Pr_{\boldsymbol{g}\sim N(0,1)}\left[\boldsymbol{g} \in \left[-\frac{1}{2\sqrt{2}}, \frac{1}{2\sqrt{2}}\right]\right] = \Omega(1),$$

which shows that conditioning on event (1) does not significantly affect the probability of (2). ∎

5 A Mildly-Exponential Lower Bound for Non-Adaptive Tolerant Testers

We will prove the following:

THEOREM 5.1. (TWO-SIDED NON-ADAPTIVE TOLERANT TESTING LOWER BOUND) *There exist absolute constants $0 < \varepsilon_1 < \varepsilon_2 < 0.5$ such that any non-adaptive $(\varepsilon_1, \varepsilon_2)$-tolerant tester for convexity over $N(0, I_n)$ (which may make two-sided errors) must use at least $2^{\Omega(n^{1/4})}$ queries.*

5.1 The $\mathcal{D}_{\text{yes}}$ and $\mathcal{D}_{\text{no}}$ Distributions

Before specifying the $\mathcal{D}_{\text{yes}}$ and $\mathcal{D}_{\text{no}}$ distributions, we first describe some necessary objects.

The Control and Action Subspaces. Throughout, we will work over $\mathbb{R}^{n+1}$ for convenience. Let $\boldsymbol{A}$ denote a random 1-dimensional subspace of $\mathbb{R}^{n+1}$, i.e.

$$\boldsymbol{A} = \{t\boldsymbol{v} : t \in \mathbb{R}\} \text{ where } \boldsymbol{v} \sim \mathbb{S}^n \text{ is a Haar-random unit vector.}$$

Let $\boldsymbol{C}$ be the orthogonal complement of $\boldsymbol{A}$; note that $\boldsymbol{C}$ is a random n-dimensional subspace of $\mathbb{R}^{n+1}$. We call $\boldsymbol{C}$ the *control subspace* and we call $\boldsymbol{A}$ the *action subspace*.

Given a vector $x \in \mathbb{R}^n$, we write $x_{\boldsymbol{C}}$ to denote the projection of x onto $\boldsymbol{C}$ and we write $x_{\boldsymbol{A}}$ to denote the projection of x onto $\boldsymbol{A}$, so every vector satisfies $x = x_{\boldsymbol{A}} + x_{\boldsymbol{C}}$. Recalling that $\boldsymbol{A}$ is a 1-dimensional subspace, when there is no risk of confusion we write $x_{\boldsymbol{A}}$ to denote the scalar value t such that $x_{\boldsymbol{A}} = t\boldsymbol{v}$.

Constants. We will use four positive absolute constants c_0, c_1, c_2 and τ in the construction. Here c_0 is the constant hidden in the statement of Lemma 3.3. We set c_1, c_2 and τ as follows:

$$c_1 = \frac{1}{100} \quad \text{and} \quad c_2 = \tau = \frac{c_0 c_1}{100} \tag{5.28}$$

so that Equation (5.29) at the end of Section 5.2 is $\Omega(1)$.

Nazarov's Body on the Control Subspace. Let $N := 2^{\sqrt{n}}$. We take r to satisfy Equation (3.4) for the absolute constant c_1 given in Equation (5.28), and draw $\boldsymbol{B} \sim \text{Naz}(r, N, \boldsymbol{C})$ (cf. Definition 2) where $\text{Naz}(r, N, \boldsymbol{C})$ is as defined in Definition 2 but with the n-dimensional control subspace $\boldsymbol{C}$ playing the role of $\mathbb{R}^n$. (This notation is as in Section 4.1.1 where we write $\text{Naz}(r, N, \boldsymbol{C})$ to mean the distribution $\text{Naz}(r, N)$ over bodies in $\boldsymbol{C}$.) Similar to Section 4, $\boldsymbol{B} \subseteq \mathbb{R}^{n+1}$ is a $\boldsymbol{C}$-subspace junta. Note in particular that a draw of $\boldsymbol{B} \sim \text{Naz}(r, N, \boldsymbol{C})$ immediately specifies $\boldsymbol{g}^i, \boldsymbol{H}_i, \boldsymbol{F}_i, \boldsymbol{U}_i$ for $i \in [N]$ (and that the sets $\boldsymbol{H}_i, \boldsymbol{F}_i, \boldsymbol{U}_i \subseteq \mathbb{R}^{n+1}$ are $\boldsymbol{C}$-subspace juntas as well).

Functions on the Action Subspace. Let c_2 be the absolute constant given in Equation (5.28). Intuitively, $2c_2$ will be the Gaussian measure of two symmetric intervals in the action subspace $\boldsymbol{A}$. More formally we define

$$\text{Curb} := \left[\Phi^{-1}\left(\frac{1-2c_2}{3}\right), \Phi^{-1}\left(\frac{1+c_2}{3}\right)\right] \cup \left[\Phi^{-1}\left(\frac{2-c_2}{3}\right), \Phi^{-1}\left(\frac{2+2c_2}{3}\right)\right].$$

Using the absolute constant τ given in Equation (5.28), we also define the interval

$$I := \left[\sqrt{n+1} - \sqrt{2\ln(2/\tau)}, \sqrt{n+1} + \sqrt{2\ln(2/\tau)}\right],$$

and we observe that a random draw of $\boldsymbol{x}$ from $N(0, I_{n+1})$ has $\|\boldsymbol{x}\| \in I$ with probability at least $1-\tau$. We write Shell_{n+1} to denote the corresponding spherical shell in $\mathbb{R}^{n+1}$, i.e.

$$\text{Shell}_{n+1} = \{x \in \mathbb{R}^{n+1} : \|x\| \in I\}.$$

Finally, let $\boldsymbol{P}$ be a uniformly random subset of $[N]$.

For a fixed setting of C (which defines the complementary A), B (which in turn defines H_i, F_i, the F_T's, and U_i), and P, we define the function $g_{C,B,P} : \mathbb{R}^{n+1} \to \{0, 1, 0^\star, 1^\star\}$ as follows:

$$g_{C,B,P}(x) = \begin{cases} 0 & \text{if } \|x_C\| \geq \sqrt{n} \text{ or } x \in \bigcup_{|T|\geq 2} F_T \text{ or } x \notin \text{Shell}_{n+1}, \\ 1 & \text{if } x \in \text{Shell}_{n+1} \text{ and } x \in B, \\ 0 & \text{if } x \in \text{Shell}_{n+1} \text{ and } x \in U_i \text{ for some } i \in [N] \text{ and } x_A \in \text{Curb}, \\ 0^\star & \text{if } x \in \text{Shell}_{n+1} \text{ and } x \in U_i \text{ for some } i \in P \text{ and } x_A \notin \text{Curb}, \\ 1^\star & \text{if } x \in \text{Shell}_{n+1} \text{ and } x \in U_i \text{ for some } i \notin P \text{ and } x_A \notin \text{Curb}. \end{cases}$$

The $\mathcal{D}_{\text{yes}}$ and $\mathcal{D}_{\text{no}}$ Distributions. To sample a set from either $\mathcal{D}_{\text{yes}}$ or $\mathcal{D}_{\text{no}}$, first draw $\boldsymbol{C}, \boldsymbol{B}, \boldsymbol{P}$ as described above; note that this induces a draw of $\boldsymbol{A}, \boldsymbol{H}_i, \boldsymbol{F}_i$, the $\boldsymbol{F}_T$'s, and $\boldsymbol{U}_i$. Draws from $\mathcal{D}_{\text{yes}}$ and $\mathcal{D}_{\text{no}}$ are identical on points $x \in \mathbb{R}^{n+1}$ where $g_{\boldsymbol{C},\boldsymbol{B},\boldsymbol{P}}(x) \in \{0,1\}$; on the other values, however,

- For functions in $\mathcal{D}_{\text{yes}}$, we set $0^\star \mapsto 0$ and $1^\star \mapsto 1$.
- For functions in $\mathcal{D}_{\text{no}}$, we set

$$0^\star \mapsto \mathbf{1}\{x_A \notin \text{Middle}\} \qquad \text{and} \qquad 1^\star \mapsto \mathbf{1}\{x_A \in \text{Middle}\}$$

  where we define

$$\text{Middle} := \left(\Phi^{-1}\left(\frac{1+c_2}{3}\right), \Phi^{-1}\left(\frac{2-c_2}{3}\right)\right) \subset \mathbb{R}.$$

See Figures 2 and 3 for illustrations of $\mathcal{D}_{\text{yes}}$ and $\mathcal{D}_{\text{no}}$.

5.2 Distance to Convexity Recall that a draw of a function from either $\mathcal{D}_{\text{yes}}$ or $\mathcal{D}_{\text{no}}$ induces a draw of $\boldsymbol{B}, \boldsymbol{C}$, and $\boldsymbol{P}$. First, we give an upper bound on the expected distance to convexity of a function drawn from $\mathcal{D}_{\text{yes}}$. Let ε_1 be the constant given by

$$\varepsilon_1 := 2c_2 + \tau + 2 \cdot \mathop{\mathbf{E}}_{\boldsymbol{B} \sim \text{Naz}(r,N)} \left[\text{Vol}\left(\bigcup_{|T|\geq 2} \boldsymbol{F}_T\right)\right].$$

PROPOSITION 5.1. *We have* $\text{dist}(\boldsymbol{f}_{\text{yes}}, \mathcal{P}_{\text{conv}}) \leq \varepsilon_1$ *with probability at least* 0.5 *when* $\boldsymbol{f}_{\text{yes}} \sim \mathcal{D}_{\text{yes}}$.

Proof. Consider a fixed choice of C, B and P. We then consider the convex set $G'_{C,B,P} \subset \mathbb{R}^{n+1}$ defined as the intersection of H_i (for $i \in P$) and the set $\{x : \|x_C\|_2 \leq \sqrt{n}\}$. By construction, the set $G'_{C,B,P}$ is a convex set. Let $g'_{C,B,P}$ denote the corresponding indicator function. We now analyze the distance between the functions $g_{C,B,P}$ (where, as for functions in the support of $\mathcal{D}_{\text{yes}}$, we identify $0^\star$ with 0 and $1^\star$ with 1) and $g'_{C,B,P}$.

First of all, by construction, if $g_{C,B,P}(x) = 1$, then $g'_{C,B,P}(x)$ is also 1. So to bound the distance, we note that there are three possible ways in which $g_{C,B,P}(x)$ can be 0 but $g'_{C,B,P}(x)$ can be 1:

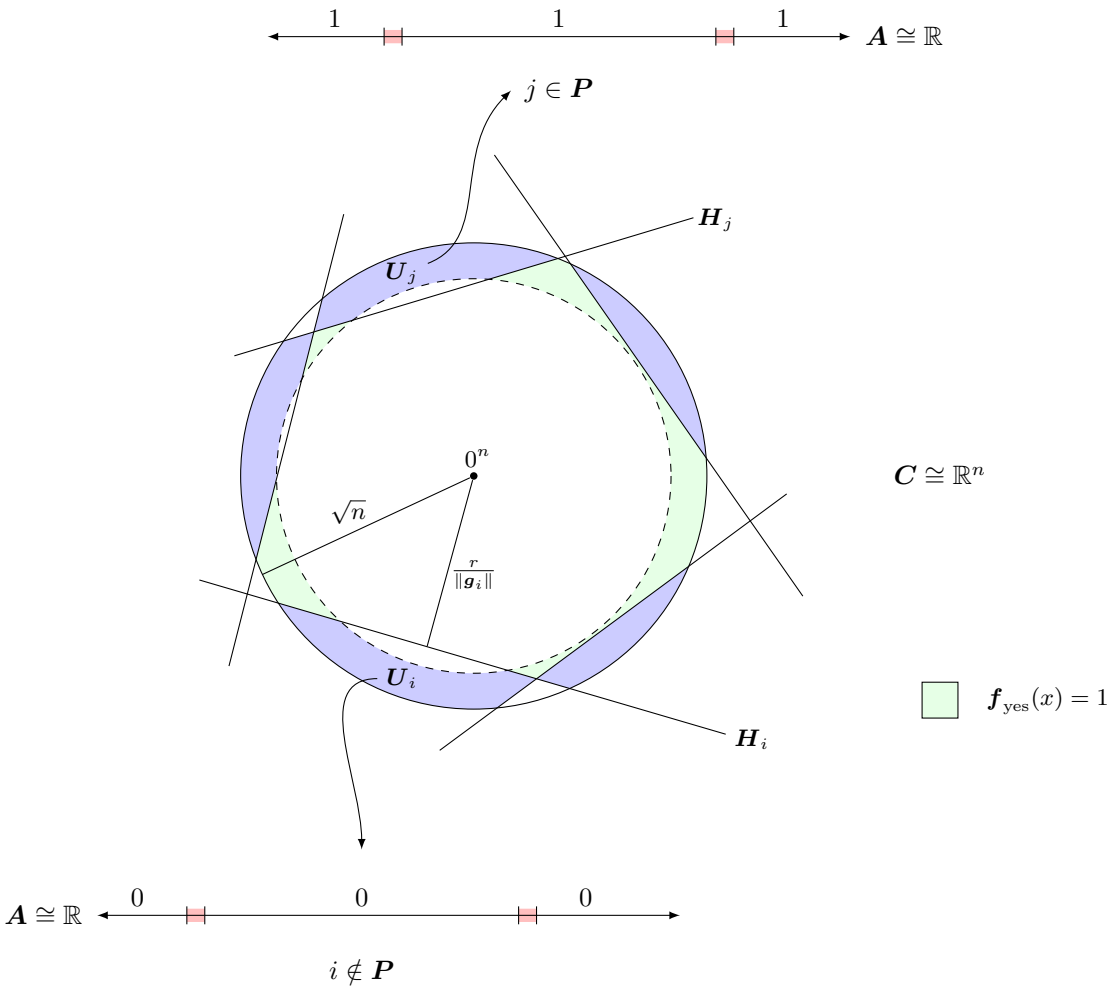

Figure 2: A depiction of $\mathcal{D}_{\text{yes}}$. We identify the control subspace $\boldsymbol{C} \cong \mathbb{R}^n$. The annulus defined by the boundary of Ball($\sqrt{n}$) and the dotted circle corresponds to points x which satisfy $x \in \text{Shell}_{n+1}$ and $\|x_{\boldsymbol{C}}\| \leq \sqrt{n}$. Finally, the red region in the action subspace $\boldsymbol{A} \cong \mathbb{R}$ corresponds to Curb.

1. $x \in \cup_{|T|\geq 2} F_T$;
2. $x \notin \text{Shell}_{n+1}$;
3. There is some $i \in [N]$ such that $x \in U_i$ and $x_A \in \text{Curb}$.

By definition, (i) the Gaussian volume of the first set is $\text{Vol}(\cup_{|T|\geq 2} F_T)$; (ii) the Gaussian volume of the second set is bounded by τ; (iii) the Gaussian volume of the third set is bounded by Vol(Curb) which by definition is $2c_2$. Thus, for a specific instantiation of C, B and P,

$$\text{dist}(g_{C,B,P}, \mathcal{P}_{\text{conv}}) \leq 2c_2 + \tau + \text{Vol}\left(\bigcup_{|T|\geq 2} F_T\right).$$

The claim follows from Markov's inequality, that the last term on the RHS above is at most twice the expectation with probability at least 1/2. ∎

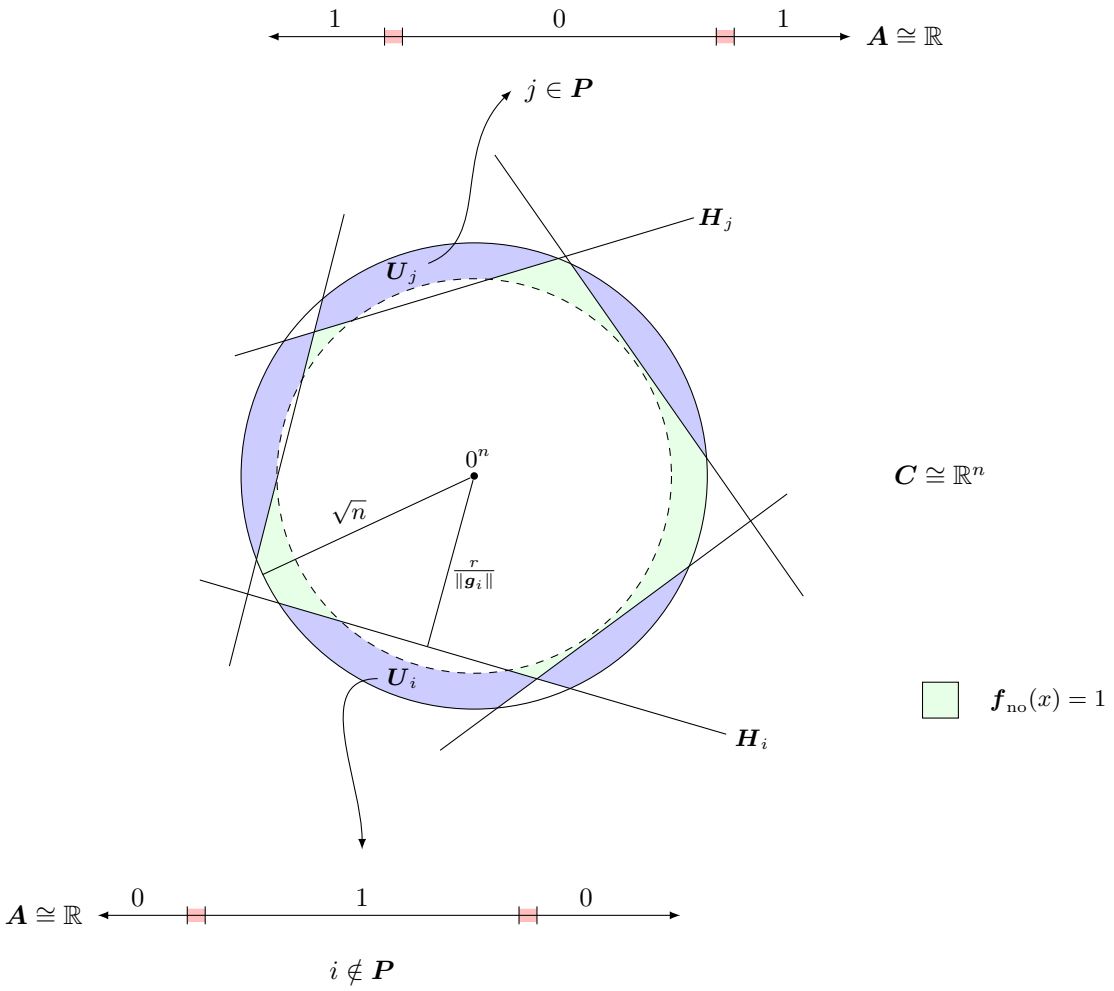

Figure 3: A depiction of $\mathcal{D}_{\text{no}}$. Our conventions are as in Figure 2.

Next, we give a lower bound on the expected distance to convexity of a function drawn from $\mathcal{D}_{\text{no}}$. Let ε_2 be the constant given as

$$\varepsilon_2 := \left(\frac{1-2c_2}{3}\right)\left(0.3 \cdot \mathop{\mathbf{E}}_{\boldsymbol{B}\sim\text{Naz}(r,N)}\left[\text{Vol}\left(\bigsqcup_{i=1}^{N} \boldsymbol{U}_i\right)\right] - \frac{\tau}{2}\right).$$

PROPOSITION 5.2. *We have* $\text{dist}(\boldsymbol{f}_{\text{no}}, \mathcal{P}_{\text{conv}}) \geq \varepsilon_2$ *with probability at least* $1 - o(1)$ *when* $\boldsymbol{f}_{\text{no}} \sim \mathcal{D}_{\text{no}}$.

Proof. As before, consider a fixed choice of C, B and P. Take any x_C such that $x_C \in U_i$ for some $i \in P$ and $\|x_C\| \geq \sqrt{n+1} - \sqrt{2\ln(2/\tau)}$ (which is the left end of I). For any such x_C, the line along x_A looks like the picture at the top of Figure 3, except that the function is set to 0 when $\|x_C\|$ makes x go outside of Shell_{n+1}. Given that $\|x_C\| \leq \sqrt{n}$, this only happens when $|x_A|$ is at least $\Omega(n^{1/4})$ given that $\tau \leq 1/10000$ in Equation (5.28).

As a result, the distance to convexity along this line in the action space (with x_C fixed) is at least $\text{Vol}(\text{Middle}) = (1-2c_2)/3$. This follows immediately from the fact that Middle is a symmetric interval about 0 and the choice of parameters in the definition of Middle.

Given that the mass of x with $\|x_C\| < \sqrt{n+1} - \sqrt{2\ln(2/\tau)}$ is at most $\tau/2$, it follows that

$$\operatorname{dist}(f_{\text{no}}, \mathcal{P}_{\text{conv}}) \geq \left(\left(\sum_{i \in P} \operatorname{Vol}(U_i)\right) - \frac{\tau}{2}\right) \cdot \left(\frac{1-2c_2}{3}\right)$$

The result follows by a straight forward modification of Lemma 3.4 to show that with probability at least $1 - o(1)$, we have $\sum_{i \in P} \operatorname{Vol}(U_i)$ is at least $0.3 \cdot \mathbf{E}[\operatorname{Vol}(\sqcup_{i \in [N]} \boldsymbol{U}_i)]$ when $\boldsymbol{B} \sim \operatorname{Naz}(r, N)$. ∎

Setting Parameters. We verify that $\varepsilon_2 - \varepsilon_1 = \Omega(1)$:

$$\varepsilon_2 - \varepsilon_1 \geq \left(\frac{1-2c_2}{3}\right)\left(0.3 \cdot \mathbf{E}\left[\operatorname{Vol}\left(\bigsqcup_{i=1}^{N} \boldsymbol{U}_i\right)\right] - \frac{\tau}{2}\right) - 2c_2 - \tau - 2 \cdot \mathbf{E}\left[\operatorname{Vol}\left(\bigcup_{|T| \geq 2} \boldsymbol{F}_T\right)\right]$$

$$\text{(Lemma 3.5)} \qquad \geq \mathbf{E}\left[\operatorname{Vol}\left(\bigsqcup_{i=1}^{N} \boldsymbol{U}_i\right)\right]\left(\frac{1-2c_2}{10} - \frac{c_1}{1-c_1}\right) - 2c_2 - \tau\left(\frac{7-2c_2}{6}\right)$$

$$(5.29) \qquad \geq c_0 c_1 \left(\frac{1-2c_2}{10} - \frac{c_1}{1-c_1}\right) - 2c_2 - \frac{7\tau}{6},$$

(where the last line is by Lemma 3.3) which is $\Omega(1)$ given choices of c_0, c_1, c_2 and τ made in Equation (5.28).

5.3 Proof of Theorem 1.2 We introduce some helpful notation and outline the high-level structure of the argument.

5.3.1 Setup and Outline of Argument We introduce the following notation:

NOTATION 7. *Given an outcome of the control subspace C and of Nazarov's body $B = H_1 \cap \cdots \cap H_N \cap \operatorname{Ball}(\sqrt{n}) \subset \mathbb{R}^{n+1}$ within C as defined earlier, for $x \in \mathbb{R}^{n+1}$ we define the set $S_B(x)$ as*

$$S_B(x) := \{\ell \in [N] : x \in F_\ell\}.$$

Note that if x and y have $x_C = y_C$, then $S_B(x) = S_B(y)$, i.e. only the C-part of x affects S_B.

We define the regions $\text{Left}, \text{Middle}, \text{Right} \subset \mathbb{R}$ as follows:

$$\text{Left} := \left(-\infty, \Phi^{-1}\left(\frac{1-2c_2}{3}\right)\right),$$

$$\text{Middle} := \left(\Phi^{-1}\left(\frac{1+c_2}{3}\right), \Phi^{-1}\left(\frac{2-c_2}{3}\right)\right),$$

$$\text{Right} := \left(\Phi^{-1}\left(\frac{2+2c_2}{3}\right), \infty\right).$$

Note that $\text{Left} \sqcup \text{Middle} \sqcup \text{Right} \sqcup \text{Curb} = \mathbb{R}$ (where as before we identify $\mathbb{R}$ with an outcome of the one-dimensional action subspace A).

To establish indistinguishability, we show that no non-adaptive deterministic algorithm $\mathcal{A}$ that makes $q = 2^{c_3 n^{1/4}}$ queries, for some sufficiently small constant c_3, can distinguish $\mathcal{D}_{\text{yes}}$ from $\mathcal{D}_{\text{no}}$. Specifically, for any nonadaptive deterministic algorithm $\mathcal{A}$ with query complexity q, we show that

$$(5.30) \qquad \mathop{\mathbf{Pr}}_{\boldsymbol{f}_{\text{yes}} \sim \mathcal{D}_{\text{yes}}}\left[\mathcal{A} \text{ accepts } \boldsymbol{f}_{\text{yes}}\right] \leq \mathop{\mathbf{Pr}}_{\boldsymbol{f}_{\text{no}} \sim \mathcal{D}_{\text{no}}}\left[\mathcal{A} \text{ accepts } \boldsymbol{f}_{\text{no}}\right] + o(1).$$

To this end, we define Bad to be the following event:

Bad: There are $x, y \in \text{Shell}_{n+1}$ queried by $\mathcal{A}$ that (i) satisfy $S_{\boldsymbol{B}}(x) = S_{\boldsymbol{B}}(y) = \{\ell\}$ for some $\ell \in [N]$ (or equivalently, $x, y \in U_\ell$ for some ℓ), and (ii) have $x_{\boldsymbol{A}}, y_{\boldsymbol{A}}$ belonging to two *distinct* sets among Left, Middle, Right.

We will first show in Lemma 5.1 that $\mathcal{A}$ can distinguish $\mathcal{D}_{\text{yes}}$ from $\mathcal{D}_{\text{no}}$ only when Bad occurs. On the other hand, in Lemma 5.2, we show Bad occurs with probability $o(1)$ when the number of queries is $q = 2^{c_3 n^{1/4}}$ and c_3 is sufficiently small. Lemmas 5.1 and 5.2 together establishe Equation (5.30); the proof of this is analogous to the proof of Theorem 1 in Section 4.2 of [CDL+24] and we refer the reader to [CDL+24] for full details. Theorem 1.2 then follows from Equation (5.30) via Yao's minimax principle (Theorem 2.1).

5.3.2 Indistinguishability of $\mathcal{D}_{\text{yes}}$ and $\mathcal{D}_{\text{no}}$ We write $\mathcal{A}(f)$ to denote the sequence of q answers to the queries made by $\mathcal{A}$ to f. We write $\text{view}_{\mathcal{A}}(\mathcal{D}_{\text{yes}})$ (respectively $\text{view}_{\mathcal{A}}(\mathcal{D}_{\text{no}})$) to be the distribution of $\mathcal{A}(\boldsymbol{f}_{\text{yes}})$ for $\boldsymbol{f}_{\text{yes}} \sim \mathcal{D}_{\text{yes}}$ (respectively $\boldsymbol{f}_{\text{no}} \sim \mathcal{D}_{\text{no}}$). The following claim asserts that conditioned on Bad not happening, the distributions $\text{view}_{\mathcal{A}}(\mathcal{D}_{\text{yes}}|_{\overline{\text{Bad}}})$ and $\text{view}_{\mathcal{A}}(\mathcal{D}_{\text{no}}|_{\overline{\text{Bad}}})$ are identical.

LEMMA 5.1. $\text{view}_{\mathcal{A}}(\mathcal{D}_{\text{yes}}|_{\overline{\text{Bad}}}) = \text{view}_{\mathcal{A}}(\mathcal{D}_{\text{no}}|_{\overline{\text{Bad}}})$.

Proof. Let Q be the set of points queried by $\mathcal{A}$. Recall that the distributions of the subspaces $\boldsymbol{C}$ and action variables $\boldsymbol{A}$ are identical for $\mathcal{D}_{\text{yes}}$ and $\mathcal{D}_{\text{no}}$. So fix an arbitrary outcome of the n-dimensional subspace C and the orthogonal one-dimensional subspace A. As the distribution of the Nazarov body $\boldsymbol{B} \sim \text{Naz}(r, N, C)$ is also identical for $\mathcal{D}_{\text{yes}}$ and $\mathcal{D}_{\text{no}}$, we fix an arbitrary outcome B of $\boldsymbol{B}$. Let $\boldsymbol{f}$ be a random function drawn from either $\mathcal{D}_{\text{yes}}$ or $\mathcal{D}_{\text{no}}$.

Note that for any point $x \in \mathbb{R}^{n+1}$ such that $|S_B(x)| \neq 1$ or $x \notin \text{Shell}_{n+1}$ or $x_A \in \text{Curb}$, by construction we have that $\boldsymbol{f}(x)$ can be determined directly in the same way for both $\mathcal{D}_{\text{yes}}$ and $\mathcal{D}_{\text{no}}$ (no query is required). So it suffices for us to consider the points x such that $|S_B(x)| = 1$, $x \in \text{Shell}_{n+1}$, and $x_A \notin \text{Curb}$. We call these points *important* points.

We divide these important points into disjoint groups according to $S_B(x)$. More precisely, for every $\ell \in [N]$, let $X_\ell = \{x \in \mathbb{R}^{n+1} \mid x \text{ is important}, S_B(x) = \{\ell\}\}$. Let $\boldsymbol{f}_\ell(x)$ denote the function $\boldsymbol{f}(x)$ restricted to X_ℓ (where as stated above, $\boldsymbol{f}$ denotes either a function drawn from $\mathcal{D}_{\text{yes}}$ or from $\mathcal{D}_{\text{no}}$). The condition that Bad does not happen implies that *either* $x_A \in$ Left for all $x \in Q \cap X_\ell$, *or* $x_A \in$ Middle for all $x \in Q \cap X_\ell$, *or* $x_A \in$ Right for all $x \in Q \cap X_\ell$. In particular, this means $\boldsymbol{f}_\ell(x) = \boldsymbol{f}_\ell(y)$ for all $x, y \in Q \cap X_\ell$, and this holds for both $\mathcal{D}_{\text{yes}}$ and $\mathcal{D}_{\text{no}}$.

Since $\boldsymbol{f}_\ell(x)$ are the same for all $x \in Q \cap X_\ell$, the distribution of $\boldsymbol{f}_\ell$ is actually one random bit. Indeed, $\boldsymbol{f}_\ell(x) = 0$ with probability $1/2$ and $\boldsymbol{f}_\ell(x) = 1$ with probability $1/2$ (because each element $\ell \in [N]$ belongs to $\boldsymbol{P}$ with probability $1/2$) independently, and this holds for both $\mathcal{D}_{\text{yes}}$ and $\mathcal{D}_{\text{no}}$. This completes the proof of the lemma. ∎

Next, we show that Bad happens with probability $o(1)$ (recall that $q = 2^{c_3 n^{1/4}}$). The proof of the following lemma follows the proof of an analogous lemma from [CDL+24]:

LEMMA 5.2. *For any fixed set of points $Q = \{x^1, \cdots, x^q\} \subset \mathbb{R}^{n+1}$, we have* $\mathbf{Pr}[\text{Bad}] = o(1)$.

Proof. Fix a pair of query points $x, y \in \mathbb{R}^{n+1}$ that belong to Q. By the definition of Bad, we may assume without loss of generality that $x, y \in \text{Shell}_{n+1}$. Let $\text{Bad}_{x,y}$ be the event that

(a) $x, y \in \boldsymbol{U}_\ell$ for some $\ell \in [N]$ (equivalently, $S_{\boldsymbol{B}}(x) = S_{\boldsymbol{B}}(y) = \{\ell\}$), and

(b) $x_{\boldsymbol{A}}, y_{\boldsymbol{A}}$ belong to two *distinct* sets among $\{\text{Left}, \text{Middle}, \text{Right}\}$.

Analogous to the argument in [CDL+24], we will show that

$$\mathop{\mathbf{Pr}}_{\boldsymbol{B}}[\text{Bad}_{x,y}] \leq \min\{\mathbf{Pr}_{\boldsymbol{B}}[(a)], \mathbf{Pr}_{\boldsymbol{B}}[(b)]\} \text{ is very small.} \tag{5.31}$$

Recall that each of the two intervals defining Curb (cf. Section 5.3.1) has the same width which we will denote $\rho(c_2)$ for succinctness, i.e.

$$\rho(c_2) := \Phi^{-1}\left(\frac{1+c_2}{3}\right) - \Phi^{-1}\left(\frac{1-2c_2}{3}\right).$$

On one hand, for (b) to happen on x, y, we must have

$$(\diamond) \qquad \left|x_{\boldsymbol{A}} - y_{\boldsymbol{A}}\right| \geq \rho(c_2).$$

On the other hand, (a) means

$$(\star) \qquad \text{There exists } \ell \in [N] \text{ such that } S_{\boldsymbol{B}}(x) = S_{\boldsymbol{B}}(y) = \{\ell\}.$$

It follows that $\mathbf{Pr}[\mathsf{Bad}_{xy}] \leq \min\{\mathbf{Pr}[\diamond], \mathbf{Pr}[\star]\}$. We will show that $\min\{\mathbf{Pr}[\diamond], \mathbf{Pr}[\star]\} \leq 2^{-4c_3 n^{1/4}}$ and will do so via the following lemmas, Lemma 5.3 and Lemma 5.4 (below $c > 0$ is a suitable positive absolute constant):

LEMMA 5.3. *If $\|x - y\| \leq cn^{3/8}$, then $\mathbf{Pr}[\diamond] \leq 2^{-4c_3 n^{1/4}}$.*

Proof. Fix x, y such that $\|x-y\| \leq cn^{3/8}$. For succinctness we write z to denote $x-y$, so $z \in \mathbb{R}^{n+1}$ and $\|z\| \leq cn^{3/8}$; our goal is to show that $|z_{\boldsymbol{A}}| \leq \rho(c_2)$ except with probability at most $2^{-4c_3 n^{1/4}}$.

Since $\boldsymbol{A}$ is a Haar-random direction in $\mathbb{R}^{n+1}$, the distribution of $z_{\boldsymbol{A}}$ is the same as the distribution of $\|z\| \cdot \boldsymbol{v}_1$ where $\boldsymbol{v} \sim \mathbb{S}^{n-1}$. Hence by standard bounds on spherical caps (Lemma 2.1),

$$\mathbf{Pr}\left[|\boldsymbol{v}_1| \geq \frac{t}{\sqrt{n}}\right] \leq e^{-t^2/2}.$$

Taking $t = \sqrt{8c_3} \cdot n^{1/8}$, this probability is at most $e^{-4c_3 n^{1/4}} < 2^{-4c_3 n^{1/4}}$. So we set our threshold as

$$\|z\| \leq \frac{\rho(c_2)}{2\sqrt{2c_3}} \cdot n^{3/8},$$

i.e. we require that $c \leq \frac{\rho(c_2)}{2\sqrt{2c_3}}$, and the lemma is established. ∎

LEMMA 5.4. *If $\|x - y\| > cn^{3/8}$, then $\mathbf{Pr}[\star] \leq 2^{-4c_3 n^{1/4}}$.*

We defer the proof of Lemma 5.4 to Section 5.3.3. Thanks to Lemmas 5.3 and 5.4, we get that

$$\mathbf{Pr}[\mathsf{Bad}_{xy}] \leq \min\left\{\mathbf{Pr}[\diamond], \mathbf{Pr}[\star]\right\} \leq 2^{-4c_3 n^{1/4}}.$$

By a union bound over all (at most $2^{2c_3 n^{1/4}}$) pairs of points x, y from Q, we get that

$$\mathbf{Pr}[\mathsf{Bad}] \leq 2^{-4c_3 n^{1/4}} \cdot 2^{2c_3 n^{1/4}} = 2^{-2c_3 n^{1/4}} = o(1),$$

which completes the proof. ∎

5.3.3 Proof of Lemma 5.4 For the remainder of this section, we will always assume that $x, y \in \mathrm{Shell}_{n+1}$ satisfy $\|x - y\| > cn^{3/8}$. Note that we can view the construction $g_{\boldsymbol{C},\boldsymbol{B},\boldsymbol{P}}$ as a two stage process:

- We first draw $\boldsymbol{C}$, which is a Haar random n dimensional subspace of $\mathbb{R}^{n+1}$.
- We then draw $\boldsymbol{B} \sim \mathrm{Naz}(r, N, \boldsymbol{C})$, and we draw $\boldsymbol{P}$ as a uniformly random subset of $[N]$.

We require the following claim:

CLAIM 8. *Suppose $x, y \in \mathrm{Shell}_{n+1}$ satisfy $\|x - y\| > cn^{3/8}$. Then with probability at least $1 - 2^{-\Omega(n^{1/4})}$ over the outcomes of $\boldsymbol{C}$, we have*

$$(5.32) \qquad \|x_{\boldsymbol{C}}\| \geq \|x\| - 1, \quad \|y_{\boldsymbol{C}}\| \geq \|y\| - 1, \quad \text{and} \quad \|(x-y)_{\boldsymbol{C}}\| \geq cn^{3/8} - 1.$$

Proof. Fix $x, y \in \mathrm{Shell}_{n+1}$ such that $\|x - y\| > cn^{3/8}$. Because $\boldsymbol{A}$ is drawn Haar-randomly (and since it defines $\boldsymbol{C}$), it follows from Lemma 2.1 that

$$\Pr_{\boldsymbol{A}}\left[\|x_{\boldsymbol{A}}\| \geq t \cdot n^{-1/2} \cdot \|x\|\right] \leq e^{-t^2/2}.$$

Let $t = \beta n^{1/8}$ for a suitable constant $\beta > 0$. The previous inequality gives

$$\Pr_{\boldsymbol{A}}\left[\|x_{\boldsymbol{A}}\| \geq \beta \cdot n^{-3/8} \cdot \|x\|\right] \leq e^{-\beta^2 n^{1/4}/2}.$$

Thus, with probability $1 - e^{-\beta^2 n^{1/4}/2}$, we have

$$\begin{aligned}
\|x_{\boldsymbol{C}}\| = \sqrt{\|x\|^2 - \|x_{\boldsymbol{A}}\|^2} &= \|x\| \cdot \sqrt{1 - \frac{\|x_{\boldsymbol{A}}\|^2}{\|x\|^2}} \\
&\geq \|x\| \cdot \left(1 - \frac{\|x_{\boldsymbol{A}}\|^2}{2\|x\|^2}\right) \\
&= \|x\| - \frac{\|x_{\boldsymbol{A}}\|^2}{2\|x\|} \\
&\geq \|x\| - \frac{\beta^2 n^{-3/4}\|x\|}{2}.
\end{aligned}$$

As $x \in \mathrm{Shell}_{n+1}$ and thus $\|x\| \leq 2\sqrt{n}$, it follows that the last expression is at least $\|x\| - 1$. Identical calculations yield the corresponding lower bounds on $\|y_{\boldsymbol{C}}\|$ and $\|(x-y)_{\boldsymbol{C}}\|$. ∎

Fix an outcome C of $\boldsymbol{C}$ such that Equation (5.32) holds. For convenience, we will write x' for x_C, y' for y_C, both of which lie in $\mathbb{R}^n$. For the rest of the argument, we will work over C, i.e. we view sets such as $\boldsymbol{H}_\ell, \boldsymbol{H}'_\ell$ and $\boldsymbol{B}$ as lying in C (which we identify with $\mathbb{R}^n$) rather than in $\mathbb{R}^{n+1}$.

The following argument is analogous to (parts of) the proof of Lemma 15 of [CDL+24]. Recall

$$S_{\boldsymbol{B}}(x') = \left\{\ell \in [N] : x' \in \boldsymbol{U}_\ell\right\}.$$

By Claim 8, we have that

$$\begin{aligned}
\mathbf{Pr}\left[\star\right] &\leq 2^{-\Omega(n^{1/4})} + \Pr_{\boldsymbol{B}}\left[S_{\boldsymbol{B}}(x') = S_{\boldsymbol{B}}(y') = \{\ell\} \text{ for some } \ell\right] \\
&= 2^{-\Omega(n^{1/4})} + \Pr_{\boldsymbol{B}}\left[{}_{\boldsymbol{B}}(x') = S_{\boldsymbol{B}}(y') \text{ and } \exists \ell \text{ s.t. } S_{\boldsymbol{B}}(y') = \{\ell\}\right] \\
&\leq 2^{-\Omega(n^{1/4})} + \Pr_{\boldsymbol{B}}\left[S_{\boldsymbol{B}}(x') = S_{\boldsymbol{B}}(y') \;\mid\; \exists \ell \text{ s.t. } S_{\boldsymbol{B}}(y') = \{\ell\}\right],
\end{aligned}$$

where x', y' satisfy Equation (5.32). We will analyze the case when $\ell = 1$, which is without loss of generality since

$$\mathbf{Pr}[\boldsymbol{X} | \sqcup_\ell E_\ell] \leq \sup_\ell \mathbf{Pr}[\boldsymbol{X}|E_\ell]$$

for disjoint events $\{E_\ell\}$ and since the probabilities

$$\Pr_{\boldsymbol{B}}\left[S_{\boldsymbol{B}}(x') = S_{\boldsymbol{B}}(y') \;\mid\; \exists \ell \text{ s.t. } S_{\boldsymbol{B}}(x') = \{\ell_1\}\right] = \Pr_{\boldsymbol{B}}\left[S_{\boldsymbol{B}}(x') = S_{\boldsymbol{B}}(y') \;\mid\; \exists \ell \text{ s.t. } S_{\boldsymbol{B}}(x') = \{\ell_2\}\right]$$

for all $\ell_1, \ell_2 \in [N]$. So our goal is to upper bound

$$\mathbf{Pr}_{\boldsymbol{B}}\left[S_{\boldsymbol{B}}(x') = S_{\boldsymbol{B}}(y') \;\mid\; S_{\boldsymbol{B}}(y') = \{1\}\right]. \tag{5.33}$$

Observe that the event "$S_{\boldsymbol{B}}(y') = \{1\}$" that we are conditioning on is an event over the random draw of $\boldsymbol{B}$, i.e. over the draw of $\boldsymbol{g}^1, \ldots, \boldsymbol{g}^N$. To analyze this event it is helpful to introduce the following notation: For $z' \in \mathbb{R}^n$, define

$$\mathrm{hfsp}(z') := \begin{cases} \left\{g \in \mathbb{R}^n : g \cdot z' \geq r\right\} & \|z'\| \leq \sqrt{n}, \\ \emptyset & \|z'\| > \sqrt{n}. \end{cases}$$

Consequently, the event "$S_{\boldsymbol{B}}(y') = \{1\}$" is the same as the event

$$\left\{\boldsymbol{g}^1 \in \text{hfsp}(y')\right\} \wedge \left\{\boldsymbol{g}^i \notin \text{hfsp}(y') \text{ for } i \in \{2,\ldots,N\}\right\}.$$

We may fix any outcome $g^{2*},\ldots,g^{N*}$ of $\boldsymbol{g}^2,\ldots,\boldsymbol{g}^N$ all of which lie outside of $\text{hfsp}(y')$, and we get (writing $\vec{\boldsymbol{g}}$ to denote $(\boldsymbol{g}^1,\ldots,\boldsymbol{g}^N)$) that

$$(5.33) = \mathop{\mathbf{Pr}}_{\vec{\boldsymbol{g}}}\left[S_{\boldsymbol{B}}(x') = S_{\boldsymbol{B}}(y') \mid (\boldsymbol{g}^1 \in \text{hfsp}(x')) \wedge \left(\boldsymbol{g}^i \notin \text{hfsp}(x') \text{ for } i \in [2:N]\right)\right]$$

$$(5.34) \qquad \leq \sup_{\substack{g^{i*} \notin \text{hfsp}(y') \\ i \neq 1}} \mathop{\mathbf{Pr}}_{\vec{\boldsymbol{g}}}\left[S_{\boldsymbol{B}}(x') = S_{\boldsymbol{B}}(y') \mid (\boldsymbol{g}^1 \in \text{hfsp}(x')) \wedge \left(\boldsymbol{g}^i = g^{i*} \text{ for } i \in [2:N]\right)\right]$$

$$(5.35) \qquad \leq \sup_{\substack{g^{i*} \notin \text{hfsp}(y') \\ i \neq 1}} \mathop{\mathbf{Pr}}_{\vec{\boldsymbol{g}}}\left[\boldsymbol{g}^1 \in \text{hfsp}(x') \cap \text{hfsp}(y') \mid (\boldsymbol{g}^1 \in \text{hfsp}(x')) \wedge \left(\boldsymbol{g}^i = g^{i*} \text{ for } i \in [2:N]\right)\right]$$

$$(5.36) \qquad = \mathop{\mathbf{Pr}}_{\boldsymbol{g} \sim N(0,I_n)}\left[\boldsymbol{g} \in \text{hfsp}(x') \cap \text{hfsp}(y') \mid \boldsymbol{g} \in \text{hfsp}(x')\right],$$

where Equation (5.34) uses $\mathbf{Pr}[\boldsymbol{X} | \sqcup_\ell E_\ell] \leq \sup_\ell \mathbf{Pr}[\boldsymbol{X}|E_\ell]$ as earlier; Equation (5.35) uses that if $\boldsymbol{g}^1 \in \text{hfsp}(y')$ then in order to have $S_{\boldsymbol{B}}(x') = S_{\boldsymbol{B}}(y')$ it must be the case that $\boldsymbol{g}^1 \in \text{hfsp}(x') \cap \text{hfsp}(y')$; and Equation (5.36) is because the event $\boldsymbol{g}^1 \in \text{hfsp}(x') \cap \text{hfsp}(y')$ is independent of the outcome of $\boldsymbol{g}^2,\ldots,\boldsymbol{g}^S$. So in what follows our goal is to upper bound (5.36). In other words, recalling that we write $\text{Vol}(K)$ to denote the Gaussian measure of the set K (cf. Section 2.2), our goal is to obtain an upper bound on

$$(5.37) \qquad (5.36) = \frac{\text{Vol}(\text{hfsp}(x') \cap \text{hfsp}(y'))}{\text{Vol}(\text{hfsp}(x'))},$$

which is a two-dimensional problem because the only thing that matters about the outcome of $\boldsymbol{g} \sim N(0,I_n)$ vis-a-vis (5.37) is the projection of $\boldsymbol{g}$ in the directions of x' and y'. Towards this goal, we recall the following tail bound for bivariate Gaussian random variables:

PROPOSITION 5.3. (EQUATION (2.11) OF [WIL05]) *Suppose* $(\boldsymbol{Z}_1, \boldsymbol{Z}_2) \sim N(0,\Sigma)$ *where*

$$\Sigma = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \qquad \text{for } \rho > 0.$$

Then for $h, k > 0$, *we have*

$$\mathop{\mathbf{Pr}}_{(\boldsymbol{Z}_1,\boldsymbol{Z}_2)\sim N(0,\Sigma)}[\boldsymbol{Z}_1 > h, \boldsymbol{Z}_2 > k] \leq \Phi(-h)\left(\Phi\left(\frac{\rho h - k}{\sqrt{1-\rho^2}}\right) + \rho e^{(h^2-k^2)/2}\Phi\left(\frac{\rho k - h}{\sqrt{1-\rho^2}}\right)\right).$$

Let $\boldsymbol{g} \sim N(0,I_n)$. Define the random variables

$$\boldsymbol{Z}_1 := \frac{\boldsymbol{g}\cdot x'}{\|x'\|} \qquad \text{and} \qquad \boldsymbol{Z}_2 := \frac{\boldsymbol{g}\cdot y'}{\|y'\|}$$

and set $h := \frac{r}{\|x'\|}$, $k := \frac{r}{\|y'\|}$. It is immediate that

$$\text{Vol}(\text{hfsp}(x')) = \mathbf{Pr}[\boldsymbol{Z}_1 > h] \quad \text{and} \quad \text{Vol}(\text{hfsp}(x') \cap \text{hfsp}(y')) = \mathbf{Pr}[\boldsymbol{Z}_1 > h, \boldsymbol{Z}_2 > k].$$

Furthermore, note that $\mathbf{Var}[\boldsymbol{Z}_1] = \mathbf{Var}[\boldsymbol{Z}_2] = 1$. We also have $\rho := \mathbf{E}[\boldsymbol{Z}_1\boldsymbol{Z}_2] = \frac{x'\cdot y'}{\|x'\|\|y'\|}$. Thanks to Claim 8, we have

$$(cn^{3/8}-1)^2 \leq \|x'-y'\|^2 = \|x'\|^2 + \|y'\|^2 - 2x'\cdot y' \leq 2(n - x'\cdot y')$$

which in turn implies that

$$\rho = \frac{x' \cdot y'}{\|x'\|\|y'\|} \leq \left(n - \frac{1}{2}(cn^{3/8} - 1)^2\right)\frac{1}{\|x'\|\|y'\|}. \tag{5.38}$$

Using Claim 8 and the fact that $x, y \in \mathrm{Shell}_{n+1}$, we have that $\|x'\|, \|y'\| \geq \sqrt{n+1} - \sqrt{2\ln(2/\tau)} - 1$ and combining this with Equation (5.38) gives

$$\rho \leq \frac{\left(n - \frac{1}{2}(cn^{3/8} - 1)^2\right)}{\left(\sqrt{n+1} - 2\sqrt{2\ln(2/\tau)} - 1\right)^2} = 1 - \Omega(n^{-1/4}). \tag{5.39}$$

Note that $\mathrm{Vol}(\mathrm{hfsp}(x')) = \Phi(-h)$. Consequently, using Proposition 5.3 we get

$$(5.37) \leq \Phi\left(\frac{\rho h - k}{\sqrt{1-\rho^2}}\right) + \rho e^{(h^2-k^2)/2}\Phi\left(\frac{\rho k - h}{\sqrt{1-\rho^2}}\right),$$

and we will obtain an upper bound on this in the remainder of this section. In particular, note that

$$\begin{aligned}\rho h - k &\leq \left(1 - \Omega(n^{-1/4})\right)\frac{r}{\|x'\|} - \frac{r}{\|y'\|} \\ &= \frac{r}{\|x'\|}\left(1 - \Omega(n^{-1/4}) - \frac{\|x'\|}{\|y'\|}\right)\end{aligned}$$

Recall that $\|x'\|, \|y'\| \leq \sqrt{n}$ and that $\|x'\| \geq \sqrt{n+1} - \sqrt{2\ln(2/\tau)} - 1$. Hence, for n large enough and an appropriate constant τ, we have

$$\frac{\|x'\|}{\|y'\|} \geq 1 - \Omega(n^{-1/2}) \tag{5.40}$$

and consequently, we get that

$$\rho h - k \leq O\left(\frac{r}{\|x'\| \cdot n^{1/4}}\left(O\left(\frac{1}{n^{1/4}}\right) - \Theta(1)\right)\right) \leq O\left(\frac{-r}{\|x'\| \cdot n^{1/4}}\right) \leq -\Omega(1)$$

for an appropriate choice of τ. The final inequality relies on the above lower bound on $\|x'\|$ and Lemma 3.1. An identical calculation gives that $\rho k - h \leq -\Omega(1)$. It follows that

$$\begin{aligned}(5.37) &\leq 2\exp\left(\frac{r^2}{\|x'\|^2}\left(1 - \frac{\|x'\|^2}{\|y'\|^2}\right)\right)\Phi\left(\frac{-\Omega(1)}{\sqrt{1-\rho^2}}\right) \\ &\leq 2\exp\big(o(1)\big)\Phi\left(\frac{-\Omega(1)}{\sqrt{1-\rho^2}}\right) &(5.41)\\ &\leq 2\exp\big(o(1)\big)\Phi\Big(-\Omega(n^{1/8})\Big) &(5.42)\\ &\leq 2^{-\Theta(n^{1/4})} &(5.43)\end{aligned}$$

where the final inequality relies on a standard Gaussian tail bound (cf. Proposition 2.1). To see Equation (5.41),

note that

$$\exp\left(\frac{r^2}{\|x'\|^2}\left(1-\frac{\|x'\|^2}{\|y'\|^2}\right)\right) \le \exp\left(\Theta\left(\frac{n^{3/2}\ln(1/c_1)}{\sqrt{n+1}-\sqrt{2\ln(2/\tau)}-1}\right)\left(1-\frac{\|x'\|^2}{\|y'\|^2}\right)\right)$$
$$\le \exp\left(\Theta\left(\frac{n^{3/2}\ln(1/c_1)}{\sqrt{n+1}-\sqrt{2\ln(2/\tau)}-1}\right)\cdot O(n^{-1})\right)$$
$$\le \exp\left(O\left(\frac{1}{\sqrt{n}}\right)\right),$$

where we used Lemma 3.1 and Equation (5.38). Equation (5.42) immediately follows from Equation (5.39). Finally, note that Equation (5.43) completes the proof.

6 Two-Sided Non-Adaptive Lower Bound

Our goal in this section is to prove Theorem 1.3 restated below:

THEOREM 6.1. (TWO-SIDED NON-ADAPTIVE LOWER BOUND) *For any constant $c > 0$, there is a constant $\varepsilon = \varepsilon_c > 0$ such that any non-adaptive ε-tester for convexity over $N(0, I_n)$ (which may make two-sided errors) must use at least $n^{1/4-c}$ queries.*

6.1 Setup

We recall some necessary tools and results from [CDST15].

6.1.1 Distributions with Matching Moments

The first results we need, stated below as Propositions 6.1 and 6.2, establish the existence of two finitely supported random variables that match the first ℓ moments of a univariate Gaussian, for any ℓ. Crucially, one of the random variables is supported entirely on non-negative reals, while the other puts nonzero probability on negative values (so if ℓ is any fixed constant, it puts a constant amount of probability on negative values):

PROPOSITION 6.1. ([CDST15] PROPOSITION 3.1) *Given an odd $\ell \in \mathbb{N}$, there exists a value $\mu = \mu(\ell) > 0$ and a real random variable $\boldsymbol{u}$ such that*

1. *$\boldsymbol{u}$ is supported on at most ℓ nonnegative real values; and*
2. *$\mathbf{E}[\boldsymbol{u}^k] = \mathbf{E}_{\boldsymbol{g}\sim N(\mu,I_k)}[\boldsymbol{g}^k]$ for all $k \in [\ell]$.*

PROPOSITION 6.2. ([CDST15] PROPOSITION 3.2) *Given $\mu > 0$ and $\ell \in \mathbb{N}$, there exists a real random variable $\boldsymbol{v}$ such that*

1. *$\boldsymbol{v}$ is supported on at most $\ell+1$ real values, with $\mathbf{Pr}[\boldsymbol{v} < 0] > 0$; and*
2. *$\mathbf{E}[\boldsymbol{v}^k] = \mathbf{E}_{\boldsymbol{g}\sim N(\mu,I_k)}[\boldsymbol{g}^k]$ for all $k \in [\ell]$.*

We will use $\boldsymbol{u}$ (respectively $\boldsymbol{v}$) to sample coefficients in our construction of the yes-distribution (respectively the no-distribution).

6.1.2 Mollifiers, CLTs, Tail Bounds and Other Tools

We recall the following basic proposition from [CDST15] and its simple proof:

PROPOSITION 6.3. ([CDST15] PROPOSITION 4.1) *Let $\mathcal{A}, \mathcal{A}_{in} \subseteq \mathbb{R}^q$ where $\mathcal{A}_{in} \subseteq \mathcal{A}$. Let $\Psi_{in} : \mathbb{R}^q \to [0,1]$ be a function satisfying $\Psi_{in}(X) = 1$ for all $X \in \mathcal{A}_{in}$ and $\Psi_{in}(X) = 0$ for all $X \notin \mathcal{A}$. Then for all random variables $\boldsymbol{S}, \boldsymbol{T}$:*

$$\left|\mathbf{Pr}[\boldsymbol{S} \in \mathcal{A}] - \mathbf{Pr}[\boldsymbol{T} \in \mathcal{A}]\right| \le \left|\mathbf{E}[\Psi_{in}(\boldsymbol{S})] - \mathbf{E}[\Psi_{in}(\boldsymbol{T})]\right| + \max\left\{\mathbf{Pr}[\boldsymbol{S} \in \mathcal{A} \setminus \mathcal{A}_{in}], \mathbf{Pr}[\boldsymbol{T} \in \mathcal{A} \setminus \mathcal{A}_{in}]\right\}.$$

Proof. Observe that $\mathbf{Pr}[\boldsymbol{S} \in \mathcal{A}] \geq \mathbf{E}[\Psi_{in}(\boldsymbol{S})]$ and $\mathbf{Pr}[\boldsymbol{S} \in \mathcal{A}] \leq \mathbf{E}[\Psi_{in}(\boldsymbol{S})] + \mathbf{Pr}[\boldsymbol{S} \in \mathcal{A} \setminus \mathcal{A}_{in}]$, and likewise for $\boldsymbol{T}$. As a result, we have

$$\mathbf{Pr}[\boldsymbol{S} \in \mathcal{A}] - \mathbf{Pr}[\boldsymbol{T} \in \mathcal{A}] \leq \mathbf{E}[\Psi_{in}(\boldsymbol{S})] + \mathbf{Pr}[\boldsymbol{S} \in \mathcal{A} \setminus \mathcal{A}_{in}] - \mathbf{E}[\Psi_{in}(\boldsymbol{T})], \quad \text{and}$$
$$\mathbf{Pr}[\boldsymbol{S} \in \mathcal{A}] - \mathbf{Pr}[\boldsymbol{T} \in \mathcal{A}] \geq \mathbf{E}[\Psi_{in}(\boldsymbol{S})] - \mathbf{Pr}[\boldsymbol{T} \in \mathcal{A} \setminus \mathcal{A}_{in}] - \mathbf{E}[\Psi_{in}(\boldsymbol{T})].$$

Combining these, we have the proposition. ∎

We adopt the following notation: for $J = (J_1, \ldots, J_q) \in \mathbb{N}^q$ a q-dimensional multi-index, we let $|J|$ denote $J_1 + \cdots + J_q$ and let $J!$ denote $J_1!J_2! \cdots J_q!$. We write $\#J$ to denote $|\{i \in [q] \colon J_i \neq 0\}|$ (and we observe that $\#J \leq |J|$). Given $X \in \mathbb{R}^q$ we write X^J to denote $\prod_{i=1}^q (X_i)^{J_i}$, and we write $X|_J \in \mathbb{R}^{\#J}$ to denote the projection of X onto the coordinates for which $J_i \neq 0$. For $f : \mathbb{R}^q \to \mathbb{R}$, we write $f^{(J)}$ to denote the J-th derivative, i.e.

$$f^{(J)} = \frac{\partial^{J_1 + \cdots + J_q} f}{\partial x_1^{J_1} \cdots \partial x_q^{J_q}}.$$

We recall the standard multivariate Taylor expansion:

Fact 6.1. (Multivariate Taylor expansion) *Given a smooth function $f : \mathbb{R}^q \to \mathbb{R}$ and $k \in \mathbb{N}$,*

$$f(X + \Delta) = \sum_{|J| \leq k} \frac{f^{(J)}(X)}{J!} \cdot \Delta^J + (k+1) \sum_{|J| = k+1} \left(\frac{\Delta^J}{J!} \mathbf{E}\left[(1 - \boldsymbol{\tau})^k f^{(J)}(X + \boldsymbol{\tau}\Delta) \right] \right),$$

for $X, \Delta \in \mathbb{R}^q$, where $\boldsymbol{\tau}$ is uniform random over the interval $[0, 1]$.

We recall the standard Berry–Esseen theorem for sums of independent real random variables (see for example, [Fel68]), which is a quantitative form of the Central Limit Theorem:

Theorem 6.2. (Berry–Esseen) *Let $\boldsymbol{s} = \boldsymbol{x}_1 + \cdots + \boldsymbol{x}_n$, where $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ are independent real-valued random variables with $\mathbf{E}[\boldsymbol{x}_j] = \mu_j$ and $\mathbf{Var}[\boldsymbol{x}_j] = \sigma_j^2$, and $\sum_{i=1}^n \mathbf{E}\left[|\boldsymbol{x}_i|^3\right] \leq \kappa$. Let $\boldsymbol{g}$ denote a Gaussian random variable with mean $\sum_{j=1}^n \mu_j$ and variance $\sum_{j=1}^n \sigma_j^2$, matching those of $\boldsymbol{s}$. Then for all $\theta \in \mathbb{R}$, we have*

$$\left| \mathbf{Pr}[\boldsymbol{s} \leq \theta] - \mathbf{Pr}[\boldsymbol{g} \leq \theta] \right| \leq \frac{O(\kappa)}{\sqrt{\sum_{j=1}^n \sigma_j^2}}.$$

For $\boldsymbol{g} \sim N(0, I_n)$, the value $\sum_{i=1}^n \boldsymbol{g}_i^2$ is distributed according to a chi-squared distribution with n degrees of freedom, denoted $\chi(n)^2$. We recall the following tail bound:

Lemma 6.1. (Tail bound for the chi-squared distribution, from [Joh01]) *Let $\boldsymbol{X} \sim \chi(n)^2$. Then we have*

$$\mathbf{Pr}\left[|\boldsymbol{X} - n| \geq tn \right] \leq e^{-(3/16)nt^2}, \quad \text{for all } t \in [0, 1/2).$$

Following [CDST15], our proof will employ a carefully chosen "mollifier," i.e. a particular smooth function which approximates the indicator function of a set (the use of such mollifiers is standard in Lindeberg-type "replacement method" analyses). We will use a specific mollifier, given in [CDST15], whose properties are tailored to our sets of interest (unions of orthants). The key properties of this mollifier are as follows:

Proposition 6.4. ([CDST15] Proposition 4.3: "product mollifier") *Let $\mathcal{O}$ be a union of orthants in $\mathbb{R}^q$. For all $\varepsilon > 0$, there exists a smooth function $\Psi_{\mathcal{O}} : \mathbb{R}^q \to [0, 1]$ with the following properties:*

1. *$\Psi_{\mathcal{O}}(X) = 0$ for all $X \notin \mathcal{O}$.*
2. *$\Psi_{\mathcal{O}}(X) = 1$ for all $X \in \mathcal{O}$ with $\min_i\{|X_i|\} \geq \varepsilon$.*
3. *For any multi-index $J \in \mathbb{N}^q$ such that $|J| = k$, $\|\Psi_{\mathcal{O}}^{(J)}\|_\infty \leq \alpha(k) \cdot (1/\varepsilon)^k$, where $\alpha(k) = k^{O(k)}$.*
4. *For any $J \in \mathbb{N}^q$, $\Psi_{\mathcal{O}}^{(J)}(X) \neq 0$ only if $X \in \mathcal{O}$ and $|X_i| \leq \varepsilon$ for all i such that $J_i \neq 0$. Equivalently, $\Psi_{\mathcal{O}}^{(J)}(X) \neq 0$ only if $X \in \mathcal{O}$ and $\|X|_J\|_\infty \leq \varepsilon$.*

6.1.3 Clipping Given $C > 0$, we define the "clipping" function $\mathrm{clip}_C : \mathbb{R}^n \to \{0,1\}$ which, on input a vector $x \in \mathbb{R}^n$, outputs 1 if and only if $\|x\| \le \sqrt{n} + C$.

6.2 The Yes- and No- Distributions Let $c > 0$ (this is the c of Theorem 1.3). Let $\boldsymbol{u}$ and $\boldsymbol{v}$ be the random variables given by Propositions 6.1 and 6.2, where we take ℓ to be the smallest odd integer that is at least $1/c$ and take $\mu = \mu(\ell)$.

A set $\boldsymbol{K}$ drawn from our "yes-distribution" $\mathcal{D}_{\mathrm{yes}}$ has indicator function defined as follows:

- First, choose a Haar random orthonormal basis normalized so that each vector has Euclidean length $1/\sqrt{n}$, and denote those vectors $\boldsymbol{a}^{(1)}, \dots, \boldsymbol{a}^{(n)}$. (So $\boldsymbol{a}^{(1)} \in \mathbb{R}^n$ is a Haar random unit vector in $\mathbb{R}^n$ scaled by $1/\sqrt{n}$; $\boldsymbol{a}^{(2)}$ is Haar random over the radius-$(1/\sqrt{n})$ sphere in the $(n-1)$-dimensional subspace of $\mathbb{R}^n$ that is orthogonal to $\boldsymbol{a}^{(1)}$; and so on.)
- Then n independent draws $\boldsymbol{u}_1, \dots, \boldsymbol{u}_n$ are made from the real random variable $\boldsymbol{u}$ of Proposition 6.1.
- The indicator function $\boldsymbol{K}(x)$ is

$$\boldsymbol{K}(x) = \mathbf{1}\Big[\boldsymbol{u}_1(\boldsymbol{a}^{(1)} \cdot x)^2 + \cdots + \boldsymbol{u}_n(\boldsymbol{a}^{(n)} \cdot x)^2 \le \mu \ \&\ \mathrm{clip}_C(x) = 1\Big]. \tag{6.44}$$

(Here $C > 0$ is a suitable constant, depending only on c but not on n, that will be fixed later in our argument.)

A set $\boldsymbol{K}$ drawn from our "no-distribution" $\mathcal{D}_{\mathrm{no}}$ is defined very similarly, with the only difference being that $\boldsymbol{v}$ takes the place of $\boldsymbol{u}$:

- The vectors $\boldsymbol{a}^{(1)}, \dots, \boldsymbol{a}^{(n)}$ are chosen exactly as in the yes-case.
- Then n independent draws $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ are made from the real random variable $\boldsymbol{v}$ of Proposition 6.2.
- The indicator function $\boldsymbol{K}(x)$ is

$$\boldsymbol{K}(x) = \mathbf{1}\Big[\boldsymbol{v}_1(\boldsymbol{a}^{(1)} \cdot x)^2 + \cdots + \boldsymbol{v}_n(\boldsymbol{a}^{(n)} \cdot x)^2 \le \mu \ \&\ \mathrm{clip}_C(x) = 1\Big]. \tag{6.45}$$

(Here $C > 0$ is the same constant as in the yes-case.)

We remark that our yes- and no- functions differ from the yes- and no- functions of [CDST15] in a number of ways: Our functions involve a random orthonormal basis, they are degree-2 polynomial threshold functions rather than linear threshold functions, and they involve clipping. (In contrast the [CDST15] functions do not involve choosing a random orthonormal basis, are LTFs, and do not incorporate any clipping.)

6.2.1 Distance to Convexity We first consider yes-functions. Since $\boldsymbol{u}$ is supported on non-negative real values and $\boldsymbol{a}^{(1)}, \dots, \boldsymbol{a}^{(n)}$ are orthogonal vectors, every outcome of $\mathbf{1}\Big[\boldsymbol{u}_1(\boldsymbol{a}^{(1)} \cdot x)^2 + \cdots + \boldsymbol{u}_n(\boldsymbol{a}^{(n)} \cdot x)^2 \le \mu\Big]$ is an ellipsoid in $\mathbb{R}^n$. Since $\mathrm{clip}_C(x)$ is the indicator function of a ball in $\mathbb{R}^n$, and the intersection of a ball and an ellipsoid is a convex set, we immediately have the following:

COROLLARY 6.1. *For every $C > 0$, every $K \subset \mathbb{R}^n$ in the support of $\mathcal{D}_{\mathrm{yes}}$ is convex.*

The following lemma shows that a constant fraction of draws of $\boldsymbol{K} \sim \mathcal{D}_{\mathrm{no}}$ are constant-far from being convex (intuitively, this is because with extremely high probability a constant fraction of the coefficients $\boldsymbol{v}_1, \dots, \boldsymbol{v}_n$ are negative, which causes the degree-2 PTF to be far from an ellipsoid):

LEMMA 6.2. *For a suitable choice of the constant $C > 0$, with probability at least $1/2$ a random $\boldsymbol{K} \sim \mathcal{D}_{\mathrm{no}}$ is κ-far from convex (where $\kappa > 0$ depends on μ and ℓ and hence only on c).*

Proof. By the rotational symmetry of the $N(0, I_n)$ distribution, we may assume that the orthonormal basis $\boldsymbol{a}^{(1)}, \dots, \boldsymbol{a}^{(n)}$ is the canonical basis $e_1, \dots, e_n$ scaled by $1/\sqrt{n}$. Thus a draw of $\boldsymbol{K} \sim \mathcal{D}_{\mathrm{no}}$ (after a suitable rotation) is

$$\boldsymbol{K}(x) = \mathbf{1}\Big[\boldsymbol{v}_1 x_1^2 + \cdots + \boldsymbol{v}_n x_n^2 \le n\mu \ \&\ \mathrm{clip}_C(x) = 1\Big].$$

Given this, it suffices to show that a random set

$$\boldsymbol{K}' := \mathbf{1}\Big[\boldsymbol{v}_1 x_1^2 + \cdots + \boldsymbol{v}_n x_n^2 \le n\mu\Big] \tag{6.46}$$

is 2κ-far from convex with probability at least $1/2$. If we have this, then since $\boldsymbol{K}$ has distance at most κ from $\boldsymbol{K}'$ (which holds for a suitable choice of the constant C, using Lemma 6.1), the lemma follows.

To analyze Equation (6.46), we begin by recalling that by Proposition 6.2, the random variable $\boldsymbol{v}$ has probability $p_i > 0$ of taking value d_i for $1 \le i \le \ell'$, where ℓ' is some value that is at most $\ell + 1$, and we have $d_1 < 0$, $d_1 < d_2 < \cdots < d_{\ell'}$, and $p_1 + \cdots + p_{\ell'} = 1$. Taking $k = 1$ in item (2) of Proposition 6.2, we have

$$p_1 d_1 + \cdots + p_{\ell'} d_{\ell'} = \mu. \tag{6.47}$$

For $i \in [\ell']$, let $\boldsymbol{n}_i$ denote the number of indices $j \in [n]$ such that $\boldsymbol{v}_j = d_i$. Since all of the values $p_1, \dots, p_{\ell'}$ are constants independent of the asymptotic parameter n, by a standard Chernoff bound and union bound, we have that for suitable constants $c_1, \dots, c_{\ell'} > 0$ (which depend on the p_i's),

$$\Pr_{\boldsymbol{v}_1,\dots,\boldsymbol{v}_n}\left[\boldsymbol{n}_i \in [p_i n - c_i\sqrt{n}, p_i n + c_i\sqrt{n}] \text{ for each } i \in [\ell']\right] \ge 1/2. \tag{6.48}$$

Fix any outcome $(v_1, \dots, v_n)$ of $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ such that the event on the LHS of Equation (6.48) is satisfied. In the rest of the proof we will argue that for such an outcome the set

$$K' = \mathbf{1}\Big[v_1 x_1^2 + \cdots + v_n x_n^2 \le n\mu\Big] \tag{6.49}$$

corresponding to Equation (6.46) is $\Omega(1)$-far from convex.

For each $i \in [\ell']$, let $S_i \subset [n]$ denote the set of indices $j \in [n]$ such that $v_j = d_i$. Let c_i' be such that $|S_i| = p_i n + c_i'\sqrt{n}$, and observe that $|c_i'| \le c_i$. For ease of notation we may suppose that S_1 consists of the first coordinates $\{1, \dots, p_1 n + c_1'\sqrt{n}\}$ (this is without loss of generality by the rotational invariance of $N(0, I_n)$).

Fix any $i \in \{2, \dots, \ell'\}$ and consider the tuple of random Gaussian coordinates $(\boldsymbol{x}_j)_{j \in S_i}$ for a draw of $\boldsymbol{x} = (\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \sim N(0, I_n)$. We have

$$\mathop{\mathbf{E}}_{\boldsymbol{x}}\left[\sum_{j \in S_i} \boldsymbol{x}_j^2\right] = p_i n + c_i'\sqrt{n},$$

and by the Berry-Esseen theorem (Theorem 6.2), we get that

$$\mathbf{Pr}\left[\sum_{j \in S_i} \boldsymbol{x}_j^2 \in \left[p_i n - A_i\sqrt{n}, p_i n + A_i\sqrt{n}\right]\right] \ge 1 - \frac{1}{10\ell'} \tag{6.50}$$

for suitable positive absolute constants $A_2, \dots, A_{\ell'}$ (depending on the p_i's and the c_i''s but not on n).

Let $A := \ell' \cdot \max\{|d_2|, \dots, |d_{\ell'}|\} \cdot \max\{A_2, \dots, A_{\ell'}\}$.. By a union bound applied to Equation (6.50) over all $i \in \{2, \dots, \ell'\}$, with probability at least $9/10$ we have that

$$\sum_{i=2}^{\ell'} \sum_{j \in S_i} d_i \boldsymbol{x}_j^2 \in \left[\left(\sum_{i=2}^{\ell'} d_i p_i n\right) - A\sqrt{n}, \left(\sum_{i=2}^{\ell'} d_i p_i n\right) + A\sqrt{n}\right]; \tag{6.51}$$

let us say that any such outcome of $(\boldsymbol{x}_j)_{j \in S_2 \cup \cdots \cup S_{\ell'}}$ is *good*. Fix any good outcome $(x_j)_{j \in S_2 \cup \cdots \cup S_{\ell'}}$ of the last $n - (p_1 n + c_1'\sqrt{n})$ coordinates of $\boldsymbol{x} \sim N(0, I_n)$, and let $A' \in [-A, A]$ be the value such that the LHS of Equation (6.51) is equal to $\left(\sum_{i=2}^{\ell'} d_i p_i n\right) + A'\sqrt{n}$. Recalling Equation (6.47), for this good outcome of the last $n - (p_1 n + c_1'\sqrt{n})$ coordinates, the set (6.49) (viewed as an indicator function of coordinates $1, \dots, p_1 n + c_1'\sqrt{n}$) becomes

$$\mathbf{1}\left[\sum_{j=1}^{p_1 n + c_1'\sqrt{n}} d_1 x_j^2 \le p_1 d_1 n - A'\sqrt{n}\right]. \tag{6.52}$$

Recalling that $d_1 < 0$, this is equivalent to

$$\mathbf{1}\left[\sum_{j=1}^{p_1 n + c_1'\sqrt{n}} x_j^2 \geq p_1 n - (A'/d_1)\sqrt{n}\right]. \tag{6.53}$$

Let q denote the probability that (6.53) holds for independent standard Gaussians $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{p_1 n + c_1'\sqrt{n}}$; the Berry-Esseen theorem implies that q is a constant in $(0,1)$ which is bounded away from both 0 and 1. By the radial symmetry of the $N(0,1)^{p_1 n + c_1'\sqrt{n}}$ distribution, it follows that the subset of $\mathbb{R}^{p_1 n + c_1'\sqrt{n}}$ whose indicator function is given by Equation (6.53) is $\Omega(1)$-far from convex, because it is $\Omega(1)$-far from convex on a "line by line" basis. In more detail, for each unit vector $v \in \mathbb{R}^{p_1 n + c_1'\sqrt{n}}$, the function (6.53) labels points on the corresponding line $\{tv : t \in \mathbb{R}\}$ as follows:

(i) if $|t| \geq \sqrt{p_1 n - (A'/d_1)\sqrt{n}}$ then (6.53) outputs 1 on tv;

(ii) if $|t| > \sqrt{p_1 n - (A'/d_1)\sqrt{n}}$ then (6.53) outputs 0 on tv.

Since this labeling corresponds to the complement of an interval, and since both (i) and (ii) have constant probability as explained above, the distance to convexity is $\Omega(1)$, and the proof of Lemma 6.2 is complete. ∎

6.3 Proof of Theorem 1.3 As is usual for a non-adaptive lower bound, we use Yao's principle. Let $\mathcal{X}$ be a $q \times n$ query matrix, so the i-th row $\mathcal{X}_{i*} = (\mathcal{X}_{i1}, \ldots, \mathcal{X}_{in})$ is a vector in $\mathbb{R}^n$ corresponding to the i-th query made by some deterministic algorithm. We will argue that the behavior of such a deterministic algorithm will be almost the same on a target function $\boldsymbol{K} \sim \mathcal{D}_{\text{yes}}$ and on a target function $\boldsymbol{K} \sim \mathcal{D}_{\text{no}}$.

First, since our analysis will only consider target functions drawn from $\mathcal{D}_{\text{yes}}$ and $\mathcal{D}_{\text{no}}$, and any draw from either of these distributions always involves clipping (the clip_C component of Equations (6.44) and (6.45)), we may suppose without loss of generality that each query vector $\mathcal{X}_{i*}$ has $\|\mathcal{X}_{i*}\| \leq \sqrt{n} + C$, i.e. it satisfies $\text{clip}_C(\mathcal{X}_{i*}) = 1$.

Let $\boldsymbol{R}_{\text{yes}}$ be the $\{0,1\}^q$-valued random variable obtained by drawing $\boldsymbol{K} \sim \mathcal{D}_{\text{yes}}$ (recall that this corresponds to drawing $\boldsymbol{u}_1, \boldsymbol{a}^{(1)}, \ldots, \boldsymbol{u}_n, \boldsymbol{a}^{(n)}$) and setting the t-th coordinate of $\boldsymbol{R}_{\text{yes}}$ to be

$$\boldsymbol{K}(\mathcal{X}_{t*}) = \mathbf{1}\Big[\boldsymbol{u}_1(\boldsymbol{a}^{(1)} \cdot \mathcal{X}_{t*})^2 + \cdots + \boldsymbol{u}_n(\boldsymbol{a}^{(n)} \cdot \mathcal{X}_{t*})^2 \leq \mu\Big].$$

Similarly, let $\boldsymbol{R}_{\text{no}}$ be the $\{0,1\}^q$-valued random variable obtained by drawing $\boldsymbol{K} \sim \mathcal{D}_{\text{no}}$ (recall that this corresponds to drawing $\boldsymbol{v}_1, \boldsymbol{a}^{(1)}, \ldots, \boldsymbol{v}_n, \boldsymbol{a}^{(n)}$) and setting the t-th coordinate of $\boldsymbol{R}_{\text{no}}$ to be

$$\boldsymbol{K}(\mathcal{X}_{t*}) = \mathbf{1}\Big[\boldsymbol{v}_1(\boldsymbol{a}^{(1)} \cdot \mathcal{X}_{t*})^2 + \cdots + \boldsymbol{v}_n(\boldsymbol{a}^{(n)} \cdot \mathcal{X}_{t*})^2 \leq \mu\Big].$$

To prove a two-sided non-adaptive lower bound of q queries, it suffices to show that for the $\boldsymbol{R}_{\text{yes}}, \boldsymbol{R}_{\text{no}}$ defined above, we have $d_{\text{TV}}(\boldsymbol{R}_{\text{yes}}, \boldsymbol{R}_{\text{no}}) = o(1)$.

Let us write $\overline{a}$ to denote $\overline{a} = (a^{(1)}, \ldots, a^{(n)})$, and let us write $\boldsymbol{R}_{\text{yes}}^{\overline{a}}$ to denote the random variable $\boldsymbol{R}_{\text{yes}}$ conditioned on having the outcome of $\boldsymbol{a}^{(1)}, \ldots, \boldsymbol{a}^{(n)}$ come out equal to $\overline{a}$, and similarly for $\boldsymbol{R}_{\text{no}}^{\overline{a}}$. Using the coupling interpretation of total variation distance and the natural coupling between $\mathcal{D}_{\text{yes}}$ and $\mathcal{D}_{\text{no}}$, we have that

$$d_{\text{TV}}(\boldsymbol{R}_{\text{yes}}, \boldsymbol{R}_{\text{no}}) \leq \mathop{\mathbf{E}}_{\overline{\boldsymbol{a}} \sim \text{Haar}}\Big[d_{\text{TV}}(\boldsymbol{R}_{\text{yes}}^{\overline{\boldsymbol{a}}}, \boldsymbol{R}_{\text{no}}^{\overline{\boldsymbol{a}}})\Big], \tag{6.54}$$

so it suffices to upper bound the RHS of Equation (6.54) by $o(1)$.

Let us say that an outcome $\overline{a} = (a^{(1)}, \ldots, a^{(n)}) \in (\mathbb{R}^n)^n$ is *bad* if there is a pair $(t,j) \in [q] \times [n]$ such that $(a^{(j)} \cdot \mathcal{X}_{t*})^2 \geq \frac{10 \ln n}{n}$. Recalling that each query vector $\mathcal{X}_{t*}$ has norm at most $\sqrt{n} + C$ and that each $\boldsymbol{a}^{(j)}$ is a Haar random unit vector scaled by $1/\sqrt{n}$, it is easy to show that bad outcomes of $\overline{\boldsymbol{a}}$ have very low probability:

LEMMA 6.3. $\mathbf{Pr}[\overline{\boldsymbol{a}}$ *is bad*$] = o(1)$.

Proof. Fix some pair $(t,j) \in [q] \times [n]$ and let $r \leq \sqrt{n} + C$ be the norm of the query vector $\mathcal{X}_{t*}$. The distribution of $\boldsymbol{a}^{(j)} \cdot \mathcal{X}_{t*}$ is precisely the distribution of the first coordinate of a Haar random point drawn from the n-dimensional

sphere of radius $r/\sqrt{n}$. Hence, writing $\boldsymbol{u} \sim \mathbb{S}^{n-1}$ to denote a Haar random point from the n-dimensional unit sphere, we have

$$\begin{aligned}
\mathbf{Pr}\left[(\boldsymbol{a}^{(j)} \cdot \mathcal{X}_{t*})^2 \geq \frac{10 \ln n}{n}\right] &= \mathop{\mathbf{Pr}}_{\boldsymbol{u} \sim \mathbb{S}^{n-1}}\left[\frac{|\boldsymbol{u}_1| r}{\sqrt{n}} \geq \frac{\sqrt{10 \ln n}}{\sqrt{n}}\right] \\
&\leq \mathop{\mathbf{Pr}}_{\boldsymbol{u} \sim \mathbb{S}^{n-1}}\left[\boldsymbol{u}_1 \geq \frac{\sqrt{10 \ln n}}{r}\right] \\
\text{(using } r \leq \sqrt{n} + C\text{)} \qquad &\leq \mathop{\mathbf{Pr}}_{\boldsymbol{u} \sim \mathbb{S}^{n-1}}\left[\boldsymbol{u}_1 \geq \frac{3\sqrt{\ln n}}{\sqrt{n}}\right] \\
&\leq e^{-(9/2)\ln n} = 1/n^{9/2},
\end{aligned}$$

using a standard bound on spherical caps (see Lemma 2.1). Since there are only $qn < n^{5/4}$ many pairs $(t, j) \in [q] \times [n]$, a union bound concludes the proof. ∎

Fix $\overline{a} = (a^{(1)}, \ldots, a^{(n)})$ to be any non-bad outcome of $\overline{\boldsymbol{a}}$. Recalling Equation (6.54), by Lemma 6.3 it suffices to show that $d_{\mathrm{TV}}(\boldsymbol{R}^{\overline{a}}_{\mathrm{yes}}, \boldsymbol{R}^{\overline{a}}_{\mathrm{no}}) \leq o(1)$; this is our goal in the rest of the proof.

Let $\boldsymbol{S} \in \mathbb{R}^q$ be the random column vector whose t-th entry is

$$\boldsymbol{u}_1 (a^{(1)} \cdot \mathcal{X}_{t*})^2 + \cdots + \boldsymbol{u}_n (a^{(n)} \cdot \mathcal{X}_{t*})^2 - \mu,$$

and let $\boldsymbol{T} \in \mathbb{R}^q$ be the random column vector whose t-th entry is

$$\boldsymbol{v}_1 (a^{(1)} \cdot \mathcal{X}_{t*})^2 + \cdots + \boldsymbol{v}_n (a^{(n)} \cdot \mathcal{X}_{t*})^2 - \mu.$$

The response vector $\boldsymbol{R}^{\overline{a}}_{\mathrm{yes}}$ is determined by the orthant of $\mathbb{R}^q$ in which $\boldsymbol{S}$ lies and the response vector $\boldsymbol{R}^{\overline{a}}_{\mathrm{no}}$ is determined by the orthant of $\mathbb{R}^q$ in which $\boldsymbol{T}$ lies. So to prove a q-query monotonicity testing lower bound for non-adaptive algorithms, it suffices to upper bound

$$d_{\mathrm{UO}}(\boldsymbol{S}, \boldsymbol{T}) \leq o(1), \tag{6.55}$$

where d_{UO} is the "union-of-orthants" distance:

$$d_{\mathrm{UO}}(\boldsymbol{S}, \boldsymbol{T}) := \max\Big\{ |\,\mathbf{Pr}[\boldsymbol{S} \in \mathcal{O}] - \mathbf{Pr}[\boldsymbol{T} \in \mathcal{O}]| \colon \mathcal{O} \text{ is a union of orthants in } \mathbb{R}^q \Big\}.$$

In what follows, we will show that $d_{\mathrm{UO}}(\boldsymbol{S}, \boldsymbol{T}) \leq o(1)$ when $q = O(n^{1/4-c})$. To this end, let $\mathcal{O}$ denote a union of orthants such that

$$d_{\mathrm{UO}}(\boldsymbol{S}, \boldsymbol{T}) = \big|\mathbf{Pr}[\boldsymbol{S} \in \mathcal{O}] - \mathbf{Pr}[\boldsymbol{T} \in \mathcal{O}]\big|. \tag{6.56}$$

Following [Mos08, GOWZ10, CDST15], we first use the Lindeberg replacement method to bound

$$\big|\mathbf{E}[\Psi_{\mathcal{O}}(\boldsymbol{S})] - \mathbf{E}[\Psi_{\mathcal{O}}(\boldsymbol{T})]\big|, \tag{6.57}$$

and then apply Proposition 6.3 to bound (6.56).

For all $i \in \{0, 1 \ldots, n\}$ we introduce the $\mathbb{R}^q$-valued hybrid random variable $\boldsymbol{Q}^{(i)}$ whose t-th coordinate is

$$(\boldsymbol{Q}^{(i)})_t = \sum_{j=1}^{i} \boldsymbol{v}_j (a^{(j)} \cdot \mathcal{X}_{t*})^2 + \sum_{j=i+1}^{n} \boldsymbol{u}_j (a^{(j)} \cdot \mathcal{X}_{t*})^2.$$

Observe that $\boldsymbol{Q}^{(0)} = \boldsymbol{S}$ and $\boldsymbol{Q}^{(n)} = \boldsymbol{T}$. Informally, we are considering a sequence of hybrid distributions between $\boldsymbol{S}$ and $\boldsymbol{T}$ obtained by swapping out each of the $\boldsymbol{u}$-summands for a corresponding $\boldsymbol{v}$-summand one by one. The main idea is to bound the difference in expectations

$$\big|\mathbf{E}[\Psi_{\mathcal{O}}(\boldsymbol{Q}^{(i-1)})] - \mathbf{E}[\Psi_{\mathcal{O}}(\boldsymbol{Q}^{(i)})]\big| \quad \text{for each } i, \tag{6.58}$$

since summing (6.58) over all $i \in [n]$ gives an upper bound on

$$\left|\mathbf{E}[\Psi_{\mathcal{O}}(\boldsymbol{S})] - \mathbf{E}[\Psi_{\mathcal{O}}(\boldsymbol{T})]\right| = \left|\mathbf{E}[\Psi_{\mathcal{O}}(\boldsymbol{Q}^{(0)})] - \mathbf{E}[\Psi_{\mathcal{O}}(\boldsymbol{Q}^{(n)})]\right| \le \sum_{i=1}^{n} \left|\mathbf{E}[\Psi_{\mathcal{O}}(\boldsymbol{Q}^{(i-1)})] - \mathbf{E}[\Psi_{\mathcal{O}}(\boldsymbol{Q}^{(i)})]\right|$$

using the triangle inequality.

To bound (6.58), we define the $\mathbb{R}^q$-valued random variable $\mathbf{R}_{-i}$ whose t-th coordinate is

$$(\mathbf{R}_{-i})_t = \sum_{j=1}^{i-1} \boldsymbol{v}_j (a^{(j)} \cdot \mathcal{X}_{t*})^2 + \sum_{j=i+1}^{n} \boldsymbol{u}_j (a^{(j)} \cdot \mathcal{X}_{t*})^2. \tag{6.59}$$

Writing $\boldsymbol{\Phi}(\boldsymbol{v}_i, a^{(i)})$ to denote the random vector in $\mathbb{R}^q$ whose t-th coordinate is $\boldsymbol{v}_i (a^{(i)} \cdot \mathcal{X}_{t*})^2$ and likewise for $\boldsymbol{\Phi}(\boldsymbol{u}_i, a^{(i)})$, we have that

$$\left|\mathbf{E}[\Psi_{\mathcal{O}}(\boldsymbol{Q}^{(i-1)})] - \mathbf{E}[\Psi_{\mathcal{O}}(\boldsymbol{Q}^{(i)})]\right| = \left|\mathbf{E}[\Psi_{\mathcal{O}}(\mathbf{R}_{-i} + \boldsymbol{\Phi}(\boldsymbol{v}_i, a^{(i)}))] - \mathbf{E}[\Psi_{\mathcal{O}}(\mathbf{R}_{-i} + \boldsymbol{\Phi}(\boldsymbol{u}_i, a^{(i)}))]\right|.$$

Truncating the Taylor expansion of $\Psi_{\mathcal{O}}$ at the ℓ-th term (Fact 6.1), we get

$$\begin{aligned} \mathbf{E}\left[\Psi_{\mathcal{O}}(\mathbf{R}_{-i} + \boldsymbol{\Phi}(\boldsymbol{v}_i, a^{(i)}))\right] = &\sum_{|J| \le \ell} \frac{1}{J!} \cdot \mathbf{E}\left[\Psi_{\mathcal{O}}^{(J)}(\mathbf{R}_{-i}) \cdot (\boldsymbol{\Phi}(\boldsymbol{v}_i, a^{(i)}))^J\right] \\ &+ \sum_{|J| = \ell+1} \frac{\ell+1}{J!} \cdot \mathbf{E}\left[(1-\boldsymbol{\tau})^{\ell} \Psi_{\mathcal{O}}^{(J)}(\mathbf{R}_{-i} + \boldsymbol{\tau}\boldsymbol{\Phi}(\boldsymbol{v}_i, a^{(i)}))(\boldsymbol{\Phi}(\boldsymbol{v}_i, a^{(i)}))^J\right] \end{aligned} \tag{6.60}$$

where $\boldsymbol{\tau}$ is a random variable uniformly distributed on the interval $[0,1]$ (so the very last expectation is with respect to $\boldsymbol{\tau}$, $\boldsymbol{v}_i$ and $\mathbf{R}_{-i}$). Writing the analogous expression for $\mathbf{E}[\Psi_{\mathcal{O}}(\boldsymbol{R}_{-i} + \boldsymbol{\Phi}(\boldsymbol{v}_i, a^{(i)}))]$, we observe that by Propositions 6.1 and 6.2 the first sums are equal term by term, i.e. we have

$$\sum_{|J| \le \ell} \frac{1}{J!} \cdot \mathbf{E}\left[\Psi_{\mathcal{O}}^{(J)}(\mathbf{R}_{-i}) \cdot (\boldsymbol{\Phi}(\boldsymbol{v}_i, a^{(i)}))^J\right] = \sum_{|J| \le \ell} \frac{1}{J!} \cdot \mathbf{E}\left[\Psi_{\mathcal{O}}^{(J)}(\mathbf{R}_{-i}) \cdot (\boldsymbol{\Phi}(\boldsymbol{u}_i, a^{(i)}))^J\right]$$

for each $|J| \le h$. Thus we may cancel all but the last terms to obtain

$$\left|\mathbf{E}[\Psi_{\mathcal{O}}(\boldsymbol{Q}^{(i-1)})] - \mathbf{E}[\Psi_{\mathcal{O}}(\boldsymbol{Q}^{(i)})]\right| \le \sum_{|J| = \ell+1} \frac{\ell+1}{J!} \|\Psi_{\mathcal{O}}^{(J)}\|_{\infty} \left(\mathbf{E}\left[|(\boldsymbol{\Phi}(\boldsymbol{v}_i, a^{(i)}))^J|\right] + \mathbf{E}\left[|(\boldsymbol{\Phi}(\boldsymbol{u}_i, a^{(i)}))^J|\right]\right).$$

Observe that there are $|\{J \in \mathbb{N}^q : |J| = \ell+1\}| = \Theta(q^{\ell+1})$ many terms in this sum. Recalling that each value of $(a^{(j)} \cdot \mathcal{X}_{t*})^2$ is at most $\frac{10 \log n}{n}$ (because $\bar{a}$ is not bad), that both $\boldsymbol{u}_i$ and $\boldsymbol{v}_i$ are supported on at most $\ell+1$ real values that depend only on ℓ (by Propositions 6.1 and 6.2), and Proposition 6.4, we have that for any $\tau > 0$ (we will choose a value for τ soon),

$$\left|\mathbf{E}[\Psi_{\mathcal{O}}(\boldsymbol{Q}^{(i-1)})] - \mathbf{E}[\Psi_{\mathcal{O}}(\boldsymbol{Q}^{(i)})]\right| = O_\ell(1) \cdot \left(\frac{q}{\tau}\right)^{\ell+1} \cdot \left(\frac{10 \log n}{n}\right)^{(\ell+1)/2}. \tag{6.61}$$

Summing over all $i \in [n]$ costs us a factor of n and so we get

$$\left|\mathbf{E}[\Psi_{\mathcal{O}}(\boldsymbol{S})] - \mathbf{E}[\Psi_{\mathcal{O}}(\boldsymbol{T})]\right| = O_\ell(1) \cdot \left(\frac{q}{\tau}\right)^{\ell+1} \cdot \frac{(10 \log n)^{(\ell+1)/2}}{n^{(\ell-1)/2}}. \tag{6.62}$$

Equation (6.62) gives us the desired bound on Equation (6.57); it remains only to apply Proposition 6.3 to finish the argument. To do this, let

$$\mathcal{B}_\tau = \left\{X \in \mathcal{O} : |X_i| \le \tau \text{ for some } i \in [q]\right\}$$

($\mathcal{B}_\tau$ corresponds to the region $\mathcal{A} \setminus \mathcal{A}_{in}$ of Proposition 6.3). Since both $\boldsymbol{v}$ and $\boldsymbol{u}$ are supported on values of magnitude $O_\ell(1)$, using the one-dimensional Berry-Esseen inequality (Theorem 6.2) and a union bound across the q coordinates we get that

$$\mathbf{Pr}[\boldsymbol{S} \in \mathcal{B}_\tau], \mathbf{Pr}[\boldsymbol{T} \in \mathcal{B}_\tau] \leq O_\ell(q\tau) + O_\ell(q/\sqrt{n}). \tag{6.63}$$

So by applying Proposition 6.3, we get that

$$d_{\mathrm{UO}}(\boldsymbol{S}, \boldsymbol{T}) \leq O_\ell(q\tau) + O_\ell(q/\sqrt{n}) + O_\ell(1) \cdot \left(\frac{q}{\tau}\right)^{\ell+1} \cdot \frac{(10 \log n)^{(\ell+1)/2}}{n^{(\ell-1)/2}}.$$

Choosing $\tau = 1/n^{1/4}$ and recalling that ℓ is the smallest odd integer that is at least $1/c$, we get that for $q = O(n^{1/4-c})$ the RHS above is $O_\ell((10 \log n)^{(\ell+1)/2} n^{-c})$. This is $o(1)$ for any constants $c > 0, \ell \in \mathbb{N}$, and the proof of Theorem 1.3 is complete.

Acknowledgements

X.C. is supported by NSF grants IIS-1838154, CCF-2106429, and CCF-2107187. A.D. is supported by NSF grants CCF-1910534 and CCF0-2045128. S.N. is supported by NSF grants CCF-2106429, CCF-2211238, CCF-1763970, and CCF-2107187. R.A.S. is supported by NSF grants CCF-2106429 and CCF-2211238. E.W. is supported by NSF grant CCF-2337993.

This work was partially completed while a subset of the authors were visiting the Simons Institute for the Theory of Computing.

References

[B+97] Keith Ball et al. An elementary introduction to modern convex geometry. *Flavors of geometry*, 31(1–58):26, 1997. 1, 7

[Bal93] K. Ball. The Reverse Isoperimetric Problem for Gaussian Measure. *Discrete and Computational Geometry*, 10:411–420, 1993. 1, 8

[Ban10] Nikhil Bansal. Constructive algorithms for discrepancy minimization. In *IEEE 51st Annual Symposium on Foundations of Computer Science (FOCS)*, pages 3–10, 2010. 2

[BB16] A. Belovs and E. Blais. A polynomial lower bound for testing monotonicity. In *Proceedings of the 48th ACM Symposium on Theory of Computing (STOC)*, pages 1021–1032, 2016. 3, 4

[BB20] Eric Blais and Abhinav Bommireddi. On testing and robust characterizations of convexity. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM*, pages 18:1–18:15, 2020. 6

[BBB20] Aleksandrs Belovs, Eric Blais, and Abhinav Bommireddi. Testing convexity of functions over finite domains. In *Proceedings of the 2020 ACM-SIAM Symposium on Discrete Algorithms, SODA*, pages 2030–2045, 2020. 6

[BBH24] Hadley Black, Eric Blais, and Nathaniel Harms. Testing and learning convex sets in the ternary hypercube. In *15th Innovations in Theoretical Computer Science Conference, ITCS*, pages 15:1–15:21, 2024. 6

[BF18] Omri Ben-Eliezer and Eldar Fischer. Earthmover resilience and testing in ordered structures. In *33rd Computational Complexity Conference, CCC*, pages 18:1–18:35, 2018. 6

[BLR93] M. Blum, M. Luby, and R. Rubinfeld. Self-testing/correcting with applications to numerical problems. *Journal of Computer and System Sciences*, 47:549–595, 1993. 1, 2

[BMR16] Piotr Berman, Meiram Murzabulatov, and Sofya Raskhodnikova. The Power and Limitations of Uniform Samples in Testing Properties of Figures. In *36th IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS)*, pages 45:1–45:14, 2016. 6

[BMR19] Piotr Berman, Meiram Murzabulatov, and Sofya Raskhodnikova. Testing convexity of figures under the uniform distribution. *Random Struct. Algorithms*, 54(3):413–443, 2019. 6

[BMR22] Piotr Berman, Meiram Murzabulatov, and Sofya Raskhodnikova. Tolerant testers of image properties. *ACM Trans. Algorithms*, 18(4):37:1–37:39, 2022. 6

[Bor75] C. Borell. The Brunn-Minkowski inequality in Gauss space. *Invent. Math.*, 30:207–216, 1975. 1

[Bor03] Christer Borell. The Ehrhard inequality. *Comptes Rendus. Mathématique*, 337(10):663–666, 2003. 1

[Bor08] C. Borell. Inequalities of the Brunn–Minkowski type for Gaussian measures. *Probability Theory and Related Fields*, 140:195–205, 2008. 1

[BRY14a] Piotr Berman, Sofya Raskhodnikova, and Grigory Yaroslavtsev. L_p-testing. In *Symposium on Theory of Computing, STOC 2014*, pages 164–173, 2014. 6

[BRY14b] Eric Blais, Sofya Raskhodnikova, and Grigory Yaroslavtsev. Lower bounds for testing properties of functions over hypergrid domains. In *IEEE 29th Conference on Computational Complexity, CCC 2014*, pages 309–320, 2014. 6

[CDL+24] Xi Chen, Anindya De, Yuhao Li, Shivam Nadimpalli, and Rocco A Servedio. Mildly exponential lower bounds on tolerant testers for monotonicity, unateness, and juntas. In *Proceedings of the 2024 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 4321–4337. SIAM, 2024. 3, 4, 5, 28, 30

[CDS19] Eshan Chattopadhyay, Anindya De, and Rocco A. Servedio. Simple and efficient pseudorandom generators from gaussian processes. In Amir Shpilka, editor, *34th Computational Complexity Conference (CCC)*, volume 137 of *LIPIcs*, pages 4:1–4:33. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019. 1

[CDST15] X. Chen, A. De, R. Servedio, and L.-Y. Tan. Boolean Function Monotonicity Testing Requires (Almost) $n^{1/2}$ Non-adaptive Queries. In *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing, STOC 2015*, pages 519–528, 2015. 5, 6, 33, 34, 35, 38

[CEFM04] Dario Cordero-Erausquin, Matthieu Fradelizi, and Bernard Maurey. The (b) conjecture for the gaussian measure of dilates of symmetric convex sets and related problems. *Journal of Functional Analysis*, 214(2):410–427, 2004. 1

[CFSS17] X. Chen, A. Freilich, R. Servedio, and T. Sun. Sample-based high-dimensional convexity testing. In *Proceedings of the 17th Int. Workshop on Randomization and Computation (RANDOM)*, pages 37:1–37:20, 2017. 1, 2, 3, 6

[CS13] Deeparnab Chakrabarty and C. Seshadhri. A $o(n)$ monotonicity tester for boolean functions over the hypercube. In *Proceedings of the 45th ACM Symposium on Theory of Computing*, pages 411–418, 2013. 2

[CST14] Xi Chen, Rocco A. Servedio, and Li-Yang Tan. New algorithms and lower bounds for testing monotonicity. In *Proceedings of the 55th IEEE Symposium on Foundations of Computer Science*, pages 286–295, 2014. 2, 3

[CWX17] Xi Chen, Erik Waingarten, and Jinyu Xie. Beyond Talagrand functions: new lower bounds for testing monotonicity and unateness. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing (STOC)*, pages 523–536, 2017. 3, 4

[DMN19] Anindya De, Elchanan Mossel, and Joe Neeman. Is your function low dimensional? In Alina Beygelzimer and Daniel Hsu, editors, *Conference on Learning Theory, COLT 2019, 25-28 June 2019, Phoenix, AZ, USA*, volume 99 of *Proceedings of Machine Learning Research*, pages 979–993. PMLR, 2019. 1

[DMN21] Anindya De, Elchanan Mossel, and Joe Neeman. Robust testing of low dimensional functions. In Samir Khuller and Virginia Vassilevska Williams, editors, *STOC '21: 53rd Annual ACM SIGACT Symposium on Theory of Computing, Virtual Event, Italy, June 21-25, 2021*, pages 584–597. ACM, 2021. 1

[DNS21] Anindya De, Shivam Nadimpalli, and Rocco A. Servedio. Quantitative correlation inequalities via semigroup interpolation. In *12th Innovations in Theoretical Computer Science Conference, ITCS 2021*, volume 185 of *LIPIcs*, pages 69:1–69:20, 2021. 3

[DNS22] Anindya De, Shivam Nadimpalli, and Rocco A. Servedio. Convex influences. In Mark Braverman, editor, *13th Innovations in Theoretical Computer Science Conference, ITCS*, volume 215 of *LIPIcs*, pages 53:1–53:21, 2022. 3

[DNS23] Anindya De, Shivam Nadimpalli, and Rocco A. Servedio. Testing Convex Truncation. In *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 4050–4082. 2023. 1

[DNS24] Anindya De, Shivam Nadimpalli, and Rocco A. Servedio. Gaussian Approximation of Convex Sets by Intersections of Halfspaces. In *Proceedings of the 65th IEEE Symposium on Foundations of Computer Science (FOCS)*, 2024. To appear. 3, 4, 8

[Dur19] Rick Durrett. *Probability: Theory and Examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 5 edition, 2019. 7

[Eld22] Ronen Eldan. Second-order bounds on correlations between increasing families. *Combinatorica*, 42:1099–1118, 2022. 2

[Fel68] William Feller. *An introduction to probability theory and its applications*, volume 1. Wiley, 3rd edition, 1968. 34

[GGL+00] O. Goldreich, S. Goldwasser, E. Lehman, D. Ron, and A. Samordinsky. Testing monotonicity. *Combinatorica*, 20(3):301–337, 2000. 2

[Glu89] Efim Davydovich Gluskin. Extremal properties of orthogonal parallelepipeds and their applications to the geometry of banach spaces. *Mathematics of the USSR-Sbornik*, 64(1):85, 1989. 2

[GOWZ10] P. Gopalan, R. O'Donnell, Y. Wu, and D. Zuckerman. Fooling functions of halfspaces under product distributions. In *IEEE Conf. on Computational Complexity (CCC)*, pages 223–234, 2010. 38

[GR11] Oded Goldreich and Dana Ron. On proximity-oblivious testing. *SIAM Journal on Computing*, 40(2):534–566, 2011. 2

[GR16] Oded Goldreich and Dana Ron. On sample-based testers. *ACM Trans. Comput. Theory*, 8(2):7:1–7:54, 2016. 2

[GW93] P. M. Gruber and J. M. Wills, editors. *Handbook of Convex Geometry*. Elsevier, 1993. 1

[HSSV22] Daniel J. Hsu, Clayton Hendrick Sanford, Rocco A. Servedio, and Emmanouil-Vasileios Vlatakis-Gkaragkounis. Near-optimal statistical query lower bounds for agnostically learning intersections of halfspaces with gaussian marginals. In Po-Ling Loh and Maxim Raginsky, editors, *Conference on Learning Theory, 2-5 July 2022, London, UK*, volume

178 of *Proceedings of Machine Learning Research*, pages 283–312. PMLR, 2022. 1

[HW20] Daniel Hug and Wolfgang Weil. *Lectures on Convex Geometry*. Springer Graduate Texts in Mathematics, 2020. 1

[HY22] Nathaniel Harms and Yuichi Yoshida. Downsampling for testing and learning in product distributions. In *49th International Colloquium on Automata, Languages, and Programming, (ICALP)*, pages 71:1–71:19, 2022. 6

[Joh01] Iain M. Johnstone. Chi-square oracle inequalities. In *State of the art in probability and statistics*, pages 399–418. Institute of Mathematical Statistics, 2001. 34

[Kan11] D. M. Kane. The Gaussian Surface Area and Noise Sensitivity of Degree-d Polynomial Threshold Functions. *Computational Complexity*, 20(2):389–412, 2011. 1

[Kan12] D. Kane. A Structure Theorem for Poorly Anticoncentrated Gaussian Chaoses and Applications to the Study of Polynomial Threshold Functions. In *53rd Annual IEEE Symposium on Foundations of Computer Science, FOCS 2012, New Brunswick, NJ, USA, October 20-23, 2012*, pages 91–100, 2012. 1

[Kan14] D. Kane. A Pseudorandom Generator for Polynomial Threshold Functions of Gaussian with Subpolynomial Seed Length. In *IEEE 29th Conference on Computational Complexity, CCC 2014, Vancouver, BC, Canada, June 11-13, 2014*, pages 217–228, 2014. 1

[Kan15] D. M. Kane. A Polylogarithmic PRG for Degree 2 Threshold Functions in the Gaussian Setting. In *30th Conference on Computational Complexity, CCC 2015, June 17-19, 2015, Portland, Oregon, USA*, pages 567–581, 2015. 1

[KK14] A. R. Klivans and P. Kothari. Embedding Hard Learning Problems Into Gaussian Space. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2014*, pages 793–809, 2014. 1

[KMS18] Subhash Khot, Dor Minzer, and Muli Safra. On monotonicity testing and boolean isoperimetric-type theorems. *SIAM J. Comput.*, 47(6):2238–2276, 2018. 2

[KNOW14] Pravesh Kothari, Amir Nayyeri, Ryan O'Donnell, and Chenggang Wu. Testing surface area. In Chandra Chekuri, editor, *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2014*. SIAM, 2014. 1

[KOS08] A. Klivans, R. O'Donnell, and R. Servedio. Learning geometric concepts via Gaussian surface area. In *Proc. 49th IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 541–550, 2008. 1, 2, 3, 8

[Lat02] Rafał Latała. On some inequalities for gaussian measures. In *Proceedings of the ICM*, volume 2, pages 813–822, 2002. 1

[LL15] I. E. Leonard and J. E. Lewis. *Geometry of Convex Sets*. Wiley, 2015. 1

[LM00] B. Laurent and P. Massart. Adaptive estimation of a quadratic functional by model selection. *Annals of Statistics*, 28(5):1302–1338, 2000. 7

[LM15] Shachar Lovett and Raghu Meka. Constructive discrepancy minimization by walking on the edges. *SIAM Journal on Computing*, 44(5):1573–1582, 2015. 2

[LO99] Rafał Latała and Krzysztof Oleszkiewicz. Gaussian measures of dilatations of convex symmetric sets. *The Annals of Probability*, 27(4):1922–1938, 10 1999. 1

[LO05] Rafał Latała and Krzysztof Oleszkiewicz. Small ball probability estimates in terms of width. *Studia Mathematica*, 169(3):305–314, 2005. 1

[LRR17] Avi Levy, Harishchandra Ramadas, and Thomas Rothvoss. Deterministic discrepancy minimization via the multiplicative weight update method. In *International Conference on Integer Programming and Combinatorial Optimization (IPCO)*, pages 380–391, 2017. 2

[McD89] C. McDiarmid. On the method of bounded differences. In *Surveys in Combinatorics 1989*, pages 148–188. London Mathematical Society Lecture Notes, 1989. 13

[MO03] Elchanan Mossel and Ryan O'Donnell. On the noise sensitivity of monotone functions. *Random Structures & Algorithms*, 23(3):333–350, 2003. 5

[MORS10] K. Matulef, R. O'Donnell, R. Rubinfeld, and R. Servedio. Testing halfspaces. *SIAM J. on Comput.*, 39(5):2004–2047, 2010. 1

[Mos08] Elchanan Mossel. Gaussian bounds for noise correlation of functions and tight analysis of long codes. *FOCS*, pages 156–165, 2008. 38

[Naz03] F. 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 (2001-2002)*, pages 169–187. Lecture Notes in Math., Vol. 1807, Springer, 2003. 1, 3, 8

[OSTK21] Ryan O'Donnell, Rocco A. Servedio, Li-Yang Tan, and Daniel Kane. Fooling Gaussian PTFs via local hyperconcentration. Preliminary version in STOC 2020. Revised version includes an appendix by Daniel Kane, 2021. 1

[OW07] Ryan O'Donnell and Karl Wimmer. Approximation by DNF: examples and counterexamples. In *International Colloquium on Automata, Languages, and Programming*, pages 195–206. Springer, 2007. 5

[PRR03] Michal Parnas, Dana Ron, and Ronitt Rubinfeld. On testing convexity and submodularity. *SIAM J. Comput.*, 32(5):1158–1184, 2003. 6

[PRR06] Michal Parnas, Dana Ron, and Ronitt Rubinfeld. Tolerant property testing and distance approximation. *J. Comput. Syst. Sci.*, 72(6):1012–1042, 2006. 3

[PRV18] Ramesh Krishnan S. Pallavoor, Sofya Raskhodnikova, and Nithin Varma. Parameterized property testing of functions. *ACM Trans. Comput. Theory*, 9(4):17:1–17:19, 2018. 6

[PRW22] R. Pallavoor, S. Raskhodnikova, and E. Waingarten. Approximating the distance to monotonicity of Boolean functions. *Random Struct. Algorithms*, 60(2):233–260, 2022. 3, 4, 5

[Ras03] Sofya Raskhodnikova. Approximate testing of visual properties. In *6th International Workshop on Approximation Algorithms for Combinatorial Optimization Problems, APPROX 2003 and 7th International Workshop on Randomization and Approximation Techniques in Computer Science, RANDOM 2003*, pages 370–381, 2003. 6

[Rot17] Thomas Rothvoss. Constructive discrepancy minimization for convex sets. *SIAM Journal on Computing*, 46(1):224–234, 2017. 2

[Rot23] Thomas Rothvoss. Lattices: CSE 599S—Winter 2023 Lecture Notes, 2023. URL: https://sites.math.washington.edu/~rothvoss/599-winter-2023/lattices.pdf. 2

[Roy14] Thomas Royen. A simple proof of the Gaussian correlation conjecture extended to multivariate gamma distributions. 2014. arXiv:1408.1028. 1

[RR23a] Victor Reis and Thomas Rothvoss. The subspace flatness conjecture and faster integer programming. In *IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 974–988, 2023. 2

[RR23b] Victor Reis and Thomas Rothvoss. Vector balancing in lebesgue spaces. *Random Structures & Algorithms*, 62(3):667–688, 2023. 2

[RS96] R. Rubinfeld and M. Sudan. Robust characterizations of polynomials with applications to program testing. *SIAM Journal on Computing*, 25:252–271, 1996. 1

[RSD24] Oded Regev and Noah Stephens-Davidowitz. A reverse Minkowski theorem. *Annals of Mathematics*, 199(1):1–49, 2024. 2

[RV05] Luis Rademacher and Santosh Vempala. Testing geometric convexity. In *FSTTCS 2004: Foundations of Software Technology and Theoretical Computer Science: 24th International Conference, 2004*, pages 469–480, 2005. 6

[Tal96] M. Talagrand. How much are increasing sets positively correlated? *Combinatorica*, 16(2):243–258, 1996. 4, 5

[Tko18] Tomasz Tkocz. Asymptotic Convex Geometry Lecture Notes, 2018. URL: https://www.math.cmu.edu/~ttkocz/teaching/1819/asympt-conv-geom-notes.pdf. 1

[Tro18] Joel A. Tropp. Lectures on convex geometry. 2018. URL: https://tropp.caltech.edu/notes/Tro18-Lectures-Convex-LN.pdf. 1

[Vem10] Santosh S. Vempala. Learning convex concepts from gaussian distributions with PCA. In *51th Annual IEEE Symposium on Foundations of Computer Science, FOCS*, pages 124–130. IEEE Computer Society, 2010. 1

[Ver18] Roman Vershynin. *High-Dimensional Probability: An Introduction with Applications in Data Science*, volume 47. Cambridge University Press, 2018. 16, 22

[Wai15] M. Wainwright. Basic tail and concentration bounds, 2015. URL: www.stat.berkeley.edu/~mjwain/stat210b/Chap2_TailBounds_Jan22_2015.pdf. 7

[Wik23] Wikipedia contributors. Chi-squared distribution. Wikipedia, The Free Encyclopedia, Accessed on September 27, 2023. URL: https://en.wikipedia.org/wiki/Chi-squared_distribution. 7

[Wil05] R Willink. Bounds on the bivariate normal distribution function. *Communications in Statistics-Theory and Methods*, 33(10):2281–2297, 2005. 31

[Yao77] A. Yao. Probabilistic computations: Towards a unified measure of complexity. In *Proc. Seventeenth Annual Symposium on Foundations of Computer Science (STOC)*, pages 222–227, 1977. 8