Topics in the Geometry of the High-Dimensional Sphere, Discrete Cube, and Convex Sets

Thesis for the degree: Doctor of Philosophy

Submitted to the Scientific Council of the Weizmann Institute of Science, Rehovot, Israel

Uri Grupel

supervised by

Prof. Bo'az Klartag

August 2018

נושאים בגיאומטריה במימדים גבוהים של הספירה, הקוביה הדיסקרטית וגופים קמורים

עבודת גמר (תזה) לתואר: דוקטור לפילוסופיה

מוגשת למועצה המדעית של מכון ויצמן למדע, רחובות, ישראל.

אורי ברוסל

מנחה: בועז קלרטג

אלול תשע"ח

Acknowledgment

This thesis was written under the supervision of Bo'az Klartag whose support, patience, suggestions and knowledge were invaluable. His guidance helped me immensely through all stages of the research that is presented here. I am very grateful to him, and could not have asked for a better advisor.

I also thank Itai Benjamini, for his support, and many enlightening discussions.

My work was supported by the European Research Council (ERC).

In addition, I would like to thank people who helped and contributed to different chapters of this work.

Chapter 2

I Thank for Sasha Sodin for useful discussions and suggestions, and Oded Regev for his remarks on an early draft of this work.

Chapter 3

I thank Alon Nishry, Liran Rotem and Ben Cousins for many useful discussions.

Chapter 4

I Thank Renan Gross, who co-wrote this work with me. Additional thanks for Itai Benjamini for proposing the question of indistinguishability and for his advice, Ronen Eldan for his suggestions on locally stable functions, and David Ellis for his comments on perfect codes. In addition, I thank Noga Alon and Peleg Michaeli for some useful discussions.

Declaration

This thesis summarize my independent research. Chapter 4 is the result of joint work with Renan Gross.

Contents

Abstract

In this work we examine several high dimensional problems, and see whether our low dimensional intuition is affected by high dimensions phenomena. The problems considered here include random intersections of the sphere by high and low dimensional linear subspaces, intersections of convex bodies by random geodesics, and reconstruction of boolean functions on the hypercube (or the inability to do so) by the distribution of the scenery viewed by a random walker.

תקציר

בעבודה זו אנו בוחנים מספר בעיות במימדים גבוהים, ובודקים האם האינטואיציה שלנו שמקורה במימדים נמוכים מתאימה לתופעות הנצפות במימד גבוה. הבעיות הנבחנות כאן כוללות חיתוכים אקראים של הספירה על ידי תתי מרחבים לינארים ממימד גבוה או נמוך, חיתוכים של גופים קמורים על ידי גיאודזים אקראיים ושיחזור נוף של פונקציות בוליאניות בקוביה הדיסקרטית (או חוסר היכולת לעשות זאת) על ידי התפלגות הנוף הנצפה על ידי מהלך אקראי.

Chapter 1

Introduction

A natural question in analysis and geometry is what happens when the number of parameters grows. On the one hand, our intuition is that there will be a positive correlation between the number of parameters and the complexity of the problem. In that case, we would expect, accurate predictions to be impossible. On the other hand, concentration of measure phenomena show us, that under certain assumptions, as the number of parameters grows, the probability of diverging from the average (or expected) result decreases. In this report we discuss examples for different phenomena on this spectrum.

An early example of measure concentration, due to Lévy, is the isoperimetric inequality on the sphere (see [3]). We denote the unit sphere by $S^{n-1} = \{x \in \mathbb{R};\ |x| = 1\}$. Let $A \subseteq S^{n-1}$, we denote by A_ε the ε extension of the set A, that is

$$A_\varepsilon = \left\{x \in S^{n-1};\ d(x, A) \leq \varepsilon\right\}.$$

We denote by σ_{n-1} the normalized surface area measure on the sphere. Let $H \subseteq S^{n-1}$ be a spherical cap of the same measure as the set A. Then

$$\sigma_{n-1}(A_\varepsilon) \geq \sigma_{n-1}(H_\varepsilon).$$

In particular, taking $\sigma_{n-1}(A) \geq 1/2$ we have

$$\sigma_{n-1}(A_\varepsilon) \geq 1 - Ce^{-\varepsilon^2 n/2}.$$

An immediate consequence of the spherical isoperimetric inequality is the concentration of Lipschitz functions on the sphere. Let $f : S^{n-1} \to \mathbb{R}$ be an L-Lipschitz

function. Then,

$$\mathbb{P}\left(|f(X) - \mathbb{E}f| \geq \varepsilon\right) \leq Ce^{-n\varepsilon^2/2L^2}.$$

Hence, Lipschitz functions in high dimensional are close to radial functions. Lévy's theorem is one of the fundamental tools for many high dimensional results such as the Johnson-Lindenstrauss lemma [18], Milman's proof of Dvoretzky's theorem [3] and Klartag's proof of the central limit theorem for convex sets [19].

The Dvoretzky-Milman theorem states that for any centrally symmetric convex body $K \subseteq \mathbb{R}^n$, and any $k \leq c\varepsilon^2 \log n$, we have

$$\mathbb{P}\left((1-\varepsilon)\mathcal{E} \subseteq K \cap F \subseteq (1+\varepsilon)\mathcal{E}\right) \geq 1 - e^{-c\varepsilon^2 k},$$

where $\mathcal{E} \subseteq \mathbb{R}^k$ is an ellipsoid and $F \subseteq \mathbb{R}^n$ is a random subspace of dimension k. This shows that for **any** normed space of dimension n, a random k-dimensional subspace will typically be close to Euclidean. Depending on the geometry of the body, it might be possible to improve the bound on k (For example, if $K = \{x \in \mathbb{R}^n;\ \sum |x_i|^p\}$ is the unit ball of ℓ_p^n and $1 \leq p \leq 2$ then we can take $k \leq cn$). By considering the cube $[-1,1]^n$, we see that in general $\log n$ cannot be improved.

The Dvoretzky-Milman theorem shows us that random intersections of convex bodies is an example to how behavior can become more regular as the dimension grows. In this report we examine what happens when we take random intersections of subsets of the sphere.

In order to prove the central limit theorem for convex sets, Klartag proved a thin shell property for convex sets. There had been some improvements by Klartag [20], Fleury [11] and Guédon and Milman [16]. They proved that for any isotropic convex body K, and a random vector X distributed uniformly inside K, we have

$$\mathbb{P}\left(\left|\frac{|X|}{\sqrt{n}} - 1\right| \geq t\right) \leq Ce^{-c\sqrt{n}\min\{t^3,t\}}, \quad \forall t \geq 0.$$

This statement shows that most of the mass of an isotropic convex body is concentrated around a thin spherical shell of radius $\sqrt{n}$ and width $n^{1/3}$. A famous conjecture by Anttila, Ball and Perissinaki [2] states that the width of the shell is constant, and does not depend on the dimension . This conjuncture was proved by Klartag [21] for unconditional convex bodies. The thin shell property can be seen as the reason behind Theorem 3.2, though we use simpler tools in the proof.

Given a measurable subset $A \subseteq S^{n-1}$ and a random subspace $H \subseteq \mathbb{R}^n$ of dimension k, the expected normalized surface area of the intersection $A \cap H$ is the normalized surface area of the set A. Writing σ_H for the normalized surface area measures on $S^{n-1} \cap H$ we have,

$$\mathbb{E}\sigma_H(A \cap H) = \sigma_{n-1}(A),$$

where H is distributed according to the $SO(n)$ invariant measure of the Grassmanian manifold $G_{n,k}$ of k-dimensional subspaces of $\mathbb{R}^n$.

In [22], Klartag and Regev showed that for a high dimensional random subspace, the random variable $\sigma_H(A \cap H)/\mathbb{E}\sigma_{n-1}(A)$ is highly concentrated around its mean. They started with the case $\dim(H) = n-1$ and by repeated applications, extended the result (with a weaker concentration probability) to $\dim(H) = n/2$.

Theorem 1.1 (Klartag and Regev)**.** *Let $A \subseteq S^{n-1}$ such that $\sigma_{n-1}(A) \geq e^{-cn^{1/3}}$, then*

$$\mathbb{P}_H\left(\left|\frac{\sigma_H(A \cap H)}{\sigma_{n-1}(A)} - 1\right| \geq \frac{1}{10}\right) \leq Ce^{-c'n^{1/3}},$$

Where $H \subseteq \mathbb{R}^n$ is a uniform subspace of dimension $n/2$.

The motivation to the result of Klartag and Regev came from a communication complexity problem called *The Vector in Subspace Problem*. Their result, shows that any protocol that solves this communication problem has to use at least $Cn^{1/3}$ bits. Raz demonstrated in [33] a protocol that solves this problem with the exchange of $C\sqrt{n}$ bits. Hence, there is a gap between the lower and the upper bound on the communication complexity of this problem.

The connection between measure concentration and communication complexity, that was used by Klartag and Regev, is through the rectangle method [34]. Any protocol for the vector in subspace problem, divides the product space $S^{n-1} \times G_{n,n/2}$ into rectangles (product sets), according to the bits that were exchanged. Each rectangle is marked as either "In" or "Out". After exchanging N bits, we have at most 2^N rectangles. Hence, if a small number of bits were exchanged, at least some of the rectangles would be "big" (of large measure). The measure concentration result shows that in "big" rectangles we cannot determine a correct result with high enough probability.

In [15] we suggested a modification to Theorem 1.1 in order to improve the lower bound of the communication complexity. The main idea is to examine both the intersection with the random subspace H and its orthogonal complement $H^{\perp}$ simultaneously. We prove that for a class of measurable sets that are defined by at most $c\sqrt{n}$ directions, we have a stronger concentration probability for the geometric average of the surface areas to diverge.

Theorem 1.2. *Let $A \subseteq S^{n-1}$ such that $\sigma_{n-1}(A) \geq e^{-c\sqrt{n}}$. Assume that $1_A(x)$ can be written as $F(\langle x, \xi_1 \rangle, \ldots, \langle x, \xi_k \rangle)$ where $k \leq C\sqrt{n}$. Then*

$$\mathbb{P}_H \left(\sqrt{\sigma_H(A \cap H)\sigma_{H^{\perp}}(A \cap H^{\perp})} \leq 0.9\sigma_{n-1}(A) \right) \leq C' e^{-c'\sqrt{n}},$$

Where $H \subseteq \mathbb{R}^n$ is a uniform subspace of dimension $n/2$.

In Chapter 2 we explain the the proof of Theorem 1.2. This gives a sharp lower bound of $C\sqrt{n}$, for a subclass of protocols to the vector in subspace problem.

Theorems 1.1 and 1.2 show us that intersection with a random high dimensional subspace enjoys strong concentration of measure. In the proof of theorem 1.1, we see that decreasing the dimension can cause a decrease in the concentration phenomenon. The next result deals with the case that the dimension of the random subspace is small. For example, if $\dim(H) = 2$ then $H \cap S^{n-1}$ is a random geodesic curve on S^{n-1}. In that case we get a dimension free result.

Theorem 1.3. *Let $A \subseteq S^{n-1}$ such that $\sigma_{n-1}(A) = 1/2$, then*

$$\mathbb{P}_L \left(\left| \frac{Length(A \cap L)}{Length(L)} - 1 \right| \geq \frac{1}{2^{1/3}} \right) \leq \frac{1}{2^{1/3}},$$

Where $L \subseteq \mathbb{R}^n$ is a uniform geodesic curve in S^{n-1}.

For a general set, the probability bound cannot be improved to a term that tends to 0 with the dimension. In Chapter 3 we discuss the proof of Theorem 1.3. This proof can be generalized to random subspaces of small dimension (does not depend on n) or to the discrete case, where we replace that sphere with the torus $(\mathbb{Z}/p\mathbb{Z})^n$, for a prime p.

The analysis of Theorem 1.3 is done by finding the singular values of the Radon transform, and use them to bound the variance of the random variable Length($A \cap$

$L)/\text{Length}(L)$. We have full understanding of the singular values of the Radon transform to any other dimension k as well. Hence we can bound the variance of the intersection with a random k-dimensional subspaces and get a bound on the probability to deviate from the mean.

Theorem 1.4. *Let $A \subseteq S^{n-1}$ be a measurable set. Let $2 \leq k \leq n-1$, and let $H \subseteq \mathbb{R}^n$ be a random subspace of dimension k. Then,*

$$Var\left(\frac{\sigma_H(A \cap H)}{\sigma_{n-1}(A)}\right) \leq \frac{n-k}{k(n-1)}\left(\frac{1-\sigma_{n-1}(A)}{\sigma_{n-1}(A)}\right).$$

There are additional ingredients needed in order to get a sharp results such as Theorem 1.1 using this method.

The question of random intersections by geodesics can be examined on other settings as well. We prove that doing a similar procedure inside a convex body, we have a very different result than that of the sphere.

Theorem 1.5. *Let $K \subseteq \mathbb{R}^n$ be a convex body. There exists a set $A \subseteq K$ such that $Vol(A)\,/\,Vol(K) = 1/2$ and*

$$\mathbb{P}\left(\frac{length(L \cap A)}{length(L \cap K)} \in \{0,1\}\right) = 1 - O^*\left(\frac{1}{\sqrt{n}}\right),$$

where $L = X + \mathbb{R}\theta$, X and θ are independent and distributed uniformly in K and S^{n-1} respectively.

The proof of Theorem 3.2 (see Chapter 3 for more details), shows that in this case, the obstacle to concentration phenomenon of random geodesic intersections is the concentration of the volume distribution inside convex bodies.

Our last example is of *scenery reconstruction* on the n dimensional hypercube. Given a graph $G = (V, E)$ and a scenery $f : V \to \{-1, 1\}$, start a simple random walk on the graph $(X_k)_{k=0}^{\infty}$. For every step of the random walk mark the label of the vertex $f(X_k)$. The sequence $f(X_0), f(X_1), \ldots$ is the scenery observed by the random walker. Can you recover the scenery function f (up to graph isometries) by the distribution of the scenery viewed by the walker?

In [5] Benjamini and Kesten proved that when the graph G is a circle $\mathbb{Z}/n\mathbb{Z}$, reconstruction is possible in polynomial time. In [25] Lindenstrauss proved that scenery reconstruction on the line is impossible. In [14], together with Gross, we

proved that scenery reconstruction on the hypercube $\{-1,1\}^n$ is impossible when $n \geq 4$. We constructed a class of functions that share the same scenery distribution, and showed that the number of non isomorphic functions in that class grows rapidly with the dimension n. We discuss this in Chapter 4. The example of scenery reconstruction on the hypercube, is an example of how a complexity of a problem can grow as the dimension increases.

One class of functions we defined is the class of locally p-biased functions. A functions in this class has the following property: for every vertex of the graph, the portion of it neighbors labeled by 1 is exactly p. On the hypercube, we give a full characterization of which values of p are possible for a given dimension n. We also raise the question of existence of such functions in other graphs. We give constructions for trees and the lattice $\mathbb{Z}^n$. In a recent paper, Van Hintum [37] built upon it and gave a full characterization for the case $\mathbb{Z}^n$, and showed that these classes are uncountable when $n \geq 2$.

Chapter 2

Intersection with Random Mutually Orthogonal Subspaces

2.1 Introduction

The Vector in Subspace Problem (*VSP*) is a communication problem where one party (Alice) receives a unit vector $u \in S^{n-1}$, and a second party (Bob) receives a subspace $H \subseteq \mathbb{R}^n$ of dimension $\lfloor n/2 \rfloor$ such that either $u \in H$ or $u \in H^{\perp}$. The goal of Alice and Bob is to determine whether $u \in H$ or not.

VSP was introduced by Kremer [23] and has been studied under both classical and quantum communication models . In the classical communication model, Alice and Bob exchange bits between them in order to determine whether $u \in H$. In the quantum communication model, Alice and Bob exchange qubits.

In this paper, the terms protocol and complexity refer to distributional complexity (see [40] [4]). That is, a protocol outputs the correct answer with probability at least 2/3. The complexity is measured according to the number of bits or qubits that are exchanged in the worst case .

It is known that VSP can be solved in the quantum model with the exchange of $O(\log n)$ qubits. In [33] Raz presented a classical protocol that solves VSP with the exchange of $O(\sqrt{n})$ bits. In [22], Klartag and Regev proved that any classical protocol for VSP has a communication complexity of at least $\Omega(n^{1/3})$ bits. Thus, VSP shows that quantum communication can be exponentially stronger than clas-

sical communication. In this paper we discuss the gap between the lower and upper bound for the classical model.

We focus on the class of protocols of *bounded total rank*. In a deterministic protocol, each decision by Alice on the value of the next bit to be sent to Bob is based on two factors: her knowledge of the communication received so far, and perhaps an additional measurement of the vector u. We define the rank of the decision to be the number of linear functionals of the vector u that Alice has to compute in order to carry out the measurement. For example, deciding the value of the next bit by using the indicator function of the set $\{\langle x, v_1\rangle \geq 0, \sin(\langle x, v_2\rangle^2) \geq 1/2\}$, where $v_1, v_2 \in \mathbb{R}^n$ are determined by the communication received so far, is a decision of rank 2. In general, the rank of a decision is an integer between 0 and n. The total rank of a protocol is the sum of all ranks of decisions made by Alice in the worst case scenario. Note that we do not count decisions by Bob. The protocol Raz introduced can be slightly modified to be of total rank $O(\sqrt{n})$ (for more details see Appendix 2.A).

We prove that any protocol of VSP, for the classical model, of total rank at most $O(\sqrt{n})$ has communication complexity of at least $\Omega(\sqrt{n})$ bits. In light of the upper bound by Raz, this lower bound is sharp. We also introduce a novel mathematical conjecture about concentration of measure in the high dimensional sphere. This conjecture implies that any classical protocol for VSP has a communication complexity of at least $\Omega(\sqrt{n})$ bits.

The lower bound by Klartag and Regev is a result of a concentration theorem for sampling on the sphere by random subspaces. They proved that for any measurable subset $A \subseteq S^{n-1}$ with $\sigma_{n-1}(A) \geq Ce^{-cn^{1/3}}$, where σ_{n-1} is the uniform probability measure on the sphere S^{n-1}, it holds that

$$\mathbb{P}_H\left(\left|\frac{\sigma_H(A \cap H)}{\sigma_{n-1}(A)} - 1\right| \leq 0.1\right) \geq 1 - e^{-c'n^{1/3}},$$

where $C, c, c' > 0$ are universal constants. Here σ_H denotes the Haar probability measure on $S^{n-1} \cap H$ and $\mathbb{P}_H$ denotes the orthogonally invariant Haar probability measure over the Grassmanian manifold of subspaces $H \subseteq \mathbb{R}^n$ of dimension $\lfloor n/2 \rfloor$.

The concentration inequality by Klartag and Regev is sharp. Taking $A = \{x \in$

$S^{n-1};\ x_1 \geq T\}$, where $T \approx n^{-1/3}$ is chosen such that $\sigma_{n-1}(A) = e^{-n^{1/3}}$ gives

$$\mathbb{P}_H\left(\left|\frac{\sigma_H(A \cap H)}{\sigma_{n-1}(A)} - 1\right| \leq 0.1\right) = 1 - e^{-c_n n^{1/3}},$$

where c_n has a finite limit $c \in (0, \infty)$ as $n \to \infty$.

Our goal is to find a concentration inequality that applies to smaller sets, that is sets with measure of the order of magnitude of $e^{-\sqrt{n}}$. Our hope is that by considering both H and $H^\perp$ simultaneously, a stronger concentration result can be achieved.

Conjecture 2.1. *Let $A \subseteq S^{n-1}$ be a measurable subset with $\sigma_{n-1}(A) \geq e^{-c\sqrt{n}}$. Then*

$$\mathbb{P}_H\left(\sqrt{\sigma_H(A \cap H)\sigma_{H^\perp}(A \cap H^\perp)} \geq 0.9\sigma_{n-1}(A)\right)$$

is at least

$$1 - Ce^{-c'\sqrt{n}},$$

where $C, c, c' > 0$ are universal constants.

Conjecture 2.1 was essentially suggested by Klartag and Regev [22], albeit with a weaker arithmetic average in place of the geometric one.

In §2.3 we prove a special case of the conjecture where the set $A \subseteq S^{n-1}$ is of the form $\{x \in S^{n-1};\ (x_1, \ldots, x_k) \in I\}$ for some Borel set $I \subseteq B_k = \{x \in \mathbb{R}^k;\ |x| \leq 1\}$, and $k = O(\sqrt{n})$. By considering the case $k = 1$, this result shows that the conjecture holds for the extremal case of the theorem by Klartag and Regev. This extremal case also shows that if the conjecture is true it is tight.

This special case of the conjecture follows from the following result:

Theorem 2.2. *Let $k \leq \alpha_1\sqrt{n}$. Let $f : S^{n-1} \to [0, \infty)$ be a measurable function such that $\|f\|_\infty \leq e^{\alpha_2\sqrt{n}}$, $\|f\|_1 = 1$ and f depends only on $x_1, \ldots, x_k$. Then,*

$$\mathbb{P}_H\left(\sqrt{\int_{S_H} f(x)d\sigma_H(x)\int_{S_{H^\perp}} f(x)d\sigma_{H^\perp}(x)} \geq 0.9\right) \geq 1 - \alpha_3 e^{-\sqrt{n}},$$

where $\alpha_1, \alpha_2, \alpha_3 > 0$ are universal constants.

In the proof we use various tools from Geometric Functional Analysis. We begin by reformulating the problem in terms of random matrices instead of random subspaces. We show that the event in Theorem 2.2 strongly depends on the singular

values of a random projection. Using results from the theory of Wishart matrices we show that these singular values are concentrated around their expected values.

Next, we use the Cauchy-Schwarz inequality to define a smaller event than the one in the theorem. The use of the Cauchy-Schwarz inequality demonstrates how considering both H and $H^{\perp}$ simultaneously can enhance the concentration results and is fundamental to our approach.

Finally, we use the concentration results, and asymptotic tools such as the Laplace method in order to show that with high probability this smaller event holds true. In this last step we present bounds for the universal constants in Theorem 2.2 which are, in principle, explicit.

In §2.2 we employ the rectangle method and show how Theorem 2.2 implies a sharp lower bound for classical protocols of total rank at most $O(\sqrt{n})$.

Corollary 2.3. *Let $\mathcal{P}$ be a protocol for the Vector in Subspace Problem of total rank at most $\alpha_1\sqrt{n}$, with probability of error which is at most a constant smaller than $\frac{1}{2}$. Then the communication complexity of $\mathcal{P}$ is $\Omega(\sqrt{n})$ bits.*

Using the same methods, a positive resolution of Conjecture 2.1 would imply a sharp lower bound for a general classical protocol to VSP.

Theorem 2.4. *Let $\mathcal{P}$ be a general protocol for the Vector in Subspace Problem with probability of error which is at most a constant smaller than $\frac{1}{2}$. If Conjecture 2.1 is true then the communication complexity of $\mathcal{P}$ is $\Omega(\sqrt{n})$ bits.*

2.2 Applications to VSP

In this section we prove that Theorem 2.2 implies a lower bound of $O(\sqrt{n})$ for classical protocols of total rank at most $O(\sqrt{n})$. We also show that Conjecture 2.1 implies a lower bound of $O(\sqrt{n})$ for any classical protocol. In light of the result of Raz, if the conjecture is true then this bound is sharp.

Theorem 2.2 implies a special case of the conjecture for sets that depend only on $\alpha_1\sqrt{n}$ directions. For any such $A \subseteq S^{n-1}$ with $\sigma_{n-1}(A) \geq e^{-\alpha_2\sqrt{n}}$, define $f(x) = 1_A(x)/\sigma_{n-1}(A)$. The function f depends only on $\alpha_1\sqrt{n}$ directions, bounded by $1/\sigma_{n-1}(A) \leq e^{\alpha_2\sqrt{n}}$ and has $\|f\|_1 = 1$. Hence, we may apply Theorem 2.2. We have,

$$\alpha_3 e^{-\sqrt{n}} \geq \mathbb{P}_H \left(\sqrt{\int_{S_H} 1_A(x)/\sigma_{n-1}(A) d\sigma_H(x)} \sqrt{\int_{S_{H^\perp}} 1_A(x)/\sigma_{n-1}(A) d\sigma_{H^\perp}(x)} \leq 0.9 \right)$$
$$= \mathbb{P}_H \left(\sqrt{\sigma_H(A \cap H)\sigma_{H^\perp}(A \cap H^\perp)}/\sigma_{n-1}(A) \leq 0.9 \right).$$

In this section, we use this consequence of Theorem 2.2.

Our argument follows the rectangle method usually attributed to Babai, Frankl and Simon [4] and Razborov [34].

For simplicity, we assume that n is even. We denote by $G_{n/2}$ the Grassmanian manifold of all subspaces of $\mathbb{R}^n$ of dimension $n/2$, equipped with the O_n invariant measure σ_G. Let μ_0 be the uniform measure on $S^{n-1} \times G_{n/2}$. Denote

$$I_1 = \left\{ (u, H) \in S^{n-1} \times G_{n/2};\ u \in H \right\},$$

and

$$I_2 = \left\{ (u, H) \in S^{n-1} \times G_{n/2};\ u \in H^\perp \right\}.$$

Let μ_i be the Haar invariant probability measure on I_i for $i = 1, 2$, with respect to the obvious O_n action. Such measure exists due to the transitive property of such action. For a rectangular set $A_i \times B_i \subseteq I_i$ we have

$$\mu_1(A_1 \times B_1) = \int_{H \in B_1} \sigma_H(A_1 \cap H),$$

and

$$\mu_2(A_2 \times B_2) = \int_{H \in B_2} \sigma_{H^\perp}(A_2 \cap H^\perp).$$

By replacing α_2 in Theorem 2.2 with $\min\{\alpha_2, 1\}$ we may assume that it is at most 1.

Proposition 2.5. *Let $Q = A \times B \subseteq S^{n-1} \times G_{n/2}$ be such that $\mu_0(A \times B) \geq C e^{-\alpha_2 \sqrt{n}}$. Assume that the set A depends on $\alpha_1 \sqrt{n}$ directions. Then*

$$\sqrt{\mu_1(Q)\mu_2(Q)} \geq 0.8 \mu_0(Q).$$

The constants $\alpha_1, \alpha_2, C > 0$ are universal constants.

Proof. Define

$$E = \left\{ H \in B;\ \ \sqrt{\sigma_H(A \cap H)\sigma_{H^\perp}(A \cap H^\perp)} \leq 0.9\sigma(A) \right\}.$$

According to our assumption $\sigma_{n-1}(A) \geq \mu_0(A \times B) \geq Ce^{-\alpha_2\sqrt{n}}$. By the Theorem 2.2, $\mathbb{P}(E) \leq \alpha_3 e^{-\sqrt{n}}$. We choose $C \geq 1$ big enough, such that

$$0.9\sigma_G(B \setminus E) \geq 0.8\sigma_G(B). \tag{2.1}$$

By the Cauchy-Schwarz inequality

$$\begin{aligned}
\sqrt{\mu_1(Q)\mu_2(Q)} &= \sqrt{\left(\int_{H\in B} \sigma_{H\in B}(A \cap H)\right)\left(\int_{H\in B} \sigma_{H^\perp}(A \cap H^\perp)\right)} \\
&\geq \int_{H\in B} \sqrt{\sigma_H(A \cap H)\sigma_{H^\perp}(A \cap H^\perp)} \\
&\geq \int_{H\in B\setminus E} \sqrt{\sigma_H(A \cap H)\sigma_{H^\perp}(A \cap H^\perp)} \\
&\geq 0.9 \int_{H\in B\setminus E} \sigma(A) = 0.9\sigma(A)\sigma_G(B \setminus E).
\end{aligned}$$

Equation (2.1) gives us

$$\sqrt{\mu_1(Q)\mu_2(Q)} \geq 0.8\sigma_G(B)\sigma(A) = 0.8\mu_0(Q).$$

□

Corollary 2.6. *Let $Q = A \times B \subseteq S^{n-1} \times G_{n/2}$. Assume that the set A depends on $\alpha_1\sqrt{n}$ directions. Then*

$$\sqrt{\mu_1(Q)\mu_2(Q)} \geq 0.8\mu_0(Q) - Ce^{-\alpha_2\sqrt{n}}.$$

Using the above propositions, we are ready to prove Corollary 2.3.

Proof. By repeated application of the protocol we may assume that the probability of error is less then $\frac{1}{9}$. By Yao's principle [39], we may assume that our protocol is a randomly chosen deterministic protocol. Let D be the number of bits exchange in the protocol. We have a partition of $S^{n-1} \times G_{n/2}$ into 2^D rectangles of the form $Q = A \times B$, each labeled as "In H" or "In $H^\perp$". Since we assume the total rank is at most $\alpha_1\sqrt{n}$, for every $Q = A \times B$ in the partition, the set A is determined by at most $\alpha_1\sqrt{n}$ directions. Denote by $\mathcal{Q}_+$ all the rectangles labeled "In H", and by $\mathcal{Q}_-$ all the rectangles labeled "In $H^\perp$". According to our assumption $\sum_{Q\in\mathcal{Q}_+} \mu_2(Q) \leq \frac{1}{9}$ and $\sum_{Q\in\mathcal{Q}_-} \mu_1(Q) \leq \frac{1}{9}$. By Corollary 2.6 and the Cauchy-Schwarz inequality, we

have

$$\sum_{Q\in\mathcal{Q}_+}(0.8\mu_0(Q)-Ce^{-\alpha_2\sqrt{n}}) \le \sum_{Q\in\mathcal{Q}_+}\sqrt{\mu_1(Q)\mu_2(Q)}$$
$$\le \sqrt{\left(\sum_{Q\in\mathcal{Q}_+}\mu_1(Q)\right)\left(\sum_{Q\in\mathcal{Q}_+}\mu_2(Q)\right)}$$
$$\le \sqrt{1\cdot\frac{1}{9}}=\frac{1}{3}.$$

Similarly we have,

$$\sum_{Q\in\mathcal{Q}_-}(0.8\mu_0(Q)-Ce^{-\alpha_2\sqrt{n}}) \le \frac{1}{3}.$$

Summing the above inequalities, we obtain

$$0.8-2^DCe^{-\alpha_2\sqrt{n}} \le \frac{2}{3} \Rightarrow D \ge C'\sqrt{n}.$$

□

If Conjecture 2.1 is true, then Proposition 2.5 and Corollary 2.6 are true without the assumption that the set A depends on $\alpha_1\sqrt{n}$ directions. Hence, we may repeat the proof of Corollary 2.3 for a general classical protocol, and deduce Theorem 2.4.

2.3 Proofs

In this section we prove Theorem 2.2. This theorem is a special case of the conjecture for functions that depend only on $O(\sqrt{n})$ directions. We assume that n is even and greater than some universal constant. In this section, when we say *uniform distribution*, we refer to the Haar probability distribution.

Throughout this section we shall use the letters $c, \tilde{c}, C$ etc. to denote various universal constants, whose value may change from one line to the next. Additionally, $\alpha_1, \alpha_2, \alpha_3, \rho > 0$ are universal constants whose value would be determined only at the end of the section. Specifically, in this section we will assume a few upper bounds for α_1 in terms of explicit positive universal constants, an upper bound for α_2 in terms of α_1, and a lower bound for α_3 in terms of α_2.

Let $\psi : B_k \times S^{m-k-1} \to S^{m-1}$ be defined by $\psi(x,y) = (x, \sqrt{1-|x|^2}y)$. The map ψ enables us to separate the dependence on the first k coordinates. The following change of variables formula is standard:

Proposition 2.7. *For any integrable $f : S^{m-1} \to \mathbb{R}$ and any $1 \leq k \leq m-1$ there exists $C_{m,k} = (m-k)\,Vol(B_{m-k})/(m\,Vol(B_m))$ such that*

$$\int_{S^{m-1}} f(x)d\sigma_{m-1}(x) = C_{m,k}\int_{B_k}\left(1-|x|^2\right)^{(m-k-2)/2} \times\left(\int_{S^{m-k-1}} f\left(x,\sqrt{1-|x|^2}\theta\right)d\sigma_{m-k-1}(\theta)\right)dx.$$

Let $E = \mathrm{span}\{e_1,\ldots,e_k\}$. Let $H \subseteq \mathbb{R}^n$ be a random subspace of dimension $n/2$ distributed uniformly. Let $\lambda_1,\ldots,\lambda_k$ be the singular values of the projection map $P : H \to E$. The singular values $\lambda_1,\ldots,\lambda_k$ are the cosines of the principal angles between H and E. The next proposition along with Proposition 2.7, allows us to study the distribution of singular values of a random matrix instead of integration on a random subspace.

Proposition 2.8. *Let H be a random subspace of dimension $n/2$, let E and $\lambda_1,\ldots,\lambda_k$ be as before. Let $\Lambda = diag(\lambda_1,\ldots,\lambda_k)$. Let $U : E \to E$ be a random orthogonal map distributed uniformly, independent of H. Let $f : S^{n-1} \to \mathbb{R}$ be a measurable function such that f depends only on the first k coordinates. Then, the random variable*

$$\sqrt{\int_{S_H} f(x)d\sigma_H(x)\int_{S_{H^\perp}} f(x)d\sigma_{H^\perp}(x)}$$

is equal in distribution to

$$C_{n/2,k}\left(\psi(\Lambda)\psi(\sqrt{I-\Lambda^2})\right)^{1/2}, \tag{2.2}$$

where,

$$\psi(A) = \int_{B_k} f\left(UAU^Tx\right)\left(1-|x|^2\right)^{(n/2-k-2)/2}dx.$$

The constant $C_{n/2,k}$ is the same as in Proposition 2.7 with $m = n/2$.

Proof. In this proof, we construct an orthogonal map $V : \mathbb{R}^n \to \mathbb{R}^n$ that maps H and $H^\perp$ to canonical subspaces that depend only on the principle angles and a rotation $U : E \to E$. The map V is chosen such that f would be invariant under V. In order to construct V we use the projection maps to define appropriate orthonormal bases for H, $H^\perp$ and E.

By the Singular Value Decomposition (*SVD*) of the projection $P : H \to E$ there

exists an orthonormal basis $x_1, \ldots, x_k$ of E and an orthonormal basis $y_1, \ldots, y_{n/2}$ of H such that

$$P = \sum_{i=1}^{k} \lambda_i x_i \otimes y_i.$$

Since $Py_j = 0$ for all $j = k+1, \ldots, n/2$ we have $y_{k+1}, \ldots, y_{n/2} \in E^\perp$. Since $Py_i = \lambda_i x_i$ for $i = 1, \ldots, k$ there exists a unit vector $v_{n/2+i} \in E^\perp$ such that

$$y_i = \lambda_i x_i + \sqrt{1-\lambda_i^2} v_{n/2+i}, \quad \forall i = 1, \ldots, k.$$

Denote $v_j = y_j$ for $j = k+1, \ldots, n/2$. Let $i' = n/2 + i$ where $i \leq k$ and let $k+1 \leq j \leq n/2$. Since $v_{k+1}, \ldots, v_{n/2+k} \in E^\perp$, we have

$$0 = \langle y_i, y_j \rangle = \left\langle \lambda_i x_i + \sqrt{1-\lambda_i^2} v_{i'}, v_j \right\rangle = \sqrt{1-\lambda_i^2} \langle v_{i'}, v_j \rangle .$$

With probability 1 we have $0 < \lambda_i < 1$, hence with probability 1

$$\langle v_{i'}, v_j \rangle = 0.$$

By the same argument we obtain $\langle v_i, v_j \rangle = 0$ for all $i \neq j$. Let $P_\perp : H^\perp \to E$ be the orthogonal projection to E. The singular values of $P_\perp$ are exactly $\sqrt{1-\lambda_1^2}, \ldots, \sqrt{1-\lambda_k^2}$. Note that

$$\sqrt{1-\lambda_1^2} x_1 - \lambda_1 v_{n/2+1}, \ldots, \sqrt{1-\lambda_k^2} x_k - \lambda_k v_{n/2+k} \in H^\perp$$

are orthogonal to each other. Since the SVD is unique, up to trivial transformations, we have

$$P_\perp = \sum_{i=1}^{k} \sqrt{1-\lambda_i^2} x_i \otimes \left(\sqrt{1-\lambda_i^2} x_i - \lambda_i v_{n/2+i} \right).$$

Hence, there exist $v_{n/2+k+1}, \ldots, v_n \in E^\perp$ such that

$$\sqrt{1-\lambda_1^2} x_1 - \lambda_1 v_{n/2+1}, \ldots, \sqrt{1-\lambda_k^2} x_k - \lambda_k v_{n/2+k}, v_{n/2+k+1}, \ldots, v_n$$

is an orthonormal basis of $H^\perp$. By the same argument as before, with probability one, $v_{n/2+1}, \ldots, v_n$ are orthogonal to each other. Since $v_{n/2+k+1}, \ldots, v_n \in H^\perp$ and $v_{k+1}, \ldots, v_{n/2} \in H$ we find that $x_1, \ldots, x_k, v_{k+1}, \ldots, v_n$ is an orthonormal basis of $\mathbb{R}^n$. Let V be the orthogonal map defined by $Vx_i = x_i$ for $i = 1, \ldots, k$ and $Vv_j = e_j$

for $j = k+1, \dots, n$. Since V is the identity map on E we have $f(Vx) = f(x)$ for all $x \in S^{n-1}$. Hence,

$$\sqrt{\int_{S_H} f(x)d\sigma_H(x) \int_{S_{H^\perp}} f(x)d\sigma_{H^\perp}(x)} = \sqrt{\int_{S_H} f(Vx)d\sigma_H(x) \int_{S_{H^\perp}} f(Vx)d\sigma_{H^\perp}(x)}$$
$$= \sqrt{\int_{S_{VH}} f(x)d\sigma_{VH}(x) \int_{S_{VH^\perp}} f(x)d\sigma_{VH^\perp}(x)}$$

Let $B_{H,k}$ be the unit ball of

$$V\left(H \cap (E^\perp \cap H)^\perp\right) =$$
$$\text{span}\{\lambda_1 x_1 + \sqrt{1-\lambda_1^2}e_{n/2+1}, \dots, \lambda_k x_k + \sqrt{1-\lambda_k^2}e_{n/2+k}\}.$$

For any

$$x = \sum_{i=1}^k t_i \left(\lambda_i x_i + \sqrt{1-\lambda_i^2}e_{n/2+i}\right) \in B_{H,k},$$

where $\sum_{i=1}^k t_i^2 \le 1$, we have

$$f(x) = f\left(\sum_{i=1}^k t_i \left(\lambda_i x_i + \sqrt{1-\lambda_i^2}e_{n/2+i}\right)\right) = f\left(\sum_{i=1}^k t_i \lambda_i x_i\right).$$

Let $U : E \to E$ be the orthogonal map defined by $Ux_i = e_i$ for $i = 1, \dots, k$. We have,

$$\int_{B_{H,k}} f(x)dx = \int_{B_k} f(U\Lambda U^T x)dx,$$

where B_k is the unit ball of E. Since the distribution of H is invariant under the action of $O(k) \times O(n-k)$, the distribution of U is uniform over the orthogonal maps of E. Let $B_{H^\perp,k}$ be the unit ball of $\text{span}\{\sqrt{1-\lambda_1^2}x_1 - \lambda_1 e_{n/2+1}, \dots, \sqrt{1-\lambda_k^2}x_k - \lambda_k e_{n/2+k}\}$. Using the same map U, we have

$$\int_{B_{H^\perp,k}} f(x)dx = \int_{B_k} f(U\sqrt{I-\Lambda^2}U^T x)dx.$$

To finish the proof we use Proposition 2.7 on S_{VH} and $S_{VH^\perp}$. □

With probability 1, the matrices Λ and $\sqrt{I-\Lambda^2}$ are invertible. Hence, using the change of variables formula, (2.2) can be written as

$$\frac{C_{n/2,k}}{\sqrt{\prod_{j=1}^k \lambda_j \sqrt{1-\lambda_j^2}}} \sqrt{\int_{\mathbb{R}^k} \varphi_\Lambda(x)dx \int_{\mathbb{R}^k} \varphi_{\sqrt{I-\Lambda^2}}(x)dx},$$

where

$$\varphi_A(x) = f(x)\left(1 - \left|UA^{-1}U^T x\right|^2\right)_+^{(n/2-k-2)/2}.$$

By the Cauchy-Schwarz inequality this is at least

$$\frac{C_{n/2,k}}{\sqrt{\prod_{j=1}^k \lambda_j\sqrt{1-\lambda_j^2}}}\int_{\mathbb{R}^k}\sqrt{\varphi_\Lambda(x)\varphi_{\sqrt{I-\Lambda^2}}(x)}dx.$$

Hence, we need to estimate the coefficient and the integral on the function

$$f(x)\left(1 - \left|\Lambda^{-1}U^T x\right|^2\right)_+^{(n/2-k-2)/4}\left(1 - \left|\left(I-\Lambda^2\right)^{-1/2}U^T x\right|^2\right)_+^{(n/2-k-2)/4}. \quad (2.3)$$

The random variables $\lambda_1, \ldots, \lambda_k$ are the singular values of a block of size $n/2 \times k$ in a random orthogonal matrix. These singular values can be described using Wishart matrices [9]

Proposition 2.9. *Let N_1, N_2 be $(n/2) \times k$ independent random matrices with independent standard Gaussian entries. Let X be a random orthogonal matrix, chosen by the Haar uniform distribution. Let*

$$X = \begin{pmatrix} X_{1,1} & X_{1,2} \\ X_{2,1} & X_{2,2} \end{pmatrix}$$

Where $X_{1,1}$ is $(n/2) \times k$ block. Then, the singular values of $X_{1,1}$ have the same distribution as the square roots of the eigenvalues of $N_1^T N_1 (N_1^T N_1 + N_2^T N_2)^{-1}$.

Upper and lower bounds for the eigenvalues of the above matrix, can be achieved using a concentration result by Gordon for singular values of Gaussian matrices [38].

Lemma 2.10. *Let A be $(n/2) \times k$ random matrix with independent standard Gaussian entries. Assume that $k \leq n/2$. Let $s_1 \leq \cdots \leq s_k$ be the singular values of A, then with probability greater than $1 - 2e^{-t^2/2}$ we have*

$$\sqrt{n/2} - \sqrt{k} - t \leq s_1 \leq s_k \leq \sqrt{n/2} + \sqrt{k} + t.$$

Combining both results, we have

Proposition 2.11. *Let $k \leq \alpha_1\sqrt{n}$ and let $\lambda_1, \ldots, \lambda_k$ be as before. Then, with probability greater then $1 - 4e^{-\sqrt{n}}$, we have*

$$\left|\lambda_i - \frac{1}{\sqrt{2}}\right| \leq \frac{C(\sqrt{\alpha_1} + \sqrt{2})}{n^{1/4}},\ \forall i = 1, \ldots, k.$$

Proof. Let N_1, N_2 be as in Proposition 2.9. By Lemma 2.10 there exists $\mu_1, \dots, \mu_k$, $\sigma_1, \dots, \sigma_k$ and U, V orthogonal matrices such that

$$N_1^T N_1 = U\mathrm{diag}(\mu_1^2, \dots, \mu_k^2)U^T,\ N_2^T N_2 = V\mathrm{diag}(\sigma_1^2, \dots, \sigma_k^2)V^T,$$

and, there exists $C' > 0$ such that, with probability greater then $1 - 4e^{-\sqrt{n}}$,

$$\left|\mu_i^2 - \frac{n}{2}\right|, \left|\sigma_i^2 - \frac{n}{2}\right| \leq C'(\sqrt{\alpha_1} + \sqrt{2})n^{3/4}, \tag{2.4}$$

for all $i = 1, \dots, k$. Assume that event (2.4) holds true. Let E_1, E_2 be defined by $N_1^T N_1 = (n/2)I + E_1$ and $N_2^T N_2 = (n/2)I + E_2$. Then $\|E_i\|_{op} \leq C_i(\sqrt{\alpha_1} + \sqrt{2})n^{3/4}$ for $i = 1, 2$. We have,

$$N_1^T N_1(N_1^T N_1 + N_2^T N_2)^{-1} = \frac{1}{2}\left(I + \frac{2}{n}E_1\right)\left(I + \frac{1}{n}(E_1 + E_2)\right)^{-1}.$$

Let $T_1 = (E_1 + E_2)/n$ and $T_2 = E_1/n$, then $\|T_i\|_{op} \leq C_i'(\sqrt{\alpha_1} + \sqrt{2})n^{-1/4}$ for $i = 1, 2$. We have

$$N_1^T N_1(N_1^T N_1 + N_2^T N_2)^{-1} = \left(\frac{1}{2}I + T_2\right)\left(I - T_1 + \sum_{j=2}^{\infty}(-1)^j T_1^j\right).$$

Hence,

$$N_1^T N_1(N_1^T N_1 + N_2^T N_2)^{-1} = \frac{1}{2}I + T,$$

where $\|T\|_{op} \leq \tilde{C}(\sqrt{\alpha_1} + \sqrt{2})n^{-1/4}$. □

Corollary 2.12. *With probability greater then* $1 - 4e^{-\sqrt{n}}$ *we have*

$$\left|\left|U\Lambda^{-1}U^T x\right|^2 + \left|U(I - \Lambda^2)^{-1/2}U^T x\right|^2 - 4|x|^2\right| \leq C(\sqrt{\alpha_1} + \sqrt{2})^2|x|^2/\sqrt{n},$$

for any $x \in \mathbb{R}^k$.

Proof. Assume that

$$\Lambda = \frac{1}{\sqrt{2}}I + T,$$

where $\|T\|_{op} \leq C'(\sqrt{\alpha_1} + \sqrt{2})/n^{1/4}$. By the above proposition, this event has probability greater than $1 - 4e^{-\sqrt{n}}$. We have,

$$\begin{aligned}\Lambda^{-2} + (I - \Lambda^2)^{-1} &= \left(\frac{1}{2}I + \sqrt{2}T + T^2\right)^{-1} + \left(\frac{1}{2}I - \sqrt{2}T - T^2\right)^{-1} \\ &= 4(I - 4(\sqrt{2}T + T^2)^2)^{-1} = 4I + \tilde{T},\end{aligned}$$

where $\left\|\tilde{T}\right\|_{op} \leq \tilde{C}\left\|T^2\right\|_{op} \leq C(\sqrt{\alpha_1}+\sqrt{2})^2/\sqrt{n}$. Hence,

$$\begin{aligned}\left|U\Lambda^{-1}U^Tx\right|^2+\left|U(I-\Lambda^2)^{-1/2}U^Tx\right|^2 &= \left\langle U(\Lambda^{-2}+(I-\Lambda^2)^{-1})U^Tx,x\right\rangle \\ &= 4|x|^2+\left\langle U\tilde{T}U^Tx,x\right\rangle.\end{aligned}$$

Hence,

$$\begin{aligned}\left|\left|U\Lambda^{-1}U^Tx\right|^2+\left|U(I-\Lambda^2)^{-1/2}U^Tx\right|^2-4|x|^2\right| &\leq \left\|\tilde{T}\right\|_{op}|x|^2 \\ &\leq C(\sqrt{\alpha_1}+\sqrt{2})^2|x|^2/\sqrt{n}.\end{aligned}$$

□

The above proof demonstrates how considering both H and $H^\perp$ simultaneously can cancel the first order term in concentration inequalities. This cancellation leads to great improvement of the estimations, and it is one of the fundamental ideas of our approach.

The concentration of the principal angles, allows us to evaluate the coefficient $C_{n/2,k}\sqrt{\prod_{j=1}^k \frac{1}{\lambda_j\sqrt{1-\lambda_j^2}}}$ and the integral on (2.3).

Proposition 2.13. *Let $C_{n,k}$ and $C_{n/2,k}$ be the constants from Proposition 2.7 with $m=n,n/2$. For $k\leq\alpha_1\sqrt{n}$, we have*

$$2^{k/2}\frac{C_{n/2,k}}{C_{n,k}} \geq Ce^{-\alpha_1^2/4}.$$

Proof. By the definition of $C_{n,k}$ and $C_{n/2,k}$, we need to estimate

$$2^{k/2}\frac{\Gamma(n/4+1/2)\Gamma(n/2-k/2+1/2)}{\Gamma(n/4-k/2+1/2)\Gamma(n/2+1/2)}$$

Using Sterling's formula and the assumption on k, this is

$$\left(1+O\left(\frac{1}{\sqrt{n}}\right)\right)\exp\left(-\frac{k^2}{4n}+O\left(\frac{1}{\sqrt{n}}\right)\right).$$

□

Hence, by choosing α_1 small enough and using the concentration result for $\lambda_i\sqrt{1-\lambda_i^2}$ (as in Corollary 2.12) we have:

Corollary 2.14. *Let $k \leq \alpha_1\sqrt{n}$ and let $\lambda_1, \ldots, \lambda_k$ be as before. Then, with probability greater than $1 - 4e^{-\sqrt{n}}$, we have*

$$C_{n/2,k}\sqrt{\prod_{j=1}^{k}\frac{1}{\lambda_j\sqrt{1-\lambda_j^2}}} \geq 0.98C_{n,k}.$$

Proof. Assume that for all $1 \leq i \leq k$ we have $\lambda_i^2 = 1/2 + t_i$ where $|t_i| \leq C'(\sqrt{\alpha_1} + \sqrt{2})/n^{1/4}$. By Proposition 2.11, this event has probability greater than $1 - 4e^{-\sqrt{n}}$. We have,

$$\frac{1}{\lambda_i\sqrt{1-\lambda_i^2}} = \sqrt{\frac{1}{(1/2+t_i)(1/2-t_i)}} = \frac{2}{\sqrt{1-4t_i^2}}.$$

Hence,

$$\left|\frac{1}{\lambda_i\sqrt{1-\lambda_i^2}} - 2\right| \leq \frac{C(\sqrt{\alpha_1}+\sqrt{2})^2}{\sqrt{n}}.$$

We have

$$\sqrt{\prod_{j=1}^{k}\frac{1}{\lambda_j\sqrt{1-\lambda_j^2}}} \geq 2^{k/2}\exp\left(-\frac{C(\sqrt{\alpha_1}+\sqrt{2})^2 k}{4\sqrt{n}} + O\left(\frac{k(\sqrt{\alpha_1}+\sqrt{2})^2}{n}\right)\right).$$

We may assume that α_1 is small enough, such that both

$$\exp\left(-\frac{C(\sqrt{\alpha_1}+\sqrt{2})^2\alpha_1}{4} + O\left(\frac{\alpha_1(\sqrt{\alpha_1}+\sqrt{2})^2}{\sqrt{n}}\right)\right) \geq 0.99,$$

and (By Proposition 2.13), $0.99C_{n/2,k}2^{k/2} \geq 0.98C_{n,k}$. Hence,

$$C_{n/2,k}\sqrt{\prod_{j=1}^{k}\frac{1}{\lambda_j\sqrt{1-\lambda_j^2}}} \geq 0.99C_{n/2,k}2^{k/2} \geq 0.98C_{n,k}.$$

□

In order to understand the integral on (2.3), we write $\mathbb{R}^k$ as $\rho n^{-1/4}B_k \cup (\mathbb{R}^k \setminus \rho n^{-1/4}B_k)$ where $\rho > 0$. Inside the ball $\rho n^{-1/4}B_k$ the integral is close to 1. Outside the ball, we show that the integral is negligible.

The estimation inside the ball of radius $\rho n^{-1/4}$ uses standard inequalities and corollary 2.12 (see Appendix 2.B for the proof). In this proposition we define an upper bound on ρ.

Proposition 2.15. *Let $k \le \alpha_1\sqrt{n}$. Let Λ and U be as before. Then with probability greater than $1 - 4e^{-\sqrt{n}}$*

$$\int_{\rho n^{-1/4}B_k} f(x)\left(1-\left|\Lambda^{-1}U^Tx\right|^2\right)_+^{(n/2-k-2)/4}\left(1-\left|\left(I-\Lambda^2\right)^{-1/2}U^Tx\right|^2\right)_+^{(n/2-k-2)/4}dx$$
$$\ge 0.95\int_{\rho n^{-1/4}B_k} f(x)\left(1-|x|^2\right)_+^{(n-k-2)/2}dx.$$

Using the Laplace method (see Appendix 2.B), we estimate the integral outside $\rho n^{-1/4}B_k$.

Proposition 2.16. *Let $k \le \alpha_1\sqrt{n}$. Then, for any $f : \mathbb{R}^n \to \mathbb{R}_+$ with $\|f\|_\infty \le e^{\alpha_2\sqrt{n}}$, we have*

$$I = C_{n,k}\int_{\mathbb{R}^k\setminus\rho n^{-1/4}B_k} f(x)\left(1-|x|^2\right)_+^{(n-k-2)/2}dx \le \frac{2\alpha_1}{\rho^2}e^{-\alpha_2\sqrt{n}}.$$

The constant $C_{n,k}$ is the same as in Proposition 2.7, and $\rho > 0$ is the same as in Proposition 2.15.

Proof of Theorem 2.2. By Proposition 2.8 and the Cauchy-Schwarz inequality, the event

$$\sqrt{\int_{S_H} f(x)d\sigma_H(x)\int_{S_{H^\perp}} f(x)d\sigma_{H^\perp}(x)} \ge 0.9$$

has the greater probability than the event that

$$\int_{\mathbb{R}^k} f(x)\left(1-\left|\Lambda^{-1}U^Tx\right|^2\right)_+^{(n/2-k-2)/4}\left(1-\left|\left(I-\Lambda^2\right)^{-1/2}U^Tx\right|^2\right)_+^{(n/2-k-2)/4}dx$$

is at least

$$0.9C_{n/2,k}^{-1}\left(\sqrt{\prod_{j=1}^k \lambda_j\sqrt{1-\lambda_j^2}}\right).$$

By Corollary 2.14 and Proposition 2.15 with probability greater than $1 - 4e^{-\sqrt{n}}$ the left hand side is at least

$$0.93C_{n,k}\int_{\rho n^{-1/4}B_k} f(x)(1-|x|^2)_+^{(n-k-2)/2}dx.$$

By Proposition 2.7 this is equal to

$$0.93\left(\int_{S^{n-1}} f(x)d\sigma_{n-1}(x) - C_{n,k}\int_{\mathbb{R}^k\setminus\rho n^{-1/4}B_k} f(x)(1-|x|^2)_+^{(n-k-2)/2}dx\right).$$

By Proposition 2.16 there exists $\hat{C} > 0$ such that for all $n > \hat{C}$ we have

$$0.93C_{n,k} \int_{\mathbb{R}^k \setminus \rho n^{-1/4} B_k} f(x)(1-|x|^2)^{(n-k-2)/4} dx \leq 0.01.$$

Hence,

$$0.93 \left(\int_{S^{n-1}} f(x) d\sigma_{n-1}(x) - C_{n,k} \int_{\mathbb{R}^k \setminus \rho n^{-1/4} B_k} f(x)(1-|x|^2)_+^{(n-k-2)/2} dx \right) \geq 0.9.$$

□

2.A Protocol for VSP

The protocol we present here is a simple modification of the one presented by Raz [33].

As before, c, c_1, C etc. denote positive universal constants.

Let $k = \lfloor e^{\sqrt{n}} \rfloor$. Let $E_1, \ldots, E_k \subseteq \mathbb{R}^n$ be independent random subspaces of dimension $\lfloor C_1 \sqrt{n} \rfloor$ chosen uniformly. For every $1 \leq i \leq k$, let $\mathcal{N}_i = \{\theta_1^i, \ldots, \theta_m^i\}$ be independent random vectors in $S^{n-1} \cap E_i$, where $m = \lfloor e^{C_2 \sqrt{n}} \rfloor$. Alice and Bob sample $(E_1, \mathcal{N}_1), \ldots, (E_k, \mathcal{N}_k)$ in advance and store the results. Each real number stored by Alice and Bob is kept with accuracy of $\log n$ bits. The protocol will be the following: Alice chooses a random index $1 \leq \hat{i} \leq k$, and then finds the index $1 \leq \hat{j} \leq m$ such that

$$\max_{1 \leq j \leq m} \left\langle \theta_j^{\hat{i}}, u \right\rangle = \left\langle \theta_{\hat{j}}^{\hat{i}}, u \right\rangle.$$

Alice sends Bob both indices $\hat{i}$ and $\hat{j}$ using at most $\log k + \log m = (1 + C_2)\sqrt{n}$ bits. Bob checks the distance of $\theta_{\hat{j}}^{\hat{i}}$ to H and $H^\perp$. If $d(\theta_{\hat{j}}^{\hat{i}}, H) > d(\theta_{\hat{j}}^{\hat{i}}, H)$ then they answer that $u \in H$ otherwise they answer that $u \in H^\perp$.

In this protocol Alice preforms one measurement. This measurement is in a subspace of dimension $O(\sqrt{n})$, hence the protocol has total rank of $O(\sqrt{n})$.

The analysis of this protocol is done in two steps. First we show that the protocol works when we replace $(E_1, \mathcal{N}_1), \ldots, (E_k, \mathcal{N}_k)$ with shared random pair $(E, \mathcal{N})$. The complexity of the public coin protocol is $\log m = C_2 \sqrt{n}$ bits. Second, we eliminate the need for shared randomness by considering $(E_1, \mathcal{N}_1), \ldots, (E_k, \mathcal{N}_k)$. This step is standard, and the cost of eliminating the shared randomness is another $\log k = \sqrt{n}$ bits. We present the main ideas of these steps.

In the analysis of the first step, we use two standard results: In the first, we use the fact that the norm of a projection of a random vector is close to Gaussian [18].

Proposition 2.17. *Let $v \in S^{d-1}$ be a random vector distributed uniformly. Let $F \subseteq \mathbb{R}^d$ be a subspace of dimension ℓ. Then,*

$$\mathbb{P}\left(\left||Proj_F v|^2 - \frac{\ell}{d}\right| \geq t\right) \leq Ce^{-ct^2 d}, \quad \forall t.$$

Note that by applying a random rotation we may assume that v is fixed and F is random.

The second standard result shows that our choice of $\mathcal{N}_i$ is typically an $1/2-$net of $S^{n-1} \cap E_i$.

Proposition 2.18. *Let $z_1, ..., z_\ell$ be independent uniformly chosen random vectors in S^{d-1}, where $\ell = e^{Cd}$. Then with probability greater than $1 - e^{-e^{c\ell}}$ they form an $1/2-$net of the sphere.*

Sketch of the proof. Let N be an $\varepsilon-$net with $\#N = e^{c_1 k}$ (e.g [32]). For any $x \in N$ we have

$$\mathbb{P}\left(|z_i - x| > \varepsilon \ \forall i\right) \leq e^{-m\mathbb{P}(|z_1 - x| \leq \varepsilon)}.$$

Since

$$\mathbb{P}\left(|z_1 - x| \leq \varepsilon\right) \approx e^{-c_2(1-\varepsilon)^2 k},$$

we have

$$\mathbb{P}\left(\exists x \in N;\ |z_i - x| > \varepsilon \ \forall i\right) \leq \exp\left(c_1 k - e^{(c - c_2(1-\varepsilon)^2)k}\right).$$

□

We are now ready to prove that the protocol works with a shared random pair $(E, \mathcal{N})$.

Proof. Let u, H be fixed, such that either $u \in H$ or $u \in H^\perp$. We have,

$$\begin{aligned} \max_{\theta \in S^{n-1} \cap E} \langle u, \theta \rangle &= \max_{\theta \in S^{n-1} \cap E} \langle u, \mathrm{Proj}_E \theta \rangle = \max_{\theta \in S^{n-1} \cap E} \langle \mathrm{Proj}_E u, \theta \rangle \\ &= |\mathrm{Proj}_E u|. \end{aligned} \tag{2.5}$$

The dimension of E is $\lfloor C_1\sqrt{n}\rfloor$. By Proposition 2.17 and (2.5) we can choose C_1 big enough such that

$$\mathbb{P}\left(\max_{\theta\in S^{n-1}\cap E}\langle u,\theta\rangle \geq \frac{100}{n^{1/4}}\right)\geq 0.91.$$

Let $\theta_{\hat{j}}\in\mathcal{N}$ be the closest point to u. By Proposition 2.18, with probability greater than 0.99, the set $\mathcal{N}=\{\theta_1,\dots,\theta_m\}$ is an $1/2-$net of $S^{n-1}\cap E$. Hence, for any $\theta\in S^{n-1}\cap E$ there exists $\theta_i\in\mathcal{N}$ such that $|\theta_i-\theta|\leq 1/2$. Hence,

$$\begin{aligned}\langle\theta,u\rangle &= \langle\theta-\theta_i,u\rangle+\langle\theta_i,u\rangle \leq |\theta-\theta_i||\mathrm{Proj}_E u|+\max_j\langle\theta_j,u\rangle\\ &\leq \frac{1}{2}|\mathrm{Proj}_E u|+\max_j\langle\theta_j,u\rangle\,.\end{aligned}$$

The right hand side does not depend on θ, hence,

$$\begin{aligned}\max_j\langle\theta_j,u\rangle &\geq \max_{\theta\in S^{n-1}\cap E}\langle\theta,u\rangle-\frac{1}{2}|\mathrm{Proj}_E u|\\ &= \frac{1}{2}\max_{\theta\in S^{n-1}\cap E}\langle\theta,u\rangle\,.\end{aligned}$$

Let $\alpha=\left\langle\theta_{\hat{j}},u\right\rangle$. With probability greater than 0.9 we have

$$\alpha\geq\frac{50}{n^{1/4}}.$$

Let

$$\theta_{\hat{j}}=\alpha u+\sqrt{1-\alpha^2}v,$$

where $v\in S^{n-1}\cap u^\perp$. By the definition v, it is distributed uniformly in $S^{n-1}\cap u^\perp$. By Proposition 2.17 we have

$$\mathbb{P}\left(\left||\mathrm{Proj}_H v|^2-\frac{1}{2}\right|\geq\frac{10}{\sqrt{n}}\right)\leq 0.1.$$

Hence, if $u\in H$, then with probability greater than 0.8 we have,

$$|\mathrm{Proj}_H\theta_{\hat{j}}|^2=\alpha^2+(1-\alpha^2)|\mathrm{Proj}_H v|^2\geq\frac{1}{2}+\frac{1000}{\sqrt{n}},$$

and

$$|\mathrm{Proj}_{H^\perp}\theta_{\hat{j}}|^2\leq|\mathrm{Proj}_{H^\perp}v|^2\leq\frac{1}{2}+\frac{10}{\sqrt{n}}.$$

Hence, with probability greater than 0.8 we have $|\mathrm{Proj}_{H^\perp}\theta_{\hat{j}}|<|\mathrm{Proj}_H\theta_{\hat{j}}|$, thus the protocol would correctly determine that $u\in H$. The case $u\in H^\perp$ is proven similarly.

□

Next we explain how to eliminate the shared randomness.

Proof. We denote by $\theta_j \in \mathcal{N}$ the closest vector to u in $\mathcal{N}$. Let

$$A = \left\{ (u, H, E, \mathcal{N}); \begin{array}{c} |\mathrm{Proj}_H \theta_j|^2 > |\mathrm{Proj}_{H^\perp} \theta_j|^2 + 10/\sqrt{n}, \text{ if } u \in H \\ |\mathrm{Proj}_{H^\perp} \theta_j|^2 > |\mathrm{Proj}_H \theta_j|^2 + 10/\sqrt{n}, \text{ if } u \in H^\perp \end{array} \right\}.$$

Let $A_{u,H}$ and $A_{E,\mathcal{N}}$ denote the corresponding sections of the set A. By the previous step of shared randomness, for any fixed (u, H) we have,

$$\mathbb{P}_{(E,\mathcal{N})} \left((u, H) \in A_{E,\mathcal{N}} \right) \geq 0.8.$$

By the Chernoff-Hoeffding inequality, for any fixed u, H we have

$$\mathbb{P}\left(\frac{\#\{i;\ (E_i, \mathcal{N}_i) \in A_{u,H}\}}{m} \leq 0.8 \right) \leq e^{-ck}.$$

Hence, by Fubini's theorem, for most choices of $(E_1, \mathcal{N}_1), \ldots, (E_k, \mathcal{N}_k)$

$$\mathbb{P}_{u,H}\left(\frac{\#\{i;\ (E_i, \mathcal{N}_i) \in A_{u,H}\}}{m} \leq 0.8 \right) \leq e^{-c'k}.$$

Recall that Alice and Bob sample in advance the list $(E_1, \mathcal{N}_1), \ldots, (E_k, \mathcal{N}_k)$. Thus, with high probability, their protocol works for a any (u, H) outside a set of measure $e^{-c'e^{\sqrt{n}}}$. Hence, for any vector u and a subspace H we can find u' and H' for which the protocol works, $|u - u'| \leq 1/\sqrt{n}$ and $\|\mathrm{Proj}_H - \mathrm{Proj}_{H'}\|_{op} \leq 1/\sqrt{n}$. Therefore $|\mathrm{Proj}_H u - \mathrm{Proj}_{H'} u'| \leq 2/\sqrt{n}$ and the protocol works for arbitrary u and H. □

2.B Asymptotic estimates

Here we present the proofs of Propositions 2.15 and 2.16.

proof of Proposition 2.15. Assume the event $\left\|\Lambda - I/\sqrt{2}\right\|_{op} \leq C(\sqrt{\alpha_1} + \sqrt{2})/n^{1/4}$ holds true. By Proposition 2.11 this event has probability greater than $1 - 4e^{-\sqrt{n}}$. Define $\psi : \mathbb{R}^n \to \mathbb{R}$ by

$$\psi(x) = \left(1 - \left|\Lambda^{-1} U^T x\right|^2\right)_+^{(n/2-k-2)/4} \left(1 - \left|\left(I - \Lambda^2\right)^{-1/2} U^T x\right|^2\right)_+^{(n/2-k-2)/4}.$$

Using the Taylor expansion

$$\log(1 - |x|^2) = -|x|^2 + O(|x|^4)$$

for any $|x| < 3/4$, we have

$$\psi(x) = \exp\left(-(n/8 - k/4 - 1/2)\left(\left|U\Lambda^{-1}U^T x\right|^2 + \left|U(I-\Lambda^2)^{-1/2}U^T x\right|^2 + O(|x|^4)\right)\right),$$

for any $|x| \leq 1/10$. By Corollary 2.12 for any $x \in \rho n^{-1/4} B_k$ we have

$$\psi(x) = \exp\Big(-(n-k-2)|x|^2/2 + O(\rho^2(\sqrt{\alpha_1}+\sqrt{2})^2) + O(\rho^4) + O((\alpha_1 + 1/\sqrt{n})\rho^2)\Big).$$

In Corollary 2.14 we assumed an upper bound on α_1. Under this upper bound assumption we can choose $\rho > 0$ small enough, independent of n and any specific choice of α_1 such that

$$\psi(x) \geq 0.95 \exp\left(-(n/2 - k/2 - 1)|x|^2\right) \geq 0.95(1 - |x|^2)^{(n-k-2)/2},$$

for all $x \in \rho n^{-1/4} B_k$. □

proof of Proposition 2.16. By the assumption on f we have

$$I \leq C_{n,k} e^{\alpha_2\sqrt{n}} \int_{\mathbb{R}^k \setminus \rho n^{-1/4} B_k} \left(1 - |x|^2\right)_+^{(n-k-2)/2} dx$$

Using $1 - x \leq e^{-x}$ and the assumption $k \leq \alpha_1\sqrt{n}$ and that α_1 is bounded by some universal constant, for n exceeding some universal constant, we have

$$I \leq C_{n,k} e^{\alpha_2\sqrt{n}} \int_{\mathbb{R}^k \setminus \rho n^{-1/4} B_k} e^{-(n/2-k/2-1)|x|^2} dx \leq C_{n,k} e^{\alpha_2\sqrt{n}} \int_{\mathbb{R}^k \setminus \rho n^{-1/4} B_k} e^{-n|x|^2/3} dx.$$

By integrating in polar coordinates, we have

$$\begin{aligned} I &\leq C_{n,k} k \mathrm{Vol}\left(B_k\right) e^{\alpha_2\sqrt{n}} \int_{\rho n^{-1/4}}^{\infty} r^{k-1} e^{-nr^2/3} dr \\ &= C_{n,k} k \mathrm{Vol}\left(B_k\right) e^{\alpha_2\sqrt{n}} \frac{1}{n^{k/2}} \int_{\rho n^{1/4}}^{\infty} r^{k-1} e^{-r^2/3} dr. \end{aligned}$$

Define $h(r) = -(k-1)\log r + r^2/3$. The function h is convex, hence

$$h(r) \geq h\left(\rho n^{1/4}\right) + h'\left(\rho n^{1/4}\right)\left(r - \rho n^{1/4}\right).$$

Assuming $\alpha_1 \leq \rho^2/6$ we have,

$$h'\left(\rho n^{1/4}\right) = -\frac{k-1}{\rho n^{1/4}} + \frac{2}{3}\rho n^{1/4} \geq \frac{1}{2}\rho n^{1/4}.$$

We have,

$$\int_{\rho n^{1/4}}^{\infty} e^{-h(r)} dr \leq e^{-h\left(\rho n^{1/4}\right)} \int_{\rho n^{1/4}}^{\infty} e^{-\rho n^{1/4}\left(r-\rho n^{1/4}\right)/2} dr = \frac{2}{\rho n^{1/4}} e^{-h\left(\rho n^{1/4}\right)}.$$

Hence,

$$I \leq 2C_{n,k} \mathrm{Vol}\left(B_k\right) \rho^{k-2} k n^{-k/4-1/2} e^{-\sqrt{n}(\rho^2/3-\alpha_2)}.$$

Using

$$C_{n,k}\mathrm{Vol}(B_k) = \frac{n-k}{n} \frac{\mathrm{Vol}(B_{n-k})\mathrm{Vol}(B_k)}{\mathrm{Vol}(B_n)} = \frac{n-k}{n} \binom{n/2}{k/2} \leq \left(\frac{n \cdot e}{k}\right)^{k/2},$$

we have

$$I \leq \frac{2}{\rho^2} (\sqrt{e}\rho)^k \frac{n^{k/4-1/2}}{k^{k/2-1}} e^{-\sqrt{n}(\rho^2/3-\alpha_2)}.$$

Assuming $\alpha_1 > 0$ is small enough such that $\rho^2/3 - \alpha_1 \log(\rho\sqrt{e/\alpha_1}) > 0$, we optimize over k, and get

$$I \leq \frac{2\alpha_1}{\rho^2} \exp\left[-\sqrt{n}\left(\rho^2/3 - \alpha_1 \log(\rho\sqrt{e/\alpha_1}) - \alpha_2\right)\right].$$

Hence, we can choose

$$2\alpha_2 < \rho^2/3 - \alpha_1 \log(\rho\sqrt{e/\alpha_1}),$$

to finish the proof. □

Chapter 3

Intersection with Random Geodesics

3.1 Introduction

The main question of this paper is whether the length of the intersection with a random geodesic curve can represent faithfully the measure of a given set. While we consider it a natural question in geometry, there are additional motivations for such an investigation. One such motivation is connected to the efficiency of algorithmic sampling results, such as *hit and run* [28].

Given a subset A of some ambient space M, the length of the intersection of A with a random geodesic curve is related to the probability of escaping the set A by choosing a uniform point on the geodesic. This is related to the notion of *conductance*, which is key in studying the effectiveness of hit and run algorithms (e.g [28]).

We start with the unit sphere $S^{n-1} = \{x \in \mathbb{R}^n;\ |x| = 1\}$ as our ambient space. In this case, we expand our scope from geodesics to higher dimensional subspaces. Let $A \subseteq S^{n-1}$ be a measurable subset of the n-dimensional sphere. Let $H \subseteq \mathbb{R}^n$ be a random k-dimensional linear subspace. We investigate the random variable

$$X = \frac{\sigma_H(A \cap H)}{\sigma_{n-1}(A)},$$

where σ_{n-1} and σ_H are the rotationally invariant probability measures on S^{n-1} and $S^{n-1} \cap H$ respectively.

The case $k = 2$ is our original question of intersection of a set with a random geodesic. In [22], Klartag and Regev used the case of $k = n/2$ in order to give a lower bound on the communication complexity of the *Vector in Subspace Problem*. The connection between the concentration phenomenon of random intersections and communication complexity is through the *rectangle method*. Here, high concentration of X shows that any protocol would require the exchange of many bits in order to distinguish between different states.

In [22] Klartag and Regev showed that when the dimension k is large, the random variable X is highly concentrated around its mean 1. Their proof consisted of two main steps. First, they dealt with the case where H is a random hyperplane ($k = n - 1$). Then they used the result of hyperplanes repeatedly in order to obtain concentration inequalities for k which depends on n linearly.

The result of Klartag and Regev is sharp, but their method proves to be more difficult when the dimension of H is low. Hence we employ a direct analysis of the affect the intersection has, by defining the Radon transform, as we discuss in §3.2. In §3.3 we build upon this analysis and obtain a dimension free estimate on the probability of X diverging from one by a fixed percentage, for random geodesics (the case $k = 2$). This result requires us to consider sets of large measure.

Theorem 3.1. *Let $A \subseteq S^{n-1}$ be a measurable subset, such that $\sigma_{n-1}(A) = 1/2$ then*

$$\mathbb{P}_L\left(\left|\frac{\sigma_L(A \cap L)}{\sigma_{n-1}(A)} - 1\right| \geq \frac{1}{2^{1/3}}\right) \leq \frac{1}{2^{1/3}},$$

where L is a uniformly chosen geodesic curve on S^{n-1} and σ_L is the uniform probability measure on $S^{n-1} \cap L$.

Remark 3.17 in §3.3 shows that we cannot improve the probability bound to a bound that tends to zero when the dimension grows.

In §3.4 we show that there is no analogous result for random geodesics inside convex sets. We show, that for any convex body, we can construct a subset of half the volume of the body, such that a random geodesic curve would either miss it or its complement with high probability.

Theorem 3.2. *Let $K \subseteq \mathbb{R}^n$ be a convex body. There exists a set $A \subseteq K$ such that*

Vol(A) / *Vol*$(K) = 1/2$ *and*

$$\mathbb{P}\left(\frac{\textit{length}(L \cap A)}{\textit{length}(L \cap K)} \in \{0,1\}\right) = 1 - O^*\left(\frac{1}{\sqrt{n}}\right),$$

where $L = X + \mathbb{R}\theta$, X *and* θ *are independent and distributed uniformly in* K *and* S^{n-1} *respectively.*

Here, O^* represent the big O notation up to logarithmic factors.

The above construction relies on the concentration of measure in convex bodies. It is known that for an isotropic convex body, most of its mass is concentrated in a thin spherical shell. Combining this with the concentrated one dimensional marginals of a uniform random vector on the sphere, we see that a typical random geodesic will miss a neighborhood of the barycenter of the body. This allows us to construct the desired subset on the convex body.

The different phenomena observed in Theorems 3.1 and 3.2, raise the question: which of the two would occur for random geodesics in different spaces? In §3.5 we show how the same tools used to analyze random geodesics on the sphere can be used on the discrete tours $\mathbb{Z}/p\mathbb{Z}$ where p is prime, and obtain a similar result to Theorem 3.1.

It would be interesting to understand which other ambient spaces M behave similarly to the sphere and the discrete torus, and which behave similarly to convex sets. A more general question is whether we can find a sufficient condition for such phenomena. A possible candidate would be a curvature condition.

Question 3.3. *Does a positive Ricci curvature implies a theorem analogous to Theorem 3.1?*

The main tool for understanding both the sphere and the discrete torus is to define the appropriate *Radon transform*, that averages functions on subspaces. In §3.2 we develop our method to calculate its singular decomposition. The calculations in §3.2 are not limited to the case of geodesics ($k = 2$), and in §3.6 we expand our analysis to all $2 \leq k \leq n - 1$.

One of the consequences of this analysis is a bound on the variance of the random variable $\sigma_H(A \cap H)/\sigma_{n-1}(A)$.

Theorem 3.4. *Let $A \subseteq S^{n-1}$ be a measurable set. Let $2 \leq k \leq n-1$, and let $H \subseteq \mathbb{R}^n$ be a random subspace of dimension k. Then,*

$$Var\left(\frac{\sigma_H(A \cap H)}{\sigma_{n-1}(A)}\right) \leq \frac{n-k}{k(n-1)}\left(\frac{1}{\sigma_{n-1}(A)} - 1\right).$$

The above analysis could possibly lead to a direct proof of the result of Klartag and Regev, or to generalize [15], where we average the measures of the intersections with a random subspace and its orthogonal complement.

3.2 The Radon Transform

In this section we introduce the Radon transform, an integral transform that averages a function along a given subspace. We denote by $G_{n,k}$ the Grassmanian manifold of all k dimensional subspaces of $\mathbb{R}^n$.

Definition 3.5. *Let $1 \leq k \leq n$. The Radon transform $R_k : L^2(S^{n-1}) \to L^2(G_{n,k})$ is defined by*

$$R_k f(E) = \int_{S^{n-1} \cap E} f(x) d\sigma_E(x),$$

where σ_E is the $SO(n)$ invariant Haar probability measure on $S^{n-1} \cap E$.

We can see that the radon transform gives a functional version of our geometric question. Taking f to be the normalized indicator of a subset $A \subseteq S^{n-1}$, and E to be a random subspace, we obtain

$$R_k f(E) = \int_{S^{n-1} \cap E} \frac{1_A(x)}{\sigma_{n-1}(A)} d\sigma_E(x) = \frac{\sigma_E(A \cap E)}{\sigma_{n-1}(A)}.$$

Hence, understanding the singular values of the radon transform, will give us a clearer picture of the behavior under random intersections.

In this section, we express the singular values of the Radon transform R_k by a one dimensional integral. In later sections we analyze this expression in order to understand random intersections on the sphere by k dimensional subspaces.

In order to find the singular values we introduce the conjugate transform of the Radon transform.

Definition 3.6. *Let $1 \leq k \leq n$. The conjugate Radon transform $R_k^* : L^2(G_{n,k}) \to L^2(S^{n-1})$ is defined by*

$$R_k^* g(\theta) = \int_{\{E;\, \theta \in E\}} g(E) d\mu_{k,\theta}(E),$$

Where $\mu_{k,\theta}$ is the Haar probability measure on

$$\{E \in G_{n,k};\ \theta \in E\},$$

invariant under the $SO(n-1)$ action of rotations on $\mathbb{R}^n$ that fix θ.

For every $1 \leq k \leq n$ we define the operator $S_k = R_k^* R_k$. By definition, if $\lambda_{k,i}^2$ is an eigenvalue of S_k then $\lambda_{k,i} > 0$ is a singular value of R_k.

We use the symmetries of S_k in order to show that the spherical harmonics are its eigenfunctions (for more details about spherical harmonics see [30]).

Proposition 3.7. *The eigenfunctions of S_k are the spherical harmonics.*

Proof. The space of spherical harmonics of fixed degree is an irreducible representation of $SO(n)$. Hence, by Schur's lemma, it is enough to show the operator S_k commutes with the $SO(n)$ action. Let $U \in SO(n)$, let $f \in L^2(S^{n-1})$ and let $g \in L^2(G_{n,k})$. We denote by T_U and T'_U the action of $SO(n)$ (the Koopman representation) on $L^2(S^{n-1})$ and $L^2(G_{n,k})$ respectively,

$$T_U f(x) = f(U^{-1}x),\ T_U g(H) = g(U^{-1}H).$$

By the definitions of the Radon transform and the measures, we have

$$\begin{aligned} T'_U (R_k f)(H) = R_k f(U^{-1}H) &= \int_{U^{-1}H \cap S^{n-1}} f(x) d\sigma_{U^{-1}H}(x) = \int_{H \cap S^{n-1}} f(U^{-1}x) d\sigma_H(x) \\ &= R_k (T_U f)(H). \end{aligned}$$

Since R_k^* is conjugate to R_k we have

$$T_U (R_k^* g)(\theta) = R_k^* (T'_U g)(\theta).$$

Hence,

$$T_U (S_k f)(\theta) = T_U (R_k R_k^* f)(\theta) = R_k R_k^* (T_U f)(\theta) = S_k (T_U f)(\theta).$$

□

By considering the map $\psi : [-1,1]\times S^{n-2} \to S^{n-1}$ defined by $\psi(s,x) = (s, \sqrt{1-s^2}x)$, we obtain the following standard formula for integrating on the sphere,

Proposition 3.8. *Let $f \in L^2\left(S^{n-1}\right)$. Then,*

$$\int_{S^{n-1}} f(x) d\sigma_{n-1}(x) = \tau_n \int_{-1}^{1} \left(\int_{S^{n-2}} f\left(t, \sqrt{1-t^2}y\right) d\sigma_{n-2}(y) \right) \left(1-t^2\right)^{(n-3)/2} dt,$$

where $\tau_n = \Gamma(n/2)/(\sqrt{\pi}\Gamma(n/2-1/2)) = \sqrt{n/(2\pi)} + O(1/\sqrt{n})$.

Let P_ℓ denote the Gegenbauer polynomial of degree ℓ with respect to the weight function $(1-t^2)^{(n-3)/2}$. We use the standard normalization [30],

$$P_\ell(1) = \binom{\ell+n-3}{\ell}.$$

It is well known (e.g [30]) that we can define a spherical harmonic of degree ℓ by using the Gegenbauer polynomial of the same degree.

Proposition 3.9. *Let $\xi \in S^{n-1}$, then $f_\ell(\theta) = P_\ell(\langle \theta, \xi \rangle)$ is a spherical harmonic of degree ℓ.*

Since the space of spherical harmonics of a fixed degree is irreducible, they all share the same eigenvalue for S_k. Hence, we can combine both propositions, and express the eigenvalues of S_k by one dimensional integrals of the Gegenbauer polynomials. We note that when ℓ is odd and f_ℓ is a spherical harmonic of degree ℓ, it is an odd function. Hence, $R_k f_\ell \equiv 0$, or equivalently, $\lambda_{k,\ell} = 0$ for odd ℓ. Therefore we need to consider only spherical harmonics of even degree.

Proposition 3.10. *Let $\ell \geq 0$ be even. Let $\lambda_{k,\ell}^2$ be the eigenvalue of S_k that corresponds to spherical harmonics of degree ℓ. Then,*

$$\lambda_{k,\ell}^2 = \frac{\tau_k \int_{-1}^{1} P_\ell(t) \left(1-t^2\right)^{(k-3)/2} dt}{\binom{\ell+n-3}{\ell}}.$$

Proof. We choose the spherical harmonic $f_\ell(x) = P_\ell\left(\langle x, e_1 \rangle\right)$, where $e_1 = (1,0,\ldots,0)$. By definition of $\lambda_{k,\ell}$ we have,

$$\left(S_k f_\ell\right)(\theta) = \lambda_{k,\ell}^2 f_\ell(\theta).$$

By choosing $\theta = e_1$ and the normalization of the Gegenbauer polynomials, we have

$$\left(S_k f_\ell\right)(e_1) = \lambda_{k,\ell}^2 P_\ell(1) = \binom{\ell+n-3}{\ell} \lambda_{k,\ell}^2.$$

Hence, we need to show that

$$\left(S_k f_\ell\right)(e_1) = \tau_k \int_{-1}^{1} P_\ell(t)\left(1-t^2\right)^{(k-3)/2} dt.$$

By definition of the operators R_k and R_k^* we have,

$$\left(S_k f_\ell\right)(e_1) = \left(R_k^* R_k f_\ell\right)(e_1) = \int_{\{E;\, e_1 \in E\}} \int_{S^{n-1}\cap E} f_\ell(x) d\sigma_E(x) d\mu_{k,e_1}(E).$$

Using Proposition (3.8) on the inner integral, we have

$$\left(S_k f_\ell\right)(e_1) = \tau_k \int_{\{E;\, e_1 \in E\}} \int_{-1}^{1} \int_{S^{n-1}\cap E\cap e_1^\perp} f_\ell(t, \sqrt{1-t^2}y) d\sigma_{E\cap e_1^\perp}(y)\left(1-t^2\right)^{(k-3)/2} dt d\mu_{k,e_1}.$$

By our choice of f_ℓ we have $f_\ell(t,y) = P_\ell(t)$. Hence, the integrals with respect to $\sigma_{E\cap e_1^\perp}$ and μ_{k,e_1} are on a constant function, and the proof is complete. □

In [15] we studied the random variable

$$\sqrt{\int_{S^{n-1}\cap H} f d\sigma_H \int_{S^{n-1}\cap H^\perp} f d\sigma_{H^\perp}},$$

where $\dim(H) = n/2$. Using a similar method as above we can calculate the second moment of this random variable.

Theorem 3.11. *Let $1 \le k \le n-1$. For any $f, g \in L^2(S^{n-1})$ we have,*

$$\left| \int R_k f(H) R_{n-k} g(H^\perp) - \int f d\sigma_{n-1} \int g d\sigma_{n-1} \right| \le \sum_{j=1}^{\infty} \binom{\ell + n/2 - 2}{\ell} \binom{2\ell + n - 3}{2\ell}^{-1} \|f_{2\ell}\| \, \|g_{2\ell}\|,$$

and

$$\int R_k f(H) R_{n-k} f(H^\perp) = \sum_{j=0}^{\infty} (-1)^\ell \binom{\ell + n/2 - 2}{\ell} \binom{2\ell + n - 3}{2\ell}^{-1} \|f_{2\ell}\|^2,$$

where $f_{2\ell}$ and $g_{2\ell}$ are the projections of f and g to the space of spherical harmonics of degree 2ℓ.

Proof. Let $\varphi : G_{n,k} \to G_{n,n-k}$ be the involution

$$\varphi(H) = H^\perp.$$

As before, the operator $R_k^* \varphi R_{n-k}$ commutes with rotations, hence it is diagonalized by the spherical harmonics. Let $P_{2\ell}$ be the Gegenbauer polynomial of degree 2ℓ

and $f_{2\ell}$ the spherical harmonics defined by $f_{2\ell}(x) = P_{2\ell}(x_1)$. Denoting by $\eta_{2\ell}$ the corresponding eigenvalue, we have

$$\eta_{2\ell} P_{2\ell}(1) = \eta_{2\ell} f_{2\ell}(e_1) = R_k^* \varphi R_{n-k} f_{2\ell}(e_1) = \int_{\{H;\, e_1 \in H\}} \int_{S^{n-1} \cap H^\perp} f_{2\ell} d\sigma_{H^\perp} d\mu_{e_1}.$$

Since $e_1 \in H$ for every $x \in H^\perp$ we have $\langle x, e_1 \rangle = 0$. Hence, the inner integral is on the constant function $f_{2\ell}(0)$, and we obtain

$$\eta_{2\ell} P_{2\ell}(1) = f_{2\ell}(0) = P_{2\ell}(0).$$

Hence, by using [36, equation 4.7.31]

$$\eta_{2\ell} = \frac{P_{2\ell}(0)}{P_{2\ell}(1)} = (-1)^\ell \binom{\ell + n/2 - 2}{\ell} \binom{2\ell + n - 3}{2\ell}^{-1}.$$

Let f and g be a spherical harmonics of degree 2ℓ, we have

$$\langle R_k f, \varphi R_{n-k} g \rangle = \langle f, R_k^* \varphi R_{n-k} g \rangle \leq |\eta_{2\ell}| \, \|f\| \, \|g\| .$$

If $f = g$ then we can get equality

$$\langle R_k f, \varphi R_{n-k} f \rangle = \eta_{2\ell} \, \|f\|^2 .$$

In order to finish the proof, we remember the spherical harmonics of different degrees are orthogonal to each other. □

Remark 3.12. *Theorem 3.11 shows that the integral over the Grassmanian*

$$\int R_k f(H) R_{n-k} f(H^\perp),$$

does not depend on the dimension of H.

Note that $P_{2\ell}(0)/P_{2\ell}(1)$ are the singular values of the Radon transform R_{n-1}. Hence, we can repeat the analysis of Klartag and Regev for estimating the norm of the projections and the sum. We get,

Corollary 3.13. *Let $f, g : S^{n-1} \to [0, \infty)$. Assume that $\int f = \int g = 1$, then for all $1 \leq k \leq n-1$*

$$\left| \int R_k f(H) R_{n-k} g(H^\perp) - 1 \right| \leq C \frac{\log(2 \|f\|_\infty) \log(2 \|g\|_\infty)}{n}.$$

3.3 Random Geodesics on the Sphere

The goal of this section is to prove Theorem 3.1. We do this by calculating the singular values of the Radon transform R_2 and use them to bound the variance of the random variable $\sigma_L(A \cap L)/\sigma_{n-1}(A)$.

There are two natural ways to get a random geodesics on the sphere. The first is intersecting the sphere with a random 2 dimensional subspace. The second is by choosing a uniform random point $X \in S^{n-1}$ and a uniform direction $Y \in T_X S^{n-1}$ of unit length, and define L as the geodesic curve that starts at X in the direction Y. We note that these two methods have the same distribution.

Proposition 3.14. *The eigenvalues of S_2 are*

$$\lambda_{2,2\ell}^2 = \binom{\ell + n/2 - 2}{\ell}^2 \binom{2\ell + n - 3}{2\ell}^{-1}.$$

Proof. By Proposition 3.10 we have

$$\lambda_{k,\ell}^2 = \frac{\tau_k \int_{-1}^{1} P_\ell(t)\,(1-t^2)^{(k-3)/2}\,dt}{\binom{\ell+n-3}{\ell}}.$$

Using the change of variables $t = \cos\theta$, we obtain

$$\lambda_{2,2\ell}^2 = \frac{\tau_2 \int_0^{\pi} P_{2\ell}(\cos t)dt}{\binom{2\ell+n-3}{2\ell}}.$$

We have

$$\tau_2 = \left(\int_{-1}^{1} \frac{1}{\sqrt{1-s^2}} ds\right)^{-1} = 1/\pi,$$

and by the zero coefficient of the trigonometric polynomial $P_{2\ell}(\cos t)$ (see [36, equation 4.9.19]), we have

$$\int_0^{\pi} P_{2\ell}(\cos t)dt = \pi \binom{\ell + n/2 - 2}{\ell}^2.$$

□

Next we prove that the biggest eigenvalue of S_2 is $\lambda_{2,2}^2$.

Proposition 3.15. *The eigenvalues of S_2, $\{\lambda_{2,2\ell}^2\}$ is a decreasing sequence.*

Proof. By Proposition 3.14, we have

$$\frac{\lambda_{2,2\ell+2}^2}{\lambda_{2,2\ell}^2} = \frac{\binom{\ell+n/2-1}{\ell+1}^2 \binom{2\ell+n-3}{2\ell}}{\binom{2\ell+n-1}{2\ell+2}\binom{\ell+n/2-2}{\ell}^2} = \frac{(2\ell+1)(n+2\ell-2)}{(2\ell+2)(n+2\ell-1)}.$$

Hence, for the sequence to be decreasing, we need the ratio to be at most one. This condition is equivalent to

$$n + 4\ell \geq 0,$$

which is always satisfied. □

By Proposition 3.14 we have

$$\lambda_{2,2}^2 = \frac{n-2}{2(n-1)}.$$

Combining this with Proposition 3.15, we can give an upper bound for the variance of a Radon transform of a function by the variance of the original function.

Corollary 3.16. *Let $f \in L^2\left(S^{n-1}\right)$ such that $\int_{S^{n-1}} f(x)d\sigma_{n-1} = 0$. Then*

$$\|R_2 f\|_{L^2(G_{n,2})} \leq \frac{1}{\sqrt{2}} \|f\|_{L^2(S^{n-1})}.$$

Using Corollary 3.16 we can prove Theorem 3.1.

Proof. Define $f(x) = 1_A(x)/\sigma_{n-1}(A) - 1$ and $X = \sigma_L(A \cap L)/\sigma_{n-1}(A)$. By Corollary 3.16 we have

$$\|R_2 f\|^2 \leq \frac{1}{2} \|f\|^2 = \frac{1}{2} \left(\frac{1}{\sigma_{n-1}(A)} - 1 \right).$$

On the other hand,

$$\|R_2 f\|^2 = \int_{G_{n,2}} \left(\int_{S^{n-1} \cap H} f(x) d\sigma_H(x) \right)^2 d\mu_2 = \mathrm{Var}\left(X\right).$$

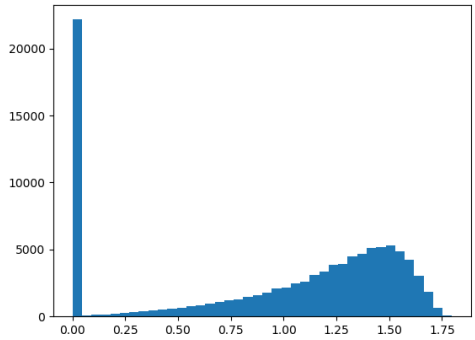
By Markov's inequality

$$\mathbb{P}_L\left(|X - 1| \geq t\right) \leq \frac{\mathrm{Var}X}{t^2} \leq \frac{1 - \sigma_{n-1}(A)}{2\sigma_{n-1}(A)t^2}.$$

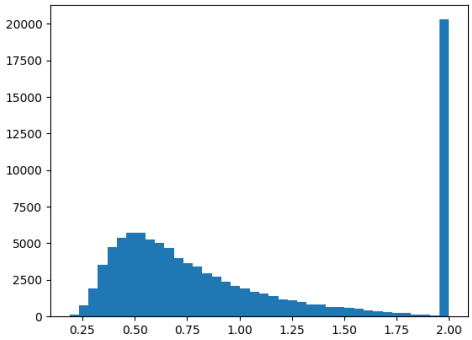
Setting $\sigma_{n-1}(A) = 1/2$ and $t = 1/2^{1/3}$ finishes the proof. □

Remark 3.17. *By looking at the set $A = \{x \in S^{n-1};\ |x_1| \geq T\}$ where $T \approx c/\sqrt{n}$ is chosen such that $\sigma_{n-1}(A) = 1/2$ we see that the probability bound in Theorem 3.1 cannot be improved to an expression that decays with the dimension n. See Figure 3.1 for simulation results of this example.*

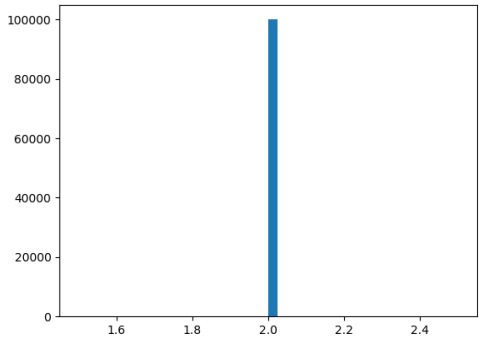
Remark 3.18. *As we can see in Figure 3.1, there are no small ball estimates on either side and there is no apparent information about the shape of the distribution.*



(a) $A = \{x \in S^{n-1};\ |x_1| \geq c/\sqrt{n}\}$



(b) $A = \{x \in S^{n-1};\ |x_1| \leq c/\sqrt{n}\}$



(c) $A = \{x \in S^{n-1};\ x_1 \geq 0\}$

Figure 3.1: Histogram of simulations on the sphere in dimension 1000 with 10^5 samples for different sets A.

3.4 Random Geodesics in a Convex Body

In this section we prove that there is no analogous theorem to Theorem 3.1 when we choose random geodesics inside a convex body. We show that for every convex body, we can find a set such that the relative length of the intersection with a random line is close to a zero-one law.

There are different distributions for geodesics inside a convex body. We focus on the following model; choose a random point $X \in K$ distributed uniformly, and a random direction $\theta \in S^{n-1}$ distributed uniformly on the sphere and independent of X. The random geodesic is defined by $L = \{X + \mathbb{R}\theta\}$.

This model of random geodesics is the *hit and run* model, commonly used in various algorithmic problems such as computing the volume of a convex body.

We say a convex body $K \subseteq \mathbb{R}^n$ is in *isotropic position* if a random vector X distributed uniformly inside K has $\mathbb{E}X = 0$ and $\mathbb{E}X \otimes X = Id$. We start by two observations about isotropic convex bodies.

Proposition 3.19. *Let $K \subseteq \mathbb{R}^n$ be an isotropic convex body. Let X be a random vector distributed uniformly in K. Then for any $a > 0$ and any $\varepsilon > 0$ we have*

$$\mathbb{P}\left(|\langle X, \xi\rangle| \geq a + \varepsilon\right) \geq \mathbb{P}\left(|\langle X, \xi\rangle| \geq a\right) - c\varepsilon,$$

and

$$\mathbb{P}\left(|\langle X, \xi\rangle| \leq a - \varepsilon\right) \geq \mathbb{P}\left(|\langle X, \xi\rangle| \leq a\right) - c\varepsilon,$$

where $c > 0$ is a universal constant and $\xi \in S^{n-1}$.

Proof. Let $f : \mathbb{R} \to \mathbb{R}_+$ be the density of $\langle X, \xi\rangle$, then f is log-concave, $\int tf(t)dt = 0$ and $\int t^2 f(t)dt = 1$. By [12] we have

$$\sup f(t) \leq ef(0).$$

By the Berwald-Borell lemma [6, 8] the functions $M^+(p) = (\Gamma(p+1))^{-1} \int_0^\infty t^p f(t)dt$ and $M^-(p) = (\Gamma(p+1))^{-1} \int_{-\infty}^0 t^p f(t)dt$ are log concave in $[-1, \infty)$. Hence,

$$M^{\pm}(0) \geq \left(M^{\pm}(-1)\right)^{2/3} \left(M^{\pm}(2)\right)^{1/3}.$$

Since f is a density function for an isotropic random variable, we have

$$M^+(-1) = M^-(-1) = f(0)$$
$$M^+(0) + M^-(0) = 1$$
$$M^+(2) + M^-(2) = 1/2.$$

Adding everything, and using $a^{1/3} + b^{1/3} \geq (a+b)^{1/3}$ for all $a, b \geq 0$, we have

$$1 = M^+(0) + M^-(0) \geq f^{2/3}(0) \left(\left(M^+(2)\right)^{1/3} + \left(M^-(2)\right)^{1/3} \right) \geq f^{2/3}(0) 2^{-1/3}.$$

Hence,

$$\sup f \leq e f(0) \leq e\sqrt{2}.$$

To conclude, we have

$$\begin{aligned} \mathbb{P}\left(|\langle X, \xi\rangle| \geq a + \varepsilon\right) &= \mathbb{P}\left(|\langle X, \xi\rangle| \geq a\right) - \mathbb{P}\left(a \leq |\langle X, \xi\rangle| \leq a + \varepsilon\right) \\ &\geq \mathbb{P}\left(|\langle X, \xi\rangle| \geq a\right) - 2\varepsilon \sup f \geq \mathbb{P}\left(|\langle X, \xi\rangle| \geq a\right) - c\varepsilon. \end{aligned}$$

The second statement follows similarly. □

Proposition 3.20. *Let $K \subseteq \mathbb{R}^n$ be an isotropic convex body. Let X be a random vector distributed uniformly in K. Let $\theta \in S^{n-1}$ and let $\ell : K \to \mathbb{R}_+$ be the length function in direction θ, that is $\ell(x) = length\left(K \cap \{x + \mathbb{R}\theta\}\right)$. Then*

$$\mathbb{P}(\ell(X) \geq t) \leq 2e^{-ct},$$

where $c > 0$ is a universal constant.

Proof. The function $\ell(x)$ depends only on the projection of x to the hyperplane $\theta^\perp$. Hence, by Fubini's theorem we have,

$$\mathbb{E}\ell^p(X) = \frac{1}{\mathrm{Vol}\left(K\right)} \int_{\mathrm{Proj}_{\theta^\perp} K} \ell^p(x)\ell(x)dx.$$

Let S_θ be the Steiner symmetrization in the θ direction and set $T = S_\theta K$ (for more details on the Steiner symmetrization see [7]). Since the Steiner symmetrization is in the θ direction the function ℓ and $\mathrm{Proj}_{\theta^\perp} K$ are preserved under it. Hence, for a random vectors X and Y distributed uniformly on K and T we have,

$$\mathbb{E}\ell^p(Y) = \mathbb{E}\ell^p(X).$$

In addition

$$\mathrm{Var}\,\langle Y,\theta\rangle \le \mathrm{Var}\,\langle X,\theta\rangle = 1.$$

Hence, by the Berwald-Borell lemma [6, 8] we have

$$\left(\mathbb{E}|\langle Y,\theta\rangle|^p\right)^{1/p} \le Cp\left(\mathbb{E}|\langle Y,\theta\rangle|^2\right)^{1/2} \le Cp.$$

Since $\langle Y,\theta\rangle$ is a symmetric random variable, we have

$$\begin{aligned}\mathbb{E}|\langle Y,\theta\rangle|^p &= \frac{1}{\mathrm{Vol}\,(T)}\int_{\mathrm{Proj}_{\theta^\perp}T}\left(\int_{-\ell(y)/2}^{\ell(y)/2}|t|^p dt\right)\ell(y)dy = \frac{1}{\mathrm{Vol}\,(K)}\int_{\mathrm{Proj}_{\theta^\perp}K}\left(\int_{-\ell(x)/2}^{\ell(x)/2}|t|^p dt\right)\ell(x)dx\\ &= \frac{1}{2^p(p+1)\mathrm{Vol}\,(K)}\int_{\mathrm{Proj}_{\theta^\perp}K}\ell^{p+1}(x)\ell(x)dx = \frac{1}{2^p(p+1)}\mathbb{E}\ell^{p+1}(X)\end{aligned}$$

We have,

$$\mathbb{E}\ell^p(X) = 2^{p-1}p\mathbb{E}|\langle Y,\theta\rangle|^{p-1} \le (2C)^{p-1}p^p \le C_1^p p!.$$

By Markov's inequality, for any $\alpha > 0$ we have

$$\mathbb{P}(\ell(X)\ge t) = \mathbb{P}(e^{\alpha\ell(X)}\ge e^{\alpha t}) \le e^{-\alpha t}\mathbb{E}e^{\alpha\ell(X)} = e^{-\alpha t}\sum_{p=0}^{\infty}\frac{\alpha^p}{p!}\mathbb{E}\ell^p(X) \le e^{-\alpha t}\sum_{p=0}^{\infty}(C_1\alpha)^p.$$

Choosing $\alpha = 1/(2C_1)$ concludes the proof. □

Another observation we use is the concentration of measure on the sphere. The following proposition can be proved by a direct computation or by the concentration of Lipschitz functions on the sphere.

Proposition 3.21. *Let $\theta \in S^{n-1}$ be a random vector chosen by the uniform distribution, and let $\xi \in S^{n-1}$ be a fixed direction. Then*

$$\mathbb{P}\left(|\langle\theta,\xi\rangle| \ge t\right) \le Ce^{-nt^2/2}, \quad \forall t > 0.$$

We are now ready to show that random geodesic sampling in any isotropic convex body can be close to a zero one law. Let $K \subseteq \mathbb{R}^n$ be a convex body. For any $\xi \in S^{n-1}$ define the set $A_\xi = \{x \in K;\ |\langle x,\xi\rangle| \ge t_\xi\}$, where $t_\xi > 0$ is chosen such that $\mathrm{Vol}\,(A_\xi)/\mathrm{Vol}\,(K) = 1/2$.

Proposition 3.22. *Let $K \subseteq \mathbb{R}^n$ be an isotropic convex body. Let X be a random vector distributed uniformly in K. Let $L = \{X + \mathbb{R}\theta\}$ where X is uniform in K and θ is uniform in S^{n-1}. For any $\xi \in S^{n-1}$ we have,*

$$\mathbb{P}\left(\frac{length(L\cap A_\xi)}{length(L\cap K)} \in \{0,1\}\right) = 1 - O^*\left(\frac{1}{\sqrt{n}}\right).$$

Proof. We define the following events:

$$D = \left\{ |\langle X, \xi \rangle| \geq t_\xi + c^{-1} \frac{\log^3 n}{\sqrt{n}} \right\},$$
$$P = \left\{ |\langle \theta, \xi \rangle| \leq \frac{\log n}{\sqrt{n}} \right\},$$
$$M = \left\{ \text{length}(L \cap K) \leq c^{-1} \log n \right\},$$

where $c > 0$ is the constant from Proposition 3.20. Assuming the event $D \cap P \cap M$, we have X inside A_ξ with distant to the inner boundary of A_ξ of at least $c^{-1} \log^3 n / \sqrt{n}$. In addition, the line passing through X can diverge in the ξ direction by at most $c^{-1} \log n \cdot \log n / \sqrt{n}$, hence it cannot escape it. Therefore,

$$\mathbb{P}\left(\frac{\text{length}(L \cap A_\xi)}{\text{length}(L \cap K)} = 1 \right) \geq \mathbb{P}(D \cap P \cap M).$$

By previous propositions

$$\mathbb{P}(P) \geq 1 - \frac{C}{n},$$
$$\mathbb{P}(M) \geq 1 - \frac{2}{n},$$

and

$$\mathbb{P}(D) = \frac{1}{2} - C' \frac{\log^3 n}{\sqrt{n}}.$$

Hence,

$$\mathbb{P}\left(\frac{\text{length}(L \cap A_\xi)}{\text{length}(L \cap K)} = 1 \right) \geq \frac{1}{2} - O^*\left(\frac{1}{\sqrt{n}} \right).$$

We can repeat the argument by replacing D with

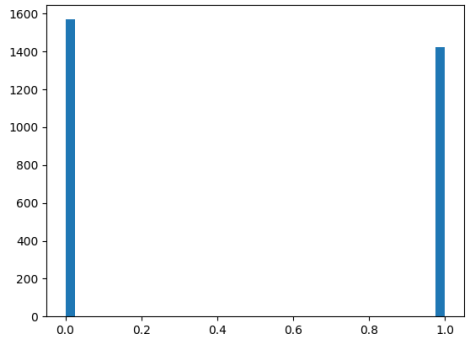
$$D' = \left\{ |\langle X, \xi \rangle| \leq t_\xi - c^{-1} \frac{\log^3 n}{\sqrt{n}} \right\},$$

and obtain

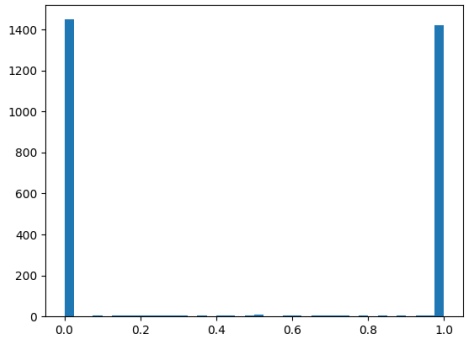
$$\mathbb{P}\left(\frac{\text{length}(L \cap A_\xi)}{\text{length}(L \cap K)} = 0 \right) \geq \frac{1}{2} - O^*\left(\frac{1}{\sqrt{n}} \right).$$

□

Remark 3.23. *There are other examples of subsets with a similar property. For example, similar analysis shows that in an isotropic convex body K with constant thin shell width (e.g the cube or the simplex) the set $A = K \cap (RB_2^n)$, where B_2^n is the euclidean ball and R is chosen such that $Vol(A) / Vol(K) = 1/2$, has a similar property (when replacing $0, 1$ with $o(1), 1 - o(1)$).*



(a) Standard simplex.



(b) Ball

Figure 3.2: Histogram of simulations on the sphere in dimension 1000 with 3000 samples for A_ξ defined on different convex bodies.

In order to complete the proof of Theorem 3.2, we note that we can repeat this argument for any convex body, but in a specific direction ξ. Let $K \subseteq \mathbb{R}^n$ be a convex body. Let X be a random vector distributed uniformly in K. Denote the covariance matrix of X by $C = \mathbb{E}X \otimes X$. Choosing $\xi = \operatorname{argmax}\{\langle Cx, x\rangle\,;\ x \in S^{n-1}\}$, will obtain the desired result.

While Theorem 3.2 shows that for some subsets of the simplex, intersection with random geodesic will have a zero one law, Theorem 3.1 shows that for the simplex there is another curve model that samples subsets of the simplex in a more representative way.

Let $\Delta_n = \{x \in \mathbb{R}^n;\ x_i \geq 0,\ \sum x_i \leq 1\}$ be the Archimedes simplex in $\mathbb{R}^n$. It is well known that the transformation $\pi_n : S^{2n+1} \to \Delta_n$ defined by

$$\pi_n(x) = (x_1^2 + x_{n+2}^2, \ldots, x_n^2 + x_{2n+1}^2)$$

pushes the surface area measure of the sphere to the volume measure of the simplex. Let $x, y \in S^{2n+1}$ be orthogonal to each other, and let $\gamma(t) = x\cos t + y \sin t$, be the geodesic curve that starts at x in the direction y, and $\delta = \pi_n \circ \gamma$. We have

$$\begin{aligned}(\delta(t))_j = ((\pi_n\gamma)(t))_j &= (c_j \cos t + y_j \sin t)^2 + (c_{n+j+1}\cos t + y_{n+j+1}\sin t)^2 \\ &= \frac{1}{2}(x_j^2 + x_{n+j+1}^2 + y_j^2 + y_{n+j+1}^2) + \frac{1}{2}(x_j^2 + x_{n+j+1}^2 - y_j^2 - y_{n+j+1}^2)\cos(2t) \\ &+ (x_j y_j + x_{n+j+1}y_{n+j+1})\sin(2t).\end{aligned}$$

Hence, the map π_n maps geodesics on a sphere to a ellipses in Δ_n. Hence, we can push forward our result to the simplex and get,

Corollary 3.24. *Let $A \subseteq \Delta_n$ be with $Vol(A) / Vol(\Delta_n) = 1/2$. Let L be a random ellipse in Δ_n chosen as the π_n image of a uniform geodesic on S^{2n-1}. We have,*

$$\mathbb{P}\left(\left|\frac{2\mu(A \cap L)}{\mu(L)} - 1\right| \geq \frac{1}{2^{1/3}}\right) \leq \frac{1}{2^{1/3}},$$

where μ is the push forward on the length by the map π_n.

Question 3.25. *Let $K \subseteq \mathbb{R}^n$ be a convex body, can we find a distribution of curves inside K that would generalize Corollary 3.24 to K?*

3.5 Arithmetic Progressions in Finite Fields

In this section we demonstrate how a similar technique to the one we used to prove Theorem 3.1, can be applied in other settings. We analyze the intersection of subsets of the n dimensional discrete torus with random arithmetic progressions.

Let $p > 0$ be a prime number and let $n > 0$ be an integer. Denote $\mathbb{Z}_p = \mathbb{Z}/p\mathbb{Z}$, and

$$G_p = \{(a, a+b, \ldots, a+(p-1)b);\ a, b \in \mathbb{Z}_p^n,\ b \neq 0\}.$$

For convenience we shall denote a sequence $(a, a+b, \ldots, a+(p-1)b)$ by the pair $(a; b)$. We define a discrete version of the Radon transform $R : L^2(\mathbb{Z}_p^n) \to L^2(G_p)$ by

$$(Rf)(a; b) = \frac{1}{p} \sum_{j=0}^{p-1} f(a + jb).$$

The conjugate transform $R^* : L^2(G_p) \to L^2(\mathbb{Z}_p^n)$ is defined by

$$(R^* g)(a) = \frac{1}{p^n - 1} \sum_{b \neq 0} g(a; b).$$

As before, we define $S = R^* R$.

Proposition 3.26. *The eigenvalues of the operator S are*

$$\lambda^2 = \frac{p^{n-1} - 1}{p^n - 1},$$

for functions with mean zero, and

$$\lambda^2 = 1,$$

for constant functions.

Proof. The operator S commutes with translations. Hence, the eigenfunctions of S are of the form $f_y(x) = e^{i2\pi\langle x,y\rangle/p}$ for $y \in \mathbb{Z}_p^n$. We have,

$$Sf_y(a) = \frac{1}{p(p^n-1)} \sum_{b\neq 0} \sum_{j=0}^{p-1} f_y(a+jb).$$

By the definition of f_y we have

$$\sum_{j=0}^{p-1} f_y(a+jb) = \sum_{j=0}^{p-1} e^{i2\pi\langle a+jb,y\rangle} = e^{i2\pi\langle a,y\rangle/p} \sum_{j=0}^{p-1} e^{i2\pi j\langle b,y\rangle/p} = f_y(a) \begin{cases} p, & \langle b,y\rangle = 0 \\ 0, & \langle b,y\rangle \neq 0 \end{cases}.$$

Hence,

$$S(f_y)(a) = f_y(a)\frac{1}{p^n-1}\#\{b;\ \langle b,y\rangle = 0, b\neq 0\} = f_y(a)\frac{1}{p^n-1} \begin{cases} p^n-1,\ y=0 \\ p^{n-1}-1,\ y\neq 0 \end{cases}.$$

We have,

$$\lambda_y^2 = \frac{p^{n-1}-1}{p^n-1}, \quad \forall y \neq 0.$$

□

Using the singular values of the Radon transform, we can bound the variance as before, and get a result analogous to Theorem 3.1.

Theorem 3.27. *Let $A \subseteq \mathbb{Z}^n$ with $m(A) = 1/2$ and let $a, b \in \mathbb{Z}_p^n$ be random vectors such that $b \neq 0$. Let $L = \{a, a+b, \ldots, a+(p-1)b\}$ be the arithmetic progression defined by a and b. Then,*

$$\mathbb{P}\left(\left|\#(A\cap L) - \frac{p}{2}\right| \geq \sqrt{\frac{p}{2}}\right) \leq \frac{1}{2}.$$

3.6 Intersection With Higher Dimensional Subspaces

In § 3.2 we saw how to express the singular values of the k dimensional Radon transform R_k by one dimensional integration of the Gegenbauer polynomials with

respect to some weight function. We used this result in § 3.3 in order to find the behavior of intersection with random geodesics on the sphere through the special case $k = 2$. The purpose of this section is to understand random intersection on the sphere by subspaces of higher dimension k by studying the singular values of R_k for $k > 2$.

It is possible to generalize the approach of Section 3.3 and use the known coefficients of the trigonometric polynomials $P_\ell(\cos t)$ (see [36]) and the Fourier series of $\sin^{k-2}(t)$. By orthogonality of the Fourier basis, this will be a finite sum, and the number of summands in $\lambda_{k,2\ell}$ is $\min\{\ell, k/2-1\}$. If either k or ℓ are small, the calculations are simple, but when both k and ℓ are large, it becomes difficult to handle.

We begin with an example of results of this method. Here, we fix a small k and calculate $\lambda^2_{k,2\ell}$ for all ℓ.

Example 3.28. *The eigenvalues of S_4 are*

$$\lambda^2_{4,\ell} = \binom{\ell+n-3}{\ell}^{-1} \binom{\ell/2+n/2-2}{\ell/2}^2 \frac{2n-8}{(\ell+n-4)(\ell+2)}.$$

We obtain $|\lambda_{4,2}| \leq 1/2$. For large ℓ we have

$$\lambda^2_{4,\ell} \approx C \left(\frac{n}{\ell(n+\ell)}\right)^{3/2}.$$

In addition, using Proposition 3.15 the sequence $\{\lambda^2_{4,2\ell}\}$ is a product of two positive decreasing sequences, hence it is also decreasing. We obtain, that for any $f \in L^2(S^{n-1})$ such that $\int_{S^{n-1}} f(x) d\sigma_{n-1} = 0$,

$$\|R_4 f\|_{L^2(G_{n,2})} \leq \frac{1}{2} \|f\|_{L^2(S^{n-1})}.$$

Similarly we can calculate the eigenvalue $\lambda^2_{k,2}$ for all even $k \geq 2$.

Example 3.29. *For any even $k \geq 4$ we have,*

$$\lambda^2_{k,2} = \tau_k \frac{\pi(n-2)}{2^{k-1}} \binom{k-2}{k/2-1} \frac{n-k}{k} \binom{n-1}{2}^{-1} = \frac{n-k}{k(n-1)}.$$

We employ a different technique in order to calculate the eigenvalues of S_k for all k and ℓ.

Proposition 3.30. *Let $2 \leq k \leq n$ and let $\ell \geq 1$. Then,*

$$\lambda^2_{k,2\ell} = \tau_k \binom{2\ell+n-3}{2\ell}^{-1} \frac{\sqrt{\pi}\Gamma(2\ell+n/2-1)\Gamma(k/2-1/2)}{\Gamma(\ell+1)\Gamma(k/2+\ell)\Gamma(n/2-1)} \frac{(-\ell-n/2+k/2+1)_\ell}{(-2\ell-n/2+2)_\ell},$$

where $(a)_b$ is the Pochhammer symbol.

When $a, b \geq 0$ and $b \in \mathbb{Z}$, we have $(a)_b = \Gamma(a+b)/\Gamma(a)$ and $(-a)_b = (-1)^b\Gamma(a+1)/\Gamma(a-b+1)$.

Proof. By Proposition 3.10, it is enough to calculate $\int_{-1}^1 P_{2\ell}(t)(1-t^2)^{k/2-3/2}dt$. By [36, equation 4.7.31] we have

$$P_{2\ell}(t) = \sum_{j=0}^{\ell} (-1)^j \frac{2^{2\ell-2j}\Gamma(2\ell-j+n/2-1)}{\Gamma(n/2-1)\Gamma(j+1)\Gamma(2\ell-2j+1)} t^{2\ell-2j}.$$

Hence, we can start by integrating only the monomials. By standard computations,

$$\int_{-1}^1 t^{2m}(1-t^2)^{k/2-3/2}dt = \frac{\Gamma(m+1/2)\Gamma(k/2-1/2)}{\Gamma(m+k/2)}.$$

Combining the two, we have

$$\int_{-1}^1 P_{2\ell}(t)(1-t^2)^{k/2-3/2}dt = \frac{\Gamma(k/2-1/2)}{\Gamma(n/2-1)} \sum_{j=0}^{\ell} (-1)^j \frac{2^{2\ell-2j}\Gamma(2\ell-j+n/2-1)\Gamma(\ell-j+1/2)}{\Gamma(j+1)\Gamma(2\ell-2j+1)\Gamma(\ell-j+k/2)}.$$

Multiplying inside the sum and diving outside the sum by

$$\frac{\Gamma(\ell+1)\Gamma(k/2+\ell)}{\sqrt{\pi}\Gamma(2\ell+n/2-1)},$$

we need to sum over

$$(-1)^j \frac{2^{2\ell-2j}\Gamma(2\ell-j+n/2-1)\Gamma(\ell-j+1/2)\Gamma(\ell+1)\Gamma(k/2+\ell)}{\Gamma(j+1)\Gamma(2\ell-2j+1)\Gamma(\ell-j+k/2)\sqrt{\pi}\Gamma(2\ell+n/2-1)}.$$

Since (see [1, equation 6.1.18])

$$\frac{2^{2\ell-2j}\Gamma(\ell-j+1/2)}{\sqrt{\pi}\Gamma(2\ell-2j+1)} = \frac{1}{\Gamma(\ell-j+1)},$$

Using the definition of the Pochhammer symbol for negative numbers, we can reduce the summands to

$$(-1)^j \frac{\Gamma(\ell+1)\Gamma(k/2+\ell)\Gamma(2\ell+n/2-j-1)}{\Gamma(j+1)\Gamma(\ell-j+1)\Gamma(k/2+\ell-j)\Gamma(2\ell+n/2-1)} = (-1)^j \binom{\ell}{j} \frac{(-k/2-\ell+1)_j}{(-2\ell-n/2+2)_j}.$$

By the Vandermonde identity (see [35, Appendix III])

$$\sum_{j=0}^{\ell}(-1)^j\binom{\ell}{j}\frac{(-k/2-\ell+1)_j}{(-2\ell-n/2+2)_j}=\frac{(-\ell-n/2+k/2-1)_\ell}{(-2\ell-n/2+2)_\ell}.$$

Hence,

$$\int_{-1}^{1}P_{2\ell}(t)(1-t^2)^{k/2-3/2}dt=\frac{\sqrt{\pi}\Gamma(2\ell+n/2-1)\Gamma(k/2-1/2)}{\Gamma(\ell+1)\Gamma(k/2+\ell)\Gamma(n/2-1)}\frac{(-\ell-n/2+k/2+1)_\ell}{(-2\ell-n/2+2)_\ell}.$$

□

Corollary 3.31. *Let $2\le k\le n-1$. The sequence $\{\lambda_{k,2\ell}\}^2$ of eigenvalues of S_k is a decreasing sequence.*

Proof. Using the functional relationships $\Gamma(x+1)/\Gamma(x)=x$ and $(x+1)_m/(x)_m=(m+x)/x$, by Proposition 3.30, we have

$$\frac{\lambda_{k,2\ell+2}^2}{\lambda_{k,2\ell}^2}=\frac{(2\ell+1)(2\ell+n-k)}{(2\ell+k)(2\ell+n-1)}. \qquad (\star)$$

We note that for any $k>1$, this ratio is strictly less than one, hence the sequence is decreasing. □

Using the above calculations, we can prove Theorem 3.4.

Proof of Theorem 3.4. By Corollary 3.31 the non trivial singular values of R_k are at most $\lambda_{k,2}$. By Proposition 3.30 we have

$$\lambda_{k,2}^2=\tau_k\binom{n-1}{2}^{-1}\frac{\sqrt{\pi}\Gamma(n/2+1)\Gamma(k/2-1/2)}{\Gamma(k/2+1)\Gamma(n/2-1)}\frac{n-k}{n}.$$

In addition,

$$\tau_k=\frac{\Gamma(k/2)}{\sqrt{\pi}\Gamma(k/2-1/2)}.$$

We have,

$$\lambda_{k,2}^2=\frac{2(n-k)\Gamma(n/2+1)\Gamma(k/2)}{n(n-1)(n-2)\Gamma(k/2+1)\Gamma(n/2-1)}=\frac{n-k}{k(n-1)}.$$

Hence, for any $f\in L^2(S^{n-1})$ we have

$$\mathrm{Var}(R_kf)\le\frac{n-k}{k(n-1)}\mathrm{Var}(f).$$

To finish the proof, we set f to be the normalized indicator of the subset $A\subseteq S^{n-1}$. □

The importance of Corollary 3.31 is not only in showing that the sequence is decreasing, but also by the rate of the decrease in $(\star)$.

Example 3.32. *When $k = n - 1$, Equation $(\star)$ gives us*

$$\lambda_{n-1,2\ell} \leq \left(C \frac{\ell}{n+\ell} \right)^{\ell}.$$

This is the rate used by Klartag and Regev in [22] to prove the result for intersection with hyper-planes.

Another interesting case is when $k = \lfloor n/2 \rfloor$. Equation $(\star)$ gives us

$$\lambda_{n/2,2\ell} \leq \left(C \frac{\ell}{n+\ell} \right)^{\ell/2}.$$

An interesting question would be to give a direct proof for Theorem 6.1 in [22] using this result. A possible starting point would be to generalize the hypercontractivity result (Lemma 5.3 in [22]) for the Grassmanian manifold $G_{n,k}$.

Chapter 4

Scenery Reconstruction in the Hypercube

4.1 Introduction

Let $f : \{-1, 1\}^n \to \{-1, 1\}$ be a Boolean function on the n-dimensional hypercube, and let S_i be a random walk on the hypercube. Can we reconstruct the function f (with probability 1, up to the hypercube's symmetries) by only observing the scenery process $\{f(S_i)\}_i$?

Similar questions have been raised for other graphs. For example, it was shown in [5] that when G is a cycle graph, the answer is yes: it is possible to reconstruct the function f (which is a string up to choice of origin) up to rotation and reflection with probability 1. It is still an open question whether any such string can be reconstructed in polynomial time. When $G = \mathbb{Z}$, reconstruction is generally impossible [25]; for random sceneries on $\mathbb{Z}$ see [29].

When G is the hypercube, such a process was studied for a specific Boolean function, the *percolation* crossing, under the notion of dynamical percolation; see [13] for details.

In the general case, however, we show that for $n \geq 4$ the answer is no. We do this by considering a pair of non-isomorphic functions f and g such that if S_i and T_i are random walks on the hypercube, then $f(S_i)$ and $g(T_i)$ have exactly the same distribution. We discuss two different classes of such functions:

- **Locally p-biased functions:** Let G be a graph. A Boolean function $f : G \to \{-1, 1\}$ is called *locally p-biased*, if for every vertex $x \in G$ we have

$$\frac{|\{y \sim x;\ f(y) = 1\}|}{deg(x)} = p.$$

  In words, f is locally p-biased if for every vertex x, f takes the value 1 on exactly a p-fraction of x's neighbors. If f is a locally p-biased function, then the random variables $\{f(S_i)\}_i$ have the same distribution as independent Bernoulli random variables with $\mathbb{P}(f(S_i) = 1) = p$.

- **Locally p-stable functions:** Let G be a graph. A Boolean function $f : G \to \{-1, 1\}$ is called *locally p-stable*, if for every vertex $x \in G$ we have

$$\frac{|\{y \sim x;\ f(x) = f(y)\}|}{deg(x)} = p.$$

  In words, f is locally p-stable if for every vertex x, f retains its value on exactly a p-fraction of x's neighbors. If f is locally p-stable, then the random variables $\{f(S_i)f(S_{i+1})\}_i$ have the same distribution as independent Bernoulli random variables with $\mathbb{P}(f(S_i)f(S_{i+1}) = 1) = p$.

We say that two Boolean functions $f, g : \{-1, 1\}^n \to \{-1, 1\}$ are *isomorphic*, if there exists an automorphism of the hypercube $\psi : \{-1, 1\}^n \to \{-1, 1\}^n$ such that $f \circ \psi = g$. Two functions are *non-isomorphic* if no such ψ exists.

The existence of two non-isomorphic locally p-biased functions, or two non-isomorphic locally p-stable functions thus render scenery reconstruction on the hypercube impossible.

It is not immediately obvious that pairs of non-isomorphic locally p-biased and pairs of non-isomorphic locally p-stable functions exist. It is then natural to ask, for which p values do they exist? If they do exist, how many of them are there?

In this paper, we characterize the possible p values on the n-dimensional hypercube, give bounds on the number of non-isomorphic pairs, and discuss results on other graphs. The paper is organized as follows.

In § 4.2 we give a full characterization of the connection between the dimension of the hypercube n and the permissible p values of locally p-biased functions, as expressed in the following theorem:

Theorem 4.1. *Let $n \in \mathbb{N}$ be a natural number and $p \in [0,1]$. There exists a locally p-biased function $f : \{-1,1\}^n \to \{-1,1\}$ if and only if $p = b/2^k$ for some integers $b \geq 0, k \geq 0$, and 2^k divides n.*

Our construction is based on the Hamming code. In this code, some of the bits of an n-bit word are designated as parity bits, and their value is a linear combination of the rest of the word such that all the words in the code have a minimum distance of 3. Combining different translations of Hamming codes, we can construct functions for all p of the above form.

In § 4.3 we inspect the class size of non-isomorphic locally p-biased functions on the hypercube. We show that the class size for $p = 1/2$ is at least $C2^{\sqrt{n}}/n^{1/4}$ for some constant $C > 0$, and for $p = 1/n$ is super-exponential in n, when such p values are permissible. Thus reconstruction is impossible for such functions. We conjecture that the number of non-isomorphic locally p-biased functions scales quickly for all permissible p values:

Conjecture 4.2. *Let $n > 0$ be even. Let $p = b/2^k$, where $1 \leq b \leq 2^k$, $k \geq 1$ and 2^k divides n. Let B_p^n be the set of non-isomorphic locally p-biased functions. Then $\left|B_p^n\right|$ is super-exponential in n.*

In § 4.4 we briefly discuss locally p-stable functions. We show that they exist for all possible p values, and that for most p values there are many non-isomorphic pairs; however, for every n, there are p values for which there is a single unique locally p-stable function. The results in this section are based on those of § 4.3.

In § 4.5 we discuss locally p-biased functions on other graphs. First, we show that when G is a regular tree of degree n, then all $p = a/n$ are permissible. Second, we show that for $G = \mathbb{Z}^n$ all the results for the hypercubes hold true. This gives us a partial answer for permissible p values for $\mathbb{Z}^n$, but there are additional values that cannot be achieved through the hypercube construction: for example, for $n = 1$ we can define a function with $p = 1/2$ and when $n = 2$ we can find a function with $p = 1/4$. We also discuss other Cayley graphs of $\mathbb{Z}$, and suggest further questions on scenerey reconstruction.

Throughout most of this paper we treat the Boolean hypercube as the set $\{-1,1\}^n$. We identify it with the $\{0,1\}^n$ hypercube by considering -1 in the first to

correspond to 0 in the second. In this context, for two functions $f, g : \{-1,1\}^n \to \{-1,1\}$, the support of f is the set $\{x \in \{-1,1\}^n ; f(x) = 1\}$, and the union of f and g is the function supported on the union of the supports of f and g.

4.2 Characterization of permissible p values for locally p-biased functions

In this section we prove Theorem 4.1. The "only if" part is achieved by a double counting argument.

Proof (of the "only if" statement of Theorem 4.1). Let $f : \{-1,1\}^n \to \{-1,1\}$ be a locally p-biased Boolean function. For a given vertex, p represents both the fraction of neighbors on which f obtains the value 1, and also the fraction of vertices of the entire graph on which f obtains the value 1. Thus $p = m/n$ for some $m \in \{0, 1, \ldots, n\}$, and also $p = l/2^n$ where $l = |\{x \in \{-1,1\}^n ;\ f(x) = 1\}|$.

$$p = \frac{l}{2^n} = \frac{m}{n}. \tag{4.1}$$

Decompose n into its prime powers, writing $n = c2^k$, where c is odd. Then by (4.1), we have that

$$l = \frac{2^{n-k} \cdot m}{c}$$

is an integer, and so c must divide m, i.e $m = bc$ for some b. But then

$$p = \frac{m}{n} = \frac{b}{2^k}$$

as stated by the theorem. □

The "if" part of Theorem 4.1 is given by an explicit construction, performed in three steps. First, we use perfect codes (introduced below) in order to obtain a locally $1/n$-biased function for n that is a power of two. Second, we extend the result to a locally m/n-biased function by taking the union of m locally $1/n$-biased functions with disjoint support. Finally, given a locally p-biased function on n bits,

we show how to manipulate its Fourier representation in order to yield a locally p-biased function on cn bits for any c.

We begin with a brief review of binary codes. We omit proofs and simply state definitions and known results; for a more thorough introduction, see e.g [26, 27].

A binary code C on the n-dimensional hypercube is simply a subset of $\{-1,1\}^n$; its elements are called *codewords*. The *distance* of a code C is defined as $\min_{x \neq y \in C} \delta_H(x,y)$, where $\delta_H(x,y) = |\{i \in \{1,\ldots,n\}\,; x_i \neq y_i\}|$ is the Hamming distance between x and y, that is, the number of coordinates in which x and y differ. A code of odd distance d is called *perfect* if the Hamming balls of radius $(d-1)/2$ around each codeword completely tile the hypercube without overlaps. A code is called *linear* if its codewords form a vector space over $\mathbb{F}_2$.

A particularly interesting code is the Hamming code with k parity bits, denoted H_k. It is a linear, distance-3 perfect code on the hypercube of dimension $n = 2^k - 1$. Its codewords are structured as follows. For $x \in H_k$ and $i \in \{1,\ldots,n\}$, the bit x_i is called a *parity bit* if i is a power of 2, and *data bit* otherwise. Thus every codeword contains k parity bits and $2^k - k - 1$ data bits. The data bits range over all possible bit-strings on $2^k - k - 1$ bits, while the parity bits are a function of the data bits:

$$x_i = \bigoplus_{j : i \wedge j \neq 0} x_j \qquad \forall i = 2^l, l \geq 0$$

where $\oplus$ denotes exclusive bitwise or (xor), and $\wedge$ denotes bitwise AND. Thus there are $2^k - k - 1$ codewords in H_k.

Armed with perfect codes, we are ready to start our proof.

Lemma 4.3. *Let $n = 2^k$ be a power of two. Then there exists a locally $1/n$-biased function on $\{-1,1\}^n$.*

Proof. In a locally $1/n$-biased function f, every point in the hypercube must have exactly 1 neighbor which is given the value 1, and $n-1$ neighbors which are given the value -1.

Let C be a distance-3 perfect code on the $n-1 = 2^k - 1$-dimensional hypercube. That is, every two code words in C are at a Hamming distance of at least 3 from each other, and the Hamming balls of radius 1 centered around each codeword completely tile the hypercube. Such codes exist for dimension $2^k - 1$; for example,

as mentioned above and shown in [26], the Hamming code is such a code. Define $f : \{-1,1\}^n \to \{-1,1\}$ to be the following function:

$$f(x) = \begin{cases} 1, & x \in C \times \{-1,1\} \\ -1, & \text{otherwise.} \end{cases} \tag{4.2}$$

In words, $f(x)$ takes the value of 1 whenever the first $n-1$ coordinates of x are a codeword in C, and otherwise takes the value of -1. Then f is a locally $1/n$-biased function:

- If $f(x) = 1$, then $x = (y, b) \in C \times \{-1,1\}$. Thus $x' = (y, -b)$ is the only neighbor of x on which $f(x') = 1$; any other neighbor differs from x in the first $n-1$ coordinates, and since C is a distance-3 code, these coordinates are not a codeword in C.

- If $f(x) = -1$, then $x = (y, b)$ where $b \in \{-1,1\}$ and y is not a codeword of C. Since C is perfect, y must fall inside some radius-1 ball of a codeword z. Then $x' = (z, b)$ is the only neighbor of x such that $f(x') = 1$; any other codeword differs from z in at least 3 coordinates since C is a distance-3 code, and so differs from y in at least 2.

□

Lemma 4.4. *Let $n = 2^k$ be a power of two. Then there exists a locally m/n-biased function on $\{-1,1\}^n$ for any $m = 0, 1, \ldots, n$.*

Proof. For $m = 0$ the statement is trivial. Let $m \in \{1, \ldots, n\}$. In order to construct a locally m/n-biased function, it is enough to find m locally $1/n$-biased functions $f_1, \ldots, f_m$ with pairwise disjoint support, i.e $\{x : f_i(x) = 1\} \cap \{x : f_j(x) = 1\} = \emptyset$ for all $i \neq j$. With these functions, we can define f in the following manner:

$$f(x) = \begin{cases} 1, & f_i(x) = 1 \text{ for some } i \\ -1, & \text{otherwise.} \end{cases}$$

Then f is a locally m/n-biased function: For every $x \in \{1,-1\}^n$, consider its neighbors on which f takes the value 1, i.e the set $\{y;\ d(x,y) = 1, \exists i\ s.t\ f_i(y) = 1\}$.

Each f_i contributes exactly one element to this set, since it is a locally $1/n$-biased function; further, these elements are all distinct, since the f_i's have pairwise disjoint supports. So x has m neighbors on which f takes the value 1.

Recall that the Hamming code on 2^k-1 bits uses 2^k-k-1 data bits (these range over all possible bit-strings on 2^k-k-1 bits) and k parity bits (these are a function of the data bits). Let C be the Hamming code on 2^k-1 bits, and rearrange the order of the bits so that the parity bits are all on the right hand side of the codeword, i.e each codeword x can be written as $x=(y,z)$ where y is a word of length 2^k-k-1 constituting the data bits and z is a word of length k constituting the parity bits.

Now, for all $1 \leq i \leq n$, define the sets $C_i = \{x \oplus (i-1);\ x \in C\}$, where $\oplus$ denotes the exclusive or (xor) operator. Then the sets C_i are all pairwise disjoint: in order for two words $x=(y,z) \in C_i$ and $x'=(y',z') \in C_j$ to be the same, we need to have both $y=y'$ and $z=z'$. But if $y=y'$ then the data bits are the same, and by construction $z \oplus z' = (i-1) \oplus (j-1)$, so $z \neq z'$ if $i \neq j$. Further, since xoring by a constant only amounts to a rotation of the hypercube, each C_i is still a perfect code.

Let f_i be the function which uses C_i as its perfect code as defined in (4.2). Then, $f_1, \ldots, f_n$ are n locally $1/n$-biased functions with pairwise disjoint supports. The combination of any m of these functions yields a locally m/n-biased function.

□

Lemma 4.5. *Let $f : \{-1,1\}^n \to \{-1,1\}$ be a locally p-biased function on the n-dimensional hypercube. Let $c \in \mathbb{N}$, and define a new function $f' : \{-1,1\}^{cn} \to \{-1,1\}$ by*

$$f'(x) \;=\; f\left(\prod_{j=0}^{c-1} x_{1+jn}, \ldots, \prod_{j=0}^{c-1} x_{n+jn}\right). \tag{4.3}$$

Then f' is a locally p-biased function.

Proof. Let $x' \in \{-1,1\}^{cn}$ and let $i \in \{1,2,\ldots,n\}$. The number of neighbors of x' that change the sign of the $i-$th coordinate of $\left(\prod_{j=0}^{c-1} x_{1+jn}, \ldots, \prod_{j=0}^{c-1} x_{n+jn}\right)$ is exactly c. Since the coordinates are independent and f is locally p-biased, the fraction of neighbors of x' where f' obtain the value 1 is pcn. □

We are now ready to prove that the condition on p is sufficient in Theorem 4.1.

Proof (of the "if" statement of Theorem 4.1). All that is left is to stitch the above lemmas together: Let $n = c2^k$. Using Lemma 4.4, create a locally p-biased function $g : \{-1,1\}^{2^k} \to \{-1,1\}$ on 2^k variables; then, using Lemma 4.5, extend it to a function f on n variables. □

4.3 Non-isomorphic locally p-biased functions

In this section we discuss the classes of non-isomorphic locally p-biased function. We show that for the hypercube of dimension n, the growth rate with respect to n is at least $\Omega\left(2^{\sqrt{n}}/\sqrt{n}\right)$ for $p = 1/2$ and super-exponential for $p = 1/n$, when such p's are permissible. We conjecture that for any permissible p the growth rate is super-polynomial.

The proof for $p = 1/2$ is based on an explicit construction of non-isomorphic locally 1/2-biased functions. In order to define these functions we use the following simple observation.

Observation 4.6. *Let $f_i : \{-1,1\}^{n_i} \to \{-1,1\}$ be locally 1/2-biased functions for $i = 1,2$ where $n_1 + n_2 = n$. Then*

$$f(x) = f_1(x_1, ..., x_{n_1}) f_2(x_{n_1+1}, ..., x_n)$$

is a locally 1/2-biased function on $\{-1,1\}^n$.

The above proposition allows us to construct examples for locally 1/2-biased functions, by combinations of such functions on lower dimensions.

We have two basic examples for locally 1/2-biased functions:

1. In any even dimension n,

$$g_n(x_1, \ldots, x_n) = x_1 \cdots x_{n/2}.$$

2. In dimension $n = 4$,

$$h(x_1, x_2, x_3, x_4) = \frac{1}{2}\left(x_1x_2 + x_2x_3 - x_3x_4 + x_1x_4\right).$$

The Fourier decomposition of a Boolean function is its expansion as a real multilinear polynomial: any Boolean function $f : \{-1,1\}^n \to \mathbb{R}$ can be written as a sum

$$f(x_1,\ldots,x_n) = \sum_{S\subseteq\{1,\ldots,n\}} \hat{f}_S \prod_{i\in S} x_i,$$

where the $\hat{f}_S$ are real coefficients. Such a representation is unique; for a proof and other properties of the Fourier decomposition, see e.g chapter 1 in [31].

Automorphisms of the hypercube are manifested on the Fourier decomposition of a Boolean function by either permutation or by a sign change to a subset of indices. Hence, we can show that two Boolean functions are not isomorphic by showing that their Fourier decompositions cannot be mapped into one another by such permutations and sign changes.

In this section, a tensor product of two functions $f(x_1,\ldots,x_n)$ and $g(x_1,\ldots,x_m)$ is a function on disjoint indices, i.e.

$$h(x_1,\ldots,x_{n+m}) = f(x_1,\ldots,x_n)\cdot g(x_{n+1},\ldots,x_{n+m}).$$

Proposition 4.7. *There exists $h_1, h_2, \ldots$ such that for any k the function h_k is locally $1/2$-biased on the $4k$-dimensional hypercube and h_k is not isomorphic to any tensor product of $h_1,\ldots,h_{k-1},g_2,g_4,g_6,....$*

Proof. We define $h_1 = h$, and

$$h_k = h\left(\prod_{i=0}^{k-1} x_{1+4i},\ldots,\prod_{i=0}^{k-1} x_{4+4i}\right),$$

where h is the function from example 2. By Lemma 4.5, h_k is locally $1/2$-biased on $\{-1,1\}^{4k}$. Assume that h_k is isomorphic to a tensor product of $h_1,\ldots,h_{k-1},g_2,...,g_{n-2}$. If there exists $1 \le i \le j < k$ such that both h_i and h_j appear in a product that is isomorphic to h_k, then the Fourier decomposition of the product would have at least 16 different monomials. But h_k has only 4 different monomials, and the functions cannot be isomorphic. Similarly, if we do not use any of the functions $h_1,\ldots,h_{k-1}$, then we get the parity function, which has only one monomial in its Fourier decomposition. Hence, we may assume that there is only one $1 \le i < k$ such that h_i is in the product. Then, up to an automorphism, this function is of the form

$$f(x) = h_i(x_1,\ldots,x_{4i})g_{4k-4i}(x_{4i+1},\ldots,x_{4k}).$$

On the one hand, by definition of h_k, its Fourier decomposition has pairs of monomials with no shared indices (e.g. the monomials that replace x_1x_2 and x_3x_4 in h_1). On the other hand, in the decomposition of f, all monomials have shared indices; for example x_{4i+1} appears in all monomials. Hence they are not isomorphic. □

Using the functions $h_1, h_2, \ldots$ we can give a lower bound for the class of non-isomorphic locally 1/2-biased functions.

Lemma 4.8. *The number of non-negative integer solutions to*

$$a_1 + 2a_2 + \ldots + ka_k \leq k \tag{4.4}$$

is at least $C4^{\sqrt{k}}/k^{1/4}$, where $C > 0$ is a universal constant.

Proof. For any $1 \leq \ell \leq k$, the number of solutions to (4.4) is at least the number of solutions to

$$\ell a_1 + \ell a_2 + \ldots + \ell a_\ell \leq k.$$

It is well known that the number of solutions to this inequality is

$$\binom{\ell + k/\ell}{\ell}.$$

This term is maximized when $\ell^2 = k$. Hence, a lower bound for the number of solutions to (4.4) is

$$\binom{2\sqrt{k}}{\sqrt{k}}.$$

By Stirling's formula, the asymptotic of this is $(1/\sqrt{\pi})4^{\sqrt{k}}/k^{1/4}$. □

Remark 4.9. *The number of integer solutions to the equality case is the famous partition function $p(n)$. Hardy and Ramanujan [17] showed precise asymptotics. Using their result it is possible to show that the number of integer solutions is*

$$\sum_{j=1}^{k} p(j) \sim Ce^{c\sqrt{k}}/\sqrt{k},$$

with explicit constants $C, c > 0$. While this result gives better bounds than Lemma 4.8, our simple estimation is enough for our purposes.

Proposition 4.10. *Let n be even. Let $B^n_{1/2}$ be a maximal class of non-isomorphic locally 1/2-biased functions. Then $\left|B^n_{1/2}\right| \geq C2^{\sqrt{n}}/n^{1/4}$, where $C > 0$ is a universal constant.*

Proof. Let $k = \lfloor n/4 \rfloor$. By Observation 4.6, we can construct locally 1/2-biased functions by tensor products of $h_1, \ldots, h_k$ and $g_1, \ldots, g_n$, as follows: choose functions $\{h_{i_j}\}$ such that $m := \sum 4i_j \leq n$. Then the tensor product $\otimes h_{i_j}$ uses m variables. This can be completed to n variables by tensoring with g_{n-m}.

If two functions use the same h_i's, then they are isomorphic (by change of indices). And if they have a different decomposition of h_i's, then by the same arguments used in Proposition 4.7, they have a different Fourier decomposition and are therefore non-isomorphic. Thus, the isomorphic class of such a function is determined by the number of times each h_i appears in the product.

Hence, the number of non-isomorphic functions we can construct in this manner is the number of solutions to

$$4a_1 + 8a_2 + \cdots + 4ka_k \leq n \tag{4.5}$$

where the $a_1, \ldots, a_k$ are non-negative integers that represent the number of copies of h_i in the product. Using Lemma 4.8, this number is at least $C4^{\sqrt{k}}/k^{1/4} = C'2^{\sqrt{n}}/n^{1/4}$

□

For a Boolean function f with Fourier coefficients $\hat{f}_S$, the Fourier weight at degree d is defined as

$$W_d(f) = \sum_{|S|=d} \hat{f}_S^2.$$

As the following proposition shows, the Fourier decomposition of a locally 1/2-biased function contains only monomials of degree $n/2$. It might be possible to obtain better bounds on the number of non-isomorphic functions using this condition.

Proposition 4.11. *Let $f : \{-1,1\}^n \to \{-1,1\}$ be a locally 1/2-biased function. Then the Fourier weight at degree $n/2$ is 1.*

It is an interesting question to find a general connection between p and the weight distribution of a locally p-biased function.

Proof. Let A_n be the adjacency matrix of the hypercube. The map

$$\varphi : f \mapsto (f(a_1), ..., f(a_{2^n})),$$

where $a_1, ..., a_{2^n}$ are the vertices of the hypercube, is a bijection between locally 1/2-biased functions and the null space of A_n. Since

$$A_n = \begin{pmatrix} A_{n-1} & I \\ I & A_{n-1} \end{pmatrix},$$

we have

$$P_n(t) = P_{n-1}(t-1)P_{n-1}(t+1),$$

where P_n is the characteristic polynomial of A_n. For A_2 the eigenvalue 0 has multiplicity 2 and ± 2 has 1. Continuing by induction, the eigenvalues of A_m are $-m, -m+2, ..., m$ with multiplicities $\binom{m}{0}, \binom{m}{2}, ..., \binom{m}{m}$. Hence, for even n the dimension of the null space is $\binom{n}{n/2}$. For any $S \subseteq \{1, 2, ..., n\}$ with $|S| = n/2$ we denote $\chi_S(x) = \prod_{i \in S} x_i$ and $v_S = \varphi(\chi_S)$. Since the functions χ_S are all locally 1/2-biased, the vectors v_S are in the null space of A_n. Note that there are $\binom{n}{n/2}$ such vectors, and they form an independent set. Hence the set $\{v_S\}_S$ is a basis of the null set. By the bijection we get that every locally 1/2-biased function is a linear combination of χ_S. □

Class sizes for locally $1/n$-biased functions can also be achieved via the following proposition.

Proposition 4.12. *Let $n = 2^k$, and let C_1 and C_2 be two non-isomorphic distance-3 perfect codes on the $n-1$-dimensional hypercube. Then the two functions f_1 and f_2 defined by equation (4.2) using the perfect codes C_1 and C_2 are non-isomorphic.*

The proof shows that in any isomorphism between two functions constructed using equation (4.2), the last coordinate must be preserved. However, this will imply that the remaining coordinates are isomorphic, in contradiction to the assumption.

Proof. Suppose to the contrary that f_1 and f_2 are isomorphic, i.e there is an automorphism $\varphi : \{-1, 1\}^n \to \{-1, 1\}^n$ such that for all $x \in \{-1, 1\}^n$, we have $f_1(x) = f_2(\varphi(x))$. Denote by $B = \left\{(y, 1);\ y \in \{-1, 1\}^{n-1}\right\}$ the $n-1$-dimensional hypercube obtained by fixing the last coordinate to 1, denote $C = \{(y, 1);\ y \in C_2\}$ and note that $\text{support}(f_2|_B) = C$ by construction. Consider $\varphi|_B$, the restriction of φ to B. This restriction is an isomorphism between B and some $n-1$-dimensional

hypercube A contained within the n- dimensional hypercube. Any sub-hypercube of dimension $n-1$ is obtained from $\{-1,1\}^n$ by fixing one of the coordinates to be either 1 or -1, and taking the span of all other coordinates. Then A must be spanned by the first $n-1$ coordinates, leaving the last coordinate fixed: otherwise, by equation (4.2), the set A would contain two neighboring points x and x' that differ only in their last coordinate such that $f_1(x) = f_1(x') = 1$. This means there are $y, y' \in C$ obeying $\varphi(y) = x, \varphi(y') = x'$; but this is a contradiction, since φ should preserve distances, and the distance between x and x' is 1 while the distance between y and y' is 3. So $A = \left\{(y,b);\ y \in \{-1,1\}^{n-1}\right\}$ for some $b \in \{-1,1\}$. But then $\varphi|_B$ is an isomorphism between C_1 and C_2, since C_1 is a perfect code in A and C_2 is a perfect code in B; a contradiction. □

Corollary 4.13. *Let $n = 2^k$. Let $B^n_{1/n}$ be the class of non-isomorphic locally $1/n$-biased functions. Then $\left|B^n_{1/n}\right|$ is super-exponential in n.*

Proof. By Proposition 4.12, any lower bound on the number of non-isomorphic perfect codes on the $n-1$-dimensional hypercube gives a lower bound to the number of locally $1/n$-biased functions on the n-dimensional hypercube. Recent constructions, such as in [24], give a super-exponential lower bound on the number of such perfect codes. □

We would have liked to apply the same argument to locally m/n-biased functions, as given by the construction in Lemma 4.4. Our argument there used the explicit construction of the Hamming code which, being linear, was easy to modify in order to obtain functions with disjoint supports. Such is not the case for the construction of non-linear codes. However, we still believe that similar estimates are true for any permissible p.

Corollary 4.14. *By Proposition 4.10, scenery reconstruction is impossible for even-dimensional hypercubes.*

For odd dimensional hypercubes, on which there are no non-trivial locally biased functions, we use locally stable functions instead, as described in the next section.

4.4 Locally p-stable functions

Unlike locally p-biased functions, there is no restriction on permissible p values for locally p-stable functions:

Observation 4.15. *Let $p = m/n$ for some $m \in \{0, 1, \ldots, n\}$. Then the parity function on $n - m$ variables,*

$$f(x_1, \ldots, x_n) = x_{m+1} x_{m+2} \ldots x_n$$

is locally p-stable.

Thus we will focus on the number of non-isomorphic pairs of locally stable functions. A negative result is attainable by a simple examination:

Proposition 4.16. *If $p = 1/n$ or $p = (n-1)/n$, then the parity function is the only locally stable p-function on the hypercube, up to isomorphisms.*

Proof. We prove only for $p = (n-1)/n$; the proof for $p = 1/n$ is similar.

We will show that f depends only on a single coordinate. Let x be an initial point in the hypercube and y its unique neighbor such that $f(x) \neq f(y)$. Denote the coordinate in which they differ by i. By local stability, every other neighbor x' of x has $f(x') = f(x)$, and every other neighbor y' of y has $f(y') = f(y)$.

Let $j \neq i$, let $\tilde{x}$ be the neighbor of x that differs from x in coordinate j, and let $\tilde{y}$ be the neighbor of y that differs y in coordinate j. Then $\tilde{x}$ is a neighbor of $\tilde{y}$, since $\tilde{x}$ and $\tilde{y}$ differ only in the i-th coordinate. Also, since $f(x) = f(\tilde{x})$ and $f(y) = f(\tilde{y})$ but $f(x) \neq f(y)$, we have $f(\tilde{x}) \neq f(\tilde{y})$.

Since f is locally $(n-1)/n$-stable, each of x's neighbors x' has exactly one neighbor y' on which f attains the opposite value. By the above, for each such x', the corresponding y' differs from it in the i-th coordinate. This reasoning can be repeated, choosing a neighbor of x as the initial starting point, showing that for all x' with the same i-th coordinate as x, $f(x) = f(x')$, while for all x' that differ in the i-th coordinate from x, $f(x) \neq f(x')$. This means that either $f(x) = x_i$ or $f(x) = -x_i$. □

Many other p values, however, have larger classes of non-isomorphic locally p-stable functions, since locally stable functions can be built out of locally 1/2-biased functions:

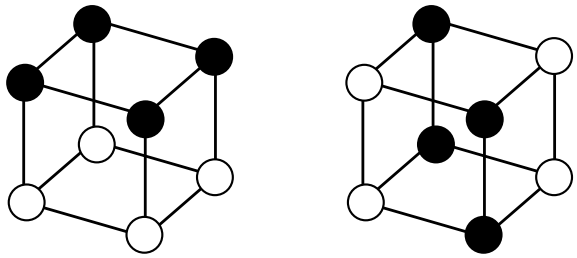


Figure 4.1: Left: The only locally $(n-1)/n$-stable function is the parity function on 1 variable. Right: The only locally $1/n$-stable function is the parity function on $n-1$ variables.

Proposition 4.17. *There is an injection φ between locally $1/2$-biased functions on $\{-1,1\}^n$ and locally $(n/2)/(n+1)$-stable functions on $\{-1,1\}^{n+1}$. Further, if f and g are two non-isomorphic locally $1/2$-biased functions, then $\varphi(f)$ and $\varphi(g)$ are also non-isomorphic.*

Proof. Define φ by

$$(\varphi(f))(x_1,\ldots,x_{n+1}) = f(x_1,\ldots,x_n)\cdot x_{n+1}.$$

Then $\varphi(f)$ is locally $(n/2)/(n+1)$-stable, since for every $x \in \{-1,1\}^{n+1}$, $\varphi(f)$ retains its value on exactly half of the neighbors which differ in the first n coordinates, but flips its value on the neighbor that differs in the last coordinate. The claim about non-isomorphism follows directly from the functions' Fourier decomposition. □

Observe that unlike locally biased functions, locally stable functions can be easily extended to higher dimension:

Observation 4.18. *Let f be a locally $(n-m)/n$-stable function. Then f can be extended to hypercubes of size $n' \geq n$ by simply ignoring all but the first n coordinates. This gives a locally $(n'-m)/n'$-stable function.*

We can use this observation to give a lower bound on the number of locally $(n'-m)/n'$-stable functions for a fixed m and any $n' \geq 2m-2$. This works as follows: first, pick any fixed $m > 1$. Using Proposition 4.17, we obtain a locally $(n-m)/n = (n/2)/(n+1)$-stable with $n = 2m-2$. This can be extended by Observation 4.18 to any $n' \geq n$, and together with Proposition 4.10 we get a lower bound of $C2^{\sqrt{2m-2}}/(2m-2)^{1/4}$ different locally $(n'-m)/n'$-stable functions.

This observation also provides us with a pair of non-isomorphic locally stable functions for all hypercubes of dimension $n \geq 5$, showing that:

Corollary 4.19. *Scenery reconstruction is impossible for n-dimensional hypercubes for* $n \geq 5$.

4.5 Other directions and open questions

In this section we discuss similar results and questions for other graphs. We also list some further questions regarding locally biased and locally stable functions on the hypercube. For other excellent open problems see [10].

4.5.1 Hypercube reconstruction

Our work shows that in general, Boolean functions on the hypercube cannot be reconstructed.

Question 4.20. *Under which conditions is it possible to reconstruct Boolean functions on the hypercube?*

Question 4.21. *Is a random Boolean function reconstructible with high probability?*

Remark 4.22. *Using the techniques of [5], it can be shown that reconstruction is always possible in the hypercube of dimension at most* 3.

4.5.2 Other graphs

Note that the necessity condition on p of Theorem 4.1 can be applied to any finite regular graph, ruling out functions based on the relation between the graph degree and the number of vertices.

Trees

Let G be an n-regular infinite tree. Then for any $p = b/n$, $b = 0, 1, \ldots, n$ there exists a locally p-biased function. Such a function can be found greedily by picking

a root vertex $v \in G$, setting $f(v) = 1$, and iteratively assigning values to vertices further away in any way that meets the constraints.

Notice that the method above requires picking some initial vertex, and that the method yields many possible functions on labeled trees (all of which are isomorphic when we remove the labels). Once the initial vertex v has been fixed, it is possible to generate a distribution on locally p-biased functions, by setting $f(v)$ to be 1 with probability b/n, and randomly expanding from there.

Question 4.23. *For an n-regular tree G, find an invariant probability measure on locally p-biased functions that commutes with the automorphisms of the tree.*

The standard lattice

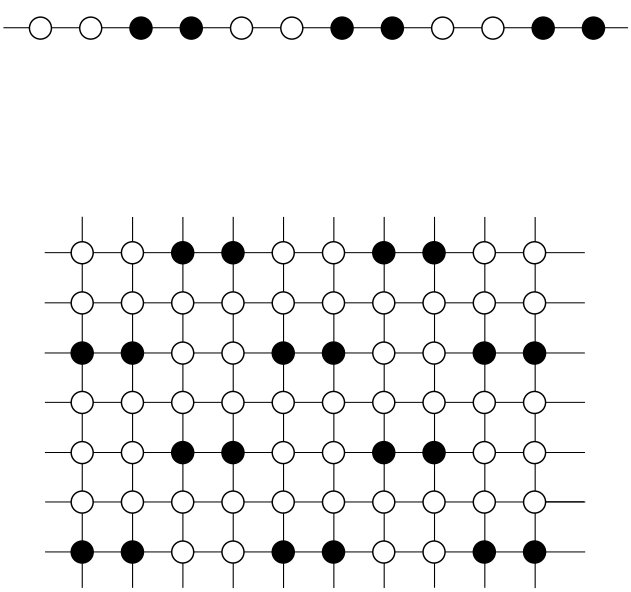


Figure 4.2: Top: a locally 1/2-biased function on $\mathbb{Z}$. Bottom: a locally 1/4-biased function on $\mathbb{Z}^2$.

The following propositions show that there is a one-to-one mapping of locally p-biased functions from the hypercube to $\mathbb{Z}^n$. Since automorphisms of the lattice can be pulled back to automorphisms of the hypercube, we get lower bounds for the size of non-isomorphic locally p-biased functions on $\mathbb{Z}^n$.

Proposition 4.24. *Let $f : \{-1,1\}^n \to \{-1,1\}$ be a locally p-biased function. Then there exists a locally p-biased extension $\tilde{f} : \mathbb{Z}^n \to \{-1,1\}$ such that $f\big|_{\{-1,1\}^n} = f$. In addition, if f and g are non-isomorphic locally p-biased functions on the hypercube, then $\tilde{f}$ and $\tilde{g}$ are non-isomorphic.*

Proof. Here we think of the hypercube as $\{0,1\}^n$ instead of $\{-1,1\}^n$. For any $x \in \mathbb{Z}^n$, define

$$\psi(x_1, \ldots, x_n) = (x_1 \bmod 2, \ldots, x_n \bmod 2)$$

and

$$\widetilde{f}(x) = f(\psi(x)).$$

Let $x \in \mathbb{Z}^n$, and let e_i be the i-th vector in the standard basis. Denote $y = \psi(x)$ and write y^i for the neighbor of y in $\{0,1\}^n$ which differs from y in the i-th direction. Then $\widetilde{f}(x + e_i) = \widetilde{f}(x - e_i) = f(y^i)$, showing that $\widetilde{f}$ is a locally p-biased function. Note that the automorphisms of $\mathbb{Z}^n$ are those of the hypercube with the addition of translations. But under the map ψ, translations in $\mathbb{Z}^n$ amount to reflections in $\{0,1\}^n$. Thus any automorphism between $\widetilde{f}$ and $\widetilde{g}$ would induce one between f and g. □

The above extension procedure gives us lower bounds on the growth rate of some classes of non-isomorphic locally p-biased functions.

Corollary 4.25. *Let $\widetilde{B}_p^n$ be the class of non-isomorphic locally p-biased functions on $\mathbb{Z}^n$.*

1. *If n is even, then $\left|\widetilde{B}_{1/2}^n\right| \geq C2^{\sqrt{n}}/n^{1/4}$, where $C > 0$ is a universal constant.*
2. *If $n = 2^k$, then $\left|\widetilde{B}_{1/n}^n\right|$ is super-exponential.*

Unlike for the hypercube, we do not have a characterization theorem for the lattice $\mathbb{Z}^n$. In fact, we have found a locally 1/2-biased function for $\mathbb{Z}$ and a locally 1/4-biased function for $\mathbb{Z}^2$; see Figure 4.2. Both of these are not the result of embedding the relevant hypercube in the lattice via Proposition 4.24.

Question 4.26. *Give a complete characterization of permissible p values for locally p-biased functions on $\mathbb{Z}^n$. When such functions exist, count how many there are.*

Cayley Graphs

In general, for a given group with a natural generating set, it is interesting to ask whether its Cayley graph admits locally biased or locally stable functions, and if so, how many. Specific examples which spring to mind for such groups are the group of permutations S_n with all transpositions $\{\sigma_{ij}\}_{i<j}$, and $\mathbb{Z}$ with any number of generators. For the latter case, the following observation shows that for any two generators, $\mathbb{Z}$ has a locally 1/2-biased function:

Observation 4.27. *Let $a > 1$ and $b > 1$ generate $\mathbb{Z}$. Then the function f defined by*

$$f(x) = \begin{cases} 1, & 0 \leq (x \mod 2(a+b)) < a+b \\ -1, & a+b \leq (x \mod 2(a+b)) < 2(a+b) \end{cases}$$

is locally 1/2-biased.

Computer search shows that for some generators, other locally biased functions exist; see Figure 4.3 for an example.

Question 4.28. *Characterize the locally biased and locally stable functions on S_n as a function of its generating set.*

Question 4.29. *Characterize the locally biased and locally stable functions on $\mathbb{Z}$ as a function of its generating set.*

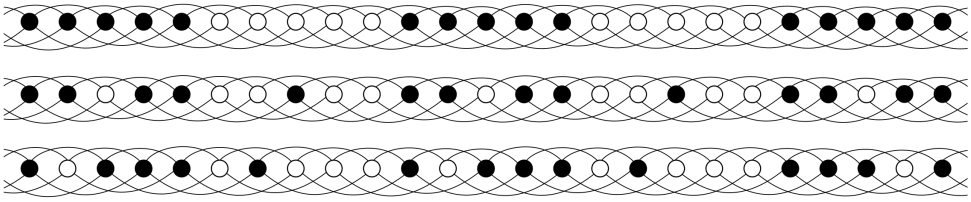


Figure 4.3: Three non-isomorphic locally 1/2-biased functions for $\mathbb{Z}$ with the generators $\{2, 3\}$. Computer search shows that these are the only ones.

4.5.3 Locally biased and locally stable functions

Section § 4.3 only gives lower bounds on the number of locally biased functions, and applies only for $p = 1/2$ and $p = 1/n$ (and $1 - 1/n$ by taking negation of functions).

Question 4.30. *What are the exact asymptotics for the number of non-isomorphic locally biased functions, for all permissible p?*

We can also ask about the robustness of the locally biased property:

Question 4.31. *How do the characterization and counting theorems for locally biased functions change, when we relax the locally biased demand for $2^{o(n)}$ of the vertices (i.e a small amount of vertices can have their neighbors labeled arbitrarily)?*

The uniqueness of locally $1/n$-stable functions is in stark contrast to the exponential size of locally $1/n$-biased functions. Our bounds in section § 4.4 for the number of $(n-m)/n$-locally stable functions are exponential in m, but not in n. We seek a better understanding of these functions:

Question 4.32. *What are the exact asymptotics for the number of non-isomorphic locally stable functions?*

Bibliography

[1] Milton Abramowitz and Irene A. Stegun. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, volume 55 of *National Bureau of Standards Applied Mathematics Series*. For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C., 1964.

[2] Milla Anttila, Keith Ball, and Irini Perissinaki. The central limit problem for convex bodies. *Trans. Amer. Math. Soc.*, 355(12):4723–4735, 2003.

[3] Shiri Artstein-Avidan, Apostolos Giannopoulos, and Vitali D. Milman. *Asymptotic geometric analysis. Part I*, volume 202 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2015.

[4] Laszlo Babai, Peter Frankl, and Janos Simon. Complexity classes in communication complexity theory. In *Proceedings of the 27th Annual Symposium on Foundations of Computer Science*, SFCS '86, pages 337–347, Washington, DC, USA, 1986. IEEE Computer Society.

[5] Itai Benjamini and Harry Kesten. Distinguishing sceneries by observing the scenery along a random walk path. *J. Anal. Math.*, 69:97–135, 1996.

[6] Ludwig Berwald. Verallgemeinerung eines Mittelwertsatzes von J. Favard für positive konkave Funktionen. *Acta Math.*, 79:17–37, 1947.

[7] Tommy Bonnesen and Werner Fenchel. *Theory of convex bodies*. BCS Associates, Moscow, ID, 1987. Translated from the German and edited by L. Boron, C. Christenson and B. Smith.

[8] Christer Borell. Convex measures on locally convex spaces. *Ark. Mat.*, 12:239–252, 1974.

[9] Alan Edelman and Brian D. Sutton. The beta-Jacobi matrix model, the CS decomposition, and generalized singular value problems. *Found. Comput. Math.*, 8(2):259–285, 2008.

[10] Hilary Finucane, Omer Tamuz, and Yariv Yaari. Scenery reconstruction on finite abelian groups. *Stochastic Processes and their Applications*, 124(8):2754 – 2770, 2014.

[11] Bruno Fleury. Concentration in a thin Euclidean shell for log-concave measures. *J. Funct. Anal.*, 259(4):832–841, 2010.

[12] Matthieu Fradelizi. Sections of convex bodies through their centroid. *Arch. Math. (Basel)*, 69(6):515–522, 1997.

[13] Christophe Garban and Jeffrey E. Steif. *Noise sensitivity of Boolean functions and percolation*. Institute of Mathematical Statistics Textbooks. Cambridge University Press, New York, 2015.

[14] Renan Gross and Uri Grupel. Indistinguishable sceneries on the boolean hypercube. *Combinatorics, Probability and Computing*, page 1–15, 2018.

[15] Uri Grupel. Sampling on the sphere by mutually orthogonal subspaces. In *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 973–983. SIAM, Philadelphia, PA, 2017.

[16] Olivier Guédon and Emanuel Milman. Interpolating thin-shell and sharp large-deviation estimates for isotropic log-concave measures. *Geom. Funct. Anal.*, 21(5):1043–1068, 2011.

[17] Godfrey Harold Hardy and Srinivasa Ramanujan. Asymptotic formulæ in combinatory analysis. *Proceedings of the London Mathematical Society*, s2-17(1):75–115, 1918.

[18] William B. Johnson and Joram Lindenstrauss. Extensions of Lipschitz mappings into a Hilbert space. In *Conference in modern analysis and probability (New Haven, Conn., 1982)*, volume 26 of *Contemp. Math.*, pages 189–206. Amer. Math. Soc., Providence, RI, 1984.

[19] Bo'az Klartag. A central limit theorem for convex sets. *Invent. Math.*, 168(1):91–131, 2007.

[20] Bo'az Klartag. Power-law estimates for the central limit theorem for convex sets. *J. Funct. Anal.*, 245(1):284–310, 2007.

[21] Bo'az Klartag. A Berry-Esseen type inequality for convex bodies with an unconditional basis. *Probab. Theory Related Fields*, 145(1-2):1–33, 2009.

[22] Bo'az Klartag and Oded Regev. Quantum one-way communication can be exponentially stronger than classical communication. In *STOC'11—Proceedings of the 43rd ACM Symposium on Theory of Computing*, pages 31–40. ACM, New York, 2011.

[23] Ilan Kremer. Quantum communication. Master's thesis, Hebrew University, Computer Science Department, 1995.

[24] Denis S. Krotov and Sergey V. Avgustinovich. On the number of -perfect binary codes: A lower bound. *IEEE Trans. Inf. Theor.*, 54(4):1760–1765, April 2008.

[25] Elon Lindenstrauss. Indistinguishable sceneries. *Random Structures Algorithms*, 14(1):71–86, 1999.

[26] Jacobus Hendricus Van Lint. A survey of perfect codes. *Rocky Mountain J. Math.*, 5(2):199–224, 06 1975.

[27] Jacobus Hendricus Van Lint. *Introduction to Coding Theory*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 3rd edition, 1998.

[28] László Lovász and Santosh Vempala. Hit-and-run from a corner. In *Proceedings of the 36th Annual ACM Symposium on Theory of Computing*, pages 310–314. ACM, New York, 2004.

[29] Heinrich Matzinger and Silke W.W. Rolles. Reconstructing a piece of scenery with polynomially many observations. *Stochastic Processes and their Applications*, 107(2):289 – 300, 2003.

[30] Claus Müller. *Spherical harmonics*, volume 17 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin-New York, 1966.

[31] Ryan O'Donnell. *Analysis of Boolean Functions*. Cambridge University Press, New York, NY, USA, 2014.

[32] Gilles Pisier. *The Volume of Convex Bodies and Banach Space Geometry:*. Cambridge University Press, Cambridge, 10 1989.

[33] Ran Raz. Exponential separation of quantum and classical communication complexity. In *Annual ACM Symposium on Theory of Computing (Atlanta, GA, 1999)*, pages 358–367. ACM, New York, 1999.

[34] Alexander A. Razborov. Communication complexity. In *An Invitation to Mathematics*, pages 97–117. Springer-Verlag Berlin Heidelberg, 2011.

[35] Lucy Joan Slater. *Generalized hypergeometric functions*. Cambridge University Press, Cambridge, 1966.

[36] Gábor Szegő. *Orthogonal polynomials*. American Mathematical Society, Providence, R.I., fourth edition, 1975. American Mathematical Society, Colloquium Publications, Vol. XXIII.

[37] Peter van Hintum. Biased partitions of $\mathbb{Z}^n$. *ArXiv e-prints*, May 2018.

[38] Roman Vershynin. Introduction to the non-asymptotic analysis of random matrices. In *Compressed sensing*, pages 210–268. Cambridge Univ. Press, Cambridge, 2012.

[39] Andrew Chi Chih Yao. Probabilistic computations: toward a unified measure of complexity (extended abstract). In *18th Annual Symposium on Foundations of Computer Science (Providence, R.I., 1977)*, pages 222–227. IEEE Comput. Sci., Long Beach, Calif., 1977.

[40] Andrew Chi-Chih Yao. Some complexity questions related to distributive computing(preliminary report). In *Proceedings of the Eleventh Annual ACM Symposium on Theory of Computing*, STOC '79, pages 209–213, New York, NY, USA, 1979. ACM.