Gaussian Isoperimetry for Multiple Sets

by

Steven Heilman

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Department of Mathematics

New York University

May 2014

Professor Assaf Naor

Acknowledgements

I would first like to thank my advisor, Assaf Naor, for guidance and encouragement. I would then like to thank all of my teachers over the years. I would particularly like to mention my high school math teacher Mr. Plymette. I then thank the faculty and staff at my undergraduate institution. I would thank Hsiao-Bing Cheng, Rick Durrett, Ravi Ramakrishnan, Alex Vladimirsky, James West, and I would especially thank my analysis teachers Leonard Gross, Camil Muscalu, Laurent Saloff-Coste, and of course my undergraduate advisor, Bob Strichartz.

I would now like to thank my teachers at NYU, including Sourav Chatterjee, Sinan Gunturk, Larry Guth, Bruce Kleiner, Robert V. Kohn, Fang-Hua Lin, Jalal Shatah and S. R. S. Varadhan. I would like to single out Jalal Shatah for pointed advice which led to my choosing NYU as a graduate school during my first visit. I would also thank my oral exam committee, including Fedor Bogomolov, Percy Deift, Pierre Germain, Bruce Kleiner, and Henry McKean. I would then thank my thesis committee of Jop Briët, Jeff Cheeger, Oded Regev, Assaf Naor, and Joel Spencer. Thanks also to the friends I have made at NYU, including but not limited to Evan Chou, Doris Dobi, Arie Israel, Aukosh Jagannath, Sean Li, Trushant Majmudar, Behzad Mehrdad, and Jim Portegies. I would also like to thank Assaf Naor for organizing the Quantitative Geometry Program at MSRI and for inviting me to attend. I also thank Elchanan Mossel for inviting me to the Simons Institute for Theoretical Computer Science. Thanks also to Jill Pipher and the Network Science and Graph Algorithms Program at ICERM. Thanks also to my collaborators Aukosh Jagannath, Elchanan Mossel, Assaf Naor and Joe Neeman, who created part of the content of this thesis.

I would now like to thank my family and especially my parents. I would also like to thank the entire Khadem family and especially Nasim. Lastly, this thesis is dedicated to Marshall and Naomi Spector.

Abstract

We discuss a few conjectural Gaussian isoperimetric inequalities where traditional symmetrization or heat flow methods fail. Generally speaking, we are trying to maximize a convex function on a convex domain. So, any method that tries to maximize a function by moving in the direction of higher gradient should fail. Symmetrization and heat flow methods seem to fall into this category, so we need to develop new approaches to our problems.

The central object of study is the Ornstein-Uhlenbeck operator, i.e. the heat operator on Euclidean space equipped with the Gaussian measure. In this setting, Fourier analysis with respect to Hermite polynomials becomes most relevant. Due to the relation of the hypercube and the Gaussian measure via Central Limit Theorems, these Gaussian isoperimetric results imply inequalities for discrete functions, which then have applications to theoretical computer science.

The first problem we discuss asks for the maximum of the sum squared first-order Hermite-Fourier coefficients over all partitions of Euclidean space. The second problem we discuss asks for the most noise stable partitions of Euclidean space. The latter problem is a generalization of the so-called Gaussian double bubble problem. We give partial solutions to these two problems, and we also give some negative results for the latter problem.

Contents

List of Figures

Chapter 1

Introduction

1.1 History of Isoperimetry

Classical isoperimetry can be traced to ancient times, though its full understanding still remains incomplete. Generally speaking, we look for an object with least perimeter among all objects of fixed volume. And we expect that the smallest perimeter object has a simple structure. For example, if our notions of volume and perimeter have some symmetries, we expect the smallest perimeter object of fixed volume to also exhibit some symmetries. Since the Euclidean volume and Euclidean surface area are highly symmetric, we expect that the set of fixed Euclidean volume and least Euclidean perimeter will be highly symmetric. However, symmetry considerations are often insufficient to find a set of least perimeter [84].

The classical isoperimetric problem asks for a Euclidean set of fixed Euclidean measure and minimal surface area. The solution of this problem, the Euclidean ball, has been known since antiquity. Steiner nearly proved this fact in 1838 using symmetrization [95]. Steiner showed that, starting from a fixed Euclidean set A, it is possible to create a more symmetric set from A with the same volume and smaller or equal surface area. This argument was made rigorous by Weierstrass in 1870 [101]. We can think of symmetrization as a gradient descent method. That is, starting from any set, there exists a way to decrease the perimeter of this set, such that we are now closer to a least perimeter object, with respect to some metric. In this way, the classical isoperimetric problem can be intuitively understood as maximizing a concave function over a convex set. The standard reference for symmetrization methods is now [21], where several previous approaches are unified.

Hurwitz solved the isoperimetric problem in the plane using the rather striking method of Fourier series in 1902. Given a non-intersecting closed rectifiable curve in the plane, the curve can be represented as a periodic function of one real variable with convergent Fourier series. Then, Fourier analysis shows that the perimeter is smallest when all higher order Fourier coefficients are zero. Surprisingly, it is still an open problem to try to extend his proof to all Euclidean spaces. Yet, for convex bodies, the essence of his proof can be used, since in that case, one can express a convex body as a function on the sphere, and then expand this function into spherical harmonics [59].

In the mid-1900s, mathematicians began to solve isoperimetric problems in spaces other than Euclidean space. In the sphere, the set of fixed volume and smallest boundary is a cap, or geodesic ball [67], as shown by Levy. A nice proof of this fact by symmetrization appears in [30]. Levy's result was generalized by Gromov [73] in the 1980s. Gromov's result says that in a smooth manifold with Ricci curvature bounded below by a positive number R, the boundary of any set is greater than the boundary of a geodesic ball on a sphere with Ricci curvature R. Gromov used the methods of comparison geometry, and he realized that a crucial result of Almgren [1] gave the existence and regularity for Gromov's isoperimetric problem.

We should mention that existence and regularity of isoperimetric problems has been an important issue. In particular, Steiner's avoidance of this issue led to the error in his proof. In the mid-1970s, Almgren achieved a breakthrough in regularity theory which would change the entire direction of isoperimetric theory [1]. As we just mentioned, this result allowed Gromov's general isoperimetric inequality. However, Almgren's result also gave the existence and regularity for a wide class of isoperimetric problems, including problems involving multiple sets. The two set isoperimetric problem or double bubble problem asks for two Euclidean sets of fixed volume and smallest combined surface area. That is, if we have two sets A_1, A_2 of fixed Euclidean volume, how can we minimize volume of $(\partial A_1)\cup(\partial A_2)$? Thanks to the rigorous foundation given by [1] to this ancient problem, the double bubble problem was finally solved in 2002 [49, 89]. Essentially nothing is known when try to solve this problem for three or more sets. That is, if we are given three sets A_1, A_2, A_3 of fixed Euclidean volume, we cannot find the minimum volume of $(\partial A_1)\cup(\partial A_2)\cup(\partial A_3)$. The three set case of this problem in the plane was solved in 2004 by Wichiramala [102] by a variational argument which considers over fifty separate cases. So, the proof does not seem to generalize to higher dimensions.

Concurrent with Almgren's existence and regularity result, even more isoperimetric inequalities were investigated. In particular, the isoperimetric problem with respect to the Gaussian measure was investigated. This measure could be considered natural in light of the Central Limit Theorem, or the fact that it is the unique rotation invariant product probability measure on Euclidean space. We now need to introduce some notation to state the Gaussian isoperimetric problem.

Let $n\geq 1$, $n\in\mathbb{Z}$, let $x=(x_1,\ldots,x_n)\in\mathbb{R}^n$, define $\|x\|_2^2:=x_1^2+\cdots+x_n^2$, and define

$$\gamma_n(x):=(2\pi)^{-n/2}\cdot e^{-\|x\|_2^2/2}. \tag{1.1}$$

Let $A\subseteq\mathbb{R}^n$ be a set with smooth boundary and let $H\subseteq\mathbb{R}^n$ be a half-space.

That is, H is the region that lies on one side of a hyperplane. Denote the Gaussian perimeter of ∂A by

$$\gamma_{n-1}(\partial A) := \liminf_{\varepsilon\to 0} (2\varepsilon)^{-1}\gamma_n(x \in \mathbb{R}^n \colon \exists\, y \in \partial A, \|x-y\|_2 < \varepsilon).$$

The Gaussian Isoperimetric Inequality says that the half-space H has the smallest Gaussian boundary among all sets of fixed Gaussian volume.

Theorem 1.1.1 (**Gaussian Isoperimetric Inequality**). *[16, 96]*

$$\gamma_n(A) = \gamma_n(H) \quad \Longrightarrow \quad \gamma_{n-1}(\partial A) \geq \gamma_{n-1}(\partial H). \tag{1.2}$$

In the 1970s, Borell and Sudakov-Tsirelson proved Theorem 1.1.1 using symmetrization methods. Since then, the result has been strengthened and extended in various ways, which we will discuss further below. We should mention that the Gaussian analogue of the Double Bubble Theorem [49] has been solved in 2008 [27] by first solving the double bubble problem on the sphere, and then using that the Gaussian measure is the limit of projections of the Haar measure on the n-dimensional sphere of radius $\sqrt{n}$, as $n \to \infty$.

We would now like to describe how Theorem 1.1.1 has been generalized. To do this, we need to introduce some operators. For $r > 0$, define $B(x, r) := \{y \in \mathbb{R}^n \colon \|x-y\|_2 \leq r\}$. Let dx and vol_n be Lebesgue measure on $\mathbb{R}^n$. Let $f \colon \mathbb{R}^n \to \mathbb{R}$ be a bounded function and let $t \geq 0$. Let $\Delta := -\sum_{i=1}^n \partial^2/\partial x_i^2$. Let $e^{-t\Delta}f(x)$ denote the **classical heat semigroup** applied to f. That is, for $x \in \mathbb{R}^n$,

$$e^{-t\Delta}f(x) := \int_{\mathbb{R}^n} f(x + y\sqrt{2t}\,)d\gamma_n(y).$$

The classical isoperimetric inequality is a consequence of the following inequality for the heat semigroup $e^{-t\Delta}$, which can be proven via symmetrization.

Theorem 1.1.2 (**Riesz-Rearrangement Inequality**). *[65, 21]. For all $t \geq 0$,*

$$\mathrm{vol}_n(A) = \mathrm{vol}_n(B(0,r)) \quad \Longrightarrow \quad \int_{\mathbb{R}^n} 1_A \cdot e^{-t\Delta}1_A dx \leq \int_{\mathbb{R}^n} 1_{B(0,r)} \cdot e^{-t\Delta}1_{B(0,r)}\, dx. \tag{1.3}$$

So, as heat flows out of a given set, the one that preserves the most of its heat is the ball. To get the classical isoperimetric inequality from Theorem 1.1.2, let $t \to 0^+$ in (1.3). Note also that both quantities in (1.3) are equal at $t = 0$, so differentiating both sides at $t = 0$ preserves the inequality, once we normalize appropriately.

Theorem 1.1.2 can also be considered a variant of the Brunn-Minkowski Inequality

Theorem 1.1.3 (**Brunn-Minkowski Inequality**). *[73, 86, 34] Let* $A, B \subseteq \mathbb{R}^n$ *be measurable and nonempty.*

(a) $\forall\ \lambda \in [0,1]$, $\mathrm{vol}_n(\lambda A + (1-\lambda)B) \geq \mathrm{vol}_n(A)^{\lambda}\mathrm{vol}_n(B)^{1-\lambda}$.

(b) $\mathrm{vol}_n(A+B)^{1/n} \geq \mathrm{vol}_n(A)^{1/n} + \mathrm{vol}_n(B)^{1/n}$.

To see the analogy between Theorem 1.1.3 and Theorem 1.1.2, let $A \subseteq \mathbb{R}^n$ and let B be a Euclidean ball with $\mathrm{vol}(A) = \mathrm{vol}(B)$. Then for all $\varepsilon > 0$, Theorem 1.1.3(b) implies that $\mathrm{vol}_n(A+\varepsilon B) \geq \mathrm{vol}_n(B+\varepsilon B)$. Moreover, $e^{-t\Delta}1_A$ is an average of 1_A over balls. There are several available proofs of Theorem 1.1.3. The quickest proof uses a measure transportation argument for the case $n = 1$, and then an induction on n.

We now provide a Gaussian version of Theorem 1.1.2. To this end, we first replace the heat semigroup by the **Ornstein-Uhlenbeck semigroup**. Define the operator L by $L := \Delta + \sum_{i=1}^{n} x_i \frac{\partial}{\partial x_i}$.

Definition 1.1.4 (**Ornstein-Uhlenbeck semigroup**)**.** For $x \in \mathbb{R}^n$ and $t \geq 0$, define the Ornstein-Uhlenbeck semigroup applied to f by

$$e^{-tL}f(x) := \int_{\mathbb{R}^n} f(xe^{-t} + y\sqrt{1-e^{-2t}})d\gamma_n(y).$$

For our purposes below, it is sometimes desirable to use a non-semigroup version of this operator. Therefore, let $\rho \in (-1,1)$, and define the operator

$$T_\rho f(x) := \int_{\mathbb{R}^n} f(x\rho + y\sqrt{1-\rho^2})d\gamma_n(y).$$

These operators are also known as the noise operators.

Definition 1.1.5 (**Noise Stability**)**.** Let $f\colon \mathbb{R}^n \to [0,1]$. For $\rho \in (-1,1)$, define the **Noise Stability** of f with parameter ρ as

$$\int_{\mathbb{R}^n} f \cdot T_\rho f d\gamma_n.$$

Then Theorem 1.1.2 can be generalized to the following inequality of Borell from 1985.

Theorem 1.1.6. *[17] For all* $t \geq 0$*,*

$$\gamma_n(A) = \gamma_n(H) \quad \Longrightarrow \quad \int_{\mathbb{R}^n} 1_A \cdot e^{-tL}1_A d\gamma_n \leq \int_{\mathbb{R}^n} 1_H \cdot e^{-tL}1_H d\gamma_n. \tag{1.4}$$

Note the similarity of Theorem 1.1.6 and Theorem 1.1.2. As in the Euclidean case, to get (1.2) from (1.4), let $t \to 0$ in (1.4). Borell's original result (1.4) from [17] has been highly influential, and my research program can be summarized as an attempt to duplicate his success in solving an isoperimetric problem that has many consequences. Since Borell's paper, Theorem 1.1.6 has been strengthened. In the work [77], the authors use monotonicity methods to prove a stability version of Theorem 1.1.6. On the one hand, they differentiate a suitable quantity in t, show that this quantity is monotone, and deduce (1.4) as a consequence. Moreover, they use the size of this differentiated quantity to show that the inequality (1.4) is almost an equality if and only if A is close a half-space. A loosely related stochastic calculus version of this result appears in [29] with tighter error bounds.

In 2003, Borell proved a Gaussian version of the Brunn Minkowski inequality, Theorem 1.1.3. He uses a rather interesting twist on the parabolic maximum principle and the heat flow method. That is, he smooths out sets $A, B \subseteq \mathbb{R}^n$ using the classical heat semigroup, and he then uses a parabolic maximum principle to prove an infinite family of inequalities, of which the following Brunn-Minkowski (or Ehrhard) Inequality is a special case. Unfortunately, the proof is so elegant that it is unclear how it can apply to other problems.

Theorem 1.1.7 (**Gaussian Brunn-Minkowski/ Ehrhard Inequality**). *[18] For $a \in \mathbb{R}$, define $\Phi\colon \mathbb{R} \to [0,1]$ by $\Phi(a) := \int_{-\infty}^{a} d\gamma_1$. Let $A, B \subseteq \mathbb{R}^n$. Then for all $\lambda \in [0,1]$,*

$$\Phi^{-1}(\gamma_n(\lambda A + (1-\lambda)B)) \geq \lambda \Phi^{-1}(\gamma_n(A)) + (1-\lambda)\Phi^{-1}(\gamma_n(B)). \tag{1.5}$$

We should mention that Gaussian isoperimetry has been developed in the context of self-shrinking solutions to the mean-curvature flow, as in the following sequence of recent papers by Colding, Minicozzi and others [24, 23, 25]. It is unclear whether or not their methods can apply to noise stability, since their methods seems special to isoperimetric problems. Nevertheless, it could be possible to extend their methods to noise stability problems for ρ close to 1. Such a strategy has at least been successful for related isoperimetric problems. See for example [36] and the references therein.

1.2 Theoretical Computer Science

One of the reasons for renewed interest in isoperimetric theory in general and in Theorem 1.1.6 in particular comes from theoretical computer science. From this perspective, Gaussian isoperimetric inequalities become the most relevant. One of the most fundamental problems in computational complexity, the MAX-CUT problem asks for the partition of the vertices of an undirected graph into two sets S and S^c that maximizes the number of edges that go from S to S^c. For a graph on n vertices, it is expected that in general we cannot solve the MAX-CUT problem in time which is polynomial in n. More specifically, MAX-CUT is an NP-complete problem, and it is also in the complexity class MAX-SNP, which is the optimization version of NP [85]. So, if $P \neq NP$, then there is no way to solve the MAX-CUT problem in general, in time polynomial in n. Fortunately, one can approximately solve the MAX-CUT problem in time polynomial in n. In fact, any problem in MAX-SNP can be approximately solved in polynomial time [85, Theorem 1]. The best known approximation for MAX-CUT appeared in a landmark paper of Goemans and Williamson in 1995 [35], where it was shown that we can find in polynomial time a two-set partition of any undirected graph which cuts at least .87856 times the maximum possible number of cut edges.

In order to describe the technique of the paper [35], we first define the MAX-CUT problem.

Definition 1.2.1 (MAX-CUT). We define the weighted MAX-CUT problem for $n \in \mathbb{N}$. In this problem, we are given a symmetric matrix $\{a_{ij}\}_{i,j=1}^{n}$ with $a_{ij} \geq 0$ for all $i, j \in \{1, \ldots, n\}$. The goal of the MAX-CUT problem is to find the quantity:

$$\max_{c\colon \{1,\ldots,n\}\to\{1,2\}} \sum_{\substack{i,j\in\{1,\ldots,n\}:\\ c(i)\neq c(j)}} a_{ij}.$$

In particular, if $V = \{1, \ldots, n\}$, we can think of a_{ij} as the weight of an undirected edge from vertex i to vertex j. Written another way, we want to compute

$$\max_{\varepsilon_1,\ldots,\varepsilon_n\in\{-1,1\}} \frac{1}{2}\sum_{i,j=1}^{n} a_{ij}(1-\varepsilon_i\varepsilon_j). \tag{1.6}$$

Goemans and Williamson approximate (1.6) by replacing the numbers ε_i by unit vectors v_i in the unit sphere $S^{n-1} \subseteq \mathbb{R}^n$, and by maximizing the same quantity.

$$\max_{v_1,\ldots,v_n\in S^{n-1}} \frac{1}{2}\sum_{i,j=1}^{n} a_{ij}(1-\langle v_i, v_j\rangle). \tag{1.7}$$

Since unit vectors in S^{n-1} include unit vectors in $\{-1, 1\}$, the quantity (1.7) exceeds the quantity (1.6).

$$\max_{v_1,\ldots,v_n\in S^{n-1}} \frac{1}{2}\sum_{i,j=1}^{n} a_{ij}(1-\langle v_i, v_j\rangle) \geq \max_{\varepsilon_1,\ldots,\varepsilon_n\in\{-1,1\}} \frac{1}{2}\sum_{i,j=1}^{n} a_{ij}(1-\varepsilon_i\varepsilon_j). \tag{1.8}$$

Crucially, the space of all $n \times n$ matrices of the form $\{\langle v_i, v_j\rangle\}_{i,j=1}^{n}$ is a cone, where $v_i \in \mathbb{R}^n$ for all $i = 1, \ldots, n$. That is, if A, B are matrices in this cone, then for any

$\lambda \in [0,1]$, $\lambda A + (1-\lambda)B$ is also in this cone. This cone is known as the cone of positive semidefinite matrices. And the restriction $v_i \in S^{n-1}$ can be written as a linear constraint in the cone of positive semidefinite matrices: $\langle v_i, v_i\rangle = 1$, for all $i = 1,\ldots,n$. So, the quantity (1.7) can be computed in polynomial time in n, since it involves a function which is linear on the cone of positive semidefinite matrices. Such functions are known as semidefinite programs, and they can be solved by, e.g. the Ellipsoid Method [100].

Finally, the vectors achieved by solving (1.7) can be related back to (1.6) using a randomized rounding method. That is, given vectors $r_1,\ldots,r_n \in S^{n-1}$ achieving the maximum in (1.7), let $\Pi \subseteq \mathbb{R}^n$ be a random hyperplane such that $0 \in \Pi$. For all $i \in \{1,\ldots,n\}$ such that r_i lies on one side of this hyperplane, define the random sign $\nu_i := -1$. For all $i \in \{1,\ldots,n\}$ such that r_i lies on the opposite side of this hyperplane, define the random sign $\nu_i := 1$. In expectation, we then compute

$$\begin{aligned}
\mathbb{E}\frac{1}{2}\sum_{i,j=1}^{n} a_{ij}(1-\nu_i\nu_j) &= \sum_{i,j=1}^{n} a_{ij}\frac{\cos^{-1}\langle r_i, r_j\rangle}{\pi} \\
&\geq \left[\min_{0\leq\theta\leq\pi} \frac{2}{\pi}\frac{\theta}{1-\cos\theta}\right]\frac{1}{2}\sum_{i,j=1}^{n} a_{ij}(1-\langle r_i, r_j\rangle) \\
&\approx [.87856]\frac{1}{2}\sum_{i,j=1}^{n} a_{ij}(1-\langle r_i, r_j\rangle) \\
&= [.87856]\max_{v_1,\ldots,v_n\in S^{n-1}}\frac{1}{2}\sum_{i,j=1}^{n} a_{ij}(1-\langle v_i, v_j\rangle).
\end{aligned} \tag{1.9}$$

Finally, combining (1.9) and (1.8),

$$\mathbb{E}\frac{1}{2}\sum_{i,j=1}^{n} a_{ij}(1-\nu_i\nu_j) \geq [.87856]\max_{\varepsilon_1,\ldots,\varepsilon_n\in\{-1,1\}}\frac{1}{2}\sum_{i,j=1}^{n}(1-\varepsilon_i\varepsilon_j). \tag{1.10}$$

In summary, [35] used semidefinite programming to approximately solve MAX-CUT in polynomial time within a multiplicative factor .87856, as shown by (1.10). Since this paper, no one has been able to improve upon this approximation factor, and there is strong evidence that it may never be improved at all. The first hint of such a result appeared in 1991 [85] (before [35]) where it was shown that there exists some constant $c > 0$ such that, if $P \neq NP$, then there is no polynomial time algorithm that approximately solves the MAX-CUT problem within the constant c. An explicit constant 71/72 was found in 1998 by[14]. Then, in 2001 Håstad [44] showed that if $P \neq NP$, then there is no polynomial time approximation algorithm for the MAX-CUT problem with approximation ratio exceeding 16/17. He also introduced discrete Fourier analysis as a tool for proving results of this type.

However, $16/17 > .87856$, so we still did not know the best possible approximation ratio for solving MAX-CUT. In 2007 [54], it was shown that the ratio .87856 should be the best possible. That is, assuming the Unique Games Conjecture (UGC), it is not possible to approximately solve MAX-CUT within a multiplicative factor greater than the Goemans-Williamson constant .87856 [78, Corollary 2.12]. The Unique Games Conjecture can be considered a contemporary proxy for the $P \neq NP$ problem. However, the Unique Games Conjecture could be false, whereas it could still be true that $P \neq NP$. The result of [54] combines the Fourier analytic methods of [44] with the semidefinite programming methods of [35], with some new ingredients as well. In particular, the hardness result of [54] uses the Majority is Stablest Theorem, which was shown to be a (nontrivial) consequence of Borell's isoperimetric result, Theorem 1.1.6.

The Majority is Stablest Theorem says the following. If we have an election between two candidates, if each candidate has an equal chance of winning, and if no one person has too much influence over the outcome of the election, then the best way to determine the winner of an election is to take the majority. Here the "best" voting method is the one which is most stable to errors in the tabulation of the votes, or noise. The exact statement of the theorem involves functions $f\colon \{1,2\}^n \to \{1,2\}$. Here the domain of f represents n votes for each of two candidates, the function f represents the voting method (e.g. taking the majority of the votes), and the range of f denotes the winner of the election, when given the votes as input. As already mentioned, the Majority is Stablest Theorem was shown to follow from Borell's result [78]. In fact, we can consider the Majority is Stablest Theorem to be the discrete version of Borell's Theorem 1.1.6. To prove the Majority is Stablest Theorem, the authors of [78] used an invariance principle, where it is shown that, for functions with low influences, swapping Boolean or Gaussian inputs of the function one at a time does not change the distribution of the function very much. This strategy, known as the Lindeberg replacement trick, had already been incorporated in an invariance principle of Chatterjee [22] and much earlier in 1979 by Rotar [91]. However, the result of [78] provides improved error bounds by also using the hypercontractive inequality.

In the interest of being precise, we now state the Majority is Stablest Theorem, the Invariance Principle, and the Unique Games Conjecture.

Let $f\colon \{-1,1\}^n \to \{-1,1\}$. For any $S \subseteq \{1,\ldots,n\}$, and $\sigma = (\sigma_1,\ldots,\sigma_n) \in \{-1,1\}^n$, define $W_S(\sigma) := \prod_{i\in S}\sigma_i$. Then $\{W_S\}_{S\subseteq\{1,\ldots,n\}}$ is an orthonormal basis for the space of functions $\{-1,1\}^n \to \mathbb{R}$ with respect to the inner product $\langle f,g\rangle := 2^{-n}\sum_{\sigma\in\{-1,1\}^n} f(\sigma)g(\sigma)$. So if we define $\widehat{f}(S) := \langle f, W_S\rangle$, then we can write $f = \sum_{S\subseteq\{1,\ldots,n\}} \widehat{f}(S)W_S$. Let $\rho \in (-1,1)$, and define the **Noise Operator**, or Ornstein-

Uhlenbeck or Bonami-Beckner operator, applied to f with parameter ρ by

$$T_\rho f := \sum_{S\subseteq\{1,\ldots,n\}} \rho^{|S|}\widehat{f}(S)W_S.$$

Let $x,y\in\{-1,1\}^n$ be random variables such that x is uniformly distributed in $\{-1,1\}^n$, and for each $i\in\{1,\ldots,n\}$, y_i satisfies $\mathbb{E}[x_iy_i]=\rho$. Then the **Noise Stability** of f with parameter ρ can be written as

$$2^{-n}\sum_{\sigma\in\{-1,1\}^n} f(\sigma)T_\rho f(\sigma)=\mathbb{E}_{x,y}[f(x)f(y)].$$

For each $i\in\{1,\ldots,n\}$, let $\alpha_i,\beta_i\in\{-1,1\}$ be i.i.d. random variables. Given $i\in\{1,\ldots,n\}$, define the **Influence** of the i^{th} variable on f by

$$\begin{aligned}\mathrm{Inf}_i f &:= \sum_{S\subseteq\{1,\ldots,n\}\colon\, i\in S} |\widehat{f}(S)|^2\\ &=\mathbb{P}[f(\alpha_1,\ldots,\alpha_n)\neq f(\alpha_1,\ldots,\alpha_{i-1},\beta_i,\alpha_{i+1},\ldots,\alpha_n)].\end{aligned}$$

Define $\mathrm{MAJ}_n\colon\{-1,1\}^n\to\{-1,1\}$ so that

$$\mathrm{MAJ}_n(\sigma_1,\ldots,\sigma_n)=\mathrm{sign}\Big(\sum_{i=1}^n\sigma_i\Big).$$

In the statement of Theorem 1.2.2, the definition of sign(0) does not matter.

Theorem 1.2.2 (Majority is Stablest). *[78] Let $0\leq\rho\leq 1$, and let $\varepsilon>0$. Then there exists $\tau>0$ such that if $f\colon\{-1,1\}^n\to\{-1,1\}$ satisfies $\#\{\sigma\in\{-1,1\}^n\colon f(\sigma)=1\}=\#\{\sigma\in\{-1,1\}^n\colon f(\sigma)=-1\}=2^{n-1}$ and $\mathrm{Inf}_i f<\tau$ for all $i=1,\ldots,n$, then*

$$\begin{aligned}2^{-n}\sum_{\sigma\in\{-1,1\}^n} f(\sigma)\cdot T_\rho f(\sigma)&\leq\frac{2}{\pi}\sin^{-1}(\rho)+\varepsilon\\ &=\lim_{m\to\infty}2^{-m}\sum_{\sigma\in\{-1,1\}^m}\mathrm{MAJ}_m(\sigma)\cdot T_\rho\mathrm{MAJ}_m(\sigma)+\varepsilon.\end{aligned}$$

Definition 1.2.3 (Γ-MAX-2LIN(k)). Let $k\in\mathbb{N}$, $k\geq 2$. We define the Γ-MAX-2LIN(k) problem. In this problem, we are given $m\in\mathbb{N}$ and $2m$ variables $x_i\in\mathbb{Z}/k\mathbb{Z}$, $i\in\{1,\ldots,2m\}$. We are also given a matrix $\{a_{ij}\}_{i,j=1}^{2m}$ with $a_{ij}\geq 0$ for all $i,j\in\{1,\ldots,2m\}$. An element (i,j) corresponds to one of m linear equations of the form $x_i-x_j=c_{ij}(\mathrm{mod}\,k)$, $i,j\in\{1,\ldots,2m\}$, $c_{ij}\in\mathbb{Z}/k\mathbb{Z}$. The goal of the Γ-MAX-2LIN(k) problem is to find the following quantity:

$$\max_{(x_1,\ldots,x_{2m})\in(\mathbb{Z}/k\mathbb{Z})^{2m}}\ \sum_{\substack{(i,j)\in E\colon\\ x_i-x_j=c_{ij}(\mathrm{mod}\,k)}} a_{ij}. \tag{1.11}$$

Definition 1.2.4 (**Unique Games Conjecture**, [52, 54]). For every $\varepsilon \in (0,1)$, there exists a prime number $p(\varepsilon)$ such that no polynomial time algorithm (with respect to the parameter m) can distinguish between the following two cases, for instances of Γ-MAX-2LIN($p(\varepsilon)$) with $a_{ij} = 1$ for all $i, j \in \{1, \ldots, 2m\}$:

(i) (1.11) is larger than $(1-\varepsilon)m$, or

(ii) (1.11) is smaller than εm.

If (1.11) were equal to m, then we could find $(x_1, \ldots, x_{2m})$ achieving the maximum in (1.11) by linear algebra. One can therefore interpret the Unique Games Conjecture as an assertion that approximate linear algebra is hard. Recently, it has been shown that instances of the Unique Games Problem can be solved in subexponential time [5], leading some to believe that the Unique Games Conjecture could be false. However, the truth or falsity of this Conjecture remains a major open problem.

The Unique Games Conjecture is more commonly written in the following equivalent way [54]. Define a **Unique game** instance as

$$\mathcal{L} = \mathcal{L}(G(V, W, E), k, \{\pi_{vw}\}_{(v,w)\in E}).$$

Here $G(V, W, E)$ is a regular bipartite graph with (left) vertex set V, (right) vertex set W, and edges E. For every edge $(v, w) \in E$ with $v \in V, w \in W$, there is a given permutation $\pi_{vw} \colon \{1, \ldots, k\} \to \{1, \ldots, k\}$. A labeling to the Unique Game instance is an assignment $\eta \colon V \cup W \to \{1, \ldots, k\}$. An edge is satisfied by a labeling η if $\eta(v) = \pi_{vw}(\eta(w))$. The goal is to find a labeling that maximizes the fraction of edges satisfied. We call this maximum $\mathrm{OPT}(\mathcal{L})$. We think of k as a constant, and the size of the graph as the size of the problem instance. The **Unique Games Conjecture** (weak form) says that, for arbitrarily small constants $\varepsilon, \delta > 0$, there exists $k = k(\varepsilon, \delta)$ such that no polynomial time algorithm can distinguish whether a Unique Games instance $\mathcal{L}$ with k labels satisfies $\mathrm{OPT}(\mathcal{L}) \geq 1-\varepsilon$ or $\mathrm{OPT}(\mathcal{L}) \leq \delta$.

1.3 Isoperimetry for More than Two Sets, MAX-k-CUT, and Plurality

We are now prepared to synthesize the above research threads and discuss the main problems of this thesis. Mathematically speaking, it is natural to consider a generalization of Theorem 1.1.6 to multiple sets. For Euclidean sets and the Lebesgue measure, there are conjectured partitions of minimal surface area when partitioning into k parts in $\mathbb{R}^n$ as long as $k \leq n+1$ [97, Proposition 2] [75, p. 153]. These conjectured partitions are the union of subsets of spheres, and the partitions are intuitively understood as the collision of soap bubbles in nature. We then seek an analogous conjecture for the Gaussian measure.

Computationally speaking, we saw that Borell's Theorem 1.1.6 implied that the Goemans-Williamson semidefinite program gives the best approximation for the MAX-CUT problem, assuming the Unique Games Conjecture. If we generalize MAX-CUT to $k \geq 3$ sets, which we refer to as MAX-k-CUT, then there is a similar assertion. That is, if we could solve an analogue of Borell's Theorem 1.1.6 for $k \geq 3$ sets, then the Frieze-Jerrum [31] semidefinite program for the MAX-k-CUT problem would give the best approximation for the MAX-k-CUT problem, assuming the Unique Games Conjecture [50]. As in the case of MAX-CUT, we know that we cannot solve MAX-k-CUT in polynomial time [31]. And there exists a constant $c > 0$ such that we cannot approximately solve MAX-k-CUT within the constant c unless $P = NP$ [85].

The Frieze-Jerrum semidefinite program follows the same ideas of (1.6), (1.7) and (1.9), though since we need to partition vectors into $k \geq 3$ sets, we can no longer use a single hyperplane. Instead, we need to randomly rotate a partition $\{A_i\}_{i=1}^k$ of $\mathbb{R}^n$ into k disjoint sets such that $\gamma_n(A_i) = 1/k$ for each $i = 1, \ldots, k$. Then, just as in the Goemans-Williamson semidefinite program, we "round" vectors $v_i \in S^{n-1}$ to elements of $\{1, \ldots, k\}$. In particular, we assign two vectors v_i, v_j to the same element of $\ell \in \{1, \ldots, k\}$ if and only if there exists $\ell \in \{1, \ldots, k\}$ such that $v_i, v_j \in A_\ell$. The details of this argument are described in [50, Theorem A.6]. In the end, we want to solve the k-set version of Borell's Theorem 1.1.6, or equivalently, we want to find the optimal approximation ratio for MAX-k-CUT. So, for our purposes, it suffices to examine the analogue of the computation (1.9) for k sets. This computation is summarized in Theorem 1.3.3 below.

Definition 1.3.1 (**MAX-k-CUT**)**.** Let $k, n \in \mathbb{N}$, $k \geq 2$. We define the weighted MAX-k-CUT problem. We are given a symmetric matrix $\{a_{ij}\}_{i,j=1}^n$ with $a_{ij} \geq 0$ for all $i, j \in \{1, \ldots, n\}$. The goal of the MAX-k-CUT problem is to find the following quantity:

$$\max_{c\colon \{1,\ldots,n\}\to\{1,\ldots,k\}} \sum_{\substack{i,j\in\{1,\ldots,n\}\colon \\ c(i)\neq c(j)}} a_{ij}.$$

We now would like to state a conjecture which is the analogue of Borell's Theorem 1.1.6 for $k \geq 3$ sets. As we will see in Theorem 1.3.3, the quantity of Conjecture 1 naturally arises in the optimal approximation ratio of the MAX-k-CUT problem. Before stating the analogue of Borell's theorem, we need to describe the sets that we believe to be optimal within this conjecture.

Definition 1.3.2 (**Regular Simplicial Conical Partition**)**.** Let $z_1, \ldots, z_k \in \mathbb{R}^n$ be the vertices of a regular k-simplex, which is centered at the origin of $\mathbb{R}^n$. For $i = 1, \ldots, k$, let $B_i := \{x \in \mathbb{R}^n \colon \langle x, z_i\rangle = \max_{j=1,\ldots,k}\langle x, z_j\rangle\}$. Then $\{B_i\}_{i=1}^k$ is a partition of $\mathbb{R}^n$ into k regular simplicial cones. We therefore call $\{B_i\}_{i=1}^k$ a **regular simplicial conical partition** of $\mathbb{R}^n$ into k pieces.

Conjecture 1 (**Standard Simplex Conjecture**, [50]). *Let $n \geq 2$, let $\rho \in (-1,1)$, and let $3 \leq k \leq n+1$.*

(a) If $\rho \in (0,1)$, then

$$\begin{aligned}&\gamma_n(A_i) = 1/k,\, \forall\, i \in \{1,\ldots,k\} \quad \wedge \quad \cup_{i=1}^k A_i = \mathbb{R}^n \\ &\qquad \Longrightarrow \quad \sum_{i=1}^k \int_{\mathbb{R}^n} 1_{A_i} \cdot T_\rho 1_{A_i} d\gamma_n \leq \sum_{i=1}^k \int_{\mathbb{R}^n} 1_{B_i} \cdot T_\rho 1_{B_i} d\gamma_n.\end{aligned} \tag{1.12}$$

(b) If $\rho \in (-1,0)$ then

$$\begin{aligned}&\cup_{i=1}^k A_i = \mathbb{R}^n \quad \wedge \quad \gamma_n(A_i \cap A_j) = 0,\ \forall\, i,j \in \{1,\ldots,k\},\, i \neq j \\ &\qquad \Longrightarrow \quad \sum_{i=1}^k \int_{\mathbb{R}^n} 1_{A_i} \cdot T_\rho 1_{A_i} d\gamma_n \geq \sum_{i=1}^k \int_{\mathbb{R}^n} 1_{B_i} \cdot T_\rho 1_{B_i} d\gamma_n.\end{aligned} \tag{1.13}$$

Theorem 1.3.3. (**Optimal Approximation for MAX-k-CUT**, [50, Theorem 1.13],[31]). *Let $k \in \mathbb{N}$, $k \geq 2$. Let $\{B_i\}_{i=1}^k \subseteq \mathbb{R}^{k-1}$ be a regular simplicial conical partition. Define*

$$\begin{aligned}\alpha_k &:= \inf_{-\frac{1}{k-1} \leq \rho \leq 1} \frac{k - k^2 \sum_{i=1}^k \int_{\mathbb{R}^n} 1_{B_i} T_\rho 1_{B_i} d\gamma_n}{(k-1)(1-\rho)} \\ &= \inf_{-\frac{1}{k-1} \leq \rho \leq 0} \frac{k - k^2 \sum_{i=1}^k \int_{\mathbb{R}^n} 1_{B_i} T_\rho 1_{B_i} d\gamma_n}{(k-1)(1-\rho)}.\end{aligned}$$

Assume Conjecture 1 and the Unique Games Conjecture. Then, for any $\varepsilon > 0$, there exists a polynomial time algorithm that approximates MAX-k-CUT within a multiplicative factor $\alpha_k - \varepsilon$, and it is NP-hard to approximate MAX-k-CUT within a multiplicative factor of $\alpha_k + \varepsilon$.

As we saw for Borell's Theorem 1.1.6, there is a corresponding statement for discrete functions, known as the Majority is Stablest Theorem 1.2.2. Once again, there is a discrete version of Conjecture 1. This discrete problem is known as the Plurality is Stablest Conjecture, and it has the following informal description. If we have an election between $k \geq 3$ candidates, if each candidate has an equal chance of winning, and if no one person has too much influence over the outcome of the election, then the most noise-stable way to determine the winner of an election is to take the plurality. The exact statement of the theorem involves functions $f \colon \{1,\ldots,k\}^n \to \{1,\ldots,k\}$. Here the domain of f represents n votes for each of k

candidates, the function f represents the voting method (e.g. taking the plurality of the votes), and the range of f denotes the winner of the election, when given the votes as input.

We now make a few definitions and then state the Plurality is Stablest conjecture. This Conjecture seems to first appear in [54]. The work [54] emphasizes the applications of this conjecture to MAX-k-CUT and to Γ-MAX-2LIN(k), from Definitions 1.3.1 and 1.2.3.

Let $n \geq 2, k \geq 3$. Let $(W_1, \ldots, W_k)$ be an orthonormal basis for the space of functions $\{g\colon \{1,\ldots,k\} \to [0,1]\}$ equipped with the inner product $\langle g, h\rangle_k := \frac{1}{k}\sum_{\sigma\in\{1,\ldots,k\}} g(\sigma)h(\sigma)$. Assume that $W_1 = 1$. By orthonormality, for every $\sigma \in \{1,\ldots,k\}$, there exists $\widehat{g}(\sigma) \in \mathbb{R}$ such that the following expression holds: $g = \sum_{\sigma\in\{1,\ldots,k\}} \widehat{g}(\sigma)W_\sigma$. Define

$$\Delta_k := \{(x_1,\ldots,x_k) \in \mathbb{R}^k \colon \forall\, 1 \leq i \leq k, 0 \leq x_i \leq 1, \textstyle\sum_{i=1}^k x_i = 1\}.$$

Let $f\colon \{1,\ldots,k\}^n \to \Delta_k$, $f = (f_1,\ldots,f_k)$, $f_i\colon \{1,\ldots,k\}^n \to [0,1]$, $i \in \{1,\ldots,k\}$. Let $\sigma = (\sigma_1,\ldots,\sigma_n) \in \{1,\ldots,k\}^n$. Define $W_\sigma := \prod_{i=1}^n W_{\sigma_i}$, and let $|\sigma| := |\{i \in \{1,\ldots,n\}\colon \sigma_i \neq 1\}|$. Then for every $\sigma \in \{1,\ldots,k\}^n$ there exists $\widehat{f_i}(\sigma) \in \mathbb{R}$ such that $f_i = \sum_{\sigma\in\{1,\ldots,k\}^n} \widehat{f_i}(\sigma)W_\sigma$, $i \in \{1,\ldots,k\}$. For $\rho \in [-1,1]$ and $i \in \{1,\ldots,k\}$, define

$$T_\rho f_i := \sum_{\sigma\in\{1,\ldots,k\}^n} \rho^{|\sigma|}\widehat{f_i}(\sigma)W_\sigma, \quad T_\rho f := (T_\rho f_1,\ldots,T_\rho f_k) \in \mathbb{R}^k.$$

Let $m \geq 2$, $k \geq 3$. For each $j \in \{1,\ldots,k\}$, let $e_j = (0,\ldots,0,1,0,\ldots,0) \in \mathbb{R}^k$ be the j^{th} unit coordinate vector. Let $\sigma \in \{1,\ldots,k\}^n$. Define the function $\mathrm{PLUR}_{m,k}\colon \{1,\ldots,k\}^m \to \Delta_k$ such that

$$\mathrm{PLUR}_{m,k}(\sigma) := \begin{cases} e_j & ,\text{if } |\{i \in \{1,\ldots,m\}\colon \sigma_i = j\}| > \\ & \qquad |\{i \in \{1,\ldots,m\}\colon \sigma_i = r\}|, \\ & \qquad\qquad \forall\, r \in \{1,\ldots,k\}\setminus\{j\} \\ \frac{1}{k}\sum_{i=1}^k e_i & ,\text{otherwise} \end{cases}$$

Conjecture 2 (**Plurality is Stablest Conjecture**, [50])**.** *Let $n \geq 2$, $k \geq 3$, $\rho \in [-\frac{1}{k-1}, 1]$, $\varepsilon > 0$. Let $\langle\cdot,\cdot\rangle$ denote the standard inner product on $\mathbb{R}^n$. Then there exists $\tau > 0$ such that, if $f\colon \{1,\ldots,k\}^n \to \Delta_k$ satisfies $\sum_{\sigma\in\{1,\ldots,k\}^n\colon \sigma_j\neq 1}(\widehat{f_i}(\sigma))^2 \leq \tau$ for all $i \in \{1,\ldots,k\}$, $j \in \{1,\ldots,n\}$, then* *(a) If* $\rho \in (0,1]$*, and if* $\frac{1}{k^n}\sum_{\sigma\in\{1,\ldots,k\}^n} f(\sigma) = \frac{1}{k}\sum_{i=1}^k e_i$*, then*

$$\begin{aligned}
&\frac{1}{k^n}\sum_{\sigma\in\{1,\ldots,k\}^n}\langle f(\sigma), T_\rho f(\sigma)\rangle \\
&\qquad \leq \lim_{m\to\infty}\frac{1}{k^m}\sum_{\sigma\in\{1,\ldots,k\}^m}\langle \mathrm{PLUR}_{m,k}(\sigma), T_\rho(\mathrm{PLUR}_{m,k})(\sigma)\rangle + \varepsilon.
\end{aligned}$$

(b) If $\rho \in [-1/(k-1), 0)$*, then*

$$\begin{aligned}
&\frac{1}{k^n}\sum_{\sigma\in\{1,\ldots,k\}^n}\langle f(\sigma), T_\rho f(\sigma)\rangle \\
&\qquad \geq \lim_{m\to\infty}\frac{1}{k^m}\sum_{\sigma\in\{1,\ldots,k\}^m}\langle \mathrm{PLUR}_{m,k}(\sigma), T_\rho(\mathrm{PLUR}_{m,k})(\sigma)\rangle - \varepsilon.
\end{aligned}$$

One of our main results from Chapter 4 concerns a very specific solution of Conjecture 1. One of our other main results from Chapter 5 states that the following natural variant of Conjecture 1 is false. Conjecture 3 is essentially Conjecture 1 with an unequal measure assumption. Actually, a statement more general than Conjecture 3 is false, but for the purposes of this introduction, we state the simpler Conjecture 3. For the more general case, see Chapter 5.

Conjecture 3 (**Shifted Simplex Conjecture**)**.** *Let* $n \geq 2$*, let* $\rho \in (-1,1)$*. Let* $\{B_i\}_{i=1}^{n+1} \subseteq \mathbb{R}^n$ *be a regular simplicial conical partition. Let* $0 < a_1, \ldots, a_{n+1} < 1$ *with* $\sum_{i=1}^{n+1} a_i = 1$*. Assume that* $(a_1, \ldots, a_{n+1}) \neq (1/(n+1), \ldots, 1/(n+1))$*.*

(a) If $\rho \in (0,1)$*, then there exists* $y \in \mathbb{R}^n$ *such that* $\gamma_n(B_i + y) = a_i$ *and such that*

$$\begin{aligned}
&\gamma_n(A_i) = a_i,\ \forall\, i = 1,\ldots,k \quad \wedge \quad \cup_{i=1}^k A_i = \mathbb{R}^n \\
&\qquad \Longrightarrow \quad \sum_{i=1}^k \int_{\mathbb{R}^n} 1_{A_i}\cdot T_\rho 1_{A_i} d\gamma_n \leq \sum_{i=1}^k \int_{\mathbb{R}^n} 1_{B_i+y}\cdot T_\rho 1_{B_i+y} d\gamma_n.
\end{aligned} \tag{1.14}$$

(b) If $\rho \in (-1,0)$ *then there exists* $y \in \mathbb{R}^n$ *such that* $\gamma_n(B_i + y) = a_i$ *and such that*

$$\begin{aligned}
&\gamma_n(A_i) = a_i,\ \forall\, i = 1,\ldots,k \quad \wedge \quad \cup_{i=1}^k A_i = \mathbb{R}^n \\
&\qquad \Longrightarrow \quad \sum_{i=1}^k \int_{\mathbb{R}^n} 1_{A_i}\cdot T_\rho 1_{A_i} d\gamma_n \geq \sum_{i=1}^k \int_{\mathbb{R}^n} 1_{B_i+y}\cdot T_\rho 1_{B_i+y} d\gamma_n.
\end{aligned} \tag{1.15}$$

1.4 Grothendieck's Inequality and Kernel Clustering

Let $\{a_{ij}\}_{i,j=1}^n$ be a real matrix. Proven first in 1953 [38], **Grothendieck's Inequality** says that there is a constant $K > 0$ which does not depend on n or on $\{a_{ij}\}_{i,j=1}^n$ such that

$$\max_{w_1,\ldots,w_n,r_1,\ldots,r_n\in S^{2n-1}} \sum_{i,j=1}^n a_{ij}\langle w_i, r_j\rangle \le K \cdot \max_{\varepsilon_1,\ldots,\varepsilon_n,\delta_1,\ldots,\delta_n\in\{-1,1\}} \sum_{i,j=1}^n a_{ij}\varepsilon_i\delta_j. \tag{1.16}$$

Grothendieck had written his inequality in terms of different norms on tensor products of Banach spaces. It was actually Lindenstrauss and Pełczyński [68] who discretized this inequality in 1968 and wrote it as (1.16). Since Grothendieck's paper, it has remained an open (and difficult) question to determine the smallest possible K which can be used in (1.16).

In the case that the matrix $\{a_{ij}\}_{i,j=1}^n$ is symmetric positive semidefinite, we know the sharp constant in the corresponding Grothendieck Inequality, as shown in (1.17), which is a result of Rietz from 1974. Note that (1.17) looks more specific than (1.16) since (1.17) is quadratic rather than bilinear. However, if $\{a_{ij}\}_{i,j=1}^n$ is symmetric positive semidefinite, then both sides of (1.16) achieve their maxima when $w_i = r_i$ and $\varepsilon_i = \delta_i$ for all $i = 1,\ldots,n$ [2], so that (1.16) reduces to (1.17).

Theorem 1.4.1 (Positive Semidefinite Grothendieck Inequality). *[90] Let $\{a_{ij}\}_{i,j=1}^n$ be symmetric positive semidefinite, i.e. a real symmetric matrix with all eigenvalues nonnegative. Then*

$$\max_{w_1,\ldots,w_n\in S^{n-1}} \sum_{i,j=1}^n a_{ij}\langle w_i, w_j\rangle \le \frac{\pi}{2} \cdot \max_{\varepsilon_1,\ldots,\varepsilon_n\in\{-1,1\}} \sum_{i,j=1}^n a_{ij}\varepsilon_i\varepsilon_j. \tag{1.17}$$

Moreover, the constant $\pi/2$ is the smallest possible constant that can appear on the right side of (1.17).

The modern interpretation of (1.16) follows the same lines as the investigation of MAX-CUT by Goemans and Williamson [35] mentioned in Section 1.2. This interpretation came in 2006 from Alon and Naor [2]. In this interpretation, the left side of (1.16) corresponds to (1.7) and the right side of (1.16) corresponds to (1.6). The left side is an optimization problem which is linear in the inner products of vectors, so it is a semidefinite program which can be solved in polynomial time in n [100]. In general, the right side of (1.16) cannot be solved in time polynomial in n, assuming $P \neq NP$. Actually, we can write (1.16) in a way that makes it exactly analogous with the inequality (1.9), as follows. Until recently [19], the following statement of (1.16) was believed to produce the sharp constant in (1.16) [2].

Theorem 1.4.2 (Grothendieck's Inequality, Algorithmic Version). *[2] There is a randomized polynomial time algorithm such that, given $\{a_{ij}\}_{i,j=1}^n$ a real matrix and vectors $\{w_i\}_{i=1}^n, \{r_j\}_{j=1}^n \subseteq S^{2n-1}$, the algorithm outputs random signs $\{u_i\}_{i=1}^n, \{\nu_j\}_{j=1}^n \subseteq \{-1, 1\}$, such that*

$$\mathbb{E} \sum_{i,j=1}^n a_{ij} u_i \nu_j = \frac{2\log(1+\sqrt{2})}{\pi} \sum_{i,j=1}^n a_{ij} \langle w_i, r_j \rangle. \tag{1.18}$$

So, if we solve the semidefinite program corresponding to the left side of (1.16) and then apply Theorem 1.4.2 to these vectors, we have then approximately solved the right side of (1.16) within a multiplicative factor of about .56. That is, we first use the analogue of (1.8).

$$\max_{w_1,\ldots,w_n,r_1,\ldots,r_n \in S^{2n-1}} \sum_{i,j=1}^n a_{ij} \langle w_i, r_j \rangle \geq \max_{\varepsilon_1,\ldots,\varepsilon_n,\delta_1,\ldots,\delta_n \in \{-1,1\}} \sum_{i,j=1}^n a_{ij} \varepsilon_i \delta_j. \tag{1.19}$$

Then, combining (1.18) with (1.19) gives the analogue of (1.10).

$$\mathbb{E} \sum_{i,j=1}^n a_{ij} u_i \nu_j \geq \frac{2\log(1+\sqrt{2})}{\pi} \cdot \max_{\varepsilon_1,\ldots,\varepsilon_n,\delta_1,\ldots,\delta_n \in \{-1,1\}} \sum_{i,j=1}^n a_{ij} \varepsilon_i \delta_j. \tag{1.20}$$

Note that (1.18) differs from (1.9) since in (1.9) the weight function $w(i,j)$ is nonnegative, whereas the matrix $\{a_{ij}\}_{i,j=1}^n$ can have positive or negative entries. The nonnegativity of $w(i,j)$ was crucial in proving (1.9), so the same proof does not immediately work for (1.18). However, by first mapping the vectors w_i, r_j to new unit vectors in a way that removes the inverse cosine term of (1.9), we can then use a random hyperplane to "round" the resulting vectors to $\{-1, 1\}$ as in (1.9). And this procedure proves Theorem 1.4.2. For around forty years, it was thought that this two-step process of mapping the unit vectors to new vectors and then rounding with a random hyperplane was the best possible way to prove Grothendieck's Inequality. However, recently, it was shown that this procedure can be changed in a surprising way to produce a smaller constant K in (1.16). That is, we can take $K < \pi/(2\log(1+\sqrt{2}))$ in (1.16) [19]. Moreover, some "rounding procedure" can find the best constant in (1.16) [79].

Finding the smallest possible constant K in (1.16) has several interpretations beyond mathematics. From the physics perspective, the best constant K in (1.16) is also the best constant in certain Bell inequalities in quantum mechanics [87]. Also, assuming the Unique Games Conjecture, it is impossible to approximate the right side of (1.16), in time polynomial in n, within a multiplicative factor smaller than K [88], where K is the smallest possible constant in the inequality (1.16). One could consider this breadth of interpretation as evidence for the difficulty and importance of finding the smallest possible value of K in (1.16).

Since (1.16) is very general, and since rounding procedures appear often in theoretical computer science [35, 31, 2, 6, 66], inequality (1.16) has many applications. Actually, (1.16) and (1.9) describe a paradigm that has been very successful in solving many combinatorial optimization problems over the last two decades. Some combinatorial optimization problem is expressed as a maximization over variables with values $+1$ or -1, as in the right side of (1.16). Then, this maximization problem is changed by replacing the $\{-1,1\}$-valued variables with variables taking values in a high dimensional sphere, as in the left side of (1.16). The vectors maximizing the left side of (1.16) can be found efficiently, using semidefinite programming. These vectors are then "rounded" to the values $+1$ or -1 in such a way that the right side of (1.16) is almost maximized. And computing the right side of (1.16) cannot be done in a reasonable amount of time, as a function of n, assuming $P \neq NP$. So, an inequality such as (1.16) describes the most reasonable way of computationally approximating the right side of (1.16).

To summarize the state of affairs for finding the optimal constant in (1.16), it seems quite difficult to determine the best constant in (1.16), but we do know the best constant in (1.17). To find the best constant in (1.16), it would be helpful to have some similar problem of intermediate difficulty between (1.16) and (1.17). Khot and Naor found such a family of problems in [55, 57], which we now describe. In an effort to understand (1.16) better, and to approximate the optima of kernel clustering problems, Khot and Naor [55, 57] investigated Grothendieck's inequality (1.16) in 2009 for $\{a_{ij}\}_{i,j=1}^{n}$ that are positive semidefinite. We will first describe the kernel clustering problem, and we will then see how a version of Grothendieck's inequality arises from the clustering problem. The kernel clustering problem from machine learning is defined as follows.

Definition 1.4.3 (**Kernel Clustering Problem**). Let $1 \leq k \leq n$. Let $\{a_{ij}\}_{i,j=1}^{n}$ be a symmetric real positive semidefinite matrix with $\sum_{i,j=1}^{n} a_{ij} = 0$. Let $\{b_{ij}\}_{i,j=1}^{k}$ be another symmetric real positive semidefinite matrix. The goal of the kernel clustering problem is to find the following quantity

$$\max_{\substack{\{S_1,\ldots,S_k\}\text{ a partition}\\ \text{of }\{1,\ldots,n\}}} \sum_{i,j=1}^{k} \left(\sum_{(p,q)\in S_i\times S_j} a_{pq} \right) b_{ij}.$$

Here we think of the matrix $\{a_{ij}\}_{i,j=1}^{n}$ as a matrix containing a large amount of data, and $\{b_{ij}\}_{i,j=1}^{k}$ is some hypothesis expressing that $\{a_{ij}\}_{i,j=1}^{n}$ should be able to be clustered into a small number of blocks (k^2 of them). And the data in each block should be similar. Definition 1.4.3 is therefore a variant of the popular clustering paradigm in machine learning. For example, the similar k-means clustering problem and algorithms for it were investigated as far back as 1957 by Lloyd [70] with renewed interest in 1967 [72] and the late 1990s [28].

However, as with many machine learning algorithms, rigorous analysis remains elusive. Interestingly, the analysis of the k-means clustering algorithm is similar to the analysis of the Kernel Clustering Problem from Definition 1.4.3. In particular, in both problems, a gradient-descent method amounts to iteratively constructing Voronoi tesselations of space, and it is unclear (and unlikely) that this procedure will converge to an optimal value. We now summarize the problems with machine learning algorithms and the problems that we discuss below. The typical situation is that we have some nonconvex optimization problem, and we want to find an optimal point. Of course, one can start at any given point, take the gradient of the given function, and then move in the direction of the gradient to try to optimize the given function. However, there is no guarantee of achieving a global optimum with this procedure.

Definition 1.4.4 (k-Means Clustering). Let $1 \leq k \leq n$. Let $a_1, \ldots, a_n \in \mathbb{R}^d$. The goal of the k-means clustering problem is to find the following quantity

$$\min_{\substack{\{S_1,\ldots,S_k\} \text{ a partition} \\ \text{of } \{1,\ldots,n\}}} \sum_{i=1}^{k} \sum_{j \in S_i} \left\| a_j - \frac{1}{|S_i|} \sum_{\ell \in S_i} a_\ell \right\|_2^2$$

There are many algorithms and analyses related to clustering problems. We only name a few, which include the nonnegative matrix factorization problem [15], and the sparsest k-partition problem [71]. The latter was influenced by the investigation into higher order Cheeger inequalities, as in [66]

As stated above, the sharp constant is known in Theorem 1.4.1, which is a result of Rietz [90]. Khot and Naor then created a more general positive semidefinite Grothendieck inequality, and they made a conjecture concerning its optimal constant for certain simple cases. Even in these simple cases, we still do not know what the optimal constant for the inequality is, though we have a guess for what it should be. So, we can therefore consider this new generalized positive semidefinite Grothendieck inequality to be of intermediate difficulty, between the "easier" Theorem 1.4.1 and the "harder" (1.16). Surprisingly, Theorem 1.4.5 yields the optimal approximation algorithm for the kernel clustering problem, assuming the Unique Games Conjecture.

Theorem 1.4.5 (The Generalized Positive Semidefinite Grothendieck Inequality). *[57, Theorem 3.1, Theorem 3.3] Let $\{a_{ij}\}_{i,j=1}^{n}$ be a positive semidefinite matrix, i.e. a real symmetric matrix with all eigenvalues nonnegative. Let $\nu_1, \ldots, \nu_k \in \mathbb{R}^k$ with $k \geq 2$, and let $B = \{b_{ij}\}_{i,j=1}^{k}$ be the symmetric positive semidefinite matrix with $b_{ij} := \langle \nu_i, \nu_j \rangle$. Then there exists $C(B)$ which does not* *depend on n or on $\{a_{ij}\}_{i,j=1}^n$ such that*

$$\max_{w_1,\ldots,w_n\in S^{n-1}} \sum_{i,j=1}^n a_{ij}\langle w_i, w_j\rangle \leq \frac{1}{C(B)}\cdot \max_{\sigma\colon \{1,2,\ldots,n\}\to\{1,2,\ldots,k\}} \sum_{i,j=1}^n a_{ij}\langle \nu_{\sigma(i)}, \nu_{\sigma(j)}\rangle. \tag{1.21}$$

Instead of rounding the vectors w_i, $i=1,\ldots,n$ to 1 or -1 as in Grothendieck's inequality, Khot and Naor prove (1.21) by rounding the vectors w_i, $i=1,\ldots,n$ to one of k vectors ν_i, $i=1,\ldots,k$. Also, the best constant $1/C(B)$ is found by finding the best rounding procedure. As in the MAX-k-CUT problem, finding the best rounding procedure in Theorem 1.4.5 reduces to finding a partition of Euclidean space that maximizes a certain isoperimetric quantity.

Let I_k denote the $k\times k$ identity matrix. In [55, 57], it is shown that $C(B)$ can be found by solving a finite-dimensional optimization problem, whose parameters depend on B. However, this optimization problem is non-convex in general, so standard methods cannot compute $C(B)$. As a special case,

$$C(I_k) = \sup_{\substack{A_1,\ldots,A_k\colon \cup_{i=1}^k A_i=\mathbb{R}^{k-1},\\ \gamma_{k-1}(A_i\cap A_j)=0,\,\forall\, i,j\in\{1,\ldots,k\},\, i\neq j}} \sum_{i=1}^k \left\| \int_{A_i} x d\gamma_{k-1}(x)\right\|_2^2.$$

The **Propeller Conjecture** guesses the value of $C(I_k)$, $k\geq 4$, in (1.21)

Conjecture 4 (**Propeller Conjecture**). *[55, 57] The following constant is the best possible in* (1.21) *for $B=I_k$ and $k\geq 4$.*

$$C(I_k) = \frac{9}{8\pi} = C(I_3). \tag{1.22}$$

The constant $C(I_3)=9/(8\pi)$ is computed in [55] using Lagrange multipliers, but $C(I_4)$ does not appear to be computable using this technique. Using some theoretical results and a brute force search, we give a computer-assisted proof of (1.22) in the case $k=4$ [46] in Chapter 3. The analytic results use a connection between the maximization problem (1.22) and discrete harmonic maps into the sphere. Intuition derived from this connection allows several ad hoc arguments to rule out candidates for partitions that maximize (1.22). Removing the computer-assistance and treating $k>4$ remains a challenge.

As an aside, it would be nice to adapt the results of [55] and [57] to the noncommutative Grothendieck inequality. Recent work [10] finds an improved constant in a semidefinite special case of the noncommutative Grothendieck inequality of [80]. However, the results of [10] do not guarantee the sharpness of their inequality. It would be desirable to look for a sharp constant in their positive semidefinite noncommutative Grothendieck inequality, and to prove a corresponding Unique Games hardness result.

Lastly, the Propeller Conjecture (1.22) is connected to the Standard Simplex Conjecture 1 by the following formula for $A \subseteq \mathbb{R}^n$, which is discussed in Chapter 4, (4.13). Thus, from (1.23), Chapters 3, 4 and 5 deal with essentially the same quantity.

$$\frac{d}{d\rho}\bigg|_{\rho=0} \int_{\mathbb{R}^n} 1_A \cdot T_\rho 1_A d\gamma_n = \left\| \int_A x d\gamma_n(x) \right\|_2^2. \tag{1.23}$$

1.5 Summary of Main Results

In this thesis we show the following things.

- The Propeller Conjecture (1.22) holds for $k = 4$. (Chapter 3)
- For all $n \geq 2$, there exists $\rho_0(n) > 0$ such that the Standard Simplex Conjecture 1 in dimension n holds for all $0 < \rho < \rho_0(n)$. (Chapter 4)
- The Shifted Standard Simplex Conjecture 3 does not hold. (Chapter 5)

Chapter 3 is derived from the paper [46], which is joint work with Aukosh Jagannath and Assaf Naor. The main idea there is to use a computer-assisted ε-net argument together with modulus of continuity bounds to determine that many partitions are not optimal. The remaining partitions are dealt with using derivative bounds and modulus of continuity bounds. Chapter 4 is derived from the paper [45]. The main idea there is an iterative argument to repeatedly improve the regions of a partition to look more and more like what we want. Iterative arguments appear many places in mathematics, e.g. Picard iteration, the Nash embedding theorem [81], PDE methods, or De Giorgi-Nash-Moser iteration [76]. Also a more contemporary result is the investigation of Krivine's functional in dimension 1 [19, Theorem 1.3]. Chapter (5) is derived from the paper [47], which is joint work with Elchanan Mossel and Joe Neeman [47]. There, an analyticity argument shows that the first variation of many natural Euclidean partitions cannot be zero, and therefore these partitions do not optimize noise stability.

1.6 A Final Comment on Intuitive Reasoning for Conjectures 4 and 1

As mentioned at the beginning of the introduction, we believe that Conjectures 4 and 1 hold since the Gaussian measure is invariant under Euclidean rotations. And these conjectures are also invariant under Euclidean rotations, and under permutation of the partition elements. So, we expect that the solution of these problems should also be symmetric and simple to describe. However, since Conjecture 3 is false, any proof of Conjecture 4 or 1 must also fail to prove Conjecture 3. But the existing proof strategies for Theorem 1.1.6 [77, 29] do not seem to make such a distinction when we try to apply them to solve Conjectures 4 and 1. Moreover, the finite dimensional optimization problem corresponding to Conjecture 1 contains at least one saddle point. We therefore infer that both Conjectures 1 and 4 cannot be solved by first variation arguments or gradient-descent methods, such as heat flow methods. In the ensuing investigation, we therefore attempt to find new and different approaches to Conjectures 1 and 4 that circumvent these issues.

Chapter 2

Preliminaries

2.1 Basic Notation

Let n be a positive integer. Let $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, with $x_i \in \mathbb{R}$ for all $i \in \{1, \ldots, n\}$. Let $r > 0$, let $A, B \subseteq \mathbb{R}^n$, let $f\colon \mathbb{R}^n \to \mathbb{R}^m$, $f_i\colon \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, m$, $f = (f_1, \ldots, f_m)$. Assume that f is differentiable and $g\colon \mathbb{R}^n \to \mathbb{R}$ is differentiable. Let $1 \leq p \leq \infty$, and let (Ω, μ) be a measure space with measure μ.

$$
\begin{aligned}
\mathbb{Z} &:= \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}, \text{ the integers}\\
\mathbb{N} &:= \{0, 1, 2, 3, 4, 5, \ldots\}, \text{ the natural numbers}\\
\mathbb{Q} &:= \{m/n\colon m, n \in \mathbb{Z}, n \neq 0\}, \text{ the rationals}\\
\mathbb{R} &\text{ denotes the set of real numbers}\\
\mathbb{C} &:= \{x + y\sqrt{-1} : x, y \in \mathbb{R}\}, \text{ the complex numbers}\\
\varnothing &\text{ denotes the empty set}\\
A \setminus B &:= \{x \in A\colon x \notin B\}\\
A^c &:= \mathbb{R}^n \setminus A, \text{ the complement of } A\\
A \cap B &\text{ denotes the intersection of } A \text{ and } B\\
A \cup B &\text{ denotes the union of } A \text{ and } B\\
A \Delta B &:= (A \setminus B) \cup (B \setminus A) \text{ the symmetric difference of } A \text{ and } B
\end{aligned}
$$

max (or min) denotes the maximum (or minimum) of a set of numbers

sup (or inf) denotes the supremum (or infimum) of a set of numbers

$\|x\|_{\ell_p^n} = \|x\|_p := (\sum_{i=1}^n |x_i|^p)^{1/p}$, the p-norm, or ℓ_p-norm for $1 \leq p < \infty$

$\|x\|_{\ell_\infty^n} := \sup_{i\in\{1,\ldots,n\}} |x_i|$, the sup-norm, or ∞-norm, or ℓ_∞ norm

$\langle x, y\rangle := \sum_{i=1}^n x_i y_i,$ the standard inner product on $\mathbb{R}^n$

$B_p(x,r) := \{y \in \mathbb{R}^n \colon \|y-x\|_{\ell_p^n} < r\}$, the open ℓ_p^n ball of radius r

$B(x,r) = B_2(x,r) := \{y \in \mathbb{R}^n \colon \|y-x\|_{\ell_2^n} < r\}$, the open ball of radius r

$\|f\|_{L_p(\mu)} := (\int_\Omega |f|^p\, d\mu)^{1/p},$ the $L_p(\mu)$ norm for $1 \leq p < \infty$

$\|f\|_{L_\infty(\mu)} := \inf\{\lambda > 0 \colon \mu(\omega \in \Omega \colon |f(\omega)| > \lambda) = 0\},$ the $L_\infty(\mu)$ norm

$L_p(\mu) := \{f \colon \Omega \to \mathbb{R}, \|f\|_{L_p(\mu)} < \infty\}$

det denotes the determinant function

$id = id_n = I_n$ denotes the identify map, or the $n\times n$ identity matrix

$e_1, \ldots, e_n$ denotes the standard basis of $\mathbb{R}^n$

$\partial g/\partial x_j = D_{e_j} g = D_j g$ denotes the partial derivative of g with respect to x_j

$\nabla g(x) = (D_1 g(x), \ldots, D_n g(x))$ denotes the gradient of g

$\mathrm{Hessian}(g) = (D_{ij} g)_{1\leq i,j\leq n},$ the Hessian matrix of second partial derivatives

$\gamma_n(x) := e^{-\|x\|_2^2/2}(2\pi)^{-n/2},$ the Gaussian density on $\mathbb{R}^n$

$\mathbb{P}$ denotes a probability measure

$\mathbb{E}$ denotes the expected value

$\Delta := \sum_{i=1}^n \partial^2/\partial x_i^2$ denotes the Laplacian on $\mathbb{R}^n$

$L := -\Delta + \sum_{i=1}^n x_i \cdot \partial/\partial x_i$, the Ornstein-Uhlenbeck operator on $\mathbb{R}^n$

$T_\rho f(x) := \int_{\mathbb{R}^n} f(x\rho + y\sqrt{1-\rho^2}\,) d\gamma_n(y), \quad \rho \in (-1,1).$

$(x_1,x_2,x_3)\times(y_1,y_2,y_3) := (x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1)$ denotes the cross product on $\mathbb{R}^3$. For $\ell = (\ell_1,\ldots,\ell_n) \in \mathbb{N}^n$, let $|\ell| := \|\ell\|_1 = \sum_{i=1}^n |\ell_i|$, and let $\ell! := \prod_{i=1}^n (\ell_i!)$.

2.2 Orthonormality of Hermite Polynomials

Recall that the Hermite polynomials of one variable $h_0(x), h_1(x), \ldots$ are defined by the following generating function. For all $0 < \lambda < 1$, $x \in \mathbb{R}$, we have

$$\sum_{\ell=0}^{\infty}\lambda^{\ell}h_{\ell}(x)=e^{\lambda x-\lambda^2/2}=\sum_{p=0}^{\infty}\frac{x^p}{p!}\lambda^p\sum_{q=0}^{\infty}\frac{(-1)^q(\lambda)^{2q}(1/2)^q}{q!} \tag{2.1}$$

$$=\sum_{\ell=0}^{\infty}\lambda^{\ell}\sum_{k=0}^{\lfloor \ell/2\rfloor}\frac{x^{\ell-2k}(-1)^k2^{-k}}{k!(\ell-2k)!} \tag{2.2}$$

Here we let $p+2q=\ell$, $k=q$. In particular, equating the terms with the same coefficients of λ^{ℓ} for a fixed $\ell\in\mathbb{N}$, we get the following explicit formula for the Hermite polynomials.

$$h_{\ell}(x)=\sum_{k=0}^{\lfloor \ell/2\rfloor}\frac{x^{\ell-2k}(-1)^k2^{-k}}{k!(\ell-2k)!}. \tag{2.3}$$

So, $h_0=1$, $h_1=x$, $h_2=(1/2)x^2-(1/2)$, $h_3=(1/6)x^3-(1/2)x$. Since $\|\sqrt{\ell!}h_{\ell}\|_{L_2(\gamma_1)}=1$, we have $1=\|x\|_{L_2(\gamma_1)}=\|(\sqrt{2}/2)(x^2-1)\|_{L_2(\gamma_1)}$. We will now derive the formula $\|\sqrt{\ell!}h_{\ell}\|_{L_2(\gamma_1)}=1$.

Recall that $e^{-\pi\xi^2}=\int_{\mathbb{R}}e^{-\pi x^2}e^{-2\pi i x\xi}dx$. So, $e^{-\xi^2/2}=\int_{\mathbb{R}}e^{-x^2/2}e^{-ix\xi}dx/\sqrt{2\pi}$. So, we write

$$e^{\lambda x-\lambda^2/2}=e^{x^2/2}e^{-(1/2)(x-\lambda)^2}=e^{x^2/2}\int_{\mathbb{R}}e^{-y^2/2}e^{-iy(x-\lambda)}dy/\sqrt{2\pi}$$

$$=e^{x^2/2}\int_{\mathbb{R}}e^{-y^2/2}e^{-iyx}e^{iy\lambda}dy/\sqrt{2\pi}. \tag{2.4}$$

Now, $e^{iy\lambda}=\sum_{\ell=0}^{\infty}(iy\lambda)^{\ell}/\ell!$, so

$$\sum_{\ell=0}^{\infty}h_{\ell}(x)\lambda^{\ell}=e^{\lambda x-\lambda^2/2}\overset{(2.4)}{=}\sum_{\ell=0}^{\infty}\frac{\lambda^{\ell}}{\ell!}e^{x^2/2}\int_{\mathbb{R}}e^{-y^2/2}(iy)^{\ell}e^{-iyx}dy/\sqrt{2\pi}.$$

Therefore,

$$h_{\ell}(x)=\frac{1}{\ell!}e^{x^2/2}\int_{\mathbb{R}}e^{-y^2/2}(iy)^{\ell}e^{-iyx}dy/\sqrt{2\pi}. \tag{2.5}$$

Written another way,

$$h_{\ell}(x)=\frac{(-1)^{\ell}}{\ell!}e^{x^2/2}(d/dx)^{\ell}\int_{\mathbb{R}}e^{-y^2/2}e^{-iyx}dy/\sqrt{2\pi}=\frac{(-1)^{\ell}}{\ell!}e^{x^2/2}(d/dx)^{\ell}e^{-x^2/2}. \tag{2.6}$$

Suppose $j>k$, and consider $\int_{\mathbb{R}}h_j(x)h_k(x)e^{-x^2/2}dx/\sqrt{2\pi}$. By substituting the differential expression (2.6) and then integrating by parts k times, we get the following orthogonality condition.

$$\int_{\mathbb{R}}h_j(x)h_k(x)e^{-x^2/2}dx/\sqrt{2\pi}=\frac{(-1)^j}{j!}\int_{\mathbb{R}}[(d/dx)^je^{x^2/2}]h_k(x)dx/\sqrt{2\pi}=0. \tag{2.7}$$

To compute the norms of the Hermite polynomials, we square the generating function (2.1) and then integrate.

$$\begin{aligned}
\int_{\mathbb{R}} (\sum_{\ell=0}^{\infty} h_\ell(x)\lambda^\ell)^2 e^{-x^2/2}dx/\sqrt{2\pi} &= \sum_{\ell=0}^{\infty}\int_{\mathbb{R}} h_\ell^2(x)\lambda^{2\ell}e^{-x^2/2}dx/\sqrt{2\pi} \\
&= \int_{\mathbb{R}} e^{2\lambda x-\lambda^2}e^{-x^2/2}dx/\sqrt{2\pi} = \int_{\mathbb{R}} e^{-(1/2)(x-2\lambda)^2+\lambda^2}dx/\sqrt{2\pi} \\
&= e^{\lambda^2} = \sum_{\ell=0}^{\infty}\frac{\lambda^{2\ell}}{\ell!}.
\end{aligned}$$

So, equating the same powers of λ on both sides gives us the $L_2(\gamma_1)$ norms of the Hermite polynomials.

$$\|h_\ell\|_{L_2(\gamma_1)}^2 = \int_{\mathbb{R}} h_\ell^2 e^{-x^2/2}dx/\sqrt{2\pi} = 1/\ell!. \tag{2.8}$$

We can also get an explicit formula for the Gaussian heat kernel (or Mehler kernel) as follows.

$$\begin{aligned}
&e^{-(x^2+y^2)/2}\sum_{k=0}^{\infty} h_k(x)h_k(y)\cdot k!r^k \\
&\overset{(2.5)}{=} \sum_{k=0}^{\infty}\int_{\mathbb{R}}\int_{\mathbb{R}}\frac{(-uvr)^k}{k!}e^{-u^2/2-v^2/2-ixu-iyv}dudv(2\pi)^{-1} \\
&= \int_{\mathbb{R}} e^{-u^2/2-ixu}\int_{\mathbb{R}} e^{-v^2/2-iyv-uvr}dvdu(2\pi)^{-1} \\
&= \int_{\mathbb{R}} e^{-u^2/2-ixu+u^2r^2/2}\left[\int_{\mathbb{R}} e^{-(v+ur)^2/2-iyv}dv\right]du(2\pi)^{-1} \\
&= \int_{\mathbb{R}} e^{-u^2/2-ixu+u^2r^2/2}\left[\int_{\mathbb{R}} e^{-v^2/2-iy(v-ur)}dv\right]du(2\pi)^{-1} \\
&= \int_{\mathbb{R}} e^{-u^2/2-ixu+u^2r^2/2+iyur}\left[\int_{\mathbb{R}} e^{-v^2/2-iyv}dv\right]du(2\pi)^{-1} \\
&= e^{-y^2/2}\int_{\mathbb{R}} e^{-u^2(1-r^2)/2-iu(x-yr)}du(2\pi)^{-1/2} \\
&= e^{-y^2/2}(1-r^2)^{-1/2}\int_{\mathbb{R}} e^{-u^2/2-i(1-r^2)^{-1/2}(x-yr)}du(2\pi)^{-1/2} \\
&= e^{-y^2/2}(1-r^2)^{-1/2}e^{-(x-yr)^2/[2(1-r^2)]} \\
&= (1-r^2)^{-1/2}e^{-y^2/2-x^2/[2(1-r^2)]-y^2r^2/[2(1-r^2)]+xyr/(1-r^2)} \\
&= (1-r^2)^{-1/2}e^{-\frac{1}{2}\frac{1+r^2}{1-r^2}(x^2+y^2-2rxy)}.
\end{aligned}$$

From (2.7) and (2.8), we know that $\{h_k\sqrt{k!}\}_{k=0}^{\infty}$ is an orthonormal set in $L_2(\gamma_1)$. To see that $\{h_k\sqrt{k!}\}_{k=0}^{\infty}$ is a basis for $L_2(\gamma_1)$, note that from (2.3), the span of $\{h_k\sqrt{k!}\}_{k=0}^{\infty}$ in $L_2(\gamma_1)$ is equal to the span of monomials in $L_2(\gamma_1)$. So, to prove that $\{h_k\sqrt{k!}\}_{k=0}^{\infty}$ is an orthonormal basis of $\{h_k\sqrt{k!}\}_{k=0}^{\infty}$ it suffices to let $f \in L_2(\gamma_1)$ with $\int_{\mathbb{R}} f(x)x^k d\gamma_1(x) = 0$ for all $k \in \mathbb{N}$, and then prove that $f = 0$. To this end, let $z \in \mathbb{C}$ and define

$$g(z) := \int_{\mathbb{R}} f(x)e^{2\pi zx}e^{-x^2/2}dx = \sum_{k\in\mathbb{N}} \frac{(2\pi)^k z^k}{k!} \int_{\mathbb{R}} f(x)x^k e^{-x^2/2}dx = 0.$$

Then for all $t \in \mathbb{R}$, $g(it) = 0$, so that the Fourier transform of $f(x)e^{-x^2/2}$ is zero. Note that $f \in L_2(\gamma_1)$ implies $f \in L_1(\gamma_1)$, so $f(x)e^{-x^2/2} \in L_1(\mathbb{R})$. So, by Fourier inversion, we conclude that $f(x)e^{-x^2/2} = 0$, and $f = 0$ as desired. In conclusion, $\{h_k\sqrt{k!}\}_{k=0}^{\infty}$ is an orthonormal basis of $L_2(\gamma_1)$.

Let $\ell = (\ell_1, \ldots \ell_n) \in \mathbb{N}^n$ and for $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, define

$$h_\ell(x) := \prod_{i=1}^{n} h_{\ell_i}(x_i).$$

For $\ell \in \mathbb{N}^n$, define $\ell! := \prod_{i=1}^{n}(\ell_i!)$. Since $\gamma_n(x) = \prod_{i=1}^{n} \gamma_1(x_1)$, we conclude that

$$\{h_\ell\sqrt{\ell!}\}_{\ell\in\mathbb{N}^n} \text{ is an orthonormal basis of } L_2(\gamma_n). \tag{2.9}$$

Taking n products of this identity,

$$\sum_{k=0}^{\infty} r^k \sum_{\substack{\ell\in\mathbb{N}^n:\\ \|\ell\|_1=k}} h_\ell(x)h_\ell(y)\ell! = (1-r^2)^{-n/2} e^{-\frac{1}{2}\frac{1+r^2}{1-r^2}(\|x\|_2^2+\|y\|_2^2-2r\langle x,y\rangle)}. \tag{2.10}$$

2.3 P and NP

Let $\mathcal{L} \subseteq \cup_{m=1}^{\infty}\{0,1\}^m$. We say that an algorithm $\mathrm{ALG}\colon \cup_{m=1}^{\infty}\{0,1\}^m \to \{0,1\}$ decides $\mathcal{L}$ if $\mathrm{ALG}(x) = 1$ for all $x \in \mathcal{L}$, and $\mathrm{ALG}(x) = 0$ for all $x \notin \mathcal{L}$. We say that ALG decides $\mathcal{L}$ in $T(n) > 0$ time if, for each $n \in \mathbb{N}$, for all $x \in \{0,1\}^n$, $\mathrm{ALG}(x)$ runs in time at most $T(n)$. Define the complexity class P as the set of all $\mathcal{L} \subseteq \cup_{m=1}^{\infty}\{0,1\}^m$ such that, for each $\mathcal{L}$ there exists an algorithm ALG and there exists a polynomial $T(n)$ such that ALG decides $\mathcal{L}$ in time at most $T(n)$. Define the complexity class NP as the set of all $\mathcal{L} \subseteq \cup_{n=1}^{\infty}\{0,1\}^n$ such that, for each $\mathcal{L}$ there exists an algorithm

$$V\colon (\cup_{m=1}^{\infty}\{0,1\}^m) \times (\cup_{m=1}^{\infty}\{0,1\}^m) \to \{0,1\},$$

and there exists a polynomial $T(n)$ such that, for all $x \in \{0,1\}^n$ with $x \in \mathcal{L}$, there exists $y \in \cup_{m=1}^{\infty}\{0,1\}^m$ such that $V(x,y) = 1$, and $V(x,y)$ runs in time at most $T(n)$. Moreover, for any $x \in \{0,1\}^n$ with $x \notin \mathcal{L}$, for all $y \in \cup_{m=1}^{\infty}\{0,1\}^m$, $V(x,y) = 0$ and $V(x,y)$ runs in time at most $T(n)$.

If $\mathcal{L} \in P$, then we are given $\mathrm{ALG}(x)$, and we then define $V(x,y) := \mathrm{ALG}(x)$ to see that $P \subseteq NP$. It is currently unknown if $P = NP$ or if $P \neq NP$.

Intuitively speaking, P represents the set of problems that can be solved in a reasonable (i.e. polynomial) amount of time, and NP represents the set of problems whose solution can be verified in a reasonable (i.e. polynomial) amount of time.

Chapter 3

The Propeller Conjecture in $\mathbb{R}^3$

This chapter is derived from the paper [46], which is joint work with Aukosh Jagannath and Assaf Naor.

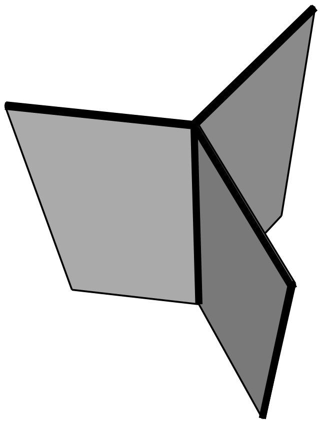


Figure 3.1: *The partition of* $\mathbb{R}^3$ *that maximizes the sum of squared lengths of Gaussian moments is a "propeller": three planar* $120°$ *sectors multiplied by an orthogonal line, with the rest of the partition elements being empty.*

3.1 Introduction

The following theorem is our main result, asserting that among all measurable partitions of $\mathbb{R}^3$, the "propeller partition" as depicted in Figure 3.1 maximizes the sum of the squared lengths of the Gaussian moments associated to the members of the partition.

Theorem 3.1.1 (Main theorem; geometric formulation)**.** *Let $\{A_1, \ldots, A_k\}$ be a partition of $\mathbb{R}^3$ into Lebesgue measurable sets. For $i \in \{1, \ldots, k\}$ let $\zeta_i = \int_{A_i} x d\gamma_3(x) \in \mathbb{R}^3$ be the Gaussian moment of the set A_i. Then*

$$\sum_{i=1}^{k} \|\zeta_i\|_2^2 \leqslant \frac{9}{8\pi}. \tag{3.1}$$

Let $\{P_1, P_2, P_3\}$ be the partition of $\mathbb{R}^2$ into $120°$ sectors centered at the origin. The bound (3.1) cannot be improved, with equality holding if $A_i = P_i \times \mathbb{R}$ for $i \in \{1, 2, 3\}$ and $A_i = \varnothing$ for $i \in \{4, \ldots, k\}$ (up to measure zero corrections, orthogonal transformations, and renumbering of the sets $\{A_1, \ldots, A_k\}$).

The Propeller Conjecture was posed by Khot and Naor in [55] as part of their investigation of the computational complexity of the Kernel Clustering problem from Machine Learning. Specifically, they conjectured the validity of the bound (3.1) for measurable partitions of $\mathbb{R}^m$ for all $m > 2$ and all k. They proved this conjecture for $m = 2$, and for $m = 1$ they showed that the right hand side of (3.1) can be improved to the sharp bound $\frac{1}{\pi}$. It is also shown in [55] that proving (3.1) for $k = 4$ and $m = 3$ implies the same conclusion for all $k \in \mathbb{N}$ and $m = 3$. Correspondingly, for partitions of $\mathbb{R}^m$ it suffices to prove this statement for $k = m + 1$.

Before giving the complexity-theoretic consequence of Theorem 3.1.1, which was the motivation of [55] for posing the Propeller Conjecture, we state two equivalent formulations of it: the first probabilistic (a sharp estimate for the expected maximum of a Gaussian vector) and the second analytic (a sharp Grothendieck inequality). The equivalence of these results to Theorem 3.1.1 was established in [55]; for Theorem 3.1.2 below see [55, Lem. 3.7] and for Theorem 3.1.3 below see [55, Thm. 1.1].

Theorem 3.1.2 (Main theorem; probabilistic formulation)**.** *Let $(g_1, g_2, g_3, g_4) \in \mathbb{R}^4$ be a mean zero Gaussian vector (with arbitrary covariance matrix). then*

$$\left| \mathbb{E}\left[\max_{i \in \{1,2,3,4\}} g_i \right] \right| \leqslant \frac{3}{2\sqrt{2\pi}} \sqrt{\sum_{i=1}^{4} \mathbb{E}\left[g_i^2\right]}. \tag{3.2}$$

The bound (3.2) *cannot be improved, with equality holding when the covariance matrix of* (g_1, g_2, g_3, g_4) *equals*

$$\frac{1}{2}\begin{pmatrix} 2 & -1 & -1 & 0 \\ -1 & 2 & -1 & 0 \\ -1 & -1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The following analytic formulation of Theorem 3.1.1 is in terms of a sharp Grothendieck inequality. Grothendieck [38] pioneered in 1953 the use of inequalities of this type in functional analysis, and ever since then such inequalities have permeated many mathematical disciplines. See [55, 57, 56] for an explanation of how the result below relates to a natural extension of the Grothendieck inequality and to approximation algorithms.

Theorem 3.1.3 (Main theorem; analytic formulation)**.** *Let* (a_{ij}) *be an* $n \times n$ *positive semidefinite matrix with* $\sum_{i=1}^n \sum_{j=1}^n a_{ij} = 0$*. For every* $\{v_1, v_2, v_3, v_4\} \subseteq S^3$ *with* $\sum_{i=1}^4 v_i = 0$ *we have*

$$\max_{x_1,\dots,x_n \in S^{n-1}} \sum_{i=1}^n \sum_{j=1}^n a_{ij} \langle x_i, x_j \rangle \leqslant \frac{2\pi}{3} \max_{y_1,\dots,y_n \in \{v_1,v_2,v_3,v_4\}} \sum_{i=1}^n \sum_{j=1}^n a_{ij} \langle y_i, y_j \rangle. \qquad (3.3)$$

The bound (3.3) *cannot be improved, with asymptotic equality holding when the vectors* $\{v_1, v_2, v_3, v_4\}$ *are the rows of the following matrix.*

$$\frac{1}{2\sqrt{3}}\begin{pmatrix} 3 & -1 & -1 & -1 \\ -1 & 3 & -1 & -1 \\ -1 & -1 & 3 & -1 \\ -1 & -1 & -1 & 3 \end{pmatrix}.$$

See [57, Sec. 3.1] for a description of a family of $n \times n$ *matrices* (a_{ij}) *for which equality in* (3.3) *is asymptotically attained (as* $n \to \infty$*).*

Inequality (3.3) was proved in [55]. Our new contribution is the assertion that (3.3) cannot be improved, a statement that is equivalent to inequality (3.1) of Theorem 3.1.1. This is not the first time that attempts to prove sharpness of a Grothendieck inequality led to an extremal geometric partitioning question. Notably, see König's conjecture [61] as a step towards Krivine's conjecture [63] on the sharpness of his version of the classical Grothendieck inequality. Unlike the Propeller Conjecture in $\mathbb{R}^3$, these conjectures turned out to be false [19], but they do indicate the interconnection between Grothendieck inequalities and extremal geometric partitioning problems.

The main consequence (and motivation) of Theorem 3.1.1 is complexity theoretic. To explain it we briefly recall Khot's Unique Games Conjecture (UGC), which asserts that for every $\varepsilon \in (0,1)$ there exists a prime $p = p(\varepsilon) \in \mathbb{N}$ such that no polynomial time algorithm can perform the following task. The input is a system of m linear equations in n variables $x_1, \ldots, x_n$, each of which has the form $x_i - x_j \equiv c_{ij} \mod p$. The algorithm must determine whether there exists an assignment of an integer value to each variable x_i such that at least $(1-\varepsilon)m$ of the equations are satisfied, or no assignment of such values can satisfy more than εm of the equations. If neither of these possibilities occurs, then an arbitrary output of the algorithm is allowed. The UGC was introduced by Khot in [52], though the above formulation of it, which is equivalent to the original one, is due to [54]. The use of the UGC as a hardness hypothesis has become popular over the past decade; we refer to Khot's survey [53] for more information on this topic. Saying that the UGC hardness threshold of an optimization problem $\mathscr{O}$ equals $\alpha \in [1, \infty)$ means that for every $\varepsilon \in (0,1)$ there exists a polynomial time algorithm that outputs a number that is guaranteed to be within a multiplicative factor of $\alpha + \varepsilon$ from the solution of $\mathscr{O}$, and that the existence of such an algorithm with approximation guarantee of $\alpha - \varepsilon$ would contradict the UGC.

The Kernel Clustering problem is a clustering framework for covariance matrices that originated in the work of Borgwardt, Gretton, Smola and Song [92] in the context of Machine Learning. This problem is generic in the sense that it contains well-studied optimization problems as special cases, and its versatility allows one to design a variety of algorithms tailor-made for particular applications (many of these algorithms are at present shown [92] to be successful empirically, but not rigorously). The input of the Kernel Clustering problem is an $n \times n$ symmetric positive semidefinite matrix $A = (a_{ij})$ and a $k \times k$ symmetric positive semidefinite matrix $B = (b_{ij})$ called the hypothesis matrix. Think of n as very large and k as small, the goal being to cluster the entries of A into a $k \times k$ matrix that is most correlated with the hypothesis matrix B. Formally, given a partition $\{S_1, \ldots, S_k\}$ of $\{1, \ldots, n\}$, form the associated clustered version of A by summing the entries of A over the blocks induced by the partition $\{S_1, \ldots, S_k\}$. One thus obtains a $k \times k$ matrix $C = (c_{ij})$ given by $c_{ij} = \sum_{(s,t) \in S_i \times S_j} a_{st}$. Let $\mathbf{Clust}(A|B)$ denote the maximum of $\sum_{i=1}^k \sum_{j=1}^k c_{ij} b_{ij}$ over all partitions $\{S_1, \ldots, S_k\}$ of $\{1, \ldots, n\}$. We refer to [92, 55, 57, 56] for further explanation of this clustering framework, as well as a discussion of important special cases arising from appropriate choices of the hypothesis matrix B.

In what follows, an $n \times n$ symmetric positive semidefinite matrix $A = (a_{ij})$ is called centered if $\sum_{i=1}^n \sum_{j=1}^n a_{ij} = 0$, and it is called spherical if $a_{ii} = 1$ for all $i \in \{1, \ldots, n\}$.

Theorem 3.1.4 (Main theorem; complexity theoretic formulation)**.** *Let $\mathscr{O}$ be the following optimization problem. The input is an $n \times n$ symmetric positive semidef-*

inite centered matrix $A = (a_{ij})$, and also a 4×4 symmetric centered and spherical positive semidefinite matrix $B = (b_{ij})$. The goal is to compute in polynomial time the quantity $\mathbf{Clust}(A|B)$. *Then the UGC hardness threshold of $\mathscr{O}$ equals $\frac{2\pi}{3}$.*

A polynomial time algorithm with approximation ratio $\frac{2\pi}{3}+o(1)$ for the problem $\mathscr{O}$ of Theorem 3.1.4 was designed in [55]. The fact that Theorem 3.1.1 implies the matching UGC hardness result was also proved in [55]. More generally, it was shown in [55] that the validity of the m-dimensional Propeller Conjecture for some $m \geqslant 3$ would imply that the UGC hardness threshold of the variant of $\mathscr{O}$ with B being an $m \times m$ matrix equals $\frac{8\pi}{9}\left(1 - \frac{1}{m}\right)$.

This is not the first time that extremal problems in the measure space $(\mathbb{R}^m, \gamma_m)$ arose from investigations in complexity theory; see for example the Majority is Stablest Conjecture of Khot, Kindler, Mossel and O'Donnell [54], which was solved by Mossel, O'Donnell, and Oleszkiewicz [78] via a reduction to a classical isoperimetric inequality of Borell [17]. A special feature of the Propeller Conjecture is that its conclusion is quite surprising: despite allowing for partitions of $\mathbb{R}^m$ into $m+1$ sets, the optimal partition has only three nonempty sets. Note that this degeneracy property occurs for the first time when $m = 3$, since in the result of [55] for partitions of $\mathbb{R}^2$ into three sets, all three sets are indeed present in the optimal partition. Thus Theorem 3.1.1 is a proof of the first nonintuitive case of the Propeller Conjecture: a verification of a prediction about a clean yet unexpected geometric phenomenon that arose from an investigation in computational complexity. To the best of our knowledge this is the first time that such a development occurred in this theory.

Our proof of Theorem 3.1.1 proceeds as follows. First, using reductions of the Propeller Conjecture that were obtained in [55], we show that it suffices to prove an analogous partitioning problem for the sphere S^2. Namely, given a partition of S^2 into four spherical triangles, the goal is to show that the sum of the squares of lengths of the moments (with respect to the surface area measure on S^2) of these spherical triangles cannot exceed $9\pi^2/4$. Using additional geometric arguments, we further reduce the question to an optimization problem over a three dimensional search space. We thus obtain a coordinate system with three degrees of freedom, with respect to which one must solve a rather complicated nonlinear optimization problem. After proving several additional estimates that serve to further reduce the search space and give crucial modulus of continuity estimates, we show that it suffices to check that the desired estimate holds true for a finite list of spherical partitions (arising from an appropriate net of the search space). This list of inequalities that must be proved is explicit but very large, so we proceed to check it in a fully rigorous computer-assisted fashion. That is, our computation carefully accounts for all the rounding errors; equivalently the computation can be viewed as an implementation in our setting of interval arithmetic (see [42]).

The role and implications of rigorous computer-assisted proofs has been discussed at length in the literature; many such discussions appear in papers that establish striking results via a proof that has a computer-assisted component (famous examples include [3, 4, 43, 32, 40, 33]). We see no reason to include here a new treatment of this topic. Zwick's discussion in [103] is an excellent reference for a well thought out explanation of the role of computer-assisted proofs in mathematics and computer science[1]. The essence of the argument can be best conveyed by quoting Zwick directly [103, Sec. 7]:

> "A typical computer assisted proof, like ours, is essentially composed of two steps. The first step says something like: "Here is program P. If program P, when executed by an abstract computer, outputs YES, then the theorem is true, because . . . ". The second step says something like: "I ran program P and it said YES!", or more specifically: "I compiled program P using compiler A, under operating system B, and ran it on processor C, again under operating system B. It said YES!"."

Zwick then proceeds to explain that

> "The first step is a conventional mathematical proof. It simply argues that a certain program has certain properties. Such arguments are common in computer science. Anyone who objects to this step of the proof should find a flaw, or a gap, in the arguments made."

Nevertheless, Zwick explains that the second step, i.e., the claim that the compilation of the program gave the desired result, is not of the same nature:

> "The second step is more problematic. It is certainly not a proof in the conventional mathematical sense of the word. Many things can go wrong here. Program P may not have been compiled correctly by compiler A. Or, due to some bug in operating system B, the program did not run as intended. Or, processor C may suffer from a design flaw that causes an incorrect execution of the program, or perhaps the specific processor used has a short-circuit somewhere, and so on. It is possible to reduce the likelihood of such problems by compiling the program using several different compilers, and running it on different processors. But, while we may eventually be able to produce mathematical correctness proofs for compilers, operating systems, and the hardware design of the processors, it seems that we would never be

[1] Zwick's paper [103] is another instance of a computer-assisted proof in the context of approximation algorithms. While the present article deals with an integrality gap lower bound for a semindefinite program (the left hand side of (3.3)), Zwick proves an integrality gap upper bound, i.e., he uses a rigorous computer-assisted argument to show that a certain algorithm performs well rather than to prove a hardness result.

able to produce a mathematical proof that no hardware fault occurred during a specific computation.

What we can do, is instruct the computer executing program P to print a trace of the execution. This trace is a mathematical proof, though perhaps a not so inspiring one. Like any mathematical proof it should be checked carefully to make sure that it is correct. In many cases, however, this is not humanly possible. The main problem with computer assisted proofs, therefore, is that they are usually too long. They are also, in most cases, less insightful than conventional proofs."

The bulk of the work presented in this article consists of conventional geometric and analytic proofs, resulting in a reduction of the problem to a region where the desired inequality holds with "room to spare". This extra room allows us to complete the proof by checking a finite list of concrete inequalities (that we write explicitly) on a certain finite net. Each such inequality can be checked by hand, but the number of such checks is large, and it would be unrealistic (and probably unilluminating) to complete this final check manually. The code that we used is publicly available at the following link, which contains a command-line user interface so that interacting with the code is easy.

`http://cims.nyu.edu/~aukosh/propeller/propeller.html`

The code for this proof was run many times on several different computers. The technical requirements are quite modest, with the processor speed only affecting the total run time, and the total memory requirement being less than 200 Mb. One can run the code on anything from a dedicated compute server to a laptop computer (we did both). We ran the code in two different ways. The simplest way was to run it in a serial fashion. That is, we ran the program as a single instance which performed each step of the algorithm described in this paper one after the other. This took 37 hours to complete on a single processor of a compute server with 3.3 GhZ Intel Xenon processors. The second way we ran it was to parallelize the algorithm by breaking up the domain of interest into seven distinct portions and running separate instances on each. This procedure took eight hours when split between seven 3.3 GhZ Intel Xenon processors. (We also ran the program split between seven 2.4 GhZ AMD Opteron processors; this procedure took twelve hours.)

Needless to say, our proof of the Propeller Conjecture in $\mathbb{R}^3$ leaves something to be desired, since we do not have a short explanation of the validity of the computations in the final step of our proof. It is conceptually important to have a proof of this result, despite the fact that it ends in a lengthy computation, not only because it yields interesting results in mathematics and computational complexity. Importantly, since the Propeller Conjecture is a nonintuitive prediction that arose from investigations in computational complexity, knowing that the first nonintuitive case of this conjecture is indeed true puts the full Propeller Conjecture, and consequently the link between geometry and algorithms that it describes, on a stronger footing. Thus, in addition to yielding a remarkably involved UGC hardness result, our theorem will hopefully invigorate future research on this problem that might lead to a traditional mathematical proof of the Propeller Conjecture in $\mathbb{R}^3$ that can hopefully be extended to higher dimensions.

A key feature of the Propeller Conjecture is that it asserts that a natural optimization problem exhibits intermediate-dimensional symmetry breaking. There are precedents of results of this type that have been proved in the literature, mostly using Fourier-analytic methods. For this reason we are hopeful that a clean and shorter proof of the Propeller Conjecture will eventually be found. Consider for example the problem of sharp Khinchine inequalities: if $\varepsilon_1, \ldots, \varepsilon_n$ are i.i.d. symmetric random variables taking values in $\{-1, 1\}$ and $p \in [1, 2)$, the goal is to compute the minimum of $\mathbb{E}\left[|\sum_{i=1}^n a_i\varepsilon_i|^p\right]$ over all unit vectors $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$. For $p = 1$ it was conjectured by Littlewood (see [41]) that this minimum occurs at $a = (1, 1, 0, \ldots, 0)/\sqrt{2}$. Littlewood's conjecture was solved affirmatively by Szarek [98] (see also [99, 64]). Haagerup [39] proved that the same two dimensional symmetry breaking occurs for $p \in [1, p_0]$, where $p_0 = 1.87...$ is the solution of the equation $2\Gamma((p + 1)/2) = \sqrt{\pi}$, i.e., the unit vector that minimizes $\mathbb{E}\left[|\sum_{i=1}^n a_i\varepsilon_i|^p\right]$ equals $(1, 1, 0, \ldots, 0)/\sqrt{2}$ for $p \in [1, p_0]$ and for $p \in [p_0, 2]$ this minimum occurs at $a = (1, \ldots, 1)/\sqrt{n}$. Another famous result of this type is Ball's cube slicing theorem [8] (resolving a conjecture of Hensley [48]), asserting that the hyperplane section of $[-1, 1]^n$ with maximal $(n - 1)$-dimensional volume is perpendicular to $(1, 1, 0, \ldots, 0)$; see [83] for the corresponding result for complex scalars, as well as [9, 13] for an analogous statement for projections. We refer to [82] for a unified treatment of the results of Szarek, Haagerup, and Ball. There is also a conjecture of Milman predicting a symmetry breaking phenomenon for extremal volumes of slabs in the cube $[-1, 1]^n$; see [12, 62] for partial results along these lines. We mentioned the above statements since we believe that among the literature they are most similar to the Propeller Conjecture, and perhaps (with much more work) related methods could be be used to address the Propeller Conjecture as well.

This chapter is organized as follows. In Section 3.2 we present the reduction of the Propeller Conjecture to a certain optimization problem for spherical partitions, and we explain the main ingredients of our proof, including the conventional geometric/analytic arguments, as well as the computer-assisted component. The proofs of the geometric and analytic results leading to the final computational step are contained in Section 3.3 and Section 3.4. Section 3.4.1 contains a detailed explanation of how the numerical step is implemented so as to account for all possible rounding errors. We remark that our arguments extend mutatis mutandis to higher dimensions. For the sake of simplicity we present the entire argument in $\mathbb{R}^3$, since while it is conceivable that the scheme presented here can yield a computer-assisted proof of the Propeller Conjecture in higher dimensions, at some fixed dimension the computer-assisted component of the proof will become unfeasible. Now that we know that a nonintuitive case of the Propeller Conjecture is indeed correct, the next natural step is to search for a proof that extends to all dimensions, rather than attempting to prove a few more cases in low dimensions.

3.2 An overview of the proof of Theorem 3.1.1

From now on assume for the sake of eventually obtaining a contradiction that $\{A_i\}_{i=1}^4$ is a partition of $\mathbb{R}^3$ into measurable sets that violates the propeller conjecture, i.e.,

$$\sum_{i=1}^{4} \left\| \int_{A_i} x d\gamma_3(x) \right\|_2^2 > \frac{9}{8\pi}. \tag{3.4}$$

Assume moreover that the maximum of $\sum_{i=1}^4 \left\| \int_{B_i} x d\gamma_3(x) \right\|_2^2$ over all measurable partitions $\{B_i\}_{i=1}^4$ of $\mathbb{R}^3$ is attained at $\{A_i\}_{i=1}^4$. For a proof that this maximum is indeed attained, see [55, Lem. 3,1]. Using this maximality, it follows from [55, Lem. 3.3] that (up to measure zero corrections) the A_i are cones with cusp at the origin, and if we write $\zeta_i = \int_{A_i} x d\gamma_3(x)$ then $\zeta_1 + \zeta_2 + \zeta_3 + \zeta_4 = 0$ and for each $i \in \{1, 2, 3, 4\}$ we have

$$A_i = \left\{ x \in \mathbb{R}^3 : \ \langle x, \zeta_i \rangle = \max_{j \in \{1,2,3,4\}} \langle x, \zeta_j \rangle \right\}. \tag{3.5}$$

By [55, Lem. 3.3 & Cor. 3.4] the vectors $\{\zeta_i\}_{i=1}^4$ are distinct, nonzero and not coplanar.

Since the sets in question are cones, it is beneficial to study them in terms of their intersection with the sphere S^2, i.e., define

$$T_i = A_i \cap S^2, \tag{3.6}$$

and letting σ denote the surface area measure on S^2 (thus $\sigma(S^2) = 4\pi$), define

$$z_i = \int_{T_i} x d\sigma(x) = \sqrt{2\pi^3} \cdot \zeta_i, \tag{3.7}$$

where the last equality in (3.7) follows from integration in polar coordinates. Being a constant multiple of the vectors $\{\zeta_i\}_{i=1}^4$, the vectors $\{z_i\}_{i=1}^4$ are also distinct, nonzero, not coplanar, and satisfy $z_1 + z_2 + z_3 + z_4 = 0$. Moreover,

$$T_i \overset{(3.5)\wedge(3.6)\wedge(3.7)}{=} \left\{ x \in S^2 \colon \langle x, z_i \rangle = \max_{j \in \{1,2,3,4\}} \langle x, z_j \rangle \right\}. \tag{3.8}$$

Consequently,

$$\forall i,j \in \{1,2,3,4\}, \quad x \in T_i \cap T_j \implies \langle z_i, x\rangle = \langle z_j, x\rangle. \tag{3.9}$$

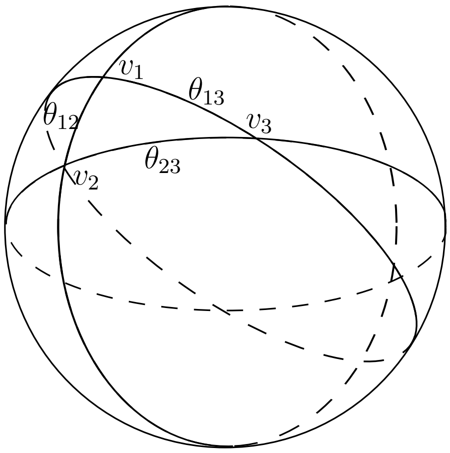
Fix $\ell \in \{1,2,3,4\}$. Since $\{z_i\}_{i=1}^4$ are not coplanar there exists a unique $v_\ell \in S^2$ satisfying

$$\forall i,j \in \{1,2,3,4\}\setminus\{\ell\}, \quad \langle z_i, v_\ell\rangle = \langle z_j, v_\ell\rangle > \langle z_\ell, v_\ell\rangle. \tag{3.10}$$

The vectors $\{v_\ell\}_{\ell=1}^4$ are the vertices of the partition $P = \{T_1, T_2, T_3, T_4\}$ of S^2. A simple argument presented in Section 3.3 shows that $v_i \notin \{-v_j, v_j\}$ if $i \neq j$ and each spherical triangle T_i is contained in an open hemisphere of S^2. Moreover, for distinct $i, j, \ell \in \{1,2,3,4\}$ we have $\det(v_i, v_j, v_\ell) \neq 0$, and if we define

$$\theta_{ij} := \arccos(\langle v_i, v_j\rangle) \qquad \text{and} \qquad \Theta_{ij\ell} := \arccos\left(\left\langle \frac{(v_i \times v_j)}{\|v_i \times v_j\|_2}, \frac{(v_j \times v_\ell)}{\|v_j \times v_\ell\|_2}\right\rangle\right), \tag{3.11}$$

then it is also argued in Section 3.3 that $\theta_{ij}, \Theta_{ij\ell} \in (0,\pi)$. Observe $\theta_{ij} = \theta_{ji}$, $\Theta_{ijk} = \Theta_{kji}$, and $\Theta_{ij\ell}$ is the spherical angle, at vertex v_j, of the spherical triangle with vertices $\{v_i, v_j, v_\ell\}$; the cosine of this angle is exactly the inner product of unit normals of the two planes containing $\{v_i, v_j, 0\}$ and $\{v_j, v_\ell, 0\}$ respectively. See Figure 3.2 for a schematic description of the notation.



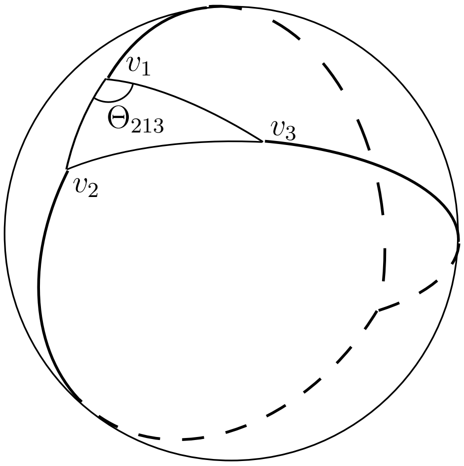


Figure 3.2: (a) Spherical triangle constructed with three great circles (b) Partition of the sphere into four spherical triangles.

In order to proceed we need to have a formula for the spherical moment z_4 in terms of the vertices $\{v_1, v_2, v_3\}$ of the spherical triangle T_4. In order to do so, assume that $\det(v_1, v_2, v_3) > 0$. Then

$$z_4 = \frac{1}{2}\left(\theta_{12}\frac{v_1 \times v_2}{\|v_1 \times v_2\|_2} + \theta_{23}\frac{v_2 \times v_3}{\|v_2 \times v_3\|_2} + \theta_{31}\frac{v_3 \times v_1}{\|v_3 \times v_1\|_2}\right). \tag{3.12}$$

Identity (3.12) was proved by Minchin in 1877; see [74, p. 259]. We also found identity (3.12) in [20], which is a military publication that is not publicly available (it isn't classified: we purchased access to it). Since both references for (3.12) are hard to find, we include a brief derivation of it in Proposition 3.3.2, using the Cauchy projection formula (see [60, p. 25]).

Using (3.12) and geometric arguments, we show in Lemma 3.3.5 that the vertex v_4 is determined by the vertices $\{v_1, v_2, v_3\}$ as follows; an analogous formula expresses each vertex of the partition in terms of the other three vertices.

$$v_4 = -\frac{1}{\sqrt{\lambda}}\left(\frac{\sin\theta_{23}}{\theta_{23}}v_1 + \frac{\sin\theta_{13}}{\theta_{13}}v_2 + \frac{\sin\theta_{12}}{\theta_{12}}v_3\right),$$

where

$$\begin{aligned}\lambda := \left\|\frac{\sin\theta_{23}}{\theta_{23}}v_1 + \frac{\sin\theta_{13}}{\theta_{13}}v_2 + \frac{\sin\theta_{12}}{\theta_{12}}v_3\right\|_2^2 = \frac{\sin^2\theta_{23}}{\theta_{23}^2} + \frac{\sin^2\theta_{13}}{\theta_{13}^2} + \frac{\sin^2\theta_{12}}{\theta_{12}^2} \\ + 2\frac{\sin\theta_{23}\sin\theta_{13}}{\theta_{23}\theta_{13}}\cos\theta_{12} + 2\frac{\sin\theta_{23}\sin\theta_{12}}{\theta_{23}\theta_{12}}\cos\theta_{13} + 2\frac{\sin\theta_{12}\sin\theta_{13}}{\theta_{12}\theta_{13}}\cos\theta_{23}.\end{aligned} \tag{3.13}$$

In Proposition 3.3.7 we derive the following restriction on products of opposite edges.

$$\frac{\sin\theta_{12}\sin\theta_{34}}{\theta_{12}\theta_{34}} = \frac{\sin\theta_{13}\sin\theta_{24}}{\theta_{13}\theta_{24}} = \frac{\sin\theta_{14}\sin\theta_{23}}{\theta_{14}\theta_{23}}.$$

Also, in Lemma 3.3.9 we obtain the following expression for the objective function at the partition $P = \{T_i\}_{i=1}^4$ as the sum of the squares of the angles between the vertices.

$$F(P) := \|z_1\|_2^2 + \|z_2\|_2^2 + \|z_3\|_2^2 + \|z_4\|_2^2 = \theta_{12}^2 + \theta_{13}^2 + \theta_{14}^2 + \theta_{23}^2 + \theta_{24}^2 + \theta_{34}^2. \tag{3.14}$$

Define

$$\gamma := \frac{\sin\theta_{23}}{\theta_{23}}\cos\theta_{12} + \frac{\sin\theta_{13}}{\theta_{13}} + \frac{\sin\theta_{12}}{\theta_{12}}\cos\theta_{23}. \tag{3.15}$$

In Lemma 3.3.8 we obtain the following useful restriction on a single spherical triangle appearing in the partition $\{T_1, T_2, T_3, T_4\}$.

$$\cos\left(\sqrt{\lambda-\gamma^2}\frac{\theta_{23}\theta_{12}}{\sin\theta_{23}\sin\theta_{12}}\right) = \cos\left(\sqrt{\theta_{12}^2 + \theta_{23}^2 + 2\theta_{12}\theta_{23}\cos\Theta_{123}}\right) = -\frac{\gamma}{\sqrt{\lambda}}, \tag{3.16}$$

where λ is given in (3.13).

Observe that (3.16) is a restriction involving data from only one triangle T_4, as we see by the definitions of γ and λ. Moreover, by symmetry, every cyclic permutation of $\{\theta_{12}, \theta_{23}, \theta_{13}\}$ must also satisfy (3.16). And, every triangle of P must satisfy (3.16), with the indices of the θ_{ij} substituted appropriately. To understand the ramifications of (3.16) assume that $\theta_{12} = \theta_{23} = \theta_{13} =: \theta$. Then (3.16) simplifies to

$$\cos\left(\sqrt{2(2\cos\theta+1)(1-\cos\theta)}\frac{\theta}{\sin\theta}\right) = -\frac{1}{\sqrt{3}}\sqrt{2\cos\theta+1}. \tag{3.17}$$

The only $\theta \in (0, 2\pi/3)$ solving (3.17) are $\theta = \arccos(-1/3) \approx 1.9106332362490187$, and $\theta \approx 1.5379684120790425$. The first solution corresponds to the partition whose vertices are those of a regular simplex inscribed in the sphere, and the second intuitively corresponds to a critical point "between" the regular partition and the propeller partition (note that $\theta = 2\pi/3$ implies $\det(v_1, v_2, v_3) = 0$, and $\theta > 2\pi/3$ cannot be achieved by a spherical triangle). The above special values of θ will play an important role in the ensuing computations.

As explained in the discussion following Lemma 3.3.9, a combination of (3.14) and (3.16) can be used to express the value of our objective function at the maximizing partition P in terms of the data from a single spherical triangle in P. Specifically, we have

$$F(P) = 3\left(\theta_{12}^2+\theta_{23}^2+\theta_{13}^2\right)+2\cos(\Theta_{213})\theta_{12}\theta_{13}+2\cos(\Theta_{123})\theta_{12}\theta_{23}+2\cos(\Theta_{231})\theta_{23}\theta_{13}. \tag{3.18}$$

In particular, it follows that the right hand side of (3.18) must be the same for each set of data for all four triangles of P.

Up to now we only discussed identities that the extremal partition P must satisfy. In order to proceed we need to prove some a priori estimates on the various parameters in question. First, let $M = \max\{\theta_{12}, \theta_{13}, \theta_{23}\}$. It follows from (3.18) that $F(P) \leqslant 14M^2$ (using the fact that at least one Θ_{ijk} must be greater than $\frac{\pi}{3}$). Our contrapositive assumption (3.4) combined with (3.7) implies that $F(P) > \frac{9\pi^2}{4}$, so we deduce that all the triangles in P must have an edge of length greater than $\frac{3\pi}{2\sqrt{14}} > \frac{5}{4}$. Additional arguments (that are significantly more involved technically) imply that for all distinct $i, j, \ell \in \{1,2,3,4\}$ we have $\theta_{ij} \leqslant \pi - \frac{1}{2}$ (Lemma 3.3.12) and $\theta_{ij}, \sin\Theta_{ij\ell} > \frac{1}{10}$ (Lemma 1.1.1). In Lemma 3.4.1 we prove that $(\theta_{12}, \theta_{13}, \theta_{23})$ must be outside the ℓ_2 ball of radius $\frac{1}{100}$ centered at (θ, θ, θ) for $\theta = \arccos\left(-\frac{1}{3}\right)$ and $\theta = 1.53796841207904$. This excludes from our search space balls centered at the two solutions of (3.17) that were discussed above. In Lemma 3.4.2 we show that, up to a relabeling of the spherical triangles T_1, T_2, T_3, T_4, we may assume that $\sqrt{\lambda} > \frac{9}{50}$.

It turns out that the above estimates suffice in order to conclude the proof of Theorem 3.1.1 via a search over a sufficiently fine net. To explain this endgame, note that by the spherical law of cosines we have

$$\cos\theta_{13} = \cos\theta_{12}\cos\theta_{23} + \sin\theta_{12}\sin\theta_{23}\cos\Theta_{123}.$$

It therefore follows from (3.13), (3.16), (3.15) that if we define $h, \lambda, \gamma : [0,\pi]^3 \to \mathbb{R}$

by

$$\lambda(x,y,z) := \frac{\sin^2 x}{x} + \frac{\sin^2 y}{y} + \frac{\sin^2 z}{z} \\ + 2\frac{\sin x \sin y}{xy}\cos z + 2\frac{\sin x \sin z}{xz}\cos y + 2\frac{\sin y \sin z}{yz}\cos x,$$

$$\gamma(x,y,z) := \frac{\sin z}{z}\cos x + \frac{\sin y}{y} + \frac{\sin x}{x}\cos z,$$

$$h(x,y,z) := \sqrt{\lambda(x,y,z)}\cos\left(\sqrt{x^2+z^2+2xz\frac{\cos y - \cos x \cos z}{\sin x \sin z}}\right) + \gamma(x,y,z),$$

then $h(\theta_{12},\theta_{13},\theta_{23}) = 0$. This identity must hold for all cyclic permutations of the indices $\{1,2,3\}$, so we also have the two identities $h(\theta_{13},\theta_{23},\theta_{12}) = h(\theta_{23},\theta_{12},\theta_{13}) = 0$. It follows that if we define $H = (H_1,H_2,H_3) : [0,\pi]^3 \to \mathbb{R}^3$ by $H(x,y,z) = (h(y,z,x), h(x,y,z), h(z,x,y))$ then at the extremal partition P, $H(\theta_{12},\theta_{13},\theta_{23}) = 0$.

Our strategy is therefore as follows. At every point q in the search space given by the constraints described above (specifically, see the system of equations (3.66), excluding also the two balls described in Lemma 3.4.1), we will show that either $H(q) \neq 0$ or $F_0(q) < 9\pi^2/4$, where F_0 is given by the right hand side of (3.18) (with the understanding that one expresses $\cos\Theta_{123}, \cos\Theta_{213}, \cos\Theta_{231}$ in terms of $\theta_{12},\theta_{23},\theta_{13}$ using the spherical law of cosines, so that F_0 is a function of $\theta_{12},\theta_{23},\theta_{13}$). Due to the above discussion, such an assertion will complete the proof of Theorem 3.1.1. Moreover, it turns out that this assertion holds "with room to spare", making a computer-assisted verification feasible.

To this end, we need to have estimates on the modulus of continuity of H and F_0. Such estimates are complicated but can be proved using elementary considerations; see Lemma 3.4.3 and the discussion immediately following it. Consequently, if we are given a point q in our restricted search space and $\tau \in (0,1)$ for which either $|H(q)| > \tau$ or $F_0(q) < 9\pi^2/4 - \tau$, then it would follow that for some $r > 0$ that depends on τ via our modulus of continuity estimates, in the entire ball of radius r centered at q either $H \neq 0$ or $F_0 < 9\pi^2/4$.

From this procedure we achieve a process by which we can iteratively remove from our search space macroscopically large regions of controlled size in which we are guaranteed that a counterexample to the Propeller Conjecture cannot exist. Our program proceeds to remove such balls until it eventually exhausts the entire search space, thus arriving at the conclusion that there is no counterexample to the Propeller Conjecture.

In order to make such a procedure rigorous, we carefully account for all possible rounding errors. We do this quite conservatively, i.e., while significantly overestimating the magnitude of all possible errors, as explained in detail in Section 3.4.1.

Note that our computer-assisted proof does not compute any computationally complicated expressions such as, say, definite integrals or implicitly defined functions: it only checks inequalities between expressions involving elementary combinations of trigonometric functions and square roots.

In summary, the validity of the Propeller Conjecture in $\mathbb{R}^3$ has mainly conceptual implications, and what remains open seems to be of a more technical nature: to perhaps find a clever transformation, e.g., as in Ball's cube slicing theorem [8], that allows one to prove the theorem analytically rather than resorting to a transversal of a net in a region where the desired inequality is actually quite weak. At present, we have a clean geometric argument which addresses directly the region where the Propeller Conjecture is most subtle, and in the remainder of the search space we do something rather crude. This has two conceptual consequences: the Propeller Conjecture as a nonintuitive link between complexity theory and geometry is correct, and moreover one has a geometric/analytic proof of this conjecture in a region where it is tightest. Nevertheless, we are lacking a technical idea that allows us to address the remaining region in a way that does not resort to "brute force". It remains a challenge to find such an idea, with the hope that it will pave the way to a proof of the Propeller Conjecture in all dimensions.

3.3 Proofs of basic identities and estimates

This section contains proofs the geometric identities and inequalities that were stated in Section 3.2. Recall that we are assuming that $P = \{T_1, T_2, T_3, T_4\}$ is a partition of S^2 into four spherical triangles, and that $\{z_i\}_{i=1}^4$, as defined in (3.7), are the corresponding spherical moments. We are also making the contrapositive assumption that our objective function

$$F(P) = \|z_1\|_2^2 + \|z_2\|_2^2 + \|z_3\|_2^2 + \|z_4\|_2^2 \tag{3.19}$$

(recall (3.14)) exceeds $9\pi^2/4$, which is the maximal value of $F(\cdot)$ that is predicted by the Propeller Conjecture. We also assume that P maximizes F. The vertices of the partition $\{v_i\}_{i=1}^4$ were defined in Section 3.2, along with the angles θ_{ij} and $\Theta_{ij\ell}$ as given in (3.11).

Proposition 3.3.1. (a) $z_1 + z_2 + z_3 + z_4 = 0$.

(b) Suppose two triangles $T_i, T_j \in P$ share an edge $E := T_i \cap T_j$. Then for all $e \in E$ we have $\langle z_i, e\rangle = \langle z_j, e\rangle$. In particular, if $\ell \notin \{i, j\}$ then $\langle z_i, v_\ell\rangle = \langle z_j, v_\ell\rangle$.

(c) For all distinct $i, j \in \{1, 2, 3, 4\}$ we have $\langle z_i, v_i\rangle = \langle -3z_j, v_i\rangle$.

(d) For all distinct $i, j \in \{1, 2, 3, 4\}$ we have $v_j \notin \{-v_i, v_i\}$.

(e) *For all distinct $i, j, \ell \in \{1,2,3,4\}$ we have $\det(v_i, v_j, v_\ell) \neq 0$, each T_i is contained in an open hemisphere, and $0 < \theta_{ij}, \Theta_{ij\ell} < \pi$.*

Proof. Parts (a) and (b) were already proved in Section 3.2; for part (b) see (3.9). A combination of (3.10) and part (a) implies part (c). For $\ell \in \{1,2,3,4\}$ let Π_ℓ be the unique plane containing $\{z_i\}_{i\in\{1,2,3,4\}\setminus\{\ell\}}$. A vector is perpendicular to Π_ℓ if and only if it has equal inner product on $\{z_i\}_{i\in\{1,2,3,4\}\setminus\{\ell\}}$, so $\mathbb{R}v_\ell$ is also perpendicular to Π_ℓ. Since $\{z_i\}_{i=1}^4$ are not coplanar the planes $\{\Pi_\ell\}_{\ell=1}^4$ have distinct perpendiculars. So v_1, v_2, v_3, v_4 are distinct with $v_i \notin \mathbb{R}v_j$ for $i \neq j$. This establishes (d). For every $\ell \in \{1,2,3,4\}$ consider the set

$$U_\ell := \left\{ x \in S^2 : \min_{i\in\{1,2,3,4\}\setminus\ell} \langle x, z_i - z_\ell\rangle > \frac{1}{2}\langle v_\ell, z_i - z_\ell\rangle \right\}.$$

It follows from (3.10) that U_ℓ is an open neighborhood of v_ℓ, and in combination with (3.8) we know that

$$i \in \{1,2,3,4\} \setminus \{\ell\} \implies U_\ell \cap T_i = \left\{ x \in U_\ell \cap S^2 : \langle x, z_i\rangle = \max_{j\in\{1,2,3,4\}\setminus\{\ell\}} \langle x, z_j\rangle \right\}. \tag{3.20}$$

Write $I = \{1,2,3,4\} \setminus \{\ell\}$. For every distinct $i, k \in \{1,2,3,4\}$ the set $\{x \in S^2 : \langle x, z_i\rangle \geqslant \langle x, z_j\rangle\}$ is a half sphere, and therefore the set $T_{i,\ell} := \{x \in S^2 : \langle x, z_i\rangle = \max_{j\in I}\langle x, z_j\rangle\}$ is strictly contained in a half sphere (since the z_j are distinct and nonzero). Note that $\partial T_{i,\ell}$ is the union of two half circles that meet at antipodal points, and v_ℓ is one of these antipodal points. So, the spherical angle of $T_{i,\ell}$ at the vertex v_ℓ is in $(0,\pi)$. By (3.8) and (3.20) this spherical angle is identical for $T_{i,\ell}$ and T_i. This shows that $\Theta_{i\ell j} \in (0,\pi)$ for all distinct $i, j, \ell \in \{1,2,3,4\}$. Also $\theta_{ij} \in (0,\pi)$ since $v_j \notin \{-v_i, v_i\}$. Next, assume for the sake of contradiction that $\det(v_i, v_j, v_\ell) = 0$. Since $v_i \notin \mathbb{R}v_j$ for distinct $i, j \in \{1,2,3,4\}$, it follows that $v_i = \alpha_j v_j + \alpha_\ell v_\ell$ with $\alpha_j, \alpha_\ell \in \mathbb{R} \setminus \{0\}$. Hence $v_i \times v_j = \alpha_\ell v_\ell \times v_j$ with $\alpha_\ell \neq 0$, contradicting $\Theta_{ij\ell} \in (0,\pi)$. We have already seen that $T_{i,\ell}$ is contained a closed hemisphere, and this containment can be chosen such that $T_{i,\ell}$ intersected with the boundary of the hemisphere is $\{v_\ell, -v_\ell\}$. But from (3.8) we know that $T_i \subseteq T_{i,\ell}$ and T_i avoids a neighborhood of $\{-v_\ell\}$, so each T_i is contained in an open hemisphere. This concludes the proof of (e). □

We now give an explicit formula for the spherical moments $\{z_i\}_{i=1}^4$; as explained in Section 3.2, this formula can be found in [74, p. 259] and [20, p. 6]. Since these references are hard to find, we provide a proof.

Proposition 3.3.2. *Suppose that the spherical triangle $T_4 \subseteq S^2$ has vertices $\{v_1, v_2, v_3\}$ satisfying $\det(v_1, v_2, v_3) > 0$. Then the vector $z_4 = \int_{T_4} x d\sigma \in \mathbb{R}^3$ satis-*

fies

$$z_4 = \frac{1}{2}\left(\theta_{12}\frac{v_1 \times v_2}{\|v_1 \times v_2\|_2} + \theta_{23}\frac{v_2 \times v_3}{\|v_2 \times v_3\|_2} + \theta_{31}\frac{v_3 \times v_1}{\|v_3 \times v_1\|_2}\right). \tag{3.21}$$

Consequently

$$\langle z_4, v_1\rangle = \left\langle \frac{\theta_{23}}{2}\frac{v_2 \times v_3}{\|v_2 \times v_3\|_2}, v_1 \right\rangle = \frac{\theta_{23}\det(v_1, v_2, v_3)}{2\sin\theta_{23}} \tag{3.22}$$

Proof. Fix $y \in T_4 \setminus \partial T_4$. By definition of z_4,

$$\langle z_4, y\rangle = \int_{T_4} \langle x, y\rangle d\sigma(x). \tag{3.23}$$

Let $H \subseteq \mathbb{R}^3$ be the convex hull of $T_4 \cup \{0\}$ and let $\{U_i\}_{i=1}^3 \subseteq \mathbb{R}^3$ be half spaces through the origin such that $\{x \in \mathbb{R}^3 \colon \|x\|_2 \leq 1\} \bigcap \left(\bigcap_{i=1}^3 U_i\right) = H$; see Figure 3.3.

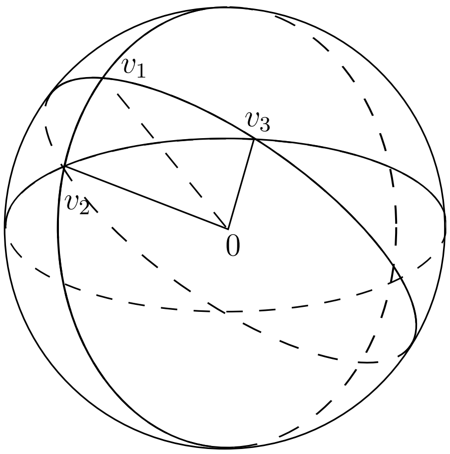


Figure 3.3: Spherical triangle together with three disc sectors, each containing $\{v_i, v_j, 0\}$.

Define $\Pi\colon S^2 \to \mathbb{R}^2$ so that $\Pi(x)$ is the orthogonal projection of $x \in S^2 \subseteq \mathbb{R}^3$ onto the unique plane that is intersecting the origin and perpendicular to y. Applying the Cauchy projection formula (see for example [60, p. 25]) to T_4, we see that

$$\begin{aligned}\int_{T_4} \langle x, y\rangle d\sigma(x) = \mathrm{Area}_{\mathbb{R}^2}\Pi(T_4 \cap \{x \in S^2 \colon \langle y, x\rangle > 0\})\\ - \mathrm{Area}_{\mathbb{R}^2}\Pi(T_4 \cap \{x \in S^2 \colon \langle y, x\rangle \leq 0\}).\end{aligned} \tag{3.24}$$

Define $A := (T_4 \setminus \partial T_4) \cap \{x \in S^2 \colon \langle y, x\rangle > 0\}$ and $B := (\partial H) \setminus A$. Note that ∂H is the union of T_4 and three sectors of discs $\{D_i\}_{i=1}^3$ such that $D_i \subseteq \partial U_i$. Since $y \in \bigcap_{i=1}^3 U_i$, the exterior normal $n(x)$ of $x \in D_i \subseteq \partial H$ has nonpositive inner product with y for $i = 1, 2, 3$. So, $\langle n(x), y\rangle > 0$ for $x \in \partial H$ if and only if $x \in A$. Since H is convex, $\Pi \colon A \to \Pi(\partial H)$ and $\Pi \colon B \to \Pi(\partial H)$ are bijections almost everywhere. Since $\partial H = T_4 \bigcup \left(\bigcup_{i=1}^3 D_i\right)$, the Cauchy projection formula applied to right side of the identity $\mathrm{Area}_{\mathbb{R}^2}(\Pi(A)) = \mathrm{Area}_{\mathbb{R}^2}(\Pi(B))$ gives

$$\begin{aligned}\mathrm{Area}_{\mathbb{R}^2}\Pi(T_4 \cap \{x \in S^2 \colon \langle y, x\rangle > 0\}) &= \mathrm{Area}_{\mathbb{R}^2}\Pi(T_4 \cap \{x \in S^2 \colon \langle y, x\rangle \le 0\}) \\ + \frac{\theta_{12}}{2}\left\langle y, \frac{v_1 \times v_2}{\|v_1 \times v_2\|_2}\right\rangle &+ \frac{\theta_{23}}{2}\left\langle y, \frac{v_2 \times v_3}{\|v_2 \times v_3\|_2}\right\rangle + \frac{\theta_{13}}{2}\left\langle y, \frac{v_3 \times v_1}{\|v_3 \times v_1\|_2}\right\rangle.\end{aligned} \tag{3.25}$$

Since (3.25) holds for any $y \in S^2 \cap T_4$, (3.23), (3.24) and (3.25) give (3.21). □

Note that the vectors in the determinant of (3.22) are assumed to have positive orientation. So, when we apply this equation below we will need to respect orientations.

Corollary 3.3.3. *Under the assumptions of Proposition 3.3.2, we have*

$$\begin{aligned}\|z_4\|_2^2 &= \frac{1}{4}\left(\theta_{12}^2 + \theta_{23}^2 + \theta_{13}^2 - 2\cos(\Theta_{213})\theta_{12}\theta_{13}\right. \\ &\qquad \left. - 2\cos(\Theta_{123})\theta_{12}\theta_{23} - 2\cos(\Theta_{231})\theta_{23}\theta_{13}\right) \qquad (3.26) \\ &= \frac{1}{4}\left(\theta_{12}^2 + \theta_{23}^2 + \theta_{31}^2 + 2\frac{\cos\theta_{12}\cos\theta_{13} - \cos\theta_{23}}{\sin\theta_{12}\sin\theta_{13}}\theta_{12}\theta_{13}\right. \\ &\quad \left. + 2\frac{\cos\theta_{12}\cos\theta_{23} - \cos\theta_{13}}{\sin\theta_{12}\sin\theta_{23}}\theta_{12}\theta_{23} + 2\frac{\cos\theta_{23}\cos\theta_{13} - \cos\theta_{12}}{\sin\theta_{23}\sin\theta_{13}}\theta_{23}\theta_{13}\right). \qquad (3.27)\end{aligned}$$

Proof. Apply (3.21), and then use either the vector identity $\langle a \times b, c \times d\rangle = \langle a, c\rangle\langle b, d\rangle - \langle a, d\rangle\langle b, d\rangle$, or the spherical law of cosines. □

Corollary 3.3.4. *For distinct $i, j \in \{1, 2, 3, 4\}$ we have $\langle z_i, v_j\rangle > 0$.*

Proof. Let k, ℓ be such that $\det(v_k, v_j, v_\ell) > 0$, and apply (3.22) and Proposition 3.3.1(e). □

Lemma 3.3.5. *Write $\langle z_4, v_1\rangle =: a$, $\langle z_4, v_2\rangle =: b$, $\langle z_4, v_3\rangle =: c$ and $\langle z_4, v_4\rangle =: -3d$. Then the following relation holds*

$$\frac{1}{a}v_1 + \frac{1}{b}v_2 + \frac{1}{c}v_3 + \frac{1}{d}v_4 = 0 \tag{3.28}$$

Proof. By Proposition 3.3.1(e) we have $\det(v_1, v_2, v_3) \neq 0$, so we may write $v_4 = \sum_{i=1}^{3} \alpha_i v_i$ for some $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$. By Proposition 3.3.1(c),

$$d = \langle z_1, v_4 \rangle = \langle z_2, v_4 \rangle = \langle z_3, v_4 \rangle.$$

Substituting in the expression $v_4 = \sum_{i=1}^{3} \alpha_i v_i$ and then applying Proposition 3.3.1(b) and (c) several times, we get

$$d = -3a\alpha_1 + b\alpha_2 + c\alpha_3 = a\alpha_1 - 3b\alpha_2 + c\alpha_3 = a\alpha_1 + b\alpha_2 - 3c\alpha_3.$$

Solving this system of equations gives

$$d = -a\alpha_1 = -b\alpha_2 = -c\alpha_3.$$

By Corollary 3.3.4, we may write $\alpha_1 = -d/a, \alpha_2 = -d/b, \alpha_3 = -d/c$. Substituting these equalities into the expression $v_4 = \sum_{i=1}^{3} \alpha_i v_i$ gives (3.28). Note that at least one α_i is nonzero, so d is nonzero as well, that is, we have not made a division by zero. $\square$

Since $\det(v_1, v_2, v_3) \neq 0$, we have $\|\frac{1}{a}v_1 + \frac{1}{b}v_2 + \frac{1}{c}v_3\|_2 \neq 0$. Since $\|v_4\|_2 = 1$, Lemma 3.3.5 implies

$$v_4 = -\frac{\frac{1}{a}v_1 + \frac{1}{b}v_2 + \frac{1}{c}v_3}{\|\frac{1}{a}v_1 + \frac{1}{b}v_2 + \frac{1}{c}v_3\|_2}. \tag{3.29}$$

A priori, we have only determined v_4 up to multiplication by ± 1. However, Proposition 3.3.1(c) and Corollary 3.3.4 imply that $d = (-1/3)\langle z_4, v_4 \rangle > 0$. So, taking the inner product of z_4 with both sides of (3.29) shows that (3.29) has the correct sign.

Applying (3.22) to (3.29) gives

$$v_4 = -\frac{\frac{\sin\theta_{23}}{\theta_{23}}v_1 + \frac{\sin\theta_{13}}{\theta_{13}}v_2 + \frac{\sin\theta_{12}}{\theta_{12}}v_3}{\left\|\frac{\sin\theta_{23}}{\theta_{23}}v_1 + \frac{\sin\theta_{13}}{\theta_{13}}v_2 + \frac{\sin\theta_{12}}{\theta_{12}}v_3\right\|_2}. \tag{3.30}$$

Define, as we did in Section 3.2,

$$\lambda := \left\|\frac{\sin\theta_{23}}{\theta_{23}}v_1 + \frac{\sin\theta_{13}}{\theta_{13}}v_2 + \frac{\sin\theta_{12}}{\theta_{12}}v_3\right\|_2^2 \tag{3.31}$$

$$\begin{aligned} &= \frac{\sin^2\theta_{23}}{\theta_{23}^2} + \frac{\sin^2\theta_{13}}{\theta_{13}^2} + \frac{\sin^2\theta_{12}}{\theta_{12}^2} + 2\frac{\sin\theta_{23}\sin\theta_{13}}{\theta_{23}\theta_{13}}\cos\theta_{12} \\ &\quad + 2\frac{\sin\theta_{23}\sin\theta_{12}}{\theta_{23}\theta_{12}}\cos\theta_{13} + 2\frac{\sin\theta_{12}\sin\theta_{13}}{\theta_{12}\theta_{13}}\cos\theta_{23}. \end{aligned} \tag{3.32}$$

Proposition 3.3.6. (a) *All four vertices of P are not contained in any closed half sphere.*

(b) The function $F(P)$ is well-defined as a function of the vertices $\{v_i\}_{i=1}^4$.

Proof. (a) By Proposition 3.3.1(e) we have $0 < \theta_{ij} < \pi$. (3.30) says that v_4 is in the complement of all closed half spheres containing $\{v_1, v_2, v_3\}$.

(b) From Proposition 3.3.1(e), the edges $T_i \cap T_j$ of P are geodesics of length less than π, so the vertices uniquely determine the edges. Finally, the edges determine P itself. □

Proposition 3.3.7 (Restrictions on products of opposite edges)**.**

$$\frac{\sin\theta_{12}\sin\theta_{34}}{\theta_{12}\theta_{34}} = \frac{\sin\theta_{13}\sin\theta_{24}}{\theta_{13}\theta_{24}} = \frac{\sin\theta_{14}\sin\theta_{23}}{\theta_{14}\theta_{23}}. \tag{3.33}$$

Proof. From Proposition 3.3.1(c) we have $\langle z_3, v_1\rangle = \langle z_2, v_1\rangle$. For each side of the equality $\langle z_3, v_1\rangle = \langle z_2, v_1\rangle$, apply (3.22) for $\langle z_i, v_1\rangle$ and then substitute in (3.30) to get

$$-\frac{\det(v_1, v_3, v_2)}{2\sqrt{\lambda}}\frac{\sin\theta_{12}}{\theta_{12}}\frac{\theta_{24}}{\sin\theta_{24}} = -\frac{\det(v_1, v_3, v_2)}{2\sqrt{\lambda}}\frac{\sin\theta_{13}}{\theta_{13}}\frac{\theta_{34}}{\sin\theta_{34}}.$$

Canceling terms gives the first equality of (3.33). The second follows similarly: Proposition 3.3.1(c) says $\langle z_3, v_2\rangle = \langle z_1, v_2\rangle$, and so on. □

Lemma 3.3.8 (Restrictions on a single triangle)**.** *Let λ be defined as in* (3.32) *and let γ be defined as in* (3.15)*, i.e.,*

$$\gamma := \frac{\sin\theta_{23}}{\theta_{23}}\cos\theta_{12} + \frac{\sin\theta_{13}}{\theta_{13}} + \frac{\sin\theta_{12}}{\theta_{12}}\cos\theta_{23}. \tag{3.34}$$

Then the following holds

$$\cos\left(\sqrt{\lambda - \gamma^2}\frac{\theta_{23}\theta_{12}}{\sin\theta_{23}\sin\theta_{12}}\right) = -\frac{\gamma}{\sqrt{\lambda}}. \tag{3.35}$$

$$\cos\left(\sqrt{\theta_{12}^2 + \theta_{23}^2 + 2\theta_{12}\theta_{23}\cos\Theta_{123}}\right) = -\frac{\gamma}{\sqrt{\lambda}}. \tag{3.36}$$

Proof. Arguing as in Proposition 3.3.7, we start with Proposition 3.3.1(c) and the equality $\langle z_4, v_1\rangle = \langle z_3, v_1\rangle$. For each side of this equality, apply (3.22) for $\langle z_i, v_1\rangle$ and substitute in (3.30) to get

$$\frac{\theta_{23}\theta_{12}}{\sin\theta_{23}\sin\theta_{12}} = \frac{\theta_{24}}{\sin\theta_{24}}\frac{1}{\sqrt{\lambda}}. \tag{3.37}$$

Substituting $\theta_{24} = \arccos(\langle v_2, v_4\rangle)$ into (3.37) and rearranging gives

$$\cos\left(\sqrt{\lambda}\sqrt{1-\langle v_2, v_4\rangle^2}\frac{\theta_{23}\theta_{12}}{\sin\theta_{23}\sin\theta_{12}}\right) = \langle v_2, v_4\rangle.$$

From (3.30) and the definitions of λ and γ ((3.32) and (3.34)), we have $\langle v_2, v_4\rangle = -\gamma/\sqrt{\lambda}$, so

$$\cos\left(\sqrt{\lambda}\sqrt{1-\frac{\gamma^2}{\lambda}}\frac{\theta_{23}\theta_{12}}{\sin\theta_{23}\sin\theta_{12}}\right) = -\frac{\gamma}{\sqrt{\lambda}}.$$

yielding (3.35). To derive (3.36), observe that $\lambda - \gamma^2$ allows some cancelation as follows.

$$\begin{aligned}
\lambda - \gamma^2 &= \frac{\sin^2\theta_{23}}{\theta_{23}^2}(1-\cos^2\theta_{12}) + \frac{\sin^2\theta_{12}}{\theta_{12}^2}(1-\cos^2\theta_{23}) \\
&\quad + 2\frac{\sin\theta_{23}\sin\theta_{12}}{\theta_{23}\theta_{12}}(\cos\theta_{31} - \cos\theta_{23}\cos\theta_{12}) \\
&= \frac{\sin^2\theta_{23}}{\theta_{23}^2}(\sin^2\theta_{12}) + \frac{\sin^2\theta_{12}}{\theta_{12}^2}(\sin^2\theta_{23}) \\
&\quad + 2\frac{\sin\theta_{23}\sin\theta_{12}}{\theta_{23}\theta_{12}}(\cos\theta_{31} - \cos\theta_{23}\cos\theta_{12}).
\end{aligned}$$

Thus,

$$\begin{aligned}
(\lambda-\gamma^2)\frac{\theta_{23}^2\theta_{12}^2}{\sin^2\theta_{23}\sin^2\theta_{12}} &= \theta_{12}^2 + \theta_{23}^2 + 2\theta_{12}\theta_{23}\frac{\cos\theta_{31} - \cos\theta_{23}\cos\theta_{12}}{\sin\theta_{12}\sin\theta_{23}} \\
&= \theta_{12}^2 + \theta_{23}^2 + 2\theta_{12}\theta_{23}\cos\Theta_{123} \quad \text{, by the spherical law of cosines.}
\end{aligned}$$

Substituting this equation into (3.35) gives (3.36). ◻

Since $-\frac{\gamma}{\sqrt{\lambda}} = \langle v_4, v_2\rangle = \cos\theta_{24}$, (3.36) says that

$$\theta_{12}^2 + \theta_{23}^2 + 2\theta_{12}\theta_{23}\cos\Theta_{123} = \theta_{24}^2. \tag{3.38}$$

Substituting (3.38) back into (3.26) and (3.19) gives a simplified form of $F(P)$.

Lemma 3.3.9.

$$F(P) = \sum_{i<j}\theta_{ij}^2 = \sum_{i<j}(\arccos(\langle v_i, v_j\rangle))^2 \tag{3.39}$$

Proof. Applying (3.27) to the definition $F(P) = \sum_{i=1}^4 \|z_i\|_2^2$ gives

$$F(P) = \frac{1}{2}\sum_{i<j}\theta_{ij}^2 - \frac{1}{2}\sum_{i,j,\ell}\cos(\Theta_{ij\ell})\theta_{ij}\theta_{j\ell}. \tag{3.40}$$

Here the first sum in (3.40) runs over all (i,j) satisfying $1 \leq i < j \leq 4$, and the second sum runs over all equivalence classes of 3-tuples (i,j,ℓ) under the equivalence $(i,j,\ell) \sim (\ell,j,i)$, where i,j,ℓ are distinct elements of $\{1,2,3,4\}$. From (3.38) we have, for $r \notin \{i,j,\ell\}$,

$$\theta_{ij}^2 + \theta_{j\ell}^2 + 2\theta_{ij}\theta_{j\ell}\cos\Theta_{ij\ell} = \theta_{jr}^2. \tag{3.41}$$

Substituting (3.41) into (3.40) gives

$$F(P) = \frac{1}{2}\sum_{i<j}\theta_{ij}^2 - \frac{1}{4}\sum_{i,j,\ell}\left(\theta_{jr}^2 - \theta_{ij}^2 - \theta_{j\ell}^2\right).$$

In the right summation, for a fixed (i',j') with $i' \neq j'$, the term $\theta_{i'j'}^2$ appears four times with a minus sign and two times with a plus sign. Therefore,

$$F(P) = \frac{1}{2}\sum_{i<j}\theta_{ij}^2 - \frac{1}{4}\sum_{i,j,\ell}\left(\theta_{jr}^2 - \theta_{ij}^2 - \theta_{j\ell}^2\right) = \frac{1}{2}\sum_{i<j}\theta_{ij}^2 - \frac{1}{4}(\sum_{i<j} -2\theta_{ij}^2) = \sum_{i<j}\theta_{ij}^2,$$

proving (3.39). $\square$

In (3.39) replace θ_{jr}^2 with the left side of (3.41) for $(j,r) \in \{(1,4),(2,4),(3,4)\}$ to get

$$F(P) = 3\left(\theta_{12}^2 + \theta_{23}^2 + \theta_{13}^2\right) + 2\cos(\Theta_{213})\theta_{12}\theta_{13} + 2\cos(\Theta_{123})\theta_{12}\theta_{23} + 2\cos(\Theta_{231})\theta_{23}\theta_{13}. \tag{3.42}$$

Definition 3.3.10. By the spherical law of cosines, note that $\Theta_{ij\ell}$ is a function of $\{\theta_{ij}, \theta_{j\ell}, \theta_{i\ell}\}$. Let

$$\begin{aligned} F_0(\theta_{12},\theta_{23},\theta_{13}) := 3\left(\theta_{12}^2 + \theta_{23}^2 + \theta_{13}^2\right) &+ 2\cos(\Theta_{213})\theta_{12}\theta_{13} \\ &+ 2\cos(\Theta_{123})\theta_{12}\theta_{23} + 2\cos(\Theta_{231})\theta_{23}\theta_{13}. \end{aligned}$$

By (3.42), $F(P) = F_0$ can be computed by the data from only one triangle of P. That is,

$$F_0(\theta_{12},\theta_{23},\theta_{13}) = F_0(\theta_{12},\theta_{24},\theta_{14}) = F_0(\theta_{34},\theta_{24},\theta_{23}) = F_0(\theta_{34},\theta_{14},\theta_{13}).$$

Since the spherical propeller partition, i.e., the partition of S^2 corresponding to the intersection of the propeller partition of $\mathbb{R}^3$ with S^2, satisfies $F = (9/4)\pi^2$, we may bound (3.42) in terms of $M := \max\{\theta_{12},\theta_{13},\theta_{23}\}$. In particular, we get $F(P) \leq 14M^2$, which is less than $(9/4)\pi^2$ if $M < \pi\frac{3}{2\sqrt{14}}$. (In our bound on $F(P)$, we use that one Θ_{ijk} must be larger than $\pi/3$ by the pigeonhole principle applied to $\Theta_{123} + \Theta_{231} + \Theta_{312} - \pi = \mathrm{Area}_{S^2}(T_4) > 0$, so one cosine term satisfies $\cos\Theta_{ijk} \leq 1/2$.) So, this bound on F gives

Corollary 3.3.11. *Any triangle T_i of P must have an edge of length greater than $\pi\frac{3}{2\sqrt{14}} > \frac{5}{4}$.*

Iteratively applying (3.33) now gives an improvement to Proposition 3.3.1(e). This improvement will be important in our numerical calculations in Section 3.4 and in the proof of Theorem 3.1.1, since we eventually require a bound on the derivative of $\theta_{ij}/\sin\theta_{ij}$.

Lemma 3.3.12 (Restrictions on edge length)**.** *For every distinct $i, j \in \{1, 2, 3, 4\}$,*

$$\theta_{ij} < \pi - \frac{1}{2}. \tag{3.43}$$

Proof. Suppose $\theta_{12} \geq \pi - 1/2$. We are eventually going to derive a contradiction, which will allow us to conclude that $\theta_{12} < \pi - 1/2$. We will essentially only need (3.33), the monotonicity of $x \mapsto (\sin x)/x$ for $x \in (0, \pi)$, the fact that the perimeter $\theta_{ij} + \theta_{j\ell} + \theta_{\ell i}$ of a spherical triangle is bounded by 2π, and that $(\sin x)/x \leq 1$ on $[0, \pi]$. Relabeling the edges if necessary, we may assume that $\theta_{13} = \max\{\theta_{13}, \theta_{24}\}$. Using $\theta_{12} \geq \pi - 1/2$, (3.33) gives

$$\frac{\sin(\pi - 1/2)}{\pi - 1/2} \geq \frac{\sin\theta_{12}}{\theta_{12}} \geq \frac{\sin\theta_{12}\sin\theta_{34}}{\theta_{12}\theta_{34}} = \frac{\sin\theta_{13}\sin\theta_{24}}{\theta_{13}\theta_{24}} \geq \left(\frac{\sin\theta_{13}}{\theta_{13}}\right)^2. \tag{3.44}$$

Therefore, $\theta_{13} \geq X$ where $X \in (0, \pi)$ satisfies

$$\left(\frac{\sin X}{X}\right)^2 = \frac{\sin(\pi - 1/2)}{\pi - 1/2}. \tag{3.45}$$

So, $\theta_{13} \geq X > 2.065$. Now, since $\theta_{12} + \theta_{13} + \theta_{23} \leq 2\pi$ the bounds on θ_{12} and θ_{13} give

$$\theta_{23} \leq 2\pi - (\pi - 1/2 + X) =: Y < 1.577.$$

But then (3.33) and the bounds on θ_{12} and θ_{23} yield

$$\frac{\sin\theta_{14}}{\theta_{14}} = \frac{\sin\theta_{12}\sin\theta_{34}}{\theta_{12}\theta_{34}}\frac{\theta_{23}}{\sin\theta_{23}} \leq \frac{\sin(\pi - 1/2)}{\pi - 1/2}\frac{Y}{\sin Y}.$$

So, $\theta_{14} \geq Z$ where $Z \in (0, \pi)$ is defined by

$$\frac{\sin Z}{Z} = \frac{\sin(\pi - 1/2)}{\pi - 1/2}\frac{Y}{\sin Y}.$$

Then $\theta_{14} \geq Z > 2.388$. But since $\theta_{14} + \theta_{24} + \theta_{12} \leq 2\pi$ the bounds on θ_{14} and θ_{12} give

$$\theta_{24} \leq 2\pi - (\pi - 1/2 + Z) =: W < 1.254$$

We now iterate the above argument. Using (3.44),

$$\frac{\sin\theta_{13}}{\theta_{13}} = \frac{\sin\theta_{12}\sin\theta_{34}}{\theta_{12}\theta_{34}}\frac{\theta_{24}}{\sin\theta_{24}} \leq \frac{\sin(\pi - 1/2)}{\pi - 1/2}\frac{W}{\sin W}. \tag{3.46}$$

So, $\theta_{13} \geq \tilde{X}$ with $(\sin\tilde{X})/\tilde{X} = (\sin(\pi - 1/2)/(\pi - 1/2))(W/\sin W)$. We now repeat the argument above, starting at (3.45) and using tildes to designate updated variables. Thus, $\theta_{13} \geq \tilde{X} > 2.499$, $\theta_{23} \leq \tilde{Y} < 1.14, \theta_{14} \geq \tilde{Z} > 2.52, \theta_{24} \leq \tilde{W} < 1.122$. Using $\tilde{W}$ in place of W in (3.46), we find that $\theta_{13} > 2.52$.

In summary, iterating the above procedure improves our bounds on the θ_{ij}. We find that $\theta_{13}, \theta_{14} > 2.52$. We also find that $\theta_{23}, \theta_{24} < 1.14$. Since $\theta_{13}+\theta_{14}+\theta_{34} \leq 2\pi$, $\theta_{34} \leq 2\pi - (2(2.52))$. But then $\theta_{23}, \theta_{24}, \theta_{34} < 5/4 < \pi\frac{3}{2\sqrt{14}}$, violating Corollary 3.3.11 and giving our desired contradiction. We therefore conclude that $\theta_{ij} < \pi - 1/2$. $\square$

For a partition $P = \{T_i\}_{i=1}^4$ as above, the following map $G\colon (S^2)^4 \to \mathbb{R}$ is also a function of P by Proposition 3.3.6(b)

$$G(v_1, v_2, v_3, v_4) := \sum_{1\leq i<j\leq 4} \theta_{ij}^2. \tag{3.47}$$

Proposition 3.3.13. *Suppose that the conclusion of Proposition 3.3.1(e) holds. Then Proposition 3.3.1(c) holds if and only if P is a critical point of G.*

Proof. For a manifold M with $v \in M$, T_vM denotes the tangent space of M at v. Since $G = \sum_{i<j}\theta_{ij}^2 = \sum_{i<j}(\arccos\langle v_i, v_j\rangle)^2$, we take derivatives of G with respect to v_i. For $v_2, v_3, v_4 \in S^2$ fixed and $x \in \mathbb{R}^3$ variable,

$$\nabla_{S^2} G(\cdot, v_2, v_3, v_4) = \left(\nabla_{\mathbb{R}^3} G\left(\frac{x}{\|x\|_2}, v_2, v_3, v_4\right)\right)\Big|_{S^2}$$

Recall that $T_{v_1}\mathbb{R}^3$ is isomorphic to $\mathbb{R}^3$. We denote this isomorphism by $(d/dy) \leftrightarrow y$.

Let $\frac{d}{dy}$ be a vector in the tangent space $T_{v_1}\mathbb{R}^3$. Identifying $\frac{d}{dy}$ with $y \in \mathbb{R}^3$ we compute

$$\begin{aligned}
\frac{d}{dy} G\left(\frac{x}{\|x\|_2}, v_2, v_3, v_4\right) &= \sum_{j\neq 1} \frac{-2\arccos(\langle x, v_j\rangle/\|x\|_2)}{\sqrt{1-(\langle x, v_j\rangle/\|x\|_2)^2}}\left\langle \frac{d}{dy}(x/\|x\|_2), v_j\right\rangle \\
&= \sum_{j\neq 1} \frac{-2\arccos(\langle x, v_j\rangle/\|x\|_2)}{\sqrt{1-(\langle x, v_j\rangle/\|x\|_2)^2}}\left(\frac{\|x\|_2 \frac{d}{dy}\langle x, v_j\rangle - \frac{1}{2}\langle x, v_j\rangle \|x\|_2^{-1}\frac{d}{dy}\langle x, x\rangle}{\|x\|_2^2}\right) \\
&= \sum_{j\neq 1} \frac{-2\arccos(\langle x, v_j\rangle/\|x\|_2)}{\sqrt{1-(\langle x, v_j\rangle/\|x\|_2)^2}}\left\langle\left(\frac{\|x\|_2 y - x\|x\|_2^{-1}\langle y, x\rangle}{\|x\|_2^2}\right), v_j\right\rangle.
\end{aligned}$$

Letting $\frac{d}{dy} \in T_{v_1}S^2$ and $x = v_1 \in S^2$ gives $\langle y, x\rangle = 0$, so that

$$\frac{d}{dy}G\left(\frac{x}{\|x\|_2}, v_2, v_3, v_4\right) = \sum_{j\neq 1} \frac{-2\arccos(\langle v_1, v_j\rangle)}{\sqrt{1-(\langle v_1, v_j\rangle)^2}} \langle y, v_j\rangle = \sum_{j\neq 1} \frac{-2\theta_{1j}}{\sin\theta_{1j}} \langle y, v_j\rangle .$$

So, if we take $\frac{d}{dy} := v_1 \times v_2 \in T_{v_1}S^2$ and then use (3.22), we get

$$\frac{d}{dy}G = \frac{-2\theta_{13}}{\sin\theta_{13}} \det(v_1, v_2, v_3) - \frac{2\theta_{14}}{\sin\theta_{14}} \det(v_1, v_2, v_4) = 2(-\langle z_4, v_2\rangle + \langle z_3, v_2\rangle).$$

In general, if we take the derivative at v_i in the direction $v_i \times v_j$ and set this derivative to zero, we get $\langle z_r, v_j\rangle = \langle z_\ell, v_j\rangle$ where the set $\{i, j, \ell, r\}$ is equal to the set $\{1, 2, 3, 4\}$. Since the above argument can be reversed, the proposition is proven. $\square$

Setting $\frac{d}{dy} = v_2 - \langle v_2, v_1\rangle v_1$ and $\frac{d}{dy} = v_3 - \langle v_3, v_1\rangle v_3$ in the above argument gives

$$-\theta_{12} = \theta_{13}\cos\Theta_{213} + \theta_{14}\cos\Theta_{214}. \tag{3.48}$$

Observe that

$$\begin{aligned}
\frac{d}{dy}G\left(\frac{x}{\|x\|_2}, v_2, v_3, v_4\right) &= \sum_{j\neq 1} \frac{-2\theta_{1j}}{\sin\theta_{1j}} \langle y, v_j\rangle = \sum_{j\neq 1} \frac{-2\theta_{1j}}{\sin\theta_{1j}} \langle (v_2 - \langle v_1, v_2\rangle v_1), v_j\rangle \\
&= -\frac{2\theta_{12}}{\sin\theta_{12}}(1-(\langle v_2, v_1\rangle)^2) - \frac{2\theta_{13}}{\sin\theta_{13}}(\langle v_2, v_3\rangle - (\langle v_2, v_1\rangle)(\langle v_1, v_3\rangle)) \\
&\quad - \frac{2\theta_{14}}{\sin\theta_{14}}(\langle v_2, v_4\rangle - (\langle v_2, v_1\rangle)(\langle v_1, v_4\rangle)) \\
&= -\frac{2\theta_{12}}{\sin\theta_{12}}(1-(\langle v_2, v_1\rangle)^2) - \frac{2\theta_{13}}{\sin\theta_{13}} \langle (v_1 \times v_2), (v_1 \times v_3)\rangle \\
&\quad - \frac{2\theta_{14}}{\sin\theta_{14}} \langle (v_1 \times v_2), (v_1 \times v_4)\rangle .
\end{aligned}$$

Thus, dividing by $\sin\theta_{12}$, setting $\frac{d}{dy}G = 0$, and simplifying as in Corollary 3.3.3 gives (3.48).

We now give a geometric interpretation of (3.38) and (3.48) that should clarify their meaning. Define $\exp_1^{-1} \colon \{2, 3, 4\} \to \mathbb{R}^2$ by

$$\exp_1^{-1}(j) := \theta_{1j}\frac{(v_j - \langle v_j, v_1\rangle v_1)}{\|v_j - \langle v_j, v_1\rangle v_1\|_2}. \tag{3.49}$$

Then $\exp_1^{-1}(j)$ is perpendicular to v_1 and parallel to $v_j - \mathrm{Proj}_{v_1} v_j$, with length θ_{1j}.

Lemma 3.3.14. *Suppose that the conclusion of Proposition 3.3.1(e) holds. If* $\{v_i\}_{i=1}^4$ *is a critical point of* G, *then*

$$\exp_1^{-1}(2) + \exp_1^{-1}(3) + \exp_1^{-1}(4) = 0. \tag{3.50}$$

Proof. Let $\Pi_{ij} \subseteq \mathbb{R}^3$ be the plane with $\Pi_{ij} \supseteq \{0, v_i, v_j\}$. For $j \in \{2,3,4\}$, $\exp_1^{-1}(j) \in \Pi_{1j}$ is perpendicular to v_1 and in Π_{1j}, so $\langle \exp_1^{-1}(2), \exp_1^{-1}(3)\rangle = \theta_{12}\theta_{13}\cos\Theta_{213}$. We can therefore rewrite (3.38) as

$$\left\|\exp_1^{-1}(2) + \exp_1^{-1}(3)\right\|_2^2 = \left\|\exp_1^{-1}(4)\right\|_2^2. \tag{3.51}$$

Using the definition of $\exp_1^{-1}(j)$, we similarly rewrite (3.48) (with its indices permuted) as

$$-\exp_1^{-1}(4) = \mathrm{Proj}_{\exp_1^{-1}(4)}(\exp_1^{-1}(2)) + \mathrm{Proj}_{\exp_1^{-1}(4)}(\exp_1^{-1}(3)). \tag{3.52}$$

Since $\exp_1^{-1}(j)$ is perpendicular to v_1 for $j \in \{2,3,4\}$, $\{\exp_1^{-1}(j)\}_{j\in\{2,3,4\}}$ lies in a plane containing the origin. Combining (3.51) and (3.52) therefore gives (3.50). $\square$

In particular, orthogonally projecting (3.50) orthogonal to $\exp_1^{-1}(2)$ gives

$$\theta_{13}\sin\Theta_{213} = \theta_{14}\sin\Theta_{214}. \tag{3.53}$$

The similarity of (3.38),(3.48) and (3.53) is no coincidence. For example, note that (3.48) and (3.53) express $\sum_{j\in\{2,3,4\}} \exp_1^{-1}(j) = 0$, tangential and perpendicular to $\exp_1^{-1}(2)$, respectively. So, combining Proposition 3.3.13, (3.48), (3.53) and Lemma 3.3.14, we deduce the equivalence

Corollary 3.3.15. *Suppose the conclusion of Proposition 3.3.1(e) holds. Then*

$$\left\{\begin{matrix}\textit{Prop. 3.3.1(c)}\\ \textit{holds}\end{matrix}\right\} \Leftrightarrow \left\{\begin{matrix}P \textit{ is a}\\ \textit{critical point of } G\end{matrix}\right\} \Leftrightarrow \left\{\begin{matrix}\text{(3.50) } \textit{holds at each } v_i\textit{. That is,}\\ \textit{for each } i,\ \sum_{j:\, j\neq i}\exp_i^{-1}(j) = 0\end{matrix}\right\}$$

Lemma 3.3.16. $\theta_{ij} > 1/10$ *and* $\sin\Theta_{ij\ell} > 1/10$.

Proof. After recording a few identities, we consider a few cases in proving $\theta_{ij} > 1/10$. In all cases, we argue by contradiction and assume $0 < \theta_{12} \le 1/10$. Also, by relabeling if necessary, we may assume that $\max\{\theta_{13}, \theta_{23}, \theta_{14}, \theta_{24}\} = \max\{\theta_{13}, \theta_{23}\}$.

$$\cos\Theta_{213} = \frac{\cos\theta_{23} - \cos\theta_{12}\cos\theta_{13}}{\sin\theta_{12}\sin\theta_{13}}. \tag{3.54}$$

$$\cos\Theta_{132} = \frac{\cos\theta_{12} - \cos\theta_{13}\cos\theta_{23}}{\sin\theta_{13}\sin\theta_{23}}. \tag{3.55}$$

$$\begin{aligned}\cos\theta_{34} &= \cos\theta_{13}\cos\theta_{14} + \sin\theta_{13}\sin\theta_{14}\cos\Theta_{314} \\ &= \cos\theta_{13}\cos\theta_{14} + \sin\theta_{13}\sin\theta_{14}\cos(2\pi - \Theta_{213} - \Theta_{214}).\end{aligned} \tag{3.56}$$

$$|\arccos(x+y) - \arccos(x)| \leq \arccos(1-y), \quad x, y, x+y \in [-1, 1]. \tag{3.57}$$

(Case 1) $\pi/2 \leq \max\{\theta_{13}, \theta_{23}\} \leq \pi/2 + 1/10$. Without loss of generality, label the edges so that $\theta_{13} \geq \theta_{23}$, so $\pi/2 \leq \theta_{13} \leq \pi/2 + 1/10$. Then $\theta_{14} \geq \theta_{24}$, from Proposition 3.3.1(b). The Proposition shows that $\mathrm{Proj}_{\mathrm{span}\{v_1,v_2\}}(z_4 - z_3) = 0$. Suppose $\theta_{14} < \theta_{24}$. Let e be the midpoint of the edge $T_4 \cap T_3$ and let $n \in \mathrm{span}\{v_1, v_2\}$ be perpendicular to e such that $\langle n, v_2\rangle > 0$. Since $\theta_{13} \geq \theta_{23}$, $\langle z_4, n\rangle \geq 0$. To see this, let $\Pi\colon S^2 \to \mathbb{R}^2$ denote projection onto the unique plane intersecting the origin and perpendicular to n. Without loss of generality, the vertices are oriented so that $\det(v_1, v_2, v_3) > 0$. Let m be the point of intersection of this plane with $(\partial T_4) \setminus (\partial T_3)$. The spherical triangle T_4' with vertices $\{v_1, v_2, m\}$ then has two equal length edges $\{v_1, m\}$ and $\{v_2, m\}$, and $T_4' \subseteq T_4$. Also, $\theta_{13} \geq \theta_{23}$ implies that $T_4 \cap \{x \in S^2 \colon \langle x, n\rangle \leq 0\} = T_4' \cap \{x \in S^2 \colon \langle x, n\rangle \leq 0\}$. Since $T_4 \supseteq T_4'$ and T_4' is isosceles,

$$\begin{aligned}\Pi\left(T_4 \cap \{x \in S^2 \colon \langle x, n\rangle \geq 0\}\right) &\supseteq \Pi\left(T_4' \cap \{x \in S^2 \colon \langle x, n\rangle \geq 0\}\right) \\ &= \Pi\left(T_4 \cap \{x \in S^2 \colon \langle x, n\rangle \leq 0\}\right).\end{aligned}$$

So (3.24) shows $\langle z_4, n\rangle \geq 0$. Similarly, $\theta_{14} < \theta_{24}$, implies $\langle z_3, n\rangle < 0$, a contradiction. Now

$$\begin{aligned}\cos\Theta_{213} &= \frac{\cos\theta_{23} - \cos\theta_{13}\cos\theta_{12}}{\sin\theta_{13}\sin\theta_{12}} \geq \frac{\cos\theta_{13}(1 - \cos\theta_{12})}{\sin\theta_{13}\sin\theta_{12}} \\ &= \frac{\tan(\theta_{12}/2)}{\tan\theta_{13}} \geq \frac{\tan(1/20)}{\tan(\pi/2 + 1/10)}.\end{aligned}$$

So, $\Theta_{213} \leq \arccos(\frac{\tan(1/20)}{\tan(\pi/2+1/10)}) < \pi/2 + .01$. By (3.38), $(\theta_{13} - \theta_{12})^2 \leq \theta_{14}^2 \leq (\theta_{12} + \theta_{13})^2$, so $\pi/2 - 1/10 \leq \theta_{14} \leq \theta_{13} + \theta_{12} \leq \pi/2 + 1/10 + 1/10$. So we similarly conclude that $\Theta_{214} \leq \arccos(\frac{\tan(1/20)}{\tan(\pi/2+1/5)}) < \pi/2 + .02$. By (3.56),

$$\cos\theta_{34} = \cos(\theta_{13} + \theta_{14}) + \sin\theta_{13}\sin\theta_{14}(1 + \cos(2\pi - \Theta_{213} - \Theta_{214})). \tag{3.58}$$

Since $\Theta_{314} < \pi$ by Proposition 3.3.1(e),

$$|1 + \cos\Theta_{314}| = |1 + \cos(2\pi - \Theta_{213} - \Theta_{214})| \leq |1 + \cos(2\pi - \pi - .03)| < .0005.$$

Then (3.57) applied to (3.58) gives a dichotomy: (i) if $\theta_{13} + \theta_{14} \leq \pi$, then $|\theta_{34} - (\theta_{13} + \theta_{14})| < .04$. Or (ii) if $\theta_{13} + \theta_{14} > \pi$, then $|\theta_{34} - (2\pi - \theta_{13} + \theta_{14})| < .04$. In case (i), $\theta_{34} \geq \pi/2 + \pi/2 - 1/10 - .04 > \pi - 1/2$. In case (ii), $\theta_{34} \geq 2\pi - (\pi/2 + 1/10 + \pi/2 + 1/5) - .04 > \pi - 1/2$. So, in either case, Lemma 3.3.12 is violated. Therefore Case 1 cannot hold.

(Case 2) $\pi/2 + 1/10 \leq \max\{\theta_{13}, \theta_{23}\} \leq 2.9$. Without loss of generality, label the edges so that $\theta_{13} \geq \theta_{23}$, so $\pi/2 + 1/10 \leq \theta_{13} \leq 2.9$. By the triangle inequality on the sphere, $\theta_{23} \geq \theta_{13} - \theta_{12} \geq \theta_{13} - 1/10 \geq \pi/2$, so

$$\begin{aligned}\theta_{13}\cos\Theta_{132} &= \theta_{13}\left(\frac{\cos\theta_{12} - \cos\theta_{13}\cos\theta_{23}}{\sin\theta_{13}\sin\theta_{23}}\right)\\ &\geq \theta_{13}\left(\frac{\cos(1/10) - \cos\theta_{13}\cos\theta_{13}}{\sin\theta_{13}\sin(\theta_{13} - 1/10)}\right) > \frac{\pi}{2}.\end{aligned}$$

Since $\theta_{23}, \theta_{13} \geq \pi/2$, (3.38) gives

$$\theta_{34} = \sqrt{\theta_{13}^2 + \theta_{23}^2 + 2\theta_{13}\theta_{23}\cos\Theta_{132}} > \sqrt{2(\pi/2)^2 + 2(\pi/2)(\pi/2)} = \pi,$$

so (3.43) is violated. Therefore Case 2 cannot occur.

(Case 3) The case $\max\{\theta_{13}, \theta_{23}\} > 2.9$ cannot occur, by (3.43). Also, note that $\max\{\theta_{13}, \theta_{23}, \theta_{14}, \theta_{24}\} < \pi/2$ implies that T_4 and T_3 are contained in separate but adjacent quarter spheres, so $T_4 \cup T_3$ is contained in a half sphere. To see this, consider the edge $T_4 \cap T_3$. Let Π_{12} be the plane such that $\Pi_{12} \cap S^2 \supseteq T_4 \cap T_3$. Temporarily suppose $\theta_{13} = \theta_{23} = \pi/2$, and let $R \colon \mathbb{R}^3 \to \mathbb{R}^3$ be the origin fixing rotation that maps the midpoint of $T_4 \cap T_3$ to v_3. Then T_4 is contained in a quarter sphere whose boundary is contained in $\Pi_{12} \cap S^2$ and $(R(\Pi_{12})) \cap S^2$. If $\theta_{13}, \theta_{23} \leq \pi/2$, then T_4 remains in this quarter sphere, since otherwise $v_3 \notin \{x \in S^2 \colon d_{S^2}(x, v_1) \leq \pi/2, d_{S^2}(x, v_2) \leq \pi/2\}$. So, Proposition 3.3.6(a) is violated. All cases therefore produce a contradiction, concluding the proof that $\theta_{12} > 1/10$.

We now prove that $\sin\Theta_{ijk} > 1/10$. We argue by contradiction and assume that $\sin\Theta_{132} < 1/10$. The procedure is similar to the calculations above. The theme is: if some Θ_{ijk} is very small or very large, this creates additive relations among the edges, up to small errors. Before we begin, we list a few consequences of the spherical law of cosines, for i, j, k distinct elements of $\{1, 2, 3, 4\}$.

$$\cos\theta_{ij} = \cos(\theta_{ik} - \theta_{kj}) + \sin\theta_{ik}\sin\theta_{kj}(\cos\Theta_{ikj} - 1). \tag{3.59}$$

$$\cos\theta_{ij} = \cos(\theta_{ik} + \theta_{kj}) + \sin\theta_{ik}\sin\theta_{kj}(\cos\Theta_{ikj} + 1). \tag{3.60}$$

Now, assume $\Theta_{132} < \pi/2$. Since $\Theta_{132} + \Theta_{234} + \Theta_{134} = 2\pi$ and $-\Theta_{134} > -\pi$ from Proposition 3.3.1(e), $\Theta_{132} + \Theta_{234} \geq \pi$. So $\Theta_{234} > \pi/2$ and $\sin\Theta_{234} \leq \sin(\pi - \Theta_{132}) = \sin(\Theta_{132}) < 1/10$. So, by relabeling edges if necessary, we may assume that $\Theta_{132} > \pi/2$. Since $\Theta_{132} > \pi/2$, $\cos\Theta_{132} \leq -.9949$. Without loss of generality, label the edges so that $\theta_{23} = \max\{\theta_{13}, \theta_{23}\}$ and $\theta_{13} = \min\{\theta_{13}, \theta_{23}\}$.

(Case 1') Assume $\max\{\theta_{13}, \theta_{23}\} < \pi/3$. By (3.38), $\theta_{34} \leq \theta_{13} + \theta_{23} \leq 2\pi/3$. But then Proposition 3.3.6(a) is violated. (Let $v \in T_1 \cap T_2$ satisfy $d_{S^2}(v, v_4) = \min\{\theta_{34}, \pi/2\}$, and note that $\{v_i\}_{i=1}^4 \subseteq \{x \in S^2 \colon d(x, v) \leq \pi/2\}$.) So Case 1' cannot occur.

(Case 2') Assume $\max\{\theta_{13}, \theta_{23}\} \geq \pi/3$ and $1/10 < \min\{\theta_{13}, \theta_{23}\} \leq 1/2$. Then $\theta_{13} + \theta_{23} < \pi$ by Lemma 3.3.12, so (3.56) and (3.57) give $|\theta_{12} - (\theta_{13} + \theta_{23})| < .11$. By the spherical law of sines and Lemma 3.3.12, $\sin\Theta_{123} = \sin\Theta_{132}\sin\theta_{13}/\sin\theta_{12} \leq \sin\Theta_{132}\sin(1/2)/\sin(1/2) < 1/10$. Since $\theta_{13} \leq \theta_{23}$ and $\theta_{13} \leq 1/2$, Lemma 3.3.12 gives $\cos\theta_{13} \geq |\cos\theta_{23}| \geq \cos\theta_{23}\cos\theta_{12}$. So, the formula $\cos\Theta_{123} = (\cos\theta_{13} - \cos\theta_{23}\cos\theta_{12})/(\sin\theta_{12}\sin\theta_{23})$ shows $\cos\Theta_{123} \geq 0$, i.e. $\cos\Theta_{123} > .9949$.

Since $\theta_{23} < \pi - 1/2$ from Lemma 3.3.12, $\theta_{13} + \theta_{23} < \pi$. Combining this with $\cos\Theta_{132} < -.9949$, (3.60) and (3.57), we see that $\theta_{12} \geq \theta_{13} + \theta_{23} - .11$. Since $\cos\Theta_{123} > .9949$, (3.38) and $\theta_{12} > .1$ imply

$$\theta_{24} \geq \sqrt{(\pi/3)^2 + (.1)^2 + 2(.1)(\pi/3)(.9949)} > \pi/3.$$

Now, since $\cos\Theta_{123} > .9949$, $\Theta_{123} + \Theta_{124} + \Theta_{324} = 2\pi$, and $-\Theta_{123} > -\pi$ from Proposition 3.3.1(e), $\Theta_{123} + \Theta_{324} \geq \pi$, so $\cos\Theta_{324} \leq \cos(\pi - \Theta_{123}) < -.9949$. Since $\theta_{23}, \theta_{24} > \pi/3$, by relabeling the edges if necessary we may assume that $\cos\Theta_{123} < -.9949$, $\max\{\theta_{13}, \theta_{23}\} \geq \pi/3$ and $1/2 < \pi/3 < \min\{\theta_{13}, \theta_{23}\}$. Thus, we enter into the following case.

(Case 3') Assume $\max\{\theta_{13}, \theta_{23}\} \geq \pi/3$ and $1/2 < \min\{\theta_{13}, \theta_{23}\}$. Recall that we may label the edges so that $\theta_{23} = \max\{\theta_{13}, \theta_{23}\}$ and $\theta_{13} = \min\{\theta_{13}, \theta_{23}\}$. We claim that $\theta_{23} > 1.3$. If not, then (3.38) says $\theta_{34}^2 = \theta_{23}^2 + \theta_{13}^2 + 2\theta_{23}\theta_{13}\cos\Theta_{132}$, so $\theta_{34} \leq \sqrt{2(1.3)^2 - 2(.994)(\pi/3)(1/2)} < 1.53 < \pi/2$ and $\{v_i\}_{i=1}^4 \subseteq \{x \in S^2 \colon d_{S^2}(x, v_3) \leq \pi/2\}$. Thus, Proposition 3.3.6(a) is violated, and our claim is proven.

As in Case 1 in the proof that $\theta_{ij} > 1/10$, (3.60) gives a dichotomy

$$(i) \quad |\theta_{12} - (\theta_{13} + \theta_{23})| < .11, \qquad \text{or} \quad (ii) \quad |\theta_{12} + \theta_{13} + \theta_{23} - 2\pi| < .11$$

In case (i), $\theta_{12} \geq \theta_{23} + \theta_{13} - .11 > 1/2$, and in case (ii), Lemma 3.3.12 says $\theta_{ij} < \pi - 1/2$, so $\theta_{12} \geq .89 > 1/2$. So in either case, $1/2 \leq \theta_{12} \leq \pi - 1/2$, using Lemma 3.3.12 again. By the spherical law of sines, $\sin\Theta_{123} \leq \sin\Theta_{132}/\sin\theta_{12} \leq (1/10)(1/\sin(1/2)) < 21/100$. Similarly, $\sin\Theta_{312} < 21/100$.

Assume (i) holds. If $\Theta_{123} < \pi/2$, then $\cos\Theta_{123} > .977$. By (i), $\theta_{12} \geq 1.3 + 1/2 - .11 = 1.69$. By (3.38), $\theta_{24}^2 = \theta_{12}^2 + \theta_{23}^2 + 2\theta_{12}\theta_{23}\cos\Theta_{123}$, so

$$\theta_{24} \geq \sqrt{(1.69)^2 + (\pi/3)^2 + 2(1.69)(\pi/3)(.977)} > 2.72 > \pi - \frac{1}{2}.$$

contradicting Lemma 3.3.12. If $\Theta_{123} \geq \pi/2$ then $\cos\Theta_{123} < -.977$. Then (3.60) gives a dichotomy: we either have $(i)'$ $|\theta_{13} - (\theta_{12} + \theta_{23})| < .26$ or the following holds $(ii)'$ $|\theta_{13} + \theta_{12} + \theta_{23} - 2\pi| < .26$. If $(i)'$ holds, then $\theta_{13} \geq \theta_{12} + \theta_{23} - .26$, but (i) says $\theta_{12} \geq \theta_{13} + \theta_{23} - .11$. So $\theta_{13} \geq \theta_{13} + 2\theta_{23} - .37$, i.e. $\theta_{23} < .185$, contradicting that $\theta_{23} > \pi/3$. So $(ii)'$ must hold. But then (i) implies $2\theta_{12} \geq \theta_{12} + \theta_{13} + \theta_{23} - .11 \geq 2\pi - .37$, so $\theta_{12} \geq \pi - .185$, contradicting Lemma 3.3.12. Therefore (i) does not hold.

Assume (ii) holds. Suppose $\Theta_{123} < \pi/2$ so $\cos\Theta_{123} > .977$. Then (3.59) and (3.57) show that $|\theta_{13} - |\theta_{12} - \theta_{23}|| < .26$. If $\theta_{12} \geq \theta_{23}$, then $|\theta_{12} - (\theta_{13} + \theta_{23})| < .26$. So by (ii) $2\theta_{12} \geq \theta_{12} + \theta_{13} + \theta_{23} - .26 \geq 2\pi - .37$, contradicting Lemma 3.3.12. If $\theta_{23} \geq \theta_{12}$, then the same argument shows $2\theta_{23} \geq 2\pi - .37$, contradicting Lemma 3.3.12. We may therefore assume that $\cos\Theta_{123} < -.977$. If $\cos\Theta_{312} > .977$, we get the same contradiction, so we may assume $\cos\Theta_{312} < -.977$. In summary, $\cos\Theta_{123}, \cos\Theta_{312}, \cos\Theta_{132} < -.977$.

For i,j distinct, $i,j \in \{1,2,3,4\}$ let $E_{ij} = T_i \cap T_j$. Let i',j' so that $\theta_{i'j'} := \max\{\theta_{13}, \theta_{23}, \theta_{12}\}$. Let $\Pi_{i'j'}$ be the plane containing $E_{i'j'}$ and the origin, and let n be the unit normal to $\Pi_{i'j'}$ such that $\langle n, z_4\rangle < 0$. We claim that the edges $E_{i'4}, E_{j'4}$ satisfy

$$E_{i'4}, E_{j'4} \subseteq \{x \in S^2 \colon 0 \leq \langle x, n\rangle \leq .65\}. \tag{3.61}$$

Also, the union of $E_{i'4}, E_{j'4}, E_{i'j'}$ forms a noncontractible loop in this topological annulus.

To prove $E_{i'4} \subseteq \{x \in S^2 \colon \langle x, n\rangle \leq .65\}$, it suffices to show

$$\Theta_{j'i'4} > \pi - \sin^{-1}(.65), \quad \text{or} \quad \theta_{i'4} < \sin^{-1}(.65). \tag{3.62}$$

To see this consequence, let $x \in S^2$ be contained in the great circle containing $E_{i'4}$ and let $\{i', j', k\} = \{1,2,3\}$. Since $\langle n, v_{i'}\rangle = 0$ and $\langle n, z_k\rangle > 0$, $\langle x, n\rangle$ is maximized when $d_{S^2}(x, v_{i'}) = \pi/2$ and $\langle x, n\rangle \geq 0$. For such an x, $\langle x, n\rangle = \sin\Theta_{j'i'4}$. So, in the case that $\Theta_{j'i'4} > \pi - \sin^{-1}(.65)$, $E_{i'4} \subseteq \{x \in S^2 \colon \langle x, n\rangle \leq .65\}$. Let now $x \in \{y \in S^2 \colon d_{S^2}(y, v_{i'}) \leq \theta_{i'4}\} \supseteq E_{i'4}$. For such an x, $\langle x, n\rangle$ is maximized for x in the plane containing $\{v_{i'}, n, 0\}$ with $d_{S^2}(x, v_{i'}) = \theta_{i'4}$. So $\langle x, n\rangle = \sin\theta_{i'4}$. So, in the case that $\theta_{i'4} < \sin^{-1}(.65)$, $E_{i'4} \subseteq \{x \in S^2 \colon \langle x, n\rangle \leq .65\}$, as desired. The containment $E_{j'4} \subseteq \{x \in S^2 \colon \langle x, n\rangle \leq .65\}$ follows similarly if one shows: $\Theta_{i'j'4} > \pi - \sin^{-1}(.65)$ or $\theta_{j'4} < \sin^{-1}(.65)$. Specifically, in the above paragraph we switch all indices of the form i' to j' and all j' to i', and the containment follows.

We now prove (3.62). Let $\{i', j', k\} = \{1,2,3\}$. Since $\Theta_{j'i'k}, \Theta_{i'j'k} < \pi$ by Proposition 3.3.1(e) and $E_{i'j'} \subseteq \{x \in S^2 \colon \langle x, n\rangle = 0\}$, we conclude that $E_{i'k}, E_{j'k} \subseteq \{x \in S^2 \colon \langle x, n\rangle \leq 0\}$. By Lemma 3.3.14, $\sum_{j \colon j \neq i'} \exp_{i'}^{-1}(j) = 0$. Moreover, $\sum_{j \colon j \neq j'} \exp_{j'}^{-1}(j) = 0$, so we conclude $E_{i'4}, E_{j'4} \subseteq \{x \in S^2 \colon 0 \leq \langle x, n\rangle\}$. By the definition of $\theta_{i'j'}$ and Lemma 3.3.14, $\Theta_{j'i'4} \geq \Theta_{ki'4}$. In particular, since $\Theta_{j'i'4} + \Theta_{ki'4} \geq \pi$, we conclude that $\Theta_{j'i'4} \geq \pi/2$. Similarly, $\Theta_{i'j'4} \geq \pi/2$. If $E_{i'4} \cup E_{i'j'} \cup E_{j'4}$ were contractible in $\{x \in S^2 \colon 0 \leq \langle x, n\rangle \leq .65\}$, then by (3.62), the location of v_4 is such that T_3 avoids the region $\{x \in S^2 \colon \langle x, n\rangle \geq .65 \text{ or } \langle x, n\rangle \leq 0\}$. But then either $\Theta_{i'j'4}$ or $\Theta_{j'i'4}$ is less than $\pi/2$, a contradiction. So, the noncontractibility property of $E_{i'4} \cup E_{i'j'} \cup E_{j'4}$ follows from (3.62), which remains to be proven.

Consider $\theta_{i'k}$ and $E_{i'j'}$ as fixed with $\Theta_{j'i'k}, E_{i'4}$ variable. From Lemma 3.3.14, recall that $\exp_{i'}^{-1}(4)$ and $\exp_{i'}^{-1}(j')$ have angle $\Theta_{j'i'4}$. Also $\exp_{i'}^{-1}(k)$ and $\exp_{i'}^{-1}(j')$

have angle $\Theta_{j'i'k}$, and $\exp_{i'}^{-1}(\cdot)$ is confined to a plane. By Lemma 3.3.14, scaling $\theta_{i'j'}$ and $\theta_{i'k}$ by the same constant (greater than one) leaves $\Theta_{j'i'4}$ unchanged and increases $\theta_{i'4}$. So in proving (3.62), we may assume that $\theta_{i'j'} = \pi - 1/2$, by Lemma 3.3.12.

Suppose $\theta_{i'k} < .815(\pi - 1/2)$. In this case, we will show that $\sin\Theta_{j'i'4} < .65$, recalling that $\Theta_{j'i'4} > \pi/2$. Using planar geometry and Lemma 3.3.14, one can show: decreasing $\Theta_{j'i'k}$ decreases $\Theta_{j'i'4}$. So, for the purpose of showing $\sin\Theta_{j'i'4} < .65$, we may assume that $\cos\Theta_{j'i'k} = -.977$.

Now, suppose $\theta_{i'k} \geq .815(\pi - 1/2)$. In this case we show that $\sin\theta_{i'4} < .65$, recalling that $\theta_{i'4} < \pi/2$, since (3.41) says

$$\begin{aligned}\theta_{i'4}^2 &= (\theta_{i'j'} - \theta_{i'k})^2 + 2\theta_{i'j'}\theta_{i'k}(1 + \cos\Theta_{j'i'k}) \\ &\leq (.185(\pi - 1/2))^2 + 2(\pi - 1/2)^2(.815)(.033) < .62.\end{aligned}$$

Consider $\theta_{i'k}, E_{i'j'}$ as fixed with $\Theta_{j'i'k}, E_{i'4}$ variable. By Lemma 3.3.14,

$$\theta_{i'4}^2 = \left\|\exp_{i'}^{-1}(4)\right\|_2^2 = \left\|\exp_{i'}^{-1}(j') + \exp_{i'}^{-1}(k)\right\|_2^2 = \theta_{i'j'}^2 + \theta_{i'k}^2 + 2\theta_{i'j'}\theta_{i'k}\cos\Theta_{j'i'k}.$$

Recall $\cos\Theta_{j'i'k} < -.977$, so decreasing $\Theta_{j'i'k}$ increases $\theta_{i'4}$. So, for the purpose of maximizing $\sin\theta_{i'4}$, we may assume that $\cos\Theta_{j'i'k} = -.977$. With these reductions (i.e. that $\cos\Theta_{j'i'k} = -.977$ and $\theta_{i'j'} = \pi - 1/2$) for these two cases, one can then verify (3.62) directly as a one-dimensional inequality, using Lemma 3.3.14 and treating θ_{ik} as a variable with all other quantities a function of θ_{ik}. With identical reductions, $\Theta_{i'j'4} > \pi - \sin^{-1}(.65)$ $\theta_{j'4} < \sin^{-1}(.65)$. Therefore, (3.61) holds.

Given (3.61), note that the circumference of the geodesic ball $\{x \in S^2 \colon \langle x, n\rangle = .65\}$ is given by $2\pi\sqrt{1 - (.65)^2} > 4.77$. By the noncontractible property, $\theta_{i'j'} + \theta_{i'4} + \theta_{j'4} \geq 4.77$. By (3.38), $\theta_{i'4} \geq \sqrt{(\theta_{i'j'} - \theta_{i'k})^2 + 2\theta_{i'j'}\theta_{i'k}(1 + \cos\Theta_{j'i'k})} \geq \theta_{i'j'} - \theta_{i'k}$ and similarly $\theta_{j'4} \geq (\theta_{i'j'} - \theta_{j'k})$. Substituting these inequalities into $\theta_{i'j'} + \theta_{i'4} + \theta_{j'4} \geq 4.77$ gives

$$\theta_{i'j'} + (\theta_{i'j'} - \theta_{i'k}) + (\theta_{i'j'} - \theta_{j'k}) \geq 4.77.$$

However, by (ii), $\theta_{i'j'} + \theta_{i'k} + \theta_{j'k} \geq 2\pi - .11$. So adding the inequalities gives $4\theta_{i'j'} \geq 4.77 + 2\pi - .11$, i.e. $\theta_{i'j'} > 2.73 > \pi - 1/2$, violating Lemma 3.3.12. So, all cases produce a contradiction, and the result is proven. □

3.4 Numerical Computations

We now provide a more comprehensive analysis of the system of equations resulting from Lemma 3.3.8. We begin by writing this system explicitly. For $i \in \{1, 2, 3\}$ choose j, ℓ such that $\{i, j, \ell\} = \{1, 2, 3\}$. Define

$$\gamma_i := -\sqrt{\lambda}\,\langle v_i, v_4\rangle = \frac{\sin\theta_{i\ell}}{\theta_{i\ell}}\cos\theta_{ij} + \frac{\sin\theta_{j\ell}}{\theta_{j\ell}} + \frac{\sin\theta_{ij}}{\theta_{ij}}\cos\theta_{i\ell}. \tag{3.63}$$

Then $\gamma = \gamma_2$ from (3.34). Our system follows by cyclically permuting the θ_{ij} in (3.36), i.e.,

$$\begin{cases} \sqrt{\lambda}\cos\left(\sqrt{\theta_{12}^2+\theta_{13}^2+2\theta_{12}\theta_{13}\cos\Theta_{213}}\right)+\gamma_1=0, \\ \sqrt{\lambda}\cos\left(\sqrt{\theta_{12}^2+\theta_{23}^2+2\theta_{12}\theta_{23}\cos\Theta_{123}}\right)+\gamma_2=0, \\ \sqrt{\lambda}\cos\left(\sqrt{\theta_{13}^2+\theta_{23}^2+2\theta_{13}\theta_{23}\cos\Theta_{132}}\right)+\gamma_3=0, \end{cases} \tag{3.64}$$

where $\Theta_{ij\ell}$ and λ are defined in (3.11) and (3.31) respectively.

Observe that (3.64) gives three equations in the three unknowns $\theta_{12}, \theta_{13}, \theta_{23}$. Define $H\colon [0,\pi]^3 \to \mathbb{R}^3$ so that $H(\theta_{12}, \theta_{13}, \theta_{23}) = (H_1, H_2, H_3)$ is the vector corresponding to the entries of the left side of (3.64). We now wish to find $(\theta_{12}, \theta_{13}, \theta_{23})$ such that $H = (0, 0, 0)$.

We first examine neighborhoods of the two points discussed in Section 3.3.

Lemma 3.4.1. *The global maximum of F does not occur in the ℓ_2 balls of radius $1/100$ around the points $\{\theta_{12}, \theta_{13}, \theta_{23}\} = \{\theta, \theta, \theta\} \in \mathbb{R}^3$ with $\theta = \arccos(-1/3)$ and $\theta = 1.53796841207904$.*

Proof. Let $\delta = 1/100$. We claim that $|\langle \nabla_{\mathbb{R}^3} F_0(\theta_{12}, \theta_{13}, \theta_{23}), v\rangle| \leq 20$, for all $(\theta_{12}, \theta_{13}, \theta_{23})$ in an ℓ_2 ball in $\mathbb{R}^3$ of radius δ ball around the point

$$(\arccos(-1/3), \arccos(-1/3), \arccos(-1/3)) \in \mathbb{R}^3,$$

where $v \in S^2$. Now,

$$F_0(\arccos(-1/3), \arccos(-1/3), \arccos(-1/3)) = 6(\arccos(-1/3))^2 \approx 21.9031.$$

Also, F at the spherical propeller partition evaluates to $(9/4)\pi^2 \approx 22.206$. So, given this claim, since $\delta\,|\langle \nabla_{\mathbb{R}^3} F_0\rangle| < 22.206 - 21.903$ in $B_2^3(\{\theta\}_{i=1}^3, \delta)$, no global maximum of F may occur in $B_2^3(\{\theta\}_{i=1}^3, \delta)$. Similarly, we claim $|\langle \nabla_{\mathbb{R}^3} F_0(\theta_{12}, \theta_{13}, \theta_{23}), v\rangle| \leq 15$ for $(\theta_{12}, \theta_{13}, \theta_{23})$ in an ℓ_2 ball of radius δ around $\{\theta\}_{i=1}^3 \in \mathbb{R}^3$ with $\theta \approx 1.53796841207904$. Since $F_0(\{\theta\}_{i=1}^3) \approx 21.7391$, once again, no global maximum of F occurs in $B_2^3(\{\theta\}_{i=1}^3, \delta)$.

Proof of claim: for $\theta = \arccos(-1/3)$ or $\theta = 1.537968$, we first show that $\|\text{Hessian}(F_0)\|_2 < 488$ for all points in $B_2^3(\{\theta\}_{i=1}^3, \delta) \subseteq \mathbb{R}^3$. Let $\tilde{F}(x, y, z) :=$

$F_0(x,y,z)$. Then

$$\begin{aligned}
&\tilde{F}_{xx}\\
&= 6 + 4\,\frac{\cos(y)\,y}{\sin(y)} - 2\,\frac{\cos(y)\,xy\cos(x)}{\sin(x)\sin(y)} + 4\,\frac{(\cos(z)-\cos(x)\cos(y))\,xy\,(\cos(x))^2}{(\sin(x))^3\sin(y)}\\
&- 4\,\frac{(\cos(z)-\cos(x)\cos(y))\,y\cos(x)}{(\sin(x))^2\sin(y)}\\
&+ 2\,\frac{(\cos(z)-\cos(x)\cos(y))\,xy}{\sin(x)\sin(y)} - 2\,\frac{\cos(x)\,zy}{\sin(z)\sin(y)}\\
&+ 4\,\frac{\cos(z)\,z}{\sin(z)} - 2\,\frac{\cos(z)\,xz\cos(x)}{\sin(x)\sin(z)} + 4\,\frac{(\cos(y)-\cos(x)\cos(z))\,xz\,(\cos(x))^2}{(\sin(x))^3\sin(z)}\\
&- 4\,\frac{(\cos(y)-\cos(x)\cos(z))\,z\cos(x)}{(\sin(x))^2\sin(z)} + 2\,\frac{(\cos(y)-\cos(x)\cos(z))\,xz}{\sin(x)\sin(z)}.
\end{aligned}$$

$$\begin{aligned}
&\tilde{F}_{xy}\\
&= -2\,xy - 2\,\frac{(\cos(y))^2\,xy}{(\sin(y))^2} + 2\,\frac{\cos(y)\,x}{\sin(y)} - 2\,\frac{(\cos(x))^2\,xy}{(\sin(x))^2}\\
&+ 2\,\frac{(\cos(z)-\cos(x)\cos(y))\,xy\cos(x)\cos(y)}{(\sin(x))^2(\sin(y))^2}\\
&- 2\,\frac{(\cos(z)-\cos(x)\cos(y))\,x\cos(x)}{(\sin(x))^2\sin(y)}\\
&+ 2\,\frac{\cos(x)\,y}{\sin(x)} - 2\,\frac{(\cos(z)-\cos(x)\cos(y))\,y\cos(y)}{\sin(x)(\sin(y))^2} + 2\,\frac{\cos(z)-\cos(x)\cos(y)}{\sin(x)\sin(y)}\\
&+ 2\,\frac{\sin(x)\,zy\cos(y)}{\sin(z)(\sin(y))^2} - 2\,\frac{\sin(x)\,z}{\sin(z)\sin(y)} + 2\,\frac{\sin(y)\,xz\cos(x)}{(\sin(x))^2\sin(z)} - 2\,\frac{\sin(y)\,z}{\sin(x)\sin(z)}.
\end{aligned}$$

With $\delta < 1/100$, $\theta = \arccos(-1/3)$, and $(x,y,z) \in B_2^3(\{\theta\}_{i=1}^3, \delta)$, we then have $1/|\sin(x)| < 1.1$, and similarly for y, z. Therefore, in this δ-ball we have $|\tilde{F}_{xx}| < 228$. Here we have bounded the cosine terms by 1, the x, y, z terms by 2 (since $\arccos(1/3) + \delta < 2$), and the inverted sine terms by 1.1. Using these same bounds, we find that $|\tilde{F}_{xy}| < 117$ in this same δ-ball. For a 3×3 matrix $A = a_{ij}$, let $\|A\|_2 := (\sum_{i,j=1}^3 a_{ij}^2)^{1/2}$. Since $\tilde{F}$ is symmetric in its arguments, we have the following bound in $B_2^3(\{\theta\}_{i=1}^3, \delta)$

$$\|\mathrm{Hessian}(\tilde{F})\|_2 \le \left\| \begin{pmatrix} 228 & 117 & 117 \\ 117 & 228 & 117 \\ 117 & 117 & 228 \end{pmatrix} \right\|_2 < 488,$$

as desired. And the same Hessian estimate applies for $\theta \approx 1.5379684120790425$.

Apply now the Hessian estimate to

$$\|\nabla_{\mathbb{R}^3} f(b)\|_2 \le \|\nabla_{\mathbb{R}^3} f(a)\|_2 + \delta \sup_{\vec{x}\in B_2^3(a,\delta)} \|\mathrm{Hessian}(f)(\vec{x})\|_2,$$

which holds for every sufficiently smooth $f\colon \mathbb{R}^3 \to \mathbb{R}$ if $a, b \in \mathbb{R}^3$ satisfy $\|b-a\|_2 \le \delta$. We then explicitly calculate

$$\|\nabla_{\mathbb{R}^3} F_0|_{\{\arccos(-1/3)\}_{i=1}^3}\|_2 < 13.6,$$

$$\|\nabla_{\mathbb{R}^3} F_0|_{\{1.5379684\}_{i=1}^3}\|_2 < 9.7.$$

So, in $B_2^3(\{\arccos(-1/3)\}_{i=1}^3, \delta)$ we have

$$\|\nabla_{\mathbb{R}^3} F_0\|_2 < 13.6 + 4.88.$$

And in $B_2^3(\{1.5379684\}_{i=1}^3, \delta)$ we have $\|\nabla_{\mathbb{R}^3} F_0\|_2 < 9.7 + 4.88$, proving the claim. □

Lemma 3.4.2. *We may assume that* $\sqrt{\lambda} > .18$.

Proof. Substituting in (3.30), $\det(v_2, v_1, v_4) = \det(v_1, v_2, v_3)\frac{\sin\theta_{12}}{\theta_{12}}\frac{1}{\sqrt{\lambda}}$. Using (3.43),

$$|\det(v_2, v_1, v_4)| > |\det(v_1, v_2, v_3)|\, \frac{\sin(1/2)}{\sqrt{\lambda}\,(\pi - 1/2)}. \tag{3.65}$$

Now, suppose $\sqrt{\lambda} < \sin(1/2)/(\pi - 1/2)$, and $\sqrt{\lambda'} < \sin(1/2)/(\pi - 1/2)$, with λ' defined by

$$\lambda' := \left\| \frac{\sin\theta_{12}}{\theta_{12}} v_4 + \frac{\sin\theta_{24}}{\theta_{24}} v_1 + \frac{\sin\theta_{14}}{\theta_{14}} v_2 \right\|_2^2$$

Applying (3.65) twice (with indices changed appropriately) shows $|\det(v_2, v_1, v_4)| > |\det(v_1, v_2, v_3)|$ and $|\det(v_2, v_1, v_4)| < |\det(v_1, v_2, v_3)|$, a contradiction. We may therefore assume that one triangle of P (which, up to relabeling, is T_4) satisfies $\sqrt{\lambda} > (\sin(1/2)/(\pi - 1/2)) > .18$. □

We now derive some modulus of continuity and derivative estimates for our system (3.64). Recall that $P = \{T_i\}_{i=1}^4$ satisfies Proposition 3.3.1(e). Since (3.64) must hold for the data of each individual triangle T_i, we select in (3.64) the data from the triangle T_4. Then $\theta_{12}+\theta_{23}+\theta_{13} \le 2\pi$, since the perimeter of T_4 is bounded by 2π. By relabeling the edges of T_4 and using the pigeonhole principle, we may assume $\theta_{12} \le 2\pi/3$. The following summary of (3.64) holds for any labeling of the form $\{i, j, \ell\} = \{1, 2, 3\}$.

$$\begin{cases} \theta_{ij} + \theta_{j\ell} > \theta_{i\ell} & \text{, triangle inequality on the sphere} \\ \theta_{12} + \theta_{23} + \theta_{13} \leq 2\pi & \\ \theta_{12} \leq 2\pi/3 & \\ \theta_{13}, \theta_{23} \leq \pi - 1/2 & \text{, (3.43)} \\ \max_{i,j \in \{1,2,3\}, i \neq j} \{\theta_{ij}\} \geq \pi \frac{3}{2\sqrt{14}} & \text{, Corollary 3.3.11} \\ \theta_{ij} \geq 1/10 & \text{, Lemma 3.3.16} \\ \sin \Theta_{ij\ell} \geq 1/10 & \text{, Lemma 3.3.16} \\ \sqrt{\lambda} > .18 & \text{, Lemma 3.4.2} \end{cases} \tag{3.66}$$

For a map $f\colon (X, \|\cdot\|_X) \to (Y, \|\cdot\|_Y)$ between two normed linear spaces, a modulus of continuity for f, denoted by ω_f, satisfies

$$\|f(x_1) - f(x_2)\|_Y \leq \omega_f(\|x_1 - x_2\|_X, x_1, x_2).$$

We allow ω_f to depend on x_1, with x_2 in a δ-ball around some given point $x_1 = (\theta_{12}, \theta_{13}, \theta_{23})$.

Lemma 3.4.3. *Suppose $\delta < 1/100$, and consider $H_1\colon ([0,\pi]^3, \|\cdot\|_\infty) \to \mathbb{R}$, $H_1 = H_1(\theta_{12}, \theta_{13}, \theta_{23})$. Suppose also that $\lambda = \lambda(\theta_{12}, \theta_{13}, \theta_{23}) > \eta$. Then, in an ℓ_∞ ball of radius δ centered at $(\theta_{12}, \theta_{13}, \theta_{23})$, the following holds*

$$\begin{aligned} \omega_{H_1}(\delta) \leq G(\delta, \vec{\theta}, \eta) &:= \delta \Bigg[\frac{7}{2} + \frac{15}{2\sqrt{\eta - 15\delta}} + 3\Bigg(2(\theta_{12} + \delta) + 2(\theta_{13} + \delta) \\ &+ (\theta_{12} + \delta)(\theta_{13} + \delta)\left(\frac{2}{\sin(\theta_{12} + \delta)} + \frac{2}{\sin(\theta_{13} + \delta)} + \frac{1}{\sin(\theta_{12} + \delta)\sin(\theta_{13} + \delta)}\right)\Bigg)\Bigg]. \end{aligned} \tag{3.67}$$

Or, applying Lemma 3.4.2, for δ sufficiently small,

$$\begin{aligned} \omega_{H_1}(\delta) \leq \delta &\Bigg[\frac{7}{2} + \frac{15}{2\sqrt{.0329 - 15\delta}} + 3\Bigg(2(\theta_{12} + \delta) + 2(\theta_{13} + \delta) \\ &+ (\theta_{12} + \delta)(\theta_{13} + \delta)\left(\frac{2}{\sin(\theta_{12} + \delta)} + \frac{2}{\sin(\theta_{13} + \delta)} + \frac{1}{\sin(\theta_{12} + \delta)\sin(\theta_{13} + \delta)}\right)\Bigg)\Bigg]. \end{aligned} \tag{3.68}$$

Proof. We want to compute the modulus for the following function from $(\mathbb{R}^3, \|\cdot\|_\infty)$ to $\mathbb{R}$

$$H_1(\theta_{12}, \theta_{13}, \theta_{23}) = \sqrt{\lambda} \cos\left(\sqrt{\theta_{12}^2 + \theta_{13}^2 + 2\theta_{12}\theta_{13}\cos\Theta_{213}}\right) + \gamma_1$$

where λ and γ_1 are defined in (3.32) and (3.63) respectively. In what follows, we use the product rule, chain rule, and sum rule for the modulus $\omega_{(\cdot)}$.

$$\begin{aligned}\omega_{H_1}(\delta) &\leq \omega_{\sqrt{\lambda}}(\delta) + \|\sqrt{\lambda}\|_\infty \cdot \omega_{\cos\left(\sqrt{\theta_{12}^2+\theta_{13}^2+2\theta_{12}\theta_{13}\cos\Theta_{213}}\right)}(\delta) + \omega_{\gamma_1}(\delta) \\ &\leq \omega_{\sqrt{\lambda}}(\delta) + 3\cdot\frac{1}{2}\cdot\omega_{(\theta_{12}^2+\theta_{13}^2+2\theta_{12}\theta_{13}\cos\Theta_{213})}(\delta) + \frac{7}{2}\delta.\end{aligned}$$

In these inequalities, we used $\lambda \leq 9$, $\sqrt{\lambda} \leq 3$, $\left|\frac{d}{dt}(\cos(\sqrt{t}))\right| \leq 1/2$, $|(\sin t)/t| \leq 1$, and $|\frac{d}{dt}((\sin t)/t) \leq 1/2|$. Also, the term $(7/2)\delta$ appears since $\delta^{-1}\omega_{\gamma_1}(\delta) \leq \|\nabla_{\mathbb{R}^3}\gamma_1(\theta_{12},\theta_{13},\theta_{23})\|_1 \leq 7/2$. Now, observe that

$$\begin{aligned}\frac{\partial\lambda}{\partial\theta_{12}} &= 2\frac{\sin\theta_{12}}{\theta_{12}}\left(\frac{d}{dx}\frac{\sin x}{x}\bigg|_{x=\theta_{12}}\right) + 2\frac{\sin\theta_{13}\sin\theta_{23}}{\theta_{13}\theta_{23}}(-\sin\theta_{12}) \\ &\quad + 2\left(\frac{d}{dx}\frac{\sin x}{x}\bigg|_{x=\theta_{12}}\right)\frac{\sin\theta_{13}}{\theta_{13}}\cos\theta_{23} + 2\left(\frac{d}{dx}\frac{\sin x}{x}\bigg|_{x=\theta_{12}}\right)\frac{\sin\theta_{23}}{\theta_{23}}\cos\theta_{13}.\end{aligned}$$

So, taking absolute values, and using that $|\sin(t)/t| \leq 1$ and $\left|\frac{d}{dt}(\sin(t)/t)\right| \leq 1/2$, we see that $|\partial\lambda/\partial\theta_{ij}| \leq 5$. So, if $\lambda(\vec{\theta}) = \lambda(\theta_{12},\theta_{13},\theta_{23}) > \eta > 0$ for some η, and if $\vec{x} = (x,y,z)$ satisfies $|x-\theta_{12}| < \delta, |y-\theta_{13}| < \delta, |z-\theta_{23}| < \delta$, then at $\vec{x}$,

$$\begin{aligned}\left|\frac{\partial(\sqrt{\lambda})}{\partial\theta_{12}}\right| &\leq \max_{\left\{\vec{\theta}\in\mathbb{R}^3\colon\, \|\vec{\theta}-\vec{x}\|_\infty<\delta\right\}} \frac{1}{2}\frac{1}{\sqrt{\lambda(\vec{\theta})}}\cdot\frac{\partial\lambda(\vec{\theta})}{\partial\theta_{12}} \\ &\leq \frac{1}{2}\frac{1}{\sqrt{\eta - \|\nabla_{\mathbb{R}^3}\lambda\|_1\delta}}\cdot 5 \leq \frac{1}{2}\frac{1}{\sqrt{\eta-15\delta}}\cdot 5.\end{aligned}$$

Therefore, for such $\vec{x}$,

$$\|\nabla_{\mathbb{R}^3}(\sqrt{\lambda})\|_1 \leq \frac{15}{2\sqrt{\eta-15\delta}}.$$

Combining this estimate with our estimate above, we get the following modulus estimate, that applies to $\vec{x},\vec{\theta}$ as just described

$$\begin{aligned}\omega_{H_1}(\delta) \leq \delta\Bigg[\frac{7}{2} + \frac{15}{2\sqrt{\eta-15\delta}} + 3\Bigg(&2(\theta_{12}+\delta) + 2(\theta_{13}+\delta) \\ + (\theta_{12}+\delta)(\theta_{13}+\delta)&\left(\frac{2}{\sin(\theta_{12}+\delta)} + \frac{2}{\sin(\theta_{13}+\delta)} + \frac{1}{\sin(\theta_{12}+\delta)\sin(\theta_{13}+\delta)}\right)\Bigg)\Bigg].\end{aligned}$$

We used here the monotonicity of the function $x \mapsto x/\sin x$ and the following identities for $\psi(\theta_{12},\theta_{13},\theta_{23}) := (\theta_{12}^2+\theta_{13}^2+2\frac{\theta_{12}\theta_{13}}{\sin\theta_{12}\sin\theta_{13}}(\cos\theta_{23}-\cos\theta_{12}\cos\theta_{13}))$,

$$\begin{aligned}
\frac{1}{2}\frac{\partial\psi}{\partial\theta_{12}} &= \theta_{12}+\frac{\theta_{12}\theta_{13}}{\sin\theta_{12}\sin\theta_{13}}(\sin\theta_{12}\cos\theta_{13})\\
&\qquad+(\cos\theta_{23}-\cos\theta_{12}\cos\theta_{13})(\theta_{12}/\sin\theta_{12})'\cdot(\theta_{13}/\sin\theta_{13})\\
&= \theta_{12}+\frac{\theta_{12}\theta_{13}}{\sin\theta_{13}}\cos\theta_{13}+\cos\Theta_{213}\left(1-\frac{\theta_{12}}{\tan\theta_{12}}\right)\theta_{13},\\
\frac{1}{2}\frac{\partial\psi}{\partial\theta_{13}} &= \theta_{13}+\frac{\theta_{12}\theta_{13}}{\sin\theta_{12}\sin\theta_{13}}(\sin\theta_{13}\cos\theta_{12})\\
&\qquad+(\cos\theta_{23}-\cos\theta_{12}\cos\theta_{13})(\theta_{13}/\sin\theta_{13})'\cdot(\theta_{12}/\sin\theta_{12})\\
&= \theta_{13}+\frac{\theta_{12}\theta_{13}}{\sin\theta_{12}}\cos\theta_{12}+\cos\Theta_{213}\left(1-\frac{\theta_{13}}{\tan\theta_{13}}\right)\theta_{12},\\
\frac{1}{2}\frac{\partial\psi}{\partial\theta_{23}} &= \frac{\theta_{12}\theta_{13}}{\sin\theta_{12}\sin\theta_{13}}(-\sin\theta_{23}).
\end{aligned}$$

□

Analogous estimates give a gradient bound for F_0, as defined in (3.42). Since F_0 is symmetric in its arguments, it suffices to take a derivative with respect to one variable. Fix $(\theta_{12},\theta_{13},\theta_{23})$, and let $0<\delta<1/100$. For any (x,y,z) such that $|x-\theta_{12}|<\delta, |y-\theta_{13}|<\delta, |z-\theta_{23}|<\delta$ we have the following derivative bound

$$\begin{aligned}
\left|\frac{\partial F_0}{\partial\theta_{12}}\right| &\le 6(\theta_{12}+\delta)+\frac{2(\theta_{13}+\delta)(\theta_{23}+\delta)}{\sin(\theta_{13}+\delta)\sin(\theta_{23}+\delta)}+2\frac{(\theta_{12}+\delta)(\theta_{13}+\delta)}{\sin(\theta_{13}+\delta)}\\
&\quad+2\frac{(\theta_{12}+\delta)(\theta_{23}+\delta)}{\sin(\theta_{23}+\delta)}+2(\theta_{13}+\theta_{23}+2\delta)\left(1-\frac{\theta_{12}+\delta}{\tan(\theta_{12}+\delta)}\right)
\end{aligned} \tag{3.69}$$

The restrictions of (3.66) and the bounds of (3.67) and (3.69) finally yield Theorem 3.1.1.

Proof of Theorem 3.1.1. Let P be a partition of the sphere S^2 that maximizes $\sum_{i=1}^4\|z_i\|_2^2$. Arguing by contradiction, we may assume that this partition exceeds the value of the spherical propeller partition, $9\pi^2/4$. From Proposition 3.3.1 we may assume that P is a collection of four nonempty spherical triangles along with four vertices such that $\det(v_1,v_2,v_3)>0$. Since P is a global maximum of F, it satisfies three equations in the three variables $\{\theta_{12},\theta_{13},\theta_{23}\}$ defined by (3.64). We may restrict these variables via (3.66). We may also exclude neighborhoods of the two candidate maxima found in Section 3.3, using Lemma 3.4.1. Combining the estimates of (3.67),(3.68) and (3.69) gives sufficient information for an ε-net traversal of the parameter space defined by (3.66). This numerical computation finds that no zeros of the system (3.64), within the parameter space defined by

(3.66), exceed the objective value of the propeller partition. This computation contradicts the assumed existence of a maximal P with four nonempty elements. So, recalling (3.7), $\sum_{i=1}^{4}\|z_i\|_2^2 \leq 9\pi^2/4 = 2\pi^3(9/(8\pi))$. □

3.4.1 Implementation

We describe the ε-net traversal of Theorem 3.1.1 further below. Before doing so, we describe some issues related to the numerical verification of inequalities. In particular, we discuss how a computer can rigorously verify a mathematical statement. Recall that a normal double precision floating point number in the IEEE-754-2008 standard is a number in base two of the form

$$\pm 1.a_1a_2a_3\cdots a_{52} \times 2^{b_1b_2\cdots b_{11}-2^{10}},$$

where $a_i \in \{0,1\}$ and $b_1b_2\cdots b_{11} \in [2, 2^{11}-1]\cap\mathbb{Z}$. All of our computations are done with this discrete set of numbers, with zero included as well. Note that the distance between two consecutive numbers changes according to the values of the two numbers. For example, the distance between 1 and the next largest number is 2^{-52}, whereas the distance between 2 and the next largest number is 2^{-51}. (In general, if the number 2^k and the next largest number can be represented in this system, then their distance is 2^{-52+k}.) Due to this spacing between numbers, one must round the result of any arithmetic operation. For example, the addition $1+2^{-54}$ evaluates to 1, if we choose to round to the nearest double precision number. For the sake of flexibility, our analysis below applies regardless of the rounding scheme that is chosen.

For $x, y \in [0, 2^{500}]$ where x, y are normal double precision numbers, let $\mathrm{fl}(x+y)$ denote the normal double precision number that is evaluated in a computer with the IEEE-754-2008 standard. It suffices for now to note that $(x+y) \leq (1+2\cdot 2^{-52})\mathrm{fl}(x+y)$. In this case, we say that $(x+y)$ has multiplicative error bounded by $(1+2\cdot 2^{-52})$. When we speak of numerical error, we mean multiplicative error, unless otherwise stated. Define

$$\varepsilon := 3\cdot 2^{-52}.$$

To verify inequalities in our numerical computations, we perform operations for several floating point numbers and keep track of error bounds of the form $(x+y) \leq (1+2\cdot 2^{-52})\mathrm{fl}(x+y)$. Since $(1+2\cdot 2^{-52})^N \leq (1+N\cdot 3\cdot 2^{-52}) = (1+N\varepsilon)$ for $1 \leq N \leq 30000$, we can bound the multiplicative error of N operations with a multiplicative term of the form $(1+N\varepsilon)$. (Here and below we are overly conservative in our error estimates.)

We now present the formulas that are used in our computations, delaying further discussion of numerical errors. Let $x, y > 0$ be two normal double precision floating point numbers with $x > y$. A loss of significance refers to a loss of binary significant digits in the computation of $\mathrm{fl}(x-y)$. One can show that $\mathrm{fl}(x-y)$ loses at most q and at least p significant binary digits, where $2^{-q} \le 1-(y/x) \le 2^{-p}$. Due to potential loss of significance errors, we evaluate λ as

$$\begin{aligned}\lambda = &\left(\frac{\sin^2\theta_{23}}{\theta_{23}^2} + \frac{\sin^2\theta_{13}}{\theta_{13}^2} + \frac{\sin^2\theta_{12}}{\theta_{12}^2}\right) + 2\left(\frac{\sin\theta_{23}\sin\theta_{13}}{\theta_{23}\theta_{13}}\cos\theta_{12}\right.\\ &\left.+ \frac{\sin\theta_{23}\sin\theta_{12}}{\theta_{23}\theta_{12}}\cos\theta_{13} + \frac{\sin\theta_{12}\sin\theta_{13}}{\theta_{12}\theta_{13}}\cos\theta_{23}\right).\end{aligned} \tag{3.70}$$

Our numerical computation of inequalities consists of four steps.
(Step (i)) Suppose we want to check (3.66) for a closed ℓ_∞ ball (i.e. a cube) of radius δ with center $(\theta_{12}, \theta_{13}, \theta_{23}) \in [0,\pi]^3$. To do so, we must verify the following perturbation of (3.66).

$$\begin{cases}(\theta_{ij} + \theta_{j\ell} + 3\delta) \ge \theta_{i\ell}\\ \theta_{12} + \theta_{23} + \theta_{13} \le (2\pi + 3\delta)\\ \theta_{12} \le (2\pi/3 + \delta)(1+10\varepsilon)\\ \theta_{13} + 1/2, \theta_{23} + 1/2 \le (\pi+\delta)(1+10\varepsilon)\\ (\delta + \max_{i,j\in\{1,2,3\}, i\neq j}\{\theta_{ij}\})(1+7\varepsilon) \ge \pi\frac{3}{2\sqrt{14}}(1-5\varepsilon)\\ (\theta_{ij}+\delta)(1+7\varepsilon) \ge (1/10)(1-\varepsilon)\\ -(3\delta + \frac{\sqrt{99}}{10}(\sin\theta_{ij}\sin\theta_{j\ell} + 2\delta))(1+1500\varepsilon) \le (\cos\theta_{i\ell} - \cos\theta_{ij}\cos\theta_{j\ell} + 2000\varepsilon)\\ (\cos\theta_{i\ell} - \cos\theta_{ij}\cos\theta_{j\ell} - 2000\varepsilon) \le (3\delta + \frac{\sqrt{99}}{10}(\sin\theta_{ij}\sin\theta_{j\ell} + 2\delta))(1+1500\varepsilon)\\ (15\delta + 2\cdot 10^4\varepsilon + \lambda) \ge .03293\end{cases} \tag{3.71}$$

If a given inequality of (3.71) does not hold, then the corresponding inequality of (3.66) is violated for the entire ℓ_∞ ball of radius δ.
(Step (ii)) If all of the inequalities of (3.71) are satisfied, we then check if $(\theta_{12}, \theta_{13}, \theta_{23})$ lies in a neighborhood of two of our known candidate global optima. Let $\vec{x} = (x_1, x_2, x_3) \in [0,\pi]^3$. Suppose $x_1 = x_2 = x_3 = \arccos(-1/3)$ or $x_1 = x_2 = x_3 = 1.5379684120790425$. From Lemma 3.4.1, we know that if $\vec{\theta} = (\theta_{12}, \theta_{13}, \theta_{23})$ is close to $\vec{x}$, then F is bounded by $(9/4)\pi^2$ in $\overline{B_\infty^3}((\theta_{12}, \theta_{13}, \theta_{23}), \delta)$. That is, if the following is not satisfied, then F is bounded as just stated.

$$\begin{aligned}\left[(d_{\ell_2}(\vec{\theta}, \vec{x}))^2 + 40000\varepsilon + 2\delta\left(d_{\ell_1}(\vec{\theta}, \vec{x}) + 6000\varepsilon + 3\delta\right)\right](1+20\varepsilon)\\ > ((1/100)^2)(1-2\varepsilon).\end{aligned} \tag{3.72}$$

Here d_{ℓ_2} and d_{ℓ_1} denote the usual ℓ_2 and ℓ_1 metrics on $\mathbb{R}^3$.
(Step (iii)) If the verification of (3.72) fails, we use the modulus estimates of Lemma 3.4.3 to see if H is near zero at $\vec{\theta}$, with $H = (H_1, H_2, H_3)$ defined after

(3.64). Let $\alpha, \beta, \rho, \sigma \in \{\pm 1\}$. We first check if the following expression has constant sign, over the 2^4 such choices of $\alpha, \beta, \rho, \sigma$.

$$\left[\cos\left(\sqrt{\theta_{12}^2 + \theta_{13}^2 + 2\theta_{12}\theta_{13}\left(\frac{\cos\theta_{23} - \cos\theta_{12}\cos\theta_{13} + 2000\varepsilon\alpha}{\sin\theta_{12}\sin\theta_{13}}\right)}\right) + 25000\varepsilon\beta\right] \cdot \sqrt{\lambda - 2\cdot 10^4\varepsilon\sigma} + (\gamma_1 + 3000\varepsilon\rho). \tag{3.73}$$

If (3.73) is well-defined and its sign is constant, we then consider the following inequalities involving H_i for $i = 1$ and (3.67).

$$\min_{\alpha,\beta,\rho,\sigma\in\{\pm 1\}} \left|\left[\cos\left(\sqrt{\theta_{12}^2 + \theta_{13}^2 + 2\theta_{12}\theta_{13}\left(\frac{\cos\theta_{23} - \cos\theta_{12}\cos\theta_{13} + 2000\varepsilon\alpha}{\sin\theta_{12}\sin\theta_{13}}\right)}\right) + 25000\varepsilon\beta\right] \cdot \sqrt{\lambda - 2\cdot 10^4\varepsilon\sigma} + (\gamma_1 + 3000\varepsilon\rho)\right|(1 - 3000\varepsilon) > (1 + 3000\varepsilon)G(\delta, \vec{\theta}, \lambda - 2\cdot 10^4\varepsilon). \tag{3.74}$$

$$\begin{aligned}\min_{\alpha,\beta,\rho,\sigma\in\{\pm 1\}} &\left|\left[\cos\left(\sqrt{\theta_{12}^2 + \theta_{13}^2 + 2\theta_{12}\theta_{13}\left(\frac{\cos\theta_{23} - \cos\theta_{12}\cos\theta_{13} + 2000\varepsilon\alpha}{\sin\theta_{12}\sin\theta_{13}}\right)}\right) + 25000\varepsilon\beta\right] \cdot \sqrt{\lambda - 2\cdot 10^4\varepsilon\sigma} + (\gamma_1 + 3000\varepsilon\rho)\right|(1 - 3000\varepsilon) \\ &> (1 + 5\varepsilon)\sqrt{15\delta} + \delta(1 + 3000\varepsilon)\left[\frac{7}{2} + 3\Big(2(\theta_{12} + \delta) + 2(\theta_{13} + \delta) \right. \\ &\left. + (\theta_{12} + \delta)(\theta_{13} + \delta)\left(\frac{2}{\sin(\theta_{12} + \delta)} + \frac{2}{\sin(\theta_{13} + \delta)} + \frac{1}{\sin(\theta_{12} + \delta)\sin(\theta_{13} + \delta)}\right)\Big)\right].\end{aligned} \tag{3.75}$$

The inequalities for H_2 and H_3 are constructed by permuting cyclically the indices appearing above. If at least one of (3.74) and (3.75) is satisfied (for at least one H_i), then H_i is nonzero on $\overline{B^3_\infty}((\theta_{12}, \theta_{13}, \theta_{23}), \delta)$.

(Step (iv)) If (3.74) and (3.75) are not satisfied for $\{H_i\}_{i=1}^3$, we finally check the value of F_0 directly, using a modification of (3.42). Define

$$\begin{aligned}
\cos\Theta_{ijk}^{\varepsilon} &:= \left(\frac{\cos\theta_{ik} - \cos\theta_{ij}\cos\theta_{jk} + 2000\varepsilon}{\sin\theta_{ij}\sin\theta_{jk}}\right) \\
F_0^{\varepsilon} &:= 3\left(\theta_{12}^2 + \theta_{23}^2 + \theta_{13}^2\right) + 2\cos(\Theta_{213}^{\varepsilon})\theta_{12}\theta_{13} \\
&\qquad + 2\cos(\Theta_{123}^{\varepsilon})\theta_{12}\theta_{23} + 2\cos(\Theta_{231}^{\varepsilon})\theta_{23}\theta_{13}, \\
G_{12}^{\varepsilon}(\vec{\theta},\delta) &:= 6(\theta_{12}+\delta) + \frac{4(\theta_{13}+\delta)(\theta_{23}+\delta)}{\sin(\theta_{13}+\delta)\sin(\theta_{23}+\delta)} + 2\frac{(\theta_{12}+\delta)(\theta_{13}+\delta)}{\sin(\theta_{13}+\delta)} \\
&\quad + 2\frac{(\theta_{12}+\delta)(\theta_{23}+\delta)}{\sin(\theta_{23}+\delta)} \\
&\quad + 2(\theta_{13}+\theta_{23}+2\delta)\left[1 - \left(\frac{\theta_{12}+\delta}{\tan(\theta_{12}+\delta)}(1+1200\varepsilon) + \varepsilon\right)\right].
\end{aligned}$$

If the following is satisfied, then F_0 is bounded by $(9/4)\pi^2$ in $\overline{B_\infty^3}((\theta_{12},\theta_{13},\theta_{23}),\delta)$, and therefore F satisfies the same bound for any global maximum of F in this ball.

$$(9/4)\pi^2(1-5\varepsilon) - F_0^{\varepsilon}\cdot(1+5000\varepsilon) > \delta(G_{12}^{\varepsilon} + G_{13}^{\varepsilon} + G_{23}^{\varepsilon})(1+10^4\varepsilon) + 100\varepsilon. \quad (3.76)$$

Given these estimates, we can now describe the ε-net traversal of Theorem 3.1.1. We need a discrete subset $\mathcal{N} = \{x_i\} \subseteq [0,\pi]^3$ together with a set of radii $\delta_i > 0$ such that

$$\bigcup_{x_i\in\mathcal{N}} \overline{B_\infty^3}(x_i,\delta_i) \supseteq [0,\pi]^3,$$

and such that one of the following cases occurs, for (x_i,δ_i) with $x_i \in \mathcal{N}$.

(i) One inequality of (3.71) does not hold.

(ii) (3.72) does not hold.

(iii) (3.74) or (3.75) holds, and (3.73) has constant sign over all sign choices $\alpha,\beta,\sigma,\rho \in \{\pm 1\}$.

(iv) (3.76) holds.

In case (i), $\overline{B_\infty^3}(x_i,\delta_i)$ is either outside of the domain of (3.64) or (3.64) cannot be satisfied in $\overline{B_\infty^3}(x_i,\delta_i)$. Case (ii) implies that $F = \sum_{i=1}^4 \|z_i\|_2^2$ is bounded by $(9/4)\pi^2$ on $\overline{B_\infty^3}(x_i,\delta_i)$. Cases (iii) and (iv) imply that no global maximum of F occurs in $\overline{B_\infty^3}(x_i,\delta_i)$. So, the existence of $\mathcal{N}$ as defined above completes the proof of Theorem 3.1.1.

To construct $\mathcal{N}$, we perform the following standard procedure. Begin with the grid

$$\mathcal{N}_0 := \frac{\pi}{100}([0,100]^3 \cap \mathbb{Z}^3) \subseteq [0,\pi]^3.$$

For each point $x \in \mathcal{N}_0$, check the inequalities discussed above with $\delta = \delta_0 := \pi/200$. If we discover that x satisfies one of (i) through (iv), then eliminate x from the set $\mathcal{N}_0$. Next, include the element (x, δ) in the set $\mathcal{N}$. Let $\mathcal{N}_0'$ be the set of all points $x \in \mathcal{N}_0$ that have not been included in $\mathcal{N}$. We now "refine" the grid $\mathcal{N}_0'$ by a factor of 10 and repeat the above procedure. That is, we define

$$\mathcal{N}_1 := \Big(\frac{\pi}{1000}([0, 1000]^3 \cap \mathbb{Z}^3)\Big) \cap \left(\bigcup_{x_i \in \mathcal{N}_0'} \overline{B_\infty^3}(x_i, \pi/200)\right).$$

For each point $x \in \mathcal{N}_1$, we now check (i)-(iv) with $\delta = \delta_1 := \pi/2000$, and so on. For $j \geq 1$, at the j^{th} iteration of this algorithm, we define

$$\mathcal{N}_j := \Big(\frac{\pi}{10^{j+2}}([0, 10^{j+2}]^3 \cap \mathbb{Z}^3)\Big) \cap \left(\bigcup_{x_i \in \mathcal{N}_{j-1}'} \overline{B_\infty^3}(x_i, \pi/(2 \cdot 10^{j+1}))\right).$$

We continue in this manner until the procedure terminates. Upon termination at the j^{th} step, $\mathcal{N}_j' = \varnothing$, so $\mathcal{N}$ satisfies our desired properties and proves Theorem 3.1.1.

We now return to our discussion of numerical error. To keep track of rounding errors resulting from subtractions, we use absolute error where appropriate and add the ε terms that correspond to absolute numerical errors. These additive terms are scaled according to the absolute value of the numbers being subtracted (due to the unequal spacing of floating point numbers described above).

In the description and analysis of (i) through (iv), δ terms come from modulus of continuity estimates, and ε terms account for numerical error. The latter is discussed further below. The first two inequalities of (3.71) contain no ε terms, since they can be checked in exact arithmetic. Let $x \in [0, \pi - 1/2 + 2 \times 10^{-2}]$. (We make this restriction due to the fourth inequality of (3.71).) Let us discuss the numerical computation of the sine function. If $x > \pi/2$, we replace x with $\pi - x$, and we evaluate $\sin(\pi - x)$. Thus, we only need to consider $x \in [0, \pi/2]$. For such x, we first use a 19^{th} degree (ten term) Taylor expansion around zero, summing terms in ascending degree. In perfect arithmetic, we get a multiplicative error of one plus

$$\left|\frac{x^{21}}{(21!)\sin(x)}\right| = \left|\frac{x}{\sin(x)}\frac{x^{20}}{(21!)}\right| \leq \frac{\pi}{2}\frac{(\pi/2)^{20}}{21!} < 2.6 \times 10^{-16} < \frac{\varepsilon}{2}.$$

Since $x \in [0, \pi/2]$, there are no loss of significance errors in the summation of the series. The addition of each term of the Taylor series involves one operation, and there are ten such terms. Moreover, the n^{th} term contains $n+1$ operations. In total, one evaluation therefore has less than 210 operations. So, very conservatively, the multiplicative error of $\sin(x)$ is bounded by $(1+250\varepsilon)$, if x is represented exactly. If the argument x has multiplicative error $(1+k\varepsilon)$, then this same analysis shows that the multiplicative error of $\sin(x)$ is bounded by $(1+(250+100k)\varepsilon)$, since the term x appears 100 times in the formula for sin. For example, since θ_{ij} has multiplicative error $(1+3\varepsilon)$, we conservatively bound the error of $\sin\theta_{ij}$ by $(1+550\varepsilon)$.

To compute $\cos(x)$ for $x \in [0, \pi-1/2+2\times10^{-2}]$, we instead compute $\sin(\pi/2-x)$ with the procedure outlined above. Due to our representation of the points θ_{ij}, we compute $\pi/2-x$ as a rational multiple of π, so no loss of significance errors occurs in this subtraction. Hence, the error estimates for $\cos\theta_{ij}$ are exactly the ones already used for $\sin\theta_{ij}$. However, when evaluating the cosine of more complicated expressions, we need to revert to absolute errors.

We will also need to use the square root function. Here we use the usual Newton's Method iteration, which is well known to have multiplicative error bounded by $(1+\varepsilon)$.

We now describe the error terms appearing in (3.71), which are perturbations of (3.66). The guiding principle is that we want to perturb each inequality so that it is more easily satisfied. To see why this is done, suppose an inequality of (3.71) is not satisfied. Then, a modulus of continuity estimate and a numerical error estimate, used in the definition of (3.71), show that the corresponding inequality of (3.66) is not satisfied in an ℓ_∞ ball of radius δ. For example, contrast the first inequality of (3.66) with that of (3.71). We need only add an additional δ term for each θ_{ij} term, since the modulus of continuity of the function $(x, y, z) \mapsto x+y+z$, viewed as a mapping from $(\mathbb{R}^3, \|\cdot\|_\infty) \to \mathbb{R}$, is exactly 3δ.

So, if the inequality of (3.71) is violated at a given point $\vec{\theta} = (\theta_{12}, \theta_{13}, \theta_{23})$, then the modulus of continuity bound shows that the corresponding inequality of (3.66) is violated in the ball $\overline{B^3_\infty}(\vec{\theta}, \delta)$. The same considerations apply to the second inequality of (3.71). Note that the first and second inequality are checked in exact arithmetic in our computer implementation, so we do not need to take into account numerical error. As we have mentioned above, our θ_{ij} and δ are represented as numbers of the form $\pi(a/c)$ where a and c are integers, and it suffices to perform additions of the form $\pi((a+b)/c)$, and the integer addition $a+b$ can be performed with no error.

The third inequality of (3.71) only requires a modulus of continuity term of δ. However, we must now take into account numerical errors. We count ten numerical operations on positive numbers (including the error in representing π as a floating point number three times), so in the worst case, the multiplicative error is bounded by $(1+10\varepsilon)$. Similar considerations apply in the next three inequalities. In the seventh and eighth inequalities, we use our analysis of the sine function, and recall that the multiplicative error for these operations is always bounded by $1+550\varepsilon$. So, in total, the evaluation of the term $-(3\delta+\frac{\sqrt{99}}{10}(\sin\theta_{ij}\sin\theta_{j\ell}+2\delta))$ has multiplicative error bounded by $1+1500\varepsilon$. Here we have also included δ terms corresponding to the moduli of continuity, as before.

Unfortunately, we cannot use multiplicative error to deal with the cosine terms, since loss of significance can occur in the evaluation of $(\cos\theta_{i\ell} - \cos\theta_{ij}\cos\theta_{j\ell})$. However, since the cosine terms are bounded by 1 in absolute value, the absolute error from each cosine evaluation is bounded by 550ε. Therefore the total absolute error from the right hand side of the seventh inequality is bounded by 2000ε. This bound applies similarly to the penultimate inequality of (3.71). The final inequality of (3.71) is similar to the previous cases. To get the δ terms, we examine (3.70) and use that $|(\sin(x)/x)'| < 1/2$ and $|\sin(x)/x| \leq 1$ for $x \in [0,\pi]$. To get the ε terms, note that each term $\sin^2\theta_{ij}/\theta_{ij}^2$ has absolute error bounded by 1200ε. And each term $\sin\theta_{ij}\sin\theta_{jk}\cos\theta_{ik}/\theta_{ij}\theta_{jk}$ has absolute error bounded by 1700ε. The total absolute error of λ is therefore bounded by $2\cdot 10^4\varepsilon$.

We now treat (3.72). First, recall that (3.17) yielded our two known zeros of the system (3.64). The first zero $\theta_{ij} = \theta = \arccos(-1/3)$ is known exactly (up to multiplicative error $(1\pm\varepsilon)$). However, we can only find the second zero computationally. For the purposes of our computations, we therefore show that this second zero is contained in an interval of the form $\theta_{ij} = \theta \in [x-2000\varepsilon, x+2000\varepsilon]$ where $x = 1.5379684120790425$. To see this, we take (3.17) and write it as a quotient of the left side over the right side, and we wish to find where the quotient is 1. For $\theta \in [x-2000\varepsilon, x+2000\varepsilon]$, the multiplicative error involved in calculating this quotient is bounded by $1+3000\varepsilon$. This follows by our analysis of the sine and cosine functions, and since this particular region of parameters avoids loss of significance errors. We can then check numerically that, for $\theta = x - 2000\varepsilon$, this quotient, multiplied by $(1+3000\varepsilon)$ is less than 1, while for $\theta = x+2000\varepsilon$, the quotient, multiplied by $(1-3000\varepsilon)$, is greater than 1. So, by our numerical analysis and continuity of the given function, a solution of (3.17) must lie in the interval $\theta \in [x-2000\varepsilon, x+2000\varepsilon]$.

We can now confront the errors of (3.72). Once again, the guiding principle is to add error terms so that the inequality is more easily satisfied. First, the term $d_{\ell_2}(\vec{\theta},\vec{x})^2$ is the sum of three squared differences. Due to numerical error in the operations, and the uncertainty of the zero $\theta_{ij} \approx 1.5379684120790425$, each squared difference has an absolute error bounded by $2\pi(2000\varepsilon)+100\varepsilon$. (The factor of 2π comes from the squaring of the difference, and since $|x| \leq \pi$.) So, in total, the absolute error from the term $d_{\ell_2}(\vec{\theta},\vec{x})^2$ is bounded by 40000ε. Since we want to check this inequality in an ℓ_∞ ball of radius δ, we use the modulus of continuity of $d_{\ell_2}^2$, which is $2\delta d_{\ell_1}$. We therefore add the term $2\delta d_{\ell_1}(\vec{\theta},\vec{x})$, and again add terms to this quantity to take into account its own absolute error and its own modulus of continuity. The combined effect of summing the error terms then accumulates a multiplicative error bounded by $1+20\varepsilon$.

We now explain the error terms of (3.74). From the analysis of the cosine, we know that $\cos\theta_{ij}$ for $\theta_{ij} \in [0, \pi - 1/2 + 2\times 10^{-2}]$ has an absolute error bounded by 550ε. (This follows since the multiplicative error is bounded by $(1 + 550\varepsilon)$, and $|\cos(\cdot)| \leq 1$.) So, the expression $\cos\theta_{23} - \cos\theta_{12}\cos\theta_{13}$ has absolute error bounded by 2000ε. The remaining operations under the square root of (3.74) all have multiplicative error bounded by $(1 + 1200\varepsilon)$.

When the cosine of the square root term is evaluated, we need to again take into account absolute errors (since, unlike in the $\cos\theta_{ij}$ terms, we must now account for loss of significance). Since the argument of the cosine term is bounded by π, we can bound the absolute error of the square root term by 4000ε. To take the cosine of the square root term, we need to return to our Taylor series analysis. Note that our Taylor series has Lipschitz constant bounded by 6. (This is a conservative bound that does not account for cancelations in the series summation.) Therefore, the total absolute error of the cosine of the square root in (3.74) is bounded by 25000ε. From before, the absolute error in the computation of λ in (3.70) is bounded by $2 \cdot 10^4\varepsilon$. Similarly, the computation of γ_1 has absolute error bounded by 3000ε.

Temporarily ignoring any numerical errors, note that we are computing a bound for the absolute value of the derivative of a function. So, in order to conclude anything from (3.74), the sign of (3.73) must be constant over all possible numerical errors. (If not, it may be that (3.73) has an actual value of zero.) Given that this sign is constant, we then take the minimum over all possible numerical errors. It then suffices to take the minimum over the extreme points of a suitable rectangle, due to monotonicity of (3.74) with respect to the error terms. Note that we take the minimum since we wish to make the inequality more difficult to be satisfied. Finally, the addition of all terms on the left side of the inequality incurs a multiplicative error bounded by $1 + 3000\varepsilon$, and the addition of the terms on the right side has multiplicative error bounded by $1 + 3000\varepsilon$. A similar analysis applies to (3.75).

We now briefly mention the analysis of (3.76), which is routine at this stage. The principle that we apply here is to make (3.76) more difficult to be satisfied, so we make its right side larger and its left side smaller. The multiplicative error of the F_0^ε term is bounded by $1 + 5000\varepsilon$, since each $\cos\Theta_{ijk}^\varepsilon$ term has multiplicative error bounded by $1 + 1200\varepsilon$. Each G_{ij}^ε term has multiplicative error bounded by $1 + 2500\varepsilon$, so the right side of (3.76) has multiplicative error bounded by $1 + 10^4\varepsilon$. A 100ε term is added to the right side of (3.76) to account for the error of the subtraction on the left side. The $1 + 1200\varepsilon$ term and the ε term in the definition of G_{ij}^ε account for the errors of the tangent function, and the loss of significance in the subtraction that occurs in the square brackets.

3.5 Discrete Harmonic Maps into the Sphere

We now summarize and reinterpret the results of Section 3.3. Recall that we defined $G(P) = \sum_{1 \leq i < j \leq 4} \theta_{ij}^2$ as the sum of the squared lengths of the edges of our partition P. We then saw from Propositions 3.3.13 and 3.3.1 that all maxima of F are also critical points of G. We also found that critical points of G satisfy (3.50). Given a partition P, think of its edges as rubber bands. (Recall that a rubber band has energy given by its length squared, and it exerts a force proportional to its length.) The critical points of G are given by configurations of rubber bands (arranged like the edges of Figure 3.2(b)) that are at equilibrium with respect to the forces exerted by the rubber bands. With this interpretation, (3.50) becomes intuitively clear. This equation says that, at a fixed vertex, the (tangential) forces of the incident rubber bands are in equilibrium. (For the Euclidean version of this, see e.g. [26] and [69].)

These considerations lead us naturally to the following generalization. Let (V, E) be a (finite) graph. Let $u\colon (V, E) \to S^2$ be an embedding of (V, E) into S^2 (u is sometimes called a map). To be consistent with the above, let $u(V) = \{v_i\}$, and label $\mathrm{Length}(u(\overline{v_i v_j})) = \theta_{ij}$ where $\overline{v_i v_j}$ denotes a spherical geodesic connecting v_i to v_j. Then a **harmonic map** u is a critical point of the energy functional $\sum_{\{i,j\}\in E} \theta_{ij}^2$.

Chapter 4

The Standard Simplex Conjecture for Small Positive Correlation

This chapter is derived from the paper [45].

4.1 Introduction

The Standard Simplex Conjecture [50] asks for the partition $\{A_i\}_{i=1}^k$ of $\mathbb{R}^n$ into $k \leq n+1$ sets of equal Gaussian measure of optimal noise stability. This Conjecture generalizes a seminal result of Borell, [17, 78], which corresponds to the $k=2$ case of the Standard Simplex Conjecture. Borell's result says that the two disjoint regions of fixed Gaussian measures $0<a<1$ and $1-a$ and of optimal noise stability must be separated by a hyperplane. Since two disjoint sets of total Gaussian measure 1 can be described by a single set and its complement, Borell's result can be stated as follows. Let $A \subseteq \mathbb{R}^n$ have Gaussian measure $0<a<1$ and let $\rho \in (0,1)$. Then the following quantity, which is referred to as the noise stability of A, is maximized when A is a half-space.

$$\int_{\mathbb{R}^n}\int_{\mathbb{R}^n} 1_A(x)1_A(x\rho + y\sqrt{1-\rho^2})e^{-(x_1^2+\cdots+x_n^2)/2}e^{-(y_1^2+\cdots+y_n^2)/2}dxdy. \tag{4.1}$$

When we say that A is a half-space, we mean that A is the set of points lying on one side of a hyperplane. If $\rho \in (-1,0)$, then the noise stability (4.1) of A is minimized among all sets of Gaussian measure a, when A is a half-space. We can rewrite (4.1) probabilistically as follows. Let $X=(X_1,\ldots,X_n), Y=(Y_1,\ldots,Y_n) \in \mathbb{R}^n$ be two standard Gaussian random vectors such that $\mathbb{E}(X_iY_j)=\rho\cdot 1_{(i=j)}$. Then the noise stability (4.1) of A is equal to $\mathbb{P}((X,Y)\in A\times A)$.

For modern proofs of Borell's theorem with additional stability statements, see [77, 29]. In the present work, we prove a specific case of the Standard Simplex Conjecture for $k=3$, when $0<\rho<\rho_0(n)$. Already for the case $k=3$, the methods used in the case $k = 2$ do not seem to apply, so new techniques are required to treat the case $k = 3$. We first discuss consequences of the full conjecture and we then state the conjecture precisely. The Standard Simplex Conjecture appears to be first stated explicitly in [50]. If true, this conjecture implies:

- Optimal hardness results for approximating the MAX-k-CUT problem [50, Theorem 1.13], a generalization of the MAX-CUT problem. (These hardness results are optimal, assuming the Unique Games Conjecture).

- The Plurality is Stablest Conjecture [54],[50][Theorem 1.10], an extension of the Majority is Stablest Conjecture [78] asserting that: the most noise-stable way to determine the winner of an election between k candidates is to take the plurality. (This result assumes that no one person has too much influence over the election's outcome, and each candidate has an equal probability of winning).

- The solution of a multi-bubble problem in Gaussian space [27, 50, 65]: in $\mathbb{R}^n$, minimize the total Gaussian perimeter of $k \leq n+1$ sets of Gaussian measure $1/k$.

The MAX-k-CUT problem asks for the partition of the vertices of any graph into k sets of maximum total edge perimeter. For the precise statement, see Definition 1.3.1 above. For a graph on n vertices, the MAX-k-CUT problem cannot be solved time polynomial in n, unless P=NP [31]. Yet, we can always find an approximate solution of the MAX-k-CUT problem in time polynomial in n [31]. To create this approximate solution, we label the vertices of the graph by vectors in $\mathbb{R}^n$, solve an appropriate semidefinite program for these vectors, and we then "round" these vectors into k bins. In particular, two vectors are rounded into the same bin if they lie in the same subset of a given partition $\{A_i\}_{i=1}^k$ of $\mathbb{R}^n$. The best way to perform this rounding procedure is then provided by the partition $\{A_i\}_{i=1}^k$ of optimal noise stability. That is, the Standard Simplex Conjecture exactly describes the best way to solve the MAX-k-CUT problem [50, Theorem A.6]. This connection between combinatorial optimization and geometry has been well-studied; see e.g. [88, 7, 55, 57, 19, 46]. For a survey of the complexity theoretic motivation for problems related to the Standard Simplex Conjecture, see [56], where Grothendieck inequalities are emphasized.

The Plurality is Stablest Conjecture for $k = 2$ was proven in [78], where it was found to be a consequence of Borell's theorem, after applying a nonlinear central limit theorem, which is referred to as an invariance principle. For $k = 2$, this problem is known as the Majority is Stablest Theorem. For more on the invariance principle, see also [22]. The invariance principle of [78] is proven by combining the Lindeberg replacement method with the hypercontractive inequality [37]. The Plurality is Stablest Conjecture says that the Plurality function nearly maximizes discrete noise stability over all functions $f\colon \{1,\ldots,k\}^n \to \{1,\ldots,k\}$. In this context, we think of the domain of f as n voters who vote for any one of k candidates. Given n votes $(a_1,\ldots,a_n) \in \{1,\ldots,k\}^n$, the value $f(a_1,\ldots,a_n) \in \{1,\ldots,k\}$ is the winner of the election. The Plurality is Stablest conjecture also assumes that each candidate has an equal probability of winning the election, and no one person has too much influence over the outcome of the election. It turns out that the latter assumption means that the function f can be well approximated by a function $g\colon \mathbb{R}^n \to \{1,\ldots,k\}$. That is, the noise stability of f is close to the sum of noise stabilities of the sets $g^{-1}(1),\ldots,g^{-1}(k)$. This approximation procedure, which uses an invariance principle, shows the equivalence of the Plurality is Stablest Conjecture and Standard Simplex Conjecture [50, Theorems 1.10 and 1.11]. We are therefore partially motivated to solve the Standard Simplex Conjecture to attempt to complete the picture set out by the sequence of works [17, 54, 78, 50].

The problem of minimizing Gaussian perimeter arises as an endpoint case of the Standard Simplex Conjecture. The Standard Simplex Conjecture is a statement involving a sum of terms of the form (4.1), and the Gaussian perimeter is recovered by letting $\rho \to 1^-$.

We now precisely state the Standard Simplex Conjecture. Let $\rho \in (-1,1)$, $n \geq 1$, $n \in \mathbb{Z}$, let $f\colon \mathbb{R}^n \to \mathbb{R}$ be bounded and measurable, and define $d\gamma_n(y) := e^{-(y_1^2+\cdots+y_n^2)/2}dy/(2\pi)^{n/2}$, $y = (y_1,\ldots,y_n) \in \mathbb{R}^n$. For $x \in \mathbb{R}^n$, define

$$T_\rho f(x) := \int_{\mathbb{R}^n} f(x\rho + y\sqrt{1-\rho^2})d\gamma_n(y). \tag{4.2}$$

The operator defined by (4.2) is known as the noise operator, or Bonami-Beckner operator, or Ornstein-Uhlenbeck operator. In particular, the Ornstein-Uhlenbeck operator is often written with $\rho = e^{-t}$, $t > 0$, so that $T_{e^{-t}}$ becomes a semigroup.

Definition 4.1.1. Let $A_1,\ldots,A_k \subseteq \mathbb{R}^n$ be measurable, $k \leq n+1$. We say that $\{A_i\}_{i=1}^k$ is a **partition** of $\mathbb{R}^n$ if $\cup_{i=1}^k A_i = \mathbb{R}^n$, and $\gamma_n(A_i \cap A_j) = 0$ for $i \neq j$, $i,j \in \{1,\ldots,k\}$. Let $\{z_i\}_{i=1}^k$ be the vertices of a regular simplex centered at the origin of $\mathbb{R}^n$. For each $i \in \{1,\ldots,k\}$, define $A_i := \{x \in \mathbb{R}^n\colon \langle x, z_i\rangle = \max_{j\in\{1,\ldots,k\}}\langle x, z_j\rangle\}$, the Voronoi region of z_i. We call $\{A_i\}_{i=1}^k$ a **regular simplicial conical partition**.

Conjecture 5 (**Standard Simplex Conjecture**, [50]). *Let $n \geq 2$, let $\rho \in [-1,1]$, and let $3 \leq k \leq n+1$. Let $\{A_i\}_{i=1}^k$ be a partition of $\mathbb{R}^n$.*

(a) If $\rho \in (0,1]$, and if $\gamma_n(A_i) = 1/k$, $\forall\ i \in \{1,\ldots,k\}$, then among all such partitions of $\mathbb{R}^n$, the quantity

$$J := \sum_{i=1}^k \int_{\mathbb{R}^n} 1_{A_i}(x)T_\rho(1_{A_i})(x)d\gamma_n(x) \tag{4.3}$$

is maximized by a regular simplicial conical partition.

(b) If $\rho \in [-1, 0)$ (with no restriction on the measures of the sets A_i, $i \in \{1, \ldots, k\}$), then among all partitions of $\mathbb{R}^n$, the quantity J is minimized by a regular simplicial conical partition.

The following theorem is our main result.

Theorem 4.1.2 (Main Theorem). *Fix $n \geq 2$, $k = 3$. There exists $\rho_0 = \rho_0(n, k) > 0$ such that Conjecture 5 holds for $\rho \in (0, \rho_0)$.*

Theorem 4.1.2 seems to have no direct relation to Gaussian isoperimetric problems [27], since these problems are implied by letting $\rho \to 1^-$ in Conjecture 5. Also, [50, Lemma A.4,Theorem A.6] shows that Theorem 4.1.2 seems to give no new information about the MAX-k-CUT problem, since in this problem, $\rho < 0$ is most relevant. Surprisingly, our proof strategy does not work for $\rho < 0$, as we show in Theorem 4.7.4. Lastly, Theorem 4.1.2 has a weak application to the Plurality is Stablest problem, which we describe in Corollary 4.7.3 below.

Let $X = (X_1, \ldots, X_n), Y = (Y_1, \ldots, Y_n)$ be standard normal n-dimensional Gaussian random vectors such that the covariances satisfy $\mathbb{E}(X_i Y_j) = \rho \cdot 1_{\{i=j\}}$, $i, j \in \{1, \ldots, n\}$. In [50], the quantity (4.3) is written as $\sum_{i=1}^{k} \mathbb{P}((X, Y) \in A_i \times A_i)$. To see that our formulation of Conjecture 5 is equivalent to that of [50], let $A \subseteq \mathbb{R}^n$ and note that

$$\begin{aligned}
\int_{\mathbb{R}^n} 1_A T_\rho 1_A d\gamma_n &= \int_{\mathbb{R}^n} 1_A(x) \int_{\mathbb{R}^n} 1_A(x\rho + y\sqrt{1-\rho^2}) d\gamma_n(y) d\gamma_n(x) \\
&= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} 1_A(x) 1_A(x\rho + y\sqrt{1-\rho^2}) d\gamma_n(y) d\gamma_n(x) \\
&= \mathbb{P}((X, Y) \in A \times A).
\end{aligned}$$

4.1.1 A Synopsis of the Main Theorem

We now describe the proof of Theorem 4.1.2. We first take the derivative $d/d\rho$ of the quantity J defined by (4.3). This procedure is common, and it dates back at least to the the proof of the Log-Sobolev Inequality by Gross [37]. Taking this derivative allows us to relate J to the works [55, 57]. In Section 4.3, we modify the results of [55, 57] to prove the existence of a partition that maximizes $(d/d\rho)J$. Then, in Section 4.4, we further modify results of [55, 57] to show that, if $\rho > 0$ is small, then a partition maximizing $(d/d\rho)J$ is close to a partition maximizing $(d/d\rho)|_{\rho=0}J$. And by [55], we know that the partition maximizing $(d/d\rho)|_{\rho=0}J$ is a regular simplicial conical partition, for dimension $n \geq 2$ and $k = 3$ partition elements.

So, for small $\rho > 0$, a partition maximizing $(d/d\rho)J$ is close to a regular simplicial conical partition. The structure of the operator T_ρ then permits the exploitation of a feedback loop. This feedback loop says: if our partition maximizes

$(d/d\rho)J$ for small $\rho > 0$, and if this partition is close to a regular simplicial conical partition, then this partition is even closer to a regular simplicial conical partition. This feedback loop is investigated in Section 4.5, especially in the crucial Lemma 4.6.1. A similar feedback loop was already apparent in [55][Lemma 3.3]. The full argument of Theorem 4.1.2 is then assembled in Section 4.7. By using this feedback loop, we show in Theorem 4.7.1 that a regular simplicial conical partition maximizes $(d/d\rho)J$ for small $\rho > 0$, $k = 3$, $n \geq 2$. Then, the Fundamental Theorem of Calculus allows us to relate $(d/d\rho)J$ to J, therefore completing the proof of the main theorem, Theorem 4.1.2.

Since Lemma 4.6.1 is rather lengthy and crucial to this investigation, we will further describe the idea behind it. If we know that our partition maximizes $(d/d\rho)J$, and if we also know that this partition is close to a regular simplicial conical partition, then the first variation should immediately tell us that our partition is actually a regular simplicial conical partition. Unfortunately, this intuition does not translate into a proof. The main technical problem is that the sets we are dealing with are unbounded, and we need to know precise information about the Ornstein-Uhlenbeck operator applied to these sets, for points that are very far from the origin. Since the Gaussian measure decays exponentially away from the origin, this means that it becomes hard to say something precise about the points in these sets that are very far from the origin. So, we require very precise estimates of the Ornstein-Uhlenbeck operator, and the errors that it accrues when we evaluate it far from the origin. These estimates are performed in Lemmas 4.5.2 and 4.5.3. Unfortunately, to use these estimates effectively, we need to slowly enlarge the regions where we use these estimates. The details of enlarging these regions becomes surprisingly complicated, occupying the seven steps of Lemma 4.6.1.

In Section 4.7, we also show the surprising fact that our strategy fails for small negative correlation. That is, for small $\rho < 0$, $(d/d\rho)J$ is not maximized by the regular simplicial conical partition. This result does not confirm or deny Conjecture 5 for $\rho < 0$. However, one may interpret from this result that the case of Conjecture 5 for $\rho < 0$ could be more difficult than the case $\rho > 0$.

We should also emphasize the lack of symmetrization in the proof of Theorem 4.1.2. Symmetrization is one of a few general strategies that solves many optimization problems. In our context, symmetrization would appear as follows. Recall the definition of J from (4.3). Suppose we have a partition $\{A_i\}_{i=1}^k \subseteq \mathbb{R}^n$. Change this partition into a "more symmetric" partition $\{\widetilde{A_i}\}_{i=1}^k$ such that J or $(d/d\rho)J$ is larger for $\{\widetilde{A_i}\}_{i=1}^k$. In the proof of the main theorem, it is tempting to use this symmetrization paradigm. The works [17],[78] and [50] use Gaussian symmetrization in a crucial way. However, we find this approach to be less natural for Conjecture 5, so we do not explicitly use symmetrization. Nevertheless, symmetry does play a crucial role in our proof, especially in the estimates of Section 4.4. It should also be noted that the works [55, 57] do not explicitly use symmetrization, and this lack of symmetrization is one of their novel aspects.

4.1.2 Preliminaries

For $f \in L_2(\gamma_n)$, define T_ρ as in (4.2). The operator T_ρ is a parametrization of the Ornstein-Uhlenbeck operator. The operator T_ρ is not a semigroup, but it satisfies $T_{\rho_1}T_{\rho_2} = T_{\rho_1\rho_2}$, $\rho_1, \rho_2 \in [-1,1]$, by (4.6) below. We use definition (4.2) since the usual Ornstein-Uhlenbeck operator is only defined for $\rho \in [0,1]$. Let $\lambda > 0$, $x \in \mathbb{R}$. Recall from (2.1) that the Hermite polynomials of one variable are defined by the generating function

$$e^{\lambda x - \lambda^2/2} =: \sum_{\ell \in \mathbb{N}} \lambda^\ell h_\ell(x). \tag{4.4}$$

Recall that $\int_{\mathbb{R}} h_\ell^2 d\gamma_1 = 1/\ell!$, and $\{\sqrt{\ell!}\, h_\ell\}_{\ell \in \mathbb{N}}$ is an orthonormal basis of $L_2(\gamma_1)$ from (2.8) and (2.9). Recall also that $h_0(x) = 1$ and $h_1(x) = x$ from (2.3). Set $f(x) := e^{\lambda x - \lambda^2/2}$. A routine computation shows that $T_\rho(f)(x) = e^{(\lambda\rho)x - (\lambda\rho)^2/2}$. Indeed

$$\begin{aligned}
T_\rho(f)(x) &= \int_{\mathbb{R}^n} e^{\lambda(x\rho + y\sqrt{1-\rho^2}) - \lambda^2/2} d\gamma_1(y) = \int_{\mathbb{R}^n} e^{(\lambda\rho)x + (\lambda\sqrt{1-\rho^2})y - \lambda^2/2 - y^2/2} \frac{dy}{\sqrt{2\pi}} \\
&= e^{(\lambda\rho)x - \lambda^2/2} \int_{\mathbb{R}^n} e^{-\frac{1}{2}(y - \lambda\sqrt{1-\rho^2})^2 + \lambda^2(1-\rho^2)/2} \frac{dy}{\sqrt{2\pi}} = e^{(\lambda\rho)x - \lambda^2/2 + \lambda^2(1-\rho^2)/2} \\
&= e^{(\lambda\rho)x - (\lambda\rho)^2/2}.
\end{aligned}$$

Therefore, by (4.4),

$$T_\rho f(x) = \sum_{\ell \in \mathbb{N}} \lambda^\ell \rho^\ell h_\ell(x). \tag{4.5}$$

So, by linearity, $T_\rho h_\ell(x) = \rho^\ell h_\ell(x)$.

We now extend the above observations to higher dimensions. Let $f \in L_2(\gamma_n)$, so that $f = \sum_{\ell \in \mathbb{N}^n} a_\ell h_\ell \sqrt{\ell!}$, $a_\ell \in \mathbb{R}$, where $\ell = (\ell_1, \ldots, \ell_n) \in \mathbb{N}^n$ and $h_\ell(x) = \prod_{i=1}^n h_{\ell_i}(x_i)$. Write $|\ell| := \ell_1 + \cdots + \ell_n$ and $\ell! := (\ell_1!)\cdots(\ell_n!)$. Then T_ρ satisfies $T_\rho h_\ell = \rho^{|\ell|} h_\ell$ and for $x \in \mathbb{R}^n$,

$$T_\rho f(x) = \sum_{\ell \in \mathbb{N}^n} \rho^{|\ell|} \sqrt{\ell!}\, h_\ell(x) \left(\int_{\mathbb{R}^n} \sqrt{\ell!}\, h_\ell f d\gamma_n \right) \tag{4.6}$$

Let $\Delta := \sum_{i=1}^n \partial^2/\partial x_i^2$, and define

$$L := -\Delta + \sum_{i=1}^n x_i \cdot \frac{\partial}{\partial x_i}. \tag{4.7}$$

A well-known calculation shows the following equality, which we prove in the Appendix, Section 4.9.

$$\frac{d}{d\rho}T_\rho f(x) = \rho^{-1}LT_\rho f(x) = \frac{1}{\rho}\left(\langle x, \nabla T_\rho f(x)\rangle - \Delta T_\rho f(x)\right) \tag{4.8}$$

$$= \frac{1}{\sqrt{1-\rho^2}}\Bigg[\left\langle x, \int_{\mathbb{R}^n} y f(x\rho + y\sqrt{1-\rho^2})d\gamma_n(y)\right\rangle + \frac{\rho}{\sqrt{1-\rho^2}}\int_{\mathbb{R}^n}\left(\sum_{i=1}^n (1-y_i^2)\right) f(x\rho + y\sqrt{1-\rho^2})d\gamma_n(y)\Bigg]. \tag{4.9}$$

We say that $A \subseteq \mathbb{R}^n$ is a **cone** if A is measurable and $\forall\, t > 0$, $tA = A$.

Definition 4.1.3. A **simplicial conical partition** $\{A_i\}_{i=1}^k$ is a partition of $\mathbb{R}^n$ together with a simplex $S \subseteq \mathbb{R}^{k-1}$ with $0 \leq k-1 \leq n$ and a rotation σ of $\mathbb{R}^n$ such that $0 \in S$ and such that each facet F_i of $\sigma(S \times \mathbb{R}^{n-k+1})$ generates a partition element, i.e. $A_i = \{t\sigma(F_i \times \mathbb{R}^{n-j})\colon t \in [0,\infty)\}$, $i \in \{1,\ldots,j+1\}$. Let $k-1 \leq n$ and let $\{z_i\}_{i=1}^k \subseteq \mathbb{R}^n$ be nonzero vectors that do not all lie in a $(k-1)$-dimensional hyperplane. Define a partition such that, for $i \in \{1,\ldots,k\}$, $A_i := \{x \in \mathbb{R}^n\colon \langle x, z_i\rangle = \max_{j=1,\ldots,k}\langle x, z_j\rangle\}$. Such a partition is called the simplicial conical partition induced by $\{z_i\}_{i=1}^k$.

If $\{A_i\}_{i=1}^k$ is a simplicial conical partition induced by $\{\int_{A_i} x d\gamma_n(x)\}_{i=1}^k$, then we say the partition is a **balanced conical partition**. If $\{z_i\}_{i=1}^k \subseteq \mathbb{R}^n$ are the vertices of a $(k-1)$-dimensional regular simplex in $\mathbb{R}^n$ centered at the origin, then the partition induced by $\{z_i\}_{i=1}^k$ is called a **regular simplicial conical partition**.

Let $f \in L_2(\gamma_n)$. By Plancherel and (4.6)

$$\int_{\mathbb{R}^n} fT_\rho f d\gamma_n = \sum_{\ell\in\mathbb{N}^n}\rho^{|\ell|}\left|\int_{\mathbb{R}^n} f\sqrt{\ell!}h_\ell d\gamma_n\right|^2. \tag{4.10}$$

Substituting (4.10) into (4.3) gives

$$\sum_{i=1}^k\int_{\mathbb{R}^n} 1_{A_i}T_\rho(1_{A_i})d\gamma_n = \sum_{i=1}^k\left[\gamma_n(A_i)^2 + \rho\left\|\int_{A_i} x d\gamma_n(x)\right\|_{\ell_2^n}^2 + \sum_{\substack{\ell\in\mathbb{N}^n\\|\ell|\geq 2}}\rho^{|\ell|}\left|\int_{A_i}\sqrt{\ell!}h_\ell d\gamma_n\right|^2\right]. \tag{4.11}$$

Taking the derivative $d/d\rho$ of (4.11) at $\rho = 0$, we get a quantity studied in [55, 57].

$$\frac{d}{d\rho}\sum_{i=1}^{k}\int_{\mathbb{R}^n} 1_{A_i}T_\rho(1_{A_i})d\gamma_n = \sum_{i=1}^{k}\left[\left\|\int_{A_i} xd\gamma_n(x)\right\|_{\ell_2^n}^2 + \sum_{\substack{\ell\in\mathbb{N}^n\\ |\ell|\geq 2}} |\ell|\,\rho^{|\ell|-1}\left|\int_{A_i}\sqrt{\ell!}h_\ell d\gamma_n\right|^2\right]. \tag{4.12}$$

$$\left.\frac{d}{d\rho}\right|_{\rho=0}\sum_{i=1}^{k}\int_{\mathbb{R}^n} 1_{A_i}T_\rho(1_{A_i})d\gamma_n = \sum_{i=1}^{k}\left\|\int_{A_i} xd\gamma_n(x)\right\|_{\ell_2^n}^2. \tag{4.13}$$

4.2 Noise Stability for Zero Correlation

This section concerns noise stability at the endpoint $\rho = 0$. Specifically, we will investigate the quantity (4.13), which has already been studied in [55, 57]. Using our understanding of (4.13), we will then be able to analyze the left side of (4.12) when ρ is small, using the equality (4.12). Before beginning our discussion, we first need to consider partitions of $\mathbb{R}^n$ within the convex set defined in (4.14). Definition 4.2.3 provides the metric allowing an assertion that two partitions are close to each other, and Definition 4.2.4 allows us to discuss the Gaussian measure restricted to hypersurfaces.

Definition 4.2.1. Let $H := \oplus_{i=1}^k L_2(\gamma_n)$ and define

$$\Delta_k(\gamma_n) := \{(f_1,\ldots,f_k)\in H\colon \forall\, 1\leq i\leq k, 0\leq f_i\leq 1, \textstyle\sum_{i=1}^k f_i = 1\}. \tag{4.14}$$

Definition 4.2.2. Let $\varepsilon \geq 0$. Define

$$\Delta_k^\varepsilon(\gamma_n) := \{(f_1,\ldots,f_k)\in \Delta_k(\gamma_n)\colon \frac{1}{k}-\varepsilon \leq \int_{\mathbb{R}^n} f_i d\gamma_n \leq \frac{1}{k}+\varepsilon\}.$$

Definition 4.2.3. Define a metric d_2 on partitions $\{A_i\}_{i=1}^k, \{C_i\}_{i=1}^k$ of $\mathbb{R}^n$ by the formula

$$d_2(\{A_i\}_{i=1}^k, \{C_i\}_{i=1}^k) := \inf_{\substack{\sigma\in SO(n)\\ \pi \text{ a permutation}}} \left(\sum_{i=1}^k \left\|1_{A_i} - 1_{(\sigma C_{\pi(i)})}\right\|_{L_2(\gamma_n)}^2\right)^{1/2}.$$

Definition 4.2.4. Let $A \subseteq \mathbb{R}^n$, let $\mathcal{L}$ denote Lebesgue measure on $\mathbb{R}^n$, and define the distance $d(x,y) := \|x-y\|_{\ell_2^n}$, $x,y\in\mathbb{R}^n$. Denote δ_A as the measure $\mathcal{L}$ on $\mathbb{R}^n$ restricted to A. That is, $\delta_A(B) := \liminf_{\delta\to 0}\frac{1}{2\delta}\mathcal{L}\{y\in\mathbb{R}^n\colon \exists x\in A\cap B \text{ with } d(x,y)<\delta\}$, $B\subseteq\mathbb{R}^n$. Also, we denote $\gamma_n(\delta_A) := \liminf_{\delta\to 0}\frac{1}{2\delta}\gamma_n\{y\in\mathbb{R}^n\colon \exists x\in A \text{ with } d(x,y)<\delta\}$.

The next two lemmas are derived from [55]. Lemma 4.2.6 is a quantitative variant of Lemma 4.2.5, and it will be further improved in Lemma 4.2.8 below. In particular, Lemma 4.2.6 says that, if the first variation condition for achieving the optimum value of (4.15) is nearly satisfied, then the partition is close to being simplicial.

Lemma 4.2.5. *[55, Lemma 3.3, Corollary 3.4] Let $n \geq 2$ and let $\{B_i\}_{i=1}^3$ be a regular simplicial conical partition of $\mathbb{R}^n$. Then $(1_{B_1}, 1_{B_2}, 1_{B_3})$ uniquely achieves the following supremum, up to orthogonal transformation.*

$$\sup_{(f_1,f_2,f_3)\in\Delta_3(\gamma_n)} \sum_{i=1}^{3} \left\| \int_{\mathbb{R}^n} x f_i(x) d\gamma_n(x) \right\|_{\ell_2^n}^2 . \tag{4.15}$$

Lemma 4.2.6. *Let $n \geq 2$ and let $\{B_i\}_{i=1}^3, \{C_i\}_{i=1}^2 \subseteq \mathbb{R}^n$ be regular simplicial conical partitions. Let $\{A_i\}_{i=1}^3 \subseteq \mathbb{R}^n$ be a simplicial conical partition. Let $z_i := \int_{A_i} x d\gamma_n(x)$, and let $v_{ij} \in S^{n-1} \cap A_i \cap A_j \cap \operatorname{span}\{z_i, z_j\}$. If $|\langle z_i - z_j, v_{ij}\rangle| \leq \varepsilon < 10^{-16}$ $\forall$ $i, j \in \{1,2,3\}$, $i \neq j$, and if $d_2(\{A_i\}_{i=1}^3, \{C_1, C_2, \varnothing\}) > 1/100$, then $d_2(\{A_i\}_{i=1}^3, \{B_i\}_{i=1}^3) \leq \sqrt{6\varepsilon}$.*

Proof. For $i, j \in \{1,2,3\}$, let $0 \leq \alpha_i \leq \pi$ such that A_i is a cone with angle α_i. Let $\sigma\colon \mathbb{R}^n \to \mathbb{R}^n$ be a reflection that fixes $A_i \cap A_j$. Without loss of generality, $\sigma(A_j) \subseteq A_i$. Then $z_i - z_j = \int_{A_i \setminus \sigma(A_j)} x d\gamma_n(x)$ and $\|z_i - z_j\|_2 = \sin((\alpha_i - \alpha_j)/2)/\sqrt{2\pi}$. Let $0 \leq \theta \leq \pi$ such that $\|z_i - z_j\|_2 \cos(\theta) = \langle z_i - z_j, v_{ij}\rangle$. Then either $\|z_i - z_j\|_2 \leq \sqrt{\varepsilon/18\pi}$, or $|\cos\theta| \leq \sqrt{18\pi\varepsilon}$. In the first case, $\alpha_i - \alpha_j \leq \sqrt{\varepsilon}$. So, to complete the proof, it suffices to show that the second case does not occur. We find a contradiction by assuming that the second case occurs.

If $|\cos\theta| \leq \sqrt{18\pi\varepsilon}$, then since $\theta = (\alpha_i - \alpha_j)/2$, we must have $|\alpha_i - \alpha_j - \pi| < 18\sqrt{\varepsilon}$, so $\pi - 18\sqrt{\varepsilon} < \alpha_i - \alpha_j < \pi + 18\sqrt{\varepsilon}$, i.e. $\pi - 18\sqrt{\varepsilon} < \alpha_i \leq \pi$ and $\alpha_j \leq 18\sqrt{\varepsilon}$. Then for $r \neq i, j$, $r \in \{1,2,3\}$, we have $\alpha_r = 2\pi - \alpha_i - \alpha_j > 2\pi - \pi - 18\sqrt{\varepsilon} > \pi - 18\sqrt{\varepsilon}$. Since $\alpha_i, \alpha_j > \pi - 18\sqrt{\varepsilon}$, we conclude that $d_2(\{A_i\}_{i=1}^3, \{C_1, C_2, \varnothing\}) < 18\varepsilon^{1/4} < 1/100$, a contradiction. ∎

We require the ensuing explicit calculation from [55] in Lemma 4.2.8 below. This calculation is reduced to a computation of Lagrange Multipliers in [55, Corollary 3.4]. For any $(f_1, \ldots, f_k) \in \Delta_k(\gamma_n)$, define

$$\psi_0(f_1, \ldots, f_k) := \sum_{i=1}^{k} \left\| \int_{\mathbb{R}^n} x f_i(x) d\gamma_n(x) \right\|_{\ell_2^n}^2 .$$

Lemma 4.2.7. *[55, Corollary 3.4]*

$$\sup_{(f_1,f_2)\in\Delta_2(\gamma_n)} \psi_0(f_1, f_2) = \frac{1}{\pi}, \qquad \sup_{(f_1,f_2,f_3)\in\Delta_3(\gamma_n)} \psi_0(f_1, f_2, f_3) = \frac{9}{8\pi}.$$

The following Lemma is a quantitative improvement of Lemmas 4.2.5 and 4.2.6. Combining Lemma 4.2.8 with (4.13) will show that an optimizer of the quantity $(d/d\rho)\sum_{i=1}^{k}\int_{\mathbb{R}^n} 1_{A_i}T_\rho 1_{A_i}d\gamma_n$ is almost simplicial conical for small $\rho>0$.

Lemma 4.2.8. *Let $\varepsilon>0$, $n\geq 2$. Let $\{A_i\}_{i=1}^3$ be a partition of $\mathbb{R}^n$, and let $\{B_i\}_{i=1}^3$ be a regular simplicial conical partition of $\mathbb{R}^n$. Assume that $\varepsilon<1/100$ and*

$$\psi_0(1_{A_1},1_{A_2},1_{A_3}) > \sup_{(f_1,f_2,f_3)\in\Delta_3(\gamma_n)} \psi_0(f_1,f_2,f_3)-\varepsilon. \tag{4.16}$$

Then

$$d_2(\{A_i\}_{i=1}^3,\{B_i\}_{i=1}^3)\leq 6\varepsilon^{1/8}. \tag{4.17}$$

Proof. Assume that (4.16) holds. For $i\in\{1,2,3\}$, let $z_i:=\int_{A_i}xd\gamma_n(x)$, $w_i:=\int_{B_i}xd\gamma_n(x)$. We may assume that, for all $i,j\in\{1,2,3\}$ with $i\neq j$, $\langle z_i,z_j\rangle<0$. To see this, we argue by contradiction. Suppose there exist $i,j\in\{1,2,3\}$, $i\neq j$ with $\langle z_i,z_j\rangle\geq 0$. For $p\in\{1,2,3\}$, $p\neq i,j$, let $A_p'':=A_p$, let $A_i'':=A_i\cup A_j$, and let $A_j'':=\varnothing$. For $p\in\{1,2,3\}$, let $z_p'':=\int_{A_p''}xd\gamma_n(x)$. Then

$$\sum_{p=1}^{3}\|z_p''\|_{\ell_2^n}^2-\sum_{p=1}^{3}\|z_p\|_{\ell_2^n}^2=\|z_i+z_j\|_{\ell_2^n}^2-\|z_i\|_{\ell_2^n}^2-\|z_j\|_{\ell_2^n}^2\geq 0.$$

Rewriting this inequality using the definition of ψ_0,

$$\psi_0(1_{A_1},1_{A_2},1_{A_3})\leq\psi_0(1_{A_1''},1_{A_2''},1_{A_3''}). \tag{4.18}$$

Since $\{A_p''\}_{p=1}^3$ is a partition of $\mathbb{R}^n$ with at most two nonempty elements, Lemma 4.2.7 says

$$\left(\sup_{(f_1,f_2,f_3)\in\Delta_3(\gamma_n)}\psi_0(f_1,f_2,f_3)\right)-\psi_0(1_{A_1''},1_{A_2''},1_{A_3''})\geq\frac{1}{8\pi}>10^{-2}. \tag{4.19}$$

Combining (4.18) and (4.19) contradicts (4.16). Therefore, $\langle z_i,z_j\rangle<0$ for all $i,j\in\{1,2,3\}$.

We now claim that, for each pair $i,j\in\{1,2,3\}$ with $i\neq j$, we have

$$\max_{p\in\{i,j\}}\|z_p\|_{\ell_2^n}^2\geq 1/16. \tag{4.20}$$

We again argue by contradiction. Suppose there exist $i,j\in\{1,2,3\}$ with $i\neq j$ and $\max_{p\in\{i,j\}}\|z_p\|_{\ell_2^n}^2<1/16$. Let $p\in\{1,2,3\}$, $p\neq i,j$. Then $\|z_p\|_{\ell_2^n}^2\leq 1/(2\pi)$ with equality if and only if 1_{A_p} is a half-space whose boundary contains the origin of $\mathbb{R}^n$. This follows immediately from rearrangement. Observe, if $z_p \neq 0$,

$$\begin{aligned}
\|z_p\|_{\ell_2^n}^2 &= \left\langle z_p, \int_{A_p} x d\gamma_n(x) \right\rangle \leq \left\langle z_p, \int_{A_p \cap \{x\colon \langle x, z_p\rangle \geq 0\}} x d\gamma_n(x) \right\rangle \\
&\leq \left\langle z_p, \int_{\{x\colon \langle x, z_p\rangle \geq 0\}} x d\gamma_n(x) \right\rangle \\
&= \left\langle \int_{A_p} x d\gamma_n(x), \int_{\{x\colon \langle x, z_p\rangle \geq 0\}} x d\gamma_n(x) \right\rangle \leq \left\| \int_{\{x\colon \langle x, z_p\rangle \geq 0\}} x d\gamma_n(x) \right\|_{\ell_2^n}^2 = \frac{1}{2\pi}.
\end{aligned}$$

Therefore,

$$\psi_0(1_{A_1}, 1_{A_2}, 1_{A_3}) \leq 1/8 + 1/(2\pi) \leq 1/\pi.$$

This inequality contradicts (4.16) as in (4.19), since

$$\sup_{(f_1,f_2,f_3)\in\Delta_3(\gamma_n)} \psi_0(f_1, f_2, f_3) = 9/(8\pi),$$

using Lemma 4.2.7. We conclude that (4.20) holds.

Define δ such that

$$\delta := \max_{i,j\in\{1,2,3\}, i\neq j} \gamma_n(\{x \in \mathbb{R}^n \colon \langle z_i - z_j, x\rangle \leq 0\} \cap A_i). \tag{4.21}$$

Fix $i, j \in \{1, 2, 3\}$ such that $\delta = \gamma_n(\{x \in \mathbb{R}^n \colon \langle z_i - z_j, x\rangle \leq 0\} \cap A_i)$. We want to find a bound on δ. Let $0 < h$ such that $\int_0^h d\gamma_1 = \delta$. Now, define $\{A'_r\}_{r=1}^3$ such that $A'_p = A_p$ for $p \neq i, j$, $A'_i = A_i \setminus (A_i \cap \{x \in \mathbb{R}^n \colon \langle z_i - z_j, x\rangle \leq 0\})$ and $A'_j = A_j \cup (A_i \cap \{x \in \mathbb{R}^n \colon \langle z_i - z_j, x\rangle \leq 0\})$. Let $z'_p := \int_{A'_p} x d\gamma_n(x)$, $p = 1, 2, 3$. Then

$$\begin{aligned}
&\sum_{p=1}^3 \|z'_p\|_{\ell_2^n}^2 - \sum_{p=1}^3 \|z_p\|_{\ell_2^n}^2 \\
&\qquad = 2\left\langle \int_{\{y\colon \langle z_i - z_j, y\rangle \leq 0\}\cap A_i} y d\gamma_n(y), z_j - z_i \right\rangle + 2\left\| \int_{\{y\colon \langle z_i - z_j, y\rangle \leq 0\}\cap A_i} y d\gamma_n(y) \right\|_{\ell_2^n}^2 \\
&\qquad \geq 2\left\langle \int_{\{y\colon -h \leq \langle z_i - z_j, y\rangle \leq 0\}} y d\gamma_n(y), z_j - z_i \right\rangle \\
&\qquad = 2\,\|z_i - z_j\|_{\ell_2^n} \int_0^h y d\gamma_1(y) \overset{(4.20)}{>} \delta^2/3.
\end{aligned} \tag{4.22}$$

Here we used rearrangement and also the inequality

$$\|z_i - z_j\|_{\ell_2^n} > (\max_{p\in\{i,j\}} \|z_p\|_{\ell_2^n}^2)^{1/2},$$

which itself uses $\langle z_i, z_j\rangle < 0$. By (4.16) and (4.22), $\delta^2 < 3\varepsilon$, i.e.

$$\delta < \sqrt{3\varepsilon}. \tag{4.23}$$

Now, for $p \in \{1,2,3\}$, let $\widetilde{A_p} := \{x \in \mathbb{R}^n\colon \langle x, z_p\rangle = \max_{j=1,2,3}\langle x, z_j\rangle\}$ and let $\widetilde{z_p} := \int_{\widetilde{A_p}} x d\gamma_n(x)$. By (4.23) and (4.21),

$$d_2(\{A_i\}_{i=1}^3, \{\widetilde{A_i}\}_{i=1}^3) \leq 3\sqrt{2}\,\varepsilon^{1/4}. \tag{4.24}$$

For $p \in \{1,2,3\}$, let $y_p := \widetilde{z_p} - z_p \in \mathbb{R}^n$, so that $\|y_p\|_2 \leq 3\sqrt{2}\,\varepsilon^{1/4}$ by (4.24) and Hilbert space duality. Let $x \in \mathbb{R}^n$. Then for $i, j \in \{1,2,3\}$, $i \neq j$,

$$\langle \widetilde{z_i} - \widetilde{z_j}, x\rangle = \langle z_i - z_j, x\rangle + \langle y_i - y_j, x\rangle. \tag{4.25}$$

For $i, j \in \{1,2,3\}$, $i \neq j$, let $v_{ij} = S^{n-1} \cap \widetilde{A_i} \cap \widetilde{A_j} \cap \mathrm{span}\{\widetilde{z_i}\}_{i=1}^3$. By definition of v_{ij} and $\{\widetilde{A_i}\}_{i=1}^3$, $\langle z_i - z_j, v_{ij}\rangle = 0$. So, by (4.25), $|\langle \widetilde{z_i} - \widetilde{z_j}, v_{ij}\rangle| \leq 3\sqrt{2}\,\varepsilon^{1/4}$, implying that $d_2(\{\widetilde{A_i}\}_{i=1}^3, \{B_i\}_{i=1}^3) \leq 3 \cdot 2^{3/4}\varepsilon^{1/8}$, by Lemma 4.2.6. This inequality together with (4.24) and the triangle inequality for d_2 prove (4.17). □

4.3 The First Variation

Recall (4.7). The following existence argument which uses convexity is a variant of [55, Lemma 3.1] and [57, Lemma 2.1].

Lemma 4.3.1 (First Variation). *Let $\rho \in (0,1)$. Then $\exists$ a partition $\{A_i\}_{i=1}^k$ of $\mathbb{R}^n$ such that*

$$\sum_{i=1}^k \int_{\mathbb{R}^n} 1_{A_i} \frac{d}{d\rho} T_\rho 1_{A_i} d\gamma_n = \sup_{(f_1,\ldots,f_k)\in\Delta_k(\gamma_n)} \sum_{i=1}^k \int_{\mathbb{R}^n} f_i \frac{d}{d\rho} T_\rho f_i d\gamma_n \tag{4.26}$$

Also, for each $i \in \{1,\ldots,k\}$, the following containment holds, less sets of γ_n measure zero:

$$A_i \supseteq \{x \in \mathbb{R}^n\colon LT_\rho 1_{A_i}(x) > LT_\rho 1_{A_j}(x), \forall\, j \neq i, j \in \{1,\ldots,k\}\}. \tag{4.27}$$

Proof. We show that (4.3) is maximized over $\Delta_k(\gamma_n)$, which contains the set of partitions of $\mathbb{R}^n$. Note that $\Delta_k(\gamma_n) \subseteq H$ is norm closed, convex, and norm bounded. Therefore, $\Delta_k(\gamma_n)$ is weakly closed. Also, $\Delta_k(\gamma_n)$ is weakly compact by the Banach-Alaoglu Theorem. Using (4.8), define $\psi_\rho\colon \Delta_k(\gamma_n) \to \mathbb{R}$ by

$$\psi_\rho(g_1,\ldots,g_k) := \frac{d}{d\rho}\sum_{i=1}^k \int_{\mathbb{R}^n} g_i T_\rho g_i d\gamma_n := \rho^{-1}\sum_{i=1}^k \int_{\mathbb{R}^n} g_i LT_\rho g_i d\gamma_n. \tag{4.28}$$

By (4.11), ψ_ρ is an exponentially decaying sum of uniformly bounded weakly continuous functions. Therefore, ψ_ρ is weakly continuous on the weakly compact set $\Delta_k(\gamma_n)$. So there exists $(f_1,\ldots,f_k)\in\Delta_k(\gamma_n)$ that maximizes ψ_ρ.

Since $\rho\in(0,1]$, (4.12) implies: $\forall\ f\in L_2(\gamma_n)$, $\int fLT_\rho f d\gamma_n\geq 0$. We now apply this fact to see that ψ_ρ is convex. Let $\lambda\in[0,1]$, $(g_1,\ldots,g_k),(h_1,\ldots,h_k)\in\Delta_k(\gamma_n)$. Then

$$\begin{aligned}
&\lambda\psi_\rho(g_1,\ldots,g_k)+(1-\lambda)\psi_\rho(h_1,\ldots,h_k)\\
&\qquad\qquad-\psi_\rho(\lambda g_1+(1-\lambda)h_1,\ldots,\lambda g_k+(1-\lambda)h_k)\\
&=\frac{1}{\rho}\sum_{i=1}^{k}\Big[\lambda\int_{\mathbb{R}^n}g_iLT_\rho g_i+(1-\lambda)\int_{\mathbb{R}^n}h_iLT_\rho h_i\\
&\qquad\qquad-\int_{\mathbb{R}^n}(\lambda g_i-(1-\lambda)h_i)LT_\rho(\lambda g_i-(1-\lambda)h_i)\Big]\\
&=\lambda(1-\lambda)\int_{\mathbb{R}^n}(g_i-h_i)LT_\rho(g_i-h_i)\geq 0.
\end{aligned}$$

Since ψ_ρ is convex on $\Delta_k(\gamma_n)$, ψ_ρ achieves its maximum at an extreme point of $\Delta_k(\gamma_n)$. Therefore, there exists a partition $\{A_i\}_{i=1}^k$ of $\mathbb{R}^n$ such that $(1_{A_1},\ldots,1_{A_k})\in\Delta_k(\gamma_n)$ maximizes ψ_ρ on $\Delta_k(\gamma_n)$ [57, Lemma 2.1]. Specifically, it is noted in [57, Lemma 2.1] that the extreme points of $\Delta_k(\gamma_n)$ are exactly the partitions of $\mathbb{R}^n$ into k pieces.

We now prove (4.27) by contradiction. By the Lebesgue density theorem [93][1.2.1, Proposition 1], we may assume that, for all $i\in\{1,\ldots,k\}$, if $y\in A_i$, then we have $\lim_{r\to 0}\gamma_n(A_i\cap B(y,r))/\gamma_n(B(y,r))=1$. Suppose there exist $j,m\in\{1,\ldots,k\}$ and there exists $y\in\mathbb{R}^n$, $r>0$ such that $y\in A_j$, $\gamma_n(B(y,r)\cap A_j)>0$ and $LT_\rho 1_{A_j}(y)<LT_\rho 1_{A_m}(y)$. By (4.2),

$$T_\rho 1_{A_j}(x)=\int_{\mathbb{R}^n}1_{A_j}(y)e^{-\|y-x\rho\|_2^2/[2(1-\rho^2)]}\frac{dy}{(2\pi(1-\rho^2))^{n/2}}.$$

So, $LT_\rho 1_{A_j}=\rho(d/d\rho)T_\rho 1_{A_j}(x)$ is a continuous function of x.

Therefore, there exists a ball $B(y,r)$, $r>0$ such that $\gamma_n(B(y,r)\cap A_j)>0$ and such that

$$\sup_{x\in B(y,r)}LT_\rho 1_{A_j}(x)<\inf_{x\in B(y,r)}LT_\rho 1_{A_m}(x).$$

Let $\phi(x):=1_{B(y,r)\cap A_j}(x)$. For $\lambda\in[0,1]$, note that

$$\big(1_{A_1},\ldots,1_{A_j}-\lambda\phi,\ldots,1_{A}+\lambda\phi,\ldots,1_{A_k}\big)\in\Delta_k(\gamma_n).\tag{4.29}$$

However,

$$\begin{aligned}\frac{d}{d\lambda}\Big|_{\lambda=0}&\psi_\rho(1_{A_1},\ldots,1_{A_j}-\lambda\phi,\ldots,1_{A_m}+\lambda\phi,\ldots,1_{A_k})\\&=\frac{2}{\rho}\int\phi(x)LT_\rho(1_{A_m}-1_{A_j})(x)d\gamma_n(x)>0.\end{aligned}\tag{4.30}$$

But (4.30) contradicts the maximality of $(1_{A_1},\ldots,1_{A_k})$ on $\Delta_k(\gamma_n)$, so (4.27) holds. □

4.4 Perturbative Estimates

Recalling (4.28), the following estimates allow us to relate ψ_ρ to ψ_0 for small $\rho>0$, for simplicial conical partitions. In particular, we make a close examination of the two quantities of (4.9). Since lemma 4.4.2 gives precise estimates of the two quantities of (4.9), combining Lemma 4.4.2 with (4.27) gives precise geometric information about a partition $\{A_i\}_{i=1}^k\subseteq\mathbb{R}^n$ optimizing noise stability. In particular, to see one way that we will apply Lemma 4.4.2, see (4.120) below. However, note that (4.120) below does not give sufficiently precise information to identify the sets optimizing noise stability. So, the real need for Lemma 4.4.2 will occur in the proof of the Main Lemma 4.6.1, where the precise estimate (4.50) is used.

Lemma 4.4.1. *Let $A\subseteq\mathbb{R}^n$ be a cone. Then*

$$\int_{\mathbb{R}^n}\Big(\sum_{i=1}^n(1-y_i^2)\Big)1_A(y)d\gamma_n(y)=0.$$

Proof. The assertion follows by standard equalities for the moments of a Gaussian random variable. Let $\alpha>0$. Define $f(\alpha)$ by the formula

$$f(\alpha):=\int_{\mathbb{R}^n}1_A(y)e^{-\alpha(y_1^2+\cdots+y_n^2)/2}\frac{dy}{(2\pi)^{n/2}}.$$

By changing variables, $f(\alpha)=\alpha^{-n/2}\int_{\mathbb{R}^n}1_A(y)d\gamma_n(y)$. So,

$$-\frac{1}{2}\int_{\mathbb{R}^n}\Big(\sum_{i=1}^n y_i^2\Big)1_A(y)d\gamma_n(y)=\frac{df(\alpha)}{d\alpha}\Big|_{\alpha=1}=-\frac{n}{2}\int_{\mathbb{R}^n}1_A(y)d\gamma_n(y).$$

□

Lemma 4.4.2. *Fix $k=3$, $n\geq 2$, $\rho\in(0,1)$. Let $\{C_i\}_{i=1}^k\subseteq\mathbb{R}^2$ be a simplicial conical partition. Let $\{B_i\}_{i=1}^k:=\{C_i\times\mathbb{R}^{n-2}\}_{i=1}^k$. Fix $i,j\in\{1,\ldots,k\}$. Let $\sigma\colon\mathbb{R}^n\to\mathbb{R}^n$ denote reflection across $B_i\cap B_j$. Assume that $B_i=\sigma B_j$ and that*

$B_i \subseteq \{x \in \mathbb{R}^n : x_1 \geq 0\}$. Let $e_1 = (1, 0, \ldots, 0)$, $e_2 = (0, 1, 0, \ldots, 0)$, $e_1, e_2 \in \mathbb{R}^n$. For $p \in \{1, \ldots, k\}$, let $z_p := \int_{B_p} x d\gamma_n(x)$. Note that $\mathrm{span}\{z_i, z_j\} = \mathrm{span}\{e_1, e_2\}$. Let $n_j \in \mathbb{R}^n$ be the interior unit normal of B_j so that n_j is normal to the face $(\partial B_j) \setminus (\partial B_i)$, and let $n_i \in \mathbb{R}^n$ be the interior unit normal of B_i so that n_i is normal to the face $(\partial B_i) \setminus (\partial B_j)$.

(i) If $x \in B_i \cap \{x \in \mathbb{R}^n : \langle x, n_j \rangle \leq 0\}$, then

$$\frac{1}{\rho} \left\langle x, \nabla T_\rho (1_{B_i} - 1_{B_j})(x) \right\rangle \geq 2x_1 \gamma_n \left(\delta_{\frac{(B_i \cap B_j) - x\rho}{\sqrt{1-\rho^2}}} \right) + \langle x, n_i \rangle \gamma_n \left(\delta_{\frac{((\partial B_i) \setminus B_j) - x\rho}{\sqrt{1-\rho^2}}} \right). \tag{4.31}$$

(ii) If $x \in B_i \cap \{x \in \mathbb{R}^n : \langle x, n_j \rangle \geq 0\}$, then

$$\frac{1}{\rho} \left\langle x, \nabla T_\rho (1_{B_i} - 1_{B_j})(x) \right\rangle \geq 2x_1 \gamma_n \left(\delta_{\frac{(B_i \cap B_j) - x\rho}{\sqrt{1-\rho^2}}} \right). \tag{4.32}$$

(iii) For $x \in B_i$,

$$\begin{aligned} &\left| \int_{\mathbb{R}^n} \left(\sum_{\ell=1}^n (1 - y_\ell^2) \right) (1_{B_i} - 1_{B_j})(x\rho + y\sqrt{1-\rho^2}) d\gamma_n(y) \right| \\ &\qquad \leq \frac{\rho}{\sqrt{1-\rho^2}} (\sqrt{6} + (n-1)\sqrt{2}) x_1. \end{aligned} \tag{4.33}$$

(iv) For $x \in B_i$ with $x_1 > \sqrt{n}\sqrt{1-\rho^2}/\rho$,

$$\int_{\mathbb{R}^n} \left(\sum_{\ell=1}^n (1 - y_\ell^2) \right) (1_{B_i} - 1_{B_j})(x\rho + y\sqrt{1-\rho^2}) d\gamma_n(y) \geq 0. \tag{4.34}$$

Proof of (i). Below, we use differentiation in the distributional sense. Let $x \neq 0$. For $x \in (\partial B_i) \cap B_j$, $\nabla 1_{B_i}(x) = e_1$ since $B_i \subseteq \{x \in \mathbb{R}^n : x_1 \geq 0\}$, and for $x \in (\partial B_i) \setminus B_j$, $\nabla 1_{B_i}(x) = n_i$. Similarly, for $x \in (\partial B_j) \cap B_i$, $-\nabla 1_{B_j}(x) = e_1$, and for $x \in (\partial B_j) \setminus B_i$, $-\nabla 1_{B_j}(x) = -n_j$. Then

$$\begin{aligned} &\frac{1}{\rho} \nabla T_\rho (1_{B_i} - 1_{B_j})(x) = T_\rho(\nabla(1_{B_i} - 1_{B_j}))(x) \\ &= T_\rho[2(e_1)\delta_{B_i \cap B_j} + n_i \delta_{(\partial B_i) \setminus B_j} + (-n_j)\delta_{(\partial B_j) \setminus B_i}](x) \\ &\overset{(4.2)}{=} 2e_1 \gamma_n \left(\delta_{\frac{(B_i \cap B_j) - x\rho}{\sqrt{1-\rho^2}}} \right) + n_i \gamma_n \left(\delta_{\frac{((\partial B_i) \setminus B_j) - x\rho}{\sqrt{1-\rho^2}}} \right) + (-n_j) \gamma_n \left(\delta_{\frac{((\partial B_j) \setminus B_i) - x\rho}{\sqrt{1-\rho^2}}} \right). \end{aligned} \tag{4.35}$$

Here we used

$$\begin{aligned} \int_{\mathbb{R}^n} 1_A(x\rho + y\sqrt{1-\rho^2}) d\gamma_n(y) &= \int_{\mathbb{R}^n} 1_{A - x\rho}(y\sqrt{1-\rho^2}) d\gamma_n(y) \\ &= \int_{\mathbb{R}^n} 1_{(A - x\rho)/\sqrt{1-\rho^2}}(y) d\gamma_n(y). \end{aligned}$$

Let x with $x \in B_i$ and $\langle x, (-n_j)\rangle \geq 0$. Then (4.35) immediately proves (4.31). $\square$

Proof of (ii). Let $x \in B_i \cap \{x \in \mathbb{R}^n : \langle x, n_j\rangle \geq 0\}$. By reflecting across $B_i \cap B_j$,

$$\gamma_n\left(\delta_{\frac{((\partial B_i)\setminus B_j)-x\rho}{\sqrt{1-\rho^2}}}\right) \geq \gamma_n\left(\delta_{\frac{((\partial B_j)\setminus B_i)-x\rho}{\sqrt{1-\rho^2}}}\right). \tag{4.36}$$

Define

$$w := n_i\gamma_n\left(\delta_{\frac{((\partial B_i)\setminus B_j)-x\rho}{\sqrt{1-\rho^2}}}\right) + (-n_j)\gamma_n\left(\delta_{\frac{((\partial B_j)\setminus B_i)-x\rho}{\sqrt{1-\rho^2}}}\right).$$

By (4.36), w is in the convex hull of e_1 and n_i. In particular, $\langle x, w\rangle \geq 0$, since $x \in B_i$. Combining $\langle x, w\rangle \geq 0$ with (4.35) proves (4.32). $\square$

Proof of (iii). By reflecting across $B_i \cap B_j$,

$$x \in B_i \cap B_j \quad \Longrightarrow \quad \int_{\mathbb{R}^n}\left(\sum_{\ell=1}^{n}(1-y_\ell^2)\right)(1_{B_i} - 1_{B_j})(x\rho + y\sqrt{1-\rho^2})d\gamma_n(y) = 0.$$

So, a derivative bound gives (4.33). Specifically, we apply the Fundamental Theorem of Calculus to the following identity, along with $\|(1-y_2^2)y_1\|_{L_2(\gamma_n)} = \sqrt{2}$ and $\|y_1^3 + 3y_1\|_{L_2(\gamma_n)} = \sqrt{6}$.

$$\begin{aligned}&\frac{\partial}{\partial x_1}\int_{\mathbb{R}^n}\left(\sum_{\ell=1}^{n}(1-y_\ell^2)\right)(1_{B_i} - 1_{B_j})(x\rho + y\sqrt{1-\rho^2})d\gamma_n(y)\\ &= \frac{\rho}{\sqrt{1-\rho^2}}\int_{\mathbb{R}^n}\big(-3y_1 - y_1^3 + \sum_{\ell\neq 1}(1-y_\ell^2)y_1\big)(1_{B_i} - 1_{B_j})(x\rho + y\sqrt{1-\rho^2})d\gamma_n(y).\end{aligned}$$

$\square$

Proof of (iv). Let x with $x_1 > \sqrt{n}\sqrt{1-\rho^2}/\rho$ and consider the following cone

$$A := \{0\} \cup \left\{y \in \mathbb{R}^n : y \neq 0 \wedge \sqrt{n}\frac{y}{\|y\|_2} \in \frac{B_i - x\rho}{\sqrt{1-\rho^2}}\right\}.$$

By Lemma 4.4.1, $\int_{\mathbb{R}^n} 1_A(y)\sum_{\ell=1}^{n}(1-y_\ell^2)d\gamma_n(y) = 0$. If $d(x, \partial B_i) \geq \sqrt{n}\sqrt{1-\rho^2}/\rho$, then $A = \mathbb{R}^n$ and $1_A(y)\sum_{\ell=1}^{n}(1-y_\ell^2)1_{B_i^c}(x\rho + y\sqrt{1-\rho^2}) \leq 0$, so

$$\begin{aligned}&\int_{\mathbb{R}^n}\left(\sum_{\ell=1}^{n}(1-y_\ell^2)\right)(1_{B_i} - 1_{B_j})(x\rho + y\sqrt{1-\rho^2})d\gamma_n(y)\\ &\qquad \geq \int_{\mathbb{R}^n}\left(\sum_{\ell=1}^{n}(1-y_\ell^2)\right)1_{B_i}(x\rho + y\sqrt{1-\rho^2})d\gamma_n(y)\\ &\qquad \geq \int_{\mathbb{R}^n}\left(\sum_{\ell=1}^{n}(1-y_\ell^2)\right)d\gamma_n(y) = 0.\end{aligned}$$

So, it remains to consider the case $d(x,\partial B_i) < \sqrt{n}\sqrt{1-\rho^2}/\rho$. In this case $A \neq \mathbb{R}^n$.

Since $x_1 > \sqrt{n}\sqrt{1-\rho^2}/\rho$, we have $\sum_{\ell=1}^n (1-y_\ell^2)1_{B_j}(x\rho+y\sqrt{1-\rho^2}) \leq 0$. Also, we have $\sum_{\ell=1}^n (1-y_\ell^2)1_{A^c}(y)1_{B_i}(x\rho+y\sqrt{1-\rho^2}) \geq 0$, $\sum_{\ell=1}^n (1-y_\ell^2)1_A(y)1_{B_i^c}(x\rho+y\sqrt{1-\rho^2}) \leq 0$, so

$$
\begin{aligned}
&\int_{\mathbb{R}^n} \left(\sum_{\ell=1}^n (1-y_\ell^2)\right)(1_{B_i}-1_{B_j})(x\rho+y\sqrt{1-\rho^2})d\gamma_n(y) \\
&\quad\geq \int_{\mathbb{R}^n} \left(\sum_{\ell=1}^n (1-y_\ell^2)\right)1_{B_i}(x\rho+y\sqrt{1-\rho^2})d\gamma_n(y) \\
&\quad\geq \int_{\mathbb{R}^n} \left(\sum_{\ell=1}^n (1-y_\ell^2)\right)1_A(y)1_{B_i}(x\rho+y\sqrt{1-\rho^2})d\gamma_n(y) \\
&\quad\geq \int_{\mathbb{R}^n} \left(\sum_{\ell=1}^n (1-y_\ell^2)\right)1_A(y)d\gamma_n(y) = 0.
\end{aligned}
$$

□

4.5 Iterative Estimates

The following estimates control the errors that appear in the proof of Theorem 4.1.2. Being rather technical in nature, this section could be skipped on a first reading.

Lemma 4.5.1. *For $\ell = (\ell_1,\ldots,\ell_n) \in \mathbb{N}^n$ and $x = (x_1,\ldots,x_n) \in \mathbb{R}^n$,*

$$h_\ell(x)\sqrt{\ell!} \leq |\ell|^n\, 3^{|\ell|} \prod_{i=1}^n \max\{1, |x_i|^{\ell_i}\}.$$

Proof. Let $\ell \in \mathbb{N}$. From (2.3), recall that

$$h_\ell(x) = \sum_{m=0}^{\lfloor \ell/2 \rfloor} \frac{x^{\ell-2m}(-1)^m 2^{-m}}{m!(\ell-2m)!}. \tag{4.37}$$

Using Stirling's formula, $\sqrt{2\pi}\ell^{\ell+1/2}e^{-\ell} \leq \ell! \leq e\,\ell^{\ell+1/2}e^{-\ell}$. Let ℓ, m with $m \in \{0,\ldots,\ell/2\}$, $\ell \geq 1$. Note that $\lim_{x\to 0^+} x^x = 1$ and $\min_{x\in[0,1]} x^x > 2/3$. Also, $m+\ell-2m = \ell-m \geq \ell/2$. For $m \neq 0$, write $m = \ell j$, $j \in [1/\ell, 1/2]$. Note that

$\max\{m, \ell-2m\} \geq \ell/3$. Then

$$\begin{aligned}
&\frac{\sqrt{\ell!}}{m!(\ell-2m)!} \\
&\leq \frac{\sqrt{e}\,\ell^{(1/2)\ell+1/4}e^{-\ell/2}}{2\pi m^{m+1/2}e^{-m}(\ell-2m)^{\ell-2m+1/2}e^{-(\ell-2m)}} = \frac{\sqrt{e}}{2\pi}\frac{\ell^{(1/2)\ell+1/4}}{m^{m+1/2}(\ell-2m)^{\ell-2m+1/2}}e^{\ell/2-m} \\
&= \frac{\sqrt{e}}{2\pi}\frac{\ell^{1/4}}{\sqrt{m}\sqrt{\ell-2m}}\frac{\ell^{\ell/2}}{m^m(\ell-2m)^{\ell-2m}}e^{\ell/2-m} \\
&= \frac{\sqrt{e}}{2\pi}\frac{\ell^{1/4}}{\sqrt{m}\sqrt{\ell-2m}}\frac{\ell^{\ell/2}}{(\ell j)^{\ell j}(\ell(1-2j))^{\ell(1-2j)}}e^{\frac{\ell}{2}-m} \\
&= \frac{\sqrt{e}}{2\pi}\frac{\ell^{1/4}}{\sqrt{m}\sqrt{\ell-2m}}\frac{\ell^{\ell/2}e^{\ell/2-m}}{\ell^{\ell(1-j)}j^{\ell j}(1-2j)^{\ell(1-2j)}} \\
&= \frac{\sqrt{e}}{2\pi}\frac{\ell^{1/4}}{\sqrt{m}\sqrt{\ell-2m}}\frac{\ell^{\ell/2}e^{\ell/2-m}}{\ell^{\ell/2}\ell^{\frac{\ell}{2}(1-2j)}j^{\ell j}(1-2j)^{\ell(1-2j)}} \\
&= \frac{\sqrt{e}}{2\pi}\frac{\ell^{1/4}}{\sqrt{m}\sqrt{\ell-2m}}\frac{1}{j^{\ell j}(1-2j)^{\ell(1-2j)}}\frac{e^{\frac{\ell}{2}(1-2j)}}{\ell^{\frac{\ell}{2}(1-2j)}} = \frac{\sqrt{e}}{2\pi}\frac{\ell^{1/4}}{\sqrt{m}\sqrt{\ell-2m}}\frac{(e/\ell)^{\frac{\ell}{2}(1-2j)}}{j^{\ell j}(1-2j)^{\ell(1-2j)}} \\
&\leq \frac{e}{2\pi}\frac{\ell^{1/4}}{\sqrt{m}\sqrt{\ell-2m}}\frac{1}{j^{\ell j}(1-2j)^{\ell(1-2j)}} \leq \frac{e}{2\pi}\frac{\ell^{1/4}}{\sqrt{m}\sqrt{\ell-2m}}\frac{1}{(2/3)^{2\ell}} \\
&\leq \frac{e\sqrt{3}}{\ell^{1/4}2\pi}\frac{1}{(2/3)^{2\ell}} = \frac{e\sqrt{3}}{\ell^{1/4}2\pi}(9/4)^\ell \leq \ell^{-1/4}(9/4)^\ell.
\end{aligned}$$

Here we used $(e/\ell)^{(\ell/2)(1-2j)} \leq \sqrt{e}$ for $\ell = 1, 2$.

Also, for $m = 0$ we have $\frac{\sqrt{\ell!}}{m!(\ell-2m)!} = 1$, and for $m = \ell/2$ we have

$$\begin{aligned}
\frac{\sqrt{\ell!}}{m!(\ell-2m)!} &= \frac{\sqrt{\ell!}}{(\ell/2)!} \leq \frac{\sqrt{e}\ell^{\ell/2+1/4}e^{-\ell/2}}{\sqrt{2\pi}(\ell/2)^{\ell/2+1/2}e^{-\ell/2}} = \sqrt{e}\frac{\ell^{1/4}}{\sqrt{2\pi}\ell^{1/2}2^{-\ell/2}2^{-1/2}} \\
&= \sqrt{\frac{e}{\pi}}\ell^{-1/4}2^{\ell/2} \leq \ell^{-1/4}2^{\ell/2}.
\end{aligned}$$

So, combining the above estimates with (4.37),

$$\begin{aligned}
|h_\ell(x)\sqrt{\ell!}| &\leq \sum_{m=0}^{\lfloor \ell/2 \rfloor} \ell^{-1/4}(9/4)^\ell\,|x|^{\ell-2m} \leq \sum_{m=0}^{\lfloor \ell/2 \rfloor} \ell^{-1/4}(9/4)^\ell \max\{1, |x|^{\ell-2m}\} \\
&\leq \ell\,\ell^{-1/4}(9/4)^\ell \max\{1, |x|^\ell\} \leq \ell\,3^\ell \max\{1, |x|^\ell\}.
\end{aligned}$$

Therefore, for $\ell = (\ell_1, \ldots, \ell_n) \in \mathbb{N}^n$,

$$h_\ell(x)\sqrt{\ell!} \leq \ell_1 \cdots \ell_n 3^{\ell_1+\cdots+\ell_n} \prod_{i=1}^{k} \max\{1, |x_i|^{\ell_i}\} \leq |\ell|^n\,3^{|\ell|} \prod_{i=1}^{n} \max\{1, |x_i|^{\ell_i}\}.$$

$\square$

The following Lemma uses standard tail bounds for a Gaussian random variable. We therefore omit the proof.

Lemma 4.5.2. *Let $\eta > 0, t > 0$, and let $n \geq 2$. Then*

$$\left|\int_{[-\eta,\eta]\times[t,\infty]\times\mathbb{R}^{n-2}} \sum_{\substack{\ell\in\mathbb{N}^n:\\ 0\leq|\ell|\leq 3}} \prod_{i=1}^{n} |y_i|^{\ell_i}\, d\gamma_n(y)\right| \leq 3000n^3\eta(t^2+2)e^{-t^2/2},$$

$$\left|\int_{B(0,t)^c} \sum_{\substack{\ell\in\mathbb{N}^n:\\ 0\leq|\ell|\leq 3}} \prod_{i=1}^{n} |y_i|^{\ell_i}\, d\gamma_n(y)\right| \leq 4n^3 2^{-n/2}(\Gamma(n/2))^{-1}(n+2)!(t^{n+1}+1)e^{-t^2/2}$$

$$\leq 100(n+2)!(t^{n+1}+1)e^{-t^2/2}.$$

The following Lemma says, if $\int_{\mathbb{R}^n} xf(x)d\gamma_n(x)$ is parallel to the x_1-axis, then $(d/d\rho)T_\rho f(x)$ should be bounded by a constant multiplied by $|x_1| + O(\rho)$. The precise error term (4.38) will be needed in Lemma 4.6.1 to determine the size of $(d/d\rho)T_\rho(1_{A_i} - 1_{A_j})$. The error term (4.38) will be estimated by Lemma 4.5.2, and the resulting estimate will be introduced into (4.27).

Lemma 4.5.3. *Let $\rho \in (-1,1)$, $n \geq 2$. Let $f \in L_2(\gamma_n)$ with $\int_{\mathbb{R}^n} y_2 f(y)d\gamma_n(y) = 0$. Let $x_1 \geq 0$ and $x_2 \geq 0$. Then*

$$\begin{aligned}\left|\frac{d}{d\rho}T_\rho f(x_1,x_2,0,\ldots,0)\right| &\leq \Big(|x_1| + 2\rho(|x_2|^2 + (n+1)\rho\,|x_2| + |x_1x_2| + 2n)\Big)\\ &\cdot \sup_{\substack{t_1\in[0,x_1],\,t_2\in[0,x_2]\\ \eta\in[0,\rho]}} \left|\int_{\mathbb{R}^n} \sum_{\substack{\ell\in\mathbb{N}^n:\\ 0\leq|\ell|\leq 3}} \prod_{i=1}^{n} |y_i|^{\ell_i} f((t_1,t_2,0,\ldots,0)\eta + y\sqrt{1-\eta^2})d\gamma_n(y)\right|.\end{aligned} \tag{4.38}$$

Proof. By integrating by parts, note that

$$\begin{aligned}\frac{d}{d\rho}\int_{\mathbb{R}^n} y_2 f(y\sqrt{1-\rho^2})d\gamma_n(y) &= \frac{\rho}{1-\rho^2}\int_{\mathbb{R}^n} y_2((n+1)-y_2^2)f(y\sqrt{1-\rho^2})d\gamma_n(y)\\ &\quad - \sum_{i\neq 2}\frac{\rho}{1-\rho^2}\int_{\mathbb{R}^n} y_i^2 y_2 f(y\sqrt{1-\rho^2})d\gamma_n(y).\end{aligned}$$

So, using $\int_{\mathbb{R}^n} y_2 f(y) d\gamma_n(y) = 0$ and the Fundamental Theorem of Calculus,

$$\begin{aligned}
&\int_{\mathbb{R}^n} y_2 f(y\sqrt{1-\rho^2}) d\gamma_n(y) \\
&\leq \frac{\rho^2}{1-\rho^2} \sup_{\eta \in [0,\rho]} \left(\int_{\mathbb{R}^n} y_2((n+1) - y_2^2) f(y\sqrt{1-\eta^2}) d\gamma_n(y) \right. \\
&\qquad\qquad \left. - \sum_{i \neq 2} \int_{\mathbb{R}^n} y_i^2 y_2 f(y\sqrt{1-\eta^2}) d\gamma_n(y) \right).
\end{aligned} \tag{4.39}$$

By integrating by parts again, note that

$$\begin{aligned}
&\frac{\partial}{\partial x_2} \int_{\mathbb{R}^n} y_2 f((0, x_2, 0, \ldots, 0)\rho + y\sqrt{1-\rho^2}) d\gamma_n(y) \\
&\quad = \frac{\rho}{\sqrt{1-\rho^2}} \int_{\mathbb{R}^n} f((0, x_2, 0, \ldots, 0)\rho + y\sqrt{1-\rho^2})(y_2^2 - 1) d\gamma_n(y).
\end{aligned} \tag{4.40}$$

Applying the Fundamental Theorem of Calculus to (4.40) and then using (4.39),

$$\begin{aligned}
&\left| \int_{\mathbb{R}^n} y_2 f((0, x_2, 0, \ldots, 0)\rho + y\sqrt{1-\rho^2}) d\gamma_n(y) \right| \\
&\leq |x_2| \sup_{t \in [0, x_2]} \left| \frac{\rho}{\sqrt{1-\rho^2}} \int_{\mathbb{R}^n} f((0, t, 0, \ldots, 0)\rho + y\sqrt{1-\rho^2})(y_2^2 - 1) d\gamma_n(y) \right| \\
&+ \frac{\rho^2}{1-\rho^2} \sup_{\eta \in [0,\rho]} \left(\int_{\mathbb{R}^n} y_2((n+1) - y_2^2) f(y\sqrt{1-\eta^2}) d\gamma_n(y) \right. \\
&\qquad\qquad \left. - \sum_{i \neq 2} \int_{\mathbb{R}^n} y_i^2 y_2 f(y\sqrt{1-\eta^2}) d\gamma_n(y) \right).
\end{aligned} \tag{4.41}$$

By integrating by parts as before,

$$\begin{aligned}
&\frac{\partial}{\partial x_1} [x_2 \int_{\mathbb{R}^n} y_2 f((x_1, x_2, 0, \ldots, 0)\rho + y\sqrt{1-\rho^2}) d\gamma_n(y)] \\
&\quad = x_2 \frac{\rho}{\sqrt{1-\rho^2}} \int_{\mathbb{R}^n} y_2 y_1 f((x_1, x_2, 0, \ldots, 0)\rho + y\sqrt{1-\rho^2}) d\gamma_n(y).
\end{aligned} \tag{4.42}$$

Combining (4.9), (4.41) and (4.42),

$$\begin{aligned}
|(d/d\rho)T_\rho f(x)| &\leq |x_1| \left| \int_{\mathbb{R}^n} y_1 f((x_1,x_2,0,\ldots,0)\rho + y\sqrt{1-\rho^2}) d\gamma_n(y) \right| \\
&+ |x_2|^2 \sup_{t\in[0,x_2]} \left| \frac{\rho}{\sqrt{1-\rho^2}} \int_{\mathbb{R}^n} f((0,t,0,\ldots,0)\rho + y\sqrt{1-\rho^2})(y_2^2-1) d\gamma_n(y) \right| \\
&+ \frac{\rho^2 |x_2|}{1-\rho^2} \sup_{\eta\in[0,\rho]} \Bigg(\int_{\mathbb{R}^n} y_2((n+1)-y_2^2) f(y\sqrt{1-\eta^2}) d\gamma_n(y) \\
&\qquad\qquad - \sum_{i\neq 2} \int_{\mathbb{R}^n} y_i^2 y_2 f(y\sqrt{1-\eta^2}) d\gamma_n(y) \Bigg) \\
&+ |x_1 x_2| \sup_{t\in[0,x_1]} \frac{\rho}{\sqrt{1-\rho^2}} \left| \int_{\mathbb{R}^n} y_2 y_1 f((t,x_2,0,\ldots,0)\rho + y\sqrt{1-\rho^2}) d\gamma_n(y) \right| \\
&+ \frac{\rho}{\sqrt{1-\rho^2}} \left| \int_{\mathbb{R}^n} (\sum_{i=1}^n (y_i^2-1)) f((x_1,x_2,0,\ldots,0)\rho + y\sqrt{1-\rho^2}) d\gamma_n(y) \right|.
\end{aligned} \tag{4.43}$$

We then deduce (4.38) from (4.43). $\square$

4.6 The Main Lemma

Lemma 4.6.1 below represents the main tool in the proof of the main theorem. As depicted in Figure 4.1, Lemma 4.6.1 says that, if an optimal partition is close to being simplicial conical, then it is actually much closer to being simplicial conical. So, this Lemma can be understood as a feedback loop, or as a contractive mapping type of argument. We first give an intuitive sketch of the proof of the Lemma. Let $\rho > 0$. We begin with a partition $\{A_p\}_{p=1}^3 \subseteq \mathbb{R}^n$ maximizing noise stability (4.3). We assume that there are disjoint sets $\{D_p\}_{p=1}^3$ that resemble a simplicial conical partition, as in the left side of Figure 4.1. We also assume that $A_p \supseteq D_p$ for all $p = 1, 2, 3$. We then find a sequence of sets $\{D_{p,1}\}_{p=1}^3$, $\{D_{p,2}\}_{p=1}^3$, $\ldots$ $\{D_{p,R}\}_{p=1}^3$ such that $D_{p,r} \subseteq D_{p,r+1}$ for all $1 \leq p \leq 3$, for all $r \geq 1$. This sequence of sets is chosen so that the following implication can be proven:

$$A_p \supseteq D_{p,r} \quad \Longrightarrow \quad A_p \supseteq D_{p,r+1}. \tag{4.44}$$

In order to prove (4.44), we need to show: if $A_p \supseteq D_{p,r}$, then we can get sufficiently strong estimates on $LT_\rho 1_{D_{p,r+1}}$ such that (4.27) can be verified on A_p for each $p = 1, 2, 3$. For example, in Step 1 of the proof of Lemma 4.6.1, the estimate (4.53) eventually implies (4.57). And (4.57) says that A_p must contain more points than the initial information that we assumed in (4.47).

Finally, we need to choose our sets $\{D_{p,r}\}_{p=1}^{3}$ appropriately so that, after finitely many implications of the form (4.44), we eventually get the conclusion (4.48). That is, the three sets $\{D_{p,R}\}_{p=1}^{3}$ resemble the right side of Figure 4.1, and $A_p \supseteq D_{p,R}$ for each $p = 1, 2, 3$. Thus concludes our description of the main strategy of the proof. Within the proof itself, the sets $\{D_{p,1}\}_{p=1}^{3}$, $\{D_{p,2}\}_{p=1}^{3}$, ... will not be explicitly defined. However, portions of these sets will be defined at the end of every Step of the proof. In particular, examine the sets defined by the following sequence of assertions: (4.47), (4.57), (4.62), (4.67), (4.77), (4.87), (4.94), (4.100), and finally (4.48).

Unfortunately, there are many technical obstacles that stand in the way of bringing this strategy to fruition. The first minor issue is that we cannot control small rotations of our sets. At every step of the proof, we therefore need to redefine our simplicial sets B_i, B_j to account for these small rotations. However, the main technical issue is that it is not at all obvious how to choose the sets $\{D_{p,r}\}_{p=1}^{3}$ for $r = 1, 2, 3, \ldots$ such that (4.44) can be proven for each $r = 1, 2, 3, \ldots$. Moreover, the simplest choice of these sets, namely dilations of the sets depicted in Figure 4.1, do not produce satisfactory estimates.

Ultimately, the sequence of sets defined by (4.57), (4.62), (4.67), (4.77), (4.87), (4.94), (4.100) succeeds in proving the sequence of implications (4.44) for $r = 1, 2, 3, \ldots$. Lemma 4.5.3 allows us to control the errors from our estimates, and we then make around seven modifications of the same error estimate within Lemma 4.6.1. This error estimate allows us to apply Lemma 4.4.2, so that we can improve our knowledge of the optimal partition $\{A_i\}_{i=1}^{k}$ via (4.27).

It would be preferable to write Lemma 4.6.1 as seven applications of a single Lemma, however the statement of such a Lemma would perhaps be so long and convoluted that its application would become opaque. We therefore use the longer presentation below in the hope of providing greater clarity. Finally, in the statement of Lemma 4.6.1 below, note that the plane Π exists independently of $i, j \in \{1, \ldots, k\}$.

Lemma 4.6.1. *Fix $n \geq 2$, $k = 3$. Let $0 < \eta < \rho < e^{-20(n+1)^{10^{12}} n^3 (n+2)!}$. Let $\{A_i\}_{i=1}^{k}$ be a partition of $\mathbb{R}^n$ such that (4.27) holds. Let $\Pi \subseteq \mathbb{R}^n$ be a fixed 2-dimensional plane such that $0 \in \Pi$. Assume that, for each pair $i, j \in \{1, 2, \ldots, k\}$ with $i \neq j$, there exists $\lambda' > 0$ and there exists a regular simplicial conical partition $\{B'_p\}_{p=1}^{k} \subseteq \mathbb{R}^n$ such that*

$$\int_{\mathbb{R}^n} y(1_{A_i}(y) - 1_{A_j}(y)) d\gamma_n(y) = \lambda' \int_{\mathbb{R}^n} y(1_{B'_i}(y) - 1_{B'_j}(y)) d\gamma_n(y), \tag{4.45}$$

such that

$$\int_{B'_p} x d\gamma_n(x) \in \Pi, \; \forall\, p \in \{i, j\}, \tag{4.46}$$

and such that

$$\begin{aligned}\{x \in B_i' \cup B_j' \colon 1_{A_i}(x) - 1_{A_j}(x) \neq 1_{B_i'}(x) - 1_{B_j'}(x)\} \\ \subseteq \{x \in B_i' \cup B_j' \colon |d(x, (\partial B_i') \cup (\partial B_j'))| < \eta \\ \vee \|x\|_2 \geq \sqrt{-2\log\eta} + (\rho+\eta)\sqrt{-2\log\rho}\}.\end{aligned} \tag{4.47}$$

Then, for each pair $i, j \in \{1, 2, \ldots, k\}$ with $i \neq j$, there exists $\lambda'' > 0$ and there exists a regular simplicial conical partition $\{B_p''\}_{p=1}^k \subseteq \mathbb{R}^n$ such that $\int_{\mathbb{R}^n} y(1_{A_i}(y) - 1_{A_j}(y))d\gamma_n(y) = \lambda'' \int_{\mathbb{R}^n} y(1_{B_i''}(y) - 1_{B_j''}(y))d\gamma_n(y)$, such that $\int_{B_p''} x d\gamma_n(x) \in \Pi$, $\forall\, p \in \{i, j\}$, and such that

$$\begin{aligned}\{x \in B_i'' \cup B_j'' \colon 1_{A_i}(x) - 1_{A_j}(x) \neq 1_{B_i''}(x) - 1_{B_j''}(x)\} \\ \subseteq \{x \in B_i'' \cap B_j'' \colon |d(x, (\partial B_i'') \cup (\partial B_j''))| < \rho\eta \vee \|x\|_2 \geq \sqrt{-2\log(\rho\eta)} + 1\}.\end{aligned} \tag{4.48}$$

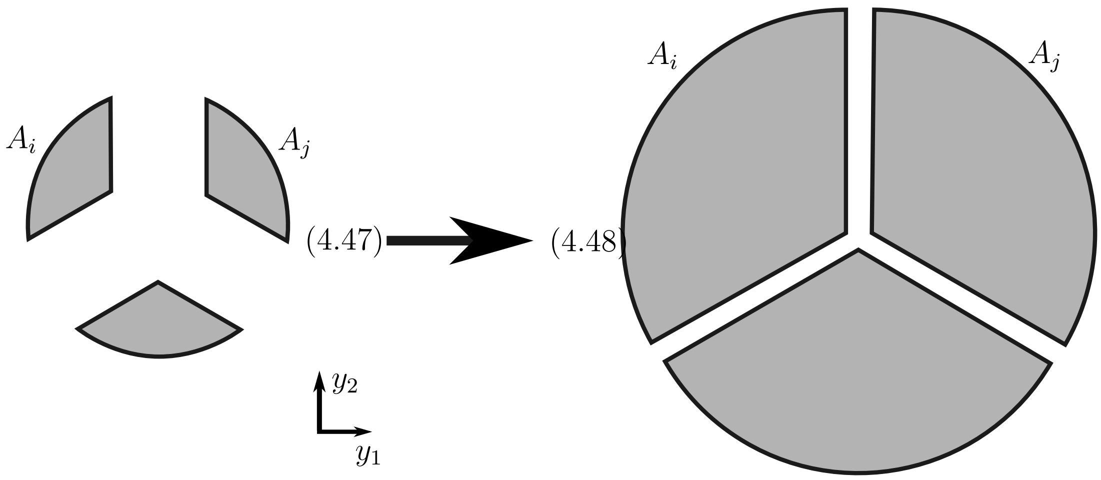


Figure 4.1: Depiction of Lemma 4.6.1

Proof. Fix $i, j \in \{1, 2, \ldots, k\}$ with $i \neq j$. By applying a rotation to $\mathbb{R}^n$, we assume that $B_i' \cap B_j' \subseteq \{x \in \mathbb{R}^n \colon x_1 = 0\}$ and $B_i' \subseteq \{x \in \mathbb{R}^n \colon x_1 \geq 0\}$. Assume that (4.47) and (4.45) hold. Let $n_i' \in \mathbb{R}^n$ denote the interior unit normal of B_i' such that n_i' is normal to $(\partial B_i') \setminus B_j'$, and let $n_j' \in \mathbb{R}^n$ denote the interior unit normal of B_j' such that n_j' is normal to $(\partial B_j') \setminus B_i'$. Define B_i, B_j such that

$$\begin{aligned}B_i = B_{i, \frac{2\eta}{\sqrt{-2\log\eta}}} &:= B_i' \cup \{x \in \mathbb{R}^n \colon x_1 \geq 0 \wedge \langle n_i', x/\|x\|_2\rangle \geq -2\eta/\sqrt{-2\log\eta}\}, \\ B_j = B_{j, \frac{2\eta}{\sqrt{-2\log\eta}}} &:= B_j' \cup \{x \in \mathbb{R}^n \colon x_1 \leq 0 \wedge \langle n_j', x/\|x\|_2\rangle \geq -2\eta/\sqrt{-2\log\eta}\}.\end{aligned} \tag{4.49}$$

Let $x = (x_1, \ldots, x_n) \in B_i \cup B_j$. If $x_1 < \sqrt{n}\sqrt{1-\rho^2}/\rho$, then

$$\gamma_n\left(\delta_{\frac{(B_i \cap B_j) - x\rho}{\sqrt{1-\rho^2}}}\right) \geq \frac{e^{-n/2}}{\sqrt{2\pi}} \int_{\sqrt{n}}^{\infty} e^{-t^2/2} dt/\sqrt{2\pi} \geq \frac{e^{-n/2}}{2\pi} \frac{1}{2\sqrt{n}} e^{-n/2} \geq \frac{1}{100\sqrt{n}} e^{-n}.$$

So, using Lemma 4.4.2, (4.8), and $\rho < 10^{-5} n^{-3/2} e^{-n}$, if $x \in B_i \cup B_j$ then

$$\rho^{-1}\mathrm{sign}(x_1) \cdot LT_\rho(1_{B_i} - 1_{B_j})(x) \geq \begin{cases} \frac{1}{9} |x_1| e^{-x_1^2 \rho^2/(2(1-\rho^2))} & , |x_1| \leq 1 \vee x_2 \geq 0 \\ \frac{1}{9} \frac{|x_1|}{\rho |x_2|} e^{-\|x\|_2^2 \rho^2/(2(1-\rho^2))} & , \rho x_2 \leq -1/\sqrt{3}. \end{cases} \tag{4.50}$$

Let $\sigma \colon \mathbb{R}^n \to \mathbb{R}^n$ be a rotation such that the x_1-axis is fixed. For any such rotation, let

$$g(x) = g_\sigma(x) := 1_{A_i}(\sigma x) - 1_{A_j}(\sigma x) - (1_{B_i}(\sigma x) - 1_{B_j}(\sigma x)). \tag{4.51}$$

By (4.45), and since $B_i \cup B_j$ is symmetric with respect to reflection across $B_i \cap B_j \subseteq \{x \in \mathbb{R}^n \colon x_1 = 0\}$, $\exists\ \lambda > 0$ such that $\int_{\mathbb{R}^n} y(1_{A_i}(y) - 1_{A_j}(y)) d\gamma_n(y) = \lambda \int_{\mathbb{R}^n} y(1_{B_i}(y) - 1_{B_j}(y)) d\gamma_n(y)$. So $\int_{\mathbb{R}^n} y_2 g(y) d\gamma_n(y) = 0$, for all such rotations σ. For $x \in \mathbb{R}^n$,

$$\left| LT_\rho(1_{A_i} - 1_{A_j})(\sigma x) - LT_\rho(1_{B_i} - 1_{B_j})(\sigma x) \right| \leq |LT_\rho g(x)|. \tag{4.52}$$

Step 1. An estimate for large x.

By (4.51), $|g| \leq 2$. Applying (4.47) and (4.49) and the inclusion-exclusion principle, $g = 0$ on the set

$$\{y \in \mathbb{R}^n \colon d(\sigma y - \rho x, (\partial B_i') \cup (\partial B_j')) > \eta + 3\eta \\ \wedge \|\sigma y - \rho x\|_2 \leq \sqrt{-2\log\eta} + (\rho + \eta)\sqrt{-2\log\rho}\}.$$

Let $x \in \mathbb{R}^n$ with $\|x\|_2^2 \leq -4\log(\eta\rho)$. Since $0 < \eta < \rho$, we have $\rho\|x\|_2 \leq -4\rho\log(\eta\rho) \leq -8\rho\log\eta$. By (4.47), (4.46) and the inclusion-exclusion principle, $g \neq 0$ only on the following sets: $\{y \in \mathbb{R}^n \colon |d(\sigma y - \rho x, (\partial B_i') \cup (\partial B_j'))| \leq 4\eta/\sqrt{1-\rho^2}\}$ and $\{y \in \mathbb{R}^n \colon \|\sigma y - \rho x\|_2 \geq \sqrt{-2\log\eta}/\sqrt{1-\rho^2}\}$. Then Lemma 4.5.2 says

$$\sup_{\substack{t_1 \in [\min(x_1,0), \max(x_1,0)] \\ t_2 \in [\min(x_2,0), \max(x_2,0)] \\ \alpha \in [0,\rho]}} \left| \int_{\mathbb{R}^n} \sum_{\substack{\ell \in \mathbb{N}^n \colon \\ 0 \leq |\ell| \leq 3}} \prod_{i=1}^n |y_i|^{\ell_i} g((t_1, t_2, 0, \ldots, 0)\alpha + y\sqrt{1-\alpha^2}) d\gamma_n(y) \right| \\ \leq 500000 n^3 4\eta + 200(n+2)!((-2(1-2\rho)\log\eta)^{(n+1)/2} + 1)\eta^{1-2\rho}.$$

Using Lemma 4.5.3,

$$
\begin{aligned}
&\|x\|_2^2 \leq -4\log(\rho\eta) \,\wedge\, |x_1| \geq (\rho\eta)^{1/3} \\
&\Longrightarrow \rho^{-1}\left|LT_\rho(1_{A_i}-1_{A_j})(x_1,x_2,0,\ldots,0) - LT_\rho(1_{B_i}-1_{B_j})(x_1,x_2,0,\ldots,0)\right| \\
&\qquad \leq [|x_1| + 2\rho(|x_2|^2 + (n+1)\rho\,|x_2| + |x_1x_2| + 2n)] \\
&\qquad\qquad \cdot [500000n^3 4\eta + 200(n+2)!((-2(1-2\rho)\log\eta)^{(n+1)/2} + 1)\eta^{1-2\rho}] \\
&\qquad < 10^7(n+2)!(-2\log\eta)^{(n+5)/2}\eta^{1-2\rho}.
\end{aligned}
\tag{4.53}
$$

Also, by (4.50), and using that $0 < \eta < \rho < 10^{-5}n^{-3/2}e^{-n}$,

$$
\begin{aligned}
&\|x\|_2^2 \leq -4\log(\rho\eta) \,\wedge\, x \in B_i \cup B_j \\
&\quad\Longrightarrow \rho^{-1}\mathrm{sign}(x_1)\cdot(LT_\rho(1_{B_i}-1_{B_j})(x)) > \\
&\qquad\qquad \begin{cases} \frac{1}{9}\,|x_1|\,(\rho\eta)^{2\rho^2/(1-\rho^2)} & ,|x_1| \leq 1 \,\vee\, x_2 \geq 0 \\ \frac{1}{9}\frac{|x_1|}{\rho|x_2|}(\rho\eta)^{2\rho^2/(1-\rho^2)} & ,\rho x_2 \leq -1/\sqrt{3} \end{cases}.
\end{aligned}
\tag{4.54}
$$

Combining (4.53) and (4.54), using (4.51) and $0 < \eta < \rho < e^{-20(n+1)10^{12}(n+2)!}$,

$$
\begin{aligned}
&|x_1| \geq (\rho\eta)^{1/3} \wedge \|x\|_2^2 \leq -4\log(\rho\eta) \wedge x \in B_i \cup B_j \\
&\quad\Longrightarrow \rho^{-1}\left|LT_\rho(1_{A_i}-1_{A_j})(x) - LT_\rho(1_{B_i}-1_{B_j})(x)\right| \leq \eta^{4/5} \\
&\quad\wedge\, \rho^{-1}\mathrm{sign}(x_1)\cdot LT_\rho(1_{B_i}-1_{B_j})(x) \\
&\qquad\qquad \geq (\rho\eta)^{2\rho^2/(1-\rho^2)}(\rho\eta)^{1/3}\min\left(1, \frac{1}{\rho\sqrt{-4\log(\eta\rho)}}\right).
\end{aligned}
\tag{4.55}
$$

By (4.55),

$$
\begin{aligned}
&|x_1| \geq (\rho\eta)^{1/3} \,\wedge\, \|x\|_2^2 \leq -4\log(\rho\eta) \,\wedge\, x \in B_i \cup B_j \\
&\qquad\Longrightarrow \rho^{-1}\mathrm{sign}(x_1)\cdot LT_\rho(1_{A_i}-1_{A_j})(x) > 0.
\end{aligned}
\tag{4.56}
$$

Finally, applying (4.56) to Lemma 4.3.1 for all $i', j' \in \{1,\ldots,k\}$, $i' \neq j'$, and using (4.49) together with the inclusion-exclusion principle,

$$
\begin{aligned}
&|x_1| \geq (\rho\eta)^{1/3} \,\wedge\, \|x\|_2^2 \leq -4\log(\rho\eta) \,\wedge\, x \in B_i' \cup B_j' \\
&\qquad\Longrightarrow \mathrm{sign}(x_1)\cdot(1_{A_i}(x) - 1_{A_j}(x)) > 0.
\end{aligned}
\tag{4.57}
$$

Step 2. An estimate for small x.

For $x = (x_1,\ldots,x_n) \in \mathbb{R}^n$ with $\|x\|_2^2 \leq -2\log\rho$, we have $\rho\,\|x\|_2 \leq \rho\sqrt{-2\log\rho}$. Suppose also that $|x_1| \leq \eta$ and $x \in B_i \cup B_j$. By (4.49), (4.47), (4.57), (4.46) and the inclusion-exclusion principle, $g \neq 0$ only on the following sets:

$$\{y \in \mathbb{R}^n \colon |d(\sigma y - \rho x, (B_i \cap B_j) \cup [(B_i \cup B_j) \setminus (B_i' \cup B_j')])| \leq \eta/\sqrt{1-\rho^2}\},$$
$$\{y \in \mathbb{R}^n \colon |d(\sigma y - \rho x, (B_i \cap B_j) \cup [(B_i \cup B_j) \setminus (B_i' \cup B_j')])| \leq (\rho\eta)^{1/4}$$
$$\wedge \left\|\sigma y - \rho x\right\|_2 \geq \sqrt{-2\log\eta}\},$$
$$\{y \in \mathbb{R}^n \colon \left\|\sigma y - \rho x\right\|_2 \geq (1+1/10)\sqrt{-3\log(\eta\rho)}/\sqrt{1-\rho^2}\}.$$

We then apply Lemma 4.5.2 to get

$$\begin{aligned}
&\sup_{\substack{t_1\in[\min(x_1,0),\max(x_1,0)]\\ t_2\in[\min(x_2,0),\max(x_2,0)]\\ \alpha\in[0,\rho]}} \left| \int_{\mathbb{R}^n} \sum_{\substack{\ell\in\mathbb{N}^n\colon\\ 0\leq|\ell|\leq 3}} \prod_{i=1}^{n} |y_i|^{\ell_i}\, g((t_1,t_2,0,\ldots,0)\alpha + y\sqrt{1-\alpha^2})d\gamma_n(y)\right| \\
&\quad\leq 500000n^3\eta + 500000n^3(\rho\eta)^{1/4}(-2(1-\rho)^2\log\eta+1)\eta^{(1-\rho)^2} \\
&\qquad + 200(n+2)!((-3\log(\rho\eta))^{(n+1)/2}+1)(\rho\eta)^{3/2} + 1600(n+2)!2\eta/\sqrt{-2\log\eta}.
\end{aligned}$$

So, using Lemma 4.5.3 and $0<\eta<\rho<e^{-20(n+1)10^{12}n^3(n+2)!}$,

$$\begin{aligned}
&\rho^{3/4}\eta \leq |x_1| \leq \eta \,\wedge\, \left\|x\right\|_2^2 \leq -2\log\rho \,\wedge\, x\in B_i\cup B_j \\
&\Longrightarrow \rho^{-1}\left|LT_\rho(1_{A_i}-1_{A_j})(x_1,x_2,0,\ldots,0) - LT_\rho(1_{B_i}-1_{B_j})(x_1,x_2,0,\ldots,0)\right| \\
&\leq [|x_1| + 2\rho(|x_2|^2 + (n+1)\rho|x_2| + |x_1x_2| + 2n)] \\
&\cdot \Big[500000n^3 3\eta + 500000n^3(\rho\eta)^{1/4}(-2(1-\rho)^2\log\eta+1)\eta^{(1-\rho)^2} \\
&+200(n+2)!((-3\log(\rho\eta))^{(n+1)/2}+1)(\rho\eta)^{3/2} + 1600(n+2)!2\eta/\sqrt{-2\log\eta}\Big] \\
&< \frac{1}{10}\eta\rho^{3/4}.
\end{aligned} \tag{4.58}$$

Also, by (4.50),

$$\begin{aligned}
&\eta\rho^{3/4} \leq |x_1| \leq \eta \,\wedge\, \left\|x\right\|_2^2 \leq -2\log\rho \,\wedge\, x\in B_i\cup B_j \\
&\qquad \Longrightarrow \rho^{-1}\mathrm{sign}(x_1)\cdot LT_\rho(1_{B_i}-1_{B_j})(x) > \frac{1}{10}|x_1|.
\end{aligned} \tag{4.59}$$

Combining (4.58) and (4.59), and using (4.51),

$$\begin{aligned}
&\eta\rho^{3/4} \leq |x_1| \leq \eta \,\wedge\, \left\|x\right\|_2^2 \leq -2\log\rho \,\wedge\, x\in B_i\cup B_j \\
&\qquad \Longrightarrow \rho^{-1}\mathrm{sign}(x_1)\cdot LT_\rho(1_{A_i}-1_{A_j})(x) > 0.
\end{aligned} \tag{4.60}$$

Similarly, by (4.58) and (4.51), we have the following estimate.

$$\eta \leq |x_1| \leq (\rho\eta)^{1/3} \wedge \left\|x\right\|_2^2 \leq 1 \wedge x\in B_i\cup B_j \Longrightarrow \rho^{-1}\mathrm{sign}(x_1)\cdot LT_\rho(1_{A_i}-1_{A_j})(x) > 0. \tag{4.61}$$

Finally, applying (4.60) to Lemma 4.3.1 for all $i', j' \in \{1, \ldots, k\}$, and using (4.49) together with the inclusion-exclusion principle, (4.56) and (4.61),

$$\begin{aligned} \eta\rho^{3/4} \leq |x_1| \leq \eta \,\wedge\, \|x\|_2^2 \leq -2\log\rho \,\wedge\, x \in B_i' \cup B_j' \\ \Longrightarrow \mathrm{sign}(x_1) \cdot (1_{A_i}(x) - 1_{A_j}(x)) > 0. \end{aligned} \tag{4.62}$$

Step 3. An estimate for intermediate values of x.

In summary, (4.62) and (4.57) improve our initial assumption (4.47). We now repeat the above procedure with the improved assumptions. Before continuing, we need to redefine B_i, B_j. Via (4.49), let

$$B_i := B_{i,\min\left(\frac{2\rho^{3/4}\eta}{\sqrt{-2\log\rho}}, \frac{2\eta}{\sqrt{-2\log\eta}}\right)}, \; B_j := B_{j,\min\left(\frac{2\rho^{3/4}\eta}{\sqrt{-2\log\rho}}, \frac{2\eta}{\sqrt{-2\log\eta}}\right)} \tag{4.63}$$

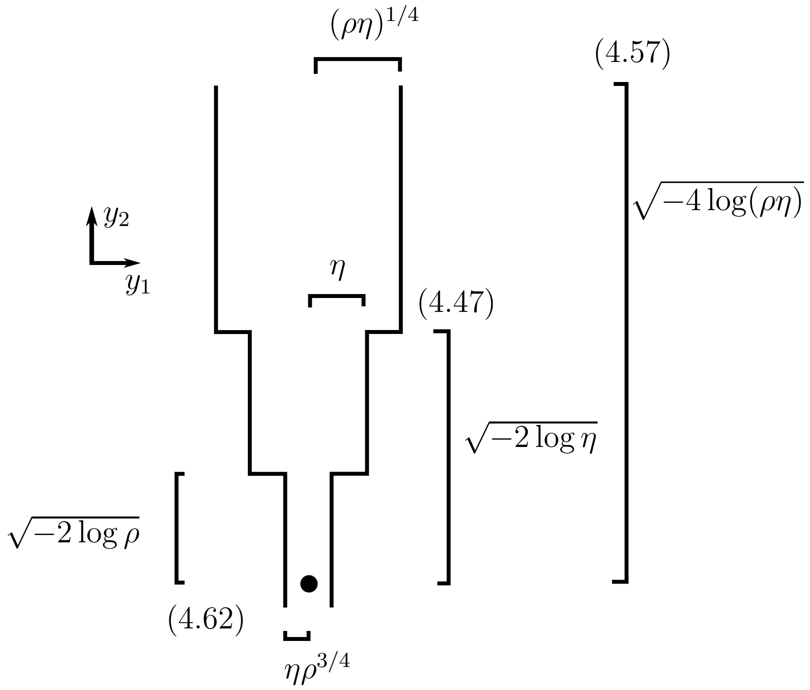


Figure 4.2: Integration regions where $g \neq 0$, near $B_i' \cap B_j'$ for (4.64).

Let x with $\|x\|_2^2 \leq -4\log\rho$, $\eta\rho \leq |x_1| \leq \eta$. Let $B := (B_i \cap B_j) \cup [(B_i \cup B_j) \setminus (B_i' \cup B_j')]$. Suppose $x \in B_i \cup B_j$ also. By (4.62), (4.47), (4.57), (4.46) and the inclusion-exclusion principle, $g\neq 0$ only on the following sets

$$\{y\in\mathbb{R}^n\colon |d(\sigma y-\rho x,B)|\leq \eta\rho^{3/4}/\sqrt{1-\rho^2}\},$$
$$\{y\in\mathbb{R}^n\colon |d(\sigma y-\rho x,B)|\leq \eta/\sqrt{1-\rho^2},\, \|\sigma y-\rho x\|_2>\sqrt{-2\log\rho}/\sqrt{1-\rho^2}\},$$
$$\{y\in\mathbb{R}^n\colon |d(\sigma y-\rho x,B)|\leq (\rho\eta)^{1/4}/\sqrt{1-\rho^2},\, \|\sigma y-\rho x\|_2\geq\sqrt{-2\log\eta}/\sqrt{1-\rho^2}\},$$
$$\{y\in\mathbb{R}^n\colon \|\sigma y-\rho x\|_2>(1+1/10)\sqrt{-3\log(\rho\eta)}/\sqrt{1-\rho^2}\}.$$

We then apply Lemma 4.5.2 to get

$$\begin{aligned}
&\sup_{\substack{t_1\in[\min(x_1,0),\max(x_1,0)]\\ t_2\in[\min(x_2,0),\max(x_2,0)]\\ \alpha\in[0,\rho]}}\left|\int_{\mathbb{R}^n}\sum_{\substack{\ell\in\mathbb{N}^n\colon\\ 0\leq|\ell|\leq 3}}\prod_{i=1}^n|y_i|^{\ell_i}\, g((t_1,t_2,0,\ldots,0)\alpha+y\sqrt{1-\alpha^2})d\gamma_n(y)\right|\\
&\quad\leq \eta\rho^{3/4}500000n^3+500000n^3\eta(-2(1-\sqrt{2}\rho)^2\log\rho+1)\rho^{(1-\sqrt{2}\rho)^2}\\
&\qquad+500000n^3(\eta\rho)^{1/4}(-2(1-\sqrt{2}\rho)^2\log\eta+1)\eta^{(1-\sqrt{2}\rho)^2}\\
&\qquad+200(n+2)!((-3\log(\eta\rho))^{(n+1)/2}+1)(\rho\eta)^{3/2}\\
&\qquad+1600(n+2)!\min(2\rho^{3/4}\eta/\sqrt{-2\log\rho},2\eta/\sqrt{-2\log\eta}\,).
\end{aligned}\tag{4.64}$$

Applying (4.64) to Lemma 4.5.3, using (4.51) and $\eta<\rho<e^{-20(n+1)^{10^{12}n^3(n+2)!}}$,

$$\begin{aligned}
&\|x\|_2^2\leq -4\log(\rho)\,\wedge\,\eta\rho\leq|x_1|\leq\eta\,\wedge x\in B_i\cup B_j\\
&\quad\Longrightarrow \rho^{-1}\left|LT_\rho(1_{A_i}-1_{A_j})(x)-LT_\rho(1_{B_i}-1_{B_j})(x)\right|<\frac{1}{10}\rho\eta.
\end{aligned}\tag{4.65}$$

Also, by (4.50),

$$\begin{aligned}
&\rho\eta\leq|x_1|\leq\eta\,\wedge\,\|x\|_2^2\leq -4\log\rho\,\wedge\,x\in B_i\cup B_j\\
&\quad\Longrightarrow \rho^{-1}\mathrm{sign}(x_1)\cdot LT_\rho(1_{B_i}-1_{B_j})(x)>\frac{1}{10}\rho\eta.
\end{aligned}\tag{4.66}$$

So, combining (4.65), (4.66) for all $i',j'\in\{1,\ldots,k\}$, $i'\neq j'$, Lemma 4.3.1, (4.63), and by applying the inclusion-exclusion principle, (4.56) and (4.61),

$$\rho\eta\leq|x_1|\leq\eta\wedge\|x\|_2^2\leq-4\log\rho\wedge x\in B_i'\cup B_j'\Longrightarrow \mathrm{sign}(x_1)\cdot(1_{A_i}(x)-1_{A_j}(x))>0.\tag{4.67}$$

Step 4. Iterating the estimate for intermediate values of x.

The estimate (4.67) now has a cascading effect on the estimates below. From (4.67),

$$\rho^{.9}\eta\leq|x_1|\leq\eta\wedge\|x\|_2^2\leq-4\log\rho\wedge x\in B_i'\cup B_j'\Longrightarrow \mathrm{sign}(x_1)\cdot(1_{A_i}(x)-1_{A_j}(x))>0.$$

This estimate can be iterated on itself. Let $K \in \mathbb{N}$, $K \geq 1$, and let $M \in \mathbb{N}$ with $0 \leq M \leq \sqrt{K}$. Suppose $\rho^{.9K} > \eta^{1/5}$. We prove by induction on K and M that

$$\begin{aligned} 2^{M^2}\eta\rho^{.9K} \leq |x_1| \leq \eta \,\wedge\, \|x\|_2^2 \leq -2^{M+2}\log\rho \,\wedge\, x \in B_i' \cup B_j' \\ \Longrightarrow \mathrm{sign}(x_1) \cdot (1_{A_i}(x) - 1_{A_j}(x)) > 0. \end{aligned} \tag{4.68}$$

We already verified the case $M = 0, K = 1$. We assume that, for $0 \leq m < M$,

$$\begin{aligned} 2^{m^2}\eta\rho^{.9K} \leq |x_1| \leq \eta \,\wedge\, \|x\|_2^2 \leq -2^{m+2}\log\rho \,\wedge\, x \in B_i' \cup B_j' \\ \Longrightarrow \mathrm{sign}(x_1) \cdot (1_{A_i}(x) - 1_{A_j}(x)) > 0. \end{aligned} \tag{4.69}$$

Assume also that, for $M \leq m \leq \sqrt{K-1}$ and $K \geq 1$,

$$\begin{aligned} 2^{m^2}\eta\rho^{.9(K-1)} \leq |x_1| \leq \eta \,\wedge\, \|x\|_2^2 \leq -2^{m+2}\log\rho \,\wedge\, x \in B_i' \cup B_j' \\ \Longrightarrow \mathrm{sign}(x_1) \cdot (1_{A_i}(x) - 1_{A_j}(x)) > 0. \end{aligned} \tag{4.70}$$

We will conclude that (4.69) holds for $m = M$, i.e.

$$\begin{aligned} 2^{M^2}\eta\rho^{.9K} \leq |x_1| \leq \eta \,\wedge\, \|x\|_2^2 \leq -2^{M+2}\log\rho \,\wedge\, x \in B_i' \cup B_j' \\ \Longrightarrow \mathrm{sign}(x_1) \cdot (1_{A_i}(x) - 1_{A_j}(x)) > 0. \end{aligned} \tag{4.71}$$

We repeat the calculations (4.63) through (4.67). Redefine B_i, B_j so that

$$B_i := B_{i,\min\left(\frac{2\eta\rho^{.9K}}{\sqrt{-4\log\rho}},\frac{2\eta}{\sqrt{-2\log\eta}}\right)}, \quad B_j := B_{j,\min\left(\frac{2\eta\rho^{.9K}}{\sqrt{-4\log\rho}},\frac{2\eta}{\sqrt{-2\log\eta}}\right)}. \tag{4.72}$$

If $M > 0$, we use (4.69) for $m = M - 1$. For any $M \geq 0$, we use (4.70) for $M \leq m \leq \sqrt{K}$. Let x, M with $\|x\|_2^2 \leq -2^{M+2}\log\rho \leq -2^{\lfloor\sqrt{K}\rfloor+2}\log\rho \leq -4\log\eta$, $2^{M^2}\eta\rho^{.9K} \leq |x_1| \leq \eta$, $x \in B_i \cup B_j$. Let $B := (B_i \cap B_j) \cup [(B_i \cup B_j) \setminus (B_i' \cup B_j')]$. Combining (4.69), (4.70), (4.47), (4.57), (4.46) and the inclusion-exclusion principle, $g \neq 0$ only on the following sets

$$\begin{aligned} &\{y \in \mathbb{R}^n \colon |d(\sigma y - \rho x, B)| \leq \min(M,1) \cdot 2^{(M-1)^2}\eta\rho^{.9K}/\sqrt{1-\rho^2}\}, \\ &\cup_{M \leq m \leq \lfloor\sqrt{K-1}\rfloor} \{y \in \mathbb{R}^n \colon |d(\sigma y - \rho x, B)| \leq 2^{m^2}\eta\rho^{.9(K-1)}/\sqrt{1-\rho^2}, \\ &\qquad\qquad \|\sigma y - \rho x\|_2 > \min(m,1) \cdot \sqrt{-2^{m+1}\log\rho}/\sqrt{1-\rho^2}\}, \\ &\{y \in \mathbb{R}^n \colon |d(\sigma y - \rho x, B)| \leq \eta, \|\sigma y - \rho x\|_2 > \sqrt{-2^{\lfloor\sqrt{K-1}\rfloor+2}\log\rho}/\sqrt{1-\rho^2}\}, \\ &\{y \in \mathbb{R}^n \colon |d(\sigma y - \rho x, B)| \leq (\rho\eta)^{1/4}/\sqrt{1-\rho^2}, \|\sigma y - \rho x\|_2 \geq \sqrt{-2\log\eta}/\sqrt{1-\rho^2}\}, \\ &\{y \in \mathbb{R}^n \colon \|\sigma y - \rho x\|_2 > (1+1/10)\sqrt{-3\log(\rho\eta)}/\sqrt{1-\rho^2}\}. \end{aligned}$$

We then apply Lemma 4.5.2 to get

$$\sup_{\substack{t_1\in[\min(x_1,0),\max(x_1,0)]\\ t_2\in[\min(x_2,0),\max(x_2,0)]\\ \alpha\in[0,\rho]}} \left| \int_{\mathbb{R}^n} \sum_{\substack{\ell\in\mathbb{N}^n:\\ 0\leq|\ell|\leq 3}} \prod_{i=1}^{n} |y_i|^{\ell_i}\, g((t_1,t_2,0,\ldots,0)\alpha + y\sqrt{1-\alpha^2})d\gamma_n(y)\right|$$
$$\begin{aligned}
&\leq \min(M,1)\cdot 2^{(M-1)^2}\eta\rho^{.9K}500000n^3\\
&\quad + \eta(-(1-\rho)^2 2^{\lfloor\sqrt{K-1}\rfloor+2}\log\rho+1)\rho^{(1-\rho)^2 2^{\lfloor\sqrt{K-1}\rfloor+1}}500000n^3\\
&\quad + 500000n^3\sum_{m=M}^{\lfloor\sqrt{K-1}\rfloor}\eta\rho^{.9(K-1)}(-(1-\rho)^2 2^{m+1}\log\rho+1)2^{m^2}\rho^{(1-\rho)^2 2^m\cdot\min(m,1)}\\
&\quad + 500000n^3(\eta\rho)^{1/4}(-2(1-\sqrt{2}\rho)^2\log\eta+1)\eta^{(1-\sqrt{2}\rho)^2}\\
&\quad + 200(n+2)!((-3\log(\eta\rho))^{(n+1)/2}+1)(\rho\eta)^{3/2}\\
&\quad + 1600(n+2)!\min(2\eta\rho^{.9K}/\sqrt{-4\log\rho}, 2\eta/\sqrt{-2\log\eta}).
\end{aligned}\tag{4.73}$$

Applying (4.73) to Lemma 4.5.3, using (4.51), $\eta<\rho<e^{-20(n+1)10^{12}n^3(n+2)!}$, and $\rho^{.9K}>\eta^{1/5}$,

$$\begin{aligned}
&\|x\|_2^2\leq -2^{M+2}\log(\rho)\,\wedge\, 2^{M^2}\eta\rho^{.9K}\leq|x_1|\leq\eta\,\wedge\, x\in B_i\cup B_j\\
&\qquad\Longrightarrow \rho^{-1}\left|LT_\rho(1_{A_i}-1_{A_j})(x)-LT_\rho(1_{B_i}-1_{B_j})(x)\right| < \frac{1}{10}2^{M^2}\eta\rho^{.9K}.
\end{aligned}\tag{4.74}$$

Also, by (4.50),

$$\begin{aligned}
&2^{M^2}\eta\rho^{.9K}\leq|x_1|\leq\eta\,\wedge\,\|x\|_2^2\leq -2^{M+2}\log\rho\,\wedge\, x\in B_i\cup B_j\\
&\qquad\Longrightarrow \rho^{-1}\mathrm{sign}(x_1)\cdot LT_\rho(1_{B_i}-1_{B_j})(x) > \frac{1}{10}2^{M^2}\eta\rho^{.9K}.
\end{aligned}\tag{4.75}$$

So, combining (4.74), (4.75) for all $i',j'\in\{1,\ldots,k\}$, $i'\neq j'$, Lemma 4.3.1, (4.72), and by applying the inclusion-exclusion principle, (4.56) and (4.61),

$$\begin{aligned}
&2^{M^2}\eta\rho^{.9K}\leq|x_1|\leq\eta\,\wedge\,\|x\|_2^2\leq -2^{M+2}\log\rho\,\wedge\, x\in B_i'\cup B_j'\\
&\qquad\Longrightarrow \mathrm{sign}(x_1)\cdot(1_{A_i}(x)-1_{A_j}(x))>0.
\end{aligned}\tag{4.76}$$

Thus, the inductive step is completed.

Let $K\in\mathbb{N}$ with $-2\log\eta\leq -2^{\lfloor\sqrt{K}\rfloor+2}\log\rho\leq -4\log\eta$. Then (4.76) and (4.47) say that

$$\begin{aligned}
&2^K\eta\rho^{.9K}\leq|x_1|\leq 1\,\wedge\,\|x\|_2^2\leq -2\log\eta\,\wedge\, x\in B_i'\cup B_j'\\
&\qquad\Longrightarrow \mathrm{sign}(x_1)\cdot(1_{A_i}(x)-1_{A_j}(x))>0.
\end{aligned}\tag{4.77}$$

Step 5. Another iterative estimate, now for larger values of x.

We perform another induction, though this time we hold K fixed and use the additional ingredient (4.77). Let $M, R \in \mathbb{N}$ with $0 \leq M \leq \sqrt{K}$, $R \geq 0$ such that $\rho^{.9(K+R)} > \eta^{1/5}$. We will induct on M and R. We assume that, for $0 \leq m < M$,

$$\begin{aligned} 2^{m^2}\eta\rho^{.9(K+R)} \leq |x_1| \leq \eta \,\wedge\, \|x\|_2^2 \leq -2^{m+2}\log\rho \,\wedge\, x \in B_i' \cup B_j' \\ \Longrightarrow \operatorname{sign}(x_1)\cdot(1_{A_i}(x) - 1_{A_j}(x)) > 0. \end{aligned} \tag{4.78}$$

We know that the case $R = 0, 0 \leq M \leq \sqrt{K}$ of (4.78) holds by (4.71). We therefore assume that $R \geq 1$. Assume also that, for $M \leq m \leq \sqrt{K}$,

$$\begin{aligned} 2^{m^2}\eta\rho^{.9(K+R-1)} \leq |x_1| \leq \eta \,\wedge\, \|x\|_2^2 \leq -2^{m+2}\log\rho \,\wedge\, x \in B_i' \cup B_j' \\ \Longrightarrow \operatorname{sign}(x_1)\cdot(1_{A_i}(x) - 1_{A_j}(x)) > 0. \end{aligned} \tag{4.79}$$

We will conclude that (4.78) holds for $m = M$, i.e.

$$\begin{aligned} 2^{M^2}\eta\rho^{.9(K+R)} \leq |x_1| \leq \eta \,\wedge\, \|x\|_2^2 \leq -2^{M+2}\log\rho \,\wedge\, x \in B_i' \cup B_j' \\ \Longrightarrow \operatorname{sign}(x_1)\cdot(1_{A_i}(x) - 1_{A_j}(x)) > 0. \end{aligned} \tag{4.80}$$

Redefine B_i, B_j so that

$$B_i := B_{i,\min\left(\frac{2\eta\rho^{.9(K+R)}}{\sqrt{-4\log\rho}},\frac{2\eta}{\sqrt{-2\log\eta}}\right)}, \qquad B_j := B_{j,\min\left(\frac{2\eta\rho^{.9(K+R)}}{\sqrt{-4\log\rho}},\frac{2\eta}{\sqrt{-2\log\eta}}\right)}. \tag{4.81}$$

If $M > 0$, we use (4.78) for $m = M - 1$. For any $M \geq 0$, we also use (4.79) for $M \leq m \leq \sqrt{K}$. Let x, M with $\|x\|_2^2 \leq -2^{M+2}\log\rho \leq -2^{\lfloor\sqrt{K}\rfloor+2}\log\rho \leq -4\log\eta$, $2^{M^2}\eta\rho^{.9K} \leq |x_1| \leq \eta$, $x \in B_i \cup B_j$. Let $B := (B_i \cap B_j) \cup [(B_i \cup B_j) \setminus (B_i' \cup B_j')]$. Combining (4.78), (4.79), (4.47), (4.57), (4.46), and the fact that $-2\log\eta \leq -2^{\lfloor\sqrt{K}\rfloor+2}\log\rho \leq -4\log\eta$, we conclude that $g \neq 0$ only on the following sets:

$$\begin{aligned}
&\{y \in \mathbb{R}^n \colon |d(\sigma y - \rho x, B)| \leq \min(M,1)\cdot 2^{(M-1)^2}\eta\rho^{.9(K+R)}/\sqrt{1-\rho^2}\},\\
&\cup_{M \leq m \leq \lfloor\sqrt{K}\rfloor}\{y \in \mathbb{R}^n \colon |d(\sigma y - \rho x, B)| \leq 2^{m^2}\eta\rho^{.9(K+R-1)}/\sqrt{1-\rho^2},\\
&\qquad\qquad\qquad\qquad \|\sigma y - \rho x\|_2 > \min(m,1)\cdot\sqrt{-2^{m+1}\log\rho}/\sqrt{1-\rho^2}\},\\
&\{y \in \mathbb{R}^n \colon |d(\sigma y - \rho x, B)| \leq (\rho\eta)^{1/4}/\sqrt{1-\rho^2}, \|\sigma y - \rho x\|_2 \geq \sqrt{-2\log\eta}/\sqrt{1-\rho^2}\},\\
&\{y \in \mathbb{R}^n \colon \|\sigma y - \rho x\|_2 > (1+1/10)\sqrt{-3\log(\rho\eta)}/\sqrt{1-\rho^2}\}.
\end{aligned}$$

We then apply Lemma 4.5.2 to get

$$\begin{aligned}
&\sup_{\substack{t_1\in[\min(x_1,0),\max(x_1,0)]\\ t_2\in[\min(x_2,0),\max(x_2,0)]\\ \alpha\in[0,\rho]}} \left|\int_{\mathbb{R}^n} \sum_{\substack{\ell\in\mathbb{N}^n:\\ 0\leq|\ell|\leq 3}} \prod_{i=1}^{n} |y_i|^{\ell_i}\, g((t_1,t_2,0,\ldots,0)\alpha + y\sqrt{1-\alpha^2})d\gamma_n(y)\right| \\
&\quad\leq \min(M,1)\cdot 2^{(M-1)^2}\eta\rho^{.9(K+R)}500000n^3 \\
&\quad\quad + 500000n^3 \sum_{m=M}^{\lfloor\sqrt{K}\rfloor} \eta\rho^{.9(K+R-1)}(-(1-\rho)^2 2^{m+1}\log\rho + 1)2^{m^2}\rho^{(1-\rho)^2 2^m} \\
&\quad\quad + 500000n^3(\eta\rho)^{1/4}(-2(1-\sqrt{2}\rho)^2\log\eta + 1)\eta^{(1-\sqrt{2}\rho)^2} \\
&\quad\quad + 200(n+2)!((-3\log(\eta\rho))^{(n+1)/2} + 1)(\rho\eta)^{3/2} \\
&\quad\quad + 1600(n+2)!\min(2\eta\rho^{.9(K+R)}/\sqrt{-4\log\rho}, 2\eta/\sqrt{-2\log\eta}).
\end{aligned} \tag{4.82}$$

Applying (4.82) to Lemma 4.5.3, using (4.51), $\eta < \rho < e^{-20(n+1)10^{12}n^3(n+2)!}$, and $\rho^{.9(K+R)} > \eta^{1/5}$,

$$\begin{aligned}
&\|x\|_2^2 \leq -2^{M+2}\log(\rho) \,\wedge\, 2^{M^2}\eta\rho^{.9(K+R)} \leq |x_1| \leq \eta \,\wedge\, x \in B_i \cup B_j \\
&\quad\Longrightarrow \rho^{-1}\left|LT_\rho(1_{A_i} - 1_{A_j})(x) - LT_\rho(1_{B_i} - 1_{B_j})(x)\right| < \frac{1}{10}2^{M^2}\eta\rho^{.9(K+R)}.
\end{aligned} \tag{4.83}$$

Also, by (4.50),

$$\begin{aligned}
&2^{M^2}\eta\rho^{.9(K+R)} \leq |x_1| \leq \eta \,\wedge\, \|x\|_2^2 \leq -2^{M+2}\log\rho \,\wedge\, x \in B_i \cup B_j \\
&\quad\Longrightarrow \rho^{-1}\mathrm{sign}(x_1)\cdot LT_\rho(1_{B_i} - 1_{B_j})(x) > \frac{1}{10}2^{M^2}\eta\rho^{.9(K+R)}.
\end{aligned} \tag{4.84}$$

So, combining (4.83), (4.84) for all $i', j' \in \{1,\ldots,k\}$, $i' \neq j'$, Lemma 4.3.1, (4.81), and by applying the inclusion-exclusion principle, (4.56) and (4.61),

$$\begin{aligned}
&2^{M^2}\eta\rho^{.9(K+R)} \leq |x_1| \leq \eta \,\wedge\, \|x\|_2^2 \leq -2^{M+2}\log\rho \,\wedge\, x \in B_i' \cup B_j' \\
&\quad\Longrightarrow \mathrm{sign}(x_1)\cdot(1_{A_i}(x) - 1_{A_j}(x)) > 0.
\end{aligned} \tag{4.85}$$

Thus, the inductive step is completed. Let $M = \lfloor\sqrt{K}\rfloor$. Let $R \in \mathbb{N}$ such that $\eta^{1/5} \leq \rho^{.9(K+R)} \leq \eta^{1/5}\rho^{-.9}$. If no such R exists, then $\rho^{.9K} < \eta^{1/5}$, so $\rho^{.45K} < \eta^{1/10}$, and (4.87) below holds by combining (4.77) and (4.47). Otherwise, $R \geq 0$, so (4.85) and (4.47) say that

$$\begin{aligned}
&2^K\eta^{6/5}\rho^{-.9} \leq |x_1| \leq 1 \,\wedge\, \|x\|_2^2 \leq -2\log\eta \,\wedge\, x \in B_i' \cup B_j' \\
&\quad\Longrightarrow \mathrm{sign}(x_1)\cdot(1_{A_i}(x) - 1_{A_j}(x)) > 0.
\end{aligned} \tag{4.86}$$

Since $\eta^{1/5} \leq \rho^{.9(K+R)} \leq \eta^{1/5}\rho^{-.9}$, note that $\eta^{1/10} \leq \rho^{.45(K+R)} \leq \eta^{1/10}\rho^{-.45}$, so for $R \geq 2$, we have $2^K\eta^{6/5}\rho^{-.9} \leq 2^K\rho^{.45K}\rho^{.45R}\eta^{11/10}\rho^{-.9} < \eta^{11/10}$. If $R = 1$, and if $K \geq 2$, note that $2^K\eta^{6/5}\rho^{-.9} \leq 2^K\rho^{.2K}\rho^{.25K}\rho^{.45}\eta^{11/10}\rho^{-.9} < \eta^{11/10}$. If $R = 0$, $K \geq 3$ then $2^K\eta^{6/5}\rho^{-.9} \leq 2^K\rho^{.1K}\rho^{.35K}\eta^{11/10}\rho^{-.9} < \eta^{11/10}$. If $1 \leq R + K \leq 3$, then $(1/5)\log\eta \leq 3\log\rho$ and $2\log\rho \leq (1/5)\log\eta$, so (4.73) directly implies (4.87). More specifically, by (4.73), Lemma 4.5.3,(4.56) and (4.61), $\mathrm{sign}(x_1)\cdot(1_{A_i} - 1_{A_j})(x) > 0$ for $x \in B_i' \cup B_j'$ with $\|x\|_2^2 \leq -2\log\eta$ and $\eta\rho^{.9K}\rho^{.9} \leq |x_1| \leq \eta$. Now, $\rho^{.45(K+R)} \leq \eta^{1/10}\rho^{-.45}$, so $\eta\rho^{.9K}\rho^{.9} = \eta\rho^{.45K}\rho^{.45K}\rho^{.9} \leq \eta\rho^{.45K}\rho^{.45(K+R)} \leq \eta^{11/10}$.

In the latter case, (4.87) follows, and in the former cases, (4.86) implies

$$\begin{aligned}&\eta^{11/10} \leq |x_1| \leq 1 \,\wedge\, \|x\|_2^2 \leq -2\log\eta \,\wedge\, x \in B_i' \cup B_j' \\ &\qquad\Longrightarrow \mathrm{sign}(x_1)\cdot(1_{A_i}(x) - 1_{A_j}(x)) > 0.\end{aligned} \tag{4.87}$$

In all cases, (4.87) holds. We can finally use (4.87) to conclude the proof.

Step 6. Using Step 5 to get an estimate for large values of x.

Redefine B_i, B_j so that

$$B_i := B_{i,2\eta^{11/10}/\sqrt{-2\log\eta}}, \quad B_j := B_{j,2\eta^{11/10}/\sqrt{-2\log\eta}}. \tag{4.88}$$

Let x with $\|x\|_2^2 \leq -4\log(\eta\rho) \leq -8\log\eta$ and $\eta^{21/20}\rho^{1/2} \leq |x_1| \leq (\eta\rho)^{1/4}$. Let $B := (B_i \cap B_j) \cup [(B_i \cup B_j) \setminus (B_i' \cup B_j')]$. Suppose $x \in B_i \cup B_j$ also. Combining (4.87), (4.57), and (4.46), $g \neq 0$ only on the following sets:

$$\begin{aligned}&\{y \in \mathbb{R}^n\colon |d(\sigma y - \rho x, B)| \leq \eta^{11/10}/\sqrt{1-\rho^2}\},\\ &\{y \in \mathbb{R}^n\colon |d(\sigma y - \rho x, B)| \leq (\rho\eta)^{1/4}/\sqrt{1-\rho^2}, \|\sigma y - \rho x\|_2 \geq \sqrt{-2\log\eta}/\sqrt{1-\rho^2}\},\\ &\{y \in \mathbb{R}^n\colon\ \|\sigma y - \rho x\|_2 > (1 + 1/10)\sqrt{-3\log(\rho\eta)}/\sqrt{1-\rho^2}\}.\end{aligned}$$

We then apply Lemma 4.5.2 to get

$$\begin{aligned}&\sup_{\substack{t_1\in[\min(x_1,0),\max(x_1,0)]\\ t_2\in[\min(x_2,0),\max(x_2,0)]\\ \alpha\in[0,\rho]}} \left|\int_{\mathbb{R}^n} \sum_{\substack{\ell\in\mathbb{N}^n\colon\\ 0\leq|\ell|\leq 3}} \prod_{i=1}^n |y_i|^{\ell_i}\, g((t_1,t_2,0,\ldots,0)\alpha + y\sqrt{1-\alpha^2})d\gamma_n(y)\right|\\ &\quad\leq 500000n^3\eta^{11/10} + 500000n^3(\rho\eta)^{1/4}(-2(1-2\rho)^2\log\eta + 1)\eta^{(1-2\rho)^2}\\ &\qquad + 200(n+2)!((-3\log(\eta\rho))^{(n+1)/2} + 1)(\rho\eta)^{3/2}\\ &\qquad\quad + 1600(n+2)!2\eta^{11/10}/\sqrt{-2\log\eta}.\end{aligned} \tag{4.89}$$

Applying (4.89) to Lemma 4.5.3, using (4.51) and $\eta < \rho < e^{-20(n+1)^{10^{12}n^3(n+2)!}}$,

$$\begin{aligned}&\|x\|_2^2 \leq -4\log(\eta\rho) \,\wedge\, \eta^{21/20}\rho^{1/2} \leq |x_1| \leq (\eta\rho)^{1/4} \,\wedge\, x \in B_i \cup B_j\\ &\qquad\Longrightarrow \rho^{-1}\left|LT_\rho(1_{A_i} - 1_{A_j})(x) - LT_\rho(1_{B_i} - 1_{B_j})(x)\right| < \frac{1}{10}\eta^{21/20}\rho^{1/2}.\end{aligned} \tag{4.90}$$

Also, by (4.50),

$$\begin{aligned}\eta^{21/20}\rho^{1/2} \leq |x_1| \leq (\eta\rho)^{1/4} \,\wedge\, \|x\|_2^2 - 4\log(\eta\rho) \,\wedge\, x \in B_i \cup B_j \\ \Longrightarrow \rho^{-1}\mathrm{sign}(x_1)\cdot LT_\rho(1_{B_i} - 1_{B_j})(x) > \frac{1}{10}\eta^{21/20}\rho^{1/2}.\end{aligned} \tag{4.91}$$

So, combining (4.90), (4.91) for all $i', j' \in \{1, \ldots, k\}$, $i' \neq j'$, Lemma 4.3.1, (4.88), and by applying the inclusion-exclusion principle, (4.56) and (4.61),

$$\begin{aligned}\eta^{21/20}\rho^{1/2} \leq |x_1| \leq (\eta\rho)^{1/4} \,\wedge\, \|x\|_2^2 \leq -4\log(\eta\rho) \,\wedge\, x \in B_i' \cup B_j' \\ \Longrightarrow \mathrm{sign}(x_1)\cdot(1_{A_i}(x) - 1_{A_j}(x)) > 0.\end{aligned} \tag{4.92}$$

So, (4.92) and (4.57) say that

$$\begin{aligned}\eta^{21/20}\rho^{1/2} \leq |x_1| \,\wedge\, \|x\|_2^2 \leq -4\log(\eta\rho) \,\wedge\, x \in B_i' \cup B_j' \\ \Longrightarrow \mathrm{sign}(x_1)\cdot(1_{A_i}(x) - 1_{A_j}(x)) > 0.\end{aligned} \tag{4.93}$$

Finally, we use (4.93) in place of (4.87) and repeat the computations (4.89) through (4.92) to get

$$\eta\rho \leq |x_1| \,\wedge\, \|x\|_2^2 \leq -4\log(\eta\rho) \,\wedge\, x \in B_i' \cup B_j' \Longrightarrow \mathrm{sign}(x_1)\cdot(1_{A_i}(x) - 1_{A_j}(x)) > 0. \tag{4.94}$$

In conclusion, (4.48) follows from (4.94) and (4.45), letting $B_i'' := B_i'$ and $B_j'' := B_j'$.

Step 7. Completing the proof.

For completeness, we derive (4.94). Redefine B_i, B_j so that

$$B_i := B_{i,2\eta^{21/20}\rho^{1/2}/\sqrt{-4\log(\eta\rho)}}, \quad B_j := B_{j,2\eta^{21/20}\rho^{1/2}/\sqrt{-4\log(\eta\rho)}}. \tag{4.95}$$

Let x with $\|x\|_2^2 \leq -4\log(\eta\rho) \leq -8\log\eta$ and $\eta\rho \leq |x_1| \leq (\eta\rho)^{1/3}$. Let $B := (B_i \cap B_j) \cup [(B_i \cup B_j) \setminus (B_i' \cup B_j')]$. Suppose $x \in B_i \cup B_j$ also. Combining (4.93), (4.57), and (4.46), $g \neq 0$ only on the following sets:

$\{y \in \mathbb{R}^n \colon |d(\sigma y - \rho x, B)| \leq \eta^{21/20}\rho^{1/2}/\sqrt{1-\rho^2}\}$,
$\{y \in \mathbb{R}^n \colon |d(\sigma y - \rho x, B)| \leq (\rho\eta)^{1/3}/\sqrt{1-\rho^2}, \|\sigma y - \rho x\|_2 \geq \sqrt{-2\log\eta}/\sqrt{1-\rho^2}\}$,
$\{y \in \mathbb{R}^n \colon \|\sigma y - \rho x\|_2 > (1 + 1/10)\sqrt{-3\log(\rho\eta)}/\sqrt{1-\rho^2}\}$.

We then apply Lemma 4.5.2 to get

$$\begin{aligned}
&\sup_{\substack{t_1\in[\min(x_1,0),\max(x_1,0)]\\ t_2\in[\min(x_2,0),\max(x_2,0)]\\ \alpha\in[0,\rho]}} \left| \int_{\mathbb{R}^n} \sum_{\substack{\ell\in\mathbb{N}^n:\\ 0\leq|\ell|\leq 3}} \prod_{i=1}^{n} |y_i|^{\ell_i}\, g((t_1,t_2,0,\ldots,0)\alpha + y\sqrt{1-\alpha^2}) d\gamma_n(y) \right| \\
&\quad\leq 500000n^3\eta^{21/20}\rho^{1/2} + 500000n^3(\rho\eta)^{1/3}(-2(1-2\rho)^2\log\eta+1)\eta^{(1-2\rho)^2} \\
&\qquad + 200(n+2)!((-3\log(\eta\rho))^{(n+1)/2}+1)(\rho\eta)^{3/2} \\
&\qquad + 1600(n+2)!2\eta^{6/5}\rho^{1/2}/\sqrt{-4\log(\eta\rho)}.
\end{aligned} \tag{4.96}$$

Applying (4.96) to Lemma 4.5.3, using (4.51) and $\eta<\rho<e^{-20(n+1)^{10^{12}n^3(n+2)!}}$,

$$\begin{aligned}
&\|x\|_2^2 \leq -4\log(\eta\rho) \,\wedge\, \eta\rho\leq|x_1|\leq(\eta\rho)^{1/3} \,\wedge\, x\in B_i\cup B_j \\
&\quad\Longrightarrow \rho^{-1}\left|LT_\rho(1_{A_i}-1_{A_j})(x) - LT_\rho(1_{B_i}-1_{B_j})(x)\right| < \frac{1}{10}\eta\rho.
\end{aligned} \tag{4.97}$$

Also, by (4.50),

$$\begin{aligned}
&\eta\rho\leq|x_1|\leq(\eta\rho)^{1/3} \,\wedge\, \|x\|_2^2 \leq -4\log(\eta\rho) \,\wedge\, x\in B_i\cup B_j \\
&\quad\Longrightarrow \rho^{-1}\mathrm{sign}(x_1)\cdot LT_\rho(1_{B_i}-1_{B_j})(x) > \frac{1}{10}\eta\rho.
\end{aligned} \tag{4.98}$$

So, combining (4.97), (4.98) for all $i',j'\in\{1,\ldots,k\}$, $i'\neq j'$, Lemma 4.3.1, (4.95), and by applying the inclusion-exclusion principle, (4.56) and (4.61),

$$\begin{aligned}
&\eta\rho\leq|x_1|\leq(\eta\rho)^{1/3} \,\wedge\, \|x\|_2^2 \leq -4\log(\eta\rho) \,\wedge\, x\in B_i'\cup B_j' \\
&\quad\Longrightarrow \mathrm{sign}(x_1)\cdot(1_{A_i}(x)-1_{A_j}(x)) > 0.
\end{aligned} \tag{4.99}$$

Then, (4.99) and (4.57) say that

$$\eta\rho\leq|x_1| \,\wedge\, \|x\|_2^2 \leq -4\log(\eta\rho) \,\wedge\, x\in B_i'\cup B_j' \Longrightarrow \mathrm{sign}(x_1)\cdot(1_{A_i}(x)-1_{A_j}(x)) > 0. \tag{4.100}$$

Finally, (4.94) follows from (4.100), completing the proof. □

4.7 Proof of the Main Theorem

We now combine the Lemmas of the previous sections, as described in Section 4.1. The main effort involves verifying the assumption of the Main Lemma 4.6.1. Once this is done, Lemma 4.6.1 can be iterated infinitely many times to complete the proof.

Theorem 4.7.1. *Fix $k=3$, $n \geq 2$. Define $\Delta_k(\gamma_n)$ as in Definition 4.2.1 and define ψ_ρ as in (4.28). Let $\{C_i\}_{i=1}^k \subseteq \mathbb{R}^n$ be a regular simplicial conical partition. Then there exists $\rho_0 = \rho_0(n,k) > 0$ such that, for all $\rho \in (0, \rho_0)$, $(1_{C_1}, \ldots, 1_{C_k})$ uniquely achieves the following supremum, up to rotation*

$$\sup_{(f_1,\ldots,f_k)\in\Delta_k(\gamma_n)} \rho^{-1} \sum_{i=1}^k \int_{\mathbb{R}^n} f_i L T_\rho f_i d\gamma_n = \sup_{(f_1,\ldots,f_k)\in\Delta_k(\gamma_n)} \psi_\rho(f_1,\ldots,f_k).$$

Proof. Within the proof, we will assert that $\rho > 0$ satisfies several upper bounds, and then at the end of the proof, we will define ρ_0 as the minimum of these upper bounds. By Lemma 4.3.1, let $\{A_i\}_{i=1}^k$ be a partition of $\mathbb{R}^n$ such that

$$\psi_\rho(1_{A_1},\ldots,1_{A_k}) = \sup_{(f_1,\ldots,f_k)\in\Delta_k(\gamma_n)} \psi_\rho(f_1,\ldots,f_k). \tag{4.101}$$

By (4.11), write

$$\rho^{-1}\sum_{i=1}^k \int_{\mathbb{R}^n} 1_{A_i} L T_\rho 1_{A_i} d\gamma_n = \sum_{i=1}^k \sum_{\ell\in\mathbb{N}^n} |\ell| \left| \int_{\mathbb{R}^n} 1_{A_i} \sqrt{\ell!} h_\ell d\gamma_n \right|^2 \rho^{|\ell|-1}. \tag{4.102}$$

Step 1. The partition $\{A_i\}_{i=1}^k$ is close to being simplicial.

For $i \in \{1,\ldots,k\}$, let $z_i := \int_{A_i} x d\gamma_n(x) \in \mathbb{R}^n$. Subtracting the $|\ell|=1$ term from both sides of (4.102), treating the remaining terms as error terms, and using that $\|1_{A_i}\|_{L_2(\gamma_n)} \leq 1$ for all $i=1,\ldots,k$,

$$\left| \rho^{-1}\sum_{i=1}^k \int_{\mathbb{R}^n} 1_{A_i} L T_\rho 1_{A_i} d\gamma_n - \sum_{i=1}^k \|z_i\|_{\ell_2^n}^2 \right| \leq 3k\rho. \tag{4.103}$$

Therefore,

$$\begin{aligned}
\sum_{i=1}^k \|z_i\|_{\ell_2^n}^2 &\overset{(4.13)}{=} \psi_0(1_{A_1},\ldots,1_{A_k}) \overset{(4.103)}{\geq} \psi_\rho(1_{A_1},\ldots,1_{A_k}) - 3k\rho \\
&\overset{(4.101)}{\geq} \psi_\rho(1_{B_1},\ldots,1_{B_k}) - 3k\rho \\
&\overset{(4.103)}{\geq} \psi_0(1_{B_1},\ldots,1_{B_k}) - 6k\rho \overset{(\text{Lemma } 4.2.5)}{=} \sup_{(f_1,\ldots,f_k)\in\Delta_k(\gamma_n)} \psi_0(f_1,\ldots,f_k) - 6k\rho.
\end{aligned}$$

Step 2. Applying a small rotation.

For $i \in \{1,\ldots,k\}$, let $w_i := \int_{B_i} x d\gamma_n(x)$. Let $\rho > 0$ such that $6k\rho < 10^{-2}$. Then by Lemma 4.2.8,

$$d_2(\{A_i\}_{i=1}^k, \{B_i\}_{i=1}^k) < 6(6k\rho)^{1/8}, \tag{4.104}$$

$$\inf_{\sigma \in SO(n)} \left(\sum_{i=1}^{k} \|\sigma z_i - w_i\|_{\ell_2^n}^2 \right)^{1/2} < 6(6k\rho)^{1/8}. \tag{4.105}$$

Note that (4.105) follows from (4.104) by Hilbert space duality and since the set of functions $\{x_i\}_{i=1}^n$ are contained in the orthonormal basis $\{h_\ell \sqrt{\ell!}\}_{\ell \in \mathbb{N}^n}$ of $L_2(\gamma_n)$.

Let $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, let $i, j \in \{1, \ldots, k\}$, $i \neq j$, and write the following equality of L_2 functions

$$1_{A_i}(x) - 1_{A_j}(x) =: \sum_{\ell \in \mathbb{N}^n} c_\ell h_\ell(x) \sqrt{\ell!}. \tag{4.106}$$

Let $\ell = (\ell_1, \ldots, \ell_n) \in \mathbb{N}^n$. By applying an orthogonal change of coordinates to $\{A_p\}_{p=1}^k$, we may assume that $c_\ell = 0$ when $|\ell| = 1$, $\ell_1 = 0$. By (4.8) and (4.5), write

$$\rho^{-1} L T_\rho (1_{A_i} - 1_{A_j})(x) = \sum_{\ell \in \mathbb{N}^n} c_\ell \, |\ell| \, \rho^{|\ell|-1} h_\ell(x) \sqrt{\ell!}. \tag{4.107}$$

Since $d_2(\{A_p\}_{p=1}^k, \{B_p\}_{p=1}^k) < 6(6k\rho)^{1/8}$, there exists $\{B_p''\}_{p=1}^k$ a regular simplicial conical partition, such that $(\sum_{p=1}^k \|1_{A_p} - 1_{B_p''}\|_{L_2(\gamma_n)}^2)^{1/2} < 6(6k\rho)^{1/8}$. In particular, by Hilbert space duality,

$$\left\| \int_{\mathbb{R}^n} x(1_{A_i}(x) - 1_{A_j}(x) - (1_{B_i''}(x) - 1_{B_j''}(x))) d\gamma_n(x) \right\|_{\ell_2^n} < 6(6k\rho)^{1/8}, \tag{4.108}$$

$$\left\| \int_{\mathbb{R}^n} x(1_{(A_i \cup A_j)^c}(x) - 1_{(B_i'' \cup B_j'')^c}(x)) d\gamma_n(x) \right\|_{\ell_2^n} < 6(6k\rho)^{1/8}. \tag{4.109}$$

Since $k = 3$, and since $\sum_{p=1}^k \int_{A_p} x d\gamma_n(x) = \int_{\mathbb{R}^n} x d\gamma_n(x) = 0$, there exists a 2-dimensional plane $\Pi \subseteq \mathbb{R}^n$ such that $0 \in \Pi$ and such that, for all $p \in \{1, \ldots, k\}$, $\int_{A_p} x d\gamma_n(x) \in \Pi$. Without loss of generality, Π contains the x_1 and x_2 axes.

Let $\{B_p'\}_{p=1}^k$ be a regular simplicial conical partition such that

$$\left(\sum_{p=1}^{k} \|1_{B_p'} - 1_{B_p''}\|_{L_2(\gamma_n)}^2 \right)^{1/2} < 10(6k\rho)^{1/16}, \tag{4.110}$$

such that for fixed $i \neq j$, $i, j \in \{1, \ldots, k\}$ and for some $\lambda' \in \mathbb{R}$,

$$\int_{\mathbb{R}^n} x(1_{A_i}(x) - 1_{A_j}(x)) d\gamma_n(x) = \lambda' \int_{\mathbb{R}^n} x(1_{B_i'}(x) - 1_{B_j'}(x)) d\gamma_n(x), \tag{4.111}$$

and such that

$$\int_{\mathbb{R}^n} x(1_{(B_i' \cup B_j')^c}) d\gamma_n(x) \in \Pi. \tag{4.112}$$

Such $\{B'_p\}_{p=1}^k$ exists by (4.108), letting $\rho > 0$ such that $\rho < (10000k)^{-8}$, so that

$$\left\|\int_{\mathbb{R}^n} x(1_{B''_i}(x) - 1_{B''_j}(x))d\gamma_n(x)\right\|_{\ell_2^n} = 3\sqrt{2}/(4\sqrt{\pi}).$$

$$\left\|\int_{\mathbb{R}^n} x(1_{(B''_i\cup B''_j)^c}(x))d\gamma_n(x)\right\|_{\ell_2^n} = \left\|\int_{\mathbb{R}^n} x(1_{B''_i}(x))\right\|_{\ell_2^n} = \sqrt{6}/(4\sqrt{\pi}).$$

So, by the triangle inequality applied to (4.108), and (4.109),

$$\left\|\int_{\mathbb{R}^n} x(1_{A_i}(x) - 1_{A_j}(x))d\gamma_n(x)\right\|_{\ell_2^n} > 3\sqrt{2}/(4\sqrt{\pi}) - 10^{-2} > 1/3. \tag{4.113}$$

$$\left\|\int_{\mathbb{R}^n} x(1_{(A_i\cup A_j)^c}(x))d\gamma_n(x)\right\|_{\ell_2^n} > \sqrt{6}/(4\sqrt{\pi}) - 10^{-2} > 1/3. \tag{4.114}$$

Specifically, we first apply a rotation to $\{B''_p\}_{p=1}^k$ such that (4.111) holds. Then, by (4.109), we then apply another rotation that fixes the x_1 axis, so that (4.112) holds. By (4.108), (4.109), (4.113) and (4.114), each of these two rotations can be chosen so that a given unit vector is moved in $\mathbb{R}^n$ a distance not more than $12(6k\rho)^{1/8}$. And since we are rotating three polygonal cones with two facets each, (4.110) holds.

Using (4.110) and the triangle inequality,

$$(\sum_{p=1}^k \|1_{A_p} - 1_{B'_p}\|_{L_2(\gamma_n)}^2)^{1/2} < 20(6k\rho)^{1/16}. \tag{4.115}$$

Also, using that $c_\ell = 0$ for $|\ell| = 1, \ell_1 = 0$, (4.111) implies that $B'_i \cap B'_j \subseteq \{x \in \mathbb{R}^n : x_1 = 0\}$, and we may assume that $B'_i \subseteq \{x \in \mathbb{R}^n : x_1 \geq 0\}$.

Let $n'_i \in \mathbb{R}^n$ denote the interior unit normal of B'_i such that n'_i is normal to $(\partial B'_i) \setminus B'_j$, and let $n'_j \in \mathbb{R}^n$ denote the interior unit normal of B'_j such that n'_j is normal to $(\partial B'_j) \setminus B'_i$. Then, define B_i, B_j such that

$$\begin{aligned} B_i &:= B'_i \cup \{x \in \mathbb{R}^n : x_1 \geq 0 \wedge \langle n'_i, x/\|x\|_2\rangle \geq -4\rho^{21/20}/\sqrt{-3\log\rho}\},\\ B_j &:= B'_j \cup \{x \in \mathbb{R}^n : x_1 \leq 0 \wedge \langle n'_j, x/\|x\|_2\rangle \geq -4\rho^{21/20}/\sqrt{-3\log\rho}\}. \end{aligned} \tag{4.116}$$

Since $B_i \cup B_j$ is symmetric with respect to reflection across $B_i \cap B_j = B'_i \cap B'_j$, equation (4.111) implies that there is a $\lambda > 0$ such that

$$\int_{\mathbb{R}^n} x(1_{A_i}(x) - 1_{A_j}(x))d\gamma_n(x) = \lambda \int_{\mathbb{R}^n} x(1_{B_i}(x) - 1_{B_j}(x))d\gamma_n(x). \tag{4.117}$$

Step 3. An estimate for small x.

Combining (4.5), (4.8), and (4.117), there exists $|b_1| < 50(6k\rho)^{1/16}$ (by Hilbert space duality, eqrefthree7.93 and (4.116)) such that

$$\rho^{-1}LT_\rho(1_{A_i}-1_{A_j})(x)-\rho^{-1}LT_\rho(1_{B_i}-1_{B_j})(x)-x_1b_1 =: \sum_{\ell\in\mathbb{N}^n\colon |\ell|\geq 2} b_\ell\,|\ell|\,\rho^{|\ell|-1}h_\ell(x)\sqrt{\ell!}. \tag{4.118}$$

Choose ρ_1 so that $0<\rho<\rho_1$ implies

$$100k^{1/16}\sum_{m=2}^{\infty} m(m+n-1)^n\rho^{m-2}m^n3^m(-\log\rho^3)^{m/2} < \rho^{-1/80}/20.$$

Recall that the number of $\ell\in\mathbb{N}^n$ such that $|\ell|=m$ is equal to $\frac{m+n-1!}{m!(n-1)!}\leq(m+n-1)^n$. Note that, $|b_\ell|<100k^{1/16}\rho^{1/16}$, for all $\ell\in\mathbb{N}^n$, $|\ell|\geq 2$, by Hilbert space duality. Let $x\in\mathbb{R}^n$ with $\|x\|_2^2\leq-\log\rho^3$. By (4.118), Lemma 4.5.1,

$$\begin{aligned}
&\left|\rho^{-1}LT_\rho(1_{A_i}-1_{A_j})(x)-\rho^{-1}LT_\rho(1_{B_i}-1_{B_j})(x)-x_1b_1\right|\\
&\quad\leq 100k^{1/8}\rho^{17/16}\sum_{\ell\in\mathbb{N}^n\colon |\ell|\geq 2}|\ell|\,\rho^{|\ell|-2}\,|h_\ell(x)|\sqrt{\ell!}\\
&\quad\leq 100k^{1/8}\rho^{17/16}\sum_{\ell\in\mathbb{N}^n\colon |\ell|\geq 2}|\ell|\,\rho^{|\ell|-2}\,|\ell|^n\,3^{|\ell|}\prod_{i=1}^n\max\{1,|x_i|^{\ell_i}\}\\
&\quad\leq 100k^{1/8}\rho^{17/16}\sum_{m=2}^{\infty}m(m+n-1)^n\rho^{m-2}m^n3^m(-\log\rho^3)^{m/2}\leq\rho^{21/20}/20.
\end{aligned} \tag{4.119}$$

From Lemma 4.4.2 and (4.8), for $x=(x_1,\ldots,x_n)$, with $B_i\cap B_j\subseteq\{x\in\mathbb{R}^n\colon x_1=0\}$,

$$x\in B_i\cup B_j\,\wedge\,\|x\|_2^2\leq-\log\rho^3\implies\rho^{-1}\mathrm{sign}(x_1)\cdot LT_\rho(1_{B_i}-1_{B_j})(x)>(1/10)\,|x_1|\,. \tag{4.120}$$

Then (4.119) and (4.120) show that

$$\begin{aligned}
&|x_1|>\rho^{21/20}\,\wedge\,\|x\|_2^2\leq-\log\rho^3\,\wedge\,x\in B_i\cup B_j\\
&\qquad\implies\rho^{-1}\mathrm{sign}(x_1)\cdot LT_\rho(1_{A_i}-1_{A_j})(x)>0.
\end{aligned} \tag{4.121}$$

By (4.101), Lemma 4.3.1, and by applying (4.121) for all $i',j'\in\{1,\ldots,k\}$, $i'\neq j'$, along with the inclusion-exclusion principle,

$$|x_1|>\rho^{21/20}\,\wedge\,\|x\|_2^2\leq-\log\rho^3\,\wedge\,x\in B_i'\cup B_j'\implies\mathrm{sign}(x_1)\cdot(1_{A_i}(x)-1_{A_j}(x))>0. \tag{4.122}$$

From (4.122),

$$\begin{aligned}
&x\in B_i'\cup B_j'\,\wedge\,|d(x,(\partial B_i')\cup(\partial B_j'))|>\rho^{21/20}\,\wedge\,\|x\|_2^2\leq-\log\rho^3\\
&\qquad\implies\mathrm{sign}(x_1)\cdot(1_{A_i}(x)-1_{A_j}(x))>0.
\end{aligned} \tag{4.123}$$

Step 4. Applying the Main Lemma.

Recall that there exists a 2-dimensional plane $\Pi \subseteq \mathbb{R}^n$ such that $0 \in \Pi$ and such that, for all $p \in \{1, \ldots, k\}$, $\int_{A_p} x d\gamma_n(x) \in \Pi$. Define

$$S := \mathrm{span}\left\{\int_{\mathbb{R}^n} (1_{B_i'}(x) - 1_{B_j'}(x)) x d\gamma_n(x), \int_{\mathbb{R}^n} (1_{B_i'}(x) - 1_{(B_i' \cup B_j')^c}(x)) x d\gamma_n(x)\right.$$
$$\left., \int_{\mathbb{R}^n} (1_{B_j'}(x) - 1_{(B_i' \cup B_j')^c}(x)) x d\gamma_n(x)\right\}.$$

Note that S is a 2-dimensional plane and $0 \in S$. By (4.111), $\int_{\mathbb{R}^n}(1_{B_i'}(x) - 1_{B_j'}(x)) d\gamma_n(x) \in \Pi$. Moreover, since $\{B_i', B_j', (B_i' \cup B_j')^c\}$ is a regular simplicial conical partition,

$$S = \mathrm{span}\left\{\int_{B_i'} x d\gamma_n(x), \int_{B_j'} x d\gamma_n(x), \int_{(B_i' \cup B_j')^c} x d\gamma_n(x)\right\}.$$

From (4.112), S and Π are 2-dimensional planes that both contain the linearly independent vectors $\int_{(B_i' \cup B_j')^c} x d\gamma_n(x)$ and $\int_{\mathbb{R}^n}(1_{B_i'}(x) - 1_{B_j'}(x)) x d\gamma_n(x)$. We therefore conclude that $S = \Pi$. In particular,

$$\int_{B_p'} x d\gamma_n(x) \in \Pi, \forall\, p \in \{i, j\}. \tag{4.124}$$

Let $\rho_0 := \min(\rho_1, 10^{-2}/6k, e^{-20(n+1)^{10^{12} n^3 (n+2)!}})$. Using (4.123), (4.111) and (4.124), we can iteratively apply Lemma 4.6.1 an infinite number of times. In particular, any time we know the conclusion (4.48), we use (4.48) in the assumption (4.47) Repeating this process infinitely many times shows that there exists a regular simplicial conical partition $\{C_i\}_{i=1}^k$ which is equal to $\{A_i\}_{i=1}^k$. □

The Main Theorem now follows from Theorem 4.7.1 and the Fundamental Theorem of Calculus.

Theorem 4.7.2 (Main Theorem). *Let $n \geq 2, k = 3$. There exists $\rho_0 = \rho_0(n, k) > 0$ such that Conjecture 5 holds for $\rho \in (0, \rho_0)$. Moreover, up to orthogonal transformation, the regular simplicial conical partition uniquely achieves the maximum of* (4.3) *in Conjecture 5.*

Proof. Choose ρ_0 via Theorem 4.7.1 and let $0 < \rho < \rho_0$. Let $\{B_i\}_{i=1}^k \subseteq \mathbb{R}^n$ be a regular simplicial conical partition. By Theorem 4.7.1 and the fact that $\Delta_k^0(\gamma_n) \subseteq \Delta_k(\gamma_n)$,

$$\psi_\rho(1_{B_1}, \ldots, 1_{B_k}) = \sup_{(f_1, \ldots, f_k) \in \Delta_k^0(\gamma_n)} \psi_\rho(f_1, \ldots, f_k). \tag{4.125}$$

Let $(f_1, \ldots, f_k) \in \Delta_k^0(\gamma_n)$. By (4.11), $\sum_{i=1}^k \int_{\mathbb{R}^n} f_i T_0 f_i d\gamma_n = k(1/k^2) = 1/k$. By the Fundamental Theorem of Calculus and (4.125),

$$\begin{aligned}
&\sum_{i=1}^{k} \int_{\mathbb{R}^n} f_i T_\rho f_i d\gamma_n = \int_0^\rho \left[\frac{d}{d\alpha} \sum_{i=1}^{k} \int_{\mathbb{R}^n} f_i T_\alpha f_i d\gamma_n\right] d\alpha + \frac{1}{k} \\
&= \int_0^\rho \psi_\alpha(f_1, \ldots, f_k) d\alpha + \frac{1}{k} \leq \int_0^\rho \psi_\alpha(1_{B_1}, \ldots, 1_{B_k}) d\alpha + \frac{1}{k} \\
&= \int_0^\rho \left[\frac{d}{d\alpha} \sum_{i=1}^{k} \int_{\mathbb{R}^n} 1_{B_i} T_\alpha 1_{B_i} d\gamma_n\right] d\alpha + \frac{1}{k} = \sum_{i=1}^{k} \int_{\mathbb{R}^n} 1_{B_i} T_\rho 1_{B_i} d\gamma_n.
\end{aligned}$$

□

By using the invariance principle of [50, Theorem 1.10,Theorem 3.6,Theorem 7.1,Theorem 7.4] which transfers results from partitions of Euclidean space to low-influence discrete functions, Theorem 4.7.2 implies a weak form of the Plurality is Stablest Conjecture. While the following result is quite far from Conjecture 2 and might not be of immediate use to complexity theory, it is included to indicate a possible application of Theorem 4.7.2. Essentially, if we modify the exact application of the invariance principle that is used in [50, Theorem 7.1], then Conjecture 2 follows. However, by avoiding [50, Theorem 7.1], we must make very restrictive assumptions on the function f in Conjecture 2. Nevertheless, [50, Theorem 7.4] shows that the class of functions f described in Corollary 4.7.3 is nonempty.

Note that the most straightforward application of Theorem 4.7.2 only gives vacuous cases of Conjecture 2, in which $0 < \rho < \rho_0(n, k)$. In particular, since Theorem 4.7.2 requires $0 < \rho < \rho_0(n, k)$, by (4.11) we must take $\varepsilon < 3k\rho$ to get a nontrivial statement in Conjecture 2. In this case, the invariance principle [50, Theorem 3.6] gives τ with $\log \tau = -C(\log(\varepsilon))^2(1/\varepsilon)$, so that τ becomes a function of ρ. Since we provide a ρ with inverse exponential dependence on n, then τ also has inverse exponential dependence on n. Thus, no function f can satisfy the assumptions of Conjecture 2 in this case. To avoid this issue, we modify Conjecture 2 as follows.

Corollary 4.7.3 (Weak Form of Plurality is Stablest)**.** *Let $\rho_0(n, k)$ be given by Theorem 4.7.2. Fix $n \geq 2$, $k = 3$, and Let $N := \log\log\log\log\log(n) \geq 1$. Let $0 < \rho < \rho_0(N, k) < 1/2$, $\varepsilon > 0$, $\tau = \tau(\varepsilon, k) > 0$. Let $f \colon \{1, \ldots, k\}^n \to \Delta_k$ with $\sum_{\sigma \in \{1,\ldots,k\}^n \colon \sigma_j \neq 0} (\widehat{f_i}(\sigma))^2 \leq \tau$ for all $i \in \{1, \ldots, k\}$, $j \in \{1, \ldots, n\}$. Assume that there exists $0 < m < N$ and $g \colon \mathbb{R}^m \to \Delta_k$ with $\int_{\mathbb{R}^m} g d\gamma_m = \frac{1}{k^n} \sum_{\sigma \in \{1,\ldots,k\}^n} f(\sigma)$, and such that*

$$\left| \int_{\mathbb{R}^n} \langle g, T_\rho g \rangle d\gamma_n - \frac{1}{k^n} \sum_{\sigma \in \{1,\ldots,k\}^n} \langle f(\sigma), T_\rho f(\sigma) \rangle \right| < \varepsilon.$$

Then part (a) of Conjecture 2 holds. From [50][Theorem 7.4], this class of f is nontrivial.

Unfortunately, the proof of Theorem 4.7.2 fails for small negative ρ, as we now show.

Theorem 4.7.4. *Fix $k = 3$, $n \geq 2$. Define $\Delta_k^0(\gamma_n)$ as in Definition 4.2.2 and define ψ_ρ as in (4.28). Let $\{B_i\}_{i=1}^k \subseteq \mathbb{R}^n$ be a regular simplicial conical partition. Then there exists $\rho_2 = \rho_2(n,k) > 0$ such that, for $\rho \in (-\rho_2, 0)$, $(1_{B_1}, \ldots, 1_{B_k})$ does not achieve the following supremum.*

$$\sup_{(f_1,\ldots,f_k)\in\Delta_k^0(\gamma_n)} \rho^{-1}\sum_{i=1}^k \int_{\mathbb{R}^n} f_i L T_\rho f_i d\gamma_n = \sup_{(f_1,\ldots,f_k)\in\Delta_k^0(\gamma_n)} \psi_\rho(f_1,\ldots,f_k).$$

Proof. Let $e_1 = (1, 0, \ldots, 0)$, $e_2 = (0, 1, 0, \ldots, 0)$. Fix $i, j \in \{1, \ldots, k\}, i \neq j$. Let $\sigma\colon \mathbb{R}^n \to \mathbb{R}^n$ denote reflection across $B_i \cap B_j$. Since $B_i = \sigma(B_j)$, by (4.30), it suffices to find $i, j \in \{1, \ldots, k\}$ and $x \in B_i$ such that $\rho^{-1} L T_\rho 1_{B_i}(x) < \rho^{-1} L T_\rho 1_{B_j}(x)$. By replacing $\{B_i\}_{i=1}^k$ with $\{\tau B_i\}_{i=1}^k$ for $\tau\colon \mathbb{R}^n \to \mathbb{R}^n$ a rotation, we may assume that $\mathrm{span}\{z_i\}_{i=1}^k = \mathrm{span}\{e_1, e_2\}$. Moreover, we may assume $B_i \cap B_j \subseteq \{x \in \mathbb{R}^n \colon x_1 = 0\}$ and $B_i \subseteq \{x \in \mathbb{R}^n \colon x_1 \geq 0\}$. Let $y := (\sqrt{3}/2)e_1 + (1/2)e_2$, $\widetilde{y} := -(1/2)e_1 + (\sqrt{3}/2)e_2$. Fix $x \in B_i$ with $\langle x, \widetilde{y}\rangle > 0$ also fixed. From (4.35) and the fact that $\rho < 0$, there exists $c = c(\langle x, \widetilde{y}\rangle) > 0$ such that

$$\left\langle x, \frac{1}{\rho}\nabla T_\rho(1_{B_i} - 1_{B_j})(x)\right\rangle = -\langle x, \widetilde{y}\rangle(c + O(e^{-\langle x, y\rangle^2/2})). \tag{4.126}$$

For $x \in \mathbb{R}^n$ with $\langle x, \widetilde{y}\rangle = 0$, we have, as in Lemma 4.5.2, and Lemma 4.4.1,

$$\begin{aligned}
&\left|\int_{\mathbb{R}^n}\left(\sum_{i=1}^n (1 - y_i^2)(1_{B_i} - 1_{B_j})(x\rho + y\sqrt{1-\rho^2})d\gamma_n(y)\right)\right| \\
&\qquad \leq 2\left|\int_{B(0,\rho\|x\|_2)} \sum_{i=1}^n (1 - y_i^2) d\gamma_n(y)\right| \leq 200(n+1)!((\rho\|x\|_2)^n + 1)e^{-\rho^2\|x\|_2^2}.
\end{aligned}$$

So, a derivative bound as in the proof of (4.33) shows

$$\begin{aligned}
&\left|\int_{\mathbb{R}^n}\left(\sum_{i=1}^n (1 - y_i^2)(1_{B_i} - 1_{B_j})(x\rho + y\sqrt{1-\rho^2})d\gamma_n(y)\right)\right| \\
&\qquad \leq \rho\langle x, \widetilde{y}\rangle 200(n+2)! + 200(n+1)!((\rho\|x\|_2)^n + 1)e^{-\rho^2\|x\|_2^2}.
\end{aligned} \tag{4.127}$$

Then,

$$\begin{aligned}
\rho^{-1}LT_\rho(1_{B_i}-1_{B_j})(x) &\overset{(4.8)}{=} \frac{1}{\rho}(\langle x, \nabla T_\rho(1_{B_i}-1_{B_j})(x)\rangle - \Delta T_\rho(1_{B_i}-1_{B_j})(x)) \\
&= \langle x, T_\rho(\nabla(1_{B_i}-1_{B_j}))(x)\rangle \\
&\quad + \frac{\rho}{1-\rho^2}\int_{\mathbb{R}^n}\left(\sum_{i=1}^{n}(1-y_i^2)(1_{B_i}-1_{B_j})(x\rho+y\sqrt{1-\rho^2})\right)d\gamma_n(y).
\end{aligned} \tag{4.128}$$

So, choose $\rho < (c/8)(200(n+2)!)^{-1}$, then choose $\langle x,y\rangle$ sufficiently large, and then combine (4.126),(4.127) and (4.128) to get

$$\rho^{-1}LT_\rho(1_{B_i}-1_{B_j})(x) < -\langle x, \widetilde{y}\rangle\frac{c}{4}.$$

□

4.8 Open Problems

There are two problems that are left open in this work. First, Conjecture 5 remains entirely open for $k \geq 4$ partition elements. Some of the results of this work hold for the case $k \geq 4$, and some do not. The first variation in Lemma 4.3.1 holds for all $k \geq 4$. Strictly speaking, the argument of Lemma 4.3.1 may not hold for $\rho < 0$ since it is not clear whether or not the functional (4.28) is convex. Also, the technical error estimate from Lemma 4.5.3 holds. One of the main issues for the case $k \geq 4$ is that Lemma 4.2.5 is no longer available. Moreover, the stability estimate in Lemma 4.2.8 would be needed for $k \geq 4$. The following conjecture summarizes the main technical issue in proving an analogue of Lemma 4.2.5 for $k = 4$, $n = 3$. If we could have a stability estimate for Conjecture 6 below, resembling the estimate of Lemma 4.2.8, then in principle the proof of the Main Lemma, Lemma 4.6.1 would go through, and therefore Theorem 4.7.2 would hold for $k \geq 4$ as well. Before we state the conjecture, recall Definition 4.2.2, (4.28) and (4.13).

Conjecture 6. *Let $k = 4$, $n = 3$. Suppose $\{A_i\}_{i=1}^k \subseteq \mathbb{R}^n$ satisfies*

$$\psi_0(1_{A_1}, \ldots, 1_{A_k}) = \sup_{(f_1,\ldots,f_k)\in\Delta_k^0(\gamma_n)} \psi_0(f_1, \ldots, f_k).$$

Then $\{A_i\}_{i=1}^k$ is a simplicial conical partition.

This result is known to be true if we replace $\Delta_k^0(\gamma_n)$ with $\Delta_k(\gamma_n)$, by [55, Lemma 3.3]. However, the volume constraint of $\Delta_k^0(\gamma_n)$ causes difficulties for the methods of [55, 57].

The second problem that remains open is Conjecture 5 for $\rho < 0$ or for ρ positive and much larger than 0. We have already discussed the issues for $\rho < 0$ in Theorem 4.7.4, where it is shown that our proof strategy surprisingly fails for $\rho < 0$. For ρ with, e.g. $\rho \in (1/2, 1)$, the error bounds that we use in the proof of Theorem 4.7.2 seem to break down, especially when we apply Lemma 4.6.1, Lemma 4.4.2, and (4.33). There is nothing special about our choice of 1/2 here, other than that it is a positive number that is sufficiently far from 0. So, it seems that our method is not applicable for $\rho \in (1/2, 1)$. For example, Lemma 4.5.3 has an error term which is estimated by Lemma 4.5.2. However, the error estimate of Lemma 4.5.2 grows exponentially in n. And to compensate for this error, we need to choose ρ to decrease exponentially in n. Even before we apply the Main Lemma 4.6.1, there is also a loss in (4.119), where we essentially need a very specific L_∞ bound on the Gaussian heat kernel (or Mehler kernel, (2.10)). We used the rather crude method of bounding each Hermite polynomial separately in Lemma 4.5.1, and then summing up these polynomials. In principle, both of these losses could be avoided with dimension independent error estimates in Lemmas 4.5.3 and Lemma 4.5.2. However, this seems to be a difficult task.

However, since the case $\rho \in (1/2, 1)$ relates to geometric multi-bubble problems, whereas the case of small ρ seems to concern entirely different geometric information, it is unclear whether or not a single method could simultaneously solve or interpolate between different values of ρ in Conjecture 5.

Acknowledgement. Thanks to Assaf Naor for guidance and encouragement, and for helpful comments concerning Corollary 4.7.3. Thanks also to Elchanan Mossel for helpful comments concerning the details of Corollary 4.7.3 and the vacuous cases of Conjecture 2. Thanks also to the anonymous reviewers' thorough reading and helpful comments.

4.9 Differentiation of the Noise Operator

We prove (4.8) and (4.9). Let $\rho \in (-1, 1)$ and let $f: \mathbb{R}^n \to \mathbb{R}$. In the following calculations, we use integration by parts freely, and we use differentiation in the distributional sense. We first calculate derivatives of $T_\rho f(x)$ with respect to $x \in \mathbb{R}^n$, and then we calculate the derivative of $T_\rho f(x)$ with respect to ρ.

$$
\begin{aligned}
\frac{\partial}{\partial x_i}T_\rho f(x) &\overset{(4.2)}{=} \int \frac{\partial}{\partial x_i}[f(x\rho + y\sqrt{1-\rho^2})]d\gamma_n(y) = \int \frac{\partial f(x\rho + y\sqrt{1-\rho^2})}{\partial z_i}d\gamma_n(y)\rho \\
&= \int \frac{\partial}{\partial y_i}[f(x\rho + y\sqrt{1-\rho^2})]d\gamma_n(y)\frac{\rho}{\sqrt{1-\rho^2}} \\
&= -\int f(x\rho + y\sqrt{1-\rho^2})\frac{\partial}{\partial y_i}[d\gamma_n(y)]\frac{\rho}{\sqrt{1-\rho^2}} \\
&= \frac{\rho}{\sqrt{1-\rho^2}}\int y_i f(x\rho + y\sqrt{1-\rho^2})d\gamma_n(y). \qquad (4.129)
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2}{\partial x_i^2}T_\rho f(x) &\overset{(4.129)}{=} \frac{\rho}{\sqrt{1-\rho^2}}\int \frac{\partial}{\partial x_i}[f(x\rho + y\sqrt{1-\rho^2})]y_i d\gamma_n(y) \\
&= \frac{\rho}{\sqrt{1-\rho^2}}\int \frac{\partial f(x\rho + y\sqrt{1-\rho^2})}{\partial z_i}y_i d\gamma_n(y)\rho \\
&= \frac{\rho^2}{1-\rho^2}\int \frac{\partial}{\partial y_i}[f(x\rho + y\sqrt{1-\rho^2})]y_i d\gamma_n(y) \\
&= -\frac{\rho^2}{1-\rho^2}\int f(x\rho + y\sqrt{1-\rho^2})\frac{\partial}{\partial y_i}[y_i d\gamma_n(y)] \\
&= -\frac{\rho^2}{1-\rho^2}\int f(x\rho + y\sqrt{1-\rho^2})(-y_i^2+1)d\gamma_n(y) \\
&= \frac{\rho^2}{1-\rho^2}\int (y_i^2-1)f(x\rho + y\sqrt{1-\rho^2})d\gamma_n(y). \qquad (4.130)
\end{aligned}
$$

$$
\begin{aligned}
&\frac{d}{d\rho}T_\rho f(x) = \frac{d}{d\rho}\int_{\mathbb{R}^n} f(x\rho + y\sqrt{1-\rho^2})d\gamma_n(y)\\
&= \int_{\mathbb{R}^n} \sum_{i=1}^n \frac{\partial f(x\rho + y\sqrt{1-\rho^2})}{\partial z_i} \left(x_i - y_i\frac{\rho}{\sqrt{1-\rho^2}}\right) d\gamma_n(y)\\
&= \int_{\mathbb{R}^n} \sum_{i=1}^n \frac{\partial}{\partial y_i}[f(x\rho + y\sqrt{1-\rho^2})] \left(x_i - y_i\frac{\rho}{\sqrt{1-\rho^2}}\right) \frac{d\gamma_n(y)}{\sqrt{1-\rho^2}}\\
&= -\int_{\mathbb{R}^n} f(x\rho + y\sqrt{1-\rho^2}) \sum_{i=1}^n \frac{\partial}{\partial y_i}\left[\left(x_i - y_i\frac{\rho}{\sqrt{1-\rho^2}}\right) \frac{d\gamma_n(y)}{\sqrt{1-\rho^2}}\right]\\
&= -\int_{\mathbb{R}^n} f(x\rho + y\sqrt{1-\rho^2}) \sum_{i=1}^n \left[\left(x_i - y_i\frac{\rho}{\sqrt{1-\rho^2}}\right)(-y_i) - \frac{\rho}{\sqrt{1-\rho^2}}\right] \frac{d\gamma_n(y)}{\sqrt{1-\rho^2}}\\
&= -\int_{\mathbb{R}^n} f(x\rho + y\sqrt{1-\rho^2}) \sum_{i=1}^n \left[(y_i^2-1)\frac{\rho}{\sqrt{1-\rho^2}} - x_i y_i\right] \frac{d\gamma_n(y)}{\sqrt{1-\rho^2}}\\
&= \frac{1}{\rho}\Bigg[\frac{\rho}{\sqrt{1-\rho^2}}\left\langle x, \int_{\mathbb{R}^n} y f(x\rho + y\sqrt{1-\rho^2})d\gamma_n(y)\right\rangle\\
&\qquad\qquad + \frac{\rho^2}{1-\rho^2}\int_{\mathbb{R}^n}\left(\sum_{i=1}^n (1-y_i^2)\right) f(x\rho + y\sqrt{1-\rho^2})d\gamma_n(y)\Bigg]\\
&\overset{(4.129)\wedge(4.130)}{=} \frac{1}{\rho}\left(\langle x, \nabla T_\rho f(x)\rangle - \Delta T_\rho f(x)\right).
\end{aligned}
$$

Chapter 5

Biased Plurality is Not Stablest

This chapter is derived from the paper [47], which is joint work with Elchanan Mossel and Joe Neeman.

5.1 Introduction

Noise stability is a natural concept which appears in the study of Gaussian processes, voting, percolation and theoretical computer science. The study of partitions of the space which are optimal with respect to noise stability may be viewed as a natural extension of isoperimetric theory; see e.g. [65, Chapter 8].

The basic case which was studied most extensively is the following: which partitions of a space into two parts (with given measures) maximize noise stability? The answer to this question follows the development of isoperimetric theory: the surface-minimizing body in $\mathbb{R}^n$ with prescribed Lebesgue measure is a ball [95, 101, 67]. In the sphere it is a cap, or geodesic ball [67] which in turn implies that in Gaussian space it is a half-space [16, 96, 18]. The answer to the noise stability question is the analog: half-spaces maximize the Gaussian noise stability among all sets of a given measure [17, 77, 29]. Using an invariance principle, this implies that the majority functions maximize noise stability among all low-influence functions on the discrete cube [78].

5.1.1 Partitions into more than two parts

A much more challenging question deals with partitions into more than two parts. For the isoperimetric question, it took more than 100 years to prove the "Double Bubble Theorem" [49] which determines the minimal area that encloses and separates two fixed volumes in $\mathbb{R}^3$. The optimal partition is given by two spheres which intersect at a $120°$ angle having a third spherical cap separating the two volumes. Such a partition is called a double bubble partition. It is further conjectured that multi-bubble partitions minimize the surface area when partitioning into k parts in $\mathbb{R}^n$ as long as $k \leq n+1$ [97, Proposition 2] [75, p. 153]. An analogous isoperimetric result was established in Gaussian space, i.e., $\mathbb{R}^n$ equipped with a standard Gaussian density. Building on the previous work on the Double Bubble Conjecture, the authors of [27] found the partition of $\mathbb{R}^n (n \geq 2)$ into three parts, each having Gaussian volume close to $\frac{1}{3}$, that minimizes the Gaussian surface area between the three volumes. Their work shows that the optimal partition is a standard simplex partition, which can be seen as the limit of the double bubble partition scaled up around one point on the intersection. The standard simplex partition is defined by taking a regular simplex with center P, so that each element of the partition is the cone of a facet of the simplex with common base point P.

The analogies with the isoperimetric problem led the authors of [50] to conjecture that standard simplices are optimal for Gaussian noise stability.

In discrete product spaces, the questions regarding noise stability are central in studying hardness of approximation and voting. The relevant question in this setting involves low-influence partitions. Using a non-linear invariance principle it was shown that the majority function maximizes noise stability [78] as conjectured both in the context of hardness of approximation [54] and in voting [51]. Since plurality is the natural generalization of majority, it was asked in [54] if plurality is the most stable low-influence partition into $k \geq 3$ parts of $[k]^n$. Furthermore it was shown in [50] that the standard simplex being optimal for Gaussian noise stability is equivalent to the fact that plurality is stablest. Therefore, the authors of [50] conjectured that plurality is stablest and demonstrated a number of applications of this result in hardness of approximation and voting.

5.1.2 Our Results

In our main results we show that *the standard simplex is not the most stable partition in Gaussian space and that plurality is not the most stable low-influence partition in discrete space* for every number of parts $k \geq 3$, for every value $\rho \neq 0$ of the noise and for every prescribed measures for the different parts as long as they are not all equal to $1/k$.

In other words, the optimal partitions for noise stability are of a different nature than the ones considered for partitions into three parts in isoperimetric theory, as in Figure 3. Thus, we now know that the extension of noise stability theory from two to three or more parts is very much different than the extension of isoperimetric theory from two to three or more parts [27]. Moreover, all existing proofs which optimize noise stability of two sets [17, 58, 50, 77, 29] must fail for more than three sets, since these proofs rely on the fact that half-spaces optimize noise stability for any measure restriction.

We note that our results do not contradict the precise conjectures stated in [50] as these are stated in the case where the measures of all partition elements are exactly $1/k$. However, the authors of [50] gave no indication that they believe there is something special about the case of equal measured partitions – rather, they made the conjectures that were needed for the applications presented in their paper.

5.2 Definition and Statement of Main Results

5.2.1 Gaussian Noise Stability

We begin by recalling the non-semigroup version of the Ornstein-Uhlenbeck operator from Definition 1.1.4.

Definition 5.2.1. Let $n \geq 1$, $n \in \mathbb{Z}$. Let $\rho \in (-1, 1)$, $x \in \mathbb{R}^n$. Let $f : \mathbb{R}^n \to [0, 1]$. The Ornstein-Uhlenbeck (or Bonami-Beckner, or noise) operator is defined by

$$T_\rho f(x) := \int_{\mathbb{R}^n} f(x\rho + y\sqrt{1 - \rho^2})d\gamma_n(y).$$

Definition 5.2.2. Let $A_1, \ldots, A_k \subseteq \mathbb{R}^n$ be measurable, $k \leq n + 1$. We say that $\{A_i\}_{i=1}^k$ is a *partition* of $\mathbb{R}^n$ if $\cup_{i=1}^k A_i = \mathbb{R}^n$, and $\gamma_n(A_i \cap A_j) = 0$ for all $i \neq j$, $i, j \in \{1, \ldots, k\}$.

Definition 5.2.3. Let $\{A_i\}_{i=1}^k$ be a partition of $\mathbb{R}^n$. We say that $\{A_i\}_{i=1}^k$ *has volumes* $(a_1, \ldots, a_k)$ if $\sum_{i=1}^k a_i = 1$, $0 \leq a_i \leq 1$ for all $i = 1, \ldots, k$, and $\gamma_n(A_i) = a_i$ for all $i = 1, \ldots, k$.

Definition 5.2.4 (Flat / Simplex Partitions). Let $\{A_i\}_{i=1}^k$ be a partition of $\mathbb{R}^n$. We say that $\{A_i\}_{i=1}^k$ is a *flat partition* if there exist $y \in \mathbb{R}^n$ and $\{y_i\}_{i=1}^k \subseteq \mathbb{R}^n \setminus \{0\}$, such that

- For all $i, j \in \{1, \ldots, k\}$, $i \neq j$, y_i is not a positive multiple of y_j, and
- for all $i \in \{1, \ldots, k\}$, $A_i = y + \{x \in \mathbb{R}^n : \langle x, y_i \rangle = \max_{j=1,\ldots,k} \langle x, y_j \rangle\}$.

If $y = 0$ we say that $\{A_i\}_{i=1}^k$ is *centered*. If $y \neq 0$, we say it is *shifted*.

A *standard simplex partition* is a flat partition where $\|y_i\|_2 = 1$ for all i and $\langle y_i, y_j \rangle = -\frac{1}{k-1}$ for all $i \neq j$, $i, j \in \{1, \ldots, k\}$. Again, we call a standard simplex partition centered if $y = 0$ and shifted if $y \neq 0$.

The main picture to keep in mind for flat partitions and standard simplex partitions is Figure 3.

Definition 5.2.5 (Gaussian Noise Stability). The Gaussian noise stability of a partition $\{A_i\}_{i=1}^k$ is given by

$$\sum_{i=1}^{k}\int_{\mathbb{R}^n} 1_{A_i}T_\rho 1_{A_i}d\gamma_n.$$

Recall the following probabilistic interpretation of noise stability. Let $X = (X_1,\ldots,X_n)$ and let $Y = (Y_1,\ldots,Y_n) \in \mathbb{R}^n$ be standard Gaussian random vectors such that $\mathbb{E}(X_iY_j) = \rho 1_{i=j}$ for all $i,j \in \{1,\ldots,n\}$. Then for any $A \subseteq \mathbb{R}^n$, $\int 1_A T_\rho 1_A d\gamma_n = \mathbb{P}((X,Y) \in A\times A)$. So, Definition 5.2.5 can be equivalently written as $\sum_{i=1}^k \mathbb{P}((X,Y) \in (A_i,A_i))$.

5.2.2 Optimal Partitions are not Flat

As mentioned earlier, the result of [27] finds the partition of $\mathbb{R}^n$ into three sets of fixed Gaussian measures a_1, a_2, a_3 with $|a_i - 1/3| < .04$ for all $i = 1,2,3$ and of minimum total Gaussian surface area. This partition is always given by three 120 degree sectors, so if the a_i are not all 1/3 then it is a shifted standard simplex partition. This, along with the standard simplex conjecture of [78] may suggest that standard simplex partitions are always optimal. In our main result we prove that this is not the case:

Theorem 5.2.6. *Let $\{A_i\}_{i=1}^{n+1} \subseteq \mathbb{R}^n$ be a shifted flat partition with volumes $a := (a_1,\ldots,a_{n+1}) \neq (1/(n+1),\ldots,1/(n+1))$. Let $0<\rho<1$. Then*

$$\sum_{i=1}^{n+1}\int_{\mathbb{R}^n} 1_{A_i}T_\rho 1_{A_i}d\gamma_n < \sup_{\substack{\{B_i\}_{i=1}^{n+1}\text{ is a partition of }\mathbb{R}^n\\ \text{with volumes }(a_1,\ldots,a_{n+1})}} \sum_{i=1}^{n+1}\int_{\mathbb{R}^n} 1_{B_i}T_\rho 1_{B_i}d\gamma_n. \tag{5.1}$$

where the supremum is taken over all partitions. Similarly for $-1<\rho<0$, we have that:

$$\sum_{i=1}^{n+1}\int_{\mathbb{R}^n} 1_{A_i}T_\rho 1_{A_i}d\gamma_n > \inf_{\substack{\{B_i\}_{i=1}^{n+1}\text{ is a partition of }\mathbb{R}^n\\ \text{with volumes }(a_1,\ldots,a_{n+1})}} \sum_{i=1}^{n+1}\int_{\mathbb{R}^n} 1_{B_i}T_\rho 1_{B_i}d\gamma_n. \tag{5.2}$$

Our results indicate in particular that there is no straightforward generalization of the any of the methods of [17, 58, 50, 77, 29] that is applicable to the study of noise stability for more than two sets. As mentioned earlier, Theorem 5.2.6 also indicates a dramatic difference between the noise stability problem with two sets (where the optimal partitions are always simplex partitions) and three or more sets. This also indicates a difference between the isoperimetric problem with three sets (where in all the known cases the partitions are simplex partitions) and the noise stability problem.

5.2.3 Discrete Noise Stability

Using the invariance principle [78], [50, Theorem 1.10], it is by now standard to deduce discrete analogues of our main theorem. For simplicity we formulate one special case for partitions into 3 parts with $0 < \rho < 1$ in Corollary 5.2.9 below, which only requires the Central Limit Theorem.

Definition 5.2.7 (Plurality). A *plurality function* $f_n : \{1,2,3\}^n \to \{1,2,3\}$ is a function satisfying $f_n(x_1,\ldots,x_n) := i$ whenever $|\{\ell \in \{1,\ldots,n\} : x_\ell = i\}| > |\{\ell \in \{1,\ldots,n\} : x_\ell = j\}|$ for all $j \neq i$, $i,j \in \{1,2,3\}$. Note that we do not prescribe any behavior in cases of ties, i.e. when $\exists\ i,j \in \{1,2,3\}$ such that $|\{\ell \in \{1,\ldots,n\} : x_\ell = i\}| = |\{\ell \in \{1,\ldots,n\} : x_\ell = j\}|$.

Since the values of f_n for ties does not concern us particularly, we will often write "plurality" or "the plurality function" to denote an arbitrary plurality function.

Definition 5.2.8 (Influences). Given a probability measure $\mathbb{P}$ on $\{1,2,3\}$, the i'th *influence* $I_i(f) \in \mathbb{R}$ of a function $f : \{1,2,3\}^n \to \{1,2,3\}$ is defined by

$$I_i(f) := \mathbb{P}[f(x_1,\ldots,x_{i-1},x_i,x_{i+1},\ldots,x_n) \neq f(x_1,\ldots,x_{i-1},y,x_{i+1},\ldots,x_n)],$$

where x_i, y are i.i.d. samples from P for all $i = 1,\ldots,n$. We say that the sequence $f_n\colon \{1,2,3\}^n \to \{1,2,3\}$ has *vanishing influences* if

$$\lim_{n\to\infty} \max_{1\leq i\leq n} I_i(f_n) = 0.$$

Note that any sequence of plurality functions has vanishing influences.

Corollary 5.2.9 (Plurality is not Stablest). *Let $\alpha,\beta \in \mathbb{R}$. Consider the space $\{1,2,3\}^n$ equipped with the probability measure $Q_n = \tilde{Q}^n$ where*

$$\tilde{Q}(1) = \frac{1}{3} + \alpha n^{-1/2}, \quad \tilde{Q}(2) = \frac{1}{3} + \beta n^{-1/2}, \quad \tilde{Q}(3) = \frac{1}{3} - (\alpha+\beta) n^{-1/2}.$$

Let $\tilde{P}$ be the distribution of correlated pairs on $\{1,2,3\}^2$ with marginals $\tilde{Q}$ where the two values agree with probability $\rho > 0$ and are otherwise independent: for all $x,y \in \{1,2,3\}$,

$$\tilde{P}(x,y) = \rho 1_{(x=y)}\tilde{Q}(x) + (1-\rho)\tilde{Q}(x)\tilde{Q}(y),$$

and let $\mathbb{P} = \tilde{P}^n$. Let $f_n : \{1,2,3\}^n \to \{1,2,3\}$ be the sequence of plurality functions and assume $(\alpha,\beta) \neq (0,0)$. Then there exist a sequence $g_n\colon \{1,2,3\}^n \to \{1,2,3\}$ of functions of vanishing influences with respect to P such that

$$\lim_{n\to\infty} \mathbb{P}[f_n = i] = \lim_{n\to\infty} \mathbb{P}[g_n = i], \quad \forall i \in \{1,2,3\}$$

and for $x,y \in \{1,2,3\}^n$ distributed according to $\mathbb{P}$,

$$\lim_{n\to\infty} \mathbb{P}[f_n(x) = f_n(y)] < \lim_{n\to\infty} \mathbb{P}[g_n(x) = g_n(y)].$$

Informally, the corollary says that plurality is not the most stable low-influence function among all functions that have the same probability of getting all outcomes, when the election is slightly biased. Here we think of n as the number of voters in an election, $\{1,2,3\}$ as possible votes between 3 candidates, and f_n, g_n as functions which input the votes and output the winner of the election. We further note that for every fixed values of α and β a sequence of functions has vanishing influences if and only if it has vanishing influences under the uniform measure.

5.2.4 Bilinear Noise Stability

Given our main results, it is natural to ask *what is* optimal partition for noise stability when partitioning into three or more parts of unequal measures. While we do not have a conjecture for what the partition is, there is a very special case where we consider two different partitions of special sizes where the optimal partitions can be found explicitly using Borell's results.

Proposition 5.2.10. *Let $\rho \in (0,1)$. Then there exist partitions $\{A_1, A_2, A_3\} \subseteq \mathbb{R}^2$ and $\{B_1, B_2, B_3\} \subseteq \mathbb{R}^2$ with volumes $a = (1/3, 1/3, 1/3)$ and $b = (1/2, 0, 1/2)$ respectively, so that*

$$\sum_{i=1}^{3} \int_{\mathbb{R}^2} 1_{A_i(a)} T_\rho 1_{B_i(b)} d\gamma_2 = \sup_{\substack{\{C_i\}_{i=1}^3 \text{ is a partition of } \mathbb{R}^2 \text{ with volumes } a \\ \{D_i\}_{i=1}^3 \text{ is a partition of } \mathbb{R}^2 \text{ with volumes } b}} \sum_{i=1}^{3} \int_{\mathbb{R}^2} 1_{C_i} T_\rho 1_{D_i} d\gamma_2. \tag{5.3}$$

Moreover the maximizing partitions (must) satisfy that A_1 and B_1 (and also A_3 and B_3) are parallel half-spaces of measures $1/3$ and $1/2$.

Note in particular that the partition $\{A_1, A_2, A_3\}$ is not a standard simplex partition. The proposition in this case follows from Borell's results as each of the terms $\int_{\mathbb{R}^2} f_1 T g_1 d\gamma_2$ and $\int_{\mathbb{R}^2} f_3 T g_3 d\gamma_2$ is individually maximized by A_1, B_1 and A_3, B_3 respectively. Moreover, given the uniqueness version of Borell's result [77, 29] these are the only partitions that achieve the maximum. We note again that in the case of partitions into two parts, even in the bilinear case of partitions $\{A_1, A_2\}$ and $\{B_1, B_2\}$ the optimal partitions are always half-space (standard simplex) partitions.

We further note that even if the measures are slightly perturbed for example by letting $\tilde{b} = (1/2 - \varepsilon, 2\varepsilon, 1/2 - \varepsilon)$ for a small $\varepsilon > 0$ then any (almost) optimal partitions $\tilde{A}, \tilde{B}$ for $a, \tilde{b}$ in (5.3) have to be close to those given by A and B in the sense that they can be matched so that $\gamma_n(\tilde{A}_i \Delta A_i)$ and $\gamma_n(\tilde{B}_i \Delta B_i)$ are small for all $1 \le i \le 3$. This follows since if A, B are almost optimal for $a, \tilde{b}$ then they are almost optimal for a, b. The robust versions of Borell's result given in [77, 29] then implies that A, B have to be close to the optimal partition.

5.2.5 Organization

In Section 5.3, we compute the first variation of the noise stability of flat partitions. This computation shows that, if a partition $\{A_i\}_{i=1}^{n+1}$ optimizes noise stability, then $T_\rho(1_{A_i}-1_{A_j})$ must be constant on $A_i\cap A_j$ for all $i,j\in\{1,\ldots,n+1\}$, $i\neq j$. We will use this statement in the contrapositive form, since in Section 5.4, we show that for many flat partitions $\{A_i\}_{i=1}^{n+1}$, there exists $i,j\in\{1,\ldots,n+1\}$ with $i\neq j$ such that $T_\rho(1_{A_i}-1_{A_j})$ is not constant on $A_i\cap A_j$. And therefore, many flat partitions $\{A_i\}_{i=1}^{n+1}$ do not optimize noise stability. The Lemmas of Sections 5.3 and 5.4 are combined in Section 5.5, where Theorem 5.2.6 is proved. Finally, Corollary 5.2.9 is shown to be a consequence of Theorem 5.2.6 in Section 5.6.

5.3 The First Variation

We compute the first variation of the noise stability, in order to understand the sets that achieve equality in (5.1) and (5.3). Lemma 5.3.2 gives a normal variation implementation of arguments from [55, Lemma 3.3], [45, Lemma 2.7]. A simpler argument works for $\rho\in(0,1)$, which we present first.

Lemma 5.3.1. *(**The First Variation**) Let $\rho\in(0,1)$ and let $\{A_i\}_{i=1}^{n+1}$ be a flat partition of $\mathbb{R}^n$ with volumes $(a_1,\ldots,a_{n+1})$. Assume that there exist $i,j\in\{1,\ldots,n+1\}$ with $i\neq j$ and there exist open balls $U_i\subseteq A_i$, $U_j\subseteq V_j$ with $\gamma_n(U_i)=\gamma_n(U_j)$ and such that, for all $u_i\in U_i, u_j\in U_j$, we have*

$$T_\rho(1_{A_i}-1_{A_j})(u_i)<T_\rho(1_{A_i}-1_{A_j})(u_j). \tag{5.4}$$

Then

$$\sum_{i=1}^{n+1}\int_{\mathbb{R}^n}1_{A_i}T_\rho 1_{A_i}d\gamma_n<\sup_{\substack{\{B_p\}_{p=1}^{n+1}\text{ is a partition of }\mathbb{R}^n\\ \text{with volumes }(a_1,\ldots,a_{n+1})}}\sum_{i=1}^{n+1}\int_{\mathbb{R}^n}1_{B_i}T_\rho 1_{B_i}d\gamma_n. \tag{5.5}$$

Proof. Let $t\in[0,1]$. Define $\{f_{p,t}\}_{p=1}^{n+1}$ such that $f_{p,t}=1_{A_p}$ for all $p\in\{1,\ldots,n+$

$1\} \setminus \{i,j\}$, $f_{i,t} = 1_{A_i} - t1_{U_i} + t1_{U_j}$, $f_{j,t} := 1_{A_j} - t1_{U_j} + t1_{U_i}$. Then

$$\begin{aligned}
&\sum_{p=1}^{n+1}\int f_{p,t}T_\rho f_{p,t}d\gamma_n - \sum_{p=1}^{n+1}\int 1_{A_p}T_\rho 1_{A_p}d\gamma_n \\
&\quad= \int (1_{A_i} - t1_{U_i} + t1_{U_j})T_\rho(1_{A_i} - t1_{U_i} + t1_{U_j})d\gamma_n \\
&\qquad+ \int (1_{A_j} - t1_{U_j} + t1_{U_i})T_\rho(1_{A_j} - t1_{U_j} + t1_{U_i})d\gamma_n \\
&\qquad- \int 1_{A_i}T_\rho 1_{A_i}d\gamma_n - \int 1_{A_j}T_\rho 1_{A_j}d\gamma_n \\
&\quad= 2t\int (1_{A_i} - 1_{A_j})T_\rho(1_{U_j} - 1_{U_i})d\gamma_n + 2t^2\int (1_{U_i} - 1_{U_j})T_\rho(1_{U_i} - 1_{U_j})d\gamma_n.
\end{aligned}$$

Therefore, noting that $f_{p,0} = 1_{A_p}$ for all $p = 1, \ldots, n+1$,

$$\frac{d}{dt}\Big|_{t=0}\sum_{p=1}^{n+1}\int f_{p,t}T_\rho f_{p,t}d\gamma_n = 2\int (1_{A_i} - 1_{A_j})T_\rho(1_{U_j} - 1_{U_i})d\gamma_n \overset{(5.4)}{>} 0. \tag{5.6}$$

Since $(f_{1,t}, \ldots, f_{n+1,t})$, satisfies, for all $0 \leq t \leq 1$, $0 \leq f_{i,t} \leq 1$, $\int_{\mathbb{R}^n} f_{i,t}d\gamma_n = a_i$ for all $1 \leq i \leq n+1$, and $\sum_{i=1}^{n+1} f_{i,t} = 1$, (5.6) implies

$$\sum_{i=1}^{n+1}\int 1_{A_i}T_\rho 1_{A_i}d\gamma_n < \sup_{\substack{(f_1,\ldots,f_{n+1})\colon 0\leq f_i\leq 1,\\ \int_{\mathbb{R}^n} f_i d\gamma_n = a_i, \forall\, 1\leq i\leq n+1, \sum_{i=1}^{n+1} f_i = 1}} \sum_{i=1}^{n+1}\int_{\mathbb{R}^n} f_i T_\rho f_i d\gamma_n. \tag{5.7}$$

Finally, the convexity argument from Lemma 4.3.1 shows that the right side of (5.7) is equal to the right side of (5.5). So, (5.7) implies (5.5). □

Unfortunately, we are unable to extend the convexity argument of Lemma 5.3.1 to the case $\rho < 0$. So, to get the conclusion of Lemma 5.3.1 for $\rho < 0$, we instead use a normal variation argument. Here the notation becomes more cumbersome, but the basic idea in the proof is the same as that of Lemma 5.3.1. In Lemma 5.3.1 we are moving mass between two sets to show that the noise stability can be improved. And we do the same thing in Lemma 5.3.2, where now the mass is being moved as defined by a vector field V.

Lemma 5.3.2 (**Normal Variation**). *Let $\rho \in (-1,0) \cup (0,1)$. Let $\{A_i\}_{i=1}^{n+1}$ be a flat partition of $\mathbb{R}^n$ with volumes $(a_1, \ldots, a_{n+1})$. Assume that there exist $i, j \in \{1, \ldots, n+1\}$ with $i \neq j$ and there exists a line segment $I \subseteq A_i \cap A_j$ such that*

$T_\rho(1_{A_i}-1_{A_j})$ *is not constant on* I. *Then*

$$\begin{aligned}
&\rho\in(0,1) \quad\Longrightarrow\\
&\qquad\sum_{i=1}^{n+1}\int_{\mathbb{R}^n}1_{A_i}T_\rho 1_{A_i}d\gamma_n<\sup_{\substack{\{B_p\}_{p=1}^{n+1}\text{ is a partition of }\mathbb{R}^n\\ \text{with volumes }(a_1,\ldots,a_{n+1})}}\sum_{i=1}^{n+1}\int_{\mathbb{R}^n}1_{B_i}T_\rho 1_{B_i}d\gamma_n.\\
&\rho\in(-1,0) \quad\Longrightarrow\\
&\qquad\sum_{i=1}^{n+1}\int_{\mathbb{R}^n}1_{A_i}T_\rho 1_{A_i}d\gamma_n>\inf_{\substack{\{B_p\}_{p=1}^{n+1}\text{ is a partition of }\mathbb{R}^n\\ \text{with volumes }(a_1,\ldots,a_{n+1})}}\sum_{i=1}^{n+1}\int_{\mathbb{R}^n}1_{B_i}T_\rho 1_{B_i}d\gamma_n.
\end{aligned}\tag{5.8}$$

Proof. Let $A\subseteq\mathbb{R}^n$ be a set with smooth boundary, and let $N\colon\partial A\to S^{n-1}$ denote the unit exterior normal to ∂A. Let $V\colon\mathbb{R}^n\to\mathbb{R}^n$ be specified later. Let $\Psi\colon\mathbb{R}^n\times(-1,1)$ such that $(d/dt)|_{t=0}\Psi(x,t)=V(x)$ for all $x\in\mathbb{R}^n$, and such that $\Psi(x,0)=x$ for all $x\in\mathbb{R}^n$. Such a Ψ exists by the fundamental theorem of ODEs. For $x,y\in\mathbb{R}^n$, define

$$G(x,y):=e^{-\|\rho x-y\|_2^2/[2(1-\rho^2)]}.$$

Define $A^{(t)}:=\Psi(A,t)$. For $f\colon\mathbb{R}^n\to\mathbb{R}$, using Definition 5.2.1, write

$$T_\rho f(x)=:\int_{\mathbb{R}^n}f(y)G(x,y)dy.\tag{5.9}$$

Let $J\Psi(y,t)$ denote $|\det(\partial\Psi_i(y,t)/\partial y_j)_{1\leq i,j\leq n}|\in\mathbb{R}$. We claim that

$$\frac{d}{dt}|_{t=0}T_\rho 1_{A^{(t)}}(x)=\int_{\partial A}G(x,y)\langle V(y),N(y)\rangle dy.\tag{5.10}$$

Indeed, using that $(d/dt)|_{t=0}\Psi(x,t)=V(x)$ and $\Psi(x,0)=x$, $J\Psi(y,0)=1$ for all y, so the chain rule and divergence theorem show that

$$\begin{aligned}
&\frac{d}{dt}|_{t=0}\int_{\mathbb{R}^n}1_{A^{(t)}}(y)G(x,y)dy\\
&\qquad=\frac{d}{dt}|_{t=0}\int_{A^{(t)}}G(x,y)dy=\frac{d}{dt}|_{t=0}\int_A G(x,\Psi(y,t))J\Psi(y,t)dy\\
&\qquad=\int_A \mathrm{div}_y(G(x,y)V(y))dy=\int_{\partial A}G(x,y)\langle V(y),N(y)\rangle dy.
\end{aligned}\tag{5.11}$$

So, using $G(x,y)\gamma_n(x) = G(y,x)\gamma_n(y)$,

$$
\begin{aligned}
\frac{d}{dt}|_{t=0}\int_{\mathbb{R}^n} 1_{A^{(t)}}T_\rho 1_{A^{(t)}}d\gamma_n &= 2\int 1_A(\frac{d}{dt}|_{t=0}T_\rho 1_{A^{(t)}})d\gamma_n \\
&\overset{(5.10)}{=} 2\int_A\int_{\partial A} G(x,y)\langle V(y),N(y)\rangle dy d\gamma_n(x) \\
&= 2\int_{\partial A}\int_A G(x,y)d\gamma_n(x)\langle V(y),N(y)\rangle dy \qquad (5.12)\\
&= 2\int_{\partial A}\int_A G(y,x)dx\langle V(y),N(y)\rangle \gamma_n(y)dy \\
&\overset{(5.9)}{=} 2\int_{\partial A} T_\rho 1_A(y)\langle V(y),N(y)\rangle\gamma_n(y)dy.
\end{aligned}
$$

Suppose $V\colon\mathbb{R}^n\to\mathbb{R}^n$ satisfies

$$\int_{\partial A}\langle V(y),N(y)\rangle\gamma_n(y)dy = 0. \tag{5.13}$$

We then claim that Ψ is volume-preserving at $t=0$. Indeed, using (5.10) for $\rho=0$,

$$\frac{d}{dt}|_{t=0}\gamma_n(A^{(t)}) = \int_{\partial A}\langle V(y),N(y)\rangle d\gamma_n(y) \overset{(5.13)}{=} 0. \tag{5.14}$$

Let N denote the exterior normal to A_i, so that $-N$ is the exterior normal to A_j. For a subset of $(\partial A_i)\cap(\partial A_j)$, we denote γ_{n-1} as the measure γ_n restricted to $(\partial A_i)\cap(\partial A_j)$. Since $T_\rho(1_{A_i}-1_{A_j})$ is continuous and not constant on $A_i\cap A_j$, there exist disjoint relatively open sets $U_1,U_2\subseteq(\partial A_i)\cap(\partial A_j)$ such that $\gamma_{n-1}(U_1)=\gamma_{n-1}(U_2)>0$, and such that

$$T_\rho(1_{A_i}-1_{A_j})(u_1) > T_\rho(1_{A_i}-1_{A_j})(u_2) \quad \forall\, u_1\in U_1,\, u_2\in U_2. \tag{5.15}$$

Let $A:=A_i$. Let $\phi_1,\phi_2\colon\mathbb{R}^n\to[0,\infty)$ be C^∞ functions such that $(\partial A_i)\cap(\partial A_j)\cap\text{support}(\phi_1)$ is disjoint from $(\partial A_i)\cap(\partial A_j)\cap\text{support}(\phi_2)$, and such that

$$\int_{U_1}\phi_1(y)\gamma_n(y)dy = \int_{U_2}\phi_2(y)\gamma_n(y)dy. \tag{5.16}$$

Define $V\colon\mathbb{R}^n\to\mathbb{R}^n$ by $V(y):=(\phi_1(y)-\phi_2(y))N$. Note that V is C^∞, and

$$\int_{\partial A}\langle V(y),N(y)\rangle\gamma_n(y)dy = \int_{U_1}\phi_1\gamma_n(y)dy - \int_{U_2}\phi_2\gamma_n(y)dy \overset{(5.16)}{=} 0.$$

So, V satisfies (5.13). We can therefore compute

$$\begin{aligned}
&\frac{d}{dt}|_{t=0}S_\rho(\{A_p^{(t)}\}_{p=1}^{n+1})\\
&=\frac{d}{dt}|_{t=0}\int 1_{A_i^{(t)}}T_\rho 1_{A_i^{(t)}}d\gamma_n+\frac{d}{dt}|_{t=0}\int 1_{A_j^{(t)}}T_\rho 1_{A_j^{(t)}}d\gamma_n\\
&\overset{(5.12)}{=}2\int_{(\partial A_i)\cap(\partial A_j)}T_\rho 1_{A_i}(y)\langle V(y),N(y)\rangle\gamma_n(y)dy\\
&\qquad-2\int_{(\partial A_i)\cap(\partial A_j)}T_\rho 1_{A_j}(y)\langle V(y),N(y)\rangle\gamma_n(y)dy\\
&=2\int_{(\partial A_i)\cap(\partial A_j)}(T_\rho 1_{A_i}(y)-T_\rho 1_{A_j}(y))\langle V(y),N(y)\rangle\gamma_n(y)dy\\
&=2\int_{(\partial A_i)\cap(\partial A_j)}(T_\rho 1_{A_i}(y)-T_\rho 1_{A_j}(y))(\phi_1(y)-\phi_2(y))\gamma_n(y)dy\overset{(5.15)\wedge(5.16)}{>}0. \qquad (5.17)
\end{aligned}$$

Finally, applying the Implicit Function Theorem as in [11, Lemma 2.4], there exists $\varepsilon>0$ such that, for every $t\in(-\varepsilon,\varepsilon)$, there exists a partition $\{\widetilde{A}_p^{(t)}\}_{p=1}^{n+1}$ of $\mathbb{R}^n$ with volumes $(a_1,\ldots,a_{n+1})$ such that

$$\frac{d}{dt}|_{t=0}\sum_{p=1}^{n+1}\int_{\mathbb{R}^n}1_{A_p^{(t)}}T_\rho 1_{A_p^{(t)}}d\gamma_n=\frac{d}{dt}|_{t=0}\sum_{p=1}^{n+1}\int_{\mathbb{R}^n}1_{\widetilde{A}_p^{(t)}}T_\rho 1_{\widetilde{A}_p^{(t)}}d\gamma_n. \qquad (5.18)$$

Then (5.17) and (5.18) imply (5.8).

To justify the use of the Implicit Function Theorem, let $U_1:=(\partial A_i)\cap(\partial A_j)\cap\text{support}(\phi_1)$, let $U_2:=(\partial A_i)\cap(\partial A_j)\cap\text{support}(\phi_2)$, and let $w\colon U_1\cup U_2\to\mathbb{R}$ satisfy $w|_{(\partial U_1)\cup(\partial U_2)}=0$ and

$$\int_{U_1\cup U_2}wd\gamma_{n-1}\neq 0. \qquad (5.19)$$

For $u\in(\partial A_i)\cap(\partial A_j)$ with $u\notin U_1\cup U_2$, set $w(u)=0$. For all $s,t\in(-1,1)$ and for all $y\in(\partial A_i)\cap(\partial A_j)$, let

$$\Phi(y,s,t):=y+tV(y)+sw(y)N(y). \qquad (5.20)$$

Let $A_i^{s,t}$ denote the interior of the region formed by $(\partial A_i)\setminus((\partial A_i)\cap(\partial A_j))$ and $\cup_{y\in(\partial A_i)\cap(\partial A_j)}\Phi(y,s,t)$. Then $A_i^{s,t}$ denotes A_i after being deformed by the normal variation defined by $\Phi(y,s,t)$. Let $h(s,t):=\gamma_n(A_i^{s,t})$. By (5.14),

$$\partial h/\partial t|_{t=0}=0. \qquad (5.21)$$

Consider the equation $h(s,t)=\gamma_n(A_i)$. Note that by (5.19),

$$\partial h/\partial s|_{s=0}\neq 0. \qquad (5.22)$$

So, by the Implicit Function Theorem there exists $\varepsilon > 0$ and there exists a smooth function $\phi\colon (-\varepsilon, \varepsilon) \to \mathbb{R}$ such that $s = \phi(t)$ and such that

$$h(\phi(t), t) = \gamma_n(A_i), \quad \forall\, t \in (-\varepsilon, \varepsilon). \tag{5.23}$$

Then, we define $\widetilde{A}_i^{(t)}$ as the interior of the region formed by $(\partial A_i) \setminus ((\partial A_i) \cap (\partial A_j))$ and $\cup_{y \in (\partial A_i) \cap (\partial A_j)} \Phi(y, \phi(t), t)$. Since $h(\phi(t), t) = \gamma_n(A_i)$, we conclude that

$$\gamma_n(\widetilde{A}_i^{(t)}) = \gamma_n(A_i), \quad \forall\, t \in (-\varepsilon, \varepsilon).$$

Also, by differentiating (5.19),

$$\phi'(0) = \left(\frac{\partial h}{\partial t}|_{t=0}\right) \cdot \left(\frac{\partial h}{\partial s}|_{s=0}\right)^{-1} \overset{(5.22)\wedge(5.21)}{=} 0. \tag{5.24}$$

Therefore, we finally have

$$\frac{d}{dt}\Phi(y, \phi(t), t) \overset{(5.20)}{=} V(y) + \phi'(0)w(y)N(y) \overset{(5.24)}{=} V(y).$$

That is, (5.18) holds, since $A_i^{(t)}$ has the same normal variation as $\widetilde{A}_i^{(t)}$ at $t = 0$.

$\square$

5.4 Analyticity of the Ornstein-Uhlenbeck Operator

Let A be a polyhedral cone, and let $\rho \in (-1, 0) \cup (0, 1)$. In this section we show, roughly speaking, that if we restrict $T_\rho 1_A$ to a line parallel to a facet of A, then we get a holomorphic function. Let $\mathbb{C} = \{x_1 + x_2\sqrt{-1}\colon x_1, x_2 \in \mathbb{R}\}$ denote the complex numbers.

The following statements establish some notation and assumptions on a partition $\{A_p\}_{p=1}^{n+1}$. Rather than repeating these statements multiple times below, we will instead repeatedly reference them as Assumption 1.

Assumption 1. (**Geometric Assumptions**) Let $\{A_p\}_{p=1}^{n+1}$ be a flat partition of $\mathbb{R}^n$ with volumes $a := (a_1, \ldots, a_{n+1})$. Fix $i, j \in \{1, \ldots, n+1\}$, $i \neq j$ and let $\Pi \subseteq \mathbb{R}^n$ be a hyperplane so that $A_i \cap A_j \subseteq \Pi$. Assume that A_i and A_j share a common facet. Let $c \in \mathbb{R}$ and let $N \in \mathbb{R}^n$ with $\|N\|_2 = 1$ so that

$$\Pi = \{x \in \mathbb{R}^n\colon \langle x, N\rangle = c\}. \tag{5.25}$$

Without loss of generality, assume that

$$A_i \subseteq \{x \in \mathbb{R}^n\colon \langle x, N\rangle \leq c\}. \tag{5.26}$$

Let $\text{int}(A_i \cap A_j)$ denote the relative interior of $A_i \cap A_j$. Note that $A_i \cap A_j$ is a nonempty $(n-1)$-dimensional set. Let $L \subseteq \Pi$ be any infinite line so that $L \cap \text{int}(A_i \cap A_j) \neq \varnothing$. Let $w \in \mathbb{R}^n$ so that $L = \{cN + tw : t \in \mathbb{R}\}$. Without loss of generality, assume $A_i \cap A_j \supseteq \{cN + tw : t \geq 1\}$.

The following is the key lemma used to prove Theorem 5.2.6.

Lemma 5.4.1. *(**Analyticity Restricted to Lines**) Let $\rho \in (-1,0) \cup (0,1)$. Suppose Assumption 1 holds. Then the function $T_\rho(1_{A_i} - 1_{A_j})$ satisfies the following two properties.*

(a) $\lim_{t\to\infty} T_\rho(1_{A_i} - 1_{A_j})(cN+tw) = \lim_{t\to-\infty} T_\rho(1_{A_i} - 1_{A_j})(cN+tw)$ if and only if $c = 0$.

(b) The function $t \mapsto T_\rho(1_{A_i} - 1_{A_j})(cN + tw)$, initially defined for $t \in \mathbb{R}$, can be extended to a holomorphic function of $t \in \mathbb{C}$.

Proof of (a). We first assume that $\rho \in (0,1)$. By Assumption 1, $A_i \cap A_j \supseteq \{cN + tw : t > 1\}$. Also, $L \cap \text{int}(A_i \cap A_j) \neq 0$, and A_i, A_j are cones, so

$$\begin{aligned}
&\lim_{t\to\infty} T_\rho(1_{A_i} - 1_{A_j})(cN + tw) = T_\rho(1_{\{z\in\mathbb{R}^n :\, \langle z,N\rangle \leq c\}} - 1_{\{z\in\mathbb{R}^n :\, \langle z,N\rangle \geq c\}})(c) \\
&\quad = \int 1_{(-\infty,c)}(c\rho + s\sqrt{1-\rho^2}\,)d\gamma_1(s) - \int 1_{(c,\infty)}(c\rho + s\sqrt{1-\rho^2}\,)d\gamma_1(s) \\
&\quad = \int_{-\infty}^{(c-c\rho)/\sqrt{1-\rho^2}} d\gamma_1(s) - \int_{(c-c\rho)/\sqrt{1-\rho^2}}^{\infty} d\gamma_1(s) = 2\,\text{sign}(c)\cdot\gamma_1([0, |c|\,\frac{1-\rho}{\sqrt{1-\rho^2}})).
\end{aligned}$$

So, $\lim_{t\to\infty} T_\rho(1_{A_i} - 1_{A_j})(cN + tw) = 0$ if and only if $c = 0$. In order to prove (a), it remains to show that $\lim_{t\to-\infty} T_\rho(cN + tw) = 0$. Since $cN + w \in \text{int}(A_i \cap A_j)$ by Assumption 1 and Definition 5.2.4, we know that $\langle y_i, w\rangle > \langle y_p, w\rangle$ and $\langle y_j, w\rangle > \langle y_p, w\rangle$ for all $p \in \{1,\ldots,n+1\} \setminus \{i,j\}$. So, $\langle y_i, -w\rangle < \langle y_p, -w\rangle$ and $\langle y_j, -w\rangle < \langle y_p, -w\rangle$ for all $p \in \{1,\ldots,n+1\} \setminus \{i,j\}$. Therefore, there exists $\delta > 0$ and $T < 0$ such that, for all $t < T$, $B(cN + tw, \delta\,|t|)$ does not intersect either A_i or A_j. Hence, there is some $\delta' > 0$ such that

$$|\limsup_{t\to-\infty} T_\rho(1_{A_i} - 1_{A_j})(cN + tw)| \leq \limsup_{t\to-\infty} \int_{B(\rho(cN+|t|w),\delta' t(1-\rho^2)^{-1/2})^c} d\gamma_n = 0. \tag{5.27}$$

The proof is therefore completed for $\rho \in (0,1)$. For $\rho \in (-1,0)$, we similarly conclude that $\lim_{t\to-\infty} T_\rho(1_{A_i} - 1_{A_j})(cN + tw) = 0$ if and only if $c = 0$, and $\lim_{t\to\infty} T_\rho(cN + tw) = 0$.

∎

Proof of (b). We prove this by induction on n. We first consider the case $n = 2$. In this case, A_p is a two-dimensional sector for $p = 1, 2, 3$. Fix $i \in \{1, 2, 3\}$. We will show that $T_\rho 1_{A_i}(cn + tw)$ is holomorphic in t. Without loss of generality, assume that one edge of A_i is parallel to the x_2 axis. Further, without loss of generality, suppose there exist $\alpha, \beta, \gamma \in \mathbb{R}$ such that $A_i = \{(x_1, x_2) \in \mathbb{R}^2 \colon x_1 \geq \alpha, x_2 \geq \beta x_1 + \gamma\}$. Then

$$\begin{aligned} T_\rho 1_{A_i}(x) &= \int_{(A_i - \rho x)/\sqrt{1-\rho^2}} d\gamma_2 \\ &= \int_{(\alpha - \rho x_1)/\sqrt{1-\rho^2}}^{\infty} \int_{(\beta s + \gamma - \rho x_2)/\sqrt{1-\rho^2}}^{\infty} e^{-r^2/2} e^{-s^2/2} dr ds / 2\pi. \end{aligned}$$

Note that the set $\{cN + tw \colon t \geq 1\} \subseteq \mathbb{R}^2$ has constant first coordinate, and the second coordinate of $\{cN + tw \colon t \geq 1\}$ is an affine function of t. So, defining a constant $\alpha' := (\alpha - \rho(cN + w)_1)/\sqrt{1 - \rho^2} \in \mathbb{R}$, we have

$$T_\rho 1_{A_i}(cN + tw) = \int_{\alpha'}^{\infty} \int_{\beta' s + \gamma' + c' t}^{\infty} e^{-r^2/2} e^{-s^2/2} dr ds / 2\pi.$$

For $s, t \in \mathbb{R}$, define

$$\phi(s, t) := \int_{\beta' s + \gamma' + c' t}^{\infty} e^{-r^2/2} dr / \sqrt{2\pi} = \frac{1}{2} - \int_0^{\beta' s + \gamma' + c' t} e^{-r^2/2} dr / \sqrt{2\pi}. \tag{5.28}$$

For fixed $s \in \mathbb{R}$, $\phi(s, t)$ is a holomorphic function in t. For $z \in \mathbb{C}$ and fixed $s \in \mathbb{R}$, define the following function, where we interpret the integral as a contour integral.

$$g(z) := \int_0^{\beta' s + \gamma' + c' z} e^{-r^2/2} dr / \sqrt{2\pi}. \tag{5.29}$$

Since $e^{-r^2/2}$ is an entire function of $r \in \mathbb{C}$ with no poles, the choice of the path in the definition of g does not matter. In particular, for $z = (x_1, x_2) \in \mathbb{C}$, if we choose the path that first moves from 0 to $(\beta' s + \gamma' + c' x_1, 0)$ along a straight line, and then from $(\beta' s + \gamma' + c' x_1, 0)$ to $(\beta' s + \gamma' + c' x_1, c' x_2)$ along a straight line, we get from (5.29) the bound

$$\begin{aligned} |g(z)| &\leq \int_0^\infty e^{-r^2/2} \frac{dr}{\sqrt{2\pi}} + \left| \int_0^{c' x_2} e^{-(\beta' s + \gamma' + c' x_1 + ir)^2/2} dr \right| \\ &\leq \frac{1}{2} + \int_0^{c' x_2} e^{r^2/2} dr \leq \frac{1}{2} + c' \left| x_2 \right| e^{c' x_2^2/2}. \end{aligned} \tag{5.30}$$

In the penultimate estimate, we used $|e^{-(x_1 + i x_2)^2}| \leq |e^{x_2^2}|$. Combining (5.30) and (5.28),

$$|\phi(s, (x_1, x_2))| \leq 1 + c' \left| x_2 \right| e^{c' x_2^2/2}. \tag{5.31}$$

Consider the function $f_\varepsilon(z) := \int_\alpha^{1/\varepsilon} \phi(s,z)e^{-s^2/2}ds$, $z \in \mathbb{C}$. This function is an entire function of $z \in \mathbb{C}$, by e.g. [94, Theorem 5.4]. To prove that $f(z) := T_\rho 1_{A_i}(cN+zw)$ is holomorphic, It now suffices to show that f_ε converges uniformly to f on any compact subset of $\mathbb{C}$, as $\varepsilon \to 0$. So, let $z = (x_1, x_2) \in \mathbb{C}$. From (5.31),

$$\begin{aligned} |f_\varepsilon(z) - f(z)| &\leq \int_{1/\varepsilon}^\infty |\phi(s,z)|\, e^{-s^2/2}ds \\ &\leq \int_{1/\varepsilon}^\infty (1 + c'\,|x_2|\, e^{c'x_2^2/2})e^{-s^2/2}ds \leq (1 + c'\,|x_2|\, e^{c'x_2^2/2})e^{-1/(2\varepsilon^2)}. \end{aligned} \tag{5.32}$$

So, (5.32) gives the desired uniform convergence property, completing the proof in the case $n = 2$. Note that, by (5.31)

$$|f(z)| \leq \int_\alpha^\infty |\phi(s,z)|\, e^{-s^2/2}ds \leq 1 + c'\,|x_2|\, e^{c'x_2^2/2} \tag{5.33}$$

For the more general case, we induct on n. We are given an n-dimensional polyhedral cone Σ_n with $n+1$ facets, and we write Σ_n as a union of translates of an $(n-1)$-dimensional simplicial cone Σ_{n-1}, which can be taken to be one of the facets of Σ_n. That is, there exists $u \in \mathbb{R}^n$ such that $\Sigma_n = \cup_{s>\alpha}(\Sigma_{n-1} + su)$. So,

$$T_\rho 1_{\Sigma_n}(cN + tw) = \int_{\alpha'}^\infty \int_{\Sigma_{n-1}+wt} d\gamma_{n-1} e^{-s^2/2}ds/\sqrt{2\pi} =: \int_{\alpha'}^\infty \psi(s,t)e^{-s^2/2}ds/\sqrt{2\pi}.$$

Now, $\psi(s,t)$ is known, by the inductive hypothesis, to be holomorphic in $t = (x_1, x_2)$, for each s. Moreover, by the inductive hypothesis, $|\psi(s,(x_1,x_2))| \leq 1 + c'\,|x_2|\, e^{c'x_2^2/2}$. We therefore truncate and make a tail estimate as in (5.32) to prove uniform convergence of $f_\varepsilon(z) := \int_{1/\varepsilon}^\infty \Psi(s,z)e^{-s^2/2}ds$ to $f(z) := T_\rho 1_{A_i}(cn + zw)$. Finally, showing the following bound completes the inductive step.

$$|f(z)| \leq \int_\alpha^\infty |\Psi(s,z)|\, e^{-s^2/2}ds/\sqrt{2\pi} \leq 1 + c'\,|x_2|\, e^{c'x_2^2/2}.$$

□

5.5 The Main Theorem

Corollary 5.5.2 below implies the Main Theorem 5.2.6.

Corollary 5.5.1. ***(An Optimal Flat Partition Must be Centered)***

(1) Let $\rho \in (0,1)$. Suppose $\{A_i\}_{i=1}^{n+1}$ is a partition of $\mathbb{R}^n$ satisfying

$$\sum_{i=1}^{n+1} \int_{\mathbb{R}^n} 1_{A_i} T_\rho 1_{A_i} d\gamma_n = \sup_{\substack{\{B_i\}_{i=1}^{n+1} \text{ is a partition of } \mathbb{R}^n \\ \text{with volumes } (a_1,\ldots,a_{n+1})}} \sum_{i=1}^{n+1} \int_{\mathbb{R}^n} 1_{B_i} T_\rho 1_{B_i} d\gamma_n.$$

Suppose Assumption 1 holds. Then $y = 0$ in Assumption 1 (and in Definition 5.2.4).

(2) Let $\rho \in (-1, 0)$. Suppose $\{A_i\}_{i=1}^{n+1}$ is a partition of $\mathbb{R}^n$ satisfying

$$\sum_{i=1}^{n+1} \int_{\mathbb{R}^n} 1_{A_i} T_\rho 1_{A_i} d\gamma_n = \inf_{\substack{\{B_i\}_{i=1}^{n+1} \text{ is a partition of } \mathbb{R}^n \\ \text{with volumes } (a_1, \ldots, a_{n+1})}} \sum_{i=1}^{n+1} \int_{\mathbb{R}^n} 1_{B_i} T_\rho 1_{B_i} d\gamma_n.$$

Suppose Assumption 1 holds. Then $y = 0$ in Assumption 1 (and in Definition 5.2.4).

Proof. Fix i and j in $\{1, \ldots, n+1\}$, and recall the definition of c from Assumption 1. It suffices to show that if $c \neq 0$ then the partition $\{A_i\}_{i=1}^{n+1}$ is not optimal. Suppose, therefore, that $c \neq 0$. Then the function $T_\rho(1_{A_i} - 1_{A_j})$ restricted to any line segment emanating from the origin is holomorphic and not constant by Lemma 5.4.1. That is, $T_\rho(1_{A_i} - 1_{A_j})$ is not constant on some line segment contained in $A_i \cap A_j$. We then apply Lemma 5.3.2 to see that the partition is not optimal. □

Corollary 5.5.2. *(**Biased Optimizers are not Flat**) Let $\{A_i\}_{i=1}^{n+1}$ be a partition of $\mathbb{R}^n$ with volumes $a = (a_1, \ldots, a_{n+1})$.*

(1) Let $\rho \in (0, 1)$. Suppose $\{A_i\}_{i=1}^{n+1}$ is a partition of $\mathbb{R}^n$ satisfying

$$\sum_{i=1}^{n+1} \int_{\mathbb{R}^n} 1_{A_i} T_\rho 1_{A_i} d\gamma_n = \sup_{\substack{\{B_i\}_{i=1}^{n+1} \text{ is a partition of } \mathbb{R}^n \\ \text{with volumes } (a_1, \ldots, a_{n+1})}} \sum_{i=1}^{n+1} \int_{\mathbb{R}^n} 1_{B_i} T_\rho 1_{B_i} d\gamma_n.$$

If Assumption 1 also holds, then we must have $a = (1/(n+1), \ldots, 1/(n+1))$.

(2) Let $\rho \in (-1, 0)$. Suppose $\{A_i\}_{i=1}^{n+1}$ is a partition of $\mathbb{R}^n$ satisfying

$$\sum_{i=1}^{n+1} \int_{\mathbb{R}^n} 1_{A_i} T_\rho 1_{A_i} d\gamma_n = \inf_{\substack{\{B_i\}_{i=1}^{n+1} \text{ is a partition of } \mathbb{R}^n \\ \text{with volumes } (a_1, \ldots, a_{n+1})}} \sum_{i=1}^{n+1} \int_{\mathbb{R}^n} 1_{B_i} T_\rho 1_{B_i} d\gamma_n.$$

If Assumption 1 also holds, then we must have $a = (1/(n+1), \ldots, 1/(n+1))$.

Proof. Suppose that $\{A_p\}_{p=1}^{n+1}$ is a flat, optimal partition; we will show that it must be unbiased. First, Corollary 5.5.1 implies that $0 \in A_p$ for all $p \in \{1, \ldots, n+1\}$. Specifically, every facet which is shared by exactly two partition elements must be contained in a hyperplane that passes through the origin. This condition implies that $0 \in A_p$ for all $p \in \{1, \ldots, n+1\}$. Then, by Lemma 5.3.2 and (5.27), we have $T_\rho(1_{A_i} - 1_{A_j})(0) = 0$ for all $i, j \in \{1, \ldots, n+1\}$ such that A_i and A_j share a facet. Since each A_p is a cone, its scale-invariance and the definition of T_ρ imply that $\gamma_n(A_p) = (T_\rho 1_{A_p})(0)$; hence, $\gamma_n(A_i) = \gamma_n(A_j)$ for all $i, j \in \{1, \ldots, n+1\}$ such A_i, A_j share a facet. By Definition 5.2.4, since $0 \in A_p$ for all $p \in \{1, \ldots, n+1\}$, there is a single equivalence class associated to the following relation: $i \sim j$ if and only if there exists a sequence $(i_1, \ldots, i_r)$ with $i_1 = i$, $i_r = j$ and A_{i_ℓ} and $A_{i_{\ell+1}}$ share a facet, for all $\ell = 1, \ldots, r-1$. Therefore, $\gamma_n(A_i) = \gamma_n(A_j)$ for all $i, j \in \{1, \ldots, n\}$. □

5.6 Voting Interpretation

As mentioned earlier our results imply that, generically speaking, plurality is not the most noise stable discrete function. Gaussian space may be obtained as a limit of discrete spaces in multiple ways, each resulting in a somewhat different statement of Plurality is not stablest, via Corollary 5.5.2. For concreteness we prove Corollary 5.2.9.

Proof of Cor. 5.2.9. The proof follows by the Central Limit Theorem. Define functions $f_n\colon \{1,2,3\}^n \to \{1,2,3\}$ as the sequence of plurality functions from Definition 5.2.7 and let $(\alpha, \beta) \neq (0,0)$. In order to prove Corollary 5.2.9 we first claim that there exists a shifted flat partition $A = \{A_1, A_2, A_3\}$ of $\mathbb{R}^2$ such that

$$\lim_{n\to\infty} \mathbb{P}[f_n = i] = \gamma_2(A_i), \forall\, i \in \{1,2,3\},$$

and moreover

$$\lim_{n\to\infty} \mathbb{P}[f_n(x) = f_n(y)] = \sum_{i=1}^{3} \mathbb{P}[X \in A_i, Y \in A_i].$$

This follows from the central limit theorem, noting that the vector $(X_i^n)_{i=1}^3 \in \mathbb{R}^3$ given by

$$X_i^n(\omega_1, \ldots, \omega_n) := n^{-1/2} \sum_{j=1}^{n} (1(\omega_j = i) - 1/3),\ \ i = 1,2,3$$

converges to a centered normal vector $(N_i)_{i=1}^3 \in \mathbb{R}^3$ with variances $2/9$ and covariances $-2/3$ as $n \to \infty$. Moreover the partition $\{(X_i^n)_{i=1}^3\colon f_n = 1\}, \{(X_i^n)_{i=1}^3\colon f_n = 2\}, \{(X_i^n)_{i=1}^3\colon f_n = 3\}$ converges in the weak $L_1(\gamma_2)$ norm to the sets given by

$$A_i = \{x \in \mathbb{R}^2 : i = \mathrm{argmax}_{j=1,2,3}(N_1(x) + \alpha, N_2(x) + \beta, N_3(x) - \alpha - \beta)\}.$$

In other words, $\{A_1, A_2, A_3\}$ is a shifted standard simplex.

Let $\{B_1, B_2, B_3\}$ be a partition of $\mathbb{R}^2$ which satisfies $\gamma_2(A_i) = \gamma_2(B_i)$ for all $i = 1, 2, 3$ and

$$\sum_{i=1}^{3} \mathbb{P}[X \in A_i, Y \in A_i] < \sum_{i=1}^{3} \mathbb{P}[X \in B_i, Y \in B_i].$$

Note that $\{B_1, B_2, B_3\}$ exists by Theorem 5.2.6. By approximating $\{B_1, B_2, B_3\}$ by a finite number of axis-parallel rectangles, we may assume that $\{B_1, B_2, B_3\}$ consists of a finite union of axis-parallel rectangles.

Consider now $\tilde{g}_n \colon \{1, 2, 3\}^n \to \{1, 2, 3\}$ satisfying $\tilde{g}_n = i$ if $(X_1^n, X_2^n, X_3^n) \in B_i$. Since $\{B_1, B_2, B_3\}$ consists of a finite number of axis-parallel rectangles, Definition 5.2.8 shows that $\max_{i=1,\ldots,n} I_i \tilde{g}_n = O(n^{-1/2})$. Moreover by the Central Limit Theorem, we see that

$$|\mathbb{P}[\tilde{g}_n = i] - \gamma_2(B_i)| = O(n^{-1/2}), \quad |\mathbb{P}[\tilde{f}_n = i] - \gamma_2(A_i)| = O(n^{-1/2}).$$

and

$$\lim_{n\to\infty} \mathbb{P}[\tilde{g}_n(x) = \tilde{g}_n(y)] = \sum_{i=1}^{3} \mathbb{P}[X \in B_i, Y \in B_i] > \sum_{i=1}^{3} \mathbb{P}[X \in A_i, Y \in A_i]$$

as needed.

□

Acknowledgement. Thanks to Ronen Eldan, Kostya Makarychev, Assaf Naor, and Oded Regev for helpful comments. This work was mostly conducted during the semester on “Real Analysis in Computer Science” at the Simons Institute for Theoretical Computer Science. We are grateful to the institute for its hospitality and support.

Bibliography

[1] F. J. Almgren, Jr. Existence and regularity almost everywhere of solutions to elliptic variational problems with constraints. *Mem. Amer. Math. Soc.*, 4(165):viii+199, 1976.

[2] N. Alon and A. Naor. Approximating the cut-norm via Grothendieck's inequality. *SIAM J. Comput.*, 35(4):787–803 (electronic), 2006.

[3] K. Appel and W. Haken. Every planar map is four colorable. I. Discharging. *Illinois J. Math.*, 21(3):429–490, 1977.

[4] K. Appel, W. Haken, and J. Koch. Every planar map is four colorable. II. Reducibility. *Illinois J. Math.*, 21(3):491–567, 1977.

[5] S. Arora, B. Barak, and D. Steurer. Subexponential algorithms for unique games and related problems. In *2010 IEEE 51st Annual Symposium on Foundations of Computer Science FOCS 2010*, pages 563–572. IEEE Computer Soc., Los Alamitos, CA, 2010.

[6] S. Arora, S. Rao, and U. Vazirani. Expander flows, geometric embeddings and graph partitioning. *J. ACM*, 56(2):Art. 5, 37, 2009.

[7] P. Austrin. Towards sharp inapproximability for any 2-CSP. *SIAM J. Comput.*, 39(6):2430–2463, 2010.

[8] K. Ball. Cube slicing in $\mathbf{R}^n$. *Proc. Amer. Math. Soc.*, 97(3):465–473, 1986.

[9] K. Ball. Mahler's conjecture and wavelets. *Discrete Comput. Geom.*, 13(3-4):271–277, 1995.

[10] A. S. Bandeira, C. Kennedy, and A. Singer. Approximating the little Grothendieck problem over the orthogonal group. Preprint, arXiv:1308.5207, 2013.

[11] J. L. Barbosa and M. do Carmo. Stability of hypersurfaces with constant mean curvature. *Math. Z.*, 185(3):339–353, 1984.

[12] F. Barthe and A. Koldobsky. Extremal slabs in the cube and the Laplace transform. *Adv. Math.*, 174(1):89–114, 2003.

[13] F. Barthe and A. Naor. Hyperplane projections of the unit ball of l_p^n. *Discrete Comput. Geom.*, 27(2):215–226, 2002.

[14] M. Bellare, O. Goldreich, and M. Sudan. Free bits, PCPs, and nonapproximability—towards tight results. *SIAM J. Comput.*, 27(3):804–915 (electronic), 1998.

[15] V. Bittorf, B. Recht, C. Re, and J. A. Tropp. Factoring nonnegative matrices with linear programs. Preprint, arXiv:1206.1270, 2012.

[16] C. Borell. The Brunn-Minkowski inequality in Gauss space. *Invent. Math.*, 30(2):207–216, 1975.

[17] C. Borell. Geometric bounds on the Ornstein-Uhlenbeck velocity process. *Z. Wahrsch. Verw. Gebiete*, 70(1):1–13, 1985.

[18] C. Borell. The Ehrhard inequality. *C. R. Math. Acad. Sci. Paris*, 337(10):663–666, 2003.

[19] M. Braverman, K. Makarychev, Y. Makarychev, and A. Naor. The Grothendieck constant is strictly smaller than Krivine's bound. *Forum Math. Pi*, 1:e4, 42, 2013.

[20] J. E. Brock. Centroid and inertia tensor of a spherical triangle. *National Technical Information Service, Naval Postgraduate School, Monterey, California*, 1974.

[21] A. Burchard and M. Schmuckenschlger. Comparison theorems for exit times. *Geometric & Functional Analysis GAFA*, 11(4):651–692, 2001.

[22] S. Chatterjee. A generalization of the Lindeberg principle. *Ann. Probab.*, 34(6):2061–2076, 2006.

[23] T. H. Colding, T. Ilmanen, W. P. Minicozzi, II, and B. White. The round sphere minimizes entropy among closed self-shrinkers. *J. Differential Geom.*, 95(1):53–69, 2013.

[24] T. H. Colding and W. P. Minicozzi, II. Minimal surfaces and mean curvature flow. In *Surveys in geometric analysis and relativity*, volume 20 of *Adv. Lect. Math. (ALM)*, pages 73–143. Int. Press, Somerville, MA, 2011.

[25] T. H. Colding and W. P. Minicozzi, II. Generic mean curvature flow I: generic singularities. *Ann. of Math. (2)*, 175(2):755–833, 2012.

[26] R. Connelly. Rigidity and energy. *Invent. Math.*, 66(1):11–33, 1982.

[27] J. Corneli, I. Corwin, S. Hurder, V. Sesum, Y. Xu, E. Adams, D. Davis, M. Lee, R. Visocchi, and N. Hoffman. Double bubbles in Gauss space and spheres. *Houston J. Math.*, 34(1):181–204, 2008.

[28] Q. Du, V. Faber, and M. Gunzburger. Centroidal Voronoi tessellations: applications and algorithms. *SIAM Rev.*, 41(4):637–676, 1999.

[29] R. Eldan. A two-sided estimate for the Gaussian noise stability deficit. Preprint, arXiv:1307.2781, 2013.

[30] T. Figiel, J. Lindenstrauss, and V. D. Milman. The dimension of almost spherical sections of convex bodies. *Acta Math.*, 139(1-2):53–94, 1977.

[31] A. Frieze and M. Jerrum. Improved approximation algorithms for MAX k-CUT and MAX BISECTION. In *Integer programming and combinatorial optimization (Copenhagen, 1995)*, volume 920 of *Lecture Notes in Comput. Sci.*, pages 1–13. Springer, Berlin, 1995.

[32] D. Gabai, G. R. Meyerhoff, and N. Thurston. Homotopy hyperbolic 3-manifolds are hyperbolic. *Ann. of Math. (2)*, 157(2):335–431, 2003.

[33] D. Gabai, R. Meyerhoff, and P. Milley. Minimum volume cusped hyperbolic three-manifolds. *J. Amer. Math. Soc.*, 22(4):1157–1215, 2009.

[34] R. J. Gardner. The Brunn-Minkowski inequality. *Bull. Amer. Math. Soc. (N.S.)*, 39(3):355–405, 2002.

[35] M. X. Goemans and D. P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *J. Assoc. Comput. Mach.*, 42(6):1115–1145, 1995.

[36] D. Goldman. Uniqueness results for critical points of a non-local isoperimetric problem via curve shortening. Preprint, arXiv:1206.5984, 2012.

[37] L. Gross. Logarithmic Sobolev inequalities. *Amer. J. Math.*, 97(4):1061–1083, 1975.

[38] A. Grothendieck. Résumé de la théorie métrique des produits tensoriels topologiques. *Bol. Soc. Mat. São Paulo*, 8:1–79, 1953.

[39] U. Haagerup. The best constants in the Khintchine inequality. *Studia Math.*, 70(3):231–283 (1982), 1981.

[40] T. C. Hales. A proof of the Kepler conjecture. *Ann. of Math. (2)*, 162(3):1065–1185, 2005.

[41] R. R. Hall. On a conjecture of Littlewood. *Math. Proc. Cambridge Philos. Soc.*, 78(3):443–445, 1975.

[42] E. Hansen. *Global optimization using interval analysis*, volume 165 of *Monographs and Textbooks in Pure and Applied Mathematics*. Marcel Dekker Inc., New York, 1992.

[43] J. Hass and R. Schlafly. Double bubbles minimize. *Ann. of Math. (2)*, 151(2):459–515, 2000.

[44] J. Håstad. Some optimal inapproximability results. *J. ACM*, 48(4):798–859, 2001.

[45] S. Heilman. Euclidean partitions optimizing noise stability. Submitted, arXiv:1211.7138v1, 2012.

[46] S. Heilman, A. Jagannath, and A. Naor. Solution of the propeller conjecture in $\mathbb{R}^3$. *Discrete & Computational Geometry*, 50(2):263–305, 2013.

[47] S. Heilman, E. Mossel, and J. Neeman. Standard simplices and pluralities are not the most noise stable. Preprint, arXiv:1403.0885, 2014.

[48] D. Hensley. Slicing the cube in $\mathbf{R}^n$ and probability (bounds for the measure of a central cube slice in $\mathbf{R}^n$ by probability methods). *Proc. Amer. Math. Soc.*, 73(1):95–100, 1979.

[49] M. Hutchings, F. Morgan, M. Ritoré, and A. Ros. Proof of the double bubble conjecture. *Ann. of Math. (2)*, 155(2):459–489, 2002.

[50] M. Isaksson and E. Mossel. Maximally stable Gaussian partitions with discrete applications. *Israel J. Math.*, 189:347–396, 2012.

[51] G. Kalai. A Fourier-theoretic perspective on the Condorcet paradox and Arrow's theorem. *Adv. in Appl. Math.*, 29(3):412–426, 2002.

[52] S. Khot. On the power of unique 2-prover 1-round games. In *Proceedings of the Thirty-Fourth Annual ACM Symposium on Theory of Computing*, pages 767–775 (electronic), New York, 2002. ACM.

[53] S. Khot. On the unique games conjecture (invited survey). In *Proceedings of the 25th Annual IEEE Conference on Computational Complexity*, pages 99–121, 2010.

[54] S. Khot, G. Kindler, E. Mossel, and R. O'Donnell. Optimal inapproximability results for MAX-CUT and other 2-variable CSPs? *SIAM J. Comput.*, 37(1):319–357, 2007.

[55] S. Khot and A. Naor. Approximate kernel clustering. *Mathematika*, 55(1-2):129–165, 2009.

[56] S. Khot and A. Naor. Grothendieck-type inequalities in combinatorial optimization. *Comm. Pure Appl. Math.*, 65(7):992–1035, 2012.

[57] S. Khot and A. Naor. Sharp kernel clustering algorithms and their associated grothendieck inequalities. *Random Structures & Algorithms*, 42(3):269–300, 2013.

[58] G. Kindler and R. O'Donnell. Gaussian noise sensitivity and Fourier tails. In *2012 IEEE 27th Conference on Computational Complexity—CCC 2012*, pages 137–147. IEEE Computer Soc., Los Alamitos, CA, 2012.

[59] B. Klartag. Rate of convergence of geometric symmetrizations. *Geom. Funct. Anal.*, 14(6):1322–1338, 2004.

[60] A. Koldobsky. *Fourier analysis in convex geometry*, volume 116 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2005.

[61] H. König. On an extremal problem originating in questions of unconditional convergence. In *Recent progress in multivariate approximation (Witten-Bommerholz, 2000)*, volume 137 of *Internat. Ser. Numer. Math.*, pages 185–192. Birkhäuser, Basel, 2001.

[62] H. König and A. Koldobsky. Volumes of low-dimensional slabs and sections in the cube. To appear in Advances of Appl. Math., 2011.

[63] J.-L. Krivine. Sur la constante de Grothendieck. *C. R. Acad. Sci. Paris Sér. A-B*, 284(8):A445–A446, 1977.

[64] R. Latała and K. Oleszkiewicz. On the best constant in the Khinchin-Kahane inequality. *Studia Math.*, 109(1):101–104, 1994.

[65] M. Ledoux. Isoperimetry and Gaussian analysis. In *Lectures on probability theory and statistics (Saint-Flour, 1994)*, volume 1648 of *Lecture Notes in Math.*, pages 165–294. Springer, Berlin, 1996.

[66] J. R. Lee, S. Oveis Gharan, and L. Trevisan. Multi-way spectral partitioning and higher-order Cheeger inequalities. In *STOC'12—Proceedings of the 2012 ACM Symposium on Theory of Computing*, pages 1117–1130. ACM, New York, 2012.

[67] P. Lévy. *Problèmes concrets d'analyse fonctionnelle. Avec un complément sur les fonctionnelles analytiques par F. Pellegrino*. Gauthier-Villars, Paris, 1951. 2d ed.

[68] J. Lindenstrauss and A. Pełczyński. Absolutely summing operators in L_p-spaces and their applications. *Studia Math.*, 29:275–326, 1968.

[69] N. Linial, L. Lovász, and A. Wigderson. Rubber bands, convex embeddings and graph connectivity. *Combinatorica*, 8(1):91–102, 1988.

[70] S. P. Lloyd. Least squares quantization in PCM. *IEEE Trans. Inform. Theory*, 28(2):129–137, 1982.

[71] A. Louis and K. Makarychev. Approximation algorithm for sparsest k-partitioning. Preprint, arXiv:1306.4384, 2014.

[72] J. MacQueen. Some methods for classification and analysis of multivariate observations. In *Proc. Fifth Berkeley Sympos. Math. Statist. and Probability (Berkeley, Calif., 1965/66)*, pages Vol. I: Statistics, pp. 281–297. Univ. California Press, Berkeley, Calif., 1967.

[73] V. D. Milman and G. Schechtman. *Asymptotic theory of finite-dimensional normed spaces*, volume 1200 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1986. With an appendix by M. Gromov.

[74] G. M. Minchin. *A treatise on statics, containing some of the fundamental propositions in electrostatics*. Longmans, Green & Co., London, 1877.

[75] F. Morgan. *Geometric measure theory*. Elsevier/Academic Press, Amsterdam, fourth edition, 2009. A beginner's guide.

[76] J. Moser. A new proof of De Giorgi's theorem concerning the regularity problem for elliptic differential equations. *Comm. Pure Appl. Math.*, 13:457–468, 1960.

[77] E. Mossel and J. Neeman. Robust optimality of Gaussian noise stability. Preprint, arXiv:1210.4126, 2012.

[78] E. Mossel, R. O'Donnell, and K. Oleszkiewicz. Noise stability of functions with low influences: invariance and optimality. *Ann. of Math. (2)*, 171(1):295–341, 2010.

[79] A. Naor and O. Regev. Krivine schemes are optimal. Preprint, arXiv:1205.6415v1, 2012.

[80] A. Naor, O. Regev, and T. Vidick. Efficient rounding for the noncommutative Grothendieck inequality. To appear, Proceedings of the 2013 ACM Symposium on Theory of Computing, arXiv:1205.6415v1.

[81] J. Nash. C^1 isometric imbeddings. *Ann. of Math. (2)*, 60:383–396, 1954.

[82] F. L. Nazarov and A. N. Podkorytov. Ball, Haagerup, and distribution functions. In *Complex analysis, operators, and related topics*, volume 113 of *Oper. Theory Adv. Appl.*, pages 247–267. Birkhäuser, Basel, 2000.

[83] K. Oleszkiewicz and A. Pełczyński. Polydisc slicing in $\mathbf{C}^n$. *Studia Math.*, 142(3):281–294, 2000.

[84] R. S. Palais. The principle of symmetric criticality. *Comm. Math. Phys.*, 69(1):19–30, 1979.

[85] C. H. Papadimitriou and M. Yannakakis. Optimization, approximation, and complexity classes. *J. Comput. System Sci.*, 43(3):425–440, 1991.

[86] G. Pisier. *The volume of convex bodies and Banach space geometry*, volume 94 of *Cambridge Tracts in Mathematics.* Cambridge University Press, Cambridge, 1989.

[87] G. Pisier. Grothendieck's theorem, past and present. *Bull. Amer. Math. Soc. (N.S.)*, 49(2):237–323, 2012.

[88] P. Raghavendra and D. Steurer. Towards computing the Grothendieck constant. In *Proceedings of the Twentieth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 525–534, Philadelphia, PA, 2009. SIAM.

[89] B. W. Reichardt. Proof of the double bubble conjecture in $\mathbf{R^n}$. *J. Geom. Anal.*, 18(1):172–191, 2008.

[90] R. E. Rietz. A proof of the Grothendieck inequality. *Israel J. Math.*, 19:271–276, 1974.

[91] V. I. Rotar′. Limit theorems for polylinear forms. *J. Multivariate Anal.*, 9(4):511–530, 1979.

[92] L. Song, A. Smola, A. Gretton, and K. A. Borgwardt. A dependence maximization view of clustering. In *Proceedings of the 24th International Conference on Machine Learning*, pages 815 – 822, 2007.

[93] E. M. Stein. *Singular integrals and differentiability properties of functions.* Princeton Mathematical Series, No. 30. Princeton University Press, Princeton, N.J., 1970.

[94] E. M. Stein and R. Shakarchi. *Complex analysis.* Princeton Lectures in Analysis, II. Princeton University Press, Princeton, NJ, 2003.

[95] J. Steiner. Einfache beweise der isoeperimetrische hauptsätze. *J. Reine Angew. Math.*, 18:281–296, 1838.

[96] V. N. Sudakov and B. S. Cirel′son. Extremal properties of half-spaces for spherically invariant measures. *Zap. Naučn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI)*, 41:14–24, 165, 1974. Problems in the theory of probability distributions, II.

[97] J. M. Sullivan and F. Morgan. Open problems in soap bubble geometry. *Internat. J. Math.*, 7(6):833–842, 1996.

[98] S. J. Szarek. On the best constants in the Khinchin inequality. *Studia Math.*, 58(2):197–208, 1976.

[99] B. Tomaszewski. A simple and elementary proof of the Kchintchine inequality with the best constant. *Bull. Sci. Math. (2)*, 111(1):103–109, 1987.

[100] L. Vandenberghe and S. Boyd. Semidefinite programming. *SIAM Rev.*, 38(1):49–95, 1996.

[101] K. Weierstrass. *Mathematische Werke. VII. Vorlesungen über Variationsrechnung.* Georg Olms Verlagsbuchhandlung, Hildesheim, 1967.

[102] W. Wichiramala. Proof of the planar triple bubble conjecture. *J. Reine Angew. Math.*, 567:1–49, 2004.

[103] U. Zwick. Computer assisted proof of optimal approximability results. In *Proceedings of the Thirteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 496–505, 2002.