Extremal Sets with Forbidden Configurations and the Independence Ratio of Geometric Hypergraphs

Inaugural-Dissertation

zur
Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Universität zu Köln

vorgelegt von

Davi de Castro Silva

aus Goiânia, Brasilien

Köln, 2022

Berichterstatter (Gutachter):

Prof. Dr. Frank Vallentin,	Universität zu Köln
Prof. Dr. David Gross,	Universität zu Köln
Prof. Dr. Fernando Mario de Oliveira Filho,	Technische Universiteit Delft

Tag der mündlichen Prüfung: 07/12/2021

Contents

Chapter One

Introduction

1.1 The central problem and its history

The central problem considered in this thesis can be phrased by the following general question: given some family of forbidden geometrical configurations, how large can a set be if it does not contain any member of this family?

1.1.1 Forbidding one distance

The simplest and most well-studied instance of this problem concerns forbidden configurations of only two points in $\mathbb{R}^d$, which are then characterized by their distance; since there clearly exist unbounded sets in $\mathbb{R}^d$ which do not span a given distance, the appropriate notion of 'largeness' must take into account their *density* rather than their cardinality or measure. We define the *upper density* $\overline{d}(A)$ of a measurable set $A \subseteq \mathbb{R}^d$ by

$$\overline{d}(A) = \limsup_{R\to\infty} \frac{\mathrm{vol}(A \cap [-R, R]^d)}{\mathrm{vol}([-R, R]^d)},$$

where vol denotes the Lebesgue measure.

Our general problem in this case becomes: What is the maximum[1] upper density that a subset of $\mathbb{R}^d$ can have if it does not contain pairs of points at distance 1? (Note that the problem is dilation-invariant, so there is no loss of generality in assuming the forbidden distance to be 1.) We shall denote this extremal density by $\mathbf{m}_{\mathbb{R}^d}(1)$.[2]

The parameter $\mathbf{m}_{\mathbb{R}^d}(1)$ can also be seen in a more combinatorial perspective as the independence ratio of the *unit distance graph*, whose vertex set is $\mathbb{R}^d$ and two vertices form an edge if their distance is 1; then $\mathbf{m}_{\mathbb{R}^d}(1)$ corresponds to the (upper) density of a maximum independent set. It is many times studied in conjunction with the *measurable chromatic number* $\chi_{\mathrm{m}}(\mathbb{R}^d)$ of the Euclidean space (see e.g. Section 3 of Székely's survey [63]), which is the minimum number of 'color classes' in a measurable partition of $\mathbb{R}^d$ with each class

[1] In order to lighten the presentation, throughout this introduction we shall informally use the word 'maximum' even when there is no guarantee that a maximum is attained.

[2] In the literature this quantity is usually denoted by $m_1(\mathbb{R}^d)$, but our notation is more adequate for the generalizations we will consider later.

being independent in the unit distance graph. The connection between these two parameters is given by the simple inequality $\chi_{\mathrm{m}}(\mathbb{R}^d) \geq 1/\mathbf{m}_{\mathbb{R}^d}(1)$; indeed, if no color class contains pairs of points at unit distance, then each of them has upper density at most $\mathbf{m}_{\mathbb{R}^d}(1)$, and it takes at least $1/\mathbf{m}_{\mathbb{R}^d}(1)$ such classes to cover the whole space. Thus, any upper bound for $\mathbf{m}_{\mathbb{R}^d}(1)$ gives a corresponding lower bound for $\chi_{\mathrm{m}}(\mathbb{R}^d)$.

A simple but very useful observation is that finite configurations of points in $\mathbb{R}^d$ can provide upper bounds for $\mathbf{m}_{\mathbb{R}^d}(1)$. Indeed, suppose $v_1, v_2, \ldots, v_M \in \mathbb{R}^d$ are M points such that any $D+1$ of them contain some two points at unit distance; in other words, the subgraph of the unit distance graph induced by the points $v_1, \ldots, v_M$ has independence number at most D. Then any set $A \subseteq \mathbb{R}^d$ which avoids distance 1 can intersect at most D points in any translated set $\{x+v_1, x+v_2, \ldots, x+v_M\}$; by averaging over the translation parameter x, we conclude that $\overline{d}(A) \leq D/M$. Larman and Rogers [45] constructed several such configurations with a low ratio D/M, in different dimensions $d \geq 2$, from which they derived bounds for $\mathbf{m}_{\mathbb{R}^d}(1)$ and other related parameters.

Despite significant research on the subject, there is still no dimension $d \geq 2$ for which the value of $\mathbf{m}_{\mathbb{R}^d}(1)$ is known. As far back as 1982, Erdős [31] conjectured that $\mathbf{m}_{\mathbb{R}^2}(1) < 1/4$, implying that any measurable planar set covering one fourth of the Euclidean plane contains pairs of points at unit distance. This conjecture is still open; the best upper bound currently known is $\mathbf{m}_{\mathbb{R}^2}(1) \leq 0.25442$, obtained by Ambrus and Matolcsi [1] using a combination of Fourier analytic and linear programming methods. On the other side, a construction given by Croft [14] (c.f. Székely [63]) provides the current best lower bound of 0.22936.

A celebrated combinatorial theorem of Frankl and Wilson [36] on intersecting set systems implies that $\mathbf{m}_{\mathbb{R}^d}(1)$ decays exponentially with the dimension, and obtains the asymptotic upper bound $\mathbf{m}_{\mathbb{R}^d}(1) \leq (1.207+o(1))^{-d}$. This is obtained by applying their intersection theorem to a suitable collection of vectors in $\mathbb{R}^d$ representing subsets of $\{1, 2, \ldots, d\}$, whose pairwise distances encode the size of the corresponding intersection; see [36, Theorem 3]. Using similar arguments to Frankl and Wilson but considering other collections of vectors in Euclidean space, Raigorodskii [51] improved their upper bound to $(1.239+o(1))^{-d}$.

The best asymptotic upper bound for $\mathbf{m}_{\mathbb{R}^d}(1)$ currently known is $(1.268+o(1))^{-d}$, which was obtained by Bachoc, Passuello and Thiery [5] using a combination of linear programming methods and Raigorodskii's constructions; as a consequence, they obtain also the asymptotic lower bound $\chi_{\mathrm{m}}(\mathbb{R}^d) \geq (1.268+o(1))^d$ for the measurable chromatic number of $\mathbb{R}^d$. On the other direction, a construction of Larman and Rogers [45] shows the upper bound $\chi_{\mathrm{m}}(\mathbb{R}^d) \leq (3+o(1))^d$, which in turn gives the current best asymptotic lower bound $\mathbf{m}_{\mathbb{R}^d}(1) \geq (3+o(1))^{-d}$.

We refer the reader to Bachoc, Passuello and Thiery [5] and to DeCorte, Oliveira and Vallentin [22] for the best known numerical bounds for $\mathbf{m}_{\mathbb{R}^d}(1)$ and $\chi_{\mathrm{m}}(\mathbb{R}^d)$, in several dimensions $d \geq 3$.

1.1.2 Forbidding several distances

The situation becomes even more complex and interesting when one forbids multiple distances $r_1, \ldots, r_n > 0$. Let us denote by $\mathbf{m}_{\mathbb{R}^d}(r_1, \ldots, r_n)$ the maximum upper density of a set in $\mathbb{R}^d$ avoiding all of these distances. This parameter was first studied by Székely [61, 62] in connection with the chromatic number of geometric graphs, and it depends not only on the dimension of the space and number of forbidden distances, but also on how these distances relate to each other.

In his first paper [61], Székely made two conjectures connecting the structure of a set $H \subset \mathbb{R}_+$ of forbidden distances to the value of $\mathbf{m}_{\mathbb{R}^d}(H)$. He conjectured that, in dimension $d \geq 2$, one has $\mathbf{m}_{\mathbb{R}^d}(H) = 0$ whenever $\sup H = \infty$; this can also be stated by saying that, whenever a set $A \subseteq \mathbb{R}^d$ has positive upper density, there is some number r_0 such that all the distances greater than r_0 occur among the points of A. (Note that this property is false in dimension 1, as the set $\bigcup_{k\in\mathbb{Z}} [2k,\, 2k+1)$ has density $1/2$ but avoids all odd distances.) This conjecture was proven by Furstenberg, Katznelson and Weiss [37] using ergodic-theoretic methods.

Székely also conjectured that, if $d \geq 2$ and $r_1, r_2, \ldots$ is a sequence of positive numbers converging to zero, then $\mathbf{m}_{\mathbb{R}^d}(r_1, \ldots, r_n) \to 0$ as $n \to \infty$; this can be seen as a kind of 'continuity property' for Steinhaus' theorem in measure theory, which implies that $\mathbf{m}_{\mathbb{R}^d}(H) = 0$ whenever $\inf H = 0$. This conjecture was first proven by Falconer [33], who also noted that it fails in dimension $d = 1$: for any integer $N \geq 1$, the set $\bigcup_{k\in\mathbb{Z}} [2k/3^N,\, (2k+1)/3^N)$ has density $1/2$ but avoids distances $1, 3^{-1}, 3^{-2}, \ldots, 3^{-N}$. We refer the reader to Székely [63] for a survey of results related to the extremal density parameter $\mathbf{m}_{\mathbb{R}^d}$ of a collection of distances, and to the measurable chromatic number of $\mathbb{R}^d$.

More recently, Bukh [8] obtained the stronger asymptotic result that $\mathbf{m}_{\mathbb{R}^d}(r_1, \ldots, r_n)$ tends to $\mathbf{m}_{\mathbb{R}^d}(1)^n$ as the ratios r_{j+1}/r_j between consecutive distances get large, in dimension $d \geq 2$. (The two examples given in the last paragraphs show that this is not true in dimension 1.) Note that Bukh's result easily implies both conjectures of Székely. Bukh also showed that, for $d \geq 2$, there always exists an extremal measurable set $A \subset \mathbb{R}^d$ avoiding a prescribed finite collection of distances; thus $\mathbf{m}_{\mathbb{R}^d}(r_1, \ldots, r_n)$ is attained as a maximum.

1.1.3 Forbidding higher-order configurations

Another interesting instance of our central problem concerns forbidden configurations of more than two points in $\mathbb{R}^d$.

Given a family of finite configurations $P_1, P_2, \ldots, P_n \subset \mathbb{R}^d$, let us define their *independence density* $\mathbf{m}_{\mathbb{R}^d}(P_1,\, P_2,\, \ldots,\, P_n)$ as the maximum upper density of a set in $\mathbb{R}^d$ which does not contain a congruent copy of any of these configurations. This parameter generalizes the earlier notion of extremal density $\mathbf{m}_{\mathbb{R}^d}(r_1, \ldots, r_n)$ from two-point to higher-order configurations; it can also be seen as the independence ratio of the geometric hypergraph with vertex set $\mathbb{R}^d$ whose edges are all isometric copies of P_j, $1 \leq j \leq n$.

Bourgain [7] was the first to consider questions related to the independent density of higher-order configurations. Making use of Fourier analytic methods, he was able to generalize the already mentioned Furstenberg-Katznelson-Weiss theorem [37] from two-point to d-point configurations in general position in $\mathbb{R}^d$, for any $d \geq 2$. For convenience, henceforth we shall say that a configuration $P \subset \mathbb{R}^d$ is *admissible* if it has at most d points and spans a $(|P| - 1)$-dimensional affine hyperplane; Bourgain showed:

Theorem 1.1 (Bourgain [7])**.** *Suppose $P \subset \mathbb{R}^d$ is admissible. If $A \subseteq \mathbb{R}^d$ has positive upper density, then there is some number t_0 such that A contains a congruent copy of $t \cdot P$ for all $t \geq t_0$.*

This result will play an important role in what follows, and will be referred to as *Bourgain's theorem*. Using our independence density notation, it can be restated as the assertion that $\mathbf{m}_{\mathbb{R}^d}((t_j P)_{j\geq 1}) = 0$ for all admissible $P \subset \mathbb{R}^d$ and all unbounded positive sequences $(t_j)_{j\geq 1}$; Bourgain's proof in fact implies the stronger result that

$$\mathbf{m}_{\mathbb{R}^d}(t_1 P,\, t_2 P,\, \ldots,\, t_n P) \to 0 \quad \text{as } n \to \infty$$

whenever the dilation sequence $(t_j)_{j\geq 1}$ tends either to zero or to infinity. His methods then provide also a generalization of Falconer's theorem [33] mentioned before.

It is interesting to note also a 'negative' result of Graham in a similar vein. A set in $\mathbb{R}^d$ is said to be *spherical* if it is contained on the surface of some sphere (with finite radius); it is easy to see that every affinely independent set of points in $\mathbb{R}^d$ (and hence every admissible configuration) is spherical. Graham [39] showed that the conclusion of Bourgain's theorem is *false* whenever the configuration $P \subset \mathbb{R}^d$ considered is *not* spherical; indeed, in this case his arguments show (for instance) that $\mathbf{m}_{\mathbb{R}^d}(P,\, \sqrt{3}P,\, \sqrt{5}P,\, \sqrt{7}P,\, \ldots) > 0$.

In another direction of research, one can fix some finite configuration P and consider how the value of $\mathbf{m}_{\mathbb{R}^d}(P)$ evolves as the dimension d grows. A famous combinatorial theorem of Frankl and Rödl [34] on forbidden intersections implies that, for any fixed $k \geq 1$, the independence density of a regular k-dimensional simplex in $\mathbb{R}^d$ decays exponentially with the dimension d; this is obtained by considering the set of points $(2d)^{-1/2}\{-1, 1\}^d \subset \mathbb{R}^d$ (with d a multiple of 4), where unit simplices correspond to families of subsets of $\{1, \ldots, d\}$ having pairwise symmetric difference $d/2$. This asymptotic result was later extended by the same authors [35] to any non-degenerate simplex (that is, any finite collection of affinely independent points). They deduced from these same methods that non-degenerate simplices are *exponentially Ramsey*, in the sense that any partition of $\mathbb{R}^d$ into sets each avoiding some fixed simplex must have exponentially many parts on the dimension d.

Despite these previous results, it seems that problems related to the independence density of higher-order configurations remain largely unexplored. The only other result we are aware of which can be stated in terms of this parameter is a lemma of Bukh [8, Lemma 6], which shows that $\mathbf{m}_{\mathbb{R}^d}$ is supermultiplicative, i.e. it satisfies $\mathbf{m}_{\mathbb{R}^d}(P_1, \ldots, P_n) \geq \prod_{i=1}^{n} \mathbf{m}_{\mathbb{R}^d}(P_i)$ for all finite configurations $P_1, \ldots, P_n \subset \mathbb{R}^d$.

1.1.4 Forbidding configurations on the sphere

One can also consider similar problems in the spherical setting, where the sets considered are required to lie on the unit sphere $S^d = \{x \in \mathbb{R}^{d+1} : \|x\| = 1\}$.

The first author to consider such questions was Witsenhausen [67], in 1974, when he posed the problem of finding the maximum density[3] of a set in S^d with no two points orthogonal to each other; we shall denote this extremal value by $\mathbf{m}_{S^d}(\pi/2)$.

In his note, Witsenhausen obtained the general upper bound $\mathbf{m}_{S^d}(\pi/2) \leq 1/(d+1)$, which is valid for all dimensions $d \geq 1$ and is sharp for $d = 1$. He also provided the example of two antipodal spherical caps, each of angular radius $\pi/4$, which avoids orthogonal pairs and shows that $\mathbf{m}_{S^d}(\pi/2) \geq (\sqrt{2} + o(1))^{-d}$; this is still the best lower bound known, and Kalai [44] has conjectured that this example is in fact optimal. This problem in now known as the *double-cap conjecture*, and it remains open for all $d \geq 2$; an interesting consequence

[3]The density of a measurable set $A \subseteq S^d$ is the Lebesgue measure of A divided by the total measure of S^d.

of this conjecture would be the strong lower bound $\chi_{\mathrm{m}}(\mathbb{R}^d) \geq (\sqrt{2}+o(1))^d$ on the measurable chromatic number, due to the easily-proven fact that $\mathbf{m}_{S^d}(\pi/2) \geq \mathbf{m}_{\mathbb{R}^{d+1}}(1)$.

As in the Euclidean setting, the Frankl-Wilson theorem [36] on intersecting set systems implies that $\mathbf{m}_{S^d}(\pi/2)$ decays exponentially fast with the dimension d, and obtains the bound $\mathbf{m}_{S^d}(\pi/2) \leq (1.13 + o(1))^{-d}$. This bound was later improved by Raigorodskii [50] to $(1.225 + o(1))^{-d}$, using a refinement of the Frankl–Wilson method.

Despite these strong asymptotic results, Witsenhausen's upper bound $\mathbf{m}_{S^2}(\pi/2) \leq 1/3$ for S^2 was only improved in 2016 by DeCorte and Pikhurko [23], who used linear programming combined with some combinatorial reasoning to show that $\mathbf{m}_{S^2}(\pi/2) \leq 0.313$. DeCorte, Oliveira and Vallentin [22] then used semidefinite programming to further improve this bound to $\mathbf{m}_{S^2}(\pi/2) \leq 0.302$, as well as improve on Witsenhausen's bound for $\mathbf{m}_{S^d}(\pi/2)$ for each $d \leq 7$.

Székely [62] considered also the problem of finding $\mathbf{m}_{S^1}(\theta)$, the maximum density of a measurable set $I \subset S^1$ with no pair of points forming angle θ, for general $0 < \theta \leq \pi$. He showed that $\mathbf{m}_{S^1}(\theta) = 1/2$ whenever θ/π is irrational, but that there exists no measurable θ-avoiding set in the circle attaining this extremal density (so the problem must be formally stated using a supremum). DeCorte and Pikhurko [23] provided a complete answer to the problem of finding $\mathbf{m}_{S^1}(\theta)$, for general $0 < \theta \leq \pi$; they showed that, if $\theta = 2\pi p/q$ with p and q coprime integers, then

$$\mathbf{m}_{S^1}(\theta) = \begin{cases} 1/2 & \text{if } q \text{ is even,} \\ (q-1)/2q & \text{if } q \text{ is odd.} \end{cases}$$

Moreover, in this case of rational θ/π, a measurable θ-avoiding set $I \subset S^1$ attaining the extremal density $\mathbf{m}_{S^1}(\theta)$ always exists.

The problem of several forbidden angles (or distances) in the sphere was first studied by Székely and Wormald [64] in 1989; for a collection of angles $\theta_1, \ldots, \theta_n \in (0, \pi]$, denote by $\mathbf{m}_{S^d}(\theta_1, \ldots, \theta_n)$ the maximum density of a set in S^d containing no pair of points realizing one of these angles. Székely and Wormald derived a recursive formula (on the dimension d) to give upper bounds for $\mathbf{m}_{S^d}(\theta_1, \ldots, \theta_n)$, and used their technique to obtain also upper bounds for $\mathbf{m}_{\mathbb{R}^d}(1)$. More recently, DeCorte and Pikhurko [23] showed that, unlike in the case of the circle S^1, in dimension $d \geq 2$ there always exists a maximal measurable set $A \subset S^d$ avoiding any prescribed collection of angles; the parameter $\mathbf{m}_{S^d}$ is then attained as a maximum in this case.

We shall also consider the independence density parameter $\mathbf{m}_{S^d}$ for higher-order configurations $P_1, \ldots, P_n \subset S^d$. The only result in this setting that we are aware of is a theorem of Frankl and Rödl [34, Theorem 1.13], stating that the independence density of k pairwise orthogonal points in the sphere S^d tends to zero exponentially fast on the dimension, for any fixed $k \geq 2$.

1.2 Outline of the thesis

In this thesis we will give an in-depth study of the independence density parameters $\mathbf{m}_{\mathbb{R}^d}$ and $\mathbf{m}_{S^d}$; in particular, we shall generalize several of the results discussed in Section 1.1. The methods employed here are a combination of harmonic analysis, functional analysis, combinatorics and conic optimization, giving the problems considered an appealing interdisciplinary flavour within mathematics.

Our arguments and results are organized in five chapters, as follows:

Chapter 2: Configuration-avoiding sets in Euclidean space

In Chapter 2 we will formally define the independence density of a family of configurations in $\mathbb{R}^d$, and start our study of this geometrical parameter.

The main tools to be used in this chapter will be a *Counting lemma* (Lemma 2.5) and a *Supersaturation theorem* (Theorem 2.10), both of which are conceptually similar to results of the same name in graph and hypergraph theory (see e.g. [53, 11, 32]), and might be of independent interest. Intuitively, the Counting lemma says that the count of admissible configurations inside a given set does not significantly change if we 'blur' the set a little; this will be proven via Fourier-analytic methods. The Supersaturation theorem states that any bounded set $A \subseteq [-R, R]^d$, which is just slightly denser than the independence density of an admissible configuration P, must necessarily contain a positive proportion of all congruent copies of P lying in $[-R, R]^d$; this is proven by functional-analytic methods, via a compactness and weak* continuity argument.

We will then use these tools to obtain several results on the independence density parameter $\mathbf{m}_{\mathbb{R}^d}$. Strengthening Bourgain's theorem, we show the 'asymptotic independence' property that $\mathbf{m}_{\mathbb{R}^d}(t_1P, t_2P, \ldots, t_nP)$ tends to $\mathbf{m}_{\mathbb{R}^d}(P)^n$ as the ratios t_{j+1}/t_j get large, whenever $P \subset \mathbb{R}^d$ is an admissible configuration; this provides a common generalization of the previously discussed theorems of Bukh [8] and Bourgain [7]. We also show that, by forbidding n distinct dilates of such a configuration P, we can obtain as independence density any real number strictly[4] between $\mathbf{m}_{\mathbb{R}^d}(P)^n$ and $\mathbf{m}_{\mathbb{R}^d}(P)$, but none smaller than $\mathbf{m}_{\mathbb{R}^d}(P)^n$ or larger than $\mathbf{m}_{\mathbb{R}^d}(P)$; moreover, the lower bound is approached whenever the dilation parameters get far apart from each other, and the upper bound whenever they get close to each other. This provides some information on how the value of $\mathbf{m}_{\mathbb{R}^d}(t_1P, t_2P, \ldots, t_nP)$ depends on the dilation parameters t_j.

We will also prove a few other results related to this parameter, namely:

- The general lower bound $\mathbf{m}_{\mathbb{R}^d}(P_1, P_2, \ldots, P_n) \geq \prod_{i=1}^n \mathbf{m}_{\mathbb{R}^d}(P_i)$, which holds for all configurations $P_1, P_2, \ldots, P_n \subset \mathbb{R}^d$; this result is due to Bukh [8].

- Existence of extremizer measurable sets (i.e. having maximal density) which avoid (multiple) admissible configurations; this generalizes a theorem of Bukh from forbidden distances to general admissible configurations.

- Continuity of the independence density function $\mathbf{m}_{\mathbb{R}^d}$ on the set of (multiple) admissible configurations.

This chapter follows the author's paper "Geometrical sets with forbidden configurations" [10].

[4] Whether these boundary values can be attained is not yet clear.

Chapter 3: Configuration-avoiding sets on the sphere

In Chapter 3 we will consider the same problems as before, but now related to the more complicated case of sets on the unit sphere S^d. We will also present (and prove) a spherical analogue of Bourgain's theorem.

Many of the arguments from the Euclidean setting will be used again in the spherical setting, in particular the reliance on our two main 'combinatorial' tools, but there are also some complications we need to solve that are intrinsic to the sphere. One of them is that some aspects of Fourier analysis on S^d are much more complicated than their analogues on $\mathbb{R}^d$, which makes our proof of the spherical Counting lemma correspondingly harder and more technical than its Euclidean counterpart.

Another complication is the lack of dilation invariance in the spherical setting, which makes some of our arguments more intricate; it also prevents us from characterizing the possible independence densities when forbidding several dilates of a given configuration, as was done in the Euclidean setting. The other results proven in the Euclidean space setting will continue to hold in a similar form for sets on the sphere.

This chapter also follows the author's paper "Geometrical sets with forbidden configurations" [10].

Chapter 4: An exact completely positive formulation

Turning towards optimization and the problem of computing the independence density parameters, in Chapter 4 we shall reformulate them into conic linear optimization programs. This will be done by extending the usual completely positive formulation for the independence number of finite graphs, first to finite hypergraphs and then to infinite geometric hypergraphs which encode (families of) admissible configurations.

The main ideas needed for showing that the formulation obtained is exact are not hard, but there are some serious technicalities involved in dealing with infinite-dimensional programs. Perhaps the key innovation here is the use of supersaturation results in order to deal with errors introduced when approximating continuous functions by finite-rank functions; this argument seems to be more general than just the problem considered, and might also be useful for other problems.

This chapter is based on yet unpublished work of the author.

Chapter 5: A recursive Lovász theta number for simplex-avoiding sets

Chapter 5 is devoted to obtaining good upper bounds for the independence density of regular simplices, both in Euclidean space and on the sphere. We shall do so by recursively extending the famous *Lovász theta number* from graphs to geometric hypergraphs encoding such configurations, and then analyzing the resulting optimization program.

Other than obtaining numerical upper bounds for these quantities in any fixed dimension, we will also show that they decay exponentially fast on the dimension of the space. For instance, denoting by $\mathcal{S}_k$ a unit k-dimensional simplex (that is, $k+1$ points whose pairwise distances are all 1), we show that

$$\mathbf{m}_{\mathbb{R}^d}(\mathcal{S}_k) \leq \left(1 - \frac{1}{9k^2} + o(1)\right)^d$$

holds for all $k \geq 1$. This reproves, in the measurable setting, Frankl and Rödl's result [34] that regular simplices are exponentially Ramsey, and significantly improves the existing bounds for the base of the exponential (see Naslund [49] and Sagdeev [54, 55]).

This chapter follows the paper "A recursive Lovász theta number for simplex-avoiding sets" [12], which is joint work of the author with Fernando de Oliveira Filho, Lucas Slot and Frank Vallentin.

Chapter 6: Generalizations of the Lovász theta number to hypergraphs

Finally, in Chapter 6 we will take a more general approach and extend the usual notions of (weighted) theta number and theta body from graphs to uniform hypergraphs. This allows us, for instance, to efficiently compute upper bounds for the independence ratio of geometric hypergraphs in *finite geometries*, such as the binary or q-ary cubes equipped with the Hamming distance.

We shall first recall the basic definitions of these classical notions as they apply to graphs, and then extend them in a recursive manner to higher-order hypergraphs. Several important properties of the weighted theta number and of the theta body of graphs also extend to these generalizations, and will be proven in this chapter; for instance, we show that the theta body of a hypergraph is a convex body encompassing its independent set polytope, and the weighted theta number of a hypergraph gives a polynomial-time computable upper bound for its weighted independence number.

We will then give an exposition on how to exploit symmetries in the considered hypergraph in order to simplify the computation of its theta number; this is a generalization of some of our arguments from Chapter 5 to arbitrary (but finite) uniform hypergraphs. As an example, we provide a detailed analysis of how this can be done in the case of 3-uniform hypergraphs encoding copies of a given equilateral triangle in the hypercube $\{0, 1\}^n$.

This chapter is based on still ongoing work done together with Fernando de Oliveira Filho, Lucas Slot and Frank Vallentin.

General notation and organizational remarks

Here we collect some general pieces of notation which will be used throughout the thesis without further comment.

The characteristic function of a set A will be denoted by χ_A; that is, $\chi_A(x) = 1$ if $x \in A$ and $\chi_A(x) = 0$ otherwise. Given a group G acting on some space X and an element x of this space, we write $\mathrm{Stab}^G(x) := \{g \in G : g.x = x\}$ for the stabilizer subgroup of x. Given a topological space V, the vector space of continuous real-valued functions on V will be denoted $C(V)$.

The averaging notation $\mathbb{E}_{x\in X}$ is used to denote the expectation when the variable x is distributed uniformly over the set X. When X is (a subset of) a compact group G, this measure is (the restriction of) the normalized Haar measure on G, which is the unique Borel probability measure on G which is invariant by both left- and right-actions of this group. Similarly, we write $\mathbb{P}_{x\in X}$ to denote the probability under this same distribution.

All other notation, as well as the necessary technical tools, will be introduced as they are needed. The questions that arose during the research reported at each chapter, and which we are yet unable to answer, are collected at the end of the corresponding chapter. The index at the end of this thesis provides the page where a mathematical term or notation is defined.

Chapter Two

Configuration-avoiding sets in Euclidean space

We will start by studying sets in Euclidean space which avoid a prescribed collection of finite configurations $P_1, \dots, P_n \subset \mathbb{R}^d$. This chapter follows the author's paper "Geometrical sets with forbidden configurations" [10].

Basic definitions and notation

Throughout this chapter, we shall fix an integer $d \geq 2$ and work on the d-dimensional Euclidean space $\mathbb{R}^d$, equipped with its usual inner product $x \cdot y$ and associated Euclidean norm $\|x\|$. We denote by vol the Lebesgue measure on $\mathbb{R}^d$ and by μ the normalized Haar measure on the orthogonal group $\mathrm{O}(d) = \{O \in \mathbb{R}^{d \times d} : O^t O = I\}$.

Given $x \in \mathbb{R}^d$ and $R > 0$, we denote by $Q(x, R)$ the axis-parallel open cube of side length R centered at x. We write $d_{Q(x,R)}(A) := \mathrm{vol}(A \cap Q(x, R))/R^d$ for the density of $A \subseteq \mathbb{R}^d$ inside the cube $Q(x, R)$. The upper density of a measurable set $A \subseteq \mathbb{R}^d$ can then be written as $\overline{d}(A) = \limsup_{R\to\infty} d_{Q(0,R)}(A)$; if the limit exists, we shall instead denote it by $d(A)$ and call it the *density* of A.

A *configuration* P is just a finite subset of $\mathbb{R}^d$, and we define its diameter $\operatorname{diam} P$ as the largest distance between two of its points. Recall from Chapter 1 that a configuration $P \subset \mathbb{R}^d$ on k points is said to be *admissible* if $k \leq d$ and if P is non-degenerate (that is, if it spans a $(k-1)$-dimensional affine hyperplane); it is easy to see that the set of admissible configurations is an open set and that it is dense inside the family of all subsets of $\mathbb{R}^d$ with at most d elements.

We say that two configurations $P, Q \subset \mathbb{R}^d$ are *congruent*, and write $P \simeq Q$, if they can be made equal using only rigid transformations;[1] that is, $P \simeq Q$ if and only if there exist $x \in \mathbb{R}^d$ and $T \in \mathrm{O}(d)$ such that $P = x + T \cdot Q$. Given a configuration $P \subset \mathbb{R}^d$, we say that a set $A \subseteq \mathbb{R}^d$ *avoids* P if there is no subset of A which is congruent to P.

[1] One could also restrict this notion and allow only translations and rotations when defining congruence, giving rise to formally different but very closely related definitions and problems as the ones we consider. All methods used and results obtained here work exactly the same way in this case, with only notational differences.

2.1 Independence density and the counting function

We can now more formally define the notion of *independence density* of a configuration or family of configurations in Euclidean space. There are in fact two closely related versions of this parameter we will need, depending on whether we are considering bounded or unbounded configuration-avoiding sets. Given $n \geq 1$ configuration $P_1, \ldots, P_n \subset \mathbb{R}^d$, we then define the quantities

$$\mathbf{m}_{\mathbb{R}^d}(P_1, \ldots, P_n) := \sup\left\{\overline{d}(A) : A \subset \mathbb{R}^d \text{ avoids } P_i,\ 1 \leq i \leq n\right\} \text{ and}$$
$$\mathbf{m}_{Q(0,R)}(P_1, \ldots, P_n) := \sup\left\{d_{Q(0,R)}(A) : A \subset Q(0, R) \text{ avoids } P_i,\ 1 \leq i \leq n\right\}.$$

Remark 2.1. *For the sake of clarity and notational convenience, whenever possible the results we give about independence density will be stated and proved in the case of only one forbidden configuration. It can be easily verified that these results also hold in the case of several (but finitely many) forbidden configurations, with essentially unchanged proofs. Whenever we need this greater generality we will mention how the corresponding statement would be in the case of several configurations.*

We start our investigations by proving a simple lemma which relates the two versions of independence density just defined:

Lemma 2.2. *For all configurations $P \subset \mathbb{R}^d$ and all $R > 0$, we have*

$$\frac{\mathbf{m}_{Q(0,R)}(P)}{\left(1 + \frac{\operatorname{diam} P}{R}\right)^d} \leq \mathbf{m}_{\mathbb{R}^d}(P) \leq \mathbf{m}_{Q(0,R)}(P).$$

Proof. For the first inequality, suppose $A \subseteq Q(0, R)$ is any set avoiding P and consider the periodic set $A' := A + (R + \operatorname{diam} P)\mathbb{Z}^d$. This set also avoids P, and it has density

$$d(A') = \frac{\operatorname{vol}(A)}{(R + \operatorname{diam} P)^d} = \frac{d_{Q(0,R)}(A)}{\left(1 + \frac{\operatorname{diam} P}{R}\right)^d}.$$

Since we can choose $d_{Q(0,R)}(A)$ arbitrarily close to $\mathbf{m}_{Q(0,R)}(P)$, the leftmost inequality follows.

Now let $A \subseteq \mathbb{R}^d$ be any set avoiding P, and note that $A \cap Q(x, R)$ also avoids P for every $x \in \mathbb{R}^d$. By fixing $\varepsilon > 0$ and then averaging over all x inside a large enough cube $Q(0, R')$ (depending on A, $\operatorname{diam} P$ and ε), we conclude there is $x \in \mathbb{R}^d$ for which

$$\operatorname{vol}(A \cap Q(x, R)) > (\overline{d}(A) - \varepsilon)R^d.$$

The rightmost inequality follows. □

As we are interested in the study of sets avoiding certain configurations, it is useful also to have a way of *counting* how many such configurations there are in a given set. For a given configuration $P = \{v_1, v_2, \ldots, v_k\} \subset \mathbb{R}^d$ and a measurable set $A \subseteq \mathbb{R}^d$, we define

$$I_P(A) := \int_{\mathbb{R}^d} \int_{\mathrm{O}(d)} \chi_A(x + Tv_1)\chi_A(x + Tv_2) \cdots \chi_A(x + Tv_k)\, d\mu(T)\, dx,$$

which represents how many (congruent) copies of P are contained in A. This quantity $I_P(A)$ can of course be infinite if the set A is unbounded, but we will use it almost exclusively for bounded sets. We can similarly define its weighted version

$$I_P(f) := \int_{\mathbb{R}^d} \int_{\mathrm{O}(d)} f(x + Tv_1)\, f(x + Tv_2) \cdots f(x + Tv_k)\, d\mu(T)\, dx,$$

whenever $f : \mathbb{R}^d \to \mathbb{R}$ is a measurable function for which this integral makes sense (say, for $f \in L^k(\mathbb{R}^d)$); note that $I_P(A) \equiv I_P(\chi_A)$ in this notation. A large part of our analysis consists in getting a better understanding of the counting function I_P.

When a measurable set $A \subseteq \mathbb{R}^d$ avoids some configuration P, it is clear from the definition that $I_P(A) = 0$; however, it is also possible for $I_P(A)$ to be zero even when A contains congruent copies of P. In intuitive terms, the condition $I_P(A) = 0$ means only that A contains a *negligible fraction* of all possible copies of P. The next result shows that this distinction is essentially irrelevant for most purposes:

Lemma 2.3 (Zero-measure removal)**.** *Suppose $P \subset \mathbb{R}^d$ is a finite configuration and $A \subseteq \mathbb{R}^d$ is measurable. If $I_P(A) = 0$, then we can remove a zero measure subset of A in order to remove all copies of P.*

Proof. By Lebesgue's Density theorem, we have that

$$\lim_{\delta \to 0} \left| \frac{1}{\delta^d} \int_{Q(x,\,\delta)} \chi_A(y)\, dy - \chi_A(x) \right| = 0 \quad \text{for almost every } x \in \mathbb{R}^d.$$

Now we remove from A all points x for which this identity does *not* hold, thus obtaining a subset $B \subseteq A$ with $\mathrm{vol}(A \setminus B) = 0$ and

$$\lim_{\delta \to 0} \frac{1}{\delta^d} \int_{Q(x,\,\delta)} \chi_B(y)\, dy = 1 \quad \text{for all } x \in B.$$

We will show that no congruent copy of P remains on this restricted set B.

Suppose for contradiction that B contains a copy $\{u_1, \dots, u_k\}$ of P. Then there exists some $\delta > 0$ such that

$$d_{Q(u_i,\,\delta)}(B) = \frac{1}{\delta^d} \int_{Q(u_i,\,\delta)} \chi_B(y)\, dy \geq 1 - \frac{1}{2^{d+1}k} \quad \text{for all } 1 \leq i \leq k; \tag{2.1}$$

fix such a value of δ. Note that, if $d_{Q(x,\,\delta)}(B) \geq 1 - 1/(2^{d+1}k)$ for some $x \in \mathbb{R}^d$, then $d_{Q(y,\,\delta/2)}(B) \geq 1 - 1/2k$ for all $y \in Q(x,\, \delta/2)$. Our hypothesis (2.1) thus implies that $d_{Q(y,\,\delta/2)}(B) \geq 1 - 1/2k$ whenever $y \in Q(u_i,\, \delta/2)$ for some $1 \leq i \leq k$.

Let $\ell := \max\{\|u_i\| : 1 \leq i \leq k\}$ be the largest length of a vector in our copy of P, and let us write $\mathcal{B}(I,\, \delta/4\ell) := \{T \in \mathrm{O}(d) : \|T - I\|_{2\to 2} \leq \delta/4\ell\}$ for the ball of radius $\delta/4\ell$ (in operator norm) centered on the identity I. Then, whenever $T \in \mathcal{B}(I,\, \delta/4\ell)$, we have that

$Tu_i \in Q(u_i, \delta/2)$ for each $1 \le i \le k$. By union bound,

$$\begin{aligned}
\int_{\mathbb{R}^d} \prod_{i=1}^{k} \chi_B(x + Tu_i)\, dx &\ge \int_{Q(0,\delta/2)} \prod_{i=1}^{k} \chi_B(x + Tu_i)\, dx \\
&= \left(\frac{\delta}{2}\right)^d \mathbb{P}_{x \in Q(0,\delta/2)}(x + Tu_i \in B \text{ for all } 1 \le i \le k) \\
&\ge \left(\frac{\delta}{2}\right)^d \left(1 - \sum_{i=1}^{k} \mathbb{P}_{x \in Q(0,\delta/2)}(x + Tu_i \notin B)\right) \\
&= \left(\frac{\delta}{2}\right)^d \left(1 - \sum_{i=1}^{k} (1 - d_{Q(Tu_i,\delta/2)}(B))\right) \\
&\ge \frac{1}{2}\left(\frac{\delta}{2}\right)^d.
\end{aligned}$$

This immediately implies that

$$I_P(B) \ge \int_{\mathbb{R}^d} \int_{\mathcal{B}(I,\delta/4\ell)} \prod_{i=1}^{k} \chi_B(x + Tu_i)\, d\mu(T)\, dx \ge \frac{\mu(\mathcal{B}(I, \delta/4\ell))}{2} \left(\frac{\delta}{2}\right)^d > 0,$$

contradicting our assumption that $I_P(A) = 0$ and finishing the proof. ◻

2.2 Fourier analysis on $\mathbb{R}^d$ and the Counting lemma

We next show that the count of copies of an admissible configuration P inside a measurable set A does not significantly change if we ignore its fine details and 'blur' the set A a little. The philosophy is similar to the famous *regularity method* in graph theory, where a large graph can be replaced by a much smaller weighted 'reduced graph' (which is an averaged version of the original graph which ignores its fine details) without significantly changing the count of copies of any small subgraph.

The methods we will use here are Fourier analytic in nature, drawing from Bourgain's arguments presented in [7]. We define the Fourier transform on $\mathbb{R}^d$ by

$$\widehat{f}(\xi) := \int_{\mathbb{R}^d} f(x)\, e^{-i\xi\cdot x}\, dx \quad \text{and} \quad \widehat{\sigma}(\xi) := \int_{\mathbb{R}^d} e^{-i\xi\cdot x}\, d\sigma(x),$$

for a suitable function f and a suitable Borel measure σ on $\mathbb{R}^d$. We recall that the convolution between two functions $f, g \in L^2(\mathbb{R}^d)$ is defined by

$$f * g(x) := \int_{\mathbb{R}^d} f(y)\, g(x - y)\, dy.$$

We recall also the basic identities $\widehat{f * g}(\xi) = \widehat{f}(\xi)\,\widehat{g}(\xi)$,

$$\int_{\mathbb{R}^d} f(x)\, g(x)\, dx = \int_{\mathbb{R}^d} \widehat{f}(\xi)\, \widehat{g}(-\xi)\, d\xi \quad \text{and} \quad \int_{\mathbb{R}^d} f(x)\, d\sigma(x) = \int_{\mathbb{R}^d} \widehat{f}(\xi)\, \widehat{\sigma}(-\xi)\, d\xi.$$

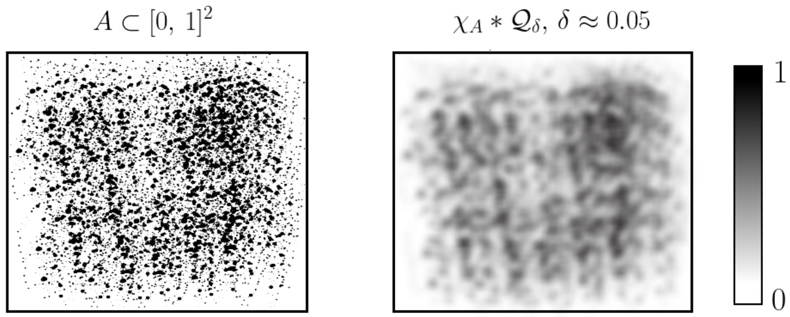

Figure 2.1: An example of a planar set A on the unit square and the corresponding function $\chi_A * Q_\delta$, for some small δ; the shades of gray represent the value this function takes at each point.

Denote $Q_\delta(x) := \delta^{-d}\chi_{Q(0,\,\delta)}(x)$. Note that $f * Q_\delta(x) = \delta^{-d}\int_{Q(x,\,\delta)} f(y)\,dy$ gives the average of a function f inside the cube $Q(x, \delta)$. Specializing to the indicator function of a measurable set $A \subseteq \mathbb{R}^d$, we obtain $\chi_A * Q_\delta(x) = d_{Q(x,\,\delta)}(A)$; this represents a 'blurring' of the set A considered (see Figure 2.1). What we wish to obtain is then an upper bound for the difference $|I_P(A) - I_P(\chi_A * Q_\delta)|$ which goes to zero as δ goes to zero, uniformly over all measurable sets $A \subseteq Q(0, R)$ (and any fixed $R > 0$).

Before delving into the details of our argument, let us present a simple telescoping sum argument which will be needed here and will be reused several times in this thesis. Suppose we wish to bound from above the expression

$$|I_P(f) - I_P(g)| = \left| \int_{\mathbb{R}^d} \int_{\mathrm{O}(d)} \left(\prod_{i=1}^{k} f(x + Tv_i) - \prod_{i=1}^{k} g(x + Tv_i) \right) d\mu(T)\,dx \right|$$

for some given functions f, g and some configuration $P = \{v_1, \dots, v_k\}$. Since we can rewrite the term inside the parenthesis above as the telescoping sum

$$\sum_{i=1}^{k} \prod_{j=1}^{i-1} f(x + Tv_j) \cdot \big(f(x + Tv_i) - g(x + Tv_i)\big) \cdot \prod_{j=i+1}^{k} g(x + Tv_j),$$

it follows from the triangle inequality that $|I_P(f) - I_P(g)|$ is at most

$$\sum_{i=1}^{k} \left| \int_{\mathbb{R}^d} \int_{\mathrm{O}(d)} \prod_{j=1}^{i-1} f(x + Tv_j) \cdot \big(f(x + Tv_i) - g(x + Tv_i)\big) \cdot \prod_{j=i+1}^{k} g(x + Tv_j)\,d\mu(T)\,dx \right|.$$

In order to obtain some bound for $|I_P(f) - I_P(g)|$, it then suffices to obtain a similar bound for an expression of the form

$$\left| \int_{\mathbb{R}^d} \int_{\mathrm{O}(d)} h_1(x + Tu_1) \cdots h_{k-1}(x + Tu_{k-1})\,\big(f(x + Tu_k) - g(x + Tu_k)\big)\,d\mu(T)\,dx \right|$$

whenever each h_i is either f or g, and whenever $(u_1, \dots, u_k)$ is a permutation of the points of P.

We shall refer to an argument of this form (breaking a difference of products into a telescoping sum, using the triangle inequality and bounding each term of the resulting expression) as the *telescoping sum trick*. It is frequently used in modern graph and hypergraph theory when estimating the number of subgraphs inside a given large (hyper)graph G with the aid of edge-discrepancy measures such as the cut norm; such results are usually known as *counting lemmas*, and are an essential part of the regularity method we have already mentioned (see the surveys [53, 11] for details).

In our arguments we will also have cause to use the following estimate. We denote by $G_{d,m}$ (for $m < d$) the Grassmanian of m-dimensional subspaces of $\mathbb{R}^d$ endowed with the normalized Haar measure, and write $\operatorname{dist}(\xi, F)$ for the Euclidean distance between a point $\xi \in \mathbb{R}^d$ and a subspace $F \in G_{d,m}$.

Lemma 2.4. *For all $m < d$ and all $\rho > 0$, we have*

$$\int_{G_{d,m}} \int_{\mathbb{R}^d} |\widehat{f}(\xi)|^2 \, |1 - \widehat{Q}_\delta(\xi)|^2 \, (1 + \operatorname{dist}(\xi, F))^{-\rho} \, d\xi \, dF \le C_d(\delta^2 + \delta^{\rho/2}) \|f\|_2^2.$$

Proof. Using the easily proven bounds $|\widehat{Q}_\delta(\xi)| \le 1$ and $|1 - \widehat{Q}_\delta(\xi)| \le C'_d \delta^2 \|\xi\|^2$, we can bound the expression in the statement by

$$\begin{aligned}
&\int_{G_{d,m}} \left\{ \int_{\|\xi\| \le \delta^{-1/2}} + \int_{\|\xi\| > \delta^{-1/2}} \right\} |\widehat{f}(\xi)|^2 \, |1 - \widehat{Q}_\delta(\xi)|^2 \, (1 + \operatorname{dist}(\xi, F))^{-\rho} \, d\xi \, dF \\
&\quad \le \int_{\|\xi\| \le \delta^{-1/2}} |\widehat{f}(\xi)|^2 \, C'^2_d \delta^2 \, d\xi + \int_{\|\xi\| > \delta^{-1/2}} |\widehat{f}(\xi)|^2 \int_{G_{d,m}} 4(1 + \operatorname{dist}(\xi, F))^{-\rho} \, dF \, d\xi \\
&\quad \le C'^2_d \delta^2 \, \|f\|_2^2 + 4 \, \|f\|_2^2 \sup_{\|\xi\| > \delta^{-1/2}} \int_{G_{d,m}} (1 + \operatorname{dist}(\xi, F))^{-\rho} \, dF \\
&\quad \le C_d(\delta^2 + \delta^{\rho/2}) \|f\|_2^2,
\end{aligned}$$

as wished. □

We are now ready to formally state and prove our main technical tool in the Euclidean setting, which by analogy with methods from graph theory we shall call the Counting lemma. Most of its proof can be read off Bourgain's paper [7] (though it is not expressed in this form), and here we will follow his arguments.

Lemma 2.5 (Counting lemma)**.** *For every admissible configuration $P \subset \mathbb{R}^d$ there exists a constant $C_P > 0$ such that the following holds. For every $R > 0$ and any measurable set $A \subseteq Q(0, R)$, we have that*

$$|I_P(A) - I_P(\chi_A * Q_\delta)| \le C_P \delta^{1/4} R^d \qquad \forall \delta \in (0, 1].$$

Moreover, the same constant C_P can be made to hold uniformly over all configurations P' inside a small neighborhood of P.

Proof. Let $(v_1, \ldots, v_k)$ be a fixed permutation of the points of P; by first translating all points by $-v_k$ and using translation invariance we may assume that $v_k = 0$, so that $P \simeq \{0, v_1, \ldots, v_{k-1}\}$. We will work a bit more generally and show that a bound as in the statement of the lemma holds for

$$\left| \int_{\mathbb{R}^d} \int_{\mathrm{O}(d)} f_0(x) f_1(x + T v_1) \cdots f_{k-2}(x + T v_{k-2}) \right. \\ \left. \times \left(f_{k-1}(x + T v_{k-1}) - f_{k-1} * Q_\delta(x + T v_{k-1}) \right) d\mu(T)\, dx \right| \tag{2.2}$$

whenever $f_0, f_1, \ldots, f_{k-1} : Q(0, R) \to [-1, 1]$ are bounded measurable functions. By our telescoping sum trick, this immediately implies the result.

By considering the variables $y_1 = T v_1, \ldots, y_{k-1} = T v_{k-1}$, we can rewrite the expression

$$\int_{\mathbb{R}^d} \int_{\mathrm{O}(d)} f_0(x) f_1(x + T v_1) \cdots f_{k-1}(x + T v_{k-1})\, d\mu(T)\, dx$$

counting weighted copies of P as

$$\int f_0(x)\, f_1(x + y_1) \cdots f_{k-1}(x + y_{k-1})\, d\sigma^{(d-1)}(y_1)\, d\sigma^{(d-2)}_{y_1}(y_2) \cdots d\sigma^{(d-k+1)}_{y_1,\ldots,y_{k-2}}(y_{k-1})\, dx,$$

where $\sigma^{(d-j)}_{y_1,\ldots,y_{j-1}}$ is the average over a $(d-j)$-dimensional sphere in $\mathbb{R}^d$ depending on the points $y_1, \ldots, y_{j-1}$ already fixed and on P. (More precisely, if $y_i = T v_i$ for $1 \le i \le j-1$, then $\sigma^{(d-j)}_{y_1,\ldots,y_{j-1}}$ is the uniform probability measure on the $(d-j)$-sphere which is the orbit of $T v_j$ under $\mathrm{Stab}^{\mathrm{O}(d)}(y_1, \ldots, y_{j-1})$.) We will not need an explicit description for these measures $\sigma^{(d-j)}_{y_1,\ldots,y_{j-1}}$, only the simple Fourier estimate

$$\left| \widehat{\sigma}^{(d-j)}_{y_1,\ldots,y_{j-1}}(\xi) \right| \le C'_P \left(1 + \mathrm{dist}(\xi, [y_1, \ldots, y_{j-1}])\right)^{-(d-j)/2} \tag{2.3}$$

which follows from the decay at infinity of the Fourier transform of the $(d-j)$-sphere on $\mathbb{R}^{d-j+1}$.

Let us denote for simplicity

$$d\Omega_j(y_1, \ldots, y_j) := d\sigma^{(d-1)}(y_1)\, d\sigma^{(d-2)}_{y_1}(y_2) \cdots d\sigma^{(d-j)}_{y_1,\ldots,y_{j-1}}(y_j).$$

Let $G := f_{k-1} - f_{k-1} * Q_\delta$ and, for $Y = (y_1, \ldots, y_{k-2}) \in (\mathbb{R}^d)^{k-2}$ fixed, denote

$$F_Y(x) := f_0(x) f_1(x + y_1) \cdots f_{k-2}(x + y_{k-2}).$$

The expression (2.2) we wish to bound may then be written as

$$\left| \iiint F_Y(x)\, G(x + y_{k-1})\, d\sigma^{(d-k+1)}_Y(y_{k-1})\, d\Omega_{k-2}(Y)\, dx \right|,$$

which can in turn be bounded by

$$\begin{aligned}
&\int \left| \iint F_Y(x)\, G(x + y_{k-1})\, d\sigma_Y^{(d-k+1)}(y_{k-1})\, dx \right| d\Omega_{k-2}(Y) \\
&\quad = \int \left| \int \widehat{F}_Y(-\xi)\, \widehat{G}(\xi)\, \widehat{\sigma}_Y^{(d-k+1)}(-\xi)\, d\xi \right| d\Omega_{k-2}(Y) \\
&\quad \le \int \|\widehat{F}_Y\|_2 \left(\int |\widehat{G}(\xi)|^2\, |\widehat{\sigma}_Y^{(d-k+1)}(\xi)|^2\, d\xi \right)^{1/2} d\Omega_{k-2}(Y) \\
&\quad \le \Big(\sup_Y \|\widehat{F}_Y\|_2 \Big) \int \left(\int |\widehat{f}_{k-1}(\xi)|^2\, |1 - \widehat{Q}_\delta(\xi)|^2\, |\widehat{\sigma}_Y^{(d-k+1)}(\xi)|^2\, d\xi \right)^{1/2} d\Omega_{k-2}(Y).
\end{aligned}$$

Since $\|\widehat{F}_Y\|_2 = \|F_Y\|_2$ and $|F_Y(x)| \le |f_0(x)|$ pointwise for all $Y \in (\mathbb{R}^d)^k$, we have that the supremum above is at most $\|f_0\|_2$.

Using Cauchy-Schwarz on the outer integral of the last expression and the Fourier estimate (2.3), we conclude that the right-hand side is at most

$$C'_P \|f_0\|_2 \left(\iint |\widehat{f}_{k-1}(\xi)|^2\, |1 - \widehat{Q}_\delta(\xi)|^2 (1 + \operatorname{dist}(\xi, [Y]))^{-(d-k+1)}\, d\xi\, d\Omega_{k-2}(Y) \right)^{1/2}.$$

Note that the expression inside the parenthesis above is equal to the left-hand side of the expression in Lemma 2.4, with $f = f_{k-1}$, $m = k-2$ and $\rho = d-k+1$. It follows that the original expression (2.2) we wanted to bound (for $d \ge k$ and $0 < \delta \le 1$) is at most $C_P \delta^{1/4} \|f_0\|_2 \|f_{k-1}\|_2$, and the desired inequality follows.

The claim that this constant C_P can be made uniform inside a small neighborhood of P follows by analyzing our proof. Indeed, the only place in our arguments where we used information on the specific positioning of the points in P was in the Fourier estimate (2.3), which can easily be made robust to small perturbations of the considered configuration. □

2.3 Continuity properties of the counting function

Given some configuration P on the space $\mathbb{R}^d$, it is sometimes important to understand how much the count of congruent copies of P on a set $A \subseteq \mathbb{R}^d$ can change if we perturb the set A a little. An instance of this problem was already considered in the Counting lemma, where the perturbation was given by blurring and it was seen that the counting function I_P is somewhat robust to small perturbations (in the case of admissible configurations).

Using our telescoping sum trick, it is easy to show that I_P is also robust to small perturbations measured by the L^∞ norm; more precisely, I_P is continuous on $L^\infty(Q(0,R))$ for any fixed $R > 0$. When P is admissible, we obtain the following significantly stronger continuity property:

Lemma 2.6 (Weak* continuity)**.** *If $P \subset \mathbb{R}^d$ is an admissible configuration, then for every fixed $R > 0$ the function I_P is weak* continuous on the unit ball of $L^\infty(Q(0,R))$.*

Proof. Denote the (closed) unit ball of $L^\infty(Q(0,R))$ by $\mathcal{B}_\infty$. Since $\mathcal{B}_\infty$ endowed with the weak* topology is metrizable, it suffices to prove that I_P is sequentially continuous (i.e. that $I_P(f_i) \xrightarrow{i\to\infty} I_P(f)$ whenever $f_i \xrightarrow{i\to\infty} f$).

Suppose then $(f_i)_{i\geq 1} \subset \mathcal{B}_\infty$ is a sequence weak* converging to $f \in \mathcal{B}_\infty$. It follows that, for every $x \in Q(0, R)$ and every $\delta > 0$, we have

$$f_i * Q_\delta(x) = \delta^{-d} \int_{Q(x,\delta)} f_i(y)\,dy \xrightarrow{i\to\infty} \delta^{-d} \int_{Q(x,\delta)} f(y)\,dy = f * Q_\delta(x).$$

Since $f * Q_\delta$ and each $f_i * Q_\delta$ are Lipschitz with the same constant (depending only on δ, as $\|f\|_\infty, \|f_i\|_\infty \leq 1$) and $Q(0, R)$ is bounded, this easily implies that

$$\|f_i * Q_\delta - f * Q_\delta\|_\infty \to 0 \quad \text{as } i \to \infty.$$

In particular, it follows that $\lim_{i\to\infty} I_P(f_i * Q_\delta) = I_P(f * Q_\delta)$.

Since P is admissible, by the Counting lemma (Lemma 2.5) we have

$$|I_P(f * Q_\delta) - I_P(f)|, \quad |I_P(f_i * Q_\delta) - I_P(f_i)| \leq C_P \delta^{1/4} R^d \quad \forall i \geq 1.$$

Choosing $i_0(\delta) \geq 1$ sufficiently large so that

$$|I_P(f_i * Q_\delta) - I_P(f * Q_\delta)| \leq C_P \delta^{1/4} R^d \qquad \forall i \geq i_0(\delta),$$

we conclude that

$$\begin{aligned} |I_P(f) - I_P(f_i)| &\leq |I_P(f) - I_P(f * Q_\delta)| + |I_P(f * Q_\delta) - I_P(f_i * Q_\delta)| \\ &\quad + |I_P(f_i * Q_\delta) - I_P(f_i)| \\ &\leq 3C_P \delta^{1/4} R^d \qquad \forall i \geq i_0(\delta). \end{aligned}$$

Since $\delta > 0$ is arbitrary, this implies that $\lim_{i\to\infty} I_P(f_i) = I_P(f)$, as wished. □

We will also need an *equicontinuity* property for the family of counting functions $P \mapsto I_P(A)$, over all bounded measurable sets $A \subseteq \mathbb{R}^d$. In what follows we shall write $\mathcal{B}(P, r) \subset (\mathbb{R}^d)^k$ for the ball of radius r centered on[2] $P = \{v_1, \ldots, v_k\}$, where the distance from P to $Q = \{u_1, \ldots, u_k\}$ is given by

$$\|Q - P\|_\infty := \max\{\|u_i - v_i\| : 1 \leq i \leq k\}.$$

Lemma 2.7 (Equicontinuity)**.** *For every admissible $P \subset \mathbb{R}^d$ and every $\varepsilon > 0$ there is $\delta > 0$ such that the following holds. If $\|P' - P\|_\infty \leq \delta$, then for all $R > 0$ we have*

$$|I_{P'}(A) - I_P(A)| \leq \varepsilon R^d \quad \textit{for all measurable } A \subseteq Q(0, R).$$

Proof. We will use the fact that the constant C_P promised in the Counting lemma can be made uniform inside a small neighborhood of P; more precisely, there is $r > 0$ and a constant $\tilde{C}_P > 0$ such that

$$|I_{P'}(A) - I_{P'}(\chi_A * Q_\rho)| \leq \tilde{C}_P \rho^{1/4} R^d \qquad \forall \rho \in (0, 1]$$

holds for all $P' \in \mathcal{B}(P, r)$, $R > 0$ and (measurable) $A \subseteq Q(0, R)$.

[2]Strictly speaking, we would need this set P to be ordered, but since the specific order chosen for its points makes no difference in our arguments we will ignore this detail.

Fix constants $R, \rho, \delta > 0$ with $\delta < \rho$. For any set $A \subseteq Q(0, R)$ and any points $x, y \in Q(0, R)$ with $\|x - y\| \leq \delta$, we have that

$$\begin{aligned}
|\chi_A * Q_\rho(x) - \chi_A * Q_\rho(y)| &= \frac{|\operatorname{vol}(A \cap Q(x, \rho)) - \operatorname{vol}(A \cap Q(y, \rho))|}{\rho^d} \\
&\leq \frac{\operatorname{vol}(Q(x, \rho) \setminus Q(y, \rho))}{\rho^d} \\
&\leq \frac{\rho^d - (\rho - \delta)^d}{\rho^d}.
\end{aligned}$$

By the usual telescoping sum argument, whenever $\|P' - P\|_\infty \leq \delta$ we conclude that

$$|I_{P'}(\chi_A * Q_\rho) - I_P(\chi_A * Q_\rho)| \leq k\frac{\rho^d - (\rho - \delta)^d}{\rho^d} R^d.$$

Take $\rho > 0$ small enough so that $\tilde{C}_P \rho^{1/4} \leq \varepsilon/3$, and for this value of ρ take $0 < \delta < \min\{r, \rho\}$ small enough so that $(\rho - \delta)^d \geq (1 - \varepsilon/3k)\rho^d$. Then, for any configuration $P' \in \mathcal{B}(P, \delta)$ and any set $A \subseteq Q(0, R)$, we obtain

$$\begin{aligned}
|I_{P'}(A) - I_P(A)| &\leq |I_{P'}(A) - I_{P'}(\chi_A * Q_\rho)| + |I_{P'}(\chi_A * Q_\rho) - I_P(\chi_A * Q_\rho)| \\
&\quad + |I_P(\chi_A * Q_\rho) - I_P(A)| \\
&\leq \tilde{C}_P \rho^{1/4} R^d + k\frac{\rho^d - (\rho - \delta)^d}{\rho^d} R^d + \tilde{C}_P \rho^{1/4} R^d \\
&\leq \left(\frac{\varepsilon}{3} + k\frac{\varepsilon}{3k} + \frac{\varepsilon}{3}\right) R^d = \varepsilon R^d,
\end{aligned}$$

as desired. □

2.4 The Supersaturation theorem

Now we wish to show that geometrical hypergraphs encoding copies of some admissible configuration P have a nice *supersaturation property*: if a set $A \subseteq \mathbb{R}^d$ is just slightly denser than the independence density of P, then it must contain a *positive proportion* of all congruent copies of P. This result is quite similar, both formally and in spirit, to an important combinatorial theorem of Erdős and Simonovits [32] in the setting of forbidden graphs and hypergraphs.

Remark 2.8. *The insight that supersaturation results can be used to better study extremal geometrical problems of the kind we are interested in is due to Bukh [8]. He introduced the notion of a 'supersaturable property' as any characteristic of measurable sets which satisfies several conditions meant to enable the proof of a supersaturation result; the prototypical and most important example of supersaturable property given in Bukh's paper is that of avoiding a finite collection of distances. Here we will obtain similar results in the case of avoiding general admissible configurations, but our method of proof is more analytical in nature and quite different from his.*

Using our 'zero-measure removal' Lemma 2.3, we can immediately obtain a weak supersaturation property which holds for any $R > 0$ and any configuration $P \subset \mathbb{R}^d$:

(WS) If $d_{Q(0,R)}(A) > \mathbf{m}_{Q(0,R)}(P)$, then $I_P(A) > 0$.

For our purposes, however, we will need to strengthen this simple property in two ways: first to obtain a uniform lower bound on $I_P(A)$ which depends only on R and the 'slackness' $d_{Q(0,R)}(A) - \mathbf{m}_{Q(0,R)}(P)$ (but not on the specific set $A \subseteq Q(0, R)$); and then to make the proportion $I_P(A \cap Q(0, R))/R^d$ of copies of P uniform also on the size R of the cube considered.

The first strengthening can be obtained from (WS) by a compactness argument, using the fact that the counting function of admissible configurations is weak* continuous:

Lemma 2.9 (Weak supersaturation)**.** *Suppose $P \subset \mathbb{R}^d$ is admissible, and fix $\varepsilon > 0$ and $R > 0$. There exists some $c_0 > 0$ so that, if $A \subseteq Q(0, R)$ satisfies*

$$d_{Q(0,R)}(A) \geq \mathbf{m}_{Q(0,R)}(P) + \varepsilon,$$

then $I_P(A) \geq c_0$.

Proof. Suppose for contradiction that the result is false. Then there exist $\varepsilon > 0$, $R > 0$ and a sequence $(A_i)_{i\geq 1}$ of subsets of $Q(0, R)$, each of density at least $\mathbf{m}_{Q(0,R)}(P) + \varepsilon$, which satisfy $\lim_{i\to\infty} I_P(A_i) = 0$.

Since the unit ball $\mathcal{B}_\infty$ of $L^\infty(Q(0, R))$ is weak* compact (and also metrizable in this topology), by possibly restricting to a subsequence we may assume that the indicator functions χ_{A_i} converge in the weak* topology of $L^\infty(Q(0, R))$; let us denote their limit by $f \in \mathcal{B}_\infty$. It is clear that $0 \leq f \leq 1$ almost everywhere, and

$$\frac{1}{R^d}\int_{Q(0,R)} f(x)\,dx = \lim_{i\to\infty} \frac{1}{R^d}\int_{Q(0,R)} \chi_{A_i}(x)\,dx \geq \mathbf{m}_{Q(0,R)}(P) + \varepsilon.$$

By weak* continuity of I_P (Lemma 2.6), we also have $I_P(f) = \lim_{i\to\infty} I_P(A_i) = 0$.

Now let $B := \{x \in Q(0, R) : f(x) \geq \varepsilon\}$. Since

$$\varepsilon \cdot \chi_B(x) \leq f(x) < \varepsilon + \chi_B(x) \quad \text{for a.e. } x \in Q(0, R),$$

we conclude that $I_P(B) \leq \varepsilon^{-k} I_P(f) = 0$ and

$$d_{Q(0,R)}(B) > \frac{1}{R^d}\int_{Q(0,R)} f(x)\,d\sigma(x) - \varepsilon \geq \mathbf{m}_{Q(0,R)}(P).$$

But this set B contradicts (WS) (or Lemma 2.3), finishing the proof. □

Our desired supersaturation result now follows from a simple averaging argument:

Theorem 2.10 (Supersaturation theorem)**.** *For every admissible configuration $P \subset \mathbb{R}^d$ and every $\varepsilon > 0$ there exist $c > 0$ and $R_0 > 0$ such that the following holds. For all $R \geq R_0$, if $A \subseteq Q(0, R)$ satisfies*

$$d_{Q(0,R)}(A) \geq \mathbf{m}_{Q(0,R)}(P) + \varepsilon,$$

then $I_P(A) \geq cR^d$.

Proof. For a given $\varepsilon > 0$, take some $R_1 > 0$ large enough so that $\mathbf{m}_{Q(0,R_1)}(P) \le \mathbf{m}_{\mathbb{R}^d}(P) + \varepsilon/4$. We will show that the conclusion of the theorem holds for $R_0 = 4dR_1/\varepsilon$ and some constant $c > 0$.

Let then $R \ge 4dR_1/\varepsilon$, and let $A \subseteq Q(0, R)$ be a set having density $d_{Q(0,R)}(A)$ at least $\mathbf{m}_{Q(0,R)}(P) + \varepsilon$. Since

$$\mathbf{m}_{Q(0,R_1)}(P) \le \mathbf{m}_{\mathbb{R}^d}(P) + \varepsilon/4 \le \mathbf{m}_{Q(0,R)}(P) + \varepsilon/4,$$

we have that $\mathrm{vol}(A) \ge (\mathbf{m}_{Q(0,R_1)}(P) + 3\varepsilon/4)R^d$.

Denote $K := \lfloor R/R_1 \rfloor$. Since

$$K^d R_1^d > \left(1 - \frac{R_1}{R}\right)^d R^d \ge \left(1 - \frac{dR_1}{R}\right) R^d \ge \left(1 - \frac{\varepsilon}{4}\right) R^d,$$

we conclude that $\mathrm{vol}\,(A \cap Q(0, KR_1)) \ge (\mathbf{m}_{Q(0,R_1)}(P) + \varepsilon/2)K^d R_1^d$. Dividing the cube $Q(0, KR_1)$ into K^d cubes of side length R_1, by averaging we conclude that at least $\varepsilon K^d/4$ of these cubes $Q(x, R_1)$ satisfy

$$\mathrm{vol}\,(A \cap Q(x, R_1)) \ge (\mathbf{m}_{Q(0,R_1)}(P) + \varepsilon/4)R_1^d.$$

By Lemma 2.9, there is some $c_0 > 0$ (independent of A) so that $I_P(A \cap Q(x, R_1)) \ge c_0$ for each of these cubes where A has high density. We conclude that

$$I_P(A) \ge \frac{\varepsilon K^d}{4} c_0 > \left(\frac{\varepsilon c_0}{2^{d+2} R_1^d}\right) R^d,$$

finishing the proof. □

Remark 2.11. *Exactly the same proof works in the case of several configurations, showing that $I_{P_i}(A) \ge c(\varepsilon)R^d$ holds for some $1 \le i \le n$ whenever the density condition $d_{Q(0,R)}(A) \ge \mathbf{m}_{Q(0,R)}(P_1, \ldots, P_n) + \varepsilon$ is satisfied (assuming R is large enough and all the configurations P_i are admissible).*

Following Bukh [8], for each $\delta > 0$ and $\gamma > 0$ we define the *zooming-out operator* $\mathcal{Z}_\delta(\gamma)$ as the map which takes a measurable set $A \subseteq \mathbb{R}^d$ to the set

$$\mathcal{Z}_\delta(\gamma)A := \{x \in \mathbb{R}^d : d_{Q(x,\delta)}(A) \ge \gamma\}.$$

Intuitively, $\mathcal{Z}_\delta(\gamma)A$ represents the points where A is not too sparse at scale δ.

Using the Supersaturation theorem together with the Counting lemma, we can now show that the existence of copies of P in a set A follows also from the weaker assumption that its *zoomed-out version* $\mathcal{Z}_\delta(\gamma)A$ has density higher than $\mathbf{m}_{\mathbb{R}^d}(P)$ (rather than A itself having this same density); this property will be important for us later.

Corollary 2.12. *For every admissible configuration $P \subset \mathbb{R}^d$ and every $\varepsilon > 0$ there exists $\delta_0 > 0$ such that the following holds. For all $\delta \le \delta_0$, if $A \subseteq \mathbb{R}^d$ satisfies*

$$\overline{d}\,(\mathcal{Z}_\delta(\varepsilon)A) \ge \mathbf{m}_{\mathbb{R}^d}(P) + \varepsilon,$$

then A contains a congruent copy of P.

Proof. Let R_0, $c > 0$ be the constants promised in the Supersaturation theorem applied to P and with ε substituted by $\varepsilon/3$. Up to substituting R_0 by some larger constant, we may also assume that $\mathbf{m}_{Q(0,R)}(P) \leq \mathbf{m}_{\mathbb{R}^d}(P) + \varepsilon/3$ for all $R \geq R_0$ (see Lemma 2.2).

Suppose $A \subseteq \mathbb{R}^d$ satisfies $\overline{d}(\mathcal{Z}_\delta(\varepsilon)A) \geq \mathbf{m}_{\mathbb{R}^d}(P) + \varepsilon$ for some $0 < \delta \leq 1$. There must then exist some $R \geq R_0$ such that

$$d_{Q(0,R)}(\mathcal{Z}_\delta(\varepsilon)A) \geq \mathbf{m}_{\mathbb{R}^d}(P) + 2\varepsilon/3,$$

and for which this last inequality remains true even if we substitute A by $A \cap Q(0, R)$. We may then assume $A \subseteq Q(0, R)$ satisfies $d_{Q(0,R)}(\mathcal{Z}_\delta(\varepsilon)A) \geq \mathbf{m}_{Q(0,R)}(P) + \varepsilon/3$ for some $R \geq R_0$, and wish to show that A contains a copy of P if $\delta > 0$ is small enough depending on P and the parameter ε.

By the Supersaturation theorem, we have $I_P(\mathcal{Z}_\delta(\varepsilon)A) \geq cR^d$. Since

$$\chi_A * Q_\delta(x) = d_{Q(x,\delta)}(A) \geq \varepsilon \cdot \chi_{\mathcal{Z}_\delta(\varepsilon)A}(x) \qquad \text{for all } x \in \mathbb{R}^d,$$

we obtain from the Counting lemma that

$$\begin{aligned} I_P(A) &\geq I_P(\chi_A * Q_\delta) - C_P\delta^{1/4}R^d \\ &\geq \varepsilon^k I_P(\mathcal{Z}_\delta(\varepsilon)A) - C_P\delta^{1/4}R^d \\ &\geq (\varepsilon^k c - C_P\delta^{1/4})R^d. \end{aligned}$$

Taking $\delta > 0$ small enough for this last expression to be positive, we conclude that $I_P(A) > 0$, and so A contains a copy of P. □

2.5 Results on the independence density

We are finally in a position to properly study the independence density parameter for a family of configurations in Euclidean space.

We start by proving a simple lower bound on the independence density of several distinct configurations; this result and the argument we use to prove it are originally due to Bukh [8].

Lemma 2.13 (Supermultiplicativity). *For all $n \geq 1$ and all configurations $P_1, \dots, P_n \subset \mathbb{R}^d$, we have that*

$$\mathbf{m}_{\mathbb{R}^d}(P_1, P_2, \dots, P_n) \geq \prod_{i=1}^n \mathbf{m}_{\mathbb{R}^d}(P_i).$$

Proof. Fix $\varepsilon > 0$ and choose R large enough so that

$$\min_{1\leq i\leq n} (R - \operatorname{diam} P_i)^d \geq (1-\varepsilon)R^d.$$

For each $1 \leq i \leq n$, let $A_i \subseteq Q(0, R - \operatorname{diam} P_i)$ be a set which avoids P_i and satisfies $d_{Q(0,R-\operatorname{diam} P_i)}(A_i) > \mathbf{m}_{\mathbb{R}^d}(P_i) - \varepsilon$ (this is possible by Lemma 2.2). We then construct the R-periodic set $A_i' := A_i + R\mathbb{Z}^d$, which also avoids P_i and has density

$$d(A_i') = \frac{(R - \operatorname{diam} P_i)^d}{R^d} d_{Q(0,R-\operatorname{diam} P_i)}(A_i) > \mathbf{m}_{\mathbb{R}^d}(P_i) - 2\varepsilon.$$

Since each set A'_i is periodic with the same 'fundamental domain' $Q(0, R)$, it follows that the average of $d\big(\bigcap_{i=1}^n (x_i + A'_i)\big)$ over independent translates $x_1, \dots, x_n \in Q(0, R)$ is equal to $\prod_{i=1}^n d(A'_i)$. There must then exist some $x_1, \dots, x_n \in Q(0, R)$ for which

$$d\left(\bigcap_{i=1}^{n}(x_i + A'_i)\right) \geq \prod_{i=1}^{n} d(A'_i) > \prod_{i=1}^{n}(\mathbf{m}_{\mathbb{R}^d}(P_i) - 2\varepsilon).$$

Since $\bigcap_{i=1}^n (x_i + A'_i)$ avoids each of the configurations P_i and $\varepsilon > 0$ was arbitrary, the desired lower bound follows. □

Intuitively, one may regard $\mathbf{m}_{\mathbb{R}^d}(P_1, P_2, \dots, P_n)$ being close to $\prod_{i=1}^n \mathbf{m}_{\mathbb{R}^d}(P_i)$ as some sort of *independence* or *lack of correlation* between the n constraints of forbidding each configuration P_i; in this case, there is no better way to choose a set avoiding all these configurations than simply intersecting optimal P_i-avoiding sets for each i (after suitably translating them). One might then expect this to happen if the *natural sizes* of each P_i are very different from each other, so that each constraint will be relevant in different and largely independent scales.

Our next result shows this is indeed the case whenever the configurations considered are all admissible. (A theorem of Graham [39] implies this is *not* necessarily true if the configurations considered aren't admissible; see Section 2.6 for a discussion.) The proof we present here is based on Bukh's arguments for supersaturable properties, and generalizes his result from two-point configurations to general admissible configurations.

Theorem 2.14 (Asymptotic independence)**.** *If $P_1, P_2, \dots, P_n \subset \mathbb{R}^d$ are admissible configurations, then*

$$\mathbf{m}_{\mathbb{R}^d}(t_1 P_1,\, t_2 P_2,\, \dots,\, t_n P_n) \to \prod_{i=1}^{n} \mathbf{m}_{\mathbb{R}^d}(P_i)$$

as the ratios $t_2/t_1,\, t_3/t_2,\, \dots,\, t_n/t_{n-1}$ tend to infinity.

Proof. We have already seen that

$$\mathbf{m}_{\mathbb{R}^d}(t_1 P_1,\, t_2 P_2,\, \dots,\, t_n P_n) \geq \prod_{i=1}^{n} \mathbf{m}_{\mathbb{R}^d}(t_i P_i) = \prod_{i=1}^{n} \mathbf{m}_{\mathbb{R}^d}(P_i)$$

always holds, so it suffices to show that $\mathbf{m}_{\mathbb{R}^d}(t_1 P_1,\, t_2 P_2,\, \dots,\, t_n P_n) \leq \prod_{i=1}^n \mathbf{m}_{\mathbb{R}^d}(P_i) + \varepsilon$ whenever $\varepsilon > 0$ and the ratios between consecutive scales t_i are large enough. We shall proceed by induction, with the case $n = 1$ being trivial.

Suppose $n \geq 2$ and the theorem holds for configurations $P_1, \dots, P_{n-1}$. Fix $\varepsilon > 0$, and let $t_1, \dots, t_{n-1} > 0$ be scales for which

$$\mathbf{m}_{\mathbb{R}^d}(t_1 P_1,\, \dots,\, t_{n-1} P_{n-1}) \leq \prod_{i=1}^{n-1} \mathbf{m}_{\mathbb{R}^d}(P_i) + \varepsilon;$$

now take $R_0 > 0$ large enough so that

$$\mathbf{m}_{Q(0,R)}(t_1 P_1,\, \dots,\, t_{n-1} P_{n-1}) \leq \mathbf{m}_{\mathbb{R}^d}(t_1 P_1,\, \dots,\, t_{n-1} P_{n-1}) + \varepsilon$$

holds for all $R \geq R_0$ (this quantity exists by Lemma 2.2).

If $A \subseteq \mathbb{R}^d$ is a measurable set avoiding $t_1 P_1, \ldots, t_{n-1} P_{n-1}$, then clearly

$$d_{Q(x,R)}(A) \leq \mathbf{m}_{Q(0,R)}(t_1 P_1, \ldots, t_{n-1} P_{n-1}) \quad \text{for all } x \in \mathbb{R}^d,\ R > 0. \tag{2.4}$$

Moreover, if A also avoids $t_n P_n$ for some $t_n > 0$, then A/t_n avoids P_n and so by Corollary 2.12 there is some $\delta_0 > 0$ (depending only on P_n and ε) for which

$$\overline{d}\,(\mathcal{Z}_\delta(\varepsilon)A/t_n) \leq \mathbf{m}_{\mathbb{R}^d}(P_n) + \varepsilon \quad \forall \delta \leq \delta_0. \tag{2.5}$$

Suppose now that $t_n \geq R_0/\delta_0$, and let $A \subseteq \mathbb{R}^d$ be any measurable set avoiding $t_1 P_1, \ldots, t_n P_n$. We conclude from (2.4) that

$$\begin{aligned} d_{Q(x,\,t_n\delta_0)}(A) &\leq \mathbf{m}_{Q(0,\,t_n\delta_0)}(t_1 P_1, \ldots, t_{n-1} P_{n-1}) \\ &\leq \mathbf{m}_{\mathbb{R}^d}(t_1 P_1, \ldots, t_{n-1} P_{n-1}) + \varepsilon \\ &\leq \prod_{i=1}^{n-1} \mathbf{m}_{\mathbb{R}^d}(P_i) + 2\varepsilon \end{aligned}$$

holds for all $x \in \mathbb{R}^d$, and from (2.5) we have

$$\overline{d}\,(\mathcal{Z}_{t_n\delta_0}(\varepsilon)A) = \overline{d}\,(\mathcal{Z}_{\delta_0}(\varepsilon)A/t_n) \leq \mathbf{m}_{\mathbb{R}^d}(P_n) + \varepsilon.$$

This means that the density of A inside cubes $Q(x, t_n\delta_0)$ of side length $t_n\delta_0$ is at most ε (when $x \notin \mathcal{Z}_{t_n\delta_0}(\varepsilon)A$) except at a set of upper density at most $\mathbf{m}_{\mathbb{R}^d}(P_n) + \varepsilon$, when it is instead no more than $\prod_{i=1}^{n-1} \mathbf{m}_{\mathbb{R}^d}(P_i) + 2\varepsilon$. Taking averages, we conclude that

$$\overline{d}(A) \leq \varepsilon + (\mathbf{m}_{\mathbb{R}^d}(P_n) + \varepsilon)\left(\prod_{i=1}^{n-1} \mathbf{m}_{\mathbb{R}^d}(P_i) + 2\varepsilon\right) \leq \prod_{i=1}^{n} \mathbf{m}_{\mathbb{R}^d}(P_i) + 6\varepsilon.$$

This inequality finishes the proof. □

As an immediate corollary of the last theorem, we conclude that

$$\mathbf{m}_{\mathbb{R}^d}(t_1 P, t_2 P, \ldots, t_n P) \to \mathbf{m}_{\mathbb{R}^d}(P)^n$$

as $t_2/t_1, t_3/t_2, \ldots, t_n/t_{n-1} \to \infty$ whenever $P \subset \mathbb{R}^d$ is admissible; let us now show how this result easily implies Bourgain's theorem given in Section 1.1:

Proof of Theorem 1.1. Suppose $A \subset \mathbb{R}^d$ is a measurable set not satisfying the conclusion of the theorem; thus there is a sequence $(t_j)_{j\geq 1}$ tending to infinity such that A does not contain a copy of any $t_j P$. This implies that $\overline{d}(A) \leq \mathbf{m}_{\mathbb{R}^d}(t_1 P, t_2 P, \ldots, t_n P)$ for all $n \in \mathbb{N}$. By taking a suitably fast-growing subsequence, we may then use Theorem 2.14 to obtain (say) $\overline{d}(A) \leq 2\mathbf{m}_{\mathbb{R}^d}(P)^n$ for any fixed $n \geq 1$. This implies that $\overline{d}(A) = 0$, as wished. □

Going back to our study of the independence density for multiple configurations, we will now consider the *opposite* situation of what we have seen before: when the constraints of forbidding each individual configuration are so strongly correlated as to be essentially redundant. One might expect this to be the case, for instance, when we are forbidding very close dilates of a given configuration P.

We will show that this intuition is indeed correct, whether or not the configuration considered is admissible, and the proof is much simpler than in the case of very distant dilates of P (in particular not needing the results from earlier sections).

Lemma 2.15 (Asymptotic redundancy)**.** *For any configuration $P \subset \mathbb{R}^d$, we have that*

$$\mathbf{m}_{\mathbb{R}^d}(t_1 P,\, t_2 P,\, \ldots,\, t_n P) \to \mathbf{m}_{\mathbb{R}^d}(P)$$

as $t_2/t_1,\, t_3/t_2,\, \ldots,\, t_n/t_{n-1} \to 1$.

Proof. Assume by dilation invariance that $t_1 = 1$. By Lemma 2.2, it suffices to show that the convergence above holds with $\mathbf{m}_{\mathbb{R}^d}$ replaced by $\mathbf{m}_{Q(0,R)}$ for all fixed $R > 0$. We will then fix an arbitrary $R > 0$ and prove that $\mathbf{m}_{Q(0,R)}(P,\, t_2 P,\, \ldots,\, t_n P) \to \mathbf{m}_{Q(0,R)}(P)$ as $t_2,\, t_3,\, \ldots,\, t_n \to 1$.

Let $(v_1, v_2, \ldots, v_k)$ be an ordering of the points of P, and consider the continuous function $g_P : (\mathbb{R}^d)^k \times \mathrm{O}(d) \to \mathbb{R}$ given by

$$g_P(x_1,\, \ldots,\, x_k,\, T) := \sum_{j=2}^{k} \|(x_j - x_1) - T(v_j - v_1)\|.$$

Note that $\min_{T \in \mathrm{O}(d)} g_P(x_1,\, \ldots,\, x_k,\, T) = 0$ if and only if $(x_1, \ldots, x_k) \simeq (v_1, \ldots, v_k)$.

Fix some $\varepsilon > 0$, and let $A \subset Q(0, R)$ be a measurable set which avoids P and has density $d_{Q(0,R)}(A) \geq \mathbf{m}_{Q(0,R)}(P) - \varepsilon$. From elementary measure theory, we know there exists a compact set $\tilde{A} \subseteq A$ with $d_{Q(0,R)}(\tilde{A}) \geq \mathbf{m}_{Q(0,R)}(P) - 2\varepsilon$. As the set $\tilde{A}^k \times \mathrm{O}(d)$ is then compact, the continuous function g_P attains a minimum on this set; let us call this minimum γ. Since $\tilde{A}$ avoids P, it follows that $\gamma > 0$.

We will now prove that $\tilde{A}$ also avoids tP whenever t is sufficiently close to 1, say when $|t - 1| < \gamma/(k \cdot \operatorname{diam} P)$. Indeed, for all $x_1, \ldots, x_k \in \tilde{A}$ and all $T \in \mathrm{O}(d)$, by the triangle inequality we have that

$$\begin{aligned}\sum_{j=2}^{k} \|(x_j - x_1) - T(tv_j - tv_1)\| &\geq \sum_{j=2}^{k} \big(\|(x_j - x_1) - T(v_j - v_1)\| - |t-1|\,\|v_j - v_1\|\big)\\ &> \sum_{j=2}^{k} \|(x_j - x_1) - T(v_j - v_1)\| - k \cdot |t-1| \operatorname{diam} P\\ &\geq \gamma - k \cdot |t-1| \operatorname{diam} P,\end{aligned}$$

which is positive if $|t - 1| < \gamma/(k \cdot \operatorname{diam} P)$. In particular, we see that

$$\mathbf{m}_{Q(0,R)}(P,\, t_2 P,\, \ldots,\, t_n P) \geq d_{Q(0,R)}(\tilde{A}) \geq \mathbf{m}_{Q(0,R)}(P) - 2\varepsilon$$

whenever $|t_j - 1| < \gamma/(k \cdot \operatorname{diam} P)$ for $2 \leq j \leq n$. Since $\mathbf{m}_{Q(0,R)}(P,\, t_2 P,\, \ldots,\, t_n P) \leq \mathbf{m}_{Q(0,R)}(P)$ clearly holds, the result follows. □

The proof of this last result actually implies a somewhat stronger and more technical property of the independence density, namely that every configuration P where $\mathbf{m}_{\mathbb{R}^d}$ is *discontinuous* must be a local minimum across the 'discontinuity barrier'. If the configuration P is admissible, then we can also prove a corresponding inequality in the reverse direction and conclude that $\mathbf{m}_{\mathbb{R}^d}$ is in fact continuous at this point; this is done in the next theorem:

Theorem 2.16 (Continuity of the independence density)**.** *For every $n \geq 1$, the function $(P_1, \ldots, P_n) \mapsto \mathbf{m}_{\mathbb{R}^d}(P_1, \ldots, P_n)$ is continuous on the set of n admissible configurations.*

Proof. For the sake of better readability, we will prove the result in the case of only one forbidden configuration; the n-variable version easily follows from the same argument.

Fix some $\varepsilon > 0$, and let R_0, $c > 0$ be the constants promised by the Supersaturation theorem (Theorem 2.10). By our 'equicontinuity' Lemma 2.7, there is some $\delta > 0$ for which the inequality

$$|I_{P'}(A) - I_P(A)| < cR^d \qquad \forall P' \in \mathcal{B}(P, \delta)$$

holds for all $R > 0$ and all measurable sets $A \subset Q(0, R)$. Finally, let $R \geq R_0$ be large enough so that $\mathbf{m}_{\mathbb{R}^d}(P') \leq \mathbf{m}_{Q(0,R)}(P') \leq \mathbf{m}_{\mathbb{R}^d}(P') + \varepsilon$ holds for all $P' \in \mathcal{B}(P, \delta)$ (this value exists by Lemma 2.2).

If $A \subset Q(0, R)$ is a measurable set avoiding P', for some configuration $P' \in \mathcal{B}(P, \delta)$, we conclude that $I_P(A) < cR^d$, and so (by the Supersaturation theorem)

$$d_{Q(0,R)}(A) < \mathbf{m}_{Q(0,R)}(P) + \varepsilon.$$

This immediately implies that

$$\mathbf{m}_{\mathbb{R}^d}(P') \leq \mathbf{m}_{Q(0,R)}(P') \leq \mathbf{m}_{Q(0,R)}(P) + \varepsilon \leq \mathbf{m}_{\mathbb{R}^d}(P) + 2\varepsilon$$

whenever $\|P' - P\|_\infty \leq \delta$.

Let now $A \subset Q(0, R)$ be a *compact* P-avoiding set with density

$$d_{Q(0,R)}(A) \geq \mathbf{m}_{Q(0,R)}(P) - \varepsilon.$$

Proceeding exactly as we did in the proof of the last lemma, we conclude that A also avoids all P' close enough to P, so for such configurations

$$\mathbf{m}_{Q(0,R)}(P') \geq d_{Q(0,R)}(A) \geq \mathbf{m}_{Q(0,R)}(P) - \varepsilon \geq \mathbf{m}_{\mathbb{R}^d}(P) - \varepsilon.$$

Since $\mathbf{m}_{Q(0,R)}(P') \leq \mathbf{m}_{\mathbb{R}^d}(P') + \varepsilon$, this finishes the proof. □

Let us denote by $\mathcal{M}_n(P)$ the set of all possible independence densities one can obtain by forbidding n distinct dilates of a given configuration P, that is

$$\mathcal{M}_n(P) := \{\mathbf{m}_{\mathbb{R}^d}(t_1 P, t_2 P, \ldots, t_n P) : \ 0 < t_1 < t_2 < \cdots < t_n < \infty\}.$$

It is natural to wonder how this set depends on the number of forbidden dilates, for each configuration P; the last results can be combined in a simple way to give an (almost complete) answer to this question in the case of admissible configurations:

Theorem 2.17 (Forbidding multiple dilates)**.** *If $P \subset \mathbb{R}^d$ is admissible, then*

$$(\mathbf{m}_{\mathbb{R}^d}(P)^n,\, \mathbf{m}_{\mathbb{R}^d}(P)) \subseteq \mathcal{M}_n(P) \subseteq [\mathbf{m}_{\mathbb{R}^d}(P)^n,\, \mathbf{m}_{\mathbb{R}^d}(P)].$$

Proof. It is clear that $\mathbf{m}_{\mathbb{R}^d}(t_1P, t_2P, \ldots, t_nP) \leq \mathbf{m}_{\mathbb{R}^d}(t_1P) = \mathbf{m}_{\mathbb{R}^d}(P)$ always holds, and we saw in Lemma 2.13 that

$$\mathbf{m}_{\mathbb{R}^d}(t_1P, t_2P, \ldots, t_nP) \geq \prod_{i=1}^{n} \mathbf{m}_{\mathbb{R}^d}(t_iP) = \mathbf{m}_{\mathbb{R}^d}(P)^n.$$

Moreover, Lemma 2.15 implies that $\mathbf{m}_{\mathbb{R}^d}(P)$ is an accumulation point of the set $\mathcal{M}_n(P)$, and (since P is admissible) Theorem 2.14 implies the same about $\mathbf{m}_{\mathbb{R}^d}(P)^n$. The result now follows from continuity of the function

$$(t_1, t_2, \ldots, t_n) \mapsto \mathbf{m}_{\mathbb{R}^d}(t_1P, t_2P, \ldots, t_nP),$$

which is an immediate consequence of Theorem 2.16. □

As the final result of this chapter, we now show the existence of extremizer measurable sets which avoid admissible configurations. This generalizes a result of Bukh (see Corollary 13 in [8]) from forbidden distances to higher-order configurations.

Theorem 2.18 (Existence of extremizers)**.** *If $P \subset \mathbb{R}^d$ is admissible, then there exists a P-avoiding set $A \subseteq \mathbb{R}^d$ with well-defined density attaining $d(A) = \mathbf{m}_{\mathbb{R}^d}(P)$. The same is true when forbidding multiple admissible configurations $P_1, \ldots, P_n \subset \mathbb{R}^d$.*

Proof. For each integer $i \geq 1$, let $A_i \subseteq Q(0, i)$ be a P-avoiding set with density $d_{Q(0,i)}(A_i) \geq \mathbf{m}_{Q(0,i)}(P) - 2^{-i}$. By restricting to a subsequence if necessary, we may assume that the indicator functions $(\chi_{A_i})_{i\geq 1}$ converge to some function $f \in \mathcal{B}_\infty$ in the weak* topology of $L^\infty(\mathbb{R}^d)$ (where $\mathcal{B}_\infty$ now denotes the unit ball of $L^\infty(\mathbb{R}^d)$).

It is easy to see from the definition that, for any fixed $R > 0$, $(\chi_{A_i\cap Q(0,R)})_{i\geq 1}$ converges to $\chi_{Q(0,R)}f$ in the weak* topology of $L^\infty(Q(0, R))$. By weak* continuity of I_P (Lemma 2.6) we conclude that $I_P(\chi_{Q(0,R)}f) = 0$, which easily implies[3] that $I_P(\operatorname{supp}\chi_{Q(0,R)}f) = 0$. Since $R > 0$ is arbitrary, this in turn implies that $I_P(\operatorname{supp} f) = 0$.

Denoting $A := \operatorname{supp} f$, we will now prove that A has density $d(A) = \mathbf{m}_{\mathbb{R}^d}(P)$. Since $I_P(A) = 0$, by Lemma 2.3 we can then remove a zero-measure subset of A in order to remove all copies of P and conclude the proof.

By Lemma 2.3, it suffices to show that $\liminf_{R\to\infty} d_{Q(0,R)}(A) \geq \mathbf{m}_{\mathbb{R}^d}(P)$. Fix some arbitrary $\varepsilon > 0$ and take $R_0 \geq 2$ large enough so that $(R_0 + 2\operatorname{diam} P)^d < (1 + \varepsilon/4)R_0^d$. For any given $R \geq R_0$, take a P-avoiding set $B_R \subseteq Q(0, R)$ with

$$d_{Q(0,R)}(B_R) > \mathbf{m}_{Q(0,R)}(P) - \varepsilon/4 \geq \mathbf{m}_{\mathbb{R}^d}(P) - \varepsilon/4.$$

[3] For instance we can proceed as in the proof of Lemma 2.9, approximating $I_P(\operatorname{supp}\chi_{Q(0,R)}f)$ by $I_P(B_\varepsilon)$ where $B_\varepsilon := \{x \in Q(0, R) :\ f(x) \geq \varepsilon\}$, and noting that $I_P(B_\varepsilon) \leq \varepsilon^{-k} I_P(\chi_{Q(0,R)}f) = 0$ for all $\varepsilon > 0$.

For all $i \geq R$, define $A_i' := B_R \cup (A_i \setminus Q(0, R + 2\operatorname{diam} P))$; note that A_i' avoids P and

$$\begin{aligned}
\operatorname{vol}(A_i') &= \operatorname{vol}(A_i) - \operatorname{vol}\left(A_i \cap (Q(0, R + 2\operatorname{diam} P) \setminus Q(0, R))\right) \\
&\qquad - \operatorname{vol}(A_i \cap Q(0, R)) + \operatorname{vol}(B_R) \\
&\geq \operatorname{vol}(A_i) - ((R + 2\operatorname{diam} P)^d - R^d) + \operatorname{vol}(B_R) - \operatorname{vol}(A \cap Q(0, R)) \\
&\qquad + \operatorname{vol}(A \cap Q(0, R)) - \operatorname{vol}(A_i \cap Q(0, R)) \\
&\geq (\mathbf{m}_{Q(0,i)}(P) - 2^{-i})i^d - \frac{\varepsilon R^d}{4} + (d_{Q(0,R)}(B_R) - d_{Q(0,R)}(A))R^d \\
&\qquad + \int_{Q(0,R)} (\chi_A(x) - \chi_{A_i}(x))\,dx \\
&\geq (\mathbf{m}_{Q(0,i)}(P) - 2^{-i})i^d - \frac{\varepsilon R^d}{2} + (\mathbf{m}_{\mathbb{R}^d}(P) - d_{Q(0,R)}(A))R^d \\
&\qquad + \int_{Q(0,R)} (f(x) - \chi_{A_i}(x))\,dx.
\end{aligned}$$

Since $\operatorname{vol}(A_i') \leq \mathbf{m}_{Q(0,i)}(P)\, i^d$ for all $i \geq R$ and $\int_{Q(0,R)} (f(x) - \chi_{A_i}(x))\,dx > -\varepsilon$ for all sufficiently large i, we conclude that for large enough i we have

$$d_{Q(0,R)}(A) > \mathbf{m}_{\mathbb{R}^d}(P) - \frac{i^d}{2^i R^d} - \frac{\varepsilon}{2} - \frac{\varepsilon}{R^d} > \mathbf{m}_{\mathbb{R}^d}(P) - \varepsilon,$$

as wished. A similar argument holds for multiple forbidden configurations. □

2.6 Remarks and open problems

Our results leave open the question of what happens when the configurations we forbid are *not* admissible. There are two different reasons for a given configuration $P \subset \mathbb{R}^d$ to not be admissible, so let us examine them separately.

The fist reason is that P is 'degenerate', meaning that its points are affinely dependent. Bourgain [7] showed an example of sets $A_d \subset \mathbb{R}^d$ (for each $d \geq 2$) which have positive density but which avoid arbitrarily large dilates of the degenerate three-point configuration $\{-1, 0, 1\}$; these sets then show that the conclusion of Bourgain's theorem (and thus also the conclusion of our Theorem 2.14) is false for this degenerate configuration.

This counterexample was later generalized by Graham [39], who showed that a result like Bourgain's theorem *can only hold* if P is contained on the surface of some sphere of finite radius (as is always the case when P is non-degenerate). In fact, Graham's result implies (for instance) that

$$\mathbf{m}_{\mathbb{R}^d}(P,\ \sqrt{3}P,\ \sqrt{5}P,\ \sqrt{7}P,\ \ldots) > 0$$

whenever $P \subset \mathbb{R}^d$ is nonspherical, that is, not contained on the surface of any sphere. Some kind of non-degeneracy hypothesis is thus necessary both for Bourgain's theorem and for our 'asymptotic independence' Theorem 2.14.

It is interesting to note, however, that a more recent result of Ziegler [70, 71] (generalizing a theorem of Furstenberg, Katznelson and Weiss [37] for three-point configurations) shows that every set $A \subseteq \mathbb{R}^d$ of positive upper density is *arbitrarily close* to containing all large enough dilates of *any* finite configuration $P \subset \mathbb{R}^d$. More precisely, denoting by A_δ the set of all points at distance at most δ from the set A, Ziegler proved the following:

Theorem 2.19 (Ziegler [71])**.** *Let $A \subseteq \mathbb{R}^d$ be a set of positive upper density and $P \subset \mathbb{R}^d$ be a finite set. Then there exists $t_0 > 0$ such that, for any $t \geq t_0$ and any $\delta > 0$, the set A_δ contains a configuration congruent to tP.*

Let us now turn to the second reason for a configuration P in $\mathbb{R}^d$ to be non-admissible, namely that it contains $d + 1$ points (if it has more than $d + 1$ points then it is obviously degenerate). In this case we cannot apply the same strategy we used to prove the Counting lemma, and it is not clear whether the analogues of this result or of Bourgain's theorem are true. We conjecture that they are whenever $d \geq 2$, so that we can remove the cardinality condition from the statement of Bourgain's result and of Theorem 2.14.

In particular, let us make more explicit the simplest case of this conjecture, which is an obvious question left open since the results of Bourgain and of Furstenberg, Katznelson and Weiss:

Conjecture 2.20. *Let $A \subset \mathbb{R}^2$ be a set of positive upper density and let $u, v, w \in \mathbb{R}^2$ be non-collinear points. Then there exists $t_0 > 0$ such that for any $t \geq t_0$ the set A contains a configuration congruent to $\{tu, tv, tw\}$.*

Chapter Three

Configuration-avoiding sets on the sphere

In this chapter, we turn to the question of whether the methods and results shown in the Euclidean space setting can also be made to work in the spherical setting. We again follow the author's paper "Geometrical sets with forbidden configurations" [10].

Basic definitions and notation

We shall fix an integer $d \geq 2$ throughout the chapter and work on the d-dimensional unit sphere $S^d \subset \mathbb{R}^{d+1}$. We denote the uniform probability measure on S^d by $\sigma^{(d)} = \sigma$, and the normalized Haar measure on $\mathrm{O}(d+1)$ by $\mu_{d+1} = \mu$.

The analogue of the axis-parallel cube in the spherical setting will be the *spherical cap*: given $x \in S^d$ and $\rho > 0$, we denote[1]

$$\mathrm{Cap}(x, \rho) := \{y \in S^d : \|x - y\|_{\mathbb{R}^{d+1}} \leq \rho\}.$$

We say that $\mathrm{Cap}(x, \rho)$ is the *spherical cap with center x and radius ρ*. Since its measure $\sigma(\mathrm{Cap}(x, \rho))$ does not depend on the center point x, we shall denote this value simply by $\sigma(\mathrm{Cap}_\rho)$. For a given (measurable) set $A \subseteq S^d$ we then write

$$d_{\mathrm{Cap}(x,\rho)}(A) := \frac{\sigma(A \cap \mathrm{Cap}(x, \rho))}{\sigma(\mathrm{Cap}_\rho)}$$

for the density of A inside this cap.

We define a (spherical) configuration on S^d as a finite subset of $\mathbb{R}^{d+1}$ which *is congruent to a set on S^d*; it is convenient to allow for configurations that are not necessarily on the sphere in order to consider dilations. Note that, if $P, Q \subset S^d$ are two configurations which *are* on the sphere, then $P \simeq Q$ if and only if there is a transformation $T \in \mathrm{O}(d+1)$ for which $P = T \cdot Q$ (translations are no longer necessary in this case).

[1] It is more customary to define the spherical cap using *angular distance* instead of Euclidean distance as we use. There is no meaningful (qualitative) difference between these two choices, but the use of the Euclidean distance will be more convenient for us.

A spherical configuration P on S^d is said to be *admissible* if it has at most d points and if it is congruent to a collection $P' \subset S^d$ which is *linearly independent*.[2] As before, we shall say that some set $A \subseteq S^d$ *avoids* P if there is no subset of A which is congruent to P.

3.1 Independence density and the counting function

The natural analogues of the independence density in the spherical setting can now be formally defined.

It is again important in our arguments to have two notions of independence density, one global (on the whole sphere S^d) and one local (for spherical caps). Given $n \geq 1$ configurations $P_1, \dots, P_n$ on S^d, we then define the quantities

$$\mathbf{m}_{S^d}(P_1, \dots, P_n) := \sup\{\, \sigma(A) : A \subset S^d \text{ avoids } P_i,\ 1 \leq i \leq n \,\} \text{ and}$$
$$\mathbf{m}_{\mathrm{Cap}(x,\rho)}(P_1, \dots, P_n) := \sup\{\, d_{\mathrm{Cap}(x,\rho)}(A) : A \subset \mathrm{Cap}(x, \rho) \text{ avoids } P_i,\ 1 \leq i \leq n \,\}.$$

Whenever convenient we will enunciate and prove results in the case of only one forbidden configuration, as the more general case of multiple forbidden configurations follows from the same arguments with only trivial modifications.

The first issue we encounter in the spherical setting is that it is not compatible with dilations: given a collection of points $P \subset S^d$ and some dilation parameter $t > 0$, it is usually *not* true that there exists a collection $Q \subset S^d$ congruent to tP. However, there is a large class of configurations (including the ones we call admissible) for which this *is* true whenever $0 < t \leq 1$; we shall say that they are *contractible*.

It is easy to show that any configuration $P \subset S^d$ having at most $d + 1$ points is contractible. Indeed, these points will all be contained in a d-dimensional affine hyperplane $\mathcal{H} \subset \mathbb{R}^{d+1}$; let $w \in \mathbb{R}^{d+1}$ be a normal vector to $\mathcal{H}$ and consider translations $sw + \mathcal{H}$ of this hyperplane in the direction of w. By elementary geometry, for any given $0 < t \leq 1$ we can find some parameter $s \geq 0$ for which the collection of points in $(sw + \mathcal{H}) \cap S^d$ which are closest to $sw + P$ is congruent to tP.

Even when the configuration we are considering is contractible, however, there is no easy relationship between the independence densities of its distinct dilates. We will then start with the following reassuring lemma, which in a sense assures us the results we wish to obtain aren't true for only trivial reasons.

Lemma 3.1. *For any fixed contractible configuration $P \subset S^d$, we have that*

$$\inf_{0<t\leq 1} \mathbf{m}_{S^d}(tP) > 0 \quad \textit{and} \quad \sup_{0<t\leq 1} \mathbf{m}_{S^d}(tP) < 1.$$

Proof. Denote by $\delta_{S^d}(\gamma)$ the packing density of S^d by caps of radius γ, i.e. the largest possible density of a collection of interior-disjoint caps each having radius γ. It is clear that $\delta_{S^d}(\gamma) \geq \sigma(\mathrm{Cap}_\gamma)$ is bounded away from zero when γ is bounded away from zero, and it is well-known that $\delta_{S^d}(\gamma)$ tends to the sphere packing density of $\mathbb{R}^d$ when $\gamma \to 0$. In particular, $\inf_{0<\gamma\leq 2} \delta_{S^d}(\gamma) > 0$.

[2] Note that this definition is different from the one in the Euclidean setting, where we required the points to be affinely independent instead of linearly independent. The reason behind this difference is that the Euclidean space is translation-invariant while the sphere is not, so affine properties on $\mathbb{R}^d$ translate to linear properties on S^d.

For any given $0 < t \leq 1$, let $\mathcal{P}_t$ denote the centers of caps on an (arbitrary) optimal cap packing of radius $\operatorname{diam} tP$, and define the set

$$A_t := \bigcup_{x \in \mathcal{P}_t} \operatorname{Cap}\Big(x, \frac{\operatorname{diam} tP}{4}\Big).$$

It is easy to see that A_t does not contain any copy of tP. Moreover, since the inequality $\sigma(\operatorname{Cap}_{\rho/4}) \geq c_d \sigma(\operatorname{Cap}_\rho)$ holds for some $c_d > 0$ and all $0 < \rho \leq 2$, we conclude that

$$\inf_{0<t\leq 1} \mathbf{m}_{S^d}(tP) \geq \inf_{0<t\leq 1} \sigma(A_t) \geq \inf_{0<\gamma\leq 2} c_d \delta_{S^d}(\gamma) > 0.$$

For the second inequality, suppose $A \subseteq S^d$ avoids $P = \{v_1, \ldots, v_k\}$. Then

$$\sum_{i=1}^{k} \chi_A(Rv_i) = |A \cap RP| \leq k - 1 \qquad \forall R \in \mathrm{O}(d+1).$$

Integrating over $\mathrm{O}(d+1)$, we obtain

$$k\sigma(A) = \int_{\mathrm{O}(d+1)} \left(\sum_{i=1}^{k} \chi_A(Rv_i) \right) d\mu(R) \leq k - 1,$$

implying that $\sigma(A) \leq 1 - 1/k$. Thus $\sup_{0<t\leq 1} \mathbf{m}_{S^d}(tP) \leq 1 - 1/|P|$. $\square$

Given some configuration $P = \{v_1, v_2, \ldots, v_k\} \subset S^d$, we define the *counting function* I_P which acts on a bounded measurable function $f : S^d \to \mathbb{R}$ by

$$I_P(f) := \int_{\mathrm{O}(d+1)} f(Rv_1) f(Rv_2) \cdots f(Rv_k)\, d\mu(R).$$

In the case where f is the indicator function of a set $A \subseteq S^d$, we shall denote $I_P(\chi_A)$ more simply by $I_P(A)$; note that

$$I_P(A) = \mathbb{P}_{R \in \mathrm{O}(d+1)}(Rv_1,\, Rv_2,\, \ldots,\, Rv_k \in A).$$

If the spherical configuration P is *not* a subset of the sphere, we define the function I_P as being equal to I_Q for any $Q \simeq P$ which is contained in S^d.

As in the Euclidean setting, one can show there is no meaningful difference between requiring that a measurable set $A \subseteq S^d$ avoids some configuration P or that it only satisfies $I_P(A) = 0$. This is proven in the next lemma:

Lemma 3.2 (Zero-measure removal)**.** *Suppose $P \subset S^d$ is a finite configuration and $A \subseteq S^d$ is measurable. If $I_P(A) = 0$, then we can remove a zero-measure subset of A in order to remove all congruent copies of P.*

Proof. It will be more convenient to change spaces and work on the orthogonal group $\mathrm{O}(d+1)$ rather than on the sphere S^d. For $\delta > 0$ and $R \in \mathrm{O}(d+1)$, denote by

$$\mathcal{B}(R, \delta) := \{T \in \mathrm{O}(d+1) : \|T - R\|_{2\to 2} \leq \delta\}$$

the ball of radius δ in operator norm centered on R, and let I denote the identity transformation. We will first show that

$$\lim_{\delta\to 0}\left|\frac{1}{\mu(\mathcal{B}(I,\delta))}\int_{\mathcal{B}(I,\delta)}\chi_A(Tx)\,d\mu(T)-\chi_A(x)\right| = 0 \text{ for almost every } x\in S^d. \tag{3.1}$$

Fix some point $e\in S^d$ and define on $\mathrm{O}(d+1)$ the (measurable) set

$$E := \{R\in \mathrm{O}(d+1) : \ Re\in A\}.$$

By Lebesgue's Density theorem on $\mathrm{O}(d+1)$, we have that

$$\lim_{\delta\to 0}\left|\frac{1}{\mu(\mathcal{B}(R,\delta))}\int_{\mathcal{B}(R,\delta)}\chi_E(T)\,d\mu(T)-\chi_E(R)\right| = 0 \text{ for } \mu\text{-a.e. } R\in \mathrm{O}(d+1).$$

But this means exactly that the measure of the set

$$F := \left\{R\in \mathrm{O}(d+1) : \ \lim_{\delta\to 0}\left|\frac{1}{\mu(\mathcal{B}(I,\delta))}\int_{\mathcal{B}(I,\delta)}\chi_A(TRe)\,d\mu(T)-\chi_A(Re)\right|\neq 0\right\}$$

of 'non-density points' is zero. It is clear from the definition of F that it is invariant under the right-action of $\mathrm{Stab}(e)$; this implies that $\sigma(\{Re : \ R\in F\}) = \mu(F) = 0$, proving (3.1).

Now we remove from A all points x for which identity (3.1) does not hold, thus obtaining a subset $B\subseteq A$ with $\sigma(A\setminus B) = 0$ and

$$\lim_{\delta\to 0}\frac{1}{\mu(\mathcal{B}(I,\delta))}\int_{\mathcal{B}(I,\delta)}\chi_B(Tx)\,d\mu(T) = 1 \qquad \text{for all } x\in B.$$

We will show that no copy of P remains on this restricted set B, which will finish the proof of the lemma.

Suppose for contradiction that B contains a copy $\{u_1,\dots,u_k\}$ of P. Then there exists $\delta > 0$ for which

$$\frac{1}{\mu(\mathcal{B}(I,\delta))}\int_{\mathcal{B}(I,\delta)}\chi_B(Tu_i)\,d\mu(T)\geq 1-\frac{1}{2k} \qquad \text{for all } 1\leq i\leq k,$$

which means that $\mathbb{P}_{T\in\mathcal{B}(I,\delta)}(Tu_i\notin B)\leq 1/2k$ for each $1\leq i\leq k$. Thus

$$\begin{aligned}
I_P(B) &= \mathbb{P}_{T\in \mathrm{O}(d+1)}(Tu_1,\ldots,Tu_k\in B)\\
&\geq \mu(\mathcal{B}(I,\delta))\cdot \mathbb{P}_{T\in\mathcal{B}(I,\delta)}(Tu_1,\ldots,Tu_k\in B)\\
&\geq \mu(\mathcal{B}(I,\delta))\left(1-\sum_{i=1}^{k}\mathbb{P}_{T\in\mathcal{B}(I,\delta)}(Tu_i\notin B)\right)\\
&\geq \frac{\mu(\mathcal{B}(I,\delta))}{2} > 0,
\end{aligned}$$

contradicting our assumption that $I_P(A) = 0$. $\square$

3.2 Harmonic analysis on S^d and the Counting lemma

The next thing we need is an analogue of the Counting lemma in the spherical setting, saying we do not significantly change the count of configurations in a given set $A \subseteq S^d$ by blurring this set a little. As in the Euclidean setting, we will use Fourier-analytic methods to prove such a result; we now give a quick overview of the definitions and results we need on harmonic analysis for our arguments.

Given an integer $n \geq 0$, write $\mathscr{H}_n^{d+1}$ for the space of real harmonic polynomials, homogeneous of degree n, on $\mathbb{R}^{d+1}$; that is,

$$\mathscr{H}_n^{d+1} = \left\{ f \in \mathbb{R}[x_1, \ldots, x_{d+1}] : f \text{ homogeneous, } \deg f = n,\ \sum_{i=1}^{d+1} \frac{\partial^2}{\partial x_i^2} f = 0 \right\}.$$

The restriction of the elements of $\mathscr{H}_n^{d+1}$ to S^d are called *spherical harmonics of degree n* on S^d. If $Y \in \mathscr{H}_n^{d+1}$, note that $Y(x) = \|x\|^n Y(x')$ where $x = \|x\|x'$ and $x' \in S^d$; we can then identify $\mathscr{H}_n^{d+1}$ with the space of spherical harmonics of degree n, which by a slight (and common) abuse of notation we also denote $\mathscr{H}_n^{d+1}$.

Harmonic polynomials of different degrees are orthogonal with respect to the standard inner product $\langle f, g\rangle_{S^d} := \int_{S^d} f(x)\, g(x)\, d\omega(x)$. Moreover, it is a well-known fact (see e.g. [16]) that the family of spherical harmonics is dense in $L^2(S^d)$, and so

$$L^2(S^d) = \bigoplus_{n=0}^{\infty} \mathscr{H}_n^{d+1}.$$

Denoting by $\operatorname{proj}_n : L^2(S^d) \to \mathscr{H}_n^{d+1}$ the orthogonal projection onto $\mathscr{H}_n^{d+1}$, this means that $f = \sum_{n=0}^{\infty} \operatorname{proj}_n f$ (with equality in the L^2 sense) for all $f \in L^2(S^d)$.

There is a family $(P_n^d)_{n\geq 0}$ of polynomials on $[-1, 1]$ which is associated to this decomposition. We use the convention that $\deg P_n^d = n$ and $P_n^d(1) = 1$. These polynomials are then uniquely characterized by the following two properties:

(i) for each fixed $y \in S^d$, the function on S^d given by $x \mapsto P_n^d(x \cdot y)$ is in $\mathscr{H}_n^{d+1}$;

(ii) the projection operator $\operatorname{proj}_n : L^2(S^d) \to \mathscr{H}_n^{d+1}$ is given by

$$\operatorname{proj}_n f(x) = \dim \mathscr{H}_n^{d+1} \int_{S^d} P_n^d(x \cdot y) f(y)\, d\omega(y). \tag{3.2}$$

Using these two facts we conclude also the useful property

$$\int_{S^d} P_n^d(x \cdot y) P_n^d(x \cdot z)\, d\omega(x) = \frac{1}{\dim \mathscr{H}_n^{d+1}} P_n^d(y \cdot z) \qquad \forall y, z \in S^d. \tag{3.3}$$

Note that, by orthogonality of the spaces $\mathscr{H}_n^{d+1}$, property (i) implies that

$$\int_{S^d} P_n^d(x \cdot y) P_m^d(x \cdot y)\, d\omega(x) = 0 \quad \text{if } n \neq m$$

(for any fixed $y \in S^d$). Using the change of variables $t = x \cdot y$, this is equivalent to saying that
$$\int_{-1}^{1} P_n^d(t) P_m^d(t)\,(1-t^2)^{(d-2)/2} dt = 0 \quad \text{if } n \neq m.$$
This shows that the polynomials P_n^d are, up to multiplicative constants, the *Gegenbauer polynomials* C_n^λ with parameter $\lambda = (d-1)/2$: $P_n^d(x) = C_n^{(d-1)/2}(x)/C_n^{(d-1)/2}(1)$; we refer the reader to Dai and Xu's book [16] for information on the Gegenbauer polynomials, and for a proof that the P_n^d thus defined indeed satisfy properties (*i*) and (*ii*). The following simple facts about P_n^d follow immediately from the corresponding properties of the Gegenbauer polynomials:

Lemma 3.3. *For all integers $d \geq 2$ and $n \geq 0$ the following hold:*

- $P_n^d(t) \in [-1, 1]$ *for all* $t \in [-1, 1]$*;*
- *For any fixed* $\gamma > 0$, $\max_{t\in[-1+\gamma,\,1-\gamma]} P_n^d(t)$ *tends to zero as* $n \to \infty$.

We will follow Dunkl [26] in defining both the convolution operation on the sphere and the spherical analogue of Fourier coefficients. For this we will need to break a little the symmetry of the sphere and distinguish an (arbitrary) point e on S^d; we think of this point as being the north pole.

For a given $x \in S^d$, we write $\mathcal{M}(S^d; x)$ for the space of Borel regular *zonal measures on S^d with pole at x*, that is, those measures which are invariant under the action of $\mathrm{Stab}^{\mathrm{O}(d+1)}(x)$. The elements of $\mathcal{M}(S^d; e)$ are referred to simply as the *zonal measures*.

Given a function $f \in L^2(S^d)$ and a zonal measure $\nu \in \mathcal{M}(S^d; e)$, we define their *convolution $f * \nu$* by
$$f * \nu(x) := \int_{S^d} f(y)\, d\varphi_x \nu(y) \qquad \forall x \in S^d,$$
where $\varphi_x : \mathcal{M}(S^d; e) \to \mathcal{M}(S^d; x)$ is the rotation operator defined by
$$\varphi_x \nu(A) = \nu(T_x^{-1} A) \qquad \text{where } T_x \in \mathrm{O}(d+1) \text{ satisfies } T_x e = x.$$
The value $f * \nu(x)$ can be thought of as the average of f according to a measure which acts with respect to x as ν acts with respect to e. It is easy to see that this operation is well-defined, independently of the choice of T_x: if $S_x e = T_x e = x$, then $S_x^{-1} T_x \in \mathrm{Stab}(e)$ and so $\nu(S_x^{-1} A) = \nu((S_x^{-1} T_x) T_x^{-1} A) = \nu(T_x^{-1} A)$.

For an integer $n \geq 0$ and a zonal measure $\nu \in \mathcal{M}(S^d; e)$, we define its *n-th Fourier coefficient* $\widehat{\nu}_n$ by
$$\widehat{\nu}_n = \int_{S^d} P_n^d(e \cdot y)\, d\nu(y).$$
The main property we will need of Fourier coefficients is the following result, which is stated in Dunkl's paper [26] and can be proven using a straightforward modification of the methods exposed in Chapter 2 of Dai and Xu's book [16]:

Theorem 3.4. *If $f \in L^2(S^d)$ and $\nu \in \mathcal{M}(S^d; e)$, then $f * \nu \in L^2(S^d)$ and*
$$\mathit{proj}_n(f * \nu) = \widehat{\nu}_n\, \mathit{proj}_n f \qquad \forall n \geq 0.$$

With this we finish our review of harmonic analysis on the sphere, so let us return to our specific problem. For a given $\delta > 0$, denote by cap_δ the uniform probability measure on the spherical cap $\mathrm{Cap}(e, \delta)$:

$$\mathrm{cap}_\delta(A) = \frac{\sigma(A \cap \mathrm{Cap}(e, \delta))}{\sigma(\mathrm{Cap}(e, \delta))} \qquad \forall A \subseteq S^d.$$

Note that each cap_δ is a zonal measure; one immediately checks that

$$(\widehat{\mathrm{cap}}_\delta)_n = \frac{1}{\sigma(\mathrm{Cap}_\delta)} \int_{\mathrm{Cap}(e,\,\delta)} P_n^d(e \cdot y)\, d\sigma(y)$$

for all $n \geq 0$, and

$$f * \mathrm{cap}_\delta(x) = \frac{1}{\sigma(\mathrm{Cap}_\delta)} \int_{\mathrm{Cap}(x,\,\delta)} f(y)\, d\sigma(y)$$

for all $f \in L^2(S^d)$. In particular, if $A \subseteq S^d$ is a measurable set, then $\chi_A * \mathrm{cap}_\delta(x) = d_{\mathrm{Cap}(x,\,\delta)}(A)$; this gives the 'blurring' of the spherical sets we shall consider.

Lemma 3.5. *For every $d \geq 2$ and $\gamma > 0$, there exists a function $c_{d,\gamma} : (0, 1] \to \mathbb{R}$ with $\lim_{\delta \to 0^+} c_{d,\gamma}(\delta) = 0$ such that the following holds. For all $f, g \in L^2(S^d)$ and all points $u, v \in S^d$ with $u \cdot v \in [-1 + \gamma,\, 1 - \gamma]$, we have that*

$$\left| \int_{\mathrm{O}(d+1)} f(Ru)\, (g(Rv) - g * \mathrm{cap}_\delta(Rv))\, d\mu(R) \right| \leq c_{d,\gamma}(\delta)\, \|f\|_2 \|g\|_2.$$

Proof. Denote by $\tilde{\mu}_e$ the Haar measure on $\mathrm{Stab}(e)$, and assume without loss of generality that u coincides with the north pole e. By symmetry, the expression we wish to bound may then be written as

$$\left| \int_{\mathrm{O}(d+1)} f(Re) h(Rv)\, d\mu(R) \right| = \left| \int_{\mathrm{O}(d+1)} f(Re) \Big(\int_{\mathrm{Stab}(e)} h(RSv)\, d\tilde{\mu}_e(S) \Big) d\mu(R) \right|,$$

where $h = g - g * \mathrm{cap}_\delta$.

Write $t_0 := e \cdot v$. Note that, when $S \in \mathrm{Stab}(e)$ is distributed uniformly according to $\tilde{\mu}_e$, the point Sv is uniformly distributed on $S_{t_0}^{d-1} := \{y \in S^d : e \cdot y = t_0\}$. Denote by $\sigma_{t_0}^{(d-1)}$ the uniform probability measure on $S_{t_0}^{d-1}$ (that is, the unique one which is invariant under the action of $\mathrm{Stab}(e)$).

Making the change of variables $y = Sv$, we see that

$$\begin{aligned} \int_{\mathrm{Stab}(e)} h(RSv)\, d\tilde{\mu}_e(S) &= \int_{S_{t_0}^{d-1}} h(Ry)\, d\sigma_{t_0}^{(d-1)}(y) \\ &= \int_{S_{t_0}^{d-1}} h(z)\, d\sigma_{t_0}^{(d-1)}(R^{-1}z) = h * \sigma_{t_0}^{(d-1)}(Re). \end{aligned} \tag{3.4}$$

The expression we wish to bound is then equal to

$$\left| \int_{\mathrm{O}(d+1)} f(Re)\, h * \sigma_{t_0}^{(d-1)}(Re)\, d\mu(R) \right| = \left| \int_{S^d} f(x)\, h * \sigma_{t_0}^{(d-1)}(x)\, d\sigma(x) \right|.$$

Using Parseval's identity, we can rewrite the right-hand side of this last equality as

$$\Big| \sum_{n=0}^{\infty} \int_{S^d} \operatorname{proj}_n f(x) \, \operatorname{proj}_n(h * \sigma_{t_0}^{(d-1)})(x) \, d\sigma(x) \Big|$$
$$\leq \sum_{n=0}^{\infty} \int_{S^d} |\operatorname{proj}_n f(x)| \, |(\widehat{\sigma_{t_0}^{(d-1)}})_n| \, |\operatorname{proj}_n h(x)| \, d\sigma(x)$$
$$\leq \sum_{n=0}^{\infty} |(\widehat{\sigma_{t_0}^{(d-1)}})_n| \, \|\operatorname{proj}_n f\|_2 \, \|\operatorname{proj}_n h\|_2.$$

As $h = g - g * \mathrm{cap}_\delta$, the expression above is equal to

$$\sum_{n=0}^{\infty} |(\widehat{\sigma_{t_0}^{(d-1)}})_n| \, |1 - (\widehat{\mathrm{cap}}_\delta)_n| \, \|\operatorname{proj}_n f\|_2 \, \|\operatorname{proj}_n g\|_2$$
$$= \sum_{n=0}^{\infty} |P_n^d(t_0)| \left| 1 - \frac{1}{\sigma(\mathrm{Cap}_\delta)} \int_{\mathrm{Cap}(e,\,\delta)} P_n^d(e \cdot y) \, d\sigma(y) \right| \|\operatorname{proj}_n f\|_2 \, \|\operatorname{proj}_n g\|_2.$$

Fix some $\varepsilon > 0$. Since $t_0 \in [-1+\gamma,\ 1-\gamma]$ (by hypothesis), from Lemma 3.3 we obtain that $|P_n^d(t_0)| \leq \varepsilon/2$ holds for all $n \geq N(\varepsilon, \gamma)$, while

$$\left| 1 - \frac{1}{\sigma(\mathrm{Cap}_\delta)} \int_{\mathrm{Cap}(e,\,\delta)} P_n^d(e \cdot y) \, d\sigma(y) \right| \leq \max_{-1 \leq t \leq 1} \left| 1 - P_n^d(t) \right| = 2$$

always holds. Moreover, since each P_n^d is a polynomial satisfying $P_n^d(1) = 1$, we can choose $\delta_0 = \delta_0(\varepsilon, \gamma) > 0$ small enough so that $|1 - P_n^d(e \cdot y)| \leq \varepsilon$ holds whenever $n < N(\varepsilon, \gamma)$ and $y \in \mathrm{Cap}(e,\ \delta_0)$. This implies that the last sum is at most

$$\sum_{n=0}^{\infty} \varepsilon \| \operatorname{proj}_n f\|_2 \, \|\operatorname{proj}_n g\|_2 \leq \varepsilon \|f\|_2 \|g\|_2$$

whenever $\delta \leq \delta_0(\varepsilon, \gamma)$, finishing the proof. □

Recall that a spherical configuration P on S^d is admissible if it has at most d points and if it is congruent to a collection $P' \subset S^d$ which is linearly independent. We can now give the spherical counterpart to the Counting lemma from last chapter:

Lemma 3.6 (Counting lemma)**.** *For every admissible configuration P on S^d there exists a function $\eta_P : (0,1] \to (0,1]$ with $\lim_{\delta \to 0^+} \eta_P(\delta) = 0$ such that the following holds for all measurable sets $A \subseteq S^d$:*

$$|I_P(A) - I_P(\chi_A * \mathrm{cap}_\delta)| \leq \eta_P(\delta) \qquad \forall \delta \in (0,1].$$

Moreover, this upper bound function η_P can be made uniform inside a small ball $\mathcal{B} \subset (\mathbb{R}^{d+1})^k$ centered on the configuration P considered.

Proof. Up to congruence, we may assume $P \subset S^d$. Similarly to what we did in the Euclidean setting, we will first obtain a uniform upper bound for

$$\left| \int_{\mathrm{O}(d+1)} f_1(Tv_1) \cdots f_{k-1}(Tv_{k-1}) \left(f_k(Tv_k) - f_k * \mathrm{cap}_\delta(Tv_k) \right) d\mu(T) \right|,$$

valid whenever $0 \le f_1, \ldots, f_k \le 1$ are measurable functions and $(v_1, v_2, \ldots, v_k)$ is a permutation of the points of P.

Denote by $G := \mathrm{Stab}^{\mathrm{O}(d+1)}(v_1, \ldots, v_{k-2})$ the stabilizer of the first $k-2$ points of P, and by $H := \mathrm{Stab}^{\mathrm{O}(d+1)}(v_1, \ldots, v_{k-2}, v_{k-1}) = \mathrm{Stab}^G(v_{k-1})$ the stabilizer of the first $k-1$ points of P. We can then bound the expression above by

$$\int_{\mathrm{O}(d+1)} \left| \int_G f_{k-1}(TSv_{k-1}) \left(f_k(TSv_k) - f_k * \mathrm{cap}_\delta(TSv_k) \right) d\mu_G(S) \right| d\mu(T), \qquad (3.5)$$

where μ_G denotes the normalized Haar measure on G.

Denote $\ell := d - k + 2 \ge 2$. Since P is non-degenerate, we see that $G \simeq \mathrm{O}(\ell + 1)$ and that both Gv_{k-1} and Gv_k are spheres of dimension ℓ. Morally, we should then be able to apply the last lemma (with $d = \ell$, $f = f_{k-1}(T\cdot)$ and $g = f_k(T\cdot)$) and easily conclude. However, the convolution in expression (3.5) above happens in S^d, while that on the last lemma would happen in S^ℓ; in particular, if $k \ge 3$ so that $\ell < d$, all of the mass on the average defined by the convolution in (3.5) lies *outside* of the ℓ-dimensional sphere Gv_k, so this argument cannot work. We will have to work harder to conclude.

Note that, since Gv_k is an ℓ-dimensional sphere while Hv_k is an $(\ell - 1)$-dimensional sphere (which happens because P is non-degenerate), it follows that there is a point $\xi \in Gv_k$ which is fixed by H; this point will work as the north pole of Gv_k.

It will be more convenient to work on the canonical unit sphere S^ℓ instead of the ℓ-dimensional sphere $Gv_k \subset S^d$. We shall then restrict ourselves to the $(\ell + 1)$-dimensional affine hyperplane $\mathcal{H}$ determined by $\mathcal{H} \cap S^d = Gv_k$, and place coordinates on it to identify $\mathcal{H}$ with $\mathbb{R}^{\ell+1}$ and Gv_k with S^ℓ, noting that G then acts as $\mathrm{O}(\ell + 1)$. More formally, let $r > 0$ be the radius of Gv_k in $\mathbb{R}^{d+1}$, so that Gv_k is isometric to rS^ℓ; take such an isometry $\psi : Gv_k \to rS^\ell$, and define $e \in S^\ell$ by $e := \psi(\xi)/r$. Now we construct a map $\phi : G \to \mathrm{O}(\ell+1)$ defined by

$$\phi(S)\psi(x) = \psi(Sx) \qquad \forall x \in Gv_k$$

for each $S \in G$. It is easy to check that this map is well-defined and gives an isomorphism between G and $\mathrm{O}(\ell + 1)$ satisfying $\phi(H) = \mathrm{Stab}^{\mathrm{O}(\ell+1)}(e)$.

For each fixed $T \in \mathrm{O}(d + 1)$, define the functions $g_T, h_T : S^\ell \to [-1, 1]$ by

$$\begin{aligned} g_T(Re) &:= f_{k-1}(T\phi^{-1}(R)v_{k-1}) \quad \text{and} \\ h_T(Re) &:= f_k(T\phi^{-1}(R)\xi) - f_k * \mathrm{cap}_\delta(T\phi^{-1}(R)\xi), \end{aligned}$$

for all $R \in \mathrm{O}(\ell + 1)$. These functions are indeed well-defined on S^ℓ, since $\mathrm{Stab}^G(v_{k-1}) = \mathrm{Stab}^G(\xi) = \phi^{-1}(\mathrm{Stab}^{\mathrm{O}(\ell+1)}(e))$. Note that h_T can also be written as a function of $x \in S^\ell$ by making use of the isometry $\psi^{-1} : rS^\ell \to Gv_k$:

$$h_T(x) = f_k(T\psi^{-1}(rx)) - f_k * \mathrm{cap}_\delta(T\psi^{-1}(rx)).$$

Denote by $u := \psi(v_k)/r$ the point in S^ℓ corresponding to v_k. Making the change of variables $R = \phi(S)$, we obtain

$$\begin{aligned}\int_G f_{k-1}(TS v_{k-1})\big(f_k(TS v_k) - f_k * \mathrm{cap}_\delta(TS v_k)\big)\, d\mu_G(S)\\ = \int_{\mathrm{O}(\ell+1)} g_T(Re)\, h_T(Ru)\, d\mu_{\ell+1}(R)\\ = \int_{\mathrm{O}(\ell+1)} g_T(Re)\Big(\int_{\mathrm{Stab}(e)} h_T(RS u)\, d\tilde{\mu}_e(S)\Big)\, d\mu_{\ell+1}(R),\end{aligned}$$

where we write $\mathrm{Stab}(e)$ for $\mathrm{Stab}^{\mathrm{O}(\ell+1)}(e)$ and $\tilde{\mu}_e$ for its Haar measure. Working as we did in the chain of equalities (3.4), we see that the expression in parenthesis is equal to $h_T * \sigma^{(\ell-1)}_{e\cdot u}(Re)$, where $\sigma^{(\ell-1)}_{e\cdot u}$ is the uniform probability measure on the $(\ell-1)$-sphere $\mathrm{Stab}(e)u = \{y \in S^\ell : e\cdot y = e\cdot u\}$ (and the convolution now takes place in S^ℓ with e as the north pole). Making the change of variables $x = Re$, we then see that the expression above is equal to

$$\int_{\mathrm{O}(\ell+1)} g_T(Re)\, h_T * \sigma^{(\ell-1)}_{e\cdot u}(Re)\, d\mu_{\ell+1}(R) = \int_{S^\ell} g_T(x)\, h_T * \sigma^{(\ell-1)}_{e\cdot u}(x)\, d\sigma^{(\ell)}(x).$$

We conclude that the expression (3.5) we wish to bound is equal to

$$\begin{aligned}\int_{\mathrm{O}(d+1)}\Big|\int_{S^\ell} g_T(x)\, h_T * \sigma^{(\ell-1)}_{e\cdot u}(x)\, d\sigma^{(\ell)}(x)\Big|\, d\mu_{d+1}(T)\\ \le \int_{\mathrm{O}(d+1)} \|h_T * \sigma^{(\ell-1)}_{e\cdot u}\|_2\, d\mu_{d+1}(T)\\ \le \Big(\int_{\mathrm{O}(d+1)} \|h_T * \sigma^{(\ell-1)}_{e\cdot u}\|_2^2\, d\mu_{d+1}(T)\Big)^{1/2}.\end{aligned}$$

Let us now compute $e\cdot u$, which will be necessary for bounding $\|h_T * \sigma^{(\ell-1)}_{e\cdot u}\|_2^2$. From the identity

$$\|re - ru\|^2_{\mathbb{R}^{\ell+1}} = \|\psi^{-1}(re) - \psi^{-1}(ru)\|^2_{\mathbb{R}^{d+1}} = \|\xi - v_k\|^2_{\mathbb{R}^{d+1}},$$

we conclude that $r^2(2 - 2\, e\cdot u) = 2 - 2\,\xi\cdot v_k$, and so

$$e\cdot u = (\xi\cdot v_k - (1 - r^2))/r^2 \notin \{-1,\ 1\}$$

depends only on P and not our later choices.

Now fix an arbitrary $\varepsilon > 0$. By Parseval's identity, we have that

$$\begin{aligned}\|h_T * \sigma^{(\ell-1)}_{e\cdot u}\|_2^2 &= \sum_{n=0}^{\infty} \|\operatorname{proj}_n(h_T * \sigma^{(\ell-1)}_{e\cdot u})\|_2^2\\ &= \sum_{n=0}^{\infty} |(\widehat{\sigma}^{(\ell-1)}_{e\cdot u})_n|^2\, \|\operatorname{proj}_n h_T\|_2^2\\ &= \sum_{n=0}^{\infty} P_n^\ell(e\cdot u)^2\, \|\operatorname{proj}_n h_T\|_2^2.\end{aligned}$$

Since $e \cdot u \notin \{-1,\ 1\}$ is a constant depending only on P, there exists[3] $N = N(\varepsilon, P) \in \mathbb{N}$ such that $|P_n^\ell(e \cdot u)| \le \varepsilon$ for all $n > N$. Using also that $-1 \le P_n^\ell(t) \le 1$ for all $-1 \le t \le 1$, we conclude that

$$\|h_T * \sigma_{e\cdot u}^{(\ell-1)}\|_2^2 \le \sum_{n=0}^{N} \|\operatorname{proj}_n h_T\|_2^2 + \sum_{n>N} \varepsilon^2 \|\operatorname{proj}_n h_T\|_2^2.$$

The second term on the right-hand side of the inequality above is bounded by $\varepsilon^2\|h_T\|_2^2 \le \varepsilon^2$, so let us concentrate on the first term.

By identities (3.2) and (3.3), we have

$$\begin{aligned}
\|\operatorname{proj}_n h_T\|_2^2 &= \int_{S^\ell} \Big(\dim \mathscr{H}_n^{\ell+1} \int_{S^\ell} h_T(y) P_n^\ell(x\cdot y)\, d\sigma(y)\Big)^2\, d\sigma(x) \\
&= (\dim \mathscr{H}_n^{\ell+1})^2 \int_{S^\ell}\int_{S^\ell} h_T(y)h_T(z) \Big(\int_{S^\ell} P_n^\ell(x\cdot y)P_n^\ell(x\cdot z)\, d\sigma(x)\Big) d\sigma(y)\, d\sigma(z) \\
&= \dim \mathscr{H}_n^{\ell+1} \int_{S^\ell}\int_{S^\ell} h_T(y)h_T(z)\, P_n^\ell(y\cdot z)\, d\sigma(y)\, d\sigma(z).
\end{aligned}$$

Since $|P_n^\ell(y \cdot z)| \le 1$ for all $y, z \in S^\ell$, we conclude that

$$\begin{aligned}
&\int_{\mathrm{O}(d+1)} \|\operatorname{proj}_n h_T\|_2^2\, d\mu_{d+1}(T) \\
&\qquad = \dim \mathscr{H}_n^{\ell+1} \int_{S^\ell}\int_{S^\ell} \Big(\int_{\mathrm{O}(d+1)} h_T(y)h_T(z)\, d\mu_{d+1}(T)\Big) P_n^\ell(y\cdot z)\, d\sigma(y)\, d\sigma(z) \\
&\qquad \le \dim \mathscr{H}_n^{\ell+1} \int_{S^\ell}\int_{S^\ell} \Big|\int_{\mathrm{O}(d+1)} h_T(y)h_T(z)\, d\mu_{d+1}(T)\Big|\, d\sigma(y)\, d\sigma(z).
\end{aligned}$$

We now divide this last double integral on the sphere into two parts, depending on whether or not $y \cdot z$ is close to the extremal points 1 or -1. Thus, for some parameter $0 < \gamma < 1$ to be chosen later, we write the double integral as

$$\begin{aligned}
&\int_{S^\ell}\int_{S^\ell} \Big|\int_{\mathrm{O}(d+1)} h_T(y)h_T(z)\, d\mu_{d+1}(T)\Big|\, \mathbb{1}\{|y\cdot z| > 1-\gamma\}\, d\sigma(y)\, d\sigma(z) \\
&\qquad + \int_{S^\ell}\int_{S^\ell} \Big|\int_{\mathrm{O}(d+1)} h_T(y)h_T(z)\, d\mu_{d+1}(T)\Big|\, \mathbb{1}\{|y\cdot z| \le 1-\gamma\}\, d\sigma(y)\, d\sigma(z).
\end{aligned}$$

Since $-1 \le h_T \le 1$, the first term is at most

$$2\int_{S^\ell}\int_{S^\ell} \mathbb{1}\{y\cdot z > 1-\gamma\}\, d\sigma(y)\, d\sigma(z) = 2\sigma^{(\ell)}(\mathrm{Cap}_{S^\ell}(e,\ \sqrt{2\gamma})).$$

To bound the second term, note that for fixed $y, z \in S^\ell$ we have

$$\begin{aligned}
&\int_{\mathrm{O}(d+1)} h_T(y)h_T(z)\, d\mu_{d+1}(T) \\
&\qquad = \int_{\mathrm{O}(d+1)} \big(f_k(T\tilde{y}) - f_k * \mathrm{cap}_\delta(T\tilde{y})\big)\big(f_k(T\tilde{z}) - f_k * \mathrm{cap}_\delta(T\tilde{z})\big)\, d\mu_{d+1}(T),
\end{aligned}$$

[3] By Lemma 3.3 this value of N can be made robust to small perturbations of the value $e\cdot u$, which is equivalent to small perturbations of the configuration P. This remark, and others in the same vein, are the reason why the bound obtained in the proof can be made to hold uniformly inside small neighborhoods of the considered configuration.

where $\tilde{y} := \psi^{-1}(ry)$ and $\tilde{z} := \psi^{-1}(rz)$. Moreover, we have

$$\|ry - rz\|_{\mathbb{R}^{\ell+1}}^2 = \|\tilde{y} - \tilde{z}\|_{\mathbb{R}^{d+1}}^2 \implies \tilde{y} \cdot \tilde{z} = 1 - r^2(1 - y \cdot z);$$

thus, whenever $y \cdot z \in [-1+\gamma,\ 1-\gamma]$, we have $\tilde{y} \cdot \tilde{z} \in [-1 + r^2\gamma,\ 1 - r^2\gamma]$. Using Lemma 3.5 (with $f = f_k - f_k * \mathrm{cap}_\delta$, $g = f_k$ and γ substituted by $r^2\gamma$) we conclude that the second term is bounded by $c_{d,r^2\gamma}(\delta)$.

Taking stock of everything, we obtain

$$\int_{\mathrm{O}(d+1)} \|h_T * \sigma_{e\cdot u}^{(\ell-1)}\|_2^2 \, d\mu_{d+1}(T)$$
$$\leq \varepsilon^2 + \sum_{n=0}^{N} \dim \mathscr{H}_n^{\ell+1} \left(2\sigma^{(\ell)}(\mathrm{Cap}_{S^\ell}(e,\ \sqrt{2\gamma})) + c_{d,r^2\gamma}(\delta)\right)$$

for any $0 < \gamma < 1$. Choosing γ small enough depending on ℓ, ε and N, and then choosing δ small enough depending on d, $r^2\gamma$, ε and N (so ultimately only on ε and P), we can bound the right-hand side by $4\varepsilon^2$; the expression (3.5) is then bounded by 2ε in this case.

For such small values of δ, we thus conclude from our telescoping sum trick (explained in Section 2.2) that $|I_P(A) - I_P(\chi_A * \mathrm{cap}_\delta)| \leq 2k\varepsilon$, proving the desired inequality since $\varepsilon > 0$ is arbitrary. The claim that the upper bound can be made uniform inside a small ball centered on P follows from analyzing our proof. □

3.3 Continuity properties of the counting function

Following the same script as in the Euclidean setting, we now consider other ways in which the counting function is robust to small perturbations.

It is again easy to show, using our telescoping sum trick, that I_P is continuous in $L^\infty(S^d)$ (and even in $L^{|P|}(S^d)$) for all spherical configurations. When the configuration considered is admissible, we obtain also the following significantly stronger continuity property of I_P when restricting to bounded functions:

Lemma 3.7 (Weak* continuity)**.** *If P is an admissible configuration on S^d, then I_P is weak* continuous on the unit ball of $L^\infty(S^d)$.*

Proof. Denote the closed unit ball of $L^\infty(S^d)$ by $\mathcal{B}_\infty$, and let $(f_i)_{i\geq 1} \subset \mathcal{B}_\infty$ be a sequence weak* converging to $f \in \mathcal{B}_\infty$. It will suffice to show that $(I_P(f_i))_{i\geq 1}$ converges to $I_P(f)$.

Note that, for every $x \in S^d$, $\delta > 0$, we have

$$f_i * \mathrm{cap}_\delta(x) = \frac{1}{\sigma(\mathrm{Cap}_\delta)} \int_{\mathrm{Cap}(x,\,\delta)} f_i(y)\, d\sigma(y)$$
$$\xrightarrow{i\to\infty} \frac{1}{\sigma(\mathrm{Cap}_\delta)} \int_{\mathrm{Cap}(x,\,\delta)} f(y)\, d\sigma(y) = f * \mathrm{cap}_\delta(x).$$

Since $f * \mathrm{cap}_\delta$ and each $f_i * \mathrm{cap}_\delta$ are Lipschitz with the same constant (depending only on δ) and S^d is compact, this easily implies that

$$\|f_i * \mathrm{cap}_\delta - f * \mathrm{cap}_\delta\|_\infty \to 0 \quad \text{as } i \to \infty.$$

In particular, we conclude $\lim_{i\to\infty} I_P(f_i * \mathrm{cap}_\delta) = I_P(f * \mathrm{cap}_\delta)$.

Since P is admissible, by the spherical Counting lemma we have

$$|I_P(f * \mathrm{cap}_\delta) - I_P(f)| \le \eta_P(\delta) \quad \text{and} \quad |I_P(f_i * \mathrm{cap}_\delta) - I_P(f_i)| \le \eta_P(\delta) \quad \forall i \ge 1.$$

Choosing $i_0(\delta) \ge 1$ sufficiently large so that

$$|I_P(f_i * \mathrm{cap}_\delta) - I_P(f * \mathrm{cap}_\delta)| \le \eta_P(\delta) \qquad \forall i \ge i_0(\delta),$$

we conclude that

$$\begin{aligned} |I_P(f) - I_P(f_i)| &\le |I_P(f) - I_P(f * \mathrm{cap}_\delta)| + |I_P(f * \mathrm{cap}_\delta) - I_P(f_i * \mathrm{cap}_\delta)| \\ &\qquad + |I_P(f_i * \mathrm{cap}_\delta) - I_P(f_i)| \\ &\le 3\eta_P(\delta) \qquad \forall i \ge i_0(\delta). \end{aligned}$$

Since $\delta > 0$ is arbitrary and $\eta_P(\delta) \to 0$ as $\delta \to 0$, this finishes the proof. □

Given some spherical configuration $P \subset \mathbb{R}^{d+1}$, fix an arbitrary ordering $(v_1, \dots, v_k)$ of its points and let us write $\mathcal{B}(P, r) \subset (\mathbb{R}^{d+1})^k$ for the ball of radius r centered on (this ordering of) P, where the distance from P to $Q = (u_1, \dots, u_k)$ is given by

$$\|Q - P\|_\infty := \max\{\|u_i - v_i\| : 1 \le i \le k\}.$$

If P is an admissible spherical configuration, note that all configurations inside a small enough ball centered on P will also be admissible.

We will later need an equicontinuity property for the family of counting functions $P \mapsto I_P(A)$, over all measurable sets $A \subseteq S^d$; this is given in the following lemma:

Lemma 3.8 (Equicontinuity)**.** *For every admissible $P \subset S^d$ and every $\varepsilon > 0$, there is some $\delta > 0$ such that*

$$\|Q - P\|_\infty \le \delta \implies |I_Q(A) - I_P(A)| \le \varepsilon \qquad \forall A \subseteq S^d.$$

Proof. We will use the fact that the function η_P obtained in the Counting lemma can be made uniform inside a small ball centered on P. In other words, there is $r > 0$ and a function $\eta'_P : (0, 1] \to (0, 1]$ with $\lim_{t\to 0} \eta'_P(t) = 0$ such that $|I_Q(A) - I_Q(\chi_A * \mathrm{cap}_\rho)| \le \eta'_P(\rho)$ for all $Q \in \mathcal{B}(P, r)$ and all (measurable) $A \subseteq S^d$.

Now, for a given $\rho > 0$ and all $0 < \delta < \rho$, we see from the triangle inequality that

$$\|x - y\| \le \delta \implies \mathrm{Cap}(x, \rho - \delta) \subset \mathrm{Cap}(x, \rho) \cap \mathrm{Cap}(y, \rho),$$

and so $\sigma(\mathrm{Cap}(x, \rho) \setminus \mathrm{Cap}(y, \rho)) \le \sigma(\mathrm{Cap}_\rho) - \sigma(\mathrm{Cap}_{\rho-\delta})$. This implies that, for any set $A \subseteq S^d$ and any $x, y \in S^d$ with $\|x - y\| \le \delta$, we have

$$\begin{aligned} |\chi_A * \mathrm{cap}_\rho(x) - \chi_A * \mathrm{cap}_\rho(y)| &= \frac{\big|\sigma(A \cap \mathrm{Cap}(x, \rho)) - \sigma(A \cap \mathrm{Cap}(y, \rho))\big|}{\sigma(\mathrm{Cap}_\rho)} \\ &\le \frac{\sigma(\mathrm{Cap}(x, \rho) \setminus \mathrm{Cap}(y, \rho))}{\sigma(\mathrm{Cap}_\rho)} \\ &\le \frac{\sigma(\mathrm{Cap}_\rho) - \sigma(\mathrm{Cap}_{\rho-\delta})}{\sigma(\mathrm{Cap}_\rho)}. \end{aligned}$$

By the usual telescoping sum argument, whenever $\|Q - P\|_\infty \leq \delta$ we then conclude

$$|I_Q(\chi_A * \mathrm{cap}_\rho) - I_P(\chi_A * \mathrm{cap}_\rho)| \leq k \frac{\sigma(\mathrm{Cap}_\rho) - \sigma(\mathrm{Cap}_{\rho-\delta})}{\sigma(\mathrm{Cap}_\rho)}.$$

Take $\rho > 0$ small enough so that $\eta'_P(\rho) \leq \varepsilon/3$, and for this value of ρ take $0 < \delta < r$ small enough so that $\sigma(\mathrm{Cap}_{\rho-\delta}) \geq (1 - \varepsilon/3k)\,\sigma(\mathrm{Cap}_\rho)$. Then, for any $Q \in \mathcal{B}(P, \delta)$ and any measurable set $A \subseteq S^d$, we have

$$\begin{aligned}
|I_Q(A) - I_P(A)| &\leq |I_Q(A) - I_Q(\chi_A * \mathrm{cap}_\rho)| + |I_Q(\chi_A * \mathrm{cap}_\rho) - I_P(\chi_A * \mathrm{cap}_\rho)| \\
&\quad + |I_P(\chi_A * \mathrm{cap}_\rho) - I_P(A)| \\
&\leq \eta'_P(\rho) + k \frac{\sigma(\mathrm{Cap}_\rho) - \sigma(\mathrm{Cap}_{\rho-\delta})}{\sigma(\mathrm{Cap}_\rho)} + \eta'_P(\rho) \\
&\leq \frac{\varepsilon}{3} + k\frac{\varepsilon}{3k} + \frac{\varepsilon}{3} = \varepsilon,
\end{aligned}$$

as wished. □

3.4 The spherical Supersaturation theorem

Having proven that the counting function for admissible spherical configurations is robust to various kinds of small perturbations, we next show that it also satisfies a useful *supersaturation property*.

This is the second main technical tool we need to study the independence density in the spherical setting, and due to the fact that the unit sphere is compact both its statement and its proof are slightly simpler than those in the Euclidean space setting.

Theorem 3.9 (Supersaturation theorem)**.** *For every admissible configuration P on S^d and every $\varepsilon > 0$ there exists $c(\varepsilon) > 0$ such that the following holds. If $A \subseteq S^d$ has measure $\sigma(A) \geq \mathbf{m}_{S^d}(P) + \varepsilon$, then $I_P(A) \geq c(\varepsilon)$.*

Proof. Suppose for contradiction that the result is false; then there exist some $\varepsilon > 0$ and some sequence $(A_i)_{i\geq 1}$ of sets, each of density at least $\mathbf{m}_{S^d}(P) + \varepsilon$, which satisfy $\lim_{i\to\infty} I_P(A_i) = 0$.

Since the unit ball $\mathcal{B}_\infty$ of $L^\infty(S^d)$ is weak* compact (and also metrizable in this topology), by possibly restricting to a subsequence we may assume that $(\chi_{A_i})_{i\geq 1}$ converges in the weak* topology of $L^\infty(S^d)$; let us denote its limit by $f \in \mathcal{B}_\infty$. It is clear that $0 \leq f \leq 1$ almost everywhere, and $\int_{S^d} f(x)\, d\sigma(x) = \lim_{i\to\infty} \sigma(A_i) \geq \mathbf{m}_{S^d}(P) + \varepsilon$. By weak* continuity of I_P (Lemma 3.7), we also have $I_P(f) = \lim_{i\to\infty} I_P(A_i) = 0$.

Now let $B := \{x \in S^d : f(x) \geq \varepsilon\}$. Since

$$\varepsilon \cdot \chi_B(x) \leq f(x) < \varepsilon + \chi_B(x) \quad \text{for a.e. } x \in S^d,$$

we conclude that $I_P(B) \leq \varepsilon^{-k} I_P(f) = 0$ and

$$\sigma(B) > \int_{S^d} f(x)\, d\sigma(x) - \varepsilon \geq \mathbf{m}_{S^d}(P).$$

But this set B contradicts Lemma 3.2, finishing the proof. □

It will be useful to also introduce a spherical analogue of the zooming-out operator, which acts on measurable spherical sets and represents the points on the sphere around which the considered set has a somewhat large density. Given quantities $\delta, \gamma > 0$, we then denote by $\mathcal{Z}_\delta(\gamma)$ the operator which takes a measurable set $A \subseteq S^d$ to the set

$$\mathcal{Z}_\delta(\gamma)A := \{x \in S^d : d_{\mathrm{Cap}(x,\delta)}(A) \geq \gamma\}.$$

The most important property of the zooming-out operator is the following result:

Corollary 3.10. *For every admissible configuration P on S^d and every $\varepsilon > 0$ there exists $\delta_0 > 0$ such that the following holds. For any $\delta \leq \delta_0$, if $A \subseteq S^d$ satisfies*

$$\sigma(\mathcal{Z}_\delta(\varepsilon)A) \geq \mathbf{m}_{S^d}(P) + \varepsilon,$$

then A contains a congruent copy of P.

Proof. By the Supersaturation theorem, we know that

$$\sigma(\mathcal{Z}_\delta(\varepsilon)A) \geq \mathbf{m}_{S^d}(P) + \varepsilon \implies I_P(\mathcal{Z}_\delta(\varepsilon)A) \geq c(\varepsilon)$$

holds for all $\delta > 0$. By the Counting lemma, we then have

$$\begin{aligned} I_P(A) \geq I_P(\chi_A * \mathrm{cap}_\delta) - \eta_P(\delta) &\geq \varepsilon^k I_P(\mathcal{Z}_\delta(\varepsilon)A) - \eta_P(\delta) \\ &\geq \varepsilon^k c(\varepsilon) - \eta_P(\delta). \end{aligned}$$

Since $\eta_P(\delta) \to 0$ as $\delta \to 0$, there is some $\delta_0 > 0$ such that for all $\delta \leq \delta_0$ we can conclude $I_P(A) > 0$; this implies that A contains a copy of P. □

3.5 From the sphere to spherical caps

We must now tackle the problem of obtaining a relationship between the independence density $\mathbf{m}_{S^d}(P)$ of a given configuration $P \subset S^d$ and its spherical cap version $\mathbf{m}_{\mathrm{Cap}(x,\rho)}(P)$, as this will be needed later.

In the Euclidean setting this was very easy to do (see Lemma 2.2), using the fact that we can tessellate $\mathbb{R}^d$ with cubes $Q(x, R)$ of any given side length $R > 0$. This is no longer the case in the spherical setting, as it is impossible to completely cover S^d using non-overlapping spherical caps of some given radius; in fact, this cannot be done even approximately if we require the radii of the spherical caps to be the same (as we did with the side length of the cubes in $\mathbb{R}^d$).

We will then need to use a much weaker 'almost-covering' result, saying that we can cover almost all of the sphere by using finitely many non-overlapping spherical caps of different radii. For technical reasons we will also want these radii to be arbitrarily small.

Lemma 3.11. *For every $\varepsilon > 0$ there is a finite cap packing*

$$\mathcal{P} = \{\mathrm{Cap}(x_I, \rho_i) : 1 \leq i \leq N\}$$

of S^d with density $\sigma(\mathcal{P}) > 1 - \varepsilon$ and with radii $\rho_i \leq \varepsilon$ for all $1 \leq i \leq N$.

Proof. We will use the same notation for both a collection of caps and the set of points on S^d which belong to (at least) one of these caps. The desired packing $\mathcal{P}$ will be constructed in several steps, starting with $\mathcal{P}_0 := \{\mathrm{Cap}(e, \varepsilon)\}$.

Now suppose $\mathcal{P}_{i-1}$ has already been constructed (and is finite) for some $i \geq 1$, and let us construct $\mathcal{P}_i$. Define

$$C_i := \{\mathrm{Cap}(x, \min\{\varepsilon, \mathrm{dist}(x, \mathcal{P}_{i-1})\}) : x \in S^d \setminus \mathcal{P}_{i-1}\},$$

and note that C_i is a covering of $S^d \setminus \mathcal{P}_{i-1}$ by caps of positive radii (since $\mathcal{P}_{i-1}$ is closed on S^d). By Vitali's Covering lemma,[4] there is a countable subcollection

$$Q_i = \bigcup_{j=1}^{\infty} \{\mathrm{Cap}(x_j, r_j)\} \subset C_i$$

of *disjoint* caps in C_i such that $S^d \setminus \mathcal{P}_{i-1} \subseteq \bigcup_{j=1}^{\infty} \mathrm{Cap}(x_j, 5r_j)$. In particular

$$1 - \sigma(\mathcal{P}_{i-1}) = \sigma(S^d \setminus \mathcal{P}_{i-1}) \leq \sum_{j=1}^{\infty} \sigma(\mathrm{Cap}(x_j, 5r_j)) \leq K_d\, \sigma(Q_i),$$

where we denote $K_d := \sup_{r>0} \sigma(\mathrm{Cap}_{5r})/\sigma(\mathrm{Cap}_r) < \infty$. Taking $N_i \in \mathbb{N}$ such that

$$\sum_{j=1}^{N_i} \sigma(\mathrm{Cap}(x_j, r_j)) \geq \sigma(Q_i) - \frac{1 - \sigma(\mathcal{P}_{i-1})}{2K_d},$$

we see that $\mathcal{P}_i' := \{\mathrm{Cap}(x_j, r_j) : 1 \leq j \leq N_i\} \subset S^d \setminus \mathcal{P}_{i-1}$ satisfies

$$\sigma(\mathcal{P}_i') \geq \frac{1 - \sigma(\mathcal{P}_{i-1})}{2K_d}.$$

Now set $\mathcal{P}_i := \mathcal{P}_{i-1} \cup \mathcal{P}_i'$; this is a finite cap packing with

$$\begin{aligned} 1 - \sigma(\mathcal{P}_i) = 1 - \sigma(\mathcal{P}_{i-1}) - \sigma(\mathcal{P}_i') &\leq (1 - \sigma(\mathcal{P}_{i-1}))\left(1 - \frac{1}{2K_d}\right) \\ &\leq (1 - \sigma(\mathrm{Cap}_\varepsilon))\left(1 - \frac{1}{2K_d}\right)^i \end{aligned}$$

(where the last inequality follows by induction). Taking $n \geq 1$ large enough so that

$$(1 - \sigma(\mathrm{Cap}_\varepsilon))\left(1 - \frac{1}{2K_d}\right)^n < \varepsilon,$$

we see that $\mathcal{P} := \mathcal{P}_n$ satisfies all requirements. □

We can now obtain our analogue of Lemma 2.2, relating the two versions of independence density in the spherical setting:

[4]Note that spherical caps are exactly the (closed) balls of the separable metric space S^d endowed with the Euclidean distance induced from $\mathbb{R}^{d+1}$.

Lemma 3.12. *For every $\varepsilon > 0$, $\rho > 0$ there is $t_0 > 0$ such that the following holds whenever $P_1, \dots, P_n \subset S^d$ have diameter at most t_0:*

$$\left|\mathbf{m}_{\mathrm{Cap}(x,\rho)}(P_1, \dots, P_n) - \mathbf{m}_{S^d}(P_1, \dots, P_n)\right| < \varepsilon.$$

Proof. If $A \subset S^d$ is a set that does not contain copies of $P_1, \dots, P_n$, then for any $x \in S^d$ the set $A \cap \mathrm{Cap}(x, \rho) \subseteq \mathrm{Cap}(x, \rho)$ also does not contain copies of $P_1, \dots, P_n$ and $\mathbb{E}_{x \in S^d}[d_{\mathrm{Cap}(x,\rho)}(A)] = \sigma(A)$. There must then exist some $x \in S^d$ such that

$$d_{\mathrm{Cap}(x,\rho)}(A \cap \mathrm{Cap}(x, \rho)) = d_{\mathrm{Cap}(x,\rho)}(A) \geq \sigma(A),$$

proving that $\mathbf{m}_{\mathrm{Cap}(x,\rho)}(P_1, \dots, P_n) \geq \mathbf{m}_{S^d}(P_1, \dots, P_n)$.

For the opposite direction, let $\gamma \leq \varepsilon/4$ be small enough so that

$$\sigma(\mathrm{Cap}_{\rho+\gamma}) \leq (1 + \varepsilon/4)\, \sigma(\mathrm{Cap}_\rho).$$

By Lemma 3.11, we know there is a cap packing $\mathcal{P} = \{\mathrm{Cap}(x_I, \rho_i) : 1 \leq i \leq N\}$ of S^d with $\sigma(\mathcal{P}) \geq 1 - \gamma$ and $0 < \rho_1, \dots, \rho_N \leq \gamma$. Now let $t_0 > 0$ be small enough so that $\sigma(\mathrm{Cap}_{\rho_i - 2t_0}) \geq (1 - \varepsilon/4)\, \sigma(\mathrm{Cap}_{\rho_i})$ for all $1 \leq i \leq N$; note that t_0 will ultimately depend only on ε and ρ.

Fixing any configurations $P_1, \dots, P_n \subset S^d$ of diameter at most t_0, let $A \subset \mathrm{Cap}(x, \rho)$ be a set which avoids all of them. We shall construct a set $\tilde{A} \subset S^d$ which also avoids $P_1, \dots, P_n$, and which satisfies $\sigma(\tilde{A}) > d_{\mathrm{Cap}(x,\rho)}(A) - \varepsilon$; this will finish the proof.

For each $1 \leq i \leq N$, denote $\tilde{\rho}_i := \rho_i - 2t_0 < \gamma$. We have that

$$\begin{aligned}
\sigma(A) &= \int_{S^d} d_{\mathrm{Cap}(y,\tilde{\rho}_i)}(A)\, d\sigma(y) \\
&= \int_{\mathrm{Cap}(x,\rho+\tilde{\rho}_i)} d_{\mathrm{Cap}(y,\tilde{\rho}_i)}(A)\, d\sigma(y) \\
&\leq \int_{\mathrm{Cap}(x,\rho)} d_{\mathrm{Cap}(y,\tilde{\rho}_i)}(A)\, d\sigma(y) + \sigma(\mathrm{Cap}_{\rho+\tilde{\rho}_i}) - \sigma(\mathrm{Cap}_\rho).
\end{aligned}$$

Since $\tilde{\rho}_i < \gamma$, dividing by $\sigma(\mathrm{Cap}_\rho)$ we obtain

$$\begin{aligned}
\mathbb{E}_{y \in \mathrm{Cap}(x,\rho)}[d_{\mathrm{Cap}(y,\tilde{\rho}_i)}(A)] &\geq \frac{\sigma(A)}{\sigma(\mathrm{Cap}_\rho)} - \frac{\sigma(\mathrm{Cap}_{\rho+\tilde{\rho}_i}) - \sigma(\mathrm{Cap}_\rho)}{\sigma(\mathrm{Cap}_\rho)} \\
&> d_{\mathrm{Cap}(x,\rho)}(A) - \frac{\varepsilon}{4}.
\end{aligned}$$

There must then exist $y_i \in \mathrm{Cap}(x, \rho)$ for which $d_{\mathrm{Cap}(y_I,\tilde{\rho}_i)}(A) > d_{\mathrm{Cap}(x,\rho)}(A) - \varepsilon/4$; fix one such y_i for each $1 \leq i \leq N$, and let $T_{y_i \to x_i} \in \mathrm{O}(d + 1)$ be any rotation taking y_i to x_i (and thus taking $\mathrm{Cap}(y_i, \tilde{\rho}_i)$ to $\mathrm{Cap}(x_i, \tilde{\rho}_i)$).

We claim that the set

$$\tilde{A} := \bigcup_{i=1}^{N} T_{y_i \to x_i}(A \cap \mathrm{Cap}(y_i, \tilde{\rho}_i))$$

satisfies our requirements. Indeed, we have

$$\begin{aligned}
\sigma(\tilde{A}) = \sum_{i=1}^{N} \sigma(A \cap \mathrm{Cap}(y_i, \tilde{\rho}_i)) &= \sum_{i=1}^{N} d_{\mathrm{Cap}(y_I, \tilde{\rho}_i)}(A) \cdot \sigma(\mathrm{Cap}_{\tilde{\rho}_i}) \\
&> \sum_{i=1}^{N} \left(d_{\mathrm{Cap}(x,\rho)}(A) - \frac{\varepsilon}{4}\right) \cdot \left(1 - \frac{\varepsilon}{4}\right) \sigma(\mathrm{Cap}_{\rho_i}) \\
&\geq \left(d_{\mathrm{Cap}(x,\rho)}(A) - \frac{\varepsilon}{2}\right) \sigma(\mathcal{P}) \\
&> d_{\mathrm{Cap}(x,\rho)}(A) - \varepsilon.
\end{aligned}$$

Moreover, since $\mathrm{diam}(P_j) \leq t_0$ and the caps $\mathrm{Cap}(x_i, \tilde{\rho}_i)$ are (at least) $2t_0$-distant from each other, we see that any copy of P_j in $\tilde{A} \subset \bigcup_{i=1}^N \mathrm{Cap}(x_i, \tilde{\rho}_i)$ must be entirely contained in one of the the caps $\mathrm{Cap}(x_i, \tilde{\rho}_i)$. But then it should also be contained (after rotation by $T^{-1}_{y_i \to x_i}$) in $A \cap \mathrm{Cap}(y_i, \tilde{\rho}_i)$; this shows that $\tilde{A}$ does not contain copies of P_j for any $1 \leq j \leq N$, since A doesn't, and we are done. □

3.6 Results on the spherical independence density

We are finally ready to start a more detailed study of the independence density parameter in the spherical setting.

We start by providing a general lower bound on the independence density of several different configurations in terms of their individual independence densities:

Lemma 3.13 (Supermultiplicativity). *For all configurations $P_1, \ldots, P_n$ on S^d, we have*

$$\mathbf{m}_{S^d}(P_1, \ldots, P_n) \geq \prod_{i=1}^{n} \mathbf{m}_{S^d}(P_i).$$

Proof. Choose, for each $1 \leq i \leq n$, a set $A_i \subset S^d$ which avoids configuration P_i. By taking independent rotations $R_i A_i$ of each set A_i, we see that

$$\begin{aligned}
\mathbb{E}_{R_1,\ldots,R_n \in \mathrm{O}(d+1)}\left[\sigma\left(\bigcap_{i=1}^{n} R_i A_i\right)\right] &= \int_{S^d} \prod_{i=1}^{n} \mathbb{E}_{R_i \in \mathrm{O}(d+1)}[\chi_{A_i}(R_i^{-1} x)]\, d\sigma(x) \\
&= \prod_{i=1}^{n} \sigma(A_i).
\end{aligned}$$

There must then exist $R_1, \ldots, R_n \in \mathrm{O}(d+1)$ for which

$$\sigma\left(\bigcap_{i=1}^{n} R_i A_i\right) \geq \prod_{i=1}^{n} \sigma(A_i).$$

Since $\bigcap_{i=1}^n R_i A_i$ avoids all configurations $P_1, \ldots, P_n$ and the sets $A_1, \ldots, A_n$ were chosen arbitrarily, the result follows. □

Using supersaturation, we can show that this lower bound is essentially tight when the configurations considered are all admissible and each one is at a different size scale. Intuitively, this happens because the constraints of avoiding each of these configurations will act at distinct scales and thus not correlate with each other.

Theorem 3.14 (Asymptotic independence)**.** *For every admissible configurations $P_1, \dots, P_n$ on S^d and every $\varepsilon > 0$ there is a positive increasing function $f : (0,1] \to (0,1]$ such that the following holds. Whenever $0 < t_1, \dots, t_n \leq 1$ satisfy $t_{i+1} \leq f(t_i)$ for $1 \leq i < n$, we have*

$$\left| \mathbf{m}_{S^d}(t_1 P_1, \dots, t_n P_n) - \prod_{i=1}^{n} \mathbf{m}_{S^d}(t_i P_i) \right| \leq \varepsilon.$$

Proof. We have already seen that $\mathbf{m}_{S^d}(t_1 P_1, \dots, t_n P_n) \geq \prod_{i=1}^{n} \mathbf{m}_{S^d}(t_i P_i)$, so it suffices to show that $\mathbf{m}_{S^d}(t_1 P_1, \dots, t_n P_n) \leq \prod_{i=1}^{n} \mathbf{m}_{S^d}(t_i P_i) + \varepsilon$ for suitably separated $t_1, \dots, t_n \leq 1$. We will do so by induction on n, with the base case $n = 1$ being trivial (and taking $f \equiv 1$).

Suppose then $n \geq 2$ and we have already proven the result for $n-1$ configurations. Let $\tilde{f} : (0,1] \to (0,1]$ be the function promised by the theorem applied to the $n-1$ configurations $P_2, \dots, P_n$ and with accuracy ε, so that whenever $0 < t_2 \leq 1$ and $0 < t_{j+1} \leq \tilde{f}(t_j)$ for each $2 \leq j < n$ we have

$$\mathbf{m}_{S^d}(t_2 P_2, \dots, t_n P_n) \leq \prod_{j=2}^{n} \mathbf{m}_{S^d}(t_j P_j) + \varepsilon.$$

By the corollary to the Supersaturation theorem (Corollary 3.10), for all $0 < t_1 \leq 1$ there is $\delta_0 = \delta_0(\varepsilon; t_1 P_1) > 0$ such that

$$\sigma(\mathcal{Z}_{\delta_0}(\varepsilon) A) \geq \mathbf{m}_{S^d}(t_1 P_1) + \varepsilon \implies A \text{ contains a copy of } t_1 P_1.$$

Applying Lemma 3.12 with radius $\rho = \delta_0$, we see there is $t_0 = t_0(\varepsilon, \delta_0) > 0$ for which

$$\mathbf{m}_{\mathrm{Cap}(x,\delta_0)}(t_2 P_2, \dots, t_n P_n) \leq \mathbf{m}_{S^d}(t_2 P_2, \dots, t_n P_n) + \varepsilon$$

holds whenever $0 < t_2, \dots, t_n \leq t_0/2$.

Let now $0 < t_1, \dots, t_n \leq 1$ be numbers satisfying

$$t_2 \leq t_0(\varepsilon, \delta_0(\varepsilon; t_1 P_1))/2 \quad \text{and} \quad t_{j+1} \leq \tilde{f}(t_j) \text{ for all } 2 \leq j < n.$$

If $A \subset S^d$ does not contain copies of $t_1 P_1, \dots, t_n P_n$, then by the preceding discussion we must have $\sigma(\mathcal{Z}_{\delta_0}(\varepsilon) A) < \mathbf{m}_{S^d}(t_1 P_1) + \varepsilon$ and, for all $x \in S^d$,

$$\begin{aligned} d_{\mathrm{Cap}(x,\delta_0)}(A) \leq \mathbf{m}_{\mathrm{Cap}(x,\delta_0)}(t_2 P_2, \dots, t_n P_n) &\leq \mathbf{m}_{S^d}(t_2 P_2, \dots, t_n P_n) + \varepsilon \\ &\leq \prod_{j=2}^{n} \mathbf{m}_{S^d}(t_j P_j) + 2\varepsilon. \end{aligned}$$

This means that, inside caps $\mathrm{Cap}(x, \delta_0)$ of radius δ_0, A has density less than ε (when $x \notin \mathcal{Z}_{\delta_0}(\varepsilon) A$) except on a set of measure at most $\mathbf{m}_{S^d}(t_1 P_1) + \varepsilon$, when it instead has density at most $\prod_{j=2}^{n} \mathbf{m}_{S^d}(t_j P_j) + 2\varepsilon$. Taking averages, we conclude that

$$\begin{aligned}\sigma(A) &= \mathbb{E}_{x\in S^d}\big[d_{\mathrm{Cap}(x,\,\delta)}(A)\big] \\ &\le \varepsilon + (\mathbf{m}_{S^d}(t_1P_1) + \varepsilon)\left(\prod_{j=2}^{n} \mathbf{m}_{S^d}(t_jP_j) + 2\varepsilon\right) \\ &\le 6\varepsilon + \prod_{i=1}^{n} \mathbf{m}_{S^d}(t_iP_i).\end{aligned}$$

It thus suffices to take the function $f : (0,1] \to (0,1]$ given by

$$f(t) = \min\left\{\tilde{f}(t),\ \frac{t_0(\varepsilon/6,\ \delta_0(\varepsilon/6;\ tP_1))}{2}\right\}$$

to conclude the induction. □

Recalling from Lemma 3.1 that $\mathbf{m}_{S^d}(tP)$ is bounded away from both zero and one for $0 < t \le 1$, an immediate consequence of this theorem is the following: if P is admissible, then $\mathbf{m}_{S^d}(t_1P,\ t_2P,\ \ldots,\ t_nP)$ decays exponentially with n as the ratios t_{j+1}/t_j between consecutive scales go to zero. By considering an infinite sequence of 'counterexamples' as we did in our proof of Bourgain's theorem last chapter, we then obtain from Theorem 3.14 the following result:

Corollary 3.15. *Let $P \subset S^d$ be an admissible configuration. If $A \subseteq S^d$ has positive measure, then there is some number $t_0 > 0$ such that A contains a congruent copy of tP for all $t \le t_0$.*

This corollary can be seen as the counterpart to Bourgain's theorem in the spherical setting, where it impossible to consider arbitrarily large dilates. (The equivalent result of containing all sufficiently small dilates of a configuration in the Euclidean setting also holds with the same proof.)

We will next prove that the independence density function $P \mapsto \mathbf{m}_{S^d}(P)$ is continuous on the set of admissible configurations on S^d. Before doing so, it is interesting to note that a similar result does *not* hold for two-point configurations on the unit circle S^1 (which can be seen as the very first instance of non-admissible configurations). Indeed, it was shown by DeCorte and Pikhurko [23] that $\mathbf{m}_{S^1}(\{u, v\})$ is *discontinuous* at a configuration $\{u, v\} \subset S^1$ whenever the arc length between u and v is a rational multiple of 2π with odd denominator.

Theorem 3.16 (Continuity of the independence density)**.** *For any $n \ge 1$, the function $(P_1, \ldots, P_n) \mapsto \mathbf{m}_{S^d}(P_1, \ldots, P_n)$ is continuous on the set of n admissible spherical configurations.*

Proof. For simplicity we will prove the result in the case of only one forbidden configuration, but the general case follows from the same argument.

Fix some $\varepsilon > 0$ and some admissible configuration P on S^d, and let $c(\varepsilon) > 0$ be the constant promised by the Supersaturation theorem (Theorem 3.9). By our 'equicontinuity' Lemma 3.8, there exists $\delta > 0$ such that, whenever $\|Q - P\|_\infty \le \delta$, we have $|I_Q(A) - I_P(A)| \le c(\varepsilon)$ for all $A \subseteq S^d$.

Suppose $Q \in \mathcal{B}(P, \delta)$ and $A \subset S^d$ is a measurable set avoiding Q; we must then have $I_P(A) \leq c(\varepsilon)$, and so (by the Supersaturation theorem) $\sigma(A) \leq \mathbf{m}_{S^d}(P) + \varepsilon$. We conclude that $\mathbf{m}_{S^d}(Q) \leq \mathbf{m}_{S^d}(P) + \varepsilon$ whenever $Q \in \mathcal{B}(P, \delta)$.

Now write $P = \{v_1, \ldots, v_k\}$, and consider the function $g_P : (S^d)^k \times \mathrm{O}(d+1) \to \mathbb{R}$ given by

$$g_P(x_1, \ldots, x_k, T) := \sum_{i=1}^{k} \|x_i - Tv_i\|.$$

Note that this function is continuous, nonnegative and that

$$\min_{T \in \mathrm{O}(d+1)} g_P(x_1, \ldots, x_k, T) = 0 \quad \text{if and only if} \quad \{x_1, \ldots, x_k\} \simeq P.$$

By elementary measure theory, we can find a compact set $A \subset S^d$ which avoids P and has measure $\sigma(A) \geq \mathbf{m}_{S^d}(P) - \varepsilon$. The continuous function g_P attains a minimum on the compact set $A^k \times \mathrm{O}(d+1)$; denote this minimum by γ, and note that $\gamma > 0$ since A avoids P.

Let us show that A also avoids Q, for all $Q \in \mathcal{B}(P, \gamma/2k)$. Indeed, writing $Q = \{u_1, \ldots, u_k\}$, for any points $x_1, \ldots, x_k \in A$ and any $T \in \mathrm{O}(d+1)$ we have that

$$\begin{aligned}\sum_{i=1}^{k} \|x_i - Tu_i\| &\geq \sum_{i=1}^{k} \big|\|x_i - Tv_i\| - \|Tu_i - Tv_i\|\big| \\ &\geq g_P(x_1, \ldots, x_k, T) - k\|Q - P\|_\infty,\end{aligned}$$

which is at least $\gamma/2 > 0$ if $\|Q - P\|_\infty \leq \gamma/2k$. For such configurations we obtain

$$\mathbf{m}_{S^d}(Q) \geq \sigma(A) \geq \mathbf{m}_{S^d}(P) - \varepsilon.$$

We conclude that $|\mathbf{m}_{S^d}(Q) - \mathbf{m}_{S^d}(P)| \leq \varepsilon$ whenever $\|Q - P\|_\infty \leq \min\{\delta, \gamma/2k\}$, finishing the proof. □

As our definition of the independence density $\mathbf{m}_{S^d}(P)$ involved a supremum over all P-avoiding measurable sets $A \subseteq S^d$, it is not immediately clear whether there actually exists a measurable P-avoiding set attaining this extremal value of density. In fact, such a result is *false* in the case where $d = 1$ and we are considering two-point configurations $\{u, v\} \subset S^1$: if the length of the arc between u and v is *not* a rational multiple of π, it was shown by Székely [62] that $\mathbf{m}_{S^1}(\{u, v\}) = 1/2$ but that there is no $\{u, v\}$-avoiding measurable set of density $1/2$.

We will now show that extremizer sets exist whenever the configuration we are forbidding is admissible. Note that the result also holds (with the same proof) when forbidding several admissible configurations; this generalizes to higher-order configurations a theorem of DeCorte and Pikhurko [23] for forbidden distances on the sphere.

Theorem 3.17 (Existence of extremizers)**.** *If $P \subset S^d$ is an admissible configuration, then there is a P-avoiding measurable set $A \subseteq S^d$ attaining $\sigma(A) = \mathbf{m}_{S^d}(P)$.*

Proof. Let $A_1, A_2, \cdots \subseteq S^d$ be a sequence of P-avoiding measurable sets whose measure converges to $\mathbf{m}_{S^d}(P)$. By passing to a subsequence if necessary, we may assume that $(\chi_{A_i})_{i \geq 1}$ converges to some function $f \in \mathcal{B}_\infty$ in the weak* topology of $L^\infty(S^d)$.

By weak* convergence we know that $0 \leq f \leq 1$ almost everywhere, and that

$$\int_{S^d} f(x)\, d\sigma(x) = \lim_{i\to\infty} \sigma(A_i) = \mathbf{m}_{S^d}(P);$$

by weak* continuity we also have $I_P(f) = \lim_{i\to\infty} I_P(A_i) = 0$. Denoting $A := \operatorname{supp} f$, we easily conclude that $I_P(A) = 0$, and also

$$\sigma(A) = \int_{S^d} \chi_A(x)\, d\sigma(x) \geq \int_{S^d} f(x)\, d\sigma(x) = \mathbf{m}_{S^d}(P). \tag{3.6}$$

But Lemma 3.2 implies we can remove a zero-measure subset from A in order to remove all copies of P. The theorem follows. □

To conclude this chapter, let us make explicit what we can say about the possible independence densities when forbidding n distinct contractions of an admissible configuration P; due to lack of dilation invariance in the spherical setting, characterizing these values in terms of simpler quantities is much harder than it is in the Euclidean setting.

Denote $\mathcal{M}_n^{S^d}(P) := \{\mathbf{m}_{S^d}(t_1 P, t_2 P, \ldots, t_n P) : 0 < t_1 < t_2 < \cdots < t_n \leq 1\}$. Due to continuity of $\mathbf{m}_{S^d}$ (Theorem 3.16) this set is an interval, and its upper extremity is $\sup_{0<t\leq 1} \mathbf{m}_{S^d}(tP)$. By supermultiplicativity (Lemma 3.13) the lower extremity of $\mathcal{M}_n^{S^d}(P)$ is at least $\inf_{0<t\leq 1} \mathbf{m}_{S^d}(tP)^n$, and by asymptotic independence (Theorem 3.14) it can be at most $\inf_{0<t\leq 1} \mathbf{m}_{S^d}(tP) \cdot \liminf_{t\to 0} \mathbf{m}_{S^d}(tP)^{n-1}$.

3.7 Remarks and open problems

It is not clear whether the results shown in this chapter should continue to hold when the configurations considered are not admissible. Our reasons for requiring such a condition are quite similar to those we had in Euclidean space, and as in the Euclidean setting we believe that only some sort of non-degeneracy condition should be necessary.

Another question we ask is related to a suspected 'compatibility condition' between the Euclidean and spherical settings. Since S^d resembles $\mathbb{R}^d$ at small scales, it seem geometrically intuitive that $\mathbf{m}_{S^d}(tP)$ should get increasingly close to $\mathbf{m}_{\mathbb{R}^d}(P)$ as $t \to 0$, whenever P is a contractible configuration on S^d. (It is easy to show that a configuration $P \subset S^d$ is contractible if and only if it is contained in a d-dimensional affine subspace, so we can embed it in $\mathbb{R}^d$.) We wish to know whether this intuition is correct:

Question 3.18. *Is it true that* $\lim_{t\to 0} \mathbf{m}_{S^d}(tP) = \mathbf{m}_{\mathbb{R}^d}(P)$ *whenever* $P \subset S^d$ *is a contractible configuration?*

Note that in dimension $d = 1$ this question has a positive answer, due to DeCorte and Pikhurko's characterization [23, Theorem 3.2] of the independence density of two points in S^1 and the easy fact that $\mathbf{m}_{\mathbb{R}^1}(\{0, 1\}) = 1/2$.

Chapter Four

An exact completely positive formulation

This chapter is devoted to obtaining conic programming formulations for the independence density parameters. We shall do so by extending the cone of completely positive matrices to multivariable functions on the Euclidean space and the sphere, and then generalizing the completely positive formulation of the independence number of graphs to certain geometric hypergraphs.

4.1 A conic linear program for the independence number of hypergraphs

We say that a symmetric matrix $M \in \mathbb{R}^{n\times n}$ is *completely positive* if it is a conic combination of rank-one, symmetric and nonnegative matrices; in other words, if there are nonnegative vectors $v_1, \ldots, v_m \in \mathbb{R}^n_+$ such that

$$M = v_1 \otimes v_1 + \cdots + v_m \otimes v_m.$$

Here the tensor product $u \otimes v$ is the matrix given by $(u \otimes v)_{ij} = u_i v_j$. The set of all completely positive matrices forms a closed convex cone of symmetric matrices.

While the problem of optimizing linear functions over the completely positive cone is NP-hard in general, there has been significant study in this area (see e.g. [6, 18, 9]) since it provides exact reformulations of interesting but hard combinatorial problems as convex optimization problems. For instance, de Klerk and Pasechnik [18] showed that the independence number of a finite graph can be formulated as a linear optimization problem over the cone of completely positive matrices; more precisely, denoting by $\alpha(G)$ the independence number of a given graph G, they showed that

$$\begin{aligned} \alpha(G) \;=\; \max \;& \textstyle\sum_{i,j \in V(G)} M_{ij} \\ & \textstyle\sum_{i \in V(G)} M_{ii} = 1, \\ & M_{ij} = 0 \quad \text{if } \{i, j\} \in E(G), \\ & M \in \mathbb{R}^{V(G)\times V(G)} \text{ is completely positive.} \end{aligned} \tag{4.1}$$

We will now extend this formulation to hypergraphs of higher uniformity, substituting matrices by higher-order arrays and suitably generalizing the completely positive cone.

For convenience, we shall use a more 'functional' notation rather than the usual 'multi-index' notation for higher-order arrays. Given an integer $k \geq 2$, we define a *completely positive k-array* $F \in \mathbb{R}^{n^k}$ as a conic combination of rank-one, symmetric and nonnegative k-arrays. In other words, a k-array $F \in \mathbb{R}^{n^k}$ is completely positive if there are nonnegative vectors $v_1, \ldots, v_m \in \mathbb{R}^n_+$ such that

$$F = v_1^{\otimes k} + \cdots + v_m^{\otimes k};$$

here $v^{\otimes k}$ denotes the *k-th tensor power* of v, which has coordinates $v^{\otimes k}(i_1, \ldots, i_k) = \prod_{j=1}^k v_{i_j}$.

Given some (finite) k-uniform hypergraph H on vertex set $V(H)$, we then consider the optimization program

$$\begin{aligned} \max\ & \textstyle\sum_{x_1,\ldots,x_k \in V(H)} F(x_1, \ldots, x_k) \\ & \textstyle\sum_{x_1,\ldots,x_{k-1} \in V(H)} F(x_1, \ldots, x_{k-1}, x_{k-1}) = 1, \\ & F(x_1, \ldots, x_k) = 0 \quad \text{if } \{x_1, \ldots, x_k\} \in E(H), \\ & F \in \mathbb{R}^{V(H)^k} \text{ is completely positive.} \end{aligned} \tag{4.2}$$

A variation of de Klerk and Pasechnik's argument for graphs can be used to show that program (4.2) above is an exact formulation for the independence number of hypergraphs:

Theorem 4.1. *Whenever H is a k-uniform hypergraph, the optimal value of program (4.2) is equal to $\alpha(H)$.*

Proof. Denote bu v the value of program (4.2), and write $V = V(H)$ for convenience. It is easy to show that $v \geq \alpha(H)$: let $I \subseteq V$ be any independent set in H, and consider the function

$$F_I(x_1, \ldots, x_k) := \frac{1}{|I|^{k-1}} \prod_{i=1}^k \chi_I(x_i).$$

This function is a feasible solution of (4.2) with value $|I|$; since I was arbitrary, the optimum value v is at least $\alpha(H)$.

In order to show the converse inequality $v \leq \alpha(H)$, consider any feasible solution F of program (4.2). Since F is completely positive, it can be written as a conic combination

$$F = \sum_{i=1}^n \lambda_i f_i^{\otimes k} \tag{4.3}$$

for some constants $\lambda_i > 0$ and nonnegative functions $f_i : V \to \mathbb{R}_+$. Let $\|\cdot\|_1$ and $\|\cdot\|_2$ denote the usual L^1 and L^2 norms on $\mathbb{R}^V$, that is

$$\|f\|_1 := \sum_{x \in V} |f(x)| \quad \text{and} \quad \|f\|_2 := \left(\sum_{x \in V} |f(x)|^2 \right)^{1/2}.$$

After a suitable normalization of (4.3), we may assume that $\|f_i\|_2 = 1$ for each $1 \leq i \leq n$; we then have

$$\begin{aligned}\sum_{x_1,\ldots,x_{k-1}\in V} F(x_1, \ldots, x_{k-1}, x_{k-1}) &= \sum_{i=1}^{n} \lambda_i \sum_{x_1,\ldots,x_{k-1}\in V} \left(\prod_{j=1}^{k-2} f_i(x_j)\right) f_i(x_{k-1})^2 \\ &= \sum_{i=1}^{n} \lambda_i \left(\sum_{x\in V} f_i(x)\right)^{k-2} \left(\sum_{x\in V} f_i(x)^2\right) \\ &= \sum_{i=1}^{n} \lambda_i \|f_i\|_1^{k-2}.\end{aligned}$$

The first constraint of program (4.2) thus implies that $\sum_{i=1}^{n} \lambda_i \|f_i\|_1^{k-2} = 1$; there must then exist some $1 \leq j \leq n$ satisfying

$$\|f_j\|_1^2 \geq \sum_{i=1}^{n} \lambda_i \|f_i\|_1^{k-2} \cdot \|f_i\|_1^2 = \sum_{i=1}^{n} \lambda_i \left(\sum_{x\in V} f_i(x)\right)^k = \sum_{x_1,\ldots,x_k\in V} F(x_1, \ldots, x_k). \tag{4.4}$$

Fix such an index j, and denote the support of f_j by $I_j \subseteq V$. Since F is zero on edges of H and each function f_i is nonnegative, it follows that $\prod_{\ell=1}^{k} f_j(x_\ell) = 0$ whenever $\{x_1, \ldots, x_k\}$ is an edge of H; this implies that its support I_j is an independent set in H. By Cauchy-Schwarz we have that

$$\|f_j\|_1 = \sum_{x\in V} f_j(x)\chi_{I_j}(x) \leq \|f_j\|_2 \, \|\chi_{I_j}\|_2 = |I_j|^{1/2},$$

so we conclude from equation (4.4) that the objective function of program (4.2) for F is at most $|I_j|$. Since I_j is independent in H and F is an arbitrary feasible solution, this implies that $\nu \leq \alpha(H)$ and concludes the proof. □

4.2 A conic linear program for the independence density

We now wish to extend program (4.2) considered above to *infinite* geometric hypergraphs, in particular obtaining completely positive formulations for the independence density parameters.

We will first need to properly extend the notion of completely positive k-arrays to functions $F : V^k \to \mathbb{R}$, with V being an infinite set. The simplest way to do so is to require all finite induced k-arrays to be completely positive; more formally, we say that a k-variable function $F : V^k \to \mathbb{R}$ is *completely positive* if for every finite subset $U \subseteq V$ the k-array $(F(u_1, \ldots, u_k))_{u_1,\ldots,u_k\in U}$ is completely positive. When the set V is endowed with a topology, it is convenient to also restrict our attention to those completely positive functions which are continuous; this requirement gives them much more structure and makes them easier to work with.

Using this notion, it is now easy to extend program (4.2) to infinite hypergraphs whose vertex set is a topological space endowed with a Borel measure. However, it is no longer clear whether the resulting program is an exact formulation for the appropriate analogue of the independence number. The next theorems show that this is still true if we consider geometric hypergraphs encoding admissible configurations:

Theorem 4.2. *Let $P \subset S^d$ be an admissible spherical configuration with k points. Then $\mathbf{m}_{S^d}(P)$ is the value of the following conic optimization program:*

$$\begin{aligned} \sup\ & \int_{(S^d)^k} F(x_1, \ldots, x_k)\, d\sigma^k(x_1, \ldots, x_k) \\ & \int_{(S^d)^{k-1}} F(x_1, \ldots, x_{k-1}, x_{k-1})\, d\sigma^{k-1}(x_1, \ldots, x_{k-1}) = 1, \\ & F(x_1, \ldots, x_k) = 0 \quad \text{if } \{x_1, \ldots, x_k\} \simeq P, \\ & F \in C((S^d)^k) \text{ is completely positive.} \end{aligned} \tag{4.5}$$

Moreover, the optimal value is attained as a maximum.

Theorem 4.3. *Let $R > 0$ and suppose $P \subset \mathbb{R}^d$ is an admissible configuration with k points. Then $\mathbf{m}_{Q(0,R)}(P)$ is the value of the following conic optimization program:*

$$\begin{aligned} \sup\ & R^{-d} \int_{Q(0,R)^k} F(x_1, \ldots, x_k)\, dx_1 \ldots dx_k \\ & \int_{Q(0,R)^{k-1}} F(x_1, \ldots, x_{k-1}, x_{k-1})\, dx_1 \ldots dx_{k-1} = 1, \\ & F(x_1, \ldots, x_k) = 0 \quad \text{if } \{x_1, \ldots, x_k\} \simeq P, \\ & F \in C(Q(0,R)^k) \text{ is completely positive.} \end{aligned} \tag{4.6}$$

Remark 4.4. *Although we have stated the results in the case of only one forbidden configuration, similar programs can be obtained when forbidding multiple admissible configurations. The proof that these programs give exact formulations for the corresponding independence density remains essentially unchanged.*

The formal proof of these results will occupy the next three sections, but we can quickly give an overview of its main ingredients. On a high level, it follows the same general strategy as the proof of the completely positive formulation for finite hypergraphs; however, the fact that the program is now infinite-dimensional introduces some additional technical complications which must be dealt with. For concreteness let us now concentrate on the spherical case, the Euclidean case being very similar.

Following the proof of Theorem 4.1, given some P-avoiding measurable set $A \subset S^d$ we are led to consider the function

$$F_A(x_1, \ldots, x_k) = \frac{1}{\sigma(A)^{k-1}} \prod_{i=1}^{k} \chi_A(x_i).$$

This function is clearly completely positive, and it is easy to see that it satisfies the first two constraints of program (4.5), and that it has objective value $\sigma(A)$; unfortunately F_A is not continuous, as is required in the third constraint. The idea to get around this problem is to consider instead an ‘averaged’ version of F_A, which has the same good properties as the original function but is now continuous; this is done in Section 4.3.

Showing that the optimal value of program (4.5) is at most $\mathbf{m}_{S^d}(P)$ is somewhat more complicated. Since the space of functions considered is infinite-dimensional, we can no longer decompose a completely positive function into a finite conic combination as we did in equation (4.3). The first thing we will show is that any *continuous* completely positive function F can be arbitrarily well-approximated (in the supremum norm) by a finite-rank completely positive function $\tilde{F}$; this is done in Section 4.4.

Unfortunately, after passing to a finite-rank approximation, we can no longer guarantee that the function $\tilde{F}$ we are dealing with is zero on edges of the hypergraph; this was the crucial property we needed to conclude the proof in the finite case. In order to deal with the errors introduced by our approximation, we shall use the Supersaturation theorem (Theorem 3.9); this is the key step in our proof. A suitable application of this result in conjunction with the arguments used for finite hypergraphs will show that, if the objective function is noticeably larger than $\mathbf{m}_{S^d}(P)$, then *on average* the value of $\tilde{F}$ on edges of H must be noticeably larger than 0; but this would contradict the fact that $\tilde{F}$ is a good approximation of F in the supremum norm, which finishes the proof. The details are shown in Section 4.5.

One can also obtain a similar completely positive program in the case of the non-compact Euclidean space $\mathbb{R}^d$, but for this we will need to consider some limits. Given a continuous function $F \in C((\mathbb{R}^d)^k)$ on $k \geq 2$ variables, we write

$$\begin{aligned} M_k(F) &:= \limsup_{T\to\infty} \frac{1}{(\operatorname{vol}[-T,T]^d)^k} \int_{([-T,T]^d)^k} F(x_1,\ldots,x_k)\,dx_1\ldots dx_k \quad \text{and} \\ M_{k-1}(F) &:= \limsup_{T\to\infty} \frac{1}{(\operatorname{vol}[-T,T]^d)^{k-1}} \int_{([-T,T]^d)^{k-1}} F(x_1,\ldots,x_{k-1},x_{k-1})\,dx_1\ldots dx_{k-1}; \end{aligned}$$

these are continuous linear functionals on $C((\mathbb{R}^d)^k)$ endowed with the supremum norm. As a corollary to Theorem 4.3 we will obtain:

Corollary 4.5. *If $P \subset \mathbb{R}^d$ is an admissible configuration on k points, then $\mathbf{m}_{\mathbb{R}^d}(P)$ is the value of the following conic optimization program:*

$$\begin{aligned} \sup\ & M_k(F) \\ & M_{k-1}(F) = 1, \\ & F(x_1,\ldots,x_k) = 0 \quad \text{if } \{x_1,\ldots,x_k\} \simeq P, \\ & F \in C((\mathbb{R}^d)^k) \text{ is completely positive.} \end{aligned} \tag{4.7}$$

Moreover, the optimal value is attained as a maximum.

4.2.1 The case of forbidden distances

Let us now turn to the special case of forbidden angles and distances, which correspond to two-point configurations. In this case, the programs (4.5) and (4.7) can be further simplified and written in a more convenient form; this was first done by DeCorte, Oliveira and Vallentin [22], who showed (using different arguments) that the corresponding completely positive formulations are exact.

In the spherical setting, we first exploit the invariance of program (4.5) under the action of $\mathrm{O}(d+1)$ to show that the optimization variable F can also be required to be invariant under this action. More precisely, if we forbid two-point configurations on S^d forming angle θ, we obtain from (4.5) the equivalent program

$$\begin{aligned} \sup\ & \textstyle\int_{S^d}\int_{S^d} F(x,y)\,d\sigma(x)\,d\sigma(y) \\ & \textstyle\int_{S^d} F(x,x)\,d\sigma(x) = 1, \\ & F(x,y) = 0 \quad \text{if } x\cdot y = \cos\theta, \\ & F \in C(S^d \times S^d) \text{ is completely positive and } \mathrm{O}(d+1)\text{-invariant.} \end{aligned}$$

The equivalence can be easily seen by transforming any given feasible solution F of program (4.5) into the 'averaged' function

$$\tilde{F}(x,y) := \int_{\mathrm{O}(d+1)} F(Tx,Ty)\,d\mu(T),$$

which is a feasible solution of the program above and has the same objective value.

As the orbit of a pair of points $(x,y) \in S^d \times S^d$ under $\mathrm{O}(d+1)$ depends only on their inner product $x\cdot y$, it follows that any $\mathrm{O}(d+1)$-invariant function $F : S^d \times S^d \to \mathbb{R}$ can be written as $F(x,y) = f(x\cdot y)$ for some suitable function $f : [-1,1] \to \mathbb{R}$; moreover, if F is continuous then its corresponding univariate function f is also continuous. We shall say that a continuous function $f : [-1,1] \to \mathbb{R}$ is of *completely positive type for* S^d if the kernel $(x,y) \mapsto f(x\cdot y)$, where $x, y \in S^d$, is completely positive. The last program can then be equivalently written in the more convenient single-variable form:

$$\begin{aligned} \sup\ & \textstyle\int_{S^d}\int_{S^d} f(x\cdot y)\,d\sigma(y)\,d\sigma(x) \\ & f(1) = 1, \\ & f(\cos\theta) = 0, \\ & f \in C([-1,1]) \text{ is of completely positive type for } S^d. \end{aligned} \tag{4.8}$$

Obtaining a single-variable program for distance-avoiding sets in Euclidean space is similar, but slightly more complicated due to the fact that $\mathbb{R}^d$ is not compact.

A continuous function $f : \mathbb{R}^d \to \mathbb{R}$ is of *completely positive type* if for every finite set $U \subset \mathbb{R}^d$ the matrix $\big(f(x-y)\big)_{x,y\in U}$ is completely positive. A continuous function $f : \mathbb{R}^d \to \mathbb{R}$ of completely positive type has a well-defined *mean value*

$$M(f) = \lim_{T\to\infty} \frac{1}{\operatorname{vol}[-T,T]^d} \int_{[-T,T]^d} f(x)\,dx;$$

this is a consequence of Bochner's theorem, as shown in [22].

Note that program (4.7) is invariant under translation, so it is natural to restrict it to functions which are also translation-invariant. Any translation-invariant completely positive function $F : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ can be written as $F(x,y) = f(x-y)$ for some function $f : \mathbb{R}^d \to \mathbb{R}$ of completely positive type; under this identification, one can easily check that $M_2(F) = M(f)$ and $M_1(F) = f(0)$. Restricting program (4.7) for sets avoiding distance 1 to translation-invariant functions, we then obtain

$$\begin{aligned} \sup\ & M(f) \\ & f(0) = 1, \\ & f(x) = 0 \quad \text{if } \|x\| = 1, \\ & f \in C(\mathbb{R}^d) \text{ is of completely positive type.} \end{aligned} \tag{4.9}$$

By Corollary 4.5, the value of this restricted program is at most $\mathbf{m}_{\mathbb{R}^d}(1)$. Moreover, it is easy to adapt the proof of this corollary in order to show that *periodic* sets give rise to translation-invariant feasible solutions of (4.7); it suffices to average over its fundamental domain when using Lemma 4.8. Since the extremal value $\mathbf{m}_{\mathbb{R}^d}(1)$ can be arbitrarily well-approximated by periodic sets, this shows that the value of program (4.9) is at least $\mathbf{m}_{\mathbb{R}^d}(1)$ (though this value might no longer be attained).

We conclude from this discussion the following result, which was first proven by DeCorte, Oliveira and Vallentin [22]:

Corollary 4.6. *The optimal value of program (4.8) is $\mathbf{m}_{S^d}(\theta)$, and it is attained as a maximum. Moreover, the optimal value of program (4.9) is $\mathbf{m}_{\mathbb{R}^d}(1)$.*

We remark that the main result of DeCorte, Oliveira and Vallentin also applies to a more general class of 'well-behaved' infinite graphs, of which distance graphs on $\mathbb{R}^d$ and S^d are the main examples. Their methods were also different from those used here in some essential ways; in particular, they did not use any supersaturation result but instead relied on deeper analytic arguments.

4.3 Averaging and continuity

In this section we prove two technical lemmas designed to 'regularize' a function by taking averages; this will be necessary when constructing feasible solutions to our optimization programs from configuration-avoiding sets, and will be used again next chapter.

The first such result we consider is for functions on the sphere:

Lemma 4.7. *For all functions $f_1, \dots, f_k \in L^k(S^d)$, the function*

$$\Lambda_k(f_1, \dots, f_k) : \ (x_1, \dots, x_k) \mapsto \int_{\mathrm{O}(d+1)} \prod_{i=1}^{k} f_i(Tx_i)\, d\mu(T)$$

is continuous on $(S^d)^k$.

Proof. For fixed points $x_1, \dots, x_k \in S^d$, we can apply Hölder's inequality $k-1$ times to obtain

$$\left| \int_{\mathrm{O}(d+1)} \prod_{i=1}^{k} f_i(Tx_i)\, d\mu(T) \right| \leq \prod_{i=1}^{k} \left(\int_{\mathrm{O}(d+1)} |f_i(Tx_i)|^k \, d\mu(T) \right)^{1/k}.$$

Since each Tx_i is uniformly distributed on S^d when T is uniformly distributed on $O(d+1)$, it follows that the right-hand side of the inequality above is equal to $\prod_{i=1}^{k} \|f_i\|_k$; we conclude that

$$\|\Lambda_k(f_1, \dots, f_k)\|_\infty \leq \prod_{i=1}^{k} \|f_i\|_k. \tag{4.10}$$

For each $1 \leq i \leq k$, take a sequence $(f_{i,n})_{n\geq 1}$ of continuous functions on S^d converging to f_i in L^k norm. Since S^d is compact, it follows that each of these functions $f_{i,n}$ is *uniformly* *continuous*; so is their product $(x_1, \ldots, x_k) \mapsto \prod_{i=1}^k f_{i,n}(x_i)$ (for any fixed $n \geq 1$). As the action of $\mathrm{O}(d+1)$ preserves distances, it follows that the average

$$\Lambda_k(f_{1,n},\, f_{2,n},\, \ldots,\, f_{k,n}) = \int_{\mathrm{O}(d+1)} \prod_{i=1}^k f_{i,n}(T\,\cdot\,)\, d\mu(T)$$

is also (uniformly) continuous.

Since Λ_k is multilinear, we have from (4.10) that

$$\begin{aligned}
&\|\Lambda_k(f_1,\, f_2,\, \ldots,\, f_k) - \Lambda_k(f_{1,n},\, f_{2,n},\, \ldots,\, f_{k,n})\|_\infty \\
&\qquad \leq \|\Lambda_k(f_1 - f_{1,n},\, f_2,\, \ldots,\, f_k)\|_\infty + \|\Lambda_k(f_{1,n},\, f_2 - f_{2,n},\, \ldots,\, f_k)\|_\infty \\
&\qquad\qquad + \cdots + \|\Lambda_k(f_{1,n},\, f_{2,n},\, \ldots,\, f_k - f_{k,n})\|_\infty \\
&\qquad \leq \sum_{i=1}^k \prod_{j=1}^{i-1} \|f_{j,n}\|_k \cdot \|f_i - f_{i,n}\|_k \cdot \prod_{j=i+1}^k \|f_j\|_k \xrightarrow{n\to\infty} 0.
\end{aligned}$$

Thus $(\Lambda_k(f_{1,n}, \ldots, f_{k,n}))_{n\geq 1}$ is a sequence of continuous functions converging uniformly to $\Lambda_k(f_1, \ldots, f_k)$; it follows that the function $\Lambda_k(f_1, \ldots, f_k)$ is also continuous. □

We will also need an analogue of this result for functions on the Euclidean space:

Lemma 4.8. *For all $R > 0$ and all functions $f_1, \ldots, f_k \in L^\infty(\mathbb{R}^d)$, the function*

$$\Lambda_k(f_1, \ldots, f_k) :\ (x_1, \ldots, x_k) \mapsto \int_{Q(0,\,R)} \prod_{i=1}^k f_i(x_i + y)\, dy$$

is continuous on $(\mathbb{R}^d)^k$.

Proof. Fix $x_1, \ldots, x_k \in \mathbb{R}^d$, and denote by F the restriction of $\Lambda_k(f_1, \ldots, f_k)$ to the set $\prod_{i=1}^k Q(x_i,\, 1) \subset (\mathbb{R}^d)^k$; it suffices to show that F is continuous. Note that F depends only on the restriction of each function f_i to $Q(x_i,\, R+1)$, so we may assume that $\operatorname{supp} f_i \subseteq Q(x_i,\, R+1)$ for $1 \leq i \leq k$.

For each $1 \leq i \leq k$, take a sequence $(f_{i,n})_{n\geq 1}$ of continuous functions supported on $Q(x_i,\, R+1)$, converging to f_i in L^k norm. Let us first show that $\Lambda_k(f_{1,n},\, f_{2,n},\, \ldots,\, f_{k,n})$ is continuous on $(\mathbb{R}^d)^k$, for each $n \geq 1$.

Fixed $n \geq 1$, since each $f_{i,n}$ is continuous and compactly supported, it follows that they are *uniformly continuous*; so is their product $\prod_{i=1}^k f_{i,n}$. Thus, for every $\varepsilon > 0$ there is $\delta_n(\varepsilon) > 0$ such that

$$\max_{1\leq i\leq k} \|z_i - z_i'\| \leq \delta_n(\varepsilon) \implies \left| \prod_{i=1}^k f_{i,n}(z_i) - \prod_{i=1}^k f_{i,n}(z_i') \right| \leq \varepsilon.$$

We conclude that, whenever $\max_{1\leq i\leq k} \|z_i - z_i'\| \leq \delta_n(\varepsilon)$, we have

$$\begin{aligned}
&|\Lambda_k(f_{1,n},\, \ldots,\, f_{k,n})(z_1, \ldots, z_k) - \Lambda_k(f_{1,n},\, \ldots,\, f_{k,n})(z_1', \ldots, z_k')| \\
&\qquad = \left| \int_{Q(0,\,R))} \left(\prod_{i=1}^k f_{i,n}(z_i + y) - \prod_{i=1}^k f_{i,n}(z_i' + y) \right) dy \right| \\
&\qquad \leq \int_{Q(0,\,R))} \varepsilon\, dy = \varepsilon R^d.
\end{aligned}$$

It follows that $\Lambda_k(f_{1,n}, \ldots, f_{k,n})$ is continuous.

Applying Hölder's inequality $k-1$ times we see that, for any functions $g_1, \ldots, g_k$ in $L^k(\mathbb{R}^d)$ and any $z_1, \ldots, z_k \in \mathbb{R}^d$, we have

$$\int_{\mathbb{R}^d} \left| \prod_{i=1}^k g_i(z_i + y) \right| dy \leq \prod_{i=1}^k \left(\int_{\mathbb{R}^d} |g_i(z_i + y)|^k \, dy \right)^{1/k} = \prod_{i=1}^k \|g_i\|_k;$$

this shows that $\|\Lambda_k(g_1, \ldots, g_k)\|_\infty \leq \prod_{i=1}^k \|g_i\|_k$. Since Λ_k is multilinear, we conclude that

$$\begin{aligned}
&\|\Lambda_k(f_1, f_2, \ldots, f_k) - \Lambda_k(f_{1,n}, f_{2,n}, \ldots, f_{k,n})\|_\infty \\
&\qquad \leq \|\Lambda_k(f_1 - f_{1,n}, f_2, \ldots, f_k)\|_\infty + \|\Lambda_k(f_{1,n}, f_2 - f_{2,n}, \ldots, f_k)\|_\infty \\
&\qquad\qquad + \cdots + \|\Lambda_k(f_{1,n}, f_{2,n}, \ldots, f_k - f_{k,n})\|_\infty \\
&\qquad \leq \sum_{i=1}^k \prod_{j=1}^{i-1} \|f_{j,n}\|_k \cdot \|f_i - f_{i,n}\|_k \cdot \prod_{j=i+1}^k \|f_j\|_k \xrightarrow{n\to\infty} 0.
\end{aligned}$$

Thus $(\Lambda_k(f_{1,n}, \ldots, f_{k,n}))_{n\geq 1}$ is a sequence of continuous functions converging uniformly to F inside $\prod_{i=1}^k Q(x_i, 1)$; it follows that F is also continuous. □

4.4 Finite-rank approximations

Our next task is to approximate continuous completely positive functions by finite-rank completely positive functions. We will do so in the setting of functions defined on compact metric spaces, though this assumption is not strictly necessary.

For the next lemma we then assume that V is a compact metric space, endowed with its Borel σ-algebra. The distance between two points $x, y \in V$ is denoted by $d(x, y)$.

Lemma 4.9. *Suppose $F \in C(V^k)$ is completely positive. Then for every $\varepsilon > 0$ there exist nonnegative functions $f_1, \ldots, f_N \in L^\infty(V)$ such that*

$$\left\| F - \sum_{j=1}^N f_j^{\otimes k} \right\|_\infty \leq \varepsilon.$$

Proof. Fix $\varepsilon > 0$. Since F is continuous on the compact metric space V^k, there exists $\delta > 0$ for which

$$\max_{1\leq i\leq k} d(x_i, x_i') \leq \delta \implies |F(x_1, \ldots, x_k) - F(x_1', \ldots, x_k')| \leq \varepsilon.$$

Now take a finite δ-net Y of V, that is, a finite subset $Y \subseteq V$ such that

$$\min\{d(x, y) : y \in Y\} \leq \delta \quad \text{for all } x \in V.$$

By assumption, the k-array $(F(x_1, \ldots, x_k))_{x_1,\ldots,x_k\in Y}$ is completely positive; there must then exist nonnegative functions $\phi_1, \ldots, \phi_N : Y \to \mathbb{R}_+$ for which

$$F(x_1, \ldots, x_k) = \sum_{j=1}^N \prod_{i=1}^k \phi_j(x_i) \quad \text{for all } x_1, \ldots, x_k \in Y.$$

Let $y_1, \dots, y_m$ be an (arbitrary) ordering of the points in Y. For each $1 \le j \le N$ we define the nonnegative function $f_j : V \to \mathbb{R}_+$ by $f_j(x) = \phi_j(y_k)$, where $y_k \in Y$ is the point in Y which is closest to x, and with ties broken in favour of the smallest index.

These functions f_j are clearly bounded and (Borel) measurable. Note that F coincides with $\sum_{j=1}^N f_j^{\otimes k}$ on Y^k, and by definition of the set Y we conclude that

$$\left| F(x_1, \dots, x_k) - \sum_{j=1}^N \prod_{i=1}^k f_j(x_i) \right| \le \varepsilon$$

for all $x_1, \dots, x_k \in V$. The lemma follows. □

4.5 Proof of the completely positive formulations

We can now show that the completely positive formulations given in Section 4.2 are exact, following the proof outline presented in that section.

We start by proving Theorem 4.2 on the independence density of admissible spherical configurations. The other two results are proven through minor variations of the arguments used in this case, and will be given afterwards.

Proof of Theorem 4.2. Denote by ν the optimum value of program (4.5); we wish to show that $\nu = \mathbf{m}_{S^d}(P)$, and that this optimum value is attained.

The proof that $\nu \ge \mathbf{m}_{S^d}(P)$ is simple. By Theorem 3.17, there exists an extremal P-avoiding measurable set $A \subset S^d$, that is, satisfying $\sigma(A) = \mathbf{m}_{S^d}(P)$; consider the function

$$F_A(x_1, \dots, x_k) := \frac{1}{\sigma(A)^{k-1}} \int_{\mathrm{O}(d+1)} \prod_{i=1}^k \chi_A(Tx_i)\, d\mu(T).$$

By Lemma 4.7, this function is continuous on $(S^d)^k$. It is also completely positive: for any finite set $U \subset S^d$, the cone of completely positive k-arrays on U is closed as a subset of $\mathbb{R}^{U^k}$, and so the array $(F_A(u_1, \dots, u_k))_{u_1, \dots, u_k \in U}$ is completely positive. Moreover, one can easily check that F_A satisfies the other two constraints of program (4.5) and has objective value $\sigma(A)$, thus showing that $\nu \ge \mathbf{m}_{S^d}(P)$ and that this lower bound can be attained.

In order to show that $\nu \le \mathbf{m}_{S^d}(P)$, fix $\varepsilon > 0$ and let $0 < \delta < \varepsilon$ be a small constant to be chosen later. Let F be a feasible solution of program (4.5) with objective value at least $\nu - \delta$. By Lemma 4.9, there are functions $f_1, \dots, f_N \ge 0$ such that

$$\left\| F - \sum_{j=1}^N f_j^{\otimes k} \right\|_\infty \le \delta.$$

From this bound and our assumptions on F, we immediately obtain:

$$\sum_{j=1}^N \|f_j\|_1^k = \int_{(S^d)^k} \left(\sum_{j=1}^N f_j^{\otimes k}(x_1, \dots, x_k) \right) d\sigma^k(x_1, \dots, x_k) \ge \nu - 2\delta,$$

$$\sum_{j=1}^N \prod_{i=1}^k f_j(x_i) = \sum_{j=1}^N f_j^{\otimes k}(x_1, \dots, x_k) \le \delta \quad \text{whenever } \{x_1, \dots, x_k\} \simeq P, \text{ and}$$

$$\sum_{j=1}^{N} \|f_j\|_1^{k-2}\|f_j\|_2^2 = \int_{(S^d)^{k-1}} \left(\sum_{j=1}^{N} f_j^{\otimes k}(x_1,\ldots,x_{k-1},x_{k-1})\right) d\sigma^{k-1}(x_1,\ldots,x_{k-1}) \in [1-\delta,\ 1+\delta].$$

It will be convenient to perform some normalizations with these functions f_j, so that the sum $\sum_{j=1}^{N} \|f_j\|_1^{k-2}\|f_j\|_2^2$ on the last expression is exactly 1 and the functions we will work with have L^2 norm equal to 1. In order to do this, let us denote $C := \sum_{j=1}^{N} \|f_j\|_1^{k-2}\|f_j\|_2^2$, and define the constants $\lambda_j := \|f_j\|_1^{k-2}\|f_j\|_2^2/C$ and the functions $g_j := f_j/\|f_j\|_2$ for each $1 \le j \le N$. Then each g_j is non-negative, satisfies $\|g_j\|_2 = 1$ and

$$\sum_{j=1}^{N} \lambda_j \|g_j\|_1^2 = \frac{1}{C}\sum_{j=1}^{N} \|f_j\|_1^k \ge \nu - 4\delta, \tag{4.11}$$

$$\sum_{j=1}^{N} \lambda_j \prod_{i=1}^{k} g_j(x_i) \le \frac{1}{C}\sum_{j=1}^{N}\prod_{i=1}^{k} f_j(x_i) \le 2\delta \qquad \text{whenever } \{x_1,\ldots,x_k\} \simeq P, \tag{4.12}$$

$$\sum_{j=1}^{N} \lambda_j = \frac{1}{C}\sum_{j=1}^{N} \|f_j\|_1^{k-2}\|f_j\|_2^2 = 1. \tag{4.13}$$

(We have used that $\|f_j\|_1 \le \|f_j\|_2$ for the first inequality in (4.12), which holds since we are in a probability space.)

Equation (4.13) can be used to define a probability measure $\mathbb{P}$ on the set of indices $\{1,2,\ldots,N\}$, given by $\mathbb{P}(j) = \lambda_j$; inequality (4.11) may then be written as $\mathbb{E}[\|g_j\|_1^2] \ge \nu - 4\delta$. For each $1 \le j \le N$, let us define the set $A_j := \{x \in S^d : g_j(x) \ge \varepsilon\}$ corresponding to points where g_j is not too small; then

$$\begin{aligned}\nu - 4\delta \le \mathbb{E}[\|g_j\|_1^2] &= \mathbb{E}[(\langle g_j, \chi_{A_j}\rangle + \langle g_j, \chi_{S^d\setminus A_j}\rangle)^2] \\ &\le \mathbb{E}[(\|g_j\|_2\,\|\chi_{A_j}\|_2 + \varepsilon)^2] \\ &\le \mathbb{E}[\sigma(A_j)] + 3\varepsilon,\end{aligned}$$

where we used that $\|g_j\|_2 = 1$. Since $\delta \le \varepsilon$, this implies that

$$\mathbb{E}[\sigma(A_j)] \ge \nu - 7\varepsilon. \tag{4.14}$$

Let us now show that $\mathbb{E}[\sigma(A_j)] \le \mathbf{m}_{S^d}(P) + 2\varepsilon$ if δ is chosen small enough. Suppose this is false, so that $\mathbb{E}[\sigma(A_j)] > \mathbf{m}_{S^d}(P) + 2\varepsilon$; by averaging, this implies that

$$\mathbb{P}(\sigma(A_j) \ge \mathbf{m}_{S^d}(P) + \varepsilon) > \varepsilon.$$

Integrating inequality (4.12) over all congruent copies of P on S^d, and then using that $g_j \ge \varepsilon \cdot \chi_{A_j}$ for all j, we obtain

$$\begin{aligned}2\delta \ge \sum_{j=1}^{N} \lambda_j I_P(g_j) &= \mathbb{E}[I_P(g_j)] \\ &\ge \mathbb{E}[\varepsilon^k I_P(A_j)] \\ &\ge \varepsilon^k\, \mathbb{P}(\sigma(A_j) \ge \mathbf{m}_{S^d}(P) + \varepsilon)\, \mathbb{E}[I_P(A_j) \mid \sigma(A_j) \ge \mathbf{m}_{S^d}(P) + \varepsilon] \\ &\ge \varepsilon^{k+1}\, \mathbb{E}[I_P(A_j) \mid \sigma(A_j) \ge \mathbf{m}_{S^d}(P) + \varepsilon].\end{aligned}$$

By the spherical Supersaturation theorem (Theorem 3.9), this last expression is at least $\varepsilon^{k+1}c(\varepsilon) > 0$. But this cannot happen if we choose $\delta := \varepsilon^{k+1}c(\varepsilon)/3$, thus proving that for this value of δ we have $\mathbb{E}[\sigma(A_j)] \leq \mathbf{m}_{S^d}(P) + 2\varepsilon$.

Combining this last bound with inequality (4.14), we conclude that $\mathbf{m}_{S^d}(P) \geq \nu - 9\varepsilon$; since $\varepsilon > 0$ is arbitrary, this finishes the proof. □

The proof of the completely positive formulation for $\mathbf{m}_{Q(0,R)}(P)$ is very similar:

Proof of Theorem 4.3. Given some $\delta > 0$ and a P-avoiding set $A \subseteq Q(0, R)$ of positive measure, consider the function

$$F_A^\delta(x_1, \dots, x_k) := \frac{1}{\operatorname{vol}(A)^{k-1}} \frac{1}{\delta^d} \int_{Q(0,\delta)} \prod_{i=1}^{k} \chi_A(x_i + y)\, dy.$$

By Lemma 4.8, this function is continuous; by the same argument as before, it is also completely positive. Note that

$$\begin{aligned}
&\int_{Q(0,R)^{k-1}} F_A^\delta(x_1, \dots, x_{k-1}, x_{k-1})\, dx_1 \dots dx_{k-1} \\
&\qquad = \frac{1}{\operatorname{vol}(A)^{k-1}} \frac{1}{\delta^d} \int_{Q(0,\delta)} \prod_{i=1}^{k-1} \left(\int_{Q(0,R)} \chi_A(x_i + y)\, dx_i \right) dy \\
&\qquad = \frac{1}{\operatorname{vol}(A)^{k-1}} \frac{1}{\delta^d} \int_{Q(0,\delta)} \operatorname{vol}(A \cap Q(y, R))^{k-1}\, dy.
\end{aligned}$$

Since $\operatorname{vol}(A) - d\delta R^{d-1} \leq \operatorname{vol}(A \cap Q(y, R)) \leq \operatorname{vol}(A)$ for all $y \in Q(0, \delta)$, it follows that the last expression lies between $(1 - d\delta R^{d-1}/\operatorname{vol}(A))^{k-1}$ and 1. Likewise, we obtain

$$\left(1 - \frac{d\delta R^{d-1}}{\operatorname{vol}(A)}\right)^k d_{Q(0,R)}(A) \leq \frac{1}{R^d} \int_{Q(0,R)^k} F_A^\delta(x_1, \dots, x_k)\, dx_1 \dots dx_k \leq d_{Q(0,R)}(A).$$

Taking $\delta > 0$ small enough, we conclude that an appropriate multiple of F_A^δ will be a feasible solution of program (4.5) with objective value as close as we wish to $d_{Q(0,R)}(A)$. This shows that $\nu \geq \mathbf{m}_{Q(0,R)}(P)$.

The proof of the converse direction is almost exactly the same as the one presented in the spherical case; the only differences are that we now use the supersaturation property given in Lemma 2.9, and that the value of δ we take in the arguments must also depend on the (fixed) parameter R. □

Finally, we now make use of Theorem 4.3 and similar arguments as before in order to show that the completely positive formulation for $\mathbf{m}_{\mathbb{R}^d}(P)$ is exact:

Proof of Corollary 4.5. We first show that there is a feasible solution of program (4.7) with objective value $\mathbf{m}_{\mathbb{R}^d}(P)$. By Theorem 2.18, there is a measurable P-avoiding set $A \subset \mathbb{R}^d$ with well-defined density $d(A)$ which attains $d(A) = \mathbf{m}_{\mathbb{R}^d}(P)$; consider the function

$$F_A(x_1, \dots, x_k) := \frac{1}{d(A)^{k-1}} \int_{Q(0,1)} \prod_{i=1}^{k} \chi_A(x_i + y)\, dy.$$

By the same argument as in the last proof, we see that F_A is a feasible solution of program (4.7), and has objective value $M_k(F_A) = d(A) = \mathbf{m}_{\mathbb{R}^d}(P)$.

For the converse direction, take any feasible solution F of program (4.7). Given $\varepsilon > 0$, let $R_0 > 0$ be a number such that

$$\frac{1}{(R^d)^{k-1}} \int_{Q(0,R)^{k-1}} F(x_1, \ldots, x_{k-1}, x_{k-1})\, dx_1 \ldots dx_{k-1} \leq M_{k-1}(F) + \varepsilon = 1 + \varepsilon \tag{4.15}$$

holds for all $R \geq R_0$. Now, for a given $R \geq R_0$, denote by $F_R : Q(0, R) \to \mathbb{R}$ the restriction of F to $Q(0, R)$, normalized so that

$$\int_{Q(0,R)^{k-1}} F_R(x_1, \ldots, x_{k-1}, x_{k-1})\, dx_1 \ldots dx_{k-1} = 1.$$

This function F_R is clearly a feasible solution of program (4.5), so by Theorem 4.2 its objective value is

$$\frac{1}{R^d} \int_{Q(0,R)^k} F_R(x_1, \ldots, x_k)\, dx_1 \ldots dx_k \leq \mathbf{m}_{Q(0,R)}(P).$$

By inequality (4.15), this implies that

$$\frac{1}{(R^d)^k} \int_{Q(0,R)^k} F(x_1, \ldots, x_k)\, dx_1 \ldots dx_k \leq (1 + \varepsilon)\, \mathbf{m}_{Q(0,R)}(P).$$

Since this holds for all $R \geq R_0$ and $\mathbf{m}_{Q(0,R)}(P) \to \mathbf{m}_{\mathbb{R}^d}(P)$ as $R \to \infty$, we conclude that $M_k(F) \leq (1 + \varepsilon)\, \mathbf{m}_{\mathbb{R}^d}(P)$. The corollary follows. □

4.6 Remarks and open problems

An obvious question that presents itself is whether the condition that the configuration considered is admissible is truly necessary. The reason we needed this assumption in our proofs was the use of a suitable Supersaturation theorem, whose proof in turn uses the corresponding Counting lemma. 2 However, the Supersaturation theorem looks weaker (and thus less stringent) than the Counting lemma, and it is conceivable that it should remain true even when the Counting lemma does not hold (for instance, when in $\mathbb{R}^d$ and the configuration considered is nonspherical). Our proof should remain essentially unchanged whenever there is a suitable supersaturation result for the configuration (or configurations) being avoided.

More generally, one can ask for which infinite hypergraphs the corresponding completely positive formulation is exact, that is, equal to the corresponding notion of independence ratio. It seems that here again the bottleneck is the existence of a supersaturation result for the edges of the hypergraph; both the process of constructing feasible completely positive solutions from measurable independent sets (albeit by making use of Urysohn's lemma and losing attainability as maximum) and the construction of finite-rank approximations can be made much more general.

We note that DeCorte, Oliveira and Vallentin [22] did not rely on any supersaturation result to obtain their completely positive formulations (4.8) and (4.9) for the independence ratio of distance graphs. Instead, they relied on a better understanding of the cone of completely positive kernels and on deeper analytic arguments. It might be possible to adapt their methods also to 'well-behaved' infinite hypergraphs, and thus attain a greater generality than that obtained from our methods.

Chapter Five

A recursive Lovász theta number for simplex-avoiding sets

Regular simplices being the simplest 'higher-dimensional' point configurations, we shall devote the present chapter to computing good upper bounds for their independence densities, both in Euclidean space and on the sphere.

This will be done by recursively extending the *Lovász theta number* from graphs to geometric hypergraphs encoding such configurations, and then analyzing the resulting optimization program. We follow the paper "A recursive Lovász theta number for simplex-avoiding sets" [12], which is joint work of the author with Fernando de Oliveira Filho, Lucas Slot and Frank Vallentin.

5.1 The theta number of graphs and generalizations

The Lovász theta number $\vartheta(G)$ of a finite graph G can be defined as a relaxation of the (exact) completely positive program (4.1) to the cone of positive semidefinite matrices; more explicitly, $\vartheta(G)$ is equal to the optimal value of

$$
\begin{aligned}
\max\ & \textstyle\sum_{i,j\in V(G)} M_{ij} \\
& \textstyle\sum_{i\in V(G)} M_{ii} = 1, \\
& M_{ij} = 0 \quad \text{if } \{i,j\} \in E(G), \\
& M \in \mathbb{R}^{V(G)\times V(G)} \text{ is positive semidefinite.}
\end{aligned}
$$

It is known that this parameter satisfies $\alpha(G) \le \vartheta(G) \le \chi(\overline{G})$, where $\alpha(G)$ is the independence number of G and $\chi(\overline{G})$ is the chromatic number of the complement of G (that is, the edges in $\overline{G}$ are the non-edges of G and vice versa); the theta number can be computed efficiently using semidefinite programming.

Originally, Lovász [46] introduced ϑ to determine the Shannon capacity of the 5-cycle. The theta number turned out to be a very versatile tool in optimization, with applications in combinatorics and geometry. It is related to spectral bounds like Hoffman's bound, as noted by Lovász in his paper, and also to Delsarte's linear programming bound in coding theory, as observed independently by McEliece, Rodemich, and Rumsey [48] and Schrijver [58].

Bachoc, Nebe, Oliveira, and Vallentin [4] extended ϑ to infinite geometric graphs on compact metric spaces. They also showed that this extension leads to the classical linear programming bound for spherical codes of Delsarte, Goethals, and Seidel [25]; the linear programming bound of Cohn and Elkies for the sphere-packing density [13] can also be seen as an appropriate extension of ϑ [19, 21]. These many applications illustrate the power of the Lovász theta number as a unifying concept in optimization; Goemans [38] even remarked that "it seems all paths lead to ϑ!".

We will show how a recursive variant of ϑ can be used to find upper bounds for the independence ratio of geometric hypergraphs encoding regular simplices on the sphere and in Euclidean space; this will lead to upper bounds for the independence densities of these configurations, and also to new bounds for a problem in Euclidean Ramsey theory.

5.1.1 Simplex-encoding hypergraphs

A set of $k+1 \geq 2$ points $\{x_1, \ldots, x_{k+1}\}$ in Euclidean space $\mathbb{R}^d$ is a *unit k-simplex* if $\|x_i - x_j\| = 1$ for all $i \neq j$. We shall denote a unit k-simplex (up to congruence) by $\mathcal{S}_k$; then $\mathbb{R}^d$ contains copies of $\mathcal{S}_k$ whenever $k \leq d$.

Similarly, we shall call a set of $k + 1 \geq 2$ points $\{x_1, \ldots, x_{k+1}\}$ in the d-dimensional unit sphere S^d a *(k, t)-simplex* if $x_i \cdot x_j = t$ for all $i \neq j$. Note that a (k, t)-simplex is congruent to $(2 - 2t)^{1/2}\mathcal{S}_k$, but it will be more convenient to work with inner products rather than Euclidean distance when we are on the unit sphere. There is a (k, t)-simplex in S^d for every $k \leq d$ and $t \in [-1/k, 1)$.

We are interested in obtaining good upper bounds on the independence density of unit k-simplices in $\mathbb{R}^d$ and of (k, t)-simplices in S^d.

On the unit sphere

Fix $d \geq k \geq 2$ and $t \in [-1/k, 1)$. For notational convenience, let us denote

$$\alpha_{S^d}(k, t) := \sup\{\, \sigma(I) : I \subseteq S^d \text{ is measurable and avoids } (k, t)\text{-simplices}\,\};$$

note that $\alpha_{S^d}(k, t) = \mathbf{m}_{S^d}((2 - 2t)^{1/2}\mathcal{S}_k)$. This is the independence ratio of the hypergraph whose vertex set is S^d and whose edges are all (k, t)-simplices.

In Section 5.2 we shall define the parameter $\vartheta(S^d, k, t)$ recursively as the optimal value of the problem

$$\begin{aligned}
\sup\ & \int_{S^d}\int_{S^d} f(x \cdot y)\, d\sigma(y)\, d\sigma(x)\\
& f(1) = 1,\\
& f(t) \leq \vartheta(S^{d-1}, k-1, t/(1+t)),\\
& f \in C([-1, 1]) \text{ is a function of positive type for } S^d
\end{aligned}$$

for $k \geq 2$. The base of the recursion is $k = 1$: $\vartheta(S^d, 1, t)$ is the optimal value of the problem above when the constraint "$f(t) \leq \vartheta(S^{d-1}, k-1, t/(1+t))$" is replaced by "$f(t) = 0$".

From Theorem 5.2 below it follows that $\vartheta(S^d, k, t) \geq \alpha_{S^d}(k, t)$. Using extremal properties of ultraspherical polynomials, an explicit formula can be computed for this bound, as we show in Theorem 5.3.

In Euclidean space

Transferring these concepts from the compact unit sphere to the non-compact Euclidean space requires a bit of care; this is done in Section 5.3.

As in the spherical setting, the parameter

$$\mathbf{m}_{\mathbb{R}^d}(\mathcal{S}_k) = \sup\{\, \overline{d}(I) : I \subseteq \mathbb{R}^d \text{ is measurable and avoids unit } k\text{-simplices}\,\}$$

has an interpretation in terms of a hypergraph on the Euclidean space, and again we can bound the independence ratio of this hypergraph from above by an appropriately defined parameter $\vartheta(\mathbb{R}^d, k)$. Theorem 5.5 gives an explicit expression for $\vartheta(\mathbb{R}^d, k)$ in terms of Bessel functions and ultraspherical polynomials.

5.1.2 The Main Theorem and combinatorial consequences

In Section 5.4, we analyze the upper bounds $\vartheta(S^d, k, t)$ for simplex-avoiding sets on the sphere and $\vartheta(\mathbb{R}^d, k)$ for simplex-avoiding sets in Euclidean space by using properties of ultraspherical polynomials, obtaining the following theorem.

Theorem 5.1. *Let $k \geq 1$ be an integer. Then:*

(i) for every $t \in (0, 1)$, there is a constant $c = c(k, t) \in (0, 1)$ such that $\vartheta(S^d, k, t) \leq (c + o(1))^d$;

(ii) there is a constant $c = c(k) \in (0, 1)$ such that $\vartheta(\mathbb{R}^d, k) \leq (c + o(1))^d$.

From this result we obtain exponentially decaying upper bounds on the independence densities $\mathbf{m}_{\mathbb{R}^d}(\mathcal{S}_k)$ and $\mathbf{m}_{S^d}(\ell\mathcal{S}_k)$, for $0 < \ell < \sqrt{2}$. Estimates for the exponential bases $c(k, t)$ and $c(k)$ will be computed in Section 5.4.1. We will also compute the value of $\vartheta(\mathbb{R}^d, k)$ for all $2 \leq k \leq d \leq 10$ (see Table 5.1), thus obtaining numerical upper bounds for the corresponding values of $\mathbf{m}_{\mathbb{R}^d}(\mathcal{S}_k)$.

Euclidean Ramsey theory

The central question of Euclidean Ramsey theory is: given a finite configuration P of points in $\mathbb{R}^d$ and an integer $r \geq 1$, does every r-coloring of $\mathbb{R}^d$ contain a monochromatic congruent copy of P?

The simplest point configurations are unit k-simplices, which are known to have the exponential Ramsey property: exponentially many colors (in the dimension) are needed to avoid monochromatic unit k-simplices in $\mathbb{R}^d$. This was first proved by Frankl and Wilson [36] for $k = 1$ and by Frankl and Rödl [34] for $k \geq 2$. Results in this area are usually proved by the linear algebra method; see also Sagdeev [54].

Recently, Naslund [49] used the slice-rank method from the work of Croot, Lev, and Pach [15] and Ellenberg and Gijswijt [30] on the cap-set problem[1] to prove that

$$\chi(\mathbb{R}^d, 2) \geq (1.01466 + o(1))^d,$$

[1] The slice-rank method is only implicit in the original works; the actual notion of slice-rank for a tensor was introduced by Terence Tao in a blog post (https://terrytao.wordpress.com/2016/05/18/a-symmetric-formulation-of-the-croot-lev-pach-ellenberg-gijswijt-capset-bound/).

where $\chi(\mathbb{R}^d, 2)$ is the minimum number of colors needed to color the points of $\mathbb{R}^d$ in such a way that the vertices of any equilateral triangle of unit side length do not all get the same color. This is the best lower bound known at the moment.

For simplices of higher dimension there are also explicit lower bounds known for the chromatic number $\chi(\mathbb{R}^d, k)$. The current best was obtained by Sagdeev [55] using a quantitative version of the Frankl-Rödl theorem, and reads

$$\chi(\mathbb{R}^d, k) \geq \left(1 + \frac{1}{2^{2^{k+4}}} + o(1)\right)^d.$$

If we restrict our colorings to have measurable color classes, then from Theorem 5.1 we get an exponential lower bound for $\chi_{\mathrm{m}}(\mathbb{R}^d, k)$, the measurable counterpart of $\chi(\mathbb{R}^d, k)$, due to the simple inequality

$$\chi_{\mathrm{m}}(\mathbb{R}^d, k) \cdot \mathbf{m}_{\mathbb{R}^d}(\mathcal{S}_k) \geq 1.$$

Rigorous estimates of the constant c in the theorem then yield significantly better lower bounds than those currently known.

Indeed, in the case $k = 2$ we obtain (see Section 5.4.1)

$$\mathbf{m}_{\mathbb{R}^d}(\mathcal{S}_2) \leq (0.95622 + o(1))^d,$$

and so

$$\chi_{\mathrm{m}}(\mathbb{R}^d, 2) \geq (1.04578 + o(1))^d.$$

We also obtain the rougher estimate

$$\mathbf{m}_{\mathbb{R}^d}(\mathcal{S}_k) \leq \left(1 - \frac{1}{9k^2} + o(1)\right)^d,$$

valid for all $k \geq 2$, which immediately implies

$$\chi_{\mathrm{m}}(\mathbb{R}^d, k) \geq \left(1 + \frac{1}{9k^2} + o(1)\right)^d.$$

5.2 Simplex-avoiding sets on the sphere

We call a continuous kernel $K\colon S^d \times S^d \to \mathbb{R}$ *positive* if for every finite set $U \subseteq S^d$ the matrix $(K(x, y))_{x,y \in U}$ is positive semidefinite. A continuous function $f\colon [-1, 1] \to \mathbb{R}$ is of *positive type for S^d* if the kernel $(x, y) \mapsto f(x \cdot y)$, where $x, y \in S^d$, is positive.

Fix $d \geq k \geq 2$ and $t \in [-1/k,\ 1)$. For any $\gamma \geq 0$, consider the optimization problem:

$$\begin{aligned} \sup \int_{S^d} \int_{S^d} & f(x \cdot y)\, d\sigma(y)\, d\sigma(x) \\ & f(1) = 1, \\ & f(t) \leq \gamma, \\ & f \in C([-1, 1]) \text{ is of positive type for } S^d. \end{aligned} \tag{5.1}$$

Theorem 5.2. *Fix $d \geq k \geq 2$, $t \in [-1/k,\ 1)$. If $\gamma \geq \alpha_{S^{d-1}}(k-1,\ t/(1+t))$, then the optimal value of* (5.1) *is an upper bound for $\alpha_{S^d}(k,\ t)$.*

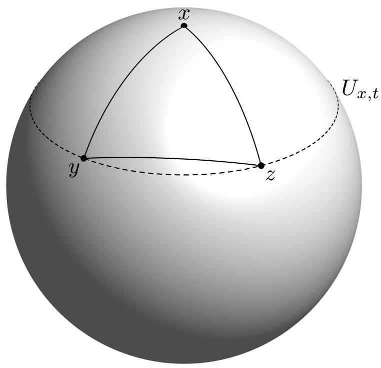

Figure 5.1: The set $U_{x,t}$ with $t = 0.5$ and a $(2, t)$-simplex on S^2.

Proof. Let $I \subseteq S^d$ be a measurable set that avoids (k, t)-simplices and assume $\sigma(I) > 0$. Consider the kernel $K\colon S^d \times S^d \to \mathbb{R}$ such that

$$K(x, y) = \int_{\mathrm{O}(d+1)} \chi_I(Tx)\chi_I(Ty)\, d\mu(T),$$

where χ_I is the characteristic function of I and where μ is the Haar measure on $\mathrm{O}(d + 1)$.

By Lemma 4.7, we have that K is continuous. By construction, K is also positive and $\mathrm{O}(d + 1)$-invariant, that is, $K(Tx, Ty) = K(x, y)$ for all $T \in \mathrm{O}(d + 1)$ and $x, y \in S^d$. Such kernels are known to be of the form $K(x, y) = g(x \cdot y)$, where $g \in C([-1, 1])$ is of positive type for S^d. Note that

$$K(x, x) = \int_{\mathrm{O}(d+1)} \chi_I(Tx)\, d\mu(T) = \sigma(I),$$

so $g(1) = \sigma(I) > 0$.

We set $f = g/g(1)$. Immediately we have that f is continuous and of positive type and that $f(1) = 1$; moreover

$$\int_{S^d}\int_{S^d} f(x \cdot y)\, d\sigma(y)\, d\sigma(x) = \sigma(I).$$

Hence, if we show that $f(t) \leq \gamma$, the theorem will follow.

If $x \in S^d$ is a point in a (k, t)-simplex, all other points in the simplex are in $U_{x,t} = \{ y \in S^d : y \cdot x = t \}$ (see Figure 5.1). Note that $U_{x,t}$ is a $(d - 1)$-dimensional sphere with radius $(1 - t^2)^{1/2}$; let ν be the surface measure on $U_{x,t}$ normalized so the total measure is 1. If $T \in \mathrm{O}(d + 1)$ is any orthogonal matrix, then TI avoids (k, t)-simplices. Hence if $x \in TI$, then $TI \cap U_{x,t}$ cannot contain k points with pairwise inner product t. After applying a transformation $U_{x,t} \to S^{d-1}$, we thus see that $\nu(TI \cap U_{x,t}) \leq \alpha_{S^{d-1}}(k - 1, t/(1 + t)) \leq \gamma$. Indeed, if two points in $U_{x,t}$ have inner product t, a small computation shows that the corresponding points in S^{d-1} have inner product $t/(1 + t)$.

Now fix $x \in S^d$ and note that

$$\begin{aligned} g(t) = \int_{U_{x,t}} K(x,y)\,d\nu(y) &= \int_{U_{x,t}} \int_{\mathrm{O}(d+1)} \chi_I(Tx)\chi_I(Ty)\,d\mu(T)\,d\nu(y) \\ &= \int_{\mathrm{O}(d+1)} \chi_I(Tx) \int_{U_{x,t}} \chi_I(Ty)\,d\nu(y)\,d\mu(T) \\ &\leq \gamma\sigma(I), \end{aligned}$$

whence $f(t) \leq \gamma$, and we are done. □

One obvious choice for γ in Problem (5.1) is the bound given by the same problem for $(k-1,\ t/(1+t))$-simplices. The base for the recursion is $k = 1$: then we need an upper bound for the measure of a set of points on the sphere that avoids pairs of points with a fixed inner product. Such a bound was given by Bachoc, Nebe, Oliveira, and Vallentin [4] and looks very similar to (5.1). They show that, for $d \geq 1$ and $t \in [-1, 1)$, the optimal value of the following optimization problem is an upper bound for $\alpha_{S^d}(1,\ t)$:

$$\begin{aligned} &\sup \textstyle\int_{S^d}\int_{S^d} f(x\cdot y)\,d\sigma(y)\,d\sigma(x) \\ &\quad f(1) = 1, \\ &\quad f(t) = 0, \\ &\quad f \in C([-1,1]) \text{ is a function of positive type for } S^d. \end{aligned} \tag{5.2}$$

(Note that this also follows immediately from our Corollary 4.6.)

Let $\vartheta(S^d,\ 1,\ t)$ denote the optimal value of the optimization problem above, so that $\vartheta(S^d,\ 1,\ t) \geq \alpha_{S^d}(1,\ t)$. For $k \geq 2$ and $t \in [-1/k,\ 1)$, let $\vartheta(S^d,\ k,\ t)$ be the optimal value of Problem (5.1) when $\gamma = \vartheta(S^{d-1},\ k-1,\ t/(1+t))$. We then have

$$\vartheta(S^d,\ k,\ t) \geq \alpha_{S^d}(k,\ t) = \mathbf{m}_{S^d}((2-2t)^{1/2}\mathcal{S}_k).$$

There is actually a simple analytical expression for $\vartheta(S^d,\ k,\ t)$, as we see now. For $d \geq 1$ and $j \geq 0$, let P_j^d denote the Jacobi polynomial with parameters $\alpha = \beta = (d-2)/2$ and degree j, normalized so that $P_j^d(1) = 1$ (for background on Jacobi polynomials, see the book by Szegö [60]).

In Theorem 6.2 of Bachoc, Nebe, Oliveira, and Vallentin [4] it is shown that for every $t \in [-1, 1)$ there is some $j \geq 0$ such that $P_j^d(t) < 0$. Theorem 8.21.8 in the book by Szegö [60] implies that, for every $t \in (-1, 1)$,

$$\lim_{j\to\infty} P_j^d(t) = 0.$$

Hence, for every $t \in (-1, 1)$ we can define

$$M_d(t) = \min\{\, P_j^d(t) : j \geq 0 \,\}, \tag{5.3}$$

and we see that $M_d(t) < 0$. With this we have [4, Theorem 6.2]

$$\vartheta(S^d,\ 1,\ t) = \frac{-M_d(t)}{1 - M_d(t)}.$$

The expression for $\vartheta(S^d,\ k,\ t)$ is very similar:

Theorem 5.3. *If $d \ge k \ge 2$ and $t \in [-1/k,\ 1)$, then*

$$\vartheta(S^d,\ k,\ t) = \frac{\vartheta(S^{d-1},\ k-1,\ t/(1+t)) - M_d(t)}{1 - M_d(t)}. \tag{5.4}$$

The proof requires the following characterization of functions of positive type due to Schoenberg [57]: a function $f\colon [-1,1] \to \mathbb{R}$ is continuous and of positive type for S^d if and only if there are nonnegative numbers $f_0, f_1, \ldots$ such that $\sum_{j=0}^{\infty} f_j < \infty$ and

$$f(t) = \sum_{j=0}^{\infty} f_j P_j^d(t), \tag{5.5}$$

with uniform convergence in $[-1,1]$.

Proof of Theorem 5.3. The orthogonality of the Jacobi polynomials P_j^d implies in particular that, if $j \ge 1$, then

$$\int_{S^d}\int_{S^d} P_j^d(x \cdot y)\, d\sigma(y)\, d\sigma(x) = 0.$$

Use this and Schoenberg's characterization of positive type functions to rewrite (5.1), obtaining the equivalent problem

$$\begin{aligned}
&\sup f_0\\
&\quad \textstyle\sum_{j=0}^{\infty} f_j = 1,\\
&\quad \textstyle\sum_{j=0}^{\infty} f_j P_j^d(t) \le \vartheta(S^{d-1},\ k-1,\ t/(1+t)),\\
&\quad f_j \ge 0 \text{ for all } j \ge 0.
\end{aligned}$$

To solve this problem, note that

$$\sum_{j=0}^{\infty} f_j P_j^d(t)$$

is a convex combination of the numbers $P_j^d(t)$. We want to keep this convex combination below $\vartheta(S^{d-1},\ k-1,\ t/(1+t))$ while maximizing f_0. The best way to do so is to concentrate all the weight of the combination on f_0 and f_{j^*}, where j^* is such that $P_{j^*}^d(t)$ is the most negative number appearing in the convex combination, that is, $P_{j^*}^d(t) = M_d(t)$. Now solve the problem using only the variables f_0 and f_{j^*} to get the optimal value as given in the statement of the theorem. □

The expression for $\vartheta(S^d,\ k,\ 0)$ is particularly simple. Indeed, for $d \ge 1$ it follows from the recurrence relation for the Jacobi polynomials that $M_d(0) = P_2^d(0) = -1/d$, whence

$$\vartheta(S^d,\ k,\ 0) = k/(d+1).$$

Figure 5.2 shows the behavior of $\vartheta(S^d,\ 2,\ t)$ for a few values of d as t changes. Plots for $k \ge 3$ are very similar.

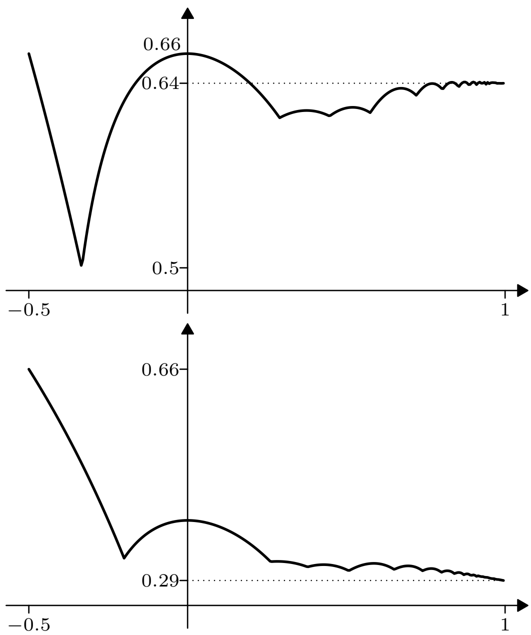

Figure 5.2: Plots of $\vartheta(S^d, 2, t)$ for $t \in [-0.5, 1)$ and $d = 2$ (top) and 4 (bottom).

5.3 Simplex-avoiding sets in Euclidean space

An optimization problem similar to (5.1) provides an upper bound for $\mathbf{m}_{\mathbb{R}^d}(\mathcal{S}_k)$. To introduce it, we need some definitions and facts from harmonic analysis on $\mathbb{R}^d$; for background, see e.g. the book by Reed and Simon [52].

A continuous function $f : \mathbb{R}^d \to \mathbb{R}$ is of *positive type* if for every finite set $U \subseteq \mathbb{R}^d$ the matrix $(f(x - y))_{x,y \in U}$ is positive semidefinite. A continuous function $f : \mathbb{R}^d \to \mathbb{R}$ of positive type has a well-defined *mean value*

$$M(f) = \lim_{T \to \infty} \frac{1}{\mathrm{vol}[-T, T]^d} \int_{[-T,T]^d} f(x)\, dx.$$

We say that a function $f : \mathbb{R}^d \to \mathbb{R}$ is *radial* if $f(x)$ depends only on $\|x\|$. In this case, for $t \geq 0$ we denote by $f(t)$ the common value of f for vectors of norm t.

Fix integers $d \geq k \geq 2$. For every $\gamma \geq 0$, consider the optimization problem

$$\begin{aligned} &\sup M(f) \\ &\quad f(0) = 1, \\ &\quad f(1) \leq \gamma, \\ &\quad f : \mathbb{R}^d \to \mathbb{R} \text{ is continuous, radial, and of positive type.} \end{aligned} \tag{5.6}$$

We have the analogue of Theorem 5.2:

Theorem 5.4. *Fix integers $d \geq k \geq 2$. If $\gamma \geq \alpha_{S^{d-1}}(k-1, 1/2)$, then the optimal value of* (5.6) *is an upper bound for $\mathbf{m}_{\mathbb{R}^d}(\mathcal{S}_k)$.*

Proof. Let $I \subseteq \mathbb{R}^d$ be a measurable set of positive upper density avoiding unit k-simplices. The first step is to see that we can assume that I is periodic with periodicity lattice $R\mathbb{Z}^d$, for some $R > 0$. Indeed, fix $R > 1/2$. Erase a border of width $1/2$ around $I \cap [-R, R]^d$ and paste the resulting set periodically in such a way that there is an empty gap of width 1 between any two pasted copies. The resulting periodic set still avoids unit k-simplices and is measurable. Its upper density is

$$\frac{\operatorname{vol}(I \cap [-R + 1/2,\, R - 1/2]^d)}{\operatorname{vol}([-R,\, R]^d)};$$

by taking R large enough, we can make this density as close as we want to the upper density of I.

Assume then that I is periodic with periodicity lattice $R\mathbb{Z}^d$, so its characteristic function χ_I is also periodic. Set

$$g(x) = \frac{1}{R^d} \int_{Q(0,\, R)} \chi_I(y)\chi_I(x + y)\, dy.$$

Lemma 4.8 applied to χ_I implies that g is continuous. Direct verification yields that g is of positive type. Moreover, $g(0) = \overline{d}(I)$ and $M(g) = \overline{d}(I)^2$.

Now set

$$f(x) = \overline{d}(I)^{-1} \int_{\mathrm{O}(d)} g(Tx)\, d\mu(T),$$

where μ is the Haar measure on $\mathrm{O}(d)$. Note that f is continuous, radial, and of positive type. Moreover, $f(0) = 1$ and $M(f) = \overline{d}(I)$. If we show that $f(1) \leq \gamma$, then f is a feasible solution of (5.6) with $M(f) = \overline{d}(I)$, and so the theorem will follow.

To see that $f(1) \leq \gamma$, note that if x is a point of a unit k-simplex in $\mathbb{R}^d$, then all the others points in the simplex lie on the unit sphere $x + S^{d-1}$ centered at x. Hence if $x \in I$, then $I \cap (x + S^{d-1})$ is a measurable subset of $x + S^{d-1}$ that avoids $(k - 1,\, 1/2)$-simplices, and so the measure of $I \cap (x + S^{d-1})$ as a subset of the unit sphere is at most $\alpha_{S^{d-1}}(k - 1,\, 1/2)$. So if $\xi \in \mathbb{R}^d$ is any unit vector, then

$$\begin{aligned}
f(1) &= \overline{d}(I)^{-1} \int_{\mathrm{O}(d)} g(T\xi)\, d\mu(T) \\
&= \overline{d}(I)^{-1} \int_{\mathrm{O}(d)} \frac{1}{R^d} \int_{Q(0,\, R)} \chi_I(x)\chi_I(T\xi + x)\, dx\, d\mu(T) \\
&= \overline{d}(I)^{-1} \frac{1}{R^d} \int_{Q(0,\, R)} \chi_I(x) \int_{\mathrm{O}(d)} \chi_I(T\xi + x)\, d\mu(T)\, dx \\
&\leq \alpha_{S^{d-1}}(k - 1,\, 1/2) \leq \gamma,
\end{aligned}$$

as we wanted. □

Denote by $\vartheta(\mathbb{R}^d,\, k)$ the optimal value of (5.6) when setting $\gamma = \vartheta(S^{d-1},\, k - 1,\, 1/2)$. Then $\vartheta(\mathbb{R}^d,\, k) \geq \mathbf{m}_{\mathbb{R}^d}(\mathcal{S}_k)$.

An expression akin to the one for $\vartheta(S^d,\, k,\, t)$ can be derived for $\vartheta(\mathbb{R}^d,\, k)$. For $d \geq 2$ and $u \geq 0$, let

$$\Omega_d(u) = \Gamma(d/2)\,(2/u)^{(d-2)/2}\, J_{(d-2)/2}(u),$$

$d \setminus k$	2	3	4	5	6	7	8	9	10
2	0.64355	—	—	—	—	—	—	—	—
3	0.42849	0.69138	—	—	—	—	—	—	—
4	0.29346	0.49798	0.73225	—	—	—	—	—	—
5	0.20374	0.36768	0.55035	0.76580	—	—	—	—	—
6	0.15225	0.28471	0.42777	0.60262	0.79563	—	—	—	—
7	0.11866	0.22740	0.34071	0.48493	0.64681	0.81972	—	—	—
8	0.09339	0.18405	0.27471	0.39559	0.53374	0.68268	0.83882	—	—
9	0.07387	0.15030	0.22864	0.33042	0.44903	0.57816	0.71431	0.85537	—
10	0.05846	0.12340	0.19194	0.27851	0.38158	0.49496	0.61521	0.74026	0.86882

Table 5.1: The bound $\vartheta(\mathbb{R}^d, k)$ for $d = 2, \dots, 10$ and $k = 2, \dots, 10$, with values of d on each row and of k on each column.

where J_α is the Bessel function of the first kind with parameter α. Let m_d be the global minimum of Ω_d, which is a negative number (cf. Oliveira and Vallentin [20]). The following theorem is the analogue of Theorem 5.3.

Theorem 5.5. *For $d \geq 2$, we have*

$$\vartheta(\mathbb{R}^d, k) = \frac{\vartheta(S^{d-1}, k-1, 1/2) - m_d}{1 - m_d}.$$

The proof uses again a theorem of Schoenberg [56], that this time characterizes radial and continuous functions of positive type on $\mathbb{R}^d$: these are the functions $f\colon \mathbb{R}^d \to \mathbb{R}$ such that

$$f(x) = \int_0^\infty \Omega_d(z\|x\|)\, d\nu(z) \tag{5.7}$$

for some finite Borel measure ν.

Proof. If f is given as in (5.7), then $M(f) = \nu(\{0\})$ (see e.g. §6.2 in DeCorte, Oliveira, and Vallentin [22]). Using Schoenberg's theorem, we can rewrite (5.6) (with γ being $\vartheta(S^{d-1}, k-1, 1/2)$) equivalently as:

$$\begin{aligned} &\sup \nu(\{0\}) \\ &\quad \nu([0,\infty)) = 1, \\ &\quad \textstyle\int_0^\infty \Omega_d(z)\, d\nu(z) \leq \vartheta(S^{d-1}, k-1, 1/2), \\ &\quad \nu \text{ is a Borel measure.} \end{aligned}$$

We are now in the same situation as in the proof of Theorem 5.3. If z^* is such that $m_d = \Omega_d(z^*)$, then the optimal ν is supported at 0 and z^*. Solving the resulting system yields the theorem. □

Table 5.1 contains some values for $\vartheta(\mathbb{R}^d, k)$, which are then upper bounds for $\mathbf{m}_{\mathbb{R}^d}(\mathcal{S}_k)$.

5.4 Exponential density decay

In this section we analyze the asymptotic behavior of $\vartheta(S^d, k, t)$ and $\vartheta(\mathbb{R}^d, k)$ as functions of d, proving Theorem 5.1.

The main step in our analysis is to understand the asymptotic behavior of

$$M_d(t) = \min\{P_j^d(t) : j \geq 0\},$$

as defined in (5.3). For $t \in [-1, 0)$, it is possible to show that $\lim_{d\to\infty} M_d(t) = t$, and so $M_d(t)$ does not approach 0. We have seen in Section 5.2 that $M_d(0) = -1/d$, so for $t = 0$ we have that $M_d(t)$ approaches 0 linearly fast as d grows. Things get interesting when $t \in (0, 1)$: then $M_d(t)$ approaches 0 exponentially fast as d grows.

Theorem 5.6. *For every $t \in (0, 1)$ there is $c \in (0, 1)$ such that $|M_d(t)| \leq (c + o(1))^d$.*

We will need the following lemma, whose proof is a refinement of the analysis carried out by Schoenberg [57].

Lemma 5.7. *If for $\theta \in (0, \pi)$ and $\delta \in (0, \pi/2)$ we write*

$$C = (\cos^2\theta + \sin^2\theta \sin^2\delta)^{1/2},$$

then $|P_j^d(\cos\theta)| \leq \pi d^{1/2} \cos^{d-3}\delta + C^j$ for all $d \geq 3$.

Proof. An integral representation of Gegenbauer for the ultraspherical polynomials (take $\lambda = (d-2)/2$ in Theorem 6.7.4 from Andrews, Askey, and Roy [2]) gives us the formula

$$P_j^d(\cos\theta) = R(d)^{-1} \int_0^\pi F(\phi)^j \sin^{d-3}\phi \, d\phi,$$

where

$$F(\phi) = \cos\theta + i\sin\theta\cos\phi \qquad \text{and} \qquad R(d) = \int_0^\pi \sin^{d-3}\phi \, d\phi.$$

Note that $|F(\phi)|^2 = \cos^2\theta + \sin^2\theta\cos^2\phi$ and that $|F(\phi)| \leq 1$. Split the integration domain into the intervals $[0, \pi/2 - \delta]$, $[\pi/2 - \delta, \pi/2 + \delta]$, and $[\pi/2 + \delta, \pi]$ to obtain

$$\begin{aligned}
|P_j^d(\cos\theta)| &\leq R(d)^{-1} \int_0^\pi |F(\phi)|^j \sin^{d-3}\phi \, d\phi \\
&\leq 2R(d)^{-1} \int_0^{\pi/2-\delta} \sin^{d-3}\phi \, d\phi + R(d)^{-1} \int_{\pi/2-\delta}^{\pi/2+\delta} |F(\phi)|^j \sin^{d-3}\phi \, d\phi.
\end{aligned}$$

For the first term above, note that

$$R(d) = \frac{\pi^{1/2}\Gamma(d/2 - 1)}{\Gamma((d-1)/2)}.$$

Use Gautschi's inequality to conclude that $R(d) > \pi^{1/2}((d-1)/2)^{-1/2}$, and hence

$$R(d)^{-1} < \pi^{-1/2}((d-1)/2)^{1/2} < d^{1/2}.$$

Now

$$\begin{aligned}
2R(d)^{-1}\int_0^{\pi/2-\delta}\sin^{d-3}\phi\,d\phi &\le 2d^{1/2}\int_0^{\pi/2-\delta}\sin^{d-3}(\pi/2-\delta)\,d\phi\\
&= 2d^{1/2}(\pi/2-\delta)\cos^{d-3}\delta\\
&\le \pi d^{1/2}\cos^{d-3}\delta.
\end{aligned}$$

For the second term we get directly

$$R(d)^{-1}\int_{\pi/2-\delta}^{\pi/2+\delta}|F(\phi)|^j\sin^{d-3}\phi\,d\phi \le R(d)^{-1}\int_{\pi/2-\delta}^{\pi/2+\delta}C^j\sin^{d-3}\phi\,d\phi \le C^j,$$

and we are done. □

We can now prove the theorem.

Proof of Theorem 5.6. Our strategy is to find a lower bound on the largest j_0 such that $P_j^d(t) \ge 0$ for all $j \le j_0$. Then we know that $M_d(t)$ is attained by some $j \ge j_0$, and we can use Lemma 5.7 to estimate $|M_d(t)|$.

Recall [60] that the zeros of P_j^d are all in $[-1, 1]$ and that the rightmost zero of P_{j+1}^d is to the right of the rightmost zero of P_j^d. Let C_j^λ denote the ultraspherical (or Gegenbauer) polynomial with parameter λ and degree j, so

$$P_j^d(t) = \frac{C_j^{(d-2)/2}(t)}{C_j^{(d-2)/2}(1)}. \tag{5.8}$$

Let x_j be the largest zero of C_j^λ. Elbert and Laforgia [29, p. 94] show that, for $\lambda \ge 0$,

$$x_j^2 < \frac{j^2+2\lambda j}{(j+\lambda)^2}.$$

If for a given j we have that

$$\frac{j^2+2\lambda j}{(j+\lambda)^2} \le t^2, \tag{5.9}$$

then we know that the rightmost zero of C_j^λ is to the left of t, and so $C_j^\lambda(t) \ge 0$.

Note that the left-hand side in (5.9) is increasing in j. Let us estimate the largest j for which (5.9) holds. We want

$$j^2 + 2\lambda j - t^2(j+\lambda)^2 \le 0.$$

The left-hand side above is quadratic in j, and since $t^2 < 1$ the coefficient of j^2 is positive. So all we have to do is to compute the largest root of the left-hand side, which is $2a(t)\lambda$, where $a(t) = ((1-t^2)^{-1/2} - 1)/2$.

Hence for $j \le 2a(t)\lambda$ we have $C_j^\lambda(t) \ge 0$. From (5.8) we see that $P_j^d(t) \ge 0$ if

$$j \le a(t)d - 2a(t).$$

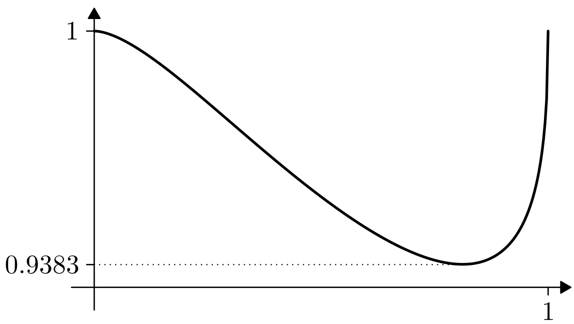

Figure 5.3: The best constant c obtained in our proof of Theorem 5.6, for each value of $t \in (0, 1)$.

Now plug the right-hand side above into the upper bound of Lemma 5.7 to get

$$\begin{aligned}|M_d(t)| &\leq (\pi d^{1/2} \cos^{-3} \delta) \cos^d \delta + C^{a(t)d-2a(t)} \\ &= O(d^{1/2}) \cos^d \delta + O(1)(C^{a(t)})^d,\end{aligned}$$

with C as defined in Lemma 5.7 with $\cos\theta = t$. For any choice of $\delta \in (0, \pi/2)$, we have that $\cos\delta$, $C \in (0, 1)$, and since $a(t) > 0$ for all $t \in (0, 1)$, the theorem follows. □

The constant c given in Theorem 5.6 depends on t. Following the proof, we can find the best constant for every $t \in (0, 1)$ by finding $\delta \in (0, \pi/2)$ such that $\cos\delta = C^{a(t)}$. This leads to the equation

$$\cos^4 \delta = (t^2 + (1 - t^2) \sin^2 \delta)^{(1-t^2)^{-1/2}-1}; \tag{5.10}$$

we can then take $c = \cos\delta > 0$. Figure 5.3 shows a plot of the best constant c for every $t \in (0, 1)$.

We now obtain exponential decay for $\vartheta(S^d, k, t)$, for any $k \geq 2$ and $t \in (0, 1)$. Indeed, consider the recurrence $F_0 = t$ and $F_i = F_{i-1}/(1 + F_{i-1})$ for $i \geq 1$, whose solution is $F_i = t/(1 + it)$. Using Theorem 5.6 to develop our analytic solution (5.4), we get

$$\vartheta(S^d, k, t) \sim \sum_{i=0}^{k-1} |M_{d-i}(F_i)| = \sum_{i=0}^{k-1} |M_{d-i}(t/(1 + it))|. \tag{5.11}$$

(We write $a_n \sim b_n$ to mean that $\lim_{n\to\infty} a_n/b_n = 1$.) Since $t/(1 + it) > 0$ for all i, each term decays exponentially fast, and so we get exponential decay for the sum.

We also get exponential decay for $\vartheta(\mathbb{R}^d, k)$ and any $k \geq 2$, since from Theorem 5.5 we have that

$$\vartheta(\mathbb{R}^d, k) \sim |m_d| + \sum_{i=0}^{k-2} |M_{d-i}(1/(2 + i))|. \tag{5.12}$$

From Theorem 5.6 we know that every term in the summation above decays exponentially fast. Oliveira and Vallentin [20] give an asymptotic bound for m_d that shows that $|m_d|$ also decays exponentially in d, namely

$$|m_d| \le (2/e + o(1))^{d/2} = (0.8577\ldots + o(1))^d.$$

This finishes the proof of Theorem 5.1.

5.4.1 Explicit bounds

We now compute explicit constants $c(k, t)$ and $c(k)$ which can serve as bases for the exponentials in Theorem 5.1, in particular obtaining the bounds advertised in Section 5.1.2.

Our main tool is the estimate given in (5.10) for the constant in Theorem 5.6. Writing c for the (unique) positive solution $\cos\delta$ of this equation and taking $\theta \in (0, \pi/2)$ such that $\cos\theta = t$, we can rewrite (5.10) in the more convenient form

$$c^{4\sin\theta/(1-\sin\theta)} = 1 - c^2 \sin^2\theta. \tag{5.13}$$

For this value of c we have from Theorem 5.6 that $|M_d(t)| \le (c + o(1))^d$. Using the asymptotics (5.11) and (5.12), we can then find good upper bounds for the constants $c(k, t)$ and $c(k)$, respectively.

For specific values of θ equation (5.13) is quite easy to solve numerically; for instance, for $\theta = \pi/3$ (corresponding to $t = 1/2$) we obtain $c = 0.95621\ldots$, and so $|M_d(1/2)| \le (0.95622 + o(1))^d$. We can also get a rougher estimate valid for all $\theta \in (0, \pi/2)$ by using Bernoulli's inequality: writing $c = 1 - x$, by Bernoulli's inequality we get

$$\begin{aligned}(1-x)^{4\sin\theta/(1-\sin\theta)} &\ge 1 - \frac{4\sin\theta}{1-\sin\theta}x,\\ 1-(1-x)^2\sin^2\theta &\le 1-(1-2x)\sin^2\theta.\end{aligned}$$

Equating the left-hand sides of both inequalities above and solving for x, we obtain

$$c = 1 - x \le 1 - \frac{\sin\theta(1-\sin\theta)}{4 + 2\sin\theta(1-\sin\theta)}.$$

In particular, when $\cos\theta = 1/k$ we obtain

$$\begin{aligned}|M_d(1/k)| &\le \left(1 - \frac{1}{4k^2(1+\sqrt{k^2/(k^2-1)})+2} + o(1)\right)^d\\ &\le \left(1 - \frac{1}{9k^2} + o(1)\right)^d\end{aligned}$$

for all $k \ge 3$ (and also for $k = 2$ by our previous bound).

We then have

$$\vartheta(\mathbb{R}^d, 2) \sim |m_d| + |M_d(1/2)| \le (0.95622 + o(1))^d$$

and, for higher values of k,

$$\vartheta(\mathbb{R}^d, k) \sim |M_{d-k+2}(1/k)| \le \left(1 - \frac{1}{9k^2} + o(1)\right)^d.$$

5.5 Remarks and open problems

Our methods to give lower bounds on the measurable chromatic numbers $\chi_{\mathrm{m}}(\mathbb{R}^d, k)$ do not immediately extend to $\chi(\mathbb{R}^d, k)$, though they easily imply the following approximate version of the general coloring problem.

Lemma 5.8. *Given any coloring of $\mathbb{R}^d$ using less than $1/\mathbf{m}_{\mathbb{R}^d}(\mathcal{S}_k)$ colors, one of the color classes must contain configurations arbitrarily close to $\mathcal{S}_k$; in fact, this color class must contain configurations arbitrarily close to $\ell\mathcal{S}_k$ for every $\ell > 0$.*

Proof. Consider any partition $\mathbb{R}^d = \bigcup_{i=1}^n V_i$ into $n < 1/\mathbf{m}_{\mathbb{R}^d}(\mathcal{S}_k)$ classes. By taking the closure of each part, we obtain a cover of $\mathbb{R}^d$ into n *measurable* sets $\overline{V_i}$; one of these sets, say $\overline{V_1}$, must have upper density at least $1/n > \mathbf{m}_{\mathbb{R}^d}(\mathcal{S}_k)$. Since $\mathbf{m}_{\mathbb{R}^d}(\ell\mathcal{S}_k) = \mathbf{m}_{\mathbb{R}^d}(\mathcal{S}_k)$ for all $\ell > 0$, it follows that $\overline{V_1}$ contains a congruent copy of each $\ell\mathcal{S}_k$. The lemma follows. $\square$

We believe that in fact $1/\mathbf{m}_{\mathbb{R}^d}(\mathcal{S}_k)$ is a lower bound also for the unrestricted chromatic number $\chi(\mathbb{R}^d, k)$:

Conjecture 5.9. $\chi(\mathbb{R}^d, k) \geq 1/\mathbf{m}_{\mathbb{R}^d}(\mathcal{S}_k)$ *holds for all integers $d \geq k \geq 1$.*

In order to obtain lower bounds for $\chi(\mathbb{R}^d, k)$, one must consider *finite* subhypergraphs of the corresponding simplex-encoding hypergraph, and then obtain lower bounds on the chromatic number of this finite hypergraph. (A result of de Bruijn and Erdős [17] implies that this chromatic number $\chi(\mathbb{R}^d, k)$ is indeed attained on a finite subhypergraph.) One way to do so is by making use of the Frankl-Rödl theorem [34] on forbidden intersections, which rather easily implies that the independence ratio of the hypergraph encoding unit simplices on $(2d)^{-1/2}\{-1, 1\}^d \subset \mathbb{R}^d$ decays exponentially with the dimension d, as long as d is divisible by four.

It might still be possible to improve the existing bounds on the unrestricted chromatic numbers $\chi(\mathbb{R}^d, k)$ by using arguments similar to those presented here, but now considering hypergraphs on the hypercube $\{-1, 1\}^d$ and using extremal properties of discrete orthogonal polynomials. We leave this as possible future work.

Chapter Six

Generalizations of the Lovász theta number to hypergraphs

In this final chapter we introduce and begin to study generalizations of the Lovász theta number to hypergraphs, extending some of the ideas given in the last chapter to a much more general setting.

The work reported in this chapter is still ongoing, and done in collaboration with Fernando de Oliveira Filho, Lucas Slot and Frank Vallentin.

6.1 The theta body and the weighted theta number

Throughout this thesis we have considered maximal independent sets in certain geometric graphs and hypergraphs. Sometimes it is also useful to consider this problem in greater generality, where the vertex set is weighted by a given nonnegative function and one wishes to know what is the maximum possible weight of an independent set.

Let G be a finite graph, and let w be a nonnegative weight function on its vertex set $V(G)$. The *weighted independence number* of G is

$$\alpha_w(G) = \max\left\{ \sum_{x \in I} w(x) : I \subseteq V(G) \text{ is independent} \right\}.$$

The task of finding this number for a given graph and weight function is a classical and very important problem in combinatorial optimization, with applications in various domains such as coding theory [69], computer vision [68] and protein structure prediction [47]. This problem is also well-known to be NP-hard, even in the case where the weight function is identically one.

A useful way of efficiently computing an upper bound for $\alpha_w(G)$ was given by Grötschel, Lovász and Schrijver [40, 41], who extended the Lovász theta number to the weighted setting. Given $w \in \mathbb{R}_+^{V(G)}$, we define the *weighted theta number* $\vartheta_w(G)$ of the graph G as the optimal value of the following optimization problem:

$$\begin{aligned} \max\ & w^T a \\ & a = \operatorname{diag} A, \\ & A(x,y) = 0 \quad \text{if } \{x,y\} \in E(G), \\ & A \in \mathbb{R}^{V\times V},\ A - a\otimes a \succeq 0. \end{aligned} \tag{6.1}$$

Here and throughout this chapter we write $M \succeq 0$ to mean that the matrix M is positive semidefinite. By taking an independent set $I \subseteq V(G)$ of maximum weight and considering the rank-one matrix $A = \chi_I \otimes \chi_I$, one immediately checks that $\alpha_w(G) \leq \vartheta_w(G)$.

As with the usual (unweighted) theta number, the weighted theta number $\vartheta_w(G)$ can be written in many different but equivalent ways; see for instance Chapter 9.3 in Grötschel, Lovász and Schrijver's book [43]. Note that the feasible region of program (6.1) does not depend on the weight function w chosen. We call this region the *theta body of* G, and denote it by $\mathrm{TH}(G)$; more explicitly,

$$\begin{aligned} \mathrm{TH}(G) = \{\, a \in \mathbb{R}^{V(G)} :\ \exists A \in \mathbb{R}^{V(G)\times V(G)} \text{ with } \operatorname{diag} A = a,\ A - a\otimes a \succeq 0, \text{ and } \\ A(x,y) = 0 \text{ for all } \{x,y\} \in E(G) \,\}. \end{aligned}$$

The theta body of a graph was first defined by Grötschel, Lovász and Schrijver [42], and it satisfies many nice properties. For instance, $\mathrm{TH}(G)$ is a convex set which encompasses the *independent set polytope* of G: this is the polytope $\mathrm{IND}(G)$ on $\mathbb{R}^{V(G)}$ defined by

$$\mathrm{IND}(G) = \operatorname{conv}\{\chi_I :\ I \subseteq V(G) \text{ is independent in } G\},$$

where conv denotes the convex hull. Such polytopes have been extensively studied in the combinatorial optimization literature (see Chapter 9 in [43]), as they encapsulate all the hardness of the weighted independence number parameter α_w. Indeed, by linearity we easily see that optimizing weight functions over all independent sets in G gives the same value as optimizing over the polytope $\mathrm{IND}(G)$, and so $\alpha_w(G) = \max\{w^T f :\ f \in \mathrm{IND}(G)\}$.

The advantage of dealing with the theta body rather than with the independent set polytope is that $\mathrm{TH}(G)$ is in some ways much easier to handle than $\mathrm{IND}(G)$. Most crucially, while optimizing linear functions over $\mathrm{IND}(G)$ is NP-hard, one can optimize linear functions over $\mathrm{TH}(G)$ in polynomial time; see Grötschel, Lovász and Schrijver [42].

6.2 Extensions to hypergraphs

We now wish to obtain similar notions of theta number and theta body which are valid for hypergraphs. In order to keep the presentation simple and avoid nonessential technicalities, we will restrict our attention to finite uniform hypergraphs. Recall that a hypergraph H is *k-uniform* if every edge of H contains exactly k vertices.

We first need to introduce some notation. Given some matrix $A \in \mathbb{R}^{V\times V}$ and an element $x \in V$, denote by A_x the function on V given by $A_x(y) := A(x,y)$; its restriction to a subset $U \subseteq V$ is denoted $A_x[U]$. Given a k-uniform hypergraph H on vertex set V and some vertex $x \in V$, we write V_x for the set of all vertices which share an edge with x; that is,

$$V_x = \{\, y \in V \setminus \{x\} :\ \exists e \in E(H) \text{ with } \{x,y\} \subseteq e \,\}.$$

We then define the *link hypergraph of H at x*, denoted H_x, as the $(k-1)$-uniform hypergraph on vertex set V_x whose edges are all sets of $k-1$ vertices which form an edge of H together with x.

It is now possible to define in a recursive manner both the theta number and the theta body of an arbitrary uniform hypergraph H. We start with the degenerate case of '1-uniform hypergraphs' (which are sets V where every element is an edge by itself), for which the theta body is defined to be $\{0\} \subset \mathbb{R}^V$.

Now suppose that the theta body of $(k-1)$-uniform hypergraphs has already been defined for some $k \geq 2$, and let H be a k-uniform hypergraph on vertex set V. Given a nonnegative weight function $w \in \mathbb{R}^V_+$, we define the *weighted theta number $\vartheta_w(H)$* of H by:

$$\begin{aligned} \max\ & w^T a \\ & a = \operatorname{diag} A, \\ & A_x[V_x] \in A(x,x)\,\mathrm{TH}(H_x) \quad \text{for } x \in V, \\ & A \in \mathbb{R}^{V\times V},\ A - a \otimes a \succeq 0. \end{aligned} \tag{6.2}$$

The theta body of H is then defined as the feasible domain of this last program, that is

$$\begin{aligned} \mathrm{TH}(H) = \{\, a \in \mathbb{R}^V :\ \exists A \in \mathbb{R}^{V\times V} \text{ with } \operatorname{diag} A = a,\ A - a \otimes a \succeq 0, \text{ and} \\ A_x[V_x] \in A(x,x)\,\mathrm{TH}(H_x) \text{ for all } x \in V \,\}. \end{aligned} \tag{6.3}$$

Note that both these definitions collapse down to the classical notions seen in the last section when considering graphs (which are 2-uniform hypergraphs).

As in the case of graphs, it is possible to show that the theta body is convex and encompasses the independent set polytope, and that our weighted theta number gives an upper bound for the weighted independence number of hypergraphs. More precisely, we have:

Theorem 6.1. *Given any uniform hypergraph H, $\mathrm{TH}(H)$ is a closed convex set satisfying the inclusions $\mathrm{IND}(H) \subseteq \mathrm{TH}(H) \subseteq [0,1]^V$. Moreover, for any nonnegative weight function $w \in \mathbb{R}^V_+$ we have that $\alpha_w(H) \leq \vartheta_w(H)$.*

Proof. We first show the inclusion $\mathrm{TH}(H) \subseteq [0,1]^V$. Take any matrix $A \in \mathbb{R}^{V\times V}$ and vector $a \in \mathbb{R}^V$ satisfying $\operatorname{diag} A = a$ and $A - a \otimes a \succeq 0$. Since $a \otimes a \succeq 0$, it follows that $A \succeq 0$ too, and so both the diagonal of A and the diagonal of $A - a \otimes a$ are nonnegative. This implies that $a(x) \geq 0$ and $a(x)(1 - a(x)) \geq 0$ for all $x \in V$; we conclude that $0 \leq a \leq 1$, and the desired inclusion follows from the definition (6.3) of the theta body.

The proof that $\mathrm{TH}(H)$ is closed, convex and contains $\mathrm{IND}(H)$ proceeds by induction on the uniformity of the hypergraph H. The base of the induction is the degenerate case of 1-uniform hypergraphs, where every vertex is a 1-edge by itself and the theta body is defined to be $\{0\} \subset \mathbb{R}^V$.

Suppose then we have already proven that $\mathrm{TH}(H')$ is a closed convex set containing $\mathrm{IND}(H')$, for some $k \geq 2$ and every $(k-1)$-uniform hypergraph H', and let H be a k-uniform hypergraph with vertex set V. Define the set of matrices

$$\begin{aligned} \mathcal{M}_H := \{\, Y \in \mathbb{R}^{(V\cup\{*\})\times(V\cup\{*\})} :\ Y(*,*) = 1,\ Y(*,x) = Y(x,x) \text{ for all } x \in V, \\ Y_x[V_x] \in Y(x,x)\,\mathrm{TH}(H_x) \text{ for all } x \in V, \text{ and } Y \succeq 0 \,\}, \end{aligned}$$

where $*$ is an additional index not contained in V. By considering principal 2×2 submatrices of some (arbitrary) matrix $Y \in \mathcal{M}_H$, one sees that all of its entries are bounded in magnitude by 1; thus $\mathcal{M}_H \subseteq [-1, 1]^{(V \cup \{*\}) \times (V \cup \{*\})}$ is bounded. Moreover, since $\mathrm{TH}(H_x)$ is convex and closed for each $x \in V$ (by the induction hypothesis), it easily follows that this set $\mathcal{M}_H$ is convex and closed as well.

Fixed some $Y \in \mathcal{M}_H$, denote by A the restriction of Y to $\mathbb{R}^{V \times V}$ and by $a \in \mathbb{R}^V$ the diagonal of A. Taking Schur complements, we see that $Y \succeq 0$ if and only if $A - a \otimes a \succeq 0$; it follows that the set of all vectors a obtained in this way, with Y ranging over $\mathcal{M}_H$, is exactly the theta body $\mathrm{TH}(H)$. We conclude that $\mathrm{TH}(H)$ is closed and convex, as the projection of the compact convex set $\mathcal{M}_H \subseteq [-1, 1]^{(V \cup \{*\}) \times (V \cup \{*\})}$ onto a linear subspace.

Now suppose $I \subseteq V$ is an independent set in H, and let us show that $\chi_I \in \mathrm{TH}(H)$. Consider the rank-one matrix $A = \chi_I \otimes \chi_I$; clearly $\operatorname{diag} A = \chi_I$ and $A - \chi_I \otimes \chi_I \succeq 0$. Since I is independent in H, it follows that $I \cap V_x$ is independent in H_x for all $x \in I$. We conclude from the induction hypothesis that $\chi_I[V_x] \in \mathrm{TH}(H_x)$ for all $x \in I$; this shows that $A_x[V_x] \in A(x, x)\,\mathrm{TH}(H_x)$ for all $x \in V$, and so $\chi_I \in \mathrm{TH}(H)$ by definition. As $\mathrm{TH}(H)$ is convex, it must also contain the convex hull of all such indicator functions χ_I with $I \subseteq V$ independent, i.e. it contains the independent set polytope $\mathrm{IND}(H)$. The inductive step (and the induction) are then completed.

Finally, since $\mathrm{IND}(H) \subseteq \mathrm{TH}(H)$, for any weight function $w \in \mathbb{R}^V_+$ it follows that

$$\alpha_w(H) = \max\{w^T f : f \in \mathrm{IND}(H)\} \le \max\{w^T f : f \in \mathrm{TH}(H)\} = \vartheta_w(H).$$

This finishes the proof of the theorem. □

The last theorem shows that $\mathrm{TH}(H)$ is a ‘relaxation’ of the independent set polytope $\mathrm{IND}(H)$, in the sense that $\mathrm{IND}(H) \subseteq \mathrm{TH}(H)$. The main advantage of working with this relaxation instead of the original polytope is that linear optimization problems over $\mathrm{TH}(H)$ can be efficiently solved, using (for instance) the ellipsoid method. More precisely, for a fixed uniformity k of the hypergraphs considered, the optimization problem $\max\{w^T f : f \in \mathrm{TH}(H)\}$ can be solved up to any accuracy parameter $\varepsilon > 0$, in time which is polynomial in the size of the input (H, w, ε).

Indeed, by the methods exposed in Chapter 4 of Grötschel, Lovász and Schrijver [41], the ‘weak optimization problem’ over $\mathrm{TH}(H)$ is polynomial-time equivalent to the ‘weak membership problem’ for $\mathrm{TH}(H)$ (note that $\mathrm{TH}(H)$ is a ‘centered convex body’ by Theorem 6.1 and the simple fact that $[0, 1/|V|]^V \subseteq \mathrm{IND}(H)$). A polynomial-time membership algorithm for $\mathrm{TH}(H)$ can be constructed recursively, using algorithms for checking membership for its link hypergraphs H_x, $x \in V$, and for checking if a given matrix is positive semidefinite.

Our next theorem informally shows that our relaxation $\mathrm{TH}(H)$ is not too much larger than $\mathrm{IND}(H)$ itself: they share the same set of integral points (i.e. vectors where all coordinates are integers).

Theorem 6.2. *For any uniform hypergraph H, the set of integral points in* $\mathrm{TH}(H)$ *coincides with the set of integral points in* $\mathrm{IND}(H)$.

Proof. Since $\mathrm{IND}(H) \subseteq \mathrm{TH}(H)$, it suffices to show that all integral points in $\mathrm{TH}(H)$ are also in $\mathrm{IND}(H)$. We will proceed by induction on the uniformity of the hypergraph H, with the base case of 1-uniform hypergraphs (whose theta body is $\{0\}$ by definition) being trivial.

Suppose we have already proven the result for some $k \geq 2$ and all $(k-1)$-uniform hypergraphs. Let H be a k-uniform hypergraph on vertex set V, and let $f \in \mathbb{Z}^V$ be an integral point of $\mathrm{TH}(H)$. By Theorem 6.1, this vector f must be $\{0,1\}$-valued; denote by $I \subseteq V$ the support of f, so that $f = \chi_I$. We will show that I is independent in H, implying that $\chi_I \in \mathrm{IND}(H)$ and finishing the proof.

As $\chi_I \in \mathrm{TH}(H)$, there is some matrix A which witnesses this fact, i.e. satisfies the conditions in the definition (6.3) of $\mathrm{TH}(H)$ with $a = \chi_I$. Since $\chi_I = \chi_I^2$, it follows that the diagonal of $A - \chi_I \otimes \chi_I$ is zero; together with the constraint $A - \chi_I \otimes \chi_I \succeq 0$, this implies that the whole matrix is zero. Thus $A(x,y) = \chi_I(x)\chi_I(y)$ for all $x, y \in V$.

The constraint $A_x[V_x] \in A(x,x)\,\mathrm{TH}(H_x)$ for $x \in V$ then translates to $\chi_I[V_x] \in \mathrm{TH}(H_x)$ for all $x \in I$. By the induction hypothesis, it follows that $I \cap V_x$ is independent in H_x for all $x \in I$; by the definition of the link hypergraphs H_x, this implies that I is independent in H, as wished. □

The next results show that our generalizations satisfy several of the same properties of the original notions of theta body and weighted theta number of graphs.

Lemma 6.3. *Let H be a uniform hypergraph. For all $u, w \in \mathbb{R}_+^V$ and all $\lambda > 0$, we have that $\vartheta_{\lambda w}(H) = \lambda \vartheta_w(H)$ and*

$$\max\{\vartheta_u(H),\, \vartheta_w(H)\} \leq \vartheta_{u+w}(H) \leq \vartheta_u(H) + \vartheta_w(H).$$

Proof. This lemma immediately follows from the definition

$$\vartheta_w(H) = \max\{w^T f : f \in \mathrm{TH}(H)\}$$

and the fact that all vectors in $\mathrm{TH}(H)$ are nonnegative. □

Lemma 6.4. *Let H be a uniform hypergraph and let $f \in \mathbb{R}_+^V$ be a nonnegative vector. If $f \in \mathrm{TH}(H)$, then also $g \in \mathrm{TH}(H)$ for all $0 \leq g \leq f$.*

Proof. We again proceed by induction on the uniformity of the hypergraph H, with the base case of 1-uniform hypergraphs being trivial.

Now suppose we have proven the result for some $k \geq 2$ and all $(k-1)$-uniform hypergraphs. Let H be a k-uniform hypergraph on vertex set V, and fix any vertex $z \in V$. Given a vector $f \in \mathbb{R}_+^V$, we denote by $\tilde{f}$ the vector satisfying $\tilde{f}(z) = 0$ and $\tilde{f}(x) = f(x)$ for all $x \in V \setminus \{z\}$; we will first show that $\tilde{f} \in \mathrm{TH}(H)$ whenever $f \in \mathrm{TH}(H)$.

If $f \in \mathrm{TH}(H)$, then there is some matrix $F \in \mathbb{R}^{V \times V}$ such that $\operatorname{diag} F = f$, $F - f \otimes f \succeq 0$ and $F_x[V_x] \in F(x,x)\,\mathrm{TH}(H_x)$ for all $x \in V$. Denote by $\tilde{F} \in \mathbb{R}^{V \times V}$ the matrix with entries

$$\tilde{F}(x,y) = \begin{cases} F(x,y) & \text{if } x \neq z \text{ and } y \neq z, \\ 0 & \text{otherwise.} \end{cases}$$

Note that $\operatorname{diag} \tilde{F} = \tilde{f}$, and that $\tilde{F} - \tilde{f} \otimes \tilde{f}$ is exactly the principal submatrix of $F - f \otimes f$ indexed by $V \setminus \{z\}$, followed by a row and column of zeros. As $F - f \otimes f \succeq 0$, it then follows that $\tilde{F} - \tilde{f} \otimes \tilde{f} \succeq 0$ too.

We also note that $\tilde{F}_z \equiv 0$, while $0 \leq \tilde{F}_x[V_x] \leq F_x[V_x]$ and $\tilde{F}(x,x) = F(x,x)$ for all $x \in V \setminus \{z\}$. By the induction hypothesis applied to each link hypergraph H_x, this implies that $\tilde{F}_x[V_x] \in \tilde{F}(x,x)\,\mathrm{TH}(H_x)$ for all $x \in V$; we conclude that $\tilde{f} \in \mathrm{TH}(H)$, as wished.

The lemma now follows by a simple recursive procedure. Label the vertices of H by $x_1, x_2, \ldots, x_{|V|}$, and fix $f \in \mathrm{TH}(H)$ and some vector $0 \leq g \leq f$. For each $0 \leq i \leq |V|$, denote by $g_i \in \mathbb{R}_+^V$ the vector given by

$$g_i(x_j) = \begin{cases} g(x_j) & \text{if } j \leq i, \\ f(x_j) & \text{if } j > i; \end{cases}$$

denote also by $\tilde{g}_i$ the vector which agrees with g_i in $V \setminus \{x_{i+1}\}$, but satisfies $\tilde{g}_i(x_{i+1}) = 0$.

We have already shown that $\tilde{g}_i \in \mathrm{TH}(H)$ whenever $g_i \in \mathrm{TH}(H)$. Since g_i, $\tilde{g}_i$ and g_{i+1} all agree on $V \setminus \{x_{i+1}\}$ while

$$g_{i+1}(x_{i+1}) = g(x_{i+1}) \in [0,\ f(x_{i+1})] = [\tilde{g}_i(x_{i+1}),\ g_i(x_{i+1})],$$

it follows that g_{i+1} is a convex combination of g_i and $\tilde{g}_i$; by convexity of $\mathrm{TH}(H)$, we conclude that $g_{i+1} \in \mathrm{TH}(H)$ whenever $g_i \in \mathrm{TH}(H)$. Starting from $g_0 = f \in \mathrm{TH}(H)$, it then follows by recursion that $g_{|V|} = g \in \mathrm{TH}(H)$ too, finishing the inductive step and the proof. □

One can also give a different formulation for the theta body, using only the weighted theta number as a black box. For some applications this formulation is more suitable than our more explicit definition.

Lemma 6.5. *The theta body of a uniform hypergraph H is equivalently given by*

$$\mathrm{TH}(H) = \{f \in \mathbb{R}_+^V :\ w^T f \leq \vartheta_w(H) \text{ for all } w \in \mathbb{R}_+^V\}.$$

Proof. Let us denote the set defined in the statement above by $B(H)$; we wish to show that $\mathrm{TH}(H) \subseteq B(H)$ and $B(H) \subseteq \mathrm{TH}(H)$.

The first inclusion is easy: since $\vartheta_w(H) = \max\{w^T f : f \in \mathrm{TH}(H)\}$, it follows that for each $f \in \mathrm{TH}(H)$ and all $w \in \mathbb{R}_+^V$ we have $w^T f \leq \vartheta_w(H)$. This means that $\mathrm{TH}(H) \subseteq B(H)$.

Let us now assume for contradiction that the second inclusion is false, so there is some vector $f \in B(H) \setminus \mathrm{TH}(H)$. Since $\mathrm{TH}(H)$ is bounded, closed and convex (by Theorem 6.1), it follows that f can be strictly separated from $\mathrm{TH}(H)$ by a hyperplane: there is a vector $u \in \mathbb{R}^V$ for which

$$u^T f > \max\{u^T g :\ g \in \mathrm{TH}(H)\}.$$

Let us denote by w the nonnegative part of u, that is

$$w(x) = \max\{0,\ u(x)\} \quad \text{for } x \in V.$$

Given a vector $g \in \mathbb{R}^V$, denote by $\overline{g}$ the vector given by

$$\overline{g}(x) = \begin{cases} g(x) & \text{if } w(x) \neq 0, \\ 0 & \text{if } w(x) = 0 \end{cases}$$

for all $x \in V$. Note that

$$w^T g = w^T \overline{g} = u^T \overline{g} \quad \text{for all } g \in \mathbb{R}^V.$$

Since $\overline{g} \in \mathrm{TH}(H)$ whenever $g \in \mathrm{TH}(H)$ (by Lemma 6.4), by maximizing over $g \in \mathrm{TH}(H)$ it follows that

$$\max\{w^T g : g \in \mathrm{TH}(H)\} \le \max\{u^T g : g \in \mathrm{TH}(H)\}.$$

Finally, since f is a nonnegative vector (by hypothesis), we have that $u^T f \le w^T f$. Putting everything together, we conclude that

$$w^T f \ge u^T f > \max\{u^T g : g \in \mathrm{TH}(H)\} \ge \max\{w^T g : g \in \mathrm{TH}(H)\} = \vartheta_w(H),$$

which contradicts our assumption that $f \in B(H)$ and finishes the proof. □

Other than the theoretical properties shown in the previous lemmas, another advantage of our notion of theta number for hypergraphs is that it lends itself well to methods of *symmetry reduction* for semidefinite programs. This allows us to efficiently compute bounds on the independence number even for some very large hypergraphs, as we next show.

6.3 Exploiting symmetry

We shall now focus on the 'unweighted' version of the theta number, where the weight function w is identically one, as this is the simplest and most common case in applications. Recall that $\vartheta_{\mathbf{1}}(H)$ gives an upper bound on the usual independence number $\alpha(H)$ of the hypergraph H.

An advantage of dealing with a constant weight function is that the optimization variables of the corresponding theta number program (6.2) can be made to satisfy the same symmetries as the hypergraph in consideration; this can in turn be used to significantly simplify the optimization program if the hypergraph admits many symmetries. We will now give an overview of how this can be accomplished.

Let us denote by Γ the *automorphism group* of H, that is, the group of permutations of its vertex set V which maps edges to edges. The first thing to note is that, given any feasible solution $(a, A) \in \mathbb{R}^V \times \mathbb{R}^{V\times V}$ of the program for $\vartheta_{\mathbf{1}}(H)$, its Γ-*symmetrization* $(\tilde{a}, \tilde{A})$ where

$$\tilde{a}(x) = \frac{1}{|\Gamma|}\sum_{\gamma\in\Gamma} a(\gamma x), \quad \tilde{A}(x,y) = \frac{1}{|\Gamma|}\sum_{\gamma\in\Gamma} A(\gamma x, \gamma y), \quad x, y \in V$$

will still be a feasible solution of this same program (by convexity, see Theorem 6.1 and its proof), and has the same objective value. This observation already allows us to reduce the search space of our optimization program to Γ-invariant functions, but with some more information on how Γ acts on the vertex set we can also simplify the formulation of this program.

For instance, let us consider the case where H is *vertex transitive*, meaning that its vertex set decomposes into a single orbit under the action of Γ. Then $\tilde{a} = \operatorname{diag}\tilde{A}$ will be a constant function, say $\tilde{a} = \lambda\mathbf{1}$; moreover, it suffices to check the theta body constraint from program (6.2) for $\tilde{A}$ for a single vertex, as the same constraint for other vertices will be equivalent up to the action of an appropriate group element. Fixing an arbitrary vertex

$z \in V$, the program for $\vartheta_{\mathbf{1}}(H)$ then becomes:

$$\begin{aligned} \max\ & \lambda|V| \\ & A(z,z) = \lambda, \\ & A_z[V_z] \in \lambda\,\mathrm{TH}(H_z), \\ & A \in \mathbb{R}^{V\times V} \text{ is } \Gamma\text{-invariant},\ A - \lambda^2 \mathbf{1}\otimes\mathbf{1} \succeq 0. \end{aligned}$$

Next we simplify the third constraint in the program above. Note that every Γ-invariant matrix $A \in \mathbb{R}^{V\times V}$ admits $\mathbf{1}$ as an eigenvector, with associated eigenvalue $|V|^{-1}\sum_{x,y\in V} A(x,y)$. Indeed, given any two vertices $x, x' \in V$, by assumption there is an element $\gamma \in \Gamma$ for which $\gamma x = x'$; by Γ-invariance of A we have

$$\sum_{y\in V} A(x,y) = \sum_{y\in V} A(\gamma x, \gamma y) = \sum_{y\in V} A(x', y),$$

and so

$$A\mathbf{1} = \Big(\sum_{y\in V} A(x',y)\Big)\mathbf{1} = \Big(\frac{1}{|V|}\sum_{x,y\in V} A(x,y)\Big)\mathbf{1}.$$

The eigenvalue of $A - \lambda^2\mathbf{1}\otimes\mathbf{1}$ associated to the eigenvector $\mathbf{1}$ is then $|V|^{-1}\sum_{x,y\in V} A(x,y) - \lambda^2|V|$, while all other eigenvalues of $A - \lambda\mathbf{1}\otimes\mathbf{1}$ are the same as those for A. It follows that a Γ-invariant matrix A satisfies $A - \lambda^2\mathbf{1}\otimes\mathbf{1} \succeq 0$ if and only if A is positive semidefinite and satisfies

$$\sum_{x,y\in V} A(x,y) \geq \lambda^2|V|^2.$$

Now we substitute A by λA in order to take out the quadratic factor λ^2, while not significantly affecting the rest of the program (note that $\lambda > 0$, as $\vartheta_{\mathbf{1}}(H) \geq \alpha(H) \geq 1$). Optimizing the resulting program over $\lambda > 0$, we immediately obtain the following result:

Lemma 6.6. *Suppose H is a vertex transitive uniform hypergraph with automorphism group Γ, and fix an arbitrary vertex $z \in V$. Then $\vartheta_{\mathbf{1}}(H)$ is equal to*

$$\begin{aligned} \max\ & |V|^{-1}\textstyle\sum_{x,y\in V} A(x,y) \\ & A(z,z) = 1, \\ & A_z[V_z] \in \mathrm{TH}(H_z), \\ & A \in \mathbb{R}^{V\times V} \textit{ is } \Gamma\textit{-invariant and positive semidefinite.} \end{aligned} \tag{6.4}$$

Next we note that both the link hypergraph H_z and the restricted function $f := A_z[V_z]$ in program (6.4) above are invariant under the induced action of the stabilizer group $\mathrm{Stab}^\Gamma(z)$. Using the characterization of theta bodies given in Lemma 6.5, to check the second constraint $f \in \mathrm{TH}(H_z)$ it then suffices to check the inequalities $w^T f \leq \vartheta_w(H_z)$ for those vectors $w \in \mathbb{R}_+^{V_z}$ which are also invariant under $\mathrm{Stab}^\Gamma(z)$; this happens because neither the value of $w^T f$ nor $\vartheta_w(H_z)$ changes if one substitutes a vector w by its symmetrized version

$$\tilde{w}(x) = \frac{1}{|\,\mathrm{Stab}^\Gamma(z)|}\sum_{\tau\in\mathrm{Stab}^\Gamma(z)} w(\tau x), \qquad x \in V_z.$$

In the simplest case where the action of $\mathrm{Stab}^\Gamma(z)$ on V_z is transitive, the space of $\mathrm{Stab}^\Gamma(z)$-invariant vectors on $\mathbb{R}^{V_z}$ comprises only the constant vectors; the theta body constraint then reduces to the simple linear constraint

$$\sum_{x \in V_z} A(z, x) \le \vartheta_{\mathbf{1}}(H_z).$$

Recall that this was the case for the geometric hypergraphs encoding regular simplices considered in the last chapter; since all link hypergraphs of each order were vertex transitive (under the action of the corresponding stabilizer subgroup), this simplification worked through the whole recursive definition of the theta numbers and we obtained very simple programs at the end.

Finally, tools from representation theory allow us to obtain a more efficient characterization of Γ-invariant positive semidefinite matrices, and can drastically reduce the dimension of the resulting optimization program. We will now quickly review how this can be accomplished; for more detailed expositions, we refer the reader to Vallentin [66] and to Bachoc, Gijswijt, Schrijver and Vallentin [3].

While the program $\vartheta_{\mathbf{1}}(H)$ we wish to analyze is real, for the purpose of carrying out this analysis via representation theory it is more convenient to consider *complex* vector spaces. We recall that a complex matrix $X \in \mathbb{C}^{n\times n}$ is *positive semidefinite* if for all $(z_1, \dots, z_n) \in \mathbb{C}^n$ we have

$$\sum_{i,j=1}^{n} z_i X_{ij} \overline{z_j} \ge 0.$$

The space of complex matrices $\mathbb{C}^{n\times n}$ is equipped with a complex inner product, defined by $\langle X, Y\rangle = \sum_{i,j=1}^{n} X_{ij}\overline{Y_{ij}}$.

Fix a finite set V, and a finite group Γ acting on V. We will work on the vector space $\mathbb{C}^V$, which is equipped with inner product

$$(f, g) := \frac{1}{|V|} \sum_{x \in V} f(x)\overline{g(x)}.$$

The group Γ also naturally acts on the vector space $\mathbb{C}^V$ by $(\gamma f)(x) := f(\gamma^{-1}x)$; note that the inner product on $\mathbb{C}^V$ is invariant under this group action.

We can decompose $\mathbb{C}^V$ under the action of Γ into a direct sum of irreducible Γ-invariant subspaces. Write $\{R_k : 1 \le k \le d\}$ for a complete set of representatives of the isomorphism classes of irreducible Γ-invariant subspaces which are direct summands of $\mathbb{C}^V$. Then

$$\mathbb{C}^V = \bigoplus_{k=1}^{d} \bigoplus_{i=1}^{m_k} H_{k,i},$$

where m_k denotes the multiplicity of the irreducible representation R_k in $\mathbb{C}^V$ and, for $1 \le i \le m_k$, the subspaces $H_{k,i}$ are pairwise orthogonal and Γ-isomorphic to R_k.

Note that the decomposition $\mathbb{C}^V = \bigoplus_{k=1}^{d} \mathcal{I}_k$ into isotypic components is unique, but the decomposition of each isotypic component $\mathcal{I}_k = \bigoplus_{i=1}^{m_k} H_{k,i}$ is not unique unless $m_k = 1$. To this specific decomposition of $\mathcal{I}_k$ we associate a *zonal matrix* $E_k : V \times V \to \mathbb{C}^{m_k \times m_k}$ as follows.

For each $1 \le k \le d$ we choose an orthonormal basis $(e_{k,0,1}, \ldots, e_{k,0,h_k})$ of R_k, where $h_k = \dim(R_k)$, and for each $1 \le i \le m_k$ let $\phi_{k,i} : R_k \to H_{k,i}$ be a Γ-isomorphism which preserves the inner product. For $1 \le s \le h_k$ denote $e_{k,i,s} := \phi_{k,i}(e_{k,0,s})$, so that $(e_{k,i,1}, \ldots, e_{k,i,h_k})$ is an orthonormal basis of $H_{k,i}$. For each $x, y \in V$ we denote by $E_k(x, y)$ the $m_k \times m_k$ matrix given by

$$E_{k,ij}(x,y) = \frac{1}{h_k} \sum_{s=1}^{h_k} e_{k,i,s}(x)\overline{e_{k,j,s}(y)}. \tag{6.5}$$

These matrices E_k are invariant under the action of Γ, and one can show that they are in fact independent of our choice of basis for the spaces R_k; see [66, Proposition 3.1].

The zonal matrices just defined allow us to efficiently characterize Γ-invariant positive semidefinite matrices $A \in \mathbb{C}^{V\times V}$. This is the content of the next result, which is (a special case of) Theorem 3.3 in [3]:

Theorem 6.7 (Bachoc, Gijswijt, Schrijver and Vallentin [3])**.** *Let $A \in \mathbb{C}^{V\times V}$ be a complex matrix and let Γ be a finite group acting on V. Then A is Γ-invariant and positive semidefinite if and only if there exist positive semidefinite matrices $F_k \in \mathbb{C}^{m_k\times m_k}$, $1 \le k \le d$, such that*

$$A(x,y) = \sum_{k=1}^{d} \langle F_k, \overline{E_k(x,y)} \rangle \quad \textit{for all } x, y \in V.$$

The main advantage of using this characterization is that one exchanges a 'large' matrix $A \in \mathbb{C}^{V\times V}$ for several 'small' matrices $F_k \in \mathbb{C}^{m_k\times m_k}$, $1 \le k \le d$. Depending on how Γ acts on V, this can constitute a drastic reduction in the dimension of the considered problem.

6.3.1 Example: triangle-encoding hypergraphs on the hypercube

In order to illustrate the methods discussed in this section, let us take a look at hypergraphs encoding equilateral triangles on the hypercube $\{0,1\}^n$. More explicitly, take integers $n \ge 2$, $1 \le m \le n/2$ and consider the 3-uniform hypergraph H on vertex set $V = \{0,1\}^n$, whose edges are all collections of 3 vertices having pairwise Hamming distance m:

$$E(H) = \{\, \{x, y, z\} \subset \{0,1\}^n : \ d(x,y) = d(x,z) = d(y,z) = m \,\},$$

where $d(x, y)$ denotes the number of coordinates where x and y differ.

Note that this hypergraph H contains no edges if m is odd. Indeed, denoting by $|x|$ the Hamming weight of $x \in \{0,1\}^n$, it is easy to see that $|x| + |y| - d(x, y)$ is equal to twice the number of indices i where $x_i = y_i = 1$; thus $d(x, y) \equiv |x| + |y| \mod 2$. This implies that

$$d(x,y) + d(x,z) + d(y,z) \equiv 0 \mod 2 \quad \text{for all } x, y, z \in \{0,1\}^n,$$

and so not all of the distances $d(x, y)$, $d(x, z)$, $d(y, z)$ can be odd. We then assume from now on that m is an even integer.

We wish to compute the theta number $\vartheta_{\mathbf{1}}(H)$ of this hypergraph H, and we will do so by exploiting its rich symmetry as was previously discussed. Note that the automorphism group of H is the group Γ generated by permutations of the n indices in $\{0,1\}^n$ followed by independent switches $0 \leftrightarrow 1$ at each of the n positions; it has order $2^n n!$.

Denote by W_m the set of all points on the hypercube $\{0,1\}^n$ which have Hamming weight m, and let G_m denote the graph on vertex set W_m whose edges are all pairs $\{x,y\} \subseteq W_m$ having Hamming distance m. In our earlier notation, W_m is exactly the vertex set V_0 of the link hypergraph of H at $0 = (0,\dots,0) \in \{0,1\}^n$, and G_m is exactly this same link hypergraph H_0. By Lemma 6.6, $\vartheta_{\mathbf{1}}(H)$ is the optimal value of program

$$\begin{aligned} \max\ & 2^{-n} \textstyle\sum_{x,y\in\{0,1\}^n} A(x,y) \\ & A(0,0) = 1, \\ & A_0[W_m] \in \mathrm{TH}(G_m), \\ & A \in \mathbb{R}^{\{0,1\}^n\times\{0,1\}^n} \text{ is } \Gamma\text{-invariant and positive semidefinite.} \end{aligned} \tag{6.6}$$

Note that the stabilizer subgroup $\mathrm{Stab}^{\Gamma}(0)$ is the group $\mathfrak{S}_n$ of permutation of the n indices. Both $A_0[W_m]$ in the program above and the graph G_m are invariant under the action of $\mathfrak{S}_n$; moreover, $\mathfrak{S}_n$ acts transitively on W_m. In order to check whether $A_0[W_m] \in \mathrm{TH}(G_m)$, it then suffices (by Lemma 6.5 and symmetrization, as previously discussed) to check whether $w^T A_0[W_m] \le \vartheta_w(G_m)$ for the constant functions $w \in \mathbb{R}_+^{W_m}$. More explicitly, the constraint $A_0[W_m] \in \mathrm{TH}(G_m)$ in the last program is equivalent to the simple inequality

$$\sum_{x\in W_m} A(0,x) \le \vartheta_{\mathbf{1}}(G_m).$$

Now we must find a more efficient characterization of Γ-invariant positive semidefinite matrices on $\{0,1\}^n$, as the one given in Theorem 6.7. It is a classical result (see for instance Dunkl [27]) that $\mathbb{C}^{\{0,1\}^n}$ decomposes under the action of Γ as

$$\mathbb{C}^{\{0,1\}^n} = \bigoplus_{k=0}^{n} H_k,$$

where each H_k denotes the subspace spanned by the *characters* χ_y with weight $|y| = k$, where $\chi_y(x) := (-1)^{x\cdot y}$; these spaces H_k are all Γ-irreducible and pairwise non-equivalent.

It is a simple matter to check that the characters are all unit vectors and pairwise orthogonal under the inner product $(f,g) = 2^{-n}\sum_{x\in\{0,1\}^n} f(x)\overline{g(x)}$. It follows that, for each $0 \le k \le n$, the vectors $\{\chi_y : |y| = k\}$ form an orthonormal basis of the space H_k, which has dimension $\binom{n}{k}$.

Since the multiplicity of each irreducible space H_k in $\mathbb{C}^{\{0,1\}^n}$ is 1, the zonal matrix defined in equation (6.5) becomes the *zonal function*

$$E_k(x,y) = \frac{1}{\binom{n}{k}} \sum_{\substack{z\in\{0,1\}^n \\ |z|=k}} \chi_z(x)\chi_z(y).$$

These zonal functions are Γ-invariant, so the value of $E_k(x,y)$ depends only on the Hamming distance $d(x,y)$ (which characterizes the orbit of the pair (x,y) under Γ); let us write $K_k(t)$ for the value of $E_k(x,y)$ when $d(x,y) = t$.

We can easily obtain an explicit formula for these functions $K_k(t)$, $0 \le k \le n$. Let 1^t0^{n-t} denote the vector whose first t indices are 1 and the rest are 0, so $d(0, 1^t0^{n-t}) = t$. Then

$$K_k(t) = E_k(0, 1^t0^{n-t}) = \frac{1}{\binom{n}{k}} \sum_{\substack{z\in\{0,1\}^n \\ |z|=k}} \chi_z(0)\chi_z(1^t0^{n-t}) = \frac{1}{\binom{n}{k}} \sum_{\substack{z\in\{0,1\}^n \\ |z|=k}} (-1)^{z\,\cdot\, 1^t0^{n-t}}.$$

Fix some $0 \le i \le t$. The vectors $z \in \{0,1\}^n$ with $|z| = k$ and satisfying $z \cdot 1^t 0^{n-t} = i$ are those where exactly i of the first t coordinates are 1, and exactly $k-i$ of the last $n-t$ coordinates are 1. There are $\binom{t}{i}\binom{n-t}{k-i}$ of those vectors; it then follows from the last expression that

$$K_k(t) = \frac{1}{\binom{n}{k}} \sum_{i=0}^{k} (-1)^i \binom{t}{i}\binom{n-t}{k-i}.$$

These functions K_k just defined are exactly the *Krawtchouk polynomials*, normalized so that $K_k(0) = 1$ for all $1 \le k \le n$. These polynomials are pairwise orthogonal with respect to the weight function $w(t) = \binom{n}{t}$, meaning that

$$\sum_{t=0}^{n} \binom{n}{t} K_k(t) K_\ell(t) = 0 \quad \text{if } k \ne \ell. \tag{6.7}$$

For more information on Krawtchouk polynomials and their role in representation theory, see Dunkl [27].

By Theorem 6.7, we conclude that $A \in \mathbb{C}^{\{0,1\}^n \times \{0,1\}^n}$ is Γ-invariant and positive semidefinite if and only if there are nonnegative quantities $f_0, f_1, \dots, f_n$ for which

$$A(x,y) = \sum_{k=0}^{n} f_k K_k(d(x,y)), \qquad x, y \in \{0,1\}^n.$$

Note that all such matrices are real-valued, so we can immediately substitute this last expression into our program (6.6). The objective function $2^{-n} \sum_{x,y \in \{0,1\}^n} A(x,y)$ can then be expressed as

$$\frac{1}{2^n} \sum_{k=0}^{n} f_k \sum_{t=0}^{n} \sum_{\substack{x,y \in \{0,1\}^n \\ d(x,y)=t}} K_k(t) = \sum_{k=0}^{n} f_k \sum_{t=0}^{n} \binom{n}{t} K_k(t).$$

By the orthogonality relations (6.7), this is equal to

$$f_0 \sum_{t=0}^{n} \binom{n}{t} = 2^n f_0.$$

Moreover, since $K_k(0) = 1$ for $1 \le k \le n$, the first constraint of program (6.6) translates to

$$A(0,0) = \sum_{k=0}^{n} f_k = 1.$$

Finally, as $d(0,x) = m$ for all $x \in W_m$, we have

$$\sum_{x \in W_m} A(0,x) = \binom{n}{m} \sum_{k=0}^{n} f_k K_k(m).$$

We conclude that the program for $\vartheta_{\mathbf{1}}(H)$ can be simplified to

$$\begin{aligned} \max\ & 2^n f_0 \\ & \textstyle\sum_{k=0}^{n} f_k = 1, \\ & \textstyle\binom{n}{m} \sum_{k=0}^{n} f_k K_k(m) \le \vartheta_{\mathbf{1}}(G_m), \\ & f_0, f_1, \dots, f_n \ge 0. \end{aligned}$$

The optimal value of this last program admits a simple analytical expression. Indeed, the sum $\sum_{k=0}^{n} f_k K_k(m)$ on the second constraint is a convex combination of the numbers $K_k(m)$, and we wish to keep this convex combination below $\binom{n}{m}^{-1} \vartheta_{\mathbf{1}}(G_m)$ while maximizing f_0. The best way to do so is to concentrate all the weight of the combination on f_0 and f_{k^*}, where k^* is such that $K_{k^*}(m)$ is the most negative number appearing in the convex combination; that is, denoting

$$M_{n,m} := \min_{1\leq k\leq n} K_k(m),$$

we choose $1 \leq k^* \leq n$ for which $K_{k^*}(t) = M_{n,m}$.[1] Solving the optimization problem using only the variables f_0 and f_{k^*}, we obtain the optimal value

$$\vartheta_{\mathbf{1}}(H) = 2^n \frac{\binom{n}{m}^{-1} \vartheta_{\mathbf{1}}(G_m) - M_{n,m}}{1 - M_{n,m}}. \tag{6.8}$$

Remark 6.8. *It is interesting to compare this last formula to those in Theorem 5.3 and Theorem 5.5 from last chapter. These formulas were all obtained as the optimal value of 'theta-like' programs for geometric hypergraphs in different spaces, which admit very different groups of symmetries; their analysis through symmetry reduction techniques, however, proceeded in a very similar manner. The main difference between all three formulas is the choice of orthogonal polynomials we are minimizing, which in turn comes from representation theory of the corresponding automorphism group; that we obtain such similar formulas at the end is ultimately due to the fact that the considered decompositions into irreducible subspaces are all multiplicity-free. We refer the reader to Stanton [59] for a discussion on the similarities between multiplicity-free actions of finite and infinite groups, and in particular the analogy between the action of* Γ *on the hypercube* $\{0,1\}^n$ *and the action of* $\mathrm{O}(n)$ *on the sphere* S^{n-1}.

Now we must compute $\vartheta_{\mathbf{1}}(G_m)$. Recall that G_m is the graph whose vertex set W_m comprises all binary sequences $x \in \{0,1\}^n$ with Hamming weight m, and two vertices are adjacent if their Hamming distance is m. This graph is invariant under the action of the symmetric group $\mathfrak{S}_n$ permuting the coordinates of each vertex, and under this action it is vertex transitive. By Lemma 6.6, $\vartheta_{\mathbf{1}}(G_m)$ is then equal to

$$\begin{aligned} \max\ & \tbinom{n}{m}^{-1} \textstyle\sum_{x,y\in W_m} A(x,y) \\ & \textstyle\sum_{x\in W_m} A(x,x) = \tbinom{n}{m}, \\ & A(x,y) = 0 \quad \text{if } d(x,y) = m, \\ & A \in \mathbb{R}^{W_m\times W_m} \text{ is } \mathfrak{S}_n\text{-invariant and positive semidefinite.} \end{aligned}$$

Under the action of $\mathfrak{S}_n$ the vector space $\mathbb{C}^{W_m}$ decomposes into $\bigoplus_{k=0}^{m} U_k$, where each subspace U_k has dimension $h_k = \binom{n}{k} - \binom{n}{k-1}$ and corresponds to the irreducible representation of $\mathfrak{S}_n$ given by the partition $(n-k,k)$. These vector spaces were described explicitly by

[1] One can show that these values $M_{n,m}$ are indeed negative, for all integers $n \geq m \geq 1$. For instance, it is known that $\left\{\binom{n}{m} K_k(m) : 0 \leq k \leq n\right\}$ is exactly the set of eigenvalues of the adjacency matrix of the graph on $\{0,1\}^n$ where two vertices form an edge if their distance is m; this is shown e.g. in Chapter 10 of Terras' book [65]. Since the trace of the adjacency matrix is zero and $K_0(m) = 1$, its minimum eigenvalue $\binom{n}{m} M_{n,m}$ must be negative.

Delsarte [24] (where they are denoted Harm(k)), who also computed the corresponding zonal functions

$$E_k(x,y) := \frac{1}{h_k}\sum_{s=1}^{h_k} e_{k,s}(x)\overline{e_{k,s}(y)}, \quad x,y \in W_m,$$

where $\{e_{k,1},\ldots,e_{k,h_k}\}$ denotes an arbitrary orthonomal basis of U_k.

Define the *Hahn polynomials* Q_k, $0 \le k \le m$, by

$$Q_k(t) = \sum_{i=0}^{k}(-1)^i \frac{\binom{k}{i}\binom{n+1-k}{i}\binom{t}{i}}{\binom{m}{i}\binom{n-m}{i}}.$$

These polynomials Q_k are normalized so that $Q_k(0) = 1$, and they are orthogonal with respect to the weight function

$$w(t) = \binom{m}{t}\binom{n-m}{t}, \quad 0 \le t \le m,$$

meaning that $\sum_{t=0}^m w(t)Q_k(t)Q_\ell(t) = 0$ whenever $k \ne \ell$. Delsarte [24, Theorem 5] showed that the zonal functions E_k can be expressed in terms of the Hahn polynomials by

$$E_k(x,y) = Q_k(d(x,y)/2), \quad 0 \le k \le m;$$

see also Dunkl [28]. Using this result we then obtain from Theorem 6.7 the following characterization: $A \in \mathbb{R}^{W_m\times W_m}$ is $\mathfrak{S}_n$-invariant and positive semidefinite if and only if

$$A(x,y) = \sum_{k=0}^{m} f_k Q_k(d(x,y)/2), \quad x,y \in W_m$$

holds for some nonnegative quantities $f_0, f_1, \ldots, f_m \ge 0$.

The program defining $\vartheta_{\mathbf{1}}(G_m)$ can now be simplified. First we note that

$$\binom{n}{m}^{-1}\sum_{x,y\in W_m} A(x,y) = \sum_{k=0}^{m} f_k \sum_{t=0}^{m}\binom{n}{m}^{-1}\sum_{\substack{x,y\in W_m\\ d(x,y)=2t}} Q_k(t) = \sum_{k=0}^{m} f_k \sum_{t=0}^{m}\binom{m}{t}\binom{n-m}{t}Q_k(t);$$

by the orthogonality relations for Hahn polynomials, this is equal to

$$f_0\sum_{t=0}^{m}\binom{m}{t}\binom{n-m}{t} = \binom{n}{m}f_0.$$

Moreover, we have

$$\sum_{x\in W_m} A(x,x) = \binom{n}{m}\sum_{k=0}^{m} f_k,$$

and $A(x,y) = \sum_{k=0}^m f_k Q_k(m/2)$ whenever $d(x,y) = m$. We conclude that $\vartheta_{\mathbf{1}}(G_m)$ is equal to

$$\begin{aligned}
\max\ & \tbinom{n}{m} f_0\\
& \textstyle\sum_{k=0}^m f_k = 1,\\
& \textstyle\sum_{k=0}^m f_k Q_k(m/2) = 0,\\
& f_0, f_1, \ldots, f_m \ge 0.
\end{aligned}$$

As before, this last program admits a simple analytical solution. Denote

$$N_{n,m} := \min_{0 \le k \le m} Q_k(m/2);$$

proceeding as we did when computing $\vartheta_{\mathbf{1}}(H)$, one can easily show that

$$\vartheta_{\mathbf{1}}(G_m) = \binom{n}{m} \frac{-N_{n,m}}{1 - N_{n,m}}.$$

(It is interesting to note that this is precisely Hoffman's spectral bound for the independence number of G_m.) Combining this expression with equation (6.8), we finally conclude that

$$\vartheta_{\mathbf{1}}(H) = 2^n \Big(1 - \frac{1}{(1 - N_{n,m})(1 - M_{n,m})} \Big).$$

It is now a simple matter to numerically compute the value of $\vartheta_{\mathbf{1}}(H)$, which then provides an upper bound for the independence number of this hypergraph.

We note that these same methods continue to work in the more general setting of *q-ary cubes* $\{1, \ldots, q\}^n$, for any integer $q \ge 2$. In this case we must make use of Krawtchouk polynomials with weight $(q-1)/q$ (see Dunkl [27]), and *q-ary Hahn polynomials* (see Delsarte [24]). One can also obtain similar programs when forbidding *non-equilateral* triangles, though in such cases the resulting optimization program will contain as variables semidefinite matrices of size 2 (for isosceles triangles) or 3 (for scalene triangles), due to the fact that the link graphs are no longer vertex transitive.

6.4 Remarks and open problems

The study of our notions of theta number and theta body for hypergraphs has only just begun, and there are several natural questions still to be answered. For instance:

Question 6.9. *For which hypergraphs H is the theta body* $\mathrm{TH}(H)$ *a polytope? Also, for which hypergraphs is the theta body* $\mathrm{TH}(H)$ *equal to the independent set polytope* $\mathrm{IND}(H)$*?*

It turns out that both of these questions have the same answer in the case of graphs, namely the class of *perfect graphs*; this was proven by Grötschel, Lovász and Schrijver [42], who determined all facet-defining inequalities for the theta body of graphs. Perhaps the answer to our questions give rise to interesting notions of perfect hypergraphs.

Another important property of the (unweighted) theta number for graphs is that it is *multiplicative* when using the strong product of graphs: it satisfies $\vartheta(F \bullet G) = \vartheta(F)\vartheta(G)$ for all finite graphs F and G, where $F \bullet G$ denotes their strong product.[2] This property was already proven by Lovász in his seminal paper [46] where he first defined the theta number. Our last question is whether such a property extends to hypergraphs:

Question 6.10. *Is there a natural notion of hypergraph product for which the hypergraph theta number $\vartheta_{\mathbf{1}}$ is multiplicative?*

[2]The strong product of F and G is the graph whose vertex set is $V(F) \times V(G)$, and where $(x, y), (x', y') \in V(F) \times V(G)$ are joined by an edge if: either $x = x'$ and $\{y, y'\} \in E(G)$, or $\{x, x'\} \in E(F)$ and $y = y'$, or $\{x, x'\} \in E(F)$ and $\{y, y'\} \in E(G)$.

Bibliography

[1] G. Ambrus and M. Matolcsi. Density estimates of 1-avoiding sets via higher order correlations. *Discrete & Computational Geometry*, pages 1–12, 2020.

[2] G. E. Andrews, R. Askey, and R. Roy. *Special functions*, volume 71 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1999.

[3] C. Bachoc, D. C. Gijswijt, A. Schrijver, and F. Vallentin. Invariant semidefinite programs. In *Handbook on semidefinite, conic and polynomial optimization*, volume 166 of *Internat. Ser. Oper. Res. Management Sci.*, pages 219–269. Springer, New York, 2012.

[4] C. Bachoc, G. Nebe, F. M. de Oliveira Filho, and F. Vallentin. Lower bounds for measurable chromatic numbers. *Geom. Funct. Anal.*, 19(3):645–661, 2009.

[5] C. Bachoc, A. Passuello, and A. Thiery. The density of sets avoiding distance 1 in Euclidean space. *Discrete Comput. Geom.*, 53(4):783–808, 2015.

[6] I. M. Bomze, M. Dür, E. de Klerk, C. Roos, A. J. Quist, and T. Terlaky. On copositive programming and standard quadratic optimization problems. volume 18, pages 301–320. 2000. GO'99 Firenze.

[7] J. Bourgain. A Szemerédi type theorem for sets of positive density in $\mathbf{R}^k$. *Israel J. Math.*, 54(3):307–316, 1986.

[8] B. Bukh. Measurable sets with excluded distances. *Geom. Funct. Anal.*, 18(3):668–697, 2008.

[9] S. Burer. On the copositive representation of binary and continuous nonconvex quadratic programs. *Math. Program.*, 120(2, Ser. A):479–495, 2009.

[10] D. Castro-Silva. Geometrical sets with forbidden configurations. *ArXiv preprint 2102.10018*, 2021.

[11] D. Castro-Silva. Quasirandomness in additive groups and hypergraphs. *ArXiv preprint 2107.01463*, 2021.

[12] D. Castro-Silva, F. M. de Oliveira Filho, L. Slot, and F. Vallentin. A recursive Lovász theta number for simplex-avoiding sets. *ArXiv preprint 2106.09360*, 2021.

[13] H. Cohn and N. Elkies. New upper bounds on sphere packings. I. *Ann. of Math. (2)*, 157(2):689–714, 2003.

[14] H. T. Croft. Incidence incidents. *Eureka*, 30:22–26, 1967.

[15] E. Croot, V. F. Lev, and P. P. Pach. Progression-free sets in $\mathbb{Z}_4^n$ are exponentially small. *Ann. of Math. (2)*, 185(1):331–337, 2017.

[16] F. Dai and Y. Xu. *Approximation theory and harmonic analysis on spheres and balls*. Springer Monographs in Mathematics. Springer, New York, 2013.

[17] N. G. de Bruijn and P. Erdős. A colour problem for infinite graphs and a problem in the theory of relations. *Nederl. Akad. Wetensch. Proc. Ser. A.* **54** = *Indagationes Math.*, 13:369–373, 1951.

[18] E. de Klerk and D. V. Pasechnik. Approximation of the stability number of a graph via copositive programming. *SIAM J. Optim.*, 12(4):875–892, 2002.

[19] D. de Laat, F. M. de Oliveira Filho, and F. Vallentin. Upper bounds for packings of spheres of several radii. *Forum Math. Sigma*, 2:Paper No. e23, 42, 2014.

[20] F. M. de Oliveira Filho and F. Vallentin. Fourier analysis, linear programming, and densities of distance avoiding sets in $\mathbb{R}^n$. *J. Eur. Math. Soc. (JEMS)*, 12(6):1417–1428, 2010.

[21] F. M. de Oliveira Filho and F. Vallentin. Computing upper bounds for the packing density of congruent copies of a convex body. In *New trends in intuitive geometry*, volume 27 of *Bolyai Soc. Math. Stud.*, pages 155–188. János Bolyai Math. Soc., Budapest, 2018.

[22] E. DeCorte, F. M. de Oliveira Filho, and F. Vallentin. Complete positivity and distance-avoiding sets. *Mathematical Programming*, 2020.

[23] E. DeCorte and O. Pikhurko. Spherical sets avoiding a prescribed set of angles. *Int. Math. Res. Not. IMRN*, (20):6095–6117, 2016.

[24] P. Delsarte. Hahn polynomials, discrete harmonics, and t-designs. *SIAM J. Appl. Math.*, 34(1):157–166, 1978.

[25] P. Delsarte, J. M. Goethals, and J. J. Seidel. Spherical codes and designs. *Geometriae Dedicata*, 6(3):363–388, 1977.

[26] C. F. Dunkl. Operators and harmonic analysis on the sphere. *Trans. Amer. Math. Soc.*, 125:250–263, 1966.

[27] C. F. Dunkl. A Krawtchouk polynomial addition theorem and wreath products of symmetric groups. *Indiana Univ. Math. J.*, 25(4):335–358, 1976.

[28] C. F. Dunkl. An addition theorem for Hahn polynomials: the spherical functions. *SIAM J. Math. Anal.*, 9(4):627–637, 1978.

[29] A. Elbert and A. Laforgia. Upper bounds for the zeros of ultraspherical polynomials. *J. Approx. Theory*, 61(1):88–97, 1990.

[30] J. S. Ellenberg and D. Gijswijt. On large subsets of $\mathbb{F}_q^n$ with no three-term arithmetic progression. *Ann. of Math. (2)*, 185(1):339–343, 2017.

[31] P. Erdős. Problems and results in combinatorial geometry. In *Discrete geometry and convexity (New York, 1982)*, volume 440 of *Ann. New York Acad. Sci.*, pages 1–11. New York Acad. Sci., New York, 1985.

[32] P. Erdős and M. Simonovits. Supersaturated graphs and hypergraphs. *Combinatorica*, 3(2):181–192, 1983.

[33] K. J. Falconer. The realization of small distances in plane sets of positive measure. *Bull. London Math. Soc.*, 18(5):475–477, 1986.

[34] P. Frankl and V. Rödl. Forbidden intersections. *Trans. Amer. Math. Soc.*, 300(1):259–286, 1987.

[35] P. Frankl and V. Rödl. A partition property of simplices in Euclidean space. *J. Amer. Math. Soc.*, 3(1):1–7, 1990.

[36] P. Frankl and R. M. Wilson. Intersection theorems with geometric consequences. *Combinatorica*, 1(4):357–368, 1981.

[37] H. Furstenberg, Y. Katznelson, and B. Weiss. Ergodic theory and configurations in sets of positive density. In *Mathematics of Ramsey theory*, volume 5 of *Algorithms Combin.*, pages 184–198. Springer, Berlin, 1990.

[38] M. X. Goemans. Semidefinite programming in combinatorial optimization. volume 79, pages 143–161. 1997. Lectures on mathematical programming (ismp97) (Lausanne, 1997).

[39] R. L. Graham. Recent trends in Euclidean Ramsey theory. *Discrete Math.*, 136(1-3):119–127, 1994.

[40] M. Grötschel, L. Lovász, and A. Schrijver. The ellipsoid method and its consequences in combinatorial optimization. *Combinatorica*, 1(2):169–197, 1981.

[41] M. Grötschel, L. Lovász, and A. Schrijver. Polynomial algorithms for perfect graphs. In *Topics on perfect graphs*, volume 88 of *North-Holland Math. Stud.*, pages 325–356. North-Holland, Amsterdam, 1984.

[42] M. Grötschel, L. Lovász, and A. Schrijver. Relaxations of vertex packing. *J. Combin. Theory Ser. B*, 40(3):330–343, 1986.

[43] M. Grötschel, L. Lovász, and A. Schrijver. *Geometric algorithms and combinatorial optimization*, volume 2 of *Algorithms and Combinatorics: Study and Research Texts*. Springer-Verlag, Berlin, 1988.

[44] G. Kalai. Some old and new problems in combinatorial geometry I: around Borsuk's problem. In *Surveys in combinatorics 2015*, volume 424 of *London Math. Soc. Lecture Note Ser.*, pages 147–174. Cambridge Univ. Press, Cambridge, 2015.

[45] D. G. Larman and C. A. Rogers. The realization of distances within sets in Euclidean space. *Mathematika*, 19:1–24, 1972.

[46] L. Lovász. On the Shannon capacity of a graph. *IEEE Trans. Inform. Theory*, 25(1):1–7, 1979.

[47] F. Mascia, E. Cilia, M. Brunato, and A. Passerini. Predicting structural and functional sites in proteins by searching for maximum-weight cliques. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 24, 2010.

[48] R. J. McEliece, E. R. Rodemich, and H. C. Rumsey, Jr. The Lovász bound and some generalizations. *J. Combin. Inform. System Sci.*, 3(3):134–152, 1978.

[49] E. Naslund. Monochromatic equilateral triangles in the unit distance graph. *Bull. Lond. Math. Soc.*, 52(4):687–692, 2020.

[50] A. M. Raĭgorodskiĭ. On a bound in the Borsuk problem. *Uspekhi Mat. Nauk*, 54(2(326)):185–186, 1999.

[51] A. M. Raĭgorodskiĭ. On the chromatic number of a space. *Uspekhi Mat. Nauk*, 55(2(332)):147–148, 2000.

[52] M. Reed and B. Simon. *Methods of modern mathematical physics. II. Fourier analysis, self-adjointness*. Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London, 1975.

[53] V. Rödl and M. Schacht. Regularity lemmas for graphs. In *Fete of combinatorics and computer science*, volume 20 of *Bolyai Soc. Math. Stud.*, pages 287–325. János Bolyai Math. Soc., Budapest, 2010.

[54] A. A. Sagdeev. An improved Frankl-Rödl theorem and some of its geometric consequences. *Problemy Peredachi Informatsii*, 54(2):45–72, 2018.

[55] A. A. Sagdeev. On a Frankl-Rödl theorem. *Izv. Ross. Akad. Nauk Ser. Mat.*, 82(6):128–157, 2018.

[56] I. J. Schoenberg. Metric spaces and completely monotone functions. *Ann. of Math. (2)*, 39(4):811–841, 1938.

[57] I. J. Schoenberg. Positive definite functions on spheres. *Duke Math. J.*, 9:96–108, 1942.

[58] A. Schrijver. A comparison of the Delsarte and Lovász bounds. *IEEE Trans. Inform. Theory*, 25(4):425–429, 1979.

[59] D. Stanton. An introduction to group representations and orthogonal polynomials. In *Orthogonal polynomials (Columbus, OH, 1989)*, volume 294 of *NATO Adv. Sci. Inst. Ser. C: Math. Phys. Sci.*, pages 419–433. Kluwer Acad. Publ., Dordrecht, 1990.

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

[61] L. A. Székely. Remarks on the chromatic number of geometric graphs. In *Graphs and other combinatorial topics (Prague, 1982)*, volume 59 of *Teubner-Texte Math.*, pages 312–315. Teubner, Leipzig, 1983.

[62] L. A. Székely. Measurable chromatic number of geometric graphs and sets without some distances in Euclidean space. *Combinatorica*, 4(2-3):213–218, 1984.

[63] L. A. Székely. Erdös on unit distances and the Szemerédi-Trotter theorems. In *Paul Erdös and his mathematics, II (Budapest, 1999)*, volume 11 of *Bolyai Soc. Math. Stud.*, pages 649–666. János Bolyai Math. Soc., Budapest, 2002.

[64] L. A. Székely and N. C. Wormald. Bounds on the measurable chromatic number of $\mathbf{R}^n$. volume 75, pages 343–372. 1989. Graph theory and combinatorics (Cambridge, 1988).

[65] A. Terras. *Fourier analysis on finite groups and applications*, volume 43 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 1999.

[66] F. Vallentin. Symmetry in semidefinite programs. *Linear Algebra Appl.*, 430(1):360–369, 2009.

[67] H. S. Witsenhausen. Spherical sets without orthogonal point pairs. *Amer. Math. Monthly*, 81:1101–1102, 1974.

[68] D. Zhang, O. Javed, and M. Shah. Video object co-segmentation by regulated maximum weight cliques. In *European Conference on Computer Vision*, pages 551–566. Springer, 2014.

[69] H. Zhian, M. Sabaei, N. T. Javan, and O. Tavallaie. Increasing coding opportunities using maximum-weight clique. In *2013 5th Computer Science and Electronic Engineering Conference (CEEC)*, pages 168–173. IEEE, 2013.

[70] T. Ziegler. An application of ergodic theory to a problem in geometric Ramsey theory. *Israel J. Math.*, 114:271–288, 1999.

[71] T. Ziegler. Nilfactors of $\mathbb{R}^m$-actions and configurations in sets of positive upper density in $\mathbb{R}^m$. *J. Anal. Math.*, 99:249–266, 2006.

Index

Summary

The central problem considered in this thesis can be phrased by the following question: given some family of forbidden geometrical configurations, how large can a set be if it does not contain any member of this family?

More concretely, given finite point configurations $P_1, \ldots, P_n \subset \mathbb{R}^d$ in Euclidean space, we denote by $\mathbf{m}_{\mathbb{R}^d}(P_1, \ldots, P_n)$ the maximum upper density a set $A \subseteq \mathbb{R}^d$ can have without containing congruent copies of any P_i. Similarly, if $P_1, \ldots, P_n$ are finite point configurations in the d-dimensional unit sphere S^d, we denote by $\mathbf{m}_{S^d}(P_1, \ldots, P_n)$ the maximum density a set in S^d can have if it avoids congruent copies of each of these configurations.

The geometrical parameters $\mathbf{m}_{\mathbb{R}^d}$ and $\mathbf{m}_{S^d}$ are called the *independence density* of the considered configurations (on the corresponding space $\mathbb{R}^d$ or S^d). They can also be seen in a more combinatorial perspective, as the independence ratio of the infinite geometric hypergraphs whose vertices are all points in the corresponding space, and whose edges are all congruent copies of each forbidden configuration.

While these parameters have been extensively studied in the case of two-point configurations, which correspond to forbidden distances in Euclidean space or forbidden angles in the sphere, the more general case of higher-order configurations remains largely unexplored. In this thesis we initiate their study, obtaining several interesting results and providing combinatorial, analytic and optimization perspectives on the general problem.

Erklärung

„Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation selbstständig und ohne die Benutzung anderer als der angegebenen Hilfsmittel und Literatur angefertigt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten und nicht veröffentlichten Werken dem Wortlaut oder dem Sinn nach entnommen wurden, sind als solche kenntlich gemacht. Ich versichere an Eides statt, dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie - abgesehen von unten angegebenen Teilpublikationen und eingebundenen Artikeln und Manuskripten - noch nicht veröffentlicht worden ist sowie, dass ich eine Veröffentlichung der Dissertation vor Abschluss der Promotion nicht ohne Genehmigung des Promotionsausschusses vornehmen werde. Die Bestimmungen dieser Ordnung sind mir bekannt. Darüber hinaus erkläre ich hiermit, dass ich die Ordnung zur Sicherung guter wissenschaftlicher Praxis und zum Umgang mit wissenschaftlichem Fehlverhalten der Universität zu Köln gelesen und sie bei der Durchführung der Dissertation zugrundeliegenden Arbeiten und der schriftlich verfassten Dissertation beachtet habe und verpflichte mich hiermit, die dort genannten Vorgaben bei allen wissenschaftlichen Tätigkeiten zu beachten und umzusetzen. Ich versichere, dass die eingereichte elektronische Fassung der eingereichten Druckfassung vollständig entspricht."

Davi de Castro Silva

Teilpublikationen:

- *Geometrical sets with forbidden configurations.* D. Castro-Silva, 41 pages, preprint available on ArXiv, arXiv:2102.10018, 2021.
- *A recursive Lovász theta number for simplex-avoiding sets.* D. Castro-Silva, F. M. de Oliveira Filho, L. Slot and F. Vallentin, 13 pages, preprint available on ArXiv, arXiv:2106.09360, 2021.