TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 379, Number 1, January 2026, Pages 33–70
https://doi.org/10.1090/tran/9578
Article electronically published on October 17, 2025

OPTIMIZATION HIERARCHIES FOR DISTANCE-AVOIDING SETS IN COMPACT SPACES

BRAM BEKKER, OLGA KURYATNIKOVA, FERNANDO MÁRIO DE OLIVEIRA FILHO, AND JUAN C. VERA

Abstract. Witsenhausen's problem asks for the maximum fraction α_n of the $(n-1)$-dimensional unit sphere that can be covered by a measurable set containing no pairs of orthogonal points. The best upper bounds for α_n are given by extensions of the Lovász theta number. In this paper, optimization hierarchies based on the Lovász theta number, like the Lasserre hierarchy, are extended to Witsenhausen's problem and similar problems. These hierarchies are shown to converge and are used to compute the best upper bounds for α_n in low dimensions.

Contents

1. Introduction

A set of points on the unit sphere $S^{n-1} = \{ x \in \mathbb{R}^n : \|x\| = 1 \}$ *avoids orthogonal pairs* if it does not contain pairs of orthogonal vectors. Witsenhausen's problem [46] asks for the maximum density that a measurable subset of S^{n-1} can have if it avoids orthogonal pairs; namely, we want to determine the parameter

$$\alpha_n = \sup\{ \omega(I)/\omega_n : I \subseteq S^{n-1} \text{ is measurable and avoids orthogonal pairs} \},$$

where ω is the surface measure on the sphere and ω_n is the total measure of the sphere.

Denote by $x \cdot y$ the Euclidean inner product between x, $y \in \mathbb{R}^n$. Fix $e \in S^{n-1}$. Witsenhausen [46] observed that the union of two open antipodal spherical caps of spherical radius $\pi/4$, namely the set

$$\{ x \in S^{n-1} : |e \cdot x| > \cos(\pi/4) \},$$

Received by the editors May 22, 2023.
2020 *Mathematics Subject Classification.* Primary 46N10, 52C10, 51K99, 90C22, 90C34.
The first author was supported by the grant OCENW.KLEIN.024 of the Dutch Research Council (NWO).

avoids orthogonal pairs, hence α_n is at least the density of this set, which is $O(n^{-1/2}2^{-n/2})$. Kalai [20, Conjecture 2.8] conjectured that this construction is optimal, that is, α_n is exactly the density of these two caps; this is known as the *double-cap conjecture*. A version of the double-cap conjecture for the complex sphere has an interpretation in quantum information theory [33].

The canonical basis vectors in $\mathbb{R}^n$ are n pairwise-orthogonal unit vectors. Any set that avoids orthogonal pairs can contain at most one of these vectors. It then follows from a simple averaging argument that $\alpha_n \leq 1/n$. This upper bound was also given by Witsenhausen [46]; it is quite far from the lower bound of the double-cap conjecture for all $n \geq 3$. For $n = 2$, the lower and upper bounds coincide. Frankl and Wilson [15] were the first to give an asymptotic upper bound for α_n that decreases exponentially fast with the dimension n.

On the sphere, distance and inner product are related: a set of points on the sphere avoids orthogonal pairs if and only if it avoids pairs of points at distance $\sqrt{2}$. More generally, let V be a metric space with metric d and let $D \subseteq (0, \infty)$ be a set of *forbidden* distances. We say that a set $I \subseteq V$ *avoids* the distances in D or that it is a *D-avoiding set* if $d(x, y) \notin D$ for all x, $y \in I$. One is usually interested in the maximum size of distance-avoiding sets, which has to be appropriately defined depending on V. Witsenhausen's problem asks for the maximum density of a $\sqrt{2}$-avoiding set on the sphere equipped with the Euclidean distance.

In Euclidean space, 1-avoiding sets are of particular interest. The maximum density of 1-avoiding sets in $\mathbb{R}^n$ has been extensively studied (see DeCorte, Oliveira, and Vallentin [10] for references), especially in its relation to the chromatic number of Euclidean space.

Distance-avoiding sets can be modeled as independent sets of graphs. Let $G = (V, E)$ be a graph. A subset of V is *independent* if it does not contain any edges. The *independence number* of G is

$$\alpha(G) = \max\{\, |I| : I \subseteq V \text{ is independent}\,\}.$$

Computing the independence number of a graph is a well-known NP-hard problem.

Given a metric space V with metric d and a set D of forbidden distances, let G be the graph with vertex set V in which x, $y \in V$ are adjacent if $d(x, y) \in D$; we call G a *distance graph*. The independent sets of G are exactly the D-avoiding sets.

Witsenhausen's problem can therefore be seen as an independent-set problem on a distance graph over the sphere. There are several optimization hierarchies based on the Lovász theta number [31] that are used to bound the independence number of finite graphs from above. In this paper, we extend some of these hierarchies from finite graphs to infinite graphs like the one related to Witsenhausen's problem; among the hierarchies we extend is the Lasserre hierarchy. We prove convergence of these hierarchies and use them to compute better upper bounds for α_n; see Table 1.

Optimization hierarchies have been used to obtain upper bounds for geometrical parameters before; the first and perhaps most famous example is the 3-point bound of Bachoc and Vallentin [3] for the kissing number problem. Hierarchies provide some of the best, and often sharp, bounds in many circumstances. This is the first use of optimization hierarchies to obtain upper bounds for Witsenhausen's problem.

1.1. **Infinite geometrical graphs.** Since we work with distance graphs on the sphere, it is more natural to talk about forbidden inner products than forbidden

TABLE 1. Lower and upper bounds for Witsenhausen's parameter α_n. The lower bound is given by the double-cap conjecture. The simple upper bound was given by Witsenhausen [46] and is just $1/n$; the best previous upper bounds are by DeCorte, Oliveira, and Vallentin [10]. The new upper bounds are derived in §6; see also Table 2.

		Previous upper bound		
n	*Lower bound*	*Simple*	*Best*	*New upper bound*
3	0.2928...	0.3333...	0.30153	0.297742
4	0.1816...	0.25	0.21676	0.194297
5	0.1161...	0.2	0.16765	0.134588
6	0.0755...	0.1666...	0.13382	0.098095
7	0.0498...	0.1428...	0.11739	0.075751
8	0.0331...	0.125	0.09981	0.061178

distances. For a set $D \subseteq [-1, 1]$ of forbidden inner products, denote by $G(S^{n-1}, D)$ the graph whose vertex set is S^{n-1} and in which two vertices x, $y \in S^{n-1}$ are adjacent if $x \cdot y \in D$.

The independent sets of $G(S^{n-1}, D)$ are precisely those sets that do not contain pairs of points with inner product in D. For instance, the independent sets of $G = G(S^{n-1}, (1/2, 1))$ are exactly the sets of points on the sphere any two of which form an angle of at least $\pi/3$. If we take an independent set of G and place on each point in the set a unit sphere touching it, we get a collection of nonoverlapping unit spheres that touch the central unit sphere. Conversely, any set of nonoverlapping unit spheres touching the central unit sphere corresponds to an independent set of G. Hence the independence number of G is precisely the *kissing number* of $\mathbb{R}^n$, which is the maximum number of nonoverlapping unit spheres that can simultaneously touch a central unit sphere.

The independent sets of the graph $G = G(S^{n-1}, \{0\})$, which we call *Witsenhausen's graph*, are precisely the subsets of S^{n-1} that avoid orthogonal pairs. These sets can be infinite, and so the independence number of G is infinite. In this case we consider the *measurable independence number*

$$\alpha_\omega(G) = \sup\{\, \omega(I) : I \subseteq S^{n-1} \text{ is measurable and independent} \,\}.$$

Witsenhausen's parameter α_n is precisely $\alpha_\omega(G)/\omega_n$.

The graphs $G(S^{n-1}, (1/2, 1))$ and $G(S^{n-1}, \{0\})$ are fundamentally different. In the first, every vertex is contained in an open clique, and the independence number is finite. In the second, every vertex is contained in an open independent set, and the independence number is infinite.

A *topological graph* is a graph whose vertex set is a topological space. The graph $G(S^{n-1}, (1/2, 1))$ is an example of a *topological packing graph* [29]; these are topological graphs in which every finite clique is a subset of an open clique. Topological packing graphs with compact vertex sets always have finite independence number.

The graph $G(S^{n-1}, \{0\})$ is, in contrast, an example of a *locally independent graph* [10]; these are topological graphs in which every compact independent set is a subset of an open independent set. If the topology on V is not discrete, then locally independent graphs have infinite independent sets, and concepts like the measurable independence number are used instead of the independence number in these cases. In this paper we focus on optimization hierarchies for upper bounds for the measurable independence number of locally independent graphs with compact vertex sets.

1.2. **Hierarchies.** The *Lovász theta number* of a finite graph $G = (V, E)$ gives an upper bound for the independence number of G. One of its many definitions is as the optimal value of a semidefinite programming problem:

$$\begin{aligned} \vartheta(G) = \max\ & \textstyle\sum_{x,y \in V} A(x, y) \\ & \textstyle\sum_{x \in V} A(x, x) = 1, \\ & A(x, y) = 0 \quad \text{if } xy \in E, \\ & A\colon V^2 \to \mathbb{R} \text{ is positive semidefinite.} \end{aligned} \tag{1}$$

Lovász [31] defined the theta number and showed that $\alpha(G) \leq \vartheta(G)$. Indeed, if $I \subseteq V$ is a nonempty independent set, then the matrix $A(x, y) = |I|^{-1}\chi_I(x)\chi_I(y)$, where χ_I is the characteristic vector of I, is a feasible solution of (1) with objective value $|I|$, as we wanted. Since (1) can be solved in polynomial time [17, Theorem 9.3.30], it provides a polynomial-time computable upper bound for the independence number.

There are basically two approaches to strengthen the Lovász theta number, bringing it closer to the independence number. The first approach is to consider higher-order interactions: in the theta number, we consider interactions between pairs of vertices; in higher-order bounds, we consider interactions between more than two vertices at once. The second approach is to change the cone where A lies: in the theta number, A lies in the cone of positive-semidefinite matrices, but we can consider subcones of this cone instead.

The Lasserre hierarchy [30] is an example of the first approach. In it, instead of considering matrices indexed by vertices, we consider matrices indexed by subsets of vertices. Let $\mathrm{Sub}(V, k)$ denote the set of subsets of V with cardinality at most k. We say that a matrix $A\colon \mathrm{Sub}(V, k)^2 \to \mathbb{R}$ is a *moment matrix* if $A(S, T)$ depends only on $S \cup T$. The *kth level* of the *Lasserre hierarchy* for G is the optimization problem

$$\begin{aligned} \mathrm{lass}_k(G) = \max\ & \textstyle\sum_{x \in V} A(\{x\}, \{x\}) \\ & A(\emptyset, \emptyset) = 1, \\ & A(S, T) = 0 \quad \text{if } S \cup T \text{ is not independent}, \\ & A\colon \mathrm{Sub}(V, k)^2 \to \mathbb{R} \text{ is a positive-semidefinite} \\ & \qquad \text{moment matrix.} \end{aligned} \tag{2}$$

Each level of the Lasserre hierarchy gives an upper bound for the independence number. Indeed, for an independent set $I \subseteq V$ let $A(S, T) = 1$ if $S \cup T \subseteq I$ and $A(S, T) = 0$ otherwise. Then A is a feasible solution of (2) with objective value $|I|$. Moreover, $\mathrm{lass}_k(G)$ can be computed in polynomial time for any fixed k.

The first level coincides with the Lovász theta number: $\mathrm{lass}_1(G) = \vartheta(G)$. Each following level is at least as strong as the one before, and the hierarchy *converges*, even in a finite number of steps: if $k \geq \alpha(G)$, then $\mathrm{lass}_k(G) = \alpha(G)$.

The completely positive formulation is an example of the second approach. In it, we replace in (1) the cone of positive-semidefinite matrices by the cone of *completely* *positive* matrices, namely

$$\mathrm{CP}(V) = \mathrm{cone}\{\, f f^{\mathsf{T}} : f \in \mathbb{R}^V,\ f \geq 0 \,\};$$

call the optimal value of the resulting problem $\vartheta(G, \mathrm{CP}(V))$. The same proof that $\vartheta(G) \geq \alpha(G)$ given above shows that $\vartheta(G, \mathrm{CP}(V)) \geq \alpha(G)$.

De Klerk and Pasechnik [21] observed that a theorem of Motzkin and Straus [34] implies that $\vartheta(G, \mathrm{CP}(V)) = \alpha(G)$. From a computational perspective, at least at first glance, this result is not that interesting: computing the independence number is NP-hard, and so all this tells us is that the completely positive cone is computationally hard. There are several hierarchies of tractable outer approximations of the completely positive cone, however, and using these we obtain converging hierarchies of upper bounds for the independence number.

Bachoc, Nebe, Oliveira, and Vallentin [2] observed that the linear programming bound of Delsarte, Goethals, and Seidel [11] is an extension of the Lovász theta number[1] to the topological packing graph $G(S^{n-1}, (1/2, 1))$. The linear programming bound of Cohn and Elkies for the sphere-packing problem [6], recently used to determine the optimal sphere packings in dimensions 8 and 24 [7, 44], is likewise an extension of the Lovász theta number to the topological packing graph with vertex set $\mathbb{R}^n$ in which x, $y \in \mathbb{R}^n$ are adjacent if $0 < \|x - y\| < 1$.

Bachoc and Vallentin [3] were the first to consider higher-order bounds for geometric problems: their *3-point bound* for the kissing number problem can be seen [26] as an intermediate step between the first and second levels of the Lasserre hierarchy. Later, de Laat and Vallentin [29] extended the Lasserre hierarchy to compact topological packing graphs and proved convergence; more recently, de Laat, de Muinck Keizer, and Machado [27] computed higher levels of the Lasserre hierarchy for specific topological packing graphs arising from the equiangular-lines problem. De Laat, Machado, Oliveira, and Vallentin [26] introduced a *k-point bound* for topological packing graphs that lies between the steps of the Lasserre hierarchy and that for $k = 3$ reduces to the 3-point bound of Bachoc and Vallentin.

Cohn, de Laat, and Salmon [8] introduced a 3-point bound for the sphere-packing problem, extending the 3-point bound of Bachoc and Vallentin to topological packing graphs on the Euclidean space, which is not compact. Cohn and Salmon [9] extended the Lasserre hierarchy to these graphs.

Dobre, Dür, Frerick, and Vallentin [12] defined an infinite-dimensional analogue of the copositive cone, which is the dual of the completely positive cone, and used it to give an exact formulation for the independence number of topological packing graphs. In this way, they extended the result of Motzkin and Straus [34] mentioned above. Kuryatnikova and Vera [23] defined a hierarchy of inner approximations of the infinite-dimensional copositive cone and used it to give a converging hierarchy of upper bounds for the independence number of topological packing graphs.

The picture for topological packing graphs is thus quite complete: on the theoretical side, we have appropriate extensions of the Lasserre hierarchy and of the completely positive/copositive formulations for such graphs, both in the compact and noncompact settings. On the practical side, we can use some of these formulations to obtain better bounds in many cases of interest.

[1] Actually, of the *theta prime* number, which is (1) with the extra constraint that A should be entrywise nonnegative.

For locally independent graphs, the picture is less clear. Bachoc, Nebe, Oliveira, and Vallentin [2] extended the Lovász theta number to locally independent graphs on the sphere, like $G(S^{n-1}, \{0\})$, and used it to compute new bounds for the measurable chromatic number of Euclidean space.[2] Oliveira and Vallentin [38] then extended the theta number to locally independent graphs on $\mathbb{R}^n$, namely to the *unit-distance graph* whose vertex set is $\mathbb{R}^n$ and in which x, $y \in \mathbb{R}^n$ are adjacent if $\|x - y\| = 1$, and obtained better lower bounds for the measurable chromatic number.

DeCorte, Oliveira, and Vallentin [10] extended the completely positive formulation to locally independent graphs on compact spaces and on Euclidean space, and proved that the formulation is exact, that is, that it gives exactly the independence number. The best upper bounds for Witsenhausen's parameter α_n are obtained by this technique.

In this paper, we extend the Lasserre hierarchy and the related k-point bound hierarchy of de Laat, Machado, Oliveira, and Vallentin [26] to locally independent graphs on compact spaces. We also extend the copositive/completely positive hierarchy of Kuryatnikova and Vera [23], and show that it converges for many locally independent graphs, and in particular for Witsenhausen's graph $G(S^{n-1}, \{0\})$. By comparing the k-point bound and the Lasserre hierarchy to the completely positive hierarchy, we show that they also converge.

This is the first time a converging optimization hierarchy is proposed for locally independent graphs. We use the completely positive hierarchy to compute the best upper bounds for Witsenhausen's parameter α_n known to date; see Table 1 and the more detailed Table 2.

2. Preliminaries

Independent sets and the independence number of a graph are defined in §1. A *topological graph* is a graph $G = (V, E)$ where V is a topological space. When G is a topological graph, we see E as a symmetric subset of $V \times V$. Topological packing graphs and locally independent graphs are defined in §1.1.

Let $G = (V, E)$ be a graph where V is a measure space with measure ω. The *measurable independence number* of G is

$$\alpha_\omega(G) = \sup\{\, \omega(I) : I \subseteq V \text{ is independent and measurable} \,\}.$$

We denote by $\operatorname{Aut}(G)$ the group of automorphisms (that is, edge-preserving bijections $V \to V$) of G.

Given a set S and $X \subseteq S$, we denote by $\chi_X \colon S \to \mathbb{R}$ the characteristic function of X, which is such that $\chi_X(x) = 1$ if $x \in X$ and $\chi_X(x) = 0$ otherwise.

We denote by $\mathbf{1}$ the constant-one function or vector and by J the constant-one kernel (see below) or matrix.

2.1. Functional analysis.

All function spaces we consider are real valued. Let V be a measure space equipped with a measure ω. Unless otherwise noted, we always equip $L^2(V)$ with the inner product

$$\langle f, g \rangle = \int_V f(x) g(x) \, d\omega(x).$$

[2]The *measurable chromatic number* of $\mathbb{R}^n$ is the minimum number of colors needed to color the points of $\mathbb{R}^n$ so no two points at distance 1 have the same color and so the sets of points having a same color are Lebesgue measurable.

The L^p norm of a function f is denoted by $\|f\|_p$. If we say that a measurable function is nonnegative, we mean that it is nonnegative almost everywhere.

The functions in $L^2(V^2)$ are called *kernels*. A kernel $K \in L^2(V^2)$ corresponds to the operator $K\colon L^2(V) \to L^2(V)$ such that

$$(Kf)(x) = \int_V K(x,y)f(y)\,d\omega(y)$$

for all $f \in L^2(V)$. The operator K is self-adjoint if and only if K is *symmetric*, that is, $K(x,y) = K(y,x)$ almost everywhere. We denote the set of symmetric kernels by $L^2_{\text{sym}}(V)$.

We say that a kernel $K \in L^2_{\text{sym}}(V)$ is *positive semidefinite* if $\langle Kf, f\rangle \geq 0$ for all $f \in L^2(V)$. Bochner [5] gives a useful characterization of continuous positive semidefinite kernels that we use throughout. Namely, if V is a compact Hausdorff space and if ω is a Radon measure that is positive on open sets, then K is positive semidefinite if and only if $\big(K(x,y)\big)_{x,y\in U}$ is a positive-semidefinite matrix for every finite set $U \subseteq V$.

We denote by $C(V)$ the space of real-valued continuous functions on V and by $C_{\text{sym}}(V)$ the space of continuous symmetric kernels on V^2.

Given functions $f\colon U \to \mathbb{R}$ and $g\colon V \to \mathbb{R}$, we denote by $f \otimes g$ the function from $U \times V$ to $\mathbb{R}$ such that $(f \otimes g)(x,y) = f(x)g(y)$.

There are two natural topologies to consider on $L^2(V)$. The first is the *weak topology*. In this topology, a net (f_α) in $L^2(V)$ converges to $f \in L^2(V)$ if $(\langle f_\alpha, g\rangle)$ converges to $\langle f, g\rangle$ for all $g \in L^2(V)$. The second is the *L^2-norm topology* given by the L^2 norm $\|\cdot\|_2$. The set of continuous linear functionals is the same for both these topologies. As a consequence, if $S \subseteq L^2(V)$ is a convex set, then its closure is the same whether taken in the weak topology or the L^2-norm topology [42, Theorem 5.2(iv)].

2.2. **Spaces of subsets.** Let V be a set. For an integer $k \geq 0$, denote by $\text{Sub}(V,k)$ the collection of all subsets of V with cardinality at most k. For $k \geq 1$ and $v = (v_1,\ldots,v_k) \in V^k$, denote by $[\![v]\!]$ the set $\{v_1,\ldots,v_k\}$. So, for every $k \geq 1$ we have that $[\![\cdot]\!]$ maps V^k to $\text{Sub}(V,k)$. Note that k is superfluous in the definition of $[\![\cdot]\!]$, but not in the definition of the preimage. Hence, given a set $S \subseteq \text{Sub}(V,k)$, we write

$$[\![S]\!]_k^{-1} = \{\, v \in V^k : [\![v]\!] \in S \,\}.$$

If V is a topological space, then we can introduce on $\text{Sub}(V,k) \setminus \{\emptyset\}$ the quotient topology of $[\![\cdot]\!]$ by declaring a set S open if $[\![S]\!]_k^{-1}$ is open in V^k. We define a topology on $\text{Sub}(V,k)$ by taking the disjoint union with $\{\emptyset\}$. For background on the topology on the space of subsets, see Handel [19].

If V is a measure space equipped with a measure ω, then we can turn $\text{Sub}(V,k) \setminus \{\emptyset\}$ into a measure space by considering the pushforward of ω through $[\![\cdot]\!]$. Namely, we declare a set S measurable if $[\![S]\!]_k^{-1}$ is measurable in V^k with respect to the product measure, and we set the measure of S to be the measure of $[\![S]\!]_k^{-1}$. We always define the measure on $\text{Sub}(V,k)$ by setting the measure of $\{\emptyset\}$ to be 1, that is, if $f\colon \text{Sub}(V,k) \to \mathbb{R}$ is a continuous function, then

$$\int_{\text{Sub}(V,k)} f(S)\,d\omega(S) = f(\emptyset) + \int_{V^k} f([\![v]\!])\,d\omega(v).$$

2.3. **Topological groups.** A *topological group* is a group Γ equipped with a topology in which the group operations — multiplication and inversion — are continuous, that is, $(\sigma, \tau) \mapsto \sigma\tau$ is a continuous function from $\Gamma \times \Gamma$ to Γ and $\sigma \mapsto \sigma^{-1}$ is a continuous function from Γ to Γ. We denote the identity element by 1. For us, topological groups are always Hausdorff spaces.

Let Γ be a locally compact group and V be a locally compact Hausdorff space. An *action* of Γ on V is a continuous map $(\sigma, x) \mapsto \sigma x$ from $\Gamma \times V$ to V such that (i) $x \mapsto \sigma x$ is a homeomorphism of V for every $\sigma \in \Gamma$ and (ii) $\sigma(\tau x) = (\sigma\tau)x$ for all σ, $\tau \in \Gamma$ and $x \in V$. We call V a Γ*-space* when there is an action of Γ on V.

The action is called *transitive*, and V is called a *transitive* Γ*-space*, if for every x, $y \in V$ there is $\sigma \in \Gamma$ such that $\sigma x = y$. Let V be a transitive Γ-space and fix $e \in V$; consider the map $p\colon \Gamma \to V$ such that $p(\sigma) = \sigma e$. Note that p is surjective; if it is also open, mapping open subsets of Γ to open subsets of V, then we call V a *homogeneous* Γ*-space*. If Γ is σ-compact, then p is open [13, Proposition 2.44].

If Γ acts on V, then it also acts on functions $f\colon V^k \to \mathbb{R}$ by

$$(\sigma f)(x_1, \ldots, x_k) = f(\sigma^{-1}x_1, \ldots, \sigma^{-1}x_k).$$

We always consider this action on function spaces, unless otherwise noted. We say that a function $f\colon V^k \to \mathbb{R}$ is Γ*-invariant* (or simply *invariant* when the group is clear from context) if $\sigma f = f$ for all $\sigma \in \Gamma$.

There is a useful relation between functions on Γ^k and functions on V^k when V is a homogeneous Γ-space. Namely, fix $e \in V$. From a function $f\colon V^k \to \mathbb{R}$ we can construct a function $g\colon \Gamma^k \to \mathbb{R}$ by setting

$$g(\sigma_1, \ldots, \sigma_k) = f(\sigma_1 e, \ldots, \sigma_k e).$$

If f is continuous, then so is g. Moreover,

$$\text{if } \sigma_i e = \tau_i e \text{ for } i = 1, \ldots, k, \text{ then } g(\sigma_1, \ldots, \sigma_k) = g(\tau_1, \ldots, \tau_k). \tag{3}$$

Conversely, given a function g satisfying (3), we can define $f\colon V^k \to \mathbb{R}$ by setting

$$f(x_1, \ldots, x_k) = g(\sigma_1, \ldots, \sigma_k),$$

where the σ_i are any elements of Γ such that $\sigma_i e = x_i$. Like before, if g is continuous, then f is also continuous. Indeed, given an open set $A \subseteq \mathbb{R}$, we want to show that $f^{-1}(A)$ is open in V^k. Consider the map $p^k\colon \Gamma^k \to V^k$ such that $p^k(\sigma_1, \ldots, \sigma_k) = (\sigma_1 e, \ldots, \sigma_k e)$. Note that $f^{-1}(A) = p^k(g^{-1}(A))$; since p^k is an open map, we are done.

Let Γ be a compact group. We always normalize the Haar measure μ on Γ so $\mu(\Gamma) = 1$. For a compact group, the Haar measure is both left and right invariant, that is, $\mu(\sigma X \tau) = \mu(X)$ for all σ, $\tau \in \Gamma$ and measurable $X \subseteq \Gamma$; moreover, the measure μ_{-1} such that $\mu_{-1}(X) = \mu(\{\, \sigma^{-1} : \sigma \in X \,\})$ coincides with μ.

If V is a homogeneous Γ-space, then V is in particular compact. The pushforward of μ to V is the Radon measure ω such that

$$\int_V f(x)\, d\omega(x) = \int_\Gamma f(\sigma e)\, d\mu(\sigma)$$

for every integrable function $f\colon V \to \mathbb{R}$ and every $e \in V$. This is a Γ-invariant measure: $\omega(\sigma X) = \omega(X)$ for all $\sigma \in \Gamma$ and measurable $X \subseteq V$; it is moreover the unique Γ-invariant Radon measure on V, up to a constant factor [13, Theorem 2.49].

Say Γ is a compact group acting on a compact Hausdorff space V; let μ be the Haar measure on Γ. The *Reynolds operator* maps a measurable function $f\colon V \to \mathbb{R}$ to the Γ-invariant function $\mathcal{R}_\Gamma f$ such that

$$(\mathcal{R}_\Gamma f)(x) = \int_\Gamma f(\sigma x)\, d\mu(\sigma).$$

When using the Reynolds operator, the action of Γ will always be clear from the context. Finally, if V is equipped with a finite measure invariant under the action of Γ, then the Reynolds operator is self-adjoint, that is, $\langle \mathcal{R}_\Gamma f, g\rangle = \langle f, \mathcal{R}_\Gamma g\rangle$.

3. The Lasserre hierarchy and the k-point bound

Let $G = (V, E)$ be a topological graph and ω be a Borel measure on V. For an integer $k \geq 2$, let $M\colon C(\mathrm{Sub}(V, 2k)) \to C(\mathrm{Sub}(V, k)^2)$ be the operator such that

$$(M\nu)(S, T) = \nu(S \cup T).$$

Note that $M\nu$ is indeed continuous since the union map $(S, T) \mapsto S \cup T$ is continuous [19, Proposition 2.14]. The *kth level* of the *Lasserre hierarchy* for G is the optimization problem

$$\begin{aligned} \mathrm{lass}_k(G) = \sup \quad & \textstyle\int_V \nu(\{x\}) d\omega(x) \\ & \nu(\emptyset) = 1, \\ & \nu(S) = 0 \quad \text{if } S \in \mathrm{Sub}(V, 2k) \text{ is not independent}, \\ & M\nu \text{ is positive semidefinite}, \\ & \nu \in C(\mathrm{Sub}(V, 2k)). \end{aligned}$$

Depending on G, this problem could be infeasible. We denote by $\mathrm{lass}_k(G)$ both the optimal value of the problem above and the problem itself, and we follow the same convention for other optimization problems below.

For an integer $k \geq 2$ and a set $Q \subseteq V$ with $|Q| \leq k-2$, let $M_Q\colon C(\mathrm{Sub}(V, k)) \to C(\mathrm{Sub}(V, 1)^2)$ be the operator such that

$$(M_Q\nu)(S, T) = \nu(Q \cup S \cup T);$$

again, $M_Q\nu$ is indeed continuous. The *k-point bound* for G is the optimization problem

$$\begin{aligned} \Delta_k(G) = \sup \quad & \textstyle\int_V \nu(\{x\})\, d\omega(x) \\ & \nu(\emptyset) = 1, \\ & \nu(S) = 0 \quad \text{if } S \in \mathrm{Sub}(V, k) \text{ is not independent}, \\ & M_Q\nu \text{ is positive semidefinite for every independent} \\ & \qquad \text{set } Q \in \mathrm{Sub}(V, k-2), \\ & \nu \in C(\mathrm{Sub}(V, k)). \end{aligned} \tag{4}$$

The restriction of a feasible solution of $\mathrm{lass}_{k+1}(G)$ to $\mathrm{Sub}(V, 2k)$ is continuous [19, Proposition 2.4], hence it is a feasible solution of $\mathrm{lass}_k(G)$. The same can be said about $\Delta_{k+1}(G)$ and $\Delta_k(G)$, and so we have

$$\mathrm{lass}_1(G) \geq \mathrm{lass}_2(G) \geq \cdots \qquad \text{and} \qquad \Delta_2(G) \geq \Delta_3(G) \geq \cdots .$$

Moreover, it is immediate that $\mathrm{lass}_k(G) \leq \Delta_{k+1}(G)$, since if $\nu \in C(\mathrm{Sub}(V, 2k))$ is a feasible solution of $\mathrm{lass}_k(G)$, then for every independent set $Q \subseteq V$ with $|Q| \leq k-1$ and every S, $T \in \mathrm{Sub}(V, 1)$ we have $(M_Q\nu')(S, T) = (M\nu)(Q \cup S, Q \cup T)$, where ν' is the restriction of ν to $\mathrm{Sub}(V, k+1)$; it follows that $M_Q\nu'$ is positive semidefinite.

When V is finite with the discrete topology and the counting measure, $\mathrm{lass}_k(G)$ is simply the Lasserre hierarchy (2) for the independence number, whereas $\Delta_k(G)$ is closely related to restrictions of the Lasserre hierarchy proposed by Gvozdenović, Laurent, and Vallentin [18] and de Laat, Machado, Oliveira, and Vallentin [26]. In this case, we have $\mathrm{lass}_k(G) \geq \alpha(G)$ and $\Delta_k(G) \geq \alpha(G)$ for all k.

The proof is simple. Given an independent set $I \subseteq V$, let $\nu(S) = 1$ if $S \subseteq I$ and $\nu(S) = 0$ otherwise for every $S \in \mathrm{Sub}(V, 2k)$. Then $\nu(\emptyset) = 1$ and $\nu(S) = 0$ if S is not independent. Moreover, since $\nu(\{x_1, \ldots, x_t\}) = \chi_I(x_1) \cdots \chi_I(x_t)$, we see that $M\nu$ is positive semidefinite. So ν is a feasible solution of $\mathrm{lass}_k(G)$ with objective value $|I|$, and since I is any independent set we get $\mathrm{lass}_k(G) \geq \alpha(G)$. In the same way one shows that $\Delta_k(G) \geq \alpha(G)$.

This same proof fails in general for infinite topological graphs, since the function ν constructed above will often not be continuous. This issue can be overcome with extra assumptions.

Theorem 3.1. *Let $G = (V, E)$ be a topological graph and let $\Gamma \subseteq \mathrm{Aut}(G)$ be a compact group. If V is a homogeneous Γ-space and if ω is the pushforward to V of the Haar measure on Γ, then $\mathrm{lass}_k(G) \geq \alpha_\omega(G)$ and $\Delta_k(G) \geq \alpha_\omega(G)$ for all k.*

The main step in the proof uses Lemma 3.2.

Lemma 3.2. *Let Γ be a compact group and let V be a homogeneous Γ-space equipped with the pushforward of the Haar measure μ on Γ. If $f_1, \ldots, f_k \in L^k(V)$, then the function*

$$\mathcal{R}_\Gamma(f_1 \otimes \cdots \otimes f_k)(x_1, \ldots, x_k) = \int_\Gamma f_1(\alpha x_1) \cdots f_k(\alpha x_k)\, d\mu(\alpha)$$

from V^k to $\mathbb{R}$ is continuous.

Proof. In view of §2.3, it suffices to consider $V = \Gamma$. If $f_1, \ldots, f_k$ are continuous, then they are left and right uniformly continuous [13, Proposition 2.6], and the result is immediate.

Thus assume $f_i \in L^k(\Gamma)$ and fix $\epsilon > 0$. Since continuous functions are dense in $L^k(\Gamma)$, let $g_i \in C(\Gamma)$ be such that $\|f_i - g_i\|_k < \epsilon$. Our goal is to find an upper bound for

$$\begin{aligned} &|\mathcal{R}_\Gamma(f_1 \otimes \cdots \otimes f_k)(\sigma_1, \ldots, \sigma_k) - \mathcal{R}_\Gamma(g_1 \otimes \cdots \otimes g_k)(\sigma_1, \ldots, \sigma_k)| \\ &\qquad = \left| \int_\Gamma f_1(\alpha\sigma_1) \cdots f_k(\alpha\sigma_k) - g(\alpha\sigma_1) \cdots g(\alpha\sigma_k)\, d\mu(\alpha) \right| \end{aligned} \tag{5}$$

that is uniform on $\sigma_1, \ldots, \sigma_k \in \Gamma$. If this bound goes to zero with ϵ, then $\mathcal{R}_\Gamma(f_1 \otimes \cdots \otimes f_k)$ is a uniform limit of continuous functions, being therefore continuous, as we want.

Using the triangle inequality, we see that (5) is at most

$$\begin{aligned} \int_\Gamma &|f_1(\alpha\sigma_1) \cdots f_k(\alpha\sigma_k) - f_1(\alpha\sigma_1) \cdots f_{k-1}(\alpha\sigma_{k-1}) g_k(\alpha\sigma_k) \\ &\quad + f_1(\alpha\sigma_1) \cdots f_{k-1}(\alpha\sigma_{k-1}) g_k(\alpha\sigma_k) - g_1(\alpha\sigma_1) \cdots g_k(\alpha\sigma_k)|\, d\mu(\alpha) \\ &\leq \int_\Gamma |f_1(\alpha\sigma_1) \cdots f_{k-1}(\alpha\sigma_{k-1})||f_k(\alpha\sigma_k) - g_k(\alpha\sigma_k)|\, d\mu(\alpha) \\ &\quad + \int_\Gamma |f_1(\alpha\sigma_1) \cdots f_{k-1}(\alpha\sigma_{k-1}) - g_1(\alpha\sigma_1) \cdots g_{k-1}(\alpha\sigma_{k-1})||g_k(\alpha\sigma_k)|\, d\mu(\alpha). \end{aligned}$$

By repeating this procedure, we see that (5) is bounded from above by a sum of k terms of the form

$$\int_\Gamma |f_i(\alpha\sigma_i) - g_i(\alpha\sigma_i)||h_1(\alpha)|\cdots|h_{k-1}(\alpha)|\,d\mu(\alpha), \tag{6}$$

where each h_j is one of the functions $\alpha \mapsto f_l(\alpha\sigma_l)$ or $\alpha \mapsto g_l(\alpha\sigma_l)$ for some l.

A recursive application of Hölder's inequality shows that each summand (6) is at most

$$\|f_i - g_i\|_k \|h_1\|_k \cdots \|h_{k-1}\|_k \le \epsilon M^{k-1},$$

where $M = \epsilon + \max\{\|f_1\|_k, \ldots, \|f_k\|_k\}$. So (5) is bounded from above by $\epsilon k M^{k-1}$, thus $\mathcal{R}_\Gamma(f_1 \otimes \cdots \otimes f_k)$ is a uniform limit of continuous functions and hence continuous. □

Proof of Theorem 3.1. Fix an integer $k \ge 1$ and let $I \subseteq V$ be a measurable independent set. From Lemma 3.2 we know that $F = \mathcal{R}_\Gamma \chi_I^{\otimes 2k}$ is continuous. Moreover, since χ_I is 0–1, we know that $F(v)$ depends only on $[\![v]\!]$: for all u, $v \in V^{2k}$, if $[\![u]\!] = [\![v]\!]$, then $F(u) = F(v)$. So we can define $\nu\colon \mathrm{Sub}(V, 2k) \to \mathbb{R}$ by setting $\nu(\emptyset) = 1$ and $\nu([\![v]\!]) = F(v)$ for all $v \in V^{2k}$.

If we show that ν is continuous, then it is clear that ν is a feasible solution of $\mathrm{lass}_k(G)$, and so $\mathrm{lass}_k(G) \ge \int_V \nu(\{x\})\,d\omega(x) = \omega(I)$. Since I is any measurable independent set, it follows that $\mathrm{lass}_k(G) \ge \alpha_\omega(G)$.

It suffices to show that the restriction $\overline{\nu}$ of ν to $\mathrm{Sub}(V, 2k) \setminus \{\emptyset\}$ is continuous. So let $A \subseteq \mathbb{R}$ be an open set; we want to show that $\overline{\nu}^{-1}(A)$ is open. This is the case, by definition,[3] if $[\![\overline{\nu}^{-1}(A)]\!]_{2k}^{-1}$ is open in V^{2k}. Now note that $[\![\overline{\nu}^{-1}(A)]\!]_{2k}^{-1} = F^{-1}(A)$. Since F is continuous, $F^{-1}(A)$ is open, and we are done.

The proof that $\Delta_k(G) \ge \alpha_\omega(G)$ is similar. □

4. A completely positive hierarchy

Let V be a compact Hausdorff space and ω be a Radon measure on V. The *copositive cone* on V is the convex cone

$$\mathrm{COP}(V) = \{\, A \in L^2_{\mathrm{sym}}(V) : \langle Af, f\rangle \ge 0 \text{ for all } f \in L^2(V) \text{ with } f \ge 0 \,\}.$$

We say that a kernel is *copositive* if it belongs to $\mathrm{COP}(V)$.

Note that $\mathrm{COP}(V)$ is closed in the L^2-norm topology. Since the weak topology and the L^2-norm topology have the same continuous linear functionals, and since $\mathrm{COP}(V)$ is convex, it is also closed in the weak topology (see §2.1).

The dual of $\mathrm{COP}(V)$ is the *completely positive cone* on V, namely

$$\mathrm{CP}(V) = \mathrm{COP}(V)^* = \{\, A \in L^2_{\mathrm{sym}}(V) : \langle A, Z\rangle \ge 0 \text{ for all } Z \in \mathrm{COP}(V) \,\}.$$

We say that a kernel is *completely positive* if it belongs to $\mathrm{CP}(V)$.

Let $\mathcal{K} \subseteq L^2_{\mathrm{sym}}(V)$ be any closed convex cone containing $\mathrm{CP}(V)$. Given a graph $G = (V, E)$, consider the optimization problem

$$\begin{aligned} \vartheta(G, \mathcal{K}) = \sup\ & \langle A, J\rangle \\ & \textstyle\int_V A(x, x)\,d\omega(x) = 1, \\ & A(x, y) = 0 \quad \text{if } xy \in E, \\ & A \in \mathcal{K} \text{ is continuous,} \end{aligned}$$

where J is the constant-one kernel.

[3] Recall that for $S \subseteq \mathrm{Sub}(V, n)$ we write $[\![S]\!]_n^{-1} = \{\, v \in V^n : [\![v]\!] \in S \,\}$.

DeCorte, Oliveira, and Vallentin [10] showed that, if G is locally independent, then $\vartheta(G, \mathcal{K}) \geq \alpha_\omega(G)$. They also showed that $\vartheta(G, \mathrm{CP}(V)) = \alpha_\omega(G)$ under some extra assumptions, extending a result of Motzkin and Straus [22,34] to locally independent graphs. An obvious choice for $\mathcal{K}$ is the cone of positive semidefinite kernels. In this case, we recover the extension of the Lovász theta number introduced by Bachoc, Nebe, Oliveira, and Vallentin [2].

For an integer $r \geq 1$, consider the set

$$\mathcal{C}_r(V) = \{\, A \in L^2_{\mathrm{sym}}(V) : \mathcal{R}_{\mathcal{S}_{r+2}}(A \otimes \mathbf{1}^{\otimes r}) \geq 0 \,\}.$$

Here, the symmetric group $\mathcal{S}_{r+2}$ acts on V^{r+2} by permuting coordinates, so for $F\colon V^{r+2} \to \mathbb{R}$ we have

$$(\mathcal{R}_{\mathcal{S}_{r+2}} F)(x_1, \ldots, x_{r+2}) = \frac{1}{(r+2)!} \sum_{\pi \in \mathcal{S}_{r+2}} F(x_{\pi(1)}, \ldots, x_{\pi(r+2)}).$$

Note that $\mathcal{C}_r(V)$ is a convex cone.

For $A \in L^2_{\mathrm{sym}}(V)$, write $\mathcal{T}_r A = \mathcal{R}_{\mathcal{S}_{r+2}}(A \otimes \mathbf{1}^{\otimes r})$. Use the triangle inequality to show that $\|\mathcal{R}_{\mathcal{S}_{r+2}} F\|_2 \leq \|F\|_2$ for $F \in L^2(V^2)$; it follows that $\mathcal{T}_r A \in L^2(V^{r+2})$ for all $A \in L^2_{\mathrm{sym}}(V)$, and so $\mathcal{T}_r$ is a linear transformation from $L^2_{\mathrm{sym}}(V)$ to $L^2(V^{r+2})$.

Consider the linear transformation $\mathcal{T}_r^*\colon L^2(V^{r+2}) \to L^2_{\mathrm{sym}}(V)$ such that

$$(\mathcal{T}_r^* F)(x, y) = \int_{V^r} (\mathcal{R}_{\mathcal{S}_{r+2}} F)(x, y, v)\, d\omega(v)$$

for $F \in L^2(V^{r+2})$ and x, $y \in V$, where $F(x, y, v)$ is short for $F(x, y, v_1, \ldots, v_r)$ and where we consider on V^r the product measure. One easily verifies that $\mathcal{T}_r^* F \in L^2_{\mathrm{sym}}(V)$. Direct computation shows that $\langle \mathcal{T}_r A, F\rangle = \langle A, \mathcal{T}_r^* F\rangle$.

Both $\mathcal{T}_r$ and $\mathcal{T}_r^*$ are continuous in the weak topology (and also in the norm topology, though we do not need this fact). Indeed, if (A_α) is a net in $L^2_{\mathrm{sym}}(V)$ that converges to A and if $F \in L^2(V^{r+2})$, then

$$\langle \mathcal{T}_r A_\alpha, F\rangle = \langle A_\alpha, \mathcal{T}_r^* F\rangle \to \langle A, \mathcal{T}_r^* F\rangle = \langle \mathcal{T}_r A, F\rangle,$$

and we see that $\mathcal{T}_r A_\alpha \to \mathcal{T}_r A$. In the same way we show that $\mathcal{T}_r^*$ is continuous. It follows that $\mathcal{T}_r^*$ is the adjoint of $\mathcal{T}_r$.

Theorem 4.1 was shown by Kuryatnikova and Vera [23] for the continuous counterpart of $\mathcal{C}_r(V)$.

Theorem 4.1. *Let V be a compact Hausdorff space and let ω be a Radon measure on V. The cone $\mathcal{C}_r(V)$ is closed for every $r \geq 1$, both in the L^2-norm topology and the weak topology, and*

$$\mathcal{C}_1(V) \subseteq \mathcal{C}_2(V) \subseteq \cdots \subseteq \mathrm{COP}(V).$$

The proof requires Lemma 4.2.

Lemma 4.2. *Let V be a Hausdorff space and ω be a Radon measure on V. A function $F \in L^2(V^k)$ is nonnegative if and only if $\langle F, g_1 \otimes \cdots \otimes g_k\rangle \geq 0$ for all nonnegative $g_1, \ldots, g_k \in L^2(V)$.*

Proof. Necessity is easy, so let us prove sufficiency.

Since F is L^2, if it is not nonnegative, then there is a set $X \subseteq V^k$ of finite positive measure such that $\langle F, \chi_X \rangle < 0$. Since ω is outer regular, we can approximate X in measure arbitrarily well by open sets containing X, and so by using the Cauchy-Schwarz inequality we see that there is an open set A of finite measure such that $\langle F, \chi_A \rangle < 0$.

Now ω is also inner regular on open sets. So fix $\epsilon > 0$ and let $C \subseteq A$ be a compact set such that $\omega(A \setminus C) < \epsilon$. Sets of the form $S_1 \times \cdots \times S_k$, where S_1, $\ldots$, $S_k \subseteq V$ are open, are a base of the product topology on V^k. Together with the compactness of C, this means that there is a finite cover $\mathcal{C}$ of C by sets of the form $S_1 \times \cdots \times S_k \subseteq A$, where $S_1, \ldots, S_k \subseteq V$ are open.

Let $\mathcal{B} = \{ S_i : S_1 \times \cdots \times S_k \in \mathcal{C} \text{ and } i = 1, \ldots, k \}$. For $\mathcal{S} \subseteq \mathcal{B}$, write

$$E_{\mathcal{S}} = \bigcap_{S \in \mathcal{S}} S \cap \bigcap_{S \in \mathcal{B} \setminus \mathcal{S}} V \setminus S \tag{7}$$

and $\mathcal{E} = \{ E_{\mathcal{S}} : \mathcal{S} \subseteq \mathcal{B} \text{ and } E_{\mathcal{S}} \neq \emptyset \}$. Note that the sets in $\mathcal{E}$ are pairwise disjoint and that sets of the form $T_1 \times \cdots \times T_k$ with $T_i \in \mathcal{E}$ cover C.

Now let $\mathcal{F} = \{ T_1 \times \cdots \times T_k : T_i \in \mathcal{E} \text{ and } T_1 \times \cdots \times T_k \cap C \neq \emptyset \}$, so $\mathcal{F}$ is a cover of C. We claim that every set in $\mathcal{F}$ is contained in A. Indeed, if $T_1 \times \cdots \times T_k \in \mathcal{F}$, then there is $S_1 \times \cdots \times S_k \in \mathcal{C}$ such that $T_1 \times \cdots \times T_k \cap S_1 \times \cdots \times S_k \neq \emptyset$. This implies that $T_i \cap S_i \neq \emptyset$ for all i. But since T_i is of the form (7), if $T_i \cap S_i \neq \emptyset$, then $T_i \subseteq S_i$, and so $T_1 \times \cdots \times T_k \subseteq S_1 \times \cdots \times S_k \subseteq A$.

Setting $C' = \bigcup \mathcal{F}$ we get

$$\begin{aligned} \langle F, \chi_{C'} \rangle &= \langle F, \chi_C \rangle + \langle F, \chi_{C' \setminus C} \rangle \\ &\leq \langle F, \chi_C \rangle + \|F\|_2 \|\chi_{C' \setminus C}\|_2 \\ &\leq \langle F, \chi_C \rangle + \|F\|_2 \|\chi_{A \setminus C}\|_2 \\ &< \langle F, \chi_C \rangle + \|F\| \epsilon^{1/2}. \end{aligned}$$

By taking $\epsilon \to 0$ we get $\langle F, \chi_C \rangle \to \langle F, \chi_A \rangle < 0$, and so for small enough ϵ we have $\langle F, \chi_{C'} \rangle < 0$. Since the sets in $\mathcal{F}$ are disjoint, there must then be $T_1 \times \cdots \times T_k \in \mathcal{F}$ such that $\langle F, \chi_{T_1 \times \cdots \times T_k} \rangle < 0$. But then setting $g_i = \chi_{T_i}$ we see that $g_i \geq 0$ and that $\langle F, g_1 \otimes \cdots \otimes g_k \rangle < 0$. □

Proof of Theorem 4.1. To see that $\mathcal{C}_r(V)$ is weakly closed, note that $\mathcal{C}_r(V) = \{ A \in L^2_{\mathrm{sym}}(V) : \mathcal{T}_r A \geq 0 \}$. If (A_α) is a net in $\mathcal{C}_r(V)$ converging to A, then since $\mathcal{T}_r$ is weakly continuous we know that $\mathcal{T}_r A_\alpha \to \mathcal{T}_r A$. Finally, since the set of nonnegative functions in $L^2(V^{r+2})$ is weakly closed, we see that $\mathcal{T}_r A \geq 0$, and so $A \in \mathcal{C}_r(V)$. Since $\mathcal{C}_r(V)$ is convex and since the L^2-norm topology and the weak topology have the same continuous functionals, it follows that $\mathcal{C}_r(V)$ is closed in the norm topology as well (see §2.1).

We now show the chain of inclusions. Given a kernel $A \in \mathcal{C}_r(V)$ for some r, we want to show that $A \in \mathcal{C}_{r+1}(V)$, that is, we want to show $\mathcal{R}_{\mathcal{S}_{r+3}}(A \otimes \mathbf{1}^{\otimes(r+1)}) \geq 0$, and for that we use Lemma 4.2. So let $f_1, \ldots, f_{r+3} \in L^2(V)$ be any nonnegative functions. For $i = 1, \ldots, r+3$, let g_i be the tensor product of the functions

$f_1, \ldots, f_{r+3}$, except for f_i, in some arbitrary order. Then

$$\begin{aligned}
&\langle \mathcal{R}_{\mathcal{S}_{r+3}}(A \otimes \mathbf{1}^{\otimes(r+1)}), f_1 \otimes \cdots \otimes f_{r+3}\rangle \\
&\quad = \langle A \otimes 1^{\otimes(r+1)}, \mathcal{R}_{\mathcal{S}_{r+3}}(f_1 \otimes \cdots \otimes f_{r+3})\rangle \\
&\quad = \frac{1}{(r+3)!} \sum_{\pi \in \mathcal{S}_{r+3}} \langle A \otimes \mathbf{1}^{\otimes(r+1)}, f_{\pi(1)} \otimes \cdots \otimes f_{\pi(r+3)}\rangle \\
&\quad = \frac{1}{r+3} \sum_{i=1}^{r+3} \langle A \otimes \mathbf{1}^{\otimes r} \otimes \mathbf{1}, \mathcal{R}_{\mathcal{S}_{r+2}} g_i \otimes f_i\rangle \\
&\quad = \frac{1}{r+3} \sum_{i=1}^{r+3} \langle \mathcal{R}_{\mathcal{S}_{r+2}}(A \otimes \mathbf{1}^{\otimes r}), g_i\rangle \langle \mathbf{1}, f_i\rangle \\
&\quad \geq 0,
\end{aligned}$$

as we wanted.

To see that $\mathcal{C}_r(V) \subseteq \mathrm{COP}(V)$, take any $A \in \mathcal{C}_r(V)$. For every $f \in L^2(V)$ such that $f \geq 0$ and $\langle \mathbf{1}, f\rangle > 0$ we have

$$\begin{aligned}
0 &\leq \langle \mathcal{R}_{\mathcal{S}_{r+2}}(A \otimes \mathbf{1}^{\otimes r}), f^{\otimes(r+2)}\rangle \\
&= \langle A \otimes \mathbf{1}^{\otimes r}, \mathcal{R}_{\mathcal{S}_{r+2}}(f \otimes f \otimes f^{\otimes r})\rangle \\
&= \langle A \otimes \mathbf{1}^{\otimes r}, f \otimes f \otimes f^{\otimes r}\rangle \\
&= \langle A, f \otimes f\rangle \langle \mathbf{1}, f\rangle^r,
\end{aligned}$$

whence $\langle Af, f\rangle = \langle A, f \otimes f\rangle \geq 0$, and so $A \in \mathrm{COP}(V)$. □

The cones $\mathcal{C}_r(V)$ are thus a hierarchy of inner approximations of $\mathrm{COP}(V)$. We will see in §5 some sufficient conditions for this hierarchy to converge, in the sense that the closure of the union of the cones $\mathcal{C}_r(V)$ is $\mathrm{COP}(V)$.

The dual cones $\mathcal{C}_r^*(V)$ provide a hierarchy of outer approximations of $\mathrm{CP}(V)$, namely

$$\mathcal{C}_1^*(V) \supseteq \mathcal{C}_2^*(V) \supseteq \cdots \supseteq \mathrm{CP}(V).$$

Given a graph $G = (V, E)$, set

$$\begin{aligned}
\gamma_r(G) = \sup\ &\langle A, J\rangle \\
&\textstyle\int_V A(x,x)\, d\omega(x) = 1, \\
&A(x,y) = 0 \quad \text{if } xy \in E, \\
&A \in \mathcal{C}_r^*(V) \text{ is continuous and positive semidefinite.}
\end{aligned}$$

This gives a hierarchy of bounds for the measurable independence number, namely

$$\gamma_1(G) \geq \gamma_2(G) \geq \cdots \geq \alpha_\omega(G).$$

In §5 we will prove that this hierarchy converges to the measurable independence number for some classes of graphs. For finite graphs, this hierarchy was proposed (without the positive-semidefiniteness constraint) by de Klerk and Pasechnik [22].

4.1. **Comparison to the Lasserre hierarchy and the k-point bound.** Both the Lasserre hierarchy and the k-point bound of §3 are eventually at least as good as the completely positive hierarchy, hence if the latter converges, so do the other two.

Theorem 4.3. *If $G = (V, E)$ is a topological graph where V is a compact Hausdorff space and if ω is a Radon measure on V, then* $\mathrm{lass}_{r+2}(G) \leq \Delta_{r+3}(G) \leq \gamma_r(G)$ *for every* $r \geq 1$.

Proof. Fix an integer $r \geq 1$. Given a feasible solution of $\Delta_{r+3}(G)$, we will construct a feasible solution of $\gamma_r(G)$ with at least the same objective value, thus proving that $\Delta_{r+3}(G) \leq \gamma_r(G)$. Since $\mathrm{lass}_k(G) \leq \Delta_{k+1}(G)$ for all $k \geq 1$, as shown in §3, the other inequality in the statement will follow as well.

Thus let $\nu \in C(\mathrm{Sub}(V, r+3))$ be a feasible solution of $\Delta_{r+3}(G)$ with positive objective value. Consider the function $F\colon V^{r+2} \to \mathbb{R}$ such that

$$F(x, y, v) = \nu(\llbracket v \rrbracket \cup \{x, y\}) = (M_{\llbracket v \rrbracket}\nu)(\{x\}, \{y\}),$$

where x, $y \in V$ and $v \in V^r$. Note that F is continuous and invariant under $\mathcal{S}_{r+2}$, so $\mathcal{T}_r^* F$ is continuous and given by

$$\begin{aligned}(\mathcal{T}_r^* F)(x, y) &= \int_{V^r} F(x, y, v)\, d\omega(v)\\ &= \int_{V^r} \nu(\llbracket v \rrbracket \cup \{x, y\})\, d\omega(v)\\ &= \int_{V^r} (M_{\llbracket v \rrbracket}\nu)(\{x\}, \{y\})\, d\omega(v).\end{aligned}$$

Since $M_{\llbracket v \rrbracket}\nu$ is positive semidefinite for every $v \in V^r$, it is clear that $\mathcal{T}_r^* F$ is positive semidefinite. We claim that $\mathcal{T}_r^* F \in \mathcal{C}_r^*(V)$. Indeed, if $\llbracket v \rrbracket \cup \{x, y\}$ is not independent, then $F(x, y, v) = 0$; otherwise we have $F(x, y, v) = (M_{\llbracket v \rrbracket \cup \{y\}}\nu)(\{x\}, \{x\}) \geq 0$ since $M_Q \nu$ is positive semidefinite for every independent set Q with $|Q| \leq r+1$. So $F \geq 0$ and for every kernel $Z \in \mathcal{C}_r(V)$ we have $\langle Z, \mathcal{T}_r^* F \rangle = \langle \mathcal{T}_r Z, F \rangle \geq 0$, whence $\mathcal{T}_r^* F \in \mathcal{C}_r^*(V)$.

We clearly have $(\mathcal{T}_r^* F)(x, y) = 0$ if $xy \in E$. Let $\tau = \int_V (\mathcal{T}_r^* F)(x, x)\, d\omega(x)$; we will see soon that $\tau > 0$. Then $A = \tau^{-1} \mathcal{T}_r^* F$ will be a feasible solution of $\gamma_r(G)$; our goal is then to estimate its objective value.

To simplify notation, we define $V^0 = \{\emptyset\}$ and set $\llbracket \emptyset \rrbracket = \emptyset$; we also denote by ω the counting measure on V^0. For an integer $0 \leq t \leq r+3$ write

$$\Phi_t = \int_{V^t} \nu(\llbracket v \rrbracket)\, d\omega(v).$$

We claim that the matrix

$$\begin{pmatrix} \Phi_t & \Phi_{t+1} \\ \Phi_{t+1} & \Phi_{t+2} \end{pmatrix} \tag{8}$$

is positive semidefinite for all $0 \leq t \leq r+1$.

Indeed, fix $0 \leq t \leq r+1$ and let $B\colon \operatorname{Sub}(V,1)^2 \to \mathbb{R}$ be such that

$$B(S,T) = \int_{V^t} \nu([\![v]\!] \cup S \cup T)\, d\omega(v) = \int_{V^t} (M_{[\![v]\!]}\nu)(S,T)\, d\omega(v).$$

It is at once clear that B is a positive-semidefinite kernel and that $B(S,T)$ depends only on $S \cup T$. Moreover,

$$\begin{aligned} B(\emptyset,\emptyset) &= \int_{V^t} \nu([\![v]\!])\, d\omega(v) = \Phi_t, \\ \int_V B(\emptyset,\{x\})\, d\omega(x) &= \int_V \int_{V^t} \nu([\![v]\!] \cup \{x\})\, d\omega(v) d\omega(x) = \Phi_{t+1}, \text{ and} \\ \int_V \int_V B(\{x\},\{y\})\, d\omega(y) d\omega(x) &= \int_V \int_V \int_{V^t} \nu([\![v]\!] \cup \{x,y\})\, d\omega(v) d\omega(y) d\omega(x) \\ &= \Phi_{t+2}, \end{aligned}$$

and so the matrix in (8) is positive semidefinite. It follows that, since $\Phi_0 = 1$ and since ν has objective value $\Phi_1 > 0$, we need to have $\Phi_2 > 0$ as well. Repeating the argument, we see that $\Phi_t > 0$ for all t.

Hence for every fixed t we have $\Phi_t\Phi_{t+2} - \Phi_{t+1}^2 \geq 0$, whence $\Phi_{t+2}\Phi_{t+1}^{-1} \geq \Phi_{t+1}\Phi_t^{-1}$. Apply this inequality repeatedly to get $\Phi_{r+2}\Phi_{r+1}^{-1} \geq \Phi_1\Phi_0^{-1} = \Phi_1$. Now $\tau = \Phi_{r+1}$, and so $\tau > 0$. Moreover, $\Phi_{r+2} = \langle \mathcal{T}_r^* F, J\rangle$, hence $\langle A, J\rangle \geq \Phi_1 = \int_V \nu(\{x\})\, d\omega(x)$, as we wanted. □

5. Convergence

The Lasserre hierarchy converges for finite graphs, that is, if G is any finite graph then $\operatorname{lass}_k(G) = \alpha(G)$ for all $k \geq \alpha(G)$. De Klerk and Pasechnik [22] showed that the completely positive hierarchy γ_r also converges for finite graphs, namely $\gamma_r(G) \to \alpha(G)$ as $r \to \infty$. Since $\operatorname{lass}_k(G) = \operatorname{lass}_{k'}(G)$ for all k, $k' \geq \alpha(G)$, the convergence of the Lasserre hierarchy then follows from Theorem 4.3. Actually, we also get in this way convergence for the k-point bound.

De Laat and Vallentin [29] showed that an extension of the Lasserre hierarchy converges also for topological packing graphs. Kuryatnikova and Vera [23] extended the completely positive hierarchy γ_r from finite graphs to topological packing graphs and showed that it converges.

In this section we will discuss sufficient conditions under which the completely positive hierarchy γ_r converges for locally independent graphs and, using Theorem 4.3, we will obtain convergence results for the Lasserre hierarchy and the k-point bound as consequences. The first step is to better understand the copositive cone $\operatorname{COP}(V)$ and its inner approximations $\mathcal{C}_r(V)$.

5.1. **The copositive cone and its inner approximations.** Let X be a real vector space and let $S \subseteq X$. We say that $x \in S$ is in the *algebraic interior* of S if 0 is in the interior of $\{\, \lambda \in \mathbb{R} : x + \lambda y \in S \,\}$ for every $y \in X$. If X is a topological vector space, then the interior of S is a subset of its algebraic interior.

Let V be a compact Hausdorff space equipped with a Radon measure ω. Write

$$\operatorname{COP}_c(V) = \{\, A \in \operatorname{COP}(V) : A \text{ is continuous} \,\}.$$

In §2.1 we mentioned the observation of Bochner [5] that if ω is positive on open sets, then a continuous kernel is positive semidefinite if and only if each of its finite principal submatrices is positive semidefinite. The same holds for copositive kernels [10, Theorem 4.7]:

Theorem 5.1. *Let V be a compact Hausdorff space equipped with a Radon measure ω that is positive on open sets. A kernel $A \in C_{\mathrm{sym}}(V)$ is copositive if and only if for every finite set $U \subseteq V$ the matrix $\big(A(x,y)\big)_{x,y\in U}$ is copositive.*

As a subset of the vector space $C_{\mathrm{sym}}(V)$, the cone $\mathrm{COP}_{\mathrm{c}}(V)$ has nonempty algebraic interior: the constant-one kernel J is, for instance, in its algebraic interior. If we equip $C_{\mathrm{sym}}(V)$ with the supremum norm, then J is also in the interior of $\mathrm{COP}_{\mathrm{c}}(V)$. In contrast, $\mathrm{COP}(V) \subseteq L^2_{\mathrm{sym}}(V)$ has, in general, empty algebraic interior and hence empty interior, as we will see later.

Kuryatnikova and Vera [23] showed the following result, extending a theorem of de Klerk and Pasechnik [21] (for completeness, a proof can be found in Appendix A):

Theorem 5.2. *If V is a compact Hausdorff space equipped with a Radon measure that is positive on open sets, then every kernel in the algebraic interior of $\mathrm{COP}_{\mathrm{c}}(V) \subseteq C_{\mathrm{sym}}(V)$ belongs to a cone $\mathcal{C}_r(V)$ for some $r \geq 1$.*

Therefore the cones $\mathcal{C}_r(V)$ somehow capture $\mathrm{COP}_{\mathrm{c}}(V)$, since every kernel in $\mathrm{COP}_{\mathrm{c}}(V)$ can be approximated in supremum norm arbitrarily well by kernels in the algebraic interior of $\mathrm{COP}_{\mathrm{c}}(V)$. We need an analogous result for $\mathrm{CP}(V)$ and the outer approximations $\mathcal{C}_r^*(V)$.

Theorem 5.3. *Let Γ be a compact group and V be a homogeneous Γ-space equipped with the pushforward of the Haar measure on Γ. If $A \in L^2_{\mathrm{sym}}(V)$ is such that $\langle A, Z\rangle \geq 0$ for all $Z \in \bigcup_{r\geq 1} \mathcal{C}_r(V)$, then A is completely positive.*

Why is the theorem stated only for kernels on homogeneous spaces? In the proof we use that $\mathrm{COP}_{\mathrm{c}}(V)$ is dense in $\mathrm{COP}(V)$, and though this may well be true in general, the proof provided in Lemma 5.4 below uses the group action to approximate a copositive kernel by a continuous copositive kernel. The analogous result for positive-semidefinite (instead of copositive) kernels, namely that a positive-semidefinite kernel in $L^2_{\mathrm{sym}}(V)$ can be approximated arbitrarily well by continuous positive-semidefinite kernels, follows from the spectral theorem; it is the lack of such a characterization of copositive kernels that explains the less general result above.

Lemma 5.4. *If Γ is a compact group and if V is a homogeneous Γ-space equipped with the pushforward ω of the Haar measure μ on Γ, then $\mathrm{COP}_{\mathrm{c}}(V)$ is dense in $\mathrm{COP}(V)$ in the L^2-norm topology.*

Proof. We first prove the statement for $V = \Gamma$. Let $\mathcal{B}$ be a neighborhood base of $1 \in \Gamma$ consisting of compact symmetric sets [13, Proposition 2.1]. Then $\psi_U = \mu(U)^{-2}\chi_U \otimes \chi_U$ for $U \in \mathcal{B}$ is an approximate identity for the direct product $\Gamma \times \Gamma$ (see Folland [13, §2.5]). Note that $\psi_U(\sigma^{-1}, \tau^{-1}) = \psi_U(\sigma, \tau)$ for all σ, $\tau \in \Gamma$, since the sets in $\mathcal{B}$ are symmetric.

Take $A \in \mathrm{COP}(\Gamma)$. For every $U \in \mathcal{B}$, the convolution

$$\begin{aligned}(\psi_U * A)(\sigma,\tau) &= \int_{\Gamma\times\Gamma} \psi_U(\alpha,\beta)A(\alpha^{-1}\sigma,\beta^{-1}\tau)\,d\mu(\alpha,\beta)\\ &= \int_{\Gamma\times\Gamma} \psi_U(\sigma\alpha,\tau\beta)A(\alpha^{-1},\beta^{-1})\,d\mu(\alpha,\beta)\\ &= \int_{\Gamma\times\Gamma} \psi_U(\alpha^{-1}\sigma^{-1},\beta^{-1}\tau^{-1})A(\alpha^{-1},\beta^{-1})\,d\mu(\alpha,\beta)\\ &= \int_{\Gamma\times\Gamma} \psi_U(\alpha\sigma^{-1},\beta\tau^{-1})A(\alpha,\beta)\,d\mu(\alpha,\beta)\end{aligned}$$

is continuous [13, Proposition 2.39(d)]. Moreover, $\|\psi_U * A - A\|_2 \to 0$ as $U \to \{1\}$, implying that for every $\epsilon > 0$ there is $U \in \mathcal{B}$ such that $\|\psi_U * A - A\|_2 < \epsilon$ (see Folland [13, Proposition 2.42]).

Finally, $\psi_U * A$ is copositive for all $U \in \mathcal{B}$. Indeed, given $f \in L^2(\Gamma)$ with $f \geq 0$, let $g \in L^2(\Gamma)$ be given by

$$g(\alpha) = \int_\Gamma f(\sigma)\chi_U(\alpha\sigma^{-1})\,d\mu(\sigma);$$

note that $g \geq 0$. Since A is copositive we then have

$$\begin{aligned}\langle(\psi_U * A)f, f\rangle &= \mu(U)^{-2}\int_\Gamma\int_\Gamma\int_{\Gamma\times\Gamma} A(\alpha,\beta)\chi_U(\alpha\sigma^{-1})\chi_U(\beta\tau^{-1})\,d\mu(\alpha,\beta)\\ &\qquad\cdot f(\sigma)f(\tau)\,d\tau d\sigma\\ &= \mu(U)^{-2}\int_\Gamma\int_\Gamma A(\alpha,\beta)g(\alpha)g(\beta)\,d\mu(\beta)d\mu(\alpha)\\ &\geq 0,\end{aligned}$$

as we wanted.

So we see that $\mathrm{COP}_c(\Gamma)$ is dense in $\mathrm{COP}(\Gamma)$. Now suppose V is a homogeneous Γ-space and let $A \in \mathrm{COP}(V)$. Use the trick from §2.3: fix $e \in V$ and consider $F \in L^2(\Gamma^2)$ given by $F(\sigma,\tau) = A(\sigma e, \tau e)$. Since A is copositive, so is F; moreover, for every $\epsilon > 0$ there is $U \in \mathcal{B}$ with $\|\psi_U * F - F\|_2 < \epsilon$.

If $\sigma_1 e = \sigma_2 e$ and $\tau_1 e = \tau_2 e$, then

$$\begin{aligned}(\psi_U * F)(\sigma_1,\tau_1) &= \int_{\Gamma\times\Gamma} \psi_U(\alpha,\beta)A(\alpha^{-1}\sigma_1 e,\beta^{-1}\tau_1 e)\,d\mu(\alpha,\beta)\\ &= \int_{\Gamma\times\Gamma} \psi_U(\alpha,\beta)A(\alpha^{-1}\sigma_2 e,\beta^{-1}\tau_2 e)\,d\mu(\alpha,\beta)\\ &= (\psi_U * F)(\sigma_2,\tau_2),\end{aligned}$$

hence there is a function $A_U\colon V^2 \to \mathbb{R}$ such that $(\psi_U * F)(\sigma,\tau) = A_U(\sigma e,\tau e)$; since $\psi_U * F$ is continuous, so is A_U, and since $\psi_U * F$ is copositive, A_U is also copositive (use Theorem 5.1). Moreover, $\|A_U - A\|_2 = \|\psi_U * F - F\|_2 < \epsilon$, and we are done. □

Proof of Theorem 5.3. If $Z \in \mathrm{COP}_c(V)$, then for every $\lambda \in (0,1]$ the kernel $\lambda J + (1-\lambda)Z$ is in the algebraic interior of $\mathrm{COP}_c(V)$. The theorem then follows from Theorem 5.2 together with Lemma 5.4. □

Let us close this section by showing that $\mathrm{COP}(V)$ may have empty algebraic interior. Let V be any measure space equipped with a finite measure ω such that V can be partitioned into infinitely many sets P_1, P_2, ... of positive measure. Take any $A \in \mathrm{COP}(V)$.

For an integer $n \geq 1$ write

$$f(n) = \frac{1}{\omega(P_n)^2} \int_{P_n} \int_{P_n} A(x, y)\, d\omega(y) d\omega(x).$$

Note that, since A is copositive, $f(n) \geq 0$ for all n. If for some n we have $f(n) = 0$, then for every $\lambda > 0$ we have

$$\begin{aligned}\int_{P_n} \int_{P_n} (A - \lambda J)(x, y)\, d\omega(y) d\omega(x) &= f(n)\omega(P_n)^2 - \lambda\omega(P_n)^2 \\ &= (f(n) - \lambda)\omega(P_n)^2 \\ &= -\lambda\omega(P_n)^2 < 0,\end{aligned}$$

and so $A - \lambda J$ is not copositive and hence A is not in the algebraic interior of $\mathrm{COP}(V)$. So assume $f(n) > 0$ for all n.

Use Cauchy-Schwarz to get

$$\left(\int_{P_n} \int_{P_n} A(x, y)\, d\omega(y) d\omega(x)\right)^2 \leq \omega(P_n)^2 \int_{P_n} \int_{P_n} A(x, y)^2\, d\omega(y) d\omega(x),$$

whence $\sum_{n \geq 1} f(n)^2 \omega(P_n)^2 < \infty$.

Now let $g(n)$ be such that $g(n) \to \infty$ and $\sum_{n \geq 1} f(n)^2 g(n)^2 \omega(P_n)^2 < \infty$. Set

$$K = \sum_{n \geq 1} f(n) g(n) \chi_{P_n} \otimes \chi_{P_n}.$$

Note that $\|K\|_2^2 = \sum_{n \geq 1} f(n)^2 g(n)^2 \omega(P_n)^2 < \infty$, so $K \in L^2_{\mathrm{sym}}(V)$. Moreover, for every $\lambda > 0$ we have

$$\begin{aligned}\int_{P_n} \int_{P_n} (A - \lambda K)(x, y)\, d\omega(y) d\omega(x) &= f(n)\omega(P_n)^2 - \lambda f(n) g(n) \omega(P_n)^2 \\ &= f(n)\omega(P_n)^2 (1 - \lambda g(n)),\end{aligned}$$

and since $g(n) \to \infty$ and $f(n) > 0$ for all n we see that for every $\lambda > 0$ the kernel $A - \lambda K$ is not copositive, and hence A is not in the algebraic interior of $\mathrm{COP}(V)$.

5.2. **Convergence for distance graphs.** Given a locally independent graph $G = (V, E)$ and a measure ω on V, our goal is to show that $\gamma_r(G)$ converges to $\alpha_\omega(G)$. Then, using Theorem 4.3, we get convergence of both the Lasserre hierarchy and the k-point bound.

Our strategy is as follows. For $r \geq 1$, let A_r be a feasible solution of $\gamma_r(G)$. Note that the feasible region of $\gamma_r(G)$ is contained in the feasible region of $\gamma_1(G)$. So, if the feasible region of $\gamma_1(G)$ is contained in a compact set, then the sequence (A_r) will have a converging subsequence, say with limit A. The goal is then to show that A is a feasible solution of the completely positive formulation $\vartheta(G, \mathrm{CP}(V))$, so we can finish by using the result of DeCorte, Oliveira, and Vallentin [10].

Summarizing, we need to

(1) show that the feasible region of $\gamma_1(G)$ is contained in some compact set in a convenient topology and
(2) show that the limit solution A is feasible for $\vartheta(G, \mathrm{CP}(V))$.

There is an interplay between (1) and (2), since the topology influences what we know about the limit A, and so if we choose a bad topology, A will not be feasible. As for (2), the main issue seems to be the constraints "$A(x, y) = 0$ for $xy \in E$"; depending on the graph, A may not satisfy these constraints even though each A_r does.

For distance graphs on the sphere (and related spaces; see below), this strategy works.

Theorem 5.5. *Let ω be the surface measure on S^{n-1}. If $n \geq 3$ and if $D \subseteq (-1, 1)$ is closed, then $\lim_{r\to\infty} \gamma_r(G) = \alpha_\omega(G)$ where $G = G(S^{n-1}, D)$.*

So the γ_r hierarchy converges for Witsenhausen's graph $G = G(S^{n-1}, \{0\})$, as long as $n \geq 3$. For $n = 2$, Theorem 5.5 does not provide any information, even though in this case $\gamma_1(G) = \alpha_\omega(G)$ (cf. Oliveira [36, §3.5a]).

The proof requires some facts about harmonic analysis on the sphere; see Andrews, Askey, and Roy [1] for background. Let ω be the surface measure on S^{n-1} and write $\omega_n = \omega(S^{n-1})$.

For $k \geq 0$, let $H_k^n \subseteq C(S^{n-1})$ be the space of n-variable spherical harmonics of degree k; denote the dimension of H_k^n by h_k^n and let $S_{k,1}^n, \ldots, S_{k,h_k^n}^n$ be an orthonormal basis of H_k^n. The set of all spherical harmonics $S_{k,i}^n$ for all k and i forms a complete orthonormal system of $L^2(S^{n-1})$.

Let P_k^n be the Jacobi polynomial of degree k and parameters $\alpha = \beta = (n-3)/2$ normalized so $P_k^n(1) = 1$. The addition formula [1, Theorem 9.6.3] states that

$$P_k^n(x \cdot y) = \frac{\omega_n}{h_k^n} \sum_{i=1}^{h_k^n} S_{k,i}^n(x) S_{k,i}^n(y)$$

for all $x, y \in S^{n-1}$.

Write $E_k^n(x, y) = P_k^n(x \cdot y)$ and let $\mathrm{O}(n)$ be the orthogonal group on $\mathbb{R}^n$. The kernels E_k^n are $\mathrm{O}(n)$-invariant and form a complete orthogonal system of the space $L^2((S^{n-1})^2)^{\mathrm{O}(n)}$ of $\mathrm{O}(n)$-invariant kernels. We have $\langle E_k^n, E_l^n \rangle = 0$ if $k \neq l$ and $\langle E_k^n, E_k^n \rangle = \omega_n^2 / h_k^n$. So if $A = \sum_{k=0}^\infty f(k) E_k^n$ and $B = \sum_{k=0}^\infty g(k) E_k^n$, then

$$\langle A, B \rangle = \sum_{k=0}^\infty f(k) g(k) \omega_n^2 / h_k^n. \tag{9}$$

From the addition formula, each E_k^n is positive semidefinite, hence $\sum_{k=0}^\infty f(k) E_k^n$ is positive semidefinite if and only if $f(k) \geq 0$ for all k. Schoenberg's theorem [41] states that if a kernel $K \colon (S^{n-1})^2 \to \mathbb{R}$ is continuous, $\mathrm{O}(n)$-invariant, and positive semidefinite, then there is a nonnegative sequence $f \colon \mathbb{N} \to \mathbb{R}$ with $\sum_{k=0}^\infty f(k) < \infty$ such that

$$K(x, y) = \sum_{k=0}^\infty f(k) E_k^n(x, y) = \sum_{k=0}^\infty f(k) P_k^n(x \cdot y) \tag{10}$$

with absolute and uniform convergence over $S^{n-1} \times S^{n-1}$.

Conversely, if $f \colon \mathbb{N} \to \mathbb{R}$ is a nonnegative sequence such that $\sum_{k=0}^\infty f(k) < \infty$, then the series in (10) converges absolutely and uniformly over $S^{n-1} \times S^{n-1}$ and the kernel K it defines is continuous, $\mathrm{O}(n)$-invariant, and positive semidefinite.

Proof of Theorem 5.5. We know from §4 that $\lim_{r\to\infty} \gamma_r(G) \geq \vartheta(G, \mathrm{CP}(S^{n-1}))$; let us show the reverse inequality, thus establishing the result by using Theorem 5.1 from DeCorte, Oliveira, and Vallentin [10].

The cone $\mathcal{C}_r(S^{n-1})$ is invariant under $\mathrm{O}(n) \subseteq \mathrm{Aut}(G)$, so if A is a feasible solution of $\gamma_r(G)$, then so is its symmetrization $\mathcal{R}_{\mathrm{O}(n)}A$. It follows that we may restrict ourselves in $\gamma_r(G)$ to $\mathrm{O}(n)$-invariant kernels.

For $r \geq 1$, let A_r be any $\mathrm{O}(n)$-invariant feasible solution of $\gamma_r(G)$ such that $\langle A_r, J\rangle \geq \alpha_\omega(G)/2 > 0$. For each $r \geq 1$, use Schoenberg's theorem to get a nonnegative sequence $f_r\colon \mathbb{N} \to \mathbb{R}$ such that $A_r = \sum_{k=0}^\infty f_r(k) E_k^n$. Note that

$$\omega_n \sum_{k=0}^{\infty} f_r(k) = \int_{S^{n-1}} A_r(x,x)\, d\omega(x) = 1 \tag{11}$$

for every $r \geq 1$.

Let ℓ^1 be the space of all sequences $f\colon \mathbb{N} \to \mathbb{R}$ such that $\|f\|_1 = \sum_{k=0}^\infty |f(k)| < \infty$ and let c_0 be the space of all sequences $f\colon \mathbb{N}\to\mathbb{R}$ that vanish at infinity. Under the supremum norm, c_0 is a Banach space and its dual is ℓ^1; the duality pairing is $(f,g) = \sum_{k=0}^\infty f(k)g(k)$.

Let $\mathcal{B} = \{\, f \in \ell^1 : \|f\|_1 \leq 1\,\}$; from (11) we know that $\omega_n f_r \in \mathcal{B}$ for all r. Alaoglu's theorem [14, Theorem 5.18] says that $\mathcal{B}$ is weak* compact. Since the weak* topology on $\mathcal{B}$ is also metrizable [14, p. 171, Exercise 50], the sequence (f_r) has a converging subsequence; assume that (f_r) itself converges to f. This means that, if $g \in c_0$, then $\lim_{r\to\infty}(f_r, g) = (f,g)$.

It follows immediately that f is nonnegative and, using (11), that $\omega_n\|f\|_1 \leq 1$. So by Schoenberg's theorem the kernel $A = \sum_{k=0}^\infty f(k)E_k^n$ is continuous, $\mathrm{O}(n)$-invariant, and positive semidefinite. Moreover, if $B \in L^2((S^{n-1})^2)^{\mathrm{O}(n)}$, then $B = \sum_{k=0}^\infty g(k) E_k^n$ for some sequence g, and since $k \mapsto g(k)\omega_n^2/h_k^n$ vanishes at infinity we have from (9) that

$$\lim_{r\to\infty} \langle A_r, B\rangle = \langle A, B\rangle.$$

We then have that $\langle A, J\rangle = \lim_{r\to\infty}\langle A_r, J\rangle > 0$, and so A is nonzero and hence $0 < \tau = \int_V A(x,x)\, d\omega(x) = \omega_n \|f\|_1 \leq 1$. We show that A is completely positive and that $A(x,y) = 0$ whenever $x\cdot y \in D$, thus showing that $\tau^{-1}A$ is a feasible solution of $\vartheta(G, \mathrm{CP}(S^{n-1}))$ and hence that $\vartheta(G, \mathrm{CP}(S^{n-1})) \geq \lim_{r\to\infty}\gamma_r(G)$, finishing the proof.

If $Z \in \mathcal{C}_s(S^{n-1})$ for some s, then

$$\langle A, Z\rangle = \langle A, \mathcal{R}_{\mathrm{O}(n)}Z\rangle = \lim_{r\to\infty}\langle A_r, \mathcal{R}_{\mathrm{O}(n)}Z\rangle \geq 0,$$

and using Theorem 5.3 we see that A is completely positive. Next, since $n \geq 3$ the asymptotic formula for the Jacobi polynomials [43, Theorem 8.21.8] implies that $k \mapsto P_k^n(t)$ vanishes at infinity for every $t \in D \subseteq (-1,1)$. So, if x, $y \in S^{n-1}$ are such that $x \cdot y \in D$, then

$$A(x,y) = \sum_{k=0}^\infty f(k) P_k^n(x\cdot y) = \lim_{r\to\infty} \sum_{k=0}^\infty f_r(k) P_k^n(x \cdot y) = 0,$$

and $A(x,y) = 0$ whenever $x\cdot y \in D$, as we wanted. □

The reason we had to require $n \geq 3$ in Theorem 5.5 is to get the sequence $k \mapsto P_k^n(t)$ to vanish at infinity for $t \in D$ so the limit A would satisfy the edge constraints. The same trick works also for manifolds similar to the sphere, namely the continuous, compact, two-point homogeneous spaces [45]: the sphere, the real, complex, and quaternionic projective spaces, and the octonionic projective plane. For these spaces, a theorem like Schoenberg's theorem [37, Theorem 3.1] gives a characterization of continuous, invariant, and positive-semidefinite kernels in terms of an expansion like (10) involving Jacobi polynomials. As long as the real dimension of the space is at least 2, we get sequences vanishing at infinity as we like. So for these spaces both Theorem 5.5 and its proof can be easily adapted.

5.3. **Convergence on homogeneous spaces.** Let X be a topological space and μ be a Borel measure on X. We say that a measurable set $U \subseteq X$ is *μ-thick* if for every open set $A \subseteq X$ such that $A \cap U \neq \emptyset$ we have $\mu(A \cap U) > 0$.

Theorem 5.6. *Let $G = (V, E)$ be a locally independent graph and let $\Gamma \subseteq \operatorname{Aut}(G)$ be a compact group. If V is a homogeneous Γ-space, if ω is the pushforward to V of the Haar measure on Γ, and if E is ω-thick, then $\lim_{r\to\infty} \gamma_r(G) = \vartheta(G, \mathrm{CP}(V))$.*

An example of a graph satisfying the hypotheses of Theorem 5.6 is $G(S^{n-1}, [a, b])$ where $-1 \leq a < b < 1$. Another example, not covered by Theorem 5.5, is the graph on the torus $\mathbb{R}^n/\mathbb{Z}^n$ where two points are adjacent if their geodesic distance lies in $[a, b]$. In contrast, Witsenhausen's graph $G(S^{n-1}, \{0\})$ does not satisfy the hypotheses.

Note that the theorem does not state that $\gamma_r(G) \to \alpha_\omega(G)$. To prove this we need to know that $\vartheta(G, \mathrm{CP}(V)) = \alpha_\omega(G)$, and for this we need a few extra technical assumptions [10, Theorem 5.1] on Γ and G.

For the proof we will need Lemma 5.7.

Lemma 5.7. *Let Γ be a compact group, V be a homogeneous Γ-space, and ω be the pushforward to V of the Haar measure on Γ. If $A \in L^2_{\mathrm{sym}}(V)$ is trace class, then $\mathcal{R}_\Gamma A$ is continuous.*

Proof. Consider the spectral decomposition of A, namely

$$A = \sum_{k=1}^{\infty} \lambda_k f_k \otimes f_k$$

with L^2 convergence, where $f_k \in L^2(V)$ are pairwise orthogonal and $\|f_k\|_2 = 1$ for all k. Since A is trace class, $\sum_{k=1}^\infty |\lambda_k| < \infty$.

With $E_k = \mathcal{R}_\Gamma(f_k \otimes f_k)$ we have $\mathcal{R}_\Gamma A = \sum_{k=1}^\infty \lambda_k E_k$ with L^2 convergence. From Lemma 3.2, each E_k is continuous. Since Γ acts transitively, we also have $E_k(x, x) = \langle f_k, f_k \rangle = 1$ for all k.

Each kernel E_k is continuous and positive semidefinite, so given x, $y \in V$ the matrix

$$\begin{pmatrix} E_k(x,x) & E_k(x,y) \\ E_k(x,y) & E_k(y,y) \end{pmatrix}$$

is positive semidefinite, whence $|E_k(x, y)| \leq 1$. Given $\epsilon > 0$, let N be such that $\sum_{k=N}^\infty |\lambda_k| < \epsilon$. Then for all x, $y \in V$ we have

$$\left| \sum_{k=N}^{\infty} \lambda_k E_k(x,y) \right| \leq \sum_{k=N}^{\infty} |\lambda_k||E_k(x,y)| \leq \sum_{k=N}^{\infty} |\lambda_k| < \epsilon.$$

So the series $\sum_{k=1}^\infty \lambda_k E_k(x, y)$ converges absolutely and uniformly over $V \times V$, and A is continuous. □

Proof of Theorem 5.6. We know from §4 that $\lim_{r\to\infty} \gamma_r(G) \geq \vartheta(G, \mathrm{CP}(V))$; let us show the reverse inequality.

The cone $\mathcal{C}_r(V)$ is invariant under Γ, so if A is a feasible solution of $\gamma_r(G)$, then so is its symmetrization $\mathcal{R}_\Gamma A$. It follows that we may restrict ourselves in $\gamma_r(G)$ to Γ-invariant kernels.

For $r \geq 1$, let A_r be any Γ-invariant feasible solution of $\gamma_r(G)$ such that $\langle A, J\rangle \geq \alpha_\omega(G)/2 > 0$. Since A_r is Γ-invariant, $A_r(x,x) = \int_V A_r(y,y)\, d\omega(y) = 1$ for all x. Since A_r is continuous and positive semidefinite, for all x, $y \in V$ the matrix

$$\begin{pmatrix} A_r(x,x) & A_r(x,y) \\ A_r(x,y) & A_r(y,y) \end{pmatrix}$$

is positive semidefinite, whence $|A_r(x,y)| \leq 1$. So for $r \geq 1$ we have $\|A_r\|_2 \leq 1$, thus $A_r \in \mathcal{B} = \{\, F \in L^2_{\mathrm{sym}}(V) : \|F\|_2 \leq 1 \,\}$ for all $r \geq 1$.

Alaoglu's theorem [14, Theorem 5.18] says that $\mathcal{B}$ is compact in the weak topology of $L^2_{\mathrm{sym}}(V)$. The weak topology on $\mathcal{B}$ is moreover metrizable [14, p. 171, Exercise 50], so the sequence (A_r) has a weakly converging subsequence; assume that the sequence itself converges to A.

This means that for all $B \in L^2_{\mathrm{sym}}(V)$ we have

$$\lim_{r\to\infty} \langle A_r, B\rangle = \langle A, B\rangle.$$

In particular, $\langle A, J\rangle = \lim_{r\to\infty}\langle A_r, J\rangle > 0$, and so $A \neq 0$.

If $Z \in \mathcal{C}_s(V)$ for some s, then $\langle A, Z\rangle = \lim_{r\to\infty}\langle A_r, Z\rangle \geq 0$, and so from Theorem 5.3 it follows that A is completely positive. We claim that A is trace class and $\operatorname{tr} A \leq 1$.

Indeed, consider the spectral decomposition

$$A = \sum_{k=1}^{\infty} \lambda_k f_k \otimes f_k$$

of A, where $f_k \in L^2(V)$ are pairwise orthogonal and $\|f_k\|_2 = 1$ for all k. Since A is positive semidefinite (as it is completely positive), we have $\lambda_k \geq 0$ for all k. Since every A_r is positive semidefinite as well, for every $N \geq 1$ we have

$$\begin{aligned} \sum_{k=1}^{N} \lambda_k &= \Big\langle A, \sum_{k=1}^{N} f_k \otimes f_k \Big\rangle \\ &= \lim_{r\to\infty} \Big\langle A_r, \sum_{k=1}^{N} f_k \otimes f_k \Big\rangle \\ &\leq \lim_{r\to\infty} \operatorname{tr} A_r = 1, \end{aligned}$$

hence $\operatorname{tr} A \leq 1$.

From Lemma 5.7 we see that $\mathcal{R}_\Gamma A$ is continuous. Then, since E is ω-thick, $(\mathcal{R}_\Gamma A)(x,y) = 0$ for all $xy \in E$ if and only if $\langle \mathcal{R}_\Gamma A, B\rangle = 0$ for all $B \in L^2_{\mathrm{sym}}(V)$ with support contained in E. For any such B, since also $\Gamma \subseteq \mathrm{Aut}(G)$, we have

$$\langle \mathcal{R}_\Gamma A, B\rangle = \langle A, \mathcal{R}_\Gamma B\rangle = \lim_{r\to\infty} \langle A_r, \mathcal{R}_\Gamma B\rangle = 0,$$

and so $(\mathcal{R}_\Gamma A)(x,y) = 0$ for all $xy \in E$.

Finally, since A is nonzero and positive semidefinite we have $\operatorname{tr} A > 0$. From Mercer's theorem we see that $\int_V (\mathcal{R}_\Gamma A)(x,x)\, d\omega(x) = \operatorname{tr} A \leq 1$. So $(\operatorname{tr} A)^{-1}(\mathcal{R}_\Gamma A)$

is a feasible solution of $\vartheta(G, \mathrm{CP}(V))$ and hence $\lim_{r\to\infty} \gamma_r(G) \le \vartheta(G, \mathrm{CP}(V))$, as we wanted. □

6. Computations

We have seen three hierarchies for the independence number of locally independent graphs and conditions under which they converge, but can we use them to get better bounds for, say, Witsenhausen's problem?

It turns out that the Lasserre hierarchy and the k-point bound are hard to compute in this case, as we will see in §6.1. This motivates the introduction of another completely positive hierarchy, stronger than the one of §4, which we use to obtain better bounds for Witsenhausen's problem (see Tables 1 and 2).

In this section, we write $\langle X, Y\rangle = \operatorname{tr} X^{\mathsf{T}} Y$ for the trace inner product between matrices X, $Y \in \mathbb{R}^{n\times n}$.

6.1. The 3-point bound.

Our goal is to compute $\Delta_3(G_n)$ for Witsenhausen's graph $G_n = G(S^{n-1}, \{0\})$. To simplify the discussion, we use an alternative normalization for (4), namely

$$
\begin{aligned}
\sup\ & \textstyle\int_{S^{n-1}} \int_{S^{n-1}} \nu(\{x, y\})\, d\omega(y) d\omega(x)\\
& \textstyle\int_{S^{n-1}} \nu(\{x\}) = 1,\\
& \nu(S) = 0 \quad \text{if } S \in \mathrm{Sub}(V, 3) \text{ is not independent},\\
& M_\emptyset \nu \text{ is positive semidefinite},\\
& M_{\{u\}}\nu \text{ is positive semidefinite for every } u \in S^{n-1},\\
& \nu \in C(\mathrm{Sub}(V, 3)).
\end{aligned}
$$

With this normalization, $\Delta_3(G_n)$ is at most the optimal value of the problem above. It is not hard to give a direct proof of this assertion; the relation between different normalizations is developed in the proof of Theorem 4.3.

We simplify the problem further by disregarding the empty set in $M_Q\nu$. More precisely, we change the operator M_Q, where $|Q| \le 1$, so that it maps the function ν to the kernel $K \in C((S^{n-1})^2)$ such that $K(x, y) = \nu(Q \cup \{x, y\})$.

When solving the resulting problem we may restrict ourselves to $\mathrm{O}(n)$-invariant functions ν, since we can apply the Reynolds operator to any feasible solution ν and obtain as a result an $\mathrm{O}(n)$-invariant feasible solution with the same objective value. Let now $e \in S^{n-1}$ be the north pole (or any other fixed point). Given any $u \in S^{n-1}$ there is an orthogonal transformation T such that $Tu = e$, hence if ν is an $\mathrm{O}(n)$-invariant feasible solution, then

$$
(M_{\{u\}}\nu)(x, y) = \nu(\{u, x, y\}) = \nu(\{e, Tx, Ty\}) = (M_{\{e\}}\nu)(Tx, Ty).
$$

It follows that if $M_{\{e\}}\nu$ is positive semidefinite, then so is $M_{\{u\}}\nu$ for every $u \in S^{n-1}$. Note moreover that, since ν is $\mathrm{O}(n)$-invariant, then $M_{\{e\}}\nu$ is invariant under the stabilizer $\mathrm{Stab}(\mathrm{O}(n), e)$ of e. This allows us to rewrite the problem by considering the two kernels $A = M_\emptyset \nu$ and $K = M_{\{e\}}\nu$:

(12)

$$\begin{array}{ll} \sup & \int_{S^{n-1}} \int_{S^{n-1}} A(x,y)\, d\omega(y) d\omega(x) \\ & \int_{S^{n-1}} A(x,x)\, d\omega(x) = 1, \\ & A(x,y) = K(e,Ty) \quad \text{for every } x,\, y \in S^{n-1} \text{ and } T \in \mathrm{O}(n) \text{ with } Tx = e, \\ & K(e,x) = K(x,x) \quad \text{for every } x \in S^{n-1}, \\ & K(x,y) = 0 \quad \text{if } \{e,x,y\} \text{ is not independent}, \\ & A \in C((S^{n-1})^2) \text{ is } \mathrm{O}(n)\text{-invariant and positive semidefinite}, \\ & K \in C((S^{n-1})^2) \text{ is } \mathrm{Stab}(\mathrm{O}(n),e)\text{-invariant and positive semidefinite.} \end{array}$$

Since A is $\mathrm{O}(n)$-invariant, Schoenberg's theorem can be used to express A in terms of Jacobi polynomials as in (10). The kernel K is invariant under the stabilizer subgroup $\mathrm{Stab}(\mathrm{O}(n), e)$; the parametrization of such a kernel was given by Bachoc and Vallentin [3] and is as follows.

With P_k^n being the Jacobi polynomial used in (10), for $n \geq 2$ and $k \geq 0$ consider the polynomial

$$Q_k^n(u,v,t) = (1-u^2)^{k/2}(1-v^2)^{k/2} P_k^n\left(\frac{t-uv}{(1-u^2)^{1/2}(1-v^2)^{1/2}}\right)$$

and let $Y_{k,d}^n$ be the $(d-k+1) \times (d-k+1)$ matrix given by

$$(Y_{k,d}^n)_{ij}(u,v,t) = u^i v^j Q_k^{n-1}(u,v,t) \tag{13}$$

for $0 \leq i, j \leq d-k$.

If $K \in C((S^{n-1})^2)$ is $\mathrm{Stab}(\mathrm{O}(n), e)$-invariant, then $K(x,y)$ depends only on $e \cdot x$, $e \cdot y$, and $x \cdot y$. Bachoc and Vallentin showed that, for any d and any choice of positive semidefinite matrices $F_k \in \mathbb{R}^{(d-k+1)\times(d-k+1)}$, the kernel

$$K(x,y) = \sum_{k=0}^{d} \langle F_k, Y_{k,d}^n(e\cdot x, e\cdot y, x\cdot y)\rangle \tag{14}$$

is $\mathrm{Stab}(\mathrm{O}(n), e)$-invariant and positive semidefinite. (Note that K is continuous by construction, since it is a polynomial on the three inner products above.)

We would like to say that, conversely, any continuous, positive-semidefinite, and $\mathrm{Stab}(\mathrm{O}(n), e)$-invariant kernel can be expressed as in (14) with pointwise uniform convergence, like Schoenberg's theorem asserts for $\mathrm{O}(n)$-invariant kernels. This cannot be the case with (14) as it is given, since it only captures kernels coming from polynomials. Even if we use in (14) an infinite sum and allow the matrices F_k to be infinite, we can still not achieve pointwise convergence. What can be shown [24, Theorem A.8] however is that any continuous, positive-semidefinite, and $\mathrm{Stab}(\mathrm{O}(n), e)$-invariant kernel can be uniformly approximated by kernels of the form (14) by taking larger and larger degree d.

To solve (12) we parameterize A using Schoenberg's theorem and K using (14). A first problem pops up: even if we take $d = \infty$ and use (14) with infinite matrices, since we do not have pointwise convergence the constraint "$K(x,y) = 0$ if $\{e,x,y\}$ is not independent" cannot be equivalently rewritten in terms of the expansion (14), hence it is unclear that the resulting problem gives an upper bound for $\alpha_\omega(G_n)$. The same happens with other constraints in (12).

We could circumvent this problem by requiring that $K(x,y) \in [-\epsilon, \epsilon]$ for some fixed $\epsilon > 0$, rewriting other constraints in a similar fashion. Still, using the resulting problem to get a rigorous upper bound for the measurable independence number remains difficult.

Indeed, to solve the modified problem (12) we have to fix the degrees of the polynomials at some point. Note that we have to solve the problem (or a relaxation of it) to optimality to get an upper bound. To do so rigorously we have to use polynomials of high degree, and since K is parameterized by a 3-variable polynomial, the variable matrices become prohibitively large.

6.2. **Slice-positive functions and another hierarchy.** Let V be a measure space equipped with a measure ω. A function $F \in L^2(V^{k+2})$ is *slice-positive* if for almost all $v \in V^k$ the kernel $(x, y) \mapsto F(x, y, v)$ is positive semidefinite. If V is a compact Hausdorff space and ω is a Radon measure that is positive on open sets, then a continuous function $F\colon V^{k+2} \to \mathbb{R}$ is slice-positive if and only if $(x, y) \mapsto F(x, y, v)$ is a positive-semidefinite kernel for every $v \in V^k$.

For an integer $r \geq 1$, let

$$\mathcal{Q}_r(V) = \{\, A \in L^2_{\mathrm{sym}}(V) : \mathcal{R}_{\mathcal{S}_{r+2}}(A \otimes \mathbf{1}^{\otimes r} - F) \geq 0 \\ \text{for some slice-positive } F \in L^2(V^{r+2}) \,\}.$$

It is immediate that $\mathcal{C}_r(V) \subseteq \mathcal{Q}_r(V)$ for all r. Moreover, $\mathcal{Q}_r(V) \subseteq \mathrm{COP}(V)$ for all r. Indeed, take $A \in \mathcal{Q}_r(V)$ and $F \in L^2(V^{r+2})$ with $\mathcal{R}_{\mathcal{S}_{r+2}}(A \otimes \mathbf{1}^{\otimes r} - F) \geq 0$. Given a nonnegative $f \in L^2(V)$ with $\langle \mathbf{1}, f\rangle > 0$ we have

$$\begin{aligned} 0 &\leq \langle \mathcal{R}_{\mathcal{S}_{r+2}}(A \otimes \mathbf{1}^{\otimes r} - F), f^{\otimes(r+2)}\rangle \\ &= \langle A \otimes \mathbf{1}^{\otimes r}, \mathcal{R}_{\mathcal{S}_{r+2}} f^{\otimes(r+2)}\rangle - \langle F, \mathcal{R}_{\mathcal{S}_{r+2}} f^{\otimes(r+2)}\rangle \\ &= \langle A, f \otimes f\rangle \langle \mathbf{1}, f\rangle^r - \langle F, f^{\otimes(r+2)}\rangle. \end{aligned}$$

Now, since F is slice-positive,

$$\begin{aligned} \langle F, f^{\otimes(r+2)}\rangle &= \int_{V^r}\int_V\int_V F(x, y, v) f(x) f(y)\, d\omega(y) d\omega(x)\, f(v_1)\cdots f(v_r)\, d\omega(v) \\ &\geq 0, \end{aligned}$$

and we see that $\langle A, f \otimes f\rangle \geq 0$, and so A is copositive.

If $A \in \mathcal{Q}_r(V)$, that is, if $\mathcal{R}_{\mathcal{S}_{r+2}}(A \otimes \mathbf{1}^{\otimes r} - F) \geq 0$ for some slice-positive F, then as in the proof of Theorem 4.1 one shows that $\mathcal{R}_{\mathcal{S}_{r+2}}(A \otimes \mathbf{1}^{\otimes r} \otimes \mathbf{1} - F \otimes \mathbf{1}) \geq 0$, and so $A \in \mathcal{Q}_{r+1}(V)$. Hence

$$\mathcal{Q}_1(V) \subseteq \mathcal{Q}_2(V) \subseteq \cdots \subseteq \mathrm{COP}(V)$$

is a hierarchy of inner approximations of $\mathrm{COP}(V)$ stronger than the $\mathcal{C}_r(V)$ hierarchy; it was proposed in the finite setting by Peña, Vera, and Zuluaga [39] and then extended to the infinite setting by Kuryatnikova and Vera [23].

Given a locally independent graph $G = (V, E)$, set

$$\begin{aligned} \xi_r(G) = \sup\ & \langle A, J\rangle \\ & \textstyle\int_V A(x, x)\, d\omega(x) = 1, \\ & A(x, y) = 0 \quad \text{if } xy \in E, \\ & A \in \mathcal{Q}_r^*(V) \text{ is continuous and positive semidefinite.} \end{aligned}$$

This gives a hierarchy of bounds for the measurable independence number, namely

$$\xi_1(G) \geq \xi_2(G) \geq \cdots \geq \alpha_\omega(G),$$

that is at least as strong as the γ_r hierarchy of §4. In particular, the same convergence results of §5 hold for this hierarchy.

6.3. **Invariant slice-positive functions.** Let Γ be a group acting on a topological space V. For each $v \in V^k$, let $\langle v \rangle$ be an arbitrary representative of the orbit $[v]$ of v under the action of Γ.

Suppose $F \in C(V^{k+2})$ is slice-positive and Γ-invariant. Consider the function $K\colon (V^k/\Gamma) \times V^2 \to \mathbb{R}$ such that

$$K([v], x, y) = F(x, y, \langle v \rangle)$$

and for every orbit $[v]$ let $K_{[v]}(x, y) = K([v], x, y)$; note that K depends on the choice of representatives. The kernel $K_{[v]}$ is continuous and positive semidefinite for every orbit $[v]$. Moreover, since F is Γ-invariant, $K_{[v]}$ is invariant under the stabilizer subgroup $\mathrm{Stab}(\Gamma, \langle v \rangle)$ of $\langle v \rangle$.

If we equip V^k/Γ with the quotient topology, then we may ask whether K given above is continuous. This is the case as long as the function mapping each orbit $[v]$ to its representative $\langle v \rangle$ is continuous.

Conversely, say $K\colon (V^k/\Gamma) \times V^2 \to \mathbb{R}$ is a continuous function such that $K_{[v]}$ is a positive-semidefinite $\mathrm{Stab}(\Gamma, \langle v \rangle)$-invariant kernel for every orbit $[v]$. Then we may define a function $F\colon V^{k+2} \to \mathbb{R}$ by letting

$$F(x, y, v) = K([v], \sigma x, \sigma y)$$

for any $\sigma \in \Gamma$ such that $\sigma v = \langle v \rangle$. Note that F is well defined: if $\tau v = \langle v \rangle$, then $\sigma\tau^{-1}\langle v \rangle = \langle v \rangle$, and so from the invariance of $K_{[v]}$ we get $K([v], \tau x, \tau y) = K([v], \sigma x, \sigma y)$.

By construction, F is slice-positive and Γ-invariant. Indeed, slice-positivity is clear. As for invariance, given $\sigma \in \Gamma$, let $\tau \in \Gamma$ be such that $\tau\sigma v = \langle v \rangle$. Then $F(\sigma x, \sigma y, \sigma v) = K([v], \tau\sigma x, \tau\sigma y) = F(x, y, v)$, as we wanted. Moreover, if the σs can be chosen to vary continuously, then F is also continuous. More precisely, if there is a continuous function $s\colon V^k \to \Gamma$ such that $s(v)v = \langle v \rangle$ for every $v \in V^k$, then F is continuous.

Our aim is to compute something like $\xi_1(G_n)$, where $G_n = G(S^{n-1}, \{0\})$ is Witsenhausen's graph. We see now how we can use the Bachoc-Vallentin kernels (14) to get a subset of $\mathcal{Q}_1(S^{n-1})$ and therefore a superset of $\mathcal{Q}_1^*(S^{n-1})$, obtaining in this way a relaxation of $\xi_1(G_n)$.

If $Z \in \mathcal{Q}_1(S^{n-1})$, then there is a continuous slice-positive $F\colon (S^{n-1})^3 \to \mathbb{R}$ such that $\mathcal{R}_{\mathcal{S}_3}(Z \otimes \mathbf{1} - F) \geq 0$. If Z is $\mathrm{O}(n)$-invariant, then we can assume that also F is $\mathrm{O}(n)$-invariant, otherwise we simply take $\mathcal{R}_{\mathrm{O}(n)}F$ instead, which is a continuous slice-positive function.

There is only one orbit for the action of the orthogonal group on the sphere; we pick the north pole $e \in S^{n-1}$ as its representative. So the invariant function F is continuous and slice-positive if and only if there is a continuous, positive-semidefinite, and $\mathrm{Stab}(\mathrm{O}(n), e)$-invariant kernel $K\colon S^{n-1} \times S^{n-1} \to \mathbb{R}$ such that $F(x, y, z) = K(Tx, Ty)$, where T is any orthogonal matrix such that $Tz = e$. So the value of $F(x, y, z)$ depends only on $e \cdot Tx = x \cdot z$, $e \cdot Ty = y \cdot z$, and $Tx \cdot Ty = x \cdot y$.

The Bachoc-Vallentin kernels (14) are positive-semidefinite and $\mathrm{Stab}(\mathrm{O}(n), e)$-invariant. Fix an integer $d \geq 1$. Say Z is the $\mathrm{O}(n)$-invariant kernel given by

$$Z(x, y) = \sum_{k=0}^{2d} f(k) P_k^n(x \cdot y). \tag{15}$$

Let $\overline{Y}_{k,d}^n = \mathcal{R}_{\mathcal{S}_3} Y_{k,d}^n$ be the matrix obtained from $Y_{k,d}^n$ of (13) by averaging over all permutations of (u, v, t).

If there are positive-semidefinite matrices $F_k \in \mathbb{R}^{(d-k+1)\times(d-k+1)}$ for $k = 0$, $\ldots$, d such that

$$\sum_{k=0}^{2d} f(k)(1/3)(P_k^n(u) + P_k^n(v) + P_k^n(t)) - \sum_{k=0}^{d} \langle F_k, \overline{Y}_{k,d}^n(u, v, t)\rangle \geq 0 \tag{16}$$

for all $(u, v, t) \in \Delta = \{\, (x \cdot z, y \cdot z, x \cdot y) : x, y, z \in S^{n-1} \,\}$, then $Z \in \mathcal{Q}_1(S^{n-1})$.

Indeed, for x, y, $z \in S^{n-1}$ with $u = x \cdot z$, $v = y \cdot z$, and $t = x \cdot y$ we have

$$\mathcal{R}_{\mathcal{S}_3}(Z \otimes \mathbf{1})(x, y, z) = \sum_{k=0}^{2d} f(k)(1/3)(P_k^n(u) + P_k^n(v) + P_k^n(t)).$$

The function F given by

$$F(x, y, z) = \sum_{k=0}^{d} \langle F_k, Y_{k,d}^n(u, v, t)\rangle$$

is slice-positive and continuous and

$$(\mathcal{R}_{\mathcal{S}_3} F)(x, y, z) = \sum_{k=0}^{d} \langle F_k, \overline{Y}_{k,d}^n(u, v, t)\rangle.$$

Putting it all together, we see that $Z \in \mathcal{Q}_1(S^{n-1})$.

Note that the left side of (16) is a polynomial $p \in \mathbb{R}[u, v, t]$ of degree at most $2d$ that should be nonnegative on Δ. The polynomial p is also invariant under the action of $\mathcal{S}_3$ which permutes the variables. The domain Δ is also invariant under $\mathcal{S}_3$; it is a semialgebraic set: $\Delta = \{\, (u, v, t) \in \mathbb{R}^3 : g_i(u, v, t) \geq 0 \text{ for } i = 1, \ldots, 4 \,\}$, where $g(w) = 1 - w^2$ and

$$\begin{aligned} g_1 &= g(u) + g(v) + g(t), & g_2 &= g(u)g(v) + g(u)g(t) + g(v)g(t),\\ g_3 &= g(u)g(v)g(t), & g_4 &= 1 + 2uvt - u^2 - v^2 - t^2. \end{aligned}$$

(See Lemma 3.1 in Machado and Oliveira [32].) So if there are sums-of-squares polynomials q_0, $\ldots$, $q_4 \in \mathbb{R}[u, v, t]$ such that

$$p = q_0 + g_1 q_1 + g_2 q_2 + g_3 q_3 + g_4 q_4, \tag{17}$$

then p is nonnegative on Δ. Moreover, since p and the g_i are all invariant under $\mathcal{S}_3$, we may assume without loss of generality that the q_i are also invariant.

Let V_r be the matrix indexed by all monomials on u, v, and t of degree at most $\lfloor r/2 \rfloor$ such that $(V_r)(m_1, m_2) = m_1 m_2$ for any two such monomials. Note that every entry of V_r is a polynomial of degree at most r. A polynomial q of degree $2k$ is a sum of squares if and only if there is a positive-semidefinite matrix Q such that $q = \langle Q, V_{2k}\rangle$.

Using this equivalence and by restricting the degrees of the polynomials q_i appearing in (17), we can write a sufficient condition for p to be nonnegative on Δ in terms of positive-semidefinite matrices. Namely, if there are positive-semidefinite matrices F_k and Q_i such that

$$(18)\quad \sum_{k=0}^{2d} f(k)(1/3)(P_k^n(u)+P_k^n(v)+P_k^n(t)) - \sum_{k=0}^{d}\langle F_k, \overline{Y}^n_{k,d}(u,v,t)\rangle \\ = \langle Q_0, V_{2d}\rangle + \langle Q_1, g_1V_{2d-2}\rangle + \langle Q_2, g_2V_{2d-4}\rangle \\ + \langle Q_3, g_3V_{2d-6}\rangle + \langle Q_4, g_4V_{2d-3}\rangle,$$

then Z given in (15) is in $\mathcal{Q}_1(S^{n-1})$. This leads us to the definition of the following cone for every fixed d:

(19)

$$\mathcal{Q}_1^d = \{\,(f(0),\ldots,f(2d),0,\ldots)\in\mathbb{R}^{\mathbb{N}} : \text{there are positive-semidefinite matrices } F_k \text{ and } Q_i \text{ such that (18) holds}\,\}.$$

We were careful to describe the domain Δ with invariant polynomials so we could assume that all polynomials q_i are likewise invariant. This can be used to simplify (18) so we can work with block-diagonal positive-semidefinite matrices Q_i. The original idea was presented by Gatermann and Parrilo [16]; see also Machado and Oliveira [32] and de Laat and Leijenhorst [25, §4] for more recent descriptions of the method and an application to this exact situation. This use of symmetry to reduce the problem size is essential to reach high degrees.

6.4. **Witsenhausen's problem: Setup.** Our goal is to compute an upper bound for Witsenhausen's number $\alpha_n = \alpha_\omega(G_n)/\omega_n$, where $G_n = G(S^{n-1},\{0\})$ is Witsenhausen's graph. We do so by combining the cone $\mathcal{Q}_1^d$ with constraints from the Boolean quadratic polytope, which for a finite set V is defined as

$$\mathrm{BQP}(V) = \mathrm{conv}\{\, xx^{\mathsf{T}} : x\in\{0,1\}^V\,\}.$$

Such constraints were used before by DeCorte, Oliveira, and Vallentin [10]. The bound we compute is thus a hybrid between $\xi_1(G_n)$ and the bound by DeCorte, Oliveira, and Vallentin.

Given a measurable independent set $I\subseteq S^{n-1}$ of G_n, let $A = \mathcal{R}_{\mathrm{O}(n)}(\chi_I\otimes\chi_I)$. Then:

(i) A is an $\mathrm{O}(n)$-invariant continuous kernel (by Lemma 3.2);
(ii) $A(x,y)=0$ for all orthogonal x, $y\in S^{n-1}$;
(iii) A is positive semidefinite and $A\in\mathcal{Q}_r^*(S^{n-1})$ for all $r\geq 1$;
(iv) $\big(A(x,y)\big)_{x,y\in U}\in\mathrm{BQP}(U)$ for every finite $U\subseteq S^{n-1}$;
(v) $\int_{S^{n-1}} A(x,x)\,d\omega(x) = \omega(I)$ and $\langle A, J\rangle = \omega(I)^2$.

We use Schoenberg's theorem to express A in terms of Jacobi polynomials as in (10), so

$$A(x,y) = \sum_{k=0}^{\infty} a(k)P_k^n(x\cdot y)$$

for some sequence $a\geq 0$. Recall that we normalize the polynomials so $P_k^n(1)=1$; together with the addition formula (§5.2) this gives

$$(20)\qquad \int_{S^{n-1}} A(x,x)\,d\omega(x) = \omega_n\sum_{k=0}^{\infty}a(k) \qquad\text{and}\qquad \langle A,J\rangle = \omega_n^2 a(0),$$

where $\omega_n = \omega(S^{n-1})$.

Let $U \subseteq S^{n-1}$ be a finite set and let $L \in \mathbb{R}^{U \times U}$ and $\beta \in \mathbb{R}$ be such that $\langle L, X \rangle \leq \beta$ for all $X \in \mathrm{BQP}(U)$. Then defining $r \colon \mathbb{N} \to \mathbb{R}$ by

$$r(k) = \sum_{x,y \in U} L(x,y) P_k^n(x \cdot y), \tag{21}$$

we have

$$\sum_{k=0}^{\infty} a(k) r(k) \leq \beta. \tag{22}$$

We call (r, β) a $\mathrm{BQP}(S^{n-1})$-*inequality* and we call the points in U the *support points* of the inequality.

Finally, if $(f(0), \ldots, f(2d), 0, \ldots) \in \mathcal{Q}_1^d$, where $\mathcal{Q}_1^d$ is defined in (19), and if $Z(x,y) = \sum_{k=0}^{2d} f(k) P_k^n(x \cdot y)$, then $Z \in \mathcal{Q}_1(S^{n-1})$ and from (9) we get

$$\sum_{k=0}^{2d} (a(k)/h_k^n) f(k) = \omega_n^{-2} \langle A, Z \rangle \geq 0,$$

that is, $k \mapsto a(k)/h_k^n$ belongs to $(\mathcal{Q}_1^d)^*$.

Let (r_1, β_1), ..., (r_N, β_N) be any $\mathrm{BQP}(S^{n-1})$-inequalities and fix some integer $d \geq 1$. Put together, our developments lead us to the following optimization problem, whose optimal value gives an upper bound for α_n:

$$\begin{aligned}
&\sup \textstyle\sum_{k=0}^{\infty} a(k)\\
&\quad \textstyle\sum_{k=0}^{\infty} a(k) P_k^n(0) = 0,\\
&\quad \textstyle\sum_{k=0}^{\infty} a(k) r_i(k) \leq \beta_i \quad \text{for } i = 1, \ldots, N,\\
&\quad \begin{pmatrix} 1 & \omega_n \sum_{k=0}^{\infty} a(k) \\ \omega_n \sum_{k=0}^{\infty} a(k) & \omega_n^2 a(0) \end{pmatrix} \text{ is positive semidefinite,}\\
&\quad a \geq 0 \text{ and } k \mapsto a(k)/h_k^n \in (\mathcal{Q}_1^d)^*.
\end{aligned} \tag{23}$$

The 2×2 matrix comes from (20) and (v) and is used to normalize the problem. Note moreover that the objective function is divided by ω_n, ensuring that we get a bound for $\alpha_n = \alpha_\omega(G_n)/\omega_n$. Finally, our problem has infinitely many variables a, but only the first $2d+1$ of them appear in the cone constraint with $(\mathcal{Q}_1^d)^*$. Contrast this with the situation of the 3-point bound from §6.1.

The dual of this problem is:

$$\begin{aligned}
&\inf z_{11} + \textstyle\sum_{i=1}^{N} y_i \beta_i\\
&\quad \lambda + \textstyle\sum_{i=1}^{\infty} y_i r_i(0) - \omega_n z_{12} - \omega_n^2 z_{22} - f(0) \geq 1,\\
&\quad \lambda P_k^n(0) + \textstyle\sum_{i=1}^{N} y_i r_i(k) - \omega_n z_{12} - f(k) \geq 1 \quad \text{for all } k = 1, \ldots, 2d,\\
&\quad \lambda P_k^n(0) + \textstyle\sum_{i=1}^{N} y_i r_i(k) - \omega_n z_{12} \geq 1 \quad \text{for all } k \geq 2d+1,\\
&\quad \begin{pmatrix} z_{11} & z_{12}/2 \\ z_{12}/2 & z_{22} \end{pmatrix} \text{ is positive semidefinite,}\\
&\quad y \geq 0 \text{ and } k \mapsto h_k^n f(k) \in \mathcal{Q}_1^d.
\end{aligned} \tag{24}$$

It is easy to show that the objective value of any feasible solution of (24) is greater or equal than the objective value of any feasible solution of (23). So any feasible solution of the dual gives an upper bound for α_n. Note moreover that the constraint

"$k \mapsto h_k^n f(k) \in \mathcal{Q}_1^d$" is expressed in terms of (18), namely we require there to be positive-semidefinite matrices F_k and Q_i such that

$$\begin{aligned}(25)\quad \sum_{k=0}^{2d} f(k)(h_k^n/3)(P_k^n(u) + P_k^n(v) + P_k^n(t)) - \sum_{k=0}^{d} \langle F_k, \overline{Y}_{k,d}^n(u,v,t)\rangle \\ - \langle Q_0, V_{2d}\rangle - \langle Q_1, g_1 V_{2d-2}\rangle - \langle Q_2, g_2 V_{2d-4}\rangle \\ - \langle Q_3, g_3 V_{2d-6}\rangle - \langle Q_4, g_4 V_{2d-3}\rangle = 0.\end{aligned}$$

So (24) is a semidefinite programming problem with finitely many variables but infinitely many constraints.

6.5. **Witsenhausen's problem: Solution and verification.** To solve (24) we use the `ClusteredLowRankSolver` of de Laat and Leijenhorst [25]; the input for the solver is generated by a Julia program. The program and all data files used are available in the Harvard Dataverse repository [35].

To find good BQP(S^{n-1})-inequalities, we use a separation heuristic described by DeCorte, Oliveira, and Vallentin [10]. The inequalities used are also included in the repository and need not be recomputed.

Table 2 contains a detailed account of all the bounds computed from (24). Solving the problem for $d = 14$ and 18 takes time and memory, and so files with the corresponding solutions are also available in the repository.

Since (24) has infinitely many linear constraints, to solve it we select some finite set $S \subseteq \{0, 1, \ldots\}$ and consider only the constraints for $k \in S$. After a solution is found it has to be verified, that is, we need to check that all constraints are indeed satisfied.

Let (λ, y, z, f, F, Q) be a candidate solution to (24), where F and Q are as in (25), returned by the solver. The first step is to certify ourselves that f, F, and Q indeed satisfy (25).

This is certainly not true: the solver works with floating-point arithmetic, and so (25) will not hold. Rather, the left side of (25) will be a polynomial with coefficients close to 0. Since the `ClusteredLowRankSolver` uses high-precision floating-point arithmetic, the coefficients will be quite small; let η be the largest absolute value of any such coefficient.

It is always possible to perturb the matrices Q_i in order to satisfy the constraint; the order of the perturbation depends on η. We want to do so and keep the Q_i positive semidefinite, and as long as the minimum eigenvalues of the matrices Q_i are large enough compared to η, this is always possible. The solutions stored in the repository have large minimum eigenvalues, several orders of magnitude larger than η, and so this perturbation of the Q_i can always be carried out. Note that we do not have to actually change the Q_i; it suffices to know that such a perturbation is possible, since then we know we can get a feasible solution if we want to. This procedure was used before by de Laat, Oliveira, and Vallentin [28].

Checking that the linear constraints for all $k \geq 0$ are satisfied is more difficult; we use the approach outlined in DeCorte, Oliveira, and Vallentin [10].

The idea is as follows. Let $\operatorname{lhs}(k)$ be the left side of the kth linear constraint in (24) and write $\operatorname{lhs}(\infty) = \lim_{k\to\infty} \operatorname{lhs}(k)$; we will see that this limit exists. Then we follow the steps:

TABLE 2. Lower and upper bounds for Witsenhausen's parameter α_n. The lower bound is given by the double-cap conjecture. The simple upper bound was given by Witsenhausen [46] and is just $1/n$; the best previous upper bounds are by DeCorte, Oliveira, and Vallentin [10]. The table gives upper bounds obtained from solving (24) with and without BQP(S^{n-1})-inequalities and for several values of d.

		Previous upper bound			*New upper bound*	
n	*Lower bound*	*Simple*	*Best*	d	*No BQP*	*With BQP*
3	0.2928...	0.3333...	0.30153	6	0.316925	0.300708
				10	0.309298	0.298998
				14	0.305627	0.298341
				18	0.303294	0.297742
4	0.1816...	0.25	0.21676	6	0.223633	0.207617
				10	0.211825	0.199402
				14	0.205479	0.196162
				18	0.201445	0.194297
5	0.1161...	0.2	0.16765	6	0.167357	0.151541
				10	0.153819	0.141539
				14	0.146612	0.137142
				18	0.142349	0.134588
6	0.0755...	0.1666...	0.13382	6	0.130829	0.116599
				10	0.116509	0.105200
				14	0.109989	0.100374
				18	0.106727	0.098095
7	0.0498...	0.1428...	0.11739	6	0.106059	0.093031
				10	0.091477	0.081221
				14	0.086656	0.077278
				18	0.084787	0.075751
8	0.0331...	0.125	0.09981	6	0.088750	0.076801
				10	0.074309	0.064919
				14	0.071676	0.063287
				18	0.070607	0.061178

(i) As long as $z_{22} > 0$, we can change z_{11} and z_{12} to get $\text{lhs}(\infty) \geq 1 + \eta$ for some $\eta > 0$. The more we change z_{12}, the more we have to change z_{11}, and the worse the bound gets.

(ii) Next, for some $\epsilon < \eta$ we find a k_0 such that $|\text{lhs}(k) - \text{lhs}(\infty)| \leq \epsilon$ for all $k \geq k_0$. Then $\text{lhs}(k) \geq \text{lhs}(\infty) - \epsilon \geq 1 + \eta - \epsilon > 1$, and so all constraints are satisfied for $k \geq k_0$.

(iii) Finally, we check the constraints for $k = 0, \ldots, k_0$, and by changing z again we can make all these constraints satisfied.

If we choose our initial sample S well, then all constraints will be almost satisfied, and we will not have to change z too much in order to get a feasible solution. This is the procedure implemented by the `fix_linear_constraints` function in the Julia program in the repository [35].

Let us see the details of the procedure. The asymptotic formula for the Jacobi polynomials [43, Theorem 8.21.8] implies that $P_k^n(t) \to 0$ as $k \to \infty$ for all $t \in (-1, 1)$. We make sure that all the BQP(S^{n-1})-inequalities (22) we use have support points U such that distinct $x, y \in U$ have inner product $x \cdot y$ bounded away from ± 1. So if r is given as in (21), then

$$r(\infty) = \lim_{k \to \infty} r(k) = \operatorname{tr} L$$

and

$$|r(k) - r(\infty)| \leq \sum_{\substack{x,y \in U \\ x \neq y}} |L(x,y)||P_k^n(x \cdot y)|. \tag{26}$$

We also have

$$\text{lhs}(\infty) = \sum_{i=1}^{N} y_i r_i(\infty) - \omega_n z_{12}.$$

Given $\epsilon > 0$ we want to get k_0 as in (ii). Note that

$$|\,\text{lhs}(k) - \text{lhs}(\infty)| \leq |\lambda||P_k^n(0)| + \sum_{i=1}^{N} y_i |r_i(k) - r_i(\infty)|.$$

Fix k_0. Using (26), we see that to find an upper bound for the left side above for all $k \geq k_0$, it suffices to find for all $k \geq k_0$ an upper bound on $|P_k^n(t)|$ for $t = 0$ and all other $t \in (-1, 1)$ that occur as inner products between distinct support points of the $\text{BQP}(S^{n-1})$-inequalities we use.

To do so rigorously, we use an integral representation for the ultraspherical polynomials due to Gegenbauer (take $\lambda = (n-2)/2$ in Theorem 6.7.4 from Andrews, Askey, and Roy [1]):

$$P_k^n(\cos\theta) = R(n)^{-1} \int_0^{\pi} F(\phi)^k \sin^{n-3}\phi \, d\phi,$$

where

$$F(\phi) = \cos\theta + i \sin\theta \cos\phi \qquad \text{and} \qquad R(n) = \int_0^{\pi} \sin^{n-3}\phi \, d\phi.$$

Note that $|F(\phi)|^2 = \cos^2\theta + \sin^2\theta \cos^2\phi$, so

$$|P_k^n(\cos\theta)| \leq R(n)^{-1} \int_0^{\pi} (\cos^2\theta + \sin^2\theta\cos^2\phi)^{k/2} \sin^{n-3}\phi \, d\phi.$$

The right side is decreasing in k, and we can estimate the integrals rigorously using interval arithmetic.

For (iii) we need to compute $\text{lhs}(k)$ for all $k \leq k_0$. We would like to do this rigorously, using for instance interval arithmetic. The most time-consuming step here is to compute the r_i functions. In practice, this step involves evaluating the polynomials P_k^n for values of k that can exceed 100,000.

The Jacobi polynomials P_k^n are given by a simple recurrence, namely

$$P_k^n(u) = a_k^n(u) P_{k-1}^n(u) + b_k^n P_{k-2}^n(u)$$

for $k \geq 2$ with $P_1^n(u) = u$ and $P_0^n(u) = 1$, where

$$a_k^n(u) = \frac{2k + 2\alpha - 1}{k + 2\alpha} u \qquad \text{and} \qquad b_k = -\frac{k-1}{k+2\alpha}$$

with $\alpha = (n-3)/2$. This recurrence comes from formula (4.5.1) in Szegö [43], adapted to our normalization of $P_k^n(1) = 1$.

The recurrence is very stable: even using double-precision floating-point arithmetic it is possible to accurately evaluate the polynomial for very high degrees for points in $[-1, 1]$. If we use this recurrence with interval arithmetic though, the error estimation quickly gets out of hand: if $\lim_{k\to\infty} a_k(u) > 1$, then the error bound grows exponentially.

Using interval arithmetic then requires very high precision and is very slow, though not prohibitively so. In any case, we can trust floating-point computations. Using the recurrence amounts to solving by backward substitution a linear system with a triangular matrix whose entries are the numbers $a_k(u)$, b_k, and 1, and this matrix is well conditioned, so the error we make in solving the system is very small. The error was analyzed for instance by Barrio [4]. The Julia program that performs the verification uses high-precision floating-point arithmetic.

Appendix A. Pólya's theorem and a proof of Theorem 5.2

Let V be a finite set. A *symmetric k-tensor* is a function $T\colon V^k \to \mathbb{R}$ invariant under the action of $\mathcal{S}_k$, that is, invariant under permutations of the variables. Symmetric tensors correspond to homogeneous polynomials: if $[n] = \{1, \ldots, n\}$ and if $T\colon [n]^k \to \mathbb{R}$ is a symmetric k-tensor, then the polynomial

$$p(w) = \langle T, w^{\otimes k}\rangle, \tag{27}$$

where $w \in \mathbb{R}^n$ and where $\langle \cdot, \cdot\rangle$ is the Euclidean inner product in the space of symmetric k-tensors, is an n-variable homogeneous polynomial of degree k. Conversely, if p is an n-variable homogeneous polynomial of degree k, then there is a unique symmetric k-tensor T such that (27) holds.

Let $p \in \mathbb{R}[w]$, where $w = (w_1, \ldots, w_n)$, be a homogeneous polynomial. If for some integer $r \geq 0$ all the coefficients of the polynomial

$$(w_1 + \cdots + w_n)^r p(w) = (\mathbf{1}^{\mathsf T} w)^r p(w) \tag{28}$$

are nonnegative, then $p(w) \geq 0$ for all $w \geq 0$. Pólya's theorem gives a converse: if $p(w) > 0$ for all nonzero $w \geq 0$, then there is r such that all coefficients of (28) are nonnegative.

Let V be a finite set and say $A\colon V^2 \to \mathbb{R}$ is a matrix in the algebraic interior of the copositive cone. This means in particular that the homogeneous polynomial $p(w) = w^{\mathsf T} A w$ is positive for all nonzero $w \geq 0$. By Pólya's theorem, there is r such that all coefficients of the polynomial $(\mathbf{1}^{\mathsf T} w)^r (w^{\mathsf T} A w)$ are nonnegative. Now note that

$$(\mathbf{1}^{\mathsf T} w)^r (w^{\mathsf T} A w) = \langle A \otimes \mathbf{1}^{\otimes r}, w^{\otimes (r+2)}\rangle = \langle \mathcal{R}_{\mathcal{S}_{r+2}}(A \otimes \mathbf{1}^{\otimes r}), w^{\otimes (r+2)}\rangle,$$

so the coefficients of our polynomial are given by the symmetric $(r + 2)$-tensor $\mathcal{R}_{\mathcal{S}_{r+2}}(A \otimes \mathbf{1}^{\otimes r})$, which is therefore nonnegative. It follows that $A \in \mathcal{C}_r(V)$, proving Theorem 5.2 for finite V.

For a function $T\colon V^k \to \mathbb{R}$ and any set $U \subseteq V$, we denote by $T[U]$ the restriction of T to U^k. To prove Theorem 5.2 for infinite V, we use the following infinite-dimensional version of Pólya's theorem, which is interesting in itself.

Theorem A.1. *Let V be a compact Hausdorff space equipped with a Radon measure ω that is positive on open sets. Given a symmetric $T \in C(V^k)$, let $M = \max\{\, |T(v)| : v \in V^k \,\}$ and $\lambda = \inf\{\, \langle T, f^{\otimes k}\rangle : f \in L^2(V), f \geq 0,$ and $\langle \mathbf{1}, f\rangle = 1 \,\}$. If $\lambda > 0$, then for every $r > k(k-1)M/(2\lambda) - k$ we have $\mathcal{R}_{\mathcal{S}_{r+k}}(T \otimes \mathbf{1}^{\otimes r}) \geq 0$.*

Proof. Let $U \subseteq V$ be any finite nonempty set. Given a function $w\colon U \to \mathbb{R}_+$ with $\mathbf{1}^{\mathsf{T}} w = 1$, fix $\epsilon > 0$ and let $P(x)$ for $x \in U$ be disjoint sets of positive measure such that $x \in P(x)$ for all $x \in U$ and such that T varies at most ϵ in any set of the form $P(x_1) \times \cdots \times P(x_k)$ for $x_1, \ldots, x_k \in U$. Such sets always exist: see for instance Theorem 4.4 in DeCorte, Oliveira, and Vallentin [10], where this is shown for $k = 2$, though the proof immediately generalizes for $k \geq 3$; the proof of Lemma 4.2 uses the same idea.

Set $f = \sum_{x \in U} w(x)\omega(P(x))^{-1}\chi_{P(x)}$; note that $f \geq 0$ and $\langle \mathbf{1}, f \rangle = 1$. For x_1, $\ldots$, $x_k \in U$, write $P(x_1, \ldots, x_k) = P(x_1) \times \cdots \times P(x_k)$. We have

$$\begin{aligned}
\langle T, f^{\otimes k} \rangle &= \sum_{x_1,\ldots,x_k \in U} \left(\prod_{i=1}^{k} w(x_i)\omega(P(x_i))^{-1} \right) \int_{P(x_1,\ldots,x_k)} T(y_1, \ldots, y_k)\, d\omega(y) \\
&\leq \sum_{x_1,\ldots,x_k \in U} \left(\prod_{i=1}^{k} w(x_i)\omega(P(x_i))^{-1} \right) \int_{P(x_1,\ldots,x_k)} T(x_1, \ldots, x_k) + \epsilon\, d\omega(y) \\
&= \langle T[U], w^{\otimes k} \rangle + \epsilon,
\end{aligned}$$

hence by taking $\epsilon \to 0$ we see that $\langle T[U], w^{\otimes k} \rangle \geq \langle T, f^{\otimes k} \rangle \geq \lambda$. This holds for every U and w.

Fix a finite nonempty set $U \subseteq V$. We now apply a theorem of Powers and Reznick [40, Theorem 1] to the homogeneous polynomial $p(w) = \langle T[U], w^{\otimes k} \rangle$ of degree k. Since $p(w) \geq \lambda > 0$ for all $w \geq 0$ with $\mathbf{1}^{\mathsf{T}} w = 1$, the theorem of Powers and Reznick says that all the coefficients of the polynomial $(\mathbf{1}^{\mathsf{T}} w)^r p(w)$ are positive for all $r > k(k-1)M'/(2\lambda) - k$, where $M' = \max\{\, |T(v)| : v \in U^k \,\}$. This means that all the entries of $\mathcal{R}_{\mathcal{S}_{r+k}}(T[U] \otimes \mathbf{1}^{\otimes r})$ are nonnegative. Since U is any finite subset and since $M \geq M'$, we are done. ☐

Proof of Theorem 5.2. If A is in the algebraic interior of $\mathrm{COP_c}(V)$, then there is $\lambda > 0$ such that $A - \lambda J$ is copositive. So if $f \in L^2(V)$ is such that $f \geq 0$ and $\langle \mathbf{1}, f \rangle = 1$, then

$$0 \leq \langle A - \lambda J, f \otimes f \rangle = \langle A, f \otimes f \rangle - \lambda,$$

and so $\langle A, f \otimes f \rangle \geq \lambda$. Using Theorem A.1 we see that there is r such that $\mathcal{R}_{\mathcal{S}_{r+2}}(A \otimes \mathbf{1}^{\otimes r}) \geq 0$, and so $A \in \mathcal{C}_r(V)$, proving the theorem. ☐

Acknowledgments

We would like to thank David de Laat, Nando Leijenhorst, Fabrício Caluza Machado, and Willem de Muinck Keizer for fruitful discussions. David de Laat and Nando Leijenhorst also gave some much needed technical support regarding the `ClusteredLowRankSolver`. The optimization problems were solved in a computational cluster at TU Delft maintained by Joffrey Wallaart.

References

[1] George E. Andrews, Richard Askey, and Ranjan Roy, *Special functions*, Encyclopedia of Mathematics and its Applications, vol. 71, Cambridge University Press, Cambridge, 1999, DOI 10.1017/CBO9781107325937. MR1688958

[2] Christine Bachoc, Gabriele Nebe, Fernando Mário de Oliveira Filho, and Frank Vallentin, *Lower bounds for measurable chromatic numbers*, Geom. Funct. Anal. **19** (2009), no. 3, 645–661, DOI 10.1007/s00039-009-0013-7. MR2563765

[3] Christine Bachoc and Frank Vallentin, *New upper bounds for kissing numbers from semidefinite programming*, J. Amer. Math. Soc. **21** (2008), no. 3, 909–924, DOI 10.1090/S0894-0347-07-00589-9. MR2393433

[4] Roberto Barrio, *Rounding error bounds for the Clenshaw and Forsythe algorithms for the evaluation of orthogonal polynomial series*, J. Comput. Appl. Math. **138** (2002), no. 2, 185–204, DOI 10.1016/S0377-0427(01)00382-X. MR1876230

[5] S. Bochner, *Hilbert distances and positive definite functions*, Ann. of Math. (2) **42** (1941), 647–656, DOI 10.2307/1969252. MR5782

[6] Henry Cohn and Noam Elkies, *New upper bounds on sphere packings. I*, Ann. of Math. (2) **157** (2003), no. 2, 689–714, DOI 10.4007/annals.2003.157.689. MR1973059

[7] Henry Cohn, Abhinav Kumar, Stephen D. Miller, Danylo Radchenko, and Maryna Viazovska, *The sphere packing problem in dimension 24*, Ann. of Math. (2) **185** (2017), no. 3, 1017–1033, DOI 10.4007/annals.2017.185.3.8. MR3664817

[8] H. Cohn, D. de Laat, and A. Salmon, *Three-point bounds for sphere packing*, `arXiv:2206.15373`, 2022, 37 pp.

[9] H. Cohn and A. Salmon, *Sphere packing bounds via rescaling*, `arXiv:2108.10936`, 2021, 43 pp.

[10] Evan DeCorte, Fernando Mário de Oliveira Filho, and Frank Vallentin, *Complete positivity and distance-avoiding sets*, Math. Program. **191** (2022), no. 2, Ser. A, 487–558, DOI 10.1007/s10107-020-01562-6. MR4384753

[11] P. Delsarte, J. M. Goethals, and J. J. Seidel, *Spherical codes and designs*, Geometriae Dedicata **6** (1977), no. 3, 363–388, DOI 10.1007/bf03187604. MR485471

[12] Cristian Dobre, Mirjam Dür, Leonhard Frerick, and Frank Vallentin, *A copositive formulation for the stability number of infinite graphs*, Math. Program. **160** (2016), no. 1-2, Ser. A, 65–83, DOI 10.1007/s10107-015-0974-2. MR3555382

[13] Gerald B. Folland, *A course in abstract harmonic analysis*, Studies in Advanced Mathematics, CRC Press, Boca Raton, FL, 1995. MR1397028

[14] Gerald B. Folland, *Real analysis*, 2nd ed., Pure and Applied Mathematics (New York), John Wiley & Sons, Inc., New York, 1999. Modern techniques and their applications; A Wiley-Interscience Publication. MR1681462

[15] P. Frankl and R. M. Wilson, *Intersection theorems with geometric consequences*, Combinatorica **1** (1981), no. 4, 357–368, DOI 10.1007/BF02579457. MR647986

[16] Karin Gatermann and Pablo A. Parrilo, *Symmetry groups, semidefinite programs, and sums of squares*, J. Pure Appl. Algebra **192** (2004), no. 1-3, 95–128, DOI 10.1016/j.jpaa.2003.12.011. MR2067190

[17] Martin Grötschel, László Lovász, and Alexander Schrijver, *Geometric algorithms and combinatorial optimization*, Algorithms and Combinatorics: Study and Research Texts, vol. 2, Springer-Verlag, Berlin, 1988, DOI 10.1007/978-3-642-97881-4. MR936633

[18] Nebojša Gvozdenović, Monique Laurent, and Frank Vallentin, *Block-diagonal semidefinite programming hierarchies for 0/1 programming*, Oper. Res. Lett. **37** (2009), no. 1, 27–31, DOI 10.1016/j.orl.2008.10.003. MR2488078

[19] David Handel, *Some homotopy properties of spaces of finite subsets of topological spaces*, Houston J. Math. **26** (2000), no. 4, 747–764. MR1823966

[20] Gil Kalai, *Some old and new problems in combinatorial geometry I: around Borsuk's problem*, Surveys in combinatorics 2015, London Math. Soc. Lecture Note Ser., vol. 424, Cambridge Univ. Press, Cambridge, 2015, pp. 147–174. MR3497269

[21] E. de Klerk and D. V. Pasechnik, *A linear programming reformulation of the standard quadratic optimization problem*, J. Global Optim. **37** (2007), no. 1, 75–84, DOI 10.1007/s10898-006-9037-9. MR2284559

[22] E. de Klerk and D. V. Pasechnik, *Approximation of the stability number of a graph via copositive programming*, SIAM J. Optim. **12** (2002), no. 4, 875–892, DOI 10.1137/S1052623401383248. MR1922500

[23] O. Kuryatnikova, *The many faces of positivity to approximate structured optimization problems*, Ph.D. Thesis, Tilburg University, 2019.

[24] David de Laat, *Moment methods in energy minimization: new bounds for Riesz minimal energy problems*, Trans. Amer. Math. Soc. **373** (2020), no. 2, 1407–1453, DOI 10.1090/tran/7976. MR4068268

[25] Nando Leijenhorst and David de Laat, *Solving clustered low-rank semidefinite programs arising from polynomial optimization*, Math. Program. Comput. **16** (2024), no. 3, 503–534, DOI 10.1007/s12532-024-00264-w. MR4800956

[26] David de Laat, Fabrício Caluza Machado, Fernando Mário de Oliveira Filho, and Frank Vallentin, *k-point semidefinite programming bounds for equiangular lines*, Math. Program. **194** (2022), no. 1-2, Ser. A, 533–567, DOI 10.1007/s10107-021-01638-x. MR4445463

[27] D. de Laat, W. de Muinck Keizer, and F. C. Machado, *The Lasserre hierarchy for equiangular lines with a fixed angle*, `arXiv:2211.16471`, 2022, 25 pp.

[28] David de Laat, Fernando Mário de Oliveira Filho, and Frank Vallentin, *Upper bounds for packings of spheres of several radii*, Forum Math. Sigma **2** (2014), Paper No. e23, 42, DOI 10.1017/fms.2014.24. MR3264261

[29] David de Laat and Frank Vallentin, *A semidefinite programming hierarchy for packing problems in discrete geometry*, Math. Program. **151** (2015), no. 2, Ser. B, 529–553, DOI 10.1007/s10107-014-0843-4. MR3348162

[30] Jean B. Lasserre, *Global optimization with polynomials and the problem of moments*, SIAM J. Optim. **11** (2000/01), no. 3, 796–817, DOI 10.1137/S1052623400366802. MR1814045

[31] László Lovász, *On the Shannon capacity of a graph*, IEEE Trans. Inform. Theory **25** (1979), no. 1, 1–7, DOI 10.1109/TIT.1979.1055985. MR514926

[32] Fabrício Caluza Machado and Fernando Mário de Oliveira Filho, *Improving the semidefinite programming bound for the kissing number by exploiting polynomial symmetry*, Exp. Math. **27** (2018), no. 3, 362–369, DOI 10.1080/10586458.2017.1286273. MR3857670

[33] A. Montina, *Communication cost of classically simulating a quantum channel with subsequent rank-1 projective measurement*, Phys. Rev. A **84** (2011), 4 pp.

[34] T. S. Motzkin and E. G. Straus, *Maxima for graphs and a new proof of a theorem of Turán*, Canadian J. Math. **17** (1965), 533–540, DOI 10.4153/CJM-1965-053-6. MR175813

[35] F. M. de Oliveira Filho, *Data files and code related to the upper bound for Witsenhausen's parameter*, Harvard Dataverse, 2023, V1, DOI 10.7910/DVN/TWI1SR.

[36] F. M. de Oliveira Filho, *New bounds for geometric packing and coloring via harmonic analysis and optimization*, Ph.D. Thesis, University of Amsterdam, 2009.

[37] Fernando Mário de Oliveira Filho and Frank Vallentin, *A quantitative version of Steinhaus' theorem for compact, connected, rank-one symmetric spaces*, Geom. Dedicata **167** (2013), 295–307, DOI 10.1007/s10711-012-9814-1. MR3128781

[38] Fernando Mário de Oliveira Filho and Frank Vallentin, *Fourier analysis, linear programming, and densities of distance avoiding sets in* $\mathbb{R}^n$, J. Eur. Math. Soc. (JEMS) **12** (2010), no. 6, 1417–1428, DOI 10.4171/JEMS/236. MR2734347

[39] Javier Peña, Juan Vera, and Luis F. Zuluaga, *Computing the stability number of a graph via linear and semidefinite programming*, SIAM J. Optim. **18** (2007), no. 1, 87–105, DOI 10.1137/05064401X. MR2299675

[40] Victoria Powers and Bruce Reznick, *A new bound for Pólya's theorem with applications to polynomials positive on polyhedra*, J. Pure Appl. Algebra **164** (2001), no. 1-2, 221–229, DOI 10.1016/S0022-4049(00)00155-9. Effective methods in algebraic geometry (Bath, 2000). MR1854339

[41] I. J. Schoenberg, *Positive definite functions on spheres*, Duke Math. J. **9** (1942), 96–108. MR5922

[42] Barry Simon, *Convexity*, Cambridge Tracts in Mathematics, vol. 187, Cambridge University Press, Cambridge, 2011. An analytic viewpoint, DOI 10.1017/CBO9780511910135. MR2814377

[43] Gábor Szegő, *Orthogonal polynomials*, 4th ed., American Mathematical Society Colloquium Publications, Vol. XXIII, American Mathematical Society, Providence, RI, 1975. MR372517

[44] Maryna S. Viazovska, *The sphere packing problem in dimension 8*, Ann. of Math. (2) **185** (2017), no. 3, 991–1015, DOI 10.4007/annals.2017.185.3.7. MR3664816

[45] Hsien-Chung Wang, *Two-point homogeneous spaces*, Ann. of Math. (2) **55** (1952), 177–191, DOI 10.2307/1969427. MR47345

[46] H. S. Witsenhausen, *Spherical sets without orthogonal point pairs*, Amer. Math. Monthly **81** (1974), 1101–1102, DOI 10.2307/2319047. MR358578

Delft Institute of Applied Mathematics, Delft University of Technology, Mekelweg 4, 2628 CD Delft, The Netherlands
Email address: B.Bekker@tudelft.nl

Department of Econometrics, Erasmus School of Economics, Burg. Oudlaan 50, 3062 PA Rotterdam, The Netherlands
Email address: kuryatnikova@ese.eur.nl

Delft Institute of Applied Mathematics, Delft University of Technology, Mekelweg 4, 2628 CD Delft, The Netherlands
Email address: F.M.deOliveiraFilho@tudelft.nl

Department of Econometrics and Operations Research, Tilburg University, Tilburg, The Netherlands
Email address: J.C.VeraLizcano@tilburguniversity.edu