Random homomorphisms into the orthogonality graph

Dávid Kunszenti-Kovács*, László Lovász† and Balázs Szegedy‡
Alfréd Rényi Institute of Mathematics
Budapest, Hungary

May 11, 2021

Contents

*Research supported by ERC Consolidator Grant 648017
†Research supported by ERC Synergy Grant No. 810115.
‡Research was partially supported by the NKFIH "Élvonal" KKP 133921 grant.

Abstract

Subgraph densities have been defined, and served as basic tools, both in the case of graphons (limits of dense graph sequences) and graphings (limits of bounded-degree graph sequences). While limit objects have been described for the "middle ranges", the notion of subgraph densities in these limit objects remains elusive. We define subgraph densities in the orthogonality graphs on the unit spheres in dimension d, under appropriate sparsity condition on the subgraphs. These orthogonality graphs exhibit the main difficulties of defining subgraphs the "middle" range, and so we expect their study to serve as a key example to defining subgraph densities in more general Markov spaces.

The problem can also be formulated as defining and computing random orthogonal representations of graphs. Orthogonal representations have played a role in information theory, optimization, rigidity theory and quantum physics, so to study random ones may be of interest from the point of view of these applications as well.

1 Introduction

Let H_d denote the orthogonality graph on S^{d-1}, i.e., the infinite graph whose node set is the unit sphere S^{d-1}, and two nodes are adjacent if they are orthogonal (as vectors in $\mathbb{R}^d$). For a finite graph G, we call a homomorphism of G into H an *ortho-homomorphism* of G (in dimension d).

Our motivation for studying ortho-homomorphisms comes from graph limit theory. This theory is rather well worked out for dense graphs on one end of scale (where the limit objects are graphons), and bounded degree graphs on the other (where the limit objects are graphings). In spite of several efforts to extend the theory to the intermediate cases, no complete theory has been developed.

One basic question is: what structures can serve as limit objects for "convergent" graph sequences? Here at least we seem to have a common ground: symmetric probability measures on the unit square (or on any other standard probability space; these measures are essentially equivalent to time-reversible Markov chains with a stationary distribution). These structures, which we call *Markov spaces*, capture most special cases of interest, including limit objects for L_p-convergence [3], shape convergence [9] and action convergence [2].

However, all these limit notions are defined through a global (right) convergence. To characterize them by local (left) convergence, we need to define the density of subgraphs in Markov spaces. At this time, we have a definition beyond the the two extreme cases in rather special cases only.

Our main goal in this paper is to define subgraph densities in the orthogonality graphs H_d (which have a natural Markov space associated with them). These spaces exhibit the main difficulties of the "middle" range, and so we expect their study to serve as a key example to defining subgraph densities in more general Markov spaces.

To justify this special choice, let us describe a somewhat unexpected further connection. An ortho-homomorphism of G in dimension d is the same thing as an orthonormal representation of the complementary graph $\overline{G}$ (see [12]). Such representations have played a role in information theory [10], graph algorithms [7, 8], rigidity of frameworks [1], and quantum physics [4]. Our results in this paper could be thought of as establishing further connections with probability and measure theory.

A related question is to define a *random* homomorphism of G into H_d. The notion of a random edge (the uniform distribution on orthogonal pairs of vectors) is trivial, but for more complicated graphs, it is not obvious what "random" should mean. Ortho-homomorphisms from a given graph form a real algebraic variety $\mathrm{Hom}(G, d)$, which can have a very complicated topology; but ortho-homomorphisms in general position (see below) form a smooth semialgebraic variety $\Sigma_{G,d}$. We could consider the surface measure inherited from the ambient space $(\mathbb{R}^d)^V$; however, this does not seem to have really useful properties. Natural conditions to impose are invariance under orthogonal transformations of $\mathbb{R}^d$ and the *Markov property* (see Section 2.3).

The example of the 4-cycle in dimension 3 should be a warning. Obviously, for every homomorphism $C_4 \to H_3$, one pair of nonadjacent nodes will be mapped onto parallel vectors (the other pair can form an arbitrary angle). But which one? The variety $\mathrm{Hom}(G, d)$ splits into two, and $\Sigma_{C_4,3} = \emptyset$.

In this paper we show that for several classes of graphs satisfying appropriate sparsity conditions, a measure on their ortho-homomorphisms in a given dimension d can be defined, with good properties. The measure we define is always a Radon measure, but finiteness is not guaranteed. Indeed, we'll give examples where this measure is finite, and so it can be scaled to a probability measure (defining a "random ortho-homomorphism"); unfortunately, we also have examples where the measure is infinite. The combinatorial significance of this finiteness (depending on G and d) remains an interesting unsolved problem. When this measure is finite, then its value on the set of all ortho-homomorphisms appears to be good substitute for the homomorphism density.

We describe three methods for defining subgraph densities in H_d.

Sequential mapping. One of our constructions works for graphs not containing a complete bipartite graph $K_{a,b}$ with $a + b > d$. This condition is equivalent to saying that $\overline{G}$ is $(n-d)$-connected. We'll call such graphs *d-sparse*. It implies, in particular, that every node has degree at most $d - 1$. We say that a mapping $x : V \to \mathbb{R}^d$ is in *general position*, if any d elements of V are mapped onto linearly independent vectors. The following fact was proved in [13] (Theorem 2.1).

Proposition 1 *A graph G has an ortho-homomorphism in $\mathbb{R}_d$ in general position if and only if it is d-sparse.*

The main tool in the proof of Proposition 1 was the following. Let us order the nodes of G in some way, and choose the images of the nodes one-by-one. At every step, the new node is restricted to unit vectors orthogonal to those neighbors that are already mapped. By the degree condition, the available vectors form a nonempty sphere of some dimension, and we choose a next vector on this sphere randomly and uniformly. Repeating this for all nodes, we get an ortho-homomorphism, which we call a *random sequential ortho-homomorphism* of G. The fact that this ortho-homomorphism is in general position almost surely is the main result in [13].

The distribution of the random sequential ortho-homomorphism may depend on the ordering of the nodes. If G is a tree, then we get the same distribution for every search order $(1,\dots,n)$ of the nodes, but not for other orders. However, we can define a density function for which the modified distribution will be independent of the ordering. One of our main results can be stated as follows:

Theorem 2 *For every simple d-sparse graph G, there exists a nonzero Radon measure on ortho-homomorphisms in dimension d with a Markovian conditioning.*

The measure of all homomorphisms is a good generalization of the notion of homomorphism density, a basic tool in the theory of dense graph limits. The Markov property is usually defined for probability measures, and we cannot always normalize our measure on ortho-homomorphisms to a probability measure. We'll describe the formal definition later.

Spectral methods. Our other construction is based on functional analysis. The orthogonality graph H_d defines a compact linear operator $\mathbf{A}_d :\ L^2(S^{d-1},\pi) \to L^2(S^{d-1},\pi)$, where π is the uniform probability measure on S^{d-1}, and $(\mathbf{A}_d f)(x)$ is the average of f on the $(d-2)$-dimensional sphere orthogonal to x. Taking the k-th power of this operator corresponds to subdividing each edge of G by $k-1$ nodes. It turns out that the square of this operator is smooth enough so that random subgraphs and subgraph densities can be defined by "classical" formulas. Also, the trace of $\mathbf{A}_d^k$ gives the density of k-cycles (at least for sufficiently large k).

Using the spectral decomposition of $\mathbf{A}_d$, we derive explicit formulas for the densities of cycles in H_d. As an interesting fact, cycle densities in H_4 can be expressed by the zeta-function.

Approximation by graphs and graphons. The third method of defining and calculating subgraph densities in H_d is based on approximating H_d by graphons and finite graphs, and calculating the density in H_d as the limit of densities in these approximations.

A consequence of our results is that H_d *is* the limit of finite graphs in the left-convergence sense. While this property is easy for graphons, it is not known in the bounded-degree case whether all graphings can be approximated by finite graphs (this is equivalent for the famous soficity problem for finitely generated groups). So the fact that H_d is "sofic" in this sense has some independent interest.

Finally, it should be noted that a good part of the results of this paper extend to more general Markov spaces. In particular, a general version of the operator $\mathbf{A}_d$ is called a *graphop* and it arises in the theory of action convergence [2] and, in an equivalent form, in the theory of shape convergence [9].

2 Preliminaries

2.1 Notation

We consider the unit sphere S^{d-1} in $\mathbb{R}^d$. (The cases $d \leq 2$ are very simple, so to avoid trivial complications, we assume throughout that $d \geq 3$.) For two real quantities (depending on a choice of points in S^{d-1}), let $A \lesssim B$ denote that there is a constant $c > 0$ such that $A \leq cB$. Here the constant may depend on the dimension and on the graph denoted by G, but not on other variables. Let A_k denote the surface area of S^k. It is well known that

$$A_k = \begin{cases} \dfrac{2(2\pi)^{k/2}}{(k-1)!!} & \text{if } k \text{ is even,} \\ \dfrac{(2\pi)^{(k+1)/2}}{(k-1)!!} & \text{if } k \text{ is odd,} \end{cases} \tag{1}$$

and for $a, b \in \mathbb{Z}_+$,

$$\int_0^{\pi/2} (\sin\theta)^a (\cos\theta)^b \, d\theta = \left(\frac{\pi}{2}\right)^{e(a,b)} \frac{(a-1)!!(b-1)!!}{(a+b)!!} = \frac{A_{a+b+1}}{A_a A_b}, \tag{2}$$

where $e(a,b) = (a-1)(b-1) \mod 2$, and $(-1)!! = 0!! = 1!! = 1$.

When we talk about a "random" point of a sphere, we mean a random point from the uniform distribution on the sphere.

2.2 Generalized determinants

For a finite set $X = \{x_1, \ldots, x_m\} \subseteq \mathbb{R}^d$, we define the quantity

$$\mathrm{Det}(X) = \mathrm{Det}(x_1, \ldots, x_m) = |x_1 \wedge \cdots \wedge x_m| = \sqrt{\det\left((x_i^\mathsf{T} x_j)_{i,j=1}^m\right)}.$$

We define $\mathrm{Det}(\emptyset) = 1$. For $m = 1$, $\mathrm{Det}(X) = \mathrm{Det}(x_1) = |x_1|$. Note that $\mathrm{Det}(X) \geq 0$, and $\mathrm{Det}(X) > 0$ if and only if X consists of linearly independent vectors.

Lemma 3 *Let $n, d \in \mathbb{N}$ and $p \in \mathbb{R}$ such that $1 \leq n < d$, and let $x_1, \ldots, x_n$ be independent random points on S^{d-1}. Then*

$$\mathsf{E}\Big(\frac{\mathrm{Det}(x_1, \ldots, x_n)^p}{\mathrm{Det}(x_1, \ldots, x_{n-1})^p}\Big) \qquad \textit{and} \qquad \mathsf{E}\big(\mathrm{Det}(x_1, \ldots, x_n)^p\big)$$

are finite if and only if $p > n - d - 1$. If p is an integer, then we have the explicit formulas

$$\mathsf{E}\Big(\frac{\mathrm{Det}(x_1, \ldots, x_n)^p}{\mathrm{Det}(x_1, \ldots, x_{n-1})^p}\Big) = \frac{A_{d+p-1} A_{d-n}}{A_{d-1} A_{d-n+p}}.$$

and

$$\mathsf{E}\big(\mathrm{Det}(x_1,\dots,x_n)^p\big) = \Big(\frac{A_{d+p-1}}{A_{d-1}}\Big)^{n-1} \frac{A_{d-2}\cdots A_{d-n}}{A_{d+p-2}\cdots A_{d+p-n}}.$$

In these expectations, we could condition on (say) a fixed x_1, by the symmetry of the sphere. Note that p may be negative, but if $p \le n-d-1$, then the expectations are infinite.

Proof. For $n = 1$ the identities are trivial, so we assume that $n \ge 2$. The ratio $\mathrm{Det}(x_1,\dots,x_n)/\mathrm{Det}(x_1,\dots,x_{n-1})$ is the (unsigned) distance of x_n from the subspace $L = \mathrm{lin}(x_1,\dots,x_{n-1})$, which has dimension $n-1$ with probability 1. The distribution of this distance is independent of $x_1,\dots,x_{n-1}$, so we may fix L and just take expectation in x_n.

Let θ be the angle between x_n and L ($0 \le \theta \le \pi/2$), then

$$\frac{\mathrm{Det}(x_1,\dots,x_n)}{\mathrm{Det}(x_1,\dots,x_{n-1})} = \sin\theta.$$

For a fixed θ, points at this distance from L form the direct product of the two spheres $L \cap (\cos\theta)S^{d-1}$ and $L^\perp \cap (\sin\theta)S^{d-1}$, and so their density is proportional to $(\cos\theta)^{n-2}(\sin\theta)^{d-n}$. Hence

$$\mathsf{E}\Big(\frac{\mathrm{Det}(x_1,\dots,x_n)^p}{\mathrm{Det}(x_1,\dots,x_{n-1})^p}\Big) = \frac{\int_0^{\pi/2} (\sin\theta)^{d-n+p}(\cos\theta)^{n-2}\,d\theta}{\int_0^{\pi/2} (\sin\theta)^{d-n}(\cos\theta)^{n-2}\,d\theta}.$$

Using that $2\theta/\pi \le \sin\theta \le \theta$, it follows that the numerator is finite if and only if $d_n + p > -1$, proving the first assertion. Substituting from (2) for integral p, we get the first formula in the lemma.

To prove the second identity, we use the telescopic product decomposition

$$\mathrm{Det}(x_1,\dots,x_n)^p = \prod_{r=2}^{n} \frac{\mathrm{Det}(x_1,\dots,x_r)^p}{\mathrm{Det}(x_1,\dots,x_{r-1})^p}.$$

As remarked above, the factors are independent random variables, and hence

$$\begin{aligned}\mathsf{E}(\mathrm{Det}(x_1,\dots,x_n)^p) &= \prod_{r=2}^{n} \mathsf{E}\Big(\frac{\mathrm{Det}(x_1,\dots,x_r)^p}{\mathrm{Det}(x_1,\dots,x_{r-1})^p}\Big) = \prod_{r=2}^{n} \frac{A_{d+p-1}A_{d-r}}{A_{d-1}A_{d-r+p}}\\ &= \Big(\frac{A_{d+p-1}}{A_{d-1}}\Big)^{n-1} \frac{A_{d-2}\cdots A_{d-n}}{A_{d+p-2}\cdots A_{d+p-n}}.\end{aligned}$$

□

For small values of $|p|$, we can cancel most of the terms on the right hand side of the second equality. The most important special case for us will be $p = -1$:

$$\mathsf{E}\Big(\frac{1}{\mathrm{Det}(x_1,\ldots,x_n)}\Big) = \frac{A^n_{d-2}}{A^{n-1}_{d-1}A_{d-n-1}}.$$

This identity makes sense for $n = 0$ as well, and it is trivially valid. In particular $1/\mathrm{Det}$ is integrable provided $n \leq d-1$. We shall make use of the following one-sided bound on averages of such inverses of determinants.

Lemma 4 *Let $x_1, \ldots, x_n \in S^{d-1}$, and let $B_r(x)$ denote the r-neighborhood of x on S^{d-1}. Let $D_r(x_1, \ldots, x_n)$ denote the average of $1/\mathrm{Det}(x_1, \ldots, x_n)$ over $B_r(x_1) \times \cdots \times B_r(x_n)$. Then*

$$D_r(x_1,\ldots,x_n)\mathrm{Det}(x_1,\ldots,x_n) < C_{n,d}, \tag{3}$$

where $C_{n,d} > 0$ may depend on d and n, but not on r and $(x_1, \ldots, x_n)$.

Proof. We shall fix d, and proceed by induction on n. The case $n = 1$ is trivial, as the determinant is constant 1. So let us now assume $C_{n-1,d}$ exists, and show that $C_{n,d}$ exists as well ($1 < n < d$). Since $1/\mathrm{Det}(x_1, \ldots, x_n)$ is positive, integrable and continuous outside of the null-set of its singularities, the map

$$(x_1,\ldots,x_n) \mapsto D_r(x_1,\ldots,x_n)\mathrm{Det}(x_1,\ldots,x_n)$$

is continuous, with a maximum M_r, and also $r \mapsto M_r$ is continuous on $(0, \infty)$. To show that M_r is bounded above note that it is constant once $r \geq \pi$, and so we only need to show that it remains bounded above on some finite interval $(0, \varepsilon]$, where ε may be chosen arbitrarily small. Let $q := 1/(\sqrt[n]{1,5} - 1)$, and set $\varepsilon := 1/(10q)$.

We distinguish two cases based on the relative positions of the points. We may assume without loss of generality that the minimal distance R from an x_j to the subspace generated by the other $n-1$ points is realized for $j = n$.

<u>Case 1</u>: $R \geq qr$.

Note that R is also a lower bound on any distance from one of the x_j's to any subspace generated by some selection of the other points. In particular, we have that for any $J \subseteq [n]$, $\left|\bigwedge_{j\in J} x_j\right| \leq \mathrm{Det}(x_1, \ldots, x_n)/R^{n-|J|}$. Therefore for any $\rho \in (\mathbb{R}^d)^n$

with $|\rho_k| < r$ for all $1 \le k \le n$,

$$\begin{aligned}
\mathrm{Det}(x_1+\rho_1,\dots,x_n+\rho_n) \ge\ & \mathrm{Det}(x_1,\dots,x_n) - \sum_{J\subsetneq[n]} \left| \bigwedge_{j\in J} x_j \bigwedge_{i\in[n]\setminus J} \rho_i \right| \\
&\ge \mathrm{Det}(x_1,\dots,x_n) - \sum_{J\subsetneq[n]} \mathrm{Det}(x_1,\dots,x_n)\frac{r^{n-|J|}}{R^{n-|J|}} \\
&\ge \mathrm{Det}(x_1,\dots,x_n)\left(1-\sum_{j=1}^{n}\frac{1}{q^j}\binom{n}{j}\right) \\
&= \mathrm{Det}(x_1,\dots,x_n)\left(2-\left(1+\frac{1}{q}\right)^n\right) = \frac{\mathrm{Det}(x_1,\dots,x_n)}{2}.
\end{aligned}$$

Consequently $D_r(x_1,\dots,x_n)\mathrm{Det}(x_1,\dots,x_n) \le 2$.

Case 2: $R < qr$.

In this case $\mathrm{Det}(x_1,\dots,x_{n-1},x_n) \le qr\mathrm{Det}(x_1,\dots,x_{n-1})$. Fix any choice of linearly independent points $y_i \in S^{d-1}$ $(1 \le i \le n-1)$. Then the set of points z such that $\mathrm{Det}(y_1,\dots,y_{n-1},z) = t\mathrm{Det}(y_1,\dots,y_{n-1})$ form a $(d-n+1)$-dimensional sphere of radius t around the $(n-2)$-dimensional subspace $\mathrm{lin}(y_1,\dots,y_{n-1})$. After intersecting with S^{d-1}, the dimension of suitable z's is reduced to $d-n>0$. Now $\int_0^r (t^{d-n})/t\,dt = r^{d-n}/(d-n)$, and so for any $y \in S^{d-1} \cap\ \mathrm{lin}(x_1,\dots,x_{n-1})$ we obtain that

$$\int\limits_{B_r(y)} \frac{1}{\mathrm{Det}(y_1,\dots,y_{n-1},z)}\,d\pi(z) \le \frac{C'_{d,n}r^{d-n}r^{n-2}}{\mathrm{Det}(y_1,\dots,y_{n-1})} = \frac{C'_{d,n}r^{d-2}}{\mathrm{Det}(y_1,\dots,y_{n-1})},$$

where $C'_{d,n}$ does not depend on r or y (recall that $r \in (0, 1/(10q)]$). Consequently,

$$\begin{aligned}
\mathbb{E}_{z\in B_r(y)}\left(\frac{1}{\mathrm{Det}(y_1,\dots,y_{n-1},z)}\right) &\le \frac{1}{\pi(B_r(y))}\frac{C'_{d,n}r^{d-2}}{r^{d-1}\mathrm{Det}(y_1,\dots,y_{n-1})} \\
&= \frac{C''_{d,n}}{r\mathrm{Det}(y_1,\dots,y_{n-1})}.
\end{aligned} \tag{4}$$

Also note that replacing y by any point not on $\mathrm{lin}(y_1,\dots,y_{n-1})$ will actually increase the expectation (the distribution of the values of t within the r-neighborhood gets shifted away from 0). Since the set of points $(y_1,\dots,y_{n-1}) \in \prod_{j=1}^{n-1} B(x_j,r)$ that are not a linearly independent $(n-1)$-tuple is of measure zero, we have the following.

$$\begin{aligned}
D_r(x_1,\dots,x_n) &\le qr\, D_r(x_1,\dots,x_{n-1}) \\
&= qr\,\mathbb{E}_{y_j\in B_r(x_j)}\mathbb{E}_{z\in B_r(x_n)}\left(\frac{1}{\mathrm{Det}(y_1,\dots,y_{n-1},z)}\right) \\
&\le qr\,\mathbb{E}_{y_j\in B_r(x_j)}\mathbb{E}_{z\in B_r(y_{n-1})}\left(\frac{1}{\mathrm{Det}(y_1,\dots,y_{n-1},z)}\right) \\
&\le qr\,\mathbb{E}_{y_j\in B_r(x_j)}\left(\frac{C''_{d,n}}{r\mathrm{Det}(y_1,\dots,y_{n-1})}\right) \\
&= q\,\mathbb{E}_{y_j\in B_r(x_j)}\left(\frac{C''_{d,n}}{\mathrm{Det}(y_1,\dots,y_{n-1})}\right) \\
&\le q\,\frac{C_{n-1,d}\,C''_{d,n}}{\mathrm{Det}(x_1,\dots,x_{n-1})}
\end{aligned}$$

(we have used (4) in the fourth step and the induction hypothesis in the last). This implies the inequality in the lemma. □

The following lemma connects these reciprocals of determinants to the orthogonality graph.

Lemma 5 *Let $0 \le n < d$, and let $(x_1,\dots,x_n)$ be obtained by selecting a random point y on S^{d-1}, then selecting n independent random points $x_1,\dots,x_n$ from the "equator" $y^\perp \cap S^{d-1}$, then forgetting y. Then the density function of $(x_1,\dots,x_n)$ is*

$$s_{d,n}(x_1,\dots,x_n) = \frac{A_{d-1}^{n-1}A_{d-n-1}}{A_{d-2}^{n}}\,\frac{1}{\mathrm{Det}(x_1,\dots,x_n)} \tag{5}$$

As remarked above in a different language (cf. Lemma 3), $s_{d,n} \in L_p(S^{d-1},\pi)$ $(p \ge 1)$ if and only if $p < d-n+1$.

Proof. Similarly as in the proof of Lemma 3, we use induction on n. For $n \le 1$ the assertion is trivial. Let $n \ge 2$, and let L denote the linear space spanned by $x_1,\dots,x_{n-1}$. With probability 1, $\dim(L) = n-1$. Clearly L is uniformly distributed among all $(n-1)$-dimensional subspaces of $y^\perp$, and so $y^\perp$ is uniformly distributed among all linear hyperplanes containing L. So we construct x_n by (a) choosing a random $(n-1)$-dimensional subspace L, (b) choosing a random hyperplane H containing L, and (c) choosing a random point from $H \cap S^{d-1}$. Let us fix L, and let θ be the angle between x_n and L. It is clear by symmetry that the density of x_n depends only on θ.

For every choice of H, the distribution of θ is the same, and the density of this distribution (in $[0,\pi]$ is proportional to $(\cos\theta)^{n-2}(\sin\theta)^{d-n-1}$, as we have seen in the proof of Lemma 3. By the same argument, for a uniform random point $x'_n \in S^{d-1}$, the angle θ' between x'_n and L has density proportional to $(\cos\theta)^{n-2}(\sin\theta)^{d-n}$. It follows that the density of x_n, relative to the uniform distribution on S^{d-1}, is proportional to

$$\frac{(\cos\theta)^{n-2}(\sin\theta)^{d-n-1}}{(\cos\theta)^{n-2}(\sin\theta)^{d-n}} = \frac{1}{\sin\theta} = \frac{\mathrm{Det}(x_1,\dots,x_{n-1})}{\mathrm{Det}(x_1,\dots,x_n)}.$$

Since the density of $(x_1,\dots,x_{n-1})$ is proportional to $1/\mathrm{Det}(x_1,\dots,x_{n-1})$ by the induction hypothesis, it follows that the density of $(x_1,\dots,x_n)$ is proportional to

$$\frac{1}{\mathrm{Det}(x_1,\dots,x_{n-1})}\frac{\mathrm{Det}(x_1,\dots,x_{n-1})}{\mathrm{Det}(x_1,\dots,x_n)} = \frac{1}{\mathrm{Det}(x_1,\dots,x_n)}.$$

The coefficient of proportionality can be computed by Lemma 3. $\square$

2.3 Conditioning and Markov property

Let V be a finite set, and let $(\Omega,\mathcal{A})$ be a measurable space (for most of the paper, $\Omega = S^{d-1}$, and $\mathcal{A}$ is the sigma-algebra of Borel sets). Let Ω^{*V} denote the set of partial mappings $z:\ S\to J$, $S\subseteq V$. Let φ be a measure on $(\Omega^V,\mathcal{A}^V)$, and let φ^S denote the marginal of φ on $(\Omega^S,\mathcal{A}^S)$.

A family $(\varphi_z:\ z\in\Omega^{*V})$ is a *conditioning of* φ, if

(C1) for every $S\subseteq V$ and $z\in J^S$, φ_z is a measure on $(\Omega^{V\setminus S},\mathcal{A}^{V\setminus S})$;

(C2) for every $S\subseteq V$ and $B\in\mathcal{A}^{V\setminus S}$, the value $\varphi_z(B)$ is a measurable function of $z\in\Omega^S$;

(C3) for every $T\subseteq S\subseteq V$, for every $z\in\Omega^T$, $B\in\mathcal{A}^{S\setminus T}$ and $C\in\mathcal{A}^{V\setminus S}$,

$$\varphi_z(B\times C) = \int\limits_B \varphi_{zy}(C)\, d\varphi_z^{S\setminus T}(y).$$

As extreme cases, $\varphi_\emptyset=\varphi$, and $\varphi_z=\delta_z$ (the Dirac measure) for $z\in\Omega^V$.

For a fixed set $S\subseteq V$, the family $\{\varphi_z:\ z\in\Omega^S\}$ is a disintegration of the measure φ according to the marginal φ^S. The conditioning as defined above means a bit more: first, it is well-defined for all $z\in\Omega^S$, not just almost everywhere; second, it is defined simultaneously for all marginals φ^S, with compatibility condition (C3).

If V is the set of nodes of a graph G, we can define an important probabilistic property of conditionings. A conditioning (φ_z) is *Markovian*, if for every $S\subseteq V$ and $z\in\Omega^S$, the measure φ_z is multiplicative over the connected components of $G\setminus S$. If, in particular, φ is a probability distribution, and $G\setminus S$ has connected components $G_1,\dots,G_r$, then $\varphi_z|_{G_1},\dots,\varphi_z|_{G_1}$ are independent.

2.4 Graphons

We conclude this section with a brief survey of related constructions for graphons, partly as analogues for the orthogonality graph (which is not a graphon), but also for later reference. A *graphon* is a symmetric integrable function $W:\ \Omega^2 \to \mathbb{R}_+$, where $(\Omega, \mathcal{A}, \pi)$ is a standard Borel probability space. In the theory of dense graph limits, graphons are bounded by 1, but since then much of the theory has been extended to the unbounded case.

Given a graphon W and a finite simple graph $G = (V, E)$, we define a function $W^G:\ \Omega^V \to \mathbb{R}_+$ for $x = (x_i:\ i \in V)$ by

$$W^G(x) = \prod_{ij \in E} W(x_i, x_j). \tag{6}$$

Sometimes it will be convenient to use this notation for a single edge $e = ij \in E$: $W^e(x) = W(x_i, x_j)$. The function W^G defines a measure η_W^G as its density function:

$$\eta_W^G(A) = \int\limits_A W^G(x)\, d\pi^V(x) \qquad (A \in \mathcal{A}^V),$$

and the *subgraph density*

$$t(G, W) = \eta_W^G(\Omega^V) = \int\limits_{\Omega^V} W^G(x)\, d\pi^V(x). \tag{7}$$

For bounded graphons this is always finite, but in general, it may be infinite.

We call a graphon 1-*regular*, if $\int_\Omega W(x, y)\, d\pi(y) = 1$ for every x. For a 1-regular graphon, the function $W(x, .)$ can be considered as the density function of a probability distribution ν_x on $(\Omega, \mathcal{A})$, which defines a step from $x \in \Omega$ of a time-reversible Markov chain. Let us make n independent steps, each from the same point x, to points $x_1, \dots, x_n$. The joint distribution of $(x, x_1, \dots, x_n)$ has density function $W(x, x_1) \dots W(x, x_n)$, and if x is chosen randomly from π, then the analogue of Lemma 5 says that the joint distribution of $(x_1, \dots, x_n)$ has density function

$$s(x_1, \dots, x_n) = \int\limits_\Omega W(x, x_1) \dots W(x, x_n)\, d\pi(x). \tag{8}$$

A Markovian conditioning of η_W^G can be constructed as the family of measures $\{\eta^z:\ z \in \Omega^S,\ S \subseteq V\}$, with density functions

$$t_z(G, W) = \int\limits_{\Omega^{V \setminus S}} W^G(y, z)\, d\pi^{V \setminus S}(y).$$

3 The main construction

3.1 Swapping lemmas

For the next (main) lemma, we need some geometric preparation. We fix the dimension d. Let L_i $(i = 1, 2)$ be linear subspaces of $\mathbb{R}^d$ of dimension $d_i \geq 2$. Let $x_i \in L_i \cap S^{d-1}$, and let $\widehat{x}_1$ and $\widehat{x}_2$ be the orthogonal projections of x_1 onto L_2 and of x_2 onto L_1, respectively. Define

$$\Omega = \{(x_1, x_2) \in L_1 \times L_2 : \; x_1 \perp x_2, \; , \widehat{x}_1, \widehat{x}_2 \neq 0\}.$$

Lemma 6 *Let X_i be a random vector from $L_i \cap S^{d-1}$, and let X_1' and X_2' be random vectors from the spheres $L_1 \cap S^{d-1} \cap X_2^\perp$ and $L_2 \cap S^{d-1} \cap X_1^\perp$, respectively. Let ρ_1 and ρ_2 be the distributions of (X_1, X_2') and (X_1', X_2), respectively. Then ρ_1 and ρ_2 are mutually absolutely continuous on Ω, and*

$$\frac{d\rho_2}{d\rho_1}(x_1, x_2) = \frac{A_{d_1-1}A_{d_2-2}}{A_{d_2-1}A_{d_1-2}} \frac{|\widehat{x}_2|}{|\widehat{x}_1|} = \frac{A_{d_1-2}}{A_{d_2-2}} \frac{|\widehat{x}_2|}{|\widehat{x}_1|}.$$

Proof. The first assertion follows by the considerations in [13, 12], and also from the computations below.

For a nonzero vector $u \in \mathbb{R}^d$, we set $u^0 = u/|u|$. Let $(x_1, x_2) \in \Omega$, let $u_1 = x_1, u_2 = \widehat{x}_2^0, u_3, \dots, u_{d_1}$ be an orthonormal basis in L_1, and select an orthonormal basis $v_1, \dots, v_{d_2}$ in L_2 analogously. Let $\|.\|_\infty$ denote the ℓ_∞ norm on each of L_1 and L_2 in these bases. Let T_i be the tangent space of the unit sphere U_i of L_i at x_i (as an affine subspace of L_i containing x_i). Fix an $\varepsilon > 0$. Let B_i be the cube $x \in T_i : \; \|x - x_i\|_\infty \leq \varepsilon$, and let B_i' denote the projection of B_i onto the sphere U_i from the origin.

For $y_1 \in B_1$, consider the linear subspace $H = H(y_1) = \{y_2 \in L_2 : \; y_1^\mathsf{T} y_2 = 0\}$ and the affine subspace $H' = H'(y_1) = \{y_2 \in L_2 : \; x_1^\mathsf{T} y_2 + x_2^\mathsf{T} y_1 = 0\}$. Note that the equation defining $H'(y_1)$ can be written as $H'(y_1) = \{y \in L_2 : \; \widehat{x}_1^\mathsf{T} y + \widehat{x}_2^\mathsf{T} y_1 = 0\}$, since $x_1 - \widehat{x}_1 \perp y$ and $x_2 - \widehat{x}_2 \perp y_1$. Furthermore, $x_2 - \widehat{x}_2 \perp x_1$ by the orthogonality of the projection, so $\widehat{x}_2 = x_2 - (x_2 - \widehat{x}_2) \perp x_1$. We claim that these two subspaces are almost the same:

Claim 1 *There is a constant $C > 0$ independent of ε such that $d(y_2, H') < C\varepsilon^2$ for every $y_2 \in H \cap B_2$, and $d(y_1, H') < C\varepsilon^2$ for every $y_1 \in H \cap B_2$.*

We use the identity

$$y_1^\mathsf{T} y_2 - (x_1^\mathsf{T} y_2 + x_2^\mathsf{T} y_1) = (y_1 - x_1)^\mathsf{T} (y_2 - x_2)$$

(all asymptotic statements concern $\varepsilon \to 0$). Here $\|y_i - x_i\| = O(\varepsilon)$, so the right hand side is $O(\varepsilon^2)$. Up to sign, the first term on the left is $\|y_1\| \, d(y_2, H)$, while the second term is $\|\widehat{x}_1\| d(y_2, H')$. If either one of these is 0, the other one is $O(\varepsilon^2)$.

Let X_i and X_i' be generated as in the statement of the Lemma. Then

$$\mathsf{P}\big((X_1, X_2') \in B_1' \times B_2'\big) = \mathsf{P}(X_2' \in B_2' \mid X_1 \in B_1')\mathsf{P}(X_1 \in B_1').$$

Here

$$\mathsf{P}(X_1 \in B_1') = \frac{\lambda_{d_1-1}(B_1')}{A_{d_1-1}} \sim \frac{\lambda_{d_1-1}(B_1)}{A_{d_1-1}} = \frac{(2\varepsilon)^{d_1-1}}{A_{d_1-1}}.$$

(where λ_k denotes the k-dimensional volume in $\mathbb{R}^d$). The first factor is more complicated. For a fixed $y_1 \in B_1$, we have

$$\mathsf{P}(X_2' \in B_2' \mid X_1 = y_1^0) = \frac{\lambda_{d_2-2}(B_2' \cap H(y_1))}{A_{d_2-2}} \sim \frac{\lambda_{d_2-2}(B_2 \cap H(y_1))}{A_{d_2-2}}.$$

We want to compare $B_2 \cap H(y_1)$ and $B_2 \cap H'(y_1)$. The hyperplane $H'(y_1)$ in L_2 is orthogonal to the edge v_2 of the cube B_2, and hence it either avoids B_2 or intersects it in a set isometric with a facet. Using the fact that $H(y_1)$ and $H'(y_1)$ are very close, we get that there is a $C > 0$, independent of ε, such that

$$\text{if } d(x_2, H'(y_1)) < \varepsilon - C\varepsilon^2, \text{ then } \lambda_{d_2-2}(B_2 \cap H(y_1)) \sim (2\varepsilon)^{d_2-2}, \tag{9}$$

and

$$\text{if } d(x_2, H'(y_1)) > \varepsilon + C\varepsilon^2, \text{ then } \lambda_{d_2-2}(B_2 \cap H(y_1)) = 0. \tag{10}$$

In the modified equation defining H', the coefficient vector $\widehat{x}_1 \in L_2$, and hence

$$d(x_2, H'(y_1)) = \frac{1}{|\widehat{x}_1|}\Big|\widehat{x}_1^{\mathsf{T}} x_2 + \widehat{x}_2^{\mathsf{T}} y_1\Big| = \frac{|\widehat{x}_2|}{|\widehat{x}_1|}|u_1^{\mathsf{T}} y_1|.$$

Hence

$$\mathsf{P}(X_2' \in B_2' \mid X_1 = y_1^0) \sim \begin{cases} \dfrac{(2\varepsilon)^{d_2-2}}{A_{d_2-2}}, & \text{if } |u_1^{\mathsf{T}} y_1| < (\varepsilon - C\varepsilon^2)|\widehat{x}_1|/|\widehat{x}_2|, \\ 0, & \text{if } |u_1^{\mathsf{T}} y_1| > (\varepsilon + C\varepsilon^2)|\widehat{x}_1|/|\widehat{x}_2|, \\ O(\varepsilon^{d_2-2}), & \text{otherwise.} \end{cases}$$

The first option applies for a fraction of $\min\{1, (1 - C\varepsilon)|\widehat{x}_2|/|\widehat{x}_1|\}$ of points of B_1. The third possibility occurs for a negligible fraction of the points of B_1. Since the distribution of X_1 in B_1 is almost uniform, we get

$$\mathsf{P}\big((X_1, X_2') \in B_1 \times B_2\big) \sim \frac{(2\varepsilon)^{d_1-1}}{A_{d_1-1}} \frac{(2\varepsilon)^{d_2-2}}{A_{d_2-2}} \frac{\min(|\widehat{x}_1|, |\widehat{x}_2|)}{|\widehat{x}_2|}.$$

Similarly,

$$\mathsf{P}\big((X_1', X_2) \in B_1 \times B_2\big) \sim \frac{(2\varepsilon)^{d_2-1}}{A_{d_2-1}} \frac{(2\varepsilon)^{d_1-2}}{A_{d_1-2}} \frac{\min(|\widehat{x}_1|, |\widehat{x}_2|)}{|\widehat{x}_1|}.$$

and so

$$\frac{d\rho_2}{d\rho_1}(x_1, x_2) \sim \frac{\mathsf{P}\big((X_1', X_2) \in B_1 \times B_2\big)}{\mathsf{P}\big((X_1, X_2') \in B_1 \times B_2\big)} \sim \frac{A_{d_1-1}A_{d_2-2}}{A_{d_2-1}A_{d_1-2}} \frac{|\widehat{x}_2|}{|\widehat{x}_1|}.$$

Letting $\varepsilon \to 0$, the lemma follows. $\square$

Let $p: V \to [n]$ be a bijection, defining an ordering of the nodes of a graph $G = (V, E)$, let $N_p(u) = \{w : p(w) < p(u), uw \in E\}$, and let $d_p(u) = |N_p(u)|$. To simplify notation, for a map $x: V \to \mathbb{R}^d$, we write $x_p(u) = x|_{N_p(u)}$.

We recall more formally the construction of an ortho-homomorphism from the Introduction. Let $v \in V$, $S = \{u \in V : p(u) < p(v)\}$, and suppose that we already have an ortho-homomorphism $(x_u : u \in S)$ in general position for the subgraph $G[S]$. The vectors in S^{d-1} orthogonal to every x_i with $i \in N_p(v)$ form a sphere of dimension at least $(d-1) - d_p(v) \geq 0$; we choose a vector x_v on this sphere randomly. Repeating this until x_v is defined for every $v \in V$, we get an ortho-homomorphism, which we call a *random sequential ortho-homomorphism* of G. Let ρ_p be the distribution of this ortho-homomorphism. By [13], this ortho-homomorphism is in general position almost surely. The main step in the proof was that flipping two consecutive nodes in the ordering, we may get a possibly different distribution on ortho-homomorphisms, but this new new distribution is absolutely continuous with respect to the previous one. In the next lemma, we give an explicit formula showing this.

Lemma 7 *Let r be obtained from the ordering p by flipping two consecutive adjacent nodes u and v, where $p(v) = p(u) + 1$. Then*

$$\frac{d\rho_r}{d\rho_p}(x) = \frac{A_{d-d_p(u)-1}A_{d-d_p(v)-1}}{A_{d-d_r(u)-1}A_{d-d_r(v)-1}} \frac{\mathrm{Det}(x_r(u))\mathrm{Det}(x_r(v))}{\mathrm{Det}(x_p(u))\mathrm{Det}(x_p(v))}.$$

Proof. We apply Lemma 6 with $L_1 = N_p(u)^\perp$ and $L_2 = N_r(v)^\perp$. Then $\dim(L_1) = d - d_p(u) \geq d - \deg(u) + 1 \geq 2$ (since v is not counted in $d_p(u)$), and similarly $\dim(L_2) = d - d_r(v) \geq 2$. Also note that $d_r(u) = d_p(u) + 1$, $d_r(v) = d_p(v) - 1$, and $d_r(w) = d_p(w)$ for every $w \neq u, v$. Since $x_p(u)$ is a basis in $L_1^\perp$ and $x_r(v) = x_p(v) \setminus \{u\}$ is a basis in $L_2^\perp$, the length of the orthogonal projection of x_u onto L_2 is

$$\frac{\mathrm{Det}(x_r(v) \cup \{u\})}{\mathrm{Det}(x_r(v))} = \frac{\mathrm{Det}(x_p(v))}{\mathrm{Det}(x_r(v))}$$

and the length of orthogonal projection of x_v onto L_1 is

$$\frac{\mathrm{Det}(x_p(u) \cup \{v\})}{\mathrm{Det}(x_p(u))} = \frac{\mathrm{Det}(x_r(u))}{\mathrm{Det}(x_p(u))}.$$

This implies, in particular, that these projections are nonzero, and so we can apply Lemma 6. Since the order in which x_u and x_v are chosen does not influence the distribution of $(x_w : \ p(w) < p(u))$ and the distribution of $(x_w : \ p(w) > p(v))$ conditional on $(x_w : \ p(w) \le p(v))$, we get that

$$\frac{d\rho_p}{d\rho_r}(x) = \frac{A_{d-d_p(u)-1}A_{d-d_p(v)-1}}{A_{d-d_r(v)-1}A_{d-d_r(u)-1}} \left(\frac{\mathrm{Det}(x_r(u))}{\mathrm{Det}(x_p(u))} \Big/ \frac{\mathrm{Det}(x_p(v))}{\mathrm{Det}(x_r(v))} \right),$$

proving the lemma. □

3.2 Order-independent measure

Let $p : \ V \to [n]$ be an ordering of the nodes of a graph G, and let $x : \ V \to \mathbb{R}^d$ be an orthogonal representation in general position. Using the functions defined in (5), let

$$\begin{aligned} f_p(x) &= \prod_{v \in V} s_{d,n}(x_p(v)) = \prod_{v \in V} \frac{A_{d-1}^{d_p(v)-1} A_{d-d_p(v)-1}}{A_{d-2}^{d_p(v)}} \frac{1}{\mathrm{Det}(x_p(v))} \\ &= \frac{A_{d-1}^{|E|-|V|}}{A_{d-2}^{|E|}} \prod_{v \in V} \frac{A_{d-d_p(v)-1}}{\mathrm{Det}(x_p(v))}. \end{aligned} \tag{11}$$

We define a measure $\varphi_p = f_p \cdot \rho_p$ on Σ; more explicitly,

$$\varphi_p(A) = \int_A f_p \, d\rho_p. \tag{12}$$

The following lemma is the main property of this construction.

Lemma 8 *The measure φ_p is independent of the ordering p.*

By this lemma, we can denote φ_p simply by φ or φ_G. We can think of φ either as a measure on Σ_G, or as a measure on $(\mathbb{R}^d)^V$ concentrated on Σ_G.

Proof. It suffices to check that if r is the permutation obtained from p by swapping two consecutive nodes u and v, then $\varphi_p = \varphi_r$. If u and v are nonadjacent, then this is trivial: $\rho_p = \rho_r$ and $N_p(w) = N_r(w)$ for every node w, and hence $f_p = f_r$. So suppose that $uv \in E$, and (say) $p(v) = p(u) + 1$. Then

$$\frac{d\varphi_p}{d\varphi_r}(x) = \frac{f_p(x)}{f_r(x)} \cdot \frac{d\rho_p}{d\rho_r}(x).$$

Here

$$\frac{f_p(x)}{f_r(x)} = \frac{A_{d-d_p(u)-1}A_{d-d_p(v)-1}}{A_{d-d_r(v)-1}A_{d-d_r(u)-1}} \frac{\mathrm{Det}(x_r(u))\mathrm{Det}(x_r(v))}{\mathrm{Det}(x_p(u))\mathrm{Det}(x_p(v))}$$

by definition (11). Substituting for the second factor from Lemma 7, we get

$$\frac{d\varphi_p}{d\varphi_r}(x) = 1.$$

Since this holds for all $x \in \Sigma$, this proves that $\varphi_r = \varphi_p$. $\square$

The measure φ is not always finite: in Section 3.4 we show that it is finite for every even cycle longer than 4 in dimension 3, but infinite for the 3-cube in dimension 4. The measure is, however, finite on compact subsets of Σ_G: the denominator in (11) remains bounded away from zero. It is easy to see that φ is a Radon measure.

We will also be interested in the *ortho-homomorphism number* (of graph G in dimension d)

$$t(G,d) = \varphi(\Sigma_G) = \frac{A_{d-1}^{|E|-|V|}}{A_{d-2}^{|E|}} \int\limits_{(S^{d-1})^V} \prod_{v \in V} \frac{A_{d-d_p(v)-1}}{\mathrm{Det}(x_p(v))}\, d\pi^V. \tag{13}$$

Let us note that $t(G,d)$ is positive for every d-sparse graph G; but it may be infinite. If $t(G,d)$ is finite, then we can scale φ to get a probability measure on ortho-homomorphisms of G into S^{d-1}.

The measure φ has a natural conditioning. We can think of the construction of the measure φ as follows: Choose the vectors x_i in any given order according to the random sequential rule; whenever x_i is chosen, we multiply the density function by $s_{d,n}(x_p(v))$ (which is determined by the previous nodes). An important consequence of this fact is that if we stop when a subset S of nodes has been processed, the vectors selected and the density function computed up to this point define an ortho-homomorphism from the measure $\varphi_{G[S]}$.

For $z \in \Sigma_{G[S]}$, we construct a measure φ_z on $\Sigma_{G[V\setminus S]}$ by continuing the random sequential choice. Formally, let p be any ordering of the nodes of G starting with S; extend z to an ortho-homomorphism x of G in $\mathbb{R}^d$ by random sequential choice; let $\rho_{z,p}$ be the distribution of this extension. Define the density function the measure φ_z on $\Sigma_{G[V\setminus S]}$ by

$$\varphi_z(A) = \int\limits_A \prod_{u \in V\setminus S} s_{d,n}(x_p(u))\, d\rho_{z,p}. \tag{14}$$

Lemma 9 *The family $(\varphi_z :\ z \in \Sigma_{G[S]},\ S \subseteq V)$ is a Markovian conditioning of φ.*

Proof. The fact that the family (φ_z) is a conditioning follows from the construction of φ as described above.

The Markov property is easy as well. Let $S \subseteq V$, and let $G_1, \dots, G_r$ be the connected components of $G \setminus S$. Let $z \in \Sigma_{G[S]}$. Constructing the random extension of z sequentially, we see $\varphi_z|_{G_i}$ is independent of the vectors and density function values of the other components G_j. $\square$

Lemmas 8 and 9 imply Theorem 2.

The following related fact was observed and used (implicitly) in [13].

Proposition 10 *Let $S \subseteq V(G)$, let $G' = G[S]$, and let p be an ordering of the nodes of G starting with S. Then $\rho_p^S = \rho_{G',p}$, and φ_G^S is absolutely continuous with respect to $\varphi_{G'}$, and vice versa.*

Proof. The first assertion is obvious from the sequential construction of ρ_p. By construction, $\varphi_{G'}$ and $\rho_{G',p}$ are mutually absolutely continuous, and so are φ_G and ρ_p. The second assertion implies that their marginals $\rho_p^S = \rho_{G',p}$ and φ_G^S are mutually absolutely continuous, and hence so are φ_G^S and $\varphi_{G'}$. $\square$

3.3 Explicitly computable examples

Example 1 (Trees) A simple example is a tree F. Let p be a search order from a root u. Then $d_p(v) = 1$ for every $v \neq u$, and $\mathrm{Det}(x_p(u)) = 1$ for every u. Hence $f_p(x) \equiv 1$, ρ_p is the same distribution for every search order, and $\varphi = \rho_p$. Thus the measure $\varphi(F, d)$ is a well defined probability distribution, and $t(G, d) = 1$. The more general example of bipartite graphs will be discussed in Section 3.4.

Example 2 (Triangles) Let $d = 3$ and $G = K_3$, with the nodes labeled $1, 2, 3$. Then (x_1, x_2) is uniformly distributed on orthogonal pairs. It follows that $\mathrm{Det}(x_1, x_2) = 1$, and so all the Det's in the denominator of (13) are 1. Hence

$$t(K_3, d) = \frac{A_{d-1}A_{d-3}}{A_{d-2}^2} = \frac{(\pi/2)^{(-1)^d}\big((d-3)!!\big)^2}{(d-2)!!(d-4)!!}. \qquad (15)$$

In particular, $t(K_3, 3) = 2/\pi$ and $t(K_3, 4) = \pi/4$. Other cycles will be discussed in Sections 3.4 and 5.2.

Example 3 (Rigid Circuit Graphs) We can get rid of the integration for all *rigid circuit graphs*, which contain no induced cycles other than triangles. A well-known characterization of these graphs is that their nodes can be ordered so that the neighbors of any node v preceding it spans a complete subgraph. Using this ordering p to compute $t(G, d)$ (where d is large enough so that G is d-sparse), we see that the vectors in every $x(N_p(v))$ are mutually orthogonal, and so the Det's in the denominator are 1. Hence we get

$$t(G, d) = \frac{A_{d-1}^{|E|-|V|}}{A_{d-2}^{|E|}} \prod_{v \in V} A_{d-d_p(v)-1}. \qquad (16)$$

In particular, we get a formula for complete graphs:

$$t(K_r, d) = \frac{A_{d-1}^{r(r-3)/2}}{A_{d-2}^{r(r-1)/2}} \prod_{i=1}^{r} A_{d-i}. \qquad (17)$$

We could use (1) to express (16) as $a\pi^b$, where a is rational and b is an integer. Let q denote the number of odd "backward" degrees $d_p(v)$. Then straightforward computation gives that

$$b = \begin{cases} \dfrac{|E|-q}{2}, & \text{if } d \text{ is even,} \\ \dfrac{q-|E|}{2}, & \text{if } d \text{ is odd.} \end{cases}$$

Surprisingly, this exponent does not depend on d except for its sign. The combinatorial significance of the rational coefficient a would be interesting to determine.

Example 4 (Complete bipartite graphs) Let $G = K_{a,b}$, where $a+b \le d$. Then, using (13) and Lemma 3,

$$t(G,d) = \frac{A_{d-1}^{ab-b} A_{d-a-1}^{b}}{A_{d-2}^{ab}} \int\limits_{(S^{d-1})^a} \frac{1}{\mathrm{Det}(x)^b}\, d\pi^a(x)$$
$$= \frac{A_{d-1}^{ab-b-a} A_{d-b-1}^{a} A_{d-a-1}^{b} A_{d-1} \cdots A_{d-a}}{A_{d-2}^{ab} A_{d-b-1} \cdots A_{d-b-a}}. \tag{18}$$

This implies that $t(G,d)$ can again be expressed as a rational multiple of an integer power of π, where the exponent of π depends on the parity of d only.

3.4 Bipartite graphs

Let G be a bipartite graph with bipartition $V = U \cup W$. The construction of the measure φ can be carried out by ordering the nodes starting with U, to get the reference ordering p of V. If x is a random point from ρ_p, then x_u ($u \in U$) are independent random vectors in S^d, and f_p depends only on these variables x_u. Furthermore, $s_{d,n}(x_p(u)) = A_{d-1}$ for $u \in U$. Hence (canceling $A_{d-1}^{|U|}$)

$$f_p(x) = \frac{A_{d-1}^{|E|-|W|}}{A_{d-2}^{|E|}} \prod_{v \in W} \frac{A_{d-d(v)-1}}{\mathrm{Det}(x(N(v)))},$$

and

$$t(G,d) = \int\limits_{(S^{d-1})^V} f_p\, d\rho_p = \int\limits_{(S^{d-1})^U} \prod_{v \in W} s_{d,d_p(v)}(x)\, d\pi^U$$
$$= \frac{A_{d-1}^{|E|-|W|}}{A_{d-2}^{|E|}} \int\limits_{(S^{d-1})^U} \prod_{v \in W} \frac{A_{d-d(v)-1}}{\mathrm{Det}(x(N(v)))}\, d\pi^U \tag{19}$$

We'll see examples where this number is finite and also where this number is infinite.

Remark 11 An ortho-homomorphism of a bipartite graph has the following geometric interpretation. Consider the points of U as vectors in $\mathbb{R}^d$ (as before), but the points of W as normal vectors of hyperplanes. Orthogonality translates to incidence. For example, a representation of C_6 in $\mathbb{R}^3$ is a simplicial cone, with three rays (the vectors in their direction having unit length), and the normals of the three faces (again, of unit length).

Remark 12 We'll see (Corollary 27) that Sidorenko's conjecture would imply the inequality

$$t(G,d) \geq 1 \tag{20}$$

for every d-sparse bipartite graph G. It would be interesting to prove this inequality at least in this special case.

As a simple but important special class of bipartite graphs, the subdivision $G = H'$ of a simple graph H by one node on each edge is a bipartite graph. Then

$$t(H',d) = \frac{A_{d-1}^{|E|} A_{d-3}^{|E|}}{A_{d-2}^{2|E|}} \int\limits_{(S^{d-1})^U} \prod_{ij\in E} \frac{1}{\mathrm{Det}(x_i,x_j)}\, d\pi^V. \tag{21}$$

We can use this special case to justify considering $t(G,d)$ as the density of G in the orthogonality graph. The function

$$s_{d,2}(x,y) = \frac{A_{d-1}A_{d-3}}{A_{d-2}^2} \frac{1}{\mathrm{Det}(x,y)} \tag{22}$$

defines a graphon $(S^{d-1}, \pi, s_{d,2})$. Let $n = |V(H)|$, $m = |E(H)|$, and assume that all degrees of H are bounded by $d-1$. Then the ortho-homomorphism number can be expressed as follows.

$$t(H',d) = \int\limits_{(S^{d-1})^V} \prod_{ij\in E(H)} s_{d,2}(x(N(v)))\, d\pi^V(x) = t(H, s_{d,2}). \tag{23}$$

Since $t(H',W) = t(H, W\circ W)$ for any graphon W, this justifies to consider $t(G,d)$ as the homomorphism density of G in the orthogonality graph (with the edge measure scaled to a probability measure).

Example 5 Let $d = 3$ and $G = C_6 = K_3'$. Then the conditions above are satisfied, and (19) gives that

$$t(C_6,3) = \frac{8}{\pi^3} \int\limits_{(S^2)^3} \frac{1}{\mathrm{Det}(x_1,x_2)\mathrm{Det}(x_2,x_3)\mathrm{Det}(x_3,x_1))}\, d\pi^3(x).$$

Let $\measuredangle(x_1, x_3) = \alpha$, $\measuredangle(x_2, x_3) = \beta$, and $\measuredangle(x_1, x_2) = \gamma$. Let θ denote the (unsigned) angle between the planes $\mathrm{lin}(x_1, x_3)$ and $\mathrm{lin}(x_2, x_3)$. By the spherical cosine theorem, $\gamma = \gamma(\alpha, \beta, \theta)$ is given by

$$\cos\gamma = \cos\alpha\cos\beta - \sin\alpha\sin\beta\cos\theta. \tag{24}$$

Fixing x_3, it is easy to see that the angles α, β and θ are independent random variables with values in $[0, \pi]$; their density functions are $\frac{1}{2}\sin\alpha$, $\frac{1}{2}\sin\beta$ and $1/\pi$, respectively. Hence

$$t(C_6, 3) = \frac{8}{\pi^3} \int\limits_{[0,\pi]^3} \frac{\frac{1}{2}\sin\alpha\,\frac{1}{2}\sin\beta\,\frac{1}{\pi}}{\sin\alpha\,\sin\beta\,\sin\gamma}\,d\alpha\,d\beta\,d\theta = \frac{2}{\pi^4} \int\limits_{[0,\pi]^3} \frac{1}{\sin\gamma}\,d\alpha\,d\beta\,d\theta.$$

Substituting from (24),

$$t(C_6, 3) = \frac{2}{\pi^4} \int\limits_{[0,\pi]^3} \frac{d\alpha\,d\beta\,d\theta}{\sqrt{1 - \big(\cos\alpha\cos\beta - \sin\alpha\sin\beta\cos\theta\big)^2}}. \tag{25}$$

4 Finiteness

The value $t(G, d)$, as defined by the integral in (13), may be finite or infinite even for d-sparse graphs, as we will show below. In this section, we study the issue of finiteness. We only address this issue for bipartite graphs. Further exact formulas, based on spectral methods, will be given in the next section.

4.1 A general bound

A general upper bound on $t(G, d)$ can be obtained by applying the following generalized Hölder inequality [5].

Lemma 13 *Let $f_1, \ldots, f_m : \Omega^n \to \mathbb{R}$ be measurable n-variable functions on some probability space $(\Omega, \mathcal{A}, \pi)$, such that f_i depends only on a subset $B_i \subseteq [n]$ of the variables. Let $p_1, \ldots, p_m \geq 1$ such that*

$$\sum_{i:\, B_i \ni j} \frac{1}{p_i} \leq 1 \qquad (j = 1, \ldots, n).$$

Then

$$\int\limits_{\Omega^n} f_1 \ldots f_m \, d\pi^n \leq \|f_1\|_{p_1} \ldots \|f_m\|_{p_m}.$$

Of course, the lemma is only interesting if the right hand side is finite, i.e., if $f_i \in L_{p_i}(\Omega^m)$ for every i. Applying it to the expression (19), we get

$$t(G,d) \le \prod_{v\in W} \|s_{d,\deg(v)}\|_{p_v}, \tag{26}$$

where the numbers p_v must satisfy

$$\sum_{v\in N(u)} \frac{1}{p_v} \le 1 \tag{27}$$

for all $(u \in U)$. The bound is finite when $s_{d,\deg(v)} \in L_{p_v}$ for all $v \in W$; as noted in Section 2.2, this happens if and only if

$$p_v \le d - \deg(v) \qquad (v \in W). \tag{28}$$

The upper bound in (26) can be expressed explicitly using Lemma 3, but it is not really appealing. However, the finiteness result is worth stating:

Theorem 14 *Let G be a bipartite graph with bipartition (U,W), and suppose that*

$$\sum_{v\in N(u)} \frac{1}{d-\deg(v)} \le 1 \tag{29}$$

for all $u \in U$. Then $t(G,d)$ is finite.

Proof. The condition implies that G is d-sparse, and so $t(G,d)$ is well defined. Choosing $p_v = d - \deg(v)$, (27) and (28) are satisfied. □

A special case when this condition is satisfied and that is easier to handle is the following.

Corollary 15 *Let G be a bipartite graph with bipartition (U,W), and suppose that all degrees in U are bounded by a $(1 \le a \le d-2)$, and all degrees is W are bounded by $d-a$. Then $t(G,d)$ is finite.*

4.2 Subdivisions

In this section we show:

Theorem 16 *If G is the subdivision (with one node on each edge) of a simple graph H with maximum degree $d-1$, then $t(G,d)$ is finite.*

A notable special case for $d=3$ is the cycle C_{2k}, as the subdivision of C_k $(k \ge 3)$. Note that this Proposition does not follow from Corollary 15 (only if the degrees are strictly smaller than $d-1$).

We need a simple lemma in elementary graph theory.

Lemma 17 *Let $G = (V, E)$ be a simple connected graph on $n \geq 2$ nodes, with all degrees at most D, let $e_1 \in E$ and $w: E \to \mathbb{R}_+$. Then there is a spanning tree F of G, and integers $(k_e : e \in E(F)$ such that $k_{e_1} = 1$, $k_e \leq D$ for all $e \in E(F)$, $\sum_e k_e = |E(G)|$, and*

$$\sum_{e \in E} w(e) \leq \sum_{e \in E(F)} k_e w(e).$$

Proof. Let F be a maximum weight spanning tree. It suffices to define a map $\phi: E(G) \to E(F)$ such $w(\phi(e)) \geq w(e)$, $k_e = |\phi^{-1}(e)| \leq D$, and $|\phi^{-1}(e_1)| = 1$. For any search order $(v_1, \ldots, v_n)$ of F starting with $e_1 = v_1 v_2$, we map each edge $v_i v_j$ $(i < j)$ to the (unique) edge $v_j v_{j'}$ of F with $j' < j$. It is easy to see that this map satisfies our requirements: at most D edges are mapped onto any edge of F, no edge other than itself is mapped onto e_1, and if $v_i v_j$ is mapped onto $v_{j'} v_j$, then $w(v_i v_j) \leq w(v_{j'} v_j)$, because otherwise replacing the edge $v_{j'} v_j$ by $v_i v_j$, we would get a tree with larger weight than F. □

Proof of Theorem 16. By identity (23), we have

$$t(G, d) = t(H, W),$$

where $W = s_{d,2}$ defines a graphon on S^{d-1}. Lemma 17, applied to the logarithm of W, gives that for every $x \in (S^{d-1})^V$ there is a spanning tree F of G and integers $(k_e : e \in E(F))$ such that $k_e \leq d - 1$, $k_{e_1} = 1$, $\sum_e k_e = |E(G)|$, and

$$W^G(x) \leq \prod_{e \in E(F)} (W^e(x))^{k_e}.$$

Since W is bounded from below, this implies that

$$W^G(x) \leq C_0 W^{e_1}(x) W^{F \setminus e}(x)^{d-1} \tag{30}$$

for some constant C_0 independent of x. For a spanning tree F of G with an ordered edge set $E(F) = \{e_1, e_2, \ldots, e_{n-1}\}$, let Y_F denote the set of points $x \in (S^{d-1})^V$ for which $W^{e_1}(x) \geq W^{e_2}(x) \geq \ldots$ and (30) is satisfied. By the above, $\cup_F Y_F = (S^{d-1})^V$, and so

$$\begin{aligned} t(G, W) = \int\limits_{(S^{d-1})^V} W^G \, d\pi^V &\leq \sum_F \int\limits_{Y_F} W^G \, d\pi^V \\ &\lesssim \sum_F \int\limits_{Y_F} W^{e_1}(x) W^{F \setminus e_1}(x)^{d-1} \, d\pi^V(x) \\ &\lesssim \max_F \int\limits_{Y_F} W^{e_1}(x) W^{F \setminus e_1}(x)^{d-1} \, d\pi^V(x) \end{aligned}$$

So it suffices to prove that

$$\int\limits_{Y_F} W^{e_1}(x) W^{F\setminus e_1}(x)^{d-1}\, d\pi^V(x) = \int\limits_{Y_F} W^{e_1}(x) W^{e_2}(x)^{d-1} \dots W^{e_{n-1}}(x)^{d-1}\, d\pi^V(x) \quad (31)$$

is finite for every edge-ordered spanning tree F.

Disregarding the condition on the ordering of the edges, the random variables $W^{e_i}(x)$ are independent. Indeed, selecting the images of the nodes in a search order of the tree, each $W^{e_i}(x)$ will have the same distribution even with one endpoint of e_i already fixed, by symmetry. Let $\vartheta_i(x)$ be the angle between x_u and x_v, where $e_i = uv$. Then

$$W^{e_i}(x) \lesssim \frac{1}{\sin\vartheta_i},$$

and the density function of each $\vartheta_i(x)$ is

$$f(\vartheta) \lesssim (\sin\vartheta)^{d-2}$$

Let $T(F)$ denote the set of vectors $(\vartheta_1, \dots, \vartheta_{n-1})$ with $0 \le \vartheta_i \le \pi$ and $\sin\vartheta_1 \le \dots \le \sin\vartheta_{n-1}$. Then

$$\begin{aligned}
&\int\limits_{Y_F} W^{e_1}(x) W^{e_2}(x)^{d-1} \dots W^{e_{n-1}}(x)^{d-1}\, d\pi^V(x) \\
&\quad \lesssim \int\limits_{T(F)} \frac{(\sin\vartheta_1)^{d-2} \dots (\sin\vartheta_{n-1})^{d-2}}{\sin\vartheta_1 (\sin\vartheta_2)^{d-1} \dots (\sin\vartheta_{n-1})^{d-1}}\, d\vartheta_1 \dots d\vartheta_{n-1} \\
&\quad = \int\limits_{T(F)} \frac{(\sin\vartheta_1)^{d-3}}{\sin\vartheta_2 \dots \sin\vartheta_{n-1}}\, d\vartheta_1 \dots d\vartheta_{n-1} \\
&\quad \le \int\limits_{T(F)} \frac{1}{\sin\vartheta_2 \dots \sin\vartheta_{n-1}}\, d\vartheta_1 \dots d\vartheta_{n-1}
\end{aligned}$$

Introducing the variables $\phi_i = \min(\vartheta_i, \pi - \vartheta_i)$ and the set

$$T'(F) = \{(\phi_1, \dots, \phi_{n-1}) : \ 0 \le \phi_i \le \pi/2, \ \phi_1 \le \dots \le \phi_{n-1}\},$$

we can go on as follows:

$$\int\limits_{T'(F)} \frac{1}{\sin\phi_2\cdots\sin\phi_{n-1}}\,d\phi_1\dots d\phi_{n-1} \lesssim \int\limits_{T'(F)} \frac{1}{\phi_2\cdots\phi_{n-1}}\,d\phi_1\dots d\phi_{n-1}$$

$$= \int\limits_0^{\pi/2}\int\limits_0^{\phi_{n-2}}\dots\int\limits_0^{\phi_2} \frac{1}{\phi_2\cdots\phi_{n-1}}\,d\phi_1\dots d\phi_{n-1}$$

$$= \int\limits_0^{\pi/2}\int\limits_0^{\phi_{n-2}}\dots\int\limits_0^{\phi_3} \frac{1}{\phi_3\cdots\phi_{n-1}}\,d\phi_2\dots d\phi_{n-1}$$

$$= \dots = \int\limits_0^{\pi/2} 1\,d\phi_{n-1} = \pi/2.$$

This proves the theorem. □

Remark 18 Theorem 16 asserts that the subdivision of any simple graph with all degrees at most $d-1$ has a finite Markovian probability distribution on its ortho-homomorphisms. (This does not remain true for multigraphs, as shown by the multigraph on two nodes connected by $d-1$ edges.)

As remarked after Lemma 5, $s_{d,2} \in L_p(S^{d-1}, \pi)$ if and only if $p < d-1$, so $W = s_{d,2}$ is an L_p-graphon for every $p < d-1$, as defined by Borgs et al. in [3]. By one of the results of that paper, all simple graphs with all degrees at most $d-2$ have a finite density in W; our analysis shows that this remains valid for graphs with degrees bounded by $d-1$ in the special case of $s_{d,2}$.

4.3 Paths and cycles

Theorem 16 implies that every even cycle C_{2k} of length at least 6 has finite density in every dimension $d \ge 3$. We are going to show that this holds for odd cycles without exception. But for later reference, we start with discussing properties of ortho-homomorphisms of paths.

Let P_k denote the path of length k. As we have seen (Example 1), the ortho-homomorphism measure of paths is a probability distribution. Let η_d^k denote the marginal of this distribution on the pair of endpoints ($k \ge 1$). In particular, $\eta_d^1 = \eta_d$. Easy properties of η_d^k are summarized in the next lemma.

Lemma 19 *The distribution η_d^k is absolutely continuous with respect to π^2 for all $d \ge 3$ and $k \ge 2$. The density function $u_{d,k}(x,y) = (d\eta_d^k)/d\pi^2(x,y)$ is continuous for $k \le 5$ if $d = 3$ and for $k \ge 3$ if $d \ge 4$. For $k = 2$, it has a singularity when $x \| y$; for $d = 3$ and $k = 3$, it has a singularity when $x \perp y$; for $d = 3$ and $k = 4$, it has a singularity when $x \| y$.*

Proof. By Lemma 5,

$$u_{d,2}(x,y) = s_{d,2}(x,y) = \frac{A_{d-1}A_{d-3}}{A_{d-2}^2}\frac{1}{\sin(\measuredangle(x,y))},$$

from which statements for $k=2$ are easily verified. It is easy to check that

$$u_{d,k+m}(x,y) = \int\limits_{S^{d-1}} u_{d,k}(x,z)u_{d,m}(z,y)\,d\pi(z) \tag{32}$$

for $k,m \geq 2$, and

$$u_{d,k+1}(x,y) = \int\limits_{S^{d-1}\cap x^\perp} s_{d,k}(z,y)\,d\pi_0(z), \tag{33}$$

where π_0 is the uniform distribution on the $(d-2)$-dimensional sphere $x^\perp \cap S^{d-1}$. Using this formula, we see that $u_{d,3}$ is a continuous function for non-orthogonal pairs of points (x,y), and it is not hard to check that if $\varepsilon = \measuredangle(x,y) - \pi/2 \to 0$, then

$$u_{d,3}(x,y) = \begin{cases} O(\log\varepsilon) & \text{if } d=3, \\ O(1), & \text{if } d>3. \end{cases}$$

From this it follows that $u_{d,k}$ is bounded (even continuous) for all $k \geq 3$ if $d \geq 4$. If $d=3$, then (32) implies that $u_{d,4}$ still has singularity if $x=y$; for $\varepsilon = \measuredangle(x,y) \to 0$, we have $u_{d,4}(x,y) = O(\log\varepsilon)$. Using (32) again, we see that $u_{d,k}$ is bounded and continuous on $S^{d-1}\times S^{d-1}$ for all $k \geq 5$. □

Theorem 20 *The ortho-homomorphism density $t(C_k,d)$ is finite except if $d=3$ and $k=4$.*

Proof. For even cycles longer than 4 we already know this by Theorem 16; also for $k=3$, by the computations of Example 2. Let $k=2r+1$, $r\geq 2$. Then, using any ordering of the nodes, we get that

$$t(C_{2r+1},d) = \int\limits_{S^{d-1}} u_{d,2}u_{d,2r-1}\,d\pi^2.$$

Here $u_{d,2} < C_1$ on $S_1 = \{(x,y):\ \pi/4 \leq \measuredangle(x,y) \leq 3\pi/4\}$ and $u_{d,2r-1} < C_2$ on $S_2 = (S^{d-1})^2 \setminus S_1$, by Lemma 19, for some constants C_i. Thus

$$\begin{aligned} t(C_{2r+1},d) &= \int\limits_{S^{d-1}} u_{d,2}u_{d,2r-1}\,d\pi^2 \leq C_1\int\limits_{S_1} u_{d,2r-1}\,d\pi^2 + C_2\int\limits_{S_2} u_{d,2}\,d\pi^2 \\ &\leq C_1 t(P_2,d) + C_2 t(P_1,d), \end{aligned}$$

which is finite.

For $d = 3$ and $k = 4$, the graph C_4 does not satisfy the 3-sparsity condition, and indeed, as we have seen, the ortho-homomorphism measure has no natural definition. Formula (19) applies but the integral is infinite. □

These computations imply that for $k \geq 5$, $u_{d,k}$ is a continuous function on $S^{d-1} \times S^{d-1}$, so $u_{d,k}(x,x)$ is well defined, and

$$t(C_k, d) = \int_{S^{d-1}} u_{d,k}(x,x)\, d\pi(x). \tag{34}$$

More explicit formulas for these densities will be given in Section 5 based on the spectrum of the graphop $\mathbf{A}$.

4.4 Crowns

For $n \geq 4$, we define the *n-crown* Cr_n as the bipartite graph with bipartition $U \cup W$, where $U = \{u_1, \dots, u_n\}$, $W = \{w_1, \dots, w_n\}$, and w_i is connected to u_{i-1}, u_i, and u_{i+1} (subscripts modulo n). The 4-crown is the skeleton of the 3-dimensional cube. For odd n, the n-crown is also known as the "Möbius ladder", for even n, as the "prism". The n-crown is 4-sparse if $n \geq 4$.

Proposition 21 (a) *If $d \geq 4$, $n \geq 4$ and $(d,n) \notin \{(4,4),(4,5),(4,6),(5,4)\}$, then $t(\mathrm{Cr}_n, d)$ is finite.* (b) *$t(\mathrm{Cr}_4, 4)$ is infinite.*

Proof. (a) Let $x_i = x_{u_i}$ be independent random points of S^{d-1}. Let α_i be the angle between x_i and x_{i+1}, and let ϑ_i be the angle between the planes $\mathrm{lin}(x_{i-1}, x_i)$ and $\mathrm{lin}(x_i, x_{i+1})$. Let $Y_i = \sin\alpha_i = \mathrm{Det}(x_i, x_{i+1})$, $Z_i = \sin\vartheta_i$, $D_i = \mathrm{Det}(x_{i-1}, x_i, x_{i+1}) = Y_{i-1}Y_iZ_i$ and $W = Y_1^2 \dots Y_n^2 Z_1 \dots Z_n$. Then

$$t(G,d) = C\mathsf{E}(W^{-1}),$$

where the constant C is computable by (19), but we don't need this here. Let B_i denote the event that $D_iD_{i+1} \geq D_jD_{j+1}$ for all $j = 1, \dots, n$. Clearly $\mathsf{P}(B_i) = 1/n$ and $\mathsf{E}(W^{-1} \mid B_i)$ is independent of i, which implies that $\mathsf{E}(W^{-1} \mid B_i) = \mathsf{E}(W^{-1})$.

Assume that B_n occurs, and let $W_0 = D_2 \cdots D_{n-1}$. Then $D_nD_1 \geq W^{2/n}$, and hence $W_0 = W/(D_1D_2) \leq W^{(n-2)/n}$. Thus $W \geq W_0^{n/(n-2)}$. Hence

$$\begin{aligned}\mathsf{E}(W^{-1}) = \mathsf{E}(W^{-1} \mid B_n) \leq \mathsf{E}\Big(W_0^{-n/(n-2)} \mid B_n\Big) &= \frac{\mathsf{E}\Big(W_0^{-n/(n-2)}\mathbb{1}_{B_n}\Big)}{\mathsf{P}(B_n)} \\ = n\mathsf{E}\Big(W_0^{-n/(n-2)}\mathbb{1}_{B_n}\Big) &\leq n\mathsf{E}\Big(W_0^{-n/(n-2)}\Big).\end{aligned}$$

The advantage of considering W_0 is that we can write it as

$$W_0 = Y_1Y_2^2 \cdots Y_{n-2}^2 Y_{n-1} Z_2 \cdots Z_{n-1},$$

and here all of the factors are independent random variables. So the expectation of $W_0^{-n/(n-2)}$ is finite if and only if $\mathsf{E}(Y_1^{-n/(n-2)})$, $\mathsf{E}(Y_i^{-2n/(n-2)})$ and $\mathsf{E}(Z_i^{-n/(n-2)})$ are finite. Clearly, finiteness of the second expectation implies finiteness of the first one.

The expectations of powers of Y_1 and Z_1 are easy to compute: the density function of (say) α_1 is proportional to $(\sin\alpha_1)^{d-2}$, and so

$$\mathsf{E}(Y_1^{-2n/(n-2)}) = \frac{\int_0^\pi (\sin\alpha)^{d-2-2n/(n-2)}\,d\alpha}{\int_0^\pi (\sin\alpha)^{d-2}\,d\alpha}.$$

The integral in the numerator is finite if the exponent of $\sin\alpha$ is larger than -1; this means that

$$d-2-2n/(n-2) > -1. \tag{35}$$

Similarly, the density function of Z_1 is proportional to $(\sin\vartheta_1)^{d-3}$ (the density of the angle between two random points on the equator), and hence

$$\mathsf{E}(Z_1^{-n/(n-2)}) = \frac{\int_0^\pi (\sin\alpha)^{d-3-n/(n-2)}\,d\alpha}{\int_0^\pi (\sin\alpha)^{d-2}\,d\alpha}.$$

As before, this is finite if and only if $d-3-n/(n-2) > -1$. It is not hard to see that (35) is stronger. Rewriting (35) as $(d-3)(n-2) > 4$, we see that this holds for $d=4$ and $n\ge 7$, $d=5$ and $n\ge 5$ and $d\ge 6$, $n\ge 4$.

(b) For any three unit vectors y_1, y_2, y_3, we have

$$|y_1\wedge y_3| + |y_2\wedge y_3| \ge |y_1\wedge y_2|.$$

For vectors of arbitrary length, this gives

$$|z_1\wedge z_3|\,|z_2| + |z_2\wedge z_3|\,|z_1| \ge |z_1\wedge z_2|\,|z_3|.$$

Applying this to the vectors $z_i = x_i/x_4$, we get

$$|x_1\wedge x_3\wedge x_4|\,|x_2\wedge x_4| + |x_2\wedge x_3\wedge x_4|\,|x_1\wedge x_4| \ge |x_1\wedge x_2\wedge x_4|\,|x_3\wedge x_4|.$$

Setting $Y_i = |x_i\wedge x_3\wedge x_4|/|x_3\wedge x_4|$ (this is the distance of x_i from the plane $\mathrm{lin}(x_3,x_4)$, we get from this

$$|x_1\wedge x_2\wedge x_4| \le Y_1|x_2\wedge x_4| + Y_2\,|x_1\wedge x_4| \le Y_1+Y_2.$$

The same upper bound can be given on $|x_1\wedge x_2\wedge x_3|$, and trivially

$$|x_i\wedge x_3\wedge x_4| = Y_i|x_3\wedge x_4| \le Y_i.$$

The denominator in (19) can be estimated as

$$|x_1\wedge x_2\wedge x_3|\,|x_1\wedge x_2\wedge x_4|\,|x_1\wedge x_3\wedge x_4|\,|x_2\wedge x_3\wedge x_4| \le (Y_1+Y_2)^2Y_1Y_2 \le 4\max(Y_1,Y_2)^4.$$

Note that the distributions of Y_1 and Y_2 do not depend on x_3 and x_4, and so we can fix $L=\mathrm{lin}(x_3,x_4)$. Similarly as in the proof of Lemma 3, let ϑ_i be the angle between x_i and L $(0\le\vartheta\le\pi/2)$, then $Y_i=\sin\vartheta_i$. For a fixed ϑ, points at this distance $\sin\vartheta$ from L form the direct product of two circles $L\cap(\cos\vartheta)S^1$ and $L^\perp\cap(\sin\vartheta)S^1$, and so their density is proportional to $\cos\vartheta\sin\vartheta$. Hence $\mathsf{E}\big(\max(Y_1,Y_2)^{-4}\big)$ is proportional to

$$\int\limits_{[0,\pi/2]^2}\frac{\sin\vartheta_1\cos\vartheta_1\sin\vartheta_2\cos\vartheta_2}{\max(\sin\vartheta_1,\sin\vartheta_2)^4}\,d\vartheta_2\,d\vartheta_1 = 2\int\limits_{\vartheta_2\le\vartheta_1}\frac{\cos\vartheta_1\sin\vartheta_2\cos\vartheta_2}{(\sin\vartheta_1)^3}\,d\vartheta_2\,d\vartheta_1$$
$$=\int\limits_{[0,\pi/2]}\frac{\cos\vartheta_1(\sin\vartheta_1)^2}{(\sin\vartheta_1)^3}\,d\vartheta_1=\int\limits_{[0,\pi/2]}\frac{\cos\vartheta_1}{\sin\vartheta_1}\,d\vartheta_1,$$

which is infinite. □

5 Spectral formulas

5.1 Powers of the graphop

Let $\mathcal{H}_d$ denote the function space $L^2(S^{d-1},\pi)$ where π is the uniform measure on S^{d-1}. If Q is an element in the orthogonal group $\mathsf{O}(d)$, then it also acts naturally on $\mathcal{H}_d$ by $(fQ)(x)=f(Q(x))$ where $f\in\mathcal{H}_d$ and $x\in S^{d-1}$. We say that a linear operator $\mathbf{T}$ on $\mathcal{H}_d$ is *rotation invariant*, if $Q\mathbf{T}Q^{-1}=\mathbf{T}$ holds for every $Q\in\mathsf{O}(d)$.

Under the general correspondence between measures and linear operators, we can define a bounded linear operator $\mathbf{A}=\mathbf{A}_d:\ \mathcal{H}_d\to\mathcal{H}_d$ by letting $(\mathbf{A}_df)(x)$ be the average of f on the $(d-2)$-subsphere $x^\perp\cap S^{d-1}$. (It is not hard to see that this is well-defined for almost all x; see [2].)

This operator satisfies

$$\langle\mathbf{A}f,g\rangle=\int\limits_{S^{d-1}\times S^{d-1}} f(x)g(y)\,d\eta(x,y). \tag{36}$$

for every $f\in L_p(S^{d-1},\pi)$ and $g\in L_q(S^{d-1},\pi)$ (see [2]). This implies that it is self-adjoint. It is trivial that $\mathbf{A}$ is *monotone*: if $f\ge0$ then $\mathbf{A}f\ge0$. We also note that $\mathbf{A}$ is 1-regular: $\mathbf{A}\mathbb{1}_{S^{d-1}}=\mathbb{1}_{S^{d-1}}$. This operator also has the geometric property that it is *rotation invariant*, i.e.,

We say that an operator $\mathbf{T}:\ L_2(S^{d-1})\to L_2(S^{d-1})$ is *represented* by a measurable function $u:\ S^{d-1}\to\mathbb{R}$ if

$$(\mathbf{T}f)(x)=\int\limits_{S^{d-1}} u(x,y)f(y)\,d\pi(y).$$

Then $\mathbf{T}$ is a Hilbert–Schmidt integral operator. The operator $\mathbf{A}$ cannot be represented by any function, but its higher powers can: the operator $\mathbf{A}^k$ is represented by the function $u_{d,k}$ for all $k \geq 2$.

5.2 Spherical harmonics

In this section we study the spectrum of the orthogonality operator $\mathbf{A} = \mathbf{A}_d$ for a fixed $d \geq 3$. As an application, we obtain formulas for the ortho-homomorphism densities $t(C_k, d)$ in the next section.

The fact that $\mathbf{A}^k$ is a Hilbert–Schmidt operator for $k \geq 2$ implies that $\mathbf{A}$ is a compact operator. Let λ_n ($n = 0, 1, 2, \dots$) be the distinct nonzero eigenvalues of $\mathbf{A}$, and let $\mathbf{T}_n$ denote the orthogonal projection onto the eigensubspace W_n belonging to λ_n. Then the expansion

$$\mathbf{A}^k = \sum_{n=0}^{\infty} \lambda_n^k \mathbf{T}_n \tag{37}$$

is convergent in operator norm for $d = 3$ and $k \geq 5$. Our goal is to give more explicit formulas for λ_n and W_n.

It is well-known that the action of the orthogonal group $\mathsf{O}(d)$ on the Hilbert space $\mathcal{H}_d = L^2(S^{d-1})$ has a unique decomposition into distinct irreducible representations. These representations are carried by subspaces $W_0, W_1, W_2, \dots$ of $L^2(S^{d-1})$, where W_n consists of polynomials of degree n and has dimension

$$\dim(W_n) = \binom{d+n-1}{d-1} - \binom{d+n-3}{d-1}. \tag{38}$$

Since the operator $\mathbf{A}$ is rotation invariant, standard arguments show that each eigenspace of $\mathbf{A}$ is invariant under $\mathsf{O}(d)$. Hence each W_n is contained in one of the eigenspaces of $\mathbf{A}$ and thus elements in W_n are eigenvectors of $\mathbf{A}$ with identical eigenvalue λ_n.

The Gegenbauer polynomials $C_n^{(\alpha)}(x)$ (also called ultraspherical polynomials) are orthogonal polynomials on $[-1, 1]$ with respect to the weight function $(1 - x^2)^{\alpha - 1/2}$. (In particular, $C_n^{(1/2)}(x)$ is the n-th Legendre polynomial.) The significance of these polynomials for us is that if $\alpha = d/2 - 1$, $n \in \mathbb{N}$ and $y \in \mathbb{R}^d$ is a fixed unit vector, then the function $x \mapsto C_n^{(\alpha)}(x \cdot y)$ defined for $x \in S^{d-1}$ is an eigenfunction of the operator $\mathbf{A}$ (called a *zonal spherical harmonic function*). Furthermore, the corresponding eigenvalues (with appropriate multiplicities) describe all the eigenvalues of $\mathbf{A}$.

It is not hard to calculate the eigenvalues corresponding to these functions. It is clear that $f_n(y) = C_n^{(\alpha)}(1)$ and that $(\mathbf{A} f_n)(y) = C_n^{(\alpha)}(0)$, and so the eigenvalue is $C_n^{(\alpha)}(0)/C_n^{(\alpha)}(1)$. Fortunately, these special values of the Gegenbauer polynomials are easily derived from the classical series expansion [6]

$$\sum_{n=0}^{\infty} C_n^{(\alpha)}(x) y^n = \frac{1}{(1 - 2xy + y^2)^\alpha}. \tag{39}$$

Substituting $x = 0$ and $x = 1$, we get

$$C_n^{(\alpha)}(0) = \begin{cases} (-1)^r \binom{r+\alpha-1}{r}, & \text{if } n = 2r \text{ is even,} \\ 0, & \text{if } n \text{ is odd,} \end{cases} \tag{40}$$

and

$$C_n^{(\alpha)}(1) = \binom{n+2\alpha-1}{n}. \tag{41}$$

In our case when $\alpha = d/2 - 1$, both quantities are rational numbers. From these formulas we obtain that the eigenvalue λ_n of $\mathbf{A}$ corresponding to n-th zonal harmonic function is

$$\lambda_n = \frac{C_n^{(d/2-1)}(0)}{C_n^{(d/2-1)}(1)} = \begin{cases} (-1)^r \dfrac{(d-3)!!\,(2r-1)!!}{(2r+d-3)!!}, & \text{if } n = 2r \text{ is even,} \\ 0, & \text{if } n \text{ is odd.} \end{cases} \tag{42}$$

For even d, the formula for λ_n can be simplified:

$$\lambda_n = \begin{cases} (-1)^{n/2} \dfrac{(d-3)!!}{(n+1)(n+3)\dots(n+d-3)}, & \text{if } n \text{ is even,} \\ 0, & \text{if } n \text{ is odd.} \end{cases} \tag{43}$$

Note that in this case the numerator is a constant (we consider d fixed), and the denominator is a polynomial in n. If $d = 4$, then $\lambda_n = (-1)^{n/2}/(n+1)$ for even n and $\lambda_n = 0$ for odd n.

The projections $\mathbf{T}_n$ to these subspaces can be described as well. The fact that W_n is finite dimensional implies that $\mathbf{T}_n$ is an integral kernel operator representable by some measurable function $Q_n : S^{d-1} \times S^{d-1} \to \mathbb{R}$. Furthermore, since W_n is an eigenspace of $\mathbf{A}$, each Q_n is invariant under the natural action of the orthogonal group $\mathsf{O}(d)$. This implies $Q_n(x, y)$ depends on the scalar product of x and y only. In other words, there is a measurable function $f_n : [-1, 1] \to \mathbb{R}$ such that $Q_n(x, y) = f_n(x \cdot y)$. This also means the for every fixed $y \in S^{d-1}$, the map $x \mapsto f_n(x \cdot y)$ is in W_n and thus these functions are the zonal spherical harmonic functions. We obtain that $f_n(x) = C_n^{(\alpha)}(x) c_n$ for some constant c_n where $\alpha = (d-2)/2$. The fact that Q_n represents a projection onto W_n implies that $\dim(W_n) = \mathbb{E}_x Q_n(x, x) = C_n^{(\alpha)}(1) c_n$ and thus $c_n = \dim(W_n)/C_n^{(\alpha)}(1)$. So

$$f_n(x) = \dim(W_n) C_n^{(\alpha)}(x)/C_n^{(\alpha)}(1),$$

and for $x \in S^{d-1}$ and $f \in \mathcal{H}_d$,

$$(\mathbf{T}_n f)(x) = \int_{S^{d-1}} f_n(x \cdot y) f(y)\, d\pi(y). \tag{44}$$

We can apply (37) (formally) for $k = 1$:

$$\mathbf{A}f(x) = \sum_{n=0}^{\infty} \lambda_n (\mathbf{T}_n f)(x) = \sum_{n=0}^{\infty} \lambda_n \int_{S^{d-1}} f_n(x \cdot y) f(y)\, d\pi(y). \tag{45}$$

It is not clear when this infinite sum converges.

5.3 Cycle densities

We start with the expansion

$$t(C_k, d) = \operatorname{tr}(\mathbf{A}_d^k) = \sum_{n=0}^{\infty} \lambda_n^k \dim(W_n), \tag{46}$$

convergent for $d = 3$ and $k \geq 5$, and for $d \geq 4$ and $k \geq 4$. Substituting values computed above, we get the formula

$$t(C_k, d) = \sum_{r=0}^{\infty} \left(\binom{d+2r-1}{d-1} - \binom{d+2r-3}{d-1} \right) \left((-1)^r \frac{(d-3)!!(2r-1)!!}{(2r+d-3)!!} \right)^k. \tag{47}$$

If $d = 4$, we obtain much nicer formulas:

$$t(C_k, 4) = \sum_{r=0}^{\infty} (2r+1)^{2-k} = \zeta(k-2)(1 - 2^{2-k})$$

if k is even and

$$t(C_k, 4) = \sum_{r=0}^{\infty} (2r+1)^{2-k} (-1)^r$$

if k is odd. In particular, $t(C_4, 4) = \pi^2/8$. The case of a triangle is interesting: the formula specializes to

$$t(C_3, 4) = 1 - \frac{1}{3} + \frac{1}{5} - \cdots = \frac{\pi}{4}.$$

The series is not absolute convergent, and we have no good argument to justify the order in which it is summed; but the computations in Example 2 show that this is the "right" order.

If $d = 3$, then the eigenvalues of $\mathbf{A}$ are $(-1/4)^r \binom{2r}{r}$ with multiplicity $4r + 1$ for $r = 0, 1, 2, \ldots$. This leads to

$$t(C_k, 3) = \sum_{r=0}^{\infty} (4r+1)(-1/4)^{rk} \binom{2r}{r}^k.$$

Comparing with (25), we get the identity

$$\int\limits_{[0,\pi]^3} \frac{d\alpha\, d\beta\, d\theta}{\sqrt{1 - \big(\cos\alpha \cos\beta - \sin\alpha \sin\beta \cos\theta\big)^2}} = \frac{\pi^4}{2} \sum_{r=0}^{\infty} (4r+1) 4^{-6r} \binom{2r}{r}^6. \tag{48}$$

6 Approximations by graphons and graphs

In this chapter we explain how the orthogonality graphs H_d and ortho-homomorphism densities fit into graph limit theory. Our goal is to find sequences of graphons and finite graphs which approximate H_d (or more precisely the operator $\mathbf{A}_d$) in the sense that ortho-homomorphism densities become limits of classical subgraph densities. As a consequence we obtain that ortho-homomorphism densities behave a lot like subgraph densities. They satisfy a variety of inequalities that are known in the graph theoretic framework. A very interesting example is Sidorenko's conjecture, which has been proved for quite a few classes of graphs. The ortho-homomorphism version of this conjecture is especially nice: It says that the ortho-homomorphism density of any bipartite graph in H_d (for $d \geq 3$) is at least 1. Our results in this chapter will imply this for every bipartite graph that satisfies the finite version of the conjecture.

For $x \in S^{d-1}$ and $0 < r < \pi$, let $B_r(x)$ be the set of points $y \in S^{d-1}$ such that $d(x,y) < r$ (where d is the spherical distance), and let $V_r = \pi(B_r(x))$. Let $f_{x,r} = \mathbb{1}_{B_{x,r}}$ be the indicator function of $B_{r,x}$.

Lemma 22 *If A is a rotation invariant bounded operator on $\mathcal{H}$, then A is self-adjoint, i.e., $A^* = A$. Any two rotation invariant bounded operators commute.*

Proof. For any two points $x, y \in S^{d-1}$ and $0 < r < \pi$, there is a reflection $R \in \mathsf{O}(d)$ in a hyperplane such that $R(x) = y$ and $R(y) = x$. For this reflection we have $f_{x,r} = f_{y,r}R$ and $f_{y,r} = f_{x,r}R$ for every $r > 0$. It follows that

$$\langle f_{x,r}A, f_{y,r}\rangle = \langle f_{x,r}RAR, f_{y,r}\rangle = \langle f_{x,r}RA, f_{y,r}R\rangle = \langle f_{y,r}A, f_{x,r}\rangle. \tag{49}$$

For $r > 0$ let $K_r := \{f_{x,r} : x \in S^{d-1}\}$ and let W_r denote the space of finite linear combinations of elements in K_r. From equation (49) and the bilinearity of the scalar product, we obtain that $\langle fA, g\rangle = \langle f, gA\rangle$ holds for any two functions $f, g \in W_r$.

Now let f, g be arbitrary functions in $\mathcal{H}_d$. It is easy to see that for every $\epsilon > 0$ there is an $r > 0$ and two functions $f', g' \in W_r$ such that $\|f - f'\|_2 < \epsilon$, $\|g - g'\|_2 < \epsilon$. Then

$$|\langle f, gA\rangle - \langle f', g'A\rangle \leq |\langle f - f', gA\rangle| + |\langle f', (g - g')A\rangle| < \epsilon\|g\|_2\|A\|_2 + (\|f\|_2 + \epsilon)\epsilon\|A\|_2$$

and similarly $|\langle fA, g\rangle - \langle f'A, g'\rangle| < \epsilon\|A\|_2(\|f\|_2 + \|g\|_2 + \epsilon)$. From $\langle f', g'A\rangle = \langle f'A, g'\rangle$ and with $\epsilon \to 0$ we obtain that $\langle f, gA\rangle = \langle fA, g\rangle$, showing that $A^* = A$.

To show the second claim, let A, B be bounded rotation invariant operators. Then AB is also rotation invariant and so $AB = (AB)^* = B^*A^* = BA$ using the first statement. $\square$

Next we introduce a set of operators $\mathbf{M}_r$ on $\mathcal{H}_d$ defined by

$$(\mathbf{M}_r f)(x) = \frac{1}{V_r} \int\limits_{B_r(x)} f(y)\, d\pi(y).$$

It is clear that $\mathbf{M}_r$ is rotation invariant, so by Lemma 22 $\mathbf{M}_r$ and $\mathbf{A}_d$ commute and the product $\mathbf{C}_r := \mathbf{A}_d\mathbf{M}_r$ is a self-adjoint operator on $\mathcal{H}_d$.

The operator $\mathbf{C}_r$ is a Hilbert–Schmidt operator with a nonnegative, symmetric, bounded, measurable kernel $W_r:\ S^{d-1}\times S^{d-1}\to\mathbb{R}$:

$$(\mathbf{C}_r f)(x) = \int\limits_{S^{d-1}} W_r(x,y)f(y)\,d\pi(y).$$

It is easy to see that for a fixed x, $W_r(x,y)$ is the density function of the random point y obtained by moving from x in a random direction by $\pi/2$ to a point x', and then moving to a uniform random point of $B_r(x')$.

Clearly $\int_{S^{d-1}} W_r(x,y)\,d\pi(y) = 1$ for every x, so W_r is a 1-regular graphon, and $t(H,W_r)$ is well-defined by (7). Our main goal in this chapter is to prove that for a class of bipartite graphs H we have

$$t(H,d) = \lim_{r\to 0} t(H,W_r) \tag{50}$$

Our main tool is a rather explicit formula for the value of $t(H,W_r)$.

Lemma 23 *For every $d\ge 3$ and $n\in\mathbb{N}$,*

$$\int\limits_{S^{d-1}} \prod_{i=1}^{n} W_r(z,x_i)\ d\pi(z) = \frac{A_{d-1}^{n-1}A_{d-n-1}}{A_{d-2}^{n}} D_r(x_1,x_2,\dots,x_n). \tag{51}$$

Proof. Let z be a uniform random point on S^{d-1} and let $z_1,z_2,\dots,z_n$ be independent uniform elements on S^{d-1} orthogonal to z. Let x_i be chosen uniformly from $B_r(z_i)$. By (8), the density function of the joint distribution of $(x_1,\dots,x_n)$ is just the function on the right hand side of (51). On the other hand, by (5) the joint distribution of $(z_1,\dots,z_n)$ has density function

$$\frac{A_{d-1}^{n-1}A_{d-n-1}}{A_{d-2}^{n}} D(z_1,x_2,\dots,z_n). \tag{52}$$

Since $(x_1,\dots,x_n)$ is a random point in $B_r(z_1)\times\cdots\times B_r(z_n)$, the density function of $(x_1,\dots,x_n)$ is the average of (52) on $B_r(z_1)\times\cdots\times B_r(z_n)$. □

Lemma 24 *Let $G=(V,E)$ be a d-sparse bipartite graph ($d\ge 3$). Then*

$$t(G,W_r) = \frac{A_{d-1}^{|E|-|W|}}{A_{d-2}^{|E|}} \int\limits_{(S^{d-1})^U} \prod_{v\in W} A_{d-d(v)-1} D_r(x(N(v)))\,d\pi^U(x). \tag{53}$$

Proof. Let $U \cup W$ be a bipartition of V, then using (51),

$$\begin{aligned} t(G, W_r) &= \int\limits_{(S^{d-1})^V} \prod_{i \in W, j \in N(i)} W(x_i, x_j)\, d\pi^V(x) \\ &= \int\limits_{(S^{d-1})^U} \prod_{i \in W} \left(\int\limits_{S^{d-1}} \prod_{j \in N(i)} W(x_i, x_j)\, d\pi(x_j) \right) d\pi^U \\ &= \int\limits_{(S^{d-1})^U} \prod_{i \in W} \left(\frac{A_{d-1}^{n-1} A_{d-n-1}}{A_{d-2}^n} D_r(x_1, x_2, \ldots, x_n) \right) d\pi^U. \end{aligned}$$

Simplifying, we get (53). □

Theorem 25 *If G is a bipartite graph that satisfies the sparsity condition, and $t(G, d) < \infty$, then*

$$t(G, d) = \lim_{r \to 0} t(G, W_r).$$

Proof. According to Lemma 24 and the formula (19), it is enough to prove that

$$\lim_{r \to 0} \int\limits_{(S^{d-1})^U} \prod_{v \in W} D_r(x(N(v)))\, d\pi^U(x) = \int\limits_{(S^{d-1})^U} \prod_{v \in W} D(x(N(v)))\, d\pi^U(x). \quad (54)$$

Let

$$\widehat{D}_r(x) = \prod_{v \in W} D_r(x(N(v))) \quad \text{and} \quad \widehat{D}(x) = \prod_{v \in W} D(x(N(v))).$$

It is clear that $\widehat{D}_r(x) \to \widehat{D}(x)$ as $r \to 0$ for almost all $x \in (S^{d-1})^U$. By Lemma 4 we have that $\widehat{D}_r(x) \le C_d \widehat{D}(x)$ for some $c > 0$ independent from r and x. Since $t(G, d)$ is finite, the function $\widehat{D}$ is integrable, and so $c\widehat{D}$ is an integrable upper bound on $\widehat{D}_r$. Thus (54) follows by Lebesgue's Dominated Convergence Theorem. □

The next theorem is a corollary of Theorem 25.

Theorem 26 *For every $d \ge 3$ there is a sequence of finite graphs $\{G_i\}_{i=1}^\infty$ such that if a finite bipartite graph H satisfies the sparsity condition, then*

$$\lim_{i \to \infty} t(H, G_i)/t(e, G_i)^{|E(H)|} = t(H, d).$$

Proof. For $n \in \mathbb{N}$ let $U_n := W_{d,1/n}/\|W_{d,1/n}\|_\infty$, then U_n is a symmetric measurable function with values in $[0, 1]$. It follows from the results in [14] that there is a finite graph G_n such that

$$\left| t(H, G_n)/t(e, G_n)^{|E(H)|} - t(H, U_n)/t(e, U_n)^{|E(H)|} \right| \le 1/n.$$

Since $t(e, W_{d,1/n}) = 1$, we also have that $t(H, U_n)/t(e, U_n)^{|E(H)|} = t(H, W_{d,1/n})$. Together with Theorem 25, this completes the proof. $\square$

Theorem 26 shows that in some sense H_d is a limit of finite graphs. It is interesting to mention that the sequence $\{G_i\}_{i=1}^{\infty}$ given by the proof of the theorem is a sparse graph sequence. We also have an interesting corollary of Theorem 26.

Corollary 27 *If H is a d-sparse bipartite graph that satisfies Sidorenko's conjecture, then $t(H, d) \geq 1$.*

Sidorenko's conjecture is verified for large families of bipartite graphs, and thus Corollary 27 implies several non-trivial inequalities for ortho-homomorphism densities. Some other graph theoretic inequalities can also be transported to ortho-homomorphism densities with the help of Theorem 26; but we omit the details here.

7 Open problems

Let us conclude with some special and more general problems left open by our work.

Problem 28 Decide the finiteness of $t(\mathrm{Cr}_n, d)$ the open cases $(d, n) \in \{(4, 5), (4, 6), (5, 4)\}$ in Proposition 21.

Problem 29 Characterize graphs G and dimensions d for which $t(G, d)$ is finite. As an interesting example: if G is the incidence graph of the Fano plane, is $t(G, 4)$ finite?

Problem 30 The fact that the cube graph Cr_4 is, in a sense, exceptional among crowns, may be related to the fact that for 4-sparse graphs, the real algebraic variety of all ortho-homomorphisms in dimension 4 is irreducible, except for the cube. Is there a more substantial connection?

Problem 31 Make sense of the identity (48), perhaps generalized to all cycles and all dimensions.

Problem 32 Let G be a d-sparse graph, and let μ be a probability measure on $\Sigma_{G,d}$ with Markovian conditioning. Is μ uniquely determined by G?

Problem 33 Are there natural graph sequences converging to the orthogonality graph? The orthogonality graph $H_{p,d}$ of $\mathbb{F}_p^d \setminus \{0\}$ (more exactly, the conjugacy graph in the projective space $\mathbb{P}_p^{d-1}$) is a natural example, but it does not work: cf. [2], Section 12.5, from which it follows that conjugacy graphs of finite projective spaces tend to a trivial limit in the sense of action convergence (a form of right convergence). From the other side, it is easy to compute that $t^*(K_3, H_{p,3}) = (p^2+p+1)/(p+1)^2 \sim 1$, while we have seen that $t(K_3, H_3) = 2/\pi$, showing that $H_{p,3}$ does not tend to H_3 in the local sense either.

Problem 34 Instead of random unit vectors, we could consider other probability distributions; Gaussian would be a natural choice. In the sequentially constructed random map, we map each node v onto a random vector from the standard Gaussian distribution on the subspace orthogonal to the previously chosen images of neighbors of v. We expect that a density function making this mapping independent of the order of the nodes can be constructed along the same lines as in this paper. This construction may have even nicer properties than our random ortho-homomorphism; but this is not discussed in this paper.

As another natural generalization, we could determine subgraph densities in the uniform measure on pairs of points of a unit sphere at any given distance (different from $\pi/2$). Even more generally, perhaps the methods above can be applied to any probability measure on pairs of points in $\mathbb{R}^d$ invariant under the orthogonal group.

Problem 35 Based on (44) and (45), one can (formally) derive the following formula:

$$t(G,d) = \sum_{\tau\colon E(H)\to\mathbb{N}} \int_{(S^{d-1})^V} \prod_{ij\in E} \lambda_{\tau(ij)} f_{\tau(ij)}(x_i \cdot x_j)\, d\pi^V(x). \tag{55}$$

Note that the product in the formula is a multivariate polynomial on $\mathbb{R}^{dn}$ with rational coefficients which depends on the edge labeling τ. It is not clear when this infinite sum converges and when the equality holds.

References

[1] A.Y. Alfakih: Graph connectivity and universal rigidity of bar frameworks, *Discr. Applied Math.* **217** (2017), 707–710.

[2] Á. Backhausz, B. Szegedy: Action convergence of operators and graphs, https://arxiv.org/abs/1811.00626

[3] C. Borgs, J. Chayes, H. Cohn and Y. Zhao: An L^p theory of sparse graph convergence I: Limits, sparse random graph models, and power law distributions, *Trans. Amer. Math. Soc.* **372** (2019), 3019–3062.

[4] A. Cabello, S. Severini, A. Winter: Graph-Theoretic Approach to Quantum Correlations, *Phys. Rev. Lett.* **112** (2014), 040401.

[5] H. Finner: A generalization of Hölder's inequality and some probability inequalities, *Ann. Probab.* **20** (1992), 1893–1901.

[6] L. Gegenbauer: Über einige bestimmte Integrale, *Sitzungsber. Kaiser. Akademie der Wissenschaften. Math.-Naturwiss.* **70** (1874), 433–443.

[7] M. Grötschel, L. Lovász and A. Schrijver: The ellipsoid method and its consequences in combinatorial optimization, *Combinatorica* **1** (1981), 169-197.

[8] M. Grötschel, L. Lovász, A. Schrijver: Geometric Algorithms and Combinatorial Optimization, Springer (1988).

[9] D. Kunszenti-Kovács, L. Lovász, B. Szegedy: Measures on the square as sparse graph limits, *J. Combin. Theory B* (2019), online:
https://www.sciencedirect.com/science/article/pii/S009589561930005X

[10] L. Lovász: On the Shannon capacity of graphs, *IEEE Trans. Inform. Theory* **25** (1979), 1–7.

[11] L. Lovász: *Large networks and graph limits*, Amer. Math. Soc., Providence, RI (2012).

[12] L. Lovász: *Graphs and Geometry*, Amer. Math. Soc. Coll. Publ. **65**, Providence, R.I. (2019).

[13] L. Lovász, M. Saks and A. Schrijver: : Orthogonal representations and connectivity of graphs, *Linear Algebra Appl.* **114/115** (1989), 439–454. A correction: *Linear Algebra Appl.* **313** (2000), 101–105.

[14] L. Lovász, B. Szegedy: Limits of dense graph sequences, *J. Combin. Theory B* **96** (2006), 933–957.