INTERVAL GRAPH LIMITS

PERSI DIACONIS, SUSAN HOLMES, AND SVANTE JANSON

Abstract. We work out the graph limit theory for dense interval graphs. The theory developed departs from the usual description of a graph limit as a symmetric function $W(x, y)$ on the unit square, with x and y uniform on the interval $(0, 1)$. Instead, we fix a W and change the underlying distribution of the coordinates x and y. We find choices such that our limits are continuous. Connections to random interval graphs are given, including some examples. We also show a continuity result for the chromatic number and clique number of interval graphs. Some results on uniqueness of the limit description are given for general graph limits.

1. Introduction

A graph G is an *interval graph* if there exists a collection of intervals $\{I_i\}_{i\in V(G)}$ such that there is an edge $ij \in E(G)$ if and only if $I_i \cap I_j \neq 0$, for all pairs $(i, j) \in V(G)^2$ with $i \neq j$.

Example 1.1. Figure 1(a) shows published confidence intervals for the astronomical unit (roughly the length of the semi-major axis of the earth's elliptical orbit about the sun). Figure 1(b) shows the corresponding interval graph (data from Youden [43]).

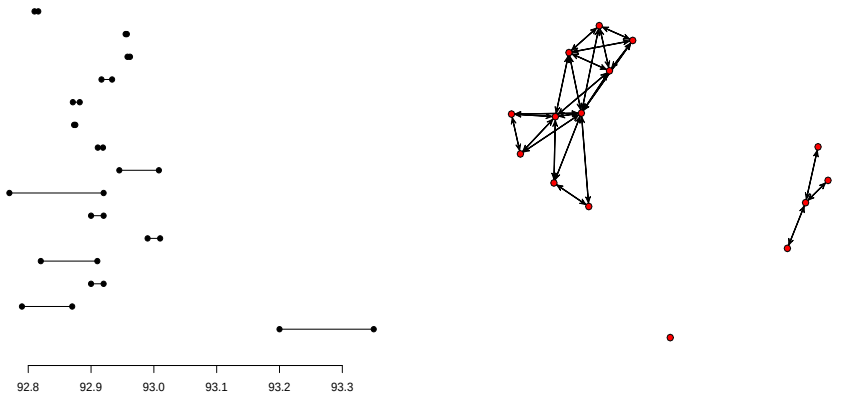
It is surprising how many missing edges there are in this graph as these correspond to disjoint confidence intervals for this basic unit of astronomy. Even in the large component, the biggest clique only has size 4.

The literature on interval graphs and further examples are given in Section 2.1–2.3 below. Section 2.4 reviews the emerging literature on graph limits. Roughly, a sequence of graphs G_n is said to converge if the proportion of edges, triangles and other small subgraphs tends to a limit. The limiting object is not usually a graph but is represented as a symmetric function $W(s, t)$ and a probability measure μ on a space $\mathcal{S}$. Again roughly $W(s, t)$ is the chance that the limiting graph has an edge from s to t, more details will be provided in Section 2.4.

The main results in this paper combine these two sets of ideas and work out the graph limit theory for interval graphs. The intervals in the definition above may be arbitrary intervals of real numbers $[a, b]$, that without loss can be considered inside $[0, 1]$. Thus an interval can be identified with a point

Date: February 11, 2011.

2000 *Mathematics Subject Classification.* 05C99.



(A) Confidence Intervals (B) Interval Graph

FIGURE 1. Building the interval graph for the Youden astronomical constant confidence intervals [43].

in the triangle $\mathcal{S} := \{[a,b] : 0 \le a \le b \le 1\}$, see Figure 2. An interval graph G_n is defined by a set of intervals $\{[a_i,b_i]\}_{i=1}^n$, which may be identified with the empirical measure $\mu_n = \frac{1}{n}\sum \delta_{(a_i,b_i)}$. In Section 3 we show that a sequence of graphs G_n converges if the empirical measures μ_n converge to a limiting probability μ in the usual weak star topology, provided μ satisfies a technical condition which we show may be assumed. The limit of the graphs is specified by a function W defined by

$$W(a,b;a',b') := \begin{cases} 1 & \text{if } [a,b]\cap[a',b'] \neq \emptyset \\ 0 & \text{otherwise} \end{cases} \qquad \text{and the limiting } \mu.$$

We thus fix W and and simply vary μ with μ specifying the graph limit; this gives all interval graph limits, but note that several μ may give the same graph limit. With a naïve choice of μ, the assignment of μ to a graph limit is not usually continuous (as a map from probabilities on $\mathcal{S}$ to graph limits). We show that there are several natural choices of μ that lead to the same graph limit and result in continuous assignments.

The main theorem is stated in Section 3, and results on the chromatic number and clique number are given in Section 4. Some important preliminaries on continuity of the mapping $\mu \mapsto \Gamma_\mu$ are dealt with in Section 5, and Section 6 gives the proof of the main theorem. Section 7 discusses some examples of interval graph limits and the corresponding random interval graphs. The parametrization of graph limits is highly non unique; this is seen in some of the examples in Section 7. Section 8 gives a portemanteau theorem which clarifies the connections between various uniqueness results. This is developed for the general case, not just interval graphs. The problem of finding a unique "canonical" representing measure is still open in general. Section 9 gives the proofs of the results on clique numbers. Finally,

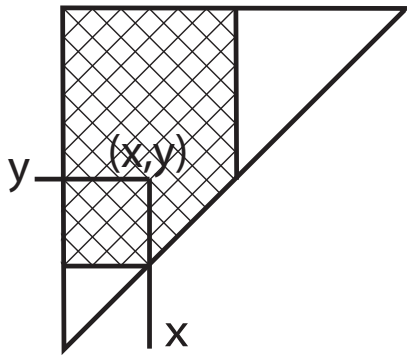


FIGURE 2. For a given (x, y) in $\mathcal{S}$, the relevant (x', y') that will give an edge in the intersection graph are in the hatched area.

Section 10 discusses extensions to other classes of intersection graphs, in particular circular-arc graphs, circle graphs, permutation graphs and unit interval graphs.

2. Background

This section gives background and references, treating interval graphs in Sections 2.1 and 2.2, random interval graphs in Section 2.3 and graph limits in Section 2.4–2.5.

2.1. **Interval Graphs.** Interval graphs and the closely associated subject of interval orders are a standard topic in combinatorics. A book length treatment of the subject is given by Fishburn [14]. Among many other results, we mention that interval graphs are perfect graphs, i.e., the chromatic number equals the size of the largest clique (for the graph and all induced subgraphs).

Interval graphs are a special case of intersection graphs; more generally, we may consider a collection $\mathcal{A}$ of subsets of some universe and the class of graphs that can be defined by replacing intervals by elements of $\mathcal{A}$ in the definition above. (We may call such graphs $\mathcal{A}$-intersection graphs.)

McKee and McMorris [32]'s book on Intersection Graphs establishes the relation with intervals. Further literature on the connections between these various graph classes is in Brandstädt, Le and Spinrad [9] and Golumbic [18].

2.2. **Applications of Interval Graphs.** The original question for which interval graphs saw their first application was in the structure of genetic DNA. Waterman and Griggs [42] and Klee [29] cite Benzer's original paper from 1959 [2]. This is also developed in the papers by Karp [28] and Golumbic et al. [19]. Interval graphs are used for censored and truncated data; the interval indicating for instance observed lifetime (see Gentleman and Vandal [15] and the R packages `MLEcens` and `lcens`). They also come in when restricting data - like permutations - to certain observable intervals, this was the motivation behind the astrophysics paper Efron and Petrosian [13] and the followup paper Diaconis et al. [10]. For an application of rectangle intersections, see Rim and Nakajima [36] and for sphere intersections see Ghrist [16].

2.3. **Random Interval Graphs.** A natural model of *random* interval graphs has $[a_i, b_i]$ chosen uniformly at random inside [0,1]. Scheinerman [38] shows that

$$\#\text{edges} = \frac{n^2}{3} + o_p(n^2), \tag{2.1}$$

$$\mathbb{P}\Big\{\frac{\min_v \deg(v)}{\sqrt{n}} \le x\Big\} \longrightarrow 1 - e^{-x^2/2}, \qquad x > 0, \tag{2.2}$$

and, if v is a fixed vertex,

$$\mathbb{P}\Big\{\frac{\deg(v)}{n} \le x\Big\} \to \begin{cases} 1-(1-x)\frac{\pi}{2}, & x \ge \frac{1}{2}; \\ 1-(1-x)\left\{\frac{\pi}{2} - 2\cos^{-1}[\frac{1}{\sqrt{2-2x}}]\right\} - \sqrt{1-2x}, & x < \frac{1}{2}. \end{cases} \tag{2.3}$$

He further shows that most such graphs are connected, indeed Hamiltonian, the chromatic number is $\frac{n}{2} + o_p(n)$ and several other things; see also Justicz, Scheinerman and Winkler [24] where it is shown that the maximum degree is $n-1$ with probability exactly $2/3$ for any $n > 1$. The chromatic number equals, as said in Section 2.1, the size of the largest clique, and this is equivalent to the random sock sorting problem studied by Steinsaltz [41] and Janson [20] where more refined results are shown, including asymptotic normality which for the random interval graph G_n considered here can be written $(\chi(G_n) - n/2)/\sqrt{n} \to N(0, 1/4)$.

We connect this random interval graph to graph limits in Example 7.4, where also other models of random interval graphs are considered.

There has been some followup on this work with Scheinerman [39] introducing an evolving family of models, and Godehardt and Jaworski [17] studying independence numbers of random interval graphs for cluster discovery. Pippenger [35] has studied other models with application to allocation

in multi-sever queues. Here, customers arrive according to a Poisson process, the service time distribution determines an interval length distribution and the intervals, falling into a given window give an interval graph. For natural models, those graphs are sparse, in contrast to our present study of dense graphs.

Finally we give a pointer to an emerging literature on random intersection graphs where subsets of size d from a finite set are chosen uniformly for each vertex and there is an edge between two vertices if the subsets have a non empty intersection. See results and references in [27] and [40].

2.4. **Graph Limits.** This paper studies limits of interval graphs, using the theory of graph limits introduced by Lovász and Szegedy [30] and further developed in Borgs, Chayes, Lovász, Sós and Vesztergombi [7, 8] and other papers by various combinations of these and other authors; see also Austin [1] and Diaconis and Janson [12]. We refer to these papers for the detailed definitions, which may be summarized as follows (using the notation of [12]).

If F and G are two graphs, then $t(F,G)$ denotes the probability that a random mapping $\phi : V(F) \to V(G)$ defines a graph homomorphism, i.e., that $\phi(v)\phi(w) \in E(G)$ when $vw \in E(F)$. (By a random mapping we mean a mapping uniformly chosen among all $|G|^{|F|}$ possible ones; the images of the vertices in F are thus independent and uniformly distributed over $V(G)$, i.e., they are obtained by random sampling with replacement.) The basic definition is that a sequence G_n of graphs converges if $t(F,G_n)$ converges for every graph F; we will use the version in [12] where we further assume $|G_n| \to \infty$. More precisely, the (countable and discrete) set $\mathcal{U}$ of all unlabeled graphs can be embedded in a compact metric space $\overline{\mathcal{U}}$ such that a sequence $G_n \in \mathcal{U}$ of graphs with $|G_n| \to \infty$ converges in $\overline{\mathcal{U}}$ to some limit $\Gamma \in \overline{\mathcal{U}}$ if and only if $t(F,G_n)$ converges for every graph F. Let $\mathcal{U}_\infty := \overline{\mathcal{U}} \setminus \mathcal{U}$ be the set of proper graph limits. The functionals $t(F,\cdot)$ extend to continuous functions on $\overline{\mathcal{U}}$, and an element $\Gamma \in \mathcal{U}_\infty$ is determined by the numbers $t(F,\Gamma)$. Hence, $G_n \to \Gamma \in \mathcal{U}_\infty$ if and only if $|G_n| \to \infty$ and $t(F,G_n) \to t(F,\Gamma)$ for every graph F. (See [7; 8] for other, equivalent, characterizations of $G_n \to \Gamma$.)

We say that a graph limit $\Gamma \in \mathcal{U}_\infty$ is an *interval graph limit* if $G_n \to \Gamma$ for some sequence of interval graphs. The purpose of the present paper is to study this class of graph limits.

Remark 2.1. In Diaconis, Holmes and Janson [11], the corresponding problem for the class $\mathcal{T}$ of threshold graphs is studied. Recall that a graph is a threshold graph [33] if there are real valued vertex labels v_i and a threshold t such that (i,j) is an edge if and only if $v_i + v_j \le t$. Equivalently, the graph can be built up sequentially by adding vertices which are either dominating (connected to all previous vertices) or isolated (disjoint from all previous vertices). Threshold graphs are a subclass of interval graphs; this can be seen from the sequential description by choosing a sequence of intervals overlapping all previous or disjoint from all previous intervals as required. Thus every threshold graph limit is an interval graph limit. The description of

threshold graph limits in [11] uses special properties of threshold graphs, and is of a somewhat different type than the descriptions of interval graph limits in the present paper. Thus, a threshold graph limit may be represented both as in [11] and as in the present paper, and the representations will not be the same. (This is nothing strange, since the representations typically are non-unique.)

Let $\mathcal{I} \subset \mathcal{U}$ be the set of all interval graphs, and let $\mathcal{I}_\infty \subset \mathcal{U}_\infty$ be the set of all interval graph limits; further, let $\overline{\mathcal{I}} \subset \overline{\mathcal{U}}$ be the closure of $\mathcal{I}$ in $\overline{\mathcal{U}}$. Then $\mathcal{I}_\infty = \overline{\mathcal{I}} \cap \mathcal{U}_\infty = \overline{\mathcal{I}} \setminus \mathcal{I}$. Clearly, $\mathcal{I}_\infty$ is a closed subset of $\mathcal{U}_\infty$ and thus a compact metric space.

A graph limit $\Gamma \in \mathcal{U}_\infty$ may be represented as follows [30], see also [12; 1] for connections to the Aldous–Hoover representation theory for exchangeable arrays [26]. Let $(\mathcal{S}, \mu)$ be an arbitrary probability space and let $W : \mathcal{S} \times \mathcal{S} \to [0,1]$ be a symmetric measurable function. (W is sometimes called *graphon* [7; 8], we will use the alternative *kernel* denomination [4].) Let $X_1, X_2, \dots$, be an i.i.d. sequence of random elements of $\mathcal{S}$ with common distribution μ. Then there is a (unique) graph limit $\Gamma \in \mathcal{U}_\infty$ with, for every graph F,

$$\begin{aligned} t(F,\Gamma) &= \mathbb{E} \prod_{ij \in E(F)} W(X_i, X_j) \\ &= \int_{\mathcal{S}^{|F|}} \prod_{ij \in E(F)} W(x_i, x_j) \, \mathrm{d}\mu(x_1) \cdots \mathrm{d}\mu(x_{|F|}). \end{aligned} \tag{2.4}$$

Further, let, for every $n \geq 1$, $G(n, W, \mu)$ be the random graph obtained by first taking random $X_1, X_2, \dots, X_n$, and then, conditionally given $X_1, X_2, \dots, X_n$, for each pair (i,j) with $i < j$ letting the edge ij appear with probability $W(X_i, X_j)$, (conditionally) independently for all pairs (i,j) with $i < j$. Then the random graph $G_n = G(n, W, \mu)$ converges to Γ a.s. as $n \to \infty$.

Conversely, every graph limit $\Gamma \in \mathcal{U}_\infty$ can be represented in this way by some such $(\mathcal{S}, \mu)$ and W. (The representation is not unique, see Section 8.)

Remark 2.2. For any random graph $G(n, W, \mu)$ (not just interval graphs) the number of copies of any fixed subgraph (e.g. triangles) is a U-statistic, perhaps with extra randomization if W takes on values other than 0 or 1. Thus central limit theorems with error estimates and correction terms as well as large deviation results are available.

It is usually convenient to fix $(\mathcal{S}, \mu)$ and let $W : \mathcal{S}^2 \to [0,1]$ vary; the standard choice of $(\mathcal{S}, \mu)$ is the unit interval $[0,1]$ with Lebesgue measure λ. (Every graph limit can be represented as in (2.4) using this space.) For interval graphs, however, we find it more natural and convenient to instead fix $\mathcal{S}$ and W as follows, and let μ vary.

There is some flexibility in the definition above of interval graphs. The intervals in the definition may be arbitrary intervals of real numbers, or more generally intervals in any totally ordered set, but we may without changing the class of interval graphs restrict the intervals to be, for example, closed.

We may also suppose that all intervals are subsets of $[0,1]$. It may sometimes be convenient to allow an empty interval $\emptyset$ (for isolated vertices), but we find it more convenient (at least notationally) to abstain from this and consider non-empty intervals only. We will, however, allow "intervals" $[a,a]=\{a\}$ of length 0.

Consequently, from now and throughout the paper (except where stated otherwise) we let $\mathcal{S}:=\{[a,b]:0\le a\le b\le 1\}$ be the set of closed subintervals of $[0,1]$ (non-empty, but allowing intervals of length 0). $\mathcal{S}$ is naturally identified with a closed triangle in the plane, and is thus a compact metric space. (It is the compactness that makes this space better for our purposes than, for example, the space of all closed intervals in $\mathbb{R}$.) Further, we let from now on $W:\mathcal{S}\times\mathcal{S}\to\{0,1\}$ be the function

$$W(I,J)=\mathbf{1}[I\cap J\neq\emptyset]. \tag{2.5}$$

Then, a graph $G=(V,E)$ is an interval graph if and only if there exist intervals $I_v\in\mathcal{S}$, $v\in V$, such that the edge indicators $\mathbf{1}[vw\in E]=W(I_v,I_w)$, $v\neq w$.

Every probability measure μ on $\mathcal{S}$ defines a graph limit $\Gamma\in\mathcal{U}_\infty$ by (2.4); we denote this graph limit by Γ_μ. Similarly, we denote the random graph $G(n,W,\mu)$ constructed from $(\mathcal{S},\mu)$ and W by $G(n,\mu)$; this is simply the random interval graph defined by a random i.i.d. sequence of intervals $X_1,X_2,\dots,X_n$ with distribution μ; we further allow $n=\infty$ here, and let $G(\infty,\mu)$ be the random infinite graph defined in the same way by $X_1,X_2,\dots$. (In [12], the standard situation when $\mathcal{S}$ and μ are fixed, we instead use the notations Γ_W and $G(n,W)$; we will also use that notation when we discuss general functions W again in Section 8.) Hence, by the general results quoted above, $G(n,\mu)\to\Gamma_\mu$ a.s. as $n\to\infty$. In particular, Γ_μ is an interval graph limit: $\Gamma_\mu\in\mathcal{I}_\infty$ for every probability measure μ on $\mathcal{S}$.

Remark 2.3. A graph $G\in\mathcal{U}$ corresponds to a 'ghost' $\Gamma_G\in\mathcal{U}_\infty$ with $t(F,\Gamma_G)=t(F,G)$ for all F [30; 12]. If G is an interval graph represented by a sequence $I_1,\dots,I_n$ of intervals in $\mathcal{S}$ (with $n=|G|$), then it follows easily from (2.4) that $\Gamma_G=\Gamma_\mu$, where $\mu=\frac{1}{n}\sum_1^n\delta_{I_i}$ is the distribution of a random interval chosen uniformly from $I_1,\dots,I_n$.

Our main theorem (Theorem 3.1) gives a converse: every interval graph limit can be represented by a probability μ on $\mathcal{S}$; moreover, we may impose a normalization. (In fact, we have a choice between three different normalizations.) However, even with one of these normalizations, the representing measure is not always unique.

Remark 2.4. For every measure μ on $\mathcal{S}$ we thus have a model $G(n,\mu)$ of random interval graphs. Different measures μ give the same model (i.e., with the same distribution for every n) if and only if they give the same graph limit Γ_μ, see Section 8. We may thus construct a large number of different

models of random interval graphs in this way. We give a few examples in Section 7.

2.5. **Degree distribution.** Suppose that G_n is a sequence of graphs with, for convenience, $|G_n| = n$, such that $G_n \to \Gamma$ for a graph limit Γ which is represented by a kernel W on a probability space $(\mathcal{S}, \mu)$. (In this subsection W and $\mathcal{S}$ may be arbitrary.) Let $\bar{d}(G_n) = 2e(G_n)/n$ be the average degree of G_n. It follows immediately $\bar{d}(G_n)/n$ converges to the average $\int_{\mathcal{S}^2} W(x_1, x_2)\,\mathrm{d}\mu(x_1)\,\mathrm{d}\mu(x_2)$; in fact, $\bar{d}(G_n)/n = 2e(G_n)/n^2 = t(K_2, G_n) \to t(K_2, \Gamma) = \int_{\mathcal{S}^2} W$. (Equivalently, the edge density $e(G_n)/\binom{n}{2} \to \int_{\mathcal{S}^2} W$.)

Moreover, let $\nu(G_n)$ be the normalized degree distribution of G_n, defined as the distribution of the random variable d_i/n, where i is a uniformly random vertex in G_n and d_i its degree. Then $\nu(G_n)$ converges weakly (as a probability measure on $[0,1]$) to the distribution of the random variable $W_1(X) := \int_{\mathcal{S}} W(X, z)\,\mathrm{d}\mu(z)$, where X is a random element of $\mathcal{S}$ with distribution μ; note that $W_1(X) \in [0,1]$ and that its mean is $\int_{\mathcal{S}^2} W$. We can thus regard the distribution of this random variable $W_1(X)$ as the degree distribution of the graph limit; we denote it by $\nu(\Gamma)$ or (in our case, where W is fixed) $\nu(\mu)$. See for example [11].

In particular, for any given μ on our standard $\mathcal{S}$, this applies a.s. to the random interval graphs $G(n, \mu)$, since $G(n, \mu) \to \Gamma_\mu$ as said above.

3. Interval Graph Limits, Theorems

Let $\mathcal{P}(\mathcal{S})$ be the set of probability measures on $\mathcal{S} := \{[a, b] : 0 \le a \le b \le 1\}$, equipped with the standard topology of weak convergence, which makes $\mathcal{P}(\mathcal{S})$ a compact metric space. If $\mu \in \mathcal{P}(\mathcal{S})$, let μ_L and μ_R be the marginals of μ (regarding $\mathcal{S}$ as a subset of $\mathbb{R}^2$), i.e., the probability measures on $[0,1]$ induced by μ and the mappings $\mathcal{S} \to [0,1]$ given by $[a, b] \mapsto a$ and $[a, b] \mapsto b$, respectively.

We further consider, both as normalizations and for reasons of continuity, see Corollary 5.2 below, three subsets of $\mathcal{P}(\mathcal{S})$: (as above, λ denotes Lebesgue measure, i.e., the uniform distribution)

$$\mathcal{P}_L(\mathcal{S}) := \{\mu \in \mathcal{P}(\mathcal{S}) : \mu_L = \lambda\}, \tag{3.1}$$

$$\mathcal{P}_R(\mathcal{S}) := \{\mu \in \mathcal{P}(\mathcal{S}) : \mu_R = \lambda\}, \tag{3.2}$$

$$\mathcal{P}_m(\mathcal{S}) := \{\mu \in \mathcal{P}(\mathcal{S}) : \tfrac{1}{2}(\mu_L + \mu_R) = \lambda\}. \tag{3.3}$$

We have the following result, which is proved in Section 6.

Theorem 3.1. *$\mathcal{I}_\infty = \{\Gamma_\mu : \mu \in \mathcal{P}(\mathcal{S})\}$. Moreover, every $\Gamma \in \mathcal{I}_\infty$ may be represented as Γ_μ where we further may impose any one of the normalization conditions in (3.1)–(3.3). In other words,*

$$\mathcal{I}_\infty = \{\Gamma_\mu : \mu \in \mathcal{P}_L(\mathcal{S})\} = \{\Gamma_\mu : \mu \in \mathcal{P}_R(\mathcal{S})\} = \{\Gamma_\mu : \mu \in \mathcal{P}_m(\mathcal{S})\}.$$

Furthermore, the mapping $\mu \to \Gamma_\mu$ is a continuous map of each of $\mathcal{P}_L(\mathcal{S})$, $\mathcal{P}_R(\mathcal{S})$ and $\mathcal{P}_m(\mathcal{S})$ onto $\mathcal{I}_\infty$.

The mappings $\mathcal{P}_L(\mathcal{S}) \to \mathcal{I}_\infty$, $\mathcal{P}_R(\mathcal{S}) \to \mathcal{I}_\infty$, $\mathcal{P}_m(\mathcal{S}) \to \mathcal{I}_\infty$ are not injective. We return to this question in Sections 7 and 8.

The proof in Section 6 also shows the following, which gives an interpretation of the measure μ.

Theorem 3.2. *Let G_n be an interval graph, for convenience with n vertices, defined by intervals $I_{ni} = [a_{ni}, b_{ni}] \subseteq [0,1]$, $i = 1, \dots, n$. Suppose that, as $n \to \infty$, the empirical measure*

$$\mu_n := \frac{1}{n} \sum_{i=1}^{n} \delta_{I_{ni}} \in \mathcal{P}(\mathcal{S}) \tag{3.4}$$

converges weakly to a measure $\mu \in \mathcal{P}(\mathcal{S})$, and suppose further that μ_L and μ_R have no common atom. Then G_n converges to the graph limit Γ_μ.

Instead of probability measures $\mu \in \mathcal{P}(\mathcal{S})$, we may equivalently consider $\mathcal{S}$-valued random variables, i.e., random intervals $[L, R] \in \mathcal{S}$. Each such random interval is given by a pair of random variables (L, R) with $0 \le L \le R \le 1$ (a.e.), and conversely. (Of course, we then only care about the (joint) distribution of (L, R).) Note that the distribution of $[L, R]$ belongs to $\mathcal{P}_L(\mathcal{S})$ $[\mathcal{P}_R(\mathcal{S})]$ if and only if $L \sim \mathsf{U}(0,1)$ $[R \sim \mathsf{U}(0,1)]$.

Example 3.3. A natural model for a collection of confidence intervals for a basic physical constant (as Youden's data in the introduction or the speed of light or the gravitational constant) has intervals of the form $[\mu_i - c\sigma_i, \mu_i + c\sigma_i]$ with μ_i and σ_i independently chosen, μ_i from a normal (μ, σ^2) distribution and σ_i^2 from a Chi-squared distribution, c is computed from the normal quantile q and the sample size $c = q_{1-\frac{\alpha}{2}}/\sqrt{n}$, where α is the target type I error. Here the intervals are not constrained to $\mathcal{S}$. A natural transformation using the distribution function $F(x)$ of $\mu_i + c\sigma_i$ yields the random intervals $[F^{-1}(\mu_i - c\sigma_i), F^{-1}(\mu_i + c\sigma_i)]$, which correspond to points from a distribution on $\mathcal{S}$ belonging to our $\mathbb{P}_R(\mathcal{S})$.

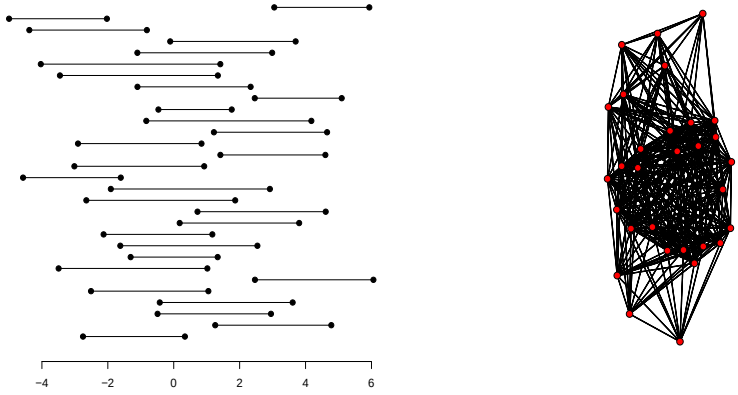
An example, with $\mu_i \sim N(0,4)$ and $\sigma_i^2 \sim \frac{4}{19}\chi^2_{19}$ is given in Figure 3. Although the graph is not the complete graph as it should be if all 30 intervals overlapped, the degrees are high and quite even. The degree distribution is:

```
22 20 25 26 23 14 23 23 27 23 26 25 26 27 13
23 17 23 20 27 23 14 24 25 25 26 26 17 11 10
```

Remark 3.4. There is an obvious reflection map of $\mathcal{S}$ onto itself given by $[a, b] \mapsto [1-b, 1-a]$; we denote the corresponding map of $\mathcal{P}(\mathcal{S})$ onto itself by $\mu \mapsto \check{\mu}$. (It terms of random intervals $[L, R]$, this is $[L, R] \mapsto [1-R, 1-L]$.) The reflection map preserves W, and it follows that $\Gamma_{\check{\mu}} = \Gamma_\mu$.

Note that $\mu \in \mathcal{P}_L(\mathcal{S}) \iff \check{\mu} \in \mathcal{P}_R(\mathcal{S})$, and conversely, which means that we can transfer results from $\mathcal{P}_L(\mathcal{S})$ to $\mathcal{P}_R(\mathcal{S})$, and conversely, by the reflection map; hence it is enough to consider one of $\mathcal{P}_L(\mathcal{S})$ and $\mathcal{P}_R(\mathcal{S})$.

Remark 3.5. As a corollary to Theorem 3.1, we see that every limit of interval graphs may be represented by a kernel that is 0/1-valued. (This implies



(A) Confidence Intervals (30) for random normal data generated with $\mu = 0, \sigma = 2$.

(B) Interval Graph

FIGURE 3. The interval graph for a sets of Normal intervals.

that every representing kernel is 0/1-valued, see [23] for details.) Graph classes with this property are called *random-free* by Lovász and Szegedy [31], who among other results gave a graph-theoretic characterization of such classes. We have thus shown that the class of interval graphs is random-free. We will see in Sections 10.1–10.4 that so are the graph classes considered there.

4. Cliques and chromatic number

If G is a graph, let $\chi(G)$ be its chromatic number and $\omega(G)$ its clique number, i.e., the maximal size of a clique. As said in Section 2.1, interval graphs are perfect and $\chi(G) = \omega(G)$ for them. If G is an interval graph defined by a collection of intervals $\{I_i\}$, it is easily seen that $\omega(G) = \max_x \#\{i : x \in I_i\}$. We define the corresponding quantity for measures $\mu \in \mathcal{P}(\mathcal{S})$ by

$$\omega(\mu) := \sup_{a\in[0,1]} \mu\{I : a \in I\} = \sup_{a\in[0,1]} \mu\big([0,a]\times[a,1]\big). \tag{4.1}$$

Thus, if G is an interval graph defined by intervals $I_1, \dots, I_n$ in $\mathcal{S}$, and $\mu = \frac{1}{n}\sum_{i=1}^n \delta_{I_i}$, then $\omega(G) = n\omega(\mu)$.

It is easy to see that $a \mapsto \mu\big([0,a]\times[a,1]\big)$ is upper semicontinuous; this implies that the supremum in (4.1) is attained.

We will prove the following results in Section 9.

Lemma 4.1. *If μ_1 and μ_2 are probability measures on $\mathcal{S}$ that are equivalent in the sense that $\Gamma_{\mu_1} = \Gamma_{\mu_2}$, then $\omega(\mu_1) = \omega(\mu_2)$.*

This shows that we can define the clique number $\omega(\Gamma)$ for every interval graph limit Γ by $\omega(\Gamma_\mu) = \omega(\mu)$ for $\mu \in \mathcal{P}(\mathcal{S})$.

Theorem 4.2. *Let G_n be an interval graph, for convenience with n vertices, and suppose that $G_n \to \Gamma$ as $n \to \infty$ for some graph limit Γ. Then*

$$\frac{1}{n}\chi(G_n) = \frac{1}{n}\omega(G_n) \to \omega(\Gamma). \tag{4.2}$$

Remark 4.3. Neither $\frac{1}{n}\chi$ nor $\frac{1}{n}\omega$ are continuous functions on the space of all graphs. This may be seen by the following construction: a sequence of dense graphs which tend to the limiting Erdös-Renyi graph with $p = 1$ (complete graph) but with $\frac{1}{n}\chi$ and $\frac{1}{n}\omega$ converging to limits different from one. For the construction, let G_n be an Erdös-Renyi graph with $p = 1 - \frac{1}{\sqrt{n}}$. This converges to the same limit as the sequence of complete graphs K_n. However, an easy argument shows that $\frac{1}{n}\omega$ converges to zero. The same example can be used to show that $\frac{1}{n}\chi$ is not continuous. For this we use the following

Lemma 4.4. *For any graph G with n vertices, $\chi(G) \leq (n + \omega(G))/2$*

Proof. Color by picking two non adjacent vertices, giving both the same new color. Repeat until a connected subgraph of size m remains and give each remaining vertex a separate color. This uses $(n - m)/2 + m = (n + m)/2$ colors and $m \leq \omega(G)$. □

For the random graphs constructed above, $\omega(G) = o(n)$ implies $\chi(G) \leq \frac{n}{2} + o(n)$. Thus $\frac{1}{n}\chi$ is discontinuous.

5. Continuity

The mapping $\mu \mapsto \Gamma_\mu$ of $\mathcal{P}(\mathcal{S})$ into $\mathcal{U}_\infty$ is not continuous. However, the following holds, as we will prove below.

Theorem 5.1. *The mapping $\mu \mapsto \Gamma_\mu$ of $\mathcal{P}(\mathcal{S})$ into $\mathcal{U}_\infty$ is continuous at every $\mu \in \mathcal{P}(\mathcal{S})$ such that μ_L and μ_R have no common atom. Conversely, it is continuous only at these μ.*

In particular, Γ_μ is a continuous function of μ at every μ such that either μ_L or μ_R is continuous, which yields the following corollary.

Corollary 5.2. *The mapping $\mu \mapsto \Gamma_\mu$ is a continuous map $\mathcal{P}_L(\mathcal{S}) \to \mathcal{U}_\infty$, $\mathcal{P}_R(\mathcal{S}) \to \mathcal{U}_\infty$ and $\mathcal{P}_m(\mathcal{S}) \to \mathcal{U}_\infty$.*

To prove Theorem 5.1, we begin by letting $D_W \subset \mathcal{S}^2$ be the set of discontinuity points of W.

Lemma 5.3. $D_W = \big\{([a,b],[c,d]) : b = c \text{ or } a = d\big\}$.

Proof. Obvious. □

Proof of Theorem 5.1. Suppose that $\mu_n \to \mu$ in $\mathcal{P}(\mathcal{S})$ and that μ_L and μ_R have no common atom. Then Lemma 5.3 implies that $\mu \times \mu(D_W) = 0$, and it follows that if $F \in \mathcal{U}$ and $k = |F|$, then $\prod_{ij \in E(F)} W(x_i, x_j) : \mathcal{S}^k \to \{0,1\} \subset \mathbb{R}$ is μ^k-a.e. continuous. Further, $\mu_n^k \to \mu^k$ in $\mathcal{P}(\mathcal{S}^k)$, and thus

$t(F,\Gamma_{\mu_n}) \to t(F,\Gamma_\mu)$ by (2.4), see [3, Theorem 5.2]. Hence, $\Gamma_{\mu_n} \to \Gamma_\mu$ by the definition of $\mathcal{U}_\infty$.

For the converse (which we will not use), assume that a is a common atom of μ_L and μ_R. By symmetry we may suppose that $a > 0$. Let, for $n > 1/a$, $a_n := a - 1/n$. If μ has an atom at $[a,a]$, we define μ_n by moving half of that atom to $[a_n, a_n]$. Otherwise, we replace every interval $[c,a]$ with $c \le a_n$ by $[c,a_n]$; this yields a map $\mathcal{S} \to \mathcal{S}$ which maps μ to a measure μ_n. It is easy to see, in both cases, that $\mu_n \to \mu$ but, using (2.4),

$$t(K_2,\Gamma_{\mu_n}) = \int_{\mathcal{S}^2} W \,\mathrm{d}\mu_n \times \,\mathrm{d}\mu_n \not\to \int_{\mathcal{S}^2} W \,\mathrm{d}\mu \times \,\mathrm{d}\mu = t(K_2,\Gamma_\mu).$$

Hence $\Gamma_{\mu_n} \not\to \Gamma_\mu$. □

6. Proof of Theorems 3.1 and 3.2

Proof of Theorem 3.2. It follows from (2.4), see Remark 2.3, that

$$t(F,G_n) = t(F,\Gamma_{\mu_n}), \qquad F \in \mathcal{U}. \tag{6.1}$$

Theorem 5.1 shows that $\Gamma_{\mu_n} \to \Gamma_\mu$, i.e., $t(F,\Gamma_{\mu_n}) \to t(F,\Gamma_\mu)$ for every $F \in \mathcal{U}$. By (6.1), this implies $t(F,G_n) \to t(F,\Gamma_\mu)$, $F \in \mathcal{U}$, and thus $G_n \to \Gamma_\mu$. □

Proof of Theorem 3.1. If $\mu \in \mathcal{P}(\mathcal{S})$, then as said in Section 2.4, Γ_μ is the limit a.s. of the sequence $G(n,\mu)$ of interval graphs, and thus $\Gamma_\mu \in \mathcal{I}_\infty$.

Conversely, if G_n is a sequence of interval graphs and $G_n \to \Gamma \in \mathcal{U}_\infty$, then each G_n is represented by some sequence of closed intervals $I_{ni} = [a_{ni}, b_{ni}] \subset \mathbb{R}$, $i = 1,\dots,n$. By, if necessary, increasing the lengths of these interval by small (and, e.g., random) amounts, we may further assume that for each n, the $2n$ endpoints $\{a_{ni}, b_{ni} : 1 \le i \le n\}$ are distinct.

Using an increasing homeomorphism φ_n of $\mathbb{R}$ onto itself, we may further assume that the left endpoints $\{a_{ni} : 1 \le i \le n\}$ are the points $\{j/n : 0 \le j < n\}$ in some order, and further that all endpoints $b_{ni} \le 1$. Thus $I_{ni} \in \mathcal{S}$ for every i. Let $\mu_n \in \mathcal{P}(\mathcal{S})$ be the corresponding probability measure given by (3.4).

Since $\mathcal{S}$ is compact, the sequence μ_n is automatically tight, and there exists a probability measure $\mu \in \mathcal{P}(\mathcal{S})$ such that, at least along a subsequence, $\mu_n \to \mu$. As a consequence, $\mu_{nL} \to \mu_L$, and since we have forced μ_{nL} to be the uniform measure on the set $\{j/n : j = 0,\dots,n-1\}$, the limit $\mu_L = \lambda$. Hence $\mu \in \mathcal{P}_L(\mathcal{S})$.

Consequently, Theorem 3.2 applies and shows that (along the subsequence) $G_n \to \Gamma_\mu$. Hence $\Gamma = \Gamma_\mu$.

This shows that every $\Gamma \in \mathcal{I}_\infty$ equals Γ_μ for some $\mu \in \mathcal{P}_L(\mathcal{S})$. The same argument but choosing the homeomorphism φ_n of $\mathbb{R}$ onto itself such that the right endpoints or all $2n$ endpoints are evenly spaced in $[0,1]$ similarly yields $\Gamma = \Gamma_\mu$ with $\mu \in \mathcal{P}_R(\mathcal{S})$ or $\mu \in \mathcal{P}_m(\mathcal{S})$.

This, combined with Corollary 5.2, completes the proof of Theorem 3.1. □

7. Examples

As is well-known, representations as in Section 1 of graph limits by symmetric measurable functions W on a probability space are far from unique, see e.g., [30; 7; 12] and Section 8.

In particular, an interval graph limit $\Gamma \in \mathcal{I}$ may be represented as Γ_μ for many different $\mu \in \mathcal{P}(\mathcal{S})$. For example, any monotone (increasing or decreasing) homeomorphism $[0,1] \to [0,1]$ induces a homeomorphism of $\mathcal{S}$ onto itself which preserves W, and hence maps any $\mu \in \mathcal{P}(\mathcal{S})$ to a measure μ' with $\Gamma_\mu = \Gamma_{\mu'}$. (One example of such a homeomorphism of $\mathcal{S}$ onto itself is the reflection map in Remark 3.4, induced by the map $x \to 1-x$.)

If we use one of the normalizations in (3.1)–(3.3) and consider only $\mathcal{P}_L(\mathcal{S})$, $\mathcal{P}_R(\mathcal{S})$ or $\mathcal{P}_m(\mathcal{S})$, the possibilities are severly restricted, and we have uniqueness in some cases, but not all.

Example 7.1. The complete graph K_n is an interval graph, and can be represented by any family of intervals that contain a common point. The sequence converges to a graph limit $\Gamma \in \mathcal{I}$. On the standard space $[0,1]$, Γ is simply represented by the function $[0,1]^2 \to [0,1]$ that is identically 1, but we are are interested in representations as Γ_μ for $\mu \in \mathcal{P}(\mathcal{S})$. Clearly, $\Gamma = \Gamma_\mu$ for any $\mu \in \mathcal{P}(\mathcal{S})$ such that there exists a point $c \in [0,1]$ with μ supported on the set $\{[a,b] : a \le c \le b\}$.

It is easily seen that there is a unique representation with $\mu \in \mathcal{P}_L(\mathcal{S})$; μ is the distribution of $[U,1]$ with $U \sim \mathsf{U}(0,1)$).

Similarly (and equivalently by reflection), there is a unique representation with $\mu \in \mathcal{P}_R(\mathcal{S})$; μ is the distribution of $[0,U]$ with $U \sim \mathsf{U}(0,1)$.

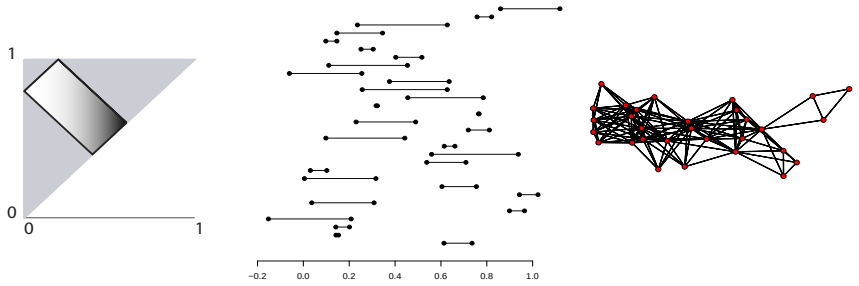
However, there are many representations with $\mu \in \mathcal{P}_m(\mathcal{S})$; these are given by random intervals $[L,R]$ where (L,R) has any joint distribution with the marginals $L \sim \mathsf{U}(0,\frac{1}{2})$ and $R \sim \mathsf{U}(\frac{1}{2},1)$.

Example 7.2. Consider the disjoint union of two complete graphs with $\lfloor an \rfloor$ and $n - \lfloor an \rfloor$ vertices, where $0 < a < 1/2$. This sequence of graphs converges as $n \to \infty$ to a graph limit that is represented by two measures in $\mathcal{P}_L(\mathcal{S})$, with corresponding random intervals $[L,R]$ where $L \sim \mathsf{U}(0,1)$ and R is given by either

$$R := \begin{cases} a, & L \le a, \\ 1, & L > a, \end{cases}$$

or the same formula with a replaced by $1-a$. It can be seen that these two measures are the only measures in $\mathcal{P}_L(\mathcal{S})$ representing the graph limit. (This is an example of a sum of two graph limits; see [21] for general results on such sums and decompositions.)

Example 7.3. More generally, let $(p_i)_1^m$ be a finite or infinite sequence of positive numbers with sum 1. Let G_n be the interval graph consisting of disjoint complete graphs of orders $\lfloor np_1 \rfloor$, $\lfloor np_2 \rfloor$, ... (Hence, $|G_n| = n - o(n)$.) It is easily seen that $G_n \to \Gamma$ for some $\Gamma \in \mathcal{U}_\infty$; thus $\Gamma \in \mathcal{I}$. (Again, cf. [21].)



(A) Support (B) Intervals (C) Interval Graph

FIGURE 4. (A) shows support of μ: tilted rectangle within $\mathcal{S}$. (B) shows the intervals with choice of parameters: 30 intervals, $r = 0.2$. (C) shows the corresponding interval graph.

To represent Γ as Γ_μ with $\mu \in \mathcal{P}_L(\mathcal{S})$, let $(J_i)_1^m$ be a partition of $(0,1]$ into disjoint intervals $J_i = (a_i, b_i]$ with $\lambda(J_i) = p_i$. Then, if $L \sim \mathsf{U}(0,1)$ and R is defined by $R := b_i$ when $L \in J_i$, the random interval $[L,R]$ represents Γ. If $m < \infty$ and $p_1, \dots, p_m$ are distinct, this gives $m!$ different measures $\mu \in \mathcal{P}_L(\mathcal{S})$ representing the same Γ_μ, since the intervals J_i may come in any order. If $m = \infty$, we have an infinite number of different representations.

Example 7.4. The random interval graph studied by Scheinerman [38], see Section 2.3, is defined as $G(n,\mu)$ where $\mu \in \mathcal{P}(\mathcal{S})$ is the uniform measure on $\mathcal{S}$; thus μ has the density $2\,\mathrm{d}x\,\mathrm{d}y$ on $0 \le x \le y \le 1$. Note that the marginal distributions μ_L and μ_R have densities $2(1-x)$ and $2x$ on [0,1], and are thus not uniform. Hence $\mu \notin \mathcal{P}_L(\mathcal{S})$ and $\mu \notin \mathcal{P}_R(\mathcal{S})$; however, $\mu \in \mathcal{P}_m(\mathcal{S})$.

The integral $W_1([x,y]) := \int_{\mathcal{S}} W([x,y],J)\,\mathrm{d}\mu(J) = 1 - x^2 - (1-y)^2$; this leads by Section 2.5 and a calculation to the degree distribution (2.3) found by Scheinerman [38].

It is easily seen that $\omega(\mu) = 1/2$, and thus Theorem 4.2 yields Scheinerman's result that $\chi(G(n,\mu))/n \to 1/2$ (with convergence a.s.).

To obtain an equivalent representing measure $\mu' \in \mathcal{P}_R(\mathcal{S})$, we apply the homeomorphism $x \mapsto x^2$ of $[0,1]$ onto itself; this measure μ' has the density $(2\sqrt{xy})^{-1}\,\mathrm{d}x\,\mathrm{d}y$ on $\mathcal{S} = \{[x,y] : 0 \le x \le y \le 1\}$.

Example 7.5. Scheinerman [39] studies another random interval graph model, defined by random intervals $[x_i - \rho_i, x_i + \rho_i]$ where $x_i \sim \mathsf{U}(0,1)$ and $\rho_i \sim \mathsf{U}(0,r)$ are independent, and $r > 0$ is a parameter. This is $G(n,\mu)$ where μ is the uniform distribution on the tilted rectangle with vertices in $(0,0)$, $(1,1)$, $(1-r,1+r)$, $(-r,r)$; this rectangle does not lie inside our standard triangle (i.e., the intervals are not necessarily inside $[0,1]$), but we may scale it to, for example, the rectangle with vertices $(0, \frac{2r}{1+2r})$, $(\frac{r}{1+2r}, \frac{r}{1+2r})$, $(\frac{r+1}{1+2r}, \frac{r+1}{1+2r})$, $(\frac{1}{1+2r}, 1)$. See Figure 4 for an example.

Example 7.6. Let $0 < r \le 1$ and let μ be uniform on the line $\{(x, x+r) : 0 \le x \le 1-r\}$. This is the set of intervals of length r inside [0,1], so by scaling we obtain a random set of intervals of length 1 in $\mathbb{R}$; hence the random graph $G(n,\mu)$ is in this case a *unit interval graph*, see Section 10.4.

The degree distribution $\nu(\mu)$, i.e., the asymptotic degree destribution of the random graph $G(n,\mu)$, is easily found from Section 2.5. For example, if $r \le \frac{1}{3}$, then $\nu(\mu)$ is the distribution of $W(X)$ with $X \sim \mathsf{U}(0, 1-r)$ and

$$W_1(x) = \begin{cases} \frac{x+r}{1-r}, & 0 \le x < r, \\ \frac{2r}{1-r}, & r \le x \le 1-2r, \\ \frac{1-x}{1-r}, & 1-2r < x \le 1-r. \end{cases}$$

Thus, $\nu(\mu)$ has a density 2 on $[\frac{r}{1-r}, \frac{2r}{1-r})$ and a point mass $\frac{1-3r}{1-r}$ at $\frac{2r}{1-r}$. If $\frac{1}{3} \le r \le \frac{1}{2}$, then similarly $\nu(\mu)$ has a density 2 on $[\frac{r}{1-r}, 1)$ and a point mass $\frac{3r-1}{1-r}$ at 1. If $r \ge \frac{1}{2}$, then $G(n,\mu)$ is the complete graph and $\nu(\mu)$ is a point mass at 1.

The chromatic number is by Theorem 4.2 a.s. $\frac{r}{1-r}n + o(n)$ for $r \le \frac{1}{2}$ (and trivially n for $r \ge \frac{1}{2}$).

Example 7.7. Theorem 3.1 shows that we can build any interval graph limit from a probability distribution μ on $\mathcal{S} := \{[x,y] : 0 \le x \le y \le 1\}$, with the marginal distribution of μ on the y axis being uniform, i.e., $\mu \in \mathcal{P}_R(\mathcal{S})$. Here is a hierarchy of examples of building such measures

(i) As in Example 7.1 for the complete graph K_n. We take μ to be the uniform distribution on the y axis. Repeated picks from μ correspond to intervals $[0, u_i]$ which all intersect.

(ii) The empty graph E_n is an interval graph corresponding to disjoint intervals. Let μ be the uniform distribution on the $x = y$ diagonal. Repeated picks from μ yield intervals $[u_i, u_i]$ which are disjoint with probability 1.

(iii) We may interpolate between these two examples, choosing a with $0 \le a \le 1$ and μ_a uniform on the line $\ell_a = \{(x,y) : x = ay,\ 0 \le y \le 1\}$. This is done by picking intervals $[aU, U]$ with $U \sim \mathsf{U}(0,1)$ so the y-margin U is uniform on $[0,1]$. Now, some pairs of points on the line ℓ_a will result in edges and some not:
For $[x_1, y_1]$, $[x_2, y_2]$ in $\mathcal{S}$, the intervals overlap iff $x_1 \le x_2 \le y_1$ or $x_2 \le x_1 \le y_2$. Equivalently if $x_1 \le x_2$ then $x_2 \le y_1$, or if $x_1 \ge x_2$ then $y_2 \le x_1$. Here, the points on ℓ_a are $[ay_1, y_1]$ and $[ay_2, y_2]$, so there is overlap iff $ay_1 \le ay_2 \le y_1$ or $ay_2 \le ay_1 \le y_2$, or equivalently,

$$ay_1 \le y_2 \le y_1/a.$$

Thus the chance of an edge in this model is $\mathbb{P}(aU_1 \le U_2 \le U_1/a) = 2\,\mathbb{P}(aU_1 \le U_2 \le U_1) = 1 - a$.

Moreover, by Section 2.5, the asymptotic degree distribution $\nu(\mu_a)$ is the distribution of $W_1(U)$, where $U \sim \mathsf{U}(0,1)$ and

$$W_1(u) = \begin{cases} \bigl(\frac{1}{a} - a\bigr)u, & u \ge a, \\ 1 - au, & a \le u \le 1. \end{cases}$$

This distribution has density $a/(1-a^2)$ on $[0, 1-a]$ and $1/(a(1-a^2))$ on $[1-a, 1-a^2]$ (for $a < 1$). The chromatic number is by Theorem 4.2 a.s. asymptotic to $n\omega(\mu_a) = (1-a)n$.

(iv) The next example of $\mu \in \mathcal{P}_R(\mathcal{S})$ is a mixture of uniforms on ℓ_a, where a has a distribution on $[0,1]$. The prescription:
- Pick $a = A$ at random from some distribution on $[0,1]$ and
- independently pick uniformly on ℓ_a,

means that we pick intervals $[A, UA]$ with A and U independent and $U \sim \mathsf{U}(0,1)$, while A has any given distribution.

(v) As an extreme example, consider the measure μ which is a $(\theta, 1-\theta)$ mixture of uniform on ℓ_0, ℓ_1, with $\theta \in [0,1]$. Then, identifying the vertices of $G(n,\mu)$ with the picked points in $\mathcal{S}$:
- None of the points on ℓ_1 have an edge between them.
- All of the points on the line ℓ_0 have edges between them.
- Pairs of points, one from ℓ_0, one from ℓ_1 have an edge with probability $1/2$, but not independently. More precisely, there is an edge between $(0,u_1)$ and (u_2,u_2) iff $u_1 \ge u_2$.

It is easily seen that in this case, the random interval graph $G(n,\mu)$ is a threshold graph, see Remark 2.1; we may give $(0,u)$ label u and (u,u) label $-u$ and take the threshold $t = 0$. (By [11, Corollary 6.7], $G(n,\mu)$ equals the random graph $T_{n,\theta}$ defined in [11].) Hence Γ_μ is a threshold graph limit in this case. (It is an open problem to characterize all $\mu \in \mathcal{P}(\mathcal{S})$ such that Γ_μ is a threshold graph limit.)

(vi) Uniform intervals: As said in Example 7.4, the uniform distribution on $\mathcal{S}$ does not belong to $\mathcal{P}_R(\mathcal{S})$, but it is equivalent to the distribution with density $(2\sqrt{xy})^{-1}\,\mathrm{d}x\,\mathrm{d}y$ which does. A change of variables to $(a,y) \in [0,1]^2$ with $a = x/y$ yields the density $\frac{1}{2}a^{-1/2}\,\mathrm{d}a\,\mathrm{d}y$, so this is of the type studied here, with a having the $B(\frac{1}{2},1)$ distribution with density $\frac{1}{2}a^{-1/2}\,\mathrm{d}a$.

8. Uniqueness

We state a general equivalence theorem for representation of graph limits (not necessarily interval graph limits) by symmetric measurable functions. We therefore allow rather general probability spaces $(\mathcal{S}_1,\mu_1) = (\mathcal{S}_2,\mu_2)$ and general symmetric functions $W_i : \mathcal{S}_i^2 \to [0,1]$ on them. In the standard case $(\mathcal{S}_1,\mu_1) = (\mathcal{S}_2,\mu_2) = ([0,1],\lambda)$, parts (i)–(vii) of the theorem are given in [12] as a consequence of Hoover's equivalence theorem for representations of exchangeable arrays Kallenberg [26, Theorem 7.28]. Other similar results are given by Bollobás and Riordan [5] and Borgs, Chayes and Lovász [6]; in

particular, (viii) and (ix) below are modelled after similar results in Borgs, Chayes and Lovász [6]. A similar theorem is stated in Janson [23], and an almost identical theorem in the related case of partial orders is given in Janson [22].

We first introduce more notation. If $W_2 : \mathcal{S}_2^2 \to [0,1]$ and $\varphi : \mathcal{S}_1 \to \mathcal{S}_2$, then $W_2^\varphi(x,y) := W_2(\varphi(x),\varphi(y))$.

A *Borel space* is a measurable space $(\mathcal{S},\mathcal{F})$ that is isomorphic to a Borel subset of $[0,1]$, see e.g. [25, Appendix A1] and Parthasarathy [34]. In fact, a Borel space is either isomorphic to $([0,1],\mathcal{B})$ or it is countable infinite or finite. Moreover, every Borel subset of a Polish topological space (with the Borel σ-field) is a Borel space. A *Borel probability space* is a probability space $(\mathcal{S},\mathcal{F},\mu)$ such that $(\mathcal{S},\mathcal{F})$ is a Borel space.

If W' is a symmetric function $\mathcal{S}^2 \to [0,1]$, where $\mathcal{S}$ is a probability space, we say following [6] that $x_1, x_2 \in \mathcal{S}$ are *twins* (for W') if $W'(x_1,y) = W'(x_2,y)$ for a.e. $y \in \mathcal{S}$. We say that W' is *almost twin-free* if there exists a null set $N \subset \mathcal{S}$ such that there are no twins $x_1, x_2 \in \mathcal{S} \setminus N$ with $x_1 \neq x_2$.

In the theorem and its proof, we assume that $[0,1]$ is equipped with the measure λ, and $\mathcal{S}_j$ with μ_j; for simplicity we do not always repeat this.

Theorem 8.1. *Suppose that $(\mathcal{S}_1,\mu_1)$ and $(\mathcal{S}_2,\mu_2)$ are two Borel probability spaces and that $W_1 : \mathcal{S}_1^2 \to [0,1]$ and $W_2 : \mathcal{S}_2^2 \to [0,1]$ are two symmetric measurable functions, and let $\Gamma_1, \Gamma_2 \in \mathcal{U}_\infty$ be the corresponding graph limits. Then the following are equivalent.*

(i) $\Gamma_1 = \Gamma_2$ *in* $\mathcal{U}_\infty$.
(ii) $t(F,\Gamma_1) = t(F,\Gamma_2)$ *for every graph* F.
(iii) *The exchangeable random infinite graphs $G(\infty,W_1)$ and $G(\infty,W_2)$ have the same distribution.*
(iv) *The random graphs $G(n,W_1)$ and $G(n,W_2)$ have the same distribution for every finite n.*
(v) *There exist measure preserving maps $\varphi_j : [0,1] \to \mathcal{S}_j$, $j = 1,2$, such that $W_1^{\varphi_1} = W_2^{\varphi_2}$ a.e., i.e., $W_1\bigl(\varphi_1(x),\varphi_1(y)\bigr) = W_2\bigl(\varphi_2(x),\varphi_2(y)\bigr)$ a.e. on $[0,1]^2$.*
(vi) *There exists a measurable mapping $\psi : \mathcal{S}_1 \times [0,1] \to \mathcal{S}_2$ that maps $\mu_1 \times \lambda$ to μ_2 such that $W_1(x,y) = W_2\bigl(\psi(x,t_1),\psi(y,t_2)\bigr)$ for a.e. $x,y \in \mathcal{S}_1$ and $t_1,t_2 \in [0,1]$.*
(vii) *$\delta_\square(W_1,W_2) = 0$, where $\delta_\square$ is the cut metric defined in [7] (see also [5]).*

If further W_2 is almost twin-free, then these are also equivalent to:

(viii) *There exists a measure preserving map $\varphi : \mathcal{S}_1 \to \mathcal{S}_2$ such that $W_1 = W_2^\varphi$ a.s., i.e. $W_1(x,y) = W_2\bigl(\varphi(x),\varphi(y)\bigr)$ a.e. on $\mathcal{S}_1^2$.*

If both W_1 and W_2 are almost twin-free, then these are also equivalent to:

(ix) *There exists a measure preserving map $\varphi : \mathcal{S}_1 \to \mathcal{S}_2$ such that φ is a bimeasurable bijection of $\mathcal{S}_1 \setminus N_1$ onto $\mathcal{S}_2 \setminus N_2$ for some null sets $N_1 \subset \mathcal{S}_1$ and $N_2 \subset \mathcal{S}_2$, and $W_1 = W_2^\varphi$ a.s., i.e. $W_1(x,y) =$*

$W_2\bigl(\varphi(x),\varphi(y)\bigr)$ *a.e. on* $\mathcal{S}_1^2$. *If further* $(\mathcal{S}_2,\mu_2)$ *has no atoms, for example if* $\mathcal{S}_2=[0,1]$, *then we may take* $N_1=N_2=\emptyset$.

Note that (i) $\implies$ (iv) implies that we can uniquely define the random graphs $G(n,\Gamma)$ for any graph limit Γ.

Proof. (i) $\iff$ (ii) holds by our definition of graph limits.

Next, the equivalences (i) $\iff$ (ii) $\iff$ (iii) $\iff$ (iv) $\iff$ (v) $\iff$ (vi) $\iff$ (vii) where shown in [12] in the special (but standard) case $(\mathcal{S}_1,\mu_1)=(\mathcal{S}_2,\mu_2)=([0,1],\lambda)$. Since every Borel space is either finite, countably infinite or (Borel) isomorphic to $[0,1]$, it is easily seen that there exist measure preserving maps $\gamma_j:[0,1]\to\mathcal{S}_j$, $j=1,2$. Then $W_j^{\gamma_j}:[0,1]^2\to[0,1]$, and it is easily seen that $\Gamma_j:=\Gamma_{W_j}=\Gamma_{W_j^{\gamma_j}}$ and $G(n,W_j)\overset{\mathrm{d}}{=}G(n,W_j^{\gamma_j})$ for $n\le\infty$, and further $\delta_\square(W_j,W_j^{\gamma_j})=0$; hence (i) $\iff$ (ii) $\iff$ (iii) $\iff$ (iv) $\iff$ (vii) by the corresponding results for $[0,1]$.

If (i)–(iv) hold, then by (v) for $[0,1]$, there exist measure preserving functions $\varphi_j':[0,1]\to[0,1]$ such that $W_1^{\gamma_1}\bigl(\varphi_1'(x),\varphi_1'(y)\bigr)=W_2^{\gamma_2}\bigl(\varphi_2'(x),\varphi_2'(y)\bigr)$ a.e., and thus (v) holds with $\varphi_j:=\gamma_j\circ\varphi_j'$.

Conversely, if (v) holds, then $G(n,W_1)\overset{\mathrm{d}}{=}G(n,W_1^{\varphi_1})=G(n,W_2^{\varphi_2})\overset{\mathrm{d}}{=}G(n,W_2)$ for every $n\le\infty$; thus (v) $\implies$ (iii),(iv).

(vi) $\implies$ (iii),(iv) is similar.

(iii) $\implies$ (vi): Assume (iii). Then $G(\infty,W_1^{\gamma_1})\overset{\mathrm{d}}{=}G(\infty,W_2^{\gamma_2})$, so by the result for $[0,1]$, there exists a measure preserving function $h:[0,1]^2\to[0,1]$ such that $W_1^{\gamma_1}(x,y)=W_2^{\gamma_2}\bigl(h(x,z_1),h(y,z_2)\bigr)$ for a.e. $x,y,z_1,z_2\in[0,1]$. By [22, Lemma 7.2] (applied to $(\mathcal{S}_1,\mu_1)$ and γ_1), there exists a measure preserving map $\alpha:\mathcal{S}_1\times[0,1]\to[0,1]$ such that $\gamma_1(\alpha(s,u))=s$ a.e. Hence, for a.e. $x,y\in\mathcal{S}_1$ and $u_1,u_2,z_1,z_2\in[0,1]$,

$$\begin{aligned}W_1(x,y)&=W_1\bigl(\gamma_1\circ\alpha(x,u_1),\gamma_1\circ\alpha(y,u_2)\bigr)=W_1^{\gamma_1}\bigl(\alpha(x,u_1),\alpha(y,u_2)\bigr)\\&=W_2^{\gamma_2}\bigl(h(\alpha(x,u_1),z_1),h(\alpha(y,u_2),z_2)\bigr)\\&=W_2\bigl(\gamma_2\circ h(\alpha(x,u_1),z_1),\gamma_2\circ h(\alpha(y,u_2),z_2)\bigr).\end{aligned}$$

Finally, let $\beta=(\beta_1,\beta_2)$ be a measure preserving map $[0,1]\to[0,1]^2$, and define $\psi(x,t):=\gamma_2\circ h\bigl(\alpha(x,\beta_1(t)),\beta_2(t)\bigr)$.

(vi) $\implies$ (viii): Since, for a.e. x,y,t_1,t_2,t_1',

$$W_2\bigl(\psi(x,t_1),\psi(y,t_2)\bigr)=W_1(x,y)=W_2\bigl(\psi(x,t_1'),\psi(y,t_2)\bigr)$$

and ψ is measure preserving, it follows that for a.e. x,t_1,t_1', $\psi(x,t_1)$ and $\psi(x,t_1')$ are twins for W_2. If W_2 is almost twin-free, with exceptional null set N, then further $\psi(x,t_1),\psi(x,t_1')\notin N$ for a.e. x,t_1,t_1', since ψ is measure preserving, and consequently $\psi(x,t_1)=\psi(x,t_1')$ for a.e. x,t_1,t_1'. It follows that we can choose a fixed t_1' (almost every choice will do) such that $\psi(x,t)=\psi(x,t_1')$ for a.e. x,t. Define $\varphi(x):=\psi(x,t_1')$. Then $\psi(x,t)=\varphi(x)$ for a.e. x,t, which in particular implies that φ is measure preserving, and (vi) yields $W_1(x,y)=W_2\bigl(\varphi(x),\varphi(y)\bigr)$ a.e.

(viii) $\Longrightarrow$ (ix): Let $N' \subset \mathcal{S}_1$ be a null set such that if $x \notin N'$, then $W_1(x,y) = W_2(\varphi(x),\varphi(y))$ for a.e. $y \in \mathcal{S}_1$. If $x, x' \in \mathcal{S}_1 \setminus N'$ and $\varphi(x) = \varphi(x')$, then x and x' are twins for W_1. Consequently, if W_1 is almost twin-free with exceptional null set N'', then φ is injective on $\mathcal{S}_1 \setminus N_1$ with $N_1 := N' \cup N''$. Since $\mathcal{S}_1 \setminus N_1$ and $\mathcal{S}_2$ are Borel spaces, the injective map $\varphi : \mathcal{S}_1 \setminus N_1 \to \mathcal{S}_2$ has measurable range and is a bimeasurable bijection $\varphi : \mathcal{S}_1 \setminus N_1 \to \mathcal{S}_2 \setminus N_2$ for some measurable set $N_2 \subset \mathcal{S}_2$. Since φ is measure preserving, $\mu_2(N_2) = 0$.

If $\mathcal{S}_2$ has no atoms, we may take an uncountable null set $N_2' \subset \mathcal{S}_2 \setminus N_2$. Let $N_1' := \varphi^{-1}(N_2')$. Then $N_1 \cup N_1'$ and $N_2 \cup N_2'$ are uncountable Borel spaces so there is a bimeasurable bijection $\eta : N_1 \cup N_1' \to N_2 \cup N_2'$. Redefine φ on $N_1 \cup N_1'$ so that $\varphi = \eta$ there; then φ becomes a bijection $\mathcal{S}_1 \to \mathcal{S}_2$.

(viii),(ix) $\Longrightarrow$ (v): Trivial. □

We apply this general theorem to the case $\mathcal{S}_1 = \mathcal{S}_2 = \mathcal{S}$ and $W_1 = W_2 = W$.

Corollary 8.2. *Let $\mu_1, \mu_2 \in \mathcal{P}(\mathcal{S})$. Then, $\Gamma_{\mu_1} = \Gamma_{\mu_2}$ if and only if there exists a measurable map $\psi : \mathcal{S} \times [0,1] \to \mathcal{S}$ that maps $\mu_1 \times \lambda \to \mu_2$ such that for μ_1-a.e. intervals $I, J \in \mathcal{S}$ and a.e. $t, u \in [0,1]$,*

$$I \cap J \neq \emptyset \iff \psi(I,t) \cap \psi(J,u) \neq \emptyset. \tag{8.1}$$

This result is still not completely satisfactory, and it leads to a number of open questions:

Problems 8.3. (i) The simple case is when the mapping ψ in Corollary 8.2 does not depend on the second variable at all; in other words, when there exists a measurable map $\varphi : \mathcal{S} \to \mathcal{S}$ that maps μ_1 to μ_2 such that for μ_1-a.e. intervals $I, J \in \mathcal{S}$,

$$I \cap J \neq \emptyset \iff \varphi(I) \cap \varphi(J) \neq \emptyset. \tag{8.2}$$

When is this possible, and when is the extra randomization in (8.1) really needed?

(ii) To simplify the condition further, when is it possible to choose ψ or φ such that (8.1) or (8.2) hold for *all* I and J, and not just almost all? Note that in Example 7.2, the two different representing measures are related by the map φ defined by $\varphi([x,y]) = [x+1-a, y+1-a]$ for $y \leq a$ and $\varphi([x,y]) = [x-a, y-a]$ for $y > x \geq a$, and arbitrarily for $x < a < y$; this φ satisfies (8.2) for a.e. I and J, but not for all.

(iii) One way to obtain a map $\varphi : \mathcal{S} \to \mathcal{S}$ that satisfies (8.1) for all I and J is to take $\varphi([a,b]) = [f(a), f(b)]$ for a (strictly) increasing map $f : [0,1] \to [0,1]$, or $\varphi([a,b]) = [f(b), f(a)]$ for a (strictly) decreasing map $f : [0,1] \to [0,1]$. Are there any other such maps φ? Again, note that in Examples 7.2 and 7.3 there are natural maps φ that satisfy (8.2) for a.e. I and J, but these are given by functions f that permute subintervals of $[0,1]$, and are not monotone. It seems that this problem is related to connectedness of the random interval graphs $G(n, \mu_1)$, and also to the question whether

there are several orientations of the complement of these interval graphs, cf. [14].

Problem 8.4. Is there some additional condition on μ that leads to a unique "canonical" representing measure $\mu \in \mathcal{P}(\mathcal{S})$ for each interval graph limit Γ?

Note that requiring $\mu \in \mathcal{P}_L(\mathcal{S})$ yields uniqueness in Example 7.1 but not in Example 7.2.

9. Proof of Theorem 4.2

We begin by proving a special case.

Lemma 9.1. *Let $\mu \in \mathcal{P}(\mathcal{S})$. Then $\frac{1}{n}\omega(G(n,\mu)) \xrightarrow{\text{a.s.}} \omega(\mu)$ as $n \to \infty$.*

Proof. Recall the construction of $G(n,\mu)$ using i.i.d. random intervals $I_1, \dots, I_n$ with distribution μ, and let again $\mu_n = \frac{1}{n}\sum_1^n \delta_{I_i}$ be the corresponding empirical measure.

Let $\varepsilon > 0$. Choose a such that $\omega(\mu) = \mu\bigl([0,a]\times[a,1]\bigr)$. By the law of large numbers, a.s. for all large n,

$$\mu_n\bigl([0,a]\times[a,1]\bigr) = \frac{1}{n}\#\bigl\{i \le n : I_i \in [0,a]\times[a,1]\bigr\} > \omega(\mu) - \varepsilon. \tag{9.1}$$

In the opposite direction, for every $a \in [0,1]$, $\mu\bigl([0,a]\times[a,1]\bigr) < \omega(\mu)+\varepsilon$, and thus, for some $\delta = \delta(a) > 0$, $\mu\bigl([0,a+\delta]\times[a-\delta,1]\bigr) < \omega(\mu)+\varepsilon$. The open intervals $(a-\delta(a), a+\delta(a))$ cover the compact set $[0,1]$, so we can choose a finite subcover $(a_j-\delta_j, a_j+\delta_j)$, $j = 1,\dots,m$. By the law of large numbers, a.s. for all large n, $\#\{i \le n : I_i \in [0,a_j+\delta_j]\times[a_j-\delta_j,1]\} < n(\omega(\mu)+\varepsilon)$ for each $j = 1,\dots,m$, which implies that

$$\mu_n\bigl([0,a]\times[a,1]\bigr) = \frac{1}{n}\#\bigl\{i \le n : I_i \in [0,a]\times[a,1]\bigr\} < \omega(\mu) + \varepsilon \tag{9.2}$$

for every $a \in [0,1]$. Combining (9.1) and (9.2), we see that $\omega(\mu) - \varepsilon < \omega(\mu_n) < \omega(\mu) + \varepsilon$, and the result follows since $\frac{1}{n}\omega(G(n,\mu)) = \omega(\mu_n)$. □

Proof of Lemma 4.1. By Theorem 8.1(i) $\Longrightarrow$ (iv), the random graphs $G(n,\mu_1)$ and $G(n,\mu_2)$ have the same distribution and the result follows by Lemma 9.1. □

A direct analytic proof of Lemma 4.1 using e.g. Corollary 8.2 seems more difficult than this argument using random graphs.

Proof of Theorem 4.2. As in the proof of Theorem 3.1, we may (by considering a subsequence) assume that G_n is defined by intervals $I_{ni} = [a_{ni}, b_{ni}] \subseteq [0,1]$, $i = 1,\dots,n$ such that the corresponding empirical measures μ_n given by (3.4) converge to a measure $\mu \in \mathcal{P}_m(\mathcal{S})$. By Theorem 3.2, $\Gamma_\mu = \Gamma$.

Let $a_n \in [0,1]$ be such that $\omega(\mu_n) = \mu_n\bigl([0,a_n]\times[a_n,1]\bigr)$. By considering a further subsequence we may assume that $a_n \to a$ for some $a \in [0,1]$. Since $\mu \in \mathcal{P}_m(\mathcal{S})$, μ_L and μ_R are continuous measures and thus $\mu\bigl(\partial([0,b]\times$

$[b,1])\bigr) = \mu\bigl([0,b]\times\{b\}\cup\{b\}\times[b,1]\bigr) = 0$ for every $b\in[0,1]$. Together with $\mu_n\to\mu$, this implies

$$\mu\bigl([0,b]\times[b,1]\bigr) = \lim_{n\to\infty}\mu_n\bigl([0,b]\times[b,1]\bigr) \le \liminf_{n\to\infty}\omega(\mu_n). \tag{9.3}$$

Moreover, a routine argument shows that

$$\omega(\mu_n) = \mu_n\bigl([0,a_n]\times[a_n,1]\bigr) \to \mu\bigl([0,a]\times[a,1]\bigr). \tag{9.4}$$

Consequently, $\omega(\mu) = \mu\bigl([0,a]\times[a,1]\bigr)$ and $\omega(\mu_n)\to\omega(\mu)=\omega(\Gamma)$. The result follows for the subsequence since $\chi(G_n)=\omega(G_n)=n\omega(\mu_n)$. The same argument applies to every subsequence of G_n, which thus has a subsubsequence such that (4.2) holds; this implies that (4.2) holds for the full sequence. □

10. Other intersection graphs

The methods above can be used also for some other classes of intersection graphs. In general, for $\mathcal{A}$-intersection graphs defined using a collection $\mathcal{A}$ of sets, we define $W=W_{\mathcal{A}}:\mathcal{A}\times\mathcal{A}\to\{0,1\}$ by

$$W(A,B) = \begin{cases} 1 & \text{if } A\cap B\neq\emptyset,\\ 0 & \text{if } A\cap B=\emptyset. \end{cases} \tag{10.1}$$

We take $\mathcal{S}=\mathcal{A}$ (equipped with some suitable σ-field) and use this fixed function W, just as for the case of interval graphs above. If μ is any probability measure on $\mathcal{S}=\mathcal{A}$, then the random graphs $G(n,\mu)$ are random $\mathcal{A}$-intersection graphs (and each μ gives a model of such random graphs); thus the graph limit Γ_μ is an $\mathcal{A}$-intersection graph limit. The problem whether the converse holds, i.e., whether every $\mathcal{A}$-intersection graph limit can be represented as Γ_μ for some such μ, is more subtle; we have proved it for interval graphs above, and our methods apply also to some other cases, see Sections 10.1–10.3 below; however, the converse is not true in general, see Section 10.4. (For a more trivial counterexample, let $\mathcal{A}$ be the countable family of all finite subsets of $\mathbb{N}$; then every graph is an $\mathcal{A}$-intersection graph, but not every graph limit can be represented by Γ_μ for a measure μ on $\mathcal{A}$, since this would imply that the class of all graphs is random-free, see Remark 3.5, a contradiction.)

We leave the general case as an open problem and remark that our methods seem to work best when the set $\mathcal{A}$ has a compact topology; however, even in that case there are problems because the map $\mu\to\Gamma_\mu$ is in general not continuous, as seen in Theorem 5.1.

Problem 10.1. Find general conditions on $\mathcal{A}$ that guarantee that every $\mathcal{A}$-intersection graph limit is Γ_μ for some $\mu\in\mathcal{P}(\mathcal{A})$.

We study a few cases individually. Note that the function W depends on the graph class by the general formula (10.1). For each class one can ask questions similar to Problems 8.3–8.4, study random graphs $G(n,\Gamma)$ generated by suitable graph limits, and so on; we leave this to the readers.

10.1. **Circular-arc graphs.** *Circular-arc graphs* are the intersection graphs defined by letting $\mathcal{A}$ be the collection of arcs on the unit circle $\mathbb{T}$, see [9; 18; 29]. As for interval graphs, we may assume that the arcs are closed, and we allow arcs of length 0. We also allow the whole circle as an arc; this is special since it has no endpoint. This class obviously contain the interval graphs, and the containment is strict. (For example, the cycle C_n with $n \ge 4$ is a circular-arc graph but not an interval graph.)

For technical reasons, we first regard the whole circle as having two coinciding (and otherwise arbitrary) endpoints. The space of arcs may then be identified with $\mathcal{S}^0_{\mathsf{CA}} := [0, 2\pi] \times \mathbb{T}$, with $(\ell, e^{\mathrm{i}\theta})$ corresponding to the arc $\{e^{\mathrm{i}t} : t \in [\theta, \theta + \ell]\}$ of length ℓ. The argument in the proof of Theorem 3.1 shows that every circular-arc graph limit may be represented as Γ_μ for some measure $\mu \in \mathcal{P}(\mathcal{S}^0_{\mathsf{CA}})$, for example with the marginal distribution of θ uniform on $\mathbb{T}$.

To get rid of the artificial endpoints for the full circle, we identify all points $(2\pi, e^{\mathrm{i}\theta})$ in $\mathcal{S}^0_{\mathsf{CA}}$ and let $\mathcal{S}_{\mathsf{CA}}$ be the resulting quotient space; $\mathcal{S}_{\mathsf{CA}}$ is homeomorphic to the unit disc $D := \{z \in \mathbb{C} : |z| \le 1\}$ with $re^{\mathrm{i}\theta} \in D$ corresponding to $(2\pi(1-r), e^{\mathrm{i}\theta}) \in \mathcal{S}^0_{\mathsf{CA}}$ and thus $0 \in D$ corresponding to the full circle. (This gives a unique representation of the closed arcs on $\mathbb{T}$.) The quotient map $\mathcal{S}^0_{\mathsf{CA}} \to \mathcal{S}_{\mathsf{CA}}$ preserves W, so by mapping μ from $\mathcal{S}^0_{\mathsf{CA}}$ to $\mathcal{S}_{\mathsf{CA}}$, we see that the circular-arc graph limits are exactly the graph limits Γ_μ for $\mu \in \mathcal{P}(\mathcal{S}_{\mathsf{CA}})$, in analogy with Theorem 3.1 for interval graphs. (The main reason that we do not use $\mathcal{S}_{\mathsf{CA}}$ directly in the proof is that W is not continuous at pairs (I, J) where $I = \mathbb{T}$ and J has length 0.)

10.2. **Circle graphs.** *Circle graphs* are the intersection graphs defined by the collection of chords of the unit circle $\mathbb{T}$ [18, Chapter 11]. We represent a chord by its two endpoints, and first for convenience consider the endpoints as an ordered pair of points. We thus consider the space $\mathcal{S}^0_{\mathsf{CG}} := \mathbb{T} \times \mathbb{T}$ (allowing chords of length 0). The argument in the proof of Theorem 3.1 shows that every circle graph limit may be represented as Γ_μ for some measure $\mu \in \mathcal{P}(\mathcal{S}^0_{\mathsf{CG}})$, for example with the average of the two marginal distributions on $\mathbb{T}$ being uniform (in analogy with $\mathcal{P}_m(\mathcal{S})$).

The space of all chords on $\mathbb{T}$ really is the quotient space $\mathcal{S}_{\mathsf{CG}}$ of $\mathcal{S}^0_{\mathsf{CG}}$ obtained by identifying (a, b) and (b, a) for any $a, b \in \mathbb{T}$. (The resulting compact space is homeomorphic to a Möbius strip.) Again, the quotient mapping preserves W, so we can map $\mu \in \mathcal{P}(\mathcal{S}^0_{\mathsf{CG}})$ to a measure on $\mathcal{S}_{\mathsf{CG}}$. Consequently, the circle graph limits are the graph limits Γ_μ for $\mu \in \mathcal{P}(\mathcal{S}_{\mathsf{CG}})$.

10.3. **Permutation graphs.** A graph is a *permutation graph* if we can label the vertices by $1, \dots, n$ and there is a permutation π of $\{1, \dots n\}$ such that for $i < j$ there is an edge ij if and only if $\pi(i) > \pi(j)$. It is easy to see that the permutation graphs are the intersection graphs defined by the collection of all line segments with one endpoint on each of two parallel lines; we may

take $\mathcal{A} = [0,1]\times[0,1]$ with (a,b) representing the line segment between $(a,0)$ and $(b,1)$ [18, Chapter 7].

The argument in the proof of Theorem 3.1 shows that every permutation graph limit may be represented as Γ_μ for some measure $\mu \in \mathcal{P}([0,1]^2)$, for example with the two marginal distributions on $[0,1]$ both being uniform.

10.4. **Unit interval graphs.** *Unit interval graphs* are the intersection graphs defined by the collection $\mathcal{A} = \{[x, x+1] : x \in \mathbb{R}\}$ of unit intervals in $\mathbb{R}$. (Again, we choose the intervals as closed; the collection of open unit intervals defines the same class of graphs.) This class coincides with the class of *proper interval graphs*, defined by collections of intervals $I_1, \dots, I_n$ in $\mathbb{R}$, with the additional requirement that no I_i is a proper subinterval of another. (Or, equivalently, that $I_i \not\subseteq I_j$ for all i, j.) They are also called *indifference graphs*. See [9; 18; 37]. This is a subclass of all interval graphs and the containment is strict since $K_{1,3}$ is an interval graph but not a unit interval graph.

The set $\mathcal{A}$ above is naturally identified with $\mathbb{R}$, with $W(x,y) = 1$ when $|x-y| \le 1$; thus every probability measure on $\mathbb{R}$ defines a unit interval graph limit. However, this mapping is *not* onto. In fact, the empty graph E_n is a unit interval graph, so the limit as $n \to \infty$ is a unit interval graph limit; this graph limit Γ_0 is defined by the kernel 0 on any probability space and has $t(K_2, \Gamma_0) = 0$, but if $\mu \in \mathcal{P}(\mathbb{R})$, then the corresponding graph limit Γ_μ has by (2.4)

$$t(K_2, \Gamma_\mu) = \iint_{|x-y|\le 1} \mathrm{d}\mu(x_1)\,\mathrm{d}\mu(x_2) > 0.$$

Thus $\Gamma_0 \neq \Gamma_\mu$. (Note that if $\mu_n \in \mathcal{P}(\mathbb{R})$ is a measure representing E_n, then necessarily the sequence μ_n is not tight, and in fact converges vaguely to 0, so this problem is connected to the non-compactness of $\mathbb{R}$.)

Another approach to unit interval graph limits is to regard them as special cases of interval graph limits and use the theory developed above to characterize them using special measures on the triangle $\mathcal{S} = \{[a,b] : 0 \le a \le b \le 1\}$. This yields the following theorem.

Theorem 10.2. *A graph limit Γ is a unit interval graph limit if and only if $\Gamma = \Gamma_\mu$ for a measure $\mu \in \mathcal{P}(\mathcal{S})$ that has support on some curve $t \mapsto \gamma(t) = (\gamma_1(t), \gamma_2(t)) \in \mathcal{S}$ such that $\gamma_1(t)$ and $\gamma_2(t)$ are weakly increasing.*

Proof. Suppose that G_n is a sequence of unit interval graphs with $G_n \to \Gamma$. In the proof of Theorem 3.1, the interval representations are modified by homeomorphisms, and the results are, of course, not unit interval representations, but they are proper interval representations, i.e., no interval is a subinterval of another. Thus, the measures μ_n have the property that for each $(a,b) \in \mathcal{S}$, $\mu_n\bigl([0,a)\times(b,1]\bigr)\cdot\mu_n\bigl((a,1]\times[0,b)\bigr) = 0$. Since $\mu_n \to \mu$ (for a subsequence), the same holds for μ, which implies that if $a_1 < a_2$ and $b_1 > b_2$, then (a_1, b_1) and (a_2, b_2) cannot both belong to $\operatorname{supp}\mu$. (Choose $a = (a_1+a_2)/2$ and $b = (b_1+b_2)/2$.)

Let $E = \{a+b : (a,b) \in \operatorname{supp}\mu\}$. Then E is a closed subset of $[0,2]$ and for each $t \in E$ there is exactly one $(a,b) \in \operatorname{supp}\mu$ with $a+b=t$; we define $f(t)=a$ and $g(t)=b$ so f and g are functions $E \to [0,1]$. Note that $f(t)+g(t)=t$. If $t_1 < t_2$ and $g(t_1) > g(t_2)$, then $f(t_1) < f(t_2)$; thus $(f(t_1),g(t_1))$ and $(f(t_2),g(t_2))$ are two points in $\operatorname{supp}\mu$ violating the condition above. Consequently, if $t_1 < t_2$ then $g(t_1) \le g(t_2)$, and similarly $f(t_1) \le f(t_2)$. (Since $f(t)+g(t)=t$ this further implies $f(t_2)-f(t_1) \le t_2-t_1$ and $g(t_2)-g(t_1) \le t_2-t_1$.) We may now extend f and g to the complement $[0,1]\setminus E$, e.g. linearly in each component, and define $\gamma(t)=(f(t),g(t))$.

For the converse, consider the random graph $G(n,\mu)$. This is an interval graph represented by intervals $I_1,\dots,I_n \in \mathcal{S}$ that lie on the curve γ. This is not necessarily a proper interval representation, since two of the intervals may lie on the same horizontal or vertical part of γ, but it is easily seen that it is always possible to obtain a proper interval representation of the same graph by moving some of the endpoints a little. Thus $G(n,\mu)$ is a proper interval graph, and thus a unit interval graph, whence Γ_μ is a unit interval graph limit. □

Again, the representation by such a measure μ is not unique.

Problem 10.3. Is it possible to make a canonical choice in some way? Is it possible to use a fixed curve γ?

Remark 10.4. Γ_μ may happen to be a unit interval graph limit also if μ is not of the type in Theorem 10.2; for example if μ is any measure supported on $[0,\frac12]\times[\frac12,1]$ when each $G(n,\mu)$ is the complete graph K_n. To characterize all measures $\mu \in \mathcal{P}(\mathcal{S})$ such that Γ_μ is a unit interval graph limit is a different, and open, problem.

The unit interval graphs can also be characterized as the intervals graphs G that do not contain $K_{1,3}$ as an induced subgraph [9; 18; 37]. In general, for two graphs F and G with $|F| \le |G|$, let $t_{\mathrm{ind}}(F,G)$ be the probability that the induced subgraph of G obtained by selecting $|F|$ vertices uniformly at random is isomorphic to F; this number is closely connected to $t(F,G)$ defined in Section 2.4 (which loosely speaking counts subgraphs of G and not just induced subgraphs), see [7], [30] or [12] for details. For any fixed F, $t_{\mathrm{ind}}(F,\cdot)$ extends to graph limits Γ and we have $t_{\mathrm{ind}}(F,G_n) \to t_{\mathrm{ind}}(F,\Gamma)$ if $G_n \to \Gamma$; moreover, $t_{\mathrm{ind}}(F,\Gamma)$ is a continuous function of Γ. Using this notation, G is a unit interval graph if and only if G is an interval graph with $t_{\mathrm{ind}}(K_{1,3},G)=0$.

Theorem 10.5. *Let Γ be a graph limit. Then the following are equivalent:*

(i) *Γ is a unit interval graph limit.*
(ii) *Γ is an interval graph limit and $t_{\mathrm{ind}}(K_{1,3},\Gamma)=0$.*
(iii) *The random graphs $G(n,\Gamma)$ are unit interval graphs.*

Proof. (i) $\implies$ (ii) is clear by the comments above.

(ii) $\implies$ (iii). Use Theorem 3.1 and choose a measure $\mu \in \mathcal{P}(\mathcal{S})$ representing Γ. There is a formula analoguous to (2.4) for $t_{\mathrm{ind}}(F,\Gamma)$, with $\prod_{ij\in E(F)} W(x_i,x_j)$ replaced by $\prod_{ij\in E(F)} W(x_i,x_j)\prod_{ij\notin E(F)}(1-W(x_i,x_j))$, and it follows easily that for any $n \ge |F|$,

$$\mathbb{E}\, t_{\mathrm{ind}}\bigl(F,G(n,\Gamma)\bigr) = \mathbb{E}\, t_{\mathrm{ind}}\bigl(F,G(|F|,\Gamma)\bigr) = t_{\mathrm{ind}}(F,\Gamma).$$

Hence, (ii) implies that $G(n,\Gamma)$ a.s. is an interval graph G with $t_{\mathrm{ind}}(K_{1,3},G) = 0$, i.e., a unit interval graph. (The case $n<4$ is trivial.)

(iii) $\implies$ (i) follows since $G(n,\Gamma)\to\Gamma$ a.s. □

Finally, we mention that a related characterization of unit interval graphs is that they are the graphs that contain no induced subgraph isomorphic to C_k for any $k\ge 4$, $K_{1,3}$, S_3 or $\overline{S}_3$, where S_3 is the graph on 6 vertices $\{1,\dots,6\}$ with edge set $\{12,13,23,14,25,36\}$, and $\overline{S}_3$ is its complement [9]. The same argument as in the proof of Theorem 10.5 yields (see [11, Theorem 3.2] for a more general result):

Theorem 10.6. *A graph limit Γ is a unit interval graph limit if and only if $t_{\mathrm{ind}}(F,\Gamma)=0$ for every $F\in\{C_k\}_{k\ge4}\cup\{K_{1,3},S_3,\overline{S}_3\}$.* □

Acknowledgements. Part of this research was done during the 2007 Conference on Analysis of Algorithms (AofA'07) in Juan-les-Pins, France, and in a bus back to Nice after the conference.

References

[1] T. Austin. On exchangeable random variables and the statistics of large graphs and hypergraphs. *Probab. Surv.*, 5:80–145, 2008.

[2] S. Benzer. On the topology of the genetic fine structure. *Proceedings of the National Academy of Sciences*, **45** (1959), 1607–1620. `http://www.pnas.org/cgi/reprint/45/11/1607.pdf`.

[3] P. Billingsley, *Convergence of Probability Measures*. Wiley, New York, 1968.

[4] B. Bollobas, S. Janson and O. Riordan, Monotone graph limits and quasimonotone graphs. Preprint, 2011. `http://arxiv.org/1101.4296`

[5] B. Bollobás and O. Riordan, Metrics for sparse graphs. *Surveys in Combinatorics 2009*, LMS Lecture Notes Series 365, Cambidge Univ. Press, 2009, pp. 211–287. `http://arxiv.org/0708.1919`

[6] C. Borgs, J. T. Chayes and L. Lovász, Moments of two-variable functions and the uniqueness of graph limits. *Geom. Funct. Anal.* **19** (2010), no. 6, 1597–1619.

[7] C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós and K. Vesztergombi, Convergent sequences of dense graphs. I. Subgraph frequencies, metric properties and testing. *Adv. Math.* **219** (2008), no. 6, 1801–1851.

[8] C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós and K. Vesztergombi, Convergent sequences of dense graphs II: Multiway cuts and statistical physics. Preprint, 2007. `http://research.microsoft.com/~borgs/`

[9] A. Brandstädt, V. B. Le and J. P. Spinrad, *Graph Classes: a Survey*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1999.

[10] P. Diaconis, R. Graham, and S. Holmes. Statistical problems involving permutations with restricted positions. *State of the Art in Probability and Statistics (Leiden, 1999)*. IMS Lecture Notes Monogr. Ser. 36, Inst. Math. Statist., Beachwood, OH, 2001, pp. 195–222. http://www.jstor.org/stable/4356113.

[11] P. Diaconis, S. Holmes and S. Janson, Threshold graph limits and random threshold graphs. *Internet Mathematics* **5** (2009), no. 3, 267–318.

[12] P. Diaconis and S. Janson, Graph limits and exchangeable random graphs. *Rend. Mat. Appl. (7)* **28** (2008), 33–61.

[13] B. Efron and V. Petrosian. Nonparametric methods for doubly truncated data. *J. Amer. Statist. Assoc.* **94** (1999), no. 447, 824–834.

[14] P. C. Fishburn, *Interval Orders and Interval Graphs. A Study of Partially Ordered Sets.* Wiley, Chichester, 1985.

[15] R. Gentleman and A. C. Vandal. Computational algorithms for censored-data problems using intersection graphs. *Journal of Computational and Graphical Statistics*, **10** (2001), no. 3, 403–421. http://www.jstor.org/stable/1391096.

[16] R. Ghrist. Barcodes: The persistent topology of data. *Bull. Amer. Math. Soc.*, **45** (2008), no. 1, 61–75.

[17] E. Godehardt and J. Jaworski. Two models of random intersection graphs for classification. *Exploratory Data Analysis in Empirical Research: Proc. 25th Annual Conference of the Gesellschaft für Klassifikation e.V., University of Munich, March 14–16, 2001*, M. Schwaiger and O. Opitz, eds., Springer-Verlag, Berlin, 2003, pp. 67–81.

[18] M. C. Golumbic, *Algorithmic Graph Theory and Perfect Graphs*. 2nd ed. Annals of Discrete Mathematics, 57. Elsevier, Amsterdam, 2004.

[19] M. Golumbic, H. Kaplan, and R. Shamir. On the complexity of DNA physical mapping. *Adv. in Appl. Math.* **15** (1994), no. 3, 251–261. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.12.5001&rep=rep1&type=pdf

[20] S. Janson. Sorting using complete subintervals and the maximum number of runs in a randomly evolving sequence. *Annals of Combinatorics* **12** (2009), no. 4, 417–447.

[21] S. Janson, Connectedness in graph limits. Preprint, 2008. http://arxiv.org/0802.3795

[22] S. Janson, Poset limits and exchangeable random posets. *Combinatorica*, to appear. http://arxiv.org/0902.0306

[23] S. Janson, Graphons, cut norm and distance, couplings and rearrangements. Preprint, 2010. http://arxiv.org/1009.2376

[24] J. Justicz, E. R. Scheinerman and P. Winkler. Random intervals. *Amer. Math. Monthly* **97** (1990), no. 10, 881–889.

[25] O. Kallenberg, *Foundations of Modern Probability*. 2nd ed., Springer-Verlag, New York, 2002.
[26] O. Kallenberg, *Probabilistic Symmetries and Invariance Principles*. Springer, New York, 2005.
[27] M. Karoński, E. R. Scheinerman, and K. B. Singer-Cohen. On random intersection graphs: the subgraph problem. *Combin. Probab. Comput.* **8** (1999), no. 1–2, 131–159.
[28] R. Karp. Mapping the genome: some combinatorial problems arising in molecular biology. *Proceedings of the 25th Annual ACM Symposium on the Theory of Computing (STOC'93)*, ACM, New York, NY, 1993, pp. 278–285.
[29] V. Klee. What are the intersection graphs of arcs in a circle? *Amer. Math. Monthly* **76** (1969), no. 7, 810–813.
[30] L. Lovász and B. Szegedy, Limits of dense graph sequences. *J. Comb. Theory B* **96**, 933–957, 2006.
[31] L. Lovász and B. Szegedy, Regularity partitions and the topology of graphons. Preprint, 2010. `http://arxiv.org/1002.4377v1`
[32] T. A. McKee and F. R. McMorris. *Topics in Intersection Graph Theory*. SIAM Monographs on Discrete Mathematics and Applications. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1999.
[33] N. Mahadev and U. Peled. *Threshold Graphs and Related Topics*. Annals of Discrete Math. 56, North-Holland, Amsterdam, 1995.
[34] K. R. Parthasarathy, *Probability Measures on Metric Spaces*. Academic Press, New York, 1967.
[35] N. Pippenger. Random interval graphs. *Random Structures and Algorithms* **12** (1998), no. 4, 361–380.
[36] C. S. Rim and K. Nakajima. On rectangle intersection and overlap graphs. *IEEE Trans. Circuits Systems I Fund. Theory Appl.* **42** (1995), no. 9, 549553.
[37] F. S. Roberts, Indifference graphs. *Proof Techniques in Graph Theory (Proc. Second Ann Arbor Graph Theory Conf., Ann Arbor, Mich., 1968)*, Academic Press, New York, 1969, pp. 139–146.
[38] E. R. Scheinerman. Random interval graphs. *Combinatorica* **8** (1988), no. 4, 357–371.
[39] E. R. Scheinerman. An evolution of interval graphs. *Discrete Math.* **82** (1990), no. 3, 287–302.
[40] D. Stark. The vertex degree distribution of random intersection graphs. *Random Structures Algorithms* **24** (2004), no. 3, 249–258.
[41] D. Steinsaltz. Random time changes for sock-sorting and other stochastic process limit theorems. *Electron. J. Probab.* **4** (1999), no. 14, 25 pp.
[42] M. S. Waterman and J. R. Griggs. Interval graphs and maps of DNA. *Bulletin of Mathematical Biology* **48** (1986), no. 2, 189–195.
[43] W. J. Youden. Enduring values. *Technometrics* **14** (1972), no. 1, 1–11.

[44] X. J. Zhou, M.-C. J. Kao, H. Huang, A. Wong, J. Nunez-Iglesias, M. Primig, O. M. Aparicio, C. E. Finch, T. E. Morgan, and W. H. Wong. Functional annotation and network reconstruction through cross-platform integration of microarray data. *Nature Biotechnology* **23** (2005), no. 2, 238–243.

Department of Mathematics, Stanford University, CA 94305, USA
E-mail address: diaconis@math.stanford.edu
URL: http://www-stat.stanford.edu/∼CGATES/persi

Department of Statistics, Stanford, CA 94305, USA
E-mail address: susan@stat.stanford.edu
URL: http://www-stat.stanford.edu/∼susan/

Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden
E-mail address: svante.janson@math.uu.se
URL: http://www.math.uu.se/∼svante/