CONNECTEDNESS IN GRAPH LIMITS

SVANTE JANSON

Abstract. We define direct sums and a corresponding notion of connectedness for graph limits. Every graph limit has a unique decomposition as a direct sum of connected components. As is well-known, graph limits may be represented by symmetric functions on a probability space; there are natural definitions of direct sums and connectedness for such functions, and there is a perfect correspondence with the corresponding properties of the graph limit. Similarly, every graph limit determines an infinite random graph, which is a.s. connected if and only if the graph limit is connected. There are also characterizations in terms of the asymptotic size of the largest component in the corresponding finite random graphs, and of minimal cuts in sequences of graphs converging to a given limit.

1. Introduction and main results

A deep and beautiful theory of limit objects of dense graphs has in recent years been created by Lovász and Szegedy [17] and Borgs, Chayes, Lovász, Sós and Vesztergombi [7, 8], and further developed in a series of further papers by these and other authors. (In particular, Bollobas, Borgs, Chayes and Riordan [4] contains some ideas similar to the ones in the present paper.) Some basic components of the theory are reviewed below and in Section 2, following the presentation in Diaconis and Janson [11].

We let $\mathcal{U}$ be the set of unlabelled graphs; as is explained in Section 2, this is embedded in a compact metric space $\overline{\mathcal{U}}$, and we let $\mathcal{U}_\infty := \overline{\mathcal{U}} \setminus \mathcal{U}$, the (compact) set of graph limits.

We denote the vertex and edge sets of a graph G by $E(G)$ and $V(G)$, and let $v(G) := |V(G)|$ and $e(G) := |E(G)|$ be the numbers of vertices and edges. All graphs in this paper are simple (i.e., without multiple edges or loops) and undirected; we further assume $0 < v(G) < \infty$, except sometimes when we explicitly consider infinite graphs or (sub)graphs that may be without vertices.

A well-known basic property for graphs is connectedness. We write the disjoint union of two graphs G_1 and G_2 as $G_1 \oplus G_2$; thus $V(G_1 \oplus G_2) = V(G_1) \cup V(G_2)$ and $E(G_1 \oplus G_2) = E(G_1) \cup E(G_2)$, where we assume that $V(G_1) \cap V(G_2) = \emptyset$. (We usually work with unlabelled graphs, which can be regarded as equivalence classes of labelled graphs, and then $V(G_1) \cap V(G_2) = \emptyset$ can always be assumed by relabelling the vertices when necessary.) A

Date: February 26, 2008; revised April 10, 2008.

graph is *connected* if and only if it cannot be written as $G_1 \oplus G_2$ for two graphs G_1 and G_2. It is well-known, and easy to see, that every graph G can be uniquely decomposed as a disjoint union $G_1 \oplus G_2 \oplus \cdots \oplus G_M$ of connected graphs, the *components* of G.

The first aim of the present paper is to define an analoguous notion of direct sum of graph limits and to establish analogues of these standard results for graphs. The natural definition seems to be that of a weighted sum. In Section 4, we define for any two graph limits Γ_1 and Γ_2 in $\mathcal{U}_\infty$ and any $\alpha \in [0,1]$ a graph limit $\alpha\Gamma_1 \oplus (1-\alpha)\Gamma_2$. One way these sums can be characterized is the following, which shows the connection with disjoint unions of graphs, and further yields an explanation of the weight α.

All unspecified limits below are as $n \to \infty$.

Theorem 1.1. *Suppose that G_n and G'_n, $n \geq 1$, are graphs that satisfy $v(G_n) \to \infty$, $v(G'_n) \to \infty$, $G_n \to \Gamma \in \mathcal{U}_\infty$ and $G'_n \to \Gamma' \in \mathcal{U}_\infty$. If further $v(G_n)/\big(v(G_n) + v(G'_n)\big) \to \alpha \in [0,1]$, then $G_n \oplus G'_n \to \alpha\Gamma \oplus (1-\alpha)\Gamma'$.*

The proof is given in Section 4, while other results stated in this section are proven in Section 5.

We use the analogy with graphs to define connectedness for graph limits.

Definition 1.2. A graph limit $\Gamma \in \mathcal{U}_\infty$ is *disconnected* if there exist graph limits $\Gamma_1, \Gamma_2 \in \mathcal{U}_\infty$ and $\alpha \in (0,1)$ such that $\Gamma = \alpha\Gamma_1 \oplus (1-\alpha)\Gamma_2$. If this is not the case, Γ is *connected.*

It is not obvious that there are any connected graph limits at all, but this follows from the theorems below; an explicit example is given in Example 1.17.

Remark 1.3. Graph limits are defined in a rather abstract way, by taking the completion of $\mathcal{U}$ in a suitable metric, and they are not constructed as topological spaces, nor can they (as far as I know) be seen as topological spaces in any meaningful way (as graphs can). Hence, the connectedness defined and studied here is not the usual topological notion.

Definition 1.4. Let $\Gamma \in \mathcal{U}_\infty$. A *component* of Γ is a connected graph limit $\Gamma_1 \in \mathcal{U}_\infty$ such that $\Gamma = \alpha\Gamma_1 \oplus (1-\alpha)\Gamma_2$ for some $\alpha \in (0,1]$ and $\Gamma_2 \in \mathcal{U}_\infty$.

It is not obvious that components exist, but we will see that every graph limit has at least one component, with one exception discussed in Remark 1.8.

A graph limit may have an infinite number of components. (An example is given in Example 4.4.) Hence, in order to exhibit a graph limit as the sum of its components, we need to extend the direct sum to the case of an infinite number of terms. This too is done in Section 4, where we define the direct sum $\bigoplus_{i=1}^m \alpha_i\Gamma_i$ for any finite or infinite sequence $(\Gamma_i)_{i=1}^m$ of graph limits and corresponding weights $(\alpha_i)_{i=1}^m$ with $\alpha_i \geq 0$ and $\sum_{i=1}^m \alpha_i \leq 1$. Let $\mathcal{A}$ be the set of all such sequences $(\alpha_i)_{i=1}^m$ with $0 \leq m \leq \infty$, every $\alpha_i \geq 0$ and $\sum_{i=1}^m \alpha_i \leq 1$. Further, let $\mathcal{A}_+$ be the subset of $\mathcal{A}$ of such sequences with every $\alpha_i > 0$, and let $\mathcal{A}_1$ be the subset of $\mathcal{A}$ of sequences with $\sum_{i=1}^m \alpha_i = 1$. (Although we allow $\alpha_i = 0$ in the definition of $\bigoplus_{i=1}^m \alpha_i\Gamma_i$, all such terms may be deleted without changing the direct sum, cf. Definition 4.3 and (4.2), and thus there is no essential restriction to consider $\mathcal{A}_+$ only.)

Theorem 1.5. *Every graph limit $\Gamma \in \mathcal{U}_\infty$ can be written as a finite or infinite direct sum $\bigoplus_{i=1}^M \alpha_i\Gamma_i$ for some M with $0 \le M \le \infty$ and some sequences $(\Gamma_i)_{i=1}^M$ and $(\alpha_i)_{i=1}^M$ with every $\Gamma_i \in \mathcal{U}_\infty$ connected and $(\alpha_i)_{i=1}^M \in \mathcal{A}_+$.*

The number M and the sequences $(\Gamma_i)_1^M$ and $(\alpha_i)_1^M$ are uniquely determined by Γ, up to simultaneous permutations of the two sequences.

The set of components of Γ equals $\{\Gamma_i\}_{i=1}^M$, i.e., a graph limit Γ' is a component of Γ if and only if $\Gamma' = \Gamma_i$ for some i.

We call the number M in Theorem 1.5 the *number of components* of Γ.

Remark 1.6. There may be repetitions in the sequence $(\Gamma_i)_1^M$; hence the number of distinct components, i.e., the number of graph limits Γ' that are components of Γ, is not necessarily equal to M. (For example, take any connected Γ_1 and let $\Gamma = \frac{1}{2}\Gamma_1 \oplus \frac{1}{2}\Gamma_1$.)

Remark 1.7. It may seem surprising that we allow $\sum_i \alpha_i < 1$ in the definition of direct sums and in Theorem 1.5. However, such *deficient* direct sums are equivalent to *complete* ones (i.e., $\sum_i \alpha_i = 1$) as follows. There is a special (and trivial) graph limit $\mathbf{0} \in \mathcal{U}_\infty$, which is the limit of the empty graphs E_n, see Example 2.1, and for any sequences $(\Gamma_i)_1^M$ and $(\alpha_i)_1^M$ with $\alpha_i \ge 0$ and $\sum_{i=1}^M \alpha_i < 1$, if we let $\alpha_0 := 1 - \sum_{i=1}^M \alpha_i$ and $\Gamma_0 = \mathbf{0}$, then $\bigoplus_{i=1}^M \alpha_i\Gamma_i = \bigoplus_{i=0}^M \alpha_i\Gamma_i$, where the latter direct sum is complete.

Conversely, any summand $\alpha_i\mathbf{0}$ in a direct sum may always be deleted.

It is thus a matter of taste whether we want to allow deficient direct sums or not, in the latter case instead allowing a term $\alpha_0\mathbf{0}$ in the decomposition in Theorem 1.5. We prefer the version above, partly because $\mathbf{0}$ is disconnected and thus not a component of Γ.

Remark 1.8. We allow $M = 0$ in Theorem 1.5, but this occurs only in a trivial case: $M = 0 \iff \Gamma$ has no components $\iff \Gamma = \mathbf{0}$.

Remark 1.9. A graph limit Γ is connected if and only if it has $M = 1$ and $\alpha_1 = 1$ in Theorem 1.5. Note that a graph limit with $M = 1$ and $\alpha_1 < 1$ is *not* connected since it equals the direct sum $\alpha_1\Gamma_1 \oplus (1-\alpha_1)\mathbf{0}$, cf. Remark 1.7.

1.1. **Connectedness in graphs and their limits.** It should be obvious that connectedness (or its opposite) of the graphs G_n in a convergent sequence does not say anything about the limit. In fact, the convergence of G_n to a limit Γ is not sensitive to addition or deletion of $o(v(G_n)^2)$ edges to/from G_n, and such changes might create or destroy connectedness. For example, assuming $G_n \to \Gamma$, we may add edges from a given vertex, say 1, in G_n, to all other vertices, thus creating connected graphs G'_n with $G'_n \to \Gamma$.

Similarly, we may delete all edges having vertex 1 as an endpoint, thus making 1 isolated and creating disconnected graphs G''_n with $G''_n \to \Gamma$.

Instead, the connectedness of the limit of a sequence of graphs is connected to quantitative connectedness properties of the graphs, more precisely the sizes of minimal cuts. Given two subsets V', V'' of the vertex set $V(G)$ of a graph G, we let $e(V', V'') = e_G(V', V'')$ denote the number $|\{ij \in E(G) : i \in V',\ j \in V''\}|$ of edges between the two sets.

Theorem 1.10. *Suppose that G_n are graphs with $v(G_n) \to \infty$ and that $G_n \to \Gamma \in \mathcal{U}_\infty$.*

(i) *If Γ is connected, then for every $\delta > 0$ there exists $\varepsilon > 0$ such that, for all large n, if $V(G_n) = V' \cup V''$ is a partition with $|V'|, |V''| \geq \delta v(G_n)$, then $e(V', V'') \geq \varepsilon v(G_n)^2$.*

(ii) *If Γ is disconnected, then there exists $\delta > 0$ and, for all large n, a partition $V(G_n) = V' \cup V''$ with $|V'|, |V''| \geq \delta v(G_n)$ such that $e(V', V'') = o\big(v(G_n)^2\big)$.*

Theorem 1.10(ii) (in a version using kernels, cf., Theorem 1.16 below, is given (although not as a formal theorem) by Bollobas, Borgs, Chayes and Riordan [4].

1.2. **Kernels.** As was shown by Lovász and Szegedy [17], see Section 2 for details, every graph limit may be represented by a *kernel* W on a probability space $(\mathcal{S}, \mu)$, i.e., a symmetric measurable function $W : \mathcal{S} \times \mathcal{S} \to [0, 1]$. (Furthermore, $(\mathcal{S}, \mu)$ may be, and often is, chosen as $([0, 1], \lambda)$, where λ is Lebesgue measure, but we will not insist on that.) We let $\Gamma_W \in \mathcal{U}_\infty$ denote the graph limit represented by W. Note that the representation is not unique; different W, even on $([0, 1], \lambda)$, may give the same Γ_W; again see Section 2.

Remark 1.11. Borgs, Chayes, Lovász, Sós and Vesztergombi [7, 8] use the term *graphon* for such functions W. However, they also consider the graphons as the graph limits, thus identifying equivalent graphons. We prefer for our purposes to be more specific, and in order to avoid possible confusion between the two uses of "graphon", we will not use this term here. As just said, we use instead "kernel" for such functions W.

There are natural definitions of connectedness and sums for kernels, which, as we shall see in Theorems 1.16 and 1.18, correspond directly to the corresponding notions for graph limits. (The non-uniqueness of the representation does not cause any complications in the results, although we have to worry about it in some proofs.)

Definition 1.12. A kernel W on $(\mathcal{S}, \mu)$ is *disconnected* if either $W = 0$ a.e. on $\mathcal{S} \times \mathcal{S}$ or there exists a subset $A \subset \mathcal{S}$ with $0 < \mu(A) < 1$ such that $W = 0$ a.e. on $A \times (\mathcal{S} \setminus A)$. Otherwise W is *connected.*

The exceptional (and trivial) case $W = 0$ a.e. has to be treated separately only when $\mathcal{S}$ is an atom, i.e. if $\mu(A) = 0$ or 1 for every measurable A.

The same properties of kernels were defined and studied in [5] (in a somewhat more general situation), but there called *reducible* and *irreducible*; the same terms were used in [4], where these properties where studied further (in a similar context as the present paper). (The trivial case $W = 0$ a.e. was not treated separately in [5]; this made no difference there, but was with hindsight perhaps a mistake.)

By the *disjoint union* $\dot{\bigcup}\mathcal{S}_i$ of sets $\mathcal{S}_i$, we mean their usual union if the sets are disjoint; if not, we first make them disjoint by replacing $\mathcal{S}_i$ by $\mathcal{S}_i \times \{i\}$.

Definition 1.13. Let $0 \le m \le \infty$ and let $(\alpha_i)_1^m \in \mathcal{A}$; further, for each i, let W_i be a kernel on a probability space $(\mathcal{S}_i, \mu_i)$.

(i) In the complete case, $(\alpha_i)_1^m \in \mathcal{A}_1$, let $\mathcal{S}$ be the disjoint union $\dot{\bigcup}\mathcal{S}_i$, let μ be the probability measure on $\mathcal{S}$ given by $\mu(A) := \sum_{i=1}^m \alpha_i \mu_i(A \cap \mathcal{S}_i)$, and let the direct sum $\bigoplus_{i=1}^m \alpha_i W_i$ be the kernel W on $(\mathcal{S}, \mu)$ defined by

$$W(x,y) = \begin{cases} W_i(x,y), & x, y \in \mathcal{S}_i; \\ 0, & x \in \mathcal{S}_i,\ y \in \mathcal{S}_j \text{ with } i \neq j. \end{cases}$$

(ii) In the deficient case, $(\alpha_i)_1^m \in \mathcal{A} \setminus \mathcal{A}_1$, take any probability space $(\mathcal{S}_0, \mu_0)$, let $W_0 := 0$ (on $\mathcal{S}_0 \times \mathcal{S}_0$) and $\alpha_0 := 1 - \sum_{i=1}^m \alpha_i$, and define $\bigoplus_{i=1}^m \alpha_i W_i := \bigoplus_{i=0}^m \alpha_i W_i$.

For the definition of $\bigoplus_{i=1}^m \alpha_i W_i$ in the deficient case, cf. Remark 1.7 and note that $\mathbf{0}$ is represented by $W_0 = 0$ on any $(\mathcal{S}, \mu)$. Our definition does not specify $\mathcal{S}_0$ and μ_0 and is thus formally not a proper definition, but any choice will do for our purposes and the flexibility is convenient.

Remark 1.14. If $(\mathcal{S}_i, \mu_i) = ([0,1], \lambda)$ for every i, it is natural to make linear changes of variables to replace $\mathcal{S}_i$ by the interval $I_i := [\sigma_{i-1}, \sigma_i)$ of length α_i, where $\sigma_i := \sum_{j \le i} \alpha_j$; note that $(I_i)_1^m$ form a partition of $[0,1)$ and we obtain $(\mathcal{S}, \mu) = ([0,1), \lambda)$ or, if we prefer, $([0,1], \lambda)$.

Remark 1.15. The notation for deficient sums of kernels has to be used with some care (in particular in the case $m = 1$): the summands $\alpha_i W_i$ do not mean the usual product.

Theorem 1.16. *Let the graph limit Γ be represented by a kernel W on a probability space $(\mathcal{S}, \mu)$. Then Γ is connected if and only if W is.*

As said above, the representation is not unique, but the theorem implies that all representing kernels are connected or disconnected simultaneously.

Example 1.17. Let Γ_p be the graph limit given by the constant kernel $W(x,y) = p$, for some $p \in [0,1]$ and some $(\mathcal{S}, \mu)$. (The graph limit in this case depends on p only, as a consequence of (2.1) below, which yields $t(F, \Gamma_p) = p^{e(F)}$.) By Theorem 1.16, Γ_p is connected for $p > 0$. (These graph limits are the limits of quasi-random sequences of graphs [9], see [17].)

Theorem 1.18. *Let $0 \leq m \leq \infty$ and let $(\Gamma_i)_1^m$ be graph limits and $(\alpha_i)_1^m \in \mathcal{A}$. Suppose that, for each i, the graph limit Γ_i is represented by a kernel W_i on a probability space $(\mathcal{S}_i, \mu_i)$. Then $\bigoplus_{i=1}^m \alpha_i \Gamma_i$ is represented by $\bigoplus_{i=1}^m \alpha_i W_i$.*

1.3. **Random graphs.** A graph limit $\Gamma \in \mathcal{U}_\infty$ defines an infinite random graph $G(\infty, \Gamma)$, which has vertex set $\mathbb{N} = 1, 2, \dots$ and is uniquely determined in the sense that its distribution is. (Again, see Section 2.) Taking the subgraph induced by $[n] := \{1, \dots, n\}$ we obtain finite random graphs $G(n, \Gamma)$, $n = 1, 2, \dots$. These random graphs have the property that $G(n, \Gamma) \to \Gamma$ in $\overline{\mathcal{U}}$ a.s.

For the infinite random graph $G(\infty, \Gamma)$, connectedness is equivalent to connectedness of Γ.

Theorem 1.19. *Let $\Gamma \in \mathcal{U}_\infty$.*

(i) *If Γ is connected, then $G(\infty, \Gamma)$ is a.s. connected.*
(ii) *If Γ is disconnected, then $G(\infty, \Gamma)$ is a.s. disconnected.*

For the (finite) random graph $G(n, \Gamma)$, we cannot expect that connectedness is determined by the connectedness of Γ, not even asymptotically. For example, it is easy to see that if $W = p > 0$ is constant as in Example 1.17, then $G(n, \Gamma) = G(n, p)$ is connected whp, i.e., with probability tending to 1 as $n \to \infty$, (this is one of the earliest results in random graph theory [12], [3]); on the other hand, if $W(x, y) = x^2 y^2$ on $([0, 1], \lambda)$, then $G(n, \Gamma)$ whp has many isolated vertices (at least $n^{1/4}$, considering only those with $X_i < 1.1 n^{-3/4}$ in the construction in Section 2). Instead, there is a corresponding result on the asymptotic size of the largest component. Let $\mathcal{C}_1(G)$ denote the largest component of a graph G (splitting ties arbitrarily).

Theorem 1.20. *Let $\Gamma \in \mathcal{U}_\infty$.*

(i) *Γ is connected if and only if $|\mathcal{C}_1(G(n, \Gamma))|/n \xrightarrow{\text{p}} 1$ as $n \to \infty$.*
(ii) *More precisely, for every Γ, $|\mathcal{C}_1(G(n, \Gamma))|/n \xrightarrow{\text{p}} \rho$ for some $\rho \in [0, 1]$; if Γ is connected, then $\rho = 1$, while if Γ is disconnected, then $0 \leq \rho < 1$. Furthermore, $\rho = 0$ if and only if $\Gamma = \mathbf{0}$ (in which case $G(n, \Gamma) = E_n$, the empty graph). In fact, if Γ has the decomposition $\bigoplus_{i=1}^M \alpha_i \Gamma_i$ into components as in Theorem 1.5, then $\rho = \max_i \{\alpha_i\}$ (with $\rho = 0$ if $M = 0$).*

Remark 1.21. Bollobás, Janson and Riordan [5] study the component sizes in the much sparser random graphs with edge probabilities $O(1/n)$ obtained from $G(n, \Gamma)$ by randomly deleting edges, keeping each edge only with probability c/n for some constant c (the range where a component of order n appears), and Bollobas, Borgs, Chayes and Riordan [4] extend this to a deterministic sequence of graphs $G_n \to \Gamma$; these much more intricate results easily imply (i).

We can also give a complete description of the components of $G(\infty, \Gamma)$. We let $G|_V$, where G is a graph and $V \subseteq V(G)$, denote the induced subgraph of G with vertex set V. (We allow here the possibility that $V = \emptyset$, when the induced subgraph has no vertices. Such cases may be ignored in (ii) in the theorem below, since they only occur for small n, and thus do not affect the limit.)

Theorem 1.22. *Let $\Gamma \in \mathcal{U}_\infty$. Let the random infinite graph $G(\infty, \Gamma)$ have components G_j, $j = 1, \dots, N$ (with $1 \le N \le \infty$), listed in increasing order of the smallest element, say, and let $V_j := V(G_j)$ be the vertex set of G_j. Then, a.s., the following hold for every component G_j.*

- (i) *G_j is either infinite or an isolated vertex. I.e., $|V_j| = v(G_j) \in \{1, \infty\}$.*
- (ii) *If G_j is infinite, then $G_j|_{V_j \cap [n]} \to \Gamma'_j$ as $n \to \infty$ for some (random) $\Gamma'_j \in \mathcal{U}_\infty$.*
- (iii) *The asymptotic density $\nu_j := \lim_{n\to\infty} |V_j \cap [n]|/n$ exists; furthermore, if $|V_j| = \infty$, then $\nu_j > 0$ (and trivially conversely).*
- (iv) *Consider the (random, and finite or infinite) sequence $\{(\Gamma'_j, \nu_j)\}_{j=1}^N$ defined by* (ii) *and* (iii)*. Let Γ have the decomposition $\bigoplus_{i=1}^M \alpha_i \Gamma_i$ into components as in Theorem 1.5. Then, the subsequence $\{(\Gamma'_j, \nu_j) : \nu_j > 0\}$ equals a permutation of $\{(\Gamma_i, \alpha_i)\}_{i=1}^M$.*
- (v) *The set $V_0 := \bigcup_{|V_j|=1} V_j$ of all isolated vertices in $G(\infty, \Gamma)$ has a.s. a density $\nu_0 := \lim_{n\to\infty} |V_0 \cap [n]|/n$, and $\nu_0 = \alpha_0 := 1 - \sum_{i=1}^M \alpha_i$ with α_i as in* (iv)*. Furthermore, $V_0 = \emptyset \iff \alpha_0 = 0 \iff (\alpha_i)_1^M \in \mathcal{A}_1$, i.e., the direct sum $\bigoplus_{i=1}^M \alpha_i \Gamma_i$ is complete.*

The components of the infinite random graph $G(\infty, \Gamma)$ define a random partition of $\mathbb{N}$. Since $G(\infty, \Gamma)$ is exchangeable (i.e., its distribution is invariant under permutations of $\mathbb{N}$, see further [11]), this yields an exchangeable random partition Π_Γ of $\mathbb{N}$. Kingman [15] has shown, see also Bertoin [2, Section 2.3], that the blocks of an exchangeable random partition a.s. have asymptotic densities, so an exchangeable random partition has a (generally random) sequence $(p_i)_1^\infty$ of asymptotic densities of the blocks in the partition; we assume that these are ordered in decreasing order (possibly ignoring or adding 0's), and thus $(p_i)_1^\infty \in \mathcal{A}^\downarrow := \{(p_i)_1^\infty \in \mathcal{A} : p_1 \ge p_2 \ge \dots\}$. Moreover, the (distribution of) $(p_i)_1^\infty$ determines the (distribution of) the exchangeable random partition and every sequence $(p_i)_1^\infty \in \mathcal{A}^\downarrow$ corresponds to an exchangeable random partition of $\mathbb{N}$; the exchangeable random partition can be constructed from $(p_i)_1^\infty$ by the paint-box construction [2, Section 2.3].

Theorem 1.23. *Let the graph limit Γ have the decomposition $\bigoplus_{i=1}^M \alpha_i \Gamma_i$ into components as in Theorem 1.5. Then the sequence $(p_i)_1^\infty$ of asymptotic densities of the exchangeable partition Π_Γ equals a.s. $(\alpha_i)_1^M$ arranged in decreasing order, and extended by 0's if $M < \infty$.*

In this case the asymptotic densities thus are (a.s.) deterministic.

It is also interesting to consider the component of $G(\infty,\Gamma)$ containing a given vertex, which we by symmetry can take to be 1.

Theorem 1.24. *Let $\Gamma \in \mathcal{U}_\infty$, and suppose that Γ has the decomposition $\bigoplus_{i=1}^M \alpha_i\Gamma_i$ into components as in Theorem 1.5.*

(i) *Let G_1 be the component of $G(\infty,\Gamma)$ that contains vertex 1. Then $|G_1|$ is a.s. either 1 or ∞, with $\mathbb{P}(|G_1| = \infty) = \sum_{i=1}^M \alpha_i$ and $\mathbb{P}(|G_1| = 1) = \alpha_0 := 1 - \sum_{i=1}^M \alpha_i$.*

(ii) *Let H be the infinite random graph with $V(H) = \mathbb{N}$ obtained from G_1 by relabelling the vertices in increasing order when $|G_1| = \infty$, and simply taking $H := E_\infty$, the infinite graph on $\mathbb{N}$ with no edges, when $|G_1| = 1$. Then H is an exchangeable infinite random graph, and H has the same distribution as the mixture $\sum_{i=0}^M \alpha_i\mathcal{L}(G(\infty,\Gamma_i))$, where $\Gamma_0 = \mathbf{0}$. In other words, for any measurable set A of infinite graphs on $\mathbb{N}$, $\mathbb{P}(H \in A) = \sum_{i=0}^M \alpha_i \mathbb{P}(G(\infty,\Gamma_i) \in A)$.*

Remark 1.25. The graph limit theory has versions for bipartite graphs and directed graphs too, see [11]. We presume that the definitions and results in the present paper have analogues for these cases too, but we have not pursued this. Furthermore, Bollobás and Riordan [6] have recently started to extend the theory to limits of sparse graphs. We do not know whether our results can be extended to that case.

2. Graph limits

We summarize some basic facts about graph limits that we will use as follows (using the notation of [11]), see Lovász and Szegedy [17], Borgs, Chayes, Lovász, Sós and Vesztergombi [7, 8] and Diaconis and Janson [11] for details and further results.

If F and G are two graphs, let $t(F,G)$ denote 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 $v(G)^{v(F)}$ possible ones; the images of the vertices in F are thus independent and uniformly distributed over $V(G)$.) The basic definition [17; 7] 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 [11] where we further assume $v(G_n) \to \infty$. More precisely, the (countable and discrete) set $\mathcal{U}$ of all unlabelled graphs can be embedded as a dense subspace of a compact metric space $\overline{\mathcal{U}}$ such that a sequence $G_n \in \mathcal{U}$ of graphs with $v(G_n) \to \infty$ 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. We let $\mathcal{U}_\infty := \overline{\mathcal{U}} \setminus \mathcal{U}$ be the set of proper limit elements, and define $v(\Gamma) := \infty$ for $\Gamma \in \mathcal{U}_\infty$. 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 $v(G_n) \to v(\Gamma) = \infty$ and $t(F,G_n) \to t(F,\Gamma)$ for every graph F. (See [7; 8] for deep results giving several other, equivalent, characterizations of $G_n \to \Gamma$; for example the fact that $\overline{\mathcal{U}}$ may be metrized by the cut distance $\delta_\square$ defined there.)

Graph limits $\Gamma \in \mathcal{U}_\infty$ may be represented (non-uniquely) by functions as follows [17], see also [11] for connections to the Aldous–Hoover representation theory for exchangeable arrays [14]. (See [18] and [1] for related results.)

Let $(\mathcal{S}, \mu)$ be an arbitrary probability space and let $W : \mathcal{S} \times \mathcal{S} \to [0, 1]$ be a kernel, i.e., a symmetric measurable function. 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 $\Gamma_W \in \mathcal{U}_\infty$ such that, for every graph F,

$$t(F, \Gamma_W) = \mathbb{E} \prod_{ij \in \mathbb{E}(F)} W(X_i, X_j) = \int_{\mathcal{S}^{v(F)}} \prod_{ij \in \mathbb{E}(F)} W(x_i, x_j) \,\mathrm{d}\mu(x_1) \cdots \mathrm{d}\mu(x_{v(F)}). \tag{2.1}$$

Further, for every $n \geq 1$, let $G(n, W)$ be the random graph with vertex set $[n]$ and edges obtained by, conditionally given $X_1, X_2, \dots, X_n$, taking an edge ij with probability $W(X_i, X_j)$, (conditionally) independently for all pairs (i, j) with $i < j$. Then the random graph $G(n, W)$ converges to Γ_W a.s. as $n \to \infty$.

We may in this construction also take $n = \infty$, with $[\infty] = \mathbb{N}$; this gives a random infinite graph $G(\infty, W)$. Note that $G(n, W)$ is the induced subgraph $G(\infty, W)|_{[n]}$.

Every graph limit in $\mathcal{U}_\infty$ equals Γ_W for some such kernel $W : \mathcal{S} \times \mathcal{S} \to [0, 1]$ on a suitable probability space $(\mathcal{S}, \mu)$; in fact [17], see also [11] and [13], we can always choose $\mathcal{S} = [0, 1]$ equipped with Lebesgue measure λ. (This choice of $(\mathcal{S}, \mu)$ is the standard choice, and often the only one considered, but we find it useful in this paper to be more general.) Note, however, that even if we restrict ourselves to $([0, 1], \lambda)$ only, the representing function W is in general not unique. It is trivial that replacing W by W' with $W' = W$ $\mu \times \mu$-a.e. does not affect the integral in (2.1) and thus not Γ_W. It is equally obvious that a measure-preserving change of variables will not affect Γ. Moreover, this can be extended a little further, and the complete characterization of functions W yielding the same Γ_W is rather subtle, see [6; 7; 11; 14] for details.

In view of this non-uniqueness, we will thus distinguish between the graph limits, being elements of $\mathcal{U}_\infty$, and the functions W that represent them; cf. Remark 1.11.

Given a graph limit $\Gamma \in \mathcal{U}_\infty$, it can, as just said, be represented as Γ_W for some kernel W. Although W is not unique, the distribution of the random graph $G(n, W)$ is the same for all representing W, for every $n \leq \infty$; in fact, for every graph F with vertex set $[k]$ and $k \leq n$, $\mathbb{P}(G(n, W) \supseteq F) = t(F, \Gamma)$. Consequently, for every graph limit Γ there is a well-defined random graph $G(n, \Gamma)$ with n vertices, for every n with $1 \leq n \leq \infty$; this includes the case

$n = \infty$ when $G(\infty, \Gamma)$ is an infinite random graph. (See further Diaconis and Janson [11, Theorem 7.1 and Corollary 5.4]; that paper treats only the case $\mathcal{S} = [0,1]$, but the general case can be proved the same way or by first transferring to $[0,1]$ as in [13].)

We thus have, for every graph F with $V(F) = [k]$ where $1 \le k \le n \le \infty$,

$$\mathbb{P}(G(n,\Gamma) \supseteq F) = t(F,\Gamma). \tag{2.2}$$

If W is a kernel representing Γ, then $G(n,W) \to \Gamma$ a.s., as said above. Hence, for every graph limit Γ,

$$G(n,\Gamma) \to \Gamma \qquad \text{a.s., as } n \to \infty. \tag{2.3}$$

One corollary of this, or of (2.2), is that the distribution of $G(\infty,\Gamma)$ determines Γ: if $G(\infty,\Gamma_1) \overset{\mathrm{d}}{=} G(\infty,\Gamma_2)$ for two graph limits Γ_1, Γ_2, then $\Gamma_1 = \Gamma_2$.

Example 2.1. Let E_n be the empty graph with $V(E_n) = [n]$ and $E(E_n) = \emptyset$. Then $t(F,E_n) = 0$ for any F with $e(F) > 0$, while, as always, $t(F,E_n) = 1$ when $e(F) = 0$. Hence the sequence $(E_n)_n$ converges in $\overline{\mathcal{U}}$, and there is a graph limit $\mathbf{0} \in \mathcal{U}_\infty$ such that $E_n \to \mathbf{0}$; this graph limit is characterized by

$$t(F,\mathbf{0}) = \begin{cases} 0, & e(F) \ge 1, \\ 1, & e(F) = 0. \end{cases} \tag{2.4}$$

The graph limit $\mathbf{0}$ is represented by the kernel $W = 0$ (on any probability space). Hence, $G(n,\mathbf{0}) = E_n$ a.s., for every $n \le \infty$. It is easily seen that $\mathbf{0} = \alpha\mathbf{0} \oplus (1-\alpha)\mathbf{0}$ for any $\alpha \in [0,1]$; hence $\mathbf{0}$ is disconnected.

We may call $\mathbf{0}$ the *empty* or *trivial* graph limit; nevertheless it is useful and important, as is seen in Section 1.

3. Connected test graphs suffice

In Section 2, we defined graph limits and convergence to them using the functionals $t(F,\cdot)$ where F ranges over the set of all (unlabelled) graphs. It turns out that it suffices to use connected graphs F. More precisely, let $\mathcal{U}_{\mathsf{c}} \subset \mathcal{U}$ be the set of all connected unlabelled graphs, and let $\mathcal{U}'_{\mathsf{c}} := \{G \in \mathcal{U}_{\mathsf{c}} : e(G) > 0\} = \mathcal{U}_{\mathsf{c}} \setminus \{K_1\}$ be the subset of unlabelled connected graphs with at least one edge. As the next lemma shows, the functionals $t(F,\cdot)$ for $F \in \mathcal{U}'_{\mathsf{c}}$ are sufficient to characterize graph limits as well as convergence to them.

Lemma 3.1. *A graph limit Γ is uniquely determined by the numbers $t(F,\Gamma)$ for $F \in \mathcal{U}'_{\mathsf{c}}$.*

Moreover, if $\Gamma_1, \Gamma_2, \dots \in \overline{\mathcal{U}}$ is a sequence of graphs or graph limits with $v(\Gamma_n) \to \infty$, and for every $F \in \mathcal{U}'_{\mathsf{c}}$, $t(F,\Gamma_n) \to t_F$ as $n \to \infty$ for some number $t_F \in [0,1]$, then $\Gamma_n \to \Gamma$, where $\Gamma \in \mathcal{U}_\infty$ is the unique graph limit with $t(F,\Gamma) = t_F$, $F \in \mathcal{U}'_{\mathsf{c}}$.

Proof. We begin by observing that if $F = \bigoplus_1^m F_i$, then

$$t(F,\Gamma) = \prod_{i=1}^{m} t(F_i,\Gamma), \qquad \Gamma \in \overline{\mathcal{U}}; \tag{3.1}$$

if $\Gamma = G \in \mathcal{U}$, this follows directly from the definition of $t(F,G)$, and the general case $\Gamma \in \overline{\mathcal{U}}$ follows by continuity.

Now, suppose that $\Gamma, \Gamma' \in \mathcal{U}_\infty$ and that $t(F,\Gamma) = t(F,\Gamma')$ for all $F \in \mathcal{U}'_{\mathsf{c}}$. Since trivially $t(K_1,\Gamma) = 1 = t(K_1,\Gamma')$, the equality holds for all $F \in \mathcal{U}_{\mathsf{c}}$, and thus by decomposing a graph F into components and (3.1), $t(F,\Gamma) = t(F,\Gamma')$ for every graph F, i.e., $\Gamma = \Gamma'$.

This proves the first statement. The second statement is an immediate consequence by compactness and a well-known general argument: Since $\overline{\mathcal{U}}$ is compact, there exist at least subsequences of (Γ_n) that converge. If $\Gamma \in \overline{\mathcal{U}}$ is the limit of such a subsequence, then $v(\Gamma) = \lim v(\Gamma_n) = \infty$, so $\Gamma \in \mathcal{U}_\infty$, and further, $t(F,\Gamma_n) \to t(F,\Gamma)$ along the subsequence for every F, so $t(F,\Gamma) = t_F$ for $F \in \mathcal{U}'_{\mathsf{c}}$. The first part now shows that any two convergent subsequences have the same limit, which in a compact space implies that the entire sequence converges. □

Remark 3.2. It is an interesting and still very much open question to study when a subset of all unlabelled graphs is sufficient to determine graph limits and convergence to them. Lemma 3.1 gives one general result. Other results are known in special cases. For example, if Γ_p is the graph limit in Example 1.17 determined by a kernel W that is constant $p \in [0,1]$, or, equivalently, $t(F,\Gamma_p) = p^{e(F)}$ for every F, and G_n is a sequence of graphs with $v(G_n) \to \infty$, then it suffices that $t(F,G_n) \to t(F,\Gamma_p)$ for the two graphs $F = K_2$ and $F = C_4$ in order that $G_n \to \Gamma_p$ (which is equivalent to the well-known property that G_n is quasirandom) [9], [17]; this is generalized by Lovász and Sós [16] to a larger class of graph limits where a finite set of F suffices. Another example is given by restricting G_n to threshold graphs; in this case it suffices to consider stars F [10].

4. Direct sums

We first see how the functionals $t(F,\cdot)$ behave for direct sums of graphs. This is easier for connected F.

Lemma 4.1. *Suppose that F, G_1 and G_2 are graphs with F connected. Then*

$$t(F,G_1\oplus G_2) = \Bigl(\frac{v(G_1)}{v(G_1)+v(G_2)}\Bigr)^{v(F)} t(F,G_1) + \Bigl(\frac{v(G_2)}{v(G_1)+v(G_2)}\Bigr)^{v(F)} t(F,G_2).$$

Proof. Since F is connected, a mapping $\phi : V(F) \to V(G_1 \oplus G_2)$ is a graph homomorphism $F \to G_1 \oplus G_2$ if and only if ϕ maps $V(F)$ into either $V(G_1)$ or $V(G_2)$, and further ϕ is a graph homomorphism $F \to G_1$ or $F \to G_2$, respectively. If ϕ is a uniformly random mapping $V(F) \to V(G_1 \oplus G_2)$, and $j = 1$ or 2, then the probability that ϕ maps $V(F)$ into $V(G_j)$ is

$\big(v(G_j)/(v(G_1)+v(G_2))\big)^{v(F)}$, and conditioned on this event, ϕ is a graph homomorphism $F \to G_j$ with probability $t(F, G_j)$. $\square$

We use this lemma both as an inspiration and as a tool to define direct sums of graph limits.

Theorem 4.2. (i) *If $\Gamma_1, \Gamma_2 \in \mathcal{U}_\infty$ and $0 \le \alpha \le 1$, then there exists a unique graph limit $\Gamma \in \mathcal{U}_\infty$ such that*

$$t(F,\Gamma) = \alpha^{v(F)} t(F,\Gamma_1) + (1-\alpha)^{v(F)} t(F,\Gamma_2), \qquad F \in \mathcal{U}_{\mathsf{c}}. \tag{4.1}$$

(ii) *More generally, if $(G_i)_{i=1}^m$, where $1 \le m \le \infty$, is a finite or infinite sequence of elements of $\mathcal{U}_\infty$, and $(\alpha_i)_{i=1}^m \in \mathcal{A}$ is a sequence of weights, then there exists a unique graph limit $\Gamma \in \mathcal{U}_\infty$ such that*

$$t(F,\Gamma) = \sum_{i=1}^m \alpha_i^{v(F)} t(F,\Gamma_i), \qquad F \in \mathcal{U}'_{\mathsf{c}}. \tag{4.2}$$

Definition 4.3. The graph limit Γ determined by (4.1) is denoted $\alpha\Gamma_1 \oplus (1-\alpha)\Gamma_2$. More generally, the graph limit Γ determined by (4.2) is denoted $\bigoplus_{i=1}^m \alpha_i \Gamma_i$.

Example 4.4. We can now construct disconnected graphs, even with infinitely many components, as direct sums. For example, let $\Gamma = \bigoplus_1^\infty 2^{-i}\Gamma_{1/i}$, with $\Gamma_{1/i}$ the connected graph limit defined in Example 1.17.

Proof of Theorems 4.2 and 1.1. The uniqueness in Theorem 4.2 follows from Lemma 3.1.

For the existence, we give a proof based on taking limits of graphs, which also proves Theorem 1.1. (An alternative construction is given by kernels and Theorem 1.18.)

Assume first that $m < \infty$ and $\sum_{i=1}^m \alpha_i = 1$. Assume further that $G_{in} \in \mathcal{U}$, $1 \le i \le m$ and $n \ge 1$, are such that, as $n \to \infty$, $v(G_{in}) \to \infty$ and

$$G_{in} \to \Gamma_i, \qquad 1 \le i \le m, \tag{4.3}$$

$$\frac{v(G_{in})}{\sum_j v(G_{jn})} \to \alpha_i, \qquad 1 \le i \le m. \tag{4.4}$$

Note that such graphs G_{in} always may be found. For example, by (2.3), there exist $H_{in} \in \mathcal{U}$ with $v(H_{in}) = n$ and $H_{in} \to \Gamma_i$ as $n \to \infty$; we may then take $G_{in} := H_{i,\lceil n\alpha_i \rceil}$ if $\alpha_i > 0$ and, e.g., $G_{in} := H_{i,\lfloor \log n \rfloor + 1}$ if $\alpha_i = 0$.

Lemma 4.1 extends immediately (e.g. by induction) to disjoint sums of several graphs, which yields, for every $F \in \mathcal{U}_{\mathsf{c}}$, using (4.3) and (4.4),

$$t\Big(F, \bigoplus_{i=1}^m G_{in}\Big) = \sum_{i=1}^m \bigg(\frac{v(G_{in})}{\sum_j v(G_{jn})}\bigg)^{v(F)} t(F, G_{in}) \to \sum_{i=1}^m \alpha_i^{v(F)} t(F,\Gamma_i).$$

This shows by Lemma 3.1 that $\bigoplus_{i=1}^m G_{in} \to \Gamma$ for some $\Gamma \in \overline{\mathcal{U}}$ that satisfies (4.2).

The special case $m = 2$ yields (4.1) and, taking G_{1n} and G_{2n} as the given G_n and G'_n, Theorem 1.1.

In general, we have shown the existence of $\Gamma = \bigoplus_{i=1}^m \alpha_i \Gamma_i$ whenever $m < \infty$ and $\sum_{i=1}^m \alpha_i = 1$.

Next, assume $m < \infty$ and $\sum_{i=1}^m \alpha_i < 1$. Let $\alpha_0 := 1 - \sum_{i=1}^m \alpha_i$ and let $\Gamma_0 := \mathbf{0}$ be as in Example 2.1. By the case just shown, $\Gamma = \sum_{i=0}^m \alpha_i \Gamma_i$ exists in $\mathcal{U}_\infty$. However, $t(F, \Gamma_0) = 0$ for $F \in \mathcal{U}'_{\mathsf{c}}$, so (4.2) shows that $\Gamma = \sum_{i=1}^m \alpha_i \Gamma_i$ too.

Finally, if $m = \infty$, define $\Gamma_{(n)} := \sum_{i=1}^n \alpha_i \Gamma_i$. By (4.2), as $n \to \infty$,

$$t\bigl(F, \Gamma_{(n)}\bigr) = \sum_{i=1}^n \alpha_i^{v(F)} t(F, \Gamma_i) \to \sum_{i=1}^\infty \alpha_i^{v(F)} t(F, \Gamma_i), \qquad F \in \mathcal{U}'_{\mathsf{c}}. \tag{4.5}$$

Hence, using Lemma 3.1, $\Gamma_{(n)} \to \Gamma$ for some Γ satisfying (4.2), i.e., $\Gamma = \sum_{i=1}^\infty \alpha_i \Gamma_i$. □

5. Remaining proofs

Proof of Theorem 1.18. Let $W := \bigoplus_{i=1}^m \alpha_i W_i$. If $\sum_{i=1}^m \alpha_i < 1$, we first rewrite this as a complete direct sum $W = \bigoplus_{i=0}^m \alpha_i W_i$ by Definition 1.13(ii).

For any $F \in \mathcal{U}'_{\mathsf{c}}$, it follows by Definition 1.13 that the integrand in (2.1) is non-zero only if all x_k belong to the same $\mathcal{S}_i$, and further $i \neq 0$; moreover, each $i > 0$ gives a contribution $\alpha_i^{v(F)} t(F, \Gamma_{W_i})$. Thus $t(F, \Gamma_W) = \sum_{i=1}^m \alpha_i^{v(F)} t(F, \Gamma_{W_i})$ when $F \in \mathcal{U}'_{\mathsf{c}}$, which completes the proof by (4.2) and Definition 4.3. □

Lemma 5.1. *Let W be a kernel on a probability space $(\mathcal{S}, \mu) = (\mathcal{S}, \mathcal{F}, \mu)$. Define an operator $\mathcal{N}$ on the σ-algebra $\mathcal{F}$ by*

$$\mathcal{N}(A) := \Bigl\{x : \int_A W(x, y) \,\mathrm{d}\mu(y) > 0\Bigr\}, \qquad A \in \mathcal{F}.$$

Further, for a point $x \in \mathcal{S}$, define

$$\mathcal{N}_x := \{y : W(x, y) > 0\}.$$

If W is connected, the following holds.

(i) *For μ-a.e. $x \in \mathcal{S}$, $\mu(\mathcal{N}_x) > 0$.*
(ii) *If $\mu(A) > 0$, then $\mu(\mathcal{N}(A)) > 0$.*
(iii) *If $\mathcal{N}(A) \subseteq A$, then $\mu(A) = 0$ or $\mu(A) = 1$.*
(iv) *If $A_1, A_2, \ldots \in \mathcal{F}$, then $\mathcal{N}\bigl(\bigcup_{n=1}^\infty A_n\bigr) = \bigcup_{n=1}^\infty \mathcal{N}(A_n)$.*
(v) *If $\mu(A) > 0$, then $\mu\bigl(\bigcup_{n=1}^\infty \mathcal{N}^n(A)\bigr) = 1$.*

Proof. First, note that for every $A \in \mathcal{F}$, if $x \notin \mathcal{N}(A)$, then $W(x, y) = 0$ for a.e. $y \in A$, and thus by Fubini $W = 0$ a.e. on $(\mathcal{S} \setminus \mathcal{N}(A)) \times A$. Hence, by the symmetry of W,

$$W = 0 \text{ a.e. on } A \times (\mathcal{S} \setminus \mathcal{N}(A)), \qquad A \in \mathcal{F}. \tag{5.1}$$

Next, suppose that $A \in \mathcal{F}$ is a subset of $\mathcal{S}$ such that $W = 0$ a.e. on $A \times \mathcal{S}$. If $\mu(A) = 1$, then $W = 0$ a.e. on $\mathcal{S} \times \mathcal{S}$, which by Definition 1.12 contradicts the assumption that W is connected. Furthermore, $W = 0$ a.e. on $A \times (\mathcal{S} \setminus A)$, which if $0 < \mu(A) < 1$ again contradicts Definition 1.12. Consequently:

$$A \in \mathcal{F} \text{ and } W = 0 \text{ a.e. on } A \times \mathcal{S} \implies \mu(A) = 0. \tag{5.2}$$

(i): Let $B := \{x : \mu(\mathcal{N}_x) = 0\}$. If $x \in B$, then $W(x, y) = 0$ for a.e. $y \in \mathcal{S}$, and thus, by Fubini, $W = 0$ a.e. on $B \times \mathcal{S}$. Hence, (5.2) yields $\mu(B) = 0$.

(ii): If $\mu(\mathcal{N}(A)) = 0$, then (5.1) implies $W = 0$ a.e. on $A \times \mathcal{S}$, and thus (5.2) yields $\mu(A) = 0$.

(iii): By (5.1), $W = 0$ a.e. on $A \times (\mathcal{S} \setminus \mathcal{N}(A)) \supseteq A \times (\mathcal{S} \setminus A)$. If $0 < \mu(A) < 1$, this means by Definition 1.12 that W is disconnected, a contradiction.

(iv): Clearly, for every x,

$$\begin{aligned} x \notin \mathcal{N}\Bigl(\bigcup_{n=1}^{\infty} A_n\Bigr) &\iff W(x, y) = 0 \text{ for a.e. } y \in \bigcup_n A_n \\ &\iff \text{for every } n,\ W(x, y) = 0 \text{ for a.e. } y \in A_n \\ &\iff \text{for every } n,\ x \notin \mathcal{N}(A_n). \end{aligned}$$

(v): Let $\widehat{A} := \bigcup_{n=1}^{\infty} \mathcal{N}^n(A)$. Then, by (iv), $\mathcal{N}(\widehat{A}) = \bigcup_{n=1}^{\infty} \mathcal{N}^{n+1}(A) \subseteq \widehat{A}$. Further, by (ii), $\mu(\widehat{A}) \geq \mu(\mathcal{N}(A)) > 0$. Hence (iii) implies $\mu(\widehat{A}) = 1$. $\square$

Lemma 5.2. *If W is a connected kernel, then $G(\infty, W)$ is a.s. connected.*

Proof. Recall the construction of $G(\infty, W)$ in Section 2 using an i.i.d. sequence $(X_i)_1^{\infty}$; we now do this construction by adding one vertex i at a time, each time randomly choosing first X_i and then, for each $j < i$, whether ji is an edge or not.

Let $x := X_1$. By Lemma 5.1(i), a.s. $\mu(\mathcal{N}_x) > 0$; we assume this in the sequel. Then, by Lemma 5.1(v), $\mu\bigl(\bigcup_{n=1}^{\infty} \mathcal{N}^n(\mathcal{N}_x)\bigr) = 1$, and thus a.s. $X_2 \in \bigcup_{n=1}^{\infty} \mathcal{N}^n(\mathcal{N}_x)$. We assume that this happens and choose an n (depending on X_1 and X_2) such that $X_2 \in \mathcal{N}^n(\mathcal{N}_x)$. Let $i_n := 2$ and $x_n := X_2$.

Assume first that $n > 1$. For each new vertex i, the probability that $X_i \in \mathcal{N}^{n-1}(\mathcal{N}_x)$ and that there is an edge $i_n i$ equals $\int_{\mathcal{N}^{n-1}(\mathcal{N}_x)} W(x_n, y) \,\mathrm{d}\mu(y)$, which is > 0 because $x_n \in \mathcal{N}^n(\mathcal{N}_x)$. Hence, a.s., there exists some such $i > i_n$; let i_{n-1} be the first such i and let $x_{n-1} := X_{i_{n-1}}$.

Repeating $n - 1$ times, we a.s. find vertices $2 = i_n < i_{n-1} < \cdots < i_1$ that are connected to a path by edges in $G(\infty, W)$, and with $x_1 := X_{i_1} \in \mathcal{N}(\mathcal{N}_x)$.

Finally, for each new vertex $i > i_1$, the probability that there are edges $i_1 i$ and $1i$ equals $\int_{\mathcal{S}} W(x_1, y) W(x, y) \,\mathrm{d}\mu(y)$, which is > 0 because $x_1 \in \mathcal{N}(\mathcal{N}_x)$, and thus there is a set $A \subseteq \mathcal{N}_x$ of positive measure with $W(x_1, y) > 0$ for $y \in A$, and $W(x, y) > 0$ too for $y \in A$ by the definition of $\mathcal{N}_x$. Consequently, a.s. there exists some such vertex i, which means that there is a path $1, i, i_1, \ldots, i_n = 2$ in $G(\infty, W)$.

We have shown that a.s. the vertices 1 and 2 can be connected by a path in $G(\infty, W)$. By symmetry (exchangeability), the same is true for any given pair of vertices i and j. Since there is only a countable number of vertices, a.s. $G(\infty, W)$ is connected. $\square$

Lemma 5.3. *If W is a disconnected kernel, then Γ_W is disconnected.*

Proof. If $W = 0$ a.e., then $\Gamma_W = \mathbf{0}$, which is disconnected, see Example 2.1.

By Definition 1.12, the other possibility is that there exists $A \subset \mathcal{S}$ with $0 < \mu(A) < 1$ and $W = 0$ a.e. on $A \times (\mathcal{S} \setminus A)$. Let $\mathcal{S}_1 := A$ and $\mathcal{S}_2 := \mathcal{S} \setminus A$; let further, for $j = 1, 2$, $\alpha_j := \mu(\mathcal{S}_j)$ and $\mu_j := \alpha_j^{-1} \mu|_{\mathcal{S}_j}$, and let $W_j := W|_{\mathcal{S}_j \times \mathcal{S}_j}$, considered as a kernel on $(\mathcal{S}_j, \mu_j)$. Then $W^* := \alpha_1 W_1 \oplus \alpha_2 W_2$ is a kernel defined on $(\mathcal{S}, \mu)$ and $W^* = W$ a.e. (More precisely, W^* equals W modified to be identically 0 on $A \times (\mathcal{S} \setminus A)$ and $(\mathcal{S} \setminus A) \times A$.) Hence, by (2.1), $\Gamma_W = \Gamma_{W^*}$. Consequently, using Theorem 1.18,

$$\Gamma_W = \Gamma_{W^*} = \Gamma_{\alpha_1 W_1 \oplus \alpha_2 W_2} = \alpha_1 \Gamma_{W_1} \oplus \alpha_2 \Gamma_{W_2},$$

and thus Γ_W is disconnected. (Recall that $\alpha_1 = \mu(A) \in (0, 1)$.) $\square$

Lemma 5.4. *If $\Gamma \in \mathcal{U}_\infty$ is disconnected, then $G(\infty, \Gamma)$ is disconnected a.s.*

Proof. By Definition 1.2, $\Gamma = \alpha \Gamma_1 \oplus (1 - \alpha)\Gamma_2$ where $0 < \alpha < 1$. Choose a kernel W_j, on a probability space $(\mathcal{S}_j, \mu_j)$, that represents Γ_j $(j = 1, 2)$ and assume as we may that $\mathcal{S}_1$ and $\mathcal{S}_2$ are disjoint. By Theorem 1.18, Γ is represented by the kernel $W := \alpha W_1 \oplus (1 - \alpha) W_2$ on $\mathcal{S} := \mathcal{S}_1 \cup \mathcal{S}_2$, and thus

$$G(\infty, \Gamma) = G(\infty, W) = G(\infty, \alpha W_1 \oplus (1 - \alpha) W_2).$$

However, it is evident from the construction of $G(\infty, W)$ in Section 2 that there are no edges between $V_1 := \{i : X_i \in \mathcal{S}_1\}$ and $V_2 := \{i : X_i \in \mathcal{S}_2\}$, and that these sets are a.s. non-empty; hence $G(\infty, \Gamma) = G(\infty, W)$ is a.s. disconnected. $\square$

Proof of Theorem 1.16. If W is connected, then $G(\infty, \Gamma) = G(\infty, W)$ is a.s. connected by Lemma 5.2, and thus by Lemma 5.4, Γ cannot be disconnected, i.e., Γ is connected.

Conversely, if W is disconnected, then $\Gamma = \Gamma_W$ is disconnected by Lemma 5.3. $\square$

Proof of Theorem 1.19. Let W be a kernel representing Γ. If Γ is connected, then W is connected by Theorem 1.16 and thus $G(\infty, \Gamma) = G(\infty, W)$ is a.s. connected by Lemma 5.2.

If Γ is disconnected, then $G(\infty, \Gamma)$ is a.s. disconnected by Lemma 5.4. $\square$

Lemma 5.5. *Let W be a kernel on a probability space $(\mathcal{S}, \mu)$. Then there is a decomposition $\mathcal{S} = \bigcup_{i=0}^M \mathcal{S}_i$ into disjoint measurable sets $\mathcal{S}_i$, where $0 \leq M \leq \infty$ and $\mathcal{S}_0$ may be empty but $\alpha_i := \mu(\mathcal{S}_i) > 0$ for $i \geq 1$, such that if $W_i := W|_{\mathcal{S}_i \times \mathcal{S}_i}$ and $\mu_i := \alpha_i^{-1} \mu|_{\mathcal{S}_i}$ $(i \geq 1)$, then W_i is a connected kernel on $(\mathcal{S}_i, \mu_i)$ for $i \geq 1$, and $W = 0$ a.e. on $(\mathcal{S} \times \mathcal{S}) \setminus \bigcup_{i=1}^M (\mathcal{S}_i \times \mathcal{S}_i)$; hence $\bigoplus_{i=1}^M \alpha_i W_i = W$ a.e.*

Proof. This is Bollobás, Janson and Riordan [5, Lemma 5.17] expressed in the present terminology and notation, except that the definitions there allow $W_i = 0$ a.e. on $\mathcal{S}_i \times \mathcal{S}_i$ for some $i \geq 1$; however, any such term may be deleted and $\mathcal{S}_i$ included in $\mathcal{S}_0$. $\square$

Proof of Theorem 1.5, existence. Represent Γ by a kernel W. By Lemma 5.5, $W = \bigoplus_{i=1}^M \alpha_i W_i$ a.e. for some sequences $(W_i)_1^M$ of connected kernels and $(\alpha_i)_1^M \in \mathcal{A}_+$, and thus Γ is represented by $\bigoplus_{i=1}^M \alpha_i W_i$ too.

Let $\Gamma_i := \Gamma_{W_i}$, and note that Γ_i is connected by Theorem 1.16. By Theorem 1.18, $\bigoplus_{i=1}^M \alpha_i \Gamma_i$ is represented by the same kernel $\bigoplus_{i=1}^M \alpha_i W_i$ as Γ, and thus $\Gamma = \bigoplus_{i=1}^M \alpha_i \Gamma_i$. $\square$

The proof of uniqueness is more complicated because the kernel W is not unique, and consequently it is not enough to show uniqueness of the decomposition of W in Lemma 5.5. To avoid having to use the deep and rather subtle criteria for equivalence of kernels, we use instead the random infinite graph $G(\infty, \Gamma)$, which is uniquely determined by Γ (in the sense that its distribution is determined); we thus study this further first.

Proof of Theorem 1.22. Let $\Gamma = \bigoplus_{i=1}^M \alpha_i \Gamma_i$ be a decomposition as in Theorem 1.5; recall that we have proved existence of this (but not yet uniqueness). Choose a kernel W_i, on a probability space $(\mathcal{S}_i, \mu_i)$, that represents Γ_i $(i \geq 1)$. By Theorem 1.16, W_i is connected, and by Theorem 1.18, Γ is represented by $W := \bigoplus_{i=1}^M \alpha_i W_i$ on $(\mathcal{S}, \mu)$, where $\mathcal{S} = \bigcup_0^\infty \mathcal{S}_i$ (assuming as we may that the sets $\mathcal{S}_i$ are disjoint, and allowing $\mathcal{S}_0 = \emptyset$). In the construction of $G(\infty, \Gamma) = G(\infty, W)$ in Section 2, let $\widetilde{V}_i := \{k : X_k \in \mathcal{S}_i\}$ and let $\widetilde{G}_i$ be the induced subgraph $G(\infty, W)|_{\widetilde{V}_i}$. (We here allow graphs with empty vertex set.) Since $W(x,y) = 0$ when $x \in \mathcal{S}_i$ and $y \in \mathcal{S}_j$ with $i \neq j$, there are no edges in $G(\infty, W)$ between $\widetilde{V}_i$ and $\widetilde{V}_j$; thus $G(\infty, W) = \bigoplus_{i=0}^M \widetilde{G}_i$. Furthermore, $W(x,y) = 0$ for $x, y \in \mathcal{S}_0$ too, so $E(\widetilde{G}_0) = \emptyset$ and all vertices in $\widetilde{V}_0$ (if any) are isolated.

By the law of large numbers, a.s. each $\widetilde{V}_i$ has an asymptotic density

$$\lim_{n\to\infty} |\widetilde{V}_i \cap [n]|/n = \mathbb{P}(X_1 \in \mathcal{S}_i) = \mu(\mathcal{S}_i) = \alpha_i, \qquad i \geq 0.$$

In particular, for $i \geq 1$, since then $\alpha_i > 0$, $|\widetilde{V}_i| = \infty$ a.s. Moreover, the subsequence $\{X_k : k \in \widetilde{V}_i\}$ is a sequence of i.i.d. elements of $\mathcal{S}_i$ with the distribution μ_i; hence the induced subgraph $\widetilde{G}_i$ $(i \geq 1)$ equals (in distribution), if we relabel the vertices in increasing order as $1, 2, \dots$, the infinite random graph $G(\infty, W_i) = G(\infty, \Gamma_i)$. In particular, by Theorem 1.19, a.s. each $\widetilde{G}_i$, $i \geq 1$, is connected. Consequently, the components of $G(\infty, \Gamma) = G(\infty, W)$ are a.s. given by $\widetilde{G}_i$, $i \geq 1$, and the vertices in $\widetilde{V}_0$, the latter being the isolated vertices of $G(\infty, W)$. Moreover, for $i \geq 1$, with $n_i(n) := |\widetilde{V}_i \cap [n]|$, as $n \to \infty$ and thus $n_i(n) \to \infty$,

$$\widetilde{G}_i|_{\widetilde{V}_i \cap [n]} = G\bigl(n_i(n), W_i\bigr) \to \Gamma_i.$$

We have shown that $\{(G_j, V_j)\}$ a.s. is a permutation of

$$\{(\widetilde{G}_i, \widetilde{V}_i) : i \geq 1\} \cup \{(K_1(x), \{x\}) : x \in \widetilde{V}_0\},$$

where $K_1(x)$ denotes the graph with vertex set $\{x\}$ (and thus no edges). The results follow from the results just proven for $\widetilde{G}_i$ and $\widetilde{V}_i$. □

Proof of Theorem 1.5, uniqueness and components. By Theorem 1.22(iv), the sequence $\{(\Gamma_i, \alpha_i)\}_{i=1}^M$ is a.s. a permutation of the sequence $\{(\Gamma'_j, \nu_j) : \nu_j > 0\}$ constructed there from the components of $G(\infty, \Gamma)$; hence the sequence is determined by Γ up to permutation. (In particular, M is determined as being a.s. the number of infinite components in $G(\infty, \Gamma)$.)

Since it follows from (4.2) that the direct sum operation for graph limits is associative in the natural way, we have $\Gamma = \bigoplus_{i=1}^M \alpha_i \Gamma_i = \alpha_1 \Gamma_1 \oplus \bigoplus_{i=2}^M \alpha_i \Gamma_i$, and thus Γ_1 is a component of Γ, and similarly every Γ_i, $i \geq 1$.

Conversely, if Γ' is a component of Γ, then $\Gamma = \alpha\Gamma' \oplus (1-\alpha)\Gamma''$ for some $\alpha > 0$ and some $\Gamma'' \in \mathcal{U}_\infty$. Decomposing $\Gamma'' = \bigoplus_{i=2}^{M''} \alpha''_i \Gamma''_i$ (by the existence part already proven and relabelling), we find (again by associativity) a decomposition $\Gamma = \bigoplus_{i=1}^{M''} \alpha'_i \Gamma'_i$ with $\Gamma'_1 = \Gamma'$, $\Gamma'_i = \Gamma''_i$ for $i \geq 2$, $\alpha'_1 = \alpha$ and $\alpha'_i = (1-\alpha)\alpha''_i$, $i \geq 2$; hence all $\alpha'_i > 0$ and all Γ'_i are connected. By the uniqueness just proved, $\Gamma' = \Gamma'_1 = \Gamma_i$ for some i. □

Proof of Theorem 1.23. A corollary of Theorem 1.22(iii)(iv), since the partition is $\{V_j\}$. □

Proof of Theorem 1.24. Another corollary of Theorem 1.22 and its proof: In the notation above, $G_1 = \widetilde{G}_i$ and $H \overset{\mathrm{d}}{=} G(\infty, \Gamma_i)$ if $X_1 \in \mathcal{S}_i$, $i \geq 1$, and $G_1 = K_1$ and $H = E_\infty = G(\infty, \Gamma_0)$ if $X_1 \in \mathcal{S}_0$; furthermore, $\mathbb{P}(X_1 \in \mathcal{S}_i) = \mu(\mathcal{S}_i) = \alpha_i$. □

Lemma 5.6. *If $\Gamma \in \mathcal{U}_\infty$ is connected, then $|\mathcal{C}_1(G(n, \Gamma))|/n \xrightarrow{\mathrm{p}} 1$.*

Proof. Let $\mathcal{E}_n$ be the event that vertices 1 and 2 are connected by a path in $G(n, W)$. Then $\mathcal{E}_n \uparrow \mathcal{E}_\infty$ and $\mathbb{P}(\mathcal{E}_\infty) = 1$ by Theorem 1.19, and thus $\mathbb{P}(\mathcal{E}_n) \to 1$.

Let $\mathcal{C}_1(n) := \mathcal{C}_1(G(n, \Gamma))$ and let $G_1(n)$ be the component of $G(n, W)$ that contains vertex 1. Then $|\mathcal{C}_1(n)| \geq |G_1(n)|$ and, using the symmetry,

$$\mathbb{E}(n - |\mathcal{C}_1(n)|) \leq \mathbb{E}(n - |G_1(n)|) = (n-1)\,\mathbb{P}(\mathcal{E}_n^c) = (n-1)(1 - \mathbb{P}(\mathcal{E}_n)) = o(n).$$

Thus $\mathbb{E}(1 - |\mathcal{C}_1(n)|/n) \to 0$. Hence $1 - |\mathcal{C}_1(n)|/n \xrightarrow{\mathrm{p}} 0$, i.e. $|\mathcal{C}_1(n)|/n \xrightarrow{\mathrm{p}} 1$. □

Proof of Theorem 1.20. Let $\widetilde{V}_i$, $\widetilde{G}_i$ and $n_i(n)$ be as in the proof of Theorem 1.22, and let $\widetilde{G}_i[n] := \widetilde{G}_i|_{[n]}$. (We allow this subgraph to lack vertices.) Then $G(n, \Gamma)$ is a.s. the direct sum of $\widetilde{G}_i[n]$, $i \geq 1$, together with the isolated vertices in $\widetilde{V}_0 \cap [n]$ (if any); hence

$$|\mathcal{C}_1(G(n, \Gamma))| = 1 \vee \max_{i \geq 1} |\mathcal{C}_1(\widetilde{G}_i[n])|. \tag{5.3}$$

(Note that $\widetilde{G}_i[n]$ does not have to be connected.)

Since $\widetilde{G}_i \overset{\mathrm{d}}{=} G(\infty,\Gamma_i)$ and $v(\widetilde{G}_i[n]) = n_i(n)$ by the proof of Theorem 1.22, we have, conditioned on $n_i(n)$, $\widetilde{G}_i[n] \overset{\mathrm{d}}{=} G(n_i(n),\Gamma_i)$, and thus by Lemma 5.6 applied to the connected graph limit Γ_i (and considering only n with $n_i(n) \geq 1$)

$$\frac{|\mathcal{C}_1(\widetilde{G}_i[n])|}{n} = \frac{|\mathcal{C}_1(\widetilde{G}_i[n])|}{n_i(n)} \cdot \frac{n_i(n)}{n} \overset{\mathrm{p}}{\longrightarrow} 1 \cdot \alpha_i = \alpha_i. \tag{5.4}$$

Let $\rho := \max_i \alpha_i$ (with $\rho = 0$ if $M = 0$) and let $\varepsilon > 0$. If $\rho > 0$, choose i such that $\alpha_i = \rho$; then (5.3) and (5.4) yield

$$|\mathcal{C}_1(G(n,\Gamma))|/n \geq |\mathcal{C}_1(\widetilde{G}_i[n])|/n > \rho - \varepsilon \qquad \text{whp.} \tag{5.5}$$

(If $\rho = 0$, this is trivial.)

On the other hand, for every i, $n_i(n)/n \overset{\mathrm{p}}{\longrightarrow} \alpha_i$, and thus whp

$$|\mathcal{C}_1(\widetilde{G}_i[n])| \leq v(\widetilde{G}_i[n]) = n_i(n) < (\alpha_i + \varepsilon)n \leq (\rho+\varepsilon)n. \tag{5.6}$$

Choose $M' < \infty$ such that $\sum_{i>M'} \alpha_i < \varepsilon$. (If $M < \infty$ we may simply take $M' = M$.) Then, similarly, by the law of large numbers, $\sum_{i>M'} n_i(n)/n \overset{\mathrm{p}}{\longrightarrow} \sum_{i>M'} \alpha_i < \varepsilon$, and thus whp, for all $i > M'$ simultaneously,

$$|\mathcal{C}_1(\widetilde{G}_i[n])| \leq v(\widetilde{G}_i[n]) = n_i(n) \leq \sum_{i>M'} n_i(n) < \varepsilon n.$$

Since whp (5.6) holds for every $i \leq M'$, we find that whp $|\mathcal{C}_1(\widetilde{G}_i[n])|/n \leq \rho + \varepsilon$ for all i, and thus (5.3) implies $|\mathcal{C}_1(G(n,\Gamma))|/n \leq \rho+\varepsilon$ whp.

Since $\varepsilon > 0$ is arbitrary, this and (5.5) show that $|\mathcal{C}_1(G(n,\Gamma))|/n \overset{\mathrm{p}}{\longrightarrow} \rho$ as asserted. Furthermore, it is clear that $\rho = 1$ if and only if $M = 1$ and $\alpha_1 = 1$, which holds if and only if Γ is connected, cf. Remark 1.9. Similarly, $\rho = 0$ if and only if $M = 0$, which holds if and only if $\Gamma = \mathbf{0}$, cf. Remark 1.8. ☐

Finally we consider the minimal sizes of cuts, Theorem 1.10. This proof differs from the others in this section and does not use kernels or $G(\infty,\Gamma)$; instead it uses, not surprisingly, the cut distance [7].

Proof of Theorem 1.10. (i): Suppose that the conclusion fails; we will show that then Γ is disconnected. Thus we assume that there exists $\delta > 0$ such that for every $\varepsilon > 0$ there exists a subsequence $\mathcal{N}_\varepsilon \subseteq \mathbb{N}$ along which there exists partitions $V(G_n) = V' \cup V''$ with $|V'|, |V''| \geq \delta v(G_n)$ and $e(V',V'') \leq \varepsilon v(G_n)^2$. Choosing n_k in the subsequence $\mathcal{N}_{1/k}$, $k \geq 1$, and such that $n_k > n_{k-1}$ for $k \geq 2$, we obtain a subsequence (G_{n_k}) which we relabel as (G_n); then for every n there is a partition $V(G_n) = V' \cup V''$ with $|V'|, |V''| \geq \delta v(G_n)$ and $e(V',V'') = o\big(v(G_n)^2\big)$.

Let $G'_n := G_n|_{V'}$, $G''_n := G_n|_{V''}$, and $G^*_n := G'_n \oplus G''_n$, i.e., G_n with the edges between V' and V'' deleted. Then $V(G^*_n) = V(G_n)$ and

$$|E(G^*_n)\Delta E(G_n)| = e(V',V'') = o\big(v(G_n)^2\big). \tag{5.7}$$

It is easy to see that this implies, for any graph F,

$$|t(F, G_n^*) - t(F, G_n)| \leq v(F)^2 \frac{|E(G_n^*)\Delta E(G_n)|}{v(G_n)^2} = o(1).$$

Since $G_n \to \Gamma$, i.e. $t(F, G_n) \to t(F, \Gamma)$, we obtain $G_n^* \to \Gamma$ in $\overline{\mathcal{U}}$. (Alternatively, (5.7) immediately yields, in the notation of [7], $\delta_\square(G_n^*, G_n) \leq d_\square(G_n^*, G_n) = o(1)$, and thus $G_n^* \to \Gamma$ by [7, Theorem 2.6].)

Note that $v(G_n'), v(G_n'') \to \infty$ and $\delta \leq v(G_n')/v(G_n) \leq 1-\delta$. By the compactness of $\overline{\mathcal{U}}$ and $[\delta, 1-\delta]$, we may select a subsequence such that, along the subsequence, $G_n' \to \Gamma'$ and $G_n'' \to \Gamma''$ for some $\Gamma', \Gamma'' \in \mathcal{U}_\infty$, and further $v(G_n')/v(G_n) \to \alpha \in [\delta, 1-\delta]$. By Theorem 1.1, we thus have (still along the subsequence)

$$G_n^* := G_n' \oplus G_n'' \to \alpha\Gamma' \oplus (1-\alpha)\Gamma''.$$

Since also, as shown above, $G_n^* \to \Gamma$, we conclude that $\Gamma = \alpha\Gamma' \oplus (1-\alpha)\Gamma''$, and thus Γ is disconnected.

(ii): Suppose that Γ is disconnected; then $\Gamma = \alpha\Gamma' \oplus (1-\alpha)\Gamma''$ for some $\Gamma', \Gamma'' \in \mathcal{U}_\infty$. Choose graphs H_n', H_n'' with $v(H_n') = v(H_n'') = n$ and $H_n' \to \Gamma'$, $H_n'' \to \Gamma''$ as $n \to \infty$. Further, let $n'(n) := \lfloor \alpha v(G_n) \rfloor$ and $n''(n) := v(G_n) - n'(n)$, and consider only n so large that $n'(n), n''(n) \geq 1$. Then, $H_n := H'_{n'(n)} \oplus H''_{n''(n)}$ has the same number of vertices as G_n and, by Theorem 1.1, $H_n \to \alpha\Gamma' \oplus (1-\alpha)\Gamma'' = \Gamma$.

It follows that the combined sequence $G_1, H_1, G_2, H_2, \dots$ converges (to Γ), and thus by Borgs, Chayes, Lovász, Sós and Vesztergombi [7, Theorem 2.6] $\delta_\square(G_n, H_n) \to 0$, where $\delta_\square$ is the cut distance defined in [7]. Moreover, by [7, Theorem 2.3], this implies $\widehat{\delta}_\square(G_n, H_n) \to 0$, where

$$\begin{aligned}\widehat{\delta}_\square(G_n, H_n) &:= \min_{\widetilde{H}_n \cong H_n} d_\square(G_n, H_n) \\ &:= \min_{\widetilde{H}_n \cong H_n} \max_{S,T \subseteq V(G_n)} \frac{|e_{G_n}(S,T) - e_{\widetilde{H}_n}(S,T)|}{v(G_n)^2}, \end{aligned} \tag{5.8}$$

taking the minima over $\widetilde{H}_n \cong H_n$ with the same vertex set $V(G_n)$ as G_n. Fix a $\widetilde{H}_n$ that achives the minimum in (5.8). Since H_n can be partitioned into two parts with $n'(n)$ and $n''(n)$ vertices and no edges in between, the same holds for $\widetilde{H}_n$, and thus there exists a partition $V(G_n) = V' \cup V''$ with $|V'|/v(G_n) = n'(n)/v(G_n) \to \alpha > 0$, $|V''|/v(G_n) = n''(n)/v(G_n) \to 1-\alpha > 0$, $e_{\widetilde{H}_n}(V', V'') = 0$ and thus, by (5.8),

$$e_{G_n}(V', V'') \leq d_\square(G_n, \widetilde{H}_n) v(G_n)^2 = o\big(v(G_n)^2\big),$$

which proves the result for any $\delta < \min(\alpha, 1-\alpha)$. $\square$

References

[1] T. D. Austin, On exchangeable random variables and the statistics of large graphs and hypergraphs. Preprint, 2007. http://arxiv.org/0801.1698v1

[2] J. Bertoin, *Random fragmentation and coagulation processes.* Cambridge Univ. Press, Cambridge, 2006.

[3] B. Bollobás, *Random Graphs*, 2nd ed., Cambridge Univ. Press, Cambridge, 2001.

[4] B. Bollobas, C. Borgs, J. Chayes & O. Riordan, Percolation on dense graph sequences. Preprint, 2007. http://arxiv.org/math/0701346v2

[5] B. Bollobás, S. Janson and O. Riordan, The phase transition in inhomogeneous random graphs. *Random Struct. Alg.* **31** (2007), 3–122.

[6] B. Bollobás and O. Riordan, Sparse graphs: metrics and random models. Preprint, 2007. http://arxiv.org/0708.1919v1

[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. Preprint, 2007. http://arxiv.org/math.CO/0702004

[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] F. R. K. Chung, R. L. Graham & R. M. Wilson, Quasi-random graphs. *Combinatorica* **9** (1989), no. 4, 345–362.

[10] P. Diaconis, S. Holmes and S. Janson, Threshold graph limits and random threshold graphs. In preparation.

[11] P. Diaconis & S. Janson, Graph limits and exchangeable random graphs. *Rendiconti di Matematica*, to appear. http://arxiv.org/0712.2749v1

[12] E. N. Gilbert, Random graphs. *Ann. Math. Statist.* **30** (1959), 1141–1144.

[13] S. Janson, Standard representation of multivariate functions on a general probability space. Preprint, 2007. http://arxiv.org/0801.0196v1

[14] O. Kallenberg, *Probabilistic Symmetries and Invariance Principles.* Springer, New York, 2005.

[15] J. F. C. Kingman, The coalescent. *Stochastic Process. Appl.* **13** (1982), no. 3, 235–248.

[16] L. Lovász and V. T. Sós, Generalized quasirandom graphs. *J. Comb. Theory B.* **98** (2008), no. 1, 146–163.

[17] L. Lovász and B. Szegedy, Limits of dense graph sequences. *J. Comb. Theory B* **96**, 933–957, 2006.

[18] T. Tao, A correspondence principle between (hyper)graph theory and probability theory, and the (hyper)graph removal lemma. *J. d'Analyse Math.*, to appear. http://arxiv.org/math/0602037v2

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/