ASYMPTOTIC EQUIVALENCE AND CONTIGUITY OF SOME RANDOM GRAPHS

SVANTE JANSON

ABSTRACT. We show that asymptotic equivalence, in a strong form, holds between two random graph models with slightly differing edge probabilities under substantially weaker conditions than what might naively be expected.

One application is a simple proof of a recent result by van den Esker, van der Hofstad and Hooghiemstra on the equivalence between graph distances for some random graph models.

1. INTRODUCTION

There are many different models of random graphs. Sometimes, the differences are minor, and it can be guessed that the asymptotic behaviour of two models are the same (for all or at least for some interesting properties). This note concerns some cases where it is possible to actually prove such results in a strong form. We begin by defining the two types of asymptotic equality that we will study. All unspecified limits are as $n \to \infty$.

Definition 1.1. Let $(\mathcal{X}_n, \mathcal{A}_n)$, $n \geq 1$, be a sequence of arbitrary measurable spaces and let P_n and Q_n be two probability measures on $(\mathcal{X}_n, \mathcal{A}_n)$.

(i) The sequence $(P_n)_n$ is *asymptotically equivalent* to $(Q_n)_n$, denoted by $(P_n)_n \cong (Q_n)_n$, if for every sequence of measurable sets A_n (i.e., $A_n \in \mathcal{A}_n$), we have $P_n(A_n) - Q_n(A_n) \to 0$.

(ii) The sequence $(P_n)_n$ is *contiguous* with respect to $(Q_n)_n$, denoted by $(P_n)_n \lhd (Q_n)_n$, if for every sequence of measurable sets A_n such that $Q_n(A_n) \to 0$, we also have $P_n(A_n) \to 0$.

We use the same terminology and notations for sequences of random variables X_n and Y_n with values in the same space $\mathcal{X}_n$, meaning that these properties hold for their distributions $\mathcal{L}(X_n)$ and $\mathcal{L}(Y_n)$. For example, $(X_n)_n \cong (Y_n)_n$ means that $\mathbb{P}(X_n \in A_n) - \mathbb{P}(Y_n \in A_n) \to 0$ for every sequence $(A_n)_n$. We will also use the simpler notations $X_n \cong Y_n$ and $X_n \lhd Y_n$, etc.

Note that asymptotic equivalence is a symmetric relation while contiguity is not; we say that $(P_n)_n$ and $(Q_n)_n$ are (mutually) contiguous, $(P_n)_n \lhd\rhd (Q_n)_n$, if both $(P_n)_n \lhd (Q_n)_n$ and $(Q_n)_n \lhd (P_n)_n$, i.e., if $P_n(A_n) \to 0 \iff$

Date: February 12, 2008; revised September 23, 2009.
2000 *Mathematics Subject Classification.* 60C05; 05C80; 60G30.

$Q_n(A_n) \to 0$ for any sequence of measurable sets $A_n \subseteq \mathcal{X}_n$. (And similarly for sequences of random variables X_n and Y_n.)

Asymptotic equivalence implies contiguity, but not conversely (see e.g. Example 1.2 and Remark 1.6), so contiguity is a weaker property.

We illustrate these notions by two simple examples.

Example 1.2. In the special case of two constant sequences, $P_n = P$ and $Q_n = Q$ where P and Q are two probability measures defined on the same space $(\mathcal{X}_n, \mathcal{A}_n) = (\mathcal{X}; \mathcal{A})$, $(P_n)_n \cong (Q_n)_n$ if and only $P = Q$, and $(P_n)_n \lhd (Q_n)_n$ if and only if $P \ll Q$, i.e., P is absolutely continuous with respect to Q. Hence asymptotic equivalence and contiguity can be thought of as asymptotic versions of equality and absolute continuity, respectively.

Example 1.3. Let X_n be random elements in some spaces $\mathcal{X}_n$ and let $\mathcal{E}_n$ be events that depend on X_n only, i.e., $\mathcal{E}_n = \{X_n \in E_n\}$ for some (measurable) sets $E_n \subseteq \mathcal{X}_n$, and suppose that $\liminf \mathbb{P}(\mathcal{E}_n) > 0$. Let $Y_n := (X_n \mid \mathcal{E}_n)$ be X_n conditioned on $\mathcal{E}_n$ (possibly ignoring some small n with $\mathbb{P}(\mathcal{E}_n) = 0$). Then, for any A_n, and some $C < \infty$,

$$\mathbb{P}(Y_n \in A_n) = \frac{\mathbb{P}(X_n \in A_n \cap E_n)}{\mathbb{P}(\mathcal{E}_n)} \le C\,\mathbb{P}(X_n \in A_n \cap E_n) \le C\,\mathbb{P}(X_n \in A_n)$$

and thus $(Y_n)_n \lhd (X_n)_n$

An important random graph example of this is when Y_n is a random graph with a given degree sequence $d_1, \dots, d_n$, uniformly chosen among all such graphs, and X_n is the random multigraph constructed by the configuration model (see, e.g., Bollobás [3]); then $Y_n \overset{\mathrm{d}}{=} (X_n \mid X_n \text{ is a simple graph})$ so $(Y_n)_n \lhd (X_n)_n$ provided $\liminf_{n\to\infty} \mathbb{P}(X_n \text{ is simple}) > 0$. This is the case when $\sum_{i=1}^n d_i \to \infty$ and $\sum_{i=1}^n d_i^2 = O\big(\sum_{i=1}^n d_i\big)$, see Janson [13] (with several earlier partial results by various authors), which makes it possible to transfer many results from X_n to Y_n. Indeed, this is a standard method to study random graphs with a given degree sequence, and in particular random regular graphs, see e.g., [3], [14], [18].

Remark 1.4. Suppose that $X_n \cong Y_n$. If $\mathbb{P}(X_n \in A_n) \to \alpha$ for some sequence of (measurable) sets A_n and some $\alpha \in [0,1]$, then $\mathbb{P}(Y_n \in A_n) \to \alpha$ too. Hence, any result for X_n that can be stated in terms of convergence of some probabilities holds for Y_n too; for example, this includes any result of the type $\varphi_n(X_n) \xrightarrow{\mathrm{p}} a$ and $\varphi_n(X_n) \xrightarrow{\mathrm{d}} W$ for some functionals $\varphi_n : \mathcal{X}_n \to \mathbb{R}$ (and a number a or a random variable W). However, results that are sensitive to events with small probabilities, such as moment convergence or large deviation estimates, do not transfer automatically. For example, if $X_n \cong Y_n$ and we know that $\mathbb{E}\,\varphi_n(X_n) \to a$, we may guess that $\mathbb{E}\,\varphi_n(Y_n) \to a$ too, but we cannot conclude it without further information (for example uniform integrability of $\varphi_n(X_n)$ and $\varphi_n(Y_n)$).

If instead only $X_n \rhd Y_n$, then results of the type $\varphi_n(X_n) \xrightarrow{\mathrm{p}} a$ still transfer to Y_n, but not result on convergence in distribution. (If $\varphi_n(X_n) \xrightarrow{\mathrm{d}}$

W, then the sequence $\varphi_n(Y_n)$ is tight, but does not have to converge to W, or at all. Typically, $\varphi(Y_n) \xrightarrow{\mathrm{d}} W'$ for some $W' \neq W$, see e.g. Example 1.2 and several examples of cycle counts in [14, Chapter 9].)

Suppose now that G_n and G'_n are random graphs on the vertex set $[n] := \{1, \dots, n\}$. By the standard Theorem 4.2 below, $G_n \cong G'_n$ if and only if it is possible to couple G_n and G'_n, i.e., to define them simultaneously on some probability space, such that $\mathbb{P}(G_n \neq G'_n) \to 0$. (We assume that we are interested only in the distributions of G_n and G'_n, so we may replace them by any random graphs with the same distributions.)

In particular, we will study random graphs of the following type. If p_{ij}, $1 \le i < j \le n$, are given probabilities in [0,1], let $G(n, \{p_{ij}\})$ be the random graph on $[n]$ where the edge ij appears with probability p_{ij} and the indicators $I_{ij} := \mathbf{1}[\text{edge } ij \text{ appears}]$, $1 \le i < j \le n$, are independent. (We will later also consider an extension to random p_{ij}, see Section 2.)

Consider two sequences of such graphs, defined by probabilities $\{p_{ij}\}_{1\le i<j\le n}$ and $\{p'_{ij}\}_{1\le i<j\le n}$; p_{ij} and p_{ij} may depend on n too, but to simplify the notation we do not show this explicitly. It is obvious that we may couple the edge indicators I_{ij} and I'_{ij} of ij in $G(n, \{p_{ij}\})$ and $G(n, \{p'_{ij}\})$ such that $\mathbb{P}(I_{ij} \neq I'_{ij}) = |p_{ij} - p'_{ij}|$, and by taking independent pairs (I_{ij}, I'_{ij}) we obtain a coupling of the random graphs $G(n, \{p_{ij}\})$ and $G(n, \{p'_{ij}\})$ with

$$\mathbb{P}\big(G(n, \{p_{ij}\}) \neq G(n, \{p'_{ij}\})\big) \le \sum_{i<j} |p_{ij} - p'_{ij}|. \tag{1.1}$$

Consequently, $G(n, \{p_{ij}\}) \cong G(n, \{p'_{ij}\})$ if $\sum_{i<j} |p_{ij} - p'_{ij}| \to 0$; a simple fact that has been used by many authors. It may be believed that this is essentially best possible, but, somewhat surprisingly, this is not so. In fact, by Corollary 2.12 below, provided $p_{ij} \le 0.9$, say, $G(n, \{p_{ij}\}) \cong G(n, \{p'_{ij}\})$ if $\sum_{i<j} (p_{ij} - p'_{ij})^2/p_{ij} \to 0$. (Moreover, Theorem 2.2(i) shows that this is best possible if, for example, $p'_{ij} \le 2p_{ij}$.)

For a particular case, suppose that $p'_{ij} = p_{ij} + O(p_{ij}^2)$, Then Corollary 2.13 shows that $G(n, \{p_{ij}\}) \cong G(n, \{p'_{ij}\})$ if $\sum_{i<j} p_{ij}^3 \to 0$, while (1.1) implies this only under the stronger condition $\sum_{i<j} p_{ij}^2 \to 0$. A typical case where this is an important improvement is when all $p_{ij} = \Theta(1/n)$ and $|p'_{ij} - p_{ij}| = \Theta(1/n^2)$. See further the examples in Section 3.

To see that such an improvement of (1.1) might be possible at all, consider as an example the case when all p_{ij} are the same, so we consider the random graph $G(n, p)$:

Example 1.5. Let $p = p(n)$ and $p' = p'(n)$ be given in $[0, 1]$ and consider $G(n, p)$ and $G(n, p')$. Let $N := \binom{n}{2}$ be the number of possible edges and let $M \sim \mathrm{Bi}(N, p)$ and $M' \sim \mathrm{Bi}(N, p')$ be the number of edges in $G(n, p)$ and $G(n, p')$. The conditional distribution of $G(n, p)$ given $M = m$ is uniform over all graphs on $[n]$ with m edges, and the conditional distribution of $G(n, p')$ given $M' = m$ is the same. It follows that any coupling of M

and M' may be extended to a coupling of $G(n,p)$ and $G(n,p')$ such that $G(n,p) = G(n,p')$ when $M = M'$; as a consequence, using (4.6) below, $d_{\mathrm{TV}}\bigl(G(n,p),G(n,p')\bigr) = d_{\mathrm{TV}}(M,M')$, and in particular, as $n\to\infty$, using also Theorem 4.2,

$$G(n,p) \cong G(n,p') \iff M \cong M'.$$

Since M and M' have binomial distributions with the same n, $\mathbb{P}(M = k)/\mathbb{P}(M' = k)$ is monotone in k, and it follows that the maximum of $|\mathbb{P}(M \in A) - \mathbb{P}(M' \in A)|$ over subsets A of $\mathbb{Z}$ is attained for a set of the form $[0,1,\dots,j]$.

Now suppose that $p\to 0$ and $N(p'-p)/\sqrt{Np} \to \alpha \in [-\infty,\infty]$. Suppose first that α is finite. By the central limit theorem, $(M-Np)/\sqrt{Np} \xrightarrow{\mathrm{d}} \mathrm{N}(0,1)$ and $(M'-Np)/\sqrt{Np} \xrightarrow{\mathrm{d}} \mathrm{N}(\alpha,1)$, and it follows easily that

$$\begin{aligned}
d_{\mathrm{TV}}\bigl(G(n,p),G(n,p')\bigr) &= d_{\mathrm{TV}}(M,M') = \sup_j |\,\mathbb{P}(M\le j) - \mathbb{P}(M'\le j)| \\
&= \sup_x |\,\mathbb{P}\bigl(\mathrm{N}(0,1)\le x\bigr) - \mathbb{P}\bigl(\mathrm{N}(\alpha,1)\le x\bigr)| + o(1) \\
&\to \Phi(\alpha/2) - \Phi(-\alpha/2).
\end{aligned}$$

It follows, using Theorem 4.2 again, that $G(n,p) \cong G(n,p')$ if and only if $\alpha = 0$, i.e. $N(p'-p)/\sqrt{Np} \to 0$, which is equivalent to $\sum_{i<j}(p'-p)^2/p = N(p'-p)^2/p \to 0$. For example, if $p = 1/n$ and $p' = 1 - e^{-1/n} = p - \frac12 n^{-2} + O(n^{-3})$, then $N(p'-p)^2/p = O(1/n)$, so $G(n,p) \cong G(n,p')$, but $N|p-p'| \to 1/4$, so (1.1) is not enough to show this.

We see that in this example, the trick to improve the simple and 'obvious' edgewise coupling used in (1.1), is to first ignore the positions of the edges and couple their numbers only; this then is extended to a coupling of the random graphs by randomly reinserting the positions. Corollary 2.12(i) shows that couplings improving the simple edgewise coupling exist also when the edge probabilities are unequal, but in that case we do not know any explicit construction of such couplings.

We give the main results in Section 2 and a number of examples in Section 3; this includes an application to a recent result by van den Esker, van der Hofstad and Hooghiemstra (Example 3.6). Proofs are given in Section 5, after some preliminaries in Section 4.

We use the standard notations o_p and O_p, see e.g. [14, Section 1.2], and we write whp (with high probability) for events with probability tending to 1 as $n\to\infty$.

Remark 1.6. There are also interesting examples of contiguity among random graphs of other types than $G(n,\{p_{ij}\})$. In particular, several different constructions of random regular graphs (or multigraphs) are known to yield distributions that are (mutually) contiguous but not asymptotically equivalent, see e.g. [12], [14, Section 9.5], [10]. These examples are not covered by the present paper.

2. Results

Recall that an indicator variable is a random variable with values in $\{0, 1\}$, and that the Bernoulli distribution $\mathrm{Be}(p)$ is the distribution of an indicator variable I with $\mathbb{P}(I = 1) = p$ and $\mathbb{P}(I = 0) = 1 - p$, and thus $\mathbb{E}\, I = p$.

We defined above the random graph $G(n, \mathbf{p})$, where $\mathbf{p} = \{p_{ij}\}_{1 \le i<j \le n}$ is a vector of probabilities. We extend the definition of $G(n, \mathbf{p})$ to the case when $\mathbf{p}$ is a random vector (with entries in [0,1]) by conditioning on $\mathbf{p}$, i.e., given $\mathbf{p} = \{p_{ij}\}$, the edge indicators I_{ij} are independent random variables with $I_{ij} \sim \mathrm{Be}(p_{ij})$. Random graphs of this type have been studied in many papers, see for example Bollobás, Janson and Riordan [4] and the further references given there.

We now state our main results on asymptotic equivalence and contiguity of such random graphs. Actually, the results have nothing to do with the graph structure and the way the indicator variables are indexed by pairs ij. It therefore seems more natural to consider the more general situation of a (finite or infinite) sequence $(I_i)_1^N$ of indicator variables. The results for random graphs then follow by relabelling the indicators.

We define a function $\rho : [0, 1]^2 \to [0, \infty)$ in Definition 2.1, where we also give some equivalent (within constant factors) alternative formulas that often are more convenient. Since the results below are not affected by changing ρ within constant factors, we could use any of these alternative formulas (and several other similar ones) as our definition. (The motivation for the definition comes in Lemma 4.3.)

We write $x \asymp y$ (where $x, y \ge 0$) to denote that $cy \le x \le Cy$ for some positive constants c, C, i.e., that $x = \Theta(y)$ (or, equivalently, $x = O(y)$ and $y = O(x)$). Further, we use $x \vee y$ and $x \wedge y$ for the maximum and minimum, respectively, of x and y. We interpret $0/0$ as 0.

Definition 2.1. Let

$$\rho(p, q) := \left(\sqrt{p} - \sqrt{q}\right)^2 + \left(\sqrt{1 - p} - \sqrt{1 - q}\right)^2 \tag{2.1}$$

$$\asymp \frac{(p - q)^2}{p + q} + \frac{(p - q)^2}{1 - p + 1 - q} \tag{2.2}$$

$$\asymp \frac{(p - q)^2}{(p \vee q) \wedge ((1 - p) \vee (1 - q))} \tag{2.3}$$

$$\asymp \frac{(p - q)^2}{p \wedge (1 - p)} \wedge |p - q|. \tag{2.4}$$

In particular, if $p \le 0.9$, then

$$\rho(p, q) \asymp \frac{(p - q)^2}{p} \wedge |p - q|. \tag{2.5}$$

Of course, the constant 0.9 here and below is arbitrary and could be replaced by any number < 1.

Proof. The first equivalence follows from

$$\left(\sqrt{p}-\sqrt{q}\right)^2 = \frac{(p-q)^2}{(\sqrt{p}+\sqrt{q})^2} \asymp \frac{(p-q)^2}{p+q},$$

together with the similar result with $1-p$ and $1-q$. The second follows from $x+y \asymp x \vee y$ for $x,y \geq 0$ (used thrice). The third is equivalent to

$$(p \vee q) \wedge \bigl((1-p) \vee (1-q)\bigr) \asymp \bigl(p \wedge (1-p)\bigr) \vee |p-q|, \tag{2.6}$$

which is easily verified, for example by assuming (by the symmetry $p \mapsto 1-p$, $q \mapsto 1-q$) that $p \leq 1/2$, in which case (2.6) easily reduces to $p \vee q \asymp p \vee |p-q|$. $\square$

We state our results first for the simpler case of sequences of independent indicator variables with given (non-random) probabilities. The following theorem gives necessary and sufficient conditions for asymptotical equivalence and contiguity. (The asymptotical equivalence criterion follows by a simple and standard type of calculation with Hellinger distances, see the proof in Section 5 and, e.g., [16, p. 158], although we have not seen it stated in this form before. The contiguity criterion is a special case of a result by Oosterhoff and van Zwet [16] for general sequences of independent variables.) The proofs of the theorems are given in Section 5.

Theorem 2.2. *Let $1 \leq N(n) \leq \infty$ and let $X_n = (I_{ni})_{i=1}^{N(n)}$ and $X'_n = (I'_{ni})_{i=1}^{N(n)}$ be finite or infinite random vectors consisting of independent indicator variables $I_{ni} \sim \mathrm{Be}(p_{ni})$ and $I'_{ni} \sim \mathrm{Be}(p'_{ni})$.*

(i) *$X_n \cong X'_n$ if and only if*

$$\sum_{i=1}^{N(n)} \rho(p_{ni}, p'_{ni}) \to 0. \tag{2.7}$$

(ii) *$X_n \lhd X'_n$ if and only if*

$$\sum_{i=1}^{N(n)} \rho(p_{ni}, p'_{ni}) = O(1) \tag{2.8}$$

and, with $q_{ni} := 1-p_{ni}$ and $q'_{ni} := 1-p'_{ni}$,

$$\lim_{C\to\infty} \limsup_{n\to\infty} \left(\sum_{i:p_{ni}>Cp'_{ni}} p_{ni} + \sum_{i:q_{ni}>Cq'_{ni}} q_{ni} \right) = 0. \tag{2.9}$$

Remark 2.3. By symmetry, $X_n \rhd X'_n$ is equivalent to (2.8) and

$$\lim_{C\to\infty} \limsup_{n\to\infty} \left(\sum_{i:p'_{ni}>Cp_{ni}} p'_{ni} + \sum_{i:q'_{ni}>Cq_{ni}} q'_{ni} \right) = 0, \tag{2.10}$$

and thus $X_n \lhd\rhd X'_n$ is equivalent to (2.8), (2.9) and (2.10).

Remark 2.4. Often $p_{ni} \le 0.9$ for all n and i, and then the second sum in (2.9) vanishes for $C > 10$ and can thus be omitted.

Remark 2.5. The condition (2.9) is only needed to take care of cases when p_{ni} and p'_{ni} (or q_{ni} and q'_{ni}, in case p_{ni} and p'_{ni} are close to 1) are not of the same order. If no such p_{ni} and p'_{ni} appear, which is the typical case, then (2.8) is thus enough.

We may rewrite (2.9) in several ways. For example, it is equivalent to (following the formulation in [16] in a more general case): for every sequence $C_n \to \infty$,

$$\sum_{i:p_{ni}>C_n p'_{ni}} p_{ni} + \sum_{i:q_{ni}>C_n q'_{ni}} q_{ni} \to 0. \tag{2.11}$$

It is also equivalent to: For every $\varepsilon > 0$, there exist C and n_0 such that if $n \ge n_0$, then

$$\sum_{i:p_{ni}>C p'_{ni}} p_{ni} < \varepsilon \qquad \text{and} \qquad \sum_{i:q_{ni}>C q'_{ni}} q_{ni} < \varepsilon. \tag{2.12}$$

Remark 2.6. As pointed out by Oosterhoff and van Zwet [16], (2.8) does not imply (2.9) in general. A simple counter example is provided by $N(n) = n$, $p_{ni} = n^{-1}$, $p'_{ni} = n^{-2}$. (On the other hand, it is easy to see, and also follows by the theorem, that (2.7) implies (2.9) and (2.10).)

Remark 2.7. In the very special case when $N(n)$, p_{ni} and p'_{ni} do not depend on n (and we omit the subscript n), it is easily shown that if $0 < p'_i < 1$ for all i, then (2.8) implies (2.9), and thus $(X) \lhd (X')$, which by Example 1.2 says that the distribution of X is absolutely continuous with respect to the distribution of X'. If we further assume also $0 < p_i < 1$, by symmetry the distributions are thus mutually absolutely continuous. This is part of the dichotomy by Kakutani for product measures, see e.g. [11, Corollary IV.2.38], which in our case says that either $\sum_i \rho(p_i, p'_i) < \infty$ and the distributions are mutually absolutely continuous, or $\sum_i \rho(p_i, p'_i) = \infty$ and the distributions are mutually singular.

Returning to the general case in Theorem 2.2, it is easy to show that, analoguously, if $\sum_i \rho(p_{ni}, p'_{ni}) \to \infty$, then the distributions of X_n and X'_n are asymptotically mutually singular in the sense that there exist sets A_n with $\mathbb{P}(X_n \in A_n) \to 1$ and $\mathbb{P}(X'_n \in A_n) \to 0$, cf. [11, Theorem V.2.32].

Remark 2.8. We have stated Theorem 2.2 in terms of sequences of pairs of random vectors. It is possible (at least partly) to rephrase it in terms of estimates for a single pair (X, X'), see Lemmas 5.1 and 5.2 below. Similar reformulations may be made for Theorem 2.9, but we leave these to the reader.

We extend Theorem 2.2 to the case of random probabilities p_{ni}. In this case we cannot expect conditions that are both necessary and sufficient, so we give only sufficient conditions, which are more important in applications.

(An important obstacle to finding necessary conditions is that different distributions of the probabilities may give the same distribution of the indicators. For example, using the notation of Theorem 2.9, if p_{ni} are i.i.d. with $p_{ni} \sim U(0,1)$ and $p'_{ni} = 1/2$, then $X_n \overset{\mathrm{d}}{=} X'_n$.)

Theorem 2.9. *Let $1 \le N(n) \le \infty$ and suppose that $\mathbf{p}_n = \{p_{ni}\}$ and $\mathbf{p}'_n = \{p'_{ni}\}$ are random vectors in $[0,1]^{N(n)}$. Let $X_n = (I_{ni})_{i=1}^{N(n)}$ and $X'_n = (I'_{ni})_{i=1}^{N(n)}$ be random vectors of indicator variables such that the conditioned random vectors $(X_n \mid \mathbf{p}_n)$ and $(X'_n \mid \mathbf{p}'_n)$ are sequences of independent indicator variables with $(I_{ni} \mid \mathbf{p}_n) \sim \mathrm{Be}(p_{ni})$ and $(I'_{ni} \mid \mathbf{p}'_n) \sim \mathrm{Be}(p'_{ni})$.*

(i) *If*

$$\sum_{i=1}^{N(n)} \rho(p_{ni}, p'_{ni}) = o_p(1), \tag{2.13}$$

then $X_n \cong X'_n$.

(ii) *If*

$$\sum_{i=1}^{N(n)} \rho(p_{ni}, p'_{ni}) = O_p(1) \tag{2.14}$$

and, with $q_{ni} := 1 - p_{ni}$ and $q'_{ni} := 1 - p'_{ni}$, for every $\varepsilon > 0$,

$$\lim_{C\to\infty} \limsup_{n\to\infty} \mathbb{P}\left(\sum_{i: p_{ni} > C p'_{ni}} p_{ni} + \sum_{i: q_{ni} > C q'_{ni}} q_{ni} > \varepsilon \right) = 0, \tag{2.15}$$

then $X_n \lhd X'_n$.

Remark 2.10. Recall that if S_n denotes the random sum on the left-hand side of (2.13), then (2.13) can also be written $S_n \xrightarrow{\mathrm{p}} 0$. Similarly, the $O_p(1)$ notation in (2.14) means that for every $\varepsilon > 0$, there exists C such that $\mathbb{P}(S_n > C) < \varepsilon$ for all n; this is also known as stochastic boundedness or tightness of the sequence $\{S_n\}$, and is equivalent to $\mathbb{P}(S_n > C_n) \to 0$ for every sequence $C_n \to \infty$.

Remark 2.11. In analogy to (2.11), the condition (2.15) is equivalent to: For every sequence $C_n \to \infty$,

$$\sum_{i: p_{ni} > C_n p'_{ni}} p_{ni} + \sum_{i: q_{ni} > C_n q'_{ni}} q_{ni} \xrightarrow{\mathrm{p}} 0. \tag{2.16}$$

As said above, Theorems 2.2 and 2.9 apply immediately to random graphs $G(n, \mathbf{p})$. We state a version of Theorem 2.9 for this case, where we have added some simplifying assumptions. Recall that p_{ij} and p'_{ij} may (and typically do) depend on n, although we do not show that in our notation. Of course, as a special case, the results immediately apply to non-random p_{ij} and p'_{ij} too.

Corollary 2.12. *Let, for each n, $\mathbf{p} = \{p_{ij}\}$ and $\mathbf{p}' = \{p'_{ij}\}$ be random vectors of probabilities and suppose that whp* $\max_{i,j} p_{ij} \le 0.9$.

(i) *If*
$$\sum_{i<j} \frac{(p_{ij} - p'_{ij})^2}{p_{ij}} = o_p(1), \tag{2.17}$$
then $G(n,\mathbf{p}) \cong G(n,\mathbf{p}')$.

(ii) *If*
$$\sum_{i<j} \frac{(p_{ij} - p'_{ij})^2}{p_{ij}} = O_p(1), \tag{2.18}$$
then $G(n,\mathbf{p}) \rhd G(n,\mathbf{p}')$.

(iii) *If (2.18) holds, and further, for some constant $c > 0$, whp $cp_{ij} \le p'_{ij} \le 0.9$ for all i,j, then $G(n,\mathbf{p}) \lhd\rhd G(n,\mathbf{p}')$.*

We specialize further to an important case.

Corollary 2.13. *Let, for each n, $\mathbf{p} = \{p_{ij}\}$ and $\mathbf{p}' = \{p'_{ij}\}$ be random vectors of probabilities and suppose that $p'_{ij} = p_{ij} + O(p_{ij}^2)$.*

(i) *If $\sum_{i<j} p_{ij}^3 = o_p(1)$, then $G(n,\mathbf{p}) \cong G(n,\mathbf{p}')$.*

(ii) *If $\sum_{i<j} p_{ij}^3 = O_p(1)$, and further, for some constant $c > 0$, whp $\max_{i,j} p_{ij} \le 0.9$, $\max_{i,j} p'_{ij} \le 0.9$ and $p'_{ij} \ge cp_{ij}$ for all i,j, then $G(n,\mathbf{p}) \lhd\rhd G(n,\mathbf{p}')$.*

3. Examples

Example 3.1. Bollobás, Janson and Riordan [4] study a general class of sparse random graphs $G(n,\kappa)$ which include many cases studied earlier by various authors. These random graphs are defined as $G(n,\{p_{ij}\})$ with
$$p_{ij} := p_{ij}^{(1)} := \min\Bigl(\frac{\kappa(x_i,x_j)}{n}, 1\Bigr) = \hat{p}_{ij} \wedge 1, \tag{3.1}$$
with
$$\hat{p}_{ij} := \frac{\kappa(x_i,x_j)}{n}, \tag{3.2}$$
where $\kappa : \mathcal{S} \times \mathcal{S} \to [0,\infty)$ is a symmetric measurable function defined on some measurable space $\mathcal{S}$ and $x_1,\dots,x_n$ is a random sequence of elements of $\mathcal{S}$, not necessarily i.i.d. but such that the empirical distribution of $x_1,\dots,x_n$ converges to a probability measure μ on $\mathcal{S}$; see [4] for details. (Some further technical conditions are assumed in [4]; they are not relevant here.) Typically, whp $\kappa(x_i,x_j) \le n$ for all i,j, and then p_{ij} equals the simpler $\hat{p}_{ij}$. Two natural variations, also treated in [4] and used in various cases by various authors, are obtained by replacing (3.1) by
$$p_{ij}^{(2)} := 1 - \exp\Bigl(-\frac{\kappa(x_i,x_j)}{n}\Bigr) = 1 - \exp\bigl(-\hat{p}_{ij}\bigr) \tag{3.3}$$

or

$$p_{ij}^{(3)} := \frac{\kappa(x_i, x_j)}{n + \kappa(x,x_j)} = \frac{\hat{p}_{ij}}{1 + \hat{p}_{ij}}. \tag{3.4}$$

(Thus $p_{ij}^{(3)}/(1 - p_{ij}^{(3)}) = \hat{p}_{ij}$; at least in the case studied in Example 3.5 below, this is in some sense simpler and more natural, see Britton, Deijfen and Martin-Löf [6].) In all cases $p_{ij}^{(\ell)} = \hat{p}_{ij} + O(\hat{p}_{ij}^2)$, which is the essential estimate for our purposes; the results below extend to the general case

$$p_{ij} := h(\hat{p}_{ij}) \quad \text{for a function } h \text{ with} \quad h(x) = x + O(x^2). \tag{3.5}$$

It was shown in [4] that the same asymptotic results hold for these three versions for the properties studied there. We can now show that, under an extra condition, the three versions are asymptotically equivalent, and thus have the same asymptotic behaviour for *any* property. Indeed, Corollary 2.13 applies immediately and shows that if

$$\sum_{1 \le i < j \le n} \kappa(x_i, x_j)^3 = o_p(n^3), \tag{3.6}$$

then all three $G(n, p_{ij}^{(\ell)})$, $\ell = 1, 2, 3$, are asymptotically equivalent; similarly, if the weaker

$$\sum_{1 \le i < j \le n} \kappa(x_i, x_j)^3 = O_p(n^3) \tag{3.7}$$

holds together with $\max_{i,j} \hat{p}_{ij} \le 0.9$ whp, then all three $G(n, p_{ij}^{(\ell)})$ are mutually contiguous.

In fact, (3.7) alone suffices for $G(n, p_{ij}^{(2)}) \lhd\rhd G(n, p_{ij}^{(3)})$ because (3.7) implies $\max_{i,j} \hat{p}_{ij} = O_p(1)$ so by conditioning we may assume that $\max_{i,j} \hat{p}_{ij} \le C_1$ for some constant C_1, and then, for $\ell = 2, 3$, $p_{ij}^{(\ell)} \le C_2 < 1$ and $c\hat{p}_{ij} \le p_{ij}^{(\ell)} \le \hat{p}_{ij}$. Furthermore, by the same conditioning and Corollary 2.12(ii), (3.7) implies $G(n, p_{ij}^{(2)}) \rhd G(n, p_{ij}^{(1)})$ and $G(n, p_{ij}^{(3)}) \rhd G(n, p_{ij}^{(1)})$. However, if, for example $\kappa(x_1, x_2) \ge n$ whp, then $I_{12} = 1$ whp in $G(n, p_{ij}^{(1)})$ but not in $G(n, p_{ij}^{(2)})$ or $G(n, p_{ij}^{(3)})$, and we do not have contiguity in the opposite direction.

We study some special cases in the following examples.

Example 3.2. One common case of the construction in Example 3.1 uses $x_1, \dots, x_n$ that are i.i.d. on $\mathcal{S}$ with distribution μ. In this case, we show that the condition

$$\mu \times \mu\{(x, y) : \kappa(x, y) > t)\} = o\big(t^{-2}\big) \qquad \text{as } t \to \infty \tag{3.8}$$

implies (3.6) and thus asymptotic equivalence of the three versions. In particular, this holds if $\int_{\mathcal{S}\times\mathcal{S}} \kappa(x, y)^2 \, d\mu(x) \, d\mu(y) < \infty$.

In fact, if $G(t) := \mu \times \mu\{(x,y) : \kappa(x,y) > t)\} = o(t^{-2})$, then, using a standard Fubini argument,

$$\begin{aligned}
\mathbb{E}\bigl(\kappa(x_1,x_2)\wedge n\bigr)^3 &= \int_{\mathcal{S}\times\mathcal{S}} \bigl(\kappa(x,y)\wedge n\bigr)^3 \, d\mu(x)\, d\mu(y) \\
&= \int_{\mathcal{S}\times\mathcal{S}} \int_0^{\kappa(x,y)\wedge n} 3t^2 \, dt \, d\mu(x)\, d\mu(y) \\
&= \int_0^n 3t^2 G(t)\, dt = n \int_0^1 3(ns)^2 G(ns)\, ds = o(n) \qquad (3.9)
\end{aligned}$$

by (3.8) and dominated convergence. Hence, $\mathbb{E}\sum_{i<j}(\kappa(x_i,x_j)\wedge n)^3 = o(n^3)$, so $\sum_{i<j}(\kappa(x_i,x_j)\wedge n)^3 = o_p(n^3)$. Moreover,

$$\begin{aligned}
\mathbb{P}\Bigl(\sum_{1\le i<j\le n} (\kappa(x_i,x_j)\wedge n)^3 \neq \sum_{1\le i<j\le n} \kappa(x_i,x_j)^3\Bigr) \\
\le \sum_{1\le i<j\le n} \mathbb{P}(\kappa(x_i,x_j) > n) \le n^2 G(n) = o(1),
\end{aligned}$$

and (3.6) follows.

Similarly, we can easily shown that (3.7), and thus at least partial contiguity, follows from

$$\mu \times \mu\{(x,y) : \kappa(x,y) > t)\} = O\bigl(t^{-2}\bigr) \qquad \text{as } t \to \infty. \qquad (3.10)$$

In this case, given $\varepsilon > 0$, there exists C_1 such that

$$\mathbb{P}\Bigl(\sum_{1\le i<j\le n} (\kappa(x_i,x_j)\wedge C_1 n)^3 \neq \sum_{1\le i<j\le n} \kappa(x_i,x_j)^3\Bigr) \le \varepsilon;$$

further, a calculation as in (3.9) yields $\mathbb{E}\sum_{1\le i<j\le n}(\kappa(x_i,x_j)\wedge C_1 n)^3 = O(n^3)$ and thus $\mathbb{P}\bigl(\sum_{1\le i<j\le n}(\kappa(x_i,x_j)\wedge C_1 n)^3 > C_2 n^3\bigr) < \varepsilon$ for some C_2; we omit the details.

Example 3.3. Another case of the construction in Example 3.1 uses $\mathcal{S} = (0,1]$ with $\mu =$ Lebesgue measure and the deterministic $x_i = i/n$, $i = 1,\dots,n$. The homogeneous case $\kappa(x,y) = c/(x\vee y)$ yielding $\hat{p}_{ij} = c/(i\vee j)$, where $c > 0$ is a constant, is particularly interesting and related to the CHKNS model, see Bollobás, Janson and Riordan [4], Sections 16.1, Durrett [7, 8] and Riordan [17] and the references given there.

In this case, $\sum_{1\le i<j<\infty} p_{ij}^3 \le c^3 \sum_{j\ge 2} j\cdot j^{-3} < \infty$, and thus $\sum_{i<j} p_{ij}^3 = O(1)$; if we further for simplicity assume $c < 2$ and thus $\max_{ij}\hat{p}_{ij} < 1$, then Corollary 2.13(iii) implies that $G(n,p_{ij}^{(1)}) \lhd\rhd G(n,p_{ij}^{(2)}) \lhd\rhd G(n,p_{ij}^{(3)})$.

Note that in this case, $p_{12}^{(1)}$, $p_{12}^{(2)}$ and $p_{12}^{(3)}$ are constant and different, which shows that the three random graphs are *not* asymptotically equivalent (for a trivial reason).

We have $p_{ij}^{(3)} = c/(i\vee j + c)$; the same results hold for the further variation $p_{ij} = c/(i\vee j + d)$ for any $d > c - 2$.

In this example, the infinite random graphs $G(\infty, p_{ij}^{(\ell)})$, $\ell = 1, 2, 3$, are well-defined too, and it follows from Kakutani's criterion discussed in Remark 2.7 that (still provided $c < 2$) these three infinite random graphs have mutually absolutely continuous distributions, which is the infinite graph version of the contiguity result just given for finite n, cf. Example 1.2. (The infinite random graph $G(\infty, p_{ij}^{(1)})$ was studied before the finite version, see [15] and [7], [8] with further references.)

Example 3.4. A related case uses the same $\mathcal{S} = (0, 1]$, $\mu =$ Lebesgue measure and $x_i = i/n$, $i = 1, \dots, n$, as Example 3.3, now with the homogeneous $\kappa(x, y) = c/\sqrt{xy}$ yielding $\hat{p}_{ij} = c/\sqrt{ij}$; this case is a mean-field version of the preferential attachment model by Barabási and Albert [1], see Bollobás, Janson and Riordan [4], 16.2 and Riordan [17] and the references given there.

Also in this case, $\sum_{1 \le i < j < \infty} p_{ij}^3 < \infty$, and thus $\sum_{i<j} p_{ij}^3 = O(1)$ (in spite of the fact that (3.10) does not hold); if we further for simplicity assume $c < \sqrt{2}$, and thus $\max_{ij} \hat{p}_{ij} < 1$, we obtain the same results as in Example 3.3.

Example 3.5. A common case of Example 3.1 is when $\kappa(x, y) = \psi(x)\psi(y)$ for some function $\psi : \mathcal{S} \to [0, \infty)$, see [4, Section 16.4] for discussion and references to previous papers.

In this case, $\sum_{i<j} \kappa(x_i, x_j)^3 \le \bigl(\sum_i \psi(x_i)^3\bigr)^2$, so (3.6) and (3.7) may be replaced by

$$\sum_{i=1}^{n} \psi(x_i)^3 = o_p\bigl(n^{3/2}\bigr), \tag{3.11}$$

and

$$\sum_{i=1}^{n} \psi(x_i)^3 = O_p\bigl(n^{3/2}\bigr). \tag{3.12}$$

If we combine this choice of κ with the i.i.d. choice of x_i in Example 3.2, it is easily seen, arguing as in (3.9) but now with $\sum_i \bigl(\psi(x_i) \wedge n^{1/2}\bigr)$, that

$$\mu\{x : \psi(x) > t)\} = o\bigl(t^{-2}\bigr) \qquad \text{as } t \to \infty \tag{3.13}$$

implies (3.6) and thus asymptotic equivalence of the three versions; in particular this holds if $\int \psi(x)^2 \, d\mu(x) < \infty$. Similarly,

$$\mu\{x : \psi(x) > t)\} = O\bigl(t^{-2}\bigr) \qquad \text{as } t \to \infty \tag{3.14}$$

implies (3.7) and thus at least partial contiguity.

Example 3.6. van den Esker, van der Hofstad and Hooghiemstra [9] study a minor variation of the construction in Example 3.5; they let $\lambda_1, \dots, \lambda_n$ be positive *capacities* and define p_{ij} by (in our notation) (3.1), (3.3), (3.4) or more generally (3.5) with

$$\hat{p}_{ij} := \frac{\lambda_i \lambda_j}{\sum_1^n \lambda_i}. \tag{3.15}$$

(This too can be seen as an instance of the general construction in Example 3.1, see [4, Section 16.4].) They consider both the case when the capacities $\lambda_1, \dots, \lambda_n$ are deterministic, and when $\lambda_i = \Lambda_i$ for a sequence $\Lambda_1, \Lambda_2, \dots$ of positive i.i.d. random variables with some fixed distribution. (In both cases, some further conditions are imposed in [9].)

Consider for example the random case and assume that $\mathbb{P}(\Lambda_1 > t) = o(t^{-2})$ (which is the case in [9]). Then, just as (3.11) follows from (3.13), $\sum_1^n \Lambda_i^3 = o_p(n^{3/2})$. Since further $\sum_1^n \Lambda_i/n \xrightarrow{\text{p}} \mathbb{E}\,\Lambda > 0$ by the law of large numbers, it follows from (3.15) that $\sum_{i<j} p_{ij}^3 \xrightarrow{\text{p}} 0$. Hence, if we compare two random graphs $G(n, \mathbf{p})$ and $G(n, \mathbf{p}')$ defined by this method for two different functions h and h', we obtain $G(n, \mathbf{p}) \cong G(n, \mathbf{p}')$ by Corollary 2.13.

The same applies to the deterministic case studied in [9]. Then $\sum_i \lambda_i \sim \mu n$ for some $\mu > 0$, $\sum_i \lambda_i^2 = O(n)$ and $\max_i \lambda_i = o(n^{1/2})$, and thus $\sum_i \lambda_i^3 = o(n^{3/2})$ and (3.15) implies $\sum_{i<j} p_{ij}^3 = O\big((\sum_i \lambda_i^3)^2/n^3\big) = o(1)$.

van den Esker, van der Hofstad and Hooghiemstra [9] study the distance H_n between two random points, and (a minor) part of their proof consists in showing that the choice of h does not matter (enabling them to consider only the version (3.3) in the main part of the proof): the variables H_n and H_n' obtained by two different functions h and h' in (3.5) can be coupled such that $\mathbb{P}(H_n \neq H_n') = o(1)$, or in our notation $H_n \cong H_n'$, see Theorem 4.2. We thus obtain this as an immediate consequence of the stronger statement $G(n, \mathbf{p}) \cong G(n, \mathbf{p}')$, which by Theorem 4.2 means that the random graphs can be coupled with $\mathbb{P}\big(G(n, \mathbf{p}) \neq G(n, \mathbf{p}')\big) \to 0$.

Example 3.7. Our results are stated for graphs with a deterministic number of vertices, but can be extended to graphs with random vertex set too by conditioning on the vertex set. One interesting such case is obtained from Example 3.1 by letting $x_1, \dots, x_n$ be the points of a Poisson process on $\mathcal{S}$ with intensity $\lambda\mu$, where $\lambda > 0$ is our parameter and we consider asymptotics as $\lambda \to \infty$; thus n is random with the distribution $\mathrm{Po}(\lambda)$.

Conditioned on n, we have the situation in Example 3.2. It follows, for example, that if (3.8) holds, then the random graphs defined in this way using (3.1), (3.3) and (3.4) are asymptotically equivalent; we omit the details.

Example 3.8. Bollobás, Janson and Riordan [5] study a generalization of the model in Example 3.1 where small sets of edges are added at once, thus allowing a certain degree of clustering; more precisely, for every subgraph F of the complete graph K_n, we have a certain probability of adding (the edges of) F, and these events are independent for different F. While this introduces dependencies between the edge indicators, the results of the present paper are still applicable to the sequence of indicators I_F describing the added sets of edges, and asymptotic equivalence or contiguity for two versions of this sequence obviously implies asymptotic equivalence or contiguity for the resulting random graphs too.

We leave the explicit statement of results in this case to the reader.

Example 3.9. In this final example, let us return to the case of deterministic $\mathbf{p} = \{p_{ij}\}$ and let us change all p_{ij} proportionately to $p'_{ij} := (1+\delta_n)p_{ij}$ for some δ_n. Assume for simplicity that all $p_{ij} \le 0.9$ and that $|\delta_n| \le 1$.

By Corollary 2.12(i), if $\delta_n^2 \sum_{i<j} p_{ij} \to 0$, then $G(n,\mathbf{p}) \cong G(n,\mathbf{p}')$. Further, by Corollary 2.12(iii), if $\delta_n^2 \sum_{i<j} p_{ij} = O(1)$ and, for simplicity, $\delta_n \to 0$, then $G(n,\mathbf{p}) \lhd\rhd G(n,\mathbf{p}')$.

In fact, by (2.5), $\rho(p_{ij}, p'_{ij}) \asymp \delta_n^2 p_{ij} \wedge |\delta_n| p_{ij} = \delta_n^2 p_{ij}$, and thus by Theorem 2.2 the conditions $\delta_n^2 \sum_{i<j} p_{ij} \to 0$ and $\delta_n^2 \sum_{i<j} p_{ij} = O(1)$ are necessary too for asymptotic equivalence and contiguity, respectively. (The necessity can also be checked by considering the total number of edges, as in the special case in Example 1.5.)

Remark 3.10. As in Example 3.9, necessity in Theorem 2.2 can in many cases where $p'_{ij} \le p_{ij}$ for all i and j (or conversely) be proved by considering the total numbers $\sum_i I_{ni}$ and $\sum_i I'_{ni}$, but this method does not suffice in all cases. A simple counter example is given by $N(n) = n^4 + n^8$, $p_{ni} = n^{-1}$ for $1 \le i \le n^4$ and $p_{ni} = n^{-3}$ for $i > n^4$, and $p'_{ni} = p_{ni} - p_{ni}^2$; it is easily checked that then (2.7) and (2.8) do not hold, and thus we do not have asymptotic equivalence or even contiguity, but, using [2, Theorems 2.M and 1.C],

$$d_{\mathrm{TV}}\Bigl(\sum_i I_{ni}, \sum_i I'_{ni}\Bigr) = d_{\mathrm{TV}}\bigl(\mathrm{Po}(n^5+n^3), \mathrm{Po}(n^5+n^3-2n^2)\bigr) + o(1) \to 0.$$

4. More on asymptotic equivalence and contiguity

We will use two metrics to measure the distance between probability distributions. We state some well-known definitions and facts, see e.g. [2, Appendix A.1] and [11, Chapter IV.1 and V.4a].

Definition 4.1. If P and Q are two probability measures on the same measurable space $(\mathcal{X}, \mathcal{A})$, and R is any σ-finite measure on $(\mathcal{X}, \mathcal{A})$ such that $P \ll R$ and $Q \ll R$, define the *total variation distance*

$$d_{\mathrm{TV}}(P,Q) := \sup_{A\in\mathcal{A}} |P(A) - Q(A)| = \frac{1}{2}\int_{\mathcal{X}} \left|\frac{dP}{dR} - \frac{dQ}{dR}\right| dR \tag{4.1}$$

and the *Hellinger distance*

$$d_{\mathrm{H}}(P,Q) := \left(\frac{1}{2}\int_{\mathcal{X}} \left(\sqrt{\frac{dP}{dR}} - \sqrt{\frac{dQ}{dR}}\right)^2 dR\right)^{1/2} = \bigl(1 - H(P,Q)\bigr)^{1/2} \tag{4.2}$$

where $H(P,Q)$ is the *Hellinger integral*

$$H(P,Q) := \int_{\mathcal{X}} \sqrt{\frac{dP}{dR}}\sqrt{\frac{dQ}{dR}}\, dR. \tag{4.3}$$

(We can, at least symbolically, write (4.3) as $H(P,Q) := \int_{\mathcal{X}} \sqrt{dP\,dQ}$.) Note that these quantities do not depend on the choice of R. (We may thus take, e.g., $R = P + Q$.)

We have $d_{\mathrm{TV}}(P,Q) = \frac{1}{2}\|P - Q\|$, using the standard norm on real-valued measures. (The factor $\frac{1}{2}$ is conventional and convenient but unimportant, as is the factor $\frac{1}{2}$ in the definition of d_{H}.)

We use the same notations for two random variables X and Y with values in $\mathcal{X}$, with $d_{\mathrm{TV}}(X,Y) := d_{\mathrm{TV}}(\mathcal{L}(X), \mathcal{L}(Y))$ and similarly for d_{H} and H. (Thus $d_{\mathrm{TV}}(X,Y) = 0 \iff d_{\mathrm{H}}(X,Y) = 0 \iff X \overset{\mathrm{d}}{=} Y$.) In particular,

$$d_{\mathrm{TV}}(X,Y) := \sup_{A \in \mathcal{A}} |\mathbb{P}(X \in A) - \mathbb{P}(Y \in A)|. \tag{4.4}$$

It is easily seen that d_{TV} and d_{H} are metrics on the set of all probability measures on $(\mathcal{X}, \mathcal{A})$; further, $0 \le d_{\mathrm{TV}} \le 1$, $0 \le d_{\mathrm{H}} \le 1$ and $0 \le H \le 1$, and

$$d_{\mathrm{H}}^2(P,Q) \le d_{\mathrm{TV}}(P,Q) \le \sqrt{2}\, d_{\mathrm{H}}(P,Q); \tag{4.5}$$

hence d_{TV} and d_{H} are equivalent metrics. Furthermore, $d_{\mathrm{TV}}(P,Q) = 1 \iff d_{\mathrm{H}}(P,Q) = 1 \iff P \perp Q$, i.e., P and Q are mutually singular.

Recall that a *coupling* of two random variables X and Y with values in the same space is a pair of random variables (X', Y'), defined together on the same probability space, with $X' \overset{\mathrm{d}}{=} X$ and $Y' \overset{\mathrm{d}}{=} Y$. It is well-known that

$$d_{\mathrm{TV}}(X,Y) = \min_{(X',Y')} \mathbb{P}(X' \neq Y'), \tag{4.6}$$

taking the minimum over all couplings (X', Y') of X and Y.

Theorem 4.2. *Let X_n and Y_n be random variables with values in $\mathcal{X}_n$. Then the following are equivalent.*

- (i) $X_n \cong Y_n$.
- (ii) $d_{\mathrm{TV}}(X_n, Y_n) \to 0$.
- (iii) $d_{\mathrm{H}}(X_n, Y_n) \to 0$.
- (iv) $H(X_n, Y_n) \to 1$.
- (v) *There exist couplings (X_n', Y_n') of X_n and Y_n such that $\mathbb{P}(X_n' \neq Y_n') \to 0$.*

Proof. This too is well-known and easy: (i) $\iff$ (ii) by (4.4) and Definition 1.1; (ii) $\iff$ (iii) by (4.5); (iii) $\iff$ (iv) by (4.2); (ii) $\iff$ (v) by (4.6). ∎

We calculate the Hellinger distance and integral for two Bernoulli distributions. (This is the origin of our function ρ in Definition 2.1.)

Lemma 4.3. *For any $p, q \in [0,1]$,*

$$d_{\mathrm{H}}\big(\mathrm{Be}(p), \mathrm{Be}(q)\big) = 2^{-1/2} \rho(p,q)^{1/2},$$
$$H\big(\mathrm{Be}(p), \mathrm{Be}(q)\big) = 1 - \tfrac{1}{2}\rho(p,q).$$

Proof. Use (4.2) with $P = \mathrm{Be}(p) = p\delta_0 + (1-p)\delta_1$, $Q = \mathrm{Be}(q) = q\delta_0 + (1-q)\delta_1$ and $R = \delta_0 + \delta_1$, together with the definition (2.1). $\square$

An important, and well-known, property of Hellinger distances and integrals is that they behave simple for product measures. Let $[n] := \{1, \dots, n\}$ when $n < \infty$ and $[\infty] := \mathbb{N} = \{1, 2, \dots\}$.

Lemma 4.4. *Let $1 \le N \le \infty$ and let, for $i \in [N]$, P_i and Q_i be probability measures on the same measurable space $(\mathcal{X}_i, \mathcal{A}_i)$. If $P = \prod_{i=1}^N P_i$ and $Q = \prod_{i=1}^N Q_i$, then $H(P, Q) = \prod_{i=1}^N H(P_i, Q_i)$.*

Proof. This is stated in, e.g., [11, Proposition IV.1.73], but for completeness we give the simple proof.

If $N < \infty$, the result is an immediate consequence of (4.3) and Fubini's theorem, choosing e.g. $R_i = P_i + Q_i$ and $R = \prod_{i=1}^N R_i$.

If $N = \infty$, let $\mathcal{F}_n$ be the σ-field on $\prod_{i=1}^\infty \mathcal{X}_i$ given by $\{A \times \prod_{n+1}^\infty \mathcal{X}_i : A \in \prod_1^n \mathcal{A}_i\}$, and let $\overline{P}_n := P|_{\mathcal{F}_n}$ and $\overline{Q}_n := Q|_{\mathcal{F}_n}$. Then, using the finite case,

$$H(\overline{P}_n, \overline{Q}_n) = H\Bigl(\prod_1^n P_i, \prod_1^n Q_i\Bigr) = \prod_1^n H(P_i, Q_i).$$

Furthermore, choosing $R = (P + Q)/2$ on $\mathcal{X} := \prod_{i=1}^\infty \mathcal{X}_i$, $d\overline{P}_n/dR = \mathbb{E}(dP/dR \mid \mathcal{F}_n)$ with respect to R, so $(d\overline{P}_n/dR)$ is a bounded R-martingale and $d\overline{P}_n/dR \to dP/dR$ R-a.s., and similarly for $\overline{Q}_n$. Hence, (4.3) and dominated convergence yields $H(P, Q) = \lim_{n\to\infty} H(\overline{P}_n, \overline{Q}_n) = \prod_{i=1}^\infty H(P_i, Q_i)$. $\square$

5. Proofs

Proof of Theorem 2.2. (i): By Lemmas 4.4 and 4.3,

$$H(X_n, X_n') = \prod_1^{N(n)} H(I_{ni}, I_{ni}') = \prod_1^{N(n)} \bigl(1 - \tfrac12 \rho(p_{ni}, p_{ni}')\bigr).$$

Hence,

$$1 - \tfrac12 \sum_1^{N(n)} \rho(p_{ni}, p_{ni}') \le H(X_n, X_n') \le \exp\Bigl(-\tfrac12 \sum_1^{N(n)} \rho(p_{ni}, p_{ni}')\Bigr),$$

and thus $H(X_n, X_n') \to 1 \iff \sum_1^{N(n)} \rho(p_{ni}, p_{ni}') \to 0$, which yields the result by Theorem 4.2.

(ii): This is, in view of Lemma 4.3 and the equivalence of (2.9) and (2.11), a special case of [16, Theorem 1], to which we refer for a complete proof. Nevertheless, for completeness, we sketch a proof of the more important "if" direction.

First, we can by a simpler version of the argument in the proof of Theorem 2.9 below assume that $p_{ni} \le C_2 p_{ni}'$ and $q_{ni} \le C_2 q_{ni}'$ for some constant C_2 and all $i \le N(n)$. (We define p_{ni}''' by (5.3) with $p_{ni}'' := p_{ni}$ and use

(5.4)–(5.5).) Under this assumption, if we let $P_{ni} := \mathcal{L}(I_{ni}) = \mathrm{Be}(p_{ni})$, $P_n := \prod_i P_{ni}$, $P'_{ni} := \mathcal{L}(I'_{ni}) = \mathrm{Be}(p'_{ni})$, $P'_n := \prod_i P'_{ni}$, we have by Fubini, using $\int (dP_{ni}/dP'_{ni})\, dP'_{ni} = 1$ and (2.2),

$$\begin{aligned}
\int \Bigl(\frac{dP_n}{dP'_n}\Bigr)^2 dP'_n &= \prod_i \int \Bigl(\frac{dP_{ni}}{dP'_{ni}}\Bigr)^2 dP'_{ni} = \prod_i \Bigl(1 + \int \Bigl(\frac{dP_{ni}}{dP'_{ni}} - 1\Bigr)^2 dP'_{ni}\Bigr) \\
&= \prod_i \Bigl(1 + \Bigl(\frac{p_{ni}-p'_{ni}}{p'_{ni}}\Bigr)^2 p'_{ni} + \Bigl(\frac{q_{ni}-q'_{ni}}{q'_{ni}}\Bigr)^2 q'_{ni}\Bigr) \\
&= \prod_i \Bigl(1 + \frac{(p_{ni}-p'_{ni})^2}{p'_{ni}} + \frac{(p_{ni}-p'_{ni})^2}{1-p'_{ni}}\Bigr) \\
&\le \prod_i \Bigl(1 + (C_2+1)\frac{(p_{ni}-p'_{ni})^2}{p_{ni}+p'_{ni}} + (C_2+1)\frac{(p_{ni}-p'_{ni})^2}{1-p_{ni}+1-p'_{ni}}\Bigr) \\
&\le \prod_i \bigl(1 + (C_2+1)C\rho(p_{ni},p'_{ni})\bigr) \\
&\le \exp\Bigl((C_2+1)C\sum_i \rho(p_{ni},p'_{ni})\Bigr) \\
&= O(1)
\end{aligned}$$

and thus for any sets A_n, by the Cauchy–Schwarz inequality,

$$P_n(A_n) = \int_{A_n} dP_n \le \Bigl(\int \Bigl(\frac{dP_n}{dP'_n}\Bigr)^2 dP'_n \cdot \int_{A_n} dP'_n\Bigr)^{1/2} = O\bigl(P'_n(A_n)^{1/2}\bigr)$$

and thus $P_n \lhd P'_n$, which is the same as $X_n \lhd X'_n$. ☐

We say that a finite or infinite random vectors of indicator variables $X = (I_i)_{i=1}^N$ has distribution $\mathrm{Be}(\mathbf{p})$, where $\mathbf{p} = \{p_i\}_{i=1}^N$ is a deterministic vector with elements in $[0,1]$, if the random variables I_i are independent indicator variables with $I_i \sim \mathrm{Be}(p_i)$.

More generally, if $\mathbf{p} = \{p_i\}_{i=1}^N$ is a random vector with elements in $[0,1]$, with $N \le \infty$, we say that random vectors of indicator variables $X = (I_i)_{i=1}^N$ has distribution $\mathrm{Be}(\mathbf{p})$ if the conditioned random vector $(X \mid \mathbf{p})$ is a sequence of independent indicator variables with $(I_i \mid \mathbf{p}) \sim \mathrm{Be}(p_i)$.

We next give two results comparing two random vectors with distributions $\mathrm{Be}(\mathbf{p})$ and $\mathrm{Be}(\mathbf{p}')$ with deterministic $\mathbf{p}$ and $\mathbf{p}'$. The first result is easily seen to be equivalent to the "if" direction of Theorem 2.2(i), while the second is equivalent to a special case of the "if" direction of Theorem 2.2(ii).

Lemma 5.1. *For every $\varepsilon > 0$, there exists $\delta > 0$ such that if $X \sim \mathrm{Be}(\mathbf{p})$ and $X' \sim \mathrm{Be}(\mathbf{p}')$ for two deterministic vectors $\mathbf{p} = \{p_i\}_{i=1}^N$ and $\mathbf{p}' = \{p'_i\}_{i=1}^N$ of the same length $N \le \infty$, and these satisfy $\sum_{i=1}^N \rho(p_i, p'_i) < \delta$, then $d_{\mathrm{TV}}(X, X') < \varepsilon$.*

Proof. Suppose not. Then there exist $\varepsilon > 0$ and such random vectors $X_n \sim \operatorname{Be}(\mathbf{p}_n)$ and $X'_n \sim \operatorname{Be}(\mathbf{p}'_n)$ such that $\sum_1^{N(n)} \rho(p_{ni}, p'_{ni}) < 1/n$ but $d_{\mathrm{TV}}(X_n, X'_n) \geq \varepsilon$, but this contradicts Theorems 2.2(i) and 4.2. □

Lemma 5.2. *For every constants C_1, C_2 and $\varepsilon > 0$, there exists $\delta > 0$ such that if $X \sim \operatorname{Be}(\mathbf{p})$ and $X' \sim \operatorname{Be}(\mathbf{p}')$ for two deterministic vectors $\mathbf{p} = \{p_i\}_{i=1}^N$ and $\mathbf{p}' = \{p'_i\}_{i=1}^N$ of the same length $N \leq \infty$, and these satisfy $\sum_{i=1}^N \rho(p_i, p'_i) \leq C_1$ and further, for every $i \in [N]$, $p_i \leq C_2 p'_i$ and $(1-p_i) \leq C_2(1-p'_i)$, then for every set A with $\mathbb{P}(X' \in A) < \delta$, we have $\mathbb{P}(X \in A) < \varepsilon$.*

Proof. If not, it would be possible to find, for some fixed C_1, C_2 and ε, sequences X_n and X'_n of such random vectors and sets A_n such that $\mathbb{P}(X'_n \in A_n) < 1/n$ and $\mathbb{P}(X_n \in A_n) \geq \varepsilon$. In particular, $X_n \ntriangleleft X'_n$.

On the other hand, (2.8) and (2.9) hold for these random vectors (since the sums in (2.9) vanish for any $C \geq C_2$), and thus Theorem 2.2(ii) yields $X_n \triangleleft X'_n$, which is a contradiction. □

Proof of Theorem 2.9. (i): Let $\varepsilon > 0$ and choose $\delta > 0$ as in Lemma 5.1. Then, by Lemma 5.1 applied to the conditioned variables $(X_n \mid \mathbf{p}_n)$ and $(X'_n \mid \mathbf{p}'_n)$, if $\sum_i \rho(p_{ni}, p'_{ni}) < \delta$, then $d_{\mathrm{TV}}\bigl((X_n \mid \mathbf{p}_n), (X'_n \mid \mathbf{p}'_n)\bigr) < \varepsilon$. Since

$$\mathbb{P}(X_n \in A) - \mathbb{P}(X'_n \in A) = \mathbb{E}\bigl(\mathbb{P}(X_n \in A \mid \mathbf{p}_n) - \mathbb{P}(X'_n \in A \mid \mathbf{p}'_n)\bigr)$$

for every measurable $A \subseteq \mathcal{X}_n = \{0,1\}^{N(n)}$, it follows that

$$d_{\mathrm{TV}}(X_n, X'_n) \leq \mathbb{E}\, d_{\mathrm{TV}}\bigl((X_n \mid \mathbf{p}_n), (X'_n \mid \mathbf{p}'_n)\bigr) \leq \varepsilon + \mathbb{P}\Bigl(\sum_i \rho(p_{ni}, p'_{ni}) \geq \delta\Bigr).$$

The latter probability tends to 0 by assumption, and since ε is arbitrary, this yields $d_{\mathrm{TV}}(X_n, X'_n) \to 0$.

(ii): Let $(A_n)_n$ be an arbitrary sequence of measurable sets with $A_n \subseteq \mathcal{X}_n = \{0,1\}^{N(n)}$ and let $\varepsilon > 0$.

By (2.14), there exists C_1 such that $\mathbb{P}\bigl(\sum_i \rho(p_{ni}, p'_{ni}) > C_1\bigr) < \varepsilon$ for all n. Similarly, by (2.15), there exist $C_2 \geq 1$ and n_0 such that for $n \geq n_0$,

$$\mathbb{P}\left(\sum_{i:p_{ni} > C_2 p'_{ni}} p_{ni} + \sum_{i:q_{ni} > C_2 q'_{ni}} q_{ni} > \varepsilon\right) < \varepsilon;$$

in the sequel we consider only $n \geq n_0$.

Define $\mathbf{p}''_n = \{p''_{ni}\}_{i=1}^{N(n)}$ by

$$\mathbf{p}''_n := \begin{cases} \mathbf{p}'_n, & \sum_i \rho(p_{ni}, p'_{ni}) > C_1 \text{ or } \sum_{i:p_{ni} > C_2 p'_{ni}} p_{ni} + \sum_{i:q_{ni} > C_2 q'_{ni}} q_{ni} > \varepsilon; \\ \mathbf{p}_n, & \text{otherwise.} \end{cases}$$

By our choices of C_1 and C_2, $\mathbb{P}(\mathbf{p}''_n \neq \mathbf{p}_n) < 2\varepsilon$, and we may thus define $X''_n = (I''_{ni})_{i=1}^{N(n)} \sim \operatorname{Be}(\mathbf{p}''_n)$ such that

$$\mathbb{P}(X''_n \neq X_n) \leq \mathbb{P}(\mathbf{p}''_n \neq \mathbf{p}_n) < 2\varepsilon. \tag{5.1}$$

Moreover, by the construction, with $q''_{ni} := 1 - p''_{ni}$,

$$\sum_i \rho(p''_{ni}, p'_{ni}) \le C_1 \qquad \text{and} \qquad \sum_{i: p''_{ni} > C_2 p'_{ni}} p''_{ni} + \sum_{i: q''_{ni} > C_2 q'_{ni}} q''_{ni} \le \varepsilon. \tag{5.2}$$

Next, define $p'''_{ni} = \{p'''_{ni}\}_{i=1}^{N(n)}$ by

$$p'''_{ni} := \begin{cases} p'_{ni}, & p''_{ni} > C_2 p'_{ni} \text{ or } q''_{ni} > C_2 q'_{ni}; \\ p''_{ni}, & \text{otherwise.} \end{cases} \tag{5.3}$$

We can construct $X'''_n \sim \mathrm{Be}(\mathbf{p}'''_n)$ using maximal couplings of $(I'''_{ni} \mid \mathbf{p}'''_n)$ and $(I''_{ni} \mid \mathbf{p}''_n)$ so that, using (5.3) and (5.2),

$$\begin{aligned} d_{\mathrm{TV}}\bigl((X'''_n \mid \mathbf{p}'''_n), (X''_n \mid \mathbf{p}''_n)\bigr) &\le \sum_i d_{\mathrm{TV}}\bigl((I'''_{ni} \mid p'''_{ni}), (I''_{ni} \mid p''_{ni})\bigr) \\ &\le \sum_i |p'''_{ni} - p''_{ni}| \\ &\le \sum_{i: p''_{ni} > C_2 p'_{ni}} p''_{ni} + \sum_{i: q''_{ni} > C_2 q'_{ni}} q''_{ni} \le \varepsilon. \end{aligned} \tag{5.4}$$

Consequently,

$$d_{\mathrm{TV}}(X'''_n, X''_n) \le \mathbb{E}\, d_{\mathrm{TV}}\bigl((X'''_n \mid \mathbf{p}'''_n), (X''_n \mid \mathbf{p}''_n)\bigr) \le \varepsilon. \tag{5.5}$$

Furthermore, by (5.3), $p'''_{ni} \le C_2 p'_{ni}$ and $q'''_{ni} := 1 - p'''_{ni} \le C_2 q'_{ni}$ and by (5.3) and (5.2),

$$\sum_i \rho(p'''_{ni}, p'_{ni}) \le \sum_i \rho(p''_{ni}, p'_{ni}) \le C_1.$$

We can thus apply Lemma 5.2 to the conditioned variables $(X'''_n \mid \mathbf{p}'''_n)$ and $(X'_n \mid \mathbf{p}'_n)$ for all values of $\mathbf{p}'''_n$ and $\mathbf{p}'_n$. Consequently there exists $\delta > 0$ such that if $\mathbb{P}(X'_n \in A_n \mid \mathbf{p}'_n) < \delta$, then $\mathbb{P}(X'''_n \in A_n \mid \mathbf{p}'''_n) < \varepsilon$. Hence, using Markov's inequality,

$$\begin{aligned} \mathbb{P}(X'''_n \in A_n) = \mathbb{E}\,\mathbb{P}(X'''_n \in A_n \mid \mathbf{p}'''_n) &\le \varepsilon + \mathbb{P}\bigl(\mathbb{P}(X'_n \in A_n \mid \mathbf{p}'_n) \ge \delta\bigr) \\ &\le \varepsilon + \delta^{-1}\,\mathbb{E}\,\mathbb{P}(X'_n \in A_n \mid \mathbf{p}'_n) = \varepsilon + \delta^{-1}\,\mathbb{P}(X'_n \in A_n). \end{aligned}$$

Using (5.1) and (5.5), we thus obtain

$$\begin{aligned} \mathbb{P}(X_n \in A_n) &\le \mathbb{P}(X_n \ne X''_n) + d_{\mathrm{TV}}(X''_n, X'''_n) + \mathbb{P}(X'''_n \in A_n) \\ &\le 4\varepsilon + \delta^{-1}\,\mathbb{P}(X'_n \in A_n). \end{aligned}$$

If we assume that $\mathbb{P}(X'_n \in A_n) \to 0$, it follows that $\limsup \mathbb{P}(X_n \in A_n) \le 4\varepsilon$, and since ε is arbitrary thus $\mathbb{P}(X_n \in A_n) \to 0$, which shows that $X_n \lhd X'_n$. ∎

Proof of Corollary 2.12. In order to apply Theorem 2.9, we reorder $\{p_{ij}\}_{i<j}$ to $\{p_{ni}\}_{i=1}^{N(n)}$; we do this without further comment and keep the notation p_{ij}. We also let $q_{ij} := 1 - p_{ij}$ and $q'_{ij} := 1 - p'_{ij}$.

(i): By (2.5), whp $\rho(p_{ij}, p'_{ij}) \le C_0(p_{ij} - p'_{ij})^2/p_{ij}$ for some C_0, and thus (2.17) implies (2.13), and the conclusion follows by Theorem 2.9(i).

(ii): Similarly, by (2.5) again, (2.18) implies (2.14). Moreover, for any $C \ge 2$,

$$\sum_{i<j:\, p'_{ij}>Cp_{ij}} p'_{ij} \le \frac{1}{C} \sum_{i<j:\, p'_{ij}>Cp_{ij}} \frac{(p'_{ij})^2}{p_{ij}} \le \frac{4}{C} \sum_{i<j} \frac{(p'_{ij} - p_{ij})^2}{p_{ij}}. \tag{5.6}$$

Hence, for any sequence $C_n \to \infty$, (2.18) implies that $\sum_{i:p'_{ij}>C_n p_{ij}} p'_{ij} \xrightarrow{\mathrm{p}} 0$. Moreover, for any $C \ge 10$, whp, $Cq_{ij} \ge 1$ and thus $q'_{ij} \le Cq_{ij}$ for all i,j. It follows that (2.16) with $\mathbf{p}$ and $\mathbf{p}'$ interchanged holds, and thus (2.15) with $\mathbf{p}$ and $\mathbf{p}'$ interchanged holds. (The latter is also easily proved directly using (5.6).) Consequently, Theorem 2.9(ii) yields $G(n,\mathbf{p}') \lhd G(n,\mathbf{p})$.

(iii): The extra assumptions allow us to interchange $\mathbf{p}$ and $\mathbf{p}'$ in the assumptions. Hence (ii) yields both $G(n,\mathbf{p}) \rhd G(n,\mathbf{p}')$ and $G(n,\mathbf{p}') \rhd G(n,\mathbf{p})$. □

Proof of Corollary 2.13. An immediate consequence of Corollary 2.12, since now $(p_{ij} - p'_{ij})^2/p_{ij} = O(p_{ij}^3)$; note also that the assumption in (i) implies $\max_{i,j} p_{ij} = o_p(1)$ and thus $\max_{i,j} p_{ij} < 0.9$ whp. □

References

[1] A.-L. Barabási & R. Albert, Emergence of scaling in random networks, *Science* **286** (1999), 509–512.

[2] A. Barbour, L. Holst & S. Janson, *Poisson Approximation.* Oxford University Press, Oxford, 1992.

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

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

[5] B. Bollobás, S. Janson & O. Riordan, Sparse random graphs with clustering. *Random Struct. Alg.*, to appear.

[6] T. Britton, M. Deijfen & A. Martin-Löf, Generating simple random graphs with prescribed degree distribution, *J. Stat. Phys.* **124** (2006), no. 6, 1377–1397.

[7] R. Durrett, Rigorous result for the CHKNS random graph model, *Proceedings, Discrete Random Walks 2003 (Paris, 2003)*, eds. C. Banderier & Chr. Krattenthaler, Discrete Mathematics and Theoretical Computer Science **AC** (2003), pp. 95–104, `http://dmtcs.loria.fr/proceedings/`

[8] R. Durrett, *Random Graph Dynamics.* Cambridge Univ. Press, Cambridge, 2007.

[9] H. van den Esker, R. van der Hofstad & G. Hooghiemstra, Universality for the distance in finite variance random graphs. *J. Stat. Phys.* **133** (2008), no. 1, 169–202.

[10] C. Greenhill, S. Janson, J. H. Kim & N. C. Wormald. Permutation pseudographs and contiguity. *Combin. Probab. Comput.* **11** (2002), no. 3, 273–298.

[11] J. Jacod & A. N. Shiryaev, *Limit Theorems for Stochastic Processes.* Springer-Verlag, Berlin, 1987.

[12] S. Janson, Random regular graphs: asymptotic distributions and contiguity. *Combin. Probab. Comput.* **4** (1995), 369–405.

[13] S. Janson, The probability that a random multigraph is simple. *Combin. Probab. Comput.* **18** (2009), no. 1-2, 205–225.

[14] S. Janson, T. Łuczak & A. Ruciński, *Random Graphs*, Wiley, New York, 2000.

[15] S. Kalikow & B. Weiss, When are random graphs connected? *Israel J. Math.* **62** (1988), 257–268.

[16] J. Oosterhoff & W. R. van Zwet, A note on contiguity and Hellinger distance. *Contributions to statistics*, pp. 157–166, Reidel, Dordrecht, 1979.

[17] O. Riordan, The small giant component in scale-free random graphs, *Combin. Probab. Comput.* **14** (2005), 897–938.

[18] N. C. Wormald, Models of random regular graphs. *Surveys in combinatorics, 1999 (Canterbury)*, 239–298, London Math. Soc. Lecture Note Ser., 267, Cambridge Univ. Press, Cambridge, 1999.

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/`