J. Appl. Prob. **55,** 112–136 (2018)
*doi:*10.1017/jpr.2018.9
© *Applied Probability Trust* 2018

INHOMOGENEOUS RANDOM GRAPHS, ISOLATED VERTICES, AND POISSON APPROXIMATION

MATHEW D. PENROSE,* *University of Bath*

Abstract

Consider a graph on randomly scattered points in an arbitrary space, with any two points x, y connected with probability $\phi(x, y)$. Suppose the number of points is large but the mean number of isolated points is $O(1)$. We give general criteria for the latter to be approximately Poisson distributed. More generally, we consider the number of vertices of fixed degree, the number of components of fixed order, and the number of edges. We use a general result on Poisson approximation by Stein's method for a set of points selected from a Poisson point process. This method also gives a good Poisson approximation for U-statistics of a Poisson process.

Keywords: Inhomogeneous random graph; random connection model; stochastic block model; latent variable model; random geometric graph; Poisson approximation; Stein's method; U-statistic

2010 Mathematics Subject Classification: Primary 05C80
Secondary 60F05

1. Introduction

In the *inhomogeneous random graph* (IRG), each vertex has one of several possible types or *states*, where the space of possible states may be infinite. Given the states of the vertices, each possible edge is present with a probability that depends on the states of the two endpoints, independently of the other edges. This provides a very flexible class of random graph models; for example, the state of a vertex could represent spatial location, or it could represent the time at which a vertex is born, for a randomly evolving graph.

Such models are popular in statistical network modelling, where they go under names such as *stochastic block model* [24] (in which case the state space is usually taken to be finite) and *latent space model* [11] or *latent variable model* [23]. The IRG terminology dates from [25], and is prevalent in the more probabilistic literature, where the model has been studied in depth, for example, in [2] and [4]. Much of this literature is concerned with the birth of the giant component, but full connectivity has also been studied in [6].

For our purposes, the IRG is defined as follows. Let $(\mathbb{X}, \mathcal{F}, \mu)$ be a probability space (the state space). For $s > 0$, let $\mathcal{P}_s$ be the (random) set of points of a Poisson point process on $\mathbb{X}$ with mean measure (that is, intensity measure) $s\mu$. Also, for $n \in \mathbb{N}$, let $\mathcal{X}_n$ be the binomial point process consisting of n independent random elements of $\mathbb{X}$ with common distribution μ.

Suppose that $\phi : \mathbb{X} \times \mathbb{X} \to [0, 1]$ is a measurable symmetric function; we call such a function a *connection function*. Given finite $\mathcal{V} \subset \mathbb{X}$ (possibly with multiplicity), let $G(\mathcal{V}, \phi)$ be the random (undirected) graph with vertex set $\mathcal{V}$, with each pair $\{x, y\}$ of points of $\mathcal{V}$ connected by

Received 17 January 2017; revision received 20 September 2017.
* Postal address: Department of Mathematical Sciences, University of Bath, Bath BA2 7AY, UK.
Email address: m.d.penrose@bath.ac.uk

https://doi.org/10.1017/jpr.2018.9 Published online by Cambridge University Press

an edge with probability $\phi(x, y)$, independently of all other pairs. We are particularly interested in $G(\mathcal{P}_s, \phi)$ and $G(\mathcal{X}_n, \phi)$; we define these graphs more formally in Section 4.

In the special case where $\mathbb{X}$ is a region of Euclidean space and $\phi(x, y)$ is determined by the displacement $x - y$ (typically via $\|x - y\|$, where $\|\cdot\|$ denotes the Euclidean norm), the IRG is also known as the *soft random geometric graph* [22] or *random connection model* (RCM) [18]. If, in the Euclidean setting, we have $\phi(x, y) = \mathbf{1}_{\{\|x-y\|\le r\}}$ the IRG is known as the *random geometric graph* (RGG) [21] or *Gilbert graph.* These models are important in applications to wireless communications; see, for example, [7], [8], [10], [17], and [26].

For any graph G and any $j \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$, let $D_j(G)$ denote the number of vertices in G of degree j; also set $D_{\le j}(G) := \sum_{i=0}^{j} D_i(G)$. In particular, $D_0(G)$ is the number of isolated vertices. Of interest is the question of whether $D_{\le k}(G(\mathcal{P}_s, \phi))$ and $D_{\le k}(G(\mathcal{X}_{\lfloor s \rfloor}, \phi))$ are approximately Poisson distributed for large s, with k fixed, for example, $k = 0$. One reason for this interest is that $D_0(G) = 0$ (respectively, $D_{\le k}(G) = 0$) is clearly a necessary condition for G to be connected (respectively, $(k + 1)$-connected, assuming G has at least $k + 2$ vertices), and for many choices of $\mathbb{X}$ and ϕ_n this condition is asymptotically sufficient (in probability) when $G = G(\mathcal{X}_n, \phi_n)$ with large n; see [9], [12], [19], [20], and [22]. In such cases, $\mathbb{P}[D_{\le k}(G(\mathcal{X}_n, \phi_n)) = 0]$ is a good approximation for the probability that $G(\mathcal{X}_n, \phi_n)$ is $(k + 1)$-connected, so if we can estimate the former via Poisson approximation, we may have a useful estimate for the probability of $(k + 1)$-connectivity.

The references just mentioned show that in many cases where s is large and ϕ is chosen so that $\mathbb{E}[D_j(G(\mathcal{P}_s, \phi))] = O(1)$, the distributions of $D_j(G(\mathcal{P}_s, \phi))$ and of $D_j(G(\mathcal{X}_{\lfloor s \rfloor}, \phi))$ are approximately Poisson. In this paper we give a general criterion for this to be the case, without making any geometrical or topological assumptions on the space $\mathbb{X}$ (Theorem 2.1). We give similar results for the number of components of order k with k fixed (Theorems 2.2 and 2.3), and for the number of edges (Theorem 2.4). These theorems also incorporate asymptotic normality, when the mean of the variable in question grows slowly as a function of s.

We prove Theorems 2.1–2.4 in Sections 5–7. The proofs use Theorem 3.1, a general result on Poisson approximation for functionals of Poisson processes, which is of independent interest. This theorem also gives us further results on the Poisson approximation for the number of edges under different assumptions from those of Theorem 2.4. See Theorem 7.1 and the subsequent discussion. It also provides an alternative way to derive (and slightly improve) a result of [5] on Poisson approximation of U-statistics of a Poisson process; see Section 8.

2. Results on IRGs

Let Φ be the class of all measurable symmetric functions from $\mathbb{X} \times \mathbb{X}$ to $[0, 1]$, and, for $\varepsilon > 0$, let Φ_ε be the class of all $\phi \in \Phi$ that satisfy

$$\inf_{x\in\mathbb{X}} \left(\int_{\mathbb{X}} \phi(x, y)\mu(\mathrm{d}y) \right) \ge \varepsilon \sup_{x\in\mathbb{X}} \left(\int_{\mathbb{X}} \phi(x, y)\mu(\mathrm{d}y) \right).$$

If $\phi \in \Phi_\varepsilon$, we say the connection function ϕ is *ε-homogeneous*. Note that 1-homogeneity, according to our definition, is the same as homogeneity as defined in [6].

Several of our results require ε-homogeneity. There are many interesting classes of connection function which satisfy this condition. For example, suppose that $\mathbb{X}$ is a bounded convex region in Euclidean space $\mathbb{R}^d$, and μ has a density (with respect to Lebesgue measure on that region) that is bounded away from 0 and ∞. Then there exists $\varepsilon > 0$ such that all connection functions of the form $\phi(x, y) = \psi(|x - y|)$, with ψ nonincreasing, lie in Φ_ε. In particular, we do not require any exponential decay condition on ψ, such as that imposed in [22].

https://doi.org/10.1017/jpr.2018.9 Published online by Cambridge University Press

For $\phi \in \Phi$, set $\overline{\phi} = \sup_{x,y} \phi(x, y)$. As well as ε-homogeneity, some of our results also require that $\overline{\phi}$ be bounded away from 1. This condition is annoying since it excludes from consideration the standard RGG, and also some cases of the IRG which have been considered elsewhere in the literature [4], [6]. On the other hand, 'soft' versions of these models, which do satisfy this condition, are easily defined and arguably will often be reasonable from the point of view of applications (we will discuss this further below). Without this condition, it may be unrealistic to expect to find a simple argument for the Poisson approximation of D_0 covering a large class of RGG densities without getting involved in geometrical details of any particular probability density or even assuming any Euclidean structure at all, which is what our condition on $\overline{\phi}$ allows us to do.

Suppose we have connection functions ϕ_s defined for each $s > 0$ with $\mathbb{E}[D_{\le j}(G(\mathcal{P}_s, \phi_s))]$ bounded as $s \to \infty$, for some fixed $j \in \mathbb{N}_0$. See (5.1) below for a formula for $\mathbb{E}[D_j(G(\mathcal{P}_s, \phi_s))]$. Our first result shows that $D_j(G(\mathcal{P}_s, \phi_s))$ and $D_j(G(\mathcal{X}_{\lfloor s \rfloor}, \phi_{\lfloor s \rfloor}))$ are approximately Poisson and $D_k(G(\mathcal{P}_s, \phi_s))$ is approximately normal for $k > j$, for large s under the condition that $\phi_s \in \Phi_\varepsilon$ for all s and $\overline{\phi}_s$ is bounded away from 1.

For $\alpha \in (0, \infty)$, let Z_α denote a random variable having the Poisson distribution with parameter α. Let $\mathcal{N}$ denote a random variable having the standard normal distribution in $\mathbb{R}$, that is, with probability density function $(2\pi)^{-1/2} \exp(-x^2/2)$, $x \in \mathbb{R}$. For any graph G, we define $D_{\le -1}(G) := 0$. For any two $(0, \infty)$-valued functions $u(s)$ and $v(s)$, we say that $u(s) = O(v(s))$ if $\limsup_{s\to\infty} u(s)/v(s) < \infty$, $u(s) = o(v(s))$ if $\limsup_{s\to\infty} u(s)/v(s) = 0$, and $u(s) = \Theta(v(s))$ if $u(s) = O(v(s))$ and $v(s) = O(u(s))$.

Theorem 2.1. *Let $j, k \in \mathbb{N}_0$, $\varepsilon > 0$, and $\phi_s \in \Phi_\varepsilon$ with $\overline{\phi}_s \le 1 - \varepsilon$ for $s \in (0, \infty)$, satisfying $\lim_{s\to\infty} \mathbb{E}[D_{\le j}(G(\mathcal{P}_s, \phi_s))] = \alpha \in (0, \infty)$. Then, as $s \to \infty$,*

$$D_j(G(\mathcal{P}_s, \phi_s)) \xrightarrow{\mathrm{D}} Z_\alpha, \qquad D_j(G(\mathcal{X}_{\lfloor s \rfloor}, \phi_{\lfloor s \rfloor})) \xrightarrow{\mathrm{D}} Z_\alpha, \tag{2.1}$$

$$D_{\le j-1}(G(\mathcal{P}_s, \phi_s)) \xrightarrow{\mathbb{P}} 0, \qquad D_{\le j-1}(G(\mathcal{X}_{\lfloor s \rfloor}, \phi_{\lfloor s \rfloor})) \xrightarrow{\mathbb{P}} 0, \tag{2.2}$$

and, moreover,

$$\mathbb{E}[D_k(G(\mathcal{P}_s, \phi_s))] = \Theta((\log s)^{k-j}) \tag{2.3}$$

and if also $k > j$ then

$$\frac{D_k(G(\mathcal{P}_s, \phi_s)) - \mathbb{E}[D_k(G(\mathcal{P}_s, \phi_s))]}{\sqrt{\mathbb{E}[D_k(G(\mathcal{P}_s, \phi_s))]}} \xrightarrow{\mathrm{D}} \mathcal{N}. \tag{2.4}$$

It is interesting to compare the conclusions (2.1) and (2.2) of this result with the example of [21, p. 55]. In that case, for a certain sequence of RGGs one can arrange for the number of vertices of degree two to be asymptotically *compound* Poisson, whereas here it is asymptotically Poisson.

For $k \in \mathbb{N}$, and for any graph G, we refer to the components of G of order k (that is, with k vertices) as the *k-components* of G. Let $N_k(G)$ denote the number of k-components in G. In particular, $N_1(G) = D_0(G)$.

Suppose we have connection functions ϕ_s defined for each $s > 0$ with $\mathbb{E}[N_k(G(\mathcal{P}_s, \phi_s))]$ bounded as $s \to \infty$ for some fixed $k \in \mathbb{N}$. In Theorem 2.3 below we show that $N_k(G(\mathcal{P}_s, \phi_s))$ and $N_k(G(\mathcal{X}_{\lfloor s \rfloor}, \phi_{\lfloor s \rfloor}))$ are approximately Poisson under the condition that $\phi_s \in \Phi_\varepsilon$ and $\overline{\phi}_s \le 1 - \varepsilon$ for all s and some fixed $\varepsilon > 0$; even without this condition, Theorem 2.2 shows that the Poisson approximation holds provided that $\overline{\phi}_s = o(1/\log s)$ (in the case of $G(\mathcal{P}_s, \phi_s)$), or provided that $\overline{\phi}_{\lfloor s \rfloor} = o(s^{-1/2})$ (in the case of $G(\mathcal{X}_{\lfloor s \rfloor}, \phi_{\lfloor s \rfloor})$). These rates of decay imposed

https://doi.org/10.1017/jpr.2018.9 Published online by Cambridge University Press

on $\overline{\phi}_s$ are significantly milder than the condition $p_n = \Theta((\log n)/n)$ for the Erdős–Rényi random graph $G(n, p_n)$ to be at the threshold for having no isolated vertices.

We also give a result on asymptotic normality, when $\mathbb{E}[N_k(G(\mathcal{P}_s, \phi_s))]$ grows slowly as $s \to \infty$. All asymptotics in the next two theorems are as $s \to \infty$.

Theorem 2.2. *Let $k \in \mathbb{N}$ and $\phi_s \in \Phi$ for $s > 0$, with $\mathbb{E}[N_k(G(\mathcal{P}_s, \phi_s))] \to \alpha \in (0, \infty)$.*

(i) *If $\overline{\phi}_s = o(1/\log s)$ then $N_k(G(\mathcal{P}_s, \phi_s)) \xrightarrow{\mathrm{D}} Z_\alpha$.*

(ii) *If $\overline{\phi}_s = o(s^{-1/2})$ then $N_k(G(\mathcal{X}_{\lfloor s \rfloor}, \phi_{\lfloor s \rfloor})) \xrightarrow{\mathrm{D}} Z_\alpha$.*

Theorem 2.3. *Let $k \in \mathbb{N}$, $\varepsilon > 0$, and $\phi_s \in \Phi_\varepsilon$ for $s > 0$, with $\overline{\phi}_s \le 1 - \varepsilon$ for all s.*

(i) *If $\mathbb{E}[N_k(G(\mathcal{P}_s, \phi_s))] \to \alpha \in (0, \infty)$ then*

$$N_k(G(\mathcal{P}_s, \phi_s)) \xrightarrow{\mathrm{D}} Z_\alpha, \qquad N_k(G(\mathcal{X}_{\lfloor s \rfloor}, \phi_{\lfloor s \rfloor})) \xrightarrow{\mathrm{D}} Z_\alpha. \tag{2.5}$$

(ii) *If $\mathbb{E}[N_k(G(\mathcal{P}_s, \phi_s))] \to \infty$, but $\mathbb{E}[N_k(G(\mathcal{P}_s, \phi_s))] = o(s)$, then by setting $\alpha_s = \mathbb{E}[N_k(G(\mathcal{P}_s, \phi_s))]$, we have $(N_k(G(\mathcal{P}_s, \phi_s)) - \alpha_s)\sqrt{\alpha_s} \xrightarrow{\mathrm{D}} \mathcal{N}$.*

Among other things, the $k = 1$ case of Theorem 2.2(i) generalizes Lemma 3.2 of [22] to a more general class of $(\mathbb{X}, \phi_s)$ than is considered in [22].

Given a connection function $\phi \in \Phi$, we define

$$\kappa(\phi) := \sup_{x \in \mathbb{X}} \int_{\mathbb{X}} \phi(x, y) \mu(\mathrm{d}y). \tag{2.6}$$

If ϕ is ε-homogeneous for fixed $\varepsilon > 0$ then, for any $x \in \mathbb{X}$, the expected degree of a vertex of $G(\mathcal{P}_s, \phi)$ located at x is of the order of $s\kappa(\phi)$. Our proofs of Theorems 2.1 and 2.3 actually show that under conditions of ε-homogeneity and $\overline{\phi}_s \le 1 - \varepsilon$, if $s\kappa(\phi_s) \to \infty$ (so the 'typical degrees' become large) then, for any fixed $j \ge 0$ and $k \ge 2$, both D_j and N_k are approximately Poisson distributed (so if the mean of this Poisson distribution goes to ∞ then they are approximately normally distributed). Moreover, under the conditions of Theorem 2.1 or Theorem 2.3(i), we will show that $s\kappa(\phi_s) = \Theta(\log s)$ (see (5.2) and (6.12)).

In Theorem 2.1 we do not address asymptotic normality $D_j(G(\mathcal{P}_s, \phi_s))$ when its expected value grows to ∞ more slowly than s, except for the $j = 0$ case which is covered by the $k = 1$ case of Theorem 2.3(ii). In attempting to adapt the proof of the latter to the $j > 0$ case, the difficulty is that, in general, expression (5.8) does not vanish in the case where $s\kappa(\phi_s)$ tends to 0. We conjecture that asymptotic normality of $D_j(G(\mathcal{P}_s, \phi_s))$ can be proved by other means but this is beyond the scope of this paper.

Our next result is concerned with Poisson or normal approximation for (a generalization of) the number of edges of $G(\mathcal{P}_s, \phi_s)$. This is of interest in itself, and will also be of use in the proof of Theorem 2.3. Given $k \in \mathbb{N}$ with $k \ge 2$, for any graph G, let $H_k(G)$ denote the number of connected induced subgraphs of G of order k (so, in particular, $H_2(G)$ is the number of edges.) Again, asymptotics are as $s \to \infty$.

Theorem 2.4. *Let $\varepsilon > 0$, $\phi_s \in \Phi_\varepsilon$ for all $s > 0$, and $k \in \mathbb{N}$ with $k \ge 2$.*

(i) *If $\mathbb{E}[H_k(G(\mathcal{P}_s, \phi_s))] \to \alpha \in (0, \infty)$ then $H_k(G(\mathcal{P}_s, \phi_s)) \xrightarrow{\mathrm{D}} Z_\alpha$.*

(ii) *If $\mathbb{E}[H_k(G(\mathcal{P}_s, \phi_s))] \to \infty$, but $\mathbb{E}[H_k(G(\mathcal{P}_s, \phi_s))] = o(s)$, then on setting $\alpha_s := \mathbb{E}[H_k(G(\mathcal{P}_s, \phi_s))]$, we have $(H_k(G(\mathcal{P}_s, \phi_s)) - \alpha_s)/\sqrt{\alpha_s} \xrightarrow{\mathrm{D}} \mathcal{N}$.*

https://doi.org/10.1017/jpr.2018.9 Published online by Cambridge University Press

The proofs of Theorems 2.1, 2.3, and 2.4 provide information about the rates of convergence in these results. Theorem 2.2 will be proved by the method of moments, which does not provide any information about rates. The reason part (ii) of this result requires a stronger condition on $\overline{\phi}_s$ than part (i) is because when bounding these moments from below, for part (i) we have a sequence of the form $\exp(-nx_n)$ (with x_n bounded by a constant times $\overline{\phi}_n$), while for (ii) we have a sequence of the form $(1-x_n)^n$ which is asymptotic to $\exp(-nx_n)$ only when $x_n = o(n^{-1/2})$.

We now discuss some of the literature related to these results. Much of this concerns the RGG. Suppose that $(\mathbb{X}, \mathcal{F}, \mu)$ is the d-dimensional Euclidean space with the Borel σ-algebra and with μ having a density f with respect to Lebesgue measure, while the connection function is given by $\phi_s(x, y) = \mathbf{1}_{\{|x-y|\le r_s\}}$, with r_s chosen in a way that $\mathbb{E}[D_0(G(\mathcal{P}_s, \phi_s))] \to \alpha \in (0, \infty)$. It has been shown that $D_0 \xrightarrow{\mathrm{D}} Z_\alpha$ when μ is uniform on the unit cube [20], and for certain special types of density with unbounded support [9], [12], [19]. However, D_0 is not always asymptotically Poisson; see the last paragraph of [12, Section 2], where it was suggested that the Poisson limit is 'the exception rather than the rule' in dimension $d = 1$. We are a long way from having any complete characterisation of distributions for which the number of isolated vertices in the RGG is asymptotically Poisson.

Suppose that $(\mathbb{X}, \mathcal{F}, \mu)$ is the d-dimensional unit cube equipped with Lebesgue measure. If ϕ_s is of the form $\phi_s(x, y) = p_s \mathbf{1}_{\{|x-y|\le r_s\}}$, with p_s and r_s chosen so that $\mathbb{E}[D_0(G(\mathcal{P}_s, \phi_s))] \to \alpha$, then $D_0 \xrightarrow{\mathrm{D}} Z_\alpha$. More generally, if ϕ_s is of the form $\phi_s(x, y) = \psi_s(|x - y|)$ with ψ_s a decreasing function satisfying an exponential decay condition, then the Poisson convergence of D_0 is known to hold; see [22]. Our results enable us to relax the exponential decay condition and allow for other distributions, for example with density bounded away from 0 and ∞ on a convex compact region in $\mathbb{R}^d$.

Here is a simple example of a sequence of connection functions ϕ_s where $D_0(G(\mathcal{P}_s, \phi_s))$ has bounded mean, but is *not* asymptotically Poisson. Let $(\mathbb{X}, \mathcal{F}, \mu)$ be the unit interval equipped with Lebesgue measure, and set $\phi_s(x, y) = 1$ if $\max(x, y) \le s^{-1}$ or $\min(x, y) > s^{-1}$, otherwise setting $\phi_s(x, y) = 0$. Then, in the large s limit, the random variable $D_0(G(\mathcal{P}_s, \phi_s))$ is Bernoulli distributed with parameter e^{-1}, not Poisson distributed. The condition of ε-homogeneity, which appears in many of our results, rules out this sort of example.

Devroye and Fraiman [6] considered $D_0(G(\mathcal{X}_n, \phi_n))$ on a general space $\mathbb{X}$ in the case where $\phi_n = (a(n)\phi_0) \wedge 1$ for a fixed function $\phi_0 : \mathbb{X} \times \mathbb{X} \to \mathbb{R}_+$ (such a function ϕ_0 is called a *kernel*), and some sequence $a(n)$ (where $\wedge$ denotes minimum). This is a common assumption for the IRG; see, for example, [4]. Their results may be interpreted as saying that (under certain conditions) the threshold value of a above which $G(\mathcal{X}_n, a\phi_0 \wedge 1)$ is free of isolated vertices (which may be viewed as a random variable) satisfies a weak law of large numbers; it is asymptotic to a constant times $(\log n)/n$, in probability. They also derive a similar law of large numbers for the threshold for the graph to be connected.

When ϕ_0 is bounded, our Theorem 2.2(ii) yields a possibility of deriving convergence in distribution for this threshold (suitably transformed). For example, suppose that $\mathbb{X} = [0, 1)$ and μ has a density f with respect to Lebesgue measure, with f bounded away from 0. Suppose also that $\phi_0(x, y)$ is determined via the absolute value of $x - y$ (mod 1) (so we are in the one-dimensional torus). If $f \equiv 1$, or if f is smooth with $g(x) := \int \phi_0(x, y) f(y) (\mathrm{d}y)$ having a unique minimum, then it should be possible to derive a distributional limit law for this threshold, since in these cases it should be possible, for any $\alpha \in (0, \infty)$, to determine a sequence a_s such that $\mathbb{E}[D_0(G(\mathcal{P}_s, a_s\phi_0))] \to \alpha$, and then apply Theorem 2.2(ii).

https://doi.org/10.1017/jpr.2018.9 Published online by Cambridge University Press

If ϕ_0 is unbounded but shift-invariant on the torus, for example, if $\phi_0(x, y) = ((x - y) \bmod 1)^{-\gamma}$ for some fixed positive γ and $f \equiv 1$, then for any $\varepsilon \in (0, 1)$, it may be possible to find a limiting distribution for a suitable transformation of the threshold value of a above which $G(\mathcal{X}_n, a\phi_0 \wedge (1 - \varepsilon))$ is free of isolated vertices, now using Theorem 2.3(i). Indeed, in this case the connection functions are all 1-homogeneous, and by using connection functions $\phi_n = a_n\phi_0 \wedge (1 - \varepsilon)$ rather than the more standard $\phi_n = a_n\phi_0 \wedge 1$, we ensure that the extra condition $\overline{\phi}_n \leq 1 - \varepsilon$ is also satisfied.

It would be interesting to fully work out and extend these examples; to obtain a similar results for the connectivity threshold; and to improve the weak law of [6] to a strong law.

In this paper we consider only undirected graphs. Analogous directed graph models can be defined similarly; these have been considered in the random geometric graphs literature [13] and in the statistical literature [11]. In this case, the connection function $\phi(x, y)$ represents the probability that there is a directed edge from a vertex at x to a vertex at y. It is not required to be symmetric, and can be adapted so that all vertices have the same expected outdegree, as in [13]. It would be interesting to try to derive a similar result to Theorem 2.1 for the number of vertices with outdegree j in such a model.

3. A general result on Poisson approximation

Let $(\mathbb{M}, \mathcal{M}, \boldsymbol{m})$ be a probability space (known as a *mark space*). Assume that the probability measure $\boldsymbol{m}$ on $\mathbb{M}$ is *diffuse*, by which we mean that there is a product measurable set $A \subset \mathbb{M} \times \mathbb{M}$ with $(\boldsymbol{m} \otimes \boldsymbol{m})(A) = 0$, such that the diagonal $\{(t, t) : t \in \mathbb{M}\}$ is contained in A. For example, if $\{t\} \in \mathcal{M}$ and $\boldsymbol{m}(\{t\}) = 0$ for all $t \in \mathbb{M}$, then $\boldsymbol{m}$ is diffuse.

Suppose that, on a suitable probability space, we have a sequence $((X_i, T_i),\ i = 1, 2, 3, \ldots)$ of independent and identically distributed (i.i.d.) random elements of $\mathbb{X} \times \mathbb{M}$ with common distribution $\mu \otimes \boldsymbol{m}$, and an independent unit-rate Poisson counting process $(Z_s,\ s > 0)$, so that the random variable Z_s has Poisson(s) distribution for each s, and also a further independent sequence $(\tau, \tau_1, \tau_2, \tau_3, \ldots)$ of independent random elements of $\mathbb{M}$ with common distribution $\boldsymbol{m}$. By our assumption that $\boldsymbol{m}$ is diffuse, the values of $\tau, T_1, \tau_1, T_2, \tau_2, T_3, \tau_3, \ldots$ are almost surely distinct.

A *finite point process* in $\mathbb{X}$ is defined as a random element of the space $\boldsymbol{S}(\mathbb{X})$ of all finite subsets of $\mathbb{X}$, where $\boldsymbol{S}(\mathbb{X})$ is equipped with the smallest σ-algebra $\mathcal{S}(\mathbb{X})$ containing the sets $\{\xi \in \boldsymbol{S}(\mathbb{X}) : \xi(B) = k\}$ for all $B \in \mathcal{F}$ and all $k \in \mathbb{N}_0$, where $\xi(B) := |\xi \cap B|$ and $|\xi|$ denotes the number of elements of ξ.

A finite point process in $\mathbb{X} \times \mathbb{M}$ is defined similarly as a random element of the space $\boldsymbol{S} := \boldsymbol{S}(\mathbb{X} \times \mathbb{M})$, where $\mathbb{X} \times \mathbb{M}$ is equipped with the product σ-algebra $\mathcal{F} \otimes \mathcal{M}$. For $k \in \mathbb{N}$, let $\boldsymbol{S}_k := \{\xi \in \boldsymbol{S} : |\xi| = k\}$.

Given $s > 0$, $n \in \mathbb{N}$, define the following point processes in $\mathbb{X} \times \mathbb{M}$:

$$\eta_s := \bigcup_{i=1}^{Z_s} \{(X_i, T_i)\}, \qquad \xi_n := \bigcup_{i=1}^{n} \{(X_i, T_i)\}. \tag{3.1}$$

Then η_s is a Poisson point process in $\mathbb{X} \times \mathbb{M}$ with mean measure $s\mu \times \boldsymbol{m}$. Similarly, ξ_n is a binomial point process in $\mathbb{X} \times \mathbb{M}$.

Let d_{TV} and d_{W} denote total variation distance and Wasserstein distance, respectively, between probability measures on the nonnegative integers. That is, for $\mathbb{N}_0$-valued random

https://doi.org/10.1017/jpr.2018.9 Published online by Cambridge University Press

variables X, Y with distribution $\mathcal{L}(X)$, $\mathcal{L}(Y)$, respectively, we set

$$d_{\mathrm{TV}}(\mathcal{L}(X), \mathcal{L}(Y)) = \sup_{A \subset \mathbb{N}_0} (\mathbb{P}[X \in A] - \mathbb{P}[Y \in A]),$$
$$d_{\mathrm{W}}(\mathcal{L}(X), \mathcal{L}(Y)) = \sup\{|\mathbb{E}[h(X)] - \mathbb{E}[h(Y)]| \colon \|\Delta h\|_\infty \le 1\},$$

where, for $h \colon \mathbb{N}_0 \to \mathbb{R}$, we set $\Delta h(i) := h(i+1) - h(i)$ for $i \in \mathbb{N}_0$, and $\|h\|_\infty := \sup_{i \in \mathbb{N}_0} |h(i)|$.

The following theorem is related to a well-known result on the Poisson approximation of a sum of Bernoulli random variables by Stein's method via coupling (see Theorem II.24.3 of [16] or Theorem 1.B of [1]). Here the terms in the sum are themselves indexed by k-subsets of the set of points of a (marked) Poisson point process.

Let $k \in \mathbb{N}$ and let $f \colon \boldsymbol{S}_k \times \boldsymbol{S} \to \{0, 1\}$ be a measurable function. For $\xi \in \boldsymbol{S}$, set

$$F(\xi) := \sum_{\{\psi \subset \xi \colon |\psi| = k\}} f(\psi, \xi \setminus \psi). \tag{3.2}$$

We can think of f as a mechanism for selecting some of the k-subsets of ξ, and $F(\xi)$ as the total number of k-subsets selected.

Theorem 3.1. *Let $s > 0$. Let $W := F(\eta_s)$ with η_s and F as described above. For $x_1, \dots, x_k \in \mathbb{X}$, set $p(x_1, \dots, x_k) := \mathbb{E}[f(\{(x_1, \tau_1), \dots, (x_k, \tau_k)\}, \eta_s)]$, and set $\lambda := s\mu$.*

Suppose that $w \colon \mathbb{X}^k \to [0, \infty)$ is a measurable function, and that, for λ^k-almost every $\boldsymbol{x} = (x_1, \dots, x_k) \in \mathbb{X}^k$ with $p(x_1, \dots, x_k) > 0$, we can find coupled random variables $U_{\boldsymbol{x}}$, $V_{\boldsymbol{x}}$ such that

- $\mathcal{L}(U_{\boldsymbol{x}}) = \mathcal{L}(W)$;
- $\mathcal{L}(1 + V_{\boldsymbol{x}}) = \mathcal{L}(F(\bigcup_{i=1}^k \{(x_i, \tau_i)\} \cup \eta_s) \mid f(\bigcup_{i=1}^k \{(x_i, \tau_i)\}, \eta_s) = 1)$;
- $\mathbb{E}[|U_{\boldsymbol{x}} - V_{\boldsymbol{x}}|] \le w(\boldsymbol{x})$.

Set $\alpha = \mathbb{E}[W] = (1/k!) \int p(\boldsymbol{x}) \lambda^k(\mathrm{d}\boldsymbol{x})$. Then

$$d_{\mathrm{TV}}(\mathcal{L}(W), \mathcal{L}(Z_\alpha)) \le \frac{(1 \wedge \alpha^{-1})}{k!} \int w(\boldsymbol{x}) p(\boldsymbol{x}) \lambda^k(\mathrm{d}\boldsymbol{x}) \tag{3.3}$$

and

$$d_{\mathrm{W}}(\mathcal{L}(W), \mathcal{L}(Z_\alpha)) \le \frac{3(1 \wedge \alpha^{-1/2})}{k!} \int w(\boldsymbol{x}) p(\boldsymbol{x}) \lambda^k(\mathrm{d}\boldsymbol{x}). \tag{3.4}$$

The proof uses the (multivariate) *Mecke formula* (see, for example, [15] or [21, Theorem 1.6]), which says that if $g \colon \boldsymbol{S}_k \times \boldsymbol{S} \to \mathbb{R}$ is a bounded measurable function, then

$$\mathbb{E}\Big[\sum_{\{\psi \subset \eta_s \colon |\psi| = k\}} g(\psi, \eta_s \setminus \psi)\Big] = \frac{1}{k!} \int \mathbb{E}[g(\{(x_1, \tau_1), \dots, (x_k, \tau_k)\}, \eta_s)] \lambda^k(\mathrm{d}\boldsymbol{x}). \tag{3.5}$$

The assertion in the statement of the theorem that $\mathbb{E}[W] = (1/k!) \int p(\boldsymbol{x}) \lambda^k(\mathrm{d}\boldsymbol{x})$ follows from this fact.

Theorem 3.1 still holds in the case where the measure λ is taken to be σ-finite but infinite, and η_s is replaced by a Poisson point process η with mean measure $\lambda \otimes \boldsymbol{m}$. The proof is essentially unchanged.

https://doi.org/10.1017/jpr.2018.9 Published online by Cambridge University Press

Proof of Theorem 3.1. Let $h\colon \mathbb{N}_0 \to \mathbb{R}$ be bounded. Then

$$\begin{aligned}\mathbb{E}[Wh(W)] &= \mathbb{E}\bigg[\sum_{\{\psi\subset\eta_s\colon |\psi|=k\}} f(\psi, \eta_s\setminus\psi)h(F(\eta_s))\bigg]\\ &= \frac{1}{k!}\int \mathbb{E}\bigg[f(\{(x_1,\tau_1),\dots,(x_k,\tau_k)\},\eta_s)h\bigg(F\bigg(\bigcup_{i=1}^k\{(x_i,\tau_i)\}\cup\eta_s\bigg)\bigg)\bigg]\lambda^k(\mathrm{d}\boldsymbol{x})\\ &= \frac{1}{k!}\int \mathbb{E}\bigg[h\bigg(F\bigg(\bigcup_{i=1}^k\{(x_i,\tau_i)\}\cup\eta_s\bigg)\bigg)\;\bigg|\; f\bigg(\bigcup_{i=1}^k\{(x_i,\tau_i)\},\eta_s\bigg)=1\bigg]p(\boldsymbol{x})\lambda^k(\mathrm{d}\boldsymbol{x}).\end{aligned}$$

Also, $\mathbb{E}[\alpha h(W+1)] = (1/k!)\int \mathbb{E}[h(W+1)]p(\boldsymbol{x})\lambda^k(\mathrm{d}\boldsymbol{x})$ and, therefore,

$$\begin{aligned}&|\mathbb{E}[\alpha h(W+1) - Wh(W)]|\\ &\quad\le \frac{1}{k!}\int p(\boldsymbol{x})\lambda^k(\mathrm{d}\boldsymbol{x})\\ &\qquad\times\bigg|\mathbb{E}[h(W+1)] - \mathbb{E}\bigg[h\bigg(F\bigg(\bigcup_{i=1}^k\{(x_i,\tau_i)\}\cup\eta_s\bigg)\bigg)\;\bigg|\; f\bigg(\bigcup_{i=1}^k\{(x_i,\tau_i)\},\eta_s\bigg)=1\bigg]\bigg|.\end{aligned}$$

Since $|h(i)-h(j)| \le \|\Delta h\|_\infty |i-j|$ for $i, j \in \mathbb{N}_0$, we obtain, for each $\boldsymbol{x} = (x_1,\dots,x_k) \in \mathbb{X}^k$,

$$\begin{aligned}&\bigg|\mathbb{E}[h(W+1)] - \mathbb{E}\bigg[h\bigg(F\bigg(\bigcup_{i=1}^k\{(x_i,\tau_i)\}\cup\eta_s\bigg)\bigg)\;\bigg|\; f\bigg(\bigcup_{i=1}^k\{(x_i,\tau_i)\},\eta_s\bigg)=1\bigg]\bigg|\\ &\quad\le |\mathbb{E}[h(U_{\boldsymbol{x}}+1)] - \mathbb{E}[h(V_{\boldsymbol{x}}+1)]|\\ &\quad\le \|\Delta h\|_\infty \mathbb{E}[|U_{\boldsymbol{x}} - V_{\boldsymbol{x}}|]\\ &\quad\le \|\Delta h\|_\infty w(\boldsymbol{x})\end{aligned}$$

and, therefore,

$$|\mathbb{E}[\alpha h(W+1) - Wh(W)]| \le \frac{\|\Delta h\|_\infty}{k!}\int w(\boldsymbol{x})p(\boldsymbol{x})\lambda^k(\mathrm{d}\boldsymbol{x}).$$

Given $A \subset \mathbb{N}_0$, set $g = \mathbf{1}_A$ and choose $h\colon \mathbb{N}_0 \to \mathbb{R}$ so that $h(0) = 0$ and

$$\alpha h(i+1) - ih(i) = g(i) - \mathbb{E}[G(Z_\alpha)], \qquad i \in \mathbb{N}_0. \tag{3.6}$$

Then (see Lemma 1.1.1 of [1]) h is bounded and $\|\Delta h\|_\infty \le 1 \wedge \alpha^{-1}$, and, hence,

$$|\mathbb{P}[W \in A] - \mathbb{P}[Z_\alpha \in A]| \le \frac{(1\wedge\alpha^{-1})}{k!}\int w(\boldsymbol{x})p(\boldsymbol{x})\lambda^k(\mathrm{d}\boldsymbol{x}).$$

The result (3.3) follows.

We can obtain (3.4) similarly by choosing, for any given g with $\|\Delta g\|_\infty \le 1$, a solution h to (3.6) with $h(0) = 0$, and using Lemma 1.1.5 of [1]. □

We now provide an overview of how we will use Theorem 3.1 to prove Theorems 2.1 and 2.3. As explained in the next section, we may view $G(\mathcal{P}_s, \phi_s)$ as being determined by a marked Poisson point process in $\mathbb{X}$, that is, a Poisson point process in a product space $\boldsymbol{X} \times \mathbb{M}^*$ (where $\mathbb{M}^*$ is a certain mark space).

https://doi.org/10.1017/jpr.2018.9 Published online by Cambridge University Press

Then the functional $D_k(G(\mathcal{P}_s, \phi_s))$ may be viewed as a sum of the form (3.2) for a suitable f which selects those points with degree k. For each $x \in \mathbb{X}$, we need to find coupled variables U_x and V_x such that U_x has the distribution of $D_k(G(\mathcal{P}_s, \phi_s))$ and $1 + V_x$ has the conditional distribution of $D_k(G(\mathcal{P}_s \cup \{x\}, \phi_s))$ given that x has degree k, and such that $|U_x - V_x|$ is small (in probability). To do this we note that by the thinning theorem (see, for example, [15]), the point process of the points of $\mathcal{P}_s$ connected to x, and the point process of the points of $\mathcal{P}_s$ not connected to x, are independent. To generate V_x we need to condition the first point process to have exactly k points, which we can do by adding or removing points from it, while leaving the second point process unchanged. Since we modify only the first point process (that is, the points connected to x), we may hope that the score V_x obtained from the modified (overall) point process is similar to the score U_x obtained from the original (overall) point process. This is how we will prove Theorem 2.1.

For Theorem 2.3, we view $N_k(G(\mathcal{P}_s, \phi_s))$ as a sum of the form (3.2) for a suitable f which selects those k-tuples of points forming a component. For each $\boldsymbol{x} \subset \mathbb{X}$ with k elements, we need to find coupled variables $U_{\boldsymbol{x}}$ and $V_{\boldsymbol{x}}$ such that $U_{\boldsymbol{x}}$ has the distribution of $N_k(G(\mathcal{P}_s, \phi_s))$ and $1 + V_{\boldsymbol{x}}$ has the conditional distribution of $N_k(G(\mathcal{P}_s \cup \boldsymbol{x}, \phi_s))$ given that the points of $\boldsymbol{x}$ form a component. Again by the thinning theorem, the point process of the points of $\mathcal{P}_s$ connected to $\boldsymbol{x}$ is independent of the rest of $\mathcal{P}_s$, so to obtain $V_{\boldsymbol{x}}$ we condition on this point process having no elements, simply by removing those points. Again this is a small change, so the difference between $U_{\boldsymbol{x}}$ and $V_{\boldsymbol{x}}$ again is small (in probability).

4. Formal constructions of the IRG

Let $s \in (0, \infty)$ and $n \in \mathbb{N}$. Let $\phi \in \Phi$. We now provide a more formal definition of the graphs $G(\mathcal{P}_s, \phi)$ and $G(\mathcal{X}_n, \phi)$,

We make the following particular choice of mark space $(\mathbb{M}^*, \mathcal{M}^*, \boldsymbol{m}^*)$. Let 'Leb' denote Lebesgue measure on $[0, 1)$. Let $\mathbb{M}^* = [0, 1)^{\mathbb{N}_0}$ with $\mathcal{M}^*$ the product Borel σ-algebra, and with $\boldsymbol{m}^* := \otimes_{n=0}^{\infty}\text{Leb}$, so that a random element of $\mathbb{M}^*$ with distribution $\boldsymbol{m}^*$ is a sequence of independent uniform$(0, 1)$ random variables indexed by $\mathbb{N}_0$.

Now (and for the rest of this paper) taking $(\mathbb{M}, \mathcal{M}, \boldsymbol{m}) = (\mathbb{M}^*, \mathcal{M}^*, \boldsymbol{m}^*)$, let $(X_i, T_i)_{i\geq 1}$, $(Z_s)_{s>0}$, and $(\tau, \tau_1, \tau_2, \ldots)$ be as in the preceding section. Given $n \in \mathbb{N}$ and $s > 0$, let the point processes η_s and ξ_n be as given by (3.1). Thus, η_s is a Poisson point process in $\mathbb{X}$ with mean measure $s\mu$ and with each point marked with a sequence of independent uniform$[0, 1)$ variables indexed by the nonnegative integers. Similarly, ξ_n is a binomial point process in $\mathbb{X} \times \mathbb{M}^*$. We write $\boldsymbol{S}^*$ for $\boldsymbol{S}(\mathbb{X} \times \mathbb{M}^*)$.

To ease notation, we will also assume from now on that the probability measure μ on $\mathbb{X}$ is *diffuse*. This ensures that the values of $X_1, X_2, \ldots$ are almost surely distinct. However, this assumption is for notational convenience only; even without it, one can make sense of our results either by allowing the set $\{X_1, \ldots, X_n\}$ to have multiplicities or by using the attached marks T_i (which are almost surely distinct) to distinguish between different points X_i.

Set $\mathcal{X}_n := \{X_1, \ldots, X_n\}$ and $\mathcal{P}_s := \{X_1, \ldots, X_{Z_s}\}$, the canonical projections of ξ_n and η_s, respectively, onto $\mathbb{X}$. Define the graph $G(\mathcal{X}_n, \phi)$ to have vertex set $\mathcal{X}_n$ and to have an edge between vertices X_i and X_j, for $i, j \in [n] := \{1, 2, \ldots, n\}$ with $i < j$, if and only if $T_{i,j} \leq \phi(X_i, X_j)$, where $T_i = (T_{i,0}, T_{i,1}, T_{i,2} \ldots)$. Let $G(\mathcal{P}_s, \phi)$ be the graph $G(\mathcal{X}_{Z_s}, \phi)$.

This is one way to formally define the random graphs with the properties described more informally in the introduction. It has the advantage that $G(\mathcal{X}_n, \phi)$ is a subgraph of $G(\mathcal{P}_s, \phi)$ whenever $n \leq Z_s$ and $G(\mathcal{P}_s, \phi)$ is a subgraph of $G(\mathcal{X}_n, \phi)$ whenever $Z_s \leq n$, which is useful for coupling arguments. However, it has the disadvantage that the edge-set of the

https://doi.org/10.1017/jpr.2018.9 Published online by Cambridge University Press

graph $G(\mathcal{P}_s, \phi)$ is not invariant under permutation of the order in which the marked points $(X_1, T_1), \ldots, (X_{Z_s}, T_{Z_s})$ are listed. Therefore, we define a further graph which has the same distribution but also satisfies this permutation invariance. This will be useful in applying Theorem 3.1 in the proofs of Theorems 2.1, 2.3, and 2.4.

Given $\xi \in \boldsymbol{S}^*$ and $\phi \in \Phi$, define the graph $G_\phi(\xi)$ as follows. If there exist distinct $(x, t_0, t_1 \ldots) \in \xi$ and $(y, u_0, u_1, \ldots) \in \xi$ with $t_0 = u_0$, then set $G_\phi(\xi)$ to be the empty graph. Otherwise, we can write ξ uniquely as

$$\xi = \bigcup_{i=1}^{|\xi|} \{(x_i, t_{i,0}, t_{i,1}, t_{i,2}, \ldots)\}$$

with each $x_i \in \mathbb{X}$ and with $t_{1,0} < t_{2,0} < \cdots < t_{|\xi|,0}$. Let $G_\phi(\xi)$ have vertex set $\{x_1, \ldots, x_{|\xi|}\}$, and, for each $i < j \le |\xi|$, let $G_\phi(\xi)$ have an edge connecting x_i to x_j if and only if $t_{i,j} \le \phi(x_i, x_j)$. In other words, we use the first coordinate of the marks to determine the order in which we enumerate the points of ξ; having done so, for $i < j$, we use the $(j+1)$th component of the mark attached to the ith point to decide whether to connect it to the jth point.

We say that two random graphs G and G' have the same distribution if any graph invariant of G has the same distribution as the same graph invariant evaluated on G'. The following is immediate from the independence of the components of the marks τ_i.

Proposition 4.1. *For any finite $\mathcal{X} = \{x_1, \ldots, x_m\} \subset \mathbb{X}$, the distribution of the random graph $G_\phi(\{(x_1, \tau_1), \ldots, (x_m, \tau_m)\})$ is the same as that of $G(\mathcal{X}, \phi)$.*

In particular, the distribution of $G_\phi(\eta_s)$ is the same as that of $G(\mathcal{P}_s, \phi)$, although they are not the same graph because the set of edges is defined differently for the two graphs. Likewise, $G_\phi(\xi_n)$ has the same distribution as $G(\mathcal{X}_n, \phi)$.

In the following lemmas, we check measurability of functions which will feature in the proofs of Theorems 2.1 and 2.3, respectively. We use the following notation. For $\ell \in \mathbb{N}$ and $i \in [\ell]$, let $I_{\ell,i}$ denote the set of $(x, t_0, t_1, t_2, \ldots) \in \mathbb{X} \times [0,1)^{\mathbb{N}_0}$ such that $(i-1)/\ell \le t_0 < i/\ell$.

Lemma 4.1. *Let $j \in \mathbb{N}_0$. Then the function $f : \mathbb{X} \times \mathbb{M}^* \times \boldsymbol{S}^* \to \{0, 1\}$ given by*

$$f(x, \boldsymbol{t}, \xi) := \mathbf{1}_{\{x \text{ has degree } j \text{ in } G_\phi(\xi \cup \{(x, \boldsymbol{t})\})\}}, \qquad (x, \boldsymbol{t}, \xi) \in \mathbb{X} \times \mathbb{M}^* \times \boldsymbol{S}^*,$$

is measurable.

Proof. For $\ell, n, m, i_1 \ldots, i_n \in \mathbb{N}$ with $m \le n \le \ell$, and $i_1 < i_2 < \cdots < i_n \le \ell$, let $A_{\ell,n,m,i_1,\ldots,i_n}$ denote the set of $(x, \boldsymbol{t}, \xi) \in \mathbb{X} \times \mathbb{M}^* \times \boldsymbol{S}^*$ such that

- $(\xi \cup \{(x, \boldsymbol{t})\})(I_{\ell,i_k}) = 1$ for $1 \le k \le n$;
- $(\xi \cup \{(x, \boldsymbol{t})\})(I_{\ell,i}) = 0$ for $i \in [\ell] \setminus \{i_1, \ldots, i_n\}$;
- $(x, \boldsymbol{t}) \in I_{\ell,i_m}$;
- $(\sum_{k=1}^{m-1} \xi(\{(y, \boldsymbol{u}) \in I_{\ell,i_k} : u_m \le \phi(x, y)\})) + \sum_{k=m+1}^{n} \xi(\{(y, \boldsymbol{u}) \in I_{\ell,i_k} : t_k \le \phi(x, y)\}) = j$, where $\boldsymbol{u} = (u_0, u_1, \ldots)$ and $\boldsymbol{t} = (t_0, t_1, \ldots)$.

Then each $A_{\ell,n,m,i_1,\ldots,i_n}$ is measurable in $\mathbb{X} \times \mathbb{M}^* \times \boldsymbol{S}^*$, and

$$f(x, \boldsymbol{t}, \xi) = \mathbf{1}_{\{(x,\boldsymbol{t},\xi) \in \bigcup_{\ell,n,m,i_1,\ldots,i_n \in \mathbb{N}:\, m \le n \le \ell,\, 1 \le i_1 < i_2 < \cdots < i_n \le \ell} A_{\ell,n,m,i_1,\ldots,i_n}\}},$$

which is a measurable function. □

https://doi.org/10.1017/jpr.2018.9 Published online by Cambridge University Press

Lemma 4.2. *Let $k \in \mathbb{N}$. Suppose that the function $\tilde{f}\colon (\mathbb{X} \times \mathbb{M}^*)^k \times \boldsymbol{S}^* \to \{0,1\}$ is given, for $(x_1, \boldsymbol{t}_1, \dots, x_k, \boldsymbol{t}_k) \in (\mathbb{X} \times \mathbb{M}^*)^k$ and $\xi \in \boldsymbol{S}^*$, by*

$$\tilde{f}(x_1, \boldsymbol{t}_1, \dots, x_k, \boldsymbol{t}_k, \xi) := \mathbf{1}_{\{\{x_1,\dots,x_k\} \text{ induces a component of } G_\phi(\bigcup_{i=1}^k \{(x_i,\boldsymbol{t}_i)\} \cup \xi)\}}.$$

Then $\tilde{f}$ is measurable.

Proof. For $\ell, n, m_1, \dots, m_k, i_1 \dots, i_n \in \mathbb{N}$, with $\max_{1\le i\le k} m_i \le n \le \ell$, $m_1, \dots, m_k$ distinct, and $1 \le i_1 < i_2 < \cdots < i_n \le \ell$, let $A_{\ell,n,m_1,\dots,m_k,i_1,\dots,i_n}$ denote the set of $(x_1, \boldsymbol{t}_1, \dots, x_k, \boldsymbol{t}_k, \xi) \in (\mathbb{X} \times \mathbb{M}^*)^k \times \boldsymbol{S}^*$ such that, for some connected graph Γ on vertex set $[k]$,

- $(\bigcup_{h=1}^k \{(x_h, \boldsymbol{t}_h)\} \cup \xi)(I_{\ell,i_j}) = 1$ for $1 \le j \le n$;
- $(\bigcup_{h=1}^k \{(x_h, \boldsymbol{t}_h)\} \cup \xi)(I_{\ell,i}) = 0$ for $i \in [\ell] \setminus \{i_1, \dots, i_n\}$;
- $(x_h, \boldsymbol{t}_h) \in I_{\ell,i_{m_h}}$ for $1 \le h \le k$;
- for all $h \in [k]$ and $1 \le j < m_h$ with $j \notin \{m_1, \dots, m_k\}$, we have $\xi(\{(y, \boldsymbol{u}) \in I_{i_j} : u_{m_h} > \phi(x_h, y)\}) = 1$, where $\boldsymbol{u} = (u_0, u_1, \dots)$;
- for all $h \in [k]$ and $m_h < j \le n$ with $j \notin \{m_1, \dots, m_k\}$, we have $\xi(\{(y, \boldsymbol{u}) \in I_{i_j} : t_{h,j} > \phi(x_h, y)\}) = 1$, where $\boldsymbol{t}_h = (t_{h,0}, t_{h,1}, \dots)$, and
- for all $h, h' \in [k]$ with $m_h < m_{h'}$ and $\{h, h'\}$ an edge of Γ, we have $t_{h,m_h} \le \phi(x_h, x_{h'})$.

Then each $A_{\ell,n,k,m_1,\dots,m_k,i_1,\dots,i_n}$ is measurable in $(\mathbb{X} \times \mathbb{M})^k \times \boldsymbol{S}^*$, and setting $[n]^k_{\ne}$ to be the set of $(m_1, \dots, m_k) \in [n]^k$ such that $m_1, \dots, m_k$ are distinct, it follows that $\tilde{f}$ is the indicator of the set

$$\bigcup_{\ell,n,i_1,\dots,i_n \in \mathbb{N},\, (m_1,\dots,m_k) \in [n]^k_{\ne} :\, n \le \ell, 1 \le i_1 < i_2 < \cdots < i_n \le \ell} A_{\ell,n,m_1,\dots,m_k,i_1,\dots,i_n},$$

which is measurable. □

5. Proof of Theorem 2.1

For $s > 0$, let η_s, $\mathcal{P}_s$, Z_s, τ, τ_1, τ_2, ... be as in the preceding section. For $\phi \in \Phi$ and $i \in \mathbb{N}_0$, we can obtain, from the Mecke equation (3.5),

$$\mathbb{E}[D_i(G_\phi(\eta_s))] = s \int_{\mathbb{X}} \frac{(s \int \phi(x,y)\mu(\mathrm{d}y))^i}{i!} \exp\left(-s \int \phi(x,y)\mu(\mathrm{d}y)\right) \mu(\mathrm{d}x). \tag{5.1}$$

Here is a brief explanation of (5.1). In (3.5), we take $k = 1$ and use the space $\boldsymbol{S}^*$ rather than $\boldsymbol{S}$. For $(x, \boldsymbol{t}, \xi) \in \mathbb{X} \times \mathbb{M}^* \times \boldsymbol{S}^*$, our function $g((x, \boldsymbol{t}), \xi)$ takes the value 1 if x has degree i in the graph $G_\phi((x, \boldsymbol{t}) \cup \xi)$, and otherwise takes the value 0. Given $Z_s = n$, the point process $\mathcal{P}_s$ has n points, each of which is independently connected to x with probability $\int \phi(x, y)\mu(y)\mathrm{d}y$. Hence, by the thinning property of the Poisson distribution [15, Proposition 1.3], the number of points of $\mathcal{P}_s$ connected to x in $G_\phi((x, \tau) \cup \eta_s)$ is Poisson with mean $s \int \phi(x, y)\mu(\mathrm{d}y)$.

In the sequel, other formulae for expectations of numbers of vertices, or k-tuples of vertices, having certain properties in terms of the graph $G_\phi(\eta_s)$ (or, equivalently, the graph $G(\mathcal{P}_s, \phi)$), will also be justified by the Mecke formula. These arguments also justify Equation (3.5) of [22].

Recall from (2.6) that, for $\phi \in \Phi$, we set $\kappa(\phi) := \sup_{x \in \mathbb{X}} \int \phi(x, y)\mu(\mathrm{d}y)$.

https://doi.org/10.1017/jpr.2018.9 Published online by Cambridge University Press

Lemma 5.1. *Suppose the assumptions of Theorem 2.1 hold. Then (2.3) holds. Also,*

$$s\kappa(\phi_s) = \Theta(\log s), \tag{5.2}$$

$$\mathbb{E}[D_j(G(\mathcal{P}_s, \phi_s))] \to \alpha. \tag{5.3}$$

Proof. For all $s > 0$, since we assume that $\phi_s \in \Phi_\varepsilon$, by (5.1), we have, for $i \in \mathbb{N}$,

$$\frac{\varepsilon s\kappa(\phi_s)}{i} \le \frac{\mathbb{E}[D_i(G(\mathcal{P}_s, \phi_s))]}{\mathbb{E}[D_{i-1}(G(\mathcal{P}_s, \phi_s))]} \le \frac{s\kappa(\phi_s)}{i}. \tag{5.4}$$

We are assuming, for some fixed $j \in \mathbb{N}_0$, that as $s \to \infty$, we have

$$\mathbb{E}[D_{\le j}(G(\mathcal{P}_s, \phi_s))] \to \alpha \in (0, \infty). \tag{5.5}$$

Hence, $\mathbb{E}[D_0(G(\mathcal{P}_s, \phi_s))]$ is bounded, but also $\mathbb{E}[D_0(G(\mathcal{P}_s, \phi_s))] \ge s\mathrm{e}^{-s\kappa(\phi_s)}$ by (5.1), so $s\kappa(\phi_s) \to \infty$ as $s \to \infty$.

By (5.5), $\mathbb{E}[D_j(G(\mathcal{P}_s, \phi_s))]$ remains bounded, and since also $s\kappa(\phi_s) \to \infty$, if $j \ge 1$ then using (5.4) we have $\mathbb{E}[D_{j-1}(G(\mathcal{P}_s, \phi_s))] \to 0$, and repeating the argument we also have $\mathbb{E}[D_i(G(\mathcal{P}_s, \phi_s))] \to 0$ for $i = 0, 1, \ldots, j-1$. Hence, by (5.5), we have (5.3).

Using (5.3) and (5.1), the assumed ε-homogeneity of the ϕ_s, and the fact that $s\kappa(\phi_s) \to \infty$, it is straightforward to show that (5.2) holds. Then, using (5.4) repeatedly, and (5.3), we obtain (2.3). □

Proof of Theorem 2.1. Assume the assumptions of Theorem 2.1 apply. Then (2.3) follows from Lemma 5.1, and (2.3) gives us the first part of (2.2).

Next we show that $D_j(G(\mathcal{P}_s, \phi_s)) \xrightarrow{\mathrm{D}} Z_\alpha$. To carry out the strategy outlined at the end of Section 3, we will apply Theorem 3.1 to the $\phi = \phi_s$ case of the function f considered in Lemma 4.1. That is, for $(x, \boldsymbol{t}, \xi) \in \mathbb{X} \times \mathbb{M}^* \times \boldsymbol{S}^*$, we set $f_s(x, \boldsymbol{t}, \xi)$ to be the indicator of the statement that x has degree j in $G_{\phi_s}(\xi \cup \{(x, \boldsymbol{t})\})$. Then $D_j(G_{\phi_s}(\eta_s)) = F_s(\eta_s)$, where F_s is the function F obtained by using $f \equiv f_s$ in (3.2) (with $k = 1$).

Let $s > 0$ and $x \in \mathbb{X}$. If $j \ge 1$, suppose we also have an extra sequence $(Y, Y_1, \ldots, Y_j)$ of i.i.d. random elements of $\mathbb{X}$ with

$$\mathbb{P}[Y \in \mathrm{d}y] = \frac{\phi_s(x, y)\mu(\mathrm{d}y)}{\int \phi_s(x, z)\mu(\mathrm{d}z)},$$

independent of $((X_i, T_i))_{i\ge1}$, $(Z_s)_{s>0}$, and $(\tau, \tau_1, \tau_2, \ldots)$. Let $\mathcal{G}_s$ be the graph $G_{\phi_s}(\eta_s \cup \{(x, \tau), (Y_1, \tau_1), \ldots, (Y_j, \tau_j)\})$, with added edges from x to each of $Y_1, \ldots, Y_j$ (if not already included). Let $\mathcal{P}_{s,x}$ be the set of points of $\mathcal{P}_s$ that are connected to x in this graph, and set $\mathcal{P}_s^x := \mathcal{P}_s \setminus \mathcal{P}_{s,x}$.

Let U_x denote the number of vertices of degree j in the subgraph of $\mathcal{G}_s$ induced by vertex set $\mathcal{P}_s$. By Proposition 4.1, this graph has the distribution of $G(\mathcal{P}_s, \phi_s)$, so U_x has the distribution of $F_s(\eta_s)$.

Now consider the subgraph of $\mathcal{G}_s$ induced by $\{x\} \cup \mathcal{P}_s$. This has the distribution of $G(\mathcal{P}_s \cup \{x\}, \phi_s)$, and $\mathcal{P}_{s,x}$ is the set of vertices in this graph lying adjacent to x. Conditioning on x having degree j amounts to conditioning on $|\mathcal{P}_{s,x}| = j$. We define a coupled point process $\mathcal{P}^*$ (a subset of $\mathcal{P}_s \cup \{Y_1, \ldots, Y_j\}$) with the distribution of $\mathcal{P}_s$ conditioned on x having degree j, as follows.

If $|\mathcal{P}_{s,x}| > j$ then we select $|\mathcal{P}_{s,x}| - j$ elements of $\mathcal{P}_{s,x}$ uniformly at random and discard them from $\mathcal{P}_s$ to obtain a point process $\mathcal{P}^*$. If $|\mathcal{P}_{s,x}| < j$, we set $\mathcal{P}^* := \mathcal{P}_s \cup \mathcal{Y}_x$, where we set

$$\mathcal{Y}_x := \{Y_1, \ldots, Y_{j-|\mathcal{P}_{s,x}|}\}.$$

https://doi.org/10.1017/jpr.2018.9 Published online by Cambridge University Press

Let $\mathcal{G}_s^*$ denote the subgraph of $\mathcal{G}_s$ induced by $\mathcal{P}^* \cup \{x\}$. Then x has degree j in $\mathcal{G}_s^*$. Let V_x be the number of vertices in $\mathcal{G}_s^*$ having degree j, other than x. This has the conditional distribution of $F_s(\{(x,\tau)\} \cup \eta_s) - 1$ given that $f_s(\{(x,\tau)\}, \eta_s) = 1$. This is because $\mathcal{P}_{s,x}$ and $\mathcal{P}_s^x$ are independent Poisson processes, and conditioning on $f_s(\{(x,\tau)\}, \eta_s) = 1$ amounts to conditioning on the first of these two Poisson processes having j points.

If $|\mathcal{P}_{s,x}| > j$ then $|U_x - V_x| \le U_x' + V_x'$, where we set U_x' to be the number of $y \in \mathcal{P}_{s,x}$ such that y has j neighbours in $\mathcal{P}_s$, and V_x' to be the number of pairs (y,z) with $y \in \mathcal{P}_{s,x}$, $z \in \mathcal{P}_s$, such that $z \neq y$, z is connected to y and z has at most j neighbours in $\mathcal{P}_s^x$. By the Mecke formula, and the assumption that $\phi_s \in \Phi_\varepsilon$ for all s, writing $\kappa_s := \kappa(\phi_s)$ and recalling that $s\kappa_s \to \infty$ by (5.2), we have

$$\begin{aligned}\mathbb{E}[U_x'] &= \int s\phi_s(x,y)\left(\frac{(s\int \phi_s(y,z)\mu(\mathrm{d}z))^j}{j!}\right)\exp\left(-\int s\phi_s(y,z)\mu(\mathrm{d}z)\right)\mu(\mathrm{d}y)\\ &= O((s\kappa_s)^{j+1})\exp(-\Theta(s\kappa_s))\\ &= o(1),\end{aligned} \tag{5.6}$$

uniformly over $x \in \mathbb{X}$. Also, using the fact that $\overline{\phi}_s \le 1 - \varepsilon$ for all s, we have

$$\begin{aligned}\mathbb{E}[V_x'] &= \sum_{i=0}^{j}\int s\phi_s(x,y)\int s\phi_s(y,z)\left(\frac{(s\int \phi_s(z,w)(1-\phi_s(x,w))\mu(\mathrm{d}w))^i}{i!}\right)\\ &\qquad\qquad \times \exp\left(-\int s\phi_s(z,w)(1-\phi_s(x,w))\mu(\mathrm{d}w)\right)\mu(\mathrm{d}z)\mu(\mathrm{d}y)\\ &= O((s\kappa_s)^{j+2})\exp(-\Theta(s\kappa_s))\\ &\to 0,\end{aligned} \tag{5.7}$$

uniformly over $x \in \mathbb{X}$.

If $|\mathcal{P}_{s,x}| < j$ then $|U_x - V_x| \le \tilde{U}_x + \tilde{V}_x$, where we set $\tilde{U}_x$ to be the number of $y \in \mathcal{Y}_x$ having at most j neighbours in $\mathcal{P}_s$, and $\tilde{V}_x$ is the number of pairs (y,z) with $y \in \mathcal{Y}_x$, $z \in \mathcal{P}_s$ connected to y, and z having at most j neighbours in $\mathcal{P}_s$. Then

$$\mathbb{E}[\tilde{U}_x] \le j\sum_{i=0}^{j}\int (s\kappa_s)^i\exp(-\Theta(s\kappa_s))\left(\frac{\phi_s(x,y)}{\int \phi_s(x,w)\mu(\mathrm{d}w)}\right)\mu(\mathrm{d}y) \to 0, \tag{5.8}$$

uniformly over $x \in \mathbb{X}$. Also,

$$\begin{aligned}\mathbb{E}[\tilde{V}_x] &\le \sum_{i=0}^{j}\int s\left(s\int \phi_s(z,w)\mu(\mathrm{d}w)\right)^i\exp\left(-\int s\phi_s(z,w)\mu(\mathrm{d}w)\right)\\ &\qquad\qquad \times j\int\left(\frac{\phi_s(x,y)\phi_s(z,y)}{\int \phi_s(x,w)\mu(\mathrm{d}w)}\right)\mu(\mathrm{d}y)\mu(\mathrm{d}z)\\ &= O((s\kappa_s)^{j+1})\exp(-\Theta(s\kappa_s))\\ &\to 0,\end{aligned} \tag{5.9}$$

uniformly over $x \in \mathbb{X}$.

Combining the estimates (5.6)–(5.9), and using Theorem 3.1, gives us the first part of (2.1), namely, $D_j(G(\mathcal{P}_s, \phi_s)) \xrightarrow{\mathrm{D}} Z_\alpha$.

https://doi.org/10.1017/jpr.2018.9 Published online by Cambridge University Press

Now suppose that $k > j$. Set $\beta_{k,s} := \mathbb{E}[D_k(G(\mathcal{P}_s, \phi_s))]$. The argument just given, with j replaced by k, shows that $d_{\mathrm{TV}}(D_k(G(\mathcal{P}_s, \phi_s)), Z_{\beta_{k,s}}) \to 0$ as $s \to \infty$. Also, $\beta_{k,s} \to \infty$ by (2.3), so that $(Z_{\beta_{k,s}} - \beta_{k,s})/\sqrt{\beta_{k,s}} \xrightarrow{\mathrm{D}} \mathcal{N}$. Hence, $(D_k(G(\mathcal{P}_s, \phi_s)) - \beta_{k,s})/\sqrt{\beta_{k,s}} \xrightarrow{\mathrm{D}} \mathcal{N}$, which is (2.4).

It remains to prove the second parts of (2.1) and (2.2). For $n \in \mathbb{N}$, let $s(n) = n - n^{3/4}$ and $t(n) := n + n^{3/4}$. By Chebyshev's inequality, with high probability (that is, with probability tending to 1), we have $Z_{s(n)} \le n \le Z_{t(n)}$ so that $\mathcal{P}_{s(n)} \subset \mathcal{X}_n \subset \mathcal{P}_{t(n)}$. Moreover, when this happens, $G(\mathcal{X}_n, \phi)$ is the subgraph of $G(\mathcal{P}_{t(n)}, \phi)$ induced by $\mathcal{X}_n$, and $G(\mathcal{P}_{s(n)}, \phi)$ is the subgraph of $G(\mathcal{X}_n, \phi)$ induced by $\mathcal{P}_{s(n)}$.

By (5.1), for $i \in \mathbb{N}_0$,

$$\left(\frac{s(n)}{n}\right)^{i+1} \le \frac{\mathbb{E}[D_i(G(\mathcal{P}_{s(n)}, \phi_n))]}{\mathbb{E}[D_i(G(\mathcal{P}_n, \phi_n))]} \le \sup_{x \in \mathbb{X}} \left(\exp\left(n^{3/4} \int \phi_n(x, y) \mu(\mathrm{d}y) \right) \right)$$

and, by (5.2), both the upper and the lower bound tend to 1. Therefore, by (2.3) and (5.3), we have, as $n \to \infty$,

$$\mathbb{E}[D_i(G(\mathcal{P}_{s(n)}, \phi_n))] \to 0, \quad i = 0, 1, \ldots, j-1, \qquad \mathbb{E}[D_j(G(\mathcal{P}_{s(n)}, \phi_n))] \to \alpha. \quad (5.10)$$

Let $k \in \mathbb{N}_0$ with $k \le j$. If $\mathcal{P}_{s(n)} \subset \mathcal{X}_n$ then

$$D_{\le k}(G(\mathcal{X}_n, \phi_n)) - D_{\le k}(G(\mathcal{P}_{s(n)}, \phi_n)) = S_n - R_n,$$

where S_n denotes the number of points of $\mathcal{X}_n \setminus \mathcal{P}_{s(n)}$ with degree at most k in $G(\mathcal{X}_n, \phi_n)$, and R_n is the number of points of $\mathcal{P}_{s(n)}$ with degree at most k in $G(\mathcal{P}_{s(n)}, \phi_n)$ but with degree at least $k+1$ in $G(\mathcal{X}_n, \phi_n)$.

Let S_n' denote the number of points of $\mathcal{P}_{t(n)} \setminus \mathcal{P}_{s(n)}$ that are connected to at most k points of $\mathcal{P}_{s(n)}$, and let R_n' be the number of points of $\mathcal{P}_{s(n)}$ with degree at most k in $G(\mathcal{P}_{s(n)}, \phi_n)$ but with degree at least $k+1$ in $G(\mathcal{P}_{t(n)}, \phi_n)$. If $\mathcal{P}_{s(n)} \subset \mathcal{X}_n \subset \mathcal{P}_{t(n)}$ then $S_n \le S_n'$ and $R_n \le R_n'$. Hence, by Markov's inequality,

$$\mathbb{P}[D_{\le k}(G(\mathcal{X}_n, \phi_n)) \ne D_{\le k}(G(\mathcal{P}_{s(n)}, \phi_n))] \le \mathbb{P}[\{Z_{s(n)} \le n \le Z_{t(n)}\}^c] + \mathbb{E}[S_n'] + \mathbb{E}[R_n'].$$

Now

$$\begin{aligned}\mathbb{E}[S_n'] &= 2n^{3/4} \sum_{i=0}^{k} \int_{\mathbb{X}} \frac{(s(n) \int \phi_n(x, y) \mu(\mathrm{d}y))^i}{i!} \exp\left(-s(n) \int \phi_n(x, y)\, \mathrm{d}y \right) \mu(\mathrm{d}x) \\ &= \left(\frac{2n^{3/4}}{s(n)} \right) \mathbb{E}[D_{\le k}(G(\mathcal{P}_{s(n)}, \phi_n))],\end{aligned}$$

which tends to 0 by (5.10). Also,

$$\begin{aligned}\mathbb{E}[R_n'] &\le s(n) \sum_{i=0}^{k} \int \frac{(s(n) \int \phi_n(x, y) \mu(\mathrm{d}y))^i}{i!} \\ &\qquad \times \exp\left(-s(n) \int \phi_n(x, y) \mu(\mathrm{d}y) \right) \left\{ 2n^{3/4} \int \phi_n(x, y) \mu(\mathrm{d}y) \right\} \mu(\mathrm{d}x) \\ &\le 2n^{3/4} a_n \mathbb{E}[D_{\le k}(\mathcal{P}_{s(n)}, \phi_n)],\end{aligned}$$

which tends to 0 by (5.2) and (5.10).

https://doi.org/10.1017/jpr.2018.9 Published online by Cambridge University Press

Therefore, with high probability, we have $D_{\le k}(G(\mathcal{X}_n, \phi_n)) = D_{\le k}(G(\mathcal{P}_{s(n)}, \phi_n))$. This holds both for $k = j$ and for $k = j - 1$. Hence, using the first parts of (2.1) and (2.2), we obtain the second parts of (2.1) and (2.2). □

6. Proofs of Theorems 2.2 and 2.3

Given $\phi \in \Phi$, $k, \ell \in \mathbb{N}$, and $\boldsymbol{x} = (x_1 \dots, x_k) \in \mathbb{X}^k$ and $\boldsymbol{y} = (y_1, \dots, y_\ell) \in \mathbb{X}^\ell$, set

$$\phi(\boldsymbol{x}, \boldsymbol{y}) := 1 - \prod_{i=1}^{k} \prod_{j=1}^{\ell} (1 - \phi(x_i, y_j)). \tag{6.1}$$

We also write $\phi(\{x_1, \dots, x_k\}, \{y_1, \dots, y_\ell\})$ for $\phi(\boldsymbol{x}, \boldsymbol{y})$ (allowing multiplicities in the sets $\{x_1, \dots, x_k\}$ and $\{y_1, \dots, y_\ell\}$); it is the probability that there is at least one edge in the random graph $G(\{x_1, \dots, x_k, y_1, \dots, y_\ell\}, \phi)$ connecting one of the vertices x_i to one of the vertices y_j. If $k = 1$, we write $\phi(x_1, \{y_1, \dots, y_\ell\})$ for $\phi(\{x_1\}, \{y_1, \dots, y_\ell\})$. Also, let $h_\phi(\boldsymbol{x})$ or $h_\phi(x_1, \dots, x_k)$ denote the probability that $G(\{x_1, \dots, x_k\}; \phi)$ is connected; more precisely, let $h_\phi(\boldsymbol{x}) := 1$ if $k = 1$ and, otherwise, let

$$h_\phi(x_1, \dots, x_k) := \sum_{\Gamma} \prod_{\{\{i,j\}:\ \{i,j\} \in E(\Gamma)\}} \phi(x_i, x_j) \prod_{\{\{i,j\}:\ i,j \notin E(\Gamma)\}} (1 - \phi(x_i, x_j)), \tag{6.2}$$

where the sum is over all connected graphs Γ on vertex set $\{1, \dots, k\}$, and $E(G)$ denotes the set of edges of a graph G. By the Mecke formula (3.5), and the equality in distribution of $G(\mathcal{P}_s, \phi)$ and $G_\phi(\eta_s)$ as discussed in Section 4,

$$\begin{aligned} \mathbb{E}[N_k(G(\mathcal{P}_s, \phi))] &= \mathbb{E}[N_k(G_\phi(\eta_s))] \\ &= \frac{s^k}{k!} \int_{\mathbb{X}^k} h_\phi(x_1, \dots, x_k) \exp\left(-s \int \phi(z;\ \{x_1, \dots, x_k\}) \mu(\mathrm{d}z)\right) \\ &\qquad \times \mu^k(\mathrm{d}(x_1, \dots, x_k)). \end{aligned} \tag{6.3}$$

Now fix $k \in \mathbb{N}$ and $\alpha \in (0, \infty)$. Assume throughout this section that $\phi_s \in \Phi$ for $s > 0$, and (unless explicitly stated otherwise) that

$$\lim_{s \to \infty} \mathbb{E}[N_k(G(\mathcal{P}_s, \phi_s))] = \alpha. \tag{6.4}$$

Proof of Theorem 2.2. (i) Assume that $\overline{\phi}_s = o(1/\log s)$. We will use the method of moments.

For $n, \ell \in \mathbb{N}$, we write $(n)_\ell$ for the descending factorial $n(n-1)\cdots(n-\ell+1)$. Then $(N_k(G(\mathcal{P}_s, \phi_s)))_\ell$ is the number of ordered ℓ-tuples of distinct k-components of $G(\mathcal{P}_s, \phi_s)$. This equals the sum over all ordered $k\ell$-tuples $x_{1,1}, \dots, x_{1,k}, \dots, x_{\ell,1}, \dots, x_{\ell,k}$ of distinct points of $\mathcal{P}_s$, of the indicator of the event that for each $i \le \ell$ the subgraph of $G(\mathcal{P}_s, \phi_s)$ induced by $x_{i,1} \dots, x_{i,k}$ is connected and these vertices are not connected to any other vertices of $G(\mathcal{P}_s, \phi_s)$, divided by $(k!)^\ell$. Hence, by the Mecke formula,

$$\begin{aligned} &\mathbb{E}[[(N_k(G(\mathcal{P}_s, \phi_s)))_\ell] \\ &\quad = \frac{s^{k\ell}}{(k!)^\ell} \int_{\mathbb{X}^{k\ell}} \left(\prod_{i=1}^{\ell} h_{\phi_s}(x_{i,1}, \dots, x_{i,k})\right) u_s(x_{1,1} \dots, x_{\ell,k}) \\ &\qquad\qquad \times \exp\left(-s \int \phi_s(z, \{x_{1,1}, \dots, x_{\ell,k}\}) \mu(\mathrm{d}z)\right) \mu^{k\ell}(\mathrm{d}(x_{1,1}, \dots, x_{\ell,k})), \end{aligned} \tag{6.5}$$

https://doi.org/10.1017/jpr.2018.9 Published online by Cambridge University Press

where we set $u_s(\{x_{1,1}, \dots, x_{\ell,k}\})$ to be the probability that the graph $G(\{x_{1,1}, \dots, x_{\ell,k}\}, \phi_s)$ has no edge between any $x_{i,j}$ and $x_{i',j'}$ such that $i \neq i'$, that is,

$$u_s(x_{1,1} \dots, x_{\ell,k}) := \prod_{(i_1,j_1),(i_2,j_2)\in[\ell]\times[k]:\, i_1<i_2} (1 - \phi_s(x_{i_1,j_1}, x_{i_2,j_2})). \tag{6.6}$$

Applying our condition on $\overline{\phi}_s$, we see that the value of $u_s(x_{1,1}, \dots, x_{\ell,k})$ tends to 1, uniformly over $(x_{1,1}, \dots, x_{\ell,k})$. Also, by the union bound, $\phi_s(z, \{x_{1,1}, \dots, x_{\ell,k}\})$ is bounded by $\sum_{i=1}^{\ell} \phi_s(z, \{x_{i,1}, \dots, x_{i,k}\})$. Therefore, by (6.5), writing just N_k for $N_k(G(\mathcal{P}_s, \phi_s))$, we have

$$\mathbb{E}[(N_k)_\ell] \geq (1 + o(1))\bigg(\frac{s^k}{k!} \int_{\mathbb{X}^k} h_{\phi_s}(x_1, \dots, x_k) \exp\bigg(-s \int \phi_s(z, \{x_1, \dots, x_k\})\mu(\mathrm{d}z)\bigg) \times \mu^k(\mathrm{d}(x_1, \dots, x_k))\bigg)^\ell.$$

By (6.4) and (6.3), this lower bound for $\mathbb{E}[(N_k)_\ell]$ tends to α^ℓ as $s \to \infty$.

By the Bonferroni bound and the union bound, we have, for $(z, x_{1,1}, \dots, x_{\ell,k}) \in \mathbb{X}^{1+\ell k}$,

$$\begin{aligned}\phi_s(z, \{x_{1,1}, \dots, x_{\ell,k}\}) &\geq \sum_{i=1}^{\ell} \phi_s(z, \{x_{i,1}, \dots, x_{i,k}\})\bigg(1 - \sum_{j=i+1}^{\ell} \phi_s(z, \{x_{j,1}, \dots, x_{j,k}\})\bigg) \\ &\geq \sum_{i=1}^{\ell} \phi_s(z, \{x_{i,1}, \dots, x_{i,k}\})(1 - k\ell\overline{\phi}_s). \end{aligned} \tag{6.7}$$

Therefore, by (6.5),

$$\begin{aligned}\mathbb{E}[(N_k)_\ell] &\leq \frac{s^{k\ell}}{(k!)^\ell} \int_{\mathbb{X}^{k\ell}} \bigg(\prod_{i=1}^{\ell} h_{\phi_s}(x_{i,1}, \dots, x_{i,k})\bigg) \\ &\qquad \times \exp\bigg(-s \sum_{i=1}^{\ell} \int \phi_s(z, \{x_{i,1}, \dots, x_{i,k}\})(1 - k\ell\overline{\phi}_s)\mu(\mathrm{d}z)\bigg) \\ &\qquad \times \mu^{k\ell}(\mathrm{d}(x_{1,1}, \dots, x_{\ell,k})) \\ &= \bigg(\frac{s^k}{k!} \int_{\mathbb{X}^k} h_{\phi_s}(x_1, \dots, x_k) \exp\bigg(-s \int \phi_s(z, \{x_1, \dots, x_k\})(1 - k\ell\overline{\phi}_s)\mu(\mathrm{d}z)\bigg) \\ &\qquad \times \mu^k(\mathrm{d}(x_1, \dots, x_k))\bigg)^\ell. \end{aligned}$$

Since μ is a probability measure and $x^{1-k\ell\overline{\phi}_s}$ is a concave function on $x \geq 0$, we obtain, by Jensen's inequality and the fact that $h_{\phi_s}(\cdot)^{1/(1-k\ell\overline{\phi}_s)} \leq h_{\phi_s}(\cdot)$,

$$\mathbb{E}[(N_k)_\ell] \leq \frac{s^{k^2\ell^2\overline{\phi}_s}}{k!^{\ell^2 k\overline{\phi}_s}} \bigg(\frac{s^k}{k!} \int h_{\phi_s}(x_1, \dots, x_k) \exp\bigg(-s \int \phi_s(z, \{x_1, \dots, x_k\})\mu(\mathrm{d}z)\bigg) \times \mu^k(\mathrm{d}(x_1, \dots, x_k))\bigg)^{\ell(1-k\ell\overline{\phi}_s)},$$

which tends to α^ℓ by (6.3), (6.4), and our assumption on $\overline{\phi}_s$. Thus, $\mathbb{E}[(N_k)_\ell] \to \alpha^\ell$, so by the method of moments (see, for example, Theorem 1.22 of [3]), the result (i) follows.

https://doi.org/10.1017/jpr.2018.9 Published online by Cambridge University Press

(ii) Assume now that $\overline{\phi}_s = o(s^{-1/2})$. Again we use the method of moments. Write n for $\lfloor s \rfloor$, and set $N'_k = N_k(G(\mathcal{X}_n, \phi_n))$. Then, by (6.6) and the union bound,

$$\mathbb{E}[(N'_k)_\ell] = \frac{(n)_{k\ell}}{k!^\ell} \int \cdots \int \Big(\prod_{i=1}^{\ell} h_{\phi_n}(x_{i,1}, \ldots, x_{i,k}) \Big) u_n(\{x_{1,1}, \ldots, x_{\ell,k}\}) \times \left(1 - \int \phi_n(z; \{x_{1,1}, \ldots, x_{\ell,k}\}) \mu(\mathrm{d}z)\right)^{n-k\ell} \times \mu(\mathrm{d}x_{1,1}) \cdots \mu(\mathrm{d}x_{\ell,k})$$

$$\geq (1 + o(1)) \left(\frac{n^k}{k!}\right)^\ell \int \cdots \int \Big(\prod_{i=1}^{\ell} h_{\phi_n}(x_{i,1}, \ldots, x_{i,k}) \Big) \times \left(1 - \sum_{i=1}^{\ell} \int \phi_n(z, \{x_{i,1}, \ldots, x_{i,k}\}) \mu(\mathrm{d}z)\right)^n \times \mu(\mathrm{d}x_{1,1}) \cdots \mu(\mathrm{d}x_{\ell,k}).$$

Using the bound $1 - x \geq \exp(-x - x^2)$ for small positive x, we have, for large n,

$$\mathbb{E}[(N'_k)_\ell] \geq (1 + o(1)) \left(\frac{n^k}{k!}\right)^\ell \int \cdots \int \Big(\prod_{i=1}^{\ell} h_{\phi_n}(x_{i,1}, \ldots, x_{i,k}) \Big) \times \exp\left(-n \left(\sum_{i=1}^{\ell} \int \phi_n(z, \{x_{i,1}, \ldots, x_{i,k}\}) \mu(\mathrm{d}z) \right) - n \left(\sum_{i=1}^{\ell} \sum_{j=1}^{k} \int \phi_n(z, x_{i,j}) \mu(\mathrm{d}z) \right)^2 \right) \times \mu(\mathrm{d}x_{1,1}) \cdots \mu(\mathrm{d}x_{\ell,k}),$$

so that

$$\mathbb{E}[(N'_k)_\ell] \geq (1 + o(1)) \left(\frac{n^k}{k!} \int h_{\phi_n}(x_1, \ldots, x_k) \exp\left(-n \int \phi_n(z, \{x_1, \ldots, x_k\}) \mu(\mathrm{d}z) \right) \times \mu^k(\mathrm{d}(x_1, \ldots, \mathrm{d}x_k)) \right)^\ell \exp(-nk^2\ell^2 \overline{\phi}_n^2),$$

which tends to α^ℓ by (6.3), (6.4), and the assumption that $\overline{\phi}_n = o(n^{-1/2})$.

Conversely, by the bound $\mathrm{e}^x \geq 1 + x$ for all $x \in \mathbb{R}$, we also have

$$\mathbb{E}[(N'_k)_\ell] \leq \left(\frac{n^k}{k!}\right)^\ell \int \cdots \int \Big(\prod_{i=1}^{\ell} h_{\phi_n}(x_{i,1}, \ldots, x_{i,k}) \Big) \times \exp\left(-(n - \ell k) \int \phi_n(z; \{x_{1,1}, \ldots, x_{\ell,k}\}) \mu(\mathrm{d}z) \right) \times \mu(\mathrm{d}x_{1,1}) \cdots \mu(\mathrm{d}x_{\ell,k})$$

https://doi.org/10.1017/jpr.2018.9 Published online by Cambridge University Press

and using the Bonferroni bound as in (6.7), we obtain

$$\begin{aligned}\mathbb{E}[(N_k')_\ell] \le \mathrm{e}^{k^2\ell^2\overline{\phi}_n}\left(\frac{n^k}{k!}\right)^\ell &\times \int\cdots\int\left(\prod_{i=1}^{\ell} h_{\phi_n}(x_{i,1},\dots,x_{i,k})\right)\\ &\times \exp\left(-n\sum_{i=1}^{\ell}\int \phi_n(z;\,\{x_{i,1},\dots,x_{i,k}\})(1-k\ell\overline{\phi}_n)\mu(\mathrm{d}z)\right)\\ &\times \mu(\mathrm{d}x_{1,1})\cdots\mu(\mathrm{d}x_{\ell,k}).\end{aligned}$$

By Jensen's inequality, since $x^{1-k\ell\overline{\phi}_n}$ is a concave function on $x \ge 0$ and $h_{\phi_n}(\cdot) \le 1$, we have

$$\begin{aligned}\mathbb{E}[(N_k')_\ell] \le \mathrm{e}^{k^2\ell^2\overline{\phi}_n} n^{k^2\ell^2\overline{\phi}_n}\left(\frac{1}{k!}\right)^\ell \Bigg(n^k\int\cdots\int h_{\phi_n}(x_1,\dots,x_k)&\\ \times \exp\left(-n\int \phi_n(z;\,\{x_1,\dots,x_k\})\mu(\mathrm{d}z)\right)&\\ \times \mu(\mathrm{d}x_1)\cdots\mu(\mathrm{d}x_k)\Bigg)^{\ell(1-k\ell\overline{\phi}_n)}&.\end{aligned}$$

Hence, since our assumption $\overline{\phi}_n = o(n^{-1/2})$ implies that $\overline{\phi}_n = o(1/(\log n))$, we have

$$\limsup \mathbb{E}[(N_k')_\ell] \le \alpha^\ell.$$

Therefore, $\mathbb{E}[(N_k')_\ell] \to \alpha^\ell$, and the method of moments gives us part (ii). □

Lemma 6.1. *Let $k \in \mathbb{N}$, $\varepsilon > 0$, and $\phi_s \in \Phi_\varepsilon$ for all $s > 0$. Set*

$$\kappa_s := \kappa(\phi_s) = \sup_{x\in\mathbb{X}}\int \phi_s(x,y)\mu(\mathrm{d}y).$$

Then

$$\mathbb{E}[N_k(G(\mathcal{P}_s,\phi_s))] = \Theta(s^k\kappa_s^{k-1})\exp(-\Theta(s\kappa_s)). \tag{6.8}$$

Proof. Our starting point is (6.3). We first bound the exponent in the exponential factor. Let $z, x_1, \dots, x_k \in \mathbb{X}$ and $s > 0$. Then, by the union bound,

$$\phi_s(z,x_1) \le \phi_s(z,\{x_1,\dots,x_k\}) \le \sum_{i=1}^{k}\phi_s(z,x_i),$$

so integrating over z and using the assumed ε-homogeneity, we have

$$\varepsilon\kappa_s \le \int_{\mathbb{X}} \phi_s(z,\{x_1,\dots,x_k\})\mu(\mathrm{d}z) \le k\kappa_s. \tag{6.9}$$

Using (6.3), this already proves the result for $k = 1$, so from now on assume that $k \ge 2$.

https://doi.org/10.1017/jpr.2018.9 Published online by Cambridge University Press

If x_{i-1} is connected to x_i for $2 \le i \le k$, then $G(\{x_1, \dots, x_k\}; \phi_s)$ is connected; hence, $h_{\phi_s}(x_1, \dots, x_k) \ge \prod_{i=2}^{k} \phi_s(x_{i-1}, x_i)$, so using (6.3) and (6.9), we obtain

$$\mathbb{E}[N_k(G(\mathcal{P}_s, \phi_s))] \ge \frac{s^k}{k!} \exp(-ks\kappa_s) \int_{\mathbb{X}} \cdots \int_{\mathbb{X}} \prod_{i=2}^{k} \phi_s(x_{i-1}, x_i) \mu(\mathrm{d}x_k) \cdots \mu(\mathrm{d}x_1)$$
$$\ge \frac{\varepsilon^{k-1} \kappa_s^{k-1} s^k}{k!} \exp(-ks\kappa_s). \tag{6.10}$$

For an upper bound, observe that if $G(\{x_1, \dots, x_k\}; \phi_s)$ is connected, then there exists a permutation σ of $\{1, \dots, k\}$ with $\sigma(1) = 1$ such that for $2 \le i \le k$ this graph has an edge from $x_{\sigma(i)}$ to $\{x_{\sigma(1)}, \dots, x_{\sigma(i-1)}\}$. Therefore, setting

$$h^*_{\phi_s}(x_1, \dots, x_k) := \prod_{i=2}^{k} \phi_s(x_i, \{x_1, \dots, x_{i-1}\}) \le \prod_{i=2}^{k} \sum_{j=1}^{i-1} \phi_s(x_i, x_j), \tag{6.11}$$

we have by the union bound that $h_{\phi_s}(x_1, \dots, x_k) \le \sum_\sigma h^*_{\phi_s}(x_{\sigma(1)}, \dots, x_{\sigma(k)})$, where the sum is over all such permutations. Hence, using (6.3) and (6.9), setting $y_j = x_{\sigma(j)}$, and taking the integrals in order $y_k, \dots, y_1$ for each permutation σ, we have

$$\mathbb{E}[N_k(G(\mathcal{P}_s, \phi_s))] \le \frac{s^k (k-1)!}{k!} \exp(-\varepsilon s\kappa_s) \int_{\mathbb{X}} \cdots \int_{\mathbb{X}} h^*_{\phi_s}(y_1, \dots, y_k) \mu(\mathrm{d}y_k) \cdots \mu(\mathrm{d}y_1),$$

and then using the inequality in (6.11), we have

$$\mathbb{E}[N_k(G(\mathcal{P}_s, \phi_s))] \le s^k (k-1)! \exp(-\varepsilon s\kappa_s) \kappa_s^{k-1}.$$

Combined with (6.10) this gives us (6.8). □

Proof of Theorem 2.3. Assume there exists $\varepsilon > 0$ such that $\phi_s \in \Phi_\varepsilon$ and $\overline{\phi}_s \le 1 - \varepsilon$ for all s. Set $\kappa_s := \kappa(\phi_s)$. If $s\kappa_s$ remains bounded away from 0 and ∞, then, by (6.8), we have $\mathbb{E}[N_k(G(\mathcal{P}_s, \phi_s))] \to \infty$, contradicting (6.4). Hence, for any sequence of values of s tending to ∞, there is a subsequence such that either $s\kappa_s \to 0$ or $s\kappa_s \to \infty$ as $s \to \infty$ along the subsequence.

Consider first the case with $s\kappa_s \to 0$. In this case, by (6.8) and (6.4), we have $k \ge 2$ and $s^k \kappa_s^{k-1} = \Theta(1)$. Recalling that $H_k(G)$ denotes the number of connected induced subgraphs of a graph G of order k, we have

$$\mathbb{E}[H_{k+1}(G(\mathcal{P}_s, \phi_s))] = O(s^{k+1} \kappa_s^k) = o(1).$$

Since $0 \le H_k(G(\mathcal{P}_s, \phi_s)) - N_k(G(\mathcal{P}_s, \phi_s)) \le (k+1) H_{k+1}(G(\mathcal{P}_s, \phi_s))$, we have

$$\mathbb{E}[H_k(G(\mathcal{P}_s, \phi_s)) - N_k(G(\mathcal{P}_s, \phi_s))] \to 0,$$

so by (6.4), we have $\mathbb{E}[H_k(G(\mathcal{P}_s, \phi_s))] \to \alpha$. Hence, by Theorem 2.4(i), it follows that $H_k(G(\mathcal{P}_s, \phi_s)) \xrightarrow{\mathrm{D}} Z_\alpha$. By Markov's inequality, $\mathbb{P}[H_k(G(\mathcal{P}_s, \phi_s)) - N_k(G(\mathcal{P}_s, \phi_s)) \ge 1] \to 0$, so we also have $N_k(G(\mathcal{P}_s, \phi_s)) \xrightarrow{\mathrm{D}} Z_\alpha$, which is the first part of (2.5).

Suppose now that $s\kappa_s \to \infty$. Then, by (6.8) and (6.4), one may readily deduce that

$$s\kappa_s = \Theta(\log s). \tag{6.12}$$

https://doi.org/10.1017/jpr.2018.9 Published online by Cambridge University Press

We seek to apply Theorem 3.1. For $s > 0$, let η_s, $\mathcal{P}_s$, Z_s, τ, τ_1, τ_2, ... be as in Section 4. Let $\tilde{f}_s$ be the function $\tilde{f}$ considered in Lemma 4.2, using the connection function $\phi \equiv \phi_s$. That is, let $\tilde{f}_s(x_1, \boldsymbol{t}_1, \ldots, x_k, \boldsymbol{t}_k, \xi)$ be the indicator of the statement that $\{x_1, \ldots, x_k\}$ induces a component of $G_{\phi_s}(\bigcup_{i=1}^k \{(x_i, \boldsymbol{t}_i)\} \cup \xi)$. Then with $\tilde{F}_s$ denoting the function F obtained by taking $f \equiv \tilde{f}_s$ in the definition (3.2), we have $\tilde{F}_s(\eta_s) = N_k(G_{\phi_s}(\eta_s))$, which has the same distribution as $N_k(G(\mathcal{P}_s, \phi_s))$.

For $\boldsymbol{x} = (x_1, \ldots, x_k) \in \mathbb{X}^k$, set $\mathcal{X} = \{x_1, \ldots, x_k\}$ (allowing multiplicities). Define the graph

$$\tilde{\mathcal{G}}_s := G_{\phi_s}(\eta_s \cup \{(x_1, \tau_1), \ldots, (x_k, \tau_k)\}). \tag{6.13}$$

Let $\mathcal{P}_{s,\boldsymbol{x}}$ be the set of points of $\mathcal{P}_s$ connected to at least one point of $\mathcal{X}$ in $\tilde{\mathcal{G}}_s$ and let $\mathcal{P}_s^{\boldsymbol{x}} = \mathcal{P}_s \setminus \mathcal{P}_{s,\boldsymbol{x}}$.

The subgraph of $\tilde{\mathcal{G}}_s$ induced by vertex set $\mathcal{P}_s$ has the same distribution as $G(\mathcal{P}_s, \phi_s)$, and we refer to this subgraph as $G'(\mathcal{P}_s, \phi_s)$. Likewise, we refer to the subgraph of $\tilde{\mathcal{G}}_s$ induced by vertex set $\mathcal{P}_s^{\boldsymbol{x}}$ as $G'(\mathcal{P}_s^{\boldsymbol{x}}, \phi_s)$, and we refer to the subgraph of $\tilde{\mathcal{G}}_s$ induced by vertex set $\mathcal{X}$ as $G'(\mathcal{X}, \phi_s)$.

Let $U_{\boldsymbol{x}} = N_k(G'(\mathcal{P}_s, \phi_s))$. This has the same distribution as $\tilde{F}_s(\eta_s)$.

Let $V_{\boldsymbol{x}} = N_k(G'(\mathcal{P}_s^{\boldsymbol{x}}, \phi_s))$. We claim that this has the same distribution as the conditional distribution of $\tilde{F}_s(\bigcup_{i=1}^k \{(x_i, \tau_i)\} \cup \eta_s) - 1$ given that $\tilde{f}_s(\bigcup_{i=1}^k \{(x_i, \tau_i)\}, \eta_s) = 1$. This is because by the marking theorem for Poisson processes (see, for example, [14]), the point processes $\mathcal{P}_{s,\boldsymbol{x}}$ and $\mathcal{P}_s^{\boldsymbol{x}}$ are independent and the statement that $\tilde{\mathcal{G}}_s$ has $\mathcal{X}$ as the vertex set of a component is equivalent to the statement that

(i) $\mathcal{P}_{s,\boldsymbol{x}}$ has no points, and

(ii) the graph $G'(\mathcal{X}, \phi_s)$ is connected, which is independent of the outcome of $G'(\mathcal{P}_s^{\boldsymbol{x}}, \phi_s)$.

Also, $U_{\boldsymbol{x}} - V_{\boldsymbol{x}} = U'_{\boldsymbol{x}} - V'_{\boldsymbol{x}}$, where $U'_{\boldsymbol{x}}$ denotes the number of k-components of $G'(\mathcal{P}_s, \phi_s)$ with at least one vertex in $\mathcal{P}_{s,\boldsymbol{x}}$, and $V'_{\boldsymbol{x}}$ is the number of k-components of $G'(\mathcal{P}_s^{\boldsymbol{x}}, \phi_s)$ with at least one neighbour in $\mathcal{P}_{s,\boldsymbol{x}}$.

By the Mecke formula,

$$\mathbb{E}[U'_{\boldsymbol{x}}] = \frac{s^k}{k!} \int_{\mathbb{X}^k} \phi_s(\boldsymbol{x}, \boldsymbol{y}) h_{\phi_s}(\boldsymbol{y}) \exp\left(-s \int \phi_s(z, \boldsymbol{y}) \mu(\mathrm{d}z)\right) \mu^k(\mathrm{d}\boldsymbol{y}),$$

where $\phi_s(\boldsymbol{x}, \boldsymbol{y})$ is given by (6.1). Since we can choose the elements $y_1 \ldots, y_k$ of $\boldsymbol{y}$ in an order such that y_1 is connected to $\mathcal{X}$ and for each $j \geq 2$, y_j is connected to $\{y_1, \ldots, y_{j-1}\}$, by a similar argument to the proof of Lemma 6.1, we have

$$\mathbb{E}[U'_{\boldsymbol{x}}] = O(s^k \kappa_s^k) \times \exp(-\Theta(s\kappa_s)), \tag{6.14}$$

which tends to 0 (uniformly over $\boldsymbol{x}$).

Now $V'_{\boldsymbol{x}}$ is bounded by the number of pairs $(y, \boldsymbol{z})$ with $y \in \mathcal{P}_{s,\boldsymbol{x}}$ and $\boldsymbol{z} = (z_1, \ldots, z_k)$ with $\{z_1, \ldots, z_k\}$ inducing a k-component of $G'(\mathcal{P}_s^{\boldsymbol{x}}, \phi_n)$ and $\boldsymbol{z}$ connected to y. Hence, by the Mecke equation,

$$\begin{aligned}\mathbb{E}[V'_{\boldsymbol{x}}] &\leq s^{k+1} \int_{\mathbb{X}} \int_{\mathbb{X}^k} \phi_s(y, \boldsymbol{x}) \phi_s(y, \boldsymbol{z})(1 - \phi_s(\boldsymbol{x}, \boldsymbol{z})) h_{\phi_s}(\boldsymbol{z}) \\ &\qquad \times \exp\left(-s \int \phi_s(w, \boldsymbol{z})(1 - \phi_s(w, \boldsymbol{x}))\mu(\mathrm{d}w)\right) \mu^k(\mathrm{d}\boldsymbol{z}) \mu(\mathrm{d}y) \\ &= O((s\kappa_s)^{k+1}) \exp(-\Theta(s\kappa_s)),\end{aligned} \tag{6.15}$$

https://doi.org/10.1017/jpr.2018.9 Published online by Cambridge University Press

which tends to 0, uniformly over $\boldsymbol{x}$; here we have used the assumption that $\overline{\phi}_s \le 1-\varepsilon$, so that $1-\phi_s(w,\boldsymbol{x}) \ge \varepsilon^k$ for all $s, w, \boldsymbol{x}$. Therefore, $\mathbb{E}[|U_{\boldsymbol{x}} - V_{\boldsymbol{x}}|] = \mathbb{E}[|U'_{\boldsymbol{x}} - V'_{\boldsymbol{x}}|] \to 0$, uniformly over $\boldsymbol{x} \in \mathbb{X}^k$. Then we can use Theorem 3.1 to obtain the first part of (2.5).

Before completing the proof of part (i) of Theorem 2.3, we prove part (ii), so now instead of (6.4) we assume $\alpha_s := \mathbb{E}[N_k(G(\mathcal{P}_s,\phi_s))] \to \infty$, but $\alpha_s = o(s)$. Then, by (6.8), for every sequence of values of s tending to ∞, there is a subsequence such that either $s\kappa_s \to \infty$ or $k \ge 2$ and $s\kappa_s \to 0$ as $s \to \infty$ along the subsequence.

In both cases, the estimates (6.14) and (6.15) hold so, by Theorem 3.1, we have

$$d_{\mathrm{TV}}(\mathcal{L}(N_k(G(\mathcal{P}_s,\phi_s))), \mathcal{L}(Z_{\alpha_s})) \to 0.$$

Since also $(Z_{\alpha_s} - \alpha_s)/\sqrt{\alpha_s} \xrightarrow{\mathrm{D}} \mathcal{N}$, it follows that $(N_k(G(\mathcal{P}_s,\phi_s)) - \alpha_s)/\sqrt{\alpha_s} \xrightarrow{\mathrm{D}} \mathcal{N}$. That is, for every sequence of values of s tending to ∞, there exists a subsequence such that $(N_k(G(\mathcal{P}_s,\phi_s)) - \alpha_s)/\sqrt{\alpha_s} \xrightarrow{\mathrm{D}} \mathcal{N}$ as $s \to \infty$ along the subsequence. Hence, $(N_k(G(\mathcal{P}_s,\phi_s)) - \alpha_s)/\sqrt{\alpha_s}$ converges in distribution to $\mathcal{N}$ as $s \to \infty$. This completes the proof of part (ii).

Now we return to part (i), so we go back to assuming (6.4). As in the corresponding part of the proof of Theorem 2.1, for $n \in \mathbb{N}$ set $s(n) = n - n^{3/4}$ and $t(n) = n + n^{3/4}$. Then $\mathcal{P}_{s(n)} \subset \mathcal{X}_n \subset \mathcal{P}_{t(n)}$ with high probability, and also the point process $\mathcal{P}_{t(n)} \setminus \mathcal{P}_{s(n)}$ is a Poisson point process with mean measure $2n^{3/4}\mu(\cdot)$, independent of $\mathcal{P}_{s(n)}$. By (6.12), $n^{3/4}\int \phi_n(y,x)\mu(\mathrm{d}y) = o(1)$, uniformly over $x \in \mathbb{X}$, and, therefore, by (6.4) the sequence $(\phi_n)_{n\in\mathbb{N}}$ satisfies

$$\frac{s(n)^k}{k!}\int_{\mathbb{X}^k} \exp\left(-s(n)\int_{\mathbb{X}} \phi_n(z,\boldsymbol{x})\mu(\mathrm{d}z)\right) h_{\phi_n}(\boldsymbol{x})\mu^k(\mathrm{d}\boldsymbol{x}) \to \alpha. \tag{6.16}$$

For $n, \ell \in \mathbb{N}$ with $1 \le \ell \le k$, let $A_{n,\ell}$ be the event that at least one collection of ℓ of the added vertices of $\mathcal{X}_n \setminus \mathcal{P}_{s(n)}$ lies in a k-component of $G(\mathcal{X}_n,\phi_n)$. Let B_n be the event that at least one of the added vertices of $\mathcal{P}_{t(n)} \setminus \mathcal{P}_{s(n)}$ is connected to one of the k-components of $G(\mathcal{P}_{s(n)},\phi_n)$.

If $A_{n,\ell}$ occurs and $\mathcal{P}_{s(n)} \subset \mathcal{X}_n \subset \mathcal{P}_{t(n)}$, then there is at least one pair $(\mathcal{X},\mathcal{Y})$, such that $\mathcal{X} \subset \mathcal{P}_{s(n)}$ has $k-\ell$ elements, $\mathcal{Y} \subset \mathcal{P}_{t(n)} \setminus \mathcal{P}_{s(n)}$ has ℓ elements, $\mathcal{X} \cup \mathcal{Y}$ induces a connected subgraph of $G(\mathcal{P}_{t(n)},\phi_n)$, and there is no connection between any vertex of $\mathcal{X} \cup \mathcal{Y}$ and any vertex of $\mathcal{P}_{s(n)} \setminus \mathcal{X}$ (however, we do allow other connections between vertices of $\mathcal{X} \cup \mathcal{Y}$ and other vertices of $\mathcal{P}_{t(n)} \setminus \mathcal{P}_{s(n)}$). By the Mecke equation, the expected number of such pairs is equal to

$$\begin{aligned}\frac{(2n^{3/4})^\ell}{\ell!}\int \frac{(s(n))^{k-\ell}}{(k-\ell)!}\int & h_{\phi_n}(x_1,\dots,x_{k-\ell},y_1,\dots,y_\ell) \\ & \times \exp\left(-s(n)\int \phi_n(z,\{x_1,\dots,x_{k-\ell},y_1,\dots,y_\ell\})\mu(\mathrm{d}z)\right) \\ & \times \mu^{k-\ell}(\mathrm{d}(x_1,\dots,x_{k-\ell}))\mu^\ell(\mathrm{d}(y_1,\dots,y_\ell)),\end{aligned}$$

which tends to 0 by (6.16), so $\mathbb{P}[A_{n,\ell}] \to 0$. Also, the expected number of k-components in $G(\mathcal{P}_{s(n)},\phi_n)$ which are connected to at least one vertex of $\mathcal{P}_{t(n)} \setminus \mathcal{P}_{s(n)}$ is at most

$$\frac{(s(n))^k}{k!}\int \exp\left(-s(n)\int \phi_n(z,\boldsymbol{x})\mu(\mathrm{d}z)\right) h_{\phi_n}(\boldsymbol{x}) 2n^{3/4}ka_n\mu^k(\mathrm{d}\boldsymbol{x}),$$

https://doi.org/10.1017/jpr.2018.9 Published online by Cambridge University Press

and by (6.16) and (6.12) this tends to 0. Hence, $\mathbb{P}[B_n] \to 0$. By the first part of (2.5), we have, for $k \in \mathbb{N}$, $N_k(G(\mathcal{P}_{s(n)}, \phi_n)) \xrightarrow{\mathrm{D}} Z_\alpha$. Also,

$$\mathbb{P}[N_k(G(\mathcal{X}_n, \phi_n)) \neq N_k(G(\mathcal{P}_{s(n)}, \phi_n))] \le \mathbb{P}\Big[\bigcup_{\ell=1}^{k} A_{n,\ell}\Big] + \mathbb{P}[B_n] + \mathbb{P}[\{Z_{s(n)} \le n \le Z_{t(n)}\}^c],$$

which tends to 0, and the second part of (2.5) follows. □

7. Number of edges

For any graph G, according to our earlier notation, $H_2(G)$ denotes the number of edges of G. Let $\phi \in \Phi$. Then, by the Mecke formula (3.5),

$$\mathbb{E}[H_2(G(\mathcal{P}_s, \phi))] = \mathbb{E}[H_2(G_\phi(\eta_s))] = \frac{1}{2}\int_{\mathbb{X}}\int_{\mathbb{X}} \phi(x, y)s^2\mu(\mathrm{d}x)\mu(\mathrm{d}y). \tag{7.1}$$

Theorem 7.1. *Suppose that $\phi \in \Phi$. Set $\alpha := \mathbb{E}[H_2(G(\mathcal{P}_s, \phi))]$. Then*

$$d_{\mathrm{TV}}(H_2(G(\mathcal{P}_s, \phi)), Z_\alpha) \le (1 \wedge \alpha^{-1}) \int \left(\int \phi(x, y)s\mu(\mathrm{d}y)\right)^2 s\mu(\mathrm{d}x). \tag{7.2}$$

Proof. We use Theorem 3.1. Given $x, y \in \mathbb{X}$ with $\phi(x, y) > 0$, set $\tilde{\mathcal{G}}_s = G_\phi(\eta_s \cup \{(x, \tau_1), (y, \tau_2)\})$. Let $G'(\mathcal{P}_s, \phi)$ denote the subgraph of $\tilde{\mathcal{G}}_s$ induced by the vertex set $\mathcal{P}_s$, which has the same distribution as $G(\mathcal{P}_s, \phi)$.

Set $U_{x,y} = H_2(G'(\mathcal{P}_s, \phi))$ and $V_{x,y} = H_2(\tilde{\mathcal{G}}_s) - \mathbf{1}_{\{\{x,y\} \in E(\tilde{\mathcal{G}}_s)\}}$, where $E(G)$ denotes the set of edges of a graph G. Then $V_{x,y}$ has the conditional distribution of $H_2(\tilde{\mathcal{G}}_s) - 1$ given that $\{x, y\} \in E(\tilde{\mathcal{G}}_s)$. Also, $V_{x,y} \ge U_{x,y}$ and

$$\mathbb{E}[V_{x,y} - U_{x,y}] = \int_{\mathbb{X}} (\phi(x, z) + \phi(y, z))s\mu(\mathrm{d}z) =: w(x, y).$$

Hence, by Theorem 3.1, the left-hand side of (7.2) is bounded by the expression

$$\begin{aligned}&\left(\frac{1}{2}\right)(1 \wedge \alpha^{-1}) \int\int w(x, y)\phi(x, y)s^2\mu(\mathrm{d}x)\mu(\mathrm{d}y)\\ &\qquad = (1 \wedge \alpha^{-1}) \int \left(\int \phi(x, y)s\mu(\mathrm{d}y)\right)^2 s\mu(\mathrm{d}x).\end{aligned}$$

□

For example, consider the geometric setting with μ having bounded, almost everywhere continuous density f with respect to Lebesgue measure on $\mathbb{R}^d$. Suppose that $\phi_s(x, y) = \phi(r_s^{-1}(x - y))$ for some fixed integrable, symmetric and almost everywhere continuous ϕ, and $r_s > 0$ satisfying $s^2 r_s^d \to \beta$ for some $\beta > 0$. Set $\alpha_s := \mathbb{E}[H_2(G(\mathcal{P}_s, \phi_s))]$. Then, by (7.1),

$$\alpha_s \to \frac{\beta}{2}\left(\int \phi(z)\,\mathrm{d}z\right)\int f(x)^2\,\mathrm{d}x =: \alpha.$$

By (7.2), we have $d_{\mathrm{TV}}(\mathcal{L}(H_2(G(\mathcal{P}_s, \phi_s))), \mathcal{L}(Z_{\alpha_s})) = O(s^3 r_s^{2d}) = O(s^{-1})$.

In particular, $H_2(G(\mathcal{P}_s, \phi_s)) \xrightarrow{\mathrm{D}} Z_\alpha$. This could possibly also be proved by deriving a Poisson limit for $H_2(G(\mathcal{X}_n, \phi_n))$ by adapting the argument of [21, Theorem 3.4] to the RCM, and Poissonizing. However, the Poissonization would seem to introduce an error of at least $s^{-1/2}$ in the total variation distance, so the rate of convergence would not be as good.

https://doi.org/10.1017/jpr.2018.9 Published online by Cambridge University Press

Proof of Theorem 2.4. Let $k \in \mathbb{N}$ with $k \geq 2$. Let $\varepsilon > 0$ and assume that $\phi_s \in \Phi_\varepsilon$ for all $s > 0$. Set $G_s := G(\mathcal{P}_s, \phi_s)$. Set $\kappa_s = \sup_{x \in \mathbb{X}} \int \phi_s(x, y)\mu(\mathrm{d}y)$. Then with $h_\phi(\cdot)$ defined in (6.2), by (3.5), we have

$$\mathbb{E}[H_k(G_s)] = \frac{1}{k!} \int s^k h_{\phi_s}(x_1, \ldots, x_k)\mu^k(\mathrm{d}(x_1, \ldots, x_k)) = \Theta(s^k \kappa_s^{k-1}), \tag{7.3}$$

where the second relation is obtained similarly to the proof of Lemma 6.1, using the fact that for each connected graph Γ on $\{1, \ldots, k\}$ we can integrate the variables $x_1, \ldots, x_k$ in an order $(x_{\sigma(1)}, \ldots, x_{\sigma(k)})$ such that for $2 \leq i \leq k$ each successive $\sigma(i)$ is connected in Γ to one or more of $\sigma(1), \ldots, \sigma(i-1)$. Then each successive integral gives another factor of $\Theta(\kappa_s)$.

Let $\boldsymbol{x} = (x_1, \ldots, x_k) \in \mathbb{X}^k$, with $x_1, \ldots, x_k$ distinct and with $h_{\phi_s}(x_1, \ldots, x_k) > 0$. Let $\tilde{\mathcal{G}}_s$ be the graph defined by (6.13), but now conditioned on the subgraph induced by $\{x_1, \ldots, x_k\}$ being connected. Denote by $G'(\mathcal{P}_s, \phi_s)$ the subgraph of $\tilde{\mathcal{G}}_s$ induced by $\mathcal{P}_s$. Set $U_{\boldsymbol{x}} = H_k(G'(\mathcal{P}_s, \phi_s))$ and $V_{\boldsymbol{x}} = H_k(\tilde{\mathcal{G}}_s) - 1$. Then $U_{\boldsymbol{x}}$ has the distribution of $H_k(G_s)$ and $1 + V_{\boldsymbol{x}}$ has the conditional distribution of $H_k(G(\mathcal{P}_s \cup \{x_1, \ldots, x_k\}, \phi_s))$ given that $\{x_1, \ldots, x_k\}$ induces a connected subgraph of this graph.

Now $V_{\boldsymbol{x}} \geq U_{\boldsymbol{x}}$ and we assert that

$$\mathbb{E}[V_{\boldsymbol{x}} - U_{\boldsymbol{x}}] = O\Big(\sum_{j=1}^{k-1} (s^j \kappa_s^j)\Big), \tag{7.4}$$

uniformly over $\boldsymbol{x} \in \mathbb{X}^k$. To see this, observe that $V_{\boldsymbol{x}} - U_{\boldsymbol{x}}$ is the number of pairs $(\mathcal{X}, \mathcal{Y})$ with $\mathcal{X}$ a nonempty subset of $\{x_1, \ldots, x_k\}$ and $\mathcal{Y}$ a nonempty subset of $\mathcal{P}_s$, such that the subgraph of $\tilde{\mathcal{G}}_s$ induced by vertex set $\mathcal{X} \cup \mathcal{Y}$ is a connected graph of order k (to ease notation we ignore the issue of multiplicities in this notation). But then, similarly to the proof of Lemma 6.1, we can take the successive elements y_i of $\mathcal{Y}$ in an order such that each of them is connected to at least one existing vertex from $\mathcal{X} \cup \{y_1, \ldots, y_{i-1}\}$. Then each successive integral gives another factor of $O(s\kappa_s)$.

For part (i), assume that $\mathbb{E}[H_k(G_s)] \to \alpha \in (0, \infty)$. Then, by (7.3) and (7.4), we have $\mathbb{E}[|V_{\boldsymbol{x}} - U_{\boldsymbol{x}}|] = o(1)$, uniformly over $\boldsymbol{x} \in \mathbb{X}^k$. Hence, by Theorem 3.1, $H_k(G_s) \xrightarrow{\mathrm{D}} Z_\alpha$, which is part (i).

For part (ii), set $\alpha_s = \mathbb{E}[H_k(G_s)]$, and assume that $\alpha_s \to \infty$ but $\alpha_s = o(s)$. Then, by (7.3), the assumption $\alpha_s = o(s)$ implies that (recalling $k \geq 2$) we have $s\kappa_s \to 0$. Then, by (7.4) and Theorem 3.1, we have $d_{\mathrm{TV}}(\mathcal{L}(H_k(G_s)), \mathcal{L}(Z_{\alpha_s})) \to 0$. Since also $(Z_{\alpha_s} - \alpha_s)/\sqrt{\alpha_s} \xrightarrow{\mathrm{D}} \mathcal{N}$, we therefore have $(H_k(G_s) - \alpha_s)/\sqrt{\alpha_s} \xrightarrow{\mathrm{D}} \mathcal{N}$. □

8. U-statistics of a Poisson process

Let $k \in \mathbb{N}$, and let $\boldsymbol{S}_k(\mathbb{X}) := \{\xi \in \boldsymbol{S}(\mathbb{X}) : |\xi| = k\}$. Let $h : \boldsymbol{S}_k(\mathbb{X}) \to \{0, 1\}$ be measurable. For $\xi \in \boldsymbol{S}(\mathbb{X})$, set

$$F(\xi) = \sum_{\psi \subset \xi \,:\, |\psi| = k} h(\psi).$$

Let $s > 0$, and let η be a Poisson process on $\mathbb{X}$ with mean measure $\lambda := s\mu$. We seek to apply Theorem 3.1 to $W := F(\eta)$ for this class of choices of F, called U-statistics of the Poisson process η. Assume μ is diffuse; once again, this assumption is for notational convenience only.

https://doi.org/10.1017/jpr.2018.9 Published online by Cambridge University Press

For $\boldsymbol{x} = (x_1, \ldots, x_k) \in \mathbb{X}^k$ with $h(\{x_1, \ldots, x_k\}) = 1$, set $U_{\boldsymbol{x}} = F(\eta)$, and

$$V_{\boldsymbol{x}} = F\Big(\bigcup_{i=1}^{k} \{x_i\} \cup \eta\Big) - 1.$$

Clearly, $V_{\boldsymbol{x}}$ and $U_{\boldsymbol{x}}$ have the required distributional properties in the statement of Theorem 3.1. Also, $V_{\boldsymbol{x}} \geq U_{\boldsymbol{x}}$, and

$$|V_{\boldsymbol{x}} - U_{\boldsymbol{x}}| = \sum_{J \subset [k]:\ 1 \leq |J| < k} \sum_{\psi \subset \eta:\ |\psi| = k - |J|} h(\psi \cup \{x_i : i \in J\}),$$

so using the Mecke formula and integrating over $\boldsymbol{x}$, we obtain

$$\begin{aligned}
&\int (\mathbb{E}[|V_{\boldsymbol{x}} - U_{\boldsymbol{x}}|]) h(\boldsymbol{x}) \lambda^k(\mathrm{d}\boldsymbol{x}) \\
&\quad = \sum_{\ell=1}^{k-1} \frac{1}{(k-\ell)!} \binom{k}{\ell} \int \int h(\{x_1, \ldots, x_\ell, y_1, \ldots, y_{k-\ell}\}) \\
&\qquad\qquad \times \lambda^{k-\ell}(\mathrm{d}(y_1, \ldots, y_{k-\ell})) h(\{x_1, \ldots, x_k\}) \lambda^k(\mathrm{d}\boldsymbol{x}) \\
&\quad = \sum_{\ell=1}^{k-1} \frac{k!}{\ell!(k-\ell)!^2} \int \bigg(\int h(\{x_1, \ldots, x_\ell, y_1, \ldots, y_{k-\ell}\}) \lambda^{k-\ell}(\mathrm{d}(y_1, \ldots, y_{k-\ell})) \bigg)^2 \\
&\qquad\qquad \times \lambda^{\ell}(\mathrm{d}(x_1, \ldots, x_\ell)) \\
&\quad =: \gamma(h, \lambda).
\end{aligned}$$

By Theorem 3.1, if $\mathbb{E}[F(\eta)] = \alpha$ then $d_{\mathrm{TV}}(\mathcal{L}(F(\eta)), \mathcal{L}(Z_\alpha)) \leq (1 \wedge \alpha^{-1})\gamma(h, \lambda)/k!$. Also, $d_{\mathrm{W}}(\mathcal{L}(F(\eta)), \mathcal{L}(Z_\alpha)) \leq 3(1 \wedge \alpha^{-1/2})\gamma(h, \lambda)/k!$. This bound is comparable to the one obtained in Theorem 7.1 of [5]. Our bound has an extra factor $1 \wedge \alpha^{-1/2}$ in front, which may make it better when α is large. Also, unlike [5] we do not make any topological assumptions on the measurable space $\mathbb{X}$. As remarked just after the statement of Theorem 3.1, it is possible to extend that result to the case where the measure λ is σ-finite, and, hence, to extend the above argument likewise, but we do not go into details here.

Acknowledgements

We thank the anonymous referees for some valuable comments on a previous version of this paper. Part of this work was carried out while attending the programme Theoretical Foundations for Statistical Network Analysis at the Isaac Newton Institute for Mathematical Sciences, Cambridge (EPSRC grant number EP/K032208/1).

References

[1] Barbour, A. D., Holst, L. and Janson, S. (1992). *Poisson Approximation* (Oxford Studies Prob. **2**). Oxford University Press.

[2] Bhamidi, S., van der Hofstad, R. and van Leeuwaarden, J. S. H. (2012). Novel scaling limits for critical inhomogeneous random graphs. *Ann. Prob.* **40,** 2299–2361.

[3] Bollobás, B. (2001). *Random Graphs* (Camb. Studies Adv. Math. **73**), 2nd edn. Cambridge University Press.

https://doi.org/10.1017/jpr.2018.9 Published online by Cambridge University Press

[4] Bollobás, B., Janson, S. and Riordan, O. (2007). The phase transition in inhomogeneous random graphs. *Random Structures Algorithms* **31,** 3–122.
[5] Decreusefond, L., Schulte, M. and Thäle C. (2016). Functional Poisson approximation in Kantorovich-Rubinstein distance with applications to U-statistics and stochastic geometry. *Ann. Prob.* **44,** 2147–2197.
[6] Devroye, L. and Fraiman, N. (2014). Connectivity of inhomogeneous random graphs. *Random Structures Algorithms* **45,** 408–420.
[7] Díaz, J., Petit, J. and Serna, M. (2000). Faulty random geometric networks. *Parallel Process. Lett.* **10,** 343–357.
[8] Franceschetti, M. and Meester, R. (2007). *Random Networks for Communication.* Cambridge University Press.
[9] Gupta, B. and Iyer, S. K. (2010). Criticality of the exponential rate of decay for the largest nearest-neighbor link in random geometric graphs. *Adv. Appl. Prob.* **42,** 631–658.
[10] Gupta, P. and Kumar, P. R. (1999). Critical power for asymptotic connectivity in wireless networks. In *Stochastic Analysis, Control, Optimization and Applications*, Birkhäuser, Boston, MA, pp. 547–566.
[11] Hoff, P. D., Raftery, A. E. and Handcock, M. S. (2002). Latent space approaches to social network analysis. *J. Amer. Statist. Assoc.* **97,** 1090–1098.
[12] Hsing, T. and Rootzén, H. (2005). Extremes on trees. *Ann. Prob.* **33,** 413–444.
[13] Iyer, S. K. and Thacker, D. (2012). Nonuniform random geometric graphs with location-dependent radii. *Ann. Appl. Prob.* **22,** 2048–2066.
[14] Kingman, J. F. C. (1993). *Poisson Processes* (Oxford Studies Prob. **3**). Oxford University Press.
[15] Last, G. and Penrose, M. (2018). *Lectures on the Poisson Process* (IMS Textbooks **7**). Cambridge University Press.
[16] Lindvall, T. (1992). *Lectures on the Coupling Method*. John Wiley, New York.
[17] Mao, G. and Anderson, B. D. O. (2012). Towards a better understanding of large-scale network models. *IEEE/ACM Trans. Networking* **20,** 408–421.
[18] Meester, R. and Roy, R. (1996). *Continuum Percolation*. (Camb. Tracts Math. **119**). Cambridge University Press.
[19] Penrose, M. D. (1998). Extremes for the minimal spanning tree on normally distributed points. *Adv. Appl. Prob.* **30,** 628–639.
[20] Penrose, M. D. (1999). On k-connectivity for a geometric random graph. *Random Structures Algorithms* **15,** 145–164.
[21] Penrose, M. (2003). *Random Geometric Graphs* (Oxford Studies Prob. **5**). Oxford University Press.
[22] Penrose, M. D. (2016). Connectivity of soft random geometric graphs. *Ann. Appl. Prob.* **26,** 986–1028.
[23] Rastelli, R., Friel, N. and Raftery, A. E. (2016). Properties of latent variable network models. *Network Sci.* **4,** 407–432.
[24] Snijders, T. A. B. and Nowicki, K. (1997). Estimation and prediction for stochastic blockmodels for graphs with latent block structure. *J. Classification* **14,** 75–100.
[25] Söderberg, B. (2002). General formalism for inhomogeneous random graphs. *Phys Rev. E* **66,** 066121.
[26] Yi, C.-W., Wan, P.-J., Lin, K.-W. and Huang, C.-H. (2006). Asymptotic distribution of the number of isolated nodes in wireless ad hoc networks with unreliable nodes and links. In *Proc. Global Telecommun. Conf., 2006*, IEEE.

https://doi.org/10.1017/jpr.2018.9 Published online by Cambridge University Press