ON SPARSITY, POWER-LAW, AND CLUSTERING PROPERTIES OF GRAPHEX PROCESSES[‡]

FRANÇOIS CARON,* ** *University of Oxford*
FRANCESCA PANERO,*** *London School of Economics and Political Science*
JUDITH ROUSSEAU,* **** *University of Oxford*

Abstract

This paper investigates properties of the class of graphs based on exchangeable point processes. We provide asymptotic expressions for the number of edges, number of nodes, and degree distributions, identifying four regimes: (i) a dense regime, (ii) a sparse, almost dense regime, (iii) a sparse regime with power-law behaviour, and (iv) an almost extremely sparse regime. We show that, under mild assumptions, both the global and local clustering coefficients converge to constants which may or may not be the same. We also derive a central limit theorem for subgraph counts and for the number of nodes. Finally, we propose a class of models within this framework where one can separately control the latent structure and the global sparsity/power-law properties of the graph.

Keywords: Networks; sparsity; Poisson processes; community structure; power law; generalised graphon; transitivity; subgraph counts

2020 Mathematics Subject Classification: Primary 05C80
Secondary 60F15; 60G55

1. Introduction

The ubiquitous availability of large, structured network data in various scientific areas ranging from biology to social sciences has been a driving force in the development of statistical network models [23, 30]. Vertex-exchangeable random graphs, also known as *W*-random graphs or graphon models [1, 15, 19, 28], offer in particular a flexible and tractable class of random graph models. It includes many models, such as the stochastic block-model [31], as special cases. Various parametric and nonparametric model-based approaches [25, 26, 33] and nonparametric estimation procedures [14, 16, 41] have been developed within this framework. Although the framework is very flexible, it is known that vertex-exchangeable random graphs are dense [28, 32]; that is, the number of edges scales quadratically with the number of nodes. This property is considered unrealistic for many real-world networks.
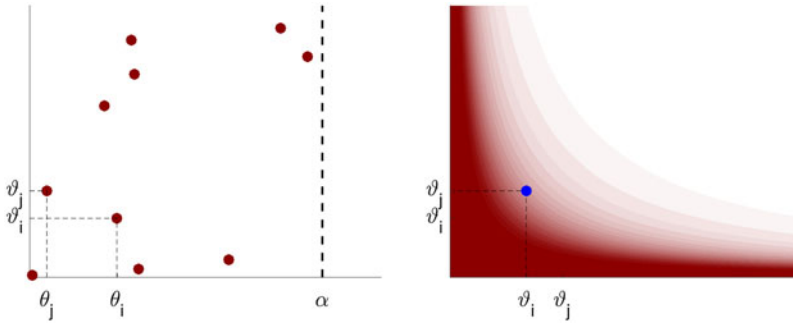
FIGURE 1. Illustration of the graph model based on exchangeable point processes. Left: A unit-rate Poisson process (θ_i, ϑ_i), $i \in \mathbb{N}$, on $(0, \alpha] \times \mathbb{R}_+$. Right: For each pair $i \leq j$, set $Z_{ij} = Z_{ji} = 1$ with probability $W(\vartheta_i, \vartheta_j)$. Here, W is indicated by the red shading (darker shading indicates higher value). Similar to [12, Figure 5].

To achieve sparsity, rescaled graphon models have been proposed in the literature [4, 5, 8, 41]. While these models can capture sparsity, they are not projective; additionally, standard rescaled graphon models cannot simultaneously capture sparsity and a clustering coefficient bounded away from zero (see Section 5).

These limitations have been overcome by another line of works initiated by [9, 12, 38]. They showed that, by modelling the graph as an exchangeable point process, one can naturally extend the classical vertex-exchangeable/graphon framework to the sparse regime, while preserving its flexibility and tractability. In such a representation, introduced by [12], nodes are embedded at some location $\theta_i \in \mathbb{R}_+$, and the set of edges is represented by a point process on the plane,

$$\sum_{i,j} Z_{ij} \delta_{(\theta_i, \theta_j)}, \tag{1}$$

where $Z_{ij} = Z_{ji}$ is a binary variable indicating whether there is an edge between node θ_i and node θ_j. Finite-size graphs are obtained by restricting the point process (1) to points (θ_i, θ_j) such that $\theta_i, \theta_j \leq \alpha$, with α a positive parameter controlling the size of the graph. Focusing on a particular construction as a case study, [12] showed that one can obtain sparse and exchangeable graphs within this framework; it also pointed out that exchangeable random measures admit a representation theorem due to [22], giving a general construction for such graph models. The papers [18, 37] developed sparse graph models with (overlapping) community structure within this framework. The papers [9, 38] showed how such a construction naturally generalises the dense exchangeable graphon framework to the sparse regime, and analysed some of the properties of the associated class of random graphs, called *graphex processes*. (The paper [38] introduced the term *graphex* and referred to the class of random graphs as *Kallenberg exchangeable graphs*, but the term *graphex processes* is now more commonly used.) Further properties were derived by [10, 20, 21, 39]. Following the notation of [38], and ignoring additional terms corresponding to stars and isolated edges, the graph is then parametrised by a symmetric measurable function $W : \mathbb{R}_+^2 \to [0, 1]$, where for each $i \leq j$,

$$Z_{ij} \mid (\theta_k, \vartheta_k)_{k=1,2,\dots} \sim \text{Bernoulli}\{W(\vartheta_i, \vartheta_j)\}, \tag{2}$$

where $(\theta_k, \vartheta_k)_{k=1,2,\dots}$ is a unit-rate Poisson process on $\mathbb{R}_+^2$. See Figure 1 for an illustration of the model construction. The function W is a natural generalisation of the graphon for dense exchangeable graphs [9, 38], and we refer to it as the graphon function.

Power-law Degree distribution   Average local and global clustering coefficients   Local clustering coefficient per degree
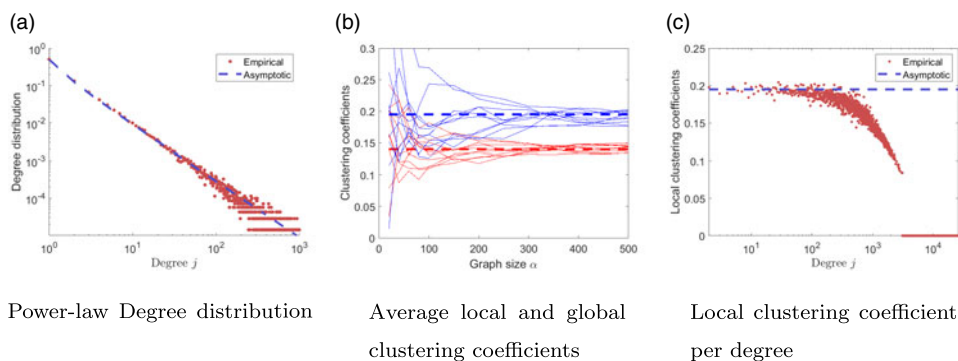
FIGURE 2. Illustration of some of the asymptotic results developed in this paper, applied to the generalised graphon model defined by Equations (41) and (46) with $\sigma_0 = 0.2$ and $\tau_0 = 2$. (a) Empirical degree distribution for a graph of size $\alpha = 1000$ (red) and asymptotic degree distribution (dashed blue; see Corollary 1). (b) Average local (blue) and global (red) clustering coefficients for 10 graphs of growing sizes. Limit values are represented by dashed lines (see Propositions 5 and 6). (c) Local clustering coefficient for nodes of a given degree j, for a graph of size $\alpha = 1000$. The limit value is represented by a dashed line (see Proposition 6).

This paper investigates asymptotic properties of the general class of graphs based on exchangeable point processes defined by Equations (1) and (2). Our findings can be summarised as follows.

(i) We relate the sparsity and power-law properties of the graph to the tail behaviour of the marginal of the graphon function W, identifying four regimes: (a) a dense regime, (b) a sparse (almost dense) regime without power-law behaviour, (c) a sparse regime with power-law behaviour, and (d) an almost extremely sparse regime. In the sparse, power-law regime, the power-law exponent is in the range $(1, 2)$.

(ii) We derive the asymptotic properties of the global and local clustering coefficients, two standard measures of the transitivity of the graph.

(iii) We give a central limit theorem for subgraph counts and for the number of nodes in the graph.

(iv) We introduce a parametrisation that allows us to model separately the global sparsity structure and other local properties such as community structure. Such a framework enables us to sparsify any dense graphon model, and to characterise its sparsity properties.

(v) We show that the results apply to a wide range of sparse and dense graphex processes, including the models studied by [12, 18, 37].

Some of the asymptotic results are illustrated in Figure 2 for a specific graphex process in the sparse, power-law regime.

The article is organised as follows. In Section 2 we give the notation and the main assumptions. In Section 3, we derive the asymptotic results for the number of nodes, degree distribution, and clustering coefficients. In Section 4, we derive central limit theorems for subgraphs and for the number of nodes. Section 5 discusses related work. In Section 6 we provide specific examples of sparse and dense graphs and show how to apply the results of the

previous section to those models. In Section 7 we describe a generic construction for graphs with local/global structure and adapt some results of Section 3 to this setting. Most of the proofs are given in the main text, with some longer proofs in the appendix, together with some technical lemmas and background material. Other, more technical proofs are given in supplementary material [13].

Throughout the document, we use the expressions $X_\alpha \sim Y_\alpha$ and $X_\alpha = o(Y_\alpha)$ respectively for $X_\alpha/Y_\alpha \to 1$ and $X_\alpha/Y_\alpha \to 0$. Both $X_\alpha \lesssim Y_\alpha$ and $X_\alpha = O(Y_\alpha)$ are used to express $\limsup X_\alpha/Y_\alpha < \infty$. The notation $X_\alpha \asymp Y_\alpha$ means both $X_\alpha \lesssim Y_\alpha$ and $Y_\alpha \lesssim X_\alpha$ hold. All unspecified limits are when α tends to infinity. When X_α and/or Y_α are random quantities, the asymptotic relation is meant to hold almost surely.

2. Notation and assumptions

2.1. Notation

Let $M = \sum_i \delta_{(\theta_i, \vartheta_i)}$ be a unit-rate Poisson random measure on $(0, +\infty)^2$, and let $W : [0, +\infty)^2 \to [0, 1]$ be a symmetric measurable function such that $\lim_{x\to\infty} W(x, x)$ and $\lim_{x\to 0} W(x, x)$ both exist (by (3), this implies that $\lim_{x\to\infty} W(x, x) = 0$) and

$$0 < \overline{W} = \int_{\mathbb{R}_+^2} W(x, y)dxdy < \infty, \qquad \int_0^\infty W(x, x)dx < \infty. \tag{3}$$

Let $(U_{ij})_{i,j\in\mathbb{N}^2}$ be a symmetric array of independent random variables, with $U_{ij} \sim U(0, 1)$ if $i \le j$ and $U_{ij} = U_{ji}$ for $i > j$. Let $Z_{ij} = \mathbb{1}_{U_{ij} \le W(\vartheta_i, \vartheta_j)}$ be a binary random variable indicating whether there is a link between i and j, where $\mathbb{1}_A$ denotes the indicator function.

Restrictions of the point process $\sum_{ij} Z_{ij}\delta_{(\theta_i, \theta_j)}$ to squares $[0, \alpha]^2$ then define a growing family of random graphs $(\mathcal{G}_\alpha)_{\alpha \ge 0}$, called a graphex process, where $\mathcal{G}_\alpha = (\mathcal{V}_\alpha, \mathcal{E}_\alpha)$ denotes a graph of size $\alpha \ge 0$ with vertex set $\mathcal{V}_\alpha$ and edge set $\mathcal{E}_\alpha$, defined by

$$\mathcal{V}_\alpha = \{\theta_i \mid \theta_i \le \alpha \text{ and } \exists \theta_k \le \alpha \text{ s.t. } Z_{ik} = 1\}, \tag{4}$$

$$\mathcal{E}_\alpha = \left\{\{\theta_i, \theta_j\} \mid \theta_i, \theta_j \le \alpha \text{ and } Z_{ij} = 1\right\}. \tag{5}$$

The connection between the point process and graphex process is illustrated in Figure 3. The conditions (3) are sufficient (though not necessary) conditions for $|\mathcal{E}_\alpha|$ (hence $|\mathcal{V}_\alpha|$) to be almost surely finite, and for the graphex process to be well defined [38, Theorem 4.9]. Note crucially that the graphs $\mathcal{G}_\alpha$ have no isolated vertices (that is, no vertices of degree 0), and that the number of nodes $|\mathcal{V}_\alpha|$ and the number of edges $|\mathcal{E}_\alpha|$ are both random variables.

We now define a number of summary statistics of the graph $\mathcal{G}_\alpha$. For $i \ge 1$, let

$$D_{\alpha,i} = \sum_k Z_{ik}\mathbb{1}_{\theta_k \le \alpha}.$$

If $\theta_i \in \mathcal{V}_\alpha$, then $D_{\alpha,i} \ge 1$ corresponds to the degree of the node θ_i in the graph $\mathcal{G}_\alpha$ of size α; otherwise $D_{\alpha,i} = 0$. Let $N_\alpha = |\mathcal{V}_\alpha|$ and $N_{\alpha,j}$ be the number of nodes and the number of nodes of degree j, $j \ge 1$, respectively,

$$N_\alpha = \sum_i \mathbb{1}_{\theta_i \le \alpha} \mathbb{1}_{D_{\alpha,i} \ge 1}, \qquad N_{\alpha,j} = \sum_i \mathbb{1}_{\theta_i \le \alpha} \mathbb{1}_{D_{\alpha,i} = j}, \tag{6}$$
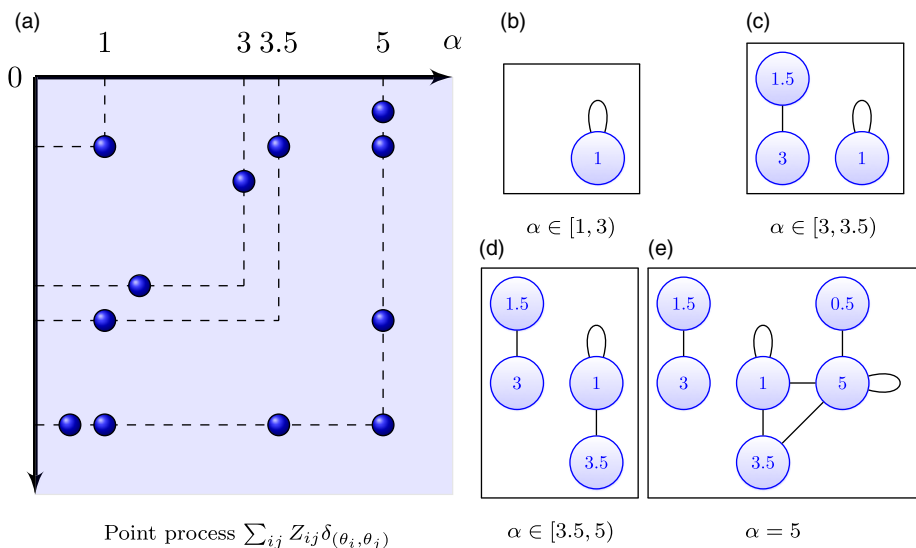
FIGURE 3. Illustration of the connection between the point process on the plane and the graphex process. (a) Point process $\sum_{ij} Z_{ij} \delta_{(\theta_i, \theta_j)}$ on the plane. (b–e) Associated graphs $\mathcal{G}_\alpha$ for (b) $\alpha \in [1, 3)$, (c) $\alpha \in [3, 3.5)$, (d) $\alpha \in [3.5, 5)$, and (e) $\alpha = 5$. Note that the graph is empty for $\alpha < 1$.

and $N_\alpha^{(e)} = |\mathcal{E}_\alpha|$ the number of edges

$$N_\alpha^{(e)} = \frac{1}{2} \sum_{i \neq j} Z_{ij} \mathbb{1}_{\theta_i \leq \alpha} \mathbb{1}_{\theta_j \leq \alpha} + \sum_i Z_{ii} \mathbb{1}_{\theta_i \leq \alpha}. \tag{7}$$

For $i \geq 1$, let

$$T_{\alpha,i} = \frac{1}{2} \sum_{j,k | j \neq k \neq i} Z_{ij} Z_{jk} Z_{ik} \mathbb{1}_{\theta_i \leq \alpha} \mathbb{1}_{\theta_j \leq \alpha} \mathbb{1}_{\theta_k \leq \alpha}. \tag{8}$$

If $\theta_i \in \mathcal{V}_\alpha$, $T_{\alpha,i}$ corresponds to the number of triangles containing node θ_i in the graph $\mathcal{G}_\alpha$; otherwise $T_{\alpha,i} = 0$. Let

$$T_\alpha = \frac{1}{3} \sum_i T_{\alpha,i} = \frac{1}{6} \sum_{i \neq j \neq k} Z_{ij} Z_{jk} Z_{ik} \mathbb{1}_{\theta_i \leq \alpha} \mathbb{1}_{\theta_j \leq \alpha} \mathbb{1}_{\theta_k \leq \alpha} \tag{9}$$

denote the total number of triangles and

$$A_\alpha = \sum_i \frac{D_{\alpha,i}(D_{\alpha,i} - 1)}{2} = \frac{1}{2} \sum_{i \neq j \neq k} Z_{ij} Z_{jk} \mathbb{1}_{\theta_i \leq \alpha} \mathbb{1}_{\theta_j \leq \alpha} \mathbb{1}_{\theta_k \leq \alpha} \tag{10}$$

the total number of adjacent edges in the graph $\mathcal{G}_\alpha$. The global clustering coefficient, also known as the transitivity coefficient, is defined as

$$C_\alpha^{(g)} = \frac{3T_\alpha}{A_\alpha} \tag{11}$$

if $A_\alpha \geq 1$ and 0 otherwise. The global clustering coefficient counts the proportion of closed connected triplets over all the connected triplets, or equivalently the fraction of pairs of nodes

connected to the same node that are themselves connected, and is a standard measure of the transitivity of a network [30, Section 7.9]. Another measure of the transitivity of the graph is the local clustering coefficient. For any degree $j \geq 2$, define

$$C_{\alpha,j}^{(\ell)} = \frac{2}{j(j-1)N_{\alpha,j}} \sum_i T_{\alpha,i} \mathbb{1}_{D_{\alpha,i}=j} \tag{12}$$

if $N_{\alpha,j} \geq 1$ and 0 otherwise. Then $C_{\alpha,j}^{(\ell)}$ corresponds to the proportion of pairs of neighbours of nodes of degree j that are connected. The average local clustering coefficient is obtained by

$$\overline{C}_\alpha^{(\ell)} = \frac{1}{N_\alpha - N_{\alpha,1}} \sum_{j \geq 2} N_{\alpha,j} C_{\alpha,j}^{(\ell)} \tag{13}$$

if $N_\alpha - N_{\alpha,1} \geq 1$ and $\overline{C}_\alpha^{(\ell)} = 0$ otherwise.

2.2. Assumptions

We will make use of the following three assumptions. Assumption 1 characterises the behaviour of the small-degree nodes. Assumption 2 is a technical assumption to obtain the almost sure results. Assumption 3 characterises the behaviour of large-degree nodes.

A central quantity of interest in the analysis of the asymptotic properties of graphex processes is the marginal generalised graphon function $\mu : (0, \infty) \to \mathbb{R}_+$, defined for $x > 0$ by

$$\mu(x) = \int_0^\infty W(x, y)dy. \tag{14}$$

The integrability of the generalised graphon W implies that μ is integrable. Ignoring loops (self-edges), the expected number of connections of a node with parameter ϑ is proportional to $\mu(\vartheta)$. Therefore, assuming μ is monotone decreasing, its behaviour at infinity controls the small-degree nodes, while its behaviour at zero controls the large-degree nodes.

For mathematical convenience, it will be easier to work with the generalised inverse μ^{-1} of μ. The behaviour at zero of μ^{-1} then controls the small-degree nodes, while the behaviour of μ^{-1} at infinity controls large-degree nodes.

The following assumption characterises the behaviour of μ at infinity or, equivalently, of μ^{-1} at zero. We require μ^{-1} to behave approximately as a power function $x^{-\sigma}$ around zero, for some $\sigma \in [0, 1]$. This behaviour, known as regular variation, has been extensively studied (see, e.g., [6]) and we provide some background on it in Appendix C.

Assumption 1. *Assume μ is non-increasing, with generalised inverse $\mu^{-1}(x) = \inf\{y > 0 \mid \mu(y) \leq x\}$, such that*

$$\mu^{-1}(x) \sim \ell(1/x)x^{-\sigma} \text{ as } x \to 0, \tag{15}$$

where $\sigma \in [0, 1]$ and ℓ is a slowly varying function at infinity: for all $c > 0$, $\lim_{t\to\infty} \ell(ct)/\ell(t) = 1$.

Examples of slowly varying functions ℓ include functions converging to a strictly positive constant, and powers of logarithms. Note that Assumption 1 implies that, for $\sigma \in (0, 1)$, $\mu(t) \sim \overline{\ell}(t)t^{-1/\sigma}$ as $t \to \infty$ for some slowly varying function $\overline{\ell}$. We can differentiate four cases, as will be formally derived in Corollary 1:

(i) Dense case: $\sigma = 0$ and $\lim_{t \to \infty} \ell(t) < \infty$. In this case, $\lim_{x \to 0} \mu^{-1}(x) < \infty$, and hence μ has bounded support. The other three cases are all sparse cases.

(ii) Almost dense case: $\sigma = 0$ and $\lim_{t \to \infty} \ell(t) = \infty$. In this case μ has full support and super-polynomially decaying tails.

(iii) Sparse case with power law: $\sigma \in (0, 1)$. In this case μ has full support and polynomially decaying tails (up to a slowly varying function).

(iv) Very sparse case: $\sigma = 1$. In this case μ has full support and very light tails. In order for μ^{-1} (and hence W) to be integrable, we need ℓ to go to zero sufficiently fast.

Now define, for $x, y > 0$,

$$\nu(x, y) = \int_0^\infty W(x, z)W(y, z)dz. \tag{16}$$

The expected number of common neighbours of nodes with parameters $(\vartheta_1, \vartheta_2)$ is proportional to $\nu(\vartheta_1, \vartheta_2)$.

The following assumption is a technical assumption needed in order to obtain the almost sure results on the number of nodes and degrees. The paper [38] made a similar assumption to obtain results in probability; see the discussion section for further details.

Assumption 2. *Assume that there exist C_1, $a > 0$ and $x_0 \geq 0$ such that for all $x, y > x_0$,*

$$\nu(x, y) \leq C_1 \mu(x)^a \mu(y)^a, \quad \mu(x_0) > 0, \quad \begin{cases} a > \max\left(\frac{1}{2}, \sigma\right) & \text{if } \sigma \in [0, 1), \\ a = 1 & \text{if } \sigma = 1. \end{cases} \tag{17}$$

Remark 1. Assumption 2 is trivially satisfied when the function W is separable, $W(x, y) = \mu(x)\mu(y)/\overline{W}$. Assumptions 1 and 2 are also satisfied if

$$W(x, y) = 1 - e^{-f(x)f(y)/\bar{f}} \tag{18}$$

for some positive, non-increasing, measurable function f with $\bar{f} = \int_0^\infty f(x)dx < \infty$ and generalised inverse f^{-1} satisfying $f^{-1}(x) \sim \ell(1/x)x^{-\sigma}$ as x tends to 0. In this case, μ is monotone non-increasing. We have

$$\mu\{f^{-1}(x)\} = \int_0^\infty \left\{1 - e^{-xf(y)/\bar{f}}\right\}dy = x \int_0^\infty e^{-xu/\bar{f}}f^{-1}(u)/\bar{f}\,du \sim x$$

as x tends to 0 by dominated convergence. Hence $f\{\mu^{-1}(x)\} \sim x$ as x tends to 0 and $f^{-1}[f\{\mu^{-1}(x)\}] \sim \ell(1/x)x^{-\sigma}$. Assumption 2 follows from the inequality $W(x, y) \leq f(x)f(y)/\bar{f}$. Other examples are considered in Section 6.

The following assumption is used to characterise the asymptotic behaviour of both small- and large-degree nodes.

Assumption 3. *Assume $\mu^{-1}(t) = \int_t^\infty f(x)dx$ where f is continuous on $(0, \infty)$ and the following hold:*

$$(a) \qquad f(x) \sim \tau x^{-\tau-1}\ell_2(x) \text{ as } x \to \infty,$$

$$(b) \qquad f(x) \sim x^{-\tilde{\sigma}-1}\tilde{\ell}_2(1/x) \text{ as } x \to 0,$$

where $\tau > 0$, $\tilde{\sigma} \leq 1$, and ℓ_2, $\tilde{\ell}_2$ are slowly varying functions.

Note that Assumption 3 implies that $\mu^{-1}(x) \sim x^{-\tau} \ell_2(x)$ as $x \to \infty$, and $\mu(t) \sim \overline{\ell}_2(t)t^{-1/\tau}$ as $t \to 0$, for some slowly varying function $\overline{\ell}_2$. Assumption 3 also implies Assumption 1 with $\sigma = \max(\tilde{\sigma}, 0)$, $\ell(x) = \frac{1}{\sigma} \tilde{\ell}_2(x)$ if $\tilde{\sigma} \neq 0$, and $\ell(x) = o(\tilde{\ell}_2(x))$ if $\tilde{\sigma} = 0$.

Finally, we state an assumption on $\nu(x, y)$, the quantity proportional to the expected number of common neighbours of two nodes with parameters x and y, defined in Equation (16). This technical assumption is used to prove a result on the asymptotic behaviour of the variance of the number of nodes (Proposition 3) and the central limit theorem for sparse graphs enunciated in Section 4.3.

Assumption 4. *Assume that there exist* $0 < C_0 \leq C_1$ *and* $x_0 \geq 0$ *such that for all* $x, y > x_0$,

$$C_0 \mu(x)\mu(y) \leq \nu(x, y) \leq C_1 \mu(x)\mu(y).$$

Assumption 4 holds when W is separable, as well as in the model of [12] under some moment conditions (see Section 6.5). Obviously, Assumption 4 implies that Assumption 2 is satisfied with $a = 1$.

3. Asymptotic behaviour of various statistics of the graph

3.1. Asymptotic behaviour of the number of edges, number of nodes, and degree distribution

In this section we characterise the almost sure and expected behaviour of the number of nodes N_α, number of edges $N_\alpha^{(e)}$, and number of nodes with j edges $N_{\alpha,j}$. These results allow us to provide precise statements about the sparsity of the graph and the asymptotic power-law properties of its degree distribution.

We first recall existing results on the asymptotic growth of the number of edges. The growth of the mean number of edges has been shown by [38], and the almost sure convergence follows from [9, Proposition 56].

Proposition 1. (Number of edges [9, 38].) *As* α *goes to infinity, almost surely*

$$N_\alpha^{(e)} \sim E(N_\alpha^{(e)}) \sim \alpha^2 \overline{W}/2. \tag{19}$$

The following two theorems provide a description of the asymptotic behaviour of the terms N_α, $N_{\alpha,j}$ in expectation and almost surely.

Theorem 1. *For* $\sigma \in [0, 1]$, *let* ℓ_σ *be slowly varying functions defined as*

$$\ell_1(t) = \int_t^\infty y^{-1} \ell(y) dy \qquad and \quad \ell_\sigma(t) = \ell(t)\Gamma(1 - \sigma) \text{ for } \sigma \in [0, 1). \tag{20}$$

Under Assumption 1, for all $\sigma \in [0, 1]$,

$$E(N_\alpha) \sim \alpha^{1+\sigma} \ell_\sigma(\alpha). \tag{21}$$

If $\sigma = 0$ *then for* $j \geq 1$

$$E(N_{\alpha,j}) = o\{\alpha \ell(\alpha)\}.$$

If $\sigma \in (0, 1)$ *then for* $j \geq 1$

$$E(N_{\alpha,j}) \sim \frac{\sigma \Gamma(j - \sigma)}{j!} \alpha^{1+\sigma} \ell(\alpha).$$

Finally, if $\sigma = 1$, then

$$E(N_{\alpha,j}) \sim \begin{cases} \alpha^2 \ell_1(\alpha), & j = 1, \\ \frac{\alpha^2}{j(j-1)} \ell(\alpha), & j \geq 2. \end{cases}$$

Theorem 1 follows rather directly from asymptotic properties of regularly varying functions [17], recalled in Lemmas B.2 and B.3 in the appendix. Details of the proof are given in Appendix A.1. Note that $\ell(\alpha) = o(\ell_1(\alpha))$; hence, for $\sigma = 1$, $E(N_{\alpha,j}) = o\{E(N_{\alpha,1})\}$ for all $j \geq 2$.

The paper [38] shows that, under Assumption 2 with $a = 1$, we have, in probability,

$$N_\alpha \sim E(N_\alpha), \quad \sum_{k \geq j} N_{\alpha,k} \sim E\left(\sum_{k \geq j} N_{\alpha,k}\right) \quad \text{for } j \geq 1.$$

The next theorem shows that the asymptotic equivalence holds almost surely under Assumptions 1 and 2. Additionally, combining these results with Theorem 1 allows us to characterise the almost sure asymptotic behaviour of the number of nodes and the number of nodes of a given degree. The proof of Theorem 2 is given in Section 3.2.

Theorem 2. *Under Assumptions 1 and 2, we have almost surely as α tends to infinity*

$$N_\alpha \sim E(N_\alpha), \quad \sum_{k \geq j} N_{\alpha,k} \sim E\left(\sum_{k \geq j} N_{\alpha,k}\right) \quad \text{for } j \geq 1. \tag{22}$$

Combining this with Theorem 1, we obtain that, for all $\sigma \in [0, 1]$,

$$N_\alpha \sim \alpha^{1+\sigma} \ell_\sigma(\alpha).$$

Moreover, for $j \geq 1$, if $\sigma = 0$ then $N_{\alpha,j} = o\{\alpha \ell(\alpha)\}$, while if $0 < \sigma < 1$ then

$$N_{\alpha,j} \sim \frac{\sigma \Gamma(j - \sigma)}{j!} \alpha^{1+\sigma} \ell(\alpha).$$

If $\sigma = 1$, then $N_{\alpha,1} \sim \alpha^2 \ell_1(\alpha)$ and for all $j \geq 2$ we also have $N_{\alpha,j} = o\{\alpha^2 \ell_1(\alpha)\}$.

The following result is a corollary of Theorem 2 which shows how the parameter σ relates to the sparsity and power-law properties of the graphs. We denote by $\ell^\#$ the de Bruijn conjugate (see Definition C.2 in the appendix) of the slowly varying function ℓ.

Corollary 1. (Sparsity and power-law degree distribution.) *Assume Assumptions 1 and 2. For $\sigma \in [0, 1]$, almost surely as α tends to infinity,*

$$N_\alpha^{(e)} \sim \frac{\overline{W}}{2} N_\alpha^{2/(1+\sigma)} \ell_\sigma^*(N_\alpha), \quad \ell_\sigma^*(y) = \left[\left\{\ell_\sigma^{1/(1+\sigma)}(y^{1/1+\sigma})\right\}^\#\right]^2.$$

The function $\ell_\sigma^*(y)$ is slowly varying and the graph is dense if $\sigma = 0$ and $\lim_{t \to \infty} \ell(t) = C < \infty$, as $N_\alpha^{(e)}/N_\alpha^2 \to C^2 \overline{W}/2$ almost surely. Otherwise, if $\sigma > 0$ or $\sigma = 0$ and $\lim_t \ell(t) = \infty$, the graph is sparse, as $N_\alpha^{(e)}/N_\alpha^2 \to 0$. Additionally, for $\sigma \in [0, 1)$, for any $j = 1, 2, \ldots$,

$$\frac{N_{\alpha,j}}{N_\alpha} \to \frac{\sigma \Gamma(j - \sigma)}{j! \Gamma(1 - \sigma)} \tag{23}$$

almost surely. If $\sigma > 0$, this corresponds to a degree distribution with power-law behaviour, as, for j large,

$$\frac{\sigma\,\Gamma(j-\sigma)}{j!\,\Gamma(1-\sigma)} \sim \frac{\sigma}{\Gamma(1-\sigma)j^{1+\sigma}}.$$

For $\sigma = 1$, $N_{\alpha,1}/N_\alpha \to 1$ and $N_{\alpha,j}/N_\alpha \to 0$ for $j \geq 2$; hence the nodes of degree 1 dominate in the graph.

Remark 2. If $\sigma = 0$ and $\lim_{t \to \infty} \ell(t) = \infty$, the graph is almost dense; that is, $N_\alpha^{(e)}/N_\alpha^2 \to 0$ and $N_\alpha^{(e)}/N_\alpha^{2-\epsilon} \to \infty$ for any $\epsilon > 0$. If $\sigma = 1$, the graph is almost extremely sparse [8], as $N_\alpha^{(e)}/N_\alpha \to \infty$ and $N_\alpha^{(e)}/N_\alpha^{1+\epsilon} \to 0$ for any $\epsilon > 0$.

The above results are important in terms of modelling aspects, since they allow a precise description of the degrees and number of edges as a function of the number of nodes. They can also be used to conduct inference on the parameters of the statistical network model, since the behaviour of most estimators will depend heavily on the behaviour of N_α, $N_\alpha^{(e)}$, and possibly $N_{\alpha,j}$. For instance the naive estimator of σ given by

$$\hat{\sigma} = \frac{2\log N_\alpha}{\log N_\alpha^{(e)}} - 1 \tag{24}$$

is almost surely consistent. Note that, following an earlier version of the present paper, [29] proposed an alternative estimator for σ, with better statistical properties. Indeed, under Assumptions 1 and 2, using Theorems 1 and 2, we have almost surely $N_\alpha^2 \sim \alpha^{2+2\sigma}\ell_\sigma(\alpha)^2$ and $N_\alpha^{(e)} \sim \alpha^2\overline{W}/2$. Hence

$$\log\frac{N_\alpha^2}{N_\alpha^{(e)}} \sim 2\sigma\,\log(\alpha) + \log\left(\ell_\sigma(\alpha)^2 2/\overline{W}\right)$$

and the result follows as $\log \ell_\sigma(\alpha)/\log\alpha \to 0$.

All of the above results concern the behaviour of small-degree nodes, where the degree j is fixed as the size of the graph goes to infinity. It is also of interest to look at the number of nodes of degree j as both α and j tend to ∞. We show in the next proposition that this is controlled by the behaviour of the function f, introduced in Assumption 3, at 0 or ∞.

Proposition 2. (Power law for high-degree nodes.) *Assume that Assumption 3 holds. Then when $j \to \infty$, $\log\alpha = o(j)$, and $j/\alpha \to c_0 \in [0,\infty]$, we have*

$$E\big(N_{\alpha,j}\big) \sim f(j/\alpha).$$

Note that Proposition 2 implies that when $j/\alpha \to \infty$,

$$E\big(N_{\alpha,j}\big) \sim \frac{\tau\alpha^{1+\tau}\ell_2(j/\alpha)}{j^{1+\tau}},$$

which corresponds to power-law behaviour with exponent $1 + \tau$. If $j/\alpha \to 0$ then

$$E\big(N_{\alpha,j}\big) \sim \frac{\alpha^{1+\tilde{\sigma}}\,\tilde{\ell}_2(\alpha/j)}{j^{1+\tilde{\sigma}}}.$$

This is similar to the asymptotic results for j fixed, stated in Theorem 1, noting that

$$\frac{\Gamma(j-\sigma)}{j!} \sim j^{-1-\sigma}$$

as $j \to \infty$. Finally, if $j/\alpha \to c_0 \in (0, \infty)$, then $E(N_{\alpha,j}) \sim f(c_0) \in (0, \infty)$.

Proof. Under Assumption 3, we have $\mu^{-1}(t) = \int_t^\infty f(x)dx$ with

$$f(x) \sim \tau x^{-\tau-1}\ell_2(x) \text{ as } x \to \infty, \quad f(x) \sim \tau x^{-\sigma-1}\tilde{\ell}_2(x) \text{ as } x \to 0.$$

From [38, Theorem 5.5] we have, assuming that $W(x, x) = 0$ for the sake of simplicity,

$$\begin{aligned}
E(N_{\alpha,j}) &= \alpha \int_0^\infty e^{-\alpha\mu(\vartheta)}\frac{(\alpha\mu(\vartheta))^j}{j!}d\vartheta \\
&= \alpha \int_0^\infty e^{-\alpha x}\frac{(\alpha x)^j}{j!}f(x)dx \\
&= E[f((j+1)X_j/\alpha)],
\end{aligned}$$

where X_j is a gamma random variable with rate $j+1$ and inverse scale $j+1$. We split the above expectation into $X_j < 1/2$, $X_j \in [1/2, 3/2]$, and $X_j > 3/2$. The idea is that the third and the first expectations are small because X_j concentrates fast to 1, while the middle expectation ($X_j \in [1/2, 3/2]$) uses the fact that $f((j+1)X_j/\alpha) \approx f((j+1)/\alpha)$. More precisely, using Stirling's approximation, for every $\epsilon > 0$ there exists $c > 0$ such that

$$\begin{aligned}
E[f((j+1)X_j/\alpha)\mathbb{1}_{X_j<1/2}] &= \frac{(j+1)^{j+1}}{\Gamma(j+1)}\int_0^{1/2}f((j+1)x/\alpha)x^j e^{-(j+1)x}dx \\
&\lesssim \sqrt{j}\int_0^{1/2}\left(1 + \left(\frac{(j+1)x}{\alpha}\right)^{-1-\tilde{\sigma}-\epsilon}\right)e^{-j(x-\log x-1)}dx \\
&\lesssim e^{-cj}\left(1 + \left(\frac{j}{\alpha}\right)^{-1-\tilde{\sigma}-\epsilon}\right) = o(1),
\end{aligned}$$

since $\alpha/j = o(e^{cj})$ for any $c > 0$. The expectation over $X_j > 3/2$ is treated similarly. We now study the expectation over $[1/2, 3/2]$. We have that if $j/\alpha \to \infty$, then uniformly in $x \in [1/2, 3/2]$, under Assumption 3,

$$\left|\frac{f((j+1)x/\alpha)}{f((j+1)/\alpha)} - x^{-1-\tau}\right| = o(1),$$

and similarly when $j/\alpha \to 0$, with τ replaced by $\tilde{\sigma}$; if $j/\alpha \to c_0 \in (0, \infty)$, then uniformly in $x \in [1/2, 3/2]$,

$$\left|\frac{f((j+1)x/\alpha)}{f((j+1)/\alpha)} - \frac{f(c_0 x)}{f(c_0)}\right| = o(1).$$

Moreover, since X_j converges almost surely to 1, we finally obtain that

$$E\left[\frac{f((j+1)X_j/\alpha)}{f((j+1)/\alpha)}\mathbb{1}_{X_j\in[1/2,3/2]}\right] \to 1,$$

which terminates the proof. □

3.2. Proof of Theorem 2

The proof follows similarly to that of [38, Theorem 6.1], by bounding the variance. The paper [38] showed that $\mathrm{var}(N_\alpha) = o(E(N_\alpha)^2)$ and $\mathrm{var}(N_{\alpha,j}) = o(E(N_{\alpha,j})^2)$ and used this result to prove that (22) holds in probability; we need a slightly tighter bound on the variances to obtain the almost sure convergence. This is stated in the next two propositions.

Proposition 3. *Let N_α be the number of nodes. We have*

$$\mathrm{var}(N_\alpha) = E(N_\alpha) + 2\alpha^2 \int_{\mathbb{R}_+} \mu(x)\{1 - W(x,x)\}e^{-\alpha\mu(x)}dx$$
$$+ \alpha^2 \int_{\mathbb{R}_+^2} \{1 - W(x,y)\}\{1 - W(x,x)\}\{1 - W(y,y)\}$$
$$\left\{e^{\alpha\nu(x,y)} - 1 + W(x,y)\right\} e^{-\alpha\mu(x) - \alpha\mu(y)}dxdy. \tag{25}$$

Under Assumptions 1 and 2, with $\sigma \in [0, 1]$, slowly varying function ℓ, and positive scalar a satisfying (17), we have

$$\mathrm{var}(N_\alpha) = O\{\alpha^{3+2\sigma-2a}\ell_\sigma(\alpha)^2\}, \tag{26}$$

where the slowly varying functions ℓ_σ are defined in Equation (20). Additionally, under Assumptions 1 and 4, we have, for any $\sigma \in [0, 1]$ and any slowly varying function ℓ,

$$\mathrm{var}(N_\alpha) \asymp \alpha^{1+2\sigma}\ell_\sigma^2(\alpha). \tag{27}$$

Sketch of the proof. We give here the ideas behind the proof, deferring its completion to Section A.1 of the supplementary material [13]. Equation (25) is immediately obtained using the Slivnyak–Mecke and Campbell theorems. Applying the inequality $e^x - 1 \le xe^x$ and Lemmas B.2 and B.6 to the right-hand side of Equation (25), we obtain the upper bound of Equation (26). Finally, if Assumption 4 holds, then Assumption 2 holds as well with $a = 1$. Combining this with Assumption 1, we can therefore specialise the upper bound of Equation (26) to the case $a = 1 : O(\alpha^{1+2\sigma}\ell_\sigma^2(\alpha))$. The lower bound with the same order is found using the inequality $e^x - 1 \ge x$ and Lemmas B.2 and B.3. □

Proposition 3 and Theorem 1 imply in particular that, under Assumptions 1 and 2,

$$\mathrm{var}(N_\alpha) = O\{E(N_\alpha)^2\alpha^{-\kappa}\}$$

for some $\kappa > 0$. Here N_α is a positive, monotone increasing stochastic process. Using Lemma B.1 in the appendix, we obtain that $N_\alpha \sim E(N_\alpha)$ almost surely as α tends to ∞.

Proposition 4. *Let $N_{\alpha,j}$ be the number of nodes of degree j. Then, under Assumptions 1 and 2, with $\sigma \in [0, 1]$, slowly varying function ℓ, and positive scalar a satisfying (17), we have*

$$\mathrm{var}(N_{\alpha,j}) = O\{\alpha^{3+2\sigma-2a}\ell_\sigma(\alpha)^2\}.$$

where the slowly varying functions ℓ_σ are defined in Equation (20). In the case $\sigma = 0$ and $a = 1$, we have the stronger result

$$\mathrm{var}(N_{\alpha,j}) = o\{\alpha\ell(\alpha)^2\}.$$

Sketch of the proof. While the complete proof of Proposition 4 is given in Section A.2 in the supplementary material [13], we explain here its main passages. We start by evaluating

the expectation of $N_{\alpha,j}^2$ and $N_{\alpha,j}$ conditional on the unit-rate Poisson random measure $M = \sum_i \delta_{(\theta_i, \vartheta_i)}$:

$$E\left(N_{\alpha,j}^2 \mid M\right) - E\left(N_{\alpha,j} \mid M\right)$$

$$= \sum_{i_1 \neq i_2} \mathbb{1}_{\theta_{i_1} \leq \alpha} \mathbb{1}_{\theta_{i_2} \leq \alpha} \mathbb{P}\left\{ \sum_k \mathbb{1}_{\theta_k \leq \alpha} Z_{i_1 k} = j \text{ and } \sum_k \mathbb{1}_{\theta_k \leq \alpha} Z_{i_2, k} = j \mid M \right\}$$

$$= \sum_{b \in \{0,1\}^3} \sum_{j_1 = 0}^{j} \sum_{i_1 \neq i_2} \mathbb{1}_{\theta_{i_1} \leq \alpha} \mathbb{1}_{\theta_{i_2} \leq \alpha}$$

$$\times \mathbb{P}\left\{ \sum_k \mathbb{1}_{\theta_k \leq \alpha} Z_{i_1 k} = j \text{ and } \sum_k \mathbb{1}_{\theta_k \leq \alpha} Z_{i_2, k} = j \text{ and } \sum_k \mathbb{1}_{\theta_k \leq \alpha} Z_{i_1 k} Z_{i_2 k} = j - j_1 \right.$$

$$\left. \text{and } Z_{i_1 i_1} = b_{11}, Z_{i_1 i_2} = b_{12}, Z_{i_2 i_2} = b_{22} \mid M \right\},$$

where $b = (b_{11}, b_{12}, b_{22}) \in \{0, 1\}^3$. We then use the Slivnyak–Mecke theorem to obtain $E(N_{\alpha,j}^2) - E(N_{\alpha,j})$, which can be bounded by a sum of terms of the form

$$\alpha^2 \int_{\mathbb{R}^2} [\alpha \mu(x)]^{k_1} [\alpha \mu(y)]^{k_2} (\alpha \nu(x, y))^r e^{-\alpha \mu(x) - \alpha \mu(y) + \alpha \nu(x,y)} dx dy \tag{28}$$

for $k_1, k_2, r \in \{0, \ldots, j\}$. For terms with $r \geq 1$, we use Lemma A.1 (enunciated and proved, using Lemmas B.2 and B.4, in Section A.2 of the supplementary material). The lemma states that, under Assumptions 1 and 2, the integral in (28) is in $O\left(\alpha^{r-2ar+2\sigma} \ell_\sigma^2(\alpha)\right)) = O\left(\alpha^{1-2a+2\sigma} \ell_\sigma^2(\alpha)\right))$ for any $r \geq 1$, $k_1, k_2 \geq 0$. For terms with $r = 0$ in (28), we use the inequality $e^x \leq 1 + x e^x$, the Cauchy–Schwarz inequality, and Lemma B.4 to show that these terms are in $O\{\alpha^{3+2\sigma-2a} \ell_\sigma^2(\alpha)\}$, which completes the proof. $\qquad \square$

Define $\widetilde{N}_{\alpha,j} = \sum_{k \geq j} N_{\alpha,k}$, the number of nodes of degree at least j. Note that $\widetilde{N}_{\alpha,j}$ is a positive, monotone increasing stochastic process in α, with $\widetilde{N}_{\alpha,j} = N_\alpha - \sum_{k=1}^{j-1} N_{\alpha,k}$. We then have, using the Cauchy–Schwarz and Jensen inequalities, that

$$E\left(\widetilde{N}_{\alpha,j}\right) = E(N_\alpha) - \sum_{k=1}^{j-1} E(N_{\alpha,k}), \quad \mathrm{var}\left(\widetilde{N}_{\alpha,j}\right) \leq j \left\{ \mathrm{var}(N_\alpha) + \sum_{k=1}^{j-1} \mathrm{var}(N_{\alpha,k}) \right\}.$$

Consider first the case $\sigma \in [0, 1)$. Since Theorem 1 implies, for $j \geq 2$, $\alpha^{1+\sigma} \ell(\alpha) \lesssim E\left(\widetilde{N}_{\alpha,j}\right)$ as α goes to infinity, using Propositions 3 and 4, we obtain $\mathrm{var}\left(\widetilde{N}_{\alpha,j}\right) = O\{\alpha^{-\tau} E\left(\widetilde{N}_{\alpha,j}\right)^2\}$ for some $\tau > 0$. Combining this with Lemma B.1 leads to $\widetilde{N}_{\alpha,j} \sim E\left(\widetilde{N}_{\alpha,j}\right)$ almost surely as α goes to infinity.

The almost sure results for $N_{\alpha,j}$ then follow from the fact that, for all $j \geq 2$, $E\left(\widetilde{N}_{\alpha,j}\right) \asymp E(N_\alpha)$ if $\sigma \in (0, 1)$, $E\left(\widetilde{N}_{\alpha,j}\right) \sim E(N_\alpha)$ if $\sigma = 0$, and $E\left(\widetilde{N}_{\alpha,j}\right) = o\{E(N_\alpha)\}$ if $\sigma = 1$.

3.3. Asymptotic behaviour of the clustering coefficients

The following proposition is a direct corollary of [9, Proposition 56], which showed the almost sure convergence of subgraph counts in graphex processes.

Proposition 5. (Global clustering coefficient [9].) *Assume $\int_0^\infty \mu(x)^2 dx < \infty$. Recall that T_α and A_α are respectively the number of triangles and the number of adjacent edges in the graph of size α. We have*

$$T_\alpha \sim E(T_\alpha) = \frac{\alpha^3}{6} \int_{\mathbb{R}_+^3} W(x, y)W(x, z)W(y, z)dxdydz,$$

$$A_\alpha \sim E(A_\alpha) = \frac{\alpha^3}{2} \int_0^\infty \mu(x)^2 dx$$

almost surely as $\alpha \to \infty$. Therefore, if $\int_0^\infty \mu(x)^2 dx > 0$, the global clustering coefficient defined in Equation (11) converges to a constant

$$C_\alpha^{(g)} \to \frac{\int_{\mathbb{R}_+^3} W(x, y)W(x, z)W(y, z)dxdydz}{\int_0^\infty \mu(x)^2 dx} \quad \text{almost surely as } \alpha \to \infty.$$

Note that if μ is monotone decreasing, as $\overline{W} < \infty$, we necessarily have $\int_a^\infty \mu(x)^2 dx < \infty$ for any $a > 0$. Hence the condition $\int_0^\infty \mu(x)^2 dx < \infty$ in Proposition 5 requires additional assumptions on the behaviour of μ at 0 (or equivalently the behaviour of μ^{-1} at ∞), which drives the behaviour of large-degree nodes. If the graph is dense, μ is bounded and thus $\int_0^\infty \mu(x)^2 dx < \infty$.

Proposition 6. (Local clustering coefficient.) *Assume Assumptions 1 and 2 hold with $\sigma \in (0, 1)$. Assume additionally that*

$$\lim_{x \to \infty} \frac{\int_{\mathbb{R}_+^2} W(x, y)W(x, z)W(y, z)dydz}{\mu(x)^2} \to b \tag{29}$$

for some $b \in [0, 1]$. Then the local clustering coefficients converge in probabiltiy as $\alpha \to \infty$:

$$C_{\alpha,j}^{(\ell)} \to b \quad \forall j \geq 2.$$

If $b > 0$, the above result holds almost surely, and the average local clustering coefficient satisfies

$$\lim_{\alpha \to \infty} \overline{C}_\alpha^{(\ell)} \to b, \quad \text{almost surely.}$$

In general,

$$\lim_{x \to \infty} \frac{1}{\mu(x)^2} \int W(x, y)W(x, z)W(y, z)dydz \neq \frac{\int W(x, y)W(x, z)W(y, z)dxdydz}{\int \mu(x)^2 dx},$$

and the global clustering and local clustering coefficients converge to different limits. A notable exception is the separable case where $W(x, y) = \mu(x)\mu(y)/\overline{W}$, since in this case

$$\int W(x, y)W(x, z)W(y, z)dydz = \overline{W}^{-3}\mu(x)^2 \left(\int \mu(y)^2 dy \right)^2, \quad b = \frac{\left(\int \mu(y)^2 dy \right)^2}{\overline{W}^3}$$

and

$$\int W(x, y)W(x, z)W(y, z)dydzdx = \overline{W}^{-3} \left(\int \mu(y)^2 dy \right)^3.$$

Sketch of the proof. Full details are given in Appendix A.2; here we give only a sketch of the proof, which is similar to that of Theorem 2. We have

$$C_{\alpha,j}^{(\ell)} = \frac{2R_{\alpha,j}}{j(j-1)N_{\alpha,j}}, \quad \text{where} \quad R_{\alpha,j} = \sum_i T_{\alpha,i} \mathbb{1}_{D_{\alpha,i}=j}.$$

Here $R_{\alpha,j}$ corresponds to the number of triangles having a node of degree j as a vertex; triangles having $k \leq 3$ degree-j nodes as vertices are counted k times.

We obtain an asymptotic expression for $E(R_{\alpha,j})$, and show that $\text{var}(R_{\alpha,j}) = O(\alpha^{1-2a}[E(R_{\alpha,j})]^2)$. We then prove that $R_{\alpha,j}/E(R_{\alpha,j})$ goes to 1 almost surely. The latter is obtained by proving that $R_{\alpha,j}$ is nearly monotone increasing by constructing an increasing sequence α_n going to infinity such that $E(R_{\alpha_n,j})/E(R_{\alpha_{n+1},j})$ goes to 1 and such that for all $\alpha \in (\alpha_n, \alpha_{n+1})$,

$$R_{\alpha_n,j} - \tilde{R}_{n,j} \leq R_{\alpha,j} \leq R_{\alpha_{n+1},j} + \tilde{R}_{n,j}, \qquad \tilde{R}_{n,j} = o(E(R_{\alpha_n,j})).$$

Roughly speaking, $\tilde{R}_{n,j}$ (defined in Equation (60)) corresponds to the sum of the number of triangles $T_{\alpha_{n+1}i}$, over the set of nodes i such that $D_{n,i} \leq j$ and i has at least one connection with some i' such that $\theta_{i'} \in (\alpha_n, \alpha_{n+1})$. The result for the local clustering coefficient then follows from Toeplitz's lemma (see e.g. [27, p. 250]).

4. Central limit theorems

We now present central limit theorems (CLTs) for subgraph counts (numbers of edges, triangles, etc.) and for the number of nodes N_α. Subgraph counts can be expressed as U-statistics of Poisson random measures (up to an asymptotically negligible term). A CLT then follows rather directly from CLT on U-statistics of Poisson random measures [35].

Obtaining a CLT for quantities like N_α is more challenging, since these cannot be reduced to U-statistics. In this section we prove the CLT for N_α; we separate the dense and sparse cases because the techniques of the respective proofs are very different. The proof of the sparse case requires additional assumptions and is much more involved. We believe that the same technique of proof can be used for other quantities of interest, such as the number $N_{\alpha,j}$ of nodes of degree j, with more tedious computations.

4.1. CLT for subgraph counts

4.1.1. *Statement of the result.* Let F be a given subgraph which has neither isolated vertices nor loops. Denote by $|F|$ the number of nodes, $\{1, \cdots, |F|\}$ the set of vertices, and $e(F)$ the set of edges. Let $N_\alpha^{(F)}$ be the number of subgraphs F in the graph $\mathcal{G}_\alpha$:

$$N_\alpha^{(F)} = k_{(F)} \sum_{(v_1, \cdots, v_{|F|})}^{\neq} \prod_{(i,j) \in e(F)} Z_{v_i, v_j} \mathbb{1}_{\theta_{v_i} \leq \alpha} \mathbb{1}_{\theta_{v_j} \leq \alpha},$$

where $k_{(F)}$ is a constant accounting for the multiple counts of F, which we can omit in the rest of the discussion since it does not depend on α. Note that this statistic covers the number of edges (excluding loops) if $|F| = 2$ and the number of triangles if $|F| = 3$ and $e(F) = \{(1, 2), (1, 3), (2, 3)\}$. It is known in the graph literature as the number of injective adjacency maps from the vertex set of F to the vertex set of $\mathcal{G}_\alpha$; see [9, Section 2.5].

Proposition 7. *Let F be a subgraph without loops or isolated vertices. Assume that $\int_0^\infty \mu(x)^{2|F|-2}dx < \infty$. Then*

$$\frac{N_\alpha^{(F)} - E\big(N_\alpha^{(F)}\big)}{\sqrt{\mathrm{var}\big(N_\alpha^{(F)}\big)}} \to \mathcal{N}(0, 1), \tag{30}$$

as α goes to infinity, where

$$E\big(N_\alpha^{(F)}\big) = k_{(F)}\alpha^{|F|} \int_{\mathbb{R}_+^{|F|}} \prod_{(i,j)\in e(F)} W(x_i, x_j)dx_1 \cdots dx_{|F|} < \infty \tag{31}$$

and

$$\mathrm{var}\big(N_\alpha^{(F)}\big) \sim c_F\alpha^{2|F|-1}$$

for some positive constant c_F that depends only on F.

Remark 3. If the graph is dense, then μ is a bounded function with bounded support and therefore $\int_0^\infty \mu(x)^p dx < \infty$ for any p. In the sparse case, if μ is monotone, we necessarily have $\int_a^\infty \mu(x)^p dx < \infty$ for any $p > 1$. The condition $\int_0^\infty \mu(x)^{2|F|-2}dx < \infty$ therefore requires additional assumptions on the behaviour of μ at 0, which drives the behaviour of large-degree nodes.

4.1.2. *Proof.* Recall that $M = \sum_i \delta_{(\theta_i, \vartheta_i)}$. The main idea of the proof is to use the decomposition

$$N_\alpha^{(F)} - E\big(N_\alpha^{(F)}\big) = E\big(N_\alpha^{(F)}|M\big) - E\big(N_\alpha^{(F)}\big) + N_\alpha^{(F)} - E\big(N_\alpha^{(F)}|M\big), \tag{32}$$

and to show that $E\big(N_\alpha^{(F)}|M\big)$ is a geometric U-statistic of a Poisson process, for which a CLT has been derived by [35].

In this section, denote by $K = |F| \geq 2$ the number of nodes of the subgraph F. The subgraph counts are

$$N_\alpha^{(F)} = k_{(F)} \sum_{(v_1,\cdots,v_K)}^{\neq} \left(\prod_{k=1}^K \mathbb{1}_{\theta_{v_k}\leq\alpha}\right) \frac{1}{|\mathbb{S}_K|} \sum_{\pi\in\mathbb{S}_K} \prod_{(i,j)\in e(F)} Z_{v_{\pi_i},v_{\pi_j}},$$

where $\mathbb{S}_K$ denotes the set of permutations of $\{1, \ldots, K\}$.

Using the extended Slivnyak–Mecke theorem, we have

$$E\big(N_\alpha^{(F)}\big) = k_{(F)}\alpha^K \int_{\mathbb{R}_+^K} \prod_{(i,j)\in e(F)} W(x_i, x_j)dx_1 \cdots dx_K. \tag{33}$$

As $\int_0^\infty \mu(x)^{K-1}dx < \infty$, [9, Lemma 62] implies that $E\big(N_\alpha^{(F)}\big) < \infty$. For any $K \geq 2$, define the symmetric function

$$f(x_1, \ldots, x_K) = \frac{1}{|\mathbb{S}_K|} \sum_{\pi\in\mathbb{S}_K} \prod_{(i,j)\in e(F)} W\big(x_{\pi_i}, x_{\pi_j}\big);$$

using the condition (3) and the fact that $\int_0^\infty \mu(x)^{K-1}dx < \infty$, it satisfies $0 < \int_{\mathbb{R}_+^K} f(x_1, \ldots, x_K)dx_1 \ldots dx_K < \infty$.

We state the following useful lemma.

Lemma 1. *The function f satisfies, for all $x_K \geq 0$,*

$$g(x_K) := \int_{\mathbb{R}_+^{K-1}} f(x_1, \ldots, x_{K-1}, x_K) dx_1 \ldots dx_{K-1} \leq C_0 \max\left(\mu(x_K), \mu(x_K)^{K-1}\right)$$

for some constant C_0.

Proof. Let $\pi \in \mathbb{S}_K$ and $r_K \in \{1, \ldots, K\}$ be such that $\pi_{r_K} = K$. Denote by $S \subseteq \{1, \ldots, K-1\}$ the set of indices i such that $(i, r_K) \in e(F)$ and i has no other connections in F. Then

$$\int_{\mathbb{R}_+^{K-1}} \prod_{(i,j) \in e(F)} W(x_{\pi_i}, x_{\pi_j}) dx_1 \ldots dx_{K-1} \leq C_1 \int_{\mathbb{R}_+^{|S|}} \left[\prod_{i \in S} W(x_{\pi_i}, x_K) dx_i\right]$$

$$= C_1 \mu(x_K)^{|S|} \leq C_1 \max\left(\mu(x_K), \mu(x_K)^{K-1}\right)$$

for some constant C_1. $\qquad\square$

It follows from Lemma 1 and from the fact that $\int_0^\infty \mu(x) dx < \infty$ that, if $\int_0^\infty \mu(x)^{2K-2} dx < \infty$, then

$$\int_0^\infty \left(\int_{\mathbb{R}_+^{K-1}} f(x_1, \ldots, x_{K-1}, y) dx_1 \ldots dx_{K-1}\right)^2 dy < \infty.$$

We are now ready to derive the asymptotic expression for the variance of $N_\alpha^{(F)}$. Using the extended Slivnyak–Mecke theorem again,

$$E\left((N_\alpha^{(F)})^2\right) = E\left(E\left((N_\alpha^{(F)})^2 \mid M\right)\right)$$

$$= k_{(F)}^2 E\left(\sum_{\left(v_1, \cdots, v_K, v_1', \ldots, v_K'\right)}^{\neq} f\left(\vartheta_{v_1}, \ldots, \vartheta_{v_K}\right) f\left(\vartheta_{v_1'}, \ldots, \vartheta_{v_K'}\right) \prod_{k=1}^K \mathbb{1}_{\theta_{v_k} \leq \alpha} \mathbb{1}_{\theta_{v_k'} \leq \alpha}\right)$$

$$+ k_{(F)}^2 K^2 E\left(\sum_{\substack{(v_1, \cdots, v_K, \\ v_1', \ldots, v_{K-1}')}}^{\neq} f\left(\vartheta_{v_1}, \ldots, \vartheta_{v_K}\right) f\left(\vartheta_{v_1'}, \ldots, \vartheta_{v_{K-1}'}, \vartheta_{v_K}\right) \mathbb{1}_{\theta_{v_K} \leq \alpha} \prod_{k=1}^{K-1} \mathbb{1}_{\theta_{v_k} \leq \alpha} \mathbb{1}_{\theta_{v_k'} \leq \alpha}\right)$$

$$+ O\left(\alpha^{2K-2}\right)$$

$$= k_{(F)}^2 K^2 \alpha^{2K-1} \int_{\mathbb{R}_+^{2K-1}} f\left(x_1, \ldots, x_K\right) f\left(x_1', \ldots, x_{K-1}', x_K\right) dx_1, \ldots dx_K dx_1' \ldots dx_{K-1}'$$

$$+ E\left(N_\alpha^{(F)}\right)^2 + O\left(\alpha^{2K-2}\right).$$

It follows that

$$\mathrm{var}\left(N_\alpha^{(F)}\right) \sim k_{(F)}^2 K^2 \alpha^{2K-1} \sigma_F^2$$

as α tends to infinity, where

$$\sigma_F^2 = \int_0^\infty \left(\int_{\mathbb{R}_+^{K-1}} f(x_1, \ldots, x_{K-1}, y) dx_1 \ldots dx_{K-1} \right)^2 dy < \infty.$$

We now prove the CLT. The first term of the right-hand side of Equation (32) takes the form

$$E\big(N_\alpha^{(F)} | M\big) = k_{(F)} \sum_{(v_1, \cdots, v_K)}^{\neq} f\big(\vartheta_{v_1}, \ldots, \vartheta_{v_K}\big) \prod_{i=1}^K \mathbb{1}_{\theta_{v_i} \leq \alpha}. \tag{34}$$

By the superposition property of Poisson random measures, we have

$$E\big(N_\alpha^{(F)} | M\big) \stackrel{d}{=} k_{(F)} \sum_{(v_1, \cdots, v_K)}^{\neq} f\big(\widetilde{\vartheta}_{v_1}, \ldots, \widetilde{\vartheta}_{v_K}\big) \prod_{i=1}^K \mathbb{1}_{\widetilde{\theta}_{v_i} \leq 1},$$

where the right-hand side is a geometric U-statistic [35, Definition 5.1] of the Poisson point process $\big\{\big(\widetilde{\theta}_i, \widetilde{\vartheta}_i\big)_{i \geq 1}\big\}$ with mean measure $\alpha d\widetilde{\theta} d\widetilde{\vartheta}$ on $[0, 1] \times \mathbb{R}_+$. Theorem 5.2 in [35] therefore implies that

$$\frac{E\big(N_\alpha^{(F)} | M\big) - E\big(N_\alpha^{(F)}\big)}{\sqrt{\mathrm{var}\big(E\big(N_\alpha^{(F)} | M\big)\big)}} \to \mathcal{N}(0, 1), \tag{35}$$

where $\mathrm{var}\big(E\big(N_\alpha^{(F)} | M\big)\big) \sim \mathrm{var}\big(N_\alpha^{(F)}\big) \sim k_{(F)}^2 |F|^2 \alpha^{2|F|-1} \sigma_F^2$. One can show similarly (proof omitted) that $\mathrm{var}\big(N_\alpha^{(F)} - E\big(N_\alpha^{(F)} | M\big)\big) = o\big(\alpha^{2|F|-1}\big)$. It follows from Equations (32) and (35) and the Chebyshev inequality that

$$\frac{N_\alpha^{(F)} - E\big(N_\alpha^{(F)}\big)}{\sqrt{\mathrm{var}\big(N_\alpha^{(F)}\big)}} \to \mathcal{N}(0, 1)$$

as α tends to infinity.

4.2. CLT for N_α (dense case)

4.2.1. *Statement of the result.* In the dense case, μ has bounded support. If it is monotone decreasing, then Assumption 1 is satisfied with $\sigma = 0$, and $\ell(t) = \sup\{x > 0 \mid \mu(x) > 0\}$ is constant. In this case a CLT applies, as described in the following theorem.

Theorem 3. (Dense case.) *Assume that Assumption 1 holds with $\sigma = 0$ and $\ell(t) = C \in (0, \infty)$, where $C = \sup\{x > 0 \mid \mu(x) > 0\}$ (dense case). Also assume that Assumption 2 holds with $a = 1$. Then*

$$\frac{N_\alpha - E(N_\alpha)}{\sqrt{\alpha C}} \to \mathcal{N}(0, 1). \tag{36}$$

Moreover, $E(N_\alpha) = \alpha C - m_{\alpha,0}$, where

$$m_{\alpha,0} = \alpha \int_0^C e^{-\alpha\mu(x)} (1 - W(x, x)) dx = o(\alpha). \tag{37}$$

The quantity $m_{\alpha,0}$ can be interpreted as the expected number of degree-0 nodes, and is finite in the dense case. As shown in the following examples, $m_{\alpha,0}$ can either diverge or converge to a constant as α tends to infinity.

Example 1. Consider $\mu(x) = \mathbb{1}_{x\in[0,1]}$, $\mu(x) = (1-x)^2\mathbb{1}_{x\in[0,1]}$, and $\mu(x) = (1-x)^3\mathbb{1}_{x\in[0,1]}$. We respectively have $m_{\alpha,0} \to 0$, $m_{\alpha,0} \sim \frac{\sqrt{\pi}}{2}\alpha^{1/2}$, and $m_{\alpha,0} \sim \Gamma(4/3)\alpha^{2/3}$.

The above CLT for N_α can be generalised to $\widetilde{N}_{\alpha,j} = \sum_{k\geq j} N_{\alpha,k}$, the number of nodes of degree at least j.

Theorem 4. *Assume that Assumption* 1 *holds with* $\sigma = 0$ *and* $\ell(t) = C \in (0, \infty)$, *where* $C = \sup\{x > 0 \mid \mu(x) > 0\}$ *(dense case). Also assume that Assumption* 2 *holds with* $a = 1$. *Then for any* $j \geq 1$,

$$\frac{\widetilde{N}_{\alpha,j} - E(\widetilde{N}_{\alpha,j})}{\sqrt{\alpha C}} \to \mathcal{N}(0, 1). \tag{38}$$

Moreover, $E(\widetilde{N}_{\alpha,1}) = E(N_\alpha) = \alpha C - m_{\alpha,0}$, *and for* $j \geq 2$, $E(\widetilde{N}_{\alpha,j}) = \alpha C - m_{\alpha,0} - \sum_{k=1}^{j-1} E(N_{\alpha,j})$, *where* $m_{\alpha,0}$ *is defined in Equation* (37) *and* $E(N_{\alpha,j})$ *is defined in Equation* (53). *Note that* $m_{\alpha,0} = o(\alpha)$, *and for any* $j \geq 1$, $E(N_{\alpha,j}) = o(\alpha)$.

4.2.2. Proof. Given a point (θ, ϑ) such that $\vartheta > C$, its degree is necessarily equal to zero, as $\mu(\vartheta) = 0$. Write

$$N_\alpha = Q_\alpha - N_{\alpha,0}, \quad \text{where} \quad Q_\alpha = \sum_i \mathbb{1}_{\theta_i \leq \alpha} \mathbb{1}_{\vartheta_i \leq C};$$

Q_α is the total number of nodes i with $\theta_i \leq \alpha$ that could have a connection (hence such that $\mu(\vartheta_i) > 0$), and

$$N_{\alpha,0} = \sum_i \mathbb{1}_{\theta_i \leq \alpha} \mathbb{1}_{\vartheta_i \leq C} \mathbb{1}_{D_{\alpha,i}=0}$$

is the set of nodes i with degree 0, but for which $\theta_i \leq \alpha$, $\mu(\vartheta_i) > 0$. In the dense regime, both Q_α and $N_{\alpha,0}$ are almost surely finite. Furthermore, $(Q_\alpha)_{\alpha\geq 0}$ is a homogeneous Poisson process with rate C. By the law of large numbers, $Q_\alpha \sim \alpha C \sim N_\alpha$ almost surely as α tends to infinity. Using Campbell's theorem, the Slivnyak–Mecke formula, and monotone convergence, we have $E(N_{\alpha,0}) = \alpha \int_0^C (1 - W(x, x))e^{-\alpha\mu(x)}dx = o(\alpha)$. We also have that

$$E(N_{\alpha,0}^2) - E(N_{\alpha,0}) = \alpha^2 \int_0^C \int_0^C (1 - W(x, x))(1 - W(y, y))(1 - W(x, y))$$
$$\times e^{-\alpha\mu(x)-\alpha\mu(y)+\alpha v(x,y)}dxdy.$$

Hence, using the inequality $e^x - 1 \leq xe^x$, we obtain

$$\text{var}(N_{\alpha,0}) = \alpha^2 \int_0^C \int_0^C (1 - W(x, x))(1 - W(y, y))(1 - W(x, y))e^{-\alpha\mu(x)-\alpha\mu(y)+\alpha v(x,y)}dxdy$$
$$- \alpha^2 \left(\int_0^C (1 - W(x, x))e^{-\alpha\mu(x)}dx\right)^2 + E(N_{\alpha,0})$$
$$\leq E(N_{\alpha,0}) + \alpha^3 \int_0^C \int_0^C v(x, y)e^{-\alpha\mu(x)-\alpha\mu(y)+\alpha v(x,y)}dxdy.$$

Using Lemma B.6 in the appendix and Assumption 2 with $a = 1$, we have

$$\int_0^C \int_0^C v(x, y)e^{-\alpha/2\mu(x)-\alpha/2\mu(y)}dxdy = o(\alpha^{-2}).$$

It follows that $\operatorname{var}(N_{\alpha,0}) = o(\alpha)$. This implies, by Chebyshev's inequality, the CLT for Poisson processes, and Slutsky's theorem, that

$$\frac{N_\alpha - E(N_\alpha)}{\sqrt{\alpha C}} = \frac{Q_\alpha - \alpha C}{\sqrt{\alpha C}} - \frac{N_{\alpha,0} - E(N_{\alpha,0})}{\sqrt{\alpha C}} \to \mathcal{N}(0, 1).$$

This concludes the proof of Theorem 3. The proof of Theorem 4 follows similarly. Note that the case $j = 1$ in Theorem 4 corresponds to Theorem 3. For any $j \geq 2$, $\widetilde{N}_{\alpha,j} = Q_\alpha - N_{\alpha,0} - \sum_{k=1}^{j-1} N_{\alpha,k}$. We have, using the Cauchy–Schwarz inequality and Proposition 4,

$$\operatorname{var}\left(N_{\alpha,0} + \sum_{k=1}^{j-1} N_{\alpha,k}\right) \leq j\left(\operatorname{var}(N_{\alpha,0}) + \sum_{k=1}^{j-1} \operatorname{var}(N_{\alpha,k})\right) = o(\alpha).$$

This implies

$$\frac{\widetilde{N}_{\alpha,j} - E(\widetilde{N}_{\alpha,j})}{\sqrt{\alpha C}} = \frac{Q_\alpha - \alpha C}{\sqrt{\alpha C}} - \frac{N_{\alpha,0} + \sum_{k=1}^{j-1} N_{\alpha,k} - E(N_{\alpha,0} + \sum_{k=1}^{j-1} N_{\alpha,k})}{\sqrt{\alpha C}} \to \mathcal{N}(0, 1).$$

4.3. CLT for N_α (sparse case)

4.3.1. *Statement of the result.* We now assume that we are in the sparse regime; that is, μ has unbounded support. We make the following additional assumption in order to prove the asymptotic normality. This holds when W is separable, as well as in the model of [12] under some moment conditions (see Section 6.5).

Assumption 5. *Assume that for any $j \leq 6$ and any $(x_1, \ldots, x_j) \in \mathbb{R}_+^j$,*

$$\int_0^\infty \prod_{i=1}^j W(x_i, y)dy \leq \prod_{i=1}^j L(x_i)\mu(x_i),$$

where L is a locally integrable, slowly varying function converging to a (strictly positive) constant, such that

$$\int_0^\infty L(x)\mu(x)dx < \infty.$$

We now state the CLT for N_α under the sparse regime. Recall that in this case, when Assumption 1 holds, we have either $\sigma = 0$ and $\ell(t) \to \infty$ or $\sigma \in (0, 1]$.

Theorem 5. (Sparse case.) *Assume that μ has unbounded support (sparse regime). Under Assumptions 1, 4, and 5, we have*

$$\frac{N_\alpha - E(N_\alpha)}{\sqrt{\operatorname{var}(N_\alpha)}} \to \mathcal{N}(0, 1).$$

Remark 4. As detailed in Proposition 3, under Assumptions 1 and 4, for any $\sigma \in [0, 1]$ and any slowly varying function ℓ we have $\operatorname{var}(N_\alpha) \asymp \alpha^{1+2\sigma} \ell_\sigma^2(\alpha)$, where the slowly varying function ℓ_σ is defined in Equation (20).

4.3.2. *Proof.* The proof uses the recent results of [24] on normal approximations of nonlinear functions of a Poisson random measure. We have the decomposition

$$N_\alpha - E(N_\alpha) = (N_\alpha - E(N_\alpha \mid M)) + (E(N_\alpha \mid M) - E(N_\alpha))$$
$$= (N_\alpha - E(N_\alpha \mid M)) + (M(h_\alpha) - E(N_\alpha)) + f_\alpha(M),$$

where

$$f_\alpha(M) = \sum_i \mathbb{1}_{\theta_i \le \alpha} \left[(1 - W(\vartheta_i, \vartheta_i)) e^{-\alpha\mu(\vartheta_i)} - e^{-M\left(g_{\alpha,\vartheta_i}\right)} \right]$$

is a nonlinear functional of the Poisson random measure M, and

$$M(h_\alpha) = \sum_i \mathbb{1}_{\theta_i \le \alpha} \left[1 - (1 - W(\vartheta_i, \vartheta_i)) e^{-\alpha\mu(\vartheta_i)} \right]$$

is a linear functional of M with $h_\alpha(\theta, \vartheta) = \mathbb{1}_{\theta \le \alpha} \left[1 - (1 - W(\vartheta, \vartheta)) e^{-\alpha\mu(\vartheta)} \right]$. Theorem 5 is a direct consequence of the following three propositions and of Slutsky's theorem.

Proposition 8. *Under Assumptions* 1 *and* 4*, we have*

$$N_\alpha - E(N_\alpha \mid M) = \begin{cases} O\left(\alpha^{1/2+\sigma/2}\ell_\sigma^{1/2}(\alpha)\right) & \text{if } \sigma \in [0, 1), \\ o\left(\alpha^{1/2}\ell^{1/2}(\alpha)\right) & \text{if } \sigma = 0, \end{cases} \quad \text{in probability;}$$

hence

$$\frac{N_\alpha - E(N_\alpha \mid M)}{\sqrt{\mathrm{var}(N_\alpha)}} \to 0 \text{ in probability.}$$

Proposition 9. *Under Assumptions* 1 *and* 4*, we have*

$$M(h_\alpha) - E(N_\alpha) = O\left(\alpha^{1/2+\sigma/2}\ell^{1/2}(\alpha)\right) \text{ in probability;}$$

hence, if μ has unbounded support,

$$\frac{M(h_\alpha) - E(N_\alpha)}{\sqrt{\mathrm{var}(N_\alpha)}} \to 0 \text{ in probability.}$$

The above two propositions are proved in Section B of the supplementary material [13].

Proposition 10. *Assume μ has unbounded support. Under Assumptions* 1*,* 4*, and* 5*, we have*

$$\frac{f_\alpha(M)}{\sqrt{\mathrm{var}(N_\alpha)}} \to \mathcal{N}(0, 1).$$

Sketch of the proof. To prove Proposition 10 we resort to [24, Theorem 1.1] on the normal approximation of nonlinear functionals of Poisson random measures. Define

$$F_\alpha = \frac{f_\alpha(M)}{\sqrt{v_\alpha}}, \tag{39}$$

where $v_\alpha = \mathrm{var}(f_\alpha(M)) \sim \mathrm{var}(N_\alpha) \asymp \alpha^{1+2\sigma}\ell_\sigma^2(\alpha)$. Note that $E(F_\alpha) = 0$ and $\mathrm{var}(F_\alpha) = 1$. Consider the difference operator $D_z F_\alpha$ defined by

$$D_z F_\alpha = \frac{1}{\sqrt{v_\alpha}}(f_\alpha(M + \delta_z) - f_\alpha(M)),$$

and also

$$D_{z_1,z_2}^2 F_\alpha = D_{z_2}(D_{z_1} F_\alpha) = D_{z_2}\left(\frac{1}{\sqrt{v_\alpha}}(f_\alpha(M+\delta_{z_1}) - f_\alpha(M))\right)$$

$$= \frac{1}{\sqrt{v_\alpha}}\left(f_\alpha(M+\delta_{z_1}+\delta_{z_2}) - f_\alpha(M+\delta_{z_1}) - f_\alpha(M+\delta_{z_2}) + f_\alpha(M)\right).$$

Define

$$\gamma_{\alpha,1} := 2\left(\int_{\mathbb{R}_+^6}\sqrt{\mathbb{E}(D_{z_1}F_\alpha)^2(D_{z_2}F_\alpha)^2}\sqrt{\mathbb{E}(D_{z_1,z_3}^2 F_\alpha)^2(D_{z_2,z_3}^2 F_\alpha)^2}\,dz_1 dz_2 dz_3\right)^{1/2},$$

$$\gamma_{\alpha,2} := \left(\int_{\mathbb{R}_+^6}\mathbb{E}\left[(D_{z_1,z_3}^2 F_\alpha)^2(D_{z_2,z_3}^2 F_\alpha)^2\right]dz_1 dz_2 dz_3\right)^{1/2},$$

$$\gamma_{\alpha,3} := \int_{\mathbb{R}_+^2}\mathbb{E}|D_z F_\alpha|^3 dz.$$

In Section B.3 of the supplementary material [13], we prove that under Assumptions 1, 4, and 5, we have $\gamma_{\alpha,1}$, $\gamma_{\alpha,2}$, $\gamma_{\alpha,3} \to 0$. The proof is rather lengthy, and makes repeated use of Hölder's inequality and of properties of integrals involving regularly varying functions (in particular Lemma B.5). An application of [24, Theorem 1.1] then implies that $F_\alpha \to \mathcal{N}(0,1)$. □

5. Related work and discussion

Veitch and Roy [38] proved that Equation (22) holds in probability, under slightly different assumptions: they assume that Assumption 2 holds with $a = 1$ and that μ is differentiable, with some conditions on the derivative, but do not make any assumption on the existence of σ or ℓ. We note that for all the examples considered in Section 6, Assumptions 1 and 2 are always satisfied, but Assumption 2 does not hold with $a = 1$ for the non-separable graphon function (40). Additionally, the differentiability condition does not hold for some standard graphon models, such as the stochastic block-model. Borgs *et al.* [9] proved, amongst other results, the almost sure convergence of the subgraph counts in graphex models (Theorem 156). For the subclass of graphon models defined by Equation (41), [12] provided a lower bound on the growth in the number of nodes, and therefore an upper bound on the sparsity rate, using assumptions of regular variation similar to Assumption 1. Applying the results derived in this section, we show in Section 6.5 that the bound is tight, and we derive additional asymptotic properties for this particular class.

As mentioned in the introduction, another class of (non-projective) models that can produce sparse graphs are sparse graphons [4, 5, 8, 41]. In particular, a number of authors have considered the sparse graphon model in which two nodes i and j in a graph of size n connect with probability $\rho_n W(U_i, U_j)$, where $W : [0,1]^2 \to [0,1]$ is the graphon function, measurable and symmetric, and $\rho_n \to 0$. Although such a model can capture sparsity, it has rather different properties from graphex models. For example, the global clustering coefficient for this sparse graphon model converges to 0, while the clustering coefficient converges to a positive constant, as shown in Proposition 5.

Also, graphex processes include as a special case dense vertex-exchangeable random graphs [1, 15, 19, 28], that is, models based on a graphon on [0, 1]. They also include as a special case

the class of graphon models over more general probability spaces [7]; see [9, p. 21] for more details. Some other classes of graphs, such as geometric graphs arising from Poisson processes in different spaces [34], cannot be cast in this framework.

6. Examples of sparse and dense models

We provide here some examples of the four different cases: dense, almost dense, sparse, and almost extremely sparse. We also show that the results of the previous section apply to the particular model studied by [12].

6.1. Dense graph

Let us consider the graphon function

$$W(x, y) = (1 - x)(1 - y) \, \mathbb{1}_{x \leq 1} \, \mathbb{1}_{y \leq 1}$$

which has bounded support. The corresponding marginal graphon function $\mu(x) = \mathbb{1}_{x \leq 1}(1 - x)/2$ has inverse $\mu^{-1}(x) = \ell(1/x)$, where $\ell(1/x) = (1 - 2x)\mathbb{1}_{x \leq 1/2}$ is slowly varying since $\ell(1/x) \to 1$. Assumptions 1 and 2 are satisfied, so by Theorem 2 and Corollary 1,

$$N_\alpha \sim \alpha, \qquad N_\alpha^{(e)} \sim \alpha^2/8, \qquad N_\alpha^{(e)} \sim N_\alpha^2/8, \qquad \frac{N_{\alpha,j}}{N_\alpha} \to 0, \quad j \geq 1,$$

almost surely as $\alpha \to \infty$. The function W is separable and $C_\alpha^{(g)} \to 4/9$.

6.2. Sparse, almost dense graph without power-law behaviour

Consider the graphon function

$$W(x, y) = e^{-x-y},$$

considered by [38], which has full support. The corresponding function $\mu(x) = e^{-x}$ has inverse $\mu^{-1}(x) = \ell(1/x) = \log(1/x)\mathbb{1}_{0<x<1}$, which is a slowly varying function. We have $\ell_0^*(x) = 1/\log(x)^2$. Assumptions 1 and 2 are satisfied, and

$$N_\alpha \sim \alpha \log(\alpha), \qquad N_\alpha^{(e)} \sim \alpha^2/2, \qquad N_\alpha^{(e)} \sim \frac{N_\alpha^2}{2\log(N_\alpha)^2}, \qquad \frac{N_{\alpha,j}}{N_\alpha} \to 0 \text{ for all } j = 1, 2, \ldots.$$

The function W is separable, and $C_\alpha^{(g)} \to 1/4$.

6.3. Sparse graphs with power-law behaviour

We consider two examples here, one separable and one non-separable. Interestingly, while the degree distributions in the two examples have similar power-law behaviours, the clustering properties are very different. In the first example, the local clustering coefficient converges to a strictly positive constant, while in the second example it converges to 0.

Separable example. First, consider the function

$$W(x, y) = (x + 1)^{-1/\sigma}(y + 1)^{-1/\sigma}$$

with $\sigma \in (0, 1)$. We have $\mu(x) = \sigma(x+1)^{-1/\sigma}/(1-\sigma)$, $\mu^{-1}(x) = x^{-\sigma}(1/\sigma - 1)^{-\sigma} - 1$, $\ell(t) \sim (1/\sigma - 1)^{-\sigma}$, and $\ell_\sigma^*(t) \sim \left\{ (1/\sigma - 1)^{-\sigma}\Gamma(1-\sigma) \right\}^{-2/(1+\sigma)}$. Assumptions 1 and 2 are satisfied. We have $N_\alpha \sim \alpha^{1+\sigma}\Gamma(1-\sigma)(1/\sigma - 1)^{-\sigma}$, $N_\alpha^{(e)} \sim \alpha^2\sigma^2/\{2(1-\sigma)^2\}$, and

$$N_\alpha^{(e)} \sim \frac{\sigma^2 \left\{ \Gamma(1-\sigma)\left(\frac{1}{\sigma} - 1\right)^{-\sigma} \right\}^{-\frac{2}{1+\sigma}}}{2(1-\sigma)^2} N_\alpha^{2/(1+\sigma)}, \qquad \frac{N_{\alpha,j}}{N_\alpha} \to \frac{\sigma\Gamma(j-\sigma)}{j!\Gamma(1-\sigma)}, \quad j \geq 1.$$

The function is separable, and for $\sigma \in (0, 1)$ we obtain

$$\lim_{\alpha \to \infty} C_\alpha^{(g)} = \left(\frac{1-\sigma}{2-\sigma}\right)^2 \text{ and } \lim_{\alpha \to \infty} C_{\alpha,j}^{(\ell)} = \left(\frac{1-\sigma}{2-\sigma}\right)^2 \text{ almost surely.}$$

Non-separable example. Consider now the non-separable function

$$W(x, y) = (x + y + 1)^{-1/\sigma - 1} \tag{40}$$

where $\sigma \in (0, 1)$. We have $\mu(x) = \sigma(x+1)^{-1/\sigma}$, $\mu^{-1}(x) = \sigma^\sigma x^{-\sigma} - 1$, $\ell(t) \sim \sigma^\sigma$, and $\ell_\sigma^*(t) \sim \{\sigma^\sigma\Gamma(1-\sigma)\}^{-2/(1+\sigma)}$. Assumptions 1 and 2 are satisfied as for all $(x, y) \in \mathbb{R}_+^2$,

$$W(x, y) \leq (x+1)^{-1/(2\sigma)-1/2}(y+1)^{-1/(2\sigma)-1/2} = \sigma^{-1-\sigma}\mu(x)^{\frac{1+\sigma}{2}}\mu(y)^{\frac{1+\sigma}{2}}.$$

We have $N_\alpha \sim \alpha^{1+\sigma}\Gamma(1-\sigma)\sigma^\sigma$, $N_\alpha^{(e)} \sim \alpha^2\sigma^2/\{2(1-\sigma)\}$, and

$$N_\alpha^{(e)} \sim \frac{\sigma^2 [\Gamma(1-\sigma)\sigma^\sigma]^{-\frac{2}{1+\sigma}}}{2(1-\sigma)} N_\alpha^{2/(1+\sigma)}, \qquad \frac{N_{\alpha,j}}{N_\alpha} \to \frac{\sigma\Gamma(j-\sigma)}{j!\Gamma(1-\sigma)}, \quad j \geq 1.$$

We have $\int \mu(x)^2 dx = \frac{\sigma^3}{2-\sigma}$. There is no analytical expression for $\int W(x, y)W(y, z)$ $W(x, z)dxdydz$, but this quantity can be evaluated numerically, and is non-zero, so the global clustering coefficient converges almost surely to a non-zero constant for any $\sigma \in (0, 1)$. For the local clustering coefficient, we have $\mu(x)^2 \sim \sigma^2 x^{-2/\sigma}$ as $x \to \infty$ and

$$\int W(x, y)W(x, z)W(y, z)dydz \leq x^{-2/\sigma - 2} \int (y + z + 1)dydz = o(\mu(x)^2).$$

Hence the local clustering coefficients $C_{\alpha,j}^{(\ell)}$ converge in probability to 0 for all j.

6.4. Almost extremely sparse graph

Consider the function

$$W(x, y) = \frac{1}{(x+1)(1+\log(1+x))^2} \frac{1}{(y+1)(1+\log(1+y))^2}.$$

We have $\overline{W} = 1$, $\mu(x) = (x+1)^{-1}(1+\log(1+x))^{-2}$, and, using properties of inverses of regularly varying functions, $\mu^{-1}(x) \sim x^{-1}\ell(1/x)$ as $x \to 0$, where $\ell(t) = \log(t)^{-2}$ is a slowly varying function. For $t > 1$ we have $\ell_1(t) = \int_t^\infty x^{-1}\ell(x)dx = 1/\log(t)$ and $\ell_1^*(t) \sim \log(t)/2$. Assumptions 1 and 2 are satisfied, and almost surely

$$N_\alpha^{(e)} \sim \alpha^2/2, \qquad N_\alpha \sim \frac{\alpha^2}{\log(\alpha)}, \qquad N_\alpha^{(e)} \sim \frac{1}{4}N_\alpha \log(N_\alpha),$$

$$\frac{N_{\alpha,1}}{N_\alpha} \to 1, \qquad \frac{N_{\alpha,j}}{N_\alpha} \to 0 \text{ for all } j \geq 2.$$

We have $\int \mu(x)^2 dx = \frac{1}{6}(2 + e\mathrm{Ei}(-1)) \simeq 0.24$, where Ei is the exponential integral; hence $C_\alpha^{(g)} \to 0.0576$ almost surely.

6.5. Model of [12]

The paper [12] studied a particular subclass of non-separable graphon models. This class is very flexible and allows one to span the whole range of sparsity and power-law behaviours described in Section 3. As shown by [12], efficient Monte Carlo algorithms can be developed for estimating the parameters of this class of models. Additionally, [11, Corollary 1.3] recently showed that this class is the limit of some sparse configuration models, providing further motivation for the study of their mathematical properties.

Let ρ be a Lévy measure on $(0, +\infty)$ and $\overline{\rho}(x) = \int_x^\infty \rho(dw)$ the corresponding tail Lévy intensity with generalised inverse $\overline{\rho}^{-1}(x) = \inf\{u > 0 | \overline{\rho}(u) < x\}$. The paper [12] introduced the model defined by

$$W(x, y) = \begin{cases} 1 - e^{-2\overline{\rho}^{-1}(x)\overline{\rho}^{-1}(y)}, & x \neq y, \\ 1 - e^{-\{\overline{\rho}^{-1}(x)\}^2}, & x = y. \end{cases} \tag{41}$$

The quantity $w = \overline{\rho}^{-1}(x)$ can be interpreted as the sociability of a node with parameter x. The larger this value, the more likely the node is to connect to other nodes. The tail Lévy intensity $\overline{\rho}$ is a monotone decreasing function; its behaviour at zero will control the behaviour of low-degree nodes, while its behaviour at infinity will control the behaviour of high-degree nodes.

The following proposition formalises this and shows how the results of Sections 3 and 4 apply to this model. Its proof is given in Section 6.6.

Proposition 11. *Consider the graphon function W defined by Equation* (41) *with Lévy measure ρ and tail Lévy intensity $\overline{\rho}$. Assume $m = \int_0^\infty w\rho(dw) < \infty$ and*

$$\overline{\rho}(x) \sim x^{-\sigma}\widetilde{\ell}(1/x) \, as \, x \to 0 \tag{42}$$

for some $\sigma \in [0, 1]$ and some slowly varying function $\widetilde{\ell}$. Then Equation (3) *and Assumptions* 1 *and* 2 *hold, with $a = 1$ and $\ell(x) = (2m)^\sigma \widetilde{\ell}(x)$. Proposition* 1, *Theorems* 1 *and* 2, *and Corollary* 1 *therefore hold. If $\int_0^\infty \psi(2w)^2 \rho(dw) < \infty$, where $\psi(t) = \int (1 - e^{-wt})\rho(dw)$ is the Laplace exponent, then the global clustering coefficient converges almost surely,*

$$\lim_{\alpha \to \infty} C_\alpha^{(g)} = \frac{\int_{\mathbb{R}_+^3} \left(1 - e^{-2xy}\right)\left(1 - e^{-2xz}\right)\left(1 - e^{-2yz}\right)\rho(dx)\rho(dy)\rho(dz)}{\int_0^\infty \psi(2w)^2 \rho(dw)},$$

and when $\sigma \in (0, 1)$, Proposition 6 *holds and for any $j \geq 2$ we have*

$$\lim_{\alpha \to \infty} C_{\alpha,j}^{(\ell)} = \lim_{\alpha \to \infty} \overline{C}_\alpha^{(\ell)} = 1 - \frac{\int_{\mathbb{R}_+^2} yze^{-2yz}\rho(dy)\rho(dz)}{m^2}$$

almost surely. For a given subgraph F, the CLT for the number of such subgraphs (Proposition 7*) holds if $\int \psi\left(2\overline{\rho}^{-1}(x)\right)^{2|F|-2} dx < \infty$. Under Assumption* 1, *this condition always holds if $\sigma = 0$; for $\sigma \in (0, 1]$, it holds if $\overline{\rho}(x) = O\left(x^{-(2|F|-2)\sigma-\epsilon}\right)$ as $x \to \infty$ for some $\epsilon > 0$. In this case, we have*

$$\frac{N_\alpha^{(F)} - E\left(N_\alpha^{(F)}\right)}{\sqrt{\mathrm{var}\left(N_\alpha^{(F)}\right)}} \to \mathcal{N}(0, 1). \tag{43}$$

Moreover, if $\int w^6 \rho(dw) < \infty$, *then Assumptions* 4 *and* 5 *also hold. It follows that Theorems* 3, 4, *and* 5 *apply, and for any* $\sigma \in [0, 1]$ *and any* ℓ,

$$\frac{N_\alpha - E(N_\alpha)}{\sqrt{\text{var}(N_\alpha)}} \to \mathcal{N}(0, 1). \tag{44}$$

Finally, assume $\sigma \in (0, 1)$ *and* $\widetilde{\ell}(t) = c > 0$. *If additionally*

$$\overline{\rho}(x) \sim c_0 x^{-\sigma\tau} \text{ as } x \to \infty \tag{45}$$

for some $\tau > 0$, $c_0 > 0$, *then Assumption* 3 *is also satisfied with* $\tau > 0$, $\ell_2(x) = \frac{c_0}{2^{\sigma\tau} c^\tau \Gamma(1-\sigma)^\tau}$, *and Proposition* 2 *applies; that is, for fixed* α,

$$E\big(N_{\alpha,j}\big) \sim \frac{\alpha^{1+\tau} \tau \ell_2(j)}{j^{1+\tau}} \text{ as } j \to \infty.$$

We consider below two specific choices of mean measures ρ. The two measures have similar properties for large graph size α, but different properties for large degrees j.

Generalised gamma measure. Let ρ be the generalised gamma measure

$$\rho(dw) = 1/\Gamma(1-\sigma_0) w^{-1-\sigma_0} e^{-\tau_0 w} dw \tag{46}$$

with $\tau_0 > 0$ and $\sigma_0 \in (-\infty, 1)$. The tail Lévy intensity satisfies

$$\overline{\rho}(x) \sim \begin{cases} \frac{1}{\Gamma(1-\sigma_0)\sigma_0} x^{-\sigma_0}, & \sigma_0 > 0, \\[2mm] \log(1/x), & \sigma_0 = 0, \\[2mm] -\frac{\tau_0^{\sigma_0}}{\sigma_0}, & \sigma_0 < 0, \end{cases}$$

as $x \to 0$. Then for $\sigma_0 \in (0, 1)$ (sparse with power law),

$$N_\alpha^{(e)} \asymp N_\alpha^{2/(1+\sigma_0)}, \qquad \frac{N_{\alpha,j}}{N_\alpha} \to \frac{\sigma_0 \Gamma(j-\sigma_0)}{j! \Gamma(1-\sigma_0)}, \quad j \geq 1.$$

For $\sigma_0 = 0$ (sparse, almost dense), we have $N_\alpha^{(e)} \asymp N_\alpha^2 / \log(N_\alpha)^2$ and $N_{\alpha,j}/N_\alpha \to 0$ for $j \geq 1$; for $\sigma_0 < 0$ (dense), we have $N_\alpha^{(e)} \asymp N_\alpha^2$ and $N_{\alpha,j}/N_\alpha \to 0$ for $j \geq 1$, almost surely, as α tends to infinity. The constants in the asymptotic results are omitted for simplicity of exposition, but they can also be obtained from the results of Section 3. We have $\int w^p \rho(dw) < \infty$ for all $p \geq 1$; hence the global clustering coefficient converges, and the CLT applies for the number of subgraphs and the number of nodes. Note that Equation (45) is not satisfied, as the Lévy

measure has exponentially decaying tails, and Proposition 2 does not apply. The asymptotic properties of this model are illustrated in Figure 2 for $\sigma_0 = 0.2$ and $\tau_0 = 2$ (sparse, power-law regime).

Generalised gamma-Pareto measure. Consider the generalised gamma-Pareto measure, introduced by [2, 3]:

$$\rho(dw) = \frac{1}{\Gamma(1-\sigma)} w^{-1-\sigma\tau} \gamma(\sigma(\tau-1), \beta w) dw,$$

where $\gamma(s, x) = \int_0^x u^{s-1} e^{-u} du$ is the lower incomplete gamma function, $c > 0$, $\tau > 1$, $\sigma \in (0, 1)$. The tail Lévy intensity satisfies

$$\overline{\rho}(x) \sim cx^{-\sigma} \text{ as } x \to 0,$$
$$\overline{\rho}(x) \sim c_0 x^{-\sigma\tau} \text{ as } x \to \infty,$$

where

$$c = \frac{\beta^{\sigma(\tau-1)}}{\sigma^2(\tau-1)\Gamma(1-\sigma)}, \qquad c_0 = \frac{\Gamma(\sigma(\tau-1))}{\sigma\tau\Gamma(1-\sigma)}.$$

It is regularly varying at both zero and infinity, and it satisfies (42) and (45). We therefore have, almost surely,

$$N_\alpha^{(e)} \asymp N_\alpha^{2/(1+\sigma)}, \qquad \frac{N_{\alpha,j}}{N_\alpha} \to \frac{\sigma_0\Gamma(j-\sigma)}{j!\Gamma(1-\sigma)}, \quad j \geq 1.$$

Proposition 2 applies, and for large-degree nodes,

$$E(N_{\alpha,j}) \sim \frac{\tau\alpha^{1+\tau} c_0}{2^{\sigma\tau} c^\tau \Gamma(1-\sigma)^\tau} \frac{1}{j^{1+\tau}} \text{ as } j \to \infty.$$

The global clustering coefficient converges if $\tau > 2$; the CLT applies for the number of subgraphs F if $\tau > 2|F| - 2$, and for the number of nodes if $\sigma\tau > 6$.

6.6. Proof of Proposition 11

The marginal graphon function is given by $\mu(x) = \psi(2\overline{\rho}^{-1}(x))$ where $\psi(t) = \int_0^\infty (1 - e^{-wt})\rho(dw)$ is the Laplace exponent. Its generalised inverse is given by $\mu^{-1}(x) = \overline{\rho}(\psi^{-1}(x)/2)$. The Laplace exponent satisfies $\psi(t) \sim mt$ as $t \to 0$. It therefore follows that μ^{-1} satisfies Assumption 1 with $\ell(x) = (2m)^\sigma \widetilde{\ell}(x)$. Ignoring loops, the model is of the form given by Equation (18) with $f(x) = 2m\overline{\rho}^{-1}(x)$. Assumption 2 is therefore satisfied. Regarding the global clustering coefficient, $\int \psi(2w)^2 \rho(dw) \leq 4 \int w^2 \rho(dw) < \infty$, so its limit is finite. For the local clustering coefficient, using dominated convergence and the inequality $\frac{1-e^{-2\overline{\rho}^{-1}(x)y}}{2\overline{\rho}^{-1}(x)} \leq y$, we obtain

$$\int W(x, y)W(y, z)W(x, z)dydz = \int \left(1 - e^{-2\overline{\rho}^{-1}(x)y}\right)\left(1 - e^{-2\overline{\rho}^{-1}(x)z}\right)\left(1 - e^{-2yz}\right)\rho(dy)\rho(dz)$$

$$\sim 4\overline{\rho}^{-1}(x)^2 \int yz\left(1 - e^{-2yz}\right)\rho(dy)\rho(dz).$$

Using the fact that $\mu(x) = \psi(2\overline{\rho}^{-1}(x)) \sim 2m\overline{\rho}^{-1}(x)$ as $x \to \infty$, we obtain the result. Finally, if $\overline{\rho}$ satisfies (42), then $\psi(t) \sim \Gamma(1-\sigma)\widetilde{\ell}(t)t^\sigma$ as $t \to \infty$. Using [6, Proposition 1.5.15], we have

$$\psi^{-1}(t) \sim \Gamma(1-\sigma)^{-1/\sigma}\widetilde{\ell}^{\#1/\sigma}\left(t^{1/\sigma}\right)t^{1/\sigma}$$

as $t \to \infty$, where $\widetilde{\ell}^\#$ is the de Bruijn conjugate of $\widetilde{\ell}$. We obtain $\psi^{-1}(t) = \ell_3(t^{1/\sigma})t^{\frac{1}{\sigma}}$, where ℓ_3 is a slowly varying function with $\ell_3(t^{1/\sigma}) \sim \widetilde{\ell}^{\#1/\sigma}(t^{1/\sigma})\Gamma(1-\sigma)^{-1/\sigma}$ as $t \to \infty$. We therefore have $\mu^{-1}(t) \sim c_0 2^{-\tau\sigma}\ell_3(t^{1/\sigma})^{\sigma\tau}t^\tau$ as $t \to \infty$. If $\widetilde{\ell}(t) = c$, then $\ell_3(t) = (c\Gamma(1-\sigma))^{-1/\sigma}$.

For the CLT for the number of subgraphs F to hold, we need $\int_0^\infty \mu(x)^{2|F|-2}dx < \infty$. As μ is monotone decreasing and integrable, we only need $\mu(x)^{2|F|-2} = \psi(2\overline{\rho}^{-1}(x))^{2|F|-2}$ to be integrable in a neighbourhood of 0. In the dense case, $\psi(t)$ is bounded, and the condition holds. If $\overline{\rho}$ satisfies (42), then $\psi(t) \sim \Gamma(1-\sigma)\widetilde{\ell}(t)t^\sigma$ as $t \to \infty$. For $\sigma \in (0,1]$ (sparse regime), the condition holds if $\overline{\rho}(x) = O(x^{-(2|F|-2)\sigma-\epsilon})$ as $x \to \infty$ for some $\epsilon > 0$.

We now check the assumptions for the CLT for the number of nodes. Noting again that $\mu(x) \sim 2m\overline{\rho}^{-1}(x)$ as $x \to \infty$, we have, using the inequality $1 - e^{-x} \leq x$,

$$v(x,y) = \int \left(1 - e^{-2\overline{\rho}^{-1}(x)w}\right)\left(1 - e^{-2\overline{\rho}^{-1}(y)w}\right)\rho(dw)$$
$$\leq L(x)L(y)\mu(x)\mu(y),$$

where

$$L(x) = 2\frac{\overline{\rho}^{-1}(x)}{\mu(x)}\sqrt{\int w^2\rho(dw)} \to \sqrt{\int w^2\rho(dw)/m}$$

as $x \to \infty$. Using now the inequality $1 - e^{-x} \geq xe^{-x}$, we have

$$v(x,y) \geq 4\overline{\rho}^{-1}(x)\overline{\rho}^{-1}(y)\int w^2 e^{-2(\overline{\rho}^{-1}(x)+\overline{\rho}^{-1}(y))w}\rho(dw).$$

As

$$\int w^2 e^{-2(\overline{\rho}^{-1}(x)+\overline{\rho}^{-1}(y))w}\rho(dw) \to \int w^2\rho(dw)$$

as $\min(x,y) \to \infty$, there exist $C_0 = 2\int w^2\rho(dw)$ and x_0 such that for all $x,y > x_0$, $v(x,y) \geq C_0\mu(x)\mu(y)$.

More generally, if $\int w^6\rho(dw) < \infty$, then for any $j \leq 6$,

$$\int_0^\infty \prod_{i=1}^j W(x_i,y)dy \leq \prod_{i=1}^j L(x_i)\mu(x_i),$$

where

$$L(x) = 2\frac{\overline{\rho}^{-1}(x)}{\mu(x)}\max\left(1, \max_{j=1,\dots,6}\int w^j\rho(dw)\right) \to \max\left(1, \max_{j=1,\dots,6}\int w^j\rho(dw)\right)/m$$

as $x \to \infty$. Note also that

$$\int L(x)\mu(x)dx = 2\max\left(1, \max_{j=1,\dots,6}\int w^j\rho(dw)\right)\int w\rho(dw) < \infty.$$

7. Sparse and dense models with local structure

In this section, we develop a class of models which allows us to control separately the local structure, for example the presence of communities or particular subgraphs, and the global sparsity/power-law properties. The class of models introduced can be used as a way of sparsifying any dense graphon model.

7.1. Statement of the results

By Kallenberg's representation theorem, any exchangeable point process can be represented by Equation (2). However, it may be more suitable to use a different formulation where the function W is defined on a general space, not necessarily $\mathbb{R}_+^2$, as discussed by [9]. Such a construction may lead to more interpretable parameters and easier inference methods. Indeed, a few sparse vertex-exchangeable models, such as the models of [18] or [37], are written in such a way that it is not straightforward to express them in the form given by (2).

In this section we show that the above results easily extend to models expressed in the following way. Let F be a probability space. Writing $\vartheta = (u, v) \in \mathbb{R}_+ \times F$, let $\xi(d\vartheta) = du \, G(dv)$ where G is some probability distribution on F. Consider models expressed as in (1) with

$$Z_{ij} \mid (\theta_k, \vartheta_k)_{k=1,2,\dots} \sim \text{Bernoulli}\{W(\vartheta_i, \vartheta_j)\}, \quad W : (\mathbb{R}_+ \times F)^2 \to [0, 1], \qquad (47)$$

where $(\theta_k, \vartheta_k)_{k=1,2,\infty}$ are the points of a Poisson point process with mean measure $d\theta \, \xi(d\vartheta)$ on $\mathbb{R}_+ \times (\mathbb{R}_+ \times F)$. Let us assume additionally that the function W factors in the following way:

$$W((u_i, v_i), (u_j, v_j)) = \omega(v_i, v_j)\eta(u_i, u_j), \qquad (48)$$

where $\omega : F \times F \to [0, 1]$ and the function $\eta : \mathbb{R}_+ \times \mathbb{R}_+ \to [0, 1]$ is integrable. In this model, ω can capture the local structure, as in the classical dense graphon, and η the sparsity behaviour of the graph. Let $\mu_\eta(u) = \int_0^\infty \eta(u, u')du'$, $\mu_\omega(v) = \int_F \omega(v, v')G(dv')$, and $\nu_\eta(x, y) = \int_{\mathbb{R}_+^2} \eta(x, z)\eta(y, z)dz$. The results presented in Section 3 remain valid when μ_η and ν_η satisfy Assumptions 1 and 2. The proof of Proposition 12 is given in Section 7.2.

Proposition 12. *Consider the model defined by Equations* (47) *and* (48) *and assume that the functions μ_η and ν_η satisfy Assumptions* 1 *and* 2. *Then the conclusions of Proposition* 1 *hold, and so do the conclusions of Theorems* 1 *and* 2, *with $\ell(\alpha)$ and $\ell_1(\alpha)$ replaced respectively by*

$$\tilde{\ell}(\alpha) = \ell(\alpha) \int_F \mu_\omega(v)^\sigma G(dv), \qquad \tilde{\ell}_1(\alpha) = \ell_1(\alpha) \int_F \mu_\omega(v)^\sigma G(dv).$$

Consider for example the following class of models for dense and sparse stochastic block-models.

Example 2. (*Dense and sparse stochastic block-models.*) Consider $F = [0, 1]$ and G the uniform distribution on $[0, 1]$. We choose for ω the graphon function associated to a (dense) stochastic block-model. For some partition $A_1, \dots, A_p$ of $[0, 1]$ and any $v, v' \in [0, 1]$, let

$$\omega(v, v') = B_{k,\ell} \qquad (49)$$

with $v \in A_k$, $v' \in A_\ell$, and B a $p \times p$ matrix where $B_{k,\ell} \in [0, 1]$ denotes the probability that a node in community k forms a link with a node in community ℓ. Then ω defines the community structure of the graph, and η will tune its sparsity properties. Choosing $\eta(x, y) = \mathbb{1}_{x \leq 1} \mathbb{1}_{y \leq 1}$ yields the dense, standard stochastic block-model; choosing $\eta(x, y) = \exp(-x - y)$ yields a sparse
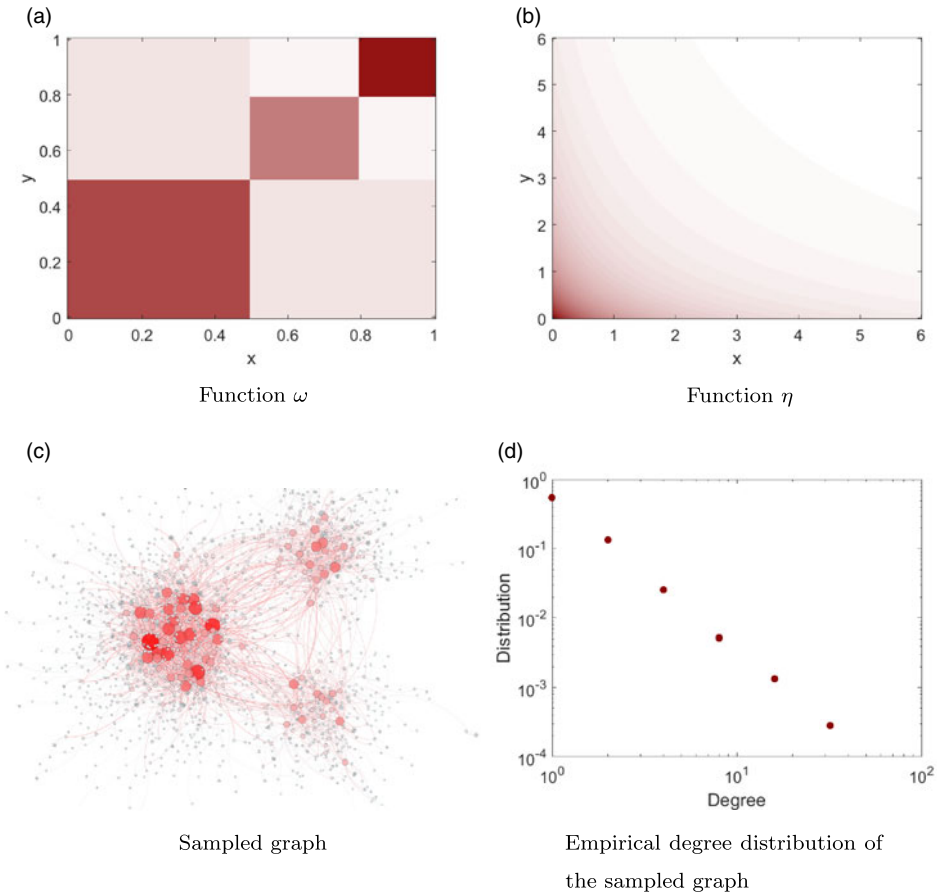
(a) Function ω

(b) Function η

(c) Sampled graph

(d) Empirical degree distribution of the sampled graph

FIGURE 4. Illustration of a sparse stochastic block-model with three communities. (a) The function ω, which controls the local community structure. A darker colour represents a higher value. (b) The function η, which controls the sparsity. (c) Graph sampled from the sparse stochastic block-model using $\alpha = 50$. The size of each node is proportional to its degree. (d) Empirical degree distribution of the sampled graph.

stochastic block-model without power-law behaviour; and so on. Figure 4 gives an illustration of the use of this model to obtain sparse stochastic block-models with power-law behaviour, generalising the model of Section 6.3. The function ω is defined by

$$A_1 = [0, 0.5), \quad A_2 = [0.5, 0.8), \quad A_3 = [0.8, 1],$$
$$B_{11} = 0.7, \quad B_{22} = 0.5, \quad B_{33} = 0.9, \quad B_{12} = B_{13} = 0.1, \quad B_{23} = 0.05,$$

and $\eta(x, y) = (1 + x)^{-1/\sigma}(1 + y)^{-1/\sigma}$, with $\sigma = 0.8$.

More generally, one can build on the large literature on (dense) graphon/exchangeable graph models, and combine these models with a function η satisfying Assumptions 1 and 2, such as

those described in the previous section, in order to sparsify a dense graphon and control its sparsity/power-law properties.

Remark 5. We can also obtain asymptotic results for those functions W that do not satisfy the separability condition (48). Let $\mu(u, v) = \int_{\mathbb{R}_+ \times F} W((u, v), (u', v'))du'dv'$. Assume that for each fixed v there exists $u_0(v) > 0$ such that, for $u > u_0$,

$$C_3 \tilde{\mu}_\eta(u)\tilde{\mu}_\omega(v) \leq \mu(u, v) \leq C_4 \tilde{\mu}_\eta(u)\tilde{\mu}_\omega(v), \tag{50}$$

where $\tilde{\mu}_\omega : F \to \mathbb{R}_+$, $\tilde{\mu}_\eta : \mathbb{R}_+ \to \mathbb{R}_+$ with $\tilde{\mu}_\eta(u) = \int_0^\infty \tilde{\eta}(u, u')du'$ for some positive function $\tilde{\eta}$, and $C_3 > 0$ and $C_4 > 0$. Assume that $\tilde{\mu}_\eta$ and $\tilde{\nu}_\eta$ satisfy Assumptions 1 and 2. Then the results of Theorems 1 and 2 and Corollary 1 hold up to a constant. For example, for $\sigma \in [0, 1]$ we have $N_\alpha^{(e)} \asymp N_\alpha^{2/(1+\sigma)} \ell_\sigma^*(N_\alpha)$ almost surely as α tends to infinity. In particular, the inequality from (50) is satisfied if

$$W((u_i, v_i), (u_j, v_j)) = 1 - e^{-\tilde{\omega}(v_i, v_j)\tilde{\eta}(u_i, u_j)}. \tag{51}$$

The models developed by [18, 37] for capturing (overlapping) communities fit into this framework. Ignoring loops, both models can be written in the form given by Equation (51) with $\tilde{\eta}(u, u') = 2\overline{\rho}^{-1}(u)\overline{\rho}^{-1}(u')$, where ρ is a Lévy measure on $(0, +\infty)$ and $\overline{\rho}(x) = \int_x^\infty \rho(dw)$ is the tail Lévy intensity with generalised inverse $\overline{\rho}^{-1}(x)$. When $\tilde{\omega}$ is given by Equation (49), it corresponds to the (dense) stochastic block-model graphon of [18], and when $\tilde{\omega}(v_i, v_j) = v_i^T v_j$ with $v_i \in \mathbb{R}_+^p$, it corresponds to the model of [37]. For instance, let ρ be the mean measure from Equation (46) with parameters $\tau_0 > 0$ and $\sigma_0 \in (-\infty, 1)$. Then for $\sigma_0 \in (0, 1)$, the corresponding sparse regime with power law for this graph is given by

$$N_\alpha^{(e)} \asymp N_\alpha^{2/(1+\sigma_0)}, \qquad \frac{C_3}{C_4} \frac{\sigma_0 \Gamma(j - \sigma_0)}{j! \Gamma(1 - \sigma_0)} \leq \lim_{\alpha \to \infty} \frac{N_{\alpha, j}}{N_\alpha} \leq \frac{C_4}{C_3} \frac{\sigma_0 \Gamma(j - \sigma_0)}{j! \Gamma(1 - \sigma_0)}, \quad j \geq 1$$

For $\sigma_0 = 0$ (sparse, almost dense regime), $N_\alpha^{(e)} \asymp N_\alpha^2 / \log(N_\alpha)^2$ and $N_{\alpha, j}/N_\alpha \to 0$ for $j \geq 1$; for $\sigma_0 < 0$ (dense regime), $N_\alpha^{(e)} \asymp N_\alpha^2$ and $N_{\alpha, j}/N_\alpha \to 0$ for $j \geq 1$, almost surely, as α tends to infinity.

7.2. Proof of Proposition 12

The proofs of Proposition 1 and Theorems 1 and 2 hold with x replaced by $(u, v) \in \mathbb{R}_+ \times F$, $dx = duG(dv)$, and $\mu(x) = \mu_\eta(u)\mu_\omega(v)$. We thus need only prove that if η satisfies Assumptions 1 and 2, then Lemmas B.2, B.3, and B.4 in the appendix hold. Recall that $\mu(x) = \mu_\eta(u)\mu_\omega(v)$, for $x = (u, v)$. Then, for all v such that $\mu_\omega(v) > 0$, we apply Lemma B.2 to get

$$g_0(t) = \int_0^\infty (1 - e^{-t\mu_\eta(u)})du, \qquad g_r(t) = \int_0^\infty \mu_\eta(u)^r e^{-t\mu_\eta(u)}du, \qquad t = \alpha\mu_\omega(v).$$

For all v such that $\mu_\omega(v) > 0$, this leads to

$$\int_0^\infty (1 - e^{-\alpha\mu_\omega(v)\mu_\eta(u)})du = \Gamma(1 - \sigma)\alpha^\sigma \ell(\alpha)\mu_\omega(v)^\sigma \frac{\ell\{\alpha\mu_\omega(v)\}}{\ell(\alpha)}\{1 + o(1)\}$$

$$= \Gamma(1 - \sigma)\alpha^\sigma \ell(\alpha)\mu_\omega(v)^\sigma \{1 + o(1)\}.$$

To prove that there is convergence in $L_1(G)$, note that if $\mu_\omega(v) > 0$, since $\mu_\omega \leq 1$, we have

$$\int_0^\infty \left(1 - e^{-\alpha\mu_\omega(v)\mu_\eta(u)}\right)du = \int_0^\infty \mu_\eta^{-1}\left\{\frac{z}{\alpha\mu_\omega(v)}\right\}e^{-z}dz \leq \int_0^\infty \mu_\eta^{-1}\left(\frac{z}{\alpha}\right)e^{-z}dz.$$

Moreover,

$$\sup_{\alpha \geq 1} \frac{1}{\alpha^\sigma \ell(\alpha)} \int_0^\infty \mu_\eta^{-1}\left(\frac{z}{\alpha}\right) e^{-z} dz < +\infty;$$

thus the Lebesgue dominated convergence theorem implies

$$\int_F \int_0^\infty \left(1 - e^{-\alpha \mu_\omega(v)\mu_\eta(u)}\right) du G(dv) \sim \Gamma(1-\sigma)\alpha^\sigma \ell(\alpha) \int_F \mu_\omega(v)^\sigma G(dv)$$

when $\sigma < 1$, while when $\sigma = 1$,

$$\int_F \int_0^\infty (1 - e^{-\alpha \mu_\omega(v)\mu_\eta(u)}) du G(dv) \sim \alpha \ell_1(\alpha) \int_F \mu_\omega(v) G(dv).$$

The same reasoning is applied to the integrals

$$\int_F \mu_\omega(v)^r \int_0^\infty \mu_\eta(u)^r e^{-\alpha \mu_\omega(v)\mu_\eta(u)} du G(dv).$$

To verify Lemma B.3, note that

$$h_0(\alpha) = \int_F \omega(v, v) \int_0^\infty \eta(u, u)\left(1 - e^{-\alpha \mu_\omega(v)\mu_\eta(u)}\right) du G(dv),$$

$$h_r(\alpha) = \int_F \omega(v, v)\mu_\omega(v)^r \int_0^\infty \eta(u, u)\mu_\eta(u)^r e^{-\alpha \mu_\omega(v)\mu_\eta(u)} du G(dv),$$

so that the Lebesgue dominated convergence theorem also leads to

$$h_0(\alpha) \sim \int_F \omega(v, v) \int_0^\infty \eta(u, u) du G(dv), \qquad h_r(\alpha) = o(\alpha^{-r})$$

and the control of the integrals $\int_{\mathbb{R}_+ \times F} \{t\mu(u, v)\} e^{-t\mu(u,v)} du G(dv)$ as in Lemma B.4.

8. Conclusion

In this article, we derived a number of properties of graphs based on exchangeable random measures. We related the sparsity and power-law properties of the graphs to the regular variation properties of the marginal graphon function, identifying four different regimes, from dense to almost extremely sparse. We derived asymptotic results for the global and local clustering coefficients. We derived a central limit theorem for the number of nodes N_α in the sparse and dense regimes, and for the number of nodes of degree greater than j in the dense regime. We conjecture that a CLT also holds for $N_{\alpha,j}$ in the sparse regime, under assumptions similar to Assumptions 4 and 5, and that a (lengthy) proof similar to that of Theorem 5 could be used. We leave this for future work.

Appendix A. Proofs of Theorem 1 and Proposition 6

Let $g_{\alpha,x}(\theta, \vartheta)$ be defined, for any $\alpha, x, \theta, \vartheta > 0$, by

$$g_{\alpha,x}(\theta, \vartheta) = -\log\{1 - W(x, \vartheta)\}\mathbb{1}_{\theta \leq \alpha}. \tag{52}$$

A.1. Proof of Theorem 1

The mean number of nodes is

$$E(N_\alpha) = \alpha \int_{\mathbb{R}_+} \{1 - e^{-\alpha\mu(x)}\}dx + \alpha \int_{\mathbb{R}_+} W(x, x)e^{-\alpha\mu(x)}dx;$$

see [38, Theorem 5.4]. By the Lebesgue dominated convergence theorem, we have $\alpha \int_{\mathbb{R}_+} W(x, x)e^{-\alpha\mu(x)}dx = o(\alpha)$. Using Lemma B.2, as α goes to infinity we have $\int_{\mathbb{R}_+} (1 - e^{-\alpha\mu(x)})dx \sim \alpha^\sigma \ell(\alpha)\Gamma(1-\sigma)$ for $\sigma \in [0, 1)$ and $\int_{\mathbb{R}_+} \{1 - e^{-\alpha\mu(x)}\}dx \sim \alpha\ell_1(\alpha)$ for $\sigma = 1$. It follows that as α goes to infinity,

$$E(N_\alpha) \sim \begin{cases} \alpha^{\sigma+1}\ell(\alpha)\Gamma(1-\sigma) & \text{if } \sigma \in [0, 1), \\ \alpha^2\ell_1(\alpha) & \text{if } \sigma = 1. \end{cases}$$

The mean number of nodes of degree j is

$$E(N_{\alpha,j}) = \frac{\alpha^{j+1}}{j!} \int_{\mathbb{R}_+} (1 - W(\vartheta, \vartheta))e^{-\alpha\mu(\vartheta)}\mu(\vartheta)^j d\vartheta \qquad (53)$$
$$+ \frac{\alpha^j}{j-1!} \int_{\mathbb{R}_+} e^{-\alpha\mu(\vartheta)}W(\vartheta, \vartheta)\mu(\vartheta)^{j-1}d\vartheta;$$

see [38, Theorem 5.5]. Lemma B.3 implies that

$$-\frac{\alpha^{j+1}}{j!} \int_{\mathbb{R}_+} W(\vartheta, \vartheta)e^{-\alpha\mu(\vartheta)}\mu(\vartheta)^j d\vartheta + \frac{\alpha^j}{j-1!} \int_{\mathbb{R}_+} e^{-\alpha\mu(\vartheta)}W(\vartheta, \vartheta)\mu(\vartheta)^{j-1}d\vartheta = o(\alpha),$$

and from Lemma B.2, when $\sigma \in [0, 1)$, we have

$$\frac{\alpha^{j+1}}{j!} \int_{\mathbb{R}_+} e^{-\alpha\mu(\vartheta)}\mu(\vartheta)^j d\vartheta \sim \frac{\sigma\Gamma(j-\sigma)}{j!}\alpha^{1+\sigma}\ell(\alpha).$$

If $\sigma = 1$, from Lemma B.2 we have

$$\alpha^2 \int_{\mathbb{R}_+} e^{-\alpha\mu(\vartheta)}\mu(\vartheta)d\vartheta \sim \alpha^2\ell_1(\alpha),$$

and for $j \geq 2$,

$$\frac{\alpha^{j+1}}{j!} \int_{\mathbb{R}_+} e^{-\alpha\mu(\vartheta)}\mu(\vartheta)^j d\vartheta \sim \frac{1}{j(j-1)}\alpha^2\ell(\alpha).$$

Finally, for $\sigma \in [0, 1)$ we obtain

$$E(N_{\alpha,j}) \sim \frac{\sigma\Gamma(j-\sigma)}{j!}\alpha^{1+\sigma}\ell(\alpha),$$

and for $\sigma = 1$ we obtain $E(N_{\alpha,1}) \sim \alpha^2\ell_1(\alpha)$ and $E(N_{\alpha,j}) \sim \alpha^2/\{j(j-1)\}\ell(\alpha)$ for $j \geq 2$.

A.2. Proof of Proposition 6

For $j \geq 1$, define

$$R_{\alpha j} = \sum_i T_{\alpha i} \mathbb{1}_{D_{\alpha i} = j}. \tag{54}$$

Then $R_{\alpha j}$ corresponds to the number of triangles having a node of degree j as a vertex, where triangles having $k \leq 3$ degree-j nodes as vertices are counted k times. We therefore have

$$C_{\alpha,j}^{(\ell)} = \frac{2}{j(j-1)} \frac{R_{\alpha j}}{N_{\alpha,j}}.$$

The proof for the asymptotic behaviour of the local clustering coefficients $C_{\alpha,j}^{(\ell)}$ is organised as follows. We first derive a convergence result for $E(R_{\alpha j})$. This result is then extended to an almost sure result. The extension requires some additional work as $R_{\alpha j}$ is not monotone, and $\sum_{j \geq k} R_{\alpha k}$ is monotone but not of the same order as $R_{\alpha j}$; hence a proof similar to that for $N_{\alpha,j}$ (see Section 3.2) cannot be used. The almost sure convergence results for $C_{\alpha,j}^{(\ell)}$ and $\overline{C}_\alpha^{(\ell)}$ then follow from the almost sure convergence result for $R_{\alpha j}$.

We have

$$R_{\alpha j} = \sum_i T_{\alpha i} \mathbb{1}_{D_{\alpha i} = j} = \frac{1}{2} \sum_{i \neq l \neq k} Z_{il} Z_{ik} Z_{lk} \mathbb{1}_{\sum_s Z_{is} = j} \mathbb{1}_{\theta_s \leq \alpha} \mathbb{1}_{\theta_i \leq \alpha} \mathbb{1}_{\theta_l \leq \alpha} \mathbb{1}_{\theta_k \leq \alpha}$$

and

$$E\left(R_{\alpha j} \mid M\right) = \frac{1}{2} \sum_{i \neq l \neq k} W(\vartheta_i, \vartheta_l) W(\vartheta_i, \vartheta_k) W(\vartheta_l, \vartheta_k) \frac{1}{(j-2)!}$$

$$\times \sum_{i_1 \neq i_2 \ldots \neq i_{j-2} \neq l \neq k} \left[\prod_{s=1}^{j-2} W(\vartheta_i, \vartheta_s) \right] e^{- \sum_{s \neq l, k, i_1, \ldots, i_{j-2}} g_{\alpha, \vartheta_i}(\theta_s, \vartheta_s)}$$

$$= \frac{1}{2(j-2)!} \sum_{i \neq l \neq k \neq i_1 \neq i_2 \ldots \neq i_{j-2}} W(\vartheta_i, \vartheta_l) W(\vartheta_i, \vartheta_k) W(\vartheta_l, \vartheta_k)(1 - W(\vartheta_i, \vartheta_i))$$

$$\times \left[\prod_{s=1}^{j-2} W(\vartheta_i, \vartheta_s) \right] e^{- \sum_{s \neq l, k, i_1, \ldots, i_{j-2}} g_{\alpha, \vartheta_i}(\theta_s, \vartheta_s)}$$

$$+ \frac{1}{2(j-3)!} \sum_{i \neq l \neq k \neq i_1 \neq i_2 \ldots \neq i_{j-3}} W(\vartheta_i, \vartheta_i) W(\vartheta_i, \vartheta_l) W(\vartheta_i, \vartheta_k) W(\vartheta_l, \vartheta_k)$$

$$\times \prod_{s=1}^{j-3} W(\vartheta_i, \vartheta_s) e^{- \sum_{s \neq i, l, k, i_1, \ldots, i_{j-3}} g_{\alpha, \vartheta_i}(\theta_s, \vartheta_s)},$$

where $g_{\alpha, x}(\theta, \vartheta)$ is defined in Equation (52). Applying the Slivnyak–Mecke theorem, we obtain

$$E\left(R_{\alpha j}\right) = \frac{\alpha^{j+1}}{2(j-2)!} \int_{\mathbb{R}_+^3} W(x, y) W(x, z) W(y, z)(1 - W(x, x)) \mu(x)^{j-2} e^{-\alpha \mu(x)} dx dy dz$$

$$+ \frac{\alpha^j}{2(j-3)!} \int_{\mathbb{R}_+^3} W(x, y) W(x, z) W(y, z) W(x, x) \mu(x)^{j-3} e^{-\alpha \mu(x)} dx dy dz. \tag{55}$$

Note that under Assumption 1 with $\sigma \in (0, 1)$, $\mu(x) > 0$ for all x. The leading term in the right-hand side of Equation (A.2) is the first term. We have therefore

$$E\left(R_{\alpha j}\right) \sim \frac{\alpha^{j+1}}{2(j-2)!} \int_{\mathbb{R}_+^3} L(x)\mu(x)^j e^{-\alpha\mu(x)} dxdydz,$$

where

$$L(x) = \frac{(1 - W(x, x)) \int_{\mathbb{R}_+^2} W(x, y)W(x, z)W(y, z)dydz}{\mu(x)^2}.$$

As $\lim_{x \to \infty} W(x, x) = 0$, the condition (29) implies $\lim_{x \to \infty} L(x) = b$.

Case $b > 0$. Assume first that $b > 0$. In this case, L is a slowly varying function by assumption. Therefore, using Lemma B.5, we have, under Assumption 1, for $\sigma \in (0, 1)$,

$$\int_0^\infty L(x)\mu(x)^j e^{-\alpha\mu(x)} dx \sim \sigma b\ell(\alpha)\Gamma(j - \sigma)\alpha^{\sigma - j}$$

as α tends to infinity. Hence

$$E\left(R_{\alpha j}\right) \sim \frac{b\sigma\,\Gamma(j - \sigma)}{2(j-2)!}\alpha^{1+\sigma}\ell(\alpha) \tag{56}$$

as α tends to infinity. In order to obtain a convergence in probability, we state the following proposition, whose proof is given in Section A.3 in the supplementary material [13] and is similar to that of Proposition 4.

Proposition A.1. *Under Assumptions 1 and 2, with $\sigma \in [0, 1]$, slowly varying function ℓ, and positive scalar a satisfying (17), we have*

$$\mathrm{var}\left(\sum_i T_{\alpha i}\mathbb{1}_{D_{\alpha i}=j}\right) = O\{\alpha^{3+2\sigma-2a}\ell_\sigma(\alpha)^2\}\text{as } \alpha \to \infty,$$

and for any sequence α_n going to infinity such that $\alpha_{n+1} - \alpha_n = o(\alpha_n)$,

$$\mathrm{var}\left(\sum_i T_{\alpha_{n+1}i}\mathbb{1}_{D_{\alpha n i}=j}\mathbb{1}_{\sum_{i'}\mathbb{1}_{\alpha_n<\theta_{i'}\leq\alpha_{n+1}}Z_{ii'}=1}\right) = O\left(\alpha_n^{3+2\sigma-2a}\ell_\sigma(\alpha_n)^2\right) \text{ as } n \to \infty.$$

We now want to find a subsequence α_n along which the convergence is almost sure. Using Chebyshev's inequality and the first part of Proposition A.1, there exist $n_0 \geq 0$ and $C \geq 0$ such that for all $n > n_0$,

$$\mathbb{P}\left(\left|\frac{R_{\alpha_n j}}{E(R_{\alpha_n j})} - 1\right| > \epsilon\right) \leq \frac{C\alpha_n^{3+2\sigma-2a}\ell_\sigma(\alpha_n)^2}{\epsilon^2\left(\frac{b\sigma\Gamma(j-\sigma)}{2(j-2)!}\alpha_n^{1+\sigma}\ell(\alpha_n)\right)^2}.$$

Now, if Assumption 2 is satisfied for a given $a > 1/2$, consider the sequence

$$\alpha_n = \left(n\log^2 n\right)^{1/(2a-1)}, \tag{57}$$

so that $\sum_n \alpha_n^{1-2a} < +\infty$ and

$$\sum_n \mathbb{P}\left(\left|\frac{R_{\alpha_n j}}{Ec(R_{\alpha_n j}(n \log^2 n))} - 1\right| > \epsilon\right) < \infty.$$

Therefore, using the Borel–Cantelli lemma, we have

$$R_{\alpha_n j} \sim \frac{b\sigma\,\Gamma(j-\sigma)}{2(j-2)!}\alpha_n^{1+\sigma}\ell(\alpha_n)$$

almost surely as $n \to \infty$.

The goal is now to extend this result to $R_{\alpha j}$, by sandwiching. Let $I_\alpha := \{i : \theta_i \leq \alpha\}$. We have the following upper and lower bounds for $R_{\alpha j}$:

$$\sum_{i \in I_{\alpha_n}} T_{\alpha_n i}\mathbb{1}_{D_{\alpha i}=j} \leq \sum_{i \in I_\alpha} T_{\alpha i}\mathbb{1}_{D_{\alpha i}=j} \leq \sum_{i \in I_{\alpha_{n+1}}} T_{\alpha_{n+1} i}\mathbb{1}_{D_{\alpha i}=j}. \tag{58}$$

Considering the upper bound of (58), we have

$$\sum_{i \in I_{\alpha_{n+1}}} T_{\alpha_{n+1} i}\mathbb{1}_{D_{\alpha i}=j} \leq \sum_{i \in I_{\alpha_{n+1}}} T_{\alpha_{n+1} i}\mathbb{1}_{D_{\alpha_{n+1} i}=j} + \sum_{i \in I_{\alpha_{n+1}}} T_{\alpha_{n+1} i}\mathbb{1}_{D_{\alpha i}=j}\mathbb{1}_{D_{\alpha_{n+1} i}>j}$$

$$\leq R_{\alpha_{n+1} j} + \widetilde{R}_{nj}, \tag{59}$$

where

$$\widetilde{R}_{nj} = \sum_{i \in I_{\alpha_{n+1}}} T_{\alpha_{n+1} i}\mathbb{1}_{D_{\alpha_n i}\leq j}\mathbb{1}_{\sum_{i'}\mathbb{1}_{\alpha_n < \theta_{i'} \leq \alpha_{n+1}} Z_{ii'} \geq 1}. \tag{60}$$

We can bound the lower bound of (58) by

$$\sum_{i \in I_{\alpha_n}} T_{\alpha_n i}\mathbb{1}_{D_{\alpha i}=j} \geq \sum_{i \in I_{\alpha_n}} T_{\alpha_n i}\mathbb{1}_{D_{\alpha_n i}=j}\mathbb{1}_{D_{\alpha i}=j}$$

$$\geq \sum_{i \in I_{\alpha_n}} T_{\alpha_n i}\mathbb{1}_{D_{\alpha_n i}=j} - \sum_{i \in I_{\alpha_n}} T_{\alpha_n i}\mathbb{1}_{D_{\alpha_n i}=j}\mathbb{1}_{D_{\alpha_{n+1} i}>j}$$

$$\geq \sum_{i \in I_{\alpha_n}} T_{\alpha_n i}\mathbb{1}_{D_{\alpha_n i}=j} - \sum_{i \in I_{\alpha_{n+1}}} T_{\alpha_{n+1} i}\mathbb{1}_{D_{\alpha_n i}\leq j}\mathbb{1}_{\sum_{i'}\mathbb{1}_{\alpha_n < \theta_{i'} \leq \alpha_{n+1}} Z_{ii'} \geq 1}$$

$$= R_{\alpha_n j} - \widetilde{R}_{nj}. \tag{61}$$

The following lemma, proved in Section A.4 of the supplementary material [13], provides an asymptotic bound for the remainder term $\widetilde{R}_{nj}$.

Lemma A.1. *Let* $\widetilde{R}_{nj}$ *be defined as in Equation* (60). *If Assumptions* 1 *and* 2 *hold with* $\sigma \in (0, 1)$ *and slowly varying function* ℓ, *and the condition* (29) *is satisfied with* $b > 0$, *then we have*

$$\widetilde{R}_{nj} = o(n \log^2 n)(\alpha_n^{1+\sigma}\ell(\alpha_n)(n \log^2 n))$$

almost surely as α *tends to infinity*

Combining Lemma A.1 with the inequalities (58), (59), and (61), and the fact that $R_{\alpha_n j} \sim R_{\alpha_{n+1} j} \asymp \alpha_n^{1+\sigma}\ell(\alpha_n)$ almost surely as $n \to \infty$, we obtain by sandwiching

$$R_{\alpha j} \sim \frac{b\sigma\,\Gamma(j-\sigma)}{2(j-2)!}\alpha^{1+\sigma}\ell(\alpha) \text{ almost surely as } \alpha \text{ tends to infinity.}$$

Recalling that $N_{\alpha,j} \sim \frac{\sigma\Gamma(j-\sigma)}{j!}\alpha^{1+\sigma}\ell(\alpha)$ almost surely, we have, for any $j \geq 1$,

$$C_{\alpha,j}^{(\ell)} = \frac{2R_{\alpha j}}{j(j-1)N_{\alpha,j}} \to b \text{ almost surely as } \alpha \text{ tends to infinity.}$$

Finally, as $\frac{N_{\alpha,j}}{N_\alpha - N_{\alpha,1}}$ converges to a constant $\pi_j \in (0, 1)$ almost surely for any j, we have, using Toeplitz's lemma,

$$\overline{C}_\alpha^{(\ell)} = \frac{1}{N_\alpha - N_{\alpha,1}} \sum_{j\geq 2} N_{\alpha,j} C_{\alpha,j}^{(\ell)} \to b$$

almost surely as α tends to infinity.

Case $b = 0$. In the case $L(x) \to 0$, Lemma B.5 gives $\int_0^\infty L(x)\mu(x)^j e^{-\alpha\mu(x)}dx = o(\alpha^{\sigma-j})$; hence, by Markov's inequality,

$$R_{\alpha j} = o(\alpha^{1+\sigma}\ell(\alpha))$$

and $C_{\alpha j}^{(\ell)} \to 0$ in probability as α tends to infinity.

Appendix B. Technical lemma

The proof of the following lemma is similar to the proof of Proposition 2 in [17], and is omitted here.

Lemma B.1. *Let $(X_t)_{t\geq 0}$ be some positive monotone increasing stochastic process with finite first moment $(E(X_t))_{t\geq 0} \in RV_\gamma$, where $\gamma \geq 0$ (see Definition C.1). Assume*

$$\text{var}(X_t) = O\{t^{-a}E(X_t)^2\}$$

for some $a > 0$. Then

$$\frac{X_t}{E(X_t)} \to 1 \text{ almost surely as } t \to \infty.$$

The following lemma is a compilation of results from Propositions 17, 18, and 19 in [17].

Lemma B.2. *Let $\mu : \mathbb{R}_+ \to \mathbb{R}_+$ be a positive, right-continuous, and monotone decreasing function with $\int_0^\infty \mu(x)dx < \infty$ and generalised inverse $\mu^{-1}(x) = \inf\{y > 0 \,|\, f(y) \leq x\}$ satisfying*

$$\mu^{-1}(x) = x^{-\sigma}\ell(1/x), \tag{62}$$

where $\sigma \in [0, 1]$ and ℓ is a slowly varying function. Consider

$$g_0(t) = \int_0^\infty \left(1 - e^{-t\mu(x)}\right)dx, \qquad g_r(t) = \int_0^\infty e^{-t\mu(x)}\mu(x)^r dx, \quad r \geq 1.$$

Then, for any $\sigma \in [0, 1)$,

$$g_0(t) \sim \Gamma(1-\sigma)t^\sigma\ell(t) \text{ as } t \to \infty,$$

and for $r \geq 1$,

$$\begin{cases} g_r(t) \sim t^{\sigma-r}\ell(t)\sigma\Gamma(r-\sigma) & \text{if } \sigma \in (0, 1), \\ g_r(t) = o\{t^{\sigma-r}\ell(t)\} & \text{if } \sigma = 0, \end{cases}$$

as $t \to \infty$. For $\sigma = 1$, as $t \to \infty$,

$$g_0(t) \sim t\ell_1(t), \quad g_1(t) \sim \ell_1(t), \quad g_r(t) \sim t^{1-r}\ell(t)\Gamma(r-1),$$

where $\ell_1(t) = \int_t^\infty x^{-1}\ell(x)dx$. Note that $\ell(t) = o(\ell_1(t))$; hence $g_r(t) = o\{t^{1-r}\ell_1(t)\}$.

Lemma B.3. *Let* $\mu : \mathbb{R}_+ \to \mathbb{R}_+$ *be a positive, monotone decreasing function, and* $u : \mathbb{R}_+ \to [0, 1]$ *a positive and integrable function with* $\int_0^\infty u(x)dx < \infty$. *Consider* $h_0(t) = \int_0^\infty u(x)(1 - e^{-t\mu(x)})dx$ *and, for* $r \geq 1$, $h_r(t) = \int_0^\infty u(x)e^{-t\mu(x)}\mu(x)^r dx$. *Then, as* $t \to \infty$,

$$h_0(t) \sim \int_0^\infty u(x)dx, \quad h_r(t) = o(t^{-r}), \quad r \geq 1.$$

Proof. We have $h_0(t) \to \int_0^\infty u(x)dx$ by dominated convergence. Using Proposition C.3, we have

$$\frac{th_1(t)}{\int_0^\infty u(x)dx} \to 0.$$

We proceed by induction to obtain the final result.   □

Lemma B.4. *Let* μ *be a non-negative, non-increasing function on* $\mathbb{R}_+$, *with* $\int_0^\infty \mu(x)dx < \infty$, *whose generalised inverse* μ^{-1} *satisfies* $\mu^{-1}(x) \sim x^{-\sigma}\ell(1/x)$ *as* $x \to 0$, *with* $\sigma \in [0, 1]$ *and* ℓ *a slowly varying function. Then as* $t \to \infty$, *for all* $r > \sigma$,

$$\int_{\mathbb{R}_+} \mu(x)^r e^{-t\mu(x)}dx = O\{t^{\sigma-r}\ell(t)\}$$

Proof. Let $r > \sigma$. Let $U(y) = \mu^{-1}(1/y)$. Then U is non-negative and non-decreasing, with $U(y) \sim y^\sigma \ell(y)$ as $y \to \infty$. Making the change of variable $x = U(y)$, one obtains

$$\int_0^\infty \{\mu(x)\}^r e^{-t\mu(x)}dx = \int_0^\infty y^{-r}e^{-t/y}dU(y).$$

We follow part of the proof in [6, p. 37]. Note that $y \to y^{-r}\exp(-t/y)$ is monotone increasing on $[0, t/r]$ and monotone decreasing on $[t/r, \infty)$. We have

$$\int_0^\infty y^{-r}e^{-t/y}dU(y) = \left\{\int_0^{t/r} + \sum_{n=1}^\infty \int_{2^{n-1}t/r}^{2^n t/r}\right\} y^{-r}e^{-t/y}dU(y)$$

$$\leq t^{-r}e^{-r}r^r U(t/r) + t^{-r}r^r \sum_{n=1}^\infty 2^{-r(n-1)}U\left(2^n t/r\right)$$

$$\leq 2t^{\sigma-r}e^{-r}r^r\ell(t/r) + 2t^{-r}r^r \sum_{n=1}^\infty 2^{-r(n-1)}(2^n t/r)^\sigma \ell\left(2^n t/r\right)$$

$$\leq 2t^{\sigma-r}e^{-r}r^r\ell(t/r) + 2^{r+1}t^{\sigma-r}r^{r-\sigma} \sum_{n=1}^\infty 2^{-n(r-\sigma)}\ell\left(2^n t/r\right)$$

for t large, using the regular variation property of U. Using Potter's bound [6, Theorem 1.5.6], for any $\delta > 0$ and for t large we have

$$\ell(2^n t/r) \leq 2\ell(t)\max\left(1, 2^{n\delta}/r^\delta\right).$$

Hence, for t large,

$$\int_0^\infty y^{-r}e^{-t/y}dU(y) \lesssim t^{\sigma-r}\ell(t)\left(1 + \sum_{n=1}^\infty 2^{-n(r-\sigma)}\max\left(r^\delta, 2^{n\delta}\right)\right).$$

Taking $0 < \delta < \frac{r-\sigma}{2}$, we conclude that the series in the right-hand side converges.   □

The next lemma is a slight variation of Lemma B.2, with the addition of a slowly varying function in the integrals. Note that the case where $\sigma = 0$ and ℓ tends to a constant is not covered.

Lemma B.5. *Let $f : \mathbb{R}_+ \to \mathbb{R}_+$ be a positive, right-continuous, and monotone decreasing function with $\int_0^\infty f(x)dx < \infty$ and generalised inverse $f^{-1}(x) = \inf\{y > 0 \,|\, f(y) \le x\}$ satisfying*

$$f^{-1}(x) = x^{-\sigma}\ell(1/x), \tag{63}$$

where $\sigma \in [0, 1]$ and ℓ is a slowly varying function, with $\lim_{t\to\infty}\ell(t) = \infty$ if $\sigma = 0$. Consider

$$\widetilde{g}_0(t) = \int_0^\infty \left(1 - e^{-tf(x)}\right)L(x)dx$$

and, for $r \ge 1$,

$$\widetilde{g}_r(t) = \int_0^\infty e^{-tf(x)}f(x)^r L(x)dx,$$

where $L : \mathbb{R}_+ \to (0, \infty)$ is a locally integrable function with $\lim_{t\to\infty} L(t) = b \in [0, \infty)$. Then, for any $\sigma \in [0, 1)$,

$$\begin{cases} \widetilde{g}_0(t) \sim b\Gamma(1-\sigma)t^\sigma\ell(t) & \text{if } b > 0, \\ \widetilde{g}_0(t) = o(t^\sigma\ell(t)) & \text{if } b = 0, \end{cases}$$

and for $r \ge 1$,

$$\begin{cases} \widetilde{g}_r(t) \sim bt^{\sigma-r}\ell(t)\sigma\Gamma(r-\sigma) & \text{if } \sigma \in (0, 1), \ b > 0, \\ \widetilde{g}_r(t) = o\{t^{\sigma-r}\ell(t)\} & \text{if } \sigma = 0 \text{ or } b = 0, \end{cases}$$

as $t \to \infty$. For $\sigma = 1$, $b > 0$, as $t \to \infty$,

$$\widetilde{g}_0(t) \sim bt\ell_1(t), \qquad \widetilde{g}_1(t) \sim b\ell_1(t), \qquad \widetilde{g}_r(t) \sim bt^{1-r}\ell(t)\Gamma(r-1),$$

where $\ell_1(t) = \int_t^\infty x^{-1}\ell(x)dx$. Note that $\ell(t) = o(\ell_1(t))$; hence $\widetilde{g}_r(t) = o\{t^{1-r}\ell_1(t)\}$.

Proof. Let $g_0(t) = \int_0^\infty \left(1 - e^{-tf(x)}\right)dx$. Let $\ell_1(t) = \int_t^\infty x^{-1}\ell(x)dx$ and $\ell_\sigma(t) = \Gamma(1-\sigma)\ell(t)$ if $\sigma \in [0, 1)$. Using Lemma B.2, we have $g_0(t) \sim t^\sigma\ell_\sigma(t)$ as $t \to \infty$, and in particular $g_0(t) \to \infty$. By dominated convergence, for any $x_0 > 0$, we have

$$\int_0^{x_0} \left(1 - e^{-tf(x)}\right)L(x)dx \to \int_0^{x_0} L(x)dx < \infty;$$

hence, $\widetilde{g}_0(t) \sim \int_{x_0}^\infty \left(1 - e^{-tf(x)}\right)L(x)dx$ as $t \to \infty$.

Let $\epsilon > 0$. There exists x_0 such that for all $x \ge x_0$, $|L(x) - b| \le \epsilon$ and so

$$(b - \epsilon)\int_{x_0}^\infty \left(1 - e^{-tf(x)}\right)dx \le \int_{x_0}^\infty \left(1 - e^{-tf(x)}\right)L(x)dx \le (b + \epsilon)\int_{x_0}^\infty \left(1 - e^{-tf(x)}\right)dx.$$

Hence, by sandwiching, we have

$$\lim_{t\to\infty} \frac{\widetilde{g}_0(t)}{t^\sigma\ell_\sigma(t)} = \lim_{t\to\infty} \frac{\int_{x_0}^\infty \left(1 - e^{-tf(x)}\right)L(x)dx}{t^\sigma\ell_\sigma(t)} \in (b - \epsilon, b + \epsilon).$$

As this is true for any $\epsilon > 0$, we obtain $\widetilde{g}_0(t) \sim bt^\sigma\ell_\sigma(t)$ as $t \to \infty$ if $b > 0$ and $\widetilde{g}_0(t) = o(t^\sigma\ell_\sigma(t))$ if $b = 0$. The asymptotic results for $\widetilde{g}_r(t)$ then follow from Proposition C.3. □

The following is a corollary of [40, Theorem 2.1].

Corollary B.1. [40, Theorem 2.1]. *Assume that*

$$f(x) \sim \ell(x)x^\alpha e^{-\beta x}$$

where ℓ is a slowly varying, locally bounded function on $(0, \infty)$, and where either $\beta \geq 0$ and $\alpha \in \mathbb{R}$, or $\alpha < -1$ and $\beta = 0$. Then, as $n \to \infty$,

$$\int_0^\infty \frac{(\lambda x)^n e^{-\lambda x}}{n!} f(x)dx \sim \frac{\ell(n)}{(\lambda+\beta)^{\alpha+1}} \left(\frac{\lambda}{\lambda+\beta}\right)^n n^\alpha \tag{64}$$

and

$$\int_0^\infty \frac{(\lambda x)^n e^{-\lambda x}}{n!} u(x)f(x)dx = o\left(\frac{\ell(n)}{(\lambda+\beta)^{\alpha+1}} \left(\frac{\lambda}{\lambda+\beta}\right)^n n^\alpha\right) \tag{65}$$

for any locally bounded function u vanishing at infinity.

Proof. Equation (64) is proved in [40, Theorem 2.1]. For any $x_0 > 0$, we have

$$\int_0^\infty \frac{(\lambda x)^n e^{-\lambda x}}{n!} u(x)f(x)dx \sim \int_{x_0}^\infty \frac{(\lambda x)^n e^{-\lambda x}}{n!} u(x)f(x)dx.$$

For any $\epsilon > 0$, there is x_0 such that $u(x) < \epsilon$ for all $x > x_0$; hence

$$\int_{x_0}^\infty \frac{(\lambda x)^n e^{-\lambda x}}{n!} u(x)f(x)dx \leq \epsilon \int_0^\infty \frac{(\lambda x)^n e^{-\lambda x}}{n!} f(x)dx,$$

and (65) follows from (64) by sandwiching. $\square$

The following lemma is useful to bound the variance and for the proof of the central limit theorem.

Lemma B.6. *Assume the functions μ and ν satisfy Assumptions 1 and 2, for some $\sigma \in [0, 1]$ and slowly varying function ℓ, with $a > \min(1/2, \sigma)$ if $\sigma < 1$ and $a = 1$ if $\sigma = 1$. Then*

$$\int_{\mathbb{R}_+^2} \nu(x, y)e^{-\alpha\mu(x)-\alpha\mu(y)+\alpha\nu(x,y)}dxdy = O\left(\alpha^{2\sigma-2a}\ell_\sigma^2(\alpha)\right),$$

where ℓ_σ is defined in Equation (20). If $a = 1$ and $\sigma = 0$ we have the stronger result

$$\int_{\mathbb{R}_+^2} \nu(x, y)e^{-\alpha\mu(x)-\alpha\mu(y)+\alpha\nu(x,y)}dxdy = o\left(\alpha^{-2}\ell^2(\alpha)\right).$$

Proof. Using the fact that $\nu(x, y) \leq \sqrt{\mu(x)\mu(y)} \leq (\mu(x) + \mu(y))/2$ and Assumption 2, we have

$$\int_{\mathbb{R}_+^2} \nu(x, y)e^{-\alpha\mu(x)-\alpha\mu(y)+\alpha\nu(x,y)}dxdy \leq \int_{\mathbb{R}_+^2} \nu(x, y)e^{-\alpha\mu(x)/2-\alpha\mu(y)/2}dxdy$$

$$\leq C_1 \left(\int_{x_0}^\infty \mu(x)^a e^{-\alpha\mu(x)/2}\right)^2 + 2\int_0^{x_0}\int_0^\infty \nu(x, y)e^{-\alpha\mu(x)/2-\alpha\mu(y)/2}dxdy,$$

where $a > \min(1/2, \sigma)$ if $\sigma < 1$ and $a = 1$ if $\sigma = 1$. Using $\int_0^{x_0} \nu(x, y) dx \leq x_0 \mu(y)$, we have

$$\int_0^{x_0} \int_0^\infty \nu(x, y) e^{-\alpha\mu(x)/2 - \alpha\mu(y)/2} dx dy \leq e^{-\alpha\mu(x_0)/2} x_0 \int_0^\infty \mu(y) e^{-\alpha\mu(y)/2} dy$$

if $x_0 > 0$ (otherwise the bound is trivial). Since $\mu(x_0) > 0$, the right-hand side is in $o(\alpha^{-p})$ for any $p > 0$. Using Lemma B.4 (for $\sigma < 1$) or 10 (for $\sigma = 1$) together with Assumption 1, we therefore obtain

$$\int_{\mathbb{R}_+^2} \nu(x, y) e^{-\alpha\mu(x) - \alpha\mu(y) + \alpha\nu(x,y)} dx dy = O\{\alpha^{2\sigma - 2a} \ell_\sigma^2(\alpha)\}.$$

In the case $\sigma = 0$ and $a = 1$, Lemma B.2 and Assumption 1 give

$$\int_{\mathbb{R}_+^2} \nu(x, y) e^{-\alpha\mu(x) - \alpha\mu(y) + \alpha\nu(x,y)} dx dy = o(\alpha^{-2} \ell^2(\alpha)).$$

$\square$

Appendix C. Background on regular variation and some technical lemmas about regularly varying functions

Definition C.1. *A measurable function $U : \mathbb{R}_+ \to \mathbb{R}_+$ is regularly varying at ∞ with index $\rho \in \mathbb{R}$ if, for $x > 0$, $\lim_{t \to \infty} U(tx)/U(t) = x^\rho$. We denote this by $U \in RV_\rho$. If $\rho = 0$, we call U slowly varying.*

Proposition C.1. *If $U \in RV_\rho$, then there exists a slowly varying function $\ell \in RV_0$ such that*

$$U(x) = x^\rho \ell(x). \tag{66}$$

Definition C.2. *The de Bruijn conjugate $\ell^\#$ of the slowly varying function ℓ, which always exists, is uniquely defined up to asymptotic equivalence [6, Theorem 1.5.13] by*

$$\ell(x)\ell^\#\{x\ell(x)\} \to 1, \qquad \ell^\#(x)\ell\{x\ell^\#(x)\} \to 1,$$

as $x \to \infty$. Then $(\ell^\#)^\# \sim \ell$. For example, $(\log^a x)^\# \sim \log^{-a} x$ for $a \neq 0$ and $\ell^\#(x) \sim 1/c$ if $\ell(x) \sim c$.

Proposition C.2. *([36, Proposition 0.8, Chapter 0].) If $U \in RV_\rho$, $\rho \in \mathbb{R}$, and the sequences (a_n) and (a_n') satisfy $0 < a_n \to \infty$, $0 < a_n' \to \infty$ and $a_n \sim c a_n'$ for some $0 < c < \infty$, then*

$$U(a_n) \sim c^\rho U(a_n') \text{ as } n \to \infty.$$

Proposition C.3. *([36, Proposition 0.7, p. 21].) Let $U : \mathbb{R}_+ \to \mathbb{R}_+$ be absolutely continuous with density u, so that $U(x) = \int_0^x u(t) dt$. If $U \in RV_\rho$, $\rho \in \mathbb{R}$, and u is monotone, then*

$$\lim_{x \to \infty} \frac{xu(x)}{U(x)} = \rho;$$

furthermore, if $\rho \neq 0$, then $\text{sign}(\rho) u(x) \in RV_{\rho-1}$.

Acknowledgements

Funding information

The project leading to this work received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 834175). At the start of the project, F. Panero was funded by the EPSRC and MRC Centre for Doctoral Training in Statistical Science (grant code EP/L016710/1).

Data access statement

The data simulated to produce the figures of the paper can be found in the code repository https://github.com/francescapanero/OnSparsity_graphex.git.

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

Supplementary material

The supplementary material for this article can be found at https://dx.doi.org/10.1017/apr.2022.75.

References

[1] ALDOUS, D. J. (1981). Representations for partially exchangeable arrays of random variables. *J. Multivariate Anal.* **11**, 581–598.

[2] AYED, F., LEE, J. AND CARON, F. (2019). Beyond the Chinese restaurant and Pitman–Yor processes: statistical models with double power-law behavior. *Proc. Machine Learning Res.* **97**, pp. 395–404.

[3] AYED, F., LEE, J. AND CARON, F. (2020). The normal-generalised gamma-Pareto process: a novel pure-jump Lévy process with flexible tail and jump-activity properties. Preprint. Available at https://arxiv.org/abs/2006.10968.

[4] BICKEL, P. J. AND CHEN, A. (2009). A nonparametric view of network models and Newman–Girvan and other modularities. *Proc. Nat. Acad. Sci. USA* **106**, 21068–21073.

[5] BICKEL, P. J., CHEN, A. AND LEVINA, E. (2011). The method of moments and degree distributions for network models. *Ann. Statist.* **39**, 2280–2301.

[6] BINGHAM, N. H., GOLDIE, C. M. AND TEUGELS, J. L. (1987). *Regular Variation*. Cambridge University Press.

[7] BOLLOBÁS, B., JANSON, S. AND RIORDAN, O. (2007). The phase transition in inhomogeneous random graphs. *Random Structures Algorithms* **31**, 3–122.

[8] BOLLOBÁS, B. AND RIORDAN, O. (2009). Metrics for sparse graphs. In *Surveys in Combinatorics 2009*, eds S. Huczynska, J. Mitchell and C. Roney-Dougal, Cambridge University Press, pp. 211–288.

[9] BORGS, C., CHAYES, J. T., COHN, H. AND HOLDEN, N. (2018). Sparse exchangeable graphs and their limits via graphon processes. *J. Machine Learning Res.* **18**, 1–71.

[10] BORGS, C., CHAYES, J. T., COHN, H. AND VEITCH, V. (2019). Sampling perspectives on sparse exchangeable graphs. *Ann. Prob.* **47**, 2754–2800.

[11] BORGS, C., CHAYES, J. T., DHARA, S. AND SEN, S. (2019). Limits of sparse configuration models and beyond: graphexes and multi-graphexes. Preprint. Available at https://arxiv.org/abs/1907.01605.

[12] CARON, F. AND FOX, E. (2017). Sparse graphs using exchangeable random measures. *J. R. Statist. Soc. B [Statist. Methodology]* **79**, 1295–1366.

[13] CARON, F., PANERO, F. AND ROUSSEAU, J. (2022). On sparsity, power-law, and clustering properties of graphs based of graphex processes: supplementary material. Tech. Rep., University of Oxford.

[14] CHATTERJEE, S. (2015). Matrix estimation by universal singular value thresholding. *Ann. Statist.* **43**, 177–214.

[15] DIACONIS, P. AND JANSON, S. (2008). Graph limits and exchangeable random graphs. *Rend. Mat. Appl.* **28**, 33–61.

[16] GAO, C., LU, Y. AND ZHOU, H. (2015). Rate-optimal graphon estimation. *Ann. Statist.* **43**, 2624–2652.

[17] GNEDIN, A., HANSEN, B. AND PITMAN, J. (2007). Notes on the occupancy problem with infinitely many boxes: general asymptotics and power laws. *Prob. Surveys* **4**, 146–171.

[18] HERLAU, T., SCHMIDT, M. N. AND MØRUP, M. (2016). Completely random measures for modelling block-structured sparse networks. In *NIPS'16: Proceedings of the 30th International Conference on Neural Information Processing Systems*, Association for Computing Machinery, New York, pp. 4267–4275.

[19] HOOVER, D. N. (1979). Relations on probability spaces and arrays of random variables. Preprint, Institute for Advanced Study, Princeton, NJ.

[20] JANSON, S. (2016). Graphons and cut metric on sigma-finite measure spaces. Preprint. Available at https://arxiv.org/abs/1608.01833.

[21] JANSON, S. (2017). On convergence for graphexes. Preprint. Available at https://arxiv.org/abs/1702.06389.

[22] KALLENBERG, O. (1990). Exchangeable random measures in the plane. *J. Theoret. Prob.* **3**, 81–136.

[23] KOLACZYK, E. D. (2009). *Statistical Analysis of Network Data*. Springer, New York.

[24] LAST, G., PECCATI, G. AND SCHULTE, M. (2016). Normal approximation on Poisson spaces: Mehler's formula, second order Poincaré inequalities and stabilization. *Prob. Theory Relat. Fields* **165**, 667–723.

[25] LATOUCHE, P. AND ROBIN, S. (2016). Variational Bayes model averaging for graphon functions and motif frequencies inference in *W*-graph models. *Statist. Comput.* **26**, 1173–1185.

[26] LLOYD, J., ORBANZ, P., GHAHRAMANI, Z. AND ROY, D. (2012). Random function priors for exchangeable arrays with applications to graphs and relational data. In *NIPS'12: Proceedings of the 25th International Conference on Neural Information Processing Systems*, Association for Computing Machinery, New York, pp. 998–1006.

[27] LOÈVE, M. (1977). *Probability Theory I*, 4th edn. Springer, New York.

[28] LOVÁSZ, L. AND SZEGEDY, B. (2006). Limits of dense graph sequences. *J. Combinatorial Theory B* **96**, 933–957.

[29] NAULET, Z., SHARMA, E., VEITCH, V. AND ROY, D. M. (2017). An estimator for the tail-index of graphex processes. Preprint. Available at https://arxiv.org/abs/1712.01745.

[30] NEWMAN, M. E. J. (2010). *Networks: An Introduction*. Oxford University Press.

[31] NOWICKI, K. AND SNIJDERS, T. (2001). Estimation and prediction for stochastic blockstructures. *J. Amer. Statist. Assoc.* **96**, 1077–1087.

[32] ORBANZ, P. AND ROY, D. M. (2015). Bayesian models of graphs, arrays and other exchangeable random structures. *IEEE Trans. Pattern Anal. Machine Intellig.* **37**, 437–461.

[33] PALLA, G., LOVÁSZ, L. AND VICSEK, T. (2010). Multifractal network generator. *Proc. Nat. Acad. Sci. USA* **107**, 7640–7645.

[34] PENROSE, M. (2003). *Random Geometric Graphs*. Oxford University Press.

[35] REITZNER, M. AND SCHULTE, M. (2013). Central limit theorems for *U*-statistics of Poisson point processes. *Ann. Prob.* **41**, 3879–3909.

[36] RESNICK, S. (1987). *Extreme Values, Point Processes and Regular Variation*. Springer, New York.

[37] TODESCHINI, A., MISCOURIDOU, X. AND CARON, F. (2020). Exchangeable random measures for sparse and modular graphs with overlapping communities. *J. R. Statist. Soc. B [Statist. Methodology]* **82**, 487–520.

[38] VEITCH, V. AND ROY, D. M. (2015). The class of random graphs arising from exchangeable random measures. Preprint. Available at https://arxiv.org/abs/1512.03099.

[39] VEITCH, V. AND ROY, D. M. (2019). Sampling and estimation for (sparse) exchangeable graphs. *Ann. Statist.* **47**, 3274–3299.

[40] WILLMOT, G. E. (1990). Asymptotic tail behaviour of Poisson mixtures by applications. *Adv. Appl. Prob.* **22**, 147–159.

[41] WOLFE, P. J. AND OLHEDE, S. C. (2013). Nonparametric graphon estimation. Preprint. Available at https://arxiv.org/abs/1309.5936.