Electronic Journal of Graph Theory and Applications **11** (1) (2023), 97–109

Electronic Journal of
Graph Theory and Applications

Connectivity of Poissonian inhomogeneous random multigraphs

Lorenzo Federico

Luiss Data Lab, Department of Political Sciences, LUISS Guido Carli,
Viale Pola 12, Roma, Italy

lfederico@luiss.it

Abstract

We introduce a model for inhomogeneous random graphs designed to have a lot of flexibility in the assignment of the degree sequence and the individual edge probabilities while remaining tractable. To achieve this we run a Poisson point process over the square $[0,1]^2$, with an intensity proportional to a kernel $W(x,y)$ and identify every couple of vertices of the graph with a subset of the square, adding an edge between them if there is a point in such subset. This ensures unconditional independence among edges and makes many statements much easier to prove in this setting than in other similar models. Here we prove sharpness of the connectivity threshold under mild integrability conditions on $W(x,y)$.

Keywords: graphons, connectivity threshold, inhomogeneous random graphs
Mathematics Subject Classification: 05C40, 05C80, 60C05
DOI: 10.5614/ejgta.2023.11.1.8

1. Introduction and Model Description

In this paper we introduce a model to generate inhomogeneous random multigraphs on n vertices, inspired by the corresponding representations as checkerboard (multi)-

Received: 4 October 2021, Revised: 6 January 2023, Accepted: 29 January 2023.

graphons, in which edges are sampled independently according to two parameters:

- A sequence $(t_n)_{n\geq 2}$ that controls the expected total number of edges in the multigraph.
- A symmetric kernel W, that is, a function $W(x,y) : [0,1]^2 \to \mathbb{R}_{\geq 0}$ such that $W(x,y) = W(y,x)$ that indicates which edges have higher probability to be present.

We define $(G_n(W,t))_{n\geq 2}$ as the sequence of graphs whose vertex and edge sets are created as follows. The vertex set of the graph $G_n(W,t)$ is created from its representation as a checkerboard graphon (see e.g [13]), that is, we define it as $V_n := \{v_i; i \in [n]\}$ and for every $i \in [n]$ we define the interval $S_i := ((i-1)/n, i/n]$. To sample the edge set of $(G_n(W,t))_{n\geq 2}$, we run a Poisson point process over the square $[0,1]$ with intensity $t_n W(x,y)$ and add an edge between v_i and v_j, $i \leq j$, for every point in the square $S_i \times S_j$.

This is equivalent to adding between any couple of vertices $\{v_i, v_j\}$, $i \leq j$, a number of edges distributed as a Poisson random variable, whose parameter λ_{ij} is given by

$$\lambda_{ij} := \int_{S_i \times S_j} t_n W(x,y) dx dy, \tag{1.1}$$

independent of each other. We also define the random graph $\tilde{G}_n(W,t)$ obtained from the multigraph $G_n(W,t)$ by erasing the multiedges and self-loops. In $\tilde{G}_n(W,t)$ every edge $\{v_i, v_j\}$ is present with probability

$$p_{ij} := 1 - \exp\Big\{ - \int_{S_i \times S_j} t_n W(x,y) dx dy \Big\}, \tag{1.2}$$

independent of the others. It is straightforward to see that $\tilde{G}_n(W,t)$ is connected if and only if $G_n(W,t)$ is connected. We define for every subset $F \subseteq [0,1]^2$ and any $q \geq 1$ the space $\mathbb{L}_q(F)$ as the set of the kernels W such that $\int_F W(x,y)^q dy < \infty$ and the space $\mathbb{L}_q^{loc}(F)$ as the set of all kernels W such that for every compact subset $D \subseteq F$, $\int_D W(x,y)^q dy < \infty$. We also define $\|W\|_q = \Big(\int_{[0,1]^2} W(x,y)^q dy \Big)^{1/q}$. We will use the abbreviation *whp* (with high probability) to mean with probability converging to 1. If $W(x,y) \in \mathbb{L}_1([0,1]^2)$, then $t_n \|W\|_1/2$ is the expected number of edges in $G_n(W,t)$, so tuning opportunely the sequence $(t_n)_{n\geq 2}$ we can use this procedure to generate multigraphs with any given density of edges. For a given constant c, taking $t_n = cn^2$ results in a dense graph, while taking $t_n = cn$ we obtain a sparse graph with finite average degree. Note that $\tilde{G}_n(W,t)$ in some special cases is asymptotically equivalent to many well-known models, such as the Erdős-Rényi random graph [5], if the kernel $W(x,y)$ is constant, the Norros-Reittu random graph [12], if $W(x,y) = f(x)f(y)$ for some function $f : [0,1] \to \mathbb{R}_{\geq 0}$, and the stochastic block model [10], if the kernel is piecewise constant.

Note however that there are also several popular models that cannot be expressed in terms of a sequence $\tilde{G}_n(W,t)$ for any kernel W, such as percolation on sparse graphs or random intersection graphs (see [6, 9, 11, 14] for some connectivity results about those models).

This model is closely related to the general inhomogeneous random graph model defined by Bollobás, Janson and Riordan in [2] which is also defined by a kernel $W(x,y)$, but is sampled by placing n points at random iid positions $(x_i)_{i\in[n]}$ on the interval $[0,1]$ and then adding the edge $\{v_i,v_j\}$ to the graph with a probability that is given by some function of $W(x_i,x_j)$. In such setting the connectivity threshold was computed in [4], finding similar results to those we present here, under stricter conditions on the kernel. The main advantage of our definition of the sampling method is that it allows us to use only one layer of randomness, contained entirely in the inhomogeneous spatial Poisson Point Process with intensity $W(x,y)$, instead of two subsequent randomizations first of the vertex set and then of the edge set based on the positions of the vertices. Consequently after we fix the parameters t,W, there is true independence between the existence and multiplicity of different edges in $G_n(t,W)$, while said independence holds in the model defined in [2] only conditionally on the positions $(x_i)_{i\in[n]}$ of the vertices. We see in the proof of the main theorem of this paper how this property, besides being of genuine theoretical interest, makes many arguments much easier, allows us to greatly relax the integrability conditions compared to those required in [4] and sometimes allows for a completely different approach. Another way to build inhomogeneous random graphs are the percolation models introduced by Bollobás, Borgs, Chayes and Riordan in [1], in which the sequence of random graphs is created by removing with uniform probability edges from a sequence of deterministic dense graphs, which converge to a graphon W. In this case, a connectivity threshold cannot be established generally, as connectivity can be determined locally by the properties of a finite number of vertices in the original dense graph (e.g. there could be a few isolated or very low-degree nodes), which would not be captured by the graphon limit.

The fact that we are sampling our graphs from a kernel $W(x,y) : [0,1]^2 \mapsto \mathbb{R}_{\geq 0}$ suggests that this model might converge to some graphon, in the sense described in [3], in the dense regime (i.e. when $t_n = cn^2$). This is the case, with the limit graphon given by $1-\mathrm{e}^{-cW(x,y)}$, as indicated by (1.2).

2. The Main Theorem

In this section we formulate the main theorem of this paper, about the connectivity threshold of the inhomogeneous multigraph we described, and discuss the conditions required to prove it.

We take $t_n = cn\log n$ for some $c \geq 0$, as this is the density scale at which the phase transition for connectivity happens. We define

$$H(x) := \int_{[0,1]} W(x,y)dy, \quad \nu_0 := \operatorname*{ess\,inf}_{[0,1]} H(x), \tag{2.1}$$

where by ess inf we denote the infimum of the set $U_H = \{z : \mu(\{x : H(x) < z\}) = 0\}$, where with μ we denote the Lebesgue measure over $[0,1]^2$. We also require the kernel $W(x,y)$ to be irreducible, which means that there is no set $B \subset [0,1]$ such that $0 < \mu(B) < 1$ and $\int_{B\times B^c} W(x,y)dxdy = 0$. What follows is the main theorem of the present paper:

Theorem 2.1. *Consider a sequence of graphs $G_n(W,t)$, with $t_n = cn\log n$ and $W(x,y)$ irreducible.*

If $c < 1/\nu_0$ and $H(x) \in \mathbb{L}_1^{loc}(F)$ for some open set $F \subseteq [0,1]$ such that $\mu(F) = 1$, then

$$\lim_{n\to\infty} \mathbb{P}(G_n(W,t) \text{ is connected}) = 0. \tag{2.2}$$

If $c > 1/\nu_0$ and $W(x,y) \in \mathbb{L}_q([0,1]^2)$ for some $q > 2$, then

$$\lim_{n\to\infty} \mathbb{P}(G_n(W,t) \text{ is connected}) = 1. \tag{2.3}$$

In other words, under some integrability conditions, the graph is connected whp if there are no vertices with an expected degree lower than $\log n$ and if there are not two sets of vertices which are deterministically separated. We divide the proof in several steps. First we analyze the threshold for the existence of isolated vertices and prove that it coincides with what we claim to be the connectivity threshold. Then, we prove that when $c > \nu_0$ the graph is actually connected, providing two separate arguments for the non existence whp of small and large components.

Note that for the upper bound to hold we require the condition $W(x,y) \in \mathbb{L}_q([0,1]^2)$, which might seem counterintuitive since in most connectivity proofs (see [4, 5, 8]) the most important role is played by the vertices of low degree, while the vertices of high degree are almost irrelevant, since they tend to be always part of the giant component. This condition is necessary because a vertex might have a high expected degree just because it is given a very large number of self loops or multiple edges, which do not actually contribute to connectivity. The $\mathbb{L}_q$ condition is required to ensure that this effect is not too drastic. It is easy to see, by stochastic domination, that if a kernel $W(x,y) \notin \mathbb{L}_q$ can be lower bounded by another kernel $W'(x,y) \in \mathbb{L}_q$ such that $G_n(W',t)$ satisfies the conditions we require for it to be connected whp, then also $G_n(W,t)$ is connected whp.

In this paper we do not discuss what happens if $c = 1/\nu_0$ or more in general if $\lim_{n\to\infty} \frac{t_n}{n\log n} = 1/\nu_0$, because in that regime the asymptotic probability of connectivity of the graph behaves differently based on the specific shape of the kernel W and it is hard to give general formulas stated in term of relatively easy and natural conditions.

3. Connection Probabilities

We first give a simple formula for the probability that a given set of vertices in $G_n(W,t)$ has no outgoing edges. For every $A \subset [n]$ we define the set $B_A \subset [0,1]$ as $B_A := \bigcup_{v_i \in A} S_i$, and the event C_A as the event that all the edges between A and A^c are vacant, i.e., that A is the union of connected components. This is a crucial notion for the present paper, as the graph $G_n(W,t)$ is connected if and only if there is no proper subset A of $[n]$ such that the event C_A happens. Thus, we need a compact formula for the probability of C_A. Define the set $[0,1]^2_x := \{(x,y) \in [0,1]^2 : x < y\}$. By the definition of $G_n(W,t)$ we write, using the symmetry of $W(x,y)$

$$\begin{aligned}\mathbb{P}(C_A) &= \exp\Big\{-t_n \int_{(B_A \times B_A^c \cup B_A^c \times B_A) \cap [0,1]^2_x} W(x,y)dxdy\Big\} \\ &= \exp\Big\{-t_n \int_{B_A \times B_A^c} W(x,y)dxdy\Big\} \qquad (3.1)\end{aligned}$$

Applying the definition of $H(x)$ from (2.1), we rewrite

$$\mathbb{P}(C_A) = \exp\Big\{-t_n\Big(\int_{B_A} H(x)dx - \int_{B_A \times B_A} W(x,y)dydx\Big)\Big\}. \qquad (3.2)$$

4. The Lower Bound: the Number of Isolated Vertices

As in most connectivity proofs, the relevant parameter for connectivity of the graph is the number Y_n of isolated vertices. We first prove that the limit behavior of Y_n is mainly determined by its expectation in a much more general setting, requiring only that edges are sampled independently, without assuming that the edge probabilities are defined using $W(x,y)$ and t_n. Then we compute bounds on $\mathbb{E}[Y_n]$ in the specific case of $G_n(W,t)$.

4.1. Law of large number for isolated vertices

We first define a more general inhomogeneous random graph in which edges are sampled independently but we do not ask for any regularity on the edge probabilities p_{ij}. Given a number n and an array $\mathbf{P} = (p_{ij})_{i<j\le n} \in [0,1]^{\binom{[n]}{2}}$, we define the random graph $G_n(\mathbf{P})$ with vertex set $\{v_i, i \in [n]\}$, in which each edge $e_{ij} := \{v_i, v_j\}$ is present with probability p_{ij} independent of the others.

We will prove that the existence of isolated points in $G_n(\mathbf{P})$ is regulated mainly by the first moment of their number. This result can be deduced from the main theorem of [7] with some effort, seeing the edge addition process as a coupon collector over the vertices, but we provide here a short and direct proof to improve readability of the paper.

The graph $\tilde{G}_n(W,t)$ is a special case of $G_n(\mathbf{P})$ in which the probabilities p_{ij} are defined by (1.2), moreover, every vertex is isolated in $\tilde{G}_n(W,t)$ if and only if it is isolated in $G_n(W,t)$ (for our purposes we consider vertices with only self loops as isolated), so the following theorem can be applied to both models. Define Y_n as the number of isolated vertices in $G_n(\mathbf{P})$. We prove the following result about the concentration of the number of isolated points:

Theorem 4.1. *Consider a sequence of random graphs $G_n(\mathbf{P})$ such that*

$$\lim_{n\to\infty} \mathbb{E}[Y_n]/\log n = \infty,$$

then

$$\lim_{n\to\infty} \mathrm{Var}(Y_n)/\mathbb{E}[Y_n]^2 = 0. \tag{4.1}$$

Proof. We write, defining the event $I_i = \{v_i \text{ is isolated}\}$ for every $i \in [n]$,

$$\begin{aligned}\mathrm{Var}(Y_n) = \mathbb{E}[Y_n^2] - \mathbb{E}[Y_n]^2 &= \sum_{i,j} \mathbb{P}(I_i \cap I_j) - \sum_{i,j} \mathbb{P}(I_i)\mathbb{P}(I_j) \qquad (4.2)\\ &= \sum_{i,j} \mathbb{P}(I_i)\mathbb{P}(I_j \mid I_i) - \sum_{i,j} \mathbb{P}(I_i)\mathbb{P}(I_j) = \sum_i \mathbb{P}(I_i) \sum_j \big(\mathbb{P}(I_j \mid I_i) - \mathbb{P}(I_j)\big).\end{aligned}$$

We take care of the elements of the sum such that $i = j$ with the following upper bound:

$$\sum_i \mathbb{P}(I_i)\big(\mathbb{P}(I_i \mid I_i) - \mathbb{P}(I_i)\big) = \sum_i \mathbb{P}(I_i)(1 - \mathbb{P}(I_i)) \le \sum_i \mathbb{P}(I_i) = \mathbb{E}[Y_n]. \tag{4.3}$$

If $i \neq j$ we note that $\mathbb{P}(I_j \mid I_i) = \mathbb{P}(I_j)/(1-p_{ij})$, so that we can rewrite

$$\mathbb{P}(I_j \mid I_i) - \mathbb{P}(I_j) = \frac{\mathbb{P}(I_j)}{1-p_{ij}} - \mathbb{P}(I_j) = \mathbb{P}(I_j)\frac{p_{ij}}{1-p_{ij}}. \tag{4.4}$$

We thus obtain

$$\mathrm{Var}(Y_n) \le \mathbb{E}[Y_n] + \sum_i \mathbb{P}(I_i) \sum_{j\neq i} \mathbb{P}(I_j)\frac{p_{ij}}{1-p_{ij}}. \tag{4.5}$$

Define the expected degree of the vertex v_i as $\overline{d}_i = \sum_{j\neq i} p_{ij}$. We divide the sum between vertices of low and high expected degree and prove separate bounds for the two cases. We choose as a boundary function for which the computations work out $\overline{d}_i = 3\log n$. To take care of the vertices v_i such that $\overline{d}_i \ge 3\log n$, we obtain

$$\sum_{i:\bar{d}_i\geq 3\log n} \mathbb{P}(I_i)\sum_{j\neq i}\mathbb{P}(I_j)\frac{p_{ij}}{1-p_{ij}} \leq \sum_{i:\bar{d}_i\geq 3\log n}\Big(1-\frac{\bar{d}_i}{n-1}\Big)^{n-1}\sum_{j\neq i}\mathbb{P}(I_j)\frac{p_{ij}}{1-p_{ij}} \\ \leq \mathrm{e}^{-3\log n(1+o(1))}n^2, \tag{4.6}$$

using that for every i, j

$$\mathbb{P}(I_j)\frac{p_{ij}}{1-p_{ij}} = p_{ij}\prod_{h\neq i,j}(1-p_{jh}) \leq 1. \tag{4.7}$$

To control the vertices such that $\bar{d}_i < 3\log n$ instead, we write, for an arbitrary $\varepsilon > 0$,

$$\sum_{i:\bar{d}_i<3\log n} \mathbb{P}(I_i)\sum_{j\neq i}\mathbb{P}(I_j)\frac{p_{ij}}{1-p_{ij}} \tag{4.8}$$
$$= \sum_{i:\bar{d}_i<3\log n}\mathbb{P}(I_i)\Big(\sum_{j:p_{ij}\leq\varepsilon}\mathbb{P}(I_j)\frac{p_{ij}}{1-p_{ij}} + \sum_{j:p_{ij}>\varepsilon}\mathbb{P}(I_j)\frac{p_{ij}}{1-p_{ij}}\Big).$$

We again bound

$$\sum_{i:\bar{d}_i<3\log n}\mathbb{P}(I_i)\sum_{j:p_{ij}\leq\varepsilon}\mathbb{P}(I_j)\frac{p_{ij}}{1-p_{ij}} \leq \sum_i\mathbb{P}(I_i)\sum_j\mathbb{P}(I_j)\frac{\varepsilon}{1-\varepsilon} = \mathbb{E}[Y_n]^2\frac{\varepsilon}{1-\varepsilon}. \tag{4.9}$$

On the other hand, if $\bar{d}_i \leq 3\log n$, then there are at most $(3/\varepsilon)\log n$ distinct js such that $p_{ij} > \varepsilon$, so, using again (4.7),

$$\sum_{i:\bar{d}_i<3\log n}\mathbb{P}(I_i)\sum_{j:p_{ij}>\varepsilon}\mathbb{P}(I_j)\frac{p_{ij}}{1-p_{ij}} \leq \sum_{i:\bar{d}_i<3\log n}\mathbb{P}(I_i)\sum_j\mathbb{1}_{\{p_{ij}>\varepsilon\}} \tag{4.10}$$
$$\leq \mathbb{E}[Y_n](3/\varepsilon)\log n.$$

Consequently, summing (4.6), (4.9) and (4.10) and substituting into (4.5), we obtain that for every $\varepsilon > 0$

$$\mathrm{Var}(Y_n) \leq \mathbb{E}[Y_n] + \mathbb{E}[Y_n]^2\frac{\varepsilon}{1-\varepsilon} + \mathbb{E}[Y_n](3/\varepsilon)\log n + \mathrm{e}^{-3\log n(1+o(1))}n^2, \tag{4.11}$$

so that, since we assumed $\lim_{n\to\infty}\mathbb{E}[Y_n]/\log n = \infty$,

$$\limsup_{n\to\infty}\mathrm{Var}(Y_n)/\mathbb{E}[Y_n]^2 \leq \frac{\varepsilon}{1-\varepsilon}, \tag{4.12}$$

from the fact that ε is arbitrary the claim follows. □

4.2. The expected number of isolated vertices

We now study the asymptotic of $\mathbb{E}[Y_n]$ for $G_n(W,t)$, to prove that the threshold for the existence of isolated vertices is indeed the claimed threshold for connectivity in Theorem 2.1.

We prove that when $c < 1/\nu_0$, $\mathbb{E}[Y_n] \gg \log n$, so that we can apply Theorem 4.1. If $c < 1/\nu_0$, then for some $\varepsilon > 0$, there exists a set $A \subseteq [0,1]$ such that

$$\mu(A) > \varepsilon; \quad \sup_{x\in A} t_n H(x) < (1-\varepsilon)n\log n. \tag{4.13}$$

We next define the sequence of functions $H_n(x)$ as

$$H_n(x) = n\int_{[\lfloor xn\rfloor/n, \lceil xn\rceil/n]} H(x)dx. \tag{4.14}$$

Note that $H_n(x)$ is constant over the intervals $((i-1)/n, i/n)$ and is not properly defined for $x = i/n$ for some i. To solve this issue we extend $H_n(x)$ so that it is left continuous. Since for every n, $\mu(\{i/n; i\in[n]\}) = 0$, this choice does not impact any of the following arguments. We recall that, for every vertex v_i, $I_i = C(\{i\})$. By (1.2), for every $x \in S_i$, recalling (3.2),

$$\mathbb{P}(I_i) \geq \exp\Big\{ -t_n \int_{S_i} H(x)dx\Big\} \geq \mathrm{e}^{-t_n H_n(x)/n}. \tag{4.15}$$

We assumed the existence of an open set $F \subseteq [0,1]$ such that $\mu(F) = 1$ and $H(x) \in \mathbb{L}_1^{loc}(F)$. By Lebesgue's differentiation theorem, we have that $H_n(x) \to H(x)$ almost everywhere in F and thus almost everywhere in $[0,1]$. Consequently, by Egorov's theorem, there exist a set B such that $\mu(B) < \varepsilon/2$ and a number m such that, for every $n > m$,

$$\sup_{[0,1]\setminus B} |H_n(x) - H(x)| < \varepsilon/2. \tag{4.16}$$

Consequently,

$$\mu(A\setminus B) \geq \varepsilon/2; \quad \sup_{A\setminus B} t_n H_n(x) < (1-\varepsilon/2)n\log n. \tag{4.17}$$

We define the set $M_n(\varepsilon) := \{x : t_n H_n(x) < (1-\varepsilon/2)n\log n\}$. We know that for every $n > m$, $\mu(M_n(\varepsilon)) > \varepsilon/2$, and that $M_n(\varepsilon)$ is the disjoint union of intervals of the form $((i-1)/n, i/n]$. We write

$$V_n(\varepsilon) := \{i \in [n] : ((i-1)/n, i/n] \subseteq M_n(\varepsilon)\}, \tag{4.18}$$

for every $n > m$. Using (4.15), we obtain

$$|V_n(\varepsilon)| > n\varepsilon/2; \tag{4.19}$$
$$\min_{i\in V_n(\varepsilon)} \mathbb{P}(I_i) \ge \min_{i\in V_n(\varepsilon)} \mathrm{e}^{-t_n H_n(x)/n} \ge n^{1-\varepsilon/2}. \tag{4.20}$$

Thus, we conclude

$$\mathbb{E}[Y_n] \ge \sum_{i\in V_n(\varepsilon)} \mathbb{P}(I_i) \ge \frac{n\varepsilon}{2} n^{1-\varepsilon/2} \gg \log n, \tag{4.21}$$

and consequently, using Theorem 4.1 we obtain, by Chebyshev's inequality,

$$\mathbb{P}(Y_n = 0) = \mathbb{P}(Y_n \le 0) \le \frac{\mathrm{Var}(Y_n)}{\mathbb{E}[Y_n]^2}, \tag{4.22}$$

so that $\lim_{n\to\infty} \mathbb{P}(Y_n = 0) = 0$.

5. The Upper Bound

In this section we prove the upper bound on the connectivity threshold, that is, that when $c > 1/\nu_0$ the graph is connected whp. The proof is divided in two steps, first we show that whp there are no small components and then that there are not multiple giant components.

5.1. No small components

We next prove that if $c > 1/\nu_0$, then exists an $\varepsilon > 0$ such that whp every component of $G_n(W,t)$ has size at least $n\varepsilon$.

Proposition 5.1. *Consider a sequence of graphs $G_n(W,t)$, with an irreducible kernel $W(x,y) \in \mathbb{L}_q([0,1]^2)$ for some $q > 2$ and $t_n = cn\log n$ with $c > 1/\nu_0$. Then there exists $\varepsilon > 0$ such that*

$$\lim_{n\to\infty} \mathbb{P}\Big(\bigcup_{A:|A|<\varepsilon n} C_A\Big) = 0. \tag{5.1}$$

Proof. We prove the claim using the union bound, that is, computing that

$$\lim_{n\to\infty} \sum_{A:|A|\le\varepsilon n} \mathbb{P}(C_A) = 0. \tag{5.2}$$

We recall the formula for $\mathbb{P}(C_A)$ from (3.2). We lower bound, by the definition of ν_0, using that $\mu(B_A) = |A|/n$,

$$\int_{B_A} H(x)dx \ge \frac{|A|}{n}\nu_0 \tag{5.3}$$

We next bound for all A, using Hölder's inequality,

$$\int_{B_A\times B_A} W(x,y)dydx \le \mu(B_A)^{2-2/q}\Big(\int_{B_A\times B_A} W(x,y)^q dydx\Big)^{1/q}, \tag{5.4}$$

for any $q>1$, so that, for every B_A such that $\mu(B_A)\le\varepsilon$,

$$\int_{B_A\times B_A} W(x,y)dydx \le \frac{|A|}{n}\varepsilon^{1-2/q}\|W\|_q. \tag{5.5}$$

We choose ε, q such that $\varepsilon^{1-2/q}\|W\|_q < \frac{\nu_0-1/c}{2}$, which is possible because of the assumptions we made on $W(x,y)$ in Section 1. Substituting the bounds from (5.3) and (5.5) into (3.2), we obtain, for every A such that $|A|<\varepsilon n$,

$$\begin{aligned}\mathbb{P}(C_A) &\le \exp\Big\{-t_n\Big(\frac{|A|}{n}\nu_0 - \frac{|A|}{n}\frac{\nu_0-1/c}{2}\Big)\Big\} \qquad (5.6)\\ &\le \exp\Big\{-cn\log n\frac{|A|}{n}\frac{\nu_0+1/c}{2}\Big\} = \exp\Big\{-|A|\log n\frac{c\nu_0+1}{2}\Big\} = n^{-|A|(1+\delta)},\end{aligned}$$

for some appropriate $\delta>0$. Finally

$$\lim_{n\to\infty}\sum_{A:|A|<\varepsilon n}\mathbb{P}(C_A) \le \lim_{n\to\infty}\sum_{i=1}^{\varepsilon n}\binom{n}{i}n^{-i(1+\delta)} \le \lim_{n\to\infty}\sum_{i=1}^{\varepsilon n} n^{-i\delta} = 0. \tag{5.7}$$

□

5.2. No multiple giants

Next, we prove that for every $\varepsilon>0$, there cannot be a set of vertices of size at least εn that is not connected to its complementary.

Proposition 5.2. *Consider a sequence of graphs $G_n(W,t)$, with an irreducible kernel $W(x,y)\in\mathbb{L}_q([0,1]^2)$ for some $q>2$ and $t_n=cn\log n$ with $c\nu_0>1$. Then for every $\varepsilon>0$,*

$$\lim_{n\to\infty}\mathbb{P}\Big(\bigcup_{A:\varepsilon n<|A|\le n/2} C_A\Big)=0. \tag{5.8}$$

Proof. Recall the definitions of B_A and C_A given at the beginning of Section 3. Even when $|A|>\varepsilon n$, the equality in (3.1) applies. By definition $\mu(B_A)=|A|/n$. By [1, Lemma 7][1], we know that for every $\varepsilon>0$, if $W(x,y)$ is irreducible,

[1]The result is originally proved for bounded kernels, but if (5.9) holds for the kernel $W'(x,y) := \max\{W(x,y),1\}$ it holds also for $W(x,y)$ by domination, and $W'(x,y)$ is irreducible if and only if $W(x,y)$ is irreducible.

$$\inf_{B:\varepsilon\le\mu(B)\le 1/2)}\int_{B\times B^c} W(x,y)dxdy=\delta(W,\varepsilon)>0. \tag{5.9}$$

This in particular holds for all B_A for A such that $\varepsilon<|A|/n\le 1/2$. Consequently, by (3.1)

$$\begin{aligned}\max_{A:\varepsilon n\le|A|\le n/2)}\mathbb{P}(C_A)&\le\sup_{B:\varepsilon\le\mu(B)\le 1/2)}\exp\Big\{-t_n\int_{B\times B^c}W(x,y)dxdy\Big\}\\&\le \mathrm{e}^{-t_n\delta(W,\varepsilon)}.\end{aligned} \tag{5.10}$$

Thus, we bound using again the first moment method

$$\begin{aligned}\lim_{n\to\infty}\mathbb{P}\Big(\bigcup_{A:\varepsilon n<|A|<n/2}C_A\Big)&\le\lim_{n\to\infty}\sum_{A:\varepsilon n<|A|\le n/2}\mathbb{P}(C_A)\le\lim_{n\to\infty}2^n\mathrm{e}^{-t_n\delta(W,\varepsilon)}\\&=\lim_{n\to\infty}\mathrm{e}^{-n(c\delta(W,\varepsilon)\log n-\log 2)}=0.\end{aligned} \tag{5.11}$$

□

We can finally use all the results we obtained to prove Theorem 2.1.

Proof of Theorem 2.1. We know that

$$\mathbb{P}(G_n(W,t)\text{ is connected})\le\mathbb{P}(Y_n=0), \tag{5.12}$$

so by (4.22) it follows that if $c\le 1/\nu_0$, then $\lim_{n\to\infty}\mathbb{P}(G_n(W,t)\text{ is connected})=0$.

On the other hand, for $G_n(W,t)$ to be disconnected, there must exist a set A of at most $n/2$ vertices such that C_A happens. By Propositions 5.1 ad 5.2, we obtain that for $c>1/\nu_0$,

$$\lim_{n\to\infty}\mathbb{P}\Big(\bigcup_{A:|A|\le n/2}C_A\Big)\le\lim_{n\to\infty}\Big(\mathbb{P}\Big(\bigcup_{A:|A|<\varepsilon n}C_A\Big)+\mathbb{P}\Big(\bigcup_{A:\varepsilon n<|A|\le n/2}C_A\Big)\Big)=0, \tag{5.13}$$

so the claim follows. □

Acknowledgments

The work in this paper is supported by the European Research Council (ERC) through Starting Grant Random Graph, Geometry and Convergence 639046. The author would like to thank Agelos Georgakopoulos for introducing him to the model and Christoforos Panagiotis for the help with some analytic details of the proofs.

References

[1] B. Bollobás, C. Borgs, J. Chayes, and O. Riordan, Percolation on dense graph sequences, *Ann. Probab.*, 38(1):150–183, (2010), https://0-doi-org.pugwash.lib.warwick.ac.uk/10.1214/09-AOP478.

[2] B. Bollobás, S. Janson, and O. Riordan, The phase transition in inhomogeneous random graphs, *Random Structures Algorithms* **31**(1) (2007), 3–122, https://doi.org/10.1002/rsa.20168.

[3] C. Borgs, J.T. Chayes, L. Lovász, V.T. Sós, and K. Vesztergombi, Convergent sequences of dense graphs. I. Subgraph frequencies, metric properties and testing, *Adv. Math.* **219**(6) (2008), 1801–1851. https://0-doi-org.pugwash.lib.warwick.ac.uk/10.1016/j.aim.2008.07.008.

[4] L. Devroye and N. Fraiman, Connectivity of inhomogeneous random graphs, *Random Structures Algorithms*, **45**(3) (2014), 408–420, https://doi.org/10.1002/rsa.20490.

[5] P. Erdős and A.Rényi, On random graphs. I, *Publ. Math. Debrecen* **6** (1959), 290–297.

[6] P. Erdős and J. Spencer, Evolution of the n-cube, *Comput. Math. Appl.* **5**(1) (1979), 33–39.

[7] V. Falgas-Ravry, J. Larsson, and K. Markström, Speed and concentration of the covering time for structured coupon collectors, *arXiv:1601.04455* (2016).

[8] L. Federico and R. van der Hofstad, Critical window for connectivity in the configuration model, *Combin. Probab. Comput.*, **26**(5) (2017), 660–680, https://doi.org/10.1017/S0963548317000177.

[9] J.A. Fill, E.R. Scheinerman, and K.B. Singer-Cohen, Random intersection graphs when $m = \omega(n)$: an equivalence theorem relating the evolution of the $G(n, m, p)$ and $G(n, p)$ models, *Random Structures Algorithms* **16**(2) (2000), 156–176, https://doi.org/10.1002/(SICI)1098-2418(200003)16:2¡156::AID-RSA3¿3.3.CO;2-8.

[10] P.W. Holland, K.B. Laskey, and S. Leinhardt, Stochastic blockmodels: first steps, *Social Networks* **5**(2) (1983), 109–137, https://doi.org/10.1016/0378-8733(83)90021-7.

[11] F. Joos, Random subgraphs in sparse graphs, *SIAM J. Discrete Math.*, **29**(4) (2015), 2350–2360, https://doi.org/10.1137/140976340.

[12] I. Norros and H. Reittu, On a conditionally Poissonian graph process, *Adv. in Appl. Probab.*, **38**(1) (2006), 59–75, https://0-doi-org.pugwash.lib.warwick.ac.uk/10.1239/aap/114393614.

[13] C. Radin and L. Sadun, Phase transitions in a complex network, *J. Phys. A* **46**(30) (2013), 305002, 12, https://doi.org/10.1088/1751-8113/46/30/305002.

[14] K. Rybarczyk, Diameter, connectivity, and phase transition of the uniform random intersection graph, *Discrete Math.* **311**(17) (2011), 1998–2019, https://0-doi-org.pugwash.lib.warwick.ac.uk/10.1016/j.disc.2011.05.029.