J. Appl. Probab. **62,** 298–318 (2025)
*doi:*10.1017/jpr.2024.66

THE *K*-CORE IN PERCOLATED DENSE GRAPH SEQUENCES

ERHAN BAYRAKTAR ,* *University of Michigan*
SUMAN CHAKRABORTY,**
XIN ZHANG,*** *New York University*

Abstract

We determine the order of the k-core in a large class of dense graph sequences. Let G_n be a sequence of undirected, n-vertex graphs with edge weights $\{a^n_{i,j}\}_{i,j\in[n]}$ that converges to a graphon $W\colon [0, 1]^2 \to [0, +\infty)$ in the cut metric. Keeping an edge (i,j) of G_n with probability $a^n_{i,j}/n$ independently, we obtain a sequence of random graphs $G_n(1/n)$. Using a branching process and the theory of dense graph limits, under mild assumptions we obtain the order of the k-core of random graphs $G_n(1/n)$. Our result can also be used to obtain the threshold of appearance of a k-core of order n.

Keywords: Random graphs; branching process; k-core; phase transition; dense graph limit; graphon; percolation

2020 Mathematics Subject Classification: Primary 05C80
Secondary 60C05; 82B43

1. Introduction

For an integer $k \geq 2$, the k-core of a graph G is the largest induced subgraph of G with minimum degree at least k. It was first introduced in [5] to find large k-connected subgraphs, and since then several studies have been devoted to investigating the existence and the order of the k-core. Apart from theoretical interest, the k-core has been applied to the study of social networks [4, 19], graph visualizing [1, 13], and biology [32]. See also [24] for an extensive discussion on its applications. The seminal paper [30] determined the threshold for the appearance of a non-empty k-core in binomial random graphs and uniform random graphs. The order of the k-core has been studied in different random graph ensembles such as binomial random graphs [28], uniformly chosen random graphs and hypergraphs with specified degree sequence [17, 18, 21, 22, 29], the Poisson cloning model [23], and the pairing-allocation model [12]. While almost all the previous work focused on the k-core of homogeneous random graphs, [31] determined the asymptotic order of the k-core for a large class of inhomogeneous random graphs introduced in [7].

In this article we study the asymptotic order of the k-core in random subgraphs of convergent dense graph sequences. Let G_n be a sequence of undirected weighted graphs on n vertices with

Received 29 March 2023; accepted 21 July 2024.
* Postal address: Department of Mathematics, University of Michigan. Email: erhan@umich.edu
** Email address: contact@sumanc.com
*** Postal address: Department of Finance and Risk Engineering, New York University. Email: xz1662@nyu.edu

edge weights $\{a^n_{i,j}\}$ that converges to a graphon W in the cut metric (see Section 2 for the definition of the cut metric). For some $c > 0$, we keep an edge (i,j) of G_n with probability $ca^n_{i,j}/n$ independently, and denote the resulting random graph by $G_n(c/n)$. For any graphon W, we can associate it with a branching process X^W, i.e. the number of children of a particle with type x has a Poisson distribution with parameter $\int W(x, y)\, \mathrm{d}y$ (see Section 2 for a precise definition). As usual, if A_n is a sequence of random variables, we write $A_n = o_{\mathrm{p}}(n)$ if $A_n/n \to 0$ in probability. Under some mild conditions, we show that

$$\text{order of } k\text{-core of } G_n(c/n) = n\mathbb{P}_{X^{cW}}(\mathcal{A}) + o_{\mathrm{p}}(n), \tag{1.1}$$

where $\mathcal{A}$ is the event that the initial particle has at least k children, each of which has at least $k-1$ children, each of which has at least $k-1$ children, and so on, and $\mathbb{P}_{X^W}(\mathcal{A})$ represents the probability of the event $\mathcal{A}$ occurring for the branching process X^{cW}.

Our contribution is two-fold. First, thanks to [27], every dense graph sequence has a convergent subsequence in the cut metric, and therefore our result applies to a large class of dense graph sequences. An important example is quasi-random graphs (see, e.g., [16]). Roughly, they are dense graph sequences that converge to a constant (non-zero) limit, such as Paley graphs (see Remark 2.3). It is worth pointing out that while the construction of various quasi-random graphs involves very different techniques (see [25]), this article first determines the order of the k-core in its random subgraphs up to the first-order term in a unified way. Actually, besides quasi-random graphs, there are plenty of examples of dense random graph models that converge to bounded graphons (see [2, 3, 14, 15]), where our result can be applied to get the size of the k-core. Second, as a by-product of our proof of the main result, for any sequence of kernels W_n satisfying some mild assumptions that converges to W we have

$$\mathbb{P}_{X^{W_n}}(\mathcal{A}) \to \mathbb{P}_{X^W}(\mathcal{A}),$$

a new continuity result concerning branching processes, which we believe is of independent interest. Even though the theory of graph limits has received enormous attention in the last two decades, the only similar result we can find is [8, Theorem 1.9], which concerns with the survival probability of a branching process.

Let us now point out the differences between [31] and the main ideas of our proof. First, while [31] obtains the size of the k-core of random graphs sampled from a kernel W, our result applies to the percolation of an arbitrary sequence G_n that converges to W in the cut metric. Noticeably, our result, together with [8, Lemma 1.6], recovers the main result in [31] for bounded graphons, whereas the result in [31] is not applicable to quasi-random graphs. In other words, our result additionally shows that the size of the k-core in percolated random graphs is stable with respect to the underlying kernel W in the cut metric. Second, the proof of the upper bound in (1.1) is done by upper-bounding the order of the k-core in terms of the probability of the event $\mathcal{A}$. In the case of random graphs generated by W [31] this probability can be directly expressed in terms of W. For us, the approximation of the probability of the event $\mathcal{A}$ is subtle: we first approximate the probability of $\mathcal{A}$ in terms of homomorphism densities of appropriate subgraphs, and then estimate this probability using the fact that homomorphism densities are continuous in the cut metric; see, e.g., [6, 27]. Third, the proof of the lower bound in (1.1) is more delicate. As a first step, we approximate W by a sequence of finitary kernels F_m as in [7]. Then, we show that, for each fixed m, the branching process X^{G_n} associated with G_n contains $X^{(1-\varepsilon_m)F_m}$ as a subset for some ε_m with $0 < \varepsilon_m < 1/m$ when n is large enough. To conclude the lower bound, we prove a continuity property which is non-trivial and relies on the properties of the cut metric, and then invoke a result (minor variant) from [31].

The rest of the paper is organized as follows. In Section 2, we present our main results with some discussions. In Sections 3 and 4, we prove the upper bound and lower bound respectively of the order of the k-core.

2. Main results and discussions

We now recall a few definitions in order to state our results. Denote [0,1] by I, and its Borel sigma-algebra by $\mathcal{B}(I)$. For any symmetric measurable function $W\colon I \times I \to \mathbb{R}$, we define its cut norm via

$$\|W\|_{\square} := \sup_{S,T\in\mathcal{B}(I)} \left| \int_{S\times T} W(u, v)\,\mathrm{d}u\,\mathrm{d}v \right|.$$

Such a symmetric measurable function W is said to be a graphon if it take values in $[0, \infty)$, and the cut metric between two graphons W_1 and W_2 is defined by $d_{\square}(W_1, W_2) := \|W_1 - W_2\|_{\square}$. An undirected finite graph G_n with non-negative adjacency matrix $(a^n_{i,j})^n_{i,j=1}$ can be associated with a graphon in a natural way so that the cut metric gives rise to a proper distance between two finite graphs,

$$W_{G_n}(x, y) = \sum_{i,j=1}^{n} a^n_{i,j}\mathbf{1}_{J^n_i}(x)\mathbf{1}_{J^n_j}(y), \tag{2.1}$$

where $J^n_1 = [0, 1/n]$ and, for $i = 2, 3, \ldots, n$, $J^n_i = ((i-1)/n, i/n]$. We refer readers to [26, Chapter 8] for detailed discussions of the cut metric and (2.1).

Let G_n be a sequence of simple graphs on n vertices with edge weights $\{a^n_{i,j}\}$ that converges to a graphon W with respect to the cut metric. For some $c > 0$, we keep an edge (i,j) of G_n with probability $ca^n_{i,j}/n$ independently, and denote the resulting random graph by $G_n(c/n)$. Here and throughout the paper we assume that the edge weights $a^n_{i,j}$ are uniformly bounded above by $\bar{a} > 0$, and therefore for sufficiently large n we will have $ca^n_{i,j}/n \leq 1$. Since retaining every edge independently is nothing but the bond-percolation on the graph, we call $G_n(c/n)$ a percolated graph sequence (bond-percolation on arbitrary dense graph sequences was first studied in [6]). Note that if we percolate on a dense graph sequence, where the number of edges is of order n^2, the resulting graph becomes sparse, i.e. the number of edges is linear in n. Our aim is to study the order of the k-core of the random graph sequence $G_n(c/n)$.

We will heavily use the branching process X^W associated with the graphon W. The process starts with a single particle with type x_0, which is chosen uniformly from [0,1]. Conditionally on generation t, each member in generation t has children in the next generation, independently of each other and everything else. The number of children with types in a set $A \subset [0, 1]$ is Poisson with parameter $\int_A W(x, y)\,\mathrm{d}y$, and these numbers are independent for disjoint sets.

Let $\mathcal{A}_d$ be the event in X^W that the initial particle has at least k children, each of which has at least $k-1$ children, each of which has at least $k-1$ children, and so on until the dth generation. Define $\mathcal{A} = \cap_{d=1}^{\infty}\mathcal{A}_d$. Let $C_k(G)$ denote the order of the k-core of a graph G. We are now ready to discuss our main result, which provides the asymptotic order of the k-core in percolated dense graph sequences. As usual, if A_n is a sequence of random variables, we write $A_n = o_{\mathrm{p}}(f(n))$ if $A_n/f(n) \to 0$ in probability; if A_n is a sequence of real numbers, we write $A_n = o(f(n))$ if $A_n/f(n) \to 0$. First, we make the following assumption.

Assumption 2.1. *The map $\lambda \to \mathbb{P}_{X^{\lambda W}}(\mathcal{A})$ is continuous from below at $\lambda = c$.*

Theorem 2.1. *Let G_n be a sequence of graphs with non-negative edge weights which are bounded above by a constant $\bar{a} > 0$. Suppose that G_n converges to a graphon W as $n \to \infty$, and that Assumption 2.1 holds. Then*

$$C_k(G_n(c/n)) = n\mathbb{P}_{X^{cW}}(\mathcal{A}) + o_{\mathrm{p}}(n). \tag{2.2}$$

It suffices to prove the case $c = 1$ in Theorem 2.1. To see this, let G_n be a graph with edge weights $\{a^n_{i,j}\}$ and consider another graph G'_n with edge weights $\{ca^n_{i,j}\}$. Therefore, the random subgraphs $G_n(c/n)$ and $G'_n(1/n)$ are equal in distribution. Finally, by our assumption G_n converges to W and thus G'_n converges to cW. The result (2.2) then follows from the result with $c = 1$.

Our proof of (2.2) is divided into two parts, which will be given in the next two sections. We should remark that for the proof of the upper bound, we only need the assumption that the edge weights of G_n are uniformly bounded above by $\bar{a}$ and that $G_n \to W$ in the cut metric. The uniform boundedness of G_n is used in Sections 3.1 and 3.2. Roughly speaking, the probability of event $\mathcal{A}$ can be written as an infinite sum of homomorphism densities. It is known that homomorphism densities are continuous in the cut metric. Thus, to show the inequality $\limsup \mathbb{P}_{X^{W_{G_n}}}(\mathcal{A}) \leq \mathbb{P}_{X^W}(\mathcal{A})$, we need a dominating term provided by $\bar{a}$. It is possible to avoid this assumption if we get the quantitative convergence rate of the homomorphism densities.

Assumption 2.1 is used only in the proof of the lower bound in Section 4, to approximate W by finitary graphons from below.

Remark 2.1. In Theorem 2.1, note that $\mathbb{P}_{X^{cW}}(\mathcal{A})$ could be zero, in which case we will only be able to say that there is no 'giant' k-core (as usual, by 'giant' we mean 'of order linear in n'). From Theorem 2.1 we can also obtain the emergence threshold for the giant k-core from the function $c \to \mathbb{P}_{X^{cW}}(\mathcal{A})$. More precisely, if there is a point $c_0 > 0$ such that for $0 \leq c < c_0$, $\mathbb{P}_{X^{cW}}(\mathcal{A}) = 0$ and for $c > c_0$, $\mathbb{P}_{X^{cW}}(\mathcal{A}) > 0$, then c_0 will be the threshold for the appearance of a giant k-core.

Remark 2.2. It is not difficult to show that the function $\lambda \to \mathbb{P}_{X^{\lambda W}}(\mathcal{A})$ is continuous from above (see [31, Section 3.2]), therefore under Assumption 2.1 we have continuity of $\lambda \to \mathbb{P}_{X^{\lambda W}}(\mathcal{A})$ at $\lambda = c$. We now discuss why it is not possible to get rid of Assumption 2.1 in Theorem 2.1. First, let c_k be the threshold of the appearance of a k-core in the binomial random graph on n vertices with edge probability c/n. Then, from [22, Theorem 1.3(ii)], the assertion in Theorem 2.1 does not hold at the threshold (discontinuity point), which tells us that Assumption 2.1 is optimal. More precisely, at the threshold the k-core is empty with probability bounded away from 0 and 1 for large n.

Secondly, there could be more than one discontinuity point, and they could appear in different (non-trivial) ways. Let us now explain roughly how such discontinuities could appear; we refer the interested reader to the end of [31, Section 3.1] for the details. Consider the graphon

$$W(x, y) = \begin{cases} 2000, & (x, y) \in \left[0, \tfrac{1}{2}\right]^2, \\ 2, & (x, y) \in \left(\tfrac{1}{2}, 1\right]^2, \\ 1/100, & \text{otherwise.} \end{cases}$$

It is not difficult to show that the 3-core first appears in the subgraph induced by the vertices that correspond to the bottom-left part, $\left[0, \frac{1}{2}\right]^2$, of the graphon, and this emerges near $c_3/1000$, where $c = c_3$ is the threshold of appearance of the 3-core in the binomial random graph on n

vertices with edge probability c/n. It could also be shown that $\mathbb{P}_{X^{\lambda W}}(\mathcal{A})$ has another discontinuity near $\lambda = c_3$, i.e. when the subgraph induced by the vertices that correspond to the top-right part, $\left(\frac{1}{2}, 1\right]^2$, of the graphon will have a 3-core. There is another, more straightforward, way in which a discontinuity could appear in $\mathbb{P}_{X^{\lambda W}}(\mathcal{A})$. Consider the graphon

$$W(x, y) = \begin{cases} 2, & (x, y) \in \left[0, \frac{1}{2}\right]^2, \\ 1, & (x, y) \in \left(\frac{1}{2}, 1\right]^2, \\ 0, & \text{otherwise.} \end{cases}$$

It is easy to see that the emergence threshold of the k-core in the subgraph induced by the vertices that correspond to the bottom-left, $\left[0, \frac{1}{2}\right]^2$, part of the graphon and the subgraph induced by the vertices that correspond to the top-right part, $\left(\frac{1}{2}, 1\right]^2$, of the graphon will be different.

Remark 2.3. (*Paley graphs*) Let us give a concrete example where Theorem 2.1 is applicable. Let q be a prime number of the form $4z + 1$ with $z \in \mathbb{N}_+$, $\mathrm{F}_q = \{0, \ldots, q-1\}$ be the finite field, and $\mathrm{F}_q^* = \mathrm{F}_q \setminus \{0\}$. Consider the sequence of Paley graphs G_q, where the vertex set is $V = \mathrm{F}_q$ and the edge set is given by $E = \{(a, b) \in V \times V : a - b \in \mathrm{F}_q^* \times \mathrm{F}_q^*\}$. Due to [20, Example 10.10], G_q converges to the constant graphon $\frac{1}{2}$ in the cut metric as $q \to \infty$. We can therefore apply Theorem 2.1 and conclude that the size of the k-core in $G_q(2c/q)$ is asymptotically the same as the size of the k-core in the binomial random graph with q vertices and parameter c/q (for $c > 0$), except for the threshold of appearance of the k-core. Let us emphasize that the Paley graph is an example of a quasi-random graph, and our result is applicable to any quasi-random graph.

As a by-product of the proof of Theorem 2.1, we also obtain a result regarding branching processes that might be of independent interest.

Proposition 2.1. *Let W_n be a sequence of graphons such that $d_\square(W_n, W) \to 0$. Also, suppose $\lambda \to \mathbb{P}_{X^{\lambda W}}(\mathcal{A})$ is continuous from below at $\lambda = 1$. Then $\mathbb{P}_{X^{W_n}}(\mathcal{A}) \to \mathbb{P}_{X^W}(\mathcal{A})$ as $n \to \infty$.*

Proof. This is proved in Propositions 3.5 and 4.1. □

Note that Proposition 2.1 has the following important consequence. The function $\lambda \to \mathbb{P}_{X^{\lambda W}}(\mathcal{A})$ is non-decreasing, and therefore it can have at most countably many discontinuity points. Hence, for almost every positive c, the next corollary provides a way to approximate the order of the k-core using only G_n.

Corollary 2.1. *Let G_n be a sequence of graphs with non-negative edge weights which are bounded above by a constant $\bar{a} > 0$. Suppose that G_n converges to a graphon W as $n \to \infty$, where $\lambda \to \mathbb{P}_{X^{\lambda W}}(\mathcal{A})$ is continuous at $\lambda = c$. Then*

$$C_k(G_n(c/n)) = n\mathbb{P}_{X^{cW_{G_n}}}(\mathcal{A}) + o_{\mathrm{p}}(n).$$

Proof of Corollary 2.1. The proof is immediate using Theorem 2.1 and Proposition 2.1. □

3. Proof of the upper bound in Theorem 2.1

We will first prove the upper bound, i.e. $C_k(G_n(1/n)) \leq n\mathbb{P}_{X^W}(\mathcal{A}) + o_{\mathrm{p}}(n)$. The idea is as follows: if a vertex v of a graph is in the k-core, then for any $d > 0$ either v has property $\mathcal{A}_d$ or v is contained in a cycle of length smaller than $2d$. Since the probability of occurrence of short

https://doi.org/10.1017/jpr.2024.66 Published online by Cambridge University Press

cycles is small for large enough n, the probability that v is in the k-core is bounded above by the probability of having property $\mathcal{A}_d$, ignoring an $o_{\rm p}(1)$ term. Therefore, to prove the upper bound, we explicitly calculate the probability of event $\mathcal{A}_d$ using homomorphism density, and a tightness argument. Finally, by letting $d \to \infty$, we obtain $C_k(G_n(1/n)) \leq n\mathbb{P}_{X^W}(\mathcal{A}) + o_{\rm p}(n)$. Note that we do not need the continuity assumption to show the upper bound.

Let us construct a branching process ${}^*X^n$ associated with the random graph $G_n(1/n)$. ${}^*X^n$ has n-types of children, $1, 2, \ldots, n$. It starts with a single particle whose type is chosen uniformly from $1, 2, \ldots, n$. Conditioning on generation t, each member of generation t has children in the next generation independently of each other, and of everything else. For any particle of type i, its number of children of type j is distributed as Bernoulli$(a^n_{i,j}/n)$.

With some $\rho_n \geq 1/n$ to be determined, we will also use another branching process where, for any particle of type i, its number of children of type j is distributed as Poisson$(a^n_{i,j}\rho_n)$. This is actually the branching process associated with the graphon $n\rho_n W_{G_n}$, and we denote it by X^{n,ρ_n}. By taking $\rho_n = 1/(n - \bar{a})$, the Poisson branching process X^{n,ρ_n} stochastically dominates, in the first order, ${}^*X^n$ for $n > 3\bar{a}$. To see this, it is sufficient to show the following inequality for any $i, j \in [n]$:

$$\mathbb{P}\big(\text{Poisson}\big(a^n_{i,j}\rho_n\big) > t\big) \geq \mathbb{P}\big(\text{Bernoulli}\big(a^n_{i,j}/n\big) > t\big) \quad \text{for all } t \in \mathbb{Z}.$$

It is trivial for $t \geq 1$ and $t < 0$. We need to check only for $t = 0$. It can be easily verified that the above inequality is equivalent to

$$\frac{n\rho_n\big(1 - \mathrm{e}^{-a^n_{i,j}\rho_n}\big)}{a^n_{i,j}\rho_n} \geq 1.$$

For $n > 3\bar{a}$, we have $a^n_{i,j}\rho_n = a^n_{i,j}/(n - \bar{a}) < \frac{1}{2}$, and hence, according to the Taylor expansion of $\mathrm{e}^{-a^n_{i,j}\rho_n}$,

$$\frac{n\rho_n\big(1 - \mathrm{e}^{-a^n_{i,j}\rho_n}\big)}{a^n_{i,j}\rho_n} > n\rho_n(1 - a^n_{i,j}\rho_n/2) \geq (1 + \bar{a}\rho_n)(1 - \bar{a}\rho_n/2) \geq 1.$$

Note that we can write

$$C_k(G_n(1/n)) = \sum_{v\in[n]} \mathbf{1}\{v \in k\text{-core of } G_n(1/n)\}.$$

If a vertex v is in the k-core, then one of the following two things must be true:

(I) v is in a cycle within the d-neighborhood of v (this implies v is in a cycle of length at most $2d$);

(II) starting from v there is a tree such that v has k neighbors, each of these k neighbors has at least $k - 1$ neighbors, and this happens up to generation d. In this case we say vertex v has property $\mathcal{A}_d$.

Therefore,

$$\begin{aligned} C_k(G_n(1/n)) \leq \sum_{v\in[n]} &\mathbf{1}\{v \text{ is in a cycle of length at most } 2d\} \\ &+ \sum_{v\in[n]} \mathbf{1}\{v \text{ has property } \mathcal{A}_d\} := \text{Term I} + \text{Term II}. \end{aligned} \tag{3.1}$$

https://doi.org/10.1017/jpr.2024.66 Published online by Cambridge University Press

Let V_n be a uniformly random variable on $\{1, 2, \ldots, n\}$, independent of everything else. Then, according to the construction of $G_n(1/n)$ and the fact that X^{n,ρ_n} stochastically dominates ${}^*X^n$, we obtain

$$\begin{aligned}\mathbb{E}(\text{Term II}) &\le n\mathbb{P}({}^*X^n \text{ with root } V_n \text{ has property } \mathcal{A}_d)\\ &\le n\mathbb{P}(X^{n,\rho_n} \text{ with root } V_n \text{ has property } \mathcal{A}_d).\end{aligned} \tag{3.2}$$

Before presenting our first proposition, we state an auxiliary result, the van den Berg–Kesten–Reimer (BKR) inequality (see, e.g., [11]). Consider a product space Ω of finite sets $\Omega_1, \ldots, \Omega_k$, $\Omega = \Omega_1 \times \cdots \times \Omega_k$. Let $\mathcal{F} = 2^\Omega$, and μ be a product of k probability measures $\mu_1, \ldots, \mu_k$. For any configuration $\omega = (\omega_1, \ldots, \omega_k) \in \Omega$, and any subset S of $[k] := \{1, \ldots, k\}$, we define the cylinder $[\omega]_S$ by $[\omega]_S := \{\hat{\omega}\colon \hat{\omega}_i = \omega_i \text{ for all } i \in S\}$. For any two subsets $A, B \subset \Omega$, define $A \circ B := \{\omega\colon \text{there exists some } S = S(\omega) \subset [k] \text{ such that } [\omega]_S \subset A,\ [\omega]_{S^c} \subset B\}$.

Lemma 3.1. *For any product space Ω of finite sets, product probability measure μ on Ω, and $A, B \subset \Omega$, we have the inequality $\mu(A \circ B) \le \mu(A)\mu(B)$.*

In this paper, to apply the BKR inequality we always take $\Omega^n_{i,j} = \{0, 1\}$, $i \neq j \in \{1, \ldots, n\}$, and $\Omega = \prod_{i\neq j\in[n]} \Omega^n_{i,j}$. Then $\omega^n_{i,j} = 1$ means that nodes i and j are linked in the random graph $G_n(1/n)$. According to our construction, we also have $\mu_i(\{1\}) = a^n_{i,j}/n$.

Proposition 3.1. *Let G_n be a sequence of graphs with non-negative edge weights which are bounded above by a constant $\bar{a} > 0$. Then, for any fixed d,*

$$C_k(G_n(1/n)) \le n\mathbb{P}(X^{n,\rho_n} \text{ with root } V_n \text{ has property } \mathcal{A}_d) + o_{\mathrm{p}}(n).$$

Proof. According to (3.1) and (3.2), it suffices to show that Term~II $= \mathbb{E}$(Term~II) $+ o_{\mathrm{p}}(n)$ and Term~I $= o_{\mathrm{p}}(n)$. In the first two steps, we show the concentration of Term II and its variance, and in the last step we prove that Term I is small.

Step 1. For any two independently and uniformly chosen vertices U and V of $G_n(1/n)$,

$$\mathbb{P}(d(U, V) \le 2d) = \frac{1}{n^2}\sum_{u,v\in[n]} \mathbb{P}(d(u, v) \le 2d) = o(1),$$

where $d(\cdot, \cdot)$ is the graph distance. To see this, note that $d(U, V) \le 2d$ implies there is a path from U to V of length at most $2d$. Thus,

$$\mathbb{P}(d(U, V) \le 2d) \le \sum_{i=1}^{2d} \mathbb{P}(\#\{\text{paths of length } i \text{ from } U \text{ to } V\} \ge 1).$$

Using Markov's inequality we get

$$\mathbb{P}(d(U, V) \le 2d) \le \frac{1}{n^2}\sum_{u,v\in[n]}\sum_{i=1}^{2d} \mathbb{E}(\#\{\text{paths of length } i \text{ from } u \text{ to } v\}).$$

We can get a crude upper bound as

$$\mathbb{P}(d(U, V) \le 2d) \le \frac{1}{n^2}\sum_{u,v\in[n]}\sum_{i=1}^{2d} n^{i-1}\left(\frac{\bar{a}}{n}\right)^i = o(1).$$

Step 2. Let $G_n^d[v]$ be the subgraph of $G_n(1/n)$ induced by the vertices within distance d of $v \in [n]$, and define $B_v = \{\text{root } v \text{ has property } \mathcal{A}_d \text{ in} G_n^d[v]\}$. It can be easily verified that

$$\begin{aligned}\mathbb{E}(\text{Term II}^2) &= \sum_{v,v' \in [n]} \mathbb{P}(\text{root } v \text{ and } v' \text{ have property } \mathcal{A}_d) \\ &= \sum_{v \in [n]} \mathbb{P}(B_v) + \sum_{v \neq v'} \mathbb{P}(B_v \cap B_{v'}).\end{aligned} \tag{3.3}$$

For two different vertices v and v', we break the probability into two parts:

$$\mathbb{P}(B_v \cap B_{v'}) = \mathbb{P}(B_v \cap B_{v'},\, d(v,\, v') \leq 2d) + \mathbb{P}(B_v \cap B_{v'},\, d(v,\, v') > 2d).$$

For the second term, it can be easily seen that

$$\{d(v,\, v') > 2d\} \cap B_v \cap B_{v'} \subset \{d(v,\, v') > 2d\} \cap B_v \circ B_{v'}.$$

Therefore,

$$\begin{aligned}\mathbb{P}(B_v \cap B_{v'}) &= \mathbb{P}(d(v,\, v') \leq 2d) + \mathbb{P}(B_v \circ B_{v'},\, d(v,\, v') > 2d) \\ &\leq \mathbb{P}(d(v,\, v') \leq 2d) + \mathbb{P}(B_v \circ B_{v'}).\end{aligned}$$

According to Lemma 3.1, we get $\mathbb{P}(B_v \cap B_{v'}) \leq \mathbb{P}(d(v,\, v') \leq 2d) + \mathbb{P}(B_v)\mathbb{P}(B_{v'})$. Combining this with (3.3), we get

$$\mathbb{E}\big(\text{Term II}^2\big) \leq n^2 \mathbb{P}(d(U,\, V) \leq 2d) + (\mathbb{E}(\text{Term II}))^2 + \sum_{v \in [n]} \big(\mathbb{P}(B_v) - \mathbb{P}(B_v)^2\big).$$

Therefore, using Step 1, we get $\mathbb{V}(\text{Term II}) = o(n^2)$. Using Chebyshev's inequality we conclude that $\text{Term II} = \mathbb{E}(\text{Term II}) + o_{\mathrm{p}}(n)$.

Step 3. We write $C_v := \{v \text{ is in a cycle of length at most } 2d\}$. The first moment of Term I is given by

$$\sum_{v \in [n]} \mathbb{P}(C_v) \leq n \sum_{l=3}^{2d} \frac{(n-1)!}{(n-l)!} \left(\frac{\bar{a}}{n}\right)^l \leq \sum_{l=3}^{2d} \bar{a}^l = o(n), \tag{3.4}$$

where $(n-1)!/(n-l)!$ is the number of all possible cycles of length l that contain v, and $((n-1)!/(n-l)!)(\bar{a}/n)^l$ is an upper bound of the probability that v is in a cycle of length l. Therefore, by Markov's inequality, $\text{Term I} = o_{\mathrm{p}}(n)$. □

3.1. Recursive formula

Let us first fix some notation. For any graphon W, we denote the initial particle of its associated branching process X^W by X_0^W, the type of X_0^W by $T(W)_0$, and the first generation by $X_{\{1\}}^W, \ldots, X_{\{N(W)_0\}}^W$, where $N(W)_0$ is the number of children of X_0^W. For $d \geq 2$, we denote each element in the dth generation by $X_{\{i_1|i_2|\cdots|i_d\}}^W$ if it is the i_dth child of $X_{\{i_1|i_2|\cdots|i_{d-1}\}}^W$. Denote the number of children of $X_{\{i_1|i_2|\cdots|i_d\}}^W$ by $N(W)_{\{i_1|i_2|\cdots|i_d\}}$, and the type of $X_{\{i_1|i_2|\cdots|i_d\}}^W$ by $T(W)_{\{i_1|i_2|\cdots|i_d\}}$. Let $\mathbf{N}(W)^d$ denote the tree formed by the first d generations of X^W.

https://doi.org/10.1017/jpr.2024.66 Published online by Cambridge University Press

Let $\mathbf{K}$ be a tree and $\mathrm{h}(\mathbf{K})$ the height of the tree $\mathbf{K}$. We define $g(x, \mathbf{K}) := \mathbb{P}(\mathbf{N}(W)^{\mathrm{h}(\mathbf{K})} = \mathbf{K} \mid T(W)_0 = x)$. It is clear that $\mathbb{P}(\mathbf{N}(W)^{\mathrm{h}(\mathbf{K})} = \mathbf{K}) = \int g(x, \mathbf{K})\, \mathrm{d}x$.

Proposition 3.2.

$$g(x, \mathbf{K}) = \frac{\mathrm{e}^{-\int W(x,y)\,\mathrm{d}y}}{k_0!} \prod_{j=1}^{k_0} \bigg(\int W(x, y) g(y, \mathbf{K}_{\{j\}})\, \mathrm{d}y \bigg), \tag{3.5}$$

where k_0 is the number of the first generation in $\mathbf{K}$, and $\mathbf{K}_{\{j\}}$ is the subtree of $\mathbf{K}$ whose root is the jth member of the elements of the first generation of $\mathbf{K}$.

Proof. When $\mathrm{h}(\mathbf{K}) = 1$, it can be easily seen that

$$g(x, \mathbf{K}) = \frac{1}{k_0!} \mathrm{e}^{-\int W(x,y)\,\mathrm{d}y} \bigg(\int W(x, y)\, \mathrm{d}y \bigg)^{k_0}.$$

For $\mathrm{h}(\mathbf{K}) \geq 2$, we get

$$\begin{aligned} g(x, \mathbf{K}) &= \mathbb{P}(\mathbf{N}(W)^{\mathrm{h}(\mathbf{K})} = \mathbf{K} \mid T(W)_0 = x) \\ &= \mathbb{P}(N(W)_0 = k_0 \mid T(W)_0 = x) \\ &\quad \times \mathbb{P}(\mathbf{N}(W)^{\mathrm{h}(\mathbf{K}_{\{j\}})} = \mathbf{K}_{\{j\}}, j = 1, \ldots, k_0 \mid N(W)_0 = k_0, T(W)_0 = x) \\ &= g(x, k_0) \int_{y_1} \cdots \int_{y_{k_0}} \prod_{j=1}^{k_0} g(y_j, \mathbf{K}_{\{j\}}) \mathbb{P}(T(W)_{\{j\}} \in \mathrm{d}y_j \mid N(W)_0 = k_0, T(W)_0 = x). \end{aligned}$$

Here we have used $\prod_{j=1}^{k_0} g(y_j, \mathbf{K}_{\{j\}}) \mathbb{P}(T(W)_{\{j\}} \in \mathrm{d}y_j \mid N(W)_0 = k_0, T(W)_0 = x)$ to denote the following density:

$$\prod_{j=1}^{k_0} \mathbb{P}(T(W)_{\{j\}} \in \mathrm{d}y_j \mid N(W)_0 = k_0, T(W)_0 = i) = \frac{\prod_{j=1}^{k_0} W(x, y_j)\, \mathrm{d}y_j}{\big(\int_y W(x, y)\, \mathrm{d}y \big)^{k_0}}.$$

Therefore, we obtain the recursive formula

$$g(x, \mathbf{K}) = \frac{\mathrm{e}^{-\int W(x,y)\,\mathrm{d}y}}{k_0!} \prod_{j=1}^{k_0} \bigg(\int W(x, y) g(y, \mathbf{K}_{\{j\}})\, \mathrm{d}y \bigg). \qquad \square$$

3.2. Convergence

Let W_n be a sequence of graphons such that $d_\square(W_n, W) \to 0$ and $\sup_{n,x,y} W_n(x, y) \leq \bar{a}$ for some positive constant $\bar{a}$. Let X^{W_n} be the associated branching process of W_n, and $g_n(x, \mathbf{K}) = \mathbb{P}(\mathbf{N}(W_n)^{\mathrm{h}(\mathbf{K})} = \mathbf{K} \mid T(W_n)_0 = x)$.

We want to show that, as $n \to \infty$, $\int g_n(x, \mathbf{K})\, \mathrm{d}x \to \int g(x, \mathbf{K})\, \mathrm{d}x$. To see this, for any graphon W, any finite tree T with root 0, and any $x \in [0, 1]$, we define the vertex-prescribed homomorphism density

$$t^x(T, W) = \int_{[0,1]^{|V(T)|-1}} \prod_{i \in E(T)} W(x, x_i) \prod_{ij \in E(T), i,j \geq 1} W(x_i, x_j)\, \mathrm{d}x_1 \cdots \mathrm{d}x_{|V(T)|-1}$$

https://doi.org/10.1017/jpr.2024.66 Published online by Cambridge University Press

and the homomorphism density $t(T, W) = \int_{[0,1]} t^x(T, W)\,\mathrm{d}x$. It is well known that, for finite T, $t(T, W_n) \to t(T, W)$ as long as $d_\square(W_n, W) \to 0$; see, e.g., [9, 10, 27]. We will rewrite $\int g_n(x, \mathbf{K})\,\mathrm{d}x$ and $\int g(x, \mathbf{K})\,\mathrm{d}x$ as $\sum_{m\geq 0} \lambda_m t(T_m, W_n)$ and $\sum_{m\geq 0} \lambda_m t(T_m, W)$ respectively for a sequence of trees T_m.

Proposition 3.3. *For any $\bar{a} > 0$ and any $\mathbf{K}$ with $\mathrm{h}(\mathbf{K}) = d < \infty$, there exists a sequence of finite trees $(T_m)_{m\geq 0}$ and a sequence of real numbers $(\lambda_m)_{m\geq 0}$ such that, for any graphon W with $\sup_{x,y} W(x, y) \leq \bar{a}$,*

(i) $\sum_{m\geq 0} |\lambda_m| \bar{a}^{|E(T_m)|} < +\infty$;

(ii) $g(x, \mathbf{K}) = \sum_{m\geq 0} \lambda_m t^x(T_m, W)$.

Proof. We proceed by induction. For $d = 1$,

$$g(x, \mathbf{K}) = \frac{1}{k_0!} \mathrm{e}^{-\int W(x,y)\,\mathrm{d}y} \left(\int W(x, y)\,\mathrm{d}y \right)^{k_0} = \frac{1}{k_0!} \sum_{m=0} \frac{(-1)^m}{m!} \left(\int W(x, y)\,\mathrm{d}y \right)^{m+k_0},$$

where k_0 is the number of the first generation in $\mathbf{K}$. For any $m \in \mathbb{N}$, take T_m to be an $(m + k_0)$-star, i.e. a tree of height 1 with $(m + k_0)$ leaves. Define $\lambda_m := (-1)^m/(k_0!m!)$. Then it can be easily seen that

$$\sum_{m\geq 0} |\lambda_m| \bar{a}^{|E(T_m)|} = \sum_{m\geq 0} \frac{\bar{a}^{m+k_0}}{k_0!m!} < +\infty, \qquad g(x, k_0) = \sum_{m\geq 0} \lambda_m t^x(T_m, W).$$

Now suppose that our claim is true for any configuration $\mathbf{K}^{d-1}$. Using our recursive formula in (3.5), we expand the exponential term and obtain

$$g(x, \mathbf{K}) = \frac{1}{k_0!} \sum_{m\geq 0} \frac{(-1)^m}{m!} \left(\int W(x, y)\mathrm{d}y \right)^m \prod_{j=1}^{k_0} \left(\int W(x, y) g(y, \mathbf{K}_{\{j\}})\,\mathrm{d}y \right).$$

For each $\mathbf{K}_{\{j\}}$, $j = 1, \ldots, k_0$, we have sequences $\left(\lambda_m^j\right)_{m\geq 0}$, $\left(T_m^j\right)_{m\geq 0}$ such that our claim is satisfied. For each $m = (m_0, m_1, \ldots, m_{k_0}) \in \mathbb{N}^{k_0+1}$, we define

$$\lambda_m = \frac{(-1)^{m_0}}{k_0!m_0!} \prod_{j=1}^{k_0} \lambda_{m_j}^j$$

and tree T_m as in Figure 1. In T_m, m_0 stands for the number of leaves attached to the root, and the additional trees $T_{m_i}^i$, $i = 1, \ldots, k_0$, are linked to the root. It is then clear that

$$\sum_{m\in\mathbb{N}^{k_0+1}} |\lambda_m| \bar{a}^{|E(T_m)|} \leq \sum_{m_0\in\mathbb{N}} \frac{\bar{a}^{k_0+m_0}}{k_0!m_0!} \prod_{j=1}^{k_0} \left(\sum_{m_j\in\mathbb{N}} |\lambda_{m_j}^j| \bar{a}^{\left|E\left(T_{m_j}^j\right)\right|} \right) < +\infty.$$

According to our induction, we have $g(y, \mathbf{K}_{\{j\}}) = \sum_{m_j\geq 0} \lambda_{m_j}^j t^y\left(T_{m_j}^j, W\right)$. Therefore, we obtain

$$g(x, \mathbf{K}) = \sum_{m\in\mathbb{N}^{k_0+1}} \lambda_m \left(\int W(x, y)\,\mathrm{d}y \right)^{m_0} \prod_{j=1}^{k_0} \left(\int W(x, y) t^y\left(T_{m_j}^j, W\right)\mathrm{d}y \right).$$

https://doi.org/10.1017/jpr.2024.66 Published online by Cambridge University Press

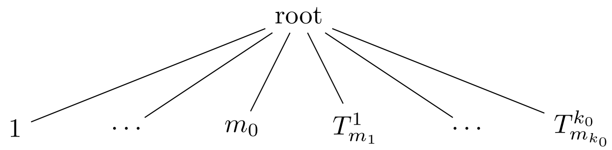

FIGURE 1. Tree T_m.

It can be easily verified that, for each $m \in \mathbb{N}^{k_0+1}$,

$$t^x(T_m, \mathbf{K}) = \left(\int W(x, y)dy \right)^{m_0} \prod_{j=1}^{k_0} \left(\int W(x, y)\, t^y\big(T^j_{m_j}, W\big)\, dy \right).$$

Thus, we conclude that $g(x, \mathbf{K}) = \sum_{m \in \mathbb{N}^{k_0+1}} \lambda_m t^x(T_m, W)$. □

Proposition 3.4. *Suppose W_n is a sequence of graphons such that $d_\square(W_n, W) \to 0$ satisfying $\sup_{n,x,y} W_n(x, y) \le \bar{a}$ for some positive constant $\bar{a}$. Let $\mathbf{K}^d$ be a tree of height d. Then $\lim_{n\to\infty} \mathbb{P}\big(\mathbf{N}(W_n)^d = \mathbf{K}^d\big) = \mathbb{P}(\mathbf{N}(W)^d = \mathbf{K}^d)$.*

Proof. According to Proposition 3.3, we get

$$\mathbb{P}\big(\mathbf{N}(W_n)^d = \mathbf{K}^d\big) = \int g_n\big(x, \mathbf{K}^d\big)\, \mathrm{d}x = \sum_{m \ge 1} \lambda_m t(T_m, W_n),$$
$$\mathbb{P}\big(\mathbf{N}(W)^d = \mathbf{K}^d\big) = \int g\big(x, \mathbf{K}^d\big)\, \mathrm{d}x = \sum_{m \ge 1} \lambda_m t(T_m, W).$$

Since W_n converges to W in that cut norm, $t(T_m, W_n) \to t(T_m, W)$ as $n \to \infty$. Due to the uniform bound

$$\sum_{m \ge 1} \lambda_m t(T_m, W_n) \le \sum_{m \ge 1} \lambda_m \bar{a}^{|E(T_m)|} < +\infty,$$

we apply the dominated convergence theorem and conclude that $\mathbb{P}\big(\mathbf{N}(W_n)^d = \mathbf{K}^d\big)$ converges to $\mathbb{P}\big(\mathbf{N}(W)^d = \mathbf{K}^d\big)$ as $n \to \infty$. □

3.3. Tightness

Notice that $X^W \in \mathcal{A}_d$ is equivalent to $\mathbf{N}(W)^d \in \mathcal{A}_d$. To make our computation clear, we will sometimes adopt the latter notation. Recall that we want to show that

$$\mathbb{P}\big(\mathbf{N}(W)^d \in \mathcal{A}_d\big) = \lim_{n\to\infty} \mathbb{P}\big(\mathbf{N}(W_n)^d \in \mathcal{A}_d\big).$$

To apply Proposition 3.4, we need a tightness result. For $K \in \mathbb{N}$, we define $\mathbf{N}(W)^d \le K$ if $N(W)_{\{i_1|i_2|\cdots|i_j\}} \le K$ for any $X^W_{\{i_1|i_2|\cdots|i_j\}}$ in the first d generations.

Lemma 3.2. *Suppose $\sup_{x,y} W(x, y) \le \bar{a}$ for some positive constant $\bar{a}$. Then, for any $\alpha > 0$, $d \in \mathbb{N}$, there exists a large enough $K_0 \in \mathbb{N}$ uniformly for $x \in [0, 1]$ such that $K \ge K_0$ implies $\mathbb{P}(\mathbf{N}(W)^d \le K \mid T(W)_0 = x) > 1 - (1/K)^\alpha$. Here, the choice of K_0 only depends on α, d, and $\bar{a}$.*

https://doi.org/10.1017/jpr.2024.66 Published online by Cambridge University Press

Proof. We prove the result by induction. Recall that when $\mathrm{h}(\mathbf{K})=1$, we have

$$g(x, \mathbf{K}) = \frac{1}{k_0!}\mathrm{e}^{-\int W(x,y)\,\mathrm{d}y}\left(\int W(x, y)\,\mathrm{d}y\right)^{k_0},$$

where k_0 is the number of the first generation in $\mathbf{K}$. For any $k \in \mathbb{N}$, we define, for $c \in \mathbb{R}_+$, $\psi_k(c) := \sum_{l=k+1}^{\infty} (1/l!)\mathrm{e}^{-c}c^l$. Thus we have

$$\mathbb{P}(N(W)_0 \le k \mid T(W)_0 = x) = 1 - \psi_k\left(\int W(x, y)\,\mathrm{d}y\right).$$

It can be easily verified that $\psi_k'(c) = \mathrm{e}^{-c}c^k/k! \ge 0$, and hence $\psi_k\big(\int W(x, y)\,\mathrm{d}y\big) \le \psi_k(\bar{a})$. Choosing K large enough that $\psi_K(\bar{a}) < K^{-\alpha}$ is indeed possible, it is clear that

$$\mathbb{P}(N(W)_0 \le K \mid T(X^n)_0 = x) = 1 - \psi_K\left(\int W(x, y)\,\mathrm{d}y\right) > 1 - (1/K)^{\alpha}.$$

Assume our claim is true for $d-1$. Then, for any $\beta > 0$, there exists a K such that

$$\mathbb{P}(\mathbf{N}(W)^d_{\{j\}} \le K \mid T(W)^d_{\{j\}} = y) \ge 1 - (1/K)^{\beta}.$$

Note that

$$\begin{aligned}\mathbb{P}\big(\mathbf{N}(W)^d \le K \mid T(W)_0 = x\big) = \sum_{k=0}^{K} &\mathbb{P}\big(\mathbf{N}(W)^d_{\{j\}} \le K, j = 1, \ldots, k \mid N(W)_0 = k, T(W)_0 = x\big)\\ &\times \mathbb{P}(N(W)_0 = k \mid T(W)_0 = x).\end{aligned}$$

As in the proof of Proposition 3.2, we have

$$\begin{aligned}&\mathbb{P}\big(\mathbf{N}(W)^d \le K \mid T(W)_0 = x\big)\\ &\qquad= \sum_{k=0}^{K}\left(\frac{\mathrm{e}^{-\int W(x,y)\,\mathrm{d}y}}{k!}\prod_{j=1}^{k}\left(\int_y W(x, y)\mathbb{P}\big(\mathbf{N}(W)^d_{\{j\}} \le K \mid T(W)_{\{j\}} = y\big)\,\mathrm{d}y\right)\right)\\ &\qquad> \sum_{k=0}^{K}\left(\frac{\mathrm{e}^{-\int W(x,y)\,\mathrm{d}y}}{k!}\prod_{j=1}^{k}\left(\int_y W(x, y)\big(1 - (1/K)^{\beta}\,\mathrm{d}y\big)\right)\right)\\ &\qquad= \sum_{k=0}^{K}\left(\frac{\mathrm{e}^{-\int W(x,y)\,\mathrm{d}y}}{k!}\left(\int W(x, y)\,\mathrm{d}y\right)^k\big(1 - (1/K)^{\beta}\big)^k\right).\end{aligned}$$

Since $(1 - (1/K)^{\beta})^K > 1 - (1/K)^{\beta-2}$ for large K, we have

$$\begin{aligned}\mathbb{P}(\mathbf{N}(W)^d \le K \mid T(W)_0 = x) &> \sum_{k=0}^{K}\left(\frac{\mathrm{e}^{-\int W(x,y)\,\mathrm{d}y}}{k!}\left(\int W(x, y)\,\mathrm{d}y\right)^k\right)\big(1 - (1/K)^{\beta-2}\big)\\ &> (1 - \psi_K(\bar{a}))\big(1 - (1/K)^{\beta-2}\big).\end{aligned}$$

Therefore, by taking $\beta = \alpha + 3$, and large K such that $\psi_K(\bar{a}) < (1/K)^{\alpha+1}$, we conclude that $\mathbb{P}(\mathbf{N}(W)^d \le K \mid T(W)_0 = x) > 1 - (1/K)^{\alpha}$. □

https://doi.org/10.1017/jpr.2024.66 Published online by Cambridge University Press

Proposition 3.5. *Suppose W_n is a sequence of graphons such that $d_\square(W_n, W) \to 0$ and satisfying $\sup_{n,x,y} W_n(x, y) \le \bar{a}$ for some positive constant $\bar{a}$. Then, for any fixed d, we have $\lim_{n\to\infty} \mathbb{P}(\mathbf{N}(W_n)^d \in \mathcal{A}_d) = \mathbb{P}(\mathbf{N}(W)^d \in \mathcal{A}_d)$, from which we conclude that* $\limsup_{n\to\infty} \mathbb{P}_{X^{W_n}}(\mathcal{A}) \le \mathbb{P}_{X^W}(\mathcal{A})$.

Proof. Due to Proposition 3.4, it can be seen that, for fixed d, K,

$$\lim_{n\to\infty} \mathbb{P}\big(\mathbf{N}(W_n)^d \in \mathcal{A}_d, \mathbf{N}(W_n)^d \le K\big) = \mathbb{P}\big(\mathbf{N}(W)^d \in \mathcal{A}_d, \mathbf{N}(W)^d \le K\big).$$

Applying Lemma 3.2, we let $K \to \infty$ and obtain $\lim_{n\to\infty} \mathbb{P}\big(\mathbf{N}(W_n)^d \in \mathcal{A}_d\big) = \mathbb{P}\big(\mathbf{N}(W)^d \in \mathcal{A}_d\big)$.

For any $\epsilon > 0$, there exists a d such that $\mathbb{P}\big(\mathbf{N}(W)^d \in \mathcal{A}_d\big) = \mathbb{P}\big(X^W \in \mathcal{A}_d\big) \le \mathbb{P}_{X^W}(\mathcal{A}) + \epsilon$. It can then be easily verified that

$$\limsup_{n\to\infty} \mathbb{P}_{X^{W_n}}(\mathcal{A}) \le \limsup_{n\to\infty} \mathbb{P}\big(\mathbf{N}(W_n)^d \in \mathcal{A}_d\big) = \mathbb{P}\big(\mathbf{N}(W)^d \in \mathcal{A}_d\big) \le \mathbb{P}_{X^W}(\mathcal{A}) + \epsilon.$$

Therefore, we obtain $\limsup_{n\to\infty} \mathbb{P}_{X^{W_n}}(\mathcal{A}) \le \mathbb{P}_{X^W}(\mathcal{A})$. □

3.4. Completing the proof of the upper bound

Recalling Proposition 3.1, we have $C_k(G_n(1/n) \le n\mathbb{P}(X^{n,\rho_n} \in \mathcal{A}_d) + o_{\mathrm{p}}(n)$. Note that X^{n,ρ_n} is the branching process associated with the graphon $n\rho_n W_{G_n}$, $d_\square(W_{G_n}, W) \to 0$, and $n\rho_n \to 1$. Applying Proposition 3.5 with $W_n = n\rho_n W_{G_n}$, we obtain that $C_k(G_n(1/n)) \le n\mathbb{P}_{X^W}(\mathcal{A}_d) + o_{\mathrm{p}}(n)$. Letting $d \to \infty$ in the above inequality, we have our result.

4. The proof of the lower bound in Theorem 2.1

Before proceeding to our proof of the lower bound, we argue that it suffices to prove it for irreducible W. A graphon W is said to be irreducible if there is no measurable $A \subset [0, 1]$ such that $\mathrm{Leb}(A) \in (0, 1)$ and $W = 0$ almost everywhere on $A \times A^{\mathrm{c}}$, where $\mathrm{Leb}(A)$ stands for the Lebesgue measure of A.

According to [7, Lemma 5.17], for any graphon W, there exists a partition $[0, 1] = \cup_{i=1}^N I_i$ with $0 \le N \le \infty$ such that $\mathrm{Leb}(I_i) > 0$ for each $i \ge 1$, the restriction of W on $I_i \times I_i$ is irreducible for $i \ge 1$, and $W = 0$ almost everywhere on $([0, 1] \times [0, 1]) \setminus \cup_{i=1}^N (I_i \times I_i)$. Let W be the graphon in Theorem 2.1, and let $I_1, \ldots, I_N$ be the corresponding partition. Without loss of generality, we assume that, for each $i \ge 1$, I_i is an interval. Denote the restriction $W|_{I_i\times I_i}$ by W_i. To properly define the branching process of W_i, we identify it with a graphon $\widetilde{W}_i \colon I \times I \to [0, \infty)$ such that $\widetilde{W}_i(x, y) = W(x, y)$ for $x, y \in I_i$ and $\widetilde{W}_i(x, y) = 0$ otherwise. It can then be easily verified that $\mathbb{P}_{X^{\lambda W}}(\mathcal{A}) = \sum_{i=1}^N \mathbb{P}_{X^{\lambda \widetilde{W}_i}}(\mathcal{A})$. Hence, under Assumption 2.1, $\lambda \mapsto \mathbb{P}_{X^{\lambda W_i}}(\mathcal{A})$ is continuous at $\lambda = 1$ for each $i \ge 1$. Fix $N_0 \le N$. There exists a partition $V_1^n, \ldots, V_{N_0}^n, [n] \setminus (\cup_i V_i^n)$, of $[n]$ such that G_n^i converges to $W|_{I_i\times I_i}$ in the cut norm, where G_n^i is the subgraph of G_n induced by V_i^n. *Assuming Theorem 2.1 holds for irreducible graphons*, we obtain, after appropriate normalization of G_n^i and W_i, $C_k(G_n^i(1/n)) \ge n\mathbb{P}_{X^{W_i}}(\mathcal{A}) + o_{\mathrm{p}}(n)$, and thus

$$C_k(G_n(1/n)) \ge \sum_{i=1}^{N_0} C_k(G_n^i(1/n)) \ge n \sum_{i=1}^{N_0} \mathbb{P}_{X^{W_i}}(\mathcal{A}) + o_{\mathrm{p}}(n).$$

Letting $N_0 \to N$, we can conclude the proof for the lower bound. From now on, let us assume that *W is irreducible*.

https://doi.org/10.1017/jpr.2024.66 Published online by Cambridge University Press

We say a graphon F is finitary if there exist finitely many disjoint intervals I_{t_i}, $i = 1, \ldots, M$, such that $\cup_{i=1}^{M} I_{t_i} = [0, 1]$ and the restriction of F on $I_{t_i} \times I_{t_j}$ is a constant for any $1 \leq i, j \leq M$. As we will see, the indexes $t_1, \ldots, t_M$ will be used to label vertices of G_n as well as particles of X^n. According to [7, Lemma 7.3], the graphon W can be approximated pointwise from below by finitary graphons. More precisely, we have the following result.

Lemma 4.1. *There exists a sequence of finitary graphons* $(F_m)_{m\in\mathbb{N}}$ *such that* $F_m \leq W$ *and* $\lim_{m\to\infty} F_m(x, y) = W(x, y)$ *almost surely.*

Under Assumption 2.1, we prove in Section 4.1 that, for any $\varepsilon > 0$, $m \in \mathbb{N}$,

$$C_k(G_n(1/n)) \geq (1 - 2\varepsilon)n\mathbb{P}_{X^{(1-2\varepsilon)F_m}}(\mathcal{A}) + o_{\mathrm{p}}(n). \tag{4.1}$$

Then, in Section 4.2, we show the continuity property

$$\liminf_{\varepsilon\to 0, m\to\infty} \mathbb{P}_{X^{(1-2\varepsilon)F_m}}(\mathcal{A}) \geq \mathbb{P}_{X^W}(\mathcal{A}). \tag{4.2}$$

It is clear that (4.1) and (4.2) together prove the lower bound part of Theorem 2.1, i.e. $C_k(G_n(1/n)) \geq n\mathbb{P}_{X^W}(\mathcal{A}) + o_{\mathrm{p}}(n)$.

4.1. Proof of (4.1)

Fix $m \in \mathbb{N}$ and $\varepsilon \in (0, 1/m)$ such that $\lambda \to \mathbb{P}_{X^{\lambda(1-\varepsilon)F_m}}(\mathcal{A})$ is continuous at $\lambda = 1$. Since F_m is finitary, there exists a partition I_{t_j}, $j = 1, \ldots, M$ of [0,1] such that $F_m|_{I_{t_j}\times I_{t_k}}$ is a constant denoted by $F_m(t_j, t_k)$. We will mark particles of the branching process X^{G_n} by labels t_h, $h = 1, \ldots, M$.

Before proceeding to the rigorous proof, let us first give the main ideas of our argument. For each G_n, we prove there exist M disjoint subsets $\mathbf{Good}_{n,t_1}, \ldots, \mathbf{Good}_{n,t_M}$ of vertices such that, for any vertex $i \in \mathbf{Good}_{n,t_h}$ and $k = 1, \ldots, M$,

$$\frac{\widetilde{d}^n_{i,t_k}}{n} := \sum_{j\in\mathbf{Good}_{n,t_k}} \frac{a^n_{i,j}}{n} \geq (1 - \varepsilon)F_m(t_h, t_k)|I_k|.$$

Therefore, we can heuristically consider G_n as a 'finitary' graph by labelling vertices in $\mathbf{Good}_{n,t_h}$ by t_h, $h = 1, \ldots, M$. Due to the above inequality, the branching process X^{G_n} associated with G_n stochastically dominates, in the first order, the branching process $X^{(1-\varepsilon)F_m}$ associated with $(1 - \varepsilon)F_m$. Take $F^{\varepsilon}_m(1/n)$ to be an n-vertex random graph sampled from $(1 - \varepsilon)F_m$, i.e. independently uniformly select vertices $v_i \in [0, 1]$ and then connect v_i, v_j independently with probability $(1 - \varepsilon)F_m(v_i, v_j)/n$. By the standard exploration argument (see, e.g., [7, Section 9]), locally the random graph $G_n(1/n)$ (respectively $F^{\varepsilon}_m(1/n)$) is almost the branching process X^{G_n} (respectively $X^{(1-\varepsilon)F_m}$). Thus, heuristically the random graph $G_n(1/n)$ is more connected than the random graph $F^{\varepsilon}_m(1/n)$, and thus has a k-core of larger order. Therefore, the inequality (4.1) follows from [31, Theorem 3.1], which says that

$$C_k(F^{\varepsilon}_m(1/n)) = n\mathbb{P}_{X^{(1-\varepsilon)F_m}}(\mathcal{A}) + o_{\mathrm{p}}(n).$$

The following simple lemma will be used to label vertices of G_n. Therein, a vertex is defined to be '**Bad**' if its degree is not bounded from below by that of F_m, and '**Good**' vertices are the complement of '**Bad**' ones.

https://doi.org/10.1017/jpr.2024.66 Published online by Cambridge University Press

Lemma 4.2. *Suppose $\varepsilon > 0$ is a fixed constant, and $\|W_{G_n} - W\|_\square \to 0$. Then there exists a collection of disjoint subsets $\widetilde{\mathbf{Bad}}_{n,t_j} \subset I_{t_j}$, $j = 1, \ldots, M$, such that [(i)]*

(i) $|\widetilde{\mathbf{Bad}}_{n,t_j}| = o(1)$, $j = 1, \ldots, M$;

(ii) *for any* $x \in I_{t_j} \setminus \widetilde{\mathbf{Bad}}_{n,t_j}$,

$$\int_{I_{t_k}} W_{G_n}(x, y)\, \mathrm{d}y \geq (1 - \varepsilon/2) F_m(t_j, t_k)|I_{t_k}|, \qquad k = 1, \ldots, M. \tag{4.3}$$

Proof. First, recall that we can also write

$$\|W\|_\square = \sup_{0 \leq f, g \leq 1 \text{ measurable}} \left| \int f(x) g(y) W(x, y)\, \mathrm{d}x\, \mathrm{d}y \right|. \tag{4.4}$$

For any $1 \leq j, k \leq M$, define

$$\widetilde{\mathbf{Bad}}_{n,t_j,t_k} = \left\{ x \in I_{t_j} \colon \int_{I_{t_k}} W_{G_n}(x, y)\, \mathrm{d}y < (1 - \varepsilon/2) F_m(t_j, t_k)|I_{t_k}| \right\}.$$

In the case that $F_m(t_j, t_k) = 0$, it is clear that $\widetilde{\mathbf{Bad}}_{n,t_j,t_k} = \emptyset$. Let us assume that $F_m(t_j, t_k) > 0$.

Taking $f(x) = \mathbf{1}_{\{x \in \widetilde{\mathbf{Bad}}_{n,t_j,t_k}\}}$ and $g(y) = \mathbf{1}_{\{y \in I_{t_k}\}}$ in (4.4), we obtain

$$\int_{\widetilde{\mathbf{Bad}}_{n,t_j,t_k}} \mathrm{d}x \int_{I_{t_k}} (W_{G_n}(x, y) - W(x, y))\, \mathrm{d}y \geq -\|W_{G_n} - W\|_\square = -o(1),$$

and, since $F_m \leq W$,

$$\int_{\widetilde{\mathbf{Bad}}_{n,t_j,t_k}} \mathrm{d}x \int_{I_{t_k}} (W_{G_n}(x, y) - F_m(x, y))\, \mathrm{d}y \geq -o(1). \tag{4.5}$$

Since, for any $x \in \widetilde{\mathbf{Bad}}_{n,t_j,t_k}$,

$$\int_{I_{t_k}} (W_{G_n}(x, y) - F_m(x, y))\, \mathrm{d}y < -\varepsilon/2 F_m(t_j, t_k)|I_{t_k}|,$$

together with (4.5) it follows that $\varepsilon F_m(t_j, t_k)|I_{t_k}|\, |\widetilde{\mathbf{Bad}}_{n,t_j,t_k}| \leq o(1)$. As a result of $F_m(t_j, t_k) > 0$, we obtain $|\widetilde{\mathbf{Bad}}_{n,t_j,t_k}| = o(1)$.

If we take $\widetilde{\mathbf{Bad}}_{n,t_j} = \bigcup_{k=1}^{M} \widetilde{\mathbf{Bad}}_{n,t_j,t_k}$, it is clear that $|\widetilde{\mathbf{Bad}}_{n,t_j}| = o(1)$ and satisfies (4.3). □

Before proving the main result in this subsection, we would like to point out that our main contribution here is the observation that we can label vertices of G_n so that heuristically it dominates the finitary graphon $(1 - \varepsilon)F_m$. The remaining part of the proof is just a modification of [31, Theorem 3.1]. We summarize it as the following lemma, and refer the reader to [31] for a detailed argument.

Lemma 4.3. *Suppose F_m is a finitary graphon with M labels $t_1, \ldots, t_M$, and $\lambda \to \mathbb{P}_{X^{\lambda F_m}}(\mathcal{A})$ is continuous at $\lambda = 1$. Let G_n be a sequence of graphs such that $\sup_{i,j,n}\{a^n_{i,j}\} < +\infty$. Denote by $X_i^{G_n}$ (respectively $X_{t_h}^{F_m}$) the branching process associated with G_n (respectively F_m) that has the initial particle with type i (respectively label t_h). If, for each G_n, there exist M disjoint subsets $\mathbf{G}_{n,t_h}$, $h = 1, \ldots, M$, of vertices such that, for some $\varepsilon \in (0, 1)$,*

https://doi.org/10.1017/jpr.2024.66 Published online by Cambridge University Press

(i) $|\mathbf{G}_{n,t_h}|/n \geq (1-\varepsilon)|I_{t_h}|$, $h = 1, \ldots, M$*;*

(ii) for each vertex $i \in \mathbf{G}_{n,t_h}$*, the branching process* $X_i^{G_n}$ *stochastically dominates, in the first order, the branching process* $X_{t_h}^{F_m}$*, then* $C_k(G_n(1/n)) \geq (1-\varepsilon)n\mathbb{P}_{X^{(1-\varepsilon)F_m}}(\mathcal{A}) + o_{\mathrm{p}}(n)$*.*

Completing the proof of (4.1). Since $\lambda \to \mathbb{P}_{X^{\lambda F_m}}(\mathcal{A})$ is non-decreasing with respect to λ, it has only countably many discontinuity points. Therefore, we can choose an arbitrarily small ε such that $\lambda \to \mathbb{P}_{X^{\lambda(1-\varepsilon)F_m}}(\mathcal{A})$ is continuous at $\lambda = 1$. We choose $0 < \varepsilon < 1/m$, and take $\widetilde{\mathbf{Bad}}_{n,t_h}$ as in Lemma 4.2. For $h = 1, \ldots, M$, define

$$\mathbf{Good}_{n,t_h} := \{i \in [n] \colon ((i-1)/n, i/n] \subset I_{t_h} \setminus \widetilde{\mathbf{Bad}}_{n,t_h}\}.$$

Due to the construction of W_{G_n} in (2.1), for any $((i-1)/n, i/n] \subset I_{t_h}$, we have either $((i-1)/n, i/n] \subset \widetilde{\mathbf{Bad}}_{n,t_h}$ or $((i-1)/n, i/n] \cap \widetilde{\mathbf{Bad}}_{n,t_h} = \emptyset$. Therefore, it can be easily verified that $|\mathbf{Good}_{n,t_h}|/n \geq |I_{t_h}| - o(1)$. For any $i \in \mathbf{Good}_{n,t_h}$, define $\widetilde{d}^n_{i,t_k} := \sum_{j \in \mathbf{Good}_{n,t_k}} a^n_{i,j}$, $k = 1, \ldots, M$. As a result of (4.3), we obtain

$$\frac{\widetilde{d}^n_{i,t_k}}{n} \geq \int_{I_{t_k}} W_{G_n}(i/n, y)\,\mathrm{d}y - \bar{a}(|I_{t_k}| - |\mathbf{Good}_{n,t_k}|/n) \geq (1-\varepsilon/2)F_m(t_h, t_k)|I_{t_k}| - o(1)\bar{a}.$$

We conclude that, for large enough n, there exists a collection of disjoint $\mathbf{Good}_{n,t_h} \subset [n]$, $h = 1, \ldots, M$, which satisfies the following:

(i) For all $h = 1, \ldots, M$,

$$|\mathbf{Good}_{n,t_h}|/n \geq (1-\varepsilon)|I_{t_h}|. \tag{4.6}$$

(ii) For any $i \in \mathbf{Good}_{n,t_h}$,

$$\frac{\widetilde{d}^n_{i,t_k}}{n} \geq (1-\varepsilon)F_m(t_h, t_k)|I_{t_k}|, \qquad k = 1, \ldots, M. \tag{4.7}$$

For vertices in $\mathbf{Good}_{n,t_h}$, $h = 1, \ldots, M$, we label them with t_h. Let us define $\mathbf{Good}_n := \bigcup_{h=1}^M \mathbf{Good}_{n,t_h}$, $\tilde{n} := |\mathbf{Good}_n|$. Let $\widetilde{G}_{\tilde{n}}$ be the graph with vertices $\mathbf{Good}_n$ such that $\widetilde{a}^n_{i,j} := \tilde{n}a^n_{i,j}/n$ for all $i, j \in \mathbf{Good}_n$. It is clear that $C_k(G_n(1/n)) \geq C_k(\widetilde{G}_{\tilde{n}}(1/\tilde{n}))$.

Take $\widetilde{X}^{\tilde{n}}$ to be a branching process sampled from $\widetilde{G}_{\tilde{n}}$. For any $i \in \mathbf{Good}_{n,t_h}$, take $\widetilde{X}^{\tilde{n}}_i$ to be a branching process sampled from graph $\widetilde{G}_{\tilde{n}}$ with root i. For any t_h, take $X^{(1-\varepsilon)F_m}_{t_h}$ to be a branching process sampled from graphon $(1-\varepsilon)F_m$ with root of label t_h. Suppose a particle in generation t of $\widetilde{X}^{\tilde{n}}_i$ is of type j with label t_h. As a result of (4.7) the number of its t_k-labelled children has Poisson distribution with parameter $\widetilde{d}^n_{j,t_k}/n$, which is larger than $(1-\varepsilon)F_m(t_h, t_k)|I_{t_k}|$. Therefore, for any $i \in \mathbf{Good}_{n,t_h}$, we can consider $X^{(1-\varepsilon)F_m}_{t_h}$ as a subset of $\widetilde{X}^{\tilde{n}}_i$. Thus, for any increasing event $\mathcal{I}$, we have $\mathbb{P}_{\widetilde{X}^{\tilde{n}}_i}(\mathcal{I}) \geq \mathbb{P}_{X^{(1-\varepsilon)F_m}_{t_h}}(\mathcal{I})$.

Now we apply Lemma 4.3 to $(1-\varepsilon)F_m$ and $\widetilde{G}_{\tilde{n}}$ to conclude that

$$\begin{aligned} C_k(G_n(1/n)) \geq C_k(\widetilde{G}_{\tilde{n}}(1/\tilde{n})) &\geq (1-\varepsilon)\tilde{n}\mathbb{P}_{X^{(1-2\varepsilon)F_m}}(\mathcal{A}) + o_{\mathrm{p}}(n) \\ &\geq (1-2\varepsilon)n\mathbb{P}_{X^{(1-2\varepsilon)F_m}}(\mathcal{A}) + o_{\mathrm{p}}(n), \end{aligned}$$

where in the last inequality we used the fact that $\tilde{n}/n \geq 1-\varepsilon$ due to (4.6). □

https://doi.org/10.1017/jpr.2024.66 Published online by Cambridge University Press

4.2. Proof of (4.2)

Note that if F_m converges to W pointwise from below, by the dominated convergence theorem it can be easily seen that $\lim_{\varepsilon\to 0, m\to\infty} d_\square((1-2\varepsilon)F_m, W)=0$. It is therefore sufficient to show that $\lim_{n\to\infty} \mathbb{P}_{X^{W_n}}(\mathcal{A}) \geq \mathbb{P}_{X^W}(\mathcal{A})$ if $\lim_{n\to\infty} d_\square(W_n, W)=0$, which we prove in Proposition 4.1.

We say a branching process has property $\mathcal{B}_d$ if its root has at least $k-1$ children, each of these $k-1$ children has at least $k-1$ children, and this occurs up to generation d, and let $\mathcal{B}=\cap_{d=1}^{\infty}\mathcal{B}_d$. Define the function $\mathbb{P}si_k(\lambda) := \mathbb{P}(\mathrm{Poi}(\lambda)\geq k)$. For any graphon W, define

$$\beta_W(x, d) := \mathbb{P}\big(X^W \in \mathcal{B}_d \mid X_0^W = x\big), \qquad \beta_W(x) := \mathbb{P}\big(X^W \in \mathcal{B} \mid X_0^W = x\big), \tag{4.8}$$

where X_0^W stands for the type of the root. For $W=W_n$, we simply write $\beta_n(x, d) := \beta_{W_n}(x, d)$, $\beta_n(x) := \beta_{W_n}(x)$.

Lemma 4.4. *Let $(W_n)_{n\in\mathbb{N}}$ be a sequence of graphons such that $\|W_n - W\|_\square \to 0$. Suppose that W is irreducible, and $\alpha\colon [0, 1]\to[0, 1]$ is a strictly positive measurable function, i.e.* $\mathrm{Leb}(\{x\colon \alpha(x)>0\})=1$. *Fix $\varepsilon>0$ and $\delta>0$. Define a subset $\mathbf{Bad}_n\subset[0, 1]$ via*

$$\mathbf{Bad}_n v := \left\{x\colon (1-\varepsilon/2)\int \alpha(y)W(x, y)\,\mathrm{d}y > \int \alpha(y)W_n(x, y)\,\mathrm{d}y\right\}.$$

Then $\mathrm{Leb}(\mathbf{Bad}_n)\leq\delta$ *for large enough n.*

Proof. According to the definition of the cut norm (4.4),

$$\|W-W_n\|_\square \geq \int_{\mathbf{Bad}_n}\mathrm{d}x\int \alpha(y)(W(x, y)-W_n(x, y))\,\mathrm{d}y \geq \frac{\varepsilon}{2}\int_{\mathbf{Bad}_n}\mathrm{d}x\int \alpha(y)W(x, y)\,\mathrm{d}y.$$

Note that $\widetilde{W}(x, y)=\alpha(x)\alpha(y)W(x, y)$ is still irreducible, and

$$\|W-W_n\|_\square \geq \frac{\varepsilon}{2}\int_{\mathbf{Bad}_n}\mathrm{d}x\int \alpha(y)W(x, y)\,\mathrm{d}y \geq \frac{\varepsilon}{2}\int_{\mathbf{Bad}_n}\mathrm{d}x\int \widetilde{W}(x, y)\,\mathrm{d}y.$$

Applying [6, Lemma 7] to the irreducible $\widetilde{W}$, we get that, given any $\delta\in\big(0, \frac{1}{2}\big)$, there exists a positive number $b(\widetilde{W}, \delta)>0$ such that, for any $A\subset[0, 1]$ with $\delta\leq\mathrm{Leb}(A)\leq 1-\delta$, $\int_{A\times A^c}\widetilde{W}(x, y)\,\mathrm{d}x\,\mathrm{d}y>b$. Now, since

$$\int_{\mathbf{Bad}_n\times\mathbf{Bad}_n^c}\widetilde{W}(x, y)\,\mathrm{d}x\,\mathrm{d}y \leq \int_{\mathbf{Bad}_n}\mathrm{d}x\int\widetilde{W}(x, y)\,\mathrm{d}y \to 0,$$

we conclude that $\mathrm{Leb}(\mathbf{Bad}_n)\leq\delta$ for large enough n. □

Lemma 4.5. *Let $k\in\mathbb{N}$, and W be an irreducible graphon such that*

$$\alpha(x)=\mathbb{P}si_k\left(\int W(x, y)\alpha(y)\,\mathrm{d}y\right) \tag{4.9}$$

has a solution $\alpha(x)$ with $\mathrm{Leb}(\{x\colon \alpha(x)=0\})<1$. *Then α is strictly positive, i.e.* $\mathrm{Leb}(\{x\colon \alpha(x)>0\})=1$.

https://doi.org/10.1017/jpr.2024.66 Published online by Cambridge University Press

Proof. Suppose α is a non-zero solution of (4.9). Define $\mathcal{Z} := \{x\colon \alpha(x) = 0\}$. For any $x \in \mathcal{Z}$, $0 = \mathbb{P}si_k\big(\int W(x, y)\alpha(y)\,\mathrm{d}y\big)$, which can only happen if $\int_{\mathcal{Z}^c} W(x, y)\alpha(y)\,\mathrm{d}y = 0$. Integrating this equation over $x \in \mathcal{Z}$, we obtain

$$\int_{\mathcal{Z}} \mathrm{d}x \int_{\mathcal{Z}^c} W(x, y)\alpha(y)\,\mathrm{d}y = 0.$$

Since $\alpha > 0$ over $\mathcal{Z}^c$, this violates the irreducibility of W. □

Proposition 4.1. *Let $(W_n)_{n\in\mathbb{N}}$ be a sequence of graphons such that $d_\square(W_n, W) \to 0$ with irreducible W. Then, for any $\varepsilon > 0$,*

$$\mathbb{P}_{X^{W_n}}(\mathcal{A}) \geq \mathbb{P}_{X^{(1-\epsilon)W}}(\mathcal{A}) - \varepsilon^2 \tag{4.10}$$

for large enough n. Then, under Assumption 2.1, $\liminf_{n\to\infty} \mathbb{P}_{X^{W_n}}(\mathcal{A}) \geq \mathbb{P}_{X^W}(\mathcal{A})$.

Proof. We will only prove (4.10), since the second statement follows directly from this. Assume there exists some $\varepsilon_0 > 0$ such that $\mathbb{P}_{X^{(1-\varepsilon_0)W}}(\mathcal{A}) > 0$ (otherwise, there is nothing to prove). Recalling the definition of $\beta_{(1-\varepsilon_0)W}(x)$ as in (4.8), we have the equality

$$\mathbb{P}_{X^{(1-\varepsilon_0)W}}(\mathcal{A}) = \int \mathbb{P}si_k\bigg(\int (1-\varepsilon_0)W(x, y)\beta_{(1-\varepsilon_0)W}(y)\,\mathrm{d}y\bigg)\,\mathrm{d}x,$$

and hence $\mathrm{Leb}\{x\colon \beta_{(1-\varepsilon_0)W}(x) > 0\} > 0$. Since $\beta_{(1-\varepsilon_0)W}(x)$ is a solution of

$$\alpha(x) = \mathbb{P}si_{k-1}\bigg(\int (1-\varepsilon_0)W(x, y)\alpha(y)\,\mathrm{d}y\bigg),$$

$\beta_{(1-\varepsilon_0)W}(x)$ is strictly positive according to Lemma 4.5. Fix any $\varepsilon \in (0, \varepsilon_0)$. We first prove the following statement by induction: for each $d \geq 1$ and $\delta > 0$, there exist subsets $\mathbf{Bad}_{n,d} \subset [0, 1]$ for large enough n such that $\mathrm{Leb}(\mathbf{Bad}_{n,d}) < \delta$ and $\beta_n(x, d) \geq \beta_{(1-\varepsilon)W}(x, d)$, for any $x \in \mathbf{Bad}^c_{n,d}$, $d \geq 1$.

Applying Lemma 4.4 with $\alpha(y) = 1$, for all $y \in [0, 1]$, we obtain some $\mathbf{Bad}_{n,1}$ with $\mathrm{Leb}(\mathbf{Bad}_{n,1}) \leq \delta$ such that $x \in \mathbf{Bad}^c_{n,1}$ implies $\int W_n(x, y)\,\mathrm{d}y \geq (1-\varepsilon/2)\int W(x, y)\,\mathrm{d}y$. It follows that

$$\begin{aligned}\beta_n(x, 1) &= \mathbb{P}si_{k-1}\bigg(\int W_n(x, y)\,\mathrm{d}y\bigg)\\ &\geq \mathbb{P}si_{k-1}\bigg(\bigg(1-\frac{\varepsilon}{2}\bigg)\int W(x, y)\,\mathrm{d}y\bigg)\\ &\geq \mathbb{P}si_{k-1}\bigg((1-\varepsilon)\int W(x, y)\,\mathrm{d}y\bigg) = \beta_{(1-\varepsilon)W}(x, 1).\end{aligned}$$

Assume our assertion holds for $d-1$ and $\delta' > 0$, where δ' is to be chosen later. Let us now prove our claim for d and $\delta > 0$. Noting that

$$\begin{aligned}\beta_n(x, d) &= \mathbb{P}si_{k-1}\bigg(\int W_n(x, y)\beta_n(y, d-1)\,\mathrm{d}y\bigg),\\ \beta_{(1-\varepsilon)W}(x, d) &= \mathbb{P}si_{k-1}\bigg(\int (1-\varepsilon)W(x, y)\beta_{(1-\varepsilon)W}(y, d-1)\,\mathrm{d}y\bigg),\end{aligned}$$

it is enough to show there exists $\mathbf{Bad}_{n,d}$ such that $x \in \mathbf{Bad}^{\mathrm{c}}_{n,d}$ implies

$$\int W_n(x, y)\beta_n(y, d-1)\,\mathrm{d}y \geq \int (1-\varepsilon)W(x, y)\beta_{(1-\varepsilon)W}(y, d-1)\,\mathrm{d}y. \tag{4.11}$$

Applying Lemma 4.4 with $\alpha(y) = \beta_{(1-\varepsilon)W}(y, d-1) > 0$, we obtain, for large enough n, a subset $\widetilde{\mathbf{Bad}}_{n,d}$ with $\mathrm{Leb}(\widetilde{\mathbf{Bad}}_{n,d}) \leq \delta/2$ such that $x \in \widetilde{\mathbf{Bad}}^{\mathrm{c}}_{n,d}$ implies

$$\int W_n(x, y)\beta_{(1-\varepsilon)W}(y, d-1)\,\mathrm{d}y \geq \left(1 - \frac{\varepsilon}{2}\right) \int W(x, y)\beta_{(1-\varepsilon)W}(y, d-1)\,\mathrm{d}y.$$

Since $\beta_{(1-\varepsilon)W}(x)$ is strictly positive, $\lim_{\xi \searrow 0} \mathrm{Leb}(\{x \colon \beta_{(1-\varepsilon)W}(x) \geq \xi\}) = 1$. Therefore, there exists a subset $\mathbf{Bad} \subset [0, 1]$ and a positive constant ξ such that $\mathrm{Leb}(\mathbf{Bad}) \leq \delta/4$ and, for any $x \in \mathbf{Bad}^{\mathrm{c}}$,

$$\int W(x, y)\beta_{(1-\varepsilon)W}(y, d-1)\,\mathrm{d}y \geq \beta_{(1-\varepsilon)W}(x) \geq \xi.$$

By induction, it follows that, for $x \in \mathbf{Bad}^{\mathrm{c}}_{n,d} := \mathbf{Bad}^{\mathrm{c}}_{n,d-1} \cap \widetilde{\mathbf{Bad}}^{\mathrm{c}}_{n,d} \cap \mathbf{Bad}^{\mathrm{c}}$,

$$\begin{aligned}
\int W_n(x, y)\beta_n(y, d-1)\,\mathrm{d}y &\geq \int_{y \in \mathbf{Bad}^{\mathrm{c}}_{n,d-1}} W_n(x, y)\beta_{(1-\varepsilon)W}(y, d-1)\,\mathrm{d}y \\
&\geq \int W_n(x, y)\beta_{(1-\varepsilon)W}(y, d-1)\,\mathrm{d}y - \delta'\bar{a} \\
&\geq \left(1 - \frac{\varepsilon}{2}\right) \int W(x, y)\beta_{(1-\varepsilon)W}(y, d-1)\,\mathrm{d}y - \delta'\bar{a},
\end{aligned}$$

where the first two inequalities are due to our induction hypothesis and the third inequality follows from our choice of $\widetilde{\mathbf{Bad}}_{n,d}$. Choosing a small enough $\delta' \in (0, \delta/4)$ that $\delta'\bar{a} \leq \varepsilon\xi/2$, it is clear that (4.11) holds for $x \in \mathbf{Bad}^{\mathrm{c}}_{n,d}$ with $\mathrm{Leb}(\mathbf{Bad}_{n,d}) \leq \delta$.

Now we prove that $\mathbb{P}_{X^{W_n}}(\mathcal{A}) \geq \mathbb{P}_{X^{(1-\epsilon)W}}(\mathcal{A}) - \varepsilon^2$. Note that

$$\mathbb{P}_{X^{W_n}}(\mathcal{A}_d) = \int \mathbb{P}si_k\left(\int W_n(x, y)\beta_n(y, d-1)\,\mathrm{d}y\right)\mathrm{d}x,$$
$$\mathbb{P}_{X^{(1-\varepsilon)W}}(\mathcal{A}_d) = \int \mathbb{P}si_k\left(\int (1-\varepsilon)W(x, y)\beta_{(1-\varepsilon)W}(y, d-1)\,\mathrm{d}y\right)\mathrm{d}x.$$

Choosing $\delta = \varepsilon^2$, there exists, for large enough n, a subset $\mathbf{Bad}_{n,d}$ with $\mathrm{Leb}(\mathbf{Bad}_{n,d}) \leq \delta$ such that $x \in \mathbf{Bad}^{\mathrm{c}}_{n,d}$ implies

$$\mathbb{P}si_k\left(\int W_n(x, y)\beta_n(y, d-1)\,\mathrm{d}y\right) \geq \mathbb{P}si_k\left(\int (1-\varepsilon)W(x, y)\beta_{(1-\varepsilon)W}(y, d-1)\,\mathrm{d}y\right).$$

Since $\mathbb{P}si_k(x) \leq 1$ for all $x \geq 0$, it can be easily verified that $\mathbb{P}_{X^{W_n}}(\mathcal{A}_d) \geq \mathbb{P}_{X^{(1-\varepsilon)W}}(\mathcal{A}_d) - \varepsilon^2$. Letting $d \to \infty$ in this inequality, we conclude the result. □

Acknowledgements

We wish to thank Remco van der Hofstad for his comments connecting our work to the emergence of a giant k-core, which led to Remark 2.1.

Funding information

E. B. is partially supported by the National Science Foundation under grant DMS-2106556 and by the Susan M. Smith chair. S. C. is partially supported by the Netherlands Organisation for Scientific Research (NWO) through Gravitation-grant NETWORKS-024.002.003.

Competing interests

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

References

[1] Alvarez-Hamelin, J. I., Dall'Asta, L., Barrat, A. and Vespignani, A. (2005). Large scale networks fingerprinting and visualization using the k-core decomposition. In *Proc. 18th Int. Conf. Neural Information Processing Systems*. MIT Press, Cambridge, MA, pp. 41–50.

[2] Basak, A., Bhamidi, S., Chakraborty, S. and Nobel, A. (2016). Large subgraphs in pseudo-random graphs. Preprint, arXiv:1610.03762.

[3] Bhamidi, S., Chakraborty, S., Cranmer, S. and Desmarais, B. (2018). Weighted exponential random graph models: Scope and large network limits. *J. Statist. Phys.* **173**, 704–735.

[4] Bhawalkar, K., Kleinberg, J., Lewi, K., Roughgarden, T. and Sharma, A. (2015). Preventing unraveling in social networks: The anchored k-core problem. *SIAM J. Discrete Math.* **29**, 1452–1475.

[5] Bollobás, B. (1984). The evolution of sparse graphs. In *Graph Theory and Combinatorics*, ed. B. Bollobás. Academic Press, London, pp. 35–57.

[6] Bollobás, B., Borgs, C., Chayes, J. and Riordan, O. (2010). Percolation on dense graph sequences. *Ann. Prob.* **38**, 150–183.

[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., Janson, S. and Riordan, O. (2010). The cut metric, random graphs, and branching processes. *J. Statist. Phys.* **140**, 289–335.

[9] Borgs, C., Chayes, J. T., Lovász, L., Sós, V. T. and Vesztergombi, K. (2008). Convergent sequences of dense graphs I. Subgraph frequencies, metric properties and testing. *Adv. Math.* **219**, 1801–1851.

[10] Borgs, C., Chayes, J. T., Lovász, L., Sós, V. T. and Vesztergombi, K. (2012). Convergent sequences of dense graphs II. Multiway cuts and statistical physics. *Ann. Math.* **176**, 151–219.

[11] Borgs, C., Chayes, J. T. and Randall, D. (1999). The van den Berg–Kesten–Reimer inequality: A review. In *Perplexing Problems in Probability* (Progr. Prob. 44), eds. M. Bramson and R. Durrett. Birkhäuser, Boston, MA, pp. 159–173.

[12] Cain, J. and Wormald, N. (2006). Encores on cores. *Electron. J. Combinatorics* **13**, R81.

[13] Carmi, S., Havlin, S., Kirkpatrick, S., Shavitt, Y. and Shir, E. (2007). A model of internet topology using k-shell decomposition. *Proc. Nat. Acad. Sci.* 104, 11150–11154.

[14] Chatterjee, S. *et al.* (2013). Estimating and understanding exponential random graph models. *Ann. Statist.* **41**, 2428–2461.

[15] Chatterjee, S. *et al.* (2011). Random graphs with a given degree sequence. *Ann. Appl. Prob.* **21**, 1400–1435.

[16] Chung, F. R. K., Graham, R. L. and Wilson, R. M. (1989). Quasi-random graphs. *Combinatorica* **9**, 345–362.

[17] Cooper, C. (2004). The cores of random hypergraphs with a given degree sequence. *Random Structures Algorithms* **25**, 353–375.

[18] Fernholz, D. and Ramachandran, V. (2004). Cores and connectivity in sparse random graphs. UTCS Technical Report TR04-13. University of Texas at Austin.

[19] Huang, X., Cheng, H., Qin, L., Tian, W. and Yu, J. X. (2014). Querying k-truss community in large and dynamic graphs. In *Proc. 2014 ACM SIGMOD Int. Conf. Management of Data*. Association for Computing Machinery, New York, NY, pp. 1311–1322.

[20] Janson, S. (2013). *Graphons, Cut Norm and Distance, Couplings and Rearrangements* (NY *J. Math.* Monographs **4**). State University of New York at Albany.

[21] Janson, S. and Luczak, M. J. (2007). A simple solution to the k-core problem. *Random Structures Algorithms* **30**, 50–62.

[22] Janson, S. and Luczak, M. J. (2008). Asymptotic normality of the k-core in random graphs. *Ann. Appl. Prob.* **18**, 1085–1137.

[23] Kim, J. H. (2007). Poisson cloning model for random graphs. *Expositions Current Math.* **2007**, 104–120.

[24] KONG, Y.-X., SHI, G.-Y., WU, R.-J. AND ZHANG, Y.-C. (2019). k-core: Theories and applications. *Phys. Rep.* **832**, 1–32.

[25] KRIVELEVICH, M. AND SUDAKOV, B. (2006). Pseudo-random graphs. In *More Sets, Graphs and Numbers*, eds E. GYÖRI *et al.* Springer, New York, pp. 199–262.

[26] LOVÁSZ, L. (2012). *Large Networks and Graph Limits* (Am. Math. Soc. Colloq. Pub. 60). American Mathematical Society, Providence, RI.

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

[28] ŁUCZAK, T. (1991). Size and connectivity of the k-core of a random graph. *Discrete Math.* **91**, 61–68.

[29] MOLLOY, M. (2005). Cores in random hypergraphs and boolean formulas. *Random Structures Algorithms* **27**, 124–135.

[30] PITTEL, B., SPENCER, J. AND WORMALD, N. (1996). Sudden emergence of a giant k-core in a random graph. *J. Combinatorial Theory B* **67**, 111–151.

[31] RIORDAN, O. (2008). The k-core and branching processes. *Combinatorics Prob. Comput.* **17**, 111–136.

[32] YOU, Z.-H., LEI, Y.-K., ZHU, L., XIA, J. AND WANG, B. (2013). Prediction of protein–protein interactions from amino acid sequences with ensemble extreme learning machines and principal component analysis. *BMC Bioinformatics* **14**, S10.