On random bipartite graphs evolving by degrees

Neeladri Maitra*

Department of Mathematics
University of Illinois Urbana-Champaign

Abstract

In this paper, we study a bipartite analogue of the 'random graphs evolving by degrees' process. We are given a bipartitioned set of vertices V into two disjoint parts $\mathcal{L}$ and $\mathcal{R}$ and possibly unequal positive constants α and β. The graph evolves starting from B_0, the empty graph (with only isolated vertices). Given B_t, a non-adjacent vertex pair $u \in \mathcal{L}, v \in \mathcal{R}$ is sampled with probability proportional to $(d_u(t) + \alpha)(d_v(t) + \beta)$, and the edge $\{u, v\}$ is included to B_t to form B_{t+1}, where $d_u(t)$ is the degree of u in B_t. For this model, we establish the threshold for the appearance of a giant component, the connectivity threshold for the associated multigraph variant, and provide a superlinear lower bound on the connectivity threshold for the simple graph case.

For the proof of the giant component result, our methods involve setting up an exact coupling of the multigraph case with a bipartite configuration model and using existing results on the giant of bipartite configuration models. This is an adaptation of the technique of Janson and Warnke [8] where they treat the unipartite case similarly. For the connectivity results, we first set up a formula for the exact probability of the occurrence of certain connectivity events in the multigraph process, which is interesting in its own right. To then derive the connectedness threshold, we analyze a particular case of it à la Pittel [10]. For the superlinear connectivity lower bound in the simple graph case, we establish and use a tail bound on the number of isolated vertices in the multigraph process, together with a change of measure statement to go from the multi to the simple graph process.

1 Introduction

This paper is about a random graph process, dynamically evolving over time. Such models are motivated by different perspectives. For example, we observe most real-world networks grow over time, imagine social media networks of friendships, or research article networks, where new articles come into the literature, and one citing the other represent connections between them. Random graph models designed to mimic growing real-world networks include the Barabási-Albert preferential attachment model [2] and related variants; we point the reader to [1, 6] for detailed discussions on the topic. However, sometimes it is also worth looking at variants of existing random graph evolutions, which shed light on the formation of certain structures in them. Consider the work [3] for example, where the authors study a variant of the classical Erdős-Rényi random graph process, with an attempt to *delay* the formation of a giant component. We also refer the reader to the works [9, 11, 12, 3, 10, 8] where variants of the Erdős-Rényi process are studied.

In a by-now celebrated paper [10], Pittel initiates the study of an α-variant of the Erdős-Rényi random graph evolution process, $\alpha > 0$ a constant. In this process, one starts with an empty graph (with only isolated vertices) G_0 on n vertices, and given the graph G_t, a currently non-adjacent vertex pair $\{u, v\}$ is sampled with a probability proportional to $(d_u(t) + \alpha)(d_v(t) + \alpha)$, where

---

*e-mail: nmaitra@illinois.edu

1

$\alpha > 0$ is a fixed real, and the edge $\{u, v\}$ is added to G_t to obtain G_{t+1} from it. Let us informally call this the α-*dynamics*. Observe that this dynamics can be interpreted as a generalization of the classical Erdős-Rényi (E-R) dynamics, where the sampling is instead done uniformly, in that, letting $\alpha \to \infty$ at least intuitively recovers the latter. Indeed, as Pittel [10] proves, many known features of the E-R dynamics are recovered from the analogous α expressions, by letting $\alpha \to \infty$.

There are certain qualitative features of the E-R dynamics that persist when considering bipartite, or in general multipartite variants. For example, in the bipartite case with equal sized partitions, say of size n each, a giant component emerges at a linear scale, namely at around $t = n$ edges (e.g., by [5, Exercise 2.4.7]), which is of the same order as of the giant component emergence threshold for the original unipartite E-R graph, around $n/2$ edges [4]. Note that even though the thresholds are both linear, they differ by a constant factor $1/2$, a consequence of the bipartite structure imposed. Not just the emergence of the giant component, even the connectivity threshold for both these processes appear at the same scale (see [5, Chapter 4]), around $(n \ln n)/2$ edges for the unipartite case, while around $n \ln n$ edges for the bipartite case; the factor $1/2$ difference also persisting here.

For the unipartite case, it is believed that when the number of edges included so far satisfies $t = \Theta(n)$, the final graph of the α-dynamics, in many ways, belong in the same universality class as the final graph of the E-R dynamics. Pittel [10] proves for example, a giant component emerges in the former at around $t = f(\alpha)n$ many edges, with $f(\alpha) = \frac{1}{2(1+\alpha^{-1})} \to 1/2$ as $\alpha \to \infty$, recovering E-R behavior, and in fact the critical window for the appearance of a giant component has width of order $n^{4/3}$ about $f(\alpha)n$ just like the E-R dynamics, see [10, 8]. In contrast, when the number of edges is superlinear, the effect of α should be strong enough to cause a change in behavior from the E-R case. The first instance of a behavior exhibiting this was proved by Pittel, who showed in [10, Theorem 2] that the connectivity threshold for the multigraph version of the α-dynamics occur at about $n^{1+\alpha^{-1}}$ many edges, much larger than $(n \ln n)/2$ for the E-R dynamics. He conjectures that this should also be the connectivity threshold for the simple graph α-dynamics case, something that still stands widely open today. Towards this, we take a small first step by showing that the connectivity threshold for the simple graph case is *strictly superlinear*, in that there is an explicit $\delta > 0$, such that the simple graph is disconnected with probability tending to 1 after the inclusion of at most $n^{1+\delta}$ edges, see Theorem 1.5 below and the discussion in Section 4.1.

Model definition, and some motivating questions. In this paper, we are motivated by the similarity in behavior between the uni and the bipartite case for the E-R dynamics, as discussed above. We study a natural bipartite analogue of the α-dynamics. We are given a set of vertices V and a partition of it into two disjoint parts $\mathcal{L}$ and $\mathcal{R}$. One starts with the empty graph B_0 (with only isolated vertices), and given B_t, one constructs B_{t+1} from it by randomly sampling a pair $\{u, v\}$ of currently non-adjacent vertices with $u \in \mathcal{L}$, $v \in \mathcal{R}$, and including the edge $\{u, v\}$ in B_t. This pair is sampled with a probability proportional to

$$(d_u(t) + \alpha)(d_v(t) + \beta),$$

where α, β are fixed positive reals (possibly unequal). Let us call this the (α, β)-dynamics. We have been driven by an attempt to answer the following questions.

- Does the bipartite (α, β)-dynamics for $\alpha = \beta$ behave similarly to the bipartite E-R dynamics, in that it exhibits some similarity with the unipartite analogue?

- What sort of difference does taking $\alpha \neq \beta$ make?

- Can we prove a result for the *simple* graph (α, β)-dynamics that shows that when the number of edges is superlinear, already the model is quite different from the E-R case with the same number of edges? As remarked before, for the multigraph unipartite version, [10, Theorem 2] establishes this.

- Are there constant factor gaps between the thresholds in the (α, β)-dynamics and its unipartite analogue, like in the E-R case?

We answer some of these questions partially in this paper. To formally state our theorems, we need to set up some notations, which let us get into next.

Some notations. We use the standard order notations $O(\cdot), \Omega(\cdot), o(\cdot), \omega(\cdot), \Theta(\cdot)$ in their usual sense. For non-negative real seqeuences $(x_n)_{n\geq 1}, (y_n)_{n\geq 1}$, we sometimes write $x_n \ll y_n$ which is the same as $x_n = o(y_n)$, and $x_n \gg y_n$, which is the same as $x_n = \omega(y_n)$. We write **whp** for 'with probability tending to 1'. By $\text{Poi}(\lambda)$, we denote a Poisson random variable with mean λ. By $\text{Bin}(n, p)$, we denote a binomial random variable with n trials and success probability p of each trial. By $\text{Exp}(\lambda)$ we define an exponential random variable with rate λ, i.e., with density function $\lambda e^{-\lambda x}$, $x \geq 0$. For a set E, by $|E|$ we simply denote its size.

1.1 Main results

Key assumptions. We work with the following key assumptions throughout the paper.

- Denoting $L_n = |\mathcal{L}|$ and $R_n = |\mathcal{R}|$, where L_n, R_n are sequences diverging to infinity, we assume that the parts $\mathcal{L}$ and $\mathcal{R}$ are *comparable* in size, i.e.,

$$\lim_{n\to\infty} \frac{R_n}{L_n} = \gamma, \text{ for some } \gamma \in \mathbb{R}_+ = (0, \infty). \tag{1.1}$$

- Throughout the note we assume $\alpha, \beta > 0$.

The following definitions are useful to phrase some of our results concisely. Define maps $\rho : \{\mathcal{L}, \mathcal{R}\} \to \{\alpha, \beta\}$ and $\zeta : \{\mathcal{L}, \mathcal{R}\} \to \{1 + \gamma, 1 + 1/\gamma\}$ by

$$\rho(\mathcal{L}) = \alpha, \rho(\mathcal{R}) = \beta, \text{ and } \zeta(\mathcal{L}) = 1 + \gamma, \zeta(\mathcal{R}) = 1 + 1/\gamma. \tag{1.2}$$

When $\alpha \neq \beta$, ρ is invertible, and we denote by ρ^{-1} its inverse map. Similarly when $\gamma \neq 1$, we denote the inverse map of ζ by ζ^{-1}. For any $\epsilon > 0$, define constants

$$p_L(\epsilon) = \frac{\sqrt{\gamma^{-1}(1 + \frac{1}{\alpha})(1 + \frac{1}{\beta})}}{\sqrt{\gamma^{-1}(1 + \frac{1}{\alpha})(1 + \frac{1}{\beta})} + \frac{\alpha}{1+\epsilon}}, \quad p_R(\epsilon) = \frac{\sqrt{\gamma(1 + \frac{1}{\alpha})(1 + \frac{1}{\beta})}}{\sqrt{\gamma(1 + \frac{1}{\alpha})(1 + \frac{1}{\beta})} + \frac{\beta}{1+\epsilon}}. \tag{1.3}$$

1.1.1 Giant component threshold

Theorem 1.1 (Emergence of a giant component). *Consider the bipartite random graph process* $(B_t)_{t\geq 0}$. *Assume* (1.1). *Define*

$$t_c := \frac{\sqrt{\gamma}}{(\gamma + 1)\sqrt{(1 + 1/\alpha)(1 + 1/\beta)}}. \tag{1.4}$$

For any $t \geq 0$, let $\mathcal{C}_1(t)$ and $\mathcal{C}_2(t)$ respectively denote the largest and the second largest connected components of the graph B_t.

- ***Supercritical regime.*** *If $t = t_\epsilon^+ = t_c(1 + \epsilon)(L_n + R_n)$ for some constant $\epsilon > 0$, then*

$$\frac{|\mathcal{C}_1(t_\epsilon^+)|}{L_n + R_n} \xrightarrow{\mathbb{P}} \frac{\xi_L(\epsilon) + \gamma\xi_R(\epsilon)}{1 + \gamma} > 0.$$

To describe the limit, recall (1.3), *and let $\eta_L = \eta_L(\epsilon)$ be the smallest solution to the equation*

$$p_R(\epsilon)\left(\frac{1 - p_L(\epsilon)}{1 - \eta_L p_L(\epsilon)}\right)^{\alpha+1} = 1 - (1 - p_R(\epsilon))(1 - \eta_L)^{-(\beta+1)^{-1}}.$$

Then, writing simply ξ_L and ξ_R for respectively $\xi_L(\epsilon)$ and $\xi_R(\epsilon)$,

$$\xi_L = 1 - \left(\frac{1 - p_L(\epsilon)}{1 - \eta_L p_L(\epsilon)}\right)^\alpha; \quad and \quad \xi_R = 1 - \left(\frac{1 - p_R(\epsilon)}{1 - \eta_R p_R(\epsilon)}\right)^\beta \quad with \quad \eta_R = \left(\frac{1 - p_L(\epsilon)}{1 - p_L(\epsilon)\eta_L}\right)^{\alpha+1}.$$

Furthermore, the second largest component $\mathcal{C}_2(t_\epsilon^+)$ satisfies $\frac{|\mathcal{C}_2(t_\epsilon^+)|}{L_n + R_n} \xrightarrow{\mathbb{P}} 0$.

- **Subcritical regime.** *If $t = t_\epsilon^- = t_c(1 - \epsilon)(L_n + R_n)$ for some $\epsilon > 0$, then $\frac{|\mathcal{C}_1(t_\epsilon^-)|}{L_n + R_n} \xrightarrow{\mathbb{P}} 0$.*

1.1.2 Connectivity properties

Let us introduce a closely related multigraph version $(B_t^*)_{t \geq 0}$ of the process $(B_t)_{t \geq 0}$. At step $t = 0$ we begin from the empty graph B_0^* as usual, and at each step, to construct the graph B_{t+1}^* from B_t^*, we sample an edge $\{u, v\}$ with $u \in \mathcal{L}$ and $v \in \mathcal{R}$ with probability

$$\frac{(d_u(t) + \alpha)(d_v(t) + \beta)}{\sum_{x \in \mathcal{L}, y \in \mathcal{R}}(d_x(t) + \alpha)(d_y(t) + \beta)} = \left(\frac{(d_u(t) + \alpha)}{\sum_{x \in \mathcal{L}}(d_x(t) + \alpha)}\right)\left(\frac{(d_v(t) + \beta)}{\sum_{y \in \mathcal{R}}(d_y(t) + \beta)}\right). \tag{1.7}$$

In contrast, recall that given B_t, to construct B_{t+1}, we add a (yet non-existing) edge between $u \in \mathcal{L}$, $v \in \mathcal{R}$ with probability

$$\frac{(d_u(t) + \alpha)(d_v(t) + \beta)}{\sum_{x \in \mathcal{L}, y \in \mathcal{R}, \{x,y\} \notin E(\mathcal{G}_t)}(d_x(t) + \alpha)(d_y(t) + \beta)}.$$

Thus, in the multigraph process $(B_t^*)_{t \geq 0}$, an edge that has been sampled already before is allowed to be sampled again, potentially giving rise to multiple edges between a pair of vertices, while in the simple process $(B_t)_{t \geq 0}$, such with replacement samplings are barred. Consider the bipartite multigraph process $(B_t^*)_{t \geq 0}$. Recall the maps ρ and ζ from (1.2).

Theorem 1.2 (Multigraph connectedness threshold). *Consider the bipartite multigraph process $(B_t^*)_{t \geq 0}$. Define*

$$\tau_{c,n} := (L_n + R_n)^{1 + (\alpha \wedge \beta)^{-1}}. \tag{1.9}$$

Assume $t = t_n$ is a sequence of n. Then,

$$\lim_{n \to \infty} \mathbb{P}(B_t^* \text{ is connected}) = \begin{cases} 0, & \text{if } t \ll \tau_{c,n}, \\ 1, & \text{if } t \gg \tau_{c,n}. \end{cases}$$

Furthermore, if $\lim_{n \to \infty} \frac{t}{\tau_{c,n}} = x \in (0, \infty)$,

$$\lim_{n \to \infty} \mathbb{P}(B_t^* \text{ is connected}) = \begin{cases} \exp\left(-(\frac{\alpha \wedge \beta}{x})^{\alpha \wedge \beta}[\zeta(\rho^{-1}(\alpha \wedge \beta))]^{-1 - \alpha \wedge \beta}\right), & \text{if } \alpha \neq \beta, \\ \exp\left(-\left(\frac{\alpha}{x}\right)^\alpha \left((1 + \gamma)^{-1 - \alpha} + (1 + 1/\gamma)^{-1 - \alpha}\right)\right), & \text{if } \alpha = \beta. \end{cases} \tag{1.10}$$

Remark 1.3 (Effect of partition sizes). Consider the situation when γ tends to either 0 or ∞, thus one partition is much larger in size than the other one. In this case recalling (1.4), $t_c \to 0$, so that a giant component emerges after the inclusion of $\epsilon(L_n + R_n)$ many edges for any $\epsilon > 0$. On the other hand, note that the connectivity threshold (1.9) is independent of γ. However, looking at (1.10), for $\alpha \neq \beta$, if $\rho^{-1}(\alpha \wedge \beta)$ is the *smaller* partition, so that $\zeta(\rho^{-1}(\alpha \wedge \beta)) \to \infty$, we see that the limit approaches 1, which suggests in this case connectivity is *easier* to achieve when t is of the order of $\tau_{c,n}$.

Following Pittel [10], we leave the following conjecture.

4

Conjecture 1.4 (Simple graph connectedness threshold). *Recall $\tau_{c,n}$ from (1.9). When $\alpha \wedge \beta > 1$, the simple graph $\mathcal{G}_t$ is **whp** connected for $t \gg \tau_{c,n}$, while it is **whp** disconnected for $t \ll \tau_{c,n}$.*

As the simple graph can ever have at most $L_n R_n = \Theta(L_n^2)$ many edges, it is interesting to think about the location of its connectivity threshold when $\alpha \wedge \beta \leq 1$. It is suggestive that the simple graph stays disconnected till the inclusion of the *very last few* edges, when it essentially becomes a complete bipartite graph, see also the discussion by Pittel around the last paragraph of [10, pg-621]. Note that the last result shows that the behavior of the (α, β)-dynamics in the multigraph setting, when the number of edges is superlinear (in the total number of vertices) can be markedly different from the E-R dynamics with the same number of edges, as in the latter case connectivity happens around order $L_n \ln L_n$ many edges, both for the simple and multigraph case. Conjecture 1.4 further suggests that the simple graph (α, β)-dynamics should also exhibit behavior analogous to the multigraph counterpart. Towards this, we ask the following easier question: can we prove *some* result for the simple graph process that shows that the connectivity behavior of it is quite different from the E-R case? Indeed, as the next result shows, this is possible.

Theorem 1.5 (Simple graph connectedness lower bound). *Consider the bipartite simple graph process $(B_t)_{t \geq 0}$. Define*

$$\mathcal{Z}(\alpha, \beta) := \left(\frac{1}{2(1+\alpha)} \wedge \frac{1}{2(1+\beta)} \right) \wedge \left(\left(\frac{1}{4+\alpha} \wedge \frac{1}{\beta} \right) \vee \left(\frac{1}{4+\beta} \wedge \frac{1}{\alpha} \right) \right). \qquad (1.11)$$

*For any $\delta \in (0, \mathcal{Z}(\alpha, \beta))$, the graph $B_{t(\delta)}$ is disconnected **whp**, where $t(\delta) = (L_n + R_n)^{1+\delta} = \Theta(L_n^{1+\delta})$.*

As a consequence of the last theorem, we note that the threshold for connectivity of the simple graph process $(B_t)_{t \geq 0}$ is thus at least (in order) $L_n^{1+\mathcal{Z}(\alpha, \beta)}$ many edges, much larger in order than $L_n \ln L_n$, the corresponding connectivity threshold for the simple random E-R bipartite graph process.

Organization of the rest of the paper. In Section 2 we discuss the strategy and give the proof of the giant component result Theorem 1.1. In Section 3, we provide the necessary required results and the proofs of Theorems 1.2 and 1.5. We close with a small discussion in Section 4.

2 Strategy and proof of emergence of giant

This proof is an adaptation of techniques developed by Janson and Warnke [8] for the unipartite setting.

Approximating with the multigraph process. The first step is to approximate the simple graph process $(B_t)_{t \geq 0}$ with the multigraph process $(B_t^*)_{t \geq 0}$.

Proposition 2.1. *Assume $t = t_n \leq C(L_n + R_n)$ for some constant $C > 0$. Further assume $\alpha, \beta > 0$. Consider a deterministic sequence of bipartite multigraphs $(H_i)_{0 \leq i \leq t}$ on the vertex set $V = \mathcal{L} \cup \mathcal{R}$, with the property that H_{i+1} is obtained from H_i by including a single edge. Then there exists a constant $K = K(C, \alpha, \beta) > 0$ such that as $n \to \infty$,*

$$\mathbb{P}\left(B_i = H_i \ \forall \ 1 \leq i \leq t \right) \leq K \cdot \mathbb{P}\left(B_i^* = H_i \ \forall \ 1 \leq i \leq t \right) + o(1).$$

Thus, when the number of edges is at most linear, rare events for the multigraph process are also rare events for the simple graph process. This result is useful, if we have some control over the multigraph process. Indeed, as we will prove, it is a bipartite configuration model.

Bipartite configuration models. Consider a bipartition $V = \mathcal{L} \cup \mathcal{R}$ of a vertex set V. Assume that we are given non-negative integer sequences $\boldsymbol{D}^{(\mathcal{L})} = (d_u)_{u \in \mathcal{L}}$ and $\boldsymbol{D}^{(\mathcal{R})} = (d_v)_{v \in \mathcal{R}}$, satisfying

$$\sum_{u \in \mathcal{L}} d_u = \sum_{v \in \mathcal{R}} d_v. \tag{2.1}$$

We form a bipartite multigraph with $\boldsymbol{D} = (\boldsymbol{D}_\mathcal{L}, \boldsymbol{D}_\mathcal{R})$ as its bi-degree sequence as follows. Associate to each vertex $u \in V$ a set of d_u many half-edges $H_u := \{l_1^{(u)}, \ldots, l_{d_u}^{(u)}\}$. Take a uniformly random pairing $\mathcal{P}$ of the *left* half-edges $\bigcup_{u \in \mathcal{L}} H_u$ with the *right* half-edges $\bigcup_{v \in \mathcal{R}} H_v$. Given the pairing, to construct a bipartite multigraph on V, let u and v share $n_\mathcal{P}(u, v)$ many edges between them where $n_\mathcal{P}(u, v)$ is the number of elements of the form $(l_i^{(u)}, r_j^{(v)})$ in the pairing $\mathcal{P}$. We call the obtained bipartite multigraph the *bipartite configuration model* corresponding to vertex partitions $\mathcal{L}$ and $\mathcal{R}$, and bi-degree sequence $\boldsymbol{D} = (\boldsymbol{D}_\mathcal{L}, \boldsymbol{D}_\mathcal{R})$, and denote it as $\mathrm{BCM}(\boldsymbol{D}_\mathcal{L}, \boldsymbol{D}_\mathcal{R})$. The key result for the multigraph process $(B_t^*)_{t \geq 0}$ is the following.

Proposition 2.2. *For any $t \geq 0$, and any bi-degree sequence $(\boldsymbol{D}_\mathcal{L}, \boldsymbol{D}_\mathcal{R})$, the bipartite multigraph B_t^* conditioned on having the bi-degree sequence $(\boldsymbol{D}_\mathcal{L}, \boldsymbol{D}_\mathcal{R})$, has the same distribution as a $\mathrm{BCM}(\boldsymbol{D}_\mathcal{L}, \boldsymbol{D}_\mathcal{R})$.*

With the last result at hand, to conclude Theorem 1.1, our task reduces to analyzing the giant component of bipartite configuration models (BCMs), and obtaining some control over the behavior of the bi-degree sequence of $(B_t^*)_{t \geq 0}$, when t is at most linear.

The giant in bipartite configuration models. The giant in bipartite configuration models was analyzed in [7]. We discuss one of their main results in this section. Consider a bi-degree sequence $\boldsymbol{D} = (\boldsymbol{D}_\mathcal{L}, \boldsymbol{D}_\mathcal{R})$ satisfying (2.1), and a $\mathrm{BCM}(\boldsymbol{D}_\mathcal{L}, \boldsymbol{D}_\mathcal{R})$ on the bipartitioned vertex set $V = \mathcal{L} \cup \mathcal{R}$. Let $U_\mathcal{L}$ and $U_\mathcal{R}$ be respectively uniformly sampled elements from $\mathcal{L}$ and $\mathcal{R}$, and let $|\mathcal{L}| = L_n$ and $|\mathcal{R}| = R_n$ be both diverging sequences of n.

Assumption 2.3. [7, Assumption 2.3]. Let $p_n(k) = \mathbb{P}(d_{U_\mathcal{L}} = k)$ and $q_n(k) = \mathbb{P}(d_{U_\mathcal{R}} = k)$.

- There exists limiting random variables $D_\mathcal{L}, D_\mathcal{R}$ with respective probability mass functions $p = (p_0, p_1, \ldots)$ and $q = (q_0, q_1, \ldots)$ such that $p_n(k) \to p_k$ and $q_n(k) \to q_k$ for each $k \geq 0$. In other words, $d_{U_\mathcal{L}} \xrightarrow{d} D_\mathcal{L}$ and $d_{U_\mathcal{R}} \xrightarrow{d} D_\mathcal{R}$.

- Furthermore, $\mathbb{E}[D_\mathcal{L}], \mathbb{E}[D_\mathcal{R}] < \infty$, and $\mathbb{E}[d_{U_\mathcal{L}}] \to \mathbb{E}[D_\mathcal{L}]$ and $\mathbb{E}[d_{U_\mathcal{R}}] \to \mathbb{E}[D_\mathcal{R}]$.

To state the next result, we need the following useful definitions. For a random variable X, denote for $z \in [0, 1]$ its probability generating function evaluated at z as

$$G_X(z) := \mathbb{E}\left[z^X\right].$$

For a $\mathbb{Z}_{>0} := \{1, 2, 3, \ldots\}$ valued random variable X, let us introduce the *shifted size-biased* random variable $\underline{X}$ which has mass function given by,

$$\mathbb{P}(\underline{X} = k) = \frac{(k+1)\mathbb{P}(X = k+1)}{\mathbb{E}[X]}, \quad \text{for } k \in \mathbb{Z}_{\geq 0} := \{0, 1, 2, \ldots\}. \tag{2.3}$$

We can now state the main result on the largest component of BCMs. Denote by $\mathcal{C}_1$ and $\mathcal{C}_2$ respectively the first and second largest components of a $\mathrm{BCM}(D_\mathcal{L}, \boldsymbol{D}_\mathcal{R})$.

Theorem 2.4. [7, Theorem 2.11, Corollary 2.12]. *Let Assumption 2.3 hold. Further assume $p_2 + q_2 < 2$.*

- ***Supercritical regime.*** *Assume further the following* supercriticality *condition holds,*

$$\mathbb{E}\left[\underline{D_\mathcal{L}}\right] \mathbb{E}\left[\underline{D_\mathcal{R}}\right] > 1. \tag{2.4}$$

6

Let $\xi_L := 1 - G_{D_{\mathcal{L}}}(\eta_L)$, with η_L being the smallest solution of the fixed point equation $\eta_L = G_{\underline{D_{\mathcal{R}}}}(G_{\underline{D_{\mathcal{L}}}}(\eta_L))$. Further, let $\xi_R = 1 - G_{D_{\mathcal{R}}}(\eta_R)$, with $\eta_R := G_{\underline{D_{\mathcal{L}}}}(\eta_L)$. Then,

$$\frac{|\mathcal{C}_1|}{L_n + R_n} \xrightarrow{\mathbb{P}} \frac{\xi_L + \gamma\xi_R}{1 + \gamma} > 0 \ \ and \ \ \frac{|\mathcal{C}_2|}{L_n + R_n} \xrightarrow{\mathbb{P}} 0, \ \ where \ \ \gamma := \frac{\mathbb{E}[D_{\mathcal{L}}]}{\mathbb{E}[D_{\mathcal{R}}]}. \tag{2.5}$$

- **Subcritical regime.** If the supercriticality condition (2.4) does not hold, then $\frac{|\mathcal{C}_1|}{L_n + R_n} \xrightarrow{\mathbb{P}} 0$.

We remark that the limit $\frac{\xi_L + \gamma\xi_R}{1+\gamma}$ of the scaled size of the giant component appearing in (2.5) is strictly positive, and that is part of the assertion of the theorem. The similarity of the statement of the last theorem with the statement of Theorem 1.1 should be noted. We next discuss the behavior of the bi-degree sequence of the random multigraph B_t^*.

Negative binomial bi-degree sequence of B_t^*. Let us begin by recalling the negative binomial distribution, and some of its properties. The *negative binomial distribution* with shape parameter $\alpha > 0$ and success probability p, has probability mass function

$$f(k) = \frac{\prod_{0 \le j \le k}(\alpha + j)}{k!}(1-p)^\alpha p^k, \ \ \text{for} \ \ k \in \mathbb{Z}_{\ge 0} = \{0, 1, 2, \dots\}. \tag{2.6}$$

The following lemma is [8, Lemma 3.1] and we don't include a proof. We have added the third statement, it easily follows from (2.6). Recall the *shifted size-biased* distribution from (2.3). For $x \in \mathbb{R}$ and an integer $k \ge 0$, define $\langle x \rangle_k = \prod_{j=0}^{k-1}(x-j)$, with the convention that $\langle x \rangle_0 = 1$.

Lemma 2.5 (Properties of negative binomial). *Let $Y \sim \text{NB}(\alpha, p)$ with $\alpha \in (0, \infty)$ and $p \in (0, 1)$.*

a. $\mathbb{E}[Y] = \frac{\alpha p}{1-p}$, *and for $k \ge 0$, $\mathbb{E}[\langle Y \rangle_k] = (\mathbb{E}[Y])^k \prod_{j=1}^{k}(1 + j/\alpha)$.*

b. $G_Y(z) = \left(\frac{1-p}{1-pz}\right)^\alpha$.

c. $\underline{Y} \sim \text{NB}(\alpha + 1, p)$.

Let $U_{\mathcal{L}}$ and $U_{\mathcal{R}}$ respectively be uniformly random elements from $\mathcal{L}$ and $\mathcal{R}$, and $p_t(k) := \mathbb{P}(d_{U_{\mathcal{L}}}(t) = k)$, $q_t(k) := \mathbb{P}(d_{U_{\mathcal{R}}}(t) = k)$, where for any $u \in V$, $d_u(t)$ denotes its degree in B_t^*.

Proposition 2.6. *Assume $t = \Theta(L_n + R_n)$ and $\alpha, \beta > 0$. Consider random variables $D_{n,\mathcal{L}} \sim \text{NB}(\alpha, t/(t + \alpha L_n)), D_{n,\mathcal{R}} \sim \text{NB}(\beta, t/(t + \beta R_n))$. Then for any $k \ge 0$,*

$$p_t(k) = \mathbb{P}(D_{n,\mathcal{L}} = k) + o(1), \ \ q_t(k) = \mathbb{P}(D_{n,\mathcal{R}} = k) + o(1).$$

Furthermore, for any $k \ge 1$ and $\square \in \{\mathcal{L}, \mathcal{R}\}$, $\mathbb{E}\left[\left(d_{U_\square}(t)\right)^k\right] = \mathbb{E}\left[D_{n,\square}^k\right] + o(1)$.

All these results at hand, we can straightaway provide the proof of Theorem 1.1.

Proof of Theorem 1.1, subject to Theorem 2.4, and Propositions 2.1, 2.2 and 2.6. For any $\epsilon > 0$ constant, define random variables $D_{\mathcal{L}}(\epsilon) \sim \text{NB}(\alpha, p_L(\epsilon)), \ D_{\mathcal{R}}(\epsilon) \sim \text{NB}(\beta, p_R(\epsilon))$, where recall $p_L(\epsilon)$ and $p_R(\epsilon)$ from (1.3). Let us only prove the supercritical statement of Theorem 1.1, the proof of the subcritical statement is completely analogous. To this end, reall t_c from (1.4), and let $t = (1 + \epsilon)t_c(L_n + R_n)$ for some $\epsilon > 0$.

For this value of t, thanks to Proposition 2.6, we have $D_{n,\mathcal{L}} \xrightarrow{d} D_{\mathcal{L}}(\epsilon)$ and $D_{n,\mathcal{R}} \xrightarrow{d} D_{\mathcal{R}}(\epsilon)$ with convergence of the respective first moments. For any $\delta > 0$, let us define the event

$$\mathcal{A}_\delta := \left\{\left|\frac{|\mathcal{C}_1(t_\epsilon^+)|}{L_n + R_n} - \frac{\xi_L + \gamma\xi_R}{1 + \gamma}\right| > \delta\right\},$$

7

where $\gamma = \frac{\mathbb{E}[D_{\mathcal{L}}(\epsilon)]}{\mathbb{E}[D_{\mathcal{R}}(\epsilon)]}$, and where $\xi_L := 1 - G_{D_{\mathcal{L}}(\epsilon)}(\eta_L)$, with η_L being the smallest solution of the fixed point equation $\eta_L = G_{\underline{D_{\mathcal{R}}(\epsilon)}}(G_{\underline{D_{\mathcal{L}}(\epsilon)}}(\eta_L))$, and further, $\xi_R = 1 - G_{D_{\mathcal{R}}(\epsilon)}(\eta_R)$, with $\eta_R := G_{\underline{D_{\mathcal{L}}(\epsilon)}}(\eta_L)$.

For any $t > 0$, let us denote by $(D_{\mathcal{L}}(t), D_{\mathcal{R}}(t))$ the (random) bi-degree sequence of the bipartite multigraph B_t^*. Thus $D_{\mathcal{L}}(t) = (d_u(t))_{u \in \mathcal{L}}$ and $D_{\mathcal{R}}(t) = (d_v(t))_{v \in \mathcal{R}}$. We note that

$$
\begin{aligned}
&\mathbb{P}\left(\mathcal{A}_\delta \text{ holds for } B_t\right) \\
&\leq B \cdot \mathbb{P}\left(\mathcal{A}_\delta \text{ holds for } B_t^*\right) + o(1) && \text{(Theorem 2.1)} \\
&\leq B \cdot \mathbb{E}\left[\mathbb{P}\left(\mathcal{A}_\delta \text{ holds for } B_t^* \mid D_{\mathcal{L}}(t), D_{\mathcal{R}}(t)\right)\right] + o(1) \\
&\leq B \cdot \mathbb{P}\left(\mathcal{A}_\delta \text{ holds for } \mathrm{BCM}(D_{\mathcal{L}}(t), D_{\mathcal{R}}(t))\right) + o(1). && \text{(Proposition 2.2)} \quad (2.7)
\end{aligned}
$$

As we have already observed, the random bi-degree sequence $(D_{\mathcal{L}}(t), D_{\mathcal{R}}(t))$ satisfies Assumption 2.3 with limiting variables $D_{\mathcal{L}}(\epsilon)$ and $D_{\mathcal{L}}(\epsilon)$. Observe that thanks to (2.1), $\gamma = \mathbb{E}[D_{\mathcal{L}}(\epsilon)]/\mathbb{E}[D_{\mathcal{R}}(\epsilon)]$ is in fact the limit of $|\mathcal{R}|/|\mathcal{L}| = R_n/L_n$. Also, $p_2 + q_2 = \mathbb{P}(D_{\mathcal{L}}(\epsilon) = 2) + \mathbb{P}(D_{\mathcal{R}}(\epsilon) = 2) < 2$ by (2.6). Next, we verify the supercriticality condition $\mathbb{E}\left[\underline{D_{\mathcal{L}}(\epsilon)}\right] \mathbb{E}\left[\underline{D_{\mathcal{R}}(\epsilon)}\right] > 1$. Note that the condition can be equivalently written as,

$$
\frac{\mathbb{E}\left[D_{\mathcal{L}}(\epsilon)(D_{\mathcal{L}}(\epsilon) - 1)\right]}{\mathbb{E}\left[D_{\mathcal{L}}(\epsilon)\right]} \frac{\mathbb{E}\left[D_{\mathcal{R}}(\epsilon)(D_{\mathcal{R}}(\epsilon) - 1)\right]}{\mathbb{E}\left[D_{\mathcal{R}}(\epsilon)\right]} > 1.
$$

From Lemma 2.5 recall that if $X \sim \mathrm{NB}(\alpha, p)$, then $\mathbb{E}[X] = \frac{\alpha p}{1-p}$ and $\mathbb{E}[X(X-1)] = \left(\frac{\alpha p}{1-p}\right)^2 (1 + \alpha^{-1})$. Also recall the definition of t_c from (1.4). Recalling the definitions of the variables $D_{\mathcal{L}}(\epsilon)$ and $D_{\mathcal{R}}(\epsilon)$ from (2.6), some easy algebra then yields that the last condition is equivalent to $t > t_c(L_n + R_n)$, which is true as $\epsilon > 0$.

Finally, using Lemma 2.5 again, and some easy rewriting, it can be verified that the description of $\xi_L, \xi_R, \eta_L, \eta_R$ as given in Theorem 1.1 matches with the description in Theorem 2.4, with $D_\square$ replaced by $D_\square(\epsilon)$ for $\square \in \{\mathcal{L}, \mathcal{R}\}$. Thus, applying Theorem 2.4, the RHS of (2.7) is $o(1)$. $\qquad\square$

In the following sections, we provide the proofs of Propositions 2.1, 2.2 and 2.6.

2.1 Proof of Proposition 2.1

For any bipartite (multi or simple) graph H on the vertex bipartition $V = \mathcal{L} \cup \mathcal{R}$, where any vertex u has degree $d_u(H)$, denote

$$
Q(H) = \sum_{u \in \mathcal{L}} d_u(H)^3 + \sum_{v \in \mathcal{R}} d_v(H)^3. \tag{2.8}
$$

Lemma 2.7 (Measure change lemma)**.** *Consider a deterministic sequence of bipartite multigraphs $(H_i)_{0 \leq i \leq t}$ on the vertex set $V = \mathcal{L} \cup \mathcal{R}$, with the property that H_{i+1} is obtained from H_i by including a single edge. Then, denoting $\eta = \alpha^2 + \beta^2$, we have*

$$
\mathbb{P}\left((B_j)_{0 \leq j \leq t} = (H_j)_{0 \leq j \leq t}\right) \leq \prod_{i=0}^{t-1} \left(1 - \frac{4(Q(H_i) + \eta i)}{(i + \alpha L_n)(i + \beta R_n)}\right)^{-1} \cdot \mathbb{P}\left((B_j^*)_{0 \leq j \leq t} = (H_j)_{0 \leq j \leq t}\right).
$$
$$
\tag{2.9}
$$

Proof. Let us denote by $\{u_i, v_i\}$ the unique edge by which H_i differs from H_{i+1}, where $u_i \in \mathcal{L}$ and $v_i \in \mathcal{R}$. Observe,

$$
\begin{aligned}
&\mathbb{P}\left((B_j)_{0 \leq j \leq t} = (H_j)_{0 \leq j \leq t}\right) \\
&= \prod_{i=0}^{t-1} \frac{(i + \alpha L_n)(i + \beta R_n)}{\sum_{\{u,v\} \notin E(H_i)} (d_u(H_i) + \alpha)(d_v(H_i) + \beta)} \mathbb{P}\left((B_j^*)_{0 \leq j \leq t} = (H_j)_{0 \leq j \leq t}\right)
\end{aligned}
$$

8

$$= \prod_{i=0}^{t-1} \left(1 - \frac{\sum_{\{u,v\} \in E(H_i)} (d_u(H_i) + \alpha)(d_v(H_i) + \beta)}{(i + \alpha L_n)(i + \beta R_n)} \right)^{-1} \mathbb{P}\left((B_j^*)_{0 \le j \le t} = (H_j)_{0 \le j \le t} \right).$$

Note that for any $x, y > 0$, $(x + \alpha)(y + \beta) \le 4x^2 + 4y^2 + 4\eta$. Thus,

$$\mathbb{P}\left((B_j)_{0 \le j \le t} = (H_j)_{0 \le j \le t} \right) \le \prod_{i=0}^{t-1} \left(1 - \frac{4(Q(H_i) + \eta i)}{(i + \alpha L_n)(i + \beta R_n)} \right)^{-1} \cdot \mathbb{P}\left((B_j^*)_{0 \le j \le t} = (H_j)_{0 \le j \le t} \right),$$

finishing the proof. $\qquad \square$

We can now prove Proposition 2.1.

Proof of Proposition 2.1. Since $i \le t \le C(L_n + R_n)$ and by (1.1) $L_n = \Theta(R_n)$, we immediately have from the last display that for any constant $A > 0$ and any set $\mathcal{A}$ of graph sequences $(H_i)_{0 \le i \le t}$,

$$\mathbb{P}\left((B_j)_{0 \le j \le t} \in \mathcal{A} \right) \le \Theta(1) \cdot \mathbb{P}\left((B_j^*)_{0 \le j \le t} \in \mathcal{A} \right) + \mathbb{P}\left(\text{for some } i \le t-1, \ Q(B_i) > A(L_n + R_n) \right), \tag{2.10}$$

(where the $\Theta(1)$ term depends on the constant A) so that it is enough to check that the second term on the RHS above is $o(1)$ for some A sufficiently large. A union bound, and using the facts that $Q(B_0) = 0$ and $Q(B_t) \ge Q(B_{t'})$ if $t \ge t'$, gives this quantity is at most

$$\sum_{i=0}^{t} \mathbb{P}\left(Q(B_{i-1}) < A(L_n + R_n) \ \text{while} \ Q(B_i) > A(L_n + R_n) \right)$$

which using that $t \le C(L_n + R_n)$ and (2.10) is at most (again using the monotonicity of Q)

$$\Theta(1) \cdot C(L_n + R_n) \cdot \mathbb{P}\left(Q(B_t^*) > A(L_n + R_n) \right).$$

Note that A can be chosen a large enough constant such that $\mathbb{P}\left(Q(B_t^*) > A(L_n + R_n) \right) = o((L_n + R_n)^{-1})$. This is possible using the second assertion of Proposition 2.6 for some large k ($k \ge 5$ suffices) combined with a Markov's inequality, and noting that since $t \le C(L_n + R_n)$, both $\mathbb{E}\left[D_{n,L}^k \right], \mathbb{E}\left[D_{n,R}^k \right] = O(1)$ (with the $O(1)$ term depending on C) for any $k \ge 1$. $\qquad \square$

2.2 Proof of Proposition 2.2

2.2.1 A convenient construction of the BCM

Recall the definition of the BCM from Section 2. Consider a bi-degree sequence $\boldsymbol{D} = (\boldsymbol{D}_{\mathcal{L}}, \boldsymbol{D}_{\mathcal{R}})$ on a bipartitioned vertex set $V = \mathcal{L} \cup \mathcal{R}$, where $\boldsymbol{D}_{\mathcal{L}} = (d_u)_{u \in \mathcal{L}}$ and $\boldsymbol{D}_{\mathcal{R}} = (d_v)_{v \in \mathcal{R}}$ satisfies $m = \sum_{u \in \mathcal{L}} d_u^{(L)} = \sum_{v \in \mathcal{L}} d_v^{(R)}$. The following is a convenient way to construct the BCM with bi-degree sequence $\boldsymbol{D}$.

- Let $\mathcal{S}(\boldsymbol{D}_{\mathcal{L}})$ be the set of all sequences of the form $\mathbf{u} := (u_1, u_2, \ldots, u_m)$ where each $u_i \in \mathcal{L}$, and any $u \in \mathcal{L}$ appears exactly d_u many times in $\mathbf{u}$. Similarly, let $\mathcal{S}(\boldsymbol{D}_{\mathcal{R}})$ be the set of all sequences $\mathbf{v} = (v_1, \ldots, v_m)$ with each $v \in \mathcal{R}$ appearing exactly d_v many times in $\mathbf{v}$.

- Sample uniformly at random $\mathbf{U} = (U_1, \ldots, U_m)$ from $\mathcal{S}(\boldsymbol{D}_L)$, and independently, uniformly at random $\mathbf{V} = (V_1, \ldots, V_m)$ from $\mathcal{S}(\boldsymbol{D}_R)$. Include the edges $\{U_i, V_i\}$ for all $1 \le i \le m$ to form a bipartite multigraph on $V = \mathcal{L} \cup \mathcal{R}$, and note that it has the same distribution as $\mathrm{BCM}(\boldsymbol{D}_L, \boldsymbol{D}_R)$.

2.2.2 A convenient construction of the multigraph process

Observe the following related construction of the multigraph $(B_t^*)_{t \geq 0}$. The product structure on the right hand side of (1.7) implies the following independence that we exploit. Given B_{t-1}^*, sample a vertex $u \in \mathcal{L}$ with probability $\frac{(d_u(t-1)+\alpha)}{\sum_{x \in \mathcal{L}}(d_x(t-1)+\alpha)}$, and independently sample $v \in \mathcal{R}$ with probability $\frac{(d_v(t-1)+\beta)}{\sum_{y \in \mathcal{R}}(d_y(t-1)+\beta)}$. Include the edge $\{u, v\}$ in B_{t-1}^* to obtain B_t^*.

Keeping this observation in mind, we introduce the following random sequence of vertices from $\mathcal{L}$ and $\mathcal{R}$. Let X_1 and Y_1 be independently uniformly sampled from respectively $\mathcal{L}$ and $\mathcal{R}$, and for $i \geq 1$, given $X_1, \ldots, X_i$ and $Y_1, \ldots, Y_i$, sample $X_{i+1} \in \mathcal{L}$ and $Y_{i+1} \in \mathcal{R}$ conditionally independently with probability masses

$$\mathbb{P}\left(X_{i+1} = u \mid (X_j)_{j=1}^i\right) = \frac{\sum\limits_{1 \leq j \leq i}(\mathbb{1}_{\{X_j=u\}} + \alpha)}{i + \alpha L_n} \quad \text{and} \quad \mathbb{P}\left(Y_{i+1} = v \mid (Y_j)_{j=1}^i\right) = \frac{\sum\limits_{1 \leq j \leq i}(\mathbb{1}_{\{Y_j=v\}} + \beta)}{i + \beta R_n}.$$

Note that $\{\{X_i, Y_i\} : 1 \leq i \leq t\}$ has the same distribution as the edge set of B_t^*.

We can now provide the proof of Proposition 2.2. Consider a bi-degree sequence $\boldsymbol{D} = (\boldsymbol{D}_{\mathcal{L}}, \boldsymbol{D}_{\mathcal{R}})$, and recall the sets $\mathcal{S}(\boldsymbol{D}_{\mathcal{L}})$ and $\mathcal{S}(\boldsymbol{D}_{\mathcal{R}})$ from Section 2.2.1. Let us use the notation $\mathrm{BP}(\mathbf{u}, \mathbf{v})$ to denote the bipartite multigraph on $V = \mathcal{L} \cup \mathcal{R}$ with edge set $\{u_i, v_i\}$, for fixed $\mathbf{u} \in \mathcal{S}(\boldsymbol{D}_{\mathcal{L}}), \mathbf{v} \in \mathcal{S}(\boldsymbol{D}_{\mathcal{R}})$.

Proof of Proposition 2.2. Thanks to the constructions described in Sections 2.2.1 and 2.2.2, for a given bipartite multigraph G with bi-degree sequence $(\boldsymbol{D}_{\mathcal{L}}, \boldsymbol{D}_{\mathcal{R}})$ we have

$$\mathbb{P}\left(B_t^* = G \mid (D_{\mathcal{L}}(t), D_{\mathcal{R}}(t)) = (\boldsymbol{D}_{\mathcal{L}}, \boldsymbol{D}_{\mathcal{R}})\right)$$

$$= \frac{\mathbb{P}\left(B_t^* = G\right)}{\mathbb{P}\left((D_{\mathcal{L}}(t), D_{\mathcal{R}}(t)) = (\boldsymbol{D}_{\mathcal{L}}, \boldsymbol{D}_{\mathcal{R}})\right)}$$

$$= \frac{\sum\limits_{\substack{\mathbf{u} \in \mathcal{S}(\boldsymbol{D}_{\mathcal{L}}), \mathbf{v} \in \mathcal{S}(\boldsymbol{D}_{\mathcal{R}}): \\ \mathrm{BP}(\mathbf{u},\mathbf{v})=G}} \mathbb{P}\left(X_i = u_i \ \forall \ 1 \leq i \leq t\right)\mathbb{P}\left(Y_i = v_i \ \forall \ 1 \leq i \leq t\right)}{\left(\sum\limits_{\mathbf{u} \in \mathcal{S}(\boldsymbol{D}_{\mathcal{L}})} \mathbb{P}\left(X_i = u_i \ \forall \ 1 \leq i \leq t\right)\right)\left(\sum\limits_{\mathbf{v} \in \mathcal{S}(\boldsymbol{D}_{\mathcal{R}})} \mathbb{P}\left(Y_i = v_i \ \forall \ 1 \leq i \leq t\right)\right)}$$

$$= \frac{\sum\limits_{\substack{\mathbf{u} \in \mathcal{S}(\boldsymbol{D}_{\mathcal{L}}), \mathbf{v} \in \mathcal{S}(\boldsymbol{D}_{\mathcal{R}}): \\ \mathrm{BP}(\mathbf{u},\mathbf{v})=G}} \frac{\prod_{u \in \mathcal{L}}\prod_{p=0}^{d_u}(p+\alpha)}{\prod_{j=0}^{t-1}(j+\alpha L_n)} \frac{\prod_{v \in \mathcal{R}}\prod_{p=0}^{d_v}(p+\beta)}{\prod_{j=0}^{t-1}(j+\beta R_n)}}{\left(\sum\limits_{\mathbf{u} \in \mathcal{S}(\boldsymbol{D}_{\mathcal{L}})} \frac{\prod_{u \in \mathcal{L}}\prod_{p=0}^{d_u}(p+\alpha)}{\prod_{j=0}^{t-1}(j+\alpha L_n)}\right)\left(\sum\limits_{\mathbf{v} \in \mathcal{S}(\boldsymbol{D}_{\mathcal{R}})} \frac{\prod_{v \in \mathcal{R}}\prod_{p=0}^{d_v}(p+\beta)}{\prod_{j=0}^{t-1}(j+\beta R_n)}\right)},$$

the last line following by the definition of the random variables $(X_i)_{1 \leq i \leq t}$ and $(Y_i)_{1 \leq i \leq t}$. Note that the denominator in the last expression is simply a function of the bi-degree sequence of G, and so is the summand in the numerator. Thus, the last probability is simply the probability of the event $\{\mathrm{BP}(\mathbf{U}, \mathbf{V}) = G\}$, where $\mathbf{U}, \mathbf{V}$ are respectively uniformly distributed independent random elements from $\mathcal{S}(\boldsymbol{D}_{\mathcal{L}})$ and $\mathcal{S}(\boldsymbol{D}_{\mathcal{R}})$. Recalling the construction of the $\mathrm{BCM}(\boldsymbol{D}_{\mathcal{L}}, \boldsymbol{D}_{\mathcal{R}})$ from Section 2.2.1, we thus obtain that conditional on the event $\{(D_{\mathcal{L}}(t), D_{\mathcal{R}}(t)) = (\boldsymbol{D}_{\mathcal{L}}, \boldsymbol{D}_{\mathcal{R}})\}$, B_t^* is distributed as a $\mathrm{BCM}(\boldsymbol{D}_{\mathcal{L}}, \boldsymbol{D}_{\mathcal{R}})$. $\square$

2.3 Proof of Proposition 2.6

2.3.1 Continuous-time embedding

The following continuous-time embedding is due to Janson and Warnke [8]. Recall the sequence of random variables $(X_i)_{i=1}^t$ and $(Y_i)_{i=1}^t$ from section 2.2.2, and note that for any $u \in \mathcal{L}$ and $v \in \mathcal{R}$, $d_u(t) \overset{d}{=} \sum_{i=1}^t \mathbb{1}_{\{X_i=u\}}$ and $d_v(t) \overset{d}{=} \sum_{i=1}^t \mathbb{1}_{\{Y_i=v\}}$. To sample the random variables $(X_i)_{i=1}^t$ and $(Y_i)_{i=1}^t$, it is thus natural to introduce independent pure birth processes $(D_u(r))_{u \in \square}$

for $\square \in \{\mathcal{L}, \mathcal{R}\}$, where recalling the map ρ from (1.2), for each $u \in \square$, the process $(D_u(r))_{r \geq 0}$ gives the next birth at rate $k + \rho(\square)$ given that it has already given k births, $k \geq 0$. Then if we introduce stopping times $\tau_{\square,1}, \tau_{\square,2}, \ldots$ where the i-th birth in the union of all the processes $(D_u(r))_{u \in \square}$ takes place at time $\tau_{\square,i}$, we note that X_i has the same distribution as the vertex in $\mathcal{L}$ that has given birth at time $\tau_{\mathcal{L},i}$, and similarly $Y_i \in \mathcal{R}$ has the same distribution as the vertex that has given birth at time $\tau_{\mathcal{R},i}$. For fixed $r \geq 0$ and any $\square \in \{\mathcal{L}, \mathcal{R}\}$ it is standard that the distribution of the number of points in $D_u(r)$ for any $u \in \square$ is distributed as $\mathrm{NB}(\rho(\square), 1 - e^{-r})$, see e.g., [8, Section 3.2]. The following lemma is essentially [8, Theorem 2.5]. We have stated the result slightly more generally for our purposes, letting the number of edges to be superlinear and providing an explicit error bound, but the same proof works with minor modifications.

Lemma 2.8 (Stopping time lemma). *Let $t = t_n$ be a sequence of n satisfying $\liminf\limits_{n \to \infty} \frac{t}{L_n + R_n} \in (0, \infty]$. For any $\square \in \{\mathcal{L}, \mathcal{R}\}$, and for any sequence $s = s_n$ satisfying $s_n \leq t$, we have*

$$\mathbb{P}\left(\tau_n^{\square,-}(t,s) \leq \tau_{\square,t} \leq \tau_n^{\square,+}(t,s)\right) \geq 1 - \left(\frac{4}{s^2}\left(\frac{t^2}{\rho(\square)|\square|}\right) + t\right),$$

where

$$\tau_n^{\square,\pm}(t,s) := \log\left(1 + \frac{t \pm s}{\rho(\square)|\square|}\right). \tag{2.11}$$

In particular, $\liminf_{n \to \infty} \mathbb{P}\left(\tau_n^{\square,-}(t,s) \leq \tau_{\square,t} \leq \tau_n^{\square,+}(t,s)\right) = 1$, if $\frac{t}{\sqrt{L_n + R_n}} \ll s_n \leq t$.

The proof of Proposition 2.6 is now straightforward using Lemma 2.8:

Proof of Proposition 2.6. Assume $t = \Theta(L_n + R_n)$. For $\square \in \{\mathcal{L}, \mathcal{R}\}$, note that we have to conclude about $D_{U_\square}(\tau_{\square,t})$, where $U_\square$ is uniformly distributed on $\square$. But thanks to Lemma 2.8, in fact for any $u \in \square$, using the monotonicity of $D_u(\cdot)$, **whp**

$$D_u(\tau_n^{\square,-}(s,t)) \leq D_u(\tau_{\square,t}) \leq D_u(\tau_n^{\square,+}(s,t)),$$

where say $s = s_n = \ln(L_n + R_n)\sqrt{L_n + R_n}$. As for any $r > 0$ we have $D_u(r) \sim \mathrm{NB}(\rho(\square), 1 - e^{-r})$, it is a matter of simple verification that the upper and lower bounding variables in the above display satisfy the conclusions of the Proposition, and hence so does the middle variable. $\square$

3 Proof of connectivity properties

In this section we give the proofs of Theorems 1.2 and 1.5. We begin with a result on the number of isolated vertices.

Lemma 3.1 (Number of isolated vertices). *Fix $\square \in \{\mathcal{L}, \mathcal{R}\}$. Let $t = t_n$ satisfy $\lim\limits_{n \to \infty} \frac{t_n}{(L_n + R_n)^{1 + \rho(\square)^{-1}}} = x \in [0, \infty]$. Denote by $\mathcal{I}(\square, t)$ the number of isolated vertices from $\square$ in the graph B_t^*.*

- *If $x = 0$, $\liminf\limits_{n \to \infty} \mathbb{P}(\mathcal{I}(\square, t) \geq 1) \to 1$.*

- *If $x \in (0, \infty)$, as $n \to \infty$, $\mathcal{I}(\square, t) \xrightarrow{d} \mathrm{Poi}\left(\left(\frac{\rho(\square)}{x}\right)^{\rho(\square)}(\zeta(\square))^{-1-\rho(\square)}\right)$.*

- *If $x = \infty$, $\limsup\limits_{n \to \infty} \mathbb{P}(\mathcal{I}(\square, t) \geq 1) = 0$.*

Proof. We work with the independent continuous-time processes $(D_u(r))_{r \geq 0}$ for $u \in \square$. Note that denoting by $\mathcal{B}_u(s)$ the event that no birth has taken place in the process $(D_u(r))_{r \geq 0}$ up to time $s \geq 0$, we need to conclude about the random variable $\sum_{u \in \square} \mathbb{1}_{\mathcal{B}_u(\tau_{\square,t})}$. Using Lemma 2.8 and the monotonicity of the processes $D_u(\cdot)$, **whp**

$$\sum_{u \in \square} \mathbb{1}_{\mathcal{B}_u(\tau_n^{\square,+}(t,s))} \leq \sum_{u \in \chi(\eta)} \mathbb{1}_{\mathcal{B}_u(\tau_{\square,t})} \leq \sum_{u \in \square} \mathbb{1}_{\mathcal{B}_u(\tau_n^{\square,-}(t,s))}, \tag{3.1}$$

11

where let $s = (L_n + R_n)^{3/4 + \rho(\square)^{-1}}$. As the first birth in any of the processes $(D_u(r))_{r \geq 0}$ takes place after an $\text{Exp}(\rho(\square))$ amount of time,

$$\sum_{u \in \square} \mathbb{1}_{\mathcal{B}_u(\tau_n^{\square, \pm}(t,s))} \sim \text{Bin}\left(|\square|, p_\pm\right), \quad \text{with } p_\pm = \exp\left(-\rho(\square) \log\left(1 + \frac{t \pm s}{|\square| \rho(\square)}\right)\right). \quad (3.2)$$

Using (3.1), it is thus enough to prove $\mathbb{P}\left(\sum_{u \in \square} \mathbb{1}_{\mathcal{B}_u(\tau_n^{\square, -}(t,s))} \geq 1\right) \to 1$ to conclude the first assertion of the lemma. Since $\rho(\square) > 0$ and $t \ll |\square|^{1 + \rho(\square)^{-1}}$, let us write $1 + \frac{t-s}{|\square| \rho(\square)} = (s_n |\square|)^{\rho(\square)^{-1}}$ where $s_n = o(1)$. Using (3.2), note that $\mathbb{E}\left[\sum_{u \in \square} \mathbb{1}_{\mathcal{B}_u(\tau_n^{\square, -}(t,s))}\right] = s_n^{-1} \gg 1$, and $\text{Var}\left(\sum_{u \in \square} \mathbb{1}_{\mathcal{B}_u(\tau_n^{\square, -}(t,s))}\right) = \frac{1}{s_n}\left(1 - \frac{1}{s_n |\square|}\right)$, so that a standard second-moment method finishes the proof.

The proof of the second and the third assertions are similar. For the second one, we simply note that the upper and lower bounding variables in (3.1) converge in distribution to the required Poisson variable, thus so does the middle one. For the third one, a mean computation shows that $\mathbb{E}\left[\sum_{u \in \square} \mathbb{1}_{\mathcal{B}_u(\tau_n^{\square, +}(t,s))}\right] = o(1)$, and the proof is finished using (3.1) and Markov's inequality. $\square$

Remark 3.2. In [10], Pittel proves a similar Poisson approximation statement for the number of isolated vertices in the unipartite case using combinatorial techniques. We instead note the usefulness of the continuous-time embedding of Section 2.3.1. We believe using this continuous description, one should also be able to prove similar Poisson approximation statements for counts of *small* subgraphs, in particular, identify their appearance thresholds. This is something we do not address in this paper and leave for future work.

We continue by deriving useful expressions for the probabilities of certain events of interest. The analysis of combinatorial flavor that follows is inspired by Pittel [10]. For a bipartite multigraph M on vertex bipartition $V = \mathcal{L} \cup \mathcal{R}$, and for any $A \subset \mathcal{L}$, denote the *neighborhood* of A in M by

$$N_M(A) := \{v \in \mathcal{R} : \exists\, u \in A \text{ satisfying } \{u,v\} \in E(M)\},$$

where $E(M)$ is the multiset of edges in M. For convenience, we write $\mathcal{N}_t(A) = \mathcal{N}(A, B_t^*)$ for any $t \geq 0$. Let $\mathcal{M}_t$ be the set of all bipartite multigraphs on $V = \mathcal{L} \cup \mathcal{R}$ with t edges, and further for any $A \subset \mathcal{L}$, $B \subset \mathcal{R}$, $t_1, y \geq 0$ with $t_1 + y \leq t$, let

$$\mathcal{B}_t(A, B, t_1, y) := \left\{M \in \mathcal{M}_t : \mathcal{N}(A, M) \subset B, \sum_{u \in A} d_u(M) = t_1, \sum_{v \in B} d_v(M) = t_1 + y\right\}, \quad (3.3)$$

where $d_u(M)$ denotes the degree of $u \in \mathcal{L}$ in the multigraph M. For any $\alpha > 0$ and $m \in \mathbb{N}$, we denote the *rising factorial* as $(\alpha)_m := \alpha(\alpha + 1) \ldots (\alpha + m - 1)$.

Proposition 3.3 (Edge-partition formula). *Letting by convention $\binom{t-t_1}{0} = 1$,*

$$\mathbb{P}\left(\mathcal{G}_t^* \in \mathcal{B}_t(A, B, t_1, y)\right)$$
$$= \binom{t}{t_1}\binom{t - t_1}{y} \cdot \frac{(|A|\alpha)_{t_1}((L_n - |A|)\alpha)_{t - t_1}}{(\alpha L_n)_t} \cdot \frac{(|B|\beta)_{t_1 + y}((R_n - |B|)\beta)_{t - t_1 - y}}{(\beta R_n)_t}. \quad (3.4)$$

We need a couple of useful tools to prove this proposition, which we state as the following two lemmas.

Lemma 3.4. *For any bipartite multigraph M on $V = \mathcal{L} \cup \mathcal{R}$, where any $u \in \mathcal{L}, v \in \mathcal{R}$ has exactly $n_M(u,v)$ many edges between them, the degree of any vertex u is $d_u(M)$, and the total number of edges $\sum_{u \in \mathcal{L}, v \in \mathcal{R}} n_M(u,v) = t$,*

$$\mathbb{P}\left(B_t^* = M\right) = \frac{\prod_{u \in \mathcal{L}} (\alpha)_{d_u(M)}}{(\alpha L_n)_t} \cdot \frac{\prod_{v \in \mathcal{R}} (\beta)_{d_v(M)}}{(\beta R_n)_t} \cdot \frac{t!}{\prod_{u \in \mathcal{L}, v \in \mathcal{R}} n_M(u,v)!}.$$

Proof. The combinatorial factor $\frac{t!}{\prod_{u \in \mathcal{L}, v \in \mathcal{R}} n_M(u,v)!}$ counts the number of *graph paths* $(G_0, G_1, \ldots, G_t)$ where G_0 is the empty graph, $G_t = M$, and G_{i+1} is obtained from G_i by the inclusion of exactly one edge for all $0 \le i \le t - 1$. Finally, for a given such graph path $(G_0, \ldots, G_t)$, recalling the convenient construction of B_t^* from Section 2.2.1, we have

$$\mathbb{P}\left(B_j^* = G_j \ \forall \ 0 \le j \le t\right) = \frac{\prod_{u \in \mathcal{L}}(\alpha)_{d_u(M)}}{(\alpha L_n)_t} \cdot \frac{\prod_{v \in \mathcal{R}}(\beta)_{d_v(M)}}{(\beta R_n)_t}.$$

$\square$

Lemma 3.5. *For any integer $n \ge 1$, consider reals $p_i \ge 0$ for $1 \le i \le n$ and let $\mathbf{Y}_k$ denote the set of all non-negative integer sequences $\mathbf{y} = (y_i)_{1 \le i \le n}$ satisfying $\sum_{1 \le i \le n} y_i = k$ for some integer $k \ge 0$. Then,*

$$\sum_{\mathbf{y} \in \mathbf{Y}_k} \prod_{i=1}^n \frac{(p_i)_{y_i}}{y_i!} = \frac{(\sum_{i=1}^n p_i)_k}{k!} \tag{3.5}$$

Proof. Recall the negative binomial theorem: $\sum_{p \ge 0} \frac{(\alpha)_p}{p!} z^p = (1 - z)^{-\alpha}$ holds for any $z \in \mathbb{R}$ with $|z| < 1$. As a consequence, the quantity on the LHS of (3.5) is simply the coefficient of z^k in the expansion of $\prod_{i=1}^n (1 - z)^{-p_i} = (1 - z)^{-(\sum_{i=1}^n p_i)}$. $\square$

We can now provide the proof of Proposition 3.3.

Proof of Proposition 3.3. To conclude Proposition 3.3, we need to sum the probability formula provided by Lemma 3.4, over all multigraphs in $\mathcal{B}_t(A, B, t_1, y)$. Note that for any fixed $M \in \mathcal{B}_t(A, B, t_1, y)$,

$$
\begin{aligned}
&\mathbb{P}\left(B_t^* = M\right) \\
&= \binom{t}{t_1}\binom{t - t_1}{y} \cdot \frac{\prod_{u \in \mathcal{L}}(\alpha)_{d_u(M)}}{(\alpha L_n)_m} \cdot \frac{t_1!}{\prod_{u \in A, v \in B} n_M(u,v)!} \\
&\qquad \cdot \frac{\prod_{v \in \mathcal{R}}(\beta)_{d_v(M)}}{(\beta R_n)_m} \cdot \frac{y!}{\prod_{u \in A^c, v \in B} n_M(u,v)!} \cdot \frac{(t - t_1 - y)!}{\prod_{u \in A^c, v \in B^c} n_M(u,v)!}, \quad (3.6)
\end{aligned}
$$

where $A^c = \mathcal{L} \setminus A$ and $B^c = \mathcal{R} \setminus B$. Recall the sets $\mathcal{S}(\boldsymbol{D}_\mathcal{L})$ and $\mathcal{S}(\boldsymbol{D}_\mathcal{R})$ from Section 2.2.1, where $\boldsymbol{D}_\mathcal{L} = (d_u(M))_{u \in \mathcal{L}}, \boldsymbol{D}_\mathcal{R} = (d_v(M))_{v \in \mathcal{R}}$, and where for any $u \in \mathcal{L}, v \in \mathcal{R}$ we have the degrees $d_u(M) := \sum_{v \in \mathcal{R}} n_M(u,v), d_v(M) := \sum_{u \in \mathcal{L}} n_M(u,v)$. Note that for any fixed $M \in \mathcal{B}_t(A, B, t_1, y)$,

$$
\begin{aligned}
&|\{(\mathbf{u}, \mathbf{v}) \in \mathcal{S}(\boldsymbol{D}_\mathcal{L}) \times \mathcal{S}(\boldsymbol{D}_\mathcal{R}) : \mathrm{BP}(\mathbf{u}, \mathbf{v}) = M\}| \\
&= \frac{t!}{\prod_{u \in \mathcal{L}, v \in \mathcal{R}} n_M(u,v)!} \\
&= \binom{t}{t_1}\binom{t - t_1}{y} \cdot \frac{t_1!}{\prod_{u \in A, v \in B} n_M(u,v)!} \cdot \frac{y!}{\prod_{u \in A^c, v \in B} n_M(u,v)!} \cdot \frac{(t - t_1 - y)!}{\prod_{u \in A^c, v \in B^c} n_M(u,v)!}, (3.7)
\end{aligned}
$$

where the first equality above is true simply because given a fixed $(\mathbf{u}, \mathbf{v}) \in \{(\mathbf{u}, \mathbf{v}) \in \mathcal{S}(\boldsymbol{D}_\mathcal{L}) \times \mathcal{S}(\boldsymbol{D}_\mathcal{R}) : \mathrm{BP}(\mathbf{u}, \mathbf{v}) = M\}$, all such other $(\mathbf{u}, \mathbf{v})$ can be obtained by permuting the edges $\{(u_i, v_i)\}_{1 \le i \le t}$, where the multi-edges between any pair of vertices are considered identical.

For any $\square \in \{\mathcal{L}, \mathcal{R}\}$, $H \subset \square$, $m \ge 0$, and a non-negative integer sequence $\mathbf{d} = (d_h)_{h \in H}$ satisfying $\sum_{h \in H} d_h = m$, we write $\mathcal{S}(H, \mathbf{d}, m)$ to be the set of all sequences $(w_1, \ldots, w_m)$ with each $w_i \in H$, and with the property that any $h \in H$ has multiplicity d_h in the multiset $\{w_1, \ldots, w_m\}$. This is essentially the analog of $\mathcal{S}(\mathcal{D}_\square)$, but restricted to the vertex subset H. Observe that

$$|\mathcal{S}(H, \mathbf{d}, m)| = \frac{m!}{\prod_{h \in H} d_h!}. \tag{3.8}$$

Consider the size of the set $\{(\mathbf{u}, \mathbf{v}) \in \mathcal{S}(\boldsymbol{D}_\mathcal{L}) \times \mathcal{S}(\boldsymbol{D}_\mathcal{R}) : \mathrm{BP}(\mathbf{u}, \mathbf{v}) = M\}$ that appears in the LHS of (3.7). Sum this size over all multigraphs M with the following properties:

13

Properties 3.6.

1. Bi-degree sequence $(\boldsymbol{D}_{\mathcal{L}}, \boldsymbol{D}_{\mathcal{R}})$.

2. $\mathcal{N}(A, M) \subset B$.

3. $\sum_{u \in A} d_u(M) = t_1$.

4. $\sum_{v \in B} d_v(M) = t_1 + y$.

5. Degree sequence from A is $\boldsymbol{D}(A) = (d(A)_v)_{v \in \mathcal{R}}$. This simply means for each $v \in \mathcal{R}$, it has exactly $d(A)_v$ many neighbors from A incident to it. Note that by (2.) above, for any $v \in \mathcal{R} \setminus B$, we must have $d(A)_v = 0$, and by (3.) above, we must have $\sum_{v \in B} d(A)_v = t_1$.

The obtained sum thus gives the size of the set of all pairs $(\mathbf{u}, \mathbf{v}) \in \mathcal{S}(\boldsymbol{D}_{\mathcal{L}}) \times \mathcal{S}(\boldsymbol{D}_{\mathcal{R}})$ such that $BP(\mathbf{u}, \mathbf{v}) \in \mathcal{B}_t(A, B, t_1, y)$ with bi-degree sequence $(\boldsymbol{D}_{\mathcal{L}}, \boldsymbol{D}_{\mathcal{R}})$ and degree sequence of B from A being $\boldsymbol{D}(A)$.

On the other hand, for any $U \subset \square \in \{\mathcal{L}, \mathcal{R}\}$, denoting the restriction of the sequence $\boldsymbol{D}_\square$ to U by $\boldsymbol{D}_{\square, U}$ and using the notation $\boldsymbol{D}(A)^c$ to denote the sequence $\boldsymbol{D}_{\mathcal{R}, B} - \boldsymbol{D}(A) = (d_v(M) - d(A)_v)_{v \in \mathcal{R}}$, we note that this size also has to equal

$$\binom{t}{t_1}\binom{t - t_1}{y} \cdot |(\mathcal{S}(A, \boldsymbol{D}_{\mathcal{L},A}, t_1) \times \mathcal{S}(A^c, \boldsymbol{D}_{\mathcal{L},A^c}, t - t_1))|$$
$$\cdot |(\mathcal{S}(B, \boldsymbol{D}(A)^c, y) \times \mathcal{S}(B, \boldsymbol{D}(A), t_1) \times \mathcal{S}(B^c, \boldsymbol{D}_{\mathcal{R},B^c}, t - t_1 - y))| . \quad (3.9)$$

Here the first binomial coefficient $\binom{t}{t_1}$ above simply chooses the coordinate indices of $\mathbf{u}$ corresponding to vertices from A. Since we require $\mathcal{N}(A, M) \subset B$, the same indices in $\mathbf{v}$ correspond to coordinates that are vertices from B. However, B has exactly y incidences from A^c. Thus the second binomial coefficient $\binom{t - t_1}{y}$ above chooses the coordinate indices of $\mathbf{v}$ other than the already chosen ones, where a vertex from B can be present.

In other words, using (3.9), and recalling (3.8) and (3.7), the following identity is true:

$$\sum_M \frac{t!}{\prod_{u \in \mathcal{L} v \in \mathcal{R}} n_M(u, v)!}$$
$$= \sum_M \binom{t}{t_1}\binom{t - t_1}{y} \cdot \frac{t_1!}{\prod_{u \in A, v \in B} n_M(u, v)!} \cdot \frac{y!}{\prod_{u \in A^c, v \in B} n_M(u, v)!} \cdot \frac{(t - t_1 - y)!}{\prod_{u \in A^c, v \in B^c} n_M(u, v)!}$$
$$= \binom{t}{t_1}\binom{t - t_1}{y} \cdot \left(\frac{t_1!}{\prod_{u \in A}(d_u(M))!} \cdot \frac{(t - t_1)!}{\prod_{v \in A^c}(d_u(M))!} \right)$$
$$\cdot \left(\frac{y!}{\prod_{v \in B}(d_v(M) - d(A)_v)!} \cdot \frac{t_1!}{\prod_{v \in B}(d(A)_v)!} \cdot \frac{(t - t_1 - y)!}{\prod_{v \in B^c}(d_v(M))!} \right), \quad (3.10)$$

where the sum above is over all $M \in \mathcal{B}_t(A, B, t_1, y)$ satisfying the five properties listed in Properties 3.6. Using (3.6), (3.7) and (3.10), recalling the (random) bi-degree sequence of B_t^* is $(D_{\mathcal{L}}(t), D_{\mathcal{R}}(t))$, and denoting by $D(A, t)$ the (random) degree sequence in B_t^* from A to vertices of $\mathcal{R}$, we thus note that

$$\mathbb{P}\left(B_t^* \in \mathcal{B}_t(A, B, t_1, y), (D_{\mathcal{L}}(t), D_{\mathcal{R}}(t)) = (\boldsymbol{D}_{\mathcal{L}}, \boldsymbol{D}_{\mathcal{R}}), D(A, t) = \boldsymbol{D}(A)\right)$$
$$= \frac{\binom{t}{t_1}\binom{t - t_1}{y}}{(\alpha L_n)_t (\beta R_n)_t} \cdot \left(\frac{(\prod_{u \in \mathcal{L}}(\alpha)_{d_u(M)}) \cdot t_1! \cdot (t - t_1)!}{(\prod_{u \in A}(d_u(M))!)(\prod_{u \in A^c}(d_u(M))!)} \right)$$
$$\cdot \left(\frac{(\prod_{v \in \mathcal{R}}(\beta)_{d_v(M)}) \cdot y! \cdot t_1! \cdot (t - t_1 - y)!}{(\prod_{v \in B}(d_v(M) - d(A)_v)!)(\prod_{v \in B}(d(A)_v)!)(\prod_{v \in B^c}(d_v(M))!)} \right)$$
$$= \frac{t! \cdot t_1! \cdot (t - t_1)!}{(\alpha L_n)_t (\beta R_n)_t} \cdot \left(\prod_{u \in A} \frac{(\alpha)_{d_u(M)}}{d_u(M)!} \right) \cdot \left(\prod_{u \in A^c} \frac{(\alpha)_{d_u(M)}}{d_u(M)!} \right) \cdot \left(\prod_{v \in B^c} \frac{(\beta)_{d_v(M)}}{d_v(M)!} \right)$$

$$\cdot \left(\prod_{v \in B} \frac{(\beta)_{d(A)_v}}{d(A)_v!} \right) \cdot \left(\prod_{v \in B} \frac{(\beta + d(A)_v)_{(d_v(M) - d(A)_v)}}{(d_v(M) - d(A)_v)!} \right). \tag{3.11}$$

To conclude (3.4), we want to take a sum of the last expression over all bi-degree sequences $(\boldsymbol{D}_{\mathcal{L}}, \boldsymbol{D}_{\mathcal{R}})$ and sequences $\boldsymbol{D}(A)$, satisfying $\sum_{u \in A} d_u(M) = t_1 = \sum_{v \in B} d(A)_v$ and $\sum_{v \in B} (d_v(M) - d(A)_v) = y$, and show that the result agrees with (3.4).

Keeping the sequence $(d(A)_v)_{v \in B}$ fixed, writing $z_v = (d_v(M) - d(A)_v)$ for $v \in B$, we note that using (3.5), taking a sum of (3.11) first over all non-negative sequences $(z_v)_{v \in B}$ with sum $\sum_{v \in B} z_v = y$ leads to

$$\frac{t! \cdot t_1! \cdot (t - t_1)!}{(\alpha L_n)_t (\beta R_n)_t} \cdot \left(\prod_{u \in A} \frac{(\alpha)_{d_u(M)}}{d_u(M)!} \right) \cdot \left(\prod_{u \in A^c} \frac{(\alpha)_{d_u(M)}}{d_u(M)!} \right) \cdot \left(\prod_{v \in B^c} \frac{(\beta)_{d_v(M)}}{d_v(M)!} \right)$$
$$\cdot \left(\prod_{v \in B} \frac{(\beta)_{d(A)_v}}{d(A)_v!} \right) \cdot \frac{(\beta|B| + \sum_{v \in B} d(A)_v)_y}{y!}.$$

Thus, taking a sum of the last expression over all non-negative integer sequences $(d(A)_v)_{v \in B}$ with sum $\sum_{v \in B} d(A)_v = t_1$ and again using (3.5), we obtain

$$\frac{t! \cdot t_1! \cdot (t - t_1)!}{(\alpha L_n)_t (\beta R_n)_t} \cdot \left(\prod_{u \in A} \frac{(\alpha)_{d_u(M)}}{d_u(M)!} \right) \cdot \left(\prod_{u \in A^c} \frac{(\alpha)_{d_u(M)}}{d_u(M)!} \right) \cdot \left(\prod_{v \in B^c} \frac{(\beta)_{d_v(M)}}{d_v(M)!} \right) \cdot \frac{(\beta|B|)_{t_1}}{t_1!} \cdot \frac{(\beta|B| + t_1)_y}{y!}.$$

Finally, employing (3.5) again, and taking iteratively sums of the last expression over non-negative integer sequences $(d_u(M))_{u \in A}$, $(d_u(M))_{u \in A^c}$ and $(d_v(M))_{v \in B^c}$ having respectively sums equal to t_1, $t - t_1$ and $t - t_1 - y$, we obtain

$$\frac{t! \cdot t_1! \cdot (t - t_1)!}{(\alpha L_n)_t (\beta R_n)_t} \cdot \frac{(\alpha|A|)_{t_1}}{t_1!} \cdot \frac{(\alpha(L_n - |A|))_{t - t_1}}{(t - t_1)!} \cdot \frac{(\beta(R_n - |B|))_{t - t_1 - y}}{(t - t_1 - y)!} \cdot \frac{(\beta|B|)_{t_1}}{t_1!} \cdot \frac{(\beta|B| + t_1)_y}{y!}$$
$$= \binom{t}{t_1} \binom{t - t_1}{y} \cdot \frac{(|A|\alpha)_{t_1}((L_n - |A|)\alpha)_{t - t_1}}{(\alpha L_n)_t} \cdot \frac{(|B|\beta)_{t_1 + y}((R_n - |B|)\beta)_{t - t_1 - y}}{(\beta R_n)_t},$$

finishing the proof. $\qquad\square$

Remark 3.7. As will be evident in the next section, to prove the connectivity threshold as claimed by Theorem 1.2, we use this result for the special case $y = 0$. However, we still prove this general result, as it is of independent interest. We hope to use the full strength of this result in future works where we tackle counts of small subgraphs, or the k-connectivity threshold.

3.1 Connectivity threshold of the multigraph process

In this section we prove Theorem 1.2. Note that if $t \ll \tau_{c,n}$, then Lemma 3.1 guarantees that there are isolated vertices with high probability, so that B_t^* is disconnected. Hence we need to prove that when $t \gg \tau_{c,n}$, B_t^* is connected with high probability. Towards this, we need the following structural result. Pittel [10] makes a similar observation in the unipartite case.

Proposition 3.8 (Multigraph structure beyond connectedness threshold). *Assume* $\liminf\limits_{n \to \infty} \frac{t}{\tau_{c,n}} \in (0, \infty]$. *Then* **whp**, B_t^* *has the following structure. There is one big connected component* $\mathcal{C}_t$, *and if this component does not cover all the vertices, then the rest of the components are simply isolated vertices.*

To prove this proposition, we need a few technical results.

Lemma 3.9. *Let M be a non-negative integer and c, N are non-negative reals. Then $\prod_{i=0}^{M} \left(1 - \frac{c}{N+i} \right) \leq \left(\frac{N}{N+M} \right)^c$, and $\prod_{i=0}^{M} \left(1 + \frac{c}{N+i} \right) \leq \left(\frac{M+N}{N} \right)^c$.*

Proof. Using the bounds $1 - x \leq e^{-x}$ and $\sum_{i=0}^{M} \frac{1}{N+i} \leq \int_0^M \frac{dy}{N+y}$,

$$\prod_{i=0}^{M} \left(1 - \frac{c}{N+i}\right) \leq \exp\left(-c \int_0^M \frac{dy}{N+y}\right) = \left(\frac{N}{N+M}\right)^c.$$

The argument for the second assertion is identical, now using $1 + x \leq e^x$. $\qquad\square$

Next, we state a couple of results assuming which we first finish the proof of Proposition 3.8 and Theorem 1.2. For a sequence x_n, we say it goes to zero *superpolynomially fast*, if $(L_n + R_n)^c x_n = o(1)$ for any $c > 0$. We denote this by $x_n = o_{\mathrm{sp}}(1)$

Lemma 3.10. *Let $t_1 = t_1(n)$ be a sequence of n satisfying $t_1 \gg (L_n + R_n)$. For any k, j integers with $1 \leq k \leq L_n$ and $1 \leq j \leq R_n$, and $t = t_n$ satisfying $\liminf_{n \to \infty} \frac{t}{\tau_{c,n}} \in (0, \infty]$,*

$$\binom{t}{t_1} \cdot \frac{(k\alpha)_{t_1}((L_n - k)\alpha)_{t-t_1}}{(L_n\alpha)_t} \cdot \frac{(j\beta)_{t_1}((R_n - j)\beta)_{t-t_1}}{(R_n\beta)_t} = o_{\mathrm{sp}}(1).$$

Let us not provide the proof of this result, a proof can be worked out essentially following the argument of [10, Proof of Theorem 2, Step 1].

Lemma 3.11. *Let $m = m_n$ be any sequence satisfying $m \ll \tau_{c,n}$. Then*

$$\sum_{k=1}^{L_n} \sum_{j=1}^{R_n} \sum_{t_1=k\vee j}^{m} \binom{L_n}{k} \cdot \binom{R_n}{j} \cdot \binom{t}{t_1} \cdot \frac{(k\alpha)_{t_1}((L_n - k)\alpha)_{t-t_1}}{(L_n\alpha)_t} \cdot \frac{(j\beta)_{t_1}((R_n - j)\beta)_{t-t_1}}{(R_n\beta)_t} = o(1).$$

We can now prove Proposition 3.8.

Proof of Proposition 3.8. First consider the case $\alpha \neq \beta$. To conclude the proposition, it is enough to show that there is a component comprising of *all* the vertices from $\rho^{-1}(\alpha \wedge \beta)$. Because of the bipartite structure, any remaining component must be an isolated vertex. We note that using Lemma 3.1, since $t \geq (L_n + R_n)^{1+(\alpha\wedge\beta)^{-1}} \gg (L_n + R_n)^{1+(\alpha\vee\beta)^{-1}}$, $\mathcal{I}(\rho^{-1}(\alpha \vee \beta)) = 0$ **whp**. Hence, we need to rule out that there is a component of B_t^* with $t_1 \in [1, t/2]$ many edges, with at most $|\rho^{-1}(\alpha \vee \beta)| - 1$ many vertices from $\rho^{-1}(\alpha \vee \beta)$. Let us further consider the case $\alpha > \beta$, the argument for the other possibility is identical.

When $\alpha > \beta$, if there is a component of B_t^* with $k \in [1, L_n - 1]$ many vertices from $\mathcal{L}$ with $t_1 \in [1, t/2]$ many edges, recalling (3.3), we note that $B_t^* \in \mathcal{B}_t(A, B, t_1, 0)$ for some $A \subset \mathcal{L}$ with size k, $B \subset \mathcal{R}$ is a proper subset of $\mathcal{R}$ (meaning not either ϕ or $\mathcal{R}$) and $k \vee |B| \leq t_1 < t/2$. Using Proposition 3.3 with $y = 0$, the probability of the event of interest is therefore at most

$$\sum_{k=1}^{L_n-1} \sum_{j=1}^{R_n-1} \sum_{t_1=k\vee j}^{t/2} \binom{L_n}{k} \cdot \binom{R_n}{j} \cdot \binom{t}{t_1} \cdot \frac{(k\alpha)_{t_1}((L_n - k)\alpha)_{t-t_1}}{(\alpha L_n)_t} \cdot \frac{(|B|\beta)_{t_1}((R_n - |B|)\beta)_{t-t_1}}{(\beta R_n)_t}.$$

$$(3.12)$$

Consider any term in the above sum corresponding to $t_1 \geq m_n := (L_n + R_n)^{1+\epsilon} \gg L_n + R_n$ for some $0 < \epsilon < (\alpha \wedge \beta)^{-1}$. Thanks to Lemma 3.10, as there are at most polynomial in $(L_n + R_n)$ many terms in the sum, the total contribution coming from these terms in the summand is $o(1)$. The remaining sum,

$$\sum_{k=1}^{L_n-1} \sum_{j=1}^{R_n-1} \sum_{t_1=k\vee j}^{m_n} \binom{L_n}{k} \cdot \binom{R_n}{j} \cdot \binom{t}{t_1} \cdot \frac{(k\alpha)_{t_1}((L_n - k)\alpha)_{t-t_1}}{(\alpha L_n)_t} \cdot \frac{(|B|\beta)_{t_1}((R_n - |B|)\beta)_{t-t_1}}{(\beta R_n)_t} = o(1)$$

applying Lemma 3.11 with $m = (L_n + R_n)^{1+\epsilon} \ll \tau_{c,n}$. Finally, consider the case $\alpha = \beta$. In this case, by Lemma 3.1, it is possible to have isolated vertices from both $\mathcal{L}$ and $\mathcal{R}$. However, a similar argument using Lemmas 3.10 and 3.11, one can show that the term (3.12) for $\alpha = \beta$ is $o(1)$. $\qquad\square$

Remark 3.12. Our argument shows that when $\alpha \neq \beta$, if $(L_n + R_n)^{1+(\alpha \vee \beta)^{-1}} \ll t \ll (L_n + R_n)^{1+(\alpha \wedge \beta)^{-1}}$, the multigraph B_t^* has the following structure **whp**: there is big connected component $\mathcal{C}_t$ comprising of all vertices from $\rho^{-1}(\alpha \vee \beta)$, and the remaining components are all isolated vertices from $\rho^{-1}(\alpha \wedge \beta)$.

Having Proposition 3.8 in hand, the proof of Theorem 1.2 is straightforward.

Proof of Theorem 1.2 for $\alpha \neq \beta$. Using Proposition 3.8, the only way for $\mathcal{G}_t^*$ to be disconnected is to have isolated vertices. The result now follows using the mutual independence of the processes $(D_u(\cdot))_{u \in V}$ and using Lemma 3.1. $\qquad\square$

Remark 3.13. Thus, the nature of connectivity for the (α, β)-multigraph dynamics is similar to the E-R multigraph dynamics, in that both are determined by isolated vertices. This is also true in the unipartite case, as was shown by Pittel [10, Theorem 2].

Finally, let us prove Lemma 3.11. This is an extension of [10, Proof of Theorem 2, Step 2]. Let us remark that we were hoping to omit out this proof as well, but the argument, even though in spirit similar to the mentioned reference, requires more subtle analysis, which is why we include a proof for completeness.

Proof. Denote the sum in question by $\sum_{k=1}^{L_n} \sum_{j=1}^{R_n} \sum_{t_1=k \vee j}^{m} \mathcal{S}(k,j,t_1)$. We split

$$\sum_{1 \leq k \leq L_n} \sum_{\substack{1 \leq j \leq R_n, \\ j \geq k}} \sum_{j \leq t_1 \leq m} \mathcal{S}(k,j,t_1) + \sum_{1 \leq k \leq L_n} \sum_{\substack{1 \leq j \leq L_n, \\ j < k}} \sum_{k \leq t_1 \leq m} \mathcal{S}(k,j,t_1), \tag{3.13}$$

with the convention that the sum over an empty set equals zero. Let us prove that the first term in (3.13) is $o(1)$, the analysis of the second term being similar, we omit it. Note that

$$\frac{\mathcal{S}(k,j,t_1+1)}{\mathcal{S}(k,j,t_1)} = \frac{t-t_1}{t_1+1} \cdot \frac{(k\alpha+t_1)(j\beta+t_1)}{((L-k)\alpha+t-t_1-1)((R-j)\beta+t-t_1-1)} < \delta,$$

for some $0 < \delta < 1$, whenever $t_1 \leq m_n \ll t$, so that the first sum in (3.13) equals up to a constant factor to

$$\sum_{1 \leq k \leq L_n} \sum_{\substack{1 \leq j \leq R_n, \\ j \geq k}} \binom{L_n}{k} \cdot \binom{R_n}{j} \cdot \binom{t}{j} \cdot \frac{(k\alpha)_j ((L_n-k)\alpha)_{t-j}}{(L_n\alpha)_t} \frac{(k\beta)_j ((R_n-j)\beta)_{t-j}}{(L_n\beta)_t}$$

$$= \sum_{1 \leq k \leq L_n} \sum_{\substack{1 \leq j \leq R_n, \\ j \geq k}} \mathcal{S}_1(k,j), \quad \text{say.}$$

Next, consider the ratio

$$\frac{\mathcal{S}_1(k,j+1)}{\mathcal{S}_1(k,j)}$$

$$= \frac{(R_n-j)(t-j)}{(j+1)^2} \frac{(k\alpha+j)}{((L_n-k)\alpha+t-j-1)} \frac{(k\beta+j)}{((R_n-j)\beta+t-j-1)} \prod_{p=0}^{t-j-2} \left(\frac{(R_n-j-1)\beta+p}{(R_n-j)\beta+p} \right).$$

Bounding the product term from above by 1, we obtain $\frac{\mathcal{S}_1(k,j+1)}{\mathcal{S}_1(k,j)} = O(\frac{R_n}{t}) \ll 1$, so that the sum

$$\sum_{1 \leq k \leq L_n} \sum_{1 \leq j \leq R_n, j \geq k} \mathcal{S}_1(k,j)$$

$$= \left(\sum_{1 \leq k \leq L_n} \binom{L_n}{k} \binom{R_n}{k} \binom{t}{k} \frac{(k\alpha)_k ((L_n-k)\alpha)_{t-k}}{(L_n\alpha)_t} \frac{(k\beta)_k ((R_n-k)\beta)_{t-k}}{(R_n\beta)_t} \right) \cdot (1+O(R_n/t))^{-1}$$

17

$$= \left(\sum_{1 \leq k \leq L_n} \mathcal{S}_2(k) \right) \cdot (1 + O(R_n/t))^{-1}, \quad \text{say.}$$

Finally, consider the ratio

$$\frac{\mathcal{S}_2(k+1)}{\mathcal{S}_2(k)} = \frac{L-k}{k+1} \frac{R-k}{k+1} \frac{t-k}{k+1} \prod_{\square \in \{L,R\}} \psi(\square),$$

where for $\square \in \{\mathcal{L}, \mathcal{R}\}$ fixed,

$$\psi(\square) = \frac{(k+1)\rho(\square) + k}{(|\square| - k)\rho(\square) + t - k - 1} \cdot \prod_{i=0}^{k-1} \left(1 + \frac{\rho(\square)}{k \cdot \rho(\square) + j} \right) \cdot \prod_{j=0}^{t-k-2} \left(1 - \frac{\rho(\square)}{(|\square| - k)\rho(\square) + j} \right).$$

Using Lemma 3.9 for the product terms in the last display, we note that

$$\psi(\square) \leq \frac{(k+1)\rho(\square) + k}{(|\square| - k)\rho(\square) + t - k - 1} \cdot \left(\frac{k - 1 + k \cdot \rho(\square)}{k \cdot \rho(\square)} \right)^{\rho(\square)} \cdot \left(\frac{(|\square| - k)\rho(\square)}{(|\square| - k)\rho(\square) + t - k - 2} \right)^{\rho(\square)}.$$

Overall, for any $\square \in \{\mathcal{L}, \mathcal{R}\}$, $\psi(\square) = O\left(\frac{k|\square|^{\rho(\square)}}{t^{1+\rho(\square)}} \right)$. Hence, recalling $L_n = \Theta(R_n)$, we note that the product term

$$\prod_{\square \in \{\mathcal{L}, \mathcal{R}\}} \psi(\square) = O\left(\frac{k^2 L_n^{\alpha+\beta}}{t^{2+\alpha+\beta}} \right).$$

As a consequence,

$$\frac{\mathcal{S}_2(k+1)}{\mathcal{S}_2(k)} = O\left(\frac{L_n^{2+\alpha+\beta}}{t^{1+\alpha+\beta}} \right) = o(1),$$

since $t = \Omega\left(L_n^{1+(\alpha \wedge \beta)^{-1}} \right)$. Therefore, we obtain

$$\sum_{k=1}^{L_n} \sum_{j=1}^{R_n} \sum_{t_1=k \vee j}^{t} \mathcal{S}(k, j, t_1) \leq 2\mathcal{S}_2(1) = 2(\alpha \beta L_n R_n t) \prod_{\square \in \{\mathcal{L}, \mathcal{R}\}} \left(\frac{((|\square| - 1) \cdot \rho(\square))_{t-1}}{(|\square| \cdot \rho(\square))_t} \right).$$

For fixed $\square \in \{\mathcal{L}, \mathcal{R}\}$, we note that

$$\frac{((|\square| - 1) \cdot \rho(\square))_{t-1}}{(|\square| \cdot \rho(\square))_t} = \frac{1}{|\square| \cdot \rho(\square) + t - 1} \cdot \prod_{j=0}^{t-2} \left(1 - \frac{\rho(\square)}{|\square| \cdot \rho(\square) + j} \right) = O\left(\frac{|\square|^{\rho(\square)}}{t^{1+\rho(\square)}} \right),$$

where to obtain the the last equality above we again use Lemma 3.9. Hence,

$$2\mathcal{S}_2(1) = O\left(\frac{L_n^{2+\alpha+\beta}}{t^{1+\alpha+\beta}} \right) = o(1),$$

which finishes the proof. $\qquad \square$

3.2 Simple graph connectedness lower bound

In this section we prove Theorem 1.5. The idea is to use the measure change lemma, Lemma 2.7. Let $\mathcal{A}$ be the event of connectedness, then to prove the simple graph is disconnected after the inclusion of $(L_n + R_n)^{1+\delta}$ many edges, we first estimate the prefactor in the RHS of (2.10), and then try to choose $\delta > 0$ carefully so that the postfactor $\mathbb{P}(B_t^*$ is connected$)$ goes to zero fast enough to beat this prefactor. Here are the details:

18

Proof of Theorem 1.5. Let $(H_i)_{0 \leq i \leq t}$ be any multigraph sequence where H_{i+1} is obtained from H_i after the inclusion of exactly one edge. By (2.9) and the monotonicity of $Q(H_i)$,

$$\mathbb{P}\left((B_j)_{0 \leq j \leq t} = (H_j)_{0 \leq j \leq t}\right) \leq \left(1 - \frac{4(Q(H_{t-1}) + \gamma(t-1))}{(\alpha\beta L_n R_n)}\right)^{-(t-1)} \cdot \mathbb{P}\left((B_j^*)_{0 \leq j \leq t} = (H_j)_{0 \leq j \leq t}\right).$$

Similar to (2.10), for any event $\mathcal{A}$ and $M > 0$, we derive from the last display

$$\mathbb{P}\left((B_j)_{0 \leq j \leq t} \in \mathcal{A}\right) \leq \left(1 - \frac{4M + \gamma(t-1))}{(\alpha\beta L_n R_n)}\right)^{-(t-1)} \mathbb{P}\left((B_j^*)_{0 \leq j \leq t} \in \mathcal{A}\right) + \mathbb{P}\left(Q(B_{t-1}^*) > M\right). \tag{3.14}$$

Let us first show that the second term on the RHS above is $o(1)$ for a right choice of M that we also make along the way. Recalling the definition of Q from (2.8), we bound this term from above as

$$\mathbb{P}\left(\sum_{u \in \mathcal{L}}(d_u(t))^3 > M/2\right) + \mathbb{P}\left(\sum_{v \in \mathcal{R}}(d_v(t))^3 > M/2\right). \tag{3.15}$$

We employ the continuous time embedding of Section 2.3.1. Recall the continuous time processes $D_v(\cdot)$ corresponding to any vertex $v \in \mathcal{L} \cup \mathcal{R}$, for any $\square \in \{\mathcal{L}, \mathcal{R}\}$ the stopping times $\tau_{\square,i}$ for $i \geq 1$, and the fact that for any $u \in \square$, $d_u(t)$ has the same distribution as $D_u(\tau_{\square,t})$. Thus, using time monotonicity of the processes $D_u(\cdot)$, and using Lemma 2.8, for $t = t(\delta) := (L_n + R_n)^{1+\delta}$ for any $\delta > 0$, we note that up to an additive error of $o(1)$, (3.15) is at most

$$\mathbb{P}\left(\sum_{u \in \mathcal{L}}(D_u(\tau_n^{\mathcal{L},+}(t,s)))^3 > M/2\right) + \mathbb{P}\left(\sum_{u \in \mathcal{R}}(D_u(\tau_n^{\mathcal{R},+}(t,s)))^3 > M/2\right), \tag{3.16}$$

for any s satisfying $\frac{t}{\sqrt{L}} \ll s \leq t$, and where recall $\tau_n^{\square,\pm}(t,s)$ from (2.11). From now on let us fix $s = (L_n + R_n)^{1/2+\delta+\delta_1}$ for a suitable $\delta_1 > 0$ satisfying $\delta + \delta_1 < 1/2$ which we will suitably choose later. Now, recall that for any $r > 0$, the distribution of $D_u(r)$ is $\text{NB}(\rho(\square), 1 - e^{-r})$ for $u \in \square$ for any $\square \in \{\mathcal{L}, \mathcal{R}\}$. Thus, using Lemma 2.5, we note that $\mathbb{E}\left[D_u(\tau_n^{\square,+}(s,t))\right] = \Theta\left((L_n + R_n)^\delta\right)$. In particular, using Lemma 2.5 for $k = 3$, and a Markov's inequality, we note that for $M = (L_n + R_n)^{3\delta+1+\epsilon}$ for arbitrary $\epsilon > 0$, the term (3.16) is $o(1)$.

We now turn to the first term on the RHS of (3.14). Observe that for the choices $M = (L_n + R_n)^{3\delta+1+\epsilon}$ and $t = (L_n + R_n)^{1+\delta}$, using the bound $1 - x > \exp(-x^{1-\delta'})$ for any $\delta' > 0$ and when x is small,

$$\left(1 - \frac{4M + \gamma(t-1))}{(\alpha\beta L_n R_n)}\right)^{-(t-1)} = \exp\left(O\left((L_n + R_n)^{4\delta+\epsilon}\right)\right). \tag{3.17}$$

Now, consider any $\delta \in (0, \mathcal{Z}(\alpha, \beta))$. Using (3.14) and (3.17), we write

$$\mathbb{P}\left(B_{t(\delta)} \text{ is connected}\right) \leq \exp\left(O\left((L_n + R_n)^{4\delta+\epsilon}\right)\right) \cdot \mathbb{P}\left(B_{t(\delta)}^* \text{ has no isolated vertices}\right) + o(1).$$

We now estimate the term $\mathbb{P}\left(B_{t(\delta)}^* \text{ has no isolated vertices}\right)$ using the continuous processes $D_u(\cdot)$. Note that by independence,

$$\mathbb{P}\left(B_{t(\delta)}^* \text{ has no isolated vertices}\right) = \mathbb{P}\left(D_u(\tau_{t(\delta)}) > 0\right)^{L_n} \cdot \mathbb{P}\left(D_v(\tau_{t(\delta)}) > 0\right)^{R_n}. \tag{3.18}$$

For any $\square \in \{\mathcal{L}, \mathcal{R}\}$ and $u \in \square$, using the monotonicity of $D_u(\cdot)$, we bound

$$\mathbb{P}\left(D_u(\tau_t) > 0\right) \leq \mathbb{P}\left(D_u(\tau_n^{\square,+}(t,s)) > 0\right) + \mathbb{P}\left(\tau_t > \tau_n^{\square,+}(t,s)\right), \tag{3.19}$$

19

and note that since $D_u(\tau_n^{\square,+}(t,s)) \sim \mathrm{NB}\left(\rho(\square), \frac{|\square|^{1+\delta}+s}{\rho(\square)|\square|+|\square|^{1+\delta}+s}\right)$, with some easy analysis, the first term on the RHS of (3.19) is at most $\exp\left(-|\square|^{-\rho(\square)\delta(1+\epsilon_1)}\right)$ for arbitrary $\epsilon_1 > 0$. Next, from Lemma 2.8, note that $\mathbb{P}\left(\tau_t > \tau_n^{\square,+}(t,s)\right)$ is at most $C\frac{|\square|^{1+2\delta}}{s^2} = C|\square|^{-2\delta_1}$ for some constant $C > 0$, where recall $s = (L_n + R_n)^{1/2+\delta+\delta_1} = \Theta\left(|\square|^{1/2+\delta+\delta_1}\right)$. Thus, choosing $\epsilon_1 > 0$ small enough so that $\delta < [2(1+\rho(\square)(1+\epsilon_1))]^{-1}$, made possible by the fact $\delta < [2(1+\rho(\square))]^{-1}$ (implied by $\delta < \mathcal{Z}(\alpha,\beta)$, recall (1.11)), and then choosing $\delta_1 > \rho(\square)\delta(1+\epsilon_1)$, which satisfies $\delta + \delta_1 < 1/2$, we note that

$$\mathbb{P}\left(D_u(\tau_n^{\square,+}(t,s)) > 0\right) + \mathbb{P}\left(\tau_t > \tau_n^{\square,+}(t,s)\right)$$
$$\leq \exp\left(-|\square|^{-\rho(\square)\delta(1+\epsilon_1)}\right) + |\square|^{-2\delta_1} \leq \exp\left(-(1-c)|\square|^{-\rho(\square)\delta(1+\epsilon_1)}\right),$$

for any constant $c > 0$. As a consequence, from (3.18) we note that for constants $c, c' > 0$,

$$\mathbb{P}\left(\mathcal{G}_{t(\delta)}^* \text{ has no isolated vertices}\right) \leq \exp\left(-(1-c)L^{1-\alpha\delta(1+\epsilon_1)} - (1-c')L^{1-\beta\delta(1+\epsilon_2)}\right),$$

Recalling (3.17), we thus note that $\mathbb{P}\left(B_{t(\delta)} \text{ is disconnected}\right) = o(1)$ if we can further choose $\epsilon, \epsilon_1, \epsilon_2 > 0$ sufficiently small such that

$$\exp\left(L^{4\delta+\epsilon} - (1-c)L^{1-\alpha\delta(1+\epsilon_1)} - (1-c')L^{1-\beta\delta(1+\epsilon_2)}\right) = o(1),$$

which is indeed possible when either $\delta < \frac{1}{4+\alpha} \wedge \frac{1}{\beta}$, or $\delta < \frac{1}{4+\beta} \wedge \frac{1}{\alpha}$, and this is implied by the fact that $\delta < \mathcal{Z}(\alpha,\beta)$, recall (1.11). $\qquad\square$

4 Discussion

4.1 Connectivity threshold lower bound for the unipartite case

For the unipartite case as was considered in [10] and [8], an argument similar to the proof of Theorem 1.5 gives that the corresponding simple graph process is disconnected **whp** after the inclusion of $n^{1+\delta}$ edges for any $\delta \in (0, 1/(4+\alpha))$, here n is the total number of vertices, and recall edges are included with probability proportional to $(d_u(t)+\alpha)(d_v(t)+\alpha)$. Thus, the simple graph connectivity threshold is at least $n^{1+1/(4+\alpha)}$ many edges. As mentioned in the introduction, this result is a small first step towards Pittel's conjecture that the connectivity threshold for the simple graph is $n^{1+\alpha^{-1}}$.

4.2 Multipartite case

The natural multipartite case, where the vertex set V is multipartitioned as $V = V_1 \cup \cdots \cup V_k$, and edges between non-adjacent vertices appear with probability proportional to $(d_u(t)+\alpha_i)(d_v(t)+\alpha_j)$ for $u \in V_i$ and $v \in V_j$, turns out to be more difficult. The analog of Proposition 2.2 can be proved, where in this case one has to consider a multipartite configuration model (MCM). This is a generalization of the BCM. Here for each vertex its degree into each of the partitions is provided. Then one constructs BCM's between each partition pair, and the union of all of these edges form the MCM. For a brief discussion of this model, see van der Hofstad [6, Chapter 9]. Proving an analog of Theorem 1.1 following the recipe of this paper (which in turn is inspired by the route-map of the paper [8]) requires – firstly, establishing the giant component threshold for the multipartite configuration model when the degree sequences between partition pairs are given; secondly, and more importantly, for a given partition pair, to understand the behavior of these degrees between the said pair. The latter seems the harder step as these degrees are not negative binomials anymore (unless all the α_i are equal), however, the former is an interesting question in its own right. We leave these explorations for future work.

References

[1] A.-L. Barabási. Network science. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 371(1987):20120375, 2013.

[2] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *science*, 286(5439):509–512, 1999.

[3] T. Bohman and A. Frieze. Avoiding a giant component. *Random Structures & Algorithms*, 19(1):75–85, 2001.

[4] P. Erdös and A. Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hungar. Acad. Sci*, 5:17–61, 1960.

[5] A. Frieze and M. Karoński. *Introduction to random graphs*. Cambridge University Press, 2015.

[6] R. van der Hofstad. *Random graphs and complex networks*, volume 2. Cambridge university press, 2024.

[7] R. van der Hofstad, J. Komjáthy, and V. Vadon. Phase transition in random intersection graphs with communities. *Random Structures & Algorithms*, 60(3):406–461, 2022.

[8] S. Janson and L. Warnke. Preferential attachment without vertex growth: Emergence of the giant component. *The Annals of Applied Probability*, 31(4):1523–1547, 2021.

[9] M. Molloy, E. Surya, and L. Warnke. The degree-restricted random process is far from uniform. *arXiv preprint arXiv:2211.00835*, 2022.

[10] B. Pittel. On a random graph evolving by degrees. *Advances in Mathematics*, 223(2):619–671, 2010.

[11] O. Riordan and L. Warnke. Explosive percolation is continuous. *Science*, 333(6040):322–324, 2011.

[12] O. Riordan and L. Warnke. Achlioptas process phase transitions are continuous. *The Annals of Applied Probability*, pages 1450–1464, 2012.