

## Phase transition in random intersection graphs with communities

van der Hofstad, Remco; Komjáthy, Júlia; Vadon, Viktória

**DOI**

[10.1002/rsa.21063](https://doi.org/10.1002/rsa.21063)

**Publication date**

2021

**Document Version**

Final published version

**Published in**

Random Structures and Algorithms

**Citation (APA)**

van der Hofstad, R., Komjáthy, J., & Vadon, V. (2021). Phase transition in random intersection graphs with communities. *Random Structures and Algorithms*, 60 (2022)(3), 406-461. <https://doi.org/10.1002/rsa.21063>

**Important note**

To cite this publication, please use the final published version (if applicable). Please check the document version above.

**Copyright**

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

**Takedown policy**

Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.

## RESEARCH ARTICLE

WILEY

# Phase transition in random intersection graphs with communities

Remco van der Hofstad<sup>1</sup> | Júlia Komjáthy<sup>2</sup> | Viktória Vadon<sup>3</sup>

<sup>1</sup>Department of Mathematics and Computer Science, Eindhoven University of Technology, Eindhoven, The Netherlands

<sup>2</sup>Department of Mathematics and Computer Science, Delft University of Technology, Delft, The Netherlands

<sup>3</sup>Institute of Mathematics, University of Miskolc, Miskolc, Hungary

## Correspondence

Júlia Komjáthy, Department of Mathematics and Computer Science, Delft University of Technology, Postbus 5, 2600 AA Delft, The Netherlands.

Email: J.Komjathy@tudelft.nl

## Abstract

The “random intersection graph with communities” (RIGC) models networks with communities, assuming an underlying bipartite structure of groups and individuals. Each group has its own internal structure described by a (small) graph, while groups may overlap. The group memberships are generated by a bipartite configuration model. The model generalizes the classical random intersection graph model, a special case where each community is a complete graph. The RIGC model is analytically tractable. We prove a phase transition in the size of the largest connected component in terms of the model parameters. We prove that percolation on RIGC produces a graph within the RIGC family, also undergoing a phase transition with respect to size of the largest component. Our proofs rely on the connection to the bipartite configuration model. Our related results on the bipartite configuration model are of independent interest, since they shed light on interesting differences from the unipartite case.

## KEYWORDS

bipartite configuration model, community structure, overlapping communities, percolation, phase transition, random intersection graphs, random networks

**Abbreviations:** BCM, bipartite configuration model; BP, branching process; CM, configuration model; iid, independent and identically distributed; lhs, left-hand side; pmf, probability mass function; rhs, right-hand side; RIG, random intersection graph; RIGC, random intersection graph with communities; uar, uniformly at random; UI, uniformly integrable; whp, with high probability; wlog, without loss of generality; wrt, with respect to.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2021 The Authors. *Random Structures and Algorithms* published by Wiley Periodicals LLC.

## 1 | INTRODUCTION

Real-world networks often exhibit a higher amount of clustering (also called transitivity [41, Chapter 7.9, 11]), that is, a larger amount of triangles than expected by pure chance [51]. A possible explanation for this phenomenon are *communities*: (small) subgraphs that are (significantly) denser than the network average. Such a community structure has been observed in several real-life networks [23], such as the Internet, collaboration networks, and social networks. In particular, many networks are well-explained [27, 28] by an underlying (possibly hidden) structure of individuals and groups that the individuals are part of. This is the kind of model we focus on in this article. The prime example is collaboration networks, such as the Internet movie database IMDb or the ArXiv, where the “individuals” are the actors and actresses or the authors, and the “groups” are the movies or articles they collaborate in. We can also model a social network in a similar fashion, giving “groups” the interpretation of families, common interests, workplaces, or cities. We take most of our terminology from social networks, however the model we present below is more widely applicable, for any network that builds on a group structure.

Due to the complexity of real-world networks, they are often modeled using *random graphs* [12, 21, 31]. Models are built to mimic some empirically observed properties of the network that we consider as defining features, such as degree structure, clustering, small-world property, and so forth. Then one may study further properties and processes of interest, such as network evolution and information or epidemic spreading processes, on the model to make predictions for real-life networks.

The traditional random graph model for networks with a group structure is the random intersection graph (RIG) [6, 7, 9–11, 20, 24, 40, 42, 43, 52, 53]. In this model, the underlying group structure mentioned above is represented by a bipartite graph, where the two partitions correspond to the individuals (people) and the groups (or attributes), and an edge represents a group membership. The group memberships, that is, connections in this bipartite graph, are random. Individuals are then connected in the intersection graph when they are together in a group, that is, share at least one group as neighbor in the bipartite graph. As a result, the members of a group form a complete subgraph, though in some variations of the model this complete graph is thinned [32, 40]. An other direction of research focuses on networks where the building blocks are communities that have an arbitrary internal structure, however no overlap [3, 4, 49, 50]. In [48] (see also [47] for an extended version), we have introduced a generalization of the RIG model that also incorporates an *arbitrary* internal structure for each group, which we call the random intersection graph with communities (RIGC). The RIGC thus combines the efforts to model networks with overlapping communities, as well as using arbitrary communities as building blocks. The RIGC is applicable to real-life network data, while we also derive rigorous analytic results.

In [48] (see also [34] for related local weak convergence results for RIGs), we have studied “local” properties of the RIGC model, such as local weak convergence (convergence of subgraph counts), degrees, local clustering coefficient, and the overlapping structure of communities. In this article, we instead focus on “global” properties, in particular the *giant component problem* and *percolation* (defined shortly), as well as the relation between “local” and “global” properties.

The *giant component problem* studies whether there exists a component containing a *linear* proportion of vertices. It has garnered quite some interest in the random graph literature since the seminal work of Erdős and Rényi [22], and has been studied on several other models (e.g., Chung–Lu model [17, 18] and configuration model [14, 30, 38, 39], see also the survey by Spencer [44] and the references therein). We prove that as we vary the parameters of the RIGC, the size of the largest component undergoes a *phase transition*: either all components have sublinear sizes as the network size grows, or there exists a unique *giant component* containing a constant fraction of the vertices, while the rest of the

components have sublinear sizes. When a giant component exists, we are able to further characterize it. To solve the giant component problem for the RIGC, we also solve it for a bipartite version of the configuration model (BCM), which is of independent interest.

*Percolation* means keeping each edge of a graph independently with a fixed probability, and it has applications related to epidemiology and attack vulnerability of networks. We give a detailed introduction to percolation in Section 2.5.1. We prove that percolation on the RIGC model can be represented as an RIGC with different parameters; we thus find that percolation on the RIGC undergoes a similar phase transition, and identify the asymptotic critical point implicitly.

### 1.1 | Main contribution and novelty of methodology

The main innovation of this article is the methodology required to solve the giant component problem for the bipartite configuration model. We make a non-trivial adaptation of the continuous-time exploration algorithm of the configuration model, originating from Janson and Luczak [30], to the bipartite case. The analysis of the bipartite case involves the study of a death process where jumps occasionally happen with *infinite* rate.

Showing that percolation on the RIGC stays within the family of models suggests that the RIGC is also a natural way to study percolation on the classical random intersection graph.

Finally, we combine the results on the giant component of the RIGC from the continuous-time exploration with the local weak convergence results from [48] (see also [34]), to obtain further properties of the giant component.

### 1.2 | Organization of the article

In Sections 2.1 and 2.2, we introduce the model and our assumptions. In Sections 2.3–2.5, we state our results on the largest component of the RIGC, the largest component of the BCM and percolation on the RIGC, respectively. In Section 3, we prove the results on the phase transition in the BCM as well as the RIGC using the adapted continuous-time exploration, and in particular, in Section 3.3, we analyze the death process with occasional instantaneous jumps. In Section 4, we prove the phase transition of (bond) percolation in the RIGC as a consequence of the largest component phase transition. Finally, in Section 5, we study further properties of the largest component of the RIGC.

### 1.3 | Notational conventions

To study the asymptotic behavior of various quantities, we will consider a sequence of graphs and consequently, a sequence of input parameters, both indexed by  $n \in \mathbb{N}$ . We note that  $n$  does not necessarily mean the size or any other parameter of the graph, and to keep the notation light, we often omit indicating the dependence on  $n$ , as long as it does not cause confusion. Throughout this article, we denote the set of positive integers as  $\mathbb{Z}^+$  and the set of non-negative integers as  $\mathbb{N}$ . The notions  $\xrightarrow{\mathbb{P}}$  and  $\xrightarrow{d}$  stand for convergence in probability and convergence in distribution (weak convergence), respectively. We write  $X \stackrel{d}{=} Y$  to mean that the random variables  $X$  and  $Y$  have the same distribution. For an  $\mathbb{N}$ -valued random variable  $X$  such that  $\mathbb{E}[X] < \infty$ , we define its *size-biased* distribution  $X^*$  and its shifted version  $\tilde{X} := X^* - 1$  with the following probability mass functions (pmf): for all  $k \in \mathbb{N}$ ,

$$\mathbb{P}(X^* = k) = k \mathbb{P}(X = k) / \mathbb{E}[X], \quad \mathbb{P}(\tilde{X} = k) = \mathbb{P}(X^* - 1 = k). \quad (1.1)$$

For a random variable  $X$  taking values in  $\mathbb{N}$ , we denote its probability generating function by  $G_X : [0, 1] \rightarrow [0, 1]$ , given by

$$G_X(z) := \mathbb{E} [z^X] = \sum_{k=0}^{\infty} \mathbb{P}(X = k)z^k. \tag{1.2}$$

Note that  $G_{X^*}(z) = zG'_X(z)/\mathbb{E}[X]$  and  $G_{\tilde{X}}(z) = G_{X^*}(z)/z = G'_X(z)/\mathbb{E}[X]$ . We say that a sequence of events  $(A_n)_{n \in \mathbb{N}}$  occurs with high probability (whp), when  $\lim_{n \rightarrow \infty} \mathbb{P}(A_n) = 1$ . For two (possibly) random sequences  $(X_n)_{n \in \mathbb{N}}$  and  $(Y_n)_{n \in \mathbb{N}}$ , we say that  $X_n = o_{\mathbb{P}}(Y_n)$  if  $X_n/Y_n \xrightarrow{\mathbb{P}} 0$  as  $n \rightarrow \infty$ . We say that the collection  $(X_a)_{a \in \mathcal{A}}$ , indexed by a set  $\mathcal{A}$ , of  $\mathbb{R}$ -valued random variables is uniformly integrable (UI) if  $\lim_{K \rightarrow \infty} \sup_{a \in \mathcal{A}} \mathbb{E} [X_a \mathbb{1}_{\{|X_a| \geq K\}}] = 0$ , where  $\mathbb{1}_A$  denotes the indicator of the event  $A$ . We denote  $[n] := \{1, 2, \dots, n\}$ . For a graph  $G$ , we denote its vertex set by  $\mathcal{V}(G)$  and its edge set by  $\mathcal{E}(G)$ . We sometimes write  $|G|$  for the number of vertices in  $G$ .

## 2 | MODEL AND RESULTS

In this section, we introduce our model and study some of its global properties.

### 2.1 | Definition of the random intersection graph with communities

This section is a more concise transcription of the model definition from the companion paper [48] on the local properties of the RIGC model. For more details on the construction, we direct the reader to that paper. Given the parameters  $\mathbf{d}^\ell$ ,  $\mathbf{Com}$ , whose meaning is explained below, we construct the random graph  $\text{RIGC}(\mathbf{d}^\ell, \mathbf{Com})$  in two steps. First, we construct the bipartite matching that determines the group memberships, then we show how to obtain the RIGC based on the group memberships.

#### 2.1.1 | The parameters

We start with a bipartite graph comprising of individuals and groups. We call the set of individuals the left-hand side (lhs) partition  $\mathcal{V}^\ell = [N_n]$ , where  $N_n \rightarrow \infty$  is the number of individuals. We may refer to an individual  $v \in \mathcal{V}^\ell$  as an  $\ell$ -vertex. We call its number of group memberships its  $\ell$ -degree (left-degree), and denote it by  $d_v^\ell = \ell\text{-deg}(v)$ . The parameter  $\mathbf{d}^\ell = (d_v^\ell)_{v \in \mathcal{V}^\ell}$  is the vector of  $\ell$ -degrees. Without loss of generality (wlog), we assume that  $\mathbf{d}^\ell \geq 1$  (element-wise).

Analogously, we call the set of groups the right-hand side (rhs) partition  $\mathcal{V}^r = [M_n]$ , where also the number of groups  $M_n$  satisfies  $M_n \rightarrow \infty$ , and we may refer to groups as  $r$ -vertices. Each  $r$ -vertex  $a$  is associated with a community graph  $\text{Com}_a$ , with properties explained below, and  $\mathbf{Com} = (\text{Com}_a)_{a \in \mathcal{V}^r}$  is the vector of community graphs. Let  $\mathcal{H}$  be the set of possible community graphs, that is, the set of all simple, finite, connected graphs. We label the vertices in each  $H \in \mathcal{H}$  arbitrarily by  $[|H|]$ , where  $|H|$  denotes the number of vertices in  $H$ , so that any two community graphs that are isomorphic are also labeled identically. We assume that each assigned community graph  $\text{Com}_a \in \mathcal{H}$  satisfies  $|\text{Com}_a| \geq 1$ . We call the number of vertices  $|\text{Com}_a|$  in  $\text{Com}_a$  the  $r$ -degree (right-degree) of group  $a$  and denote it by  $d_a^r = r\text{-deg}(a)$ . We collect all  $r$ -degrees in the vector  $\mathbf{d}^r := (d_a^r)_{a \in \mathcal{V}^r}$ .

#### 2.1.2 | Community memberships

In the bipartite graph of group memberships, the  $\ell$ - and  $r$ -degrees act as degrees. We refer to them together as  $\ell$ -degrees (bipartite degrees). To ensure the existence of a bipartite graph with these given

degrees, we assume and denote

$$\ell_n := \sum_{v \in \mathcal{V}^\ell} d_v^\ell = \sum_{a \in \mathcal{V}^r} d_a^r. \quad (2.1)$$

With the given  $\ell$ -degrees, we construct the group memberships according to a bipartite matching, described as follows. Denote the disjoint union of all vertices in community graphs by  $\mathcal{V}(\mathbf{Com})$ , which we call the set of community vertices (community roles). To each  $r$ -vertex  $a$ , we assign  $r$ -deg( $a$ )  $r$ -half-edges, labeled by  $(a, l)_{l \in [r\text{-deg}(a)]}$ . We think of  $(a, l)$  as the membership token corresponding to the community vertex  $j \in \mathcal{V}(\mathbf{Com})$  in  $\text{Com}_a$  with label  $l$ , which gives a natural correspondence between the set of  $r$ -half-edges and  $\mathcal{V}(\mathbf{Com})$ . To each  $\ell$ -vertex  $v$ , we assign  $\ell$ -deg( $v$ )  $\ell$ -half-edges, labeled by  $(v, i)_{i \in [\ell\text{-deg}(v)]}$ . In contrast to  $r$ -half-edges, the  $\ell$ -half-edges incident to the same  $\ell$ -vertex are interchangeable as membership tokens.

Denote the set of all bijections between the set of  $\ell$ -half-edges  $(v, i)_{i \in [\ell\text{-deg}(v)], v \in \mathcal{V}^\ell}$  to the set of  $r$ -half-edges  $(a, l)_{l \in [r\text{-deg}(a)], a \in \mathcal{V}^r}$  by  $\Omega_n$ . (Equivalently, these correspond to bijections between the set of  $\ell$ -half-edges and  $\mathcal{V}(\mathbf{Com})$ ). Let  $\omega_n \sim \text{Unif}[\Omega_n]$  denote a bipartite matching (or bipartite configuration) chosen uniformly at random (uar). Note that, by reindexing the half-edges, we can think of  $\omega_n$  as a permutation of  $[\ell_n]$ , thus  $|\Omega_n| = \ell_n!$ . If the  $\ell$ -half-edge  $(v, i)$  and the  $r$ -half-edge  $(a, l)$  are paired by  $\omega_n$ , then this intuitively means one of the community roles taken by  $v$  is the community vertex  $j \in \mathcal{V}(\mathbf{Com})$  in  $\text{Com}_a$  with label  $l$ , and we denote the indicator of this event by  $\mathbb{1}_{\{v \leftarrow j\}} = \mathbb{1}_{\{v \leftarrow j\}}(\omega_n)$ . Note that (almost surely)

$$\sum_{v \in \mathcal{V}^\ell} \mathbb{1}_{\{v \leftarrow j\}} = 1, \quad \sum_{j \in \mathcal{V}(\mathbf{Com})} \mathbb{1}_{\{v \leftarrow j\}} = \ell\text{-deg}(v). \quad (2.2)$$

The community memberships of the RIGC model (see Figure 1A) are determined by the uniform(ly random) bipartite matching  $\omega_n$  in the above fashion. Before we define the RIGC graph based on the community memberships, we make some observations about  $\omega_n$ :

*Remark 2.1* (Algorithmic pairing [48, Definition 2.1]). The uniform bipartite matching  $\omega_n$  can be produced sequentially, as follows. In each step, we pick an arbitrary unpaired half-edge, and match it to a uniform unpaired half-edge in the opposite partition. As the choices are arbitrary, they may even depend on the past of the pairing process.

*Remark 2.2* (The underlying BCM [48, Definition 2.1]). We may view the half-edges as tokens to form edges (rather than group membership tokens as above), as usual in the configuration model. Then the bipartite matching  $\omega_n$  also determines a bipartite (multi)graph, defined as follows. If the  $\ell$ -half-edge  $(v, i)$  and the  $r$ -half-edge  $(a, l)$  are matched, then we replace them by an edge labeled by  $(i, l)$  between  $v \in \mathcal{V}^\ell$  and  $a \in \mathcal{V}^r$ . Note that with the edge labels, we can recover the matched half-edges, thus this multigraph provides an equivalent representation of  $\omega_n$ . We refer to this labeled bipartite (multi)graph as the underlying bipartite configuration model.

Deleting the edge labels introduced above, we obtain the (classic) bipartite configuration model  $\text{BCM}(\mathbf{d}^\ell, \mathbf{d}^r)$  with degree sequences  $\mathbf{d}^\ell$  and  $\mathbf{d}^r$ .

### 2.1.3 | Community-projection

In the following, we complete the construction of the RIGC, based on the community structure defined by  $\omega_n$ . This step is entirely deterministic and we can think of it as an operator  $\mathcal{P}$  from  $\Omega_n$  into the set of *multigraphs*. For a discussion on why multigraphs arise and why we have chosen to work with

PHASE TRANSITION IN THE RIGC

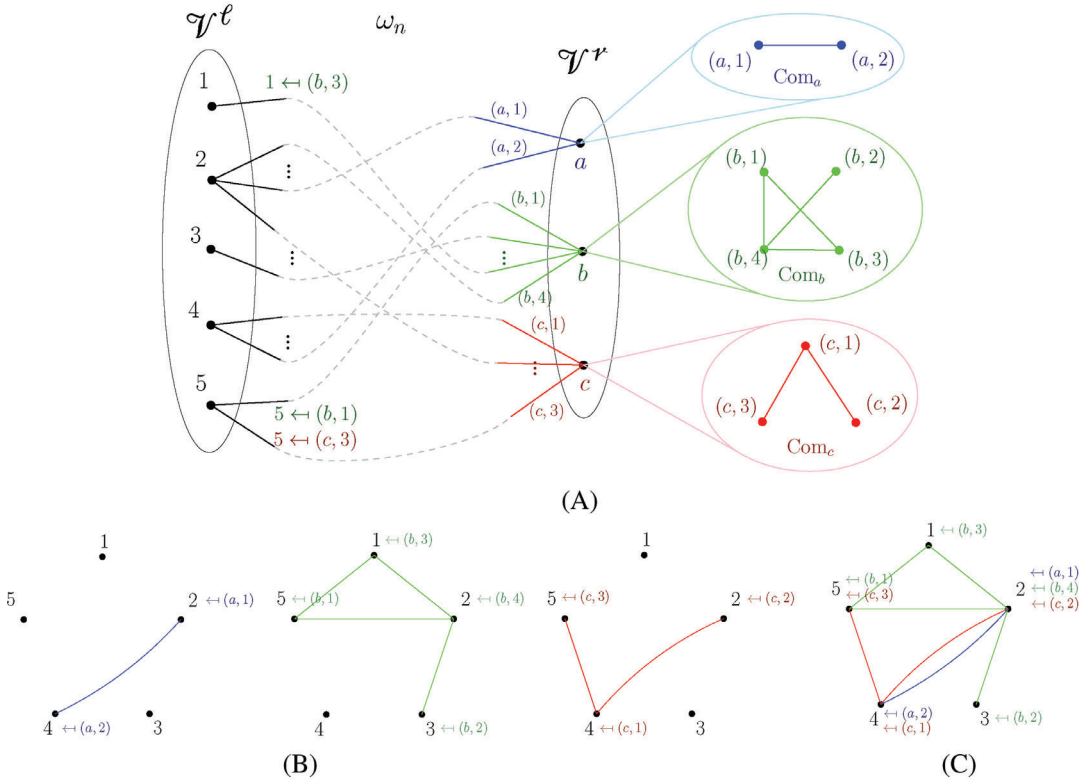


FIGURE 1 Construction of the RIGC graph. (A) Community roles assigned by the bipartite matching. (B) The projection of each community. Each edge between community roles is copied to the corresponding individuals (which are assigned the community roles forming the edge). (C) The resulting RIGC. The multigraph combining the communities

them, see the companion paper [48, Section 2.4]. In the following, we define the RIGC by its (random,  $\omega_n$ -dependent) edge multiplicities  $(X(v, w))_{v, w \in \mathcal{V}^\ell}$ .

Recall that we write  $v \leftarrow j$  when an individual  $v \in \mathcal{V}^\ell$  takes on community role  $j \in \mathcal{V}^r(\mathbf{Com})$ . Let us denote the disjoint union of edges in all community graphs by  $\mathcal{E}(\mathbf{Com})$ , which we refer to as the set of community edges. Intuitively, we construct the RIGC by copying each community edge  $(j_1, j_2)$  to the individuals  $(v, w)$  such that  $v \leftarrow j_1$  and  $w \leftarrow j_2$ . Formally, we define the multiplicity of the edge  $(v, w)$  as

$$X(v, w) = X(v, w; \omega_n) := \sum_{(j_1, j_2) \in \mathcal{E}(\mathbf{Com})} \mathbb{1}_{\{v \leftarrow j_1, w \leftarrow j_2\} \cup \{v \leftarrow j_2, w \leftarrow j_1\}}. \tag{2.3}$$

Then, the (random) degree of  $v \in \mathcal{V}^\ell$  in the resulting RIGC, sometimes referred to as projected degree ( $\rho$ -degree) for clarification, is given by

$$d_v^\rho = \rho\text{-deg}(v) := X(v, v) + \sum_{w \in \mathcal{V}^\ell} X(v, w) = 2X(v, v) + \sum_{w \in \mathcal{V}^\ell, w \neq v} X(v, w). \tag{2.4}$$

2.2 | Assumptions on the parameters

In this section, we introduce some notation and state the assumptions necessary for our results in Sections 2.3 and 2.5. We remark that the notation and assumptions are identical to those introduced in [48, Section 2.2].

### 2.2.1 | The bipartite degrees

We define uniformly chosen  $\ell$ - and  $r$ -vertices

$$V_n^\ell \sim \text{Unif}[\mathcal{V}^\ell], \quad V_n^r \sim \text{Unif}[\mathcal{V}^r], \quad (2.5)$$

and denote their degrees by

$$D_n^\ell := \ell\text{-deg}(V_n^\ell), \quad D_n^r := r\text{-deg}(V_n^r). \quad (2.6)$$

Further, denote the sets of  $\ell$ -vertices and  $r$ -vertices with degree  $k$ , respectively, by

$$\mathcal{V}_k^\ell := \{v \in \mathcal{V}^\ell : \ell\text{-deg}(v) = k\}, \quad \mathcal{V}_k^r := \{a \in \mathcal{V}^r : r\text{-deg}(a) = k\}. \quad (2.7)$$

Then the probability mass functions (pmf), for  $k \in \mathbb{Z}^+$ ,

$$p_k^{(n)} := |\mathcal{V}_k^\ell|/N_n, \quad q_k^{(n)} := |\mathcal{V}_k^r|/M_n, \quad (2.8)$$

describe the distribution of the variables  $D_n^\ell$  and  $D_n^r$ , as well as the empirical distribution of  $\mathbf{d}^\ell$  and  $\mathbf{d}^r$ , respectively. We collect the pmfs in the (infinite-dimensional) probability vectors  $\mathbf{p}^{(n)} = (p_k^{(n)})_{k \in \mathbb{Z}^+}$ ,  $\mathbf{q}^{(n)} = (q_k^{(n)})_{k \in \mathbb{Z}^+}$ .

### 2.2.2 | The empirical community distribution

Recall that the possible community graphs are  $\mathcal{H}$ , the set of simple, finite, connected graphs  $H$  with an arbitrary but fixed vertex labeling by  $[|H|]$ . For a fixed  $H \in \mathcal{H}$ , define

$$\mathcal{V}_H^r := \{a \in \mathcal{V}^r : \text{Com}_a = H\}. \quad (2.9)$$

We introduce the pmf

$$\mu_H^{(n)} := |\mathcal{V}_H^r|/M_n, \quad \boldsymbol{\mu}^{(n)} = (\mu_H^{(n)})_{H \in \mathcal{H}}. \quad (2.10)$$

Thus  $\boldsymbol{\mu}^{(n)}$  describes the empirical pmf of **Com**, as well as the pmf of  $\text{Com}_{V_n^r}$ , with  $V_n^r \sim \text{Unif}[\mathcal{V}^r]$ . Define

$$\mathcal{H}_k := \{H \in \mathcal{H} : |H| = k\}. \quad (2.11)$$

Note that  $q_k^{(n)} = \sum_{H \in \mathcal{H}_k} \mu_H^{(n)}$  (with  $q_k^{(n)}$  from (2.8)).

### 2.2.3 | The community degrees

For  $a \in \mathcal{V}^r$  and  $j \in \mathcal{V}(\text{Com}_a)$ , define the community degree ( $c$ -degree) of  $j$ , denoted by  $d_j^c$ , as the number of connections  $j$  has within  $\text{Com}_a$ . Recall that  $\mathcal{V}(\mathbf{Com})$  denotes the disjoint union of vertices in all communities, and let  $J_n \sim \text{Unif}[\mathcal{V}(\mathbf{Com})]$  be a uniformly chosen community vertex, among all  $\mathcal{H}_n$  possibilities. This is equivalent to choosing  $\text{Com}_a$  in a *size-biased fashion*, that is, picking  $\text{Com}_a$  with probability  $|\text{Com}_a|/\mathcal{H}_n$ , then picking  $J_n | \text{Com}_a \sim \text{Unif}[\mathcal{V}(\text{Com}_a)]$ .

Introduce the random variable  $D_n^c := d_{J_n}^c$ . We define the pmf that describes  $D_n^c$  as well as the empirical distribution of the collection  $(d_j^c)_{j \in \mathcal{V}(\mathbf{Com})}$ , by

$$\rho_k^{(n)} := \frac{1}{\mathcal{H}_n} \sum_{j \in \mathcal{V}(\mathbf{Com})} \mathbb{1}_{\{d_j^c = k\}}, \quad \boldsymbol{\rho}^{(n)} := (\rho_k^{(n)})_{k \in \mathbb{Z}^+}. \quad (2.12)$$



## 2.2.4 | Assumptions

Recall (2.6), (2.8), and (2.10). We can now summarize our assumptions on the model parameters:

**Assumption 2.3** ([48, Assumption 2.2]). The conditions for the empirical distributions are summarized as follows:

(A) There exists a random variable  $D^\ell$  with pmf  $\mathbf{p}$  s.t.  $\mathbf{p}^{(n)} \rightarrow \mathbf{p}$  pointwise as  $n \rightarrow \infty$ , that is,

$$D_n^\ell \xrightarrow{d} D^\ell. \quad (2.13)$$

(B)  $\mathbb{E}[D^\ell]$  is finite, and as  $n \rightarrow \infty$ ,

$$\mathbb{E}[D_n^\ell] \rightarrow \mathbb{E}[D^\ell]. \quad (2.14)$$

(C) There exists a probability mass function  $\boldsymbol{\mu}$  on  $\mathcal{H}$  such that  $\boldsymbol{\mu}^{(n)} \rightarrow \boldsymbol{\mu}$  pointwise as  $n \rightarrow \infty$ .

(1) Consequently, by  $q_k^{(n)} = \sum_{H \in \mathcal{H}_k} \mu_H^{(n)}$ , with the finite set  $\mathcal{H}_k$  from (2.11), there exists a random variable  $D^r$  with pmf  $\mathbf{q}$  such that  $\mathbf{q}^{(n)} \rightarrow \mathbf{q}$  pointwise as  $n \rightarrow \infty$ , or equivalently,

$$D_n^r \xrightarrow{d} D^r. \quad (2.15)$$

(D)  $\mathbb{E}[D^r]$  is finite, and as  $n \rightarrow \infty$ ,

$$\mathbb{E}[D_n^r] \rightarrow \mathbb{E}[D^r]. \quad (2.16)$$

*Remark 2.4* (Consequences of Assumption 2.3 [48, Remark 2.3]). We note the following:

(i) Recall  $N_n = |\mathcal{V}^\ell|$  and  $M_n = |\mathcal{V}^r|$ . By its definition in (2.1),  $\ell_n = N_n \mathbb{E}[D_n^\ell] = M_n \mathbb{E}[D_n^r]$ . By Assumption 2.3(B,D),

$$M_n/N_n = \mathbb{E}[D_n^\ell]/\mathbb{E}[D_n^r] \rightarrow \mathbb{E}[D^\ell]/\mathbb{E}[D^r] =: \gamma \in \mathbb{R}^+. \quad (2.17)$$

(ii) Since  $\boldsymbol{\rho}^{(n)}$  (see (2.12)) can be obtained from  $\boldsymbol{\mu}^{(n)}$ , Assumption 2.3(C) also implies that there exists a random variable  $D^e$  with pmf  $\boldsymbol{\rho}$  such that  $\boldsymbol{\rho}^{(n)} \rightarrow \boldsymbol{\rho}$  pointwise as  $n \rightarrow \infty$ , or equivalently,  $D_n^e \xrightarrow{d} D^e$ .

(iii) Assumption 2.3(A,B) imply that  $d_{\max}^\ell := \max_{v \in \mathcal{V}^\ell} d_v^\ell = o(\ell_n)$ , and similarly, conditions (C1-D) imply that  $d_{\max}^r := \max_{a \in \mathcal{V}^r} d_a^r = o(\ell_n)$ .

*Remark 2.5* (Random parameters [48, Remark 2.4]). The results in Section 2.3 remain valid when the sequence of parameters  $(\mathbf{d}^\ell, \mathbf{Com})$  (resp.,  $(\mathbf{d}^\ell, \mathbf{d}^r)$ ) is random itself. In this case, we require that  $N_n \xrightarrow{\mathbb{P}} \infty$  (i.e., for all  $K \in \mathbb{R}^+$ ,  $\mathbb{P}(N_n > K) \rightarrow 1$  as  $n \rightarrow \infty$ ) and  $M_n \xrightarrow{\mathbb{P}} \infty$ , and we replace Assumption 2.3(A–D) (resp., Assumption 2.3 (A,B,C1,D)) by the conditions  $\mathbf{p}^{(n)} \xrightarrow{\mathbb{P}} \mathbf{p}$  pointwise,  $\mathbb{E}[D_n^\ell | \mathbf{d}^\ell] \xrightarrow{\mathbb{P}} \mathbb{E}[D^\ell]$ ,  $\boldsymbol{\mu}^{(n)} \xrightarrow{\mathbb{P}} \boldsymbol{\mu}$  pointwise (resp.,  $\mathbf{q}^{(n)} \xrightarrow{\mathbb{P}} \mathbf{q}$ ) and  $\mathbb{E}[D_n^r | \mathbf{d}^r] \xrightarrow{\mathbb{P}} \mathbb{E}[D^r]$ . For a similar setting in the configuration model, see [45, Remark 7.9], where this is spelled out in more detail.

Note that analogously to Remark 2.4(i), the conditions of Remark 2.5 imply that  $M_n/N_n \xrightarrow{\mathbb{P}} \gamma$ .

### 2.3 | Results on the largest component of the random intersection graph with communities

In this article, we study global properties of the RIGC model; its local properties have been studied in the companion paper [48] (see also [34]). In particular, in this section we study the largest connected component of the RIGC. We prove a phase transition in the size of the largest component in terms of the model parameters, and explicitly identify the conditions under which a unique linear-sized component exists. We study further properties of this component, that is, its degree distribution and its number of edges. Denote the largest connected component (the component containing the most  $\ell$ -vertices, breaking ties arbitrarily) by  $\mathcal{C}_1 = \mathcal{C}_1^{(n)}$ , and the second largest by  $\mathcal{C}_2 = \mathcal{C}_2^{(n)}$ . Recall (1.1), (1.2), (2.6), and (2.8).

**Theorem 2.6** (Size of the largest component). *Consider  $\text{RIGC}(d^\ell, \mathbf{Com})$  under Assumption 2.3, and further assume that  $p_2 + q_2 < 2$ . Then, there exists  $\eta_\ell \in [0, 1]$ , the smallest solution of the fixed point equation*

$$\eta_\ell = G_{\tilde{D}^\ell}(G_{\tilde{D}^\ell}(\eta_\ell)), \quad (2.18)$$

and  $\xi_\ell := 1 - G_{D^\ell}(\eta_\ell) \in [0, 1]$  such that

$$|\mathcal{C}_1|/N_n \xrightarrow{\mathbb{P}} \xi_\ell. \quad (2.19)$$

Furthermore,  $\xi_\ell > 0$  exactly when

$$\mathbb{E}[\tilde{D}^\ell] \mathbb{E}[\tilde{D}^{r^\ell}] > 1, \quad (2.20)$$

which we call the supercritical case. By convention, (2.20) holds when  $\mathbb{E}[\tilde{D}^\ell] \mathbb{E}[\tilde{D}^{r^\ell}] = \infty$ . In the case where (2.20) holds,  $\mathcal{C}_1$  is unique in the sense that  $|\mathcal{C}_2| = o_{\mathbb{P}}(N_n)$ , and we call  $\mathcal{C}_1$  the giant component.

We prove Theorem 2.6 in Section 3 subject to the upcoming Theorem 2.11, and we discuss the relevance of the condition  $p_2 + q_2 < 2$  in Section 2.4.2. Note that the size of the largest connected component only depends on  $D^\ell$  and  $D^{r^\ell} = |\mathcal{H}|$ , where  $\mathcal{H} \in \mathcal{H}$  follows distribution  $\mu$  and  $|\mathcal{H}|$  denotes the number of vertices in  $\mathcal{H}$ ; this is because the communities are connected. Consequently, Theorem 2.6 applies to the classical RIG, which is the special case of RIGC with complete graph communities.

We continue by studying the degree distribution and the number of edges in the giant component. Naturally, these quantities depend more sensitively on  $\mu$  and are non-trivial: our results show that the degree distribution in the giant component is considerably different from the degree distribution of the whole graph (unless  $\xi_\ell = 1$  and the giant component contains almost all vertices). The reason for this is a size-biasing effect of the giant (see Theorem 2.11 and Corollary 2.12). Recall (2.4) and (2.7) and define

$$\mathcal{V}_d^{r^\ell} := \{v \in \mathcal{V}^\ell : r\text{-deg}(v) = d\}. \quad (2.21)$$

For  $H \in \mathcal{H}$  and  $c \in \mathbb{Z}^+$ , define  $\nu(c | H) := |\{j \in \mathcal{V}(H) : c\text{-deg}(j) = c\}|$ , the number of vertices in  $H$  with  $c$ -degree  $c$ .

**Theorem 2.7** (Degrees in the giant). *Consider  $\text{RIGC}(d^\ell, \mathbf{Com})$  under Assumption 2.3, additionally assuming the supercriticality condition (2.20). Define  $\eta_{r^\ell} := G_{\tilde{D}^{r^\ell}}(\eta_\ell)$ , with  $\eta_\ell$  from Theorem 2.6. For*

$k \in \mathbb{Z}^+, d \in \mathbb{N}$ , define

$$A(k, d) := p_k \sum_{H_1, \dots, H_k \in \mathcal{H}} \sum_{\substack{c_1, \dots, c_k \in \mathbb{N} \\ c_1 + \dots + c_k = d}} \left( 1 - \eta_r^{\sum_{i=1}^k (|H_i| - 1)} \right) \prod_{i=1}^k \frac{\nu(c_i | H_i) \mu_{H_i}}{\mathbb{E}[D^r]}. \tag{2.22}$$

Then, the proportion of individuals that have  $k$  group memberships, total degree  $d$  and are in the giant component converges as  $n \rightarrow \infty$ :

$$\frac{|\mathcal{V}_k^\ell \cap \mathcal{V}_d^r \cap \mathcal{C}_1|}{N_n} \xrightarrow{\mathbb{P}} A(k, d). \tag{2.23}$$

The proof of Theorem 2.7 is deferred to Section 5.2, where we also give a heuristic interpretation of  $A(k, d)$ . The quantities  $\eta_\ell$  and  $\eta_r$  appearing in Theorems 2.6 and 2.7 have the nice interpretation as the asymptotic extinction probabilities of corresponding branching processes started from left and right half-edges, respectively. As a result,  $\xi_\ell$  and  $\xi_r := 1 - G_{D^r}(\eta_r)$  are the asymptotic survival probabilities of uniform left and right vertices. See Sections 5.1 and 5.2 for more details.

It follows (via a truncation argument) from Theorem 2.7 and (2.19) that the empirical degree distribution in the giant converges:

$$\frac{|\mathcal{V}_d^r \cap \mathcal{C}_1|}{|\mathcal{C}_1|} = \sum_{k \in \mathbb{Z}^+} \frac{|\mathcal{V}_k^\ell \cap \mathcal{V}_d^r \cap \mathcal{C}_1|}{N_n} \frac{N_n}{|\mathcal{C}_1|} \xrightarrow{\mathbb{P}} \sum_{k \in \mathbb{Z}^+} A(k, d) / \xi_\ell. \tag{2.24}$$

While the expression in (2.22) seems quite involved, the following remark shows that in fact, it is closely related to the limiting degree distribution of the whole graph. Recall  $D^\ell$  from Assumption 2.3(A), and  $D^c$  from Remark 2.4(ii). We recall from [48, (2.20)] that the limiting degree is  $D^{\mathcal{R}} = \sum_{i=1}^d D_{(i)}^c$ , where  $(D_{(i)}^c)_{i \in \mathbb{Z}^+}$  are independent, identically distributed (iid) copies of  $D^c$ , and also independent of  $D^\ell$ .

*Remark 2.8* (Relation of Theorem 2.7 and [48, Corollary 2.8]). We now explore the relation between the degree distribution in the giant and the degree distribution in the entire graph. First, note that for any fixed  $c$ , with  $\rho_c$  from Remark 2.4(ii),

$$\sum_{H \in \mathcal{H}} \frac{\nu(c | H) \mu_H}{\mathbb{E}[D^r]} = \rho_c = \mathbb{P}(D^c = c). \tag{2.25}$$

Here, the denominator  $\mathbb{E}[D^r]$  only serves for renormalization, since  $\mu^{(n)}$  is a distribution on  $\mathcal{V}^r$  with size  $M_n$ , while  $\rho^{(n)}$  is a distribution on  $\mathcal{V}(\mathbf{Com})$  with size  $\ell_n$ . The factor  $\left( 1 - \eta_r^{\sum_{i=1}^k (|H_i| - 1)} \right)$  in (2.22) heuristically corresponds to belonging to the giant, which is later justified by (2.36). By (2.25), omitting the factor  $\left( 1 - \eta_r^{\sum_{i=1}^k (|H_i| - 1)} \right)$  from (2.22), its rhs becomes a convolution:

$$\begin{aligned} p_k \sum_{H_1, \dots, H_k \in \mathcal{H}} \sum_{\substack{c_1, \dots, c_k \in \mathbb{N} \\ c_1 + \dots + c_k = d}} \prod_{i=1}^k \frac{\nu(c_i | H_i) \mu_{H_i}}{\mathbb{E}[D^r]} &= p_k \sum_{\substack{c_1, \dots, c_k \in \mathbb{N} \\ c_1 + \dots + c_k = d}} \prod_{i=1}^k \rho_{c_i} \\ &= p_k \mathbb{P} \left( \sum_{i=1}^k D_{(i)}^c = d \right) = \mathbb{P}(D^\ell = k) \mathbb{P}(D^{\mathcal{R}} = d \mid D^\ell = k) = \mathbb{P}(D^{\mathcal{R}} = d, D^\ell = k), \end{aligned} \tag{2.26}$$

which is the asymptotic joint distribution of  $\ell$ - and  $\rho$ -degrees in the whole graph. Indeed, combining [48, Corollary 2.8 and (2.20)] implies that

$$\frac{|\mathcal{V}_k^\ell \cap \mathcal{V}_d^\rho|}{N_n} \xrightarrow{\mathbb{P}} \mathbb{P}(D^\rho = d, D^\ell = k). \tag{2.27}$$

Next, we state our result regarding the number of edges in the giant component. Recall  $D_n^c$  with pmf  $\phi^{(n)}$  from (2.12),  $\eta_r$  from Theorem 2.7 and  $\gamma$  from (2.17).

**Theorem 2.9** (Edges in the giant). *Consider  $\text{RIGC}(d^\ell, \mathbf{Com})$  under Assumption 2.3 and the supercriticality condition (2.20), and additionally assume that*

$$(D_n^c)_{n \in \mathbb{N}} \text{ is uniformly integrable (UI),} \tag{2.28}$$

which in particular implies that  $\mathbb{E}[D_n^c] \rightarrow \mathbb{E}[D^c] < \infty$  as  $n \rightarrow \infty$ . Let  $\mathcal{H}$  denote a random graph with pmf  $\mu$ . Then the number of edges in the giant component  $\mathcal{E}_1$  of the RIGC satisfies, as  $n \rightarrow \infty$ ,

$$\frac{|\mathcal{E}(\mathcal{E}_1)|}{N_n} \xrightarrow{\mathbb{P}} \gamma \mathbb{E} \left[ |\mathcal{E}(\mathcal{H})| \left( 1 - \eta_r^{|\mathcal{H}|} \right) \right]. \tag{2.29}$$

We sketch the proof shortly below, which requires the following lemma:

**Lemma 2.10** (Uniform integrability). *The following statements are equivalent:*

- (i)  $(D_n^c)_{n \in \mathbb{N}}$  is UI;
- (ii)  $(D_n^\rho)_{n \in \mathbb{N}}$  is UI;
- (iii)  $(|\mathcal{E}(\mathcal{H}_n)|)_{n \in \mathbb{N}}$  is UI.

The proof of Lemma 2.10 is rather technical and somewhat tedious and we postpone it to Appendix A.1, but we discuss the relevance of Lemma 2.10 and the condition (2.28) now. The statement in Lemma 2.10(iii), or equivalently, the condition (2.28), is the necessary and sufficient condition for the rhs of (2.29) to be finite. Since this condition takes the community structure into account, it is more refined than moment conditions on the community size  $D_n^c$ . By our assumption that community graphs are simple and connected,  $|H| - 1 \leq |\mathcal{E}(H)| \leq |H|(|H| - 1)/2$ , which implies  $\mathbb{E}[D_n^c] - 1 \leq \mathbb{E}[|\mathcal{E}(\mathcal{H}_n)|] \leq \mathbb{E}[D_n^c(D_n^c - 1)]/2$ . Thus the condition Lemma 2.10(iii) is weaker than  $\mathbb{E}[(D_n^c)^2] \rightarrow \mathbb{E}[(D^\rho)^2] < \infty$  (which is sufficient, but not necessary), but stronger than  $\mathbb{E}[D_n^c] \rightarrow \mathbb{E}[D^c] < \infty$ , that is Assumption 2.3 (D). In the general case under Assumption 2.3, it is still possible that  $\mathbb{E}[|\mathcal{E}(\mathcal{H}_n)|]$  diverges, which might imply that  $|\mathcal{E}(\mathcal{E}_1)|/N_n$  diverges and Theorem 2.9 does not hold.

*Sketch of proof of Theorem 2.9 subject to Lemma 2.10.* Theorem 2.9 follows from Theorem 2.7 under the extra condition of uniform integrability in (2.28). Denote a uniform  $\ell$ -vertex  $V_n^\ell \sim \text{Unif}[\mathcal{V}^\ell]$  and its (projected) degree  $D_n^\rho = \rho\text{-deg}(V_n^\ell)$  (see (2.4)). Let  $\mathcal{H}_n$  denote a random graph with pmf  $\mu^{(n)}$  from (2.10). In [48, Corollary 2.8], the distributional limit of  $D_n^\rho$  is established as  $D^\rho \stackrel{d}{=} \sum_{i=1}^{D^\ell} D_{(i)}^c$ , where  $D_{(i)}^c$  are iid copies of  $D^c$  that are independent of  $D^\ell$ . Under condition (2.28), Lemma 2.10(ii) ensures that  $\mathbb{E}[D^\rho] < \infty$ , which implies that the normalized degree in the giant  $2|\mathcal{E}(\mathcal{E}_1)|/N_n$  also has uniformly finite mean (recall also Remark 2.8). Thus we can show that under (2.28), as  $n \rightarrow \infty$ ,

$$\frac{|\mathcal{E}(\mathcal{E}_1)|}{N_n} = \frac{1}{2} \sum_{d \in \mathbb{N}} d \sum_{k \in \mathbb{Z}^+} \frac{|\mathcal{V}_k^\ell \cap \mathcal{V}_d^\rho \cap \mathcal{E}_1|}{N_n} \xrightarrow{\mathbb{P}} \frac{1}{2} \sum_{d \in \mathbb{N}} d \cdot \sum_{k \in \mathbb{Z}^+} A(k, d). \tag{2.30}$$

Under condition (2.28), Lemma 2.10(iii) implies that  $\mathbb{E}[|\mathcal{E}(\mathcal{H})|] < \infty$ , so that the rhs of (2.29) is finite. Then, we can show that the average degree is in fact related to the number of edges in the community graphs:

$$\frac{1}{2} \sum_{d \in \mathbb{N}} d \cdot \sum_{k \in \mathbb{Z}^+} A(k, d) = \gamma \cdot \mathbb{E} \left[ |\mathcal{E}(\mathcal{H})| \left( 1 - \eta_r^{|\mathcal{H}|} \right) \right]. \tag{2.31}$$

We provide the details of the proof in Appendix A.2.

### 2.4 | The largest component of the bipartite configuration model

In this section, we introduce our results on the largest component of the BCM (see Remark 2.2), which are of independent interest, and we further apply them to prove our results on the RIGC. Denote the largest component (the component containing the largest total number of vertices, with ties broken arbitrarily) of the  $\text{BCM}_n(\mathbf{d}^\ell, \mathbf{d}^r)$  by  $\mathcal{C}_{1,\delta} = \mathcal{C}_{1,\delta}^{(n)}$ , and the second largest by  $\mathcal{C}_{2,\delta} = \mathcal{C}_{2,\delta}^{(n)}$ . Recall  $\eta_\ell, \xi_\ell$  and  $\eta_r$  from Theorems 2.6 and 2.7, respectively, and  $\mathcal{V}_k^\ell$  from (2.7). Our main result on the BCM is as follows:

**Theorem 2.11** (The largest component of the BCM). *Consider  $\text{BCM}(\mathbf{d}^\ell, \mathbf{d}^r)$  under Assumption 2.3 (A,B,C1,D), and further assume that  $p_2 + q_2 < 2$ . Under the supercriticality condition (2.20), that we call the supercritical case of the BCM, we have that  $\xi_\ell > 0, \eta_\ell < 1$  and  $\eta_r < 1$ . Then, as  $n \rightarrow \infty$ ,*

$$\frac{|\mathcal{C}_{1,\delta} \cap \mathcal{V}^\ell|}{N_n} \xrightarrow{\mathbb{P}} \xi_\ell, \tag{2.32}$$

$$\frac{|\mathcal{C}_{1,\delta} \cap \mathcal{V}_k^\ell|}{N_n} \xrightarrow{\mathbb{P}} p_k (1 - \eta_\ell^k), \tag{2.33}$$

$$\frac{|\mathcal{E}(\mathcal{C}_{1,\delta})|}{N_n} \xrightarrow{\mathbb{P}} \mathbb{E}[D^\ell] (1 - \eta_\ell \eta_r). \tag{2.34}$$

In this case,  $\mathcal{C}_{1,\delta}$  is unique in the sense that  $|\mathcal{C}_{2,\delta}|/(N_n + M_n) \xrightarrow{\mathbb{P}} 0$ , and we refer to  $\mathcal{C}_{1,\delta}$  as the giant component of the BCM. When (2.20) does not hold,  $|\mathcal{C}_{1,\delta}|/(N_n + M_n) \xrightarrow{\mathbb{P}} 0$ .

We prove Theorem 2.11 in Section 3, and highlight the main ideas behind the proof shortly below. First, we provide some comments on, and corollaries of, Theorem 2.11. We note that while (2.34) looks ‘‘asymmetric,’’ since  $\mathcal{E}_n = N_n \mathbb{E}[D_n^\ell] = M_n \mathbb{E}[D_n^r]$  (see Remark 2.4(i)), we can rephrase it as  $|\mathcal{E}(\mathcal{C}_{1,\delta})|/M_n \xrightarrow{\mathbb{P}} \mathbb{E}[D^r] (1 - \eta_\ell \eta_r)$ , as well as  $|\mathcal{E}(\mathcal{C}_{1,\delta})|/\mathcal{E}_n \xrightarrow{\mathbb{P}} 1 - \eta_\ell \eta_r$ . In Section 2.4.2, we discuss why the condition  $p_2 + q_2 < 2$  is needed. Recall  $\mathcal{V}_k^r$  from (2.7) and  $\gamma$  from Remark 2.4(i).

**Corollary 2.12** (The rhs partition). *Under the conditions of Theorem 2.11 and the supercriticality condition (2.20), with  $\xi_r := 1 - G_{D^r}(\eta_r) \in [0, 1]$ , as  $n \rightarrow \infty$ ,*

$$\frac{|\mathcal{C}_{1,\delta} \cap \mathcal{V}^r|}{M_n} \xrightarrow{\mathbb{P}} \xi_r, \tag{2.35}$$

$$\frac{|\mathcal{C}_{1,\delta} \cap \mathcal{V}_k^r|}{M_n} \xrightarrow{\mathbb{P}} q_k (1 - \eta_r^k), \tag{2.36}$$

$$\frac{|\mathcal{E}_{1,\delta}|}{N_n + M_n} \xrightarrow{\mathbb{P}} \frac{\xi_\ell + \gamma \xi_r}{1 + \gamma}. \tag{2.37}$$

*Proof of Corollary 2.12.* Observe that the role of the lhs and rhs partitions, and in particular, the role of the quantities  $\eta_\ell$  and  $\eta_r$ , as well as that of  $\xi_\ell$  and  $\xi_r$  are symmetric; we formally establish this symmetry below in Section 5.1. Thus, by switching left and right, (2.35)–(2.36) follow from (2.32)–(2.33), and combining (2.32) and (2.35) yields (2.37). Thus, Corollary 2.12 follows from Theorem 2.11. ■

#### 2.4.1 | Overview of the proof of Theorem 2.11

The proof relies on a continuous-time exploration algorithm of the BCM. This algorithm is based on the continuous-time exploration algorithm for the (traditional, unipartite) configuration model proposed in [30]. However, the algorithm in [30] must be modified significantly, since without modification it only yields the average density in (2.37). We give a brief explanation of our modified algorithm and highlight the challenges in its analysis here; the details are provided in Section 3.

The new algorithm builds and explores the graph simultaneously and unveils the connected components one by one. We start the exploration of each component by picking an unexplored  $\ell$ -vertex. During the exploration of a component, each *round* of the algorithm is a *double-step*, described as follows. We first match a free half-edge incident to an  $\ell$ -vertex  $v$  in the component that we are currently exploring, and reach an  $r$ -vertex neighbor  $a$ . We then match all remaining half-edges of  $a$ , using it as a bridge to reach second neighbors of  $v$  that are again  $\ell$ -vertices. This constitutes one round (or double-step). By the end of each round, all unmatched half-edges belonging to the component being explored are incident to  $\ell$ -vertices. Thus, the component is fully explored when a round is completed and there are no more unmatched  $\ell$ -half-edges.

The asymmetric roles of the left and right partitions are necessary for obtaining the more refined results on the size of the giant within each partition. However, it leads to a more complex analysis, as the  $r$ -vertex to explore in each round is chosen *randomly*. (In particular, we find this vertex by matching the chosen  $\ell$ -half-edge to a uniform unmatched  $r$ -half-edge, thus we choose the  $r$ -vertex in a size-biased fashion.) Consequently the degree of the  $r$ -vertex is also random, and by matching all its remaining half-edges, we create a random number of edges in the BCM in one round. In contrast, in the original algorithm exactly one edge is created in each round.

Studying our modified algorithm, only a small portion of the analysis remains the same as for the original algorithm in [30]. The novelty and mathematical challenge lies in studying the evolution of the number of unmatched  $\ell$ -half-edges. It evolves as a pure death process, where most jumps happen with rate  $i$  from state  $i$ , however some randomly chosen jumps happen *instantaneously* (with rate infinity). The reason for this is exactly the random number of edges created in one round of the algorithm: one instantaneous jump happens in each round due to adding the new  $r$ -vertex, and the “regular” jumps happen when we match the rest of the half-edges of this  $r$ -vertex. See Section 3.1 for more explanation. To analyze this death process with two types of jumps, we compare it to a well-understood “standard” death process, where there are no instantaneous jumps. The comparison is carried out through hitting times, allowing us to think of the effect of the instantaneous jumps as the *time saved*. We are able to analyze the *time saved* in an elegant way by giving it a new probabilistic interpretation in terms of a size-biased reordering of  $r$ -degrees.

#### 2.4.2 | Discussion and open problems

For a discussion on the RIGC model (about its applicability, overlapping structure, and simplicity), see the companion paper [48, Section 2.4]. In this section, we discuss relations of our results to the literature, and provide a discussion on the extra condition  $p_2 + q_2 < 2$  and the use of the BCM to generate simple bipartite graphs with a given degree sequence.

### Relation to the RIG and BCM literature

In the RIG setting, there are several papers that identify the size of the giant component for various degree distributions as in Theorem 2.6, usually arising from RIGs with (conditionally) independent edges. Newman [40] establishes Theorem 2.6 for the RIG in a physics-style way. The work closest to ours is of Ball, Sirl, and Trapman [5], which also studies the alternating branching process limit that underlies the proof of Theorem 2.11 (see in particular [5, Section 3.2.1]). Ball, Sirl, and Trapman [5] focus on epidemics on such random graphs, see Section 2.5 for related results on *percolation* on RIGs. Ball, Sirl, and Trapman [5, Theorem 3.5] establish a limiting law for the proportion of vertices in the final epidemic, which can be 0 or the survival probability of the corresponding backwards branching process. When performing this for an SI epidemic, this corresponds to a law of large numbers for the component of a uniform individual, which, conditionally on survival, equals the law of large numbers for the giant as in Theorem 2.6. Lagerås and Lindholm [35] treat the case where both  $D^r$  and  $D^\ell$  are Poisson distributed, Bloznelis [8] studies the case where  $D^r$  is Poisson, and  $D^\ell$  is a Poisson mixture with an arbitrary mixing distribution, while [7] covers the case where  $D^r$  is Poisson mixture with a rather general mixing distribution, and  $D^\ell$  is Poisson. Thus, applied to the RIG, Theorem 2.6 nicely complements these results for the RIG to the setting of *fully general* degree distributions. Coupechoux and Lelarge [19], in turn, study an alternating branching process that approximates the local neighborhoods in the BCM. This is also the intuition behind Theorem 2.11, even though our proof is organized rather differently.

### The condition $p_2 + q_2 < 2$

We briefly explain why the almost-2-regular graph, arising when  $p_2 + q_2 = 2$ , is excluded. First, we show that the CM can be obtained from the BCM as a special case, then we recall from the literature why the general results are not applicable for the almost-2-regular case of the CM.

Assume that  $\mathcal{V}_2^r = \mathcal{V}^r$  for all  $n$ , that is, all  $r$ -vertices have degree 2. Then each  $r$ -vertex  $a$  only serves as connecting two  $\ell$ -vertices, say,  $v$  and  $w$  through the 2-length path  $(v, a, w)$ . We can construct a unipartite graph on  $\mathcal{V}^\ell$  by contracting each of these  $(v, a, w)$  paths into an edge  $(v, w)$ . We show that this unipartite graph has the distribution of the configuration model  $\text{CM}_n(d^\ell)$ . For each (unipartite) matching  $\widehat{\omega}_n = \{(v, i_1), (w, i_2)\}$  of the  $\ell$ -half-edges corresponding to a unipartite graph, there are exactly  $|\mathcal{V}^r|! 2^{|\mathcal{V}^r|}$  bipartite matchings  $\omega_n$  that are mapped into  $\widehat{\omega}_n$  by the above contraction. The reason is that we can permute all  $r$ -vertices, as well as each pair of  $r$ -half-edges attached to the same  $r$ -vertex. Since  $\omega_n$  is a uniform bipartite matching, necessarily  $\widehat{\omega}_n$  is a uniform (unipartite) matching.

In [30], the  $p_2 = 1$  case of the CM (where most vertices have degree 2) is excluded for the reason that the size of the giant component is *not concentrated*: it shows diverse behavior depending on the more refined asymptotics of the degree structure. In particular, if there are only degree-2 vertices, then the density of the largest component converges to a *non-degenerate* distribution, rather than a constant. However, for a CM with a small proportion of degree-1 vertices, aside from the  $n - o_{\mathbb{P}}(n)$  vertices of degree 2, the size of the giant component becomes sublinear in  $n$ . In contrast, when almost all vertices have degree 2 and a small proportion has degree 4, the giant component constitutes almost all vertices. For a more detailed discussion see [30, 46].

By the contraction described above, the  $p_2 + q_2 = 2$  case of the BCM includes the ambiguous  $p_2 = 1$  case of the CM. In particular, when  $\mathcal{V}_2^r = \mathcal{V}^r$  for all  $n$  and  $p_2 = 1$ , the BCM is equivalent to the CM with  $p_2 = 1$ . This shows that not only the proof fails for this case, but Theorem 2.11 itself does not always hold.

**Uniform simple bipartite graphs with given degrees**

It is well known that the (traditional, unipartite) configuration model (CM) conditioned on being simple is a uniform simple graph with the given degree sequence. Not surprisingly, the corresponding statement is also true for the BCM. We provide a brief justification below. Let  $G$  be an arbitrary bipartite multigraph with  $\ell$ -degree and  $r$ -degree sequences  $d^\ell$  and  $d^r$ , and for  $v \in \mathcal{V}^\ell, a \in \mathcal{V}^r$ , let  $x_{\beta}(v, a)$  denote the multiplicity of the edge  $(v, a)$  in  $G$ . Then the number of (bipartite) matchings  $\omega_n$  that realize  $G$  is

$$\frac{\prod_{v \in \mathcal{V}^\ell} d_v^\ell! \prod_{a \in \mathcal{V}^r} d_a^r!}{\prod_{v \in \mathcal{V}^\ell, a \in \mathcal{V}^r} x_{\beta}(v, a)!} \tag{2.38}$$

We justify the formula, as follows. The numerator arises since all half-edges attached to the same vertex are equivalent, hence permuting them leads to the same graph, but a different matching. The denominator in turn arises since all instances of a multi-edge are equivalent, and by permuting both  $\ell$ - and  $r$ -half-edges, the same set of pairs appears in all possible orderings. Then all simple bipartite graphs, that is, where all the multiplicities  $x_{\beta}(v, a)$  are 0 or 1, arise from  $\prod_{v \in \mathcal{V}^\ell} d_v^\ell! \prod_{a \in \mathcal{V}^r} d_a^r!$  matchings and thus have the same probability. Thus conditioning the BCM on being simple indeed leads to a uniform simple bipartite graph with the given  $\ell$ - and  $r$ -degree sequences.

Note that the probability of obtaining a simple graph might tend to 0 as  $n \rightarrow \infty$ . Whether the asymptotic probability of obtaining a simple graph is positive is a non-trivial question and falls out of the scope of this article. Partial results are known, for example, the condition  $\mathbb{E}[(D^\ell)^2] < \infty, \mathbb{E}[(D^r)^2] < \infty$  guarantees a positive simplicity probability, as shown in [1].

We remark that using the above observed relation of the BCM to uniform random graphs with given degree sequences, our results can be extended beyond the scope of the BCM. It is known that the generalized random graph (GRG) conditioned on its degree sequence yields a uniform random graph with those degrees [15]. One can define a bipartite version of the model, with lhs partition  $[N_n]$  and rhs partition  $[M_n]$  with weight sequences  $(w_i^\ell)_{i \in [N_n]}$  and  $(w_j^r)_{j \in [M_n]}$  such that  $\ell_n := \sum_{i \in [N_n]} w_i^\ell = \sum_{j \in [M_n]} w_j^r$ . Then the edge probability can be defined as  $p_{ij} = \frac{w_i^\ell w_j^r}{\ell_n + w_i^\ell w_j^r}$  for  $i \in [N_n], j \in [M_n]$ . A similar argument as in [15, Section 3] shows that conditionally on the lhs and rhs degree sequences, this model also yields a uniform bipartite graph with the given degree sequences. Consequently, the bipartite version of the GRG also undergoes a phase transition as in Theorem 2.11. We omit further details.

**Open problems and future research directions**

For such a young model as the RIGC, there are obviously plenty of open questions. It would be really interesting to fit the model to real-world network data to gain more insight into what type of network it is a good fit for, as well as study its performance for finite network sizes in comparison with the asymptotic theoretical results. Another exciting but challenging problem is to study graph distances: due to the community structures added, distances in the RIGC can be significantly different from the underlying BCM. The *homogeneous bond percolation* (retaining each edge independently with the same probability), that we study in the next section, also leaves open problems and plenty of room for generalizations. The question of *robustness*, formally defined in Section 2.5, informally speaking the ability of the network to withstand random attacks, is explored further in a manuscript in preparation [33]. One can consider *inhomogeneous* percolation (with different retention probabilities), for example make the retention probability dependent on the degrees of the endpoints or the community graph the edge is part of. (The methods we present in Section 4 would work for the latter case, however not the former.) Another common generalization is site percolation, that is, percolating *vertices* rather than edges, or even combining the two approaches.



## 2.5 | Results on percolation on the random intersection graph with communities

In this section, we introduce the percolation model and state our results on percolation on the random intersection graph with communities.

### 2.5.1 | Introduction to percolation

In this section, we motivate and introduce the percolation model, and prove that percolation on the RIGC exhibits a phase transition (to be defined later) as we vary the percolation parameter.

*Percolation* [13, 26] is a probabilistic model introduced in [16] to study a group of physical phenomena related to a “fluid” spreading through a “porous medium” in a unified, abstract way. Examples and motivations given in [16] include absorption of gas or liquid into a porous rock and spreading of a disease through a social network. Percolation processes differ from diffusion processes in that a diffusion process is largely determined by properties of the fluid, while in the case of percolation, the spreading behavior is largely determined by properties of the *medium*. In the mathematical model of percolation, we represent the “porous medium” by a graph and define a random environment where edges (bond percolation) or vertices (site percolation) of this graph are randomly removed. The “fluid” can then spread through all retained edges (resp., vertices). Many variations of the model exist, but here we focus on bond percolation and the Bernoulli case: each edge is removed with the same probability, independently of each other.

The notion of *phase transition* also has its roots in physics and refers to the phenomenon when a model shows significantly different behavior depending on a specific parameter. The most common example is the different states of matter, sometimes referred to as *phases*, that the same material assumes at different temperatures. The parameter value (or interval, in the case of finite systems) where the behavior change occurs is referred to as the *critical point* (or critical window for large finite systems).

Percolation was extensively studied first on infinite (deterministic) lattices, where the phase transition is characterized by the presence or absence of an infinite connected component in the percolated graph, that is, after the removal of edges. It is straightforward to apply the percolation model for finite as well as random graphs, but less straightforward to define a phase transition. Phase transition on finite graphs is commonly reinterpreted in the large graph limit, as whether or not a linear proportion of the graph is connected after percolation.

We are motivated to study percolation on the RIGC model by possible applications in epidemiology and large-scale randomized attacks on the network. The correspondence between random removal of edges and a randomized attack on the network is quite intuitive. For a virus spread, whether a computer or biological virus, the percolation model is able to capture the final infected cluster of an information cascade [25, 29] or an SIR-epidemic [36, 37] (more specifically, a Reed–Frost epidemic) as defined below.

In the Reed–Frost SIR-epidemic, individuals have three possible states: susceptible, infected, and recovered. Initially, all individuals are susceptible, and at time 0, we infect a single individual, the source. In each (discrete) time step, all the individuals that became infected in the previous step attempt to transmit the infection through all incident edges, and afterwards become recovered. Each transmission succeeds with probability  $\pi$ , independently of each other. If a successful transmission is made to a susceptible neighbor, then it becomes infected. This continues until there is a time when no new individual becomes infected, and then the process stops. In a system of size  $N$ , the process is terminated at the latest by time  $N - 1$ . It is easy to see that the set of infected individuals has the same distribution as the individuals in the percolated component of the source.

Formally, we define (bond) percolation on the RIGC as follows. Let  $\pi \in [0, 1]$  be a parameter called the edge *retention* probability. Given a realization of the RIGC, we retain each edge, independently of each other, with probability  $\pi$ , and otherwise delete it. We call the remaining subgraph (with two layers of randomness) the percolated RIGC and denote it by  $\text{RIGC}(\pi)$ . Note that  $\text{RIGC}(1) = \text{RIGC}$ , and  $\text{RIGC}(0)$  is the empty graph.

2.5.2 | Phase transition of bond percolation

Recall  $D^\ell$  and  $D^r$  from Assumption 2.3(A) and (C1), respectively. Recall  $\xi_\ell$  from Theorem 2.6 and  $N_n = |\mathcal{V}^\ell|$ . Denote the largest connected component (the component containing the most vertices, with ties broken arbitrarily) of  $\text{RIGC}(\pi)$  by  $\mathcal{C}_1(\pi) = \mathcal{C}_1^{(n)}(\pi)$ , and the second largest by  $\mathcal{C}_2(\pi) = \mathcal{C}_2^{(n)}(\pi)$ .

**Theorem 2.13** (Percolation phase transition on the RIGC). *Consider (bond) percolation with edge retention probability  $\pi \in [0, 1]$  on  $\text{RIGC}(d^\ell, \mathbf{Com})$  under Assumption 2.3 and the supercriticality condition (2.20). Then there exists a function  $\pi \mapsto \xi_\ell(\pi)$  and a threshold  $\pi_c \in [0, 1]$  such that  $|\mathcal{C}_1(\pi)|/N_n \xrightarrow{\mathbb{P}} \xi_\ell(\pi)$  and*

- (i) if  $\pi < \pi_c$ , then  $\xi_\ell(\pi) = 0$ ;
- (ii) if  $\pi > \pi_c$ , then  $\xi_\ell(\pi) \in (0, \xi_\ell]$  and  $\mathcal{C}_1(\pi)$  is whp unique:  $|\mathcal{C}_2(\pi)|/N_n \xrightarrow{\mathbb{P}} 0$ .

We prove Theorem 2.13 as a consequence of Theorem 2.6 in Section 4. We refer to the behavior in case (i) as subcritical percolation and in case (ii) as supercritical percolation. We assume the supercriticality condition (2.20) since when this condition fails, case (ii) becomes impossible, thus there is no phase transition. In the following, we characterize the threshold  $\pi_c$ . Recall (1.1) and Assumption 2.3(A).

**Proposition 2.14** (Characterization of the threshold  $\pi_c$ ). *Let  $\mathcal{H}$  denote a random graph with pmf  $\mu$  and let  $U_{\mathcal{H}}|\mathcal{H} \sim \text{Unif}[\mathcal{V}(\mathcal{H})]$ . Let  $\mathcal{C}^{\mathcal{H}}(U_{\mathcal{H}}, \pi)$  denote the percolated component of  $U_{\mathcal{H}}$  within  $\mathcal{H}$  with edge retention probability  $\pi$ . The threshold of the edge retention probability in Theorem 2.13 is given by*

$$\pi_c = \inf \left\{ \pi : \mathbb{E}[\tilde{D}^\ell] \cdot \mathbb{E} [|\mathcal{H}| (|\mathcal{C}^{\mathcal{H}}(U_{\mathcal{H}}, \pi)| - 1)] / \mathbb{E} [|\mathcal{H}|] > 1 \right\}, \tag{2.39}$$

where  $\mathbb{E}[\cdot]$  denotes total expectation (with respect to all sources of randomness). Furthermore,  $\pi_c < 1$ .

We prove Proposition 2.14 in Section 4.4. As discussed in Section 2.4.2, related results for epidemics on RIGs with conditionally independent edges were proved in [5]. The paper [5] also covers percolation on such RIGs by setting the infectious period constant, see [5, Section 3.2.1] for more details on this case. We remark that  $\pi_c < 1$  ensures that the set of supercritical percolation parameters is always non-empty. However, the set of subcritical parameters  $\pi < \pi_c$  may be empty. The phenomenon when  $\pi_c = 0$  is called *robustness*, and we explore it further in a manuscript in preparation [33].

3 | THE GIANT COMPONENT OF THE RIGC AND THE BCM

In this section, we prove Theorem 2.6 on the phase transition of the RIGC as a corollary of Theorem 2.11 on the phase transition of the BCM, and prove Theorem 2.11 itself. Note that, while the

connected components in the RIGC are closely related to those in the BCM, ordering the connected components of the underlying BCM by size generally *does not* ensure that the corresponding connected components of the RIGC are also ordered by size, and this needs proper attention. The proof of Theorem 2.11 makes use of the continuous-time exploration algorithm sketched in Section 2.4, that we describe in more detail later in this section, then analyze it.

*Proof of Theorem 2.6 subject to Theorem 2.11.* For some  $v \in \mathcal{V}^\ell$ , let us denote its connected component in the RIGC by  $\mathcal{C}^\rho(v)$ , and its connected component in the *underlying* BCM (see Remark 2.2) by  $\mathcal{C}^\beta(v)$ . Since every community graph is connected, two  $\ell$ -vertices are connected within the RIGC exactly when they are connected within the underlying BCM. Consequently,  $\mathcal{C}^\rho(v) = \mathcal{V}^\ell \cap \mathcal{C}^\beta(v)$ , and each connected component of the RIGC is exactly the set of  $\ell$ -vertices in the corresponding connected component of the underlying BCM. In the subcritical and critical case,  $|\mathcal{C}_{1,\beta}^{(n)}| = o_{\mathbb{P}}(N_n)$  by Theorem 2.11 (recall that  $M_n = \gamma N_n + o(N_n)$  by Remark 2.4(i)). Since  $|\mathcal{C}^\rho(v)| \leq |\mathcal{C}^\beta(v)|$  for any  $v \in \mathcal{V}^\ell$ , we conclude that

$$|\mathcal{C}_1| = \max_{v \in \mathcal{V}^\ell} |\mathcal{C}^\rho(v)| \leq \max_{v \in \mathcal{V}^\ell} |\mathcal{C}^\beta(v)| = |\mathcal{C}_{1,\beta}| = o_{\mathbb{P}}(N_n). \tag{3.1}$$

Under the supercriticality condition (2.20),  $|\mathcal{C}_{1,\beta} \cap \mathcal{V}^\ell| / N_n \xrightarrow{\mathbb{P}} \xi_\ell$  by Theorem 2.11, and for *any other* component  $\mathcal{C}'$  of the BCM,  $|\mathcal{C}' \cap \mathcal{V}^\ell| \leq |\mathcal{C}'| \leq |\mathcal{C}_{2,\beta}| = o_{\mathbb{P}}(N_n)$ . Thus necessarily,

$$\mathcal{C}_1 = \mathcal{C}_{1,\beta} \cap \mathcal{V}^\ell \text{ whp,} \tag{3.2}$$

which implies that  $|\mathcal{C}_1| / N_n \xrightarrow{\mathbb{P}} \xi_\ell$ , and analogously to (3.1),  $|\mathcal{C}_2| = o_{\mathbb{P}}(N_n)$ . This concludes the proof of Theorem 2.6 subject to Theorem 2.11.

### 3.1 | Global exploration

We prove our results regarding the giant component of the BCM with the aid of the exploration algorithm sketched in Section 2.4. The algorithm is an adaptation of the exploration algorithm of the CM proposed by Janson and Luczak [30], however the analysis poses new challenges. Below, we introduce the required terminology and notation, and formalize the algorithm in the form of pseudo-code.

We call two (or more) half-edges *siblings* (a family of half-edges) if they are incident to the same vertex. To keep notation simple, we do not always explicitly indicate the dependence on  $n$ , however it is always meant. Instead, we add the superscripts  $\ell$  or  $r$  to emphasize which partition each quantity is related to. We define the algorithm focusing on the lhs partition to obtain the statements in Theorem 2.11. We could analogously define and analyze the algorithm focusing on the rhs partition and obtain the statements in Corollary 2.12 instead. Note that the number of paired half-edges in the two partitions must always be equal. All the quantities below are defined to be right-continuous, that is, if the algorithm updates a quantity at time  $t$ , then the value *at* time  $t$  is the updated value.

At any given time,  $\mathcal{V}^\ell$  is partitioned into the time-dependent set of *sleeping* and *awake* vertices. Initially, all  $\ell$ -vertices are sleeping, then they are later moved one by one to the awake set and never return to sleeping. Intuitively, an awake vertex is at least partially explored. We denote the number of sleeping  $\ell$ -vertices of degree  $k$  at time  $t$  by  $\mathcal{V}_k^\ell(t)$ . Similarly,  $\mathcal{V}^r$  is partitioned into the sleeping set and awake set, and each  $r$ -vertex starts in the sleeping set and later progresses into the awake set.

The set of  $\ell$ -half-edges, at any given time, is partitioned as follows: the sleeping set of size  $S^\ell(t)$ , the active set of size  $A^\ell(t)$  and the paired (dead) set. Intuitively, active half-edges are those half-edges that we already know to belong to the component we are currently exploring and are still unpaired. We can thus use them to progress the exploration. Note that

$$S^\ell(t) = \sum_{k=1}^{\infty} k \mathcal{V}_k^\ell(t). \quad (3.3)$$

Each  $\ell$ -half-edge progresses from sleeping to active to paired, or directly from sleeping to paired. Sometimes we say a half-edge “dies” to mean that we pair it or that we must pair it immediately. We thus refer to the union of the sleeping and active sets of half-edges as the living (unmatched) set, which has size  $L^\ell(t) = A^\ell(t) + S^\ell(t)$ . Further, we assign iid  $\text{Exp}(1)$  random variables to each  $\ell$ -half-edge, that we call the *alarm clocks* of the half-edges. Once the exploration time reaches the value of this variable, the alarm goes off, and if the half-edge is still unpaired, then it dies and must be paired immediately. When an  $\ell$ -half-edge dies, if the incident  $\ell$ -vertex is sleeping, then we set it awake, and set all sibling half-edges active. (If the incident  $\ell$ -vertex is already awake, then we do not change the status of the vertex or the sibling half-edges.) When we set an  $\ell$ -vertex awake for a different reason, we set each incident half-edge active.

The  $r$ -half-edges are partitioned into the sleeping set, the *waiting-to-be-paired* set of size  $W^r(t)$ , and the paired (dead) set. Half-edges may progress from sleeping to paired directly, or through the waiting-to-be-paired status, but never move backwards. While the waiting-to-be-paired set on the rhs plays a role analogous to those of active half-edges on the lhs, we use a different notion to emphasize their different roles in the algorithm: while the set of active  $\ell$ -half-edges is allowed to grow large, the waiting-to-be-paired set must always be exhausted immediately. When an  $r$ -half-edge is paired, the incident  $r$ -vertex is set to awake, and all sibling half-edges are set to be waiting-to-be-paired. (By the design of the algorithm, this is the only way to set an  $r$ -vertex awake.)

The unit of the algorithm we often focus on is one *iteration* of the outer *while* loop, that is, the conditional execution of `step1`, the execution of `step2` and the internal *while* loop of `step3s`, which corresponds to discovering an  $r$ -vertex and matching all its remaining half-edges. By construction, the lists  $\mathcal{S}_1$  and  $\mathcal{S}_2$  contain the time stamps of all executions of `step1` and `step2`, respectively. Noting that in each iteration, `step2` is executed once while `step1` is executed once only if the condition is satisfied and is otherwise not executed,  $\mathcal{S}_1$  must be a sublist of  $\mathcal{S}_2$ . We also remark that both  $\mathcal{S}_1$  and  $\mathcal{S}_2$  may contain duplicates of the same time stamp, as the time variable is only increased in `step3`, which is not executed in those iterations when the condition of the internal *while* loop fails, that is, when the  $r$ -vertex found has degree one, so that the chosen  $r$ -half-edge does not have any sibling half-edges.

*Remark 3.2* (Original algorithm as special case). In Section 2.4.2, we have shown that when each  $r$ -vertex has degree 2, the bipartite configuration model  $\text{BCM}_n(\mathbf{d}^\ell, \mathbf{d}^r)$  is equivalent to  $\text{CM}_n(\mathbf{d}^\ell)$ . In this case, `step3` is executed exactly once in each iteration, and our algorithm gives back the exploration for the CM in [30].

### 3.2 | Analysis of the exploration algorithm

In this section, we study Algorithm 3.1. The results obtained serve as ingredients to the proof of Theorem 2.11 in Section 3.4. Recall that we begin the exploration of a new component exactly when

**Algorithm 1.** Continuous-time exploration of the BCM

Initially,  $t = 0$ , all vertices and half-edges are sleeping, and all half-edges are unpaired. Let  $\mathcal{S}_1$  and  $\mathcal{S}_2$  be initially empty lists.

**while** there are unpaired  $\ell$ -half-edges ( $L^\ell(t) > 0$ ) **do**

**if** there are no active  $\ell$ -half-edges ( $A^\ell(t) = 0$ ) **then**

    step1: Starting the exploration of a new component.

      pick a *sleeping*  $\ell$ -half-edge  $x$  uar.

      set the incident  $\ell$ -vertex  $v$  as *awake*. ( $\mathcal{V}_{\ell\text{-deg}(v)}^\ell(t) := \mathcal{V}_{\ell\text{-deg}(v)}^\ell(t) - 1$ )

      set  $x$  and all sibling half-edges as *active*. ( $A^\ell(t) := \ell\text{-deg}(v)$ )

      append  $t$  to the list  $\mathcal{S}_1$ . ( $t$  remains unchanged.)

**end** step1

**end if**

  step2: Discovering a new  $r$ -vertex.

    pick an *active*  $\ell$ -half-edge  $x$  arbitrarily and a *sleeping*  $r$ -half-edge  $y$  uar.

    match  $x$  and  $y$  to form an edge and set both as *paired*. ( $A^\ell(t) := A^\ell(t) - 1$ )

    set the  $r$ -vertex  $a$  incident to  $y$  as *awake*.

    set sibling half-edges of  $y$  as *waiting-to-be-paired*. ( $W^r(t) := r\text{-deg}(a) - 1$ )

    append  $t$  to the list  $\mathcal{S}_2$ . ( $t$  remains unchanged.)

**end** step2

**while** there are waiting-to-be-paired  $r$ -half-edges ( $W^r(t) > 0$ ) **do**

    step3: Exploring further connections of the  $r$ -vertex.

      pick a *waiting-to-be-paired*  $r$ -half-edge  $y$  arbitrarily.

      wait  $dt$  time until the first alarm clock of an *unpaired*  $\ell$ -half-edge  $x$  rings.

**if**  $x$  is sleeping **then**

      set the  $\ell$ -vertex  $v$  incident to  $x$  as *awake*. ( $\mathcal{V}_{\ell\text{-deg}(v)}^\ell(t + dt) := \mathcal{V}_{\ell\text{-deg}(v)}^\ell(t) - 1$ )

      set sibling half-edges of  $x$  as *active*. ( $A^\ell(t + dt) := A^\ell(t) + \ell\text{-deg}(v) - 1$ )

**end if**

    match  $x$  and  $y$  to form an edge and set both as *paired*. ( $L^\ell(t + dt) := L^\ell(t) - 1$ ,  $W^r(t + dt) := W^r(t) - 1$ )

    increase time  $t := t + dt$ .]

**end** step3

**end while**

**end while**

step1 is executed, for which  $A^\ell(t) = 0$  is a *necessary* condition. Thus our aim is to understand the behavior of  $t \mapsto A^\ell(t)$  during the course of the exploration, in particular, to determine the zeros of this function. Our analysis, as in [30], is based on the simple observation that  $A^\ell(t) = L^\ell(t) - S^\ell(t)$ . We move on to studying the quantities  $S^\ell(t)$  and  $L^\ell(t)$  separately.

The dynamics of  $S^\ell(t)$ , similarly to the corresponding quantity in the algorithm in [30], are the following. Note that step2 does not affect  $S^\ell(t)$ . Regularly,  $\ell$ -half-edges are removed from the sleeping set when the alarm clock of the half-edge itself or one of its siblings rings, due to step3. However, sometimes a family of  $\ell$ -half-edges is removed from the sleeping set due to step1, when we start the exploration of a new component by picking a uniform  $\ell$ -half-edge and set it active together with its siblings, and set the incident  $\ell$ -vertex awake. Let  $\hat{\mathcal{V}}_k^\ell(t)$  denote the number of  $\ell$ -vertices of degree

$k$  such that the alarm clocks of all  $\ell$ -half-edges show a time greater than  $t$ , and define

$$\widehat{S}^\ell(t) := \sum_{k=1}^{\infty} k \widehat{V}_k^\ell(t). \tag{3.4}$$

Comparing with (3.3), we intuitively think of  $\widehat{S}^\ell(t)$  as the number of sleeping  $\ell$ -half-edges ignoring the contribution of `step1`, and it serves as an approximation for  $S^\ell(t)$ . We recall the following result that holds unchanged for the bipartite case:

**Lemma 3.3** (Sleeping vertices and half-edges [30, Lemma 5.2]). *Define*

$$h_1(z) := \mathbb{E}[D^\ell] z G_{D^\ell}(z) = \sum_{k=1}^{\infty} k \mathbb{P}(D^\ell = k) z^k \tag{3.5}$$

for  $z \in [0, 1]$ . For any  $t_0$  fixed, as  $n \rightarrow \infty$ ,

$$\forall k \geq 1, \quad \sup_{t \leq t_0} \left| \frac{1}{N_n} \widehat{V}_k^\ell(t) - p_k e^{-kt} \right| \xrightarrow{\mathbb{P}} 0; \tag{3.6}$$

$$\sup_{t \leq t_0} \left| \frac{1}{N_n} \sum_{k=1}^{\infty} \widehat{V}_k^\ell(t) - G_{D^\ell}(e^{-t}) \right| \xrightarrow{\mathbb{P}} 0; \tag{3.7}$$

$$\sup_{t \leq t_0} \left| \frac{1}{N_n} \widehat{S}^\ell(t) - h_1(e^{-t}) \right| \xrightarrow{\mathbb{P}} 0. \tag{3.8}$$

We introduce

$$\widehat{A}^\ell(t) := L^\ell(t) - \widehat{S}^\ell(t), \tag{3.9}$$

that serves as our approximation for  $A^\ell(t)$ . The next lemma, recalled from [30], bounds the error that we make with this approximation:

**Lemma 3.4** (The effect of `step1` [30, Lemma 5.3]). *With  $d_{\max}^\ell$  from Remark 2.4(iii),*

$$0 \leq \widehat{S}^\ell(t) - S^\ell(t) < \sup_{s \leq t} \left( \widehat{S}^\ell(s) - L^\ell(s) \right) + d_{\max}^\ell. \tag{3.10}$$

*The above bound can be rewritten in the more convenient form*

$$0 \leq A^\ell(t) - \widehat{A}^\ell(t) = \widehat{S}^\ell(t) - S^\ell(t) < -\inf_{s \leq t} \widehat{A}^\ell(s) + d_{\max}^\ell. \tag{3.11}$$

Recall the sequence  $\mathcal{S}_1$  from Algorithm 3.1 that contains the time stamps of all executions of `step1`, and that it may contain the same time stamp several times. Since this does not occur in the original algorithm in [30], we reprove the lemma to show that this does not cause an issue.

*Proof.* First, we study what happens at a time  $t \in \mathcal{S}_1 \subseteq \mathcal{S}_2$ . As explained after Algorithm 3.1,  $t$  may appear several times in the sequences  $\mathcal{S}_1 \subseteq \mathcal{S}_2$ , as finding a degree-1  $\nu$ -vertex uses up its single half-edge in `step2`, which results in not executing `step3` in that iteration and not increasing the time variable. However, the number of active  $\ell$ -half-edges changes with each execution of `step1` and

step2, thus we overwrite (redefine)  $A^\ell(t)$  each time until an  $r$ -vertex with degree at least 2 is found in step2. This step2 then necessarily corresponds to the last instance of  $t$  in  $\mathcal{S}_2$  and sets the final value of  $A^\ell(t)$ , as step3 must be executed next and the time variable will increase. Consider the last execution of step1 at  $t$ , which is either in the same or an earlier iteration than the last step2, and denote the  $\ell$ -vertex woken up by this step1 by  $v$ . As step2 is executed at least once afterwards, we have  $A^\ell(t) \leq \ell\text{-deg}(v) - 1 < d_{\max}^\ell$ . Recall that  $L^\ell(t) = A^\ell(t) + S^\ell(t)$ , hence

$$\widehat{S}^\ell(t) - S^\ell(t) = \widehat{S}^\ell(t) - L^\ell(t) + A^\ell(t) < \widehat{S}^\ell(t) - L^\ell(t) + d_{\max}^\ell. \tag{3.12}$$

By the definition of  $\widehat{S}^\ell(t)$ ,  $\widehat{S}^\ell(t) - S^\ell(t) \geq 0$  and the difference grows only due to step1, while it might decrease due to step2 or step3, for example, when a clock of an  $\ell$ -half-edge rings that was woken up in step1 previously. Hence for a time  $t' \notin \mathcal{S}_1$ ,  $\widehat{S}^\ell(t') - S^\ell(t') \leq \widehat{S}^\ell(s) - S^\ell(s)$ , where  $s := \max\{t \in \mathcal{S}_1 : t < t'\}$ . Then, using (3.12) and that the supremum is actually a finite maximum over a subset of  $\mathcal{S}_1$ ,

$$\widehat{S}^\ell(t') - S^\ell(t') \leq \sup_{s \leq t'} \left( \widehat{S}^\ell(s) - S^\ell(s) \right) < \sup_{s \leq t'} \left( \widehat{S}^\ell(s) - L^\ell(s) \right) + d_{\max}^\ell, \tag{3.13}$$

which concludes the proof. ■

Next, we state our novel result on the process of living (unmatched) half-edges  $L^\ell(t)$ . As remarked in the sketch of the proof of Theorem 2.11 in Section 2.4, the dynamics of this process differ significantly from the corresponding process in [30]. The analysis of the new process, carried out in Section 3.3 below, is our major innovation to generalizing the algorithm to the bipartite case.

We introduce some notation necessary to state our result. For an arbitrary invertible function  $f$ , let  $f^{(-1)}$  denote the inverse function of  $f$ , that is,  $f^{(-1)}(f(z)) = z$  for any  $z$  in the domain of  $f$  and  $f(f^{(-1)}(z)) = z$  for any  $z$  in the domain of  $f^{(-1)}$  (i.e., the range of  $f$ ). Recall (1.1), (1.2), as well as  $D^\ell$  and  $D^r$  from Assumption 2.3(A) and (C1). Since the generating function  $G_X$  of a random variable  $X$  taking values from  $\mathbb{N}$  (such that  $\mathbb{P}(X = 0) < 1$ ) is continuous and strictly increasing,  $G_X^{(-1)}$  exists on the interval  $[\mathbb{P}(X = 0), 1]$ .

**Proposition 3.5** (Living half-edges). *Define the function*

$$h_2(z) := \mathbb{E}[D^\ell] z G_{\widetilde{D}^r}^{(-1)}(z) \tag{3.14}$$

on  $[\widetilde{q}_0, 1]$ , where  $\widetilde{q}_0 := \mathbb{P}(\widetilde{D}^r = 0) = q_1 / \mathbb{E}[D^r]$ . The process of living half-edges  $L^\ell(t)$  satisfies, for any  $0 < t_0 < -\log \widetilde{q}_0$ ,

$$\sup_{t \leq t_0} \left| \frac{1}{N_n} L^\ell(t) - h_2(e^{-t}) \right| \xrightarrow{\mathbb{P}} 0. \tag{3.15}$$

We prove Proposition 3.5 in Section 3.3. We remark that for the RIGC and its underlying BCM, postulating  $q_1 = 0$  implies  $\widetilde{q}_0 = 0$  and  $-\log \widetilde{q}_0 = \infty$ . It is hard to intuitively interpret the appearance of an *inverse* generating function in (3.14)–(3.15). The deeper analysis of the process  $L^\ell(t)$  in Section 3.3 reveals that it is due to step2 happening instantaneously. The proof of Theorem 2.11 in Section 3.4 provides additional justification that the inverse must appear here in order to obtain (2.18), the fixed point equation for the *composition* of the generating functions, which we give an intuitive interpretation for later in Section 5.1.

### 3.3 | Living half-edges

In this section, we carry out the analysis of the process of living half-edges in Algorithm 3.1 and in particular, prove Proposition 3.5. We first introduce our approach and the ingredients of the proof and complete the proof before proving our lemmas.

#### 3.3.1 | Asymptotics for the living half-edges: Proof of Proposition 3.5

As remarked in the sketch of the proof in Section 2.4, the process of living half-edges is significantly more complex in the bipartite case, as it is a death process with *occasional* instantaneous jumps. In light of Algorithm 3.1, we can now explain precisely how this death process arises. Note that in each execution of `step2`, as well as in each execution of `step3`, one  $\ell$ -half-edge is paired, thus both steps correspond to a jump of size  $-1$ . As `step2` does not increase the time variable, it corresponds to an instantaneous jump (a jump with infinite rate). In `step3`, we wait for the first  $\text{Exp}(1)$  alarm clock (see Section 3.1) of an *unmatched*  $\ell$ -half-edge to ring, which corresponds to each living half-edge dying at rate 1. Thus in the death process, the regular jumps, corresponding to iterations of `step3`, happen with rate  $i$  from position  $i$ .

To determine how often the instantaneous jumps happen, consider that in `step2` of each iteration (of the outer *while* loop), we pick a sleeping  $\nu$ -vertex in a *size-biased* fashion, since a uniform unmatched  $\nu$ -half-edge is drawn. It is the remaining degree of the chosen  $\nu$ -vertex that determines the number of iterations of `step3` in the inner *while* loop. In contrast, in the original algorithm (see Remark 2.1) `step3` is executed exactly once in each iteration, thus the two jumps can be “merged” into a single jump of size  $-2$  with rate  $2i$  from position  $2i$ . Consequently, in the original algorithm it suffices to study a death process with only regular jumps (of size 2). However in our case, such a merging is not possible at all, due to the fact that  $\nu$ -vertices are used up in an order determined by a size-biased reordering (formally defined below), hence the degree distribution is continuously changing throughout the course of the algorithm. Thus, we take an alternative approach, instead using *hitting times*. Note that the initial value of the death process  $L^\ell(0) = L^{\ell,(n)}(0) = \ell_n$  is deterministic. For  $c \in [0, 1]$ , we define the hitting time process

$$\tau(c) := \min\{t : L^\ell(t) \leq c\ell_n\}. \quad (3.16)$$

The following claim ensures that studying the hitting times is essentially equivalent to studying the death process:

**Claim 3.6** (Concentration of a death process and its hitting times). For each  $n \in \mathbb{N}$ , let  $(X^{(n)}(t))_{t \geq 0}$  be a pure death process with deterministic initial condition  $a_n := X^{(n)}(0) \rightarrow \infty$  as  $n \rightarrow \infty$ . For  $c \in (0, 1]$ , let  $\mathcal{T}^{(n)}(c) := \min\{t : a_n^{-1}X^{(n)}(t) \leq c\}$  and let  $f : [0, \infty) \rightarrow (0, 1]$  be a strictly decreasing function such that  $f(0) = 1$  and both  $f$  and its inverse  $f^{(-1)}$  are continuous. Then the following two statements are equivalent:

- (i) for any  $t_0 < \infty$ ,  $\sup_{t \leq t_0} |a_n^{-1}X^{(n)}(t) - f(t)| \xrightarrow{\mathbb{P}} 0$ ,
- (ii) for any  $c_0 \in (0, 1)$ ,  $\sup_{c \geq c_0} |\mathcal{T}^{(n)}(c) - f^{(-1)}(c)| \xrightarrow{\mathbb{P}} 0$ .

We prove Claim 3.6 in Appendix B. Claim 3.6 is straightforwardly tailored to be applicable to  $L^\ell(t)$ . It is stated in slightly more generality to allow application for similar processes that we define shortly and are necessary for the analysis.



To understand  $L^\ell(t)$ , we compare it to a “standard” process  $L_{\text{stn}}^\ell(t)$ , defined as the pure death process where each individual dies independently with rate 1. That is, in the process  $L_{\text{stn}}^\ell(t)$  each jump happens with rate  $i$  from state  $L_{\text{stn}}^\ell(t) = i$ , and we set the same initial condition  $L_{\text{stn}}^\ell(0) := \ell_n$ . The processes  $L^\ell(t)$  and  $L_{\text{stn}}^\ell(t)$  can be coupled in an intuitive way by using the same realization of jumps, however  $L_{\text{stn}}^\ell(t)$  “forgets” about the occasional infinite rates; in other words, all jumps happen with rate  $i$  from position  $i$ . Due to its simpler dynamics, the behavior of  $L_{\text{stn}}^\ell(t)$  is well understood, and hence so is the behavior of its hitting times

$$\tau_{\text{stn}}(c) := \min\{t : L_{\text{stn}}^\ell(t) \leq c\ell_n\}. \tag{3.17}$$

However, in the process  $L^\ell(t)$ , the instantaneous jumps due to `step2save us time`, which gives rise to a crucial correction term. We define the saved time as

$$\tau_{\text{skip}}(c) := \tau_{\text{stn}}(c) - \tau(c) > 0, \tag{3.18}$$

with  $\tau(c)$  and  $\tau_{\text{stn}}(c)$  defined in (3.16) and (3.17). Recall (1.1), (1.2),  $D^\ell$  and  $D^r$  from Assumption 2.3(A) and (C1), and that  $f^{(-1)}$  denotes the inverse of a function  $f$ . We can summarize the asymptotics of  $\tau(c)$ ,  $\tau_{\text{stn}}(c)$ , and  $\tau_{\text{skip}}(c)$  in the following lemma:

**Lemma 3.7** (Concentration of  $\tau(c)$ ). *For any  $c_0 > 0$ , as  $n \rightarrow \infty$ ,*

$$\sup_{c \geq c_0} |\tau_{\text{stn}}(c) + \log(c)| \xrightarrow{\mathbb{P}} 0, \tag{3.19}$$

$$\sup_{c \geq c_0} \left| \tau_{\text{skip}}(c) + \log\left(G_{D^{r,*}}^{(-1)}(c)\right) \right| \xrightarrow{\mathbb{P}} 0. \tag{3.20}$$

Consequently, by (3.18),

$$\sup_{c \geq c_0} \left| \tau(c) + \log(c) - \log\left(G_{D^{r,*}}^{(-1)}(c)\right) \right| \xrightarrow{\mathbb{P}} 0. \tag{3.21}$$

We prove Lemma 3.7 in Section 3.3.2. We point out the appearance of the inverse generating function in the asymptotics of the time saved  $\tau_{\text{skip}}(c)$ , which is related to the size-biased reordering. However, also note that the inverse generating function is of the distribution  $D^{r,*} = \tilde{D}^r + 1$ . Next, we prove Proposition 3.5 subject to Claim 3.6 and Lemma 3.7.

*Proof of Proposition 3.5 subject to Claim 3.6 and Lemma 3.7.* By (3.21) in Lemma 3.7,  $\tau(c)$  concentrates around

$$f^{(-1)}(c) = -\log c + \log\left(G_{D^{r,*}}^{(-1)}(c)\right). \tag{3.22}$$

Thus, by Claim 3.6,  $L^\ell(t)/\ell_n$  concentrates around  $f$ . We claim that  $f$  can be expressed as

$$c = f(t) = e^{-t} G_{\tilde{D}^r}^{(-1)}(e^{-t}). \tag{3.23}$$

We show that the inverse of the above function  $f$  is indeed  $f^{(-1)}$  by rearranging for  $t$  in a clever way. Let  $s := G_{\tilde{D}^r}^{(-1)}(e^{-t})$ , then  $e^{-t} = G_{\tilde{D}^r}(s)$ , and  $c = G_{\tilde{D}^r}(s) \cdot s = G_{D^{r,*}}(s)$ , by (1.1) and (1.2) (see in particular the two identities below (1.2)). Hence

$$e^{-t} = G_{\tilde{D}^r}(s) = \frac{G_{D^{r,*}}(s)}{s} = \frac{c}{G_{D^{r,*}}^{(-1)}(c)}. \tag{3.24}$$

Applying the function  $-\log(\cdot)$  on both sides of (3.24), and noting that  $t = f^{(-1)}(c)$ , yields (3.22) as required, and we conclude that  $\sup_{t \leq t_0} |L^\ell(t)/\ell_n - f(t)| \xrightarrow{\mathbb{P}} 0$ . By (3.14),  $h_2(e^{-t}) = \mathbb{E}[D^\ell]f(t)$ , and  $\ell_n/N_n = \mathbb{E}[D_n^\ell] \rightarrow \mathbb{E}[D^\ell]$  by Remark 2.4(i) and Assumption 2.3(B), thus (3.15) follows. This concludes the proof of Proposition 3.5 subject to Claim 3.6 and Lemma 3.7.

3.3.2 | Concentration of the hitting times

This section is dedicated to proving Lemma 3.7. We prove (3.21), (3.19), and (3.20) in this order.

*Proof of (3.21), subject to (3.19) and (3.20).* Combining (3.19) and (3.20) through the triangle inequality yields that, for any  $\varepsilon > 0$  fixed,

$$\begin{aligned} & \mathbb{P}\left(\sup_{c \geq c_0} \left| \tau(c) + \log(c) - \log\left(G_{D^{r,*}}^{(-1)}(c)\right) \right| > \varepsilon\right) \\ & \leq \mathbb{P}\left(\sup_{c \geq c_0} \left| \tau_{\text{stn}}(c) + \log(c) \right| + \left| -\tau_{\text{skip}}(c) - \log\left(G_{D^{r,*}}^{(-1)}(c)\right) \right| > \varepsilon\right) \\ & \leq \mathbb{P}\left(\sup_{c \geq c_0} \left| \tau_{\text{stn}}(c) + \log(c) \right| > \varepsilon/2\right) + \mathbb{P}\left(\sup_{c \geq c_0} \left| \tau_{\text{skip}}(c) + \log\left(G_{D^{r,*}}^{(-1)}(c)\right) \right| > \varepsilon/2\right) \rightarrow 0 \end{aligned} \quad (3.25)$$

as  $n \rightarrow \infty$ . That is, by the definition of convergence in probability, (3.21) holds.

*Proof of (3.19).* Recall the “standard” pure death process  $L_{\text{stn}}^\ell(t)$  and its hitting times (3.19) from Section 3.3.1. Also recall that the process jumps from state  $i$  to state  $i - 1$  at rate  $i$ . Using that  $\text{Exp}(i) \stackrel{d}{=} \text{Exp}(1)/i$ ,

$$\tau_{\text{stn}}(c) = \tau_{\text{stn}}\left(\frac{\lfloor c\ell_n \rfloor}{\ell_n}\right) \stackrel{d}{=} \sum_{i=\lfloor c\ell_n \rfloor+1}^{\ell_n} \frac{E_i^{(n)}}{i}, \quad (3.26)$$

where for any fixed  $n$ ,  $(E_i^{(n)})_{i \in \mathbb{Z}^+}$  are independent  $\text{Exp}(1)$  random variables. For convenience, we define the index set

$$\mathcal{I}_c = \mathcal{I}_c^{(n)} := \{\lfloor c\ell_n \rfloor + 1 \leq i \leq \ell_n\}. \quad (3.27)$$

Then, for any  $c$  fixed, using (3.26) and recognizing the Riemann-approximation for the arising sum,

$$\mathbb{E}[\tau_{\text{stn}}(c)] = \sum_{i \in \mathcal{I}_c} \frac{1}{i} = \log(\ell_n) - \log(c\ell_n) + O(N_n^{-1}) = -\log(c) + O(N_n^{-1}), \quad (3.28)$$

and

$$\text{Var}(\tau_{\text{stn}}(c)) = \sum_{i \in \mathcal{I}_c} \frac{1}{i^2} < \sum_{i=\lfloor c\ell_n \rfloor+1}^{\infty} \frac{1}{i^2} \rightarrow 0 \quad (3.29)$$

as  $n \rightarrow \infty$ , since  $\ell_n \rightarrow \infty$  (and  $c > 0$ ). For  $s \geq 0$ , define the process

$$Y(s) = Y^{(n)}(s) := \tau_{\text{stn}}(e^{-s}) - \mathbb{E}[\tau_{\text{stn}}(e^{-s})] = \sum_{i \in \mathcal{I}_{\exp\{-s\}}} \frac{E_i^{(n)} - 1}{i}. \quad (3.30)$$

Note that  $Y$  is a zero-mean martingale and thus  $Y^2(s)$  is a non-negative submartingale. We apply Doob's martingale inequality and (3.29) to obtain, for any fixed  $\varepsilon > 0$  and  $c_0 > 0$  with  $s_0 := -\log(c_0) < \infty$ ,

$$\begin{aligned} \mathbb{P}\left(\sup_{c \geq c_0} \{\tau_{\text{stn}}(c) - \mathbb{E}[\tau_{\text{stn}}(c)]\}^2 \geq \varepsilon\right) &= \mathbb{P}\left(\sup_{s \leq s_0} Y^2(s) \geq \varepsilon\right) \\ &\leq \frac{\mathbb{E}[Y^2(s_0)]}{\varepsilon} = \frac{\text{Var}(Y(s_0))}{\varepsilon} = \frac{\text{Var}(\tau_{\text{stn}}(c_0))}{\varepsilon} \rightarrow 0 \end{aligned} \tag{3.31}$$

as  $n \rightarrow \infty$ . It follows that

$$\sup_{c \geq c_0} |\tau_{\text{stn}}(c) - \mathbb{E}[\tau_{\text{stn}}(c)]| \xrightarrow{\mathbb{P}} 0. \tag{3.32}$$

Consequently, by (3.28), we can bound

$$\sup_{c \geq c_0} |\tau_{\text{stn}}(c) + \log(c)| \leq \sup_{c \geq c_0} |\tau_{\text{stn}}(c) - \mathbb{E}[\tau_{\text{stn}}(c)]| + o_{\mathbb{P}}(1) + O(N_n^{-1}) \xrightarrow{\mathbb{P}} 0. \tag{3.33}$$

This concludes the proof of (3.19).

We remark that [30, Lemma 6.1] is applicable to  $L_{\text{stn}}^\ell(t)$ , which provides a shorter alternative proof for (3.19). However, we adopted the proof above to shed light on the decomposition (3.26), preparing for the proof of (3.20), which is much more interesting and insightful.

*Proof of (3.20).* Recall the definition of the process  $L_{\text{stn}}^\ell(t)$  and its hitting times  $\tau_{\text{stn}}(c)$  from Section 3.3.1. The decomposition in (3.26) is equivalent to

$$\tau_{\text{stn}}(c) \stackrel{d}{=} \sum_{i \in \mathcal{I}_c} \frac{E_i^{(n)}}{i}. \tag{3.34}$$

Next, we derive a similar decomposition for  $\tau(c)$ . Let  $\mathcal{J}_c \subseteq \mathcal{I}_c$  denote the set of such indices  $i \in \mathcal{I}_c$  that the jump from position  $i$  to position  $i - 1$  in the process  $L^\ell(t)$  happened instantaneously, that is, due to `step2`. (We provide a formal definition of the set  $\mathcal{J}_c$  later.) Clearly, since both processes are defined using the same realization of jumps, the difference in  $\tau(c)$  and  $\tau_{\text{stn}}(c)$  only arises due to the different jump rates from positions  $i \in \mathcal{J}_c$ . While rate  $i$  in  $L_{\text{stn}}^\ell(t)$  results in the term  $E_i^{(n)}/i$ , the instantaneous jump in  $L^\ell(t)$  results in a 0 term. That is, we can write

$$\tau(c) = \sum_{i \in \mathcal{I}_c \setminus \mathcal{J}_c} \frac{E_i^{(n)}}{i} + \sum_{i \in \mathcal{J}_c} 0 = \sum_{i \in \mathcal{I}_c \setminus \mathcal{J}_c} \frac{E_i^{(n)}}{i}, \tag{3.35}$$

and necessarily the saved time is

$$\tau_{\text{skip}}(c) = \tau_{\text{stn}}(c) - \tau(c) = \sum_{i \in \mathcal{J}_c} \frac{E_i^{(n)}}{i}. \tag{3.36}$$

We analyze  $\tau_{\text{skip}}(c)$  through the index set  $\mathcal{J}_c$ . Recall that we discover a new  $\nu$ -vertex exactly when `step2` is executed. This happens exactly when all half-edges of the previous  $\nu$ -vertex have been paired. (Here, we ignore the potential `step1` in between, as `step1` does not pair any half-edges

and consequently does not correspond to any jump in the process  $L^\ell(t)$ .) Cumulatively, we execute `step2` for the  $(j + 1)$ st time when all half-edges of the first  $j$   $r$ -vertices are paired. Let us denote the  $r$ -degree of the  $j$ th explored  $r$ -vertex by  $d_{\pi(j)}^r$ . Clearly,  $(d_{\pi(j)}^r)_{j \in [M_n]}$  is a random reordering of  $\mathbf{d}^r$ , or equivalently,  $(\pi(j))$  is a random permutation. We pick the next  $r$ -vertex to explore by choosing a uniform unpaired  $r$ -half-edge, thus  $r$ -vertices are always chosen in a *size-biased* fashion wrt their degrees. That is,  $r$ -vertices are explored in the order defined by a size-biased reordering. Define  $\mathcal{U}_j := \{\pi(1), \dots, \pi(j)\} \subset [M_n]$ , the random set of indices chosen (used) in the first  $j$  steps, then the distribution of  $\pi$  is given by

$$\mathbb{P}(\pi(j) = k \mid \mathcal{U}_{j-1}) = \begin{cases} 0 & \text{for } k \in \mathcal{U}_{j-1}, \\ \frac{d_k^r}{\sum_{i \in [M_n] \setminus \mathcal{U}_{j-1}} d_i^r} & \text{for } k \in [M_n] \setminus \mathcal{U}_{j-1}. \end{cases} \tag{3.37}$$

Denote the partial sums of the first  $j$   $r$ -degrees in this reordering by

$$\Sigma_j := \sum_{i=1}^j d_{\pi(i)}^r, \tag{3.38}$$

where the empty sum  $\Sigma_0 = 0$  by convention. Then  $\ell_n - \Sigma_j$  gives the state of  $L^\ell(t)$  after we finish exploring the  $j$ th  $r$ -vertex, thus `step2` must be executed again and from this position, an instantaneous jump happens. We can now give an alternative, formal definition of the index set

$$\mathcal{J}_c = \mathcal{J}_c^{(n)} = \{\ell_n - \Sigma_j, j \in [M_n]\} \cap \mathcal{I}_c. \tag{3.39}$$

Define

$$j_{\max}(c) := \max \{j : \ell_n - \Sigma_j > c\ell_n\}, \tag{3.40}$$

then we can rewrite (3.36) as

$$\tau_{\text{skip}}(c) = \sum_{j=0}^{j_{\max}(c)} \frac{E_j^{(m)}}{\ell_n - \Sigma_j}, \tag{3.41}$$

where the set  $(E_j^{(m)})_{j \in \mathbb{N}}$  is a (possibly reordered) subset of  $(E_i^{(n)})_{i \in \mathbb{Z}^+}$ , hence it is composed of iid  $\text{Exp}(1)$  random variables. We give a convenient alternative probabilistic interpretation to the decomposition in (3.41), allowing us to relate it to a process that we already understand.

We define a process  $Z^r(s) = Z^{r,(n)}(s)$  in continuous time  $s \geq 0$  on the  $r$ -half-edges, completely independent of the exploration algorithm. The process  $Z^r(s)$  follows dynamics analogous to  $\widehat{S}^\ell(t)$ , however we avoid the intuitive notion  $\widehat{S}^r(s)$  and use a separate time variable  $s$  rather than  $t$  to emphasize that  $Z^r(s)$  is *not* related to the exploration algorithm. We formally define  $Z^r(s)$  as follows. Initially, all  $r$ -vertices and  $r$ -half-edges are sleeping, and we assign independent  $\text{Exp}(1)$  alarm clocks to each  $r$ -half-edge. An  $r$ -vertex and *all* its half-edges are woken up (and never return to sleeping) when the alarm clock on *any* of the half-edges goes off. The process  $Z^r(s)$  keeps track of the number of sleeping  $r$ -half-edges. The hitting times of this process correspond to  $\tau_{\text{skip}}(c)$ , formally,

$$(\min \{s : Z^r(s) \leq c\ell_n\})_{1 \geq c > 0} \stackrel{d}{=} (\tau_{\text{skip}}(c))_{1 \geq c > 0}, \tag{3.42}$$

where the distributional equality is meant as stochastic processes. We prove (3.42) by induction on the number of awake  $\nu$ -vertices. Clearly,  $Z''(0) = \ell_n = \ell_n - \Sigma_0$ . Assume the number of sleeping  $\nu$ -half-edges to be  $Z''(s) = \ell_n - \Sigma_j$ . Since the alarm clocks of awake  $\nu$ -half-edges can be ignored, the time we have to wait for the next  $\nu$ -half-edge  $y$  to wake up has distribution  $E_j^{(m)}/(\ell_n - \Sigma_j)$ . The  $\nu$ -half-edge  $y$  is chosen uar among the sleeping ones, hence the incident  $\nu$ -vertex  $a$  is chosen in a size-biased fashion. That is,  $\nu\text{-deg}(a) = d_{\pi(j+1)}^\nu$ , where  $\pi$  is a random permutation with distribution (3.37). Also note that all  $\nu$ -half-edges incident to  $a$  are woken up at once, thus the change in  $Z''(s)$  is  $-\nu\text{-deg}(a)$ . Hence

$$\min \{s : Z''(s) = \ell_n - \Sigma_{j+1}\} - \min \{s : Z''(s) = \ell_n - \Sigma_j\} = \frac{E'_j}{\ell_n - \Sigma_j}, \tag{3.43}$$

where  $E'_j$  is an  $Exp(1)$  random variable, independent of everything else. Then by induction,

$$\min \{s : Z''(s) = \ell_n - \Sigma_{k+1}\} = \sum_{j=0}^k E'_j / (\ell_n - \Sigma_j). \tag{3.44}$$

To determine the hitting time  $\min\{s : Z''(s) \leq c\ell_n\}$ , we want the smallest  $k$  such that  $\ell_n - \Sigma_{k+1} \leq c\ell_n$ . Since  $Z''(s)$  is non-increasing, this is equivalent to finding the largest  $k$  such that  $\ell_n - \Sigma_k > c\ell_n$ , which is straightforwardly  $j_{\max}(c)$  by (3.40). Thus

$$\min\{s : Z''(s) \leq c\ell_n\} = \min \{s : Z''(s) = \ell_n - \Sigma_{j_{\max}(c)+1}\} = \sum_{j=0}^{j_{\max}(c)} \frac{E'_j}{\ell_n - \Sigma_j}, \tag{3.45}$$

where we recognize a decomposition analogous to that of  $\tau_{\text{skip}}(c)$  from (3.41), with  $(E'_j)_{j \in \mathbb{N}}$  rather than  $(E_j^{(m)})_{j \in \mathbb{N}}$ . However, as both random vectors consist of iid  $Exp(1)$  random variables, the two processes evolve in the exact same way and the distributional identity (3.42) follows.

Now all that is left is to determine the asymptotics of  $Z''(s)$  and apply Claim 3.6 to translate it into the asymptotics of  $\tau_{\text{skip}}(c)$ . As  $Z''(s)$  is defined analogously to  $\widehat{S}^\ell(t)$ , following the same dynamics on the opposite partition, we can use the results in Lemma 3.3 for  $Z''(s)$ , with the exchange of lhs and rhs quantities. Replacing the  $\ell$ -degree distribution by the  $\nu$ -degree distribution in (3.5) and (3.8) yields that for any  $s_0$  fixed, as  $M_n \rightarrow \infty$ ,

$$\sup_{s \leq s_0} \left| M_n^{-1} Z''(s) - \mathbb{E}[D^\nu] e^{-s} G_{D^\nu}(e^{-s}) \right| \xrightarrow{\mathbb{P}} 0. \tag{3.46}$$

Since  $z \cdot G_{D^\nu}(z) = G_{D^{\nu,*}}(z)$  by (1.1) and (1.2),  $\mathbb{E}[D_n^\nu] \rightarrow \mathbb{E}[D^\nu]$  by Assumption 2.3 (D) and  $\ell_n = M_n \mathbb{E}[D_n^\nu]$  by (2.1), we can rewrite (3.46) as

$$\sup_{s \leq s_0} \left| \ell_n^{-1} Z''(s) - G_{D^{\nu,*}}(e^{-s}) \right| \xrightarrow{\mathbb{P}} 0, \tag{3.47}$$

for any  $s_0$  fixed. If  $c = f(s) = G_{D^{\nu,*}}(e^{-s})$ , then  $s = f^{(-1)}(c) = -\log \left( G_{D^{\nu,*}}^{(-1)}(c) \right)$ . Then by Claim 3.6 and (3.42), for any  $c_0$  fixed,

$$\sup_{c \geq c_0} \left| \tau_{\text{skip}}(c) + \log \left( G_{D^{\nu,*}}^{(-1)}(c) \right) \right| \xrightarrow{\mathbb{P}} 0. \tag{3.48}$$

This concludes the proof of (3.20).

### 3.4 | The giant of the BCM: Proof of Theorem 2.11

In this section, we combine the insight gained above and prove Theorem 2.11. For technical reasons, our proof requires that  $\eta_\ell > 0$ , with  $\eta_\ell$  defined in Theorem 2.6. Note that the composition  $G_{\tilde{D}^r} \circ G_{\tilde{D}^r}$  from (2.18) is the generating function of the random sum

$$N^r := \sum_{i=1}^{\tilde{D}^r} \tilde{D}_{(i)}^\ell, \tag{3.49}$$

where  $\tilde{D}_{(i)}^\ell$  are iid copies of  $\tilde{D}^\ell$  and independent of  $\tilde{D}^r$ . (This random variable will reappear in Section 5.1 where it receives an intuitive explanation.) Note that, by properties of generating functions,  $\eta_\ell = 0$  exactly when  $G_{N^r}(0) = \mathbb{P}(N^r = 0) = 0$ , which is equivalent to  $\mathbb{P}(\tilde{D}^\ell = 0) = \mathbb{P}(\tilde{D}^r = 0) = 0$  by (3.49), which in turn is equivalent to  $p_1 = q_1 = 0$  by (1.1). In the proof, we shall impose the condition  $q_1 > 0$ , with  $q_1$  defined in Assumption 2.3(C1), to ensure that  $\eta_\ell > 0$ . Hence we first show that proving Theorem 2.11 for  $q_1 > 0$  is sufficient:

*Claim 3.8* (Reduction to the case  $q_1 > 0$ ). Theorem 2.11 with  $q_1 > 0$  implies Theorem 2.11 for  $q_1 = 0$ .

*Proof.* Assume that Theorem 2.11 holds for  $q_1 > 0$ , and we are given a graph sequence with  $q_1 = 0$ . In the following, we introduce a modification of the graph sequence, parametrized by  $\varepsilon$ , such that  $q_1(\varepsilon) > 0$  for all  $\varepsilon > 0$ , while we get better approximations of the original graph sequence as  $\varepsilon \rightarrow 0$ . Let  $d_{\min}^r := \min\{k \in \mathbb{Z}^+ : q_k > 0\} \geq 2$ , that is, the minimal degree of the asymptotic  $r$ -degree distribution, and fix  $\varepsilon$  such that  $0 < \varepsilon < q_{d_{\min}^r}$ . Then (for  $n$  large enough) we cut  $\varepsilon M_n$   $r$ -vertices of degree  $d_{\min}^r$  into vertices of degree 1, that is, we replace each of them by  $d_{\min}^r$   $r$ -vertices of degree 1. The empirical  $\ell$ -degrees  $D_\varepsilon^{r,(n)}$  then converge as  $n \rightarrow \infty$  to a modified limit  $D_\varepsilon^r$  with pmf given by

$$q_k(\varepsilon) := \begin{cases} \frac{\varepsilon d_{\min}^r}{1 + \varepsilon(d_{\min}^r - 1)} & \text{for } k = 1, \\ \frac{q_{d_{\min}^r} - \varepsilon}{1 + \varepsilon(d_{\min}^r - 1)} & \text{for } k = d_{\min}^r, \\ \frac{q_k}{1 + \varepsilon(d_{\min}^r - 1)} & \text{otherwise.} \end{cases} \tag{3.50}$$

Denote the smallest fixed point of  $G_{\tilde{D}_\varepsilon^r} \circ G_{\tilde{D}_\varepsilon^r}$  by  $\eta_\ell(\varepsilon) > 0$  and define  $\xi_\ell(\varepsilon) := 1 - G_{D_\varepsilon^r}(\eta_\ell(\varepsilon)) < 1$ . (Recall  $\xi_\ell$  from Theorem 2.6.) By our assumptions, Theorem 2.11 holds for the modified graph sequence, and consequently formulas (2.32)–(2.34) hold with  $\eta_\ell(\varepsilon) > 0$  and  $\xi_\ell(\varepsilon) < 1$ . We now let  $\varepsilon \rightarrow 0$ , then  $D_\varepsilon^r \xrightarrow{d} D^r$ , thus  $G_{\tilde{D}_\varepsilon^r} \rightarrow G_{\tilde{D}^r}$  pointwise on  $[0, 1]$ , which implies  $\eta_\ell(\varepsilon) \rightarrow \eta_\ell = 0$  and  $\xi_\ell(\varepsilon) \rightarrow \xi_\ell = 1$ . The above cutting operation can only decrease the number of  $\ell$ -vertices in each connected component: we are duplicating  $r$ -vertices only and components may become disconnected. Then considering that  $\xi_\ell(\varepsilon) \rightarrow 1$ , (2.32) must extend to  $\xi_\ell = 1$  as well, and (2.33)–(2.34) follow for  $\eta_\ell = 0$ . ■

#### 3.4.1 | Identifying components in the exploration

In the following, wlog we assume that  $q_1 > 0$  or equivalently,  $\tilde{q}_0 > 0$ , with  $\tilde{q}_0 = q_1/\mathbb{E}[D^r]$  defined in Proposition 3.5. Recall (3.5) and (3.14) and define, for  $z \in [\tilde{q}_0, 1]$ ,

$$H(z) := h_2(z) - h_1(z) = \mathbb{E}[D^\ell]z \left( G_{\tilde{D}^r}^{(-1)}(z) - G_{\tilde{D}^r}(z) \right). \tag{3.51}$$

For convenience, denote

$$h(z) := G_{\tilde{D}^r}^{(-1)}(z) - G_{\tilde{D}^r}(z). \tag{3.52}$$

Recall the process  $\hat{A}^\ell(t)$  from (3.9) that approximates  $A^\ell(t)$ . By Lemma 3.3 and Proposition 3.5, for any  $t_0 < -\log \tilde{q}_0$ ,  $\hat{A}^\ell(t)$  satisfies

$$\sup_{t \leq t_0} \left| \frac{1}{N_n} \hat{A}^\ell(t) - H(e^{-t}) \right| \xrightarrow{\mathbb{P}} 0. \tag{3.53}$$

Recall that we start exploring a new component when `step1` is executed, for which  $A^\ell(t) = 0$  is a necessary condition. By the intuition that  $A^\ell(t)/N_n \approx \hat{A}^\ell(t)/N_n$  from Lemma 3.4, and (3.53), we want to find the zero(s) of  $t \mapsto H(e^{-t})$  on  $\mathbb{R}^+$ . By (3.51)–(3.52), the zeros of this function are described by  $h(e^{-t}) = 0$ . Rearranging leads to the fixed point equation  $G_{\tilde{D}^r}(G_{\tilde{D}^r}(e^{-t})) = e^{-t}$  for some  $t \in \mathbb{R}^+$ , or equivalently,  $G_{\tilde{D}^r} \circ G_{\tilde{D}^r}(z) = z$  for some  $z \in (0, 1)$ , which is the generating function of  $N^r$  defined in (3.49). We always have the trivial fixed point 1, however whether a second fixed point exists or not depends on whether the derivative  $G'_{N^r}(1) = \mathbb{E}[N^r] = \mathbb{E}[\tilde{D}^r] \mathbb{E}[\tilde{D}^\ell] > 1$ , which is exactly the supercriticality condition (2.20). In the following, we study the two cases separately, and show that a giant component exists if and only if (2.20) holds.

### 3.4.2 | The supercritical case

First, we study the case when (2.20) holds. Recall  $N^r$  from (3.49) and (1.2). Since  $G'_{N^r}(1) > 1$ , there exists a second fixed point  $\eta_\ell < 1$  in the interval  $[0, 1]$ . In fact,  $\eta_\ell > \tilde{q}_0$  with  $\tilde{q}_0$  from Proposition 3.5, by the following reasoning. By the definition of  $N^r$ ,  $\mathbb{P}(N^r = 0) \geq \mathbb{P}(\tilde{D}^r = 0) = \tilde{q}_0 = q_1/\mathbb{E}[D^r]$ , which is positive by assumption. Thus the fixed point  $\eta_\ell$  cannot be 0, and consequently by the strict monotonicity of  $G_{N^r}$ ,  $\eta_\ell = G_{N^r}(\eta_\ell) > G_{N^r}(0) > \mathbb{P}(N^r = 0) \geq \tilde{q}_0$ . Define

$$t^* := -\log \eta_\ell, \tag{3.54}$$

which lies in  $(0, -\log \tilde{q}_0)$ , and consequently  $t^*$  is the *unique* value of  $t \in \mathbb{R}^+$  such that  $H(e^{-t^*}) = 0$ . In the following, we work toward showing that the exploration of the giant component lasts from time  $0 + o_{\mathbb{P}}(1)$  to time  $t^* \pm o_{\mathbb{P}}(1)$ . Define

$$t_0 := -\log((\eta_\ell + \tilde{q}_0)/2), \tag{3.55}$$

so that  $t^* < t_0 < -\log \tilde{q}_0$ , and denote the “good event”

$$\mathcal{E}_1(\delta) = \mathcal{E}_1^{(n)}(\delta) := \left\{ \sup_{t \leq t_0} \left| N_n^{-1} \hat{A}^\ell(t) - H(e^{-t}) \right| < \delta \right\}. \tag{3.56}$$

Note that since  $t_0 < -\log \tilde{q}_0$ , both Lemma 3.3 and Proposition 3.5 are applicable for this choice of  $t_0$ . Consequently, for any fixed  $\delta$ , by (3.53) the good event happens whp, that is,

$$\mathbb{P}(\mathcal{E}_1^{(n)}(\delta)) \rightarrow 1 \tag{3.57}$$

as  $n \rightarrow \infty$ . By properties of the generating function  $G_{N^r}$  and (3.51), rearranging yields that  $z \mapsto H(z)$  is positive for  $z \in (\eta_\ell, 1)$ , thus  $t \mapsto H(e^{-t})$  is positive for  $t \in (0, t^*)$ . In fact, we have the following analytical properties of  $t \mapsto H(e^{-t})$ :

*Claim 3.9.* For any  $\varepsilon > 0$  small enough, there exists  $\delta = \delta(\varepsilon) > 0$  such that  $t \mapsto H(e^{-t}) > \delta$  on  $t \in (\varepsilon, t^* - \varepsilon)$  and  $H(e^{-(t^* + \varepsilon)}) < -2\delta$ .

*Proof.* Recall (3.51) and (3.52) and note that  $0 < e^{-t_0} < e^{-t} < 1$  is bounded for  $t \in (0, t_0)$ . It is sufficient to show that, for some  $\delta' > 0$ ,

$$\begin{cases} h(e^{-t}) > \delta' & \text{on } (\varepsilon, t^* - \varepsilon), \\ h(e^{-t}) < -2\delta' & \text{at } t = t^* + \varepsilon, \end{cases} \tag{3.58}$$

then the required statement follows for  $\delta := \delta' \mathbb{E}[D^\ell] e^{-t_0}$ . By the strict monotonicity of the mapping  $t \mapsto e^{-t}$ , (3.58) is equivalent to

$$\begin{cases} h(z) > \delta' & \text{for } z \in (e^{-(t^* - \varepsilon)}, e^{-\varepsilon}) = (\eta_\ell + \varepsilon_1, 1 - \varepsilon_2), \\ h(z) < -2\delta' & \text{for } z = e^{-(t^* + \varepsilon)} < \eta_\ell. \end{cases} \tag{3.59}$$

Recall (3.52). Note that by  $q_1 > 0$ ,  $z \mapsto h(z)$  is strictly concave on its domain  $[\tilde{q}_0, 1]$  and positive exactly on  $(\eta_\ell, 1)$ , hence for any  $\varepsilon$  fixed, we can choose  $\delta' > 0$  appropriately such that (3.59) holds. This concludes the proof of Claim 3.9. ■

**Finding the largest component**

In the following, we aim to characterize those executions of `step1` where we start exploring the giant component and the component thereafter, that is, when we finish exploring the giant. Recall the definition of  $\mathcal{S}_1$  from Algorithm 3.1. Denote the last element of  $\mathcal{S}_1$  that is less than  $t^*/2$  by  $T_1$ , and denote the next element after  $T_1$  by  $T_2$ , that is,  $T_2$  is the first element of  $\mathcal{S}_1$  that is at least  $t^*/2$ . Formally,

$$T_1 = \max\{t \in \mathcal{S}_1 : t \leq t^*/2\}, \quad T_2 = \min\{t \in \mathcal{S}_1 : t > t^*/2\}, \tag{3.60}$$

with the convention that the minimum over an empty set is  $+\infty$ . Later, we show that the exploration of the largest component lasts from  $T_1$  to  $T_2$ . We first show the following:

**Lemma 3.10** (Exploration time of the ‘‘giant’’). *As  $n \rightarrow \infty$ ,*

$$T_1 \xrightarrow{\mathbb{P}} 0, \quad T_2 \xrightarrow{\mathbb{P}} t^*. \tag{3.61}$$

*Proof.* Note that  $A^\ell(t) - \widehat{A}^\ell(t) = \widehat{S}^\ell(t) - S^\ell(t) > 0$  by definition. By (3.56) and Claim 3.9, on the event  $\mathcal{E}_1(\delta)$  for  $t \in (\varepsilon, t^* - \varepsilon)$ ,

$$A^\ell(t) \geq \widehat{A}^\ell(t) > \delta N_n > 0. \tag{3.62}$$

Recall that executing `step1` requires  $A^\ell(t) = 0$ . Consequently, on the event  $\mathcal{E}_1(\delta)$ , `step1` could not have been executed within the time interval  $(\varepsilon, t^* - \varepsilon)$ , hence on this event,

$$T_1 \leq \varepsilon, \quad T_2 \geq t^* - \varepsilon. \tag{3.63}$$

Noting that  $0 \in \mathcal{S}_1$ , thus  $0 \leq T_1$ , it follows that  $T_1 \xrightarrow{\mathbb{P}} 0$  by (3.57) and (3.63). We have yet to give an upper bound on  $T_2$  to prove that  $T_2 \xrightarrow{\mathbb{P}} t^*$ . We do so by proving that `step1` must have been executed



between  $t^* - \varepsilon$  and  $t^* + \varepsilon$ . In fact, we show that the error  $A^\ell - \widehat{A}^\ell$  has increased on the smaller interval between  $t^*$  and  $t^* + \varepsilon$ , which can only happen due to `step1`, as discussed in the proof of Lemma 3.4. Recall that  $H(e^{-t})$  is positive on  $(0, t^*)$ , hence on the event  $\mathcal{E}_1(\delta)$ ,  $N_n^{-1} \widehat{A}^\ell(t) > -\delta$  on  $(0, t^*)$ . Lemma 3.4 is applicable for our choice of  $t_0$  (see (3.55)) and  $t^* < t_0$ , thus for any fixed  $\delta$ ,

$$A^\ell(t^*) - \widehat{A}^\ell(t^*) \leq -\inf_{t \leq t^*} \widehat{A}^\ell(t) + d_{\max}^\ell < \delta N_n + \delta N_n / 2 = (3/2)\delta N_n, \tag{3.64}$$

as  $d_{\max}^\ell < \delta N_n / 2$  for  $n$  large enough by Remark 2.4(iii). However, on the event  $\mathcal{E}_1(\delta)$ ,

$$N_n^{-1} \widehat{A}^\ell(t^* + \varepsilon) \leq -2\delta \tag{3.65}$$

by Claim 3.9, while  $A^\ell(t) \geq 0$  for any  $t$ . Thus

$$A^\ell(t^* + \varepsilon) - \widehat{A}^\ell(t^* + \varepsilon) \geq 2\delta N_n. \tag{3.66}$$

Comparing (3.64) and (3.66), we see that  $A^\ell(t) - \widehat{A}^\ell(t)$  increased between times  $t^*$  and  $t^* + \varepsilon$ , which is only possible when `step1` is executed. Consequently  $T_2 \leq t^* + \varepsilon$  on the event  $\mathcal{E}_1(\delta)$  that happens whp. Combining this with (3.63), we obtain that  $T_2 \xrightarrow{\mathbb{P}} t^*$ , concluding the proof of Lemma 3.10. ■

**Properties of the giant candidate**

Recall that the exploration of each component starts with an execution of `step1`, thus by (3.60), only one component is explored in the time interval  $(T_1, T_2)$ . Let us denote this component by  $\mathcal{C}_\star = \mathcal{C}_\star^{(n)}$ . We study some properties of  $\mathcal{C}_\star$  that will help us in showing that  $\mathcal{C}_\star$  is whp the largest component. Recall from Section 3.1 that  $\ell$ -vertices can be sleeping or awake and  $\ell$ -half-edges can be sleeping, active or paired. Also recall that  $\mathcal{V}_k^\ell(t)$  denotes the number of  $\ell$ -vertices of degree  $k$  still sleeping at time  $t$ . Since  $T_1, T_2 \in \mathcal{S}_1$ , we have  $A^\ell(T_1) = A^\ell(T_2) = 0$ , thus all  $\ell$ -half-edges that are removed from the sleeping set between  $T_1$  and  $T_2$  must be paired by time  $T_2$ . In turn, therefore, all  $\ell$ -vertices and  $\ell$ -half-edges that are removed from the sleeping set between  $T_1$  and  $T_2$  are part of the component  $\mathcal{C}_\star$ . Hence, with  $\mathcal{V}_k^\ell$  defined in (2.7),

$$|\mathcal{V}_k^\ell \cap \mathcal{C}_\star| = \mathcal{V}_k^\ell(T_1) - \mathcal{V}_k^\ell(T_2), \tag{3.67}$$

$$|\mathcal{E}(\mathcal{C}_\star)| = S^\ell(T_1) - S^\ell(T_2). \tag{3.68}$$

Recall that  $t^*$  is defined in (3.54) so that  $H(e^{-t^*}) = 0$ , and further,  $H(e^{-t}) > 0$  for  $t \in (0, t^*)$ . By Lemma 3.10 and the continuity of  $H$ ,  $\inf_{t \leq T_2} H(e^{-t}) \xrightarrow{\mathbb{P}} \inf_{t \leq t^*} H(e^{-t}) = 0$ . Thus whp (3.53) applies to  $T_2$  and yields  $N_n^{-1} \inf_{t \leq T_2} \widehat{A}^\ell(t) \xrightarrow{\mathbb{P}} 0$  as well. Note that  $\widehat{\mathcal{V}}_k^\ell(t) \geq \mathcal{V}_k^\ell(t)$  for all  $t$  and  $k$ , with  $\widehat{\mathcal{V}}_k^\ell(t)$  defined in Section 3.2. Recall (3.3) and (3.4). By Lemma 3.4,

$$\frac{1}{N_n} \sup_{t \leq T_2} \left( \widehat{\mathcal{V}}_k^\ell(t) - \mathcal{V}_k^\ell(t) \right) \leq \frac{1}{N_n} \sup_{t \leq T_2} \left( \widehat{S}^\ell(t) - S^\ell(t) \right) \leq \frac{1}{N_n} \inf_{t \leq T_2} \widehat{A}^\ell(t) + \frac{d_{\max}^\ell}{N_n} \xrightarrow{\mathbb{P}} 0. \tag{3.69}$$

Combining (3.67) and (3.69) with (3.6) from Lemma 3.3,

$$N_n^{-1} \left| \mathcal{V}_k^\ell \cap \mathcal{C}_\star \right| - (p_k e^{-kT_1} - p_k e^{-kT_2})$$

$$\begin{aligned}
 &= N_n^{-1} (\mathcal{V}_k^\ell(T_1) - \mathcal{V}_k^\ell(T_2)) - N_n^{-1} (\widehat{\mathcal{V}}_k^\ell(T_1) - \widehat{\mathcal{V}}_k^\ell(T_2)) \\
 &\quad + \left( N_n^{-1} \widehat{\mathcal{V}}_k^\ell(T_1) - p_k e^{-kT_1} \right) - \left( N_n^{-1} \widehat{\mathcal{V}}_k^\ell(T_2) - p_k e^{-kT_2} \right) \xrightarrow{\mathbb{P}} 0.
 \end{aligned} \tag{3.70}$$

Since the function  $t \mapsto p_k e^{-kt}$  is continuous, by Lemma 3.10 and (3.54),

$$p_k (e^{-kT_1} - e^{-kT_2}) \xrightarrow{\mathbb{P}} p_k (e^{-k0} - e^{-kt^*}) = p_k (1 - \eta_\ell^k). \tag{3.71}$$

Then combining (3.70) and (3.71) yields

$$N_n^{-1} \left| \mathcal{V}_k^\ell \cap \mathcal{E}_\star \right| \xrightarrow{\mathbb{P}} p_k (1 - \eta_\ell^k). \tag{3.72}$$

Similarly, by summation and (3.7), as well as (3.68) and (3.8), respectively,

$$N_n^{-1} \left| \mathcal{V}^\ell \cap \mathcal{E}_\star \right| \xrightarrow{\mathbb{P}} G_{D^\ell} (e^{-0}) - G_{D^\ell} (e^{-t^*}) = 1 - G_{D^\ell}(\eta_\ell) = \xi_\ell, \tag{3.73}$$

$$N_n^{-1} |\mathcal{E}(\mathcal{E}_\star)| \xrightarrow{\mathbb{P}} \mathbb{E}[D^\ell] (1 - \eta_\ell G_{D^\ell}(\eta_\ell)) = \mathbb{E}[D^\ell] (1 - \eta_\ell \eta_r). \tag{3.74}$$

In particular,  $\mathcal{E}_\star$  contains a linear proportion of edges and  $\ell$ -vertices.

**Uniqueness**

Next, we prove that whp there is no other component containing a linear proportion of edges and vertices, hence  $\mathcal{E}_\star$  must be  $\mathcal{E}_{1,\delta}$  and further, the giant component is unique. Since  $T_1 \xrightarrow{\mathbb{P}} 0$ , by (3.8), the total number of  $\ell$ -half-edges explored before  $\mathcal{E}_\star$  is  $o_{\mathbb{P}}(N_n)$ . Consequently, whp no linear-sized component is explored before  $\mathcal{E}_\star$ . Let us define  $T_3$  as the element in  $\mathcal{S}_1$  right after  $T_2$  (and  $\infty$  if there is no such element). It may occur that  $T_3 = T_2$ , due to the multiplicities in the sequence  $\mathcal{S}_1$ . The time of  $T_3$  is given by

$$T_3 = \min \{ t \in \mathcal{S}_1 \setminus \{T_2\} : t > t^*/2 \}. \tag{3.75}$$

Recall (3.64) and (3.66) that we have used to prove that `step1` must have been executed between  $t^*$  and  $t^* + \varepsilon$ , since the difference  $A^\ell - \widehat{A}^\ell$  can only increase due to `step1`. In fact, we have shown that on the event  $\mathcal{E}_1(\delta)$ , the difference increased by at least  $\delta N_n/2$ , that is, *linearly* with  $N_n$ ; however, each execution of `step1` can only increase the difference by  $d_{\max}^\ell$ , which is  $o(N_n)$  by Remark 2.4(iii). Thus, `step1` must have been executed not once, but many times between  $t^*$  and  $t^* + \varepsilon$ ; in particular,  $T_3 \leq t^* + \varepsilon$  on the event  $\mathcal{E}_1(\delta)$ , which happens whp. Combining this with  $T_3 \geq T_2$  and  $T_2 \xrightarrow{\mathbb{P}} t^*$  yields that  $T_3 \xrightarrow{\mathbb{P}} t^*$ . Hence the component  $\mathcal{E}'$  explored between  $T_2$  and  $T_3$  has  $o_{\mathbb{P}}(N_n)$  edges by (3.8).

Now assume that for some  $\alpha > 0$ , there exists a component  $\widehat{\mathcal{E}}$  with  $\alpha N_n$  many edges, that was not explored before  $\mathcal{E}_\star$ . Then, since we pick a new vertex by choosing a *uniform* sleeping  $\ell$ -half-edge in `step1`, we find  $\widehat{\mathcal{E}}$  at  $T_2$  with *positive* probability, that is,  $\mathbb{P}(\mathcal{E}' = \widehat{\mathcal{E}}) > 0$  (where  $\mathcal{E}'$  is the component explored between  $T_2$  and  $T_3$ ), which implies that  $\liminf_{n \rightarrow \infty} \mathbb{P}(|\mathcal{E}(\mathcal{E}')|/N_n \geq \alpha) > 0$ . This contradicts that  $|\mathcal{E}(\mathcal{E}')|/N_n \xrightarrow{\mathbb{P}} 0$ , and thus  $\widehat{\mathcal{E}}$  cannot exist. We conclude that whp no component containing a linear proportion of edges was explored before or after  $\mathcal{E}_\star$ . Note that if a connected component has linearly many vertices, it must also have linearly many edges. Hence whp  $\mathcal{E}_{1,\delta} = \mathcal{E}_\star$  is the largest component, and is unique in the sense that there is no other linear-sized component. Then

the properties proven for  $\mathcal{E}_*$  in (3.72)-(3.74) verify the claimed properties of the giant in (2.32)-(2.34). This concludes the proof of the *supercritical case* of Theorem 2.11.

3.4.3 | The non-supercritical case

We now study the case when (2.20) does not hold. With  $N^r$  from (3.49), we now have  $\mathbb{E}[N^r] = \mathbb{E}[\tilde{D}^\ell] \mathbb{E}[\tilde{D}^r] \leq 1$ , thus  $G_{N^r} = G_{\tilde{D}^r} \circ G_{\tilde{D}^\ell}$  only has the trivial fixed point 1. It is straightforward to check (rearranging (3.51)) that  $t \mapsto H(e^{-t})$  is then negative on  $\mathbb{R}^+$ , its last and only zero is  $t_{\text{subcrit}}^* = 0$ .

Let us denote the first two elements of  $\mathcal{S}_1$  by  $T'_1 = t_{\text{subcrit}}^* = 0$  and  $T'_2 = \min \{t \in \mathcal{S}_1 \setminus \{T'_1\}\}$ . By Lemma 3.4 and its proof, we have that  $A^\ell(0) - \hat{A}^\ell(0) = o_{\mathbb{P}}(N_n)$  and since the error can only increase due to step1,  $A^\ell(T'_2) - \hat{A}^\ell(T'_2) \leq A^\ell(0) - \hat{A}^\ell(0) + d_{\text{max}}^\ell = o_{\mathbb{P}}(N_n)$ . On the other hand, for any  $\varepsilon > 0$ ,  $N_n^{-1} \hat{A}^\ell(\varepsilon) \xrightarrow{\mathbb{P}} H(e^{-\varepsilon}) < 0$  by (3.53). Noting that  $A^\ell(\varepsilon) \geq 0$ ,  $N_n^{-1}(A^\ell(\varepsilon) - \hat{A}^\ell(\varepsilon)) > 0$  whp, hence  $T'_2 < \varepsilon$  whp, that is,  $T'_2 \xrightarrow{\mathbb{P}} 0$ . Denote the component explored between  $T'_1$  and  $T'_2$  by  $\mathcal{E}^0$ , so that  $|\mathcal{E}(\mathcal{E}^0)| = o_{\mathbb{P}}(N_n)$  by (3.8). With an analogous argument to the proof of uniqueness in the supercritical case, no linear-sized component can exist, since we would find it at time  $T'_1$  with positive probability. Hence  $|\mathcal{E}_{1,\ell}^{(n)}| = o_{\mathbb{P}}(N_n)$ . This concludes the proof of Theorem 2.11. ■

We remark that the non-supercritical case could alternatively be proved by showing that whp  $|\mathcal{E}_1|/N_n \leq \xi_\ell + \varepsilon$  by combining the local weak convergence result [48, Theorem 2.7] with arguments similar to those in Section 5.

4 | PERCOLATION PHASE TRANSITION AND THE GIANT OF THE RIGC

In this section, we prove Theorem 2.13 as a consequence of Theorem 2.6.

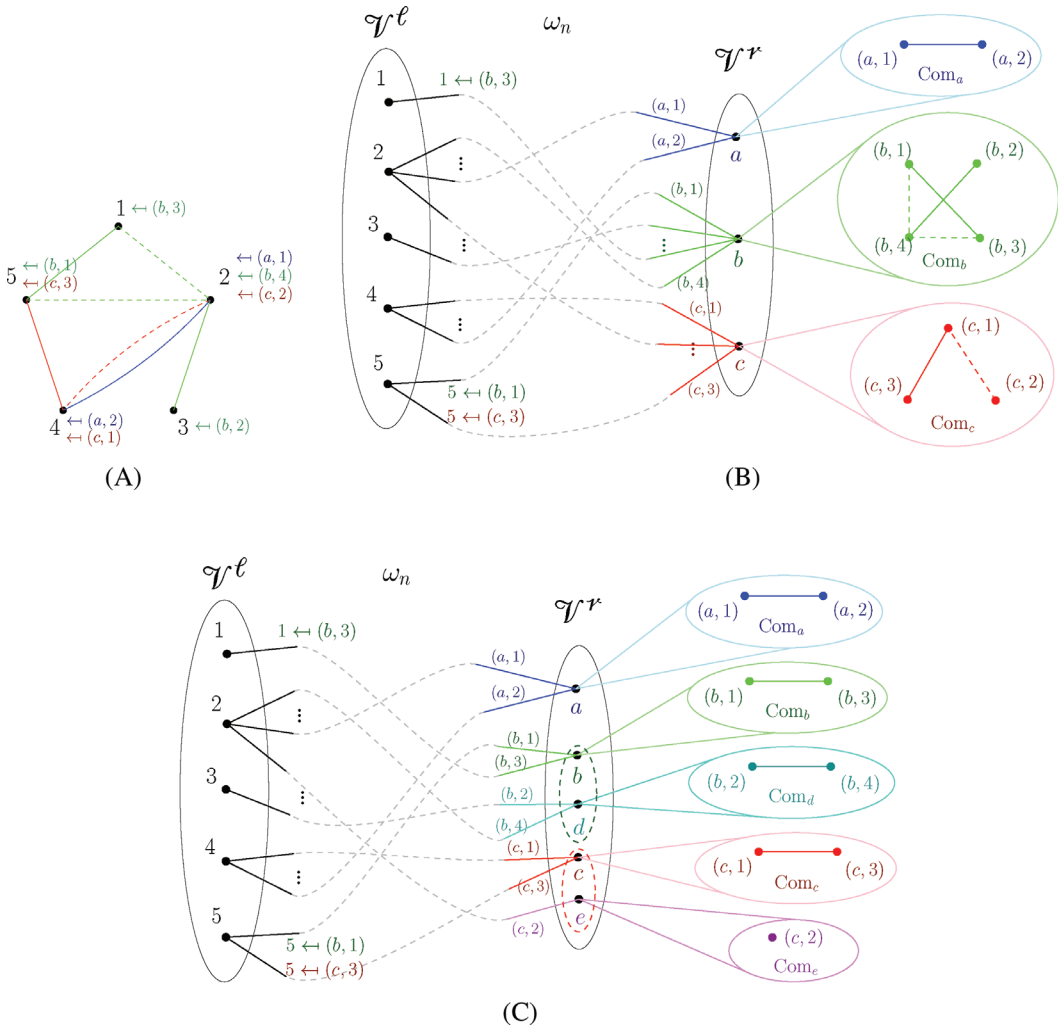
4.1 | Percolation on the RIGC represented as an RIGC with random parameters

First, we focus on a qualitative understanding of the bond percolation model.

Recall the construction of the RIGC from Section 2.1. Recall that  $\mathcal{E}(\mathbf{Com})$  denotes the disjoint union of edges in all community graphs. Further, denote the probability measure of the bipartite matching  $\omega_n$  by  $\mathbb{P}_{\omega_n}$ . For a given  $\omega_n$ , denote the edge set of the corresponding realization of the RIGC by  $\mathcal{E}(\omega_n)$ . Note that by construction and by our choice of treating the RIGC as a multigraph, for any given  $\omega_n$ , there is a one-to-one correspondence between  $\mathcal{E}(\mathbf{Com})$  and  $\mathcal{E}(\omega_n)$ . For  $e \in \mathcal{E}(\mathbf{Com})$ , we denote the corresponding edge  $e' = e(\omega_n) \in \mathcal{E}(\omega_n)$ .

Recall that percolation is defined conditionally on the realization of the random graph RIGC, as follows. Given  $\omega_n$ , each edge  $e' \in \mathcal{E}(\omega_n)$  is assigned an independent Bernoulli random variable  $X_{e'}$  with success probability  $\pi$ , and we denote this conditional measure by  $\mathbb{P}_\pi(\cdot|\omega_n)$ . Together with the measure  $\mathbb{P}_{\omega_n}$  of  $\omega_n$ , this determines the joint measure  $\mathbb{P}_\pi$  of the percolated graph  $\text{RIGC}(\pi)$ . In the following, we establish an alternative representation as a *product measure*. Intuitively, we make use of the correspondence between  $\mathcal{E}(\mathbf{Com})$  and  $\mathcal{E}(\omega_n)$  to define percolation on the communities, rather than on the RIGC, which can be done *independently* of the bipartite matching (Figure 2).

We define percolation on the communities and the percolated community list  $\mathbf{Com}(\pi)$ , as follows. With each  $e \in \mathcal{E}(\mathbf{Com})$ , we associate an independent Bernoulli random variable  $X_e$ ;  $e$  is retained exactly when  $X_e = 1$ . Denote by  $\text{Com}_a(\pi)$  the random graph produced by percolation on  $\text{Com}_a$ . Note that  $\text{Com}_a(\pi)$  is not necessarily connected, which conflicts with our initial assumptions. Thus, we need to replace  $\text{Com}_a(\pi)$  by the random list of its connected components



**FIGURE 2** Reducing percolation on the RIGC to percolation on the communities. (A) The percolated RIGC. The removed edges are represented as dashed. (B) Percolation on the communities. Using the group memberships, we can “trace back” each removed edge to a community edge. (C) The new, percolated community list. If communities become disconnected, we separate each connected component as its own community, for example,  $b$  is separated into  $b$  and  $d$

$(\text{Com}_{a,i}(\pi))_{i \in [c(\text{Com}_a(\pi))]}$ , where  $c(\text{Com}_a(\pi))$  denotes the number of connected components of  $\text{Com}_a(\pi)$ . Then  $(\text{Com}_{a,i}(\pi))_{a \in [M_n], i \in [c(\text{Com}_a(\pi))]}$  is the new list of communities. We introduce the new number of communities  $M_n(\pi) := \sum_{a \in \mathcal{V}^r} c(\text{Com}_a(\pi))$ , so that the new rhs partition is  $[M_n(\pi)]$ . By reindexing, we can now write and define  $\mathbf{Com}(\pi) := (\text{Com}_{a'}^{\pi})_{a' \in [M_n(\pi)]}$ . With these new parameters, the above intuition can be formalized as follows:

**Proposition 4.1** (Percolation on the RIGC is still an RIGC). *Bond percolation with edge retention probability  $\pi$  on an RIGC with parameters  $d^\ell$  and  $\mathbf{Com}$  is equivalent to an RIGC with parameters  $d^\ell$  and  $\mathbf{Com}(\pi)$ . Formally,*

$$\text{RIGC}(d^\ell, \mathbf{Com})(\pi) \stackrel{d}{=} \text{RIGC}(d^\ell, \mathbf{Com}(\pi)). \tag{4.1}$$

We refer to  $\text{RIGC}(d^\ell, \mathbf{Com}(\pi))$  as the RIGC representation of  $\text{RIGC}(\pi)$ .

*Proof.* Recall that given  $\omega_n$ , percolation on the RIGC is described by the iid Bernoulli( $\pi$ ) random variables  $(X_{e'})_{e' \in \mathcal{E}(\omega_n)}$ . Also recall that each  $e' \in \mathcal{E}(\omega_n)$  can be written as  $e' = e(\omega_n)$  for a unique  $e \in \mathcal{E}(\mathbf{Com})$  and define  $X_{e(\omega_n)} := X_{e'}$  for each  $e \in \mathcal{E}(\mathbf{Com})$ .

A given realization of RIGC( $\pi$ ) can be characterized by its (unpercolated) edge set  $\mathcal{E}(\omega_n) = \mathcal{E}$  and the outcomes of the Bernoulli variables,  $x_{e'} \in \{0, 1\}$  for  $e' \in \mathcal{E}$ . Define  $x_e := x_{e'}$ , with  $e' = e(\omega_n)$ . We have that, for any given edge set  $\mathcal{E}$  and  $(x_{e'})_{e' \in \mathcal{E}}$ ,

$$\begin{aligned} & \mathbb{P}_\pi \left( \mathcal{E}(\omega_n) = \mathcal{E}, X_{e'} = x_{e'} \forall e' \in \mathcal{E} \right) \\ &= \mathbb{P}_{\omega_n} \left( \mathcal{E}(\omega_n) = \mathcal{E} \right) \mathbb{P}_\pi \left( X_{e'} = x_{e'} \forall e' \in \mathcal{E}(\omega_n) \mid \omega_n \right) \\ &= \mathbb{P}_{\omega_n} \left( \mathcal{E}(\omega_n) = \mathcal{E} \right) \mathbb{P}_\pi \left( X_e(\omega_n) = x_e \forall e \in \mathcal{E}(\mathbf{Com}) \mid \omega_n \right) \\ &= \mathbb{P}_{\omega_n} \left( \mathcal{E}(\omega_n) = \mathcal{E} \right) \mathbb{P} \left( X_e = x_e \forall e \in \mathcal{E}(\mathbf{Com}) \right), \end{aligned} \tag{4.2}$$

where in the last step we have used that for any  $\omega_n$ ,  $(X_e(\omega_n))_{e \in \mathcal{E}(\mathbf{Com})}$  are independent Bernoulli( $\pi$ ) random variables, thus the collection has the same law as  $(X_e)_{e \in \mathcal{E}(\mathbf{Com})}$ . We conclude that the law of the percolated graph can indeed be written as a product measure.

Noting that  $\omega_n$  did not change throughout (4.2), we conclude that the new measure is still an RIGC. Similarly, as  $\mathbf{d}^\ell$  did not even appear in the formulas, it necessarily remains unchanged. As intuition has predicted, percolation can be executed on the communities before constructing the random graph, the formulas indeed contain the random variables  $X_e$  corresponding to  $e \in \mathcal{E}(\mathbf{Com})$ , meaning that the new RIGC must use  $\mathbf{Com}(\pi)$ . This concludes the proof of Proposition 4.1. ■

Next, we show that RIGC( $\mathbf{d}^\ell, \mathbf{Com}(\pi)$ ) still satisfies our assumptions, in the sense of Remark 2.5. Denote the (random) empirical distribution of  $\mathbf{Com}(\pi)$  by  $\mu^{(n)}(\pi)$ . Then, we have the following convergence result, in which we recall that  $\mathcal{H}$  denotes the set of all simple, finite and connected graphs:

**Lemma 4.2** (Convergence of percolated community list). *Assume that the original  $\mathbf{Com}$  sequence satisfies Assumption 2.3(C). Then for the sequence of  $\mathbf{Com}(\pi)$ , there exists a mass function  $\mu(\pi)$  on  $\mathcal{H}$ , such that, for each  $H \in \mathcal{H}$ , as  $n \rightarrow \infty$ ,*

$$\mu_H^{(n)}(\pi) \xrightarrow{\mathbb{P}} \mu_H(\pi). \tag{4.3}$$

Denote the empirical and limiting community-size distributions corresponding to  $\mu^{(n)}(\pi)$  and  $\mu(\pi)$ , respectively, by  $D_n^r(\pi)$  and  $D^r(\pi)$ . If the original  $\mathbf{Com}$  sequence also satisfies Assumption 2.3 (D), then

$$\mathbb{E} \left[ D_n^r(\pi) \mid \mathbf{Com}(\pi) \right] \xrightarrow{\mathbb{P}} \mathbb{E}[D^r(\pi)] < \infty. \tag{4.4}$$

We prove Lemma 4.2 in Section 4.3.1. Recall that  $\mathcal{V}(\mathbf{Com})$  denotes the disjoint union of vertices in all community graphs, and recall  $J_n \sim \text{Unif}[\mathcal{V}(\mathbf{Com})]$  and (1.1). For  $j \in \mathcal{V}(\mathbf{Com})$ , let  $\mathcal{C}^e(j, \pi)$  denote the percolated component of  $j$  within its community. The following statement provides insight into the percolated community sizes and it is also instrumental to the proof of Proposition 2.14:

*Claim 4.3* (Representation of size-biased percolated community size). We have the following identity of distributions:

$$\widetilde{D_n^r(\pi)} \stackrel{d}{=} |\mathcal{C}^e(J_n, \pi)| - 1. \tag{4.5}$$

We give the proof of Claim 4.3 in Section 4.3.2.

4.2 | Proof of Theorem 2.13

We now prove Theorem 2.13 subject to Lemma 4.2.

*Proof.* By Proposition 4.1, studying percolated component sizes is equivalent to studying component sizes in  $\text{RIGC}(\mathbf{d}^\ell, \mathbf{Com}(\pi))$ , that is, the RIGC representation of  $\text{RIGC}(\pi)$ . By a slight abuse of notation, we use  $\mathcal{C}_1(\pi)$  and  $\mathcal{C}_2(\pi)$  to denote the largest and second largest component of  $\text{RIGC}(\mathbf{d}^\ell, \mathbf{Com}(\pi))$ , respectively. Recall  $\mu(\pi)$  from Lemma 4.2, and the corresponding limiting community-size distribution  $D^r(\pi)$ .

By Lemma 4.2 and Remark 2.5, our results apply to  $\text{RIGC}(\mathbf{d}^\ell, \mathbf{Com}(\pi))$ , the RIGC representation of  $\text{RIGC}(\pi)$ . In particular, applying Theorem 2.6 yields that

there exists  $\eta_\ell(\pi) \in [0, 1]$ , the smallest solution of the fixed point equation

$$\eta_\ell(\pi) = G_{\overline{D^r(\pi)}}(G_{D^r}(\eta_\ell(\pi))), \tag{4.6}$$

and  $\xi_\ell(\pi) := 1 - G_{D^r}(\eta_\ell(\pi)) \in [0, 1]$  such that

$$|\mathcal{C}_1(\pi)|/N_n \xrightarrow{\mathbb{P}} \xi_\ell(\pi). \tag{4.7}$$

Furthermore,  $\xi_\ell(\pi) > 0$  exactly when

$$\mathbb{E}[\widetilde{D}^\ell] \mathbb{E}[\widetilde{D^r(\pi)}] > 1, \tag{4.8}$$

which we call supercritical percolation. In this case,  $\mathcal{C}_1(\pi)$  is unique in the sense that  $|\mathcal{C}_2(\pi)| = o_{\mathbb{P}}(N_n)$ , and we call  $\mathcal{C}_1(\pi)$  the percolated giant component.

In the following, we show that there exists  $\pi_c \in [0, 1]$  such that the set of supercritical parameters  $\pi$  can be (almost exactly, as explained shortly) characterized by  $\pi > \pi_c$ . We do so by proving that  $\xi_\ell(\pi)$  is a non-decreasing function of  $\pi$ . Subject to this statement, clearly there exists a threshold  $\pi_c$  given by

$$\pi_c := \inf\{\pi : \xi_\ell(\pi) > 0\} = \inf\left\{\pi : \mathbb{E}[\widetilde{D}^\ell] \mathbb{E}[\widetilde{D^r(\pi)}] > 1\right\}. \tag{4.9}$$

This shows that the characterization is almost exact:  $\pi > \pi_c$  implies that  $\pi$  is supercritical, and  $\pi < \pi_c$  implies that  $\pi$  is not supercritical. Whether  $\pi_c$  itself is supercritical, that is, whether  $\eta_\ell(\pi_c) < 1$ , depends on continuity properties of  $\pi \mapsto \xi_\ell(\pi)$ , which are nontrivial in some cases. We conjecture that  $\pi = \pi_c$  is not supercritical; in [33], we show that this is true in special cases.

We now prove the required monotonicity of  $\xi_\ell(\pi)$ , by showing that for any  $n$  fixed,  $|\mathcal{C}_1(\pi)|/N_n$  is non-decreasing in  $\pi$ , in the sense of stochastic domination: we say that  $Y$  stochastically dominates  $X$  and denote  $X \leq Y$  if  $\mathbb{P}(Y \geq x) \geq \mathbb{P}(X \geq x)$  for all  $x \in \mathbb{R}$ . Subject to  $|\mathcal{C}_1(\pi)|/N_n$  being non-decreasing in  $\pi$ , clearly the required monotonicity follows for the limit  $\xi_\ell(\pi)$  as well, by (4.7). We prove monotonicity for fixed  $n$  through the so-called Harris-coupling, defined as follows. To each edge  $e \in \mathcal{E}$  (RIGC), we assign independent standard uniform random variables  $U_e$ , and for any  $\pi$ , define  $X_e^\pi := \mathbb{1}_{\{U_e \leq \pi\}}$ . The edges retained in  $\text{RIGC}(\pi)$ , that is, in  $\pi$ -percolation on RIGC, are exactly the edges  $e$  such that  $X_e^\pi = 1$ . Clearly, for  $\pi_1 < \pi_2$ , we have  $X_e^{\pi_1} \leq X_e^{\pi_2}$  for any edge  $e$ , thus  $\text{RIGC}(\pi_1)$  is a subgraph (edge-subgraph) of  $\text{RIGC}(\pi_2)$ . Denote by  $\mathcal{C}(v, \pi)$  the component of  $v$  in  $\text{RIGC}(\pi)$ . Suppose  $v \in \mathcal{C}_1(\pi_1)$ , then

$$|\mathcal{C}_1(\pi_1)| = |\mathcal{C}(v, \pi_1)| \leq |\mathcal{C}(v, \pi_2)| \leq |\mathcal{C}_1(\pi_2)|. \tag{4.10}$$

That is,  $|\mathcal{C}_1(\pi_1)| \leq |\mathcal{C}_1(\pi_2)|$  holds almost surely under this coupling, which implies the stochastic domination  $|\mathcal{C}_1(\pi_1)|/N_n \leq |\mathcal{C}_1(\pi_2)|/N_n$ , as required. This concludes the proof of Theorem 2.13 subject to Lemma 4.2. ■

### 4.3 | Percolated community list: Convergence and percolated community sizes

In this section, we provide the proof for Lemma 4.2 and Claim 4.3.

#### 4.3.1 | Convergence of the percolated community list

*Proof of Lemma 4.2.* We first prove (4.3). Recall that  $\mathbf{Com}(\pi) = (\text{Com}_a^\pi)_{a \in [M_n(\pi)]}$  (see Section 4.1). Recall (2.9) and for any possible community graph  $H \in \mathcal{H}$ , we introduce

$$\mathcal{V}_H^r(\pi) := \{a \in [M_n(\pi)] : \text{Com}_a^\pi \simeq H\}. \tag{4.11}$$

Our aim is to prove that the following quantity converges in probability to some constant:

$$\mu_H^{(n)}(\pi) = |\mathcal{V}_H^r(\pi)|/M_n(\pi) = \frac{|\mathcal{V}_H^r(\pi)|}{M_n} \cdot \frac{M_n}{M_n(\pi)}. \tag{4.12}$$

In the following, we prove convergence in probability to respective *constants* for both factors separately. This implies convergence in probability for the product, despite the dependence.

Recall that each new community in  $\mathbf{Com}(\pi)$  is a connected component under percolation on some original community in  $\mathbf{Com}$ . Thus, to count the frequencies and total number of new communities, we break it down with respect to the original communities, and first study percolation on an arbitrary community and the frequency of each outcome.

We introduce some notation. Recall that  $\mathcal{H}$  denotes the set of possible community graphs: simple, finite, connected graphs, with a fixed arbitrary labeling. For an arbitrary  $F \in \mathcal{H}$ , denote bond percolation on  $F$  by  $F(\pi)$ . We introduce an object to compare realizations of this random graph to: let  $\mathcal{G}$  denote the set of simple, finite, not necessarily connected, unlabeled graphs. Denote by  $\text{supp}(F(\pi))$  the set of all  $G \in \mathcal{G}$  that are isomorphic to some possible realization of  $F(\pi)$ , which is exactly the set of edge-subgraphs of  $F$ . Note that for a fixed  $F$ ,  $\text{supp}(F(\pi))$  is a finite set. Recall (2.9). For  $F \in \mathcal{H}$  and  $G \in \text{supp}(F(\pi))$ , define the random subset of original  $r$ -vertices with original community graph  $F$  that become isomorphic to  $G$  under percolation by

$$\mathcal{V}_{F(\pi) \simeq G}^r := \{a \in \mathcal{V}_F^r : \text{Com}_a(\pi) \simeq G\}. \tag{4.13}$$

Note that

$$\sum_{G \in \text{supp}(F(\pi))} |\mathcal{V}_{F(\pi) \simeq G}^r| = |\mathcal{V}_F^r|, \tag{4.14}$$

with  $\mathcal{V}_F^r$  defined as in (2.9). Since percolation on different communities is independent,  $(|\mathcal{V}_{F(\pi) \simeq G}^r|)_{G \in \text{supp}(F(\pi))}$  has a multinomial distribution with number of trials  $|\mathcal{V}_F^r|$  and probability vector  $(\mathbb{P}(F(\pi) \simeq G))_{G \in \text{supp}(F(\pi))}$ . Thus by multinomial concentration and  $|\mathcal{V}_F^r|/M_n \rightarrow \mu_F$  (Assumption 2.3(C)),

$$M_n^{-1} \cdot (|\mathcal{V}_{F(\pi) \simeq G}^r|)_{G \in \text{supp}(F(\pi))} \xrightarrow{\mathbb{P}} \mu_F \cdot (\mathbb{P}(F(\pi) \simeq G))_{G \in \text{supp}(F(\pi))}. \tag{4.15}$$

In the following, we study how one or more copies of  $H \in \mathcal{H}$  can be produced by percolating some  $F \in \mathcal{H}$ . We define the multiplicity of  $H$  in any  $G \in \mathcal{G}$ , denoted by  $\kappa(H | G) \geq 0$ , as the number of distinct connected components in  $G$  that are isomorphic to  $H$ . Note that there exists some (possibly more than

one)  $G \in \text{supp}(F(\pi))$  such that  $\kappa(H | G) \geq 1$  exactly when  $H$  is isomorphic to an edge-subgraph of  $F$ . We compute, for arbitrary  $H \in \mathcal{H}$ ,

$$M_n^{-1} |\mathcal{V}_H^r(\pi)| = \sum_{F \in \mathcal{H}} \sum_{G \in \text{supp}(F(\pi))} M_n^{-1} |\mathcal{V}_{F(\pi) \simeq G}^r| \cdot \kappa(H | G). \tag{4.16}$$

We claim that

$$M_n^{-1} |\mathcal{V}_H^r(\pi)| \stackrel{\mathbb{P}}{\rightarrow} \sum_{F \in \mathcal{H}} \sum_{G \in \text{supp}(F(\pi))} \mu_F \cdot \mathbb{P}(F(\pi) \simeq G) \cdot \kappa(H | G) < \infty. \tag{4.17}$$

For convenience, we denote the inner sums by

$$\begin{aligned} T_n^\pi(F, H) &:= \sum_{G \in \text{supp}(F(\pi))} M_n^{-1} |\mathcal{V}_{F(\pi) \simeq G}^r| \cdot \kappa(H | G), \\ T^\pi(F, H) &:= \sum_{G \in \text{supp}(F(\pi))} \mu_F \cdot \mathbb{P}(F(\pi) \simeq G) \cdot \kappa(H | G). \end{aligned} \tag{4.18}$$

Since  $\text{supp}(F(\pi))$  is a finite set, and  $\kappa(H | G)$  are constants, by (4.15) the linear combinations  $T_n^\pi(F, H) \stackrel{\mathbb{P}}{\rightarrow} T^\pi(F, H)$  as  $n \rightarrow \infty$ , for each  $H, F \in \mathcal{H}$  and  $\pi$ . Next, we use a truncation argument to prove the convergence of the infinite sum over  $F \in \mathcal{H}$ . The same argument also reveals the rhs of (4.17) to be finite.

Note that, as  $\kappa(H | G)$  counts *components* of  $G$  that are isomorphic to  $H$ ,  $\kappa(H | G) \leq |G|/|H| = |F|/|H|$ . Thus, for a fixed  $F$ , by (4.14) almost surely

$$T_n^\pi(F, H) \leq \sum_{G \in \text{supp}(F(\pi))} M_n^{-1} |\mathcal{V}_{F(\pi) \simeq G}^r| \cdot \frac{|F|}{|H|} = M_n^{-1} |\mathcal{V}_F^r| \cdot \frac{|F|}{|H|} = \frac{\mu_F^{(n)} |F|}{|H|}. \tag{4.19}$$

By Assumption 2.3 (D),

$$\sum_{F \in \mathcal{H}} \mu_F^{(n)} |F| = \mathbb{E}[D_n^r] \rightarrow \mathbb{E}[D^r] = \sum_{F \in \mathcal{H}} \mu_F |F| < \infty. \tag{4.20}$$

Thus for arbitrary  $\varepsilon > 0$ , there exist  $K = K(\varepsilon)$  and  $n_0 = n_0(\varepsilon)$  such that for all  $n \geq n_0$ ,

$$\sum_{F \in \mathcal{H}, |F| > K} \mu_F |F| < \varepsilon/6, \quad \sum_{F \in \mathcal{H}, |F| > K} \mu_F^{(n)} |F| < \varepsilon/3. \tag{4.21}$$

Combining (4.19)–(4.21), we obtain that for  $n \geq n_0$ , almost surely

$$0 \leq \sum_{\substack{F \in \mathcal{H} \\ |F| > K}} T_n^\pi(F, H) \leq \sum_{\substack{F \in \mathcal{H} \\ |F| > K}} \frac{\mu_F^{(n)} |F|}{|H|} < \varepsilon/3. \tag{4.22}$$

Analogously, using (4.20) and the identity  $\sum_{G \in \text{supp}(F(\pi))} \mathbb{P}(F(\pi) \simeq G) = 1$ ,

$$0 \leq \sum_{\substack{F \in \mathcal{H} \\ |F| > K}} T^\pi(F, H) \leq \sum_{\substack{F \in \mathcal{H} \\ |F| > K}} \sum_{G \in \text{supp}(F(\pi))} \mu_F \mathbb{P}(F(\pi) \simeq G) \frac{|F|}{|H|} = \sum_{\substack{F \in \mathcal{H} \\ |F| > K}} \mu_F \frac{|F|}{|H|} < \varepsilon/6. \tag{4.23}$$



Note that the number of  $F \in \mathcal{H}$  such that  $|F| \leq K$  is finite, thus by (4.15), it follows that the truncated sum converges in probability, that is, as  $n \rightarrow \infty$ ,

$$\mathbb{P} \left( \left| \sum_{\substack{F \in \mathcal{H} \\ |F| \leq K}} T_n^\pi(F, H) - \sum_{\substack{F \in \mathcal{H} \\ |F| \leq K}} T^\pi(F, H) \right| < \varepsilon/2 \right) \rightarrow 1. \tag{4.24}$$

Combining (4.22)–(4.24) yields the convergence in probability claimed in (4.17). We see that the rhs of (4.17) can be written as the sum of a finite sum  $\sum_{F \in \mathcal{H}, |F| \leq K} T^\pi(F, H)$  in (4.24) and a bounded quantity  $\sum_{F \in \mathcal{H}, |F| > K} T^\pi(F, H)$  in (4.23) and is thus finite.

Next, we study  $M_n(\pi)/M_n$ , that is the reciprocal of the second factor in (4.12). Recall that for  $G \in \mathcal{G}$ , the number of connected components in  $G$  is denoted by  $c(G)$ . Recall (4.13) and compute

$$M_n(\pi)/M_n = \sum_{F \in \mathcal{H}} \sum_{G \in \text{supp}(F(\pi))} M_n^{-1} |\mathcal{V}_{F(\pi) \simeq G}^r| \cdot c(G). \tag{4.25}$$

We note the similarity between this formula and (4.16), as well as  $c(G) \leq |G| = |F|$ . Thus, with analogous arguments and the same truncation as above, we conclude that

$$M_n(\pi)/M_n \xrightarrow{\mathbb{P}} \sum_{F \in \mathcal{H}} \sum_{G \in \text{supp}(F(\pi))} \mu_F \cdot \mathbb{P}(F(\pi) \simeq G) \cdot c(G) < \infty. \tag{4.26}$$

Combining (4.12), (4.16), and (4.26) yields

$$\begin{aligned} \mu_H^{(n)}(\pi) &= \frac{|\mathcal{V}_H^r(\pi)|}{M_n} \Big/ \frac{M_n(\pi)}{M_n} \\ &\xrightarrow{\mathbb{P}} \frac{\sum_{F \in \mathcal{H}} \sum_{G \in \text{supp}(F(\pi))} \mu_F \cdot \mathbb{P}(F(\pi) \simeq G) \cdot \kappa(H | G)}{\sum_{F \in \mathcal{H}} \sum_{G \in \text{supp}(F(\pi))} \mu_F \cdot \mathbb{P}(F(\pi) \simeq G) \cdot c(G)} =: \mu_H(\pi). \end{aligned} \tag{4.27}$$

This concludes the proof of (4.3). Note that community sizes cannot increase under percolation, thus  $D_n^r(\pi) \leq D_n^r$ , and  $(D_n^r)_{n \in \mathbb{N}}$  is tight (since it is also UI). Consequently, by  $\mathbb{P}(D_n^r(\pi) \geq K) \leq \mathbb{P}(D_n^r \geq K)$ ,  $(D_n^r(\pi))_{n \in \mathbb{N}}$  is also tight, thus  $\sum_{H \in \mathcal{H}} \mu_H^{(n)}(\pi) = 1$  for each  $n \in \mathbb{N}$  implies that  $\sum_{H \in \mathcal{H}} \mu_H(\pi) = 1$ .

Next, we prove (4.4). By definition, we compute

$$\mathbb{E} \left[ D_n^r(\pi) \mid \mathbf{Com}(\pi) \right] = \frac{1}{M_n(\pi)} \sum_{a \in [M_n(\pi)]} |\text{Com}_a^\pi| = \frac{|\mathcal{V}(\mathbf{Com}(\pi))|}{M_n(\pi)}. \tag{4.28}$$

Recall that, by the definition of  $\mathbf{Com}(\pi)$ ,  $|\mathcal{V}(\mathbf{Com}(\pi))| = |\mathcal{V}(\mathbf{Com})| = \ell_n$ . Thus,

$$\begin{aligned} \mathbb{E} \left[ D_n^r(\pi) \mid \mathbf{Com}(\pi) \right] &= \frac{\ell_n}{M_n(\pi)} = \frac{\ell_n}{M_n} \cdot \frac{M_n}{M_n(\pi)} = \mathbb{E}[D_n^r] \cdot \frac{M_n}{M_n(\pi)} \\ &\xrightarrow{\mathbb{P}} \frac{\mathbb{E}[D_n^r]}{\sum_{F \in \mathcal{H}} \sum_{G \in \text{supp}(F(\pi))} \mu_F \cdot \mathbb{P}(F(\pi) \simeq G) \cdot c(G)} =: C < \infty, \end{aligned} \tag{4.29}$$

by Assumption 2.3 (D) and (4.26). From (4.3), it follows that  $D_n^r(\pi) \xrightarrow{d} D^r(\pi)$  as  $n \rightarrow \infty$ . Further, the stochastic domination  $0 \leq D_n^r(\pi) \leq D_n^r$  also holds conditionally on  $\mathbf{Com}(\pi)$ . By Assumption 2.3 (D),  $(D_n^r)_{n \in \mathbb{N}}$  is UI. For any fixed  $K$ , by stochastic domination,  $\mathbb{E} \left[ D_n^r(\pi) \mathbb{1}_{\{D_n^r(\pi) > K\}} \right] \leq$

$\mathbb{E} [D_n^r \mathbb{1}_{\{D_n^r > K\}}]$ , thus  $(D_n^r(\pi))_{n \in \mathbb{N}}$  is also UI, and consequently  $\mathbb{E}[D_n^r(\pi)] \rightarrow \mathbb{E}[D^r(\pi)] < \infty$ . Noting that  $\mathbb{E} [D_n^r(\pi) \mid \mathbf{Com}(\pi)] \leq \mathbb{E}[D_n^r]$  by stochastic domination, and  $(\mathbb{E}[D_n^r])_{n \in \mathbb{N}}$  is bounded,  $\mathbb{E} [D_n^r(\pi) \mid \mathbf{Com}(\pi)] \xrightarrow{\mathbb{P}} C$  implies  $\mathbb{E} [\mathbb{E}[D_n^r(\pi) \mid \mathbf{Com}(\pi)]] = \mathbb{E}[D_n^r(\pi)] \rightarrow C$ . Since the limit is unique, we obtain that necessarily  $C = \mathbb{E}[D^r(\pi)]$ . This concludes the proof of (4.4), and consequently the proof of Lemma 4.2.  $\blacksquare$

4.3.2 | Representation of size-biased community sizes

*Proof of Claim 4.3.* Recall that for  $j \in \mathcal{V}(\text{Com}_a)$ ,  $\mathcal{C}^\circ(j, \pi)$  denotes its connected component under  $\pi$ -percolation on  $\text{Com}_a$ . Recall the percolated community list  $\mathbf{Com}(\pi) = (\text{Com}_a^\pi)_{a \in [M_n(\pi)]}$ . Using the definition of the empirical distribution  $D_n^r(\pi)$  and its transform by (1.1), we compute the empirical mass function of  $\widetilde{D_n^r(\pi)}$

$$\begin{aligned} \mathbb{P} \left( \widetilde{D_n^r(\pi)} = k \mid \mathbf{Com}(\pi) \right) &= \frac{(k+1) \mathbb{P} \left( D_n^r(\pi) = k+1 \mid \mathbf{Com}(\pi) \right)}{\mathbb{E} \left[ D_n^r(\pi) \mid \mathbf{Com}(\pi) \right]} \\ &= \frac{(k+1) \frac{1}{M_n(\pi)} \sum_{a \in [M_n(\pi)]} \mathbb{1}_{\{|\text{Com}_a^\pi| = k+1\}}}{1 M_n(\pi) \sum_{a \in [M_n(\pi)]} |\text{Com}_a^\pi|}. \end{aligned} \tag{4.30}$$

Note that in the numerator, we can replace  $(k+1) \cdot \mathbb{1}_{\{|\text{Com}_a^\pi| = k+1\}} = |\text{Com}_a^\pi| \cdot \mathbb{1}_{\{|\text{Com}_a^\pi| = k+1\}}$ . Also note that, by construction of  $\mathbf{Com}(\pi)$  (see Section 4.1), the sum in the denominator equals  $\sum_{a \in [M_n(\pi)]} |\text{Com}_a^\pi| = |\mathcal{V}(\mathbf{Com})| = \sum_{a \in [M_n]} |\text{Com}_a| = \ell_n$ , with  $\ell_n$  from (2.1). Thus,

$$\begin{aligned} \mathbb{P} \left( \widetilde{D_n^r(\pi)} = k \mid \mathbf{Com}(\pi) \right) &= \frac{\sum_{a \in [M_n(\pi)]} |\text{Com}_a^\pi| \cdot \mathbb{1}_{\{|\text{Com}_a^\pi| = k+1\}}}{\ell_n} \\ &= \frac{1}{\ell_n} \sum_{a \in [M_n(\pi)]} \sum_{j \in \mathcal{V}(\text{Com}_a^\pi)} \mathbb{1}_{\{|\text{Com}_a^\pi| = k+1\}} = \frac{1}{\ell_n} \sum_{a \in [M_n(\pi)]} \sum_{j \in \mathcal{V}(\text{Com}_a^\pi)} \mathbb{1}_{\{|\mathcal{C}^\circ(j, \pi)| = k+1\}}, \end{aligned} \tag{4.31}$$

where we have used that for  $j \in \mathcal{V}(\text{Com}_a^\pi)$ , its percolated component is exactly  $\text{Com}_a^\pi$ , thus  $|\mathcal{C}^\circ(j, \pi)| = |\text{Com}_a^\pi|$ . Once again by  $\cup_{a \in [M_n(\pi)]} \mathcal{V}(\text{Com}_a^\pi) = \mathcal{V}(\mathbf{Com})$ ,

$$\begin{aligned} \mathbb{P} \left( \widetilde{D_n^r(\pi)} = k \mid \mathbf{Com}(\pi) \right) &= \frac{1}{\ell_n} \sum_{j \in \mathcal{V}(\mathbf{Com})} \mathbb{1}_{\{|\mathcal{C}^\circ(j, \pi)| = k+1\}} \\ &= \mathbb{P} \left( |\mathcal{C}^\circ(J_n, \pi)| = k+1 \mid \mathbf{Com}(\pi) \right) = \mathbb{P} \left( |\mathcal{C}^\circ(J_n, \pi)| - 1 = k \mid \mathbf{Com}(\pi) \right). \end{aligned} \tag{4.32}$$

It follows that  $\mathbb{P} \left( \widetilde{D_n^r(\pi)} = k \right) = \mathbb{P} (|\mathcal{C}^\circ(J_n, \pi)| - 1 = k)$  for all  $k$ , which implies (4.5). This concludes the proof of Claim 4.3.  $\blacksquare$

4.4 | Proof of Proposition 2.14

First, we prove (2.39). Recall that in (4.9), we have identified  $\pi_c = \inf \{ \pi : \mathbb{E}[\widetilde{D}^\circ] \mathbb{E}[\widetilde{D^r(\pi)}] > 1 \}$ . Thus, it is sufficient to show that

$$\mathbb{E} \left[ \widetilde{D^r(\pi)} \right] = \mathbb{E} \left[ |\mathcal{H}| (|\mathcal{C}^\mathcal{H}(U_\mathcal{H}, \pi)| - 1) \right] / \mathbb{E}[D^r], \tag{4.33}$$

where  $\mathcal{H}$  is a random graph with pmf  $\mu$ ,  $U_{\mathcal{H}}|\mathcal{H} \sim \text{Unif}[\mathcal{H}]$ , and  $\mathbb{E}[\cdot]$  on the right hand side denotes total expectation (wrt the joint measure of  $\mathcal{H}$ ,  $U_{\mathcal{H}}$  and the percolation).

To express its expectation, we analyze the distribution of  $\widetilde{D^r(\pi)}$ , using that  $\widetilde{D_n^r(\pi)} \xrightarrow{d} \widetilde{D^r(\pi)}$  by Lemma 4.2 and  $|\mathcal{C}^e(J_n, \pi)| - 1 \stackrel{d}{=} \widetilde{D_n^r(\pi)}$  by Claim 4.3. As before,  $\mathcal{C}^e(j, \pi)$  denotes the component of  $j \in \text{Com}_a$  under  $\pi$ -percolation on  $\text{Com}_a$ , and  $J_n \sim \text{Unif}[\mathcal{V}(\mathbf{Com})]$ , where  $\mathcal{V}(\mathbf{Com})$  denotes the disjoint union of all vertices in community graphs.

Note that  $J_n$  is chosen uniformly at random among all community roles, which is equivalent to choosing a community in a *size-biased* fashion, then choosing a uniform member of the chosen community. In the following, we use this observation to analyze the distribution of  $|\mathcal{C}^e(J_n, \pi)|$ . Recall from Section 2.1 that all isomorphic community graphs are labeled in the same way and community roles in  $H$  are described by a label  $l \in [|H|]$ . Intuitively, for the distribution of  $\mathcal{C}^e(J_n, \pi)$ , only the community graph and label of  $J_n$  matters, thus we introduce the concept of *type* to represent this pair. Recall that a community role  $j \in \mathcal{V}(\mathbf{Com})$  that is in  $\text{Com}_a$  and has label  $l$  can be represented by the pair  $(a, l)$ , and define its type as  $\text{Type}(j) := (\text{Com}_a, l)$ . Using that  $J_n \sim \text{Unif}[\mathcal{V}(\mathbf{Com})]$ , we compute the distribution of its random type  $\text{Type}(J_n) =: (\mathcal{H}_n^*, I_n)$ . Using (2.10) and that  $|\mathcal{V}(\mathbf{Com})| = \ell_n = M_n \mathbb{E}[D_n^r]$  by Remark 2.4(i),

$$\mathbb{P}(\mathcal{H}_n^* = H) = \mathbb{P}(J_n \in \cup_{a \in \mathcal{V}_H^r} \mathcal{V}(\text{Com}_a)) = \frac{|\mathcal{V}_H^r| \cdot |H|}{|\mathcal{V}(\mathbf{Com})|} = \frac{|\mathcal{V}_H^r| \cdot |H|}{M_n \mathbb{E}[D_n^r]} = \frac{\mu_H^{(n)} \cdot |H|}{\mathbb{E}[D_n^r]} =: \mu_H^{(n),*}. \tag{4.34}$$

Indeed, as intuition suggests,  $\mathcal{H}_n^*$  is chosen in a size-biased fashion. It is also intuitive that conditionally on  $\mathcal{H}_n^*, I_n$  is uniform on  $[|\mathcal{H}_n^*|]$ . Noting that in each community graph  $\text{Com}_a \simeq H$  there is one vertex with label  $l$  and  $|H|$  vertices in total, we indeed obtain

$$\mathbb{P}(I_n = l \mid \mathcal{H}_n^* = H) = \frac{|\mathcal{V}_H^r| \cdot 1}{|\mathcal{V}_H^r| \cdot |H|} = \frac{1}{|H|}. \tag{4.35}$$

By Assumption 2.3(C) and (D), the joint mass function converges:

$$\mathbb{P}((\mathcal{H}_n^*, I_n) = (H, l)) = \frac{\mu_H^{(n)} \cdot |H|}{\mathbb{E}[D_n^r]} \frac{1}{|H|} \rightarrow \frac{\mu_H \cdot |H|}{\mathbb{E}[D^r]} \frac{1}{|H|} =: \mathbb{P}((\mathcal{H}^*, I) = (H, l)), \tag{4.36}$$

so that  $(\mathcal{H}_n^*, I_n) \xrightarrow{d} (\mathcal{H}^*, I)$ . We now return to studying  $|\mathcal{C}^e(J_n, \pi)|$ . Let us we write  $\mathcal{C}^e(j, \pi)$  in terms of  $\text{Type}(j) = (H, l)$  as  $\mathcal{C}^H(l, \pi) \stackrel{d}{=} \mathcal{C}^e(j, \pi)$  (see Proposition 2.14), and consider  $\mathcal{C}^e(J_n, \pi) \stackrel{d}{=} \mathcal{C}^{\mathcal{H}_n^*}(I_n, \pi)$  as a *mixture*. We have shown that the mixing variable  $(\mathcal{H}_n^*, I_n) \xrightarrow{d} (\mathcal{H}^*, I)$ , and since it has countably many values, this implies convergence in total variation distance. This implies, since the pointwise difference of mass functions of the mixtures can be bounded in terms of the total variation distance of the mixing variables, that the mixture also converges in distribution:  $|\mathcal{C}^e(J_n, \pi)| \stackrel{d}{=} |\mathcal{C}^{\mathcal{H}_n^*}(I_n, \pi)| \xrightarrow{d} |\mathcal{C}^{\mathcal{H}^*}(I, \pi)|$ . Recall that  $\widetilde{D_n^r(\pi)} + 1 \stackrel{d}{=} |\mathcal{C}^e(J_n, \pi)|$  by Claim 4.3 and  $\widetilde{D_n^r(\pi)} \xrightarrow{d} \widetilde{D^r(\pi)}$  by Lemma 4.2. Necessarily, by the uniqueness of limit,  $\widetilde{D^r(\pi)} + 1 \stackrel{d}{=} |\mathcal{C}^{\mathcal{H}^*}(I, \pi)|$ . We compute

$$\mathbb{E}[\widetilde{D^r(\pi)}] = \mathbb{E}[|\mathcal{C}^{\mathcal{H}^*}(I, \pi)| - 1] = \mathbb{E}\left[\mathbb{E}\left[|\mathcal{C}^{\mathcal{H}^*}(I, \pi)| - 1 \mid (\mathcal{H}^*, I)\right]\right]$$

$$\begin{aligned}
 &= \sum_{H \in \mathcal{H}} \sum_{i \in [|H|]} \mathbb{E} [|\mathcal{C}^H(i, \pi)| - 1] \frac{1}{|H|} \frac{\mu_H |H|}{\mathbb{E}[D^r]} \\
 &= \sum_{H \in \mathcal{H}} \mu_H |H| \left( \frac{1}{|H|} \sum_{i \in [|H|]} \mathbb{E} [|\mathcal{C}^H(i, \pi)| - 1] \right) / \mathbb{E}[D^r].
 \end{aligned} \tag{4.37}$$

Recall that  $\mathcal{H}$  denotes a random graph with pmf  $\mu$  and  $U_H | \mathcal{H} \sim \text{Unif}[\mathcal{H}]$ ; in particular,  $U_H \sim \text{Unif}[H]$  for fixed  $H \in \mathcal{H}$ . Then

$$\begin{aligned}
 \mathbb{E}[\widetilde{D^r(\pi)}] &= \sum_{H \in \mathcal{H}} \mu_H |H| \cdot \mathbb{E} [|\mathcal{C}^H(U_H, \pi)| - 1] / \mathbb{E}[D^r] \\
 &= \mathbb{E} \left[ |\mathcal{H}| \cdot \mathbb{E} \left[ |\mathcal{C}^H(U_H, \pi)| - 1 \mid \mathcal{H} \right] \right] / \mathbb{E}[D^r],
 \end{aligned} \tag{4.38}$$

which is equivalent to (4.33) by the tower property of conditional expectation. This concludes the proof of (2.39).

Next, we prove that  $\pi_c < 1$  by showing that there exists some  $\pi < 1$  such that  $\mathbb{E}[\widetilde{D^r(\pi)}] > 1/\mathbb{E}[\widetilde{D^\ell}]$ , which is equivalent to (4.8). Using (4.37), for some  $K \in \mathbb{Z}^+$  to be specified later, we bound

$$\mathbb{E}[\widetilde{D^r(\pi)}] \geq \sum_{\substack{H \in \mathcal{H} \\ |H| \leq K}} \frac{\mu_H}{\mathbb{E}[D^r]} \sum_{i \in [|H|]} \mathbb{E} [|\mathcal{C}^H(i, \pi)| - 1] =: S_{\leq K}(\pi). \tag{4.39}$$

We show that, for appropriately chosen  $K$  and  $\pi$ ,  $S_{\leq K}(\pi) > 1/\mathbb{E}[\widetilde{D^\ell}]$  by comparing both to  $S_{\leq K}(1)$  as an intermediate step. In fact, noting that  $\mathcal{C}^H(i, 1)$  is the unpercolated component of  $i$ , that is, the entire graph  $H$ , we have

$$\begin{aligned}
 S_{\leq K}(1) &= \sum_{\substack{H \in \mathcal{H} \\ |H| \leq K}} \frac{\mu_H}{\mathbb{E}[D^r]} \sum_{i \in [|H|]} \mathbb{E} [|\mathcal{C}^H(i, 1)| - 1] = \sum_{\substack{H \in \mathcal{H} \\ |H| \leq K}} \frac{\mu_H}{\mathbb{E}[D^r]} \sum_{i \in [|H|]} (|H| - 1) \\
 &= \sum_{\substack{H \in \mathcal{H} \\ |H| \leq K}} \frac{\mu_H}{\mathbb{E}[D^r]} |H| (|H| - 1) = \frac{\mathbb{E} [|\mathcal{H}| (|\mathcal{H}| - 1) \mathbb{1}_{\{|\mathcal{H}| \leq K\}}]}{\mathbb{E}[D^r]},
 \end{aligned} \tag{4.40}$$

where  $\mathcal{H}$  denotes a random graph with pmf  $\mu$ . Then, by Assumption 2.3(C1) and (1.1),

$$S_{\leq K}(1) = \frac{\mathbb{E} [D^r (D^r - 1) \mathbb{1}_{\{D^r \leq K\}}]}{\mathbb{E}[D^r]} = \mathbb{E} \left[ \widetilde{D^r} \cdot \mathbb{1}_{\{\widetilde{D^r} < K\}} \right]. \tag{4.41}$$

Clearly, the partial sums  $S_{\leq K}(1) \rightarrow \mathbb{E}[\widetilde{D^r}]$  as  $K \rightarrow \infty$ . By the supercriticality condition (2.20), we know that  $\mathbb{E}[\widetilde{D^r}] > 1/\mathbb{E}[\widetilde{D^\ell}]$ , thus there exists  $K$  large enough so that  $S_{\leq K}(1) > 1/\mathbb{E}[\widetilde{D^\ell}]$ . We fix such a  $K$ . Next, we compare  $S_{\leq K}(\pi)$  and  $S_{\leq K}(1)$ . Note that  $S_{\leq K}(\pi)$  is a *finite* sum, as  $\{H \in \mathcal{H} : |H| \leq K\}$  is a finite set. Further, for any fixed  $(H, i)$ ,  $\mathbb{E} [|\mathcal{C}^H(i, \pi)|]$  is a polynomial in  $\pi$ , thus it is *continuous*. Consequently,  $\pi \mapsto S_{\leq K}(\pi)$  is continuous, and by  $S_{\leq K}(1) > 1/\mathbb{E}[\widetilde{D^\ell}]$ , we can choose  $\pi < 1$  sufficiently close to 1 so that  $S_{\leq K}(\pi) > 1/\mathbb{E}[\widetilde{D^\ell}]$ . We have thus shown that for some  $K \in \mathbb{Z}^+$  and  $\pi < 1$ , as chosen above,

$$\mathbb{E}[\widetilde{D^r(\pi)}] \geq S_{\leq K}(\pi) > 1/\mathbb{E}[\widetilde{D^\ell}], \tag{4.42}$$

and we conclude that indeed  $\pi_c < 1$ . This concludes the proof of Proposition 2.14. ■

## 5 | FURTHER PROPERTIES OF THE GIANT OF THE RIGC

In this section, we explore the relation between “local” and “global” properties of the RIGC and the (underlying) BCM. In [48], we have shown that neighborhoods in the BCM and RIGC are, respectively, well-approximated by a branching process (that we recall shortly) and an appropriate transform of it (see [48, Sections 3.2 and 3.3, respectively]). In Section 3, we have analyzed the largest component of each model using the continuous-time exploration introduced in Section 3.1. As for the configuration model [38, 39] or the Erdős-Rényi random graph [22], we show below that also in the BCM and RIGC, the proportion of the giant component is given by the survival probability of the branching process approximating the neighborhoods. However, it is not entirely trivial to prove results on the largest component using *solely* the branching process approximation, which is why we turn to the continuous-time exploration for an independent proof, and make the connection afterwards. This connection then allows us to derive “local” properties of the giant component. As a demonstration of this more generally applicable technique, we present the degree distribution within the giant and as a (not entirely trivial) consequence, the first order asymptotics of the number of edges in the giant. (The difficulty lies in identifying when the obtained degree distribution has finite or infinite mean.)

### 5.1 | The approximating BPs

In this section, we describe the branching processes (see, e.g., [2] for an introduction on branching processes) related to the local weak limit of the BCM and the RIGC (intuitively speaking, the local weak limit describes the neighborhood of a typical vertex; for a precise definition, see [48, Definition 2.6 and Section 3.1]). Further, we show how these branching processes (BPs) are related to the constants  $\eta_\ell$ ,  $\eta_r$ ,  $\xi_\ell$ , and  $\xi_r$  (defined in Theorems 2.6 and 2.7 and Corollary 2.12) describing the size and properties of the giant component of the BCM and RIGC.

#### 5.1.1 | Description of the approximating BPs

We first describe the discrete-time branching process  $\text{BP}_\ell$  defined in [48, Section 3.2] that approximates the neighborhood of  $\ell$ -vertices in the underlying BCM. Recall  $D^\ell$  and  $D^r$  from Assumption 2.3(A) and (C1), respectively, and further recall (1.1). We start with a single root in generation 0 that produces offspring distributed as  $D^\ell$ . Every other individual in an even generation has offspring distributed as  $\tilde{D}^\ell$ , while every individual in an odd generation has offspring distributed as  $\tilde{D}^r$ . The offspring of any two individuals are independent. In [48, Section 3.2], the branching process  $\text{BP}_r$  that approximates the neighborhood of  $r$ -vertices in the underlying BCM is defined analogously by reversing the roles of  $\ell$  and  $r$ .

In the following, we show the relation between the supercriticality of the BCM and the supercriticality of the approximating BPs. Just like in Theorems 2.6 and 2.11, we exclude the special case  $\mathbb{P}(D^\ell = 2) = \mathbb{P}(D^r = 2) = 1$ , or equivalently,  $\mathbb{P}(\tilde{D}^\ell = 1) = \mathbb{P}(\tilde{D}^r = 1) = 1$ , for the following reason. We say that an individual in the process survives if it produces an infinite (sub)tree, and we say that the BP survives if the root as an individual survives. The above special case guarantees exactly one offspring for both  $\text{BP}_\ell$  and  $\text{BP}_r$  in each generation (except the root that has offspring 2), and thus both BPs are guaranteed infinite survival in an *atypical* way. Thus we exclude this special case from now on.

Next, we show that  $\text{BP}_\ell$  and  $\text{BP}_r$  are supercritical exactly when the supercriticality condition (2.20) of the graph holds; and the survival probabilities are  $\xi_\ell$  (from Theorem 2.6) and  $\xi_r$  (from Corollary 2.12), respectively. We prove the statement for  $\text{BP}_\ell$  first. Consider the subprocess  $\widehat{\text{BP}}_\ell$  formed by the descendants in odd generations of the first child of the root, so that offspring in

$\widehat{\text{BP}}_\ell$  is defined as grandchildren in  $\text{BP}_\ell$ . Consequently,  $\widehat{\text{BP}}_\ell$  is a Galton–Watson process, with offspring distribution  $N^r \stackrel{d}{=} \sum_{i=1}^{\tilde{D}^r} \tilde{D}_{(i)}^\ell$ , where  $\tilde{D}_{(i)}^\ell$  are iid copies of  $\tilde{D}^\ell$  and are independent from  $\tilde{D}^r$ . Note that  $N^r$  has been introduced in (3.49) to make sense of the *composite generating function*  $G_{N^r}(z) = G_{\tilde{D}^r}(G_{\tilde{D}^\ell}(z))$ . We know from branching process literature (see, e.g., [2]) that the extinction probability of  $\widehat{\text{BP}}_\ell$  is the smallest fixed point of  $G_{N^r}$ , defined as  $\eta_\ell$  in (2.18). Further,  $\eta_\ell < 1$  exactly when  $\mathbb{E}[N^r] = \mathbb{E}[\tilde{D}^r]\mathbb{E}[\tilde{D}^\ell] > 1$  (by Wald’s identity), that is, when (2.20) holds. Each child of the root in  $\text{BP}_\ell$  survives exactly when their corresponding  $\widehat{\text{BP}}_\ell$  survives. Using that the root of  $\text{BP}_\ell$  produces offspring distributed as  $D^\ell$ , the survival probability of  $\text{BP}_\ell$  is  $1 - G_{D^\ell}(\eta_\ell)$ , defined as  $\xi_\ell$  in Theorem 2.6. By properties of  $G_{D^\ell}$ ,  $\xi_\ell > 0$  exactly when  $\eta_\ell < 1$ , that is, under the supercriticality condition (2.20).

5.1.2 | Left–right correspondence

We define the analogous Galton–Watson process  $\widehat{\text{BP}}_r$  (consisting of odd generations in  $\text{BP}_r$ ) with offspring distribution  $N^\ell \stackrel{d}{=} \sum_{i=1}^{\tilde{D}^\ell} \tilde{D}_{(i)}^r$ , where  $\tilde{D}_{(i)}^r$  are iid random variables with distribution  $\tilde{D}^r$  and independent from  $\tilde{D}^\ell$ . This process is supercritical when  $\mathbb{E}[N^\ell] = \mathbb{E}[\tilde{D}^r]\mathbb{E}[\tilde{D}^\ell] > 1$ , where we again recognize the supercriticality condition (2.20). Thus the extinction probability of  $\widehat{\text{BP}}_r$  is the smallest solution to the fixed point equation

$$G_{\tilde{D}^\ell}(G_{\tilde{D}^r}(z)) = z. \tag{5.1}$$

We now verify that the smallest solution is  $\eta_r := G_{\tilde{D}^\ell}(\eta_\ell)$  (defined in Theorem 2.7). Applying  $G_{\tilde{D}^\ell}$  to both sides of (2.18) implies that  $\eta_r$  indeed satisfies (5.1). Further, when (2.20) holds, we have that  $\eta_\ell < 1$ , which implies  $\eta_r < 1$ . Analogously with  $\xi_\ell$ ,  $\xi_r = 1 - G_{D^r}(\eta_r)$  is indeed the survival probability of  $\text{BP}_r$ , and  $\xi_r > 0$  exactly when  $\eta_r < 1$ , that is, under the supercriticality condition (2.20). We also note that applying  $G_{\tilde{D}^r}$  to both sides of (5.1) yields  $\eta_\ell = G_{\tilde{D}^r}(\eta_r)$ , showing that the roles of lhs and rhs are indeed symmetric, however the corresponding quantities are generally *not equal*.

5.2 | The relation of local and global properties and degrees in the giant component

Thinking of  $\xi_\ell$  in (2.32) and  $\xi_r$  in (2.35) as the probability for an  $\ell$ -vertex and  $r$ -vertex, respectively, to be in the giant, we find that the same quantities equal the probabilities of survival of the respective branching processes  $\text{BP}_\ell$  and  $\text{BP}_r$ . Thus, we have established the “asymptotic equivalence,” as formalized below, of the following events: a vertex being in the giant and the survival of the corresponding branching process.

For two sets  $A$  and  $B$ , let  $A \triangle B := (A \setminus B) \cup (B \setminus A)$  denote their symmetric difference. For  $v \in \mathcal{V}^\ell$ , denote by  $\mathcal{C}(v)$  the connected component of  $v$  in the RIGC. For  $K \in \mathbb{Z}^+$ , introduce the set

$$\mathcal{X}_{\geq K} := \{v \in \mathcal{V}^\ell : |\mathcal{C}(v)| \geq K\}. \tag{5.2}$$

**Lemma 5.1** (Relation of the BP-approximation and the giant component). *Consider the RIGC( $\mathbf{d}^\ell, \mathbf{Com}$ ) under Assumption 2.3. Then, for any  $\varepsilon > 0$ ,*

$$\lim_{K \rightarrow \infty} \lim_{n \rightarrow \infty} \mathbb{P}(N_n^{-1} |\mathcal{X}_{\geq K} \triangle \mathcal{C}_1| > \varepsilon) = 0. \tag{5.3}$$

This notion is closely related to convergence in probability.

*Proof.* Note that, for any  $K$  fixed, if  $|\mathcal{X}_{\geq K} \triangle \mathcal{C}_1| > \varepsilon N_n$ , then  $|\mathcal{X}_{\geq K}| \geq 1$ , so that  $\mathcal{C}_1 \subseteq \mathcal{X}_{\geq K}$ . Thus  $|\mathcal{X}_{\geq K} \triangle \mathcal{C}_1| = |\mathcal{X}_{\geq K} \setminus \mathcal{C}_1| = |\mathcal{X}_{\geq K}| - |\mathcal{C}_1|$  whp, and (5.3) is equivalent to

$$\lim_{K \rightarrow \infty} \lim_{n \rightarrow \infty} \mathbb{P} \left( N_n^{-1} |\mathcal{X}_{\geq K}| - N_n^{-1} |\mathcal{C}_1| > \varepsilon \right) = 0. \tag{5.4}$$

By Theorem 2.6, as  $n \rightarrow \infty$ , independently of  $K$

$$\mathbb{P} \left( \left| N_n^{-1} |\mathcal{C}_1| - \xi_\ell \right| > \varepsilon/2 \right) \rightarrow 0. \tag{5.5}$$

We denote by  $(\text{CP}, o)$ , introduced in detail in [48, Section 3.3], the local weak limit of the RIGC, that is, the random graph CP with root  $o$  that approximates neighborhoods in the RIGC. By the triangle inequality,

$$\lim_{K \rightarrow \infty} \lim_{n \rightarrow \infty} \mathbb{P} \left( \left| N_n^{-1} |\mathcal{X}_{\geq K}| - \xi_\ell \right| > \varepsilon/2 \right) \leq \lim_{K \rightarrow \infty} \lim_{n \rightarrow \infty} \mathbb{P} \left( \left| N_n^{-1} |\mathcal{X}_{\geq K}| - \mathbb{P}(|\text{CP}| \geq K) \right| > \varepsilon/4 \right) \tag{5.6a}$$

$$+ \lim_{K \rightarrow \infty} \lim_{n \rightarrow \infty} \mathbb{P} \left( \left| \mathbb{P}(|\text{CP}| \geq K) - \xi_\ell \right| > \varepsilon/4 \right). \tag{5.6b}$$

By the local weak convergence stated in [48, Theorem 2.7], the inner limit in (5.6a) equals 0 for any fixed  $K$ , thus (5.6a) equals 0. Note that the limit in  $n$  in (5.6b) can be omitted, and also note that the survival probability is  $\xi_\ell = \mathbb{P}(|\text{BP}_\ell| = \infty)$ . From the construction of CP in [48, Section 3.3] using  $\text{BP}_\ell$ , we know that  $\mathbb{P}(|\text{CP}| = \infty) = \mathbb{P}(|\text{BP}_\ell| = \infty)$ . As  $\mathbb{P}(|\text{CP}| \geq K) \rightarrow \mathbb{P}(|\text{CP}| = \infty) = \xi_\ell$ , (5.6b) equals 0 as well. Combining this with (5.4)–(5.5) through the triangle inequality yields (5.3), concluding the proof of Lemma 5.1. ■

The above equivalence can be extended to give a heuristic interpretation for further results, namely, the formulas in Theorem 2.11 and Corollary 2.12. Knowing the extinction probability of a child of the root and the degree of the root results in (2.33) and (2.36). For (2.34), note that choosing an (instance of an) edge uar *in the random graph* is equivalent to picking the comprising  $\ell$ - and  $r$ -half-edges uar independently. Then the two endpoints can be viewed as a child of the root in  $\text{BP}_\ell$  and  $\text{BP}_r$ , respectively, and at least one of them has to survive, which has probability  $1 - \eta_\ell \eta_r$ . In the following, we prove our results on degrees and edges in the giant as a consequence of local weak convergence and Theorem 2.6.

*Proof of Theorem 2.7.* As we recalled above from [48, Section 3.3], we denote by  $(\text{CP}, o)$  the local weak limit of the RIGC, that is, the random graph CP with root  $o$  that approximates neighborhoods in the RIGC. Denote by  $\text{deg}(o)$  the degree of  $o$  in CP, and by  $\beta\text{-deg}(o)$  the number of communities that  $o$  is part of in CP. Let  $V_n^\ell \sim \text{Unif}[\mathcal{V}^\ell]$  and note that

$$N_n^{-1} |\mathcal{V}_k^\ell \cap \mathcal{V}_d^r \cap \mathcal{C}_1| = \mathbb{P} \left( p\text{-deg}(V_n^\ell) = d, \ell\text{-deg}(V_n^\ell) = k, V_n^\ell \in \mathcal{C}_1 \right). \tag{5.7}$$

Intuitively, as  $(\text{CP}, o)$  approximates  $(\text{RIGC}, V_n^\ell)$ , the limit of the above quantity must be  $\mathbb{P}(\beta\text{-deg}(o) = k, \text{deg}(o) = d, |\mathcal{V}(\text{CP})| = \infty)$ . In the following, we prove this formally. Recall (2.22) and that  $\nu(c | H)$  denotes the number of vertices in  $H$  with  $c$ -degree  $c$ . From the construction of  $(\text{CP}, 0)$  in [48, Section 3.3], it is straightforward to see (by conditioning on the community graphs  $H_1, \dots, H_k$  assigned to the  $k$  communities that the root  $o$  is part of) that

$$\begin{aligned} & \mathbb{P}(\beta\text{-deg}(o) = k, \text{deg}(o) = d, |\mathcal{V}(\text{CP})| = \infty) \\ &= p_k \sum_{H_1, \dots, H_k \in \mathcal{H}} \sum_{\substack{c_1, \dots, c_k \in \mathbb{Z}^+ \\ c_1 + \dots + c_k = d}} \left( 1 - \eta_r^{\sum_{i=1}^k (H_i - 1)} \right) \prod_{i=1}^k \frac{\nu(c_i | H_i) \mu_{H_i}}{\mathbb{E}[D^r]} = A(k, d). \end{aligned} \tag{5.8}$$

For convenience, denote, with  $K \in \mathbb{Z}^+$ ,

$$A_K(k, d) := \mathbb{P}(\ell\text{-deg}(o) = k, \text{deg}(o) = d, |\mathcal{V}(\text{CP})| \geq K). \tag{5.9}$$

Keep in mind that our goal is to prove (2.23), that is, that  $N_n^{-1} |\mathcal{V}_k^\ell \cap \mathcal{V}_d^p \cap \mathcal{C}_1| \xrightarrow{\mathbb{P}} A(k, d)$ . We compute

$$\begin{aligned} & \lim_{n \rightarrow \infty} \mathbb{P} \left( \left| N_n^{-1} |\mathcal{V}_k^\ell \cap \mathcal{V}_d^p \cap \mathcal{C}_1| - A(k, d) \right| > \varepsilon \right) \\ & \leq \lim_{K \rightarrow \infty} \lim_{n \rightarrow \infty} \mathbb{P} \left( \left| N_n^{-1} |\mathcal{V}_k^\ell \cap \mathcal{V}_d^p \cap \mathcal{C}_1| - N_n^{-1} |\mathcal{V}_k^\ell \cap \mathcal{V}_d^p \cap \mathcal{X}_{\geq K}| \right| > \varepsilon/3 \right) \end{aligned} \tag{5.10a}$$

$$+ \lim_{K \rightarrow \infty} \lim_{n \rightarrow \infty} \mathbb{P} \left( \left| N_n^{-1} |\mathcal{V}_k^\ell \cap \mathcal{V}_d^p \cap \mathcal{X}_{\geq K}| - A_K(k, d) \right| > \varepsilon/3 \right) \tag{5.10b}$$

$$+ \lim_{K \rightarrow \infty} \lim_{n \rightarrow \infty} \mathbb{P} (|A_K(k, d) - A(k, d)| > \varepsilon/3). \tag{5.10c}$$

Clearly,  $A_K(k, d) \rightarrow A(k, d)$  as  $K \rightarrow \infty$  (cf. (5.8) and (5.9)), thus the double limit in (5.10c) equals 0. Next, we look at (5.10b) and note that  $N_n^{-1} |\mathcal{V}_k^\ell \cap \mathcal{V}_d^p \cap \mathcal{X}_{\geq K}| = \mathbb{P}(\rho\text{-deg}(V_n^\ell) = d, \ell\text{-deg}(V_n^\ell) = k, V_n^\ell \in \mathcal{X}_{\geq K} | \mathcal{X}_{\geq K})$ . Thus for  $K$  fixed, by (5.9) and the local weak convergence stated in [48, Theorem 2.7], the inner limit in (5.10b) is 0, thus (5.10b) equals 0. Removing some of the conditions, we can bound (5.10a) as

$$\begin{aligned} & \mathbb{P} \left( \left| N_n^{-1} |\mathcal{V}_k^\ell \cap \mathcal{V}_d^p \cap \mathcal{C}_1| - N_n^{-1} |\mathcal{V}_k^\ell \cap \mathcal{V}_d^p \cap \mathcal{X}_{\geq K}| \right| > \varepsilon/3 \right) \\ & \leq \mathbb{P} (N_n^{-1} |\mathcal{V}_k^\ell \cap \mathcal{V}_d^p \cap (\mathcal{C}_1 \Delta \mathcal{X}_{\geq K})| \geq \varepsilon/3) \\ & \leq \mathbb{P} (N_n^{-1} |\mathcal{C}_1 \Delta \mathcal{X}_{\geq K}| > \varepsilon/3), \end{aligned} \tag{5.11}$$

which tends to 0 as first  $n \rightarrow \infty$  followed by  $K \rightarrow \infty$ , by Lemma 5.1. Thus (5.10a) is also 0, and combining everything above, indeed  $N_n^{-1} |\mathcal{V}_k^\ell \cap \mathcal{V}_d^p \cap \mathcal{C}_1| \xrightarrow{\mathbb{P}} A(k, d)$ , that is, (2.23) holds. This concludes the proof of Theorem 2.7. ■

### ACKNOWLEDGMENTS

This work is supported by the Netherlands Organisation for Scientific Research (NWO) through VICI grant 639.033.806 (RvdH), VENI grant 639.031.447 (JK), the Gravitation Networks grant 024.002.003 (RvdH), and TOP grant 613.001.451 (VV). We thank the reviewers for useful comments that helped to improve the presentation. JK and VV carried out most of the research related to this project at the Eindhoven University of Technology.

### REFERENCES

1. O. Angel, R. van der Hofstad, and C. Holmgren, *Limit laws for self-loops and multiple edges in the configuration model*, *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques* **55** (2019), no. 3, 1509–1530.
2. K. B. Athreya and P. E. Ney, *Branching processes*, Vol **196**, Springer Science & Business Media, Berlin, Germany, 2012.
3. F. Ball, D. Sirl, and P. Trapman, *Threshold behaviour and final outcome of an epidemic on a random network with household structure*, *Adv. Appl. Probab.* **41** (2009), no. 3, 765–796.
4. F. Ball, D. Sirl, and P. Trapman, *Analysis of a stochastic SIR epidemic on a random network incorporating household structure*, *Math. Biosci.* **224** (2010), no. 2, 53–73.
5. F. G. Ball, D. J. Sirl, and P. Trapman, *Epidemics on random intersection graphs*, *Ann. Appl. Probab.* **24** (2014), no. 3, 1081–1128.



6. S. R. Blackburn and S. Gerke, *Connectivity of the uniform random intersection graph*, *Discrete Math.* **309** (2009), no. 16, 5130–5140.
7. M. Bloznelis, *Component evolution in general random intersection graphs*, *SIAM J. Discrete Math.* **24** (2010), no. 2, 639–654.
8. M. Bloznelis, *The largest component in an inhomogeneous random intersection graph with clustering*, *Electron. J. Comb.* **17** (2010), no. 1, 110.
9. M. Bloznelis, *Degree and clustering coefficient in sparse random intersection graphs*, *Ann. Appl. Probab.* **23** (2013), no. 3, 1254–1289.
10. M. Bloznelis, *Degree-degree distribution in a power law random intersection graph with clustering*, *Internet Math.* **1** (2017), no. 1, 2017. <https://doi.org/10.24166/im.03.2017>
11. M. Bloznelis, E. Godehardt, J. Jaworski, V. Kurauskas, and K. Rybarczyk, “Recent progress in complex network analysis: Properties of random intersection graphs,” in B. Lausen, S. Krolak-Schwerdt, M. Böhmer, Eds., *Data science, learning by latent structures, and knowledge discovery*, *Studies in Classification Data Analysis Knowledge Organization*, Springer, Heidelberg, Germany, 2015, pp. 79–88.
12. B. Bollobás, *Random graphs*, *Cambridge Studies in Advanced Mathematics*, 2nd ed., Vol **73**, Cambridge University Press, Cambridge, UK, 2001.
13. B. Bollobás and O. Riordan, *Percolation*, Cambridge University Press, Cambridge, UK, 2006.
14. B. Bollobás and O. Riordan, *An old approach to the giant component problem*, *J. Comb. Theory Ser. B* **113** (2015), 236–260.
15. T. Britton, M. Deijfen, and A. Martin-Löf, *Generating simple random graphs with prescribed degree distribution*, *J. Stat. Phys.* **124** (2006), no. 6, 1377–1397.
16. S. R. Broadbent and J. M. Hammersley, *Percolation processes. I. Crystals and mazes*, *Proc. Cambridge Philos. Soc.* **53** (1957), 629–641.
17. F. Chung and L. Lu, *Connected components in random graphs with given expected degree sequences*, *Ann. Comb.* **6** (2002), no. 2, 125–145.
18. F. Chung and L. Lu, *The volume of the giant component of a random graph with given expected degrees*, *SIAM J. Discrete Math.* **20** (2006), no. 2, 395–411.
19. E. Coupechoux and M. Lelarge, *Contagions in random networks with overlapping communities*, *Adv. Appl. Probab.* **47** (2015), no. 4, 973–988.
20. M. Deijfen and W. Kets, *Random intersection graphs with tunable degree distribution and clustering*, *Probab. Eng. Inf. Sci.* **23** (2009), no. 4, 661–674.
21. R. Durrett, *Random graph dynamics*, *Cambridge Series in Statistical and Probabilistic Mathematics*, Vol **20**, Cambridge University Press, Cambridge, UK, 2007.
22. P. Erdős and A. Rényi, *On the evolution of random graphs*, *Publ. Math. Inst. Hung. Acad. Sci* **5** (1960), no. 1, 17–60.
23. M. Girvan and M. E. J. Newman, *Community structure in social and biological networks*, *Proc. Natl. Acad. Sci.* **99** (2002), no. 12, 7821–7826.
24. E. Godehardt and J. Jaworski, *Two models of random intersection graphs for classification*, Springer, Berlin/Heidelberg, Germany, 2003, 67–81.
25. J. Goldenberg, B. Libai, and E. Muller, *Talk of the network: A complex systems look at the underlying process of word-of-mouth*, *Mark. Lett.* **12** (2001), no. 3, 211–223.
26. G. Grimmett, *Percolation*, *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*, 2nd ed., Vol **321**, Springer-Verlag, Berlin, Germany, 1999.
27. J.-L. Guillaume and M. Latapy, *Bipartite structure of all complex networks*, *Inf. Process. Lett.* **90** (2004), no. 5, 215–221.
28. J.-L. Guillaume and M. Latapy, *Bipartite graphs as models of complex networks*, *Phys. A Stat. Mech. Appl.* **371** (2006), no. 2, 795–813.
29. A. Guille, H. Hacid, C. Favre, and D. A. Zighed, *Information diffusion in online social networks: A survey*, *SIGMOD Rec.* **42** (2013), no. 2, 17–28.
30. S. Janson and M. J. Luczak, *A new approach to the giant component problem*, *Random Struct. Algorithms*, **34** (2009), no. 2, 197–216.
31. S. Janson, T. Łuczak, and A. Ruciński, *Random graphs*, *Wiley-Interscience Series in Discrete Mathematics and Optimization*, Wiley-Interscience, New York, NY, 2000.
32. J. Karjalainen, J. S. H. van Leeuwen, and L. Leskelä, *Parameter estimators of sparse random intersection graphs with thinned communities*, *Proc. Int. Workshop Algorithms Models Web-Graph*, Springer, 2018, pp. 44–58.
33. J. Komjáthy and V. Vadon, *Robustness of percolation on random intersection graphs with communities*, 2020. In preparation.
34. V. Kurauskas, *On local weak limit and subgraph counts for sparse random graphs*, 2015. arXiv 1504.08103.
35. A. N. Lagerås and M. Lindholm, *A note on the component structure in random intersection graphs with tunable clustering*, *Electron. J. Comb.*, **15** (2008), no. 1), 10.
36. A. Martin-Löf, *Symmetric sampling procedures, general epidemic processes and their threshold limit theorems*, *J. Appl. Probab.* **23** (1986), no. 2, 265–282.

37. A. Martin-Löf, *The final size of a nearly critical epidemic, and the first passage time of a Wiener process to a parabolic barrier*, J. Appl. Probab. **35** (1998), no. 3, 671–682.
38. M. Molloy and B. Reed, *A critical point for random graphs with a given degree sequence*, Proc. 6th Int. Seminar Random Graphs Probab. Methods Comb. Comput. Sci., “Random Graphs ’93” (Poznań, 1993), Vol. 6, 1995, pp. 161–179.
39. M. Molloy and B. Reed, *The size of the giant component of a random graph with a given degree sequence*, Comb. Probab. Comput. **7** (1998), no. 3, 295–305.
40. M. E. J. Newman, *Properties of highly clustered networks*, Phys. Rev. E **68** (2003), no. 2, 026121.
41. M. E. J. Newman, *Networks*, Oxford University Press, Oxford, UK, 2010.
42. K. Rybarczyk, *Diameter, connectivity, and phase transition of the uniform random intersection graph*, Discrete Math. **311** (2011), no. 17, 1998–2019.
43. K. B. Singer, *Random intersection graphs*, Ph.D. thesis, Johns Hopkins University, ProQuest LLC, Ann Arbor, MI, 1996.
44. J. Spencer, *The giant component: The golden anniversary*, Not. Am. Math. Soc. **57** (2010), no. 6, 720–724.
45. R. van der Hofstad, *Random graphs and complex networks*, Cambridge Series in Statistical and Probabilistic Mathematics, Vol 1, Cambridge University Press, Cambridge, UK, 2017, 43.
46. R. van der Hofstad, *Random graphs and complex networks*. Vol. 2. In preparation. <http://www.win.tue.nl/~rhofstad/NotesRGCNII.pdf>, 2020.
47. R. van der Hofstad, J. Komjáthy, and V. Vadon, *Random intersection graphs with communities, extended version*. ArXiv 1809.02514. <https://arxiv.org/abs/1809.02514>, 2018.
48. R. van der Hofstad, J. Komjáthy, and V. Vadon, *Random intersection graphs with communities*, Adv. Appl. Probab. **53** (2021), no. 4, 1061–1089.
49. R. van der Hofstad, J. S. H. van Leeuwen, and C. Stegehuis, *Epidemic spreading on complex networks with community structures*, Sci. Rep. **6** (2016), 29748.
50. R. van der Hofstad, J. S. H. van Leeuwen, and C. Stegehuis, *Hierarchical configuration model*, Internet Math. (2016). <https://doi.org/10.24166/im.01.2017>.
51. D. J. Watts and S. H. Strogatz, *Collective dynamics of ‘small-world’ networks*, Nature **393** (1998), no. 6684, 440.
52. O. Yağan, *Zero-one laws for connectivity in inhomogeneous random key graphs*, Inst. Electr. Electron. Eng. Trans. Inf. Theory **62** (2016), no. 8, 4559–4574.
53. O. Yağan and A. M. Makowski, *Zero-one laws for connectivity in random key graphs*, Inst. Electr. Electron. Eng. Trans. Inf. Theory **58** (2012), no. 5, 2983–2999.

**How to cite this article:** R. van der Hofstad, J. Komjáthy, and V. Vadon, *Phase transition in random intersection graphs with communities*, Random Struct. Algorithms. **60** (2022), 406–461. <https://doi.org/10.1002/rsa.21063>

## APPENDIX A: EDGES IN THE GIANT COMPONENT OF THE RIGC

In this section, we provide the details for the sketch proof of Theorem 2.9 from Section 2.3, by proving both Lemma 2.10 and the remaining statements (2.30) and (2.31).

Recall  $D_n^\ell$  with pmf  $\rho^{(n)}$  from (2.12) and its limit  $D^\ell$  with pmf  $\rho$  from Remark 2.4(ii). Denote a uniform  $\ell$ -vertex  $V_n^\ell \sim \text{Unif}[\mathcal{V}^\ell]$  and its (projected) degree (see (2.4))  $D_n^\rho = \rho\text{-deg}(V_n^\ell)$ . In [48, Corollary 2.8], the distributional limit of  $D_n^\rho$  is established as  $D^\rho \stackrel{d}{=} \sum_{i=1}^{D^\ell} D_{(i)}^\ell$ , where  $D_{(i)}^\ell$  are iid copies of  $D^\ell$  independently of  $D^\ell$ . Let  $\mathcal{H}_n$  and  $\mathcal{H}$  denote random graphs with pmfs  $\mu^{(n)}$  from (2.10) and  $\mu$  from Assumption 2.3(C), respectively.

### A.1 | Proof of Lemma 2.10

We prove Lemma 2.10 by showing that statement (i) is equivalent to both statement (ii) and statement (iii).

*Proof of Lemma 2.10(i) implies (ii).* We prove that  $(D_n^\rho)_{n \in \mathbb{N}}$  is uniformly integrable by showing that for any  $\varepsilon > 0$ ,  $\mathbb{E}[D_n^\rho \mathbb{1}_{\{D_n^\rho \geq K\}}] < \varepsilon$  for  $K$  large enough, uniformly in  $n$ . Recall that  $\mathcal{V}(\mathbf{Com})$  denotes the disjoint union of all vertices in community graphs and that for  $j \in \mathcal{V}(H)$ ,  $d_j^\circ$  denotes the degree of  $j$  within  $H$ . Recall that  $v \leftarrow j$  denotes the event that  $j \in \mathcal{V}(\mathbf{Com})$  is one of the community roles assigned to  $v \in \mathcal{V}^\ell$ , so that we can write the *random* degree of  $v$  (see (2.4)), depending on the bipartite matching, as  $d_v^\rho = \sum_{j \in \mathcal{V}(\mathbf{Com})} d_j^\circ \mathbb{1}_{\{v \leftarrow j\}}$ . We compute, for  $K \in \mathbb{N}$ , by taking the empirical average first,

$$\begin{aligned} \mathbb{E}[D_n^\rho \mathbb{1}_{\{D_n^\rho \geq K\}}] &= \mathbb{E} \left[ N_n^{-1} \sum_{v \in [N_n]} \sum_{j \in \mathcal{V}(\mathbf{Com})} d_j^\circ \cdot \mathbb{1}_{\{v \leftarrow j\}} \mathbb{1}_{\{d_v^\rho \geq K\}} \right] \\ &= N_n^{-1} \sum_{j \in \mathcal{V}(\mathbf{Com})} d_j^\circ \cdot \mathbb{E} \left[ \sum_{v \in [N_n]} \mathbb{1}_{\{v \leftarrow j\}} \mathbb{1}_{\{d_v^\rho \geq K\}} \right] \\ &= N_n^{-1} \sum_{j \in \mathcal{V}(\mathbf{Com})} d_j^\circ \sum_{v \in [N_n]} \mathbb{P}(v \leftarrow j, d_v^\rho \geq K). \end{aligned} \tag{A1}$$

We further rewrite the probability as

$$\mathbb{P}(v \leftarrow j, d_v^\rho \geq K) = \mathbb{P}(d_v^\rho \geq K \mid v \leftarrow j) \cdot \mathbb{P}(v \leftarrow j). \tag{A2}$$

Note that  $\mathbb{P}(v \leftarrow j) = d_j^\circ / \ell_n$ . We now split the sum over  $j$  according to whether  $d_j^\circ$  is smaller or larger than  $\sqrt{K}$ . For  $d_j^\circ \geq \sqrt{K}$ , we use the trivial bound  $\mathbb{P}(d_v^\rho \geq K \mid v \leftarrow j) \leq 1$ .

$$\mathbb{E}[D_n^\rho \mathbb{1}_{\{D_n^\rho \geq K\}}] \leq N_n^{-1} \sum_{j \in \mathcal{V}(\mathbf{Com})} d_j^\circ \mathbb{1}_{\{d_j^\circ \geq \sqrt{K}\}} \sum_{v \in [N_n]} \frac{d_v^\circ}{\ell_n} \tag{A3a}$$

$$+ N_n^{-1} \sum_{j \in \mathcal{V}(\mathbf{Com})} d_j^\circ \mathbb{1}_{\{d_j^\circ < \sqrt{K}\}} \sum_{v \in [N_n]} \mathbb{P}(d_v^\rho \geq K \mid v \leftarrow j) \cdot \frac{d_v^\circ}{\ell_n}. \tag{A3b}$$

The term corresponding to large values of  $d_j^\circ$ , that is, the rhs of (A3a) equals

$$N_n^{-1} \sum_{v \in N_n} d_v^\circ \frac{1}{\ell_n} \sum_{j \in \mathcal{V}(\mathbf{Com})} d_j^\circ \mathbb{1}_{\{d_j^\circ \geq \sqrt{K}\}} = \mathbb{E}[D_n^\circ] \cdot \mathbb{E}[D_n^\circ \mathbb{1}_{\{D_n^\circ \geq \sqrt{K}\}}], \tag{A4}$$

where  $\mathbb{E}[D_n^\circ]$  is bounded due to Assumption 2.3(B) and since  $(D_n^\circ)_{n \in \mathbb{N}}$  is UI by assumption,  $\mathbb{E}[D_n^\circ \mathbb{1}_{\{D_n^\circ \geq \sqrt{K}\}}]$  can be made arbitrarily small uniformly in  $n$  by choosing  $K$  large enough. In the term corresponding to small values of  $d_j^\circ$ , that is, (A3b), we further analyze  $\mathbb{P}(d_v^\rho \geq K \mid v \leftarrow j)$ . Note that conditionally on  $v \leftarrow j$ ,  $d_v^\rho = d_j^\circ + d_{J_2}^\circ + \dots + d_{J_{\ell-\deg(v)}}^\circ$ , where  $J_i$  is chosen uar from  $\mathcal{V}(\mathbf{Com}) \setminus \{j, J_2, \dots, J_{i-1}\}$ . Using that  $d_j^\circ < \sqrt{K}$ ,

$$\begin{aligned} \mathbb{P}(d_v^\rho \geq K \mid v \leftarrow j) &= \mathbb{P}(d_v^\rho - d_j^\circ = d_{J_2}^\circ + \dots + d_{J_{\ell-\deg(v)}}^\circ \geq K - d_j^\circ) \\ &\leq \mathbb{P}(d_{J_2}^\circ + \dots + d_{J_{\ell-\deg(v)}}^\circ \geq K - \sqrt{K}) \\ &\leq \frac{\mathbb{E}[d_{J_2}^\circ] + \dots + \mathbb{E}[d_{J_{\ell-\deg(v)}}^\circ]}{K - \sqrt{K}}, \end{aligned} \tag{A5}$$

by Markov's inequality. For any fixed  $v$ ,  $\ell$ -deg( $v$ ) is a constant and is  $o(n)$  uniformly in  $v$ , thus the depletion of community vertices becomes negligible as  $n \rightarrow \infty$ , and  $\mathbb{E} \left[ d_{j_i}^\ell \right] = \mathbb{E}[D_n^\ell](1 + o(1))$  for all  $i$ . Substituting, and using that  $d_j^\ell \mathbb{1}_{\{d_j^\ell < \sqrt{K}\}} \leq \sqrt{K} \mathbb{1}_{\{d_j^\ell < \sqrt{K}\}} \leq \sqrt{K}$ , we can bound (A3b) by

$$\begin{aligned} & N_n^{-1} \sum_{j \in \mathcal{V}(\mathbf{Com})} d_j^\ell \mathbb{1}_{\{d_j^\ell < \sqrt{K}\}} \sum_{v \in [N_n]} \frac{\mathbb{E}[D_n^\ell - 1] \cdot \mathbb{E}[D_n^\ell](1 + o(1))}{K - \sqrt{K}} \cdot \frac{d_v^\ell}{\ell_n} \\ &= \mathbb{E}[D_n^\ell - 1] \mathbb{E}[D_n^\ell](1 + o(1)) \frac{1}{N_n} \sum_{v \in [N_n]} d_v^\ell \frac{1}{\ell_n} \sum_{j \in \mathcal{V}(\mathbf{Com})} d_j^\ell \mathbb{1}_{\{d_j^\ell < \sqrt{K}\}} \frac{1}{K - \sqrt{K}} \\ &\leq \mathbb{E}[D_n^\ell - 1] \mathbb{E}[D_n^\ell](1 + o(1)) \mathbb{E}[D_n^\ell] \frac{\ell_n \sqrt{K}}{\ell_n (K - \sqrt{K})}, \end{aligned} \tag{A6}$$

where  $\mathbb{E}[D_n^\ell - 1] \mathbb{E}[D_n^\ell](1 + o(1)) \mathbb{E}[D_n^\ell]$  is bounded due to Assumption 2.3(B) and our assumption that  $(D_n^c)_{n \in \mathbb{N}}$  is UI, and  $\sqrt{K}/(K - \sqrt{K})$  can be made arbitrarily small for  $K$  large enough. We conclude that by choosing  $K$  sufficiently large,  $\mathbb{E} \left[ D_n^\ell \mathbb{1}_{\{D_n^\ell \geq K\}} \right]$  can be made arbitrarily small, uniformly in  $n$ , that is,  $(D_n^\ell)_{n \in \mathbb{N}}$  is uniformly integrable.

*Proof of Lemma 2.10(ii) implies (i).* Recall that for  $v \in \mathcal{V}^\ell$  and  $j \in \mathcal{V}(\mathbf{Com})$ ,  $v \leftarrow j$  denotes the event that  $j$  is one of the community roles assigned to  $v$ . Note that each  $j$  is assigned to a *unique*  $v$ , thus  $\sum_{v \in [N_n]} \mathbb{1}_{\{v \leftarrow j\}} = 1$  (almost surely). We calculate, for some  $K \in \mathbb{N}$ ,

$$\begin{aligned} \mathbb{E} \left[ D_n^c \mathbb{1}_{\{D_n^c \geq K\}} \right] &= \frac{1}{\ell_n} \sum_{j \in \mathcal{V}(\mathbf{Com})} d_j^c \mathbb{1}_{\{d_j^c \geq K\}} \\ &= \frac{1}{\ell_n} \sum_{j \in \mathcal{V}(\mathbf{Com})} d_j^c \mathbb{1}_{\{d_j^c \geq K\}} \mathbb{E} \left[ \sum_{v \in [N_n]} \mathbb{1}_{\{v \leftarrow j\}} \right] \\ &= \mathbb{E} \left[ \frac{1}{\ell_n} \sum_{v \in [N_n]} \sum_{j \in \mathcal{V}(\mathbf{Com})} d_j^c \mathbb{1}_{\{v \leftarrow j\}} \mathbb{1}_{\{d_j^c \geq K\}} \right]. \end{aligned} \tag{A7}$$

We recognize the  $\rho$ -degree (see (2.4)) of  $v$  written as  $d_v^\rho = \sum_{j \in \mathcal{V}(\mathbf{Com})} d_j^c \mathbb{1}_{\{v \leftarrow j\}}$ , and note that on the event  $v \leftarrow j$ ,  $d_j^c \geq K$  implies that  $d_v^\rho \geq K$ . Thus,

$$\begin{aligned} \mathbb{E} \left[ D_n^c \mathbb{1}_{\{D_n^c \geq K\}} \right] &\leq \mathbb{E} \left[ \frac{1}{\ell_n} \sum_{v \in [N_n]} \sum_{j \in \mathcal{V}(\mathbf{Com})} d_j^c \mathbb{1}_{\{v \leftarrow j\}} \mathbb{1}_{\{d_v^\rho \geq K\}} \right] \\ &= \mathbb{E} \left[ \frac{1}{\mathbb{E}[D_n^\ell] N_n} \sum_{v \in [N_n]} d_v^\rho \mathbb{1}_{\{d_v^\rho \geq K\}} \right] = \frac{1}{\mathbb{E}[D_n^\ell]} \mathbb{E} \left[ D_n^\rho \mathbb{1}_{\{D_n^\rho \geq K\}} \right]. \end{aligned} \tag{A8}$$

This can be made arbitrarily small, uniformly in  $n$ , by choosing  $K$  large enough, since  $\mathbb{E}[D_n^\ell]$  is bounded due to Assumption 2.3(B) (and bounded away from 0), and  $(D_n^\rho)_{n \in \mathbb{N}}$  is UI by assumption. This implies that  $(D_n^c)_{n \in \mathbb{N}}$  is also UI.

*Proof of Lemma 2.10(i) implies (iii).* Note that with an  $r$ -vertex  $V_n^r \sim \text{Unif}[M_n]$  chosen uar,  $\text{Com}_{V_n^r} \stackrel{d}{=} \mathcal{H}_n$ . Thus, for  $K \in \mathbb{N}$ , we compute

$$\mathbb{E} \left[ |\mathcal{E}(\mathcal{H}_n)| \mathbb{1}_{\{|\mathcal{E}(\mathcal{H}_n)| \geq K\}} \right] = \frac{1}{M_n} \sum_{a \in [M_n]} |\mathcal{E}(\text{Com}_a)| \cdot \mathbb{1}_{\{|\mathcal{E}(\text{Com}_a)| \geq K\}}. \tag{A9}$$

For any graph  $H$ ,  $|\mathcal{E}(H)| = \frac{1}{2} \sum_{v \in \mathcal{V}(H)} \deg(v)$ . Noting that for any graph  $H$ ,  $|\mathcal{E}(H)| \leq |H|(|H| - 1)/2$ , we have that  $|\mathcal{E}(H)| \geq K$  implies  $|H| \geq \sqrt{2K}$ , thus

$$\begin{aligned} \mathbb{E} [|\mathcal{E}(\mathcal{H}_n)| \mathbb{1}_{\{|\mathcal{E}(\mathcal{H}_n)| \geq K\}}] &= \frac{1}{M_n} \sum_{a \in [M_n]} \frac{1}{2} \sum_{j \in \mathcal{V}(\text{Com}_a)} d_j^c \cdot \mathbb{1}_{\{|\mathcal{E}(\text{Com}_a)| \geq K\}} \\ &\leq \frac{1}{2M_n} \sum_{a \in [M_n]} \sum_{j \in \mathcal{V}(\text{Com}_a)} d_j^c \cdot \mathbb{1}_{\{|\text{Com}_a| \geq \sqrt{2K}\}} = \frac{1}{2M_n} \sum_{j \in \mathcal{V}(\mathbf{Com})} d_j^c \cdot \mathbb{1}_{\{|\text{Com}_{a(j)}| \geq \sqrt{2K}\}}, \end{aligned} \quad (\text{A10})$$

where  $a(j)$  is the community that  $j$  is a vertex in. To estimate (A10), we first compute the number of terms, that is, the number of vertices in large enough communities, using that  $\mathcal{V}(\mathbf{Com}) = \cup_{a \in [M_n]} \mathcal{V}(\text{Com}_a)$ :

$$\begin{aligned} \kappa_n &:= \sum_{j \in \mathcal{V}(\mathbf{Com})} \mathbb{1}_{\{|\text{Com}_{a(j)}| \geq \sqrt{2K}\}} = \sum_{a \in [M_n]} |\text{Com}_a| \cdot \mathbb{1}_{\{|\text{Com}_a| \geq \sqrt{2K}\}} \\ &= M_n \mathbb{E} \left[ D_n^c \mathbb{1}_{\{D_n^c \geq \sqrt{2K}\}} \right]. \end{aligned} \quad (\text{A11})$$

By Assumption 2.3 (D),  $(D_n^c)_{n \in \mathbb{N}}$  is UI, thus  $\mathbb{E} \left[ D_n^c \mathbb{1}_{\{D_n^c \geq \sqrt{2K}\}} \right]$  can be made arbitrarily small uniformly in  $n$  by choosing  $K$  large enough. That is, we can make  $\kappa_n \leq \delta' M_n \leq \delta \ell_n$ , since  $\ell_n/M_n = \mathbb{E}[D_n^c]$  is bounded. Let  $d_{(i)}^c$  denote the  $i$ th largest element in the sequence  $(d_j^c)_{j \in \mathcal{V}(\mathbf{Com})}$ . Since there are at most  $\delta \ell_n$  community roles in communities larger than  $\sqrt{2K}$ , we can bound the sum of their  $c$ -degrees on the rhs of (A10) by taking  $\delta \ell_n$  (or more, if there are several terms equal to the one with rank  $\delta \ell_n$ ) of the largest  $c$ -degrees:

$$\mathbb{E} [|\mathcal{E}(\mathcal{H}_n)| \mathbb{1}_{\{|\mathcal{E}(\mathcal{H}_n)| \geq K\}}] \leq \frac{1}{2M_n} \sum_{j \in \mathcal{V}(\mathbf{Com})} d_j^c \cdot \mathbb{1}_{\{d_j^c \geq d_{(\delta \ell_n)}^c\}}. \quad (\text{A12})$$

Note that  $d_{(\delta \ell_n)}^c$  equals the upper- $\delta$ -quantile  $Q_n^\delta$  of the empirical distribution  $D_n^c$ , thus it must converge to the upper- $\delta$ -quantile  $Q^\delta$  of the distribution  $D^c$ . Consequently, there exists  $n_0$  such that for  $n \geq n_0$ ,  $Q_n^\delta \geq Q^\delta - \varepsilon =: K'$ , and

$$\begin{aligned} \mathbb{E} [|\mathcal{E}(\mathcal{H}_n)| \mathbb{1}_{\{|\mathcal{E}(\mathcal{H}_n)| \geq K\}}] &\leq \frac{\mathbb{E}[D_n^c]}{2\ell_n} \sum_{j \in \mathcal{V}(\mathbf{Com})} d_j^c \mathbb{1}_{\{d_j^c \geq Q^\delta - \varepsilon\}} \\ &= \frac{1}{2} \mathbb{E}[D_n^c] \mathbb{E} [D_n^c \mathbb{1}_{\{D_n^c \geq K'\}}]. \end{aligned} \quad (\text{A13})$$

Since  $\mathbb{E}[D_n^c]$  is bounded due to Assumption 2.3 (D) and  $(D_n^c)_{n \in \mathbb{N}}$  is UI, the upper bound can be made arbitrarily small for  $n \geq n_0$  by choosing  $K'$  sufficiently large. This is possible by choosing the earlier  $K$  sufficiently large so that  $\delta$  is sufficiently small. Further, since  $n < n_0$  is a finite set, we can increase  $K$  so that  $\max_{n < n_0} \mathbb{E} [|\mathcal{E}(\mathcal{H}_n)| \mathbb{1}_{\{|\mathcal{E}(\mathcal{H}_n)| \geq K\}}]$  is sufficiently small as well. We conclude that indeed,  $(|\mathcal{E}(\mathcal{H}_n)|)_{n \in \mathbb{N}}$  is UI.

*Proof of Lemma 2.10(iii) implies (i).* Recall that  $\mathcal{V}(\mathbf{Com})$  denotes the union of vertices in all  $\text{Com}_a \in \mathbf{Com}$ . We calculate, for some  $K \in \mathbb{N}$ ,

$$\mathbb{E} [D_n^c \mathbb{1}_{\{D_n^c \geq K\}}] = \frac{1}{\ell_n} \sum_{j \in \mathcal{V}(\mathbf{Com})} d_j^c \mathbb{1}_{\{d_j^c \geq K\}} = \frac{1}{\ell_n} \sum_{a \in [M_n]} \sum_{j \in \mathcal{V}(\text{Com}_a)} d_j^c \mathbb{1}_{\{d_j^c \geq K\}}. \quad (\text{A14})$$

Note that  $|\mathcal{E}(\text{Com}_a)| = 2 \sum_{j \in \mathcal{V}(\text{Com}_a)} d_j^c$ , and further, that for  $j \in \mathcal{V}(\text{Com}_a)$ ,  $d_j^c \geq K$  implies  $|\mathcal{E}(\text{Com}_a)| \geq K$ . Thus

$$\begin{aligned} \mathbb{E} [D_n^c \mathbb{1}_{\{D_n^c \geq K\}}] &\leq \frac{1}{\ell_n} \sum_{a \in [M_n]} \sum_{j \in \mathcal{V}(\text{Com}_a)} d_j^c \cdot \mathbb{1}_{\{|\mathcal{E}(\text{Com}_a)| \geq K\}} \\ &= \frac{1}{\mathbb{E}[D_n^c] M_n} \sum_{a \in [M_n]} 2|\mathcal{E}(\text{Com}_a)| \cdot \mathbb{1}_{\{|\mathcal{E}(\text{Com}_a)| \geq K\}} \\ &= \frac{2}{\mathbb{E}[D_n^c]} \mathbb{E} [|\mathcal{E}(\mathcal{H}_n)| \mathbb{1}_{\{|\mathcal{E}(\mathcal{H}_n)| \geq K\}}]. \end{aligned} \tag{A15}$$

This can be made arbitrarily small, uniformly in  $n$ , by choosing  $K$  large enough, since  $\mathbb{E}[D_n^c]$  is bounded due to Assumption 2.3 (D) (and bounded away from 0), and  $|\mathcal{E}(\mathcal{H}_n)|$  is UI. This implies that  $(D_n^c)_{n \in \mathbb{N}}$  is also UI.

**A.2 | Proof of Theorem 2.9**

Recall from the sketch proof that in order to complete the proof of Theorem 2.9, we need to prove (2.30) and (2.31), that we complete below one by one.

*Proof of (2.30).* Recall (2.4) and  $D_n^\rho = \rho\text{-deg}(V_n^\ell)$ , with  $V_n^\ell \sim \text{Unif}[N_n]$ . Note that  $D_n^\rho$  has double randomness: the choice of  $V_n^\ell$  and the bipartite matching  $\omega_n$  that determines  $\rho\text{-deg}(v)$  (see (2.4)) for each  $\ell$ -vertex  $v$ . We rewrite

$$\frac{|\mathcal{E}(\mathcal{C}_1)|}{N_n} = \frac{1}{2} \mathbb{E} \left[ \rho\text{-deg}(V_n^\ell) \mathbb{1}_{\{V_n^\ell \in \mathcal{C}_1\}} \mid \omega_n \right] = \frac{1}{2N_n} \sum_{v \in [N_n]} \rho\text{-deg}(v) \mathbb{1}_{\{v \in \mathcal{C}_1\}}. \tag{A16}$$

Under the condition (2.28), by Lemma 2.10, we have that  $(D_n^\rho)_{n \in \mathbb{N}}$  is UI. Noting that  $\rho\text{-deg}(V_n^\ell) \mathbb{1}_{\{V_n^\ell \in \mathcal{C}_1\}} \mid \omega_n \leq \rho\text{-deg}(V_n^\ell) \mid \omega_n$ ,

$$\mathbb{E} \left[ \rho\text{-deg}(V_n^\ell) \mathbb{1}_{\{V_n^\ell \in \mathcal{C}_1\}} \mathbb{1}_{\{\rho\text{-deg}(V_n^\ell) > K\}} \mid \omega_n \right] \leq \mathbb{E} \left[ \rho\text{-deg}(V_n^\ell) \mathbb{1}_{\{\rho\text{-deg}(V_n^\ell) > K\}} \mid \omega_n \right], \tag{A17}$$

which implies that also  $(\rho\text{-deg}(V_n^\ell) \mathbb{1}_{\{V_n^\ell \in \mathcal{C}_1\}})_{n \in \mathbb{N}}$  is UI. Recall that by (2.23), the empirical mass function of  $\rho\text{-deg}(V_n^\ell) \mathbb{1}_{\{V_n^\ell \in \mathcal{C}_1\}}$  converges in probability, which combined with the uniform integrability implies the convergence of the empirical average to the mean of the limit:

$$\begin{aligned} \frac{|\mathcal{E}(\mathcal{C}_1)|}{N_n} &= \frac{1}{2} \mathbb{E} \left[ \rho\text{-deg}(V_n^\ell) \mathbb{1}_{\{V_n^\ell \in \mathcal{C}_1\}} \mid \omega_n \right] \\ &= \frac{1}{2} \sum_{d \in \mathbb{N}} d \cdot \mathbb{P} \left( \rho\text{-deg}(V_n^\ell) = d, V_n^\ell \in \mathcal{C}_1 \mid \omega_n \right) = \frac{1}{2} \sum_{d \in \mathbb{N}} d \cdot \frac{|\mathcal{V}_d^\rho \cap \mathcal{C}_1|}{N_n} \\ &= \frac{1}{2} \sum_{d \in \mathbb{N}} d \cdot \sum_{k \in \mathbb{Z}^+} \frac{|\mathcal{V}_k^\ell \cap \mathcal{V}_d^\rho \cap \mathcal{C}_1|}{N_n} \xrightarrow{\mathbb{P}} \frac{1}{2} \sum_{d \in \mathbb{N}} \sum_{k \in \mathbb{Z}^+} d \cdot A(k, d), \end{aligned} \tag{A18}$$

by (2.23). This concludes the proof of (2.30).

*Proof of (2.31).* Recall that  $\mathcal{H}$  denotes a random graph with pmf  $\mu$ . Note that under condition (2.28), by Lemma 2.10,  $\mathbb{E} \left[ |\mathcal{E}(\mathcal{H})| (1 - \eta_r^{|\mathcal{H}|-1}) \right] \leq \mathbb{E} [|\mathcal{E}(\mathcal{H})|] \leq \limsup_{n \rightarrow \infty} \mathbb{E} [|\mathcal{E}(\mathcal{H}_n)|] < \infty$ . Substituting (2.22), and noting that under the condition  $d = c_1 + \dots + c_k$ , we can replace  $d$  by this sum, we compute

$$\sum_{d \in \mathbb{N}} \sum_{k \in \mathbb{Z}^+} d \cdot A(k, d) \tag{A19a}$$

$$\begin{aligned}
 &= \sum_{d \in \mathbb{N}} \sum_{k \in \mathbb{Z}^+} p_k \sum_{H_1, \dots, H_k \in \mathcal{H}} \sum_{\substack{c_1, \dots, c_k \in \mathbb{N} \\ c_1 + \dots + c_k = d}} \left( \sum_{j=1}^k c_j \right) \left( 1 - \eta_r^{\sum_{i=1}^k (|H_i| - 1)} \right) \prod_{i=1}^k \frac{v(c_i | H_i) \mu_{H_i}}{\mathbb{E}[D^r]} \\
 &= \sum_{k \in \mathbb{Z}^+} p_k \sum_{H_1, \dots, H_k \in \mathcal{H}} \sum_{c_1, \dots, c_k \in \mathbb{N}} \sum_{j=1}^k c_j \prod_{i=1}^k \frac{v(c_i | H_i) \mu_{H_i}}{\mathbb{E}[D^r]} \\
 &- \sum_{k \in \mathbb{Z}^+} p_k \sum_{H_1, \dots, H_k \in \mathcal{H}} \sum_{c_1, \dots, c_k \in \mathbb{N}} \sum_{j=1}^k c_j \cdot \eta_r^{\sum_{i=1}^k (|H_i| - 1)} \prod_{i=1}^k \frac{v(c_i | H_i) \mu_{H_i}}{\mathbb{E}[D^r]}, \tag{A19b}
 \end{aligned}$$

where we have used that summing over  $d$  removes the restriction  $c_1 + \dots + c_k = d$ . In the following, we simplify (A19a) and (A19b) separately, starting with (A19a).

For fixed  $k$  and  $H_1, \dots, H_k \in \mathcal{H}$ , we compute

$$\begin{aligned}
 \sum_{c_1, \dots, c_k \in \mathbb{N}} \sum_{j=1}^k c_j \prod_{i=1}^k \frac{v(c_i | H_i) \mu_{H_i}}{\mathbb{E}[D^r]} &= \sum_{c_1, \dots, c_k \in \mathbb{N}} \sum_{j=1}^k \frac{c_j \cdot v(c_j | H_j) \mu_{H_j}}{\mathbb{E}[D^r]} \prod_{\substack{i \in [k] \\ i \neq j}} \frac{v(c_i | H_i) \mu_{H_i}}{\mathbb{E}[D^r]} \\
 &= \sum_{j=1}^k \frac{\sum_{c_j \in \mathbb{N}} c_j \cdot v(c_j | H_j) \mu_{H_j}}{\mathbb{E}[D^r]} \prod_{\substack{i \in [k] \\ i \neq j}} \frac{\sum_{c_i \in \mathbb{N}} v(c_i | H_i) \mu_{H_i}}{\mathbb{E}[D^r]}. \tag{A20}
 \end{aligned}$$

Recall that  $v(c_i | H_i)$  counts the number of  $v \in \mathcal{V}(H_i)$  with degree  $c_i$ . Thus

$$\sum_{c_i \in \mathbb{N}} v(c_i | H_i) = |H_i|, \quad \sum_{c_j \in \mathbb{N}} c_j \cdot v(c_j | H_j) = \sum_{v \in \mathcal{V}(H_j)} \deg(v) = 2|\mathcal{E}(H_j)|. \tag{A21}$$

Again, denote by  $\mathcal{H}$  a random graph with pmf  $\mu$ . Combining (A20)–(A21) and substituting, (A19a) equals

$$\begin{aligned}
 &\sum_{k \in \mathbb{Z}^+} p_k \sum_{H_1, \dots, H_k \in \mathcal{H}} \sum_{j=1}^k \frac{2|\mathcal{E}(H_j)| \mu_{H_j}}{\mathbb{E}[D^r]} \prod_{\substack{i \in [k] \\ i \neq j}} \frac{|H_i| \mu_{H_i}}{\mathbb{E}[D^r]} \\
 &= \sum_{k \in \mathbb{Z}^+} p_k \sum_{j=1}^k \frac{\sum_{H_j \in \mathcal{H}} 2|\mathcal{E}(H_j)| \mu_{H_j}}{\mathbb{E}[D^r]} \prod_{\substack{i \in [k] \\ i \neq j}} \frac{\sum_{H_i \in \mathcal{H}} |H_i| \mu_{H_i}}{\mathbb{E}[D^r]} \\
 &= \sum_{k \in \mathbb{Z}^+} p_k \sum_{j=1}^k \frac{2\mathbb{E}[|\mathcal{E}(\mathcal{H})|]}{\mathbb{E}[D^r]} \prod_{\substack{i \in [k] \\ i \neq j}} \frac{\mathbb{E}[|\mathcal{H}|]}{\mathbb{E}[D^r]} = \sum_{k \in \mathbb{Z}^+} p_k \cdot k \cdot \frac{2\mathbb{E}[|\mathcal{E}(\mathcal{H})|]}{\mathbb{E}[D^r]}, \tag{A22}
 \end{aligned}$$

since  $\mathbb{E}[|\mathcal{H}|] / \mathbb{E}[D^r] = 1$ , so the product over  $i$  disappears, which yields  $k$  identical terms in the sum over  $j$ . Note that  $\sum_{k \in \mathbb{Z}^+} p_k \cdot k = \mathbb{E}[D^\ell]$  and recall from Remark 2.4(i) that  $\mathbb{E}[D^\ell] / \mathbb{E}[D^r] = \gamma$ . We conclude that (A19a) equals

$$\frac{2\mathbb{E}[|\mathcal{E}(\mathcal{H})|]}{\mathbb{E}[D^r]} \mathbb{E}[D^\ell] = 2\gamma \mathbb{E}[|\mathcal{E}(\mathcal{H})|]. \tag{A23}$$

Next, we study (A19b). Using (A20)–(A21) and substituting,

$$\begin{aligned}
 & \sum_{k \in \mathbb{Z}^+} p_k \sum_{H_1, \dots, H_k \in \mathcal{H}} \eta_r^{\sum_{i=1}^k (|H_i| - 1)} \sum_{j=1}^k \frac{2|\mathcal{E}(H_j)| \mu_{H_j}}{\mathbb{E}[D^r]} \prod_{\substack{i \in [k] \\ i \neq j}} \frac{|H_i| \mu_{H_i}}{\mathbb{E}[D^r]} \\
 &= \sum_{k \in \mathbb{Z}^+} p_k \sum_{H_1, \dots, H_k \in \mathcal{H}} \sum_{j=1}^k \frac{2|\mathcal{E}(H_j)| \mu_{H_j} \cdot \eta_r^{|H_j| - 1}}{\mathbb{E}[D^r]} \prod_{\substack{i \in [k] \\ i \neq j}} \frac{|H_i| \mu_{H_i} \cdot \eta_r^{|H_i| - 1}}{\mathbb{E}[D^r]} \\
 &= \sum_{k \in \mathbb{Z}^+} p_k \sum_{j=1}^k \frac{\sum_{H_j \in \mathcal{H}} 2|\mathcal{E}(H_j)| \mu_{H_j} \cdot \eta_r^{|H_j| - 1}}{\mathbb{E}[D^r]} \prod_{\substack{i \in [k] \\ i \neq j}} \frac{\sum_{H_i \in \mathcal{H}} |H_i| \mu_{H_i} \cdot \eta_r^{|H_i| - 1}}{\mathbb{E}[D^r]}. \tag{A24}
 \end{aligned}$$

Recall (1.1), (1.2), and  $q$  from Assumption 2.3(C1), and note that

$$\begin{aligned}
 \frac{\sum_{H_i \in \mathcal{H}} |H_i| \mu_{H_i} \cdot \eta_r^{|H_i| - 1}}{\mathbb{E}[D^r]} &= \sum_{m \in \mathbb{Z}^+} \frac{m \cdot q_m}{\mathbb{E}[D^r]} \eta_r^{m-1} \\
 &= \sum_{m \in \mathbb{Z}^+} \mathbb{P}(\tilde{D}^r = m - 1) \eta_r^{m-1} = G_{\tilde{D}^r}(\eta_r) = \eta_\ell, \tag{A25}
 \end{aligned}$$

as we have shown in Section 5.1. Consequently in (A24), the last product over  $i$  equals  $\eta_\ell^{k-1}$ , and in the factor before, we recognize an expectation wrt the random graph  $\mathcal{H}$  with pmf  $\mu$ . Recall  $p$  from Assumption 2.3(A). Combining (A24)–(A25), we have that (A19b) without the minus sign equals

$$\begin{aligned}
 & \sum_{k \in \mathbb{Z}^+} p_k \sum_{j=1}^k \frac{2\mathbb{E}\left[|\mathcal{E}(\mathcal{H})| \eta_r^{|\mathcal{H}| - 1}\right]}{\mathbb{E}[D^r]} \eta_\ell^{k-1} = \frac{2\mathbb{E}\left[|\mathcal{E}(\mathcal{H})| \eta_r^{|\mathcal{H}| - 1}\right]}{\mathbb{E}[D^r]} \sum_{k \in \mathbb{Z}^+} p_k \cdot k \cdot \eta_\ell^{k-1} \\
 &= 2\mathbb{E}\left[|\mathcal{E}(\mathcal{H})| \eta_r^{|\mathcal{H}| - 1}\right] \frac{\mathbb{E}[D^\ell]}{\mathbb{E}[D^r]} \sum_{k \in \mathbb{Z}^+} \frac{k \cdot p_k}{\mathbb{E}[D^\ell]} \eta_\ell^{k-1} = 2\gamma \mathbb{E}\left[|\mathcal{E}(\mathcal{H})| \eta_r^{|\mathcal{H}| - 1}\right] G_{\tilde{D}^\ell}(\eta_\ell) \\
 &= 2\gamma \mathbb{E}\left[|\mathcal{E}(\mathcal{H})| \eta_r^{|\mathcal{H}| - 1}\right] \eta_r = 2\gamma \mathbb{E}\left[|\mathcal{E}(\mathcal{H})| \eta_r^{|\mathcal{H}|}\right], \tag{A26}
 \end{aligned}$$

where we have used that we have defined  $\eta_r = G_{\tilde{D}^\ell}(\eta_\ell)$  in Theorem 2.7. Combining (A23) and (A26), we obtain that

$$\sum_{k \in \mathbb{Z}^+} \sum_{d \in \mathbb{N}} d \cdot A(k, d) = 2\gamma \mathbb{E}\left[|\mathcal{E}(\mathcal{H})|\right] - 2\gamma \mathbb{E}\left[|\mathcal{E}(\mathcal{H})| \eta_r^{|\mathcal{H}|}\right], \tag{A27}$$

which is equivalent to (2.31).

**APPENDIX B: DEATH PROCESSES AND THEIR HITTING TIMES**

In this section, we prove Claim 3.6. We only show that statement (i) implies statement (ii), the proof of the reverse implication is analogous.

*Proof: Statement (i) implies statement (ii).* For  $\delta > 0, T < \infty$  and  $\varepsilon > 0, c_0 > 0$ , we define the events

$$\mathcal{E}_1^{(n)}(\delta, T) := \left\{ \sup_{t \leq T} \left| a_n^{-1} X^{(n)}(t) - f(t) \right| < \delta \right\}, \tag{B1}$$



$$\mathcal{E}_2^{(n)}(\varepsilon, c_0) := \left\{ \sup_{c \geq c_0} |\mathcal{T}^{(n)}(c) - f^{(-1)}(c)| < \varepsilon \right\}. \tag{B2}$$

Statement (i) implies that for any  $\delta > 0$  and  $T < \infty$  fixed,

$$\mathbb{P} \left( \mathcal{E}_1^{(n)}(\delta, T) \right) \rightarrow 1. \tag{B3}$$

Note that statement (ii) is equivalent to  $\mathcal{E}_2^{(n)}(\varepsilon, c_0)$  happening whp for any  $\varepsilon > 0, c_0 > 0$ . We prove this by finding a convenient correspondence of the parameters such that the (known to be whp) event  $\mathcal{E}_1^{(n)}(\delta, T)$  implies  $\mathcal{E}_2^{(n)}(\varepsilon, c_0)$ . We fix  $\varepsilon > 0$  and  $c_0 > 0$  and denote the random function  $x^{(n)}(t) := a_n^{-1}X^{(n)}(t)$ . For some  $\delta$  and  $T$  yet to be chosen, on the event  $\mathcal{E}_1(\delta, T)$ , for each  $t \leq T$ ,

$$f(t) - \delta < x^{(n)}(t) < f(t) + \delta. \tag{B4}$$

Let  $t_1 := f^{(-1)}(c + \delta)$  and  $t_2 := f^{(-1)}(c - \delta)$ . We choose  $T := f^{(-1)}(c_0 - \delta)$ , so that  $t_1, t_2 \leq T$  for any choice of  $c \geq c_0$ . Then, on the event  $\mathcal{E}_1(\delta, T)$ ,

$$x^{(n)}(t_1) > f(t_1) - \delta = c, \tag{B5a}$$

$$x^{(n)}(t_2) < f(t_2) + \delta = c. \tag{B5b}$$

Recall that  $\mathcal{T}^{(n)}(c) = \inf\{t : x^{(n)}(t) \leq c\}$ . Since  $x^{(n)}(t)$  is non-increasing, necessarily

$$f^{(-1)}(c + \delta) = t_1 \leq \mathcal{T}^{(n)}(c) \leq t_2 = f^{(-1)}(c - \delta). \tag{B6}$$

Since  $f^{(-1)}$  is continuous on  $(0, 1]$ , it is uniformly continuous on  $[c_0/2, 1]$ . Hence for our fixed  $\varepsilon$ , we can choose  $0 < \delta = \delta(\varepsilon) < c_0/2$  such that that for any  $s_1, s_2 \in [c_0/2, 1]$ , if  $|s_1 - s_2| < \delta$ , then  $|f^{(-1)}(s_1) - f^{(-1)}(s_2)| < \varepsilon$ . Hence for this choice of  $\delta$ ,

$$f^{(-1)}(c - \delta) < f^{(-1)}(c) + \varepsilon, \tag{B7a}$$

$$f^{(-1)}(c + \delta) > f^{(-1)}(c) - \varepsilon. \tag{B7b}$$

Combining (B6) and (B7), we obtain that on the event  $\mathcal{E}_1(\delta, T)$ ,  $|\mathcal{T}^{(n)}(c) - f^{(-1)}(c)| < \varepsilon$  holds uniformly in  $c \in [c_0, 1]$ . That is, with the chosen  $T$  and  $\delta$ ,  $\mathcal{E}_1^{(n)}(\delta, T)$  implies  $\mathcal{E}_2^{(n)}(\varepsilon, c_0)$  for the given  $\varepsilon, c_0$ . Then

$$1 \geq \mathbb{P} \left( \mathcal{E}_2^{(n)}(\varepsilon, c_0) \right) \geq \mathbb{P} \left( \mathcal{E}_1^{(n)}(\delta, T) \right) \rightarrow 1 \tag{B8}$$

as  $n \rightarrow \infty$ . We conclude that statement (i) implies statement (ii).