Connectivity of a general class of inhomogeneous random digraphs

Junyu Cao
University of California, Berkeley

Mariana Olvera-Cravioto
University of North Carolina, Chapel Hill

July 26, 2019

Abstract

We study a family of directed random graphs whose arcs are sampled independently of each other, and are present in the graph with a probability that depends on the attributes of the vertices involved. In particular, this family of models includes as special cases the directed versions of the Erdős-Rényi model, graphs with given expected degrees, the generalized random graph, and the Poissonian random graph. We establish a phase transition for the existence of a giant strongly connected component and provide some other basic properties, including the limiting joint distribution of the degrees and the mean number of arcs. In particular, we show that by choosing the joint distribution of the vertex attributes according to a multivariate regularly varying distribution, one can obtain scale-free graphs with arbitrary in-degree/out-degree dependence.

Keywords: random digraphs, inhomogeneous random graphs, kernel-based random graphs, scale-free graphs, multi-type branching processes, couplings.
MSC: Primary 05C80; Secondary 90B15, 60C05

1 Introduction

Complex networks appear in essentially all branches of science and engineering, and since the pioneering work of Erdős and Rényi in the early 1960s [13, 14], people from various fields have used random graphs to model, explain and predict some of the properties commonly observed in real-world networks. Until the last decade or so, most of the work had been mainly focused on the study of undirected graphs, however, some important networks, such as the World Wide Web, Twitter, and ResearchGate, to name a few, are directed. The present paper describes a framework for analyzing a large class of directed random graphs, which includes as special cases the directed versions of some of the most popular undirected random graph models.

Specifically, we study directed random graphs where the presence or absence of an arc is independent of all other arcs. This independence among arcs is the basis of the classical Erdős-Rényi model [13, 14], where the presence of an edge is determined by the flip of coin, with all possible edges having the same probability of being present. However, it is well-known that the Erdős-Rényi model tends to produce very homogeneous graphs, that is, where all the vertices have close to the same number of neighbors, a property that is almost never observed in real-world networks. In the undirected setting, a number of models have been proposed to address this problem while preserving the independence among edges. Some of the best known models include the Chung-Lu model [8, 9, 10, 23], the generalized random graph [6, 5, 16], and the Norros-Reittu model or Poissonian random graph [26, 5, 34]. In the undirected case, all of these models were simultaneously studied in [5] under a broader class of graphs, which we will refer to as kernel-based models. In all of these models the inhomogeneity of the degrees is accomplished by assigning to each vertex a *type*, which is used to make the edge probabilities different for each pair of vertices. From a modeling perspective, the types correspond to vertex attributes that influence how likely a vertex is to have neighbors, and inhomogeneity among the types translates into inhomogeneous degrees.

Our proposed family of directed random graphs, which we will refer to as *inhomogeneous random digraphs*, provides a uniform treatment of essentially any model where arcs are present independently of each other, in the same spirit as the work in [5] written for the undirected case. The main results in this paper establish some of the basic properties studied on random graphs, including the expected number of arcs, the joint distribution of the in-degree and out-degree, and the phase transition for the size of the largest strongly connected component. We pay special attention to the so-called *scale-free* property, which states that the tail degree distribution(s) decay according to a power law. Since many real-world directed complex networks exhibit the scale-free property in either their in-degrees, their out-degrees, or both, we provide a theorem stating how the family of random directed graphs studied here can be used to model such networks. Our main result on the connectivity properties of the graphs produced by our model shows that there exists a phase transition, determined by the types, after which the largest strongly connected component contains (with high probability) a positive fraction of all the vertices in the graph, i.e., the graph contains a "giant" strongly connected component.

That the undirected models mentioned above satisfy these basic properties (e.g., scale-free degree distribution, existence of a giant connected component, etc.) constitutes a series of classical results within the random graph literature. Closely related to the results presented here for directed graphs, are the existence of a giant strongly connected component and giant weak-component in the directed configuration model [11, 18, 19], the existence of a giant strongly-connected component in the deterministic directed kernel model with a finite number of types [3], the scale-free property on a directed preferential attachment model [28, 31], and the limiting degree distributions in the directed configuration model [7][1]. From a computational point of view, the work in [33] provides numerical algorithms to identify secondary structures on directed graphs. Our present work includes as a special case the main theorem in [3] and extends it to a larger family of directed random graphs, and it also compiles several results for the number of arcs and the joint distribution of the degrees. It is also worth pointing out that the directed nature of our framework introduces some non-trivial challenges that are not present in the undirected setting, which is the reason we chose to provide a different approach from the one used in [5] for establishing some of our main results. We refer the reader to Section 3.3 for more details on these challenges and what they imply.

The paper is organized as follows. In Section 2 we specify a class of directed random graphs via their arc probabilities, and explain how the models mentioned above fit into this framework. In Section 3 we provide our main results on the basic properties of the graphs produced by our model, and in Section 4 we give all the proofs.

[1]Neither the configuration model nor the preferential attachment model have independent arcs, and therefore fall outside the scope of this paper.

2 The Model

As mentioned in the introduction, we study directed random graphs with independent arcs. Since we are particularly interested in graphs with inhomogeneous degrees, each vertex in the graph will be assigned a *type*, which will determine how large its in-degree and out-degree are likely to be. In applications, the type of a vertex can also be used to model other vertex attributes not directly related to its degrees. We will assume that the types take values in a separable metric space $\mathcal{S}$, which we will refer to as the "type space".

In order to describe our family of directed random graphs, we start by defining the type sequence $\{\mathbf{x}_1^{(n)}, \dots, \mathbf{x}_n^{(n)}\}$, where $\mathbf{x}_i^{(n)}$ denotes the type of vertex i in a graph on the vertex set $[n] := \{1, \dots, n\}$. Note that, depending on how we construct the type sequence, it is possible for $\mathbf{x}_i^{(n)}$ to be different from $\mathbf{x}_i^{(m)}$ for $n \neq m$. Define $G_n(\kappa(1 + \varphi_n))$ to be the graph on the vertex set $[n]$ whose arc probabilities are given by

$$p_{ij}^{(n)} = \left(\frac{\kappa(\mathbf{x}_i^{(n)}, \mathbf{x}_j^{(n)})}{n} (1 + \varphi_n(\mathbf{x}_i^{(n)}, \mathbf{x}_j^{(n)})) \right) \wedge 1, \qquad 1 \leq i \neq j \leq n, \tag{2.1}$$

where κ is a nonnegative function on $\mathcal{S} \times \mathcal{S}$,

$$\varphi_n(\mathbf{x}, \mathbf{y}) = \varphi\left(n, \{\mathbf{x}_k^{(n)} : 1 \leq k \leq n\}, \mathbf{x}, \mathbf{y}\right) > -1 \qquad \text{for all } \mathbf{x}, \mathbf{y} \in \mathcal{S},$$

and $x \wedge y = \min\{x, y\}$ ($x \vee y = \max\{x, y\}$). In other words, $p_{ij}^{(n)}$ denotes the probability that there is an arc from vertex i to vertex j in $G_n(\kappa(1+\varphi_n))$. The presence or absence of arc (i, j) is assumed to be independent of all other arcs. Note that the function $\varphi_n(\mathbf{x}, \mathbf{y})$ may depend on n, on the types of the two vertices involved, or on the entire type sequence; however, to simplify the notation, we emphasize only the arguments $(\mathbf{x}, \mathbf{y})$ of the two types involved. Following the terminology used in [5] and [3], we will refer to κ as the kernel of the graph. Note that we have decoupled the dependence on n and on the type sequence by including it in the term $\varphi_n(\mathbf{x}, \mathbf{y})$, which implies that with respect to the notation used in [5], $\kappa_n(\mathbf{x}, \mathbf{y})$ there corresponds to $\kappa(\mathbf{x}, \mathbf{y})(1 + \varphi_n(\mathbf{x}, \mathbf{y}))$ here.

Throughout the paper, we will refer to any directed random graph generated through our model as an *inhomogeneous random digraph* (IRD).

We end this section by explaining how the directed versions of the Erdős-Rényi graph [13, 14, 15, 4], the Chung-Lu (or "given expected degrees") model [8, 9, 10, 23], the generalized random graph [6, 5, 16], and the Norros-Reittu model (or "Poissonian random graph") [26, 5, 34], as well as the directed deterministic kernel model in [3], fit into our framework. The first four examples fall into the category of so-called rank-1 kernels, where the graph kernel is of the form $\kappa(\mathbf{x}, \mathbf{y}) = \kappa_+(\mathbf{x})\kappa_-(\mathbf{y})$ for some nonnegative continuous functions κ_- and κ_+ on $\mathcal{S}$.

Example 2.1 *Directed versions of some well-known inhomogeneous random graph models. All of them, with the exception of the last one, are defined on the space* $\mathcal{S} = \mathbb{R}_+$ *for a type of the form* $\mathbf{x} = (x^-, x^+)$*, and correspond to rank-1 kernels with* $\kappa_-(\mathbf{x}) = x^-/\sqrt{\theta}$ *and* $\kappa_+(\mathbf{x}) = x^+/\sqrt{\theta}$*, with* $\theta > 0$ *a constant. For convenience, we have dropped the superscript* $^{(n)}$ *from the type sequence, i.e.,* $\{\mathbf{x}_1, \dots, \mathbf{x}_n\} = \{\mathbf{x}_1^{(n)}, \dots, \mathbf{x}_n^{(n)}\}$.

a.) Directed Erdős-Rényi Model: *the arc probabilities are given by*

$$p_{ij}^{(n)} = \lambda/n$$

where λ *is a given constant and* n *is the total number of vertices;* $\varphi_n(\mathbf{x}_i, \mathbf{x}_j) = 0$.

b.) Directed Given Expected Degree Model (Chung-Lu): *the arc probabilities are given by*

$$p_{ij}^{(n)} = \frac{x_i^+ x_j^-}{l_n} \wedge 1,$$

where $l_n = \sum_{i=1}^n (x_i^- + x_i^+)$. *In terms of* (2.1), *it satisfies* $\varphi_n(\mathbf{x}_i, \mathbf{x}_j) = \frac{\theta n - l_n}{l_n}$, *where* $\theta = \lim_{n\to\infty} l_n/n$.

c.) Generalized Directed Random Graph: *the arc probabilities are given by*

$$p_{ij}^{(n)} = \frac{x_i^+ x_j^-}{l_n + x_i^+ x_j^-},$$

which implies that $\varphi_n(\mathbf{x}_i, \mathbf{x}_j) = \frac{\theta n - l_n - x_i^+ x_j^-}{l_n + x_i^+ x_j^-}$, *with* l_n *and* θ *defined as above.*

d.) Directed Poissonian Random Graph (Norros-Reittu): *the arc probabilities are given by*

$$p_{ij}^{(n)} = 1 - e^{-x_i^+ x_j^- / l_n},$$

which implies that $\varphi_n(\mathbf{x}_i, \mathbf{x}_j) = \left(n\theta(1 - e^{-x_i^+ x_j^- / l_n}) - x_i^+ x_j^- \right) / (x_i^+ x_j^-)$, *with* l_n *and* θ *defined as above.*

e.) Deterministic Kernel Model: *the arc probabilities are given by*

$$p_{ij}^{(n)} = \frac{\kappa(\mathbf{x}_i, \mathbf{x}_j)}{n} \wedge 1,$$

for a finite type space $\mathcal{S} = \{\mathbf{s}_1, \ldots, \mathbf{s}_M\}$, *and a strictly positive function* κ *on* $\mathcal{S} \times \mathcal{S}$; *in terms of* (2.1), $\varphi_n(\mathbf{x}_i, \mathbf{x}_j) = 0$. *This model is also known as the stochastic block model.*

3 Main Results

We now present our main results for the family of inhomogeneous random digraphs defined through (2.1). As mentioned in the introduction, we focus on establishing some of the basic properties of this family, including the distribution of the degrees, the mean number of arcs, and the size of the largest strongly connected component. When analyzing the degree distributions, we specifically explain how to obtain the scale-free property under degree-degree correlations.

As mentioned in the previous section, we assume throughout the paper that the nth graph in the sequence is constructed using the types $\{\mathbf{X}_1, \ldots, \mathbf{X}_n\} = \{\mathbf{X}_1^{(n)}, \ldots, \mathbf{X}_n^{(n)}\}$, where we will often drop the superscript $^{(n)}$ to simplify the notation. From now on we will use upper case letters to emphasize the possibility that the $\{\mathbf{X}_i\}$ may themselves be generated through a random process. To distinguish between these two levels of randomness, let P be a probability measure on a space large enough to construct all the type sequences $\{\{\mathbf{X}_i^{(n)}, 1 \leq i \leq n\} : n \geq 1\}$, as well as the random graphs $G_n(\kappa(1+\varphi_n))$, simultaneously. Define $\mathscr{F} = \sigma(\mathbf{X}_i^{(n)}, 1 \leq i \leq n)$ and the corresponding conditional probability and expectation $\mathbb{P}(\cdot) = P(\cdot|\mathscr{F})$ and $\mathbb{E}[\cdot] = E[\cdot|\mathscr{F}]$, respectively.

Our first assumption will be to ensure that the $\{\mathbf{X}_i^{(n)}\}$ converge in distribution under the unconditional probability P. As is to be expected from the work in [5] for the undirected case, we will also need to impose some regularity conditions on the kernel κ, as well as on the function φ_n. Our main assumptions are summarized below.

Assumption 3.1 *a.) There exists a Borel probability measure μ on $\mathcal{S}$ such that for any μ-continuity set $A \subseteq \mathcal{S}$,*

$$\mu_n(A) := \frac{1}{n}\sum_{i=1}^{n} 1(\mathbf{X}_i^{(n)} \in A) \xrightarrow{P} \mu(A) \qquad n \to \infty,$$

where $\xrightarrow{P}$ denotes convergence in probability. Note that μ_n is a random probability measure, whereas μ is not random.

b.) κ is nonnegative and continuous a.e. on $\mathcal{S} \times \mathcal{S}$.

c.) For any sequences $\{\mathbf{x}_n\}, \{\mathbf{y}_n\} \subseteq \mathcal{S}$ such that $\mathbf{x}_n \to \mathbf{x}$ and $\mathbf{y}_n \to \mathbf{y}$ as $n \to \infty$, we have $\varphi_n(\mathbf{x}_n, \mathbf{y}_n) \xrightarrow{P} 0$ as $n \to \infty$.

d.) The following limits hold :

$$\lim_{n\to\infty} \frac{1}{n^2} E\left[\sum_{i=1}^{n}\sum_{j=1}^{n} \kappa(\mathbf{X}_i^{(n)}, \mathbf{X}_j^{(n)})\right] = \lim_{n\to\infty} \frac{1}{n} E\left[\sum_{i=1}^{n}\sum_{j\neq i} p_{ij}^{(n)}\right] = \iint_{\mathcal{S}^2} \kappa(\mathbf{x},\mathbf{y})\, \mu(d\mathbf{x})\mu(d\mathbf{y}) < \infty.$$

Remark 3.2 *The pair $(\mathcal{S}, \mu)$, where $\mathcal{S}$ is a separable metric space and μ is a Borel probability measure, is referred to in [5] as a* generalized ground space. *For convenience, we will adopt the same terminology here. Throughout the paper, we use "a.e." to mean "almost everywhere with respect to the (non-random) measure μ".*

3.1 Number of arcs

Our assumption that the types $\{\mathbf{X}_i\}$ converge in distribution as the size of the graph grows implies that the graphs produced by our model are sparse, in the sense that the mean number of arcs is of the same order as the number of vertices. Our first result provides an expression for the exact ratio between the number of arcs and the number of vertices.

Proposition 3.3 *Define $e(G_n(\kappa(1+\varphi_n)))$ to be the number of arcs in $G_n(\kappa(1+\varphi_n))$. Then, under Assumption 3.1(a)-(d) we have*

$$\frac{1}{n} e(G_n(\kappa(1+\varphi_n))) \longrightarrow \iint_{\mathcal{S}^2} \kappa(\mathbf{x},\mathbf{y})\, \mu(d\mathbf{x})\mu(d\mathbf{y}) \qquad \text{in } L^1(P)$$

as $n \to \infty$.

3.2 Distribution of vertex degrees

We now move on to describing the vertex degree distribution, which is best accomplished by looking at the properties of a typical vertex, i.e., one chosen uniformly at random. In particular, if $D_{n,i}^-$ and $D_{n,i}^+$ denote the in-degree and out-degree, respectively, of vertex $i \in [n]$, and we let ξ be a uniform random variable in $\{1, 2, \dots, n\}$, then we study the distribution of $(D_{n,\xi}^-, D_{n,\xi}^+)$. We point out that the distribution of $(D_{n,\xi}^-, D_{n,\xi}^+)$ also allows us to compute the proportion of vertices in the graph having in-degree k and out-degree l for any $k, l \geq 0$. In the sequel, $\Rightarrow$ denotes weak convergence with respect to P.

Theorem 3.4 *Under Assumption 3.1 we have*

$$\left(D_{n,\xi}^-, D_{n,\xi}^+\right) \Rightarrow (Z^-, Z^+), \quad E[D_{n,\xi}^\pm] \to E[Z^\pm], \qquad \text{as } n \to \infty,$$

where Z^- and Z^+ are conditionally independent (given $\mathbf{X}$) mixed Poisson random variables with mixing parameters

$$\lambda_-(\mathbf{X}) := \int_{\mathcal{S}} \kappa(\mathbf{y}, \mathbf{X})\, \mu(d\mathbf{y}) \quad \text{and} \quad \lambda_+(\mathbf{X}) := \int_{\mathcal{S}} \kappa(\mathbf{X}, \mathbf{y})\, \mu(d\mathbf{y}),$$

respectively, and $\mathbf{X}$ is distributed according to μ.

As mentioned earlier, we are particularly interested in models capable of creating scale-free graphs, perhaps with a significant correlation between the in-degree and out-degree of the same vertex. To see that our family of inhomogeneous random digraphs can accomplish this, we first introduce the notion of non-standard regular variation (see [28, 31]), which extends the definition of regular variation on the real line to multiple dimensions, with each dimension having potentially different tail indexes. In our setting we only need to consider two dimensions, so we only give the bivariate version of the definition.

Definition 3.5 *A nonnegative random vector $(X, Y) \in \mathbb{R}^2$ has a distribution that is non-standard regularly varying if there exist scaling functions $a(t) \nearrow \infty$ and $b(t) \nearrow \infty$ and a non-zero limit measure $\nu(\cdot)$, called the limit or tail measure, such that*

$$tP\left((X/a(t), Y/b(t)) \in \cdot\right) \xrightarrow{v} \nu(\cdot), \qquad t \to \infty,$$

where $\xrightarrow{v}$ denotes vague convergence of measures in $M_+([0,\infty]^2 \backslash \{\mathbf{0}\})$, the space of Radon measures on $[0,\infty]^2 \setminus \{\mathbf{0}\}$.

In particular, if the scaling functions $a(t)$ and $b(t)$ are regularly varying at infinity with indexes $1/\alpha$ and $1/\beta$, respectively, that is $a(t) = t^{1/\alpha} L_a(t)$ and $b(t) = t^{1/\beta} L_b(t)$ for some $\alpha, \beta > 0$ and slowly varying functions L_a and L_b, then the marginal distributions $P(X > t)$ and $P(Y > t)$ are regularly varying with tail indexes $-\alpha$ and $-\beta$, respectively (see Theorem 6.5 in [29]). Throughout the paper we use the notation $\mathcal{R}_\alpha$ to denote the family of regularly varying functions with index α.

To see how our family of IRDs can be used to model complex networks where both the in-degrees and the out-degrees possess the scale-free property, perhaps with different tail indexes, we give a theorem stating that the non-standard regular variation of the limiting degrees (Z^-, Z^+) follows from that of the vector $(\lambda_-(\mathbf{X}), \lambda_+(\mathbf{X}))$. Moreover, for the models (a)-(d) in Example 2.1, we have

$$(\lambda_-(\mathbf{X}), \lambda_+(\mathbf{X})) = \left(\kappa_-(\mathbf{X}) \int_{\mathcal{S}} \kappa_+(\mathbf{y})\mu(d\mathbf{y}),\, \kappa_+(\mathbf{X}) \int_{\mathcal{S}} \kappa_-(\mathbf{y})\mu(d\mathbf{y}) \right) = \left(cX^-,\, (1-c)X^+ \right),$$

where $c = E[X^+]/\theta$ and $\theta = E[X^- + X^+]$, so the non-standard regular variation of (Z^-, Z^+) can be easily obtained by choosing a non-standard regularly varying type distribution μ.

Theorem 3.6 *Let $\mathbf{X}$ denote a random vector in the type space $\mathcal{S}$ distributed according to μ. Suppose that μ is such that $(\lambda_-(\mathbf{X}), \lambda_+(\mathbf{X}))$ is non-standard regularly varying with scaling functions $a(t) \in \mathcal{R}_{1/\alpha}$ and $b(t) \in \mathcal{R}_{1/\beta}$ and limiting measure $\nu(\cdot)$. Then, (Z^-, Z^+) is non-standard regularly varying with scaling functions $a(t)$ and $b(t)$ and limiting measure $\nu(\cdot)$ as well.*

To illustrate our result, we give below an example that shows how our family of random digraphs along with Theorem 3.6 can be used to model real-world networks.

Example 3.7 *As discussed in [36], many real-world networks exhibit both heavy-tailed in-degrees and heavy-tailed out-degrees. In many of those cases there also appears to be a relationship between the vertices with very high in-degrees and those with very high out-degrees, as is shown in [36] for portions of the Web graph and the English Wikipedia graph (this dependence was computed using the angular measure in [36]). Suppose we want to model such graphs using an inhomogeneous random digraph. Interesting levels of dependence ranging from the case where the in-degree and out-degree are independent to where they are essentially the same can be obtained by choosing $\mathbf{X} = (X^-, X^+)$, $P(X^- > x) \sim k_- x^{-\alpha}$ as $x \to \infty$ and $X^+ = r(X^-)^\gamma + (1-r)Y$, where Y is independent of X^- and satisfies $P(Y > y) \sim k' y^{-\beta}$, $\alpha, \beta, k_-, k' > 0$, $r \in [0,1]$ and $0 \le \gamma \le \alpha/\beta$. This choice leads to $P(X^+ > x) \sim k_+ x^{-\beta}$ for some other constant $k_+ > 0$, and covers the independent case when $r = 0$, and the perfectly dependent case when $r = 1$ and $\gamma = \alpha/\beta$. Now choose $\kappa(\mathbf{x}, \mathbf{y}) = x^+ y^-$ and note that $(\lambda_-(\mathbf{X}), \lambda_+(\mathbf{X})) = (cX^-, (1-c)X^+)$, where $c = E[X^+]/E[X^- + X^+]$. It follows from Theorems 3.4 and 3.6 that $(D^-_{n,\xi}, D^+_{n,\xi}) \Rightarrow (Z^-, Z^+)$ as $n \to \infty$, where (Z^+, Z^-) is non-standard regularly varying. In particular, $P(Z^- > z) \sim k_- c^\alpha z^{-\alpha}$ and $P(Z^+ > z) \sim k_+(1-c)^\beta z^{-\beta}$ as $z \to \infty$, and the angular measure between Z^- and Z^+ will mimic that of X^- and X^+.*

3.3 Phase transition for the largest strongly connected component

Our last result in the paper establishes a phase transition for the existence of a giant strongly connected component in $G_n(\kappa(1+\varphi_n))$. That is, we provide a critical threshold for a functional of the kernel κ and the type distribution μ, such that above this threshold the graph will have a giant strongly connected component with high probability, and below it will not. Before stating the corresponding theorem, we give a brief overview of some basic definitions.

For any two vertices i, j in the graph, we say that there is a directed path from i to j if the graph contains a set of arcs $\{(i, k_1), (k_1, k_2), \ldots, (k_t, j)\}$ for some $t \ge 0$. A set of vertices $V \subseteq [n]$ is *strongly connected*, if for any two vertices $i, j \in V$ we have that there exists a directed path from i to j and one from j to i. Moreover, we say that a *giant* strongly connected component exists for our family of random digraphs if $\liminf_{n\to\infty} |\mathcal{C}_1(G_n(\kappa(1+\varphi_n)))|/n > \epsilon$ for some $\epsilon > 0$, where $\mathcal{C}_1(G_n(\kappa(1+\varphi_n))$ is the largest strongly connected component of $G_n(\kappa(1+\varphi_n))$ and $|A|$ denotes the cardinality of set A.

For undirected graphs, the phase transition for the Erdős-Rényi model ($p_{ij}^{(n)} = \lambda/n$ for some $\lambda > 0$) dates back to the classical work of Erdős and Rényi in [14], where the threshold for the existence of a giant connected component is $\lambda = 1$. The critical case, i.e., $\lambda = 1$, was studied in [22] using edge probabilities of the form $p_{ij}^{(n)} = (1 + cn^{-1/3})/n$ for some $c > 0$, in which case the size of the largest connected component was shown to be of order $n^{2/3}$. Somewhat unrelated, the corresponding phase transition was established for the (undirected) configuration model in [25], where the threshold was shown to be $E[D(D-1)]/E[D] = 1$, with D distributed according to the limiting degree distribution (as the number of vertices grows to infinity). Back to the (undirected) inhomogeneous random graph setting, i.e., $p_{ij}^{(n)} = \kappa(\mathbf{x}_i, \mathbf{x}_j)(1 + \varphi_n(\mathbf{x}_i, \mathbf{x}_j))/n$ with κ symmetric, the phase transition was first proven for various forms of rank-1 kernels. In particular, Chung and Lu established in [9] the phase transition for the existence of a giant connected component in the so-called "given expected degree" model. The same authors also give in [8] a phase transition for the average distance between vertices when the type distribution μ follows a power-law. Norros and Reittu proved the phase transition for the existence of a giant connected component for the Poissonian random graph in [26], along with a characterization of the distance between two randomly chosen vertices, and Riordan proved it in [30] for the $c/\sqrt{ij}$ model, which is equivalent to the rank-1 kernel $\kappa(\mathbf{x}, \mathbf{y}) = \psi(\mathbf{x})\psi(\mathbf{y})$ with $\psi(\mathbf{x}) = \sqrt{c\mathbf{x}}$ and μ the distribution of a Pareto(2,1). More generally, the work in [5] gives the phase transition for the giant connected component in the general kernel case, along with some other properties (e.g., second largest connected component, distances between vertices, and stability). The threshold for the existence of a giant connected component is $\|T_\kappa\|_{op} = 1$, with $\|\cdot\|_{op}$ the operator norm[2], where T_κ is a linear operator induced by κ, which in the rank-1 case becomes $\|T_\kappa\|_{op}^2 = E[\psi(\mathbf{X})^2] = 1$, with $\mathbf{X}$ distributed according to μ.

For the directed case, the phase transition for the existence of a giant strongly connected component was proven for the directed Erdős-Rényi model ($p_{ij}^{(n)} = \lambda/n$ for some $\lambda > 0$) in [17] and for the "given number of arcs" version of the Erdős-Rényi model (number of arcs $= \lambda n$ for some $\lambda > 0$) in [20], with the threshold being $\lambda = 1$. The work in [21] studies a related model where each vertex i can have three types of arcs: up arcs for $j > i$, down arcs for $j < i$, and bidirectional arcs, and proved the corresponding phase transition for the appearance of a giant strongly connected component. For the directed configuration model the phase transition for the existence of a giant strongly connected component was given in [16] under the assumption that the limiting degrees have finite variance and satisfy some additional conditions on the growth of the maximum degree, and can also be indirectly obtained from the results in [35] under only finite covariance between the in-degree and out-degree. The threshold for the directed configuration model is $E[D^-D^+]/E[D^- + D^+] = 1$, where (D^-, D^+) are the limiting in-degree and out-degree. A hybrid model where the out-degree has a general distribution with finite mean and the destinations of the arcs are selected uniformly at random among the vertices (which gives Poisson in-degrees) was studied in [27] and was shown to have a phase transition at $E[D^+] = 1$. Finally, for general inhomogeneous random digraphs such as those studied here, the main theorem in [3] establishes the phase transition for the deterministic kernel in Example 2.1(d) with finite type space $\mathcal{S} = \{1, 2, \ldots, M\}$, without characterizing the strict

[2] $\|T\|_{op} := \sup\{\|Tf\|_2 : f \geq 0, \|f\|_2 \leq 1\}$ and $\|f\|_2^2 = \int_{\mathcal{S}} f(\mathbf{x})^2 \mu(d\mathbf{x})$.

positivity of the survival probability. The authors in [3] also suggest that the general case can be obtained using the same techniques used in [5] to go from a finite type space to the general one, however, the proof in [5] requires a critical step that does not hold for directed graphs; see Section 4.3 for more details.

Our Theorem 3.10 provides the full equivalent of the main theorem in [5] (Theorem 3.1) for the directed case, and its proof is based on a coupling argument between the exploration of both the inbound and outbound components of a randomly chosen vertex and a double multi-type branching process with a finite number of types. Our approach differs from that of [5], done for undirected graphs, in the order in which the couplings are done, and it leverages on the main theorem in [3] to obtain a lower bound for the size of the strongly connected component. We give more details on how our proof technique compares to that used in [5] in Section 4.3.

As in the undirected case, the size of the largest strongly connected component is related to the survival probability of a suitably constructed double multi-type branching process. To define it, let $\mathcal{T}_\mu^-(\kappa)$ and $\mathcal{T}_\mu^+(\kappa)$ denote two conditionally independent (given their common root) multi-type branching processes defined on the type space $\mathcal{S}$ whose roots are chosen according to μ and such that the number of offspring having types in a subset $A \subseteq \mathcal{S}$ that an individual of type $\mathbf{x} \in \mathcal{S}$ can have, is Poisson distributed with means

$$\int_A \kappa(\mathbf{y}, \mathbf{x})\mu(d\mathbf{y}) \quad \text{for } \mathcal{T}_\mu^-(\kappa) \qquad \text{and} \qquad \int_A \kappa(\mathbf{x}, \mathbf{y})\mu(d\mathbf{y}) \quad \text{for } \mathcal{T}_\mu^+(\kappa), \tag{3.1}$$

respectively. Next, let $\rho_-(\kappa; \mathbf{x})$ and $\rho_+(\kappa; \mathbf{x})$ denote the survival probabilities of $\mathcal{T}_\mu^-(\kappa; \mathbf{x})$ and $\mathcal{T}_\mu^+(\kappa; \mathbf{x})$, respectively, where $\mathcal{T}_\mu^-(\kappa; \mathbf{x})$ and $\mathcal{T}_\mu^+(\kappa; \mathbf{x})$ denote the trees whose root has type $\mathbf{x}$. We recall that a branching process is said to survive if its total population is infinite. We refer the reader to [24, 2] for more details on multi-type branching processes, including those with uncountable type spaces as the ones defined above.

In order to state our result for the phase transition in IRDs we first need to introduce the following definitions.

Definition 3.8 *A kernel κ defined on a separable metric space $\mathcal{S}$ with respect to a Borel probability measure μ is said to be* irreducible *if for any subset $A \subseteq \mathcal{S}$ satisfying $\kappa = 0$ a.e. on $A \times A^c$, we have either $\mu(A) = 0$ or $\mu(A^c) = 0$. We say that κ is* quasi-irreducible *if there is a μ-continuity set $\mathcal{S}' \subseteq \mathcal{S}$ with $\mu(\mathcal{S}') > 0$ such that the restriction of κ to $\mathcal{S}' \times \mathcal{S}'$ is irreducible, and $\kappa(\mathbf{x}, \mathbf{y}) = 0$ if $\mathbf{x} \notin \mathcal{S}'$ or $\mathbf{y} \notin \mathcal{S}'$.*

Definition 3.9 *A kernel κ on a separable metric space $\mathcal{S}$ with respect to a Borel probability measure μ is* regular finitary *if $\mathcal{S}$ has a finite partition into sets $\mathcal{J}_1, ..., \mathcal{J}_r$ such that κ is constant on each $\mathcal{J}_i \times \mathcal{J}_j$, and each $\mathcal{J}_i$ is a μ-continuity set, i.e., it is measurable and has $\mu(\partial \mathcal{J}_i) = 0$.*

To give the condition under which a giant strongly connected component exists we also need to define the operators induced by kernel κ, i.e.,

$$T_\kappa^+ f(\mathbf{x}) = \int_\mathcal{S} \kappa(\mathbf{x}, \mathbf{y}) f(\mathbf{y})\mu(d\mathbf{y}) \qquad \text{and} \qquad T_\kappa^- f(\mathbf{x}) = \int_\mathcal{S} \kappa(\mathbf{y}, \mathbf{x}) f(\mathbf{y})\mu(d\mathbf{y}).$$

Note that T_κ^+ and T_κ^- are integral linear operators on $(\mathcal{S}, \mu)$ equipped with the norm

$$\left\|T_\kappa^\pm\right\|_{op} = \sup\{\left\|T_\kappa^\pm f\right\|_2 : f \geq 0, \, \|f\|_2 \leq 1\} \leq \infty,$$

which makes them (potentially) unbounded operators in $L^2(\mathcal{S}, \mu)$. We also define their corresponding spectral radii $r(T_\kappa^+)$ and $r(T_\kappa^-)$, where the spectral radius of operator T in $L^2(\mathcal{S}, \mu)$ is defined as

$$r(T) = \sup\{|\lambda| : \lambda \in \sigma(T)\},$$

where $\sigma(T) = \{\lambda \in \mathbb{C} : T - \lambda I \text{ is not boundedly invertible}\}$ is the spectrum of T and I is the operator that maps f onto itself.[3]

The phase transition result for the largest strongly connected component is given below.

Theorem 3.10 *Suppose Assumption 3.1 is satisfied and κ is irreducible. Let $\mathcal{C}_1(G_n(\kappa(1+\varphi_n)))$ denote the largest strongly connected component of $G_n(\kappa(1+\varphi_n))$. Then,*

$$\frac{|\mathcal{C}_1(G_n(\kappa(1+\varphi_n)))|}{n} \xrightarrow{P} \rho(\kappa) \qquad n \to \infty,$$

where

$$\rho(\kappa) = \int_{\mathcal{S}} \rho_-(\kappa; \mathbf{x}) \rho_+(\kappa; \mathbf{x}) \mu(d\mathbf{x}).$$

Furthermore, if $\rho(\kappa) > 0$ then $r(T_\kappa^-) > 1$ and $r(T_\kappa^+) > 1$, and if there exists a regular finitary quasi-irreducible kernel $\tilde{\kappa}$ such that $\tilde{\kappa} \leq \kappa$ a.e. and $r(T_{\tilde{\kappa}}^-) > 1$ (equivalently, $r(T_{\tilde{\kappa}}^+) > 1$), then $\rho(\kappa) > 0$.

Moreover, when $\rho(\kappa) > 0$ we can characterize the "bow-tie" structure defined by the giant strongly connected component, $\mathcal{C}_1(G_n(\kappa(1+\varphi_n)))$, the set of vertices that can reach it (its fan-in), and the set of vertices that can be reached from it (its fan-out). The following result makes this precise.

Theorem 3.11 *Suppose Assumption 3.1 is satisfied and κ is irreducible. For each vertex $v \in [n]$ define its in-component and out-component as:*

$$R^-(v) = \{i \in [n] : v \text{ is reachable from } i \text{ by a directed path in } G_n(\kappa(1+\varphi_n))\}$$
$$R^+(v) = \{i \in [n] : i \text{ is reachable from } v \text{ by a directed path in } G_n(\kappa(1+\varphi_n))\}.$$

Define $L_n^- = \{v \in [n] : |R^-(v)| \geq (\log n)/n\}$ and $L_n^+ = \{v \in [n] : |R^+(v)| \geq (\log n)/n\}$. Then, if $\rho(\kappa) > 0$,

$$\lim_{n\to\infty} P\left(\mathcal{C}_1(G_n(\kappa(1+\varphi_n))) = L_n^+ \cap L_n^-\right) = 1,$$

and

$$\frac{|L_n^+|}{n} \xrightarrow{P} \int_{\mathcal{S}} \rho_+(\kappa; \mathbf{x}) \mu(d\mathbf{x}) \qquad \text{and} \qquad \frac{|L_n^-|}{n} \xrightarrow{P} \int_{\mathcal{S}} \rho_-(\kappa; \mathbf{x}) \mu(d\mathbf{x})$$

as $n \to \infty$.

[3] If T is not closed then $\sigma(T) = \mathbb{C}$, and therefore, $r(T) = \infty$ (see section 21 in [1]).

Remark 3.12 *We point out that we do not have a full if and only if condition for the strict positivity of* $\rho(\kappa)$*, since our operators* T_κ^- *and* T_κ^+ *may be unbounded, in which case the continuity of the spectral radius is not guaranteed. However, when* κ *satisfies*

$$\int_{\mathcal{S}}\int_{\mathcal{S}} \kappa(x,y)^2 \mu(d\mathbf{x})\mu(d\mathbf{y}) < \infty,$$

then the operators T_κ^- *and* T_κ^+ *are compact (see Lemma 5.15 in [5]), and Theorem 2.1(a) in [12] gives the continuity of the spectral radius for a sequence of quasi-irreducible kernels* $\kappa_m \nearrow \kappa$ *as* $m \to \infty$*, ensuring the existence of* $\tilde{\kappa}$ *in Theorem 3.10. Interestingly, for the rank-1 case we can indeed provide a full characterization even when the operators* T_κ^- *and* T_κ^+ *are unbounded, as Proposition 3.13 shows.*

We end the expository part of the paper with a compilation of all our results for the rank-1 case, which includes the first four models in Example 2.1.

Proposition 3.13 (IRDs with rank-1 kernel) *Suppose that Assumption 3.1 is satisfied with* κ *irreducible and of the form* $\kappa(\mathbf{x},\mathbf{y}) = \kappa_+(\mathbf{x})\kappa_-(\mathbf{y})$*. Let* $\mathbf{X}$ *denote a random variable distributed according to* μ*. Then, the following properties hold:*

a.) **Number of arcs:** *let* $e(G_n(\kappa(1+\varphi_n)))$ *denote the number of arcs in* $G_n(\kappa(1+\varphi_n))$*, then*

$$\frac{e(G_n(\kappa(1+\varphi_n)))}{n} \to E[\kappa_-(\mathbf{X})]E[\kappa_+(\mathbf{X})] \quad \text{in } L^1(P) \qquad \text{as } n \to \infty.$$

b.) **Distribution of vertex degrees:** *let* $(D_{n,\xi}^-, D_{n,\xi}^+)$ *denote the in-degree and out-degree of a randomly chosen vertex in* $G_n(\kappa(1+\varphi_n))$*. Set* $\lambda_+(\mathbf{x}) = \kappa_+(\mathbf{x})E[\kappa_-(\mathbf{X})]$ *and* $\lambda_-(\mathbf{x}) = \kappa_-(\mathbf{x})E[\kappa_+(\mathbf{X})]$*. Then,*

$$(D_{n,\xi}^-, D_{n,\xi}^+) \Rightarrow (Z^-, Z^+), \qquad E[D_{n,\xi}^\pm] \to E[Z^\pm],$$

as $n \to \infty$*, where* Z^- *and* Z^+ *are conditionally independent (given* $\mathbf{X}$*) mixed Poisson random variables with mixing parameters* $\lambda_-(\mathbf{X})$ *and* $\lambda_+(\mathbf{X})$*.*

c.) **Scale-free degrees:** *suppose that* $(\kappa_-(\mathbf{X}), \kappa_+(\mathbf{X}))$ *is non-standard regularly varying with scaling functions* $a(t) \in \mathcal{RV}(1/\alpha)$ *and* $b(t) \in \mathcal{RV}(1/\beta)$ *and limiting measure* $\tilde{\nu}(\cdot)$*. Then,* (Z^-, Z^+) *is non-standard regularly varying with scaling functions* $a(t)$ *and* $b(t)$ *and limiting measure* $\nu(\cdot)$ *satisfying*

$$\nu((x,\infty] \times (y,\infty]) = \tilde{\nu}\left(\left(\frac{x}{E[\kappa_-(\mathbf{X})]}, \infty \right] \times \left(\frac{y}{E[\kappa_+(\mathbf{X})]}, \infty \right] \right).$$

d.) **Phase transition for the largest strongly connected component:** *suppose* κ *is irreducible and let* $\mathcal{C}_1(G_n(\kappa(1+\varphi_n)))$ *denote the largest strongly connected component of* $G_n(\kappa(1+\varphi_n))$*. Then,*

$$\frac{|\mathcal{C}_1(G_n(\kappa(1+\varphi_n)))|}{n} \xrightarrow{P} \rho(\kappa), \qquad n \to \infty,$$

with $\rho(\kappa) > 0$ *if and only if* $E[\kappa_+(\mathbf{X})\kappa_-(\mathbf{X})] > 1$*.*

The remainder of the paper is devoted to the proofs of all the results mentioned above.

4 Proofs

This section contains all the proofs of the theorems in Section 3. They are organized according to the order in which their corresponding statements appear. Throughout this section we use the notation

$$q_{ij}^{(n)} = \frac{\kappa(\mathbf{X}_i, \mathbf{X}_j)}{n} \qquad 1 \leq i, j \leq n,$$

to denote the asymptotic limit of the arc probabilities in the graph, and to avoid having to explicitly exclude possible self-loops, we define $p_{ii}^{(n)} = 0$ for all $1 \leq i \leq n$. We also use $f(x) = O(g(x))$ as $x \to \infty$ to mean that $\limsup_{x\to\infty} |f(x)/g(x)| < \infty$.

4.1 Number of Arcs

The first result we prove corresponds to Proposition 3.3, which gives the asymptotic number of edges in $G_n(\kappa(1+\varphi_n))$. Before we do so, we state and prove two preliminary technical lemmas that will be used several times throughout the paper.

Lemma 4.1 *Assume Assumption 3.1 holds and define for any* $0 < \epsilon < 1/2$ *the events*

$$B_{ij} = \left\{ (1-\epsilon) q_{ij}^{(n)} \leq p_{ij}^{(n)} \leq (1+\epsilon) q_{ij}^{(n)},\, q_{ij}^{(n)} \leq \epsilon \right\}. \tag{4.1}$$

Then,

$$\lim_{n\to\infty} E\left[\frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n \left(p_{ij}^{(n)} + q_{ij}^{(n)} \right) 1(B_{ij}^c) \right] = 0.$$

Proof. We start by defining $A_{ij} = \{q_{ij}^{(n)} \leq \epsilon\}$ and noting that the expression inside the expectation is bounded from above by

$$\frac{2}{n} \sum_{i=1}^n \sum_{j=1}^n q_{ij}^{(n)} 1\left(p_{ij}^{(n)} < (1-\epsilon) q_{ij}^{(n)}, A_{ij} \right) + \frac{2}{n} \sum_{i=1}^n \sum_{j=1}^n p_{ij}^{(n)} 1\left(p_{ij}^{(n)} > (1+\epsilon) q_{ij}^{(n)},\, A_{ij} \right) \tag{4.2}$$

$$+ \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n (1 + q_{ij}^{(n)}) 1(A_{ij}^c). \tag{4.3}$$

To show that (4.3) converges to zero, let $\mathbf{X}^{(n)} = \mathbf{X}_I$ and $\mathbf{Y}^{(n)} = \mathbf{Y}_J$ where I and J are mutually independent and uniformly distributed in $\{1, \cdots, n\}$, and independent of everything else. Note that

$$\begin{aligned} \frac{1}{n} E\left[\sum_{i=1}^n \sum_{j=1}^n (1 + q_{ij}^{(n)}) 1(A_{ij}^c) \right] &\leq \frac{1}{n} E\left[\sum_{i=1}^n \sum_{j=1}^n (\epsilon^{-1} + 1) q_{ij}^{(n)} 1(A_{ij}^c) \right] \\ &= (\epsilon^{-1} + 1) E\left[\kappa(\mathbf{X}^{(n)}, \mathbf{Y}^{(n)}) 1(\kappa(\mathbf{X}^{(n)}, \mathbf{Y}^{(n)}) > \epsilon n) \right]. \end{aligned}$$

Note that Assumption 3.1(a)-(b) imply that $\kappa(\mathbf{X}^{(n)}, \mathbf{Y}^{(n)}) \Rightarrow \kappa(\mathbf{X}, \mathbf{Y})$ as $n \to \infty$, where $\mathbf{X}$ and $\mathbf{Y}$ are i.i.d. with distribution μ. Moreover, Assumption 3.1(d) gives $E[\kappa(\mathbf{X}^{(n)}, \mathbf{Y}^{(n)})] \to E[\kappa(\mathbf{X}, \mathbf{Y})]$ as $n \to \infty$. Hence, we can construct $(\{\mathbf{X}^{(n)}, \mathbf{Y}^{(n)}\}_{n\geq 1}, \mathbf{X}, \mathbf{Y})$ on a common probability space such that $(\mathbf{X}^{(n)}, \mathbf{Y}^{(n)}) \to (\mathbf{X}, \mathbf{Y})$ P-a.s. and $\kappa(\mathbf{X}^{(n)}, \mathbf{Y}^{(n)}) \to \kappa(\mathbf{X}, \mathbf{Y})$ P-a.s. Fatou's lemma then gives

$$\begin{aligned}
&\limsup_{n\to\infty} E\left[\kappa(\mathbf{X}^{(n)}, \mathbf{Y}^{(n)}) 1(\kappa(\mathbf{X}^{(n)}, \mathbf{Y}^{(n)}) > \epsilon n)\right] \\
&= \lim_{n\to\infty} E\left[\kappa(\mathbf{X}^{(n)}, \mathbf{Y}^{(n)})\right] - \liminf_{n\to\infty} E\left[\kappa(\mathbf{X}^{(n)}, \mathbf{Y}^{(n)}) 1(\kappa(\mathbf{X}^{(n)}, \mathbf{Y}^{(n)}) \leq \epsilon n)\right] \\
&\leq E\left[\kappa(\mathbf{X}, \mathbf{Y})\right] - E\left[\kappa(\mathbf{X}, \mathbf{Y})\right] = 0.
\end{aligned}$$

To analyze the expectation of the first sum in (4.2), note that

$$\begin{aligned}
&\frac{1}{n} E\left[\sum_{i=1}^n \sum_{j=1}^n q_{ij}^{(n)} 1\left(p_{ij}^{(n)} < (1-\epsilon) q_{ij}^{(n)}, A_{ij}\right)\right] \\
&= \frac{1}{n} E\left[\sum_{i=1}^n \sum_{j=1}^n q_{ij}^{(n)} 1\left(q_{ij}^{(n)}(1+\varphi_n(\mathbf{X}_i, \mathbf{X}_j)) < (1-\epsilon) q_{ij}^{(n)} \leq \epsilon(1-\epsilon)\right)\right] \\
&\leq \frac{1}{n^2} E\left[\sum_{i=1}^n \sum_{j=1}^n \kappa(\mathbf{X}_i, \mathbf{X}_j) 1\left(\varphi_n(\mathbf{X}_i, \mathbf{X}_j) < -\epsilon\right)\right] \\
&= E\left[\kappa(\mathbf{X}^{(n)}, \mathbf{Y}^{(n)})\right] - E\left[\kappa(\mathbf{X}^{(n)}, \mathbf{Y}^{(n)}) 1(\varphi_n(\mathbf{X}^{(n)}, \mathbf{Y}^{(n)}) \geq -\epsilon)\right]. \qquad (4.4)
\end{aligned}$$

Similarly, the expectation of the second sum in (4.2) can be bounded as follows

$$\begin{aligned}
&\frac{1}{n} E\left[\sum_{i=1}^n \sum_{j=1}^n p_{ij}^{(n)} 1\left(p_{ij}^{(n)} > (1+\epsilon) q_{ij}^{(n)},\, A_{ij}\right)\right] \\
&= \frac{1}{n} E\left[\sum_{i=1}^n \sum_{j=1}^n p_{ij}^{(n)} 1\left(q_{ij}^{(n)}(1+\varphi_n(\mathbf{X}_i, \mathbf{X}_j)) > (1+\epsilon) q_{ij}^{(n)},\, q_{ij}^{(n)} \leq \epsilon\right)\right] \\
&\leq \frac{1}{n} E\left[\sum_{i=1}^n \sum_{j=1}^n p_{ij}^{(n)} 1(\varphi_n(\mathbf{X}_i, \mathbf{X}_j) > \epsilon)\right] \\
&= \frac{1}{n} E\left[\sum_{i=1}^n \sum_{j=1}^n p_{ij}^{(n)}\right] - E\left[\left(\left\{\kappa(\mathbf{X}^{(n)}, \mathbf{Y}^{(n)})(1+\varphi_n(\mathbf{X}^{(n)}, \mathbf{Y}^{(n)}))\right\} \wedge n\right) 1(\varphi_n(\mathbf{X}^{(n)}, \mathbf{Y}^{(n)}) \leq \epsilon)\right]. \\
&\qquad (4.5)
\end{aligned}$$

Using Fatou's lemma again and Assumption 3.1(c) (which implies that $\varphi_n(\mathbf{X}^{(n)}, \mathbf{Y}^{(n)}) \xrightarrow{P} 0$ as $n \to \infty$), we have that

$$\liminf_{n\to\infty} E\left[\kappa(\mathbf{X}^{(n)}, \mathbf{Y}^{(n)}) 1(\varphi_n(\mathbf{X}^{(n)}, \mathbf{Y}^{(n)}) \geq -\epsilon)\right] \geq E\left[\kappa(\mathbf{X}, \mathbf{Y})\right]$$

and

$$\liminf_{n\to\infty} E\left[\left(\left\{\kappa(\mathbf{X}^{(n)},\mathbf{Y}^{(n)})(1+\varphi_n(\mathbf{X}^{(n)},\mathbf{Y}^{(n)}))\right\}\wedge n\right) 1(\varphi_n(\mathbf{X}^{(n)},\mathbf{Y}^{(n)})\leq \epsilon)\right] \geq E\left[\kappa(\mathbf{X},\mathbf{Y})\right].$$

It follows then from Assumption 3.1(d) that both (4.4) and (4.5) converge to zero. This completes the proof. ∎

The next result establishes the convergence in probability of the expected number of edges in the graph.

Lemma 4.2 *Under Assumption 3.1 we have*

$$\frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n \kappa(\mathbf{X}_i,\mathbf{X}_j) \to \iint_{\mathcal{S}^2} \kappa(\mathbf{x},\mathbf{y})\mu(d\mathbf{x})\mu(d\mathbf{y}) \quad \textit{and} \quad \frac{1}{n}\sum_{i=1}^n\sum_{j\neq i} p_{ij}^{(n)} \to \iint_{\mathcal{S}^2} \kappa(\mathbf{x},\mathbf{y})\mu(d\mathbf{x})\mu(d\mathbf{y})$$

in $L^1(P)$ as $n\to\infty$.

Proof. As in the proof of Lemma 4.1, note that

$$\frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n \kappa(\mathbf{X}_i,\mathbf{X}_j) = \mathbb{E}\left[\kappa(\mathbf{X}^{(n)},\mathbf{Y}^{(n)})\right],$$

where $\mathbf{X}^{(n)}$ and $\mathbf{Y}^{(n)}$ are conditionally i.i.d. given $\mathscr{F}$ with distribution μ_n (constructed as in Lemma 4.1). Let $\mathbf{X}$ and $\mathbf{Y}$ be i.i.d. with distribution μ and note that

$$\iint_{\mathcal{S}^2} \kappa(\mathbf{x},\mathbf{y})\mu(d\mathbf{x})\mu(d\mathbf{y}) = E[\kappa(\mathbf{X},\mathbf{Y})].$$

Next, note that for any fixed $M>0$ we have that $\kappa(\mathbf{x},\mathbf{y})\wedge M$ is bounded and continuous, so by Lemma A.2 in [5] we have that

$$\mathbb{E}[\kappa(\mathbf{X}^{(n)},\mathbf{Y}^{(n)})\wedge M] \xrightarrow{P} E[\kappa(\mathbf{X},\mathbf{Y})\wedge M]$$

as $n\to\infty$. Next, fix $\epsilon>0$ and choose $M>0$ such that $E[(\kappa(\mathbf{X},\mathbf{Y})-M)^+]\leq \epsilon/2$. Then,

$$\begin{aligned}
&P\left(\left|\mathbb{E}\left[\kappa(\mathbf{X}^{(n)},\mathbf{Y}^{(n)})\right] - E[\kappa(\mathbf{X},\mathbf{Y})]\right| > \epsilon\right)\\
&= P\left(\left|\mathbb{E}\left[\kappa(\mathbf{X}^{(n)},\mathbf{Y}^{(n)})\wedge M + (\kappa(\mathbf{X}^{(n)},\mathbf{Y}^{(n)})-M)^+\right] - E[\kappa(\mathbf{X},\mathbf{Y})\wedge M + (\kappa(\mathbf{X},\mathbf{Y})-M)^+]\right| > \epsilon\right)\\
&\leq P\left(\left|\mathbb{E}\left[\kappa(\mathbf{X}^{(n)},\mathbf{Y}^{(n)})\wedge M\right] - E[\kappa(\mathbf{X},\mathbf{Y})\wedge M]\right| + \mathbb{E}\left[(\kappa(\mathbf{X}^{(n)},\mathbf{Y}^{(n)})-M)^+\right] > \epsilon/2\right)\\
&\leq P\left(\left|\mathbb{E}\left[\kappa(\mathbf{X}^{(n)},\mathbf{Y}^{(n)})\wedge M\right] - E[\kappa(\mathbf{X},\mathbf{Y})\wedge M]\right| > \epsilon/4\right) + P\left(\mathbb{E}\left[(\kappa(\mathbf{X}^{(n)},\mathbf{Y}^{(n)})-M)^+\right] > \epsilon/4\right)\\
&\leq P\left(\left|\mathbb{E}\left[\kappa(\mathbf{X}^{(n)},\mathbf{Y}^{(n)})\wedge M\right] - E[\kappa(\mathbf{X},\mathbf{Y})\wedge M]\right| > \epsilon/4\right) + \frac{4}{\epsilon}E\left[(\kappa(\mathbf{X}^{(n)},\mathbf{Y}^{(n)})-M)^+\right].
\end{aligned}$$

Furthermore, the same arguments used in the proof of Lemma 4.1 give that

$$\limsup_{n\to\infty} E\left[(\kappa(\mathbf{X}^{(n)},\mathbf{Y}^{(n)})-M)^+\right] = E\left[\kappa(\mathbf{X},\mathbf{Y})\right] - \liminf_{n\to\infty} E\left[\kappa(\mathbf{X}^{(n)},\mathbf{Y}^{(n)})\wedge M\right]$$

$$\leq E\left[(\kappa(\mathbf{X},\mathbf{Y})-M)^+\right].$$

Therefore,

$$\limsup_{n\to\infty} P\left(\left|\mathbb{E}\left[\kappa(\mathbf{X}^{(n)},\mathbf{Y}^{(n)})\right]-E[\kappa(\mathbf{X},\mathbf{Y})]\right|>\epsilon\right)\leq \frac{4}{\epsilon}E\left[(\kappa(\mathbf{X},\mathbf{Y})-M)^+\right],$$

and taking $M\to\infty$ gives $\mathbb{E}\left[\kappa(\mathbf{X}^{(n)},\mathbf{Y}^{(n)})\right]\xrightarrow{P} E[\kappa(\mathbf{X},\mathbf{Y})]$ as $n\to\infty$. Since by Assumption 3.1(d) we have $E\left[\kappa(\mathbf{X}^{(n)},\mathbf{Y}^{(n)})\right]\to E[\kappa(\mathbf{X},\mathbf{Y})]$, then

$$\mathbb{E}\left[\kappa(\mathbf{X}^{(n)},\mathbf{Y}^{(n)})\right]\to E[\kappa(\mathbf{X},\mathbf{Y})] \quad \text{in } L^1(P) \qquad n\to\infty. \tag{4.6}$$

For the second result recall that $p_{ii}^{(n)}=0$ and $q_{ij}^{(n)}=\kappa(\mathbf{X}_i,\mathbf{X}_j)/n$, so it suffices to show that

$$\frac{1}{n}\sum_{i=1}^n\sum_{j=1}^n\left(p_{ij}^{(n)}-q_{ij}^{(n)}\right)\to 0 \quad \text{in } L^1(P) \qquad n\to\infty. \tag{4.7}$$

To see that this is the case fix $0<\epsilon<1/2$ and define B_{ij} according to Lemma 4.1. Next, note that by (4.6) and Lemma 4.1 we have

$$\begin{aligned}
E\left[\left|\frac{1}{n}\sum_{i=1}^n\sum_{j=1}^n\left(p_{ij}^{(n)}-q_{ij}^{(n)}\right)\right|\right] &\leq \frac{1}{n}E\left[\sum_{i=1}^n\sum_{j=1}^n \epsilon q_{ij}^{(n)}\right]+\frac{1}{n}E\left[\sum_{i=1}^n\sum_{j=1}^n\left(p_{ij}^{(n)}+q_{ij}^{(n)}\right)1(B_{ij}^c)\right]\\
&=\epsilon E\left[\kappa(\mathbf{X}^{(n)},\mathbf{Y}^{(n)})\right]+\frac{1}{n}E\left[\sum_{i=1}^n\sum_{j=1}^n\left(p_{ij}^{(n)}+q_{ij}^{(n)}\right)1(B_{ij}^c)\right]\\
&\to \epsilon E[\kappa(\mathbf{X},\mathbf{Y})]
\end{aligned}$$

as $n\to\infty$. Taking $\epsilon\to 0$ establishes (4.7), which completes the proof. ■

We are now ready to prove Proposition 3.3.

Proof of Proposition 3.3. We start by defining W_n to be the average number of arcs in the graph $G_n(\kappa(1+\varphi_n))$ given the types, that is, $W_n:=\mathbb{E}[e(G_n(\kappa(1+\varphi_n)))]/n=\frac{1}{n}\sum_{i=1}^n\sum_{j=1}^n p_{ij}^{(n)}$. Note that by Lemma 4.2 we have that $W_n\to E\left[\kappa(\mathbf{X},\mathbf{Y})\right]<\infty$ in $L^1(P)$ as $n\to\infty$, where $\mathbf{X}$ and $\mathbf{Y}$ are i.i.d. with common distribution μ. Therefore, it suffices to show that $e(G_n(\kappa(1+\varphi_n)))/n-W_n\to 0$ in $L^1(P)$ as $n\to\infty$.

To do this, let Y_{ij} denote the indicator of whether arc (i,j) is present in $G_n(\kappa(1+\varphi_n))$ and note that

$$e(G_n(\kappa(1+\varphi_n)))=\sum_{i=1}^n\sum_{j\neq i}Y_{ij},$$

where the $\{Y_{ij}\}$ are Bernoulli random variables with means $\{p_{ij}^{(n)}\}$, conditionally independent given $\mathscr{F}$. It follows that

$$\mathrm{Var}(e(G_n(\kappa(1+\varphi_n)))|\mathscr{F})=\sum_{i=1}^n\sum_{j=1}^n\mathrm{Var}(Y_{ij}|\mathscr{F})\leq\sum_{i=1}^n\sum_{j=1}^n\mathbb{E}[Y_{ij}]=\sum_{i=1}^n\sum_{j=1}^n p_{ij}^{(n)}=nW_n.$$

Therefore,

$$\begin{aligned} E\left[(e(G_n(\kappa(1+\varphi_n)))/n - W_n)^2\right] &= E\left[\mathbb{E}\left[(e(G_n(\kappa(1+\varphi_n)))/n - W_n)^2\right]\right] \\ &= E\left[n^{-2}\mathrm{Var}(e(G_n(\kappa(1+\varphi_n)))|\mathscr{F})\right] \\ &\leq n^{-2}E\left[nW_n\right] \xrightarrow{P} 0, \end{aligned}$$

as $n \to \infty$. Hence, $e(G_n(\kappa(1+\varphi_n)))/n - W_n \to 0$ in $L^2(P)$, which completes the proof. ■

4.2 Distribution of Vertex Degrees

We now move on to the proof of Theorem 3.6. The proof of Theorem 3.4 is given in Section 4.3, since it can be obtained as a corollary to Theorem 4.6. We will show that (Z^-, Z^+) has a non-standard regularly varying distribution whenever their conditional means $(\lambda_-(\mathbf{X}), \lambda_+(\mathbf{X}))$ have a non-standard regularly varying distribution. Throughout the proof we use the notation $[\mathbf{a}, \mathbf{b}] = \{\mathbf{x} \in \mathbb{R}^2 : \mathbf{a} \leq \mathbf{x} \leq \mathbf{b}\}$ to denote the rectangles in $\mathbb{R}^2$.

Proof of Theorem 3.6. To simplify the notation, let $\mathbf{W} = (W^-, W^+) = (\lambda_-(\mathbf{X}), \lambda_+(\mathbf{X}))$, and recall that we need to show that $\tilde{\nu}_t(\cdot) = tP((Z^-/a(t) \in du,\, Z^+/b(t)) \in \cdot)$ converges vaguely to $\nu(\cdot)$ in $M_+([0,\infty]^2 \setminus \{\mathbf{0}\})$ as $t \to \infty$. Note that by Lemma 6.1 in [29], it suffices to show that $\tilde{\nu}_t([\mathbf{0}, \mathbf{x}]^c) \to \nu([\mathbf{0}, \mathbf{x}]^c)$ as $t \to \infty$ for any continuity point $\mathbf{x} \in [\mathbf{0}, \boldsymbol{\infty}) \setminus \{\mathbf{0}\}$ of $\nu([\mathbf{0}, \cdot]^c)$.

To start, fix $(p, q) \in [\mathbf{0}, \boldsymbol{\infty}) \setminus \{\mathbf{0}\}$ to be a continuity point of $\nu([\mathbf{0}, \cdot]^c)$ and note that

$$\begin{aligned} \tilde{\nu}_t((p,\infty] \times (q,\infty]) &= \int_p^\infty \int_q^\infty tP\left(\frac{Z^-}{a(t)} \in du,\, \frac{Z^+}{b(t)} \in dv\right) \\ &= tP\left(\frac{Z^-}{a(t)} > p,\, \frac{Z^+}{b(t)} > q\right) \\ &= tE\left[P\left(\frac{Z^-}{a(t)} > p,\, \frac{Z^+}{b(t)} > q \middle| \mathbf{W}\right)\right] \\ &= tE\left[P\left(Z^- > pa(t) \middle| \mathbf{W}\right) P\left(Z^+ > qb(t) \middle| \mathbf{W}\right)\right]. \end{aligned}$$

It follows that we need to show that

$$\lim_{t\to\infty} tE\left[P\left(Z^- > pa(t) \middle| \mathbf{W}\right) P\left(Z^+ > qb(t) \middle| \mathbf{W}\right)\right] = \nu((p,\infty] \times (q,\infty]).$$

To this end, define $e(t) = \sqrt{\gamma a(t) \log a(t)}$ and $d(t) = \sqrt{\eta b(t) \log b(t)}$ with $\gamma > 2q\beta$, $\eta > 2p\alpha$, and use them to define the events

$$A_t = \{W^- > pa(t) - e(t)\} \qquad \text{and} \qquad B_t = \{W^+ > qb(t) - d(t)\}.$$

Now note that

$$\begin{aligned} & tE\left[P\left(Z^- > pa(t) \middle| \mathbf{W}\right) P\left(Z^+ > qb(t) \middle| \mathbf{W}\right)\right] \\ &= tE\left[P\left(Z^- > pa(t) \middle| \mathbf{W}\right) P\left(Z^+ > qb(t) \middle| \mathbf{W}\right) 1(A_t \cap B_t)\right] \qquad (4.8) \\ &\quad + tE\left[P\left(Z^- > pa(t) \middle| \mathbf{W}\right) P\left(Z^+ > qb(t) \middle| \mathbf{W}\right) 1(A_t^c \cup B_t^c)\right]. \qquad (4.9) \end{aligned}$$

To see that (4.9) vanishes in the limit, use the bound $P(\text{Poi}(\lambda) \geq p) \leq e^{-\lambda}(e\lambda/p)^p$ for $p > \lambda$, where $\text{Poi}(\lambda)$ is Poisson random variable with mean λ, to obtain that

$$\begin{aligned}
&tE\left[P\left(Z^- > pa(t)\middle|\,\mathbf{W}\right)P\left(Z^+ > qb(t)\middle|\,\mathbf{W}\right)1(A_t^c)\right] \\
&\leq tE\left[P\left(Z^- > pa(t)\middle|\,\mathbf{W}\right)1(A_t^c)\right] \\
&\leq tE\left[\exp\left\{-W^- + pa(t)\left(1+\log(W^-) - \log(pa(t))\right)\right\}1(A_t^c)\right] \\
&\leq t\exp\{-(pa(t)-e(t)) + pa(t)(1+\log(pa(t)-e(t)) - \log(pa(t)))\} \\
&= t\exp\left\{e(t) + pa(t)\log\left(1 - \frac{e(t)}{pa(t)}\right)\right\} \\
&= t\exp\left(-\frac{e(t)^2}{2pa(t)} + O\left(\frac{e(t)^3}{(pa(t))^2}\right)\right) = ta(t)^{-\frac{\gamma}{2p}}\left(1 + O\left(\frac{(\log a(t))^{3/2}}{a(t)^{1/2}}\right)\right),
\end{aligned}$$

where in the third inequality we used the observation that $g(u) = -u + pa(t)\log u$ is concave with a unique maximizer at $u^* = pa(t)$. Similarly,

$$\begin{aligned}
&tE\left[P\left(Z^- > pa(t)\middle|\,\mathbf{W}\right)P\left(Z^+ > qb(t)\middle|\,\mathbf{W}\right)1(B_t^c)\right] \\
&\leq tb(t)^{-\frac{\eta}{2q}}\left(1 + O\left(\frac{(\log b(t))^{3/2}}{b(t)^{1/2}}\right)\right).
\end{aligned}$$

Our choice of γ, η guarantees that both terms converge to zero as $t \to \infty$, hence showing that (4.9) does so as well.

It remains to show that (4.8) converges to $\nu((p,\infty] \times (q,\infty])$ as $t \to \infty$. To do this, we first note that (4.8) is equal to

$$tP(A_t \cap B_t) - tE\left[\left(1 - P\left(Z^- > pa(t)\middle|\,\mathbf{W}\right)P\left(Z^+ > qb(t)\middle|\,\mathbf{W}\right)\right)1(A_t \cap B_t)\right],$$

where

$$\begin{aligned}
&tE\left[\left(1 - P\left(Z^- > pa(t)\middle|\,\mathbf{W}\right)P\left(Z^+ > qb(t)\middle|\,\mathbf{W}\right)\right)1(A_t \cap B_t)\right] \\
&\leq tE\left[P\left(Z^- \leq pa(t)\middle|\,\mathbf{W}\right)1(A_t \cap B_t)\right] + tE\left[P\left(Z^+ \leq qb(t)\middle|\,\mathbf{W}\right)1(A_t \cap B_t)\right] \\
&\leq tE\left[P\left(Z^- \leq pa(t)\middle|\,\mathbf{W}\right)1(\tilde{A}_t \cap B_t)\right] + tE\left[P\left(Z^+ \leq qb(t)\middle|\,\mathbf{W}\right)1(A_t \cap \tilde{B}_t)\right] \\
&\quad + tP(\tilde{A}_t^c \cap A_t \cap B_t) + tP(A_t \cap B_t \cap \tilde{B}_t^c)
\end{aligned}$$

with

$$\tilde{A}_t = \{W^- > pa(t) + e(t)\} \subseteq A_t \qquad \text{and} \qquad \tilde{B}_t = \{W^+ > qb(t) + d(t)\} \subseteq B_t.$$

Now note that the inequality $P(\text{Poi}(\lambda) \leq p) \leq e^{-\lambda}(e\lambda/p)^p$ for $0 \leq p < \lambda$ gives that

$$\begin{aligned}
&tE\left[P\left(Z^- \leq pa(t)\middle|\,\mathbf{W}\right)1(\tilde{A}_t \cap B_t)\right] \\
&\leq tE\left[\exp\left\{-W^- + pa(t)\left(1+\log(W^-) - \log(pa(t))\right)\right\}1(\tilde{A}_t)\right] \\
&\leq t\exp\left\{-(pa(t)+e(t)) + pa(t)\left(1+\log(pa(t)+e(t)) - \log(pa(t))\right)\right\} \\
&= t\exp\left\{-e(t) + pa(t)\log\left(1 + \frac{e(t)}{pa(t)}\right)\right\}
\end{aligned}$$

$$= t\exp\left(-\frac{e(t)^2}{2pa(t)} + O\left(\frac{e(t)^3}{(pa(t))^2}\right)\right) = ta(t)^{-\frac{\gamma}{2p}}\left(1 + O\left(\frac{(\log a(t))^{3/2}}{a(t)^{1/2}}\right)\right),$$

where we used again the concavity of $g(u) = -u + pa(t)\log u$. Similarly,

$$tE\left[P\left(Z^+ \leq qb(t)\middle|\, \mathbf{W}\right) 1(A_t \cap \tilde{B}_t)\right] \leq tb(t)^{-\frac{\eta}{2q}}\left(1 + O\left(\frac{(\log b(t))^{3/2}}{b(t)^{1/2}}\right)\right),$$

and our choice of γ, η give again that

$$\lim_{t\to\infty}\left\{tE\left[P\left(Z^- \leq pa(t)\middle|\, \mathbf{W}\right) 1(\tilde{A}_t \cap B_t)\right] + tE\left[P\left(Z^+ \leq qb(t)\middle|\, \mathbf{W}\right) 1(A_t \cap \tilde{B}_t)\right]\right\} = 0. \tag{4.10}$$

Next, let $\nu_t(du, dv) = tP(W^-/a(t) \in du,\, W^+/b(t) \in dv)$ and note that for any $0 < \epsilon < p \wedge q$, we have that

$$\begin{aligned}
&\limsup_{t\to\infty}\left\{tP(\tilde{A}_t^c \cap A_t \cap B_t) + tP(A_t \cap B_t \cap \tilde{B}_t^c)\right\} \\
&= \limsup_{t\to\infty}\{\nu_t\left((p - e(t)/a(t), p + e(t)/a(t)] \times (q - d(t)/b(t), \infty]\right) \\
&\qquad\qquad + \nu_t\left((p - e(t)/a(t), \infty] \times (q - d(t)/b(t), q + d(t)/b(t)]\right)\} \\
&\leq \limsup_{t\to\infty}\{\nu_t\left((p-\epsilon, p+\epsilon] \times (q-\epsilon, \infty]\right) + \nu_t\left((p-\epsilon, \infty] \times (q-\epsilon, q+\epsilon]\right)\} \\
&= \nu\left((p-\epsilon, p+\epsilon] \times (q-\epsilon, \infty]\right) + \nu\left((p-\epsilon, \infty] \times (q-\epsilon, q+\epsilon]\right).
\end{aligned}$$

Moreover, since (p, q) is a continuity point of ν, then

$$\lim_{\epsilon\downarrow 0}\{\nu\left((p-\epsilon, p+\epsilon] \times (q-\epsilon, \infty]\right) + \nu\left((p-\epsilon, \infty] \times (q-\epsilon, q+\epsilon]\right)\} = 0.$$

It follows that

$$\lim_{t\to\infty}\left\{tP(\tilde{A}_t^c \cap A_t \cap B_t) + tP(A_t \cap B_t \cap \tilde{B}_t^c)\right\} = 0,$$

which combined with (4.10) gives that

$$\lim_{t\to\infty} tE\left[\left(1 - P\left(Z^- > pa(t)\middle|\, \mathbf{W}\right) P\left(Z^+ > qb(t)\middle|\, \mathbf{W}\right)\right) 1(A_t \cap B_t)\right] = 0.$$

Finally, the continuity of ν at (p, q) also yields that

$$\lim_{t\to\infty} tP(A_t \cap B_t) = \lim_{t\to\infty} \nu_t\left((p - e(t)/a(t), \infty] \times (q - d(t)/b(t), \infty]\right) = \nu\left((p, \infty] \times (q, \infty]\right).$$

■

4.3 Phase transition for the largest strongly connected component

The last part of the paper considers the connectivity properties of the graph, in particular, the size of the largest strongly connected component. As mentioned in Section 3.3, our Theorem 3.10 provides the directed version of Theorem 3.1 in [5]. However, our proof approach differs from the

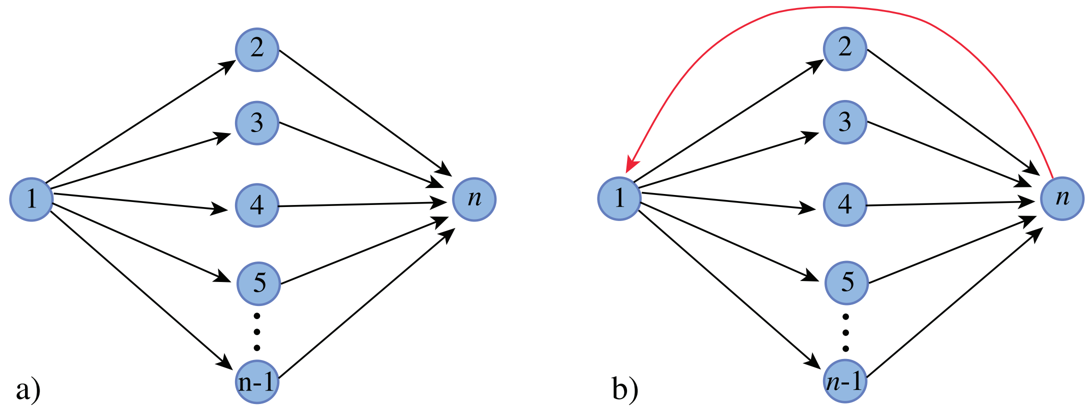

Figure 1: Directed graph with n vertices. a) There is no strongly connected component. b) The same graph with one additional arc; the largest strongly connected component is *giant* of size n.

one used in [5] in the order in which we construct the different couplings involved. Specifically, in [5] the authors first couple the graph $G_n(\kappa(1+\varphi_n))$ with another graph $G_n(\kappa_m)$, where κ_m is a piecewise constant kernel taking at most a finite number of different values and such that $\kappa_m \nearrow \kappa$ as $m \to \infty$. Then, they provide a coupling between the exploration of the component of a randomly chosen vertex in $G_n(\kappa_m)$ and that of a multi-type branching process, $\mathcal{T}_\mu(\kappa_m)$, whose offspring distribution is determined by κ_m. The phase transition result is then obtained by relating the survival probability of $\mathcal{T}_\mu(\kappa_m)$ with the survival probability of its limiting tree $\mathcal{T}_\mu(\kappa)$. Our proof leverages on the work done in [3], which applies to a related graph $G_{n'}(\kappa_m)$, to establish a lower bound for the size of the largest strongly connected component. For the upper bound, we give a new direct coupling between the exploration of the in-component and out-component of a randomly chosen vertex in $G_n(\kappa(1+\varphi_n))$ and a double tree $(\mathcal{T}_\mu^-(\kappa_m), \mathcal{T}_\mu^+(\kappa_m))$, where $\kappa_m \nearrow \kappa$ as $m \to \infty$. We then relate the survival probabilities of $(\mathcal{T}_\mu^-(\kappa_m), \mathcal{T}_\mu^+(\kappa_m))$ with those of their limiting trees $(\mathcal{T}_\mu^-(\kappa), \mathcal{T}_\mu^+(\kappa))$ as $m \to \infty$.

Interestingly, trying to adapt the approach used in [5] to the directed case leads to a phenomenon that does not occur when analyzing undirected graphs. Namely, if we consider two coupled undirected graphs $G_n(\kappa(1+\varphi_n))$ and $G_n(\kappa'(1+\varphi'_n))$ such that every edge in the first graph is also present in the second one but not the other way around (e.g., when $\kappa(\mathbf{x},\mathbf{y})(1+\varphi_n(\mathbf{x},\mathbf{y})) \leq \kappa'(\mathbf{x},\mathbf{y})(1+\varphi'_n(\mathbf{x},\mathbf{y}))$ for all $\mathbf{x},\mathbf{y} \in \mathcal{S}$), then, the difference in the sizes of the components of a vertex present in both graphs can be bounded by the difference in their number of edges (see Lemma 9.4 in [5]). However, in the directed case, this is no longer true, as Figure 1 illustrates. In other words, the existence of a (giant) strongly connected component can be determined by a single arc. For this reason, a coupling of the graphs $G_n(\kappa(1+\varphi_n))$ and $G_n(\kappa_m)$, such as the one used in [5], does not provide an upper bound for the size of the strongly connected component in the directed case. This may be a notable observation considering the folklore that exists around the equivalence of undirected and directed networks.

To ease its reading, we have subdivided this section into two subsections. In the first one we provide our coupling theorem between the exploration of the in-component and out-component of a randomly chosen vertex in $G_n(\kappa(1+\varphi_n))$ and the double tree $(\mathcal{T}_\mu^-(\kappa_m), \mathcal{T}_\mu^+(\kappa_m))$. The second subsection gives the proof of Theorem 3.10, which establishes the phase transition for the size of the largest strongly connected component.

4.3.1 Coupling with a double multi-type branching process

Starting with a randomly chosen vertex in $G_n(\kappa(1+\varphi_n))$, say vertex i, we will perform a double exploration process that we will couple with a double multi-type branching process $\{\hat{\mathbf{Z}}_t^{(n)} : t \geq 0\}$ having "types" $\{1, \ldots, n\}$. Note that these "types" are actually the *identities* of the vertices in $[n]$, so to avoid confusion with the actual *types* of each of the vertices, i.e., $\{\mathbf{X}_1, \ldots, \mathbf{X}_n\}$, we will say that a vertex in the double tree has an *identity*, not a "type". The double tree is started at $\hat{\mathbf{Z}}_0^{(n)} = (\hat{Z}_{1,0}, \hat{Z}_{2,0}, \ldots, \hat{Z}_{n,0})$, and is such that for $t \geq 1$, $\hat{\mathbf{Z}}_t^{(n)} = (\hat{Z}_{1,t}^-, \hat{Z}_{2,t}^-, \ldots, \hat{Z}_{n,t}^-, \hat{Z}_{1,t}^+, \hat{Z}_{2,t}^+, \ldots, \hat{Z}_{n,t}^+) \in \mathbb{N}^{2n}$, where $\hat{Z}_{j,t}^-$ denotes the number of individuals of *identity* j in the tth inbound generation of the double tree and $\hat{Z}_{j,t}^+$ denotes the number of individuals of *identity* j in the tth outbound generation of the double tree. Moreover, the number of offspring that each node in the double tree has is independent of all other nodes in the double tree, conditionally on the identity of the node. The initial vector $\hat{\mathbf{Z}}_0^{(n)}$ is set to equal $\mathbf{e}_i$, where $\mathbf{e}_i$ is the unit vector that has a one in position i and zeros elsewhere; note also that it does not have a $+/-$ superscript since it is at the center of the double tree.

In order to define the offspring distribution of nodes in the double tree, we fix a regular finitary kernel κ_m on $\mathcal{S} \times \mathcal{S}$ (see Definition 3.9) satisfying

$$0 \leq \kappa_m(\mathbf{x}, \mathbf{y}) \leq \kappa(\mathbf{x}, \mathbf{y}) \qquad \text{for all } \mathbf{x}, \mathbf{y} \in \mathcal{S},$$

and such that

$$\kappa_m(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{M_m} \sum_{j=1}^{M_m} c_{ij}^{(m)} 1(\mathbf{x} \in \mathcal{J}_i^{(m)}, \mathbf{y} \in \mathcal{J}_j^{(m)}),$$

for some partition $\{\mathcal{J}_i^{(m)} : 1 \leq i \leq M_m\}$ of $\mathcal{S}$ and some nonnegative constants $\{c_{ij}^{(m)} : 1 \leq i, j \leq M_m\}$, $M_m < \infty$. Now let the number of offspring of *identity* j that a node of *identity* i in the inbound tree, respectively outbound tree, has, be Poisson distributed with mean $r_{ji}^{(m,n)}$, resp. $\tilde{r}_{ij}^{(m,n)}$, where:

$$r_{ji}^{(m,n)} = \frac{\kappa_m(\mathbf{X}_j, \mathbf{X}_i)\mu(\mathcal{J}_{\theta(j)}^{(m)})}{n\mu_n(\mathcal{J}_{\theta(j)}^{(m)})} \qquad \text{and} \qquad \tilde{r}_{ij}^{(m,n)} = \frac{\kappa_m(\mathbf{X}_i, \mathbf{X}_j)\mu(\mathcal{J}_{\theta(j)}^{(m)})}{n\mu_n(\mathcal{J}_{\theta(j)}^{(m)})},$$

and $\theta(i) = j$ if and only if $\mathbf{X}_i \in \mathcal{J}_j^{(m)}$. We denote $\mathcal{T}_\mu^-(\kappa_m; \mathbf{X}_i)$ and $\mathcal{T}_\mu^+(\kappa_m; \mathbf{X}_i)$ the inbound and outbound trees, respectively, whose root is vertex i. Note that the trees $\mathcal{T}_\mu^-(\kappa_m; \mathbf{X}_i)$ and $\mathcal{T}_\mu^+(\kappa_m; \mathbf{X}_i)$ are conditionally independent (given $\mathscr{F}$) by construction.

Note: We point out that in the double tree *identities* can appear multiple times, unlike in the graph where they appear only once. In either case, *identities* take values in the set $[n] = \{1, 2, \ldots, n\}$.

Remark 4.3 *An important observation that will be used later is that the double tree* $\hat{\mathbf{Z}}_t^{(n)} = (\hat{Z}_{1,t}^-, \ldots, \hat{Z}_{n,t}^-, \hat{Z}_{1,t}^+, \ldots, \hat{Z}_{n,t}^+) \in \mathbb{N}^{2n}$ *defined above has the same law as the double tree* $\tilde{\mathbf{Z}}_t^{(m)} =$

$(\tilde{Z}^-_{1,t}, \ldots, \tilde{Z}^-_{M_m,t}, \tilde{Z}^+_{1,t}, \ldots, \tilde{Z}^+_{M_m,t}) \in \mathbb{N}^{2M_m}$ $(\tilde{\mathbf{Z}}^{(m)}_0 = \mathbf{e}_{\theta(i)}$ *if* $\hat{\mathbf{Z}}^{(n)}_0 = \mathbf{e}_i)$*, whose offspring distributions are Poisson with means*

$$m^-_{ij} := c^{(m)}_{ji} \mu(\mathcal{J}^{(m)}_j) \qquad \text{and} \qquad m^+_{ij} := c^{(m)}_{ij} \mu(\mathcal{J}^{(m)}_j), \qquad 1 \leq i, j \leq M_m.$$

Moreover, the latter is the same as $(\mathcal{T}^-_\mu(\kappa_m; \mathbf{x}), \mathcal{T}^+_\mu(\kappa_m; \mathbf{x}))$ *for any* $\mathbf{x} \in \mathcal{J}^{(m)}_i$.

Recall that $Y_{ij} = 1(\text{arc } (i,j) \text{ is present in } G_n(\kappa(1 + \varphi_n)))$ is a Bernoulli random variable with success probability

$$p^{(n)}_{ij} = \frac{\kappa(\mathbf{X}_i, \mathbf{X}_j)(1 + \varphi_n(\mathbf{X}_i, \mathbf{X}_j))}{n} \wedge 1, \quad 1 \leq i \neq j \leq n, \qquad p^{(n)}_{ii} = 0.$$

We will couple Y_{ij} with a Poisson random variable Z_{ij} having mean $r^{(m,n)}_{ij}$ on the inbound side, and with a Poisson random variable $\tilde{Z}_{ij}$ having mean $\tilde{r}^{(m,n)}_{ij}$ on the outbound side, using a sequence $\{U_{ij} : 1 \leq i, j \leq n\}$ of i.i.d. Uniform$(0,1)$ random variables.

The exploration of the graph and the construction of the double tree are done by choosing a vertex uniformly at random among those which have not been explored. Starting with vertex i, we fix the number of vertices to explore in the in-component of i, say k_{in}, and the number of vertices to explore in the out-component of i, say k_{out}. A *step* in the exploration of the in-component (out-component) corresponds to identifying the inbound (outbound) neighbors of the vertex being explored. The exploration of the in-component continues until we have explored k_{in} vertices or until there are no more vertices to reveal, after which we proceed to explore the out-component for k_{out} steps or until there are no more vertices to reveal. Moreover, we allow k_{in} and k_{out} to be stopping times with respect to the history of the exploration process.

Vertices in the graph can have one of two labels: {inactive, active}, or they may be unlabelled. Active vertices are those that have been identified to be in the in-component, respectively out-component, of vertex i but whose inbound, respectively outbound, neighbors have not been revealed. Inactive vertices are all other vertices that have been revealed through the exploration process but that are not active; again, there is an inbound inactive set and an outbound inactive set. Inactive vertices on the inbound side have revealed all its inbound neighbors, but not necessarily all their outbound ones; symmetrically, inactive nodes on the outbound side have revealed all their outbound neighbors but not necessarily all their inbound ones.

In the double tree we will say that a node is "active" if we have not yet sampled its offspring, and "inactive" if we have.

Notation: For $r = 0, 1, 2, \ldots$, and assuming the chosen vertex is i, let

A^-_r (A^+_r) = set of inbound (outbound) "active" vertices after having explored the first r vertices in the in-component (out-component) of vertex i.

I^-_r (I^+_r) = set of inbound (outbound) "inactive" vertices after having explored the first r vertices in the in-component (out-component) of vertex i.

T^-_r (T^+_r) = *identity* of the vertex being explored in step r, $r \geq 1$, of the exploration of the in-component (out-component) of vertex i.

$\hat{A}_r^-$ ($\hat{A}_r^+$) = set of "active" nodes in $\mathcal{T}_\mu^-(\kappa_m;\mathbf{X}_i)$ ($\mathcal{T}_\mu^+(\kappa_m;\mathbf{X}_i)$) after having sampled the offspring of the first r nodes in $\mathcal{T}_\mu^-(\kappa_m;\mathbf{X}_i)$ ($\mathcal{T}_\mu^+(\kappa_m;\mathbf{X}_i)$).

$\hat{I}_r^-$ ($\hat{I}_r^+$) = set of *identities* belonging to "inactive" nodes in $\mathcal{T}_\mu^-(\kappa_m;\mathbf{X}_i)$ ($\mathcal{T}_\mu^+(\kappa_m;\mathbf{X}_i)$) after having sampled the offspring of the first r nodes in $\mathcal{T}_\mu^-(\kappa_m;\mathbf{X}_i)$ ($\mathcal{T}_\mu^+(\kappa_m;\mathbf{X}_i)$).

$\hat{T}_r^-$ ($\hat{T}_r^+$) = *identity* of the node in $\mathcal{T}_\mu^-(\kappa_m;\mathbf{X}_i)$ ($\mathcal{T}_\mu^+(\kappa_m;\mathbf{X}_i)$) whose offspring are being sampled in step r; $r \geq 1$.

Exploration of the components of vertex i in the graph:

Fix k_{in} and k_{out}.

1) For the exploration of the in-component:

   Step 0: Label vertex i as "active" on the inbound side and set $A_0^- = \{i\}$, $I_0^- = \varnothing$.

   Step r, $1 \leq r \leq k_{in}$:

   Choose, uniformly at random, a vertex in A_{r-1}^-; let $T_r^- = i$ denote its *identity*.

   a) For $j = 1, 2, \ldots, n$, $j \neq i$:
      i. Realize $Y_{ji} = 1(U_{ji} > 1 - p_{ji}^{(n)})$. If $Y_{ji} = 0$ go to 1(a).
      ii. If $Y_{ji} = 1$ and vertex $j \in I_{r-1}^- \cup A_{r-1}^-$, do nothing. Go to 1(a).
      iii. If $Y_{ji} = 1$ and vertex j had no label, label it "active" on the inbound side. Go to 1(a).
   b) Once all the new inbound neighbors of vertex i have been identified and labeled "active", label vertex i as "inactive" on the inbound side.
   c) Define the sets $A_r^- = A_{r-1}^- \cup \{\text{new "active" vertices created in 1(a)(iii)}\} \setminus \{i\}$ and $I_r^- = I_{r-1}^- \cup \{i\}$. This completes Step r on the inbound side.

2) For the exploration of the out-component:

   Step 0: Label vertex i as "active" on the outbound side and set $A_0^+ = \{i\}$, $I_0^+ = \varnothing$.

   Step r, $1 \leq r \leq k_{out}$:

   Choose, uniformly at random, a vertex in A_{r-1}^+; let $T_r^+ = i$ denote its *identity*.

   a) For $j = 1, 2, \ldots, n$, $j \neq i$, $j \notin I_{k_{in}}^- \cup A_{k_{in}}^-$:
      i. Realize $Y_{ij} = 1(U_{ij} > 1 - p_{ij}^{(n)})$. If $Y_{ij} = 0$ go to 2(a).
      ii. If $Y_{ij} = 1$ and vertex $j \in I_{r-1}^+ \cup A_{r-1}^+$, do nothing. Go to 2(a).
      iii. If $Y_{ij} = 1$ and vertex j had no label, label it "active" on the outbound side. Go to 2(a).
   b) Once all the new outbound neighbors of vertex i have been identified and labeled "active", label vertex i as "inactive" on the outbound side.
   c) Define the sets $A_r^+ = A_{r-1}^+ \cup \{\text{new "active" vertices created in 2(a)(iii)}\} \setminus \{i\}$ and $I_r^+ = I_{r-1}^+ \cup \{i\}$. This completes Step r on the outbound side.

Note that by setting $k_{in} = \inf\{r \geq 1 : A_r^- = \varnothing\}$ and $k_{out} = \inf\{r \geq 1 : A_r^+ = \varnothing\}$ we can fully explore the in-component and out-component of vertex i. We now explain how the coupled double tree is constructed.

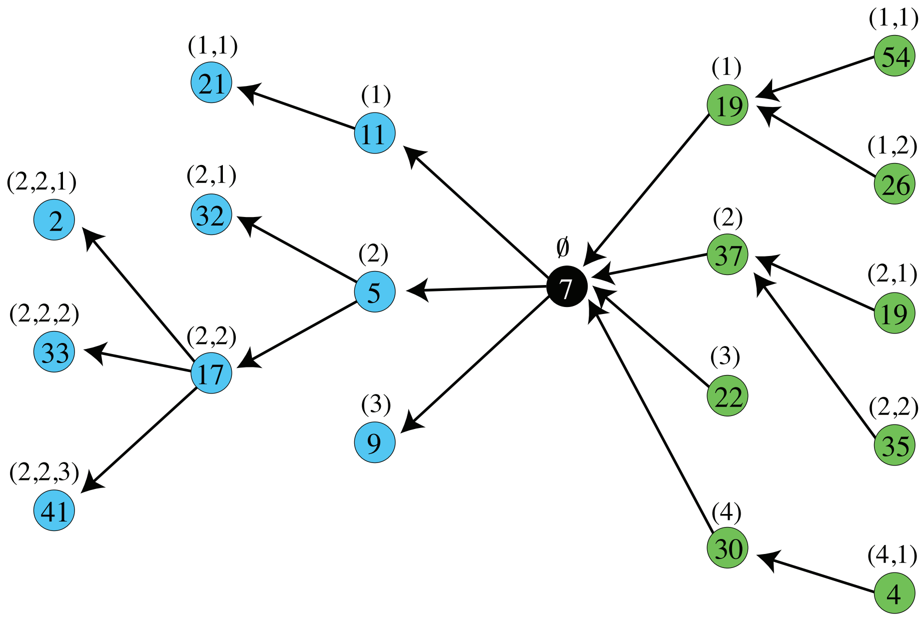

Figure 2: Exploration of the graph and coupled tree. We explore vertex 7 in the graph, which means the root of the double tree has *identity* $T_\emptyset = 7$. Node *identities* in the double tree are depicted inside the circles, whereas tree *labels* are right on top, e.g., node $(2, 2, 3)$ on the outbound tree has *identity* $T_{(2,2,3)} = 41$, whereas node (3) on the inbound tree has *identity* $T_{(3)} = 22$.

Coupled construction of the double multi-type branching process:

Let $g^{-1}(u)$ denote the pseudo inverse of function g, i.e., $g^{-1}(u) = \inf\{x : u \leq g(x)\}$. Let G_{ji} and $\tilde{G}_{ij}$ be the distribution functions of Poisson random variables having means $r_{ji}^{(m,n)}$ and $\tilde{r}_{ij}^{(m,n)}$, respectively. On the double tree we use the index notation $\mathbf{i} = (i_1, \ldots, i_r)$ to identify nodes in the rth generation (inbound/outbound) of the double tree. Let $T_\mathbf{i}$ denote the *identity* of node $\mathbf{i}$; see Figure 2.

1) Construction of the inbound tree:

   Step 0: Set $\hat{\mathbf{Z}}_0^{(n)} = \mathbf{e}_i$. Let $\hat{A}_0^- = \{\emptyset\}$, $T_\emptyset = i$, $\hat{I}_0^- = \varnothing$.

   Step r, $1 \leq r \leq k_{in}$:

   Choose a node in $\mathbf{i} \in \hat{A}_{r-1}^-$, uniformly at random; set $\hat{T}_r^- = T_\mathbf{i}$.

   I. If this is the first time *identity* $T_\mathbf{i}$ appears in the inbound tree, do as follows:

      a) For $j = 1, 2, \ldots, n$, $j \notin \{T_\mathbf{i}\}$:

         i. Realize $Z_{j,T_\mathbf{i}} = G_{j,T_\mathbf{i}}^{-1}(U_{j,T_\mathbf{i}})$. If $Z_{j,T_\mathbf{i}} = 0$ go to 1(I)(a).

ii. If $Z_{j,T_{\mathbf{i}}} \geq 1$ label each of the newly created nodes as "active" on the inbound side. Go to 1(I)(a).

b) For $j = T_{\mathbf{i}}$:

i. Sample $Z^*_{j,T_{\mathbf{i}}}$ to be a Poisson random variable with mean $r^{(m,n)}_{j,T_{\mathbf{i}}}$, independently of everything else. If $Z^*_{j,T_{\mathbf{i}}} = 0$ go to 1(I)(c).

ii. If $Z^*_{j,T_{\mathbf{i}}} \geq 1$ label each of the newly created nodes as "active" on the inbound side. Go to 1(I)(c).

c) Once all the inbound offspring of node $\mathbf{i}$ have been identified, label *identity* $T_{\mathbf{i}}$ as "inactive" on the inbound side.

d) Define the sets $\hat{A}^-_r = \hat{A}^-_{r-1} \cup \{$new "active" nodes created in 1(I)(a)(ii) and 1(I)(b)(ii)$\} \setminus \{\mathbf{i}\}$ and $\hat{I}^-_r = \hat{I}^-_{r-1} \cup \{T_{\mathbf{i}}\}$. This completes Step r on the inbound side.

II. Else:

a) For $j = 1, 2, \ldots, n$:

i. Sample $Z^*_{j,T_{\mathbf{i}}}$ to be a Poisson random variable with mean $r^{(m,n)}_{j,T_{\mathbf{i}}}$, independently of everything else. If $Z^*_{j,T_{\mathbf{i}}} = 0$ go to 1(II)(a).

ii. If $Z^*_{j,T_{\mathbf{i}}} \geq 1$ label each of the newly created nodes as "active" on the inbound side. Go to 1(II)(a).

b) Once all the inbound offspring of node $\mathbf{i}$ have been identified, label *identity* $T_{\mathbf{i}}$ as "inactive" on the inbound side.

c) Define the sets $\hat{A}^-_r = \hat{A}^-_{r-1} \cup \{$new "active" nodes created in 1(II)(a)(ii)$\} \setminus \{\mathbf{i}\}$ and $\hat{I}^-_r = \hat{I}^-_{r-1} \cup \{T_{\mathbf{i}}\}$. This completes Step r on the inbound side.

2) Construction of the outbound tree:

Step 0: Set $\hat{A}^+_0 = \{\emptyset\}$, $T_\emptyset = i$, $\hat{I}^+_0 = \varnothing$.

Step r, $1 \leq r \leq k_{out}$:

Choose a node $\mathbf{i} \in \hat{A}^+_{r-1}$, uniformly at random; set $\hat{T}^+_r = T_{\mathbf{i}}$.

I. If this is the first time *identity* $T_{\mathbf{i}}$ appears in the outbound tree, do as follows:

a) For $j = 1, 2, \ldots, n$, $j \notin \{T_{\mathbf{i}}\} \cup \{T_{\mathbf{j}} : T_{\mathbf{j}} \in \hat{I}^-_{k_{in}}$ or $\mathbf{j} \in \hat{A}^-_{k_{in}}\}$:

i. Realize $\tilde{Z}_{T_{\mathbf{i}},j} = \tilde{G}^{-1}_{T_{\mathbf{i}},j}(U_{T_{\mathbf{i}},j})$. If $\tilde{Z}_{T_{\mathbf{i}},j} = 0$ go to 2(I)(a).

ii. If $\tilde{Z}_{T_{\mathbf{i}},j} \geq 1$ label each of the newly created nodes as "active" on the outbound side. Go to 2(I)(a).

b) For $j \in \{T_{\mathbf{i}}\} \cup \{T_{\mathbf{j}} : T_{\mathbf{j}} \in \hat{I}^-_{k_{in}}$ or $\mathbf{j} \in \hat{A}^-_{k_{in}}\}$:

i. Sample $\tilde{Z}^*_{T_{\mathbf{i}},j}$ to be a Poisson random variable with mean $\tilde{r}^{(m,n)}_{T_{\mathbf{i}},j}$, independently of everything else. If $\tilde{Z}^*_{T_{\mathbf{i}},j} = 0$ go to 2(I)(b).

ii. If $\tilde{Z}^*_{T_{\mathbf{i}},j} \geq 1$ label each of the newly created nodes as "active" on the outbound side. Go to 2(I)(b).

c) Once all the outbound offspring of node $\mathbf{i}$ have been identified, label *identity* $T_{\mathbf{i}}$ as "inactive" on the outbound side.

d) Define the sets $\hat{A}^+_r = \hat{A}^+_{r-1} \cup \{$new "active" nodes created in 2(I)(a)(ii) and 2(I)(b)(ii)$\} \setminus \{\mathbf{i}\}$ and $\hat{I}^-_r = \hat{I}^-_{r-1} \cup \{T_{\mathbf{i}}\}$. This completes Step r on the outbound side.

II. Else:

a) For $j = 1, 2, \ldots, n$:

i. Sample $\tilde{Z}^*_{T_{\mathbf{i}},j}$ to be a Poisson random variable with mean $\tilde{r}^{(m,n)}_{T_{\mathbf{i}},j}$, independently of everything else. If $\tilde{Z}^*_{T_{\mathbf{i}},j} = 0$ go to 2(II)(a).

ii. If $\tilde{Z}^*_{j,T_{\mathbf{i}}} \geq 1$ label each of the newly created nodes as "active" on the outbound side. Go to 2(II)(a).

b) Once all the outbound offspring of node $\mathbf{i}$ have been identified, label *identity* $T_{\mathbf{i}}$ as "inactive" on the outbound side.

c) Define the sets $\hat{A}^+_r = \hat{A}^+_{r-1} \cup \{\text{new "active" nodes created in 2(II)(a)(ii)}\} \setminus \{\mathbf{i}\}$ and $\hat{I}^-_r = \hat{I}^-_{r-1} \cup \{T_{\mathbf{i}}\}$. This completes Step r on the outbound side.

Note: As long as the active sets in the graph and the double tree are the same, the chosen nodes in steps (1)(I) and (2)(I) are the same as the vertices chosen in steps (1) and (2) of the graph exploration process.

Definition 4.4 *We say that the coupling of the graph and the double multi-type branching process holds up to Step r on the inbound side if*

$$A^-_t = \{T_{\mathbf{j}} : \mathbf{j} \in \hat{A}^-_t\} \qquad \text{and} \qquad |A^-_t| = |\hat{A}^-_t| \quad \text{for all } 0 \leq t \leq r,$$

and up to Step r on the outbound side if

$$A^+_t = \{T_{\mathbf{j}} : \mathbf{j} \in \hat{A}^+_t\} \qquad \text{and} \qquad |A^+_t| = |\hat{A}^+_t| \quad \text{for all } 0 \leq t \leq r.$$

Let $T_\emptyset$ denote the identity *of the vertex whose in and out-components we want to explore. Define the stopping time τ^- to be the step in the graph exploration process of vertex $T_\emptyset$ during which the coupling breaks on the inbound side and τ^+ to be the step during which it breaks on the outbound side.*

Remark 4.5 *Note that $\tau^- = r$ if and only if either:*

a. For any $j = 1, 2, \ldots, n$, $j \notin \{T^-_r\} \cup A^-_{r-1} \cup I^-_{r-1}$, we have $Z_{j,T^-_r} \neq Y_{j,T^-_r}$ in step (1)(I)(a)(i),

b. For any $j \in A^-_{r-1} \cup I^-_{r-1}$ we have $Z_{j,T^-_r} \geq 1$ in step (1)(I)(a)(i),

c. $Z^*_{T^-_r,T^-_r} \geq 1$ in step (1)(I)(b)(i),

and $\tau^+ = r$ if and only if either:

d. For any $j = 1, 2, \ldots, n$, $j \notin \{T^+_r\} \cup I^-_{k_{in}} \cup A^-_{k_{in}} \cup A^+_{r-1} \cup I^+_{r-1}$, we have $\tilde{Z}_{T^+_r,j} \neq Y_{T^+_r,j}$ in step (2)(I)(a)(i),

e. For any $j \in A^+_{r-1} \cup I^+_{r-1}$ we have $\tilde{Z}_{T^+_r,j} \geq 1$ in step (2)(I)(a)(i),

f. For any $j \in \{T^+_r\} \cup I^-_{k_{in}} \cup A^-_{k_{in}}$, we have $\tilde{Z}^*_{T^+_r,j} \geq 1$ in step (2)(I)(b)(i).

We are now ready to state our main coupling result, which provides an explicit upper bound for the probability that the coupling breaks before we can determine whether both the in-component and the out-component of the vertex being explored have at least k vertices each or are fully explored.

Throughout the remainder of the paper, we use the notation $\mathbb{P}_i(\cdot) = \mathbb{E}[1(\cdot)|A_0 = \{i\}]$ and $\mathbb{E}_i[\cdot] = \mathbb{E}[\cdot|A_0 = \{i\}]$; also, $\|\mathbf{x}\|_1 = \sum_i |x_i|$ for any $\mathbf{x} \in \mathbb{R}^n$. Similarly to the definition of $\lambda_-(\mathbf{x})$ and $\lambda_+(\mathbf{x})$, define

$$\lambda_-^{(m)}(\mathbf{x}) = \int_{\mathcal{S}} \kappa_m(\mathbf{y}, \mathbf{x})\mu(d\mathbf{y}) \qquad \text{and} \qquad \lambda_+^{(m)}(\mathbf{x}) = \int_{\mathcal{S}} \kappa_m(\mathbf{x}, \mathbf{y})\mu(d\mathbf{y}),$$

$$\lambda_{m,n}^-(\mathbf{x}) = \int_{\mathcal{S}} \kappa_m(\mathbf{y}, \mathbf{x})\mu_n(d\mathbf{y}) \qquad \text{and} \qquad \lambda_{m,n}^+(\mathbf{x}) = \int_{\mathcal{S}} \kappa_m(\mathbf{x}, \mathbf{y})\mu_n(d\mathbf{y}),$$

and

$$\lambda_n^-(\mathbf{x}) = \int_{\mathcal{S}} \kappa(\mathbf{y}, \mathbf{x})\mu_n(d\mathbf{y}) \qquad \text{and} \qquad \lambda_n^+(\mathbf{x}) = \int_{\mathcal{S}} \kappa(\mathbf{x}, \mathbf{y})\mu_n(d\mathbf{y}).$$

Theorem 4.6 *Consider the exploration process described above along with its coupled double tree construction. Define for any fixed $k \in \mathbb{N}_+$ the stopping times $\sigma_k^- = \inf\{t \geq 1 : |A_t^-| + |I_t^-| \geq k \text{ or } A_t^- = \varnothing\}$ and $\sigma_k^+ = \{t \geq 1 : |A_t^+| + |I_t^+| \geq k \text{ or } A_t^+ = \varnothing\}$. For any $0 < \epsilon < 1/2$ and any $n, m \in \mathbb{N}_+$,*

$$\frac{1}{n}\sum_{i=1}^n \mathbb{P}_i\left(\{\tau^- \leq \sigma_k^-\} \cup \{\tau^+ \leq \sigma_k^+\}\right) \leq H(n, m, k, \epsilon),$$

where

$$H(n, m, k, \epsilon) = 1(\Omega_{m,n}^c) + 4\epsilon k^2 + 2\epsilon k^2 \left(1 + \sup_{\mathbf{x}\in\mathcal{S}} \lambda_-^{(m)}(\mathbf{x})\right) + 1(\Omega_{m,n}) \sum_{r=1}^k \sum_{s=0}^{r-1} \binom{r-1}{s} 2^{r-1-s}$$
$$\cdot \left\{ \int_{\mathcal{S}} (\Gamma_-^{(m,n)})^s g_{m,n,\epsilon}^-(\mathbf{x})\mu_n(d\mathbf{x}) + \int_{\mathcal{S}} (\Gamma_+^{(m,n)})^s g_{m,n,\epsilon}^+(\mathbf{x})\mu_n(d\mathbf{x}) \right\},$$

the linear integral operators $\Gamma_-^{(m,n)}$ and $\Gamma_+^{(m,n)}$ are defined in Lemma 4.9, the functions $g_{m,n,\epsilon}^-$ and $g_{m,n,\epsilon}^+$ are defined according to

$$g_{m,n,\epsilon}^-(\mathbf{X}_i) = \min\left\{1,\ (1+5\epsilon)\lambda_n^-(\mathbf{X}_i) - \lambda_{m,n}^-(\mathbf{X}_i) + (1+\epsilon)\sum_{j=1}^n (p_{ji}^{(n)} + q_{ji}^{(n)})1(B_{ji}^c)\right\},$$

$$g_{m,n,\epsilon}^+(\mathbf{X}_i) = \min\left\{1,\ (1+5\epsilon)\lambda_n^+(\mathbf{X}_i) - \lambda_{m,n}^+(\mathbf{X}_i) + (1+\epsilon)\sum_{j=1}^n (p_{ij}^{(n)} + q_{ij}^{(n)})1(B_{ij}^c)\right\},$$

$$\Omega_{m,n} = \bigcap_{t=1}^{M_m} \left\{ \left| \frac{\mu(\mathcal{J}_t^{(m)})}{\mu_n(\mathcal{J}_t^{(m)})} - 1 \right| 1(\mu_n(\mathcal{J}_t^{(m)}) > 0) < \epsilon \right\},$$

$$B_{ij} = \left\{(1-\epsilon)q_{ij}^{(n)} \leq p_{ij}^{(n)} \leq (1+\epsilon)q_{ij}^{(n)},\ q_{ij}^{(n)} \leq \epsilon\right\}.$$

Moreover, $H(n, m, k, \epsilon) \xrightarrow{P} \hat{H}(m, k, \epsilon)$ (defined in Lemma 4.10) as $n \to \infty$ and satisfies

$$\lim_{m\nearrow\infty} \lim_{\epsilon\downarrow 0} \hat{H}(m, k, \epsilon) = 0$$

for any fixed $k \geq 1$.

Before proving the theorem, we will state and prove several preliminary results. The first one below gives an upper bound for the number of offspring sampled in each side of the double-tree $(\mathcal{T}_\mu^-(\kappa_m), \mathcal{T}_\mu^+(\kappa_m))$ up to step $\hat{\sigma}_k^-$ and step $\hat{\sigma}_k^+$, respectively.

Lemma 4.7 *Let* $\hat{\sigma}_k^- = \inf\{t \geq 1 : |\hat{A}_t^-| + |\hat{I}_t^-| \geq k$ *or* $\hat{A}_t^- = \varnothing\}$ *and* $\hat{\sigma}_k^+ = \inf\{t \geq 1 : |\hat{A}_t^+| + |\hat{I}_t^+| \geq k$ *or* $\hat{A}_t^+ = \varnothing\}$*. Then,*

$$\frac{1}{n}\sum_{i=1}^n \mathbb{E}_i\left[\left|\hat{I}^-_{\hat{\sigma}_k^-}\right| + \left|\hat{A}^-_{\hat{\sigma}_k^-}\right|\right] \leq k + k\sup_{\mathbf{x}\in\mathcal{S}} \lambda_-^{(m)}(\mathbf{x}) \quad \text{and} \quad \frac{1}{n}\sum_{i=1}^n \mathbb{E}_i\left[\left|\hat{I}^+_{\hat{\sigma}_k^+}\right| + \left|\hat{A}^+_{\hat{\sigma}_k^+}\right|\right] \leq k + k\sup_{\mathbf{x}\in\mathcal{S}} \lambda_+^{(m)}(\mathbf{x}).$$

Proof. Define $\mathcal{G}_r^-$ to be the sigma-algebra containing all the information of the exploration process of the in-component of vertex i up to the end of Step r and including the *identity* of the active node T_{r+1}^-. Note that

$$\begin{aligned}
&\mathbb{E}_i\left[\left|\hat{I}^-_{\hat{\sigma}_k^-}\right| + \left|\hat{A}^-_{\hat{\sigma}_k^-}\right|\right] \\
&= \mathbb{E}_i\left[\left|\hat{I}^-_{\hat{\sigma}_k^- -1}\right| + \left|\hat{A}^-_{\hat{\sigma}_k^- -1}\right| + \sum_{j=1}^n Z_{j,\hat{T}^-_{\hat{\sigma}_k^-}}\right] \\
&\leq k - 1 + \sum_{r=1}^k \mathbb{E}_i\left[1(\hat{\sigma}_k^- = r)\sum_{j=1}^n Z_{j,\hat{T}_r^-}\right] \\
&= k - 1 + \sum_{r=1}^k \mathbb{E}_i\left[1(\hat{\sigma}_k^- > r-1)\mathbb{E}\left[1\left(\sum_{j=1}^n Z_{j,\hat{T}_r^-} \geq k - \left|\hat{A}_{r-1}^-\right| - \left|\hat{I}_{r-1}^-\right|\right)\sum_{j=1}^n Z_{j,\hat{T}_r^-}\middle|\mathcal{G}_{r-1}^-\right]\right].
\end{aligned}$$

Note that in the last equality the term that would correspond to $\{\hat{A}_r^- = \varnothing\}$ in the description of the event $\{\hat{\sigma}_k^- = r\}$ vanishes since $\sum_{j=1}^n Z_{j,\hat{T}_r^-} = 0$ in that case. Now use the observation that $\sum_{j=1}^n Z_{ji}$ is a Poisson random variable with mean $\sum_{j=1}^n r_{ji}^{(m,n)} = \lambda_-^{(m)}(\mathbf{X}_i)$, and the identity $E[X1(X \geq j)] \leq E[X] = \lambda$ when X is Poisson(λ), to obtain that

$$\begin{aligned}
&\sum_{r=1}^k \mathbb{E}_i\left[1(\hat{\sigma}_k^- > r-1)\mathbb{E}\left[1\left(\sum_{j=1}^n Z_{j,\hat{T}_r^-} \geq k - \left|\hat{A}_{r-1}^-\right| - \left|\hat{I}_{r-1}^-\right|\right)\sum_{j=1}^n Z_{j,\hat{T}_r^-}\middle|\mathcal{G}_{r-1}^-\right]\right] \\
&\leq \sum_{r=1}^k \mathbb{E}_i\left[1(\hat{\sigma}_k^- > r-1)\lambda_-^{(m)}(\mathbf{X}_{\hat{T}_r^-})\right] \\
&\leq k\sup_{\mathbf{x}\in\mathcal{S}} \lambda_-^{(m)}(\mathbf{x}).
\end{aligned}$$

The proof for the outbound tree is essentially the same and is therefore omitted. ■

The next result is a technical lemma giving an explicit upper bound for the ratio of independent Poisson random variables.

Lemma 4.8 *Let X, Y be independent Poisson random variables with means λ and μ, respectively. Let $a, b \geq 0$. Then,*

$$E\left[\frac{a+X}{b+X+Y}\cdot 1(b+X+Y\geq 1)\right] \leq \frac{2a}{b+1}+\frac{\lambda}{\lambda+\mu}(1-e^{-\lambda-\mu}).$$

Proof. Recall that X given $X+Y=n$ is a Binomial$(n, \lambda/(\lambda+\mu))$. Hence,

$$\begin{aligned}
&E\left[\frac{a+X}{b+X+Y}\cdot 1(b+X+Y\geq 1)\right] \\
&= E\left[\frac{a}{b}\cdot 1(X+Y=0, b\geq 1)\right] + E\left[\frac{a+X}{b+X+Y}\cdot 1(X+Y\geq 1)\right] \\
&= \frac{a}{b}1(b\geq 1)P(X+Y=0) + \sum_{n=1}^{\infty}\frac{E[a+X|X+Y=n]}{b+n}P(X+Y=n).
\end{aligned}$$

Now use the observation that X given $X+Y=n$ is a binomial with parameters $(n, \lambda/(\mu+\lambda))$ to obtain that

$$\begin{aligned}
&\sum_{n=1}^{\infty}\frac{E[a+X|X+Y=n]}{b+n}P(X+Y=n) \\
&= \sum_{n=1}^{\infty}\frac{a+n\lambda/(\mu+\lambda)}{b+n}P(X+Y=n) \\
&= a\sum_{n=1}^{\infty}\frac{1}{b+n}P(X+Y=n) + \frac{\lambda}{\mu+\lambda}\sum_{n=1}^{\infty}\frac{n}{b+n}P(X+Y=n) \\
&\leq \frac{a}{b+1}P(X+Y\geq 1) + \frac{\lambda}{\mu+\lambda}P(X+Y\geq 1) \\
&= \left(\frac{a}{b+1}+\frac{\lambda}{\lambda+\mu}\right)P(X+Y\geq 1).
\end{aligned}$$

Using the observation that $(a/b)1(b\geq 1)\leq 2a/(b+1)$ gives that

$$E\left[\frac{a+X}{b+X+Y}\cdot 1(b+X+Y\geq 1)\right] \leq \frac{2a}{b+1}+\frac{\lambda}{\lambda+\mu}P(X+Y\geq 1),$$

which completes the proof. ■

The following result constitutes a key step of the proof of Theorem 4.6 by providing an upper estimate for the distribution of the *identities* of the active nodes $\hat{T}_r^-$ and $\hat{T}_r^+$.

Lemma 4.9 *Let h be a nonnegative function on $\mathcal{S}$, then*

$$\mathbb{E}_i\left[1(\hat{A}_{r-1}^- \neq \varnothing)h(\mathbf{X}_{\hat{T}_r^-})\right] \leq \sum_{s=0}^{r-1}\binom{r-1}{s}2^{r-1-s}(\Gamma_-^{(m,n)})^s h(\mathbf{X}_i)$$

and

$$\mathbb{E}_i\left[1(\hat{A}_{r-1}^+ \neq \varnothing)h(\mathbf{X}_{\hat{T}_r^+})\right] \leq \sum_{s=0}^{r-1}\binom{r-1}{s}2^{r-1-s}(\Gamma_+^{(m,n)})^s h(\mathbf{X}_i),$$

where $\Gamma_-^{(m,n)}$ *and* $\Gamma_+^{(m,n)}$ *are the following linear integral operators:*

$$\Gamma_-^{(m,n)}h(\mathbf{x}) = \int_{\mathcal{S}} \frac{\mu(\mathcal{J}_{\vartheta(\mathbf{y})}^{(m)})}{\mu_n(\mathcal{J}_{\vartheta(\mathbf{y})}^{(m)})} \cdot \frac{(1-e^{-\lambda_-^{(m)}(\mathbf{x})})}{\lambda_-^{(m)}(\mathbf{x})} \cdot \kappa_m(\mathbf{y},\mathbf{x})h(\mathbf{y})\,\mu_n(d\mathbf{y})$$

and

$$\Gamma_+^{(m,n)}h(\mathbf{x}) = \int_{\mathcal{S}} \frac{\mu(\mathcal{J}_{\vartheta(\mathbf{y})}^{(m)})}{\mu_n(\mathcal{J}_{\vartheta(\mathbf{y})}^{(m)})} \cdot \frac{(1-e^{-\lambda_+^{(m)}(\mathbf{x})})}{\lambda_+^{(m)}(\mathbf{x})} \cdot \kappa_m(\mathbf{x},\mathbf{y})h(\mathbf{y})\,\mu_n(d\mathbf{y}),$$

with $\vartheta(\mathbf{x}) = t$ *if and only if* $\mathbf{x} \in \mathcal{J}_t^{(m)}$.

Proof. Let $\mathbf{W}_t^- = (W_{t,1}^-, \dots, W_{t,n}^-)$ denote the process that keeps track of the *identities* of the vertices in the active set $\hat{A}_t^-$ for $t \geq 0$; that is, $W_{t,j}^-$ denotes the number of tree nodes with *identity* j in $\hat{A}_t^-$. Then,

$$\begin{aligned}
\mathbb{P}_i(\hat{A}_{r-1}^- \neq \varnothing, \hat{T}_r^- = l) &= \mathbb{E}_i\left[1\left(\|\mathbf{W}_{r-1}^-\|_1 \geq 1\right)\mathbb{P}(\hat{T}_r^- = l|\mathbf{W}_{r-1}^-)\right] \\
&= \mathbb{E}_i\left[\frac{W_{r-1,l}^-}{\|\mathbf{W}_{r-1}^-\|_1} \cdot 1\left(\|\mathbf{W}_{r-1}^-\|_1 \geq 1\right)\right] \\
&= \mathbb{E}_i\left[\mathbb{E}\left[\left.\frac{W_{r-1,l}^-}{\|\mathbf{W}_{r-1}^-\|_1} \cdot 1\left(\|\mathbf{W}_{r-1}^-\|_1 \geq 1\right)\right| \mathbf{W}_{r-2}^-\right] 1(\|\mathbf{W}_{r-2}^-\|_1 \geq 1)\right],
\end{aligned}$$

where $\|(x_1,\dots,x_n)\|_1 = |x_1| + \cdots + |x_n|$. Now let $(h_1,\dots,h_n) = (h(\mathbf{X}_1),\dots,h(\mathbf{X}_n))$ and note that

$$\begin{aligned}
\mathbb{E}_i\left[1(\hat{A}_{r-1}^- \neq \varnothing)h(\mathbf{X}_{\hat{T}_r^-})\right] &= \sum_{l=1}^n h_l \mathbb{P}_i(\hat{A}_{r-1}^- \neq \varnothing, \hat{T}_r^- = l) \\
&= \mathbb{E}_i\left[\sum_{l=1}^n h_l \mathbb{E}\left[\left.\frac{W_{r-1,l}^-}{\|\mathbf{W}_{r-1}^-\|_1} \cdot 1\left(\|\mathbf{W}_{r-1}^-\|_1 \geq 1\right)\right| \mathbf{W}_{r-2}^-\right] 1(\|\mathbf{W}_{r-2}^-\|_1 \geq 1)\right].
\end{aligned}$$

Moreover, provided $\|\mathbf{W}_{r-2}^-\|_1 \geq 1$, we have

$$\begin{aligned}
&\mathbb{E}\left[\left.\frac{W_{r-1,l}^-}{\|\mathbf{W}_{r-1}^-\|_1} \cdot 1\left(\|\mathbf{W}_{r-1}^-\|_1 \geq 1\right)\right| \mathbf{W}_{r-2}^-\right] \\
&= \sum_{s=1}^n \mathbb{P}(\hat{T}_{r-1}^- = s|\mathbf{W}_{r-2}^-)\,\mathbb{E}\left[\left.\frac{W_{r-1,l}^-}{\|\mathbf{W}_{r-1}^-\|_1} \cdot 1\left(\|\mathbf{W}_{r-1}^-\|_1 \geq 1\right)\right| \mathbf{W}_{r-2}^-, \{\hat{T}_{r-1}^- = s\}\right] \\
&= \sum_{s=1}^n \frac{W_{r-2,s}^-}{\|\mathbf{W}_{r-2}^-\|_1}\mathbb{E}\left[\left.\frac{W_{r-2,l}^- + Z_{ls} - 1(s=l)}{\sum_{j=1}^n (W_{r-2,j}^- + Z_{js}) - 1} \cdot 1\left(\sum_{j=1}^n (W_{r-2,j}^- + Z_{js}) \geq 2\right)\right| \mathbf{W}_{r-2}^-\right] \\
&\leq \sum_{s=1}^n \frac{W_{r-2,s}^-}{\|\mathbf{W}_{r-2}^-\|_1}\mathbb{E}\left[\left.\frac{W_{r-2,l}^- + Z_{ls}}{\sum_{j=1}^n (W_{r-2,j}^- + Z_{js}) - 1} \cdot 1\left(\sum_{j=1}^n (W_{r-2,j}^- + Z_{js}) \geq 2\right)\right| \mathbf{W}_{r-2}^-\right].
\end{aligned}$$

Now use Lemma 4.8 with $a = W^-_{r-2,l}$, $b = \sum_{j=1}^n W^-_{r-2,j} - 1$, $X = Z_{ls}$ and $Y = \sum_{j\neq l} Z_{js}$ to obtain that

$$\begin{aligned}
&\sum_{s=1}^n \frac{W^-_{r-2,s}}{\|\mathbf{W}^-_{r-2}\|_1}\mathbb{E}\left[\frac{W^-_{r-2,l} + Z_{ls}}{\sum_{j=1}^n (W^-_{r-2,j} + Z_{js}) - 1} \cdot 1\left(\sum_{j=1}^n (W^-_{r-2,j} + Z_{js}) \geq 2 \right) \middle| \mathbf{W}^-_{r-2} \right] \\
&\leq \sum_{s=1}^n \frac{W^-_{r-2,s}}{\|\mathbf{W}^-_{r-2}\|_1} \left(\frac{2W^-_{r-2,l}}{\|\mathbf{W}^-_{r-2}\|_1} + \frac{r^{(m,n)}_{ls}}{\sum_{j=1}^n r^{(m,n)}_{js}} \left(1 - e^{-\sum_{j=1}^n r^{(m,n)}_{js}}\right)\right) \\
&=: \frac{2W^-_{r-2,l}}{\|\mathbf{W}^-_{r-2}\|_1} + \sum_{s=1}^n \frac{W^-_{r-2,s}}{\|\mathbf{W}^-_{r-2}\|_1} \cdot \gamma^{(m,n)}_{ls},
\end{aligned}$$

where

$$\gamma^{(m,n)}_{ls} = \frac{r^{(m,n)}_{ls}}{\sum_{j=1}^n r^{(m,n)}_{js}} \left(1 - e^{-\sum_{j=1}^n r^{(m,n)}_{js}}\right) = r^{(m,n)}_{ls} \frac{(1 - e^{-\lambda^{(m)}_-(\mathbf{X}_s)})}{\lambda^{(m)}_-(\mathbf{X}_s))},$$

and we use the convention that $(1 - e^{-0})/0 \equiv 1$. It follows that

$$\begin{aligned}
&\mathbb{E}_i \left[1(\hat{A}^-_{r-1} \neq \varnothing) h(\mathbf{X}_{\hat{T}^-_r}) \right] \\
&\leq \mathbb{E}_i \left[\sum_{l=1}^n h_l \left\{ \frac{2W^-_{r-2,l}}{\|\mathbf{W}^-_{r-2}\|_1} + \sum_{s=1}^n \frac{W^-_{r-2,s}}{\|\mathbf{W}^-_{r-2}\|_1} \cdot \gamma^{(m,n)}_{ls} \right\} 1(\|\mathbf{W}^-_{r-2}\|_1 \geq 1) \right] \\
&= 2\mathbb{E}_i \left[1(\hat{A}^-_{r-2} \neq \varnothing) h_{\hat{T}^-_{r-1}} \right] + \mathbb{E}_i \left[\sum_{s=1}^n \frac{W^-_{r-2,s}}{\|\mathbf{W}^-_{r-2}\|_1} 1(\|\mathbf{W}^-_{r-2}\|_1 \geq 1) \sum_{l=1}^n h_l \cdot \gamma^{(m,n)}_{ls} \right] \\
&= 2\mathbb{E}_i \left[1(\hat{A}^-_{r-2} \neq \varnothing) h_{\hat{T}^-_{r-1}} \right] + \mathbb{E}_i \left[\sum_{s=1}^n P(\hat{T}^-_{r-1} = s | \mathbf{W}^-_{r-2}) 1(\|\mathbf{W}^-_{r-2}\|_1 \geq 1) \Gamma^{(m,n)}_- h(\mathbf{X}_s) \right] \\
&= 2\mathbb{E}_i \left[1(\hat{A}^-_{r-2} \neq \varnothing) h(\mathbf{X}_{\hat{T}^-_{r-1}}) \right] + \mathbb{E}_i \left[1(\hat{A}_{r-2} \neq \varnothing) \Gamma^{(m,n)}_- h(\mathbf{X}_{\hat{T}^-_{r-1}}) \right].
\end{aligned}$$

Letting $a_{r,s} = \mathbb{E}_i \left[1(\hat{A}^-_{r-1} \neq \varnothing) (\Gamma^{(m,n)}_-)^s h(\mathbf{X}_{\hat{T}^-_r}) \right]$, and iterating $r - 2$ times we obtain that

$$a_{r,0} \leq 2a_{r-1,0} + a_{r-1,1} \leq \sum_{s=0}^{r-1} \binom{r-1}{s} 2^{r-1-s} a_{1,s},$$

which yields

$$\mathbb{E}_i \left[1(\hat{A}^-_{r-1} \neq \varnothing) h(\mathbf{X}_{\hat{T}^-_r}) \right] \leq \sum_{s=0}^{r-1} \binom{r-1}{s} 2^{r-1-s} (\Gamma^{(m,n)}_-)^s h(\mathbf{X}_i).$$

The proof for $\mathbb{E}_i \left[1(\hat{A}^+_{r-1} \neq \varnothing) h(\mathbf{X}_{\hat{T}^+_r}) \right]$ is essentially the same and is therefore omitted. ■

Lemma 4.10 *Let $H(n, m, k, \epsilon)$ be defined as in Theorem 4.6, then*

$$H(n, m, k, \epsilon) \xrightarrow{P} \hat{H}(m, k, \epsilon) \qquad n \to \infty,$$

where

$$\hat{H}(m,k,\epsilon) = 4\epsilon k^2 + 2\epsilon k^2 \left(1 + \sup_{\mathbf{x}\in\mathcal{S}} \lambda_-^{(m)}(\mathbf{x})\right) + \sum_{r=1}^{k} \sum_{s=0}^{r-1} \binom{r-1}{s} 2^{r-1-s}$$
$$\cdot \left\{ \int_{\mathcal{S}} (\Gamma_-^{(m)})^s g_{m,\epsilon}^-(\mathbf{x})\mu(d\mathbf{x}) + \int_{\mathcal{S}} (\Gamma_+^{(m)})^s g_{m,\epsilon}^+(\mathbf{x})\mu(d\mathbf{x}) \right\},$$

where

$$g_{m,\epsilon}^-(\mathbf{x}) = \min\left\{1,\ (1+5\epsilon)\lambda_-(\mathbf{x}) - \lambda_-^{(m)}(\mathbf{x})\right\},$$
$$g_{m,\epsilon}^+(\mathbf{x}) = \min\left\{1,\ (1+5\epsilon)\lambda_+(\mathbf{x}) - \lambda_+^{(m)}(\mathbf{x})\right\},$$

and the linear integral operators $\Gamma_-^{(m)}$ *and* $\Gamma_+^{(m)}$ *are given by*

$$\Gamma_-^{(m)} h(\mathbf{x}) = \int_{\mathcal{S}} \frac{(1 - e^{-\lambda_-^{(m)}(\mathbf{x})})}{\lambda_-^{(m)}(\mathbf{x})} \cdot \kappa_m(\mathbf{y},\mathbf{x}) h(\mathbf{y})\mu(d\mathbf{y}),$$
$$\Gamma_+^{(m)} h(\mathbf{x}) = \int_{\mathcal{S}} \frac{(1 - e^{-\lambda_+^{(m)}(\mathbf{x})})}{\lambda_+^{(m)}(\mathbf{x})} \cdot \kappa_m(\mathbf{x},\mathbf{y}) h(\mathbf{y})\mu(d\mathbf{y}).$$

Furthermore, for any fixed $k \geq 1$,

$$\lim_{m\nearrow\infty} \lim_{\epsilon\downarrow 0} \hat{H}(m,k,\epsilon) = 0.$$

Proof. Start by noting that Assumption 3.1(a) implies that $1(\Omega_{m,n}) \xrightarrow{P} 1$ as $n \to \infty$, so the convergence of $H(n,m,k,\epsilon)$ will follow once we show that

$$\int_{\mathcal{S}} (\Gamma_-^{(m,n)})^s g_{m,n,\epsilon}^-(\mathbf{x})\mu_n(d\mathbf{x}) \xrightarrow{P} \int_{\mathcal{S}} (\Gamma_-^{(m)})^s g_{m,\epsilon}^-(\mathbf{x})\mu(d\mathbf{x}) \tag{4.11}$$

and

$$\int_{\mathcal{S}} (\Gamma_+^{(m,n)})^s g_{m,n,\epsilon}^+(\mathbf{x})\mu_n(d\mathbf{x}) \xrightarrow{P} \int_{\mathcal{S}} (\Gamma_+^{(m)})^s g_{m,\epsilon}^+(\mathbf{x})\mu(d\mathbf{x}) \tag{4.12}$$

as $n \to \infty$ for any fixed $s \in \mathbb{N}$. Let

$$w_m^-(\mathbf{y},\mathbf{x}) = \frac{(1 - e^{-\lambda_-^{(m)}(\mathbf{x})})}{\lambda_-^{(m)}(\mathbf{x})} \cdot \kappa_m(\mathbf{y},\mathbf{x}) \qquad \text{and} \qquad r_n(\mathbf{y}) = \frac{\mu(\mathcal{J}_{\vartheta(\mathbf{y})}^{(m)})}{\mu_n(\mathcal{J}_{\vartheta(\mathbf{y})}^{(m)})},$$

and note that for any function h and $\mathbf{x} \in \mathcal{J}_j^{(m)}$, we have

$$\Gamma_-^{(m,n)} h(\mathbf{x}) = \int_{\mathcal{S}} r_n(\mathbf{y}) w_m^-(\mathbf{y},\mathbf{x}) h(\mathbf{y})\mu_n(d\mathbf{y}) =: \sum_{i=1}^{M_m} \mathcal{I}_i^{(m,n)}(h) d_{i,j}^{(m,n)},$$

where

$$d_{i,j}^{(m,n)} = r_n(\mathbf{y}) w_m^-(\mathbf{y},\mathbf{x}) \quad \text{for all } \mathbf{y} \in \mathcal{J}_i^{(m)}, \mathbf{x} \in \mathcal{J}_j^{(m)},$$
$$\mathcal{I}_i^{(m,n)}(h) = \int_{\mathcal{J}_i^{(m)}} h(\mathbf{y}) \mu_n(d\mathbf{y}).$$

In general, for $s \geq 1$, we have that

$$(\Gamma_-^{(m,n)})^s h(\mathbf{x}) = \sum_{i=1}^{M_m} \mathcal{I}_i^{(m,n)}(h)((\mathbf{D}^{(m,n)})^s)_{i,j} \qquad \text{for } \mathbf{x} \in \mathcal{J}_j^{(m)}$$

and

$$\int_{\mathcal{S}} (\Gamma_-^{(m,n)})^s h(\mathbf{x}) \mu_n(d\mathbf{x}) = \sum_{j=1}^{M_m} \sum_{i=1}^{M_m} \mathcal{I}_i^{(m,n)}(h)((\mathbf{D}^{(m,n)})^s)_{i,j} \mu_n(\mathcal{J}_j^{(m)}),$$

where $\mathbf{D}^{(m,n)}$ is the $M_m \times M_m$ matrix whose (i,j)th component is $d_{i,j}^{(m,n)}$. Define $\mathbf{D}^{(m)}$ to be the matrix whose (i,j)th component is $d_{i,j}^{(m)} = w_m^-(\mathbf{y},\mathbf{x})$ for all $\mathbf{y} \in \mathcal{J}_i^{(m)}, \mathbf{x} \in \mathcal{J}_j^{(m)}$. Since by Assumption 3.1(a) we have that $\mu_n(\mathcal{J}_j^{(m)}) \xrightarrow{P} \mu(\mathcal{J}_j^{(m)})$ and $d_{i,j}^{(m,n)} \xrightarrow{P} d_{i,j}^{(m)}$ as $n \to \infty$ for all $1 \leq i,j \leq M_m$, it follows that

$$\lim_{n\to\infty} \int_{\mathcal{S}} (\Gamma_-^{(m,n)})^s g_{m,n,\epsilon}^-(\mathbf{x}) \mu_n(d\mathbf{x}) = \sum_{j=1}^{M_m} \sum_{i=1}^{M_m} \left(\lim_{n\to\infty} \mathcal{I}_i^{(m,n)}(g_{m,n,\epsilon}^-) \right) ((\mathbf{D}^{(m)})^s)_{i,j} \mu(\mathcal{J}_j^{(m)}),$$

assuming the last limit exists for each $1 \leq i \leq M_m$. To see that it does let $\hat{g}_{m,n,\epsilon}^-(\mathbf{x}) = \min\left\{1, \, (1+5\epsilon)\lambda_n^-(\mathbf{x}) - \lambda_{m,n}^-(\mathbf{x})\right\}$ and note that by Lemma 4.1,

$$\left| \mathcal{I}_i^{(m,n)}(g_{m,n,\epsilon}^-) - \mathcal{I}_i^{(m,n)}(\hat{g}_{m,n,\epsilon}^-) \right| \leq (1+\epsilon) \frac{1}{n} \sum_{l \in \mathcal{J}_i^{(m)}} \sum_{j=1}^{n} (p_{jl}^{(n)} + q_{jl}^{(n)}) 1(B_{jl}^c) \xrightarrow{P} 0$$

as $n \to \infty$. Now let $\mathbf{X}^{(n)}$ and $\mathbf{Y}^{(n)}$ be conditionally i.i.d. random variables (given $\mathscr{F}$) having distribution μ_n (as constructed in Lemma 4.1). Assumption 3.1 implies that $(\mathbf{X}^{(n)}, \mathbf{Y}^{(n)}) \Rightarrow (\mathbf{X}, \mathbf{Y})$ as $n \to \infty$, where $\mathbf{X}$ and $\mathbf{Y}$ are i.i.d. with distribution μ, and Lemma 4.2 gives $\mathbb{E}\left[\kappa(\mathbf{X}^{(n)}, \mathbf{Y}^{(n)})\right] \xrightarrow{P} E[\kappa(\mathbf{X}, \mathbf{Y})]$. Therefore, by bounded convergence,

$$\begin{aligned}
\mathcal{I}_i^{(m,n)}(\hat{g}_{m,n,\epsilon}^-) &= \mathbb{E}\left[1(\mathbf{X}^{(n)} \in \mathcal{J}_i^{(m)}) \min\left\{1, \, (1+5\epsilon)\mathbb{E}\left[\kappa(\mathbf{X}^{(n)}, \mathbf{Y}^{(n)}) - \kappa_m(\mathbf{X}^{(n)}, \mathbf{Y}^{(n)}) \middle| \mathbf{X}^{(n)}\right]\right\}\right] \\
&\xrightarrow{P} E\left[1(\mathbf{X} \in \mathcal{J}_i^{(m)}) \min\left\{1, \, (1+5\epsilon)E\left[\kappa(\mathbf{X}, \mathbf{Y}) - \kappa_m(\mathbf{X}, \mathbf{Y}) | \, \mathbf{X}\right]\right\}\right] \\
&=: \mathcal{I}_i^{(m)}(g_{m,\epsilon}^-),
\end{aligned}$$

as $n \to \infty$. We conclude that

$$\mathcal{I}_i^{(m,n)}(g_{m,n,\epsilon}^-) \xrightarrow{P} \mathcal{I}_i^{(m)}(g_{m,\epsilon}^-)$$

as $n \to \infty$, and noting that

$$\sum_{j=1}^{M_m} \sum_{i=1}^{M_m} \mathcal{I}_i^{(m)}(g_{m,\epsilon}^-)((\mathbf{D}^{(m)})^s)_{i,j} \mu(\mathcal{J}_j^{(m)}) = \int_{\mathcal{S}} (\Gamma_-^{(m)})^s g_{m,\epsilon}^-(\mathbf{x}) \mu(d\mathbf{x})$$

completes the proof of (4.11). The proof for (4.12) is essentially the same and is therefore omitted. This concludes the proof that $H(n,m,k,\epsilon) \xrightarrow{P} \hat{H}(m,k,\epsilon)$ as $n \to \infty$. To compute the limit of $\hat{H}(m,k,\epsilon)$ note that by monotone convergence,

$$\lim_{\epsilon\downarrow 0} \hat{H}(m,k,\epsilon) = \sum_{k=1}^{k}\sum_{s=0}^{r-1} \binom{r-1}{s} 2^{r-1-s} \left\{ \int_{\mathcal{S}} (\Gamma_-^{(m)})^s g_m^-(\mathbf{x})\mu(d\mathbf{x}) + \int_{\mathcal{S}} (\Gamma_+^{(m)})^s g_m^+(\mathbf{x})\mu(d\mathbf{x}) \right\},$$

where $g_m^{\pm}(\mathbf{x}) = \min\left\{1, \lambda_{\pm}(\mathbf{x}) - \lambda_{\pm}^{(m)}(\mathbf{x})\right\}$. Now let Γ_- and Γ_+ be the linear integral operators defined by

$$\Gamma_- h(\mathbf{x}) = \int_{\mathcal{S}} \frac{1-e^{-\lambda_-(\mathbf{x})}}{\lambda_-(\mathbf{x})} \cdot \kappa(\mathbf{y},\mathbf{x}) h(\mathbf{y})\mu(d\mathbf{y}) \quad \text{and} \quad \Gamma_+ h(\mathbf{x}) = \int_{\mathcal{S}} \frac{1-e^{-\lambda_+(\mathbf{x})}}{\lambda_+(\mathbf{x})} \cdot \kappa(\mathbf{x},\mathbf{y}) h(\mathbf{y})\mu(d\mathbf{y})$$

and note that by monotone convergence,

$$\lim_{m\nearrow\infty} \Gamma_{\pm}^{(m)} h(\mathbf{x}) = \Gamma_{\pm} h(\mathbf{x})$$

for any nonnegative function h. Moreover, for any $h : \mathcal{S} \to [0,1]$, we have that $\Gamma_{\pm}^{(m)} h(\mathbf{x}), \Gamma_{\pm} h(\mathbf{x}) \in [0,1]$, and therefore, the bounded convergence theorem gives

$$\lim_{m\nearrow\infty} \int_{\mathcal{S}} (\Gamma_{\pm}^{(m)})^s g_m^{\pm}(\mathbf{x})\mu(d\mathbf{x}) = \int_{\mathcal{S}} (\Gamma_{\pm})^s \left(\lim_{m\nearrow\infty} g_m^{\pm} \right)(\mathbf{x})\mu(d\mathbf{x}) = 0.$$

This completes the proof. ■

We are now ready to give the proof of Theorem 4.6.

Proof of Theorem 4.6. To start, note that

$$\begin{aligned} &\mathbb{P}_i\left(\{\tau^- \leq \sigma_k^-\} \cup \{\tau^+ \leq \sigma_k^+\}\right) \\ &\leq \left\{ \mathbb{P}_i\left(\tau^- \leq \sigma_k^-\right) + \mathbb{P}_i\left(\{\tau^- > \sigma_k^-\} \cap \{\tau^+ \leq \sigma_k^+\}\right)\right\} 1(\Omega_{m,n}) + 1(\Omega_{m,n}^c), \end{aligned}$$

where the event $\Omega_{m,n}$ is defined in the statement of the theorem. To analyze the two probabilities, define $\mathcal{G}_m^-$ to be the sigma-algebra containing all the information of the exploration process of the in-component of vertex i up to the end of Step m and including the *identity* of the active node T_{m+1}^-, and let $\mathcal{G}_m^+$ be the sigma-algebra containing all the information of the exploration process of the in-component of vertex i up to Step σ_k^-, and of its out-component up to the end of Step m, including the *identity* of the active node T_{m+1}^+; note that $\mathcal{G}_m^- \subseteq \mathcal{G}_r^+$ for all $0 \leq m \leq \sigma_k^-$ and any $r \geq 0$. Next, for any $r \geq 1$ define the events

$$\begin{aligned} E_r^- &= \{|I_r^-| + |A_r^-| < k\}, \\ E_r^+ &= \{|I_r^+| + |A_r^+| < k\}, \\ C_i^-(r) &= \left\{ \max_{j\in[n], j\notin\{i\}\cup A_{r-1}^-\cup I_{r-1}^-} |Z_{ji} - Y_{ji}| + \max_{j\in A_{r-1}^-\cup I_{r-1}^-} Z_{ji} + Z_{ii}^* = 0 \right\}, \end{aligned}$$

$$C_i^+(r) = \left\{ \max_{j\in[n], j\notin\{i\}\cup I^-_{\sigma_k^-}\cup A^-_{\sigma_k^-}\cup A^+_{r-1}\cup I^+_{r-1}} |\tilde{Z}_{ij} - Y_{ij}| + \max_{j\in A^+_{r-1}\cup I^+_{r-1}} \tilde{Z}_{ij} + \max_{j\in\{i\}\cup I^-_{\sigma_k^-}\cup A^-_{\sigma_k^-}} \tilde{Z}^*_{ji} = 0 \right\}.$$

Now, use Remark 4.5 to obtain that on the event $\Omega_{m,n}$,

$$\begin{aligned}
&\mathbb{P}_i\left(\tau^- \le \sigma_k^-\right) + \mathbb{P}_i\left(\{\tau^- > \sigma_k^-\}\cap\{\tau^+ \le \sigma_k^+\}\right) \\
&= \sum_{r=1}^k \left\{\mathbb{P}_i\left(r = \tau^- \le \sigma_k^-\right) + \mathbb{P}_i\left(\{\tau^- > \sigma_k^-\}\cap\{r = \tau^+ \le \sigma_k^+\}\right)\right\} \\
&\le \sum_{r=1}^k \mathbb{P}_i\left(\tau^- > r-1, A^-_{r-1} \neq \varnothing, E^-_{r-1}, (C^-_{T_r^-}(r))^c\right) \\
&\quad + \sum_{r=1}^k \mathbb{P}_i\left(\tau^- > \sigma_k^-,\, \tau^+ > r-1, A^+_{r-1} \neq \varnothing, E^+_{r-1}, (C^+_{T_r^+}(r))^c\right) \\
&= \sum_{r=1}^k \mathbb{E}_i\left[1(\tau^- > r-1, A^-_{r-1} \neq \varnothing, E_{r-1})\mathbb{P}\left((C^-_{T_r^-}(r))^c \middle| \mathcal{G}^-_{r-1}\right)\right] \\
&\quad + \sum_{r=1}^k \mathbb{E}_i\left[1(\tau^- > \sigma_k^-,\, \tau^+ > r-1, A^+_{r-1} \neq \varnothing, E_{r-1})\mathbb{P}\left((C^+_{T_r^+}(r))^c \middle| \mathcal{G}^+_{r-1}\right)\right].
\end{aligned}$$

To analyze the two conditional probabilities in the last expressions, note that the union bound and the independence of the $\{U_{ij} : 1 \le i, j \le n\}$ from everything else give

$$\mathbb{P}\left((C^-_{T_r^-}(r))^c \middle| \mathcal{G}^-_{r-1}\right) \le \sum_{j\in[n], j\notin\{T_r^-\}\cup A^-_{r-1}\cup I^-_{r-1}} \mathbb{P}(|Z_{j,T_r^-} - Y_{j,T_r^-}| > 0 | T_r^-) \tag{4.13}$$

$$+ \sum_{j\in\{T_r^-\}\cup A^-_{r-1}\cup I^-_{r-1}} \mathbb{P}(Z_{j,T_r^-} > 0 | T_r^-), \tag{4.14}$$

and

$$\mathbb{P}\left((C^+_{T_r^+}(r))^c \middle| \mathcal{G}^+_{r-1}\right) \le \sum_{j\in[n], j\notin\{T_r^+\}\cup I^-_{\sigma_k^-}\cup A^-_{\sigma_k^-}\cup A^+_{r-1}\cup I^+_{r-1}} \mathbb{P}(|\tilde{Z}_{T_r^+,j} - Y_{T_r^+,j}| > 0 | T_r^+) \tag{4.15}$$

$$+ \sum_{j\in\{T_r^+\}\cup I^-_{\sigma_k^-}\cup A^-_{\sigma_k^-}\cup A^+_{r-1}\cup I^+_{r-1}} \mathbb{P}(\tilde{Z}_{T_r^+,j} > 0 | T_r^+). \tag{4.16}$$

To analyze (4.13) note that on the event B_{ji} we have that $(1-\epsilon)q_{ji}^{(n)} \le p_{ji}^{(n)} < (1+\epsilon)q_{ji}^{(n)} \le (1+\epsilon)\epsilon < 1$, which implies that on the event B_{ji} we have

$$\begin{aligned}
\mathbb{P}(|Y_{ji} - Z_{ji}| > 0) &= (p_{ji}^{(n)} - r_{ji}^{(m,n)})1(p_{ji}^{(n)} > r_{ji}^{(m,n)}) \\
&\quad + (e^{-r_{ji}^{(m,n)}} - 1 + p_{ji}^{(n)})1(1 - e^{-r_{ji}^{(m,n)}} < p_{ji}^{(n)} \le r_{ji}^{(m,n)}) \\
&\quad + (1 - p_{ji}^{(n)} - e^{-r_{ji}^{(m,n)}})1(p_{ji}^{(n)} < 1 - e^{-r_{ji}^{(m,n)}}) + 1 - e^{-r_{ji}^{(m,n)}} - e^{-r_{ji}^{(m,n)}} r_{ji}^{(m,n)}
\end{aligned}$$

$$
\begin{aligned}
&\le (p_{ji}^{(n)} - r_{ji}^{(m,n)})1(p_{ji}^{(n)} > r_{ji}^{(m,n)}) + (p_{ji}^{(n)} - r_{ji}^{(m,n)})1(1 - e^{-r_{ji}^{(m,n)}} < p_{ji}^{(n)} \le r_{ji}^{(m,n)}) \\
&\quad + (r_{ji}^{(m,n)} - p_{ji}^{(n)})1(p_{ji}^{(n)} < 1 - e^{-r_{ji}^{(m,n)}}) + (r_{ji}^{(m,n)})^2 \\
&= |p_{ji}^{(n)} - r_{ji}^{(m,n)}| + (r_{ji}^{(m,n)})^2,
\end{aligned}
$$

where we have used the inequalities $e^{-x} - 1 \le -x + x^2/2$, $1 - e^{-x} \le x$, and $1 - e^{-x} - e^{-x}x \le x^2/2$ for $x \ge 0$. It follows that if we let $q_{ji}^{(m,n)} = \kappa_m(\mathbf{X}_j, \mathbf{X}_i)/n$, then, on the event $\Omega_{m,n}$, where we have $(1-\epsilon)q_{ji}^{(m,n)} \le r_{ji}^{(m,n)} \le (1+\epsilon)q_{ji}^{(m,n)}$, we have

$$
\begin{aligned}
&\mathbb{P}(|Y_{ji} - Z_{ji}| > 0)1(B_{ji}) \\
&\le \left(|p_{ji}^{(n)} - r_{ji}^{(m,n)}| + (r_{ji}^{(m,n)})^2\right) 1(B_{ji}) \\
&\le \left(|p_{ji}^{(n)} - q_{ji}^{(n)}| + q_{ji}^{(n)} - q_{ji}^{(m,n)} + |q_{ji}^{(m,n)} - r_{ji}^{(m,n)}| + (1+\epsilon)^2 (q_{ji}^{(m,n)})^2\right) 1(B_{ji}) \\
&\le \epsilon q_{ji}^{(n)} + q_{ji}^{(n)} - q_{ji}^{(m,n)} + \epsilon q_{ji}^{(m,n)} + (1+\epsilon)^2 \epsilon q_{ji}^{(m,n)} \\
&\le (1+5\epsilon) q_{ji}^{(n)} - q_{ji}^{(m,n)}.
\end{aligned}
$$

On the other hand, note that on the event $\Omega_{m,n}$ we have

$$
\begin{aligned}
\mathbb{P}(|Y_{ji} - Z_{ji}| > 0)1(B_{ji}^c) &\le \mathbb{P}(Y_{ji} + Z_{ji} > 0)1(B_{ji}^c) \\
&\le \min\left\{1, p_{ji}^{(n)} + r_{ji}^{(m,n)}\right\} 1(B_{ji}^c) \\
&\le (1+\epsilon)(p_{ji}^{(n)} + q_{ji}^{(n)})1(B_{ji}^c) =: \mathcal{B}_n(\mathbf{X}_j, \mathbf{X}_i).
\end{aligned}
$$

Hence, on the event $\Omega_{m,n}$, (4.13) is bounded from above by

$$
\begin{aligned}
&\sum_{j \in [n]} \left\{(1+5\epsilon) q_{j,T_r^-}^{(n)} - q_{j,T_r^-}^{(m,n)}\right\} + \sum_{j \in [n], j \notin \{T_r^-\} \cup A_{r-1}^- \cup I_{r-1}^-} \mathcal{B}_n(\mathbf{X}_j, \mathbf{X}_{T_r^-}) \\
&\le (1+5\epsilon)\lambda_n^-(\mathbf{X}_{T_r^-}) - \lambda_{m,n}^-(\mathbf{X}_{T_r^-}) + \sum_{j \in [n], j \notin \{T_r^-\} \cup A_{r-1}^- \cup I_{r-1}^-} \mathcal{B}_n(\mathbf{X}_j, \mathbf{X}_{T_r^-}),
\end{aligned}
$$

where we have used the observation that $\sum_{j=1}^n q_{ji}^{(n)} = \lambda_n^-(\mathbf{X}_i)$ and $\sum_{j=1}^n q_{ji}^{(m,n)} = \int_{\mathcal{S}} \kappa_m(\mathbf{y}, \mathbf{X}_i)\mu_n(d\mathbf{y}) = \lambda_{m,n}^-(\mathbf{X}_i)$.

To analyze (4.14), note that on the event $\Omega_{m,n}$,

$$
\mathbb{P}(Z_{ji} \ge 1) = 1 - e^{-r_{ji}^{(m,n)}} \le r_{ji}^{(m,n)} \le (1+\epsilon) q_{ji}^{(m,n)} 1(B_{ji}) + (1+\epsilon) q_{ji}^{(m,n)} 1(B_{ji}^c) \le 2\epsilon + \mathcal{B}_n(\mathbf{X}_j, \mathbf{X}_i).
$$

We have thus obtained that, on the event $\Omega_{m,n}$,

$$
\begin{aligned}
\mathbb{P}\left((C_{T_r^-}^-)^c \middle| \mathcal{G}_{r-1}^-\right) &\le (1+5\epsilon)\lambda_n^-(\mathbf{X}_{T_r^-}) - \lambda_{m,n}^-(\mathbf{X}_{T_r^-}) + \sum_{j \in [n], j \notin \{T_r^-\} \cup A_{r-1}^- \cup I_{r-1}^-} \mathcal{B}_n(\mathbf{X}_j, \mathbf{X}_{T_r^-}) \\
&\quad + \sum_{j \in \{T_r^-\} \cup A_{r-1}^- \cup I_{r-1}^-} \left(2\epsilon + \mathcal{B}_n(\mathbf{X}_j, \mathbf{X}_{T_r^-})\right)
\end{aligned}
$$

$$\leq (1+5\epsilon)\lambda_n^-(\mathbf{X}_{T_r^-}) - \lambda_{m,n}^-(\mathbf{X}_{T_r^-}) + 2\epsilon\left|\{T_r^-\}\cup A_{r-1}^-\cup I_{r-1}^-\right| + \sum_{j\in[n]}\mathcal{B}_n(\mathbf{X}_j,\mathbf{X}_{T_r^-}).$$

The same arguments yield that, on the event $\Omega_{m,n}$, (4.15) is bounded by

$$(1+5\epsilon)\lambda_n^+(\mathbf{X}_{T_r^+}) - \lambda_{m,n}^+(\mathbf{X}_{T_r^+}) + \sum_{j\in[n], j\notin\{T_r^+\}\cup I^-_{\sigma_k^-}\cup A^-_{\sigma_k^-}\cup A^+_{r-1}\cup I^+_{r-1}} \mathcal{B}_n(\mathbf{X}_{T_r^+},\mathbf{X}_j),$$

and (4.16) is bounded by

$$\sum_{j\in\{T_r^+\}\cup I^-_{\sigma_k^-}\cup A^-_{\sigma_k^-}\cup A^+_{r-1}\cup I^+_{r-1}} \left(2\epsilon + \mathcal{B}_n(\mathbf{X}_{T_r^+},\mathbf{X}_j)\right).$$

Hence, on the event $\Omega_{m,n}$,

$$\begin{aligned}\mathbb{P}\left((C^+_{T_r^+})^c\middle|\,\mathcal{G}^+_{r-1}\right) &\leq (1+5\epsilon)\lambda_n^+(\mathbf{X}_{T_r^+}) - \lambda_{m,n}^+(\mathbf{X}_{T_r^+}) + 2\epsilon\left|\{T_r^+\}\cup I^-_{\sigma_k^-}\cup A^-_{\sigma_k^-}\cup A^+_{r-1}\cup I^+_{r-1}\right| \\ &\quad + \sum_{j\in[n]}\mathcal{B}_n(\mathbf{X}_{T_r^+},\mathbf{X}_j).\end{aligned}$$

To simplify the notation, define the functions:

$$g^-_{m,n,\epsilon}(\mathbf{X}_l) = \min\left\{1,\ (1+5\epsilon)\lambda_n^-(\mathbf{X}_l) - \lambda_{m,n}^-(\mathbf{X}_l) + \sum_{j\in[n]}\mathcal{B}_n(\mathbf{X}_j,\mathbf{X}_l)\right\} \quad \text{and}$$
$$g^-_{m,n,\epsilon}(\mathbf{X}_l) = \min\left\{1,\ (1+5\epsilon)\lambda_n^+(\mathbf{X}_l) - \lambda_{m,n}^+(\mathbf{X}_l) + \sum_{j\in[n]}\mathcal{B}_n(\mathbf{X}_l,\mathbf{X}_j)\right\},$$

and note that by using the inequality $\min\{1, x+y\} \leq x + \min\{1,y\}$, we obtain

$$\begin{aligned}&\mathbb{P}\left((C^-_{T_r^-})^c\middle|\,\mathcal{G}^-_{r-1}\right) \leq g^-_{m,n,\epsilon}(\mathbf{X}_{T_r^-}) + 2\epsilon\left|\{T_r^-\}\cup A^-_{r-1}\cup I^-_{r-1}\right| \qquad \text{and} \\ &\mathbb{P}\left((C^+_{T_r^+})^c\middle|\,\mathcal{G}^+_{r-1}\right) \leq g^+_{m,n,\epsilon}(\mathbf{X}_{T_r^+}) + 2\epsilon\left|\{T_r^+\}\cup I^-_{\sigma_k^-}\cup A^-_{\sigma_k^-}\cup A^+_{r-1}\cup I^+_{r-1}\right|.\end{aligned}$$

It follows that on the event $\Omega_{m,n}$ we have

$$\begin{aligned}&\mathbb{P}_i\left(\tau^-\leq\sigma_k^-\right) + \mathbb{P}_i\left(\{\tau^->\sigma_k^-\}\cap\{\tau^+\leq\sigma_k^+\}\right) \\ &\leq \sum_{r=1}^k \mathbb{E}_i\left[1(\tau^->r-1, A^-_{r-1}\neq\varnothing, E_{r-1})\left(g^-_{m,n,\epsilon}(\mathbf{X}_{T_r^-}) + 2\epsilon\left|\{T_r^-\}\cup A^-_{r-1}\cup I^-_{r-1}\right|\right)\right] \\ &\quad + \sum_{r=1}^k \mathbb{E}_i\left[1(\tau^->\sigma_k^-,\, \tau^+>r-1, A^+_{r-1}\neq\varnothing, E_{r-1})\right. \\ &\qquad\qquad \left.\cdot\left(g^+_{m,n,\epsilon}(\mathbf{X}_{T_r^+}) + 2\epsilon\left|\{T_r^+\}\cup I^-_{\sigma_k^-}\cup A^-_{\sigma_k^-}\cup A^+_{r-1}\cup I^+_{r-1}\right|\right)\right]\end{aligned}$$

$$
\begin{aligned}
&\le \sum_{r=1}^{k} \mathbb{E}_i \left[1(\hat{A}^-_{r-1} \neq \varnothing) \left(g^-_{m,n,\epsilon}(\mathbf{X}_{\hat{T}^-_r}) + 2\epsilon k \right) \right] \\
&\quad + \sum_{r=1}^{k} \mathbb{E}_i \left[1(\hat{A}^+_{r-1} \neq \varnothing) \left(g^+_{m,n,\epsilon}(\mathbf{X}_{\hat{T}^+_r}) + 2\epsilon \left(k + \left| \hat{I}^-_{\sigma^-_k} \cup \left\{ T_{\mathbf{i}} : \mathbf{i} \in \hat{A}^-_{\hat{\sigma}^-_k} \right\} \right| \right) \right) \right] \\
&\le 4\epsilon k^2 + 2\epsilon k \mathbb{E}_i \left[\left| \hat{I}^-_{\hat{\sigma}^-_k} \right| + \left| \hat{A}^-_{\hat{\sigma}^-_k} \right| \right] \\
&\quad + \sum_{r=1}^{k} \mathbb{E}_i \left[1(\hat{A}^-_{r-1} \neq \varnothing) g^-_{m,n,\epsilon}(\mathbf{X}_{\hat{T}^-_r}) \right] + \sum_{r=1}^{k} \mathbb{E}_i \left[1(\hat{A}^+_{r-1} \neq \varnothing) g^+_{m,n,\epsilon}(\mathbf{X}_{\hat{T}^+_r}) \right],
\end{aligned}
$$

where $\hat{T}^-_r$ and $\hat{T}^+_r$ are the *identities* of the rth "active" nodes to be explored in the inbound and outbound multi-type branching processes, respectively, and $\hat{\sigma}^\pm_k = \inf\{t \ge 1 : |\hat{A}^\pm_t| + |\hat{I}^\pm_t| \ge k \text{ or } \hat{A}^\pm_t = \varnothing\}$.

Next, use Lemma 4.9 to obtain that for $r \ge 1$,

$$
\mathbb{E}_i \left[1(\hat{A}^-_{r-1} \neq \varnothing) g^-_{m,n,\epsilon}(\mathbf{X}_{\hat{T}^-_r}) \right] \le \sum_{s=0}^{r-1} \binom{r-1}{s} 2^{r-1-s} (\Gamma^{(m,n)}_-)^s g^-_{m,n,\epsilon}(\mathbf{X}_i)
$$

and

$$
\mathbb{E}_i \left[1(\hat{A}^-_{r-1} \neq \varnothing) g^+_{m,n,\epsilon}(\mathbf{X}_{\hat{T}^+_r}) \right] \le \sum_{s=0}^{r-1} \binom{r-1}{s} 2^{r-1-s} (\Gamma^{(m,n)}_+)^s g^+_{m,n,\epsilon}(\mathbf{X}_i),
$$

where $\Gamma^{(m,n)}_-$ and $\Gamma^{(m,n)}_+$ are the linear integral operators defined in Lemma 4.9.

Averaging over all $1 \le i \le n$ and using Lemma 4.7 to bound $n^{-1} \sum_{i=1}^n \mathbb{E}_i \left[\left| \hat{I}^-_{\hat{\sigma}^-_k} \right| + \left| \hat{A}^-_{\hat{\sigma}^-_k} \right| \right]$, we obtain

$$
\begin{aligned}
&\frac{1}{n} \sum_{i=1}^{n} \left(\mathbb{P}_i \left(\tau^- \le \sigma^-_k \right) + \mathbb{P}_i \left(\{\tau^- > \sigma^-_k\} \cap \{\tau^+ \le \sigma^+_k\} \right) \right) 1(\Omega_{m,n}) \\
&\le 4\epsilon k^2 + 2\epsilon k^2 \left(1 + \sup_{\mathbf{x} \in \mathcal{S}} \lambda^{(m)}_-(\mathbf{x}) \right) \\
&\quad + 1(\Omega_{m,n}) \sum_{r=1}^{k} \sum_{s=0}^{r-1} \binom{r-1}{s} 2^{r-1-s} \left\{ \int_{\mathcal{S}} (\Gamma^{(m,n)}_-)^s g^-_{m,n,\epsilon}(\mathbf{x}) \mu_n(d\mathbf{x}) + \int_{\mathcal{S}} (\Gamma^{(m,n)}_+)^s g^+_{m,n,\epsilon}(\mathbf{x}) \mu_n(d\mathbf{x}) \right\} \\
&= H(n, m, k, \epsilon) - 1(\Omega^c_{m,n}).
\end{aligned}
$$

The upper bound for the limit of $H(n, m, k, \epsilon)$ as $n \to \infty$ is given in Lemma 4.10. This completes the proof. ■

As a last proof in this section, we use Theorem 4.6 to prove Theorem 3.4, the result establishing the limiting distribution of the degrees in $G_n(\kappa(1 + \varphi_n))$. The latter can also be proven directly using similar arguments as those used in the proof of Theorem 4.6, but we choose to do it this way to avoid repetition.

Proof of Theorem 3.4. Let

$$D^-_{n,i} = \sum_{j\neq i} Y_{ji} \qquad \text{and} \qquad D^+_{n,i} = \sum_{j\neq i} Y_{ij}$$

and define

$$Z^-_{n,i} = \sum_{j=1}^n Z_{ji} \qquad \text{and} \qquad Z^+_{n,i} = \sum_{j=1}^n \tilde{Z}_{ij},$$

where Z_{ji} is Poisson with mean $r_{ji}^{(m,n)}$ and $\tilde{Z}_{ij}$ is Poisson with mean $\tilde{r}_{ij}^{(m,n)}$. Then,

$$\left(D^-_{n,\xi}, D^+_{n,\xi}\right) = \left(D^-_{n,\xi} - Z^-_{n,\xi},\, D^+_{n,\xi} - Z^+_{n,\xi}\right) + \left(Z^-_{n,\xi},\, Z^+_{n,\xi}\right),$$

where since $\sum_{j=1}^n r_{ji}^{(m,n)} = \lambda_-^{(m)}(\mathbf{X}_i)$ and $\sum_{j=1}^n \tilde{r}_{ij}^{(m,n)} = \lambda_+^{(m)}(\mathbf{X}_i)$, we obtain that

$$\begin{aligned}\mathbb{P}\left(Z^-_{n,\xi} = k, Z^+_{n,\xi} = l\right) &= \frac{1}{n}\sum_{i=1}^n \frac{e^{-\lambda_-^{(m)}(\mathbf{X}_i)}(\lambda_-^{(m)}(\mathbf{X}_i))^k}{k!} \cdot \frac{e^{-\lambda_+^{(m)}(\mathbf{X}_i)}(\lambda_+^{(m)}(\mathbf{X}_i))^l}{l!} \\ &\xrightarrow{P} \int_{\mathcal{S}} \frac{e^{-\lambda_-^{(m)}(\mathbf{x})}(\lambda_-^{(m)}(\mathbf{x}))^k}{k!} \cdot \frac{e^{-\lambda_+^{(m)}(\mathbf{x})}(\lambda_+^{(m)}(\mathbf{x}))^l}{l!}\, \mu(d\mathbf{x})\end{aligned}$$

for any $k, l \geq 0$, as $n \to \infty$ (by the bounded convergence theorem). Moreover, by Theorem 4.6,

$$\mathbb{P}\left(|D^-_{n,\xi} - Z^-_{n,\xi}| + |D^+_{n,\xi} - Z^+_{n,\xi}| > 0\right) = \frac{1}{n}\sum_{i=1}^n \mathbb{P}_i\left(\{\tau^- \leq \sigma_1^-\} \cup \{\tau^+ \leq \sigma_1^+\}\right) \leq H(n, m, 1, \epsilon)$$

for any $0 < \epsilon < 1/2$. Therefore, for $(Z^-_{(m)}, Z^+_{(m)})$ constructed on the same probability space as $\left(Z^-_{n,\xi}, Z^+_{n,\xi}\right)$, with $Z^-_{(m)}$ and $Z^+_{(m)}$ conditionally independent (given $\mathbf{X}$) Poisson random variables with parameters $\lambda_-^{(m)}(\mathbf{X})$ and $\lambda_+^{(m)}(\mathbf{X})$, and $\mathbf{X}$ distributed according to μ, we obtain that

$$\limsup_{n\to\infty} P\left(|D^-_{n,\xi} - Z^-_{(m)}| + |D^+_{n,\xi} - Z^+_{(m)}| > 0\right) \leq \limsup_{n\to\infty} E\left[H(n, m, 1, \epsilon) \wedge 1\right] = \hat{H}(m, 1, \epsilon),$$

where $\lim_{m\nearrow\infty}\lim_{\epsilon\downarrow 0} \hat{H}(m, 1, \epsilon) = 0$ by Lemma 4.10. Taking the limit as $\epsilon \downarrow 0$ followed by $m \nearrow \infty$ and noting that $(Z^-_{(m)}, Z^+_{(m)}) \to (Z^-, Z^+)$ a.s., where (Z^-, Z^+) are conditionally independent (given $\mathbf{X}$) Poisson random variables with parameters $\lambda_-(\mathbf{X})$ and $\lambda_+(\mathbf{X})$, gives the weak convergence statement of the theorem.

To obtain the convergence of the expectations note that

$$E[D^-_{n,\xi}] = E[D^+_{n,\xi}] = \frac{1}{n}E\left[\sum_{i=1}^n\sum_{j=1}^n p_{ji}^{(n)}\right] \to \iint_{\mathcal{S}^2} \kappa(\mathbf{x}, \mathbf{y})\mu(d\mathbf{x})\mu(d\mathbf{y})$$

as $n \to \infty$ by Assumption 3.1(d). Now note that

$$\iint_{\mathcal{S}^2} \kappa(\mathbf{x}, \mathbf{y})\mu(d\mathbf{x})\mu(d\mathbf{y}) = E[\lambda_-(\mathbf{X})] = E[\lambda_+(\mathbf{X})] = E[Z^-] = E[Z^+].$$

This completes the proof. ■

4.3.2 Size of the Largest Strongly Connected Component

This last section of the paper contains the proof of Theorem 3.10, the phase transition for the existence of a giant strongly connected component. As mentioned earlier, the idea is to use Theorem 4.6 to couple the exploration of the graph $G_n(\kappa(1+\varphi_n))$ starting from a given vertex with a double tree $(\mathcal{T}_\mu^-(\kappa_m), \mathcal{T}_\mu^+(\kappa_m))$ for a kernel κ_m that takes at most a finite number of different values.

Recall from Section 3.3 that $(\mathcal{T}_\mu^-(\kappa;\mathbf{x}), \mathcal{T}_\mu^+(\kappa;\mathbf{x}))$ denotes the double multi-type Galton-Watson process having root of type $\mathbf{x} \in \mathcal{S}$, and whose offspring distributions are given by (3.1). Let $\rho_-^{\geq k}(\kappa;\mathbf{x})$ (respectively, $\rho_+^{\geq k}(\kappa;\mathbf{x})$) be the probability that the total population of $\mathcal{T}_\mu^-(\kappa;\mathbf{x})$ (respectively, $\mathcal{T}_\mu^+(\kappa;\mathbf{x})$) is at least k. Define also $\rho_-(\kappa;\mathbf{x})$ (respectively, $\rho_+(\kappa;\mathbf{x})$) to be its survival probability, i.e., the probability that its total population is infinite. The averaged joint survival probability is defined as

$$\rho(\kappa) = \int_{\mathcal{S}} \rho_+(\kappa;\mathbf{x})\rho_-(\kappa;\mathbf{x})\mu(d\mathbf{x}).$$

Similarly, for any $k \in \mathbb{N}_+$, we define $\rho^{\geq k}(\kappa) = \int_{\mathcal{S}} \rho_+^{\geq k}(\kappa;\mathbf{x})\rho_-^{\geq k}(\kappa;\mathbf{x})\mu(d\mathbf{x})$.

In addition, we will require from here on that the kernel κ_m be regular finitary (see Definition 3.9) and quasi-irreducible (see Definition 3.8). The following lemma is taken from [5] and it provides the existence of a sequence of partitions $\{\mathscr{J}_m\}_{m\geq 1}$ of $\mathcal{S}$ over which we can define a sequence of regular finitary kernels.

Lemma 4.11 (Lemma 7.1 in [5]) *There exists a sequence of partitions $\{\mathscr{J}_m : m \geq 1\}$ of $\mathcal{S}$, with $\mathscr{J}_m = \{\mathcal{J}_1^{(m)}, \dots, \mathcal{J}_{M_m}^{(m)}\}$, such that*

- *i) each $\mathcal{J}_i^{(m)}$ is measurable and $\mu(\partial \mathcal{J}_i^{(m)}) = 0$,*
- *ii) for each m, $\mathscr{J}_{m+1}$ refines $\mathscr{J}_m$, i.e., each $\mathcal{J}_i^{(m)} = \bigcup_{j \in I_i^{(m)}} \mathcal{J}_j^{(m+1)}$ for some index set $I_i^{(m)}$,*
- *iii) for a.e. $\mathbf{x} \in \mathcal{S}$, diam$(\mathcal{J}_{\vartheta(\mathbf{x})}^{(m)}) \to 0$ as $m \to \infty$, where $\vartheta(\mathbf{x}) = j$ if and only if $\mathbf{x} \in \mathcal{J}_j^{(m)}$.*

Before we construct the sequence of quasi-irreducible regular finitary kernels that we need, we define for notational convenience the following relation.

Definition 4.12 *Let $\tilde{\kappa}$ be a kernel on $\mathcal{S} \times \mathcal{S}$ and let $\mathscr{J} = \{\mathcal{J}_1, \dots, \mathcal{J}_M\}$ be a finite partition of $\mathcal{S}$. Then, we say that set $A \subseteq \mathcal{S}$ is* inbound-accessible *(respectively,* outbound-accessible*) from $\mathbf{x} \in \mathcal{S}$ with respect to $(\tilde{\kappa}, \mathscr{J})$, denoted $\mathbf{x} \to A$ (respectively, $\mathbf{x} \leftarrow A$), if there exists $\{u_1, \dots, u_k\} \subseteq \{1, \dots, M\}$ such that:*

- *i) $\tilde{\kappa}(\mathbf{x}, \mathbf{y}) > 0$ for all $\mathbf{y} \in \mathcal{J}_{u_1}$,*
- *ii) $\tilde{\kappa} > 0$ on $\mathcal{J}_{u_i} \times \mathcal{J}_{u_{i+1}}$ (respectively, $\tilde{\kappa} > 0$ on $\mathcal{J}_{u_{i+1}} \times \mathcal{J}_{u_i}$) for all $1 \leq i < k$,*
- *iii) $\mu(\mathcal{J}_{u_i}) > 0$ for all $1 \leq i \leq k$, and*
- *iv) $\mathcal{J}_{u_k} \subseteq A$.*

Remark 4.13 *Note that if we take* $\mathscr{J}_m = \{\mathcal{J}_1^{(m)}, \ldots, \mathcal{J}_{M_m}^{(m)}\}$ *as constructed in Lemma 4.11, and we let* $\tilde{\kappa}_m$ *satisfy* $\tilde{\kappa}_m \leq \tilde{\kappa}_{m+1}$ *a.e., then if* $\mathbf{x} \to A$ $(\mathbf{x} \leftarrow A)$ *with respect to* $(\tilde{\kappa}_{m_0}, \mathscr{J}_{m_0})$ *for some* $m_0 \geq 1$*, then* $\mathbf{x} \to A$ $(\mathbf{x} \leftarrow A)$ *with respect to* $(\tilde{\kappa}_m, \mathscr{J}_m)$ *for any* $m \geq m_0$*, since each* $\mathcal{J}_{u_i}^{(m)}$ *in part (iii) of Definition 4.12 must contain at least one subset* $\mathcal{J}_t^{(m+1)} \subseteq \mathcal{J}_{u_i}^{(m)}$ *with* $\mu(\mathcal{J}_t^{(m+1)}) > 0$.

We now give a result that states that we can always find a sequence of quasi-irreducible regular finitary kernels which converges monotonically to κ and can be used to approximate from below $\kappa(1 + \varphi_n)$. Its proof follows that of Lemma 7.3 in [5], with some variations due to the directed nature of our kernels.

Lemma 4.14 *For any continuous kernel* κ *and any* φ_n *satisfying Assumption 3.1, there exists a sequence* $\{\tilde{\kappa}_m\}_{m\geq 1}$ *of regular finitary kernels on* $\mathcal{S} \times \mathcal{S}$*, measurable with respect to* $\mathscr{F}$*, with the following properties.*

a.) $\tilde{\kappa}_m(\mathbf{x}, \mathbf{y}) \nearrow \kappa(\mathbf{x}, \mathbf{y})$ *in probability as* $m \to \infty$ *for a.e.* $(\mathbf{x}, \mathbf{y}) \in \mathcal{S} \times \mathcal{S}$

b.) $\tilde{\kappa}_m(\mathbf{x}, \mathbf{y}) \leq \inf_{n\geq m} \kappa(\mathbf{x}, \mathbf{y})(1 + \varphi_n(\mathbf{x}, \mathbf{y}))$ *for a.e.* $(\mathbf{x}, \mathbf{y}) \in \mathcal{S} \times \mathcal{S}$.

c.) If κ *is quasi-irreducible, then so is* κ_m *for all large* m.

Proof. We may assume that $\kappa > 0$ on a set of positive measure, as otherwise we may take $\kappa_m \equiv 0$ for every m and there is nothing to prove. We will construct the sequence $\{\kappa_m : m \geq 1\}$ in two stages. First, we construct a sequence $\{\tilde{\kappa}_m : m \geq 1\}$ where each $\tilde{\kappa}_m$ is regular finitary and satisfies conditions (a) and (b); then we use this sequence to obtain $\{\kappa_m : m \geq 1\}$ satisfying (c).

To this end, construct the sequence of partitions $\{\mathscr{J}_m\}_{m\geq 1}$ according to Lemma 4.11 and define

$$\tilde{\kappa}_m(\mathbf{x}, \mathbf{y}) := \inf \left\{ \kappa(\mathbf{x}', \mathbf{y}') \wedge \inf_{n\geq m} \kappa(\mathbf{x}', \mathbf{y}')(1 + \varphi_n(\mathbf{x}', \mathbf{y}')) : \mathbf{x}' \in \mathcal{J}_{\vartheta(\mathbf{x})}^{(m)}, \mathbf{y}' \in \mathcal{J}_{\vartheta(\mathbf{y})}^{(m)} \right\}.$$

Note that the properties of $\{\mathscr{J}_m : m \geq 1\}$, and the assumption on φ_n imply that

$$\tilde{\kappa}_m(\mathbf{x}, \mathbf{y}) \nearrow \kappa(\mathbf{x}, \mathbf{y}) \quad \text{in probability} \quad \text{as } m \to \infty, \quad \text{for a.e. } (\mathbf{x}, \mathbf{y}) \in \mathcal{S} \times \mathcal{S}.$$

Moreover, for $n \geq m$ we have that

$$\tilde{\kappa}_m(\mathbf{x}, \mathbf{y}) \leq \kappa(\mathbf{x}, \mathbf{y})(1 + \varphi_n(\mathbf{x}, \mathbf{y})) \qquad \text{for a.e. } (\mathbf{x}, \mathbf{y}) \in \mathcal{S} \times \mathcal{S}.$$

Hence, $\kappa_m = \tilde{\kappa}_m$ satisfies conditions (a) and (b) in the statement of the lemma.

To prove (c) assume from now on that κ is quasi-irreducible. In fact, without loss of generality we may assume that κ is irreducible, since it suffices to construct κ_m to be quasi-irreducible on the restriction $\mathcal{S}' \times \mathcal{S}'$ where κ is irreducible and then set it to be zero outside of $\mathcal{S}' \times \mathcal{S}'$.

The first step of the proof ensures the existence of a directed cycle $\mathcal{C} \subseteq \mathcal{S}$ for some $m_1 \geq 1$. The second step uses $\mathcal{C}$ to construct a set on which $\tilde{\kappa}_m$ is irreducible. To establish the existence of $\mathcal{C}$, note that if $\tilde{\kappa}_m = 0$ a.e. for all $m \geq 1$, it would imply that $\kappa = 0$ a.e., which would contradict the irreducibility of κ. Therefore, there must exist some $m_0 \geq 1$ and indexes $1 \leq r, s, t \leq M_{m_0}$ such that $\tilde{\kappa}_{m_0} > 0$ on $(\mathcal{J}_t^{(m_0)} \times \mathcal{J}_r^{(m_0)})$ and on $(\mathcal{J}_r^{(m_0)} \times \mathcal{J}_s^{(m_0)})$, with $\mu(\mathcal{J}_t^{(m_0)})\mu(\mathcal{J}_r^{(m_0)})\mu(\mathcal{J}_s^{(m_0)}) > 0$.

Claim: for any set $A \subseteq \mathcal{S}$ for which there exists a set $D \subseteq \mathcal{S}$ such that $\mu(D) > 0$ and $\tilde{\kappa}_m > 0$ on $D \times A$ (respectively, $A \times D$), the sequence of sets $\{B_m(A)\}_{m\geq 1}$ (respectively, $\{\tilde{B}_m(A)\}_{m\geq 1}$) defined according to $B_m(A) = \{\mathbf{x} \in \mathcal{S} : \mathbf{x} \to A \text{ w.r.t. } (\tilde{\kappa}_m, \mathscr{J}_m)\}$ (respectively, $\tilde{B}_m(A) = \{\mathbf{x} \in \mathcal{S} : \mathbf{x} \leftarrow A \text{ w.r.t. } (\tilde{\kappa}_m, \mathscr{J}_m)\}$) satisfy: 1) $B_m(A) \subseteq B_{m+1}(A)$ (respectively, $\tilde{B}_m(A) \subseteq \tilde{B}_{m+1}(A)$), and 2) $\mu\left(\bigcup_{m=1}^\infty B_m(A)\right) = 1$ $\left(\text{respectively, } \mu\left(\bigcup_{m=1}^\infty \tilde{B}_m(A)\right) = 1\right)$.

To prove the claim note that Remark 4.13 implies (1). To see that (2) holds, let $B(A) = \bigcup_{m=1}^\infty B_m(A)$ and note that from the definition of $B(A)$ we have $\kappa = 0$ a.e. on $B(A)^c \times B(A)$, and the irreducibility of κ implies that either $\mu(B(A)^c) = 0$ or $\mu(B(A)) = 0$; since $\mu(B(A)) \geq \mu(D) > 0$, it must be that $\mu(B(A)^c) = 0$, which implies that $\mu(B(A)) = 1$. The symmetric arguments yield the claim for $\{\tilde{B}_m(A)\}$.

Now apply the inbound part of the claim to $A = \mathcal{J}_r^{(m_0)}$ and $D = \mathcal{J}_t^{(m_0)}$ to obtain that there exists $m_1 \geq m_0$ such that $\mu(B_{m_1}(\mathcal{J}_r^{(m_0)}) \cap \mathcal{J}_s^{(m_0)}) > 0$, which in turn implies there exists a set $\mathcal{J}_{s'}^{(m_1)} \subseteq \mathcal{J}_s^{(m_0)}$ such that $\mu(\mathcal{J}_{s'}^{(m_1)}) > 0$ and $\mathbf{x} \to \mathcal{J}_r^{(m_0)}$ for all $\mathbf{x} \in \mathcal{J}_{s'}^{(m_1)}$. In other words, there exist sets $\{\mathcal{J}_{u_0}^{(m_1)}, \dots, \mathcal{J}_{u_k}^{(m_1)}\}$ satisfying $\mu(\mathcal{J}_{u_i}^{(m_1)}) > 0$ for all $0 \leq i \leq k$, $\mathcal{J}_{u_0}^{(m_1)} = \mathcal{J}_{s'}^{(m_1)}$, $\mathcal{J}_{u_k}^{(m_1)} \subseteq \mathcal{J}_r^{(m_0)}$, and $\tilde{\kappa}_{m_1} > 0$ on $\mathcal{J}_{u_i}^{(m_1)} \times \mathcal{J}_{u_{i+1}}^{(m_1)}$ for all $0 \leq i < k$. Since $0 < \tilde{\kappa}_{m_0} \leq \tilde{\kappa}_{m_1}$ on $\mathcal{J}_{u_k}^{(m_1)} \times \mathcal{J}_{u_0}^{(m_1)}$ by construction, we have that the set $\mathcal{C} = \bigcup_{i=0}^k \mathcal{J}_{u_i}^{(m_1)}$ defines a directed cycle.

Next, construct the sequences $\{B_m(\mathcal{C})\}_{m\geq 1}$ and $\{\tilde{B}_m(\mathcal{C})\}_{m\geq 1}$ according to the claim, and define

$$\kappa_m(\mathbf{x}, \mathbf{y}) = \tilde{\kappa}_m(\mathbf{x}, \mathbf{y}) 1(\mathbf{x} \in (B_m(\mathcal{C}) \cap \tilde{B}_m(\mathcal{C})), \mathbf{y} \in (B_m(\mathcal{C}) \cap \tilde{B}_m(\mathcal{C}))).$$

Note that $\kappa_m \nearrow \kappa$ in probability as $m \to \infty$ since $\tilde{\kappa}_m \nearrow \kappa$ in probability and

$$\mu\left(\bigcup_{m=1}^\infty (B_m(\mathcal{C}) \cap \tilde{B}_m(\mathcal{C}))\right) \geq 1 - \mu\left(\bigcap_{m=1}^\infty B_m(\mathcal{C})^c\right) - \mu\left(\bigcap_{m=1}^\infty \tilde{B}_m(\mathcal{C})^c\right) = 1.$$

It remains to show that κ_m restricted to $(B_m(\mathcal{C}) \cap \tilde{B}_m(\mathcal{C})) \times (B_m(\mathcal{C}) \cap \tilde{B}_m(\mathcal{C}))$ is irreducible. To see this, let $A \subseteq (B_m(\mathcal{C}) \cap \tilde{B}_m(\mathcal{C}))$ and suppose $\kappa_m = 0$ on $A \times (A^c \cap B_m(\mathcal{C}) \cap \tilde{B}_m(\mathcal{C}))$. Note that since $\tilde{\kappa}_{m_1} > 0$ on each $\mathcal{J}_{u_i}^{(m_1)} \times \mathcal{J}_{u_{i+1}}^{(m_1)}$, then it must be that either $\mathcal{C} \subseteq A$ or $\mathcal{C} \subseteq A^c$. Suppose that it is the former, and note that for any $\mathbf{x} \in A^c \cap B_m(\mathcal{C}) \cap \tilde{B}_m(\mathcal{C})$ there exist indexes $\{v_1, \dots, v_l\}$ and $\{w_1, \dots, w_j\}$ such that

$$\tilde{\kappa}_{m_1} > 0 \text{ on } \mathcal{J}_{v_i}^{(m_1)} \times \mathcal{J}_{v_{i+1}}^{(m_1)},\ 0 \leq i \leq l,\ \mu(\mathcal{J}_{v_i}^{(m_1)}) > 0,\ 1 \leq i \leq l,\ \mathcal{J}_{v_l}^{(m_1)} \subseteq \mathcal{C},$$

and

$$\tilde{\kappa}_{m_1} > 0 \text{ on } \mathcal{J}_{w_{i+1}}^{(m_1)} \times \mathcal{J}_{w_i}^{(m_1)},\ 0 \leq i \leq j,\ \mu(\mathcal{J}_{w_i}^{(m_1)}) > 0,\ 1 \leq i \leq j,\ \mathcal{J}_{w_j}^{(m_1)} \subseteq \mathcal{C},$$

where $\mathcal{J}_{v_0}^{(m_1)} = \mathcal{J}_{w_0}^{(m_1)} = \mathcal{J}_{\vartheta(\mathbf{x})}^{(m_1)}$. Moreover, $\mu(\mathcal{J}_{\vartheta(\mathbf{x})}^{(m_1)}) > 0$ would imply that $\mathcal{J}_{v_i}^{(m_1)} \subseteq B_m(\mathcal{C}) \cap \tilde{B}_m(\mathcal{C})$ for all $1 \leq i \leq l$ and $\mathcal{J}_{w_h}^{(m_1)} \subseteq B_m(\mathcal{C}) \cap \tilde{B}_m(\mathcal{C})$ for all $1 \leq h \leq j$, since they would all lie on a directed cycle of positive measure, but this contradicts our assumption that $\tilde{\kappa}_{m_1} = 0$ on $A \times A^c \cap B_m(\mathcal{C}) \cap \tilde{B}_m(\mathcal{C})$. Hence, it must be that $\mu(\mathcal{J}_{\vartheta(\mathbf{x})}^{(m_1)}) = 0$ for all $\mathbf{x} \in A^c \cap B_m(\mathcal{C}) \cap \tilde{B}_m(\mathcal{C})$, and therefore, $\mu(A^c \cap B_m(\mathcal{C}) \cap \tilde{B}_m(\mathcal{C})) = 0$. The same argument gives that if $\mathcal{C} \subseteq A^c \cap B_m(\mathcal{C}) \cap \tilde{B}_m(\mathcal{C})$ then $\mu(A) = 0$. We conclude that κ_m restricted to $(B_m(\mathcal{C}) \cap \tilde{B}_m(\mathcal{C})) \times (B_m(\mathcal{C}) \cap \tilde{B}_m(\mathcal{C}))$ is irreducible. This completes the proof. ■

The following lemma establishes the relationships between $\rho(\kappa_m)$, $\rho^{\geq k}(\kappa_m)$, $\rho^{\geq k}(\kappa)$, and $\rho(\kappa)$.

Lemma 4.15 *Let $\{\kappa_m\}_{m\geq 1}$ be a sequence of kernels on $(\mathcal{S},\mu)$ increasing a.e. to κ. Then, the following limits hold:*

a.) $\rho^{\geq k}(\kappa;\mathbf{x}) \searrow \rho(\kappa;\mathbf{x})$ for a.e. $\mathbf{x}$ and $\rho^{\geq k}(\kappa) \searrow \rho(\kappa)$ as $k\to\infty$.

b.) For every $k\geq 1$, $\rho^{\geq k}(\kappa_m;\mathbf{x}) \nearrow \rho^{\geq k}(\kappa;\mathbf{x})$ for a.e. $\mathbf{x}$ and $\rho^{\geq k}(\kappa_m) \nearrow \rho^{\geq k}(\kappa)$ as $m\to\infty$.

c.) $\rho(\kappa_m;\mathbf{x}) \nearrow \rho(\kappa;\mathbf{x})$ for a.e. $\mathbf{x}$ and $\rho(\kappa_m) \nearrow \rho(\kappa)$ as $m\to\infty$.

Proof. By Lemma 9.5 in [5], we have that $\rho_+^{\geq k}(\kappa;\mathbf{x}) \searrow \rho_+(\kappa;\mathbf{x})$ and $\rho_-^{\geq k}(\kappa;\mathbf{x}) \searrow \rho_-(\kappa;\mathbf{x})$ as $k\to\infty$ for a.e. $\mathbf{x}$. Then, by the monotone convergence theorem, we have

$$\begin{aligned}\lim_{k\to\infty}\rho^{\geq k}(\kappa) &= \lim_{k\to\infty}\int_{\mathcal{S}}\rho_+^{\geq k}(\kappa;\mathbf{s})\rho_-^{\geq k}(\kappa;\mathbf{s})\mu(d\mathbf{s})\\ &= \int_{\mathcal{S}}\lim_{k\to\infty}\rho_+^{\geq k}(\kappa;\mathbf{s})\rho_-^{\geq k}(\kappa;\mathbf{s})\mu(d\mathbf{s})\\ &= \int_{\mathcal{S}}\rho_+(\kappa;\mathbf{s})\rho_-(\kappa;\mathbf{s})\mu(d\mathbf{s}) = \rho(\kappa),\end{aligned}$$

which establishes (a).

By Theorem 6.5(i) in [5] we have that for any fixed $k\geq 1$, $\rho_+^{\geq k}(\kappa_m;\mathbf{x}) \nearrow \rho_+^{\geq k}(\kappa;\mathbf{x})$ and $\rho_-^{\geq k}(\kappa_m;\mathbf{x}) \nearrow \rho_-^{\geq k}(\kappa;\mathbf{x})$ as $m\to\infty$ for a.e. $\mathbf{x}$, which together with monotone convergence as above implies (b).

Part (c) follows from part (a) applied to the kernel κ_m, followed by part (b), to obtain that

$$\lim_{m\to\infty}\rho(\kappa_m;\mathbf{x}) = \lim_{m\to\infty}\lim_{k\to\infty}\rho^{\geq k}(\kappa_m;\mathbf{x}) = \lim_{k\to\infty}\lim_{m\to\infty}\rho^{\geq k}(\kappa_m;\mathbf{x}) = \lim_{k\to\infty}\rho^{\geq k}(\kappa;\mathbf{x}) = \rho(\kappa;\mathbf{x})$$

for a.e. $\mathbf{x}$. Then use monotone convergence as above. ■

Recall the definition of the operators T_κ^- and T_κ^+ given in Section 3.3, as well as of their spectral radii $r(T_\kappa^-)$ and $r(T_\kappa^+)$. The strict positivity of $\rho(\kappa)$, which ensures the existence of a giant strongly connected component, is characterized below. As a preliminary result, we establish the phase transition for regular finitary, quasi-irreducible kernels first.

Proposition 4.16 *Suppose that $\tilde{\kappa}$ is a regular finitary, quasi-irreducible, kernel on the type-space $\mathcal{S}$ with respect to measure μ. Then, $r(T_{\tilde{\kappa}}^-) = r(T_{\tilde{\kappa}}^+)$ and we have that $\rho(\tilde{\kappa}) > 0$ if and only if $r(T_{\tilde{\kappa}}^-) > 1$. Moreover, there exist nonnegative, non-zero eigenfunctions f_- and f_+, such that $T_{\tilde{\kappa}}^- f_- = r(T_{\tilde{\kappa}}^-)f_-$ and $T_{\tilde{\kappa}}^+ f_+ = r(T_{\tilde{\kappa}}^+)f_+$, and they are the only (up to multiplicative constants and sets of measure zero) nonnegative, non-zero eigenfunctions of $T_{\tilde{\kappa}}^-$ and $T_{\tilde{\kappa}}^+$, respectively.*

Proof. Since $\tilde{\kappa}$ is quasi-irreducible, there exists $\mathcal{S}^*\subseteq\mathcal{S}$ such that $\tilde{\kappa}$ restricted to $\mathcal{S}^*$ is irreducible and $\mu(\mathcal{S}^*)>0$. Also, since $\tilde{\kappa}$ is regular finitary, there exists a finite partition $\{\mathcal{J}_i : 1\leq i\leq M\}$ such that $\tilde{\kappa}$ is constant on $\mathcal{J}_i\times\mathcal{J}_j$. Next, define

$$\mathcal{S}' = \bigcup_{i=1}^{M}\{\mathcal{J}_i\cap\mathcal{S}^* : \mu(\mathcal{J}_i\cap\mathcal{S}^*)>0\},$$

and define the kernel $\kappa'(\mathbf{x},\mathbf{y}) = \mu(\mathcal{S}')\tilde{\kappa}(\mathbf{x},\mathbf{y})$ for $\mathbf{x},\mathbf{y} \in \mathcal{S}'$. Note that κ' is regular finitary and irreducible on $\mathcal{S}'$ and $\mu(\mathcal{S}') = \mu(\mathcal{S}^*)$. Moreover, if we let $\mu'(A) = \mu(A)/\mu(\mathcal{S}')$ for $A \subseteq \mathcal{S}'$, and let $\{\mathcal{J}_i' : 1 \leq i \leq M'\}$ denote the partition of $\mathcal{S}'$ such that κ' is constant on $\mathcal{J}_i' \times \mathcal{J}_j'$, then $\mu'(\mathcal{J}_i') > 0$ for all $1 \leq i \leq M'$.

Next, consider the double tree $(\mathcal{T}_{\mu'}^{-}(\kappa'), \mathcal{T}_{\mu'}^{+}(\kappa'))$ on the type-space $\mathcal{S}'$ with respect to measure μ'. Note that each of these trees can be thought of as a multi-type branching process with M' types (one associated to each of the $\mathcal{J}_i'$) each having positive probability. We will show that:

a.) the survival probability $\rho(\tilde{\kappa}) = \mu(\mathcal{S}')\rho'(\kappa')$, where

$$\rho'(\kappa') = \int_{\mathcal{S}'} \rho_-'(\kappa';\mathbf{x})\rho_+'(\kappa';\mathbf{x})\mu'(d\mathbf{x}),$$

and $\rho_-'(\kappa';\mathbf{x})$, $\rho_+'(\kappa';\mathbf{x})$ are the survival probabilities of the trees $\mathcal{T}_{\mu'}^{-}(\kappa';\mathbf{x})$ and $\mathcal{T}_{\mu'}^{+}(\kappa';\mathbf{x})$, respectively; and

b.) the spectral radii of the operators $T_{\tilde{\kappa}}^{\pm}$ on $\mathcal{S}$ and $T_{\kappa'}^{\pm}$ on $\mathcal{S}'$ are the same.

To prove (a), note that since types $\mathbf{x} \in (\mathcal{S}^*)^c$ are isolated (since $\tilde{\kappa}(\mathbf{x},\mathbf{y}) = 0$ for $\mathbf{x} \in (\mathcal{S}')^c$ or $\mathbf{y} \in (\mathcal{S}')^c$) and $\mathcal{S}^* \cap (\mathcal{S}')^c$ has measure zero, then they do not contribute to the survival probabilities of $\mathcal{T}_{\mu}^{-}(\tilde{\kappa})$ and $\mathcal{T}_{\mu}^{+}(\tilde{\kappa})$, which implies that

$$\rho(\tilde{\kappa}) = \int_{\mathcal{S}} \rho_+(\tilde{\kappa};\mathbf{x})\rho_-(\tilde{\kappa};\mathbf{x})\mu(d\mathbf{x}) = \mu(\mathcal{S}') \int_{\mathcal{S}'} \rho_+(\tilde{\kappa};\mathbf{x})\rho_-(\tilde{\kappa};\mathbf{x})\mu'(d\mathbf{x}).$$

Now note that the trees $\mathcal{T}_{\mu}^{\pm}(\tilde{\kappa})$ and $\mathcal{T}_{\mu'}^{\pm}(\kappa')$ have the same law when their roots belong to $\mathcal{S}'$ since the number of offspring of type $\mathbf{y} \in \mathcal{S}'$ that an individual of type $\mathbf{x} \in \mathcal{S}'$ on the tree $\mathcal{T}_{\mu'}^{-}(\kappa')$ has, is Poisson distributed with mean

$$\int_{\mathcal{S}'} \kappa'(\mathbf{y},\mathbf{x})\mu'(d\mathbf{x}) = \int_{\mathcal{S}'} \mu(\mathcal{S}')\tilde{\kappa}(\mathbf{y},\mathbf{x})\mu(d\mathbf{x})/\mu(\mathcal{S}') = \int_{\mathcal{S}} \tilde{\kappa}(\mathbf{y},\mathbf{x})\mu(d\mathbf{x}),$$

which is equal to the corresponding distribution in $\mathcal{T}_{\mu}(\tilde{\kappa})$. The same argument yields the result for $\mathcal{T}_{\mu}^{+}(\tilde{\kappa})$ and $\mathcal{T}_{\mu'}^{+}(\kappa')$. Hence, we have that $\rho_{\pm}(\tilde{\kappa};\mathbf{x}) = \rho_{\pm}(\kappa';\mathbf{x})$ for $\mathbf{x} \in \mathcal{S}'$, and therefore,

$$\rho(\tilde{\kappa}) = \mu(\mathcal{S}')\rho'(\kappa').$$

To establish (b), note that if $f_{\pm}'$ is the nonnegative eigenfunction associated to $r(T_{\kappa'}^{\pm})$ on $\mathcal{S}'$, then $f_{\pm}(\mathbf{x}) = f_{\pm}'(\mathbf{x})1(\mathbf{x} \in \mathcal{S}')$ satisfies

$$(T_{\tilde{\kappa}}^{-} f_-)(\mathbf{x}) = \int_{\mathcal{S}} \tilde{\kappa}(\mathbf{y},\mathbf{x}) f_-(\mathbf{y})\mu(d\mathbf{y}) = \int_{\mathcal{S}'} \kappa'(\mathbf{y},\mathbf{x}) f_-'(\mathbf{y})\mu'(d\mathbf{y}) = r(T_{\kappa'}^{-}) f_-'(\mathbf{x}) = r(T_{\kappa'}^{-}) f_-(\mathbf{x})$$

for $\mathbf{x} \in \mathcal{S}'$, while for $\mathbf{x} \in (\mathcal{S}')^c$ we have $(T_{\tilde{\kappa}}^{-} f_-)(\mathbf{x}) = 0$ since $\tilde{\kappa}(\mathbf{y},\mathbf{x}) = 0$ for all $\mathbf{y} \in \mathcal{S}$. Therefore, $r(T_{\kappa'}^{-})$ is an eigenvalue of $T_{\tilde{\kappa}}^{-}$, which implies that $r(T_{\kappa'}^{-}) \leq r(T_{\tilde{\kappa}}^{-})$; similarly, $r(T_{\kappa'}^{+})$ is an eigenvalue of $T_{\tilde{\kappa}}^{+}$ and $r(T_{\kappa'}^{+}) \leq r(T_{\tilde{\kappa}}^{+})$. For the opposite inequality, suppose $f_{\pm}$ is a nonnegative eigenvector associated to $r(T_{\tilde{\kappa}}^{\pm})$ and set $f_{\pm}'$ to be its restriction to $\mathcal{S}'$. Then note that for $\mathbf{x} \in \mathcal{S}'$,

$$(T_{\kappa'}^{-} f_-')(\mathbf{x}) = \int_{\mathcal{S}'} \kappa'(\mathbf{y},\mathbf{x}) f_-'(\mathbf{y})\mu'(d\mathbf{y}) = \int_{\mathcal{S}} \tilde{\kappa}(\mathbf{y},\mathbf{x}) f_-(\mathbf{y})\mu(d\mathbf{y}) = r(T_{\tilde{\kappa}}^{-}) f_-(\mathbf{x}) = r(T_{\tilde{\kappa}}^{-}) f_-'(\mathbf{x}),$$

and therefore, $r(T^-_{\tilde{\kappa}})$ is an eigenvalue of $T^-_{\kappa'}$ and therefore $r(T^-_{\tilde{\kappa}}) \leq r(T^-_{\kappa'})$. Similarly, $r(T^+_{\tilde{\kappa}}) \leq r(T^+_{\kappa'})$. We conclude that

$$r(T^{\pm}_{\tilde{\kappa}}) = r(T^{\pm}_{\kappa'}).$$

To see that $r(T^-_{\kappa'}) = r(T^+_{\kappa'})$ we first point out that $\mathcal{T}^-_{\mu'}(\kappa')$ and $\mathcal{T}^+_{\mu'}(\kappa')$ can be thought of as irreducible multi-type Galton-Watson processes with a finite number of types and mean progeny matrices $\mathbf{M}^- = (m^-_{ij})$ and $\mathbf{M}^+ = (m^+_{ij})$, respectively, where $m^-_{ij} = c_{ji}\mu'(\mathcal{J}'_j)$, $m^+_{ij} = c_{ij}\mu'(\mathcal{J}'_j)$, and $\kappa'(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{M'} \sum_{j=1}^{M'} c_{ij} 1(\mathbf{x} \in \mathcal{J}'_i, \mathbf{y} \in \mathcal{J}'_j)$. Moreover, the operators $T^-_{\kappa'}$ and $T^+_{\kappa'}$ satisfy

$$T^{\pm}_{\kappa'} f = \mathbf{M}^{\pm}\mathbf{v} \qquad \text{for } \mathbf{v} = (v_1, \dots, v_{M'})^T \in \mathbb{R}^{M'} \text{ and } f(\mathbf{x}) = v_i 1(\mathbf{x} \in \mathcal{J}'_i),\ \mathbf{x} \in \mathcal{S}'.$$

That $\mathbf{M}^-$ and $\mathbf{M}^+$ have the same spectral radius follows from noting that $\mathbf{M}^- = \mathbf{CD}$ and $\mathbf{M}^+ = \mathbf{C}^T\mathbf{D} = (\mathbf{DC})^T$ for $\mathbf{D} = \operatorname{diag}(\mu'(\mathcal{J}'_1), \dots, \mu'(\mathcal{J}'_{M'}))$ and $\mathbf{C} = (c_{ij})$, which implies that the eigenvalues of $\mathbf{M}^+$ are the complex conjugates of those of $\mathbf{DC}$, which in turn are the same as those of $\mathbf{CD}$.

The if and only if statement for the survival probabilities now follows from Theorem 8 in [2] (see also Theorems 2.1 and 2.2 in Chapter 2 of [24]), which states that

$$\rho'_{\pm}(\kappa'; \mathbf{x}) > 0 \text{ for all } \mathbf{x} \in \mathcal{S}' \qquad \text{if and only if} \qquad r(\mathbf{M}^{\pm}) > 1,$$

where $r(\mathbf{M}^{\pm}) = r(T^{\pm}_{\kappa'})$ is the spectral radius of $\mathbf{M}^{\pm}$.

The existence of the eigenfunctions f_- and f_+ on $\mathcal{S}$ follows from the Perron-Frobenius theorem (see Theorem 1.5 in [32]), which guarantees the existence of strictly positive eigenfunctions f'_- and f'_+ on $\mathcal{S}'$ such that $T^{\pm}_{\kappa'} f'_{\pm} = r(T^{\pm}_{\kappa'}) f'_{\pm}$, by setting $f_{\pm}(\mathbf{x}) = f'_{\pm}(\mathbf{x}) 1(\mathbf{x} \in \mathcal{S}')$. Moreover, f'_- and f'_+ are the only (up to multiplicative constants) nonnegative, non-zero eigenfunctions of the operators $T^-_{\kappa'}$ and $T^+_{\kappa'}$, respectively. To see that the nonnegative eigenfunctions f_- and f_+ are also unique (up to multiplicative constants and sets of measure zero) note that any other nonnegative eigenfunction g_- of $T^-_{\tilde{\kappa}}$ associated to a positive eigenvalue λ would have to satisfy

$$(T^-_{\tilde{\kappa}} g_-)(\mathbf{x}) = \int_{\mathcal{S}} \tilde{\kappa}(\mathbf{y}, \mathbf{x}) g_-(\mathbf{y}) \mu(d\mathbf{y}) = 0 \qquad \text{for } \mathbf{x} \in (\mathcal{S}^*)^c,$$

since $\tilde{\kappa}(\mathbf{x}, \mathbf{y}) = 0$ for $\mathbf{x} \in (\mathcal{S}^*)^c$, and

$$(T^-_{\tilde{\kappa}} g_-)(\mathbf{x}) = \int_{\mathcal{S}'} \kappa'(\mathbf{y}, \mathbf{x}) g_-(\mathbf{y}) \mu'(d\mathbf{y}) = \lambda g_-(\mathbf{x}) \qquad \text{for } \mathbf{x} \in \mathcal{S}'$$

which would imply λ is a positive eigenvalue of $T^-_{\kappa'}$ with a nonnegative, non-zero, eigenfunction. The uniqueness of f'_- then gives that $g_-(\mathbf{x}) = \alpha f'_-(\mathbf{x})$ for $\mathbf{x} \in \mathcal{S}'$ for some constant $\alpha > 0$. Finally, since $\mu(\mathcal{S}^* \cap (\mathcal{S}')^c) = 0$, we conclude that $g_-(\mathbf{x}) = \alpha f_-(\mathbf{x})$ a.e. The same arguments give that any other nonnegative eigenfunction g_+ of $T^+_{\tilde{\kappa}}$ would have to satisfy $g_+(\mathbf{x}) = \beta f_+(\mathbf{x})$ a.e. for some constant $\beta > 0$. This completes the proof. ■

We now use the regular finitary and quasi-irreducible case to establish the result for general irreducible kernels. As pointed out in Remark 3.12, the result does not provide a full if and only if condition for the strict positivity of $\rho(\kappa)$, since when the operators T^-_{κ} and T^+_{κ} are unbounded we cannot guarantee the continuity of the spectral radii of the sequence of operators $T^-_{\kappa_m}$ and $T^+_{\kappa_m}$.

Lemma 4.17 *Suppose that κ is irreducible on the type-space $\mathcal{S}$ with respect to measure μ. Then, if $\rho(\kappa) > 0$ we have $r(T_\kappa^-) > 1$ and $r(T_\kappa^+) > 1$. Moreover, if there exists a regular finitary quasi-irreducible kernel $\tilde{\kappa}$ such that $\tilde{\kappa} \leq \kappa$ a.e. and $r(T_{\tilde{\kappa}}^-) > 1$ (equivalently, $r(T_{\tilde{\kappa}}^+) > 1$), then $\rho(\kappa) > 0$.*

Proof. Suppose first that $\rho(\kappa) > 0$. Now use Lemma 4.14 and Lemma 4.15 to obtain that $\rho(\kappa_m) > 0$ for some quasi-irreducible, regular finitary, kernel κ_m such that $\kappa_m(\mathbf{x}, \mathbf{y}) \leq \kappa(\mathbf{x}, \mathbf{y})$ for all $\mathbf{x}, \mathbf{y} \in \mathcal{S}$. By Proposition 4.16 we have that the spectral radii of the operators $T_{\kappa_m}^-$ and $T_{\kappa_m}^+$ satisfy $r(T_{\kappa_m}^-) = r(T_{\kappa_m}^+) > 1$. By monotonicity of the spectral radius, we conclude that $r(T_\kappa^-) \geq r(T_{\kappa_m}^-) > 1$ and $r(T_\kappa^+) \geq r(T_{\kappa_m}^+) > 1$.

For the converse, note that if $\tilde{\kappa} \leq \kappa$ a.e. and $r(T_{\tilde{\kappa}}^-) > 1$, then by Proposition 4.16 we have that $\rho(\tilde{\kappa}) > 0$. Since $\rho(\tilde{\kappa}) \leq \rho(\kappa)$, the result follows. ■

The last preliminary result before proving Theorem 3.10 provides the key estimates obtained through Theorem 4.6, since it relates the indicator random variables for each vertex i to have in-component and out-component of size at least k with the corresponding probabilities in the double-tree $(\mathcal{T}_\mu^-(\kappa_m; \mathbf{X}_i), \mathcal{T}_\mu^+(\kappa_m; \mathbf{X}_i))$.

Proposition 4.18 *For any $k \geq 1$ and $i \in [n]$, define $\chi_{n,i}^{\geq k}$ to be the indicator function of the event that vertex i has in-component and out-component both of size at least k in the graph $G_n(\kappa(1+\varphi_n))$. Then, for any $0 < \epsilon < 1/2$, we have*

$$\left| \frac{1}{n} \sum_{i=1}^n \mathbb{E}\left[\chi_{n,i}^{\geq k}\right] - \frac{1}{n} \sum_{i=1}^n \rho_-^{\geq k}(\kappa_m; \mathbf{X}_i) \rho_+^{\geq k}(\kappa_m; \mathbf{X}_i) \right| \leq H(n, m, k, \epsilon),$$

$$\frac{1}{n^2} \sum_{i=1}^n \sum_{j \neq i} \mathbb{E}\left[\left(\chi_{n,i}^{\geq k} - \mathbb{E}\left[\chi_{n,i}^{\geq k}\right]\right)\left(\chi_{n,j}^{\geq k} - \mathbb{E}\left[\chi_{n,j}^{\geq k}\right]\right)\right] \leq K(n, m, k) + 3H(n, m, k, \epsilon),$$

where

$$K(n, m, k) := \frac{4(k+1)\log n}{n} \sup_{\mathbf{x}, \mathbf{y} \in \mathcal{S}} \kappa_m(\mathbf{x}, \mathbf{y}) + \frac{k}{\log n} \left(2 + \sup_{\mathbf{x} \in \mathcal{S}} \lambda_-^{(m)}(\mathbf{x}) + \sup_{\mathbf{x} \in \mathcal{S}} \lambda_+^{(m)}(\mathbf{x})\right),$$

and $H(n, m, k, \epsilon)$ is defined in Theorem 4.6.

Proof. To derive the first bound construct a coupling between the graph exploration processes of the in-component and out-component of vertex i and the double tree $(\mathcal{T}_\mu^-(\kappa_m; \mathbf{X}_i), \mathcal{T}_\mu^+(\kappa_m; \mathbf{X}_i))$, as described in Section 4.3.1. Define τ^- and τ^+ to be the steps in the construction when the coupling breaks on the inbound, respectively outbound, sides, and let $\sigma_k^- = \inf\{t \geq 1 : |A_t^-| + |I_t^-| \geq k \text{ or } A_t^- = \varnothing\}$ and $\sigma_k^+ = \inf\{t \geq 1 : |A_t^+| + |I_t^+| \geq k \text{ or } A_t^+ = \varnothing\}$. Note that at time $\sigma_k^- \vee \sigma_k^+$ it is possible to determine whether both the in-component and out-component of vertex i have at least k vertices or not. To simplify the notation, let $\rho^{\geq k}(\kappa_m; \mathbf{x}) = \rho_-^{\geq k}(\kappa_m; \mathbf{x}) \rho_+^{\geq k}(\kappa_m; \mathbf{x})$.

$$\begin{aligned}
\frac{1}{n} \sum_{i=1}^n \mathbb{E}\left[\chi_{n,i}^{\geq k}\right] &= \frac{1}{n} \sum_{i=1}^n \mathbb{P}\left(\chi_{n,i}^{\geq k} = 1\right) \\
&\leq \frac{1}{n} \sum_{i=1}^n \mathbb{P}_i\left(\chi_{n,i}^{\geq k} = 1,\, \tau^- \geq \sigma_k^-, \tau^+ \geq \sigma_k^+\right) + \frac{1}{n} \sum_{i=1}^n \mathbb{P}_i\left(\{\tau^- < \sigma_k^-\} \cup \{\tau^+ < \sigma_k^+\}\right)
\end{aligned}$$

$$\leq \frac{1}{n}\sum_{i=1}^{n} \mathbb{P}\left(\text{both } \mathcal{T}_\mu^-(\kappa_m;\mathbf{X}_i) \text{ and } \mathcal{T}_\mu^+(\kappa_m;\mathbf{X}_i) \text{ have at least } k \text{ nodes}\right) + H(n,m,k,\epsilon)$$
$$= \frac{1}{n}\sum_{i=1}^{n} \rho^{\geq k}(\kappa_m;\mathbf{X}_i) + H(n,m,k,\epsilon),$$

where we used Theorem 4.6 to obtain that $n^{-1}\sum_{i=1}^n \mathbb{P}_i\left(\{\tau^- < \sigma_k^-\} \cup \{\tau^+ < \sigma_k^+\}\right) \leq H(n,m,k,\epsilon)$. The other direction follows because

$$\frac{1}{n}\sum_{i=1}^{n} \mathbb{E}\left[\chi_{n,i}^{\geq k}\right] \geq \frac{1}{n}\sum_{i=1}^{n} \mathbb{P}_i\left(\chi_{n,i}^{\geq k} = 1,\, \tau^- \geq \sigma_k^-, \tau^+ \geq \sigma_k^+\right)$$
$$\geq \frac{1}{n}\sum_{i=1}^{n} \mathbb{P}\left(\text{both } \mathcal{T}_\mu^-(\kappa_m;\mathbf{X}_i) \text{ and } \mathcal{T}_\mu^+(\kappa_m;\mathbf{X}_i) \text{ have at least } k \text{ nodes}\right)$$
$$- \frac{1}{n}\sum_{i=1}^{n} \mathbb{P}_i\left(\{\tau^- < \sigma_k^-\} \cup \{\tau^+ < \sigma_k^+\}\right)$$
$$\geq \frac{1}{n}\sum_{i=1}^{n} \rho^{\geq k}(\kappa_m;\mathbf{X}_i) - H(n,m,k,\epsilon).$$

For the second inequality, first note that

$$\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j\neq i} \mathbb{E}\left[\left(\chi_{n,i}^{\geq k} - \mathbb{E}\left[\chi_{n,i}^{\geq k}\right]\right)\left(\chi_{n,j}^{\geq k} - \mathbb{E}\left[\chi_{n,j}^{\geq k}\right]\right)\right]$$
$$= \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j\neq i} \mathbb{E}\left[\chi_{n,i}^{\geq k}\chi_{n,j}^{\geq k}\right] - \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j\neq i} \mathbb{E}\left[\chi_{n,i}^{\geq k}\right]\mathbb{E}\left[\chi_{n,j}^{\geq k}\right].$$

To estimate $\mathbb{E}\left[\chi_{n,i}^{\geq k}\chi_{n,j}^{\geq k}\right]$ we will assume that we first explore the inbound and outbound neighborhood of vertex i up to the time both its in-component and out-component have at least k vertices or there are no more vertices to explore, i.e., we will explore the in-component of vertex i up to time $\sigma_{k,i}^-$ and its out-component up to time $\sigma_{k,i}^+$. Note that we have added the subscript i, relative to the notation introduced in Section 4.3.1, to emphasize that the exploration starts at vertex i. Next, define $\mathcal{F}_{k,i}$ to be the sigma-algebra generated by the exploration of the in-component and out-component of vertex i, as described in Section 4.3.1, up to Step $\sigma_{k,i}^-$ on the inbound side and up to Step $\sigma_{k,i}^+$ on the outbound side. Define $\mathcal{N}_i^{(k)} = I_{\sigma_{k,i}^+}^+ \cup I_{\sigma_{k,i}^-}^- \cup A_{\sigma_{k,i}^-}^- \cup A_{\sigma_{k,i}^+}^+$ to be the set of vertices discovered during that exploration. Now explore the in-component and out-component of vertex j, as described in Section 4.3.1, up to Step $\sigma_{k,j}^-$ on the inbound side and up to Step $\sigma_{k,j}^+$ on the outbound side; let $\mathcal{N}_j^{(k)}$ be the corresponding set of vertices discovered during the exploration of vertex j.

Define $C_{ij} = \left\{\mathcal{N}_i^{(k)} \cap \mathcal{N}_j^{(k)} = \varnothing\right\}$ and note that,

$$\mathbb{E}\left[\chi_{n,i}^{\geq k}\chi_{n,j}^{\geq k}\right] \leq \mathbb{E}\left[\chi_{n,i}^{\geq k}\chi_{n,j}^{\geq k}1(C_{ij})\right] + \mathbb{E}\left[1(C_{ij}^c)\right] = \mathbb{E}\left[\chi_{n,i}^{\geq k}\mathbb{E}\left[\chi_{n,j}^{\geq k}1(C_{ij})\middle|\mathcal{F}_{k,i}\right]\right] + \mathbb{P}(C_{ij}^c).$$

To analyze the conditional expectation, observe that

$$\mathbb{E}\left[\chi_{n,j}^{\geq k}1(C_{ij})\middle|\mathcal{F}_{k,i}\right] = \mathbb{E}\left[\chi_{n,j}^{\geq k}\middle|\mathcal{F}_{k,i}, C_{ij}\right]\mathbb{P}(C_{ij}|\mathcal{F}_{k,i}),$$

where, due to the independence among the arcs, we have that conditionally on $\mathcal{F}_{k,i}$ and C_{ij}, the random variable $\chi_{n,j}^{\geq k}$ has the same distribution as the indicator function of the event that vertex j has in-component and out-component both of size at least k on the graph $G_n(\kappa_{n,i})$, with

$$\kappa_{n,i}(\mathbf{X}_s, \mathbf{X}_t) = \kappa(\mathbf{X}_s, \mathbf{X}_t)(1 + \varphi_n(\mathbf{X}_s, \mathbf{X}_t))1(s \notin \mathcal{N}_i^{(k)}, t \notin \mathcal{N}_i^{(k)}).$$

Now note that since $\kappa_{n,i} \leq \kappa(1 + \varphi_n)$ for any realization of $\mathcal{N}_i^{(k)} \subseteq [n]$, we have

$$\mathbb{E}\left[\chi_{n,j}^{\geq k}\middle|\mathcal{F}_{k,i}, C_{ij}\right] \leq \mathbb{E}\left[\chi_{n,j}^{\geq k}\right],$$

from where it follows that

$$\frac{1}{n^2}\sum_{i=1}^n\sum_{j\neq i}\mathbb{E}\left[\chi_{n,i}^{\geq k}\chi_{n,j}^{\geq k}\right] \leq \frac{1}{n^2}\sum_{i=1}^n\sum_{j\neq i}\left(\mathbb{E}\left[\chi_{n,i}^{\geq k}\right]\mathbb{E}\left[\chi_{n,j}^{\geq k}\right] + \mathbb{P}(C_{ij}^c)\right),$$

which in turn implies that

$$\frac{1}{n^2}\sum_{i=1}^n\sum_{j\neq i}\mathbb{E}\left[\left(\chi_{n,i}^{\geq k} - \mathbb{E}\left[\chi_{n,i}^{\geq k}\right]\right)\left(\chi_{n,j}^{\geq k} - \mathbb{E}\left[\chi_{n,j}^{\geq k}\right]\right)\right] \leq \frac{1}{n^2}\sum_{i=1}^n\sum_{j\neq i}\mathbb{P}(C_{ij}^c).$$

Similarly to what was done on the graph, define $\hat{\mathcal{N}}_i^{(k)} = \hat{I}^+_{\hat{\sigma}^+_{k,i}} \cup \hat{I}^-_{\hat{\sigma}^-_{k,i}} \cup \left\{T_{\mathbf{i}} : \mathbf{i} \in \hat{A}^-_{\hat{\sigma}^-_{k,i}}\right\} \cup \left\{T_{\mathbf{i}} : \mathbf{i} \in \hat{A}^+_{\hat{\sigma}^+_{k,i}}\right\}$ to be the set of *identities* that appear during the construction of the double tree $(\mathcal{T}_\mu^-(\kappa_m; \mathbf{X}_i), \mathcal{T}_\mu^+(\kappa_m; \mathbf{X}_i))$ up to Step $\hat{\sigma}^-_{k,i}$ on the inbound side, and up to Step $\hat{\sigma}^+_{k,i}$ on the outbound side. Let $\hat{C}_{ij} = \left\{\hat{\mathcal{N}}_i^{(k)} \cap \hat{\mathcal{N}}_j^{(k)} = \varnothing\right\}$. We then have

$$\begin{aligned}
\mathbb{P}(C_{ij}^c) &\leq \mathbb{P}(C_{ij}^c, \tau_i^- > \sigma_{k,i}^-, \tau_i^+ > \sigma_{k,i}^+, \tau_j^- > \sigma_{k,j}^-, \tau_j^+ > \sigma_{k,j}^+)1(\Omega_{m,n}) + 1(\Omega_{m,n}^c) \\
&\quad + \mathbb{P}(\{\tau_i^- \leq \sigma_{k,i}^-\} \cup \{\tau_i^+ \leq \sigma_{k,i}^+\}) + \mathbb{P}(\{\tau_j^- \leq \sigma_{k,j}^-\} \cup \{\tau_j^+ \leq \sigma_{k,j}^+\}) \\
&\leq 1(\Omega_{m,n}^c) + \mathbb{P}(\hat{C}_{ij}^c, |\hat{\mathcal{N}}_i^{(k)}| \leq \log n)1(\Omega_{m,n}) + \mathbb{P}(|\hat{\mathcal{N}}_i^{(k)}| > \log n) \\
&\quad + \mathbb{P}_i(\{\tau^- \leq \sigma_k^-\} \cup \{\tau^+ \leq \sigma_k^+\}) + \mathbb{P}_j(\{\tau^- \leq \sigma_k^-\} \cup \{\tau^+ \leq \sigma_k^+\}),
\end{aligned}$$

where the event $\Omega_{m,n}$ is defined in Theorem 4.6.

To bound the first probability on the right-hand side, define $\hat{\mathcal{F}}_{k,i}$ to be the sigma-algebra generated by the construction of the double tree whose root has *identity* i, up to Step $\hat{\sigma}^-_{k,i}$ on the inbound side and up to Step $\hat{\sigma}^+_{k,i}$ on the outbound side. Now note that

$$\hat{C}_{ij} = \{j \notin \hat{\mathcal{N}}_i^{(k)}\} \cap \left(\bigcap_{r=1}^{\hat{\sigma}^+_{k,j}}\bigcap_{t\in\hat{\mathcal{N}}_i^{(k)}}\{\tilde{Z}_{\hat{T}^+_{r,j},t} = 0\}\right) \cap \left(\bigcap_{r=1}^{\hat{\sigma}^-_{k,j}}\bigcap_{t\in\hat{\mathcal{N}}_i^{(k)}}\{Z_{t,\hat{T}^-_{r,j}} = 0\}\right)$$

where $\hat{T}^-_{r,j}$ and $\hat{T}^+_{r,j}$ are the *identities* of the rth active nodes to have their offspring sampled in the double tree whose root is j. Moreover, if we define $B_s = \bigcap_{t\in\hat{\mathcal{N}}_i^{(k)}}\{Z_{ts}=0\}$ and $\tilde{B}_s = \bigcap_{t\in\hat{\mathcal{N}}_i^{(k)}}\{\tilde{Z}_{st}=0\}$, then

$$\{j\notin\hat{\mathcal{N}}_i^{(k)}\} = B_j\cap\tilde{B}_j \qquad \text{and} \qquad \hat{C}_{ij} = B_j\cap\tilde{B}_j\cap\left(\bigcap_{r=1}^{\hat{\sigma}^+_{k,j}}\tilde{B}_{\hat{T}^+_{r,j}}\right)\cap\left(\bigcap_{r=1}^{\hat{\sigma}^-_{k,j}}B_{\hat{T}^-_{r,j}}\right),$$

and therefore, since $\hat{\sigma}^+_{k,j},\hat{\sigma}^-_{k,j}\le k$, the union bound gives

$$\begin{aligned}
\mathbb{P}(\hat{C}^c_{ij}|\hat{\mathcal{F}}_{k,i}) &\le \mathbb{P}\left(B^c_j\cup\left(\bigcup_{r=1}^{\hat{\sigma}^-_{k,j}}B^c_{\hat{T}^-_{r,j}}\right)\middle|\hat{\mathcal{F}}_{k,i}\right) + \mathbb{P}\left(\tilde{B}^c_j\cup\left(\bigcup_{r=1}^{\hat{\sigma}^+_{k,j}}\tilde{B}^c_{\hat{T}^+_{r,j}}\right)\middle|\hat{\mathcal{F}}_{k,i}\right)\\
&\le \mathbb{E}\left[1(B^c_j)+\sum_{r=1}^{\hat{\sigma}^-_{k,j}}1\left(B_j\cap\bigcap_{s=1}^{r-1}B_{\hat{T}^-_{s,j}}\cap B^c_{\hat{T}^-_{r,j}}\right)\middle|\hat{\mathcal{F}}_{k,i}\right]\\
&\quad+\mathbb{E}\left[1(\tilde{B}^c_j)+\sum_{r=1}^{\hat{\sigma}^+_{k,j}}1\left(\tilde{B}_j\cap\bigcap_{s=1}^{r-1}\tilde{B}_{\hat{T}^+_{s,j}}\cap\tilde{B}^c_{\hat{T}^+_{r,j}}\right)\middle|\hat{\mathcal{F}}_{k,i}\right]\\
&\le \mathbb{E}\left[1(B^c_j)+\sum_{r=1}^{k}1\left(\hat{A}^-_{r-1,j}\ne\varnothing,\ \bigcap_{s=1}^{r}\{\hat{T}^-_{s,j}\notin\hat{\mathcal{N}}_i^{(k)}\}, B^c_{\hat{T}^-_{r,j}}\right)\middle|\hat{\mathcal{F}}_{k,i}\right] &&(4.17)\\
&\quad+\mathbb{E}\left[1(\tilde{B}^c_j)+\sum_{r=1}^{k}1\left(\hat{A}^+_{r-1,j}\ne\varnothing,\ \bigcap_{s=1}^{r}\{\hat{T}^+_{s,j}\notin\hat{\mathcal{N}}_i^{(k)}\}, \tilde{B}^c_{\hat{T}^+_{r,j}}\right)\middle|\hat{\mathcal{F}}_{k,i}\right], &&(4.18)
\end{aligned}$$

where $\hat{A}^-_{r,j}$ and $\hat{A}^+_{r,j}$ are the rth inbound and outbound active sets in the construction of the double tree started at j. Now note that the event $\bigcap_{s=1}^r\{\hat{T}^-_{s,j}\notin\hat{\mathcal{N}}_i^{(k)}\}$ implies that none of the $\{U_{s,\hat{T}^-_{r,j}}:1\le s\le n\}$ have been used in the construction of the double tree started at i, hence

$$\mathbb{P}\left(\hat{A}^-_{r-1,j}\ne\varnothing,\ \bigcap_{s=1}^{r}\{\hat{T}^-_{s,j}\notin\hat{\mathcal{N}}_i^{(k)}\}, B^c_{\hat{T}^-_{r,j}}\middle|\hat{\mathcal{F}}_{k,i}\right)\le\mathbb{E}\left[1(\hat{A}^-_{r-1,j}\ne\varnothing)Q(\hat{\mathcal{N}}_i^{(k)},\hat{T}^-_{r,j})\middle|\hat{\mathcal{N}}_i^{(k)}\right],$$

where for any set $V\subseteq[n]$ and any $s\in[n]$ we define

$$\begin{aligned}
Q(V,s) = \mathbb{P}\left(\bigcup_{t\in V}\{Z_{ts}\ge 1\}\right) &\le \sum_{t\in V}P(Z_{ts}\ge 1) = \sum_{t\in V}(1-e^{-r^{(m,n)}_{ts}})\\
&\le \sum_{t\in V}r^{(m,n)}_{ts}\le\frac{R_n}{n}\sum_{t\in V}\kappa_m(\mathbf{X}_t,\mathbf{X}_s)\le\frac{R_n}{n}|V|\sup_{\mathbf{x},\mathbf{y}\in\mathcal{S}}\kappa_m(\mathbf{x},\mathbf{y}),
\end{aligned}$$

and $R_n = \max_{1\le t\le M_m}1(\mu_n(\mathcal{J}^{(m)}_t)>0)\mu(\mathcal{J}^{(m)}_t)/\mu_n(\mathcal{J}^{(m)}_t)$. Since $\mathbb{P}(B^c_j|\hat{\mathcal{F}}_{k,i})\le Q(\hat{\mathcal{N}}_i^{(k)},j)$, we obtain that (4.17) is bounded from above by

$$Q(\hat{\mathcal{N}}_i^{(k)},j)+\sum_{r=1}^{k}\mathbb{E}\left[1(\hat{A}^-_{r-1,j}\ne\varnothing)Q(\hat{\mathcal{N}}_i^{(k)},\hat{T}^-_{r,j})\middle|\hat{\mathcal{N}}_i^{(k)}\right]\le\frac{R_n(k+1)}{n}|\hat{\mathcal{N}}_i^{(k)}|\sup_{\mathbf{x},\mathbf{y}\in\mathcal{S}}\kappa_m(\mathbf{x},\mathbf{y}).$$

Similarly, (4.18) is bounded from above by

$$\frac{R_n(k+1)}{n}|\hat{\mathcal{N}}_i^{(k)}| \sup_{\mathbf{x},\mathbf{y}\in\mathcal{S}} \kappa_m(\mathbf{x},\mathbf{y}).$$

It follows that

$$\mathbb{P}(\hat{C}_{ij}^c|\hat{\mathcal{F}}_{k,i}) \leq \frac{2R_n(k+1)}{n}|\hat{\mathcal{N}}_i^{(k)}| \sup_{\mathbf{x},\mathbf{y}\in\mathcal{S}} \kappa_m(\mathbf{x},\mathbf{y}),$$

which in turn implies that for any $i, j \in [n]$,

$$\begin{aligned}\mathbb{P}(\hat{C}_{ij}^c, |\hat{\mathcal{N}}_i^{(k)}| < \log n)1(\Omega_{m,n}) &= \mathbb{E}\left[\mathbb{P}(\hat{C}_{ij}^c|\hat{\mathcal{F}}_{k,i})1(|\hat{\mathcal{N}}_i^{(k)}| < \log n)\right] 1(\Omega_{m,n}) \\ &\leq \frac{4(k+1)\log n}{n} \sup_{\mathbf{x},\mathbf{y}\in\mathcal{S}} \kappa_m(\mathbf{x},\mathbf{y}),\end{aligned}$$

and we have used the observation that on $\Omega_{m,n}$ we have $R_n \leq 1+\epsilon \leq 2$.

Using this estimate we obtain that

$$\begin{aligned}\frac{1}{n^2}\sum_{i=1}^n\sum_{j\neq i}\mathbb{P}(C_{ij}^c) &\leq \frac{1}{n^2}\sum_{i=1}^n\sum_{j\neq i}\left\{1(\Omega_{m,n}^c) + \mathbb{P}(\hat{C}_{ij}^c,\, |\hat{\mathcal{N}}_i^{(k)}| \leq \log n)1(\Omega_{m,n}) + \mathbb{P}(|\hat{\mathcal{N}}_i^{(k)}| > \log n)\right\} \\ &\quad + \frac{2}{n}\sum_{i=1}^n \mathbb{P}_i(\{\tau^- \leq \sigma_k^-\}\cup\{\tau^+ \leq \sigma_k^+\}) \\ &\leq 1(\Omega_{m,n}^c) + \frac{4(k+1)\log n}{n}\sup_{\mathbf{x},\mathbf{y}\in\mathcal{S}}\kappa_m(\mathbf{x},\mathbf{y}) + \frac{1}{n}\sum_{i=1}^n\mathbb{P}(|\hat{\mathcal{N}}_i^{(k)}| > \log n) \\ &\quad + \frac{2}{n}\sum_{i=1}^n \mathbb{P}_i(\{\tau^- \leq \sigma_k^-\}\cup\{\tau^+ \leq \sigma_k^+\}).\end{aligned}$$

To complete the proof, apply Theorem 4.6 to obtain

$$1(\Omega_{m,n}^c) + \frac{2}{n}\sum_{i=1}^n \mathbb{P}_i(\{\tau^- \leq \sigma_k^-\}\cup\{\tau^+ \leq \sigma_k^+\}) \leq 3H(n,m,k,\epsilon),$$

and Markov's inequality followed by Lemma 4.7 to get

$$\frac{1}{n}\sum_{i=1}^n\mathbb{P}(|\hat{\mathcal{N}}_i^{(k)}| > \log n) \leq \frac{1}{n\log n}\sum_{i=1}^n\mathbb{E}\left[|\hat{\mathcal{N}}_i^{(k)}|\right] \leq \frac{k}{\log n}\left(2 + \sup_{\mathbf{x}\in\mathcal{S}}\lambda_-^{(m)}(\mathbf{x}) + \sup_{\mathbf{x}\in\mathcal{S}}\lambda_+^{(m)}(\mathbf{x})\right).$$

■

We now use Proposition 4.18 to show that the number of vertices with in-component and out-component both of size at least k converges in probability.

Proposition 4.19 *Let $R^-(v)$ and $R^+(v)$ denote the in-component and out-component of vertex $v \in [n]$, as defined in Theorem 3.11. Let $N_n^{\geq k} = \{v \in [n] : |R^-(v)| \geq k$ and $|R^+(v)| \geq k\}$. Then,*

$$\frac{|N_n^{\geq k}|}{n} \xrightarrow{P} \rho^{\geq k}(\kappa), \qquad n \to \infty.$$

Proof. Define $\{\chi_{n,i}^{\geq k}\}_{i\in[n]}$ as in Proposition 4.18. We start by noting that for any $m \geq 1$ we have

$$\begin{aligned}\left|\frac{|N_n^{\geq k}|}{n} - \rho^{\geq k}(\kappa)\right| \leq & \left|\frac{|N_n^{\geq k}|}{n} - \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[\chi_{n,i}^{\geq k}\right]\right| \\ & + \left|\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[\chi_{n,i}^{\geq k}\right] - \frac{1}{n}\sum_{i=1}^{n}\rho_-^{\geq k}(\kappa_m;\mathbf{X}_i)\rho_+^{\geq k}(\kappa_m;\mathbf{X}_i)\right| \\ & + \left|\frac{1}{n}\sum_{i=1}^{n}\rho_-^{\geq k}(\kappa_m;\mathbf{X}_i)\rho_+^{\geq k}(\kappa_m;\mathbf{X}_i) - \rho^{\geq k}(\kappa_m)\right| + \left|\rho^{\geq k}(\kappa_m) - \rho^{\geq k}(\kappa)\right|.\end{aligned} \tag{4.19}$$

Moreover, by Proposition 4.18 we have that for any $0 < \epsilon < 1/2$, (4.19) is bounded by $H(n, m, k, \epsilon)$, where $H(n, m, k, \epsilon)$ is defined in Theorem 4.6 and satisfies $H(n, m, k, \epsilon) \xrightarrow{P} \hat{H}(m, k, \epsilon)$ for some other function $\hat{H}(m, k, \epsilon)$ (defined in Lemma 4.10) as $n \to \infty$, where for any fixed $k \geq 1$ we have

$$\lim_{m\nearrow\infty}\lim_{\epsilon\downarrow 0}\hat{H}(m, k, \epsilon) = 0.$$

Also, by the bounded convergence theorem we have that for any $m, k \in \mathbb{N}_+$,

$$\frac{1}{n}\sum_{i=1}^{n}\rho_-^{\geq k}(\kappa_m;\mathbf{X}_i)\rho_+^{\geq k}(\kappa_m;\mathbf{X}_i) = \int_{\mathcal{S}}\rho_-^{\geq k}(\kappa_m;\mathbf{x})\rho_+^{\geq k}(\kappa_m;\mathbf{x})\mu_n(d\mathbf{x}) \xrightarrow{P} \rho^{\geq k}(\kappa_m) \qquad n \to \infty,$$

and by Lemma 4.15 we have that

$$\lim_{m\nearrow\infty}\rho^{\geq k}(\kappa_m) = \rho^{\geq k}(\kappa)$$

in probability, since $\kappa_m \nearrow \kappa$ in probability. Therefore, for any $m \geq 1$ and $0 < \epsilon < 1/2$ we have

$$\limsup_{n\to\infty}\left|\frac{|N_n^{\geq k}|}{n} - \rho^{\geq k}(\kappa)\right| \leq \limsup_{n\to\infty}\left|\frac{|N_n^{\geq k}|}{n} - \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[\chi_{n,i}^{\geq k}\right]\right| + \hat{H}(m, k, \epsilon) + \left|\rho^{\geq k}(\kappa_m) - \rho^{\geq k}(\kappa)\right|.$$

and by taking $\epsilon \downarrow 0$ followed by $m \nearrow \infty$ we obtain that the following limit holds in probability

$$\limsup_{n\to\infty}\left|\frac{|N_n^{\geq k}|}{n} - \rho^{\geq k}(\kappa)\right| \leq \limsup_{n\to\infty}\left|\frac{|N_n^{\geq k}|}{n} - \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[\chi_{n,i}^{\geq k}\right]\right|.$$

It remains to show that this last limit is zero. To do this, start by using Proposition 4.18 again to obtain that for any $m \geq 1$ and $0 < \epsilon < 1/2$, we have that on the event $\Omega_{m,n}$,

$$\begin{aligned}&\mathbb{E}\left[\left(\frac{N_n^{\geq k}}{n} - \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[\chi_{n,i}^{\geq k}\right]\right)^2\right] \\ &= \frac{1}{n^2}\left\{\sum_{i=1}^{n}\mathbb{E}\left[\left(\chi_{n,i}^{\geq k} - \mathbb{E}\left[\chi_{n,i}^{\geq k}\right]\right)^2\right] + \sum_{i=1}^{n}\sum_{j\neq i}\mathbb{E}\left[\left(\chi_{n,i}^{\geq k} - \mathbb{E}\left[\chi_{n,i}^{\geq k}\right]\right)\left(\chi_{n,j}^{\geq k} - \mathbb{E}\left[\chi_{n,j}^{\geq k}\right]\right)\right]\right\}\end{aligned}$$

$$\leq \frac{1}{n^2}\sum_{i=1}^{n} \mathbb{E}\left[\left(\chi_{n,i}^{\geq k} - \mathbb{E}\left[\chi_{n,i}^{\geq k}\right]\right)^2\right] + K(n,m,k) + 3H(n,m,k,\epsilon)$$

$$\leq \frac{1}{n} + K(n,m,k) + 3H(n,m,k,\epsilon),$$

where $K(n,m,k)$ is defined in Proposition 4.18 and satisfies $K(n,m,k) \xrightarrow{P} 0$ as $n \to \infty$ for any fixed m, ϵ. The bounded convergence theorem now gives

$$\lim_{n\to\infty} E\left[\left(\frac{N_n^{\geq k}}{n} - \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[\chi_{n,i}^{\geq k}\right]\right)^2\right] = E\left[\lim_{n\to\infty} \mathbb{E}\left[\left(\frac{N_n^{\geq k}}{n} - \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}\left[\chi_{n,i}^{\geq k}\right]\right)^2\right]\right] \leq 3\hat{H}(m,k,\epsilon),$$

and taking the limit as $\epsilon \downarrow 0$ followed by $m \nearrow \infty$ completes the proof. ■

We are now ready to prove Theorem 3.10, the phase transition for the existence of a giant strongly connected component in $G_n(\kappa(1+\varphi_n))$.

Proof of Theorem 3.10. By Lemma 4.14, there exists a sequence of kernels $\{\kappa_m : m \geq 1\}$ defined on $\mathcal{S} \times \mathcal{S}$, measurable with respect to $\mathscr{F}$, such that κ_m is quasi-irreducible, regular finitary, and such that for any $n \geq m$, we have

$$\kappa_m(\mathbf{x},\mathbf{y}) \leq \kappa(\mathbf{x},\mathbf{y})(1+\varphi_n(\mathbf{x},\mathbf{y})) \qquad \text{for all } \mathbf{x},\mathbf{y} \in \mathcal{S}.$$

Proof of the lower bound: We will start by proving a lower bound for the largest strongly connected component of $G_n(\kappa(1+\varphi_n))$. To this end, note that we can construct a coupling between $G_n(\kappa(1+\varphi_n))$ and $G_n(\kappa_m)$ such that every arc in $G_n(\kappa_m)$ is also in $G_n(\kappa(1+\varphi_n))$ P-a.s. It follows that

$$|\mathcal{C}_1(G_n(\kappa(1+\varphi_n)))| \geq |\mathcal{C}_1(G_n(\kappa_m))| \quad P\text{-a.s.}$$

The idea is now to apply Theorem 1 in [3] to $G_n(\kappa_m)$, however, that theorem requires that the kernel κ_m be irreducible, whereas κ_m is only quasi-irreducible. To address this issue, we construct a third graph as follows. Let $\mathcal{S}^*$ be the restriction of $\mathcal{S}$ where κ_m is irreducible and set

$$\mathcal{S}' = \bigcup_{i=1}^{M_m} \left\{ \mathcal{J}_i^{(m)} \cap \mathcal{S}^* : \mu(\mathcal{J}_i^{(m)}) > 0 \right\}.$$

To avoid trivial cases, assume from now on that $\mu(\mathcal{S}') > 0$.

Now let $V_{n'} = \{1 \leq i \leq n : \mathbf{X}_i \in \mathcal{S}'\}$ denote the set of vertices in $G_n(\kappa_m)$ that have types in $\mathcal{S}'$ and let n' denote its cardinality. Note that n' is random, but measurable with respect to $\mathscr{F}$. Next, fix $0 < \delta < 1$ and define the kernel $\kappa'(\mathbf{x},\mathbf{y}) = (1-\delta)\mu(\mathcal{S}')\kappa_m(\mathbf{x},\mathbf{y})$ and the graph $G_{n'}(\kappa')$ whose arc probabilities are given by

$$p_{ij}^{(n')} = \frac{(1-\delta)\mu(\mathcal{S}')\kappa_m(\mathbf{X}_i,\mathbf{X}_j)}{n'} \wedge 1, \qquad i,j \in [n'],\ i \neq j.$$

Note that $G_{n'}(\kappa')$ is a graph on the type space $\mathcal{S}'$ whose types are distributed according to measure $\mu_n'(A) := \mu_n(A)/\mu_n(\mathcal{S}')$ for any $A \subseteq \mathcal{S}'$. Moreover, κ' is irreducible on $\mathcal{S}'$ with each of its induced types, i.e., the sets $\mathcal{J}_i^{(m)} \cap \mathcal{S}'$, having strictly positive measure. Now note that since $n\mu_n(\mathcal{S}') = n'$ and $\mu_n(\mathcal{S}') \xrightarrow{P} \mu(\mathcal{S}')$ as $n \to \infty$, then

$$p_{ij}^{(n')} = \frac{(1-\delta)\mu(\mathcal{S}')\kappa_m(\mathbf{X}_i, \mathbf{X}_j)}{n\mu_n(\mathcal{S}')} \wedge 1 \leq \frac{\kappa_m(\mathbf{X}_i, \mathbf{X}_j)}{n} \wedge 1, \qquad i, j \in [n'],\, i \neq j,$$

for all sufficiently large n. Therefore, there exists a coupling such that every arc in $G_{n'}(\kappa')$ is also in $G_n(\kappa_m)$, and therefore, for all sufficiently large n,

$$|\mathcal{C}_1(G_n(\kappa_m))| \geq |\mathcal{C}_1(G_{n'}(\kappa'))| \quad P\text{-a.s.}$$

Now use Theorem 1 in [3] to obtain that for every $\epsilon > 0$

$$P\left(\left| \frac{|\mathcal{C}_1(G_{n'}(\kappa'))|}{n'} - \rho'(\kappa') \right| > \epsilon \right) \to 0 \qquad n \to \infty,$$

where

$$\rho'(\kappa') = \int_{\mathcal{S}'} \rho'_-(\kappa'; \mathbf{x}) \rho'_+(\kappa'; \mathbf{x}) \mu'(d\mathbf{x}),$$

and $\rho'_-(\kappa'; \mathbf{x}), \rho'_+(\kappa'; \mathbf{x})$ are the survival probabilities of the trees $\mathcal{T}_{\mu'}^-(\kappa')$ and $\mathcal{T}_{\mu'}^+(\kappa')$, respectively, defined on the type space $\mathcal{S}'$ with respect to the measure $\mu'(A) = \mu(A)/\mu(\mathcal{S}')$ for $A \subseteq \mathcal{S}'$.

By the arguments in the proof of Proposition 4.16, we have that $\rho((1-\delta)\kappa_m) = \mu(\mathcal{S}')\rho'(\kappa')$, where

$$\rho((1-\delta)\kappa) = \int_{\mathcal{S}} \rho_-((1-\delta)\kappa_m; \mathbf{x}) \rho_+((1-\delta)\kappa; \mathbf{x}) \mu(d\mathbf{x}),$$

and $\rho_-((1-\delta)\kappa_m; \mathbf{x})$, $\rho_+((1-\delta)\kappa_m; \mathbf{x})$ are the survival probabilities of the trees $\mathcal{T}_\mu^-((1-\delta)\kappa_m)$ and $\mathcal{T}_\mu^+((1-\delta)\kappa_m)$, defined on the type space $\mathcal{S}$.

Hence,

$$\frac{|\mathcal{C}_1(G_n(\kappa(1+\varphi_n)))|}{n} \geq \frac{|\mathcal{C}_1(G_n((1-\delta)\kappa_m))|}{n} \geq \frac{|\mathcal{C}_1(G_{n'}(\kappa'))|}{n'} \cdot \frac{n'}{n} \xrightarrow{P} \rho'(\kappa')\mu(\mathcal{S}') = \rho((1-\delta)\kappa_m),$$

as $n \to \infty$. Now use Lemma 4.15 to obtain that the following limits hold in probability

$$\lim_{m \nearrow \infty} \lim_{\delta \downarrow 0} \rho((1-\delta)\kappa_m) = \lim_{\delta \downarrow 0} \lim_{m \nearrow \infty} \rho((1-\delta)\kappa_m) = \rho(\kappa),$$

from where we conclude that for any $\epsilon > 0$,

$$P\left(\frac{|\mathcal{C}_1(G_n(\kappa(1+\varphi_n)))|}{n} - \rho(\kappa) < -\epsilon \right) \to 0 \qquad n \to \infty.$$

Proof of the upper bound: For any $k, m \geq 1$ let $\rho_-^{\geq k}(\kappa_m; \mathbf{x})$ ($\rho_+^{\geq k}(\kappa_m; \mathbf{x})$) denote the probability that the tree $\mathcal{T}_\mu^-(\kappa_m; \mathbf{x})$ ($\mathcal{T}_\mu^+(\kappa_m; \mathbf{x})$) has a population of at least k nodes. Define for $k \geq 1$ the set $N_n^{\geq k}$ as in Proposition 4.19; $N_n^{\geq k}$ is the set of vertices in $G_n(\kappa(1+\varphi_n))$ with both large in-component and large out-component. Now note that provided $\liminf_{n\to\infty} |\mathcal{C}_1(G_n(\kappa(1+\varphi_n)))| = \infty$, we have that for any fixed $k \geq 1$,

$$|\mathcal{C}_1(G_n(\kappa(1+\varphi_n)))| \leq |N_n^{\geq k}| \qquad \text{for all sufficiently large } n.$$

It follows that

$$\frac{|\mathcal{C}_1(G_n(\kappa(1+\varphi_n)))|}{n} - \rho(\kappa) \leq \frac{|N_n^{\geq k}|}{n} - \rho^{\geq k}(\kappa) + \rho^{\geq k}(\kappa) - \rho(\kappa).$$

Now use Proposition 4.19 to obtain that $|N_n^{\geq k}|/n \xrightarrow{P} \rho^{\geq k}(\kappa)$ as $n \to \infty$ for any fixed $k \geq 1$. Now use Lemma 4.15 to obtain that $\rho^{\geq k}(\kappa) \nearrow \rho(\kappa)$ as $k \nearrow \infty$, which completes the proof of the upper bound.

Proof of the phase transition: It follows from Lemma 4.17. ∎

We now prove Theorem 3.11, which provides a more detailed description of the giant strongly connected component and of the bow-tie structure it induces.

Proof of Theorem 3.11. In view of Theorem 3.10, we know that $\mathcal{C}_1(G_n(\kappa(1+\varphi_n)))$ contains asymptotically $n\rho(\kappa)$ vertices, and therefore, $\mathcal{C}_1(G_n(\kappa(1+\varphi_n))) \subseteq L_n^+ \cap L_n^-$ with high probability. To show the reverse subset relation fix $0 < \delta < 1$ and $m \geq 1$ and construct the kernel $\kappa'(\mathbf{x},\mathbf{y}) = (1-\delta)\mu(\mathcal{S}')\kappa_m(\mathbf{x},\mathbf{y})$ on the type space $\mathcal{S}' \subseteq \mathcal{S}$ just as in the proof of Theorem 3.10, so that κ' is regular finitary and irreducible on $\mathcal{S}'$. Now construct the graph $G_{n'}(\kappa')$ using a coupling ensuring that every arc in $G_{n'}(\kappa')$ is also in $G_n(\kappa(1+\varphi_n))$, as was done in the proof of Theorem 3.10. Note that we must have

$$\mathcal{C}_1(G_{n'}(\kappa')) \subseteq \mathcal{C}_l(G_n(\kappa(1+\varphi_n))) \qquad P\text{-a.s.},$$

for some $l \geq 1$, where $\mathcal{C}_l(G_n(\kappa(1+\varphi_n)))$ is the lth largest strongly connected component of $G_n(\kappa(1+\varphi_n))$. Now define the sets $L^-_{m,\delta,n} = \{v \in [n'] : |R^-(v)| \geq (\log n)/n\}$ and $L^+_{m,\delta,n} = \{v \in [n'] : |R^+(v)| \geq (\log n)/n\}$ relative to graph $G_{n'}(\kappa')$. By Theorem 3.10 we again have that $G_{n'}(\kappa')$ contains a number of vertices of order n', with $n/n' = O(1)$, and therefore, $\mathcal{C}_1(G_{n'}(\kappa')) \subseteq L^-_{m,\delta,n} \cap L^+_{m,\delta,n}$ with high probability. Moreover, a close inspection of the proof of Theorem 1 in [3] shows that $L^-_{m,\delta,n} \cap L^+_{m,\delta,n}$ (denoted $B'(\omega_1)$ in [3]) is strongly connected with high probability (specifically, see the proof of (22) in [3]). It follows that

$$\lim_{n\to\infty} P\left(L^+_{m,\delta,n} \cap L^-_{m,\delta,n} = \mathcal{C}_1(G_{n'}(\kappa'))\right) = 1.$$

Now note that $L^-_{m,\delta,n} \nearrow L_n^-$ and $L^+_{m,\delta,n} \nearrow L_n^+$ as $m \to \infty$ and $\delta \to 0$, which implies that, with high probability, $L_n^+ \cap L_n^- \subseteq \mathcal{C}_l(G_n(\kappa(1+\varphi_n)))$ for some $l \geq 1$, and therefore, $L_n^+ \cap L_n^-$ is strongly connected. The first sentence in the proof now implies that $l = 1$, and therefore,

$$\lim_{n\to\infty} P\left(\mathcal{C}_1(G_n(\kappa(1+\varphi_n))) = L_n^+ \cap L_n^-\right) = 1.$$

To establish the limits for $|L_n^+|$ and $|L_n^-|$, let $L^{\geq k}_{-,n} = \{v \in [n] : |R^-(v)| \geq k\}$ and $L^{\geq k}_{+,n} = \{v \in [n] : |R^+(v)| \geq k\}$ and note that $|L_n^{\pm}| \leq |L^{\geq k}_{\pm,n}|$ for any $1 \leq k \leq (\log n)/n$ and $|L^{\geq k}_{\pm,n}| \leq |L_n^{\pm}|$ for any

$k \geq (\log n)/n$. A straightforward adaptation of Proposition 4.19 can be used to obtain

$$\frac{|L_{-,n}^{\geq k}|}{n} \xrightarrow{P} \int_{\mathcal{S}} \rho_{-}^{\geq k}(\kappa; \mathbf{x})\mu(d\mathbf{x}) \qquad \text{and} \qquad \frac{|L_{+,n}^{\geq k}|}{n} \xrightarrow{P} \int_{\mathcal{S}} \rho_{+}^{\geq k}(\kappa; \mathbf{x})\mu(d\mathbf{x})$$

as $n \to \infty$. Monotone convergence gives $\int_{\mathcal{S}} \rho_{\pm}^{\geq k}(\kappa; \mathbf{x})\mu(d\mathbf{x}) \searrow \int_{\mathcal{S}} \rho_{\pm}(\kappa; \mathbf{x})\mu(d\mathbf{x})$ as $k \nearrow \infty$, which yields the result. ■

We end the paper with the proof of Proposition 3.13, which states the main results for the rank-1 kernel case.

Proof of Proposition 3.13. The first two statements follow immediately from noting that $E[\kappa_+(\mathbf{X})]E[\kappa_-(\mathbf{X})] = \iint_{\mathcal{S}^2} \kappa(\mathbf{x}, \mathbf{y})\mu(d\mathbf{x})\mu(d\mathbf{y})$. The third one follows from noting that $\lambda_-(\mathbf{X}) = \kappa_-(\mathbf{X})E[\kappa_+(\mathbf{X})]$ and $\lambda_+(\mathbf{X}) = \kappa_+(\mathbf{X})E[\kappa_-(\mathbf{X})]$.

To establish (d) assume first that $\rho(\kappa) > 0$. Now use Lemma 4.14 (applied to $\kappa(\mathbf{x}, \mathbf{y}) = \kappa_-(\mathbf{y})$ and $\kappa(\mathbf{x}, \mathbf{y}) = \kappa_+(\mathbf{x})$ separately) to obtain that there exists a sequence of kernels $\{\kappa_m^+(\mathbf{x}) : m \geq 1\}$ and $\{\kappa_m^-(\mathbf{x}) : m \geq 1\}$ such that: 1) $0 \leq \kappa_m^{\pm}(\mathbf{x}) \leq \kappa_{\pm}(\mathbf{x})$ for all $\mathbf{x} \in \mathcal{S}$, 2) each is piecewise constant taking only a finite number of values, and 3) $\kappa_m^{\pm}(\mathbf{x}) \nearrow \kappa_{\pm}(\mathbf{x})$ in probability for a.e. $\mathbf{x} \in \mathcal{S}$ as $m \to \infty$. Now set $B_m = \{\mathbf{x} \in \mathcal{S} : \kappa_m^-(\mathbf{x}) > 0, \kappa_m^+(\mathbf{x}) > 0\}$ and define

$$\kappa_m(\mathbf{x}, \mathbf{y}) = \kappa_m^+(\mathbf{x})\kappa_m^-(\mathbf{y})1(\mathbf{x} \in B_m, \mathbf{y} \in B_m).$$

Note that κ_m is regular finitary and is strictly positive on $B_m \times B_m$. Hence, the only set $A \subseteq B_m$ satisfying $\kappa_m = 0$ on $A \times (A^c \cap B_m)$ is $A = \varnothing$ or $A^c \cap B_m = \varnothing$, implying the irreducibility of κ_m on $B_m \times B_m$. Moreover, since $\kappa_- > 0$ and $\kappa_+ > 0$ a.e. in order for κ to be irreducible, we have that $\kappa_m \nearrow \kappa$ in probability as $m \to \infty$.

Next, use Lemma 4.15 to obtain that $\rho(\kappa) = \lim_{m\to\infty} \rho(\kappa_m)$, and therefore, $\rho(\kappa_m) > 0$ for some m sufficiently large. By Proposition 4.16 this implies that the spectral radii of the operators $T_{\kappa_m}^-$ and $T_{\kappa_m}^+$ are strictly larger than one. Now note that the functions $f_m^-(\mathbf{x}) = \kappa_m^-(\mathbf{x})$ and $f_m^+(\mathbf{x}) = \kappa_m^+(\mathbf{x})$ are nonnegative and satisfy

$$\begin{aligned} T_{\kappa_m}^- f_m^-(\mathbf{x}) &= \int_{\mathcal{S}} \kappa_m^+(\mathbf{y})\kappa_m^-(\mathbf{x}) f_m^-(\mathbf{y})\mu(d\mathbf{x}) = \kappa_m^-(\mathbf{x}) \int_{\mathcal{S}} \kappa_m^+(\mathbf{y}) f_m^-(\mathbf{y})\mu(d\mathbf{y}) \\ &= f_m^-(\mathbf{x}) \int_{\mathcal{S}} \kappa_m^+(\mathbf{y})\kappa_m^-(\mathbf{y})\mu(d\mathbf{y}), \end{aligned}$$

and therefore, $r_m := \int_{\mathcal{S}} \kappa_m^+(\mathbf{y})\kappa_m^-(\mathbf{y})\mu(d\mathbf{y})$ is an eigenvalue of $T_{\kappa_m}^-$. Similarly, r_m is an eigenvalue of $T_{\kappa_m}^+$ associated to the nonnegative eigenfunction f_m^+. Since we may assume that $\kappa_m^-(\mathbf{x})$ and $\kappa_m^+(\mathbf{x})$ are different from zero for sufficiently large m, then Proposition 4.16 gives that $r_m = r(T_{\kappa_m}^{\pm}) > 1$. Taking the limit as $m \to \infty$ gives that

$$E[\kappa_+(\mathbf{X})\kappa_-(\mathbf{X})] = \lim_{m\to\infty} r_m > 1.$$

For the converse, note that $E[\kappa_+(\mathbf{X})\kappa_-(\mathbf{X})] > 1$ and the monotone convergence theorem imply that $r_m > 1$ for some m sufficiently large. For this m, Proposition 4.16 gives that r_m is the spectral radius of $T_{\kappa_m}^-$ and $T_{\kappa_m}^+$, and also that $\rho(\kappa_m) > 0$. Lemma 4.15 now gives that $1 < \rho(\kappa_m) \nearrow \rho(\kappa)$ in probabiity as $m \to \infty$. ■

References

[1] Alfred D Andrew and William L Green. Spectral theory of operators on Hilbert space. *School of Mathematics Georgia Institute of Technology Atlanta, GA*, pages 30332–0160, 2002.

[2] Krishna B Athreya and Peter E Ney. *Branching Processes.* Dover, 1972.

[3] Mindaugas Bloznelis, Friedrich Götze, and Jerzy Jaworski. Birth of a strongly connected giant in an inhomogeneous random digraph. *Journal of Applied Probability*, 49(03):601–611, 2012.

[4] Béla Bollobás. Random graphs. In *Modern Graph Theory*, pages 215–252. Springer, 1998.

[5] Béla Bollobás, Svante Janson, and Oliver Riordan. The phase transition in inhomogeneous random graphs. *Random Structures & Algorithms*, 31(1):3–122, 2007.

[6] Tom Britton, Maria Deijfen, and Anders Martin-Löf. Generating simple random graphs with prescribed degree distribution. *Journal of Statistical Physics*, 124(6):1377–1397, 2006.

[7] Ningyuan Chen and Mariana Olvera-Cravioto. Directed random graphs with given degree distributions. *Stochastic Systems*, 3(1):147–186, 2013.

[8] Fan Chung and Linyuan Lu. The average distances in random graphs with given expected degrees. *Proceedings of the National Academy of Sciences*, 99(25):15879–15882, 2002.

[9] Fan Chung and Linyuan Lu. Connected components in random graphs with given expected degree sequences. *Annals of Combinatorics*, 6(2):125–145, 2002.

[10] Fan Chung, Linyuan Lu, and Van Vu. Spectra of random graphs with given expected degrees. *Proceedings of the National Academy of Sciences*, 100(11):6313–6318, 2003.

[11] Colin Cooper and Alan Frieze. The size of the largest strongly connected component of a random digraph with a given degree sequence. *Combinatorics, Probability and Computing*, 13(3):319–337, 2004.

[12] Guy Degla. An overview of semi-continuity results on the spectral radius and positivity. *Journal of Mathematical Analysis and Applications*, 338(1):101–110, 2008.

[13] Paul Erdős and Alfréd Rényi. On random graphs I. *Publ. Math. Debrecen*, 6:290–297, 1959.

[14] Paul Erdős and Alfréd Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, 5(1):17–60, 1960.

[15] Paul Erdős and Alfréd Rényi. On the strength of connectedness of a random graph. *Acta Mathematica Academiae Scientiarum Hungarica*, 12(1-2):261–267, 1964.

[16] Alan Frieze and Michał Karoński. *Introduction to random graphs.* Cambridge University Press, 2015.

[17] Richard M Karp. The transitive closure of a random digraph. *Random Structures & Algorithms*, 1(1):73–93, 1990.

[18] Ivan Kryven. Emergence of the giant weak component in directed random graphs with arbitrary degree distributions. *Physical Review E*, 94(1):012315, 2016.

[19] Ivan Kryven. Finite connected components in infinite directed and multiplex networks with arbitrary degree distributions. *Physical Review E*, 96(5):052304, 2017.

[20] Tomasz Łuczak. The phase transition in the evolution of random digraphs. *Journal of Graph Theory*, 14(2):217–223, 1990.

[21] Tomasz Łuczak and Joel E Cohen. Giant components in three-parameter random directed graphs. *Advances in Applied Probability*, 24(4):845–857, 1992.

[22] Tomasz Łuczak, Boris Pittel, and John C Wierman. The structure of a random graph at the point of the phase transition. *Transactions of the American Mathematical Society*, 341(2):721–748, 1994.

[23] Joel C Miller, Anja C Slim, and Erik M Volz. Edge-based compartmental modelling for infectious disease spread. *Journal of the Royal Society Interface*, 9(70):890–906, 2012.

[24] Charles J Mode. *Multitype branching processes: Theory and applications*, volume 34. American Elsevier Pub. Co., 1971.

[25] Michael Molloy and Bruce Reed. A critical point for random graphs with a given degree sequence. *Random Structures & Algorithms*, 6(2-3):161–180, 1995.

[26] Ilkka Norros and Hannu Reittu. On a conditionally Poissonian graph process. *Advances in Applied Probability*, 38(01):59–75, 2006.

[27] Mathew D Penrose. The strong giant in a random digraph. *Journal of Applied Probability*, 53(1):57–70, 2016.

[28] Sidney Resnick and Gennady Samorodnitsky. Tauberian theory for multivariate regularly varying distributions with application to preferential attachment networks. *Extremes*, 18(3):349–367, 2015.

[29] Sidney I Resnick. *Heavy-tail phenomena: Probabilistic and statistical modeling.* Springer Science & Business Media, 2007.

[30] Oliver Riordan. The small giant component in scale-free random graphs. *Combinatorics, Probability and Computing*, 14(5-6):897–938, 2005.

[31] Gennady Samorodnitsky, Sidney Resnick, Don Towsley, Richard Davis, Amy Willis, and Phyllis Wan. Nonstandard regular variation of in-degree and out-degree in the preferential attachment model. *Journal of Applied Probability*, 53(1):146–161, 2016.

[32] Eugene Seneta. *Non-negative matrices and Markov chains.* Springer Science & Business Media, 2006.

[33] Gabor Timár, Alexander V Goltsev, Sergey N Dorogovtsev, and José FF Mendes. Mapping the structure of directed networks: Beyond the bow-tie diagram. *Physical Review Letters*, 118(7):078301, 2017.

[34] Remco van der Hofstad. *Random graphs and complex networks*, volume 43. Cambridge University Press, 2016.

[35] Pim van der Hoorn and Mariana Olvera-Cravioto. Typical distances in the directed configuration model. *Annals of Applied Probability*, 28(3):1739–1792, 2018.

[36] Yana Volkovich, Nelly Litvak, and Bert Zwart. Extremal dependencies and rank correlations in power law networks. In *International Conference on Complex Sciences*, pages 1642–1653. Springer, 2009.