Submitted to the Annals of Probability

RANDOM NETWORKS WITH SUBLINEAR PREFERENTIAL ATTACHMENT: THE GIANT COMPONENT

By Steffen Dereich and Peter Mörters*

Philipps-Universität Marburg and University of Bath

We study a dynamical random network model in which at every construction step a new vertex is introduced and attached to every existing vertex independently with a probability proportional to a concave function f of its current degree. We give a criterion for the existence of a giant component, which is both necessary and sufficient, and which becomes explicit when f is linear. Otherwise it allows the derivation of explicit necessary and sufficient conditions, which are often fairly close. We give an explicit criterion to decide whether the giant component is robust under random removal of edges. We also determine asymptotically the size of the giant component and the empirical distribution of component sizes in terms of the survival probability and size distribution of a multitype branching random walk associated with f.

1. Introduction.

1.1. *Motivation and background.* Since the publication of the highly influential paper of Barabási and Albert [BA99] the preferential attachment paradigm has captured the imagination of scientists across the disciplines and has led to a host of, from a mathematical point of view mostly non-rigorous, research. The underlying idea is that the topological structure of large networks, such as the World-Wide-Web, social interaction or citation networks, can be explained by the principle that these networks are built dynamically, and new vertices prefer to be attached to vertices which have already a high degree in the existing network.

Barabási and Albert [BA99] and their followers argue that, by building a network in which every new vertex is attached to a number of old vertices with a probability proportional to a linear function of the current degree, we obtain networks whose degree distribution follows a power law. This

*The second author would like to thank EPSRC for support through an *Advanced Research Fellowship.*

AMS 2000 subject classifications: Primary 05C80; secondary 60C05, 90B15

Keywords and phrases: Barabási-Albert model, Erdős-Rényi model, power law, scale-free network, nonlinear preferential attachment, dynamic random graph, giant component, cluster, multitype branching random walk

degree distribution is consistent with that observed in large real networks, but quite different from the one encountered in the Erdős-Rényi model, on which most of the mathematical literature was focused by this date. Soon after that, Krapivsky and Redner [KR01] suggested to look at more general models, in which the probability of attaching a new vertex to a current one could be an arbitrary function f of its degree, called the *attachment rule*.

In this paper we investigate the properties of preferential attachment networks with general concave attachment rules. There are at least two good reasons to do this: On the one hand it turns out that global features of the network can depend in a very subtle fashion on the function f and only the possibility to vary this parameter gives sufficient leeway for statistical modelling and allows a critical analysis of the robustness of the results. On the other hand we are interested in the transitions between different qualitative behaviours as we pass from *absence* of preferential attachment, the case of constant attachment rules f, effectively corresponding to a variant of the Erdős-Rényi model, to *strong forms* of preferential attachment as given by linear attachment rules f. In a previous paper [DM09] we have studied degree distributions for such a model. We found the exact asymptotic degree distributions, which constitute the crucial tool for comparison with other models. The main result of [DM09] showed the emergence of a *perpetual hub*, a vertex which from some time on remains the vertex of maximal degree, when the tail of f is sufficiently heavy to ensure convergence of the series $\sum 1/f(n)^2$. In the present paper, which is independent of [DM09], we look at the global connectivity features of the network and ask for the emergence of a *giant component*, i.e. a connected component comprising a positive fraction of all vertices present.

Our first main result gives a necessary and sufficient criterion for the existence of a giant component in terms of the spectral radii of a family of compact linear operators associated with f, see Theorem 1.1. An analysis of this result shows that a giant component can exist for two separate reasons: *either* the tail of f at infinity is sufficiently heavy so that due to the strength of the preferential attachment mechanism the topology of the network enforces existence of a giant component *or* the bulk of f is sufficiently large to ensure that the edge density of the network is high enough to connect a positive proportion of vertices. We show that in the former case the giant component is robust under random deletion of edges, whereas it is not in the latter case. In Theorem 1.6 we characterise the robust networks by a completely explicit criterion.

The general approach to studying the connectivity structure in our model is to analyse a process that systematically explores the neighbourhood of a vertex in the network. Locally this neighbourhood looks approximately like a tree, which is constructed using a spatial branching process. The properties of this random tree determine the connectivity structure. We show that the asymptotic size of the giant component is determined by the survival probability, see Theorem 1.8, and the proportion of components with a given size is given by the distribution of the total number of vertices in this tree, see Theorem 1.9. It should be mentioned that although the tree approximation holds only *locally* it is sufficiently powerful to give *global* results through a technique called *sprinkling*.

This approach as such is not new, for example it has been carried out for the class of inhomogeneous random graphs by Bollobás, Janson and Riordan in the seminal paper [BJR07]. What is new here is that the approach is carried forward very substantially to treat the much more complex situation of a preferential attachment model with a wide range of attachment functions including nonlinear ones. The increased complexity originates in the first instance from the fact that the presence of two potential edges in our model is *not independent* if these have the same left end vertex. This is reflected in the fact that in the spatial branching process underlying the construction the offspring distributions are not given by a Poisson process. Additionally, due to the nonlinearity of the attachment function, information about parent vertices has to be retained in the form of a type chosen from an infinite type space. Hence, rather than being a relatively simple Galton-Watson tree, the analysis of our neighbourhoods has to be built on an approximation by a multitype branching random walk, which involves an infinite number of offspring and an uncountable type space. In the light of this it is rather surprising that we are able to get very fine explicit results even in the fully nonlinear case, in particular the explicit characterisation of robustness, see Theorem 1.6. Moreover, in the nonlinear case the abstract criterion for the existence of a giant component can be approximated and allows explicit necessary or sufficient estimates, which are typically rather close, see Proposition 1.10.

Although our results focus on the much harder case of nonlinear attachment rules, they are also new in the case of linear attachment rules f and so represent very significant progress on several fronts of research. Indeed, while the criterion for existence of a giant component is abstract for a general attachment function, it becomes completely explicit if this function is linear, see Proposition 1.3. Similarly our formula for the percolation threshold becomes explicit in the linear case and our result also includes behaviour at criticality, see Remark 1.7. Fine results like this are currently unavailable for the most studied variants of preferential attachment models with linear attachment rules, in particular those reviewed by Dommers et al. [DHH10].

1.2. *The model.* We call a concave function $f\colon \{0,1,2,\dots\} \longrightarrow (0,\infty)$ with $f(0) \leqslant 1$ and

$$\Delta f(k) := f(k+1) - f(k) < 1 \qquad \text{for all } k \geqslant 0,$$

an *attachment rule*. With any attachment rule we associate the parameters $\gamma^+ := \max_{k\geq 0} \Delta f(k)$ and $\gamma^- := \min_{k\geq 0} \Delta f(k)$, which satisfy $0 \leqslant \gamma^- \leqslant \gamma^+ < 1$. By concavity the limit

$$\gamma := \lim_{n\to\infty} \frac{f(n)}{n} \quad \text{exists and } \gamma = \gamma^-. \tag{1}$$

Observe also that any attachment rule f is non-decreasing with $f(k) \leqslant k+1$ for all $k \geqslant 0$.

Given an attachment rule f, we define a growing sequence $(\mathcal{G}_N)_{N\in\mathbb{N}}$ of random networks by the following iterative scheme:

- The network $\mathcal{G}_1$ consists of a single vertex (labeled 1) without edges,
- at each time $N \geqslant 1$, given the network $\mathcal{G}_N$, we add a new vertex (labeled $N+1$) and
- insert for each old vertex M a directed edge $N+1 \to M$ with probability
$$\frac{f(\text{indegree of } M \text{ at time } N)}{N},$$
to obtain the network $\mathcal{G}_{N+1}$.

The new edges are inserted independently for each old vertex. Note that our conditions on f guarantee that in each evolution step the probability for adding an edge is smaller or equal to 1. Edges in the random network $\mathcal{G}_N$ are *dependent* if they point towards the same vertex and independent otherwise. Formally we are dealing with directed networks, but indeed, by construction, all edges are pointing from the younger to the older vertex, so that the directions can trivially be recreated from the undirected (labeled) graph. All the notions of connectedness, which we discuss in this paper, are based on the *undirected* networks.

Our model differs from that studied in the majority of publications in one respect: We do not add a fixed number of edges in every step but a random number, corresponding formally to the outdegree of vertices in the directed network. It turns out, see Theorem 1.1 (b) in [DM09], that this random number is asymptotically Poisson distributed and therefore has very light tails. The formal universality class of our model is therefore determined by its asymptotic indegree distribution which, by Theorem 1.1 (a) in [DM09], is given by the probability weights

$$\mu_k = \frac{1}{1+f(k)} \prod_{l=0}^{k-1} \frac{f(l)}{1+f(l)} \qquad \text{for } k \in \mathbb{N} \cup \{0\}.$$

Note that these are power laws when $f(k)$ is of order k (but f need not be linear). More precisely, as $k \uparrow \infty$,

$$\frac{f(k)}{k} \to \gamma \in (0,1) \quad \Longrightarrow \quad \frac{-\log \mu_k}{\log k} \to 1 + \frac{1}{\gamma},$$

so that the LCD-model of Bollobás et al. [BRST01, BR03] compares to the case $\gamma = \frac{1}{2}$.

1.3. *Statement of the main results.* Fix an attachment rule f and define a pure birth Markov process $(Z_t \colon t \geqslant 0)$ started in zero with generator

$$Lg(k) = f(k)\,\Delta g(k),$$

which means that the process leaves state k with rate $f(k)$. Given a suitable $0 < \alpha < 1$ we define a linear operator A_α on the Banach space $\mathbf{C}(\mathcal{S})$ of continuous, bounded functions on $\mathcal{S} := \{\ell\} \cup [0,\infty]$ with ℓ being a (non-numerical) symbol, by

$$A_\alpha g(\tau) := \int_0^\infty g(t)\, e^{\alpha t}\, d\mathsf{M}(t) + \int_0^\infty g(\ell)\, e^{-\alpha t}\, d\mathsf{M}^\tau(t),$$

where the increasing functions M, resp. M^τ, are given by

$$\mathsf{M}(t) = \int_0^t e^{-s}\, \mathbb{E}\big[f(Z_s)\big]\, ds, \qquad \mathsf{M}^\ell(t) = \mathbb{E}\big[Z_t\big],$$
$$\mathsf{M}^\tau(t) = \mathbb{E}\big[Z_t \,\big|\, \Delta Z_\tau = 1\big] - \mathbb{1}_{[\tau,\infty)}(t) \quad \text{for } \tau \in [0,\infty).$$

We shall see in Remark 2.6 that $\mathsf{M}^\tau \leqslant \mathsf{M}^{\tau'}$ for all $\tau \geqslant \tau' \geqslant 0$ and therefore $\mathsf{M}^\infty = \lim_{\tau\to\infty} \mathsf{M}^\tau$ is well-defined. We shall see in Lemma 3.1 that

$$A_\alpha 1(0) < \infty \quad \Longleftrightarrow \quad A_\alpha \text{ is a well-defined compact operator.}$$

In particular, the set $\mathcal{I}$ of parameters where A_α is a well-defined (and therefore also compact) linear operator is a (possibly empty) subinterval of $(0,1)$.

Recall that we say that *a giant component exists* in the sequence of networks $(\mathcal{G}_N)_{N\in\mathbb{N}}$ if the proportion of vertices in the largest connected component $\mathcal{C}_N \subset \mathcal{G}_N$ converges, for $N \uparrow \infty$, in probability to a positive number.

THEOREM 1.1 (Existence of a giant component). *No giant component exists if and only if there exists* $0 < \alpha < 1$ *such that* A_α *is a compact operator with spectral radius* $\rho(A_\alpha) \leqslant 1$.

EXAMPLE 1.2. A sufficient but not necessary criterion for existence of a giant component is that $\gamma \geqslant \frac{1}{2}$, where γ is as defined in (1), see Remark 1.11 below for the proof.

The most important example is the linear case $f(k) = \gamma k + \beta$. In this case the family of operators A_α can be analysed explicitly, see Section 1.4.2. We obtain the following result.

PROPOSITION 1.3 (Existence of a giant component: linear case). *If* $f(k) = \gamma k + \beta$ *for some* $0 \leqslant \gamma < 1$ *and* $0 < \beta \leqslant 1$, *then there exists a giant component if and only if*

$$\gamma \geqslant \frac{1}{2} \text{ or } \beta > \frac{(\frac{1}{2} - \gamma)^2}{1 - \gamma}.$$

This result corresponds to the following intuition: If the preferential attachment is sufficiently strong (i.e. $\gamma \geqslant \frac{1}{2}$), then there exists a giant component in the network for purely topological reasons and regardless of the edge density. However if the preferential attachment is weak (i.e. $\gamma < \frac{1}{2}$) then a giant component exists only if the edge density is sufficiently large.

EXAMPLE 1.4. If $\gamma = 0$ the model is a dynamical version of the Erdős-Rényi model sometimes called *Dubins'model*. Observe that in this case there is *no* preferential attachment. The criterion for existence of a giant component is $\beta > \frac{1}{4}$, a fact which is essentially known from work of Shepp [She89], see Bollobás, Janson and Riordan [BJR05, BJR07] for more details.

EXAMPLE 1.5. If $\gamma = \frac{1}{2}$ the model is conjectured to be in the same universality class as the LCD-model of Bollobás et al. [BRST01, BR03]. In this case we obtain that a giant component exists regardless of the value of β, i.e. of the overall edge density. This is closely related to the robustness of the giant component under random removal of edges, obtained in [BR03].

As the last example indicates, in some situations the giant component is robust and survives a reduction in the edge density. To make this precise in a general setup, we fix a parameter $0 < p < 1$, remove every edge in the network independently with probability $1 - p$ and call the resulting network the *percolated network*. We say the giant component in a network is *robust*, if, for every $0 < p < 1$, the percolated network has a giant component.

THEOREM 1.6 (Percolation). *Suppose f is an arbitrary attachment rule and recall the definition of the parameter γ from* (1)*. Then the giant component in the preferential attachment network with attachment rule f is robust if and only if $\gamma \geqslant \frac{1}{2}$.*

REMARK 1.7. The criterion $\gamma \geqslant \frac{1}{2}$ is equivalent to the fact that the size biased indegree distribution, with weights proportional to $k\mu_k$, has infinite first moment. Precise criteria for the existence of a giant component in the percolated network can be given in terms of the operators $(A_\alpha \colon \alpha \in \mathcal{I})$.

(i) The giant component in the network is robust if and only if $\mathcal{I} = \emptyset$. Otherwise the percolated network has a giant component if and only if
$$p > \frac{1}{\min_{\alpha \in \mathcal{I}} \rho(A_\alpha)}.$$

(ii) In the linear case $f(k) = \gamma k + \beta$, for $\gamma > 0$, the network is robust if and only if $\gamma \geqslant \frac{1}{2}$. Otherwise, the percolated network has a giant component if and only if
$$p > \left(\tfrac{1}{2\gamma} - 1\right)\left(\sqrt{1 + \tfrac{\gamma}{\beta}} - 1\right). \tag{2}$$

Observe that running percolation with retention parameter p on the network $\mathcal{G}_N$ with attachment rule f leads to a network which stochastically dominates the network with attachment rule pf. Only if f is constant, say $f(k) = \beta$, these random networks coincide and the obvious criterion for existence of a giant component in this case is $p > \frac{1}{4\beta}$. This is in line with the formal criterion obtained by letting $\gamma \downarrow 0$ in (2).

We now define a multitype branching random walk, which represents an idealization of the exploration of the neighbourhood of a vertex in the infinite network $\mathcal{G}_\infty$ and which is at the heart of our results on the sizes of connected components in the network. A heuristic explanation of the approximation of the local neighbourhoods of typical points in the networks by this branching random walk will be given at the beginning of Section 6.

In the multitype branching random walk particles have *positions* on the real line and *types* in the space $\mathcal{S}$.[1] The initial particle is of type ℓ with arbitrary starting position. Recall the definition of the pure birth Markov

[1] Although the destinction of type and space appears arbitrary at this point, it turns out that the resulting structure of a branching random walk with a compact typespace, rather than a multitype branching process with noncompact typespace, is essential for the analysis.

process $(Z_t\colon t \geqslant 0)$. For $\tau \geqslant 0$, let $(Z_t^{[\tau]}\colon t \geqslant 0)$ be the same process conditioned to have a birth at time τ.

Each particle of type ℓ in position x generates offspring

- to its right of type ℓ with relative positions at the jumps of the process $(Z_t\colon t \geqslant 0)$;
- to its left with relative positions distributed according to the Poisson point process Π on $(-\infty, 0]$ with intensity measure
$$e^t\, \mathbb{E}[f(Z_{-t})]\; dt,$$
and type being the distance to the parent particle.

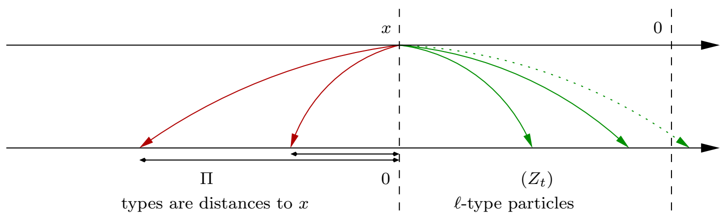

Fig 1. *Offspring of an ℓ-type particle in the branching random walk. A particle generates finitely many offspring to its left, but infinitely many offspring to its right.*

Each particle of type $\tau \geqslant 0$ in position x generates offspring

- to its left in the same manner as with a parent of type ℓ;
- to its right of type ℓ with relative positions at the jumps of $(Z_t^{[\tau]} - \mathbb{1}_{[\tau,\infty)}(t)\colon t \geqslant 0)$.

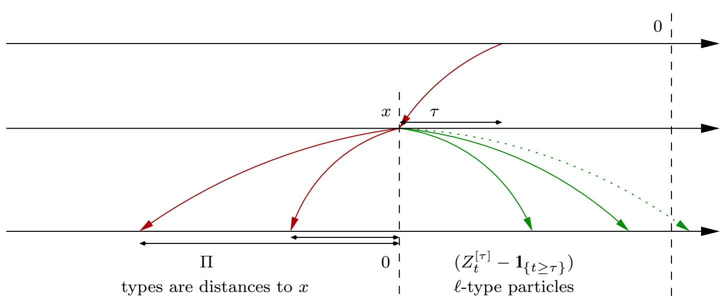

Fig 2. *Offspring of a particle of type $\tau \in [0,\infty)$ in the branching random walk. Offspring to the right have type ℓ, offspring to the left have type given by the distance to the parent.*

This branching random walk with infinitely many particles is called the *idealized branching random walk* (IBRW). Note that the functions M featuring in the definition of our operators A_α are derived from the IBRW: $\mathsf{M}(t)$ is the expected number of particles within distance t to the left of any given particle, and $\mathsf{M}^\tau(t)$ is the expected number of particles within distance t to the right of a given particle of type τ.

Equally important to us is the process representing an idealization of the exploration of the neighbourhood of a typical vertex in a large but finite network. This is the *killed* branching random walk obtained from the IBRW by removing all particles which have a position $x > 0$ together with their entire descendancy tree. Starting this process with one particle in position $x_0 < 0$ (the root), where $-x_0$ is standard exponentially distributed, we obtain a random rooted tree called the *idealized neighbourhood tree* (INT) and denoted by $\mathfrak{T}$. The genealogical structure of the tree approximates the relative neighbourhood of a typical vertex in a large but finite network. We denote by $\#\mathfrak{T}$ the total number of vertices in the INT and say that the INT *survives* if this number is infinite.

The rooted tree $\mathfrak{T}$ is the weak local limit in the sense of Benjamini and Schramm [BS01] of the sequence of graphs in our preferential attachment model. An interesting result about weak local limits for a different variant of the preferential attachment network with a linear attachment function, including the LCD-model, was recently obtained by Berger et al. [BBCS09]. In the present paper we shall not make the abstract notion of weak local limit explicit in our context. Instead, we go much further and give some fine results based on our neighbourhood approximation, which cannot be obtained from weak limit theorems alone. The following two theorems show that the INT determines the connectivity structure of the networks in a strong sense.

THEOREM 1.8 (Size of the giant component). *Let f be an attachment rule and denote by $p(f)$ the survival probability of the INT. We denote by $\mathcal{C}_N^{(1)}$ and $\mathcal{C}_N^{(2)}$ the largest and second largest connected component of $\mathcal{G}_N$. Then*

$$\frac{\#\mathcal{C}_N^{(1)}}{N} \to p(f) \quad \text{and} \quad \frac{\#\mathcal{C}_N^{(2)}}{N} \to 0, \quad \text{in probability.}$$

In particular, there exists a giant component if and only if $p(f) > 0$.

The final theorem shows the cluster size distribution in the case that no giant component exists. In this case typical connected components, or clusters, are of finite size.

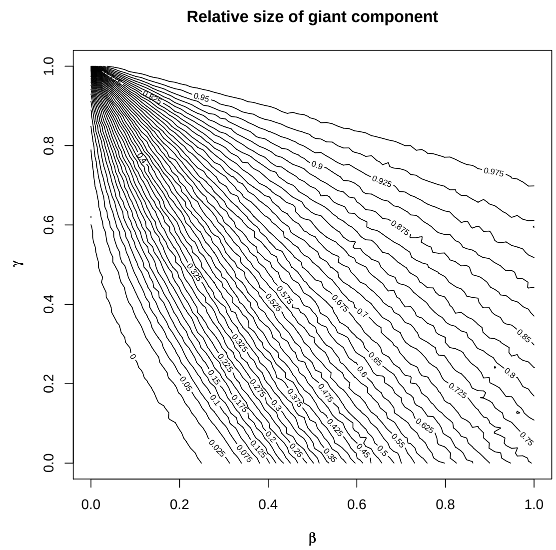

Fig 3. *Simulation of the proportion of vertices in the giant component in the linear case. The curve forming the lower envelope is determined explicitly in Proposition 1.3. The plot is based on* 15.000 *Monte Carlo simulations of the branching process for* 80 *times* 80 *gridpoints in the* (β, γ)*-plane.*

Theorem 1.9 (Empirical distribution of component sizes). *Let f be an attachment rule and denote by $\mathcal{C}_N(v)$ the connected component containing the vertex $v \in \mathcal{G}_N$. Then, for every $k \in \mathbb{N}$,*

$$\frac{1}{N} \sum_{v=1}^{N} \mathbb{1}\{\#\mathcal{C}_N(v) = k\} \longrightarrow \mathbb{P}(\#\mathfrak{T} = k) \qquad \text{in probability.}$$

1.4. *Examples.*

1.4.1. *Explicit criteria for general attachment rules.* The necessary and sufficient criterion for the existence of a giant component given in terms of the spectral radius of a compact operator on an infinite dimensional space appears unwieldy. However a small modification gives upper and lower bounds, which allow very explicit necessary or sufficient criteria that are close in many cases, see Figure 4.

Proposition 1.10. *Suppose f is an arbitrary attachment rule and let*

$$a[f] := \sum_{k=0}^{\infty} \prod_{j=0}^{k} \frac{f(j)}{\frac{1}{2} + f(j)} \quad \text{and} \quad c[f] := \sum_{k=0}^{\infty} \prod_{j=0}^{k} \frac{f(j+1)}{\frac{1}{2} + f(j+1)} \geqslant a[f].$$

(i) If $a[f] > \frac{1}{2}$*, then there exists a giant component.*
(ii) If $\frac{1}{2}\left(a[f] + \sqrt{a[f]c[f]}\right) \leqslant \frac{1}{2}$ *then there exists no giant component.*

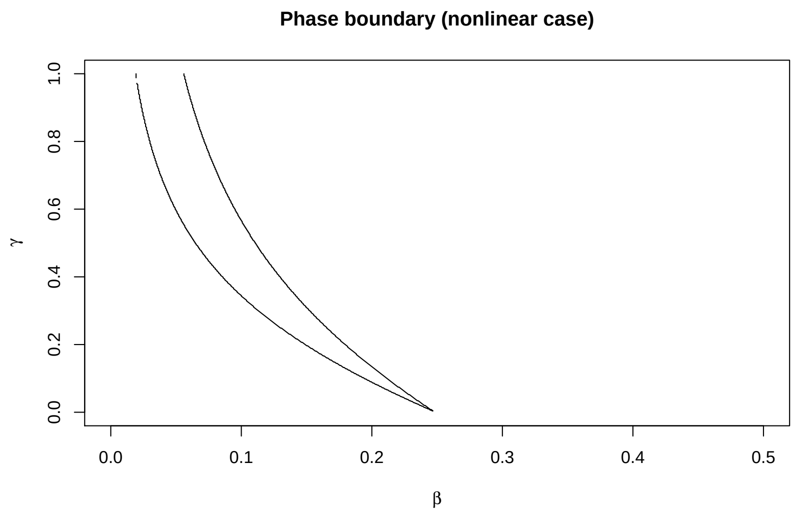

FIG 4. *For the attachment function* $f(k) = \gamma\sqrt{k} + \beta$ *the figure shows the curves* $a[f] = \frac{1}{2}$ *and* $a[f] + \sqrt{a[f]c[f]} = 1$*, which form lower and upper bound for the boundary between the two phases, nonexistence and existence of the giant component, in the* (β, γ)*–plane.*

REMARK 1.11.

- The term $\frac{1}{2}(a[f]+\sqrt{a[f]c[f]})$ differs from $a[f]$ by no more than a factor of
$$\frac{1}{2}\left(1 + \sqrt{\frac{\frac{1}{2} + f(0)}{f(0)}}\right).$$
- $a[f]$ converges if and only if $\gamma < \frac{1}{2}$. Hence a giant component exists if $\gamma \geqslant \frac{1}{2}$, as announced in Example 1.2. Otherwise there exists $\varepsilon > 0$ depending on $f(1), f(2), \ldots$ such that no giant component exists if $f(0) < \varepsilon$.

PROOF OF PROPOSITION 1.10. (i) For a lower bound on the spectral radius we recall that $\mathsf{M}^\tau \geqslant \mathsf{M}^\ell$ and therefore we may replace M^τ in the definition of A_α by M^ℓ. Then $A_\alpha g(\tau)$ no longer depends on the value of $\tau \in [0, \infty]$ but only on the fact whether $\tau = \ell$ or otherwise. Hence the operator collapses to become a 2×2 matrix of the form

$$\underline{A} = \begin{pmatrix} a(\alpha) & a(1-\alpha) \\ a(\alpha) & a(1-\alpha) \end{pmatrix}$$

with

$$a(\alpha) = \int_0^\infty e^{-\alpha t}\mathbb{E} f(Z_t)\, dt.$$

Recalling that $(Z_t\colon t\geqslant 0)$ is a pure birth process with jump rate in state k given by $f(k)$, we can simplify this expression, using T_k as the entry time into state k, as follows

$$\begin{aligned}\int_0^\infty e^{-\alpha t}\mathbb{E} f(Z_t)\, dt &= \mathbb{E}\sum_{k=0}^\infty f(k)\int_{T_k}^{T_{k+1}} e^{-\alpha t}\, dt \\ &= \sum_{k=0}^\infty f(k)\tfrac{1}{\alpha}\big[\mathbb{E}e^{-\alpha T_k} - \mathbb{E}e^{-\alpha T_{k+1}}\big].\end{aligned}$$

Recalling that T_k is the sum of independent exponential random variables with parameter $f(j)$, $j=0,\dots,k-1$, we obtain

$$\mathbb{E}e^{-\alpha T_k} = \prod_{j=0}^{k-1}\frac{f(j-1)}{f(j-1)+\alpha},$$

and hence

$$a(\alpha) = \sum_{k=0}^\infty\prod_{j=0}^{k}\frac{f(j)}{f(j)+\alpha}.$$

Now note that $\rho(\underline{A}) = a(\alpha)+a(1-\alpha)$ and since a is convex this is minimal for $\alpha=\frac12$, whence $\rho(\underline{A})\geqslant 2a(\frac12) = 2a[f]$. This shows that the given criterion is sufficient for the existence of a giant component.

(ii) For an upper bound on the spectral radius we use Lemma 2.5 to see that $\mathsf{M}^\tau\leqslant\mathsf{M}^0$ and therefore we may replace M^τ in the definition of A_α by M^0, again reducing the operator A_α to a 2×2 matrix which now has the form

$$\overline{A} = \begin{pmatrix} a(\alpha) & a(1-\alpha) \\ c(\alpha) & a(1-\alpha) \end{pmatrix},$$

with $a(\alpha)$ as before and

$$c(\alpha) = \int_0^\infty e^{-\alpha t}\mathbb{E}^1[f(Z_t)]\, dt,$$

where $\mathbb{E}^1$ is the expectation with respect to the Markov process $(Z_t : t \geqslant 0)$ started with $Z_0 = 1$. As before we obtain

$$\begin{aligned} c(\alpha) &= \mathbb{E}^1\Big[\sum_{k=1}^{\infty} f(k) \int_{T_k}^{T_{k+1}} e^{-\alpha t}\, dt\Big] = \sum_{k=1}^{\infty} f(k) \tfrac{1}{\alpha} \left[\mathbb{E}^1[e^{-\alpha T_k}] - \mathbb{E}^1[e^{-\alpha T_{k+1}}]\right] \\ &= \sum_{k=1}^{\infty} f(k) \tfrac{1}{\alpha} \Big[\prod_{j=2}^{k} \frac{f(j-1)}{f(j-1)+\alpha} - \prod_{j=2}^{k+1} \frac{f(j-1)}{f(j-1)+\alpha} \Big] \\ &= \sum_{k=0}^{\infty} \prod_{j=0}^{k} \frac{f(j+1)}{f(j+1)+\alpha}. \end{aligned}$$

Choosing $\alpha = \frac{1}{2}$, we get $\rho(\overline{A}) = a[f] + \sqrt{a[f]c[f]}$, which finishes the proof. □

1.4.2. *The case of linear attachment rules.* We show how in the linear case $f(k) = \gamma k + \beta$ the operators $(A_\alpha : \alpha \in \mathcal{I})$ can be analysed explicitly and allow to infer Proposition 1.3 from Theorem 1.1. We write $\mathbb{P}^k$ and $\mathbb{E}^k$ for probability and expectation with respect to the Markov process $(Z_t : t \geqslant 0)$ started with $Z_0 = k$.

LEMMA 1.12. *For $f(k) = \gamma k + \beta$ we have, for all $k \geqslant 0$,*

$$\mathbb{E}^k[f(Z_t)] = f(k) e^{\gamma t}, \quad \mathbb{E}^k\big[f(Z_t)^2\big] = \big(f(k)^2 + f(k)\gamma\big)\, e^{2\gamma t} - f(k)\gamma e^{\gamma t},$$

and therefore

$$d\mathsf{M}(t) = \beta e^{(\gamma-1)t}\, dt, \quad d\mathsf{M}^\ell(t) = \beta e^{\gamma t}\, dt, \quad d\mathsf{M}^\tau(t) = (\beta+\gamma) e^{\gamma t}\, dt,$$

for $\tau \in [0, \infty]$.

PROOF. Recall the definition of the generator L of $(Z_t : t \geqslant 0)$. The process $(X_t : t \geqslant 0)$ given by

$$X_t = f(Z_t) - \int_0^t Lf(Z_s)\, ds = f(Z_t) - \gamma \int_0^t f(Z_s)\, ds$$

is a local martingale. Let $(\tau_n)_{n \in \mathbb{N}}$ be a localising sequence of stopping times and note that

$$\begin{aligned} \mathbb{E}^k[f(Z_t)] &= \lim_{n\to\infty} \mathbb{E}^k f(Z_{t\wedge\tau_n}) = f(k) + \gamma \lim_{n\to\infty} \mathbb{E}^k \int_0^{t\wedge\tau_n} f(Z_s)\, ds \\ &= f(k) + \gamma \int_0^t \mathbb{E}^k[f(Z_s)]\, ds. \end{aligned}$$

We obtain the unique solution $\mathbb{E}^k[f(Z_t)] = f(k)e^{\gamma t}$. The analogous approach with f replaced by f^2 gives

$$\begin{aligned}\mathbb{E}^k\big[f^2(Z_t)\big] &= \gamma^2 \int_0^t \mathbb{E}^k f(Z_s)\,ds + 2\gamma \int_0^t \mathbb{E}^k\big[f^2(Z_s)\big]\,ds + f(k)^2 \\ &= f(k)\gamma\,\big(e^{\gamma t}-1\big) + 2\gamma\int_0^t \mathbb{E}^k\big[f^2(Z_s)\big]\;ds + f(k)^2,\end{aligned}$$

and we obtain the unique solution

$$\mathbb{E}\big[f^2(Z_t)\big] = \big(f(k)^2 + f(k)\gamma\big)\,e^{2\gamma t} - f(k)\gamma e^{\gamma t}.$$

The results for M and M^ℓ follow directly from these formulas. To characterize M^τ for $\tau \in [0,\infty)$, we observe that, for $t \geqslant \tau$,

$$\begin{aligned}\mathbb{E}[f(Z_t)\mid \Delta Z_\tau = 1] &= \sum_{k=0}^\infty \mathbb{P}(Z_\tau = k)\,\frac{f(k)}{\mathbb{E}f(Z_\tau)}\,\mathbb{E}^{k+1}[f(Z_{t-\tau})] \\ &= \frac{e^{\gamma(t-2\tau)}}{\beta}\,\sum_{k=0}^\infty \mathbb{P}(Z_\tau = k)\,f(k)f(k+1) \\ &= \frac{e^{\gamma(t-2\tau)}}{\beta}\,\big(\mathbb{E}f^2(Z_\tau) + \gamma\,\mathbb{E}f(Z_\tau)\big) \\ &= \frac{e^{\gamma(t-2\tau)}}{\beta}\,(\beta^2+\beta\gamma)e^{2\gamma\tau} = (\gamma+\beta)e^{\gamma t}\end{aligned}$$

and, for $t < \tau$,

$$\begin{aligned}\mathbb{E}[f(Z_t)\mid \Delta Z_\tau = 1] &= \sum_{k=0}^\infty \mathbb{P}(Z_t = k)\,f(k)\,\frac{\mathbb{E}^k[f(Z_{\tau-t})]}{\mathbb{E}f(Z_\tau)} \\ &= \sum_{k=0}^\infty \mathbb{P}(Z_t = k)\,f(k)\,\frac{f(k)}{f(0)}\,e^{-\gamma t} = \frac{e^{-\gamma t}}{\beta}\mathbb{E}\big[f^2(Z_t)\big] = (\gamma+\beta)e^{\gamma t} - \gamma.\end{aligned}$$

From this we obtain

$$\mathsf{M}^\tau(t) = \mathbb{E}\big[Z_t^{[\tau]}\big] - \mathbb{1}_{[\tau,\infty)}(t) = \big(\tfrac{\beta}{\gamma}+1\big)e^{\gamma t} - 1 - \tfrac{\beta}{\gamma},$$

and, by differentiating, this implies $d\mathsf{M}^\tau(t) = (\beta+\gamma)e^{\gamma t}\,dt$. ☐

Proof of Proposition 1.3. As M^τ depends only on whether $\tau = \ell$ or not, the state space $\mathcal{S}$ can be collapsed into a space with just two points. The operator A_α becomes a 2×2-matrix which, as we see from the formulas below, has finite entries if and only if $\gamma < \alpha < 1-\gamma$. This implies that there exists a giant component if $\gamma \geqslant \frac{1}{2}$, as in this case the operator A_α is never well-defined. Otherwise, denoting the collapsed state of $[0,\infty)$ by the symbol $\mathfrak{r}$, the matrix equals

$$\begin{aligned} A_\alpha^{q,\mathfrak{r}} &= \beta \int_0^\infty e^{(\gamma+\alpha-1)t}\, dt = \frac{\beta}{1-\gamma-\alpha}, \quad \text{for } q \in \{\mathfrak{r}, \ell\}, \\ A_\alpha^{\ell,\ell} &= \beta \int_0^\infty e^{(\gamma-\alpha)t}\, dt = \frac{\beta}{\alpha-\gamma}, \\ A_\alpha^{\mathfrak{r},\ell} &= (\beta+\gamma) \int_0^\infty e^{(\gamma-\alpha)t}\, dt = \frac{\beta+\gamma}{\alpha-\gamma}. \end{aligned}$$

Then $\rho(A_\alpha)$ is the (unique) positive solution of the quadratic equation

$$x^2(1-\gamma-\alpha)(\alpha-\gamma) - x(\beta - 2\beta\gamma) - \beta\gamma = 0.$$

This function is minimal when the factor in front of x^2 is maximal, i.e. when $\alpha = \frac{1}{2}$. We note that

$$\rho\big(A_{\frac{1}{2}}\big) = \frac{\sqrt{\beta^2+\beta\gamma}+\beta}{\frac{1}{2}-\gamma},$$

which indeed exceeds one if and only if

$$\beta > \frac{(\frac{1}{2}-\gamma)^2}{1-\gamma}. \qquad \square$$

1.5. *Overview.* The remainder of this paper is devoted to the proofs of the main results. In *Section 2* we discuss the process describing the indegree evolution of a fixed vertex in the network and compare it to the process $(Z_t \colon t \geqslant 0)$. The results of this section will be frequently referred to throughout the main parts of the proof. *Section 3* is devoted to the study of the idealized branching random walk and explores its relation to the properties of the family of operators $(A_\alpha \colon \alpha \in \mathcal{I})$. The main result of this section is Lemma 3.3 which shows how survival of the killed IBRW can be characterised in terms of these operators. Two important tools in the proof of Theorem 1.1 are provided in *Section 4*, namely the sprinkling argument that enables us to make statements about the giant component from local information, see Proposition 4.1, and Lemma 4.2 which ensures by means of a soft argument that the oldest vertices are always in large connected components.

The core of the proof of all our theorems is provided in Sections 5 and 6. In *Section 5* we introduce the exploration process, which systematically explores the neighbourhood of a given vertex in the network. We couple this process with an analogous exploration on a random labelled tree and show that with probability converging to one both find the same local structure, see Lemma 5.2. This random labelled tree, introduced in Subsection 5.1, is still dependent on the network size N, but significantly easier to study than the exploration process itself. *Section 6* uses further coupling arguments to relate the random labelled tree of Subsection 5.1 for large N with the idealized branching random walk. The main result of these core sections is summarised in Proposition 6.1.

In *Section 7* we use a coupling technique similar to that in Section 5 to produce a variance estimate for the number of vertices in components of a given size, see Proposition 7.1. Using the machinery provided in Sections 4 to 7 the proof of Theorem 1.8 is completed in *Section 8* and the proof of Theorem 1.9 is completed in *Section 9*. Recall that Theorem 1.8 provides a criterion for the existence of a giant component given in terms of the survival probability of the killed idealized branching random walk. In Theorem 1.1 this criterion is formulated in terms of the family of operators $(A_\alpha \colon \alpha \in \mathcal{I})$, and the proof of this result therefore follows by combining Theorem 1.8 with Lemma 3.3.

The proof of the percolation result, Theorem 1.6, requires only minor modifications of the arguments leading to Theorem 1.1 and is sketched in *Section 10*. In a short *appendix* we have collected some auxiliary coupling lemmas of general nature, which are used in Section 6. Throughout the paper we use the convention that the value of positive, finite constants c, C can change from line to line, but more important constants carry an index corresponding to the lemma or formula line in which they were introduced.

2. Properties of the degree evolution process. For $m \leqslant n$, we denote by $\mathcal{Z}[m,n]$ the indegree of vertex m at time n. Then, for each $m \in \mathbb{N}$, the degree evolution process $(\mathcal{Z}[m,n] \colon n \geqslant m)$ is a time inhomogeneous Markov process with transition probabilities in the time-step $n \to n+1$ given by

$$p^{(n)}_{k,k+1} = \frac{f(k)}{n} \wedge 1 \quad \text{and } p^{(n)}_{k,k} = 1 - p^{(n)}_{k,k+1} \qquad \text{for integers } k \geqslant 0.$$

Moreover, the evolutions $(\mathcal{Z}[m,\cdot\,] \colon m \in \mathbb{N})$ are independent. We suppose that under $\mathbb{P}^k$ the evolution $(\mathcal{Z}[m,n] \colon n \geqslant m)$ starts in $\mathcal{Z}[m,m] = k$. We write

$$P_{m,n} g(k) = \mathbb{E}^k[g(\mathcal{Z}[m,n])] \qquad \text{for any } g \colon \{0,1,\ldots\} \to (0,\infty).$$

We provide several preliminary results for the process $(\mathcal{Z}[m,n] \colon n \geqslant m)$ and its continuous-time analogue $(Z_t \colon t \geqslant 0)$ in this section. These form the basis for the computations in the network. We start by analysing the pure birth process $(Z_t : t \geqslant 0)$ and its associated semigroup $(P_t : t \geqslant 0)$ in Section 2.1, and then give the analogous results for the processes $(\mathcal{Z}[m,n] : n \geqslant m)$ in Section 2.2. We then compare the processes in Section 2.3.

2.1. *Properties of the pure birth process* $(Z_t : t \geqslant 0)$. We start with a simple upper bound.

LEMMA 2.1. *Suppose that f is an attachment rule. Then, for all $s, t \geqslant 0$ and integers $k \geqslant 0$,*

$$\mathbb{E}^k[f(Z_t)] \leqslant f(k)\, e^{\gamma^+ t} \qquad \textit{and} \qquad P_{t+s}f(k) \leqslant e^{\gamma^+ t} P_s f(k).$$

PROOF. Note that $(Z_t : t \geqslant 0)$ is stochastically increasing in f. We can therefore obtain the result for fixed $k \geqslant 0$ by using that $f(n) \leqslant f(k) + \gamma^+(n-k)$ for $n \geqslant k$, and comparing with the linear model described in Lemma 1.12. □

We now look at the conditioned process $(Z_t^{[\tau]} : t \geqslant 0)$. The next two lemmas allow a comparison of the processes $(Z_t^{[\tau]} : t \geqslant 0)$ for different values of τ.

LEMMA 2.2. *For an attachment rule f, an integer $k \geqslant 0$ and $t \geqslant 0$, one has*

$$\frac{P_t f(k+1)}{P_t f(k)} \leqslant \frac{f(k+1)}{f(k)}$$

for all $t \geqslant 0$. Moreover, if f is linear, then equality holds in the display above.

PROOF. In the following, we work under the measure $\mathbb{P} = \mathbb{P}^{k+1}$, and we suppose that $(U_j : j \geqslant 0)$ is a sequence of independent random variables, uniformly distributed in $[0,1]$, that are independent of $(Z_t : t \geqslant 0)$. We denote by $T_1, T_2, \ldots$ the random jump times of $(Z_t : t \geqslant 0)$ in increasing order, set $T_0 = 0$, and consider the process $(Y_t : t \geqslant 0)$ starting in k that is constant on each interval $[T_j, T_{j+1})$ and satisfies

$$Y_{T_{j+1}} = Y_{T_j} + \mathbb{1}\{U_j \leqslant f(Y_{T_j})/f(Z_{T_j})\}. \tag{3}$$

It is straightforward to verify that $(Y_t : t \geqslant 0)$ has the same distribution as $(Z_t : t \geqslant 0)$ under $\mathbb{P}^k$. By the concavity of f we conclude that

$$\frac{f(Y_{T_j})}{f(Z_{T_j})} \geqslant \frac{f(k) + (Y_{T_j} - k)\frac{f(Z_{T_j}) - f(k)}{Z_{T_j} - k}}{f(k) + (Z_{T_j} - k)\frac{f(Z_{T_j}) - f(k)}{Z_{T_j} - f(k)}}$$

and $\frac{f(Z_{T_j})-f(k)}{Z_{T_j}-k} \leqslant \Delta f(k)$, so that

$$\frac{f(Y_{T_j})}{f(Z_{T_j})} \geqslant \frac{Y_{T_j} + \frac{f(k)}{\Delta f(k)} - k}{Z_{T_j} + \frac{f(k)}{\Delta f(k)} - k}. \tag{4}$$

Next, we couple the processes $(Y_{T_j} : j \geqslant 0)$ and $(Z_{T_j} : j \geqslant 0)$ with a Pólya urn model. Initially the urn contains balls of two colours, blue balls of weight $B_0 = \xi := f(k)/\Delta f(k)$, and red balls of weight one. In each step a ball is picked with probability proportional to its weight and a ball of the same colour is inserted to the urn which increases its weight by one. Recalling that the total weight after j draws is $j + \xi + 1$, it is straightforward to see that we can choose the weight of the blue balls after j steps as

$$B_{j+1} = B_j + \mathbb{1}\{U_j \leqslant \tfrac{B_j}{j+\xi+1}\}.$$

Now (3) and (4) imply that whenever we pick a blue ball in the jth step, the evolution $(Y_t : t \geqslant 0)$ increases by one at time T_j. Note that $(Z_t : t \geqslant 0)$ is independent of $(U_j : j \geqslant 0)$ so that

$$\begin{aligned} \mathbb{E}\big[Y_t \big| Z_t = n+k+1\big] - k \geqslant \mathbb{E}[B_n - B_0] &= \frac{\xi}{1+\xi}(n+\xi+1) - \xi \\ &= \frac{\xi n}{1+\xi} = \frac{f(k)}{f(k+1)}\, n, \end{aligned}$$

and, by the concavity of f,

$$\begin{aligned} &\mathbb{E}\big[f(Y_t)\big| Z_t = n+k+1\big] \\ &\quad \geqslant f(k) + \frac{f(n+k+1) - f(k+1)}{n}\, \big(\mathbb{E}[Y_t|Z_t = n+k+1] - k\big) \\ &\quad \geqslant f(k) + (f(n+k+1) - f(k+1))\frac{f(k)}{f(k+1)} = f(k)\, \frac{f(n+k+1)}{f(k+1)}, \end{aligned} \tag{5}$$

so that

$$\frac{P_t f(k+1)}{P_t f(k)} = \frac{\mathbb{E}[f(Z_t)]}{\mathbb{E}[f(Y_t)]} \leqslant \frac{f(k+1)}{f(k)}.$$

If f is linear all inequalities above become equalities. □

Next, we show that the semigroup (P_t) preserves concavity.

Lemma 2.3. *For every concave and monotonically increasing g and every $t \geqslant 0$, the function $P_t g$ is concave and monotonically increasing.*

PROOF. We use an urn coupling argument similar to the one of the proof of Lemma 2.2. Fix $k \geqslant 0$ and let $(Y_t^{(2)}: t \geqslant 0)$ be the pure birth process $(Z_t: t \geqslant 0)$ started in $Z_0 = k+2$. Denote $T_0 = 0$ and let $(T_j: j = 1, 2, \ldots)$ be the breakpoints of the process in increasing order. Suppose $(U_j: j \geqslant 0)$ is a sequence of independent random variables that are uniformly distributed on $[0,1]$. For $i \in \{0,1\}$, we now denote by $(Y_t^{(i)}: t \geqslant 0)$ the step functions starting in $k+i$ which have jumps of size one precisely at those times T_{j+1}, $j \geqslant 0$, where

$$U_j \leqslant \frac{f(Y_{T_j}^{(i)})}{f(Y_{T_j}^{(2)})}.$$

By concavity of f we get

$$\mathbb{P}(\Delta Y_{T_{j+1}}^{(1)} = 1 \,|\, \Delta Y_{T_{j+1}}^{(0)} = 0) = \frac{f(Y_{T_j}^{(1)}) - f(Y_{T_j}^{(0)})}{f(Y_{T_j}^{(2)}) - f(Y_{T_j}^{(0)})} \geqslant \frac{Y_{T_j}^{(1)} - Y_{T_j}^{(0)}}{Y_{T_j}^{(2)} - Y_{T_j}^{(0)}}. \tag{6}$$

Let $(\bar{T}_j: j = 1, 2, \ldots)$ denote the elements of the possibly finite set $\{T_j: j \geqslant 1, \Delta Y_{T_j}^{(0)} = 0\}$ in increasing order. We consider a Pólya urn model starting with one blue and one red ball. We denote by B_n the number of blue balls after n steps. By (6) we can couple the urn model with our indegree evolutions such that

$$\Delta B_j \leqslant \Delta Y_{\bar{T}_j}^{(1)},$$

and such that the sequence $(B_j)_{j \in \mathbb{N}}$ is independent of $(Y_t^{(2)}: t \geqslant 0)$ and $(Y_t^{(0)}: t \geqslant 0)$. Let $\bar{g}$ be the linear function on $[l, l+2+m]$ with $\bar{g}(l) = g(l)$ and $\bar{g}(l+2+m) = g(l+2+m)$. Then

$$\begin{aligned}
\mathbb{E}\big[g(Y_t^{(1)}) \,|\, Y_t^{(0)} = l, Y_t^{(2)} = l+2+m\big] \\
&\geqslant \bar{g}\big(\mathbb{E}[Y_t^{(1)} \,|\, Y_t^{(0)} = l, Y_t^{(2)} = l+2+m]\big) \geqslant \bar{g}\big(l-1+\mathbb{E}B_{2+m}\big) \\
&= \bar{g}(l+1+\tfrac{m}{2}) = \tfrac{1}{2}\big[g(l) + g(l+2+m)\big].
\end{aligned}$$

Therefore,

$$P_t g(k+1) = \mathbb{E}[g(Y_t^{(1)})] \geqslant \tfrac{1}{2}\big[\mathbb{E}[g(Y_t^{(0)})] + \mathbb{E}[g(Y_t^{(2)})]\big] = \tfrac{1}{2}\big[P_t g(k) + P_t g(k+2)\big],$$

which implies the concavity of $P_t g$. □

The fact that the semigroup preserves concavity allows us to generalise Lemma 2.2.

LEMMA 2.4. *For an attachment rule f and integers $k \geqslant 0$ and $s, t \geqslant 0$, one has*

$$\frac{P_{t+s} f(k+1)}{P_{t+s} f(k)} \leqslant \frac{P_s f(k+1)}{P_s f(k)}.$$

PROOF. The statement follows by a slight modification of Lemma 2.2. We use Z and Y as in the proof of the latter lemma and observe that by Lemma 2.3 the function

$$g(k) := \frac{P_s f(k+1)}{P_s f(k)}$$

is concave and increasing. Similarly as in (5) we get

$$\begin{aligned}
&\mathbb{E}[g(Y_t)|Z_t = n+k+1] \\
&\quad \geqslant g(k) + \frac{g(n+k+1) - g(k+1)}{n}(\mathbb{E}[Y_t|Z_t = n+k+1] - k) \\
&\quad \geqslant g(k) + (g(n+k+1) - g(k+1))\frac{f(k)}{f(k+1)} \\
&\quad \geqslant g(k) + (g(n+k+1) - g(k+1))\frac{g(k)}{g(k+1)} = g(n+k+1)\frac{g(k)}{g(k+1)}.
\end{aligned}$$

The rest of the proof is in line with the proof of Lemma 2.2. □

LEMMA 2.5 (Stochastic domination). *One can couple the process $(Z_t^{[\tau]} \colon t \geqslant 0)$ with start in $Z_0^{[\tau]} = k$ and the process $(Z_t \colon t \geqslant 0)$ with start in $Z_0 = k+1$ in such a way that*

$$\{t > 0 \colon \Delta Z_t^{[\tau]} = 1\} \subset \{t > 0 \colon \Delta Z_t = 1\} \cup \{\tau\}.$$

In particular, this implies that $Z_t^{[\tau]} + \mathbb{1}\{t < \tau\} \leqslant Z_t$ for all $t \geqslant 0$. In the linear case we have equality in both formulas.

PROOF. Suppose $(Y_t^{(2)} \colon t \geqslant 0)$ has the distribution of $(Z_t \colon t \geqslant 0)$ with start in $Z_0 = k+1$, let $T_0 = 0$ and $(T_j \colon j = 1, 2, \ldots)$ the times of discontinuities of $(Y_t^{(2)} \colon t \geqslant 0)$ in increasing order. Denote by $(U_j \colon j \geqslant 0)$ a sequence of independent random variables that are uniformly distributed on $[0, 1]$. Now define $(Y_t^{(1)} \colon t \geqslant 0)$ as the step function starting in k which increases by one (i) at time $T_{j+1} < \tau$ if

$$U_j \leqslant \frac{f(Y_{T_j}^{(1)})}{f(Y_{T_j}^{(2)})} \frac{P_{\tau - T_{j+1}} f(Y_{T_j}^{(1)} + 1)}{P_{\tau - T_{j+1}} f(Y_{T_j}^{(1)})}, \tag{7}$$

(ii) at time τ, and (iii) at time $T_{j+1} > \tau$ if

$$U_j \leqslant \frac{f(Y_{T_j \vee \tau}^{(1)})}{f(Y_{T_j}^{(2)})}. \tag{8}$$

Clearly, we have $Y_t^{(1)} + 1 \leqslant Y_t^{(2)}$ for all $t \in [0, \tau)$ and $Y_t^{(1)} \leqslant Y_t^{(2)}$ for general $t \geqslant 0$. Moreover, by Lemma 2.2, the right hand sides of the inequalities (7) and (8) are not greater than one and it is straightforward to verify that $(Y_t^{(1)} \colon t \geqslant 0)$ has the same law as the process $(Z_t^{[\tau]} \colon t \geqslant 0)$ with start in $Z_0^{[\tau]} = k$. □

REMARK 2.6. In analogy to above, one can use Lemma 2.4 to prove that two evolutions $Z^{[\sigma]}$ and $Z^{[\tau]}$ started in k with $0 < \sigma \leqslant \tau$ can be coupled such that

$$\{t \geqslant 0 : Z^{[\tau]}\} \backslash \{\tau\} \subset \{t \geqslant 0 : Z^{[\sigma]}\} \backslash \{\sigma\}.$$

2.2. *Properties of the degree evolutions* $(\mathcal{Z}[m,n] \colon n \geqslant m)$. For the processes $(\mathcal{Z}[m,n] \colon n \geqslant m)$ we get an analogous version of Lemma 2.1.

LEMMA 2.7. *For any attachment rule f, and all integers $k \geqslant 0$ and $0 < m \leqslant n$,*

$$\mathbb{E}^k[f(\mathcal{Z}[m,n])] \leqslant f(k) \left(\frac{n}{m}\right)^{\gamma^+}.$$

PROOF. Note that $(Y_n \colon n \geqslant m)$ with $Y_n := f(\mathcal{Z}[m,n]) \prod_{i=m}^{n-1} (1 + \frac{\gamma^+}{i})^{-1}$ is a supermartingale. Hence

$$\mathbb{E}^k \big[f(\mathcal{Z}[m,n])\big] \leqslant f(k) \prod_{i=m}^{n-1} (1 + \tfrac{\gamma^+}{i}) \leqslant f(k) \left(\frac{n}{m}\right)^{\gamma^+}.$$ □

We also get the following analogue of Lemma 2.2.

LEMMA 2.8. *For an attachment rule f and integers $k \geqslant 0$ and $0 < m \leqslant n$ one has*

$$\frac{P_{m,n} f(k+1)}{P_{m,n} f(k)} \leqslant \frac{f(k+1)}{f(k)}.$$

If f is linear and $f(k+1+l) \leqslant m + l$ for all $l \in \{0, \ldots, n-m-1\}$, then equality holds.

PROOF. The statement follows by a slight modification of the proof of Lemma 2.2. □

We now provide two lemmas on stochastic domination of the degree evolutions.

LEMMA 2.9 (Stochastic domination I). *For any integers $0 < m \le n_1 < \cdots < n_j$ the process $(\mathcal{Z}[m,n]\colon n \geqslant m)$ conditioned on the event $\Delta\mathcal{Z}[m,n_i] = 0$ for all $i \in \{1,\dots,j\}$ is stochastically dominated by the unconditional process.*

PROOF. First suppose that $m < n_1$. For any $k \geqslant 0$, we have

$$\mathbb{P}^k\big(\Delta\mathcal{Z}[m,m] = 1\big|\Delta\mathcal{Z}[m,n_i] = 0\,\forall i \in \{1,\dots,j\}\big) = \frac{f(k)}{m}\,\frac{\mathbb{P}^{k+1}(\Delta\mathcal{Z}[m+1,n_i] = 0\,\forall i)}{\mathbb{P}^k(\Delta\mathcal{Z}[m,n_i] = 0\,\forall i)}.$$

The denominator on the right is equal to

$$\tfrac{f(k)}{m}\,\mathbb{P}^{k+1}(\Delta\mathcal{Z}[m+1,n_i] = 0\,\forall i) + \big(1 - \tfrac{f(k)}{m}\big)\,\mathbb{P}^k(\Delta\mathcal{Z}[m+1,n_i] = 0\,\forall i) \geqslant \mathbb{P}^{k+1}(\Delta\mathcal{Z}[m+1,n_i] = 0\,\forall i),$$

and hence we get

$$\mathbb{P}^k\big(\Delta\mathcal{Z}[m,m] = 1\big|\Delta\mathcal{Z}[m,n_i] = 0\,\forall i \in \{1,\dots,j\}\big) \leqslant \frac{f(k)}{m} = \mathbb{P}^k\big(\Delta\mathcal{Z}[m,m] = 1\big), \tag{9}$$

which is certainly also true if $m = n_1$. The result follows by induction. □

The next lemma is the analogue of Lemma 2.5.

LEMMA 2.10 (Stochastic domination II). *For integers $0 \leqslant k < m < n$ there exists a coupling of the process $(\mathcal{Z}[m,n]\colon n \geqslant m)$ started in $\mathcal{Z}[m,m] = k$ and conditioned on $\Delta\mathcal{Z}[m,n] = 1$ and the unconditional process $(\mathcal{Z}[m,n]\colon n \geqslant m)$ started in $\mathcal{Z}[m,m] = k+1$ such that for the coupled random evolutions, say $(\mathcal{Y}^{(1)}[l]\colon l \geqslant m)$ and $(\mathcal{Y}^{(2)}[l]\colon l \geqslant m)$, one has*

$$\Delta\mathcal{Y}^{(1)}[l] \leqslant \Delta\mathcal{Y}^{(2)}[l] + \mathbb{1}\{l = n\},$$

and therefore in particular $\mathcal{Y}^{(1)}[l] \leqslant \mathcal{Y}^{(2)}[l]$ for all $l \geqslant m$.

PROOF. Note that

$$\mathbb{P}^k(\Delta\mathcal{Z}[m,m] = 1|\Delta\mathcal{Z}[m,n] = 1) = \frac{\mathbb{P}^k(\Delta\mathcal{Z}[m,m] = 1, \Delta\mathcal{Z}[m,n] = 1)}{\mathbb{P}^k(\Delta\mathcal{Z}[m,n] = 1)}$$
$$= \frac{\frac{f(k)}{m}\mathbb{E}^{k+1}[f(\mathcal{Z}[m+1,n])]\frac{1}{n}}{\mathbb{E}^k[f(\mathcal{Z}[m,n])]\frac{1}{n}} = \frac{f(k)}{m}\,\frac{P_{m+1,n}f(k+1)}{P_{m,n}f(k)}.$$

By Lemma 2.8, we get

$$\mathbb{P}^k(\Delta\mathcal{Z}[m,m]=1|\Delta\mathcal{Z}[m,n]=1) \leqslant \frac{f(k)}{m}\frac{P_{m+1,n}f(k+1)}{P_{m+1,n}f(k)} \leqslant \frac{f(k+1)}{m}.$$

Now the coupling of the processes can be established as in Lemma 2.5. □

LEMMA 2.11. *For all $m \leqslant n \leqslant n'$ one has*

$$\mathbb{P}(\Delta\mathcal{Z}[m,n]=1) \geqslant \mathbb{P}(\Delta\mathcal{Z}[m,n']=1).$$

PROOF. It suffices to prove the statement for $n' = n+1$ and $n \geqslant m$ arbitrary. The statement follows immediately from

$$\mathbb{P}(\Delta\mathcal{Z}[m,n]=1) = \frac{1}{n}\mathbb{E}[f(\mathcal{Z}[m,n])] = \frac{1}{n}\sum_{k=0}^{\infty}\mathbb{P}(\mathcal{Z}[m,n]=k)f(k),$$

and

$$\begin{aligned}
&\mathbb{P}(\Delta\mathcal{Z}[m,n+1]=1)\\
&\quad= \frac{1}{n+1}\sum_{k=0}^{\infty}\mathbb{P}(\mathcal{Z}[m,n]=k)\big[\frac{f(k)}{n}f(k+1)+(1-\frac{f(k)}{n})f(k)\big]\\
&\quad= \frac{1}{n}\sum_{k=0}^{\infty}\frac{n+\Delta f(k)}{n+1}f(k)\mathbb{P}(\mathcal{Z}[m,n]=k).
\end{aligned}$$ □

We finally look at degree evolutions $(\mathcal{Z}[m,n]\colon n \geqslant m)$ conditioned on both the existence and nonexistence of some edges. In this case we cannot prove stochastic domination and comparison requires a constant factor.

LEMMA 2.12. *Suppose that $(c_N)_{N\in\mathbb{N}}, (n_N)_{N\in\mathbb{N}}$ are sequences of integers such that $\lim_{N\to\infty} n_N = \infty$ and $c_N^2 n_N^{\gamma^+-1}$ is bounded from above. Then there exists a constant $C_{2.12} > 0$, such that for all $\mathcal{I}_0, \mathcal{I}_1$ disjoint subsets of $\{n_N, \dots, N\}$ with $\#\mathcal{I}_0 \leqslant c_N$ and $\#\mathcal{I}_1 \leqslant 1$ and, for any $m \in \{1,\dots,N\}$ with $n \geqslant m$, we have*

$$\begin{aligned}
&\mathbb{P}\big(\Delta\mathcal{Z}[m,n-1]=1\big|\,\Delta\mathcal{Z}[m,i]=1\forall i\in\mathcal{I}_1,\,\Delta\mathcal{Z}[m,i]=0\forall i\in\mathcal{I}_0\big)\\
&\qquad\qquad \leqslant C_{2.12}\,\mathbb{P}\big(\Delta\mathcal{Z}[m,n-1]=1\big|\,\Delta\mathcal{Z}[m,i]=1\forall i\in\mathcal{I}_1\big).
\end{aligned}$$

PROOF. We have

$$\begin{aligned}
&\mathbb{P}\big(\Delta\mathcal{Z}[m,n-1]=1\big|\,\Delta\mathcal{Z}[m,i]=1\forall i\in\mathcal{I}_1,\,\Delta\mathcal{Z}[m,i]=0\forall i\in\mathcal{I}_0\big)\\
&\qquad\qquad \leqslant \frac{\mathbb{P}(\Delta\mathcal{Z}[m,n-1]=1|\,\Delta\mathcal{Z}[m,i]=1\forall i\in\mathcal{I}_1)}{\mathbb{P}(\Delta\mathcal{Z}[m,i]=0\forall i\in\mathcal{I}_0|\,\Delta\mathcal{Z}[m,i]=1\forall i\in\mathcal{I}_1)},
\end{aligned}$$

and it remains to bound the denominator from below by a positive constant.

Using Lemma 2.10 and denoting $k = \#\mathcal{I}_1$ we obtain that

$$\begin{aligned}\mathbb{P}\big(\Delta\mathcal{Z}[m,i] = 0 \forall i \in \mathcal{I}_0 \,\big|\, \Delta\mathcal{Z}[m,i] = 1 \forall i \in \mathcal{I}_1\big) \\ \geqslant \mathbb{P}^1\big(\Delta\mathcal{Z}[m,i] = 0 \forall i \in \mathcal{I}_0\big) \geqslant \prod_{j\in\mathcal{I}_0} \mathbb{P}^1\big(\Delta\mathcal{Z}[m,j] = 0\big) \\ = \prod_{j\in\mathcal{I}_0} \Big\{1 - \frac{\mathbb{E}^1[f(\mathcal{Z}[m,j])]}{j}\Big\}.\end{aligned}$$

By Lemma 2.7 the expectation is bounded from above by $f(k)j^{\gamma^+}$ and moreover $f(k) \leqslant k+1 \leqslant 2c_N$ for N large enough. Hence we get,

$$\prod_{j\in\mathcal{I}_0} \Big\{1 - \frac{\mathbb{E}^1[f(\mathcal{Z}[m,j])]}{j}\Big\} \geqslant \prod_{j\in\mathcal{I}_0} \Big\{1 - 2\,c_N\, j^{\gamma^+-1}\Big\} \geqslant \Big(1 - 2\,c_N\, {n_N}^{\gamma^+-1}\Big)^{c_N},$$

using that $\#\mathcal{I}_0 \leqslant c_N$. As $c_N^2 {n_N}^{\gamma^+-1}$ is bounded from above, the expression on the right is bounded from zero. This implies the statement. ☐

2.3. *Comparing the degree evolution and the pure birth process.* The aim of this section is to show that the processes $(\mathcal{Z}[m,n] \colon n \geqslant m)$ and $(Z_t \colon t \geqslant 0)$ are intimately related. To this end, we set

$$t_n := \sum_{k=1}^{n-1} \frac{1}{k} \qquad \text{and } \Delta t_n := t_{n+1} - t_n = \frac{1}{n}. \tag{10}$$

LEMMA 2.13. *For fixed $n \in \mathbb{N}$, one can couple the random variables $Z_{\Delta t_n}$ and $\mathcal{Z}[n,n+1]$ under $\mathbb{P}^k$ such that, almost surely,*

$$\mathbb{P}\big(Z_{\Delta t_n} \neq \mathcal{Z}[n,n+1]\big) \leqslant (f(k+1)\,\Delta t_n)^2 \quad \textit{and} \quad (k+1) \wedge Z_{\Delta t_n} \leqslant \mathcal{Z}[n,n+1].$$

PROOF. Note that

$$\begin{aligned}\mathbb{P}^k(Z_{\Delta t_n} = k+1) &= f(k)\Delta t_n e^{-f(k)\Delta t_n} \frac{1}{\Delta t_n} \int_0^{\Delta t_n} e^{-\Delta f(k)\, u}\, du \\ &\geqslant f(k)\Delta t_n e^{-f(k+1)\Delta t_n}.\end{aligned}$$

The same lower bound is valid for the probability $\mathbb{P}^k(\mathcal{Z}[n,n+1] = k+1)$. Moreover,

$$\mathbb{P}^k(Z_{\Delta t_n} = k) = e^{-f(k)\Delta t_n} \geqslant (1 - f(k)\Delta t_n) \vee 0 = \mathbb{P}^k(\mathcal{Z}[n,n+1] = k).$$

Hence, we can couple $Z_{\Delta t_n}$ and $\mathcal{Z}[n,n+1]$ under $\mathbb{P}^k$ such that that they differ with probability less than

$$
\begin{aligned}
(11)\qquad & 1-\big[f(k)\Delta t_n e^{-f(k+1)\Delta t_n}+1-f(k)\Delta t_n\big] \\
&\qquad = f(k)\Delta t_n(1-e^{-f(k+1)\Delta t_n}) \leqslant (f(k+1)\,\Delta t_n)^2,
\end{aligned}
$$

and moreover we have $(k+1)\wedge Z_{\Delta t_n} \leqslant \mathcal{Z}[n,n+1]$. □

PROPOSITION 2.14. *There exist constants $n_0 \in \mathbb{N}$ and $C_{2.14} > 0$ such that for all integers $n_0 \leqslant m \leqslant n$ and $0 \leqslant k < m$,*

$$\big|P_{m,n}f(k) - P_{t_n-t_m}f(k)\big| \leqslant C_{2.14}\,\frac{f(k)}{m}\,P_{m,n}f(k).$$

The proof of the proposition uses several preliminary results on the semigroups $(P_t\colon t\geqslant 0)$ and $(P_{m,n}\colon n\geqslant m)$, which we derive first. For a stochastic domination argument we introduce a further time inhomogeneous Markov process. For integers $n,k\geqslant 0$, we suppose that

$$
\begin{aligned}
\tilde{\mathbb{P}}^k(\mathcal{Z}[n,n+1]=k+1) &= 1-\tilde{\mathbb{P}}^k(\mathcal{Z}[n,n+1]=k) \\
&= \Big(\frac{f(k)}{n}+\frac{1}{2}f(k)\,\Delta f(0)\,e^{\Delta f(0)}\frac{1}{n^2}\Big)\wedge 1.
\end{aligned}
$$

The corresponding semigroup is denoted by $(\tilde{P}_{m,n})_{m\,\leqslant\,n}$.

LEMMA 2.15. *Assume that there exists $n_0\in\mathbb{N}$ such that, for all integers $n\geqslant n_0$,*

$$(12)\qquad \frac{f(n)}{n}+\frac{1}{2}f(n)\,\Delta f(0)\,e^{\Delta f(0)}\frac{1}{n^2}\leqslant 1.$$

Then, for all integers $n\geqslant n_0$ and $0\leqslant k\leqslant n$, and an increasing concave $g\colon\{0,1,2,\dots\}\to\mathbb{R}$,

$$P_{\Delta t_n}g(k)\leqslant \tilde{P}_{n,n+1}g(k).$$

PROOF. Consider $\bar{f}(l)=f(k)+\Delta f(k)(l-k)$. Note that by comparison with the linear model

$$f(k)+\Delta f(k)(\mathbb{E}^k[Z_t]-k)=\mathbb{E}^k[\bar{f}(Z_t)]\leqslant f(k)e^{\Delta f(k)\,t}.$$

Hence, for $t\in[0,1]$, using that $e^x\leqslant 1+x+\frac{1}{2}\,x^2e^x$ for $x\geqslant 0$,

$$\mathbb{E}^k[Z_t]-k\leqslant\frac{f(k)}{\Delta f(k)}(e^{\Delta f(k)\,t}-1)\leqslant f(k)\,t+\tfrac{1}{2}f(k)\,\Delta f(k)\,e^{\Delta f(k)\,t}\,t^2.$$

Therefore, $\mathbb{E}^k[Z_{\Delta t_n}] \leqslant \tilde{\mathbb{E}}^k[\mathcal{Z}[n,n+1]]$ for all $n \geqslant n_0$. As g is increasing and concave and $\mathcal{Z}$ has only increments of size one, we get

$$\begin{aligned}\mathbb{E}^k[g(Z_{\Delta t_n})] &\leqslant g(k) + \big(g(k+1)-g(k)\big)\mathbb{E}^k\big[Z_{\Delta t_n} - k\big] \\ &\leqslant g(k) + \big(g(k+1)-g(k)\big)\tilde{\mathbb{E}}^k\big[\mathcal{Z}[n,n+1]-k\big] = \tilde{\mathbb{E}}^k[g(\mathcal{Z}[n,n+1])],\end{aligned}$$

as required to complete the proof. □

LEMMA 2.16. *There exists a constant $C_{2.16} > 0$, depending on f, such that for all integers $0 \leqslant k \leqslant m$ and $0 < m \leqslant n$, we have*

$$\tilde{P}_{m,n} f(k) \leqslant C_{2.16}\, P_{m,n} f(k).$$

PROOF. For $n, m \in \mathbb{N}$ with $n \geqslant m$ let $c_{m,n} := \prod_{l=m}^{n-1}(1+\frac{\kappa}{l^2})$ where $\kappa := \frac{1}{2}(\Delta f(0))^2 e^{\Delta f(0)}$. We prove by induction (over $n-m$) that for all $0 < m \leqslant n$ and $0 \leqslant k \leqslant m$,

$$\tilde{P}_{m,n} f(k) \leqslant c_{m,n}\, P_{m,n} f(k).$$

Certainly the statement is true if $n = m$. Moreover, we have

$$\tilde{P}_{m,n+1} f(k) = P_{m,m+1}\tilde{P}_{m+1,n+1} f(k) + (\tilde{P}_{m,m+1} - P_{m,m+1})\tilde{P}_{m+1,n+1} f(k),$$

and applying the induction hypothesis we get

$$\tilde{P}_{m,n+1} f(k) \leqslant c_{m+1,n+1} P_{m,n+1} f(k) + (\tilde{P}_{m,m+1} - P_{m,m+1})\tilde{P}_{m+1,n+1} f(k).$$

Moreover, for a function $g\colon \{0,1,2,\dots\} \to \mathbb{R}$, we have

$$(\tilde{P}_{m,m+1} - P_{m,m+1})\, g(k) \leqslant \frac{1}{2} f(k)\, \Delta f(0)\, e^{\Delta f(0)} \frac{1}{m^2} \Delta g(k). \tag{13}$$

Note that the transition probabilities of the new inhomogeneous Markov process have a particular product structure: For all integers $a \geqslant 1$ and $b \geqslant 0$, one has

$$\tilde{\mathbb{P}}^b(\mathcal{Z}[a,a+1] = b+1) = (\psi_a \cdot f(b)) \wedge 1, \text{ for } \psi_a := \tfrac{1}{a} + \tfrac{1}{2}\,\Delta f(0)\, e^{\Delta f(0)} \tfrac{1}{a^2}.$$

This structure allows one to literally translate the proof of Lemma 2.8 and to obtain

$$\frac{\tilde{P}_{a_1,a_2} f(b_2)}{\tilde{P}_{a_1,a_2} f(b_1)} \leqslant \frac{f(b_2)}{f(b_1)},$$

for integers $a_1, a_2 \geqslant 1$ and $b_1, b_2 \geqslant 0$ with $a_1 \leqslant a_2$ and $b_1 \leqslant b_2$. Consequently, using (13) and the induction hypothesis,

$$\begin{aligned}(\tilde{P}_{m,m+1} - P_{m,m+1})\tilde{P}_{m+1,n+1}f(k) \\ \leqslant \frac{1}{2}f(k)\,\Delta f(0)\,e^{\Delta f(0)}\frac{1}{m^2}\frac{\Delta f(k)}{f(k)}\tilde{P}_{m+1,n+1}f(k) \\ \leqslant \frac{\kappa}{m^2}\tilde{P}_{m+1,n+1}f(k) \leqslant \frac{\kappa}{m^2}c_{m+1,n+1}P_{m+1,n+1}f(k).\end{aligned} \tag{14}$$

Altogether, we get

$$\tilde{P}_{m,n+1}f(k) \leqslant \big(1 + \frac{\kappa}{m^2}\big)c_{m+1,n+1}P_{m,n+1}f(k) = c_{m,n+1}P_{m,n+1}f(k),$$

and the statement follows since all constants are uniformly bounded by $\prod_{l=1}^{\infty}(1 + \frac{\kappa}{l^2}) < \infty$. □

PROOF OF PROPOSITION 2.14. We choose n_0 as in Lemma 2.15 and let k, m, n be integers with $n_0 \leqslant m \leqslant n$ and $0 \leqslant k \leqslant m$. We represent $\mathbb{E}^k[f(\mathcal{Z}[m,n])] - \mathbb{E}^k[f(Z_{t_n - t_m})]$ as the telescoping sum

$$P_{m,n}f(k) - P_{t_n - t_m}f(k) = \sum_{l=m}^{n-1} \underbrace{P_{m,l}(P_{l,l+1} - P_{t_{l+1}-t_l})P_{t_n - t_{l+1}}f(k)}_{=:\Sigma_l}. \tag{15}$$

In the following, we fix $l \in \{m, \dots, n-1\}$ and analyse the summand Σ_l. First note that by Lemma 2.2, one has for arbitrary integers $0 \leqslant a \leqslant b$,

$$\begin{aligned}\varphi(a,b) := \mathbb{E}^b[f(Z_{t_n - t_{l+1}})] - \mathbb{E}^a[f(Z_{t_n - t_{l+1}})] \\ \leqslant \frac{f(b) - f(a)}{f(a)}\,\mathbb{E}^a[f(Z_{t_n - t_{l+1}})].\end{aligned} \tag{16}$$

In the first part of the proof, we provide an upper bound for

$$\psi(a) := \big|(P_{l,l+1} - P_{t_{l+1}-t_l})P_{t_n - t_{l+1}}f(a)\big|, \qquad \text{for } 0 \leqslant a < l.$$

We couple $Z_{\Delta t_l}$ and $\mathcal{Z}[l, l+1]$ under $\mathbb{P}^a$ as in Lemma 2.13 and denote by $\Upsilon^{(1)}$ and $\Upsilon^{(2)}$ the respective random variables. There are two possibilities for the coupling to fail: either $\Upsilon^{(1)} \geqslant a+2$ and $\Upsilon^{(2)} = a+1$, or $\Upsilon^{(1)} = a$ and $\Upsilon^{(2)} = a+1$. Consequently,

$$\begin{aligned}\psi(a) \leqslant \mathbb{P}(\Upsilon^{(1)} = a, \Upsilon^{(2)} = a+1)\,\varphi(a, a+1) \\ + \mathbb{E}\big[\mathbb{1}_{\{\Upsilon^{(1)} \geqslant a+1\}}\varphi(a+1, \Upsilon^{(1)})\big].\end{aligned} \tag{17}$$

Since, by Taylor's formula,

$$\mathbb{P}(\Upsilon^{(1)}=a,\Upsilon^{(2)}=a+1)=e^{-f(a)\Delta t_l}-(1-f(a)\Delta t_l)\leqslant\frac{1}{2}(f(a)\Delta t_l)^2,$$

we get for the first term of (17), using (16),

$$\begin{aligned}&\mathbb{P}(\Upsilon^{(1)}=a,\Upsilon^{(2)}=a+1)\,\varphi(a,a+1)\\ &\leqslant\frac{1}{2}(f(a)\Delta t_l)^2\frac{\Delta f(a)}{f(a)}\mathbb{E}^a[f(Z_{t_n-t_{l+1}})]\\ &\leqslant f(a)\,(\Delta t_l)^2\,\mathbb{E}^a[f(Z_{t_n-t_{l+1}})].\end{aligned}\tag{18}$$

Now consider the second term in (17). We have

$$\mathbb{E}\big[\mathbb{1}_{\{\Upsilon^{(1)}\geqslant a+1\}}\varphi(a+1,\Upsilon^{(1)})\big]\leqslant\underbrace{\mathbb{P}(\Upsilon^{(2)}=a+1)}_{\leqslant f(a)\,\Delta t_l}\,\mathbb{E}^{a+1}[\varphi(a+1,Z_{\Delta t_l})].\tag{19}$$

By Lemma 2.1 we have $\mathbb{E}^{a+1}[f(Z_{\Delta t_l})]\leqslant f(a+1)\,e^{\Delta f(a+1)\,\Delta t_l}$, so that we conclude with (16) that

$$\begin{aligned}\mathbb{E}^{a+1}[\varphi(a+1,Z_{\Delta t_l})]&\leqslant(e^{\Delta f(a+1)\Delta t_l}-1)\,\mathbb{E}^{a+1}[f(Z_{t_n-t_{l+1}})]\\ &\leqslant 2\,\Delta t_l\,\mathbb{E}^{a+1}[f(Z_{t_n-t_{l+1}})],\end{aligned}$$

where we used in the last step that $\Delta f(a+1)<1$ and that $e^x\leqslant 1+2x$ for $x\in[0,1]$. We combine this with the estimates (17), (18), and (19), and get

$$\psi(a)\leqslant 3\,f(a)\,(\Delta t_l)^2\,\mathbb{E}^{a+1}[f(Z_{t_n-t_{l+1}})].$$

In the next step, we deduce an estimate for $|\Sigma_l|$ defined in (15). One has

$$\begin{aligned}|\Sigma_l|&\leqslant P_{m,l}\psi(k)\leqslant 3\Delta t_l\,\mathbb{E}^k\big[\Delta t_l\,f(\mathcal{Z}[m,l])\,\mathbb{E}^{\mathcal{Z}[m,l]+1}[f(Z_{t_n-t_{l+1}})]\big]\\ &=3\Delta t_l\,\mathbb{E}^k\big[\mathbb{1}_{\{\Delta\mathcal{Z}[m,l]=1\}}\,\mathbb{E}^{\mathcal{Z}[m,l+1]}[f(Z_{t_n-t_{l+1}})]\big].\end{aligned}$$

By Lemma 2.10 we get

$$\begin{aligned}|\Sigma_l|&\leqslant 3\Delta t_l\,\mathbb{P}^k(\Delta\mathcal{Z}[m,l]=1)\,\mathbb{E}^{k+1}\big[\mathbb{E}^{\mathcal{Z}[m,l+1]}[f(Z_{t_n-t_{l+1}})]\big]\\ &=3(\Delta t_l)^2\,\mathbb{E}^k[f(\mathcal{Z}[m,l])]\,\mathbb{E}^{k+1}\big[\mathbb{E}^{\mathcal{Z}[m,l+1]}[f(Z_{t_n-t_{l+1}})]\big]\\ &=3(\Delta t_l)^2P_{m,l}f(k)\,P_{m,l+1}P_{t_n-t_{l+1}}f(k+1).\end{aligned}\tag{20}$$

We write $P_{t_n-t_{l+1}}f(k+1)=P_{t_{l+2}-t_{l+1}}P_{t_n-t_{l+2}}f(k+1)$ and note that, by Lemma 2.3, $P_{t_n-t_{l+2}}f$ is concave. Therefore, we get with Lemma 2.15 that

$P_{t_n-t_{l+1}}f(k+1) \leqslant \tilde{P}_{l+1,l+2}P_{t_n-t_{l+2}}f(k+1)$. Successive applications of this estimate and Lemma 2.16 yield

$$P_{m,l+1}P_{t_n-t_{l+1}}f(k+1) \leqslant \tilde{P}_{m,n}f(k+1) \leqslant C_{2.16}\, P_{m,n}f(k+1). \tag{21}$$

Recall from Lemma 2.7 that $P_{m,l}f(k) \leqslant (\frac{l}{m})^{\gamma^+} f(k)$. Combining with (15), (20) and (21) yields

$$\begin{aligned} &|P_{m,n}f(k) - P_{t_n-t_m}f(k)| \\ &\qquad \leqslant 3\, C_{2.16}\, f(k)\, P_{m,n}f(k+1) m^{-\gamma^+} \sum_{l=m}^{n-1} l^{-2+\gamma^+} \\ &\qquad \leqslant C_{2.14}\, \frac{f(k)}{m}\, P_{m,n}f(k), \end{aligned} \tag{22}$$

for a suitably defined constant $C_{2.14}$ depending only on f, as required. □

3. Properties of the family (A_α: $0 < \alpha < 1$) of operators. The objective of this section is to study the operators A_α and relate them to the tree INT. We start with two lemmas on the functional analytic nature of the family $(A_\alpha \colon \alpha \in \mathcal{I})$.

LEMMA 3.1.

(a) *For any* $0 < \alpha < 1$ *the following are equivalent*
 - (i) $A_\alpha 1(0) < \infty$;
 - (ii) $A_\alpha g \in \mathbf{C}(\mathcal{S})$ *for all* $g \in \mathbf{C}(\mathcal{S})$.

 The set of α *where these conditions hold is denoted by* $\mathcal{I}$.

(b) *For any* $\alpha \in \mathcal{I}$ *the operator* A_α *is strongly positive.*

(c) *For any* $\alpha \in \mathcal{I}$ *the operator* A_α *is compact.*

PROOF. Recalling the Arzelà-Ascoli theorem, the only nontrivial claim is that, if $A_\alpha 1(0) < \infty$, then the family $(A_\alpha g \colon \|g\|_\infty < 1)$ is equicontinuous. To this end recall that, for $\tau \leqslant \sigma \leqslant \infty$, by Remark 2.6, we have $\mathsf{M}^\tau \geqslant M^\sigma$ and hence

$$\big|A_\alpha g(\tau) - A_\alpha g(\sigma)\big| \leqslant \int_0^\infty e^{-\alpha t} d(\mathsf{M}^\tau - \mathsf{M}^\sigma)(t)\,.$$

Equicontinuity at ∞ follows from this by recalling the definition $\mathsf{M}^\infty = \lim_{\tau\uparrow\infty} \mathsf{M}^\tau$. Elsewhere, for $\sigma < \infty$, we use the straightforward coupling of the processes $(Z_t^{[\tau]} \colon t \geqslant 0)$ and $(Z_t^{[\sigma]} \colon t \geqslant 0)$ with the property that *if* $Z_{\sigma-\tau}^{[\sigma]} = 0$ then $Z_t^{[\tau]} = Z_{t+\sigma-\tau}^{[\sigma]}$.

Hence we get,

$$\begin{aligned}\int_0^\infty e^{-\alpha t}\, d(\mathsf{M}^\tau - \mathsf{M}^\sigma)(t) \leqslant \big(1 - e^{-\alpha(\sigma-\tau)}\big) \int_0^\infty e^{-\alpha t} d\mathsf{M}^\tau(t) \\ + \mathbb{E}\Big[\int_0^\infty e^{-\alpha t} dZ_t^{[\tau]} \mathbb{1}\{Z_{\sigma-\tau}^{[\sigma]} > 0\}\Big]\,. \end{aligned} \tag{23}$$

Note that $\int_0^\infty e^{-\alpha t} d\mathsf{M}^\tau(t) \leqslant \mathbb{E}[\int_0^\infty e^{-\alpha t} dZ^{[\tau]}(t)] \leqslant A_\alpha 1(0) < \infty$, and that $\mathbb{P}(Z_{\sigma-\tau}^{[\sigma]} > 0) \leqslant \mathbb{P}^1(Z_{\sigma-\tau} > 1) \downarrow 0$ as $\sigma \downarrow \tau$. Hence, both terms on the right hand side of (23) can be made small by making $\sigma - \tau$ small, proving the claim. □

LEMMA 3.2. *The function $\alpha \mapsto \log \rho(A_\alpha)$ is convex on $\mathcal{I}$.*

PROOF. By Theorem 2.5 of [Kat82] the function $\alpha \mapsto \log \rho(A_\alpha)$ is convex, if for each positive $g \in \mathbf{C}(\mathcal{S})$, $\varepsilon > 0$ and triplet $\alpha_1 \leqslant \alpha_0 \leqslant \alpha_2$ in $\mathcal{I}$, there are finitely many positive $g_j \in \mathbf{C}(\mathcal{S})$ and functions $\phi_j \colon \mathcal{I} \to \mathbb{R}$, $j \in \{1, \ldots, m\}$, with $\log \phi_j$ convex, such that

$$\Big\| A_{\alpha_k} g - \sum_{j=1}^m \phi_j(\alpha_k) g_j \Big\| \leqslant \varepsilon \qquad \text{for all } k \in \{0, 1, 2\}.$$

This criterion is easily checked using the explicit form of A_α, $0 < \alpha < 1$. □

With the help of the following lemma, Theorem 1.1 follows from Theorem 1.8. The result is a variant of a standard result in the theory of branching random walks adapted to our purpose, see, e.g., Hardy and Harris [HH09] for a good account of the general theory.

LEMMA 3.3. *The* INT *dies out almost surely if and only if there exists $0 < \alpha < 1$ such that A_α is a compact linear operator with spectral radius $\rho(A_\alpha) \leqslant 1$.*

PROOF. Suppose that such an α exists. By the Krein–Rutman theorem (see, e.g., Theorem 1.3 in Section 3.2 of [Pin95]) there exists a eigenvector $v \colon \mathcal{S} \to [0, \infty)$ corresponding to the eigenvalue $\rho(A_\alpha)$. Our operator A_α is strongly positive, i.e. for every $g \geqslant 0$ which is positive somewhere, we have

$$\min_{\tau \in \mathcal{S}} A_\alpha g(\tau) > 0,$$

so that v is also bounded away from zero. Let $Y_\tau^{(n)}(dt\, dx)$ be the empirical measure of types and positions of all the offspring in the nth generation of an IBRW started by a single particle of type τ positioned at the origin. With every generation of particles in the IBRW we associate a score

$$X_n := \int Y_\tau^{(n)}(dt\,dx)\, e^{-\alpha x}\, \frac{v(t)}{v(\tau)}.$$

The assumption $\rho(A_\alpha) \leqslant 1$ implies that $(X_n \colon n \in \mathbb{N})$ is a supermartingale and thus almost surely convergent. Now fix some $N > 1$, an integer $n \geqslant 2$ and the state at generation $n-1$. Suppose there is a particle with location $x < N$ in the $(n-1)$st generation. Then there is a positive probability (depending on N but not on n) that $X_n - X_{n-1} > 1$ and, as $(X_n \colon n \in \mathbb{N})$ converges, this can only happen for finitely many n. Hence the location of the leftmost particle in the IBRW diverges to $+\infty$ almost surely. This implies that the INT dies out almost surely.

Conversely, we assume that $\mathcal{I}$ is nonempty and fix $\alpha \in \mathcal{I}$. The Krein-Rutman theorem gives the existence of an eigenvector of the dual operator, which is a positive, finite measure ν on the type space $\mathcal{S}$ such that $\int v(t)\,\nu(dt) = 1$ and, for all continuous, bounded $f \colon \mathcal{S} \to \mathbb{R}$,

$$\int A_\alpha f(t)\,\nu(dt) = \rho(A_\alpha) \int f(t)\,\nu(dt)\,.$$

Because A_α is a strongly positive operator, the Krein-Rutman theorem implies that there exists $\lambda_0 < \rho(A_\alpha)$ such that $|\lambda| \leqslant \lambda_0$ for all $\lambda \in \sigma(A_\alpha) \setminus \{\rho(A_\alpha)\}$, where $\sigma(A_\alpha)$ denotes the spectrum of the operator. Hence $\rho(A_\alpha)$ is separated from the rest of the spectrum and by Theorem IV.3.16 in [Kat76] this holds for all parameters in a small neighbourhood of α. Hence, arguing as in Note 3 on Chapter II in [Kat76, pp.568-569], the mapping $\alpha \mapsto \rho(A_\alpha)$ is differentiable and its derivative equals

$$\rho'(A_\alpha) := \frac{d}{d\alpha} \int A_\alpha v(t)\,\nu(dt) = \int \frac{\partial}{\partial\alpha} A_\alpha v(t)\,\nu(dt), \tag{24}$$

where the second equality can be inferred from the minimax characterisation of eigenvalues, see e.g. Theorem 1 in [Ram83]. Given $\tau \in \mathcal{S}$ we define a martingale by

$$W_\tau^{(n)} = \rho(A_\alpha)^{-n} \iint \frac{v(t)}{v(\tau)}\, e^{-\alpha x} Y_\tau^{(n)}(dt\,dx),$$

and argue as in Theorem 1 of [KRS01] that it converges almost surely to a strictly positive limit W_τ if

$$\log \rho(A_\alpha) - \frac{\alpha\,\rho'(A_\alpha)}{\rho(A_\alpha)} > 0 \quad \text{and} \quad \sup_{\tau \in \mathcal{S}} \mathbb{E}\big[W_\tau^{(1)} \log W_\tau^{(1)}\big] < \infty. \tag{25}$$

Let us assume for the moment that the second condition holds true for all $\alpha \in \mathcal{I}$. Then, if α is such that the limit W_τ exists and is positive, it also exists for the offspring of any particle of type τ in position x, and we denote it by $W_\tau(x)$. By decomposing the population in the mth generation according to their ancestor in the nth generation, and then letting $m \to \infty$, we get

$$W_\tau = \rho(A_\alpha)^{-n} \int \frac{v(t)}{v(\tau)} \, e^{-\alpha x} \, W_t(x) \, Y_\tau^{(n)}(dt\, dx).$$

Denoting by P_τ the law of the IBRW started with a particle at the origin of type τ, we now look at the IBRW under the changed measure

$$dQ = \int \nu(d\tau) \, v(\tau) \, W_\tau \, dP_\tau \, .$$

Given a sample IBRW we build a measure μ on the set of all infinite sequences

$$((x_0, t_0), (x_1, t_1), \ldots) \, ,$$

where x_j is the location and t_j the type of a particle in the jth generation, which is a child of a particle in position x_{j-1} of type t_{j-1}, for all $j \geqslant 1$. This measure is determined by the requirement that, for any permissible sequence

$$\begin{aligned} &\mu\big\{((y_0, s_0), (y_1, s_1), \ldots) : y_0 = x_0, s_0 = t_0 \ldots, y_n = x_n, s_n = t_n\big\} \\ &\qquad\qquad = \rho(A_\alpha)^{-n} \, \frac{v(t_n)}{v(t_0)} \, \exp\{-\alpha x_n\} \, \frac{W_{t_n}(x_n)}{W_{t_0}(x_0)}. \end{aligned}$$

Looking unconditionally at the random sequence of particle types thus generated, we note that it is a stationary Markov chain on $\mathcal{S}$ with invariant distribution $v(t) \, \nu(dt)$ and transition kernel given by

$$\begin{aligned} P_{t_0}(\ell) &= \rho(A_\alpha)^{-1} \, \frac{v(\ell)}{v(t_0)} \int_0^\infty e^{-\alpha t} d\mathsf{M}^{t_0}(t), \\ P_{t_0}(dt) &= \rho(A_\alpha)^{-1} \, \frac{v(t)}{v(t_0)} \, e^{\alpha t} d\mathsf{M}(t) \qquad \text{for } t \geqslant 0. \end{aligned}$$

Using first Birkhoff's ergodic theorem and then (24) we see that, Q-almost surely, μ-almost every path has speed

$$\begin{aligned} \lim_{n\to\infty} \frac{x_n}{n} &= \frac{1}{\rho(A_\alpha)} \int \mathbb{E}\Big[\int Y_{t_0}^{(1)}(dt\, dx) \, x e^{-\alpha x} \, \frac{v(t)}{v(t_0)} \Big] \, v(t_0) \, \nu(dt_0) \\ &= -\frac{1}{\rho(A_\alpha)} \int \frac{\partial}{\partial \alpha} \frac{A_\alpha v(t_0)}{v(t_0)} \, v(t_0) \, \nu(dt_0) = -\frac{\rho'(A_\alpha)}{\rho(A_\alpha)} = -\frac{d}{d\alpha} \log \rho(A_\alpha). \end{aligned}$$

Suppose that $\alpha_0 \in \mathcal{I}$ is such that

$$\rho\big(A_{\alpha_0}\big) = \min_{\alpha \in \mathcal{I}} \rho(A_\alpha) > 1.$$

From Lemma 3.2 we can infer that there exists $\alpha > \alpha_0$ such that the first condition in (25) holds and

$$-\frac{d}{d\alpha} \log \rho(A_\alpha) < 0.$$

This implies that, Q-almost surely, there exists an ancestral line of particles diverging to $-\infty$. For the IBRW started with a particle at the origin of type ℓ we therefore have a positive probability that an ancestral line goes to $-\infty$. This implies that the INT has a positive probability of survival.

To ensure that the second condition in (25) holds we can use a cut-off procedure, and replace the offspring distribution $Y^{(1)}(dt\,dx)$ by one that takes only the first N children to the right and left into account. It is easy to see that, for fixed $0 < \alpha < 1$ and sufficiently large N, we can ensure that the modified operator $A_\alpha^{(N)}$ is close to the original one in the operator norm, and as large as we wish if the original operator is ill-defined. Hence the continuity of the spectral radius in the operator norm ensures that $\lim_{N\to\infty} \rho(A_\alpha^{(N)}) = \rho(A_\alpha)$, with the spectral radius of an ill-defined operator being infinity. Using Lemma 3.2 and the fact that a sequence of convex functions, which converges pointwise, converges uniformly on every closed set, we can choose N so that for all $0 < \alpha < 1$ the modified operators satisfy $\rho(A_\alpha^{(N)}) > 1$, while the cut-off ensures that the second criterion in (25) automatically holds. The argument above can now be applied and yields the existence of an ancestral line of particles diverging to $-\infty$, which then automatically also exists in the original IBRW. □

Our proofs, in particular the crucial sprinkling technique, relies on the following continuity property of the survival probability $p(f)$ of the INT for the attachment rule f.

LEMMA 3.4. *One has*

$$\lim_{\varepsilon \downarrow 0} p(f - \varepsilon) = p(f).$$

PROOF. We only need to consider the case where $p(f) > 0$, as otherwise both sides of the equation are zero. We denote by $\rho(\alpha, f)$ the spectral radius of the operator A_α formed with respect to the attachment function f, setting it equal to infinity if the operator is ill-defined. The assumption $p(f) > 0$

implies, by Lemma 3.3, that for all $0 < \alpha < 1$ we have $\rho(\alpha, f) > 1$. As the operator norm $\|A_\alpha\|$ for the operator formed with respect to the attachment function $f - \varepsilon$ depends continuously on $\varepsilon \geqslant 0$, we can use the continuous dependence of the spectral radius on the operator norm to obtain, for fixed α,

$$\lim_{\varepsilon \downarrow 0} \rho(\alpha, f - \varepsilon) = \rho(\alpha, f).$$

As a sequence of convex functions, which converges pointwise, converges uniformly on every closed set, we find $\varepsilon > 0$ such that $\rho(A_\alpha, f - \varepsilon) > 1$ for all $0 < \alpha < 1$. Thus, using again Lemma 3.3, we have $p(f - \varepsilon) > 0$.

Now we look at the IBRW started with one particle of type ℓ in position t, constructed using the attachment rule $f - \varepsilon$, such that any particle with position > 0 is killed along with its offspring. We denote by $E(\varepsilon, t)$ the event this process survives forever, and by $V(\varepsilon, t, \kappa)$ the probability that a particle reaches a site $< \kappa$. Then we have

$$\lim_{\kappa \to -\infty} \inf_{t \leq \kappa} \mathbb{P}\big(E(\varepsilon, t)\big) = 1.$$

For fixed $\kappa < 0$ and $0 \leqslant \varepsilon \leqslant \varepsilon_0$ we have

$$\mathbb{P}\big(E(\varepsilon, t)\big) \geqslant \mathbb{P}\big(V(\varepsilon, t, \kappa)\big)\, \mathbb{P}\big(E(\varepsilon_0, \kappa)\big) \xrightarrow{\varepsilon \downarrow 0} \mathbb{P}\big(V(0, t, \kappa)\big)\, \mathbb{P}\big(E(\varepsilon_0, \kappa)\big).$$

Note that the first probability on the right is greater or equal to $p(f)$ and that the second probability tends to one, as κ tends to $-\infty$. □

4. The giant component. This section provides two crucial tools: A tool to obtain global results from our local approximations of neighbourhoods given by the 'sprinkling' argument in Proposition 4.1, and an a priori lower bound on the size of the connected components of the oldest vertices in the system given in Lemma 4.2. We follow the convention that a sequence of events depending on the index N holds *with high probability* if the probability of these events goes to one as $N \uparrow \infty$.

PROPOSITION 4.1 (Sprinkling argument). *Let $\varepsilon \in (0, f(0))$, $\kappa > 0$, and $\bar{f}(k) = f(k) - \varepsilon$ for integers $k \geqslant 0$. Suppose that $(c_N)_{N \in \mathbb{N}}$ is a sequence of integers with*

$$\lim_{N \uparrow \infty} \big[\tfrac{1}{2} \kappa \varepsilon c_N - \log N\big] = \infty \text{ and } \lim_{N \to \infty} \frac{c_N^2}{N} = 0,$$

and that, for the preferential attachment graphs $(\bar{\mathcal{G}}_N)_{N \in \mathbb{N}}$ with attachment rule $\bar{f}$, we have

$$\sum_{v=1}^{N} \mathbb{1}\{|\bar{\mathcal{C}}_N(v)| \geqslant 2c_N\} \geqslant \kappa N \qquad \text{with high probability,}$$

where $\bar{\mathcal{C}}_N(v)$ denotes the connected component of the vertex v in $\bar{\mathcal{G}}_N$. Then there exists a coupling of the graph sequences $(\mathcal{G}_N)_{N\in\mathbb{N}}$ with $(\bar{\mathcal{G}}_N)_{N\in\mathbb{N}}$ such that $\bar{\mathcal{G}}_N \leqslant \mathcal{G}_N$ and all connected components of $\bar{\mathcal{G}}_N$ with at least $2c_N$ vertices belong to one connected component in $\mathcal{G}_N$ with at least κN vertices, with high probability.

PROOF. Note that we can couple $\bar{\mathcal{G}}_N$ and an independent Erdős-Rényi graph $\mathcal{G}_N^{\mathsf{ER}}$ with edge probability ε/N with $\mathcal{G}_N$ such that

$$\bar{\mathcal{G}}_N \leqslant \bar{\mathcal{G}}_N \vee \mathcal{G}_N^{\mathsf{ER}} \leqslant \mathcal{G}_N. \tag{26}$$

Here, $\bar{\mathcal{G}}_N \vee \mathcal{G}_N^{\mathsf{ER}}$ denotes the graph in which all edges are open that are open in at least one of the two graphs, and $\mathcal{G}' \leqslant \mathcal{G}''$ means that all edges that are open in $\mathcal{G}'$ are also open in $\mathcal{G}''$. We denote by V_N' the vertices in $\bar{\mathcal{G}}_N$ that belong to components of size at least $2c_N$ and write V_N' as the disjoint union $C_1 \cup \cdots \cup C_M$, where $C_1, \dots, C_M$ are sets of vertices such that,

- $|C_j| \in [c_N, 2c_N]$ and
- C_j belongs to one component in $\bar{\mathcal{G}}_N$, for each $j = 1, \dots, M$.

Recall (26), and note that given $\bar{\mathcal{G}}_N$ and the sets $C_1, \dots, C_M$, the Erdős-Rényi graph $\mathcal{G}_N^{\mathsf{ER}}$ connects two distinct sets C_i and C_j with probability at least

$$p_N := 1 - \big(1 - \tfrac{\varepsilon}{N}\big)^{c_N^2} \geqslant 1 - e^{-\frac{\varepsilon}{N} c_N^2} \sim \frac{\varepsilon}{N}\, c_N^2.$$

By identifying the individual sets as one vertex and interpreting the $\mathcal{G}_N^{\mathsf{ER}}$-connections as edges, we obtain a new random graph. Certainly, this dominates an Erdős-Rényi graph with M vertices and success probability p_N, which has edge intensity Mp_N. By assumption, $\frac{1}{2}\frac{\kappa N}{c_N} \leqslant M \leqslant N$ with high probability. Hence $M \to \infty$ and $Mp_N - \log M \to \infty$ in probability as $N \uparrow \infty$. By [Hof09, Thm. 5.6], the new Erdős-Rényi graph is connected with high probability. Hence, all vertices of V_N' belong to one connected component in $\mathcal{G}_N$, with high probability. □

We need an ‘a priori’ argument asserting that the connected components of the old vertices are large with high probability. This will in particular ensure that the connected component of any vertex connected to an old vertex is large.

LEMMA 4.2 (A priori estimate). *Let $(c_N)_{N\in\mathbb{N}}$ and $(n_N)_{N\in\mathbb{N}}$ be sequences of positive integers such that*

$$\lim_{N\to\infty} \frac{c_N}{\log N \log\log N} = 0 \ \textit{and} \ \lim_{N\to\infty} \frac{\log n_N}{\log N} = 0.$$

Denote by $\mathcal{C}_N(v) \subset \mathcal{G}_N$ *the connected component containing* $v \in \{1, \dots, N\}$. *Then*

$$\mathbb{P}\big(\#\mathcal{C}_N(v) < c_N \text{ for any } v \in \{1, \dots, n_N\}\big) \longrightarrow 0.$$

PROOF. We only need to show this for the case when f is constant, say equal to $\beta > 0$, as all other cases stochastically dominate this one. Note that in this case *all* edge probabilities are independent. We first fix a vertex $v \in \{1, \dots, n_N\}$ and denote by $X_1 = X_1(v)$ the number of its direct neighbours in $(n_N, N/\log N]$. We obtain, for any $\lambda > 0$,

$$\mathbb{E}e^{-\lambda X_1} = \prod_{j=n_N}^{\lfloor N/\log N\rfloor - 1} \Big(\tfrac{\beta}{j} e^{-\lambda} + \big(1 - \tfrac{\beta}{j}\big)\Big),$$

and hence, for sufficiently large N,

$$\log \mathbb{E}e^{-\lambda X_1} \leqslant -\beta\left(1 - e^{-\lambda}\right) \sum_{j=n_N}^{\lfloor N/\log N\rfloor - 1} \frac{1}{j} \leqslant -\tfrac{3}{4}\beta\left(1 - e^{-\lambda}\right)\log N\,.$$

By the exponential Chebyshev inequality we thus get for sufficiently large N,

$$\mathbb{P}\big(X_1 < \tfrac{\beta}{2}\log N\big) \leqslant N^{\lambda\frac{\beta}{2} - \frac{3\beta}{4}(1-e^{-\lambda})} \leqslant N^{-\frac{\beta}{32}}, \tag{27}$$

choosing $\lambda = \frac{1}{2}$ and using that $1 - e^{-x} \geqslant x - \frac{1}{2}x^2$ for $x \geqslant 0$ in the last step. Now let $X_2 = X_2(v)$ be the number of direct neighbours in $(N/\log N, N]$ of any of the $X_1(v)$ vertices who are direct neighbours of v in $(n_N, N/\log N]$. Since by assumption $f(k) = \beta$ for all k, we obtain, for any $\lambda > 0$,

$$\mathbb{E}\big[e^{-\lambda X_2}\big|X_1\big] = \prod_{j=\lfloor N/\log N\rfloor}^{N-1} \Big(1 + (e^{-\lambda} - 1)\big(1 - (1 - \tfrac{\beta}{j})^{X_1}\big)\Big),$$

and hence, for sufficiently large N, on the event $\{X_1 \geqslant \frac{\beta}{2}\log N\}$,

$$\begin{aligned}\log \mathbb{E}\big[e^{-\lambda X_2}\big|X_1\big] &\leqslant -(1 - e^{-\lambda})\,\tfrac{3\beta}{4}\,X_1 \sum_{j=\lfloor N/\log N\rfloor}^{N-1} \frac{1}{j} \\ &\leqslant -(1 - e^{-\lambda})\,\tfrac{\beta^2}{4}\,\log N \log\log N.\end{aligned}$$

By (27) and the exponential Chebyshev inequality (with $\lambda = 1$) we thus get for sufficiently large N,

$$\begin{aligned}\mathbb{P}\big(X_2(v) < c_N\big) &\leqslant \mathbb{P}\big(X_1 < \tfrac{\beta}{2}\log N\big) + \mathbb{P}\big(X_2(v) < c_N \big|\, X_1 \geqslant \tfrac{\beta}{2}\log N\big) \\ &\leqslant N^{-\frac{\beta}{32}} + N^{-\frac{\beta^2}{8}\log\log N + c_N/\log N}.\end{aligned}$$

Let $\lambda = \frac{1}{2}$. By our assumptions on $(c_N)_{N\in\mathbb{N}}$ and $(n_N)_{N\in\mathbb{N}}$ the sum of the right hand sides over all $v \in \{1, \dots, n_N\}$ goes to zero, ensuring that $\#\mathcal{C}_N(v) \geqslant X_2(v) \geqslant c_N$ for all $v \in \{1, \dots, n_N\}$ with high probability. □

5. The exploration process. Our aim is to 'couple' certain aspects of the network to an easier object, namely a random tree. To each of these objects we associate a dynamic process called the exploration process. In general, an *exploration process* of a graph successively collects information about the connected component of a fixed vertex by following edges emanating from already discovered vertices in a well-defined order, so that at each instance the explored part of the graph is a connected subgraph of the cluster. We show that the exploration processes of the network and the labelled tree can be defined on the same probability space in such a way that up to a stopping time, which is typically large, the explored part of the network and the tree coincide.

5.1. *A random labelled tree.* We now describe a tree $\mathbb{T}(w)$ which informally describes the neighbourhood of a vertex $w \in \mathcal{G}_N$. Any vertex in the tree is labelled by two parameters: its *location*, an element of $\{1, \dots, N\}$, and its *type*, an element of $\{\ell\} \cup \{1, \dots, N\}$. The root is given as a vertex with location w and type ℓ. A vertex v with location i and type ℓ produces independently descendants in the locations $1, \dots, i-1$ (i.e. to its left) of type i with probability

$$\mathbb{P}(v \text{ has a descendant in } j \text{ of type } i) = \mathbb{P}(\Delta\mathcal{Z}[j, i-1] = 1).$$

Moreover, independently it produces descendants to its right, which are all of type ℓ, in such a way that the cumulative sum of these descendants is distributed according to the law of $(\mathcal{Z}[i,j] \colon i+1 \leqslant j \leqslant n)$. A vertex v of type k produces descendants to the left in the same way as a vertex of type ℓ, and independently it produces descendants to the right, which are all of type ℓ, in such a way that the cumulative sum of these descendants is distributed as $(\mathcal{Z}[i,j] - \mathbb{1}_{[k,\infty)}(j) \colon i+1 \leqslant j \leqslant n)$ conditioned on $\Delta\mathcal{Z}[i, k-1] = 1$.

Observe that, given the tree and the locations of the vertices, we may reconstruct the types of the vertices in a deterministic way: any vertex whose parent is located to its left has the type ℓ, otherwise the type of the vertex is the location of the parent.

The link between this labelled tree and our network is given in the following proposition, which will be proved in Section 5.3.

PROPOSITION 5.1. *Suppose that* $(c_N)_{N\in\mathbb{N}}$ *is a sequence of integers with*

$$\lim_{N\to\infty} \frac{c_N}{\log N \log\log N} = 0.$$

Then one can couple the pair $(V, \mathcal{G}_N)$ *consisting of the network and a uniformly chosen vertex* V *with* $\mathbb{T}(V)$ *such that with high probability*

$$\#\mathcal{C}_N(V) \wedge c_N = \#\mathbb{T}(V) \wedge c_N.$$

5.2. *Exploration of the network.* We now specify how we explore a graph like our network or the tree described above, i.e., we specify the way we collect information about the connected component, or cluster, of a particular vertex v. In the first step, we explore all immediate neighbours of v in the graph. To explain a general exploration step we classify the vertices in three categories:

- *veiled vertices*: vertices for which we have not yet found connections to the cluster of v;
- *active vertices*: vertices for which we already know that they belong to the cluster, but for which we have not yet explored all its immediate neighbours;
- *dead vertices*: vertices which belong to the cluster and for which all immediate neighbours have been explored.

After the first exploration step the vertex v is marked as dead, its immediate neighbours as active and all the remaining vertices as veiled. In a general exploration step, we choose the *leftmost* active vertex, set its state to *dead*, and explore its immediate neighbours. The newly found *veiled* vertices are marked as *active*, and we proceed with another exploration step until there are no active vertices left.

In the following, we couple the exploration processes of the network and the random labelled tree started with a particle at position v and type ℓ up to a stopping time T. Before we introduce the coupling explicitly, let us quote adverse events which stop the coupling. Whenever the exploration process of the network revisits an active vertex we have found a cycle in the network. We call this event (E1) and stop the exploration so that, before time T, the explored part of the neighbourhood of v is a tree with each node having a unique location. Additionally, we stop once the explored part of the network differs from the explored part of the random labelled tree, calling this event (E2), we shall see in Section 5.3 how this can happen. In cases (E1) and (E2) we say that *the coupling fails*.

Further reasons to stop the exploration are, for parameters $n_N, c_N \in \mathbb{N}$ with $1 \leqslant n_N, c_N \leqslant N$,

(A) the number of dead and active vertices exceeds c_N,
(B) one vertex in $\{1, \dots, n_N\}$ is activated, and
(C) there are no more active vertices left.

If we stop the exploration without (E1) and (E2) being the case, we say that the *coupling succeeds*. Once the exploration has stopped, the veiled parts of the random tree and the network may be generated independently of each other with the appropriate probabilities. Hence, if we succeed in coupling the explorations, we have coupled the random labelled tree and the network.

5.3. *Coupling the explorations.* To distinguish both exploration processes, we use the term *descendant* for a child in the labelled random tree and the term *immediate neighbour* in the context of the neighbourhood exploration in the network. In the initial step, we explore all immediate neighbours of v and all the descendants of the root. Both explorations are identically distributed and they therefore can be perfectly coupled. Suppose now that we have performed k steps and that we have not yet stopped the exploration. In particular, this means that both explored subgraphs coincide and that any unveiled (i.e. active or dead) element of the labelled random tree can be uniquely referred to by its location. We now explore the descendants and immediate neighbours of the *leftmost* active vertex, say n.

First, we explore the descendants to the left (veiled and dead) and immediately check whether they themselves have right descendants in the set of dead vertices. If we discover no dead descendants, the set of newly found left descendants is identically distributed to the immediate left neighbours in the network. Thus we can couple both explorations such that they agree in this case. Otherwise we stop the exploration due to (E2).

Second, we explore the descendants to the right. If the vertex n is *not of type* ℓ, then we know already that n has no *right* descendants that were marked as dead as n itself was discovered. Since we always explore the leftmost active vertex there are no new dead vertices to the right of n. Therefore, the explorations to the right in the network and the random labelled tree are identically distributed and we stop if we find right neighbours in the set of active vertices due to (E1). If the vertex n is *of type* ℓ, then we have not gained any information about its right *descendants* yet. If we find no right descendants in the set of dead vertices, it is identically distributed to the immediate right neighbours of n in the network. We stop if right descendants are discovered that were marked as dead, corresponding to (E2), or if right descendants are discovered in the set of active vertices, corresponding to (E1).

LEMMA 5.2. *Suppose that $(c_N)_{N\in\mathbb{N}}$, $(n_N)_{N\in\mathbb{N}}$ are sequences of integers such that*

$$\lim_{N\to\infty} \frac{c_N^2}{n_N^{1-\gamma^+}} = 0.$$

Then the coupling of the exploration processes satisfies

$$\lim_{N\to\infty} \sup_{v\in\{n_N+1,\dots,N\}} \mathbb{P}\big(\textit{coupling with initial vertex } v \textit{ ends in } \text{(E1)} \textit{ or } \text{(E2)}\big) = 0,$$

i.e. the coupling succeeds with high probability.

PROOF. We analyse one exploration step in detail. Let $\mathfrak{a}$ and $\mathfrak{d}$ denote the active and dead vertices of a feasible configuration at the beginning of an exploration step, that is $\mathfrak{a}, \mathfrak{d}$ denote two disjoint subsets of $\{n_N+1,\dots,N\}$ with $\#(\mathfrak{a}\cup\mathfrak{d}) < c_N$ and $\mathfrak{a} \neq \emptyset$.

The exploration of the minimal vertex n in the set $\mathfrak{a}$ may only fail for one of the following reasons:

- (Ia) the vertex n has left descendants in $\mathfrak{d}$,
- (Ib) the vertex n has left descendants which themselves have right descendants in $\mathfrak{d}$, or
- (II) the vertex n has right descendants in $\mathfrak{a}\cup\mathfrak{d}$.

Indeed, if (Ia) and (Ib) do not occur then the exploration to the left ends neither in state (E1) nor (E2), and if (II) does not happen the exploration to the right does not fail.

Conditionally on the configuration $(\mathfrak{a}, \mathfrak{d})$, the probability for the event (Ia) equals

$$\mathbb{P}(\exists a \in \mathfrak{d} \text{ such that } \Delta\mathcal{Z}[a, n-1] = 1) \leqslant \sum_{\substack{a\in\mathfrak{d}\\ a<n}} \mathbb{P}(\Delta\mathcal{Z}[a,n-1]=1),$$

whereas the probability for (Ib) is by Lemma 2.10 equal to

$$\begin{aligned}
&\mathbb{P}(\exists a \in \mathfrak{d}^c \text{ and } b \in \mathfrak{d} \text{ such that } \Delta\mathcal{Z}[a,n-1] = \Delta\mathcal{Z}[a,b-1] = 1)\\
&\quad\leqslant \sum_{\substack{a\in\mathfrak{d}^c\\ a<n}} \sum_{\substack{b\in\mathfrak{d}\\ b>a}} \mathbb{P}(\Delta\mathcal{Z}[a,n-1] = \Delta\mathcal{Z}[a,b-1] = 1)\\
&\quad\leqslant \sum_{\substack{a\in\mathfrak{d}^c\\ a<n}} \sum_{\substack{b\in\mathfrak{d}\\ b>a}} \mathbb{P}(\Delta\mathcal{Z}[a,n-1] = 1)\, \mathbb{P}^1(\Delta\mathcal{Z}[a,b-1]=1).
\end{aligned}$$

If the vertex n is of type $\tau \neq \ell$, then the conditional probability of (II) is

$$\begin{aligned}&\mathbb{P}\big(\exists a \in \mathfrak{a} \text{ such that } \Delta\mathcal{Z}[n, a-1] = 1 \\ &\qquad\qquad \big| \, \mathcal{Z}[n, \tau-1] = 1, \Delta\mathcal{Z}[n, b-1] = 0 \, \forall b \in \mathfrak{d} \backslash \{\tau\}\big) \\ &\leqslant C_{2.12} \sum_{\substack{a \in \mathfrak{a} \cup \mathfrak{d} \\ a > n}} \mathbb{P}^1(\Delta\mathcal{Z}[n, a-1] = 1),\end{aligned}$$

using first Lemma 2.12 and then Lemma 2.10.

If the vertex n is of type ℓ, the conditional probability of (II) is

$$\mathbb{P}\big(\exists a \in \mathfrak{a} \cup \mathfrak{d} \text{ such that } \Delta\mathcal{Z}[n, a-1] = 1) \leqslant \sum_{\substack{a \in \mathfrak{a} \cup \mathfrak{d} \\ a > n}} \mathbb{P}(\Delta\mathcal{Z}[n, a-1] = 1).$$

Since, by Lemma 2.11, for any $a > n$,

$$\mathbb{P}^1\big(\Delta\mathcal{Z}[n, a-1] = 1\big) \leqslant \mathbb{P}^1\big(\Delta\mathcal{Z}[n_N + 1, n_N + 1] = 1\big),$$

we conclude that the probabilities of the events (Ia) and (II) are bounded by

$$(2 + C_{2.12})\, c_N \, \mathbb{P}^1\big(\Delta\mathcal{Z}[n_N + 1, n_N + 1] = 1\big),$$

independently of the type τ. Moreover, the probability of (Ib) is bounded by

$$c_N \, \mathbb{P}^1(\Delta\mathcal{Z}[1, n_N] = 1) \sum_{a=1}^{n-1} \mathbb{P}(\Delta\mathcal{Z}[a, n-1] = 1).$$

The sum is the expected outdegree of vertex n, which, by Lemma 2.7, is uniformly bounded and, hence, one of the events (Ia), (Ib), or (II) occurs in one step with probability less than a constant multiple of $c_N \, \mathbb{P}^1(\Delta\mathcal{Z}[1, n_N] = 1)$. As there are at most c_N exploration steps until we end in one of the states (A), (B), or (C), the coupling fails due to (E1) or (E2) with a probability bounded from above by a constant multiple of

$$c_N^2 \, \mathbb{P}^1(\Delta\mathcal{Z}[1, n_N] = 1) \leqslant C_{2.7} \, f(1) \, \frac{c_N^2}{{n_N}^{1-\gamma^+}} \to 0 \,,$$

in other words, the coupling succeeds with high probability. □

Proof of Proposition 5.1. Apply the coupling of Lemma 5.2 with $(n_N)_{N \in \mathbb{N}}$ satisfying

$$\lim_{N \to \infty} \frac{\log n_N}{\log N} = 0 \quad \text{and} \quad \lim_{N \to \infty} \frac{(\log N \log\log N)^2}{n_N^{1-\gamma^+}} = 0.$$

Then, by Lemma 4.2, we get that with high probability

$$\text{coupling ends in (B)} \implies \#\mathcal{C}_N(V) \geq c_N. \tag{28}$$

As in the proof of Lemma 4.2 one gets

$$\lim_{N\to\infty} \max_{v=1,\dots,n_N} \mathbb{P}(\#\mathbb{T}(v) < c_N) = 0$$

so that implication (28) is also valid for $\#\mathcal{C}_N(V)$ replaced by $\#\mathbb{T}(V)$. Since the coupling succeeds we have, with high probability,

$$\text{coupling ends in (A) or (B)} \iff \#\mathcal{C}_N(V) \wedge \#\mathbb{T}(V) \geq c_N,$$

and the statement follows immediately. □

6. The idealized exploration process. We now have the means to explain heuristically the approximation of the local neighbourhood of a randomly chosen vertex $V \in \mathcal{G}_N$ by the idealized random tree $\mathfrak{T}$ featuring in our main theorems. Vertices in the network $\mathcal{G}_N$ are mapped onto particles on the negative halfline in such a way that the vertex with index $n \in \{1, \dots, N\}$ is mapped onto position $t_n - t_N$, recall (10). Note that the youngest vertex is placed at the origin, and older vertices are placed to the left with decreasing intensity. In particular the position of the particle corresponding to a vertex with fixed index will move to the left as N is increasing.

Looking at a fixed observation window $[a, b]$ on the negative halfline, as $N \uparrow \infty$, we see that the number of particles in the window is increasing. At the same time the age of the vertex corresponding to a particle closest to a fixed position in the window is increasing, which means that the probability of edges between two such vertices is decreasing. As we shall see below, the combination of these two effects leads to convergence of the distribution of offspring locations on the halfline. In particular, thanks to the independence of edges with a common right endpoint, offspring to the left converges to a Poisson process by the law of small numbers, while offspring to the right converges to the point processes corresponding to the pure birth process $(Z_t : t \geqslant 0)$ if there is no dependence on previous generations.

The considerations of Section 5 suggest that the only form of dependence of the offspring distribution of a vertex on previous generations, is via the relative position of its father. This information is encoded in the type of a particle, where type ℓ indicates that its father is to the left of the particle, and a numerical type τ indicates that the father is positioned τ units to its right. It should be noted that the *relative* positions of offspring particles only depend on the absolute position of the reproducing particle via the removal of particles whose position is not in the left halfline, and which therefore do not correspond to vertices in the network $\mathcal{G}_N$. This fact produces the random walk structure, which is crucial for the analysis of the underlying tree. Our main aim now is to prove the following result.

PROPOSITION 6.1. *Suppose that $(c_N)_{N\in\mathbb{N}}$ is a sequence of integers with*

$$\lim_{N\to\infty} \frac{c_N}{\log N \log\log N} = 0$$

Then each pair $(V, \mathcal{G}_N)$ can be coupled with $\mathfrak{T}$ such that with high probability

$$\#\mathcal{C}_N(V) \wedge c_N = \#\mathfrak{T} \wedge c_N.$$

We have seen so far that the neighbourhood of a vertex v in a large network is similar to the random tree $\mathbb{T}(v)$ constructed in Section 5.1. To establish the relationship between $\mathbb{T}(V)$, for an initial vertex V chosen uniformly from $\{1,\dots,N\}$, and the idealized neighbourhood tree $\mathfrak{T}$ we apply the projection the projection

$$\pi_N \colon (-\infty, 0] \to \{1,\dots,N\},$$

which maps $t \leqslant 0$ onto the smallest $m \in \{1,\dots,N\}$ with $t \leqslant -t_N + t_m$, to each element of the INT $\mathfrak{T}$. We obtain a branching process with location parameters in $\{1,\dots,N\}$, which we call π_N-projected INT. We need to show, using a suitable coupling, that when the INT is started with a vertex $-X$, where X is standard exponentially distributed, then this projection is close to the random tree $\mathbb{T}(V)$. Again we apply the concept of an exploration process.

To this end we show that, for every $v \leqslant 0$, the π_N-projected descendants of v have a similar distribution as the descendants of a vertex in location $\pi_N(v)$ in the labelled tree of Section 5.1. We provide couplings of both distributions and control the probability of them to fail.

Coupling the evolution to the right for ℓ-type vertices. We fix $v \leqslant 0$ and $N \in \mathbb{N}$, and suppose that $m := \pi_N(v) \geqslant 2$. For an ℓ-type vertex in v the cumulative sum of π_N-projected right descendants is distributed as $(Z_{t_n - t_N - v})_{m \leqslant n \leqslant N}$. This distribution has to be compared with the distribution of $(\mathcal{Z}[m,n])_{m \leqslant n \leqslant N}$, which is the cumulative sum of right descendants of m in $\mathbb{T}(v)$.

LEMMA 6.2. *Fix $T, N \in \mathbb{N}$ and $v \leqslant 0$ with $\pi_N(v) = m \in \{2, 3, \dots, N\}$. We can couple the processes $(Z_{t_n - t_N - v}\colon n \geqslant m)$ and $(\mathcal{Z}[m, n]\colon n \geqslant m)$ such that for the coupled processes $(\mathcal{Y}^{(1)}[n]\colon n \geqslant m)$ and $(\mathcal{Y}^{(2)}[n]\colon n \geqslant m)$ we have*

$$\mathbb{P}\big(\mathcal{Y}^{(1)}[n] \neq \mathcal{Y}^{(2)}[n] \text{ for some } n \leqslant \tau\big) \leqslant (f(0) + f(T)^2)\, \frac{1}{m-1},$$

where τ is the first time when one of the processes reaches or exceeds T.

PROOF. We define the process $\mathcal{Y} = ((\mathcal{Y}^{(1)}[n], \mathcal{Y}^{(2)}[n])\colon n \geqslant m)$ to be the Markov process with starting distribution $\mathcal{L}(Z_{t_m - t_N - v}) \otimes \delta_0$ and transition kernels $p^{(n)}$ such that the first and second marginal are the respective transition probabilities of $(Z_{t_n - t_N - v}\colon n \geqslant m)$ and $(\mathcal{Z}[m, n]\colon n \geqslant m)$ and, for any integer $a \geqslant 0$, the law $p^{(n)}((a, a), \cdot\,)$ is the coupling of the laws of $Z_{\Delta t_n}$ and $\mathcal{Z}[n, n+1]$ under $\mathbb{P}^a$ provided in Lemma 2.13. Then the processes $(\mathcal{Y}^{(1)}[n]\colon n \geqslant m)$ and $(\mathcal{Y}^{(2)}[n]\colon n \geqslant m)$ are distributed as stated in the lemma. Moreover, letting σ denote the first time when they disagree, we get

$$\begin{aligned}\mathbb{P}(\sigma \leqslant \tau) &= \sum_{n=m}^{\infty} \mathbb{P}(\tau \geqslant n, \sigma = n) \\ &\leqslant \mathbb{P}(\sigma = m) + \sum_{n=m}^{\infty} \mathbb{P}\big(\sigma = n+1 \,\big|\, \tau > n, \sigma > n\big)\end{aligned}$$

and, by Lemma 2.13,

$$\mathbb{P}\big(\sigma = n+1 \,|\, \tau > n, \sigma > n\big) \leqslant \Big(f(T)\frac{1}{n}\Big)^2 \qquad \text{for } n \in \{m, m+1, \dots\}.$$

Moreover, $\mathbb{P}(\sigma = m) = \mathbb{P}(\mathcal{Y}^{(1)}[m] > 0) = 1 - e^{-(t_m - v) f(0)} \leqslant \frac{f(0)}{m-1}$. Consequently,

$$\mathbb{P}(\sigma \leqslant \tau) \leqslant \frac{f(0)}{m-1} + f(T)^2 \sum_{n=m}^{\infty} \frac{1}{n^2} \leqslant (f(0) + f(T)^2)\, \frac{1}{m-1}. \qquad \square$$

Coupling the evolution to the left. Recall that a vertex $v \leqslant 0$ produces a Poissonian number of π_N-projected descendants at the location $m \leqslant \pi_N(v)$ with parameter

$$\lambda := \int_{-t_N + t_{m-1}}^{(-t_N + t_m) \wedge v} e^{-(v-u)}\, \mathbb{E}[f(Z_{v-u})]\, du. \tag{29}$$

Here we adopt the convention that $t_0 = -\infty$. A vertex in location $n := \pi_N(v)$ in $\mathbb{T}[v]$ produces a Bernoulli distributed number of descendants in m with success probability $\mathbb{P}(\Delta\mathcal{Z}[m, n-1] = 1)$ for $m < n$ and success probability zero for $m = n$. The following lemma provides a coupling of both distributions.

LEMMA 6.3. *There exists a constant $C_{6.3} > 0$ such that the following holds: Let $m, N \in \mathbb{N}$ and $v \leqslant 0$ with $m \leqslant n := \pi_N(v)$ and define λ as in (29). If $m < n$, one can couple a* $\mathrm{Poiss}(\lambda)$ *distributed random variable with $\Delta\mathcal{Z}[m, n-1]$, such that the coupled random variables $\Upsilon^{(1)}$ and $\Upsilon^{(2)}$ satisfy*

$$\mathbb{P}(\Upsilon^{(1)} \neq \Upsilon^{(2)}) \leqslant C_{6.3} \frac{1}{m^{1+\gamma^+}} \frac{1}{n^{1-\gamma^+}}.$$

If $m = n$, a $\mathrm{Poiss}(\lambda)$ *distributed random variable $\Upsilon^{(1)}$ satisfies*

$$\mathbb{P}(\Upsilon^{(1)} \neq 0) \leqslant C_{6.3} \frac{1}{n}.$$

PROOF. It suffices to prove the second statement for $m = n \geqslant 2$. Note that $u \mapsto e^{-u}\mathbb{E}[f(Z_u)]$ is decreasing so that

$$\lambda \leqslant \int_{-t_N+t_{n-1}}^{v} e^{-(v-u)}\, \mathbb{E}[f(Z_{v-u})]\, du \leqslant f(0)\tfrac{1}{n-1},$$

which leads directly to the second statement of the lemma. Next, consider the case where $2 \leqslant m < n$. Note that for $u \in (-t_N + t_{m-1}, -t_N + t_m]$, one has $v - u \in (t_{n-1} - t_m, t_n - t_{m-1})$ which, using again that $u \mapsto e^{-u}\mathbb{E}[f(Z_u)]$ is decreasing, implies that

$$\tfrac{1}{m-1}\, e^{-(t_n - t_{m-1})}\, \mathbb{E}[f(Z_{t_n - t_{m-1}})] \leqslant \lambda \leqslant \tfrac{1}{m-1}\, e^{-(t_{n-1} - t_m)}\, \mathbb{E}[f(Z_{t_{n-1} - t_m})].$$

Next, note that by definition of t_n we have $\log \frac{n}{m} \leqslant t_n - t_m \leqslant \log \frac{n-1}{m-1}$ so that

$$(1 - \tfrac{1}{m-1})\, \tfrac{1}{n-1}\, \mathbb{E}[f(Z_{t_{n-1}-t_m})] \leqslant \lambda \leqslant \big(1 + \tfrac{1}{m-1}\big)\, \tfrac{1}{n-1}\, \mathbb{E}[f(Z_{t_{n-1}-t_m})]. \tag{30}$$

On the other hand, $\Delta\mathcal{Z}[m, n-1]$ is a Bernoulli random variable with success probability

$$p := \tfrac{1}{n-1}\mathbb{E}[f(\mathcal{Z}[m, n-1])].$$

By Lemma A.1 it suffices to control λ^2 and $|\lambda - p|$. By Proposition 2.14 and (30),

$$|\lambda - p| \leqslant C \frac{1}{m-1} \frac{1}{n-1} \big(\mathbb{E}[f(Z_{t_{n-1}-t_m})] + \mathbb{E}[f(\mathcal{Z}[m, n-1])]\big), \tag{31}$$

and

$$\lambda^2 \leqslant 4\big(\tfrac{1}{n-1}\big)^2 \mathbb{E}[f(Z_{t_{n-1}-t_m})]^2. \tag{32}$$

Since $t_{n-1} - t_m \leqslant \log \frac{n-2}{m-1}$, we get with Lemma 2.1 and Lemma 2.7 that

$$\mathbb{E}[f(Z_{t_{n-1}-t_m})] + \mathbb{E}[f(\mathcal{Z}[m, n-1])] \leqslant C \Big(\frac{n}{m}\Big)^{\gamma^+}.$$

Recalling that $n > m \geqslant 2$, it is now straightforward to deduce the statement from equations (31) and (32). It remains to consider the case where $1 = m < n$. Here, we apply Lemma 2.1 and $t_{n-1} \geqslant \log(n-1)$ to deduce that

$$\begin{aligned}\lambda &\leqslant \int_{-\infty}^{-t_N+t_1} e^{-(v-u)}\, \mathbb{E}[f(Z_{v-u})]\, du \\ &\leqslant C \int_{t_{n-1}}^{\infty} e^{-(1-\gamma^+)u}\, du \leqslant \frac{C}{1-\gamma^+}\,(n-1)^{\gamma^+-1},\end{aligned}$$

while, by Lemma 2.7, $\mathbb{P}(\Delta\mathcal{Z}[1, n-1] = 1) \leqslant f(0)\,(n-1)^{\gamma^+-1}$, so that a Poiss($\lambda$) distributed random variable can be coupled with $\Delta\mathcal{Z}[1, n-1]$ so that they disagree with probability less than a constant multiple of n^{γ^+-1}. □

REMARK 6.4. Lemma 6.3 provides a coupling for the mechanisms with which both trees produce *left* descendants. Since the number of descendants in individual locations form an independent sequence of random variables, we can apply the coupling of the lemma sequentially for each location and obtain a coupling of the π_N-projected left descendants of a vertex v and the left descendants of $n := \pi_N(v)$ in $\mathbb{T}[v]$. Indeed, under the assumptions of Lemma 6.3, one finds a coupling of both processes such that

$$\mathbb{P}(\text{left descendants disagree}) \leqslant C_{6.3}\,\frac{1}{n} + C_{6.3}\,\frac{1}{n^{1-\gamma^+}} \sum_{m=1}^{n-1} \frac{1}{m^{1+\gamma^+}} \leqslant C_{6.4} \frac{1}{n^{1-\gamma^+}},$$

where $C_{6.4}$ is a suitable positive constant.

Coupling the evolution to the right for particles of type $\tau \neq \ell$. We fix $v \leqslant 0$ and $N \in \mathbb{N}$, and suppose that $m := \pi_N(v) \geqslant 2$. Also fix a type $\tau < -v$ with $l := \pi_N(v+\tau) > m$. The cumulative sum of π_N-projected right descendants of a vertex v of type τ (including its predecessor) is distributed according to $(Z_{-t_N+t_n-v} \colon m \leqslant n \leqslant N)$ conditioned on $\Delta Z_\tau = 1$. The cumulative sum of right descendants in $\mathbb{T}[v]$ of a vertex in m of type l (including the predecessor) is distributed according to the law of $(\mathcal{Z}[m, n] \colon m \leqslant n \leqslant N)$ conditioned on $\Delta\mathcal{Z}[m, l-1] = 1$. Both processes are Markov processes and we provide a coupling of their transition probabilities.

LEMMA 6.5. *There exists a constant* $C_{6.5} > 0$ *such that the following holds: Let* $k \geqslant 0$, $m, n \geqslant 1$ *be integers with* $k+1 < m < n$, *and let* $\tau \in (t_n - t_m, t_{n+1} - t_m]$. *Then the random variables* $Z_{\Delta t_m}$ *under* $\mathbb{P}^k(\,\cdot\,|\,\Delta Z_\tau = 1)$ *and* $\mathcal{Z}[m, m+1]$ *under* $\mathbb{P}^k(\,\cdot\,|\,\Delta\mathcal{Z}[m,n] = 1)$ *can be coupled such that the resulting random variables* $\Upsilon^{(1)}$ *and* $\Upsilon^{(2)}$ *satisfy*

$$\mathbb{P}(\Upsilon^{(1)} \neq \Upsilon^{(2)}) \leqslant C_{6.5} \Big(\frac{f(k)}{m}\Big)^2.$$

PROOF. We couple $\Upsilon^{(1)}$ and $\Upsilon^{(2)}$ by plugging a uniform random variable on $(0,1)$ in the generalised inverses of the respective distribution functions and conclude that

$$\mathbb{P}(\Upsilon^{(1)} \neq \Upsilon^{(2)}) = |\mathbb{P}(\Upsilon^{(1)} = k) - \mathbb{P}(\Upsilon^{(2)} = k)| + \mathbb{P}(\Upsilon^{(1)} \geqslant k+2).$$

The second error term is of the required order since, by Lemma 2.5,

$$\mathbb{P}(\Upsilon^{(1)} \geqslant k+2) \leqslant \mathbb{P}^{k+1}(Z_{1/m} \geqslant k+3) \leqslant \Big(\frac{f(k+2)}{m}\Big)^2.$$

It remains to analyse the first error term. We have

$$\mathbb{P}(\Upsilon^{(2)} = k) = 1 - f(k)\Delta t_m \frac{P_{m+1,n} f(k+1)}{P_{m,n} f(k)},$$

and, representing $(Z_t^{[\tau]} : t \geqslant 0)$ by its compensator,

$$\mathbb{P}(\Upsilon^{(1)} = k) = \exp\Big\{-f(k)\int_0^{\Delta t_m} \frac{P_{\tau-u} f(k+1)}{P_{\tau-u} f(k)}\, du\Big\}.$$

We need to compare

$$\frac{P_{m+1,n} f(k+1)}{P_{m,n} f(k)} \quad \text{and} \quad \frac{P_u f(k+1)}{P_u f(k)} \quad \text{for } u \in [t_n - t_{m+1}, t_{n+1} - t_m].$$

By Lemma 2.1 and Proposition 2.14, one has, for $a \in \{k, k+1\}$ and sufficiently large m,

$$\begin{aligned} P_u f(a) &\leqslant P_{t_{n+1}-t_m} f(a) \leqslant e^{\gamma^+(\frac{1}{m}+\frac{1}{n})} P_{t_n - t_{m+1}} f(a) \\ &\leqslant e^{\gamma^+(\frac{1}{m}+\frac{1}{n})} \big(1 + C_{2.14} \tfrac{f(a)}{m}\big)\, P_{m+1,n} f(a) \\ &\leqslant e^{\gamma^+(\frac{1}{m}+\frac{1}{n}) + C_{2.14}\frac{f(a)}{m}} P_{m+1,n} f(a). \end{aligned}$$

Conversely,

$$\begin{aligned} P_u f(a) &\geqslant P_{t_n - t_{m+1}} f(a) \geqslant e^{-\gamma^+ \frac{1}{m}} P_{t_n - t_m} f(a) \\ &\geqslant e^{-\gamma^+ \frac{1}{m}} \big(1 - C_{2.14} \tfrac{f(a)}{m}\big)\, P_{m,n} f(a). \end{aligned}$$

We only need to consider large m and we may assume that $C_{2.14}\frac{f(k+1)}{m}\leqslant\frac12$, as otherwise we may choose $C_{6.5}$ large to ensure that the right hand side in the display of the lemma exceeds one. Then

$$P_u f(a)\geqslant e^{-\gamma^+\frac1m-2C_{2.14}\frac{f(a)}{m}}P_{m,n}f(a),$$

since $e^{-2y}\leqslant 1-y$ for $y\in[0,1/2]$. Consequently,

$$\begin{aligned}e^{-\gamma^+(2\frac1m+\frac1n)-3C_{2.14}\frac{f(k+1)}{m}}\,\frac{P_{m+1,n}f(k+1)}{P_{m,n}f(k)}&\leqslant\frac{P_uf(k+1)}{P_uf(k)}\\&\leqslant e^{\gamma^+(2\frac1m+\frac1n)+3C_{2.14}\frac{f(k+1)}{m}}\,\frac{P_{m+1,n}f(k+1)}{P_{m,n}f(k)}.\end{aligned}$$

Recall that, by Lemma 2.2, $\frac{P_{m+1,n}f(k+1)}{P_{m,n}f(k)}$ is uniformly bounded over all k so that we arrive at

$$\frac{P_{m+1,n}f(k+1)}{P_{m,n}f(k)}-C\,\frac{f(k)}{m}\leqslant\frac{P_uf(k+1)}{P_uf(k)}\leqslant\frac{P_{m+1,n}f(k+1)}{P_{m,n}f(k)}+C\,\frac{f(k)}{m},$$

for an appropriate constant $C>0$. Therefore,

$$\begin{aligned}&\mathbb{P}(\Upsilon^{(1)}=k)-\mathbb{P}(\Upsilon^{(2)}=k)\\&\qquad\leqslant 1\wedge\exp\big\{-f(k)\Delta t_m\big(\tfrac{P_{m+1,n}f(k+1)}{P_{m,n}f(k)}-C\,\tfrac{f(k)}{m}\big)\big\}\\&\qquad\qquad\qquad\qquad-\big(1-f(k)\Delta t_m\tfrac{P_{m+1,n}f(k+1)}{P_{m,n}f(k)}\big)\\&\qquad\leqslant C\,\big(\tfrac{f(k)}{m}\big)^2+\frac12\big(f(k)\Delta t_m\big(\tfrac{P_{m+1,n}f(k+1)}{P_{m,n}f(k)}-C\,\tfrac{f(k)}{m}\big)\big)^2\leqslant C_{6.5}\big(\tfrac{f(k)}{m}\big)^2.\end{aligned}$$

Similarly, one finds that

$$\mathbb{P}(\Upsilon^{(2)}=k)-\mathbb{P}(\Upsilon^{(1)}=k)\leqslant C_{6.5}\,\big(\tfrac{f(k)}{m}\big)^2,$$

and putting everything together yields the assertion. □

From Lemma 6.5 we get the following analogue of Lemma 6.2.

LEMMA 6.6. *Fix a level $T\in\mathbb{N}$. For any $v\leqslant 0$ and $\tau\leq -v$ with $\pi_N(v)=m\in\{2,3,\dots,N\}$ and $m<l:=\pi_N(v+\tau)$ we can couple the processes $(Z_{t_n-t_N-v}\colon n\geqslant m)$ conditioned on $\Delta Z_\tau=1$ and $(\mathcal{Z}[m,n]\colon n\geqslant m)$ conditioned on $\Delta\mathcal{Z}[m,l-1]=1$ such that the coupled processes $(\mathcal{Y}^{(1)}[n]\colon n\geqslant m)$ and $(\mathcal{Y}^{(2)}[n]\colon n\geqslant m)$ satisfy*

$$\mathbb{P}(\mathcal{Y}^{(1)}[n]\neq\mathcal{Y}^{(2)}[n]\textit{ for some }n\leqslant\sigma)\leqslant C_{6.6}\,(f(T)^2+1)\,\frac1m,$$

where σ is the first time when one of the processes reaches or exceeds level T.

PROOF. We define the process $\mathcal{Y} = ((\mathcal{Y}^{(1)}[n], \mathcal{Y}^{(2)}[n]) \colon n \geqslant m)$ to be the Markov process with starting distribution $\mathcal{L}(Z_{t_m - t_N - v} | \, \Delta Z_\tau = 1) \otimes \delta_0$ and transition kernels $p^{(n)}$ such that the first and second marginal are the conditioned transition probabilities of $(Z_{t_n - t_N - v} \colon n \geqslant m)$ and $(\mathcal{Z}[m,n] \colon n \geqslant m)$ as stated in the lemma. In the case where $n < l-1$, we demand that, for any integer $a \geq 0$, the law $p^{(n)}((a,a), \cdot\,)$ is the coupling of the laws of $Z_{\Delta t_n}$ under $\mathbb{P}^a(\,\cdot\,|\,\Delta Z_{\tau-(t_n - t_N - v)} = 1)$ and $\mathcal{Z}[n, n+1]$ under $\mathbb{P}^a(\,\cdot\,|\,\Delta\mathcal{Z}[n, l-1] = 1)$ provided in Lemma 6.5. Conversely, we apply the unconditioned coupling of Lemma 6.2 for $n \geq l$. Letting ϱ denote the first time when both evolutions disagree, we get

$$\begin{aligned}\mathbb{P}(\varrho \leqslant \sigma) &= \sum_{n=m}^{\infty} \mathbb{P}(\sigma \geqslant n, \varrho = n) \\ &\leqslant \mathbb{P}(\varrho = m) + \sum_{n=m}^{\infty} \mathbb{P}\big(\varrho = n+1 \big| \, \sigma > n, \varrho > n\big)\end{aligned}$$

and, by Lemma 6.2 and Lemma 6.5,

$$\mathbb{P}\big(\varrho = n+1 |\, \sigma > n, \varrho > n\big) \leqslant C_{6.5} \Big(\frac{f(T)}{n}\Big)^2 \;\text{ for } n \in \{m, m+1, \dots\} \backslash \{l-1\}.$$

Moreover, $\mathbb{P}(\varrho = m) \leqslant \mathbb{P}^1(Z_{t_m - t_N - v} > 0) = 1 - e^{-(t_m - v) f(1)} \leqslant \frac{f(1)}{m-1}$ and $\mathbb{P}(\varrho = l | \varrho \geq l, \sigma \geq l) \leq \mathbb{P}^T(Z_{\Delta t_{l-1}} > T) \leq f(T) \frac{1}{m}$. Consequently,

$$\mathbb{P}(\varrho \leqslant \sigma) \leqslant \frac{f(1)}{m-1} + \frac{f(T)}{m} + C_{6.5}\, f(T)^2 \sum_{n=m}^{\infty} \frac{1}{n^2} \leqslant C_{6.6}\, (f(T)^2 + 1)\, \frac{1}{m}. \qquad \square$$

PROOF OF PROPOSITION 6.1. We couple the labelled tree $\mathbb{T}(V)$ and the π_N-projected INT, starting with a coupling of the position of the initial vertex V and $\pi_N(-X)$, which fails with probability going to zero, by Lemma A.2.

Again we apply the concept of an exploration process. As before we categorise vertices as *veiled*, if they have not yet been discovered, *active*, if they have been discovered, but if their descendants have not yet been explored, and *dead*, if they have been discovered and all their descendants have been explored. In one exploration step the leftmost active vertex is picked and its descendants are explored in increasing order with respect to the location parameter. We stop immediately once one of the events (A), (B) or (C) happens. Note that in that case the exploration of the last vertex might not be completed. Moreover, when coupling two explorations, we also stop in the adverse event (E) that the explored graphs disagree. In event (B), the parameters $(n_N)_{N\in\mathbb{N}}$ are chosen such that

$$\lim_{N\to\infty} \frac{\big(\log N \log\log N\big)^{\alpha}}{n_N} = 0 \quad \text{and} \quad \lim_{N\to\infty} \frac{\log n_N}{\log N} = 0,$$

for $\alpha := (1-\gamma^+)^{-1} \vee 3$. Noting that we never need to explore more than c_N vertices, we see from Lemma 6.2, Remark 6.4 and Lemma 6.6 that the probability of a failure of this coupling is bounded by a constant multiple of

$$c_N(1+f(c_N)^2)\,\frac{1}{n_N} + c_N\,\frac{1}{{n_N}^{1-\gamma^+}} \leqslant \frac{c_N^3}{n_N} + \frac{c_N}{{n_N}^{1-\gamma^+}} \longrightarrow 0.$$

Consequently, the coupling succeeds with high probability. As in Lemma 4.2 it is easy to see that, with high probability, event (B) implies that

$$\#\mathbb{T}(V) \geqslant c_N \quad \text{and } \#\mathfrak{T} \geqslant c_N.$$

Hence we have

$$\#\mathbb{T}(V) \wedge c_N = \#\mathfrak{T} \wedge c_N \qquad \text{with high probability,}$$

and the statement follows by combining this with Proposition 5.1. □

7. The variance of the number of vertices in large clusters. In this section we provide the second moment estimate needed to show that our key empirical quantity, the number of vertices in connected components of a given size, concentrate asymptotically near their mean.

PROPOSITION 7.1. *Suppose that $(c_N)_{N\in\mathbb{N}}$ and $(n_N)_{N\in\mathbb{N}}$ are sequences of integers satisfying $1 \leqslant c_N, n_N \leqslant N$ such that $c_N^2 n_N^{\gamma^+-1}$ is bounded from above. Then, for a constant $C_{7.1} > 0$ depending on these sequences and on f, we have*

$$\mathrm{var}\Big(\frac{1}{N}\sum_{v=1}^{N} \mathbb{1}\{\#\mathcal{C}_N(v) \geqslant c_N\}\Big)$$
$$\leqslant 2\,\mathbb{P}\big(\#\mathcal{C}_N(V) < c_N \text{ and } \mathcal{C}_N(V)\cap\{1,\dots,n_N\} \neq \emptyset\big) + \frac{c_N}{N} + C_{7.1}\,\frac{c_N^2}{n_N^{1-\gamma^+}},$$

where V is independent of $\mathcal{G}_N$ and uniformly distributed on $\{1,\dots,N\}$.

PROOF. Let v, w be two distinct vertices of $\mathcal{G}_N$. We start by exploring the neighbourhood of v similarly as in Section 5. As before we classify the vertices as veiled, active and dead, and in the beginning only v is active and the remaining vertices are veiled. In one exploration step we pick the leftmost active vertex and consecutively (from the left to the right) explore its immediate neighbours in the set of veiled vertices only. Newly found vertices are activated and the vertex to be explored is set to dead after the exploration. We immediately stop the exploration once one of the events

(A) the number of unveiled vertices in the cluster reaches c_N,
(B) one vertex in $\{1, \ldots, n_N\}$ is activated, or
(C) there are no more active vertices left,

happens. Note that when we stop due to (A) or (B) the exploration of the last vertex might not be finished. In that case we call this vertex *semi-active*.

We proceed with a second exploration process, namely the exploration of the cluster of w. This exploration follows the same rules as the first exploration process, treating vertices that remained active or semi-active at the end of the first exploration as veiled. In addition to the stopping in the cases (A), (B), (C) we also stop the exploration once a vertex is unveiled which was also unveiled in the first exploration, calling this event (D). We consider the following events:

E^v : the first exploration started with vertex v ends in (A) or (B);
$E_1^{v,w}$: w is unveiled during the first exploration (that of v);
$E_2^{v,w}$: w remains veiled in the first exploration and the second exploration ends in (A) or (B) but not in (D);
$E_3^{v,w}$: w remains veiled in the first exploration and the second exploration ends in (D).

We have

$$\begin{aligned}\sum_{v=1}^N \sum_{w=1}^N \mathbb{P}(\#\mathcal{C}_N(v) \geqslant c_N, \#\mathcal{C}_N(w) \geqslant c_N) &\leqslant \sum_{v=1}^N \sum_{w=1}^N \sum_{k=1}^3 \mathbb{P}(E^v \cap E_k^{v,w}) \\ &= \sum_{v=1}^N \mathbb{P}(E^v) \sum_{k=1}^3 \sum_{w=1}^N \mathbb{P}(E_k^{v,w} \mid E^v).\end{aligned} \tag{33}$$

As the first exploration immediately stops once one has unveiled c_N vertices, we conclude that, for fixed v,

$$\sum_{w=1}^N \mathbb{P}(E_1^{v,w} \mid E^v) = \mathbb{E}\Big[\sum_{w=1}^N \mathbb{1}_{E_1^{v,w}} \Big| E^v\Big] \leqslant c_N. \tag{34}$$

To analyse the remaining terms, we fix distinct vertices v and w and note that the configuration after the first exploration can be formally described by an element $\mathfrak{k}$ of

$$\{\text{open, closed, unexplored}\}^{E_N},$$

where $E_N := \{(a,b) \in \{1,\dots,N\}^2 : i < j\}$ denotes the set of possible edges. We pick a feasible configuration $\mathfrak{k}$ and denote by $\mathcal{E}_\mathfrak{k}$ the event that the first exploration ended in this configuration. On the event $\mathcal{E}_\mathfrak{k}$ the status of each vertex (veiled, active, semi-active or dead) at the end of the first exploration is determined. Suppose $\mathfrak{k}$ is such that w remained veiled in the first exploration, which means that $\mathcal{E}_\mathfrak{k}$ and $E_1^{v,w}$ are disjoint events. Next, we note that

$$\mathbb{P}(E_2^{v,w} \,|\, \mathcal{E}_\mathfrak{k}) \leqslant \mathbb{P}(E^w). \tag{35}$$

Indeed, if in the exploration of w we encounter an edge which is open in the configuration $\mathfrak{k}$, we have unveiled a vertex which was also unveiled in the exploration of v, the second exploration ends in (D) and hence $E_2^{v,w}$ does not happen. Otherwise, the event $\mathcal{E}_\mathfrak{k}$ influences the exploration of w only in the sense that in the degree evolution of some vertices some edges may be conditioned to be closed. By Lemma 2.9 this conditional probability is bounded by the unconditional probability and hence we obtain (35).

Finally, we analyse the probability $\mathbb{P}(E_3^{v,w} \,|\, \mathcal{E}_\mathfrak{k})$. If the second exploration process ends in state (D) we have discovered an edge connecting the exploration started in w to an active or semi-active vertex a from the first exploration. Recall that in each exploration we explore the immediate neighbourhoods of at most c_N vertices. Let $\mathfrak{K} \in E_N$ be a feasible configuration at the beginning of the neighbourhood exploration of a vertex $n > n_N$ and note that this implies every edge which is open (resp. closed) in $\mathfrak{k}$ is also open (resp. closed) in $\mathfrak{K}$. Recall that $\mathcal{E}_\mathfrak{K}$ denotes the event that this configuration is seen in the combined exploration processes. We denote by $\mathfrak{a}$ and $\mathfrak{s}$ the set of active and semi-active vertices of the *first exploration* induced by $\mathfrak{k}$ (or, equivalently, by $\mathfrak{K}$). Moreover, we denote by $\mathfrak{d}$ the set of dead vertices of the combined exploration excluding the father of n, and, for $a \in \mathfrak{a} \cup \mathfrak{s}$, we let $\mathfrak{d}_a$ denote the set of dead vertices of the ongoing exploration excluding the father of n, plus the vertices that were marked as dead in the first exploration at the time the vertex a was discovered. We need to distinguish several cases.

First, consider the case $a \in \mathfrak{a}$ with $a < n$. By definition of the combined exploration process, we know that a has no jumps in its indegree evolution at times associated to the vertices $\mathfrak{d}_a$. If a was explored from the right, say with father in b, we thus get

$$\mathbb{P}\big(\exists \text{ edge between } a \text{ and } n \big| \mathcal{E}_{\mathfrak{K}}\big) = \mathbb{P}\big(\Delta\mathcal{Z}[a, n-1] = 1 \\ \big| \Delta\mathcal{Z}[a, b-1] = 1 \text{ and } \Delta\mathcal{Z}[a, d-1] = 0 \, \forall d \in \mathfrak{d}_a\big). \tag{36}$$

If a was explored from the left, then

$$\mathbb{P}\big(\exists \text{ edge between } a \text{ and } n \big| \mathcal{E}_{\mathfrak{K}}\big) \\ = \mathbb{P}\big(\Delta\mathcal{Z}[a, n-1] = 1 \big| \Delta\mathcal{Z}[a, d-1] = 0 \, \forall d \in \mathfrak{d}_a\big). \tag{37}$$

Second, consider the case $a \in \mathfrak{a}$ with $n < a$. By definition of the combined exploration process, the indegree evolution of n has no jumps that can be associated to edges connecting to $\mathfrak{d}$. Hence, if n was explored from the right, say with father in b, then

$$\mathbb{P}\big(\exists \text{ edge between } a \text{ and } n \big| \mathcal{E}_{\mathfrak{K}}\big) = \mathbb{P}\big(\Delta\mathcal{Z}[n, a-1] = 1 \\ \big| \Delta\mathcal{Z}[n, b-1] = 1 \text{ and } \Delta\mathcal{Z}[n, d-1] = 0 \, \forall d \in \mathfrak{d}\big), \tag{38}$$

and, if n was explored from the left, then

$$\mathbb{P}\big(\exists \text{ edge between } a \text{ and } n \big| \mathcal{E}_{\mathfrak{K}}\big) \\ = \mathbb{P}\big(\Delta\mathcal{Z}[n, a-1] = 1 \big| \Delta\mathcal{Z}[n, d-1] = 0 \, \forall d \in \mathfrak{d}\big). \tag{39}$$

Third, consider $a \in \mathfrak{s}$ and denote by a' the last vertex which was unveiled in the first exploration. If $a' > n$ then the existence of an edge between a and n was already explored in the first exploration and no edge was found. If $a' < n < a$, we find estimates (38), (39) again. If $a < n$ and the father b of a satisfies $b > a' \vee a$,

$$\mathbb{P}\big(\exists \text{ edge between } a \text{ and } n \big| \mathcal{E}_{\mathfrak{K}}\big) \leqslant \sup_{0 \leq k \leq c_N - 1} \mathbb{P}^k\big(\Delta\mathcal{Z}[a \vee a', n-1] = 1 \, | \\ \Delta\mathcal{Z}[a \vee a', b-1] = 1 \text{ and } \Delta\mathcal{Z}[a \vee a', d-1] = 0 \, \forall d \in \mathfrak{d}_a\big), \tag{40}$$

and if $a = v$ or the father b of $a \vee a'$ satisfies $b < a \vee a'$,

$$\mathbb{P}\big(\exists \text{ edge between } a \text{ and } n \big| \mathcal{E}_{\mathfrak{K}}\big) \\ \leqslant \sup_{0 \leq k \leq c_N} \mathbb{P}^k\big(\Delta\mathcal{Z}[a \vee a', n-1] = 1 \big| \Delta\mathcal{Z}[a \vee a', d-1] = 0 \, \forall d \in \mathfrak{d}_a\big). \tag{41}$$

Using first Lemma 2.12, then Lemma 2.10 and Lemma 2.11 we see that the terms (36)–(39) are bounded by

$$C_{2.12} \, \mathbb{P}^1\big(\Delta\mathcal{Z}[a, n-1] = 1\big) \leqslant C_{2.12} \, \frac{P_{1,n_N} f(1)}{n_N},$$

and similarly, the terms (40)–(41) are bounded by

$$C_{2.12}\,\mathbb{P}^{c_N}\big(\Delta\mathcal{Z}[a,n-1]=1\big)\leqslant C_{2.12}\,\frac{P_{1,n_N}f(c_N)}{n_N}.$$

Note that there are at most c_N vertices $a\in\mathfrak{a}\cup\mathfrak{s}$ and at most one of those is semi-active. For each of these a we have to test the existence of edges no more than c_N times. Hence, using also Lemma 2.7 and the boundedness of $f(n)/n$, we find $C_{7.1}>0$ such that

$$\mathbb{P}\big(E_3^{v,w}\big|E^v\big)\leqslant C_{2.12}\,c_N^2\,\frac{P_{1,n_N}f(1)}{n_N}+C_{2.12}\,c_N\,\frac{P_{1,n_N}f(c_N)}{n_N}\leqslant C_{7.1}\,\frac{c_N^2}{n_N^{1-\gamma^+}}.$$

Summarising our steps, we have

$$\begin{aligned}
&\operatorname{var}\Big(\frac{1}{N}\sum_{v=1}^N\mathbb{1}\{\#\mathcal{C}_N(v)\geqslant c_N\}\Big)\\
&\leqslant\mathbb{E}\Big[\frac{1}{N^2}\sum_{v=1}^N\sum_{w=1}^N\mathbb{1}\{\#\mathcal{C}_N(v)\geqslant c_N,\#\mathcal{C}_N(w)\geqslant c_N\}\Big]\\
&\qquad-\frac{1}{N^2}\sum_{v=1}^N\sum_{w=1}^N\mathbb{P}(E^v)\mathbb{P}(E^w)\\
&\qquad+2\,\frac{1}{N}\sum_{v=1}^N\mathbb{P}\big(\#\mathcal{C}_N(v)<c_N\text{ and }\mathcal{C}_N(v)\cap\{1,\dots,n_N\}\neq\emptyset\big)\\
&\leqslant 2\,\mathbb{P}\big(\#\mathcal{C}_N(V)<c_N\text{ and }\mathcal{C}_N(V)\cap\{1,\dots,n_N\}\neq\emptyset\big)+\frac{c_N}{N}+C_{7.1}\,\frac{c_N^2}{n_N^{1-\gamma^+}},
\end{aligned}$$

as required to complete the proof. □

8. Proof of Theorem 1.8. We start by proving the lower bound for $\mathcal{C}_N^{(1)}$. Suppose therefore that $p(f)>0$, fix $\delta>0$ arbitrarily small and use Lemma 3.4 to choose $\varepsilon>0$ such that the survival probability of $\bar f=f-\varepsilon$ is larger than $p(f)-\delta$. We denote by $(\bar{\mathcal{G}}_N)_{N\in\mathbb{N}}$ a sequence of random networks with attachment rule $\bar f$ and let $\bar{\mathcal{C}}_N(v)$ the connected component of v in $\bar{\mathcal{G}}_N$. Suppose a vertex V is chosen uniformly at random from $\{1,\dots,N\}$. We choose $c_N:=\big\lfloor\log N\sqrt{\log\log N}\big\rfloor$ and observe that by Proposition 6.1

$$\begin{aligned}
\mathbb{E}\Big[\frac{1}{N}\sum_{v=1}^N\mathbb{1}\{\#\bar{\mathcal{C}}_N(v)\geqslant c_N\}\Big]&=\mathbb{P}\{\#\bar{\mathcal{C}}_N(V)\geqslant c_N\}\\
&\longrightarrow\mathbb{P}\{\#\mathfrak{T}=\infty\}\geqslant p(f)-\delta,
\end{aligned}\tag{42}$$

as N tends to infinity. By Proposition 7.1 with $n_N := \lfloor (\log N)^{\frac{4}{1-\gamma^+}} \rfloor$, we have

$$\begin{aligned} &\operatorname{var}\Big(\frac{1}{N}\sum_{v=1}^{N} \mathbb{1}\{\#\bar{\mathcal{C}}_N(v) \geqslant c_N\}\Big) \\ &\quad \leqslant 2\,\mathbb{P}\big(\#\bar{\mathcal{C}}_N(V) < c_N \text{ and } \bar{\mathcal{C}}_N(V) \cap \{1,\dots,n_N\} \neq \emptyset\big) + \frac{c_N}{N} + C_{7.1}\,\frac{c_N^2}{n_N^{1-\gamma^+}}. \end{aligned}$$

The first summand goes to zero by Lemma 4.2 and so do the remaining terms by the choice of our parameters. Hence

$$\liminf_{N\to\infty} \frac{1}{N}\sum_{v=1}^{N} \mathbb{1}\{\#\bar{\mathcal{C}}_N(v) \geqslant c_N\} \geqslant p(f) - \delta \qquad \text{in probability,}$$

and Proposition 4.1 implies that, with high probability, there exists a connected component comprising at least a proportion $p(f)$ of all vertices, proving the lower bound.

To see the upper bound we work with the original attachment function f. In analogy to (42) we obtain

$$\lim_{N\to\infty} \mathbb{E}\Big[\frac{1}{N}\sum_{v=1}^{N} \mathbb{1}\{\#\mathcal{C}_N(v) \geqslant c_N\}\Big] = p(f).$$

As in the lower bound, the variance goes to zero, and hence we have

$$\lim_{N\to\infty} \frac{1}{N}\sum_{v=1}^{N} \mathbb{1}\{\#\mathcal{C}_N(v) \geqslant c_N\} = p(f), \qquad \text{in probability.}$$

From this we infer that, in probability,

$$\limsup_{N\to\infty} \frac{\#\mathcal{C}_N^{(1)}}{N} \leqslant \limsup_{N\to\infty} \frac{c_N}{N} \vee \Big(\frac{1}{N}\sum_{v=1}^{N} \mathbb{1}\{\#\mathcal{C}_N(v) \geqslant c_N\}\Big) \leqslant p(f),$$

proving the upper bound.

Finally, to prove the result on the size of the second largest connected component, note that we have seen in particular that

$$\lim_{N\to\infty} \frac{1}{N}\sum_{v=1}^{N} \mathbb{1}\{\#\mathcal{C}_N(v) \geqslant c_N\} = p(f), \qquad \text{in probability,}$$

so that, with high probability, the proportion of vertices in clusters of size greater or equal c_N is asymptotically equal to the proportion of vertices in the giant component. This implies that the proportion of vertices, which are not in the giant component but in components of size at least c_N goes to zero in probability, which is a stronger result than the stated claim.

9. Proof of Theorem 1.9. We fix $k \in \mathbb{N}$ and choose $c_N := k+1$. By Proposition 6.1, we have

$$\lim_{N\to\infty} \frac{1}{N}\,\mathbb{E}\Big[\sum_{v=1}^{N} \mathbb{1}\{\#\mathcal{C}_N(v) \leqslant k\}\Big] = \lim_{N\to\infty} \mathbb{P}(\#\mathcal{C}_N(V) \leqslant k) = \mathbb{P}(\#\mathfrak{T} \leqslant k)$$

and Proposition 7.1 yields

$$\mathrm{var}\Big(\frac{1}{N}\mathbb{1}\{\#\mathcal{C}_N(v) \leqslant k\}\Big) = \mathrm{var}\Big(\frac{1}{N}\mathbb{1}\{\#\mathcal{C}_N(v) \geqslant c_N\}\Big) \to 0.$$

This implies the statement, as k is arbitrary.

10. Proof of Theorem 1.6. The equivalence of the divergence of the sequence in Theorem 1.6 and the criterion $\mathcal{I} = \emptyset$ stated in (i) of Remark 1.7 follows from the bounds on the spectral radius of the operators A_α given in the proof of Proposition 1.10. Moreover, it is easy to see from the arguments of Section 3 that the survival of the INT under percolation with retention parameter p is equivalent to the existence of $0 < \alpha < 1$ such that

$$\rho(pA_\alpha) = p\rho(A_\alpha) \leqslant 1.$$

Hence, to complete the proof of Theorem 1.6 and Remark 1.7, it suffices to show that, for a fixed retention parameter $0 < p < 1$, the existence of a giant component for the percolated network is equivalent to the survival of the INT under percolation with retention parameter p. We now give a sketch of this by showing how the corresponding arguments in the proof of Theorem 1.8 have to be modified.

As in the proof of Theorem 1.8 the main part of the argument consists of couplings of the exploration process of the neighbourhood of a vertex in the network to increasingly simple objects. To begin with we have to couple the exploration of vertices in the percolated network and the percolated labelled tree, using arguments as in Section 5. We only modify the exploration processes a little: Whenever we find a new vertex, instead of automatically declaring it active, we declare it *active* with probability p and *passive* otherwise. We do this independently for each newly found vertex. We still explore at every step the leftmost active vertex, but we change the stopping criterion (E1): we now stop the process when we rediscover an active *or passive* vertex. We also stop the process when we have discovered more than $2\frac{1-p}{p}\,c_N$ passive vertices, calling this event (E3). All other stopping criteria are retained literally.

By a simple application of the strong law of large numbers we see that the probability of stopping in the event (E3) converges to zero. The proof of

Lemma 5.2 carries over to our case, as it only uses that the number of dead, active and passive vertices is bounded by a constant multiple of c_N. Hence the coupling of explorations is successful with high probability.

Similarly, the coupling of the exploration processes for the random labelled tree and the idealised neighbourhood tree constructed in Section 6 can be performed so that under the assumption on the parameters given in Proposition 6.1, we have

$$\#\mathcal{C}_N^*(V) \wedge c_N = \#\mathfrak{T}^* \wedge c_N \quad \text{with high probability,}$$

where $\mathcal{C}_N^*(v)$ denotes the connected component in the percolated network, which contains the vertex v, and $\mathfrak{T}^*$ is the percolated INT.

In order to analyse the variance of the number of vertices in large clusters of the percolated network we modify the exploration processes described in the proof of Proposition 7.1 a little: In the first exploration we activate newly unveiled vertices with probability p and declare them passive otherwise. We always explore the neighbourhood of the leftmost active vertex and investigate its links to the set of veiled *or passive* vertices from left to right, possibly activating a passive vertex when it is revisited. We stop the exploration in the events (A), (B), and (C) as before, and additionally if the number of passive vertices exceeds $2\frac{1-p}{p}\,c_N$, calling this event (A'). As before, the probability of stopping in (A') goes to zero by the strong law of large numbers.

The exploration of the second cluster follows the same rules as that of the first, treating vertices that were left active, semi-active or passive in the first exploration as veiled. In addition to the stopping events (A), (A'), (B) and (C) we also stop in the event (D) when a vertex is unveiled which was also unveiled in the first exploration. This vertex may have been active, semi-active or passive at the end of the first exploration. We then introduce the event E^v that the first exploration ends in events (A), (A') or (B), events $E_1^{v,w}$ and $E_3^{v,w}$ as before, and event $E_2^{v,w}$ that w remained veiled in the first exploration and the second exploration ends in (A), (A') or (B). We can write

$$\sum_{v=1}^N \sum_{w=1}^N \mathbb{P}(\#\mathcal{C}'_N(v) \geqslant c_N, \#\mathcal{C}'_N(w) \geqslant c_N) \leqslant \sum_{v=1}^N \mathbb{P}(E^v) \sum_{k=1}^3 \sum_{w=1}^N \mathbb{P}(E_k^{v,w} \mid E^v),$$

where $\mathcal{C}'_N(v)$ denotes the connected component of v in the percolated network. The summand corresponding to $k = 1$ can be estimated as before. For the other summands we describe the configuration after the first exploration as an element $\mathfrak{k}$ of

$$\{\text{open, closed, removed, unexplored}\}^{E_N},$$

where edges corresponding to the creation of passive vertices are considered as 'removed'. We again obtain that $\mathbb{P}(E_2^{v,w} | \mathcal{E}_{\mathfrak{k}}) \leqslant \mathbb{P}(E^w)$ using the fact that if in the second exploration we ever encounter an edge which is open or removed in the configuration $\mathfrak{k}$ the second exploration ends in (D) and $E_2^{v,w}$ does not occur. Finally, the estimate of $\mathbb{P}(E_3^{v,w} | \mathcal{E}_{\mathfrak{k}})$ carries over to our situation as it relies only on the fact that the number of unveiled vertices in the first exploration is bounded by a constant multiple of c_N. We thus obtain a result analogous to Proposition 7.1.

Using straightforward analogues of the results in Section 4 we can now show that the existence of a giant component for the percolated network is equivalent to the survival of the INT under percolation with retention parameter p using the argument of Section 8. This completes the proof of Theorem 1.6.

APPENDIX

In this appendix we provide two auxiliary coupling lemmas.

LEMMA A.1. *Let $\lambda \geqslant 0$ and $p \in [0,1]$, $X^{(1)}$ Poisson distributed with parameter λ, and $X^{(2)}$ Bernoulli distributed with parameter p. Then there exists a coupling of these two random variables such that*

$$\mathbb{P}(X^{(1)} \neq X^{(2)}) \leqslant \lambda^2 + |\lambda - p|.$$

PROOF. We only need to consider the case where $\lambda \in [0,1]$. Then $X^{(1)}$ can be coupled to a Bernoulli distributed random variable X with parameter λ, such that $\mathbb{P}(X^{(1)} \neq X) = \lambda - \lambda e^{-\lambda} \leqslant \lambda^2$. Moreover, X and $X^{(2)}$ can be coupled such that $\mathbb{P}(X \neq X^{(2)}) = |p - \lambda|$. The two facts together imply the statement. □

LEMMA A.2. *Let Y be standard exponentially distributed and X uniformly distributed on $\{1, \dots, N\}$. Then X and Y can be coupled in such a way that*

$$\mathbb{P}\big(X \neq \pi_N(-Y)\big) \leqslant C_{A.2} \frac{\log N}{N},$$

for the function π_N defined at the beginning of Section 6.

PROOF. For $2 \leqslant k \leqslant N$ we have

$$\begin{aligned}\mathbb{P}(\pi_N(-Y) = k) &= \mathbb{P}\Big(\sum_{j=k}^{N-1} \frac{1}{j} \leqslant Y < \sum_{j=k-1}^{N-1} \frac{1}{j} \Big) \\ &= \exp\Big\{ - \sum_{j=k}^{N-1} \frac{1}{j} \Big\} - \exp\Big\{ - \sum_{j=k-1}^{N-1} \frac{1}{j} \Big\}.\end{aligned}$$

Since $\sum_{j=k-1}^{N-1} \frac{1}{j} \geqslant \log \frac{N}{k-1}$ and $e^x \leqslant 1+x+x^2$ for $x \in [1,2]$, we get

$$\begin{aligned}\mathbb{P}(\pi_N(-Y)=k) &= \exp\Big\{-\sum_{j=k-1}^{N-1} \frac{1}{j}\Big\}\big(e^{\frac{1}{k-1}}-1\big)\\ &\leqslant \frac{k-1}{N}\big(\frac{1}{k-1}+\frac{1}{(k-1)^2}\big) \leqslant \frac{1}{N}+\frac{2}{Nk}.\end{aligned}$$

Similarly, one obtains that $\mathbb{P}(\pi_N(-Y)=k) \geqslant \frac{1}{N} - \frac{2}{Nk}$. Hence we can couple the random variables so that, for a suitable constant $C_{A.2} > 0$,

$$\mathbb{P}\big(X \neq \pi_N(-Y)\big) \leqslant \sum_{k=2}^{N} \big|\mathbb{P}(\pi_N(-Y)=k) - \tfrac{1}{N}\big| \leqslant C_{A.2}\,\frac{\log N}{N}. \qquad \square$$

Acknowledgements: We would like to thank Christian Mönch and two referees for heplful comments on the first version of this paper.

References.

[BA99] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.

[BBCS09] N. Berger, C. Borgs, J. T. Chayes, and A. Saberi. A weak local limit for preferential attachment graphs. *Preprint*, 2009.

[BJR05] B. Bollobás, S. Janson, and O. Riordan. The phase transition in the uniformly grown random graph has infinite order. *Random Structures Algorithms*, 26:1–36, 2005.

[BJR07] B. Bollobás, S. Janson, and O. Riordan. The phase transition in inhomogeneous random graphs. *Random Structures Algorithms*, 31:3–122, 2007.

[BR03] B. Bollobás and O. Riordan. Robustness and vulnerability of scale-free random graphs. *Internet Math.*, 1:1–35, 2003.

[BRST01] B. Bollobás, O. Riordan, J. Spencer and G. Tusnády. The degree sequence of a scale-free random graph process. *Random Structures Algorithms*, 18:279–290, 2001.

[BS01] I. Benjamini and O. Schramm. Recurrence of distributional limits of finite planar graphs. *Electron. J. Probab.*, 6:Paper 23, 2001.

[DHH10] S. Dommers, R. van der Hofstad, and G. Hooghiemstra. Diameters in preferential attachment models. *J. Stat. Phys.*, 139:72–107, 2010.

[DM09] S. Dereich and P. Mörters. Random networks with sublinear preferential attachment: degree evolutions. *Electron. J. Probab.*, 14:1222–1267, 2009.

[HH09] R. Hardy and S.C. Harris. A spine approach to branching diffusions with applications to L^p-convergence of martingales. *Séminaire de Probabilités XLII*, pp 281–330, 2009.

[Hof09] R. van der Hofstad. *Random graphs and complex networks.* Book in preparation, see `http://www.win.tue-nl/~rhofstad/NotesRGCN2010.pdf`, 2009.

[Kat76] T. Kato. *Perturbation theory for linear operators.* Springer, Berlin, 1976.

[Kat82] T. Kato. Superconvexity of the spectral radius, and convexity of the spectral bound and the type. *Math. Zeitschrift*, 180:265–274, 1982.

[KR01] P. L. Krapivsky and S. Redner. Organization of growing random networks. *Phys. Rev. E*, 63:Paper 066123, 2001.

[KRS01] A. Kyprianou and A. Rahimzadeh Sani. Martingale convergence and the functional equation in the multi-type branching random walk. *Bernoulli*, 7:593–604, 2001.

[Pin95] R. G. Pinsky. *Positive harmonic functions and diffusion*. Cambridge University Press, Cambridge, 1995.

[Ram83] A. G. Ramm. Variational principles for eigenvalues of compact nonselfadjoint operators. II. *J. Math. Anal. Appl.*, 91(1):30–38, 1983.

[She89] L. A. Shepp. Connectedness of certain random graphs. *Israel J. Math.*, 67:23–33, 1989.

Fachbereich Mathematik und Informatik
Philipps–Universität Marburg
Hans–Meerwein Strasse
35032 Marburg
Germany
E-mail: dereich@mathematik.uni-marburg.de
URL: http://www.mathematik.uni-marburg.de/ dereich/

Department of Mathematical Sciences
University of Bath
Claverton Down
Bath BA2 7AY
United Kingdom
E-mail: maspm@bath.ac.uk
URL: http://people.bath.ac.uk/maspm/