The Annals of Applied Probability
2020, Vol. 30, No. 1, 40–79
https://doi.org/10.1214/19-AAP1494

LOCAL WEAK CONVERGENCE FOR PAGERANK

By Alessandro Garavaglia[1,*], Remco van der Hofstad[1,**] and Nelly Litvak[2]

[1]Department of Mathematics and Computer Science, Eindhoven University of Technology, [*]a.garavaglia@tue.nl; [**]r.v.d.hofstad@tue.nl

[2]Department of Applied Mathematics, University of Twente, n.litvak@utwente.nl

PageRank is a well-known algorithm for measuring centrality in networks. It was originally proposed by Google for ranking pages in the World Wide Web. One of the intriguing empirical properties of PageRank is the so-called 'power-law hypothesis': in a scale-free network, the PageRank scores follow a power law with the same exponent as the (in-)degrees. To date, this hypothesis has been confirmed empirically and in several specific random graphs models. In contrast, this paper does not focus on one random graph model but investigates the existence of an asymptotic PageRank distribution, when the graph size goes to infinity, using local weak convergence. This may help to identify general network structures in which the power-law hypothesis holds. We start from the definition of local weak convergence for sequences of (random) undirected graphs, and extend this notion to directed graphs. To this end, we define an exploration process in the directed setting that keeps track of in- and out-degrees of vertices. Then we use this to prove the existence of an asymptotic PageRank distribution. As a result, the limiting distribution of PageRank can be computed directly as a function of the limiting object. We apply our results to the directed configuration model and continuous-time branching processes trees, as well as to preferential attachment models.

1. Introduction and main results.

1.1. *Definition of PageRank.* PageRank, first introduced in [49], is an algorithm that generates a centrality measure on finite graphs. Originally introduced to rank World Wide Web pages, PageRank has a wide range of applications including citation analysis [23, 47, 53], community detection [5] or social networks analysis [12, 54].

Consider a finite directed (multi-)graph G of size n. We write $[n] = \{1, \dots, n\}$. Let $e_{j,i}$ be the number of directed edges from j to i. Denote the in-degree of vertex $i \in [n]$ by $d_i^{(\mathrm{in})}$ and the out-degree by $d_i^{(\mathrm{out})}$. Fix a parameter $c \in (0, 1)$, which is called the *damping factor*, or teleportation parameter. PageRank is the unique vector $\boldsymbol{\pi}(n) = (\pi_1(n), \dots, \pi_n(n))$ that satisfies, for every $i \in [n]$,

$$\pi_i(n) = c \sum_{j \in [n]} \frac{e_{j,i}}{d_j^{(\mathrm{out})}} \pi_j(n) + \frac{1-c}{n}. \tag{1}$$

PageRank has the natural interpretation of being the invariant measure of a random walk with restarts on G. With probability c, the random walk takes a simple random walk step on G, while with probability $(1 - c)$ it moves to a uniformly chosen vertex. Here, by simple random walk we mean the random walk that chooses, at every step, an outgoing edge from the current position uniformly at random. When $d_j^{(\mathrm{out})} > 0$ for all $j \in [n]$, then the invariant measure of

Received March 2018; revised February 2019.
MSC2010 subject classifications. Primary 60B20; secondary 05C80.
Key words and phrases. Directed random graphs, PageRank, local weak convergence.

this random walk is given exactly by (1). The interpretation is easily extended to the case when some vertices j have $d_j^{(\text{out})} = 0$ by introducing a random jump from such vertices; in this case, the stationary distribution will be the solution of (1) renormalized to sum up to one [44].

In this paper, we consider the *graph-normalized version* of PageRank, which is the vector defined as $\boldsymbol{R}(n) = n\boldsymbol{\pi}(n)$. We call both the algorithm and the vector $\boldsymbol{R}(n)$ PageRank, the meaning will always be clear from the context. The graph-normalized version of (1) is the unique solution $\boldsymbol{R}(n)$ to

$$R_i(n) = c \sum_{j\in[n]} \frac{e_{j,i}}{d_j^{(\text{out})}} R_j(n) + (1-c). \tag{2}$$

PageRank has numerous generalizations. For example, after a random jump, the random walk might not restart from a uniformly chosen vertex, but rather choose vertex i with probability b_i, where $\sum_{i=1}^n b_i = 1$. Equation (1) then becomes

$$R_i(n) = c \sum_{j\in[n]} \frac{e_{j,i}}{d_j^{(\text{out})}} R_j(n) + (1-c)nb_i. \tag{3}$$

This generalized version of PageRank is sometimes called *topic-sensitive* [36] or *personalized*. We note that the term *personalized PageRank* often refers to the case when the vector $\mathbf{b} = (b_1, \dots, b_n)$ has one of its coordinates equal to 1, and the rest equal to zero, so that the random walk always restarts from the same vertex. One can generalize further, for example, allow the probability c to be random as well. The literature [20, 43, 45, 52] usually studies the following graph-normalized equation:

$$R_i(n) = \sum_{j:e_{j,i}\geq 1} A_j R_j(n) + B_i, \quad i \in [n], \tag{4}$$

where $(A_i)_{i\in[n]}$ and $(B_i)_{i\in[n]}$ are values assigned to the vertices in the graph. In this paper, for simplicity of the argument, we will focus on the basic model (2), and then in Section 5, extend the results to the more general model (4) with $A_j = C_j/d_j^{(\text{out})}$, where C_j's are random variables bounded by $c < 1$, and $(B_i)_{i\in[n]}$ are i.i.d. across vertices.

1.2. *Power-law hypothesis for PageRank.* It has been observed [46, 50] that in real-world networks with power-law (in-)degree distributions, PageRank follows a power-law with the same exponent. Figure 1 illustrates this phenomenon in citation networks. Empirical studies suggest that the power-law hypothesis holds rather generally. However, proving this appears to be challenging. Some progress has been made in [10] for the average PageRank in a preferential attachment model. In a series of papers [46, 52], the result was proved provided that PageRank satisfies a branching-type of recursion. Then, in fact, a network is modeled as a branching process, with independent labels representing out-degrees. The full proof for PageRank defined on finite random graphs has been obtained in [20] for the directed configuration model, and recently in [45] for directed generalized random graphs.

The motivation of this paper is in finding general conditions for the existence of an asymptotic PageRank distribution. We prove the convergence (in some sense) of PageRank for a large class of models. Our results also shed light on the power-law hypothesis. Indeed, when the limit is a branching tree, this directly implies the power-law hypothesis, based on the above mentioned results in the literature. When the limit is different, for example, the tree generated by a continuous-time branching process, proving the power-law hypothesis remains an open problem. Our results imply however that it is sufficient to study PageRank on the limiting object, which hopefully is simpler since the graph-size asymptotics no longer interfere.

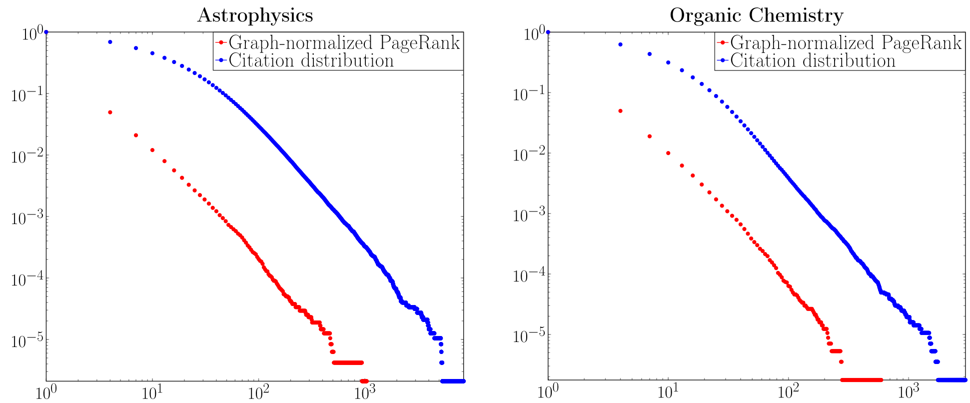

FIG. 1. *Citation networks from* Web of Science. *Citation networks can be seen as directed graphs where references are directed edges. We considered papers from Astrophysics* (*left*) *and Organic Chemistry* (*right*). *The two loglog scale plots show two different distributions each. The blue data represent the tail distribution of the in-degree* (*so number of citations*) *of a uniformly chosen vertex. The red data represents the tail distribution of the graph-normalized PageRank of a uniform vertex. Notice that, in both cases, the two distributions show a remarkably similar power-law exponent.*

1.3. *Overview of the paper.* In Section 2, we explain our general methodology and the main ideas behind the proofs and the results. In Section 2.1, we introduce the notion of local weak convergence, which is crucial in our approach, and explain how we extend it to directed graphs. Section 2.2 describes the major steps in proving weak convergence for PageRank. In Section 2.3, we present three examples that illustrate our results: directed configuration models (DCMs) (and the extension to directed inhomogeneous random graphs), continuous-time branching processes (CTBPs) and (directed) preferential attachment models (DPAs). In Section 2.4, we list some open problems.

Sections 3–6 contain formal proofs. In Section 3, we explain local weak convergence for undirected graph sequences (Section 3.1) and introduce our construction for directed graph sequences (Section 3.2), which is tailored to our PageRank application. In Section 4, we formally prove the main result for the standard definition of PageRank (2). In Section 5, we extend our main result to the generalized PageRank (4). The examples, that is, DCM (and inhomogeneous random graphs), CTBPs and DPA, are analyzed, respectively, in Sections 6.1 (and 6.2), 6.3 and 6.4.

2. Main result and methodology. Note that for any deterministic graph, PageRank is a deterministic vector. We are interested in the PageRank associated to *random* graphs. In particular, we want to investigate the asymptotic behavior of the PageRank value of a uniformly chosen vertex V_n, as the size of the graph grows. In this case, we have two sources of randomness: the choice of the vertex and the randomness of the graph itself. Our main result shows that, for a nice enough sequence of directed graphs $(G_n)_{n\in\mathbb{N}}$, $R_{V_n}(n)$ converges in distribution to a limiting random variable:

THEOREM 2.1 (Existence of asymptotic PageRank distribution). *Consider a sequence of directed random graphs* $(G_n)_{n\in\mathbb{N}}$. *Then the following hold*:

(1) *If* G_n converges in distribution in the local weak sense, *then there exists a limiting distribution* $R_\varnothing$, *with* $\mathbb{E}[R_\varnothing] \le 1$, *such that*

$$R_{V_n}(n) \xrightarrow{d} R_\varnothing;$$

(2) *If* G_n converges in probability in the local weak sense, *then there exists a limiting distribution* $R_\varnothing$, *with* $\mathbb{E}[R_\varnothing] \le 1$, *such that, for every* $r > 0$,

$$\frac{1}{n}\sum_{i\in[n]} \mathbb{1}\{R_i(n) > r\} \xrightarrow{\mathbb{P}} \mathbb{P}(R_\varnothing > r).$$

Theorem 2.1 establishes that, whenever a sequence of directed random graphs converges in the *local weak sense*, then the distribution $R_{V_n}(n)$ admits a limit in distribution, $R_\varnothing$. This limit has the interpretation of PageRank on the (possibly infinite) limiting graph. Theorem 2.1 can be extended to personalized PageRank defined in (4) under additional conditions on the random variables $(A_i)_{i\in\mathbb{N}}$ and $(B_i)_{i\in\mathbb{N}}$. The precise formulation that requires more notation is given in Theorem 5.4.

REMARK 2.2 (Stochastic lower bound for PageRank). Theorem 2.1 gives a rough lower bond on the tail of the asymptotic PageRank distribution for a graph sequence. In simple words, we can write

$$R_\varnothing \ge (1-c)\left(1 + c\sum_{i=1}^{D_\varnothing^{(\mathrm{in})}} \frac{1}{m_i^{(\mathrm{out})}}\right), \tag{5}$$

where $\varnothing$ is a vertex called root in the local weak limit of the graph sequence $(G_n)_{n\in\mathbb{N}}$, $D_\varnothing^{(\mathrm{in})}$ is the graph limiting in-degree distribution, and $m_i^{(\mathrm{out})}$ represent the out-degree in the LW limit. All the notation in (5) is introduced in Sections 3.2 and 4. In particular, (5) implies that $R_\varnothing > 1 - c$ a.s.. Since $m^{(\mathrm{out})}$ represents the limiting out-degree distribution, it follows that, if $(G_n)_{n\in\mathbb{N}}$ has out-degrees uniformly bounded by a constant $A < \infty$,

$$R_\varnothing \ge (1-c)\left(1 + \frac{c}{A} D_\varnothing^{(\mathrm{in})}\right).$$

As a consequence, if the limiting in-degree distribution obeys a power law, then the tail of the distribution $R_\varnothing$ is bounded from below by a multiple of the tail of the in-degree. This establishes a power-law lower bound for $R_\varnothing$. This is a partial solution of the power-law hypothesis mentioned in Section 1.2.

We next explain the ingredient of our main result, which is local weak convergence.

2.1. *Local weak convergence for directed graphs.* Local weak (LW) convergence is a concept that was first introduced in [3, 4, 14] for undirected graphs. In this framework, a sequence of undirected random graphs, under relatively weak conditions, converges to a (possibly random) *rooted graph*, that is, a graph where one of the vertices is labeled as root. In simple words, the limiting graph resembles the *neighborhood of a typical vertex* in the graph sequence. This methodology has been shown to be useful to investigate *local* properties of a graph sequence—the properties that depend on the *local neighborhood* of vertices.

In the literature, limits of different types of random graphs have been investigated (Aldous and Steele give a survey in [3]). Grimmett [35] obtained the LW limit for the uniform random tree. Generalized random graphs [17, 24, 25, 31] also converge in the LW sense under some regularity conditions on the weight distribution. Convergence of undirected configuration model is proved in [38], Chapter 2. In many random graph contexts, the LW limit is a branching process, and LW convergence provides a method to compare neighborhoods in random graphs to branching processes.

A recent work by Berger et al. [15] investigates the LW limit for preferential attachment models, in the case of a fixed number of edges and no self-loops allowed. In particular, their proof covers the case of a power-law distribution with exponent $\tau \geq 3$. Dereich and Morters [27–29] establish the LW limit in the case of preferential attachment models with conditionally independent edges.

In the local weak convergence setting, a sequence of graphs $(G_n)_{n\in\mathbb{N}}$ converges to a (possibly) random rooted graph $(G, \varnothing)$ that is a *rooted graph* $(G, \varnothing)$. Here, $\varnothing \in V(G)$ denotes the root.

Heuristically, $G_n \to (G, \varnothing)$ in the LW sense when the law of the neighborhood of a typical vertex in G_n converges to the law of the neighborhood of the root in G. We give now an intuitive formulation of this concept (for a precise definition, see Section 3.1). For a vertex i in a graph G_n, denote the neighborhood of i up to distance k by $U_{\leq k}(i)$. Then, for a random rooted graph $(G, \varnothing)$, we say that $G_n \to (G, \varnothing)$ if, for any finite rooted graph (H, y), and any $k \in \mathbb{N}$,

$$\frac{1}{n}\sum_{i\in[n]} \mathbb{1}\{U_{\leq k}(i) \cong (H, y)\} \longrightarrow \mathbb{P}\big(U_{\leq k}(\varnothing) \cong (H, y)\big), \tag{6}$$

where $\mathbb{1}\{\cdot\}$ is an indicator of event $\{\cdot\}$, and $U_{\leq k}(\varnothing)$ is the k-neighborhood of $\varnothing$ in G. The event $\{U_{\leq k}(i) \cong (H, y)\}$ means that the k-neighborhood of i is structured as (H, y), ignoring the precise labeling of the vertices. Notice that the left-hand term in (6) is just the probability that the k-neighborhood of a uniformly chosen vertex in G_n is structured as (H, y).

Equation (6) is formulated for a deterministic graph sequence $(G_n)_{n\in\mathbb{N}}$. When $(G_n)_{n\in\mathbb{N}}$ is a sequence of *random* graphs, the left-hand term in (6) is a random variable. In this case, there are different modes of convergence, as stated in Definition 3.6. For example, we say that $G_n \to (G, \varnothing)$ in probability if, for any finite rooted graph (H, y), and any $k \in \mathbb{N}$,

$$\frac{1}{n}\sum_{i\in[n]} \mathbb{1}\{U_{\leq k}(i) \cong (H, y)\} \xrightarrow{\mathbb{P}} \mathbb{P}\big(U_{\leq k}(\varnothing) \cong (H, y)\big). \tag{7}$$

In Section 3.2, we will extend the definition of LW convergence to directed graphs. Here, we introduce the main ideas behind the construction.

The major problem in the construction is that the *exploration of neighborhoods* is not uniquely defined in directed graphs. Indeed, in the exploration process (rigorous definition is given in Definition 3.10), motivated by the PageRank problem, we naturally explore directed edges *only in their opposite direction*. In other words, a directed edge (j, i) is only explored from i to j. Clearly, since edges are not explored in both directions, starting from the root we might not be able to explore all the graph. Heuristically, from the point of view of the root $\varnothing$, *only part of the graph has influence on the incoming neighborhood of* $\varnothing$. This is very different from the undirected case, where the exploration process continues until the entire graph is explored (when the graph is connected). We resolve this by introducing so-called marks to track the explored and not-explored out-edges in the graph. The precise definition of LW convergence in directed graphs is given in Section 3.2.

We point out that our construction is one of many possible ways to define LW convergence for directed graphs. For instance, Aldous and Steele [4] allow edge weights. This might be sufficient to define an inclusion of directed graphs in the space of undirected graphs with edge weights, and use the notion for undirected graphs to define an exploration process for directed graphs. The advantage of our construction is that it requires the minimum amount of information, sufficient to prove the convergence of PageRank, which is the main problem we aim to resolve.

Definition 3.11 below, together with Remark 3.12, gives a criterion for the convergence of a sequence of directed random graphs that can be presented as marked graphs by just assigning *marks equal to out-degrees*. The precise formulation requires heavy notation that we have not introduced yet; therefore, we do not state it here.

The advantage of having a LW limit $(G, \varnothing)$ is that a whole family of local properties of the graph sequence *can pass to the limit*, and the limit is given by a local property of $(G, \varnothing)$ itself. More precisely, in the construction of LW convergence, one defines a *distance* between (marked directed) rooted graphs (see Definition 3.3). Then, any function f from the space of rooted graphs to $\mathbb{R}$ that is *bounded and continuous* with respect to the distance function can pass to the limit, that is, for V_n a uniformly chosen vertex in G_n,

$$\lim_{n\to\infty} \mathbb{E}\big[f(G_n, V_n)\big] = \mathbb{E}\big[f(G, \varnothing)\big].$$

This can be rather useful in understanding the asymptotic behavior of local properties of a graph sequence. As a toy example, in the undirected setting, take the function $f(G, \varnothing) = \mathbb{1}\{d_\varnothing = k\}$. It is easy to show, using Definition 3.3, that f is a continuous function. Assume that a sequence of graphs $G_n \to (G, \varnothing)$ locally weakly, where $(G, \varnothing)$ is random rooted graph. Then, for every $n \in \mathbb{N}$,

$$\mathbb{E}\big[f(G_n, V_n)\big] = \frac{1}{n} \sum_{i\in[n]} \mathbb{P}(d_i = k) = \mathbb{P}(d_{V_n} = k),$$

that is, f evaluated on a random root is just the probability that a uniformly chosen vertex has degree k. As a consequence, the sequence $(G_n)_{n\in\mathbb{N}}$ has a limiting degree distribution given by

$$\lim_{n\to\infty} \mathbb{P}(d_{V_n} = k) = \mathbb{P}(d_\varnothing = k),$$

where $\varnothing$ is the root of G. Other examples of continuous functions in the undirected setting are the nearest-neighbor average degree of a uniform vertex, the finite-distance neighborhood of a uniform vertex and the average pressure per particle in the Ising model. In our directed setting, it follows that, if $G_n \to (G, \varnothing, M(G))$,

$$\big(m_{V_n}^{(\mathrm{out})}, d_{V_n}^{(\mathrm{in})}\big) \xrightarrow{d} \big(m_\varnothing^{(\mathrm{out})}, d_\varnothing^{(\mathrm{in})}\big),$$

where $M(G)$ is the set of marks of the limiting graph, $(m_{V_n}^{(\mathrm{out})}, d_{V_n}^{(\mathrm{in})})$ are the mark and the in-degree of a uniformly chosen vertex V_n, and $(m_\varnothing^{(\mathrm{out})}, d_\varnothing^{(\mathrm{in})})$ are the mark and the in-degree of the root $\varnothing$ in the limiting directed graph G. The notation $m^{(\mathrm{out})}$ hints on the relation between marks and out-degrees. When marks are assigned that are equal to the out-degree, this implies the convergence of the in- and out-degree of a uniformly chosen vertex. One of the surprises in our version of LW convergence is that in the limiting graph, the mark of the root $m_\varnothing^{(\mathrm{out})}$ is not necessarily equal to the out-degree of the root.

2.2. *Application of LW convergence to PageRank.* The proof of Theorem 2.1 is given in Section 4. Here, we describe the structure of the proof, explaining why the LW convergence for directed graphs is useful. Schematically, the structure of our proof of Theorem 2.1 is presented in Figure 2. The implication (A), denoted by the dashed red arrow, is the one we aim to prove. We split it in three steps (a), (b), (c), denoted by the solid black arrows. We will now explain each step.

Step (*a*): *Finite approximations.* It is well known [5, 10, 16, 20] that PageRank can be written as

$$R_i(n) = (1-c)\left(1 + \sum_{k=1}^{\infty} c^k \sum_{\ell\in\mathrm{path}_i(k)} \prod_{h=1}^{k} \frac{e_{\ell_h,\ell_{h+1}}}{d_{\ell_h}^{(\mathrm{out})}}\right),$$

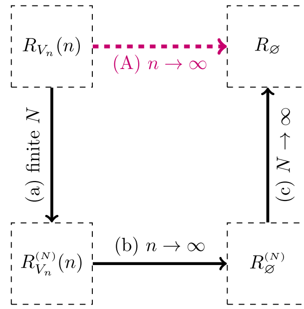

FIG. 2. *Structure of the proof of Theorem 2.1. The (A) convergence is what we are after, the convergence in distribution of $R_{V_n}(n)$ to a limiting random variable. To prove that, we need the three different steps (a), (b), (c) given by the other arrows.*

where $\mathrm{path}_i(k)$ is the set of directed paths of k steps that end at i. In other words, $R_i(n)$ is a *weighted sum of all the directed paths that end at i*. In particular, we can write finite approximations for PageRank as

$$R_i^{(N)}(n) = (1-c)\left(1 + \sum_{k=1}^{N} c^k \sum_{\ell \in \mathrm{path}_i(k)} \prod_{h=1}^{k} \frac{e_{\ell_h,\ell_{h+1}}}{d^{(\mathrm{out})}_{\ell_h}}\right),$$

where now the sum is taken over all paths of length at most $N \in \mathbb{N}$. We use the sequence of finite approximations $(R^{(N)}_{V_n}(n))_{n\in\mathbb{N}}$ to estimate the PagreRank of a random vertex with *exponentially small precision* by its finite approximations. We prove that, for any $\varepsilon > 0$,

$$\mathbb{P}\big(\big|R_{V_n}(n) - R^{(N)}_{V_n}(n)\big| \geq \varepsilon\big) \leq \frac{c^{N+1}}{\varepsilon}.$$

Notice that the bound is *independent of the graph size that we consider*. This bound is true for any directed graph of any size, so it does not require any assumption on the graph sequence.

Step (*b*): *LW convergence*. The finite approximations of PageRank are *continuous* with respect to the local weak topology. Furthermore, by definition, the Nth approximation of PageRank depends only on the incoming neighborhood of a vertex *up to distance N*. Note that $R_i(n)$ and $R_i^{(N)}(n)$ are not bounded. However, for any $r \geq 0$, the function $\mathbb{1}\{R^{(N)}_{V_n} > r\}$ is a continuous and bounded function on marked directed rooted graphs, therefore, we can pass to the limit for any $N \in \mathbb{N}$. It follows that, if G_n converges LW (in some sense),

$$\lim_{n\to\infty} \mathbb{E}\big[\mathbb{1}\{R^{(N)}_{V_n} > r\}\big] = \lim_{n\to\infty} \mathbb{P}\big(R^{(N)}_{V_n}(n) > r\big) = \mathbb{P}\big(R^{(N)}_\varnothing > r\big),$$

where in the last term $\varnothing$ is the root of the limiting random marked directed rooted graph $(G, \varnothing, M(G))$. As a consequence, every term of the sequence $(R^{(N)}_{V_n}(n))_{n\in\mathbb{N}}$ converges in distribution. Notice that similar arguments apply for Theorem 2.1(b).

Step (*c*): *Finite approximations on the limiting graph*. On the limiting random marked directed rooted graph $(G, \varnothing, M(G))$, the sequence $(R^{(N)}_\varnothing)_{N\in\mathbb{N}}$ is a monotonically increasing sequence of random variables. Therefore, there exists an almost sure limiting random variable $R_\varnothing$. Using the fact that $(G, \varnothing, M(G))$ is a local weak limit of a sequence of random directed graphs, and $\mathbb{E}[R_{V_n}] = 1$ for every $n \geq 1$, it is possible to prove that $\mathbb{E}[R_\varnothing] \leq 1$, so that $\mathbb{P}(R_\varnothing < \infty) = 1$.

REMARK 2.3. We emphasize that the above strategy is meant just to give the intuition behind the proof. In particular, in the proof it is necessary to be careful and specify with respect to which randomness we take expectations. In fact, when we consider local weak convergence of random graphs, we have two sources of randomness: the choice of the root and the randomness of the graphs. All these are made rigorous in Section 4.

2.3. *Examples*. We consider examples of directed random graphs, for which we prove LWC and find the limiting random graph. Thus, PageRank in these models converges to PageRank on the limiting graph. The following theorem makes this precise for several random graph models that have been studied in the literature. For precise definitions of the models, as well as the proof, we refer to Section 6.

THEOREM 2.4 (Examples of convergence). *The following models converge in the directed local weak sense*:

(1) *the* directed configuration model *and the* directed inhomogeneous random graph *converge* in probability;
(2) continuous-time branching processes *converge* almost surely;
(3) *the* directed preferential attachment model *converges* in probability.

As a consequence, *for these models there exists a limiting PageRank distribution*, *and the convergence holds as specified*.

REMARK 2.5 (Power-law lower bound). The directed preferential attachment model and continuous-time branching processes both have constant out-degree. Therefore, they satisfy the condition in Remark 2.2. Thus, their limiting PageRank distributions are stochastically bounded from below by a multiple of the limiting in-degree distributions. The directed configuration model satisfies Remark 2.2 whenever the out-degree distribution has bounded support.

We point out that the directed configuration model and the directed inhomogeneous random graph are listed together since the two graphs are closely related, as explained in Section 6.2.

The proof of Theorem 2.4 is divided into three propositions, respectively Proposition 6.2 for directed configuration model, Proposition 6.6 for continuous-time branching processes and Proposition 6.10 for directed preferential attachment model.

2.4. *Open problems*.

Extension to exploration of outgoing edges. In this paper, we extend the definition of local weak convergence to directed graphs. Moved by the interest in PageRank algorithms on random graphs, we build our definition on the *exploration of incoming edges* in their opposite direction, that is, an edge (i, j) is explored from i to j. The outgoing edges are considered as marks and we do not explore them. In the same way, it is possible to define the exploration process according to the natural direction of the edges. In this case, we consider *outgoing neighborhoods* instead. The definition of LW convergence would just be a consequence of symmetry. This second interpretation might be useful, for instance, in the study of diffusion processes on graphs, such as epidemic spread. An interesting and more complex extension would be to explore the incoming and outgoing neighborhoods *at the same time*.

PageRank on limiting graphs. We are able to prove that, under relatively general assumptions, a sequence of random directed graphs admits a limiting distribution for the PageRank of a uniformly chosen vertex. In this way, we have moved the analysis of a graph's PageRank distribution from a whole sequence of graphs to a single (possibly infinite) rooted directed marked graph. Note that we prove the existence of such distribution, but we do not always have a convenient description of it. It will be interesting to investigate the behavior of this limiting distribution. In particular, it is interesting to investigate the conditions under which the rank of the root in the limiting graph shows a *power-law tail*, and thus confirm the power-law hypothesis.

The remainder of the paper provides formal proofs of what has been discussed above.

3. Local weak convergence.

3.1. *Preliminaries*: *LWC of undirected graphs.* We present the definition of LWC for undirected graphs first, since the construction for directed graphs is similar. We start by defining what a rooted graph is.

DEFINITION 3.1 (Rooted graph). Let G be a locally finite graph with vertex set $V(G)$ (finite or countable), and edge set $E(G)$. Fix a vertex $\varnothing \in G$ and call it the *root*. The pair $(G, \varnothing)$ is called a *rooted graph*.

We are not interested in the labeling of the vertices, but only in the graph structure. For this, we define isomorphisms between rooted graphs as follows.

DEFINITION 3.2 (Isomorphism). An isomorphism between two rooted graphs $(G, \varnothing)$ and $(G', \varnothing')$ is a bijection $\gamma : V(G) \to V(G')$ such that:

(1) $(j, i) \in E(G)$ if and only if $(\gamma(j), \gamma(i)) \in E(G)$;
(2) $\gamma(\varnothing) = \varnothing'$.

We write $(G, \varnothing) \cong (G', \varnothing')$ to denote that $(G, \varnothing)$ and $(G', \varnothing')$ are isomorphic rooted graphs.

Denote the space of all rooted graphs (up to isomorphisms) by $\mathcal{G}_\star$. Formally, $\mathcal{G}_\star$ is the quotient space of the set of all locally finite rooted graphs with respect to the equivalence relation given by isomorphisms.

For a rooted graph $(G, \varnothing) \in \mathcal{G}_\star$, we let $U_{\le k}(\varnothing)$ denote the subgraph of G of all vertices at graph distance at most k away from $\varnothing$. Formally, this means that $U_{\le k}(\varnothing) = (V(U_{\le k}(\varnothing)), E(U_{\le k}(\varnothing)))$, where

$$V\big(U_{\le k}(\varnothing)\big) = \big\{i : d_G(i, \varnothing) \le k\big\}, \qquad E\big(U_{\le k}(\varnothing)\big) = \big\{\{j, i\} \colon j, i \in V\big(U_{\le k}(\varnothing)\big)\big\}.$$

We call $U_{\le k}(\varnothing)$ the k-neighborhood around $\varnothing$. We use this notion to define the distance between two rooted graphs.

DEFINITION 3.3 (Local distance). The function $d_{\mathrm{loc}}((G, \varnothing), (G', \varnothing')) = 1/(1 + \kappa)$, where

$$\kappa = \inf_{k \ge 1} \big\{U_{\le k}(\varnothing) \ncong U_{\le k}\big(\varnothing'\big)\big\},$$

is called the *local distance* on the space of rooted graphs $\mathcal{G}_\star$.

It is possible to prove that d_{loc} is an actual distance on the space of rooted graphs. In particular, the space $(\mathcal{G}_\star, d_{\mathrm{loc}})$ is a Polish space (see [32], Appendix A, for the proof for an equivalent definition of a distance). The function d_{loc} measures how distant two rooted graphs are *from the point of view of the root*. In many graphs though, there is no vertex that can be naturally chosen as a root, for instance in configuration models or Erdős–Rényi random graph. For this reason, it is useful to choose the root *at random*. Define, for any graph G,

$$\mathcal{P}(G) = \frac{1}{n} \sum_{i \in [n]} \delta_{(G,i)}. \tag{8}$$

Given a graph G of size n, $\mathcal{P}(G)$ is a probability measure that chooses the root uniformly at random among the n vertices. When we consider a sequence of graphs $(G_n)_{n\in\mathbb{N}}$, we denote $\mathcal{P}(G_n)$ simply by $\mathcal{P}_n$. With this notion, we are ready to define LWC for undirected deterministic graphs.

DEFINITION 3.4 (Local weak convergence). Consider a deterministic sequence of locally finite graphs $(G_n)_{n\in\mathbb{N}}$. We say that $(G_n)_{n\in\mathbb{N}}$ *converges in the local weak sense* to a (possibly) random element $(G, \varnothing)$ of $\mathcal{G}_\star$ with law $\mathcal{P}$, if, for any bounded continuous function $f : \mathcal{G}_\star \to \mathbb{R}$,

$$\mathbb{E}_{\mathcal{P}_n}[f] \longrightarrow \mathbb{E}_{\mathcal{P}}[f],$$

where $\mathbb{E}_{\mathcal{P}_n}$ and $\mathbb{E}_{\mathcal{P}}$ denote the expectation with respect to $\mathcal{P}_n$ and $\mathcal{P}$, respectively.

In particular, this means that the probability converges over open sets of the topology. Fix (H, y) finite, then

$$\begin{aligned} B_R(H, y) &= \{(G, \varnothing) \in \mathcal{G}_\star : d_{\mathrm{loc}}((H, y), (G, \varnothing)) \le R\} \\ &= \{(G, \varnothing) \in \mathcal{G}_\star : U_{\le \lfloor 1/R \rfloor}(\varnothing) \cong (H, y)\}. \end{aligned} \tag{9}$$

Elements in this open ball are determined by the neighborhood of the root up to distance $\lfloor 1/R \rfloor$. As a consequence, the probability $\mathcal{P}_n$ of the ball $B_R(H, y)$ is given by

$$\mathcal{P}_n(B_R(H, y)) = \frac{1}{n} \sum_{i\in[n]} \mathbb{1}\{U_{\le\lfloor 1/R\rfloor}(i) \cong (H, y)\}.$$

This implies that it suffices to look at the local structure of the neighborhood of a typical vertex to obtain the probability $\mathcal{P}_n$ of any open ball. We now state a criterion for a sequence of deterministic graphs to converge in the LW sense as in Definition 3.4:

THEOREM 3.5 (Criterion for local weak convergence). *Let $(G_n)_{n\in\mathbb{N}}$ be a sequence of graphs. Then G_n converges in the local weak sense to $(G, \varnothing)$ with law $\mathcal{P}$ when, for every finite rooted graph (H, y),*

$$\mathcal{P}_n(H) = \frac{1}{n} \sum_{i\in[n]} \mathbb{1}\{U_{\le k}(i) \cong (H, y)\} \longrightarrow \mathcal{P}(U_{\le k}(\varnothing) \cong (H, y)). \tag{10}$$

The proof Theorem 3.5 can be found in [38], Section 1.4. Notice that for $(H, y) \in \mathcal{G}_\star$, the functions $\mathbb{1}\{U_{\le k}(\varnothing) \cong (H, y)\}$ are continuous with respect to the local weak topology and uniquely identify the limit.

So far we have considered sequences of deterministic graphs. Whenever we consider a *random graph* G_n, we have two sources of randomness. First, we have the randomness of the choice of the root, and then the randomness of the graph itself. For this reason, it is necessary to specify the randomness we take expectation with respect to, giving rise to different ways of convergence. We specify this in the following definition.

DEFINITION 3.6 (Local weak convergence). Consider a sequence of random graphs $(G_n)_{n\in\mathbb{N}}$, and a probability $\mathcal{P}$ on $\mathcal{G}_\star$. Denote by $\mathcal{P}_n$ the probability associated to G_n as in (8).

(1) We say that G_n *converges in distribution in the local weak sense* to $\mathcal{P}$ if, for any bounded continuous function $f : \mathcal{G}_\star \to \mathbb{R}$,

$$\mathbb{E}\big[\mathbb{E}_{\mathcal{P}_n}[f]\big] \longrightarrow \mathbb{E}_{\mathcal{P}}[f]; \tag{11}$$

(2) We say that G_n *converges in probability in the local weak sense* to $\mathcal{P}$ if, for any bounded continuous function $f : \mathcal{G}_\star \to \mathbb{R}$,

$$\mathbb{E}_{\mathcal{P}_n}[f] \xrightarrow{\mathbb{P}} \mathbb{E}_{\mathcal{P}}[f]; \tag{12}$$

(3) We say that G_n *converges almost surely in the local weak sense* to $\mathcal{P}$ if, for any bounded continuous function $f : \mathcal{G}_\star \to \mathbb{R}$,

$$\mathbb{E}_{\mathcal{P}_n}[f] \xrightarrow{\mathbb{P}\text{-a.s.}} \mathbb{E}_{\mathcal{P}}[f]. \tag{13}$$

Notice that the left-hand term in (12) is a random variable, while the right-hand side is deterministic. In fact, (12) implies (11), but the opposite is not true. Similarly, (13) implies (12).

Similar to Theorem 3.5, we can give a criterion for the convergence of a sequence of random graphs.

THEOREM 3.7 (Criterion for local weak convergence of random graphs). *Let $(G_n)_{n\in\mathbb{N}}$ be a sequence random graphs. Let $(G, \varnothing)$ be a random variable on $\mathcal{G}_\star$ having law $\mathcal{P}$. Then G_n converges to $(G, \varnothing)$ in distribution (in probability, almost surely) if* (11) ((12), (13), *resp.*) *holds for every function of the type $\mathbb{1}\{U_{\leq k}(\varnothing) \cong (H, y)\}$, where $k \in \mathbb{N}$ and (H, y) is a finite element of $\mathcal{G}_\star$.*

The proof of Theorem 3.7 follows immediately from Theorem 3.5.

3.2. *Directed graphs.* The construction of local weak convergence for directed graphs is similar to the undirected case. It is necessary though to define an exploration process to construct the neighborhood of the root and keep track of in- and out-degrees of vertices. To keep notation as simple as possible, we use the same notation as in Section 3.1, while here we refer to directed graphs. We start giving the definition of rooted marked directed graphs.

DEFINITION 3.8 (Rooted marked directed graph). Let G be a directed graph with vertex set $V(G)$ and edge set $E(G)$. Let $\varnothing \in V(G)$ be a vertex called the *root*. Assume that for every $i \in V(G)$, the in-degree $d_i^{(\text{in})}$ and the out-degree $d_i^{(\text{out})}$ of the vertex i are finite. Assign to every $i \in V(G)$ an integer value $m_i^{(\text{out})}$ called a *mark*, such that $d_i^{(\text{out})} \leq m_i^{(\text{out})} < \infty$. Denote the set of marks by $M(G) = (m_i^{(\text{out})})_{i\in V(G)}$. We call the triplet $(G, \varnothing, M(G))$ a *rooted marked directed graph.*

To simplify the notation in Definition 3.8, we will specify the marks only when necessary. In simple words, a rooted marked directed graph is a locally finite directed graph where one of the vertices is marked as root, and to every vertex we assign a mark, which is larger than the out-degree of the vertex. If $m_i^{(\text{out})} = d_i^{(\text{out})}$, we keep i intact, and if $m_i^{(\text{out})} - d_i^{(\text{out})} > 0$ then we attach to i exactly $m_i^{(\text{out})} - d_i^{(\text{out})}$ outgoing arrows pointing nowhere. This is illustrated in Figure 3. We call a directed graph with marks, without specifying the root, a *marked graph*.

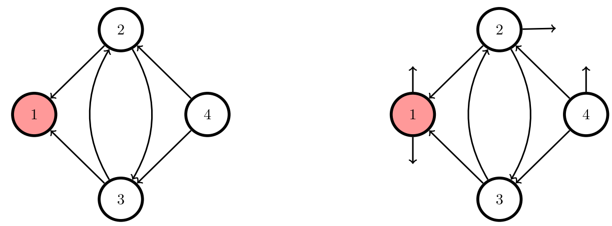

FIG. 3. *Two examples of rooted marked directed graphs. The graph on the left is considered with marks equal to the out-degree, while in the example on the right we have assigned marks larger than the out-degree. The difference between the mark and the out-degree of a vertex can be visualized as the number of arrows starting at the vertex and pointing nowhere.*

Every directed graph can be seen as a rooted marked directed graph, with marks equal to the out-degrees and a root picked from the set of vertices. In what follows, sometimes we specify the marks, and sometimes we specify the out-degree and the number of edges pointing nowhere.

As in the undirected case, we are not interested in the precise labeling of the vertices. This leads us to define the notion of isomorphism, including the presence of marks.

DEFINITION 3.9 (Isomorphism of rooted marked directed graphs). Two rooted marked directed graphs $(G, \varnothing, M(G))$ and $(G', \varnothing', M(G'))$ are *isomorphic* if and only if there exists a bijection $\gamma : V(G) \to V(G')$ such that:

(1) $(i, j) \in E(G)$ if and only if $(\gamma(i), \gamma(j)) \in E(G')$;
(2) $\gamma(\varnothing) = \varnothing'$;
(3) for every $i \in V(G)$, $m_i^{(\text{out})} = m_{\gamma(i)}^{(\text{out})}$.

We write $(G, \varnothing, M(G)) \cong (G', \varnothing', M(G'))$ to denote that $(G, \varnothing, M(G))$ and $(G', \varnothing', M(G'))$ are isomorphic rooted marked directed graphs.

Denote the space of rooted marked directed graphs by $\mathcal{G}_\star$, which is again a quotient space with respect to the equivalence given by isomorphisms. We now define the exploration process that identifies the neighborhood of the root; see Figure 4 for an example.

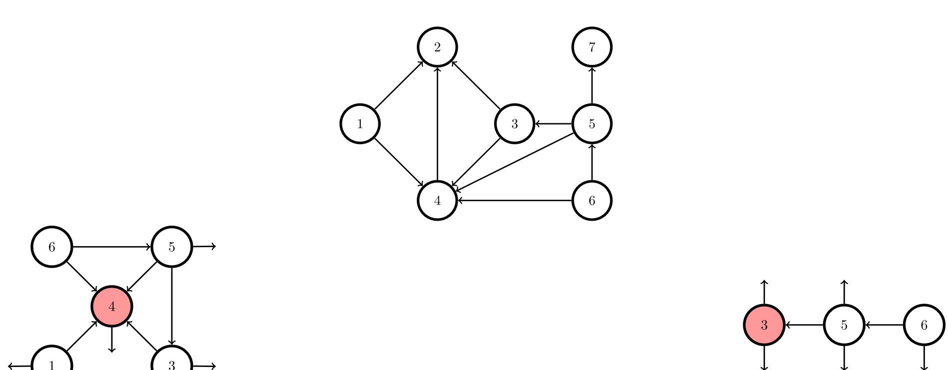

FIG. 4. *Example of two root neighborhoods in the same graph above, where we have assigned marks equal to the out-degrees, with a different choice of the root. The root on the left is vertex* 4, *and vertex* 3 *on the right. We explore the root neighborhood up to the maximum possible distance. Notice that the graph is only partially explored in this example.*

DEFINITION 3.10 (Root neighborhood). Consider a rooted marked directed graph $(G, \varnothing, M(G))$. Fix $k \in \mathbb{N}$. The *k-neighborhood of root* $\varnothing$ is a rooted marked directed graph $(U_{\leq k}(\varnothing), \varnothing, M(U_{\leq k}(\varnothing)))$ constructed as follows:

▷ for $k = 0$, $U_{\leq k}(\varnothing)$ is a graph with a single vertex $\varnothing$, no edges, and mark $m_{\varnothing}^{(\text{out})}$;
▷ for $k > 0$, consider $\varnothing$ as active, and proceed recursively as follows, for $h = 1, \dots, k$:

(1) for every vertex active at step $h - 1$, explore the incoming edges to the vertices in the opposite direction, finding the source of the edges;
(2) label the vertices that were active to be explored, and label the vertices just found as active, but only if they were not already found in the exploration process;
(3) for every vertex i (explored or active), assign the mark $m_i^{(\text{out})}$ to it, that is equal to the mark in the original graph $(G, \varnothing, M(G))$. In addition, draw every edge between two vertices that are already found in the exploration process;
(4) if there are no more active vertices, then stop the process.

In this way, we explore the *incoming neighborhood* of the root. As stated in Definition 3.10, we explore edges in the opposite direction: if $(j, i) \in E(G)$ is a directed edge, then the exploration process goes from vertex i to vertex j. Notice that it is possible that we do not explore the entire graph in this process, because we do not explore edges in all directions. This is different to the undirected case, where for k large enough, we always explore the entire graph (if connected).

We can define a local distance d_{loc} on $\mathcal{G}_\star$ as in Definition 3.3, but this time for rooted marked directed graphs, using Definitions 3.9 and 3.10. As in the undirected setting, the function d_{loc} tells us up to what distance the neighborhoods of two roots in two different rooted marked directed graphs are isomorphic. However, in the directed setting the function d_{loc} is *not* a metric on $\mathcal{G}_\star$, but it is a *pseudonorm*.

Note that d_{loc} is positive by definition, and obviously symmetric. It is not hard to prove that it satisfies the triangle inequality. The reason that d_{loc} is not a metric is that two rooted marked directed graphs can be at distance 0 without being isomorphic. This is due to the fact that the edges can be explored only in one direction, possibly leaving parts of the graph unexplored, as mentioned above. If the *explorable parts* or *incoming neighborhoods* of two graphs from the roots are isomorphic, then the two rooted marked directed graphs are at distance zero, while these graphs still might not be isomorphic. An example is given in Figure 5. Denote the explorable neighborhood of the root by $U_\infty(\varnothing)$, that is, the (possibly infinite) subgraph of a rooted marked directed graph that can be explored from the root. Then

$$\begin{aligned} &d_{\text{loc}}\big((G_1, \varnothing_1, M(G_1)), (G_2, \varnothing_2, M(G_2))\big) = 0 \quad \Longleftrightarrow \\ &U_\infty(\varnothing_1) \cong U_\infty(\varnothing_2). \end{aligned} \tag{14}$$

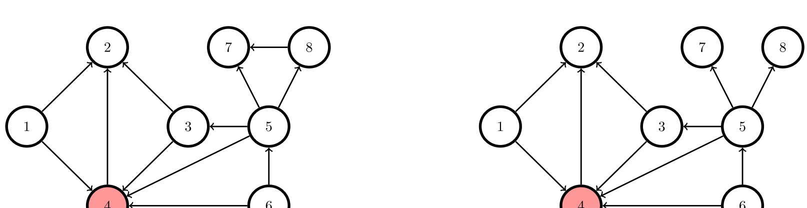

FIG. 5. *Example of two rooted marked directed graphs that are at distance zero, but are not isomorphic. The distance between the two graphs is zero since the explorable parts of the graphs from vertex* 4 (*including vertices* 1–6) *are isomorphic, but there exists no isomorphisms between the two graphs.*

Formally, $(\mathcal{G}_\star, d_{\text{loc}})$ is a complete and separable space, so every Cauchy sequence has a limiting point. Although the limiting point might not be unique, the explorable neighborhood of the root is unique. The proof that the space $(\mathcal{G}_\star, d_{\text{loc}})$ is a complete pseudometric space is a minor adaptation of the proofs in [32], Appendix A.

We can define the space $\tilde{\mathcal{G}}_\star$ as the quotient space of $\mathcal{G}_\star$ using the equivalence relation $\sim_\star$, where

$$(G_1, \varnothing_1, M(G_1)) \sim_\star (G_2, \varnothing_2, M(G_2)) \quad \Leftrightarrow$$
$$d_{\text{loc}}\big((G_1, \varnothing_1, M(G_1)), (G_2, \varnothing_2, M(G_2))\big) = 0.$$

On $\tilde{\mathcal{G}}_\star$, d_{loc} is a metric. Any equivalence class in $\tilde{\mathcal{G}}_\star$ is composed by directed marked rooted graphs whose neighborhoods of the root are isomorphic. Heuristically, everything that is in the part of the graph that is not explorable from the root *does not have any influence on the incoming neighborhood of the root*. This means that any function on $\tilde{\mathcal{G}}_\star$ is well defined if and only if it is a function of the incoming neighborhood of the root.

As in the undirected sense, we denote

$$\mathcal{P}(G) = \frac{1}{|V(G)|} \sum_{i \in V(G)} \delta_{(G,i,M(G))}. \tag{15}$$

When we consider a sequence of marked graphs $((G_n, M(G_n)))_{n\in\mathbb{N}}$, we denote $\mathcal{P}(G_n)$ by $\mathcal{P}_n$. From the definition, we have that $\mathcal{P}(G)$ is a probability on $\tilde{\mathcal{G}}_\star$, that assigns a uniformly chosen root to the marked directed finite graph. Notice that the mark set is fixed. In fact, the triplet $(G, i, M(G))$ is mapped to the equivalence class of the explorable neighborhood $U_\infty(i)$ of i in G with the same set of marks.

Since we are interested in sequences of random graphs, we give the definition of LW convergence only for random graphs.

DEFINITION 3.11 (Local weak convergence—directed). Consider a sequence of random marked directed graphs $(G_n, M(G_n))_{n\in\mathbb{N}}$. Let $(G, \varnothing, M(G))$ be a random element of $\tilde{\mathcal{G}}_\star$ with law $\mathcal{P}$. We say that G_n converges *in distribution* (*in probability, almost surely*) to $\mathcal{P}$ if (11) ((12), (13), resp.) holds for any bounded continuous function $f : \tilde{\mathcal{G}}_\star \to \mathbb{R}$.

REMARK 3.12 (Criterion for directed LW convergence). The reader can observe that, once the notion of exploration process and isomorphisms in the directed case are introduced, the construction of the definition of local weak convergence for directed graphs is the same as in the undirected case. With the presence of marks we are able to keep track of the out-degrees of vertices, while we explore the incoming edges.

It is easy to prove that Theorem 3.7 can be extended to random marked directed graphs. In other words, it is sufficient to prove the convergence for functions of the type $\mathbb{1}\{U_{\le k}(\varnothing) \cong (H, y, M(H))\}$, where $k \in \mathbb{N}$ and $(H, y, M(H))$ is a finite marked directed rooted graph.

4. Convergence of PageRank. The main result on PageRank is Theorem 2.1. It states that, for a locally weakly convergent sequence of directed random graphs $(G_n)_{n\in\mathbb{N}}$, there exists a random variable $R_\varnothing$ such that the PageRank value of a uniformly chosen vertex $R_{V_n}(n)$ satisfies

$$R_{V_n}(n) \xrightarrow{d} R_\varnothing.$$

The random variable $R_\varnothing$ is defined in Proposition 4.3 below. Notice that, even though local weak convergence is defined in terms of *local properties* of the graph, it is sufficient for the existence of the limiting distribution for a global property such as PageRank. This depends on the fact that vertices that are far away from a vertex $i \in [n]$ have, on average, small influence on the PageRank score $R_i(n)$. This is formulated in Lemma 4.1, that states that the contribution to the PageRank score of a uniformly chosen vertex V_n of other vertices decreases exponentially with the distance from V_n itself.

The existence of $R_\varnothing$ for a sequence $(G_n)_{n\in\mathbb{N}}$ is assured by the convergence in *distribution* in the local weak sense. If $(G_n)_{n\in\mathbb{N}}$ converges in probability (or almost surely), then the fraction of vertices whose PageRank value exceeds a fixed value $r > 0$ converges in probability (or almost surely) to a deterministic value.

4.1. *Finite approximation of PageRank.* Consider a directed graph G_n, and define the matrix $\boldsymbol{Q}(n)$, where $\boldsymbol{Q}(n)_{i,j} = e_{i,j}/d_i^{(\mathrm{out})}$, for $e_{i,j}$ the number of directed edges from i to j, and $\boldsymbol{Q}(n)_{i,j} = 0$ if $d_i^{(\mathrm{out})} = 0$. For $c \in (0, 1]$, the PageRank vector $\boldsymbol{\pi}(n) = (\pi_1, \dots, \pi_n)$ is the unique solution of

$$\boldsymbol{\pi}(n) = \boldsymbol{\pi}(n)\big[c\,\boldsymbol{Q}(n)\big] + \frac{1-c}{n}\mathbf{1}_n \quad \text{and} \quad \sum_{i=1}^{n} \pi_i = 1, \tag{16}$$

where $c \in (0, 1)$ and $\mathbf{1}_n$ is the vector of all ones of size n. We are interested in the graph-normalized version of PageRank, so $\boldsymbol{R}(n) = n\boldsymbol{\pi}(n)$, which is just the PageRank vector rescaled with the size of the graph. The vector $\boldsymbol{R}(n)$ satisfies

$$\boldsymbol{R}(n) = \boldsymbol{R}(n)\big[c\,\boldsymbol{Q}(n)\big] + (1-c)\mathbf{1}_n. \tag{17}$$

Denote Id_n the identity matrix of size n. We can solve (17) to obtain the well-known expression [5, 10, 16, 20]

$$\boldsymbol{R}(n) = (1-c)\mathbf{1}_n\big[\mathrm{Id}_n - c\,\boldsymbol{Q}(n)\big]^{-1}. \tag{18}$$

In practice, the inversion operation on the matrix $\mathrm{Id}_n - c\,\boldsymbol{Q}(n)$ is inefficient, therefore, power expansion is used to approximate the matrix in (18) (see, e.g., [5]), as

$$\big[\mathrm{Id}_n - c\,\boldsymbol{Q}(n)\big]^{-1} = \sum_{k=0}^{\infty} c^k \boldsymbol{Q}(n)^k. \tag{19}$$

Notice that $\boldsymbol{Q}(n)^k_{i,j} > 0$ if and only if there exists a path of length exactly k from i to j, possibly with repetition of vertices and edges. Define, for $k \in \mathbb{N}$,

$$\mathrm{path}_i(k) = \big\{\text{directed path } \ell = (\ell_0, \ell_1, \ell_2, \dots, \ell_k = i)\big\}.$$

With this notation, we can write, for $i \in [n]$,

$$R_i(n) = (1-c)\left(1 + \sum_{k=1}^{\infty} c^k \sum_{\ell \in \mathrm{path}_i(n)} \prod_{h=1}^{k} \frac{e_{\ell_h,\ell_{h+1}}}{d^+_{\ell_h}}\right), \tag{20}$$

while the Nth finite approximation of PageRank is

$$R_i^{(N)}(n) = (1-c)\left(1 + \sum_{k=0}^{N} c^k \sum_{\ell \in \mathrm{path}_i(n)} \prod_{h=1}^{k} \frac{e_{\ell_h,\ell_{h+1}}}{d^+_{\ell_h}}\right). \tag{21}$$

Heuristically, the PageRank formulation in (20) includes paths of every length, while the Nth approximation in (21) discards the paths of length $N+1$ or higher. In particular, for every $i \in [n]$, $R_i^{(N)}(n) \uparrow R_i(n)$. One can write the difference between the PageRank and its finite approximation as

$$\big|R_i(n) - R_i^{(N)}(n)\big| = (1-c)\mathbf{1}_n \sum_{k=N+1}^{\infty} \big(c\,\boldsymbol{Q}(n)\big)^k_i. \tag{22}$$

We can prove that we can approximate the PageRank value of a randomly chosen vertex by a finite approximation with an exponentially small error, that is independent of the size of the graph.

LEMMA 4.1 (Finite iterations). *Consider a directed graph G_n and denote a uniformly chosen vertex by V_n. Then*

$$\mathbb{E}\big[R_{V_n}(n) - R_{V_n}^{(N)}(n)\big] \le c^{N+1},$$

where the bound is independent of n.

PROOF. Consider (22) for a uniformly chosen vertex. We have

$$\mathbb{E}\big[R_{V_n}(n) - R_{V_n}^{(N)}(n)\big] = \frac{1-c}{n} \sum_{i=1}^{n} \sum_{k=N+1}^{\infty} \big[\mathbf{1}_n\big(c\boldsymbol{Q}(n)\big)^k\big]_i. \tag{23}$$

We write $\boldsymbol{Q}(n)^k_{j,i}$ to denote the element (j,i) of the matrix $\boldsymbol{Q}(n)^k$. We write

$$\big[\mathbf{1}_n\big(c\boldsymbol{Q}(n)\big)^k\big]_i = c^k \sum_{j=1}^{n} \boldsymbol{Q}(n)^k_{j,i}. \tag{24}$$

Substituting (24) in (23), we obtain

$$\mathbb{E}\big[R_{V_n}(n) - R_{V_n}^{(N)}(n)\big] = (1-c) \sum_{k=N+1}^{\infty} c^k \frac{1}{n} \sum_{i,j} \boldsymbol{Q}(n)^k_{j,i}.$$

Since $\boldsymbol{Q}(n)^k$ is a (sub)stochastic matrix,

$$\sum_{i=1}^{n} \boldsymbol{Q}(n)^k_{j,i} \le 1$$

for every $j \in [n]$. It follows that

$$\begin{aligned}\mathbb{E}\big[R_{V_n}(n) - R_{V_n}^{(N)}(n)\big] &\le (1-c) \sum_{k=N+1}^{\infty} c^k \frac{1}{n} \sum_{i=1}^{n} 1 \\ &= (1-c) \sum_{k=N+1}^{\infty} c^k \\ &= c^{N+1}.\end{aligned}$$

□

Lemma 4.1 means that we can approximate the PageRank value of a uniformly chosen vertex with an arbitrary precision in a finite number of iterations that is independent of the graph size. This is the starting point of our analysis. We point out that the proof of Lemma 4.1 is contained in [20], Section 4.2. We write it here for completeness of the argument and because we refer to it in Section 5.

4.2. *PageRank on marked directed graphs.* In this section, we show how the graph-normalized version of PageRank of a uniformly chosen vertex in a sequence of directed graphs $(G_n)_{n\in\mathbb{N}}$ admits a limiting distribution whenever G_n converges in the local weak sense to a distribution $\mathcal{P}$. The advantage is that such a limiting distribution is expressed in terms of functions of $\mathcal{P}$.

The first step is to write PageRank as functions of marked directed rooted graphs that are *bounded and continuous* with respect to the topology given by d_{loc}. In this way, by the definition of local weak convergence, we can pass to the limit and find the limiting distribution.

Fix $n \in \mathbb{N}$. Consider a marked rooted directed graph $(G, \varnothing, M(G)) \in \mathcal{G}_\star$ of size n. Denote as before, for $k \in \mathbb{N}$,

$$\mathrm{path}_\varnothing(k) = \{\text{directed paths } \ell = (\ell_0, \ell_1, \ell_2, \dots, \ell_k = \varnothing)\},$$

that is, the set of directed paths in $(G, \varnothing, M(G))$ of length exactly $k+1$ whose endpoint is the root $\varnothing$. It is clear that this set is completely determined by $U_{\leq k}(\varnothing)$ in $(G, \varnothing, M(G))$.

Consider a directed marked graph $(G_n, M(G_n))$, where we consider marks equal to the out-degrees. We have that

$$\begin{aligned}R_{V_n}^{(N)}(n) &= \sum_{i \in [n]} \mathbb{1}_{\{V_n = i\}}(1-c)\left(1 + \sum_{k=1}^{N} \sum_{\pi \in \mathrm{path}_i(k)} \prod_{h=1}^{k} c \frac{e_{\pi_h, \pi_{h+1}}}{d_{\pi_h}^{(\mathrm{out})}}\right) \\ &=: R^{(N)}[(G_n, V_n, M(G_n))],\end{aligned} \tag{25}$$

where the last term in (25) is a function of a marked rooted graph, evaluated on $(G_n, V_n, M(G_n))$, with V_n a uniformly chosen root. In particular, we can see the Nth approximation of PageRank as a function of the marked rooted graph. We call the function $R^{(N)} : \tilde{\mathcal{G}}_\star \to \mathbb{R}$ the *root N-PageRank*.

Clearly, the root N-PageRank $R^{(N)}$ is a function of $U_{\leq N}(\varnothing)$ only. It depends in fact on the vertices, edges and marks that are considered when exploring the graph from the root up to distance N. Notice that, since the dependence on the marked directed rooted graph is given only by $U_{\leq k}(\varnothing)$, the function $R^{(N)}$ is well defined on any equivalence class in $\tilde{\mathcal{G}}_\star$.

In addition, the function $R^{(N)}$ is *continuous* with respect to the topology generated by d_{loc}. In fact, since $R^{(N)}$ depends only on the root neighborhood up to distance N, whenever two elements $(G, \varnothing, M(G))$ and $(G', \varnothing', M(G'))$ are at distance less than $1/(1+N)$, their roots neighborhoods are isomorphic up to distance $N+1$, which implies that $R^{(N)}[(G, \varnothing, M(G))] = R^{(N)}[(G', \varnothing', M(G'))]$.

The problem is that $R^{(N)}$ *is not bounded*, so LWC does not assure that we can pass to the limit. To resolve this, we introduce a different type of function.

DEFINITION 4.2 (Root N-PageRank tail). Fix $N \in \mathbb{N}$. For $r > 0$, define $\Psi_{r,N} : \tilde{\mathcal{G}}_\star \to \{0, 1\}$ by

$$\Psi_{r,N}[(G, \varnothing, M(G))] := \mathbb{1}\{R^{(N)}[(G, \varnothing, M(G))] > r\}.$$

We call the function $\Psi_{r,N}$ the *root N-PageRank tail* at r.

The function $\Psi_{r,N}$ is clearly bounded, and it depends only on the neighborhood of the root $\varnothing$ up to distance N through the function $R^{(N)}$. This means that, for any $r > 0$, $\Psi_{r,N}$ is continuous.

Since the root N-PageRank on $\tilde{\mathcal{G}}_\star$ represents the Nth approximation of PageRank on directed graphs, it follows that

$$\mathbb{E}_{\mathcal{P}_n}[\Psi_{r,N}] = \frac{1}{n} \sum_{i \in [n]} \mathbb{1}\{R_i^{(N)}(n) > r\}, \tag{26}$$

that is, $\mathbb{E}_{\mathcal{P}_n}[\Psi_{r,N}]$ is the empirical fraction of vertices in G such that the Nth approximation of PageRank exceeds r. In particular, for every $r \geq 0$, if $G_n \to \mathcal{P}$ in distribution,

$$\mathbb{P}(R_{V_n}^{(N)}(n) > r) = \mathbb{E}\left[\frac{1}{n} \sum_{i \in [n]} \mathbb{1}\{R_i^{(N)}(n) > r\}\right] \longrightarrow \mathcal{P}(R_\varnothing^{(N)} \geq r), \tag{27}$$

while for convergence in probability (or almost surely), the limit in (27) exists in probability (or almost surely). Consider the sequence of random variables $(R_\varnothing^{(N)})_{N\in\mathbb{N}}$, where

$$R_\varnothing^{(N)} := R^{(N)}\big[(G, \varnothing, M(G))\big],$$

where $(G, \varnothing, M(G))$ is a random directed rooted graph with law $\mathcal{P}$. From (27), it follows that $R_{V_n}^{(N)}(n) \to R_\varnothing^{(N)}$ in distribution.

We have just proved that, for a sequence of directed graphs $(G_n)_{n\in\mathbb{N}}$ that converges locally weakly to $\mathcal{P}$, any finite approximation of the PageRank value of a uniformly chosen vertex converges in distribution to a limiting random variable, which is given by a function of $\mathcal{P}$.

4.3. *The limit of finite root ranks.* Assume that the sequence $(G_n)_{n\in\mathbb{N}}$ of directed graphs converges to a directed rooted marked graph $(G, \varnothing, M(G))$ with law $\mathcal{P}$. In principle, such limiting $(G, \varnothing, M(G))$ can be an infinite directed rooted marked graph. Because of this, we cannot simply take the limit as $N \to \infty$ of the sequence $(R_\varnothing^{(N)})_{N\in\mathbb{N}}$, where $\varnothing$ is the root of $(G, \varnothing, M(G))$, because the PageRank vector, as the invariant measure of a random walk as in (1), is not defined on an infinite graph. Nevertheless, if $\mathcal{P}$ is a LW limit of some sequence of directed graphs, it admits a such limit.

PROPOSITION 4.3 (Existence of limiting root rank). *Let $\mathcal{P}$ be a probability on $\tilde{\mathcal{G}}_\star$. If $\mathcal{P}$ is the LW limit in distribution of a sequence of marked directed graphs $(G_n)_{n\in\mathbb{N}}$, then there exists a random variable $R_\varnothing$ with $\mathbb{E}_\mathcal{P}[R_\varnothing] \le 1$, such that $\mathcal{P}$-a.s. $R_\varnothing^{(N)} \to R_\varnothing$. As a consequence, $\mathcal{P}(R_\varnothing < \infty) = 1$.*

PROOF. Clearly, the sequence $(R_\varnothing^{(N)})_{N\in\mathbb{N}}$ is $\mathcal{P}$-a.s. increasing. Therefore, the almost sure limit $R_\varnothing = \lim_{N\to\infty} R_\varnothing^{(N)}$ exists. This is independent of the fact that $\mathcal{P}$ is a LW limit.

By LW convergence, we know that $R_{V_n}^{(N)}(n) \to R_\varnothing^{(N)}$ in distribution. For every $N \in \mathbb{N}$, by Fatou's lemma we can bound

$$\mathbb{E}_\mathcal{P}\big[R_\varnothing^{(N)}\big] \le \liminf_{n\in\mathbb{N}} \mathbb{E}\big[R_{V_n}^{(N)}(n)\big] \le \liminf_{n\in\mathbb{N}} \mathbb{E}\big[R_{V_n}(n)\big] = 1,$$

where the second bound comes from the fact that any N-finite approximation of PageRank is less than the actual PageRank value, and the fact that the graph-normalized PageRank has expected value 1. Since $(R_\varnothing^{(N)})_{N\in\mathbb{N}}$ is increasing, we conclude that there exists $z \le 1$ such that

$$\mathbb{E}_\mathcal{P}[R_\varnothing] = \lim_{N\to\infty} \mathbb{E}_\mathcal{P}\big[R_\varnothing^{(N)}\big] = z. \qquad \square$$

4.4. *Proof of Theorem 2.1.* We start with implication (1) of Theorem 2.1. We want to prove that $R_{V_n}(n)$ converges to $R_\varnothing$ in distribution. So, for every $r \ge 0$ and $\varepsilon > 0$ there exists $M(\varepsilon) \in \mathbb{N}$ such that, for every $n \ge M(\varepsilon)$,

$$\big|\mathbb{P}\big(R_{V_n}(n) > r\big) - \mathcal{P}(R_\varnothing > r)\big| \le \varepsilon. \tag{28}$$

We can write, using the triangle inequality,

$$\begin{aligned} \big|\mathbb{P}\big(R_{V_n}(n) > r\big) - \mathcal{P}(R_\varnothing > r)\big| \le{}& \big|\mathbb{P}\big(R_{V_n}(n) > r\big) - \mathbb{E}\big[\mathcal{P}_n\big(R_\varnothing^{(N)} > r\big)\big]\big| \\ &+ \big|\mathbb{E}\big[\mathcal{P}_n\big(R_\varnothing^{(N)} > r\big)\big] - \mathcal{P}\big(R_\varnothing^{(N)} > r\big)\big| \\ &+ \big|\mathcal{P}\big(R_\varnothing^{(N)} > r\big) - \mathcal{P}(R_\varnothing > r)\big|. \end{aligned} \tag{29}$$

We show that (28) holds by proving that every term in the left-hand side of (29) can be bounded by $\varepsilon/3$.

By Lemma 4.1, we can bound the first term with c^{N+1} (independently of n). Therefore, defining $N_1 = \log_c(\varepsilon/3)$ and taking $N > N_1$, the first term is bounded by $\varepsilon/3$.

For the last term, we apply Proposition 4.3, so we can find $N_2 = N_2(\varepsilon) \in \mathbb{N}$ such that, for every $N \geq N_2$,

$$\big|\mathcal{P}\big(R_\varnothing^{(N)} > r\big) - \mathcal{P}(R_\varnothing > r)\big| \leq \varepsilon/3.$$

Set $N_0(\varepsilon) = \max(N_1, N_2)$. For any $N \geq N_0$, both the first and third terms are bounded by $\varepsilon/3$. Using LW convergence in distribution, we can find $M(N_0, \varepsilon) \in \mathbb{N}$ such that, for every $n \geq M$, the second term is bounded by $\varepsilon/3$. This completes the proof of statement (1).

For statement (2), we need to show that, for every $r > 0$, as $n \to \infty$,

$$\frac{1}{n}\sum_{i=1}^{n} \mathbb{1}\{R_i(n) > r\} \xrightarrow{\mathbb{P}} \mathcal{P}(R_\varnothing > r).$$

For every $N \in \mathbb{N} \cup \{\infty\}$ and $r \geq 0$, we denote the empirical fraction of vertices whose Nth approximation of PageRank in G_n exceeds r by

$$\bar{R}(n; r, N) := \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}\{R_i^{(N)}(n) > r\}. \tag{30}$$

If $N = \infty$, then $\bar{R}(n; r, N) = \bar{R}(n; r)$ is the empirical tail distribution of PageRank. By LW convergence in probability of $(G_n)_{n\in\mathbb{N}}$, we know that, for every $N \in \mathbb{N}$ and $r > 0$,

$$\bar{R}(n; r, N) \xrightarrow{\mathbb{P}} \mathcal{P}\big(R_\varnothing^{(N)} > r\big). \tag{31}$$

Fix $r > 0$, $\varepsilon > 0$. We need to show that for every $\delta > 0$ there exists $n_0(\delta) \in \mathbb{N}$ such that, for any $n \geq n_0$, $\mathbb{P}(|\bar{R}(n; r) - \mathcal{P}(R_\varnothing > r)| \geq \varepsilon) \leq \delta$. We can write, for N to be fixed,

$$\begin{aligned}\mathbb{P}\big(\big|\bar{R}(n; r) - \mathcal{P}(R_\varnothing > r)\big| \geq \varepsilon\big) \leq \frac{1}{\varepsilon}\big[&\mathbb{E}\big[\bar{R}(n; r) - \bar{R}(n; r, N)\big] \\ &+ \mathbb{E}\big[\big|\bar{R}(n; r, N) - \mathcal{P}\big(R_\varnothing^{(N)} > r\big)\big|\big] \\ &+ \big|\mathcal{P}\big(R_\varnothing^{(N)} > r\big) - \mathcal{P}(R_\varnothing > r)\big|\big].\end{aligned} \tag{32}$$

Similar to (29), we can find n and N large enough such that every term in the right-hand side of (32) is less than $\delta\varepsilon/3$.

For the first term, we apply Lemma 4.1, so we can find N_1 large enough such that $c^{N_1+1} \leq \delta\varepsilon/3$. For the last term, we apply Proposition 4.3, so we can find N_2 such that the last term is less than $\delta\varepsilon/3$.

Take $N_0 = \max\{N_1, N_2\}$. Then, by (31) and the fact that $\{\bar{R}(n; r, N)\}_{n\in\mathbb{N}}$ is uniformly integrable (since $\bar{R}(n; r, N) \leq 1$), we can find n_0 big enough such that

$$\mathbb{E}\big[\big|\bar{R}(n; r, N) - \mathcal{P}\big(R_\varnothing^{(N)} > r\big)\big|\big] \leq \delta\varepsilon/3$$

for all $n > n_0$, $N > N_0$. As a consequence, we conclude that, for any $n \geq n_0$,

$$\mathbb{P}\big(\big|\bar{R}(n; r) - \mathcal{P}(R_\varnothing > r)\big| \geq \varepsilon\big) \leq \delta,$$

which proves the convergence in probability.

4.5. *Undirected graphs.* Undirected graphs are in fact a special case of directed graphs, where each link is reciprocated. Theorem 2.1 does not make any assumption concerning link reciprocation, and thus it simply holds for undirected graphs as well. In that case, we may use the standard notion of the LWC for undirected graphs, as described in Section 3.1, and it is not hard to see that our notion of directed LW convergence reduces to this.

Let us explain why the special case of undirected graphs deserves our attention. Indeed, usually, undirected graphs are easier to analyze than directed ones. For example, the adjacency matrix of an undirected graph is symmetric, which implies many nice properties. However, PageRank is based on directed paths, and its analysis is greatly simplified when these paths do not contain cycles, with high probability.

For example, PageRank can be written as a product of three terms, one of which is the expected number of visits to i, starting from i, by a simple random walk, which terminates at each step with probability c [11]. Now notice that in undirected graphs, each edge can be traversed in both directions; hence, a path starting at i may return to i in only two steps, so the average number of visits to i will be a random variable that depends on the entire neighborhood. In contrast, for example, in the directed configuration model, returning to i is highly unlikely. This makes PageRank in undirected graphs hard to analyze, and only few results have been obtained so far (see, e.g., [9]).

Our result simultaneously covers the directed and the undirected cases because we only state the equivalence between the behavior of PageRank on a graph and on its limiting object. In this setting, the difficulties that arise in the analysis of PageRank on undirected graphs are, in fact, 'postponed' to the (undirected) limiting random graph.

5. Generalized PageRank.

In this section, we will show that Theorem 2.1 extends to generalized PageRank as given in (4). More precisely, we consider a sequence of directed random graphs $(G_n)_{n\in\mathbb{N}}$, and a sequence of generalized coefficients $((C_i^{(n)}, B_i^{(n)})_{i\in[n]})_{n\in\mathbb{N}}$ for the PageRank definition. In particular, for $j \in [n]$, the coefficient A_j in (4) is given by $A_j^{(n)} = C_j^{(n)}/D_j^{(\text{out})}$, where $D_j^{(\text{out})}$ is the out-degree of vertex j. The generalized coefficients are always assumed to be *nonnegative*.

The main result of this section is stated in Theorem 5.4. We divide its proof into two parts. The first part establishes the *exponential bound* on the error made by finite approximations, as given in Lemma 4.1 for the standard PageRank, holds in the generalized setting as well. This result is formulated in Lemma 5.1, and is proved in Section 5.1. The second part deals with *assigning generalized coefficients C and B to vertices in the directed LW limit*. Since the generalized coefficient can depend on the graph itself, some regularity conditions are necessary in order to be able to define the distribution of the coefficient on the limiting rooted graph. We explain this in Section 5.2. In Section 5.3, we compare out results to the literature, and in Section 5.4 we complete the proof of Theorem 5.4.

5.1. *Universality of finite approximations.* We first focus on extending Lemma 4.1 to the generalized setting, that is, proving that the PageRank score of a uniformly chosen vertex in the graph can be approximated by a finite number of iterations of the stochastic matrix of the random walk associated to PageRank, with arbitrary precision. We formulate the result in the following lemma.

LEMMA 5.1 (Universality of finite approximation). *Let G_n be a (random) graph of size $n \in \mathbb{N}$, and let $(C_i^{(n)}, B_i^{(n)})_{i\in\mathbb{N}}$ be coefficients for the generalized PageRank as in (4). Assume that*:

(a) *there exists a constant $c \in (0,1)$ such that, for every $n \in \mathbb{N}$ and $i \in [n]$, $0 \le C_i^{(n)} \le c < 1$ almost surely*;

(b) *there exists a constant* $0 < b < \infty$ *such that, for every* $n \in \mathbb{N}$, $\sup_{i\in[n]} \mathbb{E}[|B_i^{(n)}|] < b$.

Then, for every $\varepsilon > 0$ *and for every* $N \in \mathbb{N}$, *independently of* n,

$$\mathbb{P}\big(R_{V_n}(n) - R_{V_n}^{(N)}(n) \geq \varepsilon\big) \leq \frac{b}{\varepsilon(1-c)} c^{N+1}. \tag{33}$$

PROOF. Similar to the expression in (20) for the standard PageRank, we have that, given a sequence of generalized coefficients $(C_i^{(n)}, B_i^{(n)})_{i\in\mathbb{N}}$, for every $i \in [n]$,

$$R_i(n) = B_i^{(n)} + \sum_{k=1}^{\infty} \sum_{\ell\in\text{path}_i(k)} B_{\ell_k} \prod_{j=1}^{k} \frac{C_{\ell_j}^{(n)}}{D_{\ell_j}^{(\text{out})}}. \tag{34}$$

In other words, (34) shows that the generalized PageRank score of a vertex $i \in [n]$ is, as the standard one, the weighted sum of all paths ending at i. Let then $\boldsymbol{A}$ be a matrix such that $A(n)_{i,j} = C_i^{(n)} e_{i,j}/D_i^{(\text{out})}$. As in Section 4.1, define $Q(n)_{i,j} = e_{i,j}/D_i^{(\text{out})}$, and $Q(n)_{i,j} = 0$ if $D_i^{(\text{out})} = 0$. Using condition *(a)* of the lemma, we obtain that

$$\begin{aligned}
\mathbb{E}\big[R_{V_n}(n) - R_{V_n}^{(N)}(n)\big] &= \frac{1}{n}\mathbb{E}\Bigg[\sum_{i=1}^{n}\sum_{k=N+1}^{\infty} \big[\mathbf{B}_n\big(\boldsymbol{A}(n)\big)^k\big]_i\Bigg] \\
&\leq \frac{1}{n}\mathbb{E}\Bigg[\sum_{i=1}^{n}\sum_{k=N+1}^{\infty} c^k\big[\mathbf{B}_n \boldsymbol{Q}(n)^k\big]_i\Bigg] \\
&= \frac{1}{n}\mathbb{E}\Bigg[\sum_{i=1}^{n}\sum_{k=N+1}^{\infty} c^k \sum_{j=1}^{n} B_j^{(n)} \boldsymbol{Q}(n)_{j,i}^k\Bigg] \\
&= \frac{1}{n}\mathbb{E}\Bigg[\sum_{k=N+1}^{\infty} c^k \sum_{j=1}^{n} B_j^{(n)} \sum_{i=1}^{n} \boldsymbol{Q}(n)_{j,i}^k\Bigg].
\end{aligned} \tag{35}$$

By definition, $\boldsymbol{Q}(n)$ is a substochastic matrix, therefore, we have

$$\sum_{i=1}^{n} \boldsymbol{Q}(n)_{j,i}^k \leq 1.$$

Using this and condition *(b)* of the lemma, we obtain that the last expression in (35) is bounded by

$$\frac{1}{n}\mathbb{E}\Bigg[\sum_{k=N+1}^{\infty} c^k \sum_{j=1}^{n} B_j^{(n)}\Bigg] = \sum_{k=N+1}^{\infty} c^k \frac{1}{n}\sum_{j=1}^{n}\mathbb{E}\big[B_j^{(n)}\big] < \frac{b}{1-c}c^{N+1}. \tag{36}$$

Now, using the Markov inequality on the probability in (33) and the bound given in (36), the proof is complete. $\square$

We point out that Lemma 5.1 holds without any assumption on the dependence or independence of the generalized coefficients and the graph. The bound in (33) relies on the fact that $\boldsymbol{Q}(n)$ is a (sub)stochastic matrix, and the fact that V_n is a uniformly chosen vertex.

5.2. *Bringing coefficients to the limit.* In a graph G_n with coefficients $(C_i^{(n)}, B_i^{(n)})_{i\in[n]}$, the generalized PageRank score $R_i(n)$ of a vertex $i \in [n]$ is again given by a weighted sum of all paths ending at i, as in (34). Notice that the standard PageRank score is retrieved when we set $C_j^{(n)} \equiv c$ and $B_j^{(n)} \equiv 1 - c$ for all $n \in \mathbb{N}$ and $j \in [n]$. In the generalized case, however, the coefficients $(C_i^{(n)}, B_i^{(n)})_{i\in[n]}$ are not necessarily independent of each other and/or of the graph G_n. Hence, assuming $G_n \to (G, \varnothing, M(G))$ in the directed LW sense, it is not obvious *how to bring the generalized coefficients to the limit*, that is, how to assign a pair of coefficients (C_v, B_v) to every vertex v in $(G, \varnothing, M(G))$. In the case of the standard PageRank, this problem is trivial, since the coefficients are deterministic. Furthermore, if the coefficients are i.i.d., the solution is to assign i.i.d. coefficients in the limit as well. Beyond these two simplified scenarios, we need to impose some regularity conditions on $(C_i^{(n)}, B_i^{(n)})$, under which Theorem 2.1 holds for the generalized PageRank. In the remainder of this section, we first formally state our quite general regularity conditions; see Condition 5.2(a)–(c). Then we discuss motivation behind these conditions and their possible generalizations. Finally, we conclude the section by stating our main result.

For two random variables X, Y, we denote the random variable X conditioned on Y by $X|_Y$. We formulate the regularity conditions as follows.

CONDITION 5.2 (Regularity of generalized coefficients). Let $(G_n)_{n\in\mathbb{N}}$ be a sequence of directed random graphs, and let $((C_i^{(n)}, B_i^{(n)})_{i\in[n]})_{n\in\mathbb{N}}$ be a sequence of sequences of coefficients for the generalized PageRank as in (4). Then the *regularity conditions* for the generalized coefficients are:

(a) For every $n \in \mathbb{N}$, $(C_1^{(n)}, B_1^{(n)})|_{G_n}, \dots, (C_n^{(n)}, B_n^{(n)})|_{G_n}$ are independent of each other;
(b) For every $n \in \mathbb{N}$ and every $i \in [n]$,

$$(C_i^{(n)}, B_i^{(n)})\big|_{G_n} \overset{d}{=} (C_i^{(n)}, B_i^{(n)})\big|_{(D_i^{(\mathrm{in})}, D_i^{(\mathrm{out})})}; \tag{37}$$

(c) There exists a distribution $(C, B)_{[a,b]}$, with two integer parameters $a, b \in \mathbb{N}$, such that, for every $d_1, d_2 \in \mathbb{N}$,

$$(C_{V_n}^{(n)}, B_{V_n}^{(n)})\big|_{(D_{V_n}^{(\mathrm{in})}=d_1, D_{V_n}^{(\mathrm{out})}=d_2)} \xrightarrow{d} (C, B)_{[d_1,d_2]}. \tag{38}$$

Condition 5.2(a) says that, conditionally on the graph, the coefficients are independent across vertices. Condition 5.2(b) also specifies that $C_i^{(n)}$ and $B_i^{(n)}$ depend of the graph only through the in- and out-degree of the vertex i itself. Finally, Condition 5.2(c) states that joint distributions of $(C_i^{(n)}, B_i^{(n)})$, conditioned on $(D_i^{(\mathrm{in})}, D_i^{(\mathrm{out})})$ converges to a limit in distribution. Note that we do not assume anything on the joint distribution of $C_i^{(n)}$ and $B_i^{(n)}$.

The motivation for Condition 5.2 is to take advantage of representation (4) as follows. Note that for every $n \in \mathbb{N}$, and independently of the graph G_n, the weight of a path in (4) can be always factorized as a product of two terms: one term depending of the path itself and the out-degree of vertices along it, and the other term depending on a certain number of random variables (the generalized coefficients). Since in the limit the term depending of the path is a function of $(G, \varnothing, M(G))$, we can now assign coefficients to vertices in $(G, \varnothing, M(G))$, by sampling (C_v, B_v) independently for each vertex v from the limiting distribution $(C, B)_{[D_v^{(\mathrm{in})}, m_v^{(\mathrm{out})}]}$. This will complete the construction of the limiting PageRank score.

REMARK 5.3 (Further extensions of Condition 5.2). We point out that Condition 5.2 can be generalized to allow some dependence between vertices of the graph. We restrict ourselves to Condition 5.2 because it is already quite general and easy to explain, moreover, it resembles the conditions used in earlier work [20, 21, 45] (see more detailed discussion in the next section). As a possible extension, for example, we could relax Condition 5.2(b), by replacing the dependence on the degree with the dependence on finite neighborhoods. More specifically, we believe that Theorem 5.4 stated below still holds, if we replace (37) with

$$(C_i^{(n)}, B_i^{(n)})\big|_{G_n} \overset{d}{=} (C_i^{(n)}, B_i^{(n)})\big|_{U_{\le K}(i)} \quad \text{for some fixed } K \in \mathbb{N},$$

even though some extra work is required to formally prove it. We do not investigate this further in the present paper.

With Lemma 5.1 and Condition 5.2, we are ready to state the convergence result for generalized PageRank.

THEOREM 5.4 (Asymptotic generalized PageRank distribution). *Let $(G_n)_{n\in\mathbb{N}}$ be a sequence of directed random graphs, and let $((C_i^{(n)}, B_i^{(n)})_{i\in[n]})_{n\in\mathbb{N}}$ be a sequence of generalized PageRank coefficients such that*:

(i) *there exists a constant $c \in (0, 1)$ such that, for every $n \in \mathbb{N}$ and $i \in [n]$, $C_i^{(n)} \le c < 1$*;
(ii) *there exists a constant $0 < b < \infty$ such that, for every $n \in \mathbb{N}$, $\sup_{i\in[n]} \mathbb{E}[|B_i^{(n)}|] < b$*;
(iii) *Condition 5.2 is satisfied.*

Then Theorem 2.1 holds for the generalized PageRank defined in (4).

5.3. *Comparison to related results in the literature.* Chen, Litvak and Olvera-Cravioto [21] investigate generalized PageRank on directed configuration models, while Lee and Olvera-Cravioto [45] analyze it on directed inhomogeneous random graphs. In these two works, the limiting distribution of PageRank is obtained as a solution of a stochastic fixed-point equation. In particular, it is proved that this solution obeys a power-law distribution with the same power-law exponent as the in-degree distribution. We will demonstrate in more detail how our method applies to these models in Section 6.1 (configuration model) and Section 6.2 (generalized random graph). This section compares the assumptions in [21, 45] to ones in our this paper.

Analogously to our assumptions in Lemma 5.1, in [21, 45] it is assumed that there exist $c \in (0, 1)$ such that $\mathbb{P}(C_i^{(n)} \le c) = 1$ for every $i \in [n]$, $n \in \mathbb{N}$, and the expectation of B is finite. In [21, 45], independence of (C_v, B_v) across vertices v in the limiting graph follows from the coupling of a random graph with a weighted branching tree. Similarly, we need such independence in order to assign coefficients in the limiting rooted graph. Note that we in fact assume independence of (C_i, B_i), $i \in G_n$, for simplicity of the argument; we could incorporate asymptotic independence or even dependence on finite neighborhoods, as discussed above.

Let us now compare the regularity conditions in [21, 45] and in our work. In [21], the directed configuration model is constructed from a so-called *bi-degree distributions* $(D_i^{(\mathrm{in})}, D_i^{(\mathrm{out})}, C_i^{(n)}, B_i^{(n)})_{i\in[n]}$. Then [21], Assumption 5.1, contains some conditions on the moments, such as the first moment of in- and out-degrees, and it is assumed that there exist two distributions F, $F^\star$ such that

$$\begin{aligned}&(D_{V_n}^{(\mathrm{in})}, D_{V_n}^{(\mathrm{out})}, C_{V_n}^{(n)}, B_{V_n}^{(n)}) \xrightarrow{d} F := (D^{(\mathrm{in})}, D^{(\mathrm{out})}, C, B),\\ &(D_{V_n^\star}^{(\mathrm{in})}, D_{V_n^\star}^{(\mathrm{out})}, C_{V_n^\star}^{(n)}, B_{V_n^\star}^{(n)}) \xrightarrow{d} F^\star,\end{aligned} \tag{39}$$

where in (39) V_n and $V_n^\star$ are respectively a uniformly chosen vertex and a uniformly chosen incoming neighbor of V_n. Likewise, [45] makes assumptions on the moments and the limit of the *bi-weight sequences*

$$(W_i^{(\mathrm{in})}, W_i^{(\mathrm{out})}, C_i^{(n)}, B_i^{(n)})_{i\in[n]}$$

in generalized random graphs, in particular,

$$\big(W_{V_n}^{(\mathrm{in})}, W_{V_n}^{(\mathrm{out})}, C_{V_n}^{(n)}, B_{V_n}^{(n)}\big) \xrightarrow{d} \big(W^{(\mathrm{in})}, W^{(\mathrm{out})}, C, B\big).$$

Notably, additional assumptions are made on the limiting distributions: in [21], Assumption 6.2, in the limit, $C/D^{(\mathrm{out})}$ is independent of $(D^{(\mathrm{in})}, B)$, and in [45], Assumption 3.1, in the limit, $(W^{(\mathrm{in})}, B)$ is independent of $(W^{(\mathrm{out})}, C)$. These additional independence assumptions are necessary in order to prove the convergence of PageRank to the endogenous solution of a stochastic fixed-point equation. In Condition 5.2, we do not need to assume this, because we are interested only in convergence of PageRank and not in a specific form of its limiting distribution.

In summary, in [21, 45], for specific random graph models, and under some additional conditions, the limiting PageRank distribution was completely characterized, and the power-law hypothesis was proved for these two types of random graphs. The argument in this paper, on the other hand, is designed to work on *every sequence of LW convergent graphs*, and shows, under more general conditions, that the asymptotic PageRank distribution is a function of the LW limit. While the strength of [21, 45] is in complete analysis of specific models, the advantage of our setting is its universality.

5.4. *Proof of Theorem* 5.4. Consider a sequence of directed random graphs $(G_n)_{n\in\mathbb{N}}$ that converges locally weakly in distribution to $(G, \varnothing, M(G))$ with law $\mathcal{P}$. Consider also a sequence of generalized coefficients $((C_i^{(n)}, B_i^{(n)})_{i\in[n]})_{n\in\mathbb{N}}$, and assume that Condition 5.2 is satisfied.

First, we use Condition 5.2 to construct the limiting distribution $R_\varnothing$ and its finite approximations $(R_\varnothing^{(N)})_{N\in\mathbb{N}}$ as follows. Conditionally on the graph $(G, \varnothing, M(G))$, assign conditionally independent coefficients

$$\big((C, B)_{[d_v^{(\mathrm{in})}, m_v^{(\mathrm{out})}]}\big)_{v\in G}$$

to the vertices. Recall that $m_v^{(\mathrm{out})}$ stands for the mark of a node, which can be different from its out-degree in the limiting rooted graph. Then, for $N \in \mathbb{N}$, define

$$R_\varnothing^{(N)} := B_\varnothing + \sum_{k=1}^{N} \sum_{\ell\in \mathrm{path}_\varnothing(k)} B_{\ell_k} \prod_{h=1}^{k} \frac{C_{\ell_h}}{m_{\ell_h}^{(\mathrm{out})}}, \quad \text{and}$$

(40)

$$R_\varnothing := \lim_{N\to\infty} R_\varnothing^{(N)} = B_\varnothing + \sum_{k=1}^{\infty} \sum_{\ell\in \mathrm{path}_\varnothing(k)} B_{\ell_k} \prod_{h=1}^{k} \frac{C_{\ell_h}}{m_{\ell_h}^{(\mathrm{out})}}.$$

The limit $R_\varnothing$ in (40) exists and is finite since $(R_\varnothing^{(N)})_{N\in\mathbb{N}}$ is a sequence of a.s. monotonically increasing random variables, and

$$\mathbb{E}[R_\varnothing] \le \liminf_{N\in\mathbb{N}} \mathbb{E}\big[R_\varnothing^{(N)}\big] < \infty$$

by the assumptions of Lemma 5.2.

Convergence in distribution. To prove Theorem 5.4, we start showing that Theorem 2.1(1) holds for generalized PageRank. In other words, we have to show that, for every $r \ge 0$, as $n \to \infty$,

$$\big|\mathbb{P}\big(R_{V_n}(n) > r\big) - \mathcal{P}(R_\varnothing > r)\big| \to 0,$$

that is, $R_{V_n}(n) \to R_\varnothing$ in distribution. We can write, similar to (29) for the standard PageRank,

$$\begin{aligned} \bigl|\mathbb{P}\bigl(R_{V_n}(n) > r\bigr) - \mathcal{P}(R_\varnothing > r)\bigr| &\le \bigl|\mathbb{P}\bigl(R_{V_n}(n) > r\bigr) - \mathbb{E}\bigl[\mathcal{P}_n\bigl(R_\varnothing^{(N)} > r\bigr)\bigr]\bigr| \\ &\quad + \bigl|\mathbb{E}\bigl[\mathcal{P}_n\bigl(R_\varnothing^{(N)} > r\bigr)\bigr] - \mathcal{P}\bigl(R_\varnothing^{(N)} > r\bigr)\bigr| \\ &\quad + \bigl|\mathcal{P}\bigl(R_\varnothing^{(N)} > r\bigr) - \mathcal{P}(R_\varnothing > r)\bigr|. \end{aligned} \tag{41}$$

The first term in (41) is small because of Lemma 5.1, which tells us that we can approximate the generalized PageRank score of a uniformly chosen vertex with arbitrary precision as soon as conditions (i) and (ii) of Theorem 5.4 are satisfied. The last term in (41) is small since by definition $R_\varnothing^{(N)}$ converges in distribution to $R_\varnothing$.

It remains to prove that the second term on the right-hand side of (41) is small. We will prove that this term is $o(1)$. Fix then a finite marked rooted graph $(H, y, M(H))$. Using Condition 5.2(a), we write

$$\begin{aligned} &\mathbb{E}\bigl[\mathcal{P}_n\bigl(R_\varnothing^{(N)} > r, U_{\le N}(V_n) \cong \bigl(H, y, M(H)\bigr)\bigr)\bigr] \\ &\quad = \frac{1}{n}\sum_{i\in[n]} \mathbb{P}\bigl(R_i^{(N)}(n) > r \mid U_{\le N}(i) \cong \bigl(H, y, M(H)\bigr)\bigr) \\ &\qquad \times \mathbb{P}\bigl(U_{\le N}(i) \cong \bigl(H, y, M(H)\bigr)\bigr) \\ &\quad = \mathbb{E}\bigl[\Lambda_{N,r}^{(H,y,M(H))}\bigl(\bigl(C_i^{(n)}, B_i^{(n)}\bigr)_{i\in[n]}\bigr)\bigr]\frac{1}{n}\sum_{i\in[n]} \mathbb{P}\bigl(U_{\le N}(i) \cong \bigl(H, y, M(H)\bigr)\bigr) \\ &\quad = \mathbb{E}\bigl[\Lambda_{N,r}^{(H,y,M(H))}\bigl(\bigl(C_i^{(n)}, B_i^{(n)}\bigr)_{i\in[n]}\bigr)\bigr]\mathbb{E}\bigl[\mathcal{P}_n\bigl(U_{\le N}(V_n) \cong \bigl(H, y, M(H)\bigr)\bigr)\bigr], \end{aligned} \tag{42}$$

where $\Lambda_{N,r}^{(H,y,M(H))}$ is one when the Nth approximation of the PageRank score of the root y in $(H, y, M(H))$ is larger than r, and zero otherwise. The random variable $\Lambda_{N,r}^{(H,y,M(H))}$ is a function of the finite structure $(H, y, M(H))$, where the randomness is given only by a finite number of $(C_i^{(n)}, B_i^{(n)})_{i\in[n]}$. Notice that the second equality in (42) is based on Condition 5.2(a), and, more generally, exploits the fact that given the structure $(H, y, M(H))$ of the finite neighborhood, the (joint) distribution of the coefficients depends only on this structure; in fact, in our case, the coefficients depend only on the degrees, as stated in Condition 5.2(b).

With a similar argument, we can write an analogous expression for $R_\varnothing$, thus obtaining

$$\begin{aligned} &\mathcal{P}\bigl(R_\varnothing^{(N)} > r, U_{\le N}(\varnothing) \cong \bigl(H, y, M(H)\bigr)\bigr) \\ &\quad = \mathbb{E}\bigl[\Lambda_{N,r}^{(H,y,M(H))}\bigl((C_v)_{v\in G}, (B_v)_{v\in G}\bigr)\bigr]\mathcal{P}\bigl(U_{\le N}(\varnothing) \cong \bigl(H, y, M(H)\bigr)\bigr). \end{aligned} \tag{43}$$

Now, for every $(H, y, M(G))$, N and r, since $\Lambda_{N,r}^{(H,y,M(H))}(\cdot)$ is a bounded function, hence, by Condition 5.2(c) we have that, as $n \to \infty$,

$$\mathbb{E}\bigl[\Lambda_{N,r}^{(H,y,M(H))}\bigl(\bigl(C_i^{(n)}, B_i^{(n)}\bigr)_{i\in[n]}\bigr)\bigr] \longrightarrow \mathbb{E}\bigl[\Lambda_{N,r}^{(H,y,M(H))}\bigl((C_v, B_v)_{v\in G}\bigr)\bigr]. \tag{44}$$

Furthermore, by the LWC assumption, we have that, for every $(H, y, M(H))$, and $N \in \mathbb{N}$,

$$\mathbb{E}\bigl[\mathcal{P}_n\bigl(U_{\le N}(V_n) \cong \bigl(H, y, M(H)\bigr)\bigr)\bigr] \longrightarrow \mathcal{P}\bigl(U_{\le N}(\varnothing) \cong \bigl(H, y, M(H)\bigr)\bigr). \tag{45}$$

We now write

$$\begin{aligned} \mathbb{E}\bigl[\mathcal{P}_n(R_{V_n} > r)\bigr] &= \sum_{(H,y,M(H))} \mathbb{E}\bigl[\mathcal{P}_n\bigl(R_\varnothing^{(N)} > r, \bigl(H, y, M(H)\bigr)\bigr)\bigr] \\ &= \sum_{(H,y,M(H))} \mathbb{E}\bigl[\Lambda_{N,r}^{(H,y,M(H))}\bigl(\bigl(C_i^{(n)}, B_i^{(n)}\bigr)_{i\in[n]}\bigr)\bigr] \\ &\qquad \times \mathbb{E}\bigl[\mathcal{P}_n\bigl(U_{\le N}(V_n) \cong \bigl(H, y, M(H)\bigr)\bigr)\bigr]. \end{aligned} \tag{46}$$

Using (44), (45) and dominated convergence in (46), we complete the proof that the second term in (41) is $o(1)$ as $n \to \infty$. It follows that if $G_n \to (G, \varnothing, M(G))$ in distribution in the LW sense, then the generalized PageRank converges in distribution.

Convergence in probability. We now prove Theorem 2.1(2) for generalized PageRank. Assume that $G_n \to (G, \varnothing, M(G))$ in probability. We use the same notation as in Section 4.4, recall $\bar{R}(n; r, N)$ as defined in (30). Following the same steps as in Section 4.4, we can write a similar expression to (32) for the generalized PageRank. Then we just need to prove that, for every $N \in \mathbb{N}$ and $r \geq 0$, as $n \to \infty$,

$$\mathbb{E}\big[\big|\bar{R}(n; r, N) - \mathcal{P}\big(R_{\varnothing}^{(N)} > r\big)\big|\big] \longrightarrow 0. \tag{47}$$

Since LWC in probability implies convergence in distribution, by Theorem 5.4(a) we know that, for every $N \in \mathbb{N}$ and $r \geq 0$,

$$\mathbb{E}\big[\bar{R}(n; r, N)\big] \longrightarrow \mathcal{P}\big(R_{\varnothing}^{(N)} > r\big). \tag{48}$$

We want to apply the second moment method to $\bar{R}(n; r, N)$. If we prove that

$$\mathbb{E}\big[\bar{R}(n; r, N)^2\big] \leq \mathbb{E}\big[\bar{R}(n; r, N)\big]^2\big(1 + o(1)\big), \tag{49}$$

then (48) and (49) together, by the second moment method, prove (47). We write, for V_n, W_n two independent uniformly chosen vertices in $[n]$,

$$\begin{aligned}
&\mathbb{E}\big[\bar{R}(n; r, N)^2\big] \\
&\quad = \mathbb{P}\big(R_{V_n}^{(N)} > r, R_{W_n}^{(N)} > r\big) \\
&\quad = \mathbb{P}\big(R_{V_n}^{(N)} > r, R_{W_n}^{(N)} > r, U_{\leq N}(V_n) \cap U_{\leq N}(W_n) = \varnothing\big) \\
&\qquad + \mathbb{P}\big(R_{V_n}^{(N)} > r, R_{W_n}^{(N)} > r, U_{\leq N}(V_n) \cap U_{\leq N}(W_n) \neq \varnothing\big) \\
&\quad = \mathbb{P}\big(R_{V_n}^{(N)} > r, R_{W_n}^{(N)} > r, U_{\leq N}(V_n) \cap U_{\leq N}(W_n) = \varnothing\big)\big(1 + o(1)\big),
\end{aligned} \tag{50}$$

where the last equality comes from the fact that indicator functions are bounded by one and, when we look at *finite* neighborhoods $U_{\leq N}(V_n)$ and $U_{\leq N}(W_n)$ of two uniformly chosen vertices $V_n, W_n \in [n]$, the probability that the two vertices are overlapping is vanishing as $n \to \infty$.

For two vertices $V_n, W_n \in [n]$, on the event $U_{\leq N}(V_n) \cap U_{\leq N}(W_n) = \varnothing$, we have that the generalized coefficients in the expressions of $R_{V_n}^{(N)}(n)$ and $R_{W_n}^{(N)}(n)$ are all distinct. As a consequence, by Condition 5.2(b), we have that

$$\begin{aligned}
&\mathbb{P}\big(R_{V_n}^{(N)} > r, R_{W_n}^{(N)} > r, U_{\leq N}(V_n) \cap U_{\leq N}(W_n) = \varnothing\big) \\
&\quad = \sum_{\substack{(H, y, M(H)) \\ (H', y', M(H'))}} \mathbb{E}\big[\Lambda_{N,r}^{(H,y,M(H))}\big(\big(C_i^{(n)}, B_i^{(n)}\big)_{i \in [n]}\big)\big] \\
&\qquad \times \mathbb{E}\big[\Lambda_{N,r}^{(H',y',M(H'))}\big(\big(C_i^{(n)}, B_i^{(n)}\big)_{i \in [n]}\big)\big] \\
&\qquad \times \mathbb{P}\big(U_{\leq N}(V_n) \cap U_{\leq N}(W_n) = \varnothing, U_{\leq N}(V_n) \cong \big(H, y, M(H)\big), \\
&\qquad\qquad U_{\leq N}(W_n) \cong \big(H', y', M(H')\big)\big),
\end{aligned} \tag{51}$$

where $\Lambda_{N,r}^{(\cdot,\cdot,M(\cdot))}(\cdot)$ are the functions introduced in (42). The expectation of the product of two functions $\Lambda_{N,r}^{(\cdot,\cdot,M(\cdot))}(\cdot)$ is factorized in (51) since the neighborhoods of V_n and W_n use two disjoint subsets of $[n]$ as indices, and the coefficients involved are conditionally independent given the graph (Condition 5.2(a)).

Now consider the probability in the right-hand side of (51). Since V_n and W_n are independent of G_n, we can write

$$\begin{aligned}
&\mathbb{P}\big(U_{\le N}(V_n) \cap U_{\le N}(W_n) = \varnothing, U_{\le N}(V_n) \cong \big(H, y, M(H)\big),\\
&\qquad U_{\le N}(W_n) \cong \big(H', y', M\big(H'\big)\big)\big)\\
&\quad = \mathbb{E}\bigg[\frac{1}{n^2}\sum_{i,j} \mathbb{1}_{\{U_{\le N}(i)\cong(H,y,M(H)), U_{\le N}(j)\cong(H',y',M(H')), U_{\le N}(i)\cap U_{\le N}(j)=\varnothing\}}\bigg]\\
&\quad \le \mathbb{E}\bigg[\frac{1}{n^2}\sum_{i,j} \mathbb{1}_{\{U_{\le N}(i)\cong(H,y,M(H))\}}\mathbb{1}_{\{U_{\le N}(j)\cong(H',y',M(H'))\}}\bigg]\\
&\quad = \mathbb{E}\big[\mathcal{P}_n\big(\big(H, y, M(H)\big)\big)\mathcal{P}_n\big(\big(H', y', M\big(H'\big)\big)\big].
\end{aligned}$$

Since the LWC holds in probability, for every pair of marked rooted graphs $(H, y, M(H))$ and $(H', y', M(H'))$, we have that, by dominated convergence,

$$\begin{aligned}
&\mathbb{E}\big[\mathcal{P}_n\big(\big(H, y, M(H)\big)\big)\mathcal{P}_n\big(\big(H', y', M\big(H'\big)\big)\big]\\
&\qquad \longrightarrow \mathcal{P}\big(\big(H, y, M(H)\big)\big)\mathcal{P}\big(\big(H', y', M\big(H'\big)\big).
\end{aligned}$$

In fact, we have that $\mathcal{P}_n((H, y, M(H))) \xrightarrow{\mathbb{P}} \mathcal{P}((H, y, M(H)))$ and, similarly, $\mathcal{P}_n((H', y', M(H'))) \xrightarrow{\mathbb{P}} \mathcal{P}((H', y', M(H')))$. It follows that

$$\begin{aligned}
&\mathbb{P}\big(U_{\le N}(V_n) \cap U_{\le N}(W_n) = \varnothing, U_{\le N}(V_n) \cong \big(H, y, M(H)\big),\\
&\qquad U_{\le N}(W_n) \cong \big(H', y', M\big(H'\big)\big)\big)\\
&\quad \le \mathbb{P}\big(U_{\le N}(V_n) \cong \big(H, y, M(H)\big)\big)\\
&\qquad \times \mathbb{P}\big(U_{\le N}(W_n) \cong \big(H', y', M\big(H'\big)\big)\big)\big(1 + o(1)\big).
\end{aligned}\tag{52}$$

Using (52), we can bound the term on the right-hand side of (51) by

$$\begin{aligned}
&\sum_{\substack{(H,y,M(H))\\(H',y',M(H'))}} \mathbb{E}\big[\Lambda_{N,r}^{(H,y,M(H))}\big(\big(C_i^{(n)}, B_i^{(n)}\big)_{i\in[n]}\big)\big]\\
&\qquad \times \mathbb{E}\big[\Lambda_{N,r}^{(H',y',M(H'))}\big(\big(C_i^{(n)}, B_i^{(n)}\big)_{i\in[n]}\big)\big]\mathbb{P}\big(U_{\le N}(V_n) \cong \big(H, y, M(H)\big)\big)\\
&\qquad \times \mathbb{P}\big(U_{\le N}(W_n) \cong \big(H', y', M\big(H'\big)\big)\big)\big(1 + o(1)\big)\\
&\quad = \sum_{(H,y,M(H))} \mathbb{E}\big[\Lambda_{N,r}^{(H,y,M(H))}\big(\big(C_i^{(n)}, B_i^{(n)}\big)_{i\in[n]}\big)\big]\mathbb{P}\big(U_{\le N}(V_n) \cong \big(H, y, M(H)\big)\big)\\
&\qquad \times \sum_{(H',y',M(H'))} \mathbb{E}\big[\Lambda_{N,r}^{(H',y',M(H'))}\big(\big(C_i^{(n)}, B_i^{(n)}\big)_{i\in[n]}\big)\big]\\
&\qquad \times \mathbb{P}\big(U_{\le N}(W_n) \cong \big(H', y', M\big(H'\big)\big)\big)\big(1 + o(1)\big)\\
&\quad = \mathbb{E}\big[\bar{R}(n; r, N)\big]^2\big(1 + o(1)\big),
\end{aligned}$$

so that by the second moment method the proof is complete.

REMARK 5.5 (Weighted rooted graphs). Benjamini, Lyons and Schramm [13] consider undirected LWC for graphs with weighted edges. In particular, they define a different metric on the space of weighted rooted graphs, that includes the distance between edge weights. This construction can be extended to vertex weights, and it would lead to a different approach to investigate this different type of metric. This requires additional work, for example due to the fact that the metric in [13] is not a simple extension of the metrics that we consider in Sections 3.1 and 3.2. We will not further investigate these different type of metrics in this paper.

6. Examples of directed local weak convergence.

6.1. *Directed configuration model.* The directed configuration model, denoted by DCM, is a version of the configuration model where half-edges are labeled as in- and out-half-edges. In this setting, DCM_n is a directed graph of size $n \in \mathbb{N}$ with prescribed in- and out-degree sequences. We denote the in-degree sequence by $\boldsymbol{D}_n^{(\mathrm{in})} = (D_1^{(\mathrm{in})}, \dots, D_n^{(\mathrm{in})})$ and the out-degree sequence by $\boldsymbol{D}_n^{(\mathrm{out})} = (D_1^{(\mathrm{out})}, \dots, D_n^{(\mathrm{out})})$. We call $(\boldsymbol{D}_n^{(\mathrm{out})}, \boldsymbol{D}_n^{(\mathrm{in})})$ the *bi-degree sequence* of the graph.

For a precise description of the DCM, we refer to [21, 22]. The graph is defined as follows: let $n \in \mathbb{N}$ be the size of the graph, and fix a bi-degree sequence $(\boldsymbol{D}_n^{(\mathrm{out})}, \boldsymbol{D}_n^{(\mathrm{in})})$. The graph is generated by fixing a free outgoing half-edge and we pair it uniformly at random with a free incoming half edge. In this process, self-loops and multiple edges may arise. Since the pairing is made uniformly, it is not relevant in which order we choose the free outgoing half-edge. In this setting, the total in-degree and out-degree of the graph have to be equal. In the case of random in- and out-degrees, this is a rare event. The algorithm presented in [22] generates an admissible bi-degree sequence in a finite number of steps, and approximates the initial degree distributions. Our conditions on the bi-degree sequence are as follows.

CONDITION 6.1 (Bi-degree regularity conditions). Let $(\boldsymbol{D}_n^{(\mathrm{out})}, \boldsymbol{D}_n^{(\mathrm{in})})$ be a bi-degree sequence. Then, the *bi-degree regularity conditions* are as follows:

(a) There exists a distribution $(p(h,l))_{h,l\in\mathbb{N}}$ such that, for every $h, l \in \mathbb{N}$, as $n \to \infty$,

$$\frac{1}{n}\sum_{i\in[n]} \mathbb{1}_{\{D_i^{(\mathrm{out})}=h, D_i^{(\mathrm{in})}=l\}} \longrightarrow p(h,l); \tag{53}$$

(b) Denote by $(\mathcal{D}^{(\mathrm{out})}, \mathcal{D}^{(\mathrm{in})})$ a random vector with distribution $(p(h,l))_{h,l\in\mathbb{N}}$ as in (53). Then, as $n \to \infty$,

$$\begin{aligned} \frac{1}{n}\sum_{i\in[n]}\sum_{h\in\mathbb{N}} h\mathbb{1}_{\{D_i^{(\mathrm{out})}=h\}} &\longrightarrow \mathbb{E}[\mathcal{D}^{(\mathrm{out})}], \\ \frac{1}{n}\sum_{i\in[n]}\sum_{l\in\mathbb{N}} l\mathbb{1}_{\{D_i^{(\mathrm{in})}=l\}} &\longrightarrow \mathbb{E}[\mathcal{D}^{(\mathrm{in})}], \end{aligned} \tag{54}$$

and $\mathbb{E}[\mathcal{D}^{(\mathrm{in})}] = \mathbb{E}[\mathcal{D}^{(\mathrm{out})}]$;

(c) For $L_n = D_1^{(\mathrm{out})} + \cdots + D_n^{(\mathrm{out})}$, as $n \to \infty$,

$$\frac{1}{n}\sum_{i\in[n]} \frac{h}{L_n}\mathbb{1}_{\{D_i^{(\mathrm{out})}=h, D_i^{(\mathrm{in})}=l\}} \longrightarrow \frac{h}{\mathbb{E}[\mathcal{D}^{(\mathrm{out})}]}p(h,l) =: p^\star(h,l). \tag{55}$$

Denote by $(\mathcal{D}^{\star(\mathrm{out})}, \mathcal{D}^{\star(\mathrm{in})})$ the random vector with distribution given by $(p^\star(h,l))_{h,l\in\mathbb{N}}$.

Condition 6.1(a) implies that the empirical bi-degree distribution converges to a limiting distribution given by $(p(h,l))_{h,l\in\mathbb{N}}$ as in (53). Condition 6.1(b) implies that both the in- and out-degree distributions have finite first moment, equal to the one of $(p(h,l))_{h,l\in\mathbb{N}}$. Condition 6.1(c) implies that the out-degree size-biased distribution converges to a limiting distribution $(p^\star(h,l))_{h,l\in\mathbb{N}}$ as in (55).

With Condition 6.1 in hand, we are ready to state the convergence result on DCM.

PROPOSITION 6.2. *Consider a directed configuration model* DCM_n *such that the bi-degree sequence* $(\boldsymbol{D}_n^{(\mathrm{out})}, \boldsymbol{D}_n^{(\mathrm{in})})$ *satisfies Condition* 6.1. *Then* DCM_n converges in probability in the directed LW sense *to the law* $\mathcal{P}$ *of a* marked *Galton–Watson tree*, *where*:

(1) *edges are directed from children to parents*;
(2) *the mark and the in-degree of the root are distributed as* $(\mathcal{D}^{(\mathrm{out})}, \mathcal{D}^{(\mathrm{in})})$ *as in* (53);
(3) *the mark and the in-degree of any other vertex are independent across the tree vertices*, *and are distributed according to* $(\mathcal{D}^{\star(\mathrm{out})}, \mathcal{D}^{\star(\mathrm{in})})$ *as in* (55).

The proof of Proposition 6.2 is an adaptation of the proof for the undirected case as presented in [39], Section 2.2.2. The proof is divided into two parts. First, we use a coupling argument to prove that DCM_n converges in distribution to the prescribed limit. The second part consists in the application of the second moment method on the number of vertices in DCM_n with a fixed finite neighborhood structure, to prove that the number of such vertices is concentrated around its mean.

We start with the coupling argument.

LEMMA 6.3 (LW convergence of DCM in distribution). *Under the assumptions of Proposition* 6.2, *there exists a marked Galton–Watson tree* $\mathrm{GW}^{(n)}$ *such that*, *for every* finite *marked rooted tree* $(H, y, M(H))$,

$$\mathbb{P}\big(U_{\le k}(V_n) \cong (H, y, M(H))\big) = \mathbb{P}\big(\mathrm{GW}^{(n)}_{\le k} \cong (H, y, M(H))\big) + o(1), \tag{56}$$

where $\mathrm{GW}^{(n)}_{\le k}$ *denote the first k generations of* $\mathrm{GW}^{(n)}$. *Further*, $\mathrm{GW}^{(n)} \to \mathcal{P}$ *locally weakly in distribution*, *where* $\mathcal{P}$ *is the limit in Proposition* 6.2. *As a consequence*, $\mathrm{DCM}_n \to \mathcal{P}$ *locally weakly in distribution*.

PROOF. We prove that, for every finite $k \in \mathbb{N}$ and n large enough, the k-neighborhood of a uniformly chosen vertex in DCM_n has approximately the same distribution as the first k generations of a marked Galton–Watson tree GW, where the marks and offspring distributions are as in Proposition 6.2. We prove this in several steps, that also prove (56):

- We construct a coupling between DCM_n and a marked Galton–Watson tree $\mathrm{GW}^{(n)}$, where marks and offpsring distributions depend on n. In particular, the coupling holds with high probability up to the exploration of a $s = s(n) \to \infty$ number of incoming half edges, for an appropriate $s(n)$;
- The marked Galton–Watson tree $\mathrm{GW}^{(n)}$ converges locally weakly to the marked Galton–Watson tree GW that is the LW limit of the directed configuration model.

We start by constructing the coupling between DCM_n and a marked Galton–Watson tree $\mathrm{GW}^{(n)}$. Define $(p_n(h,l))_{h,l\in\mathbb{N}}$ and $(p_n^\star(h,l))_{ih,l\in\mathbb{N}}$ by

$$\begin{aligned} p_n(h,l) &= \frac{1}{n}\sum_{i\in[n]} \mathbb{1}_{\{D_i^{(\mathrm{out})}=h, D_i^{(\mathrm{in})}=l\}}, \\ p_n^*(h,l) &= \frac{1}{n}\sum_{i\in[n]} \frac{h}{L_n}\mathbb{1}_{\{D_i^{(\mathrm{out})}=h, D_i^{(\mathrm{in})}=l\}}, \end{aligned} \tag{57}$$

where $L_n = D_1^{(\text{out})} + \cdots + D_n^{(\text{out})}$. Notice that Condition 6.1 implies that $p_n(h,l) \to p(h,l)$ and $p_n^\star(h,l) \to p^\star(h,l)$ for every $k,l \geq 0$.

The coupling is constructed as follows: the mark and the in-degree of V_n and the root $\varnothing_n$ in $\mathrm{GW}^{(n)}$ are chosen according to the distribution p_n as in (57). Therefore, $U_{\leq 0}(V_n)$ and the 0-generation of $\mathrm{GW}^{(n)}$ (which both consist only of the root and its mark) are the same. Enumerate then the incoming-half edges of V_t according to any order.

Fix now a number $s \in \mathbb{N}$ of incoming half-edges to explore. We have to construct the subgraph created by the exploration of s incoming half-edges starting at V_n in DCM_n and from $\varnothing_n$ in $\mathrm{GW}^{(n)}$ at the same time.

The construction is recursive in the number of incoming half-edges explored. Order the half-edges in an arbitrary way. After t incoming half-edges are explored, the new exploration steps is made as follows: take the unpaired incoming half-edge x_{t+1} with the smallest label among the ones incident to vertices already found in the exploration. We pair x_{t+1} to a uniformly chosen outgoing half-edge. We choose this outgoing half-edge y_{t+1} uniformly at random among all outgoing half-edges, independently from the previously matched half-edges.

Let W_{t+1} be the vertex in DCM_n to which y_{t+1} is incident. Then, in $\mathrm{GW}^{(n)}$ we assign to a new vertex mark and in-degree equal to $(D_{W_{t+1}}^{(\text{out})}, D_{W_{t+1}}^{(\text{in})})$. Notice that in this case the pair $(D_{W_{t+1}}^{(\text{out})}, D_{W_{t+1}}^{(\text{in})})$ is distributed as $p_n^\star$ given in (57).

In DCM_n we have to be careful since the half-edge y_{t+1} might have already been paired. If y_{t+1} has not yet been paired, then we pair x_{t+1} to y_{t+1} to create an edge. If y_{t+1} has already been paired, then we draw a new outgoing half-edge y'_{t+1} chosen uniformly from the unpaired ones and we pair x_{t+1} to y'_{t+1}.

We repeat this procedure until we explore s half-edges (or there are no incoming half-edges to explore) We can have differences between the exploration process in DCM_n and $\mathrm{GW}^{(n)}$. Differences can happen in two ways:

(1) the outgoing half-edge that we select to create a new edge has already been paired;
(2) the outgoing half-edge that we select to create a new edge has not been paired yet, but it is incident to a vertex already found in the exploration process.

We will show that these two contributions have small probability when we pair few edges. In fact, after creating t edges, the probability that we select an outgoing half-edge that is already used equals t/L_n, where L_n is the total number of outgoing edges. This means that the probability that in the first s steps we use the same out-going half-edge twice is bounded by

$$\sum_{t=0}^{s} \frac{t}{L_n} = \frac{s(s+1)}{2L_n}. \tag{58}$$

Thanks to Condition 6.1(b), L_n is of order n, so the expression in (58) is $o(1)$ whenever $s = o(\sqrt{n})$. The probability of selecting a vertex i when choosing an outgoing half-edge is $D_i^{(\text{out})}/L_n$. Then the probability that a vertex i is selected at least twice when s edges are created is bounded by

$$\frac{s(s+1)}{2} \frac{(D_i^{(\text{out})})^2}{L_n^2}. \tag{59}$$

Using (59) and the union bound, the probability that a vertex is selected twice when s edges are created is bounded by

$$\frac{s(s+1)}{2} \sum_{i=1}^{n} \frac{(D_i^{(\text{out})})^2}{L_n^2} \leq \frac{s(s+1)}{2L_n} D_{\max}^{(\text{out})}, \tag{60}$$

where $D^{(\mathrm{out})}_{\max} = \max i \in [n] D^{(\mathrm{out})}_i$ is the maximum out-degree in the bi-degree sequence. In this case, the expression in (60) is $o(1)$ when $s = o(\sqrt{n/D^{(\mathrm{out})}_{\max}})$. Further, $D^{(\mathrm{out})}_{\max}$ under Condition 6.1 is $o(n)$.

The two bounds in (58) and (60) together holds for $s = s(n)$, with $s(n) \to \infty$ sufficiently slowly.

Consider then a finite marked rooted tree $(H, y, M(H))$ of depth $k \in \mathbb{N}$. Since any finite tree H is made by a finite number of edges $S = S(H)$, we can take n large enough such that $s(n) \geq S$, that implies that $U_{\leq k}(V_n)$ and $\mathrm{GW}^{(n)}_{\leq k}$ are isomorphic with high probability, which implies (56).

We can then prove that DCM_n converges to GW. We fix $(H, y, M(H))$ of depth k and again we take n large enough such that $s(n) \geq S(H)$. By the triangle inequality,

$$\begin{aligned} &\big|\mathbb{P}\big(U_{\leq k}(V_n) \cong \big(H, y, M(H)\big)\big) - \mathbb{P}\big(\mathrm{GW}_{\leq k} \cong \big(H, y, M(H)\big)\big)\big| \\ &\quad \leq \big|\mathbb{P}\big(U_{\leq k}(V_n) \cong \big(H, y, M(H)\big)\big) - \mathbb{P}\big(\mathrm{GW}^{(n)}_{\leq k} \cong \big(H, y, M(H)\big)\big)\big| \\ &\qquad + \big|\mathbb{P}\big(\mathrm{GW}^{(n)}_{\leq k} \cong \big(H, y, M(H)\big)\big) - \mathbb{P}\big(\mathrm{GW}_{\leq k} \cong \big(H, y, M(H)\big)\big)\big|. \end{aligned} \tag{61}$$

The first term in the right-hand side of (61) is $o(1)$ by (56). The second term is $o(1)$, since the two rooted trees $\mathrm{GW}^{(n)}_{\leq k}$ and $\mathrm{GW}_{\leq k}$ are both the first k generations of unimodular marked Galton–Watson trees, where the difference is given by the marks and the offspring distributions. More precisely, given a finite structure $(H, y, M(H))$ for a marked rooted tree, the probability that the subtree consisting of the first k generations of a marked Galton–Watson tree is isomorphic to $(H, y, M(H))$ involves a finite number of independent random variables (i.e., marks and offsprings of a finite number of vertices). By Condition 6.1, we know that the marks and offspring distributions in $\mathrm{GW}^{(n)}_{\leq k}$ converge to the ones in $\mathrm{GW}_{\leq k}$, thus

$$\mathbb{P}\big(\mathrm{GW}^{(n)}_{\leq k} \cong \big(H, y, M(H)\big)\big) \longrightarrow \mathbb{P}\big(\mathrm{GW}_{\leq k} \cong \big(H, y, M(H)\big)\big),$$

which implies the LW convergence in distribution. This completes the proof. □

Next, we prove the convergence in probability, using the second moment method on the number of vertices in DCM_n with a prescribed neighborhood $(H, y, M(H))$. We start by giving a preliminary result on the second moment of such vertices.

LEMMA 6.4 (Second moment method). *Fix finite structure $(H, y, M(H))$ for the root neighborhood of depth $k \in \mathbb{N}$. Let $N_n(k; H, y, M(H))$ be the number of vertices $i \in [n]$ in* DCM_n *such that $U_{\leq k}(i) \cong (H, y, M(H))$. Then, as $n \to \infty$,*

$$\frac{1}{n^2}\mathbb{E}\big[N_n\big(k; H, y, M(H)\big)^2\big] \longrightarrow \mathcal{P}\big(U_{\leq k}(\varnothing) \cong \big(H, y, M(H)\big)\big)^2. \tag{62}$$

PROOF. We can rewrite $\mathbb{E}[N_n(k; H, y, M(H))^2]/n^2$ as

$$\mathbb{P}\big(U_{\leq k}(V_n) \cong \big(H, y, M(H)\big), U_{\leq k}(W_n) \cong \big(H, y, M(H)\big)\big),$$

where V_n and W_n are two vertices chosen uniformly at random in $[n]$. Since we fix $k \in \mathbb{N}$, we can take n large enough such that, with high probability, W_n is not a vertex found in the exploration up to distance $2k$ from V_n. Then we can rewrite the probability in the right-hand side of (62) as

$$\begin{aligned} &\mathbb{P}\big(U_{\leq k}(V_n) \cong \big(H, y, M(H)\big), \\ &\quad U_{\leq k}(W_n) \cong \big(H, y, M(H)\big), W_n \notin U_{\leq 2k}(V_n)\big) + o(1), \end{aligned}$$

where the factor $2k$ comes from the fact that we look at the structure $(H, y, M(H))$ for the two neighborhoods when they are disjoint. With a similar argument to the one just used, since k is fixed,

$$\begin{aligned}\mathbb{P}\big(U_{\le k}(V_n) \cong \big(H, y, M(H)\big), W_n \notin U_{\le 2k}(V_n)\big) \\ \longrightarrow \mathcal{P}\big(U_{\le k}(\varnothing) \cong \big(H, y, M(H)\big)\big).\end{aligned} \tag{63}$$

We now use the fact that, conditioning on the existence of a tree in DCM_n, the probability to have a second tree disjoint from the first one is equal to having a tree in a different configuration model with different size and bi-degree distribution.

More precisely, conditioning on $\{U_{\le k}(V_n) \cong (H, y, M(H)), W_n \notin U_{\le 2k}(V_n)\}$, we want to evaluate the probability of having a second tree $U_{\le k}(W_n) \cong (H, y, M(H))$, disjoint from $U_{\le k}(V_n) \cong (H, y, M(H))$. We have that

$$\begin{aligned}&\mathbb{P}\big(U_{\le k}(W_n) \cong \big(H, y, M(H)\big) \mid U_{\le k}(V_n) \cong \big(H, y, M(H)\big), W_n \notin U_{\le 2k}(V_n)\big) \\ &\quad = \mathbb{P}\big(\widehat{U}_{\le k}(\widehat{W}_n) \cong \big(H, y, M(H)\big), \widehat{j} \notin \widehat{U}_{\le k}(\widehat{W}_n)\big),\end{aligned} \tag{64}$$

where $\widehat{U}_{\le k}(\widehat{W}_n)$ is the k-neighborhood of a vertex $\widehat{W}_n$ chosen uniformly at random in a different configuration model $\widehat{\mathrm{DCM}}_n$, and $\widehat{j}$ is a particular vertex in $\widehat{\mathrm{DCM}}_n$ whose characteristics are specified below.

The vertices set and bi-degree sequence of $\widehat{\mathrm{DCM}}_n$ are defined as follows:

(1) if $i \notin U_{\le k}(V_n)$, then i is a vertex in $\widehat{\mathrm{DCM}}_n$ with the same in- and out-degree $(D_i^{(\mathrm{out})}, D_i^{(\mathrm{in})})$;

(2) if $i \in U_{\le k}(V_n)$, then i is not present in $\widehat{\mathrm{DCM}}_n$;

(3) define an additional vertex $\widehat{j}$ in $\widehat{\mathrm{DCM}}_n$, where we denote its in- and out-degree by $(\widehat{D}_j^{(\mathrm{out})}, \widehat{D}_j^{(\mathrm{in})})$. $\widehat{D}_j^{(\mathrm{out})}$ equals the sum of the unpaired outgoing half-edges in $U_{\le k}(V_n)$, and $\widehat{D}_j^{(\mathrm{in})}$ equals the number of unpaired ingoing half-edges in $U_{\le k}(V_n)$. We point out that $\widehat{U}_{\le k}(\widehat{W}_n)$ needs to avoid $\widehat{j}$.

Notice that the unpaired incoming half-edges in $U_{\le k}(V_n)$ are incident only to vertices at distance k from the root, while the unpaired outgoing half-edges are incident to all vertices in $U_{\le k}(V_n)$. We have that $\widehat{\mathrm{DCM}}_n$ is a graph with $n - |U_{\le k}(V_n)| + 1$ vertices, and a different bi-degree sequence.

The graph $\widehat{\mathrm{DCM}}_n$ is then created by pairing an incoming half-edge to a uniformly chosen outgoing half-edge, as usual as in the regular DCM_n. The probability to observe a structure in $\widehat{\mathrm{DCM}}_n$ that is disjoint from the vertex $\widehat{j}$ is exactly the same as in the regular DCM_n, conditioning on the structure of $U_{\le k}(V_n)$. This explains the equality in (64).

It is immediate to verify that the bi-degree sequence of $\widehat{\mathrm{DCM}}_n$ satisfies Condition 6.1, since we modify a negligible fraction of vertices (recall that k is fixed). As a consequence,

$$\mathbb{P}\big(\widehat{U}_{\le k}(\widehat{W}_n) \cong \big(H, y, M(H)\big)\big) \longrightarrow \mathcal{P}\big(U_{\le k}(\varnothing) \cong \big(H, y, M(H)\big)\big). \tag{65}$$

Using together (63) and (65), we complete the proof of (62). □

Completion of the proof of Proposition 6.2. Fix a finite marked rooted tree $(H, y, M(H))$ of depth $k \in \mathbb{N}$. As we did in Lemma 6.4, denote by $N_n(k; H, y, M(H))$ the number of vertices i in DCM_n such that $U_{\le k}(i) \cong (H, y, M(H))$. We need to show that, for every $\varepsilon > 0$, as $n \to \infty$,

$$\mathbb{P}\big(\big|N_n\big(k; H, y, M(H)\big)/n - \mathcal{P}\big(U_{\le k}(\varnothing) \cong \big(H, y, M(H)\big)\big)\big| \ge \varepsilon\big) \longrightarrow 0. \tag{66}$$

We use Markov inequality to bound the probability in (66) with

$$\frac{1}{\varepsilon}\mathbb{E}\big[\big|N_n\big(k; H, y, M(H)\big)/n - \mathcal{P}\big(U_{\leq k}(\varnothing) \cong \big(H, y, M(H)\big)\big)\big|\big].$$

By Lemma 6.3, $\mathbb{E}[N_n(k; H, y, M(H))]/n \to \mathcal{P}(U_{\leq k}(\varnothing) \cong (H, y, M(H)))$. By Lemma 6.4, we can apply the second moment method to $N_n(H, y, M(H))/n$, thus obtaining that

$$\mathbb{E}\big[\big|N_n\big(k; H, y, M(H)\big)/n - \mathcal{P}\big(U_{\leq k}(\varnothing) \cong \big(H, y, M(H)\big)\big)\big|\big] \longrightarrow 0.$$

Since this holds for any finite marked rooted tree $(H, y, M(H))$, the proof of Proposition 6.2 is compete.

DCM with independent in- and out-degrees. In [21], the limiting distribution of PageRank in DCM has been obtained when the size-biased in- and out-degrees are independent:

$$p^\star(h, l) = \frac{h}{\mathbb{E}[\mathcal{D}^{(\mathrm{out})}]}\mathbb{P}\big(\mathcal{D}^{\star(\mathrm{out})} = h\big)\mathbb{P}\big(\mathcal{D}^{(\mathrm{in})} = l\big).$$

Notice that $\mathcal{D}^{(\mathrm{out})}$ and $\mathcal{D}^{(\mathrm{in})}$ can, in general, be dependent, that is, $\mathcal{D}^{(\mathrm{in})}$ may have a different distribution conditioned on the event $\{\mathcal{D}^{(\mathrm{out})} \neq 0\}$, because the vertices with zero out-degrees do not contribute in PageRank of other vertices.

The local weak convergence for this case follows from [21], Lemma 5.4, hence, our Theorem 2.1 provides an alternative argument for the existence of the limiting PageRank distribution. It has been proved in [20, 21], under some technical assumptions, that in the limit the PageRank is distributed as

$$\mathcal{R} \stackrel{d}{=} \sum_{i=1}^{\mathcal{N}} \frac{c}{\mathcal{D}_i^{\star(\mathrm{out})}}\mathcal{R}_i^\star + (1-c), \tag{67}$$

where $\mathcal{R}^\star$ are independent realizations of the endogenous solution of the stochastic fixed-point equation

$$\mathcal{R}^\star \stackrel{d}{=} \sum_{i=1}^{\mathcal{N}^\star} \frac{c}{\mathcal{D}_i^{\star(\mathrm{out})}}\mathcal{R}_i^\star + (1-c). \tag{68}$$

The recursion (68) has been studied in a number of papers (see [43, 52]) and further references in [21]. The argument in [21] is more general, in fact the authors consider generalized PageRank as the solution of a more general equation than (68), where the $(1-c)$ is replaced by a random variable $\mathcal{B}$. In particular, if $\mathcal{D}^{(\mathrm{in})}$ is regularly varying with a tail heavier than the tail of $\mathcal{B}$, then the limiting PageRank $\mathcal{R}$ follows a power law with the same exponent as the in-degree $\mathcal{D}^{(\mathrm{in})}$.

6.2. *Inhomogeneous random graphs.* In the directed inhomogeneous random graphs, each vertex i receives an in-weight $W_i^{(\mathrm{in})}$ and an out-weight $W_i^{(\mathrm{out})}$. There is a directed edge from vertex i to vertex j with probability $w_{ij}^{(n)}$, which depends on $W_i^{(\mathrm{out})}$ and $W_j^{(\mathrm{in})}$. Lee and Olvera-Cravioto [45] study PageRank in the class of inhomogeneous random graphs that satisfy the assumption

$$w_{ij}^{(n)} = \min\left\{1, \frac{W_i^{(\mathrm{out})}W_j^{(\mathrm{in})}}{\theta n}\big(1 + \phi_{ij}(n)\big)\right\},$$

where $\phi_{ij}(n)$ satisfies some technical conditions, and is in fact vanishing as $n \to \infty$ for most natural models. This formulation includes Erdős–Rényi model, the Chung–Lu model, the

Poissonian random graph and the generalized random graph. For a detailed analysis of the properties of such directed graphs, we refer to [18].

LWC for this class of graphs follows directly from [45], Theorem 3.6, under general conditions, including that the in- and out-weights are allowed to be dependent. Hence, our results imply that PageRank converges in this model as well, to the PageRank of the limiting random graph.

In the case when the in- and out-weights are asymptotically independent, it is proved in [45] that the PageRank converges to the attracting endogenous solution of stochastic recursion (68). In particular, a power-law distribution of in-weights implies the power-law distribution of PageRank.

6.3. *Directed CTBP trees.* CTBPs are models that describe the evolution of a population composed by individuals that produce children according to i.i.d. birth processes. These models have been intensively studied in the literature [8, 42, 48]. The convergence result is stated in Proposition 6.6 below, which requires some notation from CTBPs theory that we present now.

DEFINITION 6.5 (Branching process). We define the *Ulam–Harris set* as

$$\mathcal{U} = \bigcup_{n \in \mathbb{N}} \mathbb{N}^n \quad \text{where } \mathbb{N}^0 := \{\varnothing\}. \tag{69}$$

Consider a birth process ξ. Then the continuous-time branching process is described by

$$(\Omega, \mathcal{A}, \mathbb{P}) = \prod_{x \in \mathcal{U}} (\Omega_x, \mathcal{A}_x, \mathbb{P}_x), \tag{70}$$

where $(\Omega_x, \mathcal{A}_x, \mathbb{P}_x)$ are probability spaces and $(\xi^x)_{x \in \mathcal{U}}$ are i.i.d. copies of ξ. For $x \in \mathbb{N}^n$ and $k \in \mathbb{N}$ we denote the kth child of x by $xk \in \mathbb{N}^{n+1}$. More generally, for $x \in \mathbb{N}^n$ and $y \in \mathbb{N}^m$, we denote the y descendant of x by xy. We call the branching process the triplet $(\Omega, \mathcal{A}, \mathbb{P})$ and the sequence of point processes $(\xi^x)_{x \in \mathcal{U}}$. We denote the branching process by $\boldsymbol{\xi}$.

The behavior of CTBPs is determined by properties of the birth process. Consider a jump process ξ on $\mathbb{R}^+$, that is, an integer-valued random measure on $\mathbb{R}^+$. Then we say that ξ is *supercritical and Malthusian* when there exists $\alpha^* > 0$ such that

$$\mathbb{E}[\xi_{T_{\alpha^*}}] = 1, \qquad \mu := -\frac{d}{d\alpha} \mathbb{E}[\xi_{T_\alpha}]\bigg|_{\alpha = \alpha^*} < +\infty, \tag{71}$$

where T_α is an exponential random variable with mean $1/\alpha$. The unique value α^* that satisfies (71) is called the *Malthusian parameter*.

An important class of functions of branching processes are *random characteristics*. A *random characteristic* is a real-valued process $\Phi \colon \Omega \times \mathbb{R} \to \mathbb{R}$ such that, for $x \in \mathcal{U}$, $\Phi(x, s) = 0$ for any $s < 0$, and $\Phi(x, s) = \Phi(s)$ is a deterministic bounded function for every $s \geq 0$ that only depends on x through the birth time of the individual, its birth process as well as the birth processes of its children.

Random characteristics are used to evaluate the number of individuals that at time $t \geq 0$ satisfies a property. For instance, consider $\Phi(t) = \mathbb{1}_{\mathbb{R}^+}(t)$ for $x \in \mathcal{U}$ and $t \geq 0$, that is, the characteristic that is equal to one whenever the individual is alive at time t. Then the branching process evaluated at time t with the random characteristic $\mathbb{1}_{\mathbb{R}^+}(\cdot)$ is equal to the number of individuals alive at time t. We denote the CTBP evaluated with a random characteristic Φ by $\boldsymbol{\xi}_t^{\Phi}$.

It is known that, for a random characteristic Φ, as $t \to \infty$,

$$\frac{\boldsymbol{\xi}_t^{\Phi}}{\boldsymbol{\xi}_t^{\mathbb{1}_{\mathbb{R}^+}}} \xrightarrow{\mathbb{P}\text{-a.s.}} \mathbb{E}\big[\Phi(T_{\alpha^*})\big], \tag{72}$$

where the left-hand term in (72) is the fraction of alive individuals that satisfies the property given by Φ. The right-hand side is the expectation of Φ, evaluated at an exponentially distributed time T_{α^*}, on an independent copy of the CTBP.

The convergence in (72) is a general result that is used often in the literature [7, 34, 42, 48, 51]. We refer to [51], Theorem A, for a simplified formulation of the result contained in [48].

With the notation just introduced we can formulate the convergence result.

PROPOSITION 6.6 (LWC for CTBPs trees). *Consider a supercritical and Malthusian birth process $(\xi_t)_{t\geq 0}$. Denote the corresponding CTBP by $\boldsymbol{\xi}$. Let $\mathcal{T}(t)$ be the directed random tree defined by $\boldsymbol{\xi}$ at time t, where edges are directed from children to parents. Then, on the event $\{|\mathcal{T}(t)| \to \infty\}$,* $\mathcal{T}(t)$ converges $\mathbb{P}$-a.s. in the LW sense *to the law of $\mathcal{T}(T_{\alpha^*})$, where*:

(1) *all marks are* 1;
(2) *edges are directed from children to parents*;
(3) *T_{α^*} is an exponentially distributed random variable with parameter α^* (the Mathusian parameter of the CTBP).*

PROOF. First of all, at every $t \in \mathbb{R}^+$, $\mathcal{T}(t)$ is a directed finite tree. We can equivalently prove the result on the discrete sequence $(\mathcal{T}_n)_{n\in\mathbb{N}}$, where $\mathcal{T}_n = \mathcal{T}(\tau_n)$, for $(\tau_n)_{n\in\mathbb{N}}$ the sequence of birth times of the CTBP.

Denote the vertices in $\mathcal{T}_n$ by their birth order, which means that the root of $\mathcal{T}_n$ in the sense of CTBP is vertex 1. First of all, notice that, for every $i \in [n]$ and $N \in \mathbb{N}$, the N neighborhood $U_{\leq N}(i)$ in the directed marked rooted graphs $(\mathcal{T}_n, i, 1)$ is just the subtree rooted at i composed by the descendants of i only up to generation N (from i). Notice that every vertex has out-degree 1 except for vertex 1 since it has out-degree 0.

What we need to prove is that, for any finite directed rooted tree (H, y) of depth N and with mark 1 for every vertex, we have, as $n \to \infty$,

$$\frac{1}{n}\sum_{i\in[n]} \mathbb{1}\big\{U_{\leq N}(i) \cong (H, y)\big\} \xrightarrow{\mathbb{P}\text{-a.s.}} \mathbb{P}\big(U_{\leq N}(\varnothing) \cong (H, y)\big), \tag{73}$$

where $U_{\leq N}(\varnothing)$ is the N-neighborhood of the root $\varnothing$ in the random tree $\mathcal{T}(T_{\alpha^*})$. For every $i \in [n]$, the indicator function inside the expectation satisfies the definition of random characteristic, since it is a bounded function that, for every individual i in the branching population, depends only on the birth time τ_i and on the randomness associated to i and its descendants. As a consequence, the result follows by (72). $\square$

The result in Proposition 6.6 resembles the subtree counting result in [51], Theorem 2. In the case of CTBPs, since marks do not play a relevant role (they are all equal to one as the out-degree), the incoming neighborhood of a vertex corresponds to the definition of *fringe tree* as given in [40], Section 4. In particular, the fact that the limiting distribution of the tree $\mathcal{T}(t)$ defined by a CTBP is $\mathcal{T}(T_{\alpha^*})$ is stated in this form by Holmgren and Janson in [40], Theorem 5.12, who refer to previous results by Jagers and Nerman [41, 42] and Aldous [2].

Notice that the limiting rooted graph in Proposition 6.6 is finite with probability 1. In particular, this means that the incoming neighborhood of a uniformly chosen vertex in a CTBP tree is *young with high probability*, since the limiting distribution of its neighborhood is a finite tree. This is rather different than the undirected settings, where typically the limiting rooted graph is infinite when considering a sequence of graphs with growing size.

REMARK 6.7 (Nonrecursive property of PageRank). The behavior of PageRank is often investigated starting from the recursive distributional equation in (4). In particular, the solution of (4) is constructed using a weighted Galton–Watson tree. This construction is based on the fact that the subtree rooted at every vertex is again a Galton–Watson tree with the same distribution.

In some cases, the construction is adapted to allow the root to have different degree and mark, but all other vertices have i.i.d. characteristics. As an example, we refer to [21], where PageRank on directed configuration model is investigated (in the independent case, see Section 6.1).

When we consider CTBPs, we have proved that the graph-normalized PageRank converges to the PageRank value of the root in a tree with distribution $\mathcal{T}(T_{\alpha^*})$. In particular, the processes $\{(\xi_t)^x_{t\geq 0}\}_{x\in\mathcal{U}}$ that define $\mathcal{T}(T_{\alpha^*})$ are i.i.d., but they are evaluated at random dependent times $(T_{\alpha^*}-\tau_x)_{x\in\mathcal{U}}$. Thus, the solution based on a weighted Galton–Watson tree does not apply to the PageRank in CTBPs, as the CTBP is inhomogeneous.

6.4. *Preferential attachment model.* Preferential attachment models (PAMs) are discrete-time dynamical models of random graphs. The main idea behind these models is the following: conditioning on the actual state of the graph, a new vertex is added with one (or more) edges, that are attached to existing vertices with probabilities proportional to their degree plus a constant.

There are different possible definitions of the model. See [15, 26] as well as [37], Chapter 8, for different definitions of the model. We consider a modification of the *sequential* PAM as presented in [15]. Fix $m\geq 1$ to be the initial degree of the vertices, and a constant $\delta>-m$. Then we define a sequence of graphs $(\mathrm{PA}_n(m,\delta))_{n\in\mathbb{N}}$ as follows:

(1) for $n=1$, $\mathrm{PA}_1(m,\delta)$ is composed by a single vertex with no edges;
(2) for $n=2$, $\mathrm{PA}_2(m,\delta)$ is composed by two vertices with m edges between them;
(3) for $n\geq 3$, $\mathrm{PA}_n(m,\delta)$ is defined recursively: we add a vertex to $\mathrm{PA}_{n-1}(m,\delta)$ with m edges. These m edges are attached to existing vertices with the following probability: for $l=1,\dots,m$,

$$\mathbb{P}\big(n \xrightarrow{l} i \mid \mathrm{PA}_{n-1,l-1}(m,\delta)\big)=\frac{D_i(n-1,l-1)+\delta}{2m(n-2)+(n-1)\delta+(l-1)}. \tag{74}$$

In (74), $D_i(n-1,l-1)$ denotes the degree of vertex i in the graph of size $n-1$ and after the first $l-1$ edges of the new vertex have been attached.

Notice that we allow for multiple edges but not for self-loops. In this case, we talk about PAM with *affine attachment rule*, since the attachment probabilities are proportional to an affine function of the degree. This model was first introduced in [1] for $\delta=0$. PAMs have gained a lot of attention in the last years since they show properties found in many real-world networks. In fact, PAMs shows a power-law degree distribution with exponent $\tau=3+\delta/m$ [37], Section 8.4, and they show the *small-world phenomenon*, that is, the typical distance and the diameter of the graph are small compared to the size of the graph itself [19, 26, 30].

It is known that CTBPs can embedd PAMs in continuous time [6, 7, 51]. We give a definition of the birth process that describes PAMs.

DEFINITION 6.8 (Embedding birth process). Fix $m\geq 2$ and $\delta>-m$. Consider the sequence $(k+1+\delta/m)_{k\in\mathbb{N}}$. Let $(E_k)_{k\in\mathbb{N}}$ be a sequence of independent and exponentially distributed random variables, with $E_k\sim E(k+1+\delta/m)$, and $E_{-1}=0$. We call $(\xi_t)_{t\geq 0}$ the *embedding birth process*, where $\xi_t=k$ if $t\in[E_{-1}+\cdots+E_{k-1}, E_{-1}+\cdots+E_k)$.

This construction is already used in [6, 7, 33, 51]. The embedding holds for any $m \geq 2$, but the topological description of the graph as a CTBP is used only in [33].

Originally defined as undirected graphs, PAMs have a natural direction from edges given by the recursive definition of such models. We can see every edge as directed from young to old, therefore every vertex in $\mathrm{PA}_t(m, \delta)$ has out-degree m. If we see the CTBP defined by the process in (6.8) as the continuous-time version of the PAM with out-degree 1, then the directed local weak limit is given by Proposition 6.6.

For $m \geq 2$, PAM is no longer a tree, making the analysis harder than the tree case. In [15], Berger, Borgs, Chayes and Saberi give the local weak limit in probability for the undirected version of PAM with affine attachment rule when $\delta \geq 0$. When $\delta \in (-m, 0)$, we believe that the result holds by adapting the proof in [15]. This argument is left for future work. We give a definition of the limiting graph for DPAMs.

DEFINITION 6.9 (Directed Pólya point graph). The *directed Pólya point graph* is an infinite marked rooted random tree constructed as follows: let $m \geq 2$ and $\delta > -m$ be parameters for a preferential attachment model $(\mathrm{PA}_t(m, \delta))_{t \in \mathbb{N}}$. Let

(a) $\chi = (m + \delta)/(2m + \delta)$, $\psi = (1 - \chi)/\chi$;
(b) Γ_{in} denote a Gamma distribution with parameters $m + \delta$ and 1.

Vertices in the graph have three characteristics:

(a) a *label* i in the Ulam–Harris set;
(b) a *position* $x \in [0, 1]$;
(c) a positive number γ called *strength*.

In addition, every vertex has mark m (in the sense of Definition 3.10). Assign to $\varnothing$ a position $x_\varnothing = U^\chi$, where U is a uniform random variable on $[0, 1]$, and a strength $\gamma_\varnothing \sim \Gamma_{\mathrm{in}}$. Set $\varnothing$ as unexplored. Then, recursively over the elements in the set of unexplored vertices, according to the shortlex order:

(1) let i denote the current unexplored vertex;
(2) assign to i a strength value $\gamma_i \sim \Gamma_{\mathrm{in}}$;
(3) let $u_{i1}, \ldots, u_{iD_i^{(\mathrm{in})}}$ be the random $D_i^{(\mathrm{in})}$ points given by an independent Poisson process on $[u_i, 1]$ with density

$$\rho_i(x) = \gamma_i \frac{\psi x^{\psi - 1}}{x_i^{\psi}}.$$

(4) draw an edge from each one of the vertices $i1, \ldots, iD_i^{(\mathrm{in})}$ to i;
(5) set $x_{i1}, \ldots, x_{iD_i^{(\mathrm{in})}}$ unexplored and i as explored.

Definition 6.9 is obtained by the definition of the undirected LW limit of PAM given in [15], Section 2.3.2, where the exploration of the neighborhood of a vertex is limited to the exploration of *younger vertices*. In other words, the exploration from a vertex i is made only over vertices with index $j > i$. The positions in Definition 6.9 encode the *age* of a vertex in PAM. In fact, it is possible to identify a vertex $i \in [t]$ in PAM with the point $(i/t)^\chi$ [15], Lemma 3.1, so old vertices have position closer to 0 than young vertices.

With the definition of the Directed Pólya point graph, we can state the directed LWC result for PAMs.

PROPOSITION 6.10 (LW limit of directed PAM). *Fix $m \geq 2$ and $\delta > -m$. Let* $(\mathrm{PA}_t(m, \delta))_{t \in \mathbb{N}}$ *be a PAM defined by the attachment rule in* (74). *Denote by* $(\mathrm{DPA}_t(m, \delta))_{t \in \mathbb{N}}$

the directed version of $(\mathrm{PA}_t(m,\delta))_{t\in\mathbb{N}}$, *where edges are directed from young to old vertices. Then*:

(1) *for* $\delta \geq 0$, $\mathrm{DPA}_t(m,\delta)$ converges in probability in the directed LW sense *to the directed Pólya point graph as in Definition* 6.9;

(2) *for* $\delta \in (-m, 0)$, *if* [15], *Theorem* 2.2, *can be extended*, *then the convergence holds also in this case*.

The proof of Proposition 6.10 follows immediately from [15], Theorem 2.2, and the fact that the exploration process in DPA_t corresponds to exploring only younger vertices.

REMARK 6.11 (Nonrecursive property of PageRank). Similar to Remark 6.7 about CTBPs, we point out that the PageRank value of the root of a directed Pólya point graph does not satisfy the recursive property that is necessary to consider it as a solution of (4). Notice that the Poisson point process assigned to vertex i in Definition 6.9 is defined on the interval $[x_i, 1]$, where the position x_i depends on the ancestors (in the Ulam–Harris sense) of i.

Another way to interpret this is that the family of Poisson point process in Definition 6.9 is composed by i.i.d. processes parametrized by the positions of vertices, that are dependent random variables. This suggests that the positions in the Pólya point graph play the same role as the birth times in CTBPs.

Acknowledgements. We thank the referee for very valuable comments that helped to significantly improve the paper, especially Section 5.

This work was supported in part by the Netherlands Organisation for Scientific Research (NWO) through the Gravitation NETWORKS grant 024.002.003.

The second author was supported by the Netherlands Organisation for Scientific Research (NWO) through VICI grant 639.033.806.

REFERENCES

[1] ALBERT, R. and BARABÁSI, A.-L. (2002). Statistical mechanics of complex networks. *Rev. Modern Phys.* **74** 47–97. MR1895096 https://doi.org/10.1103/RevModPhys.74.47

[2] ALDOUS, D. (1991). Asymptotic fringe distributions for general families of random trees. *Ann. Appl. Probab.* **1** 228–266. MR1102319

[3] ALDOUS, D. and LYONS, R. (2007). Processes on unimodular random networks. *Electron. J. Probab.* **12** 1454–1508. MR2354165 https://doi.org/10.1214/EJP.v12-463

[4] ALDOUS, D. and STEELE, J. M. (2004). The objective method: Probabilistic combinatorial optimization and local weak convergence. In *Probability on Discrete Structures. Encyclopaedia Math. Sci.* **110** 1–72. Springer, Berlin. MR2023650 https://doi.org/10.1007/978-3-662-09444-0_1

[5] ANDERSEN, R., CHUNG, F. and LANG, K. (2006). Local graph partitioning using PageRank vectors. In *Proceedings of the 47th Annual IEEE Symposium on Foundations of Computer Science* 475–486.

[6] ATHREYA, K. B. (2007). Preferential attachment random graphs with general weight function. *Internet Math.* **4** 401–418. MR2522950

[7] ATHREYA, K. B., GHOSH, A. P. and SETHURAMAN, S. (2008). Growth of preferential attachment random graphs via continuous-time branching processes. *Proc. Indian Acad. Sci. Math. Sci.* **118** 473–494. MR2450248 https://doi.org/10.1007/s12044-008-0036-2

[8] ATHREYA, K. B. and NEY, P. E. (2004). *Branching Processes*. Dover, Mineola, NY. Reprint of the 1972 original [Springer, New York; MR0373040]. MR2047480

[9] AVRACHENKOV, K., KADAVANKANDY, A., PROKHORENKOVA, L. O. and RAIGORODSKII, A. (2015). PageRank in undirected random graphs. In *Algorithms and Models for the Web Graph. Lecture Notes in Computer Science* **9479** 151–163. Springer, Cham. MR3500675 https://doi.org/10.1007/978-3-319-26784-5_12

[10] AVRACHENKOV, K. and LEBEDEV, D. (2006). PageRank of scale-free growing networks. *Internet Math.* **3** 207–231. MR2321830

[11] AVRACHENKOV, K. and LITVAK, N. (2006). The effect of new links on Google PageRank. *Stoch. Models* **22** 319–331. MR2220968 https://doi.org/10.1080/15326340600649052

[12] BAHMANI, B., CHOWDHURY, A. and GOEL, A. (2010). Fast incremental and personalized PageRank. *Proc. VLDB Endow.* **4** 173–184.

[13] BENJAMINI, I., LYONS, R. and SCHRAMM, O. (2015). Unimodular random trees. *Ergodic Theory Dynam. Systems* **35** 359–373. MR3316916 https://doi.org/10.1017/etds.2013.56

[14] BENJAMINI, I. and SCHRAMM, O. (2001). Recurrence of distributional limits of finite planar graphs. *Electron. J. Probab.* **6** no. 23, 13. MR1873300 https://doi.org/10.1214/EJP.v6-96

[15] BERGER, N., BORGS, C., CHAYES, J. T. and SABERI, A. (2014). Asymptotic behavior and distributional limits of preferential attachment graphs. *Ann. Probab.* **42** 1–40. MR3161480 https://doi.org/10.1214/12-AOP755

[16] BIANCHINI, M., GORI, M. and SCARSELLI, F. (2005). Inside PageRank. *ACM Trans. Internet Technol.* **5** 92–128.

[17] BRITTON, T., DEIJFEN, M. and MARTIN-LÖF, A. (2006). Generating simple random graphs with prescribed degree distribution. *J. Stat. Phys.* **124** 1377–1397. MR2266448 https://doi.org/10.1007/s10955-006-9168-x

[18] CAO, J. and OLVERA-CRAVIOTO, M. (2017). On a general class of inhomogeneous random digraphs. Preprint. Available at arXiv:1712.03319 [math.PR].

[19] CARAVENNA, F., GARAVAGLIA, A. and VAN DER HOFSTAD, R. (2019). Diameter in ultra-small scale-free random graphs. *Random Structures Algorithms* **54** 444–498. MR3938775 https://doi.org/10.1002/rsa.20798

[20] CHEN, N., LITVAK, N. and OLVERA-CRAVIOTO, M. (2014). PageRank in scale-free random graphs. In *Algorithms and Models for the Web Graph. Lecture Notes in Computer Science* **8882** 120–131. Springer, Cham. MR3297352 https://doi.org/10.1007/978-3-319-13123-8_10

[21] CHEN, N., LITVAK, N. and OLVERA-CRAVIOTO, M. (2017). Generalized PageRank on directed configuration networks. *Random Structures Algorithms* **51** 237–274. MR3683363 https://doi.org/10.1002/rsa.20700

[22] CHEN, N. and OLVERA-CRAVIOTO, M. (2013). Directed random graphs with given degree distributions. *Stoch. Syst.* **3** 147–186. MR3353470 https://doi.org/10.1214/12-SSY076

[23] CHEN, P., XIE, H., MASLOV, S. and REDNER, S. (2007). Finding scientific gems with Google's PageRank algorithm. *J. Informetr.* **1** 8–15.

[24] CHUNG, F. and LU, L. (2001). The diameter of sparse random graphs. *Adv. in Appl. Math.* **26** 257–279. MR1826308 https://doi.org/10.1006/aama.2001.0720

[25] CHUNG, F. and LU, L. (2002). Connected components in random graphs with given expected degree sequences. *Ann. Comb.* **6** 125–145. MR1955514 https://doi.org/10.1007/PL00012580

[26] DEREICH, S., MÖNCH, C. and MÖRTERS, P. (2012). Typical distances in ultrasmall random networks. *Adv. in Appl. Probab.* **44** 583–601. MR2977409 https://doi.org/10.1239/aap/1339878725

[27] DEREICH, S. and MÖRTERS, P. (2009). Random networks with sublinear preferential attachment: Degree evolutions. *Electron. J. Probab.* **14** 1222–1267. MR2511283 https://doi.org/10.1214/EJP.v14-647

[28] DEREICH, S. and MÖRTERS, P. (2011). Random networks with concave preferential attachment rule. *Jahresber. Dtsch. Math.-Ver.* **113** 21–40. MR2760002 https://doi.org/10.1365/s13291-010-0011-6

[29] DEREICH, S. and MÖRTERS, P. (2013). Random networks with sublinear preferential attachment: The giant component. *Ann. Probab.* **41** 329–384. MR3059201 https://doi.org/10.1214/11-AOP697

[30] DOMMERS, S., VAN DER HOFSTAD, R. and HOOGHIEMSTRA, G. (2010). Diameters in preferential attachment models. *J. Stat. Phys.* **139** 72–107. MR2602984 https://doi.org/10.1007/s10955-010-9921-z

[31] VAN DEN ESKER, H., VAN DER HOFSTAD, R. and HOOGHIEMSTRA, G. (2008). Universality for the distance in finite variance random graphs. *J. Stat. Phys.* **133** 169–202. MR2438903 https://doi.org/10.1007/s10955-008-9594-z

[32] GABRYSCH, K. (2016). Convergence of directed random graphs to the Poisson-weighted infinite tree. *J. Appl. Probab.* **53** 463–474. MR3514291 https://doi.org/10.1017/jpr.2016.13

[33] GARAVAGLIA, A. and VAN DER HOFSTAD, R. (2018). From trees to graphs: Collapsing continuous-time branching processes. *J. Appl. Probab.* **55** 900–919. MR3877316 https://doi.org/10.1017/jpr.2018.57

[34] GARAVAGLIA, A., VAN DER HOFSTAD, R. and WOEGINGER, G. (2017). The dynamics of power laws: Fitness and aging in preferential attachment trees. *J. Stat. Phys.* **168** 1137–1179. MR3691245 https://doi.org/10.1007/s10955-017-1841-8

[35] GRIMMETT, G. R. (1980/81). Random labelled trees and their branching networks. *J. Aust. Math. Soc. A* **30** 229–237. MR0607933

[36] HAVELIWALA, T. H. (2002). Topic-sensitive PageRank. In *Proceedings of the* 11*th International Conference on World-Wide Web* 517–526.

[37] VAN DER HOFSTAD, R. (2017). *Random Graphs and Complex Networks*. *Vol*. 1. *Cambridge Series in Statistical and Probabilistic Mathematics* **43**. Cambridge Univ. Press, Cambridge. MR3617364 https://doi.org/10.1017/9781316779422

[38] VAN DER HOFSTAD, R. (2019+). *Random Graphs and Complex Networks*. *Vol*. 2. In preparation. Available at http://www.win.tue.nl/~rhofstad/NotesRGCNII.pdf.

[39] VAN DER HOFSTAD, R. (2019+). Stochastic Processes on Random Graphs. Lecture Notes for the Saint-Flour Summer School in Probability 2017. Available at https://www.win.tue.nl/~rhofstad/SaintFlour_SPoRG.pdf.

[40] HOLMGREN, C. and JANSON, S. (2017). Fringe trees, Crump-Mode-Jagers branching processes and m-ary search trees. *Probab. Surv.* **14** 53–154. MR3626585 https://doi.org/10.1214/16-PS272

[41] JAGERS, P. (1975). *Branching Processes with Biological Applications*. *Wiley Series in Probability and Mathematical Statistics—Applied Probability and Statistics*. Wiley Interscience, London. MR0488341

[42] JAGERS, P. and NERMAN, O. (1984). The growth and composition of branching populations. *Adv. in Appl. Probab.* **16** 221–259. MR0742953 https://doi.org/10.2307/1427068

[43] JELENKOVIĆ, P. R. and OLVERA-CRAVIOTO, M. (2010). Information ranking and power laws on trees. *Adv. in Appl. Probab.* **42** 1057–1093. MR2796677 https://doi.org/10.1239/aap/1293113151

[44] LEE, C. P. C., GOLUB, G. H. and ZENIOS, S. A. (2003). A fast two-stage algorithm for computing PageRank and its extensions. *Sci. Comput. Comput. Math.* **1** 1–9.

[45] LEE, J. and OLVERA-CRAVIOTO, M. (2017). PageRank on inhomogeneous random digraphs. Preprint. Available at arXiv:1707.02492 [math.PR].

[46] LITVAK, N., SCHEINHARDT, W. R. W. and VOLKOVICH, Y. (2007). In-degree and PageRank: Why do they follow similar power laws? *Internet Math.* **4** 175–198. MR2522875

[47] MA, N., GUAN, J. and ZHAO, Y. (2008). Bringing PageRank to the citation analysis. *Inf. Process. Manag.* **44** 800–810.

[48] NERMAN, O. (1981). On the convergence of supercritical general (C-M-J) branching processes. *Z. Wahrsch. Verw. Gebiete* **57** 365–395. MR0629532 https://doi.org/10.1007/BF00534830

[49] PAGE, L., BRIN, S., MOTWANI, R. and WINOGRAD, T. (1999). The PageRank citation ranking: Bringing order to the web. Stanford InfoLab, 1999.

[50] PANDURANGAN, G., RAGHAVAN, P. and UPFAL, E. (2002). Using PageRank to characterize Web structure. In *Computing and Combinatorics*. *Lecture Notes in Computer Science* **2387** 330–339. Springer, Berlin. MR2064528 https://doi.org/10.1007/3-540-45655-4_36

[51] RUDAS, A., TÓTH, B. and VALKÓ, B. (2007). Random trees and general branching processes. *Random Structures Algorithms* **31** 186–202. MR2343718 https://doi.org/10.1002/rsa.20137

[52] VOLKOVICH, Y. and LITVAK, N. (2010). Asymptotic analysis for personalized web search. *Adv. in Appl. Probab.* **42** 577–604. MR2675117 https://doi.org/10.1239/aap/1275055243

[53] WALTMAN, L. and VAN ECK, N. J. (2010). The relation between eigenfactor, audience factor, and influence weight. *J. Am. Soc. Inf. Sci. Technol.* **61** 1476–1486.

[54] WANG, R., ZHANG, W., DENG, H., WANG, N., MIAO, Q. and ZHAO, X. (2013). Discover community leader in social network with PageRank. In *Advances in Swarm Intelligence*: 4*th International Conference* 154–162.