Barak-Erdös graphs and the infinite-bin model

Bastien Mallein, Sanjay Ramassamy

HAL Id: ensl-01651022

https://ens-lyon.hal.science/ensl-01651022v1

Submitted on 28 Nov 2017 (v1), last revised 27 Oct 2022 (v2)

Barak-Erdős graphs and the infinite-bin model

Bastien Mallein   Sanjay Ramassamy

November 12, 2017

Abstract

A Barak-Erdős graph is a directed acyclic version of an Erdős-Rényi graph. It is obtained by performing independent bond percolation with parameter p on the complete graph with vertices $\{1, ..., n\}$, in which the edge between two vertices $i < j$ is directed from i to j. The length of the longest path in this graph grows linearly with the number of vertices, at rate $C(p)$. In this article, we use a coupling between Barak-Erdős graphs and infinite-bin models to provide explicit estimates on $C(p)$. More precisely, we prove that the front of an infinite-bin model grows at linear speed, and that this speed can be obtained as the sum of a series. Using these results, we prove the analyticity of C for $p > 1/2$, and compute its power series expansion. We also obtain the first two terms of the asymptotic expansion of C as $p \to 0$, using a coupling with branching random walks.

1   Introduction

Random graphs and interacting particle systems have been two active fields of research in probability in the past decades. In 2003, Foss and Konstantopoulos [11] introduced a new interacting particle system called the infinite-bin model and established a correspondence between a certain class of infinite-bin models and Barak-Erdős random graphs, which are a directed acyclic version of Erdős-Rényi graphs.

In this article, we study the speed at which the front of an infinite-bin model drifts to infinity. These results are applied to obtain a fine asymptotic of the length of the longest path in a Barak-Erdős graph. In the remainder of the introduction, we first describe Barak-Erdős graphs, then infinite-bin models. We then state our main results on infinite-bin models, and their consequences for Barak-Erdős graphs.

1.1   Barak-Erdős graphs

Barak and Erdős introduced in [3] the following model of a random directed graph with vertex set $\{1, \dots, n\}$ (which we refer to as Barak-Erdős graphs from now on) : for each pair of vertices $i < j$, add an edge directed from i to j with probability p, independently for each pair. They were interested in the maximal size of strongly independent sets in such graphs.

However, one of the most widely studied properties of Barak-Erdős graphs has been the length of its longest path. It has applications to mathematical

ecology (food chains) [9, 22], performance evaluation of computer systems (speed of parallel processes) [14, 15] and queuing theory (stability of queues) [11].

Newman [21] studied the length of the longest path in Barak-Erdős graphs in several settings, when the edge probability p is constant (dense case), but also when it is of the form c_n/n with $c_n = o(n)$ (sparse case). In the dense case, he proved that when n gets large, the length of the longest path $L_n(p)$ grows linearly with n in the first-order approximation :

$$\lim_{n \to \infty} \frac{L_n(p)}{n} = C(p) \text{ a.s.,} \qquad (1.1)$$

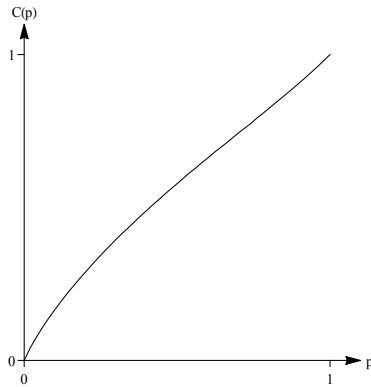where the linear growth rate C is a function of p. We plot in Figure 1 an approximation of $C(p)$.

Figure 1: Plot of a an approximation of $C(p)$, using $600,000$ iterations of an infinite-bin model, for values of p that are integer multiples of 0.02.

Newman proved that the function C is continuous and computed its derivative at $p = 0$. Foss and Konstantopoulos [11] studied Barak-Erdős graphs under the name of "stochastic ordered graphs" and provided upper and lower bounds for C, obtaining in particular that

$$C(1 - q) = 1 - q + q^2 - 3q^3 + 7q^4 + O(q^5) \text{ when } q \to 0,$$

where $q = 1 - p$ denotes the probability of the absence of an edge.

Denisov, Foss and Konstantopoulos [10] introduced the more general model of a directed slab graph and proved a law of large numbers and a central limit theorem for the length of its longest path. Konstantopoulos and Trinajstić [17] looked at a directed random graph with vertices in $\mathbb{Z}^2$ (instead of $\mathbb{Z}$ for the infinite version of Barak-Erdős graphs) and identified fluctuations following the Tracy-Widom distribution. Foss, Martin and Schmidt [12] added to the original Barak-Erdős model random edge lengths, in which case the problem of the longest path can be reformulated as a last-passage percolation question. Gelenbe, Nelson, Philips and Tantawi [14] studied a similar problem, but with random weights on the vertices rather than on the edges.

Ajtai, Komlós and Szemerédi [1] studied the asymptotic behaviour of the longest path in sparse Erdős-Rényi graphs, which are the undirected version of Barak-Erdős graphs.

1.2 The infinite-bin model

Foss and Konstantopoulos introduced the infinite-bin model in [11] as an interacting particle system which, for a right choice of parameters, gives information about the growth rate $C(p)$ of the longest path in Barak-Erdős graphs. Consider a set of bins indexed by the set of integers $\mathbb{Z}$. Each bin may contain any number of balls, finite or infinite. A configuration of balls in bins is called *admissible* if there exists $m \in \mathbb{Z}$ such that:

1. every bin with an index smaller or equal to m is non-empty ;

2. every bin with an index strictly larger than m is empty.

The largest index of a non-empty bin m is called the position of the *front*. From now on, all configurations will implicitly be assumed to be admissible. Given an integer $k \geq 1$, we define the *move of type k* as a map Φ_k from the set of configurations to itself. Given an initial configuration X, $\Phi_k(X)$ is obtained by adding one ball to the bin of index $b_k + 1$, where b_k is the index of the bin containing the k-th ball of X (the balls are counted from right to left, starting from the rightmost nonempty bin).
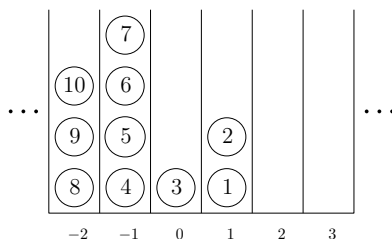
(a) A configuration X, the numbers inside the balls indicate how they are counted from right to left.
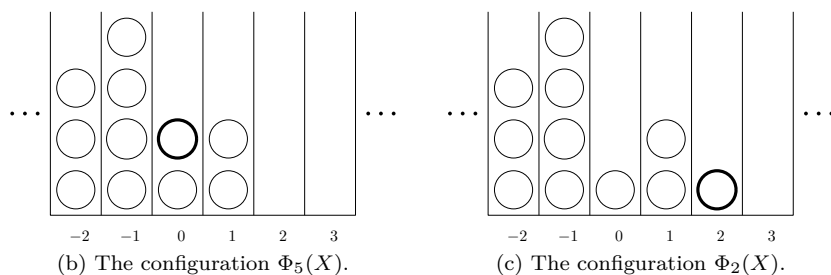
(b) The configuration $\Phi_5(X)$.

(c) The configuration $\Phi_2(X)$.

Figure 2: Action of two moves on a configuration.

Given a probability distribution μ on the set of positive integers and an initial configuration X_0, one defines the Markovian evolution of the infinite-bin model with distribution μ (or IBM(μ) for short) as the following stochastic recursive sequence:

$$X_{n+1} = \Phi_{\xi_{n+1}}(X_n) \text{ for } n \geq 0,$$

where $(\xi_n)_{n \geq 1}$ is an i.i.d. sequence of law μ. We prove in Theorem 1.1 that the front moves to the right at a speed which tends a.s. to a constant limit v_μ. We call v_μ the *speed* of the IBM(μ). Note that the model defined in [11] was slightly

more general, allowing $(\xi_n)_{n\geq 1}$ to be a stationary-ergodic sequence. We also do not adopt their convention of shifting the indexing of the bins which forces the front to always be at position 0.

Foss and Konstantopoulos [11] proved that if μ_p was the geometric distribution of parameter p then $v_{\mu_p} = C(p)$, where $C(p)$ is the growth rate of the length of the longest path in Barak-Erdős graphs with edge probability p. They also proved, for distributions μ with finite mean verifying $\mu(\{1\}) > 0$, the existence of renovations events, which yields a functional law of large numbers and central limit theorem for the IBM(μ). Based on a coupling result for the infinite-bin model obtained by Chernysh and Ramassamy [8], Foss and Zachary [13] managed to remove the condition $\mu(\{1\}) > 0$ required by [11] to obtain renovation events.

Aldous and Pitman [2] had already studied a special case of the infinite-bin model, namely what happens to the speed of the front when μ is the uniform distribution on $\{1, \ldots, n\}$, in the limit when n goes to infinity. They were motivated by an application to the running time of local improvement algorithms defined by Tovey [24].

1.3   Speed of infinite-bin models

The remainder of the introduction is devoted to the presentation of the main results proved in this paper. In this section we state the results related to general infinite-bin models, and in the next one we state the results related to the Barak-Erdős graphs.

We first prove that in every infinite-bin model, the front moves at linear speed. Foss and Konstantopoulos [11] had derived a special case of this result, when the distribution μ has finite expectation.

Theorem 1.1. *Let (X_n) be an infinite-bin model with distribution μ, starting from an admissible configuration X_0. For any $n \in \mathbb{N}$, we write M_n for the position of the front of X_n. There exists $v_\mu \in [0,1]$, depending only on the distribution μ, such that*

$$\lim_{n\to+\infty} \frac{M_n}{n} = v_\mu \quad a.s.$$

In the next result, we obtain an explicit formula for the speed v_μ of the IBM(μ), as a series. To give this formula we first introduce some notation. Recalling that $\mathbb{N}$ is the set of positive integers, we denote by $\mathcal{A}$ the set of words on the alphabet $\mathbb{N}$, *i.e.* the set of all finite-length sequences of elements of $\mathbb{N}$. Given a non-empty word $\alpha \in \mathcal{A}$, written $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ (where the α_i are the letters of α), we denote by $L(\alpha) = n$ the length of α. The empty word is denoted by $\emptyset$.

Fix an infinite-bin model configuration X. We define the subset $\mathcal{P}_X$ of $\mathcal{A}$ as follows: a word α belongs to $\mathcal{P}_X$ if it is non-empty, and if starting from configuration X and applying successively the moves $\Phi_{\alpha_1}, \ldots, \Phi_{\alpha_n}$, the last move Φ_{α_n} results in placing a ball in a previously empty bin.

Given a word $\alpha \in \mathcal{A}$ which is not the empty word, we set $\varpi\alpha \in \mathcal{A}$ to be the word obtained from α by removing the first letter. We also set $\varpi\emptyset = \emptyset$. We define the function $\varepsilon_X : \mathcal{A} \to \{-1, 0, 1\}$ as follows:

$$\varepsilon_X(\alpha) = \mathbf{1}_{\{\alpha \in \mathcal{P}_X\}} - \mathbf{1}_{\{\varpi\alpha \in \mathcal{P}_X\}}.$$

4

Theorem 1.2. *Let X be an admissible configuration and μ a probability distribution on $\mathbb{N}$. We define the weight of a word $\alpha = (\alpha_1, \ldots, \alpha_n)$ by*

$$W_\mu(\alpha) = \prod_{i=1}^{n} \mu(\{\alpha_i\}) = \mathbf{P}(\alpha = (\xi_1, \ldots \xi_n)).$$

If $\sum_{\alpha \in \mathcal{A}} |\varepsilon_X(\alpha)| W_\mu(\alpha) < +\infty$, then

$$v_\mu = \sum_{\alpha \in \mathcal{A}} \varepsilon_X(\alpha) W_\mu(\alpha). \tag{1.2}$$

Remark 1.3. One of the most striking features of (1.2) is that whereas for any $\alpha \in \mathcal{A}$, $X \mapsto \varepsilon_X(\alpha)$ is a non-constant function of X, v_μ does not depend of this choice of configuration. As a result, Theorem 1.2 gives in fact an infinite number of formulas for the speed v_μ of the IBM(μ).

Theorem 1.2 can be extended to prove the following result:

$$v_\mu = \lim_{n \to +\infty} \frac{1}{n} \sum_{k=1}^{n} \sum_{\alpha \in \mathcal{A}: L(\alpha) \leq k} \varepsilon_X(\alpha) W_\mu(\alpha).$$

In other words, if we define $\sum_{\alpha \in \mathcal{A}} \varepsilon_X(\alpha) W_\mu(\alpha)$ as the Cesàro mean of its partial sums (on words of finite length), (1.2) holds for any probability distribution μ and admissible configuration X.

1.4 Longest increasing paths in Barak-Erdős graphs

Using the coupling introduced by Foss and Konstantopoulos between Barak-Erdős graphs and infinite-bin models, we use the previous results to extract information on the function C defined in (1.1). Firstly, we prove that for p large enough (i.e. for dense Barak-Erdős graphs), the function C is analytic and we obtain the power series expansion of $C(p)$ centered at 1. Secondly, we provide the first two terms of the asymptotic expansion of $C(p)$ as $p \to 0$.

We deduce from Theorem 1.2 the analyticity of $C(p)$ for p close to 1. For any word $\alpha \in \mathcal{A}$, we define the height of α to be

$$H(\alpha) = \sum_{i=1}^{L(\alpha)} \alpha_i - L(\alpha).$$

For any $k \in \mathbb{N}$ and admissible configuration X, we set

$$a_k = \sum_{\alpha \in \mathcal{A}: H(\alpha) \leq k, L(\alpha) \leq k+1} \varepsilon_X(\alpha)(-1)^{k-H(\alpha)} \binom{L(\alpha)}{k - H(\alpha)}. \tag{1.3}$$

Theorem 1.4. *The function C is analytic on $\left(\frac{1}{2}, 1\right]$ and for $p \in \left(\frac{3-\sqrt{2}}{2}, 1\right]$,*

$$C(p) = \sum_{k \geq 0} a_k (1-p)^k.$$

Similarly to what has been observed in Remark 1.3, this result proves that the value of a_k does not depend on the configuration X, justifying *a posteriori* the notation.

We do not believe the bounds $\frac{1}{2}$ and $\frac{3-\sqrt{2}}{2}$ to be optimal, see Remark 1.7. They are obtained using very rough bounds on the function ε_X.

Remark 1.5. Using (1.3) and Lemma 6.2, it is possible to explicitly compute as many coefficients of the power series expansion as desired, by picking a configuration X and computing quantities of the form $\varepsilon_X(\alpha)$ for finitely many words $\alpha \in \mathcal{A}$. For example, we observe that as $q \to 0$,

$$C(1-q) = 1 - q + q^2 - 3q^3 + 7q^4 - 15q^5 + 29q^6 - 54q^7 + 102q^8 + O(q^9).$$

It is clear from formula (1.3) that (a_k) is integer-valued. Based on our computations, we conjecture that $((-1)^k a_k, k \geq 0)$ is non-negative and non-decreasing.

We now turn to the asymptotic behaviour of $C(p)$ as $p \to 0$, i.e. the length of the longest increasing path in sparse Barak-Erdős graphs. We improve the result obtained by Newman [21].

Theorem 1.6. *We have $C(p) = ep - \dfrac{p\pi^2 e}{2(\log p)^2} + o(p(\log p)^{-2})$.*

In particular, we note that $C(p)$ does not have a finite second derivative at $p = 0$.

Remark 1.7. Numerical simulations tend to suggest that the power series expansion of $C(p)$ at $p = 1$ has a radius of convergence larger than 0.5 but smaller than 1. Together with the fact that C admits no second derivative at $p = 0$, this raises the question of the existence of a phase transition in this process.

Theorem 1.6 is obtained by coupling the infinite-bin model with uniform distribution with a continuous-time branching random walk with selection (as observed by Aldous and Pitman [2]) and by extending to the continuous-time setting the results of Bérard and Gouéré [4] on the asymptotic behaviour of a discrete-time branching random walk. Assuming that a conjecture of Brunet and Derrida [7] on the speed of a branching random walk with selection holds, the next term in the asymptotic expansion should be given by $\frac{3e\pi^2 p \log(-\log p)}{(-\log p)^3}$.

Remark 1.8. With arguments similar to the ones used to prove Theorem 1.6, we expect that one can also obtain the asymptotic behaviour of $L_n(p)$ as $n \to +\infty$ and $p \to 0$ simultaneously, proving that:

$$L_n(p_n) = nep_n - n\frac{p_n\pi^2 e}{2(\log p_n)^2} + o(np_n(\log p_n)^{-2}) \text{ in probability,}$$

as long as $p_n \gg \frac{(\log n)^3}{n}$. We expect a different behaviour if $p_n \sim \lambda\frac{(\log n)^3}{n}$.

Organisation of the paper

We state more precisely the notation used to study the infinite-bin model in Section 2. We also introduce an increasing coupling between infinite-bin models, which is a key result for the rest of the article.

In Section 3, we prove that the speed of an infinite-bin model with a measure of finite support can be expressed using the invariant measure of a finite Markov

chain. This result is then used to prove Theorem 1.1 in the general case. We prove Theorem 1.2 in Section 4 using a method akin to "exact perturbative expansion".

We review in Section 5 the Foss-Konstantopoulos coupling between Barak-Erdős graphs and the infinite-bin model and use it to provide a sequence of upper and lower bounds converging exponentially fast to $C(p)$. This coupling is used in Section 6, where we prove Theorem 1.4 using Theorem 1.2. Finally, we prove Theorem 1.6 in Section 7, by extending the results of Bérard and Gouéré [4] to compute the asymptotic behaviour of a continuous-time branching random walk with selection.

2 Basic properties of the infinite-bin model

We write $\mathbb{N}$ for the set of positive integers, $\overline{\mathbb{N}} = \mathbb{N} \cup \{+\infty\}$, $\mathbb{Z}_+$ for the set of non-negative integers and $\overline{\mathbb{Z}}_+ = \mathbb{Z}_+ \cup \{+\infty\}$. We denote by

$$S = \left\{ X \in (\overline{\mathbb{Z}}_+)^{\mathbb{Z}} : \begin{array}{l} \exists m \in \mathbb{Z} : \forall j \in \mathbb{Z}, X(j) = 0 \iff j > m \quad \text{and} \\ \forall j \in \mathbb{Z}, X(j) = +\infty \Rightarrow X(j-1) = +\infty \end{array} \right\}$$

the set of admissible configurations for an infinite-bin model. Note that the definition we use here is more restrictive than the one used, as a simplification, in the introduction. Indeed, we impose here that if a bin has an infinite number of balls, every bin to its left also has an infinite number of balls. However, this has no impact on our results, as the dynamics of an infinite-bin model does not affect bins to the left of a bin with an infinite number of balls. One does not create balls in a bin at distance greater than 1 from a non-empty bin.

We wish to point out that our definition of admissible configurations has been chosen out of convenience. Most of the results of this article could easily be generalized to infinite-bin models with a starting configuration belonging to

$$S^0 = \left\{ X \in (\overline{\mathbb{Z}}_+)^{\mathbb{Z}} : \lim_{k \to +\infty} X(k) = 0 \text{ and } \sum_{k \in \mathbb{Z}} X(k) = +\infty \right\},$$

see e.g. Remark 3.7. They could even be generalized to configurations starting with a finite number of balls, if we adapt the dynamics of the infinite-bin model as follows. For any $n \in \mathbb{N}$, if ξ_n is larger than the number of balls existing at time n, then the step is ignored and the IBM configuration is not modified. However, with this definition some trivial cases might arise, for example starting with a configuration with only one ball, and using a measure μ with $\mu(\{1\}) = 0$.

For any $X \in S$ and $k \in \mathbb{Z}$, we call $X(k)$ the number of balls at position k in the configuration X. Observe that the set of non-empty bins is a semi-infinite interval of $\mathbb{Z}$. In particular, for any $X \in S$, there exists a unique integer $m \in \mathbb{Z}$ such that $X(m) \neq 0$ and $X(j) = 0$ for all $j > m$. The integer m is called the *front* of the configuration.

Let $X \in S$, $k \in \mathbb{Z}$ and $\xi \in \mathbb{N}$. We denote by

$$N(X, k) = \sum_{j=k}^{+\infty} X(j) \quad \text{and} \quad B(X, \xi) = \inf\{j \in \mathbb{Z} : N(X, j) < \xi\}$$

the number of balls to the right of k and the leftmost position such that there are less than ξ balls to its right respectively. Note that the position of the front in the configuration X is given by $B(X, 1) - 1$. Observe that for any $X \in S$,

$$\forall 1 \leq \xi \leq \xi', \ 0 \leq B(X, \xi) - B(X, \xi') \leq \xi' - \xi. \tag{2.1}$$

For $\xi \in \mathbb{N}$ and $X \in S$, we set $\Phi_\xi(X) = \left(X(j) + \mathbf{1}_{\{j = B(X, \xi)\}}, j \in \mathbb{Z} \right)$ the transformation that adds one ball to the right of the ξ-th largest ball in X. We extend the notation to allow $\xi \in \overline{\mathbb{N}}$, by setting $\Phi_\infty(X) = X$. We also introduce the shift operator $\tau(X) = (X(j-1), j \in \mathbb{Z})$. We observe that τ and Φ_ξ commute, i.e.

$$\forall X \in S, \forall \xi \in \overline{\mathbb{N}}, \Phi_\xi(\tau(X)) = \tau(\Phi_\xi(X)). \tag{2.2}$$

Recall that an infinite-bin model consists in the sequential application of randomly chosen transformations Φ_ξ, called move of type ξ. More precisely, given μ a probability measure on $\overline{\mathbb{N}}$ and $(\xi_n, n \geq 1)$ i.i.d. random variables with distribution μ, the IBM(μ) (X_n) is the Markov process on S starting from $X_0 \in S$, such that for any $n \geq 0$, $X_{n+1} = \Phi_{\xi_{n+1}}(X_n)$.

We introduce a partial order on S, which is compatible with the infinite-bin model dynamics: for any $X, Y \in S$, we write

$$X \preccurlyeq Y \iff \forall j \in \mathbb{Z}, N(X, j) \leq N(Y, j) \iff \forall \xi \in \mathbb{N}, B(X, \xi) \leq B(Y, \xi).$$

The functions (Φ_ξ) are monotone, increasing in X and decreasing in ξ for this partial order. More precisely

$$\forall X \preccurlyeq Y \in S, \ \forall 1 \leq \xi \leq \xi' \leq \infty, \ \Phi_{\xi'}(X) \preccurlyeq \Phi_\xi(Y). \tag{2.3}$$

Moreover, the shift operator τ dominates every function Φ_ξ, i.e.

$$\forall X \preccurlyeq Y \in S, \ \forall 1 \leq \xi \leq \infty, \ \Phi_\xi(X) \preccurlyeq \tau(Y). \tag{2.4}$$

As a consequence, infinite-bin models can be coupled in an increasing fashion.

Proposition 2.1. *Let μ and ν be two probabilities on $\overline{\mathbb{N}}$, and $X_0 \preccurlyeq Y_0 \in S^0$. If $\mu([1, k]) \leq \nu([1, k])$ for any $k \in \mathbb{N}$, we can couple the IBM(μ) (X_n) and the IBM(ν) (Y_n) such that for any $n \geq 0$, $X_n \preccurlyeq Y_n$ a.s.*

Proof. As for any $k \in \mathbb{N}$, $\mu([1, k]) \leq \nu([1, k])$, we can construct a couple (ξ, ζ) such that ξ has law μ, ζ has law ν and $\xi \geq \zeta$ a.s. Let (ξ_n, ζ_n) be i.i.d. copies of (ξ, ζ), we set $X_{n+1} = \Phi_{\xi_{n+1}}(X_n)$ and $Y_{n+1} = \Phi_{\zeta_{n+1}}(Y_n)$. By induction, using (2.3), we immediately have $X_n \preccurlyeq Y_n$ for any $n \geq 0$. $\qquad\square$

We extended in this section the definition of the IBM(μ) to measures with positive mass on $\{\infty\}$. As applying Φ_∞ does not modify the ball configuration, the IBM(μ) and the IBM$(\mu(.\,|. < \infty))$ are straightforwardly connected.

Lemma 2.2. *Let μ be a probability measure on $\overline{\mathbb{N}}$ with $p := \mu(\{\infty\}) < 1$. We write ν for the measure verifying $\nu(\{k\}) = \frac{\mu(\{k\})}{1-p}$. Let (X_n) be an IBM(ν) and (S_n) be an independent random walk with step distribution Bernoulli with parameter $1 - p$. Then the process $(X_{S_n}, n \geq 0)$ is an IBM(μ).*

In particular, assuming Theorem 1.1 holds, we have $v_\mu = (1-p)v_\nu$.

3   Speed of the infinite-bin model

In this section, we prove the existence of a well-defined notion of speed of the front of an infinite-bin model. We first discuss the case when the distribution μ is finitely supported and the initial configuration is simple, then we extend it to any distribution μ and finally we generalize to any admissible initial configuration.

3.1   Infinite-bin models with finite support

Let μ be a probability measure on $\overline{\mathbb{N}}$ with finite support, i.e. such that there exists $K \in \mathbb{N}$ verifying $\mu([K + 1, +\infty)) = 0$. Let (X_n) be an IBM(μ), we say that (X_n) is an infinite-bin model with support bounded by K. One of the main observations of the subsection is that such an infinite-bin model can be studied using a Markov chain on a finite set. As a consequence, we obtain an expression for the speed of this infinite-bin model.

Given $K \in \mathbb{N}$, we introduce the set

$$S_K = \left\{ x \in \mathbb{Z}_+^{K-1} : \sum_{i=1}^{K-1} x_i < K \text{ and } \forall 1 \le i \le j \le K - 1, x_i = 0 \Rightarrow x_j = 0 \right\}.$$

For any $Y \in S_K$, we write $|Y| = \sum_{j=1}^{K-1} Y(j)$. We introduce

$$\Pi_K : \begin{array}{ccl} S & \longrightarrow & S_K \\ X & \longmapsto & (X(B(X, K) + j - 1), 1 \le j \le K - 1). \end{array}$$

For any $n \in \mathbb{N}$, we write $Y_n = \Pi_K(X_n)$, that encodes the set of balls that are close to the front. As the IBM has support bounded by K, the bin in which the $(n + 1)$-st ball is added to X_n depend only on the position of the front and on the value of Y_n. This reduces the study of the dynamics of (X_n) to the study of $(Y_n, n \ge 1)$.

Lemma 3.1. *The sequence (Y_n) is a Markov chain on S_K.*

Proof. For any $1 \le \xi \le K$ and $Y \in S_K$, we denote by

$$\widetilde{B}(Y, \xi) = \begin{cases} \min\{k \ge 1 : \sum_{i=k}^{K-1} Y(i) < \xi\} & \text{if} \quad |Y| \ge \xi \\ 1 & \text{otherwise,} \end{cases}$$

$$\widetilde{\Phi}_\xi(Y) = \begin{cases} \left(Y(j) + \mathbf{1}_{\left\{ j = \widetilde{B}(Y, \xi) \right\}}, 1 \le j \le K - 1 \right) & \text{if} \quad |Y| < K - 1 \\ \left(Y(j+1) + \mathbf{1}_{\left\{ j+1 = \widetilde{B}(Y, \xi) \right\}}, 1 \le j \le K - 2, 0 \right) & \text{if} \quad |Y| = K - 1. \end{cases}$$

For any $X \in S$ and $\xi \le K$, we have $B(X, \xi) = B(X, K) + \widetilde{B}(\Pi_K(X), \xi) - 1$. Moreover, we have $\Pi_K(\Phi_\xi(X)) = \widetilde{\Phi}_\xi(\Pi_K(X))$.

Let (ξ_n) be i.i.d. random variables with law μ and $X_0 \in S$. For any $n \in \mathbb{N}$, we set $X_{n+1} = \Phi_{\xi_{n+1}}(X_n)$. Using the above observation, we have

$$Y_{n+1} = \Pi_K(X_{n+1}) = \Pi_K(\Phi_{\xi_{n+1}}(X_n)) = \widetilde{\Phi}_{\xi_{n+1}}(\Pi_K(X_n)) = \widetilde{\Phi}_{\xi_{n+1}}(Y_n),$$

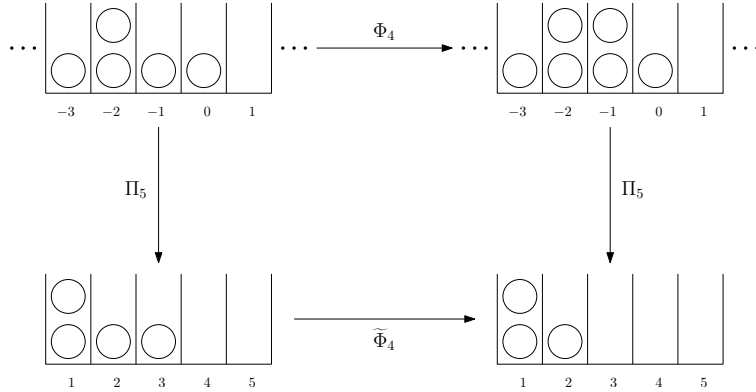thus (Y_n) is a Markov chain. $\qquad\square$

9

Figure 3: "Commutation" of Π_5 with Φ_4 and $\widetilde{\Phi}_4$.

For any $n \in \mathbb{N}$, the set of bins that are part of Y_n represents the set of "active" bins in X_n, i.e. the bins in which a ball can be added at some time in the future with positive probability. The number of balls in (Y_n) increases by one at each time step, until it reaches $K - 1$. At this time, when a new ball is added, the leftmost bin "freezes", it will no longer be possible to add balls to this bin, and the "focus" is moved one step to the right.

We introduce a sequence of stopping times defined by

$$T_0 = 0 \quad \text{and} \quad T_{p+1} = \inf\{n > T_p : |Y_{n-1}| = K - 1\}.$$

We also set $Z_p = K - |Y_{T_p}|$ the number of balls in the bin that "freezes" at time T_p. For any $n \in \mathbb{N}$, we write $\tau_n = p$ for any $T_p \leq n < T_{p+1}$.

Lemma 3.2. *Let $X_0 \in S$ such that $B(X_0, K) = 1$, then*

- *for any $p \geq 0$, $X_\infty(p) = Z_p$,*

- *for any $n \geq 0$ and $\xi \leq K$, $B(X_n, \xi) = \tau_n + B(Y_n, \xi)$.*

Proof. By induction, for any $p \geq 0$, $B(X_{T_p}, K) = p + 1$. Consequently, for any $n \geq T_p$, we have $X_n(p) = X_{T_p}(p) = K - |Y_{T_p}| = Z_p$. Moreover, as

$$B(X_n, K) = \tau_n + 1 \quad \text{and} \quad B(X_n, \xi) = B(X_n, K) + B(Y_n, \xi) - 1,$$

we have the second equality. $\qquad\square$

Using the above result, we prove that the speed of an infinite-bin model with finite support does not depend on the initial configuration. We also obtain a formula for the speed v_μ, that can be used to compute explicit bounds.

Proposition 3.3. *Let μ be a probability measure with finite support and X be an IBM(μ) with initial configuration $X_0 \in S$. There exists $v_\mu \in [0, 1]$ such that for any $\xi \in \mathbb{N}$, we have*

$$\lim_{n \to +\infty} \frac{B(X_n, \xi)}{n} = v_\mu \quad a.s.$$

Moreover, setting π for the invariant measure of (Y_n) we have

$$v_\mu = \frac{1}{\mathbf{E}_\pi(T_2 - T_1)} = \frac{1}{\mathbf{E}_\pi(Z_1)}. \tag{3.1}$$

10

Proof. Let $X_0 \in S$, we can assume that $B(X_0, K) = 1$, up to a deterministic shift. At each time n, a ball is added in a bin with a positive index, thus for any $n \in \mathbb{N}$, we have

$$\sum_{j=1}^{+\infty} X_n(j) = n + \sum_{j=1}^{+\infty} X_0(j).$$

Using the notation of Lemma 3.2, we rewrite it $\sum_{j=1}^{\tau_n} Z_j + |Y_n| = n + \sum_{j=1}^{+\infty} X_0(j)$. Moreover, as $0 \le |Y_n| \le K$ and $0 \le \sum_{j=1}^{+\infty} X_0(j) \le K$, we have

$$1 - \frac{K}{n} \le \frac{\sum_{j=1}^{\tau_n} Z_j}{n} \le 1 + \frac{K}{n},$$

yielding $\lim_{n \to +\infty} \frac{\sum_{j=1}^{\tau_n} Z_j}{n} = 1$ a.s. As $\lim_{p \to +\infty} T_p = +\infty$ a.s., we obtain

$$\lim_{p \to +\infty} \frac{\sum_{j=1}^{p} Z_j}{T_p} = 1 \quad \text{a.s.}$$

Moreover $\lim_{p \to +\infty} \frac{1}{p} \sum_{j=1}^{p} Z_j = \mathbf{E}_\pi(Z_1)$ and $\lim_{p \to +\infty} \frac{T_p}{p} = \mathbf{E}_\pi(T_2 - T_1)$ by ergodicity of (Y_n). Consequently, if we set $v_\mu := \frac{1}{\mathbf{E}_\pi(T_2 - T_1)} = \frac{1}{\mathbf{E}_\pi(Z_1)}$, the constant v_μ is well-defined.

We apply Lemma 3.2, we have

$$\frac{B(X_n, 1)}{n} = \frac{\tau_n}{n} + \frac{B(Y_n, 1)}{n} \in \left[\frac{\tau_n}{n}, \frac{\tau_n}{n} + \frac{K}{n} \right].$$

Moreover, we have $\lim_{n \to +\infty} \frac{\tau_n}{n} = \lim_{p \to +\infty} \frac{p}{T_p} = v_\mu$ a.s. This yields

$$\lim_{n \to +\infty} \frac{B(X_n, 1)}{n} = v_\mu \quad \text{a.s.} \tag{3.2}$$

Using (2.1), this convergence is extended to $\lim_{n \to +\infty} \frac{B(X_n, \xi)}{n} = v_\mu$ a.s. $\qquad \square$

Remark 3.4. If the support of μ is included in $[1, K] \cup \{+\infty\}$, it follows from Lemma 2.2 that the IBM(μ) also has a well-defined speed v_μ.

3.2 Extension to arbitrary distributions

We now use Proposition 3.3 to prove Theorem 1.1.

Proposition 3.5. *Let μ be probability measure on $\mathbb{N}$ and (X_n) an IBM(μ) with initial configuration $X_0 \in S$. There exists $v_\mu \in [0, 1]$ such that for any $\xi \in \mathbb{N}$, we have $\lim_{n \to +\infty} \frac{B(X_n, \xi)}{n} = v_\mu$ a.s.*

Moreover, if ν is another probability measure we have

$$\forall k \in \mathbb{N}, \nu([1, k]) \le \mu([1, k]) \Rightarrow v_\nu \le v_\mu. \tag{3.3}$$

Proof. Let $X_0 \in S$. We write $(\xi_n, n \ge 1)$ for an i.i.d. sequence of random variables of law μ. For any $n, K \ge 1$, we set $\xi_n^K = \xi_n \mathbf{1}_{\{\xi_n \le K\}} + \infty \mathbf{1}_{\{\xi_n > K\}}$. We then define the processes $(\underline{X}_n^K)$ and $(\overline{X}_n^K)$ by $\underline{X}_0^K = \overline{X}_0^K = X_0$ and

$$\underline{X}_{n+1}^K = \Phi_{\xi_{n+1}^K}(\underline{X}_n^K) \quad \text{and} \quad \overline{X}_{n+1}^K = \begin{cases} \Phi_{\xi_{n+1}}(\overline{X}_n^K) & \text{if } \xi_{n+1} \le K \\ \tau(\overline{X}_n^K) & \text{otherwise.} \end{cases}$$

By induction, we have $\underline{X}_n^K \preccurlyeq X_n \preccurlyeq \overline{X}_n^K$ for any $n \geq 0$, using (2.3) and (2.4).

As $(\underline{X}_n^K)$ is an infinite-bin model with support included in $[1, K] \cup \{+\infty\}$, by Remark 3.4, there exists $v_K \in [0, 1]$ such that for any $\xi \in \mathbb{N}$

$$\liminf_{n \to +\infty} \frac{B(X_n, \xi)}{n} \geq \lim_{n \to +\infty} \frac{B(\underline{X}_n^K, \xi)}{n} = v_K \quad \text{a.s.}$$

Moreover, by definition of $(\overline{X}_n^K)$ and (2.2), for any $\xi, n \geq 1$ we have

$$B(\overline{X}_n^K, \xi) = B(\underline{X}_n^K, \xi) + \sum_{j=1}^{n} \mathbf{1}_{\{K < \xi_j < +\infty\}},$$

therefore, by law of large numbers

$$\limsup_{n \to +\infty} \frac{B(X_n, \xi)}{n} \leq \lim_{n \to +\infty} \frac{B(\overline{X}_n^K, \xi)}{n} = v_K + \mu([K + 1, +\infty)) \quad \text{a.s.}$$

By Proposition 2.1, we observe immediately that (v_K) is an increasing sequence, bounded by 1, thus converges. Moreover, $\lim_{K \to +\infty} \mu([K+1, +\infty)) = 0$. We conclude that $\lim_{n \to +\infty} \frac{1}{n} B(X_n, \xi) = \lim_{K \to +\infty} v_K =: v_\mu$ a.s. By Proposition 2.1, (3.3) trivially holds. $\qquad\square$

Remark 3.6. Let μ be a probability measure on $\mathbb{N}$, we set $\mu_K = \mu(.|. \leq K)$. We observe from the proof of Proposition 3.5 and Lemma 2.2 that

$$\mu([1, K]) v_{\mu_K} \leq v_\mu \leq \mu([1, K]) v_{\mu_K} + \mu([K + 1, +\infty)).$$

As v_{μ_K} is the speed of an IBM with support bounded by K, it can be computed explicitly using (3.1). This provides tractable bounds for v_μ. For example, we have $v_\mu \geq \frac{\mu(\{K_0\})}{K_0}$, where $K_0 = \inf\{k > 0 : \mu(k) > 0\}$.

Remark 3.7. Proposition 3.5 can be extended to infinite-bin models starting with a configuration $X \in S^0$. Let μ be a probability measure and (X_n) an IBM(μ) starting with a configuration $X \in S^0$. If μ has a support bounded by K, then the projection $(\Pi_K(X_n))$ is a Markov chain, that will hit the set S_K in finite time. Therefore, we can apply Proposition 3.3, we have $\lim_{n \to +\infty} \frac{1}{n} B(X_n, 1) = v_\mu$ a.s.

If μ has unbounded support, the IBM(μ) can still be bounded, in the same way than in the proof of Proposition 3.5, by infinite-bin models with bounded support. As a consequence, Theorem 1.1 holds for any starting configuration belonging to S^0.

4 A formula for the speed of the infinite-bin model

In this section, we prove that we can write v_μ as the sum of a series, provided that this series converges. A non-rigorous heuristic for the proof goes along the following lines. Let $\eta > 0$ and μ be a probability measure such that $\mu(\{1\}) \geq 1 - \eta$, and $(\xi_n : n \in \mathbb{N})$ be i.i.d. random variables with law μ. If η is small enough then the sequence (ξ_n) consists in long time intervals such that $\xi_n = 1$ on these intervals, separated by short patterns that appear at random. Every

move of type 1 makes the front of the infinite-bin model increase by 1, and each pattern induces a delay. Therefore, we expect the value of v_μ to be close to 1 minus the sum over every possible pattern of the delay caused by this pattern to the process multiplied by its probability of occurrence.

This sum is an infinite sum and we hope that for η small enough, the contributions of the long patterns will decay fast enough so that the series converges and its sum is equal to v_μ. It appears that in fact, this series often converges, even when $\mu(1)$ is not close to 1, and whenever it converges its sum is equal to v_μ.

We recall some notation from the introduction. We denote by $\mathcal{A}$ the set of finite words on the alphabet $\mathbb{N}$. For any $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathcal{A}$, we define $L(\alpha) := n$ to be the length of α.

Let μ be a probability distribution on $\mathbb{N}$ and (ξ_j) i.i.d. random variables with law μ. We write

$$W_\mu(\alpha) := \prod_{j=1}^{L(\alpha)} \mu(\{\alpha_j\}) = \mathbf{P}((\xi_1, \ldots \xi_{L(\alpha)}) = \alpha)$$

for the weight of the word α.

If $\alpha = (\alpha_1, \ldots, \alpha_n)$ is a non-empty word, we denote by $\pi\alpha$ (respectively $\varpi\alpha$) the word $(\alpha_1, \ldots \alpha_{n-1})$ (resp. $(\alpha_2, \ldots \alpha_n)$) obtained by erasing the last (resp. first) letter of α. We use the convention $\pi\emptyset = \varpi\emptyset = \emptyset$.

Given any $X \in S$, we define the function $\varepsilon_X : \mathcal{A} \to \{-1, 0, 1\}$ by

$$\varepsilon_X(\alpha) = \mathbf{1}_{\{\alpha \in \mathcal{P}_X\}} - \mathbf{1}_{\{\varpi\alpha \in \mathcal{P}_X\}},$$

where $\mathcal{P}_X$ is the set of non-empty words α such that, starting from X and applying successively the moves $\Phi_{\alpha_1}, \ldots, \Phi_{\alpha_n}$, the last move Φ_{α_n} results in placing a ball in a previously empty bin.

For $X \in S$ and $\alpha \in \mathcal{A}$, we denote by X^α the configuration of the infinite-bin model obtained after applying successively moves of type $\alpha_1, \alpha_2, \ldots \alpha_n$ to the initial configuration X, i.e.

$$X^\alpha = \Phi_{\alpha_{L(\alpha)}} \big(\Phi_{\alpha_{L(\alpha)-1}} \big(\cdots \Phi_{\alpha_2} \big(\Phi_{\alpha_1}(X) \big) \cdots \big) \big),$$

and we set $d_X(\alpha) = B(X^\alpha, 1) - B(X, 1)$ the displacement of the front of the infinite-bin model after performing the sequence of moves in α. Using this definition, we obtain an alternative expression for $\varepsilon_X(\alpha)$.

Lemma 4.1. *For any $\alpha \in \mathcal{A}$, we have*

$$\varepsilon_X(\alpha) = d_X(\alpha) - d_X(\pi\alpha) - d_X(\varpi\alpha) + d_X(\pi\varpi\alpha). \tag{4.1}$$

Proof. Observe that $d_X(\alpha) - d_X(\pi\alpha)$ equals 0 (resp. 1) if the last move of α adds a ball in a previously non-empty (resp. empty) bin. Therefore we have $d_X(\alpha) - d_X(\pi\alpha) = \mathbf{1}_{\{\alpha \in \mathcal{P}_X\}}$. Similarly, $d_X(\varpi\alpha) - d_X(\pi\varpi\alpha) = \mathbf{1}_{\{\varpi\alpha \in \mathcal{P}_X\}}$. We conclude that

$$\varepsilon_X(\alpha) = \mathbf{1}_{\{\alpha \in \mathcal{P}_X\}} - \mathbf{1}_{\{\varpi\alpha \in \mathcal{P}_X\}} = d_X(\alpha) - d_X(\pi\alpha) - d_X(\varpi\alpha) + d_X(\pi\varpi\alpha). \quad \square$$

As a direct consequence of Lemma 4.1, for any $\alpha = (\alpha_1, \ldots \alpha_n)$ we have

$$d_X(\alpha) = \sum_{k=1}^{n} \sum_{j=1}^{n-k+1} \varepsilon_X((\alpha_k, \alpha_{k+1}, \ldots, \alpha_{k+j-1})), \qquad (4.2)$$

i.e., the displacement induced by α is the sum of $\varepsilon(\beta)$ for any consecutive subword β of α (where the subwords β are counted with multiplicity).

Remark 4.2. One could also go the other way round, start with d_X and define ε_X to be the function verifying

$$\forall \alpha \in \mathcal{A}, d_X(\alpha) = \sum_{\beta \prec \alpha} \varepsilon_X(\beta) m(\beta, \alpha),$$

where $\beta \prec \alpha$ denotes the fact that β is a factor of α (i.e. a consecutive subword of α) and $m(\beta, \alpha)$ denotes the number of times β appears as a factor of α. In that case, one would obtain formula (4.1) for ε_X as the result of a Mőbius inversion formula (see [23, Sections 3.6 and 3.7] for details on incidence algebras and Mőbius inversion formulas).

Using these notation and results, we prove the following lemma.

Lemma 4.3. *For any probability measure μ and $X \in S$, we have*

$$v_\mu = \lim_{n \to +\infty} \frac{1}{n} \sum_{k=1}^{n} \sum_{\alpha \in \mathcal{A}:L(\alpha) \leq k} \varepsilon_X(\alpha) W_\mu(\alpha).$$

This lemma straightforwardly implies Theorem 1.2 by Stolz-Cesàro theorem.

Proof. Let (X_n) be an IBM(μ) starting from the configuration $X \in S$. We have, by definition of d_X, $d_X((\xi_1, \ldots \xi_n)) = B(X_n, 1) - B(X_0, 1)$. Moreover, by Theorem 1.1 and dominated convergence,

$$\lim_{n \to +\infty} \frac{1}{n} \mathbf{E}\left(d_X((\xi_1, \ldots \xi_n))\right) = v_\mu.$$

We easily compute $\mathbf{E}\left(d_X((\xi_1, \ldots \xi_n))\right)$ using (4.2), we obtain

$$\mathbf{E}\left(d_X((\xi_1, \ldots \xi_n))\right) = \sum_{k=1}^{n} \sum_{j=1}^{n-k+1} \mathbf{E}\left(\varepsilon_X((\xi_k, \xi_{k+1}, \ldots, \xi_{k+j-1}))\right)$$

$$= \sum_{k=1}^{n} \sum_{j=1}^{n-k+1} \sum_{\alpha \in \mathcal{A}:L(\alpha)=j} W_\mu(\alpha) \varepsilon_X(\alpha)$$

$$= \sum_{k=1}^{n} \sum_{\alpha \in \mathcal{A}:L(\alpha) \leq k} W_\mu(\alpha) \varepsilon_X(\alpha),$$

which concludes the proof. $\qquad \square$

In Section 6, we study in more details the function ε_X. In particular, we give sufficient conditions on α to have $\varepsilon_X(\alpha) = 0$, which allows to prove that in some cases, the series $\sum_{\alpha \in \mathcal{A}} \varepsilon_X(\alpha) W_\mu(\alpha)$ is absolutely convergent.

5 Length of the longest path in Barak-Erdős graphs

In the rest of the article, we use the results obtained in the previous sections to study the asymptotic behaviour of the length of the longest path in a Barak-Erdős graph. Let $p \in [0,1]$, we write μ_p for the geometric distribution on $\mathbb{N}$ with parameter p, verifying $\mu_p(k) = p(1-p)^{k-1}$ for any $k \geq 1$. In this section, we present a coupling introduced by Foss and Konstantopoulos [11] between an IBM(μ_p) and a Barak-Erdős graph of size n, used to compute the asymptotic behaviour of the length of the longest path in this graph.

Recall that a Barak-Erdős graph on the n vertices $\{1, \ldots, n\}$, with edge probability p is constructed by adding an edge from i to j with probability p, independently for each pair $1 \leq i < j \leq n$. We write $L_n(p)$ for the length of the longest path in this graph. Newman [21] proved that L_n increases at linear speed. More precisely, there exists a function C such that for any $p \in [0,1]$,

$$\lim_{n \to +\infty} \frac{L_n(p)}{n} = C(p) \quad \text{in probability.}$$

Moreover, he proved that $C(p)$ is continuous and increasing on $[0,1]$, and that $C'(0) = e$.

Let $p \in (0,1)$ and (X_n) be an IBM(μ_p), we set $v_p = v_{\mu_p}$ the speed of (X_n), which is well-defined by Proposition 3.5. Foss and Konstantopoulos [11] proved that

$$C(p) = v_p = \lim_{n \to +\infty} \frac{B(X_n, 1)}{n} \quad \text{a.s.} \tag{5.1}$$

We now construct the coupling used to derive 5.1. We associate an infinite-bin model configuration in S to each acyclic directed graph on vertices $\{1, \ldots, n\}$ as follows: for each vertex $1 \leq i \leq n$, we add a ball in the bin indexed by the length of the longest path ending at vertex i, and infinitely many balls in bins with negative index (see Figure 4 for an example). We denote by l_i the length of the longest path ending at position l_i.

(a) An acyclic directed graph G.

(b) The infinite-bin model configuration corresponding to this graph.
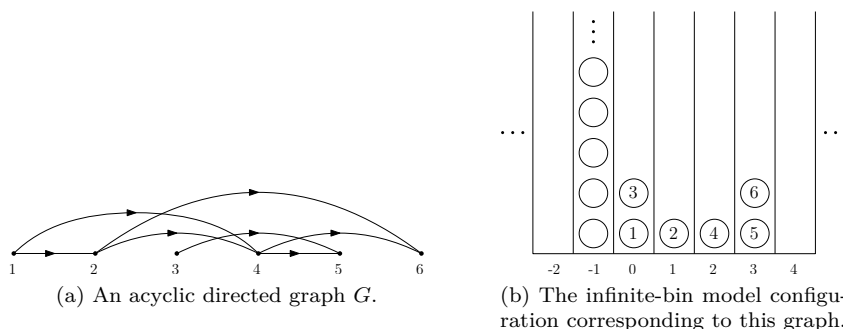
Figure 4: From a Barak-Erdős graph to an infinite-bin model configuration.

We now construct the Barak-Erdős graph as a dynamical process, which is run in parallel with its associated infinite-bin model. At time $n = 0$, we start with the Barak-Erdős graph with no vertex, the empty graph, and the infinite-bin model with infinitely many balls in bins of negative index, and no ball in

15

other bins (which is called configuration Y_0). At time $n = 1$, we add vertex 1 to the Barak-Erdős graph. As $l_1 = 0$, we also add a ball in the bin of index 0 to the configuration Y_0, to obtain the configuration Y_1.

At time $n > 1$, we add vertex n to the Barak-Erdős graph on $\{1, \ldots, n-1\}$. We compute the law of l_n conditionally on $(l_i, i \leq n-1)$. Let σ be a permutation of $\{1, \ldots, n-1\}$ such that $l_{\sigma(1)} \geq l_{\sigma(2)} \geq \cdots \geq l_{\sigma(n-1)}$. The permutation is not necessarily uniquely defined by these inequalities, but this does not matter for our purpose. For each $1 \leq i \leq n-1$, there is an edge between n and $\sigma(i)$ with probability p, independently of any other edge. In this case, there is a path of length $l_i + 1$ in the Barak-Erdős graph that end at site n. The smallest number ξ_n such that there is an edge between $\sigma(\xi_n)$ and n is distributed as a geometric random variable, where if $\xi_n > n-1$, then there is no edge between n and a previous vertex, thus $l_n = 0$ and we add a ball at position 0. As a consequence, the state associated to the graph of size n is given by $Y_n = \Phi_{\xi_n}(Y_{n-1})$.

We have coupled the IBM(μ_p) (Y_n) with a growing sequence of Barak-Erdős graphs, in such a way that for any $n \in \mathbb{N}$, the length of the longest path in the Barak-Erdős graph of size n is given by $B(Y_n, 1)$. Therefore, (5.1) is a direct consequence of Proposition 3.5.

We now use (3.3) to bound the function C. In [11], Foss and Konstantopoulos obtained upper and lower bounds for $C(p)$, that are tight enough for p close to 1 to give the first five terms of the Taylor expansion of C around $p = 1$:

$$C(1 - q) = 1 - q + q^2 - 3q^3 + 7q^4 + O(q^5) \text{ when } q \to 0.$$

We use measures with finite support to approach μ_p, as in the proof of Proposition 3.5. We obtain two sequences of functions that converge exponentially fast toward C on $[\varepsilon, 1]$ for any $\varepsilon > 0$. Let $k \geq 1$, we set

$$\underline{\mu}_p^k(\{j\}) = p(1-p)^{j-1}\mathbf{1}_{\{j \leq k\}} \text{ and } \overline{\mu}_p^k(\{j\}) = p(1-p)^{j-1}\mathbf{1}_{\{j \leq k\}} + (1-p)^k\mathbf{1}_{\{j=k\}}.$$

We write $L_k(p) = v_{\underline{\mu}_p^k}$ and $U_k(p) = v_{\overline{\mu}_p^k}$. By (3.3), for any $k \geq 1$ we have $L_k(p) \leq C(p) \leq U_k(p)$. Moreover, as a (very crude) upper bound, for any $p \in [0, 1]$ we have

$$0 \leq L_k(p) - U_k \leq (1-p)^k \wedge \frac{1}{k}. \tag{5.2}$$

Using Proposition 3.3, the functions L_k and U_k can be explicitly computed. For example, taking $k = 3$ we obtain

$$\frac{p(p^2-3p+3)^2(p^4-6p^3+14p^2-16p+8)}{3p^6-26p^5+96p^4-196p^3+235p^2-158p+47} \leq C(p) \leq \frac{p^3-2p^2+p-1}{p^5-4p^4+8p^3-9p^2+6p-3}.$$

For any $k \in \mathbb{N}$, L_k and U_k are rational functions of p. Their convergence toward C is very fast, which enables to bounds values of $C(p)$. For instance, taking $k = 9$, we obtain $C(0.5) = 0.5780338 \pm 2.10^{-8}$, improving $C(0.5) = 0.58 \pm 10^{-2}$ given by the bounds in [11].

The functions L_k and U_k are very close for p close to 1, which enables to compute the Taylor expansion of $C(1 - q)$ to any order as $q \to 0$. For example, comparing the Taylor expansion of L_6 and U_6, we obtain the first 14 terms of the Taylor expansion of C. However, Theorem 1.4 gives another way to obtain this Taylor expansion.
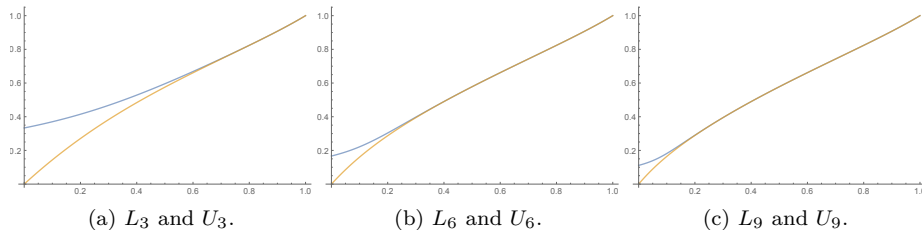
16

(a) L_3 and U_3.   (b) L_6 and U_6.   (c) L_9 and U_9.

Figure 5: Lower and upper bounds L_k and U_k for C, for $k \in \{3, 6, 9\}$.

6   Power series expansion of C in dense graphs

In this section, we prove that C is analytic for $p > 1/2$. Recall that for any word $\alpha \in \mathcal{A}$, we defined the height of α to be

$$H(\alpha) = \sum_{i=1}^{L(\alpha)} \alpha_i - L(\alpha).$$

For $p \in [0, 1]$ we set

$$W_p(\alpha) := W_{\mu_p}(\alpha) = p^{L(\alpha)}(1 - p)^{H(\alpha)}.$$

By Theorem 1.2, if $\sum_{\alpha \in \mathcal{A}} |\varepsilon_X(\alpha)| W_p(\alpha) < +\infty$, then we have

$$C(p) = \sum_{\alpha \in \mathcal{A}} \varepsilon_X(\alpha) p^{L(\alpha)}(1 - p)^{H(\alpha)}. \tag{6.1}$$

We first prove that this series is absolutely convergent. To do so, we obtain sufficient conditions on α to have $\varepsilon_X(\alpha) = 0$. We say that a word $\alpha = (\alpha_1, \ldots, \alpha_l)$ has a renovation event at position $n \geq 1$ if for all $0 \leq k \leq l - n$, $\alpha_{n+k} \leq k + 1$. This concept appears in [11], where these events are used to create time intervals on which the process starts over and is independent of its past. We first show that the existence of a renovation event in α implies $\varepsilon_X(\alpha) = 0$.

Lemma 6.1. *Let $X \in S$, if $\alpha \in \mathcal{A}$ with $L(\alpha) \geq 2$ has a renovation event at position $n \geq 2$, then $\varepsilon_X(\alpha) = 0$.*

Proof. Let $\alpha \in \mathcal{A}$ be a word of length l with a renovation event at position $n \geq 2$. When we run α starting from the configuration X, the move $\alpha_n = 1$ creates a ball in a previously empty bin, of index say b.

As $\alpha_{n+k} \leq k + 1$ for all $0 \leq k \leq l - n$, we are capable of placing the balls produced by these moves in bins of index b or greater, without knowing any information about the bins to the left of bin b (except for the fact that the bin $b - 1$ contains at least one ball).

When we run $\varpi\alpha$ starting from X, the move α_n again creates a ball in a previously empty bin, of index say b'. Running the moves $\alpha_{n+1}, \ldots, \alpha_l$ will produce the same construction as when we run α, with everything just shifted by $b' - b$. In particular, the last move of α places a ball in a previously empty bin if and only if the last move of $\varpi\alpha$ places a ball in a previously empty bin. Consequently $\mathbf{1}_{\{\alpha \in \mathcal{P}_X\}} = \mathbf{1}_{\{\varpi\alpha \in \mathcal{P}_X\}}$ so $\varepsilon_X(\alpha) = 0$. $\square$

Using Lemma 6.1, we are able to prove that for all $k \in \mathbb{N}$, the set of words of height smaller than k such that $\varepsilon_X(\alpha) \neq 0$ is finite.

Lemma 6.2. *Let $X \in S$, for any $\alpha \in \mathcal{A}$ such that $L(\alpha) > H(\alpha) + 1$, we have $\varepsilon_X(\alpha) = 0$.*

Proof. Let α be a word $(\alpha_1, \ldots, \alpha_l)$ such that $l = L(\alpha) > H(\alpha) + 1$. For any $1 \leq k \leq l$, define $S(k) = \sum_{i=1}^{k}(\alpha_i - 2)$. As $L(\alpha) > H(\alpha) + 1$ we have $S(l) < -1$. We set $n = \min\{k : S(t) < -1 \ \forall t \geq k\}$.

Observe that we have $S(1) = \alpha_1 - 2 \geq -1$, thus $n \geq 2$. By induction, for any $0 \leq k \leq l - n$, we have $S(n+k) \geq -k - 2$ and $\alpha_{n+k} \leq k + 1$. Thus α has a renovation event at position $n \geq 2$, so $\varepsilon_X(\alpha) = 0$ by Lemma 6.1. $\square$

Using Lemma 6.2, we prove the absolute convergence of the series in (6.1).

Lemma 6.3. *Let $X \in S$. The series $\sum_{\alpha \in \mathcal{A}} |\varepsilon_X(\alpha)| W_p(\alpha)$ converges for all $p > 1/2$.*

Proof. Let $p > 1/2$. Define $\mathcal{A}_l^h$ to be the set of words of length l and height h. Observe that $\mathcal{A}_l^h$ is the set of compositions of the integer $h + l$ into l parts and it is well-known that $\#\mathcal{A}_l^h = \binom{h+l-1}{l-1}$. By Lemma 6.2, if α is a word such that $|\varepsilon_X(\alpha)| = 1$, then $L(\alpha) \leq H(\alpha) + 1$, thus

$$\sum_{\alpha \in \mathcal{A}} |\varepsilon_X(\alpha)| W_p(\alpha) \leq \sum_{h \geq 0} \sum_{l=1}^{h+1} \sum_{\alpha \in \mathcal{A}_l^h} W_p(\alpha).$$

By definition of $W_p(\alpha)$, we have

$$\sum_{\alpha \in \mathcal{A}} |\varepsilon_X(\alpha)| W_p(\alpha) \leq \sum_{h \geq 0} \sum_{l=1}^{h+1} p^l (1-p)^h \#\mathcal{A}_l^h$$

$$\leq (1-p) \sum_{h \geq 0} \sum_{l=0}^{h} p^l (1-p)^h \binom{h+l}{l}.$$

Let (S_n) be a random walk on $\mathbb{Z}$ starting at 0 and doing a step $+1$ (resp. -1) with probability p (resp. $1 - p$). Then for all $p > 1/2$, we have

$$\sum_{h \geq 0} \sum_{l=0}^{h} p^l (1-p)^h \binom{h+l}{l} = \sum_{n \geq 0} \sum_{l=0}^{\lfloor n/2 \rfloor} \binom{n}{l} p^l (1-p)^{n-l}$$

$$= \sum_{n \geq 0} \sum_{l=0}^{\lfloor n/2 \rfloor} \mathbf{P}(S_n = 2l - n) = \sum_{n \geq 0} \mathbf{P}(S_n \leq 0) < +\infty.$$

Indeed, we have $\mathbf{E}(S_1) = 2p - 1 > 0$, and, $\mathbf{P}(S_n \leq 0)$ decays exponentially fast by Cramér's large deviations theorem. $\square$

Using the above lemma and Theorem 1.2, we immediately obtain the following result.

Lemma 6.4. *For any $X \in S$ and $p > 1/2$, (6.1) holds.*

We use this formula for C to prove that the function can be written as a power series around every $p > 1/2$.

Proof of Theorem 1.4. Fix $\frac{1}{2} < p \leq r \leq 1$ and write $x = r - p \geq 0$. We write $C(p) = C(r - x)$ as a power series in x and determine its radius of convergence.

$$
\begin{aligned}
C(p) &= \sum_{\alpha \in \mathcal{A}} \varepsilon_X(\alpha) p^{L(\alpha)} (1 - p)^{H(\alpha)} \\
&= \sum_{\alpha \in \mathcal{A}} \varepsilon_X(\alpha) (r - x)^{L(\alpha)} (1 - r + x)^{H(\alpha)} \\
&= \sum_{\alpha \in \mathcal{A}} \varepsilon_X(\alpha) \sum_{i=0}^{L(\alpha)} \binom{L(\alpha)}{i} (-1)^i x^i r^{L(\alpha)-i} \sum_{j=0}^{H(\alpha)} \binom{H(\alpha)}{j} x^j (1 - r)^{H(\alpha)-j}.
\end{aligned}
$$

Taking absolute values inside the last series, we obtain

$$
\begin{aligned}
&\sum_{\alpha \in \mathcal{A}} |\varepsilon_X(\alpha)| \sum_{i=0}^{L(\alpha)} \binom{L(\alpha)}{i} x^i r^{L(\alpha)-i} \sum_{j=0}^{H(\alpha)} \binom{H(\alpha)}{j} x^j (1 - r)^{H(\alpha)-j} \\
&= \sum_{\alpha \in \mathcal{A}} |\varepsilon_X(\alpha)| (r + x)^{L(\alpha)} (1 - r + x)^{H(\alpha)} \\
&= \sum_{\alpha \in \mathcal{A}} |\varepsilon_X(\alpha)| (2r - p)^{L(\alpha)} (1 - p)^{H(\alpha)}.
\end{aligned}
$$

By the same computations as in Lemma 6.3, we have

$$
\sum_{\alpha \in \mathcal{A}} |\varepsilon_X(\alpha)| (2r - p)^{L(\alpha)} (1 - p)^{H(\alpha)} \leq (1 - p) \sum_{h \geq 0} \sum_{l=0}^{h} (2r - p)^l (1 - p)^h \binom{h + l}{l}.
$$

If this quantity is finite, then the power series expansion of C around r has a radius of convergence at least $r - p$. Writing $(S_n^{p,r})$ for a random walk on $\mathbb{Z}$ starting at 0 and doing a step $+1$ (resp. -1) with probability $\frac{2r-p}{2r+1-2p}$ (resp. $\frac{1-p}{2r+1-2p}$), we have

$$
\sum_{h \geq 0} \sum_{l=0}^{h} (2r - p)^l (1 - p)^h \binom{h + l}{l} = \sum_{n \geq 0} (2r + 1 - 2p)^n \mathbf{P}\left(S_n^{p,r} \leq 0\right). \qquad (6.2)
$$

By Chernoff's bound, we obtain

$$
\begin{aligned}
\mathbf{P}\left(S_n^{p,r} \leq 0\right) &\leq \inf_{t>0} \left(\mathbf{E}\left[e^{-t S_1^{p,r}} \right] \right)^n \\
&\leq \inf_{t>0} \left(\frac{2r - p}{2r + 1 - 2p} e^{-t} + \frac{1 - p}{2r + 1 - 2p} e^t \right) \\
&\leq \left(\frac{2\sqrt{(2r - p)(1 - p)}}{2r + 1 - 2p} \right)^n.
\end{aligned}
$$

Thus the series in (6.2) converges as soon as $2\sqrt{(2r - p)(1 - p)} < 1$, i.e. if

$$
r + \tfrac{1}{2} - \sqrt{r^2 - r + \tfrac{1}{2}} < p \leq r \leq 1.
$$

For $r > 1/2$, we have $r + \frac{1}{2} - \sqrt{r^2 - r + \frac{1}{2}} < r$, thus the power series expansion of C centered at r has a positive radius of convergence. Therefore C is analytic on $\left(\frac{1}{2}, 1\right]$. In particular, for $r = 1$, expanding the expression in (6.1) in powers of $(1 - p)$, we conclude that for p larger than $\frac{3 - \sqrt{2}}{2}$, we have

$$C(p) = \sum_{k \geq 0} a_k (1 - p)^k,$$

with (a_k) defined in (1.3). $\qquad\square$

7 Longest directed path in sparse graphs

We study in this section the asymptotic behaviour of $C(p)$ as $p \to 0$. Newman proved in [21] that $C(p) \sim pe$. We link in Section 7.1 this result with the estimate obtained by Aldous and Pitman [2] for the speed of an IBM with uniform distribution. Let $k \in \mathbb{N}$, we write ν_k for the uniform distribution on $\{1, \ldots, k\}$ and w_k for the speed of the IBM(ν_k), Aldous and Pitman proved that

$$(kw_k, k \in \mathbb{N}) \text{ increases toward } e \text{ as } k \to +\infty. \tag{7.1}$$

This result is obtained by coupling the infinite-bin model with a continuous-time branching random walk with selection. Adapting the result of Bérard and Gouéré [4] on the speed of branching random walks with selection, we obtain the following estimate.

Lemma 7.1. *We have $kw_k = e - \frac{\pi^2 e}{2}(\log k)^{-2}(1 + o(1))$ as $k \to +\infty$.*

Applying Lemma 7.1 to compute the asymptotic behaviour of C, we prove Theorem 1.6 :

$$C(p) = ep\left(1 - \frac{\pi^2}{2}(\log p)^{-2}\right) + o(p(\log p)^{-2}) \text{ as } p \to 0. \tag{7.2}$$

The rest of the section is organized as follows. In Section 7.1, we prove Theorem 1.6 assuming Lemma 7.1. In Section 7.2, we prove Lemma 7.1 using the coupling with a branching random walk with selection. Some preliminary results on this model are derived in Section 7.3. The speed of the cloud of particles in a branching random walk with selection is obtained in Section 7.4.

7.1 Proof of Theorem 1.6 assuming Lemma 7.1

We use the increasing coupling of Proposition 3.5 to link the asymptotic behaviours of w_k and $C(1/k)$ as $k \to +\infty$.

Lemma 7.2. *For any $k \in \mathbb{N}$ we have*

$$\forall p \in [\tfrac{1}{k+1}, \tfrac{1}{k}], \, C(p) \leq w_k$$
$$\forall p \in [0, 1], \, C(p) \geq kp(1 - p)^k w_k.$$

Proof. Let $k \in \mathbb{N}$ and $p \in [\frac{1}{k+1}, \frac{1}{k}]$. We observe that for any $j \in \mathbb{N}$,

$$\mu_p([1,j]) = \sum_{i=1}^{j} p(1-p)^{i-1} \leq (pj) \wedge 1 \leq \nu_k([1,j]).$$

Therefore $C(p) \leq w_k$ by (3.3).

Let $p \in [0,1]$, we set $x = kp(1-p)^{k-1}$. Observe that $0 \leq x \leq 1$. For any $j \in \mathbb{N}$, we have

$$\mu_p([1,j]) = \sum_{i=1}^{j} p(1-p)^{i-1} \geq (j \wedge k)p(1-p)^{k-1} \geq k\nu_k([1,j])p(1-p)^{k-1}.$$

Therefore, writing $\nu_k^x = x\nu_k + (1-x)\delta_\infty$, we have $\mu_p([1,j]) \geq \nu_k^x([1,j])$ for any $j \in \mathbb{N}$. We apply (3.3) to μ_p and ν_x^k. By Lemma 2.2, the speed of the IBM(ν_k^x) is xw_k. We conclude that for any $k \in \mathbb{N}$ and $p \in [0,1]$, we have

$$C(p) \geq kp(1-p)^{k-1}w_k.$$

$\square$

Proof of Theorem 1.6. For any $k \in \mathbb{N}$ and $p \in [\frac{1}{k+1}, \frac{1}{k}]$, by Lemma 7.2, we have $C(p)/p \leq (k+1)w_k$, therefore Lemma 7.1 yields

$$\limsup_{p \to 0}(\log p)^2 \left(\frac{C(p)}{p} - e \right) \leq \limsup_{k \to +\infty}(\log k)^2 \left((k+1)w_k - e \right) \leq -\frac{\pi^2 e}{2}.$$

By Lemma 7.2 again, we have $C(p)/p \geq (1-p)^k(kw_k)$ for any $k \in \mathbb{N}$ and $p \in [0,1]$. Let $\delta > 0$, we set $k = \lceil 1/p^{1-\delta} \rceil$. We have $(1-p)^k - 1 \sim -p^\delta$ as $p \to 0$. This yields

$$(\log p)^2 \left(\frac{C(p)}{p} - e \right) \geq \frac{(\log k)^2}{(1-\delta)^2} \left((1-p)^k(kw_k) - e \right).$$

Using again Lemma 7.1, we have

$$\liminf_{p \to 0}(\log p)^2 \left(\frac{C(p)}{p} - e \right) \geq -\frac{\pi^2 e}{2(1-\delta)^2}.$$

Letting $\delta \to 0$ concludes the proof. $\square$

7.2 Proof of Lemma 7.1 using branching random walks

By Lemma 7.2, to obtain the asymptotic behaviour of $C(p)$ as $p \to 0$, it is enough to control the asymptotic behaviour of kw_k as $k \to +\infty$. To obtain (7.1), Aldous and Pitman compared the IBM(ν_k) with a continuous-time branching random walk with selection, that we now define.

Let $\lambda > 0$ and $\mathcal{L}$ be the law of a point process. A continuous-time branching random walk evolves as follows. Every particle in the process is associated with an independent exponential clock of parameter λ. When a clock rings, the corresponding particle dies, giving birth to children that are positioned according to a point process with law $\mathcal{L}$, shifted by the position of the dead

parent particle. For any $t \geq 0$, we write $\mathcal{N}_t$ for the set of particles alive at time t. For any $u \in \mathcal{N}_t$, we write $X_t(u)$ for the position of the particle u alive at time t.

The continuous-time branching random walk that we defined here has been introduced by Uchiyama [25]. We refer to [18] for further references of similar processes. Let L be a point process of law $\mathcal{L}$. For any $\theta > 0$ we denote by $\Lambda(\theta) = \mathbf{E}\left(\sum_{\ell \in L} e^{\theta \ell}\right) - 1$ and we set $v = \lambda \inf_{\theta > 0} \frac{\Lambda(\theta)}{\theta}$. We assume that

$$\text{there exists } \varphi^* > 0 \text{ such that } \varphi^* \Lambda'(\varphi^*) - \Lambda(\varphi^*) = 0. \tag{7.3}$$

Then we have $v = \lambda \frac{\Lambda(\varphi^*)}{\varphi^*}$ and we set $\tau^2 = \lambda \Lambda''(\varphi^*)$. Using a straightforward extension of the classical result of Biggins [6], it is well-known that

$$\lim_{t \to +\infty} \frac{1}{t} \max_{u \in \mathcal{N}_t} X_t(u) = v \quad \text{a.s.}$$

Let $N \in \mathbb{N}$. A continuous-time branching random walk with selection of the rightmost N particles or N-BRW is defined as follows. Each particle in the process reproduces independently as in a continuous-time branching random walk, but every time there are more than N particles currently alive in the process, every particle but the rightmost N are immediately killed without reproducing.

For any $t \geq 0$, we denote by $\mathcal{N}_t^N$ the set of particles alive at time t in the N-BRW, that can be defined as follows. At time 0, $\mathcal{N}_0^N$ is the set of the N rightmost particles in $\mathcal{N}_0$, with ties broken uniformly at random. The set $\mathcal{N}_t^N$ remains constant in between reproduction events. At each reproduction time T, the set $\mathcal{N}_T^N$ is defined as the rightmost N descendants of particles in $\mathcal{N}_{T-}$, with ties again broken uniformly at random. The following estimate on the speed of the cloud of particles in the N-BRW is proved in Section 7.4.

Lemma 7.3. *Let $(X_t(u), u \in \mathcal{N}_t^N)_t$ be a continuous-time branching random walk with selection of the rightmost N particles. Given L a point process of law $\mathcal{L}$, we assume there exists $\varepsilon > 0$ such that*

$$\mathbf{P}(L = \emptyset) = 0, \quad \mathbf{E}(\#L) > 1 \quad and \quad \mathbf{E}(e^{\varepsilon \# L}) < +\infty, \tag{7.4}$$

$$\mathbf{E}\left(\max_{\ell \in L} |\ell|^2\right) + \mathbf{E}\left(\left(\sum_{\ell \in L} e^{\varphi^* \ell}\right)^2\right) < +\infty. \tag{7.5}$$

For any $N \in \mathbb{N}$, there exists v_N such that

$$\lim_{t \to +\infty} \frac{\max_{u \in \mathcal{N}_t^N} X_t(u)}{t} = \lim_{t \to +\infty} \frac{\min_{u \in \mathcal{N}_t^N} X_t(u)}{t} = v_N \quad a.s.$$

$$and \quad \lim_{N \to +\infty} (\log N)^2 (v_N - v) = -\frac{\pi^2 \varphi^* \tau^2}{2}.$$

The existence of the speed v_N of the N-BRW is proved in Section 7.3, and its asymptotic behaviour is obtained in Section 7.4 by adapting the proof used to study discrete-time branching random walks with selection in [4, 20].

Using the Aldous-Pitman coupling (described below) between $\mathrm{IBM}(\nu_k)$ and continuous-time branching random walks with selection, we now derive the asymptotic behaviour of kw_k as $k \to +\infty$, assuming that Lemma 7.3 holds.

22

Proof of Lemma 7.1. Let $k \in \mathbb{N}$. We write $(N_t, t \geq 0)$ for a Poisson process of parameter k and $(X_n, n \geq 0)$ for an independent $\text{IBM}(\nu_k)$. For any $t > 0$, we denote by $\mathcal{N}_t^k$ the set consisting of the rightmost k balls in the configuration X_{N_t}, and by $Y_t(u)$ the position of the ball $u \in \mathcal{N}_t^k$.

We observe that $(Y_t(u), u \in \mathcal{N}_t^k)_t$ evolves as follows: every ball stays put until an exponential random time with parameter k. At that time T, a ball $u \in \mathcal{N}_T^k$ is chosen uniformly at random, a new ball is added at position $Y_T(u) + 1$, and the leftmost ball is erased.

By classical properties of exponential random variables, this evolution can be written in this way: to each ball is associated a clock with parameter 1. When a clock rings, the corresponding ball makes a "child" to the right of its current position, and the leftmost ball is erased. In other words, $(Y_t(u), u \in \mathcal{N}_t^k)_t$ is a continuous-time branching random walk with selection, with parameter $\lambda = 1$ and point process $L = \delta_0 + \delta_1$.

We observe that $\Lambda(\theta) = e^\theta$. By straightforward computations, this yields

$$v = e, \quad \varphi^* = 1 \quad \text{and} \quad \tau^2 = e.$$

Consequently, using Lemma 7.3, we obtain

$$\lim_{t \to +\infty} \frac{\max_{u \in \mathcal{N}_t^k} Y_t(u)}{t} = \lim_{t \to +\infty} \frac{\min_{u \in \mathcal{N}_t^k} Y_t(u)}{t} = v_k \quad \text{a.s,}$$

with $\lim_{k \to +\infty} (\log k)^2 (v_k - e) = -\frac{\pi^2 e}{2}$.

By the law of large numbers and Proposition 3.3, we have

$$\lim_{t \to +\infty} \frac{B(X_{N_t}, 1)}{t} = \lim_{n \to +\infty} \frac{B(X_n, 1)}{n} \lim_{t \to +\infty} \frac{N_t}{t} = k w_k \quad \text{a.s.}$$

We conclude the proof observing that $B(X_{N_t}, 1) - 1 = \max_{u \in \mathcal{N}_t^k} X_t(u)$. $\qquad\square$

7.3 Speed of the N-branching random walk

In this section, we present an increasing coupling introduced by Bérard and Gouéré on branching random walks with selection, and use it to prove that the speed of the N-BRW is well-defined. Loosely speaking, this coupling accounts for the following fact : the larger the population of a branching random walk with selection is, the faster it travels. To state this coupling, we extend the definition of branching random walks with selection to authorize the maximal size of the population to vary.

Let F be a càdlàg, integer-valued process, adapted to the filtration of the continuous-time branching random walk $(X_t(u), u \in \mathcal{N}_t)_t$. For any $t \geq 0$, we denote by $\mathcal{N}_t^F$ the $F(t)$ rightmost descendants of particles belonging to $\mathcal{N}_{t-}^F$. We call $(X_t(u), u \in \mathcal{N}_t^F)_t$ a branching random walk with selection of the $F(t)$ rightmost particles at each time t, or a F-BRW for short. Note that if F is a constant process $N \in \mathbb{N}$, the notation N-BRW remains consistent.

In the rest of the article, we will assume the point process law $\mathcal{L}$ satisfies the integrability conditions (7.3), (7.4) and (7.5).

Lemma 7.4. *Let F and G be two càdlàg integer-valued adapted processes, we assume that*

$$\forall x \in \mathbb{R}, \#\{u \in \mathcal{N}_0^F : X_0(u) \geq x\} \leq \#\{u \in \mathcal{N}_0^G : Y_0(u) \geq x\}.$$

There exists a coupling between an F-BRW $(X_t(u), u \in \mathcal{N}_t^F)_t$ and a G-BRW $(Y_t(u), u \in \mathcal{N}_t^G)_t$ such that a.s. for any $t > 0$, on the event $\{F_s \leq G_s, s \leq t\}$,

$$\forall x \in \mathbb{R}, \#\{u \in \mathcal{N}_t^F : X_t(u) \geq x\} \leq \#\{u \in \mathcal{N}_t^G : Y_t(u) \geq x\}. \qquad (7.6)$$

This lemma is obtained as a straightforward adaptation of [4, Lemma 1].

Proof. We write $m = \#\mathcal{N}_0^F$, $n = \#\mathcal{N}_0^G$ and $x_1 \geq \cdots \geq x_m$ (respectively $y_1 \geq \cdots \geq y_n$) the ranked values of $(X_0(u), u \in \mathcal{N}_0^F)$ (resp. $(Y_0(u), u \in \mathcal{N}_0^G)$). By assumptions, we have $m \leq F_0$, $m \leq n \leq G_0$ and $x_j \leq y_j$ for all $j \leq m$. We assume that we are on the event $\{F_s \leq G_s, s \leq t\}$, and we couple X and Y such that this property remains true at any time $s \leq t$.

We associate exponential clocks to particles in the processes in such a way that the particles in position x_j and y_j reproduce at the same time, for any $j \leq m$. We write T_b the first time one of these particles reproduces, T_a the first time a particle located at position $y_{m+1}, \ldots y_n$ reproduces,

$$S = \inf\{t > 0 : F_t \neq F_0 \quad \text{or} \quad G_t \neq G_0\} \quad \text{and} \quad R = T_a \wedge T_b \wedge S.$$

We observe that X and Y are constant processes until time R, that $R > 0$ a.s. and that $T_a \neq T_b$ a.s.

One of three things can happen at time R. First, if $R = T_a$, there is a reproduction event in Y but not in X. If we rank in the decreasing order the children of particles in $\mathcal{N}_{R-}^F$ (resp. $\mathcal{N}_{R-}^G$) as $(\widetilde{x}_j)$ (resp. $(\widetilde{y}_j)$), we again have $\widetilde{x}_j \leq \widetilde{y}_j$ for any $j \leq m$. As $F_R \leq G_R$, applying the selection procedure to both models yields

$$\forall x \in \mathbb{R}, \#\{u \in \mathcal{N}_R^F : X_R(u) \geq x\} \leq \#\{u \in \mathcal{N}_R^G : Y_R(u) \geq x\}.$$

If $R = T_b$, then there is a reproduction event in X and Y. We use the same point process to construct the children of the particle that reproduces in each process. Once again, ranking in the decreasing order these children, then applying the selection, we have

$$\forall x \in \mathbb{R}, \#\{u \in \mathcal{N}_R^F : X_R(u) \geq x\} \leq \#\{u \in \mathcal{N}_R^G : Y_R(u) \geq x\}.$$

Finally, if $R = S \notin \{T_a, T_b\}$, the maximal size of at least one of the populations is modified. Even if this implies the death of some particles in X or Y, the property (7.6) is preserved at time R.

There is a finite sequence of times (R_k) smaller than t such that X or Y is modified at each time R_k. Using this coupling on each time interval of the form $[R_k, R_{k+1}]$ yields (7.6). $\qquad \square$

Using this lemma, we prove that the cloud of particles in a N-BRW drifts at linear speed v_N.

Lemma 7.5. *For any $N \in \mathbb{N}$, there exists v_N such that*

$$\lim_{t \to +\infty} \frac{1}{t} \max_{u \in \mathcal{N}_t^N} X_t(u) = \lim_{t \to +\infty} \frac{1}{t} \min_{u \in \mathcal{N}_t^N} X_t(u) = v_N \quad a.s.$$

Moreover, if $(X_0(u), u \in \mathcal{N}_0) = (0, \ldots, 0) \in \mathbb{R}^N$, we have

$$v_N = \inf_{t > 0} \frac{\mathbf{E}\left[\max_{u \in \mathcal{N}_t^N} X_t(u)\right]}{t} = \sup_{t > 0} \frac{\mathbf{E}\left[\min_{u \in \mathcal{N}_t^N} X_t(u)\right]}{t}. \qquad (7.7)$$

The proof of this lemma is adapted from [4, Proposition 2].

Proof. Let $N \in \mathbb{N}$, we denote by $(X_t(u), u \in \mathcal{N}_t)_t$ an N-BRW starting with N particles located at 0 at time 0. We set

$$M_t = \max_{u \in \mathcal{N}_t^N} X_t(u) \quad \text{and} \quad m_t = \min_{u \in \mathcal{N}_t^N} X_t(u).$$

We prove that (M_t) (respectively (m_t)) is a sub-additive (resp. super-additive) sequence.

In effect, by Lemma 7.4, for any $s \geq 0$, we can couple $(X_{t+s}(u), u \in \mathcal{N}_{t+s})$ with an N-BRW $\widetilde{X}$ starting with N particles at position M_s in such a way that (7.6) is verified. We write $M_s + M_{s,t}$ the maximal displacement at time t of $\widetilde{X}$. For any $s \leq t$, we have $M_{0,t} \leq M_{0,s} + M_{s,t}$.

We also observe that for any $s \geq 0$, $(M_{s,s+t})_t$ has the same law as $(M_t)_t$, and $M_{s,t}$ is independent of $(M_{u,v}, u \leq v \leq s)$. Moreover, by Lemma 7.4, the maximal displacement of the N-BRW X at time t is larger than the maximal displacement of a 1-BRW Y starting with a particle located at 0 at time 0. But the process Y is a continuous-time random walk, with step distribution max L. Therefore,

$$\mathbf{E}(M_{0,1}) \geq -\mathbf{E}\left(\left|\max_{u \in \mathcal{N}_1} Y_t(u)\right|\right) > -\infty, \quad \text{by (7.4)}.$$

Applying Kingman's subadditive ergodic theorem, there exists $\overline{v}_N \in \mathbb{R}$ such that $\lim_{t \to +\infty} \frac{M_{0,t}}{t} = \overline{v}_N$ a.s. (see Kallenberg [16, Theorem 9.14]). With a similar reasoning, we obtain $\lim_{t \to +\infty} \frac{m_t}{t} = \underline{v}_N \in \mathbb{R}$. Moreover, we have

$$\overline{v}_N = \lim_{t \to +\infty} \frac{1}{t} \mathbf{E}(M_t) \quad \text{and} \quad \underline{v}_N = \lim_{t \to +\infty} \frac{1}{t} \mathbf{E}(m_t).$$

We have immediately $\overline{v}_N \geq \underline{v}_N$, we now prove these two quantities are equal.

Let $A > 0$, we define a sequence of waiting times by setting $T_0 = 0$ and T_{k+1} is the first time after time $T_k + 1$ such that only the descendants of the rightmost particle at time $T_{k+1} - 1$ reproduced between times $T_{k+1} - 1$ and T_{k+1}, every particle alive in $\mathcal{N}_{T_{k+1}}^N$ descend from the rightmost particle at time $T_{k+1} - 1$ and the distance between the rightmost and the leftmost of this offspring is smaller than A_N. As long as $A_N > 0$ is large enough, this defines a sequence of a.s. finite hitting times (as $\mathbf{E}(T_k) < +\infty$). By definition, we have

$$\limsup_{k \to +\infty} M_{T_k} - m_{T_k} \leq A_N \quad \text{a.s.}$$

therefore $\liminf_{t \to +\infty} \frac{M_t - m_t}{t} = 0$ a.s. which yields $\overline{v}_N = \underline{v}_N =: v_N$.

Finally, using Lemma 7.4 again, we can couple a N-BRW X starting from any initial condition X_0 with an N-BRW $\overline{X}$ starting from N particles in position max X_0 and another one $\underline{X}$ starting with N particles in position min X_0 in such a way that for any $t > 0$.

$$\min_{u \in \mathcal{N}_t^N} \underline{X}_t(u) \leq \min_{u \in \mathcal{N}_t^N} X_t(u) \leq \max_{u \in \mathcal{N}_t^N} X_t(u) \leq \max_{u \in \mathcal{N}_t^N} \overline{X}_t(u) \text{ a.s.}$$

As a consequence, for any N-BRW, we have

$$v_N \leq \liminf_{t \to +\infty} \frac{1}{t} \min_{u \in \mathcal{N}_t^N} X_t \leq \limsup_{t \to +\infty} \frac{1}{t} \max_{u \in \mathcal{N}_t^N} X_t \leq v_N \quad \text{a.s.}$$

which concludes the proof. $\qquad \square$

7.4   End of the proof of Lemma 7.3

In this section, we use Lemma 7.4 to compare the asymptotic behaviour of a continuous-time and a discrete-time branching random walk with selection. In the latter model, every particle in the process reproduces independently at integer-valued times. The discrete-time branching random walk with selection was introduced by Brunet and Derrida [7] to study noisy FKPP equations. In this article, they conjectured that the cloud of particles drifts at speed w_N, that satisfies, as $N \to +\infty$

$$w_N - w = -\frac{\chi}{(\log N + 3 \log \log N + o(\log \log N))^2}, \tag{7.8}$$

for some explicit constants $w \in \mathbb{R}$ and $\chi > 0$.

We describe more precisely the discrete-time branching random walk with selection model. Every particle reproduces independently at each integer time. The children are positioned around their parent according to i.i.d. point processes of law $\mathcal{M}$. Only the rightmost N children survive to form the new generation. For every $n \in \mathbb{N}$, we set $X_n(1) \geq X_n(2) \geq \ldots \geq X_n(N)$ the ranked positions of particles alive at generation n.

Let M be a point process of law $\mathcal{M}$. We write $\kappa(\theta) = \log \mathbf{E}\left(\sum_{m \in M} e^{\theta m}\right)$ for any $\theta > 0$, and $w = \inf_{\theta > 0} \frac{\kappa(\theta)}{\theta}$. We assume

$$\text{there exists } \theta^* > 0 \text{ such that } \theta^* \kappa'(\theta^*) - \kappa(\theta^*) = 0. \tag{7.9}$$

We then have $w = \frac{\kappa(\theta^*)}{\theta^*}$, and we write $\sigma^2 = \kappa''(\theta^*)$. Bérard and Gouéré proved that for a binary branching random walk with exponential moments, there exists (w_N) such that

$$\lim_{n \to +\infty} \frac{X_n(1)}{n} = \lim_{n \to +\infty} \frac{X_n(N)}{n} = w_N \quad \text{a.s.}$$

$$\text{and} \quad \lim_{N \to +\infty} (\log N)^2 (w_N - w) = -\frac{\pi^2 \theta^* \sigma^2}{2}. \tag{7.10}$$

The integrability assumptions were extended by Mallein [20] to more general reproduction laws. In particular, (7.10) holds under the following conditions:

$$\mathbf{P}(M = \emptyset) = 0, \quad \mathbf{E}(\#M) > 0 \quad \text{and} \quad \mathbf{E}\left(\left|\max_{m \in M} m\right|^2\right) < +\infty \tag{7.11}$$

$$\mathbf{E}\left(\sum_{m \in M} e^{\theta^* m} m^2\right) + \mathbf{E}\left(\sum_{m \in M} e^{\theta^* m} \log\left(\sum_{m \in M} e^{\theta^* m}\right)^2\right) < +\infty. \tag{7.12}$$

Bérard and Maillard [5] obtained a different asymptotic behaviour for branching random walks with heavy tails. Maillard [19] obtained precise results for the similar model of branching Brownian motion with selection.

Using Lemma 7.4, we extend (7.10) to obtain an upper bound for the asymptotic behaviour of the speed of a continuous-time branching random walk with selection.

Lemma 7.6. *We have* $\limsup_{N \to +\infty} (\log N)^2 (v_N - v) \leq -\dfrac{\pi^2 \varphi^* \tau^2}{2}$.

Proof. Let $(W_t(u), u \in \mathcal{N}_t)_t$ be a continuous-time branching random walk with reproduction law $\mathcal{L}$ and parameter λ, starting with a single particle at position 0 at $t = 0$. We introduce the point process $M = \sum_{u \in \mathcal{N}_1} \delta_{W_1(u)}$. This proof is based on a comparison between the N-BRW $(X_t(u), u \in \mathcal{N}_t^N)_t$ and a discrete-time branching random walk with selection $(Y_n(j), j \leq N)_n$ with reproduction law M.

Let $\theta > 0$, for any $t \geq 0$ we write $f_\theta(t) = \mathbf{E}\left(\sum_{u \in \mathcal{N}_t} e^{\theta W_t(u)}\right)$. Note that

$$\forall t \geq 0, f_\theta'(t) = \lambda \Lambda(\theta) f_\theta(t) \quad \text{with } f_\theta(0) = 1,$$

therefore $\kappa(\theta) = \log f_\theta(1) = \lambda \Lambda(\theta)$. Thus, (7.9) is verified by (7.3), and we have $w = v$, $\theta^* = \varphi^*$ and $\sigma^2 = \tau^2$. Moreover, by (7.4) and (7.5), the point process M satisfies (7.11) and (7.12). As a consequence of [20, Theorem 1.1], there exists a sequence (w_N) such that

$$\lim_{n \to +\infty} \frac{Y_n(1)}{n} = w_N \quad \text{a.s., with } \lim_{N \to +\infty} (\log N)^2 (w_N - v) = -\frac{\pi^2 \varphi^* \tau^2}{2}. \quad (7.13)$$

We now provide an alternative definition of $(Y_n(j), j \leq N)_n$ as a continuous-time branching random walk. We define a càdlàg adapted process F as follows. At any integer time $n \in \mathbb{Z}_+$, we set $F_n = N$, and if $t \in (n, n+1)$, F_t is the number of particles at time t that descend from a particle in $\mathcal{N}_n^F$. By the branching property, we easily observe that the F-BRW $(Y_t(u), u \in \mathcal{N}_t^F)_t$ satisfies

$$\left(\left(Y_n(u), u \in \mathcal{N}_n^F\right), n \geq 0\right) \overset{(d)}{=} \left(\left(Y_n(j), j \leq N\right), n \geq 0\right).$$

As a result, we can identify the two processes, and (7.13) yields

$$\lim_{n \to +\infty} \frac{1}{n} \max_{u \in \mathcal{N}_n^F} Y_n(u) = w_N \quad \text{a.s.}$$

We now extend this convergence on $\mathbb{Z}_+$ to the convergence of $(\max_{u \in \mathcal{N}_t^F} Y_t(u))_t$.

We denote by $\xi = \max_{s \in [0,1]} \max_{u \in \mathcal{N}_s} W_s(u)$ the maximal position attained before time 1 by a continuous-time branching random walk. By (7.4), we have $\mathbf{E}(\xi) < +\infty$. Moreover, given $(\xi_j^n, j \leq N, n \geq 0)$ i.i.d. copies of ξ, this sequence can easily be coupled with the process Y such that for all $n \geq 0$ and $t \in [n, n+1)$,

$$\max_{u \in \mathcal{N}_n^F} Y_n(u) + \max_{j \leq N} \xi_j \geq \max_{u \in \mathcal{N}_t^F} Y_t(u) \quad \text{a.s.}$$

Using the Borel-Cantelli lemma, we conclude that

$$\limsup_{t \to +\infty} \frac{1}{t} \max_{u \in \mathcal{N}_t^F} Y_t(u) \leq w_N \quad \text{a.s.}$$

As $F_t \geq N$ for any $t \geq 0$, by Lemma 7.4, we can couple the processes X and Y such that $\max_{u \in \mathcal{N}_t^N} X_t(u) \leq \max_{u \in \mathcal{N}_t^F} Y_t(u)$ a.s. for any $t \geq 0$. As a consequence, we have

$$\limsup_{t \to +\infty} \frac{1}{t} \max_{u \in \mathcal{N}_t^F} X_t(u) \leq w_N \quad \text{a.s.}$$

hence $v_N \leq w_N$, which concludes the proof. $\qquad\square$

The lower bound is obtained in a similar yet more involved fashion. The proof of this lemma is adapted from [20, Section 4.4].

Lemma 7.7. *We have* $\liminf_{N \to +\infty} (\log N)^2 (v_N - v) \geq -\frac{\pi^2 \varphi^* \tau^2}{2}.$

Proof. In this proof, we construct a particle process Y that evolves similarly to a continuous-time branching random walk with selection, with frequent renovation events, and that can be coupled with the N-BRW X such that its maximal displacement is smaller than the maximal displacement of X. Given $\alpha \in (0, 1)$, the process evolves typically like a discrete-time $\lfloor \alpha N \rfloor$-branching random walk, and on a time scale of order $(\log N)^3$, every particle in the process is killed and replaced by P particles starting from the smallest position in Y at that time.

Let $\alpha \in (0, 1)$, we denote by $P = \lfloor \alpha N \rfloor$. We set $(W_t(u), u \in \mathcal{N}_t)$ a continuous-time branching random walk starting from a single particle located at position 0. As $\mathbf{E}(\#L) < +\infty$, there exists $\beta > 0$ such that $\mathbf{E}(\#\mathcal{N}_\beta) < \frac{1}{\alpha} - \beta$. We introduce the point process $M^\beta = \sum_{u \in \mathcal{N}_\beta} \delta_{W_\beta(u)}$.

Let $(Y_n(j), j \leq P)_n$ be a discrete-time branching random walk with selection of the rightmost P particles, with reproduction law M^β, starting with P particles located at position 0. With the same computations as in the proof of Lemma 7.6, we have $\kappa(\theta) = \beta \lambda \Lambda(\theta)$ and therefore

$$w = \beta v, \quad \theta^* = \varphi^* \quad \text{and} \quad \sigma^2 = \beta \tau^2.$$

Let $\eta > 0$ and $\chi_N = \frac{\pi^2 \theta^* \sigma^2}{2(\log P)^2}$. By [20, Lemma 4.7], there exists $\gamma > 0$ such that for all $N \geq 1$ large enough, we have

$$\mathbf{P}\left(\forall n \leq (\log P)^3, Y_n(P) - nw \leq -n(1 + \eta)\chi_N\right) \leq \exp(-P^\gamma). \qquad (7.14)$$

We observe that, as in the proof of Lemma 7.6, $(Y_n(j), j \leq P)_n$ can be constructed as the values taken at discrete times by a continuous-time F-BRW. More precisely, we introduce the càdlàg process (F_t) defined by $F_{n\beta} = P$ for any $n \geq 0$ and for any $t \in (n\beta, (n+1)\beta)$, F_t is the number of descendants at time t of particles belonging to $\mathcal{N}_{n\beta}^F$. We have

$$\left((Y_{n\beta}(u), u \in \mathcal{N}_{n\beta}^F), n \geq 0\right) \stackrel{(d)}{=} ((Y_n(j), j \leq N), n \geq 0),$$

therefore we can identify these two processes. For any $n \in \mathbb{N}$, we introduce the event $\mathcal{A}_n^N = \{\max_{t \leq \beta n} F_t \leq N\}$. We recall that by Lemma 7.4, we can couple X and Y in such a way that

$$\forall x \in \mathbb{R}, \#\{u \in \mathcal{N}_{n\beta}^F : Y_{n\beta}(u) \geq x\} \leq \#\{u \in \mathcal{N}_{n\beta}^N : X_{n\beta}(u) \geq x\} \text{ a.s. on } \mathcal{A}_n^N.$$

We bound from below the probability for $\mathcal{A}_n^N$ to occur. As $\mathbf{P}(\#L = 0) = 0$, the process F is increasing on each interval $(n\beta, (n+1)\beta)$. Moreover, observe that $F_{\beta-}$ is the sum of P i.i.d. random variables, with the same distribution as $\#\mathcal{N}_\beta$. This random variable has mean smaller than $1/\alpha$ and exponential moments (by (7.4)). By Cramér's large deviations theorem, there exists $\rho < 1$ such that $P(F_{\beta-} > N) < \rho^N$. Therefore

$$\mathbf{P}(\mathcal{A}_n^{N\,c}) \leq \sum_{j=0}^{n-1} \mathbf{P}(F_{j\beta-} > N) \leq n\rho^N. \qquad (7.15)$$

We now construct a particle process $\widetilde{Y}$, based on the F-BRW Y that bounds from below the N-BRW X. Let $n_N = (\log P)^3$, we set $T_0 = 0$. For any $t \geq 0$, we write $\widetilde{\mathcal{N}}_t$ the set of particles in $\widetilde{Y}$ alive at time t and $\widetilde{m}_t = \min_{u \in \widetilde{\mathcal{N}}_t} \widetilde{Y}_t(u)$. The particle process $\widetilde{Y}$ behaves as Y until the waiting time

$$T_1 = \min(\beta n_N, T_1^{(1)}, T_1^{(2)}), \quad \text{where } T_1^{(1)} = \inf\{t \geq 0 : F_t \geq N\}$$

$$\text{and } T_1^{(2)} = \beta \inf\{n \in \mathbb{N} : \widetilde{m}_{n\beta} > n(w - \chi_N(1+\eta))\}.$$

At time T_1, every particle in $\widetilde{Y}$ is killed and replaced by P particles positioned at $\widetilde{m}_{T_1-}$ if $F_{T_1} > N$ (i.e. $T_1 = T_1^{(1)}$) and at position $\widetilde{m}_{T_1}$ otherwise. Observe that by Lemma 7.4, in both cases there are at time T_1 at least P particles in X that are to the right of the P newborn particles in $\widetilde{Y}$.

Let $k \in \mathbb{N}$, we assume the process $\widetilde{Y}$ has been constructed until time T_k. After this time, it evolves as an F-BRW until time

$$T_{k+1} = \min(T_k + \beta n_N, T_{k+1}^{(1)}, T_{k+1}^{(2)}), \quad \text{where } T_{k+1}^{(1)} = \inf\{t \geq T_k : F_t \geq N\}$$

$$\text{and } T_{k+1}^{(2)} = T_k + \beta \inf\{n \in \mathbb{N} : \widetilde{m}_{T_k + \beta n} - \widetilde{m}_{T_k} > n(w - \chi_N(1+\eta))\}.$$

At time T_{k+1}, every particle in $\widetilde{Y}$ is killed and replaced by P particles positioned at $\widetilde{m}_{T_{k+1}-}$ if $F_{T_{k+1}} > N$ (i.e. $T_{k+1} = T_{k+1}^{(1)}$) and at position $\widetilde{m}_{T_{k+1}}$ otherwise.

By recurrence and the construction of the process, we observe that $\widetilde{Y}$ can be coupled with X in such a way that for any $t \geq 0$, we have

$$\forall x \in \mathbb{R}, \#\{u \in \mathcal{N}_t^F : \widetilde{Y}_t(u) \geq x\} \leq \#\{u \in \mathcal{N}_t^N : X_t(u) \geq x\}.$$

As $\widetilde{m}_t \leq \max_{u \in \mathcal{N}_t^N} X_t(u)$ for any $t > 0$ we obtain $\limsup_{t \to +\infty} \dfrac{\widetilde{m}_t - tv}{t} \leq v_N - v$.

Moreover, observe that $(T_{k+1} - T_k)_k$ and $(\widetilde{m}_{T_{k+1}} - \widetilde{m}_{T_k})_k$ are i.i.d. sequences. Consequently, by the law of large numbers we conclude that

$$\frac{\mathbf{E}(\widetilde{m}_{T_1} - T_1 v)}{\mathbf{E}(T_1)} \leq v_N - v.$$

As a consequence, it is enough to bound from below $\mathbf{E}(\widetilde{m}_{T_1} - T_1 v)$ to conclude the proof.

We introduce the event $G = \{T_1 = T_1^{(2)} < T_1^{(1)}\}$. By definition of T_1,

$$\mathbf{E}(\widetilde{m}_{T_1} - T_1 v) \geq \mathbf{E}\left(-\tfrac{T_1}{\beta}\chi_N(1+\eta)\mathbf{1}_G\right) + \mathbf{E}\left((\widetilde{m}_{T_1-} - T_1 v)\mathbf{1}_{G^c}\right). \qquad (7.16)$$

Observe that until time T_1-, $\widetilde{Y}$ behaves as an F-BRW. As a consequence, using a slight modification of Lemma 7.4, similar to [20, Corollary 4.2], we can couple $\widetilde{Y}$ on $[0, T_1)$ with P independent random walks $((Z_t^j)_t, j \leq P)$, that jump at rate λ according to the law $\max L$ in such a way that

$$\forall t < T_1, \widetilde{m}_t \geq \min_{j \leq P} Z_t^j := \underline{Z_t} \quad \text{a.s.}$$

In particular $\mathbf{E}\left((\widetilde{m}_{T_1-} - T_1 v)\mathbf{1}_{G^c}\right) \geq \mathbf{E}\left((\underline{Z}_{T_1-} - T_1 v)\mathbf{1}_{G^c}\right)$. Using the Cauchy-Schwarz inequality and (7.5), (7.14) and (7.15), we have

$$
\begin{aligned}
\mathbf{E}\left((\underline{Z}_{T_1-} - T_1 v)\mathbf{1}_{G^c}\right)^2 &\leq \mathbf{E}((\underline{Z}_{T_1-} - T_1 v)^2)\mathbf{P}\left(G^c\right) \\
&\leq P\,\mathbf{E}((Z_{T_1-} - T_1 v)^2)\mathbf{P}\left(G^c\right) \\
&\leq Pn_N^2\left(e^{-P^\gamma} + n_N\rho^N\right) = o\left((\log N)^{-4}\right).
\end{aligned}
$$

As a consequence, (7.16) yields

$$
\liminf_{N\to+\infty}(\log N)^2(v_N - v) \geq \liminf_{N\to+\infty} -(\log N)^2\frac{\chi_N}{\beta}(1+\eta) \geq -\frac{\pi^2\theta^*\sigma^2}{2\beta}(1+\eta).
$$

As $\varphi^* = \theta^*$ and $\tau^2 = \frac{\sigma^2}{\beta}$, we conclude the proof by letting $\eta \to 0$. $\qquad\square$

The last statement of Lemma 7.3 is a combination of Lemmas 7.6 and 7.7.

References

[1] Miklós Ajtai, János Komlós, and Endre Szemerédi. The longest path in a random graph. *Combinatorica*, 1(1):1–12, 1981.

[2] David Aldous and Jim Pitman. The asymptotic speed and shape of a particle system. In *Probability, statistics and analysis*, volume 79 of *London Math. Soc. Lecture Note Ser.*, pages 1–23. Cambridge Univ. Press, Cambridge-New York, 1983.

[3] Amnon B. Barak and Paul Erdős. On the maximal number of strongly independent vertices in a random acyclic directed graph. *SIAM J. Algebraic Discrete Methods*, 5(4):508–514, 1984.

[4] Jean Bérard and Jean-Baptiste Gouéré. Brunet-Derrida behavior of branching-selection particle systems on the line. *Comm. Math. Phys.*, 298(2):323–342, 2010.

[5] Jean Bérard and Pascal Maillard. The limiting process of N-particle branching random walk with polynomial tails. *Electron. J. Probab.*, 19:no. 22, 17, 2014.

[6] J. D. Biggins. The first- and last-birth problems for a multitype age-dependent branching process. *Advances in Appl. Probability*, 8(3):446–459, 1976.

[7] Eric Brunet and Bernard Derrida. Shift in the velocity of a front due to a cutoff. *Phys. Rev. E (3)*, 56(3, part A):2597–2604, 1997.

[8] Ksenia Chernysh and Sanjay Ramassamy. Coupling any number of balls in the infinite-bin model. 2016. arXiv:1607.04501.

[9] Joel E Cohen and Charles M Newman. Community area and food-chain length: theoretical predictions. *American Naturalist*, pages 1542–1554, 1991.

[10] D. Denisov, S. Foss, and T. Konstantopoulos. Limit theorems for a random directed slab graph. *Ann. Appl. Probab.*, 22(2):702–733, 2012.

[11] S. Foss and T. Konstantopoulos. Extended renovation theory and limit theorems for stochastic ordered graphs. *Markov Process. Related Fields*, 9(3):413–468, 2003.

[12] Sergey Foss, James B. Martin, and Philipp Schmidt. Long-range last-passage percolation on the line. *Ann. Appl. Probab.*, 24(1):198–234, 2014.

[13] Sergey Foss and Stan Zachary. Stochastic sequences with a regenerative structure that may depend both on the future and on the past. *Adv. in Appl. Probab.*, 45(4):1083–1110, 2013.

[14] E. Gelenbe, R. Nelson, T. Philips, and A. Tantawi. An approximation of the processing time for a random graph model of parallel computation. In *Proceedings of 1986 ACM Fall Joint Computer Conference*, ACM '86, pages 691–697, Los Alamitos, CA, USA, 1986. IEEE Computer Society Press.

[15] Marco Isopi and Charles M Newman. Speed of parallel processing for random task graphs. *Communications on Pure and Applied Mathematics*, 47(3):361–376, 1994.

[16] Olav Kallenberg. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.

[17] Takis Konstantopoulos and Katja Trinajstić. Convergence to the Tracy-Widom distribution for longest paths in a directed random graph. *ALEA Lat. Am. J. Probab. Math. Stat.*, 10(2):711–730, 2013.

[18] A. E. Kyprianou. A note on branching Lévy processes. *Stochastic Process. Appl.*, 82(1):1–14, 1999.

[19] Pascal Maillard. Speed and fluctuations of N-particle branching Brownian motion with spatial selection. 2013. arXiv:1304.0562.

[20] Bastien Mallein. N-Branching random walk with α-stable spine. 2015. arXiv:1503.03762.

[21] Charles M Newman. Chain lengths in certain random directed graphs. *Random Structures and Algorithms*, 3(3):243–253, 1992.

[22] CM Newman and JE Cohen. A stochastic theory of community food webs iv: theory of food chain lengths in large webs. *Proceedings of the Royal Society of London B: Biological Sciences*, 228(1252):355–377, 1986.

[23] Richard P Stanley. Enumerative combinatorics. vol. 1, vol. 49 of cambridge studies in advanced mathematics, 1997.

[24] Craig Aaron Tovey. *Polynomial Local Improvement Algorithms in Combinatorial Optimization*. PhD thesis, Stanford, CA, USA, 1981. AAI8202049.

[25] Kōhei Uchiyama. Spatial growth of a branching process of particles living in $\mathbf{R}^d$. *Ann. Probab.*, 10(4):896–918, 1982.

Bastien Mallein, LAGA - Institut Galilée, 99 avenue Jean-Baptiste Clément 93430 Villetaneuse, France
E-mail address: mallein@math.univ-paris13.fr

Sanjay Ramassamy, Unité de Mathématiques Pures et Appliquées, École normale supérieure de Lyon, 46 allée d'Italie, 69364 Lyon Cedex 07, France
E-mail address: sanjay.ramassamy@ens-lyon.fr