Analyticity of the growth rate of the longest path in Barak-Erdos graphs

Bastien Mallein, Sanjay Ramassamy

Analyticity of the growth rate of the longest path in Barak-Erdős graphs

Bastien Mallein Sanjay Ramassamy

December 28, 2017

Abstract

In this article, we prove that for any probability distribution μ on $\mathbb{N}$ one can construct a stationary version of the infinite-bin model –an interacting particle system introduced by Foss and Konstantopoulos– with move distribution μ. Using this result, we obtain a new formula for the speed of the front of infinite-bin models, as a series of positive terms. This implies that the growth rate $C(p)$ of the longest path in a Barak-Erdős graph of parameter p is analytic on $(0, 1]$.

1 Introduction and main results

In this article, we study an interating particle system on $\mathbb{Z}$ called the infinite-bin model. Using a stationary construction of this model, we are able to show analyticity results for a directed acyclic version of Erdős-Rényi random graphs called Barak-Erdős graphs. We first introduce the latter before describing the former.

1.1 Barak-Erdős graphs

Given an integer $n \geq 1$ and a parameter $0 \leq p \leq 1$, the Barak-Erdős graph $G_{n,p}$ is the graph with vertex set $\{1, \ldots, n\}$ obtained by adding an edge directed from i to j with probability p for every pair (i, j) with $1 \leq i < j \leq n$, independently for each pair. This model was introduced by Barak and Erdős [2] and has since then been widely considered. The most studied feature of these graphs $G_{n,p}$ has been the length of their longest path $L_n(p)$, with applications including food chains [5, 15], the speed of parallel processes [10, 11], last passage percolation [8] and the stability of queues [7]. Some extensions of the model were considered in [6, 12].

Newman [14] proved that there exists a function $C : [0, 1] \to [0, 1]$ such that for any $0 \leq p \leq 1$,

$$\lim_{n\to\infty} \frac{L_n(p)}{n} = C(p) \quad \text{in probability.} \tag{1.1}$$

Moreover he showed that the function C is continuous, differentiable at 0 and that $C'(0) = e$ (see Figure 1 for a plot of $C(p)$).

Foss and Konstantopoulos [7] obtained upper and lower bounds on the function C which are tight in a neighborhood of 1. In [13], we proved that C is an

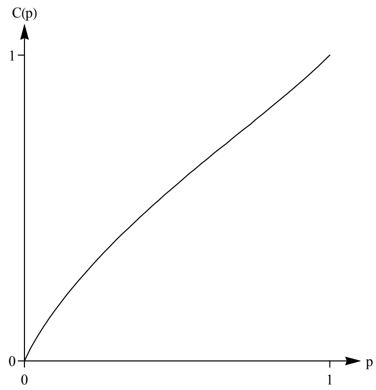

Figure 1: Plot of a simulation of $C(p)$.

analytic function on $(\frac{1}{2}, 1]$ and showed that the power series expansion of $C(p)$ centered at 1 has integer coefficients. Moreover, we proved that

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

In particular, this implied that C has no second derivative at $p = 0$. We raised the question whether there exists a phase transition for some $0 < p_0 < 1$, where the function C stops being analytic. In this paper we provide a negative answer to this question.

Theorem 1.1. *The function $p \mapsto C(p)$ is analytic on $(0, 1]$.*

The proof of this result uses a coupling, introduced by Foss and Konstantopoulos in [7], between Barak-Erdős graphs and an interacting particle system called the infinite-bin model. The infinite-bin model has been the main tool to study Barak-Erdős graphs in both [7] and [13]. Theorem 1.1 follows from a finer analysis of this particle system. In particular we obtain new formula for the speed of the front of this process. As a result we obtain a formula for $C(p)$ as the sum of a series of positive terms.

1.2 The infinite-bin model

The infinite-bin model is an interacting particle system on $\mathbb{Z}$ that can be described as follows. Consider a set of bins indexed by $\mathbb{Z}$, each bin containing a finite number of balls. There is a well-defined notion of front, which is a non-empty bin such that all the bins to its right are empty and the bins to its left are non-empty. Given an i.i.d. sequence $(\xi_n, n \geq 1)$ of positive integers, the infinite-bin model evolves as follows : at every time n, a new ball is added to the right of the ξ_nth rightmost ball.

Alternatively, the infinite-bin model can be thought of as a discrete-time branching particle system, in which the kth rightmost particle reproduces at time n with probability $\mu(k)$. More precisely, if we consider a continuous-time version of the model in which each new ball appears after an exponential random time of parameter 1, in terms of the branching process each particle reproduces independently and the kth rightmost particle creates a new child to its right at rate $\mu(k)$.

This process was introduced by Foss and Konstantopoulos [7], in order to study Barak-Erdős graphs, and further studied in [9, 4, 13]. Note that in the original description [7], the sequence (ξ_n) is only supposed to be stationary and ergodic, not necessarily i.i.d. Most of the results we state in this article could be extended to ergodic settings with mild assumptions, but we stick to the i.i.d. setting to keep the proofs simple.

We introduce some notation to define infinite-bin models more precisely. A *configuration* of balls X is defined to be any collection $(X(k), k \in \mathbb{Z}) \in \mathbb{Z}_+^{\mathbb{Z}}$ (where $X(k)$ represents the number of balls in the bin of index k) such that there exists $F(X) \in \mathbb{Z}$ satisfying the following two conditions :

1. every bin with an index smaller or equal to $F(X)$ is non-empty ;
2. every bin with an index strictly larger than $F(X)$ is empty.

The index $F(X)$ of the rightmost non-empty bin is called the position of the *front* of the configuration X. We denote by S the set of configurations. Note that one could allow bins containing infinitely many balls or even empty bins to the left of the front, provided the total number of balls in a configuration is infinite, and the results we state below would extend in straightforward way to this setting. However we will not consider these cases in this article, again for the sake of simplifying the proofs.

Let $k \geq 1$ be an integer, we define a *move of type* k as a map Φ_k from the set of configurations to itself. For any $X \in S$ and $k \in \mathbb{N}$, we set

$$N(X,k) = \sum_{j \geq k} X(j) \quad \text{and} \quad B(X,k) = \sup\{n \in \mathbb{N} : N(X,n) \geq k\},$$

which are respectively the number of balls in or to the right of bin number k and the index of the bin containing the kth rightmost ball. We define

$$\Phi_k : X \in S \mapsto (X(j) + \mathbf{1}_{\{j = B(X,k)+1\}}, j \in \mathbb{Z}) \in S.$$

In other words, $\Phi_k(X)$ is the configuration obtained from X by adding one ball to the bin to the right of the bin containing the kth rightmost ball.

Let μ be a probability distribution on $\mathbb{N}$ and X_0 be an initial configuration, which may be deterministic or random. We construct $(X_n)_{n \geq 0}$ the infinite-bin model with move distribution μ (or IBM(μ) for short) as the following stochastic recursive sequence in S:

$$\forall n \in \mathbb{N}, \quad X_{n+1} = \Phi_{\xi_{n+1}}(X_n),$$

where $(\xi_n)_{n \geq 1}$ is an i.i.d. sequence of random variables of law μ. Foss and Konstantopoulos [7] proved that when μ has finite expectation there exists a constant $v_\mu \in [0,1]$ independent of the initial configuration X_0 such that

$$\lim_{n \to \infty} \frac{F(X_n)}{n} = v_\mu \quad \text{a.s. and in } L^1. \tag{1.2}$$

In [13], we proved that this result holds without any assumption on the measure μ. The constant v_μ is called the speed of the IBM(μ).

In the case where the move distribution is μ_p, the geometric distribution of parameter $p \in [0,1]$, Foss and Konstantopoulos [7] showed that the speed of the infinite-bin model corresponds to the growth rate of the length of the longest path in Barak-Erdős graphs with parameter p, i.e.

$$\forall p \in [0,1], \quad v_{\mu_p} = C(p). \tag{1.3}$$

In this article, we express the speed v_μ as the sum of a series with positive terms for a general move distribution μ. This series representation of v_μ is based on the appearance of special patterns in the sequence $(\xi_n, n \geq 1)$. Therefore we introduce some notation about words before stating the series formula. We denote by $\mathcal{A}^* = \cup_{n\geq 0}\mathbb{N}^n$ the set of finite words with positive integer letters, where by convention $\emptyset$ is the only element of $\mathbb{N}^0$, representing the empty word. Given a word $\alpha \in \mathcal{A}^*$, we denote by $|\alpha|$ the length of α (i.e. its number of letters) and for any $1 \leq k \leq |\alpha|$, we denote by $\alpha(k)$ the kth letter of α. Furthermore, if $\alpha \in \mathcal{A}^* \setminus \{\emptyset\}$, we set $\pi\alpha$ to be the word of length $|\alpha| - 1$ obtained from α by erasing the last letter.

Recall that the map Φ_k denotes a single move of type k. We extend the notation by defining the map Φ_α for every $\alpha = (\alpha(1), \dots, \alpha(n)) \in \mathcal{A}^*$ by

$$\forall X \in S, \quad \Phi_\alpha(X) = \left(\Phi_{\alpha(n)} \circ \Phi_{\alpha(n-1)} \circ \cdots \circ \Phi_{\alpha(1)}\right)(X). \tag{1.4}$$

In other words, $\Phi_\alpha(X)$ is the configuration obtained from X by successively applying the moves of type $\alpha(1), \alpha(2), \dots, \alpha(n)$. Using this notation, we define for every $X \in S$ the set of X-*good words* as

$$\mathcal{P}_X := \{\alpha \in \mathcal{A}^*\backslash\{\emptyset\} : F(\Phi_\alpha(X)) > F(\Phi_{\pi\alpha}(X))\},$$

i.e. the set of finite sequences of moves such that, starting from X, the final move makes the front advance to the right by one unit, by adding a ball in a previously empty bin.

We define the set of *good words* to be the words that are X-good for every starting configuration X, as well as the set of *bad words* to be the words that are X-good for no initial configuration X, i.e.

$$\mathcal{G} := \bigcap_{X\in S} \mathcal{P}_X \quad \text{and} \quad \mathcal{B} := \bigcap_{X\in S} \mathcal{P}_X^c. \tag{1.5}$$

Finally, we define the set of *minimal good words* (respectively *minimal bad words*) as the good (resp. bad) words that have no good (resp. bad) strict suffix:

$$\mathcal{G}_m = \{\alpha \in \mathcal{G} : \forall k < |\alpha|, \alpha_k \notin \mathcal{G}\} \quad \text{and} \quad \mathcal{B}_m = \{\alpha \in \mathcal{B} : \forall k < |\alpha|, \alpha_k \notin \mathcal{B}\}, \tag{1.6}$$

where α_k denotes the suffix of α consisting of its last k letters[1].

The main result of the article regarding the infinite-bin model is the following formula for the speed of its front.

Theorem 1.2. *Fix a probability distribution μ on $\mathbb{N}$ which is not a Dirac. For any $\alpha \in \mathcal{A}^*$, we denote by $w_\mu(\alpha) = \prod_{j=1}^{|\alpha|} \mu(\alpha(j))$. We have*

$$v_\mu = \sum_{\alpha\in\mathcal{G}_m} w_\mu(\alpha) = 1 - \sum_{\alpha\in\mathcal{B}_m} w_\mu(\alpha). \tag{1.7}$$

[1] And not the first k letters, at it is usually the case in Ulam–Harris type encoding.

Note that for $\mu = \delta_k$, the Dirac mass at k, the process is deterministic and $v_{\delta_k} = 1/k$, but for all $k \geq 2$ the equalities (1.7) do not hold.

Remark 1.3. As a consequence of Theorem 1.2 and the Foss-Konstantopoulos coupling, by taking μ to be the geometric distribution of parameter p, we obtain a formula for the growth rate of the length of the longest path in Barak-Erdős graphs : for every $0 \leq p \leq 1$,

$$C(p) = \sum_{\alpha \in \mathcal{G}_m} p^{|\alpha|}(1-p)^{\sum_{j=1}^{|\alpha|}(\alpha(j)-1)}. \tag{1.8}$$

Formula (1.7) for the speed of the infinite-bin model can be compared to the one we obtained in [13]. For every $X \in S$, we defined the map

$$\varepsilon_X : \alpha \in \mathcal{A}^* \setminus \{\emptyset\} \mapsto \mathbf{1}_{\{\alpha \in \mathcal{P}_X\}} - \mathbf{1}_{\left\{\alpha_{|\alpha|-1} \in \mathcal{P}_X\right\}} \in \{-1, 0, 1\}.$$

We showed in [13] that whenever the series

$$\sum_{\alpha \in \mathcal{A}^*} \varepsilon_X(\alpha) w_\mu(\alpha) \tag{1.9}$$

converges absolutely, then its sum is equal to v_μ. However there was no clear condition on μ for the series (1.9) to converge. We only managed to prove its convergence for probability distributions with light enough tails, such as geometric distributions with parameter $p > 1/2$. By contrast, the new formula (1.7) is more tractable, as it only has positive terms and it holds for every non-Dirac probability distribution μ. However, formula (1.9) is still well-adapted for explicit estimates, as the computation of $\varepsilon(\alpha)$ is linear in $|\alpha|$, while verifying that a word α belongs to $\mathcal{G}$ has a complexity which is exponential in the largest letter of α.

Remark 1.4 (Speed of a branching random walk with selection). As observed in [13], the speed w_k of an infinite-bin model with measure ν_k uniform on $\{1, \ldots, k\}$ is the same as the speed of a continuous-time branching random walk with selection of the rightmost k individuals. This particular model was first studied by Aldous and Pitman in [1]. Denoting by $\mathcal{G}_{m,k}$ the set of minimal good words using letters only between 1 and k, Formula (1.7) yields

$$w_k = \sum_{\alpha \in \mathcal{G}_{m,k}} \frac{1}{k^{|\alpha|}}. \tag{1.10}$$

The asymptotic behaviour of the speed of a branching random walk with selection was conjectured by Brunet and Derrida in [3]. In this case, there would exist $\chi > 0$ such that

$$w_k = e - \frac{\chi}{(\log(k) + 3(1+o(1))\log(\log(k)))^2} \quad \text{as } k \to \infty.$$

Theorem 1.2 is based on the construction of a bi-infinite stationary version of the infinite-bin model, i.e. a process for which time takes values in $\mathbb{Z}$ rather than in $\mathbb{Z}_+$. More precisely, we prove that the following result holds.

Theorem 1.5. *Let $(\xi_n, n \in \mathbb{Z})$ be a family of i.i.d. random variables with a non-Dirac distribution μ. Almost surely, there exists a unique process $(Y_n, n \in \mathbb{Z})$ on S such that*

$$F(Y_0) = 0 \quad \text{and} \quad \forall n \in \mathbb{Z}, \quad Y_{n+1} = \Phi_{\xi_{n+1}}(Y_n).$$

Moreover, setting $\mathcal{F}_n = \sigma(\xi_k, k \leq n)$ for any $n \in \mathbb{Z}$, the configuration Y_n considered up to translation of the position of the front is $\mathcal{F}_n$-measurable for every $n \in \mathbb{Z}$.

In [7], Foss and Konstantopoulos proved the existence of a bi-infinite stationary version of the infinite-bin model in the case when μ has finite expectation. They also showed in that case that if one samples an infinite-bin model $(X_n)_{n\geq 0}$ and a bi-infinite process $(Y_n)_{n\in\mathbb{Z}}$ using the same sequence $(\xi_n)_{n\in\mathbb{Z}}$, then (X_n) coupling-converges to (Y_n) which entails the convergence of finite-depth sceneries seen from the front. Their construction was based on going back in time and searching for certain renovation events, which determine where all the balls are placed after the renovation event starts, regardless of what the configuration was before the start of the event. These renovation events have positive probability when μ has finite expectation, but have probability zero otherwise.

In order to construct $(Y_n)_{n\in\mathbb{Z}}$ even when μ has infinite expectation, we observe there exist almost surely finite patterns in the past of Y, which determine the content of a finite number of bins at the front at time 0, regardless of what the configuration was before the appearance of such a pattern.

In the remainder of the article, every probability distribution on $\mathbb{N}$ will be assumed not to be a Dirac, unless otherwise stated.

Outline of the paper In the next section, we prove Theorem 1.5 as well as a coupling-convergence result. In Section 3, we prove Theorem 1.2 by linking the speed of the infinite-bin model to the one of the bi-infinite process. Finally, we prove Theorem 1.1 in Section 4 by proving that the length of the smallest good word in the past of the bi-infinite process has an exponential tail.

2 Coupling words for the bi-infinite process

In this section we fix μ to be a probability distribution on $\mathbb{N}$. The proof of Theorem 1.5 is based on the existence of so-called coupling words, introduced by Chernysh and the second author [4] for the IBM(μ). More precisely, for every $K \in \mathbb{Z}_+$ we introduce the projection

$$\Pi_K : \begin{array}{ccc} S & \longrightarrow & \mathbb{N}^K \\ X & \longmapsto & (X(F(X) - K + 1), \ldots, X(F(X))) \end{array}$$

which associates to a configuration X its *K-scenery seen from the front*, i.e. the number of balls in each of the rightmost K non-empty bins. By convention, if $K = 0$, the target of Π_K is the singleton composed of the empty sequence. The *coupling number* $\mathfrak{C}(\gamma)$ of a word $\gamma \in \mathcal{A}^*$ is defined to be the largest integer $K \geq 0$ such that $\#(\Pi_K \circ \Phi_\gamma(S)) = 1$.

In other words, if applying successively moves of type $\gamma(1), \gamma(2), \ldots, \gamma(|\gamma|)$ brings any starting configuration to configurations with the same K-scenery seen from the front then $K \leq \mathfrak{C}(\gamma)$. For example, we have $\mathfrak{C}(2322) = 1$, as one can check by distinguishing according to the two possible relative positions of the rightmost two balls in an arbitrary initial configuration. A word γ is called *K-coupling* if $\mathfrak{C}(\gamma) \geq K$. If γ is a K-coupling word, then any word which has γ as a suffix is also a K-coupling word. Note however that a word having a K-coupling word γ as a prefix may not be K-coupling. For example, we have $\mathfrak{C}(2322) = 1$ and $\mathfrak{C}(23225) = 0$. Nevertheless, we can control the variation of $\mathfrak{C}$ when adding a suffix. If $\gamma \in \mathcal{A}$ and $a \in \mathbb{N}$, we denote by $\gamma \cdot a$ the word of length $|\gamma| + 1$ obtained by adding the letter a to the end of γ.

Lemma 2.1. *Let $\gamma \in \mathcal{A}$ and $a \in \mathbb{N}$. Then $\mathfrak{C}(\gamma \cdot a) \geq \mathfrak{C}(\gamma) - 1$. Furthermore if $a \leq \mathfrak{C}(\gamma)$ then $\mathfrak{C}(\gamma \cdot a) \geq \mathfrak{C}(\gamma)$.*

Proof. Denote by M the number of balls in the single finite configuration in the image of $\Pi_{\mathfrak{C}(\gamma)} \circ \Phi_\gamma$. We distinguish two cases, whether $a \leq M$ or $a > M$.

We first assume $a \leq M$. In that case, after executing the moves corresponding to the letters of γ, the execution of a selects a ball in the Kth rightmost bin with $K \leq \mathfrak{C}(\gamma)$, and places a ball in the bin immediately to the right of that bin. In particular, $\Pi_{\mathfrak{C}(\gamma)} \circ \Phi_{\gamma \cdot a}(S)$ is still a singleton.

We now assume that $a > M$. Then, the execution of a selects a ball in the Kth rightmost bin with $K > \mathfrak{C}(\gamma)$ (K may depend on the initial configuration before the execution of γ) and places a ball in the bin immediately to the right of that bin. Note that while it might modify the content in the $\mathfrak{C}(\gamma)$th rightmost bin, it does not change the content of any of the rightmost $\mathfrak{C}(\gamma) - 1$ bins. Thus $\Pi_{\mathfrak{C}(\gamma)-1} \circ \Phi_{\gamma \cdot a}(S)$ is a singleton.

This proves that in any case, $\mathfrak{C}(\gamma \cdot a) \geq \mathfrak{C}(\gamma) - 1$. Moreover, as there is necessarily at least one ball in each of the $\mathfrak{C}(\gamma)$ rightmost bins, we know that $M \geq \mathfrak{C}(\gamma)$. Therefore, if $a \leq \mathfrak{C}(\gamma) \leq M$, then $\mathfrak{C}(\gamma \cdot a) \geq \mathfrak{C}(\gamma)$. ☐

Let $(\xi_n, n \in \mathbb{Z})$ be a family of independent random variables with law μ. For any $n_1 < n_2$ in $\mathbb{Z}$, we denote by $\xi_{n_1}^{n_2}$ the word $(\xi_{n_1}, \xi_{n_1+1}, \ldots, \xi_{n_2})$. The following result will be the key for constructing a stationary bi-infinite version of the IBM and computing its speed.

Proposition 2.2. *For $K \in \mathbb{N}$, we set*

$$\tau_K = \inf\{n \in \mathbb{N} : \xi_{-n}^0 \text{ is a } K\text{-coupling word}\}.$$

Then τ_K is finite a.s.

Proof. Let a be the smallest integer in the support of μ. Setting $m = \frac{a(a-1)}{2} + 1$, we denote by a^m the word of length m containing only letters a. We first show that applying Φ_{a^m} to any initial configuration has the effect of making the front advance by at least 1. This can be observed using the partial order $\preccurlyeq$ on S introduced in [13, Section 2], which is such that for any $X \preccurlyeq X'$ in S, $F(X) \leq F(X')$ and for any word $\gamma \in \mathcal{A}^*$, $\Phi_\gamma(X) \preccurlyeq \Phi_\gamma(X')$. For any $n \in \mathbb{Z}$, the smallest configuration in S with the front at position n is $\underline{X}^{(n)}(k) = \mathbf{1}_{\{k \leq n\}}$ and one easily checks that applying Φ_{a^m} to $\underline{X}^{(n)}$ has the effect of making the front advance by 1. Therefore, applying Φ_{a^m} to any configuration makes the front advance by at least 1. We will need this observation towards the end of the proof.

Note that for any $K < K'$, a K'-coupling word is also a K-coupling word, hence $\tau_K \leq \tau_{K'}$. Therefore we can without loss of generality choose the integer K as large as we wish in this proof. We choose K large enough such that

$$2\mu([K,\infty)) < \mu(a)^m, \tag{2.1}$$

i.e. such that it is at least twice less likely to observe a letter larger than K than to observe the pattern a^m in the sequence (ξ_n).

We now introduce the following sequence of waiting times (backwards in time) defined by $T_0 = 0$ and

$$T_{k+1} = \sup\left\{n < T_k : \xi_n \geq K \quad \text{or} \quad (n+m-1 < T_k \text{ and } \xi_n^{n+m-1} = a^m)\right\},$$

and the random walk $S_0 = 0$ and $S_{k+1} = S_k + 2\mathbf{1}_{\left\{\xi_{T_{k+1}} \geq K\right\}} - 1$. In other words, T_{k+1} is the first time before T_k where we see either a letter larger than K or the pattern a^m and S_k counts the difference between the number of times the former versus the latter occurs. We observe that by assumption (2.1), we have $\mathbf{E}(S_1) < -1/3$, thus (S_k) drifts towards $-\infty$. As a result, we know there exists an infinite sequence of times $(R_k, k \geq 0)$ defined by $R_0 = 0$ and $R_{k+1} = \inf\{n > R_k : S_n < S_{R_k}\}$, the time at which S_n reaches its record minimum for the $(k+1)$st time.

Let b be the second largest integer in the support of μ (here we use the fact that μ is not a Dirac). It is a straightforward consequence of [4, Theorem 1.1] that there exists a $(K+1)$-coupling word γ that is written only using the letters a and b[2]. We define the new waiting time

$$L := \inf\left\{k \in \mathbb{N} : \xi_{T_{R_k}-|\gamma|}^{T_{R_k}-1} = \gamma\right\},$$

i.e. the first time that the word γ appears immediately before a time at which the random walk S hits a new minimum. As the appearance of the word γ immediately before time T_{R_k} has positive probability of occurring and is independent of everything that happens after time T_{R_k}, we observe that $L < \infty$ a.s.

Set $N := -T_{R_L} + |\gamma|$. To conclude the proof, it is enough to show that ξ_{-N}^0 is a K-coupling word, which will prove that $\tau_K \leq N < \infty$ a.s. To do so, we prove that for any $T_{R_L} - 1 \leq n < 0$, we have $\mathfrak{C}(\xi_{-N}^n) \geq K+1$. It is true for $n = T_{R_L} - 1$, since $\xi_{-N}^{T_{R_L}-1} = \gamma$ which is a $(K+1)$-coupling word. For any $k \geq 1$, define

$$T'_k := \begin{cases} T_k & \text{if } \xi_{T_k} \geq K \\ T_k + m - 1 & \text{otherwise.} \end{cases}.$$

When reading the word ξ_{-N}^0 from left to right, the time T'_k is the (R_L+1-k)th time that we read either a letter larger or equal to K or the rightmost letter of a pattern a^m. We also set $T'_0 = 0$. One shows by induction on $0 \leq k < R_L$ that for any $T'_{k+1} \leq n < T'_k$, we have

$$\mathfrak{C}\left(\xi_{-N}^n\right) \geq \mathfrak{C}\left(\xi_{-N}^{T_{R_L}-1}\right) + S_{k+1} - S_{R_L}.$$

[2] Note that when $a = 1$, the word γ can be chosen to be the word of length $K+1$ composed only of the letter 1.

This is a consequence of Lemma 2.1 and the fact that every letter at least K decreases the coupling number by at most one, every pattern a^m increases the coupling number by at least one (by the observation made at the beginning of the proof) and all the other patterns do not decrease the coupling number, since this coupling number stays above K. We conclude from the fact that $S_{k+1} \geq S_{R_L}$ if $0 \leq k < R_L$, as R_L is a time when the random walk S hits its minimum. $\square$

Using the a.s. existence of finite times τ_K for $K \in \mathbb{N}$, we deduce Theorem 1.5.

Proof of Theorem 1.5. We construct the configuration Y_0 as follows. For each $K \in \mathbb{N}$, we set the rightmost K non-empty bins of Y_0 to be the single configuration in $\Pi_K \circ \Phi_{\xi^0_{-\tau_K}}(S)$, which is a.s. well-defined as $\tau_K < \infty$ a.s.

This construction is consistent for different values of K and it produces a unique configuration Y_0 by sending K to infinity and requiring that $F(Y_0) = 0$. The variable $\Pi_K(Y_0)$ is measurable with respect to $\xi^0_{-\tau_K}$, so $Y_0 \in \mathcal{F}_0$ a.s. As a result, for any $n > 0$, $Y_n = \Phi_{\xi^n_1}(Y_0)$ is a.s. $\mathcal{F}_n$-measurable.

If $n < 0$, we can do a similar analysis as the one made for $n = 0$. For any $K \in \mathbb{N}$, $\Pi_K(Y_n) \in \mathcal{F}_n$ a.s. Choosing $K > -n$ and using Lemma 2.1, one can deduce $F(Y_n)$ from $\Pi_K(Y_n)$ and from ξ^0_n, since we know that $F(Y_0) = 0$. We conclude that the configuration Y_n is a.s. entirely determined (up to a shift) by the sequence $(\xi_k, k \leq n)$. $\square$

Remark 2.3. The bi-infinite infinite-bin model Y that we constructed using $(\xi_n, n \in \mathbb{Z})$ has a stationary behavior. Indeed, for any $k \in \mathbb{Z}$, the law of any K-scenery seen from the front of Y_k depends only on $(\xi_n, n \leq k)$, which has same law as $(\xi_n, n \leq 0)$. Hence the law of Y_k, up to a shift placing its front at position 0, is the same as the law of Y_0.

Now that we constructed the bi-infinite process Y, we observe using similar methods as in Proposition 2.2, that any infinite-bin model X ends up behaving like this bi-infinite process.

Proposition 2.4 (Coupling-convergence)**.** *Let $(\xi_n, n \in \mathbb{Z})$ be i.i.d. random variables with law μ and fix $X_0 \in S$. We denote by $(Y_n)_{n\in\mathbb{Z}}$ the bi-infinite process defined in Theorem 1.5, and by $(X_n)_{n\geq 0}$ the infinite-bin model constructed with X_0 and the random variables $(\xi_n, n \geq 1)$. For any $K \in \mathbb{N}$, for all $n \geq 0$ large enough, we have*

$$\Pi_K(X_n) = \Pi_K(Y_n) \quad a.s.$$

Note that this lemma in particular implies the convergence of the K-scenery seen from the front for any infinite-bin model X.

Proof. Let $a < b$ be the smallest two integers in the support of μ and let $K > 0$ be a number large enough such that (2.1) holds. We then define the sequence of waiting times $T_0 = 0$ and

$$T_{k+1} = \inf\left\{ n > T_k : \xi_n \geq K \quad \text{or} \quad (n - m + 1 > T_k \text{ and } \xi^n_{n-m+1} = a^m) \right\}.$$

The random walk $S_k := \sum_{j=1}^k (2\mathbf{1}_{\{\xi_{T_k} \geq K\}} - 1)$ drifts to $-\infty$. In particular, for infinitely many integers $n \geq 1$, we have $\sup_{k\geq n} S_k \leq S_n$.

Let γ be a $(K+1)$-coupling word consisting only of the letters a and b. Almost surely, there exists a time N large enough such that the first $|\gamma|$ letters after N spell the word γ and the random walk S observed after time $N + |\gamma|$ is always below its value at time $N+|\gamma|$. Hence, by an argument similar to the one used to prove Proposition 2.2, for all $n \geq N + |\gamma|$, we have $\Pi_K(X_n) = \Pi_K(Y_n)$, which concludes the proof. $\square$

3 Speed of the infinite-bin model

In this section, we use the stationary infinite-bin model Y we constructed in the previous section to obtain formula (1.7) for the speed of the infinite-bin model.

Lemma 3.1. *Let μ be a non-Dirac probability distribution on $\mathbb{N}$ and $(Y_n)_{n\in\mathbb{Z}}$ be a stationary bi-infinite infinite-bin model. We have $v_\mu = \mathbf{P}(F(Y_1) = 1)$.*

Proof. Since $(Y_n)_{n\geq 0}$ is an infinite-bin model with move distribution μ, we have

$$v_\mu = \lim_{n\to\infty} \frac{1}{n}\mathbf{E}(F(Y_n)) = \lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} \mathbf{E}(F(Y_j) - F(Y_{j-1})).$$

As Y is a stationary process, we have

$$\mathbf{E}(F(Y_j) - F(Y_{j-1})) = \mathbf{E}(F(Y_1) - F(Y_0)) = \mathbf{P}(F(Y_1) = 1),$$

thus $v_\mu = \mathbf{P}(F(Y_1) = 1)$. $\square$

We use this expression for v_μ in terms of the bi-infinite process Y to prove Theorem 1.2.

Proof of Theorem 1.2. Let $(\xi_n, n \in \mathbb{Z})$ be i.i.d. random variables with law μ. We introduce the random time

$$T := \inf\{n \geq 0 : \xi^1_{-n} \in \mathcal{G} \cup \mathcal{B}\}.$$

We first note that if ξ^0_{-n} is a 1-coupling word, then we know the value of $Y_0(0)$. In that case, ξ^1_{-n} is either a good word or a bad word, depending on whether $\xi_1 \leq Y_0(0)$ or not. We conclude that $T < \tau_1 < \infty$ a.s. by Proposition 2.2. Lemma 3.1 yields

$$v_\mu = \mathbf{P}(F(Y_1) = 1) = \mathbf{P}(\xi^1_{-T} \in \mathcal{G}) = 1 - \mathbf{P}(\xi^1_{-T} \in \mathcal{B}).$$

Note that if ξ^1_{-T} is good, then it is necessarily a minimal good word (if it had a good strict suffix, T would have been smaller). Similarly, $\xi^1_{-T} \in \mathcal{B}$ implies $\xi^1_{-T} \in \mathcal{B}_m$. Moreover the support of ξ^1_{-T} is the entire set $\mathcal{G}_m \cup \mathcal{B}_m$, as a good (resp. bad) word cannot have a bad (resp. good) suffix. Thus

$$\begin{aligned}\mathbf{P}(\xi^1_{-T} \in \mathcal{G}) = \mathbf{P}(\xi^1_{-T} \in \mathcal{G}_m) &= \sum_{\alpha\in\mathcal{G}_m} \mathbf{P}(\xi^1_{-T} = \alpha)\\ &= \sum_{\alpha\in\mathcal{G}_m} \mathbf{P}(\xi^1_{-|\alpha|+2} = \alpha) = \sum_{\alpha\in\mathcal{G}_m} w_\mu(\alpha),\end{aligned}$$

hence $v_\mu = \sum_{\alpha\in\mathcal{G}_m} w_\mu(\alpha)$. The equality $v_\mu = 1 - \sum_{\alpha\in\mathcal{B}_m} w_\mu(\alpha)$ follows from similar computations. $\square$

Remark 3.2. In order to make the formulas in Theorem 1.2 effective, one needs a criterion to find the minimal good and bad words. Given a word $\alpha \in \mathcal{A}^*$, it suffices to test it against a finite set Σ of configurations to determine whether it is good or bad : if α is X-good (resp. X-bad) for every $X \in \Sigma$, then it is good (resp. bad). Writing

$$h = \max_{1 \le i \le |\alpha|} 1 + \alpha_i - i,$$

the set Σ can be taken to be any set of 2^{h-1} configurations with the front at position 0 such that for any $X \neq X'$ in Σ, the positions of the rightmost h balls in X and X' are not all the same.

4 Analyticity of $C(p)$

Using the formula we obtained for the speed v_μ, we are now able to prove the analyticity of the growth rate C of the length of the longest path in Barak-Erdős graphs.

Proof of Theorem 1.1. For any $p, q \ge 0$, we write

$$D(p,q) := \sum_{\alpha \in \mathcal{G}_m} p^{|\alpha|} q^{\sum_{j=1}^{|\alpha|}(\alpha(j)-1)}. \tag{4.1}$$

As indicated in Remark 1.3, it follows from the Foss-Konstantopoulos coupling [7] between infinite-bin models and Barak-Erdős graphs that for any $0 < p \le 1$,

$$C(p) = v_{\mu_p} = D(p, 1-p),$$

where μ_p denotes the geometric distribution of parameter p.

To prove that C is analytic around some $p_0 \in (0,1]$, it is enough to show that the series (4.1) converges for some pair (p', q') with $p' > p_0$ and $q' > 1 - p_0$. Indeed, one would then deduce that all the series of derivatives of $D(p, 1-p)$ converge normally around p_0.

Recall that $T = \inf\{n \ge 0 : \xi^1_{-n} \in \mathcal{G} \cup \mathcal{B}\}$. For any probability distribution μ on $\mathbb{N}$, we denote by $\mathbf{E}_\mu$ the expectation associated with the IBM(μ). By simple computations similar to the proof of Theorem 1.2, for any $r > 0$ we have

$$\mathbf{E}_{\mu_p}(r^{T+2}) \ge \mathbf{E}_{\mu_p}(r^{T+2} \mathbf{1}_{\{\xi^1_{-T} \in \mathcal{G}\}}) = D(rp, 1-p),$$

As a result, to conclude the proof, it is enough to show that T admits some exponential moments uniformly in p. More precisely, we will prove that for every $s \in (0,1]$, there exists $r_s > 1$ such that

$$\forall p \in [s,1], \mathbf{E}_{\mu_p}(r_s^T) < \infty. \tag{4.2}$$

Then, for any $0 < p_0 \le 1$, choosing p such that

$$\max\left(\frac{p_0}{2}, \frac{p_0}{r_{p_0/2}}\right) < p < p_0$$

and setting $p' = p r_{p_0/2}$ and $q' = 1 - p$, one obtains the convergence of the series $D(p', q')$, which will prove Theorem 1.1.

Recall from the proof of Theorem 1.2 that T is smaller than τ_1, the smallest time such that $\xi^0_{-\tau_1}$ is a 1-coupling word. To bound $\mathbf{E}_{\mu_p}(r^{\tau_1})$ we use a construction similar to the one in the proof of Proposition 2.2. Fix $s \in (0,1]$. We choose an integer $K \geq 1$ large enough such that $2(1-s)^{1-K} \leq s$. Then for every $p \in [s,1]$, we have

$$2\mu_p([K,\infty)) \leq 2(1-s)^{1-K} \leq s \leq \mu_p(1).$$

We now introduce the sequence defined by $T_0 = 0$ and for any $k \geq 0$,

$$T_{k+1} = \sup\{n < T_k : \xi_n = 1 \quad \text{or} \quad \xi_n \geq K\}.$$

We also set $S_k = -k + 2\sum_{j=1}^{k} \mathbf{1}_{\{\xi_{T_j} \geq K\}}$. For every $p \in [s,1]$, by the choice of K, S is a nearest-neighbor random walk such that $\mathbf{E}(S_1) \leq -1/3$. We denote by $(R_k, k \geq 0)$ the sequence of strictly descending ladder times of S (i.e. R_k is the kth time when S reaches its record minimum) and by γ the word consisting in $K+1$ times the letter 1. Then, setting

$$L := \inf\left\{k \in \mathbb{N} : \xi_{T_{R_{k(K+1)}}-K-1}^{T_{R_{k(K+1)}}-1} = \gamma\right\},$$

we have $\tau_1 \leq -T_{R_L(K+1)} + K + 1$.

As (S_k) is a random walk with negative drift smaller than $-1/3$, for any $k \in \mathbb{N}$, the random variable R_k is stochastically dominated by the sum of k i.i.d. random variables U_j with the law of the first hitting time of -1 by a nearest-neighbor random walk with drift $-1/3$. Moreover, we observe that

$$\mathbf{E}(r^{U_j}) = \frac{2}{3}r + \frac{1}{3}r\,\mathbf{E}(r^{U_j})^2,$$

by decomposition with respect to the first step of the random walk. Therefore, for any $k \in \mathbb{N}$, $p \in [s,1]$ and $r < \frac{3}{2\sqrt{2}}$, we have

$$\mathbf{E}_{\mu_p}\left(r^{R_k}\right) \leq \left(\frac{3-\sqrt{9-8r^2}}{2r}\right)^k.$$

Similarly, we observe that for any $p \in [s,1]$, $-T_1$ is stochastically dominated by a geometric random variable with parameter s, as this is the minimal probability for obtaining a 1. Then $-T_k$ is stochastically dominated by the sum of k i.i.d. copies of a geometric random variable with parameter s. Thus, by conditioning with respect to R_k, as long as $1 < r < \frac{1}{1-\frac{(3-2\sqrt{2})}{3}s}$, for all $p \in [s,1]$ and $k \in \mathbb{N}$ we have

$$\mathbf{E}_{\mu_p}(r^{-T_{R_k}}) \leq \mathbf{E}_{\mu_p}\left(\left(\mathbf{E}_{\mu_s}\left(r^{-T_1}\right)\right)^{R_k}\right) \leq \left(\frac{3-\sqrt{9-8\left(\frac{sr}{1-(1-s)r}\right)^2}}{2\frac{sr}{1-(1-s)r}}\right)^k.$$

Finally, as L can be stochastically dominated by a geometric random variable with parameter s^{K+1}, which is independent of $(T_{R_k}, k \geq 1)$, it admits some finite exponential moments. We conclude that τ_1 also admits some exponential moments, uniformly in $p \in [s,1]$. □

We point out that, by computations similar to those above, one could show that for any probability distribution μ and $K \in \mathbb{N}$, there exists $r > 1$ such that $\mathbf{E}_\mu(r^{\tau_K}) < \infty$.

Acknowledgements We would like to thank an anonymous referee of our previous paper [13] for a comment inspiring us to study the set of good words.

References

[1] D. Aldous and J. 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.

[2] A. B. Barak and P. 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.

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

[4] Ksenia Chernysh and Sanjay Ramassamy. Coupling any number of balls in the infinite-bin model. *Journal of Applied Probability*, 54(2):540–549, 2017.

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

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

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

[8] S. Foss, J. B. Martin, and P. Schmidt. Long-range last-passage percolation on the line. *Ann. Appl. Probab.*, 24(1):198–234, 2014.

[9] S. Foss and S. 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.

[10] 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.

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

[12] T. Konstantopoulos and K. 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.

[13] Bastien Mallein and Sanjay Ramassamy. Barak-Erdös graphs and the infinite-bin model. *arXiv preprint arXiv:1610.04043*, 2016.

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

[15] C. M. Newman and J. E. 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.

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`