

# Estimation of the last passage percolation constant in a charged complete directed acyclic graph via perfect simulation

SERGEY FOSS, TAKIS KONSTANTOPOULOS, BASTIEN MALLEIN  
and SANJAY RAMASSAMY

## Abstract

Our object of study is the asymptotic growth of heaviest paths in a charged (weighted with signed weights) complete directed acyclic graph. Edge charges are i.i.d. random variables with common distribution  $F$  supported on  $[-\infty, 1]$  with essential supremum equal to 1 (a charge of  $-\infty$  is understood as the absence of an edge). The asymptotic growth rate is a constant that we denote by  $C(F)$ . Even in the simplest case where  $F = p\delta_1 + (1-p)\delta_{-\infty}$ , corresponding to the longest path in the Barak-Erdős random graph, there is no closed-form expression for this function, but good bounds do exist. In this paper we construct a Markovian particle system that we call “Max Growth System” (MGS), and show how it is related to the charged random graph. The MGS is a generalization of the Infinite Bin Model that has been the object of study of a number of papers. We then identify a random functional of the process that admits a stationary version and whose expectation equals the unknown constant  $C(F)$ . Furthermore, we construct an effective perfect simulation algorithm for this functional which produces samples from the random functional.

*Keywords and phrases.* PERFECT SIMULATION, COUPLING (FROM THE PAST), RANDOM GRAPH, MARKOV PROCESS, STATIONARITY, LAST PASSAGE PERCOLATION

*AMS 2010 subject classification.* Primary 82M31; secondary 60K15, 60G10, 05C80

## 1 Preliminaries

A Barak-Erdős random graph is a directed acyclic version of the standard Erdős-Rényi graph [2]. We let  $\mathbb{Z}^+$ , the set of non-negative integers, serve as the set of vertices. For each pair of vertices  $i, j$  with  $i < j$ , declare  $(i, j)$  as an edge directed from  $i$  to  $j$  with probability  $p$ , independently from any other pair. Then the maximum length  $L_n$  of all paths from vertex 0 to  $n$  satisfies a law of large numbers [20, 8]:  $\lim_{n \rightarrow \infty} L_n/n = C(p)$ , a.s., where  $C(p)$  is a certain deterministic, increasing, analytic function of  $0 < p < 1$  [18, 19]. Owing to the fact that such a graph appears as a model in various natural applications, such as in computer systems [13, 14], in mathematical ecology [20, 21] and others, information about  $C(p)$  has been the subject of a number of papers [20, 8, 6, 10, 18, 19, 9, 11].

We take interest in a generalization of this problem, considering the Last Passage Percolation problem on a complete directed acyclic graph, in which each edge has a signed charge distributed according to an independent copy of the random variable  $w$  taking values in  $\mathbb{R} \cup \{-\infty\}$ , such

that  $\inf\{z \in \mathbb{R} : \mathbb{P}(w > z) = 0\} < \infty$ . We denote by  $F$  the law of  $w$ , and let  $\{w_{i,j}, 0 \leq i < j\}$ , be a collection of i.i.d. copies of  $w$ . If  $\pi$  is a path from  $i$  to  $j$ , namely an increasing collection of vertices  $(i = i_0, i_1, \dots, i_\ell = j)$  then its charge is defined as the sum of the charges of its edges:  $w(\pi) = w_{i_0, i_1} + \dots + w_{i_{\ell-1}, i_\ell}$ , using the convention that  $-\infty + x = -\infty$  for all  $x \in \mathbb{R}$ . In other words, if a path goes through an edge with charge  $-\infty$ , then the charge of the path is  $-\infty$ . We define by convention the charge of a path consisting of a single vertex as 0.

We are concerned with the quantity

$$W_n := \sup\{w(\pi) : \pi \text{ is a path from } 0 \text{ to } n\}, \quad (1)$$

the maximum charge of all paths between 0 and  $n$ . Observe that  $(W_n, n \geq 0)$  is a superadditive random sequence as direct computations show that for all  $n, m \geq 0$ ,

$$W_{n+m} \geq W_n + W_{n,n+m},$$

with  $W_{n,n+m} = \sup\{w(\pi) : \pi \text{ is a path from } n \text{ to } n+m\}$ . As  $W_{n,n+m}$  is a copy of  $W_m$  independent of  $W_n$ , by Kingman's subadditive ergodic theorem [15] we have

$$\frac{W_n}{n} \rightarrow C(F), \quad \text{a.s. as } n \rightarrow \infty \quad (2)$$

where  $C(F)$  is a deterministic function of the law  $F$ . We refer to  $C(F)$  as the *last passage percolation constant* of  $F$ . The objective of the present article is to develop an approach to compute this constant through Monte Carlo methods.

We denote by

$$L = \inf\{z \in \mathbb{R} : \mathbb{P}(w > z) = 0\}$$

the *essential supremum* of  $F$  (this is the maximal point of the support of distribution  $F$ ).

Consider momentarily the case  $L \leq 0$ . Then edge weights are nonpositive a.s., and, since

$$-W_n = -\inf\{-w(\pi), \pi \text{ is a path from } 0 \text{ to } n\},$$

the problem is that of first passage percolation on the complete directed graph. We claim that  $C(F) = 0$ . Indeed, with  $n \geq 3$ , considering the 2-edge path  $(0, j, n)$ ,

$$0 \geq W_n \geq \sup_{1 \leq j \leq n-1} (w_{0,j} + w_{j,n}) \quad \text{a.s.},$$

hence

$$\begin{aligned} \mathbb{P}(W_n \geq 2(L-1)) &\geq \mathbb{P}(w_{0,j} + w_{j,n} \geq 2(L-1) \text{ for all } 1 \leq j \leq n-1) \\ &\geq 1 - (1 - \mathbb{P}(w > L-1))^2)^{n-1}. \end{aligned}$$

Thus, by the Borel-Cantelli lemma,  $W_n/n \rightarrow 0$  in probability, and hence  $C(F) = 0$ , as claimed. It is not hard to see that  $W_n$  itself converges weakly to the random variable  $\max(w, 2L)$ .

We only consider the case  $L > 0$  in the rest of the article. In this situation, up to replacing  $w$  by  $w/L$ , we will assume without loss of generality that the essential supremum of  $F$  is 1. Hence, we work under the following assumption for the distribution  $F$  of  $w$ :

$$\forall \varepsilon > 0, \quad F([1 - \varepsilon, 1]) > 0 \text{ and } F((1, \infty)) = 0. \quad (3)$$

The case  $F = p\delta_1 + (1 - p)\delta_{-\infty}$  formally corresponds to a Barak-Erdős graph as any edge with charge  $-\infty$  can be ignored. Such a graph was studied in [8] and a more general version of it in [6]. The constant  $C(p)$  mentioned earlier is, with an abuse of notation, the constant  $C(p\delta_1 + (1 - p)\delta_{-\infty})$ . We know that  $C(p) > 0$  for all  $p > 0$  which means that eventually, any two vertices that are far apart enough are connected by a path that has charge  $> -\infty$ . It was shown in [18, 19] that the function  $p \rightarrow C(p)$  is analytic on  $(0, 1]$  and a two-term asymptotic expansion was given in the limit  $p \rightarrow 0$ .

The case  $F = p\delta_1 + (1 - p)\delta_x$ , where  $x \in (-\infty, 1)$ , was studied in [11]. For this case, the quantity  $C(F)$  was denoted by  $C_p(x)$ , a differentiable function of  $x \in (-\infty, 1) \setminus I$ , where  $I$  is the union of nonpositive rationals and of the inverses  $1/n$ ,  $n \geq 2$ . Moreover  $C(p)$  is the decreasing limit of  $C_p(x)$ , as  $x \rightarrow -\infty$ . In the special case when  $x = 0$ , it was shown in [7] that  $C_p(0) = 1/\psi(1 - p)$  where  $\psi$  is a Ramanujan theta function.

Let  $F$  be a distribution on  $[-\infty, 1]$  with essential supremum 1. Comparing  $F$  with the distribution  $p\delta_{1/2} + (1 - p)\delta_{-\infty}$  where  $p = F([\frac{1}{2}, 1])$ , it is not hard to see that  $C(F) > 0$ . The goal of this paper is to *construct* a random variable with expectation  $C(F)$  that can be perfectly simulated via an explicit algorithm. Perfect simulation of a functional of a Markov chain in its “steady-state” is a technique that, whenever applicable, avoids the bias introduced by standard MCMC (=Markov Chain Monte Carlo) methods, in which one would approach  $C(F)$  by a realization of  $W_n/n$  for  $n$  large enough. The terminology and algorithm was introduced in [22].

A survey can be found in [16]. Its relation to the so-called backwards-coupling was studied in [12] and falls in the studies of coupling methods for stochastic recursions that may entirely lack the Markovian property [3, 5, 8].

Our perfect simulation algorithm is based on the construction of a particle system, that we call *Max Growth System* (MGS) associated to the charged complete directed graph. This particle system can be seen as an extension of the Infinite Bin Model (IBM) [8, 18, 4, 19] arising in connection to the Barak-Erdős graph. We mention *en passant* that the IBM is a particle system in discrete time introduced in [8] but one which falls in a natural class of similar particle systems, manifestations of which have appeared frequently in the literature, e.g. in [1].

In Section 2, we first define the MGS with charge distribution  $F$  and describe some properties of its dynamics. In Section 3, we show that a certain functional of the MGS is a Markov chain that admits a stationary version. In Section 4, we pull the random variable mentioned above from the stationary version, show that its expectation is  $C(F)$  and describe a perfect simulation algorithm. We conclude by suggesting further directions of research in Section 5.

## 2 The Max Growth System (MGS)

The Max Growth System is a particle system on  $\mathbb{R} \cup \{-\infty\}$  in which at every step a new atom is added to the process. This auxiliary particle system is constructed in such a way that starting from a single particle at position 0, the  $n$ th particle in the system will be placed at position  $W_n$ .

## 2.1 Deterministic dynamics of the MGS

We let  $\mathcal{N}$  be the set of locally finite point measures on  $\mathbb{R} \cup \{-\infty\}$  with a finite maximal element, namely, measures whose values are nonnegative integers, which are finite on every interval of the form  $[x, \infty)$  and which are positive on  $\mathbb{R}$ . This will be the state space on which the MGS is defined. Any such measure  $\nu \in \mathcal{N}$  is specified by the nonincreasing sequence  $\nu_1 \geq \nu_2 \geq \dots \geq -\infty$  of the locations of the points (atoms) of  $\nu$ . This sequence may be finite or infinite. For example,  $\nu = 2\delta_0 + \delta_{-1.5} + 3\delta_{-4}$  is equivalently represented by the finite sequence  $(0, 0, -1.5, -4, -4, -4)$ . We shall therefore think of any  $\nu \in \mathcal{N}$  either as a point measure  $\nu = \sum_{k \geq 1} \delta_{\nu_k}$  or as a sequence  $(\nu_1, \nu_2, \dots)$ . Note that the zero measure  $0$  is an element of  $\mathcal{N}$  and corresponds to an empty sequence of points. The total mass  $\|\nu\| = \nu(\mathbb{R})$  of  $\nu$  is the number of its points (counted with multiplicity). We let  $\inf \nu := \nu_{\|\nu\|}$  be the location of the last point of  $\nu$  if  $\|\nu\| < \infty$ . If  $\|\nu\| = \infty$ , we let  $\inf \nu = -\infty$ .

Let  $w = (w_1, w_2, \dots)$  be a sequence of elements of  $\mathbb{R} \cup \{-\infty\}$ , such that  $\sup_{k \geq 1} w_k \leq 1$ . Let  $\mathcal{W}$  be the collection of such sequences. Given  $\nu$  a non-zero element of  $\mathcal{N}$ , define the quantity

$$\mathbf{m}(\nu, w) := \sup_{k \geq 1} (\nu_k + w_k), \quad \nu \neq 0, \quad w \in \mathcal{W}.$$

Here the supremum is taken either over all  $k \geq 1$  if  $\|\nu\| = \infty$  or over  $1 \leq k \leq \|\nu\|$  if  $\|\nu\|$  is finite. Observe that  $\mathbf{m}(\delta_0, w) = w_1$  for all  $w \in \mathcal{W}$ . The map responsible for the dynamics of the MGS process is defined by

$$\Psi_w \nu := \nu + \delta_{\mathbf{m}(\nu, w)},$$

that consists in adding at every step an atom in the process at a position given by  $\mathbf{m}(\nu, w)$ .

We will later employ a ‘‘coupling from the past’’ technique. To this end, it is worth describing the MGS starting from an arbitrary point in time. Let  $(w(t), t \in \mathbb{Z})$  be a sequence of elements of  $\mathcal{W}$ ,  $\nu$  a point measure in  $\mathcal{N}$  and  $T \in \mathbb{Z}$ . The MGS starting from  $\nu$  at time  $T$  is the process  $(\nu(t), t \geq T)$  defined recursively by

$$\nu(T) = \nu \quad \text{and} \quad \nu(t+1) = \Psi_{w(t+1)} \nu(t), \quad t \geq T.$$

When  $(w_j(t), j \in \mathbb{Z}^+, t \in \mathbb{Z})$  is i.i.d. with law  $F$ , we say that  $(\nu(t), t \geq 0)$  is an MGS with charge distribution  $F$ . To simplify notation, for all  $s < t \in \mathbb{Z}$ , we write

$$\Psi_w^{s,t} = \Psi_{w(t)} \circ \Psi_{w(t-1)} \circ \dots \circ \Psi_{w(s)},$$

in which case we have  $\nu(t) = \Psi_w^{T+1,t} \nu$  for all  $t > T$ .

To consider stationary versions of the MGS process, we will sometimes need to work with the particle system seen from the rightmost particle. We denote by  $\mathcal{N}_0$  the set of  $\nu \in \mathcal{N}$  with  $\nu_1 = 0$ . For  $\nu \in \mathcal{N}$ , we define its shift  $\sigma\nu$  seen from the front by

$$\int f(x) d(\sigma\nu)(x) := \int f(x - \nu_1) d\nu,$$

for all  $\nu \in \mathcal{N}$  and all positive bounded functions  $f$ . Thus  $\sigma : \mathcal{N} \rightarrow \mathcal{N}$  and can be thought of as: ‘‘place the origin at the position of the rightmost atom’’. For example,  $\sigma(\delta_a + \delta_b) = \delta_0 + \delta_{-|a-b|}$ .

Observe that  $\sigma$  is a projection of  $\mathcal{N}$  onto  $\mathcal{N}_0$ , which is consistent with the definition of the MGS as

$$\sigma\Psi_w = \sigma\Psi_w\sigma,$$

for all sequences  $w \in \mathcal{W}$ . It is also worth mentioning that, for all  $\nu \in \mathcal{N}$ , we have

$$\mathbf{m}(\sigma\nu, w) = \mathbf{m}(\nu, w) - \nu_1. \quad (4)$$

## 2.2 Decoupling properties of the MGS

The following lemma shows that if there is a large enough gap in between the first and the second atom in the point measure  $\nu$ , and the sequence of charges satisfies a “triangular” property, then the positions of the new particles only depend on a finite number of charges.

**Lemma 1** (Decoupling property). *Fix  $\ell \in [0, 1)$  and a positive integer  $n$ . Let  $T \in \mathbb{Z}$  and  $(w(T+t), t \geq 1)$  be a sequence in  $\mathcal{W}$ . Let  $\nu$  be a point measure in  $\mathcal{N}_0$  such that  $\nu_2 \leq -\ell$ . We define the sequences*

$$\nu(t) = \Psi_{w(t)}\nu(t-1) \quad \text{and} \quad \tilde{\nu}(t) = \Psi_{w(t)}\tilde{\nu}(t-1), \quad t \geq T,$$

with  $\nu(T) = \nu$  and  $\tilde{\nu}(T) = \delta_0$ . For all  $n \in \mathbb{N}$ , if

$$\bar{w}(T; t) := \max\{w_1(T+t), \dots, w_t(T+t)\} \geq 1 - \ell \quad \text{for all } 1 \leq t \leq n, \quad (5)$$

then  $\mathbf{m}(\nu(T+n-1), w(T+n)) = \mathbf{m}(\tilde{\nu}(T+n-1), w(T+n))$ .

*Proof.* It suffices to prove this statement for  $T = 0$ . We prove, by induction, that

$$\bar{w}(0; t) \geq 1 - \ell \text{ for all } 1 \leq t \leq n \Rightarrow \nu(n)|_{\mathbb{R}^+} = \tilde{\nu}(n)|_{\mathbb{R}^+} \text{ and } \nu(n)(\mathbb{R}^+) = n + 1 \quad (6)$$

Assume first that  $n = 1$ . In this case,

$$\mathbf{m}(\nu, w(1)) = \max(w_1(1), \max_{j \geq 2}[\nu_j + w_j(1)]).$$

Since, by assumption,  $\nu_j \leq \nu_2 \leq -\ell$  for all  $j \geq 2$ , we have  $\nu_j + w_j(n) \leq -\ell + 1$  for all  $j \geq 2$ . To prove (6) for  $n = 1$  we must assume that  $w_1(1) \geq 1 - \ell$ . But then  $w_1(1) \geq \max_{j \geq 2}[\nu_j + w_j(1)]$  and so

$$\mathbf{m}(\nu, w(1)) = w_1(1) = \mathbf{m}(\tilde{\nu}(0), w(1)).$$

Hence  $\nu(1) = \nu + \delta_{w_1(1)}$  and, with  $\tilde{\nu}(0) = \delta_0$ ,  $\tilde{\nu}(1) = \delta_0 + \delta_{w_1(1)}$ . Hence (6) holds for  $n = 1$ .

Assume next that (6) holds for some  $n \geq 2$ . We prove that it also holds for  $n + 1$ . To do this, it suffices to assume that  $\nu(n)|_{\mathbb{R}^+} = \tilde{\nu}(n)|_{\mathbb{R}^+}$ ,  $\nu(n)(\mathbb{R}^+) = n + 1$ , and  $\bar{w}(0; t) \geq 1 - \ell$  for all  $1 \leq t \leq n + 1$ . In this case, we have

$$\mathbf{m}(\nu(n), w(n+1)) = \max\left(\max_{j \leq n+1}[\nu_j(n) + w_j(n+1)], \max_{j \geq n+2}[\nu_j(n) + w_j(n+1)]\right).$$

But, for all  $j \leq n+1$ ,  $\nu_j(n) \geq 0$  and so  $\nu_j(n) + w_j(n+1) \geq 1 - \ell$ . Taking into account the assumption  $\nu_2 \leq -\ell$ , we have, for all  $j \geq n+2$ ,  $\nu_j(n) \leq -\ell$  and so

$$\max_{j \leq n+1} [\nu_j(n) + w_j(n+1)] \geq 1 - \ell \geq \max_{j \geq n+2} [\nu_j(n) + w_j(n+1)],$$

which implies that

$$\mathbf{m}(\nu(n), w(n+1)) = \mathbf{m}(\tilde{\nu}(n), w(n+1)) = \max_{j \leq n+1} [\nu_j(n) + w_j(n+1)] \geq 1 - \ell > 0.$$

The configuration  $\nu(n+1)$  is thus obtained by adding a particle to  $\nu(n)$  at a positive location. Since  $\nu(n) = \tilde{\nu}(n)$  on  $\mathbb{R}^+$  and since the particle is added at the same location for both, we have  $\nu(n+1) = \tilde{\nu}(n+1)$  on  $\mathbb{R}^+$ . Clearly,  $\nu(n+1)(\mathbb{R}^+) = \nu(n)(\mathbb{R}^+) + 1 = n+2$ , so (6) holds for  $n+1$ .  $\square$

The above lemma allows us to describe a set of conditions on the sequences  $(w(t))$  so that the increments of  $\nu$  and  $\tilde{\nu}$  are algebraically independent of  $\nu$  and  $\tilde{\nu}$ .

**Corollary 1.** *Let  $T \in \mathbb{Z}$ ,  $n \in \mathbb{N}$ ,  $\ell \in [0, 1)$  and  $(w(T+t), 0 \leq t \leq n)$  a sequence such that*

$$w_1(T) \geq \ell \text{ and } \min\{\bar{w}(T; 1), \bar{w}(T; 2), \dots, \bar{w}(T; n)\} \geq 1 - \ell. \quad (7)$$

*Let  $\nu, \tilde{\nu}$  be two elements of  $\mathcal{N}_0$  and define the sequences*

$$\nu(t) = \Psi_{w(t)}\nu(t-1) \text{ and } \tilde{\nu}(t) = \Psi_{w(t)}\tilde{\nu}(t-1), \quad t \geq T,$$

*with  $\nu(T-1) = \nu$  and  $\tilde{\nu}(T-1) = \tilde{\nu}$ . Then  $\mathbf{m}(\sigma\nu(t-1), w(t)) = \mathbf{m}(\sigma\tilde{\nu}(t-1), w(t))$  for all  $T+1 \leq t \leq T+n$ .*

In other words, the sequence  $(\mathbf{m}(\sigma\nu(t-1), w(t)), T+1 \leq t \leq T+n)$  is algebraically independent of  $\nu(T-1)$  provided that  $w$  satisfies (7).

*Proof.* We observe that as  $w_1(T) \geq \ell$ , we have

$$\mathbf{m}(\nu(T-1), w(T)) \geq \nu_1(T-1) + w_1(T) \geq \ell \quad \text{and} \quad \mathbf{m}(\tilde{\nu}(T-1), w(T)) \geq \ell.$$

Therefore, the second largest atom of  $\sigma\Psi_{w(T)}\nu(T-1)$  and  $\sigma\Psi_{w(T)}\tilde{\nu}(T-1)$  are both smaller than  $-\ell$ , hence by (4) we can apply Lemma 1, which completes the proof.  $\square$

### 2.3 The MGS derived from the charged complete directed graph

Consider the charged complete directed graph with i.i.d. edge charges  $\{w_{i,j}, 0 \leq i < j\}$  of law  $F$ , a collection of i.i.d. random variables in  $\mathbb{R} \cup \{-\infty\}$  with common law  $F$  satisfying assumption (3). For all  $n \in \mathbb{N}$ , we write  $W_n$  for the length of the longest path between 0 and  $n$ . We observe that  $(W_n, n \geq 1)$  can be coupled with the charged complete graph with charge distribution  $F$ .

**Lemma 2.** *Let  $(\nu(t), t \geq 0)$  be an MGS with charge distribution  $F$  such that  $\nu(0) = \delta_0$  and let  $(W_t, t \geq 0)$  as defined in (1). We have the following equality in distribution:*

$$(\nu(t), t \geq 0) \stackrel{(d)}{=} \left( \sum_{j=0}^t \delta_{W_j}, t \geq 0 \right). \quad (8)$$

*Proof.* By definition, we have  $\nu(0) = \delta_0 = \delta_{W_0}$ , using that the path of length 0 between 1 and 1 has mass 0. Let  $t_0 \in \mathbb{Z}^+$  and assume that we can construct a coupling between  $\nu$  and  $W$  such that  $(\nu(t), t \leq t_0) = (\sum_{j=0}^t \delta_{W_j}, t \leq t_0)$  a.s. Conditionally on this coupling, let  $(w_{j,t_0+1}, j \in \mathbb{Z}^+)$  and  $(w_j(t_0 + 1), j \in \mathbb{Z}^+)$  be independent families of i.i.d. random variables with law  $F$ . By (1), decomposing all paths  $\pi$  ending at  $t_0 + 1$  according to their last step, we have

$$\begin{aligned} W_{t_0+1} &= \max\{W_j + w_{j,t_0+1}, 0 \leq j \leq t_0\} \\ &\stackrel{(d)}{=} \max\{\nu_j(t_0) + w_j(t_0 + 1), 1 \leq j \leq t_0 + 1\} = \mathbf{m}(\nu(t_0), w(t_0 + 1)), \end{aligned}$$

therefore  $\nu(t_0 + 1) \stackrel{(d)}{=} \sum_{j=0}^{t_0+1} \delta_{W_j}$ . As a result, we can couple the two sequences of random variables in such a way that the above equality holds almost surely.

Hence, by recursion, there exists a coupling between the MGS and the last passage percolation problem such that (8) holds for all times.  $\square$

A noteworthy observation is that the increments of  $W$  are the same as the relative increments of the MGS. More precisely, defining

$$M_n = \max_{0 \leq k \leq n} W_k = \sup \{w(\pi), \pi \text{ path from } 0 \text{ to } k\}, \quad (9)$$

the increments of the sequence  $(M_n, n \geq 0)$  can be connected to the relative increments of the MGS.

**Corollary 2.** *Under the foregoing assumptions,*

$$(M_n - M_{n-1}, n \geq 1) \stackrel{(d)}{=} (\mathbf{m}(\sigma\nu(n-1), w(n))^+, n \geq 1).$$

*Proof.* In the proof above, we established a coupling between  $(\nu(n), n \geq 0)$  and  $(W_k, k \geq 0)$ . Under this coupling, for  $n \in \mathbb{Z}^+$ , we have

$$W_n = \mathbf{m}(\nu(n-1), w(n)) \quad \text{and} \quad M_n = \nu_1(n).$$

As a result, under this coupling, we have

$$M_n - M_{n-1} = (W_n - M_{n-1})^+ = (\mathbf{m}(\nu(n-1), w(n)) - \nu_1(n-1))^+ = \mathbf{m}(\sigma\nu(n-1), w(n))^+. \quad \square$$

### 3 Stationarity via coupling

We recall that our aim is to compute the quantity  $C(F)$  defined by

$$C(F) := \lim_{n \rightarrow \infty} \frac{W_n}{n} \quad \text{a.s.}$$

As  $F$  has a finite essential supremum, it holds that  $\int_0^\infty xF(dx) < \infty$ . Therefore, by [10], it is known that

$$C(F) = \lim_{n \rightarrow \infty} \frac{M_n}{n} \quad \text{a.s. and in } L^1.$$

Thus, if  $F$  satisfies (3), we have

$$C(F) = \lim_{n \rightarrow \infty} \frac{\mathbb{E}(M_n)}{n}.$$

Using Corollary 2, we remark that for all  $n \in \mathbb{N}$ ,

$$\frac{M_n}{n} = \frac{1}{n} \sum_{j=1}^n (M_j - M_{j-1}) = \frac{1}{n} \sum_{j=1}^n \mathbf{m}(\sigma\nu(j-1), w(j))^+.$$

We show in this section that  $(\mathbf{m}(\sigma\nu(n-1), w(n)), n \geq 1)$  admits a stationary version (where we recall that  $\sigma\nu$  is the point measure shifted so that its rightmost element is at position 0). Since the process  $(\mathbf{m}(\sigma\nu(n-1), w(n)), n \geq 1)$  is not Markovian, the term “stationary version” should be used with caution. For us, it means that it couples with a stationary process in finite time, as in the statement of Theorem 1 below. Then, writing  $\bar{\mathbf{m}}$  the limit in distribution of  $\mathbf{m}(\sigma\nu(n-1), w(n))$  as  $n \rightarrow \infty$ , we have

$$C(F) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \mathbb{E}(\mathbf{m}(\sigma\nu(j-1), w(j))^+) = \mathbb{E}(\bar{\mathbf{m}}^+),$$

as  $\mathbf{m}(\sigma\nu, w)^+ \in [0, 1]$  a.s. In the next section, we introduce the perfect simulation algorithm, which consists in giving a realization of  $\bar{\mathbf{m}}$  without constructing the limit in distribution of  $\sigma\nu(n)$  as  $n \rightarrow \infty$ .

**Theorem 1.** *Suppose that  $F$  is a distribution satisfying (3). Let  $\ell \in [0, 1)$  be such that  $p := F([1 - \ell, 1]) \in (0, 1]$ . Let  $w = (w_i, i \in \mathbb{N})$  be i.i.d. random variables with law  $F$  and  $(w(t), t \in \mathbb{Z})$  i.i.d. copies of  $w$ . Given  $\nu(0) \in \mathcal{N}$ , we define the MGS process by*

$$\nu(t+1) = \Phi_{w(t+1)}\nu(t), \quad \text{for all } t \geq 0.$$

*There exists a stationary process  $(\bar{\mathbf{m}}(t), t \in \mathbb{Z})$  such that*

$$\mathbf{m}(\nu(t-1), w(t)) = \bar{\mathbf{m}}(t) \quad \text{a.s. for } t \text{ large enough.}$$

*In particular  $\mathbb{E}(\bar{\mathbf{m}}(0)^+) = C(F)$ .*



*Proof.* For  $T \in \mathbb{Z}$  and  $t \in \mathbb{N}$ , we recall the notation  $\bar{w}(T; t) = \max\{w_1(T+t), \dots, w_t(T+t)\}$  from Lemma 1. We introduce the event

$$R_k := \bigcap_{j=1}^{\infty} \{w_1(k) \geq \ell, \bar{w}(k; j) \geq 1 - \ell\}.$$

It is clear from its definition that  $(R_k, k \in \mathbb{Z})$  is a stationary sequence of events with

$$\mathbb{P}(R_k) = \mathbb{P}(R_0) = F([\ell, 1]) \prod_{j=1}^{\infty} (1 - (1-p)^j) > 0.$$

Consider the stationary random set  $J := \{k \in \mathbb{Z} : R_k \text{ holds}\}$ . Since  $\mathbb{P}(R_k) > 0$ , we have, by ergodicity (more specifically by the Poincaré recurrence theorem),  $\inf J = -\infty$  and  $\sup J = \infty$  a.s. We enumerate the elements of  $J$  by

$$\dots < T_{-1} < T_0 \leq 0 < T_1 < T_2 < \dots$$

We define

$$\tilde{\nu}(t) := \sigma \sum_{i \in \mathbb{Z}} \mathbb{1}_{\{T_i < t \leq T_{i+1}\}} \Psi_w^{T_i, t} \delta_0, \quad t \in \mathbb{Z}.$$

It is clear from its definition that  $(\tilde{\nu}(t), t \in \mathbb{Z})$  is stationary, as  $(w(t), t \in \mathbb{Z})$  is a stationary sequence, and  $(\tilde{\nu}(T_i + 1), i \in \mathbb{Z})$  are i.i.d. elements of  $\mathcal{N}_0$ . Next, we define

$$\bar{\mathbf{m}}(t) = \mathbf{m}(\tilde{\nu}(t-1), w(t)), \quad t \in \mathbb{Z},$$

which is again a stationary sequence.

By Corollary 1, we observe that for all  $t \geq T_1 + 1$ , the quantity  $\mathbf{m}(\sigma\nu(t-1), w(t))$  does not algebraically depend on  $\nu(T_1 - 1)$ . Hence, we have  $\mathbf{m}(\sigma\nu(t-1), w(t)) = \bar{\mathbf{m}}(t)$ , using that  $\bar{\mathbf{m}}(t)$  is the same quantity for the MGS started from  $\delta_0$  at time  $T_1 - 1$ . As  $T_1 < \infty$  a.s. this completes the proof of the first part of the theorem.

Next, using that

$$C(F) = \lim_{n \rightarrow \infty} \frac{\mathbb{E}(M_n)}{n} = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \mathbb{E}(\mathbf{m}(\sigma\nu(j-1), w(j))^+) \quad \text{a.s.},$$

and using the eventual equality between  $\mathbf{m}(\sigma\nu(t-1), w(t))$  and  $\bar{\mathbf{m}}(t)$ , we have

$$C(F) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \mathbb{E}(\bar{\mathbf{m}}(j)^+) = \mathbb{E}(\bar{\mathbf{m}}(0)^+),$$

by stationarity and ergodicity of the sequence. □

## 4 Perfect simulation

The formula (2) for  $C(F)$  suggests a straightforward method for estimating  $C(F)$ : starting from  $\nu(0) = 0$ , generate iteratively  $\nu(1), \nu(2), \dots, \nu(n)$ , and take  $\nu_1(n)$  for an estimation of  $C(F)$ . This standard (so-called MCMC) method introduces a bias. Indeed,  $\mathbb{E}(\nu_1(n))/n$  is not equal to  $C(F)$ , but merely converges to that constant.

To eliminate this bias, we produce an algorithm that constructs the variable  $\bar{\mathbf{m}}(0)$ , whose distribution is unknown. Then, by standard Monte Carlo method, an unbiased estimation of  $C(F)$  can be constructed. This is done in this case by using the construction described in the proof of Theorem 1.

This algorithm is a development of a similar construction for functionals of stochastic recursions in [8] and is based on the ideas of so-called “backward coupling”, see [12]. It is close in spirit to the coupling-from-the-past method for Markov chains [22] and to the perfect simulation construction for processes with “long memory” [5]. Note that the algorithm from [22] is applicable to either finite Markov chains or ordered monotone Markov chains possessing a unique minimal state and a unique maximal state, so it cannot be applied to our case.

**Theorem 2** (Perfect simulation). *Define*

$$T^* := \sup\{t \leq -1 : w_1(t) \geq \ell, \min_{1 \leq j \leq |t|} \bar{w}(t; j) \geq 1 - \ell\}.$$

Then  $|T^*| < \infty$  a.s., and

$$\bar{\mathbf{m}}(0) = \mathbf{m}\left(\sigma\Psi^{T^*, -1}\delta_0; w(0)\right) \text{ a.s.}$$

*Proof.* We recall that  $(T_{-j}, j \in \mathbb{N})$  are the negative elements of the random set  $J$ , with  $T_{-1} > -\infty$ . We remark that

$$w_1(T_{-1}) \geq \ell, \quad \bar{w}(T_{-1}; j) \geq 1 - \ell \text{ for all } j > 0,$$

therefore  $T_* \geq T_{-1}$ , proving its finiteness.

Moreover, since

$$w_1(T^*) \geq \ell, \quad \min_{1 \leq j \leq |t|} \bar{w}(T^*; j) \geq 1 - \ell,$$

by Corollary 1, the quantity  $\mathbf{m}\left(\sigma\Psi^{T^*, -1}\nu; w(0)\right)$  does not algebraically depend on the value of  $\nu \in \mathcal{N}_0$ . As a result, it is equal to  $\bar{\mathbf{m}}(0)$ , defined as  $\mathbf{m}\left(\sigma\Psi^{T^*, -1}\delta_0; w(0)\right)$ .  $\square$

**Remark 1.** *If the essential supremum  $L$  of  $F$  is infinite, then the perfect simulation algorithm we defined cannot apply. Indeed, in this situation, Lemma 1 does not apply and we could not find an event depending on a finite number of charges such that an analogue of this lemma would hold.*

```

Fix  $t = 0$  and  $J = 1$ ;
Generate the variable  $w_1(0)$ ;
Fix Stopping = False;
while  $Stopping = False$  do
    while  $\max_{1 \leq j \leq J} w_j(t) < 1 - \ell$  do
        Increase  $J$  by 1;
        Generate the variable  $w_J(t)$ ;
    while  $J > 1$  do
        Decrease  $J$  by 1 and  $t$  by 1;
        Generate  $w_1(t), \dots, w_J(t)$ ;
        while  $\max_{1 \leq j \leq J} w_j(t) < 1 - \ell$  do
            Increase  $J$  by 1;
            Generate the variable  $w_J(t)$ ;
        Decrease  $t$  by 1;
        Generate  $w_1(t)$ ;
        Fix Stopping =  $\{w_1(t) \geq \ell\}$ ;
Fix  $\nu = \delta_0$ ;
for  $s$  from  $t + 1$  to  $-1$  do
    Generate the variables  $w_1(s), \dots, w_{\|\nu\|}(s)$ ;
    Set  $\mathbf{m} = \max\{\nu_j + w_j(s) \text{ for } 1 \leq j \leq \|\nu\|\}$ ;
    Add  $\delta_{\mathbf{m}}$  to  $\nu$ ;
Set  $\mathbf{m} = \max\{\nu_j + w_j(0) \text{ for } 1 \leq j \leq \|\nu\|\}$ ;
Return:  $\mathbf{m} - \nu_1$ ;

```

**Algorithm 1:** Construction of a variable of law  $\bar{\mathbf{m}}(0)$ .

## The perfect simulation algorithm

We now describe more precisely the perfect simulation algorithm. Let  $F$  be a probability distribution satisfying (3), we fix  $\ell \in [0, 1)$  such that  $F([1 - \ell, 1]) \in (0, 1)$ . The algorithm requires the construction of an array of i.i.d. random variables with common distribution  $F$  until the random event  $T^*$  can be constructed.

To construct  $T^*$  as well as  $\bar{\mathbf{m}}(0)$  from the sequence  $\{w_j(t), j \in \mathbb{N}, t \in \mathbb{Z}\}$ , one only needs to consider a.s. finitely many elements of this set, as  $\{T^* = t\}$  is a measurable function of

$$\{w_1(t)\} \cup \{w_j(t+k), 1 \leq j \leq k \leq |t|\}$$

and  $\bar{\mathbf{m}}(0)$  is a measurable function of

$$\{w_1(T^*)\} \cup \{w_j(T^* + k), 1 \leq j \leq k \leq |T^*|\}.$$

Therefore, we can explore triangular arrays of the form

$$\{w_1(t)\} \cup \{w_j(t+k), 1 \leq j \leq k \leq |t|\},$$

progressively decreasing  $t$  until time  $T^*$  is detected. Once this random variable is known, we construct the random variable  $\bar{\mathbf{m}}(0)$  using the procedure described in Theorem 1 from the

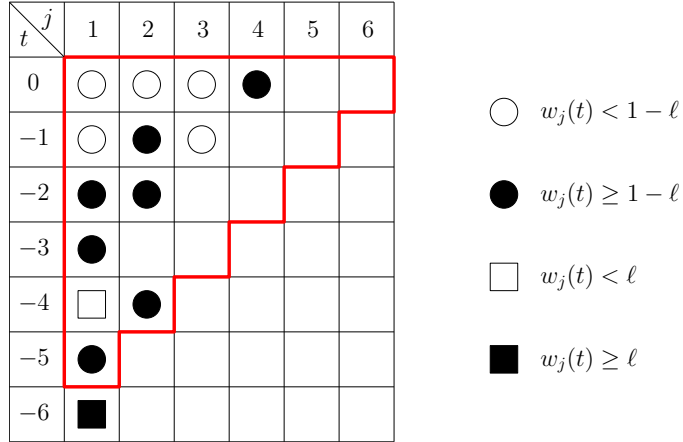


Figure 1: *Illustration of the execution of Algorithm 1 on an example, in the case where  $\ell < 1 - \ell$ . The variables sampled until the Boolean variable *Stopping* becomes *True* are pictured by black/white squares and disks. One searches for the first time  $T^*$  such that every line of index  $T^* + 1 \leq t \leq 0$  has at least one black disk between columns 1 and  $t - T^*$  and such that there is a black square in position  $(T^*, 1)$ . The full triangular array of variables used in the construction of  $\nu$  is enclosed by a red boundary.*

previously discovered random variables. A possible implementation is described in Algorithm 1. We show a graphical representation of a run of Algorithm 1 in Figure 1.

We observe that this algorithm has a complexity of  $(T^*)^2$ , as it is the number of steps needed to generate the variable  $\bar{m}(0)$ . It is worth noting that  $-T^*$  can be constructed as the first hitting time of 0 of the Markov chain  $(X_n)$  with initial state

$$X_0 = \min\{j \geq 1, w_j(0) \geq 1 - \ell\}$$

and with transition probabilities defined for all  $j \geq 2$  and  $i \geq j$  by

$$P(j, j - 1) = 1 - (1 - p)^{j-1} \text{ and } P(j, i) = p(1 - p)^{i-1}$$

where  $p = \mathbb{P}(w_1(0) \geq 1 - \ell)$ , with

$$P(1, 0) = \mathbb{P}(w_1(0) \geq \ell), \quad \mathbb{P}(1, 1) = \mathbb{P}(1 - \ell \leq w_1(0) < \ell), \quad P(1, j) = p(1 - p)^{j-1} \text{ for } j \geq 2.$$

The quantity  $X_n$  corresponds to the value of the variable  $J$  at the end of the period when  $t = -n$  in Algorithm 1. In the example shown in Figure 1, we have

$$(X_0, X_{-1}, X_{-2}, X_{-3}, X_{-4}, X_{-5}, X_{-6}) = (4, 3, 2, 1, 2, 1, 0).$$

Note that  $T^*$  has exponential tails.

The choice of the parameter  $\ell$  may have an important effect on the behaviour of the average complexity  $\mathbb{E}((T^*)^2)$  of the algorithm. We plotted  $\ell \mapsto \mathbb{E}((T^*)^2)$  in Figure 2, when the charge distribution is given by  $F(dx) = \mathbb{1}_{\{x \leq 1\}} e^{x-1} dx$ . Additionally, as  $p \rightarrow 0$ , the quantity  $\mathbb{E}((T^*)^2)$  grows to  $\infty$ . We estimated  $\mathbb{E}((T^*)^2)$  for  $F = p\delta_1 + (1 - p)\delta_{-\infty}$  and plotted this quantity as a function of  $p$  in Figure 3.

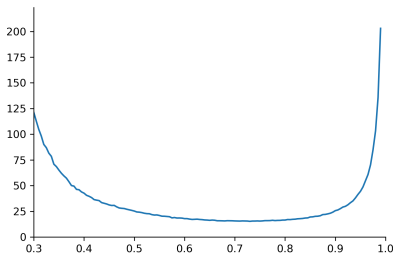


Figure 2: *Dependency in the parameter  $\ell$  of the complexity of Algorithm 1 with a charge distribution  $F(dx) = \mathbb{1}_{\{x \leq 1\}} e^{x-1} dx$ . The figure was obtained with a Monte Carlo simulation of  $N = 10^4$  copies of  $T^*$  for 100 different values of  $\ell$ . For this charge distribution, the Monte Carlo simulations give  $C(F) = 0.4432 \pm 0.0006$ .*

We observe in Figure 2 that different choices of the value  $\ell$  can have a dramatic impact on the efficiency of Algorithm 1. Choosing a value  $\ell$  too small has the effect of making the first appearance of a triangular event too late. On the other hand, if  $\ell$  is too big then with high probability, one will have  $w_1(T) \leq \ell$ , and thus the first “successful” triangular event will appear much later. For the distribution  $F$  we chose, it appears that an optimal choice of  $\ell$  seems to be around  $\ell = 0.7$ , which balances between these two extremes.

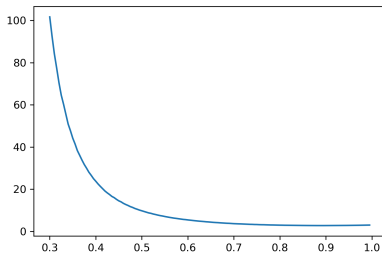


Figure 3: *Dependency in the parameter  $p$  of the complexity of Algorithm 1 applied to the detection of the longest path in the Barak-Erdős graph with parameter  $p$ . Figure obtained through Monte Carlo simulation of  $N = 10^5$  copies of  $T^*$  for 120 different values of  $p$ .*

We observe in Figure 3 that if  $F$  puts a large mass on the negative half-line, the complexity of Algorithm 1 can become quite large. The function  $p \mapsto \mathbb{E}(T^*)^2$  grows at least exponentially in  $1/p$  as  $p \rightarrow 0$  in the Barak-Erdős graph, but we were not able to obtain a good estimate of this rate of increase.

## 5 Further directions of research

In this article we considered last passage percolation on the directed complete graph, which has a total order on its vertex set. One extension of this would be to construct a perfect simulation algorithm for so-called directed slab graphs [6] where the set of vertices is only partially ordered. Another possible extension would be to add i.i.d. vertex weights with a

distribution that has a finite essential supremum.

As discussed in the previous section, the complexity of our perfect simulation algorithm may dramatically vary with  $\ell$ . In the case of  $F(dx) = \mathbb{1}_{\{x \leq 1\}} e^{x-1} dx$  presented in Figure 2, there seems to be a unique optimal choice for  $\hat{\ell}$  around 0.7. It would be interesting to find some classes of distributions  $F$  for which one has good bounds on the optimal value  $\hat{\ell}$ .

Yet another research direction would be the estimation of the constants appearing for last passage percolation on a 2-dimensional version of the Barak-Erdős directed graph on the set  $\mathbb{N} \times \mathbb{N}$  and whose edges are as follows: if  $u = (u_1, u_2), v = (v_1, v_2) \in \mathbb{N} \times \mathbb{N}$  are two vertices such that  $i_1 \leq j_1, i_2 \leq j_2$ , then declare the pair  $(u, v)$  as an edge directed from  $u$  to  $v$  with probability  $p$ , independently over all such pairs. Then maximum length  $L_n$  of all paths from  $(1, 1)$  to  $(n, n^a)$ , for a certain  $a > 0$ , rescaled appropriately, converges weakly [17] to a random variable having a Tracy-Widom distribution depending on two parameters. The estimation of these parameters is an open problem.

## Acknowledgements

SF was partially supported by the RFBR collaborative grant 19-51-15001 and TK, BM and SR were partially supported by the CNRS PRC collaborative grant CNRS-193-382 with the common title “Asymptotic and analytic properties of stochastic ordered graphs and infinite bin models”.

## References

- [1] D. Aldous and J. Pitman. The asymptotic speed and shape of a particle system. In: *Probability, statistics and analysis, Lond. Math. Soc. Lect. Notes Ser.* **79**, 1-23, 1983.
- [2] A.B. Barak and P. Erdős. On the maximal number of strongly independent vertices in a random acyclic directed graph. *SIAM J. Algebr. Discr. Methods* **5**, 508–514, 1984.
- [3] A.A. Borovkov and S.G. Foss. Stochastically recursive sequences and their generalizations. *Sib. Adv. Math.*, 2(1):16–81, 1992.
- [4] K. Chernysh and S. Ramassamy. Coupling any number of balls in the infinite-bin model. *J. Appl. Probab.* **54** no. 2, 540–549, 2017.
- [5] F. Comets, R. Fernández and P.A. Ferrari. Processes with long memory: regenerative construction and perfect simulation. *Ann. Appl. Probab.* **12**, 921–943, 2002.
- [6] D. Denisov, S. Foss and T. Konstantopoulos. Limit theorems for a random directed slab graph. *Ann. Appl. Probab.*, **22**, 702–733, 2012.
- [7] K. Dutta. On limit constants in last passage percolation in transitive tournaments. [arXiv:2005.09922](https://arxiv.org/abs/2005.09922) [math.PR]

- [8] S. Foss and T. Konstantopoulos. Extended renovation theory and limit theorems for stochastic ordered graphs. *Markov Process. Related Fields* **9**, no. 3, 413–468, 2003.
- [9] S. Foss and T. Konstantopoulos. Limiting properties of random graph models with vertex and edge weights. *J. Stat. Phys.* **173**, 626–643, 2018.
- [10] S. Foss, J. Martin and P. Schmitt. Long-range last-passage percolation on the line. *Ann. App. Prob.* **24**, no. 1, 198–234, 2014.
- [11] S. Foss, T. Konstantopoulos and A. Pyatkin. Probabilistic and analytical properties of the last passage percolation constant in a weighted random directed graph. [arXiv:2006.01727v3](https://arxiv.org/abs/2006.01727v3) [[math.PR](#)]
- [12] S. Foss and R.L. Tweedie. Perfect simulation and backward coupling. *Stoch. Models* **14**, no. 1–2, 187–203, 1998.
- [13] E. Gelenbe, R. Nelson, T. Philips and A. Tantawi. An approximation of the processing time for a random graph model of parallel computation. *Proceedings of 1986 ACM Fall Joint Computer Conference*, Los Alamitos, CA, 691–697, 1986.
- [14] M. Isopi and C.M. Newman. Speed of parallel processing for random task graphs. *Comm. Pure Appl. Math.* **47**, 261–276, 1994.
- [15] J. F. C. Kingman. Subadditive ergodic theory. *Ann. Prob.*, 1:883–909, 1973.
- [16] W.S. Kendall. Notes on perfect simulation. In: Markov chain Monte Carlo, *Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap.* **7**, 93–146, 2005.
- [17] 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**, no. 2, 711–730, 2013.
- [18] B. Mallein and S. Ramassamy. Barak-Erdős graphs and the infinite-bin model. *Ann. Inst. Henri Poincaré Probab. Stat.*, to appear. [arXiv:1610.04043](https://arxiv.org/abs/1610.04043) [[math.PR](#)].
- [19] B. Mallein and S. Ramassamy. Two-sided infinite-bin models and analyticity for Barak-Erdős graphs. *Bernoulli* **25**, no. 4B, 3479–3495, 2019.
- [20] C.M. Newman. Chain lengths in certain random directed graphs. *Random Structures Algorithms* **3**, no. 3, 243–253, 1992.
- [21] C.M. Newman and J.E. Cohen. A stochastic theory of community food webs IV: theory of food chains in large webs. *Proc. Roy. Soc. Lond. B: Biol. Sci.* **228**, 355–377, 1986.
- [22] J.G. Propp and D.B. Wilson. Exact sampling with coupled Markov chains and applications to statistical mechanics. *Random Structures Algorithms* **9**, 223–252, 1996.

Sergey Foss,  
SOBOLEV INSTITUTE OF MATHEMATICS, SIBERIAN BRANCH OF THE RUSSIAN ACADEMY OF SCIENCES, NOVOSIBIRSK, RUSSIA  
NOVOSIBIRSK STATE UNIVERSITY, NOVOSIBIRSK, RUSSIA  
SCHOOL OF MATHEMATICAL AND COMPUTER SCIENCES, HERIOT-WATT UNIVERSITY, EDINBURGH, UK  
*E-mail address:* `sergueiorfoss25@gmail.com`

Takis Konstantopoulos,  
DEPARTMENT OF MATHEMATICAL SCIENCES, UNIVERSITY OF LIVERPOOL, LIVERPOOL, UK  
*E-mail address:* `takiskonst@gmail.com`

Bastien Mallein,  
UNIVERSITÉ SORBONNE PARIS NORD, LAGA, UMR 7539, F- 93430, VILLETANEUSE, FRANCE  
ÉCOLE NORMALE SUPÉRIEURE, DMA, UMR 8553, F- 75005, PARIS, FRANCE  
*E-mail address:* `mallein@math.univ-paris13.fr`

Sanjay Ramassamy,  
UNIVERSITÉ PARIS-SACLAY, CNRS, CEA, INSTITUT DE PHYSIQUE THÉORIQUE, 91191 GIF-SUR-YVETTE, FRANCE  
*E-mail address:* `sanjay.ramassamy@ipht.fr`