



**HAL**  
open science

# Growth properties of the infinite-parent spatial Lambda-Fleming-Viot process

Apolline Louvet, Amandine Véber

► **To cite this version:**

Apolline Louvet, Amandine Véber. Growth properties of the infinite-parent spatial Lambda-Fleming-Viot process. 2022. hal-03661964

**HAL Id: hal-03661964**

**<https://hal.science/hal-03661964>**

Preprint submitted on 8 May 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Growth properties of the $\infty$ -parent spatial Lambda-Fleming Viot process

Apolline Louvet<sup>1,2</sup> and Amandine Véber<sup>1</sup>

<sup>1</sup>MAP5, Université Paris Cité, CNRS, 45 rue des Saints-Pères, 75006 Paris, France

<sup>2</sup>CMAP, Ecole Polytechnique, CNRS, Institut Polytechnique de Paris, Route de Saclay, 91128 Palaiseau Cedex, France

May 8, 2022

## Abstract

The  $\infty$ -parent spatial Lambda-Fleming Viot process, or  $\infty$ -parent SLFV, is a model for spatially expanding populations in which empty areas are filled with "ghost" individuals. The interest of this process lies in the fact that it is akin to a continuous-space version of the classical Eden growth model, while being associated to a dual process encoding ancestry and allowing one to study the evolution of the genetic diversity in such a population.

In this article, we focus on the growth properties of the  $\infty$ -parent SLFV in two dimensions. To do so, we first define the quantity that we shall use to quantify the speed of growth of the area covered with the subpopulation of real individuals. Using the associated dual process and a comparison with a first-passage percolation problem, we show that the growth of the "occupied" region in the  $\infty$ -parent SLFV is linear in time. We use numerical simulations to approximate the growth speed, and conjecture that the actual speed is higher than the speed expected from simple first-moment calculations due to the characteristic front dynamics.

We then study a toy model of two interacting growing piles of cubes in order to understand how the growth dynamics at the front edge can increase the global growth speed of the "occupied" region. We obtain an explicit formula for this speed of growth in our toy model, using the invariant distribution of a discretised version of the model. This study is of interest on its own right, and its implications are not restricted to the case of the  $\infty$ -parent SLFV.

**Running headline:** Growth properties of the  $\infty$ -parent SLFV

**Keywords:** spatial Lambda-Fleming Viot process, range expansion, duality, percolation, stochastic growth process, sub-additivity, speed of growth

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>The <math>\infty</math>-parent spatial Lambda-Fleming Viot process with elliptical reproduction events</b>	<b>6</b>
2.1	Definition of the process . . . . .	7
2.2	Dual process and duality relation . . . . .	11
2.3	Speed of growth of the occupied region in the $\infty$ -parent SLFV: Definition and main result	12
2.4	Numerical simulations . . . . .	23
<b>3</b>	<b>Lower bound on the speed of growth</b>	<b>24</b>
3.1	Sub-additivity . . . . .	25
3.2	Definition of the express chain . . . . .	27
3.3	Properties of the express chain . . . . .	29
<b>4</b>	<b>Upper bound on the speed of growth</b>	<b>32</b>
4.1	A first-passage percolation problem . . . . .	33
4.2	Discretisation of the $\infty$ -parent ancestral process . . . . .	35
4.3	Comparison to the first-passage percolation problem . . . . .	37
<b>5</b>	<b>The two column growth process</b>	<b>38</b>
5.1	The two column growth process: Definition and properties . . . . .	39
5.2	The discretised two columns growth process . . . . .	45
5.3	Invariant distribution of the discretised 2-CGP . . . . .	52
	<b>References</b>	<b>57</b>
<b>A</b>	<b>Geometrical properties of ellipses</b>	<b>60</b>

## 1 Introduction

The  $\infty$ -parent spatial  $\Lambda$ -Fleming Viot process, or  $\infty$ -parent SLFV, was introduced in [32] as a model for expanding populations in  $\mathbb{R}^2$ . Its main feature is the use of "ghost" individuals (thereafter referred

to as "type 0" individuals) to fill empty areas, adapting ideas from [13, 23]. In [13, 23, 32], ghost individuals can reproduce as well, modelling stochastic fluctuations in population sizes, but with a selective disadvantage against real individuals (thereafter referred to as "type 1" individuals), ensuring that population expansions can indeed occur. In [32], it is shown that the  $\infty$ -parent SLFV corresponds to the limit of an SLFV with strong selection when the selective advantage of type 1 individuals over type 0 individuals becomes infinitely strong. Therefore, in the limiting regime, we no longer observe local extinctions due to the stochasticity in reproduction at the front, where densities of type 1 individuals are lower.

One of the main interests of the  $\infty$ -parent SLFV lies in its dual interpretation, both as a stochastic growth model akin to models from percolation theory, and as a population genetics process with a spatial structure. Indeed, if we consider the area occupied by type 1 individuals, then its growth dynamics corresponds to a continuous space version of the Eden model [14]. Moreover, and contrary to the Eden model, the  $\infty$ -parent SLFV is equipped with tools allowing one to investigate genetic diversity patterns, assuming that real individuals are further subdivided into different types (for instance using the concept of "tracers", see [13, 23]). Therefore, it is potentially suited to study how the characteristic genetic diversity patterns observed in real expanding populations [21, 22, 24] arise. In this paper, we shall only focus on the region covered by real individuals, without specifying different subtypes of real (or type 1) individuals. However, understanding the family structure in this model is a first step towards the understanding of how genetic diversity evolves in a multitype population.

The  $\infty$ -parent SLFV process belongs to the family of *spatial  $\Lambda$ -Fleming Viot processes*, which are characterised by a reproduction dynamics driven by an exogeneous Poisson point process of *reproduction events* rather than by individual reproduction. Informally, the  $\infty$ -parent SLFV is constructed as follows. We consider that at any time  $t$ , the area occupied by type 1 individuals is represented by a measurable set  $\mathfrak{E}_t \subseteq \mathbb{R}^d$ . Reproduction events are given by a Poisson point process  $\Pi$  on  $\mathbb{R}_+ \times \mathbb{R}^2$  with intensity measure  $\alpha dt \otimes dx$ ,  $\alpha > 0$ , which encodes the sequence of times and centres of the reproduction events. For the moment, to ease the exposition, we consider that all reproduction events have the same shape: a ball with fixed radius  $\mathcal{R} > 0$ . In Section 2, we shall consider the more general case of ellipses with random bounded parameters. Then, for all  $(t, z) \in \Pi$ , if  $\mathcal{B}_{\mathcal{R}}(z)$  is the ball with centre  $z$  and radius  $\mathcal{R}$  and if  $\text{Vol}(\mathfrak{E}_{t-} \cap \mathcal{B}_{\mathcal{R}}(z)) \neq 0$ , we set  $\mathfrak{E}_t = \mathfrak{E}_{t-} \cup \mathcal{B}_{\mathcal{R}}(z)$ , and we do nothing otherwise. In other words, whenever a reproduction event occurs, if the corresponding area contains some real individuals, then one of them reproduces and completely fills the area with its descendants. When we deal with genealogies, we shall assume that the reproducing individual is chosen uniformly at random among the real individuals occupying the ball just before the event (though in fact we shall not need

to specify an ancestor in our approach below).

Formulated this way, the  $\infty$ -parent SLFV is very similar to models from continuous first-passage percolation, and in particular to the model introduced by Deijfen in [10], in which reproduction occurs when  $z \in \mathfrak{E}_{t-}$ . Contrary to the  $\infty$ -parent SLFV, this process does not jump when  $z \notin \mathfrak{E}_{t-}$  but  $\text{Vol}(\mathfrak{E}_{t-} \cap \mathcal{B}_{\mathcal{R}}(z)) > 0$ , and so we expect its growth speed to be different from that of the  $\infty$ -parent SLFV (see [10, 11]). Moreover, and perhaps more importantly, the  $\infty$ -parent SLFV is defined for more general initial conditions, which are not restricted to  $\text{Vol}(\mathfrak{E}_0) < +\infty$ . In particular, this means that the definition we just provided is only informal: when  $\text{Vol}(\mathfrak{E}_0) = +\infty$ , an infinite number of reproduction events intersect the occupied area in any time interval, and the jump rate of the process is infinite. Rigorously, the  $\infty$ -parent SLFV is a measure-valued process, which can be defined as the limit of a sequence of SLFVs with strong selection, or (under stricter conditions which will be satisfied in this article) as the unique solution to a well-posed martingale problem [32]. If  $\text{Vol}(\mathfrak{E}_0) < +\infty$ , it can also be defined as a process taking its values in the class of measurable subsets of  $\mathbb{R}^d$ . In what follows, we shall take  $\mathfrak{E}_0$  such that  $\text{Vol}(\mathfrak{E}_0) = +\infty$  and we shall use the characterisation of the process as the solution to a well-posed martingale problem (and in fact the dual process offered by this approach will be the key ingredient of most of our proofs).

While most of the studies on SLFVs with selection were carried out in a weak selection regime (see, *e.g.*, [16, 17, 18] for selection against one type, [20] for general forms of selection, [6, 8, 29] for fluctuating selection and [15] for selection against heterozygosity), the  $\infty$ -parent SLFV corresponds to a strong selection limit, motivated by the interpretation of ghost individuals as modelling empty areas. In particular, and contrary to previous works on SLFVs with selection, the results we obtain do not require to rescale time, nor space. Moreover, in most SLFVs (as well as in the original  $\infty$ -parent SLFV, see [32]), the affected areas are balls, whose radii can be fixed or random. However, biological experiments suggest that the shape of the areas impacted by reproduction events influences the observed genetic diversity patterns [22]. Moreover, this excludes the case of a preferential expansion direction, for example towards a resource by a phenomenon similar to chemotaxis. Therefore, in this article, we consider a variant of the  $\infty$ -parent SLFV in which the area affected by a reproduction event can be shaped like an ellipse rather than a ball.

In this article, we are interested in the growth properties of the occupied area in the  $\infty$ -parent SLFV and we focus on the following questions: Is the growth linear in time? To what extent is the speed of growth affected by the reproduction dynamics at the front edge? How is it affected by the shape of the reproduction events? In particular, we would like to compare the growth properties of the occupied region in the  $\infty$ -parent SLFV to those of first-passage percolation models, whose most

well-known representative is probably the Eden growth model. In order to do so, we first give a brief overview of what is known of their growth properties. First-passage percolation processes are stochastic growth models originally defined on a lattice (generally  $\mathbb{Z}^d$ ,  $d \geq 2$ ). In these models, each vertex is either occupied, or empty. If vertex  $x \in \mathbb{Z}^d$  becomes occupied at time  $t$ , and if it is connected to a vertex  $y \in \mathbb{Z}^d$  by an edge (denoted  $e$ ), then vertex  $y$  becomes occupied at time  $t + \tau_e$ , where  $\tau_e$  is independent from one edge to another (but not necessarily identically distributed). Whether the expansion is linear in time depends on the distribution of the time needed to pass through any given edge of the grid (see, *e.g.*, [4]) and whether one considers short-range percolation, such as nearest neighbour percolation, or long-range percolation, in which two vertices  $x, y \in \mathbb{Z}^d$  are connected by an edge no matter the distance between them [7, 9, 34]. In particular, when edge passing times are distributed as in the Eden growth model, growth is linear in time for short-range percolation and potentially faster for long-range percolation, depending on the relation between the distribution of  $\tau_e$  and the distance between the two vertices it connects [4, 7]. See, *e.g.*, [10, 12, 26] for examples of extensions of first-passage percolation models to a continuous setting.

When the growth is linear in time, in general it is possible to obtain lower and upper bounds on the speed of growth (see, *e.g.*, [2, 38]), or to use simulations to approximate it [1]. For other stochastic growth models, such as the corner growth model [36], which belongs to the family of *last-passage percolation models*, it is possible to obtain an explicit speed of growth for specific passage time distributions [35]. As many other growth models, the Eden model is conjectured to belong to the universality class of the Kardar-Parisi-Zhang (KPZ) equation [28]. This equation generates rough fronts, whose characteristics are similar to the ones of fronts observed in some expanding biological populations (see, *e.g.*, [27]). Such a conjecture is notably difficult to establish; to our knowledge, it has only been demonstrated in the case of the solid-on-solid growth model in [5].

The main result of this article is Theorem 2.11 in Section 2.3, and states that when the area affected by a reproduction event has bounded shape parameters, the growth of the area occupied by type 1 individuals is linear in time. In particular, this means that compared to the model of Deijfen [10], the increased reproduction rate at the front and the unbounded initial condition do not lead to superlinear growth. To prove this result, we first use sub-additivity arguments to show that the growth is *at least* linear in time. Next, we use a comparison with a short-range percolation model to show that the growth is also *at most* linear in time, which allows us to conclude. Our proof also yields an explicit lower bound on the speed of growth.

In Section 2.4, we use numerical simulations to approximate the growth speed of the area occupied by type 1 individuals, and we compare it to the lower bound we obtained. This bound was initially

conjectured to be equal to (or at least a good approximation for) the growth speed, but the simulations show that the actual speed is significantly higher. Moreover, they suggest that the growth is driven by "spikes" which occur at the front and then thicken in all directions.

In order to understand how the spike phenomenon can make the front of an expanding population advance faster, we consider a simple toy model in Section 5, composed of two interacting growing piles of cubes. Its main interest is that it is possible to obtain an explicit expression for its speed of growth. In this simplified model, new cubes can also grow sideways, at a rate which is proportional to the height difference between the two piles. We show that the interaction between the two piles leads to an increase in the speed of growth of each pile by a factor of  $\sim 1.46$  compared to the rate at which each pile grows by itself independently of the other pile, which is in line with numerical simulations. The proof relies on a coupling with a discretised version of the process, for which it is possible to compute the invariant distribution.

More precisely, the paper is structured as follows. In Section 2, we define the variant of the  $\infty$ -parent SLFV with elliptical reproduction events rigorously, introduce its dual, and define what we mean by the "speed of growth" of the occupied region in the process. We also state the main result of the paper, Theorem 2.11, and analyse numerical simulations of the  $\infty$ -parent SLFV in Section 2.4. The proof of Theorem 2.11 spans two sections. In Section 3, we use the dual process to show that the growth is *at least* linear in time, and provide a lower bound on the speed of growth of the process. In Section 4, we use a comparison with a first-passage percolation process (also based on the dual process) to show that the growth is *at most* linear in time. Combining the results from both sections yields Theorem 2.11. In Section 5, we study the toy model of interacting growing piles of cubes, and obtain an explicit expression for its speed of growth. This result is of interest in its own right, but also explains the discrepancy between the numerical approximation for the speed of growth of the occupied area in the  $\infty$ -parent SLFV and the value initially conjectured.

## 2 The $\infty$ -parent spatial Lambda-Fleming Viot process with elliptical reproduction events

In this section, we rigorously define the  $\infty$ -parent SLFV, in the version we use in this article. Indeed, the process was originally defined with reproduction events occurring in balls, while here we consider that they rather occur in ellipses, as it is straightforward to generalise the construction in [32] to our case. Then, we introduce the dual process of potential ancestors associated to the  $\infty$ -parent SLFV. We conclude by formalising what we mean by the speed of growth of the occupied area, and by explaining

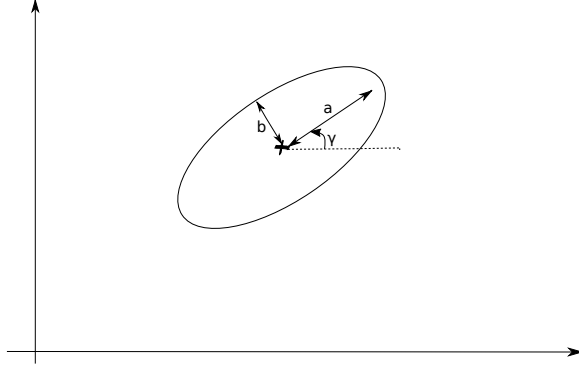


Figure 1: Ellipse with parameters  $(a, b, \gamma)$ .

how to study it using the dual process.

## 2.1 Definition of the process

All the random objects we consider in this section are defined over some probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ . We use the notation  $\mathbb{N} = \{0, 1, 2, \dots\}$ . In all that follows, let  $\tilde{\mu}$  be a finite measure on  $(0, +\infty)^2 \times (-\pi/2, \pi/2)$  such that there exists  $\mathcal{R} > 0$  satisfying

$$\tilde{\mu}\left(\left((0, \mathcal{R}] \times (0, \mathcal{R}]\right)^c \times (-\pi/2, \pi/2)\right) = 0.$$

Let  $\mathcal{R}_{\tilde{\mu}}$  be the smallest  $\mathcal{R} > 0$  such that this condition is satisfied. We also set

$$S_{\tilde{\mu}} = (0, \mathcal{R}_{\tilde{\mu}}] \times (0, \mathcal{R}_{\tilde{\mu}}] \times (-\pi/2, \pi/2).$$

**Ellipses.** We first set the notation regarding ellipses.

**Definition 2.1.** Let  $z_c = (x_c, y_c) \in \mathbb{R}^2$ ,  $(a, b) \in (0, +\infty)^2$  and  $\gamma \in (-\pi/2, \pi/2)$ . The ellipse with centre  $z_c$  and parameters  $(a, b, \gamma)$ , denoted by  $\mathfrak{B}_{a,b,\gamma}(z_c)$ , is defined by:

$$\mathfrak{B}_{a,b,\gamma}(z_c) = \left\{ \begin{pmatrix} x_c \\ y_c \end{pmatrix} + A_\gamma \begin{pmatrix} ar \cos(\theta) \\ br \sin(\theta) \end{pmatrix} : r \in [0, 1], \theta \in [0, 2\pi) \right\}$$

where

$$A_\gamma = \begin{pmatrix} \cos(\gamma) & -\sin(\gamma) \\ \sin(\gamma) & \cos(\gamma) \end{pmatrix}.$$

See Figure 1 for an illustration. We denote the volume of an ellipse with parameters  $(a, b, \gamma)$  by  $V_{a,b,\gamma} := \text{Vol}(\mathfrak{B}_{a,b,\gamma}(0))$ .



For all  $(a, b) \in (0, +\infty)^2$  and  $\gamma \in (-\pi/2, \pi/2)$ , if  $f$  is an element of the space  $C_c(\mathbb{R}^2)$  of all continuous and compactly supported functions  $\mathbb{R}^2 \rightarrow \mathbb{R}$ , let

$$\text{Supp}^{a,b,\gamma}(f) := \left\{ z \in \mathbb{R}^2 : \text{Vol}(\mathfrak{B}_{a,b,\gamma}(z) \cap \text{Supp}(f)) \neq 0 \right\},$$

where  $\text{Supp}(f)$  stands for the support of  $f$ . The set  $\text{Supp}^{a,b,\gamma}(f)$  can be interpreted as the set of all potential centres  $z \in \mathbb{R}^2$  for ellipses with parameters  $(a, b, \gamma)$  overlapping the support of  $f$ . For all  $z \in \mathbb{R}^2$  and  $\omega : \mathbb{R}^2 \rightarrow \{0, 1\}$  Lebesgue-measurable, let  $\Theta_z^{a,b,\gamma}(\omega) : \mathbb{R}^2 \rightarrow \{0, 1\}$  be the function defined by

$$\Theta_z^{a,b,\gamma}(\omega) := \mathbb{1}_{\{\mathfrak{B}_{a,b,\gamma}(z)^c\}} \times \omega.$$

If  $\omega$  represents the density of type 0 (or ghost) individuals, then  $\Theta_z^{a,b,\gamma}(\omega)$  corresponds to filling the ellipse  $\mathfrak{B}_{a,b,\gamma}(z)$  with type 1 individuals, without affecting the rest of  $\mathbb{R}^2$ .

**State space.** We now introduce the state space over which the  $\infty$ -parent SLFV is defined. Let  $\widetilde{\mathcal{M}}_\lambda$  be the space of all measures on  $\mathbb{R}^2 \times \{0, 1\}$  whose marginal distribution over  $\mathbb{R}^2$  is Lebesgue measure. In other words,  $\widetilde{\mathcal{M}}_\lambda$  is the space of all measures  $M$  on  $\mathbb{R}^2 \times \{0, 1\}$  such that the following property is satisfied:

$$\forall f \in C_c(\mathbb{R}^2), \int_{\mathbb{R}^2 \times \{0, 1\}} f(z) M(dz, dk) = \int_{\mathbb{R}^2} f(z) dz.$$

By a standard decomposition theorem, for all  $M \in \widetilde{\mathcal{M}}_\lambda$ , there exists  $\omega : \mathbb{R}^2 \rightarrow [0, 1]$  measurable such that

$$M(dz, dk) = (\omega(z)\delta_0(dk) + (1 - \omega(z))\delta_1(dk))dx. \quad (2.1)$$

Let  $\mathcal{M}_\lambda$  be the set of all measures  $M \in \widetilde{\mathcal{M}}_\lambda$  such that there exists  $\omega : \mathbb{R}^2 \rightarrow \{0, 1\}$  (instead of  $[0, 1]$ ) satisfying (2.1). For all  $M \in \mathcal{M}_\lambda$ , we refer to any measurable function  $\omega : \mathbb{R}^2 \rightarrow \{0, 1\}$  satisfying (2.1) as a *density* of  $M$ , and denote it by  $\omega_M$ . Note that  $\omega_M$  is not unique, but two densities will only differ at a Lebesgue-null set of points. Therefore, since we shall only consider integrals of continuous functions with respect to the measures describing the current state of the population, the choice of the density  $\omega_M$  used in the analysis below will not matter. Moreover,  $\omega_M$  represents the density in type 0 (or ghost) individuals. This choice is motivated by the duality formula proved in [32], which we recall in Section 2.2.

We endow  $\mathcal{M}_\lambda$  with the topology of vague convergence. Let  $D_{\mathcal{M}_\lambda}[0, +\infty)$  be the space of all càdlàg  $\mathcal{M}_\lambda$ -valued paths, endowed with the standard Skorokhod topology.

*Remark 2.2.* Assume that the state of the population at time  $t$  is encoded by  $M \in \mathcal{M}_\lambda$ . Since the

density  $\omega_M$  is not uniquely defined, for any given location  $z \in \mathbb{R}^2$ ,  $\omega_M(z)$  is not uniquely defined. Hence, it cannot be interpreted as the type of the individuals living there. Note however that just below Definition 2.9, we shall introduce a way to uniquely assign a type to any given location, which does not depend on the choice of the density.

**Martingale problem.** When reproduction events have bounded shape parameters, it is possible to define the  $\infty$ -parent SLFV as the solution to a well-posed martingale problem. In order to do so, we need to introduce some more notation. Let  $C^1(\mathbb{R})$  be the space of all continuously differentiable functions  $F : \mathbb{R} \rightarrow \mathbb{R}$ . For all  $f \in C_c(\mathbb{R}^2)$  and  $F \in C^1(\mathbb{R})$ , if  $\omega : \mathbb{R}^2 \rightarrow \{0, 1\}$  is measurable, we set

$$\langle \omega, f \rangle := \int_{\mathbb{R}^2} f(z) \omega(z) dz.$$

Moreover, we define the function  $\Psi_{F,f} : M \in \mathcal{M}_\lambda \rightarrow \Psi_{F,f}(M) \in \mathbb{R}$  by

$$\forall M \in \mathcal{M}_\lambda, \Psi_{F,f}(M) := F(\langle \omega_M, f \rangle).$$

The functions  $\Psi_{F,f}$  with  $F \in C^1(\mathbb{R})$  and  $f \in C_c(\mathbb{R}^2)$  will be used as the set of functions on which the martingale problem characterising the process of interest is defined. This martingale problem is associated to the operator  $\mathcal{L}_{\tilde{\mu}}^\infty$  defined as follows. For all  $M \in \mathcal{M}_\lambda$ ,

$$\begin{aligned} \mathcal{L}_{\tilde{\mu}}^\infty \Psi_{F,f}(M) := & \int_{S_{\tilde{\mu}}} \int_{\text{Supp}^{a,b,\gamma}(f)} \left( 1 - \delta_0 \left( \int_{\mathfrak{B}_{a,b,\gamma}(z)} (1 - \omega_M(z')) dz' \right) \right) \\ & \times \left( F(\langle \Theta_z^{a,b,\gamma}(\omega_M), f \rangle) - F(\langle \omega_M, f \rangle) \right) dz \tilde{\mu}(da, db, d\gamma). \end{aligned}$$

This operator encodes exactly the dynamics described informally in the Introduction. Indeed, reproduction events of centre  $z$  and parameters  $(a, b, \gamma)$  happen at rate  $dz \otimes \tilde{\mu}(da, db, d\gamma)$ . Whenever a reproduction event occurs, then:

1. If the affected area  $\mathfrak{B}_{a,b,\gamma}(z)$  contains a positive fraction of (real) type 1 individuals (*i.e.*, if  $\int_{\mathfrak{B}_{a,b,\gamma}(z)} (1 - \omega_M(z')) dz' \neq 0$ ), then it is filled with type 1 individuals, as encoded by the action of  $\Theta_z^{a,b,\gamma}$  on  $\omega_M$ .
2. Otherwise, it stays filled with (ghost) type 0 individuals.

We then have the following characterisation of the  $\infty$ -parent SLFV process with elliptical reproduction events.

**Theorem 2.3.** For all  $M^0 \in \mathcal{M}_\lambda$ , there exists a unique  $D_{\mathcal{M}_\lambda}[0, +\infty)$ -valued process  $(M_t)_{t \geq 0}$  such that  $M_0 = M^0$  and, for all  $F \in C^1(\mathbb{R})$  and  $f \in C_c(\mathbb{R}^2)$ ,

$$\left( \Psi_{F,f}(M_t) - \Psi_{F,f}(M_0) - \int_0^t \mathcal{L}_{\tilde{\mu}}^\infty \Psi_{F,f}(M_s) ds \right)_{t \geq 0}$$

is a martingale.

Moreover, this process is Markovian.

The proof of this theorem is a direct generalisation of the proofs of the second part of Theorem 2.14 and of Lemma 3.8 in [32], and so we omit it.

**Definition 2.4.** Let  $M^0 \in \mathcal{M}_\lambda$ . The  $\infty$ -parent spatial  $\Lambda$ -Fleming Viot process with elliptical reproduction events (or  $\infty$ -parent SLFV) with initial condition  $M^0$  associated to  $\tilde{\mu}$  is the unique solution to the martingale problem  $(\mathcal{L}_{\tilde{\mu}}^\infty, \delta_{M^0})$  stated in Theorem 2.3.

*Remark 2.5.* From a population genetics viewpoint, the  $\infty$ -parent SLFV can be seen as modelling the spread of an extremely advantageous gene, corresponding to the "real" type. For weaker strengths of selection, the most classical tool to study this question is the Fisher-KPP equation [19, 30], defined as follows. If the proportion of individuals of the favoured type at the spatial location  $x \in \mathbb{R}^d$ ,  $d \geq 1$ , and at time  $t \geq 0$  is given by  $p(t, x) \in [0, 1]$ , then  $p(t, x)$  evolves according to the Fisher-KPP equation if it is a weak solution to the equation

$$\frac{\partial p}{\partial t}(t, x) = \frac{m}{2} \Delta p(t, x) + s_0 p(t, x)(1 - p(t, x)), \quad \forall x \in \mathbb{R}^d, \forall t \geq 0, \quad (2.2)$$

where  $m \geq 0$  and  $s_0 \geq 0$  are respectively the diffusion and selection parameters.

In dimension 1, it is possible to add some stochasticity through a Wright-Fisher noise term. This cannot be done in higher dimensions. Equation (2.2) and its stochastic counterpart in dimension 1 both admit travelling wave solutions (see, e.g., [33]), corresponding to a linear speed of spread/growth. However, the spatio-temporal scales at which the spread of a weakly advantageous allele is modelled by the Fisher-KPP equation are very different from the ones in which the spread of an extremely advantageous gene is visible. The approaches and results based on the model developed here and on the Fisher-KPP equation are thus difficult to compare.

The interest of the  $\infty$ -parent SLFV as a model for expanding populations lies in the fact that it is associated to a dual process of *potential ancestors*, which can be used to study the properties of the  $\infty$ -parent SLFV. We now define this dual process, and state the duality relation.

## 2.2 Dual process and duality relation

The dual process introduced in this section is defined on a different probability space  $(\Omega, \mathcal{F}, \mathbf{P})$ .

**State space.** Let  $\mathcal{E}^c$  be the set of all subsets of  $\mathbb{R}^2$  which are Lebesgue-measurable, connected, and whose Lebesgue measure is finite and non-zero. The state space we consider for the dual process is then the set  $\mathcal{E}^{cf}$  of all finite unions of elements of  $\mathcal{E}^c$ .

**Definition of the dual process.** We can now define the dual process, called the  $\infty$ -parent ancestral process.

**Definition 2.6.** Let  $\overleftarrow{\Pi}$  be a Poisson point process with intensity measure  $dt \otimes dz \otimes \tilde{\mu}(da, db, d\gamma)$  on  $\mathbb{R}_+ \times \mathbb{R}^2 \times S_{\tilde{\mu}}$ , defined on the probability space  $(\Omega, \mathcal{F}, \mathbf{P})$ , and let  $E^0 \in \mathcal{E}^{cf}$ . The  $\mathcal{E}^{cf}$ -valued  $\infty$ -parent ancestral process  $(E_t^\infty)_{t \geq 0}$  with initial condition  $E^0$  associated to  $\tilde{\mu}$  is defined as follows.

First, we set  $E_0^\infty = E^0$ . Then, for all  $(t, z, a, b, \gamma) \in \overleftarrow{\Pi}$ , if  $E_{t-}^\infty \cap \mathfrak{B}_{a,b,\gamma}(z)$  has non-zero Lebesgue measure, we set

$$E_t^\infty = E_{t-}^\infty \cup \mathfrak{B}_{a,b,\gamma}(z).$$

**Lemma 2.7.** The process  $(E_t^\infty)_{t \geq 0}$  introduced in Definition 2.6 is well-defined and Markovian.

*Proof.* Recall that for all  $(t, z, a, b, \gamma) \in \overleftarrow{\Pi}$ , we have  $(a, b) \in (0, \mathcal{R}_{\tilde{\mu}}]^2$ . Therefore,

$$\mathfrak{B}_{a,b,\gamma}(z) \subseteq \mathcal{B}_{\mathcal{R}_{\tilde{\mu}}}(z) \text{ a.s.,}$$

where  $\mathcal{B}_{\mathcal{R}_{\tilde{\mu}}}(z)$  is the ball of radius  $\mathcal{R}_{\tilde{\mu}}$  centred at  $z$ , and so we can bound the jump rate of  $(E_t^\infty)_{t \geq 0}$  from above by that of a  $\infty$ -parent ancestral process with the same initial condition and associated to  $\tilde{\mu}(S_{\tilde{\mu}})\delta_{(\mathcal{R}_{\tilde{\mu}}, \mathcal{R}_{\tilde{\mu}}, 0)}(da, db, d\gamma)$ , which is finite (see Section 4 in [32]).  $\square$

**Duality relation.** For all  $M \in \mathcal{M}_\lambda$  and  $E \in \mathcal{E}^{cf}$ , we set

$$\tilde{D}(M, E) := \delta_0 \left( \int_E (1 - \omega_M(z)) dz \right).$$

Intuitively,  $\tilde{D}(M, E) = 0$  if the area  $E$  contains a positive fraction of real individuals when the population state is  $M$ , and  $\tilde{D}(M, E) = 1$  if the area is empty.

The  $\infty$ -parent SLFV and the  $\infty$ -parent ancestral process then satisfy the following duality relation, whose proof is similar to the one in [32].

**Proposition 2.8.** *Let  $M^0 \in \mathcal{M}_\lambda$ , and let  $(M_t^\infty)_{t \geq 0}$  be the unique solution to the martingale problem associated to  $(\mathcal{L}_{\tilde{\mu}}^\infty, \delta_{M^0})$ . Let  $E^0 \in \mathcal{E}^{cf}$ , and let  $(E_t^\infty)_{t \geq 0}$  be the  $\infty$ -parent ancestral process with initial condition  $E^0$  associated to  $\tilde{\mu}$ . Then, for all  $t \geq 0$ ,*

$$\mathbb{E}_{M^0} \left[ \tilde{D}(M_t^\infty, E^0) \right] = \mathbf{E}_{E^0} \left[ \tilde{D}(M^0, E_t^\infty) \right],$$

or equivalently,

$$\mathbb{E}_{M^0} \left[ \delta_0 \left( \int_{E^0} (1 - \omega_{M_t^\infty}(z)) dz \right) \right] = \mathbf{E}_{E^0} \left[ \delta_0 \left( \int_{E_t^\infty} (1 - \omega_{M^0}(z)) dz \right) \right].$$

This duality relation can be interpreted as follows. Whenever a reproduction event affects one area, all the individuals in the area can be considered as *potential parents*. If  $\overleftarrow{\Pi}$  encodes the reproduction events affecting the population when going *backwards in time*, then  $E_t^\infty$  encodes the locations of the potential ancestors at time 0 of the individuals living in  $E^0$  at time  $t$ . The duality relation then states that a given area  $E^0$  contains only ghost individuals at time  $t$  if, and only if all their potential ancestors at time 0 (located in  $E_t^\infty$ ) are ghost individuals.

### 2.3 Speed of growth of the occupied region in the $\infty$ -parent SLFV: Definition and main result

Let  $HP^0$  stand for the half-plane

$$HP^0 := \{(x, y) \in \mathbb{R}^2 : x \geq 0\}.$$

In order to show that the growth of the region occupied by type 1 individuals in the  $\infty$ -parent SLFV is linear in time, we consider that initially, type 1 individuals cover the half-plane

$$\overline{HP}^0 := \{(x, y) \in \mathbb{R}^2 : x < 0\}.$$

That is, we take the measure  $M^{HP}(dz) := \omega^{HP}(z)dz$  as an initial condition for the  $\infty$ -parent SLFV, where

$$\forall z \in \mathbb{R}^2, \quad \omega^{HP}(z) = \mathbf{1}_{HP^0}(z).$$

Let  $(M_t^{HP})_{t \geq 0}$  be the  $\infty$ -parent SLFV with initial condition  $M^{HP}$  associated to  $\tilde{\mu}$ . Moreover, for all  $t \geq 0$ , let  $\omega_t^{HP}$  be a density of  $M_t^{HP}$ .

**Definition 2.9.** For all  $x \in \mathbb{R}$ , let

$$\vec{\tau}_x := \min \left\{ t \geq 0 : \lim_{\epsilon \rightarrow 0} V_\epsilon^{-1} \int_{\mathcal{B}_\epsilon((x,0))} \omega_t^{HP}(z) dz < 1 \right\},$$

where we recall that  $\mathcal{B}_\epsilon((x,0))$  is the ball with centre  $(x,0)$  and radius  $\epsilon$ , and  $V_\epsilon$  is the volume of  $\mathcal{B}_\epsilon((x,0))$ .

Informally,  $\vec{\tau}_x$  is the first time at which the location  $(x,0)$  is reached by type 1 individuals. Since densities are only defined up to a Lebesgue null set, we consider that  $(x,0)$  is occupied by type 1 individuals if, and only if all neighbourhoods of  $(x,0)$  contain a non-zero fraction of type 1 individuals.

The following result means that once the location  $(x,0)$  is occupied by type 1 individuals, it cannot become empty again. However, notice that the function  $x \rightarrow \vec{\tau}_x$  is not necessarily increasing. See Figure 2 for an illustration.

**Lemma 2.10.** Let  $x \in \mathbb{R}$ . Then, for all  $t \geq 0$ ,

$$\mathbb{P} \left( \lim_{\epsilon \rightarrow 0} V_\epsilon^{-1} \int_{\mathcal{B}_\epsilon((x,0))} \omega_t^{HP}(z) dz < 1 \mid \vec{\tau}_x \leq t \right) = 1.$$

This property is obvious from the set-valued informal description of the  $\infty$ -parent SLFV given in the Introduction. However, because this description does not hold when the initial area occupied by type 1 individuals has infinite volume, we have to work a bit more to show this property by using the martingale problem characterisation.

*Proof.* We first show that for all  $\epsilon > 0$ ,

$$t \rightarrow \mathbb{P} \left( V_\epsilon^{-1} \int_{\mathcal{B}_\epsilon((x,0))} \omega_t^{HP}(z) dz < 1 \right)$$

is a non-decreasing function.

In order to do so, let  $\epsilon > 0$ , and let  $0 \leq t \leq t'$ . Then, by Proposition 2.8,

$$\begin{aligned} \mathbb{P} \left( V_\epsilon^{-1} \int_{\mathcal{B}_\epsilon((x,0))} \omega_t^{HP}(z) dz < 1 \right) &= \mathbb{P} \left( \delta_0 \left( \int_{\mathcal{B}_\epsilon((x,0))} (1 - \omega_t^{HP}(z)) dz \right) = 0 \right) \\ &= 1 - \mathbb{E} \left[ \delta_0 \left( \int_{\mathcal{B}_\epsilon((x,0))} (1 - \omega_t^{HP}(z)) dz \right) \right] \\ &= 1 - \mathbf{E} \left[ \delta_0 \left( \int_{E_t^\infty} (1 - \omega_{M^0}(z)) dz \right) \right], \end{aligned}$$

where  $(E_t^\infty)_{t \geq 0}$  is the  $\infty$ -parent ancestral process with initial condition  $\mathcal{B}_\epsilon((x,0))$  and intensity  $\tilde{\mu}$ .

Similarly,

$$\mathbb{P}\left(V_\epsilon^{-1} \int_{\mathcal{B}_\epsilon((x,0))} \omega_{t'}^{HP}(z) dz < 1\right) = 1 - \mathbf{E}\left[\delta_0\left(\int_{E_{t'}^\infty} (1 - \omega_{M^0}(z)) dz\right)\right].$$

Moreover, the sequence  $(E_t^\infty)_{t \geq 0}$  is increasing for the inclusion and since  $\omega_{M^0}$  is  $\{0, 1\}$ -valued, we have

$$\mathbf{E}\left[\delta_0\left(\int_{E_t^\infty} (1 - \omega_{M^0}(z)) dz\right)\right] \geq \mathbf{E}\left[\delta_0\left(\int_{E_{t'}^\infty} (1 - \omega_{M^0}(z)) dz\right)\right].$$

We conclude that

$$\mathbb{P}\left(V_\epsilon^{-1} \int_{\mathcal{B}_\epsilon((x,0))} \omega_t^{HP}(z) dz < 1\right) \leq \mathbb{P}\left(V_\epsilon^{-1} \int_{\mathcal{B}_\epsilon((x,0))} \omega_{t'}^{HP}(z) dz < 1\right).$$

Then, let  $t \geq 0$ , and let  $(\epsilon_n)_{n \in \mathbb{N}}$  be a sequence decreasing to 0. Since

$$\lim_{\epsilon \rightarrow 0} V_\epsilon^{-1} \int_{\mathcal{B}_\epsilon((x,0))} \omega_t^{HP}(z) dz$$

exists, we have

$$\mathbb{P}\left(\lim_{\epsilon \rightarrow 0} V_\epsilon^{-1} \int_{\mathcal{B}_\epsilon((x,0))} \omega_t^{HP}(z) dz < 1 \mid \vec{\tau}_x \leq t\right) = \mathbb{P}\left(\lim_{n \rightarrow +\infty} V_{\epsilon_n}^{-1} \int_{\mathcal{B}_{\epsilon_n}((x,0))} \omega_t^{HP}(z) dz < 1 \mid \vec{\tau}_x \leq t\right).$$

Noticing that the sequence  $(\mathcal{B}_{\epsilon_n}((x,0)))_{n \in \mathbb{N}}$  is decreasing for the inclusion, we have that if

$$V_{\epsilon_n}^{-1} \int_{\mathcal{B}_{\epsilon_n}((x,0))} \omega_t^{HP}(z) dz = 1$$

for some integer  $n$ , then this is also the case for all  $N \geq n$ . Therefore,

$$\begin{aligned} \mathbb{P}\left(\lim_{n \rightarrow +\infty} V_{\epsilon_n}^{-1} \int_{\mathcal{B}_{\epsilon_n}((x,0))} \omega_t^{HP}(z) dz = 1 \mid \vec{\tau}_x \leq t\right) &= \mathbb{P}\left(\exists n \in \mathbb{N}, V_{\epsilon_n}^{-1} \int_{\mathcal{B}_{\epsilon_n}((x,0))} \omega_t^{HP}(z) dz = 1 \mid \vec{\tau}_x \leq t\right) \\ &\leq \sum_{n \in \mathbb{N}} \mathbb{P}\left(V_{\epsilon_n}^{-1} \int_{\mathcal{B}_{\epsilon_n}((x,0))} \omega_t^{HP}(z) dz = 1 \mid \vec{\tau}_x \leq t\right) \\ &\leq \sum_{n \in \mathbb{N}} \mathbb{P}\left(V_{\epsilon_n}^{-1} \int_{\mathcal{B}_{\epsilon_n}((x,0))} \omega_{\vec{\tau}_x}^{HP}(z) dz = 1 \mid \vec{\tau}_x \leq t\right) \\ &= 0 \end{aligned}$$

by definition of  $\vec{\tau}_x$ . Here we have used the first part of the proof to pass from the second line to the third line, which allows us to conclude.  $\square$

The goal of the article is to show the following result, which tells us that the growth of the occupied area in the  $\infty$ -parent SLFV is linear in time.

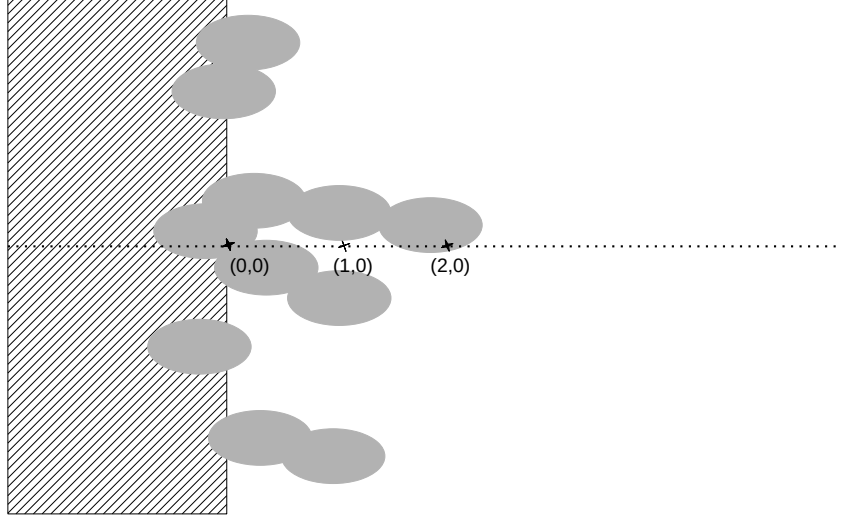


Figure 2: State of a population evolving according to the  $\infty$ -parent SLFV at time  $t$ . Initially, type 1 individuals cover the half-plane  $\overline{HP}^0$ , corresponding to the hatched area. Each grey ellipse represents a reproduction event occurring during the time interval  $[0, t]$  which overlaps an area initially empty, and resulting in the corresponding area being completely filled with type 1 individuals. Here  $\overrightarrow{\tau}_1 > t$  but  $\overrightarrow{\tau}_2 \leq t$ .

**Theorem 2.11.** *There exists  $\nu > 0$  such that*

$$\lim_{x \rightarrow +\infty} \frac{\mathbb{E}[\overrightarrow{\tau}_x]}{x} = \nu.$$

We can then interpret  $\nu^{-1}$  as the limiting speed of growth of the process. As explained in the Introduction, the proof of Theorem 2.11 spans Sections 3 and 4 and is concluded at the end of Section 4.3.

In order to show Theorem 2.11, we make use of the dual process associated to the  $\infty$ -parent SLFV, and define an equivalent of  $\overrightarrow{\tau}_x$  for the  $\infty$ -parent ancestral process. In order to do so, we first introduce a slight generalisation of the  $\infty$ -parent ancestral process which allows to take single points as initial conditions. This process is defined on the state space

$$\tilde{\mathcal{E}}^{cf} := \mathcal{E} \cup \{\{z\} : z \in \mathbb{R}^2\}.$$

**Definition 2.12.** *Let  $\overleftarrow{\Pi}$  be a Poisson point process with intensity measure  $dt \otimes dz \otimes \tilde{\mu}(da, db, d\gamma)$  on  $\mathbb{R}_+ \times \mathbb{R}^2 \times S_{\tilde{\mu}}$ , defined on the probability space  $(\Omega, \mathcal{F}, \mathbf{P})$ , and let  $z_0 \in \mathbb{R}^2$ . The  $\tilde{\mathcal{E}}^{cf}$ -valued  $\infty$ -parent ancestral process  $(E_t^{\infty, z_0})_{t \geq 0}$  with initial condition  $\{z_0\}$  and associated to  $\tilde{\mu}$  is defined as follows.*

*First, we set  $E_0^{\infty, z_0} = \{z_0\}$ . Then, for all  $(t, z, a, b, \gamma) \in \overleftarrow{\Pi}$ :*

- *If  $E_{t-}^{\infty, z_0} = \{z_0\}$  and  $z_0 \in \mathfrak{B}_{a,b,\gamma}(z)$ , we set  $E_t^{\infty, z_0} = \mathfrak{B}_{a,b,\gamma}(z)$ .*



- If  $E_{t-}^{\infty, z_0} \neq \{z_0\}$  and  $E_{t-}^{\infty, z_0} \cap \mathfrak{B}_{a,b,\gamma}(z)$  has non-zero Lebesgue measure, we set

$$E_t^{\infty, z_0} = E_{t-}^{\infty, z_0} \cup \mathfrak{B}_{a,b,\gamma}(z).$$

Using the same argument as for the initial  $\infty$ -parent ancestral process,  $(E_t^{\infty, z_0})_{t \geq 0}$  is well-defined and Markovian. Moreover, once it jumps for the first time, its behaviour is identical to that of the original  $\infty$ -parent ancestral process.

In all that follows, let  $\overleftarrow{\Pi}$  be a Poisson point process as in Definition 2.12. For all  $x > 0$  and all  $\epsilon > 0$ , let  $(E_t^{\epsilon, x})_{t \geq 0}$  be a sequence of  $\infty$ -parent ancestral processes with initial condition  $\mathcal{B}_\epsilon((x, 0))$  associated to  $\tilde{\mu}$ , all constructed using the same underlying Poisson point process  $\overleftarrow{\Pi}$ . Moreover, let  $(E_t^x)_{t \geq 0}$  be the  $\infty$ -parent ancestral process with initial condition  $\{(x, 0)\}$  associated to  $\tilde{\mu}$ , also constructed using  $\overleftarrow{\Pi}$ . That is, in the heavier notation of Definition 2.12,  $E^x = E^{\infty, (x, 0)}$ . Then,  $(E_t^{\epsilon, x})_{t \geq 0}$ ,  $\epsilon > 0$  and  $(E_t^x)_{t \geq 0}$  satisfy the following property.

**Lemma 2.13.** *For all  $x > 0$ , if  $t_0^x$  is the first time at which  $(x, 0)$  is affected by a reproduction event, then a.s. there exists  $\epsilon_0^x > 0$  such that for all  $t \geq t_0^x$ , we have*

$$\forall \epsilon < \epsilon_0^x, \quad E_t^{\epsilon, x} = E_t^x.$$

Furthermore, we have a.s. for all  $t \geq 0$ ,

$$\lim_{\epsilon \rightarrow 0} \text{Vol} \left( E_t^{\epsilon, x} \cap \overline{HP}^0 \right) = \text{Vol} \left( E_t^x \cap \overline{HP}^0 \right).$$

Concretely,  $\epsilon_0^x$  is the radius of a ball around  $(x, 0)$  small enough to be included in the area of the first event overlapping  $(x, 0)$ , and not to be overlapped by any event in  $\overleftarrow{\Pi}$  occurring before  $t_0^x$  (see Condition (2.3) below).

*Proof.* Let  $x > 0$ , and let  $(t_0^x, z_0^x, a_0^x, b_0^x, \gamma_0^x) \in \overleftarrow{\Pi}$  be the first reproduction event to affect  $(x, 0)$ . In order to show the first part of the lemma, we want to show that a.s. there exists  $\epsilon_0^x > 0$  such that

$$\mathcal{B}_{\epsilon_0^x}((x, 0)) \subseteq \mathfrak{B}_{a_0^x, b_0^x, \gamma_0^x}(z_0^x) \text{ and } \forall (t, z, a, b, \gamma) \in \overleftarrow{\Pi} \text{ s.t. } t < t_0^x, \text{Vol}(\mathcal{B}_{\epsilon_0^x}((x, 0)) \cap \mathfrak{B}_{a,b,\gamma}(z)) = 0. \quad (2.3)$$

Indeed, if such a  $\epsilon_0^x$  exists, then for all  $0 < \epsilon < \epsilon_0^x$ , the first time at which  $(E_t^{\epsilon, x})_{t \geq 0}$  jumps is  $t_0^x$ . Moreover,

$$E_{t_0^x}^{\epsilon, x} = E_0^{\epsilon, x} \cup \mathfrak{B}_{a_0^x, b_0^x, \gamma_0^x}(z_0^x)$$

$$\begin{aligned}
&= \mathcal{B}_\epsilon((x, 0)) \cup \mathfrak{B}_{a_0^x, b_0^x, \gamma_0^x}(z_0^x) \\
&= \mathfrak{B}_{a_0^x, b_0^x, \gamma_0^x}(z_0^x), \\
&= E_{t_0^x}^x,
\end{aligned}$$

yielding the first part of the lemma.

As the probability that  $(x, 0)$  belongs to the boundary of  $\mathfrak{B}_{a_0^x, b_0^x, \gamma_0^x}(z_0^x)$  is equal to zero, with probability 1 there exists  $\tilde{\epsilon} > 0$  such that

$$\mathcal{B}_{\tilde{\epsilon}}((x, 0)) \subseteq \mathfrak{B}_{a_0^x, b_0^x, \gamma_0^x}(z_0^x).$$

Let us consider the number  $N_{\tilde{\epsilon}}$  of reproduction events that affect  $\mathcal{B}_{\tilde{\epsilon}}((x, 0))$  during the time interval  $[0, t_0^x)$ , that is, all reproduction events  $(t, z, a, b, \gamma) \in \overleftarrow{\Pi}$  such that  $t \in [0, t_0^x)$  and

$$\mathcal{B}_{\tilde{\epsilon}}((x, 0)) \cap \mathfrak{B}_{a, b, \gamma}(z) \neq \emptyset.$$

Since the support of  $\tilde{\mu}$  is bounded,  $N_{\tilde{\epsilon}}$  is a.s. finite. By construction, the point  $(x, 0)$  belongs to none of the finitely many closed ellipses corresponding to these events, and so the minimal distance  $\epsilon$  between  $(x, 0)$  and these areas is a.s. positive. We can then choose any  $\epsilon_0^x \in (0, \epsilon)$  and (2.3) holds true, which concludes the proof of the first part of Lemma 2.13.

In order to show the second part of the lemma, let us first consider  $t \in [0, t_0^x)$ , and let us assume that  $\epsilon_0^x$  exists (which is almost surely satisfied). Then,

$$\text{Vol}\left(E_t^x \cap \overline{HP}^0\right) = \text{Vol}\left(\{(x, 0)\} \cap \overline{HP}^0\right) = 0.$$

Moreover, by (2.3), for all  $0 < \epsilon < \min(\epsilon_0^x, x/2)$  we have

$$\text{Vol}\left(E_t^{\epsilon, x} \cap \overline{HP}^0\right) = \text{Vol}\left(\mathcal{B}_\epsilon((x, 0)) \cap \overline{HP}^0\right) = \text{Vol}(\emptyset) = 0.$$

Therefore,

$$\lim_{\epsilon \rightarrow 0} \text{Vol}\left(E_t^{\epsilon, x} \cap \overline{HP}^0\right) = 0 = \text{Vol}\left(E_t^x \cap \overline{HP}^0\right).$$

Let us now consider  $t \geq t_0^x$ . For every  $\epsilon \in (0, \epsilon_0^x)$ , we have

$$\text{Vol}\left(E_t^{\epsilon, x} \cap \overline{HP}^0\right) = \text{Vol}\left(E_t^x \cap \overline{HP}^0\right),$$

and so

$$\lim_{\epsilon \rightarrow 0} \text{Vol} \left( E_t^{\epsilon, x} \cap \overline{HP}^0 \right) = \text{Vol} \left( E_t^x \cap \overline{HP}^0 \right).$$

□

We can now introduce an equivalent of  $\vec{\tau}_x$  for the  $\infty$ -parent ancestral process.

**Definition 2.14.** For all  $x > 0$ , let  $(t_0^x, z_0^x, a_0^x, b_0^x, \gamma_0^x) \in \overleftarrow{\Pi}$  be the first reproduction event to affect  $(x, 0)$ , and let

$$\tilde{\tau}_x := \begin{cases} \min \left\{ t \geq 0 : \lim_{\epsilon \rightarrow 0} \text{Vol} \left( E_t^{\epsilon, x} \cap \overline{HP}^0 \right) > 0 \right\} & \text{if there exists } \epsilon_0^x > 0 \text{ satisfying (2.3),} \\ 0 & \text{otherwise.} \end{cases}$$

**Lemma 2.15.** For all  $x > 0$ ,  $\tilde{\tau}_x$  is well-defined, and is almost surely equal to

$$\min \left\{ t \geq 0 : E_t^x \cap \overline{HP}^0 \neq \emptyset \right\}.$$

In particular,  $\tilde{\tau}_x \neq 0$  a.s.

*Proof.* Let  $x > 0$ . By Lemma 2.13,  $\tilde{\tau}_x$  is well-defined and almost surely equal to

$$\min \left\{ t \geq 0 : \text{Vol} \left( E_t^x \cap \overline{HP}^0 \right) > 0 \right\} \geq t_0^x,$$

since the support of  $\tilde{\mu}$  is bounded. Moreover, since for all  $(a, b, \gamma) \in S_{\tilde{\mu}}$ , the set

$$\left\{ z \in \mathbb{R}^2 : \mathfrak{B}_{a, b, \gamma}(z) \cap \overline{HP}^0 \neq \emptyset \text{ and } \text{Vol} \left( \mathfrak{B}_{a, b, \gamma}(z) \cap \overline{HP}^0 \right) = 0 \right\}$$

has zero Lebesgue measure, we have

$$\min \left\{ t \geq 0 : \text{Vol} \left( E_t^x \cap \overline{HP}^0 \right) > 0 \right\} = \min \left\{ t \geq 0 : E_t^x \cap \overline{HP}^0 \neq \emptyset \right\} \text{ a.s.,}$$

which allows us to conclude. □

In order to use  $(\tilde{\tau}_x)_{x>0}$  to derive the desired properties of  $(\vec{\tau}_x)_{x>0}$ , we shall need the following result.

**Lemma 2.16.** For all  $x > 0$ ,  $\tilde{\tau}_x$  and  $\vec{\tau}_x$  have the same distribution.

*Proof.* Recall that constructing all the  $\infty$ -parent ancestral process  $(E_t^{\epsilon, x})_{t \geq 0}$  using the same underlying

Poisson point process ensures that for all  $t \geq 0$  and all  $x > 0$ , we have

$$\forall 0 < \epsilon < \epsilon', \quad E_t^{\epsilon, x} \subseteq E_t^{\epsilon', x}. \quad (2.4)$$

Moreover, for all  $\epsilon > 0$ , we set

$$\begin{aligned} \overrightarrow{\tau}_x^\epsilon &:= \min \left\{ t > 0 : V_\epsilon^{-1} \int_{\mathcal{B}_\epsilon((x,0))} \omega_t^{HP}(z) dz < 1 \right\} \\ \text{and } \tilde{\tau}_x^\epsilon &:= \min \left\{ t > 0 : \text{Vol} \left( E_t^{\epsilon, x} \cap \overline{HP}^0 \right) > 0 \right\}. \end{aligned}$$

Using the same argument as in the proof of Lemma 2.15, we obtain that

$$\tilde{\tau}_x = \lim_{\epsilon \rightarrow 0} \tilde{\tau}_x^\epsilon \quad \text{a.s.}$$

We have a similar result for  $\overrightarrow{\tau}_x$ . Indeed, for all  $0 < \epsilon < \epsilon'$ , by definition of  $\overrightarrow{\tau}_x^\epsilon$  we can write

$$\begin{aligned} \int_{\mathcal{B}_{\epsilon'}((x,0))} \omega_{\overrightarrow{\tau}_x^\epsilon}^{HP}(z) dz &= \int_{\mathcal{B}_\epsilon((x,0))} \omega_{\overrightarrow{\tau}_x^\epsilon}^{HP}(z) dz + \int_{\mathcal{B}_{\epsilon'}((x,0)) \setminus \mathcal{B}_\epsilon((x,0))} \omega_{\overrightarrow{\tau}_x^\epsilon}^{HP}(z) dz \\ &< V_\epsilon + V_{\epsilon'} - V_\epsilon \\ &< V_{\epsilon'}. \end{aligned}$$

Therefore,  $\overrightarrow{\tau}_x^{\epsilon'} \leq \overrightarrow{\tau}_x^\epsilon$  and  $\lim_{\epsilon \rightarrow 0} \overrightarrow{\tau}_x^\epsilon$  exists. Moreover, for all  $t \geq \overrightarrow{\tau}_x$ , by Lemma 2.10 we have

$$\lim_{\epsilon \rightarrow 0} V_\epsilon^{-1} \int_{\mathcal{B}_\epsilon((x,0))} \omega_t^{HP}(z) dz < 1,$$

and so there exists  $\epsilon_t > 0$  such that for all  $0 < \epsilon \leq \epsilon_t$ ,

$$V_\epsilon^{-1} \int_{\mathcal{B}_\epsilon((x,0))} \omega_t^{HP}(z) dz < 1,$$

and hence for all  $0 < \epsilon \leq \epsilon_t$ ,  $t \geq \overrightarrow{\tau}_x^\epsilon$ . Similarly, if  $t > \lim_{\epsilon \rightarrow 0} \overrightarrow{\tau}_x^\epsilon$ , then again by Lemma 2.10, there exists  $\epsilon_t > 0$  such that for all  $0 < \epsilon \leq \epsilon_t$ ,

$$V_\epsilon^{-1} \int_{\mathcal{B}_\epsilon((x,0))} \omega_t^{HP}(z) dz < 1$$

and  $t > \overrightarrow{\tau}_x$ . Therefore,

$$\overrightarrow{\tau}_x = \lim_{\epsilon \rightarrow 0} \overrightarrow{\tau}_x^\epsilon \quad \text{a.s.}$$

Consequently, it is sufficient to show that  $\overleftarrow{\tau}_x^\epsilon$  and  $\tilde{\tau}_x^\epsilon$  have the same distribution for all  $\epsilon > 0$  in order to conclude the proof.

Let  $\epsilon > 0$ , and let  $t \geq 0$ . Since  $t \rightarrow \int_{\mathcal{B}_\epsilon((x,0))} \omega_t^{HP}(z) dz$  is non-increasing and  $[0, V_\epsilon]$ -valued, we have

$$\begin{aligned}
\mathbb{P}_{M_0^{HP}}(\overleftarrow{\tau}_x^\epsilon > t) &= \mathbb{P}_{M_0^{HP}}\left(\int_{\mathcal{B}_\epsilon((x,0))} \omega_t^{HP}(z) dz = V_\epsilon\right) \\
&= \mathbb{P}_{M_0^{HP}}\left(\delta_0\left(\int_{\mathcal{B}_\epsilon((x,0))} (1 - \omega_t^{HP}(z)) dz\right) = 1\right) \\
&= \mathbb{E}_{M_0^{HP}}\left[\delta_0\left(\int_{\mathcal{B}_\epsilon((x,0))} (1 - \omega_t^{HP}(z)) dz\right)\right] \\
&= \mathbf{E}_{\mathcal{B}_\epsilon((x,0))}\left[\delta_0\left(\int_{E_t^{\epsilon,x}} (1 - \omega_0^{HP}(z)) dz\right)\right] \\
&= \mathbf{P}_{\mathcal{B}_\epsilon((x,0))}\left(\delta_0\left(\int_{E_t^{\epsilon,x}} (1 - \omega_0^{HP}(z)) dz\right) = 1\right) \\
&= \mathbf{P}_{\mathcal{B}_\epsilon((x,0))}\left(\text{Vol}\left(E_t^{\epsilon,x} \cap \overline{HP^0}\right) = 0\right) \\
&= \mathbf{P}_{\mathcal{B}_\epsilon((x,0))}(\tilde{\tau}_x^\epsilon > t).
\end{aligned}$$

Here we have used Proposition 2.8 to pass from the third to the fourth line. □

For all  $x > 0$ , let  $HP^x$  stand for the half-plane

$$HP^x := \{(x', y) \in \mathbb{R}^2 : x' \geq x\},$$

extending the notation  $HP^0$  to the case  $x \geq 0$ .

In practice, the set of variables  $(\tilde{\tau}_x)_{x>0}$  is not very convenient to use. Indeed, if  $x < x'$ , even if  $\tilde{\tau}_x$  and  $\tilde{\tau}_{x'}$  are constructed using the same underlying Poisson point process, this is not sufficient to have  $\tilde{\tau}_x \leq \tilde{\tau}_{x'}$  a.s., as the underlying  $\infty$ -parent ancestral processes have different starting positions. Therefore, we define another sequence  $(\overleftarrow{\tau}_x)_{x>0}$  such that for all  $x > 0$ ,  $\overleftarrow{\tau}_x$  and  $\tilde{\tau}_x$  have the same distribution, but also such that all underlying  $\infty$ -parent ancestral processes start from the same location  $(0,0)$ , thereby ensuring that for all  $0 < x < x'$ , we have  $\overleftarrow{\tau}_x \leq \overleftarrow{\tau}_{x'}$  a.s. Moreover, for every  $x > 0$ , the  $\infty$ -parent ancestral process associated to  $\overleftarrow{\tau}_x$  will have the same distribution as the symmetric of the ancestral process associated to  $\tilde{\tau}_x$  with respect to the axis  $\{(x/2, y) : y \in \mathbb{R}\}$ . In other words, we shall be interested in the expansion of the new process "from left to right", starting from  $(0,0)$  and reaching abscissa  $x$ , instead of studying the expansion of  $E^x$  "from right to left", starting from  $(x,0)$  and reaching abscissa 0.

Let  $\tilde{\mu}^\leftarrow$  be the finite measure on  $(0, +\infty)^2 \times S_{\tilde{\mu}}$  defined by the property that for all measurable

subsets  $I, J$  of  $(0, +\infty)$  and for all  $-\pi/2 < \gamma_1 \leq \gamma_2 < \pi/2$ ,

$$\tilde{\mu}^{\leftarrow}(I \times J \times [\gamma_1, \gamma_2]) = \tilde{\mu}(I \times J \times [-\gamma_2, -\gamma_1]).$$

Let  $(E_t^\epsilon)_{t \geq 0}$ ,  $\epsilon > 0$ , be a sequence of  $\infty$ -parent ancestral processes associated to  $\tilde{\mu}^{\leftarrow}$  with initial condition  $\mathcal{B}_\epsilon((0, 0))$ , constructed using the same underlying Poisson point process  $\bar{\Pi}$ , and let  $(E_t)_{t \geq 0}$  be the  $\infty$ -parent ancestral process with initial condition  $\{(0, 0)\}$  associated to  $\tilde{\mu}^{\leftarrow}$ , also constructed using the Poisson point process  $\bar{\Pi}$ . For definiteness, we suppose that these new objects are also constructed on the probability space  $(\Omega, \mathcal{F}, \mathbf{P})$ . We shall call Condition (2.3)' the analogue of Condition (2.3) with  $x = 0$  and  $\bar{\Pi}$  replaced by  $\bar{\Pi}$ .

**Definition 2.17.** Let  $(t_0, z_0, a_0, b_0, \gamma_0) \in \bar{\Pi}$  be the first reproduction event to affect  $(0, 0)$ , and for all  $x > 0$ , let

$$\leftarrow{\tau}_x := \begin{cases} \min \left\{ t \geq 0 : \lim_{\epsilon \rightarrow 0} \text{Vol}(E_t^\epsilon \cap HP^x) > 0 \right\} & \text{if there exists } \epsilon_0 > 0 \text{ satisfying Cond. (2.3)',} \\ 0 & \text{otherwise.} \end{cases}$$

As for  $(\tilde{\tau}_x)_{x > 0}$ , for all  $x > 0$ ,  $\leftarrow{\tau}_x$  is well-defined and almost surely equal to

$$\min \{ t \geq 0 : E_t \cap HP^x \neq \emptyset \}.$$

Let us say that a point  $z = (x, y) \in \mathbb{R}^2$  is at *horizontal separation*  $d$  of the point  $z' = (x', y') \in \mathbb{R}^2$  if, and only if  $x - x' = d$ . Then informally,  $\leftarrow{\tau}_x$  represents the first time the  $\infty$ -parent ancestral process starting from  $(0, 0)$  reaches points at horizontal separation of at least  $x$  from the starting location. Moreover, we have the following lemma.

**Lemma 2.18.** *The function  $x \rightarrow \leftarrow{\tau}_x$  is nondecreasing.*

*Proof.* We distinguish two cases. If there does not exist  $\epsilon_0 > 0$  satisfying (2.3)', then for all  $x > 0$ ,  $\leftarrow{\tau}_x = 0$ , and we can conclude.

If there exists  $\epsilon_0 > 0$  satisfying (2.3), let  $0 < x_1 < x_2$ . By similar arguments as in Lemma 2.13 and its proof, we can write that

$$\begin{aligned} \forall 0 < \epsilon \leq \min(\epsilon_0, x_1/2), \quad \text{Vol}(E_t^\epsilon \cap HP^{x_1}) &= \text{Vol}\left(E_t^{\min(\epsilon_0, x_1/2)} \cap HP^{x_1}\right) \\ \text{and } \forall 0 < \epsilon \leq \min(\epsilon_0, x_2/2), \quad \text{Vol}(E_t^\epsilon \cap HP^{x_2}) &= \text{Vol}\left(E_t^{\min(\epsilon_0, x_2/2)} \cap HP^{x_2}\right). \end{aligned}$$

Let us set  $\epsilon' = \min(\epsilon_0, x_1/2, x_2/2)$ . Then,  $HP^{x_2} \subset HP^{x_1}$  and so

$$\text{Vol}\left(E_{\overleftarrow{\tau}_{x_2}}^{\epsilon'} \cap HP^{x_1}\right) \geq \text{Vol}\left(E_{\overleftarrow{\tau}_{x_2}}^{\epsilon'} \cap HP^{x_2}\right) = \lim_{\epsilon \rightarrow 0} \text{Vol}\left(E_{\overleftarrow{\tau}_{x_2}}^{\epsilon} \cap HP^{x_2}\right) > 0$$

by definition of  $\overleftarrow{\tau}_{x_2}$ , and so  $\overleftarrow{\tau}_{x_1} \leq \overleftarrow{\tau}_{x_2}$  and we can conclude.  $\square$

Notice that we are now studying the expansion of the backwards-in-time process in the same direction as the occupied area in the forwards-in-time process. Conversely,  $(\tilde{\tau}_x)_{x>0}$  corresponds to the expansion of the  $\infty$ -parent ancestral process in the *opposite direction*. Moreover, we recall that  $(\overleftarrow{\tau}_x)_{x>0}$  can be seen as being constructed using an underlying  $\infty$ -parent ancestral process which is the symmetric of the one used to construct  $(\tilde{\tau}_x)_{x>0}$  with respect to the axis  $\{(x/2, y) : y \in \mathbb{R}\}$ . This observation yields the following lemma.

**Lemma 2.19.** *For all  $x > 0$ ,  $\overleftarrow{\tau}_x$  and  $\tilde{\tau}_x$  have the same distribution.*

*Proof.* Let  $x > 0$ . For all  $t \geq 0$  and  $\epsilon > 0$ , let  $\text{Sym}(E_t^{\epsilon, x})$  be the symmetric of  $E_t^{\epsilon, x}$  with respect to the axis  $\{(x/2, y) : y \in \mathbb{R}\}$ . Then,  $(\text{Sym}(E_t^{\epsilon, x}))_{t \geq 0}$ ,  $\epsilon > 0$  is a sequence of  $\infty$ -parent ancestral processes with initial condition  $\mathcal{B}_\epsilon((0, 0))$ , all constructed using the same Poisson point process having the same distribution as  $\overline{\Pi}$  (that is, whose intensity measure is  $dt \otimes dz \otimes \tilde{\mu}^\leftarrow(da, db, d\gamma)$ ), which can be used to construct  $\overleftarrow{\tau}_x$ . Moreover, for all  $t \geq 0$  and  $\epsilon > 0$ ,

$$\text{Vol}(E_t^{\epsilon, x} \cap \overline{HP}^0) > 0 \quad \text{if, and only if} \quad \text{Vol}(\text{Sym}(E_t^{\epsilon, x}) \cap HP^x) > 0,$$

which allows us to conclude.  $\square$

In order to show Theorem 2.11, we shall use the following proposition, which is a direct consequence of Lemmas 2.16 and 2.19.

**Proposition 2.20.** *For all  $x > 0$ ,  $\overleftarrow{\tau}_x$  and  $\overrightarrow{\tau}_x$  have the same distribution.*

Contrary to what is usually considered for stochastic growth models,  $\overleftarrow{\tau}_x$  does not correspond to the growth of the  $\infty$ -parent ancestral process in a specific direction. In particular, this means that we cannot use the results from [10] to couple Deijfen's model to the  $\infty$ -parent ancestral process and deduce directly that its growth is at least linear in time.

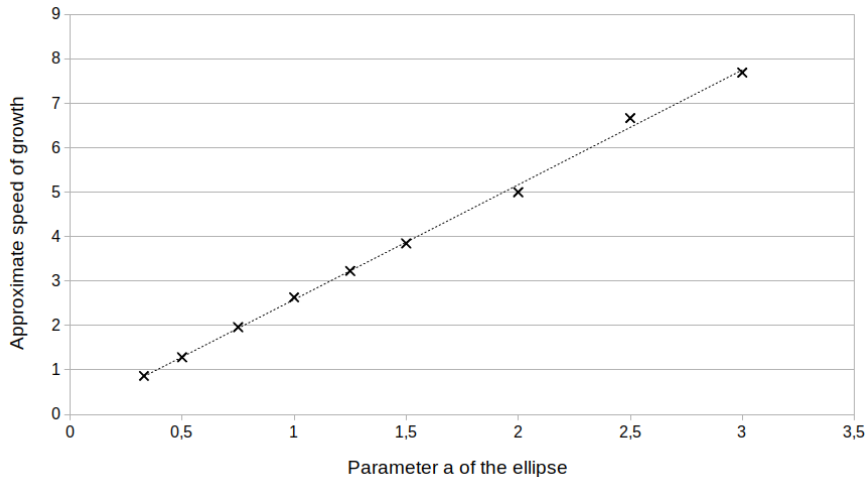


Figure 3: Approximate speed of growth of the occupied area in the  $\infty$ -parent SLFV process, as a function of  $a$ . For each value of  $a$ , 30  $\infty$ -parent ancestral processes with parameters  $(a, b, 0)$  were simulated, in order to compute  $\mathbf{E}[\overleftarrow{\tau}_x]$  for large values of  $x$ . The crosses indicate the approximate values, and the dotted line corresponds to the linear function  $\nu^{-1}(a) = 2.58a$ .

## 2.4 Numerical simulations

In order to obtain an approximation for the limiting speed of growth  $\nu^{-1}$  by means of numerical simulations, we can use the fact that

$$\nu^{-1} = \left( \lim_{x \rightarrow +\infty} x^{-1} \mathbf{E}[\overleftarrow{\tau}_x] \right)^{-1},$$

which is a consequence of Theorem 2.11 combined with Proposition 2.20. Indeed, the  $\infty$ -parent ancestral is easier to simulate than its forwards-in-time counterpart, since it jumps at a finite rate at any time, and since there are no border effects to take into account while simulating the process on an appropriately chosen compact subset of  $\mathbb{R}^2$ .

We focus on the case in which all ellipses have the same shape parameters. In order to be able to compare the speed of growth of the occupied regions in  $\infty$ -parent SLFV with different shape parameters, we assume that any given location  $z \in \mathbb{R}^2$  is affected by a reproduction event at rate 1. Therefore, we take

$$\tilde{\mu}(da, db, d\gamma) = V_{a_0, b_0, 0}^{-1} \delta_{a_0}(da) \otimes \delta_{b_0}(db) \otimes \delta_0(d\gamma),$$

where  $a_0 \in (0, +\infty)$  and  $b_0$  is chosen such that the volume of the corresponding ellipse is equal to  $\pi$ .

For 9 different values of  $a$  ranging from 0.33 to 3, we simulate 30  $\infty$ -parent ancestral processes with initial condition  $\{(0, 0)\}$  and parameters  $(a, b, 0)$ , where  $b$  is chosen as stated above. This ensures that we compare  $\infty$ -parent SLFV processes for which reproduction events have the same scale.

The results can be found in Figure 3. The numerical simulations show that the speed of growth is a linear function of  $a$  (over the range of  $a$ -values considered). This speed is around 2.6 times higher



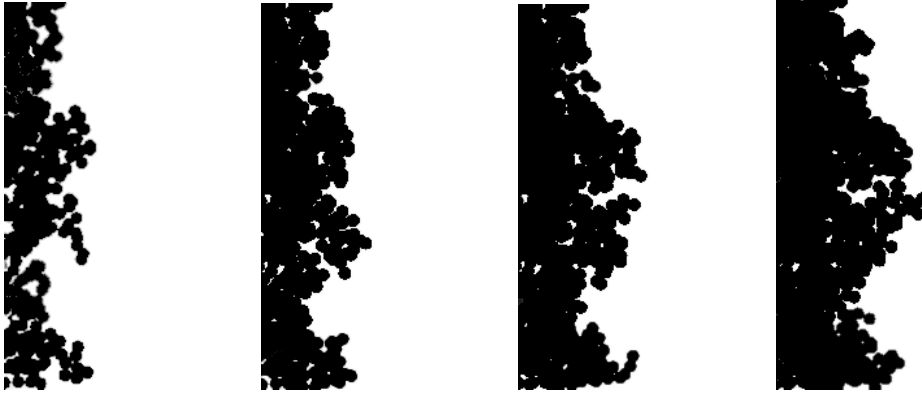


Figure 4: Illustration of the growth dynamics of the occupied region in the  $\infty$ -parent SLFV, with  $(a, b, \gamma) = (1, 1, 0)$ . The images represent the same  $\infty$ -parent SLFV at four instants  $0 < t_1 < t_2 < t_3 < t_4$ . The black area represents the area occupied by real individuals, and the white area is empty (or equivalently, filled with ghost individuals). The expansion starts from the left-most part of the image, and goes towards the right-most part of the image.

than the lower bound obtained in Section 3 (equal to  $a$  when  $\gamma = 0$ ), which was initially conjectured to be the limiting speed of the process.

Numerical simulations suggest that the growth of the process seems to be driven by "spikes" in the expansion direction, that then thicken and grow sideways, bridging the gap with the rest of the population. See Figure 4 for an illustration of this phenomenon. This motivates the study of a toy model of two interacting growing piles of cubes in Section 5.

### 3 Lower bound on the speed of growth

In this section, we show that the growth of the occupied area in the  $\infty$ -parent SLFV, as well as that of its dual, is *at least* linear in time. More precisely, we show the following result.

**Proposition 3.1.** *There exists  $\nu \geq 0$  such that*

$$\lim_{x \rightarrow +\infty} \frac{\mathbf{E}[\overleftarrow{\tau}_x]}{x} = \nu.$$

The proof of this result can be found at the end of Section 3.3. By Proposition 2.20, this result is then also true for  $\overrightarrow{\tau}_x$ .

The key difference with Theorem 2.11 is that Proposition 3.1 does not require  $\nu$  to be non-zero, and hence only means that the growth is at least linear in time (and faster if  $\nu = 0$ ). In Section 4, we show that indeed we have  $\nu \neq 0$ , or in other words, that the growth is exactly linear in time.

*Remark 3.2.* Since the limiting speed of growth is given by  $\nu^{-1}$ , a lower bound on the speed of growth amounts to an *upper bound* on  $\lim_{x \rightarrow +\infty} x^{-1} \mathbf{E}[\overleftarrow{\tau}_x]$ .

In order to prove Proposition 3.1, we first establish a few auxiliary results. In all that follows, let

$\bar{\Pi}$  be a Poisson point process on  $\mathbb{R}_+ \times \mathbb{R}^2 \times S_{\bar{\mu}}$  with intensity measure  $dt \otimes dz \otimes \tilde{\mu}^{\leftarrow}$ , and let  $(E_t)_{t \geq 0}$  be the  $\infty$ -parent ancestral process with initial condition  $\{(0, 0)\}$  and associated to  $\tilde{\mu}^{\leftarrow}$ , constructed using the Poisson point process  $\bar{\Pi}$ . Here we use the modification of the  $\infty$ -parent ancestral process introduced earlier, which allows to take single points as initial conditions.

### 3.1 Sub-additivity

We first introduce other  $\infty$ -parent ancestral processes, coupled to  $(E_t)_{t \geq 0}$  via the Poisson point process  $\bar{\Pi}$ . For all  $z \in \mathbb{R}^2$  and  $s \geq 0$ , let  $(E_t^{z,s})_{t \geq 0}$  be the  $\infty$ -parent ancestral process with initial condition  $\{z\}$  associated to  $\tilde{\mu}^{\leftarrow}$ , constructed using *only the reproduction events in  $\Pi$  occurring strictly after time  $s$* . If  $s = 0$ , then  $(E_t^{z,0})_{t \geq 0}$  is the regular  $\infty$ -parent ancestral process, but if  $s \neq 0$ , then the process is constant and equal to  $\{z\}$  during the time interval  $[0, s]$ , and only after does it start following the dynamics of an  $\infty$ -parent ancestral process.

Since all the processes  $(E_t^{z,s})_{t \geq 0}$ ,  $z \in \mathbb{R}^2$ ,  $s \geq 0$  are constructed using the same underlying Poisson point process, we have the following result.

**Lemma 3.3.** *For all  $z_1, z_2 \in \mathbb{R}^2$  and  $0 < s_1 < s_2$ , if  $z_2 \in E_{s_2}^{z_1, s_1}$ , then a.s. for all  $t \geq s_2$ , we have*

$$E_t^{z_2, s_2} \subseteq E_t^{z_1, s_1}.$$

We now introduce the following family of random variables. First, for all  $n \in \mathbb{N}$ , let

$$T_{0,n} := \min \left\{ t \geq 0 : E_t^{(0,0),0} \cap HP^{4n\mathcal{R}_{\bar{\mu}}} \neq \emptyset \right\} = \min \left\{ t \geq 0 : E_t \cap HP^{4n\mathcal{R}_{\bar{\mu}}} \neq \emptyset \right\} = \overleftarrow{\tau}_{4n\mathcal{R}_{\bar{\mu}}},$$

and let  $P_n$  be sampled uniformly at random among the points in  $E_{T_{0,n}}$  at horizontal separation of *exactly*  $4n\mathcal{R}_{\bar{\mu}}$  from  $(0, 0)$ . That is,  $P_n$  is a uniform sample from the compact set

$$E_{T_{0,n}} \cap \left\{ (x, y) \in \mathbb{R}^2 : x = 4n\mathcal{R}_{\bar{\mu}} \right\}.$$

Moreover, for all  $0 \leq m \leq n \in \mathbb{N}$ , let

$$T_{m,n} := \min \left\{ t \geq 0 : E_{t+T_{0,m}}^{P_m, T_{0,m}} \cap HP^{4n\mathcal{R}_{\bar{\mu}}} \neq \emptyset \right\},$$

where  $E_{t+T_{0,m}}^{P_m, T_{0,m}}$  corresponds to the  $\infty$ -parent ancestral process started from  $P_m$  at time  $T_{0,m}$ .

By construction, the family  $(T_{m,n})_{0 \leq m \leq n}$  satisfies the following lemma.

**Lemma 3.4.** *For all  $0 < m < n \in \mathbb{N}$ , we have*

$$T_{0,n} \leq T_{0,m} + T_{m,n}.$$

*Proof.* Let  $0 < m < n \in \mathbb{N}$ . By definition,

$$P_m \in E_{T_{0,m}} = E_{T_{0,m}}^{(0,0),0}.$$

Therefore, using Lemma 3.3, we have a.s. for all  $t \geq 0$

$$E_{t+T_{0,m}}^{P_m, T_{0,m}} \subseteq E_{t+T_{0,m}}.$$

In particular, this is true for  $t = T_{m,n}$ , hence

$$E_{T_{m,n}+T_{0,m}}^{P_m, T_{0,m}} \cap HP^{4n\mathcal{R}_{\bar{\mu}}} \subseteq E_{T_{m,n}+T_{0,m}} \cap HP^{4n\mathcal{R}_{\bar{\mu}}},$$

from which we deduce (by definition of  $T_{0,m}$ ,  $P_m$  and  $T_{m,n}$ ) that

$$E_{T_{m,n}+T_{0,m}} \cap HP^{4n\mathcal{R}_{\bar{\mu}}} \neq \emptyset.$$

Therefore,

$$T_{0,n} = \min \left\{ t > 0 : E_t \cap HP^{4n\mathcal{R}_{\bar{\mu}}} \neq \emptyset \right\} \leq T_{m,n} + T_{0,m}.$$

□

By invariance by translation of the underlying Poisson point processes, the following lemmas also hold true.

**Lemma 3.5.** *For all  $n \in \mathbb{N}$ , the joint distribution of  $(T_{n+1, n+k+1})_{k \geq 1}$  is the same as the one of  $(T_{n, n+k})_{k \geq 1}$ .*

**Lemma 3.6.** *For all  $k \in \mathbb{N} \setminus \{0\}$ , the random variables  $(T_{nk, (n+1)k})_{n \geq 1}$  are i.i.d.*

In order to use Theorem 1.10 from [31] and conclude, all we need is to show that the following lemma is true.

**Lemma 3.7.** *For all  $n \in \mathbb{N}$ , we have  $\mathbf{E}[T_{0,n}] < +\infty$ .*

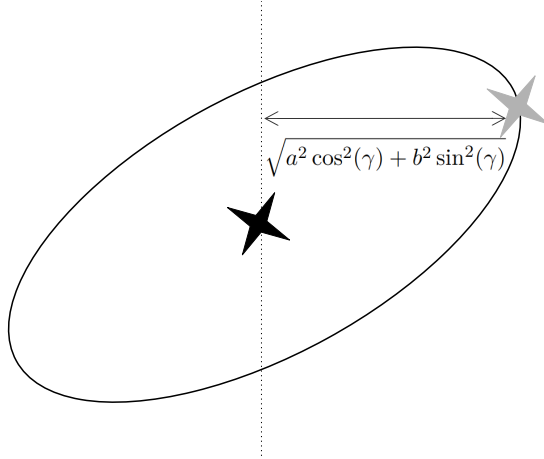


Figure 5: Ellipse with parameters  $(a, b, \gamma)$ .

Lemma 3.7 is proved at the end of Section 3.3, using the so-called *express chain* that we now define. We shall then use it to complete the proof of Proposition 3.1 (again at the end of Section 3.3).

### 3.2 Definition of the express chain

The use of what we shall call the *express chain* can be motivated by the following observation. In each ellipse with centre  $z = (x, y) \in \mathbb{R}^2$  and parameters  $(a, b, \gamma)$ , there exists exactly one point for which the horizontal separation from the centre is maximal: the point with coordinates

$$(x + a \cos(\theta_{max}) \cos(\gamma) - b \sin(\theta_{max}) \sin(\gamma), y + a \cos(\theta_{max}) \sin(\gamma) + b \sin(\theta_{max}) \cos(\gamma)),$$

where  $\theta_{max} = \arctan(-ba^{-1} \tan(\gamma))$ . This point is at horizontal separation  $\sqrt{a^2 \cos^2(\gamma) + b^2 \sin^2(\gamma)}$  from  $z$  (see Figure 5).

If we take this potential parent, wait until it is affected by a new reproduction event, and repeat, we obtain a Markov jump process, jumping at rate

$$J_{\tilde{\mu}} := \int_{S_{\tilde{\mu}}} V_{a,b,\gamma} \tilde{\mu}(da, db, d\gamma), \quad (3.1)$$

and going away from 0 at an average *horizontal speed* of  $J_{\tilde{\mu}} \mathbf{E} \left[ \sqrt{a^2 \cos^2(\gamma) + b^2 \sin^2(\gamma)} \right]$  (modulo some stochasticity due to the location of the centre of the reproduction event). See Appendix A for the proof of the geometrical properties of ellipses used throughout this section.

Formally, the express chain is defined as follows.

**Definition 3.8.** *The express chain associated to  $(E_t)_{t \geq 0}$  (constructed using  $\bar{\Pi}$ ), denoted  $(C_t^{express})_{t \geq 0}$ , is the  $\mathbb{R}^2$ -valued Markov process defined as follows.*

*First, we set  $C_0^{express} = (0, 0)$ . Then, for all  $(t, z_c, a, b, \gamma) \in \Pi$ , if  $C_{t-}^{express} \in \mathfrak{B}_{a,b,\gamma}(z_c)$  and*



Figure 6: Comparison between the express chain and the  $\infty$ -parent ancestral process at time  $t$ , for two realisations of the  $\infty$ -parent ancestral process. The ellipses indicate the reproduction events which affected the  $\infty$ -parent ancestral process until time  $t$ . The crosses represent the successive positions of the express chain, and the black point indicates the location reached by the  $\infty$ -parent ancestral process at time  $t$  which maximises the horizontal separation from the starting location. This location coincides with the one reached by the express chain in the first case, but not in the second case.

$z_c = (x_c, y_c)$ , we set:

$$C_t^{express} = (X_t^{express}, Y_t^{express})$$

where

$$X_t^{express} = x_c + a \cos(\theta_{max}) \cos(\gamma) - b \sin(\theta_{max}) \sin(\gamma),$$

$$Y_t^{express} = y_c + a \cos(\theta_{max}) \sin(\gamma) + b \sin(\theta_{max}) \cos(\gamma)$$

and  $\theta_{max} = \arctan(-ba^{-1} \tan(\gamma))$ .

See Figure 6 for an illustration of how to construct the express chain.

The interest of the express chain lies in the following observation, whose proof is a direct consequence of the definition of  $T_{0,n}$ .

**Lemma 3.9.** *Let  $n \in \mathbb{N}^*$ . For all  $t \geq 0$ , we have  $\{C_t^{express} \in HP^{4n\mathcal{R}_{\bar{\mu}}}\} \subset \{T_{0,n} \leq t\}$ .*

*Proof.* Let  $t \geq 0$ . Since  $C_t^{express} \in E_t$  by construction, the fact the  $C_t^{express} \in HP^{4n\mathcal{R}_{\bar{\mu}}}$  implies that  $E_t \cap HP^{4n\mathcal{R}_{\bar{\mu}}} \neq \emptyset$ , and thus that  $T_{0,n} \leq t$ .  $\square$

Therefore, if for all  $n \in \mathbb{N}$ , we set

$$T_{0,n}^{express} := \min \left\{ t \geq 0 : C_t^{express} \in HP^{4n\mathcal{R}_{\bar{\mu}}} \right\},$$

then for all  $n \in \mathbb{N}$  we have

$$T_{0,n} \leq T_{0,n}^{express} \quad \text{a.s.}, \quad \text{and so} \quad \mathbf{E}[T_{0,n}] \leq \mathbf{E}[T_{0,n}^{express}]. \quad (3.2)$$

In other words, in order to show Lemma 3.7, it is sufficient to obtain a similar result on  $\mathbf{E}[T_{0,n}^{express}]$ .

Before studying the properties of the express chain in the next section, we introduce the following notation. For all  $t \geq 0$ , let  $N_t^{express}$  be the number of jumps of the express chain in the time interval  $[0, t]$ . For all  $i \in \mathbb{N} \setminus \{0\}$ , let  $t_i$  be the instant of the  $i$ -th jump of the express chain, let  $R_i = (R_i^X, R_i^Y)$  be the coordinates of the centre of the reproduction event triggering this jump, and let  $(a_i, b_i, \gamma_i)$  be the parameters of the ellipse affected by the reproduction event. We then set for all  $i \in \mathbb{N} \setminus \{0\}$

$$D_i = \sqrt{a_i^2 \cos^2(\gamma_i) + b_i^2 \sin^2(\gamma_i)}.$$

The random variable  $D_i$  thus encodes the distance between the centre of the reproduction event and the right-most point in the corresponding ellipse. Note that the random variables  $(D_i)_{i \geq 1}$  are i.i.d.

### 3.3 Properties of the express chain

We now study some properties of the express chain, in order to obtain an upper bound on  $\mathbf{E}[T_{0,n}^{express}]$ .

Let  $t \geq 0$ . By construction,

$$\begin{aligned} X_t^{express} &= \sum_{i=1}^{N_t^{express}} \left( R_i^X - X_{t_i^-}^{express} + D_i \right) \\ &= \sum_{i=1}^{N_t^{express}} \left( R_i^X - X_{t_i^-}^{express} + D_i - \mathbf{E}[D_1] \right) + \mathbf{E}[D_1] N_t^{express}. \end{aligned} \quad (3.3)$$

The random variables  $(R_i^X - X_{t_i^-}^{express} + D_i - \mathbf{E}[D_1])_{i \geq 1}$  are i.i.d, bounded and with mean 0 (see Appendix A). We can then apply Hoeffding's inequality [25] and obtain the following lemma.

**Lemma 3.10.** *There exists  $C_1 > 0$  such that for all  $t \geq 0$  and  $n \in \mathbb{N} \setminus \{0\}$ , for all  $k > 4n\mathcal{R}_{\bar{\mu}}\mathbf{E}[D_1]^{-1}$ , we have*

$$\mathbf{P} \left( X_t^{express} < 4n\mathcal{R}_{\bar{\mu}} \mid N_t^{express} = k \right) \leq \exp \left( -\frac{(\mathbf{E}[D_1]k - 4n\mathcal{R}_{\bar{\mu}})^2}{C_1 k} \right).$$

Consequently, for all  $C' > 4n\mathcal{R}_{\bar{\mu}}\mathbf{E}[D_1]^{-1}$ ,

$$\mathbf{P} \left( X_t^{express} < 4n\mathcal{R}_{\bar{\mu}} \mid N_t^{express} > C' \right) \leq \exp \left( -\frac{(C'\mathbf{E}[D_1] - 4n\mathcal{R}_{\bar{\mu}})^2}{C_1 C'} \right).$$

*Proof.* First, notice that the second part of the lemma is a direct consequence of the first part, along with the variations of the function  $x \rightarrow \exp(-(xc - d)^2 x^{-1})$ ,  $c, d > 0$ . As concerns the first part of the

lemma, let  $t \geq 0$ ,  $n \in \mathbb{N} \setminus \{0\}$  and  $k > 4n\mathcal{R}_{\bar{\mu}}\mathbf{E}[D_1]^{-1}$ . Using (3.3), we have

$$\begin{aligned} & \mathbf{P}\left(X_t^{express} < 4n\mathcal{R}_{\bar{\mu}} \mid N_t^{express} = k\right) \\ & \leq \mathbf{P}\left(\sum_{i=1}^k (R_i^X - X_{t_i^-}^{express} + D_i - \mathbf{E}[D_1]) < 4n\mathcal{R}_{\bar{\mu}} - \mathbf{E}[D_1]k \mid N_t^{express} = k\right) \\ & = \mathbf{P}\left(\sum_{i=1}^k (-R_i^X + X_{t_i^-}^{express} - D_i + \mathbf{E}[D_1]) > \mathbf{E}[D_1]k - 4n\mathcal{R}_{\bar{\mu}} \mid N_t^{express} = k\right). \end{aligned}$$

We conclude by using Hoeffding's inequality together with the fact that for all  $i \geq 1$ ,

$$\left|R_i^X - X_{t_i^-}^{express}\right| < 4\mathcal{R}_{\bar{\mu}} \quad \text{and} \quad |D_i - \mathbf{E}[D_1]| \leq 2\mathcal{R}_{\bar{\mu}}.$$

□

In order to bound  $\mathbf{E}[T_{0,n}^{express}]$ , we also need to control the number of jumps made by the express chain over the time interval  $[0, t]$ . Recall the definition of  $J_{\bar{\mu}}$  given in (3.1).

**Lemma 3.11.** *There exists  $C_{\otimes} > 0$  such that for all  $t > 0$ ,*

$$\mathbf{P}(N_t^{express} \leq 0.1J_{\bar{\mu}}t) \leq \exp(-C_{\otimes}J_{\bar{\mu}}t).$$

*Proof.* The proof relies on the fact that the express chain jumps at rate  $J_{\bar{\mu}}$ . Hence,  $N_t^{express}$  is Poisson distributed with parameter  $J_{\bar{\mu}}t$ .

Let  $t > 0$ . Using a standard Chernoff bound, we obtain

$$\begin{aligned} \mathbf{P}(N_t^{express} \leq 0.1J_{\bar{\mu}}t) & \leq \frac{(eJ_{\bar{\mu}}t)^{0.1J_{\bar{\mu}}t} e^{-J_{\bar{\mu}}t}}{(0.1J_{\bar{\mu}}t)^{0.1J_{\bar{\mu}}t}} \\ & = \frac{\exp(0.1J_{\bar{\mu}}t) \exp(-J_{\bar{\mu}}t)}{\exp(0.1J_{\bar{\mu}}t \ln(0.1))} \\ & = \exp(0.1J_{\bar{\mu}}t - J_{\bar{\mu}}t - 0.1J_{\bar{\mu}} \ln(0.1)), \end{aligned}$$

and so taking  $C_{\otimes} = J_{\bar{\mu}}(0.9 - 0.1 \ln(10))$  allows us to conclude. □

Combining Lemmas 3.10 and 3.11, we obtain an upper bound for  $\mathbf{E}[T_{0,n}^{express}]$ .

**Lemma 3.12.** *There exists  $C_2 > 0$  such that for all  $n \in \mathbb{N}$ ,*

$$\mathbf{E}[T_{0,n}^{express}] \leq C_2 n.$$

*Proof.* Let  $n \in \mathbb{N}$ . Then,

$$\begin{aligned}
\mathbf{E}[T_{0,n}^{express}] &= \int_0^\infty \mathbf{P}(T_{0,n}^{express} > t) dt \\
&= \int_0^t \mathbf{P}(X_t^{express} < 4n\mathcal{R}_{\bar{\mu}}) dt \\
&\leq \frac{100n\mathcal{R}_{\bar{\mu}}}{J_{\bar{\mu}}\mathbf{E}[D_1]} + \int_{\frac{100n\mathcal{R}_{\bar{\mu}}}{J_{\bar{\mu}}\mathbf{E}[D_1]}}^\infty \mathbf{P}(X_t^{express} < 4n\mathcal{R}_{\bar{\mu}}) dt \\
&\leq \frac{100n\mathcal{R}_{\bar{\mu}}}{J_{\bar{\mu}}\mathbf{E}[D_1]} + \int_{\frac{100n\mathcal{R}_{\bar{\mu}}}{J_{\bar{\mu}}\mathbf{E}[D_1]}}^\infty \mathbf{P}\left(N_t^{express} \leq \frac{J_{\bar{\mu}}t}{10}\right) dt \\
&\quad + \int_{\frac{100n\mathcal{R}_{\bar{\mu}}}{J_{\bar{\mu}}\mathbf{E}[D_1]}}^\infty \mathbf{P}\left(\{X_t^{express} < 4n\mathcal{R}_{\bar{\mu}}\} \cap \left\{N_t^{express} > \frac{J_{\bar{\mu}}t}{10}\right\}\right) dt \\
&\leq \frac{100n\mathcal{R}_{\bar{\mu}}}{J_{\bar{\mu}}\mathbf{E}[D_1]} + \int_{\frac{100n\mathcal{R}_{\bar{\mu}}}{J_{\bar{\mu}}\mathbf{E}[D_1]}}^\infty \exp(-C_\otimes J_{\bar{\mu}}t) dt \\
&\quad + \int_{\frac{100n\mathcal{R}_{\bar{\mu}}}{J_{\bar{\mu}}\mathbf{E}[D_1]}}^\infty \exp\left(-\frac{10}{C_1 J_{\bar{\mu}}t} \left(\frac{J_{\bar{\mu}}t\mathbf{E}[D_1]}{10} - 4n\mathcal{R}_{\bar{\mu}}\right)^2\right) dt.
\end{aligned} \tag{3.4}$$

Here we have used Lemmas 3.10 and 3.11 to pass from the fourth to the fifth line. Moreover, we have

$$\begin{aligned}
&\int_{\frac{100n\mathcal{R}_{\bar{\mu}}}{J_{\bar{\mu}}\mathbf{E}[D_1]}}^\infty \exp\left(-\frac{10}{C_1 J_{\bar{\mu}}t} \left(\frac{J_{\bar{\mu}}t\mathbf{E}[D_1]}{10} - 4n\mathcal{R}_{\bar{\mu}}\right)^2\right) dt \\
&= \exp\left(\frac{2\mathbf{E}[D_1]4n\mathcal{R}_{\bar{\mu}}}{C_1}\right) \int_{\frac{100n\mathcal{R}_{\bar{\mu}}}{J_{\bar{\mu}}\mathbf{E}[D_1]}}^\infty \exp\left(-\frac{10}{C_1 J_{\bar{\mu}}t} \frac{t^2 J_{\bar{\mu}}^2 \mathbf{E}[D_1]^2}{100}\right) \exp\left(-\frac{10}{C_1 J_{\bar{\mu}}t} 16n^2 \mathcal{R}_{\bar{\mu}}^2\right) dt \\
&\leq \exp\left(\frac{8n}{C_1} \mathcal{R}_{\bar{\mu}} \mathbf{E}[D_1]\right) \int_{\frac{100n\mathcal{R}_{\bar{\mu}}}{J_{\bar{\mu}}\mathbf{E}[D_1]}}^\infty \exp\left(-\frac{J_{\bar{\mu}}t}{10C_1} \mathbf{E}[D_1]^2\right) dt \\
&\leq \exp\left(\frac{8n}{C_1} \mathcal{R}_{\bar{\mu}} \mathbf{E}[D_1]\right) \frac{10C_1}{J_{\bar{\mu}}\mathbf{E}[D_1]^2} \exp\left(-\frac{J_{\bar{\mu}}\mathbf{E}[D_1]^2}{10C_1} \frac{100n\mathcal{R}_{\bar{\mu}}}{J_{\bar{\mu}}\mathbf{E}[D_1]}\right) \\
&\leq \frac{10C_1}{J_{\bar{\mu}}\mathbf{E}[D_1]^2} \exp\left(-\frac{2n}{C_1} \mathcal{R}_{\bar{\mu}} \mathbf{E}[D_1]\right).
\end{aligned}$$

Since the first term on the r.h.s of (3.4) is proportional to  $n$ , and the second and third terms decrease exponentially fast in  $n$ , we can conclude.  $\square$

We can now conclude the proof of Lemma 3.7.

*Proof.* (Lemma 3.7) Let  $n \in \mathbb{N}$ . By Lemma 3.12,

$$\mathbf{E}[T_{0,n}^{express}] < +\infty,$$

and so by (3.2),

$$\mathbf{E}[T_{0,n}] < +\infty.$$

$\square$



We conclude this section with the proof of Proposition 3.1.

*Proof.* (Proposition 3.1) Combining Lemmas 3.12, 3.4, 3.5 and 3.6 with the fact that  $T_{0,n} \geq 0$  for all  $n$ , we obtain that the family  $(T_{m,n})_{0 \leq m \leq n}$  satisfies all the assumptions of Theorem 1.10 from [31]. Therefore, there exists  $\nu' \geq 0$  such that

$$\lim_{n \rightarrow +\infty} \frac{\mathbf{E}[T_{0,n}]}{n} = \frac{\mathbf{E}[\overleftarrow{\tau}_{4n\mathcal{R}_{\tilde{\mu}}}]}{n} = \nu'.$$

We conclude by standard upper and lower bounding arguments, using the fact that  $x \rightarrow \mathbf{E}[\overleftarrow{\tau}_x]$  is a nondecreasing function (which is a consequence of Lemma 2.18).  $\square$

*Remark 3.13.* If we see  $(X_t^{express})_{t \geq 0}$  as a cumulative process (in the sense of Chapter IV.3 in [3]), it is possible to use Theorem 3.1 from this Chapter and show that the limiting horizontal speed of advance of the express chain is equal to  $J_{\tilde{\mu}} \mathbf{E}[D_1]$ , yielding an explicit lower bound on the speed of growth of the occupied area in the  $\infty$ -parent SLFV. In particular, if all reproduction ellipses have the same shape parameters  $(a, b, \gamma)$  and if the total mass of  $\tilde{\mu}$  is  $V_{a,b,\gamma}^{-1}$ , which corresponds to the case investigated using numerical simulations, then the lower bound on the speed of growth is given by  $\sqrt{a^2 \cos^2(\gamma) + b^2 \sin^2(\gamma)}$ .

## 4 Upper bound on the speed of growth

Recall that  $(E_t)_{t \geq 0}$  is the  $\infty$ -parent ancestral process with initial condition  $\{(0, 0)\}$  associated to  $\tilde{\mu}^{\leftarrow}$ , constructed using  $\bar{\Pi}$ .

In this section, we complete the result proven in Section 3 by showing that the growth of the  $\infty$ -parent SLFV and of its dual counterpart are *at most* linear in time. This can be rewritten as a limiting property of  $\mathbf{E}[\overleftarrow{\tau}_x]$  as follows.

**Proposition 4.1.**

$$\lim_{x \rightarrow +\infty} \frac{\mathbf{E}[\overleftarrow{\tau}_x]}{x} > 0.$$

The proof can be found at the end of Section 4.3. Combining this result with Proposition 3.1 gives us the result of Theorem 2.11.

To show the result of Proposition 4.1, we first observe that it is sufficient to focus on the case in which all reproduction events are balls of fixed radius. Indeed, we have the following comparison result.

**Lemma 4.2.** Let  $\bar{\Pi}^{\mathcal{R}_{\tilde{\mu}}}$  be a Poisson point process on  $\mathbb{R}_+ \times \mathbb{R}^2 \times S_{\tilde{\mu}}$  with intensity measure

$$\tilde{\mu}^{\leftarrow}(S_{\tilde{\mu}})dt \otimes dz \otimes \delta_{\mathcal{R}_{\tilde{\mu}}}(da) \otimes \delta_{\mathcal{R}_{\tilde{\mu}}}(db) \otimes \delta_0(\gamma).$$

Let  $(E_t^{\mathcal{R}_{\tilde{\mu}}})_{t \geq 0}$  be the  $\infty$ -parent ancestral process with initial condition  $\{(0, 0)\}$  constructed using  $\bar{\Pi}^{\mathcal{R}_{\tilde{\mu}}}$ , and for all  $x > 0$ , let  $\overleftarrow{\tau}_x^{\mathcal{R}_{\tilde{\mu}}}$  be the first time  $(E_t^{\mathcal{R}_{\tilde{\mu}}})_{t \geq 0}$  reaches  $HP^x$ , defined as in Definition 2.17. Then, for all  $x > 0$ ,

$$\mathbf{E}[\overleftarrow{\tau}_x] \geq \mathbf{E}[\overleftarrow{\tau}_x^{\mathcal{R}_{\tilde{\mu}}}] .$$

*Proof.* The proof relies on the following coupling between  $(E_t)_{t \geq 0}$  and  $(E_t^{\mathcal{R}_{\tilde{\mu}}})_{t \geq 0}$ . Instead of being independent from  $\bar{\Pi}$ , the Poisson point process  $\bar{\Pi}^{\mathcal{R}_{\tilde{\mu}}}$  is constructed using the points from  $\bar{\Pi}$ , as follows:

$$\bar{\Pi}^{\mathcal{R}_{\tilde{\mu}}} := \{(t, z, \mathcal{R}_{\tilde{\mu}}, \mathcal{R}_{\tilde{\mu}}, 0) : (t, z, a, b, \gamma) \in \bar{\Pi}\}.$$

That is, we replace the elliptical area of each reproduction event in  $\bar{\Pi}$  by a ball of maximal radius  $\mathcal{R}_{\tilde{\mu}}$ . Since  $\mathfrak{B}_{a,b,\gamma}(z) \subseteq \mathcal{B}_{\mathcal{R}_{\tilde{\mu}}}(z)$ , this coupling ensures that

$$\forall t \geq 0, E_t \subseteq E_t^{\mathcal{R}_{\tilde{\mu}}} \text{ a.s.}$$

Therefore, for all  $x > 0$ ,

$$\min \{t \geq 0 : E_t \cap HP^x \neq \emptyset\} \geq \min \{t \geq 0 : E_t^{\mathcal{R}_{\tilde{\mu}}} \cap HP^x\} \text{ a.s.,}$$

which allows us to conclude.  $\square$

To alleviate the notation, we only provide the proof of Proposition 4.1 when  $\tilde{\mu}^{\leftarrow} = \pi^{-1}\delta_1 \otimes \delta_1 \otimes \delta_0$ , so that all reproduction events happen in balls of radius 1 and the rate at which a given point in space is overlapped by an event is equal to 1. The proof can be easily generalised to balls of arbitrary fixed radius and any intensity of events by a simple scaling of time and space. We can then use Lemma 4.2 to obtain the corresponding result for ellipses with bounded parameters.

#### 4.1 A first-passage percolation problem

We consider the graph  $\mathcal{G}$  on the vertex set  $\mathbb{Z}^2$ , in which  $(i, j)$  and  $(i', j')$  are connected by an edge if, and only if

$$(i', j') \in \{(i+1, j), (i-1, j), (i, j+1), (i, j-1), (i-1, j-1), (i-1, j+1), (i+1, j-1),$$

$(i + 1, j + 1)\}$ .

To each edge  $e$  of  $\mathcal{G}$ , we associate an independent exponential random variable

$$\mathcal{E}_e \sim \text{Exp}(16 \times \pi^{-1}).$$

$\mathcal{E}_e$  corresponds to the time needed to pass through the corresponding edge. Following standard terminology in first-passage percolation, we call it the *passage time* of the edge. The choice of the rate of the exponential distribution ensures we can later compare the growth of the  $\infty$ -parent ancestral process to the first-passage percolation problem we now introduce.

If  $\Gamma$  is a (potentially infinite) path formed by the edges  $e_1, \dots, e_n, \dots$ , then the passage time of the path  $\Gamma$  is defined as

$$\mathcal{E}_\Gamma = \sum_{e \in \Gamma} \mathcal{E}_e.$$

If  $z_1, z_2 \in \mathbb{Z}^2$ , we define the first-passage time  $\mathcal{E}_{z_1, z_2}$  from  $z_1$  to  $z_2$  (or equivalently, from  $z_2$  to  $z_1$  since  $\mathcal{G}$  is not oriented) as the minimum over the passage times of all the (finite) paths going from  $z_1$  to  $z_2$ .

We then define

$$\overleftarrow{\tau}_n^{fpp} := \min \left\{ t \geq 0 : \exists m \in \mathbb{Z}, \mathcal{E}_{(0,0), (n,m)} \leq t \right\}.$$

In other words,  $\overleftarrow{\tau}_n^{fpp}$  is the time needed to reach a point at horizontal separation  $n$  from the origin, starting from the origin.

The interest of this first-passage percolation problem lies in the fact that since the passage time of any given edge is a.s. strictly positive, generalising Theorem 6.7 from [37] to our lattice yields the following result.

**Lemma 4.3.** *We have*

$$\lim_{n \rightarrow +\infty} \frac{\mathbf{E}[\overleftarrow{\tau}_n^{fpp}]}{n} > 0.$$

In order to use this lemma and show Proposition 4.1, we need to be able to compare  $\overleftarrow{\tau}_x$  and  $\overleftarrow{\tau}_n^{fpp}$ . The main obstacle to this comparison lies in the fact that the  $\infty$ -parent ancestral process is continuous in space, while the first-passage percolation problem is defined on a graph. Therefore, we now introduce a way to "discretise" the  $\infty$ -parent ancestral process.

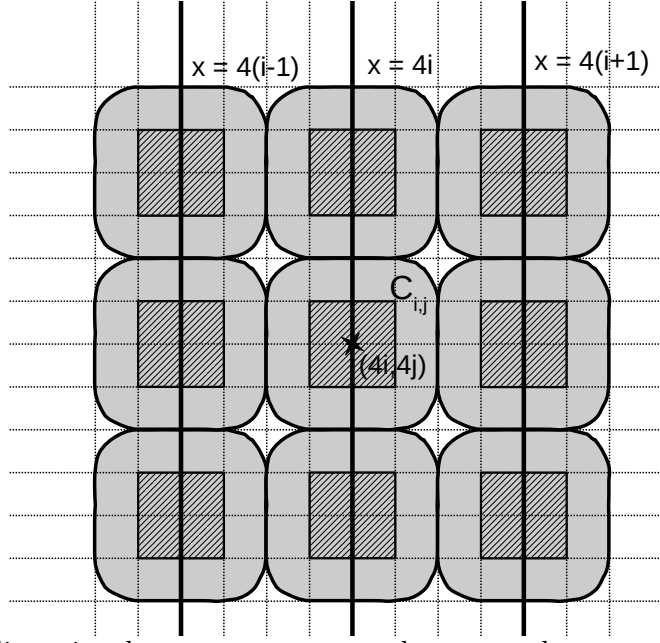


Figure 7: Grid used to discretise the  $\infty$ -parent ancestral process when reproduction events are balls. Each hatched square corresponds to a cell. The grey area around the site  $(4i, 4j)$ , indicated by a black cross, corresponds to the area in which centres of reproduction events have to fall to intersect the cell  $\mathcal{C}_{i,j}$ . The white areas contain all the centres of reproduction events which do not intersect any cell.

## 4.2 Discretisation of the $\infty$ -parent ancestral process

In order to discretise the  $\infty$ -parent ancestral process, we first place a grid on  $\mathbb{R}^2$ , and associate a cell to each site of the lattice. Let

$$\mathcal{V} := \{(4i, 4j) : (i, j) \in \mathbb{Z}^2\}$$

be the underlying grid, and for all  $(i, j) \in \mathbb{Z}^2$ , let  $\mathcal{C}_{i,j}$  be the square with centre  $(4i, 4j)$  and side length 2. Each  $\mathcal{C}_{i,j}$ ,  $i, j \in \mathbb{Z}$  corresponds to the cell associated to the site  $(4i, 4j)$  of the grid  $\mathcal{V}$ . See Figure 7 for an illustration.

This construction satisfies the two following key properties, that are consequences of the fact that all reproduction events are balls of radius 1.

1. For all  $z = (x, y) \in \mathbb{R}^2$ , unless  $x = 4i$ ,  $i \in \mathbb{Z}$  or  $y = 4j + 2$ ,  $j \in \mathbb{Z}$ , the ball  $\mathcal{B}_1(z)$  intersects at most one cell.

In other words, each reproduction events intersects at most one cell a.s. Moreover, if  $\mathcal{B}_1(z)$  does not intersect any cell, then it means that  $z$  fell into one of the areas in white on Figure 7.

2. If we refer to a sequence of reproduction events occurring in chronological order and such that each reproduction event intersects the previous one as a *path of reproduction events*, then any path of reproduction events for which none of the corresponding balls intersect a cell is almost surely confined in one of the white areas on Figure 7, in the sense that all the reproduction event centres fall into the same white area almost surely.

The discretised version of the  $\infty$ -parent ancestral process, denoted by  $(D_t)_{t \geq 0}$ , is then defined as follows. For all  $t \geq 0$ , let

$$D_t := \left\{ (i, j) \in \mathbb{Z}^2 : \mathcal{C}_{i,j} \cap E_t \neq \emptyset \right\}$$

be the set of all cells which intersect the  $\infty$ -parent ancestral process associated to  $\tilde{\mu}^{\leftarrow} = \pi^{-1} \delta_1 \otimes \delta_1 \otimes \delta_0$  (and with initial condition  $(0, 0)$ ) at time  $t$ . Let  $(\tau_n^{\text{discr}})_{n \geq 0}$  be the random variables defined for all  $n \geq 0$  by

$$\tau_n^{\text{discr}} := \min \{ t \geq 0 : \exists m \in \mathbb{Z}, (n, m) \in D_t \}.$$

Due to the structure of the grid and the size of the cells, we have the following result.

**Lemma 4.4.** *For all  $n \in \mathbb{N} \setminus \{0\}$ , we have*

$$\tau_n^{\text{discr}} \leq \tau_{4n} \quad \text{a.s.}$$

*Proof.* Let  $n \in \mathbb{N} \setminus \{0\}$ . Let  $z \in \mathbb{R}^2$  be the centre of the reproduction event which occurs at time  $\tau_{4n}^{\leftarrow}$  and makes the  $\infty$ -parent ancestral process reach for the first time a point at horizontal separation  $4n$  from the origin. Then, a.s. there exists  $j \in \mathbb{Z}$  such that

$$\mathcal{B}_1(z) \cap \mathcal{C}_{n,j} \neq \emptyset,$$

and hence  $\tau_n^{\text{discr}} \leq \tau_{4n}$ . □

In all that follows, for all  $(i, j), (i', j') \in \mathbb{Z}^2$ , we say that

- the cell  $\mathcal{C}_{i,j}$  is *active* at time  $t$  if  $(i, j) \in D_t$ ;
- the cell  $\mathcal{C}_{i,j}$  is *activated* at time  $t$  if  $(i, j) \in D_t$  and  $(i, j) \notin D_{t-}$ ;
- the cell  $\mathcal{C}_{i',j'}$  *activates*  $\mathcal{C}_{i,j}$  at time  $t$  if there exists  $s < t$  such that  $\mathcal{C}_{i',j'}$  is active at time  $s$ , and if there exists a path of reproduction events starting from  $\mathcal{C}_{i',j'}$  at time  $s$ , initially overlapping an area of  $\mathcal{C}_{i',j'}$  containing type 1 individuals, and reaching  $\mathcal{C}_{i,j}$  for the first time at time  $t$  while not intersecting any other cell in the time interval  $[s, t]$ .

Notice that under this terminology, a cell can activate another one which is already active. Moreover, with probability one the only cells that the cell  $\mathcal{C}_{i',j'}$  can activate are its nearest neighbours in the graph  $\mathcal{G}$ , that is, the cells  $\mathcal{C}_{i,j}$  such that

$$\sqrt{(i - i')^2 + (j - j')^2} \leq \sqrt{2}.$$

### 4.3 Comparison to the first-passage percolation problem

Let us now compare the growth of the discretised  $\infty$ -parent ancestral process  $(D_t)_{t \geq 0}$  to that of the first-passage percolation problem introduced earlier. Recall that

$$\tau_n^{fpp} := \min \left\{ t \geq 0 : \exists m \in \mathbb{Z}, \mathcal{E}_{(0,0),(n,m)} \leq t \right\}$$

is the time needed by the process associated to the first-passage percolation problem to reach a point at horizontal separation  $n$  from the origin, starting from the origin.

**Proposition 4.5.** *The random variable  $\tau_n^{discr}$  is stochastically bounded from below by  $\tau_n^{fpp}$ . That is, for all  $t \geq 0$ ,*

$$\mathbf{P} \left( \tau_n^{discr} \geq t \right) \geq \mathbf{P} \left( \tau_n^{fpp} \geq t \right).$$

*Proof.* First, notice that in the discretised  $\infty$ -parent ancestral process, a cell can be activated by several neighbouring cells at the same time, but a.s. an active cell can only activate neighbouring cells one after another. Conversely, in the first-passage percolation problem, the edges leading to a same cell are almost surely crossed at different times, but only the first time one of these edges is crossed matters. Therefore, it is sufficient to show that the time needed to cross an edge in the process associated to the first-passage percolation problem is stochastically bounded from above by the time needed for a cell to be activated once at least one of its neighbours is active.

Then, assume that at some time  $t$ , the cell  $\mathcal{C}_{i,j}$  was just activated, while one of its neighbours is still not active. Let  $e$  denote the edge between  $(i,j)$  and  $(i',j')$ . In order for  $\mathcal{C}_{i',j'}$  to be activated (not necessarily by  $\mathcal{C}_{i,j}$ ), there needs to be a reproduction events occurring after time  $t$  and intersecting  $\mathcal{C}_{i',j'}$ . Such reproduction events occur at a rate bounded from above by  $\pi^{-1}$  times the volume of the area of potential centres  $z$  such that  $\mathcal{B}_1(z) \cap \mathcal{C}_{i',j'} \neq \emptyset$ , which is itself bounded by  $16\pi^{-1}$ . Therefore, the time needed for  $\mathcal{C}_{i',j'}$  to be activated is stochastically bounded from below by  $\mathcal{E}_e$ , and we can conclude.  $\square$

*Remark 4.6.* Due to correlations between activations by neighbouring cells, it would be more difficult to construct a coupling with the first-passage percolation problem. However, the stochastic comparison we obtained is sufficient.

We can now show Proposition 4.1.

*Proof.* (Proposition 4.1) By Proposition 4.5 and Lemma 4.4, for all  $n \in \mathbb{N} \setminus \{0\}$ ,

$$\mathbf{E}[\tau_{4n}] \geq \mathbf{E}[\tau_n^{discr}] \geq \mathbf{E}[\tau_n^{fpp}].$$

By Lemma 4.3, we know that

$$\lim_{n \rightarrow +\infty} \frac{\mathbf{E}[\overleftarrow{\tau}_n^{fpp}]}{4n} > 0.$$

Since we know that  $\lim_{n \rightarrow +\infty} \mathbf{E}[\overleftarrow{\tau}_n] \times n^{-1}$  exists by Proposition 3.1, we obtain that

$$\lim_{n \rightarrow +\infty} \frac{\mathbf{E}[\overleftarrow{\tau}_{4n}]}{4n} > 0.$$

We conclude by using the fact that  $x \rightarrow \overleftarrow{\tau}_x$  is nondecreasing by Lemma 2.18. □

We conclude this section with the proof of Theorem 2.11.

*Proof.* (Theorem 2.11) By Proposition 3.1, we know that there exists  $\nu \geq 0$  such that

$$\lim_{x \rightarrow +\infty} \frac{\mathbf{E}[\overleftarrow{\tau}_x]}{x} = \nu.$$

Moreover, by Proposition 4.1, we know that

$$\lim_{x \rightarrow +\infty} \frac{\mathbf{E}[\overleftarrow{\tau}_x]}{x} > 0.$$

Therefore,  $\nu > 0$  and we conclude by using Proposition 2.20. □

## 5 The two column growth process

The numerical simulations performed in Section 2.4 show that the speed of growth of the occupied region in the  $\infty$ -parent SLFV is higher than the one conjectured using the express chain. Moreover, the results suggest that the growth of the process is driven by relatively unfrequent "spikes" than then thicken and grow sideways. In order to investigate how these spikes can significantly increase the growth speed, we introduce the following toy model. We consider two adjacent piles of cubes, thereafter referred to as the *left pile* and the *right pile*. Each cube has height 1. A cube is added on top of the left (*resp.*, right) pile at rate 1, independently from the other pile. Moreover, the piles can also grow sideways: if there is a cube at height  $h$  in the left (*resp.*, right) pile, and no cube at such height in the other pile, then a cube is added to the right (*resp.*, left) pile at height  $h$  at rate 1. In particular, if the left pile is one cube higher than the right (*resp.*, left) pile, then the total rate at which the height of the right (*resp.*, left) pile increases of 1 is equal to 2, as the growth can be due to a new cube falling on top of the right (*resp.*, left) pile as well as sideways growth of the left (*resp.*, right) pile. See Figure 8 for an illustration of the dynamics of the process.

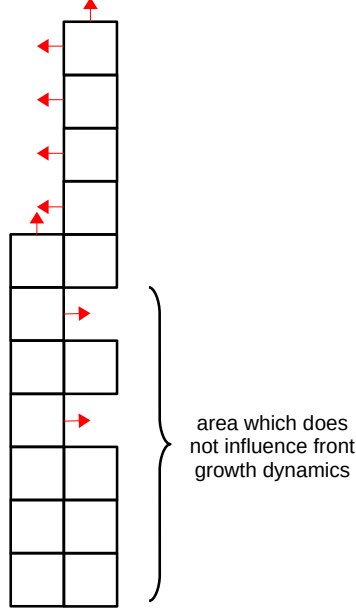


Figure 8: Growth dynamics of the two columns growth process. The red arrows indicate possible growth events occurring at rate 1.

*Remark 5.1.* The term "pile" can be slightly misleading, since it is possible to have holes in it, due to the other pile growing sideways. See Figure 8 for an illustration.

The interest of the process lies in the following observation. If  $l_t$  (*resp.*,  $r_t$ ) represents the maximal height reached by the left (*resp.*, right) pile, then the cubes at height  $h < \min(l_t, r_t)$  no longer contribute to the growth of the process. Therefore, we can "reset" the process whenever  $l_t = r_t$ .

We now define this process, called the *two columns growth process*, rigorously.

### 5.1 The two column growth process: Definition and properties

The state space over which the process is defined is the set  $\tilde{\mathcal{S}}$  defined by

$$\tilde{\mathcal{S}} := \{(i, j) \in \mathbb{N} \times \mathbb{N} : i \leq j\}.$$

If  $(i, j) \in \tilde{\mathcal{S}}$ , then  $i$  represents the height reached by the *lowest* pile, and  $j$  the height reached by the *highest* pile.

**Definition 5.2.** Let  $(i, j) \in \tilde{\mathcal{S}}$ . The two columns growth process  $(G_t)_{t \geq 0} = (m_t, M_t)_{t \geq 0}$  with initial condition  $(i, j)$ , thereafter referred to as the 2-CGP, is the continuous-time Markov chain on the state space  $\tilde{\mathcal{S}}$  whose transition rates are as follows. For all  $(i, j) \in \tilde{\mathcal{S}}$ ,

1. If  $i = j$ , then  $(i, j) \rightarrow (i, i + 1)$  at rate 2, and no other transitions are possible.



2. If  $j \geq i + 1$ , then

$$(i, j) \rightarrow \begin{cases} (i, j + 1) & \text{at rate 1,} \\ (i + 1, j) & \text{at rate 2,} \\ (i + k, j), k \in \llbracket 2, j - i \rrbracket & \text{at rate 1 if } j > i + 1, \end{cases}$$

and no other transitions are possible.

These transition rates exactly encode the dynamics described earlier. Indeed, if  $i = j$ , then both piles have the same height, and cubes fall onto each one of the piles at rate 1, yielding a total transition rate of 2 as we do not record which pile is the highest. If  $i \neq j$ , then each pile can grow upwards at rate 1, but the highest pile can also grow sideways. Depending on where the highest pile grows sideways, this results in a more or less sharp increase in the height of the lowest pile.

Earlier, we saw that it was possible to "reset" the process whenever the two piles had the same height. In order to do so, let  $\mathcal{S}_{\square}$  be the set

$$\mathcal{S}_{\square} := \{(i, i) : i \in \mathbb{N}\}$$

of all configurations such that both piles have the same height, and let  $T_{\square}$  be the time of first return of  $(G_t)_{t \geq 0}$  to a state belonging to  $\mathcal{S}_{\square}$ . Our first result is an upper bound on  $\mathbf{E}_{(0,0)} [T_{\square}]$ .

**Lemma 5.3.** *We have*

$$\mathbf{E}_{(0,0)} [T_{\square}] \leq 3/2.$$

*Proof.* In order to do so, we first notice that  $T_{\square}$  is equal to the sum of the time needed to exit  $(0, 0)$ , which follows an exponential distribution with parameter 2, and of the time needed to reach a state in  $\mathcal{S}_{\square}$  starting from  $(0, 1)$ .

The first step yields an expected contribution of  $1/2$  to  $\mathbf{E}_{(0,0)} [T_{\square}]$ . Once the process is in the state  $(0, 1)$ , assuming without loss of generality that the cube fell on the right pile (which is now the highest pile), we assign an exponential clock with parameter 1 to the highest cube of the right pile, which rings whenever the cube attempts to grow sideways. Whenever the highest cube changes, the exponential clock is assigned to the new highest cube. When the clock rings for the first time, we distinguish two cases:

- Either there is no cube at the same height in the left pile, and the cube can grow sideways. Then, we are back to a state in  $\mathcal{S}_{\square}$ , though perhaps not for the first time.
- Either there is already a cube at the same height in the left pile. Then, we know we already

went back to a state in  $\mathcal{S}_{\square}$ .

Therefore,

$$\mathbf{E}_{(0,0)}[T_{\square}] \leq 1/2 + 1,$$

allowing us to conclude.  $\square$

We can use this lemma to obtain an upper bound on  $\mathbf{E}_{(0,0)}[M_{T_{\square}}]$  and  $\mathbf{E}_{(0,0)}[m_{T_{\square}}]$ .

**Lemma 5.4.** *We have*

$$\mathbf{E}_{(0,0)}[M_{T_{\square}}] \leq 2 \quad \text{and} \quad \mathbf{E}_{(0,0)}[m_{T_{\square}}] \leq 2.$$

*Proof.* First, as  $G_{T_{\square}} \in \mathcal{S}_{\square}$ , we have  $M_{T_{\square}} = m_{T_{\square}}$ , and it is sufficient to provide an upper bound on  $\mathbf{E}_{(0,0)}[M_{T_{\square}}]$ . Let us set

$$\tilde{T} := \min \{t \geq 0 : M_t \neq m_t\},$$

which corresponds to the time needed for the process to leave the state  $(0,0)$ . For all  $t \in [\tilde{T}, T_{\square})$ , we have  $M_t > m_t$ , and the lowest pile does not contribute to the growth of the highest pile. Therefore, over the time interval  $[\tilde{T}, T_{\square})$ , the only way the highest pile can grow is by new cubes falling on top of it, whose total number is given by  $M_{T_{\square}} - M_{\tilde{T}} = M_{T_{\square}} - 1$ . Moreover, since the clocks corresponding to the upward growth of the highest pile are independent of the clocks responsible for the upward growth of the smallest pile and for the sidewise growth at the top of the highest pile, conditionally on  $T_{\square} - \tilde{T}$ , the variable  $M_{T_{\square}} - M_{\tilde{T}}$  follows a Poisson distribution with parameter  $T_{\square} - \tilde{T}$ . Therefore,

$$\begin{aligned} \mathbf{E}_{(0,0)}[M_{T_{\square}}] &= \mathbf{E}_{(0,0)}[M_{\tilde{T}}] + \mathbf{E}_{(0,0)}[M_{T_{\square}} - M_{\tilde{T}}] \\ &= 1 + \mathbf{E}_{(0,0)}[\mathbf{E}_{(0,0)}[M_{T_{\square}} - M_{\tilde{T}} | T_{\square} - \tilde{T}]] \\ &= 1 + \mathbf{E}_{(0,0)}[T_{\square} - \tilde{T}] \\ &= 1 - 1/2 + \mathbf{E}_{(0,0)}[T_{\square}] \\ &\leq 1/2 + 3/2 \end{aligned}$$

by Lemma 5.3, which allows us to conclude.  $\square$

In order to study the speed of growth of the process, we see the 2-CGP as a cumulative process, in the sense of Chapter VI in [3]. We obtain the following theoretical result.

**Theorem 5.5.** *We have*

$$\lim_{t \rightarrow +\infty} \frac{M_t}{t} = \lim_{t \rightarrow +\infty} \frac{m_t}{t} = \frac{\mathbf{E}_{(0,0)}[M_{T_{\square}}]}{\mathbf{E}_{(0,0)}[T_{\square}]} \quad a.s.$$

*Proof.* Let  $T_{\square}^0 = 0$ , and for all  $n \geq 1$ , let  $T_{\square}^n$  be the time of  $n$ -th return of  $(G_t)_{t \geq 0}$  to a state in  $\mathcal{S}_{\square}$ . For all  $t > 0$ , let  $N(t) = \max\{n \in \mathbb{N} : T_{\square}^n \leq t\}$ . We can rewrite  $m_t$  and  $M_t$  as

$$m_t = \sum_{n=1}^{N(t)} (M_{T_{\square}^n} - M_{T_{\square}^{n-1}}) + m_t - M_{T_{\square}^{N(t)}}$$

$$\text{and } M_t = \sum_{n=1}^{N(t)} (M_{T_{\square}^n} - M_{T_{\square}^{n-1}}) + M_t - M_{T_{\square}^{N(t)}}.$$

Here we have used the fact that for all  $n \in \mathbb{N}$ ,  $M_{T_{\square}^n} = m_{T_{\square}^n}$ .

Observe that the random variables  $(M_{T_{\square}^n} - M_{T_{\square}^{n-1}})_{n \geq 1}$  are i.i.d. The same is true for the random variables  $(T_{\square}^n - T_{\square}^{n-1})_{n \geq 1}$ . If we can show that

$$\mathbf{E}_{(0,0)} \left[ \max_{0 \leq t \leq T_{\square}} M_t \right] < +\infty \quad \text{and} \quad \mathbf{E}_{(0,0)} \left[ \max_{0 \leq t \leq T_{\square}} m_t \right] < +\infty, \quad (5.1)$$

then we can use Theorem 3.1 from Chapter VI in [3] and conclude. As  $t \rightarrow M_t$  and  $t \rightarrow m_t$  are non-decreasing, this amounts to showing that  $\mathbf{E}_{(0,0)} [M_{T_{\square}}] < +\infty$  and  $\mathbf{E}_{(0,0)} [m_{T_{\square}}] < +\infty$ , which is a direct consequence of Lemma 5.4.  $\square$

In order to use this result and obtain the limiting speed of growth of the 2-CGP, we need to be able to compute  $\mathbf{E}_{(0,0)} [M_{T_{\square}}]$  and  $\mathbf{E}_{(0,0)} [T_{\square}]$ . As a first step, we introduce the martingale problem satisfied by the 2-CGP, which we shall use to obtain a relation between  $\mathbf{E}_{(0,0)} [M_{T_{\square}}]$  and  $\mathbf{E}_{(0,0)} [T_{\square}]$ .

Let  $C_b(\tilde{\mathcal{S}})$  be the space of bounded functions  $f : \tilde{\mathcal{S}} \rightarrow \mathbb{R}$ . The generator  $\mathcal{G}$  of the 2-CGP acting on functions  $f \in C_b(\tilde{\mathcal{S}})$  is defined as follows. For all  $f \in C_b(\tilde{\mathcal{S}})$  and for all  $(i, j) \in \tilde{\mathcal{S}}$ ,

$$\begin{aligned} \mathcal{G}f(i, j) = & \mathbf{1}_{\{i=j\}} \times 2(f(i, i+1) - f(i, j)) \\ & + \mathbf{1}_{\{i+1=j\}} \times [2(f(i+1, j) - f(i, j)) + f(i, j+1) - f(i, j)] \\ & + \mathbf{1}_{\{j \geq i+2\}} \times \left[ 2(f(i+1, j) - f(i, j)) + f(i, j+1) - f(i, j) + \sum_{k=2}^{j-i} (f(i+k, j) - f(i, j)) \right]. \end{aligned}$$

The 2-CGP is then a solution to the following martingale problem.

**Lemma 5.6.** *Let  $(i, j) \in \tilde{\mathcal{S}}$ , and let  $(G_t)_{t \geq 0} = (m_t, M_t)_{t \geq 0}$  be the 2-CGP with initial condition  $(i, j)$ . Then, for all  $f \in C_b(\tilde{\mathcal{S}})$ ,*

$$\left( f(G_t) - f((i, j)) - \int_0^t \mathcal{G}f(G_s) ds \right)_{t \geq 0}$$

*is a martingale.*

We use the martingale problem with functions of the form  $f_d : (i, j) \rightarrow j \mathbf{1}_{\{j < d\}}$ ,  $d \in \mathbb{N} \setminus \{0, 1\}$ .

Indeed, for all  $d \in \mathbb{N} \setminus \{0, 1\}$  and  $s \geq 0$ , if  $G_s = (m_s, M_s)$  is such that  $M_s - m_s \geq 1$ , then

$$\mathcal{G}f_d(G_s) = \mathbf{1}_{\{M_s < d-1\}} - \mathbf{1}_{\{M_s = d-1\}}(d-1) \quad (5.2)$$

and if  $m_s = M_s$ , then

$$\mathcal{G}f_d(G_s) = 2\mathbf{1}_{\{M_s < d-1\}} - 2(d-1)\mathbf{1}_{\{M_s = d-1\}}. \quad (5.3)$$

We obtain the following result.

**Lemma 5.7.** *We have*

$$\mathbf{E}_{(0,0)}[M_{T_\square}] = \frac{1}{2} + \mathbf{E}_{(0,0)}[T_\square].$$

*Proof.* For all  $d \in \mathbb{N} \setminus \{0, 1\}$ , let  $T_d := \inf\{t \geq 0 : M_t \geq d\}$ . We use the martingale problem stated in Lemma 5.6 with the function  $f_d : (i, j) \rightarrow j\mathbf{1}_{\{j < d\}}$ ,  $d \in \mathbb{N} \setminus \{0, 1\}$ , and the stopping time  $T_\square \wedge T_d$ . We obtain

$$\begin{aligned} \mathbf{E}_{(0,0)}[M_{T_\square \wedge T_d}] &= \mathbf{E}_{(0,0)}\left[M_{T_\square \wedge T_d} \mathbf{1}_{\{M_{T_\square \wedge T_d} < d+2\}}\right] \\ &= 0 + \mathbf{E}_{(0,0)}\left[\int_0^{T_\square \wedge T_d} \mathcal{G}f_{d+2}(G_s) ds\right] \\ &= \mathbf{E}_{(0,0)}\left[\int_0^{T_1} \mathcal{G}f_{d+2}(G_s) ds\right] + \mathbf{E}_{(0,0)}\left[\int_{T_1}^{T_\square \wedge T_d} \mathcal{G}f_{d+2}(G_s) ds\right] \\ &= \mathbf{E}_{(0,0)}[2T_1] + \mathbf{E}_{(0,0)}[T_\square \wedge T_d - T_1] \\ &= \mathbf{E}_{(0,0)}[T_1] + \mathbf{E}_{(0,0)}[T_\square \mathbf{1}_{\{T_\square \leq T_d\}}] + \mathbf{E}_{(0,0)}[T_d \mathbf{1}_{\{T_d < T_\square\}}] \\ &= \frac{1}{2} + \mathbf{E}_{(0,0)}[T_\square \mathbf{1}_{\{T_\square \leq T_d\}}] + \mathbf{E}_{(0,0)}[T_d \mathbf{1}_{\{T_d < T_\square\}}]. \end{aligned}$$

Here we have used (5.2) and (5.3) to pass from the second to the third line. Moreover, we also have

$$\mathbf{E}_{(0,0)}[M_{T_\square \wedge T_d}] = \mathbf{E}_{(0,0)}[M_{T_\square} \mathbf{1}_{\{T_\square \leq T_d\}}] + \mathbf{E}_{(0,0)}[d \mathbf{1}_{\{T_\square > T_d\}}].$$

Therefore, if we show that

$$\lim_{d \rightarrow +\infty} \mathbf{E}_{(0,0)}[T_\square \mathbf{1}_{\{T_\square \leq T_d\}}] = \mathbf{E}_{(0,0)}[T_\square], \quad (5.4)$$

$$\lim_{d \rightarrow +\infty} \mathbf{E}_{(0,0)}[T_d \mathbf{1}_{\{T_d < T_\square\}}] = 0, \quad (5.5)$$

$$\lim_{d \rightarrow +\infty} \mathbf{E}_{(0,0)}[M_{T_\square} \mathbf{1}_{\{T_\square \leq T_d\}}] = \mathbf{E}_{(0,0)}[M_{T_\square}], \quad (5.6)$$

$$\text{and } \lim_{d \rightarrow +\infty} \mathbf{E}_{(0,0)}[M_{T_d} \mathbf{1}_{\{T_d < T_\square\}}] = 0, \quad (5.7)$$

then we shall be able to conclude.

In order to do so, recall that in order to come back to the set  $\mathcal{S}_\square$ , the process first needs to leave the state  $(0, 0)$ , which occurs at time  $T_1 \sim \text{Exp}(2)$ . Without loss of generality, we assume that the first cube falls on the right pile. As in the proof of Lemma 5.3, we assign an exponential clock  $T_\rightarrow \sim \text{Exp}(1)$  to the highest cube of the right pile, and move this exponential clock to the new highest cube in this pile whenever it grows. Then, reasoning as in the proof of Lemma 5.3, we have  $T_\square \leq T_\rightarrow + T_1$ . Moreover,  $M_{T_\square} \leq d$  if, and only if at most  $d - 1$  cubes fall on the right pile during the time interval  $[T_1, T_\square]$ . Therefore, if  $\mathcal{P}(\lambda)$ ,  $\lambda > 0$  stands for a Poisson random variable with parameter  $\lambda$ ,

$$\begin{aligned} \mathbf{P}_{(0,0)}(T_\square > T_d) &\leq \mathbf{P}(\mathcal{P}(T_\rightarrow) \geq d - 1) \\ &\leq \int_0^\infty e^{-t} \mathbf{P}(\mathcal{P}(t) \geq d - 1) dt \\ &\leq \int_0^\infty e^{-t} \frac{\mathbf{E}[\mathcal{P}(t)]}{d - 1} dt \\ &\leq \frac{1}{d - 1} \int_0^\infty t e^{-t} dt \\ &\xrightarrow{d \rightarrow +\infty} 0. \end{aligned}$$

Therefore, by the dominated convergence theorem and Lemmas 5.3 and 5.4, we can write

$$\lim_{d \rightarrow +\infty} \mathbf{E}_{(0,0)} \left[ T_\square \mathbf{1}_{\{T_d < T_\square\}} \right] = 0 \quad \text{and} \quad \lim_{d \rightarrow +\infty} \mathbf{E}_{(0,0)} \left[ M_{T_\square} \mathbf{1}_{\{T_d < T_\square\}} \right] = 0,$$

from which we deduce (5.4) and (5.6). Moreover,

$$\mathbf{E}_{(0,0)} \left[ T_d \mathbf{1}_{\{T_d < T_\square\}} \right] \leq \mathbf{E}_{(0,0)} \left[ T_\square \mathbf{1}_{\{T_d < T_\square\}} \right],$$

giving (5.5), and as  $t \rightarrow M_t$  is non-decreasing,

$$\mathbf{E}_{(0,0)} \left[ M_{T_d} \mathbf{1}_{\{T_d < T_\square\}} \right] \leq \mathbf{E}_{(0,0)} \left[ M_{T_\square} \mathbf{1}_{\{T_d < T_\square\}} \right],$$

allowing us to conclude, again by a dominated convergence argument.  $\square$

Using this result together with Theorem 5.5, we obtain a new expression for the speed of growth of the process, along with explicit lower and upper bounds.

**Proposition 5.8.** *We have*

$$\lim_{t \rightarrow +\infty} \frac{M_t}{t} = 1 + \frac{1}{2\mathbf{E}_{(0,0)}[T_\square]} \in [4/3, 2] \quad a.s.$$

*Proof.* By Theorem 5.5,

$$\lim_{t \rightarrow +\infty} \frac{M_t}{t} = \frac{\mathbf{E}_{(0,0)} [M_{T_\square}]}{\mathbf{E}_{(0,0)} [T_\square]} \text{ a.s.}$$

Moreover, by Lemma 5.7,

$$\frac{\mathbf{E}_{(0,0)} [M_{T_\square}]}{\mathbf{E}_{(0,0)} [T_\square]} = \frac{1/2 + \mathbf{E}_{(0,0)} [T_\square]}{\mathbf{E}_{(0,0)} [T_\square]} = 1 + \frac{1}{2\mathbf{E}_{(0,0)} [T_\square]}.$$

By Lemma 5.3,  $\mathbf{E}_{(0,0)} [T_\square] \leq 3/2$  and so

$$1 + \frac{1}{2\mathbf{E}_{(0,0)} [T_\square]} \geq 1 + 1/3 = 4/3.$$

Moreover, as  $T_\square \geq T_1$  and as  $T_1 \sim \text{Exp}(2)$ , we have  $\mathbf{E}_{(0,0)} [T_\square] \geq 1/2$ . Therefore,

$$1 + \frac{1}{2\mathbf{E}_{(0,0)} [T_\square]} \leq 1 + 1 = 2,$$

allowing us to conclude. □

*Remark 5.9.* Without the interaction with the other pile, the speed of growth of an isolated pile of cubes would be of 1. Therefore, this first result means that the interaction between the two piles increases the growth speed by a factor of at least 1.33, and at most 2. Considering more than two interacting piles would increase the growth speed even more, and yield a factor closer to the one obtained for the SLFV (whose value of 2.58 was obtained through numerical simulations in Section 2.4).

## 5.2 The discretised two columns growth process

The main obstacle to the study of the speed of growth of the 2-CGP lies in the fact that the process does not jump at a constant rate: the bigger the height difference between the two columns, the faster the process jumps. In order to circumvent this problem, we now introduce a discretised version of the two columns growth process. Then, we explain how to couple it to the 2-CGP, and use its invariant distribution to obtain an approximation for  $\mathbf{E}_{(0,0)} [T_\square]$  and for the speed of growth. In all that follows, let  $(G_t)_{t \geq 0} = (m_t, M_t)_{t \geq 0}$  be the 2-CGP with initial condition  $(0, 0)$ . Recall that  $T_\square$  is the time of first return of  $(G_t)_{t \geq 0}$  to a state in  $\mathcal{S}_\square$ , and for all  $d \in \mathbb{N} \setminus \{0, 1\}$ ,  $T_d = \inf\{t \geq 0 : M_t \geq d\}$ .

Moreover, let  $N \in \mathbb{N} \setminus \{0, 1\}$  and  $\epsilon > 0$ . In order to construct the discretised 2-CGP, we make the following observation. If  $t \geq 0$ , then the probability that the process  $(G_t)_{t \geq 0}$  jumps at least once in the time interval  $[t, t + \epsilon)$  is equal to

$$1 - \exp(-\epsilon(M_t - m_t + 2)) \approx \epsilon(M_t - m_t + 2),$$

and the probability that it jumps at least twice is bounded from above by

$$(1 - \exp(-\epsilon(M_t - m_t + 3)))^2 \approx \epsilon^2(M_t - m_t + 3)^2.$$

Therefore, if  $\epsilon$  is small enough and if we are able to control  $M_t - m_t$ , then we can consider that at most one growth event occurs during a time interval of length  $\epsilon$ . We use this idea to construct the discretised 2-CGP. In order to ease the notation, we only describe the dynamics of the height difference between the two piles, and define the discretised 2-CGP on the state space  $\mathbb{N}$ . Note that it is possible to recover the complete process from the evolution of the height difference. Moreover, we shall often abuse notation and say that the discretised 2-CGP starts from the state  $(0, 0)$ , or that it comes back to a state in  $\mathcal{S}_{\square}$  when it comes back to the state 0.

**Definition 5.10.** For all  $\epsilon > 0$  and  $N \in \mathbb{N} \setminus \{0, 1\}$ , the discretised 2-CGP  $(\hat{G}_n^{(N, \epsilon)})_{n \in \mathbb{N}}$  with time step  $\epsilon$  and maximal height difference  $N$  is the  $\llbracket 0, N \rrbracket$ -valued discrete-time Markov chain with initial condition  $\hat{G}_0^{(N, \epsilon)} = 0$  and whose transition probabilities  $(\hat{p}_{i,j}^{(N, \epsilon)})_{(i,j) \in \llbracket 0, N \rrbracket^2}$  are defined as follows.

1. If  $i = 0$ ,  $\hat{p}_{0,0}^{(N, \epsilon)} = \exp(-2\epsilon)$ ,  $\hat{p}_{0,1}^{(N, \epsilon)} = 1 - \exp(-2\epsilon)$ , and for all  $j \in \llbracket 2, N \rrbracket$ ,  $\hat{p}_{0,j}^{(N, \epsilon)} = 0$ .

2. For all  $i \in \llbracket 1, N - 1 \rrbracket$ ,

$$\hat{p}_{i,j}^{(N, \epsilon)} := \begin{cases} 0 & \text{if } j > i + 1, \\ \frac{2}{i+2}(1 - \exp(-(i+2)\epsilon)) & \text{if } j = i - 1, \\ \exp(-(i+2)\epsilon) & \text{if } j = i, \\ \frac{1}{i+2}(1 - \exp(-(i+2)\epsilon)) & \text{if } j = i + 1 \text{ or (if } i \neq 1) 0 \leq j \leq i - 2, \end{cases}$$

3. If  $i = N$ ,

$$\hat{p}_{N,j}^{(N, \epsilon)} := \begin{cases} 0 & \text{if } j > i + 1, \\ \frac{2}{N+2}(1 - \exp(-(N+2)\epsilon)) & \text{if } j = i - 1, \\ 1 - \frac{N+1}{N+2}(1 - \exp(-(N+2)\epsilon)) & \text{if } j = i, \\ \frac{1}{N+2}(1 - \exp(-(N+2)\epsilon)) & \text{if } 0 \leq j \leq i - 2, \end{cases}$$

This process has dynamics similar to those of the 2-CGP, except when the height difference between the two piles is equal to  $N$ : the growth of the highest pile is then blocked until the lower pile grows. This bound ensures we can compute the invariant distribution of the process.

Before studying the properties of the discretised 2-CGP, we explain how to couple it to the (continuous-time) 2-CGP. In order to do so, for all  $n \in \mathbb{N}$ , let  $t_n := n\epsilon$ . Moreover, let  $\hat{T}_p^{(\epsilon)}$  be the smallest positive integer such that  $(G_t)_{t \geq 0}$  jumps at least twice in the time interval  $[t_{\hat{T}_p^{(\epsilon)}}, t_{\hat{T}_p^{(\epsilon)} + 1})$ ,

and let  $\hat{T}_N^{(\epsilon)}$  be the smallest positive integer such that there exists  $t \in [t_{\hat{T}_N^{(\epsilon)}}, t_{\hat{T}_N^{(\epsilon)}+1})$  such that  $M_t \geq N$ .

We then construct the coupled discretised 2-CGP  $(\hat{G}_n^{(N,\epsilon)})_{n \in \mathbb{N}}$  as follows.

1. First, we set  $\hat{G}_0^{(N,\epsilon)} = 0$ .
2. For all  $n \in \llbracket 0, \min(\hat{T}_N^{(\epsilon)}, \hat{T}_p^{(\epsilon)}) - 1 \rrbracket$ , we set

$$\hat{G}_{n+1}^{(N,\epsilon)} = M_{t_{n+1}} - m_{t_{n+1}}.$$

3. If  $\hat{T}_p^{(\epsilon)} \leq \hat{T}_N^{(\epsilon)}$ ,  $\hat{G}_{\hat{T}_p^{(\epsilon)}+1}^{(N,\epsilon)}$  is taken equal to the value of  $M_t - m_t$  after the first jump of  $(G_t)_{t \geq 0}$  over the time interval  $[t_{\hat{T}_p^{(\epsilon)}}, t_{\hat{T}_p^{(\epsilon)}+1})$ . Otherwise, we set  $\hat{G}_{\hat{T}_N^{(\epsilon)}+1}^{(N,\epsilon)} = M_{t_{\hat{T}_N^{(\epsilon)}+1}} - m_{t_{\hat{T}_N^{(\epsilon)}+1}}$ .
4. For  $n > \min(\hat{T}_p^{(\epsilon)} + 1, \hat{T}_N^{(\epsilon)} + 1)$ , the coupling no longer holds, and  $(\hat{G}_n^{(N,\epsilon)})_{n \in \mathbb{N}}$  evolves according to the dynamics described in Definition 5.10.

This coupling satisfies the following property.

**Lemma 5.11.** *Let  $\epsilon > 0$  and  $N \in \mathbb{N} \setminus \{0, 1\}$ . The coupling of the discretised 2-CGP  $(\hat{G}_n^{(N,\epsilon)})_{n \in \mathbb{N}}$  with timestep  $\epsilon$  and maximal height difference  $N$  to the original 2-CGP  $(G_t)_{t \geq 0}$  holds until time  $\min(\hat{T}_p^{(\epsilon)}, \hat{T}_N^{(\epsilon)})$ . In other words, for all  $n \in \mathbb{N}$ , if  $n \leq \min(\hat{T}_p^{(\epsilon)}, \hat{T}_N^{(\epsilon)})$ , then*

$$\hat{G}_n^{(N,\epsilon)} = M_{t_n} - m_{t_n}.$$

Moreover, let  $\hat{T}_\square^{(N,\epsilon)}$  be defined as

$$\hat{T}_\square^{(N,\epsilon)} := \min \left\{ n \in \mathbb{N} \setminus \{0\} : \hat{G}_n^{(N,\epsilon)} = 0 \text{ but } \hat{G}_{n-1}^{(N,\epsilon)} \neq 0 \right\}.$$

If  $\hat{T}_\square^{(N,\epsilon)} \leq \min(\hat{T}_N^{(\epsilon)}, \hat{T}_p^{(\epsilon)})$ , then  $\epsilon \hat{T}_\square^{(N,\epsilon)} - \epsilon < T_\square \leq \epsilon \hat{T}_\square^{(N,\epsilon)}$  a.s.

Notice that contrary to  $T_\square$ , the random variable  $\hat{T}_\square^{(N,\epsilon)}$  does not correspond exactly to the time of first return of  $\hat{G}_n^{(N,\epsilon)}$  to 0, but rather to the time needed for the process to exit state 0 and then return to it. For instance, if  $\hat{G}_1^{(N,\epsilon)} = 0$ , then the time of first return to state 0 of  $(\hat{G}_n^{(N,\epsilon)})_{n \in \mathbb{N}}$  is equal to 1, while  $\hat{T}_\square^{(N,\epsilon)} > 1$ .

*Proof.* The first part of the lemma is a direct consequence of the coupling. Next, assume that  $\hat{T}_\square^{(N,\epsilon)} \leq \min(\hat{T}_N^{(\epsilon)}, \hat{T}_p^{(\epsilon)})$ . By definition of  $\hat{T}_\square^{(N,\epsilon)}$ , we know that

$$M_{t_{\hat{T}_\square^{(N,\epsilon)}}} - m_{t_{\hat{T}_\square^{(N,\epsilon)}}} = 0, \tag{5.8}$$



and for all  $n \in \llbracket 0, \hat{T}_{\square}^{(N,\epsilon)} \rrbracket$ ,

$$M_{t_n} - m_{t_n} \neq 0.$$

Moreover,  $(G_t)_{t \geq 0}$  jumps at most once over each time interval  $[t_n, t_{n+1})$ ,  $n \in \llbracket 0, \hat{T}_{\square}^{(N,\epsilon)} - 1 \rrbracket$ . Therefore, for all  $n \in \llbracket 0, \hat{T}_{\square}^{(N,\epsilon)} - 1 \rrbracket$  and  $t \in [t_n, t_{n+1})$ ,

$$M_t - m_t \in \{M_{t_n} - m_{t_n}, M_{t_{n+1}} - m_{t_{n+1}}\},$$

from which we deduce that

$$\forall t \in \left[0, t_{\hat{T}_{\square}^{(N,\epsilon)} - 1}\right], \quad M_t - m_t \neq 0.$$

Therefore,

$$T_{\square} > t_{\hat{T}_{\square}^{(N,\epsilon)} - 1} = \epsilon \left( \hat{T}_{\square}^{(N,\epsilon)} - 1 \right).$$

Moreover, by (5.8), we have

$$T_{\square} \leq t_{\hat{T}_{\square}^{(N,\epsilon)}} = \epsilon \hat{T}_{\square}^{(N,\epsilon)},$$

and we can conclude. □

The interest of the discretised 2-CGP lies in the fact that it is possible to compute  $\mathbf{E}_{(0,0)}[\hat{T}_{\square}^{(N,\epsilon)}]$  explicitly, using the invariant distribution of the process. The quantity  $\epsilon \mathbf{E}_{(0,0)}[\hat{T}_{\square}^{(N,\epsilon)}]$  is a good approximation for  $\mathbf{E}_{(0,0)}[T_{\square}]$  when  $\epsilon \rightarrow 0$  and  $N \rightarrow +\infty$ , as stated in the following result.

**Proposition 5.12.** *We have*

$$\lim_{\substack{\epsilon \rightarrow 0 \\ N \rightarrow +\infty \\ N^2 \epsilon \rightarrow 0}} \epsilon \mathbf{E}_{(0,0)} \left[ \hat{T}_{\square}^{(N,\epsilon)} \right] = \mathbf{E}_{(0,0)}[T_{\square}].$$

In order to show Proposition 5.12, we need two technical lemmas. The first one, Lemma 5.13, states that for large  $N$ , with high probability, the 2-CGP comes back to a state in  $\mathcal{S}_{\square}$  before reaching height  $N$ . The second one, Lemma 5.14, states that before coming back to a state in  $\mathcal{S}_{\square}$ , under suitable conditions on  $\epsilon$  and  $N$ , with high probability, at most one growth event occurs in each interval of length  $\epsilon$ . Notice that these two lemmas describe the properties of the *original* (non-discretised) 2-CGP. We shall use them to show that in the limiting regime, the coupling between the discretised 2-CGP and the original process still holds at time  $\hat{T}_{\square}^{(N,\epsilon)}$ , as whether the coupling breaks depends on the dynamics of the original 2-CGP.

**Lemma 5.13.** *For the non-discretised 2-CGP  $(G_t)_{t \geq 0} = (m_t, M_t)_{t \geq 0}$ , for all  $\epsilon > 0$  and  $N \in \mathbb{N} \setminus \{0, 1\}$ ,*

we have

$$\mathbf{P}_{(0,0)}(T_N \leq T_{\square} + \epsilon) \leq \frac{1}{2^{N-1}} \exp(\epsilon),$$

where we recall that  $T_N = \inf \{t \geq 0 : M_t \geq N\}$ .

*Proof.* Let  $\epsilon > 0$  and  $N \in \mathbb{N} \setminus \{0, 1\}$ . Adapting the proof of Lemma 5.7, we obtain that

$$\mathbf{P}_{(0,0)}(T_N \leq T_{\square} + \epsilon) \leq \mathbf{P}_{(0,0)}(\mathcal{P}(T_{\rightarrow} + \epsilon) \geq N - 1),$$

where  $T_{\rightarrow} \sim \text{Exp}(1)$  and where  $\mathcal{P}(\lambda)$ ,  $\lambda > 0$  stands for a Poisson random variable with parameter  $\lambda$ .

Therefore,

$$\begin{aligned} \mathbf{P}_{(0,0)}(T_N \leq T_{\square} + \epsilon) &\leq \int_0^{\infty} e^{-t} e^{-(t+\epsilon)} \left( \sum_{i=N-1}^{+\infty} \frac{(t+\epsilon)^i}{i!} \right) dt \\ &= e^{-\epsilon} \left[ \sum_{i=N-1}^{+\infty} \int_0^{\infty} e^{-2t} \frac{(t+\epsilon)^i}{i!} dt \right]. \end{aligned}$$

Moreover, for all  $i \geq N - 1$ ,

$$\begin{aligned} \int_0^{\infty} e^{-2t} \frac{(t+\epsilon)^i}{i!} dt &= \left[ -\frac{1}{2} e^{-2t} \frac{(t+\epsilon)^i}{i!} \right]_0^{\infty} + \int_0^{\infty} \frac{1}{2} e^{-2t} \frac{(t+\epsilon)^{i-1}}{(i-1)!} dt \\ &= \frac{1}{2} \frac{\epsilon^i}{i!} + \frac{1}{2} \int_0^{\infty} e^{-2t} \frac{(t+\epsilon)^{i-1}}{(i-1)!} dt \\ &= \sum_{j=0}^i \frac{1}{2^{j+1}} \frac{\epsilon^{i-j}}{(i-j)!} \\ &= \frac{1}{2^{i+1}} \left( \sum_{j=0}^i \frac{(2\epsilon)^{i-j}}{(i-j)!} \right) \\ &\leq \frac{1}{2^{i+1}} \exp(2\epsilon). \end{aligned}$$

Therefore,

$$\mathbf{P}_{(0,0)}(T_N \leq T_{\square} + \epsilon) \leq \sum_{i=N-1}^{+\infty} \frac{1}{2^{i+1}} e^{2\epsilon} e^{-\epsilon} \leq e^{\epsilon} \frac{1}{2^{N-1}},$$

which concludes the proof. □

**Lemma 5.14.** *In the notation of Lemma 5.13, we have*

$$\mathbf{P}_{(0,0)} \left( \hat{T}_p^{(\epsilon)} < \frac{T_{\square}}{\epsilon} \mid T_N > T_{\square} + \epsilon \right) \leq 2(1 - \exp(-(N+2)\epsilon)) + \frac{1}{1 - \exp(-\epsilon)} (1 - \exp(-(N+2)\epsilon))^2.$$

*Proof.* Consider the time intervals  $[t_0, t_1)$ ,  $[t_1, t_2)$ , ...,  $[t_{\lfloor T_1 \epsilon^{-1} \rfloor}, t_{\lfloor T_1 \epsilon^{-1} \rfloor + 1})$ , ...,  $[t_{\lfloor T_{\square} \epsilon^{-1} \rfloor}, t_{\lfloor T_{\square} \epsilon^{-1} \rfloor + 1})$ . We

work conditionally on  $T_N > T_\square + \epsilon$ . In order to have  $\hat{T}_p^{(\epsilon)} < T_\square \epsilon^{-1}$ , one of the following events needs to occur.

1. Another growth event occurs in the time interval  $[t_{\lfloor T_1 \epsilon^{-1} \rfloor}, t_{\lfloor T_1 \epsilon^{-1} \rfloor + 1})$ .

This occurs with probability bounded from above by  $1 - \exp(-(N+2)\epsilon)$ .

2. At least two growth events occur in at least one of the time intervals  $[t_{\lfloor T_1 \epsilon^{-1} \rfloor + 1}, t_{\lfloor T_1 \epsilon^{-1} \rfloor + 2})$ , ...,  $[t_{\lfloor T_\square \epsilon^{-1} \rfloor - 1}, t_{\lfloor T_\square \epsilon^{-1} \rfloor})$ .

For each time interval, this occurs with probability bounded from above by  $(1 - \exp(-(N+2)\epsilon))^2$ .

3. Another growth event occurs in the time interval  $[t_{\lfloor T_\square \epsilon^{-1} \rfloor}, t_{\lfloor T_\square \epsilon^{-1} \rfloor + 1})$ .

Again, this occurs with probability bounded from above by  $1 - \exp(-(N+2)\epsilon)$ .

Moreover, if only one growth event occurs during the time interval  $[t_{\lfloor T_1 \epsilon^{-1} \rfloor}, t_{\lfloor T_1 \epsilon^{-1} \rfloor + 1})$ , then the random variable  $\lfloor T_\square \epsilon^{-1} \rfloor - \lfloor T_1 \epsilon^{-1} \rfloor$  is bounded from above by a geometric law with success probability  $1 - \exp(-\epsilon)$ . Therefore,

$$\begin{aligned} & \mathbf{P}_{(0,0)} \left( \hat{T}_p^{(\epsilon)} < \frac{T_\square}{\epsilon} \mid T_N > T_\square + \epsilon \right) \\ & \leq 2(1 - \exp(-(N+2)\epsilon)) + \sum_{k=2}^{+\infty} k \exp(-\epsilon)^{k-1} (1 - \exp(-\epsilon)) (1 - \exp(-(N+2)\epsilon))^2 \\ & \leq 2(1 - \exp(-(N+2)\epsilon)) + \frac{1}{1 - \exp(-\epsilon)} (1 - \exp(-(N+2)\epsilon))^2, \end{aligned}$$

and we can conclude. □

We can now show Proposition 5.12.

*Proof.* (Proposition 5.12) By Lemma 5.11, if  $T_\square \leq \min(\hat{T}_N^{(\epsilon)}, \epsilon \hat{T}_p^{(\epsilon)})$ , then

$$\epsilon \hat{T}_\square^{(N,\epsilon)} - \epsilon < T_\square \leq \epsilon \hat{T}_\square^{(N,\epsilon)}.$$

In particular, this is true if  $T_N - \epsilon > T_\square$  and  $\hat{T}_p^{(\epsilon)} \geq T_\square \epsilon^{-1}$ . Then, by case disjunction,

$$\begin{aligned} \mathbf{E}_{(0,0)}[T_\square] &= \mathbf{E}_{(0,0)} \left[ T_\square \mathbf{1}_{\{T_N < T_\square + \epsilon\}} \right] + \mathbf{E}_{(0,0)} \left[ T_\square \mathbf{1}_{\{T_N > T_\square + \epsilon\}} \mathbf{1}_{\{\hat{T}_p^{(\epsilon)} < T_\square \epsilon^{-1}\}} \right] \\ &\quad + \mathbf{E}_{(0,0)} \left[ T_\square \mathbf{1}_{\{T_N > T_\square + \epsilon\}} \mathbf{1}_{\{\hat{T}_p^{(\epsilon)} \geq T_\square \epsilon^{-1}\}} \right], \end{aligned}$$

and similarly,

$$\mathbf{E}_{(0,0)} \left[ \hat{T}_\square^{(N,\epsilon)} \right] = \mathbf{E}_{(0,0)} \left[ \hat{T}_\square^{(N,\epsilon)} \mathbf{1}_{\{T_N < T_\square + \epsilon\}} \right] + \mathbf{E}_{(0,0)} \left[ \hat{T}_\square^{(N,\epsilon)} \mathbf{1}_{\{T_N > T_\square + \epsilon\}} \mathbf{1}_{\{\hat{T}_p^{(\epsilon)} < T_\square \epsilon^{-1}\}} \right]$$

$$+ \mathbf{E}_{(0,0)} \left[ \hat{T}_{\square}^{(N,\epsilon)} \mathbf{1}_{\{T_N > T_{\square} + \epsilon\}} \mathbf{1}_{\{\hat{T}_p(\epsilon) \geq T_{\square} \epsilon^{-1}\}} \right].$$

Therefore,

$$\begin{aligned} & \epsilon \mathbf{E}_{(0,0)} \left[ \hat{T}_{\square}^{(N,\epsilon)} \mathbf{1}_{\{T_N > T_{\square} + \epsilon\}} \mathbf{1}_{\{\hat{T}_p(\epsilon) \geq T_{\square} \epsilon^{-1}\}} \right] - \epsilon + \mathbf{E}_{(0,0)} \left[ T_{\square} \mathbf{1}_{\{T_N < T_{\square} + \epsilon\}} \right] \\ & \quad + \mathbf{E}_{(0,0)} \left[ T_{\square} \mathbf{1}_{\{T_N > T_{\square} + \epsilon\}} \mathbf{1}_{\{\hat{T}_p(\epsilon) < T_{\square} \epsilon^{-1}\}} \right] \\ & \leq \mathbf{E}_{(0,0)} [T_{\square}] \\ & \leq \epsilon \mathbf{E}_{(0,0)} \left[ \hat{T}_{\square}^{(N,\epsilon)} \mathbf{1}_{\{T_N > T_{\square} + \epsilon\}} \mathbf{1}_{\{\hat{T}_p(\epsilon) \geq T_{\square} \epsilon^{-1}\}} \right] + \mathbf{E}_{(0,0)} \left[ T_{\square} \mathbf{1}_{\{T_N < T_{\square} + \epsilon\}} \right] \\ & \quad + \mathbf{E}_{(0,0)} \left[ T_{\square} \mathbf{1}_{\{T_N > T_{\square} + \epsilon\}} \mathbf{1}_{\{\hat{T}_p(\epsilon) < T_{\square} \epsilon^{-1}\}} \right], \end{aligned}$$

and hence

$$\begin{aligned} & \epsilon \mathbf{E}_{(0,0)} \left[ \hat{T}_{\square}^{(N,\epsilon)} \right] - \epsilon + \mathbf{E}_{(0,0)} \left[ \left( T_{\square} - \epsilon \hat{T}_{\square}^{(N,\epsilon)} \right) \left( \mathbf{1}_{\{T_N < T_{\square} + \epsilon\}} + \mathbf{1}_{\{T_N > T_{\square} + \epsilon\}} \mathbf{1}_{\{\hat{T}_p(\epsilon) < T_{\square} \epsilon^{-1}\}} \right) \right] \\ & \leq \mathbf{E}_{(0,0)} [T_{\square}] \\ & \leq \epsilon \mathbf{E}_{(0,0)} \left[ \hat{T}_{\square}^{(N,\epsilon)} \right] + \mathbf{E}_{(0,0)} \left[ \left( T_{\square} - \epsilon \hat{T}_{\square}^{(N,\epsilon)} \right) \left( \mathbf{1}_{\{T_N < T_{\square} + \epsilon\}} + \mathbf{1}_{\{T_N > T_{\square} + \epsilon\}} \mathbf{1}_{\{\hat{T}_p(\epsilon) < T_{\square} \epsilon^{-1}\}} \right) \right]. \end{aligned}$$

By (5.3), we have  $\mathbf{E}_{(0,0)} [T_{\square}] \leq 3/2$ . Moreover,

$$\begin{aligned} \mathbf{E}_{(0,0)} \left[ \epsilon \hat{T}_{\square}^{(N,\epsilon)} \right] & \leq \epsilon \left( \frac{1}{\hat{p}_{0,0}^{(N,\epsilon)}} + \frac{1}{\hat{p}_{1,0}^{(N,\epsilon)}} + \max_{i \in [2, N]} \frac{1}{\hat{p}_{i,0}^{(N,\epsilon)}} \right) \\ & = \epsilon \left( \frac{1}{1 - \exp(-2\epsilon)} + \frac{3}{2(1 - \exp(-3\epsilon))} + \max_{i \in [2, N]} \frac{i+2}{1 - \exp(-(i+2)\epsilon)} \right) \\ & \leq \epsilon \left( \frac{1}{1 - \exp(-2\epsilon)} + \frac{3}{2(1 - \exp(-3\epsilon))} + \frac{N+2}{1 - \exp(-(N+2)\epsilon)} \right). \end{aligned}$$

Therefore,

$$\mathbf{E}_{(0,0)} \left[ \left| T_{\square} - \epsilon \hat{T}_{\square}^{(N,\epsilon)} \right| \right] \leq 3/2 + \epsilon \left( \frac{1}{1 - \exp(-2\epsilon)} + \frac{3}{2(1 - \exp(-3\epsilon))} + \frac{N+2}{1 - \exp(-(N+2)\epsilon)} \right)$$

and

$$\lim_{\substack{\epsilon \rightarrow 0 \\ N \rightarrow +\infty \\ N^2 \epsilon \rightarrow 0}} \mathbf{E}_{(0,0)} \left[ \left| T_{\square} - \epsilon \hat{T}_{\square}^{(N,\epsilon)} \right| \right] \leq 3/2 + 1/2 + 3/6 + 1 < +\infty.$$

Moreover, by Lemmas 5.13 and 5.14, we have

$$\mathbf{P}_{(0,0)} (T_N < T_{\square} + \epsilon) \xrightarrow[\substack{\epsilon \rightarrow 0 \\ N \rightarrow +\infty \\ N^2 \epsilon \rightarrow 0}]{\quad} 0,$$

and

$$\mathbf{P}_{(0,0)} \left( \hat{T}_p^\epsilon < T_\square \epsilon^{-1} \mid T_N > T_\square + \epsilon \right) \xrightarrow[\substack{\epsilon \rightarrow 0 \\ N \rightarrow +\infty \\ N^2 \epsilon \rightarrow 0}]{\quad} 0,$$

which allows us to conclude using the dominated convergence theorem.  $\square$

Therefore, if we compute  $\epsilon \mathbf{E}_{(0,0)}[\hat{T}_\square^{(N,\epsilon)}]$  for  $N$  large enough and  $\epsilon$  small enough, we can use the corresponding value to approximate  $\mathbf{E}_{(0,0)}[T_\square]$ . The next section is devoted to obtaining an explicit expression for  $\mathbf{E}_{(0,0)}[\hat{T}_\square^{(N,\epsilon)}]$ , using the invariant distribution of  $(\hat{G}_n^{(N,\epsilon)})_{n \geq 0}$ .

### 5.3 Invariant distribution of the discretised 2-CGP

Since the discretised 2-CGP is an irreducible and positive recurrent Markov chain, there exists a relation between its invariant distribution and the expected first return times for each of its states. We want to use this relation to obtain an explicit expression for  $\mathbf{E}_{(0,0)}[\hat{T}_\square^{(N,\epsilon)}]$ . However, as explained earlier, the definition of  $\hat{T}_\square^{(N,\epsilon)}$  does not correspond to how first return times for discrete-time Markov chains are usually defined in the literature. Therefore, the invariant distribution of  $(\hat{G}_n^{(N,\epsilon)})_{n \in \mathbb{N}}$  does not directly give access to  $\mathbf{E}_{(0,0)}[\hat{T}_\square^{(N,\epsilon)}]$ .

In order to circumvent this problem, we now introduce the *accelerated discretised 2-CGP*, denoted by  $(\tilde{G}_n^{(N,\epsilon)})_{n \geq 0}$ . The dynamics of this new process is identical to the one of the original discretised 2-CGP, *except when the process is in state 0*. In this case, the accelerated discretised 2-CGP jumps to state 1 with probability 1. Therefore, the process cannot stay in state 0 during more than one time step, and the time needed to first leave state 0, and then return to it is given by the invariant distribution of the process.

**Definition 5.15.** *The accelerated discretised 2-CGP  $(\tilde{G}_n^{(N,\epsilon)})_{n \in \mathbb{N}}$  with timestep  $\epsilon$  and maximal height difference  $N$  is the  $\llbracket 0, N \rrbracket$ -valued discrete-time Markov chain with initial condition  $\tilde{G}_0^{(N,\epsilon)} = 0$  and whose transition probabilities  $(p_{i,j}^{(N,\epsilon)})_{(i,j) \in \llbracket 0, N \rrbracket^2}$  are defined as follows.*

1. *If  $i = 0$ , then  $p_{0,0}^{(N,\epsilon)} = 0$ ,  $p_{0,1}^{(N,\epsilon)} = 1$ , and for all  $j \in \llbracket 2, N \rrbracket$ ,  $p_{0,j}^{(N,\epsilon)} = 0$ .*
2. *For all  $i \in \llbracket 1, N \rrbracket$  and for all  $j \in \llbracket 0, N \rrbracket$ ,  $p_{i,j}^{(N,\epsilon)} = \hat{p}_{i,j}^{(N,\epsilon)}$ .*

Similarly as before, we shall say that  $(\tilde{G}_n^{(N,\epsilon)})_{n \in \mathbb{N}}$  starts from the state  $(0,0)$  (*resp.*, comes back to a state in  $\mathcal{S}_{\square\square}$ ) when it starts from (*resp.*, comes back to) the state 0.

As stated before, the main difference between  $(\tilde{G}_n^{(N,\epsilon)})_{n \geq 0}$  and  $(\hat{G}_n^{(N,\epsilon)})_{n \geq 0}$  lies in the fact that  $(\tilde{G}_n^{(N,\epsilon)})_{n \geq 0}$  cannot stay in the state 0 during more than one time step. Therefore, the mean time of first return to state 0 *starting from state 1* are equal for the original process and its accelerated version. Moreover, we can compute this time for the accelerated discretised 2-CGP, since:

- Its mean time of first return to state 0 *starting from state 0* can be computed using its invariant distribution.
- If it starts from state 0, then it reaches state 1 in exactly one timestep.

Therefore, if  $\tilde{T}_{\square}^{(N,\epsilon)}$  stands for the time of first return of  $(\tilde{G}_n^{(N,\epsilon)})_{n \geq 0}$  to 0, we have the following lemma.

**Lemma 5.16.** *We have*

$$\mathbf{E}_{(0,0)}[\hat{T}_{\square}^{(N,\epsilon)}] = \mathbf{E}_{(0,0)}[\tilde{T}_{\square}^{(N,\epsilon)}] - 1 + \frac{1}{1 - \exp(-2\epsilon)}.$$

*Proof.* Indeed,  $(\tilde{G}_n^{(N,\epsilon)})_{n \geq 0}$  exits the state 0 in exactly one time step, while  $(\hat{G}_n^{(N,\epsilon)})_{n \geq 0}$  needs a number of time steps distributed as a geometrical law with probability of success  $1 - \exp(-2\epsilon)$  to do so.  $\square$

We now compute the invariant distribution of  $(\tilde{G}_n^{(N,\epsilon)})_{n \geq 0}$ . In order to do so, let  $(\tilde{\mathbf{p}}_i^{(N,\epsilon)})_{0 \leq i \leq N}$  stand for the invariant distribution of  $(\tilde{G}_n^{(N,\epsilon)})_{n \geq 0}$ . Let  $(A_i^{(N,\epsilon)})_{0 \leq i \leq N}$  be defined by backwards induction as follows.

1. First, we set  $A_N^{(N,\epsilon)} = 1$ ,  $A_{N-1}^{(N,\epsilon)} = N + 1$  and  $A_{N-2}^{(N,\epsilon)} = (N + 1)A_{N-1}^{(N,\epsilon)} - 2A_N^{(N,\epsilon)}$ .
2. Then, for all  $i \in \llbracket 2, N - 2 \rrbracket$ , we set

$$A_{i-1}^{(N,\epsilon)} = (i + 2)A_i^{(N,\epsilon)} - 2A_{i+1}^{(N,\epsilon)} - \sum_{j=i+2}^N A_j^{(N,\epsilon)}.$$

3. We conclude by setting

$$A_0^{(N,\epsilon)} = \frac{1}{2} \left( 2A_1^{(N,\epsilon)} + \sum_{j=2}^N A_j^{(N,\epsilon)} \right) (1 - \exp(-2\epsilon)).$$

Then, the sequence  $(\tilde{\mathbf{p}}_i^{(N,\epsilon)})_{0 \leq i \leq N}$  can be expressed in terms of the sequence  $(A_i^{(N,\epsilon)})_{0 \leq i \leq N}$  as follows.

**Lemma 5.17.** *For all  $i \in \llbracket 0, N \rrbracket$ , we have*

$$\tilde{\mathbf{p}}_i^{(N,\epsilon)} = \tilde{\mathbf{p}}_N^{(N,\epsilon)} A_i^{(N,\epsilon)} \frac{i + 2}{N + 2} \frac{1 - \exp(-(N + 2)\epsilon)}{1 - \exp(-(i + 2)\epsilon)}.$$

*Proof.* We show that the result is true by backwards induction. First, we check that it is true for  $i = N$ ,  $i = N - 1$  and  $i = N - 2$ .

$$A_N^{(N,\epsilon)} \frac{N + 2}{N + 2} \frac{1 - \exp(-(N + 2)\epsilon)}{1 - \exp(-(N + 2)\epsilon)} = A_N^{(N,\epsilon)} = 1,$$

and so the property is true for  $i = N$ . Then, by definition of the invariant distribution, we have

$$\tilde{\mathbf{p}}_{N-1}^{(N,\epsilon)} p_{N-1,N}^{(N,\epsilon)} + \tilde{\mathbf{p}}_N^{(N,\epsilon)} p_{N,N}^{(N,\epsilon)} = \tilde{\mathbf{p}}_N^{(N,\epsilon)},$$

and so

$$\begin{aligned} \tilde{\mathbf{p}}_{N-1}^{(N,\epsilon)} &= \frac{1}{p_{N-1,N}^{(N,\epsilon)}} \tilde{\mathbf{p}}_N^{(N,\epsilon)} (1 - p_{N,N}^{(N,\epsilon)}) \\ &= \frac{N+1}{1 - \exp(-(N+1)\epsilon)} \tilde{\mathbf{p}}_N^{(N,\epsilon)} \left( \frac{N+1}{N+2} (1 - \exp(-(N+2)\epsilon)) \right) \\ &= \tilde{\mathbf{p}}_N^{(N,\epsilon)} \frac{N+1}{N+2} (N+1) \frac{1 - \exp(-(N+2)\epsilon)}{1 - \exp(-(N+1)\epsilon)} \end{aligned}$$

and the result is true for  $i = N - 1$ . Moreover,

$$\tilde{\mathbf{p}}_{N-2}^{(N,\epsilon)} p_{N-2,N-1}^{(N,\epsilon)} + \tilde{\mathbf{p}}_{N-1}^{(N,\epsilon)} p_{N-1,N-1}^{(N,\epsilon)} + \tilde{\mathbf{p}}_N^{(N,\epsilon)} p_{N,N-1}^{(N,\epsilon)} = \tilde{\mathbf{p}}_{N-1}^{(N,\epsilon)},$$

which means that

$$\begin{aligned} \tilde{\mathbf{p}}_{N-2}^{(N,\epsilon)} &= \frac{1}{p_{N-2,N-1}^{(N,\epsilon)}} \left[ \tilde{\mathbf{p}}_{N-1}^{(N,\epsilon)} (1 - p_{N-1,N-1}^{(N,\epsilon)}) - \tilde{\mathbf{p}}_N^{(N,\epsilon)} p_{N,N-1}^{(N,\epsilon)} \right] \\ &= \frac{N}{1 - \exp(-N\epsilon)} \tilde{\mathbf{p}}_N^{(N,\epsilon)} \left[ \frac{N+1}{N+2} A_{N-1}^{(N,\epsilon)} \frac{1 - \exp(-(N+2)\epsilon)}{1 - \exp(-(N+1)\epsilon)} (1 - \exp(-(N+1)\epsilon)) \right] \\ &\quad - \frac{N}{1 - \exp(-N\epsilon)} \tilde{\mathbf{p}}_N^{(N,\epsilon)} \left[ \frac{2}{N+2} (1 - \exp(-(N+2)\epsilon)) \right] \\ &= \tilde{\mathbf{p}}_N^{(N,\epsilon)} \frac{N}{N+2} \left( (N+1) A_{N-1}^{(N,\epsilon)} - 2 A_N^{(N,\epsilon)} \right) \frac{1 - \exp(-(N+2)\epsilon)}{1 - \exp(-N\epsilon)} \\ &= \tilde{\mathbf{p}}_N^{(N,\epsilon)} \frac{N}{N+2} A_{N-2}^{(N,\epsilon)} \frac{1 - \exp(-(N+2)\epsilon)}{1 - \exp(-N\epsilon)} \end{aligned}$$

by definition of  $A_{n-2}^{(N,\epsilon)}$ .

Then, let  $i \in \llbracket 2, N-2 \rrbracket$ , and assume that the property is true for  $j \in \llbracket i, N \rrbracket$ . Again by definition of the invariance property,

$$\tilde{\mathbf{p}}_{i-1}^{(N,\epsilon)} p_{i-1,i}^{(N,\epsilon)} + \tilde{\mathbf{p}}_i^{(N,\epsilon)} p_{i,i}^{(N,\epsilon)} + \tilde{\mathbf{p}}_{i+1}^{(N,\epsilon)} p_{i+1,i}^{(N,\epsilon)} + \sum_{j=i+2}^N \tilde{\mathbf{p}}_j^{(N,\epsilon)} p_{j,i}^{(N,\epsilon)} = \tilde{\mathbf{p}}_i^{(N,\epsilon)},$$

from which we deduce

$$\begin{aligned} \tilde{\mathbf{p}}_{i-1}^{(N,\epsilon)} &= \frac{1}{p_{i-1,i}^{(N,\epsilon)}} \left( \tilde{\mathbf{p}}_i^{(N,\epsilon)} (1 - p_{i,i}^{(N,\epsilon)}) - \tilde{\mathbf{p}}_{i+1}^{(N,\epsilon)} p_{i+1,i}^{(N,\epsilon)} - \sum_{j=i+2}^N \tilde{\mathbf{p}}_j^{(N,\epsilon)} p_{j,i}^{(N,\epsilon)} \right) \\ &= \frac{i+1}{1 - \exp(-(i+1)\epsilon)} \tilde{\mathbf{p}}_i^{(N,\epsilon)} \end{aligned}$$

$$\begin{aligned}
& \times \left[ \frac{i+2}{N+2} A_i^{(N,\epsilon)} \frac{1 - \exp(-(N+2)\epsilon)}{1 - \exp(-(i+2)\epsilon)} (1 - \exp(-(i+2)\epsilon)) \right. \\
& \quad - \frac{i+3}{N+2} A_{i+1}^{(N,\epsilon)} \frac{1 - \exp(-(N+2)\epsilon)}{1 - \exp(-(i+3)\epsilon)} \frac{2}{i+3} (1 - \exp(-(i+3)\epsilon)) \\
& \quad \left. - \sum_{j=i+2}^N \frac{j+2}{N+2} A_j^{(N,\epsilon)} \frac{1 - \exp(-(N+2)\epsilon)}{1 - \exp(-(j+2)\epsilon)} \frac{1}{j+2} (1 - \exp(-(j+2)\epsilon)) \right] \\
& = \tilde{\mathbf{p}}_N^{(N,\epsilon)} \frac{i+1}{N+2} \frac{1 - \exp(-(N+2)\epsilon)}{1 - \exp(-(i+1)\epsilon)} \left( (i+2) A_i^{(N,\epsilon)} - 2A_{i+1}^{(N,\epsilon)} - \sum_{j=i+2}^N A_j^{(N,\epsilon)} \right) \\
& = \tilde{\mathbf{p}}_N A_{i-1}^{(N,\epsilon)} \frac{i+1}{N+2} \frac{1 - \exp(-(N+2)\epsilon)}{1 - \exp(-(i+1)\epsilon)}
\end{aligned}$$

by definition of  $A_{i-1}^{(N,\epsilon)}$ .

We now need to check that the property is true for  $i = 0$ . We have

$$\tilde{\mathbf{p}}_1^{(N,\epsilon)} p_{1,0}^{(N,\epsilon)} + \sum_{j=2}^N \tilde{\mathbf{p}}_j^{(N,\epsilon)} p_{j,0}^{(N,\epsilon)} = \tilde{\mathbf{p}}_0^{(N,\epsilon)},$$

and so

$$\begin{aligned}
\tilde{\mathbf{p}}_0^{(N,\epsilon)} &= \left[ \tilde{\mathbf{p}}_1^{(N,\epsilon)} p_{1,0}^{(N,\epsilon)} + \sum_{j=2}^N \tilde{\mathbf{p}}_j^{(N,\epsilon)} p_{j,0}^{(N,\epsilon)} \right] \\
&= \tilde{\mathbf{p}}_N^{(N,\epsilon)} \frac{1 - \exp(-(N+2)\epsilon)}{N+2} \\
& \quad \times \left[ 3 \frac{2}{3} \frac{1 - \exp(-3\epsilon)}{1 - \exp(-3\epsilon)} A_1^{(N,\epsilon)} + \sum_{j=2}^N \frac{(j+2) A_j^{(N,\epsilon)}}{1 - \exp(-(j+2)\epsilon)} \frac{1}{j+2} (1 - \exp(-(j+2)\epsilon)) \right] \\
&= \tilde{\mathbf{p}}_N^{(N,\epsilon)} \frac{1}{N+2} \frac{1 - \exp(-(N+2)\epsilon)}{1 - \exp(-2\epsilon)} \left( 2A_1^{(N,\epsilon)} + \sum_{j=2}^N A_j^{(N,\epsilon)} \right) (1 - \exp(-2\epsilon)) \\
&= \tilde{\mathbf{p}}_N^{(N,\epsilon)} \frac{2}{N+2} A_0^{(N,\epsilon)} \frac{1 - \exp(-(N+2)\epsilon)}{1 - \exp(-2\epsilon)},
\end{aligned}$$

which allows us to conclude.  $\square$

We can now use the invariant distribution to obtain an explicit formula for  $\mathbf{E}_{(0,0)} [\hat{T}_{\square}^{(N,\epsilon)}]$ .

**Proposition 5.18.** *We have*

$$\mathbf{E}_{(0,0)} [\hat{T}_{\square}^{(N,\epsilon)}] = \frac{1}{1 - \exp(-2\epsilon)} + \frac{2A_0^{(N,\epsilon)}}{1 - \exp(-2\epsilon)} \left( \sum_{i=1}^N \frac{(i+2) A_i^{(N,\epsilon)}}{1 - \exp(-(i+2)\epsilon)} \right).$$

*Proof.* We know that

$$\mathbf{E}_{(0,0)} [\tilde{T}_{\square}^{(N,\epsilon)}] = \frac{1}{\tilde{\mathbf{p}}_0^{(N,\epsilon)}} \quad \text{and} \quad \sum_{i=0}^N \tilde{\mathbf{p}}_i^{(N,\epsilon)} = 1.$$



Using Lemma 5.17, we obtain that

$$\tilde{\mathbf{p}}_N^{(N,\epsilon)} \frac{1 - \exp(-(N+2)\epsilon)}{N+2} \left[ \sum_{i=0}^N \frac{(i+2)A_i^{(N,\epsilon)}}{1 - \exp(-(i+2)\epsilon)} \right] = 1,$$

and so

$$\tilde{\mathbf{p}}_N^{(N,\epsilon)} = \frac{N+2}{1 - \exp(-(N+2)\epsilon)} \frac{1}{\sum_{i=0}^N \frac{(i+2)A_i^{(N,\epsilon)}}{1 - \exp(-(i+2)\epsilon)}}.$$

Using again Lemma 5.17 yields

$$\begin{aligned} \tilde{\mathbf{p}}_0^{(N,\epsilon)} &= 2A_0^{(N,\epsilon)} \frac{1}{1 - \exp(-2\epsilon)} \left( \sum_{i=0}^N \frac{(i+2)A_i^{(N,\epsilon)}}{1 - \exp(-(i+2)\epsilon)} \right)^{-1} \\ &= \frac{1}{1 + \left[ \sum_{i=1}^N \frac{(i+2)A_i^{(N,\epsilon)}}{1 - \exp(-(i+2)\epsilon)} \right]^{-1} \frac{1 - \exp(-2\epsilon)}{2A_0^{(N,\epsilon)}}}, \end{aligned}$$

and using Lemma 5.16, we obtain

$$\begin{aligned} \mathbf{E}_{(0,0)} \left[ \hat{T}_{\square}^{(N,\epsilon)} \right] &= \mathbf{E}_{(0,0)} \left[ \tilde{T}_{\square}^{(N,\epsilon)} \right] - 1 + \frac{1}{1 - \exp(-2\epsilon)} \\ &= 1 + \frac{2A_0^{(N,\epsilon)}}{1 - \exp(-2\epsilon)} \left( \sum_{i=1}^N \frac{(i+2)A_i^{(N,\epsilon)}}{1 - \exp(-(i+2)\epsilon)} \right) - 1 + \frac{1}{1 - \exp(-2\epsilon)} \\ &= \frac{1}{1 - \exp(-2\epsilon)} + \frac{2A_0^{(N,\epsilon)}}{1 - \exp(-2\epsilon)} \left( \sum_{i=1}^N \frac{(i+2)A_i^{(N,\epsilon)}}{1 - \exp(-(i+2)\epsilon)} \right). \end{aligned}$$

□

We can now use Proposition 5.12 in order to compute an approximation for  $\mathbf{E}_{(0,0)}[T_{\square}]$ .

**Proposition 5.19.**

$$\mathbf{E}_{(0,0)}[T_{\square}] = \frac{1}{2} + \lim_{\substack{\epsilon \rightarrow 0 \\ N \rightarrow +\infty \\ N^2\epsilon \rightarrow 0}} \sum_{i=1}^N \frac{(i+2)A_0^{(N,\epsilon)} A_i^{(N,\epsilon)}}{1 - \exp(-(i+2)\epsilon)}.$$

*Proof.* By Proposition 5.12,

$$\mathbf{E}_{(0,0)}[T_{\square}] = \lim_{\substack{\epsilon \rightarrow 0 \\ N \rightarrow +\infty \\ N^2\epsilon \rightarrow 0}} \epsilon \mathbf{E}_{(0,0)} \left[ \hat{T}_{\square}^{(N,\epsilon)} \right].$$

Moreover, by Proposition 5.18,

$$\epsilon \mathbf{E}_{(0,0)} \left[ \hat{T}_{\square}^{(N,\epsilon)} \right] = \frac{1}{1 - \exp(-2\epsilon)} + \frac{2A_0^{(N,\epsilon)}}{1 - \exp(-2\epsilon)} \left( \sum_{i=1}^N \frac{(i+2)A_i^{(N,\epsilon)}}{1 - \exp(-(i+2)\epsilon)} \right).$$

We conclude by using the fact that

$$\frac{\epsilon}{1 - \exp(-2\epsilon)} \xrightarrow{\epsilon \rightarrow 0} \frac{1}{2}.$$

□

We obtain that  $\mathbf{E}_{(0,0)}[T_{\square}] \simeq 1.46$ , which is higher than the lower bound of 1.33 obtained at the end of Section 5.1.

**Acknowledgements** This work was partly supported by the chair program "Mathematical Modelling and Biodiversity" of Veolia Environment-Ecole Polytechnique-National Museum of Natural History-Foundation X. AL would like to thank Laetitia Colombani and Carl Graham for helpful discussions on cumulative processes, percolation theory and stochastic geometry.

## References

- [1] S.E. Alm and M. Deijfen. First passage percolation on  $\mathbb{Z}^2$ : A simulation study. *Journal of Statistical Physics*, 161(3):657–678, 2015.
- [2] S.E. Alm and R. Parviainen. Lower and upper bounds for the time constant of first-passage percolation. *Combinatorics, Probability and Computing*, 11(5):433–445, 2002.
- [3] S. Asmussen. *Applied probability and queues*, volume 51. Springer Science & Business Media, 2008.
- [4] A. Auffinger, M. Damron, and J. Hanson. *50 years of first-passage percolation*, volume 68. American Mathematical Soc., 2017.
- [5] L. Bertini and G. Giacomin. Stochastic Burgers and KPZ equations from particle systems. *Communications in Mathematical Physics*, 183(3):571–607, 1997.
- [6] Niloy Biswas, Alison Etheridge, and Aleksander Klimek. The spatial Lambda-Fleming-Viot process with fluctuating selection. *Electronic Journal of Probability*, 26:1–51, 2021.
- [7] S. Chatterjee and P. Dey. Multiple phase transitions in long-range first-passage percolation on square lattices. *Communications on Pure and Applied Mathematics*, 69(2):203–256, 2016.
- [8] J. Chetwynd-Diggles and A. Klimek. Rare mutations in the spatial Lambda-Fleming-Viot model in a fluctuating environment and superBrownian motion. *arXiv preprint arXiv:1901.04374*, 2019.

- [9] T. Cox and R. Durrett. Some limit theorems for percolation processes with necessary and sufficient conditions. *Annals of Probability*, pages 583–603, 1981.
- [10] M. Deijfen. Asymptotic shape in a continuum growth model. *Advances in Applied Probability*, 35(2):303–318, 2003.
- [11] M. Deijfen and O. Häggström. Coexistence in a two-type continuum growth model. *Advances in Applied Probability*, 36(4):973–980, 2004.
- [12] M. Deijfen, O. Häggström, and J. Bagley. A stochastic model for competing growth on  $\mathbb{R}^d$ . *Markov Processes and Related Fields*, 10(2):217–248, 2004.
- [13] R. Durrett and W.-T. L. Fan. Genealogies in expanding populations. *Annals of Applied Probability*, 26(6):3456–3490, 2016.
- [14] M. Eden. A two-dimensional growth process. In *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability*, volume 4, pages 223–239. University of California Press Berkeley, 1961.
- [15] A. Etheridge, N. Freeman, and S. Penington. Branching Brownian motion, mean curvature flow and the motion of hybrid zones. *Electronic Journal of Probability*, 22:1–40, 2017.
- [16] A. Etheridge, N. Freeman, S. Penington, and D. Straulino. Branching Brownian motion and selection in the spatial Lambda-Fleming–Viot process. *Annals of Applied Probability*, 27(5):2605–2645, 2017.
- [17] A. Etheridge, N. Freeman, and D. Straulino. The Brownian net and selection in the spatial Lambda-Fleming-Viot process. *Electronic Journal of Probability*, 22, 2017.
- [18] A. Etheridge, A. Véber, and F. Yu. Rescaling limits of the spatial Lambda-Fleming-Viot process with selection. *Electronic Journal of Probability*, 25, 2020.
- [19] R.A. Fisher. The wave of advance of advantageous genes. *Annals of Eugenics*, 7(4):355–369, 1937.
- [20] R. Forien and S. Penington. A central limit theorem for the spatial Lambda-Fleming-Viot process with selection. *Electronic Journal of Probability*, 22:1–68, 2017.
- [21] E. Graciá, F. Botella, J.D. Anadón, P. Edelaar, D.J. Harris, and A. Giménez. Surfing in tortoises? Empirical signs of genetic structuring owing to range expansion. *Biology Letters*, 9(3):20121091, 2013.

- [22] O. Hallatschek, P. Hersen, S. Ramanathan, and D.R. Nelson. Genetic drift at expanding frontiers promotes gene segregation. *Proceedings of the National Academy of Sciences*, 104(50):19926–19930, 2007.
- [23] O. Hallatschek and D.R. Nelson. Gene surfing in expanding populations. *Theoretical Population Biology*, 73(1):158–170, 2008.
- [24] O. Hallatschek and D.R. Nelson. Life at the front of an expanding population. *Evolution: International Journal of Organic Evolution*, 64(1):193–206, 2010.
- [25] W. Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58(301):13–30, 1963.
- [26] C.D. Howard and C.M. Newman. Euclidean models of first-passage percolation. *Probability Theory and Related Fields*, 108(2):153–170, 1997.
- [27] M.A.C. Huergo, M.A. Pasquale, A.E. Bolzán, A.J. Arvia, and P.H. González. Morphology and dynamic scaling analysis of cell colonies with linear growth fronts. *Physical Review E*, 82(3):031903, 2010.
- [28] M. Kardar, G. Parisi, and Y.-C. Zhang. Dynamic scaling of growing interfaces. *Physical Review Letters*, 56(9):889, 1986.
- [29] A. Klimek and T.C. Rosati. The spatial Lambda-Fleming-Viot process in a random environment. *arXiv preprint arXiv:2004.05931*, 2020.
- [30] A.N. Kolmogorov, I.G. Petrovskii, and N.S. Piskunov. Étude de l'équation de la chaleur, de la matière et son application à un problème biologique. *Bull. Moskov. Gos. Univ. Mat. Mekh*, 1:125, 1937.
- [31] T. M. Liggett. An improved subadditive ergodic theorem. *Annals of Probability*, 13(4):1279–1285, 1985.
- [32] A. Louvet. Stochastic measure-valued models for populations expanding in a continuum. *arXiv preprint arXiv:2103.02902*, 2021.
- [33] C. Mueller and R.B. Sowers. Random travelling waves for the KPP equation with noise. *Journal of Functional Analysis*, 128:439–498, 1995.
- [34] D. Richardson. Random growth in a tessellation. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 74, pages 515–528. Cambridge University Press, 1973.

- [35] H. Rost. Non-equilibrium behaviour of a many particle process: Density profile and local equilibria. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 58(1):41–53, 1981.
- [36] T. Seppäläinen. Lecture notes on the corner growth model. *Unpublished notes*, 2009.
- [37] R.T. Smythe and J.C. Wierman. *First-passage percolation on the square lattice*, volume 671. Springer, 1978.
- [38] J. van den Berg and H. Kesten. Inequalities for the time constant in first-passage percolation. *Annals of Applied Probability*, pages 56–80, 1993.

## A Geometrical properties of ellipses

In this section, we show some geometrical properties of ellipses, which are used in other sections. In all that follows, let  $z_c = (x_c, y_c) \in \mathbb{R}^2$ ,  $(a, b) \in (0, +\infty)$  and  $\gamma \in (-\pi/2, \pi/2)$ . Recall that  $\mathfrak{B}_{a,b,\gamma}(z)$  is the ellipse defined by:

$$\mathfrak{B}_{a,b,\gamma}(z_c) = \left\{ \begin{pmatrix} x_c \\ y_c \end{pmatrix} + A_\gamma \begin{pmatrix} ar \cos(\theta) \\ br \sin(\theta) \end{pmatrix} : r \in [0, 1], \theta \in [0, 2\pi) \right\}$$

where

$$A_\gamma = \begin{pmatrix} \cos(\gamma) & -\sin(\gamma) \\ \sin(\gamma) & \cos(\gamma) \end{pmatrix}.$$

The first lemma gives the maximal *horizontal separation* between a point in the ellipse and its centre. This result is used in Section 3 to construct the express chain.

**Lemma A.1.** *Let  $f : [0, 1] \times [-\pi, \pi) \rightarrow \mathbb{R}$  be the function defined by*

$$\forall (r, \theta) \in [0, 1] \times [-\pi, \pi), f(r, \theta) = ar \cos(\theta) \cos(\gamma) - br \sin(\theta) \sin(\gamma).$$

*Then,  $f$  reaches its maximum for*

$$(r_{max}, \theta_{max}) = \left( 1, \arctan \left( -\frac{b}{a} \tan(\gamma) \right) \right),$$

*and*

$$f(r_{max}, \theta_{max}) = \sqrt{a^2 \cos^2(\gamma) + b^2 \sin^2(\gamma)}.$$

*Proof.* First, it is obvious that  $r_{max} = 1$ . Moreover,  $\cos(\theta_{max})$  must be of the same parity as  $\cos(\gamma)$ , and  $\sin(\theta)$  must be of opposite parity from  $\sin(\gamma)$ . As  $\gamma \in (-\pi/2, \pi/2)$ , we obtain that  $\theta_{max} \in (-\pi/2, \pi/2)$

too (that is, it cannot take the values  $\pm\pi/2$ ).

The function  $f_\theta : (-\pi/2, \pi/2) \rightarrow \mathbb{R}$  such that for all  $\theta \in (-\pi/2, \pi/2)$ ,

$$f_\theta(\theta) := a \cos(\theta) \cos(\gamma) - b \sin(\theta) \sin(\gamma)$$

is differentiable, and for all  $\theta \in (-\pi/2, \pi/2)$ ,

$$f'_\theta(\theta) = -a \sin(\theta) \sin(\gamma) - b \cos(\theta) \sin(\gamma).$$

Therefore,

$$f'_\theta(\theta_{max}) = 0 \iff a \sin(\theta_{max}) \cos(\gamma) = -b \cos(\theta_{max}) \sin(\gamma) \iff \theta_{max} = \arctan\left(-\frac{b}{a} \tan(\gamma)\right).$$

Moreover, since  $\cos(\gamma) > 0$ , we can write

$$\begin{aligned} & a \cos(\gamma) \cos(\theta_{max}) - b \sin(\gamma) \sin(\theta_{max}) \\ &= a \cos(\gamma) \cos\left(\arctan\left(-\frac{b}{a} \tan(\gamma)\right)\right) - b \sin(\gamma) \sin\left(\arctan\left(-\frac{b}{a} \tan(\gamma)\right)\right) \\ &= a \cos(\gamma) \frac{1}{\sqrt{1 + \frac{b^2}{a^2} \tan^2(\gamma)}} + b \sin(\gamma) \frac{\frac{b}{a} \tan(\gamma)}{\sqrt{1 + \frac{b^2}{a^2} \tan^2(\gamma)}} \\ &= \frac{a^2 \cos^2(\gamma)}{a \cos(\gamma) \sqrt{1 + \frac{b^2}{a^2} \tan^2(\gamma)}} + \frac{1}{a \cos(\gamma)} \frac{b^2 \sin^2(\gamma)}{\sqrt{1 + \frac{b^2}{a^2} \tan^2(\gamma)}} \\ &= \frac{a^2 \cos^2(\gamma) + b^2 \sin^2(\gamma)}{\sqrt{a^2 \cos^2(\gamma) + b^2 \sin^2(\gamma)}} \\ &= \sqrt{a^2 \cos^2(\gamma) + b^2 \sin^2(\gamma)}, \end{aligned}$$

which concludes the proof. □

The second lemma means that the mean horizontal separation of a point in the ellipse from its centre is equal to 0. Therefore, when a point is affected by a reproduction event, the mean horizontal separation of the centre of the corresponding ellipse from the point is equal to 0. That is:

**Lemma A.2.** *Let  $Z = (x + X, y + Y)$  be sampled uniformly at random in the ellipse  $\mathfrak{B}_{a,b,\gamma}(z)$ . Then,*

$$\mathbf{E}[X] = 0.$$