



**HAL**  
open science

# Extinction threshold and large population limit of a plant metapopulation model with recurrent extinction events and a seed bank component

Apolline Louvet

► **To cite this version:**

Apolline Louvet. Extinction threshold and large population limit of a plant metapopulation model with recurrent extinction events and a seed bank component. 2021. hal-03304276v2

**HAL Id: hal-03304276**

**<https://hal.science/hal-03304276v2>**

Preprint submitted on 16 Dec 2021 (v2), last revised 21 Feb 2022 (v3)

**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.

# Extinction threshold and large population limit of a plant metapopulation model with recurrent extinction events and a seed bank component

Apolline Louvet <sup>\*,1</sup>

<sup>1</sup>MAP5, Université de Paris, CNRS, 45 rue des Saints-Pères, 75270 Paris Cedex 06, France.

## Abstract

We introduce a new model for plant metapopulations with a seed bank component, living in a fragmented environment in which local extinction events are frequent. This model is an intermediate between population dynamics models with a seed bank component, based on the classical Wright-Fisher model, and Stochastic Patch Occupancy Models (SPOMs) used in metapopulation ecology. Its main feature is the use of "ghost" individuals, which can reproduce but with a very strong selective disadvantage against "real" individuals, to artificially ensure a constant population size. We show the existence of an extinction threshold above which persistence of the subpopulation of "real" individuals is not possible, and investigate how the seed bank characteristics affect this extinction threshold. We also show the convergence of the model to a SPOM under an appropriate scaling, bridging the gap between individual-based models and occupancy models.

**Running headline:** The  $k$ -parent WFSB metapopulation model

**Keywords:** Wright-Fisher model, seed-bank, extinction/recolonization, fluctuating population size, metapopulation, Stochastic Patch Occupancy Model, percolation, limit theorem

**MSC 2020 Subject Classification:** *Primary:* 60F99, 60J05, 92D25, *Secondary:* 60K35, 92D40

## Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
1.1	The $k$ -parent Wright-Fisher metapopulation process with seed bank . . . . .	5
1.2	The associated $k$ -parent occupancy process and its limit . . . . .	10

---

\*Email address : apolline.louvet@polytechnique.edu

1.2.1	BOA process and $k$ -parent occupancy process . . . . .	10
1.2.2	Convergence of the $k$ -parent occupancy process to the BOA process . . . . .	12
1.2.3	Critical patch extinction probability . . . . .	15
<b>2</b>	<b>Proof of the convergence of the <math>k</math>-parent occupancy process to the BOA process</b>	<b>16</b>
2.1	Upper bound on $\mathbb{P}(\text{Par}_i^{(M),n})$ . . . . .	21
2.2	Upper bound on (2.6) . . . . .	22
2.3	Proof of Theorem 1.9 . . . . .	24
<b>3</b>	<b>Extinction threshold for the <math>k</math>-parent WFSB metapopulation process</b>	<b>27</b>
3.1	Coupling between the $k$ -parent WFSB metapopulation process and the BOA process .	28
3.2	Percolation problem . . . . .	30
3.3	Proof of Theorem 1.10 . . . . .	31
<b>4</b>	<b>Appendix - Computation of <math>p_{crit}(H)</math></b>	<b>34</b>
	<b>References</b>	<b>36</b>

# 1 Introduction

Understanding how plant populations survive in fragmented landscapes is an important question in ecology and conservation biology [14]. One potential driver of plant populations' persistence is the ability to form a seed bank, which greatly influences population and community dynamics [15]. For such plant species, the seeds produced can stay dormant in the soil for up to several decades depending on the species, without losing viability [3]. See [28] for an overview of seed bank characteristics and properties, along with the emergent phenomena it can generate.

Populations living in fragmented landscapes are often modelled as metapopulations, that is, as populations distributed over a set of interconnected patches. Metapopulations are also frequently characterized by recurrent local extinction events, regional persistence being the result of a balance between colonization (from neighbouring patches or from an external source) and local extinction events [29, 32]. See [23] for a general introduction to metapopulation theory.

Many classical metapopulation models, such as the Levins model [29] or the Propagule Rain model [20], describe the occupancy of each patch (i.e whether the species of interest is present or absent in each of the patches) and do not depend on, nor model, the actual census numbers. These models are referred to as Stochastic Patch Occupancy models, or SPOMs. Since presence/absence data is easier to collect than abundance data, and since parameter inference is possible for a broad range of SPOMs (see e.g [33, 34, 36]), they are well-suited to the study of real metapopulations. Classical metapopulation models do not account for seed dormancy, but more recently models incorporating a seed bank component were also developed [9, 17, 36]. The model introduced in [36] was successfully applied to plant metapopulations in highly disturbed environments, such as weeds in agroecosystems [36] or plants in urban tree bases [31], highlighting that some plant species monitored did have a seed bank.

In population genetics, metapopulation models often describe the number and genetic types of individuals rather than the occupancy in each patch. They are usually defined by first specifying an intra-patch dynamic, and then adding migration between patches. The migration process can heavily depend on the underlying geographical structure, as in the stepping-stone model [26], or not depend on it at all, as in Wright's island model [40]. See e.g [27, 38, 39] and references therein for examples of metapopulation models based on Wright's island model, and [1, 2, 35] and references therein for examples of metapopulation models based on the stepping-stone model.

Models used to specify the intra-patch dynamic can be classical population dynamics models, without any intra-patch spatial structure, provided patches are considered as sufficiently small to neglect spatial effects in each one of them. The geographical structure in the metapopulation model

is then only contained in the localization of the patches. The intra-patch dynamic can comprise a seed bank, using population dynamics models with a seed bank component, such as the ones based on the Wright-Fisher model. In the original Wright-Fisher model, the population size (in a single patch) is constant through time and equal to  $N$ , and each individual has a genetic type, or allele. In each generation, each one of the  $N$  new individuals chooses a parent uniformly at random among the  $N$  individuals in the previous generation, and adopts its type. Including a seed bank in the Wright-Fisher model implies choosing a parent potentially not in the previous generation, but at least two generations ago, the maximal number of potentially contributing generations being bounded [25] or not [4, 5]. See [6] for a review of seed bank models in population genetics, and [11, 21, 41] for extensions of the Wright-Fisher model with a seed bank component to metapopulations.

For plant metapopulations in which extinction events are frequent, we can expect the population size of each patch to vary a lot from one generation to the next. This contradicts the constant patch population size hypothesis underlying the use of a Wright-Fisher model. In order to incorporate extinction event-induced fluctuations in a Wright-Fisher model, it is possible to adopt the approach used in [13, 22], based on a long-range biased voter model: assign a maximal population size to each patch, and fill the remaining space with "ghost", or type 0, individuals. In this framework, each patch contains both type 1 "real" individuals and type 0 "ghost" individuals, the former having a very strong selective advantage over the latter (in the spirit of [30]). As in [13, 22], the model can then be interpreted as an interacting particle system, in which a 0 corresponds to a ghost individual and a 1 to a real individual.

In this article, we introduce a new individual-based metapopulation model for plant metapopulations in which local extinction events are frequent. This model is primarily suited to annual plants living in highly disturbed patchy environments, such as urban tree bases or agroecosystems. It is also adapted to other plant species living in such environments, provided each patch is "emptied" at the end of each generation (for instance by gardeners in an urban environment or by farmers in an agroecosystem). The intra-patch dynamics will be based on a variant of a Wright-Fisher model with a seed bank component, using ghost individuals to allow for fluctuating patch population sizes. It will use the model introduced in [5], with an extra bound introduced on the number of generations a seed can stay dormant without losing viability. Indeed, for some plant species, seeds lose viability after only one or two years of dormancy [3]. This is reminiscent of the model introduced in [25], in the sense that real individuals come from parents living a bounded number of generations ago. The main difference is that in our model, individuals first choose parents living potentially arbitrarily far ago in the past, and then obtain their (real or ghost) type depending on their choice.

In order to bridge the gap between individual-based metapopulation models and SPOMs, we will show that our metapopulation process can be embedded in a SPOM. Moreover, we will prove that under an appropriate scaling of patch population size and of the "selection" strength of real individuals against ghost individuals (quantified by the parameter  $k$ ), the individual-based metapopulation process converges to this SPOM. The convergence result has two applications. First, from a theoretical viewpoint, it shows that a specific SPOM (or presence/absence-based model) is the scaling limit of an individual-based metapopulation model. Then, we will use the convergence result and the embedding in order to show the existence of an extinction threshold for metapopulation persistence, depending only on the seed bank parameters, and highlighting how the presence of a seed bank can prevent metapopulation extinction.

While the metapopulation model we will introduce and study is based on models coming from population genetics, this article will not focus on the study of the genetic diversity in such populations, which is deferred to future work. Instead, the aims of this work are threefold:

1. Introduce a general individual-based metapopulation model with a seed bank component, in which local extinction events can be frequent and patch population sizes can vary from one generation to the next.
2. Show the existence of an extinction threshold depending on the seed bank parameters.
3. Bridge the gap between SPOMs and individual-based metapopulation models by showing that in a well-chosen parameter regime, the individual-based metapopulation model we consider converges to a SPOM.

### 1.1 The $k$ -parent Wright-Fisher metapopulation process with seed bank

In the whole paper, we use the notation  $\mathbb{N} = \{0, 1, 2, \dots\}$ , and for all  $M \in \mathbb{N} \setminus \{0\}$ ,  $\llbracket 1, M \rrbracket = \{1, \dots, M\}$ .

We will consider that the metapopulation is formed by an infinite number of patches arranged in a line. A patch contains a fixed number of *seed bank compartments*, each one containing exactly one seed: either a ghost (type 0) seed, or a real (type 1) seed. In order to define the metapopulation model, we describe how in each generation, seeds germinate and grow into plants which produce new seeds and die. Concretely, the metapopulation model will only record the composition of the *seed bank* at the beginning of each generation, and not the standing vegetation in each patch in each generation.

In all that follows, let  $M \in \mathbb{N}^*$ ,  $H \in \mathbb{N}$ ,  $k \in \mathbb{N} \setminus \{0, 1\}$ ,  $g \in (0, 1)$ ,  $c \in (0, 1/2)$  and  $p \in [0, 1]$ . We assume that  $\lfloor gM \rfloor \geq 1$ . Patches will be indexed by  $i \in \mathbb{Z}$ , and seed bank compartments inside a patch by  $j \in \llbracket 1, M \rrbracket$ . The notation  $(i, j)$  will correspond to the seed bank compartment  $j$  in patch  $i$ .

The following two spaces will be used to describe the initial types and the age of the seeds occupying the seed bank compartments:

$$\mathcal{F}_M := \left\{ (\xi_{i,j})_{i \in \mathbb{Z}, j \in \llbracket 1, M \rrbracket} \in \{0, 1\}^{\mathbb{Z} \times \llbracket 1, M \rrbracket} : \text{Card}(\{(i, j) \in \mathbb{Z} \times \llbracket 1, M \rrbracket : \xi_{i,j} = 1\}) < +\infty \right\},$$

and  $\mathcal{H}_M := \left\{ (h_{i,j})_{i \in \mathbb{Z}, j \in \llbracket 1, M \rrbracket} : \forall (i, j) \in \mathbb{Z} \times \llbracket 1, M \rrbracket, h_{i,j} \in \mathbb{N} \right\}.$

$(\xi, h) \in \mathcal{F}_M \times \mathcal{H}_M$  corresponds to a metapopulation in which for all  $(i, j) \in \mathbb{Z} \times \llbracket 1, M \rrbracket$ , the seed occupying the seed bank compartment  $(i, j)$  is of age  $h_{i,j}$  and of type  $\xi_{i,j} \mathbb{1}_{\{h_{i,j} \leq H\}}$ . That is, the seed in  $(i, j)$  was originally of type  $\xi_{i,j}$  when it was produced, but may have expired since then.

The  $k$ -parent Wright-Fisher metapopulation process with seed bank is defined in the following way.

**Definition 1.1.** (*k*-parent WFSB metapopulation process) *Let  $(\xi, h) \in \mathcal{F}_M \times \mathcal{H}_M$ . The  $k$ -parent Wright-Fisher metapopulation process with seed bank, with parameters  $(M, H, g, c, p)$  and initial condition  $(\xi, h)$  and denoted by  $(\xi^n, h^n)_{n \in \mathbb{N}}$ , is the  $(\mathcal{F}_M \times \mathcal{H}_M)$ -valued Markov chain defined by  $(\xi^0, h^0) = (\xi, h)$  and for all  $n \in \mathbb{N}$ , given  $(\xi^n, h^n)$  :*

1. For each  $i \in \mathbb{Z}$ , we sample  $\lfloor gM \rfloor$  different seed bank compartments  $s_{i,1}, \dots, s_{i,\lfloor gM \rfloor} \in \llbracket 1, M \rrbracket$  uniformly at random in patch  $i$ .
2. Let  $(\text{Ext}_i)_{i \in \mathbb{Z}}$  be i.i.d  $\{0, 1\}$ -valued random variables such that  $\mathbb{P}(\text{Ext}_1 = 1) = p$ .
3. For all  $(i, j) \in \mathbb{Z} \times \llbracket 1, M \rrbracket$ , if  $j \in \{s_{i,j'} : j' \in \llbracket 1, \lfloor gM \rfloor \rrbracket\}$ , we first set  $h_{i,j}^{n+1} = 0$ . Moreover, let  $C_{1,i,j}, \dots, C_{k,i,j}$  be i.i.d  $\{-1, 0, 1\}$ -valued random variables such that

$$\mathbb{P}(C_{1,i,j} = 1) = \mathbb{P}(C_{1,i,j} = -1) = c,$$

and such that  $(C_{l,i,j})_{1 \leq l \leq k}$  is independent from  $(\text{Ext}_i)_{i \in \mathbb{Z}}$  and from  $(C_{i',j',l})_{1 \leq l \leq k}$  for  $(i', j') \neq (i, j)$ .

For all  $l \in \llbracket 1, k \rrbracket$ , if  $\text{Ext}_{i+C_{l,i,j}} = 1$ , we set  $\tilde{k}_l = 0$ , and if  $\text{Ext}_{i+C_{l,i,j}} = 0$ , we sample one seed bank compartment  $j_l$  uniformly at random among the  $\lfloor gM \rfloor$  ones sampled in the patch  $i + C_{l,i,j}$  (those in the set  $\{s_{i+C_{l,i,j},j'} : j' \in \llbracket 1, \lfloor gM \rfloor \rrbracket\}$ ), and we set

$$\tilde{k}_l = \xi_{i+C_{l,i,j},j_l}^n \mathbb{1}_{\{h_{i+C_{l,i,j},j_l}^n \leq H\}}.$$

We conclude by setting  $\xi_{i,j}^{n+1} = \max\{\tilde{k}_l : l \in \llbracket 1, k \rrbracket\}$ .

4. On the other hand, if  $j \notin \{s_{i,j'} : j' \in \llbracket 1, \lfloor gM \rfloor \rrbracket\}$ , we set  $\xi_{i,j}^{n+1} = \xi_{i,j}^n$  and  $h_{i,j}^{n+1} = h_{i,j}^n + 1$ .

Intuitively, the  $k$ -parent WFSB metapopulation process evolves as follows.

1. At each generation, exactly  $\lfloor gM \rfloor$  seeds germinate in each patch. Type 0 seeds yield (ghost) type 0 plants, while type 1 seeds yield (real) type 1 plants *only if the seed was produced less than  $H + 1$  generations ago*, i.e, only if it has not expired.
2. Then, each patch is affected by an extinction event independently from other patches and with probability  $p$ . During an extinction event, all the juvenile plants in the patch become type 0 plants.
3. In each patch, the  $\lfloor gM \rfloor$  empty seed bank compartments are filled with new seeds in the following way. For each compartment,  $k$  potential parents are chosen uniformly at random, each one of them being chosen in the same patch with probability  $1 - 2c$ , or in the patch on the left (resp. on the right) with probability  $c$ . The same potential parent may be chosen more than once for the same seed bank compartment. If all the  $k$  plants chosen as potential parents are of type 0, then the seed bank compartment is filled with a type 0 seed produced by the last plant chosen. Conversely, if at least one of the  $k$  plants chosen is of type 1, then the first type 1 plant chosen produces a seed which fills the seed bank compartment.
4. The remaining seeds stay dormant in the seed bank until the next generation.

See Figure 1 for an illustration of this dynamics. As mentioned above, observe that while the dynamics involves seeds germinating, growing into plants which produce new seeds and then die, the model only encodes the *seed bank composition*, and not the types of the plants.

In all that follows, we will refer to :

- $M$  as the *number of seeds per patch*,
- $H$  as the *maximal dormancy duration*,
- $g$  as the *germination probability*,
- $c$  as the *potential colonization probability*,
- $p$  as the *patch extinction probability*.

Moreover, we will say that a patch *goes extinct* during generation  $n$  if it is affected by an extinction event during this generation (i.e  $\text{Ext}_i = 1$ ), and that it is *empty* if it does not contain any viable type 1 seeds at the beginning of generation  $n$ . Notice that extinction events only affect standing vegetation, and not the seed bank. Therefore, in our terminology, an extinct patch is not necessarily empty. We will also say that the metapopulation *became empty* before generation  $n$  if all the patches are empty in generation  $n$ , that is, if the subpopulation of real individuals did not persist.



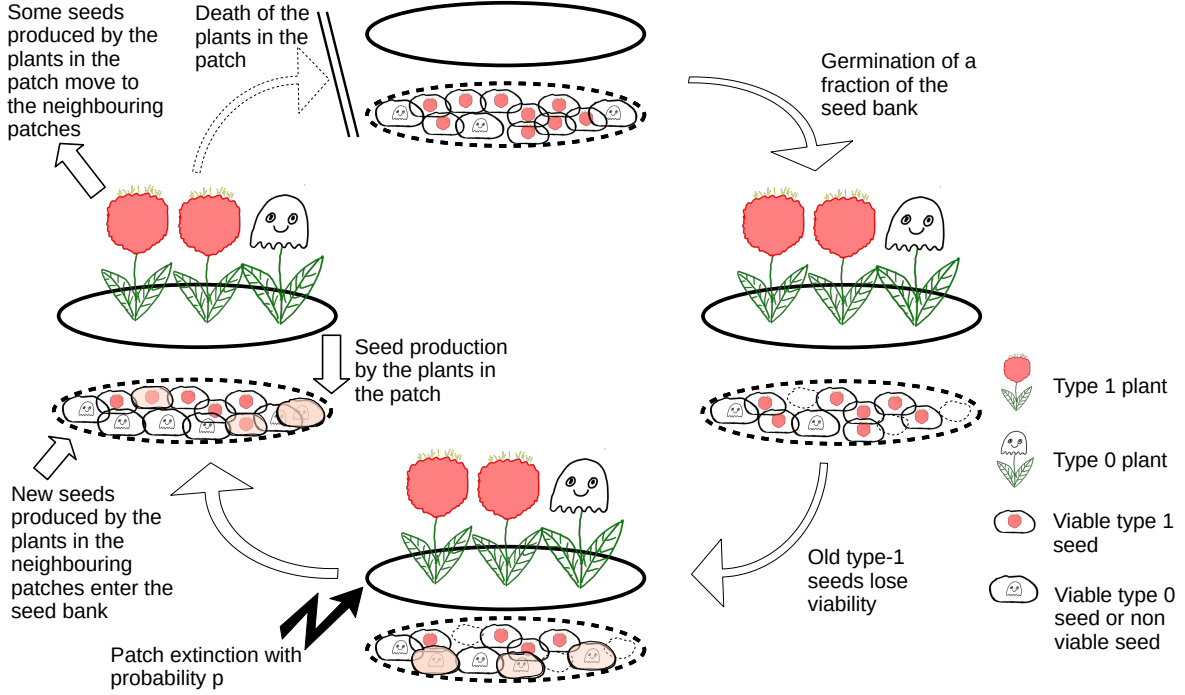


Figure 1: Illustration of the intra-patch dynamics of the  $k$ -parent WFSB metapopulation process. Here  $M = 12$  and  $\lfloor gM \rfloor = 3$ . The double line in the top line of the figure indicates the starting time of a new generation.

*Remark 1.2.* From a purely mathematical viewpoint, it is also possible to define the model for  $k = 1$ . However, in this case, real individuals do not have any selective advantage against ghost individuals, and since we assume that the number of real individuals in generation 0 is finite, the metapopulation becomes empty in finite time almost surely.

*Remark 1.3.* Even if the model is defined for  $c \in (0, 1/2)$ , in practice, since one of the assumptions behind the model is that colonization from patches which are not nearest neighbours is negligible, it is implicitly assumed that  $c$  is small. Moreover, notice that if  $c > 1/3$ , then the potential parents have higher chance of being taken from the patch on the left (or right) than in the focal patch.

Before stating our main results on the model, we give some insight on the reproduction dynamics from a *forwards in time* viewpoint. In order to do so, we first consider a real plant isolated in a patch, surrounded by empty patches. Each empty seed bank compartment in the focal patch chooses the real plant as a potential parent with probability  $1 - (1 - (1 - 2c)\lfloor gM \rfloor^{-1})^k$ , and each empty seed bank compartment in one of the two neighbouring patches chooses it with probability  $1 - (1 - c\lfloor gM \rfloor^{-1})^k$ . Therefore, from a forwards in time viewpoint, the number of offspring of the real plant is distributed as the sum of three independent random variables: one binomial random variable with parameters  $(\lfloor gM \rfloor, 1 - (1 - (1 - 2c)\lfloor gM \rfloor^{-1})^k)$ , and two binomial random variables with

parameters  $(\lfloor gM \rfloor, 1 - (1 - c\lfloor gM \rfloor^{-1})^k)$ .

If  $M \rightarrow +\infty$  while all other parameters stay constant, the number of offspring is approximately distributed as the sum of three independent Poisson random variables: one with parameter  $(1 - 2c)k$  and two with parameter  $ck$ . In particular, the average number of offspring is roughly equal to  $k$ , yielding a possible biological interpretation for the parameter  $k$ . In this article, we focus instead on another parameter regime:  $M \rightarrow +\infty$  and  $k = \lceil M^\alpha \rceil$ ,  $\alpha > 1$ . In this regime, the average number of offspring is rather equal to  $3\lfloor gM \rfloor$ . From a biological viewpoint, this parameter regime corresponds to considering that a plant can potentially produce far more viable seeds than the carrying capacity of a patch.

When several real plants are present, the average number of offspring produced by each plant decreases due to competition. In the case in which competition is the most intense (that is, when the patch and its two neighbours contain only real plants), the number of offspring of each plant in the focal patch is in average equal to 1, and is distributed as the sum of three independent random variables: one binomial random variable with parameters  $(\lfloor gM \rfloor, (1 - 2c)\lfloor gM \rfloor^{-1})$ , and two binomial random variables with parameters  $(\lfloor gM \rfloor, c\lfloor gM \rfloor^{-1})$ . When  $M \rightarrow +\infty$  and  $k$  is fixed, this can be approximated by the sum of a Poisson random variable with parameter  $1 - 2c$  and two Poisson random variables with parameter  $c$ . In particular, the random variables do not depend on the parameter  $k$ .

*Remark 1.4.* It is possible to generalize the  $k$ -parent WFSB metapopulation process by taking the potential parents of a seed in more patches than only neighbouring patches, or by having patches in a two dimensional environment instead of a one dimensional one. If the distance that seeds can travel is bounded, then all the results in this article can be extended to the generalized model (though the numerical values for the extinction thresholds will change).

*Remark 1.5.* The idea of sampling several potential parents to model selection can be found in various population genetics models, including variants of the Wright-Fisher model. See e.g [7, 8, 10, 16, 18, 19]. Usually, the models comprise both selective reproduction events, during which several potential parents are chosen, and neutral reproduction events, during which only one parent is chosen. Moreover, the mathematical analysis often involves taking selective reproduction events to be rare compared to neutral reproduction events, and to change of time scale to observe them in the limit. In contrast, the model we introduce in this article only comprises selective reproduction events, and the questions we aim at answering do not require a change of time scale.

## 1.2 The associated $k$ -parent occupancy process and its limit

### 1.2.1 BOA process and $k$ -parent occupancy process

The  $k$ -parent WFSB metapopulation process can be seen as a multi-colony Wright-Fisher model with selection and seed bank, embedded in a Stochastic Patch Occupancy Model indicating which patches are empty, extinct, or potentially occupied. Indeed, for a patch to contain real seeds, it is not sufficient for it not to go extinct. The viable seeds it contains can only come from 3 patches (the focal patch and its two neighbours), and can only have entered the seed bank during the  $H + 1$  previous generations. If all these times, the 3 patches were affected by extinction events, then the patch cannot contain viable seeds during the current generation. For instance, if  $H = 0$ , a patch which went extinct along with its two neighbours during the previous generation cannot contain non-expired type 1 seeds. In the SPOM we define just below, this patch will appear as empty. In other words, the SPOM will encode which patches *cannot* contain type 1 seeds, given the initial condition and the extinction events.

This SPOM is defined on the state space  $\mathcal{F}^\infty \times \mathcal{H}^\infty$ , with  $\mathcal{F}^\infty$  and  $\mathcal{H}^\infty$  given by:

$$\mathcal{F}^\infty := \{(O_i)_{i \in \mathbb{Z}} : \forall i \in \mathbb{Z}, O_i \in \{0, 1\} \text{ and } \text{Card}(\{i \in \mathbb{Z} : O_i = 1\}) < +\infty\}$$

and  $\mathcal{H}^\infty := \{(h_i)_{i \in \mathbb{Z}} : \forall i \in \mathbb{Z}, h_i \in \mathbb{N}\}.$

As for the  $k$ -parent WFSB metapopulation process, each patch is associated to a type (0 or 1) and an age, but now they have a different interpretation. Indeed, in the SPOM, a "type 0" patch corresponds to a patch which cannot contain nonexpired type 1 seeds, while a "type 1" patch is a patch which can potentially contain type 1 seeds, the age  $h_i$  encoding the last time type 1 seeds could have entered the seed bank.

**Definition 1.6.** (*BOA process*) Let  $(O, h) \in \mathcal{F}^\infty \times \mathcal{H}^\infty$ . The Best Occupancy Achievable process (or BOA process) with parameters  $(H, p)$  and with initial conditions  $(O, h)$  is the  $(\mathcal{F}^\infty \times \mathcal{H}^\infty)$ -valued Markov process  $(O^{\infty, n}, h^{\infty, n})_{n \in \mathbb{N}}$  defined as follows. First, we set  $(O^{\infty, 0}, h^{\infty, 0}) = (O, h)$ . Then, for all  $n \in \mathbb{N}$ , given  $(O^{\infty, n}, h^{\infty, n})$  :

1. Let  $(\text{Ext}_i)_{i \in \mathbb{Z}}$  be i.i.d  $\{0, 1\}$ -valued random variables such that  $\mathbb{P}(\text{Ext}_1 = 1) = p$ .
2. For all  $i \in \mathbb{Z}$ , if  $\text{Ext}_i = 0$  and  $O_i^{\infty, n} \mathbb{1}_{\{h_i^{\infty, n} \leq H\}} = 1$ , then we set

$$O_{i-1}^{\infty, n+1} = O_i^{\infty, n+1} = O_{i+1}^{\infty, n+1} = 1$$

and  $h_{i-1}^{\infty, n+1} = h_i^{\infty, n+1} = h_{i+1}^{\infty, n+1} = 0.$

We do nothing during this step if  $\text{Ext}_i = 1$  or  $O_i^{\infty,n} \mathbf{1}_{\{h_i^{\infty,n} \leq H\}} = 0$ .

3. For all  $i \in \mathbb{Z}$ , if  $O_i^{\infty,n+1}$  was not defined during step 2, then we set  $O_i^{\infty,n+1} = O_i^{\infty,n}$  and  $h_i^{\infty,n+1} = h_i^{\infty,n} + 1$ .

Moreover, we will say that patch  $i \in \mathbb{Z}$  is reachable at generation  $n \in \mathbb{N}$  if  $O_i^{\infty,n} \mathbf{1}_{\{h_i^{\infty,n} \leq H\}} = 1$ .

The BOA process represents all the patches which can potentially contain seeds produced by the ones initially present (as given by  $(O, h)$ ), given the extinction events. In other words, informally, the BOA process keeps track of the patches that are linked to the patches originally containing viable seeds by means of a path of reachable patches. Notice that  $O_i^{\infty,n}$  describes the composition of the seed bank, while extinction events affect the standing vegetation. Therefore, an extinction event affecting patch  $i$  during the  $n$ -th generation does not set the value of  $O_i^{\infty,n}$  to 0.

The BOA process is a best-case scenario, in the sense that using the same extinction events to construct the BOA process and the  $k$ -parent WFSB metapopulation process, it is possible to couple both processes so that all patches containing seeds in the  $k$ -parent WFSB metapopulation process are reachable patches in the BOA process. In order to formalize the coupling property, we introduce a new object associated to our metapopulation process, describing whether the seed bank in each patch contains real seeds, or only ghost seeds.

**Definition 1.7.** (*k*-parent occupancy process) Let  $(\xi, h) \in \mathcal{F}_M \times \mathcal{H}_M$ . The *k*-parent occupancy process

$$\left( O^{k,n}, h^{k,n} \right)_{n \in \mathbb{N}} = \left( \left( O_i^{k,n}, h_i^{k,n} \right)_{i \in \mathbb{Z}} \right)_{n \in \mathbb{N}}$$

associated to the *k*-parent WFSB metapopulation process  $(\xi^n, h^n)_{n \in \mathbb{N}}$  with parameters  $(M, H, g, c, p)$  and initial conditions  $(\xi, h)$  is defined as follows.

First, for all  $i \in \mathbb{Z}$ , we set

$$O_i^{k,0} := 1 - \prod_{j \in \llbracket 1, M \rrbracket} (1 - \xi_{i,j}) = \max\{\xi_{i,j} : j \in \llbracket 1, M \rrbracket\}$$

$$h_i^{k,0} := \begin{cases} \min\{h_{i,j} : j \in \llbracket 1, M \rrbracket \text{ and } \xi_{i,j} = 1\} & \text{if } O_i^{k,0} = 1 \\ 0 & \text{otherwise.} \end{cases}$$

Then, for all  $n \in \mathbb{N}^*$  and  $i \in \mathbb{Z}$ , we set

$$O_i^{k,n} := 1 - \prod_{j \in \llbracket 1, M \rrbracket} (1 - \xi_{i,j}^n)$$

$$h_i^{k,n} := \begin{cases} \min\{h_{i,j}^n : j \in \llbracket 1, M \rrbracket \text{ and } \xi_{i,j}^n = 1\} & \text{if } O_i^{k,n} = 1 \\ h_i^{k,n-1} + 1 & \text{otherwise.} \end{cases}$$

In this setting,  $O_i^{k,n} = 1$  if and only if at the beginning of generation  $n$ , before germination occurs, the patch  $i$  contains at least one (potentially expired) seed which was initially of type 1. In this case,  $h_i^{k,n}$  is the number of complete generations spent in the seed bank by the youngest of such seeds. Therefore, patch  $i$  contains at least one type 1 seed at the beginning of generation  $n$  if, and only if it contains seeds which were initially of type 1 (i.e.,  $O_i^{k,n} = 1$ ) and the youngest seeds among these ones entered the seed bank at most  $H + 1$  generations ago (i.e.,  $\mathbb{1}_{\{h_i^{k,n}\}} \leq H$ , as  $h_i^{k,n} = 0$  if the seeds entered the seed bank during the previous generation), that is, if and only if  $O_i^{k,n} \mathbb{1}_{\{h_i^{k,n} \leq H\}} = 1$ .

*Remark 1.8.* Note that the  $k$ -parent occupancy process is also defined on the state space  $\mathcal{F}^\infty \times \mathcal{H}^\infty$ . However, contrary to the BOA process, the  $k$ -parent occupancy process *cannot* be considered as a SPOM, since  $(O^{k,n+1}, h^{k,n+1})$  does not depend only on  $(O^{k,n}, h^{k,n})$ . Therefore, both processes are intrinsically different.

In all that follows, we will say that the BOA process  $(O^{\infty,n}, h^{\infty,n})_{n \in \mathbb{N}}$  associated to the  $k$ -parent WFSB metapopulation process with parameters  $(M, H, g, c, p)$  and initial condition  $(\xi, h)$  is the BOA process with parameters  $(H, p)$  and initial condition  $(O^{k,0}, h^{k,0})$ , constructed using the same extinction events as the  $k$ -parent WFSB metapopulation process. Under this coupling, the  $k$ -parent WFSB metapopulation process and its associated BOA process satisfy the following relation:

$$\forall n \in \mathbb{N}, \forall i \in \mathbb{Z}, O_i^{k,n} \leq O_i^{\infty,n} \text{ and } h_i^{k,n} \geq h_i^{\infty,n}.$$

This result will be proved in Section 3.1.

### 1.2.2 Convergence of the $k$ -parent occupancy process to the BOA process

When  $M$  and  $k$  are finite, deviations from the BOA process can occur in the following three cases:

1. Type 1 plants are present in a patch, but none of them is chosen as a potential parent.
2. Non-expired type 1 seeds are present in a patch, but none of them germinate.
3. Several type 1 seeds entered the seed bank less than  $H + 1$  generations ago, but all of them already germinated, and there is no remaining non-expired type 1 seeds in the seed bank.

However, when both  $M \rightarrow +\infty$  and  $k \rightarrow +\infty$  in an appropriate way, we can show that the occupancy process converges to the BOA process. For this convergence to occur, two conditions need

to be satisfied. First,  $k$  needs to grow to  $+\infty$  "faster" than  $M$ . We will set  $k = \lceil M \rceil^\alpha$ , with  $\alpha > 1$ , and hence define a sequence of  $\lceil M \rceil^\alpha$ -parent WFSB processes. Notice that since the  $k$  potential parents of an individual do not have to be necessarily different, it is possible to have  $k > 3\lfloor gM \rfloor$  (the number of plants in the focal patch and the two neighbouring patches). Then, we will need the following constraints on the initial conditions of the processes. Let  $(O^\infty, h^\infty) \in (\mathcal{F}^\infty \times \mathcal{H}^\infty)$  satisfying

$$\forall i \in \mathbb{Z}, \text{ if } O_i^\infty = 0, \text{ then } h_i^\infty = 0 \quad (1.1)$$

encode which patches are initially occupied, and what is the age of the youngest type 1 seeds in each of these patches. By convention, we set  $h_i^\infty = 0$  for the patches initially empty. For all  $i \in \mathbb{Z}$  such that  $O_i^\infty = 1$ , let  $g_i \in (0, 1]$ , which will represent the proportion of youngest type 1 seeds in patch  $i$ .

For all  $M \geq 2$ , let  $(\xi^{(M)}, h^{(M)}) \in (\mathcal{F}_M \times \mathcal{H}_M)$  be such that for all  $i \in \mathbb{Z}$  and  $j \in \llbracket 1, M \rrbracket$ , the following conditions are satisfied.

- (A) If  $O_i^\infty = 0$ , then  $\xi_{i,j}^{(M)} = 0$ .
- (B) If  $O_i^\infty = 1$  and  $h_{i,j}^{(M)} < h_i^\infty$ , then  $\xi_{i,j}^{(M)} = 0$ .
- (C) If  $O_i^\infty = 1$ , then  $M^{-1} \sum_{j=1}^M \xi_{i,j}^{(M)} \mathbf{1}_{\{h_{i,j}^{(M)} = h_i^\infty\}} = M^{-1} \lfloor g_i M \rfloor$ .

Intuitively, conditions (A), (B) and (C) ensure that the patches initially occupied in all  $k$ -parent WFSB metapopulation processes are the same. Conditions (B) and (C) implies that in each patch, the youngest type 1 seeds (if present) have the same age for all processes. Moreover, condition (C) quantifies the proportion of youngest type 1 seeds in the seed bank, and ensures they represent a non-negligible portion of the seed bank, even when  $M \rightarrow +\infty$ . Note that this constraint is on the proportion of the *youngest* type 1 seeds, and not on the proportion of all type 1 seeds.

**Theorem 1.9.** *Let  $\alpha > 1$ . For all  $M \geq 2$ , let  $(O^{(M),n}, h^{(M),n})_{n \in \mathbb{N}}$  be the  $\lceil M^\alpha \rceil$ -parent occupancy process associated to the  $\lceil M^\alpha \rceil$ -parent WFSB metapopulation process with parameters  $(M, H, g, c, p)$  and initial condition  $(\xi^{(M)}, h^{(M)})$ , and let  $(O^{(M),\infty,n}, h^{(M),\infty,n})_{n \in \mathbb{N}}$  be the BOA process associated to the same WFSB metapopulation process. Then, for all  $N \in \mathbb{N}$ ,*

$$\mathbb{P} \left( \bigcap_{n=0}^N \left( \left\{ \forall i \in \mathbb{Z}, O_i^{(M),n} = O_i^{(M),\infty,n} \right\} \cap \left\{ \forall i \in \mathbb{Z}, h_i^{(M),n} = h_i^{(M),\infty,n} \right\} \right) \right) \xrightarrow{M \rightarrow +\infty} 1.$$

One of the biological interpretations of this result is that in the limit considered, the metapopulation dynamics is well approximated by the BOA process. Moreover, this theorem bridges the gap between individual-based metapopulation models and SPOMs, in the sense that the BOA process is the limit

of the  $k$ -parent WFSB metapopulation process under a suitable scaling.

**Sketch of the proof.** The proof is structured as follows. First, by condition (A), type 1 seeds are initially present in only a finite number of patches. Since a plant can only send seeds in the neighbouring patches, the number of occupied patches stays finite. Then, we show that when  $k = \lceil M^\alpha \rceil$  is large enough, in each patch, the  $k$  potential parents chosen to refill any seed bank compartment in the patch span all the  $3\lfloor gM \rfloor$  possible potential parents with high probability. In this case, the  $\lfloor gM \rfloor$  new seeds entering the seed bank of the patch are all of the same type. Therefore, we can distinguish two periods:

- A transition period corresponding to the  $H+1$  first generations, during which some seeds initially present are still in the seed bank and viable, implying that some viable seeds of the same age (and in the same patch) are potentially of different types;
- What we will call the "post-transition period", corresponding to the other generations, during which all the seeds initially present and still in the seed bank are no longer viable, and during which all viable seeds of the same age and in the same patch are of the same type (with an high probability).

We first focus on the proportion of viable seeds of a given age in the seed bank after the transition period. We show that for all  $0 \leq h \leq H$ , the proportion of age  $h$  seeds (which are generally all of the same type) is roughly equal to  $g(1-g)^h$ , and that approximately  $g^2(1-g)^h M$  ( $> 1$  for  $M$  large enough) such seeds germinate during the next generation. Notice that it is what we would obtain if each seed germinated independently from others and with probability  $g$ . In particular, if type 1 seeds enter the seed bank of a patch during generation  $n$ , then with high probability at least one of them will germinate during each of the generations  $n+1, \dots, n+H+1$ , and produce new seeds if the patch is not affected by an extinction event, as in the BOA process.

We conclude by considering the transition period. At the beginning of generation  $n \in \llbracket 0, H \rrbracket$ , in each non-empty patch:

- Either no new type 1 seeds have already entered the seed bank. By conditions (B) and (C), the proportion of youngest type 1 seeds is then roughly equal to  $g_i(1-g)^n$ , and approximately  $gg_i(1-g)^n M$  ( $> 1$  for  $M$  large enough) such seeds germinate during the next generation.
- Either new type 1 seeds have already entered the seed bank. We are then in the same situation as during the post-transition period.

### 1.2.3 Critical patch extinction probability

Using the coupling with the BOA process, we will also show the existence of a critical patch extinction probability  $p_{crit}(H)$  depending only on  $H$  such that for all  $p > p_{crit}(H)$ , the metapopulation will almost surely become empty in finite time no matter the values of  $M$ ,  $g$ ,  $c$  or  $k$ .

**Theorem 1.10.** *For all  $H \in \mathbb{N}$ , there exists  $p_{crit}(H) \in (0, 1)$  such that for all  $M \in \mathbb{N}^*$ ,  $k \in \mathbb{N} \setminus \{0, 1\}$ ,  $g \in (0, 1)$  and  $c \in (0, 1/2)$ , for all  $(\xi, h) \in \mathcal{F}_M \times \mathcal{H}_M$  and  $p > p_{crit}(H)$ , if  $(O^{k,n}, h^{k,n})_{n \in \mathbb{N}}$  is the  $k$ -parent occupancy process associated to the  $k$ -parent WFSB metapopulation process with parameters  $(M, H, g, c, p)$  and initial condition  $(\xi, h)$ , then*

$$\lim_{n \rightarrow +\infty} \mathbb{P} \left( \forall i \in \mathbb{Z}, O_i^{k,n} \mathbb{1}_{\{h_i^{k,n} \leq H\}} = 0 \right) = 1.$$

The proof of this result, which can be found in Section 3, relies on the coupling between the  $k$ -parent WFSB metapopulation process and the BOA process, together with appropriate results in percolation theory.

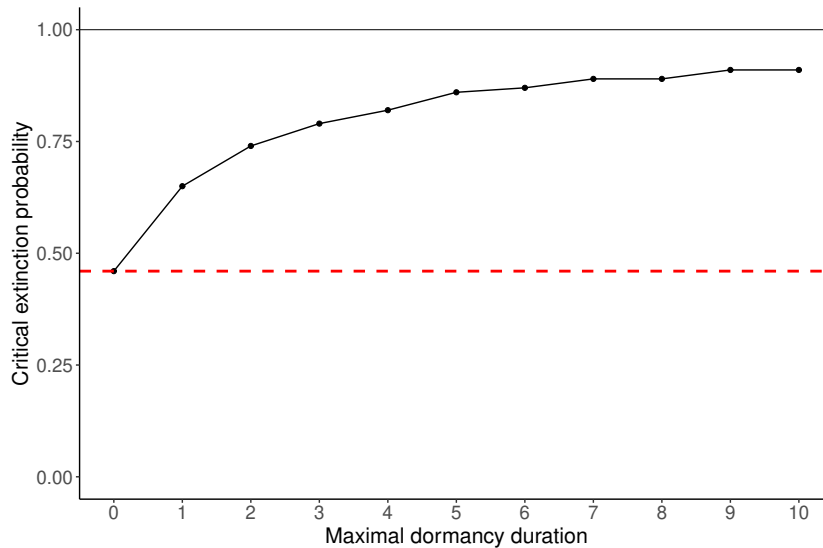


Figure 2: Approximate value of the critical extinction probability  $p_{crit}(H)$  as a function of the maximal dormancy duration  $H$ . The red dashed line indicates  $p_{crit}(0)$ , or in other words, the extinction probability above which (real) plants persistence without a seed bank is not possible. The continuous line indicates  $p_{crit}(\infty) = 1$  (see Remark 1.11). See the Appendix for details on the method used to compute  $p_{crit}(H)$ .

**Sketch of the proof.** The main idea behind the proof is the following. First, we use the coupling between the  $k$ -parent WFSB metapopulation process and the BOA process introduced earlier. This coupling ensures that each occupied patch in the  $k$ -parent WFSB metapopulation process is reachable in the BOA process. We can then focus on the study of the simpler BOA process, since the fact that no sites are reachable in the BOA process will then guarantee that the  $k$ -parent WFSB metapopulation



process has become empty. Using results from percolation theory, we obtain the existence of a critical extinction probability  $p_{crit}(H)$  for the BOA process, such that:

- For all  $p > p_{crit}(H)$ , the BOA process goes extinct in finite time almost surely ;
- For all  $p < p_{crit}(H)$ , the probability that the BOA process goes extinct in finite time is (strictly) less than 1.

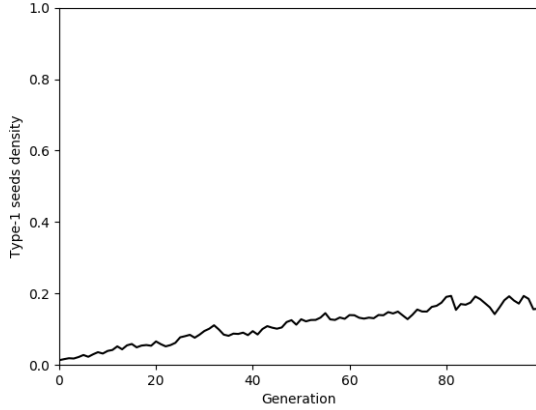
We conclude by using the fact that the  $k$ -parent WFSB metapopulation process is embedded in the BOA process.

The biological interpretation of this theorem is the following. For each maximal dormancy duration  $H \in \mathbb{N}$ , there exists a critical extinction probability  $p_{crit}(H)$  above which any metapopulation evolving according to a  $k$ -parent WFSB metapopulation process of maximal dormancy duration  $H$  will almost surely go extinct in finite time, no matter how quickly plants can invade a patch initially empty (which is quantified by  $k$  and to a lesser extent  $c$ ). In particular, no metapopulation without a seed bank can persist if the patch extinction probability is above  $p_{crit}(0)$ .  $p_{crit}(H)$  is increasing with  $H$ , so the ability to form a seed bank can potentially allow population persistence and expansion in highly disturbed fragmented environments. See Figure 2 for approximate values for  $p_{crit}(H)$ , computed using the method presented in the Appendix. Numerical simulations show the existence of parameter sets  $(M, H, g, c, p)$  with  $H > 0$  and  $p > p_{crit}(0)$  for which population persistence is indeed possible (see Figure 3). Since the  $k$ -parent occupancy process converges to the BOA process, which does not go extinct with positive probability if  $p < p_{crit}(H)$ , the critical extinction probability  $p_{crit}(H)$  we obtain is optimal, in the sense that it is not possible to obtain a better upper bound on the critical extinction probability for the  $k$ -parent WFSB metapopulation process which depends only on  $H$ , and not also on one of the other parameters.

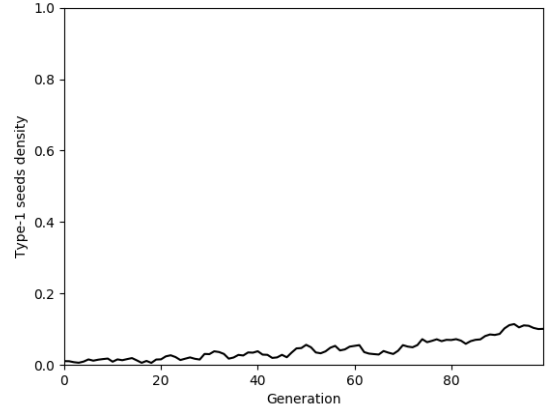
*Remark 1.11.* In the limiting case  $H = \infty$ , seeds never expire, and the number of reachable patches in the BOA process cannot decrease, implying that  $p_{crit}(\infty) = 1$ .

## 2 Proof of the convergence of the $k$ -parent occupancy process to the BOA process

The goal of this section is to show that the  $k$ -parent occupancy process converges to the BOA process in the sense of Theorem 1.9, that is, when both  $M \rightarrow +\infty$  and  $k \rightarrow +\infty$ , but with  $k$  increasing "faster" than  $M$ . In order to do so, we first focus on the post-transition period, and take an initial



(a)  $p = 0.5$  and  $H = 1$



(b)  $p = 0.7$  and  $H = 3$

Figure 3: Plant metapopulation expansion for extinction probabilities  $p_{crit}(0) < p < p_{crit}(H)$ , and for a maximal dormancy duration  $H \neq 0$ . The values taken by the other parameters are  $M = 100$ ,  $g = 0.5$ ,  $c = 0.05$  and  $k = 25$ . Initially, 5 consecutive patches contained  $gM = 50$  type 1 seeds, and all the other seed bank compartments were empty. Since only the first 100 generations were considered, the simulation was performed on a torus of 200 patches, and the density of type 1 seeds was computed over these 200 patches.

condition allowing us to skip the transition period. Then, we explain how to adapt the proof to take into account more general initial conditions.

First, we set some notation. We use the notation from Theorem 1.9 throughout this section. For all  $M \geq 2$  and  $h \in \mathbb{N}$ , let  $(\xi^{(M),n}, h^{(M),n})_{n \in \mathbb{N}}$  be the  $\lceil M^\alpha \rceil$ -parent WFSB metapopulation process with parameters  $(M, H, g, c, p)$  associated to the  $\lceil M^\alpha \rceil$ -parent occupancy process  $(O^{(M),n}, h^{(M),n})_{n \in \mathbb{N}}$ . For every  $n \in \mathbb{N}$ ,  $i \in \mathbb{Z}$  and  $h \in \mathbb{N}$ , let

$$E_{i,h}^{(M),n} := \{j \in \llbracket 1, M \rrbracket : h_{i,j}^{(M),n} = h\}$$

be the set of all seed bank compartments in patch  $i$  containing seeds of age  $h$  at the beginning of generation  $n$ , and let  $G_i^{(M),n}$  be the set of all seed bank compartments in patch  $i$  which contain the seeds germinating at the beginning of generation  $n$ .

Let  $(a_n)_{n \in \mathbb{N}}$  be the sequence defined by

$$a_0 = 0 \text{ and } \forall n \in \mathbb{N}, a_{n+1} = 3a_n + 1,$$

and for all  $0 < \epsilon < 1$ , let

$$W_h^{(M),\epsilon} := \left[ \lfloor gM \rfloor \left( \left( 1 - \frac{\lfloor gM \rfloor}{M} \right)^h - \epsilon a_h \right), \lfloor gM \rfloor \left( \left( 1 - \frac{\lfloor gM \rfloor}{M} \right)^h + \epsilon a_h \right) \right].$$

Recalling that  $(O^{(M),\infty,n}, h^{(M),\infty,n})_{n \in \mathbb{N}}$  is the BOA process associated to  $(\xi^{(M),n}, h^{(M),n})_{n \in \mathbb{N}}$ , we in-

roduce the event

$$\text{DiffBOA}_i^{(M),n} := \left\{ O_i^{(M),n} \neq O_i^{(M),\infty,n} \text{ or } h_i^{(M),n} \neq h_i^{(M),\infty,n} \right\}.$$

Given the initial condition for the associated BOA process, we already know that for  $n = 0$  the BOA process associated to  $(\xi^{(M),n}, h^{(M),n})_{n \in \mathbb{N}}$  is equal to the occupancy process associated to the same process. Therefore,

$$\mathbb{P} \left( \bigcap_{i \in \mathbb{Z}} \overline{\text{DiffBOA}_i^{(M),0}} \right) = 1.$$

Moreover, as only neighbouring sites can send colonizing seeds, if type 1 seeds are initially in a finite number of patches, then it is also the case after any arbitrary finite duration. More specifically, if we set

$$i_{min}^0 := \min\{i \in \mathbb{Z} : O_i^\infty = 1\} \text{ and } i_{max}^0 := \max\{i \in \mathbb{Z} : O_i^\infty = 1\},$$

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

$$i_{min}^{n+1} := i_{min}^n - 1 \text{ and } i_{max}^{n+1} := i_{max}^n + 1,$$

then the only patches which can potentially contain type 1 seeds after  $n$  generations are the patches with index  $i \in \llbracket i_{min}^n, i_{max}^n \rrbracket$ . In other words, for all  $M \geq 2$ ,  $n \in \mathbb{N}$  and  $i \in \mathbb{Z} \setminus \llbracket i_{min}^n, i_{max}^n \rrbracket$ , we have  $O_i^{(M),n} = O_i^{(M),\infty,n} = 0$  and  $h_i^{(M),n} = h_i^{\infty,n}$ , so that  $\overline{\text{DiffBOA}_i^{(M),n}}$  holds a.s. Therefore, for all  $N \in \mathbb{N}$ ,

$$\mathbb{P} \left( \bigcup_{n=0}^N \bigcup_{i \in \mathbb{Z}} \text{DiffBOA}_i^{(M),n} \right) = \mathbb{P} \left( \bigcup_{n=1}^N \bigcup_{i \in \llbracket i_{min}^n, i_{max}^n \rrbracket} \text{DiffBOA}_i^{(M),n} \right).$$

In all that follows, in order to ease notation, for all  $n \in \mathbb{N}$ , we set  $I_n = \llbracket i_{min}^n, i_{max}^n \rrbracket$

Let  $M \geq 2$  such that  $\lfloor gM \rfloor > 1$ , and  $N \geq 1$ . We recall that deviations from the BOA process can occur in the following cases.

1. Type 1 plants are present in a patch, but are never chosen as potential parents.
2. The seed bank contains non-expired type 1 seeds, but none of them germinate during the generation we consider.
3. Some type 1 seeds did enter the seed bank less than  $H + 1$  generations ago, but all of them already germinated, and the seed bank is now empty.

If the initial condition  $(\xi^{(M),0}, h^{(M),0})$  satisfies the extra hypothesis:

- (IC1) All the viable seeds of the same age and in the same patch are of the same type, or in formula,

$$\forall i \in \mathbb{Z}, \forall j_1, j_2 \in \llbracket 1, M \rrbracket, h_{i,j_1}^{(M),0} = h_{i,j_2}^{(M),0} \leq H \implies \xi_{i,j_1}^{(M),0} = \xi_{i,j_2}^{(M),0};$$

- (IC2) For all  $h \in \llbracket 0, H \rrbracket$ , the proportion of age  $h$  seeds is roughly equal to  $g(1-g)^h$ , in the sense that

$$\text{Card}(E_{i,h}^{(M),0}) \in W_h^{(M),\epsilon}$$

for  $\epsilon$  appropriately chosen, then the three deviation cases are covered by the following more general cases:

- (D1) There exists  $n \in \mathbb{N}$  and  $i \in I_{n+1}$  such that one of the plants in patches  $\{i-1, i, i+1\}$  is not chosen as a potential parent by at least one seed bank compartment in patch  $i$  during generation  $n$ .
- (D3) There exists  $n \in \mathbb{N}$  and  $i \in I_n$  such that  $h_i^{(M),n} \leq H$  but no seed of age  $h_i^{(M),n}$  germinates during generation  $n$ .
- (D3) There exists  $n \in \mathbb{N}$  and  $i \in I_n$  such that  $h_i^{(M),n} \leq H$  but  $\text{Card}\left(E_{i,h_i^{(M),n}}^{(M),n}\right) \notin W_{h_i^{(M),n}}^{(M),\epsilon}$ .

Indeed, initially, all the viable seeds of the same age have the same type, and this stays true until (D1) occurs. If (D1) did not occur yet before generation  $n$ , then for all  $i \in I_n$ , all seeds of age  $h_i^{(M),n}$  are of the same type. They are not necessarily type 1 seeds, but if the patch is not empty, then they are all type 1 seeds. Therefore, if the patch is not empty and if  $\epsilon$  is such that

$$\lfloor gM \rfloor \left( \left(1 - \lfloor gM \rfloor M^{-1}\right)^{h_i^{(M),n}} - \epsilon a_{h_i^{(M),n}} \right) > 1,$$

then

$$\text{Card}\left(E_{i,h_i^{(M),n}}^{(M),n}\right) \in W_{h_i^{(M),n}}^{(M),\epsilon}$$

implies that the patch contains type 1 seeds which entered less than  $H+1$  generations ago and did not germinate yet, and

$$\text{Card}\left(E_{i,h_i^{(M),n}}^{(M),n} \cap G_i^{(M),n}\right) > 0$$

implies that at least one type 1 seed germinates during generation  $n$ .

Formally, let  $\epsilon > 0$  be such that for all  $M \geq 2$  such that  $(1 - \lfloor gM \rfloor M^{-1})^H - \lfloor gM \rfloor^{-1} > 0$ ,

$$\epsilon < a_H^{-1} \left( \left(1 - \lfloor gM \rfloor M^{-1}\right)^H - \lfloor gM \rfloor^{-1} \right).$$

For all  $n \in \mathbb{N}$  and  $i \in \mathbb{Z}$ , we define the following events corresponding respectively to (D1), (D2) and (D3):

$\text{Par}_i^{(M),n} := \{\text{There exists a plant in patches } \{i-1, i, i+1\} \text{ which is not chosen as a potential parent by at least one seed bank compartment in patch } i \text{ during generation } n\},$

$\text{ErrF}_i^{(M),n} := \left\{ h_i^{(M),n} \leq H \text{ and } \text{Card} \left( E_{i,h_i^{(M),n}}^{(M),n} \right) \notin W_{h_i^{(M),n}}^{(M),\epsilon} \right\},$

and  $\text{ErrG}_i^{(M),n} := \left\{ h_i^{(M),n} \leq H \text{ and } \text{Card} \left( E_{i,h_i^{(M),n}}^{(M),n} \cap G_i^{(M),n} \right) = 0 \right\}.$

We assume that the initial condition  $(\xi^{(M),0}, h^{(M),0})$  satisfies conditions (IC1) and (IC2). Then,

$$\mathbb{P} \left( \bigcup_{n=1}^N \bigcup_{i \in I_n} \text{DiffBOA}_i^n \right) \leq \mathbb{P} \left( \bigcup_{n=0}^{N-1} \bigcup_{i \in I_{n+1}} \text{Par}_i^{(M),n} \cup \text{ErrG}_i^{(M),n} \cup \text{ErrF}_i^{(M),n+1} \right).$$

*Remark 2.1.* If  $(\xi^{(M),0}, h^{(M),0})$  does not satisfy condition (IC1) or (IC2), then we show at the end of the section that after a transition period of  $H+1$  generations,  $(\xi^{(M),H+1}, h^{(M),H+1})$  satisfies conditions (IC1) and (IC2) with high probability. We can then restart the process from  $(\xi^{(M),H+1}, h^{(M),H+1})$  and conclude.

In order to shorten the notation, let  $\text{ErrGF}_i^{(M),n}$  be the event defined as

$$\text{ErrGF}_i^{(M),n} := \text{ErrG}_i^{(M),n} \cup \text{ErrF}_i^{(M),n+1}.$$

Since the choice of the potential parents does not depend on their types and is independent from one patch (or generation) to another, we have

$$\begin{aligned} & \mathbb{P} \left( \bigcup_{n=0}^{N-1} \bigcup_{i \in I_{n+1}} \text{Par}_i^{(M),n} \cup \text{ErrGF}_i^{(M),n} \right) \\ & \leq \sum_{n=0}^{N-1} \sum_{i \in I_{n+1}} \mathbb{P} \left( \text{Par}_i^{(M),n} \right) + \mathbb{P} \left( \bigcup_{n=0}^{N-1} \bigcup_{i \in I_{n+1}} \text{ErrGF}_i^{(M),n} \middle| \left( \overline{\text{Par}_i^{(M),n'}} \right)_{\substack{0 \leq n' \leq N-1 \\ i \in I_{n'+1}}} \right) \end{aligned} \quad (2.1)$$

$$\begin{aligned} & \leq \sum_{n=0}^{N-1} \left( i_{max}^0 - i_{min}^0 + 1 + 2(n+1) \right) \mathbb{P} \left( \text{Par}_{i_{min}^0}^{(M),0} \right) + \mathbb{P} \left( \bigcup_{i \in I_1} \text{ErrGF}_i^{(M),0} \middle| \left( \overline{\text{Par}_i^{(M),n'}} \right)_{\substack{0 \leq n' \leq N-1 \\ i \in I_{n'+1}}} \right) \\ & \quad + \sum_{n=1}^{N-1} \mathbb{P} \left( \bigcup_{i \in I_{n+1}} \text{ErrGF}_i^{(M),n} \middle| \left( \overline{\text{Par}_i^{(M),n'}} \right)_{\substack{0 \leq n' \leq N-1 \\ i \in I_{n'+1}}} \cap \left( \overline{\text{ErrGF}_i^{(M),n'}} \right)_{\substack{0 \leq n' \leq n-1 \\ i \in I_{n'+1}}} \right) \\ & \leq \left[ N \left( i_{max}^0 - i_{min}^0 + 1 \right) + N(N+1) \right] \mathbb{P} \left( \text{Par}_{i_{min}^0}^{(M),0} \right) \end{aligned} \quad (2.2)$$

$$+ \mathbb{P} \left( \bigcup_{i \in I_1} \text{ErrGF}_i^{(M),0} \left| \left( \overline{\text{Par}}_i^{(M),n'} \right)_{\substack{0 \leq n' \leq N-1 \\ i \in I_{n'+1}}} \right. \right) \quad (2.3)$$

$$+ \sum_{n=1}^N \mathbb{P} \left( \bigcup_{i \in I_{n+1}} \text{ErrGF}_i^{(M),n} \left| \left( \overline{\text{Par}}_i^{(M),n'} \right)_{\substack{0 \leq n' \leq N-1 \\ i \in I_{n'+1}}} \cap \left( \overline{\text{ErrGF}}_i^{(M),n'} \right)_{\substack{0 \leq n' \leq n-1 \\ i \in I_{n'+1}}} \right. \right). \quad (2.4)$$

Bounding the quantity in (2.2) from above is straightforward, and is carried out in Section 2.1. The main obstacle to the study of the two other terms stems from the fact that all the events considered are linked together by  $(h^{(M),n})_{n \in \mathbb{N}}$ . However, we can circumvent this problem by working conditionally on  $h^{(M),n}$ . Indeed, given  $h_i^{(M),n}$ , since  $h_i^{(M),n+1}$  can only be equal to 0 or  $h_i^{(M),n} + 1$ ,

$$\begin{aligned} \text{ErrF}_i^{n+1} &:= \left\{ \text{Card} \left( E_{i, h_i^{(M),n+1}}^{(M),n+1} \right) \notin W_{h_i^{(M),n+1}}^{(M),\epsilon} \text{ and } h_i^{(M),n+1} \leq H \right\} \\ &\subseteq \left\{ \text{Card} \left( E_{i,0}^{(M),n+1} \right) \notin W_0^{(M),\epsilon} \text{ and } 0 \leq H \right\} \\ &\quad \cup \left\{ \text{Card} \left( E_{i, h_i^{(M),n} + 1}^{(M),n+1} \right) \notin W_{h_i^{(M),n} + 1}^{(M),\epsilon} \text{ and } h_i^{(M),n} + 1 \leq H \right\} \\ &= \left\{ \text{Card} \left( E_{i, h_i^{(M),n+1}}^{(M),n+1} \right) \notin W_{h_i^{(M),n+1}}^{(M),\epsilon} \text{ and } h_i^{(M),n} + 1 \leq H \right\} \end{aligned}$$

since  $W_0^{(M),\epsilon} = \{[gM]\} = \{\text{Card}(E_{i,0}^{(M),n+1})\}$ . Therefore,

$$\text{ErrGF}_i^n \subseteq \{h_i^{(M),n} \leq H\} \cap \left( \left\{ \text{Card} \left( E_{i, h_i^{(M),n}}^{(M),n} \cap G_i^{(M),n} \right) = 0 \right\} \cup \left\{ \text{Card} \left( E_{i, h_i^{(M),n+1}}^{(M),n+1} \right) \notin W_{h_i^{(M),n+1}}^{(M),\epsilon} \right\} \right). \quad (2.5)$$

In order to use (2.5), in Section 2.2, we establish an upper bound on

$$\mathbb{P} \left( \left\{ \text{Card} \left( E_{i,h}^{(M),n} \cap G_i^{(M),n} \right) = 0 \right\} \cup \left\{ \text{Card} \left( E_{i, h+1}^{(M),n+1} \right) \notin W_{h+1}^{(M),\epsilon} \right\} \mid \left\{ \text{Card} \left( E_{i,h}^{(M),n} \right) \in W_h^{(M),\epsilon} \right\} \right) \quad (2.6)$$

for all  $h \in \llbracket 0, H \rrbracket$ , which does not depend on the value of  $h$ . In Section 2.3, we use it, combined with the upper bound on (2.2) from Section 2.1, in order to complete the proof of Theorem 1.9.

## 2.1 Upper bound on $\mathbb{P}(\text{Par}_i^{(M),n})$

We set  $c^* = \min(c, 1 - 2c)$ . The goal of this section is to show the following lemma.

**Lemma 2.2.** *For all  $M \geq 2$ ,  $n \in \mathbb{N}$  and  $i \in \mathbb{Z}$ ,*

$$\mathbb{P} \left( \text{Par}_i^{(M),n} \right) \leq 3g^2 M^2 \exp \left( M^\alpha \ln \left( 1 - \frac{c^*}{gM} \right) \right).$$

A direct consequence of this lemma is the fact that since  $\alpha > 1$ ,

$$\mathbb{P}\left(\text{Par}_i^{(M),n}\right) \xrightarrow{M \rightarrow +\infty} 0.$$

*Proof.* Assume that  $c^* = c$ . Let  $\widetilde{\text{Par}}_i^{(M),n}$  be the event: {"The first seed which germinated in patch  $i+1$  was not chosen as a potential parent by the first seed bank compartment in patch  $i$  to be refilled"}. Then,

$$\mathbb{P}\left(\text{Par}_i^{(M),n}\right) \leq 3[gM]^2 \mathbb{P}\left(\widetilde{\text{Par}}_i^{(M),n}\right).$$

Indeed,  $\text{Par}_i^{(M),n}$  is the event {"at least one plant in one of the patches  $\{i-1, i, i+1\}$  is not chosen as a potential parent in order to refill at least one seed bank compartment in patch  $i$ "}. There exists  $3[gM]^2$  pairs "plant not chosen in patch  $i-1, i$  or  $i+1$ /seed bank compartment in patch  $i$ ", and as  $c^* = c$ , plants in patches  $i-1$  and  $i+1$  have less chances of being chosen as potential parents than plants in patch  $i$ .

Then, each one of the  $\lceil M^\alpha \rceil$  potential parents chosen to refill the first seed bank compartment in patch  $i$  is *not* the first plant of patch  $i+1$  with probability

$$1 - \frac{c^*}{[gM]} \leq 1 - \frac{c^*}{gM}.$$

Hence,

$$\begin{aligned} \mathbb{P}\left(\widetilde{\text{Par}}_i^{(M),n}\right) &\leq \left(1 - \frac{c^*}{gM}\right)^{\lceil M^\alpha \rceil} \\ &\leq \left(1 - \frac{c^*}{gM}\right)^{M^\alpha} \\ \text{and} \quad \mathbb{P}\left(\text{Par}_i^{(M),n}\right) &\leq 3g^2 M^2 \exp\left(M^\alpha \ln\left(1 - \frac{c^*}{gM}\right)\right). \end{aligned}$$

If  $c^* \neq c$ , then we can directly adapt this proof defining instead the event  $\widetilde{\text{Par}}_i^{(M),n}$  as the event {"The first seed which germinated in the patch  $i$  (instead of the patch  $i+1$ ) was not chosen as a potential parent by the first seed bank compartment in patch  $i$  to be refilled"}.  $\square$

## 2.2 Upper bound on (2.6)

In this section, we show an upper bound on

$$\mathbb{P}\left(\left\{\text{Card}\left(E_{i,h}^{(M),n} \cap G_i^{(M),n}\right) = 0\right\} \cup \left\{\text{Card}\left(E_{i,h+1}^{(M),n+1}\right) \notin W_{h+1}^{(M),\epsilon}\right\} \mid \left\{\text{Card}\left(E_{i,h}^{(M),n}\right) \in W_h^{(M),\epsilon}\right\}\right)$$

for all  $h \in \llbracket 0, H \rrbracket$ . In other words, we study the probability that if the number of age  $h$  seeds in patch  $i$  at the beginning of generation  $n$  is roughly equal to  $g(1-g)^h M$ :

- either no age  $h$  seeds germinate during generation  $n$
- or the number of remaining age  $h$  seeds (which become age  $h+1$  seeds in the subsequent generation) is significantly different from  $g(1-g)^{h+1} M$ .

In order to do so, we first observe that if  $E \subseteq \llbracket 1, M \rrbracket$  is a non-empty strict subset of  $\llbracket 1, M \rrbracket$ , then for all  $n \in \mathbb{N}$  and  $i \in \mathbb{Z}$ ,  $\text{Card}(E \cap G_i^{(M),n})$  follows an hypergeometric law. Using the tail inequalities from [37] yields the following lemma.

**Lemma 2.3.** *Let  $M \geq 2$  be such that  $(1 - \lfloor gM \rfloor M^{-1})^H - \lfloor gM \rfloor^{-1} > 0$ . Let  $E \subseteq \llbracket 1, M \rrbracket$  be a non-empty strict subset of  $\llbracket 1, M \rrbracket$ . Then, for all  $n \in \mathbb{N}$ ,  $i \in \mathbb{Z}$  and  $0 < \epsilon < 1$ ,*

$$\begin{aligned} \mathbb{P} \left( \text{Card}(E \cap G_i^{(M),n}) \geq \text{Card}(E)(1 + \epsilon) \frac{\lfloor gM \rfloor}{M} \right) &\leq e^{-2\epsilon^2 \text{Card}(E)^2 \lfloor gM \rfloor M^{-2}} \\ \text{and } \mathbb{P} \left( \text{Card}(E \cap G_i^{(M),n}) \leq \text{Card}(E)(1 - \epsilon) \frac{\lfloor gM \rfloor}{M} \right) &\leq e^{-2\epsilon^2 \text{Card}(E)^2 \lfloor gM \rfloor M^{-2}} \end{aligned}$$

Using this lemma, we can obtain the following upper bound.

**Lemma 2.4.** *For all  $M \geq 2$  such that  $(1 - \lfloor gM \rfloor M^{-1})^H - \lfloor gM \rfloor^{-1} > 0$ , for all  $n \in \mathbb{N}$ ,  $i \in \mathbb{Z}$  and  $h \in \llbracket 0, H \rrbracket$ ,*

$$\begin{aligned} \mathbb{P} \left( \left\{ \text{Card} \left( E_{i,h}^{(M),n} \cap G_i^{(M),n} \right) = 0 \right\} \cup \left\{ \text{Card} \left( E_{i,h+1}^{(M),n+1} \right) \notin W_{h+1}^{(M),\epsilon} \right\} \mid \left\{ \text{Card} \left( E_{i,h}^{(M),n} \right) \in W_h^{(M),\epsilon} \right\} \right) \\ \leq 2e^{-2\epsilon^2 \lfloor gM \rfloor^3 ((1 - \lfloor gM \rfloor M^{-1})^H - \epsilon a_H)^2 M^{-2}}. \end{aligned}$$

*Proof.* Let  $h \in \llbracket 0, H \rrbracket$ ,  $n \in \mathbb{N}$  and  $i \in \mathbb{Z}$ . Assume that  $\text{Card}(E_{i,h}^{(M),n}) \in W_h^{(M),\epsilon}$ . We make the following observation.

1. If less than  $\lfloor gM \rfloor M^{-1} (1 + \epsilon) \text{Card}(E_{i,h}^{(M),n})$  age  $h$  seeds germinate during generation  $n$ , the number of remaining age  $h$  seeds is bounded from below by

$$\begin{aligned} &\lfloor gM \rfloor \left[ \left( 1 - \frac{\lfloor gM \rfloor}{M} \right)^h - \epsilon a_h - \frac{\lfloor gM \rfloor}{M} (1 + \epsilon) \left( \left( 1 - \frac{\lfloor gM \rfloor}{M} \right)^h + \epsilon a_h \right) \right] \\ &= \lfloor gM \rfloor \left[ \left( 1 - \frac{\lfloor gM \rfloor}{M} \right)^h \left( 1 - \frac{\lfloor gM \rfloor}{M} \right) - \epsilon a_h - \frac{\lfloor gM \rfloor}{M} \epsilon a_h - \frac{\lfloor gM \rfloor}{M} \epsilon \left( 1 - \frac{\lfloor gM \rfloor}{M} \right)^h - \frac{\lfloor gM \rfloor}{M} \epsilon^2 a_h \right] \\ &\geq \lfloor gM \rfloor \left[ \left( 1 - \frac{\lfloor gM \rfloor}{M} \right)^{h+1} - \epsilon(3a_h + 1) \right] \\ &= \lfloor gM \rfloor \left( \left( 1 - \frac{\lfloor gM \rfloor}{M} \right)^{h+1} - \epsilon a_{h+1} \right) \end{aligned}$$



by definition of  $(a_h)_{h \geq 0}$ . By Lemma 2.3, this event happens with probability bounded from below by

$$1 - e^{-2\epsilon^2 \lfloor gM \rfloor^3 ((1 - \lfloor gM \rfloor M^{-1})^h - \epsilon a_h)^2 M^{-2}} \geq 1 - e^{-2\epsilon^2 \lfloor gM \rfloor^3 ((1 - \lfloor gM \rfloor M^{-1})^H - \epsilon a_H)^2 M^{-2}}.$$

2. If more than  $\lfloor gM \rfloor M^{-1} (1 - \epsilon) \text{Card}(E_{i,h}^{(M),n})$  age  $h$  seeds germinate, the number of remaining age  $h$  seeds is bounded from above by

$$\begin{aligned} & \lfloor gM \rfloor \left[ \left(1 - \frac{\lfloor gM \rfloor}{M}\right)^h + \epsilon a_h - \frac{\lfloor gM \rfloor}{M} (1 - \epsilon) \left( \left(1 - \frac{\lfloor gM \rfloor}{M}\right)^h - \epsilon a_h \right) \right] \\ &= \lfloor gM \rfloor \left[ \left(1 - \frac{\lfloor gM \rfloor}{M}\right)^h \left(1 - \frac{\lfloor gM \rfloor}{M}\right) + \epsilon a_h + \frac{\lfloor gM \rfloor}{M} \epsilon a_h + \frac{\lfloor gM \rfloor}{M} \epsilon \left(1 - \frac{\lfloor gM \rfloor}{M}\right)^h - \frac{\lfloor gM \rfloor}{M} \epsilon^2 a_h \right] \\ &\leq \lfloor gM \rfloor \left[ \left(1 - \frac{\lfloor gM \rfloor}{M}\right)^{h+1} + \epsilon(3a_h + 1) \right] \\ &= \lfloor gM \rfloor \left( \left(1 - \frac{\lfloor gM \rfloor}{M}\right)^{h+1} + \epsilon a_{h+1} \right). \end{aligned}$$

Again by Lemma 2.3, this event happens with probability bounded from below by

$$1 - e^{-2\epsilon^2 \lfloor gM \rfloor^3 ((1 - \lfloor gM \rfloor M^{-1})^H - \epsilon a_H)^2 M^{-2}}.$$

Therefore,

$$\begin{aligned} & \mathbb{P} \left( \left\{ \text{Card} \left( E_{i,h}^{(M),n} \cap G_i^{(M),n} \right) = 0 \right\} \cup \left\{ \text{Card} \left( E_{i,h+1}^{(M),n+1} \right) \notin W_{h+1}^{(M),\epsilon} \right\} \mid \left\{ \text{Card} \left( E_{i,h}^{(M),n} \right) \in W_h^{(M),\epsilon} \right\} \right) \\ &\leq \mathbb{P} \left( \text{Card} \left( E_{i,h}^{(M),n} \cap G_i^{(M),n} \right) < \frac{\lfloor gM \rfloor}{M} (1 - \epsilon) \text{Card} \left( E_{i,h}^{(M),n} \right) \mid \left\{ \text{Card} \left( E_{i,h}^{(M),n} \right) \in W_h^{(M),\epsilon} \right\} \right) \\ &\quad + \mathbb{P} \left( \text{Card} \left( E_{i,h}^{(M),n} \cap G_i^{(M),n} \right) > \frac{\lfloor gM \rfloor}{M} (1 + \epsilon) \text{Card} \left( E_{i,h}^{(M),n} \right) \mid \left\{ \text{Card} \left( E_{i,h}^{(M),n} \right) \in W_h^{(M),\epsilon} \right\} \right) \\ &\leq 2e^{-2\epsilon^2 \lfloor gM \rfloor^3 ((1 - \lfloor gM \rfloor M^{-1})^H - \epsilon a_H)^2 M^{-2}}, \end{aligned}$$

and we can conclude.  $\square$

### 2.3 Proof of Theorem 1.9

In order to show Theorem 1.9, we first condition on  $h^{(M),n}$  in Eq. (2.4). In order to do so, we introduce the following notation. For all  $n < N \in \mathbb{N}$ , let  $\text{Cond}^{(M),n,N}$  be the event defined by

$$\text{Cond}^{(M),n,N} := \left( \overline{\text{Par}}_i^{(M),n'} \right)_{0 \leq n' \leq N-1, i \in I_{n'+1}} \cap \left( \overline{\text{ErrGF}}_i^{(M),n'} \right)_{0 \leq n' \leq N-1, i \in I_{n'+1}}. \quad (2.7)$$

Moreover, let  $\mathcal{V}^{(M),n,N}$  be the set of all possible values for  $(h_i^{(M),n})_{i \in I_{n+1}}$  given the initial condition and  $\text{Cond}^{(M),n,N}$ . That is,

$$\mathcal{V}^{(M),n} := \left\{ (h_i)_{i \in I_{n+1}} \in \mathbb{N}^{i_{max}^{n+1} - i_{min}^{n+1} + 1} : \mathbb{P} \left( \forall i \in I_{n+1}, h_i^{(M),n} = h_i \mid \text{Cond}^{(M),n,N} \right) > 0 \right\}.$$

**Lemma 2.5.** *For all  $M \geq 2$  such that  $(1 - \lfloor gM \rfloor M^{-1})^H - \lfloor gM \rfloor^{-1} > 0$ ,*

$$\begin{aligned} & \sum_{n=1}^N \mathbb{P} \left( \bigcup_{i \in I_{n+1}} \text{ErrGF}_i^{(M),n} \mid \text{Cond}^{(M),n,N} \right) \\ & \leq 2 \left( N(i_{max}^0 - i_{min}^0 + 3) + N(N+1) \right) e^{-2\epsilon^2 \lfloor gM \rfloor^3 ((1 - \lfloor gM \rfloor M^{-1})^H - \epsilon a_H)^2 M^{-2}}. \end{aligned}$$

*Proof.* Let  $n \in \llbracket 1, N \rrbracket$ . Then,

$$\begin{aligned} & \mathbb{P} \left( \bigcup_{i \in I_{n+1}} \text{ErrGF}_i^{(M),n} \mid \text{Cond}^{(M),n,N} \right) \\ & = \sum_{(h_i)_{i \in I_{n+1}} \in \mathcal{V}^{(M),n}} \mathbb{P} \left( \forall i \in I_{n+1}, h_i^{(M),n} = h_i \mid \text{Cond}^{(M),n,N} \right) \\ & \quad \times \mathbb{P} \left( \bigcup_{i \in I_{n+1}} \text{ErrGF}_i^{(M),n} \mid \text{Cond}^{(M),n,N} \cap \left\{ \forall i \in I_{n+1}, h_i^{(M),n} = h_i \right\} \right) \\ & \leq \sum_{(h_i)_{i \in I_{n+1}} \in \mathcal{V}^{(M),n}} \sum_{i \in I_{n+1}} \mathbb{P} \left( \forall i' \in I_{n+1}, h_{i'}^{(M),n} = h_{i'} \mid \text{Cond}^{(M),n,N} \right) \\ & \quad \times \mathbb{P} \left( \text{ErrGF}_i^{(M),n} \mid \text{Cond}^{(M),n,N} \cap \left\{ \forall i' \in I_{n+1}, h_{i'}^{(M),n} = h_{i'} \right\} \right). \end{aligned}$$

Moreover, for all  $i \in I_{n+1}$ , by (2.5), if we define the event  $\text{Cond}^{+, (M), n, N}$  as

$$\text{Cond}^{+, (M), n, N} := \text{Cond}^{(M), n, N} \cap \left\{ \forall i \in I_{n+1}, h_i^{(M),n} = h_i \right\},$$

then

$$\begin{aligned} & \mathbb{P} \left( \text{ErrGF}_i^{(M),n} \mid \text{Cond}^{(M),n,N} \cap \left\{ \forall i' \in I_{n+1}, h_{i'}^{(M),n} = h_{i'} \right\} \right) \\ & \leq \mathbb{1}_{\{h_i \leq H\}} \\ & \quad \times \mathbb{P} \left( \left\{ \text{Card} \left( E_{i, h_i}^{(M),n} \cap G_i^{(M),n} \right) = 0 \right\} \cup \left\{ \text{Card} \left( E_{i, h_i+1}^{(M),n+1} \right) \notin W_{h_i+1}^{(M),\epsilon} \right\} \mid \text{Cond}^{+, (M), n, N} \right) \\ & \leq \mathbb{P} \left( \left\{ \text{Card} \left( E_{i, h_i}^{(M),n} \cap G_i^{(M),n} \right) = 0 \right\} \cup \left\{ \text{Card} \left( E_{i, h_i+1}^{(M),n+1} \right) \notin W_{h_i+1}^{(M),\epsilon} \right\} \mid \text{Card} \left( E_{i, h_i}^{(M),n} \right) \in W_{h_i}^{(M),\epsilon} \right) \\ & \leq 2e^{-2\epsilon^2 \lfloor gM \rfloor^3 ((1 - \lfloor gM \rfloor M^{-1})^H - \epsilon a_H)^2 M^{-2}} \end{aligned}$$

by Lemma 2.4. Therefore,

$$\begin{aligned}
& \mathbb{P} \left( \bigcup_{i \in I_{n+1}} \text{ErrGF}_i^{(M),n} \middle| \text{Cond}^{(M),n,N} \right) \\
& \leq 2 \left( i_{max}^{n+1} - i_{min}^{n+1} + 1 \right) e^{-2\epsilon^2 [gM]^3 ((1 - [gM]M^{-1})^H - \epsilon a_H)^2 M^{-2}} \\
& \leq 2 \left( i_{max}^0 - i_{min}^0 + 1 + 2n + 2 \right) e^{-2\epsilon^2 [gM]^3 ((1 - [gM]M^{-1})^H - \epsilon a_H)^2 M^{-2}}
\end{aligned}$$

and

$$\begin{aligned}
& \sum_{n=1}^N \mathbb{P} \left( \bigcup_{i \in I_{n+1}} \text{ErrGF}_i^{(M),n} \middle| \text{Cond}^{(M),n,N} \right) \\
& \leq \sum_{n=1}^N \left( 2(2n + i_{max}^0 - i_{min}^0 + 3) \right) e^{-2\epsilon^2 [gM]^3 ((1 - [gM]M^{-1})^H - \epsilon a_H)^2 M^{-2}} \\
& \leq \left( 2N(N+1) + 2(i_0^{max} - i_0^{min})N + 6N \right) e^{-2\epsilon^2 [gM]^3 ((1 - [gM]M^{-1})^H - \epsilon a_H)^2 M^{-2}},
\end{aligned}$$

which allows us to conclude.  $\square$

Obtaining a similar result for the quantity in (2.3) does not require conditioning, since  $h^{(M),0}$  is determined by the initial condition.

**Lemma 2.6.** *For all  $M \geq 2$  such that  $(1 - [gM]M^{-1})^H - [gM]^{-1} > 0$ ,*

$$\mathbb{P} \left( \bigcup_{i \in I_1} \text{ErrGF}_i^{(M),0} \middle| \left( \overline{\text{Par}}_i^{(M),n'} \right)_{\substack{0 \leq n' \leq N-1 \\ i \in I_{n'+1}}} \right) \leq 2 \left( i_{max}^0 - i_{min}^0 + 3 \right) e^{-2\epsilon^2 [gM]^3 ((1 - [gM]M^{-1})^H - \epsilon a_H)^2 M^{-2}}.$$

*Proof.* By 2.5 and by Lemma 2.4, given the initial condition,

$$\begin{aligned}
& \mathbb{P} \left( \bigcup_{i \in I_1} \text{ErrGF}_i^{(M),0} \middle| \left( \overline{\text{Par}}_i^{(M),n'} \right)_{\substack{0 \leq n' \leq N-1 \\ i \in I_{n'+1}}} \right) \\
& \leq \sum_{i \in I_1} \mathbf{1}_{\{h_i^{(M),0} \leq H\}} \\
& \quad \times \mathbb{P} \left( \left\{ \text{Card} \left( E_{i,h_i^{(M),0}}^{(M),0} \cap G_i^{(M),0} \right) = 0 \right\} \cup \left\{ \text{Card} \left( E_{i,h_i^{(M),0+1}}^{(M),1} \right) \notin W_{h_i^{(M),0+1}}^{(M),\epsilon} \right\} \middle| \left( \overline{\text{Par}}_i^{(M),n'} \right)_{\substack{n' \leq N-1 \\ i \in I_{n'+1}}} \right) \\
& \leq \sum_{i \in I_1} \mathbb{P} \left( \left\{ \text{Card} \left( E_{i,h_i^{(M),0}}^{(M),0} \cap G_i^{(M),0} \right) = 0 \right\} \cup \left\{ \text{Card} \left( E_{i,h_i^{(M),0+1}}^{(M),1} \right) \notin W_{h_i^{(M),0+1}}^{(M),\epsilon} \right\} \right) \\
& \leq 2 \left( i_{max}^0 - i_{min}^0 + 3 \right) e^{-2\epsilon^2 [gM]^3 ((1 - [gM]M^{-1})^H - \epsilon a_H)^2 M^{-2}}
\end{aligned}$$

given the initial condition.  $\square$

We can now show Theorem 1.9 for our specific initial condition.

*Proof.* (Theorem 1.9, post-transition period) Let  $M \geq 2$  be such that  $(1 - \lfloor gM \rfloor M^{-1})^H - \lfloor gM \rfloor^{-1} > 0$ .

The result is clear for  $N = 0$ . For  $N \in \mathbb{N} \setminus \{0\}$ , we showed earlier that

$$\begin{aligned} & \mathbb{P} \left( \bigcap_{n=0}^N \left( \left\{ \forall i \in \mathbb{Z}, O_i^{(M),n} = O_i^{(M),\infty,n} \right\} \cap \left\{ \forall i \in \mathbb{Z}, h_i^{(M),n} = h_i^{(M),\infty,n} \right\} \right) \right) \\ &= 1 - \mathbb{P} \left( \bigcup_{n=0}^N \bigcup_{i \in \mathbb{Z}} \text{DiffBOA}_i^{(M),n} \right) \end{aligned}$$

which is of the form 1– Eq.(2.2) – Eq.(2.3) – Eq.(2.4). By Lemmas 2.2, 2.6 and 2.5, we can show that each of the three terms converges to 0 when  $M \rightarrow +\infty$ , allowing us to conclude.  $\square$

We now explain how to generalize the proof of Theorem 1.9 to a more general sequence of initial conditions not necessarily satisfying (IC1) and (IC2). Since  $(\xi^{(M),n}, h^{(M),n})_{n \in \mathbb{N}}$  is Markovian for all  $M \geq 2$ , it is sufficient to show that the process does not deviate from the BOA process during the transition period, and that at the end of this period,  $(\xi^{(M),H+1}, h^{(M),H+1})$  satisfies (IC1) and (IC2).

In order to do so, let  $n \in \llbracket 0, H \rrbracket$  be a generation from the transition period, and let  $i \in \mathbb{Z}$ . We distinguish three cases.

1. If  $O_i^\infty = 0$ , then by condition (A) patch  $i$  is initially empty, so all viable  $h_i^{\infty,0}$  seeds it contains are of type 0. We are then in the same situation as during the post-transition period.
2. If  $O_i^\infty = 1$  and  $h_i^{(M),n} > n - 1$ , then the age  $h_i^{(M),n}$  seeds in patch  $i$  were already present initially. Similarly as before and using condition (C), we can show that with high probability, the number of remaining type 1 seeds of age  $h_i^{(M),n}$  is roughly equal to  $g_i(1 - g)^n M$ , and at least one of them germinates during generation  $n$ .
3. If  $O_i^\infty = 1$  and  $h_i^{(M),n} \leq n + 1$ , then the age  $h_i^{(M),n}$  seeds in patch  $i$  were not present initially. By Lemma 2.2, they are all of the same type with high probability, and we are back to the case considered during the post-transition period.

### 3 Extinction threshold for the $k$ -parent WFSB metapopulation process

This section is devoted to the proof of Theorem 1.10, that is, to the proof of the existence of a critical extinction probability  $p_{crit}(H)$  depending only on the maximal dormancy duration  $H$ . In order to do so, we will first formalize the coupling between the  $k$ -parent WFSB metapopulation process and a BOA process. Then, we will explain how the issue of occupied patches in the BOA process can be seen as a percolation problem. We will conclude using a specific case of Eq.(4) in [24].

### 3.1 Coupling between the $k$ -parent WFSB metapopulation process and the BOA process

In all that follows, let  $(\xi, h) \in \mathcal{F}_M \times \mathcal{H}_M$ , let  $(\xi^n, h^n)_{n \in \mathbb{N}}$  be the  $k$ -parent WFSB metapopulation process with parameters  $(M, H, g, c, p)$  and initial condition  $(\xi, h)$ , and let  $(O^{k,n}, h^{k,n})_{n \in \mathbb{N}}$  be the associated  $k$ -parent occupancy process. In order to couple a BOA process to  $(\xi^n, h^n)_{n \in \mathbb{N}}$ , for all  $n \in \mathbb{N}^*$ , we denote by  $(\text{Ext}_i^n)_{i \in \mathbb{Z}}$  the extinction events used to define  $(\xi^n, h^n)$  given  $(\xi^{n-1}, h^{n-1})$ . In other words, for all  $n \in \mathbb{N}^*$  and  $i \in \mathbb{Z}$ ,  $\text{Ext}_i^n = 1$  if, and only if the patch  $i$  was extinct during the  $n$ -th generation. We then define the coupled BOA process  $(O^{\infty,n}, h^{\infty,n})_{n \in \mathbb{N}}$  as the BOA process with parameters  $(H, p)$  and initial condition  $(O^{k,0}, h^{k,0})$ , constructed using the extinction events  $(\text{Ext}_i^n)_{i \in \mathbb{Z}, n \in \mathbb{N}^*}$ : for all  $n \in \mathbb{N}$ ,  $(O^{\infty,n+1}, h^{\infty,n+1})$  is constructed using  $(O^{\infty,n}, h^{\infty,n})$  and the extinction events  $(\text{Ext}_i^{n+1})_{i \in \mathbb{Z}}$ . This coupling satisfies the following property, whose proof is postponed until later in this section for the sake of clarity.

**Proposition 3.1.** *For all  $n \in \mathbb{N}$  and  $i \in \mathbb{Z}$ ,*

$$O_i^{k,n} \leq O_i^{\infty,n} \text{ and } h_i^{k,n} \geq h_i^{\infty,n}.$$

Therefore, at any generation  $n \in \mathbb{N}$ , the set of patches which contain nonexpired seeds in the  $k$ -parent WFSB metapopulation process is included in the set of reachable patches in the BOA process (that is, patches  $i$  such that  $O_i^{\infty,n} \mathbb{1}_{\{h_i^{\infty,n} \leq H\}} = 1$ ). In particular, a consequence of this coupling is the following corollary.

**Corollary 3.2.** *For all  $n \in \mathbb{N}$ ,*

$$\mathbb{P} \left( 1 - \prod_{(i,j) \in \mathbb{Z} \times \llbracket 1, M \rrbracket} (1 - \mathbb{1}_{\{h_{i,j}^n \leq H\}} \xi_{i,j}^n) = 1 \right) \leq \mathbb{P} \left( 1 - \prod_{i \in \mathbb{Z}} (1 - \mathbb{1}_{\{h_i^{\infty,n} \leq H\}} O_i^{\infty,n}) = 1 \right).$$

*Proof.* Let  $n \in \mathbb{N}$ . By definition of the  $k$ -parent occupancy process, for all  $(i, j) \in \mathbb{Z} \times \llbracket 1, M \rrbracket$ ,

$$\xi_{i,j}^n \leq O_i^{k,n}.$$

Indeed, both  $\xi_{i,j}^n$  and  $O_i^{k,n}$  are  $\{0, 1\}$ -valued, and if  $\xi_{i,j}^n = 1$ , then  $O_i^{k,n} = 1$ .

Moreover, if  $O_i^{k,n} = 1$ , then  $h_i^{k,n}$  is the age of the youngest type 1 seed in patch  $i$ . Therefore, for all  $(i, j) \in \mathbb{Z} \times \llbracket 1, M \rrbracket$ , if  $\xi_{i,j}^n = 1$ , then  $h_{i,j}^n \geq h_i^{k,n}$ . We deduce that

$$\mathbb{1}_{\{h_{i,j}^n \leq H\}} \xi_{i,j}^n \leq \mathbb{1}_{\{h_{i,j}^{k,n} \leq H\}} O_i^{k,n}.$$

By Proposition 3.1, we obtain

$$\mathbb{1}_{\{h_{i,j}^n \leq H\}} \xi_{i,j}^n \leq \mathbb{1}_{\{h_i^{\infty,n} \leq H\}} O_i^{\infty,n}.$$

Taking the product over all  $(i, j) \in \mathbb{Z} \times \llbracket 1, M \rrbracket$  yields

$$\begin{aligned} 1 - \prod_{(i,j) \in \mathbb{Z} \times \llbracket 1, M \rrbracket} \left(1 - \mathbb{1}_{\{h_{i,j}^n \leq H\}} \xi_{i,j}^n\right) &\leq 1 - \prod_{(i,j) \in \mathbb{Z} \times \llbracket 1, M \rrbracket} \left(1 - \mathbb{1}_{\{h_i^{\infty,n} \leq H\}} O_i^{\infty,n}\right) \\ &\leq 1 - \prod_{i \in \mathbb{Z}} \left(1 - \mathbb{1}_{\{h_i^{\infty,n} \leq H\}} O_i^{\infty,n}\right). \end{aligned}$$

since all the terms of the product are  $\{0, 1\}$ -valued, and we can conclude.  $\square$

We now show Proposition 3.1.

*Proof.* (Proposition 3.1) We show the result by induction. For  $n = 0$ , since  $(O^{\infty,0}, h^{\infty,0}) = (O^{k,0}, h^{k,0})$ ,

$$\begin{aligned} 1 - \prod_{(i,j) \in \mathbb{Z} \times \llbracket 1, M \rrbracket} \left(1 - \mathbb{1}_{\{h_{i,j}^0 \leq H\}} \xi_{i,j}^0\right) &= 1 - \prod_{i \in \mathbb{Z}} \prod_{j \in \llbracket 1, M \rrbracket} \left(1 - \mathbb{1}_{\{h_{i,j}^0 \leq H\}} \xi_{i,j}^0\right) \\ &= 1 - \prod_{i \in \mathbb{Z}} \left(1 - \mathbb{1}_{\{h_i^{k,0} \leq H\}} O_i^{k,0}\right) \\ &= 1 - \prod_{i \in \mathbb{Z}} \left(1 - \mathbb{1}_{\{h_i^{\infty,0} \leq H\}} O_i^{\infty,n}\right), \end{aligned}$$

so the result is true for  $n = 0$ .

Let then  $n \in \mathbb{N}$ , and we assume that for all  $i \in \mathbb{Z}$ ,

$$O_i^{k,n} \leq O_i^{\infty,n} \text{ and } h_i^{k,n} \geq h_i^{\infty,n}.$$

Let  $i \in \mathbb{Z}$ . We first show that  $O_i^{k,n+1} \leq O_i^{\infty,n+1}$ . Since  $O_i^{k,n+1} \in \{0, 1\}$ , if  $O_i^{\infty,n+1} = 1$ , then  $O_i^{k,n+1} \leq O_i^{\infty,n+1}$ . Therefore, we assume  $O_i^{\infty,n+1} = 0$ . Notice that by definition of the BOA process,  $(O_i^{\infty,n})_{n \in \mathbb{N}}$  is an increasing sequence. Indeed, for all  $n \geq 0$  and  $i \in \mathbb{Z}$ ,  $O_i^{\infty,n+1}$  is set equal to  $O_i^{\infty,n}$  or 1, so if  $O_i^{\infty,n} = 1$ , then  $O_i^{\infty,n+1} \in \{O_i^{\infty,n}, 1\} = \{1\}$ , and for all  $n' \geq n$ , we have  $O_i^{\infty,n'} = 1$ . This means that  $O_i^{\infty,n+1} = 0$  implies  $O_i^{\infty,n} = 0$  and  $O_i^{k,n} = 0$ . Moreover, it also means that both neighbouring patches were either extinct or not reachable in generation  $n$ . We deduce

$$\begin{aligned} (1 - \text{Ext}_{i+1}^{n+1}) O_{i+1}^{\infty,n} \mathbb{1}_{\{h_{i+1}^{\infty,n} \leq H\}} &= 0 \\ \text{and } (1 - \text{Ext}_{i-1}^{n+1}) O_{i-1}^{\infty,n} \mathbb{1}_{\{h_{i-1}^{\infty,n} \leq H\}} &= 0. \end{aligned}$$

Therefore, by the induction hypothesis,

$$\begin{aligned} (1 - \text{Ext}_{i+1}^{n+1}) O_{i+1}^{k,n} \mathbb{1}_{\{h_{i+1}^{k,n} \leq H\}} &= 0 \\ \text{and } (1 - \text{Ext}_{i-1}^{n+1}) O_{i-1}^{k,n} \mathbb{1}_{\{h_{i-1}^{k,n} \leq H\}} &= 0, \end{aligned}$$

which means that the patches  $i - 1$  and  $i + 1$  are either extinct or containing only ghost type 0 seeds. Combined with the knowledge that  $O_i^{k,n} = 0$ , we obtain that  $O_i^{k,n+1} = 0$ .

We now have to show that  $h_i^{k,n+1} \geq h_i^{\infty,n+1}$ . Since  $h_i^{k,n} \geq h_i^{\infty,n}$  and since  $h_i^{k,n+1}$  (resp.  $h_i^{\infty,n+1}$ ) is either equal to  $h_i^{k,n} + 1$  (resp.  $h_i^{\infty,n} + 1$ ) or equal to 0, the only potential issue is when  $h_i^{k,n+1} = 0$ . Let us assume that  $h_i^{k,n+1} = 0$ . This means that new seeds were just produced, and implies that

$$1 - \prod_{i'=i-1}^{i+1} \left( 1 - (1 - \text{Ext}_{i'}^{n+1}) O_{i'}^{k,n} \mathbb{1}_{\{h_{i'}^{k,n} \leq H\}} \right) = 1,$$

i.e, that non-expired seeds were present in at least one of the patches  $\{i - 1, i, i + 1\}$ , and that at least one of these patches was not affected by an extinction event. Moreover, if

$$\prod_{i'=i-1}^{i+1} \left( 1 - (1 - \text{Ext}_{i'}^{n+1}) O_{i'}^{\infty,n} \mathbb{1}_{\{h_{i'}^{\infty,n} \leq H\}} \right) = 0,$$

then  $h_i^{\infty,n+1} = 0$ . Using the induction hypothesis yields

$$\begin{aligned} \prod_{i'=i-1}^{i+1} \left( 1 - (1 - \text{Ext}_{i'}^{n+1}) O_{i'}^{\infty,n} \mathbb{1}_{\{h_{i'}^{\infty,n} \leq H\}} \right) &\leq \prod_{i'=i-1}^{i+1} \left( 1 - (1 - \text{Ext}_{i'}^{n+1}) O_{i'}^{k,n} \mathbb{1}_{\{h_{i'}^{k,n} \leq H\}} \right) \\ &= 0, \end{aligned}$$

hence  $h_i^{\infty,n+1} = 0 = h_i^{k,n+1}$  and we can conclude.  $\square$

### 3.2 Percolation problem

In order to show Theorem 1.10, we now link the BOA process to a percolation problem. More specifically, we rephrase the question of which patches are reachable in the BOA process as an oriented site percolation problem. Indeed, we can see patch  $i \in \mathbb{Z}$  in generation  $n \in \mathbb{N}$  as the site  $(i, n)$  of the space  $\mathbb{Z} \times \mathbb{N}$ . Each site  $(i, n) \in \mathbb{Z} \times \mathbb{N}$  is *open* (the analog of *non-extinct* in the terminology of percolation) with probability  $1 - p$ , and *closed* (i.e, extinct) otherwise. Reachable patches can be seen as sites of the space  $\mathbb{Z} \times \mathbb{N}$  linked to a site of  $\mathbb{Z} \times \{0\}$  by a path of open sites

$$(i_0, n_0) = (i_0, 0) \longrightarrow (i_1, n_1) \longrightarrow \dots \longrightarrow (i_L, n_L) = (i, n)$$

such that  $O_{i_0}^{\infty,0} \times \mathbf{1}_{\{h_{i_0}^{\infty,n} \leq H\}} = 1$ ,  $i_1 \in \{i_0 - 1, i_0, i_0 + 1\}$ ,  $n_1 - n_0 \in \llbracket 1, H - h_{i_0}^{\infty,n} + 1 \rrbracket$ , and for all  $l \in \llbracket 2, L \rrbracket$ ,

$$i_l \in \{i_{l-1} - 1, i_{l-1}, i_{l-1} + 1\} \quad \text{and} \quad n_l - n_{l-1} \in \llbracket 1, H + 1 \rrbracket. \quad (3.1)$$

For all  $n \in \mathbb{N}$ , let  $S_n(p)$  be the set of all the sites  $(i, n)$  with  $i \in \mathbb{Z}$  that are connected to  $(0, 0)$  by a path of open sites satisfying  $i_1 \in \{i_0 - 1, i_0, i_0 + 1\}$ ,  $n_1 - n_0 \in \llbracket 1, H + 1 \rrbracket$  and (3.1). Equivalently, let  $(O^{\{0\},n}, h^{\{0\},n})_{n \in \mathbb{N}}$  be the BOA process with parameters  $(H, p)$  and initial condition satisfying:

1.  $O_0^{\{0\},0} = 1$  and  $h_0^{\{0\},0} = 0$ .
2. For all  $i \in \mathbb{Z} \setminus \{0\}$ ,  $O_i^{\{0\},0} = 0$  and  $h_i^{\{0\},0} = 0$ .

We can then define  $S_n(p)$  as

$$S_n(p) := \left\{ i \in \mathbb{Z} : O_i^{\{0\},n} \mathbf{1}_{\{h_i^{\{0\},n} \leq H\}} = 1 \right\}.$$

Under this notation, a direct consequence of Eq. (4) in [24] is the following proposition.

**Proposition 3.3.** *There exists a unique  $p_{crit}(H) \in (0, 1)$  such that*

$$\begin{aligned} \forall p \in [0, p_{crit}(H)), \mathbb{P}(\forall n \in \mathbb{N}, S_n(p) \neq \emptyset) &> 0 \\ \text{and } \forall p \in (p_{crit}(H), 1], \mathbb{P}(\forall n \in \mathbb{N}, S_n(p) \neq \emptyset) &= 0. \end{aligned}$$

What remains to show is that  $p_{crit}(H)$  is indeed the extinction threshold we are looking for.

### 3.3 Proof of Theorem 1.10

In order to prove Theorem 1.10, we make three observations. First, for all  $n \in \mathbb{N}$ , the event  $\{S_n(p) \neq \emptyset\}$  is the same as the event

$$\left\{ 1 - \prod_{i \in \mathbb{Z}} \left( 1 - O_i^{\{0\},n} \mathbf{1}_{\{h_i^{\{0\},n} \leq H\}} \right) = 1 \right\}.$$

Moreover, for all finite subsets  $\mathcal{L}$  of  $\mathbb{Z}$ , let  $(O^{\mathcal{L}}, h^{\mathcal{L}}) \in \mathcal{F}^\infty \times \mathcal{H}^\infty$  satisfy the two following conditions:

- For all  $i \in \mathcal{L}$ ,  $O_i^{\mathcal{L}} = 1$  and  $h_i^{\mathcal{L}} = 0$ .
- For all  $i \in \mathbb{Z} \setminus \mathcal{L}$ ,  $O_i^{\mathcal{L}} = 0$  and  $h_i^{\mathcal{L}} = 0$ .

Let also  $(O^{\mathcal{L},n}, h^{\mathcal{L},n})_{n \in \mathbb{N}}$  be the BOA process with parameters  $(H, p)$  and initial condition  $(O^{\mathcal{L}}, h^{\mathcal{L}})$ . That is,  $(O^{\mathcal{L},n}, h^{\mathcal{L},n})_{n \in \mathbb{N}}$  is the BOA process starting from the state where all the patches in  $\mathcal{L}$  are of type 1 and all the patches in  $\mathcal{L}^c$  of type 0. Notice that if  $\mathcal{L} = \{0\}$ , then the definition of  $(O^{\{0\},n}, h^{\{0\},n})_{n \in \mathbb{N}}$  matches the one given above. We then have the following result.



**Lemma 3.4.** For all finite subset  $\mathcal{L}$  of  $\mathbb{Z}$  and for all  $n \in \mathbb{N}$ ,

$$\mathbb{P} \left( 1 - \prod_{i \in \mathbb{Z}} \left( 1 - O_i^{\mathcal{L}, n} \mathbb{1}_{\{h_i^{\mathcal{L}, n} \leq H\}} \right) = 0 \right) \geq \mathbb{P} \left( 1 - \prod_{i \in \mathbb{Z}} \left( 1 - O_i^{\{0\}, n} \mathbb{1}_{\{h_i^{\{0\}, n} \leq H\}} \right) = 0 \right)^{\text{Card}(\mathcal{L})}.$$

This lemma gives a lower bound of the probability that no patches are reachable in at least  $n$  generations in the BOA process starting from the patches in  $\mathcal{L}$ , each one of them containing type 1 seeds of age 0. This lower bound involves the probability that no patches are reachable in at least  $n$  generations starting from *only one patch*, which is used in the definition of  $p_{\text{crit}}(H)$ .

*Proof.* Let  $n \in \mathbb{N}$  and let  $\mathcal{L}$  be a finite subset of  $\mathbb{Z}$ . First, we observe that if we couple all the BOA processes considered by constructing them using the same extinction events,

$$1 - \prod_{i \in \mathbb{Z}} \left( 1 - O_i^{\mathcal{L}, n} \mathbb{1}_{\{h_i^{\mathcal{L}, n} \leq H\}} \right) = 1 - \prod_{i' \in \mathcal{L}} \left[ \prod_{i \in \mathbb{Z}} \left( 1 - O_i^{\{i'\}, n} \mathbb{1}_{\{h_i^{\{i'\}, n} \leq H\}} \right) \right].$$

Indeed, each one of the reachable patches in the BOA process with initial conditions  $(O^{\mathcal{L}}, h^{\mathcal{L}})$  is connected by a path of nonextinct patches to a patch in  $\mathcal{L}$ , and so there exists  $i_0 \in \mathcal{L}$  such as the patch is also reachable in the BOA process with initial condition  $(O^{\{i_0\}}, h^{\{i_0\}})$ . We can then use the fact that all the quantities appearing in the product are  $\{0, 1\}$ -valued.

Moreover, for  $i_0, i_1 \in \mathcal{L}$  and again using our coupling, knowing that no patch is reachable in  $n$  generations starting from  $i_0$  increases the probability that no patch is reachable in  $n$  generations starting from  $i_1$ . Indeed, informally, the fact that no patch is reachable starting from  $i_0$  "blocks" some patches, which cannot be used by a path linking  $i_1$  to other patches. Therefore,

$$\begin{aligned} \mathbb{P} \left( \prod_{i \in \mathbb{Z}} \left( 1 - O_i^{\{i_1\}, n} \mathbb{1}_{\{h_i^{\{i_1\}, n} \leq H\}} \right) = 1 \mid \prod_{i \in \mathbb{Z}} \left( 1 - O_i^{\{i_0\}, n} \mathbb{1}_{\{h_i^{\{i_0\}, n} \leq H\}} \right) = 1 \right) \\ \geq \mathbb{P} \left( \prod_{i \in \mathbb{Z}} \left( 1 - O_i^{\{i_1\}, n} \mathbb{1}_{\{h_i^{\{i_1\}, n} \leq H\}} \right) = 1 \right), \end{aligned}$$

and hence for  $i_0 \in \mathcal{L}$ ,

$$\begin{aligned} \mathbb{P} \left( 1 - \prod_{i \in \mathbb{Z}} \left( 1 - O_i^{\mathcal{L}, n} \mathbb{1}_{\{h_i^{\mathcal{L}, n} \leq H\}} \right) = 0 \right) &= \mathbb{P} \left( 1 - \prod_{i' \in \mathcal{L}} \left[ \prod_{i \in \mathbb{Z}} \left( 1 - O_i^{\{i'\}, n} \mathbb{1}_{\{h_i^{\{i'\}, n} \leq H\}} \right) \right] = 0 \right) \\ &= \mathbb{P} \left( \bigcap_{i' \in \mathcal{L}} \left\{ \prod_{i \in \mathbb{Z}} \left( 1 - O_i^{\{i'\}, n} \mathbb{1}_{\{h_i^{\{i'\}, n} \leq H\}} \right) = 1 \right\} \right) \\ &\geq \mathbb{P} \left( \prod_{i \in \mathbb{Z}} \left( 1 - O_i^{\{i_0\}, n} \mathbb{1}_{\{h_i^{\{i_0\}, n} \leq H\}} \right) = 1 \right)^{\text{Card}(\mathcal{L})} \end{aligned}$$

$$= \mathbb{P} \left( \prod_{i \in \mathbb{Z}} \left( 1 - O_i^{\{0\},n} \mathbf{1}_{\{h_i^{\{0\},n} \leq H\}} \right) = 1 \right)^{\text{Card}(\mathcal{L})},$$

where the invariance by translation of the process is used to pass from the last but first to the last line.  $\square$

We recall that the  $k$ -parent occupancy process associated to  $(\xi^n, h^n)_{n \in \mathbb{N}}$  is denoted by  $(O^{k,n}, h^{k,n})_{n \in \mathbb{N}}$ . The coupling based on the extinction events also yields the following lemma.

**Lemma 3.5.** *Let  $\mathcal{L} \subset \mathbb{Z}$  be the set defined as*

$$\mathcal{L} := \left\{ i \in \mathbb{Z} : O_i^{k,0} = 1 \right\}.$$

Then,

$$\mathbb{P} \left( 1 - \prod_{i \in \mathbb{Z}} \left( 1 - O_i^{k,n} \mathbf{1}_{\{h_i^{k,n} \leq H\}} \right) = 0 \right) \geq \mathbb{P} \left( 1 - \prod_{i \in \mathbb{Z}} \left( 1 - O_i^{\mathcal{L},n} \mathbf{1}_{\{h_i^{\mathcal{L},n} \leq H\}} \right) = 0 \right).$$

Indeed, if  $(\xi^n, h^n)_{n \in \mathbb{N}}$  (hence  $(O^{k,n}, h^{k,n})_{n \in \mathbb{N}}$ ) and  $(O^{\mathcal{L},n}, h^{\mathcal{L},n})_{n \in \mathbb{N}}$  are constructed using the same extinction events, then all the patches occupied by the  $k$ -parent WFSB metapopulation process are also reachable by the BOA process  $(O^{\mathcal{L},n}, h^{\mathcal{L},n})_{n \in \mathbb{N}}$ . Here deviations from the BOA process  $(O^{\mathcal{L},n}, h^{\mathcal{L},n})_{n \in \mathbb{N}}$  can also occur if the youngest type 1 seeds in  $(\xi^0, h^0)$  are *not* of age 0, but older.

We can now prove Theorem 1.10.

*Proof.* (Theorem 1.10) Let  $p_{\text{crit}}(H)$  be given by Proposition 3.3. We assume that  $p > p_{\text{crit}}(H)$ . Let also  $n \in \mathbb{N}$ , and let  $\mathcal{L} \subset \mathbb{Z}$  be defined as in Lemma 3.5.

By Lemma 3.5,

$$\begin{aligned} \mathbb{P} \left( \forall i \in \mathbb{Z}, O_i^{k,n} \mathbf{1}_{\{h_i^{k,n} \leq H\}} = 0 \right) &= \mathbb{P} \left( 1 - \prod_{i \in \mathbb{Z}} \left( 1 - O_i^{k,n} \mathbf{1}_{\{h_i^{k,n} \leq H\}} \right) = 0 \right) \\ &\geq \mathbb{P} \left( 1 - \prod_{i \in \mathbb{Z}} \left( 1 - O_i^{\mathcal{L},n} \mathbf{1}_{\{h_i^{\mathcal{L},n} \leq H\}} \right) = 0 \right). \end{aligned}$$

Using Lemma 3.4, we obtain

$$\begin{aligned} \mathbb{P} \left( \forall i \in \mathbb{Z}, O_i^{k,n} \mathbf{1}_{\{h_i^{k,n} \leq H\}} = 0 \right) &\geq \mathbb{P} \left( 1 - \prod_{i \in \mathbb{Z}} \left( 1 - O_i^{\{0\},n} \mathbf{1}_{\{h_i^{\{0\},n} \leq H\}} \right) = 0 \right)^{\text{Card}(\mathcal{L})} \\ &= \mathbb{P} (S_n(p) = \emptyset)^{\text{Card}(\mathcal{L})}. \end{aligned}$$

Therefore,

$$\begin{aligned} \lim_{n \rightarrow +\infty} \mathbb{P} \left( \forall i \in \mathbb{Z}, O_i^{k,n} \mathbf{1}_{\{h_i^{k,n} \leq H\}} = 0 \right) &\geq \lim_{n \rightarrow +\infty} \mathbb{P}(S_n(p) = \emptyset)^{\text{Card}(\mathcal{L})} \\ &\geq 1 \end{aligned}$$

by Proposition 3.3, and we can conclude.  $\square$

## 4 Appendix - Computation of $p_{crit}(H)$

In this section, we briefly explain how to compute  $p_{crit}(H)$ , and how to implement this approach and obtain an approximation for  $p_{crit}(H)$ . The computation method is a direct adaptation of Section 3 in [12]. Our goal here is not to obtain very precise approximations, but rather to have a rough estimate of  $p_{crit}(H)$ , and use it to assess the impact of the presence of a seed bank on the extinction threshold.

We first introduce the following notation. For all  $i \in \mathbb{Z}$  and  $n \in \mathbb{N}$ , let  $U^{i,n}$  be a random variable such that  $U^{i,n} \sim \text{Unif}([0, 1])$ . We assume that all the random variables  $(U^{i,n})_{i \in \mathbb{Z}, n \in \mathbb{N}}$  are independent. For all  $p \in [0, 1]$ , let  $\mathcal{S}_p$  be the set defined as

$$\mathcal{S}_p := \left\{ (i, n) : i \in \mathbb{Z}, n \in \mathbb{N} \text{ and } U^{i,n} \geq p \right\}.$$

$\mathcal{S}_p$  can be interpreted as the set of patches which would be non-extinct, if the extinction probability was equal to  $p$ .

For all  $x, y \in \mathbb{Z}$ ,  $n^{(x)}, n^{(y)} \in \mathbb{N}$ ,  $H \in \mathbb{N}$  and  $p \in [0, 1]$ , we will say that  $(x, n^{(x)})$  is  $(H, p)$ -reachable from  $(y, n^{(y)})$ , and denote it as  $(y, n^{(y)}) \xrightarrow{(H,p)} (x, n^{(x)})$ , if there exists  $L \in \mathbb{N}$ ,  $x_0, x_1, \dots, x_L \in \mathbb{Z}$  and  $n_0, n_1, \dots, n_L \in \mathbb{N}$  such that:

1.  $x_0 = y$ ,  $n_0 = n^{(y)}$ ,  $x_L = x$  and  $n_L = n^{(x)}$ ,
2.  $\forall l \in \llbracket 1, L \rrbracket$ ,  $x_l \in \{x_{l-1} - 1, x_{l-1}, x_{l-1} + 1\}$  and  $1 \leq n_l - n_{l-1} \leq H + 1$ ,
3.  $\forall l \in \llbracket 1, L \rrbracket$ ,  $(x_l, n_l) \in \mathcal{S}_p$ .

In other words,  $(y, n^{(y)}) \xrightarrow{(H,p)} (x, n^{(x)})$  if there exists a path of open sites going from  $(y, n^{(y)})$  to  $(x, n^{(x)})$ , spending at most  $H$  generations in each patch.

Moreover, for all  $p \in [0, 1]$  and  $n \in \mathbb{N}$ , let  $\bar{\xi}_{\mathbf{n}}(H, p)$  be the set defined as:

$$\bar{\xi}_{\mathbf{n}}(H, p) := \left\{ x \in \mathbb{Z} : \exists h_x, h_y \in \llbracket 0, H \rrbracket, \exists y \in \mathbb{Z} \setminus (\mathbb{N} \setminus \{0\}), (y, h_y) \xrightarrow{(H,p)} (x, n + h_x) \right\},$$

and let  $\bar{r}_n(H, p) := \sup \bar{\xi}_n(H, p)$ .  $\bar{\xi}_n(H, p)$  is akin to the set of patches which are reachable in  $n$  generations in a BOA process with parameters  $(H, p)$ , but starting from an infinite number of patches.

A direct adaptation of Section 3 from [12] yields the following result.

**Lemma 4.1.** *For all  $H \in \mathbb{N}$ ,*

$$p_{crit}(H) := \max \left\{ p \in [0, 1] : \lim_{n \rightarrow +\infty} \frac{\bar{r}_n(H, p)}{n} \geq 0 \right\}.$$

Therefore, in order to compute  $p_{crit}(H)$ , it is possible to simulate the random variable  $\bar{r}_n(H, p)$  for a large value of  $n$  and for different values of  $p$ .

Let  $H \in \mathbb{N}$ . In order to obtain an approximation for  $p_{crit}(H)$ , we first define some approximations for  $\bar{\xi}_n(H, p)$  and  $\bar{r}_n(H, p)$ . Let  $p \in [0, 1]$ . For all  $x, y \in \llbracket -10500, 10500 \rrbracket$  and  $n^{(x)}, n^{(y)} \in \llbracket 0, 10000 \rrbracket$ , we will say that  $(x, n^{(x)})$  is *approximatively  $(H, p)$ -reachable from  $(y, n^{(y)})$* , and denote it as

$$(y, n^{(y)}) \xrightarrow{\text{Approx}(H, p)} (x, n^{(x)}),$$

if there exists  $L \in \mathbb{N}$ ,  $x_0, \dots, x_L \in \llbracket -10500, 10500 \rrbracket$  and  $n_0, \dots, n_L \in \llbracket 0, 10000 \rrbracket$  such that:

1.  $x_0 = y$ ,  $n_0 = n^{(y)}$ ,  $x_L = x$  and  $n_L = n^{(x)}$ .
2.  $\forall l \in \llbracket 1, L \rrbracket$ ,  $x_l \in \{x_{l-1} - 1, x_{l-1}, x_{l-1} + 1\}$  and  $1 \leq n_l - n_{l-1} \leq H + 1$ .
3.  $\forall l \in \llbracket 1, L \rrbracket$ , if  $x_l \neq -10500$ , then  $(x_l, n_l) \in \mathcal{S}_p$  and  $x_l \neq 10500$ .

Therefore, in the approximation, the paths linking two sites together have to remain in the domain  $\llbracket -10500, 10500 \rrbracket$ , with extra conditions at the border of the domain. Since the value of the quantity we are interested in depends on the presence of paths staying close to the centre of the domain, we can assume that the border conditions chosen will not affect the approximate value.

We then define

$$\text{Approx}(\bar{\xi}_n(H, p)) := \left\{ x \in \mathbb{Z} : \exists h_x, h_y \in \llbracket 0, H \rrbracket, \exists y \in \mathbb{Z} \setminus (\mathbb{N} \setminus \{0\}), (y, h_y) \xrightarrow{\text{Approx}(H, p)} (x, n + h_x) \right\},$$

and let  $\text{Approx}(\bar{r}_n(H, p)) := \sup \text{Approx}(\bar{\xi}_n(H, p))$ .

In order to compute an approximate value for  $p_{crit}(H)$ , we apply the following method, starting from  $p = 0.99$ .

1. We simulate the random variable  $\text{Approx}(\bar{r}_{10000}(H, p)) \times (10000)^{-1}$ .
2. If the value obtained is larger than  $-0.005$ , we take  $p_{crit}(H) = p$ .

3. Otherwise, we substitute  $p$  with  $p - 0.01$ , and restart at Step 1.

**Acknowledgements** The author would like to thank her PhD supervisors Nathalie Machon and Amandine Véber for helpful discussions about the model and for their comments on the manuscript. The author is also grateful to the two anonymous reviewers for their helpful and constructive comments and suggestions. This work was partly supported by the chaire program "Modélisation Mathématique et Biodiversité" of Veolia Environnement-Ecole Polytechnique-Museum National d'Histoire Naturelle-Fondation X.

## References

- [1] F. Austerlitz, B. Jung-Muller, B. Godelle, and P.-H. Gouyon. Evolution of coalescence times, genetic diversity and structure during colonization. *Theoretical Population Biology*, 51(2):148–164, 1997.
- [2] N.H. Barton, A.M. Etheridge, J. Kelleher, and A. Véber. Genetic hitchhiking in spatially extended populations. *Theoretical Population Biology*, 87:75–89, 2013.
- [3] C.C. Baskin and J.M. Baskin. *Seeds: ecology, biogeography, and, evolution of dormancy and germination*. Academic Press, 2014.
- [4] J. Blath, A. González Casanova, N. Kurt, and D. Spanò. The ancestral process of long-range seed bank models. *Journal of Applied Probability*, 50(3):741–759, 2013.
- [5] J. Blath, A. González Casanova, N. Kurt, and M. Wilke-Berenguer. A new coalescent for seed-bank models. *Annals of Applied Probability*, 26(2):857–891, 2016.
- [6] J. Blath and N. Kurt. Population genetic models of dormancy. *arXiv preprint arXiv:2012.00810*, 2020.
- [7] F. Boenkost, A. González Casanova, C. Pokalyuk, and A. Wakolbinger. Haldane's formula in Cannings models: The case of moderately strong selection. *arXiv preprint arXiv:2008.02225*, 2020.
- [8] F. Boenkost, A. González Casanova, C. Pokalyuk, and A. Wakolbinger. Haldane's formula in Cannings models: the case of moderately weak selection. *Electronic Journal of Probability*, 26:1–36, 2021.

- [9] B. Borgy, X. Reboud, N. Peyrard, R. Sabbadin, and S. Gaba. Dynamics of weeds in the soil seed bank: a hidden Markov model to estimate life history traits from standing plant time series. *PloS one*, 10(10):e0139278, 2015.
- [10] F. Cordero, S. Hummel, and E. Schertzer. General selection models: Bernstein duality and minimal ancestral structures. *arXiv preprint arXiv:1903.06731*, 2019.
- [11] F. den Hollander and G. Pederzani. Multi-colony Wright–Fisher with seed-bank. *Indagationes Mathematicae*, 28(3):637–669, 2017.
- [12] R. Durrett. Oriented percolation in two dimensions. *The Annals of Probability*, 12(4):999–1040, 1984.
- [13] R. Durrett and W.-T. L. Fan. Genealogies in expanding populations. *Annals of Applied Probability*, 26(6):3456–3490, 2016.
- [14] L. Fahrig. Effects of habitat fragmentation on biodiversity. *Annual review of ecology, evolution, and systematics*, 34(1):487–515, 2003.
- [15] M. Fenner. Ecology of seed banks. *Seed development and germination*, pages 507–528, 1995.
- [16] 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.
- [17] H. Fréville, R. Choquet, R. Pradel, and P.-O. Cheptou. Inferring seed bank from hidden Markov models: new insights into metapopulation dynamics in plants. *Journal of Ecology*, 101(6):1572–1580, 2013.
- [18] A. González Casanova and C. Smadi. On Lambda-Fleming–Viot processes with general frequency-dependent selection. *Journal of Applied Probability*, 57(4):1162–1197, 2020.
- [19] A. González Casanova and D. Spanò. Duality and fixation in Xi-Wright–Fisher processes with frequency-dependent selection. *Annals of Applied Probability*, 28(1):250–284, 2018.
- [20] N.J. Gotelli. Metapopulation models: the rescue effect, the propagule rain, and the core-satellite hypothesis. *The American Naturalist*, 138(3):768–776, 1991.
- [21] A. Greven, F. den Hollander, and M. Oomen. Spatial populations with seed-bank: well-posedness, duality and equilibrium. *arXiv preprint arXiv:2004.14137*, 2020.
- [22] O. Hallatschek and D.R. Nelson. Gene surfing in expanding populations. *Theoretical Population Biology*, 73(1):158–170, 2008.

- [23] I.A. Hanski, M.E. Gilpin, and D.E. McCauley. *Metapopulation biology*, volume 454. Elsevier, 1997.
- [24] I. Hartarsky and R. Szabó. Generalised oriented site percolation, probabilistic cellular automata and bootstrap percolation. *arXiv preprint arXiv:2103.15621*, 2021.
- [25] I. Kaj, S.M. Krone, and M. Lascoux. Coalescent theory for seed bank models. *Journal of Applied Probability*, pages 285–300, 2001.
- [26] M. Kimura and G.H. Weiss. The stepping stone model of population structure and the decrease of genetic correlation with distance. *Genetics*, 49(4):561, 1964.
- [27] A. Lambert and C. Ma. The coalescent in peripatric metapopulations. *Journal of Applied Probability*, 52(2):538–557, 2015.
- [28] J. T. Lennon, F. den Hollander, M. Wilke-Berenguer, and J. Blath. Principles of seed banks: complexity emerging from dormancy. *arXiv preprint arXiv:2012.00072*, 2020.
- [29] R. Levins. Some demographic and genetic consequences of environmental heterogeneity for biological control. *American Entomologist*, 15(3):237–240, 1969.
- [30] A. Louvet. The k-parent spatial Lambda-Fleming-Viot process as a stochastic measure-valued model for an expanding population. *arXiv preprint arXiv:2103.02902*, 2021.
- [31] A. Louvet, N. Machon, J.-B. Mihoub, and A. Robert. Detecting seed bank influence on plant metapopulation dynamics. *Methods in Ecology and Evolution*, 12(4):655–664, 2021.
- [32] R.H. MacArthur and E.O. Wilson. *The theory of island biogeography*. Princeton University Press, 1967.
- [33] A. Moilanen. Patch occupancy models of metapopulation dynamics: efficient parameter estimation using implicit statistical inference. *Ecology*, 80(3):1031–1043, 1999.
- [34] A. Moilanen. SPOMSIM: software for stochastic patch occupancy models of metapopulation dynamics. *Ecological modelling*, 179(4):533–550, 2004.
- [35] S. Peischl and L. Excoffier. Expansion load: recessive mutations and the role of standing genetic variation. *Molecular Ecology*, 24(9):2084–2094, 2015.
- [36] M. Pluntz, S. Le Coz, N. Peyrard, R. Pradel, R. Choquet, and P.-O. Cheptou. A general method for estimating seed dormancy and colonisation in annual plants from the observation of existing flora. *Ecology letters*, 21(9):1311–1318, 2018.

- [37] M. Skala. Hypergeometric tail inequalities: ending the insanity. *arXiv preprint arXiv:1311.5939*, 2013.
- [38] J. Taylor and A. Véber. Coalescent processes in subdivided populations subject to recurrent mass extinctions. *Electronic Journal of Probability*, 14:242–288, 2009.
- [39] J. Wakeley and N. Aliacar. Gene genealogies in a metapopulation. *Genetics*, 159(2):893–905, 2001.
- [40] S. Wright. Evolution in Mendelian populations. *Genetics*, 16:97–159, 1931.
- [41] D. Živković and A. Tellier. Germ banks affect the inference of past demographic events. *Molecular Ecology*, 21(22):5434–5446, 2012.