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

Apolline Louvet

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

Apolline Louvet *,[1]

[1]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

*Email adress : apolline.louvet@polytechnique.edu

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 during up to several decades depending on the species, without loosing 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 [39]. See e.g [27, 37, 38] 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, 40] 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]: 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]).

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]. Although this is reminiscent of the model introduced in [25], the main difference is that in our model, even though "real" individuals do come from parents living a bounded number of generations ago, individuals of unknown types may come from a parent living arbitrarily far ago in the past.

In order to bridge the gap between individual-based metapopulation models and SPOMs, we shall show that our metapopulation process can be embedded in a SPOM. Moreover, we shall prove that

under an appropriate scaling of the selection strength and patch population size, the individual-based metapopulation process converges to this SPOM. The convergence result will have two applications. First, from a theoretical viewpoint, it will show that a specific SPOM (or presence/absence-based model) is the scaling limit of an individual-based metapopulation model. Then, we shall 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 shall 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

We shall 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 shall 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 $M \in \mathbb{N}^*$, $H \in \mathbb{N}$, $k \in \mathbb{N}\backslash\{0,1\}$, $g \in (0,1)$, $c \in (0,1/2)$ 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 $j \in [\![1, M]\!]$. 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:

$$\mathcal{F}_M := \Big\{ (\xi_{i,j})_{i\in\mathbb{Z}, j\in[\![1,M]\!]} : \forall i,j \in \mathbb{Z} \times [\![1,M]\!], \xi_{i,j} \in \{0,1\}$$

$$\text{and } \mathrm{Card}\left(\{(i,j) \in \mathbb{Z} \times [\![1,M]\!] : \xi_{i,j} = 1\}\right) < +\infty\Big\},$$

$$\text{and } \mathcal{H}_M := \Big\{(h_{i,j})_{i\in\mathbb{Z}, j\in[\![1,M]\!]} : \forall (i,j) \in \mathbb{Z} \times [\![1,M]\!], h_{i,j} \in \mathbb{N}\Big\}.$$

Here $\mathbb{N} = \{0,1,2,...\}$ and $[\![1,M]\!] = \{1,...,M\}$.

$(\xi, h) \in \mathcal{F}_M \times \mathcal{H}_M$ corresponds to a metapopulation in which for all $(i,j) \in \mathbb{Z} \times [\![1,M]\!]$, the seed occupying the seed bank compartment (i,j) is of age $h_{i,j}$ and of type $\xi_{i,j} \times \mathbb{1}_{\{h_{i,j}\leq H\}}$. 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 (ξ, 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 (ξ^n, h^n) :*

1. *For each $i \in \mathbb{Z}$, we sample $\lfloor gM \rfloor$ different seed bank compartments $s_{i,1}, ..., s_{i,\lfloor gM\rfloor} \in [\![1,M]\!]$ uniformly at random in patch i.*

2. *Let $(\mathrm{Ext}_i)_{i\in\mathbb{Z}}$ be i.i.d $\{0,1\}$-valued random variables such that $\mathbb{P}(\mathrm{Ext}_1 = 1) = p$.*

3. *For all $(i,j) \in \mathbb{Z} \times [\![1,M]\!]$, if $j \notin \{s_{i,j'} : j' \in [\![1, \lfloor gM \rfloor]\!]\}$, we set $\xi^{n+1}_{i,j} = \xi^n_{i,j}$ and $h^{n+1}_{i,j} = h^n_{i,j} + 1$.*

4. *On the other hand, if $j \in \{s_{i,j'} : j' \in [\![1, \lfloor gM \rfloor]\!]\}$, we first set $h^{n+1}_{i,j} = 0$. Moreover, let $C_1, ..., C_k$ be i.i.d $\{-1,0,1\}$-valued random variables such that*

$$\mathbb{P}(C_1 = 1) = \mathbb{P}(C_1 = -1) = c.$$

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

$$\tilde{k}_l = \xi^n_{i+C_l,j_l} \times \mathbb{1}_{\{h^n_{i+C_l,j_l}\leq H\}}.$$

We conclude by setting $\xi^{n+1}_{i,j} = \max\{\tilde{k}_l : l \in [\![1,k]\!]\}$.

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.

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.

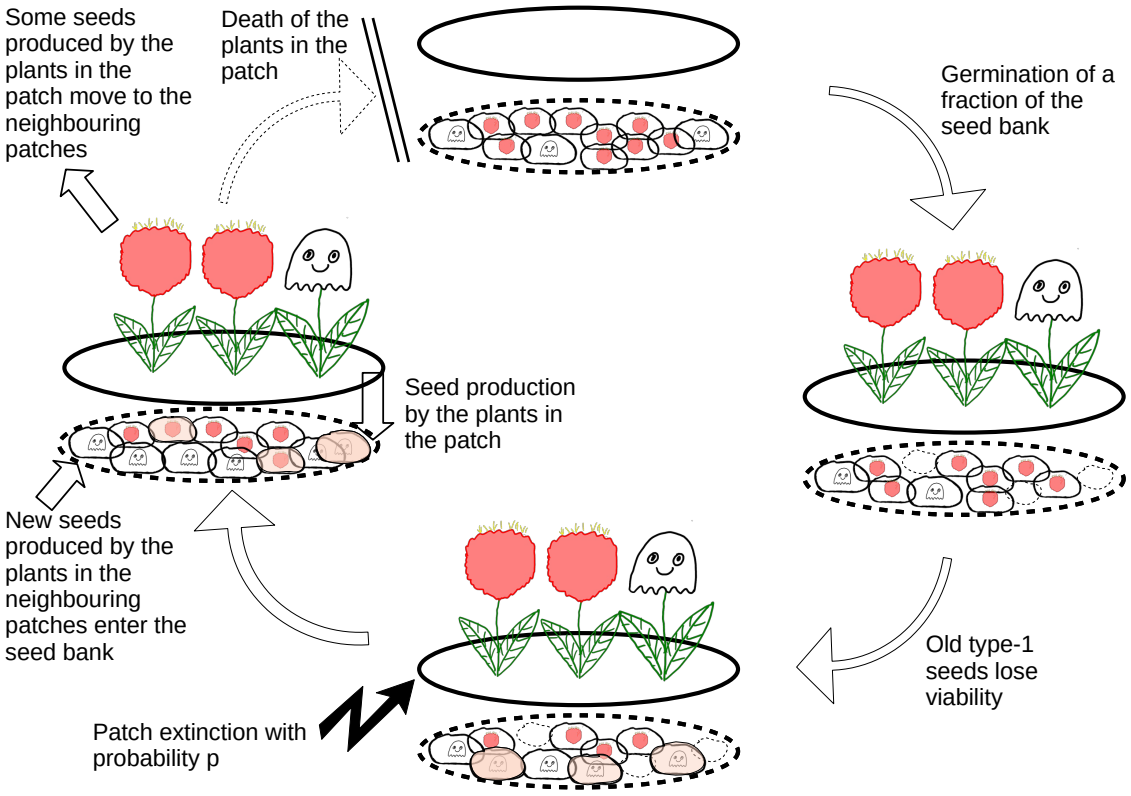


Figure 1: Illustration of the intra-patch dynamics of the k-parent WFSB metapopulation process. Here $M = 12$ and $\lfloor gM \rfloor = 3$. The double line in the top line of the figure indicates the starting time of a new generation.

In all that follows, we shall 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 colonisation probability*,
- p as the *patch extinction probability.*

Remark 1.2. 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 implicitely 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.

Remark 1.3. 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 1.4. 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 extinct, and which patches are potentially occupied. Indeed, for a patch to contain real seeds, it is not sufficient for it not to be 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 was 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:

$$\mathcal{F}^\infty := \{(O_i)_{i\in\mathbb{Z}} : \forall i \in \mathbb{Z}, O_i \in \{0,1\} \text{ and } \mathrm{Card}\left(\{i \in \mathbb{Z} : O_i = 1\}\right) < +\infty\}$$
$$\text{and } \mathcal{H}^\infty := \{(h_i)_{i\in\mathbb{Z}} : \forall i \in \mathbb{Z}, h_i \in \mathbb{N}\}.$$

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

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

   We do nothing during this step if $\mathrm{Ext}_i = 1$ or $O_i^{\infty,n} \times \mathbb{1}_{\{h_i^{\infty,n} \le H\}} = 0$.

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 shall say that patch $i \in \mathbb{Z}$ is reachable at generation $n \in \mathbb{N}$ if $O_i^{\infty,n} \times \mathbb{1}_{\{h_i^{\infty,n} \le H\}} = 1$.

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 1.6. *(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 (ξ, h) is defined as follows.

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

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

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

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

$$O_i^{k,n} := 1 - \prod_{j\in[\![1,M]\!]} \left(1 - \xi_{i,j}^n\right)$$

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

Under this setting, if the generation corresponding to the initial condition is numbered 0, $O_i^{k,n} = 1$ if and only if at the beginning of the (n+1)-th generation, 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:

$$O_i^{k,n} \times \mathbb{1}_{\{h_i^{k,n} \leq H\}} = 1.$$

Remark 1.7. 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 shall say that the BOA process *associated to* the k-parent WFSB metapopulation process with parameters (M, H, g, c, p) and initial condition (ξ, 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 the 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 $M \to +\infty$ and $k \to +\infty$ 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 shall 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 shall need the following constraints on the initial conditions of the processes.

Let $(O^\infty, h^\infty) \in (\mathcal{F}^\infty \times \mathcal{H}^\infty)$. For all $M \geq 2$, let $\xi^{(M)} \in \mathcal{F}_M$ be such that

$$\forall i \in \mathbb{Z}, \quad 1 - \prod_{j \in [\![1,M]\!]} \left(1 - \xi_{i,j}^{(M)}\right) = O_i^\infty.$$

Moreover, let $h^{(M)} \in \mathcal{H}_M$ be the random variable encoding the age of seeds, whose distribution satisfies the following conditions.

1. The vectors $(h_{i,j}^{(M)})_{j \in [\![1,M]\!]}$, $i \in \mathbb{Z}$ of the age of seeds are independent from one patch to another.
2. For all $i \in \mathbb{Z}$, the vector $(h_{i,j}^{(M)})_{j \in [\![1,M]\!]}$ of the ages of seeds in patch i is distributed according to the invariant distribution $\mu_{M,g}$ (defined in Section 2.1), conditional on

   (2A) For all $i \in \mathbb{Z}$, if $O_i^\infty = 1$, then $\min\{h_{i,j}^{(M)} : j \in [\![1, M]\!] \text{ and } \xi_{i,j}^{(M)} = 1\} = h_i^\infty$.

   (2B) For all $i \in \mathbb{Z}$ such that $O_i^\infty \times \mathbb{1}_{\{h_i^\infty \leq H\}} = 1$,

   $$\liminf_{M \to +\infty} \frac{1}{M} \sum_{i=1}^{M} \xi_{i,j}^{(M)} \times \mathbb{1}_{\{h_{i,j}^{(M)} = h_i^\infty\}} > 0 \quad \text{a.s.}$$

We shall say that the sequence of initial conditions $(\xi^{(M)}, h^{(M)})_{M \geq 2}$ satisfies **condition (C)**.

Intuitively, the first constraint ensure that the patches initially occupied for the k-parent WFSB metapopulation process and the BOA process are the same. Condition (2A) implies that in each patch, the youngest type-1 seeds (if present) have the same age for both processes, while condition (2B) means that the *youngest* type 1 seeds represent a significant portion of the seed bank, even in the large population limit. 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 1.8. *Let $\alpha > 1$. For all $M \geq 2$, let $\left(O^{(M),n}, h^{(M),n}\right)_{n \in \mathbb{N}}$ 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 $\left(O^{(M),\infty,n}, h^{(M),\infty,n}\right)_{n \in \mathbb{N}}$ be the BOA process associated to the same WFSB metapopulation process. Then, for all $N \in \mathbb{N}$,*

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

One of the biological interpretations of this result is that under 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.

1.2.3 Critical patch extinction probability

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

Theorem 1.9. *For all $H \in \mathbb{N}$, there exists $p_c(H) \in (0,1)$ such that for all $M \in \mathbb{N}^*$, $k \in \mathbb{N}\backslash\{0,1\}$, $g \in (0,1)$ and $c \in (0,1/2)$, for all $(\xi, h) \in \mathcal{F}_M \times \mathcal{H}_M$ and $p > p_c(H)$, 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 (ξ, h), then*

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

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.

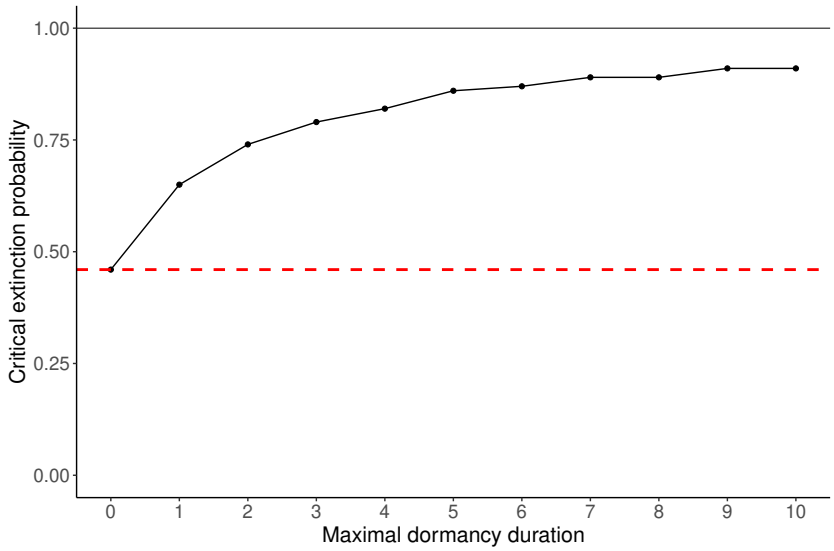
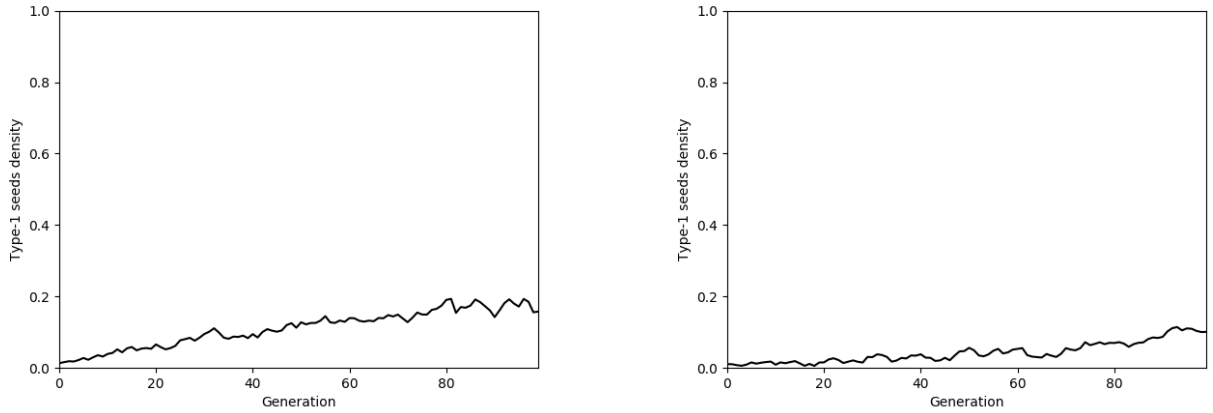


Figure 2: Approximate value of the critical extinction probability $p_c(H)$ as a function of the maximal dormancy duration H. The red dashed line indicates $p_c(0)$, or in other words, the extinction probability above which (real) plants persistence without a seed bank is not possible. See the Appendix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

The biological interpretation of this theorem is the following. For each maximal dormancy duration $H \in \mathbb{N}$, there exists a critical extinction probability $p_c(H)$ 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 $p_c(0)$. $p_c(H)$ 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 $p_c(H)$, 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 $p > p_c(0)$ for which population persistence is indeed possible (see Figure 3). Since the k-parent occupancy process converges to the BOA process, the critical extinction probability $p_c(H)$ we obtain is optimal, in the sense that it is not possible to obtain a lower critical extinction probability which depends only on H, and not also on one of the other parameters.



(a) p = 0.5 and H = 1

(b) p = 0.7 and H = 3

Figure 3: Plant metapopulation expansion for extinction probabilities $p_c(0) < p < p_c(H)$, 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.

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 1.8, that is, when both $M \to +\infty$ and $k \to +\infty$, but with k increasing "faster" than M. In order to do so, we shall first study the invariant distribution of the age of seeds in a patch being part of a metapopulation evolving according to the k-parent WFSB metapopulation process. The results will then be used in order to show that when $M \to +\infty$, with very high probability, a sample of $\lfloor gM \rfloor$ seeds contains at least one seed of age 0, 1,..., up to age H.

2.1 Invariant distribution of the age of seeds

In order to introduce the invariant distribution $\mu_{M,g}$, let us first set the following notation for the set of all partitions of $[\![1, M]\!]$:

$$\mathcal{E}_M := \{(E_v)_{v\in\mathbb{N}} : \forall v \in \mathbb{N}, E_v \subseteq [\![1, M]\!] \text{ and } \forall m \in [\![1, M]\!], \exists! v \in \mathbb{N}, m \in E_v\}.$$

An element $(E_v)_{v\in\mathbb{N}}$ of $\mathcal{E}_M$ will be interpreted as follows. Thinking of a single patch, for all $v \in \mathbb{N}$, E_v represents the list of all seed bank compartments in this patch which contain seeds of age v. The second condition in the definition of $\mathcal{E}_M$ simply says that the age of a seed is well-defined, that is, that a seed has one and only one age. Notice that all but a finite number of terms of the sequence $(E_v)_{v\in\mathbb{N}}$ are equal to the empty set, or equivalently, that there exists $V \in \mathbb{N}$ such that

$$\forall v \geq V, \quad E_v = \emptyset, \text{ or, in other words, } \bigcup_{v\in[\![0,V-1]\!]} E_v = [\![1, M]\!]. \tag{2.1}$$

We shall use the analogy between a subset of $[\![1, M]\!]$ and a subset of seeds/seed bank compartments, and between the choice of a subset of $[\![1, M]\!]$ and seed germination, throughout this section.

Consider a metapopulation having evolved according to the k-parent WFSB metapopulation process for a very long time, and take a patch in this metapopulation. The probability that the age of the seeds in the patch is given by $(E_v)_{v\in\mathbb{N}}$ can be computed using the following observation. For a seed to be of age v, the corresponding seed bank compartment needs to have been involved in a germination event $v + 1$ generations ago, but not during the more recent generations. Therefore:

1. First, going backwards in time, the first germination event has to affect all the seeds in E_0, but none of the seeds in $\cup_{v>0}E_v$. As $(E_v)_{v\in\mathbb{N}}$ is a partition of $[\![1, M]\!]$, this is only possible if $\text{Card}(E_0) = \lfloor gM \rfloor$.

   Therefore, if $\text{Card}(E_0) \neq \lfloor gM \rfloor$, then the event has a probability equal to 0 of occuring. If $\text{Card}(E_0) = \lfloor gM \rfloor$, since a germination event can be interpreted as the choice of a subset of $[\![1, M]\!]$ with cardinality $\lfloor gM \rfloor$, the probability that it is exactly the seed bank compartments in E_0 that were chosen is equal to

   $$\frac{1}{\binom{M}{\lfloor gM \rfloor}}.$$

   Considering that by convention, $\binom{a}{b}$ is equal to 0 if $a < b$ or $b < 0$, we can group the two cases together and say that the probability that the first germination affects all the seeds in E_0, but

none of the seeds in $\cup_{v>0} E_v$, is equal to

$$\delta_{\lfloor gM \rfloor, \mathrm{Card}(E_0)} = \frac{\binom{0}{\lfloor gM \rfloor - \mathrm{Card}(E_0)}}{\binom{M}{\lfloor gM \rfloor}} = \frac{\binom{M - \sum_{v \geq 0} \mathrm{Card}(E_v)}{\lfloor gM \rfloor - \mathrm{Card}(E_0)}}{\binom{M}{\lfloor gM \rfloor}}.$$

2. Then, the second germination event has to affect all the seeds in E_1, but none of the seeds in $\cup_{v>1} E_v$. This is not possible if $\mathrm{Card}(E_1) > \lfloor gM \rfloor$. If $\mathrm{Card}(E_1) < \lfloor gM \rfloor$, then the remaining $\lfloor gM \rfloor - \mathrm{Card}(E_1)$ seeds can be chosen in the seed bank compartments in E_0, since the seeds they contain are not the ones we observe at present.

   Therefore, there are

$$\binom{\mathrm{Card}(E_0)}{\lfloor gM \rfloor - \mathrm{Card}(E_1)} = \binom{M - \sum_{v \geq 1} \mathrm{Card}(E_v)}{\lfloor gM \rfloor - \mathrm{Card}(E_1)}$$

   ways of choosing a subset with cardinality $\lfloor gM \rfloor$ of $[\![1, M]\!]$ containing E_1 with the required constraint, and so the event we consider occurs with probability

$$\frac{\binom{M - \sum_{v \geq 1} \mathrm{Card}(E_v)}{\lfloor gM \rfloor - \mathrm{Card}(E_1)}}{\binom{M}{\lfloor gM \rfloor}}.$$

3. We repeat this for all $v \in \mathbb{N} \backslash \{0, 1\}$. The $(v+1)$-th germination event has to affect all the seeds in E_v, but none of the seeds in $\cup_{v'>v} E_{v'}$. This amounts to choosing a subset with cardinality $\lfloor gM \rfloor$ of $[\![1, M]\!]$, containing all E_v, and $\lfloor gM \rfloor - \mathrm{Card}(E_v)$ elements taken in $[\![1, M]\!] \backslash \cup_{v' \geq v} E_{v'}$. It occurs with probability

$$\frac{\binom{M - \sum_{v' \geq v} \mathrm{Card}(E_{v'})}{\lfloor gM \rfloor - \mathrm{Card}(E_v)}}{\binom{M}{\lfloor gM \rfloor}}.$$

Therefore, the probability that the age of the seeds in the patch is given by $(E_v)_{v \in \mathbb{N}}$ is equal to

$$\mu_{M,g}\left((E_v)_{v \in \mathbb{N}}\right) := \prod_{v \in \mathbb{N}} \frac{\binom{M - \sum_{v' \geq v} \mathrm{Card}(E_{v'})}{\lfloor gM \rfloor - \mathrm{Card}(E_v)}}{\binom{M}{\lfloor gM \rfloor}}.$$

Notice that by property (2.1), there exists $V \in \mathbb{N}$ such that for all $v \geq V$, $\mathrm{Card}(E_v) = 0$. Therefore, this product only contains a finite number of terms which are different from 1.

We shall say that $h = (h_{i,j})_{i \in \mathbb{Z}, j \in [\![1,M]\!]} \in \mathcal{H}_M$ is sampled according to the invariant distribution $\mu_{M,g}^{\otimes \mathbb{N}}$ if it is sampled as follows. Independently for all $i \in \mathbb{Z}$, we sample $(E_v^i)_{v \in \mathbb{N}} \in \mathcal{E}_M$, and for all $j \in [\![1, M]\!]$, $h_{i,j}$ is defined as the unique integer such that $j \in E^i_{h_{i,j}}$, or in other words, such that the seed in the seed bank compartment (i, j) is of age $h_{i,j}$. We shall often abuse notation and say that

such an h is sampled according to $\mu_{M,g}$.

We now assume that the age of the seeds in the focal patch is distributed according to the invariant distribution $\mu_{M,g}$, and given by $(E_v)_{v\in\mathbb{N}}$. We would like to show that out of the $\lfloor gM \rfloor$ seeds which germinate during the next generation, the proportion of seeds of age $h \in [\![0, H]\!]$ is roughly equal to $g \times (1-g)^H$. More specifically, we would like to show the following lemma.

Lemma 2.1. *Suppose the age of seeds is distributed according to $\mu_{M,g}$. Let $\epsilon > 0$. Let $(a_n)_{n\in\mathbb{N}}$ be the sequence defined by $a_0 = 0$ and $\forall n \in \mathbb{N}$, $a_{n+1} = 2a_n + 1$. For all $h \in [\![0, H]\!]$, we set:*

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

Then, for all $h \in [\![0, H]\!]$,

$$\mathbb{P}\left(\mathrm{Card}(E_h) \in W_h\right) \geq \prod_{h'=0}^{h-1} \left(1 - \left((1-g)^h \times g + \epsilon a_h\right)^{g\times\left((1-g)^h\times\lfloor gM\rfloor - \epsilon M a_h\right) - \epsilon M} - \left(1 - (1-g)^h \times \left(g - \frac{1}{M}\right) + \epsilon a_h\right)^{gM\left(1-g(1-g)^h\right) + M(1+a_h)\epsilon - 1}\right).$$

This lemma will be shown by induction, using the following technical lemma.

Lemma 2.2. *Let $E \subseteq [\![1, M]\!]$ be a non-empty strict subset of the seed bank compartments in the focal patch (or equivalently, of $[\![1, M]\!]$), and let $\epsilon > 0$. Let $G \subseteq [\![1, M]\!]$ be the random set (with cardinality $\lfloor gM \rfloor$) of all the seed bank compartments in the focal patch containing the seeds germinating during the next generation. Then,*

$$\mathbb{P}\left(\mathrm{Card}(E \cap G) \geq g \times \mathrm{Card}(E) + \epsilon M\right) \leq \left(\frac{\mathrm{Card}(E)}{M}\right)^{g\times\mathrm{Card}(E) - \epsilon M}$$
$$\mathbb{P}\left(\mathrm{Card}(E \cap G) \leq g \times \mathrm{Card}(E) - \epsilon M\right) \leq \left(1 - \frac{\mathrm{Card}(E)}{M}\right)^{gM - g\times\mathrm{Card}(E) + \epsilon M - 1}.$$

Proof. First, we assume that

$$0 < g \times \mathrm{Card}(E) - \epsilon M < g \times \mathrm{Card}(E) + \epsilon M < M.$$

In order to construct a sample of $\lfloor gM \rfloor$ seeds containing *at least* $g \times \mathrm{Card}(E) + \epsilon M$ seeds from E, one strategy, generating all possible samples, is to do as follows.

1. Choose a sequence of $\lceil g\mathrm{Card}(E) + \epsilon M \rceil$ seeds among the ones in E.
2. Choose a sequence of $\lfloor gM \rfloor - \lceil g\mathrm{Card}(E) + \epsilon M \rceil$ seeds among the $M - \lceil g\mathrm{Card}(E) + \epsilon M \rceil$ remaining ones.

3. Account for the fact that $(\lfloor gM \rfloor)!$ sequences yield the same sample.

Therefore, out of the $\binom{M}{\lfloor gM \rfloor}$ possible samples of $\lfloor gM \rfloor$ seeds, there are

$$\frac{1}{(\lfloor gM \rfloor)!} \times \frac{\mathrm{Card}(E)!}{(\mathrm{Card}(E) - \lceil g\mathrm{Card}(E) + \epsilon M \rceil)!} \times \frac{(M - \lceil g\mathrm{Card}(E) + \epsilon M \rceil)!}{(M - \lfloor gM \rfloor)!}$$

samples containing at least $g \times \mathrm{Card}(E) + \epsilon M$ seeds from E. Since the seed sample is chosen uniformly at random over all the possible ones,

$$\begin{aligned}
&\mathbb{P}\left(\mathrm{Card}(E \cap G) \geq g\mathrm{Card}(E) + \epsilon M\right) \\
&= \frac{1}{\binom{M}{\lfloor gM \rfloor}} \times \frac{1}{(\lfloor gM \rfloor)!} \times \frac{\mathrm{Card}(E)!}{(\mathrm{Card}(E) - \lceil g\mathrm{Card}(E) + \epsilon M \rceil)!} \times \frac{(M - \lceil g\mathrm{Card}(E) + \epsilon M \rceil)!}{(M - \lfloor gM \rfloor)!} \\
&= \frac{(M - \lfloor gM \rfloor)!}{M!} \times \frac{\mathrm{Card}(E)!}{(\mathrm{Card}(E) - \lceil g\mathrm{Card}(E) + \epsilon M \rceil)!} \times \frac{(M - \lceil g\mathrm{Card}(E) + \epsilon M \rceil)!}{(M - \lfloor gM \rfloor)!} \\
&= \prod_{i=0}^{\lceil g\mathrm{Card}(E) + \epsilon M \rceil - 1} \frac{\mathrm{Card}(E) - i}{M - i} \\
&\leq \left(\frac{\mathrm{Card}(E)}{M}\right)^{\lceil g\mathrm{Card}(E) + \epsilon M \rceil} \\
&\leq \left(\frac{\mathrm{Card}(E)}{M}\right)^{\lceil g\mathrm{Card}(E) - \epsilon M \rceil} \\
&\leq \left(\frac{\mathrm{Card}(E)}{M}\right)^{g\mathrm{Card}(E) - \epsilon M}
\end{aligned}$$

where the inequality on the 5^{th} line comes from the fact that $\mathrm{Card}(E) \times M^{-1} < 1$.

Then, it is possible to construct a sample of $\lfloor gM \rfloor$ seeds containing *at most* $g\mathrm{Card}(E) - \epsilon M$ seeds from E as follows:

1. Choose a sequence of $\lfloor gM \rfloor - \lfloor g\mathrm{Card}(E) - \epsilon M \rfloor$ seeds among the $M - \mathrm{Card}(E)$ ones in $[\![1, M]\!] \backslash E$.
2. Choose a sequence of $\lfloor g\mathrm{Card}(E) - \epsilon M \rfloor$ seeds among the $M - \lfloor gM \rfloor + \lfloor g\mathrm{Card}(E) - \epsilon M \rfloor$ remaining ones.
3. Account for the fact that $(\lfloor gM \rfloor)!$ sequences yield the same sample.

Again, it can be checked that this strategy generates all possible samples satisfying the desired condition. Similarly as before, we obtain

$$\begin{aligned}
&\mathbb{P}\left(\mathrm{Card}(E \cap G) \leq g\mathrm{Card}(E) - \epsilon M\right) \\
&= \frac{1}{\binom{M}{\lfloor gM \rfloor}} \times \frac{1}{(\lfloor gM \rfloor)!} \times \frac{(M - \mathrm{Card}(E))!}{(M - \mathrm{Card}(E) - \lfloor gM \rfloor + \lfloor g\mathrm{Card}(E) - \epsilon M \rfloor)!} \\
&\quad \times \frac{(M - \lfloor gM \rfloor + \lfloor g\mathrm{Card}(E) - \epsilon M \rfloor)!}{(M - \lfloor gM \rfloor)!}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{(M - \mathrm{Card}(E))!}{M!} \times \frac{(M - \lfloor gM \rfloor + \lfloor g\mathrm{Card}(E) + \epsilon M \rfloor)!}{(M - \mathrm{Card}(E) - \lfloor gM \rfloor + \lfloor g\mathrm{Card}(E) - \epsilon M \rfloor)!} \\
&= \prod_{i=0}^{\lfloor gM \rfloor - \lfloor g\mathrm{Card}(E) - \epsilon M \rfloor - 1} \frac{M - \mathrm{Card}(E) - i}{M - i} \\
&\leq \left(\frac{M - \mathrm{Card}(E)}{M} \right)^{\lfloor gM \rfloor - \lfloor g\mathrm{Card}(E) - \epsilon M \rfloor} \\
&\leq \left(1 - \frac{\mathrm{Card}(E)}{M} \right)^{gM - g\mathrm{Card}(E) + \epsilon M - 1}
\end{aligned}
$$

since $1 - \mathrm{Card}(E) \times M^{-1} < 1$.

This concludes the proof for the case $0 < g \times \mathrm{Card}(E) - \epsilon M < g \times \mathrm{Card}(E) + \epsilon M < M$. If $g \times \mathrm{Card}(E) + \epsilon M > M$, then the probability of sampling at least $g\mathrm{Card}(E) + \epsilon M$ seeds from E is equal to 0, and the upper bound remains valid. Similarly, if $g\mathrm{Card}(E) - \epsilon M < 0$, then the probability of sampling at most $g\mathrm{Card}(E) - \epsilon M$ is equal to 0 as well, which concludes the proof. □

We can now show Lemma 2.1. The proof relies on the observation that for all $h \in [\![1, H]\!]$, the number of age h seeds during the current generation is equal to the number of age $h-1$ seeds during the last generation, minus the number of such seeds which just germinated. But since the age of seeds is distributed according to the invariant distribution, the number of age $h-1$ seeds during the previous generation has the same distribution as the number of age $h-1$ seeds during the current generation.

Proof. (Lemma 2.1) First, if $h = 0$, then $\mathrm{Card}(E_0) = \lfloor gM \rfloor$ and

$$
W_0 = \Big[(1-g)^0 \times \lfloor gM \rfloor - 0, (1-g)^0 \times \lfloor gM \rfloor + 0 \Big] = \{ \lfloor gM \rfloor \}.
$$

Therefore, $\mathbb{P}(\mathrm{Card}(E_0) \in W_0) = 1$, and the result is true for $h = 0$.

In order to argue by induction, we observe that

$$
\begin{aligned}
& (1-g)^h \times \lfloor gM \rfloor - \epsilon M a_h - g \times \Big((1-g)^h \times \lfloor gM \rfloor + \epsilon M a_h \Big) - \epsilon M \\
& = (1-g)^{h+1} \times \lfloor gM \rfloor - \epsilon M \times (a_h(1+g) + 1) \\
& \geq (1-g)^{h+1} \times \lfloor gM \rfloor - \epsilon M \times (2a_h + 1) \\
& \geq (1-g)^{h+1} \times \lfloor gM \rfloor - \epsilon M a_{h+1} \\
\text{and} \quad & (1-g)^h \times \lfloor gM \rfloor + \epsilon M a_h - g \times \Big((1-g)^h \times \lfloor gM \rfloor - \epsilon M a_h \Big) + \epsilon M \\
& \leq (1-g)^{h+1} \times \lfloor gM \rfloor + \epsilon M \times (a_h \times (1+g) + 1) \\
& \leq (1-g)^{h+1} \times \lfloor gM \rfloor + \epsilon M \times (2a_h + 1) \\
& \leq (1-g)^{h+1} \times \lfloor gM \rfloor + \epsilon M a_{h+1}.
\end{aligned}
$$

Moreover, if $E \subseteq [\![1, M]\!]$ is a subset of $[\![1, M]\!]$ with cardinality $\mathrm{Card}(E) \in W_h$, we have the following properties.

1. If less than $g \times \mathrm{Card}(E) + \epsilon M$ seeds from E germinate, then the maximal number of seeds which germinate is bounded from above by

$$g \times \Big((1-g)^h \times \lfloor gM \rfloor + \epsilon M a_h \Big) + \epsilon M,$$

and the number of remaining seeds is bounded from below by

$$(1-g)^h \times \lfloor gM \rfloor - \epsilon M a_h - g \times \Big((1-g)^h \times \lfloor gM \rfloor + \epsilon M a_h \Big) - \epsilon M \geq (1-g)^{h+1} \times \lfloor gM \rfloor - \epsilon M a_{h+1}.$$

2. If more than $g \times \mathrm{Card}(E) - \epsilon M$ seeds from E germinate, then the minimal number of seeds which germinate is bounded from below by

$$g \times \Big((1-g)^h \times \lfloor gM \rfloor - \epsilon M a_h \Big) - \epsilon M,$$

and the number of remaining seeds is bounded from above by

$$(1-g)^h \times \lfloor gM \rfloor + \epsilon M a_h - g \times \Big((1-g)^h \times \lfloor gM \rfloor - \epsilon M a_h \Big) + \epsilon M \leq (1-g)^{h+1} \times \lfloor gM \rfloor + \epsilon M a_{h+1}.$$

Therefore, if

1. (Event 1) at the beginning of a given generation, the number C_h of age h seeds belongs to W_h,
2. (Event 2) during the generation, more than $gC_h - \epsilon M$ but less than $gC_h + \epsilon M$ age h seeds germinate,

then the number of remaining age h seeds, which is also the number of age $h+1$ seeds at the beginning of the next generation, belongs to W_{h+1}. Moreover, by Lemma 2.2 , the probability of Event 2 is bounded from below by

$$\begin{aligned}
\mathbb{P}(\text{Event 2}) &= 1 - \mathbb{P}\left(\{\text{more than } gC_h - \epsilon M \text{ age } h \text{ seeds germinate}\}^c \right. \\
&\qquad \left. \cup \{\text{less than } gC_h + \epsilon M \text{ age } h \text{ seeds germinate}\}^c \right) \\
&= 1 - \mathbb{P}\left(\{\text{less than } gC_h - \epsilon M \text{ age } h \text{ seeds germinate}\} \right. \\
&\qquad \left. \cup \{\text{more than } gC_h + \epsilon M \text{ age } h \text{ seeds germinate}\} \right) \\
&= 1 - \mathbb{P}\left(\{\text{less than } gC_h - \epsilon M \text{ age } h \text{ seeds germinate}\} \right)
\end{aligned}$$

$$
\begin{aligned}
&- \mathbb{P}\left(\{\text{more than } gC_h + \epsilon M \text{ age } h \text{ seeds germinate}\}\right) \\
&\geq 1 - \Big((1-g)^h \times g + \epsilon a_h\Big)^{g\times\left((1-g)^h\times\lfloor gM\rfloor - \epsilon M a_h\right)-\epsilon M} \\
&\quad - \left(1 - (1-g)^h \times \left(g - \frac{1}{M}\right) + \epsilon a_h\right)^{gM\left(1-g(1-g)^h\right)+M(1+a_h)\epsilon-1}.
\end{aligned}
$$

Let now $h \in [\![0, H-1]\!]$. We assume the induction property is true for h. Using the fact that the age of seeds is distributed according to the invariant distribution yields

$$
\begin{aligned}
\mathbb{P}\left(\mathrm{Card}(E_{h+1}) \in W_{h+1}\right) \geq\ & \mathbb{P}\left(\mathrm{Card}(E_h) \in W_h\right) \\
& \times \left(1 - \Big((1-g)^h \times g + \epsilon a_h\Big)^{g\times\left((1-g)^h\times\lfloor gM\rfloor - \epsilon M a_h\right)-\epsilon M}\right. \\
& \quad \left. - \left(1 - (1-g)^h \times \left(g - \frac{1}{M}\right) + \epsilon a_h\right)^{gM\left(1-g(1-g)^h\right)+M(1+a_h)\epsilon-1}\right)
\end{aligned}
$$

which concludes the proof. □

2.2 Proof of Theorem 1.8

In all that follows, let $\alpha > 1$, and let $(\xi^{(M)}, h^{(M)})_{M\geq 2} \in \mathcal{F}_M \times \mathcal{H}_M$ satisfy condition (C). In order to simplify the proof, we shall assume that for all $M \geq 2$ and for all $i \in \mathbb{Z}$,

$$
\forall j_1, j_2 \in [\![1, M]\!], h_{i,j_1}^{(M)} = h_{i,j_2}^{(M)} \Longrightarrow \xi_{i,j_1}^{(M)} = \xi_{i,j_2}^{(M)},
$$

or in other words, that in each patch, all the seeds produced during the same generation are of the same type. However, the proof can be generalized to the original case.

For all $M \geq 2$, 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) and initial condition $(\xi^{(M)}, h^{(M)})$. We denote the associated $\lceil M^\alpha\rceil$-parent occupancy process by $(O^{(M),n}, h^{(M),n})_{n\in\mathbb{N}}$, and the associated BOA process by $(O^{(M),\infty,n}, h^{(M),\infty,n})_{n\in\mathbb{N}}$.

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

$$
\begin{aligned}
& i_{min}^{M,0} := \min\left\{i \in \mathbb{Z} : \exists j \in [\![1, M]\!], \xi_{i,j}^{M,0} = 1\right\}, \\
\text{and} \qquad & i_{max}^{M,0} := \max\left\{i \in \mathbb{Z} : \exists j \in [\![1, M]\!], \xi_{i,j}^{M,0} = 1\right\},
\end{aligned}
$$

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

$$\begin{aligned} & i_{min}^{M,n+1} := i_{min}^{M,n} - 1, \\ \text{and} \qquad & i_{max}^{M,n+1} := i_{max}^{M,n} + 1, \end{aligned}$$

then the only patches which can potentially contain type 1 seeds after n generations are the patches of index $i \in [\![i_{min}^{M,n}, i_{max}^{M,n}]\!]$. In other words, for all $M \geq 2$, $n \in \mathbb{N}$ and $i \in \mathbb{Z} \backslash [\![i_{min}^{M,n}, i_{max}^{M,n}]\!]$,

$$\begin{aligned} & O_i^{(M),n} = O_i^{(M),\infty,n} = 0 \\ \text{and} \qquad & h_i^{(M),n} = h_i^{(M),\infty,n}. \end{aligned}$$

A consequence of this observation is the following lemma.

Lemma 2.3. *For all $M \geq 2$ and $N \in \mathbb{N}$,*

$$\begin{aligned} & \mathbb{P}\left(\cup_{n=0}^{N} \left(\left\{ \exists i \in \mathbb{Z}, O_i^{(M),n} \neq O_i^{(M),\infty,n} \text{ or } h_i^{(M),n} \neq h_i^{(M),\infty,n} \right\} \right) \right) \\ & \leq \left(N \times \left(i_{max}^{M,0} - i_{min}^{M,0} + 1 \right) + N(N+1) \right) \\ & \qquad \times \mathbb{P}\left(O_0^{(M),1} \neq O_0^{(M),\infty,1} \middle| \forall i' \in [\![-1, 1]\!], O_{i'}^{(M),0} = O_{i'}^{(M),\infty,0} \text{ and } h_{i'}^{(M),0} = h_{i'}^{(M),\infty,0} \right) \end{aligned}$$

Proof. Let $M \geq 2$ and $N \in \mathbb{N}$. First, we observe that by definition, for $n = 0$,

$$\mathbb{P}\left(\exists i \in \mathbb{Z}, O_i^{(M),0} \neq O_i^{(M),\infty,0} \text{ or } h_i^{(M),0} \neq h_i^{(M),\infty,0} \right) = 0.$$

Moreover, if we set

$$n_0 := \min \left\{ n \in \mathbb{N} : \exists i \in \mathbb{Z}, O_i^{(M),n} \neq O_i^{(M),\infty,n} \text{ or } h_i^{(M),n} \neq h_i^{(M),\infty,n} \right\},$$

if $n_0 \leq N$, then there exists $i \in \mathbb{Z}$ such that both $O_i^{(M),n_0} \neq O_i^{(M),\infty,n_0}$ and $h_i^{(M),n_0} \neq h_i^{(M),\infty,n_0}$. Therefore, we deduce

$$\begin{aligned} & \mathbb{P}\left(\cup_{n=0}^{N} \left(\left\{ \exists i \in \mathbb{Z}, O_i^{(M),n} \neq O_i^{(M),\infty,n} \text{ or } h_i^{(M),n} \neq h_i^{(M),\infty,n} \right\} \right) \right) \\ & = \sum_{n=1}^{N} \mathbb{P}\left(\left\{ \exists i \in \mathbb{Z}, O_i^{(M),n} \neq O_i^{(M),\infty,n} \text{ or } h_i^{(M),n} \neq h_i^{(M),\infty,n} \right\} \right. \\ & \qquad \left. \cap \left\{ \forall 0 \leq n' < n, \forall i \in \mathbb{Z}, O_i^{(M),n'} = O_i^{(M),\infty,n'} \text{ and } h_i^{(M),n} = h_i^{(M),\infty,n'} \right\} \right) \\ & \leq \sum_{n=1}^{N} \mathbb{P}\left(\left\{ \exists i \in \mathbb{Z}, O_i^{(M),n} \neq O_i^{(M),\infty,n} \right\} \middle| \forall i \in \mathbb{Z}, O_i^{(M),n-1} = O_i^{(M),\infty,n-1} \text{ and } h_i^{(M),n-1} = h_i^{(M),\infty,n-1} \right) \end{aligned}$$

$$\times \mathbb{P}\left(\left\{\forall 0 \leq n' < n, \forall i \in \mathbb{Z}, O_i^{(M),n'} = O_i^{(M),\infty,n'} \text{ and } h_i^{(M),n} = h_i^{(M),\infty,n'}\right\}\right)$$

$$\leq \sum_{n=1}^{N} \mathbb{P}\left(\left\{\exists i \in \mathbb{Z}, O_i^{(M),n} \neq O_i^{(M),\infty,n}\right\} \middle| \forall i \in \mathbb{Z}, O_i^{(M),n-1} = O_i^{(M),\infty,n-1} \text{ and } h_i^{(M),n-1} = h_i^{(M),\infty,n-1}\right)$$

$$\leq \sum_{n=1}^{N} \mathbb{P}\left(\bigcup_{i=i_{min}^{M,n}}^{i_{max}^{M,n}} \left\{O_i^{(M),n} \neq O_i^{(M),\infty,n}\right\} \middle| \forall i \in \mathbb{Z}, O_i^{(M),n-1} = O_i^{(M),\infty,n-1} \text{ and } h_i^{(M),n-1} = h_i^{(M),\infty,n-1}\right)$$

$$\leq \sum_{n=1}^{N} \sum_{i=i_{min}^{M,n}}^{i_{max}^{M,n}} \mathbb{P}\left(O_i^{(M),n} \neq O_i^{(M),\infty,n} \middle| \forall i' \in \mathbb{Z}, O_{i'}^{(M),n-1} = O_{i'}^{(M),\infty,n-1} \text{ and } h_{i'}^{(M),n-1} = h_{i'}^{(M),\infty,n-1}\right)$$

$$\leq \mathbb{P}\left(O_0^{(M),1} \neq O_0^{(M),\infty,1} \middle| \forall i' \in \mathbb{Z}, O_{i'}^{(M),0} = O_{i'}^{(M),\infty,0} \text{ and } h_{i'}^{(M),0} = h_{i'}^{(M),\infty,0}\right)$$
$$\times \left(\sum_{n=1}^{N} \left(i_{max}^{M,0} - i_{min}^{M,0} + 2n + 1\right)\right)$$

$$\leq \mathbb{P}\left(O_0^{(M),1} \neq O_0^{(M),\infty,1} \middle| \forall i' \in \mathbb{Z}, O_{i'}^{(M),0} = O_{i'}^{(M),\infty,0} \text{ and } h_{i'}^{(M),0} = h_{i'}^{(M),\infty,0}\right)$$
$$\times \left(N \times \left(i_{max}^{M,0} - i_{min}^{M,0} + 1\right) + N(N+1)\right)$$

$$\leq \mathbb{P}\left(O_0^{(M),1} \neq O_0^{(M),\infty,1} \middle| \forall i' \in [\![-1,1]\!], O_{i'}^{(M),0} = O_{i'}^{(M),\infty,0} \text{ and } h_{i'}^{(M),0} = h_{i'}^{(M),\infty,0}\right)$$
$$\times \left(N \times \left(i_{max}^{M,0} - i_{min}^{M,0} + 1\right) + N(N+1)\right).$$

Here we used the invariance by translation in space and in time of the distribution of the process to pass from the 6^{th} to the 7^{th} line. □

This lemma implies that in order to prove Theorem 1.8, it is sufficient to show that

$$\mathbb{P}\left(O_0^{(M),1} \neq O_0^{(M),\infty,1} \middle| \forall i' \in [\![-1,1]\!], O_{i'}^{(M),0} = O_{i'}^{(M),\infty,0} \text{ and } h_{i'}^{(M),0} = h_{i'}^{(M),\infty,0}\right) \xrightarrow[M\to+\infty]{} 0.$$

In order to do so, we recall that three different reasons can lead to deviations from the BOA process.

1. If some type 1 plants are present but are never chosen as potential parents.
2. If type 1 seeds are present in the seed bank, but do not germinate during the generation we consider.
3. If type 1 seeds were produced less than $H+1$ generations ago, but already germinated.

In particular, for the event $\{O_0^{(M),1} \neq O_0^{(M),\infty,1}\}$ to occur given the initial condition, at least one of these events need to occur:

1. There exists a plant in patches $\{-1,0,1\}$ which does not belong to the set of potential parents of at least one seed bank compartment in patch 0. This event will be denoted as $\mathbf{R^{(M)}}$.

2. In at least one of the patches $\{-1,\ 0,\ 1\}$, one age class of potentially viable seeds was not represented among the seeds which germinated. In other words, there exists $i \in \{-1, 0, 1\}$ and $h \in [\![0, H]\!]$ such that none of the seeds which germinated in the patch i were of age h. There events will be denoted respectively as $\mathbf{S_{-1}^{(M)}}$ (for patch -1), $\mathbf{S_0^{(M)}}$ (for patch 0) and $\mathbf{S_1^{(M)}}$ (for patch 1).

Note that $\mathbb{P}(\mathbf{S_{-1}^{(M)}}) = \mathbb{P}(\mathbf{S_0^{(M)}}) = \mathbb{P}(\mathbf{S_1^{(M)}})$ by invariance by translation of the process. Therefore, in order to prove Theorem 1.8, it is sufficient to show that

$$\mathbb{P}(\mathbf{R^{M)}}) \xrightarrow[M\to+\infty]{} 0$$
$$\text{and} \qquad \mathbb{P}(\mathbf{S_0^{(M)}}) \xrightarrow[M\to+\infty]{} 0.$$

In order to do so, we shall bound from above both probabilities by quantities that vanish when both M and $k = \lceil M \rceil^\alpha$ grow to $+\infty$.

2.2.1 Upper bound on $\mathbb{P}(\mathbf{R^{(M)}})$

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

Lemma 2.4. *For all $M \geq 2$,*

$$\mathbb{P}(\mathbf{R^{(M)}}) \leq 3g^2M^2 \times \exp\left(M^\alpha \times \ln\left(1 - \frac{c^*}{g} \times \frac{1}{M}\right)\right).$$

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

$$\mathbb{P}(\mathbf{R^{(M)}}) \xrightarrow[M\to+\infty]{} 0.$$

Proof. Assume that $c^* = c$. Let $\widetilde{\mathbf{R}}^{\mathbf{(M)}}$ be the event: "The first seed which germinated in patch 1 was not chosen as a potential parent by the first seed bank compartment in patch 0 to be refilled." Then,

$$\mathbb{P}(\mathbf{R^{(M)}}) \leq 3\lfloor gM \rfloor \times \lfloor gM \rfloor \times \mathbb{P}(\widetilde{\mathbf{R}}^{\mathbf{(M)}}).$$

Indeed, $\mathbf{R^{(M)}}$ is the event "at least one plant in one of the patches $\{-1,\ 0,\ 1\}$ is not chosen as a potential parent in order to refill at least one seed bank compartment in patch 0." There exists $3\lfloor gM \rfloor^2$ pairs "plant not chosen in patch -1, 0 or 1 - seed bank compartment in patch 0", and as $c^* = c$, plants in patches -1 and 1 have less chances of being chosen as potential parents than plants in patch 0.

Then, each one of the $\lceil M^\alpha \rceil$ potential parents chosen to refill the first seed bank compartment in

patch 0 is *not* the first plant of patch 1 with probability

$$1 - c^* \times \frac{1}{\lfloor gM \rfloor} \leq 1 - \frac{c^*}{g} \times \frac{1}{M}.$$

Hence,

$$\begin{aligned} \mathbb{P}(\widetilde{\mathbf{R}}^{(\mathbf{M})}) &\leq \left(1 - \frac{c^*}{g} \times \frac{1}{M}\right)^{\lceil M^\alpha \rceil} \\ &\leq \left(1 - \frac{c^*}{g} \times \frac{1}{M}\right)^{M^\alpha} \\ \text{and} \qquad \mathbb{P}(\mathbf{R}^{(\mathbf{M})}) &\leq 3g^2M^2 \times \exp\left(M^\alpha \times \ln\left(1 - \frac{c^*}{g} \times \frac{1}{M}\right)\right). \end{aligned}$$

If $c^* \neq c$, then we can directly adapt this proof defining instead the event $\widetilde{\mathbf{R}}^{(\mathbf{M})}$ as the event "The first seed which germinated in the patch 0 (instead of the patch 1) was not chosen as a potential parent by the first seed bank compartment in patch 0 to be refilled." □

2.2.2 Upper bound on $\mathbb{P}(\mathbf{S_0^{(M)}})$

The goal of this section is to prove the following lemma.

Lemma 2.5. *For all $M \geq 2$, for all $\epsilon > 0$,*

$$\begin{aligned} &\mathbb{P}(\mathbf{S_0^{(M)}}) \\ &\leq (H+1) \times \left(1 - (1-g)^H \times g\left(1 - \frac{1}{M}\right) + \epsilon a_H\right)^{gM\left(1-g(1-g)^H\right)+M(1+a_H)\epsilon-1} \\ &\quad + (H+1) \times \left(1 - \prod_{h'=0}^{H-1} \left(1 - \left((1-g)^h \times g + \epsilon a_h\right)^{g\times\left((1-g)^h\times\lfloor gM\rfloor - \epsilon M a_h\right)-\epsilon M}\right.\right. \\ &\qquad\qquad \left.\left. - \left(1 - (1-g)^h \times \left(g - \frac{1}{M}\right) + \epsilon a_h\right)^{gM\left(1-g(1-g)^h\right)+M(1+a_h)\epsilon-1}\right)\right). \end{aligned}$$

Proof. Let $M \geq 2$. In order to show this lemma, we define new events. For all $h \in [\![0, H]\!]$, let $\mathbf{S_0^{(M),h}}$ be the event : "None of the $\lfloor gM \rfloor$ seeds germinating in patch 0 during the first generation are of age h." Then,

$$\mathbb{P}(\mathbf{S_0^{(M),0}}) \leq \mathbb{P}(\mathbf{S_0^{(M),1}}) \leq ... \leq \mathbb{P}(\mathbf{S_0^{(M),H}}). \tag{2.2}$$

Indeed, informally, the age h seeds need to have avoided germination during h generations, so the expected number of age h seeds in the seed bank decreases with h. The inequality then comes from the observation that the less age h seeds there are, the easier it is to avoid all of them while choosing the $\lfloor gM \rfloor$ seeds germinating during the current generation.

Remark 2.6. Notice that this inequality would *not* be true if we had worked conditional on $(E_v)_{v\in\mathbb{N}}$, i.e, we do *not* have

$$\mathbb{P}(\mathbf{S_0^{(M),0}}\,|(E_v)_{v\in\mathbb{N}}) \leq \mathbb{P}(\mathbf{S_0^{(M),1}}\,|(E_v)_{v\in\mathbb{N}}) \leq ... \leq \mathbb{P}(\mathbf{S_0^{(M),H}}\,|(E_v)_{v\in\mathbb{N}}).$$

Indeed, for instance, consider the case $E_1 = \emptyset$ and $E_2 \neq \emptyset$. Then the seed bank does not contain any age 1 seed, and

$$\mathbb{P}(\mathbf{S_0^{(M),1}}\,|(E_v)_{v\in\mathbb{N}}) = 1.$$

However, since the seed bank contains age 2 seeds,

$$\mathbb{P}(\mathbf{S_0^{(M),h}}\,|(E_v)_{v\in\mathbb{N}}) < 1 = \mathbb{P}(\mathbf{S_0^{(M),1}}\,|(E_v)_{v\in\mathbb{N}}).$$

Eq. (2.2) yields

$$\begin{aligned}\mathbb{P}(\mathbf{S_0^{(M)}}) &\leq \sum_{h=0}^{H} \mathbb{P}(\mathbf{S_0^{(M),h}}) \\ &\leq (H+1) \times \mathbb{P}(\mathbf{S_0^{(M),H}}).\end{aligned}$$

Under the notation of Lemma 2.1, we can rewrite the event $\mathbf{S_0^{(M),H}}$ as:

$$\mathbf{S_0^{(M),H}} = \{\text{None of the seeds in } E_H \text{ germinate during the next generation}\} \cup \{\text{Card}(E_H) = 0\}.$$

Therefore,

$$\begin{aligned}&\mathbb{P}(\mathbf{S_0^{(M)}}) \\ &\leq (H+1) \times \mathbb{P}\left(\text{Card}(E_H) \notin W_H\right) \\ &\quad + (H+1) \times \mathbb{P}\left(\text{None of the seeds in } E_H \text{ germinate during the next generation}\middle|\, \text{Card}(E_H) \in W_H\right) \\ &\leq (H+1) \times \mathbb{P}\left(\text{Card}(E_H) \notin W_H\right) \\ &\quad + (H+1) \times \mathbb{P}\left(\text{less than } g\text{Card}(E_H) - \epsilon M \text{ seeds from } E_H \text{ germinate} \,|\, \text{Card}(E_H) \in W_H\right).\end{aligned}$$

We then use Lemmas 2.1 and 2.2 to conclude:

$$\begin{aligned}&\mathbb{P}(\mathbf{S_0^{(M)}}) \\ &\leq (H+1) \times \left(1 - (1-g)^H \times g\left(1 - \frac{1}{M}\right) + \epsilon a_H\right)^{gM\left(1-g(1-g)^H\right)+M(1+a_H)\epsilon-1}\end{aligned}$$

$$+ (H+1) \times \left(1 - \prod_{h'=0}^{H-1} \left(1 - \left((1-g)^h \times g + \epsilon a_h\right)^{g \times \left((1-g)^h \times \lfloor gM \rfloor - \epsilon M a_h\right) - \epsilon M} - \left(1 - (1-g)^h \times \left(g - \frac{1}{M}\right) + \epsilon a_h\right)^{gM\left(1-g(1-g)^h\right) + M(1+a_h)\epsilon - 1}\right)\right).$$

□

We can now prove Theorem 1.8.

Proof. We have seen that in order to show the theorem, it is sufficient to show that

$$\mathbb{P}(\mathbf{R}^{(\mathbf{M})}) \xrightarrow[M \to +\infty]{} 0$$
$$\text{and} \qquad \mathbb{P}(\mathbf{S_0^{(M)}}) \xrightarrow[M \to +\infty]{} 0.$$

By Lemma 2.4, it is clear that

$$\mathbb{P}(\mathbf{R}^{(\mathbf{M})}) \xrightarrow[M \to +\infty]{} 0.$$

Then, let M_0 such that $\frac{1}{M_0} < g$, and let $\epsilon > 0$ such that

$$\forall h \in [\![0, H]\!], gM\left(1 - g(1-g)^h\right) + M(1+a_h)\epsilon - 1 < 1$$
$$\text{and} \qquad g\lfloor gM \rfloor \times (1-g)^h - M(1+a_h)\epsilon < 1.$$

Applying Lemma 2.5 to ϵ and to $M \geq M_0$, we also obtain that

$$\mathbb{P}(\mathbf{S_0^{(M)}}) \xrightarrow[M \to +\infty]{} 0,$$

and we can conclude. □

3 Extinction threshold for the k-parent WFSB metapopulation process

This section is devoted to the proof of Theorem 1.9, that is, to the proof of the existence of a critical extinction probability $p_c(H)$ depending only on the maximal dormancy duration H. In order to do so, we shall first formalize the coupling between the k-parent WFSB metapopulation process and a BOA process. Then, we shall explain how the issue of occupied patches in the BOA process can be seen as a percolation problem. We shall 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 (ξ, 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 $(\mathrm{Ext}_i^n)_{i\in\mathbb{Z}}$ the extinction events used to define (ξ^n, h^n) given (ξ^{n-1}, h^{n-1}). In other words, for all $n \in \mathbb{N}^*$ and $i \in \mathbb{Z}$, $\mathrm{Ext}_i^n = 1$ 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 $(\mathrm{Ext}_i^n)_{i\in\mathbb{Z},n\in\mathbb{N}^*}$: 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 $(\mathrm{Ext}_i^{n+1})_{i\in\mathbb{Z}}$. 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. In particular, a consequence of this coupling is the following corollary.

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

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

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

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

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 $(i, j) \in \mathbb{Z} \times [\![1, M]\!]$, if $\xi_{i,j}^n = 1$, then $h_{i,j}^n \geq h_i^{k,n}$. We deduce that

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

By Proposition 3.1, we obtain

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

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

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

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

We now show Proposition 3.1.

Proof. (Proposition 3.1) We show the result by induction. The result is clear 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. 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. We deduce

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

Therefore, by the induction hypothesis,

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

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

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

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,

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

Using the induction hypothesis yields

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

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

3.2 Percolation problem

In order to show Theorem 1.9, 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

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

such that $O_{i_0}^{\infty,0}\times\mathbb{1}_{\{h_{i_0}^{\infty,n}\leq H\}}=1$ and for all $l\in[\![1,L]\!]$,

$$i_l\in\{i_{l-1}-1,i_{l-1},i_{l-1}+1\}\qquad\text{and}\qquad n_l-n_{l-1}\in[\![1,H+1]\!].\tag{3.1}$$

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 (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 $i \in \mathbb{Z}\backslash\{0\}$, $O_i^{\{0\},0} = 0$ and $h_i^{\{0\},0} = 0$.

We can then define $S_n(p)$ as

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

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

Proposition 3.3. *There exists a unique $p_c(H) \in (0,1)$ such that*

$$\forall p \in [0, p_c(H)), \mathbb{P}\left(\forall n \in \mathbb{N}, S_n(p) \neq \emptyset\right) > 0$$
$$\forall p \in (p_c(H), 1], \mathbb{P}\left(\forall n \in \mathbb{N}, S_n(p) \neq \emptyset\right) = 0.$$

What remains to show is that $p_c(H)$ is indeed the extinction threshold we are looking for.

3.3 Proof of Theorem 1.9

In order to prove Theorem 1.9, we make three observations. First, for all $n \in \mathbb{N}$, the event $\{S_n \neq \emptyset\}$ is the same as the event

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

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 $i \in \mathbb{Z}\backslash\mathcal{L}$, $O_i^{\mathcal{L}} = 0$ 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 $\left(O^{\{0\},n}, h^{\{0\},n}\right))_{n\in\mathbb{N}}$ matches the one given above. We then have the following result.

Lemma 3.4. *For all $\mathcal{L} \in \mathcal{P}_f(\mathbb{Z})$, for all $n \in \mathbb{N}$,*

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

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 $p_c(H)$.

Proof. Let $n \in \mathbb{N}$ and $\mathcal{L} \in \mathcal{P}_f(\mathbb{Z})$. First, we observe that if we couple all the BOA processes considered by constructing them using the same extinction events,

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

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,

$$\begin{aligned}
&\mathbb{P}\left(\prod_{i\in\mathbb{Z}} \left(1 - O_i^{\{i_1\},n} \times \mathbb{1}_{\{h_i^{\{i_1\},n} \le H\}}\right) = 1 \middle| \prod_{i\in\mathbb{Z}} \left(1 - O_i^{\{i_0\},n} \times \mathbb{1}_{\{h_i^{\{i_0\},n} \le H\}}\right) = 1\right) \\
&\ge \mathbb{P}\left(\prod_{i\in\mathbb{Z}} \left(1 - O_i^{\{i_1\},n} \times \mathbb{1}_{\{h_i^{\{i_1\},n} \le H\}}\right) = 1\right),
\end{aligned}$$

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

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

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

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 $\mathcal{L} \in \mathcal{P}_f(\mathbb{Z})$ be the set defined as*

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

Then,

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

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 (ξ^0, h^0) are *not* of age 0, but older.

We can now prove Theorem 1.9.

Proof. (Theorem 1.9) Let $p_c(H)$ be given by Proposition 3.3. We assume that $p > p_c(H)$. Let also $n \in \mathbb{N}$, and let $\mathcal{L} \in \mathcal{P}_f(\mathbb{Z})$ be defined as in Lemma 3.5.

By Lemma 3.5,

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

Using Lemma 3.4, we obtain

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

Therefore,

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

by Proposition 3.3, and we can conclude. □

4 Appendix - Computation of $p_c(H)$

In this section, we briefly explain how to compute $p_c(H)$, and how to implement this approach and obtain an approximation for $p_c(H)$. 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 $p_c(H)$, 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 shall 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,..., $x_L \in \mathbb{Z}$ and n_0, n_1,..., $n_L \in \mathbb{N}$ such that:

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

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:

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

and let $\bar{r}_n(H,p) := \sup \bar{\xi}_{\mathbf{n}}(H,p)$. $\bar{\xi}_{\mathbf{n}}(H,p)$ 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 $H \in \mathbb{N}$,*

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

Therefore, in order to compute $p_c(H)$, 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 $p_c(H)$, we first define some approximations for $\bar{\xi}_{\mathbf{n}}(H,p)$ and $\bar{r}_n(H,p)$. Let $p \in [0,1]$. For all $x, y \in [\![-10500, 10500]\!]$ and $n^{(x)}, n^{(y)} \in [\![0, 10000]\!]$, we shall 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,... , $x_L \in [\![-10500, 10500]\!]$ and n_0,... ,$n_L \in [\![0, 10000]\!]$ such that:

1. $x_0 = y$, $n_0 = n^{(y)}$, $x_l = x$ and $n_l = n^{(x)}$.

2. $\forall l \in [\![1, L]\!]$, $x_l \in \{x_{l-1} - 1, x_{l-1}, x_{l-1} + 1\}$ and $1 \leq n_l - n_{l-1} \leq H + 1$.

3. $\forall l \in [\![1, L]\!]$, 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 $[\![-10500, 10500]\!]$, 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

$$\text{Approx}(\bar{\xi}_{\mathbf{n}}(H,p)) := \left\{ x \in \mathbb{Z} : \exists h_x, h_y \in [\![0, H]\!], \exists y \in \mathbb{Z} \backslash (\mathbb{N} \backslash \{0\}), (y, h_y) \xrightarrow[\text{Approx}(H,p)]{} (x, n + h_x) \right\},$$

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

In order to compute an approximate value for $p_c(H)$, 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 $p_c(H) = p$.

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. 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, Frank 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] J. Taylor and A. Véber. Coalescent processes in subdivided populations subject to recurrent mass extinctions. *Electronic Journal of Probability*, 14:242–288, 2009.

[38] J. Wakeley and N. Aliacar. Gene genealogies in a metapopulation. *Genetics*, 159(2):893–905, 2001.

[39] S. Wright. Evolution in Mendelian populations. *Genetics*, 16:97–159, 1931.

[40] D. Živković and A. Tellier. Germ banks affect the inference of past demographic events. *Molecular Ecology*, 21(22):5434–5446, 2012.