Adv. Appl. Probab. **56,** 235–286 (2024)
*doi:*10.1017/apr.2023.28

ADAPTATION OF A POPULATION TO A CHANGING ENVIRONMENT IN THE LIGHT OF QUASI-STATIONARITY

AURÉLIEN VELLERET * *Université Paris-Saclay, INRAE*

Abstract

We analyze the long-term stability of a stochastic model designed to illustrate the adaptation of a population to variation in its environment. A piecewise deterministic process modeling adaptation is coupled to a Feller logistic diffusion modeling population size. As the individual features in the population become further away from the optimal ones, the growth rate declines, making population extinction more likely. Assuming that the environment changes deterministically and steadily in a constant direction, we obtain the existence and uniqueness of the quasi-stationary distribution, the associated survival capacity, and the Q-process. Our approach also provides several exponential convergence results (in total variation for the measures). From this synthetic information, we can characterize the efficiency of internal adaptation (i.e. population turnover from mutant invasions). When the latter is lacking, there is still stability, but because of the high level of population extinction. Therefore, any characterization of internal adaptation should be based on specific features of this quasi-ergodic regime rather than the mere existence of the regime itself.

Keywords: Mobile optimum; quasi-stationary distribution; evolution; ecology; jump processes; Markov process in continuous time and continuous space

2020 Mathematics Subject Classification: Primary 92D15; 37A10; 60J25
Secondary 92D25; 37A30; 60G57

1. Introduction

1.1. Eco-evolutionary motivations

In line with [20], we are interested in the relative contribution of mutations with various strong effects to the adaptation of a population. Our first goal in the present paper is to analyze a stochastic model as simple as possible in which these mutations are filtered according to the advantage they provide, and to identify the key conditions of stability. In fact, this advantage may either be immediately significant (providing a better growth rate for the mutant subpopulation) or play a role in future adaptation (the population is doomed without many of these mutations). The stochastic model considered takes both aspects into account in order to provide a mathematical framework for relating these two contributions to the biological interpretation of adaptation. Before presenting its exact definition in Subsection 1.2, let us first explain the eco-evolutionary interpretations that it is intended to illuminate.

Received 9 September 2020; accepted 2 June 2023.
* Postal address: Université Paris-Saclay, INRAE, MaIAGE, F-78350 Jouy-en-Josas, France. Email address: aurelien.velleret@nsup.org

The process extends the one introduced by [20] and described more formally in [24] and [21]. Similarly, we assume that the population is described by a certain value $\hat{x} \in \mathbb{R}^d$, hereafter referred to as its trait. For the sake of obtaining a simple theoretical model, spatial dispersion and phenotypic heterogeneity (at least for the individual features of interest) are neglected. We therefore assume that the population is monomorphic at all times and that $\hat{x}$ represents the phenotype of all individuals in the population. Nonetheless, we allow for variations of this trait $\hat{x}$ due to stochastic events, namely when a subpopulation issuing from a mutant with trait $\hat{x} + w$ manages to persist and invade the 'resident' population. In the model, such events are assumed to occur instantaneously.

The main novelty of our approach is that we couple this 'adaptive' process with a Feller diffusion process N with a logistic drift. This diffusion describes the dynamics of the population size in a limit where it is large. Here we mean that individual birth and death events have negligible impact, but that the accumulation of these events has a visible and stochastic effect. In particular, the introduction of the 'size' in the model allows us to easily translate the notion of maladaptation, in the form of a poor growth rate.

For the long-time dynamics, we are mainly interested in considering only surviving populations, that is, conditioning the process upon the fact that the population size has not decreased to 0. The implication of taking size into account is twofold. On the one hand, extinction occurs much more rapidly when adaptation is poor. Indeed, the population size then declines very rapidly. So a natural selection effect can be observed at the population level. On the other hand, the better the adaptation, the larger the population size can be, and the more frequent is the birth of new mutants in the population. Also, in our simple model, a mutant trait that is better suited for the survival of the population as a whole is characterized by a greater probability that the resident population gets invaded, once a single mutant is introduced.

Compared to the case of a fixed size as in [24] and [21], this second implication leads to a stabilizing effect for the phenotype when the population size is large enough, but also a destabilizing effect when the population size decreases. This is in contrast to natural selection at the individual level (which is the main effect detailed in [20]). Indeed, when adaptation is already nearly optimal, among the mutants that appear in the population, very few can successfully maintain themselves and eventually invade the resident trait.

Let us assume here that mutations can allow individuals to survive in these new environments. In this context, how resilient is the population to environmental changes? Is there a clear threshold to the rate of change that such a population can handle? How can we describe the interplay between the above properties?

To begin to answer these questions, and like [20], we assume for simplicity that the environmental change is given by a constant-speed translation of the profile of fitness. This speed is denoted by v, and $\mathbf{e_1}$ provides the direction of the change. In practice, this means that the growth rate of the population at time t is expressed as a function of $x := \hat{x} - v\, t\, \mathbf{e_1}$, for a monomorphic population with trait $\hat{x}$ at time t. Naturally, the phenotypic lag x becomes the main quantity of interest for varying t.

Likewise, we can express as a function of x and w the probability that a mutant individual, with mutation w, will lead to the invasion of a resident population with trait $\hat{x}$ at time t. This probability should be stated solely in terms of x and $x + w$. Furthermore, we assume that the distribution of the additive effect for the new mutations is constant over time and independent of the trait $\hat{x}$ of the population before the mutation (and thus independent of x in the moving frame of reference).

In this context, we can exploit the notion of a quasi-stationary distribution (QSD) to characterize what would be an equilibrium for these dynamics prior to extinction (see Remark 2.2.3).

The main contribution of the current paper is to ensure that this notion is unambiguously defined for the process under consideration. To the best of our knowledge, this is the first time that the existence and uniqueness of the QSD has been proved for a piecewise deterministic process coupled to a diffusion.

By our proof, we also provide a justification of the notion of typical relaxation time and extinction time. The quasi-stationary description is well suited provided the latter is much longer than the former. As can be verified by simulations, typical convergence to the QSD is exponential in such cases. However, the marginal starting from specific initial conditions may take a long time before it approximates the QSD, mainly in cases where extinction is initially very likely.

In the following subsections of the introduction, we give the precise definition of the stochastic process. In Section 2, after specifying some elementary notation, we describe the main results, starting with our main hypotheses ([H], [D], and [A]) in Subsection 2.1 and giving the key Theorem 2.1 in Subsection 2.2. In Subsection 2.3, we discuss the interpretation of the theorem in terms of ecology and evolution. Its connection to related adaptation models is given in Subsection 2.4, and its connection to the classical techniques of quasi-stationarity in Subsection 2.5. The rest of the paper is devoted to proofs. In Section 3 we prove Proposition 2.1, namely the existence and uniqueness of the process. In Section 4 we introduce the main theorems on which our key result, Theorem 2.1, is based. Two alternative hypotheses ([D] and [A]) are considered, which entail some variations in the proofs. To facilitate comparison between these variations, we have chosen to group these six theorems in the three following sections. In the appendix, we include some pieces of proofs that are only slightly adjusted from similar arguments in [31]. We also provide the definition of a specific sigma-field and present a property related to jump events that we exploit in our proofs. We conclude with some illustrations of the asymptotic profiles obtained by simulating the stochastic process, which shed new light on the biological question.

1.2. The stochastic model

As explained in the introduction, we follow [20] for the definition of the adaptive component. The system that describes the combined evolution of the population size and its phenotypic lag is then given by

$$\begin{cases} X_t = x - v\,t\,\mathbf{e_1} + \displaystyle\int_{[0,t]\times\mathbb{R}^d\times(\mathbb{R}_+)^2} w\,\varphi_0\left(X_{s^-},\ N_s,\ w,\ u_f,\ u_g\right) M\big(ds, dw, du_f, du_g\big), \\ N_t = n + \displaystyle\int_0^t \Big(r(X_s)\cdot N_s - \gamma_0\cdot (N_s)^2\Big)\,ds + \sigma \int_0^t \sqrt{N_s}\,dB_s, \end{cases} \tag{S_0}$$

where N_t describes the size of the population and X_t the phenotypic lag of this population.

Here, $v>0$ is the speed of environmental change (in direction $\mathbf{e_1}$), (B_t) is a standard $(\mathcal{F}_t)$ Brownian motion, and M is a Poisson random measure on $\mathbb{R}_+\times\mathbb{R}^d\times\mathbb{R}_+$, also adapted to $(\mathcal{F}_t)$, with intensity

$$\pi\left(ds,\ dw,\ du_f,\ du_g\right) = ds\,\nu(dw)\,du_f\,du_g,$$

where $\nu(dw)$ is a measure describing the distribution of new mutations, and

$$\varphi_0(x,\ n,\ w,\ u_f,\ u_g) = \mathbf{1}_{\{u_f\le f_0(n)\}}\cdot \mathbf{1}_{\{u_g\le g(x,w)\}}.$$

Thanks to the following proposition, the independence between M and B is automatically deduced by choosing $(\mathcal{F}_t)$ such that their increments after time t are independent from $\mathcal{F}_t$. The filtration naturally generated from M and B is the most natural such choice.

Proposition 1.1. *Any Brownian motion and any Poisson random measure that are adapted to the same filtration $(\mathcal{F}_t)$ and such that their increments after time t are independent from $\mathcal{F}_t$ are necessarily independent.*

Proof of Proposition 1.1. Thanks to [16, Theorem 2.1.8], if X_1, X_2 are additive functionals and semimartingales with respect to a common filtration, both starting from zero, such that their quadratic covariation $[X_1, X_2]$ is almost surely (a.s.) zero, then the random vector $(X_1(t) - X_1(s), X_2(t) - X_2(s))$ is independent of $\mathcal{F}_s$, for every $0 \le s \le t$. Moreover, the vector (X_1, X_2) of additive processes is independent.

Note B the Brownian motion and M the Poisson random measure on $\mathbb{R}_+ \times \mathcal{X}$. For any test function $F: \mathcal{X} \mapsto \mathbb{R}$, define $Z(t) := \int_{[0,t]\times\mathcal{X}} F(x)\, M(ds, dx)$. Both Z and B are additive functionals and semimartingales with respect to the filtration $\mathcal{F}_t$, both starting from zero. Since Z is a jump process and B is continuous, their quadratic covariation a.s. equals 0. Since it applies to any F, exploiting [16, Theorem 2.1.8] implies that B and M are independent. □

In the model of the moving optimum originally considered in [20], $X = 0$ corresponds to the optimal state in terms of some reproductive value function $R(x)$, for $x \in \mathbb{R}$. This function R is also assumed to be symmetric and decreasing with $|X|$. Here we consider a possibly multidimensional state space for X and will usually not require any assumption on the related function g.

The quantity X is described as the phenotypic lag because $X_t + v\, t\, \mathbf{e_1}$ is the character of the individuals at time t in the population, while in the model of [20], the mobile optimum is located at trait $v\, t\, \mathbf{e_1}$. These assumptions on the fitness landscape are natural, and we abide by them in our simulations. Nonetheless, they are mainly assumed for simplicity, and we have chosen here to be as general as possible in the definition of r. Thus, X_t is a lag as compared to the trait $v\, t\, \mathbf{e_1}$, which is merely a reference value.

The function $g(X_t, w)$ is the mutation kernel, which describes the rate of fixation at which a mutant subpopulation of trait $X_t + v\, t\, \mathbf{e_1} + w$ invades a resident population of trait $X_t + v\, t\, \mathbf{e_1}$. Although the rate at which the mutations occur in one individual can reasonably be assumed to be symmetric in w, this is clearly not the case for g. In a large population, the filtering of considering only fixing mutations greatly restricts the occurrence of strongly deleterious mutations, while greatly favoring strongly advantageous mutations. For mutations with little effect, there is only a slight bias. To cover both of these situations, we consider in our analysis both the case where any mutation effect is permitted and the case where only advantageous ones are. Although the latter case will raise more difficulty in terms of accessibility of the domain, the core of the argument is essentially the same, and the simulations seem to provide similar results in both cases.

The term $f(N_t)$ is introduced to model the fact that given a constant mutation rate per individual, the larger the population size, the larger the mutation rate for the population. The first reasonable choice is $f(N_t) := N_t$, but we may also be interested in introducing an effect of the population size on the fixation rate.

The quantity N follows the equation for a Feller logistic diffusion where the growth rate r at time t depends only on X_t, while the strength of competition c and the coefficient of diffusion σ are kept constant. Such a process is the most classical one for the dynamics of a large population size in a continuous-space setting such that explosion is prevented. It is described in [22] (with fixed growth rate), notably as a limit of some individual-based models. The coefficient σ is related to the proximity between two uniformly sampled individuals in terms of their filiation links: $1/\sigma^2$ scales as the population size and is sometimes described as the 'effective population size'.

From a biological perspective, X has no reason to explode. Under our assumption [H11] below, such explosion is clearly prevented. However, we will not focus on conditions ensuring non-explosion for X. Indeed, this would mean (by Assumption [H8] below) that the growth rate becomes extremely negative. It appears very natural to suppose that this would lead to the extinction of the population. We therefore define the extinction time as

$$\tau_\partial := \inf\{t \geq 0; N_t = 0\} \wedge \sup_{\{k\geq 1\}} T_X^k, \qquad \text{where } T_X^k := \inf\{t \geq 0; \|X_t\| \geq k\}. \tag{1.1}$$

Because it simplifies many of our calculations, in the following we will consider $Y_t := \frac{2}{\sigma}\sqrt{N_t}$ rather than N_t. An elementary application of the Itô formula proves the following lemma.

Lemma 1.1. *With the previous notation,* $(X,\ Y)$ *satisfies the following stochastic differential equation:*

$$\begin{cases} X_t = x - v\,t\,\mathbf{e_1} + \displaystyle\int_{[0,t]\times\mathbb{R}^d\times(\mathbb{R}_+)^2} w\,\varphi\left(X_{s-},\ Y_s,\ w,\ u_f,\ u_g\right)\, M\big(ds, dw, du_f, du_g\big), \\ Y_t = y + \displaystyle\int_0^t \psi(X_s,\ Y_s)\, ds + B_t, \end{cases} \tag{S}$$

where we define, for any $(x,\ y) \in \mathbb{R}^d\times\mathbb{R}_+$,

$$\begin{cases} \psi(x,\ y) := -\dfrac{1}{2\,y} + \dfrac{r(x)\,y}{2} - \gamma\,y^3, \qquad \text{with } \gamma := \dfrac{\gamma_0\,\sigma^2}{8}, \\ \varphi(x,\ y,\ w,\ u_f,\ u_g) := \varphi_0\left(x,\ \sigma^2 y^2/4,\ w,\ u_f,\ u_g\right). \end{cases}$$

By considering $f(y) := f_0[\sigma^2 y^2/4]$, *note that we recover*

$$\varphi(x,\ y,\ w,\ u_f,\ u_g) = \mathbf{1}_{\{u_f\leq f(y)\}}\cdot\mathbf{1}_{\{u_g\leq g(x,w)\}}.$$

The aim of the following theorems is to describe the law of the marginal of the process (X, Y) at large time t, conditionally upon the fact that the extinction has not occurred—in short, for the marginal condition on non-extinction (MCNE) at time t. Considering the conditioning at the current time leads to properties of quasi-stationarity, while conditioning much farther in the future leads to a Markov process usually referred to as the Q-process, which in some sense is the process conditioned on never going extinct. The two aspects are clearly complementary, and our approach will treat both in the same framework, in the spirit initiated by [10].

1.3. Elementary notation

In the following, the notation $k \geq 1$ is to be understood as $k \in \mathbb{N}$, while $t \geq 0$ (resp. $c > 0$) should be understood as $t \in \mathbb{R}_+ := [0, \infty)$ (resp. $c \in \mathbb{R}_+^* := (0, \infty)$). In this context (with $m \leq n$), we denote the classical sets of integers by $\mathbb{Z}_+ := \{0, 1, 2...\}$, $\mathbb{N} := \{1, 2, 3...\}$, $[\![m, n]\!] :=$

https://doi.org/10.1017/apr.2023.28 Published online by Cambridge University Press

$\{m, \ m+1, ..., n-1, \ n\}$, where the symbol $:=$ makes explicit that we are defining notation via this equality. For maxima and minima, we usually write $s \vee t := \max\{s, t\}$, $s \wedge t := \min\{s, t\}$. Accordingly, for a function φ, $\varphi^{\wedge}$ (resp. $\varphi^{\vee}$) will be the notation for a lower (resp. upper) bound of φ. By $C^0(X, Y)$ we denote the set of continuous functions from any X to any Y. By $\mathcal{B}(X)$ we denote the set of bounded functions from any X to $\mathbb{R}$. By $\mathcal{M}(X)$ and $\mathcal{M}_1(X)$ we denote the sets of positive measures and probability measures, respectively, on any state space X. Numerical indices are generally indicated in superscript, while specifying notation is often in subscript. By definition, $\{y \in \mathcal{Y}; A(y) \ , \ B(y)\}$ denotes the set of values y of $\mathcal{Y}$ such that both $A(y)$ and $B(y)$ hold true. Likewise, for two probabilistic conditions A and B on $\omega \in \Omega$, and a random variable X, we may use $\mathrm{E}(X; A \ , \ B)$ instead of $\mathrm{E}(X \mathbf{1}_{\Gamma})$, where $\Gamma := \{\omega \in \Omega; A(\omega) \ , \ B(\omega)\}$.

2. Exponential convergence to the QSD

2.1. Hypothesis

We will consider two different sets of assumptions, including or rejecting the possibility for deleterious mutations to invade the population.

First, the assumptions $[H]$ below are generally in force throughout the paper, although sometimes some of them may not be involved (we will mention when this is the case):

$[H1]$ The function $f \in \mathcal{C}^0\left(\mathbb{R}_+^*, \ \mathbb{R}_+\right)$ is positive.

$[H2]$ The function $g \in \mathcal{C}^0\left(\mathbb{R}^d \times \mathbb{R}^d, \ \mathbb{R}_+\right)$ is bounded on any $K \times \mathbb{R}^d$, where K is a compact set of $\mathbb{R}^d$.

$[H3]$ The function r is locally Lipschitz continuous on $\mathbb{R}^d$, and $r(x)$ tends to $-\infty$ as $\|x\|$ tends to ∞.

$[H4]$ We have $\nu\left(\mathbb{R}^d\right) < \infty$. Moreover, there exist $\theta, \ \nu_{\wedge} > 0$ and $\eta \in (0, \theta)$ such that

$$\nu(dw) \geq \nu_{\wedge} \, \mathbf{1}_{B(\theta+\eta)\setminus B(\theta-\eta)} \, dw,$$

where $B(R)$, for $R > 0$, denotes the open ball of radius R centered at the origin.

$[H5]$ Provided $d \geq 2$, $\nu(dw) << dw$, and the density $g(x, w)\, \nu(w)$ (for a jump from x to $x + w$) of the jump size law with respect to the Lebesgue measure satisfies

$$\forall\, x_{\vee} > 0, \qquad \sup \left\{ \frac{g(x, w)\, \nu(w)}{\int_{\mathbb{R}^d} g(x, w')\, \nu(w')\, dw'}; \, \|x\| \leq x_{\vee}, \ w \in \mathbb{R}^d \right\} < \infty.$$

When we allow deleterious mutations to invade the population, we actually mean that the rate of invasion is always positive, leading to the following assumption:

$[D]$ The function g is positive.

Otherwise, we consider the case where deleterious mutations are forbidden, in the sense that the rate is zero for mutations that would induce an increase in $\|X\|$. The invasion rate of advantageous mutations, however, is still assumed to be positive. This is stated in Assumption $[A]$, below, as the alternative to Assumption $[D]$:

$[A]$ For any $x, w \in \mathbb{R}^d$, $\|x + w\| < \|x\|$ implies $g(x, w) > 0$, while $\|x + w\| \geq \|x\|$ implies $g(x, w) = 0$.

https://doi.org/10.1017/apr.2023.28 Published online by Cambridge University Press

Remarks.

- ⋆ For $d = 1$, no condition on the density of $g \cdot \nu$ as in $[H5]$ is required.
- ⋆ It is quite natural to assume that $f(0) = 0$ and that $f(y)$ tends to ∞ as y tends to ∞, but we will not need these assumptions.
- ⋆ 1 is the natural bound with the above-mentioned biological interpretation of $[H2]$. However, an extension can be introduced where g is not exactly the fixation probability; cf. Corollary 2.1.
- ⋆ Under $[H2]$ and $[H4]$ (since $\nu\big(\mathbb{R}^d\big) < \infty$), over any finite time-interval, only a finite number of mutations can occur. We also need lower bounds on the probability of specific events which roughly prescribe the dynamics of X. This is where the lower bound on the density of ν is exploited, as well as the positivity of g, deduced from either Assumption $[D]$ or Assumption $[A]$.
- ⋆ The strong assumption $[H3]$, that $r(x)$ tends to $-\infty$ as x tends to ∞, makes it easy to prove that the process is mostly kept confined, say within the time-interval $[0, t]$ under the conditioning that $\{t < \tau_\partial\}$. However, the proof could be directly adapted to specific situations where the lim sup of $r(x)$ is only upper-bounded by $-r_\wedge$ when $\|x\|$ tends to infinity. The requirement on the large-enough value of $r_\wedge$ could then be stated in terms of the process dynamics in a well-chosen compact subset of $(x, y) \in \mathbb{R}^d \times \mathbb{R}_+^*$.

2.2. Statement of the main theorems

First we need to ensure that the model specified by Equation (S) properly defines a unique solution. This is stated in the next proposition.

Proposition 2.1. *Suppose that the assumptions [H] hold. Then, for any initial condition* $(x, y) \in \mathbb{R}^d \times \mathbb{R}_+^*$, *there is a unique strong solution* $\big(X_t, Y_t\big)_{t\geq 0}$ *in the Skorokhod space satisfying (S) for any* $t < \tau_\partial$, *and* $X_t = Y_t = 0$ *for* $t \geq \tau_\partial$, *where the extinction time is expressed as* $\tau_\partial := \sup_{\{n\geq 1\}} T_Y^n \wedge \sup_{\{n\geq 1\}} T_X^n$, *where*

$$T_Y^n := \inf\{t \geq 0,\ Y_t \leq 1/n\} \quad \text{and} \quad T_X^n := \inf\{t \geq 0,\ \|X_t\| \geq n\}.$$

Remark. This proposition makes it possible to express τ_∂ as $\inf\{t \geq 0,\ Y_t = 0\}$.

We exploit the notion of uniform exponential quasi-stationary convergence as previously introduced in [32, Section 2.3].

Definition 1. For any linear, positive, and bounded semigroup $(P_t)_{t\geq 0}$ acting on a Polish state space $\mathcal{Z}$, we say that P displays a uniform exponential quasi-stationary convergence with characteristics $(\alpha, h, \lambda) \in \mathcal{M}_1(\mathcal{Z}) \times B(\mathcal{Z}) \times \mathbb{R}$ if $\langle \alpha \mid h\rangle = 1$ and there exist $C, \gamma > 0$ such that for any $t > 0$ and for any measure $\mu \in \mathcal{M}(\mathcal{Z})$ with $\|\mu\|_{TV} \leq 1$,

$$\big\| e^{\lambda t} \mu P_t(ds) - \langle \mu \mid h\rangle \alpha(ds) \big\|_{TV} \leq C e^{-\gamma t}. \tag{2.1}$$

Remarks.

- ⋆ As shown in [32, Fact 2.3.2], this implies that for any $t > 0$, $\alpha P_t(ds) = e^{-\lambda t}\alpha(ds)$. Any measure satisfying this property is called a *quasi-stationary distribution* (QSD).
It is elementary that $h_t : x \mapsto e^{\lambda t} \langle \delta_s P_t \mid \mathbf{1}\rangle$ converges to h, in the uniform norm, as t tends to infinity. We call h the *survival capacity*, because the value $e^{\lambda t} \langle \delta_x P_t \mid \mathbf{1}\rangle = \mathrm{P}_x(t <$

https://doi.org/10.1017/apr.2023.28 Published online by Cambridge University Press

$\tau_\partial)/\mathrm{P}_\alpha(t < \tau_\partial)$ enables us to compare the likelihood of survival with respect to the initial condition.
Since $h_{t+t'} = e^{\lambda t} P_t h_{t'}$, one can then easily deduce that $e^{\lambda t} P_t h = h$. It is also obvious that h is necessarily non-negative.

- ⋆ By using the term 'characteristics', we express that they are uniquely defined.

Our main theorem is stated as follows, with $\mathcal{Z} := \mathbb{R}^d \times \mathbb{R}_+^*$.

Theorem 2.1. *Suppose that the assumptions [H] hold. Suppose that either Assumption [D] or Assumption [A] holds. Then the semigroup P associated to the process* $Z := (X, Y)$ *and extinction at time* τ_∂ *displays a uniform exponential quasi-stationary convergence with some characteristics* $(\alpha, h, \lambda) \in \mathcal{M}_1(\mathcal{Z}) \times B(\mathcal{Z}) \times \mathbb{R}_+$. *Moreover, h is positive.*

Remark. We refer to [32, Corollary 2.3.4] for the implied result on the convergence of the renormalized semigroup to α. The fact that h is positive implies that there is no other QSD in $\mathcal{M}_1(\mathcal{Z})$.

In [3a, Section 2.3.2] there is also an analysis of the so-called Q-process, whose properties are as follows.

Theorem 2.2. *Under the same assumptions as in Theorem 2.1, with* (α, h, λ) *the characteristics of exponential convergence of P, the following properties hold:*

(i) ***Existence of the Q-process:***
There exists a family $\left(\mathbb{Q}_{(x,y)}\right)_{(x,y)\in\mathcal{Z}}$ *of probability measures on* Ω *defined by*

$$\lim_{t\to\infty} \mathrm{P}_{(x,y)}(\Lambda_s \mid t < \tau_\partial) = \mathbb{Q}_{(x,y)}(\Lambda_s), \tag{2.2}$$

for all $\mathcal{F}_s$*-measurable set* Λ_s. *The process* $\left(\Omega; (\mathcal{F}_t)_{t\geq 0}; (X_t, Y_t)_{t\geq 0}; \left(\mathbb{Q}_{(x,y)}\right)_{(x,y)\in\mathcal{Z}}\right)$ *is a* $\mathcal{Z}$*-valued homogeneous strong Markov process.*

(ii) ***Weighted exponential ergodicity of the Q-process:***
The measure $\beta(dx, dy) := h(x, y)\, \alpha(dx, dy)$ *is the unique invariant probability measure under* $\mathbb{Q}$. *Moreover, for any* $\mu \in \mathcal{M}_1(\mathcal{Z})$ *satisfying* $\langle \mu \mid 1/h\rangle < \infty$ *and* $t \geq 0$,

$$\left\| \mathbb{Q}_\mu\left[(X_t, Y_t) \in (dx, dy)\right] - \beta(dx, dy) \right\|_{TV} \leq C\, \|\mu - \langle \mu \mid 1/h\rangle\, \beta\|_{1/h}\, e^{-\gamma t}, \tag{2.3}$$

where

$$\mathbb{Q}_\mu(dw) := \int_{\mathcal{Z}} \mu(dx, dy)\, \mathbb{Q}_{(x,y)}(dw), \qquad \|\mu\|_{1/h} := \left\| \frac{\mu(dx, dy)}{h(x, y)} \right\|_{TV}.$$

Remarks.

- ⋆ For the total variation norm, it is equivalent to consider either (X, Y) or (X, N).
- ⋆ The constant $\langle \mu \mid 1/h\rangle$ in (2.3) is optimal up to a factor of 2, in the sense that for any $u > 0$, we have $\|\mu - u\,\alpha\|_{1/h} \geq \|\mu - \langle \mu \mid 1/h\rangle \beta\|_{1/h}/2$ (cf. [32, Fact 2.3.8]).
- ⋆ Since r tends to $-\infty$ as $\|x\|$ tends to ∞, it is natural to assume that mutations leading X to be large have a very small probability of fixation. Notably, it means that we strongly expect the upper bound of g in $[H2]$, uniform over w.

https://doi.org/10.1017/apr.2023.28 Published online by Cambridge University Press

- ⋆ Under Assumption [A], one may expect the real probability of fixation $g(x, w)$ to be at most of order $O(\|w\|)$ for small values of w (and locally in x). In such a case, we can allow ν to satisfy a smaller integrability condition than [$H4$] while forbidding observable accumulation of mutations.

Corollary 2.1. *Suppose that the assumptions [H] and [A] hold, except that $\nu(\mathbb{R}^d) = \infty$. Suppose instead that $\int_{\mathbb{R}^d} (\|w\| \wedge 1)\, \nu(dw) < \infty$, while $\widetilde{g}:(x, w) \mapsto g(x, w)/(\|w\| \wedge 1)$ is bounded on any $K \times \mathbb{R}^d$ for K a compact set of $\mathbb{R}^d$. Then the conclusions of Theorem 2.1 and Theorem 2.2 hold true.*

Proof of Corollary 2.1. (X, Y) is a solution of (S) if and only if it is a solution of

$$\begin{cases} X_t = x - v\, t\, \mathbf{e}_1 + \displaystyle\int_{[0,t]\times\mathbb{R}^d\times\mathbb{R}_+} w\, \widetilde{\varphi}\big(X_{s^-},\ Y_s,\ w,\ u_f,\ \widetilde{u}_g\big)\ \widetilde{M}\big(ds, dw, du_f, du_g\big), \\ Y_t = y + \displaystyle\int_0^t \psi(X_s,\ Y_s)\, ds + B_t, \end{cases} \qquad (\tilde{S})$$

where $\widetilde{M}$ is a Poisson random measure of intensity $ds\, \widetilde{\nu}(dw)\, du_f\, d\widetilde{u}_g$, with

$$\widetilde{\nu}(dw) := \nu(dw)/(\|w\| \wedge 1), \qquad \widetilde{\varphi}(x,\ y,\ w,\ u_f,\ \widetilde{u}_g) = \varphi(x,\ y,\ w,\ u_f,\ \widetilde{u}_g \cdot (\|w\| \wedge 1)),$$

and with $\widetilde{\varphi}$ defined as φ with g replaced by $\widetilde{g}$.

Thanks to the condition on ν, [$H4$] holds with $\widetilde{\nu}$ instead of ν. Thanks to the condition on g, [$H2$] still holds with $\widetilde{g}$ instead of g. Assumptions [A] and [$H5$] are equivalent for the systems (g, ν) and $(\widetilde{g}, \widetilde{\nu})$. Consequently, if we prove Theorem 2.1 and Theorem 2.2 with [$H2$] and [$H4$], the results follow under the assumptions of Corollary 2.1. □

2.3. Eco-evolutionary implications of these results

One of the major motivations for the present analysis is to make a distinction, as rigorously as possible, between an environmental change to which the population can spontaneously adapt and a change that imposes too much pressure. We recall that in [24], the authors obtain a clear and explicit threshold for the speed of this environmental change. Namely, above this threshold, the Markov process that they consider is transient, whereas below the threshold it is recurrent. Thus, it might seem a bit frustrating that such a distinction (depending on the speed value v) cannot be observed in the previous theorems. At least, these results prove that the distinction is not based on the existence or the uniqueness of the QSD, nor even on the exponential convergence per se.

In fact, the reason why this threshold is so distinct in [24] is that the model of [24] is based on the following underlying assumption: the poorer the current adaptation is, the more efficiently mutations are able to fix, provided that they are then beneficial. In our case, a population that is too poorly adapted is almost certainly doomed to rapid extinction, because the population size cannot be maintained at large values. Instead, long-term survival is triggered by dynamics that maintain the population as adapted. Looking back at the history of surviving populations, it is likely that the process was mostly kept confined outside of deadly areas.

In order to establish this distinction between environmental changes that are sustainable and those that endanger the population, we need a criterion that quantifies the stability of such core regions. Our results provide two exponential rates whose comparison is enlightening: if the extinction rate is of the same order as the convergence rate, or larger, this means that the

https://doi.org/10.1017/apr.2023.28 Published online by Cambridge University Press

dynamics is strongly dependent upon the initial condition. If the convergence is much faster, the dynamics will rapidly become similar, regardless of the initial condition. At least, this is the case for initial conditions that are not too risky (i.e. where h is not too small). This criterion takes into account the intrinsic sustainability of the mechanisms involved in the adaptation to the current environmental change, but does not involve the specific initial state of adaptation.

Looking at the simulation results, the convergence in total variation indeed appears to happen at some exponential rate, provided that extinction does not abruptly wipe out a large part of the distribution at a given time. However, it appears computationally expensive and not very meaningful to use the decay in total variation to obtain a generic estimate of the exponential rate at which the effect of the initial condition is lost. Although they are not as clearly justified, it seems more practical to exploit the decay in time of the correlations of X and/or N starting from the QSD profile. On the other hand, it does not seem very difficult to compare the extinction rate from this estimate. This is especially true in the case where $\mathcal{X}$ is of dimension one, as one can directly estimate the dynamics of the density and thus the extinction rate. Furthermore, it is quite reassuring to see that the choice between including and forbidding deleterious mutations (for which the invasion probability is expected be positive but very small) is not crucial in the present proof. We do not see much difference when looking at the simulations.

Much more can be said if we look at the simulation estimates of the QSD, the quasi-ergodic distribution (QED), and the survival capacity. We plan to detail these simulation results in a later article, but let us already give some insights into the comparison between the QSDs and the QEDs provided in Appendix B.

We see that although the QSDs look very different at the three different values of mutation rates, the QEDs are in fact very similar. When extinction plays a notable role, a tail appears on the QSD from the area of concentration of the QED to an area where the population size is close to zero. From the shape of the tail and the fact that it does not appear on the QED or for larger mutation rate, we infer that it corresponds in some sense to a path towards rapid extinction.

These regions are clearly more unstable than the core areas where the Q-process is kept confined. This is probably due to this decline in population size when the level of maladaptation becomes more pronounced. This confinement caused by conditioning upon survival only weakens in the recent past. It is noticeable that the QSD may give mass to conditions (x, n) most likely leading to extinction, provided the delay is sufficiently large before extinction actually occurs.

2.4. Quasi-ergodicity of related models

The current paper completes the illustrations given in Subsection 4.2 of [31] and Sections 4–5 of [32]. Supposing the model of the current paper was in fact the original motivation for the techniques presented in those two papers, we can focus more closely on each of the difficulties identified thanks to these illustrations. In each of them, the adaptation of the population to its environment is described by some process X which is a solution of some stochastic differential equation of the form

$$X_t = x - \int_0^t V_s\, ds + \int_0^t \Sigma_s \cdot dB_s + \int_{[0,t]\times\mathbb{R}^d\times\mathbb{R}_+} w\, \mathbf{1}_{\{u\le U_s(w)\}}\, M(ds, dw, du),$$

where B is an $\mathcal{F}_t$-adapted Brownian motion and M an $\mathcal{F}_t$-adapted Poisson random measure. A priori, V_s and Σ_s depend on X_s, while U_s depends on X_{s-} and possibly on a coupled process N_t describing the population size. Like the product $f(Y_s)\, g(X_{s-}, w)$ in Equation (S), one specifies

https://doi.org/10.1017/apr.2023.28 Published online by Cambridge University Press

in $U_s(w)$ the rate at which a mutation of effect w invades the population at time s. The quantity V_s relates both to the speed of the environmental change and to the mean effects of the mutations invading the population at time s in a limit of very frequent mutations of very small effects. The quantity Σ_s then relates to the undirected fluctuations both in the environment and in the effects of this large number of small fixing mutations.

We can relate the current coupling of X and N to an approximation given by the autonomous dynamics of a process Y similar to X. For the approximation to be as valid as possible, the law of Y should be biased by some extinction rate (depending at time t on the value Y_t), and its jump rate should be adjusted. By these means, we implicitly account for what would be the fluctuations of N if X were around the value of Y_t. This approximation is particularly reasonable when the characteristic fluctuations of N around its quasi-equilibrium are much quicker than the effect of the growth rate changing over time with the adaptation. Its validity is less clear when the extinction has a strong effect on the establishment of the quasi-equilibrium.

The exponential quasi-stationary convergence is treated in Subsection 4.2 of [31] for a coupling (X, N) that behaves as an elliptic diffusion, while Sections 4–5 of [32] deal with some cases of a biased autonomous process Y that behaves as a piecewise deterministic process. For such a process with jumps, it is manageable yet technical to deal with restrictions on the allowed directions or sizes of jumps, while requiring V_t to stay at zero actually makes the proof harder than choosing $V_t := v \cdot t$.

In the current article, we treat the following two technical difficulties. Firstly, we handle the combination of techniques specific to diffusion with those specific to piecewise deterministic processes. Secondly, we treat more general restrictions on the jump effects, possibly multidimensional, even in a case where a region of the state space is transitory.

While the proofs of $(A1)$ and $(A3)$ strongly depend on such local properties of the dynamics, those of $(A2)$ for these semigroups rely on a common intuition. Although we allow X to live in an unbounded domain, the maladaptation of the process when it is far from the optimal position constrains X to be kept confined conditionally upon survival. This effect of the maladaptation has been modeled either directly on the growth rate of the coupled process N_t or using some averaged description in terms of extinction rate. Such a confinement property for the coupled process is in fact the main novelty of [31] and is notably illustrated in [31, Subsection 4.2.4].

For simplicity, we dealt there with a locally elliptic process, for which the Harnack inequality is known to greatly simplify the proof, as observed previously for instance in [13]. The proof of this confinement is actually simpler with Y behaving as an autonomous process under the pressure of a death rate, provided this rate goes to infinity outside of compact sets (by adapting the proof of $(A2)$ from [31, Subsection 4.1.2]).

Assume for now that the fluctuations of N are much quicker than the change of the growth rate in the domain where the population is well adapted. Then we conjecture that considering the autonomous process Y (including the bias by the extinction rate) instead of the coupled process (X, N) would produce very similar results: the extinction rates and the rates of stabilization to equilibrium in the two models should be close, while the QSD profile of X should be similar to that of Y.

The drop in the quality of the approximation when extinction has a crucial contribution can have only a limited effect for our purpose, which is to compare the extinction rate to the rate of stabilization to equilibrium; see Subsection 2.3. Indeed, as long as the extinction rate is not considerably larger than the rate of stabilization to equilibrium, such domains of maladaptation are strongly avoided when looking at the past of surviving populations. On the other hand, the population is almost certainly doomed when it enters these domains, so that we should be able to neglect the contribution to the extinction rate of the dynamics of the process there.

2.5. The mathematical perspective on quasi-stationarity

The subject of quasi-stationarity is now quite vast, and a considerable literature is dedicated to it, as suggested by the bibliography collected by Pollett [28]. Some insights into the subject can be found in general surveys like [9] or [29], or more specifically for population dynamics in [23]. However, it appears that much remains to be done for the study of strong Markov processes both on a continuous space and in continuous time, without any property of reversibility. For general recent results, apart from [31] and [32], which we exploit, we refer the reader to [14], [2], [8], [18], or [19]. The difficulty is increased when the process is discontinuous (because of the jumps in X) and multidimensional, since the property of reversibility becomes all the more stringent and new difficulties arise (cf. e.g. [6, Appendix A]).

Thus, ensuring the existence and uniqueness of the QSD is already a breakthrough, and we are even able to ensure an exponential rate of convergence in total variation to the QSD, as well as similar results on the Q-process. This model is in fact a very interesting illustration of the new technique which we exploit. Notably, we see how conveniently our conditions are suited for exploiting the Girsanov transform as a way to disentangle couplings (here between X and N, which are respectively the evolutionary component and the demographic one).

Our approach relies on the general result presented in [32], which, as a continuation of [31], was originally motivated by this problem. In [31], the generalization of the Harris recurrence property at the core of the results of [10] is extended to deal with exponential convergence which is not uniform with respect to the initial condition. The fine control thus obtained over the MCNE opens the way for the approach developed in [32] to deal with continuous-time and continuous-space strong Markov processes with discontinuous trajectories.

After their seminal article [10], Champagnat and Villemonais obtained quite a number of extensions, for instance with multidimensional diffusions [7], processes that are inhomogeneous in time [11], and various examples of processes in a countable space, notably with the use of Lyapunov functions; cf. [13] or [14]. Exploiting the result of [14], it may be possible to ensure the properties of exponential quasi-ergodicity for a discontinuous process such as that of this article, keeping a certain dependence on the initial condition. At least, the conditions they provide as well as the ones from [2] are necessarily implied by our convergence result (cf. [12, Theorem 2.3] or [2, Theorem 1.1]). Yet, in the approach of [14] for continuous-time and continuous-space Markov process, the rather abstract assumption ($F3$) appears tightly bound to the Harnack inequality. The similar Assumption ($A4$) in [2] is also left without further guidance, while the assumption of a strong Feller property in [18] and [19] appears too restrictive. For discontinuous processes, these two properties generally do not hold true, which is what motivated us to look for an alternative statement in [32]. This technique is very efficient here.

This dependence on the initial condition is biologically expected, although its crucial importance becomes apparent when the population is already highly susceptible to extinction. For a broader comparison of this approach with the general literature, we refer to the introduction of [14] and the comparison with the literature provided in [31] and [32].

3. Proof of Proposition 2.1

Uniqueness. *Step 1: a priori upper bound on the jump rate.* Assume that we have a solution $(X_t, Y_t)_{t\leq T}$ to (S) until some (stopping) time T (i.e. for any $t < T$) satisfying $T \leq t_\vee \wedge T_Y^m \wedge T_X^n$

https://doi.org/10.1017/apr.2023.28 Published online by Cambridge University Press

for some $t_\vee > 0$, $m, n \geq 1$ (see Equation (1.1)). We know from [$H3$] that the growth rate of the population necessarily remains upper-bounded by some $r^\vee > 0$ until T. Thus, we deduce a stochastic upper bound $\left(Y_t^\vee\right)_{t\geq 0}$ on Y, namely

$$Y_t^\vee = y + \int_0^t \psi^\vee\left(Y_s^\vee\right) ds + B_t, \quad \text{where} \quad \psi^\vee(y) = -\frac{1}{2\,y} + \frac{r^\vee\,y}{2} - \gamma\,y^3, \tag{3.1}$$

which is thus independent of M. Since $\psi^\vee(y) \leq r^\vee\,y/2$, it is classical that $Y^\vee$—and a fortiori Y—cannot explode before T; see for instance [4, Lemma 3.3] or [22], where such a process is described in detail.

Under [$H2$], the jump rate of X is uniformly bounded until T by

$$\nu\left(\mathbb{R}^d\right) \cdot \sup\left\{g(x', w); x' \in \bar{B}(0, n), w \in \mathbb{R}^d\right\} \cdot \sup\left\{f(y'); y' \leq \sup_{s\leq t_\vee} Y_s^\vee\right\} < \infty \text{ a.s.}$$

Step 2: identification until T. In any case, this means that the behavior of X until T is determined by the value of M on a (random) domain associated to an a.s. finite intensity. Thus, we need a priori to consider only a finite number K of ‘potential’ jumps, which we can describe as the points $\left(T_J^i, W^i, U_f^i, U_g^i\right)_{i\leq K}$ in increasing order of times T_J^i.

From the a priori estimates, we know that for any $t < T_J^1 \wedge T$, $X_t = x - v\,t$. By the improper notation $t < T_J^1 \wedge T$, we mean $t < T_J^1$ if $K \geq 1$ (since $T_J^1 < T$ by construction) and $t < T$ if $K = 0$, i.e. when there is no potential jump before T. We then consider the solution $\hat{Y}$ of

$$\hat{Y}_t = y + \textstyle\int_0^t \psi\left(x - v\,s, \hat{Y}_s\right) ds + B_t.$$

It is not difficult to adjust the proof of [33] to this time-inhomogeneous setting, with [$H3$], so as to prove the existence and uniqueness of such a solution until any stopping time $T \leq \hat{\tau}_\partial$, where $\hat{\tau}_\partial := \inf\left\{t \geq 0, \hat{Y}_t = 0\right\}$. Furthermore, $\hat{Y}$ is independent of M and must coincide with Y until $T_J^1 \wedge T$. Since $T \leq T_Y^m$, the event $\left\{\hat{\tau}_\partial < T_J^1 \wedge T\right\}$ is necessarily empty. If there is no potential jump before T, i.e. $K = 0$, we have identified $\left(X_t, Y_t\right)$ for $t \leq T$ as $X_t = x - v\,t$, $Y_t = \hat{Y}_t$. Otherwise, at time T_J^1, we check whether $U_f^1 \leq f\left(\hat{Y}\left(T_J^1\right)\right)$ and $U_g^1 \leq g\left(x - v\,T_J^1, W^1\right)$. If this holds, then necessarily $X\left(T_J^1\right) = x - v\,T_J^1 + W^1$; otherwise $X\left(T_J^1\right) = x - v\,T_J^1$. Doing the same inductively for the following time-intervals $\left[T_J^k, T_J^{k+1}\right]$, we identify the solution (X, Y) until T.

Step 3: uniqueness of the global solution. Now consider two solutions (X, Y) and (X', Y') of (S), respectively defined up to τ_∂ and τ'_∂ as in Proposition 2.1, with, in addition, $X_t = Y_t = 0$ for $t \geq \tau_\partial$, and $X'_t = Y'_t = 0$ for $t \geq \tau'_\partial$.

On the event $\left\{\sup_m T_y^m = \tau_\partial \wedge \tau'_\partial\right\}$, we deduce by continuity of Y' that $T_y^m = T_y'^m$, so that $\tau_\partial = \tau'_\partial$. On the event $\left\{\sup_n T_X^n = \tau_\partial \leq \tau'_\partial < \infty\right\}$, for any n and $t_\vee > 0$ there exist $m \geq 1$ and $n' \geq n$ such that

$$T_X^n \wedge t_\vee < T_Y^m \wedge T_Y'^m \quad \text{and} \quad \left\|X\left(T_X^n \wedge t_\vee\right)\right\| \vee \left\|X'\left(T_X^n \wedge t_\vee\right)\right\| < n' < \infty.$$

Thanks to Step 2, (X, Y) and (X', Y') must coincide until $T = (t_\vee + 1) \wedge T_Y^m \wedge T_Y'^m \wedge T_X^{n'} \wedge T_X'^{n'}$, where the previous definitions ensure that $T_X^n \wedge t_\vee < T$ (with the fact that X and X' are right-continuous). This proves that $T_X^n \wedge t_\vee = T_X'^n \wedge t_\vee$, and with $t_\vee, n \to \infty$ that $\tau'_\partial = \tau_\partial$.

By symmetry between the two solutions, we have that a.s. $\tau_\partial = \tau'_\partial$, $X_t = X'_t$ for all $t < \tau_\partial$, and $X_t = X'_t = 0$ for all $t \geq \tau_\partial$. This concludes the proof of the uniqueness.

Existence. We see that the identification obtained for the uniqueness clearly defines the solution (X, Y) until some $T = T(t_\vee, n)$ such that either $T = t_\vee$ or $Y_T = 0$ or $\|X_T\| \geq n$. Thanks to the uniqueness property and the a priori estimates, this solution coincides with the ones for larger values of $t_\vee$ and n. Thus, it does indeed produce a solution up to time τ_∂. □

4. Main properties leading to the proof of Theorem 2.1

4.1. General criteria for the proof of exponential quasi-stationary convergence

The proof of Theorem 2.1 relies on the set of assumptions **(AF)** presented in [32], which we recall next. The assumptions **(AF)** are stated in the general context of a process Z that is right-continuous with left limits (càdlàg) on a Polish state space $\mathcal{Z}$, with extinction at a time still denoted by τ_∂. The notation is changed from that of [32] to prevent confusion with the current notation, Z corresponding now to the couple (X, Y).

We introduce the following notation for the exit and first entry times for any set $\mathcal{D}$:

$$T_{\mathcal{D}} := \inf\{t \geq 0; Z_t \notin \mathcal{D}\}, \qquad \tau_{\mathcal{D}} := \inf\{t \geq 0; Z_t \in \mathcal{D}\}. \tag{4.1}$$

The assumptions involved in **(AF)** are the following:

($A0$) There exists a sequence $(\mathcal{D}_\ell)_{\ell \geq 1}$ of closed subsets of $\mathcal{Z}$ such that for any $\ell \geq 1$, $\mathcal{D}_\ell \subset int(\mathcal{D}_{\ell+1})$ (with $int(\mathcal{D})$ denoting the interior of $\mathcal{D}$).

($A1$) There exists a probability measure $\zeta \in \mathcal{M}_1(\mathcal{Z})$ such that, for any $\ell \geq 1$, there exist $L > \ell$ and $c, t > 0$ such that

$$\forall z \in \mathcal{D}_\ell, \qquad \mathrm{P}_z\left[Z_t \in dx; t < \tau_\partial \wedge T_{\mathcal{D}_L}\right] \geq c\,\zeta(dz).$$

($A2$) We have $\sup_{\{z \in \mathcal{Z}\}} \mathrm{E}_z\left(\exp\left[\rho\,(\tau_\partial \wedge \tau_E)\right]\right) < \infty$.

($A3_F$) For any $\epsilon \in (0,\ 1)$, there exist $t_\wedge, c > 0$ such that for any $z \in E$ there exist two stopping times U_H and V with the property

$$\mathrm{P}_z\big(Z(U_H) \in dz'; U_H < \tau_\partial\big) \leq c\,\mathrm{P}_\zeta\left(Z(V) \in dz'; V < \tau_\partial\right), \tag{4.2}$$

as well as the following conditions on U_H: $\left\{\tau_\partial \wedge t_\wedge \leq U_H\right\} = \{U_H = \infty\}$, and

$$\mathrm{P}_z\big(U_H = \infty,\ t_\wedge < \tau_\partial\big) \leq \epsilon\ \exp\left(-\rho\, t_\wedge\right). \tag{4.3}$$

We further require that there exist a stopping time U_H^∞ extending U_H in the following sense:

- ⋆ We have $U_H^\infty := U_H$ on the event $\left\{\tau_\partial \wedge U_H < \tau_E^1\right\}$, where $\tau_E^1 := \inf\{s \geq t_\wedge : Z_s \in E\}$.
- ⋆ On the event $\left\{\tau_E^1 \leq \tau_\partial \wedge U_H\right\}$ and conditionally on $\mathcal{F}_{\tau_E^1}$, the law of $U_H^\infty - \tau_E^1$ coincides with that of $\widetilde{U}_H^\infty$ for a realization $\widetilde{Z}$ of the Markov process $(Z_t, t \geq 0)$ with initial condition $\widetilde{Z}_0 := Z\big(\tau_E^1\big)$ and independent of Z conditionally on $Z\big(\tau_E^1\big)$.

The quantity ρ as stated in Assumptions ($A2$) and ($A3_F$) is required to be strictly larger than the following *survival estimate*:

$$\rho_S := \sup\left\{\gamma \geq 0;\ \sup_{L \geq 1} \inf_{t > 0}\ e^{\gamma t}\,\mathrm{P}_\zeta\left(t < \tau_\partial \wedge T_{\mathcal{D}_L}\right) = 0\right\} \vee 0.$$

https://doi.org/10.1017/apr.2023.28 Published online by Cambridge University Press

We are now in a position to state **(AF)**:

- $(A1)$ holds for some $\zeta \in \mathcal{M}_1(\mathcal{Z})$ and a sequence $(\mathcal{D}_\ell)_\ell$ satisfying $(A0)$. Moreover, there exist $\rho > \rho_S$ and a closed set E such that $E \subset \mathcal{D}_\ell$ for some $\ell \geq 1$ and such that $(A2)$ and $(A3_F)$ hold.

As stated next by gathering the results of Theorems 2.2–2.3 and Corollary 2.2.7 of [32], **(AF)** implies the convergence results that we aim for, noting that the sequence $(\mathcal{D}_\ell)_\ell$ will cover the whole space. Some additional properties of approximations are also obtained, where the process is localized to large $\mathcal{D}_L$ by extinction.

Theorem 4.1. *Provided that* **(AF)** *holds, the semigroup P_t associated to the process Z with extinction at time τ_∂ displays a uniform exponential quasi-stationary convergence with some characteristics $(\alpha, h, \lambda) \in \mathcal{M}_1(\mathcal{Z})\times B(\mathcal{Z})\times\mathbb{R}$.*

Moreover, consider for any $L \geq 1$ the semigroup P^L for which the definition of τ_∂ is replaced by $\tau_\partial^L := \tau_\partial \wedge T_{\mathcal{D}_L}$. Then, for any $L \geq 1$ sufficiently large, P^L displays a uniform exponential quasi-stationary convergence with some characteristics $\left(\alpha^L, h^L, \lambda_L\right) \in \mathcal{M}_1(\mathcal{D}_L)\times B(\mathcal{D}_L)\times\mathbb{R}_+$. The associated versions of (2.1) hold true with constants that can be chosen uniformly in L. As L tends to infinity, λ_L converges to λ and α^L, h^L converge to α, h in total variation and pointwise, respectively.

If in addition $\cup_{\ell\geq 1}\mathcal{D}_\ell = \mathcal{Z}$, then h is positive and the results of Theorem 2.2 on the Q-process also hold true.

Remark. Under **(AF)**, the Q-process can generally be defined on the set $\mathcal{H} := \{z \in \mathcal{Z}; h(z) > 0\}$, and the fact that h is positive is not required or may be proved as a second step. The proof of Theorem 4.1, however, provides a lower bound of h on any $\mathcal{D}_\ell$, so that $\mathcal{Z} = \cup_{\ell\geq 1}\mathcal{D}_\ell$ is a practical assumption for the proof that h is positive.

Remark. The assumption $(A3_F)$ appears quite technical, and its usage is the main focus of [32]. It is referred to as the 'almost perfect harvest' property; it makes it possible to upper-bound the asymptotic survival probability from initial condition z as compared to the one from initial condition ζ. To this end, a coupling is introduced between the process with initial condition z and the one with initial condition ζ. A time shift is allowed in this coupling, which is initiated at the 'harvesting time' U_H for the first process and at the related stopping time V for the other process. Thanks to (4.2) and to the Markov property, the densities of the marginals can then be compared (up to a constant factor and this time shift), in a way that is sufficient for the required comparison of survival. We simply need an upper bound on the time shift of the form of the constant $t_{\overline{\wedge}}$. Since failures where $U_H = \infty$ while $t_{\overline{\wedge}} < \tau_\partial$ are allowed, this step is to be iterated, and the probability of such failures is to be controlled through (4.3), in relation to the available estimate for the decay in the survival probability.

For the proof of Theorem 2.1, the sequence $(\mathcal{D}_\ell)_{\ell\geq 1}$ is defined as follows:

$$\mathcal{D}_\ell := \bar{B}(0, \ell) \times [1/\ell, \ell], \tag{4.4}$$

where $\bar{B}(0, \ell)$ denotes the closed ball of radius ℓ for the Euclidean norm.

Forbidding deleterious mutations in the case of unidimensional $\mathcal{X}$ will make our proof a bit more complicated. This case is thus treated later on. The expression 'with deleterious mutations' will be used a bit abusively to discuss the model under Assumption $[D]$. On the other hand, the expression 'with advantageous mutations' will refer to the case where Assumption $[A]$ holds.

https://doi.org/10.1017/apr.2023.28 Published online by Cambridge University Press

These criteria are proved to hold true under the assumptions of Theorem 2.1 in Theorems 4.2–4.6 below. We see in Subsection 4.2.1 how these theorems together with Theorem 4.1 imply Theorem 2.1. In the next subsections, we then prove Theorems 4.2–4.6. By first mentioning the mixing estimate, we wish to highlight the constraint on the reachable domain under Assumption [A]. The order of the proofs is different and chosen for clarity of presentation. The mixing estimates are handled similarly under the different sets of assumptions and are directly exploited in the proofs of the harvest properties. The escape estimates are very close to those of previously considered models, so more easily dealt with.

4.2. The whole space is accessible: with deleterious mutations or $d \geq 2$

4.2.1. *Mixing property and accessibility.* With deleterious mutations, the whole space becomes accessible. In fact, this is also the case with only advantageous mutations, provided $d \geq 2$.

Theorem 4.2. *Suppose that the assumptions [H] hold. For $d = 1$, suppose Assumption [D] holds. For $d \geq 2$, suppose either Assumption [D] or Assumption [A]. Then, for any $\ell_I, \ell_M \geq 1$, there exist $L > \ell_I \vee \ell_M$ and $c, t > 0$ such that*

$$\forall (x_I,\ y_I) \in \mathcal{D}_{\ell_I}, \quad \mathrm{P}_{(x_I,\, y_I)}\left[(X,\ Y)_t \in (dx,\ dy); t < \tau_\partial \wedge T_{\mathcal{D}_L}\right] \geq c\, \mathbf{1}_{\mathcal{D}_{\ell_M}}(x,\ y)\, dx\, dy.$$

Remarks.

- ⋆ Equation (4.1) is exploited when defining $T_{\mathcal{D}_L} := \inf\{t \geq 0; (X,\ Y)_t \notin \mathcal{D}_L\}$.
- ⋆ Theorem 4.2 implies in particular that the density with respect to the Lebesgue measure of any QSD is uniformly lower-bounded on any $\mathcal{D}_\ell$.
- ⋆ In the case where Assumption [D] holds, $L := \ell_I \vee \ell_M + \theta$ can be chosen. The choice of t cannot generally be made arbitrary, at least for $d = 1$, since the lower bound on the density of jump sizes is only valid for jumps of size close to θ. Under Assumption [A] with $d \geq 2$, the constraint that jumps must be advantageous makes the convenient choice of L less clear.

4.2.2. *Escape from the transitory domain.*

Theorem 4.3. *Suppose that the assumptions [H] hold. Then, for any $\rho > 0$, there exists $\ell_E \geq 1$ such that (A2) holds with $E := \mathcal{D}_{\ell_E}$.*

Remark. Heuristically, this means that the killing rate can be made arbitrarily large by adding a killing effect when hitting some compact $\mathcal{D}_\ell$ that sufficiently covers $\mathcal{Z} = \mathbb{R} \times \mathbb{R}_+^*$.

4.2.3. *Almost perfect harvest.* We need some reference set on which our reference measure has positive density. With the constants θ and η involved in [H4], let

$$\Delta := \bar{B}(-\theta\, \mathbf{e_1}, \eta) \times [1/2, 2]. \tag{4.5}$$

This choice (which is rather arbitrary) is made in such a way that the uniform distribution on Δ can be taken as the lower bound in the conclusions of Theorems 4.5 and 4.2.

Including deleterious mutations or with $d \geq 2$, we will exploit the following theorem for sets E of the form $E := \mathcal{D}_{\ell_E}$, where ℓ_E is determined thanks to Theorem 4.3. But the theorem holds generally for any closed subsets E of $\mathbb{R}^d \times \mathbb{R}_+^*$ for which there exists $\ell \geq 1$ such that $E \subset \mathcal{D}_\ell$, a property that, for brevity, we denote by $E \in \mathbf{D}$.

Theorem 4.4. *Suppose that the assumptions [H] hold. For $d = 1$, suppose Assumption [D]. For $d \geq 2$, suppose either Assumption [D] or Assumption [A]. Then, for any $\rho > 0$, $\epsilon \in (0,\ 1)$,*

https://doi.org/10.1017/apr.2023.28 Published online by Cambridge University Press

and $E \in \mathbf{D}$, there exist $t_{\barwedge}, c > 0$ which satisfy the following property for any $(x, y) \in E$ and $(x_\zeta, y_\zeta) \in \Delta$. There exist a stopping time U_H such that

$$\left\{\tau_\partial \wedge t_{\barwedge} \le U_H\right\} = \{U_H = \infty\} \quad \text{and} \quad \mathrm{P}_{(x,y)}(U_H = \infty,\ t_{\barwedge} < \tau_\partial) \le \epsilon\ \exp(-\rho\, t_{\barwedge}),$$

and an additional stopping time V such that

$$\mathrm{P}_{(x,y)}\left[(X(U_H),\, Y(U_H)) \in (dx', dy');\, U_H < \tau_\partial\right] \le c\, \mathrm{P}_{\left(x_\zeta, y_\zeta\right)}\left[(X(V),\, Y(V)) \in (dx', dy');\, V < \tau_\partial\right]. \quad (4.6)$$

Moreover, there exists a stopping time U_H^∞ satisfying the following properties:

- *$U_H^\infty := U_H$ on the event $\left\{\tau_\partial \wedge U_H < \tau_E^1\right\}$, where $\tau_E^1 := \inf\left\{s \ge t_{\barwedge} : (X_s, Y_s) \in E\right\}$.*
- *On the event $\left\{\tau_E^1 < \tau_\partial\right\} \cap \{U_H = \infty\}$, and conditionally on $\mathcal{F}_{\tau_E^1}$, the law of $U_H^\infty - \tau_E^1$ coincides with that of $\widetilde{U}_H^\infty$ for the solution $(\widetilde{X}, \widetilde{Y})$ of*

$$\begin{cases} \widetilde{X}_r = X(\tau_E^1) - v\, r\, \mathbf{e_1} + \displaystyle\int_{[0,r]\times\mathbb{R}^d\times(\mathbb{R}_+)^2} w\, \varphi\left(\widetilde{X}_{s-},\ \widetilde{Y}_s,\ w,\ u_f,\ u_g\right)\, \widetilde{M}\big(ds, dw, du_f, du_g\big), \\ \widetilde{Y}_r = Y(\tau_E^1) + \displaystyle\int_0^r \psi\left(\widetilde{X}_s,\ \widetilde{Y}_s\right) ds + \widetilde{B}_r, \end{cases} \quad (4.7)$$

where $r \ge 0$, and $\widetilde{M}$ and $\widetilde{B}$ are independent copies of M and B, respectively.

4.2.4. *Proof of Theorem 2.1 as a consequence of Theorems 4.2–4.3.*

- First, it is clear that the sequence $(\mathcal{D}_\ell)_\ell$ satisfies both $(A0)$ and $\cup_{\ell\ge1}\mathcal{D}_\ell = \mathcal{Z}$.
- $(A1)$ holds true thanks to Theorem 4.2, where ζ is the uniform distribution over Δ—cf. (4.5).
- Theorem 4.3 implies $(A2)$ for any ρ, and we also require that ρ be chosen so that

  $$\rho > \rho_S := \sup\left\{\gamma \ge 0;\ \sup_{L\ge1}\inf_{t>0}\, e^{\gamma t}\, \mathrm{P}_\zeta\left(t < \tau_\partial \wedge T_{\mathcal{D}_L}\right) = 0\right\} \vee 0.$$

  Thanks to [30, Lemma 3.0.2] and $(A1)$, we know that ρ_S is upper-bounded by some value $\widetilde{\rho}_S$. In order to satisfy $\rho > \rho_S$, we set $\rho := 2\widetilde{\rho}_S$. Thanks to Theorem 4.3, we deduce $E = \mathcal{D}_{\ell_E}$ such that Assumption $(A2)$ holds for this value of ρ.
- Finally, Theorem 4.4 implies that Assumption $(A3_F)$ holds true, for E and ρ. In the adaptation of (4.6) where (x_ζ, y_ζ) is replaced by ζ, V is specified by the initial condition $(x_\zeta, y_\zeta) \in \Delta$ chosen uniformly according to ζ.

This concludes the proof of the assumptions **(AF)** with $\cup_{\ell\ge1}\mathcal{D}_\ell = \mathcal{Z}$. Exploiting Theorem 4.1, this implies Theorems 2.1 and 2.2 in the case where, besides the assumptions $[H]$, either Assumption $[D]$ holds or $d \ge 2$ and Assumption $[A]$ holds. □

4.3. No deleterious mutations in the unidimensional case

4.3.1. *Mixing property and accessibility.* When only advantageous mutations are allowed and $d = 1$, as soon as the size of jumps is bounded, the process can no longer access some portion

https://doi.org/10.1017/apr.2023.28 Published online by Cambridge University Press

of the space (there is a limit in the X direction). We could prove that the limit is related to the quantity $L_A := \sup\{M; \nu[2M, +\infty) > 0\} \in (\theta/2, \infty]$.

The accessible domains with maximal extension would then be of the form $[-\ell, L_A - 1/\ell] \times [1/\ell, \ell]$, for some $\ell \geq 1$. To simplify the proof, however, the limit L_A will not appear in the statements below. We simply want to point out this potential constraint on the visited domain. In fact, the X component is assumed to be negative in the following definition of the accessibility domains:

$$\Delta_E := \{[-L, 0] \times [1/\ell, \ell]; L, \ell \geq 1\}. \tag{4.8}$$

Theorem 4.5. *Suppose $d = 1$, and that the assumptions [H] and [A] hold. Then, for any $\ell_I \geq 1$ and $E \in \Delta_E$, there exists $L > \ell_I$ and $c, t > 0$ such that the following lower bound holds for any $(x_I, y_I) \in \mathcal{D}_{\ell_I}$:*

$$\mathrm{P}_{(x_I,\, y_I)}\left[(X_t,\ Y_t) \in (dx,\ dy); t < \tau_\partial \wedge T_{\mathcal{D}_L}\right] \geq c\, \mathbf{1}_E(x,\ y)\, dx\, dy. \tag{4.9}$$

Remark. Theorem 4.5 implies that the density with respect to the Lebesgue measure of any QSD is uniformly lower-bounded on any E of the form given by (4.8).

Proof. Let α be a QSD, and $E \in \Delta_E$. Since $\mathcal{X} = \cup_\ell \mathcal{D}_\ell$ and $\alpha(\mathcal{X}) = 1$, there exists ℓ_I such that $\alpha(\mathcal{D}_{\ell_I}) > 0$. Let λ be the extinction rate of α. Let L, c, t be such that (4.9) holds for this choice of E and ℓ_I. Then

$$\alpha(dx, dy) = e^{\lambda t} \alpha P_t(dx, dy) \geq \left(e^{\lambda t} \cdot \alpha\left(\mathcal{D}_{\ell_I}\right) \cdot c\right) \cdot \mathbf{1}_E(x, y)\, dx\, dy.$$

This concludes the proof of the above remark. □

4.3.2. *Escape from the transitory domain.*

Theorem 4.6. *Suppose that $d = 1$, and the assumptions [H] and [A] hold. Then, for any $\rho > 0$, there exists $E \in \Delta_E$ such that (A2) holds.*

Remark. Heuristically, this means that the asymptotic killing rate can be made arbitrarily large by adding killing when hitting some compact E that sufficiently covers $\mathbb{R}_- \times \mathbb{R}_+^*$.

4.3.3. *Almost perfect harvest.*

Theorem 4.7. *Suppose that the assumptions [H] and [A] hold. Then, for any $\rho > 0$, $\epsilon \in (0,\ 1)$, and $E \in \Delta_E$, there exist $t_{\barwedge}$, $c > 0$ which satisfy the same property as in Theorem 4.4.*

Remark. The definition of Δ_E is chosen to apply for Theorems 4.5, 4.6, and 4.7 all together.

4.3.4. *Proof of Theorem 2.1 as a consequence of Theorems 4.5–4.6.* The argument being very similar to the one for the case $d \geq 2$ or with Assumption [D], we go through it only briefly:

- ($A1$) holds thanks to Theorem 4.5, again with the choice of ζ uniform on Δ.
- Thanks to Theorem 4.6, and similarly as in the proof exploiting Theorem 4.3 in Subsection 4.2.4, we deduce that there exists $E \in \Delta_E$ such that ($A2$) holds with some value $\rho > \rho_S$.
- Finally, ($A3_F$) holds for these choices of ρ and E, thanks to Theorem 4.4.

This concludes the proof of the assumptions **(AF)** with $\cup_{\ell \geq 1} \mathcal{D}_\ell = \mathcal{Z}$. Exploiting Theorem 4.1, it implies Theorems 2.1 and 2.2 in the case where $d = 1$ and the assumptions [H] and [A] hold. □

https://doi.org/10.1017/apr.2023.28 Published online by Cambridge University Press

4.4. Structure of the proof

To allow for fruitful comparison, the proofs are gathered according to the properties they ensure—($A1$), ($A2$), and ($A3_F$), respectively. We first prove Theorems 4.3 and 4.4 in Subsections 5.1 and 5.2 respectively. Their proofs are directly adapted from the proof of [30, Proposition 4.2.2]. We then prove Theorems 4.2 and 4.5 in Section 6, and finally Theorems 4.4 and 4.7 in Section 7.

5. Escape from the transitory domain

The most straightforward way to prove exponential integrability of first hitting times is certainly via Lyapunov methods. However, we highly doubt that this can be achieved as easily as we present next, given the interplay between the different domains on which the escape is to be justified.

5.1. With deleterious mutations or $d \geq 2$

Theorem 4.3 is a direct consequence of the following proposition, which is given as [30, Proposition 4.2.2].

Proposition 5.1. *Assume that (X, N) is a càdlàg process on $\mathbb{R}^d \times \mathbb{R}_+$ such that N is a solution to*

$$dN_t = (r(X_t) - c\, N_t)\, N_t\, dt + \sigma \sqrt{N_t}\, dB_t,$$

where B is a Brownian motion. Assume that τ_∂ is upper-bounded by $\inf\{t \geq 0; N_t = 0\}$. *Provided that* $\limsup_{\|x\| \to \infty} r(x) = -\infty$, *it holds that for any $\rho > 0$, there exists $n > 0$ such that*

$$\sup_{\{x \in \mathcal{X}\}} \mathrm{E}_x \left(\exp\left[\rho\, (\tau_\partial \wedge \tau_{\mathcal{D}_n})\right]\right) < \infty.$$

The proof developed in the next subsection extends that of this result and is sufficient to illustrate the technique.

5.2. Without deleterious mutations, $d = 1$

In this section, we prove Theorem 4.6, i.e., the following statement:

- Suppose that $d = 1$, and that the assumptions [H] and [A] hold. Then, for any $\rho > 0$, there exists some set $E \in \Delta_E$ such that the exponential moment of $\tau_E \wedge \tau_\partial$ with parameter ρ is uniformly upper-bounded as follows:

$$\sup_{(x,y) \in \mathbb{R} \times \mathbb{R}_+} \mathrm{E}_{(x,\, y)} \left(\exp\left[\rho\, (\tau_E \wedge \tau_\partial)\right]\right) < \infty.$$

5.2.1. *Decomposition of the transitory domain.* The proof is very similar to that of [30, Subsection 4.2.4] except that, by Theorem 4.7, the domain E cannot be chosen as large. We thus need to consider another subdomain of $\mathcal{T}$, which will be treated specifically thanks to Assumption [A].

The complementary $\mathcal{T}$ of E is then made up of four subdomains: '$y \approx \infty$', '$y \approx 0$', '$x > 0$', and '$\|x\| \approx \infty$', as shown in Figure 1. Thus, we make the following definitions:

- $\mathcal{T}_\infty^Y := \{(-\infty,\ -L) \cup (0,\ \infty)\} \times (y_\infty,\ \infty) \bigcup\ [-\ell,\ 0] \times [\ell,\ \infty)$ ('$y \approx \infty$'),
- $\mathcal{T}_0 := (-L,\ L) \times [0,\ 1/\ell]$ ('$y \approx 0$'),

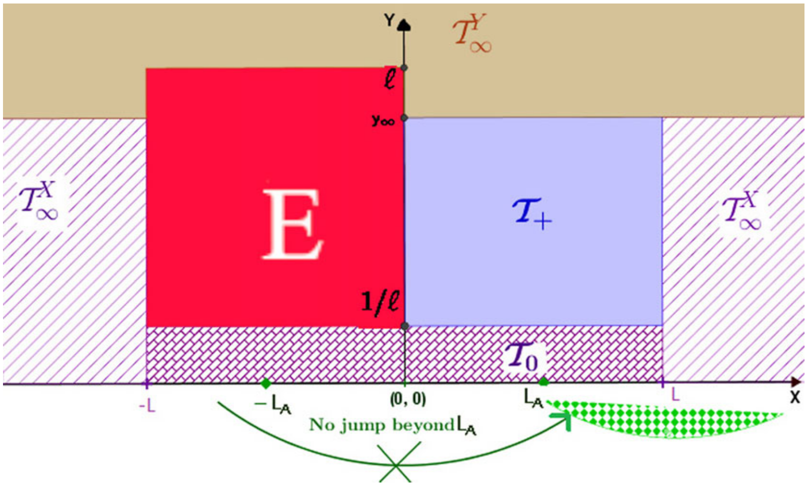


FIGURE 1. Subdomains for $(A2)$.
Until the process reaches E or extinction, it is likely to escape any region either from below or from the side into $\mathcal{T}_+$, the reverse transitions being unlikely. As long as $X_t > 0$, $\|X_t\|$ must decrease (see Fact 5.2.5 in Subsection 5.2.4). Once the process has escaped $\{x \geq L_A\}$, there is no way (via allowed jumps and ν) for it to reach it afterwards.

- $\mathcal{T}_+ := (0, L) \times (1/\ell, y_\infty]$ ('$x > 0$'),
- $\mathcal{T}_\infty^X := \{\mathbb{R} \setminus (-L,\ L)\} \times (1/\ell, y_\infty]$ ('$|x| \approx \infty$').

With some threshold $t_\vee$ (meant to ensure finiteness, but the effect will vanish as it tends to ∞), let us first introduce the exponential moments of each area (remember that τ_E is the hitting time of E):

- $\mathcal{E}_\infty^Y := \sup_{(x,\ y)\in\mathcal{T}_\infty^Y} \mathbb{E}_{(x,y)}[\exp(\rho\ V_E)]$,
- $\mathcal{E}_0 := \sup_{(x,\ y)\in\mathcal{T}_0} \mathbb{E}_{(x,y)}[\exp(\rho\ V_E)]$,
- $\mathcal{E}_\infty^X := \sup_{(x,\ y)\in\mathcal{T}_\infty^X} \mathbb{E}_{(x,y)}[\exp(\rho\ V_E)]$,
- $\mathcal{E}_X := \sup_{(x,\ y)\in\mathcal{T}_+} \mathbb{E}_{(x,y)}[\exp(\rho\ V_E)]$,

where $V_E := \tau_E \wedge \tau_\partial \wedge t_\vee$. Implicitly, $\mathcal{E}_\infty^Y$, $\mathcal{E}_\infty^X$, $\mathcal{E}_X$, and $\mathcal{E}_0$ are functions of $\rho,\ L,\ \ell,\ y_\infty$ that need to be specified.

5.2.2. *A set of inequalities.* The main ingredients for the following propositions are simple comparison properties that are specific to each of the transitory domain. By focusing on each of the domains separately (with the transitions between them), we can greatly simplify our control on the dependency of the processes.

As in [30, Subsection 4.2.4], we first state some inequalities between these quantities, summarized in Propositions 5.2, 5.3, 5.4, and 5.5 below. Using these inequalities, we prove in Subsection 5.2.3 that those quantities are bounded. This will complete the proof of Theorem 4.6.

https://doi.org/10.1017/apr.2023.28 Published online by Cambridge University Press

Proposition 5.2. *Suppose that the assumptions [H] hold. Then, given any $\rho > 0$, there exist $y_\infty > 0$ and $C^Y_\infty \geq 1$ such that for any $\ell > y_\infty$ and any $L > 0$,*

$$\mathcal{E}^Y_\infty \leq C^Y_\infty \cdot \left(1 + \mathcal{E}^X_\infty + \mathcal{E}_X\right). \tag{5.1}$$

Proposition 5.3. *Suppose that the assumptions [H] hold. Then, given any $\rho > 0$, there exists $C^X_\infty \geq 1$ which satisfies the following property for any ϵ^X, $y_\infty > 0$: there exist $L > 0$ and $\ell^X > y_\infty$ such that for any $\ell \geq \ell^X$,*

$$\mathcal{E}^X_\infty \leq C^X_\infty \cdot (1 + \mathcal{E}_0 + \mathcal{E}_X) + \epsilon^X \cdot \mathcal{E}^Y_\infty. \tag{5.2}$$

Proposition 5.4. *Suppose that the assumptions [H] and [A] hold. Then, given any ρ, $L > 0$, there exists $C_X \geq 1$ which satisfies the following property for any ϵ^+, $y_\infty > 0$: for any ℓ sufficiently large ($\ell \geq \ell^+ > y_\infty$),*

$$\mathcal{E}_X \leq C_X \cdot (1 + \mathcal{E}_0) + \epsilon^+ \cdot \mathcal{E}^Y_\infty. \tag{5.3}$$

Proposition 5.5. *Suppose that the assumptions [H] hold. Then, given any ρ, ϵ^0, $y_\infty > 0$, there exists $C_0 \geq 1$ which satisfies the following property for any L and for any ℓ sufficiently large ($\ell \geq \ell^0 > y_\infty$):*

$$\mathcal{E}_0 \leq C_0 + \epsilon^0 \cdot \left(\mathcal{E}^Y_\infty + \mathcal{E}^X_\infty + \mathcal{E}_X\right). \tag{5.4}$$

Propositions 5.2 and 5.3 are deduced from the estimates given in the following two lemmas, which are stated as Lemmas 4.2.6 and 4.2.7 in [31], on autonomous processes of the form

$$N^D_t := n + \int_0^t \left(r - c \cdot N^D_s\right) \cdot N^D_s \, ds + \int_0^t \sigma \sqrt{N^D_s} \, dB_s. \tag{5.5}$$

Proposition 5.2 *relies on the following property of descent from infinity, which is valid for any value of r.*

Lemma 5.1. *Let N^D be the solution of (5.5), for some $r \in \mathbb{R}$ and $c > 0$, with n the initial condition. Then, for any $t, \epsilon > 0$, there exists $n_\infty > 0$ such that*

$$\sup_{n>0} \mathrm{P}_n\left(t < \tau^D_\downarrow\right) \leq \epsilon \qquad \text{with} \tau^D_\downarrow := \inf\left\{s \geq 0,\ N^D_s \leq n_\infty\right\}.$$

Proposition 5.4 *relies on the strong negativity on the drift term, stated below.*

Lemma 5.2. *For any $c, t > 0$, with $\tau^D_\partial := \inf\left\{t \geq 0,\ N^D_t = 0\right\}$,*

$$\sup_{n>0} \mathrm{P}_n\left(t < \tau^D_\partial\right) \underset{r \to -\infty}{\longrightarrow} 0.$$

Moreover, for any $n, \epsilon > 0$, there exists n_c such that, for any r sufficiently low, with $T^D_\infty := \inf\left\{t \geq 0,\ N^D_t \geq n_c\right\}$, we have $\mathrm{P}_n\left(T^D_\infty \leq t\right) + \mathrm{P}_n\left(N^D_t \geq n\right) \leq \epsilon$.

On the other hand, Proposition 5.5 relies on an upper bound given as a continuous-state branching process, for which the extinction rate is much more explicit. The transition probability can clearly be made arbitrarily small by choosing a sufficiently small initial condition.

The only difference between the proofs in the current paper and those of [30, Appendices A–D] is that here we distinguish transitions into $\mathcal{T}_+$, which makes the term $\mathcal{E}_X$ appear with factors C^Y_∞, C^X_∞, and ϵ^0, respectively. These proofs are provided in Appendix A for the sake of completeness.

We prove next how to deduce Theorem 4.6 from the above set of four propositions. Then we will prove Proposition 5.4. This proof should convey both the main novelty and the common approach behind the proofs of these four propositions.

https://doi.org/10.1017/apr.2023.28 Published online by Cambridge University Press

5.2.3. *Proof that Propositions 5.2–5.4 imply Theorem 4.6.* We first prove that the inequalities (5.2), (5.3), and (5.4) given by Propositions 5.2–5.4 imply an upper bound on $\mathcal{E}_\infty^Y \vee \mathcal{E}_\infty^X \vee \mathcal{E}_X \vee \mathcal{E}_0$ for sufficiently small ϵ^X, ϵ^+ and ϵ^0.

Assuming first that $\epsilon^X \le \left(2\, C_\infty^Y\right)^{-1}$, we have

$$\mathcal{E}_\infty^X \le C_\infty^X\ (3 + 3\,\mathcal{E}_X + 2\,\mathcal{E}_0)\,, \qquad \mathcal{E}_\infty^Y \le C_\infty^Y\, C_\infty^X\ (4 + 4\,\mathcal{E}_X + 2\,\mathcal{E}_0)\,.$$

Assuming additionally that $\epsilon^+ \le \left(8\, C_\infty^Y\, C_\infty^X\right)^{-1}$, we have

$$\mathcal{E}_X \le C_X\ (2 + 3\,\mathcal{E}_0)\,, \qquad \mathcal{E}_\infty^X \le C_\infty^X\, C_X\ (9 + 11\,\mathcal{E}_0)\,, \qquad \mathcal{E}_\infty^Y \le C_\infty^Y\, C_\infty^X\ (12 + 14\,\mathcal{E}_0)\,.$$

Assuming also that $\epsilon^0 \le \left(60\, C_\infty^Y\, C_\infty^X\, C_X\right)^{-1}$ (and exploiting the fact that $2 \times [14 + 11 + 3] \le 60$), we have

$$\mathcal{E}_0 \le 50\, C_0, \qquad \mathcal{E}_X \le 152\, C_X\, C_0, \qquad \mathcal{E}_\infty^X \le 559\, C_\infty^X\, C_X\, C_0, \qquad \mathcal{E}_\infty^Y \le 712\, C_\infty^Y\, C_\infty^X\, C_0.$$

In particular,

$$\sup_{(x,y)\in\mathbb{R}\times\mathbb{R}_+} \mathrm{E}_{(x,\,y)}\left(\exp\left[\rho\,(\tau_E \wedge \tau_\partial)\right]\right) = \mathcal{E}_\infty^Y \vee \mathcal{E}_\infty^X \vee \mathcal{E}_X \vee \mathcal{E}_0 < \infty.$$

Let us now specify the choice of the various parameters involved. For any given ρ, we obtain from Proposition 5.2 the constant y_∞, and C_∞^Y, which gives us a value $\epsilon^X := \left(2\, C_\infty^Y\right)^{-1}$. We then deduce, thanks to Proposition 5.3, some value for C_∞^X, ℓ^X, and L. We can then fix $\epsilon^+ := \left(8\, C_\infty^Y\, C_\infty^X\right)^{-1}$, and deduce, according to Proposition 5.4, some value C_X and $\ell^+ > 0$. Now we fix $\epsilon^0 := \left(60\, C_\infty^Y\, C_\infty^X\, C_X\right)^{-1}$ and choose, according to Proposition 5.4, some value C_0 and $\ell^0 > 0$. To make the inequalities (5.2), (5.3), and (5.4) hold, we can just take $\ell := \ell^X \vee \ell^+ \vee \ell^0$. With the calculations above, we then conclude Theorem 4.6 with $E := [-L, 0] \times [1/\ell, \ell]$. □

5.2.4. *Proof of Proposition 5.4: phenotypic lag pushed towards the negatives.* Since the norm of X decreases at rate at least v as long as the process stays in $\widetilde{\mathcal{T}}_+ := [0, L] \times \mathbb{R}_+^*$, we know that the process cannot stay in this area during a time-interval larger than $t_\vee := L/v$. This effect will give us the bound $C_X := \exp(\rho\, L/v)$.

Moreover, we need to ensure that the transitions from $\mathcal{E}_X$ to $\mathcal{E}_\infty^Y$ are exceptional enough. This is done exactly as for [30, Proposition 4.2.2], by taking ℓ^+ sufficiently larger than y_∞. The event of having the process reach ℓ^+ in the time-interval $[0, t_\vee]$ is then exceptional enough.

More precisely, given L and $\ell > y_\infty \ge 1$ and initial condition $(x, y) \in \mathcal{T}_+$, let

$$C_X := \exp\left(\frac{\rho\, L}{v}\right), \qquad T := \inf\{t \ge 0;\, X_t \le 0\} \wedge V_E. \tag{5.6}$$

Lemma 5.3. *Suppose that the assumptions [H] and [A] hold. Then, for any initial condition $(x, y) \in \mathcal{T}_+$, we have $(X, Y)_T \notin \mathcal{T}_\infty^X$ a.s., and*

$$\forall\, t < T, \qquad X_t \le x - v\, t \le L - v\, t, \qquad \textit{so that } T \le t_\vee := L/v.$$

Thanks to Assumption [$H4$], an immediate induction on the number of jumps previous to $T \wedge t$ proves that the jumps of X can only make its value decrease (because it is positive, while the absolute value must necessarily decrease). This proves Lemma 5.3. Consequently,

$$\begin{aligned}
\mathrm{E}_{(x,y)}[\exp(\rho V_E)] &= \mathrm{E}_{(x,y)}\Big[\exp(T);\, T = V_E\Big] + \mathcal{E}_0\, \mathrm{E}_{(x,y)}\Big[\exp(T);\, (X, Y)_T \in \mathcal{T}_0\Big] \\
&\quad + \mathcal{E}_\infty^Y\, \mathrm{E}_{(x,y)}\Big[\exp(T);\, (X, Y)_T \in \mathcal{T}_\infty^Y\Big] \\
&\le C_X\ (1 + \mathcal{E}_0) + C_X\, \mathcal{E}_\infty^Y\, \mathrm{P}_{y_\infty}\left(T_\uparrow \le t_\vee\right),
\end{aligned}$$

https://doi.org/10.1017/apr.2023.28 Published online by Cambridge University Press

where $T_\uparrow := \inf\left\{t \geq 0;\, Y^\uparrow_t \geq \ell\right\}$, and $Y^\uparrow$ is the solution of

$$Y^\uparrow_t := y_\infty + \int_0^t \psi_\vee\left(Y^\uparrow_s\right)\, ds + B_t \qquad \left(\text{again } \psi_\vee(y) := -\frac{1}{2\,y} + \frac{r_\vee\, y}{2} - \gamma\, y^3\right).$$

We conclude the proof of Proposition 5.4 by noticing that $\mathrm{P}_{y_\infty}\left(T_\uparrow \leq t_\vee\right) \underset{\ell\to\infty}{\longrightarrow} 0.$ □

Given the proofs of Propositions 5.2, 5.3, and 5.5 provided in Appendix A and Subsection 5.2.3, the proof of Theorem 4.6 is now completed. The proof of Theorem 4.3 is sufficiently similar to be deduced without the need to refer to [30, Subsection 4.2.4].

6. Mixing properties and accessibility

In the following three subsections, before we turn to the proofs of Theorems 4.2 and 4.5, we describe the common elementary properties upon which they rely. The first one gives the trick for disentangling the behavior of the processes X and N up to a factor on the densities. Subsection 6.2 deals with the mixing property for the Y process. These results are exploited in Subsection 6.3 to obtain the elementary mixing properties that allow us to deduce $(A2)$. The next three subsections, starting from 6.4, deal respectively with the proof of Theorem 4.2 under Assumption $[D]$, then with the proof of Theorem 4.2 under Assumption $[A]$ and $d \geq 2$, and finally with the proof of Theorem 4.5.

General mixing properties

6.1. Construction of the change of probability under [$H4$]

The idea of this subsection is to prove that we can think of Y as a Brownian motion up to some stopping time which will bound U_H. If we get a lower bound for the probability of events in this simpler setup, this will prove that we also get a lower bound in the general setup.

The limits of our control. Let $t_G,\ x_\vee > 0,\ 0 < y_\wedge < y_\vee,\ N_J \geq 1$. Our aim is to simplify the law of $(Y_t)_{t\in[0,t_G]}$ as long as Y stays in $[y_\wedge, y_\vee]$, $\|X\|$ stays in $\bar{B}(0, x_\vee)$, and at most N_J jumps have occurred. Thus, let

$$\begin{aligned} T_X &:= \inf\left\{t \geq 0;\, \|X_t\| \geq x_\vee\right\}, & T_Y &:= \inf\left\{t \geq 0;\, Y_t \notin [y_\wedge, y_\vee]\right\}, \\ g_\vee &:= \sup\left\{g(x, w);\, \|x\| \leq x_\vee,\, w \in \mathbb{R}^d\right\}, & f_\vee &:= \sup\left\{f(y);\, y \in [y_\wedge,\ y_\vee]\right\}, \\ \mathcal{J} &:= \left\{(w, u_g, u_f) \in \mathbb{R}^d \times [0, f_\vee] \times [0, g_\vee]\right\}, \end{aligned} \tag{6.1}$$

so that $\nu \otimes du_g \otimes du_f(\mathcal{J}) = \nu\left(\mathbb{R}^d\right) g_\vee f_\vee < \infty$.

Our Girsanov transform alters the law of Y until the stopping time

$$T_G := t_G \wedge T_X \wedge T_Y \wedge U_{N_J}, \tag{6.2}$$

where

$$U_{N_J} := \inf\left\{\, t;\, M([0, t] \times \mathcal{J}) \geq N_J + 1\,\right\}. \tag{6.3}$$

Note that the $(N_J + 1)$th jump of X will then necessarily occur after T_G.

The change of probability. We define

$$L_t := -\int_0^{t\wedge T_G} \psi(X_s, Y_s) dB_s, \qquad \text{and } D_t := \exp\left[L_t - \langle L\rangle_t/2\right], \tag{6.4}$$

the exponential local martingale associated with (L_t).

https://doi.org/10.1017/apr.2023.28 Published online by Cambridge University Press

Theorem 6.1. *Suppose that the assumptions [H] hold. Then, for any t_G, $x_\vee > 0$ and $y_\vee > y_\wedge > 0$, there exists $C_G > c_G > 0$ such that, a.s. and for any $t > 0$, $c_G \le D_t \le C_G$. In particular, D_t is a uniformly integrable martingale and $\beta_t = B_t - \langle B, L\rangle_t$ is a Brownian motion under $\mathrm{P}^G_{(x,y)}$ defined as $\mathrm{P}^G_{(x,y)} := D_\infty \cdot \mathrm{P}_{(x,y)}$. We deduce the following bounds, valid for any $(x,\ y) \in \mathbb{R}^d \times \mathbb{R}_+$:*

$$c_G \cdot \mathrm{P}^G_{(x,\,y)} \le \mathrm{P}_{(x,\,y)} \le C_G \cdot \mathrm{P}^G_{(x,\,y)}.$$

On the event $\{t \le T_G\}$, $Y_t = y + \beta_t$; i.e. Y has the law of a Brownian motion under $\mathrm{P}^G_{(x,y)}$ up to time T_G. This means that we can obtain bounds on the probability of events involving Y as in our model by considering Y as a simple Brownian motion. Meanwhile, the independence between its variations as a Brownian motion and the Poisson process still holds by Proposition 1.1.

6.1.1. *Proof of Theorem* 6.1. The proof is achieved by ensuring uniform upper bounds of L_t and $\langle L\rangle_t$, which correspond to L_∞ and $\langle L\rangle_\infty$ for t_G replaced by $t \wedge t_G$.

Proof in the case where r is C^1. Let

$$\|r\|^G_\infty := \sup\left\{ |r(x)|; x \in \bar{B}(0, x_\vee)\right\}, \tag{6.5}$$

$$\|r'\|^G_\infty := \sup\left\{ |r'(x)|; x \in \bar{B}(0, x_\vee)\right\}. \tag{6.6}$$

With $\psi^\vee_G$ an upper bound of ψ on $\bar{B}(0,\ x_\vee) \times [y_\wedge,\ y_\vee]$ (deduced from $[H3]$), and recalling that (X, Y) belongs to this subset until T_G (see (6.2)), we have

$$\langle L\rangle_\infty = \int_0^{T_G} \psi(X_s, Y_s)^2 ds \le t_G \cdot \left(\psi^\vee_G\right)^2. \tag{6.7}$$

In the following, we look for bounds on $\int_0^{T_G} \psi(X_s, Y_s)dY_s$, noting that

$$L_{T_G} + \int_0^{T_G} \psi(X_s, Y_s)dY_s = \int_0^{T_G} \psi(X_s, Y_s)^2 ds \in \left[0, t_G \cdot \left(\psi^\vee_G\right)^2\right],$$
$$\int_0^{T_G} \psi(X_s, Y_s)dY_s = \int_0^{T_G} \left(-\frac{1}{2Y_s} + \frac{r(X_s)\, Y_s}{2} - \gamma \cdot (Y_s)^3\right) dY_s.$$

Now, thanks to Itô's formula,

$$\ln\left(Y_{T_G}\right) = \ln(y) + \int_0^{T_G} \frac{1}{Y_s} dY_s - \frac{1}{2}\int_0^{T_G} \frac{1}{(Y_s)^2} ds.$$

Thus,

$$\left|\int_0^{T_G} \frac{1}{Y_s} dY_s\right| \le 2\left(|\ln(y_\wedge)| \vee |\ln(y_\vee)|\right) + \frac{t_G}{2\,(y_\wedge)^2} < \infty. \tag{6.8}$$

Secondly,

$$\left(Y_{T_G}\right)^4 = y^4 + 4\int_0^{T_G} (Y_s)^3 dY_s + 6\int_0^{T_G} (Y_s)^2 ds.$$

Thus,

$$\left|\int_0^{T_G} (Y_s)^3 dY_s\right| \le (y_\vee)^4/4 + 3\, t_G\, (y_\vee)^2/2 < \infty. \tag{6.9}$$

https://doi.org/10.1017/apr.2023.28 Published online by Cambridge University Press

Thirdly,

$$r\big(X_{T_G-}\big)\cdot\big(Y_{T_G}\big)^2 = r(x)\,y^2 + 2\int_0^{T_G} r(X_s)\,Y_s\,dY_s + \int_0^{T_G} r(X_s)\,ds - v\int_0^{T_G} r'(X_s)\cdot(Y_s)^2\,ds + \int_{[0,T_G)\times\mathbb{R}^d\times\mathbb{R}_+}\big(\,r(X_{s-}+w) - r(X_{s-})\,\big)\cdot(Y_s)^2 \times \mathbf{1}_{\{u_f\le f(Y_s)\}}\,\mathbf{1}_{\{u_g\le g(X_{s-},\,w)\}}M\big(ds,\,dw,\,du_f,\,du_g\big). \tag{6.10}$$

Since $\forall s\le T_G,\ Y_s\in[y_\wedge,y_\vee]$, from [$H2$] and (6.1) we get

$$\forall\, s\le T_G,\ \forall\, w\in\mathbb{R}^d,\quad g(X_{s-},w)\le g_\vee,\quad f(Y_s)\le f_\vee,\quad \text{and } T_G\le U_{N_J}.$$

Since moreover $T_G\le T_X$,

$$\int_{[0,T_G)\times\mathbb{R}^d\times\mathbb{R}_+}\big(\,r(X_{s-}+w)-r(X_{s-})\,\big)\cdot(Y_s)^2 \times \mathbf{1}_{\{u_f\le f(Y_s)\}}\,\mathbf{1}_{\{u_g\le g(X_{s-},\,w)\}}M\big(ds,\,dw,\,du_f,\,du_g\big)\le 2\,N_J\,\|r\|_\infty^G\,(y_\vee)^2,$$

so that (6.10) leads to

$$2\left|\int_0^{T_G} r(X_s)\,Y_s\,dY_s\right| \le \big(2\,(N_J+1)\,\|r\|_\infty^G + \|r'\|_\infty^G\,v\,t_G\big)\cdot(y_\vee)^2 + \|r\|_\infty^G\,t_G<\infty. \tag{6.11}$$

The inequalities (6.8), (6.9), and (6.11) combined with (6.7) allow us to conclude that L_∞ and $\langle L\rangle_\infty$ are uniformly bounded. This proves the existence of $0<c_G<C_G$ such that a.s. $c_G\le D_\infty\le C_G$.

This statement is a priori adapted for t_G replaced by $t\wedge t_G$, yet these bounds are actually the largest for $t=t_G$. So this implies that $c_G\le D_t\le C_G$ holds uniformly in t. The rest of the proof is simply a classical application of Girsanov's transform theory.

Extension to the case where r is only Lipschitz continuous. The inequalities (6.8) and (6.9) are still true, so we show that we can find the same bound on $\left|\int_0^{T_G} r(X_s)\,Y_s\,dY_s\right|$ where we replace $\|r'\|_\infty^G$ by the Lipschitz constant $\|r\|_{Lip}^G$ of r on $\bar{B}(0,x_\vee)$, by approximating r by C^1 functions that are $\|r\|_{Lip}^G$-Lipschitz continuous.

Lemma 6.1. *Suppose r is Lipschitz continuous on* $\bar{B}(0,\ x_\vee)$ *for some* $x_\vee>0$. *Then there exists* $r_n\in C^1\big(\bar{B}(0,\ x_\vee),\mathbb{R}\big)$, $n\ge1$, *such that*

$$\|r_n-r\|_\infty^G\xrightarrow[n\to\infty]{}0\quad\textit{and}\quad\forall\,n\ge1,\ \|r_n'\|_\infty^G\le\|r\|_{Lip}^G.$$

Proof of Lemma 6.1. We begin by extending r on $\mathbb{R}^d$ with $r_G(x):=\ r\circ\Pi_G(x)$, where Π_G is the projection on $\bar{B}(0,x_\vee)$ (it is well known that r can be extended on $\bar{B}(0,x_\vee)$ with the same

https://doi.org/10.1017/apr.2023.28 Published online by Cambridge University Press

Lipschitz constant). Note that this extension r_G is still $\|r\|_{Lip}^G$-Lipschitz continuous. If we now define $r_n := r_G * \phi_n \in C^1$, where (ϕ_n) is an approximation of the identity of class C^1, then

$$\forall\, x, y,\ |r_n(x) - r_n(y)| = \left| \int_{\mathbb{R}^d} (r_G(x-z) - r_G(y-z))\phi_n(z)dz \right| \\ \leq \|r\|_{Lip}^G \, \|x-y\| \int_{\mathbb{R}^d} \phi_n(z)dz = \|r\|_{Lip}^G \, \|x-y\|.$$

It follows that

$$\forall\, n \geq 1, \quad \|r_n'\|_\infty^G \leq \|r\|_{Lip}^G, \quad \|r_n - r_G\|_\infty^G \underset{n\to\infty}{\longrightarrow} 0. \tag{6.12}$$

□

Proof that Lemma 6.1 combined with the case $r \in C^1$ proves Theorem 6.1. We just have to prove (6.11) with $\|r\|_{Lip}^G$ instead of $\|r'\|_\infty^G$. If we apply this formula for r_n and exploit Lemma 6.1, we see that there will be some $C = C(t_G, y_\vee, N_J) > 0$ such that

$$2\left| \int_0^{T_G} r_n(X_s)\, Y_s\, dY_s \right| \leq \left(2\,(N_J+1)\,\|r\|_\infty^G + \|r\|_{Lip}^G\, v\, t_G\right) (y_\vee)^2 + r_\infty\, t_G + C\,\|r - r_n\|_\infty^G.$$

Thus, it remains to bound

$$\left| \int_0^{T_G} (r_n(X_s) - r(X_s)) \cdot Y_s\, dY_s \right| \leq t_G\, y_\vee\, \psi_G^\vee\, \|r - r_n\|_\infty^G + |M_n|\,,$$

where $M_n := \int_0^{T_G} (r_n(X_s) - r(X_s))\, Y_s\, dB_s$ has mean 0 and variance

$$\mathrm{E}\big((M_n)^2\big) = \mathrm{E}\left(\int_0^{T_G} (r_n(X_s) - r(X_s))^2\, {Y_s}^2\, ds \right) \\ \leq t_G\,(y_\vee)^2 \left(\|r - r_n\|_\infty^G\right)^2 \underset{n\to\infty}{\longrightarrow} 0.$$

Thus, we can extract some subsequence $M_{\phi(n)}$ which converges a.s. towards 0, so that a.s.,

$$\left| \int_0^{T_G} r(X_s)\, Y_s\, dY_s \right| \leq \liminf_{n\to\infty} \left\{ \left| \int_0^{T_G} r_{\phi(n)}(X_s)\, Y_s\, dY_s \right| + t_G\, y_\vee\, \psi_G^\vee\, \|r - r_{\phi(n)}\|_\infty^G + \left|M_{\phi(n)}\right| \right\} \\ \leq \frac{1}{2}\left(2\,(N_J+1)\,\|r\|_\infty^G + \|r\|_{Lip}^G\, v\, t_G\right) \cdot (y_\vee)^2 + \frac{1}{2}\,\|r\|_\infty^G\, t_G < \infty.$$

The proof in the case $r \in C^1$ can then be exploited without difficulty. □

6.2. Mixing for *Y*

The proof will rely on Theorem and on the following classical property of Brownian motion.

Lemma 6.2. *Consider any constants $b_\vee > 0$, $\epsilon > 0$, and $0 < t_0 \leq t_1$. Then there exists $c_B > 0$ such that for any $b_I \in [0, b_\vee]$ and $t \in [t_0, t_1]$,*

$$\mathrm{P}_{b_I}\left(B_t \in db;\ \min_{s\leq t_1} B_s \geq -\epsilon\ ,\ \max_{s\leq t_1} B_s \leq b_\vee + \epsilon\right) \geq c_B \cdot \mathbf{1}_{[0,\, b_\vee]}(b)\, db,$$

where B under P_{b_I} has by definition the law of a Brownian motion with initial condition b_I.

https://doi.org/10.1017/apr.2023.28 Published online by Cambridge University Press

Thanks to this lemma and Theorem 6.1, we will be able to control Y to prove that it indeed diffuses and that it stays in some closed interval I_Y away from 0. We can then control the behavior of X independently of the trajectory of Y by appropriate conditioning of M—the Poisson random measure—so as to ensure the jumps we need (conditionally on its remaining in I_Y).

Proof. Consider the collection of marginal laws of B_t, with initial condition $b \in (-\epsilon, b_\vee + \epsilon)$, killed when it reaches $-\epsilon$ or $b_\vee + \epsilon$. It is classical that these laws have a density $u(t;b, b')$, $t > 0$, $b' \in [-\epsilon, b_\vee + \epsilon]$, with respect to the Lebesgue measure (cf. e.g. Bass [1, Section 2.4] for more details). We have that u is a solution to the Cauchy problem with Dirichlet boundary conditions

$$\begin{aligned} &\partial_t u(t;b_I, b) = \Delta_b u(t;b_I, b) && \text{for } t > 0,\ b_I, b \in (-\epsilon, b_\vee + \epsilon), \\ &u(t;b_I, -\epsilon) = u(t;b_I, b_\vee + \epsilon) = 0 && \text{for } t > 0. \end{aligned}$$

Thanks to the maximum principle (cf. e.g. Evans [17, Theorem 4, Subsection 2.3.3]), $u > 0$ on $\mathbb{R}_+^* \times [0, b_\vee] \times (-\epsilon, b_\vee + \epsilon)$, and since u is continuous in its three variables, it is lower-bounded by some c_B on the compact subset $[t_0, t_1] \times [0, b_\vee] \times [0, b_\vee]$. □

6.3. Mixing for X

For clarity, we decompose the 'migration' along X into different kinds of elementary steps, as already done in [32, Subsection 4.3.2]. Let

$$\mathcal{A} := \bar{B}(-\theta\, \mathbf{e_1},\ \eta/2), \qquad \tau_{\mathcal{A}} := \inf\{t \geq 0; X_t \in \mathcal{A}\,,\ Y_t \in [2, 3]\}\,, \tag{6.13}$$

where we assume without loss of generality that $\eta \leq \theta/8$ (the interval [2, 3] is chosen arbitrarily).

Under any of the three sets of assumptions considered in the following, the proof is achieved in three steps. The first step is to prove that, with a lower-bounded probability for any initial condition in $\mathcal{D}_\ell$, $\tau_{\mathcal{A}}$ is upper-bounded by some constant $t_{\mathcal{A}}$. In the second step, we prove that the process is sufficiently diffuse and that time shifts are not a problem. In the third step, we specify which sets we can reach from $\mathcal{A}$.

Recall that for any $\ell \geq 1$, $T_{\mathcal{D}_\ell} := \inf\{t \geq 0; (X,\ Y)_t \notin \mathcal{D}_\ell\} < \tau_\partial$. For $n \geq 3$, let us define $T_{(n)} := T_{\mathcal{D}_{2n}}$. For $n \geq 3$ and $t, c > 0$, let

$$\begin{aligned} \mathcal{R}^{(n)}(t, c) := \Big\{ &x_F \in \mathbb{R}^d; \forall\, (x_0, y_0) \in \mathcal{A} \times [1/n,\ n], \\ &\mathrm{P}_{(x_0,y_0)}\left[(X, Y)_t \in (dx, dy); t < T_{(n)}\right] \geq c\, \mathbf{1}_{B(x_F,\eta/2)}(x)\, \mathbf{1}_{[1/n,\, n]}(y)\, dx\, dy \Big\}\,. \end{aligned} \tag{6.14}$$

We will prove the mixing on a global scale by translating local mixing properties into certain induction properties of the sets $(\mathcal{R}^{(n)}(t, c))_{\{t,c>0\}}$.

Several local mixing properties require local lower and upper bounds on g, so that they can only be exploited in specific areas of $\mathbb{R}^d$. In order to provide a general framework for these through Proposition 6.1, let us consider the following increasing sequence of sets, indexed by $n \geq 1$:

$$\begin{aligned} \mathcal{G}_n := \{x \in \bar{B}(0, n);\ &\forall\, z \in [0, \eta/4],\ \forall\, \delta \in \bar{B}(0, \eta/2),\ \forall\, w \in \bar{B}(\theta\, \mathbf{e_1}, \eta), \\ &\qquad g(x - (\theta - z)\mathbf{e_1} + \delta, w) \geq 1/n, \\ \text{and} \quad &\forall\, z \in [-\theta, \eta/4],\ \forall\, \delta \in \bar{B}(0, \eta/2),\ \forall\, w \in \mathbb{R}^d, \\ &\qquad g(x + z\mathbf{e_1} + \delta, w) \leq n\}. \end{aligned}$$

These steps are deduced from the following elementary properties.

https://doi.org/10.1017/apr.2023.28 Published online by Cambridge University Press

Lemma 6.3. *Suppose that the assumptions [H] hold. Then for any $n \geq 1$ there exists $c_D > 0$ such that the following lower bound holds for any $(x_I, y_I) \in \mathcal{D}_n$ and*

$u \in [0, u_\vee(x_I)]$, *where* $u_\vee(x) := \sup\{u \geq 0; (x - v\,u\,\mathbf{e_1}) \in \bar{B}(0, n)\}$:

$$\mathrm{P}_{(x_I, y_I)}\left[(X_u, Y_u) \in (dx, dy); u < T_{(n)}\right] \geq c_D\, \delta_{\{x_I - v\,u\,\mathbf{e_1}\}}(dx) \cdot \mathbf{1}_{[1/n, n]}(y)\, dy.$$

In particular, for any $t, c > 0$, $n \geq 3$, the fact that x belongs to $\mathcal{R}^{(n)}(t, c)$ implies the following inclusion:

$$\forall\, u \in [0, u_\vee(x)], \quad x - v\,u\,\mathbf{e_1} \in \mathcal{R}^{(n)}(t + u, c \cdot c_D).$$

The proof of Lemma 6.3 being easily adapted from that of the next proposition, it is deferred until after the proof of the latter.

Proposition 6.1. *For any $n \geq 3$, there exist $t_P, c_P > 0$ such that for any $x_I \in \mathcal{G}_n$, for any $x_0 \in B(x_I, \eta/4)$ and $y_0 \in [1/n, n]$,*

$$\mathrm{P}_{(x_0, y_0)}\left[(X, Y)_{t_P} \in (dx, dy); t_P < T_{(n)}\right] \geq c_P\, \mathbf{1}_{B(x_I,\, 3\eta/4)}(x)\, \mathbf{1}_{[1/n,\, n]}(y)\, dx\, dy.$$

A direct application of the Markov property implies the following two results.

Corollary 6.1. *For any $n \geq 3$, there exist $t_P, c_P > 0$ such that for any $t, c > 0$, the following inclusion holds:*

$$\left\{x \in \mathbb{R}^d; d(x, \mathcal{R}^{(n)}(t, c) \cap \mathcal{G}_n) \leq \eta/4\right\} \subset \mathcal{R}^{(n)}(t + t_P, c \cdot c_P).$$

Lemma 6.4. *There exists $c_B > 0$ such that the following inclusion holds for any $t, t', c, c' > 0$ and any $n \geq 1$, provided that $-\theta\, \mathbf{e_1} \in \mathcal{R}^{(n)}(t, c)$:*

$$\mathcal{R}^{(n)}(t', c') \subset \mathcal{R}^{(n)}(t + t', c_B \cdot c \cdot c').$$

In the previous lemma, we may choose $c_B = Leb(B(x_I,\ \eta/2)) > 0$.

Corollary 6.3. as a consequence of Proposition 6.1. For $n \geq 3$, let $t_P, c_P > 0$ be prescribed by Proposition 6.1. We consider $x_I \in \mathcal{R}^{(n)}(t, c) \cap \mathcal{G}_n$, x_F such that $\|x_F - x_I\| \leq \eta/4$. Combining through the Markov property the fact that $x_I \in \mathcal{R}^{(n)}(t, c)$ and Proposition 6.1, we deduce that for any $(x_0, y_0) \in \mathcal{A} \times [1/n,\ n]$,

$$\begin{aligned}
&\mathrm{P}_{(x_0, y_0)}\left[(X, Y)_{t+t_P} \in (dx, dy); t + t_P < T_{(n)}\right] \\
&\quad \geq c \int_{B(x_I, \eta/2)} dx_0' \int_{1/n}^{n} dy_0'\, \mathrm{P}_{(x_0', y_0')}\left[(X, Y)_{t_P} \in (dx, dy); t_P < T_{(n)}\right] \\
&\quad \geq c \cdot Leb(B(x_I, \eta/4)) \cdot (n - 1/n) \cdot c_P \mathbf{1}_{B(x_I,\, 3\eta/4)}(x)\, \mathbf{1}_{[1/n,\, n]}(y)\, dx\, dy \\
&\quad \geq c \cdot c_P' \cdot \mathbf{1}_{B(x_F,\, \eta/2)}(x)\, \mathbf{1}_{[1/n,\, n]}(y)\, dx\, dy,
\end{aligned}$$

where $c_P' := Leb(B(0, \eta/4)) \cdot (n - 1/n) \cdot c_P > 0$. This means that $x_F \in \mathcal{R}^{(n)}\big(t + t_P, c \cdot c_P'\big)$. The proof of Corollary 6.1 is thus concluded with these choices of t_P and c_P', which are indeed independent from x_I, x_F. □

Proof of Proposition 6.1.

Step 1: description of the random event. For $n \geq 3$, we set $t_P := \theta/v$, $t_J := \eta/(4v)$, $y_\wedge := 1/(2\,n)$, $y_\vee := 2\,n$. Also, let

$$T^Y := \inf\{t \geq 0; Y_t \notin [y_\wedge,\ y_\vee]\}, \tag{6.15}$$

$$f_\wedge := \inf\{f(y); y \in [y_\wedge,\ y_\vee]\}, \qquad f_\vee := \sup\{f(y); y \in [y_\wedge,\ y_\vee]\}. \tag{6.16}$$

We have that $f_\vee$ is finite by [$H1$]. Thanks to [$H1$, we also know that $f_\wedge$ is positive.

On the event $\{t_P < T^Y\}$, we shall prove that the values of X on $[0, t_P]$ are prescribed as functions of M restricted to the subset

$$\mathcal{X}^M := [0, t_P] \times \mathbb{R}^d \times [0, f_\vee] \times [0, n]. \tag{6.17}$$

Let $x_0 := x_I + \delta_0$, with $x_I \in \mathcal{G}_n$ and $\delta_0 \in B(0, \eta/4)$, and $y_0 \in [1/n, \ n]$, which we consider as the initial conditions for the process (X, Y).

To ensure one jump of size around θ, at time nearly t_P, while 'deleting' the contribution of δ_0, let

$$\mathcal{J} := [t_P - t_J, t_P] \times B(\theta\, \mathbf{e_1} - \delta_0, \ \eta/2) \times [0, f_\wedge] \times [0, 1/n]. \tag{6.18}$$

We partition $\mathcal{X}^M = \mathcal{J} \cup \mathcal{N}$, where $\mathcal{N} := \mathcal{X}^M \setminus \mathcal{J}$. The main event under consideration is the following:

$$\mathcal{W} = \mathcal{W}^{(x_0,y_0)} := \left\{t_P < T^Y\right\} \cap \{M(\mathcal{J}) = 1\} \cap \{M(\mathcal{N}) = 0\}\,. \tag{6.19}$$

Thanks to Theorem 6.1 (with $x_\vee := n + 2\theta$, $t_G = t_P$, and the same values for $y_\wedge$ and $y_\vee$), there exists $c_G > 0$ such that

$$\mathrm{P}_{(x_0,y_0)}\left((X, Y)_{t_P} \in (dx, dy); \mathcal{W}\right) \geq c_G\, \mathrm{P}^G_{(x_0,y_0)}\left((X, Y)_{t_P} \in (dx, dy); \mathcal{W}\right). \tag{6.20}$$

Under the law $\mathrm{P}^G_{(x_0,y_0)}$, the condition $\{M(\mathcal{J}) = 1\}$ is independent of $\{M(\mathcal{N}) = 0\}$, of $\left\{t_P < T^Y\right\}$, and of Y_{t_P}; cf. Proposition 1.1. Thus, on the event $\mathcal{W}$, the only 'jump' coded in the restriction of M on $\mathcal{J}$ is given as $(T_J, \theta\, \mathbf{e_1} - \delta_0 + W, U_f, U_g)$, where T_J, U_f, and U_g are chosen uniformly and independently on $[t_P - t_J, t_P]$, $[0, f_\wedge]$, and $[0, 1/n]$, respectively, while $\theta\, \mathbf{e_1} - \delta_0 + W$ are chosen independently according to the restriction of ν on $B(\theta\, \mathbf{e_1} - \delta_0, \ 3\eta/4)$ (see notably [15, Chapter 2.4]). Thanks to [$H4$], W has a lower-bounded density d_W on $B(0, \ 3\eta/4)$.

The following lemma motivates this description.

Lemma 6.5. *Under $\mathrm{P}^G_{(x_0,y_0)}$, consider on the event $\mathcal{W}$ the random variable $W = W_J - \theta\mathbf{e_1} + \delta_0$, where $\left(T_J, W_J, U_f, U_g\right)$ is the only point encoded by M on $\mathcal{J}$. Then a.s. $X_{t_P} = x_I + W$ and $\mathcal{W}$ is included in $\left\{t_P < T_{(n)}\right\}$.*

Step 2: proof of Lemma 6.5.

Step 2.1. We prove that on the event $\mathcal{W}$ defined by (6.19),

$$\forall\, t < T_J, \quad X_t := x_0 - v\, t\, \mathbf{e_1}. \tag{6.21}$$

Indeed, $t_P < T^Y$ implies that for any $t \leq T_J$, $Y_t \in [y_\wedge, \ y_\vee]$. Thanks to (6.16), any 'potential jump' $\left(T'_J, W', U'_f, U'_g\right)$ such that $T'_J \leq T_J$ and either $U'_f > f_\vee$ or $U'_g > n$ will be rejected. Thanks to the definition of T_J, with (6.17), (6.18), and (6.19), no other jump can occur; thus (6.21) holds.

Note that, in order to prove this rejection very rigorously, we would like to consider the first one of such jumps. This cannot be done for (X, Y) directly, but it is easy to prove for any approximation of M where u_f and u_g are bounded. Since the result does not depend on these bounds and the approximations converge to (X, Y) (and are even equal to it, before T_J, for bounds larger than $(f_\vee, n)$), (6.21) indeed holds.

https://doi.org/10.1017/apr.2023.28 Published online by Cambridge University Press

Step 2.2. We then prove that the jump at time T_J is surely accepted.

Since $x_I \in \mathcal{G}_n$, by (6.15) and the definition of (T_J, W, U_f, U_g),

$$U_f \le f_\wedge \le f(Y_{T_J}), \qquad U_g \le 1/n \le g(x_0 - v\, T_J\, \mathbf{e_1}, \theta\, \mathbf{e_1} - \delta_0 + W) \\ = g(X_{T_J-}, \theta\, \mathbf{e_1} - \delta_0 + W).$$

Thus,

$$X_{T_J} = x_I + \delta_0 - v\, T_J\, \mathbf{e_1} + \theta\, \mathbf{e_1} - \delta_0 + W \\ = x_I + (\theta - v\, T_J)\, \mathbf{e_1} + W.$$

Step 2.3. We say that no jump can be accepted after T_J, which is proved as in Step 2.1. This means that the following equalities hold for any $t \in [T_J, t_P]$:

$$X_t = X_{T_J} - v \cdot (t - T_J)\, \mathbf{e_1} = x_I + W.$$

This concludes in particular the proof of Lemma 6.5 with $t = t_P = \theta/v$. □

Step 3. concluding the proof of Proposition 6.1.

Note that under P^G, $\{M(\mathcal{N}) = 0\}$ is also independent of $\left\{t_P < T^Y\right\}$ and of Y_{t_P}, so that

$$\mathrm{P}^G_{(x_0, y_0)}\left[(X, Y)_{t_P} \in (dx, dy); \mathcal{W}\right] \\ = \mathrm{P}(M(\mathcal{N}) = 0) \cdot \mathrm{P}(M(\mathcal{J}) = 1) \cdot \mathrm{P}^G_{y_0}\left(Y_{t_P} \in dy; t_P < T^Y\right) \\ \times d_W \cdot \mathbf{1}_{B(x_I,\, 3\eta/4)}(x)\, dx. \quad (6.22)$$

Thanks to (6.17) and (6.18),

$$\mathrm{P}(M(\mathcal{N}) = 0) \cdot \mathrm{P}(M(\mathcal{J}) = 1) \\ = (t_J \cdot f_\wedge/n) \cdot \nu\{B(\theta\, \mathbf{e_1} - \delta_0,\ 3\eta/4)\} \cdot \exp\left[-t_P \cdot f_\vee \cdot n \cdot \nu\left(\mathbb{R}^d\right)\right] \ge c_X, \quad (6.23)$$

where the lower bound c_X is chosen independently of x_0 and y_0 as follows:

$$c_X := (t_J \cdot f_\wedge \cdot d_W/n) \cdot Leb\{B(0, 3\eta/4)\} \cdot \exp\left[-t_P \cdot f_\vee \cdot n \cdot \nu\left(\mathbb{R}^d\right)\right] > 0.$$

Thanks to Lemma 6.2 (recall the definitions of $y_\wedge$ and $y_\vee$ at the beginning of this subsection),

$$\mathrm{P}^G_{y_0}\left(Y_{t_P} \in dy; t_P < T^Y\right) \ge c_B\, \mathbf{1}_{[1/n,\, n]}(y)\, dy. \quad (6.24)$$

Again, c_B is independent of x_0 and y_0.

Thanks to (6.20), (6.22), (6.23), (6.24) and Lemma 6.5, the following lower bound is valid for any $x_0 \in B(x_I, \eta/4)$ and any $y_0 \in [1/n,\ n]$ with the constant value $c_P := c_G\, c_X\, c_B\, d_W > 0$:

$$\mathrm{P}_{(x_0, y_0)}\left[(X, Y)_{t_P} \in (dx, dy); t_P < T_{(n)}\right] \ge c_P\, \mathbf{1}_{B(x_I,\, 3\eta/4)}(x)\, \mathbf{1}_{[1/n,\, n]}(y)\, dx\, dy.$$

This completes the proof of Proposition 6.1. □

Proof of Lemma 6.3. The proof of Lemma 6.3 relies on principles similar to those of Proposition 6.1. In this case, t_P is to be replaced by $u \in [0, u_\vee(x_I)]$, and the event under consideration is simply the following:

$$\mathcal{W}' := \left\{u < T^Y\right\} \cap \left\{M([0, u] \times \mathbb{R}^d \times [0, f_\vee] \times [0, n]) = 0\right\}.$$

The reasoning given for Step 2.1 can be applied to prove that for any $t \le u$, $X_t := x_0 - v\, t\, \mathbf{e_1}$. We also exploit Theorem 6.1 for the independence property between X and Y under $P^G_{(x_I, y_I)}$, and we use Lemma 6.2 to control the diffusion along the Y coordinate. Note that c_B can be taken independently of x_I, y_I, and t (noting that t is uniformly upper-bounded by $2n$). These arguments conclude the proof of the lower bound on the marginal of (X, Y) on the event $\{t < T_{(n)}\}$.

The implication in terms of the sets $\mathcal{R}^{(n)}(t, c)$ is obtained simply by exploiting the Markov property, similarly to the way in which Corollary 6.1 is deduced as a consequence of Proposition 6.1.

Application to the various sets of assumptions

6.4. Proof of Theorem 4.2 under Assumption [D]

We treat in this subsection the mixing of X when both advantageous and deleterious mutations are occurring. More precisely, each step corresponds to each of the following lemmas.

Lemma 6.6. *Suppose that the assumptions [H] and [D] hold. Then, for any $m \ge 3$, we can find $n \ge m$ and $c, t > 0$ such that $\bar{B}(0, m)$ is included in $\mathcal{R}^{(n)}(t, c)$.*

Lemma 6.7. *Suppose that the assumptions [H] and [D] hold. Then there exists $n \ge 3$ which satisfies the following property for any $t_1, t_2 > 0$: there exist $t_R > t_1$ and $c_R > 0$ such that for any $t \in [t_R, t_R + t_2]$ and $(x_0, y_0) \in \mathcal{A} \times [2, 3]$,*

$$P_{(x_0, y_0)}\left[(X, Y)_t \in (dx, dy); t < T_{(n)}\right] \ge c_R\, \mathbf{1}_{\mathcal{A}}(x)\, \mathbf{1}_{[2,3]}(y)\, dx\, dy.$$

Lemma 6.8. *Suppose that the assumptions [H] and [D] hold. Then, for any $\ell_I > 0$, there exist $c_I, t_I > 0$ and $n \ge \ell_I$ such that*

$$\forall\, (x, y) \in \mathcal{D}_{\ell_I}, \quad P_{(x,y)}\big(\tau_{\mathcal{A}} \le t_I \wedge T_{(n)}\big) \ge c_I.$$

In the following subsections, we prove these three lemmas, then explain how Theorem 4.2 is deduced as a consequence of them.

6.4.1. *Step 1: proof of Lemma* 6.6*.* Let $x_I = -\theta\, \mathbf{e_1}$. Since g is positive and continuous under Assumption [D], there exists n_0 such that $\bar{B}(x_I, \eta/2)$ is included in $\mathcal{G}_{n_0}$. With t_0, c_0 being the values associated to n_0 through Proposition 6.1, we deduce that $x_I \in \mathcal{R}^{(n_0)}(t_0, c_0)$.

For $m \ge 3$, let $K := \lfloor 4\, \|m + \theta\|/\eta \rfloor + 1$. Similarly, we can choose n_1 such that $B(0, m)$ is a subset of $\mathcal{G}_{n_1}$. Consider any $x_F \in \bar{B}(0, m)$, and for $0 \le k \le K$ let $x_k := -\theta\, \mathbf{e_1} + (k/K) \cdot (x_F + \theta\, \mathbf{e_1})$. This choice is made to ensure that $d(x_k, x_{k+1}) \le \eta/4$ and that for all $k \le K$, $x_k \in \mathcal{G}_{n_1}$. Thanks to Corollary 6.1, we deduce by immediate induction over $k \le K$ that there exist $n_2, t_k, c_k > 0$ independent of x_F such that $x_k \in \mathcal{R}^{(n_2)}(t_k, c_k)$. Furthermore, t_k and c_k are of the form $t_k := t_0 + k\, t_P$ and $c_k{:}c_0 \cdot (c_P)^k$. In particular, with $k = K$ and $n := n_2$, Lemma 6.6 is proved. □

6.4.2. *Step 2: proof of Lemma* 6.7*.* We keep $x_I := -\theta\, \mathbf{e_1}$ and $x_1 := (-\theta + \eta/2)\, \mathbf{e_1}$. Thanks to Lemma 6.6, there exist $n, t_1, c_1 > 0$ such that

$$\{\, x_I + u\, \mathbf{e_1}; u \in [\eta/6, 5\, \eta/6]\} \subset \mathcal{R}^{(n)}(t_1, c_1).$$

There exist $t_2, c_2 > 0$, thanks to Lemma 6.3, such that for all $t \in [t_2, t_2 + 2\, \eta/(3\, v)]$, we have $x_I \in \mathcal{R}^{(n)}(t, c_2)$. Applying Corollary 6.1 twice, with the knowledge that $B(x_I, \eta/2)$ is a subset

https://doi.org/10.1017/apr.2023.28 Published online by Cambridge University Press

of $\mathcal{G}_n$, we deduce that there exist $t_3, c_3 > 0$ such that

$$\forall\, t \in [t_3, t_3 + 2\,\eta/(3\,v)], \quad \mathcal{A} \subset \mathcal{R}^{(n)}(t, c_3).$$

Inductively applying Lemma 6.4, we deduce the following for any $k \geq 1$:

$$\forall\, t \in [k\,t_3, k\,t_3 + 2\,k\,\eta/(3\,v)], \quad \mathcal{A} \subset \mathcal{R}^{(n)}\big(t, c_3 \cdot [c_3 \cdot c_B]^{k-1}\big).$$

Let $t_1, t_2 > 0$ and consider $k \geq 1$ sufficiently large for $k\,t_3 > t_1$ and $2\,k\,\eta/(3\,v) > t_2$ to hold. Then Lemma 6.7 is proved with this value of n, $t_R := k\,t_3$, and $c_R := c_3 \cdot [c_3 \cdot c_B]^{k-1}$. □

6.4.3. *Step 3: proof of Lemma* 6.8. As before, we can find $n \geq \ell_I$ such that $\mathcal{D}_{\ell_I} \subset \mathcal{G}_n$. We go backwards in time from $\mathcal{A}$ by defining, for $t \geq 0$, $c > 0$,

$$\mathcal{R}'(t, c) := \left\{(x, y) \in \mathcal{G}_n; \mathrm{P}_{(x,y)}\left[\tau_{\mathcal{A}} \leq t \wedge T_{(n)}\right] \geq c\right\}.$$

It is clear that $\mathcal{A} \subset \mathcal{R}'(0, 1)$. Thanks to Proposition 6.1 and the Markov property, there exist $t_P, c_P > 0$ such that, for any $t, c > 0$,

$$\left\{x \in \mathcal{G}_n; d(x, \mathcal{R}'(t, c)) \leq \eta/4\right\} \subset \mathcal{R}'(t + t_P, c \cdot c_P).$$

Since $\mathcal{D}_{\ell_I} \subset \mathcal{G}_n$ is bounded, an immediate induction ensures that there exist $t_I, c_I > 0$ such that $\mathcal{D}_{\ell_I} \subset \mathcal{R}'(t_I, c_I)$. This concludes the proof of Lemma 6.8. □

6.4.4. *Theorem* 4.2 *as a consequence of Lemmas* 6.6–6.8. The proof is quite naturally adapted from that of Lemma 3.2.1 in [32]. Note that for any $n_1 \leq n_2$, $T_{(n_1)} \leq T_{(n_2)} \leq \tau_\partial$ holds a.s.

Let $\ell_I, \ell_M \geq 0$. According to Lemma 6.8, we can find $c_I, t_I > 0$ and $n_1 \geq \ell_I \wedge \ell_M$ such that for any $(x_I, y_I) \in \mathcal{D}_{\ell_I}$,

$$\mathrm{P}_{(x_I,y_I)}\big(\tau_{\mathcal{A}} \leq t_I \wedge T_{(n_1)}\big) \geq c_I. \tag{6.25}$$

Also, let $n_2 \geq n_1$, $c_R, t_R > 0$ be chosen, according to Lemma 6.7, to satisfy that for any $t \in [t_R, t_R + t_I]$ and $(x_0, y_0) \in \mathcal{A} \times [2, 3]$,

$$\mathrm{P}_{(x_0,y_0)}\left[(X, Y)_t \in (dx, dy); t < T_{(n_2)}\right] \geq c_R\, \mathbf{1}_{\mathcal{A}}(x)\, \mathbf{1}_{[2,3]}(y)\, dx\, dy. \tag{6.26}$$

Thanks to Lemma 6.6, since $\mathcal{D}_{\ell_M}$ is a bounded set, we know that there exist $n \geq n_2$, c_F, and $t_F > 0$ such that for any $(x_0, y_0) \in \mathcal{A} \times [2,\ 3]$,

$$\mathrm{P}_{(x_0,y_0)}\left[(X, Y)_{t_k} \in (dx, dy); t_k < T_{(n)}\right] \geq c_F\, \mathbf{1}_{\mathcal{D}_{\ell_M}}(x)\, \mathbf{1}_{[1/n,\, n]}(y)\, dx\, dy. \tag{6.27}$$

The fact that n is larger than n_1 and n_2 implies without difficulty that (6.25) and (6.26) hold with n_1 and n_2 replaced by n, which is how these statements are exploited in the following reasoning.

Let $t_M := t_I + t_R + t_F$ and $c_M := c_I \cdot c_R \cdot Leb(\mathcal{A}) \cdot c_F$. For any $(x_I, y_I) \in \mathcal{D}_{\ell_I}$, by combining (6.26), (6.27), and the Markov property, we deduce that a.s. on the event $\left\{\tau_{\mathcal{A}} \leq t_I \wedge T_{(n)}\right\}$,

$$\begin{aligned}
&\mathrm{P}_{(X,Y)[\tau_{\mathcal{A}}]}\left[\big(\widetilde{X}, \widetilde{Y}\big)[t_M - \tau_{\mathcal{A}}] \in (dx, dy); t_M - \tau_{\mathcal{A}} < \widetilde{T}_{(n)}\right] \\
&\quad \geq c_F \cdot \mathrm{P}_{(X,Y)[\tau_{\mathcal{A}}]}\left[\big(\widetilde{X}, \widetilde{Y}\big)[t_M - t_F - \tau_{\mathcal{A}}] \in \mathcal{A} \times [2, 3]; t_M - t_F - \tau_{\mathcal{A}} < \widetilde{T}_{(n)}\right] \\
&\qquad \times \mathbf{1}_{\mathcal{D}_{\ell_M}}(x)\, \mathbf{1}_{[1/n,\, n]}(y)\, dx\, dy \\
&\quad \geq c_R \cdot Leb(\mathcal{A}) \cdot c_F \cdot \mathbf{1}_{\mathcal{D}_{\ell_M}}(x)\, \mathbf{1}_{[1/n,\, n]}(y)\, dx\, dy,
\end{aligned}$$

where we have exploited the knowledge that $\tau_{\mathcal{A}} \le t_I$ to deduce that $t_M - t_F - \tau_{\mathcal{A}} \in [t_R, t_R + t_I]$. By combining this estimate with (6.25) and again the Markov property, we conclude that

$$\begin{aligned} \mathrm{P}_{(x_I,y_I)} &\left[\left(X_{t_M}, Y_{t_M}\right) \in (dx, dy); t_M < T_{(n)} \right] \\ &\ge \mathrm{P}_{(x_I,y_I)}(\tau_{\mathcal{A}} \le t_I \wedge T_{(n)}) \cdot c_R \cdot Leb(\mathcal{A}) \cdot c_F \cdot \mathbf{1}_{\mathcal{D}_{\ell_M}}(x)\, \mathbf{1}_{[1/n,\, n]}(y)\, dx\, dy \\ &\ge c_M\, \mathbf{1}_{\mathcal{D}_{\ell_M}}(x)\, \mathbf{1}_{[1/n,\, n]}(y)\, dx\, dy. \end{aligned}$$

This completes the proof of Theorem 4.2 with $L = 2n$, $c := c_M$, and $t := t_M$ under Assumption $[D]$. □

6.5. Proof of Theorem 4.2 under Assumption [A] and $d \ge 2$

The proof of Theorem 4.2 under Assumption [A] and $d \ge 2$ is handled in the same way as the proof in Subsection 6.4.4. Notably, the lemmas that replace Lemmas 6.7–6.8 have identical implications, as shown below.

Lemma 6.9. *Suppose that $d \ge 2$, and that the assumptions [H] and [A] hold. Then, for any $m \ge 3$, we can find $n \ge m$, $t, c > 0$ such that $\bar{B}(0, m)$ is included in $\mathcal{R}^{(L)}(t, c)$.*

Lemma 6.10. *Suppose that $d \ge 2$, and that the assumptions [H] and [A] hold. Then there exists $n \ge 3$ which satisfies the following property for any $t_1, t_2 > 0$: there exist $t_R > t_1$ and $c_R > 0$ such that, for any $t \in [t_R, t_R + t_2]$ and $(x_0, y_0) \in \mathcal{A} \times [2, 3]$,*

$$\mathrm{P}_{(x_0,y_0)}\left[(X, Y)_t \in (dx, dy); t < T_{(n)} \right] \ge c_R\, \mathbf{1}_{\mathcal{A}}(x)\, \mathbf{1}_{[2,3]}(y)\, dx\, dy.$$

Lemma 6.11. *Suppose that $d \ge 2$, and that the assumptions [H] and [A] hold. Then, for any $\ell_I > 0$, there exist $c_I, t_I > 0$ and $n \ge \ell_I$ such that*

$$\forall\, (x_0, y_0) \in \mathcal{D}_{\ell_I}, \quad \mathrm{P}_{(x_0,)}(\tau_{\mathcal{A}} \le t_{\mathcal{A}} \wedge T_{(n)}) \ge c_{\mathcal{A}}. \tag{6.28}$$

Since the implications are the same, the proof of Theorem 4.2 under Assumption [A] with $d \ge 2$ as a consequence of Lemmas 6.9–6.11 is mutatis mutandis the same as the proof given in Subsection 6.4.4. However, since deleterious mutations are now forbidden, the proof of Lemma 6.9 is much trickier than that of Lemma 6.6. The first step is given by the following two lemmas. To this end, given any direction $\mathbf{u}$ on the sphere S^d of radius 1, we denote its orthogonal component by

$$x^{(\perp \mathbf{u})} := x - \langle x, \mathbf{u} \rangle \mathbf{u}, \qquad \text{and specifically for } \mathbf{e_1}, \quad x^{(\perp 1)} := x - \langle x, \mathbf{e_1} \rangle \mathbf{e_1}. \tag{6.29}$$

Lemma 6.12. *Suppose that $d \ge 2$, and that the assumptions [H] and [A] hold. Then, for any $x_\vee > 0$, there exists $\epsilon \le \eta/8$ which satisfies the following property for any $n \ge 3 \vee (2\,\theta)$, $x \in B(0, n)$, and $\mathbf{u} \in S^d$ such that both $\langle x, \mathbf{u} \rangle \ge \theta$ and $\|x^{(\perp \mathbf{u})}\| \le x_\vee$: there exist $t_P, c_P > 0$ such that for any $t, c > 0$,*

$$x \in \mathcal{R}^{(n)}(t, c) \Rightarrow \bar{B}(x - \theta\, \mathbf{u}, \epsilon) \subset \mathcal{R}^{(n)}(t + t_P, c \cdot c_P).$$

Lemma 6.13. *Suppose that $d \ge 2$, and that the assumptions [H] and [A] hold. Then, for any $m \ge 3 \vee (2\,\theta)$, there exists $\epsilon \le \eta/8$ which satisfies the following property for any $x \in B(0, m)$ with $\langle x, \mathbf{e_1} \rangle \le 0$: there exist $t_P, c_P > 0$ such that*

$$\forall\, t, c > 0, \quad x \in \mathcal{R}^{(L)}(t, c) \Rightarrow \bar{B}(x, \epsilon) \subset \mathcal{R}^{(L)}(t + t_P, c \cdot c_P).$$

Lemma 6.13 is actually directly implied by Lemma 6.3 (first applied for a time-interval $[0, \theta/v]$), then Lemma 6.12 with $\mathbf{u} := \mathbf{e_1}$, combined with the Markov property. Subsection 6.5.1 is dedicated to the proof of Lemma 6.12.

https://doi.org/10.1017/apr.2023.28 Published online by Cambridge University Press

6.5.1. *Step 1: proof of Lemma 6.12*. Fix $x_\vee > 0$. Consider $\epsilon > 0$; this will be fixed later, but assume already that $\epsilon \leq \theta/8$. We recall that $\eta \leq \theta/8$ is assumed without loss of generality. Let $n \geq 3 \vee (2\theta)$, $x_0 \in B(0, n)$, and $\mathbf{u} \in S^d$ be such that both $\langle x_0, \mathbf{u}\rangle \geq \theta$ and $\|x_0^{(\perp \mathbf{u})}\| \leq x_\vee$ hold.

Compared to Proposition 6.3, the first main difference is that the jump is now almost instantaneous. The second is that, in order to have $g_\wedge > 0$, we have much less choice in the value of w when $\|x^{(\perp \mathbf{u})}\|$ is large. In particular, the variability of any particular jump will not be sufficient to wipe out the initial diffusion around x deduced from $x \in \mathcal{R}^{(n)}(t, c)$, but rather will make it even more diffuse.

To fix $\epsilon > 0$, let us first compute, for $\delta \in B(0, \eta)$ and $w \in B(-\theta\, \mathbf{u}, \epsilon)$,

$$\begin{aligned}\|x_0 + \delta\|^2 - \|x_0 + \delta + w\|^2 &= 2\,\langle x_0 + \delta, w\rangle - \|w\|^2\\ &\geq \left(\frac{7}{4} - \frac{9}{8}\times\left(\frac{1}{4} + \frac{9}{8}\right)\right)\,\theta^2 - 2\,\epsilon\, x_\vee,\end{aligned}$$

where we have exploited that $\langle \mathbf{u}, w\rangle \geq 7\,\theta/8$. We note that

$$c := \frac{7}{4} - \frac{9}{8}\times\left(\frac{1}{4} + \frac{9}{8}\right) = \frac{13}{64} > 0.$$

By taking $\epsilon := \{c\,\theta^2/(4\,x_\vee)\} \wedge \eta$ (recall that $\eta \leq \theta/8$), we thus ensure that $\|x_0 + \delta\|^2 > \|x_0 + \delta + w\|^2$. Note that ϵ does not depend on the specific choice of x_0.

Let $t_P := \epsilon/(2\,v)$. The initial condition for X, Y is taken as $x_I \in B(x, \eta/2)$ and $y_I \in [1/n,\ n]$. Let

$$\begin{aligned}g_\wedge &:= \inf\left\{\, g(x, w); x \in \bar{B}(x_0, \eta)\,,\ \ w \in \bar{B}(-\theta\, \mathbf{u}, \epsilon)\right\} > 0,\\ \mathcal{X}^M &:= [0, t_P] \times \mathbb{R}^d \times [0, f_\vee] \times [0, n],\\ \mathcal{J} &:= [0, t_P] \times B(-\theta\, \mathbf{u} + (\epsilon/2)\, \mathbf{e_1}, \epsilon/2) \times [0, f_\wedge] \times [0, g_\wedge].\end{aligned}$$

With the same reasoning as in the proof of Proposition 6.1, we obtain a change of probability $\mathrm{P}^G_{(x_I, y_I)}$ and an event $\mathcal{W}$ on which the random variable W is uniquely defined from M under $\mathrm{P}^G_{(x_I, y_I)}$, and such that it satisfies, a.s.,

$$X_{t_P} = x_I - (\epsilon/2)\, \mathbf{e_1} - \theta\, \mathbf{u} + (\epsilon/2)\, \mathbf{e_1} + W = x_I - \theta\, \mathbf{u} + W,$$

where the density of W is lower-bounded by d_W on $B(0, \epsilon/2)$, uniformly over x_I (given x) and y_I. We thus similarly obtain some constants c_P, $c'_P > 0$ independent of x_0 such that for any such x_0,

$$\begin{aligned}&\int_{B(x_0, \eta/2)} dx_I \int_{[1/n,\, n]} dy_I\, \mathrm{P}_{(x_I, y_I)}\left[(X, Y)_{t_P} \in (dx, dy); t_P < T_{(n)}\right]\\ &\qquad \geq c_P \int_{B(x_0, \eta/2)} dx_I \mathbf{1}_{B(x_I - \theta\, \mathbf{u},\, \epsilon/2)}(x) \cdot \mathbf{1}_{[1/n,\, n]}(y)\, dx\, dy\\ &\qquad \geq c'_P\, \mathbf{1}_{B(x_0 - \theta\, \mathbf{u},\, \eta/2 + \epsilon/3)}(x) \mathbf{1}_{[1/n,\, n]}(y)\, dx\, dy.\end{aligned}$$

We then reason similarly as in the proof of Corollary 6.1 as a consequence of Proposition 6.1. Assuming further that $x_0 \in \mathcal{R}^{(n)}(t, c)$ for some t, $c > 0$, we can deduce that

$$B(x - \theta\, \mathbf{u}, \epsilon/3) \subset \mathcal{R}^{(n)}(t + t_P, c \cdot c_P).$$

This is exactly the implication of Lemma 6.12, stated in terms of $\epsilon/3$ instead of ϵ. □

https://doi.org/10.1017/apr.2023.28 Published online by Cambridge University Press

6.5.2. *Step 2: Lemma 6.9 as a consequence of Lemmas 6.13 and 6.12.*

Step 2.1: $x_I \in \mathcal{R}^{(n_0)}(t_0, c_0)$. Let $x_I := -\theta \mathbf{e_1}$. We check that there exists $n_1 \geq 1$ such that $B(x_I, \eta/2)$ is a subset of $\mathcal{G}_{n_1}$. Since g is continuous, and thanks to Assumption [A], it is sufficient to prove that $\|x_I - z\mathbf{e_1} + \delta\| > \|x_I - z\mathbf{e_1} + \delta + w\|$ holds for any $z \in [0, \theta]$, $\delta \in \bar{B}(0, \eta)$, and $w \in \bar{B}(\theta\, \mathbf{e_1}, \eta)$:

$$\begin{aligned}
&\|x_I - z\mathbf{e_1} + \delta\|^2 - \|x_I - z\mathbf{e_1} + \delta + w\|^2 = 2\langle(\theta + z)\mathbf{e_1} - \delta, w\rangle - \|w\|^2\\
&\geq 2\,[\theta \cdot (\theta - \eta) - \eta \cdot (\theta + \eta)] - (\theta + \eta)^2\\
&\qquad = \theta^2 - 6\,\theta\,\eta - 3\,\eta^2 \geq \frac{13\theta^2}{64} > 0,
\end{aligned}$$

since $\eta \leq \theta/8$, as assumed above, just after (6.13). Applying Proposition 6.1 twice, we conclude that there exist $n_0 \geq 1$, t_0, $c_0 > 0$ such that $x_I \in \mathcal{R}^{(n_0)}(t_0, c_0)$.

Step 2.2: under the condition that $\langle x_F, \mathbf{e_1}\rangle := -\theta$. The purpose of this step is to prove the following lemma, in which we employ the notation $\pi_1 : x \mapsto \langle x, \mathbf{e_1}\rangle$.

Lemma 6.14. *For any $n \geq 1$ sufficiently large, there exist $t, c > 0$ such that $\pi_1^{-1}(-\theta) \cap B(0, n)$ is a subset of $\mathcal{R}^{(n)}(t, c)$.*

Let $x_F \in \pi_1^{-1}(-\theta) \cap B(0, n)$, where we assume that n is larger than n_0, 3, and 2θ. First, we define $\mathbf{u}$ as $\mathbf{e_1}$ if $x_F^{(\perp 1)} = 0$ and as $\mathbf{u} := x_F^{(\perp 1)}/\|x_F^{(\perp 1)}\|$ otherwise. Note that $\|x_F^{(\perp 1)}\| \leq n$. We consider the value of ϵ given by Lemma 6.13 for $x_\vee := n$ and make the following definitions:

$$K := \lfloor n\epsilon \rfloor + 1, \qquad \text{and for } 0 \leq k \leq K, \quad x_k := -\theta\, \mathbf{e_1} + \frac{k\, \|x_F^{(\perp 1)}\|}{K}\, \mathbf{u}.$$

This choice ensures that for any $k \in [\![0, K-1]\!]$, $x_{k+1} \in B(x_k, \epsilon)$, while $x_k \in B(0, n)$, $\langle x_k \,\big|\, \mathbf{e_1}\rangle \leq 0$, and $x_K = x_F$. Thanks to Step 2.1, $x_0 \in \mathcal{R}^{(n)}(t_0, c_0)$. Thus, by induction over $k \leq K$ with Lemma 6.13, $x_k \in \mathcal{R}^{(n)}\big(t_0 + k\, t_P,\ c_0\, [c_P]^k\big)$. In particular, there exist $t, c > 0$ such that $x \in \mathcal{R}^{(n)}(t, c)$, which concludes Step 2.2.

Step 2.3: the general case. Assume solely that $x \in \mathcal{B}(0, m)$. We consider the value of ϵ given by Lemma 6.12 for $x_\vee := m$. The choice of $\mathbf{u}$ is as in Step 2.2.

Let

$$K := \left\lfloor \frac{m + \theta}{\epsilon} \right\rfloor + 1, \qquad \text{so that } \frac{\langle x, \mathbf{e_1}\rangle + \theta}{K} \leq \epsilon, \tag{6.30}$$

and for $0 \leq k \leq K$, let

$$x_k := (-\theta + (k/K) \cdot (\langle x, \mathbf{e_1}\rangle + \theta))\, \mathbf{e_1} + (K - k)\, \theta\, \mathbf{u} + x^{(\perp 1)}.$$

In particular $\langle x_0, \mathbf{e_1}\rangle = -\theta$, and $x_K = x_F$, while for any $k \leq K - 1$, $x_{k+1} \in B(x_k, \epsilon)$, $x_k \in B(0, m + K\,\theta)$, and $\langle x_k, \mathbf{u}\rangle \leq \theta \vee \langle x, \mathbf{e_1}\rangle \leq m = x_\vee$.

Since $\langle x_0, \mathbf{e_1}\rangle = -\theta$, we can exploit Lemma 6.14 to prove that there exist $n \geq 1$ and $t_0, c_0 > 0$ independent of x_F such that $x_0 \in \mathcal{R}^{(n)}(t_0, c_0)$. Thanks to Lemma 6.12 and induction on k, we deduce that there exist $t_P, c_P > 0$ such that $x_k \in \mathcal{R}^{(n)}(t_0 + k\, t_P,\ c_0\, [c_P]^k)$. In particular, there exist $t, c > 0$ such that $x_F \in \mathcal{R}^{(n)}(t, c)$. □

https://doi.org/10.1017/apr.2023.28 Published online by Cambridge University Press

6.5.3. *Step 3: proof of Lemma* 6.10. The proof can be taken mutatis mutandis from the one given in Subsection 6.4.2. The fact that $B(x_I, \eta/2)$ is a subset of $\mathcal{G}_{n_1}$ is already proved in Step 2.1 (cf. Subsection 6.5.2), while Lemma 6.9 replaces Lemma 6.6, with identical implications. □

6.5.4. *Step 4: proof of Lemma* 6.11. **Remark.** The proof presented here efficiently exploits the lemmas we have already established but is probably very far from optimal in its estimates.

Step 4.1: study of $\mathcal{G}_n$. We look for conditions on $x \in \mathbb{R}^d$ that ensure that it belongs to $\mathcal{G}_n$ for some n. Let $x_\theta := x - (\theta - \eta/2)\mathbf{e_1}$. By definition of $\mathcal{G}_n$, it is necessary that $g(x_\theta - z\mathbf{e_1} + \delta, w) > 0$ for any $z \in [0, \eta/4]$, $\delta \in \bar{B}(0, \eta/2)$, and $w \in \bar{B}(\theta\, \mathbf{e_1}, \eta)$. The latter is equivalent, under Assumption [*A*], to $\|x_\theta - z\mathbf{e_1} + \delta\| > \|x_\theta + z\mathbf{e_1} + \delta + w\|$. We first restrict ourselves to the values of x such that $\pi_1(x) \le 0$, and we compute

$$\begin{aligned}
\|x_\theta - z\mathbf{e_1} + \delta\|^2 - \|x_\theta + z\mathbf{e_1} + \delta + w\|^2 &= -2\langle x_\theta + z\mathbf{e_1} + \delta, w\rangle - \|w\|^2 \\
&\ge 2\,(-\pi_1(x_\theta) - \eta/2)\cdot(\theta - \eta) - 2\big(\big\|x^{(\perp 1)}\big\| + \eta/2\big)\cdot\eta - (\theta + \eta)^2 \\
&\ge \big(-(7/32)\cdot\pi_1(x_\theta) - \big\|x^{(\perp 1)}\big\|/4\big)\cdot\theta + (7/4)\cdot(\theta - \eta/2)\cdot(\theta - \eta) \\
&\qquad - \eta\cdot(\theta - \eta) - \eta^2 - (\theta + \eta)^2. \\
&\ge \big(-(7/32)\cdot\pi_1(x_\theta) - \big\|x^{(\perp 1)}\big\|/4\big)\cdot\theta + (7\times 15\times 7 - 8\times 7 - 8 - 8\times 81)\cdot\theta^2/2^9 \\
&\ge \big(-(7/32)\cdot\pi_1(x_\theta) - \big\|x^{(\perp 1)}\big\|/4\big)\cdot\theta + 23\cdot\theta^2/2^9.
\end{aligned}$$

From these computations, we see that $g(x_\theta - z\mathbf{e_1} + \delta, w) > 0$ holds true provided $\pi_1(x) \le 0$ and $|\pi_1(x_\theta)| \ge (8/7)\cdot\big\|x^{(\perp 1)}\big\|$, and thus, a fortiori, if $|\pi_1(x)| \ge (8/7)\cdot\big\|x^{(\perp 1)}\big\|$. Since g is continuous, we deduce that for any $m \ge 1$, there exists $n \ge 1$ such that $\mathcal{G}_n$ contains the following set:

$$\{x \in B(0, m);\ -\pi_1(x) \ge (8/7)\cdot\big\|x^{(\perp 1)}\big\|\}.$$

Step 4.2: Let $\ell_I \ge 1$. Thanks to Step 4.1, we can find $n \ge \ell_I \vee 3$ such that $\mathcal{G}_n$ contains the following set:

$$\mathcal{A}_1 := \big\{x \in B(0, 2\ell_I);\ -\pi_1(x) \ge (8/7)\cdot\big\|x^{(\perp 1)}\big\|\big\}.$$

We go backwards in time from $\mathcal{A}$ by defining, for $t \ge 0$, $c > 0$,

$$\mathcal{R}'(t, c) := \big\{(x_I, y_I) \in \mathcal{G}_n; \mathrm{P}_{(x_I, y_I)}\left[\tau_{\mathcal{A}} \le t \wedge T_{(n)}\right] \ge c\big\}.$$

Similarly as for the proof of Lemma 6.8, by inductively applying Proposition 6.1, we deduce that $\mathcal{A}_1$ is a subset of $\mathcal{R}'(t_1, c_1)$ for some $t_1, c_1 > 0$.

Consider now any $x_I \in \bar{B}(0, \ell_I)$. If $x_I \notin \mathcal{A}_1$, let $u_* := 8\big\|x^{(\perp 1)}\big\|/(7v) + \pi_1(x)/v$ and $x_1 := x - v\,u_*\mathbf{e_1} \in \mathcal{A}_1$. If $x_I \in \mathcal{A}_1$, we simply define $x_1 := x_I$ and $u_* := 0$. Since $\big\|x^{(\perp 1)}\big\| \le \ell_I$, this choice necessarily satisfies $0 \le -\pi_1(x_1) = 8\big\|x^{(\perp 1)}\big\|/7 \le 8n/7$. In any case, $x_1 \in B(0, 2\ell_I)$, and thus $x_1 \in \mathcal{A}_1$. Since $\mathcal{A}_1 \subset \mathcal{R}'(t_1, c_1)$, and thanks to Lemma 6.3, there exists a value $c_D > 0$, uniform over x, such that $x_I \in \mathcal{R}'(t_1 + u_*, c_1\cdot c_D)$. Since u_* is upper-bounded by $2\ell_I$ and the sets $\mathcal{R}'(t, c)$ are increasing with t, we conclude that $\bar{B}(0, \ell_I)$ is a subset of $\mathcal{R}'(t_2, c_2)$ with $t_2 := t_1 + 2\ell_I$ and $c_2 := c_1\cdot c_D$. This completes the proof of Lemma 6.11. □

As mentioned at the beginning of Subsection 6.5, the last step of the proof of Theorem 4.2 can be taken mutatis mutandis from Subsection 6.4.4. With this, the proof of the theorem is complete.

https://doi.org/10.1017/apr.2023.28 Published online by Cambridge University Press

6.6. Proof of Theorem 4.5

We treat in this subsection the mixing for X when only advantageous mutations are occurring and the phenotype is unidimensional. The proof of Theorem 4.2 is handled under Assumption [A] and $d \geq 2$ in the same way as in Subsection 6.4.4, except that Lemmas 6.7–6.8 are replaced by the following ones, in the respective order. Note that only the first lemma has a different implication.

Lemma 6.15. *Suppose that $d = 1$, and that the assumptions [H] and [A] hold. Then, for any $m \geq 3$, there exist $n \geq m$, t, $c > 0$ such that $[-m, 0]$ is included in $\mathcal{R}^{(n)}(t, c)$.*

Lemma 6.16. *Suppose that $d = 1$, and that the assumptions [H] and [A] hold. Then there exists $n \geq 3$ which satisfies the following property for any $t_1, t_2 > 0$: there exist $t_R > t_1$ and $c_R > 0$ such that, for any $t \in [t_R, t_R + t_2]$ and $(x_0, y_0) \in \mathcal{A} \times [2, 3]$,*

$$\mathrm{P}_{(x_0,y_0)}\left[(X, Y)_t \in (dx, dy); t < T_{(n)}\right] \geq c_R\, \mathbf{1}_{\mathcal{A}}(x)\, \mathbf{1}_{[2,3]}(y)\, dx\, dy.$$

Lemma 6.17. *Suppose that $d = 1$, and that the assumptions [H] and [A] hold. Then, for any $\ell_I > 0$, there exist $c, t > 0$ and $n \geq \ell_I$ such that*

$$\forall\, (x_I, y_I) \in \mathcal{D}_{\ell_I}, \quad \mathrm{P}_{(x_I,y_I)}\big(\tau_{\mathcal{A}} \leq t \wedge T_{(n)}\big) \geq c. \tag{6.31}$$

Step 1: proof of Lemmas 6.15 and 6.16. Considering the calculations given in Step 4.1 (in Subsection 6.5.4), in this case where there is no contribution from $x^{(\perp 1)}$, we can conclude that for any m, there is $n \geq m$ such that $[-m, 0]$ is included in $\mathcal{G}_n$. Adapting the reasoning given in Subsections 6.4.1 and 6.4.2, respectively, we can directly conclude the proofs of Lemmas 6.15 and 6.16.

Note that the set first introduced in the proof of Lemma 6.7 here takes the form $[-\theta + \eta/6, -\theta + 5\eta/6]$. It is included in $[-m, 0]$ for any choice of $m \geq \theta$, so that Lemma 6.15 can indeed replace Lemma 6.6. □

Step 2: proof of Lemma 6.17. Let $(x_I, y_I) \in \mathcal{D}_{\ell_I}$.

Case 1: $x_I \geq -\theta$. Thanks to Lemma 6.3 with $u := x_I + \theta$, there exist $t_+, c_+ > 0$ which satisfy the following property for any $(x_I, y_I) \in \mathcal{D}_{\ell_I}$ such that $x_I \geq -\theta$:

$$\mathrm{P}_{(x_I,y_I)}\big(\tau_{\mathcal{A}} \leq t_+ \wedge T_{(n)}\big) \geq c_+.$$

Case 2: $x_I < -\theta$. We recall from the proof of Lemma 6.15 that there exists $n \geq 1$ such that $[-\ell_I, 0]$ is included in $\mathcal{G}_n$. In this set, the proof of Lemma 6.8 (given in Subsection 6.4.3) can be directly exploited to prove that there exist $t_-, c_- > 0$ which satisfy the following property for any $(x_I, y_I) \in \mathcal{D}_{\ell_I}$ such that $x_I \leq 0$:

$$\mathrm{P}_{(x_I,y_I)}\big(\tau_{\mathcal{A}} \leq t_- \wedge T_{(n)}\big) \geq c_-.$$

The combination of these two cases with $t := t_+ \vee t_-$ and $c := c_+ \wedge c_-$ concludes the proof of Lemma 6.17. □

Step 3: concluding the proof of Theorem 4.5. If we replace Lemmas 6.6, 6.7, and 6.8 by Lemmas 6.15, 6.16, and 6.17 in the proof given in Subsection 6.4.4, it is clear that the conclusion of Theorem 4.5 is reached. □

7. Almost perfect harvest

7.1. Proof of Theorem 4.4 in the case $d = 1$

7.1.1. *Definition of the stopping time and its elementary properties.* We consider a first process (X, Y) with some initial condition $(x_E,\ y_E) \in E$.

We will prove that considering $U_H = t_{\bar{\wedge}}$ is sufficient, except for exceptional behavior of the process. Given ϵ, $\rho > 0$, $t_{\bar{\wedge}}$ shall be chosen sufficiently small to ensure that, with probability close to 1 (the thresholds depending on ϵ and ρ), no jump has occurred before time $t_{\bar{\wedge}}$, and that the population size has not changed too much. We define

$$\delta y := \left(3\,\ell_E(\ell_E+1)\right)^{-1}, \quad y_\wedge := 1/(\ell_E+1) = 1/\ell_E - 3\,\delta y, \quad y_\vee := \ell_E + 1 > \ell_E + 3\,\delta y,$$
$$T_{\delta y} := \inf\{t \geq 0;\, |Y_t - y_E| \geq 2\,\delta y\} < \tau_\partial. \tag{7.1}$$

We recall that we can upper-bound the first jump time of X by

$$T_J := \inf\{t \geq 0;\, M([0, t] \times \mathcal{J}) \geq 1\}, \tag{7.2}$$

where $\mathcal{J}$ is defined as in Subsection 6.1.

- On the event $\left\{t_{\bar{\wedge}} < T_{\delta y} \wedge T_J \wedge \tau_\partial\right\}$, we set $U_H := t_{\bar{\wedge}}$.
- On the event $\left\{T_{\delta y} \wedge T_J \wedge \tau_\partial \leq t_{\bar{\wedge}}\right\}$, we set $U_H := \infty$.

Before we turn to the details of the proof of Theorem 4.4, we first give the main scheme for proving the following lemma, noting that we will not go too deeply into the details of this proof.

Lemma 7.1. *We can define a stopping time U_H^∞ extending the above definition of U_H as described in Theorem 4.4.*

7.1.2. *Step 1: main argument for the proof of Lemma 7.1.* Recall (with simplified notation) that considering the process (X, Y) with initial condition (x, y), we define, for some $t > 0$, $U_H := t$ on the event $\{t < T_{\delta y} \wedge T_J\}$, and $U_H := \infty$ otherwise, where

$$\begin{aligned}
T_{\delta y} &:= \inf\{s \geq 0;\, |Y_s - y| \geq 2\,\delta y\} < \tau_\partial && \text{for some } \delta y > 0,\\
T_J &:= \inf\{s \geq 0;\, M([0, s] \times \mathcal{J}) \geq 1\},\\
\mathcal{J} &:= \mathbb{R}^d \times [0, f_\vee] \times [0, g_\vee] && \text{for some } f_\vee,\, g_\vee > 0.
\end{aligned}$$

Recursively, we also define

$$\tau_E^{i+1} := \inf\left\{s \geq \tau_E^i + t\!:\! X_s \in E\right\} \wedge \tau_\partial, \text{ and } \tau_E^0 = 0,$$

and on the event $\left\{\tau_E^i < \tau_\partial\right\}$, for any i, we set

$$\begin{aligned}
T_{\delta y}^i &:= \inf\left\{s \geq \tau_E^i;\, \left|Y_s - Y\!\left(\tau_E^i\right)\right| \geq 2\,\delta y\right\},\\
U_j^i &:= \inf\left\{s \geq 0;\, M\!\left(\left[\tau_E^i,\, \tau_E^i + s\right] \times \mathcal{J}\right) \geq 1\right\},\\
U_H^\infty &:= \inf\left\{\tau_E^i + t;\, t \geq 0,\ \tau_E^i < \infty,\ \tau_E^i + t < T_{\delta y}^i \wedge U_j^i\right\},
\end{aligned}$$

where, in this notation, the infimum equals ∞ if the set is empty, $T_{\delta y}^i := \infty$, and $U_j^i = \infty$ on the event $\left\{\tau_\partial \leq \tau_E^i\right\}$.

https://doi.org/10.1017/apr.2023.28 Published online by Cambridge University Press

The proof that all these random times define stopping times is classical, although very technical, and the reader is spared the details. The main point is that there is a.s. a positive gap between any of these iterated stopping times. We can thus ensure recursively in I that there exists a sequence of stopping times with discrete values $(\tau_E^{i,(n)}, T_{\delta y}^{i,(n)}, U_j^{i,(n)})_{\{i\leq I, n\geq 1\}}$, such that a.s., for n sufficiently large and $1 \leq i \leq I$,

$$\tau_E^i \leq \tau_E^{i,(n)} \leq \tau_E^i + 1/n < \tau_E^i + t,$$
$$T_{\delta y}^i \leq T_{\delta y}^{i,(n)} \leq T_{\delta y}^i + 1/n, \qquad U_j^i \leq U_j^{i,(n)} \leq U_j^i + 1/n.$$

It is obvious that U_H^∞ coincides with U_H on the event $\{U_H \wedge \tau_\partial \leq \tau_E^1\}$, while the Markov property at time τ_E^1 and the way U_H^∞ is defined implies that on the event $\{\tau_E^1 < U_H \wedge \tau_\partial\}$, $U_H^\infty - \tau_E^1$ indeed has the same law as the $\widetilde{U}_H^\infty$ associated to the process $(\widetilde{X}, \widetilde{Y})$ solving the system (4.7) with initial condition $(X(\tau_E^1),\ Y(\tau_E^1))$. □

7.1.3. *Step 2: end of the proof of Theorem 4.4 when $d = 1$.* Let $\ell_E \geq 1$, ϵ, $\rho > 0$ be prescribed. We first require $t_{\overline{\wedge}} \leq 1$ to be sufficiently small.

Note that our definitions ensure that for any $t < t_{\overline{\wedge}} \wedge T_{\delta y} \wedge T_J$, we have a.s.

$$(X_t, Y_t) \in [-\ell_E - 1,\ \ell_E] \times [y_\wedge,\ y_\vee].$$

Thanks to Theorem 6.1, with some constant C_G uniform over any $(x_E, y_E) \in E$, we have

$$\begin{aligned}P_{(x_E, y_E)}\left(T_{\delta y} < t_{\overline{\wedge}} \wedge T_J\right) &\leq C_G\, P^G_{(x_E, y_E)}\left(T_{\delta y} < t_{\overline{\wedge}} \wedge T_J\right)\\ &\leq C_G\, P^G_0\left(T_{\delta y} < t_{\overline{\wedge}}\right) \to 0 \text{ as } t_{\overline{\wedge}} \to 0,\end{aligned}$$

where $T_{\delta y}$ under P_0^G denotes the first time the process $|B|$ reaches δy, with B a standard Brownian motion. Moreover,

$$P_{(x_E, y_E)}\left(T_J < t_{\overline{\wedge}} \wedge T_{\delta y}\right) \leq P\left(M([0, t_{\overline{\wedge}}] \times \mathcal{J}) \geq 1\right) \leq \nu(\mathbb{R}) \cdot f_\vee \cdot t_{\overline{\wedge}} \to 0 \text{ as } t_{\overline{\wedge}} \to 0.$$

By choosing $t_{\overline{\wedge}}$ sufficiently small, we can thus ensure the following property for any $(x_E, y_E) \in E$:

$$\begin{aligned}P_{(x_E, y_E)}(U_H = \infty, t_{\overline{\wedge}} < \tau_\partial) &\leq P_{(x_E, y_E)}\left(T_{\delta y} < t_{\overline{\wedge}} \wedge T_J\right) + P_{(x_E, y_E)}\left(T_J < t_{\overline{\wedge}} \wedge T_{\delta y}\right)\\ &\leq \epsilon\, e^{-\rho} \leq \epsilon\ \exp(-\rho\, t_{\overline{\wedge}}).\end{aligned} \tag{7.3}$$

On the event $\{t_{\overline{\wedge}} < T_{\delta y} \wedge T_J\}$, the following two properties hold: $X_{U_H} = x_E - v\, t_{\overline{\wedge}}$ and $Y_{U_H} \in [y_E - 2\delta y, y_E + 2\delta y]$. Indeed, as in the proof of Lemma 6.3, we have chosen our stopping times to ensure that no jump for X can occur before time $T_J \wedge t_{\overline{\wedge}} \wedge T_{\delta y}$. We also rely on the Girsanov transform and Theorem 6.1 to prove that, during the time-interval $[0, t_{\overline{\wedge}}]$, Y is indeed sufficiently diffused (since we are now interested in an upper bound, we can neglect the effect of assuming $t_{\overline{\wedge}} < T_{\delta y}$). This leads us to conclude that there exists $D^X > 0$ such that for any $x_E \in [-\ell_E, \ell_E]$ and $y_E \in [1/\ell_E,\ \ell_E]$,

$$\begin{aligned}&P_{(x_E, y_E)}\left[(X,\ Y)(U_H) \in (dx,\ dy); U_H < \tau_\partial\right]\\ &\qquad\qquad \leq D^X\, \mathbf{1}_{[y_E - 2\,\delta y,\, y_E + 2\,\delta y]}(y)\, \delta_{x_E - v\, t_{\overline{\wedge}}}(dx)\, dy.\end{aligned} \tag{7.4}$$

With ζ the uniform distribution over $\mathcal{D}_1$, thanks to Theorem 4.2, there exist c_M, $t_M > 0$ such that

$$P_\zeta\left[(X, Y)_{t_M} \in (dx',\ dy')\right] \geq c_M\, \mathbf{1}_{\left\{(x', y') \in \mathcal{D}_{L_E}\right\}}\, dx'\, dy'.$$

https://doi.org/10.1017/apr.2023.28 Published online by Cambridge University Press

The idea is then to let X decrease until it reaches $x_E - v\, t_{\overline{\wedge}}$, by ensuring that no jump occurs. We then identify u as the time needed for this to happen. Then, thanks to Theorem 6.1 and Lemma 6.2, we deduce a lower bound on the density of Y on $[y_E - 2\,\delta y,\ y_E + 2\,\delta y]$. We have already proved a stronger result for Lemma 6.3, which we leave it to the reader adapt to obtain the following property: for any $t_{\overline{\wedge}} > 0$, there exists d_2^X such that, for any $x_E \in [-\ell_E,\ \ell_E]$ and $y_E \in [1/\ell_E,\ \ell_E]$, there exists a stopping time V such that

$$\mathrm{P}_\zeta\left[(X,\ Y)(V) \in (dx,\ dy)\right] \geq d_2^X\, c_M\, \mathbf{1}_{[y_E-2\,\delta y,\ y_E+2\,\delta y]}(y)\, \delta_{x_E - v\, t_{\overline{\wedge}}}(dx)\, dy. \tag{7.5}$$

The proper definition of V is given by $V := t_M + t_{\overline{\wedge}} + \left(X_{t_M} - x_E\right)/v \geq t_M$ on the event $\left\{X_{t_M} \in \left[x_E,\ x_E + v\, t_{\overline{\wedge}}\right]\right\} \cap \{Y_{t_M} \in [y_E - \delta y/2,\ y_E + 2\,\delta y/2]\}$ (it can be set arbitrarily to t_M otherwise).

Thanks to Lemma 7.1, (7.3), (7.4), and (7.5), we conclude the proof of Theorem 4.4, with $c := D^X/\left(d_2^X\, c_M\right)$. □

7.2. Proof of Theorem 4.7

Except that we exploit Theorem 4.5 instead of 4.2, which constrains the shape of E, the proof is immediately adapted from Subsection 7.1. □

7.3. Proof of Theorem 4.4 in the case $d \geq 2$

The difficulty in this case is that, as long as no jump has occurred, X_t stays confined to the line $x + \mathbb{R} \cdot \mathbf{e_1}$. The 'harvest' thus cannot occur before a jump. Thus, we first wait for a jump to diffuse on $\mathbb{R}^d$ and then let Y diffuse independently in the same way as in Subsection 7.1. These two steps are summarized in the following.

Proposition 7.1. *Given any $\rho > 0$, $E \in \mathbf{D}$, and $\epsilon_X \in (0,\ 1)$, there exist $t^X,\ c^X,\ x_\vee^X > 0$ and $0 < y_\wedge^X < y_\vee^X$ which satisfy the following property for any $(x_E,\, y_E) \in E$: there exists a stopping time U^X such that*

$$\left\{\tau_\partial \wedge t^X \leq U^X\right\} = \left\{U^X = \infty\right\}, \quad \mathrm{P}_{(x_E, y_E)}\left(U^X = \infty\ ,\ t^X < \tau_\partial\right) \leq \epsilon_X\ \exp\left(-\rho\, t^X\right),$$
$$\text{and } \mathrm{P}_{(x_E, y_E)}\left(X\left(U^X\right) \in dx;\, Y\left(U^X\right) \in \left[y_\wedge^X,\, y_\vee^X\right],\ U^X < \tau_\partial\right) \leq c^X\, \mathbf{1}_{B(0, x_\vee^X)}(x)\, dx.$$

We defer the proof to Subsection 7.3.2.

Proposition 7.2. *Given any $\rho,\ x_\vee^X > 0,\ 0 < y_\wedge^X < y_\vee^X$, and $\epsilon_Y \in (0,\ 1)$, for any t^Y sufficiently small, there exist $c^Y > 0$ and $0 < y_\wedge^Y < y_\vee^Y$ which satisfy the following property for any $(x,\, y) \in B\left(0,\, x_\vee^X\right) \times \left[y_\wedge^X,\, y_\vee^X\right]$: there exists a stopping time T^Y such that*

$$\mathrm{P}_{(x,y)}\left(T^Y \leq t^Y \wedge \tau_\partial\right) \leq \epsilon_Y\ \exp\left(-\rho\, t^Y\right),$$
$$\text{and } \mathrm{P}_{(x,y)}\left((X,\, Y)\left(t^Y\right) \in (dx,\, dy);\, t^Y < T^Y \wedge \tau_\partial\right) \leq c^Y\ \delta_{\{x - v\, t^Y\, \mathbf{e_1}\}}(dx)\, \mathbf{1}_{\left[y_\wedge^Y,\, y_\vee^Y\right]}(y)\, dy.$$

The proof of Lemma 7.2 is taken mutatis mutandis from the one in Subsection 7.1.3. It leads one to define U_H as below:

- $U_H := U^X + t^Y$ on the event $\left\{U^X < t^X \wedge \tau_\partial\right\} \cap \left\{t^Y < \widetilde{\tau_\partial} \wedge \widetilde{T}^Y\right\}$, where $\widetilde{\tau_\partial}$ and $\widetilde{T}^Y$ are defined respectively as τ_∂ and T^Y for the solution $\left(\widetilde{X}_t,\, \widetilde{Y}_t\right)$, defined on the event $\left\{U^X < t^X \wedge \tau_\partial\right\}$, of

https://doi.org/10.1017/apr.2023.28 Published online by Cambridge University Press

$$\begin{cases} \widetilde{X}_t = X\big(U^X\big) - v\,t\,\mathbf{e_1} \\ \quad + \displaystyle\int_{[U^X,U^X+t]\times\mathbb{R}^d\times(\mathbb{R}_+)^2} w\,\varphi\,\big(\widetilde{X}(s-U_X-),\ \widetilde{Y}(s-U_X),\ w,\ u_f,\ u_g\big)\ M\big(ds, dw, du_f, du_g\big), \\ \widetilde{Y}_t = Y\big(U^X\big) + \displaystyle\int_0^t \psi\,\big(\widetilde{X}(s-U_X),\ \widetilde{Y}(s-U_X)\big)\,ds + \int_{U^X}^{U^X+t} dB_r. \end{cases}$$

- Otherwise, $U_H := \infty$.

Lemma 7.2. *There exists a stopping time U_H^∞ extending the above definition of U_H as described in Theorem 4.4 (with $t = t^X + t^Y$ here).*

The proof of Lemma 7.2 is technical but classical from the way we define U^X and T^Y; it is similar to the proof of Lemma 7.1. The reader is spared this proof.

7.3.1. *Proof of Theorem 4.4 as a consequence of Propositions 7.1–7.2 and Lemma 7.2.* Given $\rho > 0$, $\epsilon \in (0, 1)$, and some $E \in \mathbf{D}$, we define $\epsilon_X := \epsilon/4$ and deduce from Proposition 7.3 the values t^X, c^X, $x_\vee^X$, $y_\wedge^X$, $y_\vee^X$ and the definition for the stopping times U^X with the associated properties.

With $\epsilon_Y := \epsilon\ \exp\big(-\rho\, t^X\big)/2$, we then deduce from Proposition 7.2 the values t^Y, c^Y, $y_\wedge^Y$, $y_\vee^Y$ and the stopping time T^Y with the associated properties, with the additional requirement that $t^Y \le \ln(2)/\rho$. With U_H defined, for some $(x, y) \in E$, as in Lemma 7.2, the following bound on U_H is clearly satisfied:

$$\big\{\tau_\partial \wedge \big(t^X + t^Y\big) \le U_H\big\} = \{U_H = \infty\}\,. \tag{7.6}$$

In addition, the probability of failure in the harvesting step is upper-bounded as follows:

$$\begin{aligned} \mathrm{P}_{(x,y)}\big(U_H = \infty,\, t^X + t^Y < \tau_\partial\big) &\le \epsilon_X\ \exp\big(-\rho\, t^X\big) + \epsilon_Y\ \exp\big(-\rho\, t^Y\big) \\ &\le \epsilon\ \exp\big(-\rho\,\big[t^X + t^Y\big]\big), \end{aligned} \tag{7.7}$$

where in the last inequality we have exploited the definitions of ϵ_X, ϵ_Y and the fact that $t^Y \le \ln(2)/\rho$ $\big($i.e. $1/2\ \le \exp\big(-\rho\, t^Y\big)\big)$. The upper bound on the density of the process at harvesting time U_H is deduced as follows:

$$\begin{aligned} &\mathrm{P}_{(x,y)}\big[(X,\, Y)\,(U_H) \in (dx,\, dy);\, U_H < \tau_\partial\big] \\ &\qquad \le c^X\, c^Y\, \mathbf{1}_{B\left(0,x_\vee^X + v\,t^Y\right)}(x)\, \mathbf{1}_{\left[y_\wedge^Y, y_\vee^Y\right]}(y)\, dx\, dy. \end{aligned} \tag{7.8}$$

For the opposite upper bound, we recall first that ζ is chosen to be uniform over the compact space Δ, which is included in some $\mathcal{D}_\ell$. Exploiting Theorem 4.5 on this set $\mathcal{D}_\ell$, we deduce that there exist t, $c > 0$ such that

$$\mathrm{P}_\zeta\big[(X,\, Y)\,(t) \in (dx,\, dy);\, t < \tau_\partial\big] \ge c\,\mathbf{1}_{B\left(0,x_\vee^X + v\,t^Y\right)}(x)\, \mathbf{1}_{\left[y_\wedge^Y, y_\vee^Y\right]}(y)\, dx\, dy. \tag{7.9}$$

Combining (7.6)–(7.9) completes the proof of Theorem 4.4 in the case $d \ge 2$. □

7.3.2. *Proof of Proposition 7.1.* For readability, note that most of the subscripts 'X' (except for t^X) from Proposition 7.1 are removed in this proof.

First, observe that without any jump, $\|X\|$ tends to infinity, which makes the population almost certainly doomed to extinction. We can thus find some time-limit $t_\vee$ such that, even

https://doi.org/10.1017/apr.2023.28 Published online by Cambridge University Press

with an amplification of order $\exp(\rho\, t_\vee)$, the event that the population survives without any mutation occurring in the time-interval $[0, t_\vee]$ is sufficiently exceptional. With this time-scale, we can find an upper bound $y_\vee$ on Y: that the population reaches such size before $t_\vee$ is a sufficiently exceptional event. For the lower bound, we exploit the fact that extinction is very strong when the population size is too small. Thus, the survival of the population—at least for a bit—after it declines below this lower bound $y_\wedge$ is also a sufficiently exceptional event.

The last part is needed to ensure that this first jump is indeed diffuse in X (which is why we need $\nu(dw)$ to have a density with respect to Lebesgue measure with the bound of [$H5$]).

For $y_\vee > \ell_E > 1/\ell_E > y_\wedge > 0$, $t_\vee$, $w_\vee > 0$, and initial condition $(x, y) \in E$, let

$$T_J := \inf\{t \geq 0;\, \Delta X_t \neq 0\}, \tag{7.10}$$

$$T_Y^\vee := \inf\{t \geq 0;\, Y_t = y_\vee\}, \qquad T_Y^\wedge := \inf\{t \geq 0;\, Y_t = y_\wedge\} < \tau_\partial. \tag{7.11}$$

On the event $\left\{T_J < t_\vee \wedge T_Y^\vee \wedge T_Y^\wedge\right\} \cap \{\|\Delta X_{T_J}\| < w_\vee\}$, we define $U := T_J$. Otherwise we set $U := \infty$.

To choose $y_\wedge$, $y_\vee$, $t_\vee$, and $w_\vee$, we refer to the following lemmas, which are treated as the five first steps of the proof of Proposition 7.1, which is completed in the sixth step.

Lemma 7.3. *For any* $\rho, \epsilon_1 > 0$*, there exists* $t_\vee > 0$ *such that*

$$\forall\,(x, y) \in E, \quad \mathrm{P}_{(x,y)}(t_\vee < T_J \wedge \tau_\partial) \leq \epsilon_1\, \exp(-\rho\, t_\vee).$$

Lemma 7.4. *For any* $t_\vee, \epsilon_2 > 0$*, there exists* $y_\vee > 0$ *such that*

$$\forall\,(x, y) \in E, \quad \mathrm{P}_{(x,y)}\big(T_Y^\vee < t_\vee \wedge \tau_\partial\big) \leq \epsilon_2.$$

Lemma 7.5. *For any* $t_S, \epsilon_3 > 0$*, there exists* $y_\wedge > 0$ *such that*

$$\forall\, x \in \mathbb{R}^d, \quad \mathrm{P}_{(x,y_\wedge)}(t_S < \tau_\partial) \leq \epsilon_3.$$

Lemma 7.6. *For any* $t_\vee, \epsilon_4 > 0$*, there exists* $w_\vee > 0$ *such that*

$$\forall\,(x, y) \in E, \quad \mathrm{P}_{(x,y)}(\|\Delta X_{T_J}\| \geq w_\vee\,,\; T_J < t_\vee \wedge \tau_\partial) \leq \epsilon_4.$$

Lemma 7.7. *For any* $t_\vee > 0$*, and any* $y_\vee > \ell_E > 1/\ell_E > y_\wedge > 0$*, there exist* $c, x_\vee > 0$ *such that*

$$\forall\,(x, y) \in E, \quad \mathrm{P}_{(x,y)}\big(X(U) \in dx;\, U < \tau_\partial\big) \leq c\, \mathbf{1}_{B(0,x_\vee)}(x)\, dx.$$

Step 1: proof of Lemma 7.3. Exploiting Assumption [$H3$], as long as $\|X\|$ is sufficiently large, we can ensure that the growth rate of Y is largely negative, leading to quick extinction. The proof is similar to that of Lemma 3.2.2 in [31], where more details can be found. We consider the autonomous process Y^D as an upper bound of Y where the growth rate is replaced by r_D. For any t_D and ρ, there exists r_D (a priori negative) such that whatever y_D is the initial condition of Y^D, survival of Y^D until t_D (i.e. $t_D < \tau_\partial^D$) happens with a probability smaller than $\exp(-2\,\rho\, t_D)$. Thanks to Assumption [$H3$], we define $x_\vee$ such that for any x, $\|x\| \geq x_\vee$ implies $r(x) \leq r_D$. We then deduce that

$$\forall\,(x, y), \quad \mathrm{P}_{(x,y)}(\forall\, t \leq t_D,\ \|X_t\| \geq x_\vee;\, t_D < \tau_\partial) \leq \sup_{y_D > 0}\ \mathrm{P}_{y_D}\left(t_D < \tau_\partial^D\right) \leq \exp(-2\,\rho\, t_D).$$

Let $t_E := (x_\vee + \ell_E)/v$ and assume $t_\vee \geq t_E$. A.s. on $\{t_\vee < T_J \wedge \tau_\partial\}$, for any $(x, y) \in E$,

$$\forall\, t_E \leq t \leq t_\vee, \quad \|X(t)\| = \|x - v\, t\, \mathbf{e_1}\| \geq x_\vee.$$

https://doi.org/10.1017/apr.2023.28 Published online by Cambridge University Press

Inductively applying the Markov property at times $t_\vee := t_E + k\, t_D$ for $k \geq 1$, we obtain

$$\forall\,(x, y), \quad \exp[\rho\, t_\vee]\, \mathrm{P}_{(x,y)}(t_\vee < T_J \wedge \tau_\partial) \leq \exp(\rho\, [t_E - k\, t_D]) \underset{k\to\infty}{\longrightarrow} 0.$$

Step 2: proof of Lemma 7.4. This is an immediate consequence of the fact that Y is upper-bounded by the process $Y^\vee$ given in (3.1) with initial condition ℓ_M. This bound is uniform in the dynamics of X_t and M and uniform for any $(x, y) \in E$. It is classical that a.s. $\sup_{t\leq t_\vee} Y_t^\vee < \infty$, which proves the lemma; see e.g. [4, Lemma 3.3].

Step 3: proof of Lemma 7.5. As in the proof of Proposition 4.2.3 in [31] (cf. Appendix D), we exploit $r_\vee$ as the upper bound of the growth rate of the individuals to relate to the formulas for continuous-state branching processes. Referring for instance to [27, Subsection 4.2], notably Lemma 5, it is classical that 0 is an absorbing boundary for these processes (we even have explicit formulas for the probability of extinction). This directly implies the result of the present lemma, that the probability of extinction tends uniformly to zero as the initial population size tends to zero.

Step 4: proof of Lemma 7.6. On the event $\{T_J < t_\vee \wedge \tau_\partial\}$, for any initial condition $(x, y) \in E$, there exists a compact subset K of $\mathbb{R}^d$ that contains $X_t = x - v\, t$ for any $t \in [0, T_J)$. Thanks to Assumption $[H2]$, there exists an upper bound $g_\vee$ of g that is valid on $K \times \mathbb{R}^d$.

Let $\epsilon_4 > 0$ and $\rho_W := (-1/t_\vee). \log(1 - \epsilon_4)$. We define $w_\vee$ such that $\nu(B(0, w_\vee)^c) \leq \rho_W / g_\vee$. Then we can couple the process X to an exponential random variable T_W of mean $1/\rho_W$ such that on the event $\{T_J < t_\vee \wedge \tau_\partial\} \cap \{\|\Delta X_{T_J}\| \geq w_\vee\}$, $T_J \leq T_W$ holds a.s. We can conclude with the following upper bound, valid for any $(x, y) \in E$:

$$\begin{aligned}\mathrm{P}_{(x,y)}(\|\Delta X_{T_J}\| \geq w_\vee\,,\ T_J < t_\vee \wedge \tau_\partial) &\leq \mathrm{P}(T_W < t_\vee) = 1 - \exp(-\rho_W\, t_\vee)\\ &\leq \epsilon_4.\end{aligned}$$

Note that under Assumption $[A]$, the jump at time T_J cannot make the process escape K. This provides a deterministic upper bound $w_\vee$ such that $\|\Delta X_{T_J}\| \geq w_\vee$ a.s. on $\{T_J < t_\vee \wedge \tau_\partial\}$.

Step 5: proof of Lemma 7.7. For $x_\vee := \ell_E + v\, t_\vee$, let

$$c := \sup\left\{\frac{g(x, w)\, \nu(w)}{\int_{\mathbb{R}^d} g(x, w')\, \nu(w')\, dw'};\ \|x\| \leq x_\vee\,,\ w \in \mathbb{R}^d\right\} < \infty. \tag{7.12}$$

We exploit a sigma-field $\mathcal{F}^*_{T_J}$ that includes the whole knowledge of the process until time T_J, except for the size of the jump at this time. (It is rigorously defined and studied in Appendix B.) Conditionally on $\mathcal{F}^*_{T_J}$ on the event $\{U < \tau_\partial\} \in \mathcal{F}^*_{T_J}$, the law of $X(T_J)$ is given by

$$\frac{g(X[T_J-], x - X[T_J-]) \cdot \nu(x - X[T_J-])}{\int_{\mathbb{R}^d} g(X[T_J-], w') \cdot \nu(w')\, dw'}\, dx.$$

Note also that a.s. $\|X[T_J-]\| \leq \ell_E + v\, t_\vee = x_\vee$ (since no jump has occurred yet).

Since $\|\Delta X_{T_J}\| \leq w_\vee$ on the event $\{U < \tau_\partial\}$, with $\bar{x}_\vee := x_\vee + w_\vee$, we get the following upper bound of the law of $X(T_J)$:

$$\begin{aligned}\mathrm{P}_{(x,y)}\,(X(U) \in dx; U < \tau_\partial) &= \mathrm{P}_{(x,y)}\left(\mathrm{E}\left[X(U) \in dx \,\middle|\, \mathcal{F}^*_{T_J}\right]; U < \tau_\partial\right)\\ &\leq c\, \mathbf{1}_{B(0,\bar{x}_\vee)}(x)\, dx.\end{aligned}$$

https://doi.org/10.1017/apr.2023.28 Published online by Cambridge University Press

Step 6: concluding the proof of Proposition 7.1. Let $\ell_E, \rho, \epsilon > 0$. We first deduce the existence of $t_\vee$, thanks to Lemma 7.3, such that

$$\forall (x, y) \in E, \quad \mathrm{P}_{(x,y)}(t_\vee < T_J \wedge \tau_\partial) \leq \epsilon \, \exp(-\rho \, t_\vee)/8. \tag{7.13}$$

Thanks to Lemma 7.4, we deduce the existence of some $y_\vee > 0$ such that

$$\forall (x, y) \in E, \quad \mathrm{P}_{(x,y)}(T_Y^\vee < t_\vee \wedge \tau_\partial) \leq \epsilon \, \exp(-\rho \, t_\vee)/8. \tag{7.14}$$

We could take any value for t_S (so possibly 1), but $t_S = \log(2)/\rho$ seems somewhat more practical. We then deduce the existence of $y_\wedge$, thanks to Lemma 7.5, such that

$$\sup_{\{x \in \mathbb{R}^d\}} \mathrm{P}_{(x,y_\wedge)}(t_S < \tau_\partial) \leq \epsilon \, \exp(-\rho \, t_\vee)/8.$$

This implies that for any $(x, y) \in E$,

$$\begin{aligned}
&\mathrm{P}_{(x,y)} \left(t_\vee + t_S < \tau_\partial \, , \; T_Y^\wedge < t_\vee \wedge \tau_\partial \wedge T_Y^\vee \wedge T_J\right) \\
&\quad \leq \mathrm{E}_{(x,y)} \left(\mathrm{P}_{(X_{T_Y^\wedge}, y_\wedge)}(t_S < \tau_\partial); T_Y^\wedge < t_\vee \wedge \tau_\partial \wedge T_Y^\vee \wedge T_J\right) \\
&\quad \leq \epsilon \, \exp(-\rho \, t_\vee)/8.
\end{aligned} \tag{7.15}$$

We choose $w_\vee$, thanks to Lemma 7.6, such that

$$\forall (x, y) \in E, \quad \mathrm{P}_{(x,y)} \left(\|\Delta X_{T_J}\| \geq w_\vee \, , \; T_J < t_\vee \wedge \tau_\partial\right) \leq \epsilon \, \exp(-\rho \, t_\vee)/8. \tag{7.16}$$

Thanks to Lemma 7.7, there exist $c, x_\vee > 0$ such that

$$\forall (x, y) \in E, \quad \mathrm{P}_{(x,y)}\left(X(U) \in dx; U < \tau_\partial\right) \leq c \, \mathbf{1}_{B(0,x_\vee)}(x) \, dx.$$

Thanks to the construction of U, and noting that $t^X := t_\vee + t_S$, it is clear that $U \geq \tau_\partial \wedge t^X$ is equivalent to $U = \infty$. Combining (7.13), (7.14), (7.15), and (7.16), we have

$$\begin{aligned}
&\mathrm{P}_{(x,y)}(U = \infty \, , \; t_\vee + t_S < \tau_\partial) \\
&\quad \leq \mathrm{P}_{(x,y)}(t_\vee < T_J \wedge \tau_\partial) + \mathrm{P}_{(x,y)} \left(\|\Delta X_{T_J}\| \geq w_\vee \, , \; T_J < t_\vee \wedge \tau_\partial\right) \\
&\qquad + \mathrm{P}_{(x,y)}(T_Y^\vee < t_\vee \wedge \tau_\partial) + \mathrm{P}_{(x,y)} \left(t_\vee + t_S < \tau_\partial \, , \; T_Y^\wedge < t_\vee \wedge \tau_\partial \wedge T_Y^\vee \wedge T_J\right) \\
&\quad \leq \epsilon \, \exp(-\rho \, t_\vee)/2 = \epsilon \, \exp\left(-\rho \, t^X\right).
\end{aligned}$$

This completes the proof of Proposition 7.1. □

The proof of Theorem 4.4 in the case $d \geq 2$ is now completed. All the theorems have been proved at this point. There are three sections in the appendix. Appendix A is devoted to the elementary properties exploited to deduce $(A2)$. In Appendix B, we precisely define the filtration $\mathcal{F}^*_{T_J}$ that carries the information up to the jumping time. We conclude with Appendix C, which gives the first results of some simulations to help illustrate the discussion in Subsection 2.3.

Appendix A. The inequalities exploited for the escape

Recall that $V_E := \tau_E \wedge \tau_\partial \wedge t_\vee$, where $t_\vee > 0$ is a technical value whose only purpose is to guarantee the finiteness of the exponential moments.

https://doi.org/10.1017/apr.2023.28 Published online by Cambridge University Press

A.1. Lemma 5.1 implies Proposition 5.2

Thanks to Lemma 5.1, for n_∞ sufficiently large, we obtain by induction and the Markov property that

$$\forall\, n > 0, \quad \mathrm{P}_n(k\, t < \tau_\downarrow^D) \le \epsilon^k.$$

Thus, by choosing ϵ sufficiently small (for a given value of $t > 0$), we ensure that

$$C_\infty^N := \sup_{\{n>0\}} \left\{ \mathrm{E}_n[\, \exp(\rho \tau_\downarrow^D)] \right\} < +\infty.$$

A fortiori, with $T_\downarrow := \inf\{t, N_t \le n_\infty\} \wedge \tau_E \le \tau_\downarrow^D$,

$$\sup_{(x,n)} \left\{ \mathrm{E}_{(x,n)}[\, \exp(\rho\, T_\downarrow)] \right\} \le C_\infty^N < \infty.$$

At time $T_\downarrow$, the process is either in E, in $\mathcal{T}_+$, or in $\mathcal{T}_\infty^X$. Thus,

$$\begin{aligned} \mathrm{E}_{(x,n)}[\, \exp(\rho\, V_E)] &\le \mathrm{E}_{(x,n)}[\, \exp(\rho\, T_\downarrow)\, ; (X, N)_{T_\downarrow} \in E] \\ &\quad + \mathrm{E}_{(x,n)}\Big[\, \exp(\rho\, T_\downarrow) \mathrm{E}_{(X,N)_{T_\downarrow}}[\, \exp(\rho\, \widetilde{V}_E)]\, ; (X, N)_{T_\downarrow} \in \mathcal{T}_+ \cup \mathcal{T}_\infty^X \Big], \end{aligned}$$

with the Markov property on the event $\{(X, N)_{T_\downarrow} \in \mathcal{T}_+ \cup \mathcal{T}_\infty^X\}$ (and the fact that $(V_E - T_\downarrow)_+ \le t_\vee$). Therefore, $\mathcal{E}_\infty^N \le C_\infty^N\, (1 + \mathcal{E}_X + \mathcal{E}_\infty^X)$, which concludes the proof of Proposition 5.2. □

A.2. Lemma 5.2 implies Proposition 5.3

Let $\rho,\ \epsilon,\ n_\infty > 0$ ($c > 0$ is the same as for the definition of Z). For simplicity, we choose $t := \log(2)/\rho > 0$ (i.e. $\exp[\rho\, t] = 2$), and assume without loss of generality $t < t_h$. We choose $r_\vee \in \mathbb{R}$, according to Lemma 5.2, such that

$$\begin{aligned} &\forall\, n > 0,\ \forall\, r \le r_\vee, \quad \mathrm{P}_n\left(t < \tau_\partial^D\right) \le e^{-\rho\, t}/2 = 1/4, \\ &\forall\, r \le r_\vee, \quad \mathrm{P}_{n_\infty}\left(T_\infty^D \le t\right) + \mathrm{P}_{n_\infty}\left(N_t^D \ge n_\infty\right) \le \epsilon/4. \end{aligned}$$

Since $\limsup_{\|x\|\to\infty} r(x) = -\infty$, with n_E chosen sufficiently large, we have that

$$\forall\, x \notin B(0, n_E),\ \ r(x) \le r_\vee.$$

Let (X, N) with initial condition $(x, n) \in \mathcal{T}_\infty^X$. In the following, we define

$$\begin{aligned} &T_\infty^N := \inf\{t \ge 0,\ N_t \ge n_c\}, \qquad \tau_0 := \inf\{t > 0,\ (X,\ N)_t \in \mathcal{T}_0\}, \\ &T := t \wedge T_\infty^N \wedge \tau_0 \wedge \tau_E \wedge \tau_\partial. \end{aligned} \tag{A.1}$$

Since, on the event $\{T = t\}$, either $N_t \ge n_\infty$ or $(X, Y)_t \in \mathcal{T}_+ \cup \mathcal{T}_\infty^X$, we have

$$\begin{aligned} \mathrm{E}_{(x,n)}[\, \exp(\rho\, V_E)] &= \mathrm{E}_{(x,n)}[\, \exp(\rho\, T)\, ; T = V_E] + \mathrm{E}_{(x,n)}[\, \exp(\rho\, V_E)\, ; T = \tau_0] \\ &\quad + \mathrm{E}_{(x,n)}[\, \exp(\rho\, V_E)\, ; T = t] + \mathrm{E}_{(x,n)}\big[\, \exp(\rho\, V_E)\, ; T = T_\infty^N\big] \\ &\le \exp(\rho\, t)\ \ (1 + \mathcal{E}_0) + \exp(\rho\, t) \cdot \mathrm{P}_{(x,n)}[T = t] \cdot \left(\mathcal{E}_X + \mathcal{E}_\infty^X\right) \\ &\quad + \exp(\rho\, t) \cdot \left(\mathrm{P}_{(x,n)}\big[T = T_\infty^N\big] + \mathrm{P}_{(x,n)}[N_t \ge n_\infty,\ \ T = t]\right) \cdot \mathcal{E}_\infty^N, \end{aligned}$$

thanks to the Markov property. Now, by (A.1), N^D is an upper bound of N before T. Thus, by our definitions of t, n_E, $r_\vee$,

$$\mathrm{E}_{(x,n)}[\, \exp(\rho\, V_E)] \le 2 \cdot (1 + \mathcal{E}_0) + (1/2) \cdot (\mathcal{E}_X + \mathcal{E}_\infty^X) + (\epsilon/2) \cdot \mathcal{E}_\infty^N.$$

Taking the supremum over $(x, n) \in \mathcal{T}_\infty^X$ in the last inequality concludes the proof of Proposition 5.3, in that it yields $\mathcal{E}_\infty^X \le 4\ \ (1 + \mathcal{E}_0 + \mathcal{E}_X) + \epsilon\, \mathcal{E}_\infty^N$. □

A.3. Proof of Proposition 5.5

The equation $N_t^U = n_0 + \int_0^t r_+ N_s^U ds + \sigma \int_0^t \sqrt{N_s^U}\, dB_s^N$ defines an upper bound of N on $[0, t_D]$ provided $n \le n_0$, while N^U is a classical branching process. The survival of (X, N) beyond t_D clearly implies the survival of N^U beyond t_D. Let us define ρ_0 by the relation $\mathrm{P}_{n_0}\left(t_D < \tau_\partial^U\right) =: \exp(-\rho_0\, t_D)$. For a branching process like N^U, it is classical that $\rho_0 \to \infty$ as $n_0 \to 0$. Indeed, with $u(t, \lambda)$ the Laplace exponent of N^U (cf. e.g. [27, Subsection 4.2], notably Lemma 5), we have $\mathrm{P}_{n_0}\left(\tau_\partial^U \le t_D\right) = \exp\left[-n_0 \lim_{\lambda\to\infty} u(t_D, \lambda)\right] \to 1$ as $n_0 \to 0$.

So we can impose that $\rho_0 > \rho$, and even that $\epsilon' := 2\exp(-(\rho_0 - \rho)\, t_D)$ is sufficiently small to make transitions from $\mathcal{T}_0$ to $\mathcal{T}_0$, $\mathcal{T}_+$, $\mathcal{T}_\infty^N$, or $\mathcal{T}_\infty^X$ of little incidence. We require notably that $\epsilon' \le 1$. We have

$$\begin{aligned} \mathrm{E}_{(x,n)}[\exp(\rho V_E)] \le \mathrm{E}_{(x,n)}\Big[&\exp(\rho V_E);\, V_E < t_D\Big] \\ &+ \mathrm{E}_{(x,n)}\Big[\exp(\rho V_E); (x, n)_{t_D} \in \mathcal{T}_0 \cup \mathcal{T}_+ \cup \mathcal{T}_\infty^N \cup \mathcal{T}_\infty^X\Big] \\ &\le \exp[\rho\, t_D] + \exp(\rho\, t_D)\cdot\left(\mathcal{E}_0 + \mathcal{E}_X + \mathcal{E}_\infty^N + \mathcal{E}_\infty^X\right)\cdot \mathrm{P}_{(x,n)}(t_D < \tau_\partial). \end{aligned}$$

Thus, taking the supremum over $(x, n) \in \mathcal{T}_0$ yields

$$\mathcal{E}_0 \le e^{\rho t_D} + (\epsilon'/2)\cdot\left(\mathcal{E}_0 + \mathcal{E}_X + \mathcal{E}_\infty^N + \mathcal{E}_\infty^X\right).$$

Since $\epsilon' \le 1$, this provides the following upper bound on $\mathcal{E}_0$:

$$\mathcal{E}_0 \le 2e^{\rho t_D} + \epsilon'\cdot\left(\mathcal{E}_X + \mathcal{E}_\infty^N + \mathcal{E}_\infty^X\right).$$

Since ϵ' tends to 0 as n_0 tends to 0, this concludes the proof of Proposition 5.5. □

Appendix B. A specific filtration for jumps

This appendix extends to our case the intuition already present in [32, Subsection 5.4.2]: there exists a sigma-field $\mathcal{F}_{T_J}^*$ which informally 'includes the information carried by M and B' up to the jump time T_J, except the realization of the jump itself. We define

$$\mathcal{F}_{T_J}^* := \sigma\left(A_s \cap \{s < T_J\}\,;\, s > 0,\ A_s \in \mathcal{F}_s\right).$$

Properties of $\mathcal{F}_{T_J}^*$: If Z_s is $\mathcal{F}_s$-measurable and $s < t \in (0, \infty]$, then $Z_s\, \mathbf{1}_{\{s<T_J\le t\}}$ is $\mathcal{F}_{T_J}^*$-measurable.

Lemma B.1. *For any left-continuous and adapted process Z, Z_{T_J} is $\mathcal{F}_{T_J}^*$-measurable. Reciprocally, $\mathcal{F}_{T_J}^*$ is in fact the smallest σ-algebra generated by these random variables. In particular, for any stopping time T, $\{T_J \le T\} \in \mathcal{F}_{T_J}^*$.*

Denote by W the additive effect on X of the first jump of X, occurring at time T_J.

Lemma B.2. *For any $h:\mathbb{R} \to \mathbb{R}_+$ measurable and $(x, y) \in (-L, L) \times \mathbb{R}_+$, we have*

$$\mathrm{E}_{(x,y)}\left[h(W)\,\middle|\,\mathcal{F}_{T_J}^*\right] = \frac{\int_{\mathbb{R}} h(w) f(Y_{T_J}) g(X_{T_J-}, w)\, \nu(dw)}{\int_{\mathbb{R}} f(Y_{T_J}) g(X_{T_J-}, w')\, \nu(dw')}.$$

Proof of Lemma B.1

For any left-continuous and adapted process Z,

$$Z_{T_J} = \lim_{n\to\infty} \sum_{k\le n^2} Z_{\frac{k-1}{n}}\, \mathbf{1}_{\left\{\frac{k-1}{n} < T_J \le \frac{k}{n}\right\}},$$

https://doi.org/10.1017/apr.2023.28 Published online by Cambridge University Press

where, by the previous property and the fact that Z is adapted, we know that

$$Z_{\frac{k-1}{n}} \mathbf{1}_{\left\{\frac{k-1}{n} < T_J \le \frac{k}{n}\right\}}$$

is $\mathcal{F}^*_{T_J}$-measurable for any k, n. Reciprocally, for any $s > 0$ and $A_s \in \mathcal{F}_s$,

$$\mathbf{1}_{A_s \cup \{s < T_J\}} = \lim_{n \ge 1} Z^n_{T_J}, \qquad \text{where } Z^n_t := \{1 \wedge [n\,(t-s)_+]\} \cdot \mathbf{1}_{A_s}.$$

Now, for any stopping time T and any $t \ge 0$, we have $\{t \le T\} \in \mathcal{F}_t$ and $\{t \le T\} = \bigcap_{s<t} \{s \le T\}$; thus $\{T_J \le T\} \cap \{T_J < \infty\} \in \mathcal{F}^*_{T_J}$. Similarly,

$$\{T_J = T = \infty\} = \bigcap_{s>0} \{s < T\} \cap \{s < T_J \le \infty\} \in \mathcal{F}^*_{T_J}.$$

Proof of Lemma B.2

Let

$$Z_t := \frac{\int_{\mathbb{R}} h(w') f(Y_t) g(X_{t-}, w')\, \nu(dw')}{\int_{\mathbb{R}} f(Y_t) g(X_{t-}, w'')\, \nu(dw'')},$$

which is a left-continuous and adapted process. Thanks to Lemma B.1, Z_{T_J} is $\mathcal{F}^*_{T_J}$-measurable.

We note the following two identities:

$$h(W) = \int_{[0,t]\times\mathbb{R}^d\times(\mathbb{R}_+)^2} h(w)\, \mathbf{1}_{\{t=T_J\}}\, M(dt,\ dw,\ du_f,\ du_g),$$

$$\frac{\int_{\mathbb{R}} h(w) f(Y_{T_J}) g(X_{T_J-}, w)\, \nu(dw)}{\int_{\mathbb{R}} f(Y_{T_J}) g(X_{T_J-}, w')\, \nu(dw')} = \int_{[0,t]\times\mathbb{R}^d\times(\mathbb{R}_+)^2} \frac{\int_{\mathbb{R}} h(w') f(Y_t) g(X_{t-}, w')\, \nu(dw')}{\int_{\mathbb{R}} f(Y_t) g(X_{t-}, w'')\, \nu(dw'')}\, \mathbf{1}_{\{t=T_J\}}\, M(dt,\ dw,\ du_f,\ du_g).$$

Then we exploit the Palm formula to prove that their product with any $Z_s\, \mathbf{1}_{\{s<T_J\le r\}}$ has the same average for any $s < r$ and Z_s any $\mathcal{F}_s$-measurable random variable:

$$\begin{aligned}
&\mathrm{E}_{(x,y)}\left[\, h(W)\, Z_s;\, s < T_J \le r\,\right] \\
&= \mathrm{E}_{(x,y)}\left[Z_s \int_{[0,t]\times\mathbb{R}^d\times(\mathbb{R}_+)^2} h(w)\, \mathbf{1}_{\{t=T_J\}}\, M(dt,\ dw,\ du_f,\ du_g);\, s < T_J \le r\right] \\
&= \mathrm{E}_{(x,y)}\left[\int_{[0,t]\times\mathbb{R}^d\times(\mathbb{R}_+)^2} Z_s\, h(w)\, \mathbf{1}_{(s,\ r]}(t)\, \mathbf{1}_{\{t=T_J\}}\, M(dt,\ dw,\ du_f,\ du_g)\right] \\
&= \mathrm{E}_{(x,y)}\left[\int_{[0,t]\times\mathbb{R}^d\times(\mathbb{R}_+)^2} \mathbf{1}_{(s,\ r]}(t)\, Z_s\, h(w)\, \mathbf{1}_{\{t=\widehat{T}_J\}}\, dt\, \nu(dw)\, du_f\, du_g\right],
\end{aligned}$$

where, according to the Palm formula, $\widehat{T}_J$ is the first jump of the process $(\widehat{X},\ \widehat{Y})$ encoded by $M + \delta_{(t,w,u_f,u_g)}$ and B (cf. e.g. [15, Proposition 13.1.VII]). Since $(\widehat{X},\ \widehat{Y})$ coincides with $(X,\ Y)$ at least up to time $t > s$, Z_s is not affected by this change. Moreover,

$$\left\{t = \widehat{T}_J\right\} = \{t \le T_J\} \cap \left\{u_f \le f(Y_t)\right\} \cap \left\{u_g \le g(X_{t-}, w)\right\}.$$

https://doi.org/10.1017/apr.2023.28 Published online by Cambridge University Press

Thus,

$$\begin{aligned}
&E_{(x,y)} \left[h(W)\, Z_s; s < T_J \le r \right] \\
&= E_{(x,y)}\Big[\int_{[0,t]\times\mathbb{R}^d\times(\mathbb{R}_+)^2} \mathbf{1}_{(s,\, r]}(t)\, h(w)\, Z_s\, \mathbf{1}_{\{u_f \le f(Y_t)\}} \\
&\qquad\qquad \times \mathbf{1}_{\{u_g \le g(X_{t-}, w)\}}\, \mathbf{1}_{\{t \le T_J\}}\, dt\, \nu(dw)\, du_f\, du_g \Big], \\
&= E_{(x,y)} \left[Z_s \int_s^r \int_{\mathbb{R}} \mathbf{1}_{\{t\le T_J\}}\, h(w) f(Y_t)\, g(X_{t-}, w)\, \nu(dw)\, dt \right].
\end{aligned}$$

On the other hand, and in the same spirit, we have

$$\begin{aligned}
&E_{(x,y)} \left[\frac{\int_{\mathbb{R}} h(w') f(Y_{T_J}) g(X_{T_J-}, w')\, \nu(dw')}{\int_{\mathbb{R}} f(Y_{T_J}) g(X_{T_J-}, w'')\, \nu(dw'')} \cdot Z_s; s < T_J \le r \right] \\
&= E_{(x,y)}\Big[Z_s \int_{[0,t]\times\mathbb{R}^d\times(\mathbb{R}_+)^2} \frac{\int_{\mathbb{R}} h(w') f(Y_t) g(X_{t-}, w')\, \nu(dw')}{\int_{\mathbb{R}} f(Y_t) g(X_{t-}, w'')\, \nu(dw'')} \\
&\qquad\qquad \times \mathbf{1}_{\{t=T_J\}}\, M(dt,\ dw,\ du_f,\ du_g); s < T_J \le r \Big] \\
&= E_{(x,y)}\Big[\int_{[0,t]\times\mathbb{R}^d\times(\mathbb{R}_+)^2} Z_s\, \mathbf{1}_{(s,r]}(t)\, \frac{\int_{\mathbb{R}} h(w') f(Y_t) g(X_{t-}, w')\, \nu(dw')}{\int_{\mathbb{R}} f(Y_t) g(X_{t-}, w'')\, \nu(dw'')} \\
&\qquad\qquad \times \mathbf{1}_{\{t\le T_J\}}\, \mathbf{1}_{\{u_f \le f(Y_t)\}}\, \mathbf{1}_{\{u_g \le g(X_{t-}, w)\}}\, dt\, \nu(dw)\, du_f\, du_g \Big] \\
&= E_{(x,y)} \left[Z_s \int_s^r \int_{\mathbb{R}} \mathbf{1}_{\{t\le T_J\}}\, h(w') f(Y_t)\, g(X_{t-}, w')\, \nu(dw')\, dt \right],
\end{aligned}$$

which is indeed the same integral as for $h(W)$. □

Appendix C. Brief overview of characteristic profiles of the quasi-stationary regime obtained by simulations

We provide in this appendix some results from a particular choice of three parameter regimes, whose comparison sheds light on the discussion given in Subsection 2.3. We present the profiles of the characteristic distributions and functions of the quasi-stationary regime, namely the QSD, the quasi-ergodic distribution (QED), and the survival capacity (the limiting properties are recalled just beside the figures).

The details of the parameters used are as follows. For population size dynamics, the growth rate as a function of x is here chosen to be of the form $r(x) = 4 - 30 \cdot |x|$. A parabolic profile would give very similar results. The competition rate is $c = 0.1$, which leads to population sizes at quasi-equilibrium (carrying capacity) close to 40 (in arbitrary units). The values for the the diffusion coefficient σ and the speed of the environment v are respectively 2 and 6. Thus, there are rapid fluctuations in population size in the time-scale where adaptation changes.

The profile of additive effects of mutations is given by $\nu(dw) = \frac{1}{2w_0} \exp(-|w|/w_0)$. It is therefore symmetric exponential, with $w_0 = 0.03$, so with many small mutations. The effect of population size on the fixation rate is simply proportional: $f_N(n) = m \cdot n$. The mutation rate m is the only parameter modified here across the three simulation sets: it takes the values $m = 0.85$, $m = 0.55$, and $m = 0.25$. These values are chosen so that the adaptation is critical at $m = 0.55$: for larger m, like $m = 0.85$, extinction is kept almost negligible, so that we say that adaptation is spontaneous, whereas for smaller values of m, extinction plays a consequent role and produces differences in shape between the QSD and the QED.

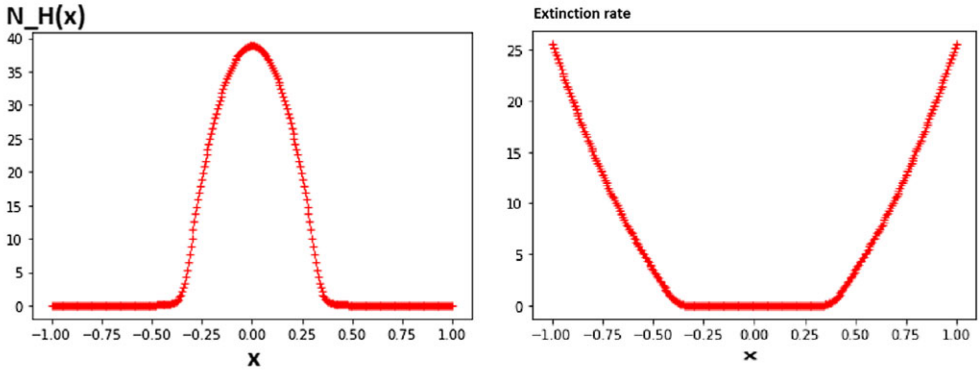


FIGURE 2. Left: $N_H(x)$, the harmonic means of the population size fluctuations of the process $(\widetilde{N}_t^x)_{t\geq 0}$ with fixed traits x (given by the associated QSD). Right: the extinction rate of the QSD.

We exploit the following expression for the probability of invasion:

$$g(x, w) := \frac{N_H(x) \cdot \Delta r/\sigma}{1 - \exp\left[-N_H(x) \cdot \Delta r/\sigma\right]},$$

where $\Delta r := r(x+w) - r(x)$ is the variation of the growth rate between the mutant and the resident, and $N_H(x)$ is the harmonic mean of a resident population with fixed trait x (averaged against its associated QSD). Deleterious mutations are allowed, but their probability of fixation is greatly reduced if they are strong in relation to population fluctuations. The values of N_H are estimated numerically, with the profile shown in Figure 2.

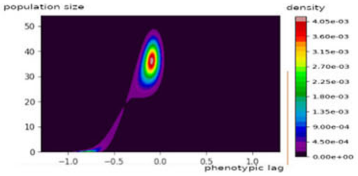
The formula relies on the Kimura diffusion approximation that has been derived in the case of fixed population size. Assuming rapid size fluctuations, we choose the harmonic mean as the reference by referring to classical approximations obtained in the case of periodically fluctuating population sizes (cf. notably [26]). More details are given (in French) in the author's doctoral thesis [30], and a subsequent paper is planned to discuss these results and the relevance of this estimation. The comparison of such a two-component stochastic model to the individual-based model through the QSD and QED will be a good test for the relevance of such a formula. The kind of dependence in the difference in growth rate seems to play a crucial role for having a QED so conserved.

These simulations were obtained by calculating the evolution of the densities themselves. This method is related to the method of finite volumes, with an explicit numerical scheme and a renormalization of density estimates at each time step. The transitions to X and N are performed successively to reduce the calculation time.

Acknowledgements

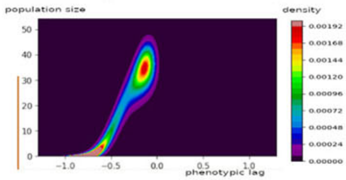
I wish mainly to thank my supervisor, Etienne Pardoux, for his great support throughout the preparation of this article. I would also like to mention the very inspiring meetings and discussions brought about by the Consortium 'Modélisation Mathématique et Biodiversité' of Veolia–École Polytechnique–Muséum National d'Histoire Naturelle. I also address my thanks to my fellow PhD students in Marseilles, with whom I shared my inspirations, and more generally to the people of the CMI with whom I interacted as I prepared this paper.

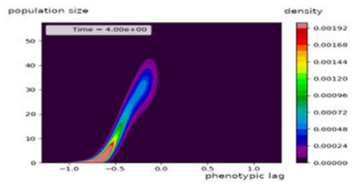
Profiles of the "QSD"



$$\mathrm{P}_{(x,\,n)}\left[(X,N)_t \in (dx,dn) \,\middle|\, t < \tau_\partial\right] \underset{t\to\infty}{\longrightarrow} \alpha(dx,dn)$$

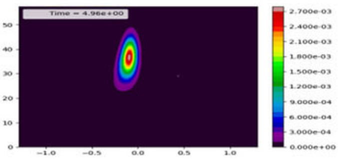
Case of a spontaneous adaptation





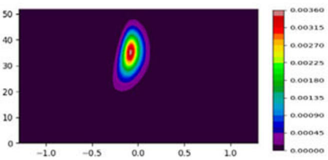
Critical regime of adaptation Selection through the extinction of populations

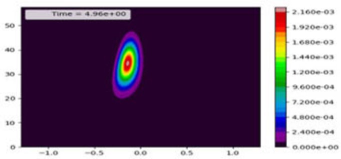
Profiles of the QED: the invariant measure of the Q-process



$$\lim_{t\to\infty} \lim_{T\to\infty} \mathrm{P}_{(x,\,n)}\left[(X,N)_t \in (dx,dn) \,\middle|\, T < \tau_\partial\right] = \beta(dx,dn)$$

Case of a spontaneous adaptation





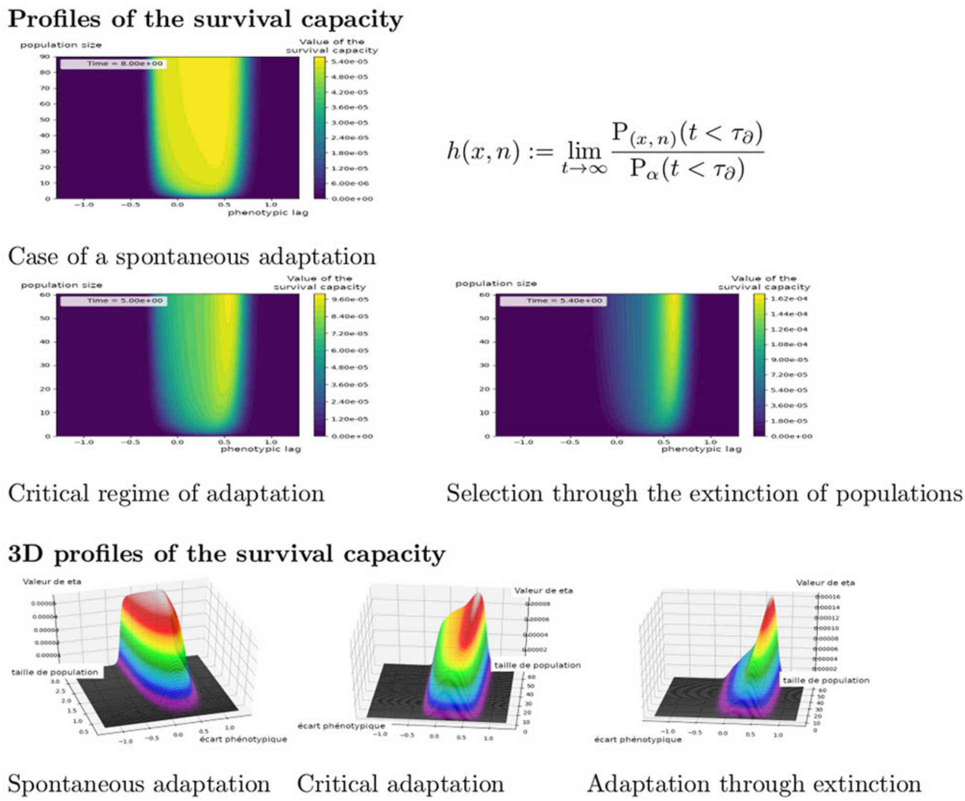
Critical regime of adaptation Selection through the extinction of populations

---

Given the very different profiles obtained for the QSD, it is quite remarkable that the quasi-ergodic measures are similar as much. In particular, we can see that the histories of the surviving populations are still shaped by the maintenance of these populations at large sizes with almost optimal traits, even when such traits are very rare according to the QSD.

---

https://doi.org/10.1017/apr.2023.28 Published online by Cambridge University Press



Funding information

There are no funding bodies to thank in relation to the creation of this article.

Competing interests

There were no competing interests to declare which arose during the preparation or publication process of this article.

References

[1] BASS, R. (1995). *Probabilistic Techniques in Analysis*. Springer, New York.

[2] BANSAYE, V., CLOEZ, B., GABRIEL, P. AND MARGUET, A. (2022). A non-conservative Harris' ergodic theorem. *J. London Math. Soc.* **106**, 2459–2510.

[3] BÜRGER, R. AND LYNCH, M.(1995). Evolution and extinction in a changing environment: a quantitative-genetic analysis. *Evolution* **49**, 151–163.

[4] BANSAYE, V. AND MÉLÉARD, S. (2015). *Stochastic Models for Structured Populations: Scaling Limits and Long Time Behavior*. Springer, Cham.

[5] CATTIAUX, P. *et al.* (2009). Quasi-stationary distributions and diffusion models in population dynamics. *Ann. Prob.* **37**, 1926–1969.

[6] CHAZOTTES, R., COLLET, P. AND MÉLÉARD, S. (2019). On time scales and quasi-stationary distributions for multitype birth-and-death processes. *Ann. Inst. H. Poincaré Prob. Statist.* **55**, 2249–2294.

[7] Champagnat, N., Coulibaly-Pasquier, K. and Villemonais, D. (2018). Criteria for exponential convergence to quasi-stationary distributions and applications to multi-dimensional diffusions. In *Séminaire de Probabilités XLIX*, eds C. Donati-Martin, A. Lejay and A. Rouault, Springer, Cham, pp. 165–182.
[8] Cloez, B. and Gabriel, P. (2020). On an irreducibility type condition for the ergodicity of nonconservative semigroups. *C. R. Math.* **358**, 733–742.
[9] Collet, P., Martnez, S. and San Martn, J. (2013). *Quasi-stationary Distributions*. Springer, Berlin, Heidelberg.
[10] Champagnat, N. and Villemonais, D. (2016). Exponential convergence to quasi-stationary distribution and Q-process. *Prob. Theory Relat. Fields* **164**, 243–283.
[11] Champagnat, N. and Villemonais, D. (2018). Uniform convergence of time-inhomogeneous penalized Markov processes. *ESAIM Prob. Statist.* **22**, 129–162.
[12] Champagnat, N. and Villemonais, D. (2020). Practical criteria for R-positive recurrence of unbounded semigroups. *Electron. Commun. Prob.* **25**, article no. 6.
[13] Champagnat, N. and Villemonais, D. (2021). Lyapunov criteria for uniform convergence of conditional distributions of absorbed Markov processes. *Stoch. Process. Appl.* **135**, 51–74.
[14] Champagnat, N. and Villemonais, D. (2023). General criteria for the study of quasi-stationarity. To appear in *Electron. J. Prob.* Preprint available at https://arxiv.org/abs/1712.08092.
[15] Daley, D. J. and Vere-Jones, D. (2008). *An Introduction to the Theory of Point Processes: Volume II: General Theory and Structure*. In *Prob. and Its Appl.*, Springer, 2nd ed., New York.
[16] Di Tella, P. (2013). On the predictable representation property of martingales associated with Lévy processes. *Stochastics* **87**, 170–184.
[17] Evans, L. C. (1998). *Partial Differential Equations*. American Mathematical Society, Providence, RI.
[18] Ferré, G., Rousset, M. and Stoltz, G. (2020). More on the long time stability of Feynman–Kac semigroups. *Stoch. Partial Diff. Equat. Anal. Comput.* **9**, 630–673.
[19] Guillin, A., Nectoux, B. and Wu, L. (2020). *Quasi-stationary distribution for strongly Feller Markov processes by Lyapunov functions and applications to hypoelliptic Hamiltonian systems*. Preprint. Available at https://hal.science/hal-03068461.
[20] Kopp, M. and Hermisson, J. (2009). The genetic basis of phenotypic adaptation II: the distribution of adaptive substitutions in the moving optimum model. *Genetics* **183**, 1453–1476.
[21] Kopp, M., Nassar, E. and Pardoux, E. (2018). Phenotypic lag and population extinction in the moving-optimum model: insights from a small-jumps limit. *J. Math. Biol.* **7**, 1431–1458.
[22] Lambert, A. (2005). The branching process with logistic growth. *Ann. Appl. Prob.* **15**, 1506–1535.
[23] Méléard, S. and Villemonais, D. (2012). Quasi-stationary distributions and population processes. *Prob. Surveys* **9**, 340–410.
[24] Nassar, E. and Pardoux, E. (2017). On the large-time behaviour of the solution of a stochastic differential equation driven by a Poisson point process. *Adv. Appl. Prob.* **49**, 344–367.
[25] Nassar, E. and Pardoux, E. (2019). Small jumps asymptotic of the moving optimum Poissonian SDE. *Stoch. Process. Appl.* **129**, 2320–2340.
[26] Otto, S. P. and Whitlock, M. C. (1997). The probability of fixation in populations of changing size. *Genetics* **146**, 723–733.
[27] Pardoux, E. (2016). *Probabilistic Models of Population Evolution: Scaling Limits, Genealogies and Interactions*. Springer, Cham.
[28] Pollett, P. K. (2015). Quasi-stationary distributions: a bibliography. Tech. Rep., University of Queensland. Available at https://people.smp.uq.edu.au/PhilipPollett/papers/qsds/qsds.html.
[29] Van Doorn, E. A. and Pollett, P. K. (2013). Quasi-stationary distributions for discrete-state models. *Europ. J. Operat. Res.* **230**, 1–14.
[30] Velleret, A. (2020). Mesures quasi-stationnaires et applications à la modélisation de l'évolution biologique. Doctoral Thesis, Aix-Marseille Université. Available at https://www.theses.fr/2020AIXM0226.
[31] Velleret, A. (2022). Unique quasi-stationary distribution, with a possibly stabilizing extinction. *Stoch. Process. Appl.* **148**, 98–138.
[32] Velleret, A. (2023). *Exponential quasi-ergodicity for processes with discontinuous trajectories*. Preprint. Available at https://arxiv.org/abs/1902.01441.
[33] Yamada, T. and Watanabe, S. (1971). On the uniqueness of solutions of stochastic differential equations. *J. Math. Kyoto Univ.* **11**, 155–167.