



**HAL**  
open science

# Sharp approximation and hitting times for stochastic invasion processes

Vincent Bansaye, Xavier Erny, Sylvie Méléard

► **To cite this version:**

Vincent Bansaye, Xavier Erny, Sylvie Méléard. Sharp approximation and hitting times for stochastic invasion processes. 2023. hal-03915479v3

**HAL Id: hal-03915479**

**<https://hal.science/hal-03915479v3>**

Preprint submitted on 11 Dec 2023 (v3), last revised 13 Jun 2024 (v4)

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

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



Distributed under a Creative Commons Attribution 4.0 International License

# Sharp approximation and hitting times for stochastic invasion processes

Vincent Bansaye\*, Xavier Erny\* and Sylvie Méléard\*

\* CMAP, CNRS, École polytechnique, Institut Polytechnique de Paris, 91120 Palaiseau, France

*This article is dedicated to Francis Comets, whose mathematical curiosity and enthusiasm have always been a source of inspiration.*

**Abstract:** We are interested in the invasion phase for stochastic processes with interactions. A single mutant with positive fitness arrives in a large resident population at equilibrium. By a now classical approach, the first stage of the invasion is well approximated by a branching process. The macroscopic phase, when the mutant population is of the same order as the resident population, is described by the limiting dynamical system. We capture the intermediate mesoscopic phase for the invasive population and obtain sharp approximations. It allows us to describe the fluctuations of the hitting times of thresholds, which inherit a large variance from the first stage. We apply our results to two models which are original motivations. In particular, we quantify the hitting times of critical values in cancer emergence and epidemics.

**MSC2020 subject classifications:** 37N25, 60F15, 60G55, 60H10.

**Keywords and phrases:** Stochastic invasion process, Branching process, Dynamical system, Stochastic processes approximation, Hitting times approximation.

## 1. Introduction and main results

We aim to finely quantify a mutant invasion in a resident population at equilibrium. Such a situation is standard in eco-evolution, in cancer emergence or in epidemiology when a single individual particularly well adapted can develop its own subpopulation and invade the global population. After a certain amount of time, the number of mutants becomes non negligible with respect to the resident population size, allowing to summarize the dynamics of the stochastic population processes by their deterministic approximations. In Champagnat [15], a systematic approach was introduced to quantify the mutant invasion success, based on the properties of the mutant birth and death process and its coupling with branching processes for which the survival probability was easily computed. That allowed to characterize the probability for the mutant process to attain a certain fixed threshold. This approach has been used by many authors, see e.g. [16, 8, 2, 9, 13].

To go in details in this invasion process, we introduce some scaling parameter  $K$ , whose values are typically large, characterizing the macroscopic population sizes and allowing to quantify each invasion step. We consider a bi-type birth and death process modeling the population size dynamics of interacting resident and mutant individuals. We denote by  $N_R^K(t)$  (resp.  $N_M^K(t)$ ) the number of resident individuals (resp. mutant individuals) at time  $t$  and the density of populations is defined by

$$X^K(t) = (X_R^K(t), X_M^K(t)) := (N_R^K(t)/K, N_M^K(t)/K).$$

The interaction is modeled through density dependence of individual birth and death rates, respectively  $b_\bullet(X^K)$ ,  $d_\bullet(X^K)$  for  $\bullet \in \{R, M\}$ . This dependence may model for instance competition for resources, which can make the death rates increase or the birth rates decrease.

The processes  $(N_R^K, N_M^K)$  are defined on a probability space  $(\Omega, \mathcal{F}, \mathbb{P})$ , as strong solutions (until potential explosion time) of the stochastic differential systems driven by Poisson point measures

$\mathcal{N}_R^b, \mathcal{N}_R^d, \mathcal{N}_M^b, \mathcal{N}_M^d$  with Lebesgue measure intensity on  $\mathbb{R}_+ \times \mathbb{R}_+$ :

$$\begin{aligned} N_{\bullet}^K(t) &= N_{\bullet}^K(0) + \int_{[0,t] \times \mathbb{R}_+} \mathbf{1}_{\{u \leq N_{\bullet}^K(s-) b_{\bullet}(X^K(s-))\}} \mathcal{N}_{\bullet}^b(ds, du) \\ &\quad - \int_{[0,t] \times \mathbb{R}_+} \mathbf{1}_{\{u \leq N_{\bullet}^K(s-) d_{\bullet}(X^K(s-))\}} \mathcal{N}_{\bullet}^d(ds, du), \end{aligned} \quad (1.1)$$

for  $\bullet \in \{R, M\}$ ,  $K > 0$  and  $t \geq 0$ . We assume that  $\mathcal{N}_R^b$  and  $\mathcal{N}_R^d$  are independent and that  $\mathcal{N}_M^b$  and  $\mathcal{N}_M^d$  are independent too. We also assume that the Poisson point measures  $(\mathcal{N}_R^b, \mathcal{N}_R^d, \mathcal{N}_M^b, \mathcal{N}_M^d)$  are independent of the initial condition  $N^K(0)$ . But  $(\mathcal{N}_R^b, \mathcal{N}_R^d)$  and  $(\mathcal{N}_M^b, \mathcal{N}_M^d)$  are not necessarily independent. We refer to Section 5.1 for an application in evolution where all these Poisson point processes are independent and to Section 5.2 for an example in epidemiology where they are dependent. The individual growth rates are defined on  $\mathbb{R}_+ \times \mathbb{R}_+$  by

$$F_R(x_R, x_M) = b_R(x_R, x_M) - d_R(x_R, x_M); \quad F_M(x_R, x_M) = b_M(x_R, x_M) - d_M(x_R, x_M).$$

It is well known (cf [27, 28, 22]) that under suitable assumptions on the parameters and when the initial conditions of the process are proportional to  $K$ , the stochastic process  $((X_R^K(t), X_M^K(t)) : t \geq 0)$  converges in probability, and even almost surely and in  $L^2$  (when  $K$  tends to infinity), on any finite time interval. The limit is the solution  $(x_R, x_M)$  of the dynamical system

$$\begin{cases} x'_R &= x_R F_R(x_R, x_M) \\ x'_M &= x_M F_M(x_R, x_M). \end{cases} \quad (1.2)$$

We assume the existence of some (non-necessary unique nor stable) equilibrium  $x_R^* > 0$  satisfying

$$F_R(x_R^*, 0) = b_R(x_R^*, 0) - d_R(x_R^*, 0) = 0.$$

In this work, we start from one mutant inside a resident population close to this equilibrium :

$$N^K(0) = (N_R^K(0), 1), \quad N_R^K(0) \sim x_R^* K. \quad (1.3)$$

Thus  $X^K(0)$  is close to  $(x_R^*, 0)$  and the mutant process  $(N_M^K(t), t \geq 0)$  will firstly be close to the branching process  $Z$  with individual birth and death rates respectively

$$b_{\star} = b_M(x_R^*, 0); \quad d_{\star} = d_M(x_R^*, 0).$$

We assume that the initial growth rate of this process is positive

$$r_{\star} = b_{\star} - d_{\star} = F_M(x_R^*, 0) > 0. \quad (1.4)$$

This ensures that the mutant population process attains values of order  $K$  with positive probability, at a time scale  $\log K$ .

Our ambition is to finely capture the dynamics of the invasion process until macroscopic levels, and to focus in particular on the intermediate scale, where the mutant process is large but negligible compared to  $K$ . We consider the hitting times of successive levels  $n \in \mathbb{N}^*$  of the mutant population process when the initial condition is  $N^K(0) = n_0 \in \mathbb{N}^2$ :

$$T_M^K(n_0, n) = \inf\{t \geq 0 : N_M^K(t) \geq n\} = \inf\{t \geq 0 : X_M^K(t) \geq n/K\}. \quad (1.5)$$

Our first result proves that the mutant process is equivalent to the branching process  $Z$  before reaching macroscopic level of order  $K$ , in the following sense, see Section 2. For any  $\xi_K \ll K/(\log K)^a$

and  $\eta > 0$ , we show that

$$\begin{aligned} & \lim_{K \rightarrow +\infty} \mathbb{P} \left( \sup_{t \leq T_M^K(N^K(0), \xi_K)} \left| \frac{N_M^K(t)}{Z(t)} - 1 \right| > \eta; W > 0 \right) \\ &= \lim_{T \rightarrow \infty} \limsup_{K \rightarrow \infty} \mathbb{P} \left( \sup_{T \leq t \leq T_M^K(N^K(0), \xi_K)} \left| \frac{N_M^K(t)}{W e^{r_* t}} - 1 \right| > \eta; W > 0 \right) = 0, \end{aligned}$$

where

$$W = \lim_{t \rightarrow \infty} Z(t) e^{-r_* t} \in [0, \infty) \quad a.s.$$

is an exponential variable with an additional atom at 0 corresponding to the extinction event (cf. Durrett [20]). The meaning of the above convergence is that, as long as the mutant population has not reached the threshold  $\xi_K$ , the size of the population  $N_M^K(t)$  is equivalent in probability to the branching process  $Z(t)$  as  $K$  goes to infinity. The value of  $a$  is 1 or 2 depending on the exponential stability of the equilibrium of the resident population alone, see forthcoming Theorem 2.2 for a precise statement. These two cases correspond to a hyperbolic or a partially hyperbolic equilibrium.

For the proof of the theorem, we use a coupling of the mutant process and the branching process and we need to control the distance between them. We use that individuals belonging to one of the two populations but not to the other in the coupling can be seen as a branching structure with an additional immigration whose intensity increases with the value of  $Z$  or  $N_M^K$ .

This approach is inspired by and complements previous works. Exact couplings have been proved both in the context of invasion of mutants in evolutionary biology and in epidemiology, where the models lead to a resident population with a hyperbolic or a partially hyperbolic equilibrium. For the SIR model in continuous time, we refer in particular to Ball and Donnelly [4] for a proof of exact coupling before the population reaches the order of magnitude  $\sqrt{K}$ . Barbour and al [7] prove that exact coupling holds before the mutant population reaches order  $K^{7/12}$  for multidimensional models in evolutionary biology. Complementary results have been obtained for epidemiological models in [5, 6]. Our result allows to compare the original mutant process and the branching process until the mutant population reaches the order  $K/(\log K)^a$ .

When the population of mutants becomes large but is not yet macroscopic, we prove (cf Section 3.2) that the deterministic approximation by the dynamical system is already valid in the following sense. Starting with a number of mutants  $1 \ll N_M^K(0) = K X_M^K(0) \ll K$ ,

$$\sup_{t \leq T_M^K(N^K(0), vK)} \left| \frac{X_M^K(t)}{x_M^K(t)} - 1 \right| \xrightarrow{K \rightarrow \infty} 0, \text{ in probability,}$$

where  $v > 0$ , and  $x_M^K$  is the solution of (1.2) with the initial condition  $1/K \ll x_M^K(0) = X_M^K(0) \ll 1$  and  $x_R^K(0)$  close to  $x_R^*$ . We refer to Theorem 3.6 for a precise statement which gives also the conditions on the initial resident population size and the admissible values of  $v$  and quantitative estimates for the convergence. The boundary of  $(\mathbb{R}_+)^2$  is invariant for the dynamical system but this domain is unstable in the neighborhood of  $\mathbb{R}_+ \times \{0\}$  in the invasion regime. This makes the problem delicate when controlling the stochastic flow. In the time window  $[0, T_M^K(N^K(0), vK)]$  whose size is of order of magnitude  $\log K$ ,  $X^K$  goes from  $X_M^K(0) \ll 1$  to  $v$ . This requires to deal with a long time scale and escape an unstable equilibrium, as in [7]. We consider in this paper different scales and techniques, which allow us to capture the expected renormalization  $X_M^K(t)/x_M^K(t)$  from low density until any reachable macroscopic value.

The simulations of Figure 1 illustrate the different approximations corresponding to the SIR example of Section 5.2, both by the branching process (from the initial time) and by the dynamical system (from hitting times of large levels).

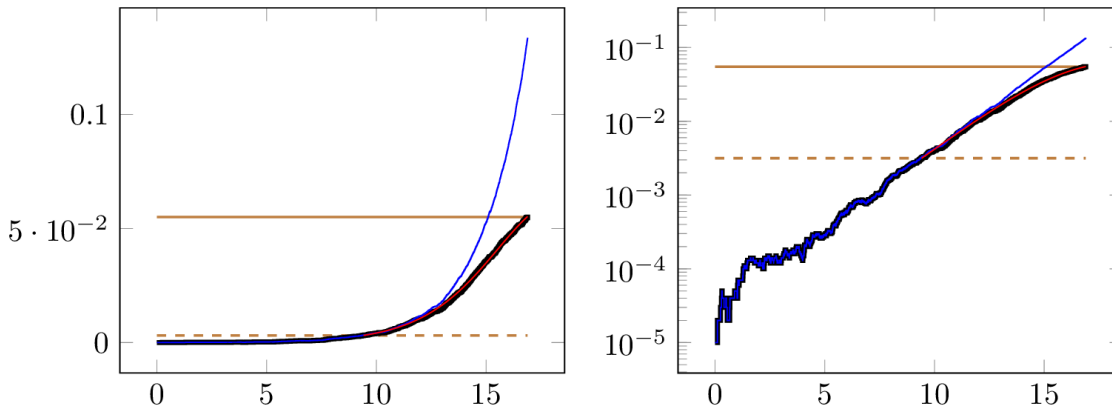


FIG 1. The black curve represents a trajectory of the process  $X_M^K(t) = N_M^K(t)/K$ , the blue curve a trajectory of  $Z(t)/K$ , and the red curve is the deterministic function  $x_M^K(t)$  starting at time  $T_M^K(N^K(0), \xi^K)$ . Both figures represent the simulations of the same trajectory in different scales (standard scale on the left, logarithmic scale on the right). These simulations have been done with the following parameters:  $K = 10^5$ ,  $F_R(x_R, x_M) = -1.5x_M$ ,  $F_M(x_R, x_M) = 1.5x_R - 1$ ,  $x_R^* = 1$ ,  $\xi_K = \sqrt{K}$  (cf the brown dashed line) and  $v = 0.055$  (cf the brown plain line).

As a byproduct, we obtain in Theorem 4.1 the following convergence in law of hitting times when  $K$  tends to infinity, on the survival event of  $N_M^K$ . Under Condition (1.3), for  $(\zeta_K)_K$  going to infinity such that  $\zeta_K/K \rightarrow v \geq 0$ , we show the following convergence in law

$$T_M^K(N^K(0), \zeta_K) - \frac{\log(\zeta_K)}{r_*} \Longrightarrow \tau(v) - \frac{\log W^*}{r_*},$$

where the function  $\tau$  is continuous,  $\tau(0) = 0$  and  $W^*$  is an exponential random variable with parameter  $r_*/b_*$ . We refer to Section 3.1 for the definition of  $\tau$  and the characterization of admissible values  $[0, v_*)$  of the level  $v$ . This result is illustrated in Figure 2 corresponding to the SIR example of Section 5.2. Thus, the hitting times of mesoscopic or macroscopic levels have Gumbel laws translated by  $(\log \zeta_K)/r_* + \tau(v)$ . We refer to [6, 7] for related results on the random time needed to see the deterministic macroscopic curve.

In the situations when  $K$  is very large, as for huge cohorts of microorganisms (bacteria, cells),  $(\log \zeta_K)/r_*$  gives a good approximation of the hitting times. But in many other cases,  $K$  is not so large. We refer to the modeling of hematopoiesis and leukemias [11, 12] where  $K$  is the number of sain stem cells, of order  $10^5$ . This is the original motivation for this work and one reason to quantify explicitly the convergence results. The intermediate phase and hitting times are indeed involved in the emergence and detection of cancer.

**Assumptions.** We consider the following assumptions on the birth and death rates. They ensure in particular that the resident population alone has a positive equilibrium point (which may be non stable) and that the mutant population can invade.

**Assumption 1.1.**

- (R) *Regularity:* the functions  $b_R, d_R, b_M, d_M$  are  $C^2(\mathbb{R}_+ \times \mathbb{R}_+, \mathbb{R}_+)$ .
- (E) *Equilibrium:* there exists  $x_R^* > 0$  such that  $F_R(x_R^*, 0) = 0$  and  $\partial_R F_R(x_R^*, 0) \leq 0$ .
- (I) *Invasion:*  $F_M(x_R^*, 0) > 0$ .

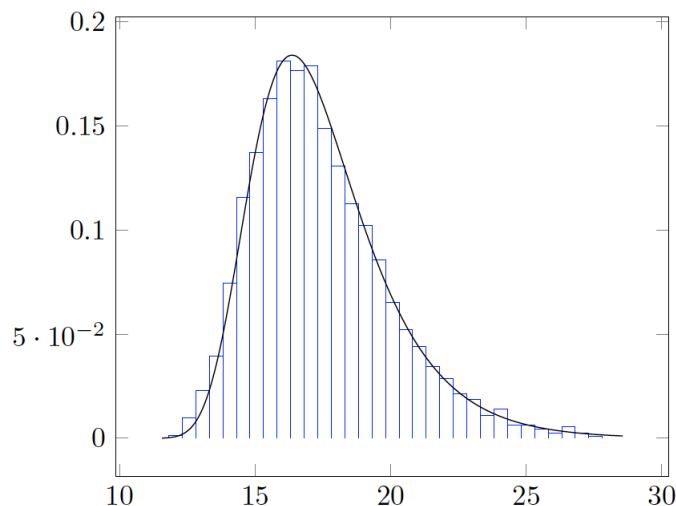


FIG 2. The histogram represents the values of  $T_M^K(N^K(0), vK)$  (conditionally to the non-extinction of  $N_M^K$ ). The simulations have been done with the following parameters:  $K = 100000$ ,  $F_R(x_R, x_M) = -1.5x_M$ ,  $F_M(x_R, x_M) = 1.5x_R - 1$ ,  $x_R^* = 1$  and  $v = 0.055$ . The size of the sample to draw the histogram is 10000. The curve is the density  $f$  of the random variable  $\log(vK)/r_* + \tau(v) - \log(W^*)/r_*$ , where the density of  $-\log(W^*)/r_*$  is equal to  $r_*^2/b_* \exp(-r_*t) \exp(-r_*/b_* e^{-r_*t})$ .

Under Assumption (R), the bi-type birth and death process  $X^K$  is well defined until its explosion time, which may be finite with positive probability. In our framework, the process is studied before hitting times of fixed levels, clearly smaller than the explosion time. In the eco-evolutionary framework,  $F_M(x_R^*, 0)$  is called invasion fitness. It quantifies the individual growth rate for a small mutant population in a resident population at the equilibrium.

#### Notation.

- We write  $x^* = (x_R^*, 0)$ .
- For a smooth function  $f : (x_R, x_M) \in \mathbb{R}^2 \mapsto f(x_R, x_M)$ , we denote by  $\partial_R f$  (resp.  $\partial_M f$ ) the derivative w.r.t. the first (resp. second) variable.
- If  $\mathcal{N}(dx)$  is a Poisson point measure with intensity  $\lambda(dx)$ , we denote by  $\tilde{\mathcal{N}}(dx) := \mathcal{N}(dx) - \lambda(dx)$  the compensated measure.
- Throughout the paper,  $C$  denotes any positive constant depending only on the model parameters. The value of  $C$  can change from line to line. If a constant depends on some (non-model) parameter  $\theta$  we write  $C_\theta$ .

## 2. Approximation by branching processes

In this section, we assume that at the initial time, the resident population is close to its equilibrium and a single mutant individual appears.

**Assumption 2.1.** *There exists  $C > 0$  such that for any  $K > 0$ ,*

$$\left| \frac{N_R^K(0)}{K} - x_R^* \right| \leq \frac{C}{K} \quad \text{and} \quad N_M^K(0) = 1 \quad \text{a.s.}$$

We introduce the branching process  $(Z(t) : t \geq 0)$  with individual birth rate  $b_\star = b_M(x^\star)$  and (individual) death rate  $d_\star = d_M(x^\star)$ , defined on the same probability space as  $N_M^K$  and coupled with  $N_M^K$  in the following way:

$$Z(t) = 1 + \int_{[0,t] \times \mathbb{R}_+} \mathbf{1}_{\{u \leq b_\star Z(s-)\}} \mathcal{N}_M^b(ds, du) - \int_{[0,t] \times \mathbb{R}_+} \mathbf{1}_{\{u \leq d_\star Z(s-)\}} \mathcal{N}_M^d(ds, du),$$

where the Poisson point measures  $\mathcal{N}_M^b$  and  $\mathcal{N}_M^d$  have been introduced in (1.1).

We recall from Assumption 1.1.(R) that  $r_\star = b_\star - d_\star > 0$  and the branching process  $Z$  is supercritical. The survival event has positive probability and as  $t$  tends to infinity, the martingale  $W(t) = Z(t) \exp(-r_\star t)$  converges a.s. to a finite random variable  $W$  which is positive on the survival event :

$$\{W > 0\} = \{\forall t > 0 : Z(t) > 0\}.$$

The random variable  $W$  is exponentially distributed with an additional atom at 0 :

$$W \stackrel{d}{=} \frac{d_\star}{b_\star} \delta_0 + \frac{r_\star}{b_\star} \mu_{r_\star/b_\star},$$

where  $\delta_0$  is a Dirac mass at 0 and  $\mu_a(dx) = \mathbf{1}_{\mathbb{R}_+}(x) a e^{-ax} dx$  and  $dx$  is the Lebesgue measure. We refer to [1] Chapter 3 or [20]-Theorem 1 for such results.

Let us study the coupling between the processes  $Z$  and  $N_M^K$ . First, when the branching process  $Z$  becomes extinct, so does the mutant process  $N_M$  and we have the following result. Under Assumptions 1.1 and 2.1,

$$\lim_{K \rightarrow \infty} \mathbb{P}(\forall t \geq 0 : N_M^K(t) = Z(t) | W = 0) = 1.$$

Indeed, the two processes coincide by coupling on a time window whose size goes to infinity with probability 1 as  $K$  goes to infinity. Besides,  $Z$  is a.s. absorbed at 0 in finite time on the event  $\{W = 0\}$ . Such arguments are classical for epidemiological models and can be adapted to our context. We refer to [4].

Let us turn our attention to the survival event and state the main result of the section. It compares the mutant process and its branching approximation as long as the process is not too close to the macroscopic scale. We prove that the ratio of the two processes converges to 1 in probability, with an explicit speed, before time  $T_M^K(N^K(0), n)$  when the population size of mutants reaches a certain threshold  $n$  (conveniently chosen), see (1.5) for definition.

**Theorem 2.2.** *Under Assumptions 1.1 and 2.1,*

*i) if  $\partial_R F_R(x^\star) < 0$ , there exists  $C > 0$  such that for any  $\eta > 0$  small enough,  $K \geq 2$  and  $n \leq K/(\log K)$ , we have*

$$\mathbb{P}\left(\sup_{t \leq T_M^K(N^K(0), n)} \left| \frac{N_M^K(t)}{Z(t)} - 1 \right| > \eta ; W > 0\right) \leq C \left( \frac{\sqrt{K \log K}}{n} + \frac{1}{K^{1/3}} + \frac{1}{\eta^{1/4}} \left( \frac{n \log K}{K} \right)^{1/10} \right),$$

*ii) if  $\partial_R F_R(x^\star) = 0$ , there exist  $L, C > 0$  such that for any  $\eta > 0$  small enough,  $K \geq 2$  and  $n \leq LK/(\log K)^2$ , we have*

$$\mathbb{P}\left(\sup_{t \leq T_M^K(N^K(0), n)} \left| \frac{N_M^K(t)}{Z(t)} - 1 \right| > \eta ; W > 0\right) \leq C \left( \sqrt{\frac{\log K}{n}} + \frac{1}{K^{1/3}} + \frac{1}{\eta^{1/4}} \left( \log K \sqrt{\frac{n}{K}} \right)^{1/10} \right).$$

We observe that  $n$  needs also to be larger than  $\sqrt{K \log(K)}$  in (i) (and larger than  $\log(K)$  in (ii)) so that the right hand side tends to 0. The choice of  $n$  allows to optimize the bounds and cover the time when the number of mutants is below  $K/\log K$  in (i) (and  $LK/(\log K)^2$  in (ii)). Let us state the immediate corollary which will be used in Section 4.

**Corollary 2.3.** *If  $\partial_R F_R(x^*) < 0$  and  $(\xi_K)_K$  satisfies  $1 \ll \xi_K \ll K/(\log K)$ , then*

$$\lim_{K \rightarrow \infty} \mathbb{P} \left( \sup_{t \leq T_M^K(N^K(0), \xi_K)} \left| \frac{N_M^K(t)}{Z(t)} - 1 \right| > \eta ; W > 0 \right) = 0.$$

*If  $\partial_R F_R(x^*) = 0$ , a similar result holds for sequences  $(\xi_K)_K$  such that  $1 \ll \xi_K \ll K/(\log K)^2$ .*

We refer to Section 4 for the asymptotic behavior of  $T_M^K(N^K(0), \xi_K)$ .

The statement depends on the value of  $\partial_R F_R(x^*)$ . Indeed, the proof is based on a comparison between  $T_M^K(N^K(0), n)$  and the first time where the resident population exits a neighborhood of its equilibrium. If  $\partial_R F_R(x^*) < 0$ , the manifold  $\{x_M = 0\}$  is stable around the fixed point  $(x_R^*, 0)$ . This stability helps for the control of the resident population and guarantees that the distance of the resident population process to its equilibrium remains smaller than  $n$  (up to some well chosen factor) until time  $T_M^K(N^K(0), \xi_K)$ . If  $\partial_R F_R(x^*) = 0$ , exponential (local) stability of residents is lost and we can (only) prove that the distance is smaller than  $\sqrt{Kn}$ . We refer to Lemma 2.8 for details.

The proof of Theorem 2.2 will be obtained in several steps that we develop below. We first introduce a coupling to compare  $N_M^K$  and  $Z$ . Let us introduce a decomposition involving the processes realizing the coupling and the additional part. The two subpopulations will be indexed respectively by  $q$  and  $r$ . We introduce

$$m^K(s) = \min(N_{M,q}^K(s)b_M(X^K(s)), Z_q^K(s)b_*), \quad p^K(s) = \max(N_{M,q}^K(s)d_M(X^K(s)), Z_q^K(s)d_*),$$

where  $(N_{M,q}^K, N_{M,r}^K)$  and  $(Z_q^K, Z_r^K)$  are defined as follows:

$$\begin{aligned} N_{M,q}^K(t) &= 1 + \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{u \leq m^K(s-)\}} \mathcal{N}_M^b(ds, du) - \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{0 < u \leq p^K(s-)\}} \mathcal{N}_M^d(ds, du) \\ N_{M,r}^K(t) &= \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{m^K(s-) < u \leq N_{M,q}^K(s-)b_M(X^K(s-))\}} \mathcal{N}_M^b(ds, du) \\ &\quad + \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{N_{M,q}^K(s-)d_M(X^K(s-)) < u \leq p^K(s-)\}} \mathcal{N}_M^d(ds, du) \\ &\quad + \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{0 \leq u - N_{M,q}^K(s-)b_M(X^K(s-)) \leq N_{M,r}^K(s-)b_M(X^K(s-))\}} \mathcal{N}_M^b(ds, du) \\ &\quad - \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{0 \leq u - N_{M,q}^K(s-)d_M(X^K(s-)) \leq N_{M,r}^K(s-)d_M(X^K(s-))\}} \mathcal{N}_M^d(ds, du) \end{aligned}$$

and in a similar way,

$$\begin{aligned} Z_q^K(t) &= N_{M,q}^K(t) \\ Z_r^K(t) &= \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{m^K(s-) < u \leq Z_q^K(s-)b_*\}} \mathcal{N}_M^b(ds, du) \\ &\quad + \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{Z_q^K(s-)d_* < u \leq p^K(s-)\}} \mathcal{N}_M^d(ds, du) \\ &\quad + \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{0 \leq u - Z_q^K(s-)b_* \leq Z_r^K(s-)b_*\}} \mathcal{N}_M^b(ds, du) \\ &\quad - \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{0 \leq u - Z_q^K(s-)d_* \leq Z_r^K(s-)d_*\}} \mathcal{N}_M^d(ds, du). \end{aligned}$$

We can easily check that



**Lemma 2.4.** *The following equalities of processes hold, almost surely:*

$$N_M^K = N_{M,q}^K + N_{M,r}^K, \quad Z = Z_q^K + Z_r^K.$$

*Proof.* That results from the fact that  $N_{M,q}^K + N_{M,r}^K$  and  $N_M^K$  are solutions of the same stochastic differential equation for which pathwise uniqueness holds. A similar argument also holds for  $Z$ .  $\square$

Let us introduce the time when the gap to the equilibrium for resident population goes beyond some level  $n \geq 0$  starting from the initial condition  $N^K(0) = n_0 \in \mathbb{N}^2$ :

$$T_R^K(n_0, n) = \inf \{t > 0 : |N_R^K(t) - x_R^*K| > n\} = \inf \{t > 0 : |X_R^K(t) - x_R^*| > n/K\}.$$

We consider the (potentially random) initial population size  $N^K(0)$  and introduce the stopping time

$$T_{R,M}^K(m, n) = T_M^K(N^K(0), m) \wedge T_R^K(N^K(0), n) \wedge \frac{4 \log K}{3r_*}. \quad (2.1)$$

The residual processes  $N_r^K, Z_r^K$  will be compared to an inhomogeneous branching process and proved to be negligible until  $T_{R,M}^K(m, n)$ , with  $m$  and  $n$  suitably chosen.

**Remark 2.5.** *This stopping time considers the first time when the population of mutants goes beyond  $m$  or the population of residents has moved of  $n$  from its equilibrium value. It is also truncated at a deterministic value of order  $\log K$ . For this truncation, we can consider any time  $\lambda \log K/r_*$  (with  $\lambda > 1$ ). We choose  $\lambda := 4/3$  in definition (2.1) to optimize the convergence speed of Theorem 2.2. This particular choice only matters in the proof of Lemma 2.8.*

**Lemma 2.6.** *Under Assumptions 1.1 and 2.1,*

i) *If  $\partial_R F_R(x^*) < 0$ , for any  $L > 0$ , there exists  $C_L > 0$  such that for any  $n \leq K/(\log K)$ ,*

$$\mathbb{E} \left[ \sup_{t \leq T_{R,M}^K(Ln, n)} (Z_r^K(t) + N_{M,r}^K(t)) e^{-r_* t} \right] \leq C_L \left( \frac{n \log K}{K} \right)^{1/2} \exp \left( C_L \frac{n \log K}{K} \right).$$

ii) *If  $\partial_R F_R(x^*) = 0$ , there exists  $C \geq 0$  such that for any  $n \leq K/(\log K)^2$ ,*

$$\mathbb{E} \left[ \sup_{t \leq T_{R,M}^K(\sqrt{Kn}, n)} (Z_r^K(t) + N_{M,r}^K(t)) e^{-r_* t} \right] \leq C \left( \log K \sqrt{\frac{n}{K}} \right)^{1/2} \exp \left( C \log K \sqrt{\frac{n}{K}} \right).$$

*Proof.* Since  $b_M$  and  $d_M$  are locally Lipschitz, there exists  $C_L > 0$  (which may change from line to line) such that for any  $s \leq T_{R,M}^K(Ln, n)$  and  $n \leq K$ ,

$$Z_q^K(s) |b_M(X^K(s)) - b_*| + Z_q^K(s) |d_M(X^K(s)) - d_*| \leq C_L Z_q^K(s) \frac{n}{K} \quad (2.2)$$

almost surely. Then, recalling that  $Z_q^K = N_{M,q}^K$ ,

$$Z_q^K(s) b_* - m^K(s) + p^K(s) - Z_q^K(s) d_* \leq C_L Z_q^K(s) \frac{n}{K}.$$

Therefore, the process  $Z_r^K$  can be stochastically dominated on the time window  $[0, T_{R,M}^K(Ln, n)]$  by a branching process with inhomogeneous immigration  $Y$ , unique strong solution of

$$\begin{aligned} Y(t) &= \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{u \leq C_L Z_q^K(s-) \frac{n}{K}\}} \mathcal{N}^I(ds, du) \\ &\quad + \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{0 \leq u \leq Y(s-) b_*\}} \mathcal{N}_Y^b(ds, du) - \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{0 \leq u \leq Y(s-) d_*\}} \mathcal{N}_Y^d(ds, du), \end{aligned}$$

where  $\mathcal{N}^I, \mathcal{N}_Y^b, \mathcal{N}_Y^d$  are independent Poisson point measures with Lebesgue intensity on  $\mathbb{R}_+^2$ . We can easily note that for any  $t \geq 0$ ,

$$\mathbb{E} [Z_q^K(t)] \leq \mathbb{E} [Y(t)] = e^{r_* t}$$

and straightforward computation leads to

$$\mathbb{E} [Y(t \wedge T_{R,M}^K(Ln, n))] \leq C_L \frac{n}{K} e^{2r_* t}. \quad (2.3)$$

Indeed, denoting  $y(t) = \mathbb{E} [Y(t \wedge T_{R,M}^K(Ln, n))]$ , the function  $y$  satisfies the following : for all  $t > 0$ ,

$$\begin{aligned} y(t) &= C_L \frac{n}{K} \int_0^{t \wedge T_{R,M}^K(Ln, n)} \mathbb{E} [Z_q^K(s)] ds + r_* \int_0^{t \wedge T_{R,M}^K(Ln, n)} y(s) ds \\ &\leq C_L \frac{n}{K} \int_0^t e^{r_* s} ds + r_* \int_0^t y(s) ds \\ &\leq C_L \frac{n}{K} e^{r_* t} / r_* + r_* \int_0^t y(s) ds. \end{aligned}$$

Hence Gronwall's inequality gives (2.3).

Let  $M(t)$  be the martingale part of the semimartingale  $Y(t)$  (i.e. defined by the same expression as  $Y$  replacing the three Poisson point measures by their compensated versions), with quadratic variation process given by

$$\langle M \rangle(t) = \int_0^t \left( C_L Z_q^K(s) \frac{n}{K} + Y(s)(b_* + d_*) \right) ds.$$

Using Doob and Cauchy Schwarz inequalities and  $T_{R,M}^K(Ln, n) \leq 4(\log K)/(3r_*)$  by (2.1), we obtain from the previous estimates

$$\mathbb{E} \left( \sup_{t \leq T_{R,M}^K(Ln, n)} \left| \int_0^t e^{-r_* s} dM(s) \right| \right) \leq C \mathbb{E} \left( \int_0^{T_{R,M}^K(Ln, n)} e^{-2r_* s} d\langle M \rangle(s) \right)^{1/2} \leq C_L \sqrt{\frac{n \log K}{K}}.$$

We compute now

$$\begin{aligned} Y(t) e^{-r_* t} &= \int_0^t e^{r_* s} dY(s) - r_* \int_0^t Y(s) e^{r_* s} ds \\ &= \int_0^t e^{-r_* s} dM(s) + \int_0^t e^{-r_* s} \left( C_L Z_q^K(s) \frac{n}{K} + Y(s)(b_* - d_*) - Y(s)(b_* - d_*) \right) ds \\ &= \int_0^t e^{-r_* s} dM(s) + \int_0^t e^{-r_* s} C_L Z_q^K(s) \frac{n}{K} ds. \end{aligned}$$

Combining the previous estimates and using again  $T_{R,M}^K(n, Ln) \leq 4 \log K / (3r_*)$ , we get

$$\mathbb{E} \left( \sup_{t \leq T_{R,M}^K(Ln, n)} Y(t) e^{-r_* t} \right) \leq C_L \left( \sqrt{\frac{n \log K}{K}} + \frac{n \log K}{K} \right).$$

Using the domination of  $Z_r^K$  by  $Y$ , we obtain

$$\mathbb{E} \left( \sup_{t \leq T_{R,M}^K(Ln,n)} Z_r^K(t) e^{-r_* t} \right) \leq C_L \sqrt{\log K} \frac{n}{K},$$

which proves the first estimate *i*) for  $Z_r$ .

The assertion *i*) for  $N_{M,r}^K$  can be proved in a similar way. Indeed the individual birth and death rates are respectively upper-bounded by  $b_* + C_L n/K$  and lower-bounded by  $d_* - C_L n/K$  for  $t \leq T_{R,M}^K(Ln,n)$ . So  $N_{M,r}^K$  is dominated on the time window  $[0, T_{R,M}^K(Ln,n)]$  by the process  $\tilde{Y}$  defined by

$$\begin{aligned} \tilde{Y}(t) &= \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{u \leq C_L Z_q(s-)n/K\}} \mathcal{N}^I(ds, du) \\ &\quad + \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{0 \leq u \leq \tilde{Y}(s-)(b_* + C_L n/K)\}} \mathcal{N}_{\tilde{Y}}^b(ds, du) - \int_0^t \int_{\mathbb{R}_+} \mathbf{1}_{\{0 \leq u \leq \tilde{Y}(s-)(d_* - C_L n/K)\}} \mathcal{N}_{\tilde{Y}}^d(ds, du). \end{aligned}$$

The conclusion follows as above by considering  $\tilde{Y}(t) \exp(-(r_* + 2C_L n/K)t)$  for  $t \leq T_{R,M}^K(Ln,n)$ .

For the second case *ii*), we proceed similarly. The only difference is the upper-bound for the resident population in the stopping time  $T_{R,M}^K$ , and (2.2) becomes

$$Z_q^K(s) |b_M(X^K(s)) - b_*| + Z_q^K(s) |d_M(X^K(s)) - d_*| \leq C Z_q^K(s) \sqrt{\frac{n}{K}},$$

almost surely for  $s \leq T_{R,M}^K(\sqrt{Kn}, n)$ , since  $n \leq K$ . Following the steps of case *i*),

$$\mathbb{E} \left( \sup_{t \leq T_{R,M}^K(\sqrt{Kn}, n)} Z_r^K(t) e^{-r_* t} \right) \leq C \left( \log K \sqrt{\frac{n}{K}} \right)^{1/2}.$$

Proceeding similarly for  $N_{M,r}^K$  ends the proof. Notice that the exponential rescaling give the exponential factor in the statement of the lemma.  $\square$

The next lemma guarantees that, in some sense,  $N^K$  is equivalent to  $Z$  in probability until suitably chosen time  $T_{R,M}^K(m, n)$ . The proof uses the fact that, by definition,  $N_{M,q}^K$  and  $Z_q^K$  are the same process, whence it is sufficient to control the residual processes  $N_{M,r}^K$  and  $Z_r^K$  normalized by  $Z$ . The previous Lemma 2.6 is important for this last step.

**Lemma 2.7.** *Under Assumptions 1.1 and 2.1,*

*i) If  $\partial_R F_R(x^*) < 0$ , for any  $L > 0$ , there exists  $C_L > 0$  such that for any small enough  $\eta > 0$  and  $n \leq K/(\log K)$ ,*

$$\mathbb{P} \left( \sup_{t \leq T_{R,M}^K(Ln,n)} \left| \frac{N_M^K(t)}{Z(t)} - 1 \right| > \eta ; W > 0 \right) \leq \frac{C_L}{\eta^{1/4}} \left( \frac{n \log K}{K} \right)^{1/10} \exp \left( C_L \frac{n \log K}{K} \right). \quad (2.4)$$

*ii) If  $\partial_R F_R(x^*) = 0$ , there exists  $C > 0$  such that for any small enough  $\eta > 0$  and  $n \leq K/(\log K)^2$ ,*

$$\mathbb{P} \left( \sup_{t \leq T_{R,M}^K(\sqrt{Kn}, n)} \left| \frac{N_M^K(t)}{Z(t)} - 1 \right| > \eta ; W > 0 \right) \leq \frac{C}{\eta^{1/4}} \left( \log K \sqrt{\frac{n}{K}} \right)^{1/10} \exp \left( C \log K \sqrt{\frac{n}{K}} \right). \quad (2.5)$$

*Proof.* Let us prove *i*) and focus on the survival event  $\{W > 0\}$  where the process  $Z$  stays positive. On this event, using Lemma 2.4, we obtain

$$\frac{|N_M^K(t) - Z(t)|}{Z(t)} = \frac{|N_{M,r}^K(t) - Z_r^K(t)|}{Z(t)} = \frac{|N_{M,r}^K(t) - Z_r^K(t)|e^{-r_*t}}{W(t)},$$

where  $W(t) = Z(t)e^{-r_*t}$  is the classical martingale associated with  $Z$ .

We use now Lemma 2.6 *i*) and the fact that  $W(t)$  converges to  $W$  a.s. as  $t$  tends to infinity. We deduce that for  $t \leq T_{R,M}^K(Ln, n)$ ,  $|N_{M,r}^K(t) - Z_r^K(t)|/Z(t)$  tends to 0 in probability on the event  $W > 0$ , as  $K$  tends to infinity.

Let us now obtain an explicit convergence speed. For convenience we define

$$U^K = \sup_{t \leq T_{R,M}^K(Ln, n)} Z_r(t)e^{-r_*t} + \sup_{t \leq T_{R,M}^K(Ln, n)} N_{M,r}^K(t)e^{-r_*t}.$$

The previous computations ensure that

$$\frac{|N_M^K(t) - Z(t)|}{Z(t)} \leq \frac{U^K}{\inf_{t \geq 0} W(t)}$$

a.s. on the event  $W > 0$ . Moreover, choosing  $\varepsilon = n \log K / K$ , for any  $\eta > 0$ ,

$$\mathbb{P}\left(U^K \geq \eta \inf_{t \geq 0} W(t); W > 0\right) \leq \mathbb{P}\left(U^K \geq \varepsilon^{2/5}\right) + \mathbb{P}\left(\inf_{t \geq 0} W(t) \leq \frac{1}{\eta} \varepsilon^{2/5}; W > 0\right).$$

By Markov's inequality and Lemma 2.6,

$$\mathbb{P}\left(U^K \geq \varepsilon^{2/5}\right) \leq C_L \varepsilon^{1/2-2/5} e^{C_L \varepsilon} = C_L \varepsilon^{1/10} e^{C_L \varepsilon}.$$

On the other hand, by Lemma 6.2 in Appendix,

$$\mathbb{P}\left(\inf_{t \geq 0} W(t) \leq \frac{1}{\eta} \varepsilon^{2/5}; W > 0\right) \leq C \frac{1}{\eta^{1/4}} \varepsilon^{2/5 \times 1/4} = C \frac{1}{\eta^{1/4}} \varepsilon^{1/10}.$$

This proves the result in the case *i*). Case *ii*) is proven similarly.  $\square$

Let us now compare the times when mutant and resident population processes reach their threshold.

**Lemma 2.8.** *We denote  $s_K = 4(\log K)/(3r_*)$ . Under Assumptions 1.1 and 2.1,*

*i) If  $\partial_R F_R(x^*) < 0$ , then there exist  $L, C > 0$  such that for any  $K \geq 2$  and  $n \leq K/\log K$ ,*

$$\mathbb{P}\left(T_M^K(N^K(0), n) \geq T_R^K(N^K(0), Ln) \wedge s_K; W > 0\right) \leq C \left( \frac{\sqrt{K \log K}}{n} + \frac{1}{K^{1/3}} + \left(\frac{n \log K}{K}\right)^{1/10} \right).$$

*ii) If  $\partial_R F_R(x^*) = 0$ , then there exist  $L, C > 0$  such that for any  $K \geq 2$  and  $n \leq LK/\log(K)^2$ ,*

$$\mathbb{P}\left(T_M^K(N^K(0), n) \geq T_R^K(N^K(0), \sqrt{Kn}) \wedge s_K; W > 0\right) \leq C \left( \sqrt{\frac{\log K}{n}} + \frac{1}{K^{1/3}} + \left(\log K \sqrt{\frac{n}{K}}\right)^{1/10} \right).$$

Observe that  $\{T_M^K(N^K(0), n) \geq T_R^K(N^K(0), m) \wedge s_K\} = \{T_R^K(N^K(0), m) = T_{R,M}^K(m, n)\}$ . We are using the result for sequences with  $m = \xi_K$  which makes the right hand side go to zero. This will ensure that the mutant population process reaches its threshold before the resident population process.

*Proof of case i).* Let us write  $T^K = T_{R,M}^K(Ln, n)$  for convenience. The first step consists in controlling the resident population. Let us prove that there exist  $L, C > 0$  such that for any  $K \geq 2$  and  $n \leq K/\log K$ ,

$$\begin{aligned} \mathbb{P}(T_R^K(N^K(0), Ln) \leq T_M^K(N^K(0), n) \wedge s_K) \\ = \mathbb{P}\left(\sup_{t \leq T^K} |X_R^K(t) - x_R^*| \geq L \frac{n}{K}\right) \leq C \frac{\sqrt{K \log K}}{n}, \end{aligned} \quad (2.6)$$

For that purpose, we first notice that, for any  $t \leq T^K$  and  $K \geq 2$ ,

$$X_M^K(t) \leq \frac{n}{K} \leq \frac{1}{\log K} \quad \text{and} \quad X_R^K(t) \leq x_R^* + L \frac{n}{K} \leq x_R^* + \frac{L}{\log K}. \quad (2.7)$$

Hence, the processes  $X_M^K$  and  $X_R^K$  are bounded up to time  $T^K$ , uniformly for  $K \geq 2$ . We define

$$\begin{aligned} M^K(t) &= \frac{1}{K} \int_{[0,t] \times \mathbb{R}_+} \mathbf{1}_{\{u \leq N_R^K(s-) b_R(X^K(s-))\}} \tilde{\mathcal{N}}_R^b(ds, du) \\ &\quad - \frac{1}{K} \int_{[0,t] \times \mathbb{R}_+} \mathbf{1}_{\{u \leq N_R^K(s-) d_R(X^K(s-))\}} \tilde{\mathcal{N}}_R^d(ds, du), \end{aligned}$$

where  $\tilde{\mathcal{N}}_R^b$  and  $\tilde{\mathcal{N}}_R^d$  are the compensated Poisson martingale measures of  $\mathcal{N}_R^b$  and  $\mathcal{N}_R^d$ . We obtain

$$X_R^K(t) = X_R^K(0) + M^K(t) + \int_0^t G_R(X^K(s)) ds,$$

where

$$G_R(x_R, x_M) = x_R(b_R - d_R)(x_R, x_M) = x_R F_R(x_R, x_M).$$

The quadratic variation process of the martingale part  $M^K$  is the following

$$\langle M^K \rangle(t) = \frac{1}{K} \int_0^t X_R^K(s) \left( b_R(X^K(s)) + d_R(X^K(s)) \right) ds. \quad (2.8)$$

Observing that  $G_R(x^*) = 0$  by definition of  $x_R^*$ , we rewrite the dynamics above as

$$X_R^K(t) = X_R^K(0) + M^K(t) + H_1^K(t) + H_2^K(t) + \int_0^t (X_R^K(s) - x_R^*) \partial_R G_R(x^*) ds,$$

where

$$H_1^K(t) = \int_0^t X_M^K(s) \partial_M G_R(x^*) ds$$

and

$$H_2^K(t) = \int_0^t \left( G_R(X^K(s-)) - G_R(x^*) - (X_R^K(s) - x_R^*) \partial_R G_R(x^*) - X_M^K(s) \partial_M G_R(x^*) \right) ds.$$

Denoting

$$Y_R^K(t) := X_R^K(t) - x_R^*,$$

we have

$$Y_R^K(t) = (X_R^K(0) - x_R^*) + H^K(t) + \int_0^t Y_R^K(s) \partial_R G_R(x^*) ds, \quad (2.9)$$

where

$$H^K(t) := M^K(t) + H_1^K(t) + H_2^K(t).$$

Recall that  $\partial_R F_R(x^*) < 0$  and so  $\partial_R G_R(x^*) < 0$ . Then Lemma 6.3 ensures that for any  $K$

$$\sup_{0 \leq s \leq T^K} |Y_R^K(s)| \leq \Gamma |X_R^K(0) - x_R^*| + \Gamma \sup_{0 \leq s \leq T^K} |H^K(s) - H^K(\lfloor s \rfloor)|,$$

with

$$\Gamma := (1 + |\partial_R G_R(x^*)|) \frac{1}{1 - e^{\partial_R G_R(x^*)}}.$$

Using that  $H^K = M^K + H_1^K + H_2^K$  and (using Taylor-Lagrange's inequality to obtain the following bound for  $H_2^K$ )

$$\sup_{t \leq T^K} |H_1^K(t) - H_1^K(\lfloor t \rfloor)| \leq \frac{n}{K} |\partial_M G_R(x^*)|, \quad \sup_{t \leq T^K} |H_2^K(t) - H_2^K(\lfloor t \rfloor)| \leq C \left(\frac{n}{K}\right)^2$$

almost surely for some constant  $C > 0$ , by definition of  $T^K$  and  $H_1^K, H_2^K$ , we get

$$\sup_{0 \leq s \leq T^K} |Y_R^K(s)| \leq \Gamma \left( \sup_{t \leq T^K} |M^K(t) - M^K(\lfloor t \rfloor)| + C \frac{n}{K} \right).$$

Besides, using (2.7) and (2.8) and  $T^K \leq C \log K$ , we get  $\mathbb{E}[\langle M^K \rangle_{T^K}] \leq C(\log K)/K$ . By Markov's and Burkholder-Davis-Gundy's inequalities, we then obtain

$$\begin{aligned} \mathbb{P} \left( \sup_{t \leq T^K} |M^K(t) - M^K(\lfloor t \rfloor)| \geq \frac{n}{K} \right) &\leq C \frac{K}{n} \mathbb{E} \left[ \sup_{t \leq T^K} |M^K(t)| \right] \\ &\leq C \frac{K}{n} \mathbb{E} [\langle M^K \rangle_{T^K}]^{1/2} \\ &\leq C \frac{K}{n} \left( \frac{\log K}{K} \right)^{1/2} \leq C \frac{\sqrt{K \log K}}{n}. \end{aligned}$$

Recalling that  $Y_R^K(t) = X_R^K(t) - x_R^*$  and  $n \leq K$ , choosing  $L$  large enough and combining these two last estimates yields the expected inequality (2.6) and ends the first step.

We turn to the second step and prove that

$$\mathbb{P}(T_M^K(N^K(0), n) \wedge T_R^K(N^K(0), Ln) \geq s_K; W > 0) \leq CK^{-1/3} + C \left( \log K \sqrt{\frac{n}{K}} \right)^{1/10}. \quad (2.10)$$

Let us notice that, a.s. on the event  $\{W > 0\}$ ,

$$\begin{aligned} &\{T_M^K(N^K(0), n) \wedge T_R^K(N^K(0), n) \geq s_K\} \\ &\subseteq \left( \{N_M^K(s_K) \leq n + 1\} \cap \left\{ \frac{Z(s_K)}{N_M^K(s_K)} \leq 2 \right\} \right) \cup \left\{ \sup_{t \leq T^K} \left| \frac{N_M^K(t)}{Z(t)} - 1 \right| > \frac{1}{2} \right\} \\ &\subseteq \{Z(s_K) \leq K + 1\} \cup \left\{ \sup_{t \leq T^K} \left| \frac{N_M^K(t)}{Z(t)} - 1 \right| > \frac{1}{2} \right\}. \end{aligned}$$

Hence, using  $Z(s_K) = W(s_K)e^{r^* s_K} = W(s_K)K^{4/3}$  (by definition of  $s_K$ ),

$$\begin{aligned} &\mathbb{P}(T_M^K(N^K(0), n) \wedge T_R^K(N^K(0), n) \geq s_K; W > 0) \\ &\leq \mathbb{P}(W(s_K) \leq 2K^{-1/3}; W > 0) + \mathbb{P} \left( \sup_{t \leq T^K} \left| \frac{N_M^K(t)}{Z(t)} - 1 \right| > \frac{1}{2}; W > 0 \right). \end{aligned}$$

The first term of the right hand side involves an estimate on the infimum of a single type branching process. It will be stated and proved in Lemma 6.1 in forthcoming Appendix. The second term of the sum in the right hand side has been controlled in Lemma 2.7 *i*). Observing that the exponential term in Lemma 2.7 is bounded since  $n \leq K/\log K$  and combining these estimates yields (2.10).

We can now conclude the case *i*) combining the two steps. More precisely we use that for any  $x, y, z \in \mathbb{R}$ ,  $(x \geq y \wedge z) \iff (x \wedge y \geq z \text{ or } x \wedge z \geq y)$ . The result follows then from (2.10) and (2.6).  $\square$

*Proof of the case ii*). We prove a similar result under the hypothesis  $\partial_R F_R(x^*) = 0$ . Let us prove that there exists  $L > 0$  such that for any  $K \geq 1$  and  $n \leq L.K/(\log K)^2$

$$\mathbb{P} \left( \sup_{t \leq T^K} |X_R^K(t) - x_R^*| \geq \sqrt{n/K} \right) \leq C \sqrt{\frac{\log K}{n}}, \quad (2.11)$$

where now  $T^K := T_{R,M}^K(\sqrt{nK}, n)$ .

For that purpose, we use again (2.9) and get directly

$$\begin{aligned} \sup_{t \leq T^K} |X_R^K(t) - x_R^*| &= \sup_{s \leq T^K} |Y_R^K(t)| \\ &\leq |X_R^K(0) - x_R^*| + \sup_{t \leq T^K} |M^K(t)| + \sup_{t \leq T^K} |H_1^K(t)| + \sup_{t \leq T^K} |H_2^K(t)|. \end{aligned} \quad (2.12)$$

Recalling that  $T_K \leq 4(\log K)/(3r_*)$  and proceeding as in the case *i*), we obtain

$$\begin{aligned} \mathbb{P} \left( \sup_{t \leq T^K} |M^K(t)| \geq \frac{1}{2} \sqrt{\frac{n}{K}} \right) &\leq C \sqrt{\frac{K}{n}} \cdot \sqrt{\frac{\log K}{K}} = C \sqrt{\frac{\log K}{n}}, \\ \sup_{t \leq T^K} |H_1^K(t)| + \sup_{t \leq T^K} |H_2^K(t)| &\leq C \frac{n}{K} T^K \leq C \frac{n}{K} \log K = C \sqrt{\frac{n}{K}} \sqrt{\frac{n(\log K)^2}{K}}. \end{aligned} \quad (2.13)$$

For any  $n \leq K/(2C \log(K))^2$ , we get

$$\sup_{t \leq T^K} |H_1^K(t)| + \sup_{t \leq T^K} |H_2^K(t)| \leq \frac{1}{2} \sqrt{\frac{n}{K}}.$$

Then (2.12) guarantees

$$\left\{ \sup_{t \leq T^K} |X_R^K(t) - x_R^*| \geq \sqrt{\frac{n}{K}} \right\} \subset \left\{ \sup_{t \leq T^K} |M^K(t)| \geq \frac{1}{2} \sqrt{\frac{n}{K}} \right\}$$

and (2.13) yields (2.11) with  $L = 1/(4C^2)$ . The proof is concluded as for the case *i*).  $\square$

Theorem 2.2 becomes now a consequence of the two previous lemmas.

*Proof of Theorem 2.2.* We recall that  $s_K = 4(\log K)/(3r_*)$  and in the case *i*) we use

$$\begin{aligned} \mathbb{P} \left( \sup_{t \leq T_M^K(N^K(0), n)} \left| \frac{N_M^K(t)}{Z(t)} - 1 \right| > \eta; W > 0 \right) \\ \leq \mathbb{P} (T_M^K(N^K(0), n) > T_R^K(N^K(0), Ln) \wedge s_K, W > 0) \\ + \mathbb{P} \left( \sup_{t \leq T_{R,M}^K(Ln, n)} \left| \frac{N_M^K(t)}{Z(t)} - 1 \right| > \eta; W > 0 \right). \end{aligned}$$

The two terms of the right hand side can be controlled using respectively Lemmas 2.8 to choose  $L$  and 2.7. This completes the proof for *i*), and *ii*) is proved similarly.  $\square$

### 3. Approximation by dynamical systems

The goal of this section is to approximate the process  $X^K = (X_R^K, X_M^K)$  defined in Section 1 by the dynamical system defined by (1.2) under degenerate initial conditions. Usually such an approximation is proved when the initial conditions are of order of magnitude one (cf. for example [22]). Here, we focus on a small deterministic initial density of mutants,  $1/K \ll x_M = N_M^K(0)/K \ll 1$  and an initial density of residents  $x_R = N_R^K(0)/K$  close to the equilibrium value  $x_R^*$ . The initial mutant density may be very small regarding  $K$  but we still enjoy a large number of mutants allowing to compare the stochastic process to its expected deterministic behavior. Starting from such sub-macroscopic level of mutants, the time for them to reach macroscopic levels will be of order  $\log K$ . We will compare the density process and its deterministic approximation on this invasion time scale. They are both small at the beginning and we will show that their ratio remains close to 1 on the full time window allowing to reach macroscopic density of mutants.

This section emphasizes the role of (small) initial conditions and we introduce the flow notation for convenience. The process  $X^K(x, \cdot)$  started from  $x$  is defined for any  $K$  and time  $t$  by

$$\begin{aligned} X_{\bullet}^K(x, t) = & x_{\bullet} + \frac{1}{K} \int_{[0, t] \times \mathbb{R}_+} \mathbf{1}_{\{u \leq K X_{\bullet}^K(x, s-) b_{\bullet}(X^K(x, s-))\}} \mathcal{N}_{\bullet}^b(ds, du) \\ & - \frac{1}{K} \int_{[0, t] \times \mathbb{R}_+} \mathbf{1}_{\{u \leq K X_{\bullet}^K(x, s-) d_{\bullet}(X^K(x, s-))\}} \mathcal{N}_{\bullet}^d(ds, du) \end{aligned} \quad (3.1)$$

for  $\bullet \in \{R, M\}$ . The flow  $\phi(x, t) = (\phi_R(x, t), \phi_M(x, t))$  is the solution of (1.2) starting from the initial value  $x$ , i.e. the unique solution of

$$\phi(x, 0) = x, \quad \frac{\partial}{\partial t} \phi(x, t) = G(\phi(x, t)), \quad \text{where } G_{\bullet}(x) = x_{\bullet}(b_{\bullet}(x) - d_{\bullet}(x)).$$

In this section we assume that the flow  $\phi$  satisfies a boundedness assumption: the solution remains in a compact set if it starts close to the equilibrium  $x^*$ .

**Assumption 3.1.** *There exists a compact domain  $\mathcal{D}$  of  $\mathbb{R}_+^2$  and  $\varepsilon > 0$  such that*

$$(x_R^* - \varepsilon, x_R^* + \varepsilon) \times [0, \varepsilon) \subset \mathcal{D} \quad \text{and} \quad \{\phi(x, t) : x \in \mathcal{D}, t \geq 0\} \text{ is bounded.}$$

#### 3.1. Properties of the dynamical system and hitting times

Let us consider the hitting time for the mutant population starting from  $x$

$$\tau_M(x, v) = \inf\{t \geq 0 : \phi_M(x, t) = v\},$$

where by convention  $\inf \emptyset = \infty$  and in that case the hitting time is infinite. The following result quantifies the hitting time  $\tau_M(x, v)$  as a function of the two variables  $x$  and  $v$ , starting close to the equilibrium  $x^* = (x_R^*, 0)$ . Let us define, for  $\eta > 0$ ,

$$\mathcal{C}(x, \eta) = \sup\{\phi_M(x, t) : t \leq \bar{\tau}_M(x, \eta)\},$$

where  $\bar{\tau}_M(x, \eta)$  is the first time when the growth rate of mutants is smaller than  $\eta$ :

$$\bar{\tau}_M(x, \eta) = \inf\{t \geq 0 : F_M(\phi(x, t)) \leq \eta\}.$$

We set

$$v_{\eta} = \lim_{r \downarrow 0} \inf_{x \in B_r(x^*) \cap (\mathbb{R}_+^*)^2} \mathcal{C}(x, \eta),$$



where  $B_r(x^*)$  denotes the ball of radius  $r$  centered in  $x^*$ . Observing that  $v_\eta$  is increasing as  $\eta$  decreases, we let  $\eta$  go to zero and define  $v_\star$  as the upper-bound of the values of mutants which can be reached in the increasing phase, starting from the neighborhood of the equilibrium:

$$v_\star := \lim_{\eta \downarrow 0} v_\eta.$$

The next result guarantees that  $v_\star$  is positive and finite. It also quantifies the hitting times  $\tau_M(x, v)$  when  $x$  goes close to  $x^*$  and  $v$  is smaller than  $v_\star$ .

**Proposition 3.2.** *Under Assumptions 1.1 and 3.1, the value  $v_\star$  is positive and finite.*

*Besides there exists a continuous increasing function  $\tau : [0, v_\star) \rightarrow \mathbb{R}_+$  such that  $\tau(0) = 0$  and for any  $\bar{v} \in (0, v_\star)$ ,*

$$\lim_{x \rightarrow x^*} \sup_{v \in (x_M, \bar{v}]} \left| \tau_M(x, v) - \frac{1}{r_\star} \log(v/x_M) - \tau(v) \right| = 0,$$

where the limit is taken for  $x \in (\mathbb{R}_+^*)^2$  if  $\partial_R F_R(x^*) < 0$ , and for

$$x \in D(\eta) = \left\{ (x_R, x_M) \in (\mathbb{R}_+^*)^2 : |x_R - x_R^*| \log(1/x_M) \leq \eta \right\}$$

for some  $\eta > 0$  if  $\partial_R F_R(x^*) = 0$ .

The role of the set  $D(\eta)$  is to guarantee a suitable control of the resident population. Indeed, as  $x_M$  decreases, the time to reach  $v$  increases (since it is of order  $\log(v/x_M)$ ), and so  $x_R$  needs to be closer to  $x_R^*$  to compensate it.

The proof of this proposition necessitates some lemmas, which guarantee the resident process to remain close to its equilibrium value as long as the mutant process remains small.

**Lemma 3.3.**

i) *If  $\partial_R F_R(x^*) < 0$ , then there exist  $u_0 \in (0, 1)$  and  $C > 0$  such that for any  $u \leq u_0$  and  $x = (x_R, x_M) \in B_u(x^*)$ ,*

$$\sup_{t \leq \tau_M(x, u)} |\phi_R(x, t) - x_R^*| \leq Cu.$$

ii) *If  $\partial_R F_R(x^*) = 0$ , then there exist  $u_0 \in (0, 1)$  and  $\eta_0 > 0$  such that, for all  $u \leq u_0$  and  $x \in B_u(x^*) \cap D(\eta_0)$ ,*

$$\sup_{t \leq \tau_M(x, u)} |\phi_R(x, t) - x_R^*| \leq Cu.$$

iii) *In both cases, for all  $u \leq u_0$  and  $x \in B_u(x^*) \cap D(\eta_0)$ ,*

$$\tau_M(x, u) \leq \frac{2}{r_\star} \log(1/x_M).$$

*Proof.* For convenience we introduce the gap between the flow with its initial position for the resident population dynamics and the first time when the flow leaves  $B_r(x^*)$  (for any  $r > 0$ ):

$$y_R(t) := \phi_R(x, t) - x_R, \quad s_r(x) := \inf\{t \geq 0 : \phi(x, t) \notin B_r(x^*)\}. \quad (3.2)$$

Note that, for the sake of notation, the initial condition  $x$  is not written explicitly in the quantity  $y_R(t)$  which still depends on it.

To begin with, let us prove that, for  $r > 0$  small enough,

$$s_r(x) \leq \frac{2}{r_\star} \log(1/x_M). \quad (3.3)$$

Recalling that  $F_M(x^*) = r_\star > 0$ , we have, for  $r$  small enough and  $t \leq s_r(x)$ ,

$$\partial_t \phi_M(x, t) = \phi_M(x, t) F_M(\phi(x, t)) \geq \phi_M(x, t) \frac{r_\star}{2}.$$

Whence, for all  $t \leq s_r(x)$ ,

$$\phi_M(x, t) \geq x_M e^{t \cdot r_\star/2}.$$

The inequality above implies that  $s_r(x)$  is finite. Then as  $t \mapsto \phi_M(x, t)$  is continuous, by choosing  $r < 1$ , we have

$$x_M e^{s_r(x) \cdot r_\star/2} \leq \phi_M(x, s_r(x)) < 1,$$

which implies

$$s_r(x) \leq \frac{2}{r_\star} \log(1/x_M).$$

Now let us prove *i*). As  $\partial_R F_R(x_\star) < 0$ , we can find  $r > 0$  such that for any  $x \in B_r(x_\star)$ ,

$$\partial_R F_R(x) \leq \partial_R F_R(x_\star)/2 < 0. \quad (3.4)$$

Let us use that by definition of the flow  $\phi_R$

$$y_R(t) = \int_0^t \phi_R(x, s) F_R(\phi(x, s)) ds = H(t) + \int_0^t \phi_R(x, s) y_R(s) \partial_R F_R(x_R, \phi_M(x, s)) ds, \quad (3.5)$$

where

$$H(t) = \int_0^t \phi_R(x, s) (F_R(\phi(x, s)) - y_R(s) \partial_R F_R(x_R, \phi_M(x, s))) ds.$$

Recalling Assumption 1.1 (R) and using Taylor expansion, there exists  $C_r > 0$  such that

$$|F_R(\phi(x, s)) - F_R(x_R, \phi_M(x, s)) - y_R(s) \partial_R F_R(x_R, \phi_M(x, s))| \leq C_r |y_R(s)|^2$$

for  $s \leq s_r(x)$ . Using that  $F_R(x_\star) = 0$ , we get also

$$|F_R(x_R, \phi_M(x, s))| = |F_R(x_R, \phi_M(x, s)) - F_R(x_R, 0)| \leq C_r (|x_R - x_\star| + |\phi_M(x, s)|)$$

and, combining these two bounds,

$$\sup_{t \leq T \wedge s_r(x)} |H(t) - H(\lfloor t \rfloor)| \leq C_r \left( |x_R - x_\star| + \sup_{t \leq T \wedge s_r(x)} (|\phi_M(x, s)| + |y_R(s)|^2) \right).$$

Thanks to (3.4), we can use Lemma 6.3 for (3.5) and we get, for any  $T > 0$ ,

$$\begin{aligned} \sup_{t \leq T \wedge s_r(x)} |y_R(t)| &\leq C_r \left( \sup_{t \leq s_r(x) \wedge T} |H(t) - H(\lfloor t \rfloor)| \right) \\ &\leq C_r \left( |x_R - x_\star| + \sup_{t \leq T \wedge s_r(x)} (|\phi_M(x, s)| + |y_R(s)|^2) \right). \end{aligned} \quad (3.6)$$

In the rest of the proof, the value of  $C_r$  is fixed. We work now for times when  $y_R$  is also smaller than  $1/(2C_r)$  and introduce

$$\bar{s}_r(x) = \inf \left\{ t \geq 0 : |y_R(s)| \geq \frac{1}{2C_r} \right\} \wedge s_r(x).$$

Then for  $s \leq \bar{s}_r(x)$ , this yields  $C_r |y_R(s)|^2 \leq |y_R(s)|/2$  and gathering terms with  $y_R$ ,

$$\sup_{t \leq T \wedge \bar{s}_r(x)} |y_R(t)| \leq 2C_r \left( |x_R - x_\star| + \sup_{t \leq T \wedge s_r(x)} |\phi_M(x, s)| \right).$$

For any  $u > 0$  and  $x \in B_u(x^*)$ ,

$$\sup_{t \leq \tau_M(x, u) \wedge \bar{s}_r(x)} |y_R(t)| \leq 4C_r u. \quad (3.7)$$

Thus, setting  $u_0 = (r/2) \wedge r/(8C_r) \wedge 1/(16C_r^2)$  ensures that for any  $x \in B_{u_0}(x^*)$ ,

$$\sup_{t \leq \tau_M(x, u_0) \wedge \bar{s}_r(x)} |y_R(t)| \leq (r/2) \wedge 1/(4C_r),$$

which guarantees  $\bar{s}_r(x) > \tau_M(x, u_0)$ . Thus for  $u \leq u_0$ , one has  $\bar{s}_r(x) > \tau_M(x, u)$  and (3.7) becomes

$$\sup_{t \leq \tau_M(x, u)} |y_R(t)| \leq 4C_r u,$$

since  $\tau_M(x, u) \leq \tau_M(x, u_0) < \bar{s}_r(x)$ .

We prove now (ii) and consider the case  $\partial_R F_R(x^*) = 0$ . Recalling that the flow is bounded from Assumption 3.1, a Taylor expansion in  $x^*$  yields now  $|F_R(\phi(x, t))| = |F_R(\phi(x, t)) - F_R(x^*)| \leq C(|\phi_R(x, t) - x_R^*|^2 + |\phi_M(x, t)|)$  and recalling that  $y_R(t) = \phi_R(x, t) - x_R$ ,

$$|y'_R(t)| = |\phi_R(x, t)F_R(\phi(x, t))| \leq C[|x_R - x_R^*|^2 + y_R(t)^2 + \phi_M(x, t)]. \quad (3.8)$$

Let us now introduce the first time when the quadratic contribution may become dominant :

$$s(x) = \inf\{t \geq 0 : y_R(t)^2 \geq |x_R - x_R^*|^2 + \phi_M(x, t)\}.$$

Before this time, we can use  $|y'_R(t)| \leq 2C[|x_R - x_R^*|^2 + \phi_M(x, t)]$  and we get for  $t \leq s(x)$ ,

$$|y_R(t)| = \left| \int_0^t y'_R(s) ds \right| \leq C \left[ |x_R - x_R^*|^2 t + \int_0^t \phi_M(x, s) ds \right].$$

Let us use again the notation  $s_r(x)$  introduced at (3.2), and fix  $r > 0$  to some small enough value such that for  $t \leq s_r(x)$ ,  $F_M(\phi(x, t)) \geq r_*/2$ . Hence, for  $t \leq s_r(x)$ ,

$$\int_0^t \phi_M(x, s) ds = \int_0^t \partial_s \phi_M(x, s) \frac{1}{F_M(\phi(x, s))} ds \leq \frac{2}{r_*} \int_0^t \partial_s \phi_M(x, s) ds \leq \frac{2}{r_*} \phi_M(x, t). \quad (3.9)$$

We obtain for  $t \leq s(x) \wedge s_r(x)$  (using (3.3)),

$$|y_R(t)| = \left| \int_0^t y'_R(s) ds \right| \leq C[|x_R - x_R^*|^2 \log(1/x_M) + \phi_M(x, t)]. \quad (3.10)$$

From now the value of  $C$  is fixed and assumed to be greater than one (possibly by increasing it) and we consider  $\eta_0 := 1/(4C^2)$ . Let  $x \in D := D(1/(4C^2)) = \{(x_R, x_M) : |x_R - x_R^*| \log(1/x_M) \leq 1/(4C^2)\}$  and  $t \leq \tau_M(x, 1/(4C^2)) \wedge s(x) \wedge s_r(x)$  such that  $\phi_M(x, t) \leq 1/(4C^2)$ . We get

$$|y_R(t)|^2 \leq 2C^2 [(|x_R - x_R^*| \log(1/x_M))^2 |x_R - x_R^*|^2 + \phi_M(x, t)^2] \leq \frac{1}{2} [|x_R - x_R^*|^2 + \phi_M(x, t)].$$

This implies that for any  $x \in D$ ,  $s(x) > \tau_M(x, 1/(4C^2)) \wedge s_r(x)$  and (3.10) becomes

$$\sup_{t \leq \tau_M(x, 1/(4C^2)) \wedge s_r(x)} |y_R(t)| \leq C[|x_R - x_R^*| + \phi_M(x, t)].$$

Finally choosing some  $u_0 \leq 1/(4C^2) \wedge r/(4C)$ , the inequality above implies that  $\tau_M(x, u_0) < s_r(x)$ , which proves the result.

*Proof of iii).* A straightforward consequence of i) and ii) is that, in both cases, it is possible to define, for any  $r_0 > 0$  small enough, some  $u_0 > 0$  small enough such that for all  $u \leq u_0$ ,

$$\tau_M(x, u) \leq s_{r_0}(x).$$

Then (3.3) gives the result.  $\square$

The following lemma is not used in the proof of Proposition 3.2. However, since it is similar to the previous one (guaranteeing a control of the cumulated gap between the resident population dynamics and its equilibrium), we state and prove it here.

**Lemma 3.4.** *i) Assume that  $\partial_R F_R(x^*) < 0$ . Then there exist  $u_0 \in (0, 1)$  and  $\eta_0 > 0$  such that for all  $u \leq u_0, \eta \leq \eta_0$  and  $x \in B_u(x^*) \cap D(\eta)$ ,*

$$\int_0^{\tau_M(x,u)} \sup_{s \leq t} |\phi_R(x, s) - x_R^*| dt \leq C(u + \eta).$$

*ii) Assume that  $\partial_R F_R(x^*) = 0$ . Then there exist some  $u_0 \in (0, 1)$  and  $\eta_0 > 0$  such that for all  $u \leq u_0$ , and  $\eta \leq \eta_0$  and  $x \in B_u(x^*) \cap D(\eta)$ ,*

$$\int_0^{\tau_M(x,u)} \sup_{s \leq t} |\phi_R(x, s) - x_R^*| dt \leq C(u + \eta^2).$$

*Proof.* Here again,  $x \in B_u(x^*)$  and we will choose  $u$  small enough to have the following estimates.

Firstly, let us prove the result under the assumption  $\partial_R F_R(x^*) < 0$ . Using (3.6) and noting

$$z(t) = \sup_{s \leq t} |y_R(s)|$$

we have for all  $t \leq \tau_M(x, u)$ ,

$$z(t) \leq C(|x_R - x_R^*| + \phi_M(x, t) + z(t)^2),$$

since the function  $t \mapsto \phi_M(x, t)$  is non-decreasing on  $[0, \tau_M(x, u)]$  for  $u$  small enough. As a consequence,

$$\int_0^{\tau_M(x,u)} z(s) ds \leq C \left( \tau_M(x, u) |x_R - x_R^*| + \int_0^{\tau_M(x,u)} \phi_M(x, s) ds + z(\tau_M(x, u)) \int_0^{\tau_M(x,u)} z(s) ds \right).$$

We use Lemma 3.3 to estimate  $z(\tau_M(x, u))$ . Moreover (3.9) gives a bound for  $\int_0^{\tau_M(x,u)} \phi_M(x, s) ds$ . Adding that  $\tau_M(x, u) \leq 2 \log(1/x_M)/r_*$  (see Lemma 3.3 iii), we obtain for any  $x \in D(\eta)$ ,

$$\int_0^{\tau_M(x,u)} z(s) ds \leq C \left( \eta + u + u \int_0^{\tau_M(x,u)} z(s) ds \right).$$

Then for  $u \in (0, 1/(2C)]$ , for  $x \in D(\eta)$ ,

$$\int_0^{\tau_M(x,u)} z(s) ds \leq 2C(\eta + u).$$

We conclude using that  $\phi_R(x, s) - x_R^* = y_R(s) + x_R - x_R^*$ .

We can proceed similarly under the assumption  $\partial_R F_R(x^*) = 0$ , using now (3.8) and

$$\int_0^T \int_0^t z(s)^2 ds dt \leq \int_0^T z(t) \int_0^t z(s) ds dt \leq \left( \int_0^T z(t) dt \right)^2.$$

The proof is completed.  $\square$

The next lemma allows to lower-bound the growth rate of the mutant population on the time-windows we are interested in. It is used to quantify the hitting times  $\tau_M(x, v)$ .

**Lemma 3.5.** *The value of  $v_*$  is positive and finite. For all  $v \in (0, v_*)$ , there exist  $\eta \in (0, r_*)$  and  $r > 0$  such that, for all  $x \in B_r(x^*) \cap (\mathbb{R}_+^*)^2$  and  $t \leq \tau_M(x, v)$ ,  $F_M(\phi(x, t)) \geq \eta$ .*

*Proof.* Recall that  $r_\star = F_M(x^\star)$ . We consider the case  $\partial_R F_R(x^\star) < 0$  and the other case can be treated similarly. Thanks to Lemma 3.3, we can choose  $u_0 > 0$  such that, for all  $u \leq u_0$ , and  $x \in B_u(x^\star)$ ,

$$|\phi_R(x, \tau_M(x, u)) - x_R^\star| \leq \sup_{t \leq \tau_M(x, u)} |\phi_R(x, t) - x_R^\star| \leq Cu.$$

Then

$$\{\phi(x, t) : t \leq \tau_M(x, u)\} \subset ([x_R^\star - Cu, x_R^\star + Cu] \times [0, u]) \cap (\mathbb{R}_+^\star)^2.$$

Since  $F_M(x^\star) > 0$ , we can choose  $u_1 \in (0, u_0]$  such that

$$\eta := \inf \{F_M(\phi(x, t)) : t \leq \tau_M(x, u_1), x \in B_{u_1}(x^\star)\} > 0. \quad (3.11)$$

This ensures that  $v_\star \geq u_1 > 0$ .

Besides, starting close to the neighborhood of  $x^\star$ , the flow is bounded by Assumption 3.1. This ensures that  $v_\star$  is finite.

Note that (3.11) only proves the second statement of the lemma for arbitrary small values of  $v < v_\star$ . For the general case, consider some  $v < v_\star$ . Then there exists  $\eta \in (0, r_\star)$  such that  $v < v_\eta$ . Hence there exists some  $r > 0$  such that

$$v < \inf_{x \in B_r(x^\star) \cap (\mathbb{R}_+^\star)^2} \mathcal{C}(x, \eta).$$

Let  $x$  belong to  $B_r(x^\star) \cap (\mathbb{R}_+^\star)^2$ . Then

$$v < \mathcal{C}(x, \eta) = \sup\{\phi_M(x, t) : t \leq \bar{\tau}_M(x, \eta)\}$$

and we get  $F_M(\phi(x, t)) \geq \eta$  for  $t \leq \tau_M(x, v)$  by continuity of  $t \rightarrow F_M(\phi(x, t))$ .  $\square$

We can now prove Proposition 3.2.

*Proof of Proposition 3.2.* Let us consider the case  $\partial_R F_R(x_R^\star) < 0$ . We start by proving the result for small values of  $v < v_\star$ .

By respectively Lemmas 3.3 and 3.5, there exist some  $0 < u_1 < v_\star$  and  $\eta > 0$ , such that, for all  $u \leq u_1$  and  $x \in B_u(x^\star)$ ,

$$|\phi_R(x, \tau_M(x, u)) - x_R^\star| \leq \sup_{t \leq \tau_M(x, u)} |\phi_R(x, t) - x_R^\star| \leq Cu, \quad (3.12)$$

and for all  $t \leq \tau_M(x, u)$ ,

$$F_M(\phi(x, t)) \geq \eta.$$

Furthermore

$$\tau_M(x, \phi_M(x, t)) = t$$

and

$$\partial_t \phi_M(x, t) = \phi_M(x, t) F_M(\phi_R(x, \tau_M(x, \phi_M(x, t))), \phi_M(x, t)).$$

By separation of variables (put  $u = \phi_M(x, t)$ ) and integration on the time interval  $[0, \tau_M(x, v)]$ , we get

$$\tau_M(x, v) = \int_{x_M}^v \frac{1}{u F_M(\phi_R(x, \tau_M(x, u)), u)} du = \frac{1}{r_\star} \log(v/x_M) + \mathcal{R}_v(x),$$

for any  $v \in (0, u_1]$  and  $x \in B_{u_1}(x^\star)$  (with  $x_M \leq v$ ) and  $t \leq \tau_M(x, v)$ , where

$$\mathcal{R}_v(x) = \int_{x_M}^v \frac{1}{u} \left( \frac{1}{F_M(\phi_R(x, \tau_M(x, u)), u)} - \frac{1}{r_\star} \right) du.$$

To get the expected estimate, we need to evaluate the term inside the integral for  $u$  close to 0. Using that  $F_M$  is locally Lipschitz and  $r_\star = F_M(x^\star)$  and (3.12), and we obtain that for any  $u \in (0, u_1]$  and  $x \in B_{u_1}(x^\star)$  (with  $x_M \leq u$ ),

$$\frac{1}{u} \left( \frac{1}{F_M(\phi_R(x, \tau_M(x, u)), u)} - \frac{1}{r_\star} \right) \leq C.$$

The integral of the right hand side is convergent at 0 and we get

$$\lim_{v \rightarrow 0} \sup_{x \in B_{u_1}(x^\star) : x_M \leq v} \mathcal{R}_v(x) = 0.$$

Besides, for any  $u_2 \in (0, u_1]$ ,  $(x, u) \rightarrow F_M(\phi_R(x, \tau_M(x, u)), u)$  is uniformly continuous on  $\{(x, u) : x \in B_{u_1}(x^\star), x_M \leq u, u \in [u_2, u_1]\}$ . This can be shown by a linearization argument at the equilibrium  $x^\star$  using Hartman-Grobman's Theorem (see e.g. Section 3.2 of [23, p. 24] and particularly Theorem 3.2.1). Combining both facts and splitting the interval  $[x_m, v] = [x_m, u_2] \cup [u_2, v]$  yield

$$\sup \{ |\mathcal{R}_v(x) - \mathcal{R}_v(x')| : x, x' \in B_r(x^\star) \cap (\mathbb{R}_+^*)^2, v \in [\max(x_M, x'_M), u_1] \} \xrightarrow{r \rightarrow 0} 0.$$

The uniform Cauchy criterion ensures then that  $\mathcal{R}_v(x)$  converges, uniformly for  $v \in [v_0, v_1]$ , to a real number  $\tau(v)$  as  $x$  goes to  $x^\star$ . Therefore,

$$\tau(v) = \lim_{x \rightarrow x^\star} \int_{x_M}^v \frac{1}{u} \left( \frac{1}{F_M(\phi_R(x, \tau_M(x, u)), u)} - \frac{1}{r_\star} \right) du \quad (3.13)$$

and the previous estimates also guarantee that  $\tau$  is continuous on  $(0, u_1]$  and that  $\tau(v)$  goes to 0 as  $v$  goes to 0. This ends the proof for  $v \in (0, u_1]$ .

To deal with  $v \in [u_1, v_\star)$ , we split the time to reach  $v$  using the time to reach first  $u_1$ :

$$\tau_M(x, v) = \tau_M(x, u_1) + \tau_M(x, u_1, v),$$

where  $\tau_M(x, u_1, v)$  yields the time to go from  $u_1$  to  $v$ :

$$\tau_M(x, u_1, v) = \tau_M((\phi_R(x, \tau_M(x, u_1)), u_1), v) = \int_{u_1}^v \frac{1}{u} \frac{1}{F_M(\phi_R(x, \tau_M(x, u)), u)} du. \quad (3.14)$$

It remains to check that this term converges as  $x$  goes to  $x^\star$  (uniformly for  $u \in [u_1, u_3]$  where  $u_3 < v_\star$ ). This can be achieved as above using uniform continuity of  $(x, u) \in \{(x, u) : x \in B_{u_1}(x^\star), u \in [u_1, u_3]\} \rightarrow F_M(\phi_R(x, \tau_M(x, u)), u)$ . It ensures that the limit is continuous and increasing with respect to  $v$ . The conclusion follows from the fact that  $\log(v/x_M) = \log(v/u_1) + \log(u_1/x_M)$ .

The case  $\partial_R F_R(x^\star) = 0$  is treated in similar way. The linearization argument for proving uniform continuity is more subtle in that partially hyperbolic case and we can use [31].  $\square$

### 3.2. Quantitative approximation of the stochastic process

The main result can now be stated. We consider a large number of mutants but authorize this number to be small compared to initial resident population size. This makes the limiting process to be null on finite time intervals and we look at large time intervals  $\tau_M(x_K, v)$  and renormalized values to describe how the mutant density escapes from the absorbing and unstable boundary. Thus the initial density of mutants is small, and the initial density of residents is close to the equilibrium.

**Theorem 3.6.** *Under Assumptions 1.1 and 3.1, for all  $x_K = (x_R^K, x_M^K) \in (\mathbb{N}^*/K)^2$ , satisfying*

$$1/K \ll x_M^K \ll 1 \quad \text{and} \quad |x_R^K - x_R^\star| \log(1/x_M^K) \ll 1,$$

there exists  $\eta_0 > 0$  such that, for any  $v \in (0, v_*)$ ,  $\eta \in (0, \eta_0)$ , there exists some  $C > 0$  satisfying for all  $K \in \mathbb{N}^*$ ,

$$\mathbb{P} \left( \sup_{t \leq \tau_M(x_K, v)} \left| \frac{X_M^K(x_K, t)}{\phi_M(x_K, t)} - 1 \right| > \eta \right) \leq \frac{C}{\sqrt{K x_M^K}}.$$

Observe that our initial conditions are satisfied as soon as the initial number of mutants  $N_M^K(0)$  goes to infinity but remains negligible compared to  $K$  and the initial density of mutants  $X_R^K(0)$  is close to the equilibrium value up to a term of order of magnitude  $1/\log K$ . Our result stops at hitting times  $\tau_M(x_K, v)$  but could be easily extended on finite time intervals and with initial conditions of order one, following classical results of [27, 28].

In all this subsection, the hypotheses of Theorem 3.6 are in force, and we denote, for  $\bullet \in \{R, M\}$ ,

$$X_{\bullet}^K(t) := X_{\bullet}^K(x_K, t), \quad x_{\bullet}^K(t) := \phi_{\bullet}(x_K, t)$$

and

$$Y_{\bullet}^K(t) := X_{\bullet}^K(t) - x_{\bullet}^K(t).$$

In addition, we introduce the following stopping times: for  $\bullet \in \{R, M\}$ ,

$$\begin{aligned} \theta_{\bullet}^K &:= \inf \left\{ t > 0 : \frac{X_{\bullet}^K(t)}{x_{\bullet}^K(t)} \geq 2 \right\}, \quad \theta^K := \theta_R^K \wedge \theta_M^K, \\ \sigma_{\eta}^K &:= \inf \left\{ t > 0 : \int_0^t \sup_{r \leq s} |Y_R^K(r)| ds \geq \eta \right\} \text{ for } \eta > 0, \\ \tilde{\sigma}_{\varepsilon}^K &:= \inf \left\{ t > 0 : \sup_{s \leq t} |Y_R^K(s)| \geq \varepsilon \right\} \text{ for } \varepsilon > 0. \end{aligned}$$

In order to prove Theorem 3.6, we need the following result providing a control in probability when  $K$  goes to infinity of the stopping times  $\theta_R^K$ ,  $\sigma_{\eta}^K$  and  $\tilde{\sigma}_{\varepsilon}^K$ . This lemma guarantees the same property as Lemma 2.8 for the previous section: with probability going to one as  $K$  goes to infinity,  $\tau_M(x_K, L\eta) \wedge \theta_M^K$  is non-greater than the stopping times defined above. In other words, the stopping condition is reached by the mutant population and not by the resident one.

**Lemma 3.7.** *There exist  $L, \eta_0 > 0$  such that for all  $0 < \eta \leq \eta_0$ ,*

$$\mathbb{P}(\sigma_{\eta}^K \leq \tau_M(x_K, L\eta) \wedge \theta^K) \leq \frac{C}{\eta} K^{-1/2} (\log(1/x_M^K))^{3/2}, \quad (3.15)$$

$$\mathbb{P}(\tilde{\sigma}_{L\eta}^K \leq \tau_M(x_K, L\eta) \wedge \theta^K \wedge \sigma_{\eta}^K) \leq \frac{C}{\eta} K^{-1/2} (\log(1/x_M^K))^{1/2}, \quad (3.16)$$

$$\mathbb{P}(\theta_R^K \leq \tau_M(x_K, L\eta) \wedge \theta_M^K) \leq \frac{C}{\eta} K^{-1/2} (\log(1/x_M^K))^{3/2}. \quad (3.17)$$

*Proof.* To begin with, let us prove (3.15). For the sake of notation, let us denote

$$S^K := \tau_M(x_K, L\eta) \wedge \theta^K \wedge \sigma_{\eta}^K.$$

By definition of  $\sigma_{\eta}^K$ ,

$$\{\sigma_{\eta}^K \leq \tau_M(x_K, L\eta) \wedge \theta^K\} = \left\{ \int_0^{S^K} \sup_{r \leq s} |Y_R^K(r)| ds \geq \eta \right\}.$$

To control the probability of this event, we use that

$$Y_R^K(t) = H^K(t) + \int_0^t Y_R^K(s) \partial_R G_R(x^K(s)) ds \quad (3.18)$$

for all  $t > 0$ , where

$$\begin{aligned} H^K(t) &:= M^K(t) + R^K(t) + \int_0^t Y_M^K(s) \partial_M G_R(x^K(s)) ds, \\ M^K(t) &:= \frac{1}{K} \int_{[0,t] \times \mathbb{R}_+} \mathbf{1}_{\{z \leq K X_R^K(s-) b_R(X^K(s-))\}} \tilde{\mathcal{N}}_R^b(ds, dz) \\ &\quad - \frac{1}{K} \int_{[0,t] \times \mathbb{R}_+} \mathbf{1}_{\{z \leq K X_R^K(s-) d_R(X^K(s-))\}} \tilde{\mathcal{N}}_R^d(ds, dz), \\ R^K(t) &:= \int_0^t [G_R(X^K(s)) - G_R(x^K(s)) - Y_M^K(s) \partial_M G_R(x^K(s)) - Y_R^K(s) \partial_R G_R(x^K(s))] dr. \end{aligned}$$

In a first time, let us treat the case  $\partial_R G_R(x^*) = x_R^* \partial_R F_R(x^*) < 0$ . Thanks to Lemma 6.3 we have

$$\sup_{s \leq t} |Y_R^K(s)| \leq C \sup_{s \leq t} |H^K(s) - H^K(\lfloor s \rfloor)| \leq C \sup_{s \leq t} |H^K(t)|.$$

On the other hand, if  $\partial_R G_R(x^*) = x_R^* \partial_R F_R(x^*) = 0$ ,

$$\int_0^{\tau_M(x_K, L\eta)} |\partial_R G_R(x^K(r))| \leq C \int_0^{\tau_M(x_K, L\eta)} (x_M^K(r) + |x_R^K(r) - x_R^*|) dr \leq CL\eta,$$

using respectively Assumption 1.1 (R) and Lemma 3.4. Besides, by Gronwall's lemma,

$$|Y_R^K(t)| \leq \sup_{s \leq t} |H^K(s)| \exp\left(\int_0^t |\partial_R G_R(x^K(r))| dr\right).$$

Combining these two estimates implies that, for all  $t \leq \tau_M(x_K, L\eta) \wedge \theta^K$ ,

$$|Y_R^K(t)| \leq \sup_{s \leq t} |H^K(s)| e^{CL\eta} \leq C \sup_{s \leq t} |H^K(s)|,$$

for  $\eta \leq \eta_0$  fixed to some small enough value.

Gathering the two cases, we have proved that if  $\partial_R F_R(x^*) \leq 0$ , for  $t \leq \tau_M(x_K, L\eta) \wedge \theta^K$ ,

$$\begin{aligned} \sup_{s \leq t} |Y_R^K(s)| &\leq C_0 \sup_{s \leq t} |H^K(s)| \\ &\leq C_0 \sup_{s \leq t} |M^K(s)| + C_0 \sup_{s \leq t} |R^K(s)| + C_0 \int_0^t |Y_M^K(s)| \cdot |\partial_M G_R(x^K(s))| ds, \end{aligned} \quad (3.19)$$

for some constant  $C_0 > 0$  whose value is fixed.

In addition, for  $t \leq \tau_M(x_K, L\eta)$ ,

$$\begin{aligned} \int_0^t x_M^K(s) ds &= \int_0^t \frac{1}{F_M(x^K(s))} (x_M^K)'(s) ds \\ &\leq C \int_0^t (x_M^K)'(s) ds = C(x_M^K(t) - x_M^K(0)) \leq Cx_M^K(t). \end{aligned} \quad (3.20)$$

Using twice this inequality and  $|Y_M^K(r)| = x_M^K(r) |X_M^K(r)/x_M^K(r) - 1| \leq x_M^K(r)$  for  $r \leq \theta^K$ ,

$$\int_0^{S^K} \int_0^s |Y_M^K(r)| \cdot |\partial_M G_R(x^K(r))| dr ds \leq C \int_0^{S^K} \int_0^s x_M^K(r) dr ds \leq Cx_M^K(S^K) \leq CL\eta. \quad (3.21)$$

By Taylor-Lagrange's inequality and Lemma 3.3 (which guarantees that the process  $x^K$  is bounded up to time  $S^K$ ), we also have

$$\int_0^{S^K} \sup_{r \leq s} |R^K(r)| ds \leq C \int_0^{S^K} \int_0^s (|Y_M^K(r)|^2 + |Y_M^K(r)| \cdot |Y_R^K(r)| + |Y_R^K(r)|^2) dr ds.$$



Since  $S^K \leq \tau_M(x_K, L\eta)$ , and using the same computation as (3.21),

$$\int_0^{S^K} \int_0^s |Y_M^K(r)|^2 dr ds \leq L\eta \int_0^{S^K} \int_0^s |Y_M^K(r)| dr ds \leq C(L\eta)^2.$$

Besides  $|Y_M^K(r)| \cdot |Y_R^K(r)| \leq |Y_M^K(r)| \sup_{r' \leq s} |Y_R^K(r')|$  for  $r \leq s$ ,

$$\int_0^{S^K} \int_0^s |Y_M^K(r)| \cdot |Y_R^K(r)| dr ds \leq CL\eta \int_0^{S^K} \sup_{r' \leq s} |Y_R^K(r')| ds \leq CL\eta^2.$$

Similarly,  $\int_0^{S^K} |Y_R^K(r)|^2 dr ds \leq C\eta \int_0^{S^K} \sup_{r' \leq s} |Y_R^K(r')| ds \leq C\eta^2$ . Combining these estimates yields

$$\int_0^{S^K} \sup_{r \leq s} |R^K(r)| ds \leq C\eta^2. \quad (3.22)$$

Plugging (3.21) and (3.22) in (3.19), we can choose  $L > 0$  and  $\eta > 0$  small enough and get

$$\mathbb{P} \left( \int_0^{\tau_M(x_K, L\eta) \wedge \theta^K} |Y_R^K(s)| ds > \eta \right) \leq \mathbb{P} \left( \int_0^{\tau_M(x_K, L\eta) \wedge \theta^K} |M^K(s)| ds > \frac{\eta}{3C_0} \right).$$

Using again Doob and Cauchy-Schwarz inequalities,

$$\begin{aligned} \mathbb{E} \left[ \sup_{t \leq \tau_M(x_K, L\eta) \wedge \theta^K} |M^K(t)| \right] &\leq C \mathbb{E} \left[ \langle M^K \rangle_{\tau_M(x_K, L\eta) \wedge \theta^K} \right]^{1/2} \\ &\leq CK^{-1/2} \mathbb{E} \left[ \int_0^{\tau_M(x_K, L\eta) \wedge \theta^K} X_R^K(s) (b_R + d_R) (X^K(s)) ds \right]^{1/2} \\ &\leq CK^{-1/2} \tau_M(x_K, L\eta)^{1/2}. \end{aligned}$$

Using Markov's inequality and the comparison of  $\tau_M(x_K, L\eta)$  with  $\log(1/x_M^K)$  (cf Lemma 3.3.iii), we get

$$\mathbb{P} \left( \int_0^{\tau_M(x_K, L\eta) \wedge \theta^K} |Y_R^K(s)| ds > \eta \right) \leq \frac{C}{\eta} K^{-1/2} (\log(1/x_M^K))^{3/2}.$$

This proves (3.15).

Now we prove (3.16). By definition of  $\tilde{\sigma}_{L\eta}^K$ ,

$$\left\{ \tilde{\sigma}_{L\eta}^K \leq \tau_M(x_K, L\eta) \wedge \theta^K \wedge \sigma_\eta^K \right\} = \left\{ \sup_{s \leq S^K \wedge \tilde{\sigma}_{L\eta}^K} |Y_R^K(s)| > (L\eta)^{1/2} \right\}.$$

For the sake of readability, let us denote

$$T^K := S^K \wedge \tilde{\sigma}_{L\eta}^K.$$

To control the probability of the event above, we use (3.19) again. We have that

$$\begin{aligned} &\int_0^{T^K} |Y_M^K(s)| \cdot |\partial_M G_R(x^K(s))| ds \\ &\leq C \int_0^{T^K} x_M^K(s) \frac{|Y_M^K(s)|}{x_M^K(s)} ds \leq C \int_0^{\tau_M(x_K, L\eta)} x_M^K(s) ds \\ &\leq C \int_0^{\tau_M(x_K, L\eta)} \frac{1}{F_M(x^K(s))} (x_M^K)'(s) ds \leq C(x_M^K(\tau_M(x_K, L\eta)) - x_M^K(0)) \leq CL\eta, \end{aligned}$$

and

$$\begin{aligned} \sup_{s \leq T^K} |R^K(s)| &\leq C \int_0^{T^K} |X_M^K(s) - x_M^K(s)|^2 ds + C \int_0^{T^K} |Y_R^K(s)|^2 ds + C \int_0^{T^K} |Y_R^K(s)| \cdot |Y_M^K(s)| ds \\ &\leq C(L\eta)^2 + CL\eta^2 + C(L\eta)^2 \leq C\eta^2. \end{aligned}$$

Since  $\eta$  is arbitrary small, we have

$$\begin{aligned} \mathbb{P} \left( \sup_{s \leq T^K} |Y_R^K(s)| > L\eta \right) &\leq \mathbb{P} \left( \sup_{s \leq T^K} |M_s^K| > L\eta/4 \right) \\ &\leq C\eta^{-1} \mathbb{E} [\langle M^K \rangle_{T^K}]^{1/2} \leq C\eta^{-1} K^{-1/2} (\log(1/x_M^K))^{1/2}, \end{aligned}$$

which proves (3.16).

Finally, we prove (3.17) using (3.15) and (3.16). Notice that

$$\left\{ \theta_R^K \leq \tau_M(x_K, L\eta) \wedge \theta_M^K \right\} \subseteq \left\{ \sup_{t \leq \tau_M(x_K, L\eta) \wedge \theta^K} \frac{|Y_R^K(t)|}{x_r^K(t)} \geq 1 \right\}.$$

Then, by Lemma 3.3, we know that for all  $t \leq \tau_M(x_K, L\eta)$ , if  $\eta > 0$  is small enough,

$$x_R^K(t) \geq x_R^* - CL\eta \geq x_R^*/2.$$

Hence, for  $\eta > 0$  small enough,

$$\begin{aligned} \left\{ \theta_R^K \leq \tau_M(x_K, L\eta) \wedge \theta_M^K \right\} &\subseteq \left\{ \sup_{t \leq \tau_M(x_K, L\eta) \wedge \theta^K} |Y_R^K(t)| \geq x_R^*/2 \right\} \\ &\subseteq \left\{ \tilde{\sigma}_{L\eta}^K \leq \tau_M(x_K, L\eta) \wedge \theta^K \right\} \\ &\subseteq \left\{ \sigma_\eta^K \leq \tau_M(x_K, L\eta) \wedge \theta^K \right\} \cup \left\{ \tilde{\sigma}_{L\eta}^K \leq \tau_M(x_K, L\eta) \wedge \theta^K \wedge \sigma_\eta^K \right\}. \end{aligned}$$

Since the probabilities of the two last events above are controlled respectively by (3.15) and (3.16), the result is proved.  $\square$

We are now able to prove the main result of this section.

*Proof of Theorem 3.6.* We decompose the proof of Theorem 3.6 in three steps. The first step is the approximation of  $X_M^K$  by  $x_M^K$  until the level of the mutant population dynamics reaches some small macroscopic value. The second step is more classical and consists in showing that the approximation still holds true during an additional time length  $T > 0$ . In the last step, we show that within these time intervals the mutant population dynamics reached levels  $v_0 < v^*$ .

*Step 1.* In this first step, we prove that there exist  $C, L, \eta_0 > 0$  such that for any  $\eta \in (0, \eta_0)$ ,

$$\mathbb{P} \left( \sup_{t \leq \tau_M(x_K, L\eta)} \left| \frac{X_M^K(t)}{x_M^K(t)} - 1 \right| > \eta \right) \leq C \cdot \frac{1}{\eta} \cdot \frac{1}{\sqrt{Kx_M^K(0)}}. \quad (3.23)$$

We use that

$$\begin{aligned} \frac{X_M^K(t)}{x_M^K(t)} &= 1 - \int_0^t \frac{X_M^K(s)}{x_M^K(s)} (b_M - d_M)(x^K(s)) ds \\ &\quad + \int_{[0, t] \times \mathbb{R}_+} \frac{1}{Kx_M^K(s)} \mathbf{1}_{\{u \leq K \cdot X_M^K(s-) \cdot b_M(X^K(s-))\}} \mathcal{N}_M^b(ds, du) \\ &\quad - \int_{[0, t] \times \mathbb{R}_+} \frac{1}{Kx_M^K(s)} \mathbf{1}_{\{u \leq K \cdot X_M^K(s-) \cdot d_M(X^K(s-))\}} \mathcal{N}_M^d(ds, du). \end{aligned}$$

Writing  $\tilde{\mathcal{N}}_M^b$  and  $\tilde{\mathcal{N}}_M^d$  the compensated measure of  $\mathcal{N}_M^b$  and  $\mathcal{N}_M^d$ , and introducing the locally square integrable martingale

$$E_M^K(t) := \int_{[0,t] \times \mathbb{R}_+} \frac{1}{Kx_M^K(s)} \mathbf{1}_{\{u \leq K \cdot X_M^K(s-) \cdot b_M(X^K(s-))\}} \tilde{\mathcal{N}}_M^b(ds, du) \\ - \int_{[0,t] \times \mathbb{R}_+} \frac{1}{Kx_M^K(s)} \mathbf{1}_{\{u \leq K \cdot X_M^K(s-) \cdot d_M(X^K(s-))\}} \tilde{\mathcal{N}}_M^d(ds, du),$$

we have

$$\frac{X_M^K(t)}{x_M^K(t)} = 1 + E_M^K(t) + \int_0^t \frac{X_M^K(s)}{x_M^K(s)} (F_M(X^K(s)) - F_M(x^K(s))) ds.$$

Let us denote  $T^K = \tau_M(x_K, L\eta) \wedge \theta^K$  for convenience. We observe that

$$\langle E_M^K \rangle_{T^K} = \int_0^{T^K} \frac{1}{Kx_M^K(s)} \cdot \frac{X_M^K(s)}{x_M^K(s)} \cdot (b_M + d_M)(X^K(s)) ds \leq \frac{2C}{K} \int_0^{T^K} \frac{1}{x_M^K(s)} ds.$$

Consequently, for  $T \leq T^K$ ,

$$\sup_{t \leq T} \left| \frac{X_M^K(t)}{x_M^K(t)} - 1 \right| \\ \leq \sup_{t \leq T} |E_M^K(t)| + C \int_0^T (|X_M^K(s) - x_M^K(s)| + |X_R^K(s) - x_R^K(s)|) ds \\ \leq \sup_{t \leq T} |E_M^K(t)| + C \int_0^T x_M^K(s) \cdot \sup_{r \leq s} \left| \frac{X_M^K(r)}{x_M^K(r)} - 1 \right| ds + C \int_0^T |X_R^K(s) - x_R^K(s)| ds.$$

So, by Gronwall's lemma, for any  $T \leq \tau_M(x_K, L\eta) \wedge \theta^K$ ,

$$\sup_{t \leq T} \left| \frac{X_M^K(t)}{x_M^K(t)} - 1 \right| \leq \left( \sup_{t \leq T} |E_M^K(t)| + C \int_0^T |X_R^K(s) - x_R^K(s)| ds \right) e^{C \int_0^T x_M^K(s) ds}. \quad (3.24)$$

Besides, by Burkholder-Davis-Gundy's inequality,

$$\mathbb{E} \left[ \sup_{t \leq T^K} |E_M^K(t)| \right] \leq C \mathbb{E} \left[ \langle E_M^K \rangle_{T^K}^{1/2} \right] \leq CK^{-1/2} \mathbb{E} \left[ \int_0^{T^K} \frac{1}{x_M^K(s)} ds \right]^{1/2},$$

where the second inequality can be obtained either by Cauchy-Schwarz' inequality or Jensen's inequality.

Besides  $F_M(x^K(t))$  is lower-bounded on  $[0, \tau_M(x_K, L\eta)]$  by some  $\kappa > 0$  thanks to Lemma 3.3 since  $F_M(x^*) > 0$  if  $\eta$  is chosen small enough. Then  $x_M^K(s) \geq e^{\kappa s} x_M^K(0)$  and

$$\mathbb{E} \left[ \sup_{t \leq T^K} |E_M^K(t)| \right] \leq C \cdot \frac{1}{\sqrt{\kappa}} \cdot \frac{1}{\sqrt{Kx_M^K(0)}}, \quad (3.25)$$

where the value of  $C$  has changed in the last inequality, but is still independent of  $K$ .

Let us recall that, thanks to (3.20), we have

$$\int_0^{\tau_M(x_K, L\eta)} x_M^K(s) ds \leq CL\eta.$$

So, (3.24) can be rewritten as

$$\sup_{t \leq \tau_M(x_K, L\eta) \wedge \theta^K} \left| \frac{X_M^K(t)}{x_M^K(t)} - 1 \right| \leq C \sup_{t \leq \tau_M(x_K, L\eta) \wedge \theta^K} |E_M^K(t)| + C \int_0^{\tau_M(x_K, L\eta) \wedge \theta^K} |Y_R^K(s)| ds.$$

By Markov's inequality and (3.25),

$$\mathbb{P} \left( \sup_{t \leq T^K} |E_M^K(t)| > \frac{\eta}{2C} \right) \leq \frac{C}{\eta \sqrt{K x_M^K(0)}}.$$

Besides, by (3.15) from Lemma 3.7,

$$\mathbb{P} \left( \int_0^{T^K} |Y_R^K(s)| ds > \frac{\eta}{2C} \right) \leq \frac{C}{\eta} K^{-1/2} \left( \log \left( \frac{1}{x_M^K(0)} \right) \right)^{3/2}.$$

We obtain from the three last inequalities that

$$\mathbb{P} \left( \sup_{t \leq T^K} \left| \frac{X_M^K(t)}{x_M^K(t)} - 1 \right| > \eta \right) \leq \frac{C}{\eta} \cdot \frac{1}{\sqrt{K x_M^K(0)}}.$$

Using now that for  $\eta \in (0, 1)$ , we obtain

$$\left\{ \theta_M^K < \infty, \theta_M^K \leq \tau_M(x_K, L\eta) \right\} \subseteq \left\{ \sup_{t \leq T^K} \left| \frac{X_M^K(t)}{x_M^K(t)} - 1 \right| > \eta \right\},$$

and (3.17) proves (3.23).

*Step 2.* We fix  $T > 0$  and we prove now that the approximation of  $X_M^K$  by  $x_M^K$  is still valid on the time interval  $[\tau_M(x_K, L\eta), \tau_M(x_K, L\eta) + T]$ . The fact that the mutant population dynamics has left the neighborhood of 0 makes such estimates more classical. More precisely, we introduce the notation

$$\tilde{X}_\bullet^K(t) := X_\bullet^K(t + \tau_M(x_K, L\eta)), \quad \tilde{x}_\bullet^K(t) := x_\bullet^K(t + \tau_M(x_K, L\eta))$$

for  $\bullet \in \{R, M\}$ . First, we observe that

$$\gamma_{T,\eta} = \inf_{t \leq T, K \geq 1} \tilde{x}_M^K(t) > 0,$$

since the boundary  $\mathbb{R}_+ \times \{0\}$  is stable for the flow  $\phi$ , which is continuously differentiable, and the initial value  $\tilde{x}^K(0)$  belongs to a compact set of  $(\mathbb{R}_+^*)^2$  (thanks to Lemma 3.3)).

Then

$$\mathbb{P} \left( \sup_{t \leq T} \left| \frac{\tilde{X}_M^K(t) - \tilde{x}_M^K(t)}{\tilde{x}_M^K(t)} \right| > \eta \right) \leq \mathbb{P} \left( \sup_{t \leq T} |\tilde{X}_M^K(t) - \tilde{x}_M^K(t)| > \eta \gamma_{T,\eta} \right),$$

and it remains to prove that for any  $\eta > 0$ ,

$$\mathbb{P} \left( \sup_{t \leq T} |\tilde{X}_M^K(t) - \tilde{x}_M^K(t)| > \eta \right) \leq CK^{-1/2}. \quad (3.26)$$

This result is a classical approximation of population processes by dynamical systems on finite time intervals, see Ethier-Kurtz [22]. For sake of completeness and to give explicit bounds, let us prove the result here. We use again a stopping time which ensures boundedness :

$$\tilde{\theta}^K := \inf \left\{ t > 0 : \tilde{X}_M^K(t) > 2\tilde{x}_M^K(t) \text{ or } \tilde{X}_R^K(t) > 2\tilde{x}_R^K(t) \right\}.$$

We have, for any  $t \leq \tilde{\theta}^K$ ,

$$\left| \tilde{X}_\bullet^K(t) - \tilde{x}_\bullet^K(t) \right| \leq |M_\bullet^K(t)| + C \int_0^t \left( |\tilde{X}_R^K(s) - \tilde{x}_R^K(s)| + |\tilde{X}_M^K(s) - \tilde{x}_M^K(s)| \right) ds,$$

where  $M_R^K, M_M^K$  are locally square integrable martingales satisfying

$$\langle M_{\bullet}^K \rangle_{t \wedge \tilde{\theta}^K} \leq CtK^{-1}.$$

Then, Gronwall's lemma implies that

$$\sup_{t \leq T \wedge \tilde{\theta}^K} \left( \left| \tilde{X}_M^K(t) - \tilde{x}_M^K(t) \right| + \left| \tilde{X}_R^K(t) - \tilde{x}_R^K(t) \right| \right) \leq \left( \sup_{t \leq T \wedge \tilde{\theta}^K} |M_R^K(t)| + \sup_{t \leq T \wedge \tilde{\theta}^K} |M_M^K(t)| \right) e^{CT}.$$

Hence, using Markov's and Burkholder-Davis-Gundy's inequalities, we obtain

$$\mathbb{P} \left( \sup_{t \leq T \wedge \tilde{\theta}^K} \left( \left| \tilde{X}_M^K(t) - \tilde{x}_M^K(t) \right| + \left| \tilde{X}_R^K(t) - \tilde{x}_R^K(t) \right| \right) > \eta' \right) \leq \frac{1}{\eta'} C_T K^{-1/2}. \quad (3.27)$$

Finally, let us now dismiss the stopping time  $\tilde{\theta}^K$ . Let  $A_{\eta'}^K(T)$  be the event

$$A_{\eta'}^K(T) := \left\{ \sup_{t \leq T} \left( \left| \tilde{X}_M^K(t) - \tilde{x}_M^K(t) \right| + \left| \tilde{X}_R^K(t) - \tilde{x}_R^K(t) \right| \right) > \eta' \right\}.$$

Choosing  $\eta'$  smaller than

$$\inf_{t \leq T, K \geq 1} \tilde{x}_R^K(t) \wedge \inf_{t \leq T, K \geq 1} \tilde{x}_R^M(t),$$

guarantees that

$$A_{\eta'}^K(T) \subseteq \left\{ \tilde{\theta}^K \geq T \right\},$$

whence

$$\mathbb{P} \left( A_{\eta'}^K(T) \right) = \mathbb{P} \left( A_{\eta'}^K(T) \cap \{ \tilde{\theta}^K \geq T \} \right) \leq \mathbb{P} \left( A_{\eta'}^K(T \wedge \tilde{\theta}^K) \right).$$

Since (3.27) provides a control of  $\mathbb{P} \left( A_{\eta'}^K(T \wedge \tilde{\theta}^K) \right)$ , (3.26) is proved.

*Step 3.* We can now conclude the proof of the theorem using the two steps. Combining the two results yields the expect control until time  $\tau_M(x_K, L\eta) + T$ , for any  $\eta \in (0, \eta_0)$  and  $T > 0$ . We just need to check that  $T_v$  can be chosen (independently of  $K$  and  $x_K$ ) so that  $\tau_M(x_K, L\eta) + T_v \geq \tau_M(x_K, v)$ . This is indeed the case, as can be seen from (3.14) with  $u_1 = L\eta$  for instance.  $\square$

#### 4. Thresholds hitting times

In this section, we use the results of the two previous sections to approximate the hitting times of the process  $(N_M^K(t))_t$ . Let us recall that, for any  $n \geq 1$ , we have defined in (1.5) the stopping time

$$T_M^K(n_0, n) := \inf \{ t > 0 : N_M^K(t) \geq n \}.$$

The following main result of this section complements the results of [6, 7]. It gives a quantitative information on the law of hitting times of thresholds of order  $K$  but also for thresholds of size less than  $K$ . We recall that  $v_*$  is defined at the beginning of Section 3.1.

**Theorem 4.1.** *Grant Assumptions 1.1, 2.1 and 3.1. Let  $(\zeta_K)_K$  be some sequence which tends to infinity and satisfies  $\zeta_K/K \rightarrow v$  as  $K \rightarrow \infty$ , for some  $v \in [0, v_*)$ . Then, for any  $\varepsilon > 0$ ,*

$$\lim_{K \rightarrow \infty} \mathbb{P} \left( \left| T_M^K(N^K(0), \zeta_K) - \frac{\log(\zeta_K/W)}{r_*} - \tau(v) \right| \geq \varepsilon; W > 0 \right) = 0,$$

where  $\tau$  is the increasing continuous function such that  $\tau(0) = 0$  defined in (3.13) and

$$W = \lim_{t \rightarrow \infty} e^{-r_* t} Z(t) \stackrel{d}{=} \frac{d_*}{b_*} \delta_0 + \frac{r_*}{b_*} \mu_{r_*/b_*}, \quad \mu_a(dx) = 1_{R_+}(x) a e^{-ax} dx.$$

To prove this result, we first focus on the submacroscopic phase and use the branching process approximation to determine the asymptotic behavior of the hitting times of the mutant population process. For this purpose, we need the following lemma which guarantees the hitting time  $T_M^K(N^K(0), \xi_K)$  to go to infinity, for suitable sequence  $(\xi_K)$ .

**Lemma 4.2.** *Under Assumptions 1.1 and 2.1, for any sequence  $\xi$  tending to infinity and satisfying  $\xi_K \ll K/\log K$  if  $\partial_R F_R(x^*) < 0$  (resp.  $\xi_K \ll K/(\log K)^2$  if  $\partial_R F_R(x^*) = 0$ ), the hitting time  $T_M^K(N^K(0), \xi_K)$  goes to infinity almost surely as  $K$  goes to infinity.*

*Proof.* For the sake of notation, let us write  $T_M^K := T_M^K(N^K(0), \xi_K)$  in all this proof.

By Lemma 2.8, for any  $\varepsilon > 0$ , there exists some  $K_\varepsilon \in \mathbb{N}^*$  such that, for all  $K \geq K_\varepsilon$ ,

$$\mathbb{P}(\forall t \leq T_M^K, b_M(X^K(t)) \leq b_* + 1) \geq 1 - \varepsilon.$$

In particular, we can couple the process  $N_M^K$  with a branching process  $Z_+$  starting at  $Z_+(0) = 1$  with individual birth rate  $b_+ := b_* + 1$  and without death, such that, for all  $K \geq K_\varepsilon$ ,

$$\mathbb{P}(\forall t \leq T_M^K, N_M^K(t) \leq Z_+(t)) \geq 1 - \varepsilon.$$

It is known that  $Z_+(t)e^{-b_+t}$  is a local martingale converging to some finite r.v.  $W_+$  that is finite a.s. (see e.g. Theorem 1 of [1, p. 111] and Theorem 2 of [1, p. 112] in Section III.7). As a consequence  $\bar{W}_+ := \sup_{t \geq 0} Z_+(t)e^{-b_+t}$  is also finite a.s. Introducing  $T_Z^K$  the hitting time of the value  $\xi_K$  by the process  $Z_+$ , we have

$$\{N_M^K(T_M^K) \leq Z_+(T_M^K)\} \subseteq \{T_M^K \geq T_Z^K\}.$$

Besides, by definition of  $T_Z^K$ , on the event  $\{T_Z^K < \infty\}$ ,

$$\lceil \xi_K \rceil = Z(T_Z^K) \leq \bar{W}_+ e^{b_+ T_Z^K},$$

which implies

$$T_Z^K \geq \frac{1}{b_+} \log(\lceil \xi_K \rceil / \bar{W}_+).$$

Consequently, for all  $\varepsilon > 0$ ,

$$\mathbb{P}\left(T_M^K \xrightarrow{K \rightarrow \infty} +\infty\right) \geq 1 - \varepsilon.$$

Letting  $\varepsilon$  go to zero proves the result.  $\square$

The previous lemma allows us to obtain an asymptotic expansion for the hitting time  $T_M^K(N^K(0), \xi_K)$ . One can note that the statement of Proposition 4.3 is stronger than the one of Lemma 4.2.

**Proposition 4.3.** *Under Assumptions 1.1 and 2.1, for any  $\varepsilon > 0$ ,*

$$\lim_{K \rightarrow \infty} \mathbb{P}\left(\left|T_M^K(N^K(0), \xi_K) - \frac{\log(\xi_K/W)}{r_*}\right| \geq \varepsilon; W > 0\right) = 0$$

for any sequence  $\xi$  tending to infinity and satisfying  $\xi_K \ll K/\log K$  if  $\partial_R F_R(x^*) < 0$  (resp.  $\xi_K \ll K/(\log K)^2$  if  $\partial_R F_R(x^*) = 0$ ).

*Proof.* Consider the case  $\partial_R F_R(x^*) < 0$  and let  $(\xi_K)_K$  be a sequence tending to infinity and satisfying  $\xi_K \ll K/\log K$ . We apply Corollary 2.3. Then for  $\varepsilon > 0$ , there exists  $K_0$  such that for any  $K \geq K_0$ ,

$$\mathbb{P}\left(\forall t \in [0, T_M^K], (1 - \varepsilon)Z(t) \leq N_M^K(t) \leq (1 + \varepsilon)Z(t) \mid W > 0\right) \geq 1 - \varepsilon,$$

where we write  $T_M^K = T_M^K(N^K(0), \xi_K)$  for convenience. Since  $Z(t)\exp(-r_*t)$  converges to  $W$  almost surely as  $t$  tends to infinity,

$$\mathbb{P}\left(\sup_{t \geq u} \left|\frac{Z(t)}{\exp(r_*t)W} - 1\right| \geq \varepsilon \mid W > 0\right) \rightarrow 0$$

as  $u$  goes to infinity. Then there exists  $t_0$  such that

$$\mathbb{P}\left(\forall t \geq t_0, (1 - \varepsilon)W e^{r_* t} \leq Z(t) \leq (1 + \varepsilon)W e^{r_* t} \mid W > 0\right) \geq 1 - \varepsilon.$$

In addition, by Lemma 4.2, we know that  $\mathbb{P}(T_M^K \geq t_0 \mid W > 0) \geq 1 - \varepsilon$  for  $K$  large enough. Combining these estimates,

$$\mathbb{P}\left(\forall t \in [t_0, T_M^K], (1 - \varepsilon)^2 W e^{r_* t} \leq N_M^K(t) \leq (1 + \varepsilon)^2 W e^{r_* t} \mid W > 0\right) \geq 1 - 3\varepsilon.$$

Adding that  $\xi_K \leq N_M^K(T_M^K) \leq \xi_K + 1$  by definition of  $T_M^K$ , we obtain for  $t = T_M^K$

$$\mathbb{P}\left((1 - \varepsilon)^2 W e^{r_* T_M^K} \leq \xi_K + 1, \xi_K \leq (1 + \varepsilon)^2 W e^{r_* t} \mid W > 0\right) \geq 1 - 3\varepsilon.$$

Finally, taking the log inside the last probability yields the result.

The case where  $\partial_R F_R(x_R^*, 0) = 0$  can be treated in a similar way, and is therefore omitted.  $\square$

Now, in order to prove Theorem 4.1 from the results of the previous sections, we need to guarantee that, when two processes are close in some sense, their hitting times are also close. This property is stated and proved in the two next lemmas.

**Lemma 4.4.** *Let  $f, g : [0, t_0] \rightarrow \mathbb{R}$ . Assume that  $g$  is differentiable and*

$$\alpha = \inf_{[0, t_0]} g' > 0, \quad \beta = \sup_{t \in [0, t_0]} |f(t) - g(t)|.$$

*If  $\beta < g(t_0) - g(0)$ , then for any  $x \in [g(0), g(t_0) - \beta]$ ,  $\{t \in [0, t_0] : f(t) \geq x\}$  is non empty and*

$$|\inf\{t \in [0, t_0] : f(t) \geq x\} - \inf\{t \in [0, t_0] : g(t) \geq x\}| \leq \frac{\beta}{\alpha}.$$

*Proof.* Let us note in a first time that  $\{t \in [0, t_0] : f(t) \geq x\}$  is non empty since

$$f(t_0) = g(t_0) + (f(t_0) - g(t_0)) \geq g(t_0) - \beta \geq x.$$

Let  $t(x) \in [0, t_0]$  be the (unique) time when  $g(t(x)) = x$ . Then for  $t \in [0, t(x) - \varepsilon]$ ,

$$f(t) \leq g(t) + \beta \leq g(t(x)) - \varepsilon\alpha + \beta = x - \varepsilon\alpha + \beta < x$$

where the last inequality holds for any  $\varepsilon > \beta/\alpha$ . This forces  $\inf\{t \in [0, t_0] : f(t) \geq x\} \geq t(x) - \varepsilon$ . Similarly for  $t \in [t(x) + \varepsilon, t_0]$  and  $\varepsilon \geq \beta/\alpha$ ,

$$f(t) \geq g(t) - \beta \geq g(t(x)) + \varepsilon\alpha - \beta \geq x.$$

This yields  $\inf\{t \in [0, t_0] : f(t) \geq x\} \leq t(x) + \varepsilon$  and completes the proof since  $\varepsilon$  can be chosen arbitrarily close to  $\beta/\alpha$ .  $\square$

We can now capture the behavior of hitting times for  $X_M^K(t)$  when  $X_M^K(0)$  is large compared to  $1/K$  using the approximation by the dynamical system. For  $x \in (\mathbb{N}/K)^2$  and  $v \in \mathbb{R}_+$ , we consider the hitting times

$$\tau_M(x, v) = \inf\{t > 0 : x_M(t) \geq v\}, \quad S_M^K(x, v) = \inf\{t > 0 : X_M^K(t) \geq v\}, \quad (4.1)$$

where  $x(0) = X^K(0) = x$ .

**Lemma 4.5.** *Under Assumptions 1.1 and 3.1, assume*

$$1/K \ll X_M^K(0) \ll 1 \quad \text{and} \quad |X_R^K(0) - x_R^*| \log(1/X_M^K(0)) \ll 1$$

*in probability as  $K \rightarrow \infty$ . For any  $1 < c_1$  and  $v_0 < v_*$ ,*

$$\lim_{K \rightarrow \infty} \mathbb{P} \left( \sup_{c_1 X_M^K(0) \leq v \leq v_0} |S_M^K(X^K(0), v) - \tau_M(X^K(0), v)| \geq \varepsilon \right) = 0.$$

*Proof.* There exists  $C > 0$  such that for any  $\varepsilon \in (0, 1/2]$  and  $u \in [1 - \varepsilon, 1 + \varepsilon]$ ,  $|\log(u)| \leq C|u - 1| \leq C\varepsilon$ . Using Theorem 3.6, we have

$$\lim_{K \rightarrow \infty} \mathbb{P} \left( \sup_{t \leq \tau_M(X^K(0), v_0)} \left| \log \left( \frac{X_M^K(t)}{x_M^K(t)} \right) \right| \geq C\varepsilon \right) = 0.$$

We apply now Lemma 4.4 on the event

$$\left\{ \sup_{t \leq \tau_M(X^K(0), v_0)} \left| \log \left( \frac{X_M^K(t)}{x_M^K(t)} \right) \right| < C\varepsilon \right\}$$

with

$$f(t) = \log(X_M^K(t)), \quad g(t) = \log(x_M^K(t)),$$

which is possible by choosing  $\varepsilon$  small enough such that, for all  $c_1 X_M^K(0) \leq v \leq v_0$  (with the notation of Lemma 4.4)

$$\beta < C\varepsilon \leq \log c_1 \leq \log(v/X_M^K(0)) = g(t_0) - g(0).$$

By Lemma 3.5, there exist some  $\eta > 0$  and  $K_0 \in \mathbb{N}^*$  such that, for all  $K \geq K_0$ ,

$$\mathbb{P}(\forall t \leq \tau_M(X^K(0), v_0), F_M(x^K(t)) \geq \eta) \geq 1 - \varepsilon.$$

Then, noticing that  $g'(t) = F_M(x^K(t))$  and that  $|f(t) - g(t)| = |\log(X_M^K(t)/x_M^K(t))|$  allows to conclude.  $\square$

Finally, we end this section with the proof of the main result.

*Proof of Theorem 4.1.* To begin with, if  $\zeta_K \ll K/(\log K)^2$ , then  $v = 0$  and the result has been proved in Proposition 4.3 above.

Otherwise, let  $(\xi_K)_K$  be any sequence such that

$$\xi_K < \zeta_K \quad \text{and} \quad \xi_K \ll K/(\log K)^2,$$

and let  $v_K := \zeta_K/K$ . Writing  $\tilde{X}^K(t) := X^K(t + T_M^K(N^K(0), \xi_K))$ , and using the notation of (4.1),

$$T_M^K(N^K(0), \zeta_K) = T_M^K(N^K(0), \xi_K) + S_M^K(\tilde{X}^K(0), v_K).$$

Finally, since  $\tilde{X}^K(0)$  satisfies the assumption of Lemma 4.5 (cf Lemma 2.8), the result is a direct consequence of Propositions 3.2, 4.3 and Lemma 4.5.

Note that, since  $K/(\log K)^2 \ll K/(\log K)$ , we do not need to treat separately the cases where  $\partial_R F_R(x^*, 0)$  is null or negative in the proof of Theorem 4.1.  $\square$

In this part, we have obtained an approximation of hitting times which allows to quantify the stochastic effects. Developing more deeply our approach, we could study the speed of convergence in Theorems 3.6 and 4.1. This is left for a future work.



## 5. Examples : competitive and epidemiological models

### 5.1. Lotka-Volterra competitive model

We consider the dynamics of a two-dimensional population composed of resident and mutant individuals, which have their own demographic parameters and interact in a competitive way, for example in sharing resources or niche areas. This model is included in the framework of the paper, with constant individual birth rates

$$b_{\bullet}(X^K(s-)) = b_{\bullet}$$

and individual death rates of the Lotka-Volterra form :

$$d_R(X^K(s)) = d_R + c_{1,1}X_R^K(s) + c_{1,2}X_M^K(s),$$

$$d_M(X^K(s)) = d_M + c_{2,1}X_R^K(s) + c_{2,2}X_M^K(s),$$

with individual birth and death rates  $b_R, b_M \geq 0$  and  $d_R, d_M \geq 0$  and a competition matrix  $(c_{i,j})_{1 \leq i, j \leq 2}$  with non-negative coefficients. We assume that the coefficients  $b_R - d_R$ ,  $b_M - d_M$ ,  $c_{1,1}$  and  $c_{2,2}$  are positive.

The corresponding dynamical system is the competitive Lotka-Volterra system, that is the ODE (1.2) with

$$F_R(x_R, x_M) = b_R - d_R - c_{1,1}x_R - c_{1,2}x_M, \quad F_M(x_R, x_M) = b_M - d_M - c_{2,1}x_R - c_{2,2}x_M.$$

Then

$$x_R^* = \frac{b_R - d_R}{c_{1,1}}.$$

Let us check that the previous results can be applied. First, we observe that the regularity in Assumption 1.1 is satisfied. Second,  $(x_R^*, 0)$  is an equilibrium and

$$\partial_R F_R(x_R^*, 0) = -c_{1,1} < 0.$$

This means that Assumption 1.1.(E) is satisfied for  $(x_R^*, 0)$ . Moreover,  $x_R^*$  is a stable equilibrium for the resident population alone, which corresponds to the first case considered in the branching process approximation.

Finally, we start from a population with only one mutant,  $N_R^K(0) = K - 1, N_M^K(0) = 1$  and we consider  $x_R^* = 1$ . So the invasion condition  $F_M(x_R^*, 0) > 0$  reads  $b_M - d_M - (b_R - d_R)c_{2,1}/c_{1,1} > 0$ . It is then easy to compute  $v^*$ , which is

$$v^* = x_M^* = \frac{b_M - d_M}{c_{2,2}} \text{ if } F_R(0, x_M^*) < 0, \quad \text{or } v^* = \frac{(b_M - d_M)c_{1,1} - (b_R - d_R)c_{2,1}}{c_{1,1}c_{2,2} - c_{1,2}c_{2,1}} \text{ if } F_R(0, x_M^*) > 0.$$

In this framework, it has been proved in many papers, in particular in [21], [14], [17] that new phenomena can emerge from small subpopulations whose size is of order  $K^\alpha$ ,  $\alpha < 1$ .

Our results allow to characterize the size of populations on such scales. In particular, these results emphasize the fact that, on the scales  $K^\alpha$  ( $\alpha \in (0, 1)$ ), two approximations are valid in the following sense (cf respectively to Theorems 2.2 and 3.6):

**Proposition 5.1.** *For  $c > 0$ ,*

*i) starting from  $N^K(0) = (K - 1, 1)$ , for all  $\eta > 0$  and  $\alpha \in (0, 1)$ ,*

$$\mathbb{P} \left( \sup_{t \leq T_M^K(N^K(0), \lfloor cK^\alpha \rfloor)} \left| \frac{N_M^K(t)}{Z(t)} - 1 \right| > \eta ; \inf_{t \geq 0} Z(t) \geq 1 \right) \xrightarrow{K \rightarrow \infty} 0,$$

*where  $Z$  is a birth and death process with individual birth rate  $b_M$  and individual death rate  $d_M + c_{2,1}x_R^*$ .*

ii) For  $\alpha \in (0, 1]$  and  $\beta \in [0, 1)$ , assume that, almost surely,  $N^K(0)$  satisfies

$$1 \ll N_M^K(0) \leq cK^\alpha \text{ and } |N_R^K(0) - K| \leq K^\beta.$$

Then

$$\mathbb{P} \left( \sup_{t \leq T_M^K(N^K(0), \lfloor cK^\alpha \rfloor)} \left| \frac{N_M^K(t)}{Kx_M^K(t)} - 1 \right| > \eta \right) \xrightarrow{K \rightarrow \infty} 0,$$

where  $(x_R^K(t), x_M^K(t))_t$  is the solution of

$$\begin{cases} \frac{d}{dt}x_R^K(t) &= x_R^K(t)(b_R - d_R - c_{1,1}x_R^K(t) - c_{1,2}x_M^K(t)), \\ \frac{d}{dt}x_M^K(t) &= x_M^K(t)(b_M - d_M - c_{2,1}x_R^K(t) - c_{2,2}x_M^K(t)), \end{cases}$$

with  $(x_R^K(0), x_M^K(0)) = (N_R^K(0)/K, N_M^K(0)/K)$ .

In addition, it can be essential to know at which time a population initiated by a single mutant individual could attain such size. Our results answers to the question for all  $\alpha < 1$ . More precisely, Theorem 4.1 for  $\zeta_K = cK^\alpha$  yields

**Corollary 5.2.** For any  $\alpha \in (0, 1)$  and  $c > 0$

$$\lim_{K \rightarrow \infty} \mathbb{P} \left( \left| T_M^K(N^K(0), \lfloor cK^\alpha \rfloor) - \frac{\alpha \log(K) + \log(c/W)}{r_\star} \right| \geq \varepsilon; W > 0 \right) = 0,$$

Here  $v = 0$  and  $\tau(0) = 0$ .

## 5.2. SIR model

Let  $\beta > 0$  be the infection rate (per pair of individuals) in a mixed population and  $\gamma > 0$  the individual recovery rate. Let  $K \geq 1$  be the total population size. At time  $t$ , the number of susceptible individuals is denoted by  $S^K(t)$ , the number of infected individuals is denoted by  $I^K(t)$  and  $K - S^K(t) - I^K(t)$  yields the number of recovered individuals. Each susceptible becomes infected at time  $t$  with rate  $\beta I^K(t)/K$ . Starting from one infected individual and all other individuals susceptible, the process  $N^K = (N_R^K, N_M^K) = (S^K, I^K)$  is the unique strong solution of

$$\begin{aligned} S^K(t) &= K - 1 - \int_{[0,t] \times \mathbb{R}_+} \mathbf{1}_{\{u \leq \beta S^K(s-)I^K(s-)/K\}} \mathcal{N}_I(ds, du), \\ I^K(t) &= 1 + \int_{[0,t] \times \mathbb{R}_+} \mathbf{1}_{\{u \leq \beta S^K(s-)I^K(s-)/K\}} \mathcal{N}_I(ds, du) - \int_{[0,t] \times \mathbb{R}_+} \mathbf{1}_{\{u \leq \gamma I^K(s-)\}} \mathcal{N}_G(ds, du), \end{aligned}$$

where  $\mathcal{N}_I$  and  $\mathcal{N}_G$  are Poisson point measures on  $\mathbb{R}_+^2$  with intensity  $dsdu$ .

The process  $N^K$  is a bitype birth and death process as considered in the introduction with  $\mathcal{N}_I = \mathcal{N}_M^b = \mathcal{N}_R^d$  and birth and death rates defined by

$$\begin{aligned} b_R(x_R; x_M) &= 0, & d_R(x_R; x_M) &= \beta x_M \\ b_M(x_R; x_M) &= \beta x_R, & d_M(x_R; x_M) &= \gamma. \end{aligned}$$

The ODE describing the limit of  $N^K/K$  is given by (1.2) with

$$F_R(x_R, x_M) = -\beta x_M, \quad F_M(x_R, x_M) = \beta x_R - \gamma.$$

Let us check that our results can be applied to the case where the first derivative of  $F_R$  with respect to  $x_R$  cancels. First, we observe that regularity of Assumption 1.1 is satisfied. Second, for any  $x_R > 0$ ,  $(x_R, 0)$  is an equilibrium and

$$\partial_R F_R(x_R, 0) = 0.$$

This means that the second part of Assumption 1.1 is satisfied for any  $x_R > 0$ . Moreover, there is no geometric stability for the resident population alone, which corresponds to the second case considered in the branching process approximation.

Finally, we start from a population with only one mutant,  $S^K(0) = N_R^K(0) = K-1$ ,  $I^K(0) = N_M^K(0) = 1$  and we consider  $x_R^* = 1$ . So the invasion condition  $F_M(x_R^*, 0) > 0$  reads  $\beta > \gamma$ . The branching process  $Z$  is a birth and death process with individual birth rate  $\beta$  and individual death rate  $\gamma$ .

In this example, the maximal value of the invasive population is the peak of the epidemic and can be computed. Indeed, at the peak time, the derivative of  $x_I$  is zero, so  $x_S = \gamma/\beta$ . Adding that  $\gamma \log(x_S) + \beta(1 - x_S - x_I) = 0$ , we obtain for the peak value  $v_* = x_I$  the expression:

$$v_* = 1 - \frac{\gamma}{\beta} + \frac{\gamma}{\beta} \log\left(\frac{\gamma}{\beta}\right).$$

We obtain the following results when there is an outbreak of the epidemic. The hitting times for the infected population are defined by

$$\tau_I^K(n) = \inf\{t \geq 0, I^K(t) \geq n\}.$$

**Proposition 5.3.** *If  $\beta > \gamma$ , then*  
i) *for any  $\xi_K \ll K/(\log K)^2$  and  $\eta > 0$ ,*

$$\lim_{K \rightarrow +\infty} \mathbb{P}\left(\sup_{t \leq \tau_I^K(\xi_K)} \left| \frac{I^K(t)}{I(t)} - 1 \right| > \eta; I(\tau_I^K(\xi_K)) > 0\right) = 0,$$

where  $I$  is the unique strong solution of

$$I(t) = 1 + \int_{[0,t] \times \mathbb{R}_+} \mathbf{1}_{\{u \leq \beta I(s-)\}} \mathcal{N}_I(ds, du) - \int_{[0,t] \times \mathbb{R}_+} \mathbf{1}_{\{u \leq \gamma I(s-)\}} \mathcal{N}_G(ds, du).$$

ii) *For any  $v < v_*$  and  $\varepsilon > 0$ ,*

$$\lim_{T \rightarrow \infty} \limsup_{K \rightarrow \infty} \mathbb{P}\left(\sup_{t \in [T, \tau_I^K(vK)]} \left| \frac{I^K(t)}{K x_I^{K,T}(t)} - 1 \right| \geq \varepsilon; I^K(T) > 0\right) = 0,$$

where  $(x_S^{K,T}(t), x_I^{K,T}(t))$  is the solution of

$$\begin{cases} \frac{d}{dt} x_S^{K,T}(t) &= -\beta x_S^{K,T}(t) x_I^{K,T}(t), \\ \frac{d}{dt} x_I^{K,T}(t) &= (\beta x_S^{K,T}(t) - \gamma) x_I^{K,T}(t). \end{cases}$$

such that  $(x_S^{K,T}(0), x_I^{K,T}(0)) = (S^K(T)/K, I^K(T)/K)$ .

Point i) is a consequence of Theorem 2.2 and point ii) comes from Theorem 3.6. These two parts have an intersection : in the time window when the number of infected individuals is large but negligible compared to  $K/(\log K)^2$ , we have informally

$$I^K(t) \sim I(t) \sim W e^{(\beta-\gamma)t} \sim K x_I^{K,T}(t-T).$$

We also obtain an approximation of the time when the epidemics reaches a given level (by Theorem 4.1):

**Corollary 5.4.** *Let  $(\zeta_K)_K$  be sequence which tends to infinity and satisfies  $\zeta_K/K \rightarrow v$  as  $K \rightarrow \infty$ , for some  $v \in [0, v_*)$ . Then, for any  $\varepsilon > 0$ ,*

$$\lim_{K \rightarrow \infty} \mathbb{P}\left(\left| \tau_I^K(\zeta_K) - \frac{\log(\zeta_K/W)}{r_*} - \tau(v) \right| \geq \varepsilon, W > 0\right) = 0,$$

where  $\tau$  is the increasing continuous function such that  $\tau(0) = 0$  defined in Proposition 3.2.

This question of how long it takes for an epidemic to reach a high level has several motivations. Indeed, sanitary capacities for managing the infected individuals (hospital beds, isolation...) are limited. The time at which sanitary capacities reach saturation is an important factor in managing epidemics and taking control measures (social distancing, confinement, etc.). In addition, we control here the period during which the stochasticity of the process still plays a role, before deterministic predictions are valid with sufficient accuracy.

## 6. Appendix

### 6.1. Minimum of the martingale associated with the branching process $Z$

Recalling the notation of Section 2, we are interested in the minimal value of the martingale  $W(t) := Z(t) \exp(-r_* t)$ . Such object for simple branching processes has already attracted a lot of attention. We are not aware in the literature of the estimates given here and provide the proof for completeness. This relies on classical  $L^2$  estimates to control the speed of the convergence of the martingale.

**Lemma 6.1.** *For any  $\varepsilon > 0$ , and any  $Z$ -adapted stopping time  $\tau$ ,*

$$\mathbb{P}(W(\tau) \leq \varepsilon; W > 0) \leq C \left( \varepsilon + \varepsilon^{-1} \mathbb{E} \left[ \mathbf{1}_{\{\tau < \infty\}} e^{-r_* \tau / 2} \right] \right),$$

where  $W$  is the almost sure limit of  $W(t)$  as  $t$  goes to infinity, and  $C$  is some positive constant independent of  $\varepsilon$  and  $\tau$ . We use the convention  $W(+\infty) := W$  for the inequality above to make sense.

*Proof.* To simplify the proof, we firstly treat the case where  $\tau$  is deterministic and finite. To begin with,

$$W(\tau + 1) - W(\tau) = e^{-r_* \tau} \left( \sum_{i=1}^{Z(\tau)} N_i e^{-r_*} - Z(\tau) \right),$$

where, conditionally on  $Z(\tau)$ , the  $N_i$  are i.i.d. with the same distribution as  $Z(1)$ . This implies that

$$\begin{aligned} \mathbb{E} \left[ (W(\tau + 1) - W(\tau))^2 \right] &= e^{-2r_* \tau} \mathbb{E} \left[ \text{Var} \left[ \sum_{i=1}^{Z(\tau)} (N_i e^{-r_*} - 1) \mid Z(\tau) \right] \right] \\ &= e^{-2r_* \tau} \mathbb{E} [Z(\tau)] \text{Var} [N_i e^{-r_*} - 1] \leq C e^{-r_* \tau}. \end{aligned} \quad (6.1)$$

On the other hand, we write

$$\begin{aligned} \mathbb{P}(W(\tau) \leq \varepsilon; W > 0) &\leq \mathbb{P}(W \leq 2\varepsilon; W > 0) + \mathbb{P}(W(\tau) \leq \varepsilon; W > 2\varepsilon) \\ &\leq \mathbb{P}(W \leq 2\varepsilon; W > 0) + \mathbb{P}(W - W(\tau) > \varepsilon). \end{aligned}$$

Recalling that, on the event  $\{W > 0\}$ ,  $W$  follows an exponential distribution, we can bound the first term of the sum above by  $1 - e^{-C\varepsilon} \leq C\varepsilon$ .

The second term of the sum can be handled as follows:

$$\begin{aligned} \mathbb{P}(W - W(\tau) > \varepsilon) &= \mathbb{P} \left( \sum_{n \geq 0} W(\tau + n + 1) - W(\tau + n) > \varepsilon \right) \\ &\leq \frac{1}{\varepsilon} \sum_{n \geq 0} \mathbb{E} [|W(\tau + n + 1) - W(\tau + n)|] \\ &\leq \varepsilon^{-1} \sum_{n \geq 0} \mathbb{E} \left[ |W(\tau + n + 1) - W(\tau + n)|^2 \right]^{1/2} \\ &\leq C \varepsilon^{-1} e^{-r_* \tau / 2} \sum_{n \geq 0} e^{-r_* n / 2} \leq C \varepsilon^{-1} e^{-r_* \tau / 2}. \end{aligned}$$

This ends the proof in the case where  $\tau$  is deterministic.

Now to prove the result when  $\tau$  is some almost surely finite  $Z$ -adapted stopping time, we just have to do the same computation conditionally on  $\mathcal{F}_\tau$ . Indeed, since  $Z(\tau)$  is  $\mathcal{F}_\tau$ -measurable and  $Z(\tau + 1) - Z(\tau)$  independent of  $\mathcal{F}(\tau)$ , (6.1) becomes

$$\mathbb{E} \left[ (W(\tau + 1) - W(\tau))^2 | \mathcal{F}_\tau \right] \leq C e^{-r_* \tau},$$

and the last computation of the previous case gives

$$\mathbb{P}(W - W(\tau) > \varepsilon | \mathcal{F}_\tau) \leq C \varepsilon^{-1} e^{-r_* \tau / 2}.$$

The result of the lemma is then proved by taking the expectation in the inequality above.

Finally let us treat the case where  $\tau$  is not necessary almost surely finite. We write

$$\begin{aligned} \mathbb{P}(W(\tau) \leq \varepsilon; W > 0) &= \mathbb{P}(W(\tau) \leq \varepsilon; W > 0; \tau < \infty) + \mathbb{P}(W(\tau) \leq \varepsilon; W > 0; \tau = \infty) \\ &\leq \mathbb{P}(W(\tau) \leq \varepsilon; W > 0; \tau < \infty) + \mathbb{P}(0 < W \leq \varepsilon). \end{aligned} \quad (6.2)$$

Note that, on the event  $\{\tau < \infty\}$ , we have the almost sure convergence

$$\mathbf{1}_{\{W(\tau \wedge n) \leq \varepsilon; W > 0\}} \xrightarrow[n \rightarrow \infty]{} \mathbf{1}_{\{W(\tau) \leq \varepsilon; W > 0\}}.$$

Hence, by Fatou's lemma,

$$\begin{aligned} \mathbb{P}(W(\tau) \leq \varepsilon; W > 0; \tau < \infty) &\leq \liminf_n \mathbb{P}(W(\tau \wedge n) \leq \varepsilon; W > 0; \tau < \infty) \\ &\leq \liminf_n \mathbb{P}(W(\tau \wedge n) \leq \varepsilon; W > 0). \end{aligned}$$

Then, applying the result of the lemma with the almost surely finite stopping time  $\tau \wedge n$ , we obtain

$$\mathbb{P}(W(\tau) \leq \varepsilon; W > 0; \tau < \infty) \leq C \left( \varepsilon + \varepsilon^{-1} \liminf_n \mathbb{E} \left[ e^{-r_* (\tau \wedge n) / 2} \right] \right).$$

Noticing that

$$e^{-r_* (\tau \wedge n) / 2} \leq \mathbf{1}_{\{\tau < \infty\}} e^{-r_* \tau / 2} + e^{-r_* n / 2},$$

we have

$$\mathbb{P}(W(\tau) \leq \varepsilon; W > 0; \tau < \infty) \leq C \left( \varepsilon + \varepsilon^{-1} \mathbb{E} \left[ \mathbf{1}_{\{\tau < \infty\}} e^{-r_* \tau / 2} \right] \right).$$

In addition, to control the second term of the sum in (6.2), we recall that the distribution of  $W$  is

$$\frac{d_*}{b_*} \delta_0 + \frac{r_*}{b_*} \mathcal{E}\left(\frac{r_*}{b_*}\right).$$

Hence

$$\mathbb{P}(0 < W \leq \varepsilon) = \frac{r_*}{b_*} \left( 1 - e^{-\frac{r_*}{b_*} \varepsilon} \right) \leq \left( \frac{r_*}{b_*} \right)^2 \varepsilon.$$

Recalling (6.2), the result of the lemma is proved.  $\square$

**Lemma 6.2.** *With the same notation as in Lemma 6.1, for any  $0 < \varepsilon < 1$ ,*

$$\mathbb{P} \left( \inf_{t \geq 0} W(t) \leq \varepsilon; W > 0 \right) \leq C \varepsilon^{1/4},$$

for some positive constant  $C$  independent of  $\varepsilon$ .

*Proof.* Let us introduce

$$s_\varepsilon := \frac{1}{r_\star} \log(1/\varepsilon).$$

On the event  $\{W > 0\}$ , we have  $Z(t) \geq 1$  for all  $t \geq 0$  and further, for  $t < s_\varepsilon$ ,  $e^{-r_\star t} > \varepsilon$ . As a consequence

$$\left\{ \inf_{t \geq 0} W(t) \leq \varepsilon; W > 0 \right\} \subseteq \left\{ \inf_{t \geq s_\varepsilon} W(t) \leq \varepsilon; W > 0 \right\}.$$

For some  $\lambda > 1$  whose value will be fixed later, let us denote

$$S_\varepsilon := \inf \left\{ t \geq s_\varepsilon : W(t) \leq \varepsilon^{1/\lambda} \right\}.$$

Since  $\lambda > 1$  and  $0 < \varepsilon < 1$ , we have  $\varepsilon < \varepsilon^{1/\lambda}$ , hence

$$\left\{ \inf_{t \geq s_\varepsilon} W(t) \leq \varepsilon \right\} \subseteq \{S_\varepsilon < \infty\} \subseteq \{W(S_\varepsilon) \leq \varepsilon^{1/\lambda}\}.$$

Note that the last event above is trivially satisfied only if  $S_\varepsilon < \infty$ .

Then, thanks to Lemma 6.1,

$$\mathbb{P} \left( W(S_\varepsilon) \leq \varepsilon^{1/\lambda}; W > 0 \right) \leq C \left( \varepsilon^{1/\lambda} + \varepsilon^{-1/\lambda} \mathbb{E} \left[ \mathbf{1}_{\{S_\varepsilon < \infty\}} e^{-r_\star S_\varepsilon/2} \right] \right).$$

Recall that, by definition of  $S_\varepsilon$ ,  $S_\varepsilon \geq s_\varepsilon$ , and so

$$\mathbb{E} \left[ \mathbf{1}_{\{S_\varepsilon < \infty\}} e^{-r_\star S_\varepsilon/2} \right] \leq e^{-r_\star s_\varepsilon/2} = \varepsilon^{1/2},$$

implying that

$$\mathbb{P} \left( W(S_\varepsilon) \leq \varepsilon^{1/\lambda}; W > 0 \right) \leq C \left( \varepsilon^{1/\lambda} + \varepsilon^{1/2-1/\lambda} \right).$$

Finally, choosing  $\lambda := 4$  to optimize the bound above proves the result.  $\square$

## 6.2. Perturbation of stable dynamical system

We provide a result which is strongly inspired from Lemmas 3.1 and 3.2 in [30]. We consider an exponential stable dynamical system :  $y' \leq -cy$ , with  $c > 0$ . We add a source term, which acts as a perturbation given by  $h$ . The following result allows to control the value of the dynamical system in terms of fluctuations of the source term  $h$  on a unit time interval.

**Lemma 6.3.** *Assume that there are three measurable and locally bounded functions  $y, h, \phi$  with  $h(0) = 0$  and satisfying for all  $t \geq 0$ ,*

$$y(t) = y(0) + h(t) + \int_0^t y(s)\phi(s)ds.$$

*Suppose in addition that  $\phi$  is bounded, and upper-bounded by some negative number. Then, for all  $t \geq 0$ ,*

$$\sup_{s \leq t} |y(s)| \leq \Gamma \left( |y(0)| \vee \sup_{s \leq t} |h(s) - h(\lfloor s \rfloor)| \right),$$

*where  $\Gamma = (1 + \|\phi\|_\infty)/(1 - \exp(\sup \phi))$ .*

*Proof.* To begin with, we introduce

$$\Phi(t, s) = e^{\int_s^t \phi(r) dr} \leq e^{-C(t-s)},$$

for some positive constant  $C := -\sup \phi > 0$ . Then

$$\begin{aligned} \Phi(t, t) &= 1 ; \quad \Phi(t, s)\Phi(s, u) = \Phi(t, u), \\ \partial_t \Phi(t, s) &= \Phi(t, s)\phi(t) ; \quad \partial_s \Phi(t, s) = -\Phi(t, s)\phi(s). \end{aligned}$$

By a standard constant variation argument, one can write

$$y(t) = \Phi(t, 0)y(0) + h(t) + \int_0^t \Phi(t, s)h(s)\phi(s)ds.$$

Then it results the following decomposition:

$$\begin{aligned} y(t) &= \Phi(t, 0)y(0) + \sum_{j=1}^{\lfloor t \rfloor} \Phi(t, j) \left( h(j) - h(j-1) + \int_{j-1}^j \Phi(j, s)\phi(s) (h(s) - h(j-1)) ds \right) \\ &\quad + h(t) - h(\lfloor t \rfloor) + \int_{\lfloor t \rfloor}^t \Phi(t, s)\phi(s) (h(s) - h(\lfloor t \rfloor)) ds. \end{aligned}$$

Using the upperbound of  $\phi$ , we obtain

$$|y(t)| \leq e^{-Ct}|y(0)| + \sum_{j=1}^{\lfloor t \rfloor} e^{-C(t-j)} \left( \sup_{j-1 \leq s \leq j} |h(s) - h(j-1)| \Gamma' \right) + \sup_{\lfloor t \rfloor \leq s \leq t} |h(s) - h(\lfloor t \rfloor)| \Gamma',$$

with  $\Gamma' := 1 + \|\phi\|_\infty$ . The result is proved by choosing  $\Gamma = \Gamma' \sum_{j=0}^{\infty} e^{-Cj} = \Gamma'/(1 - e^{-C})$ .  $\square$

### Acknowledgements:

- The authors are very grateful to Pierre Collet who has clarified some issue involved in the control of hitting values of the dynamical system started close to the equilibrium. They also thank very warmly the two referees of the first version of this work, whose many comments led to a major improvement of the paper.
- This work has been supported by the Chair “Modélisation Mathématique et Biodiversité of Veolia Environnement-Ecole Polytechnique-Museum national d’Histoire naturelle-Fondation X and by ITMO Cancer of Aviesan within the framework of the 2021-2023 Cancer Control Strategy, on funds administrated by Inserm.
- Funded by the European Union (ERC, SINGER, 101054787). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.

### References

- [1] K. B. Athreya, P. E. Ney. *Branching processes*. Springer-Verlag Berlin Heidelberg New-York 1972.
- [2] M. Baar, A. Bovier, N. Champagnat. From stochastic, individual-based models to the canonical equation of adaptive dynamics in one step. *Ann. Appl. Probab.* 27 (2017), no. 2.
- [3] H. Bahouri, J. Chemin and R. Danchin. *Fourier Analysis and Nonlinear Partial Differential Equations* Springer, Grundlehren der mathematischen Wissenschaften, 2011.

- [4] F. Ball, P. Donnelly. Strong approximations for epidemic models. *Stoc. Proc. Appl.* Vol. 55, 1995, Pages 1-21.
- [5] A.D. Barbour, S. Utev. Approximating the Reed-Frost epidemic process. *Stoc. Proc. Appl.* Vol. 113 (2004) p 173-197.
- [6] A. D. Barbour, G. Reinert. Approximating the epidemic curve. *Electron. J. Probab.* 18 (2013), no. 54, 1-30.
- [7] A. D. Barbour, K. Hamza, H. Kaspi and F. C. Klebaner. Escape from the boundary in Markov population processes. *Adv. Appl. Probab.* (2015), Vol. 47, No. 4, pp. 1190-1211.
- [8] S. Billiard, R. Ferrière, S. Méléard, V.C. Tran. Stochastic dynamics of adaptive trait and neutral marker driven by eco-evolutionary feedbacks. *J. Math. Biol.* 71 (2015), no. 5.
- [9] S. Billiard, C. Smadi. The interplay of two mutations in a population of varying size: a stochastic eco-evolutionary model for clonal interference. *Stochastic Process. Appl.* 127 (2017), no. 3, 701-748.
- [10] J. Blath, T. Paul and A. Tóbiás. A Stochastic Adaptive Dynamics Model for Bacterial Populations with Mutation, Dormancy and Transfer arXiv:2105.09228 (2021).
- [11] C. Bonnet, P. Gou, S. Girel, V. Bansaye, C. Lacout, K. Bailly, M.H. Schlagger, E. Lauret, S. Méléard, S. Giraudier. Multistage hematopoietic stem cell regulation in the mouse: a combined biological and mathematical approach. *iScience*, 2021 Nov 6; 24 (12):103399. Open access.
- [12] C. Bonnet, S. Méléard. Large fluctuations in multi-scale modeling for rest erythropoiesis. *J. Math. Biol.* 82 (2021), no. 6, Paper No. 58.
- [13] A. Bovier, L. Coquille, R. Neukirch. The recovery of a recessive allele in a Mendelian diploid model. *J. Math. Biol.* 77 (2018), no. 4.
- [14] A. Bovier, L. Coquille, C. Smadi. Crossing a fitness valley as a metastable transition in a stochastic population model. *Ann. Appl. Probab.* 29 no 6 (2019), 3541–3589.
- [15] N. Champagnat. A microscopic interpretation for adaptative dynamics trait substitution sequence models. *Stochastic Processes and their Applications*, 116 (2006), 1127–1160.
- [16] N. Champagnat and S. Méléard. Polymorphic evolution sequence and evolutionary branching. *Probability Theory and Related Fields*, 151 no 1-2 (2011), 45–94.
- [17] N. Champagnat, S. Méléard, S. Mirrahimi, V.C. Tran. Stochastic analysis of emergence of evolutionary cyclic behavior in population dynamics with transfer. *Ann. Appl. Probab.* 31 (2021), no.4, p. 1820–1867.
- [18] L. Coquille, A. Kraut, C. Smadi. Stochastic individual-based models with power law mutation rate on a general finite trait space. *Electron. J. Probab.* 26(123), 37 p., 2021.
- [19] O. Diekmann, P.-E. Jabin, S. Mischler, and B. Perthame. The dynamics of adaptation: an illuminating example and a Hamilton-Jacobi approach. *Theoretical Population Biology*, 67 ( 2005), 257–271.
- [20] R. Durrett. Branching process models of Cancer. Mathematical Biosciences Institute Lecture Series 1.1, Stochastics in Biological Systems, Springer, 2015.
- [21] R.Durrett, J. Mayberry. Traveling waves of selective sweeps. *Ann. Appl. Probab.* 21 (2011), no. 2, p. 699–744.
- [22] S.N. Ethier, T.G. Kurtz. Markov processes, Characterization and Convergence. John Wiley & Sons, New York, 1986.
- [23] J. Hofbauer and K. Sigmund. *Evolutionary Games and Population Dynamics*. Cambridge University Press 1998.
- [24] J. Jacod and A.N. Shiryaev. *Limit Theorems for Stochastic Processes*. Springer-Verlag, Berlin, 1987.
- [25] A. Jakubowski. On the Skorokhod topology. *Annales de l'Institut Henri Poincaré*, 22 No 3, 263-285, 1986.
- [26] A. Joffe and M. Métivier. Weak convergence of sequences of semimartingales with applications to multitype branching processes. *Advances in Applied Probability*, 18 (1986), 20–65.
- [27] T. G. Kurtz. Solutions of ordinary differential equations as limits of pure jump Markov processes. *J. Appl. Prob.* 7,49-58, 1970.
- [28] T. G. Kurtz. Strong approximation theorems. *Stoch. Proc. Appl.* 6, 223-240, 1978.



- [29] J.A.J. Metz, S.A.H. Geritz, G. Meszéna, F.A.J. Jacobs, and J.S. Van Heerwaarden. Adaptative dynamics, a geometrical study of the consequences of nearly faithful reproduction. *S.J. Van Strien & S.M. Verduyn Lunel (ed.), Stochastic and Spatial Structures of Dynamical Systems*, 45, 183–231, 1996.
- [30] A. Prodhomme. Strong Gaussian approximation of metastable density-dependent Markov chains on large time scales. arXiv:2010.06861 (2020).
- [31] F. Takens. Partially hyperbolic fixed points. *Topology*, Vol 10, 133-147, section 4. (1971)