Sharp approximation and hitting times for stochastic invasion processes

Vincent Bansaye, Xavier Erny, Sylvie Méléard

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*

Abstract: We are interested in the invasion phase for stochastic processes with interactions. A single mutant with positive fitness arrives in a 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 in the different phases. It allows us to describe the fluctuations of the hitting times of thresholds, which inherit a large variance from the first stage. These issues are in particular motivated by the quantification of the hitting times of critical values for cancer cell populations or 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 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 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 are 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 solutions 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 $\mathbb{R}_+ \times \mathbb{R}_+$:

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

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. 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

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

It is well known (cf [24, 25, 20]) that under suitable assumptions on the parameters and when initial conditions of the branching processes are of order K, the stochastic process $((X_R^K(t), X_M^K(t)) : t \geq 0)$ converges in probability (when K tends to infinity), on any finite time interval. The limit is the solution (x_R, x_M) of the dynamical system

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

We assume the existence of some (non-necessary unique) equilibrium $x_R^\star > 0$ satisfying

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

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

$$N^K(0) = (N_R^K(0), 1), \qquad N_R^K(0) \sim x_R^\star K. \tag{1.3}$$

Thus $X^K(0)$ is close to $(x_R^\star, 0)$ 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

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

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

$$r_\star = b_\star - d_\star > 0. \tag{1.4}$$

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

We consider the hitting times of successive levels n of the mutant population process when the initial condition is $N^K(0) = n_0 \in \mathbb{N}^2$:

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

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 mutation process is large but negligible compared to K.

Our first result proves that the mutation 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

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

where

$$W = \lim_{t\to\infty} Z(t)e^{-r_\star t} \in [0,\infty) \qquad a.s.$$

is an exponential variable with an additional atom in 0 corresponding to the extinction event. 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 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 $K^{2/3}$ 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 sub-macroscopic levels, more precisely before the mutant population reaches the order $K/(\log K)^a$.

When the population of mutants becomes large but is not yet macroscopic, we prove that the deterministic approximation by the dynamical system is already valid in the following sense:

$$\sup_{t\le T_M^K(N^K(0),vK)} \left| \frac{X_M^K(t)}{x_M^K(t)} - 1 \right| \underset{K\to\infty}{\longrightarrow} 0, \text{ in probability,}$$

where $v > 0$, $1 \ll N_M^K(0) = KX_M^K(0) \ll K$ is the initial number of mutant and x_M^K is the solution of (1.2) with the corresponding initial condition $1/K \ll x_M^K(0) = X_M^K(0) \ll 1$ and $x_R^K(0)$ close to $x_R^\star$. The boundary of $(\mathbb{R}_+)^2$ is invariant for the dynamical system but the latter is unstable around $\mathbb{R}_+ \times \{0\}$ 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. We refer to Section 3.2 for a precise statement which gives the admissible values of v and quantitative estimates for the convergence. This yields the main novelty of this work.

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).

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 $\zeta_K/K \to v \ge 0$, we show the following convergence in law when $K \to \infty$

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

where the function τ is continuous, $\tau(0) = 0$ and $W^\star$ is an exponential random variable with parameter $r_\star/b_\star$. We refer to Section 3.1 for the definition of τ and the characterization of admissible values $[0, v_\star)$ 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

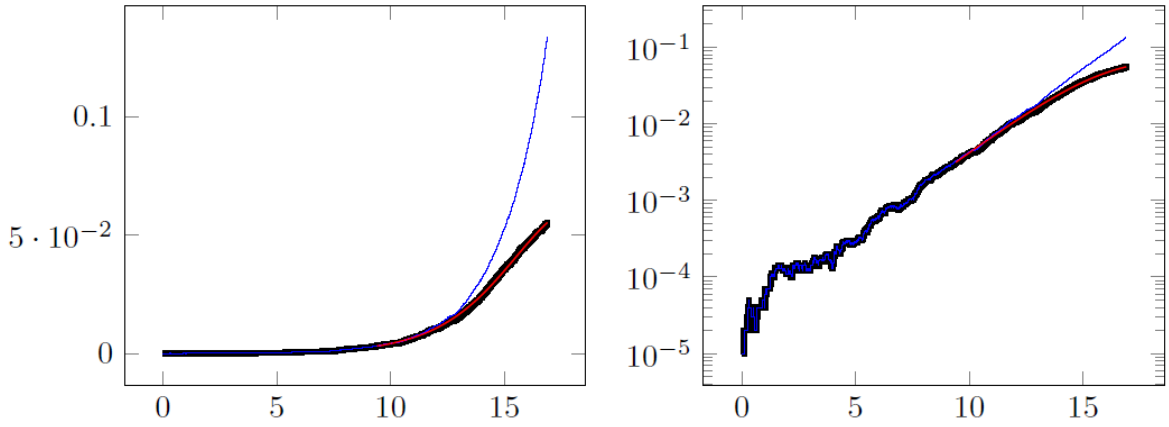

FIG 1. The black curve represents a trajectory of the process $X_M^K(t) = N_M^K(t)/K$, the blue curve a trajectory of $Z(t)/K$, and the red curve is the deterministic function $x_M^K(t)$ starting at time $T_M^K(N^K(0), \xi^K)$. Both figures represent the simulations of the same trajectory in different scales (standard scale on the left, logarithmic scale on the right). These simulations have been done with the following parameters: $K = 100000$, $F_R(x_R, x_M) = -1.5x_M$, $F_M(x_R, x_M) = 1.5x_R - 1$, $x_R^\star = 1$, $\xi_K = \sqrt{K}$ and $v = 0.055$.

by $(\log \zeta_K)/r_\star + \tau(v)$. 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), $(\log \zeta_K)/r_\star$ 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. 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 unstable) 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* $x_R^\star > 0$ *such that* $F_R(x_R^\star, 0) = 0$ *and* $\partial_R F_R(x_R^\star, 0) \leq 0$.
- (I) *Invasion:* $F_M(x_R^\star, 0) > 0$.

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, $F_M(x_R^\star, 0)$ is called invasion fitness. It quantifies the individual growth rate for a small mutant population in a resident population at the equilibrium.

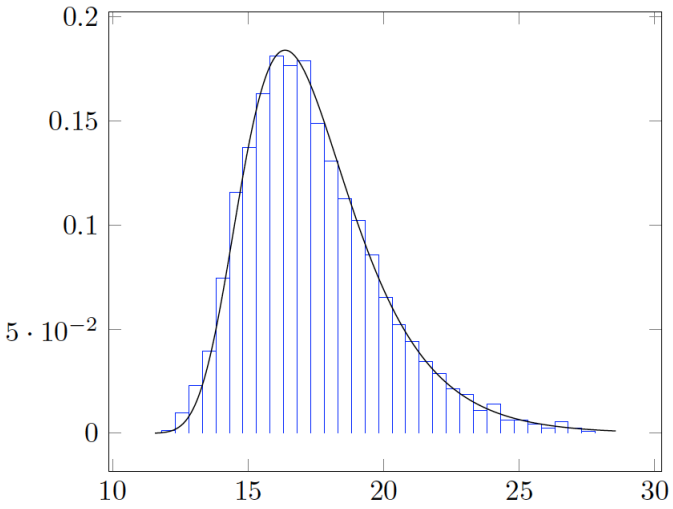

FIG 2. The histogram represents the values of $T_M^K(N^K(0), vK)$ (conditionally to the non-extinction of N_M^K). The simulations have been done with the following parameters: $K = 100000$, $F_R(x_R, x_M) = -1.5x_M$, $F_M(x_R, x_M) = 1.5x_R - 1$, $x_R^\star = 1$ and $v = 0.055$. The size of the sample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

Notation.

- We write $x^\star = (x_R^\star, 0)$.
- For $f : (x_R, x_M) \in \mathbb{R}^2 \mapsto f(x_R, x_M)$, we denote $\partial_R f$ (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 $\widetilde{\mathcal{N}}(dx) := \mathcal{N}(dx) - \lambda(dx)$ 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 θ we write C_θ.

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 $K > 0$,*

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

We introduce the branching process $(Z(t) : t \ge 0)$ 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:

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

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

We recall from (1.4) 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 in 0 :

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

where δ_0 is a Dirac mass in 0 and $\mu_a(dx) = 1_{\mathbb{R}_+}(x)ae^{-ax}dx$ and dx is the Lebesgue measure. We refer to [1] Chapter 3 or [19]-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,

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

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 from macroscopic scales. 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, there exists $C > 0$ such that the following result holds for any $\eta > 0$ and $K \geq 1$.*

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

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

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

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

Let us state the immediate corollary which will be used in Section 4.

Corollary 2.3. *If $\partial_R F_R(x^\star) < 0$, let us consider sequences $(\xi_K)_K$ such that $1 \ll \xi_K \ll K/(\log K)$. Then $N_M^K/Z - 1$ converges to 1 in probability as K goes to infinity, uniformly on the time interval $[0, T_M^K(N^K(0), \xi_K)]$.*

If $\partial_R F_R(x^\star) = 0$, a similar results 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 $\partial_R F_R(x^\star)$. 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. The choice of this neighborhood depends on $\partial_R F_R(x^\star)$. If $\partial_R F_R(x^\star) < 0$, the manifold $\{x_R = 0\}$ is stable around the fixed point $(x_R^\star, 0)$. 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 constant). If $\partial_R F_R(x^\star) = 0$, 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 consider a classical coupling to compare N_M^K and Z. Let us introduce a decomposition with the processes realizing the coupling and the additional part. The two subpopulations will be indexed respectively by q and r. We introduce

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

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

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

and in a similar way,

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

We can easily check that

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

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

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. □

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 $N^K(0) = n_0 \in \mathbb{N}^2$:

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

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

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

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

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

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

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

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 $s \leq T_{R,M}^K(n, Ln)$ and $n \leq K$,

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

almost surely. Then

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

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

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

where $\mathcal{N}^I$ is a Poisson point measure with Lebesgue intensity on $\mathbb{R}_+$, independent of $\mathcal{N}_M^b$ and $\mathcal{N}_M^d$. We can easily note that for any $t \geq 0$,

$$\mathbb{E}\left[Z_n^K(t)\right] \leq \mathbb{E}\left[Z(t)\right] = e^{r_\star t}$$

and straightforward computation leads to

$$\mathbb{E}\left[Y(t\wedge T^K_{R,M}(n,Ln))\right]\le C_L\,\frac{n}{K}\,e^{2r_\star t}.$$

Let $M(t)$ be the martingale part of the semimartingale $Y(t)$, with quadratic variation process given by

$$\langle M\rangle(t)=\int_0^t\Big(C_LZ^K_q(s)\frac{n}{K}+Y(s)(b_*+d_*)\Big)ds.$$

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

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

We compute now

$$\begin{aligned}Y(t)e^{-r_\star t}&=\int_0^t e^{-r_\star s}dM(s)+\int_0^t e^{-r_\star s}\Big(C_LZ^K_q(s)\frac{n}{K}+Y(s)(b_*-d_*)-Y(s)(b_*-d_*)\Big)ds\\&=\int_0^t e^{-r_\star s}dM(s)+\int_0^t e^{-r_\star s}\,C_LZ^K_q(s)\frac{n}{K}ds.\end{aligned}$$

Combining the previous estimates and using again $T^K_{R,M}(n,Ln)\le 4\log K/(3r_\star)$, we get

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

Using the domination of Z^K_r by Y, we obtain

$$\mathbb{E}\left(\sup_{t\le T^K_{R,M}(n,Ln)}Z^K_r(t)e^{-r_\star t}\right)\le C_L\sqrt{\log K\frac{n}{K}},$$

which proves the first estimate i) for Z_r.

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

$$\begin{aligned}\widetilde{Y}(t)\;=\;&\int_0^t\int_{\mathbb{R}_+}1_{u\le C_LZ_q(s-)n/K}\,\mathcal{N}^I(ds,du)\\&+\int_0^t\int_{\mathbb{R}_+}1_{0\le u\le Y(s-)(b_\star+C_Ln/K)}\,\mathcal{N}^b_M(ds,du)-\int_0^t\int_{\mathbb{R}_+}1_{0\le u\le Y(s-)(d_\star-C_Ln/K)}\,\mathcal{N}^d_M(ds,du).\end{aligned}$$

The conclusion follows as above by considering $\widetilde{Y}(t)\exp(-(r_\star+2C_Ln/K)t)$ for $t\le T^K_{R,M}(n,Ln)$.

For the second case ii), we proceed similarly and (2.2) becomes

$$Z^K_q(s)|b_M(X^K(s))-b_\star|+Z^K_q(s)|d_M(X^K(s))-d_\star|\le CZ^K_q(s)\sqrt{\frac{n}{K}},$$

almost surely for $s \le T^K_{R,M}(n, \sqrt{K\,n})$, since $n \le K$. Following Step 1 yields

$$\mathbb{E}\left(\sup_{t\le T^K_{R,M}(n,\sqrt{Kn})} Z^K_r(t)e^{-r_\star t}\right) \le C\left(\log K\sqrt{\frac{n}{K}}\right)^{1/2}.$$

Proceeding similarly for $N^K_{M,r}$ ends the proof. □

Lemma 2.7. *Under Assumptions 1.1 and 2.1,*

i) If $\partial_R F_R(x^\star) < 0$, for any $L > 0$, there exists $C_L > 0$ such that for any $\eta > 0$ and $n \le K$,

$$\mathbb{P}\left(\sup_{t\le T^K_{R,M}(n,Ln)} \left|\frac{N^K_M(t)}{Z(t)} - 1\right| > \eta\ ;\ W > 0\right) \le \frac{C_L}{\eta^{1/4}}\left(\frac{n\log K}{K}\right)^{1/10} \exp\left(C_L\frac{n\log K}{K}\right). \tag{2.3}$$

ii) If $\partial_R F_R(x^\star) = 0$, there exists $C > 0$ such that for any $\eta > 0$ and $n \le K$,

$$\mathbb{P}\left(\sup_{t\le T^K_{R,M}(n,\sqrt{Kn})} \left|\frac{N^K_M(t)}{Z(t)} - 1\right| > \eta\ ;\ W > 0\right) \le \frac{C}{\eta^{1/4}}\left(\log K\sqrt{\frac{n}{K}}\right)^{1/10} \exp\left(C\log K\sqrt{\frac{n}{K}}\right). \tag{2.4}$$

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

$$\frac{|N^K_M(t) - Z(t)|}{Z(t)} = \frac{|N^K_{M,r}(t) - Z^K_r(t)|}{Z(t)} = \frac{|N^K_{M,r}(t) - Z^K_r(t)|e^{-r_\star t}}{W(t)},$$

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

We use now Lemma 2.6 i) and the the fact that $W(t)$ converges to W a.s. as t tends to infinity. We deduce that for $t \le T^K_{R,M}(n, Ln)$, $|N^K_{M,r}(t) - Z^K_r(t)|/Z(t)$ 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

$$U^K = \sup_{t\le T^K_{R,M}(n,Ln)} Z_r(t)e^{-r_\star t} + \sup_{t\le T^K_{R,M}(n,Ln)} N^K_{M,r}(t)e^{-r_\star t}.$$

The previous computations ensure that

$$\frac{|N^K_M(t) - Z(t)|}{Z(t)} \le \frac{U^K}{\inf_{t\ge 0} W(t)}$$

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

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

By Markov's inequality and Lemma 2.6,

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

On the other hand, by Lemma 6.2 in Appendix,

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

This proves the result in the case i). Case ii) is proven similarly. □

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

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

i) If $\partial_R F_R(x^\star) < 0$, then there exists $L, C > 0$ 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 $\partial_R F_R(x^\star) = 0$, then there exists $L, C > 0$ 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 $\{T_M^K(N^K(0), n) \geq T_R^K(N^K(0), m) \wedge s_K\} = \{T_M^K(N^K(0), n) = T_{R,M}^K(n, m)\}$. We are using the result for sequences with $m = \xi_K$ which make 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 note $T^K = T_{R,M}^K(n, Ln)$ 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 $n \leq K/\log K$,

$$\begin{aligned}\mathbb{P}\left(T_R^K(N^K(0), Ln) \leq T_M^K(N^K(0), n) \wedge s_K\right) \\ = \mathbb{P}\left(\sup_{t \leq T^K} \left|X_R^K(t) - x_R^\star\right| \geq L\frac{n}{K}\right) \leq C\frac{\sqrt{K\log K}}{n},\end{aligned} \tag{2.5}$$

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

$$X_M^K(t) \leq \frac{n}{K} \leq \frac{1}{\log K} \text{ and } X_R^K(t) \leq x_R^\star + L\frac{n}{K} \leq x_R^\star + \frac{L}{\log K}. \tag{2.6}$$

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

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

where $\widetilde{\mathcal{N}}_R^b$ and $\widetilde{\mathcal{N}}_R^d$ 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 we recall that $G_R(x_R, x_M) = x_R(b_R - d_R)(x_R, x_M) = x_R\, F_R(x_R, x_M)$. The quadratic variation process of the martingale part M^K is the following

$$\langle M^K\rangle(t) = \frac{1}{K}\int_0^t X_R^K(s)\Big(b_R(X^K(s)) + d_R(X^K(s))\Big)ds. \tag{2.7}$$

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

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

where

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

and

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

Denoting

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

we have

$$Y_R^K(t) = (X_R^K(0) - x_R^\star) + H^K(t) + \int_0^t Y_R^K(s)\partial_R G_R(x^\star)ds, \tag{2.8}$$

where

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

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

$$\sup_{0\le s\le T^K} \left|Y_R^K(s)\right| \quad \le \quad \Gamma\left|X_R^K(0) - x_R^\star\right| + \Gamma \sup_{0\le s\le T^K} \left|H^K(s) - H^K(\lfloor s\rfloor)\right|,$$

with

$$\Gamma := (1 + |\partial_R F_R(x^\star)|)\,\frac{1}{1 - e^{\partial_R F_R(x^\star)}}.$$

Using that $H^K = M^K + H_1^K + H_2^K$ and

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

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

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

Besides, using (2.6) and (2.7) and $T^K \le C\log K$, we get $\mathbb{E}\left[\langle M^K\rangle_{T^K}\right] \le C(\log K)/K$. By Markov's and Burkholder-Davis-Gundy's inequalities, we then obtain

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

Recalling that $Y_R^K(t) = X_R^K(t) - x_R^\star$ and $n \le K$ and combining these two last estimates yields the expected inequality (2.5) and ends the first step.

We turn to the second step and prove that

$$\mathbb{P}\left(T_M^K(N^K(0), \xi_K) \wedge T_R^K(N^K(0), L\xi_K) \ge s_K\,;\, W > 0\right) \le CK^{-1/3} + \frac{C}{\eta^{1/4}}\left(\log K\sqrt{\frac{\xi_K}{K}}\right)^{1/10}. \tag{2.9}$$

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

$$\begin{aligned}\left\{T_M^K(N^K(0),n)\wedge T_R^K(N^K(0),n)\geq s_K\right\}&\\ \subseteq\left(\left\{N_M^K(s_K)\leq n+1\right\}\cap\left\{\frac{Z(s_K)}{N_M^K(s_K)}\leq 2\right\}\right)\cup\left\{\sup_{t\leq T^K}\left|\frac{N_M^K(t)}{Z(t)}-1\right|>\frac12\right\}&\\ \subseteq\left\{Z(s_K)\leq K+1\right\}\cup\left\{\sup_{t\leq T^K}\left|\frac{N_M^K(t)}{Z(t)}-1\right|>\frac12\right\}.&\end{aligned}$$

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

$$\begin{aligned}\mathbb{P}\left(T_M^K(N^K(0),n)\wedge T_R^K(N^K(0),n)\geq s_K;W>0\right)&\\ \leq\mathbb{P}\left(W(s_K)\leq 2K^{-1/3};W>0\right)+\mathbb{P}\left(\sup_{t\leq T^K}\left|\frac{N_M^K(t)}{Z(t)}-1\right|>\frac12;W>0\right).&\end{aligned}$$

The first term of the right hand side involves an estimate on the infimum of a single type branching process which we state and prove 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 (2.9).

We can now conclude the case i) combining the two steps. More precisely we use that for any $x,y,z\in\mathbb{R}$, $(x\geq y\wedge z)\Longleftrightarrow(x\wedge y\geq z$ or $x\wedge z\geq y)$. The result follows then from (2.9) and (2.5). □

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

$$\mathbb{P}\left(\sup_{t\leq T^K}\left|X_R^K(t)-x_R^\star\right|\geq L\sqrt{n/K}\right)\leq C\sqrt{\frac{\log K}{n}},\tag{2.10}$$

where now $T^K:=T^K_{R,M}(n,\sqrt{nK})$.
For that purpose, we use again (2.8) and get directly

$$\begin{aligned}\sup_{t\leq T^K}\left|X_R^K(t)-x_R^\star\right|&=\sup_{s\leq T^K}\left|Y_R^K(t)\right|\\ &\leq\left|X_R^K(0)-x_R^\star\right|+\sup_{t\leq T^K}\left|M^K(t)\right|+\sup_{t\leq T^K}\left|H_1^K(t)\right|+\sup_{t\leq T^K}\left|H_2^K(t)\right|.\end{aligned}\tag{2.11}$$

We recall that $T_K\leq 4(\log K)/(3r_\star)$ and proceeding as in the case i), the following estimates hold

$$\begin{aligned}\mathbb{P}\left(\sup_{t\leq T^K}\left|M^K(t)\right|\geq\frac12\sqrt{\frac{n}{K}}\right)&\leq C\sqrt{\frac{K}{n}}\cdot\sqrt{\frac{\log K}{K}}=C\sqrt{\frac{\log K}{n}},\\ \sup_{t\leq T^K}\left|H_1^K(t)\right|+\sup_{t\leq T^K}\left|H_2^K(t)\right|&\leq C\frac{n}{K}T^K\leq C\frac{n}{K}\log K=C\sqrt{\frac{n}{K}}\sqrt{\frac{n(\log K)^2}{K}}.\end{aligned}\tag{2.12}$$

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

$$\sup_{t\leq T^K}\left|H_1^K(t)\right|+\sup_{t\leq T^K}\left|H_2^K(t)\right|\leq\frac12\sqrt{\frac{K}{n}}.$$

Then (2.11) guarantees

$$\left\{\sup_{t\leq T^K}\left|X_R^K(t)-x_R^\star\right|\geq\sqrt{\frac{n}{K}}\right\}\subset\left\{\sup_{t\leq T^K}\left|M^K(t)\right|\geq\frac12\sqrt{\frac{n}{K}}\right\}.$$

and (2.12) yields (2.10) with $L=1/(4C^2)$. The proof is concluded as for the case i). □

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

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

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

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. □

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 [20]). 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 $x_R^\star$. 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 time scale of invasion. 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 $X^K(x,.)$ started from x is defined for any K and time t by

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

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 initial value x, i.e. the unique solution of

$$\phi(x,0)=x,\quad \frac{\partial}{\partial t}\phi(x,t)=G(\phi(x,t)),$$

where we recall that

$$G_\bullet(x)=x_\bullet(b_\bullet(x)-d_\bullet(x)).$$

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

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

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

.

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

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

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

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

We set

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

where $B_r(x^\star)$ denotes the ball of radius r centered in $x^\star$. Observing that v_η is decreasing with η, we let η 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 $x^\star$ 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 $\tau : [0, v_\star) \to \mathbb{R}_+$ such that $\tau(0) = 0$ and for any $\bar{v} \in (0, v_\star)$,

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

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

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

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

The proof of this proposition necessitates the three next 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 $\partial_R F_R(x^\star) < 0$, then there exists $u_0 \in (0,1)$ and $C > 0$ such that for any $u \leq u_0$ and $x = (x_R, x_M) \in B_u(x^\star)$,*

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

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

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

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

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

Proof. We begin by proving i). As $\partial_R F_R(x^\star) < 0$, we can find $r > 0$ such that for any $x \in B_r(x^\star)$,

$$\partial_R F_R(x) \leq \partial_R F_R(x^\star)/2 < 0. \tag{3.2}$$

For convenience we introduce the gap with initial position for the resident population dynamics and the first time when the flow leaves $B_r(x^\star)$:

$$y_R(t) := \phi_R(x,t) - x_R, \qquad s_r(x) := \inf\{t \geq 0 : \phi(x,t) \notin B_r(x^\star)\}. \tag{3.3}$$

Let us use that by definition of the flow ϕ_R

$$y_R(t) = \int_0^t \phi_R(x,s) F_R(\phi(x,s))ds = H(t) + \int_0^t \phi_R(x,s) y_R(s) \partial_R F_R(x_R, \phi_M(x,s))ds, \tag{3.4}$$

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

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

for $s \leq s_r(x)$. Using that $F_R(x_\star) = 0$, we get also $|F_R(x_R, \phi_M(x,s))| \leq C_r \left(|x_R - x_R^\star| + |\phi_M(x,s)|\right)$ and

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

Thanks to (3.2), we can use Lemma 6.3 for (3.4) and we get any $T > 0$,

$$\begin{aligned} \sup_{t \leq T \wedge s_r(x)} |y_R(t)| \leq & C_r \left(\sup_{t \leq s_r(x) \wedge T} |H(t) - H(\lfloor t \rfloor)| \right) \\ \leq & C_r \left(|x_R - x_R^\star| + \sup_{t \leq T \wedge s_r(x)} \left(|\phi_M(x,s)| + |y_R(s)|^2\right) \right). \end{aligned} \tag{3.5}$$

We work now for times when y_R is also smaller that $1/(2C_r)$ and introduce

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

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

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

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

$$\sup_{t \leq \tau_M(x,u) \wedge \overline{s}_r(x)} |y_R(t)| \leq 4C_r u. \tag{3.6}$$

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

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

which guarantees $\overline{s}_r(x) > \tau_M(x, u_r)$. Thus for $u \le u_r$, $\overline{s}_r(x) > \tau_M(x,u)$ and (3.6) becomes

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

since $\tau_M(x,u) \le \tau_M(x,u_r) < \overline{s}_r(x)$.

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

$$|y_R'(t)| = |\phi_R(x,t) F_R(\phi(x,t)| \le C\left[|x_R - x_R^\star|^2 + y_R(t)^2 + \phi_M(x,t)\right]. \tag{3.7}$$

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

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

Before this time, we can use $|y_R'(t)| \le 2C\left[|x_R - x_R^\star|^2 + \phi_M(x,t)\right]$ and we get for $t \le s(x)$,

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

Let us use again the notation $s_r(x)$ introduced at (3.3), and fix $r > 0$ to some small enough value such that for $t \le s_r(x)$, $F_M(\phi(x,t)) \ge r_\star/2$. Hence, for $t \le s_r(x)$,

$$\int_0^t \phi_M(x,s)ds = \int_0^t \partial_s \phi_M(x,s) \frac{1}{F_M(\phi(x,s))} ds \le \frac{2}{r_\star}\int_0^t \partial_s \phi_M(x,s)ds \le \frac{2}{r_\star}\phi_M(x,t). \tag{3.8}$$

We obtain for $t \le s(x) \wedge s_r(x)$,

$$|y_R(t)| = \left|\int_0^t y_R'(s)\right| \le C\left[|x_R - x_R^\star|^2 \log(1/x_M) + \phi_M(x,t)\right]. \tag{3.9}$$

Let now $x \in D := D(1/(4C^2)) = \{(x_R, x_M) : |x_R - x_R^\star| \log(1/x_M) \le 1/(4C^2)\}$ and $t \le \tau_M(x, 1/(4C^2)) \wedge s(x) \wedge s_r(x)$ such that $\phi_M(x,t) \le 1/(4C^2)$. We get

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

This implies that for any $x \in D$, $s(x) > \tau_M(x, 1/(4C^2)) \wedge s_r(x)$ and (3.9) becomes

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

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

Proof of iii). For all $t \le \tau_M(x,u)$,

$$\partial_t \phi_M(x,t) \ge \phi_M(x,t)\frac{r_\star}{2},$$

whence, for all $t \le \tau_M(x,u)$,

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

The inequality above implies that $\tau_M(x,u)$ is finite. On the other hand, since $t \mapsto \phi_M(x,t)$ is continuous, we know that $\phi_M(x, \tau_M(x,u)) = u$. Consequently

$$u \ge x_M e^{\tau_M(x,u) \cdot r_\star/2},$$

which implies

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

Considering u smaller than one, this proves the result. □

We also need an estimate of the cumulated gap :

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

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

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

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

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

Firstly, let us prove the result under the assumption $\partial_R F_R(x^\star) < 0$. Using (3.5) and noting

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

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

$$z(t) \le C\left(|x_R - x_R^\star| + \phi_M(x,t) + z(t)^2\right),$$

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,

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

We use Lemma 3.3 to estimate $z(\tau_M(x,u))$. Moreover (3.8) gives a bound for $\int_0^{\tau_M(x,u)} \phi_M(x,s)ds$. Adding that $\tau_M(x,u) \le 2\log(1/x_M)/r_\star$ (see Lemma 3.3 iii)), we obtain for any $x \in D(\eta)$,

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

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

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

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

We can proceed similarly under the assumption $\partial_R F_R(x^\star) = 0$, using now (3.7) and

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

The proof is completed. □

Lemma 3.5. *For all $v \in (0, v_\star)$, there exist $\eta \in (0, r_\star)$ and $r > 0$ such that, for all $x \in B_r(x^\star) \cap (\mathbb{R}_+^*)^2$ and $t \le \tau_M(x,v)$, $F_M(\phi(x,t)) \ge \eta$.*

Proof. Since $v < v_\star$, there exists $\eta \in (0, r_\star)$ such that $v < v_\eta$. Hence there exists $r > 0$ such that

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

Let x belong to $B_r(x^\star) \cap (\mathbb{R}_+^*)^2$. Then $v < \mathcal{C}(x,\eta) = \sup\{\phi_M(x,t) : t \le \overline{\tau}_M(x,\eta)\}$ and we get $F_M(\phi(x,t)) \ge \eta$ for $t \le \tau_M(x,v)$ by continuity of $t \to F_M(\phi(x,t))$. □

We can now prove Proposition 3.2.

Proof of Proposition 3.2. 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 $u_0 > 0$ and $\eta_0 > 0$ such that, for all $u \le u_0$, and $x \in B_u(x^\star)$,

$$|\phi_R(x, \tau_M(x,u)) - x_R^\star| \le \sup_{t \le \tau_M(x,u)} |\phi_R(x,t) - x_R^\star| \le Cu. \tag{3.10}$$

Then

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

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

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

This ensures that $v_\star \ge u_1 > 0$. Besides, starting close to the neighborhood of $x^\star$, the flow is bounded by Assumption 3.1. This ensures that $v_\star$ is finite.

Let us now prove the second part of the proposition. We start by proving the result for small values of v using the first part which guarantees a lowerbound for the growth rate of mutants. For any $v \in (0, u_1]$ and $x \in B_v(x^\star)$ and $t \le \tau_M(x,v)$,

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

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 $x_M \le v$) and $t \le \tau_M(x,v)$, 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 (3.10), and we obtain that for any $u \in (0, u_1]$ and $x \in B_{u_1}(x^\star)$ (with $x_M \le u$),

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

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

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

Besides, for any $u_0 \in (0, u_1]$, $(x,u) \to F_M(\phi_R(x, \tau_M(x,u)), u)$ is uniformly continuous on $\{(x,u) : x \in B_{u_1}(x^\star), x_M \le u, u \in [u_0, u_1]\}$. It can be shown by a linearization argument at the equilibrium $x^\star$ (Hartman Grobman Theorem). Combining both facts and splitting the interval $[x_m, v] = [x_m, u_0] \cup [u_0, v]$ yields

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

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,

$$\tau(v) = \lim_{x \to x^\star} \int_{x_M}^{v} \frac{1}{u} \left(\frac{1}{F_M(\phi_R(x, \tau_M(x,u)), u))} - \frac{1}{r_\star} \right) du \tag{3.11}$$

and the previous estimates also guarantee that τ 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 can 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:

$$\tau_M(x, u_1, v) = \tau_M((\phi_R(x, \tau_M(x, u_1)), u_1), v) = \int_{u_1}^{v} \frac{1}{u} \frac{1}{F_M(\phi_R(x, \tau_M(x,u)), u))} \, du. \tag{3.12}$$

It remains to check that this term converges as x goes to x^* (uniformly for $u \in [u_1, u_2]$ where $u_2 < v_\star$). This can be achieved as above using uniform continuity of $(x, u) \in \{(x, u) : x \in B_{u_1}(x^\star), u \in [u_1, u_2]\} \to F_M(\phi_R(x, \tau_M(x,u)), u)$. It ensures that the limit is continuous and increasing with respect to v.
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 [28]. □

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 $x_K = (x_R^K, x_M^K) \in (\mathbb{R}_+^*)^2$, satisfying*

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

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

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

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 [24, 25].

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

$$X_\bullet^K(t) := X_\bullet^K(x_K, t)\,,\; x_\bullet^K(t) := \phi_\bullet(x_K, t)$$

and

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

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

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

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 θ_R^K, σ_η^K and $\tilde{\sigma}_\varepsilon^K$.

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

$$\mathbb{P}\left(\sigma_\eta^K \leq \tau_M(x_K, L\eta) \wedge \theta^K\right) \leq \frac{C}{\eta} K^{-1/2} \left(\log\left(1/x_M^K\right)\right)^{1/2}, \tag{3.13}$$
$$\mathbb{P}\left(\tilde{\sigma}_{L\eta}^K \leq \tau_M(x_K, L\eta) \wedge \theta^K \wedge \sigma_\eta^K\right) \leq \frac{C}{\eta} K^{-1/2} \left(\log\left(1/x_M^K\right)\right)^{1/2}, \tag{3.14}$$
$$\mathbb{P}\left(\theta_R^K \leq \tau_M(x_K, L\eta) \wedge \theta_M^K\right) \leq \frac{C}{\eta} K^{-1/2} \left(\log\left(1/x_M^K\right)\right)^{1/2}. \tag{3.15}$$

Proof. To begin with, let us prove (3.13). 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 σ_η^K,

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

To control the probability of this event, we use that

$$Y_R^K(t) = H^K(t) + \int_0^t Y_R^K(s)\partial_R G_R(x^K(s))ds \tag{3.16}$$

for all $t > 0$, where

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

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

$$\sup_{s\leq t} \left|Y_R^K(s)\right| \leq \sup_{s\leq t} \left|H^K(s) - H^K(\lfloor s\rfloor)\right| \leq 2 \sup_{s\leq t} |H^K(t)|.$$

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

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

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

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

Combining these two estimates, Lemma 3.3 implies that

$$|Y_R^K(t)| \le \sup_{s\le t}|H^K(s)| e^{CL\eta} \le 2\sup_{s\le t}|H^K(s)|,$$

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

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

$$\begin{aligned}\sup_{s\le t}\left|Y_R^K(s)\right| \le& 2\sup_{s\le t}|H^K(s)| \\ \le& 2\sup_{s\le t}|M^K(s)| + 2\sup_{s\le t}|R^K(s)| + 2\int_0^t |Y_M^K(s)|\cdot|\partial_M G_R(x^K(s))|ds. \end{aligned} \tag{3.17}$$

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

$$\begin{aligned}\int_0^t x_M^K(s)ds &= \int_0^t \frac{1}{F_M(x^K(s))}(x_M^K)'(s)ds \\ &\le C\int_0^t (x_M^K)'(s)ds = C(x_M^K(t) - x_M^K(0)) \le Cx_M^K(t). \end{aligned} \tag{3.18}$$

Using twice this inequality and $|Y_M^K(r)| = x_M^K(r)|X_M^K(r)/x_M^K(r) - 1| \le 2x_M^K(r)$ for $r \le \theta^K$,

$$\int_0^{S^K}\int_0^s |Y_M^K(r)|\cdot|\partial_M G_R(x^K(r))|drds \le C\int_0^{S^K}\int_0^s x_M^K(r)drds \le Cx_M^K(S^K) \le CL\eta. \tag{3.19}$$

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

$$\int_0^{S^K}\sup_{r\le s}|R^K(r)|ds \le C\int_0^{S^K}\int_0^s \left(|Y_M^K(r)|^2 + |Y_M^K(r)|\cdot|Y_R^K(r)| + |Y_R^K(r)|^2\right)drds.$$

Since $S^K \le \tau_M(x_K, L\eta)$, and using the same computation as (3.19),

$$\int_0^{S^K}\int_0^s |Y_M^K(r)|^2 drds \le L\eta\int_0^{S^K}\int_0^s |Y_M^K(r)|drds \le C(L\eta)^2.$$

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

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

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

$$\int_0^{S^K}\sup_{r\le s}|R^K(r)|ds \le C\eta^2. \tag{3.20}$$

Plugging (3.19) and (3.20) in (3.17), we can choose $L>0$ small enough and get

$$\mathbb{P}\left(\int_0^{\tau_M(x_K,L\eta)\wedge\theta^K}|Y_R^K(s)|ds>\eta\right)\leq\mathbb{P}\left(\int_0^{\tau_M(x_K,L\eta)\wedge\theta^K}|M^K(s)|ds>\frac{\eta}{4}\right).$$

Using again Doob and Cauchy-Schwarz inequalities,

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

Using Markov's inequality and the comparison of $\tau_M(x_K,L\eta)$ with $\log\left(1/x_M^K\right)$ (see Section 3.1), we get

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

This proves (3.13).

Now we prove (3.14). 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 (3.17) again. We have that

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

and

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

Since η is arbitrary small, we have

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

which proves (3.14).

Finally, we prove (3.15) using (3.13) and (3.14). Notice that

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

Then, by Lemma 3.3, we know that for all $t \le \tau_M(x_K, L\eta)$,

$$x_R^K(t) \ge x_R^\star - CL\eta \ge x_R^\star/2,$$

if $\eta > 0$ is small enough. Hence, for $\eta > 0$ small enough,

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

Since the probabilities of the two last events above are controlled respectively by (3.13) and (3.14), the result is proved. □

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 $\eta \in (0, \eta_0)$,

$$\mathbb{P}\left(\sup_{t\le\tau_M(x_K,L\eta)} \left|\frac{X_M^K(t)}{x_M^K(t)} - 1\right| > \eta\right) \le C \cdot \frac{1}{\eta} \cdot \frac{1}{\sqrt{Kx_M^K(0)}}. \tag{3.21}$$

We use that

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

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

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

we have

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

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

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

Consequently,

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

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

$$\sup_{t \leq T^K} \left| \frac{X_M^K(t)}{x_M^K(t)} - 1 \right| \leq \left(\sup_{t \leq T^K} |E_M^K(t)| + C \int_0^T \left| X_R^K(s) - x_R^K(s) \right| ds \right) e^{C \int_0^T x_M^K(s) ds}. \tag{3.22}$$

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

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

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 $F_M(x^\star) > 0$ if η is chosen small enough. Then $x_M^K(s) \geq e^{\kappa s} x_M^K(0)$ and

$$\mathbb{E}\left[\sup_{t \leq T^K} |E_M^K(t)| \right] \leq C \cdot \frac{1}{\kappa} \cdot \frac{1}{\sqrt{K x_M^K(0)}}, \tag{3.23}$$

where the value of C has changed in the last inequality, but is still independent of K.
Let us recall that, thanks to (3.18), we have

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

So, (3.22) can be rewritten as

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

By Markov's inequality and (3.23),

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

Besides, by (3.13) from Lemma 3.7,

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

We obtain from the three last inequalities that

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

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

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

and (3.15) proves (3.21).

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 denote

$$\tilde X_\bullet^K(t):=X_\bullet^K(t+\tau_M(x_K,L\eta)),\qquad \tilde x_\bullet^K(t):=x_\bullet^K(t+\tau_M(x_K,L\eta))$$

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

$$\gamma_{T,\eta}=\inf_{t\le T,K\ge1}\tilde x_M^K(t)\;>0,$$

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

$$\mathbb{P}\left(\sup_{t\le T}\frac{\left|\tilde X_M^K(t)-\tilde x_M^K(t)\right|}{\tilde x_M^K(t)}>\eta\right)\le\mathbb{P}\left(\sup_{t\le T}\left|\tilde X_M^K(t)-\tilde x_M^K(t)\right|\;>\;\eta\,\gamma_{T,\eta}\right),$$

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

$$\mathbb{P}\left(\sup_{t\le T}\left|\tilde X_M^K(t)-\tilde x_M^K(t)\right|>\eta\right)\underset{K\to\infty}{\longrightarrow}0. \tag{3.24}$$

This result is a classical approximation of population processes by dynamical systems on finite time intervals, see Ethier-Kurtz [20]. 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 $t\le\tilde\theta^K$,

$$\left|\tilde X_\bullet^K(t)-\tilde x_\bullet^K(t)\right|\le\left|M_\bullet^K(t)\right|+C\int_0^t\left(|\tilde X_R^K(s)-\tilde x_R^K(s)|+|\tilde X_M^K(s)-\tilde x_M^K(s)|\right)ds,$$

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

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

Then, Gronwall's lemma implies that

$$\sup_{t\le T\wedge\tilde\theta^K}\left(\left|\tilde X_M^K(t)-\tilde x_M^K(t)\right|+\left|\tilde X_R^K(t)-\tilde x_R^K(t)\right|\right)\le\left(\sup_{t\le T\wedge\tilde\theta^K}|M_R^K(t)|+\sup_{t\le T\wedge\tilde\theta^K}|M_M^K(t)|\right)e^{CT}.$$

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

$$\mathbb{P}\left(\sup_{t\leq T\wedge\tilde{\theta}^K}\left(\left|\tilde{X}_M^K(t)-\tilde{x}_M^K(t)\right|+\left|\tilde{X}_R^K(t)-\tilde{x}_R^K(t)\right|\right)>\eta'\right)\leq\frac{1}{\eta'}C_TK^{-1/2}. \tag{3.25}$$

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

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

We have

$$\begin{aligned}\mathbb{P}\left(A_{\eta'}^K(T)\right)&\leq\mathbb{P}\left(A_{\eta'}^K(T)\cap\{\tilde{\theta}^K\geq T\}\right)+\mathbb{P}\left(A_{\eta'}^K(T)\cap\{\tilde{\theta}^K<T\}\right)\\&\leq\mathbb{P}\left(A_{\eta'}^K(T\wedge\tilde{\theta}^K)\right)+\mathbb{P}\left(A_{\eta'}^K(T\wedge\tilde{\theta}^K)\right)\\&\leq 2\mathbb{P}\left(A_{\eta'}^K(T\wedge\tilde{\theta}^K)\right).\end{aligned}$$

Recalling that (3.25) provides a control of $\mathbb{P}\left(A_{\eta'}^K(T\wedge\tilde{\theta}^K)\right)$, (3.24) is proved.

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 (3.12) with $u_1=L\eta$ for instance. □

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

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

The main result of this section is the following, which complements the results of [6, 7]. We recall that $v_\star$ is defined at the beginning of Section 3.1. 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.

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 $\zeta_K/K\to v$ as $K\to\infty$, for some $v\in[0,v_\star)$. Then, for any $\varepsilon>0$,*

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

where τ is the increasing continuous function such that $\tau(0)=0$ defined in (3.11) and

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

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 (ξ_K).

Lemma 4.2. *Under Assumptions 1.1 and 2.1, for any sequence ξ tending to infinity and satisfying $\xi_K \ll K/\log K$ (resp. $\xi_K \ll K/(\log K)^2$) if $\partial_R F_R(x^\star) < 0$ (resp. $\partial_R F_R(x^\star) = 0$), 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 $K \geq K_\varepsilon$,

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

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

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

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. 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 ξ_K by the process Z_+, we have

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

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

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

Consequently, for all $\varepsilon > 0$,

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

Letting ε go to zero proves the result. □

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

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

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

Proof. Consider the case $\partial_R F_R(x^\star) < 0$ 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 $K \geq K_0$,

$$\mathbb{P}\Big(\forall t \in [0, T^K],\ (1-\varepsilon)Z(t) \leq N_M^K(t) \leq (1+\varepsilon)Z(t) \,|\ W > 0\Big) \geq 1 - \varepsilon,$$

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

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

as u goes to infinity, then there exists t_0 such that

$$\mathbb{P}\Big(\forall t \geq t_0,\ (1-\varepsilon)We^{r_\star t} \leq Z(t) \leq (1+\varepsilon)We^{r_\star t} \,|\, W > 0\Big) \geq 1 - \varepsilon.$$

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

$$\mathbb{P}\Big(\forall t \in [t_0, T^K],\, (1-\varepsilon)^2 W e^{r_\star t} \leq N_M^K(t) \leq (1+\varepsilon)^2 W e^{r_\star t} \,|\, W > 0\Big) \geq 1 - 3\varepsilon.$$

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

$$\mathbb{P}\Big(\ (1-\varepsilon)^2 W e^{r_\star T_K} \leq \xi_K + 1,\, \xi_K \leq (1+\varepsilon)^2 W e^{r_\star t} \,|\, W > 0\Big) \geq 1 - 3\varepsilon.$$

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

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

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

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

$$\big|\inf\{t \geq 0 : f(t) \geq x\} - \inf\{t \geq 0 : g(t) \geq x\}\big| \leq \frac{\beta}{\alpha}.$$

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

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

where the last inequality holds for any $\varepsilon > \beta/\alpha$. This forces $\inf\{t \geq 0 : f(t) \geq x\} \geq t(x) - \varepsilon$. 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 $\inf\{t \geq 0 : f(t) \geq x\} \leq t(x) + \varepsilon$ and completes the proof since ε can be chosen arbitrarily close to β/α. □

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

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

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

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

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

in probability as $K \to \infty$. For any sequence of positive $v_K \geq X_M^K(0)$ upper-bounded by some constant $v < v_\star$, for any $\varepsilon > 0$,

$$\lim_{K\to\infty} \mathbb{P}\left(\left|S_M^K(X^K(0), v_K) - \tau_M(X^K(0), v_K)\right| \geq \varepsilon\right) = 0.$$

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

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

We apply now Lemma 4.4 with

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

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

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

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. □

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

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

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 Proposition 4.3 and Lemma 4.5. □

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 the individual birth and death rates $b_R, b_M \geq 0$ and $d_R, d_M \geq 0$ and the 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. The limiting ODE is (1.2) with

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

Then

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

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

$$\partial_R F_R(x_R^\star, 0) = -c_{2,2} < 0.$$

This means that the first part of Assumption 1.1 is satisfied for $(x_R^\star, 0)$. Moreover, $x_R^\star$ 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 $b_M - d_M - (b_R - d_R)\dfrac{c_{2,1}}{c_{1,1}} > 0$. It is then easy to compute $v^\star$, which is

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

The branching process Z is a birth and death process with individual birth rate b_M and individual death rate $d_M + c_{2,1}x_R^\star$.

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, 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 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

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

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 β and individual death rate γ.

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 and adding that $\gamma \log(x_S) + \beta(1 - x_S - x_I) = 0$, one can check that

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

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

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

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 $v < v_\star$ and $\varepsilon > 0$,

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

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

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

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

We obtain also an approximation of the time when the epidemics reaches a given level :

Corollary 5.2. *Let $(\zeta_K)_K$ be sequence which tends to infinity and satisfies $\zeta_K/K \to v$ as $K \to \infty$, for some $v \in [0, v_\star)$. Then, for any $\varepsilon > 0$,*

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

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

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 $W(t) := Z(t)\exp(-r_\star t)$. 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 τ,*

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

where W is the almost sure limit of $W(t)$ as t goes to infinity, and C is some positive constant independent of ε and τ. 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 τ is deterministic and finite. To begin with,

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

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

$$\begin{aligned}\mathbb{E}\left[\left(W(\tau+1)-W(\tau)\right)^2\right] =& e^{-2r_\star\tau}\mathbb{E}\left[\mathbb{V}ar\left[\sum_{i=1}^{Z(\tau)}\left(N_i e^{-r_\star}-1\right)|Z(\tau)\right]\right]\\ =& e^{-2r_\star\tau}\mathbb{E}\left[Z(\tau)\right]\mathbb{V}ar\left[N_i e^{-r_\star}-1\right] \leq Ce^{-r_\star\tau}. \end{aligned} \tag{6.1}$$

On the other hand, we write

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

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:

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

This ends the proof in the case where τ is deterministic.

Now to prove the result when τ 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

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

and the last computation of the previous case gives

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

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

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

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

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

$$\mathbf{1}_{\{W(\tau\wedge n)\leq\varepsilon; W>0\}} \underset{n\to\infty}{\longrightarrow} \mathbf{1}_{\{W(\tau)\leq\varepsilon; W>0\}}.$$

Hence, by Fatou's lemma,

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

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

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

Noticing that

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

we have

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

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

$$\frac{d_\star}{b_\star}\delta_0 + \frac{r_\star}{b_\star}\mathcal{E}(\frac{r_\star}{b_\star}).$$

Hence

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

Recalling (6.2), the result of the lemma is proved. □

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* ε.

Proof. Let us introduce

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

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

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

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_ε, $S_\varepsilon \geq s_\varepsilon$, 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. □

6.2. Perturbation of stable dynamical system

We provide a result which is strongly inspired from Lemmas 3.1 and 3.2 in [27]. 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, ϕ 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 ϕ is bounded, upper-bounded by some negative number and

$$-\infty < \inf \phi \leq \sup \phi < 0.$$

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

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

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:

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

Using the upperbound of ϕ, 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 $\Gamma = \Gamma' \sum_{j=0}^{\infty} e^{-Cj} = \Gamma'/(1 - e^{-C})$. □

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.
- 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 (2004). *Branching processes*. Dover Publications, Inc., Mineola, NY.
[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. Schlagetter, 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] 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.
[18] 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.
[19] R. Durrett. Branching process models of Cancer. Mathematical Biosciences Institute Lecture Series 1.1, Stochastics in Biological Systems, Springer, 2015.
[20] S.N. Ethier, T.G. Kurtz. Markov processes, Characterization and Convergence. John Wiley & Sons, New York, 1986.
[21] J. Jacod and A.N. Shiryaev. *Limit Theorems for Stochastic Processes.* Springer-Verlag, Berlin, 1987.
[22] A. Jakubowski. On the Skorokhod topology. *Annales de l'Institut Henri Poincaré*, 22 No 3, 263-285, 1986.
[23] 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.
[24] T. G. Kurtz. Solutions of ordinary differential equations as limits of pure jump Markov processes. *J. Appl. Prob.* 7,49-58, 1970.
[25] T. G. Kurtz. Strong approximation theorems. *Stoch. Proc. Appl.* 6, 223-240, 1978.
[26] 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.
[27] A. Prodhomme. Strong Gaussian approximation of metastable density-dependent Markov chains on large time scales. arXiv:2010.06861 (2020).
[28] F. Takens. Partially hyperbolic fixed points. *Topology*, Vol 10, 133-147, section 4. (1971)