SIAM J. APPL. MATH.
Vol. 80, No. 3, pp. 1567–1589

A METHOD TO DEAL WITH THE CRITICAL CASE IN STOCHASTIC POPULATION DYNAMICS*

DANG H. NGUYEN† AND EDOUARD STRICKLER‡

Abstract. In numerous papers, the behavior of stochastic population models is investigated through the sign of a real quantity which is the growth rate of the population near the extinction set. In many cases, it is proven that when this growth rate is positive, the population is persistent in the long run, while if it is negative, the population goes extinct. However, the critical case when the growth rate is null is rarely treated. The aim of this paper is to provide a method that can be applied in many situations to prove that in the critical case, the process converges in temporal average to the extinction set. A number of applications are given for stochastic differential equations and piecewise deterministic Markov processes modeling prey-predator, epidemiological or structured population dynamics.

Key words. Lyapunov exponents, stochastic persistence, piecewise deterministic Markov processes, stochastic differential equation, stochastic environment

AMS subject classifications. 60J25, 60J60, 37H15, 37A50, 92D25, 92D30

DOI. 10.1137/20M131134X

1. Introduction.

Much effort in population biology has been devoted to understanding under what conditions interacting populations, whether they be viruses, plants, or animals, coexist or go extinct. The variation of environmental factors such as temperature, precipitation, and humidity inherently affects the growth rates of the species. There is rich literature showing that the interplay of biotic interactions and environmental fluctuations can facilitate or suppress the persistence of species or disease prevalence; see [16, 9, 1, 8, 6] and the references therein. There has been intensive attention paid to modeling and analysis of ecological and epidemiological models under environmental stochasticity.

As one of the simplest models, the evolution of a single species can be modeled by the stochastic logistic equation:

$$dX_t = X_t(r - kX_t)dt + \sigma X_t dW_t. \tag{1}$$

The key quantity is the so-called stochastic growth rate $\Lambda := r - \frac{\sigma^2}{2}$. It is well-known that if $\Lambda < 0$, the population goes extinct almost surely, that is, $\lim_{t\to\infty} X(t) = 0$ a.s. In the case $\Lambda > 0$ the population is persistent and the transition probability of $(X(t))$ converges to its unique invariant probability measure on $(0, \infty)$. In the critical case $r - \frac{\sigma^2}{2} = 0$, the process is null-recurrent, and $X(t)$ does not go extinct almost surely but also does not have an invariant probability measure on $(0, \infty)$ and the time-average $\frac{1}{t}\int_0^t X(s)ds$ converges to 0 almost surely. Readers are referred to [14] for the proof and more details. However, it requires more advanced and dedicated methods to analyze higher dimensional systems, which depict interacting populations, especially when one has to analyze invariant measures instead of an equilibrium.

*Received by the editors January 7, 2020; accepted for publication (in revised form) March 24, 2020; published electronically June 25, 2020.
https://doi.org/10.1137/20M131134X
Funding: Dang H. Nguyen is supported in part by the National Science Foundation under grant DMS 1853467.
†Department of Mathematics, University of Alabama, Tuscaloosa, AL 35487-0350 (hnugyen4@ua.edu).
‡Université de Lorraine, CNRS, Inria, IECL, Nancy, France (edouard.strickler@inria.fr).

In [28], a condition for coexistence was given, which requires a certain weighted combination of populations' invasion rates to be positive for any invariant measures associated with a subcollection of populations. The results were then improved and generalized to a very general setting by Benaïm in [4], where the concept of *H-persistence* was coined and developed. With the same idea, [18] provided conditions for both extinction and persistence in a setting of stochastic differential equations. The long-term properties of some specific models are also classified in [12, 13, 6, 26, 20, 7, 17]. For many models, the conditions in the aforesaid references for extinction and persistence of a species in an interacting population are determined by a threshold Λ whose sign indicates whether the species will be persistent or extinct. Namely, the result obtained is that if $\Lambda > 0$, the species persists and if $\Lambda < 0$, extinction will happen.

The critical case $\Lambda = 0$ largely remains untreated, except for a few special cases such as the stochastic logistic model introduced above. Although the set of parameters for which $\Lambda = 0$ often has Lebesgue measure, it is of great mathematical interest to discover the dynamics of the systems in the critical cases. Analyzing the critical case not only fully classifies the long-term behaviors of the system but helps to gain more insights about the nature of the system. However, similar to (but more complicated than) the case of an equilibrium of a deterministic dynamical system whose maximum eigenvalue is 0, treating the critical cases of stochastic systems is, in general, extremely difficult, which might be the reason why the critical case usually remains open, especially for high dimensional systems.

However, population models often exhibit some certain monotone properties that can be utilized to tackle the critical cases. This paper provides some methods for treating the critical cases of population dynamics under certain conditions. It is partially inspired by the work of the first author [26], where the critical case is treated for a stochastic chemostat dynamic modeled by a switching diffusion.

The rest of the paper is organized as follows. In section 2, we formulate the model in the general setting of [4] and give a general condition for extinction in the average of stochastic populations in a critical case. Section 3 is devoted to the analysis of a number of specific models in critical cases. Different techniques are introduced so that the general result in section 2 becomes applicable for those models.

2. Notation and results.

Before we give our result, we present the very general framework of [4] for stochastic persistence and extinction. Let $(X_t)_{t\geq 0}$ be a càdlàg Markov process on a locally compact Polish metric space $(\mathcal{M}, d)$. For a distribution ν on $\mathcal{M}$, we set, as usual, $\mathbb{P}_\nu$ for the law of the process X with initial distribution ν and $\mathbb{E}_\nu$ for the associated expectation. If $\nu = \delta_x$ for some $x \in \mathcal{M}$, we write $\mathbb{P}_x$ for $\mathbb{P}_{\delta_x}$. We denote by $(P_t)_{t\geq 0}$ the semigroup of X acting on bounded measurable function $f : \mathcal{M} \to \mathbb{R}$ as

$$P_t f(x) = \mathbb{E}_x \left(f(X_t) \right).$$

An invariant distribution for the process X is a probability μ such that $\mu P_t = \mu$ for all $t \geq 0$. We let $\mathcal{P}_{inv}$ denote the set of all invariant probability measures of X and for $N \subset \mathcal{M}$ and let $\mathcal{P}_{inv}(N)$ and $\mathcal{P}_{erg}(N)$ denote the (possibly empty) sets of invariant probability measures and ergodic invariant probability measures, respectively, giving mass 1 to the set N. The following is the standing assumption.

Hypothesis 2.1. There exists a nonempty closed set $\mathcal{M}_0 \subset \mathcal{M}$ called the *extinction set* which is invariant under $(P_t)_{t\geq 0}$. That is, for all $t \geq 0$,

$$P_t \mathbb{1}_{\mathcal{M}_0} = \mathbb{1}_{\mathcal{M}_0}.$$

We set

$$\mathcal{M}_+ = \mathcal{M} \setminus \mathcal{M}_0.$$

The two following assumptions are taken from [4].

Hypothesis 2.2. The semigroup $(P_t)_{t\geq 0}$ is C_b*-Feller*, meaning that for all continuous bounded functions $f : \mathcal{M} \to \mathbb{R}$, $(t, x) \mapsto P_t f(x)$ is a continuous function.

We let $\mathcal{L}$ denote the infinitesimal generator of P_t on the space $C_b(\mathcal{M})$ of continuous bounded functions, defined for $f \in \mathcal{D}(\mathcal{L})$ by

$$\mathcal{L}f(x) = \lim_{t\to 0} \frac{P_t f(x) - f(x)}{t},$$

where the domain is the set of functions such that the above convergence holds pointwise, with the additionnal property that $\sup_{0<t\leq 1} \|t^{-1}(P_t f - f) - \mathcal{L}f\| < +\infty$. We also let $\mathcal{D}^2(\mathcal{L})$ be the set of functions such that both f and f^2 lie in $\mathcal{D}(\mathcal{L})$, and we define the carré du champ operator on $\mathcal{D}^2(\mathcal{L})$ by

$$\Gamma f = \mathcal{L}f^2 - 2f\mathcal{L}f.$$

For all $t > 0$, we let Π_t denote the empirical occupation measure of the process X up to time t. This is the random probability measure defined on $\mathcal{M}$ by

$$\Pi_t = \frac{1}{t}\int_0^t \delta_{X_s} ds.$$

When we want to emphasize the starting point, we set Π_t^x for the empirical occupation measure whenever $X_0 = x$ almost surely.

Hypothesis 2.3. For all $x \in \mathcal{M}$, the sequence $\{\Pi_t^x, t \geq 0\}$ is almost surely tight.

As proved in [4, Theorem 2.1], a sufficient condition for the tightness of the sequence of the empirical occupation measures is the existence of a suitable Lyapunov function, as defined in the following assumption. Recall that a map $f : \mathcal{M} \to \mathbb{R}_+$ is said to be *proper* if for all $R > 0$, the sublevel set $\{f \leq R\}$ is compact in $\mathcal{M}$.

Hypothesis 2.4. There exist continuous proper maps $W, \tilde{W} : \mathcal{M} \mapsto \mathbb{R}_+$ and a continuous map $LW : \mathcal{M} \mapsto \mathbb{R}$ enjoying the following properties:

(a) For all compact $K \subset \mathcal{M}$ there exists $W_K \in \mathcal{D}^2$ with $W|_K = W_K|_K$ and $(\mathcal{L}W_K)|_K = LW|_K$.

(b) For all $x \in M$, $\sup_{\{t\geq 0, K: K\subset\mathcal{M}, K \text{ compact}\}} P_t\Gamma(V_K)(x) < \infty$.

(c) $LW \leq -\tilde{W} + C$.

The latter assumption also implies that all weak-limit points of the sequence $(\Pi_t)_{t>0}$ are almost surely in $\mathcal{P}_{inv}(\mathcal{M})$ (see [4, Theorem 2.1]).

The next assumption ensures the existence of a Lyapunov function near the boundary $\mathcal{M}_0$.

Hypothesis 2.5. There exist continuous maps $V : \mathcal{M}_+ \mapsto \mathbb{R}_+$ and $H : \mathcal{M} \mapsto \mathbb{R}$ enjoying the following properties:

(a) For all compact $K \subset \mathcal{M}_+$ there exists $V_K \in \mathcal{D}^2$ with $V|_K = V_K|_K$ and $(\mathcal{L}V_K)|_K = H|_K$.

(b) For all $x \in \mathcal{M}$, $\sup_{\{K: K\subset\mathcal{M}, K \text{ compact};\, t\geq 0\}} P_t\Gamma(V_K)(x) < \infty$.

(c) The map $\frac{\tilde{W}}{1+|H|}$ is proper.

Remark 2.6. Assumption 2.4 to control the dynamics of the system near infinity. The Lyapunov function W together with the function $\tilde{W}$ with satisfying conditions (a), (b), (c) in Hypothesis 2.4 ensures that we have the tightness and boundedness in certain sense of the solution. On the other hand, the Lyapunov function V will manage the dynamics when the solution is near the boundary $\mathcal{M}_0$. Roughly speaking, analyzing the average of $H = \mathcal{L}V$ with respect to invariance measures on the boundary will determine whether the process will converge to the boundary $\mathcal{M}_0$. Condition (c) in Hypothesis 2.5 is needed to handle the scenario when the process is close to both the boundary and infinity. A quick example will be given below while more details about the main ideas and examples can be found in [4] and [18].

Example 2.7. This example is taken from [4, section 5.2]. We consider the following stochastic Rosenzweig–MacArthur model:

$$\begin{cases} dX_t = X_t\left(1 - \frac{X_t}{K} - \frac{Y_t}{1+X_t}\right)dt + \varepsilon X_t dB_t, \\ dY_t = Y_t\left(-\alpha + \frac{X_t}{1+X_t}\right)dt. \end{cases} \tag{2}$$

In this case, $\mathcal{M} = \mathbb{R}^2_+ := \{(x,y) \in \mathbb{R}^2\ : x, y \geq 0\}$, $\mathcal{M}_+ := \{(x,y) \in \mathbb{R}^2\ : x, y > 0\}$ and $\mathcal{M}_0 = \mathcal{M} \setminus \mathcal{M}_+$. It is proven in [4, Theorems 5.1 and 5.5] that Assumption 2.4 is satisfied with $W(x,y) = (x+y)^2$ and $\tilde{W} = (1+C)W$, where C is some constant. The function V can be chosen such that $V(x,y) \leq C_1 \ln(1+x+y)$ and $V(x,y) = -C_2 \ln x - \ln y$ when (x,y) is close to $\mathcal{M}_0$, where C_1, C_2 are sufficiently large positive numbers. The use of $-\ln x$ and $-\ln y$ is to manage the behavior of the systems when X and Y are small, respectively. The function H in this case is $H(x,y) = \mathcal{L}V(x,y)$. We refer to [4, Theorems 5.1 and 5.5] for more details.

From Assumption 2.5, it is possible to define the H-exponent of X as in [4, Definition 4.2].

Definition 2.8. *For V and H as in Hypothesis* 2.5*, we set*

$$\Lambda^-(H) = -\sup\{\mu H,\ \mu \in \mathcal{P}_{erg}(\mathcal{M}_0)\}$$

and

$$\Lambda^+(H) = -\inf\{\mu H,\ \mu \in \mathcal{P}_{erg}(\mathcal{M}_0)\}.$$

*We say that X is H-*persistent *if $\Lambda^-(H) > 0$ and that X is H-*nonpersistent *if $\Lambda^+(H) < 0$.*

The main results in [4] could be summed up as follows. If $\Lambda^-(H) > 0$, then $\mathcal{P}_{inv}(M_+)$ is nonempty and the family $\{\Pi_t, t \geq 0\}$ is tight in $\mathcal{M}_+$. Furthermore, the process X is *stochastically persistent* (see [27])) in the sense that, for all $\varepsilon > 0$, there exists a compact subset K of $\mathcal{M}_+$ such that, for all $x \in \mathcal{M}_+$,

$$\mathbb{P}_x\left(\liminf \Pi_t(K) \geq 1 - \varepsilon\right) = 1.$$

On the contrary, when $\Lambda^+(H) < 0$, X_t converges to $\mathcal{M}_0$ exponentially fast. (This is not yet proven in [4], but one can look at the thesis of the second author [30, section 1.3] for a proof in the special case where $\mathcal{M}_0$ is compact, relying on the proof made in [6].) However, the critical case where $\Lambda^+(H) = 0$ is not investigated. It is known from the deterministic case that in general, the information that $\Lambda^+(H) = 0$ is not sufficient to conclude on the long term behavior of the process. (One can think to the stability of an equilibrium point for a dynamical system, when the Jacobian matrix of the vector field at that point has eigenvalues with null real part.)

We now state the result of this note, which follows from a basic argument.

PROPOSITION 2.9. *Assume that if $\mathcal{P}_{inv}(\mathcal{M}_+)$ is nonempty, then there exists $\mu \in \mathcal{P}_{inv}(\mathcal{M}_+)$ and $\pi \in \mathcal{P}_{inv}(\mathcal{M}_0)$ such that*

$$\mu H > \pi H. \tag{3}$$

Then $\Lambda^+(H) > 0$.

Proof. Assume that $\mathcal{P}_{inv}(\mathcal{M}_+)$ is nonempty. Let $\mu \in \mathcal{P}_{inv}(\mathcal{M}_+)$ satisfying (3) for some $\pi \in \mathcal{P}_{inv}(\mathcal{M}_0)$, then $\mu H > -\Lambda^+(H)$. By [4, Lemma 7.5], since $\mu \in \mathcal{P}_{inv}(\mathcal{M}_+)$, we must have $\mu H = 0$. (Note that the proof of this fact in [4] does not require the process to be H-persistent.) This proves that $\Lambda^+(H) > 0$. ∎

We get the following immediate corollary.

COROLLARY 2.10. *Assume that the hypothesis in Proposition* 2.9 *holds. If $\Lambda^+(H) = 0$, $\mathcal{P}_{inv}(\mathcal{M}_+)$ is empty and all weak-* limit points of Π_t lie almost surely in $\mathcal{P}_{inv}(\mathcal{M}_0)$. In particular, if $\mathcal{P}_{inv}(\mathcal{M}_0) = \{\pi\}$, then for all bounded continuous functions $f : \mathcal{M} \to \mathbb{R}$,*

$$\lim_{t\to+\infty} \frac{1}{t}\int_0^t f(X_s)ds = \pi f. \tag{4}$$

Remark 2.11. Actually, one can prove that (4) holds for all $f : \mathcal{M} \to \mathbb{R}$ such that the map $\frac{W}{1+|f|}$ is proper, where W satisfies Assumption 2.4 (see [4, Lemma 9.1]).

Thus, the idea is that if H is strictly bigger on $\mathcal{M}_+$ than on $\mathcal{M}_0$ and if $\Lambda^+(H) = 0$, then the process goes on average to extinction. Rather than giving abstract conditions ensuring that (3) holds, we provide in the next sections five examples on which we prove (3) with different methods that can be easily reproduced for other models.

3. Applications. To illustrate the applicability of our method, we examine the critical cases in five stochastic models in ecology and epidemiology. The four first examples come from the literature, where the case $\Lambda = 0$ has not been treated. The last example is new. It should be emphasized that each model requires to be treated differently before Proposition 2.9 can be applied. New distinct techniques are therefore introduced to handle each model, especially when usual comparison arguments are not straightforwardly applicable.

3.1. SIRS model with switching. For some diseases such as influenza, an individual's immunity may wane over time after recovery. SIRS models, which are often used for this type of disease, describe the course of the transmission, recovery, and loss of immunity. Stochastic SIRS models have been studied extensively over the last decade. However, few papers have successfully classified the asymptotic behaviors of the models. In this section, we apply our method to a SIRS model with random switching that was studied in [24]. We first describe the process. Let N be a positive integer, and set $\mathcal{E} = \{1, \dots, N\}$. For $k \in \mathcal{E} = \{1, \dots, N\}$ let F^k be the vector field defined on $\mathbb{R}^3$ by

$$F^k(S, I, R) = \begin{pmatrix} \Lambda - \mu S + \lambda_k R - \beta_k S G_k(I) \\ \beta_k S G_k(I) - (\mu + \alpha_k + \delta_k) I \\ \delta_k I - (\mu + \lambda_k) R \end{pmatrix}, \tag{5}$$

where G_k is a regular function such that $G_k(0) = 0$. The reader is referred to [24] for the epidemiological interpretation of the different constants. Let $(\alpha_t)_{t\geq 0}$ be an irreducible Markov chain on $\mathcal{E}$. We denote by $p = (p_1, \dots, p_N)$ its unique invariant probability measure. We consider the process $(Z_t)_{t\geq 0} = (X_t, \alpha_t)_{t\geq 0}$ with $X_t = (S_t, I_t, R_t) \in \mathbb{R}^3_+$ evolving according to

$$\frac{dX_t}{dt} = F^{\alpha_t}(X_t). \tag{6}$$

The process Z is a *piecewise deterministic Markov process* (PDMP) as introduced in [11], and belongs to the more specific class of PDMPs recently studied in [3] and [5] (see also [6, 20, 7] and [17] for PDMP models in ecology or epidemiology).

Remark 3.1. In [24], β is the only parameter allowed to depend on k. The general case where the other constants and the function G can depend on k has been treated in [31].

We make the following assumptions, which are taken from [24].

Hypothesis 3.2.
(i) For all k, $G_k : \mathbb{R}_+ \to \mathbb{R}_+$ is C^2 with $G_k(0) = 0$ and $0 < G_k(I) \leq G_k'(0)I$ for $I > 0$.
(ii) For all k, if $\beta_k \frac{\Lambda}{\mu} G_k'(0) - (\mu + \alpha_k + \delta_k) > 0$, then F^k admits an equilibrium point $x^* \in \mathcal{M}_+$ which is accessible from $\mathcal{M}_+$.

We consider the process on the space $\mathcal{M} := K \times \mathcal{E}$, where $K = \{(s, i, r) \in \mathbb{R}^3_+\ s + i + r \leq \frac{\Lambda}{\mu}\}$. The set $K_0 = \{(s, i, r) \in K : i = 0\}$ is invariant for the F^k, and thus the set $\mathcal{M}_0 = K_0 \times E$ is invariant for Z. On this set, it is not hard to check that X converges almost surely to $(S^*, 0, 0)$, where $S^* = \frac{\Lambda}{\mu}$. Thus, the unique invariant probability measure of Z on $\mathcal{M}_0$ is $\delta^* \otimes p$, where δ^* is the Dirac mass at $(S^*, 0, 0)$. Consider the function $V : \mathcal{M}_+ \times \mathcal{E} \to \mathbb{R}_+$ given by

$$V(s, i, r, k) = \log \frac{\Lambda}{\mu} - \log i \quad \text{for all } (s, i, r, k) \in \mathcal{M}_+ \times E.$$

Define also the function $H : \mathcal{M} \times E \to \mathbb{R}$ by $H(s, i, r, k) = (\mu + \alpha_k + \delta_k - \beta_k s \tilde{G}_k(i))$, where $\tilde{G}_k$ is given by

$$\tilde{G}_k(i) = \begin{cases} \frac{G_k(i)}{i} & \text{if } i \neq 0, \\ G'(0) & \text{if } i = 0. \end{cases}$$

It is not hard to check that V and H satisfy assumption 2.5. Moreover, we have for $\pi = \delta^* \otimes p$,

$$\pi H = \sum_{k\in\mathcal{E}} p_k \left(\mu + \alpha_k + \delta_k - \beta_k \frac{\Lambda}{\mu} G_k'(0) \right)$$

for $k \in \mathcal{E} = \{1, \ldots, N\}$. As in [24], we set

$$R_0 = \frac{\sum_k p_k \beta_k \frac{\Lambda}{\mu} G_k'(0)}{\sum_k p_k(\mu + \alpha_k + \delta_k)}.$$

Note that $R_0 < 1$ (respectively, $R_0 > 1$, $R_0 = 1$) if and only if $\pi H > 0$ (respectively, $\pi H < 0$, $\pi H = 0$). The behavior of the process when $R_0 < 1$ or $R_0 > 1$ is studied in [24] (see also [31] for an alternative and more general proof). With our method, one can prove the following.

Proposition 3.3. *Assume that $R_0 = 1$. Then, for all $(s, i, r, k) \in \mathcal{M}$, $\mathbb{P}_{(s,i,r,k)}$-almost surely,*

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t S_u du = S^*,$$

and

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t (I_u+R_u)du = 0.$$

Proof. We show that when $\mathcal{P}_{inv}(\mathcal{M}_+)$ is nonempty, then for all $\mu^*\in\mathcal{P}_{inv}(\mathcal{M}_+)$, one has

$$\mu^* H > \pi H.$$

For convenience, we write C_k for $\mu+\alpha_k+\delta_k$. By Assumption 3.2, we have

$$H(s,i,r,k)\geq C_k-\beta_k G'(0)s,$$

and thus

$$\mu^* H \geq \sum_{k\in E} p_k C_k - \sum_{k\in E}\beta_k G_k'(0)\int_{M_+} s d\mu_k^*(s,i,r),$$

where μ_k^* is the measure of total mass p_k defined on $\mathcal{M}$ by $\mu_k^*(A)=\mu^*(A\times\{k\})$. Note that as $i>0$ on $\mathcal{M}_+$ and that for $(s,i,r)\in\mathcal{M}$, $s+i+r\leq S^*$, then for all $(s,i,r)\in\mathcal{M}_+$, $s<S^*$. In particular,

$$\int_{\mathcal{M}_+} s d\mu_k^*(s,i,r) < p_k S^*,$$

which yields

$$\mu^* H > \sum_{k\in\mathcal{E}} p_k C_k - \sum_{k\in\mathcal{E}} p_k\beta_k G_k'(0)S^* = \pi H.$$

This proves by Corollary 2.10 that if $R_0=1$, then $\mathcal{P}_{inv}(\mathcal{M}_+)$ is empty and for all bounded measurable function $f:\mathcal{M}\times E\to\mathbb{R}$,

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t f(S_u,I_u,R_u,r_u)du = \sum_k p_k f(S^*,0,0,k). \qquad \square$$

3.2. Stochastic Rosenzweig–MacArthur. Although the results below can also be obtained for a more general predator-prey model, we consider in detail the stochastic Rosenzweig–MacArthur predator-prey model, which was introduced in section 2, as a specific model to illustrate our method. To be precise, consider the system

$$(7)\qquad \begin{cases} dX_t = X_t\left(1-\frac{X_t}{K}-\frac{Y_t}{1+X_t}\right)dt+\varepsilon X_t dB_t,\\ dY_t = Y_t\left(-\alpha+\frac{X_t}{1+X_t}\right)dt. \end{cases}$$

In this case, $\mathcal{M}=\mathbb{R}_+^2:=\{(x,y)\in\mathbb{R}^2 : x,y\geq 0\}$. We set $\mathcal{M}_0^{\mathrm{x}}=\{(x,y)\in\mathcal{M} : x=0\}$, $\mathcal{M}_0^{\mathrm{y}}=\{(x,y)\in\mathcal{M} : y=0\}$, and $\mathcal{M}_0=\mathcal{M}_0^{\mathrm{x}}\cup\mathcal{M}_0^{\mathrm{y}}$. We also let $\mathcal{M}_+^{\mathrm{x}}=\mathcal{M}\setminus\mathcal{M}_0^{\mathrm{x}}$, $\mathcal{M}_+^{\mathrm{y}}=\mathcal{M}\setminus\mathcal{M}_0^{\mathrm{y}}$, and $\mathcal{M}_+=\mathcal{M}\setminus\mathcal{M}_0$. We also define the invasion rate of species x and y, respectively, as

$$\lambda_1(x,y)=\left(1-\frac{x}{K}-\frac{y}{1+x}\right)-\frac{\varepsilon^2}{2} \quad\text{and}\quad \lambda_2(x,y)=-\alpha+\frac{x}{1+x}.$$

By [4, Theorem 5.5], if $\varepsilon^2>2$, then for any initial condition, one has $(X_t,Y_t)\to 0$ as $t\to\infty$. Thus, we assume now that $\varepsilon^2<2$. In that case, the process is H-persistent with respect to $\mathcal{M}_0^{\mathrm{x}}$. Indeed, in that situation, $\mathcal{P}_{erg}(\mathcal{M}_0^{\mathrm{x}})=\{\delta_0\}$, where δ_0 is the

Dirac mass at 0 and $\delta_0\lambda_1 = 1 - \frac{\varepsilon^2}{2} > 0$. Hence, the condition [4, Theorem 5.1(ii)] is satisfied. In particular, every limit point of $(\Pi_t)_{t\geq 0}$ lies almost surely in $\mathcal{P}_{inv}(\mathcal{M}_+^{\mathrm{x}})$. Moreover, on $\mathcal{M}_+^{\mathrm{x}} \cap \mathcal{M}_0^{\mathrm{y}}$, the process admits a unique invariant probability measure denoted by μ_{x} (see [4, section 5.2]).

It is easily seen that $\mathcal{P}_{erg}(\mathcal{M}_0) = \{\delta_0, \mu_{\mathrm{x}}\}$. We set

$$\Lambda(\varepsilon, K, \alpha) = \mu_{\mathrm{x}}(\lambda_2) = \int_0^{+\infty} \frac{x}{1+x} d\mu_{\mathrm{x}}(x) - \alpha.$$

By [4, Theorem 5.5], if $\Lambda(\varepsilon, K, \alpha) > 0$, then the process is stochatistically persistent with respect to $\mathcal{M}_0$ and admits a unique invariant probability measure μ^* on $\mathcal{M}_+$, while if $\Lambda(\varepsilon, K, \alpha) < 0$, Y_t converges to 0. We now prove the following proposition for the critical case.

Proposition 3.4. *If $\Lambda(\varepsilon, K, \alpha) = 0$, then for all $(x, y) \in M_+$, one has $\mathbb{P}_{(x,y)}$-almost surely,*

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T Y_s ds = 0$$

and

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T X_s ds = \int_0^{+\infty} x d\mu_{\mathrm{x}}(x) = K\left(1 - \frac{\varepsilon^2}{2}\right).$$

Proof. We prove that if $\mathcal{P}_{inv}(\mathcal{M}_+^{\mathrm{y}})$ is nonempty, then for all $\mu^* \in \mathcal{P}_{inv}(\mathcal{M}_+^{\mathrm{y}})$, one has $\mu^* H > \mu_{\mathrm{x}} H$, where

$$H(x, y) = H_1(x, y) - \lambda_2(x, y)$$

with

$$H_1(x, y) = \frac{1}{1+x+y}\left(x - \alpha y - \frac{x^2}{K}\right) - \frac{\varepsilon^2 x^2}{2(1+x+y)}.$$

We set, for $(x, y) \in \mathcal{M}_+$, $V(x, y) = \log(1 + x + y) - \log x$. We can see that (V, H) satisfy Assumption 2.5. Moreover, we have $\mathcal{L}[\log(1 + x + y)] = H_1(x, y)$, and then by [4, Remark 19], we must have $\nu H_1 = 0$ for any $\nu \in \mathcal{P}_{inv}(\mathcal{M})$. As a result,

$$\nu H = -\nu\lambda_2 \text{ for any } \nu \in \mathcal{P}_{inv}(\mathcal{M}).$$

Remark 3.5. In the framework of [4], it would have been natural to take for V any function coinciding with $-\log x$ for x small enough, so that $H = -\lambda_2$ near $\mathcal{M}_0$, because it is sufficient to know H on the boundary $\mathcal{M}_0$. However, to apply our method, it is required to compare πH and μH for $\mu \in \mathcal{P}_{inv}(\mathcal{M}_+)$, and thus it is necessary to know H on the whole $\mathcal{M}_+$. Thus the idea is to take $V = V_1 + V_2$ and $H = H_1 + H_2$, with $V_2(x, y) = -\log x$, $H_2 = -\lambda_2$, V_1 defined on all $\mathcal{M}$ so that V is nonnegative, $\mathcal{L}V_1 = H_1$ and $\nu H_1 = 0$ for all $\nu \in \mathcal{P}_{inv}(\mathcal{M})$ (see [4, Remarks 11 and 19 and Proposition 4.13]). We use a similar trick in subsection 3.3.

To continue the proof, note that on $\mathcal{M}_0^{\mathrm{x}} \cap \mathcal{M}_+^{\mathrm{y}}$, Y_t converges exponentially fast to 0. Thus, it holds that $\mathcal{P}_{inv}(\mathcal{M}_+^{\mathrm{y}}) = \mathcal{P}_{inv}(\mathcal{M}_+)$. Moreover, by Theorem 5.5 in [4], if $\mathcal{P}_{inv}(\mathcal{M}_+)$ is nonempty, it reduces to a unique element, which we denote by μ^*, and μ^* has a positive density with respect to the Lebesgue measure. This implies by Birkhoff's ergodic theorem that for all $(x, y) \in \mathcal{M}_+$,

$$\mu^*\lambda_2 = \lim_{T\to\infty} \frac{1}{T} \int_0^T \lambda_2(X_s, Y_s) ds, \quad \mathbb{P}_{x,y}\text{-almost surely.}$$

We let $\hat{X}$ be the solution of the reduced system on $\mathcal{M}_0^{\mathrm{y}}$. That is,

$$d\hat{X}_t = \hat{X}_t\left(1 - \frac{\hat{X}_t}{K}\right)dt + \varepsilon dB_t. \tag{8}$$

By the comparison theorem, if $X_0 = \hat{X}_0$, then $X_t \le \hat{X}_t$ for all $t \ge 0$. The idea is now to write

$$\mu^* H = -\mu^*\lambda_2 = -\lim_{T\to\infty}\frac{1}{T}\int_0^T \lambda_2(\hat{X}_s, 0)ds - \lim_{T\to\infty}\frac{1}{T}\int_0^T \Big(\lambda_2(X_s, Y_s) - \lambda_2(\hat{X}_s, 0)\Big)\, ds$$

and to prove that the first term is $\mu_{\mathrm{x}} H$ and the second one is positive.

By [4, Theorem 5.1(i)], we have $\mu_{\mathrm{x}}(\lambda_1) = 0$. Moreover, the process $\hat{X}$ on $\mathcal{M}_0^{\mathrm{y}}$ is persistent with respect to $\mathcal{M}_0^{\mathrm{x}} \cap \mathcal{M}_0^{\mathrm{y}}$. Thus, for all $x > 0$, one has

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T \lambda_2(\hat{X}_s, 0)ds = \mu_{\mathrm{x}}\lambda_2$$

and

$$\lim_{t\to\infty}\frac{1}{T}\int_0^T \lambda_1(\hat{X}_s, 0)ds = \mu_{\mathrm{x}}(\lambda_1) = 0,$$

which gives

$$\lim_{t\to\infty}\frac{1}{T}\int_0^T \hat{X}_s = K\left(1 - \frac{\varepsilon^2}{2}\right).$$

On the other hand, since (X, Y) is persistent with respect to $\mathcal{M}_0^{\mathrm{x}}$, one has

$$\lim_{t\to\infty}\frac{1}{T}\int_0^T \lambda_1(X_s, Y_s)ds = \mu^*\lambda_1 = 0,$$

which leads to

$$\lim_{t\to\infty}\frac{1}{T}\int_0^T X_s = K\left(1 - \frac{\varepsilon^2}{2}\right) - \int_0^{+\infty}\frac{y}{1+x}d\mu^*(x, y).$$

Now, due to the fact that $\mu^*(M_+) = 1$, one has

$$\bar{y} := \int_0^{+\infty}\frac{y}{1+x}d\mu^*(x, y) > 0$$

and thus

$$\lim_{t\to\infty}\frac{1}{T}\int_0^T (\hat{X}_s - X_s)ds = \bar{y} > 0.$$

From this we have

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T \Big(\lambda_2(X_s, Y_s) - \lambda_2(\hat{X}_s, 0)\Big)\, ds < 0.$$

Indeed, let $C > 0$ such that

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T (\hat{X}_s - X_s)\mathbb{1}_{\{\hat{X}_s \le C\}}ds \ge \frac{\bar{y}}{2}.$$

Then, it is easily seen that there exists $c > 0$ such that for all $0 \leq x \leq \hat{x} \leq C$, and all $y \geq 0$, one has $\lambda_2(x, y) - \lambda_2(\hat{x}, 0) \leq -c(\hat{x} - x)$. In particular, by the monotonicity of H and the fact that $X_s \leq \hat{X}_s$ for all $s \geq 0$, we have

$$
\begin{aligned}
&\lim_{T\to\infty} \frac{1}{T} \int_0^T \left(\lambda_2(X_s, Y_s) - \lambda_2(\hat{X}_s, 0)\right) ds \\
&\leq \lim_{T\to\infty} \frac{1}{T} \int_0^T \left(\lambda_2(X_s, Y_s) - \lambda_2(\hat{X}_s, 0)\right) \mathbb{1}_{\{\hat{X}_s \leq C\}} ds \\
&\leq -\lim_{t\to\infty} \frac{c}{T} \int_0^T (\hat{X}_s - X_s) \mathbb{1}_{\{\hat{X}_s \leq C\}} ds \\
&\leq -\frac{c\bar{y}}{2}.
\end{aligned}
$$

We conclude that $\mu^* H = -\mu^* \lambda_2 \geq -\mu_x \lambda_2 + \frac{c\bar{y}}{2} = \mu_x H + \frac{c\bar{y}}{2} > \mu_{\mathrm{x}} H$. This proves that when $\Lambda(\varepsilon, K, \alpha) = 0$, $\mathcal{P}_{inv}(\mathcal{M}_+)$ is empty. Moreover, we know that the process is persistent with respect to $\mathcal{M}_0^{\mathrm{x}}$. Putting this together, the only possible limit point for $(\Pi_t)_{t\geq 0}$ is μ_{x}. Furthermore, since the maps $(x, y) \mapsto \frac{(x+y)^2}{1+y}$ and $(x, y) \mapsto \frac{(x+y)^2}{1+x}$ are proper, Corollary 2.10 and Remark 2.11 imply that

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T Y_s ds = 0$$

and

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T X_s ds = \int_0^{+\infty} x d\mu_{\mathrm{x}}(x) = K\left(1 - \frac{\varepsilon^2}{2}\right). \qquad \square$$

3.3. A stochastic model in spatially heterogeneous environments. In this section, we consider the example treated in [19] of a population submitted to random fluctuations of the environment and to spatio-temporal heterogeneity. The model aims to analyze the effect of both spatial and temporal variations to the evolution of the species; see [14] and [19] for more biological interpretations. In that setting, the space is divided into n patches, and the dynamics of the population within a patch follows a logistic SDE. There is also dispersal of the population, that is, individuals can move from one patch to the other. The precise model is the following. Let $X_t = (X_t^1, \ldots, X_t^n)$ be the vector of abundance in each patch at time t, then X satisfies the SDE

$$dX_t^i = \left[X_t^i \left(a_i - b_i(X_t^i)\right) + \sum_{j=1}^{n} D_{j,i} X_t^j\right] dt + X_t^i dE_t^i, \tag{9}$$

where $a_i > 0$ is the per capital growth rate in patch i, $b_i : \mathbb{R}_+ \to \mathbb{R}_+$ is the competition term in patch i, and $D_{j,i} \geq 0$ is for $j \neq i$, the dispersal rate of patch j to patch i and $E = \Gamma^T B$, where Γ is a square $n \times n$ matrix and $B = (B^1, \ldots, B^n)$ is a standard Brownian motion. We also set $D_{i,i} = -\sum_{j\neq i} D_{j,i}$ and $\Sigma = \Gamma^T \Gamma$.

We work under the following assumptions, made in [19].

Hypothesis 3.6.

1. For each $i \in \{1, \ldots, n\}$, $b_i : \mathbb{R}_+ \to \mathbb{R}_+$ is locally Lipschitz and vanishes only at 0, and there exist constants γ_b and M_b such that, for all $x \in \mathbb{R}_+^n$ with

$\sum_i x_i \geq M_b$, one has

$$\frac{\sum_{i=1}^n x_i(b_i(x_i) - a_i)}{\sum_{i=1}^n x_i} > \gamma_b.$$

2. The matrix D is irreducible.
3. The matrix Σ is nonsingular.

These assumptions guarantee the existence of a unique strong solution to (9), which moreover stays in $\mathbb{R}_+^n$ if $X_0 \in \mathbb{R}_+^n$. As in [19], we introduce the decomposition of the process: for any $x_0 \neq 0$ and $t \geq 0$, we set $S_t = \sum_i X_t^i$ and $Y_t^i = X_t^i / S_t$. By Ito's formula, it can be shown that (S_t, Y_t) evolves according to

$$\begin{cases} dY_t = \left[\mathrm{Diag}(Y_t) - Y_t Y_t^T\right] \Gamma^T dB_t + D^T Y_t dt \\ \qquad\qquad + \left[\mathrm{Diag}(Y_t) - Y_t Y_t^T\right] (a - \Sigma Y_t - b(S_t Y_t))\, dt, \\ dS_t = S_t \left(a - b(S_t Y_t)\right)^T Y_t dt + S_t Y_t^T \Gamma^T dB_t, \end{cases} \tag{10}$$

where $Y_t = (Y_t^1, \ldots, Y_t^n)$ lies in the simplex

$$\Delta = \{x = (x_1, \ldots, x_n) \in \mathbb{R}_+^n \ : x_1 + \cdots + x_n = 1\},$$

and $a := (a_1, \ldots, a_n)$, $b(x) := (b_1(x), \ldots, b_n(x))$. It is now possible to extend equation (10) on $\{0\} \times \Delta$ by setting $S_t = 0$ and

$$dY_t = \left[\mathrm{Diag}(Y_t) - Y_t Y_t^T\right] \Gamma^T dB_t + D^T Y_t dt + \left[\mathrm{Diag}(Y_t) - Y_t Y_t^T\right] (a - \Sigma Y_t)\, dt. \tag{11}$$

If we let $\tilde{X}_t$ be the solution to

$$d\tilde{X}_t^i = \left[a_i \tilde{X}_t^i + \sum_{j=1}^n D_{j,i} \tilde{X}_t^j\right] dt + \tilde{X}_t^i dE_t^i, \tag{12}$$

and $\tilde{S}_t = \sum_i \tilde{X}_t^i$, then

$$d\tilde{S}_t = \tilde{S}_t a^T \tilde{Y}_t dt + \tilde{S}_t \tilde{Y}_t^T \Gamma^T dB_t \tag{13}$$

with $\tilde{Y} = Y$ subjected to (11). It is proven in [15] that $\tilde{Y}$ admits a unique invariant probabilty measure π on Δ. Set

$$r = \int_\Delta \left(a^T y - \frac{1}{2} y^T \Sigma^T y\right) d\pi(y). \tag{14}$$

In [19], the authors show that the sign of r determines the long term behavior of X: if $r < 0$, then the population abundance in each patch converges to 0 exponentially fast, while if $r > 0$, the process X admits a unique invariant probability measure on $\mathbb{R}_{++}^n := \{x \in \mathbb{R}_+^n \ : x_i > 0\}$ and the law of X converges polynomially fast to this stationary distribution. The case $r = 0$ is not treated and left in the discussion as an open question.

We show now that our method enables us to handle the critical case $r = 0$.

Proposition 3.7. *If $r = 0$, then, for all i, for all $x \in \mathbb{R}_+^n$, $\mathbb{P}_x$-almost surely*

$$\lim_{t \to \infty} \frac{1}{t} \int_0^t X_s^i ds = 0.$$

Proof. First, let us write the process in our background. We consider the process $(Z_t)_{t\geq 0} = (S_t, Y_t)_{t\geq 0}$ defined on $\mathcal{M} = \mathbb{R}_+ \times \Delta$, and evolving according to (10) on $\mathcal{M}_+ = \mathbb{R}_+^* \times \Delta$ and according to (11) on $\mathcal{M}_0 = \{0\} \times \Delta$. Proposition A.1 in [19] implies that Assumption 2.2 is satisfied under Assumptions 3.6.

One can check that for a function $f : \mathcal{M} \to \mathbb{R}$, of class C^2 with bounded support, the generator $\mathcal{L}$ of Z is given by

$$\mathcal{L}f(s,y) = \frac{\partial}{\partial s} f(s,y) s\,(a - b(sy))^T\, y + \frac{1}{2}\frac{\partial^2}{\partial s^2} f(s,y) s^2 y^T \Sigma y + Af(s,y),$$

where Af is a sum of terms, each of them involving at least one derivative of f with respect to one of the coordinates of y. In particular, if $f(s,y) = g(s)$ for some function g, one has

$$\mathcal{L}f(s,y) = g'(s)s\,(a - b(sy))^T\, y + \frac{1}{2} g''(s) s^2 y^T \Sigma y.$$

Let $\varepsilon > 0$, and set $g(s) = (1+s)^{1+\varepsilon}$ and $f(s,y) = g(s)$. Then, we get (formally) that

$$\mathcal{L}f(s,y) = (1+\varepsilon) f(s,y) \left[\frac{s}{1+s}(a - b(sy))^T y + \left(\frac{s}{1+s}\right)^2 \frac{1}{2}\varepsilon y^T \Sigma y \right],$$

which by Assumption 3.6 implies that

$$\mathcal{L}f(s,y) \leq -\alpha f(s,y) + C,$$

where $\alpha = \gamma_b - \frac{1}{2}\varepsilon\|\Sigma\|$ is positive for ε small enough, and $C = \sup_{(s,y)\in[0,M_b]\times\Delta} \mathcal{L}f(s,y)$ is finite. From this, it is possible to prove that Assumption 2.4 is satisfied for $W(s,y) = (1+s)^{1+\varepsilon}$, provided ε is small enough.

Next, we prove that Assumption 2.5 is satisfied. We define two functions on $\mathcal{M}$:

$$H_1(s,y) = \frac{s}{1+s}\,(a - b(sy))^T\, y - \frac{1}{2}\frac{s^2}{(1+s)^2} y^T \Sigma y$$

and

$$H_2(s,y) = (a - b(sy))^T\, y - \frac{1}{2} y^T \Sigma y.$$

We define V on $\mathcal{M}_+$ by setting $V(s,y) = \log(1+s) - \log s$. By definition of V and Ito's formula,

$$\mathcal{L}V(s,y) = H_1(s,y) - H_2(s,y).$$

It is not hard to check that the functions V and H so defined satisfy Assumption 2.5. We have from Ito's formula that

$$\lim_{T\to\infty} \Pi_T^z H_1 = \lim_{T\to\infty} \frac{\mathbb{E}_z \log(1+S_T) - \log(1+s)}{T} = 0, z \in \mathcal{M},$$

due to [19, Lemma A.2]. As a result, $\nu H_1 = 0$ for any invariant probability measure ν on $\mathcal{M}$ of $(Z_t)_{t\geq 0}$. Subsequently, we have, $r = -\pi H_2 = \pi H$, where r is defined by (14), and by the ergodicity of $\tilde{Y}$ and (13), we have

$$r = \pi H = \lim_{t\to\infty} \frac{1}{t} \int_0^t \left[a^T \tilde{Y}_u - \tilde{Y}_u \Sigma \tilde{Y}_u \right] du = \lim_{t\to\infty} \frac{1}{t} \log \tilde{S}_t. \tag{15}$$

Now we assume that Z admits an ergodic invariant probability measure μ on $\mathcal{M}_+$. By the strong Feller property of X on $\mathbb{R}^n_{++}$, μ has to be unique, and thus the process is ergodic. In particular, we have

$$-\mu H = \mu H_2 = \lim_{t\to\infty} \frac{1}{t} \int_0^t \left[(a - b(S_u Y_u))^T Y_u - Y_u \Sigma Y_u \right] du.$$

Thus, to obtain the desired result that $\mu H > \pi H$, we will show that

$$\begin{aligned} &\lim_{t\to\infty} \frac{1}{t} \int_0^t \left[a^T \tilde{Y}_u - \tilde{Y}_u \Sigma \tilde{Y}_u \right] du \\ &\quad > \lim_{t\to\infty} \frac{1}{t} \int_0^t \left[(a - b(S_u Y_u))^T Y_u - Y_u \Sigma Y_u \right] du. \end{aligned} \tag{16}$$

While componentwise $a > a - b(S_u Y_u)$, (16) is not straightforward because $\widetilde{Y}_u - Y_u$ can be both negative and positive. The difficulty will be overcome by introducing an intermediate process to ease the comparison. For all $u \geq 0$, we set $\varsigma_u = \min_i b_i(S_u Y_u^i)$. Note that $\varsigma_u > 0$ by assumption on b. Now we introduce the process $\bar{X} = (\bar{X}^1, \dots, \bar{X}^n)$ defined by

$$d\bar{X}_t^i = \left[\bar{X}_t^i (a_i - \varsigma_t) + \sum_{j=1}^n D_{j,i} \bar{X}_t^j \right] dt + \bar{X}_t^i dE_t^i. \tag{17}$$

By a classical comparison argument for SDE (see, i.e., [10]) and the positivity of ς_t, we have $X_t^i \leq \bar{X}_t^i \leq \tilde{X}_t^i$ for all $t \geq 0$, provided the inequality holds at 0. We also set $\bar{S}_t = \bar{X}_t^1 + \cdots + \bar{X}_t^n$, and then $S_t \leq \bar{S}_t \leq \tilde{S}_t$. Finally, we introduce $\bar{Y} = \bar{X}/\bar{S}$, which is well defined as soon as $\bar{X}_0 \neq 0$. One can see that $\bar{S}$ and $\bar{Y}$ evolve according to

$$d\bar{S}_t = \bar{S}_t (a - \varsigma_t)^T \bar{Y}_t dt + \bar{S}_t \bar{Y}_t^T \Gamma^T dB_t, \tag{18}$$

$$d\bar{Y}_t = [\mathrm{Diag}(\bar{Y}_t) - \bar{Y}_t \bar{Y}_t^T] \Gamma^T dB_t + D^T \bar{Y}_t dt + [\mathrm{Diag}(\bar{Y}_t) - \bar{Y}_t \bar{Y}_t^T](a - \Sigma \bar{Y}_t - \varsigma_t \mathbb{1})\, dt, \tag{19}$$

where $\mathbb{1}$ is the vector with all components equal to 1. Now, since $\bar{Y}_t \in \Delta$, one has $(\mathrm{Diag}(\bar{Y}_t) - \bar{Y}_t Y_t^T)\mathbb{1} = 0$, and thus

$$d\bar{Y}_t = \left[\mathrm{Diag}(\bar{Y}_t) - \bar{Y}_t \bar{Y}_t^T\right] \Gamma^T dB_t + D^T \bar{Y}_t dt + \left[\mathrm{Diag}(\bar{Y}_t) - \bar{Y}_t \bar{Y}_t^T\right] \left(a - \Sigma \bar{Y}_t\right) dt, \tag{20}$$

and by the uniqueness of a strong solution to (11), $\bar{Y} = \tilde{Y}$ almost surely whenever $\bar{Y}_0 = \tilde{Y}_0$. Thus we have

$$\begin{aligned} \lim_{t\to\infty} \frac{1}{t} \log(\bar{S}_t) &= \lim_{t\to\infty} \frac{1}{t} \int_0^t \left(a^T \tilde{Y}_u - \tilde{Y}_u \Sigma \tilde{Y}_u - \varsigma_u \mathbb{1}^T \tilde{Y}_u \right) du \\ &= -\pi H - \lim_{t\to\infty} \frac{1}{t} \int_0^t \varsigma_u du \\ &= -\pi H - \int_{\mathcal{M}_+} \min_i b_i(s y_i) d\mu(y) \\ &= -\pi H - \bar{\varsigma}, \end{aligned}$$

where $\bar{\varsigma} = \int_{\mathcal{M}_+} \min_i b_i(sy_i)d\mu(y) > 0$ because $\mu(\mathcal{M}_+) = 1$. Now, since $\bar{S}_t \geq S_t$, we have

$$\begin{aligned} -\pi H - \bar{\varsigma} &= \lim_{t\to\infty} \frac{1}{t}\log(\bar{S}_t) \\ &\geq \lim_{t\to\infty} \frac{1}{t}\log(S_t) = -\mu H, \end{aligned}$$

which yields $\mu H \geq \pi H + \bar{\varsigma}$. Thus, one can apply Corollary 2.10 (and Remark 2.11): since the map $(s, y) \mapsto \frac{(1+s)^{1+\varepsilon}}{1+s}$ is proper, if $r = 0$, one has for all $(s, y) \in \mathcal{M}$, $\mathbb{P}_{s,y}$-almost surely

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t S_u du = 0,$$

or equivalently, for all i, for all $x \in \mathbb{R}^n_+$, $\mathbb{P}_x$-almost surely

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t X^i_s ds = 0. \qquad \square$$

3.4. Multigroup SIS model in a fluctuating environment. Here, we prove that the method used above also works in an SIS model with a random switching environment. More precisely, we consider a stochastic counterpart of a heterogeneous SIS model that was introduced by Lajmanovich and Yorke [23]. Note that SIS models (susceptible-infected-susceptible) describes the evolution of a disease that does not confer immunity. Suppose we are dealing with a heterogeneous population divided into d homogeneous groups. Let $x^i(t)$ be the number of susceptibles in the ith group, $y^i(t)$ the number of infecteds, $\alpha_i(t)$ the recovery rate, and c_i the total size of the ith subpopulation. Let β_{ij} be the contact rate of the ith group's susceptibles with the jth group's infecteds. Assume that (β_{ij}) is an irreducible matrix. Since $x^i(t)+y^i(t) = c_i$, we have

$$\frac{dy^i(t)}{dt} = -\alpha_i y^i(t) + \sum_{i,j} \beta_{ji} c_i y_j - \sum_{i,j} \beta_{ji} y_i y_j. \tag{21}$$

Nondimensionalizing the system by letting

$$X_t = (X^1(t), \ldots, X^d(t)) = (\frac{y^1(t)}{c_1}, \ldots, y^1(t)c_1),$$

we can transform the (21) to

$$\dot{X}_t = F(X_t),$$

where

$$F(x) = (C - \mathrm{Diag}(D))\, x - \mathrm{Diag}(x)Cx$$

and $C = (C_{i,j})$ be an irreducible $d \times d$ matrix with nonnegative entries and $D = (D_1, \ldots, D_d)$ a vector with positive entries. In [7], taking into account the fluctuations of the environment which is modeled by a switching process, we have considered a PDMP $U = (X, \alpha)$ on $[0, 1]^d \times \mathcal{E}$, where $\mathcal{E} = \{1, \ldots, N\}$ for some integer N and evolves as follows:

$$\frac{dX_t}{dt} = F^{\alpha_t}(X_t), \tag{22}$$

where α is a Markov chain on E and for all $k \in \mathcal{E}$, F^k is the vector field defined like F with C and D replaced by C^k and D^k, respectively, where C^k and D^k are a matrix and a vector as described above. We also set $A^k = C^k - \mathrm{Diag}(D^k)$. To analyze the long-term behavior of Z, we have done in [7] a polar decomposition: for $X_0 \neq 0$, we set $\rho_t = \|X_t\|$ and $\Theta_t = \frac{X_t}{\rho_t}$. Then $W = (\rho, \Theta, \alpha)$ is still a PDMP, evolving according to

$$\left\{ \begin{array}{l} \frac{d\Theta_t}{dt} = G^{\alpha_t}(\Theta_t), \\ \frac{d\rho_t}{dt} = \langle A^{\alpha_t} - \rho_t \mathrm{Diag}(\Theta_t) C\Theta_t, \Theta_t\rangle \rho_t, \end{array} \right. \tag{23}$$

where for all $i \in E$, G^i is the vector field on S^{d-1} defined by

$$G^i(\theta) = \big(A^i - \rho \mathrm{Diag}(\theta) C\big)\,\theta - \langle \big(A^i - \rho\mathrm{Diag}(\theta)C\big)\,, \theta\rangle \theta. \tag{24}$$

We set $\mathcal{M}_+ = \Psi([0,1]^d \setminus \{0\}) \times \mathcal{E}$, where $\Psi : \mathbb{R}^n \setminus \{0\} \to \mathbb{R}_+^* \times S^{d-1}$ is defined by $\Psi(x) = (\|x\|, \frac{x}{\|x\|})$. We also set $\mathcal{M}_0 = \{0\} \times S^{d-1} \times \mathcal{E}$, and then (23) can be defined on $\mathcal{M}_0$ be letting $\rho_t = 0$ for all $t \geq 0$ and

$$\frac{d\Theta_t}{dt} = A^{\alpha_t}\Theta_t - \langle A^{\alpha_t}, \Theta_t, \Theta_t\rangle \Theta_t. \tag{25}$$

We proved in [7, Proposition 2.13] that on $\mathcal{M}_0 \simeq S^{d-1} \times \mathcal{E}$, the process (Θ, α) admits a unique invariant probability π. We set

$$\Lambda = \int_{S^{d-1}\times\mathcal{E}} \langle A^i\theta, \theta\rangle d\pi(\theta, i).$$

It has also been proven that the functions $V : \mathcal{M}_+ \to \mathbb{R}_+$ and $H : \mathcal{M} \to \mathbb{R}$, defined by $V(\rho, \theta, i) = -\log(\rho)$ and by $H(\rho, \theta, i) = -\langle A^i\theta, \theta\rangle + \rho\langle \mathrm{Diag}(\theta)C\theta, \theta\rangle$, respectively, satisfy assumption 2.5. It is easily seen that $\Lambda = -\pi H$. With our method, together with the results in [7], we can now fully describe the behavior of U according to the sign of Λ.

THEOREM 3.8. *There are three possible asymptotic behaviors:*

1. *If $\Lambda < 0$, then for all $(x, i) \in [0,1]^d \times \mathcal{E}$, we have*

$$\mathbb{P}_{x,i}\left(\limsup \frac{\log \|X_t\|}{t} \leq \Lambda\right) = 1.$$

2. *If $\Lambda = 0$, then for all $(x, i) \in [0,1]^d \times \mathcal{E}$, we have*

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t \|X_s\| ds = 0 \quad \mathbb{P}_{x,i}\text{-a.s.}$$

and

$$\mathbb{P}_{x,i} - \lim_{t\to\infty} X_t = 0,$$

where $\mathbb{P}_{x,i} - \lim$ denotes the convergence in probability.

3. *If $\Lambda > 0$, then U admits a unique invariant probability measure μ on $(0,1]^d \times E$. Moreover, there exists a Wasserstein distance $\mathcal{W}$ and $r > 0$ such that, for all probability ν with $\nu(\{0\} \times \mathcal{E})$ and all $t \geq 0$,*

$$\mathcal{W}(\nu P_t, \mu) \leq e^{-rt}\mathcal{W}(\nu, \mu).$$

Proof. The case $\Lambda < 0$ is Theoerem 4.3 in [7], while $\Lambda > 0$ is Theorem 4.12 in [7].

To treat the case $\Lambda = 0$, we first prove that one can apply Proposition 2.9. We assume that W admits an invariant distribution μ on $\mathcal{M}_+$. For all $t > 0$, we define

$$\varsigma_t = \min_{1\le i\le d} X_t^i \left(\sum_j C_{i,j}^{\alpha_t} X_t^j \right),$$

and we let $\bar{X}$ be the solution to

$$\frac{d\bar{X}_t}{dt} = (A^{\alpha_t} - \varsigma_t I)\,\bar{X}_t, \tag{26}$$

where I is the identity matrix of size d. We also let Y be the solution to

$$\frac{dY_t}{dt} = A^{\alpha_t} Y_t. \tag{27}$$

By a comparison theorem for ordinary differential equations, we have $X_t^i \le \bar{X}_t^i \le Y_t$ for all $t \ge 0$, provided the inequality holds at time 0. Finally, let $\bar{\rho}_t = \|\bar{X}_t\|$, $\bar{\Theta}_t = \frac{\bar{X}_t}{\bar{\rho}_t}$, $\tilde{\rho}_t = \|Y\|_t$, and $\tilde{\Theta}_t = \frac{Y_t}{\tilde{\rho}_t}$. Then $\rho_t \le \bar{\rho}_t \le \tilde{\rho}_t$ and

$$\frac{d\bar{\Theta}_t}{dt} = (A^{\alpha_t} - \varsigma_t)\,\bar{\Theta}_t - \langle (A^{\alpha_t} - \varsigma_t)\,\bar{\Theta}_t, \bar{\Theta}_t\rangle \bar{\Theta}_t,$$

while $\tilde{\Theta}_t$ evolves according to (25). Now, since $\langle \bar{\Theta}_t, \bar{\Theta}_t\rangle = 1$ for all $t \ge 0$, we can see that $\bar{\Theta}_t$ is also driven by (25), and thus $\bar{\Theta}_t = \tilde{\Theta}_t$ for all $t \ge 0$ whenever $\bar{\Theta}_0 = \tilde{\Theta}_0$. On the other hand, one can check that

$$\lim_{t\to\infty} \frac{\log \bar{\rho}_t}{t} = \lim_{t\to\infty} \frac{1}{t}\int_0^t \langle (A^{\alpha_s} - \varsigma_t)\,\bar{\Theta}_s, \bar{\Theta}_s\rangle ds.$$

We also have

$$-\pi H = \lim_{t\to\infty} \frac{\log \tilde{\rho}_t}{t} = \lim_{t\to\infty} \frac{1}{t}\int_0^t \langle A^{\alpha_s} \tilde{\Theta}_s, \tilde{\Theta}_s\rangle ds.$$

Without loss of generality, one may assume that μ is ergodic, and therefore, one has for μ almost every $(\rho_0, \theta_0, i) \in \mathcal{M}_+$, $\mathbb{P}_{(\rho_0,\theta_0,i)}$-almost surely,

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t \varsigma_t = \int_{\mathcal{M}_+} \rho^2 \min_i \theta^i \left(\sum_j C_{i,j}^k \theta^j \right) d\mu(\rho,\theta,k) := \bar{\varsigma}.$$

Then, $\bar{\varsigma} > 0$ because on $\mathcal{M}_+$, $\rho > 0$ and $\mu(\{(\rho,\theta,i) \in \mathcal{M}_+ : \theta^i > 0\}) = 1$ since ∂S^{d-1} is transient for W. Thus, due to the fact that $\bar{\Theta}_t = \tilde{\Theta}_t$, we get for μ almost every $(\rho_0, \theta_0, i) \in \mathcal{M}_+$, $\mathbb{P}_{(\rho_0,\theta_0,i)}$-almost surely,

$$\lim_{t\to\infty} \frac{\log \bar{\rho}_t}{t} = -\pi H - \bar{\varsigma},$$

which combined with

$$\lim_{t\to\infty} \frac{\log \rho_t}{t} = -\mu H \quad \mathbb{P}_{(\rho_0,\theta_0,i)}\text{-a.s.}$$

and $\bar{\rho}_t \geq \rho_t$ gives $\mu H \geq \pi H + \varsigma > \mu H$. Thus, by Proposition 2.9, $\Lambda > 0$. Hence, if $\Lambda = 0$, the unique stationary distribution of W is π, which is concentrated on $\mathcal{M}_0$. In particular, going back to the process U, its unique invariant distribution is $\delta_0 \otimes p$, where p is the unique stationary distribution of α on $\mathcal{E}$. In particular, for all $(x, i) \in [0,1]^d \times \mathcal{E}$, one has $\mathbb{P}_{x,i}$-almost surely that

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t \|X_s\| ds = 0. \tag{28}$$

To prove that X converges in probability to 0, we use results on monotone random dynamical systems due to Chueshov [10]. Let $\Omega = \mathbb{D}(\mathbb{R}_+, \mathcal{E})$ be the Skorhokhod space of càdlàg functions $\omega : \mathbb{R}_+ \to \mathcal{E}$, endowed with its Borel sigma field $\mathcal{F}$, and on which we define the shift $\boldsymbol{\Theta} = (\boldsymbol{\Theta}_t)_{t\geq 0}$ by

$$\boldsymbol{\Theta}_t(\omega)(s) = \omega(t+s).$$

We let $\mathbb{P}_p$ be a probability measure on $(\Omega, \mathcal{F})$ such that the canonical process I has the law of α starting from its ergodic probability measure p. Then, the process $\Psi(\omega, t)$ defined by

$$\begin{cases} \frac{d\Psi(t,\omega)x}{dt} = F^{\omega(t)}(\Psi(t,\omega)x), \\ \Psi(0,\omega)x = x \end{cases} \tag{29}$$

is a random dynamical system over the ergodic dynamical system $(\Omega, \mathcal{F}, \mathbb{P}_p, \boldsymbol{\Theta})$ (see, e.g., [2] for definitions and the thesis of the second author [30, section 1.4] for more details on random dynamical systems and links with PDMPs). Moreover, the properties of F make Ψ a monotone subhogeneous random dynamical system (see [7, section 4]) for which $\mathbb{1} = (1, \ldots, 1)$ is a superequilibrium. That is, $\Psi(t,\omega)\mathbb{1} \leq \mathbb{1}$ for all $t \geq 0$ and $\omega \in \Omega$ (see [10, Definition 3.4.1]). Moreover, for all $t \geq 0$ and all $\omega \in \Omega$,

$$\Psi(t,\omega)\left([0,1]^d \setminus \{0\}\right) \subset (0,1)^d.$$

Hence, it is easily to check that we can apply Proposition 5.5.1 in [10]. According to this result, either for all $x \in [0,1]^d$,

$$\lim_{t\to\infty} \Psi(t, \boldsymbol{\Theta}_{-t}\omega)x = 0 \tag{30}$$

or there exists an equilibrium $u(\omega) \gg 0$ such that for all $x > 0$ and all $\omega \in \Omega$,

$$\lim_{t\to\infty} \Psi(t, \boldsymbol{\Theta}_{-t}\omega)x = u(\omega). \tag{31}$$

Now, assume that (31) holds. In particular, by dominated convergence and invariance of $\mathbb{P}_p$ under $\boldsymbol{\Theta}$, one has on the one hand

$$\lim_{t\to\infty} \mathbb{E}_p\left(\|\Psi(t,\omega)x\|\right) = \mathbb{E}_p(\|u\|) > 0. \tag{32}$$

On the other hand, one can check that the law of X_t under $\mathbb{P}_{x,p}$ is the same as the law of $\Psi(t,\cdot)x$ under $\mathbb{P}_p$. In particular,

$$\mathbb{E}_{(x,p)}\left(\|X_t\|\right) = \mathbb{E}_p\left(\|\Psi(t,\omega)x\|\right). \tag{33}$$

Thus, (32) and (33) imply that

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t \mathbb{E}_{(x,p)}\left(\|X_s\|\right) ds = \mathbb{E}_p(\|u\|) > 0,$$

which is in contradiction (by dominated convergence) to (28). Hence, (30) holds. This and (33) yield that for all continuous maps $f : [0,1]^d \to \mathbb{R}$,

$$\lim_{t\to\infty} \mathbb{E}_{x,p}\left(f(X_t)\right) = f(0),$$

which implies that X_t converges in law, hence in probability, to 0, under $\mathbb{P}_{x,p}$. It is easily seen that one can now replace p by any starting point i. ∎

3.5. SEIR model with switching. SEIR models describe the dynamics of an infectious disease with which individuals experience a long incubation duration (the "exposed" compartment).

Susceptible individuals, when in contact with an infectious individual, may contract the disease with a given rate, and then go into the exposed disease state when they become infected but are not yet infectious themselves. Infectious individuals can transmit the disease and, after a period of time, enter the recovered phase and have permanent immunity to the disease. The classical SEIR model consists of the following differential equations for four classes of individuals (susceptible-exposed-infectious-recovered):

$$\begin{cases} \dot S = \Lambda - \gamma S - \beta SI, \\ \dot E = \beta SI - (\gamma+\delta)E, \\ \dot I = \delta E - (\gamma+\gamma_1)I, \\ \dot R = \gamma_1 I - \gamma R, \end{cases} \tag{34}$$

where $\Lambda, \gamma, \beta, \delta, \gamma_1$ are positive constant. This system has been used to model a number of infectious diseases, such as measles, mumps, and rubella. We refer to [29, 21] for details about this model and its variants. In contrast to stochastic SIR and SIRS models, which have been studied extensively, few papers deal with stochastic SEIR models because standard arguments used to treat SIR and SIRS models do not seem effective for SEIR models due to the extra compartment E. In this section, we wish to consider an SEIR model in a switching environment and fully characterize its long-term property. The model (35) below has not been studied in the literature. Let N be a positive integer, and set $\mathcal{E} = \{1, \dots, N\}$. Let $(\alpha_t)_{t\ge0}$ be a irreducible Markov chain on $\mathcal{E}$ and consider the following system:

$$\begin{cases} \dot S = \Lambda - \gamma S - \beta(\alpha_t) SI, \\ \dot E = \beta(\alpha_t) SI - (\gamma+\delta(\alpha_t))E, \\ \dot I = \delta(\alpha_t) E - (\gamma+\gamma_1(\alpha_t))I, \end{cases} \tag{35}$$

where the component R is removed because it does not affect the dynamics of the others.

Letting $U_t = E_t + I_t$ and $V_t = \frac{I_t}{U_t}$, $Z_t = (S_t, V_t, U_t, \alpha_t)$, we can rewrite (35) as

$$\begin{cases} \dot S = f_S(Z_t), \\ \dot V = f_V(Z_t), \\ \dot U = U_t f_U(Z_t), \end{cases} \tag{36}$$

where

$$f_S(z) = \Lambda - \gamma s - \beta(k) su(1-v),$$

$$f_U(z) = (\beta(k)s - \gamma_1(k) - \gamma)v - \gamma(1-v) = (\beta(k)s - \gamma_1(k))v - \gamma,$$

$f_V(z) = (\sigma(k)(1-v)-\gamma v-\gamma_1(k)v)-vf_U(z) = \sigma(k)(1-v)-\gamma_1(k)v-(\beta(k)s-\gamma_1(k))v^2,$

and $z = (s, u, v, k)$. For this system, we have

$$\mathcal{M} := \left\{z \in \mathbb{R}^2_+ \times [0,1] \times \mathcal{E} : s + u \leq \frac{\Lambda}{\gamma}\right\} \text{ and } \mathcal{M}_0 := \{z \in \mathcal{M} : u = 0\}. \tag{37}$$

In this model, $H(z) := -f_U(z)$ and $\mathcal{V}(z) := \log\frac{\Lambda}{\gamma} - \log u$ satisfy Assumption 2.5. Unlike the arguments in subsections 3.3 and 3.4, it does not seem practically possible to treat the critical case by introducing an intermediate process. Because the function $f_U(z)$ is increasing in s while $f_V(z)$ is decreasing in s, we introduce the following function:

$$\widetilde{H}(z) = -f_U(z) - \frac{f_V(z)}{v} = -\sigma(k)\frac{1-v}{v} + \gamma + \gamma_1(k). \tag{38}$$

If $U_0 = 0$, then $U_t = 0, t \geq 0$, and $\lim_{t\to\infty} S_t = \frac{\Lambda}{\gamma}$. Let $\widetilde{V}_t$ be the solution to

$$\dot{\widetilde{V}} = f_V\left(\frac{\Lambda}{\gamma}, 0, \widetilde{V}_t, \alpha_t\right).$$

Then, one can show that $(\widetilde{V}_t, \alpha_t)$ has a unique invariant measure π_V on $[0,1]\times\mathcal{E}$ (see, e.g., [6, Proposition 2.1] or [25]). Moreover, since $f_V(z) = \sigma(k) > 0$ if $z = (s,u,v,k)$ with $v = 0$, there exists $v_0 > 0$ such that $\liminf_{t\to\infty} V_t \geq v_0 > 0$ for any initial value $z \in \mathcal{M}$. As a result,

$$\mathbb{P}_z\left\{\lim_{T\to\infty}\frac{1}{T}\int_0^T \frac{f_V(Z_t)}{V_t} = \lim_{T\to\infty}\frac{\log V_t}{T} = 0\right\}, z \in \mathcal{M}.$$

Hence, for any invariant probability measure μ of $(Z_t)_{t\geq 0}$, we have

$$\int_{\mathcal{M}} \frac{f_V(z)}{v}\mu(dz) = 0, \text{ or equivalently } \mu H = \mu\widetilde{H}. \tag{39}$$

Then $\pi := \boldsymbol{\delta}_{(\frac{\Lambda}{\gamma},0)} \otimes \pi_V$ is the unique invariant measure on $\mathcal{M}_0$. By the ergodicity of $(\widetilde{V}_t, \alpha_t)$ and (38), (39) we have

$$\begin{aligned}\widetilde{\Lambda} := -\pi H &= -\,\pi\lim_{T\to\infty}\int_0^T f_U\left(\frac{\Lambda}{\gamma}, 0, \widetilde{V}_t, \alpha_t\right)dt \\ &= \lim_{T\to\infty}\int_0^T\left(\sigma(\alpha_t)\frac{1-\widetilde{V}_t}{\widetilde{V}_t} - \gamma - \gamma_1(\alpha_t)\right)dt.\end{aligned} \tag{40}$$

With $\mathcal{M}, \mathcal{M}_0$ defined in (37), we have the following theorem.

Theorem 3.9.

1. *If $\widetilde{\Lambda} < 0$, then for all $0 < \lambda < -\widetilde{\Lambda}$, there exist $\eta > 0$ and $r > 0$ such that, for all $z \in \mathcal{M}_+ := \mathcal{M}\setminus\mathcal{M}_0$ with $u \leq r$, we have*

$$\mathbb{P}_{x,i}\left(\limsup\frac{\log U_t}{t} \leq -\lambda\right) \geq \eta.$$

2. *If* $\widetilde{\Lambda} = 0$, *then for all* $z \in \mathcal{M}_+ \times \mathcal{E}$, *we have*

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T U_t dt = 0 \quad \mathbb{P}_z\text{-a.s.}$$

3. *If* $\widetilde{\Lambda} > 0$, *then* Z *is* H*-persistent and it admits an invariant probability measure on* $\mathcal{M}_+$.

Proof. We start by proving the first and third claims. For $\alpha \in \mathcal{E}$, we define the vector field

$$F^\alpha(s,e,i) = \begin{cases} \Lambda - \gamma s + \beta(\alpha) si, \\ \beta(\alpha)si - (\gamma + \delta(\alpha))e, \\ \delta(\alpha)e - (\gamma + \gamma_1(\alpha))i. \end{cases}$$

Then, letting $(X_t)_{t\geq 0} = (S_t, E_t, I_t)_{t\geq 0}$, we have $\dot{X}_t = F^{\alpha_t}(X_t)$. Note that $(\frac{\Lambda}{\gamma}, 0, 0)$ is a common equilibrium of the vector fields F^α and that the line $\mathbb{R}_+ \times \{(0,0)\}$ is invariant for each of the vector fields. This is exactly the setting of application of the results in [31]. The Jacobian matrix of F^α at $(\frac{\Lambda}{\gamma}, 0, 0)$ is given by

$$A^\alpha = \begin{pmatrix} -\gamma & -\beta(\alpha)\frac{\Lambda}{\gamma} & 0 \\ 0 & -(\gamma+\delta(\alpha)) & \beta(\alpha)\frac{\Lambda}{\gamma} \\ 0 & \delta(\alpha) & (\gamma + \gamma_1(\alpha)) \end{pmatrix}.$$

We let $D = (\gamma)$, $C^\alpha = (-\beta(\alpha)\frac{\Lambda}{\gamma}, 0)$ and

$$B^\alpha = \begin{pmatrix} -(\gamma+\delta(\alpha)) & \beta(\alpha)\frac{\Lambda}{\gamma} \\ \delta(\alpha) & (\gamma+\gamma_1(\alpha)) \end{pmatrix},$$

so that

$$A^\alpha = \begin{pmatrix} D & C^\alpha \\ 0 & B^\alpha \end{pmatrix}.$$

Finally, we define $\Lambda_D = -\gamma$ and $\Lambda_B = \int \langle B^\alpha \theta, \theta\rangle d\pi(\alpha,\theta)$, where π is the unique invariant probability measure of the process (Θ, α), where Θ is subjected to (25) with A^α replaced by B^α. (The uniqueness of π comes from the particular form of B; see [7, Proposition 2.13].) Then, $\widetilde{\Lambda} = \Lambda_B$. Indeed, $\tilde{\Lambda}$ is defined as the growth rate of U, which is the L^1-norm of (E, I), while Λ_B is defined as the growth rate of U_2, the L^2-norm of (E, I). By equivalence of the norm on $\mathbb{R}^2$, we must have $\tilde{\Lambda} = \Lambda_B$. The third claim is hence a direct application of Theorem 2.8 in [31]. The first claim follows from Theorem 2.7 in [31].

Now, we prove the second claim. Let assume that $\mathcal{P}_{inv}(\mathcal{M}_+)$ is nonempty with an ergodic measure μ.

Since $f_V(z) + \beta(k)\left(\frac{\Lambda}{\gamma} - s\right)v^2 = f_V(\frac{\Lambda}{\gamma}, 0, v, k)$ for any $z = (s,u,v,k) \in \mathcal{M}$, we have $V_t \geq \widetilde{V}_t$ given $V_0 \geq \widetilde{V}_0$. Let Z_t have the initial distribution μ and $\widetilde{V}_0 = V_0$.

By the ergodicity we have

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T (V_t - \widetilde{V}_t)dt = \int_{\mathcal{M}} v\mu(dz) - \int_{\mathcal{M}} v\pi(dz) \text{ a.s.}$$

We will show that $\int_{\mathcal{M}} v\mu(dz) - \int_{\mathcal{M}} v\pi(dz) > 0$ by a contradiction argument. Note that, since

$$|f_V(s,u,v,k) - f_V(s,u,\widetilde{v},k)| \leq C|v - \widetilde{v}|$$

for some constant $C > 0$, we have

(41)
$$\limsup_{T\to\infty} \frac{1}{T}\int_0^T \left| f_V(Z_t)dt - f_V\left(S_t, U_t, \widetilde{V}_t, \alpha_t\right)\right| dt \le \lim_{T\to\infty}\frac{1}{T}\int_0^T C\left(V_t - \widetilde{V}_t\right) = 0$$

if $\int_{\mathcal{M}} v\mu(dz) - \int_{\mathcal{M}} v\pi(dz) = 0$. On the other hand,

$$f_V(s,u,\widetilde{v},k) = f_V\left(\frac{\Lambda}{\gamma}, 0, \widetilde{v}, k\right) + \left(\frac{\Lambda}{\gamma} - s\right)\beta(k)\widetilde{v}^2,$$

which leads to

(42)
$$\begin{aligned}\lim_{T\to\infty}\frac{1}{T}\int_0^T f_V\left(S_t, U_t, \widetilde{V}_t, \alpha_t\right)dt &= \lim_{T\to\infty}\frac{1}{T}\int_0^T f_V\left(\frac{\Lambda}{\gamma}, 0, \widetilde{V}_t, \alpha_t\right)dt \\ &\quad + \lim_{T\to\infty}\frac{1}{T}\int_0^T \left(\frac{\Lambda}{\gamma} - S_t\right)\beta(\alpha_t)\widetilde{V}_t \\ &> \lim_{T\to\infty}\frac{1}{T}\int_0^T f_V\left(\frac{\Lambda}{\gamma}, 0, \widetilde{V}_t^2, \alpha_t\right)dt,\end{aligned}$$

where we use the ergodicity of $(Z_t, \widetilde{V}_t)$ on $\mathcal{M}_+ \times (0,1)$ to have that

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T \left(\frac{\Lambda}{\gamma} - S_t\right)\beta(\alpha_t)\widetilde{V}_t > 0.$$

Combining (41) and (42) we have

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T f_V(Z_t)dt - \lim_{T\to\infty}\frac{1}{T}\int_0^T f_V\left(\frac{\Lambda}{\gamma}, 0, \widetilde{V}_t, \alpha_t\right)dt > 0$$

if $\int_{\mathcal{M}} v\mu(dz) - \int_{\mathcal{M}} v\pi(dz) = 0$.

However, it contradicts the fact that

$$\begin{aligned}\lim_{T\to\infty}\frac{1}{T}\int_0^T f_V(Z_t)dt - &\lim_{T\to\infty}\frac{1}{T}\int_0^T f_V\left(\frac{\Lambda}{\gamma}, 0, \widetilde{V}_t, \alpha_t\right)dt \\ &= \int_{\mathcal{M}} f_V(z)\mu(dz) - \int_{\mathcal{M}} f_V(z)\pi(dz) = 0 - 0,\end{aligned}$$

where the last equality is due to an argument similar to (39). Thus,

(43)
$$\lim_{T\to\infty}\frac{1}{T}\int_0^T (V_t - \widetilde{V}_t)dt > 0\text{-a.s.}$$

Since $\widetilde{H}$ is an increasing function in v with positive derivative, we can easily infer from (43) and the fact that $V_t \ge \widetilde{V}_t$ that

$$\mu\widetilde{H} - \pi\widetilde{H} = \lim_{T\to\infty}\frac{1}{T}\int_0^T \widetilde{H}(Z_t)dt - \lim_{T\to\infty}\frac{1}{T}\int_0^T \widetilde{H}\left(\frac{\Lambda}{\gamma}, 0, \widetilde{V}_t, \alpha_t\right)dt > 0.$$

In view of Corollary 2.10, we obtain the second claim of the theorem. The proof is complete. ∎

4. Conclusion. In this paper, we have given a general method to deal with the critical case in population dynamics in a random environment. We apply the method to five different models, including epidemiological, prey-predator, and population in a structured environment.

When our results apply, there is extinction in temporal average in the critical case. A natural question is whether it is possible to find other results, such that there is persistence (maybe in a weaker sense) in the critical case.

Our method consists of looking at integrals of the function $H = \mathcal{L}V$ with respect to invariant measures of the process. For some models, another method is possible, as used, for example, for some PDMPs in [22]. The idea is the following. Assume that if $\mathcal{P}_{inv}(\mathcal{M}_+)$ is nonempty, then it is possible to compute, or at least, to estimate, the density of an invariant probability $\mu \in \mathcal{P}_{inv}(\mathcal{M}_+)$. Then, this density must satisfy some integrability conditions, which can be violated if $\Lambda^+(H) = 0$ (see e.g [22, Theorem 3.1] or [17, Lemma 6]). Hence, if $\Lambda^+(H) = 0$, $\mathcal{P}_{inv}(\mathcal{M}_+)$ has to be empty. This alternative method is close in spirit to ours, since it comes to a contraction when assuming that $\mathcal{P}_{inv}(\mathcal{M}_+)$ is nonempty and $\Lambda^+(H) = 0$.

Acknowledgment. We thank two anonymous referees for valuable comments to improve the presentation of the paper.

REFERENCES

[1] P. A. Abrams, R. D. Holt, and J. D. Roth, *Apparent competition or apparent mutualism? Shared predation when populations cycle*, Ecology, 79 (1998), pp. 201–212.
[2] L. Arnold, *Random Dynamical Systems*, Springer Monogr. Math., Springer-Verlag, Berlin, 1998, https://doi.org/10.1007/978-3-662-12878-7.
[3] Y. Bakhtin and T. Hurth, *Invariant densities for dynamical systems with random switching*, Nonlinearity, 25 (2012), pp. 2937–2952, https://doi.org/10.1088/0951-7715/25/10/2937.
[4] M. Benaïm, *Stochastic Persistence*, preprint, https://arxiv.org/abs/1806.08450, 2018.
[5] M. Benaïm, S. Le Borgne, F. Malrieu, and P.-A. Zitt, *Qualitative properties of certain piecewise deterministic Markov processes*, Ann. Inst. Henri Poincaré Probab. Stat., 51 (2015), pp. 1040–1075.
[6] M. Benaïm and C. Lobry, *Lotka Volterra in fluctuating environment or "how switching between beneficial environments can make survival harder,"* Ann. Appl. Probab., 26 (2016), pp. 3754–3785.
[7] M. Benaïm and E. Strickler, *Random switching between vector fields having a common zero*, Ann. Appl. Probab., 29 (2019), pp. 326–375.
[8] P. Chesson and J. J. Kuang, *The interaction between predation and competition*, Nature, 456 (2008), 235.
[9] P. L. Chesson and R. R. Warner, *Environmental variability promotes coexistence in lottery competitive systems*, Amer. Naturalist, 117 (1981), pp. 923–943.
[10] I. Chueshov, *Monotone Random Systems Theory and Applications*, Lecture Notes in Math. 1779, Springer-Verlag, Berlin, 2002, https://doi.org/10.1007/b83277.
[11] M. H. A. Davis, *Piecewise-deterministic Markov processes: A general class of nondiffusion stochastic models*, J. Roy. Statist. Soc. Ser. B, 46 (1984), pp. 353–388.
[12] N. T. Dieu, D. H. Nguyen, N. H. Du, and G. Yin, *Classification of asymptotic behavior in a stochastic SIR model*, SIAM J. Appl. Dyn. Syst., 15 (2016), pp. 1062–1084.
[13] N. H. Du, D. H. Nguyen, and G. Yin, *Dynamics of a stochastic Lotka-Volterra model perturbed by white noise*, J. Appl. Probab., 53 (2016), pp. 187–202.
[14] S. N. Evans, A. Hening, and S. J. Schreiber, *Protected polymorphisms and evolutionary stability of patch-selection strategies in stochastic environments*, J. Math. Biol., 71 (2015), pp. 325–359.
[15] S. N. Evans, P. L. Ralph, S. J. Schreiber, and A. Sen, *Stochastic population growth in spatially heterogeneous environments*, J. Math. Biol., 66 (2013), pp. 423–476.
[16] J. H. Gillespie and H. A. Guess, *The effects of environmental autocorrelations on the progress of selection in a random environment*, Amer. Naturalist, 112 (1978), pp. 897–909.

[17] A. Guillin, A. Personne, and E. Strickler, *Persistence in the Moran Model with Random Switching*, preprint, https://arxiv.org/abs/1911.01108, 2019.

[18] A. Hening and D. Nguyen, *Coexistence and extinction for stochastic Kolmogorov systems*, Ann. Appl. Probab., 28 (2018), pp. 1893–1942.

[19] A. Hening, D. H. Nguyen, and G. Yin, *Stochastic population growth in spatially heterogeneous environments: The density-dependent case*, J. Math. Biol., 76 (2018), pp. 697–754.

[20] A. Hening and E. Strickler, *On a predator-prey system with random switching that never converges to its equilibrium*, SIAM J. Math. Anal., 51 (2019), pp. 3625–3640.

[21] H. W. Hethcote, *The mathematics of infectious diseases*, SIAM Rev., 42 (2000), pp. 599–653.

[22] T. Hurth and C. Kuehn, *Random switching near bifurcations*, Stoch. Dyn., 20 (2020), 2050008.

[23] A. Lajmanovich and J. A. Yorke, *A deterministic model for gonorrhea in a nonhomogeneous population*, Math. Biosci., 28 (1976), pp. 221–236, https://doi.org/10.1016/0025-5564(76)90125-5.

[24] D. Li, S. Liu, and J. Cui, *Threshold dynamics and ergodicity of an SIRS epidemic model with Markovian switching*, J. Differential Equations, 263 (2017), pp. 8873–8915, https://doi.org/10.1016/j.jde.2017.08.066.

[25] D. H. Nguyen and D. H. Nguyen, *Dynamics of Kolmogorov systems of competitive type under the telegraph noise*, J. Differential Equations, 250 (2011), pp. 386–409.

[26] D. H. Nguyen, N. H. Nguyen, and G. Yin, *General nonlinear stochastic systems motivated by chemostat models: Complete characterization of long-time behavior, optimal controls, and applications to wastewater treatment*, Stochastic Process. Appl., 130 (2020), pp. 4608-4642.

[27] S. Schreiber, *Persistence for stochastic difference equations: A mini review*, J. Difference Equ. Appl., 18 (2012), pp. 1381–1403, https://doi.org/10.1080/10236198.2011.628662.

[28] S. J. Schreiber, M. Benaïm, and K. A. S. Atchadé, *Persistence in fluctuating environments*, J. Math. Biol., 62 (2011), pp. 655–683, https://doi.org/10.1007/s00285-010-0349-5.

[29] I. B. Schwartz and H. L. Smith, *Infinite subharmonic bifurcation in an SEIR epidemic model*, J. Math. Biol., 18 (1983), pp. 233–253.

[30] E. Strickler, *Persistance de Processus de Markov Déterministes par Morceaux*, Ph.D thesis, Université de Neuchâtel, 2019.

[31] E. Strickler, *Randomly switched vector fields sharing a zero on a common invariant face*, Stoch. Dyn., https://doi.org/10.1142/S0219493721500076.