The Annals of Applied Probability
2018, Vol. 28, No. 3, 1893–1942
https://doi.org/10.1214/17-AAP1347

COEXISTENCE AND EXTINCTION FOR STOCHASTIC KOLMOGOROV SYSTEMS

BY ALEXANDRU HENING AND DANG H. NGUYEN[1]

Tufts University and Wayne State University

In recent years there has been a growing interest in the study of the dynamics of stochastic populations. A key question in population biology is to understand the conditions under which populations coexist or go extinct. Theoretical and empirical studies have shown that coexistence can be facilitated or negated by both biotic interactions and environmental fluctuations. We study the dynamics of n populations that live in a stochastic environment and which can interact nonlinearly (through competition for resources, predator–prey behavior, etc.). Our models are described by n-dimensional Kolmogorov systems with white noise (stochastic differential equations—SDE). We give sharp conditions under which the populations converge exponentially fast to their unique stationary distribution as well as conditions under which some populations go extinct exponentially fast.

The analysis is done by a careful study of the properties of the invariant measures of the process that are supported on the boundary of the domain. To our knowledge this is one of the first general results describing the asymptotic behavior of stochastic Kolmogorov systems in non-compact domains.

We are able to fully describe the properties of many of the SDE that appear in the literature. In particular, we extend results on two dimensional Lotka-Volterra models, two dimensional predator–prey models, n dimensional simple food chains, and two predator and one prey models. We also show how one can use our methods to classify the dynamics of any two-dimensional stochastic Kolmogorov system satisfying some mild assumptions.

CONTENTS

Received October 2016; revised August 2017.

[1]Supported in part by NSF Grant DMS-17-10827.

MSC2010 subject classifications. Primary 60K35; secondary 37H15, 60H10, 60J05, 60J99.

Key words and phrases. Kolmogorov system, ergodicity, Lotka–Volterra, Lyapunov exponent, stochastic environment, predator–prey, population dynamics.

1. Introduction. Real populations do not evolve in isolation and as a result much of ecology is concerned with understanding the characteristics that allow two species to coexist, or one species to take over the habitat of another. It is of fundamental importance to understand what will happen to an invading species. Will it invade successfully or die out in the attempt? If it does invade, will it coexist with the native population? Mathematical models for invasibility have contributed significantly to the understanding of the epidemiology of infectious disease outbreaks [Cross et al. (2005)] and ecological processes [Law and Morton (1996), Caswell (2001)]. There is widespread empirical evidence that heterogeneity, arising from abiotic (precipitation, temperature, sunlight) or biotic (competition, predation) factors, is important in determining invasibility [Davies et al. (2005), Pyšek and Hulme (2005)]. The fluctuations of the environment make the dynamics of populations inherently stochastic.

The combined effects of biotic interactions and environmental fluctuations are key when trying to determine species richness. Sometimes biotic effects can result in species going extinct. However, if one adds the effects of a random environment, extinction might be reversed into coexistence. In other instances, deterministic systems that coexist become extinct once one takes into account environmental fluctuations. A successful way of studying this interplay is by modelling the populations as discrete or continuous-time Markov processes and looking at the long-term behavior of these processes [Benaïm and Schreiber (2009), Benaïm, Hofbauer and Sandholm (2008), Blath, Etheridge and Meredith (2007), Cattiaux and Méléard (2010), Cattiaux et al. (2009), Chesson (2000), Evans, Hening and Schreiber (2015), Evans et al. (2013), Lande, Engen and Saether (2003), Schreiber, Benaïm and Atchadé (2011), Schreiber and Lloyd-Smith (2009)].

A natural way of analyzing the coexistence of species is by analyzing the average per-capita growth rate of a population when rare. Intuitively, if this growth rate is positive the respective population increases when rare, and can invade, while if it is negative the population decreases and goes extinct. If there are only two populations, then coexistence is ensured if each population can invade when it is rare and the other population is stationary [Chesson and Ellner (1989), Evans, Hening and Schreiber (2015), Turelli (1977)].

There is a general theory for coexistence for deterministic models [Hofbauer (1981), Hofbauer and So (1989), Hutson (1984)]. It is shown that a sufficient condition for persistence is the existence of a fixed set of weights associated with the interacting populations such that this weighted combination of the populations's invasion rates is positive for any invariant measure supported by the boundary (i.e., associated to a sub-collection of populations); see Hofbauer (1981).

A few recent studies have explored the effect of environmental stochasticity on continuous-time models. In Benaïm, Hofbauer and Sandholm (2008), the authors found that if a deterministic continuous-time model satisfies the above persistence criterion, then under some weak assumptions the corresponding stochastic differential equation with a small diffusion term has a positive stationary distribution concentrated on the positive global attractor of the deterministic system. For general stochastic difference and differential equations with arbitrary levels of noise on a *compact* state space, sufficient conditions for persistence are given in Schreiber, Benaïm and Atchadé (2011).

The aim of this paper is two-fold. First, we want to have a general theory that gives sharp sufficient conditions for both persistence *and* extinction for stochastic Kolmogorov systems. Second, we want our methods to work on noncompact state spaces (e.g., $\mathbb{R}^n_+$).

The criteria we present for persistence are the same as those in Schreiber, Benaïm and Atchadé (2011). However, we extend their result to noncompact state spaces and we prove that the convergence rate is exponential. We note that some of our persistence results have been announced in the 2014 Bernoulli lecture of Michel Benaïm. Furthermore, criteria for persistence for general Markov processes appear in Benaïm (2014) and we use some of those ideas in our proofs. We come up with natural assumptions under which one or more populations go extinct with nonzero probability. There do not seem to be general criteria for extinction in the literature. Results have been obtained for a Lotka–Volterra competitive system in the two-dimensional setting for SDE [Evans, Hening and Schreiber (2015)] and piecewise-deterministic Markov processes [Benaïm and Lobry (2016)]. However, in these cases there are only two or three ergodic invariant probability measures on the boundary and as such the proofs simplify significantly.

It should be noted that most of the related results in the literature are obtained by choosing a function and imposing conditions such that the function has some Lyapunov-type properties. The choice of a Lyapunov function is usually artificial and imposes unnecessary constraints on the system. The results one gets are therefore limited as the particular Lyapunov function does not reflect the true nature of the dynamical system. Our approach is to carefully analyze the dynamics of the process near the boundary of its domain. Because of this, we are able to fully characterize and classify the asymptotic behavior of the system.

As corollaries of our main theorems, we extend results on two-dimensional Lotka–Volterra models [see Evans, Hening and Schreiber (2015), Benaïm and Lobry (2016)], two-dimensional predator–prey models [see Chen and Kulperger (2005), Rudnicki (2003), Rudnicki and Pichór (2007)], two predator and one prey models [see Liu and Bai (2016)] and populations modeled by SDE in a compact state space [see Schreiber, Benaïm and Atchadé (2011)].

The paper is organized as follows. In Section 1.1, we define our framework, the problems we study, our different assumptions and the main results. In Section 2, we exhibit a few examples that fall into our general setting (Lotka–Volterra competition and predator–prey models). We also give an example of a cooperative Lotka–Volterra model that does not satisfy our assumptions. However, in this case either the solution blows up in finite time or there is no invariant probability measure supported by the interior of the domain. In Section 3, we analyze some of the properties of the SDE that models our populations. In particular, we show it has a well-defined strong solution $\mathbf{X}$ for all $t > 0$ and that this solution is pathwise unique. Section 4 is devoted to the study of conditions under which $\mathbf{X}$ converges to its unique invariant probability measure on $\mathbb{R}_+^{n,\circ} := (0, \infty)^n$. In Theorem 4.1, we show that, under some natural assumptions, $\mathbf{X}$ is strongly stochastically persistent and that the convergence in total variation of its transition probability to a unique stationary distribution on $\mathbb{R}_+^{n,\circ}$ is exponentially fast. In Section 5, we look at when one or more of the populations go extinct with a positive probability. First, we show in Theorem 5.1 that if there exists an invariant probability measure living on the boundary that is a *sink*, then the process converges to the boundary in a weak sense. Under a few extra assumptions, we show in Theorem 5.2 that for every sink invariant measure μ on the boundary the process converges with strictly positive probability to the support of μ. Finally, we present in the Appendix the proofs of some auxiliary lemmas from Section 3 and Section 5.

1.1. *Notation and results*. We work on a complete probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ with a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ satisfying the usual conditions. Consider a stochastic Kolmogorov system

$$dX_i(t) = X_i(t) f_i\big(\mathbf{X}(t)\big)\,dt + X_i(t) g_i\big(\mathbf{X}(t)\big)\,dE_i(t), \qquad i = 1, \dots, n \tag{1.1}$$

taking values in $[0, \infty)^n$. We assume $\mathbf{E}(t) = (E_1(t), \dots, E_n(t))^T = \Gamma^\top \mathbf{B}(t)$ where Γ is a $n \times n$ matrix such that $\Gamma^\top \Gamma = \Sigma = (\sigma_{ij})_{n\times n}$ and $\mathbf{B}(t) = (B_1(t), \dots, B_n(t))$ is a vector of independent standard Brownian motions adapted to the filtration $(\mathcal{F}_t)_{t\geq 0}$. The SDE (1.1) is describing the dynamics of n interacting populations $\mathbf{X}(t) = (X_1(t), \dots, X_n(t))_{t\geq 0}$. Throughout the paper, we set $\mathbb{R}_+^n := [0, \infty)^n$ and $\mathbb{R}_+^{n,\circ} := (0, \infty)^n$.

REMARK 1.1. One might wonder if one could treat the more general model

$$dX_i(t) = X_i(t) f_i\big(\mathbf{X}(t)\big)\,dt + X_i(t) g_i\big(\mathbf{X}(t)\big) \sum_{j=1}^n \Sigma_{ij}\big(\mathbf{X}(t)\big)\,dB_j(t), \qquad i = 1, \dots, n.$$

In our model (1.1), we work with a constant correlation matrix $\Sigma = (\sigma_{ij})$ but it can be seen that the proofs do not depend on whether Σ is constant or a function of $\mathbf{x}$. Thus, our results still hold if Σ depends on $\mathbf{x}$ as long as it is bounded and locally Lipschitz. Actually, we can always assume it is bounded because we can normalize Σ and absorb the necessary factors into $g_i(\mathbf{x})$.

The drift term of our system is due to the deterministic dynamics while the diffusion term is due to the effects of random fluctuations of the environment. The drift for population i is given by $X_i(t) f_i(\mathbf{X}(t))$ where f_i is its per-capita growth rate. From now on, the process given by the solution to (1.1) will be denoted by $\mathbf{X}$ or $(\mathbf{X}(t))_{t\geq 0}$.

Let $\mathcal{L}$ be the infinitesimal generator of the process $\mathbf{X}$. For smooth enough functions $F : \mathbb{R}^n_+ \to \mathbb{R}$, the generator $\mathcal{L}$ acts as

$$\mathcal{L}F(\mathbf{x}) = \sum_i x_i f_i(\mathbf{x}) \frac{\partial F}{\partial x_i}(\mathbf{x}) + \frac{1}{2} \sum_{i,j} \sigma_{ij} x_i x_j g_i(\mathbf{x}) g_j(\mathbf{x}) \frac{\partial^2 F}{\partial x_i \partial x_j}(\mathbf{x}).$$

We use the norm $\|\mathbf{x}\| = \sum_{i=1}^n |x_i|$ in $\mathbb{R}^n$. For $a, b \in \mathbb{R}$, let $a \wedge b := \min\{a, b\}$ and $a \vee b := \max\{a, b\}$. Similarly, we let $\bigwedge_{i=1}^n u_i := \min_i u_i$ and $\bigvee_{i=1}^n u_i := \max_i u_i$.

We remark that (1.1) can be seen as a generalization to noncompact state spaces of the model studied in Schreiber, Benaïm and Atchadé (2011). The following is a standing assumption throughout the paper.

ASSUMPTION 1.1. The coefficients of (1.1) satisfy the following conditions:

(1) $\operatorname{diag}(g_1(\mathbf{x}), \dots, g_n(\mathbf{x}))\Gamma^\top \Gamma \operatorname{diag}(g_1(\mathbf{x}), \dots, g_n(\mathbf{x})) = (g_i(\mathbf{x})g_j(\mathbf{x})\sigma_{ij})_{n\times n}$ is a positive definite matrix for any $\mathbf{x} \in \mathbb{R}^n_+$.

(2) $f_i(\cdot), g_i(\cdot) : \mathbb{R}^n_+ \to \mathbb{R}$ are locally Lipschitz functions for any $i = 1, \dots, n$.

(3) There exist $\mathbf{c} = (c_1, \dots, c_n) \in \mathbb{R}^{n,\circ}_+$ and $\gamma_b > 0$ such that

$$\limsup_{\|x\|\to\infty} \left[\frac{\sum_i c_i x_i f_i(\mathbf{x})}{1 + \mathbf{c}^\top \mathbf{x}} - \frac{1}{2} \frac{\sum_{i,j} \sigma_{ij} c_i c_j x_i x_j g_i(\mathbf{x}) g_j(\mathbf{x})}{(1 + \mathbf{c}^\top \mathbf{x})^2} + \gamma_b \left(1 + \sum_i \left(|f_i(\mathbf{x})| + g_i^2(\mathbf{x})\right)\right) \right] < 0. \tag{1.2}$$

REMARK 1.2. Parts (2) and (3) of Assumption 1.1 guarantee the existence and uniqueness of strong solutions to (1.1). We need part (1) of Assumption 1.1 to ensure that the solution to (1.1) is a nondegenerate diffusion. Moreover, we show later that (3) implies the tightness of the family of transition probabilities of the solution to (1.1).

REMARK 1.3. There are a few different ways to add stochastic noise to deterministic population dynamics. We assume that the environment mainly affects the growth/death rates of the populations. See Braumann (2002), Evans, Hening and Schreiber (2015), Evans et al. (2013), Gard (1988), Hening, Nguyen and Yin (2017), Schreiber, Benaïm and Atchadé (2011), Turelli (1977) for more details.

We next define what we mean by persistence and extinction in our setting.

DEFINITION 1.1. The process $\mathbf{X}$ is strongly stochastically persistent if it has a unique invariant probability measure π^* on $\mathbb{R}^{n,\circ}_+$ and

$$\lim_{t\to\infty} \|P_{\mathbf{X}}(t, \mathbf{x}, \cdot) - \pi^*(\cdot)\|_{\mathrm{TV}} = 0, \qquad \mathbf{x} \in \mathbb{R}^{n,\circ}_+, \tag{1.3}$$

where $\|\cdot, \cdot\|_{\mathrm{TV}}$ is the total variation norm and $P_{\mathbf{X}}(t, \mathbf{x}, \cdot)$ is the transition probability of $(\mathbf{X}(t))_{t\geq 0}$.

DEFINITION 1.2. If $\mathbf{X}(0) = \mathbf{x} \in \mathbb{R}_+^{n,\circ}$, we say the population X_i goes extinct with probability $p_\mathbf{x} > 0$ if

$$\mathbb{P}_\mathbf{x}\Big\{\lim_{t\to\infty} X_i(t) = 0\Big\} = p_\mathbf{x}.$$

We say the population X_i goes extinct if for all $\mathbf{x} \in \mathbb{R}_+^{n,\circ}$

$$\mathbb{P}_\mathbf{x}\Big\{\lim_{t\to\infty} X_i(t) = 0\Big\} = 1.$$

EXAMPLE 1.1. Most of the common ecological models satisfy condition (1.2).

- Consider the linear one-dimensional model

$$dX(t) = aX(t)\,dt + \sigma X(t)\,dB(t).$$

If $a - \frac{\sigma^2}{2} < 0$, then (1.2) is satisfied for any $c > 0$.

- Consider the logistic model

$$dX(t) = X(t)\big[a - bX(t)\big]\,dt + \sigma X(t)\,dB(t), \qquad b > 0.$$

Then equation (1.2) is satisfied for any $c > 0$.

- Consider the competitive Lotka–Volterra model

$$dX_i(t) = X_i(t)\Big[a_i - \sum_j b_{ji}X_j(t)\Big]\,dt + X_i(t)g_i\big(\mathbf{X}(t)\big)\,dE_i(t),$$

with $b_{ji} > 0$, $j, i = 1, \dots, n$.

If $\sum_{i=1}^n g_i^2(\mathbf{x}) < K(1 + \|\mathbf{x}\| + \bigwedge_{i=1}^n g_i^2(\mathbf{x}))$, then (1.2) is satisfied with $\mathbf{c} = (1, \dots, 1)$. We give a short argument for why this is true. Since $b_{ij} > 0$, there is $\tilde{b} > 0$ such that

$$\frac{\sum_i x_i(a_i - \sum_j b_{ji}x_j)}{1 + \sum_i x_i} < -\tilde{b}\Big(1 + \sum_i x_i\Big) \tag{1.4}$$

if $\|\mathbf{x}\|$ is sufficiently large. By the Cauchy–Schwarz inequality, there are $\tilde{\sigma}_1, \tilde{\sigma}_2 > 0$ such that

$$-\frac{1}{2}\frac{\sum_{i,j}\sigma_{ij}x_ix_jg_i(\mathbf{x})g_j(\mathbf{x})}{(1 + \sum_i x_i)^2} \le -\tilde{\sigma}_1\frac{\sum_i x_i^2 g_i^2(\mathbf{x})}{(1 + \sum_i x_i)^2} \le -\tilde{\sigma}_2\Big(\bigwedge_{i=1}^n g_i^2(\mathbf{x})\Big), \tag{1.5}$$

when $\|\mathbf{x}\|$ is sufficiently large. In light of (1.4) and (1.5), if $\sum_{i=1}^n g_i^2(\mathbf{x}) < K(1 + \|\mathbf{x}\| + \bigwedge_{i=1}^n g_i^2(\mathbf{x}))$, we can find $\gamma_b > 0$ such that

$$\begin{aligned}&\frac{\sum_i x_i(a_i - \sum_j b_{ji}x_j)(\mathbf{x})}{1 + \sum_i x_i} - \frac{1}{2}\frac{\sum_{i,j}\sigma_{ij}x_ix_jg_i(\mathbf{x})g_j(\mathbf{x})}{(1 + \sum_i x_i)^2}\\ &\quad + \gamma_b\Big(1 + \sum_i\Big(\Big|a_i - \sum_j b_{ji}x_j\Big| + g_i^2(\mathbf{x})\Big)\Big) < 0\end{aligned}$$

for sufficiently large $\|\mathbf{x}\|$. As a result, (1.2) holds.

- Consider the predator–prey Lotka–Volterra model

$$\begin{cases} dX(t) = X(t)\big[a_1 - b_1X(t) - c_1Y(t)\big]\,dt + X(t)g_1\big(X(t), Y(t)\big)\,dE_1(t), \\ dY(t) = Y(t)\big[-a_2 - b_2Y(t) + c_2X(t)\big]\,dt + Y(t)g_2\big(X(t), Y(t)\big)\,dE_2(t), \end{cases}$$

with $b_1, b_2 > 0, c_1, c_2 \geq 0, a_2 \geq 0$. If $\sum_{i=1}^{2} g_i^2(x, y) < K(1 + x + y + g_1^2(x, y) \wedge g_2^2(x, y))$, one can use arguments similar to those from the competitive Lotka–Volterra model to show that (1.2) is satisfied with $\mathbf{c} = (c_2, c_1)$.

Let $\mathcal{M}$ be the set of ergodic invariant probability measures of $\mathbf{X}$ supported on the boundary $\partial\mathbb{R}_+^n := \mathbb{R}_+^n \setminus \mathbb{R}_+^{n,\circ}$. Note that if we let $\boldsymbol{\delta}^*$ be the Dirac measure concentrated at $\mathbf{0}$ then $\boldsymbol{\delta}^* \in \mathcal{M}$ so that $\mathcal{M} \neq \varnothing$. For a subset $\widetilde{\mathcal{M}} \subset \mathcal{M}$, denote by $\mathrm{Conv}(\widetilde{\mathcal{M}})$ the convex hull of $\widetilde{\mathcal{M}}$, that is, the set of probability measures π of the form $\pi(\cdot) = \sum_{\mu \in \widetilde{\mathcal{M}}} p_\mu \mu(\cdot)$ with $p_\mu > 0, \sum_{\mu \in \widetilde{\mathcal{M}}} p_\mu = 1$.

Consider $\mu \in \mathcal{M}$. Assume $\mu \neq \boldsymbol{\delta}^*$. Since the diffusion $\mathbf{X}$ is nondegenerate in each subspace, there exist $0 < n_1 < \cdots < n_k \leq n$ such that $\mathrm{supp}(\mu) = \mathbb{R}_+^\mu$ where

$$\mathbb{R}_+^\mu := \big\{(x_1, \dots, x_n) \in \mathbb{R}_+^n : x_i = 0 \text{ if } i \in I_\mu^c\big\}$$

for $I_\mu := \{n_1, \dots, n_k\}$ and $I_\mu^c := \{1, \dots, n\} \setminus \{n_1, \dots, n_k\}$. If $\mu = \boldsymbol{\delta}^*$, then we note that $\mathbb{R}_+^{\boldsymbol{\delta}^*} = \{\mathbf{0}\}$. Let

$$\mathbb{R}_+^{\mu,\circ} := \big\{(x_1, \dots, x_n) \in \mathbb{R}_+^n : x_i = 0 \text{ if } i \in I_\mu^c \text{ and } x_i > 0 \text{ if } x_i \in I_\mu\big\}$$

and $\partial\mathbb{R}_+^\mu := \mathbb{R}_+^\mu \setminus \mathbb{R}_+^{\mu,\circ}$.

The following condition ensures strong stochastic persistence.

ASSUMPTION 1.2. For any $\mu \in \mathrm{Conv}(\mathcal{M})$, one has

$$\max_{\{i=1,\dots,n\}} \big\{\lambda_i(\mu)\big\} > 0,$$

where

$$\lambda_i(\mu) := \int_{\partial\mathbb{R}_+^n} \left(f_i(\mathbf{x}) - \frac{\sigma_{ii} g_i^2(\mathbf{x})}{2} \right) \mu(d\mathbf{x}).$$

[In view of Lemma 3.3, $\lambda_i(\mu)$ is well defined.]

THEOREM 1.1. *Suppose that Assumptions 1.1 and 1.2 hold. Then* $\mathbf{X}$ *is strongly stochastically persistent and converges exponentially fast to its unique invariant probability measure* π^* *on* $\mathbb{R}_+^{n,\circ}$.

The proof of this result is presented in detail in Section 4. The following remark gives a rough intuitive sketch of the proof.

REMARK 1.4. From a dynamical point of view, the solution in the interior domain $\mathbb{R}_+^{n,\circ}$ is persistent if every invariant probability measure on the boundary is a "repeller." In a deterministic setting, an equilibrium is a repeller if it has a positive Lyapunov exponent. In a stochastic model, ergodic invariant measures play a similar role. To determine the Lyapunov exponents of an ergodic invariant measure, one can look at the equation for $\ln X_i(t)$. An application of Itô's lemma yields that

$$\frac{\ln X_i(t)}{t} = \frac{\ln X_i(0)}{t} + \frac{1}{t}\int_0^t \left[f_i\big(\mathbf{X}(s)\big) - \frac{g_i^2(\mathbf{X}(s))\sigma_{ii}}{2}\right] ds + \frac{1}{t}\int_0^t g_i\big(\mathbf{X}(s)\big)\, dE_i(s).$$

If $\mathbf{X}$ is close to the support of an ergodic invariant measure μ for a long time, then

$$\frac{1}{t}\int_0^t \left[f_i\big(\mathbf{X}(s)\big) - \frac{g_i^2(\mathbf{X}(s))\sigma_{ii}}{2}\right] ds$$

can be approximated by the average with respect to μ

$$\lambda_i(\mu) = \int_{\partial\mathbb{R}_+^n} \left(f_i(\mathbf{x}) - \frac{g_i^2(\mathbf{x})\sigma_{ii}}{2}\right)\mu(d\mathbf{x})$$

while the term

$$\frac{\ln X_i(0)}{t} + \frac{1}{t}\int_0^t g_i\big(\mathbf{X}(s)\big)\, dE_i(s)$$

is negligible. This implies that $\lambda(\mu_i), i = 1, \dots, n$ are the Lyapunov exponents of μ [it can also be seen that $\lambda(\mu_i)$ gives the long-term growth rate of $X_i(t)$ if $\mathbf{X}$ is close to the support of μ]. As a result, if $\max_{i=1}^n\{\lambda(\mu_i)\} > 0$, then the invariant measure μ is a "repeller." Therefore, Assumption 1.2 guarantees the persistence of the population. Moreover, by evaluating the exponential rate $\frac{\ln X_i(T)}{T}$ for sufficiently large T (so that the ergodicity takes effect), we can show that the solution goes away from the boundary exponentially fast, and then obtain a geometric rate of convergence in total variation under Assumptions 1.1 and 1.2. This is achieved by constructing a suitable Lyapunov function with the help of the Laplace transform and the approximations that were mentioned above. Note that since we work on a noncompact space, Assumption 1.2 part (3) is needed to show that the solution enters a compact subset of $\mathbb{R}_+^n$ exponentially fast.

The following condition will imply extinction.

ASSUMPTION 1.3. There exists a $\mu \in \mathcal{M}$ such that

$$\max_{i\in I_\mu^c}\{\lambda_i(\mu)\} < 0. \tag{1.6}$$

If $\mathbb{R}_+^\mu \neq \{\mathbf{0}\}$, suppose further that for any $\nu \in \mathrm{Conv}(\mathcal{M}_\mu)$, we have

$$\max_{i \in I_\mu}\{\lambda_i(\nu)\} > 0, \tag{1.7}$$

where $\mathcal{M}_\mu := \{\nu' \in \mathcal{M} : \mathrm{supp}(\nu') \subset \partial\mathbb{R}_+^\mu\}$.

Define

$$\mathcal{M}^1 := \{\mu \in \mathcal{M} : \mu \text{ satisfies Assumption 1.3}\} \tag{1.8}$$

and

$$\mathcal{M}^2 := \mathcal{M} \setminus \mathcal{M}^1. \tag{1.9}$$

THEOREM 1.2. *Under Assumptions 1.1 and 1.3, for any $\delta > 0$ sufficiently small and any $\mathbf{x} \in \mathbb{R}_+^{n,\circ}$ we have*

$$\lim_{t\to\infty} \mathbb{E}_{\mathbf{x}}\left(\bigwedge_{i=1}^n X_i(t)\right)^\delta = 0. \tag{1.10}$$

The proof of this result is given in Section 5.

REMARK 1.5. If an ergodic invariant measure μ with support on the boundary is an "attractor," it will attract solutions starting nearby. Intuitively, condition (1.6) forces $X_i(t), i \in I_\mu^c$ to get close to 0 if the solution starts close to $\mathbb{R}_+^{\mu,\circ}$. We need condition (1.7) to ensure that μ is a "sink" in $\mathbb{R}_+^{\mu,\circ}$, that is, if $\mathbf{X}$ is close to $\mathbb{R}_+^{\mu,\circ}$, it is not pulled away to the boundary $\partial\mathbb{R}_+^{\mu,\circ}$ of $\mathbb{R}_+^{\mu,\circ}$ (see Remark 1.6).

To prove Theorem 1.2, using the idea above, we construct a Lyapunov function U vanishing on $\mathbb{R}_+^{\mu,\circ}$ such that $\mathbb{E}_{\mathbf{x}}U(\mathbf{X}(T)) \leq U(\mathbf{x})$ if $\mathbf{x} \in \mathbb{R}_+^{n,\circ}$ sufficiently close to $\mathbb{R}_+^{\mu,\circ}$ and T is a sufficiently large time. Then we can construct a supermartingale to show that with a large probability $X_i(t), i \in I_\mu^c$ cannot go far from 0 if the starting point of $\mathbf{X}$ is sufficiently close to $\mathbb{R}_+^{\mu,\circ}$. With some additional arguments from the theory of Markov processes, we can show that $\mathbf{X}$ has no invariant probability measure in $\mathbb{R}_+^{n,\circ}$ and approaches the boundary in some sense.

In the case when there is no persistence, one may want to know exactly which species go extinct and which survive. We answer this question in Theorem 5.2. Relying on the repulsion of invariant measures in $\mathcal{M}^2 = \mathcal{M} \setminus \mathcal{M}^1$ and properties of the randomized occupation measures, we can deduce that the process $\mathbf{X}$ must enter the "attracting" region of some invariant measure in $\mathcal{M}^1$. Finally, the attraction property of the measures from $\mathcal{M}^1$ helps us characterize the survival and extinction of each species.

REMARK 1.6. If condition (1.7) does not hold, we could have the following bad situation. Assume there exists $\nu \in \mathcal{M}_\mu$ such that

$$\max_{i\in I_\mu \setminus I_\nu} \{\lambda_i(\nu)\} = 0.$$

In this case, ν is not always a "repeller." Solutions that start near $\mathbb{R}_+^{\mu,\circ}$ will tend to stay close to R_+^μ since $\lambda_i(\mu) < 0, i \in I_\mu^c$. However, if ν is not a repeller the solutions may concentrate on $\mathbb{R}^\nu \subset \partial\mathbb{R}_+^\mu$. Now, if there exists an $i^* \in I_\mu^c$ such that $\lambda_{i^*}(\nu) > 0$ then solutions can be pushed away from $\mathbb{R}_+^\mu$ since $X_{i^*}(t)$ will tend to increase.

To characterize the extinction of specific populations, we need some additional conditions.

ASSUMPTION 1.4. Suppose that there is a $\delta_1 > 0$ such that

$$\lim_{\|\mathbf{x}\|\to\infty} \frac{\|\mathbf{x}\|^{\delta_1} \sum_i g_i^2(\mathbf{x})}{1 + \sum_i(|f_i(\mathbf{x})| + |g_i(\mathbf{x})|^2)} = 0.$$

Without loss of generality, suppose that $\delta_1 \le \delta_0$ (where δ_0 is defined at the beginning of Section 3).

REMARK 1.7. Assumption 1.4 forces the growth rates of $g_i^2(\cdot)$ to be slightly lower than those of $|f_i(\cdot)|$. This is needed in order to suppress the diffusion part so that we can obtain the tightness of the random normalized occupation measures

$$\widetilde{\Pi}_t(\cdot) := \frac{1}{t}\int_0^t \mathbf{1}_{\{\mathbf{X}(s)\in\cdot\}}\,ds, \qquad t > 0$$

as well as the convergence of $\int_{\mathbb{R}_+^n} (f_i(\mathbf{x}) - \frac{g_i^2(\mathbf{x})}{2})\widetilde{\Pi}_{t_k}(d\mathbf{x})$ to $\lambda_i(\pi)$ given that $\widetilde{\Pi}_{t_k}$ converges weakly to π for some sequence $(t_k)_{k\in\mathbb{N}}$ with $\lim_{k\to\infty} t_k = \infty$. Having these properties, we can analyze the asymptotic behavior of the sample paths of the solution.

To describe exactly which populations go extinct, we need an additional assumption which ensures that apart from those in $\mathrm{Conv}(\mathcal{M}^1)$, invariant probability measures are "repellers."

ASSUMPTION 1.5. Suppose that one of the following is true:

- $\mathcal{M}^2 = \varnothing$.
- For any $\nu \in \mathrm{Conv}(\mathcal{M}^2)$, $\max_{\{i=1,\dots,n\}}\{\lambda_i(\nu)\} > 0$.

For any initial condition $\mathbf{X}(0) = \mathbf{x} \in \mathbb{R}_+^n$, denote the weak*-limit set of the family $\{\widetilde{\Pi}_t(\cdot), t \ge 1\}$ by $\mathcal{U} = \mathcal{U}(\omega)$.

THEOREM 1.3. *Suppose that Assumptions 1.1, 1.4 and 1.5 are satisfied and $\mathcal{M}^1 \neq \varnothing$. Then for any $\mathbf{x} \in \mathbb{R}_+^{n,\circ}$*

$$\sum_{\mu \in \mathcal{M}^1} P_{\mathbf{x}}^{\mu} = 1, \tag{1.11}$$

where

$$P_{\mathbf{x}}^{\mu} := \mathbb{P}_{\mathbf{x}}\left\{\mathcal{U}(\omega) = \{\mu\} \text{ and } \lim_{t\to\infty} \frac{\ln X_i(t)}{t} = \lambda_i(\mu) < 0, i \in I_{\mu}^c\right\}$$
$$> 0, \qquad \mathbf{x} \in \mathbb{R}_+^{n,\circ}, \mu \in \mathcal{M}^1.$$

REMARK 1.8. Our results can be easily modified and applied to SDE living on smooth enough domains $D \subset \mathbb{R}^n$. We chose to work on $[0, \infty)^n$ because it was the most natural noncompact example for the dynamics of biological populations. In particular, one can recover and extend the results from Schreiber, Benaïm and Atchadé (2011) where the authors looked at the state space $D = \{(y_1, \dots, y_n) \in \mathbb{R}_+^n : y_1 + \cdots + y_n = 1\}$.

2. Examples. We present some applications of our main results. We will make use of Theorems 1.1, 1.2 and 1.3 together with the following intuitive lemma whose proof is postponed to Section 5.

LEMMA 2.1. *For any $\mu \in \mathcal{M}$ and $i \in I_\mu$, we have $\lambda_i(\mu) = 0$.*

REMARK 2.1. The intuition behind Lemma 2.1 is the following: if we are inside the support of an ergodic invariant measure μ then we are at an ‘equilibrium’ and the process does not tend to grow or decay.

EXAMPLE 2.1. Consider a stochastic Lotka–Volterra competitive model

$$\begin{cases} dX_1(t) = X_1(t)\big[a_1 - b_1 X_1(t) - c_1 X_2(t)\big]\,dt + X_1(t)\,dE_1(t), \\ dX_2(t) = X_2(t)\big[a_2 - b_2 X_2(t) - c_2 X_1(t)\big]\,dt + X_2(t)\,dE_2(t), \end{cases} \tag{2.1}$$

where $b_i, c_i > 0, i = 1, 2$. It is straightforward to see that $\lambda_i(\boldsymbol{\delta}^*) = a_i - \frac{\sigma_{ii}}{2}$, $i = 1, 2$. If $\lambda_1(\boldsymbol{\delta}^*) < 0$ [resp., $\lambda_2(\boldsymbol{\delta}^*) < 0$] there is no invariant probability measure on $\mathbb{R}_{1+}^\circ := \{(x_1, 0) : x_1 > 0\}$ [resp., $\mathbb{R}_{2+}^\circ := \{(0, x_2) : x_2 > 0\}$] in view of Theorem 1.1. If $\lambda_i(\boldsymbol{\delta}^*) > 0$, there is a unique invariant probability measure μ_i on $\mathbb{R}_{i+}^\circ$, $i = 1, 2$. By Lemma 2.1, we have

$$\lambda_i(\mu_i) = a_i - \frac{\sigma_{ii}}{2} - b_i \int_{\mathbb{R}_{i+}^\circ} x_i \mu_i(d\mathbf{x}) = 0$$

which implies

$$\int_{\mathbb{R}_{i+}^\circ} x_i \mu_i(d\mathbf{x}) = \frac{2a_i - \sigma_{ii}}{2b_i}. \tag{2.2}$$

Thus

$$\lambda_2(\mu_1) = \int_{\mathbb{R}^\circ_{1+}} \left[a_2 - \frac{\sigma_{22}}{2} - c_2 x_1\right]\mu_1(d\mathbf{x}) = a_2 - \frac{\sigma_{22}}{2} - c_2 \frac{2a_1 - \sigma_{11}}{2b_1}$$

and

$$\lambda_1(\mu_2) = \int_{\mathbb{R}^\circ_{2+}} \left[a_1 - \frac{\sigma_{11}}{2} - c_1 x_2\right]\mu_2(d\mathbf{x}) = a_1 - \frac{\sigma_{11}}{2} - c_1 \frac{2a_2 - \sigma_{22}}{2b_2}.$$

Using Theorems 1.1 and 1.3, we have the following classification:

- If $\lambda_1(\boldsymbol{\delta}^*) > 0, \lambda_2(\boldsymbol{\delta}^*) > 0$ and $\lambda_1(\mu_2) > 0, \lambda_2(\mu_1) > 0$, any invariant probability measure in $\partial\mathbb{R}^2_+$ has the form $\mu = p_0\boldsymbol{\delta}^* + p_1\mu_1 + p_2\mu_2$ with $0 \le p_0, p_1, p_2$ and $p_0 + p_1 + p_2 = 1$. It can easily be verified that $\max_{i=1,2}\{\lambda_i(\mu)\} > 0$ for any μ having the form above. As a result, there is a unique invariant probability measure π^* on $\mathbb{R}^{2,\circ}_+$ and $P(t, \mathbf{x}, \cdot), \mathbf{x} \in \mathbb{R}^{2,\circ}_+$ converges to π^* in total variation exponentially fast.
- If $\lambda_i(\boldsymbol{\delta}^*) < 0, i = 1, 2$, then $X_i(t)$ converges to 0 almost surely with the exponential rate $\lambda_i(\boldsymbol{\delta}^*)$ for any initial condition $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^{2,\circ}_+$.
- If $\lambda_i(\boldsymbol{\delta}^*) > 0, \lambda_j(\boldsymbol{\delta}^*) < 0$ for one $i \in \{1, 2\}$ and $j \in \{1, 2\} \setminus \{i\}$, then $\lambda_j(\mu_i) < 0$ and $X_j(t)$ converges to 0 almost surely with the exponential rate $\lambda_j(\mu_i)$ for any initial condition $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^{2,\circ}_+$ and the randomized occupation measure converges weakly to μ_i almost surely.
- If $\lambda_i(\boldsymbol{\delta}^*) > 0$, $i \in \{1, 2\}$ and $\lambda_1(\mu_2) < 0, \lambda_2(\mu_1) < 0$, then $p_i^{\mathbf{x}} > 0, i = 1, 2$ and $p_1^{\mathbf{x}} + p_2^{\mathbf{x}} = 1$ where

$$p_i^{\mathbf{x}} = \mathbb{P}_{\mathbf{x}}\left\{\mathcal{U}(\omega) = \{\mu_i\} \text{ and } \lim_{t\to\infty} \frac{\ln X_j(t)}{t} = \lambda_j(\mu_i), j \in \{1, 2\} \setminus \{i\}\right\}.$$

- If $\lambda_1(\boldsymbol{\delta}^*) > 0, \lambda_2(\boldsymbol{\delta}^*) > 0$, $\lambda_j(\mu_i) < 0, \lambda_i(\mu_j) > 0$ for $i, j \in \{1, 2\}, i \ne j$, then $X_j(t)$ converges to 0 almost surely with the exponential rate $\lambda_j(\mu_i)$ and the randomized occupation measure converges weakly to μ_i almost surely for any initial condition $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^{2,\circ}_+$.

This extends and generalizes the results from Evans, Hening and Schreiber (2015).

EXAMPLE 2.2. Consider a stochastic Lotka–Volterra model with two predators competing for one prey:

$$\begin{cases} dX_1(t) = X_1(t)\big[a_1 - b_1X_1(t) - c_{21}X_2(t) - c_{31}X_3(t)\big]\,dt + X_1(t)\,dE_1(t), \\ dX_2(t) = X_2(t)\big[-a_2 - b_2X_2(t) + c_{12}X_1(t) - c_{32}X_3(t)\big]\,dt + X_2(t)\,dE_2(t), \\ dX_3(t) = X_3(t)\big[-a_3 - b_3X_3(t) + c_{13}X_1(t) - c_{23}X_2(t)\big]\,dt + X_2(t)\,dE_3(t). \end{cases} \tag{2.3}$$

Assume that $a_i, b_i > 0, i = 1, 3, c_{21}, c_{31}, c_{12}, c_{21}, c_{31}, c_{32} \ge 0$. Note that, if $c_{23} > 0$ and $c_{32} > 0$, then (2.3) describes an interacting population of two predators (X_2, X_3) competing for one prey X_1. If $c_{23} < 0$ and $c_{32} > 0$, then (2.3) is a model of a tri-trophic food chain where X_3 is the top predator and X_2 is the intermediate species.

In order to analyze this model, first consider the equation on the boundary $\{(0, x_2, x_3) : x_2, x_3 \geq 0\}$. Since $\lambda_i(\boldsymbol{\delta}^*) = -a_i - \frac{\sigma_{ii}}{2} < 0, i = 2, 3$, an application of Theorem 1.2 to the space $\{(0, x_2, x_3) : x_2, x_3 \geq 0\}$ shows that there is only one invariant probability measure on $\{(0, x_2, x_3) : x_2, x_3 \geq 0\}$, which is $\boldsymbol{\delta}^*$. It indicates that without the prey, both predators die out.

Now, consider the equation on the boundaries $\mathbb{R}_{12+} := \{(x_1, x_2, 0) : x_1, x_2 \geq 0\}$ and $\mathbb{R}_{13+} := \{(x_1, 0, x_3) : x_1, x_3 \geq 0\}$. If $\lambda_1(\boldsymbol{\delta}^*) = a_1 - \frac{\sigma_{11}}{2} < 0$, $\boldsymbol{\delta}^*$ is the unique invariant probability measure on $\mathbb{R}^3_+$ by virtue of Theorem 1.2. If $\lambda_1(\boldsymbol{\delta}^*) > 0$, there is an invariant probability measure μ_1 on $\mathbb{R}^\circ_{1+} := \{(x_1, 0, 0) : x_1 > 0\}$. Similar to (2.2), we have

$$\int_{\mathbb{R}^\circ_{1+}} x_1 \mu_1(d\mathbf{x}) = \frac{2a_1 - \sigma_{11}}{2b_1}. \tag{2.4}$$

Thus

$$\begin{aligned}\lambda_i(\mu_1) &= \int_{\mathbb{R}^\circ_{1+}} \left[-a_i - \frac{\sigma_{ii}}{2} + c_{1i}x_1\right]\mu_1(d\mathbf{x}) \\ &= -a_i - \frac{\sigma_{ii}}{2} + c_{1i}\frac{2a_1 - \sigma_{11}}{2b_1}, \qquad i = 2, 3.\end{aligned}$$

If $\lambda_1(\boldsymbol{\delta}^*) > 0$ and $\lambda_i(\mu_1) < 0, i = 2, 3$, by Theorem 1.2, there is no invariant probability measure on $\mathbb{R}^\circ_{1i+}$.

If $\lambda_1(\boldsymbol{\delta}^*) > 0$ and $\lambda_2(\mu_1) > 0$, by Theorem 1.1, there is an invariant probability measure μ_{12} on $\mathbb{R}^\circ_{12+}$. In light of Lemma 2.1, we have

$$\int_{\mathbb{R}^\circ_{12+}} x_1 \mu_{12}(d\mathbf{x}) = A_1, \qquad \int_{\mathbb{R}^\circ_{12+}} x_2 \mu_{12}(d\mathbf{x}) = A_2,$$

where (A_1, A_2) be the unique solution to

$$\begin{cases} a_1 - \dfrac{\sigma_{11}}{2} - b_1 x_1 - c_{21}x_2 = 0, \\ -a_2 - \dfrac{\sigma_{22}}{2} - b_2 x_2 + c_{12}x_1 = 0. \end{cases}$$

In this case,

$$\begin{aligned}\lambda_3(\mu_{12}) &= \int_{\mathbb{R}^\circ_{12+}} \left[-a_3 - \frac{\sigma_{33}}{2} + c_{13}x_1 - c_{23}x_2\right]\mu_{12}(d\mathbf{x}) \\ &= -a_3 - \frac{\sigma_{33}}{2} + c_{13}A_1 - c_{23}A_2.\end{aligned}$$

Similarly, if $\lambda_1(\boldsymbol{\delta}^*) > 0$ and $\lambda_3(\mu_1) > 0$, by Theorem 1.1, there is an invariant probability measure μ_{13} on $\mathbb{R}^\circ_{13+}$ and

$$\begin{aligned}\lambda_2(\mu_{13}) &= \int_{\mathbb{R}^\circ_{13+}} \left[-a_2 - \frac{\sigma_{22}}{2} + c_{12}x_1 - c_{32}x_3\right]\mu_{13}(d\mathbf{x}) \\ &= -a_2 - \frac{\sigma_{22}}{2} + c_{12}\widehat{A}_1 - c_{32}\widehat{A}_3,\end{aligned}$$

where $(\widehat{A}_1, \widehat{A}_3)$ is the unique solution to

$$\begin{cases} a_1 - \dfrac{\sigma_{11}}{2} - b_1x_1 - c_{31}x_3 = 0, \\ -a_3 - \dfrac{\sigma_{33}}{2} - b_3x_3 + c_{13}x_1 = 0. \end{cases}$$

By the ergodic decomposition theorem, every invariant probability measure on $\partial\mathbb{R}^3_+$ is a convex combination of $\delta^*, \mu_1, \mu_{12}, \mu_{13}$ (when these measures exist). Some computations for the Lyapunov exponents with respect to a convex combination of these ergodic measures together with an application of Theorem 1.1 show that $P(t, \mathbf{x}, \cdot), \mathbf{x} \in \mathbb{R}^{3,\circ}_+$ converges exponentially fast to an invariant probability measure π^* on $\mathbb{R}^{3,\circ}_+$ if one of the following is satisfied:

- $\lambda_1(\boldsymbol{\delta}^*) > 0, \lambda_2(\mu_1) > 0, \lambda_3(\mu_1) < 0$ and $\lambda_3(\mu_{12}) > 0$.
- $\lambda_1(\boldsymbol{\delta}^*) > 0, \lambda_2(\mu_1) < 0, \lambda_3(\mu_1) > 0$ and $\lambda_2(\mu_{13}) > 0$.
- $\lambda_1(\boldsymbol{\delta}^*) > 0, \lambda_2(\mu_1) > 0, \lambda_3(\mu_1) > 0, \lambda_3(\mu_{12}) > 0$, and $\lambda_2(\mu_{13}) > 0$.

As an application of Theorem 1.3, we have the following classification for extinction:

- If $\lambda_1(\boldsymbol{\delta}^*) < 0$, then for any initial condition $\mathbf{x} \in \mathbb{R}^{3,\circ}_+$, $X_1(t), X_2(t), X_3(t)$, converge to 0 almost surely with the exponential rates $\lambda_i(\boldsymbol{\delta}^*), i = 1, 2, 3$, respectively.
- If $\lambda_1(\boldsymbol{\delta}^*) > 0, \lambda_i(\mu_1) < 0, i = 2, 3$ then $X_i(t), i = 2, 3$ converge to 0 almost surely with the exponential rate $\lambda_i(\mu_1), i = 2, 3$, respectively, and the occupation measure converges almost surely for any initial condition $\mathbf{x} \in \mathbb{R}^{3,\circ}_+$ to μ_1.
- If $\lambda_1(\boldsymbol{\delta}^*) > 0, \lambda_i(\mu_1) > 0, \lambda_j(\mu_{1i}) < 0, \lambda_j(\mu_1) < 0$ for $i, j \in \{2, 3\}, i \neq j$, then $X_j(t)$ converges to 0 almost surely with the exponential rate $\lambda_j(\mu_{1i})$ and the occupation measure converges almost surely for any initial condition $\mathbf{x} \in \mathbb{R}^{3,\circ}_+$ to μ_{1i}.
- If $\lambda_1(\boldsymbol{\delta}^*) > 0, \lambda_2(\mu_1) > 0, \lambda_3(\mu_1) > 0, \lambda_j(\mu_{1i}) < 0, \lambda_i(\mu_{1j}) > 0$ for $i, j \in \{2, 3\}, i \neq j$ then $X_j(t)$ converges to 0 almost surely with the exponential rate $\lambda_j(\mu_{1i})$ and the occupation measure converges almost surely for any initial condition $\mathbf{x} \in \mathbb{R}^{3,\circ}_+$ to μ_{1i}.
- If $\lambda_1(\boldsymbol{\delta}^*) > 0, \lambda_2(\mu_1) > 0, \lambda_3(\mu_1) > 0, \lambda_2(\mu_{13}) < 0, \lambda_3(\mu_{12}) < 0$, then $p^{\mathbf{x}}_i > 0, i = 2, 3$ and $p^{\mathbf{x}}_2 + p^{\mathbf{x}}_3 = 1$ where

$$p^{\mathbf{x}}_i = \mathbb{P}_{\mathbf{x}}\left\{\mathcal{U}(\omega) = \{\mu_{1i}\} \text{ and } \lim_{t\to\infty} \frac{\ln X_i(t)}{t} = \lambda_i(\mu_{1j}), i \in \{2, 3\} \setminus \{j\}\right\}.$$

Elementary but tedious computations show that our results significantly improve those in Liu and Bai (2016).

Restricting our analysis to $\mathbb{R}_{12+}$ (this describes the evolution of one predator and its prey) we get

$$\begin{cases} dX_1(t) = X_1(t)\big[a_1 - b_1X_1(t) - c_{21}X_2(t)\big]\,dt + X_1(t)\,dE_1(t), \\ dX_2(t) = X_2(t)\big[-a_2 - b_2X_2(t) + c_{12}X_1(t)\big]\,dt + X_2(t)\,dE_2(t). \end{cases} \tag{2.5}$$

In view of the analysis above, if $\lambda_1(\boldsymbol{\delta}^*) < 0$ then $X_1(t)$, $X_2(t)$ converge to 0 almost surely with the exponential rates $\lambda_1(\boldsymbol{\delta}^*)$ and $\lambda_2(\boldsymbol{\delta}^*)$, respectively. If $\lambda_1(\boldsymbol{\delta}^*) > 0$ and $\lambda_2(\mu_1) < 0$, then X_2 converges to 0 almost surely with the exponential rate $\lambda_2(\mu_1)$ and the occupation measure of the process (X_1, X_2) converges to μ_1. If $\lambda_1(\boldsymbol{\delta}^*) > 0, \lambda_2(\mu_1) > 0$, the transition probability of $(X_1(t), X_2(t))$ on $\mathbb{R}^\circ_{12+}$ converges to an invariant probability measure in total variation with an exponential rate. These results are similar to those appearing in Rudnicki (2003), Rudnicki and Pichór (2007). However, we generalize their results by obtaining a geometric rate of convergence.

REMARK 2.2. The condition for persistence in Rudnicki (2003), Rudnicki and Pichór (2007) is obtained by constructing a Lyapunov function V satisfying $\mathcal{L}V(\mathbf{x}) \le -a, \mathbf{x} \in \mathbb{R}^{2,\circ}_+$ for some $a > 0$. The papers describe how to construct the functions V rather than giving an explicit formula. It seems to us that the function V constructed in Rudnicki (2003) is not twice differentiable.

EXAMPLE 2.3. Consider a stochastic Lotka–Volterra cooperative model

$$\begin{cases} dX_1(t) = X_1(t)\big[a_1 - b_1X_1(t) + c_1X_2(t)\big]\,dt + X_1(t)\,dE_1(t), \\ dX_2(t) = X_2(t)\big[a_2 - b_2X_2(t) + c_2X_1(t)\big]\,dt + X_2(t)\,dE_2(t), \end{cases} \tag{2.6}$$

where $a_i, b_i, c_i > 0, a_i - \frac{\sigma_{ii}}{2} > 0, i = 1, 2$. As shown in Example 2.1, there exist unique invariant probability measures μ_i on $\mathbb{R}^\circ_{i+}, i = 1, 2$ (defined in Example 2.1). Moreover,

$$\lambda_2(\mu_1) = \int_{\mathbb{R}^\circ_{1+}} \left[a_2 - \frac{\sigma_{22}}{2} + c_2x_1\right]\mu_1(d\mathbf{x}) = a_2 - \frac{\sigma_{22}}{2} + c_2\frac{2a_1 - \sigma_{11}}{2b_1} > 0$$

and

$$\lambda_1(\mu_2) = \int_{\mathbb{R}^\circ_{2+}} \left[a_1 - \frac{\sigma_{11}}{2} + c_1x_2\right]\mu_2(d\mathbf{x}) = a_1 - \frac{\sigma_{11}}{2} + c_1\frac{2a_2 - \sigma_{22}}{2b_2} > 0.$$

Suppose further that

$$b_1b_2 - c_1c_2 < 0. \tag{2.7}$$

REMARK 2.3. We note that a similar example has been studied in Cattiaux and Méléard (2010). The main difference is that the authors of Cattiaux and Méléard (2010) consider demographic stochasticity instead of environmental stochasticity; their diffusion terms look like $\sqrt{X_i(t)}\,dE_i(t)$. In their setting, the diffusion hits one of the two axes in finite time almost surely and they study the existence of quasi-stationary distributions (since there are no nontrivial stationary distributions). We note however that they need $b_1b_2 - c_1c_2 > 0$ together with some other symmetry assumptions.

Standard computations show that part (3) of Assumption 1.1 is not satisfied by this model. Since $a_i - \frac{\sigma_{ii}}{2} > 0$ and $\lambda_i > 0, i = 1, 2$, Assumption 1.2 holds. However, we show that the solution either blows up in finite time almost surely or there is no invariant measure on $\mathbb{R}_+^{2,\circ}$.

We argue by contradiction. Suppose $(X_1(t), X_2(t))$ does not blow up in finite time and has an invariant measure on $\mathbb{R}_+^{2,\circ}$. As a result, $(X_1(t), X_2(t))$ is a recurrent process. It follows from Itô's formula that

$$\begin{aligned}\frac{b_2 \ln X_1(t) + b_1 \ln X_2(t)}{t} = {} & \frac{b_2 \ln X_1(0) + b_1 \ln X_2(0)}{t} \\ & + b_2\left(a_1 - \frac{\sigma_{11}}{2}\right) + b_1\left(a_2 - \frac{\sigma_{22}}{2}\right) \\ & + (c_1c_2 - b_1b_2)\frac{1}{t}\int_0^t X_1(s)\,ds \\ & + \frac{1}{t}\int_0^t \bigl(b_2\,dE_1(s) + b_1\,dE_2(s)\bigr).\end{aligned}$$

Since

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t \bigl(b_2\,dE_1(s) + b_1\,dE_2(s)\bigr) = 0 \qquad \text{a.s.},$$

$b_2(a_1 - \frac{\sigma_{11}}{2}) + b_1(a_2 - \frac{\sigma_{22}}{2}) > 0$ and $c_1c_2 - b_1b_2 > 0$, it follows that

$$\limsup_{t\to\infty} \frac{b_2 \ln X_1(t) + b_1 \ln X_2(t)}{t} > 0 \qquad \text{a.s.}$$

Thus, $(X_1(t), X_2(t))$ cannot be a recurrent process in $\mathbb{R}_+^{2,\circ}$. This is a contradiction.

EXAMPLE 2.4. Consider the two-dimensional system

$$dX_i(t) = X_i(t) f_i\bigl(\mathbf{X}(t)\bigr)\,dt + X_i(t) g_i\bigl(\mathbf{X}(t)\bigr)\,dE_i(t), \qquad i = 1, 2. \tag{2.8}$$

Suppose that Assumptions 1.1 and 1.4 hold. If $\lambda_i(\boldsymbol{\delta}^*) = f_i(\mathbf{0}) - \frac{1}{2}g_i^2(\mathbf{0})\sigma_{ii} > 0$, then $(\mathbf{X}(t))$ has a unique invariant probability measure μ_i on $\mathbb{R}_{i+}^{\circ}$ (which is defined as in Example 2.1). The density $p_i(\cdot)$ of μ_i can be found explicitly (in terms of integrals) by solving the Fokker–Plank equation

$$-\frac{d}{du}\big[p_i(u)f_i(\widehat{u})\big]+\frac{\sigma_{ii}}{2}\frac{d^2}{du^2}\big[p_i(u)g_i^2(\widehat{u})\big]=0, \qquad u>0,$$

where $\widehat{u}=(u,0)$ if $i=1$ and $\widehat{u}=(0,u)$ if $i=2$. Then $\lambda_j(\mu_i), i, j=1,2, i\neq j$ can be computed in terms of integrals. Using arguments similar to those in Examples 2.1 and 2.2, we have the following classification, which generalizes the Lotka–Volterra competitive and predator–prey models in previous examples:

- If $\lambda_i(\boldsymbol{\delta}^*)>0, i=1,2, \lambda_1(\mu_2)>0, \lambda_2(\mu_1)>0$, then there is a unique invariant probability measure π^* on $\mathbb{R}_+^{2,\circ}$ and $P(t,\mathbf{x},\cdot), \mathbf{x}\in\mathbb{R}_+^{2,\circ}$ converges to π^* in total variation exponentially fast.
- If $\lambda_i(\boldsymbol{\delta}^*)>0, \lambda_j(\boldsymbol{\delta}^*)<0, \lambda_j(\mu_i)>0$ for some $i=1,2$ and $j\neq i$, then there is a unique invariant probability measure π^* on $\mathbb{R}_+^{2,\circ}$ and $P(t,\mathbf{x},\cdot), \mathbf{x}\in\mathbb{R}_+^{2,\circ}$ converges to π^* in total variation exponentially fast.
- If $\lambda_i(\boldsymbol{\delta}^*)<0, i=1,2$ then $X_i(t)$ converges to 0 at the exponential rate $\lambda_i(\boldsymbol{\delta}^*), i=1,2$.
- If $\lambda_i(\boldsymbol{\delta}^*)>0, \lambda_j(\mu_i)<0$ and $\lambda_j(\boldsymbol{\delta}^*)<0$ for $i,j=1,2, i\neq j$, then $X_j(t)$ converges to 0 at the exponential rate $\lambda_j(\mu_i)$ and the randomized occupation measure converges weakly to μ_i.
- If $\lambda_i(\boldsymbol{\delta}^*)>0, \lambda_j(\mu_i)<0$ and $\lambda_j(\boldsymbol{\delta}^*)>0, \lambda_i(\mu_j)>0$ for $i,j=1,2, i\neq j$, then $X_j(t)$ converges to 0 at the exponential rate $\lambda_j(\mu_i)$ and the randomized occupation measure converges weakly to μ_i.
- If $\lambda_1(\boldsymbol{\delta}^*)>0, \lambda_2(\boldsymbol{\delta}^*)>0, \lambda_1(\mu_2)<0, \lambda_2(\mu_1)<0$, then $p_i^{\mathbf{x}}>0, i=1,2$ and $p_1^{\mathbf{x}}+p_2^{\mathbf{x}}=1$ where

$$p_i^{\mathbf{x}}=\mathbb{P}_{\mathbf{x}}\bigg\{\mathcal{U}(\omega)=\{\mu_i\} \text{ and } \lim_{t\to\infty}\frac{\ln X_j(t)}{t}=\lambda_j(\mu_i), j\in\{1,2\}\setminus\{i\}\bigg\}.$$

REMARK 2.4. Our methods can also be used to study the simple food chain

$$\begin{aligned}
dX_1(t) &= X_1(t)\big(a_{10}-a_{11}X_1(t)-a_{12}X_2(t)\big)\,dt+X_1(t)\,dE_1(t),\\
dX_2(t) &= X_2(t)\big(-a_{20}+a_{21}X_1(t)-a_{22}X_2(t)-a_{23}X_3(t)\big)\,dt\\
&\quad + X_2(t)\,dE_2(t),\\
&\ \vdots\\
dX_{n-1}(t) &= X_{n-1}(t)\big(-a_{n-1,0}+a_{n-1,n-2}X_{n-2}(t)\\
&\quad - a_{n-1,n-1}X_{n-1}(t)-a_{n-1,n}X_n\big)\,dt\\
&\quad + X_{n-1}(t)\,dE_{n-1}(t),\\
dX_n(t) &= X_n(t)\big(-a_{n0}+a_{n,n-1}X_{n-1}(t)-a_{nn}X_n(t)\big)\,dt\\
&\quad + X_n(t)\,dE_n(t).
\end{aligned} \tag{2.9}$$

In this model, X_1 describes a prey species, which is at the bottom of the food chain. The next $n-1$ species are predators. Species 1 has a per-capita growth rate $a_{10} > 0$ and its members compete for resources according to the intracompetition rate $a_{11} > 0$. Predator species j has a death rate $-a_{j0} < 0$, preys upon species $j-1$ at rate $a_{j,j-1} > 0$, competes with its own members at rate $a_{jj} > 0$ and is preyed upon by predator $j+1$ at rate $a_{j,j+1} > 0$. The last species, X_n, is considered to be the apex predator of the food chain. Using Theorems 1.1, 1.2 and 1.3, together with some linear algebra, we can have a sharp classification for the persistence and extinction of each species in the system (2.9). A detailed analysis of this can be found in Hening and Nguyen (2017a, 2017b).

3. Invariant measures, Lyapunov exponents and log-Laplace transforms. In this section, we explore some of the properties of the SDE (1.1). These will be used in later sections in order to prove the main results. In view of (1.2), there is an $M > 0$ such that

$$\begin{aligned}\Bigg[\frac{\sum_i c_i x_i f_i(\mathbf{x})}{1+\mathbf{c}^\top\mathbf{x}} - \frac{1}{2}\frac{\sum_{i,j}\sigma_{ij}c_ic_jx_ix_jg_i(\mathbf{x})g_j(\mathbf{x})}{(1+\mathbf{c}^\top\mathbf{x})^2} \\ + \gamma_b\Big(1+\sum_i\big(|f_i(\mathbf{x})| + g_i^2(\mathbf{x})\big)\Big)\Bigg] < 0\end{aligned} \tag{3.1}$$

if $\|\mathbf{x}\| \geq M$. Since

$$|g_i(\mathbf{x})g_j(\mathbf{x})\sigma_{ij}| \leq \frac{1}{2}|\sigma_{ij}|\big(|g_i(\mathbf{x})|^2 + |g_j(\mathbf{x})|^2\big)$$

we can find $\delta_0 \in (0, \frac{\gamma_b}{2}\wedge 1)$ such that

$$3\delta_0\sum_{i,j}|g_i(\mathbf{x})g_j(\mathbf{x})\sigma_{ij}| + \delta_0\sum_i g_i^2(\mathbf{x}) \leq \gamma_b\sum_i g_i^2(\mathbf{x}), \qquad \mathbf{x}\in\mathbb{R}_+^n. \tag{3.2}$$

By (3.1) and (3.2), we have

$$\begin{aligned}&\frac{\sum_i c_i x_i f_i(\mathbf{x})}{1+\mathbf{c}^\top\mathbf{x}} - \frac{1}{2}\frac{\sum_{i,j}\sigma_{ij}c_ic_jx_ix_jg_i(\mathbf{x})g_j(\mathbf{x})}{(1+\mathbf{c}^\top\mathbf{x})^2} \\ &\quad + \gamma_b + \delta_0\sum_i\big(|f_i(\mathbf{x})| + g_i^2(\mathbf{x})\big) \\ &\quad + 3\delta_0\sum_{i,j}|g_i(\mathbf{x})g_j(\mathbf{x})\sigma_{ij}| < 0 \qquad \text{for all } \|\mathbf{x}\| \geq M.\end{aligned} \tag{3.3}$$

For $\mathbf{p} = (p_1, \dots, p_n) \in \mathbb{R}_+^{n,\circ}$, $\|\mathbf{p}\| \leq \delta_0$, define the function $V : \mathbb{R}_+^{n,\circ} \to \mathbb{R}_+$ by

$$V(\mathbf{x}) := \frac{1+\mathbf{c}^\top\mathbf{x}}{\prod_i x_i^{p_i}}. \tag{3.4}$$

Using (3.3), one can define

$$
\begin{aligned}
H := \sup_{\mathbf{x}\in\mathbb{R}^n_+} \biggl\{ & \frac{\sum_i c_i x_i f_i(\mathbf{x})}{1+\mathbf{c}^\top \mathbf{x}} - \frac{1}{2}\frac{\sum_{i,j} \sigma_{ij} c_i c_j x_i x_j g_i(\mathbf{x}) g_j(\mathbf{x})}{(1+\mathbf{c}^\top\mathbf{x})^2} \\
& + \gamma_b + \delta_0 \sum_i \bigl(|f_i(\mathbf{x})| + g_i^2(\mathbf{x})\bigr) + 3\delta_0 \sum_{i,j} |g_i(\mathbf{x}) g_j(\mathbf{x}) \sigma_{ij}| \biggr\} \\
< \infty. &
\end{aligned} \tag{3.5}
$$

LEMMA 3.1. *For any* $\mathbf{x} \in \mathbb{R}^n_+$, *there exists a pathwise unique strong solution* $(\mathbf{X}(t))$ *to* (1.1) *with initial value* $\mathbf{X}(0) = \mathbf{x}$. *Let* $I \subset \{1, \dots, n\}$ *and* $\mathbf{x} \in \mathbb{R}^{I,\circ}_+$ *where*

$$
\mathbb{R}^{I,\circ}_+ = \{\mathbf{x} \in \mathbb{R}^n_+ : x_i = 0 \text{ if } i \notin I \text{ and } x_i > 0 \text{ if } i \in I\}.
$$

The solution $(\mathbf{X}(t))$ *with initial value* $\mathbf{x}$ *will stay forever in* $\mathbb{R}^{I,\circ}_+$ *with probability* 1. *Moreover, for* $\mathbf{x} \in \mathbb{R}^{n,\circ}_+$ *and* V *defined by* (3.4), *we have*

$$
\mathbb{E}_\mathbf{x} V^{\delta_0}\bigl(\mathbf{X}(t)\bigr) \le \exp(\delta_0 H t) V^{\delta_0}(\mathbf{x}). \tag{3.6}
$$

LEMMA 3.2. *There are* $H_1, H_2 > 0$ *such that for any* $\mathbf{x} \in \mathbb{R}^n_+$, $t > 0$

$$
\mathbb{E}_\mathbf{x}\bigl(1 + \mathbf{c}^\top \mathbf{X}(t)\bigr)^{\delta_0} \le H_1 + \bigl(1 + \mathbf{c}^\top\mathbf{x}\bigr)^{\delta_0} e^{-\delta_0 \gamma_b t} \tag{3.7}
$$

and

$$
\begin{aligned}
&\mathbb{E}_\mathbf{x} \int_0^t \bigl(1 + \mathbf{c}^\top \mathbf{X}(s)\bigr)^{\delta_0} \biggl[1 + \sum_i \bigl(|f_i(\mathbf{X}(s))| + |g_i(\mathbf{X}(s))|^2\bigr)\biggr] ds \\
&\quad \le H_2\bigl((1 + \mathbf{c}^\top\mathbf{x})^{\delta_0} + t\bigr).
\end{aligned} \tag{3.8}
$$

Moreover, the solution process $(\mathbf{X}(t))$ *is a Feller process on* $\mathbb{R}^n_+$.

REMARK 3.1. There are different possible definitions of "Feller" in the literature. What we mean by Feller is that the semigroup $(T_t)_{t\ge 0}$ of the process maps the set of bounded continuous functions $C_b(\mathbb{R}^n_+)$ into itself, that is,

$$
T_t\bigl(C_b(\mathbb{R}^n_+)\bigr) \subset C_b(\mathbb{R}^n_+), \qquad t \ge 0.
$$

Define the family of measures:

$$
\Pi^\mathbf{x}_t(\cdot) := \frac{1}{t}\int_0^t \mathbb{P}_\mathbf{x}\{\mathbf{X}(s) \in \cdot\}\, ds, \qquad \mathbf{x} \in \mathbb{R}^n_+, t > 0.
$$

LEMMA 3.3. *Let* μ *be an invariant probability measure of* $\mathbf{X}$. *Then*

$$
\int_{\mathbb{R}^n_+} (1 + \mathbf{c}^\top\mathbf{x})^{\delta_0} \biggl(1 + \sum_i \bigl(|f_i(\mathbf{x})| + |g_i(\mathbf{x})|^2\bigr)\biggr) \mu(d\mathbf{x}) \le H_2
$$

and

$$\int_{\mathbb{R}^n_+}\left(\frac{\sum_i c_i x_i f_i(\mathbf{x})}{1+\mathbf{c}^\top\mathbf{x}} - \frac{1}{2}\frac{\sum_{i,j}\sigma_{ij}c_ic_jx_ix_jg_i(\mathbf{x})g_j(\mathbf{x})}{(1+\mathbf{c}^\top\mathbf{x})^2}\right)\mu(d\mathbf{x}) = 0.$$

LEMMA 3.4. *Suppose the following*:

- *The sequences* $(\mathbf{x}_k)_{k\in N}\subset\mathbb{R}^n_+$, $(T_k)_{k\in\mathbb{N}}\subset\mathbb{R}_+$ *are such that* $\|\mathbf{x}_k\|\le M$, $T_k>1$ *for all* $k\in\mathbb{N}$ *and* $\lim_{k\to\infty}T_k=\infty$.
- *The sequence* $(\Pi^{\mathbf{x}_k}_{T_k})_{k\in\mathbb{N}}$ *converges weakly to an invariant probability measure* π.
- *The function* $h:\mathbb{R}^n_+\to\mathbb{R}$ *is any continuous function satisfying* $|h(\mathbf{x})|\le K_h(1+\mathbf{c}^\top\mathbf{x})^\delta(1+\sum_i(|f_i(\mathbf{x})|+|g_i(\mathbf{x})|^2))$, $\mathbf{x}\in\mathbb{R}^n_+$, *for some* $K_h\ge 0$, $\delta<\delta_0$.

Then one has

$$\lim_{k\to\infty}\int_{\mathbb{R}^n_+}h(\mathbf{x})\Pi^{\mathbf{x}_k}_{T_k}(d\mathbf{x}) = \int_{\mathbb{R}^n_+}h(\mathbf{x})\pi(d\mathbf{x}).$$

LEMMA 3.5. *Let* Y *be a random variable*, $\theta_0>0$ *a constant, and suppose*

$$\mathbb{E}\exp(\theta_0 Y)+\mathbb{E}\exp(-\theta_0 Y)\le K_1.$$

Then the log-Laplace transform $\phi(\theta)=\ln\mathbb{E}\exp(\theta Y)$ *is twice differentiable on* $[0,\frac{\theta_0}{2})$ *and*

$$\frac{d\phi}{d\theta}(0)=\mathbb{E}Y,$$

$$0\le\frac{d^2\phi}{d\theta^2}(\theta)\le K_2,\qquad \theta\in\left[0,\frac{\theta_0}{2}\right)$$

for some $K_2>0$ *depending only on* K_1.

REMARK 3.2. We note that we got the very nice idea of using the log-Laplace transform in the proofs of our persistence results from the manuscript Benaïm (2014).

To proceed, let us recall some technical concepts and results needed to prove the main theorem. Let $\boldsymbol{\Phi}=(\Phi_0,\Phi_1,\dots)$ be a discrete-time Markov chain on a general state space $(E,\mathcal{E})$, where $\mathcal{E}$ is a countably generated σ-algebra. Denote by $\mathcal{P}$ the Markov transition kernel for $\boldsymbol{\Phi}$. If there is a nontrivial σ-finite positive measure φ on $(E,\mathcal{E})$, such that for any $A\in\mathcal{E}$ satisfying $\varphi(A)>0$ we have

$$\sum_{n=1}^{\infty}\mathcal{P}^n(x,A)>0,\qquad x\in E,$$

where $\mathcal{P}^n$ is the n-step transition kernel of $\boldsymbol{\Phi}$, then the Markov chain $\boldsymbol{\Phi}$ is called *irreducible*. It can be shown [see Nummelin (1984)] that if $\boldsymbol{\Phi}$ is irreducible, then there exists a positive integer d and disjoint subsets $E_0, \dots, E_{d-1}$ such that for all $i = 0, \dots, d-1$ and all $x \in E_i$, we have

$$\mathcal{P}(x, E_j) = 1 \qquad \text{where } j = i + 1 \ (\text{mod } d).$$

The smallest positive integer d satisfying the above is called the period of $\mathbf{\Phi}$. An *aperiodic* Markov chain is a chain with period $d = 1$.

A set $C \in \mathcal{E}$ is called *petite*, if there exists a nonnegative sequence $(a_n)_{n\in\mathbb{N}}$ with $\sum_{n=1}^{\infty} a_n = 1$ and a nontrivial positive measure ν on $(E, \mathcal{E})$ such that

$$\sum_{n=1}^{\infty} a_n \mathcal{P}^n(x, A) \geq \nu(A), \qquad x \in C, A \in \mathcal{E}.$$

We have the following lemma.

LEMMA 3.6. *For any $T > 0$, the Markov chain $\{\mathbf{X}(kT), k \in \mathbb{N}\}$ on $\mathbb{R}_+^{n,\circ}$ is irreducible and aperiodic. Moreover, every compact set $K \subset \mathbb{R}_+^{n,\circ}$ is petite.*

The proofs of the above lemmas are collected in the Appendix.

4. Persistence. This section is devoted to finding conditions under which $\mathbf{X}$ converges to a unique invariant probability measure supported on $\mathbb{R}_+^{n,\circ}$.

It is shown in Schreiber, Benaïm and Atchadé [(2011), Lemma 4], by the min-max principle that Assumption 1.2 is equivalent to the existence of $\mathbf{p} > 0$ such that

$$\min_{\mu\in\mathcal{M}} \left\{ \sum_i p_i \lambda_i(\mu) \right\} := 2\rho^* > 0. \tag{4.1}$$

By rescaling if necessary, we can assume that $\|\mathbf{p}\| = \delta_0$.

LEMMA 4.1. *Suppose that Assumption 1.2 holds. Let $\mathbf{p}$ and ρ^* be as in (4.1). There exists a $T^* > 0$ such that, for any $T > T^*$, $\mathbf{x} \in \partial\mathbb{R}_+^n$, $\|\mathbf{x}\| \leq M$ one has*

$$\frac{1}{T} \int_0^T \mathbb{E}_{\mathbf{x}} \Phi\big(\mathbf{X}(t)\big)\, dt \leq -\rho^*, \tag{4.2}$$

where

$$\begin{aligned} \Phi(\mathbf{x}) := {} & \frac{\sum_i c_i x_i f_i(\mathbf{x})}{1 + \mathbf{c}^\top \mathbf{x}} - \frac{1}{2} \frac{\sum_{i,j} \sigma_{ij} c_i c_j x_i x_j g_i(\mathbf{x}) g_j(\mathbf{x})}{(1 + \mathbf{c}^\top \mathbf{x})^2} \\ & - \sum_i p_i \left(f_i(\mathbf{x}) - \frac{\sigma_{ii} g_i^2(\mathbf{x})}{2} \right). \end{aligned} \tag{4.3}$$

PROOF. We argue by contradiction. Suppose that the conclusion of this lemma is not true. Then we can find $\mathbf{x}_k \in \partial\mathbb{R}^n_+$, $\|\mathbf{x}_k\| \le M$ and $T_k > 0$, $\lim_{k\to\infty} T_k = \infty$ such that

$$\frac{1}{T_k}\int_0^{T_k} \mathbb{E}_{\mathbf{x}_k}\Phi\big(\mathbf{X}(t)\big)\,dt > -\rho^*, \qquad k \in \mathbb{N}. \tag{4.4}$$

Note that

$$\Pi_t^{\mathbf{x}_k}(d\mathbf{y}) := \frac{1}{t}\int_0^t \mathbb{P}_{\mathbf{x}_k}\{\mathbf{X}(s) \in d\mathbf{y}\}\,ds.$$

By Tonelli's theorem, we get that

$$\begin{aligned}\int_{\mathbb{R}^n_+}(1+\mathbf{c}^\top\mathbf{y})^{\delta_0}\Pi_t^{\mathbf{x}_k}(d\mathbf{y}) &= \int_{\mathbb{R}^n_+}(1+\mathbf{c}^\top\mathbf{y})^{\delta_0}\frac{1}{t}\int_0^t \mathbb{P}_{\mathbf{x}_k}\{\mathbf{X}(s)\in d\mathbf{y}\}\,ds \\ &= \frac{1}{t}\int_0^t \mathbb{E}_{\mathbf{x}_k}\big(1+\mathbf{c}^\top\mathbf{X}(s)\big)^{\delta_0}\,ds.\end{aligned} \tag{4.5}$$

It follows from Lemma 3.2 that

$$\begin{aligned}\sup_{k\in\mathbb{N},t\ge0}\int_{\mathbb{R}^n_+}(1+\mathbf{c}^\top\mathbf{y})^{\delta_0}\Pi_t^{\mathbf{x}_k}(d\mathbf{y}) &= \sup_{k\in\mathbb{N},t\ge0}\frac{1}{t}\int_0^t \mathbb{E}_{\mathbf{x}_k}\big(1+\mathbf{c}^\top\mathbf{X}(s)\big)^{\delta_0}\,ds \\ &\le \sup_{\|\mathbf{x}\|\le M,t\ge0}\frac{1}{t}\int_0^t\big(H_1+(1+\mathbf{c}^\top\mathbf{x})^{\delta_0}e^{-\delta_0\gamma_b s}\big)\,ds \\ &< \infty.\end{aligned} \tag{4.6}$$

This implies that the family $(\Pi_{T_k}^{\mathbf{x}_k})_{k\in\mathbb{N}}$ is tight in $\mathbb{R}^n_+$. As a result, $(\Pi_{T_k}^{\mathbf{x}_k})_{k\in\mathbb{N}}$ has a convergent subsequence in the weak*-topology. Without loss of generality, we can suppose that $\{\Pi_{T_k}^{\mathbf{x}_k} : k \in \mathbb{N}\}$ is a convergent sequence in the weak* topology. It can be shown (by Ethier and Kurtz [(2009), Theorem 9.9], or by Evans, Hening and Schreiber [(2015), Proposition 6.4]), that its limit is an invariant probability measure μ of $\mathbf{X}$. As a consequence of Lemma 3.4,

$$\lim_{k\to\infty}\frac{1}{T_k}\int_0^{T_k}\mathbb{E}_{\mathbf{x}_k}\Phi\big(\mathbf{X}(t)\big)\,dt = \int_{\mathbb{R}^n_+}\Phi(\mathbf{x})\mu(d\mathbf{x}).$$

In view of Lemma 3.3 and (4.1), we get that

$$\lim_{k\to\infty}\frac{1}{T_k}\int_0^{T_k}\mathbb{E}_{\mathbf{x}_k}\Phi\big(\mathbf{X}(t)\big)\,dt = -\sum_{i=1}^n p_i\lambda_i(\mu) \le -2\rho^*,$$

which contradicts (4.4). □

From now on, let $n^* \in \mathbb{N}$ be such that

$$\gamma_b\big(n^*-1\big) > H. \tag{4.7}$$

PROPOSITION 4.1. *Let $V(\cdot)$ be defined by* (3.4) *with* $\mathbf{p}$ *and* ρ^* *satisfying* (4.1) *and* $T^* > 0$ *satisfying the assumptions of Lemma* 4.1. *There are* $\theta \in (0, \frac{\delta_0}{2})$, $K_\theta > 0$, *such that for any* $T \in [T^*, n^*T^*]$ *and* $\mathbf{x} \in \mathbb{R}_+^{n,\circ}$, $\|\mathbf{x}\| \le M$,

$$\mathbb{E}_\mathbf{x} V^\theta\big(\mathbf{X}(T)\big) \le V^\theta(\mathbf{x}) \exp\left(-\frac{1}{2}\theta\rho^* T\right) + K_\theta.$$

PROOF. We have from Itô's formula that

$$\ln V\big(\mathbf{X}(T)\big) = \ln V\big(\mathbf{X}(0)\big) + G(T), \tag{4.8}$$

where

$$\begin{aligned} G(T) = {} & \int_0^T \Phi\big(\mathbf{X}(t)\big)\,dt \\ & + \int_0^T \left[\frac{\sum_i c_i X_i(t) g_i(\mathbf{X}(t))\,dE_i(t)}{1+\mathbf{c}^\top \mathbf{X}(t)} - \sum_i p_i g_i\big(\mathbf{X}(t)\big)\,dE_i(t)\right]. \end{aligned} \tag{4.9}$$

In view of (4.8) and (3.6),

$$\mathbb{E}_\mathbf{x} \exp\big(\delta_0 G(T)\big) = \frac{\mathbb{E}_\mathbf{x} V^{\delta_0}(\mathbf{X}(T))}{V^{\delta_0}(\mathbf{x})} \le \exp(\delta_0 H T). \tag{4.10}$$

Let $\widehat{V}(\cdot) : \mathbb{R}_+^{n,\circ} \mapsto \mathbb{R}_+$ be defined by $\widehat{V}(\mathbf{x}) = (1 + \mathbf{c}^\top \mathbf{x}) \prod_{i=1}^n x_i^{p_i}$. We can use (3.3) and some of the estimates from the proof of Lemma 3.1 to obtain

$$\frac{\mathbb{E}_\mathbf{x} \widehat{V}^{\delta_0}(\mathbf{X}(T))}{\widehat{V}^{\delta_0}(\mathbf{x})} \le \exp(\delta_0 H T). \tag{4.11}$$

Note that

$$V^{-\delta_0}(\mathbf{x}) = \widehat{V}^{\delta_0}(\mathbf{x})\big(1 + \mathbf{c}^\top \mathbf{x}\big)^{-2\delta_0} \le \widehat{V}^{\delta_0}(\mathbf{x}). \tag{4.12}$$

Applying (4.12) to (4.11) yields

$$\begin{aligned} \mathbb{E}_\mathbf{x} \exp\big(-\delta_0 G(T)\big) &= \frac{\mathbb{E}_\mathbf{x} V^{-\delta_0}(\mathbf{X}(T))}{V^{-\delta_0}(\mathbf{x})} \\ &\le \frac{\mathbb{E}_\mathbf{x} \widehat{V}^{\delta_0}(\mathbf{X}(T))}{V^{-\delta_0}(\mathbf{x})} \\ &\le \frac{\mathbb{E}_\mathbf{x} \widehat{V}^{\delta_0}(\mathbf{X}(T))}{\widehat{V}^{\delta_0}(\mathbf{x})} \big(1 + \mathbf{c}^\top \mathbf{x}\big)^{2\delta_0} \\ &\le \big(1 + \mathbf{c}^\top \mathbf{x}\big)^{2\delta_0} \exp(\delta_0 H T). \end{aligned} \tag{4.13}$$

By (4.10) and (4.13), the assumptions of Lemma 3.5 hold for $G(T)$. Therefore, there is $\tilde{K}_2 \ge 0$ such that

$$0 \le \frac{d^2 \tilde{\phi}_{\mathbf{x},T}}{d\theta^2}(\theta) \le \tilde{K}_2 \qquad \text{for all } \theta \in \left[0, \frac{\delta_0}{2}\right), \mathbf{x} \in \mathbb{R}_+^{n,\circ}, \|\mathbf{x}\| \le M, T \in [T^*, n^*T^*],$$

where

$$\tilde{\phi}_{\mathbf{x},T}(\theta) = \ln \mathbb{E}_{\mathbf{x}} \exp\bigl(\theta G(T)\bigr).$$

In view of Lemma 4.1 and the Feller property of $(\mathbf{X}(t))$, there exists a $\tilde{\delta} > 0$ such that if $\|\mathbf{x}\| \le M$, $\operatorname{dist}(\mathbf{x}, \partial\mathbb{R}^n_+) < \tilde{\delta}$ and $T \in [T^*, n^*T^*]$ then

$$\begin{aligned}\mathbb{E}_{\mathbf{x}} G(T) = {} & \int_0^T \mathbb{E}_{\mathbf{x}} \biggl(\frac{\sum_i c_i X_i(t) f_i(\mathbf{X}(t))}{1 + \mathbf{c}^\top \mathbf{X}(t)} \\ & - \frac{\sum_{i,j} c_i c_j X_i(t) X_j(t) g_i(\mathbf{X}(t)) g_j(\mathbf{X}(t)) \sigma_{ij}}{2(1 + \mathbf{c}^\top \mathbf{X}(t))^2} \biggr) dt \\ & - \sum_{i=1}^n p_i \int_0^T \mathbb{E}_{\mathbf{x}} \biggl(f_i(\mathbf{X}(t)) - \frac{\sigma_{ii} g_i^2(\mathbf{X}(t))}{2} \biggr) dt \\ \le {} & -\frac{3}{4} \rho^* T.\end{aligned} \tag{4.14}$$

Another application of Lemma 3.5 yields

$$\frac{d\tilde{\phi}_{\mathbf{x},T}}{d\theta}(0) = \mathbb{E}_{\mathbf{x}} G(T) \le -\frac{3}{4}\rho^* T \qquad \text{for } \mathbf{x} \in \mathbb{R}^{n,\circ}_+, \|\mathbf{x}\| \le M, T \in \bigl[T^*, n^*T^*\bigr].$$

By a Taylor expansion around $\theta = 0$, for $\|\mathbf{x}\| \le M, \operatorname{dist}(\mathbf{x}, \partial\mathbb{R}^n_+) < \tilde{\delta}, T \in [T^*, n^*T^*]$ and $\theta \in [0, \frac{\delta_0}{2})$ we have

$$\tilde{\phi}_{\mathbf{x},T}(\theta) \le -\frac{3}{4}\rho^* T\theta + \theta^2 \tilde{K}_2.$$

If we choose any $\theta \in (0, \frac{\delta_0}{2})$ satisfying $\theta < \frac{\rho^* T^*}{4\tilde{K}_2}$, we obtain that

$$\begin{aligned}&\tilde{\phi}_{\mathbf{x},T}(\theta) \le -\frac{1}{2}\rho^* T\theta \\ &\qquad \text{for all } \mathbf{x} \in \mathbb{R}^{n,\circ}, \|\mathbf{x}\| \le M, \operatorname{dist}\bigl(\mathbf{x}, \partial\mathbb{R}^n_+\bigr) < \tilde{\delta}, T \in \bigl[T^*, n^*T^*\bigr].\end{aligned} \tag{4.15}$$

In light of (4.15), we have for such θ and $\|\mathbf{x}\| \le M, 0 < \operatorname{dist}(\mathbf{x}, \partial\mathbb{R}^n_+) < \tilde{\delta}, T \in [T^*, n^*T^*]$ that

$$\frac{\mathbb{E}_{\mathbf{x}} V^\theta(\mathbf{X}(T))}{V^\theta(\mathbf{x})} = \exp \tilde{\phi}_{\mathbf{x},T}(\theta) \le \exp\biggl(-\frac{1}{2}\rho^* T\theta\biggr). \tag{4.16}$$

In view of (3.6), we have for $\mathbf{x}$ satisfying $\|\mathbf{x}\| \le M, \operatorname{dist}(\mathbf{x}, \partial\mathbb{R}^n_+) \ge \tilde{\delta}$ and $T \in [T^*, n^*T^*]$ that

$$\mathbb{E}_{\mathbf{x}} V^\theta\bigl(\mathbf{X}(T)\bigr) \le \exp\bigl(\theta n^* T^* H\bigr) \sup_{\|\mathbf{x}\| \le M, \operatorname{dist}(\mathbf{x}, \partial\mathbb{R}^n_+) \ge \tilde{\delta}} \bigl\{V(\mathbf{x})\bigr\} =: K_\theta < \infty. \tag{4.17}$$

The desired result follows from (4.16) and (4.17). $\square$

THEOREM 4.1. *Suppose that Assumptions 1.1 and 1.2 hold. Let θ be as in Proposition 4.1, n^* as in (4.7). There are $\kappa = \kappa(\theta, T^*) \in (0, 1)$, $\tilde{K} = \tilde{K}(\theta, T^*) > 0$ such that*

$$\mathbb{E}_{\mathbf{x}} V^\theta\big(\mathbf{X}(n^*T^*)\big) \le \kappa V^\theta(x) + \tilde{K} \qquad \text{for all } \mathbf{x} \in \mathbb{R}_+^{n,\circ}. \tag{4.18}$$

As a result, $\mathbf{X}$ is strongly persistent. Furthermore, the convergence of its transition probability in total variation to its unique probability measure π^* on $\mathbb{R}_+^{n,\circ}$ is exponentially fast. For any initial value $\mathbf{x} \in \mathbb{R}_+^{n,\circ}$ and any π^*-integrable function f, we have

$$\mathbb{P}_{\mathbf{x}}\left\{\lim_{T\to\infty} \frac{1}{T}\int_0^T f\big(\mathbf{X}(t)\big)\,dt = \int_{\mathbb{R}_+^{n,\circ}} f(\mathbf{u})\pi^*(d\mathbf{u})\right\} = 1. \tag{4.19}$$

PROOF. By direct calculation and using (3.3), we have

$$\mathcal{L}V^\theta(\mathbf{x}) \le -\theta\gamma_b V^\theta(\mathbf{x}) \qquad \text{if } \|x\| \ge M. \tag{4.20}$$

Define

$$\tau = \inf\big\{t \ge 0 : \|\mathbf{X}(t)\| \le M\big\}. \tag{4.21}$$

In view of (4.20), we can obtain from Dynkin's formula that

$$\begin{aligned}
&\mathbb{E}_{\mathbf{x}}\big[\exp\big(\theta\gamma_b(\tau \wedge n^*T^*)\big)V^\theta\big(\mathbf{X}(\tau\wedge n^*T^*)\big)\big] \\
&\quad\le V^\theta(\mathbf{x}) + \mathbb{E}_{\mathbf{x}}\int_0^{\tau\wedge n^*T^*} \exp(\theta\gamma_b s)\big[\mathcal{L}V^\theta\big(\mathbf{X}(s)\big) + \theta\gamma_b V^\theta\big(\mathbf{X}(s)\big)\big]\,ds \\
&\quad\le V^\theta(\mathbf{x}).
\end{aligned}$$

Thus,

$$\begin{aligned}
V^\theta(\mathbf{x}) &\ge \mathbb{E}_{\mathbf{x}}\big[\exp\big(\theta\gamma_b(\tau\wedge n^*T^*)\big)V^\theta\big(\mathbf{X}(\tau\wedge n^*T^*)\big)\big] \\
&= \mathbb{E}_{\mathbf{x}}\big[\mathbf{1}_{\{\tau\le(n^*-1)T^*\}}\exp\big(\theta\gamma_b(\tau\wedge n^*T^*)\big)V^\theta\big(\mathbf{X}(\tau\wedge n^*T^*)\big)\big] \\
&\quad + \mathbb{E}_{\mathbf{x}}\big[\mathbf{1}_{\{(n^*-1)T^*<\tau<n^*T^*\}}\exp\big(\theta\gamma_b(\tau\wedge n^*T^*)\big)V^\theta\big(\mathbf{X}(\tau\wedge n^*T^*)\big)\big] \\
&\quad + \mathbb{E}_{\mathbf{x}}\big[\mathbf{1}_{\{\tau\ge n^*T^*\}}\exp\big(\theta\gamma_b(\tau\wedge n^*T^*)\big)V^\theta\big(\mathbf{X}(\tau\wedge n^*T^*)\big)\big] \\
&\ge \mathbb{E}_{\mathbf{x}}\big[\mathbf{1}_{\{\tau\le(n^*-1)T^*\}}V^\theta\big(\mathbf{X}(\tau)\big)\big] \\
&\quad + \exp\big(\theta\gamma_b(n^*-1)T^*\big)\mathbb{E}_{\mathbf{x}}\big[\mathbf{1}_{\{(n^*-1)T^*<\tau<n^*T^*\}}V^\theta\big(\mathbf{X}(\tau)\big)\big] \\
&\quad + \exp\big(\theta\gamma_b n^*T^*\big)\mathbb{E}_{\mathbf{x}}\big[\mathbf{1}_{\{\tau\ge n^*T^*\}}V^\theta\big(\mathbf{X}(n^*T^*)\big)\big].
\end{aligned} \tag{4.22}$$

By the strong Markov property of $(\mathbf{X}(t))$ and Proposition 4.1, we obtain

$$\begin{aligned}
&\mathbb{E}_{\mathbf{x}}\big[\mathbf{1}_{\{\tau\le(n^*-1)T^*\}}V^\theta\big(\mathbf{X}(n^*T^*)\big)\big] \\
&\quad\le \mathbb{E}_{\mathbf{x}}\big[\mathbf{1}_{\{\tau\le(n^*-1)T^*\}}\big[K_\theta + e^{-\frac{1}{2}\theta p^*(n^*T^*-\tau)}V^\theta\big(\mathbf{X}(\tau)\big)\big]\big] \\
&\quad\le K_\theta + \exp\left(-\frac{1}{2}\theta\rho^*T^*\right)\mathbb{E}_{\mathbf{x}}\big[\mathbf{1}_{\{\tau\le(n^*-1)T^*\}}V^\theta\big(\mathbf{X}(\tau)\big)\big].
\end{aligned} \tag{4.23}$$

By making use again of the strong Markov property of $(\mathbf{X}(t))$ and (3.6), we get

$$
\begin{aligned}
&\mathbb{E}_{\mathbf{x}}\big[\mathbf{1}_{\{(n^*-1)T^*<\tau<n^*T^*\}}V^\theta\big(\mathbf{X}\big(n^*T^*\big)\big)\big]\\
&\quad\le \mathbb{E}_{\mathbf{x}}\big[\mathbf{1}_{\{(n^*-1)T^*<\tau<n^*T^*\}}e^{\theta H(n^*T^*-\tau)}V^\theta\big(\mathbf{X}(\tau)\big)\big]\\
&\quad\le \exp\big(\theta HT^*\big)\mathbb{E}_{\mathbf{x}}\big[\mathbf{1}_{\{(n^*-1)T^*<\tau<n^*T^*\}}\big(V^\theta\big(\mathbf{X}(\tau)\big)\big)\big].
\end{aligned}
\tag{4.24}
$$

Applying (4.23) and (4.24) to (4.22) yields

$$
\begin{aligned}
V^\theta(x) &\ge \mathbb{E}_{\mathbf{x}}\big[\mathbf{1}_{\{\tau\le(n^*-1)T^*\}}V^\theta\big(\mathbf{X}(\tau)\big)\big]\\
&\quad+\exp\big(\theta\gamma_b\big(n^*-1\big)T^*\big)\mathbb{E}_{\mathbf{x}}\big[\mathbf{1}_{\{(n^*-1)T^*<\tau<n^*T^*\}}V^\theta\big(\mathbf{X}(\tau)\big)\big]\\
&\quad+\exp\big(\theta\gamma_b n^*T^*\big)\mathbb{E}_{\mathbf{x}}\big[\mathbf{1}_{\{\tau\ge n^*T^*\}}V^\theta\big(\mathbf{X}\big(n^*T^*\big)\big)\big]\\
&\ge \exp\left(\frac{1}{2}\theta\rho^*T^*\right)\mathbb{E}_{\mathbf{x}}\big[\mathbf{1}_{\{\tau\le(n^*-1)T^*\}}V^\theta\big(\mathbf{X}\big(n^*T^*\big)\big)\big]\\
&\quad-\exp\left(\frac{1}{2}\theta\rho^*T^*\right)K_\theta+\exp\big(-\theta HT^*\big)\exp\big(\theta\gamma_b\big(n^*-1\big)T^*\big)\\
&\quad\times\mathbb{E}_{\mathbf{x}}\big[\mathbf{1}_{\{(n^*-1)T^*<\tau<n^*T^*\}}V^\theta\big(\mathbf{X}\big(n^*T^*\big)\big)\big]\\
&\quad+\exp\big(\theta\gamma_b n^*T^*\big)\mathbb{E}_{\mathbf{x}}\big[\mathbf{1}_{\{\tau\ge n^*T^*\}}V^\theta\big(\mathbf{X}\big(n^*T^*\big)\big)\big]\\
&\ge \exp\big(m\theta T^*\big)\mathbb{E}_{\mathbf{x}}V^\theta\big(\mathbf{X}\big(n^*T^*\big)\big)-K_\theta\exp\left(\frac{1}{2}\theta\rho^*T^*\right),
\end{aligned}
\tag{4.25}
$$

where $m=\min\{\frac{1}{2}\rho^*,\gamma_b n^*,\gamma_b(n^*-1)-H\}>0$ by (4.7). The proof of (4.18) is complete by taking $\kappa=\exp(-m\theta T^*)$ and

$$
\tilde{K}=K_\theta\exp\left(\frac{1}{2}\theta\rho^*T^*\right)\exp\big(-m\theta T^*\big).
$$

By Lemma 3.6, the Markov chain $\{\mathbf{X}(kn^*T^*):k\in\mathbb{N}\}$ is irreducible and aperiodic. Moreover, each compact subset of $\mathbb{R}_+^{n,\circ}$ is petite. Applying the second corollary of Meyn and Tweedie [(1992), Theorem 6.2], we deduce from (4.25) that

$$
\big\|P\big(kn^*T^*,\mathbf{x},\cdot\big)-\pi^*(\cdot)\big\|_{\mathrm{TV}}\le C_{\mathbf{x}}r^k,
\tag{4.26}
$$

where π^* is an invariant probability measure of $\{\mathbf{X}(kn^*T^*),k\in\mathbb{N}\}$ on $\mathbb{R}_+^{n,\circ}$, for some $r\in(0,1)$ and $C_{\mathbf{x}}>0$ a constant depending on $\mathbf{x}\in\mathbb{R}_+^{n,\circ}$.

On the other hand, it follows from (4.25) and Meyn and Tweedie [(1992), Theorem 6.2], that for any compact set $K\subset\mathbb{R}_+^{n,\circ}$, we have $\mathbb{E}_{\mathbf{x}}\tau_K^*<\infty$ where τ_K^* is the first time the Markov chain $\{\mathbf{X}(kn^*T^*),k\in\mathbb{N}\}$ enters K. Thus, the process $\mathbf{X}$ is a positive recurrent diffusion, or equivalently, $\mathbf{X}$ has a unique invariant probability measure on $\mathbb{R}_+^{n,\circ}$ [see, e.g., Khasminskii (2012), Chapter 4]. Because of (4.26), the unique invariant probability measure of the process $\mathbf{X}$ must be π^*. Moreover, it is well known that $\|P(t,\mathbf{x},\cdot)-\pi^*(\cdot)\|_{\mathrm{TV}}$ is decreasing in t (it can be shown easily using the Kolmogorov–Chapman equation). We therefore obtain an exponential upper bound for $\|P(t,\mathbf{x},\cdot)-\pi^*(\cdot)\|_{\mathrm{TV}}$. $\square$

5. Extinction. This section is devoted to the study of conditions under which some of the species will go extinct with strictly positive probability.

LEMMA 5.1. *For any $\mu \in \mathcal{M}$ and $i \in I_\mu$, we have $\lambda_i(\mu) = 0$.*

PROOF. In view of Itô's formula,

$$\frac{\ln X_i(t)}{t} = \frac{\ln X_i(0)}{t} + \frac{1}{t}\int_0^t \left[f_i(\mathbf{X}(s)) - \frac{g_i^2(\mathbf{X}(s))\sigma_{ii}}{2}\right] ds + \frac{1}{t}\int_0^t g_i(\mathbf{X}(s))\, dE_i(s).$$

In the same manner as in the second part of the proof of Lemma 3.3, we can show that if $\mathbf{X}(0) = \mathbf{x}_0 \in \mathbb{R}_+^{\mu,\circ}$ and $i \in I_\mu$, then

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t \left[f_i(\mathbf{X}(s)) - \frac{g_i^2(\mathbf{X}(s))\sigma_{ii}}{2}\right] ds = \lambda_i(\mu) \qquad \mathbb{P}_{\mathbf{x}_0}\text{-a.s.}$$

and

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t g_i(\mathbf{X}(s))\, dE_i(s) = 0 \qquad \mathbb{P}_{\mathbf{x}_0}\text{-a.s.}$$

Consequently,

$$\lim_{t\to\infty} \frac{\ln X_i(t)}{t} = \lambda_i(\mu) \qquad \mathbb{P}_{\mathbf{x}_0}\text{-a.s.}$$

On the other hand, because the diffusion is nondegenerate, and the process $\mathbf{X}$ has an ergodic invariant probability measure on $\mathbb{R}_+^{\mu,\circ}$, it follows from Khasminskii (1960) that $\mathbf{X}$ is positive recurrent on $\mathbb{R}_+^{\mu,\circ}$. If $\lambda_i(\mu) \neq 0$ for some $i \in I_\mu$, then with probability 1 we have that as $t \to \infty$ either $X_i(t) \to 0$ [if $\lambda_i(\mu) < 0$] or $X_i(t) \to \infty$ [if $\lambda_i(\mu) > 0$]. This contradicts the fact that $\mathbf{X}$ is positive recurrent on $\mathbb{R}_+^{\mu,\circ}$. As a result, $\lambda_i(\mu) = 0$ for $i \in I_\mu$. $\square$

Condition (1.7) is equivalent to the existence of $0 < \widehat{p}_i < \delta_0, i \in I_\mu$ such that for any $\nu \in \mathcal{M}_\mu$, we have

$$\sum_{i \in I_\mu} \widehat{p}_i \lambda_i(\nu) > 0.$$

Thus, there is $\check{p} \in (0, \delta_0)$ sufficiently small such that

$$\sum_{i \in I_\mu} \widehat{p}_i \lambda_i(\nu) - \check{p} \max_{i \in I_\mu^c}\{\lambda_i(\nu)\} > 0 \qquad \text{for any } \nu \in \mathcal{M}_\mu. \tag{5.1}$$

In view of (5.1), (1.6) and Lemma 5.1, there is $\rho_e > 0$ such that for any $\nu \in \mathcal{M}_\mu \cup \{\mu\}$,

$$\sum_{i \in I_\mu} \widehat{p}_i \lambda_i(\nu) - \check{p} \max_{i \in I_\mu^c}\{\lambda_i(\nu)\} > 3\rho_e. \tag{5.2}$$

LEMMA 5.2. *Suppose that Assumption 1.3 holds. Let M be as in (3.1), H as in (3.5) and $\widehat{p}_i$, $\check{p}$, ρ_e as in (5.2). Let $n_e \in \mathbb{N}$ such that $\gamma_b(n_e - 1) > H$. There are $T_e \geq 0$, $\delta_e > 0$ such that, for any $T \in [T_e, n_e T_e]$, $\|\mathbf{x}\| \leq M$, $x_i < \delta_e$, $i \in I_\mu^c$, we have*

$$
\begin{aligned}
&\frac{1}{T}\int_0^T \mathbb{E}_{\mathbf{x}}\biggl(\frac{\sum_i c_i X_i(t) f_i(\mathbf{X}(t))}{1+\mathbf{c}^\top \mathbf{X}(t)} - \frac{\sum_{i,j} c_i c_j X_i(t) X_j(t) g_i(\mathbf{X}(t)) g_j(\mathbf{X}(t))\sigma_{ij}}{2(1+\mathbf{c}^\top \mathbf{X}(t))^2}\biggr)\,dt \\
&\quad - \sum_{i\in I_\mu} \widehat{p}_i \frac{1}{T}\int_0^T \mathbb{E}_{\mathbf{x}}\biggl(f_i(\mathbf{X}(t)) - \frac{\sigma_{ii} g_i^2(\mathbf{X}(t))}{2}\biggr)\,dt \\
&\quad + \check{p} \max_{i\in I_\mu^c}\biggl\{\frac{1}{T}\int_0^T \mathbb{E}_{\mathbf{x}}\biggl(f_i(\mathbf{X}(t)) - \frac{\sigma_{ii} g_i^2(\mathbf{X}(t))}{2}\biggr)\,dt\biggr\} \\
&\quad \leq -\rho_e.
\end{aligned} \tag{5.3}
$$

PROOF. Analogous to Lemma 4.1, using (5.2), one can show there exists a $T_e > 0$ such that for any $T > T_e$, $\mathbf{x} \in \mathbb{R}_+^\mu$, $\|\mathbf{x}\| \leq M$, we have

$$
\begin{aligned}
&\frac{1}{T}\int_0^T \mathbb{E}_{\mathbf{x}}\biggl(\frac{\sum_i c_i X_i(t) f_i(\mathbf{X}(t))}{1+\mathbf{c}^\top \mathbf{X}(t)} - \frac{\sum_{i,j} c_i c_j X_i(t) X_j(t) g_i(\mathbf{X}(t)) g_j(\mathbf{X}(t))\sigma_{ij}}{2(1+\mathbf{c}^\top \mathbf{X}(t))^2}\biggr)\,dt \\
&\quad - \sum_{i\in I_\mu} \widehat{p}_i \frac{1}{T}\int_0^T \mathbb{E}_{\mathbf{x}}\biggl(f_i(\mathbf{X}(t)) - \frac{\sigma_{ii} g_i^2(\mathbf{X}(t))}{2}\biggr)\,dt \\
&\quad + \check{p} \max_{i\in I_\mu^c}\biggl\{\frac{1}{T}\int_0^T \mathbb{E}_{\mathbf{x}}\biggl(f_i(\mathbf{X}(t)) - \frac{\sigma_{ii} g_i^2(\mathbf{X}(t))}{2}\biggr)\,dt\biggr\} \\
&\quad \leq -2\rho_e.
\end{aligned} \tag{5.4}
$$

By the Feller property of $(\mathbf{X}(t))$ and compactness of the set $\{\mathbf{x} \in \mathbb{R}_+^\mu, \|\mathbf{x}\| \leq M\}$, there is a $\delta_e > 0$ such that for any $T \in [T_e, n_e T_e]$, $\|\mathbf{x}\| \leq M$, $x_i < \delta_e$, $i \in I_\mu^c$, the estimate (5.3) holds. □

PROPOSITION 5.1. *Suppose that Assumption 1.3 holds. Let $\delta_0 > 0$ be as in (3.2). There is a $\theta \in (0, \delta_0)$ such that for any $T \in [T_e, n_e T_e]$ and $\mathbf{x} \in \mathbb{R}_+^{n,\circ}$ satisfying $\|\mathbf{x}\| \leq M$, $x_i < \delta_e$, $i \in I_\mu^c$ one has*

$$
\mathbb{E}_{\mathbf{x}} U_\theta\bigl(\mathbf{X}(T)\bigr) \leq \exp\biggl(-\frac{1}{2}\theta\rho_e T\biggr) U_\theta(\mathbf{x}),
$$

where M, T_e, $\widehat{p}_i$, $\check{p}$, δ_e, n_e are as in Lemma 5.2 and

$$
U_\theta(\mathbf{x}) = \sum_{i\in I_\mu^c}\biggl[(1+\mathbf{c}^\top \mathbf{x})\frac{x_i^{\check{p}}}{\prod_{j\in I_\mu} x_j^{\widehat{p}_j}}\biggr]^\theta, \qquad \mathbf{x} \in \mathbb{R}_+^{n,\circ}.
$$

PROOF. For $i \in I_\mu^c$, let $U_i(\mathbf{x}) = (1+\mathbf{c}^\top \mathbf{x})\frac{x_i^{\check{p}}}{\prod_{j\in I_\mu} x_j^{\widehat{p}_j}}$. Similar to Proposition 4.1, by making use of Lemma 5.2, one can find a $\theta > 0$ such that for $T \in [T_e, n_e T_e]$, $\mathbf{x} \in \mathbb{R}_+^{n,\circ}$ with $\|\mathbf{x}\| \le M$, and $x_i < \delta_e$ we have

$$\mathbb{E} U_i^\theta\big(\mathbf{X}(T)\big) \le \exp\biggl(-\frac{1}{2}\theta\rho_e T\biggr) U_i^\theta(\mathbf{x}).$$

The proof is complete by noting that

$$U_\theta(\mathbf{x}) = \sum_{i\in I_\mu^c} U_i^\theta(\mathbf{x}).$$

□

LEMMA 5.3. *Let H be defined by* (3.5). *For $\theta \in [0, \delta_0]$, we have*

$$\mathbb{E}_{\mathbf{x}} U_\theta\big(\mathbf{X}(t)\big) \le \exp(\theta H t) U_\theta(\mathbf{x}), \qquad \mathbf{x} \in \mathbb{R}_+^{n,\circ}.$$

PROOF. By the arguments from the proof of (3.6), for $\theta \le \delta_0$, $i \in I_\mu^c$ we have

$$\mathbb{E}_{\mathbf{x}} U_i^\theta\big(\mathbf{X}(t)\big) \le \exp(\theta H t) U_i^\theta(\mathbf{x}), \qquad \mathbf{x} \in \mathbb{R}_+^{n,\circ}.$$

From this estimate, we can take the sum over I_μ^c to obtain the desired result. □

REMARK 5.1. It is key to note that the inequalities (A.3) and (A.4) hold if $|p_i| < \delta_0$ no matter if the p_i's are negative or positive. This then allows us to have the same kind of estimates for U_θ and V_θ.

THEOREM 5.1. *Under Assumptions 1.1 and 1.3, for any $\delta < \delta_0$ and any $\mathbf{x} \in \mathbb{R}_+^{n,\circ}$ we have*

$$\lim_{t\to\infty} \mathbb{E}_{\mathbf{x}} \bigwedge_{i=1}^{n} X_i^\delta(t) = 0, \tag{5.5}$$

where $\bigwedge_{i=1}^n a_i = \min_{i=1,\dots,n}\{a_i\}$.

PROOF. Just as in (4.20), we have

$$\mathcal{L} U_\theta(\mathbf{x}) \le -\theta\gamma_b U_\theta(\mathbf{x}) \qquad \text{if } \|x\| \ge M. \tag{5.6}$$

Let

$$\varsigma := \frac{\delta_e^{\check{p}\theta}}{C_U^\theta},$$

$$C_U := \sup\biggl\{\frac{\prod_{i\in I_\mu} x_i^{\widehat{p}_i}}{1+\mathbf{c}^\top\mathbf{x}} : \mathbf{x} \in \mathbb{R}_+^{n,\circ}\biggr\} < \infty,$$

and

$$\xi := \inf\{t \geq 0 : U^\theta(\mathbf{X}(t)) \geq \varsigma\}.$$

Clearly, if $U_\theta(\mathbf{x}) < \varsigma$, then $\xi > 0$ and for any $i \in I_\mu^c$, we get

$$X_i(t) \leq \delta_e, \qquad t \in [0, \xi). \tag{5.7}$$

Let

$$\tilde{U}_\theta(\mathbf{x}) := \varsigma \wedge U_\theta(\mathbf{x}).$$

We have from the concavity of $x \mapsto x \wedge \varsigma$ that

$$\mathbb{E}_\mathbf{x}\tilde{U}_\theta(\mathbf{X}(T)) \leq \varsigma \wedge \mathbb{E}U_\theta(\mathbf{X}(T)).$$

Let τ be defined by (4.21). By (5.6) and Dynkin's formula, we have that

$$\begin{aligned}
&\mathbb{E}_\mathbf{x}[\exp(\theta\gamma_b(\tau \wedge \xi \wedge n_e T_e))U_\theta(\mathbf{X}(\theta\gamma_b(\tau \wedge \xi \wedge n_e T_e)))] \\
&\quad \leq U_\theta(\mathbf{x}) + \mathbb{E}_\mathbf{x}\int_0^{\theta\gamma_b(\tau\wedge\xi\wedge n_e T_e)} \exp(\theta\gamma_b s)[\mathcal{L}U_\theta(\mathbf{X}(s)) + \theta\gamma_b U_\theta(\mathbf{X}(s))]\,ds \\
&\quad \leq U_\theta(\mathbf{x}).
\end{aligned}$$

As a result,

$$\begin{aligned}
U_\theta(x) &\geq \mathbb{E}_\mathbf{x}[\exp(\theta\gamma_b(\tau \wedge \xi \wedge n_e T_e))U_\theta(\mathbf{X}(\tau \wedge \xi \wedge n_e T_e))] \\
&\geq \mathbb{E}_\mathbf{x}[\mathbf{1}_{\{\tau\wedge\xi\wedge(n_e-1)T_e=\tau\}}U_\theta(\mathbf{X}(\tau))] \\
&\quad + \mathbb{E}_\mathbf{x}[\mathbf{1}_{\{\tau\wedge\xi\wedge(n_e-1)T_e=\xi\}}U_\theta(\mathbf{X}(\xi))] \\
&\quad + \exp(\theta\gamma_b(n_e-1)T_e)\mathbb{E}_\mathbf{x}[\mathbf{1}_{\{(n_e-1)T_e<\tau\wedge\xi<n_e T_e\}}U_\theta(\mathbf{X}(\tau\wedge\xi))] \\
&\quad + \exp(\theta\gamma_b n_e T_e)\mathbb{E}_\mathbf{x}[\mathbf{1}_{\{\tau\wedge\xi\geq n_e T_e\}}U_\theta(\mathbf{X}(n_e T_e))].
\end{aligned} \tag{5.8}$$

By the strong Markov property of $(\mathbf{X}(t))$ and Proposition 5.1 [which we can use because of (5.7)],

$$\begin{aligned}
&\mathbb{E}_\mathbf{x}[\mathbf{1}_{\{\tau\wedge\xi\wedge(n_e-1)T_e=\tau\}}U_\theta(\mathbf{X}(n_e T_e))] \\
&\quad \leq \mathbb{E}_\mathbf{x}\left[\mathbf{1}_{\{\tau\wedge\xi\wedge(n_e-1)T_e=\tau\}}\exp\left(-\frac{1}{2}\theta\rho_e(n_e T_e - \tau)\right)U_\theta(\mathbf{X}(\tau))\right] \\
&\quad \leq \mathbb{E}_\mathbf{x}[\mathbf{1}_{\{\tau\wedge\xi\wedge(n_e-1)T_e=\tau\}}U_\theta(\mathbf{X}(\tau))].
\end{aligned} \tag{5.9}$$

Similarly, by the strong Markov property of $(\mathbf{X}(t))$ and Lemma 5.3, we obtain

$$\begin{aligned}
&\mathbb{E}_\mathbf{x}[\mathbf{1}_{\{(n_e-1)T_e<\tau\wedge\xi<n_e T_e\}}U_\theta(\mathbf{X}(n_e T_e))] \\
&\quad \leq \mathbb{E}_\mathbf{x}[\mathbf{1}_{\{(n_e-1)T_e<\tau\wedge\xi<n_e T_e\}}\exp(\theta H(n_e T_e - \tau\wedge\xi))U_\theta(\mathbf{X}(\tau\wedge\xi))] \\
&\quad \leq \exp(\theta H T_e)\mathbb{E}_\mathbf{x}[\mathbf{1}_{\{(n_e-1)T_e<\tau\wedge\xi<n_e T_e\}}U_\theta(\mathbf{X}(\tau\wedge\xi))].
\end{aligned} \tag{5.10}$$

If $U_\theta(\mathbf{x}) < \varsigma$, then applying (5.9), (5.10) and the inequality $\tilde{U}_\theta(\mathbf{X}(n_eT_e)) \le U_\theta(\mathbf{X}(n_eT_e \wedge \xi))$ to (5.8) yields

$$
\begin{aligned}
\tilde{U}_\theta(\mathbf{x}) &= U_\theta(\mathbf{x}) \\
&\ge \mathbb{E}_{\mathbf{x}}\big[\mathbf{1}_{\{\tau\wedge\xi\wedge(n_e-1)T_e=\tau\}}U_\theta\big(\mathbf{X}(\tau)\big)\big] \\
&\quad + \mathbb{E}_{\mathbf{x}}\big[\mathbf{1}_{\{\tau\wedge\xi\wedge(n_e-1)T_e=\xi\}}U_\theta\big(\mathbf{X}(\xi)\big)\big] \\
&\quad + \exp\big(\theta\gamma_b(n_e-1)T_e\big)\mathbb{E}_{\mathbf{x}}\big[\mathbf{1}_{\{(n_e-1)T_e<\tau\wedge\xi<n_eT\}}U_\theta\big(\mathbf{X}(\tau\wedge\xi)\big)\big] \\
&\quad + \exp(\theta\gamma_b n_eT_e)\mathbb{E}_{\mathbf{x}}\big[\mathbf{1}_{\{\tau\wedge\xi\ge n_eT_e\}}U_\theta\big(\mathbf{X}(n_eT_e)\big)\big] \\
&\ge \mathbb{E}_{\mathbf{x}}\big[\mathbf{1}_{\{\tau\wedge\xi\wedge(n_e-1)T_e=\tau\}}U_\theta\big(\mathbf{X}(n_eT_e)\big)\big] \\
&\quad + \mathbb{E}_{\mathbf{x}}\big[\mathbf{1}_{\{\tau\wedge\xi\wedge(n_e-1)T_e=\xi\}}\tilde{U}_\theta\big(\mathbf{X}(n_eT_e)\big)\big] \\
&\quad + \exp\big(\theta\gamma_b(n_e-1)T_e-\theta HT_e\big) \\
&\quad \times \mathbb{E}_{\mathbf{x}}\big[\mathbf{1}_{\{(n_e-1)T_e<\tau\wedge\xi<n_eT_e\}}U_\theta\big(\mathbf{X}(n_eT_e)\big)\big] \\
&\quad + \exp(\theta\gamma_b n_eT_e)\mathbb{E}_{\mathbf{x}}\big[\mathbf{1}_{\{\tau\wedge\xi\ge n_eT_e\}}U_\theta\big(\mathbf{X}(n_eT_e)\big)\big] \\
&\ge \mathbb{E}_{\mathbf{x}}\tilde{U}_\theta\big(\mathbf{X}(n_eT_e)\big) \qquad \big(\text{since } \tilde{U}_\theta(\cdot)\le U_\theta(\cdot)\big).
\end{aligned} \tag{5.11}
$$

Clearly, if $U_\theta(\mathbf{x}) \ge \varsigma$ then

$$
\mathbb{E}_{\mathbf{x}}\tilde{U}_\theta\big(\mathbf{X}(n_eT_e)\big) \le \varsigma = \tilde{U}_\theta(\mathbf{x}). \tag{5.12}
$$

As a result of (5.11), (5.12) and the Markov property of $(\mathbf{X}(t))$, the sequence $\{Y(k) : k \in \mathbb{N}\}$ where $Y(k) := \tilde{U}_\theta(\mathbf{X}(kn_eT_e))$ is a supermartingale. For $\lambda \in (0, \varsigma)$, $\varepsilon \in (0, 1)$, let $\zeta_\lambda := \inf\{k \in \mathbb{N} : Y(k) \ge \lambda\}$. If $U_\theta(\mathbf{x}) \le \lambda\varepsilon$, we have

$$
\mathbb{E}_{\mathbf{x}}Y(k\wedge\zeta_\lambda) \le \mathbb{E}_{\mathbf{x}}Y(0) = U_\theta(\mathbf{x}) \le \lambda\varepsilon \qquad \text{for all } k\in\mathbb{N}. \tag{5.13}
$$

Subsequently, (5.13) combined with the Markov inequality yields

$$
\mathbb{P}_{\mathbf{x}}\{\zeta_\lambda < k\} \le \lambda^{-1}\mathbb{E}_{\mathbf{x}}Y(k\wedge\zeta_\lambda) \le \varepsilon, \qquad k\in\mathbb{N}, U_\theta(\mathbf{x}) \le \lambda\varepsilon.
$$

Next, let $k \to \infty$ to get

$$
\mathbb{P}_{\mathbf{x}}\{\zeta_\lambda < \infty\} \le \varepsilon \qquad \text{if } U_\theta(\mathbf{x}) \le \lambda\varepsilon. \tag{5.14}
$$

Note that for a given compact set $\mathcal{K} \subset \mathbb{R}_+^{n,\circ}$ with nonempty interior, and for any $\varepsilon > 0$ there exists a $\lambda > 0$ such that

$$
\mathbb{P}_{\mathbf{x}}\big\{X_i(t) \ge \lambda \text{ for all } t\in[0, n_eT_e], i = 1,\dots,n\big\} > 1-\varepsilon, \qquad \mathbf{x}\in\mathcal{K}. \tag{5.15}
$$

This standard fact can be shown in the same manner as (5.20), which is proved later in Lemma 5.5.

We show by contradiction that $(\mathbf{X}(t))$ is transient. If the process $(\mathbf{X}(t))$ is recurrent in $\mathbb{R}_+^{n,\circ}$, then $\mathbf{X}(t)$ will enter $\mathcal{K}$ in a finite time almost surely given that $\mathbf{X}(0) \in \mathbb{R}_+^{n,\circ}$. By the strong Markov property and (5.15), we have

$$
\mathbb{P}_{\mathbf{x}}\big\{X_i(kn_eT_e) \ge \lambda, i = 1,\dots,n \text{ for some } k\in\mathbb{N}\big\} > 1-\varepsilon, \qquad \mathbf{x}\in\mathbb{R}_+^{n,\circ}. \tag{5.16}
$$

If $\mathbf{x} \in \mathbb{R}_+^{n,\circ}$ is such that $U_\theta(\mathbf{x})$ is sufficiently small, then both (5.14) and (5.16) hold, a contradiction. Thus, $\mathbf{X}$ is transient.

As a result, any weak*limit of $P_{\mathbf{X}}(t, \mathbf{x}, \cdot)$ is a probability measure concentrated on $\partial\mathbb{R}_+^n$. Suppose that (5.5) does not hold. Then there exist $\mathbf{x}_0 \in \mathbb{R}_+^{n,\circ}$, $\widehat{\varepsilon} > 0$ and a sequence $t_k \uparrow \infty$ as $k \to \infty$ such that

$$\int_{\mathbb{R}_+^n} \left(\bigwedge_{i=1}^n x_i^\delta \right) P_{\mathbf{X}}(t_k, \mathbf{x}_0, d\mathbf{x}) > \widehat{\varepsilon} \qquad \text{for any } k \in \mathbb{N}. \tag{5.17}$$

Since the family $\{P_{\mathbf{X}}(t, \mathbf{x}_0, \cdot), t \geq 0\}$ is tight in $\mathbb{R}_+^n$ due to Lemma 3.2, there exists a subsequence of $\{t_k\}$ (still denoted by $\{t_k\}$ for convenience) such that $P_{\mathbf{X}}(t_k, \mathbf{x}_0, \cdot)$ converges weakly to π as $k \to \infty$. Similar computations to the ones from Lemma 3.4 show that and $h(\cdot)$ is a continuous function on $\mathbb{R}_+^n$ such that for all $\mathbf{x} \in \mathbb{R}_+^n$ we have $|h(\mathbf{x})| < K(1 + \|\mathbf{x}\|)^\delta$, $\delta < \delta_0$ then $h(\cdot)$ is π-integrable and $\int_{\mathbb{R}_+^n} h(\mathbf{x}) P_{\mathbf{X}}(t_k, \mathbf{x}_0, d\mathbf{x}) \to \int_{\mathbb{R}_+^n} h(\mathbf{x}) \pi(d\mathbf{x})$. Because π with $\text{supp}(\pi) \subset \partial\mathbb{R}_+^n$, we have

$$\int_{\mathbb{R}_+^n} \left(\bigwedge_{i=1}^n x_i^\delta \right) \pi(d\mathbf{x}) = 0$$

and

$$\left(\bigwedge_{i=1}^n x_i^\delta \right) \leq K\left(1 + \|\mathbf{x}\|\right)^\delta.$$

These facts imply

$$\lim_{k\to\infty} \int_{\mathbb{R}_+^n} \left(\bigwedge_{i=1}^n x_i^\delta \right) P_{\mathbf{X}}(t_k, \mathbf{x}_0, d\mathbf{x}) = 0,$$

which contradicts (5.17). As a result, (5.5) has to hold. □

To prove Theorem 1.3, we need the following lemmas.

LEMMA 5.4. *Suppose that Assumption 1.4 is satisfied. Then there is a $\widehat{K} > 0$ such that*

$$\mathbb{P}_{\mathbf{x}}\left\{ \limsup_{t\to\infty} \frac{1}{t} \int_0^t \left(1 + \mathbf{c}^\top \mathbf{X}(s)\right)^{\delta_1} \times \left(1 + \sum_i \left(\left|f_i\left(\mathbf{X}(s)\right)\right| + \left|g_i\left(\mathbf{X}(s)\right)\right|^2\right)\right) ds \leq \widehat{K} \right\} = 1, \qquad \mathbf{x} \in \mathbb{R}_+^n.$$

Moreover,

$$\mathbb{P}_{\mathbf{x}}\left\{ \lim_{t\to\infty} \frac{1}{t} \int_0^t \frac{\sum_i c_i X_i(s) g_i(\mathbf{X}(s))}{1 + \mathbf{c}^\top \mathbf{X}(s)} \, dE_i(s) = 0 \right\} = 1, \qquad \mathbf{x} \in \mathbb{R}_+^n. \tag{5.18}$$

LEMMA 5.5. *Let Assumption 1.4 be satisfied. There is a* $\widehat{K}_1 > 1$ *such that*

$$\mathbb{P}_{\mathbf{x}}\left\{\liminf_{t\to\infty}\frac{1}{t}\int_0^t \mathbf{1}_{\{\|\mathbf{X}(s)\|\le\widehat{K}_1\}}\,ds \ge \frac{1}{2}\right\} = 1, \qquad \mathbf{x}\in\mathbb{R}^n_+. \tag{5.19}$$

Moreover, for any $\varepsilon_1, \varepsilon_2 > 0$, *there is a* $\beta > 0$ *such that for each* $i = 1, \dots, n$,

$$\mathbb{P}_{\mathbf{x}}\{X_i(t) > \beta, \forall t \in [0, n_e T_e]\} > 1 - \varepsilon_1 \qquad \text{if } \mathbf{x}\in\mathbb{R}^n_+, \|\mathbf{x}\| \le \widehat{K}_1, x_i > \varepsilon_2, \tag{5.20}$$

where n_e, T_e *are as in Lemma 5.2.*

LEMMA 5.6. *Let Assumption 1.4 be satisfied and let* $\widetilde{\Pi}_t(\cdot)$ *be the random normalized occupation measure defined in Remark 1.7. Suppose we have a sample path of* $\mathbf{X}$ *satisfying*

$$\limsup_{t\to\infty}\frac{1}{t}\int_0^t (1+\mathbf{c}^\top\mathbf{X}(s))^{\delta_1}\left(1+\sum_i(|f_i(\mathbf{X}(s))|+|g_i(\mathbf{X}(s))|^2)\right)ds \le \widehat{K}$$

and that there is a sequence $(T_k)_{k\in\mathbb{N}} \subset \mathbb{R}^n_+$ *such that* $\lim_{k\to\infty} T_k = \infty$ *and* $(\widetilde{\Pi}_{T_k}(\cdot))_{k\in\mathbb{N}}$ *converges weakly to an invariant probability measure* π *of* $\mathbf{X}$ *when* $k\to\infty$. *Then for this sample path, we have* $\int_{\mathbb{R}^n_+} h(\mathbf{x})\widetilde{\Pi}_{T_k}(d\mathbf{x}) \to \int_{\mathbb{R}^n_+} h(\mathbf{x})\pi(d\mathbf{x})$ *for any continuous function* $h : \mathbb{R}^n_+ \to \mathbb{R}$ *satisfying* $|h(\mathbf{x})| < K_h(1+\mathbf{c}^\top\mathbf{x})^\delta \times (1+\sum_i(|f_i(\mathbf{x})|+|g_i(\mathbf{x})|^2))$, $\mathbf{x}\in\mathbb{R}^n_+$, *with* K_h *a positive constant and* $\delta\in[0,\delta_1)$.

The proofs of Lemmas 5.4 and 5.5 are given in the Appendix while that of Lemma 5.6 is almost the same as that of Lemma 3.4 and is left for the reader.

LEMMA 5.7. *Let Assumption 1.4 be satisfied. For any initial condition* $\mathbf{X}(0) = \mathbf{x}\in\mathbb{R}^n_+$, *the family* $\{\widetilde{\Pi}_t(\cdot), t\ge 1\}$ *is tight in* $\mathbb{R}^n_+$, *and its weak**-*limit set, denoted by* $\mathcal{U} = \mathcal{U}(\omega)$ *is a family of invariant probability measures of* $\mathbf{X}$ *with probability* 1.

PROOF. The tightness follows from Lemma 5.4. The property of the weak* limit set of normalized occupation measures was first proved in Schreiber, Benaïm and Atchadé [(2011), Theorems 4, 5] for compact state spaces and then generalized to a locally compact state space in Evans, Hening and Schreiber (2015), Theorem 4.2. Similar results for general Markov processes can be found in Benaïm (2014). □

LEMMA 5.8. *Suppose that Assumption 1.4 is satisfied. Let* $\mu\in\mathcal{M}^1$. *For any* $\mathbf{x}\in\mathbb{R}^{n,\circ}_+$,

$$\mathbb{P}_{\mathbf{x}}\{\mathcal{U}(\omega)\subset \mathrm{Conv}(\mathcal{M}_\mu\cup\{\mu\})\} = \mathbb{P}_{\mathbf{x}}\{\mathcal{U}(\omega) = \{\mu\}\}.$$

REMARK 5.2. Note that, since $\{\mathcal{U}(\omega) = \{\mu\}\} \subset \{\mathcal{U}(\omega)\subset\mathrm{Conv}(\mathcal{M}_\mu\cup\{\mu\})\}$, it would be equivalent to prove that $\{\mathcal{U}(\omega)\subset\mathrm{Conv}(\mathcal{M}_\mu\cup\{\mu\})\} = \{\mathcal{U}(\omega) = \{\mu\}\}$ $\mathbb{P}_{\mathbf{x}}$-a.s. for all $\mathbf{x}\in\mathbb{R}^{n,\circ}_+$.

PROOF OF LEMMA 5.8. Since μ satisfies Assumption 1.3, it follows from (1.7) that, there are $p_i^\mu > 0, i \in I_\mu$ such that

$$\sum_{i \in I_\mu} p_i^\mu \lambda_i(\nu) > 0, \qquad \nu \in \mathcal{M}_\mu. \tag{5.21}$$

As a result of Lemmas 3.3, 5.6 and 5.7,

$$\begin{aligned}\mathbb{P}_{\mathbf{x}}\biggl\{\lim_{t\to\infty}\frac{1}{t}\int_0^t\biggl[&\frac{\sum_i c_i X_i(s) f_i(\mathbf{X}(s))}{1+\mathbf{c}^\top\mathbf{X}(s)}\\ &-\frac{1}{2}\frac{\sum_{i,j} c_i c_j X_i(s) X_j(s) g_i(\mathbf{X}(s)) g_j(\mathbf{X}(s))}{(1+\mathbf{c}^\top\mathbf{X}(s))^2}\biggr]ds = 0\biggr\} = 1.\end{aligned} \tag{5.22}$$

In light of Itô's formula, it follows from (5.18) and (5.22) that

$$\begin{aligned}\mathbb{P}_{\mathbf{x}}\biggl\{\limsup_{t\to\infty}\frac{\ln X_i(t)}{t} \le 0, i = 1,\dots,n\biggr\} &\ge \mathbb{P}_{\mathbf{x}}\biggl\{\limsup_{t\to\infty}\frac{\ln(1+\mathbf{c}^\top\mathbf{X}(t))}{t} = 0\biggr\}\\ &= 1, \qquad \mathbf{x} \in \mathbb{R}_+^{n,\circ}.\end{aligned} \tag{5.23}$$

On the other hand, similar to (5.18), we have

$$\mathbb{P}_{\mathbf{x}}\biggl\{\lim_{t\to\infty}\frac{1}{t}\int_0^t g_i\bigl(\mathbf{X}(s)\bigr)\,dE_i(s) = 0, i = 1,\dots,n\biggr\} = 1. \tag{5.24}$$

In view of (5.24) and (5.23), to prove the lemma, it suffices to show that if the following properties:

(a) $\mathcal{U}(\omega) \subset \operatorname{Conv}(\mathcal{M}_\mu \cup \{\mu\})$;

(b)

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t g_i\bigl(\mathbf{X}(s)\bigr)\,dE_i(s) = 0, \qquad i = 1,\dots,n; \tag{5.25}$$

(c)

$$\limsup_{t\to\infty}\frac{\ln X_i(t)}{t} \le 0, \qquad i = 1,\dots,n \tag{5.26}$$

hold then $\mathcal{U}(\omega) = \{\mu\}$.

We argue by contradiction. Assume there is a sequence $\{t_k\}$ with $\lim_{k\to\infty} t_k = \infty$ such that $\widetilde{\Pi}_{t_k}(\cdot)$ converges weakly to an invariant probability of the form $\pi = (1-\rho)\pi_1 + \rho\mu$ where $\rho \in (0,1]$ and $\pi_1 \in \operatorname{Conv}(\mathcal{M}_\mu)$. It follows from Lemma 5.4

and (5.21) that

$$
\begin{aligned}
&\lim_{k\to\infty}\frac{1}{t_k}\sum_{i\in I_\mu}p_i^\mu\int_0^{t_k}\Big(f_i(\mathbf{X}(s))-\frac{\sigma_{ii}g_i^2(\mathbf{X}(s))}{2}\Big)\,ds\\
&\quad=\sum_{i\in I_\mu}p_i^\mu\lambda_i(\pi)\\
&\quad=(1-\rho)\sum_{i\in I_\mu}p_i^\mu\lambda_i(\pi_1)+\rho\sum_{i\in I_\mu}p_i^\mu\lambda_i(\mu)\\
&\quad=(1-\rho)\sum_{i\in I_\mu}p_i^\mu\lambda_i(\pi_1)\qquad\text{(due to Lemma 5.1)}\\
&\quad>0.
\end{aligned}
\tag{5.27}
$$

As a result of (5.25), (5.27) and Itô's formula,

$$
\begin{aligned}
&\lim_{k\to\infty}\sum_{i\in I_\mu}p_i^\mu\frac{\ln X_i(t_k)}{t_k}\\
&\quad=\lim_{k\to\infty}\frac{1}{t_k}\sum_{i\in I_\mu}p_i^\mu\int_0^{t_k}\Big[\Big(f_i(\mathbf{X}(s))-\frac{\sigma_{ii}g_i^2(\mathbf{X}(s))}{2}\Big)\,ds+g_i(\mathbf{X}(s))\,dE_i(s)\Big]\\
&\quad>0
\end{aligned}
$$

which contradicts (5.26). This completes the proof. □

LEMMA 5.9. *Suppose that Assumption 1.4 is satisfied. Let $\mu\in\mathcal{M}^1$. For any $k\in\mathbb{N}$, $\varepsilon>0$, there is a $\Delta>0$ such that*

$$
\mathbb{P}_{\mathbf{x}}\Big\{\mathcal{U}(\omega)=\{\mu\}\text{ and }\lim_{t\to\infty}\frac{\ln X_i(t)}{t}=\lambda_i(\mu)<0,i\in I_\mu^c\Big\}>1-\varepsilon,\qquad \mathbf{x}\in\mathcal{K}_\mu^{k,\Delta},
$$

where

$$
\mathcal{K}_\mu^{k,\Delta}:=\{\mathbf{x}\in\mathbb{R}_+^{n,\circ},k^{-1}\le x_i\le k\text{ for }i\in I_\mu,x_i<\Delta\text{ for }i\in I_\mu^c\}.
$$

PROOF. Let $\tilde U(\mathbf{x})$ be the function defined as in the proof of Theorem 5.1. In view of Lemma 5.5, there is $\beta>0$ such that

$$
\mathbb{P}_{\mathbf{x}}\Big\{\max_{i\in I_\mu^c}\{X_i(t)\}>\beta,\forall t\in[0,n_eT_e]\Big\}>\frac{1}{2},\qquad \mathbf{x}\in\mathcal{H}_\mu,
\tag{5.28}
$$

where

$$
\mathcal{H}_\mu=\Big\{\mathbf{x}\in\mathbb{R}_+^n:\|\mathbf{x}\|\le\widehat{K}_1,\max_{i\in I_\mu^c}\{x_i\}\ge 1\Big\}.
$$

It can be seen that

$$\nu(\mathcal{H}_\mu) > 0 \qquad \text{for } \nu \in \mathcal{M} \setminus \big(\mathcal{M}_\mu \cup \{\mu\}\big). \tag{5.29}$$

By the definition of $\tilde{U}(\cdot)$, there is $\Delta > 0$ sufficiently small such that

$$\sup_{\mathbf{x}\in\mathcal{K}_\mu^{k,\Delta}} \{\tilde{U}(\mathbf{x})\} \le \frac{\varepsilon}{2} \inf_{\mathbf{y}\in\mathbb{R}_+^{n,\circ}, x_i\ge\beta, i\in I_\mu^c} \{\tilde{U}(\mathbf{y})\}. \tag{5.30}$$

In view of (5.30), since $\{Y(j) := \tilde{U}_\theta(\mathbf{X}(jn_eT_e)) : j \in \mathbb{N}\}$ is a supermartingale, similar to (5.14), we can obtain

$$\mathbb{P}_\mathbf{x}\Big\{\max_{i\in I_\mu^c}\{X_i(jn_eT_e)\} < \beta \text{ for all } j \in \mathbb{N}\Big\} > 1 - \frac{\varepsilon}{2} \qquad \text{if } \mathbf{x} \in \mathcal{K}_\mu^{k,\Delta}. \tag{5.31}$$

Now, suppose that there is $\mathbf{x} \in \mathcal{K}_\mu^{k,\Delta}$ such that

$$\mathbb{P}_\mathbf{x}\Big\{\limsup_{t\to\infty} \frac{1}{t}\int_0^t \mathbf{1}_{\{\mathbf{X}(s)\in\mathcal{H}_\mu\}}\,ds > 0\Big\} > \varepsilon. \tag{5.32}$$

Then

$$\mathbb{P}_\mathbf{x}\{\tau_{\mathcal{H}_\mu} < \infty\} > \varepsilon, \tag{5.33}$$

where $\tau_{\mathcal{H}_\mu} = \inf\{t > 0 : \mathbf{X}(t) \in \mathcal{H}_\mu\}$.

By the strong Markov property of $\{\mathbf{X}(t) : t \in \mathbb{R}_+\}$, it follows from (5.28) and (5.33) that

$$\mathbb{P}_\mathbf{x}\Big(\{\tau_{\mathcal{H}_\mu} < \infty\} \cap \Big\{\max_{i\in I_\mu^c}\{X_i(t)\} \ge \beta \text{ for } t \in [\tau_{\mathcal{H}_\mu}, \tau_{\mathcal{H}_\mu} + n_eT_e]\Big\}\Big) > \frac{1}{2}\varepsilon,$$

which contradicts (5.31). Thus, (5.32) does not hold, that is, we have

$$\mathbb{P}_\mathbf{x}\Big\{\lim_{t\to\infty} \frac{1}{t}\int_0^t \mathbf{1}_{\{\mathbf{X}(s)\in\mathcal{H}_\mu\}}\,ds = 0\Big\} > 1 - \varepsilon, \qquad \mathbf{x} \in \mathcal{K}_\mu^{k,\Delta}. \tag{5.34}$$

If for an $\omega \in \Omega$, and a sequence $\{t_j\}$ with $\lim_{j\to\infty} t_j = \infty$, $\widetilde{\Pi}_{t_j}(\cdot)$ converges weakly to an invariant probability of the form $\pi = (1 - \rho)\pi_1 + \rho\pi_2$ where $\rho \in (0, 1]$ and $\pi_1 \in \operatorname{Conv}(\mathcal{M}_\mu \cup \{\mu\})$, $\pi_2 \in \operatorname{Conv}(\mathcal{M} \setminus (\mathcal{M}_\mu \cup \{\mu\}))$ then by (5.29)

$$\limsup_{j\to\infty} \frac{1}{t_j}\int_0^{t_j} \mathbf{1}_{\{\mathbf{X}(s)\in\mathcal{H}_\mu\}}\,ds \ge \pi(\mathcal{H}_\mu) \ge \rho\pi_2(\mathcal{H}_\mu) > 0.$$

This inequality, combined with Lemma 5.7 and (5.34), implies that

$$\mathbb{P}_\mathbf{x}\{\mathcal{U}(\omega) \subset \operatorname{Conv}\big(\mathcal{M}_\mu \cup \{\mu\}\big)\} > 1 - \varepsilon, \qquad \mathbf{x} \in \mathcal{K}_\mu^{k,\Delta}.$$

Lemma 5.8 and the above force

$$\mathbb{P}_\mathbf{x}\{\mathcal{U}(\omega) = \{\mu\}\} > 1 - \varepsilon, \qquad \mathbf{x} \in \mathcal{K}_\mu^{k,\Delta}. \tag{5.35}$$

In view of Lemma 5.6 and (5.35), we have for $\mathbf{x} \in \mathcal{K}_\mu^{k,\Delta}$ and for each $i = 1, \dots, n$ that

$$\mathbb{P}_\mathbf{x}\left\{\lim_{t\to\infty} \frac{1}{t}\int_0^t \left(f_i(\mathbf{X}(s)) - \frac{\sigma_{ii} g_i^2(\mathbf{X}(s))}{2}\right) ds = \lambda_i(\mu)\right\} > 1 - \varepsilon. \tag{5.36}$$

The claim of this lemma follows from (5.36), (5.24) and an application of Itô's formula. □

THEOREM 5.2. *Suppose that Assumptions 1.1, 1.4 and 1.5 are satisfied and $\mathcal{M}^1 \neq \varnothing$. Then for any $\mathbf{x} \in \mathbb{R}_+^{n,\circ}$*

$$\sum_{\mu\in\mathcal{M}^1} P_\mathbf{x}^\mu = 1, \tag{5.37}$$

where for $\mathbf{x} \in \mathbb{R}_+^{n,\circ}$, $\mu \in \mathcal{M}^1$

$$P_\mathbf{x}^\mu := \mathbb{P}_\mathbf{x}\left\{\mathcal{U}(\omega) = \{\mu\} \text{ and } \lim_{t\to\infty} \frac{\ln X_i(t)}{t} = \lambda_i(\mu) < 0, i \in I_\mu^c\right\} > 0.$$

PROOF. First, suppose that Assumption 1.5 is satisfied with nonempty $\mathcal{M}^2$. Then there is $\mathbf{q} = (q_1, \dots, q_n) \in \mathbb{R}_+^{n,\circ}$ such that $\|\mathbf{q}\| = 1$ and

$$\max_{\nu\in\mathcal{M}^2}\left\{\sum_i q_i\lambda_i(\nu)\right\} > 0. \tag{5.38}$$

Using (5.38) and arguing by contradiction, similar to the argument from Lemma 5.8, we can show that with probability 1, $\mathcal{U}(\omega)$ is a subset of $\text{Conv}(\mathcal{M}) \setminus \text{Conv}(\mathcal{M}^2)$. In other words, each invariant probability $\pi \in \mathcal{U}(\omega)$ has the form $\pi = (1-\rho)\pi_1 + \rho\pi_2$ where $\rho \in [0, 1), \pi_1 \in \text{Conv}(\mathcal{M}^1), \pi_2 \in \text{Conv}(\mathcal{M}^2)$. Let $k_0 > 1$ and for each $\mu \in \mathcal{M}^1$ define

$$\mathcal{K}_\mu^0 = \{\mathbf{x} \in \mathbb{R}_+^\mu : x_i \wedge x_i^{-1} \le k_0, i \in I_\mu\}.$$

By Lemma 5.9, there are $k > k_0$ and $\Delta > 0$ such that

$$\mathbb{P}_\mathbf{x}\left\{\lim_{t\to\infty} \frac{\ln X_i(t)}{t} = \lambda_i(\mu) < 0, i \in I_\mu^c\right\} > 1 - \varepsilon \tag{5.39}$$

for all $\mu \in \mathcal{M}^1$ and $\mathbf{x} \in \mathcal{K}_\mu^{k,\Delta}$. Let $\psi(\cdot) : \mathbb{R}_+^n \to [0, 1]$ be a continuous function satisfying

$$\psi(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x} \in \bigcup_{\mu\in\mathcal{M}^1} \mathcal{K}_\mu^0, \\ 0 & \text{if } \mathbf{x} \in \mathbb{R}_+^{n,\circ} \setminus \left(\bigcup_{\mu\in\mathcal{M}^1} \mathcal{K}_\mu^{k,\Delta}\right). \end{cases}$$

Since $\pi_1(\bigcup_{\mu\in\mathcal{M}^1}\mathcal{K}_\mu^0) > 0$ for any $\pi_1 \in \mathrm{Conv}(\mathcal{M}^1)$ and $\mathcal{U}(\omega)$ is a subset of $\mathrm{Conv}(\mathcal{M}) \setminus \mathrm{Conv}(\mathcal{M}^2)$ with probability 1, then we have from Lemma 5.7 that

$$\mathbb{P}_{\mathbf{x}}\left\{\liminf_{t\to\infty}\frac{1}{t}\int_0^t \psi\big(\mathbf{X}(s)\big)\,ds > 0\right\} = 1, \qquad \mathbf{x}\in\mathbb{R}_+^{n,\circ}. \tag{5.40}$$

Since $\psi(\mathbf{x}) = 0$ if $\mathbf{x}\in\mathbb{R}_+^{n,\circ}\setminus(\bigcup_{\mu\in\mathcal{M}^1}\mathcal{K}_\mu^{k,\Delta})$, we deduce from (5.40) that

$$\mathbb{P}_{\mathbf{x}}\left\{\liminf_{t\to\infty}\frac{1}{t}\int_0^t \mathbf{1}_{\{\mathbf{X}(s)\in\bigcup_{\mu\in\mathcal{M}^1}\mathcal{K}_\mu^{k,\Delta}\}}\,ds > 0\right\} = 1, \qquad \mathbf{x}\in\mathbb{R}_+^{n,\circ}. \tag{5.41}$$

Thus, if $\mathbf{X}(0)\in\mathbb{R}_+^{n,\circ}$ then $\{\mathbf{X}(s)\}$ will enter $\bigcup_{\mu\in\mathcal{M}^1}\mathcal{K}_\mu^{k,\Delta}$ with probability 1. This fact, combined with (5.39) and the strong Markov property of $\{\mathbf{X}(s)\}$, implies that

$$\sum_{\mu\in\mathcal{M}^1} P_{\mathbf{x}}^\mu > 1-\varepsilon, \qquad \mathbf{x}\in\mathbb{R}_+^{n,\circ},$$

where

$$P_{\mathbf{x}}^\mu = \mathbb{P}_{\mathbf{x}}\left\{\mathcal{U}(\omega) = \{\mu\} \text{ and } \lim_{t\to\infty}\frac{\ln X_i(t)}{t} = \lambda_i(\mu) < 0, i\in I_\mu^c\right\}.$$

Letting $\varepsilon\to 0$, we obtain (5.37). The positivity of $P_{\mathbf{x}}^\mu$ follows from (5.39) and the fact that $\{\mathbf{X}(s)\}$ will visit $\mathcal{K}_\mu^{k,\Delta}$ with a positive probability due to the nondegeneracy of the diffusion.

Next, we consider the case when $\mathcal{M}^1 \neq \varnothing$ and $\mathcal{M}^2 = \varnothing$. Then we claim that $\mathcal{M}^1 = \{\boldsymbol{\delta}^*\}$ where $\boldsymbol{\delta}^*$ be the Dirac measure concentrated on the origin $\mathbf{0}$. Indeed, if $\mathcal{M}^1$ contains a measure μ with $\mathbb{R}_+^\mu \neq \{\mathbf{0}\}$, then $\boldsymbol{\delta}^*\in\mathcal{M}_\mu$. Since μ satisfies Assumption 1.3, in view of (1.7), $\boldsymbol{\delta}^*\in\mathcal{M}^2$ which results in a contradiction. Thus, $\mathcal{M} = \mathcal{M}^1 = \{\boldsymbol{\delta}^*\}$. As a result, $\mathcal{U}(\omega) = \{\boldsymbol{\delta}^*\}$ with probability 1. Then we can easily deduce with probability 1 that

$$\lim_{t\to\infty}\frac{\ln X_i(t)}{t} = \lambda_i\big(\boldsymbol{\delta}^*\big) = f_i(\mathbf{0}) - \frac{\sigma_{ii}g_i^2(\mathbf{0})}{2} < 0, \qquad i = 1,\dots,n$$

since $\boldsymbol{\delta}^*$ satisfies (1.6). □

APPENDIX A: PROOFS FOR LEMMAS IN SECTION 3

PROOF OF LEMMA 3.1. We restrict our proof for the existence and uniqueness of the solution with initial value $\mathbf{x}\in\mathbb{R}_+^{n,\circ}$. If $\mathbf{x}\in\mathbb{R}_+^{I,\circ}$ for any $I\subset\{1,\dots,n\}$, the proof carries over. Let $V(\cdot)$ be defined by (3.4). Since $\|p\|\le\delta_0 < 1$, it is obvious that

$$\lim_{m\to\infty}\inf\big\{V(\mathbf{x}) : x_i \vee x_i^{-1} > m \text{ for some } i = 1,\dots,n\big\} = \infty. \tag{A.1}$$

Note that

$$\mathcal{L}V^{\delta_0}(\mathbf{x}) = \delta_0 V^{\delta_0}(\mathbf{x})\Big[\frac{\sum_i c_i x_i f_i(\mathbf{x})}{1+\mathbf{c}^\top\mathbf{x}} + \frac{\delta_0 - 1}{2}\frac{\sum_{i,j}\sigma_{ij}c_i c_j x_i x_j g_i(\mathbf{x})g_j(\mathbf{x})}{(1+\mathbf{c}^\top\mathbf{x})^2}$$

$$- \sum_i\Big(p_i f_i(\mathbf{x}) - \frac{p_i g_i^2(\mathbf{x})\sigma_{ii}}{2}\Big) + \frac{\delta_0}{2}\sum_{i,j} p_i p_j \sigma_{ij} g_i(\mathbf{x})g_j(\mathbf{x}) \tag{A.2}$$

$$- \delta_0 \sum_{i,j}\frac{c_i p_i x_i \sigma_{ij} g_i(\mathbf{x})g_j(\mathbf{x})}{(1+\mathbf{c}^\top\mathbf{x})}\Big].$$

Since $\|p\| \le \delta_0$, we have

$$\Big|\sum_i p_i f_i(\mathbf{x})\Big| \le \delta_0 \sum_i |f_i(\mathbf{x})| \tag{A.3}$$

and

$$\frac{\delta_0}{2}\frac{\sum_{i,j}\sigma_{ij}c_i c_j x_i x_j g_i(\mathbf{x})g_j(\mathbf{x})}{(1+\mathbf{c}^\top\mathbf{x})^2} + \sum_i \frac{p_i g_i^2(\mathbf{x})\sigma_{ii}}{2} + \frac{\delta_0}{2}\sum_{i,j} p_i p_j \sigma_{ij} g_i(\mathbf{x})g_j(\mathbf{x})$$

$$- \delta_0 \sum_{i,j}\frac{c_i p_i x_i \sigma_{ij} g_i(\mathbf{x})g_j(\mathbf{x})}{(1+\mathbf{c}^\top\mathbf{x})} \tag{A.4}$$

$$\le 3\delta_0 \sum_{i,j}\big(|g_i(\mathbf{x})g_j(\mathbf{x})\sigma_{ij}|\big).$$

Applying (3.3), (3.5), (A.3) and (A.4) to (A.2) one gets

$$\mathcal{L}V^{\delta_0}(\mathbf{x}) \le \delta_0 H V^{\delta_0}(\mathbf{x}), \qquad \mathbf{x} \in \mathbb{R}_+^{n,\circ}. \tag{A.5}$$

Since the coefficients of (1.1) are locally Lipschitz, using (A.1) and (A.5), it follows from Khasminskii [(2012), Theorem 3.5] that (1.1) has a unique solution $\mathbf{X}(t)$ that remains in $\mathbb{R}_+^{n,\circ}$ almost surely for all $t \ge 0$ whenever $\mathbf{X}(0) = \mathbf{x} \in \mathbb{R}_+^{n,\circ}$. The estimate (3.6) can also be derived from Khasminskii (2012), Theorem 3.5. □

PROOF OF LEMMA 3.2. Let $V_{\mathbf{c}}(\mathbf{x}) := (1+\mathbf{c}^\top\mathbf{x})^{\delta_0}$. By noting that

$$\delta_0\Big|\frac{\sum_{i,j}\sigma_{ij}c_i c_j x_i x_j g_i(\mathbf{x})g_j(\mathbf{x})}{(1+\mathbf{c}^\top\mathbf{x})^2}\Big| \le \delta_0 \sum_{i,j} -|g_i(\mathbf{x})g_j(\mathbf{x})\sigma_{ij}|$$

a direct calculation combined with (3.3) and (3.5) shows that

$$
\begin{aligned}
\mathcal{L}V_{\mathbf{c}}(\mathbf{x}) &= \delta_0 V_{\mathbf{c}}(\mathbf{x})\bigg[\frac{\sum_i c_i x_i f_i(\mathbf{x})}{1+\mathbf{c}^\top\mathbf{x}} + \frac{\delta_0 - 1}{2}\frac{\sum_{i,j}\sigma_{ij}c_i c_j x_i x_j g_i(\mathbf{x})g_j(\mathbf{x})}{(1+\mathbf{c}^\top\mathbf{x})^2}\bigg] \\
&\le \delta_0 V_{\mathbf{c}}(\mathbf{x})\bigg[\frac{\sum_i c_i x_i f_i(\mathbf{x})}{1+\mathbf{c}^\top\mathbf{x}} - \frac{1}{2}\frac{\sum_{i,j}\sigma_{ij}c_i c_j x_i x_j g_i(\mathbf{x})g_j(\mathbf{x})}{(1+\mathbf{c}^\top\mathbf{x})^2} \\
&\qquad + \gamma_b + \delta_0\sum_i\big(|f_i(\mathbf{x})| + g_i^2(\mathbf{x})\big) + \delta_0\sum_{i,j}|g_i(\mathbf{x})g_j(\mathbf{x})\sigma_{ij}|\bigg] \\
&\qquad - \delta_0 V_{\mathbf{c}}(\mathbf{x})\Big[\gamma_b + \delta_0\sum\big(|f_i(\mathbf{x})| + g_i^2(\mathbf{x})\big)\Big] \\
&\le \delta_0 H V_{\mathbf{c}}(\mathbf{x})\mathbf{1}_{\{\|\mathbf{x}\|\le M\}} - \delta_0 V_{\mathbf{c}}(\mathbf{x})\Big(\gamma_b + \delta_0\sum_i\big(|f_i(\mathbf{x})| + |g_i(\mathbf{x})|^2\big)\Big) \\
&\le \delta_0\tilde{H}_1 - \delta_0 V_{\mathbf{c}}(\mathbf{x})\Big(\gamma_b + \delta_0\sum_i\big(|f_i(\mathbf{x})| + |g_i(\mathbf{x})|^2\big)\Big) \qquad \forall \mathbf{x}\in\mathbb{R}_+^n,
\end{aligned} \tag{A.6}
$$

where $\tilde{H}_1 := H\sup_{\|\mathbf{y}\|\le M}\{V_c(\mathbf{y})\}$. Letting $\eta_k = \inf\{t > 0 : \|\mathbf{X}(t)\| \ge k\}$, we have by applying Dynkin's formula to the function $\varphi(\mathbf{x}, t) = e^{\gamma_b\delta_0 t}V_{\mathbf{c}}(\mathbf{x})$ and the stopping time $\eta_k \wedge t$ and making use of (A.6) that

$$
\begin{aligned}
&\mathbb{E}_{\mathbf{x}} e^{\delta_0\gamma_b(\eta_k\wedge t)}V_{\mathbf{c}}\big(\mathbf{X}(\eta_k\wedge t)\big) \\
&\quad = V_{\mathbf{c}}(\mathbf{x}) + \mathbb{E}_{\mathbf{x}}\int_0^{\eta_k\wedge t} e^{\delta_0\gamma_b s}\big(\delta_0\gamma_b V_{\mathbf{c}}\big(\mathbf{X}(s)\big) + \mathcal{L}V_{\mathbf{c}}\big(\mathbf{X}(s)\big)\big)\,ds \\
&\quad \le V_{\mathbf{c}}(\mathbf{x}) + \delta_0\tilde{H}_1\mathbb{E}_{\mathbf{x}}\int_0^{\eta_k\wedge t} e^{\delta_0\gamma_b s}\,ds \\
&\quad \le V_{\mathbf{c}}(\mathbf{x}) + H_1 e^{\delta_0\gamma_b t},
\end{aligned} \tag{A.7}
$$

where $H_1 := \gamma_b^{-1}\tilde{H}_1$. Letting $k\to\infty$ in (A.7) together with Fatou's lemma forces $e^{\delta_0\gamma_b t}\mathbb{E}_{\mathbf{x}}V_{\mathbf{c}}(\mathbf{X}(t)) \le V_{\mathbf{c}}(\mathbf{x}) + H_1 e^{\delta_0\gamma_b t}$, which in turn implies

$$
\mathbb{E}_{\mathbf{x}}V_{\mathbf{c}}\big(\mathbf{X}(t)\big) \le H_1 + V_{\mathbf{c}}(\mathbf{x})e^{-\delta_0\gamma_b t}
$$

as required. Another application of Dynkin's formula combined with (A.6) yields

$$
\begin{aligned}
&\mathbb{E}_{\mathbf{x}}V_{\mathbf{c}}\big(\mathbf{X}\big((\eta_k\wedge t)\big)\big) \\
&\quad = V_{\mathbf{c}}(\mathbf{x}) + \mathbb{E}_{\mathbf{x}}\int_0^{\eta_k\wedge t}\mathcal{L}V_{\mathbf{c}}\big(\mathbf{X}(s)\big)\,ds \\
&\quad \le V_{\mathbf{c}}(\mathbf{x}) + \delta_0\tilde{H}_1\mathbb{E}\int_0^{\eta_k\wedge t} ds \\
&\qquad - \delta_0^2\mathbb{E}_{\mathbf{x}}\int_0^{\eta_k\wedge t} V_{\mathbf{c}}\big(\mathbf{X}(s)\big)\Big[1 + \sum_i\big(\big|f_i\big(\mathbf{X}(s)\big)\big| + \big|g_i\big(\mathbf{X}(s)\big)\big|^2\big)\Big]\,ds.
\end{aligned}
$$

As a result,

$$\delta_0^2 \mathbb{E}_{\mathbf{x}} \int_0^{\eta_k \wedge t} V_{\mathbf{c}}(\mathbf{X}(s))\Big[1 + \sum_i (|f_i(\mathbf{X}(s))| + |g_i(\mathbf{X}(s))|^2)\Big]\,ds \le V_{\mathbf{c}}(\mathbf{x}) + \delta_0 \tilde{H}_1 t.$$

If we let $k \to \infty$, we obtain (3.8) with $H_2 = \delta_0^{-2} \vee (\delta_0^{-1}\tilde{H}_1)$.

Finally, since $\lim_{\|\mathbf{x}\|\to\infty} V_{\mathbf{c}}(\mathbf{x}) = \infty$, it follows easily from (A.7) that

$$\lim_{k\to\infty} \mathbb{P}_{\mathbf{x}}\{\eta_k < t\} = 0 \qquad \text{uniformly on each compact subset of } \mathbb{R}^n_+.$$

The above coupled with the assumption that $f_i(\cdot), g_i(\cdot), i = 1, \dots, n$ are locally Lipschitz allow us to modify the proof of Mao [(1997), Theorem 2.9.3] by a truncation argument in order to get the Markov–Feller property of $(\mathbf{X}(t))$. □

PROOF OF LEMMA 3.3. It suffices to suppose that μ is ergodic.

Let $\phi(\mathbf{x}) = [(1 + \mathbf{c}^\top\mathbf{x})^{\delta_0}(1 + \sum_i (|f_i(\mathbf{x})| + |g_i(\mathbf{x})|^2))$. Since μ is invariant, we have

$$\int_{\mathbb{R}^n_+} (k \wedge \phi(\mathbf{x}))\mu(d\mathbf{x}) = \lim_{t\to\infty} \int_{\mathbb{R}^n_+} \mathbb{E}_{\mathbf{x}}[k \wedge \phi(\mathbf{X}(t))]\mu(d\mathbf{x}). \tag{A.8}$$

In view of Lemma 3.2,

$$\limsup_{t\to\infty} \mathbb{E}_{\mathbf{x}}[k \wedge \phi(\mathbf{X}(t))] \le H_2, \qquad \mathbf{x} \in \mathbb{R}^n_+. \tag{A.9}$$

As a consequence of Fatou's lemma, it follows from (A.8) and (A.9) that

$$\int_{\mathbb{R}^n_+} (k \wedge \phi(\mathbf{x}))\mu(d\mathbf{x}) \le H_2 \qquad \text{for any } k \in \mathbb{N}.$$

Letting $k \to \infty$ and making use of Fatou's lemma again, we get

$$\int_{\mathbb{R}^n_+} \phi(\mathbf{x})\mu(d\mathbf{x}) \le H_2.$$

By the strong law of large numbers [see, e.g., Khasminskii (2012), Theorem 4.2] and the μ-integrability of $\sum_i (|f_i(\mathbf{x})| + |g_i^2(\mathbf{x})|)$ (due to the inequality above), one gets

$$\begin{aligned}
&\lim_{t\to\infty} \frac{1}{t}\int_0^t \Bigg[\frac{c_i X_i(s) f_i(\mathbf{X}(s))}{1 + \sum_i c_i X_i(s)} \\
&\qquad - \frac{1}{2}\frac{\sum_{i,j} \sigma_{ij} c_i c_j X_i(s) X_j(s) g_i(\mathbf{X}(s)) g_j(\mathbf{X}(s))}{(1 + \sum_i c_i X_i(s))^2}\Bigg]\,ds \\
&\quad = \int_{\mathbb{R}^n_+} \Bigg[\frac{c_i x_i f_i(\mathbf{x})}{1 + \mathbf{c}^\top\mathbf{x}} - \frac{1}{2}\frac{\sum_{i,j} \sigma_{ij} c_i c_j x_i x_j g_i(\mathbf{x}) g_j(\mathbf{x})}{(1 + \mathbf{c}^\top\mathbf{x})^2}\Bigg]\mu(d\mathbf{x}) \\
&\quad < \infty \qquad \mathbb{P}_\mu\text{-a.s.}
\end{aligned} \tag{A.10}$$

and

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t \frac{\sum_{i,j}\sigma_{ij}c_ic_jX_i(s)X_j(s)g_i(\mathbf{X}(s))g_j(\mathbf{X}(s))}{(1+\sum_i c_iX_i(s))^2}\,ds$$
$$=\int_{\mathbb{R}^n_+}\frac{\sum_{i,j}\sigma_{ij}c_ic_jx_ix_jg_i(\mathbf{x})g_j(\mathbf{x})}{(1+\mathbf{c}^\top\mathbf{x})^2}\mu(d\mathbf{x})<\infty \qquad \mathbb{P}_\mu\text{-a.s.}$$

The above limit tells us that if we let

$$Q_t := \langle L.,L.\rangle_t$$

be the quadratic variation of the local martingale

$$L_t := \int_0^t \frac{\sum_i c_iX_i(s)g_i(\mathbf{X}(s))\,dE_i(s)}{1+\sum_i c_iX_i(s)}$$

then

$$\limsup_{t\to\infty}\frac{Q_t}{t}=\int_{\mathbb{R}^n_+}\frac{\sum_{i,j}\sigma_{ij}c_ic_jx_ix_jg_i(\mathbf{x})g_j(\mathbf{x})}{(1+\mathbf{c}^\top\mathbf{x})^2}\mu(d\mathbf{x})<\infty \qquad \mathbb{P}_\mu\text{-a.s.}$$

Applying the strong law of large numbers for local martingales [see Mao (1997), Theorem 1.3.4], one can see that

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t \frac{\sum_i c_iX_i(s)g_i(\mathbf{X}(s))\,dE_i(s)}{1+\sum_i c_iX_i(s)}=0 \qquad \mathbb{P}_\mu\text{-a.s.} \tag{A.11}$$

In view of (A.10), (A.11) and Itô's formula,

$$\begin{aligned}\lim_{t\to\infty}&\frac{\ln(1+\mathbf{c}^\top\mathbf{X}(t))}{t}\\ &=\int_{\mathbb{R}^n_+}\biggl[\frac{\sum_i c_ix_if_i(\mathbf{x})}{1+\mathbf{c}^\top\mathbf{x}}\\ &\quad-\frac{1}{2}\frac{\sum_{i,j}\sigma_{ij}c_ic_jx_ix_jg_i(\mathbf{x})g_j(\mathbf{x})}{(1+\mathbf{c}^\top\mathbf{x})^2}\biggr]\mu(d\mathbf{x}) \qquad \mathbb{P}_\mu\text{-a.s.}\end{aligned} \tag{A.12}$$

A simple contradiction argument coupled with (A.12) forces

$$\int_{\mathbb{R}^n_+}\biggl[\frac{\sum_i c_ix_if_i(\mathbf{x})}{1+\mathbf{c}^\top\mathbf{x}}-\frac{1}{2}\frac{\sum_{i,j}\sigma_{ij}c_ic_jx_ix_jg_i(\mathbf{x})g_j(\mathbf{x})}{(1+\mathbf{c}^\top\mathbf{x})^2}\biggr]\mu(d\mathbf{x})=0. \qquad \square$$

PROOF OF LEMMA 3.4. Let $V_M=\sup\{(1+\mathbf{c}^\top\mathbf{x})^{\delta_0}:\|\mathbf{x}\|\le M\}$ and fix $\varepsilon>0$. Pick $l_\varepsilon\in\mathbb{N}$ such that

$$\frac{\varepsilon(1+\mathbf{c}^\top\mathbf{x})^{\delta_0-\delta}}{K_hH_2(1+V_M)}>1 \qquad \text{for any } \|\mathbf{x}\|\ge l_\varepsilon.$$

Let $\phi_l(\cdot):\mathbb{R}^n_+\to[0,1]$ be a continuous function with compact support satisfying $\phi_l(\mathbf{x})=1$ if $\|\mathbf{x}\|\le l_\varepsilon$. One gets the following sequence of inequalities:

$$\begin{aligned}
&\int_{\mathbb{R}^n_+}(1-\phi_l(\mathbf{x}))|h(\mathbf{x})|\Pi^{\mathbf{x}_k}_{T_k}(d\mathbf{x})\\
&\quad\le K_h\int_{\mathbb{R}^n_+}(1-\phi_l(\mathbf{x}))(1+\mathbf{c}^\top\mathbf{x})^\delta\\
&\qquad\times\Big(1+\sum_i(|f_i(\mathbf{x})|+|g_i(\mathbf{x})|^2)\Big)\Pi^{\mathbf{x}_k}_{T_k}(d\mathbf{x})\\
&\quad\le\frac{K_h\varepsilon}{K_hH_2(1+V_M)}\int_{\mathbb{R}^n_+}(1-\phi_l(\mathbf{x}))(1+\mathbf{c}^\top\mathbf{x})^{\delta_0}\\
&\qquad\times\Big(1+\sum_i(|f_i(\mathbf{x})|+|g_i(\mathbf{x})|^2)\Big)\Pi^{\mathbf{x}_k}_{T_k}(d\mathbf{x})\\
&\quad\le\varepsilon,
\end{aligned}\tag{A.13}$$

where the last inequality follows by (3.8). Similar to (A.13), we have from Lemma 3.3 that

$$\int_{\mathbb{R}^n_+}(1-\phi_l(\mathbf{x}))|h(\mathbf{x})|\pi(d\mathbf{x})\le\varepsilon.\tag{A.14}$$

Since $\Pi^{\mathbf{x}_k}_{T_k}$ converges weakly to π, we get

$$\lim_{k\to\infty}\int_{\mathbb{R}^n_+}\phi_l(\mathbf{x})h(\mathbf{x})\Pi^{\mathbf{x}_k}_{T_k}(d\mathbf{x})=\int_{\mathbb{R}^n_+}\phi_l(\mathbf{x})h(\mathbf{x})\pi(d\mathbf{x}).\tag{A.15}$$

As a consequence of (A.13), (A.14) and (A.15)

$$\limsup_{k\to\infty}\left|\int_{\mathbb{R}^n_+}h(\mathbf{x})\Pi^{\mathbf{x}_k}_{T_k}(d\mathbf{x})-\int_{\mathbb{R}^n_+}h(\mathbf{x})\pi(d\mathbf{x})\right|\le2\varepsilon.\tag{A.16}$$

The desired result follows by letting $\varepsilon\to0$. □

PROOF OF LEMMA 3.5. It is easy to show that there exists some $K_3>0$ such that

$$|y|^k\exp(\theta y)\le K_3(\exp(\theta_0y)+\exp(-\theta_0y)),\qquad k=1,2$$

for $\theta\in[0,\frac{\theta_0}{2}]$, $y\in\mathbb{R}$. For any $y\in\mathbb{R}$, let $\xi(y)$ be a number lying between y and 0 such that $\exp(\xi(y))=\frac{e^y-1}{y}$. Pick $\theta\in[0,\frac{\theta_0}{2})$ and let $h\in\mathbb{R}$ such that $0\le\theta+h\le\frac{\theta_0}{2}$. Then

$$\lim_{h\to0}\frac{\exp((\theta+h)Y)-\exp(\theta Y)}{h}=Y\exp(\theta Y)\qquad\text{a.s.}$$

and

$$\left|\frac{\exp((\theta+h)Y)-\exp(\theta Y)}{h}\right| = |Y|\exp\bigl(\theta Y+\xi(hY)\bigr) \\ \le 2K_3\bigl[\exp(\theta_0 Y)+\exp(-\theta_0 Y)\bigr].$$

By the Lebesgue dominated convergence theorem,

$$\frac{d\mathbb{E}\exp(\theta Y)}{d\theta} = \lim_{h\to 0}\mathbb{E}\frac{\exp((\theta+h)Y)-\exp(\theta Y)}{h} = \mathbb{E}Y\exp(\theta Y).$$

Similarly,

$$\frac{d^2\mathbb{E}\exp(\theta Y)}{d\theta^2} = \mathbb{E}Y^2\exp(\theta Y).$$

As a result, we obtain

$$\frac{d\phi}{d\theta} = \frac{\mathbb{E}Y\exp(\theta Y)}{\mathbb{E}\exp(\theta Y)}$$

which implies

$$\frac{d\phi}{d\theta}(0) = \mathbb{E}Y$$

and

$$\frac{d^2\phi}{d\theta^2} = \frac{\mathbb{E}Y^2\exp(\theta Y)\mathbb{E}\exp(\theta Y)-[\mathbb{E}Y\exp(\theta Y)]^2}{[\mathbb{E}\exp(\theta Y)]^2}.$$

By Hölder's inequality, we have $\mathbb{E}Y^2\exp(\theta Y)\mathbb{E}\exp(\theta Y)\ge[\mathbb{E}Y\exp(\theta Y)]^2$ and, therefore,

$$\frac{d^2\phi}{d\theta^2}\ge 0 \qquad \forall\theta\in\left[0,\frac{\theta_0}{2}\right).$$

Moreover,

$$\begin{aligned}\frac{d^2\phi}{d\theta^2} &\le \frac{\mathbb{E}Y^2\exp(\theta Y)}{\mathbb{E}\exp(\theta Y)} \\ &\le \frac{K_3(\mathbb{E}\exp(\theta_0 Y)+\mathbb{E}\exp(-\theta_0 Y))}{\exp(\theta\mathbb{E}Y)} \\ &\le \frac{K_3(\mathbb{E}\exp(\theta_0 Y)+\mathbb{E}\exp(-\theta_0 Y))}{\exp(-\theta_0|\mathbb{E}Y|)} \\ &< K_2\end{aligned}$$

for some K_2 depending only on K_1 and K_3. $\square$

PROOF OF LEMMA 3.6. Let $K \subset \mathbb{R}_+^{n,\circ}$ be a compact set and let D be an open, relatively compact subset of $\mathbb{R}_+^{n,\circ}$ with smooth boundary such that $K \subset D$. For $\mathbf{x} \in D$ and $t > 0$, define the measure

$$P_D(t, \mathbf{x}, \cdot) = \mathbb{P}_{\mathbf{x}}\big(\{\mathbf{X}(t) \in \cdot\} \cap \{\mathbf{X}(s) \in D, s \in [0, t]\}\big).$$

For a bounded continuous function $f : \mathbb{R}^n \mapsto \mathbb{R}$ vanishing outside D, let $u_f(t, x)$ be the solution to

$$\text{(A.17)} \qquad \begin{cases} \dfrac{\partial u}{\partial t} + \mathcal{L}u = 0 & \text{in } D \times [0, T), \\ u(T, x) = f(x) & \text{on } D, \\ u(t, x) = 0 & \text{on } \partial D \times [0, T]. \end{cases}$$

By the Feynman–Kac theorem [see, e.g., Mao (1997), Theorem 2.8.2],

$$u_f(t, x) = \int_D f(\mathbf{y}) P_D(T - t, \mathbf{x}, d\mathbf{y}).$$

Under the assumption of nondegeneracy [part (1) of Assumption 1.1], we deduce from Friedman [(2008), Theorem 3.16 and its corollary] that $P_D(t, \mathbf{x}, \cdot)$ has a density $p_D(t, \mathbf{x}, \mathbf{y})$ that is strictly positive and continuous in $(\mathbf{x}, \mathbf{y}) \in D \times D$. Since K is compact, $p_{K,D}(t, \mathbf{y}) := \inf_{\mathbf{x} \in K} p_D(t, \mathbf{x}, \mathbf{y})$ is strictly positive and continuous in $\mathbf{y} \in D$.

For $\mathbf{y} \notin D$, we define $p_{K,D}(t, \mathbf{y}) = 0$. Let $m_{K,D}(\cdot)$ be the measure whose density is $p_{K,D}(T, \mathbf{y})$. For any $\mathbf{x} \in K$ and a measurable $B \subset \mathbb{R}_+^{n,\circ}$, we have

$$P(t, \mathbf{x}, B) \geq P_D(T, \mathbf{x}, B) \geq m_{K,D}(B).$$

Thus, K is a petite set for the Markov chain $\{\mathbf{X}(kT), k \in \mathbb{N}\}$. On the other hand, since $p_D(t, \mathbf{x}, \mathbf{y})$ is strictly positive for any D, we note that

$$\text{(A.18)} \qquad P(T, \mathbf{x}, B) > 0 \qquad \text{for any set } B \text{ whose Lebesgue measure is nonzero.}$$

Thus, $\{\mathbf{X}(kT), k \in \mathbb{N}\}$ is irreducible. Moreover, it is easy to derive from (A.18) that there are no disjoint subsets of $\mathbb{R}_+^n \setminus \{\mathbf{0}\}$, denoted by $A_0, \dots, A_{d-1}$ with some $d > 1$ such that for any $\mathbf{x} \in A_i$,

$$P(T, \mathbf{x}, A_j) = 1 \qquad \text{where } j = i + 1 \ (\mathrm{mod}\ d).$$

As a result, the Markov chain $\{\mathbf{X}(kT), k \in \mathbb{N}\}$ is aperiodic. □

APPENDIX B: PROOFS FOR LEMMAS IN SECTION 4

PROOF OF LEMMA 5.4. Applying (A.6) with δ_0 replaced by $\delta_1 < \delta_0$, gives us

$$\text{(B.1)} \qquad \begin{aligned} \mathcal{L}(1 + \mathbf{c}^\top \mathbf{x})^{\delta_1} &\leq \delta_1 \tilde{H}_1 - \delta_1 (1 + \mathbf{c}^\top \mathbf{x})^{\delta_1} \\ &\quad \times \Big(\gamma_b + \delta_1 \sum_i \big(|f_i(\mathbf{x})| + |g_i(\mathbf{x})|^2\big)\Big), \qquad \mathbf{x} \in \mathbb{R}_+^n \end{aligned}$$

for some $\tilde{H}_1 > 0$. Equation (B.1) together with Itô's formula implies

$$
\begin{aligned}
\frac{(1+\mathbf{c}^\top\mathbf{X}(t))^{\delta_1}}{t} &= \frac{(1+\mathbf{c}^\top\mathbf{X}(0))^{\delta_1}}{t} + \frac{1}{t}\int_0^t \mathcal{L}\big(1+\mathbf{c}^\top\mathbf{X}(s)\big)^{\delta_1}\,ds \\
&\quad + \frac{1}{t}\int_0^t \frac{\sum_i c_i X_i(s) g_i(\mathbf{X}(s))\,dE_i(s)}{(1+\mathbf{c}^\top\mathbf{X}(s))^{1-\delta_1}}\,ds \\
&\le \frac{(1+\mathbf{c}^\top\mathbf{X}(0))^{\delta_1}}{t} + \delta_1\tilde{H}_2 + \frac{1}{t}\int_0^t \frac{\sum_i c_i X_i(s) g_i(\mathbf{X}(s))\,dE_i(s)}{(1+\mathbf{c}^\top\mathbf{X}(s))^{1-\delta_1}} \\
&\quad - \delta_1\frac{1}{t}\int_0^t \big(1+\mathbf{c}^\top\mathbf{X}(s)\big)^{\delta_1} \\
&\quad \times \Big(\gamma_b + \delta_1\sum_i \big(\big|f_i(\mathbf{X}(s))\big| + \big|g_i(\mathbf{X}(s))\big|^2\big)\Big)\,ds.
\end{aligned}
$$

Since $\frac{(1+\mathbf{c}^\top\mathbf{X}(t))^{\delta_1}}{t} \ge 0$, the above yields

$$
\begin{aligned}
&\frac{\delta_1}{2t}\int_0^t \big(1+\mathbf{c}^\top\mathbf{X}(s)\big)^{\delta_1}\Big(\gamma_b + \delta_1\sum_i \big(\big|f_i(\mathbf{X}(s))\big| + \big|g_i(\mathbf{X}(s))\big|^2\big)\Big) \\
&\quad \le \frac{(1+\mathbf{c}^\top\mathbf{X}(0))^{\delta_1}}{t} + \delta_1\tilde{H}_2 + \frac{1}{t}\int_0^t \frac{\sum_i c_i X_i(s) g_i(\mathbf{X}(s))\,dE_i(s)}{(1+\mathbf{c}^\top\mathbf{X}(s))^{1-\delta_1}} \\
&\quad\quad - \frac{\delta_1}{2t}\int_0^t \big(1+\mathbf{c}^\top\mathbf{X}(s)\big)^{\delta_1}\Big(\gamma_b + \delta_1\sum_i \big(\big|f_i(\mathbf{X}(s))\big| + \big|g_i(\mathbf{X}(s))\big|^2\big)\Big)\,ds.
\end{aligned} \tag{B.2}
$$

For each $i = 1, \dots, n$, the quadratic variation of

$$
\int_0^t \frac{c_i X_i(s) g_i(\mathbf{X}(s))\,dE_i(s)}{(1+\mathbf{c}^\top\mathbf{X}(s))^{1-\delta_1}}
$$

is

$$
Q_t := \int_0^t \frac{[c_i X_i(s) g_i(\mathbf{X}(s))]^2\sigma_{ii}}{(1+\mathbf{c}^\top\mathbf{X}(s))^{2-2\delta_1}}\,ds.
$$

We have the following estimate for each $i = 1, \dots, n$:

$$
\begin{aligned}
\frac{[c_i x_i g_i(\mathbf{x})]^2\sigma_{ii}}{(1+\mathbf{c}^\top\mathbf{x})^{2-2\delta_1}} &\le \big(1+\mathbf{c}^\top\mathbf{x}\big)^{2\delta_1} g_i^2(\mathbf{x})\sigma_{ii} \\
&\le K_i\Big[\big(1+\mathbf{c}^\top\mathbf{x}\big)^{\delta_1}\Big(\gamma_b + \delta_1\sum_i \big(\big|f_i(\mathbf{x})\big| + \big|g_i(\mathbf{x})\big|^2\big)\Big)\Big],
\end{aligned}
$$

where due to Assumption 1.4

$$
K_i = \sup_{\mathbf{x}\in\mathbb{R}^n_+}\left\{\frac{(1+\mathbf{c}^\top\mathbf{x})^{\delta_1} g_i^2(\mathbf{x})\sigma_{ii}}{\gamma_b + \delta_1\sum_i(|f_i(\mathbf{x})| + |g_i(\mathbf{x})|^2)}\right\} < \infty.
$$

Thus,

$$\int_0^\infty \frac{dQ_t}{(1+A_t)^2} \le \int_0^\infty K_i \frac{dA_t}{(1+A_t)^2} = K_i < \infty \qquad \text{a.s.,} \tag{B.3}$$

where

$$A_t := \int_0^u (1+\mathbf{c}^\top \mathbf{X}(s))^{\delta_1} \Big(\gamma_b + \delta_1 \sum_i (|f_i(\mathbf{X}(s))| + |g_i(\mathbf{X}(s))|^2)\Big)\, ds.$$

On the other hand,

$$\lim_{t\to\infty} A_t \ge \lim_{t\to\infty} \gamma_b t = \infty \qquad \text{a.s.} \tag{B.4}$$

By (B.3), (B.4) we can use the strong law of large numbers for local martingales (see Mao [(1997), Theorem 1.3.4]) in order to obtain for each $i = 1, \dots, n$ that

$$\begin{aligned} &\lim_{t\to\infty} \frac{\int_0^t c_i X_i(s) g_i(\mathbf{X}(s)(1+\mathbf{c}^\top \mathbf{X}(s))^{\delta_1 - 1}\, dE_i(s)}{\int_0^t (1+\mathbf{c}^\top \mathbf{X}(s))^{\delta_1}(\gamma_b + \delta_1 \sum_i (|f_i(\mathbf{X}(s))| + |g_i(\mathbf{X}(s))|^2))\, ds} \\ &\quad = 0 \qquad \mathbb{P}_{\mathbf{x}}\text{-a.s.} \end{aligned} \tag{B.5}$$

This implies

$$\begin{aligned} &\limsup_{t\to\infty} \bigg[\frac{1}{t}\int_0^t \frac{\sum_i c_i X_i(s) g_i(\mathbf{X}(s))\, dE_i(s)}{(1+\mathbf{c}^\top \mathbf{X}(s))^{1-\delta_1}} \\ &\quad - \frac{\delta_1}{2t}\int_0^t (1+\mathbf{c}^\top \mathbf{X}(s))^{\delta_1} \Big(\gamma_b + \delta_1 \sum_i (|f_i(\mathbf{X}(s))| + |g_i(\mathbf{X}(s))|^2)\Big)\, ds\bigg] \\ &\quad \le 0 \qquad \mathbb{P}_{\mathbf{x}}\text{-a.s.} \end{aligned} \tag{B.6}$$

Applying (B.6) to (B.2), we get

$$\begin{aligned} &\limsup_{t\to\infty} \frac{\delta_1}{2t}\int_0^t (1+\mathbf{c}^\top \mathbf{X}(s))^{\delta_1} \Big(\gamma_b + \delta_1 \sum_i (|f_i(\mathbf{X}(s))| + |g_i(\mathbf{X}(s))|^2)\Big)\, ds \\ &\quad \le \delta_1 \tilde{H}_1 \qquad \mathbb{P}_{\mathbf{x}}\text{-a.s.} \end{aligned} \tag{B.7}$$

Similar to the proof of (A.11), we can obtain (5.18) from (B.7) and the strong law of large numbers for local martingales. The proof is complete. □

PROOF OF LEMMA 5.5. Let $\widehat{K}_1$ be sufficiently large such that $(1+\mathbf{c}^\top \mathbf{x})^{\delta_1} > 2\widehat{K}$ if $\|\mathbf{x}\| \ge \widehat{K}_1$. By Lemma 5.4,

$$\begin{aligned} \limsup_{t\to\infty} \frac{1}{t}\int_0^t \mathbf{1}_{\{\|\mathbf{X}(s)\| > \widehat{K}_1\}}\, ds &\le \frac{1}{2\widehat{K}} \limsup_{t\to\infty} \frac{1}{t}\int_0^t (1+\mathbf{c}^\top \mathbf{X}(s))^{\delta_1}\, ds \\ &\le \frac{1}{2\widehat{K}} \widehat{K} = \frac{1}{2} \qquad \mathbb{P}_{\mathbf{x}}\text{-a.s., } \mathbf{x} \in \mathbb{R}^n_+, \end{aligned}$$

which implies (5.19).

Next, we prove (5.20). Fix $i \in \{1, \dots, n\}$ and define $\tilde{V}_i = \frac{1+\mathbf{c}^\top \mathbf{x}}{x_i^{\delta_0}}$ on $\{\mathbf{x} \in \mathbb{R}^n_+ : x_i > 0\}$. Similar to (A.5), it can be shown that

$$\mathcal{L}\tilde{V}_i^{\delta_0}(\mathbf{x}) \le \delta_0 H \tilde{V}_i^{\delta_0}(\mathbf{x}), \qquad \mathbf{x} \in \mathbb{R}^n_+, x_i > 0.$$

Let $\zeta_k := \inf\{t > 0 : X_i(t)^{-1} \vee \|\mathbf{X}(t)\| > k\}$. We have by Dynkin's formula that

$$\begin{aligned}\mathbb{E}_\mathbf{x} \tilde{V}_i^{\delta_0}\big(\mathbf{X}((n_e T_e) \wedge \zeta_k)\big) &\le \tilde{V}_i^{\delta_0}(\mathbf{x}) + \mathbb{E}_\mathbf{x} \int_0^{(n_e T_e)\wedge\zeta_k} \mathcal{L}\tilde{V}_i^{\delta_0}\big(\mathbf{X}(s)\big)\, ds \\ &\le \tilde{V}_i^{\delta_0}(\mathbf{x}) + \delta_0 H \int_0^{n_e T_e} \mathbb{E}_\mathbf{x} \tilde{V}_i^{\delta_0}\big(\mathbf{X}(s \wedge \zeta_k)\big)\, ds.\end{aligned}$$

Using Gronwall's inequality yields

$$\mathbb{E}_\mathbf{x} \tilde{V}_i^{\delta_0}\big(\mathbf{X}((n_e T_e) \wedge \zeta_k)\big) \le \tilde{V}_i^{\delta_0}(\mathbf{x}) \exp(\delta_0 H n_e T_e), \qquad \mathbf{x} \in \mathbb{R}^n_+, x_i > 0. \tag{B.8}$$

Let $k_1 \in \mathbb{N}$ sufficiently large such that

$$\begin{aligned}\tilde{V}_i^{\delta_0}(\mathbf{y}) > \frac{1}{\varepsilon_1} \sup_{\|\mathbf{x}\| \le \widehat{K}_1, x_i \ge \varepsilon_2} \big\{\tilde{V}_i^{\delta_0}(\mathbf{x})\big\} \exp(\delta_0 H n_e T_e) \\ \text{for all } \mathbf{y} \in \mathbb{R}^n_+, y_i^{-1} \vee \|\mathbf{y}\| > k_1.\end{aligned} \tag{B.9}$$

It follows from (B.8) and (B.9) that

$$\begin{aligned}\mathbb{P}_\mathbf{x}\{\zeta_{k_1} < n_e T_e\} \le \frac{\mathbb{E}_\mathbf{x} \tilde{V}_i^{\delta_0}(\mathbf{X}((n_e T_e) \wedge \zeta_{k_1}))}{\inf\{\tilde{V}_i^{\delta_0}(\mathbf{y}) : \mathbf{y} \in \mathbb{R}^n_+, y_i^{-1} \vee \|\mathbf{y}\| > k_1\}} \le \varepsilon_1 \\ \text{for } \mathbf{x} \in \mathbb{R}^n_+, \|\mathbf{x}\| \le \widehat{K}_1, x_i \ge \varepsilon_2.\end{aligned}$$

Now inequality (5.20) follows by straightforward computations. □

Acknowledgements. The authors thank Michel Benaïm for helpful discussions and for sending them his manuscript [Benaïm (2014)] which was key in proving the persistence results from this paper. We also thank two anonymous referees for their suggestions and comments which helped improve this manuscript.

REFERENCES

BENAÏM, M. (2014). Stochastic persistence. Preprint.

BENAÏM, M. and SCHREIBER, S. J. (2009). Persistence of structured populations in random environments. *Theor. Popul. Biol.* **76** 19–34.

BENAÏM, M., HOFBAUER, J. and SANDHOLM, W. H. (2008). Robust permanence and impermanence for stochastic replicator dynamics. *J. Biol. Dyn.* **2** 180–195. MR2427526

BENAÏM, M. and LOBRY, C. (2016). Lotka–Volterra with randomly fluctuating environments or "How switching between beneficial environments can make survival harder." *Ann. Appl. Probab.* **26** 3754–3785. MR3582817

BLATH, J., ETHERIDGE, A. and MEREDITH, M. (2007). Coexistence in locally regulated competing populations and survival of branching annihilating random walk. *Ann. Appl. Probab.* **17** 1474–1507. MR2358631

BRAUMANN, C. A. (2002). Variable effort harvesting models in random environments: Generalization to density-dependent noise intensities. *Math. Biosci.* **177** 229–245.

CASWELL, H. (2001). *Matrix Population Models*. Wiley, New York.

CATTIAUX, P. and MÉLÉARD, S. (2010). Competitive or weak cooperative stochastic Lotka–Volterra systems conditioned on non-extinction. *J. Math. Biol.* **60** 797–829. MR2606515

CATTIAUX, P., COLLET, P., LAMBERT, A., MARTÍNEZ, S., MÉLÉARD, S. and SAN MARTÍN, J. (2009). Quasi-stationary distributions and diffusion models in population dynamics. *Ann. Probab.* **37** 1926–1969. MR2561437

CHEN, Z. and KULPERGER, R. (2005). A stochastic competing-species model and ergodicity. *J. Appl. Probab.* **42** 738–753. MR2157517

CHESSON, P. (2000). General theory of competitive coexistence in spatially-varying environments. *Theor. Popul. Biol.* **58** 211–237.

CHESSON, P. L. and ELLNER, S. (1989). Invasibility and stochastic boundedness in monotonic competition models. *J. Math. Biol.* **27** 117–138. MR991046

CROSS, P. C., LLOYD-SMITH, J. O., JOHNSON, P. L. F. and GETZ, W. M. (2005). Duelling timescales of host movement and disease recovery determine invasion of disease in structured populations. *Ecol. Lett.* **8** 587–595.

DAVIES, K. F., CHESSON, P., HARRISON, S., INOUYE, B. D., MELBOURNE, B. and RICE, K. J. (2005). Spatial heterogeneity explains the scale dependence of the native-exotic diversity relationship. *Ecology* **86** 1602–1610.

ETHIER, S. N. and KURTZ, T. G. (2009). *Markov Processes: Characterization and Convergence* **282**. Wiley, New York.

EVANS, S. N., HENING, A. and SCHREIBER, S. J. (2015). Protected polymorphisms and evolutionary stability of patch-selection strategies in stochastic environments. *J. Math. Biol.* **71** 325–359.

EVANS, S. N., RALPH, P. L., SCHREIBER, S. J. and SEN, A. (2013). Stochastic population growth in spatially heterogeneous environments. *J. Math. Biol.* **66** 423–476. MR3010201

FRIEDMAN, A. (2008). *Partial Differential Equations of Parabolic Type*. Dover Publications, Mineola, NY.

GARD, T. C. (1988). *Introduction to Stochastic Differential Equations*. *Monographs and Textbooks in Pure and Applied Mathematics* **114**. Dekker, New York. MR0917064

HENING, A. and NGUYEN, D. (2017a). Stochastic Lotka–Volterra food chains. *J. Math. Biol.* To appear.

HENING, A. and NGUYEN, D. (2017b). Persistence in stochastic Lotka–Volterra food chains with intraspecific competition. Preprint.

HENING, A., NGUYEN, D. H. and YIN, G. (2018). Stochastic population growth in spatially heterogeneous environments: The density-dependent case. *J. Math. Biol.* To appear.

HOFBAUER, J. (1981). A general cooperation theorem for hypercycles. *Monatsh. Math.* **91** 233–240. MR0619966

HOFBAUER, J. and SO, J. W.-H. (1989). Uniform persistence and repellors for maps. *Proc. Amer. Math. Soc.* **107** 1137–1142. MR984816

HUTSON, V. (1984). A theorem on average Liapunov functions. *Monatsh. Math.* **98** 267–275. MR776353

KHASMINSKII, R. Z. (1960). Ergodic properties of recurrent diffusion processes and stabilization of the solution to the Cauchy problem for parabolic equations. *Theory Probab. Appl.* **5** 179–196.

KHASMINSKII, R. (2012). *Stochastic Stability of Differential Equations*, 2nd ed. *Stochastic Modelling and Applied Probability* **66**. Springer, Heidelberg. MR2894052

LANDE, R., ENGEN, S. and SAETHER, B.-E. (2003). *Stochastic Population Dynamics in Ecology and Conservation*. Oxford Univ. Press, Oxford.

LAW, R. and MORTON, R. D. (1996). Permanence and the assembly of ecological communities. *Ecology* **77** 762–775.

LIU, M. and BAI, C. (2016). Analysis of a stochastic tri-trophic food-chain model with harvesting. *J. Math. Biol.* **73** 597–625.

MAO, X. (1997). *Stochastic Differential Equations and Their Applications*. Horwood Publishing Limited, Chichester. MR1475218

MEYN, S. P. and TWEEDIE, R. L. (1992). Stability of Markovian processes. I. Criteria for discrete-time chains. *Adv. in Appl. Probab.* **24** 542–574. MRMR1174380

NUMMELIN, E. (1984). *General Irreducible Markov Chains and Nonnegative Operators*. *Cambridge Tracts in Mathematics* **83**. Cambridge Univ. Press, Cambridge. MR776608

PYŠEK, P. and HULME, P. E. (2005). Spatio-temporal dynamics of plant invasions: Linking pattern to process. *Ecoscience* **12** 302–315.

RUDNICKI, R. (2003). Long-time behaviour of a stochastic prey-predator model. *Stochastic Process. Appl.* **108** 93–107. MR2008602

RUDNICKI, R. and PICHÓR, K. (2007). Influence of stochastic perturbation on prey–predator systems. *Math. Biosci.* **206** 108–119. MR2311674

SCHREIBER, S. J., BENAÏM, M. and ATCHADÉ, K. A. S. (2011). Persistence in fluctuating environments. *J. Math. Biol.* **62** 655–683. MR2786721

SCHREIBER, S. J. and LLOYD-SMITH, J. O. (2009). Invasion dynamics in spatially heterogeneous environments. *Amer. Nat.* **174** 490–505.

TURELLI, M. (1977). Random environments and stochastic calculus. *Theor. Popul. Biol.* **12** 140–178.

DEPARTMENT OF MATHEMATICS
TUFTS UNIVERSITY
BROMFIELD-PEARSON HALL
503 BOSTON AVENUE
MEDFORD, MASSACHUSETTS 02155
USA
E-MAIL: alexandru.hening@tufts.edu

DEPARTMENT OF MATHEMATICS
WAYNE STATE UNIVERSITY
DETROIT, MICHIGAN 48202
USA
E-MAIL: dangnh.maths@gmail.com