Available online at www.sciencedirect.com

ScienceDirect

Stochastic Processes and their Applications 142 (2021) 319–364

stochastic processes and their applications

www.elsevier.com/locate/spa

Stochastic functional Kolmogorov equations, I: Persistence☆

Dang H. Nguyen[a], Nhu N. Nguyen[b], George Yin[b,*]

[a] Department of Mathematics, University of Alabama, Tuscaloosa, AL 35487, USA
[b] Department of Mathematics, University of Connecticut, Storrs, CT 06269, USA

Received 24 January 2021; received in revised form 21 May 2021; accepted 6 September 2021
Available online 16 September 2021

Abstract

This work (Part (I)) together with its companion (Part (II)) develops a new framework for stochastic functional Kolmogorov equations, which are nonlinear stochastic differential equations depending on the current as well as the past states. Because of the complexity of the results, it seems to be instructive to divide our contributions to two parts. In contrast to the existing literature, our effort is to advance the knowledge by allowing delay and past dependence, yielding essential utility to a wide range of applications. A long-standing question of fundamental importance pertaining to biology and ecology is: What are the minimal necessary and sufficient conditions for long-term persistence and extinction (or for long-term coexistence of interacting species) of a population? Regardless of the particular applications encountered, persistence and extinction are properties shared by Kolmogorov systems. While there are many excellent treaties of stochastic-differential-equation-based Kolmogorov equations, the work on stochastic Kolmogorov equations with past dependence is still scarce. Our aim here is to answer the aforementioned basic question. This work, Part (I), is devoted to characterization of persistence, whereas its companion, Part (II) is devoted to extinction. The main techniques used in this paper include the newly developed functional Itô formula and asymptotic coupling and Harris-like theory for infinite dimensional systems specialized to functional equations. General theorems for stochastic functional Kolmogorov equations are developed first. Then a number of applications are examined covering, improving, and substantially extending the existing literature. Furthermore, our results reduce to that in the existing literature of Kolmogorov systems when there is no past dependence.

☆ The research of D. H. Nguyen was supported in part by the National Science Foundation under grant DMS-1853467. The research of N. N. Nguyen and G. Yin was supported in part by the National Science Foundation under grant DMS-2114649.

* Corresponding author.
E-mail addresses: dangnh.maths@gmail.com (D.H. Nguyen), nguyen.nhu@uconn.edu (N.N. Nguyen), gyin@uconn.edu (G. Yin).

https://doi.org/10.1016/j.spa.2021.09.007

MSC: 34K50; 60J60; 60J70; 92B99

Keywords: Stochastic functional equation; Kolmogorov system; Ecological and biological application; Invariance measure; Extinction; Persistence

1. Introduction

This work develops a novel framework of systems of stochastic functional Kolmogorov equations. Our main motivation stems from a wide variety of applications in ecology and biology. A long-standing question of fundamental importance pertaining to biology and ecology is: What are the minimal (necessary and sufficient) conditions for long-term persistence and extinction (or for long-term coexistence of interacting species) of a population? It turns out that persistence and extinction are phenomena go far beyond biological and ecological systems. In fact, such long-term properties are shared by all processes of Kolmogorov type. We focus on the issues for such systems that involve stochastic disturbances and past dependence in the dynamics. The problems are substantially more difficult compared to systems without delay or past independence because one has to treat infinite dimensional processes.

An n-dimensional deterministic Kolmogorov system is an autonomous system of equations to depict the dynamics of n interacting populations, which takes the form

$$\dot{x}_i(t) = x_i(t) f_i(x_1(t), \dots, x_n(t)), \; i = 1, \dots, n, \tag{1.1}$$

where $f_i(\cdot)$ are functions satisfying suitable conditions. Realizing that fluctuations of the environment make the dynamics of populations inherently stochastic, much effort has been placed on the study of stochastic Kolmogorov equations. As an example, consider a simple 2-dimensional Kolmogorov equations with stochastic effects:

$$\begin{cases} dx(t) = x(t) f_1(x(t), y(t))dt + x(t) g_1(x(t), y(t)) dB_1(t), \\ dy(t) = y(t) f_2(x(t), y(t))dt + y(t) g_2(x(t), y(t)) dB_2(t), \end{cases} \tag{1.2}$$

where $B_1(t)$ and $B_2(t)$ are two Brownian motions (independent or not). The formulation readily generalizes to n-dimensional stochastic Kolmogorov equations, which are used extensively in the modeling and analysis of ecological and biological systems such as Lotka–Volterra predator–prey models, Lotka–Volterra competitive models, replicator dynamic systems, stochastic epidemic models, and stochastic chemostat models, among others. The study of such systems has encompassed the central issues of persistence and extinction as well as the existence of invariant measures. Apart from ecological and biological systems, numerous problems arising in mathematical physics, statistical mechanics, and many related fields, use Kolmogorov stochastic differential equations. We mention a simple one-dimensional generalized Ginzburg–Landau equation

$$dx(t) = x(t)[a(t) - b(t)x^k(t)] + x(t)\sigma(x(t))dB(t), \; x(0) = x_0 > 0, \tag{1.3}$$

where $k \geq 2$ is a positive integer, $B(t)$ is a real-valued Brownian motion. Such equations have been used in the theory of bistable systems, chemical turbulence, phase transitions in non-equilibrium systems, nonlinear, optics with dissipation, thermodynamics, and hydrodynamic systems, etc.

Because of its prevalence in applications, Kolmogorov systems have attracted much attention in the past decades; substantial progress has been made. To proceed, let us briefly recall some of the developments to date. Some of the early mathematical formulations were introduced by

Verhulst [62] for logistic models, by Lotka and Volterra [39,63] for Lotka–Volterra systems, and by Kermack and McKendrick [29,30] for infectious diseases modeling using ordinary differential equations in the last century. The study on mathematical models has stimulated subsequent work with attention devoted to analyzing and predicting the behavior of the populations in a longtime horizon. Subsequently, not only deterministic systems, but also stochastic systems have been studied. Resurgent effort has been devoted to finding the corresponding classification by means of threshold levels. Fast forward, Imhof studied long-run behavior of the stochastic replicator dynamics in [27], whereas Hofbauer and Imhof concentrated on time averages, recurrence, and transience for stochastic replicator dynamics in [25]. By now, Kolmogorov stochastic population systems (using stochastic differential equations or difference equations) together with their longtime behavior have been relatively well understood; see [4,54,56] for Kolmogorov stochastic systems in compact domains and [3,24] for certain general Kolmogorov systems in non-compact domains. Variants of Kolmogorov systems such as epidemic models [12,13,15,47], tumor-immune systems [61] and chemostat models [44], etc. have also been studied. In contrast to numerous papers that used Lyapunov function methods to analyze the underlying systems with limited success, Benaïm [3], Benaïm and Lobry [4], Benaïm and Strickler [5], Chesson and Ellner [8], Evans, Hening, and Schreiber [20], and Schreiber and Benaïm [56] initiated the study by examining the corresponding boundary behavior and considered the stochastic rate of growth; see also Du, Nguyen, and Yin [16]. For the most recent development and substantial progress, we refer to Benaïm [3], Hening and Nguyen [24], Schreiber and Benaïm [56], and references therein.

Our study in this work is to consider a class of n-dimensional stochastic functional Kolmogorov systems; our effort is to substantially advance the existing literature by allowing delay and past dependence, which in turn, provides essential utility to a wide range of applications. Why is it important to consider systems with delays as well as stochastic functional Kolmogorov systems? Mainly, the delays or past dependence are unavoidable in natural phenomena and dynamical systems; the framework of stochastic functional differential equations is more realistic, more effective, and more general for the population dynamics in real life than a stochastic differential equation counterpart. In population dynamics, some delay mechanisms studied in the literature include age structure, feeding times, replenishment or regeneration time for resources [11]. Although there are many excellent treatises of Kolmogorov stochastic differential equations, the work on Kolmogorov stochastic differential equations with delay is relatively scarce. A few exceptions are the study on stochastic delay Lotka–Volterra competitive models [1,32], the work on stochastic delay Lotka–Volterra predator–prey models [22,33,35, 38,66], the treatment of stochastic delay epidemic SIR models [7,34,36,37,40], and the study on stochastic delay chemostat models [57,58,67]. Nevertheless, other than the specific models and applications treated, there has not been a unified framework and a systematic treatment for Kolmogorov stochastic functional differential systems yet. Moreover, most of the existing results involving delay are not as sharp as desired. Our effort in this paper takes up the aforementioned issues.

It should be noted that from stochastic Kolmogorov differential equation-type models to that of stochastic functional differential equation models requires a big leap. There are essential difficulties. While the solutions of stochastic differential equations are Markovian processes, the solutions of stochastic differential equations with delay is non-Markov. One typically uses the so-called segment processes for the delay equations. However, such segment processes live in an infinite dimensional space. Many of the known results in the usual stochastic differential equation setup are no longer applicable. Besides, because Kolmogorov systems are highly nonlinear, analyzing such systems with delay becomes even more difficult. New methods and techniques need to be developed to carry out the analysis. This brings us to the current work.

In this paper, we set up the problem in a unified form, develop new methodology to characterize the longtime behavior of the underlying system, establish results for persistence, and demonstrate the utility in a number of applications arising in ecology and biology. Our goal is to obtain sharp results under mild and verifiable conditions, which is useful for a wide variety of stochastic functional Kolmogorov systems. In view of the progress and challenges, this work combines the techniques of functional analysis (in particular, the functional Itô formula) in [9,10], stochastic differential equations (SDEs) in infinite dimension, as well as the methods of asymptotic couplings [23], to develop a new framework for treating functional Kolmogorov systems. It will substantially generalize the methods in [3,24]. Our results will cover, improve, and advance existing results for Kolmogorov systems with and without delays. It should be mentioned that in the case of replicator dynamics, it seems to be no investigation of delayed stochastic systems to date to the best of our knowledge.

Although the models with functional stochastic differential equations are more realistic and more general, the analysis of such systems become far more difficult. Perhaps, part of the difficulties in studying stochastic delay systems is that there had been virtually no bona fide operators and functional Itô formulas except some general setup in a Banach space such as [43] before 2009. In [17], Dupire generalized the Itô formula to a functional setting by using pathwise functional derivatives. The Itô formula developed has substantially eased the difficulties and encouraged subsequent development with a wide range of applications. His work was developed further by Cont and Fournié [9,10]. Using the newly developed functional Itô formula enables us to analyze effectively the segment processes in the stochastic functional Kolmogorov equations.

Because of the non-Markovian property of the solution processes due to delay and the use of memory segment functions, one needs to analyze the corresponding stochastic equations in an infinite dimensional space. Handling occupation measures in an infinite dimensional space to obtain the tightness and characterize its limit is more challenging, so is to prove the uniqueness of the invariant probability measure. The associated Markov semigroups are often not strong Feller, even in some simple cases. Because of the absence of the strong Feller property, Doob's method to prove the uniqueness of the invariant probability measure is no longer applicable; see [53]. There are some recent works on asymptotic analysis for functional stochastic differential equations; for example, see [2] and references therein. Most notably, in [23], Hairer, Mattingly, and Scheutzow developed a necessary and sufficient condition for the uniqueness of the invariant probability measure using asymptotic couplings, provided sufficient conditions for weak convergence to the invariant probability measure, and obtained a Harris-like theory for general infinite-dimensional spaces. By using ideas from this abstract theory and our subtle estimates for certain coupled systems, we are able to prove the uniqueness of the invariant probability measure for the Kolmogorov systems. To characterize the longtime behavior of the underlying system under natural conditions and to develop a systematic method for this kind problem, we use the intuition from dynamical system theory, in which we need to examine the corresponding problem on the boundary and reveal the behavior of the process when it is close to the boundary. Nevertheless, the behavior of solutions near the boundary for functional Kolmogorov systems requires more delicate analysis than that for systems without delay. Even if the current state is close to the boundary, its history may not be.

The rest of the paper is organized as follows. Section 2 presents the formulation of the problem as well as mathematical definitions and terminologies, and state our main results.

Section 3 examines basic properties of Kolmogorov equations with delays, including well-posedness of the system, and positivity of solutions. Also obtained are the tightness of families of occupation measures and the convergence to the corresponding invariant probability measures. To obtain the desired theory, a number of key auxiliary results are provided. Then the conditions for persistence of Kolmogorov systems are given in Section 4. Finally, Section 5 provides several applications involving Kolmogorov dynamical systems and detailed account on how to use our results to treat each of the application examples.

2. Main results

To help the reading, we first provide a glossary of symbols and notation to be used in this paper.

r	a fixed positive number
$\|\cdot\|$	Euclidean norm
$\mathcal{C}([a;b];\mathbb{R}^n)$	set of $\mathbb{R}^n$-valued continuous functions defined on $[a;b]$
$\mathcal{C}$	$:= \mathcal{C}([-r;0];\mathbb{R}^n)$
$\boldsymbol{\varphi}$	$= (\varphi_1,\dots,\varphi_n) \in \mathcal{C}$
$\vec{x}$	$= (x_1,\dots,x_n) := \boldsymbol{\varphi}(0) \in \mathbb{R}^n$
$\\|\boldsymbol{\varphi}\\|$	$:= \sup\{\|\boldsymbol{\varphi}(t)\| : t \in [-r,0]\}$
$\vec{X}_t$	$:= \vec{X}_t(s) := \{\vec{X}(t+s) : -r \le s \le 0\}$ (segment function)
$X_{i,t}$	$:= X_{i,t}(s) := \{X_i(t+s) : -r \le s \le 0\}$
$\mathcal{C}_+$	$:= \{\boldsymbol{\varphi} = (\varphi_1,\dots,\varphi_n) \in \mathcal{C} : \varphi_i(s) \ge 0\ \forall s \in [-r,0], i = 1,\dots,n\}$
$\partial\mathcal{C}_+$	$:= \{\boldsymbol{\varphi} = (\varphi_1,\dots,\varphi_n) \in \mathcal{C} : \\|\varphi_i\\| = 0$ for some $i = 1,\dots,n\}$
$\mathcal{C}_+^\circ$	$:= \{\boldsymbol{\varphi} \in \mathcal{C}_+ : \varphi_i(s) > 0, \forall s \in [-r,0], i = 1,\dots,n\} \ne \mathcal{C}_+ \setminus \partial\mathcal{C}_+$
$\\|\boldsymbol{\varphi}\\|_\alpha$	$:= \\|\boldsymbol{\varphi}\\| + \sup_{-r\le s<t\le 0} \frac{\|\boldsymbol{\varphi}(t)-\boldsymbol{\varphi}(s)\|}{(t-s)^\alpha}$, for some $0<\alpha<1$
$\mathcal{C}^\alpha$	space of Hölder continuous functions endowed with the norm $\\|\cdot\\|_\alpha$
Γ	$n \times n$ matrix
$\Gamma^\top$	transpose of Γ
$\vec{B}(t)$	$= (B_1(t),\dots,B_n(t))^\top$, a n-dimensional standard Brownian motion
$\vec{E}(t)$	$= (E_1(t),\dots,E_n(t))^\top := \Gamma^\top \vec{B}(t)$
Σ	$= (\sigma_{ij})_{n\times n} := \Gamma^\top\Gamma$
$\mathcal{M}$	set of ergodic invariant probability measures of $\vec{X}_t$ supported on $\partial\mathcal{C}_+$
$\mathrm{Conv}(\mathcal{M})$	convex hull of $\mathcal{M}$
$\vec{0}$	the zero constant function in $\mathcal{C}$
$\boldsymbol{\delta}^*$	the Dirac measure concentrated at $\vec{0}$
$\mathbf{1}_A$	the indicator function of set A
$D_{\varepsilon,R}$	$:= \{\boldsymbol{\varphi} \in \mathcal{C}_+ : \\|\boldsymbol{\varphi}\\| \le R, x_i \ge \varepsilon\ \forall i; \vec{x} := \boldsymbol{\varphi}(0)\}, \varepsilon, R > 0$
$\mathbb{D}$	space of Cadlag functions mapping $[-r,0]$ to $\mathbb{R}^n$
A_0, A_1, A_2	constants satisfying Assumption 2.1
γ_0, γ_b, M	constants satisfying Assumption 2.1
$\vec{c}, h(\cdot), \mu$	vector, function and probability measure satisfying Assumption 2.1
$\widetilde{K}, b_1, b_2$	constants satisfying Assumption 2.2
$h_1(\cdot), \mu_1$	function and probability measure satisfying Assumption 2.2
D_0, d_0	constants satisfying Assumption 2.4
I	a subset of $\{1,\dots,n\}$
I^c	$:=\{1,\dots,n\} \setminus I$

$\mathcal{C}_+^I$	$:= \{\boldsymbol{\varphi} \in \mathcal{C}_+ : \|\varphi_i\| = 0 \text{ if } i \in I^c\}$
$\mathcal{C}_+^{I,\circ}$	$:= \{\boldsymbol{\varphi} \in \mathcal{C}_+ : \|\varphi_i\| = 0 \text{ if } i \in I^c \text{ and } \varphi_i(s) > 0\ \forall s \in [-r, 0] \text{ if } i \in I\}$
$\partial\mathcal{C}_+^I$	$:= \{\boldsymbol{\varphi} \in \mathcal{C}_+ : \|\varphi_i\| = 0 \text{ if } i \in I^c \text{ and } \|\varphi_i\| = 0 \text{ for some } i \in I\}$
$\mathcal{M}^I$	sets of ergodic invariant probability measures on $\mathcal{C}_+^I$
$\mathcal{M}^{I,\circ}$	sets of ergodic invariant probability measures on $\mathcal{C}_+^{I,\circ}$
$\partial\mathcal{M}^I$	sets of ergodic invariant probability measures on $\partial\mathcal{C}_+^I$
I_π	the subset of $\{1, \ldots, n\}$ such that $\pi(\mathcal{C}_+^{I_\pi,\circ}) = 1, \pi \in \mathcal{M}$
γ, p_0, A	constants satisfying the condition in Lemma 3.1
$\boldsymbol{\rho}$	$= (\rho_1, \ldots, \rho_n)$ vector satisfying the condition in Lemma 3.1
$V_{\boldsymbol{\rho}}(\boldsymbol{\varphi})$	$:= \Big(1 + \vec{c}^\top \vec{x}\Big) \prod_{i=1}^n x_i^{\rho_i} \exp\Big\{A_2 \int_{-r}^0 \mu(ds) \int_s^0 e^{\gamma(u-s)} h\big(\boldsymbol{\varphi}(u)\big)du\Big\}$
$V_{\vec{0}}(\boldsymbol{\varphi})$	$:= \Big(1 + \vec{c}^\top \vec{x}\Big) \exp\Big\{A_2 \int_{-r}^0 \mu(ds) \int_s^0 e^{\gamma(u-s)} h\big(\boldsymbol{\varphi}(u)\big)du\Big\}$
$\mathcal{C}_{V,M}$	$:= \{\boldsymbol{\varphi} \in \mathcal{C}_+ : A_2\gamma \int_{-r}^0 \mu(ds) \int_s^0 e^{\gamma(u-s)} h\big(\boldsymbol{\varphi}(u)\big)du \le A_0, \lvert\boldsymbol{\varphi}(0)\rvert \le M\}$
H_1	constant satisfying (3.29) and (3.36)
$\boldsymbol{\rho}^*, \kappa^*$	vector and constant satisfying (4.1)
n^*	constant satisfying $\gamma_0(n^* - 1) - A_0 > 0$
p_1	constant satisfying condition (3.4) and $p_1 > p_0$
H_1^*	constant determined in (4.12)
$R_{V,M}$	constant determined in (4.15)
ε^*	constant determined in (4.16)
$T^*, \widehat{\delta}$	constants determined in Lemma 4.3
$\mathcal{C}_V(\widehat{\delta})$	$:= \{\boldsymbol{\varphi} \in \mathcal{C}_+^\circ \cap \mathcal{C}_{V,M} \text{ and } \lvert\varphi_i(0)\rvert \le \widehat{\delta} \text{ for some } i\}$

Consider a stochastic delay Kolmogorov system

$$\begin{cases} dX_i(t) = X_i(t) f_i(\mathbf{X}_t)dt + X_i(t) g_i(\mathbf{X}_t) dE_i(t), \quad i = 1, \ldots, n, \\ \mathbf{X}_0 = \boldsymbol{\phi} \in \mathcal{C}_+, \end{cases} \tag{2.1}$$

and denote by $\mathbf{X}^{\boldsymbol{\phi}}(t)$ its solution. For convenience, we usually suppress the superscript "$\boldsymbol{\phi}$" and use $\mathbb{P}_{\boldsymbol{\phi}}$ and $\mathbb{E}_{\boldsymbol{\phi}}$ to denote the probability and expectation given the initial value $\boldsymbol{\phi}$, respectively. We also assume that the initial value is non-random. Denoted by $\{\mathcal{F}_t\}_{t\ge0}$ the filtration satisfying the usual conditions and assume that the n-dimensional Brownian motion $\mathbf{B}(t)$ is adapted to $\{\mathcal{F}_t\}_{t\ge0}$. Note that a segment process is also referred to as a memory segment function. Throughout the rest of the paper, we assume the following assumptions hold.

Assumption 2.1. The coefficients of (2.1) satisfy the following conditions.

(1) $\mathrm{diag}(g_1(\boldsymbol{\varphi}), \ldots, g_n(\boldsymbol{\varphi}))\Gamma^\top \Gamma \mathrm{diag}(g_1(\boldsymbol{\varphi}), \ldots, g_n(\boldsymbol{\varphi})) = (g_i(\boldsymbol{\varphi}) g_j(\boldsymbol{\varphi}) \sigma_{ij})_{n\times n}$ is a positive definite matrix for any $\boldsymbol{\varphi} \in \mathcal{C}_+$.
(2) $f_i(\cdot), g_i(\cdot) : \mathcal{C}_+ \to \mathbb{R}$ are Lipschitz continuous in each bounded set of $\mathcal{C}_+$ for any $i = 1, \ldots, n$.
(3) There exist $\mathbf{c} = (c_1, \ldots, c_n) \in \mathbb{R}^n$ with $c_i > 0, \forall i$, and $\gamma_b, \gamma_0 > 0$, $A_0 > 0$, $A_1 > A_2 > 0$, $M > 0$, a continuous function $h : \mathbb{R}^n \to \mathbb{R}_+$ and a probability measure μ concentrated

on $[-r, 0]$ such that for any $\boldsymbol{\varphi} \in \mathcal{C}_+$

$$\begin{aligned}\frac{\sum_{i=1}^n c_i x_i f_i(\boldsymbol{\varphi})}{1+\mathbf{c}^\top \mathbf{x}} - \frac{1}{2}\frac{\sum_{i,j=1}^n \sigma_{ij} c_i c_j x_i x_j g_i(\boldsymbol{\varphi}) g_j(\boldsymbol{\varphi})}{(1+\mathbf{c}^\top \mathbf{x})^2}& \\ + \gamma_b \sum_{i=1}^n (|f_i(\boldsymbol{\varphi})| + g_i^2(\boldsymbol{\varphi}))& \\ \le A_0 \mathbf{1}_{\{|\mathbf{x}|<M\}} - \gamma_0 - A_1 h(\mathbf{x}) + A_2 \int_{-r}^0 h\big(\boldsymbol{\varphi}(s)\big)\mu(ds),&\end{aligned} \tag{2.2}$$

where $\mathbf{x} := \boldsymbol{\varphi}(0)$. We assume without loss of generality that $h : \mathbb{R}^n \to [1, \infty)$, otherwise, we can always change γ_0 and A_1, A_2 to fulfill this requirement.

Assumption 2.2. One of following assumptions hold:

(a) There is a constant $\widetilde{K}$ such that for any $\boldsymbol{\varphi} \in \mathcal{C}_+$, $\mathbf{x} = \boldsymbol{\varphi}(0)$

$$\sum_{i=1}^n |f_i(\boldsymbol{\varphi})| + \sum_{i=1}^n g_i^2(\boldsymbol{\varphi}) \le \widetilde{K}\Big[h(\mathbf{x}) + \int_{-r}^0 h(\boldsymbol{\varphi}(s))\mu(ds)\Big]. \tag{2.3}$$

(b) There exist constants $b_1, b_2 > 0$, a function $h_1 : \mathbb{R}^n \to [1, \infty]$, and a probability measure μ_1 on $[-r, 0]$ such that for any $\boldsymbol{\varphi} \in \mathcal{C}_+$, $\mathbf{x} = \boldsymbol{\varphi}(0)$

$$b_1 h_1(\mathbf{x}) \le \sum_{i=1}^n |f_i(\boldsymbol{\varphi})| + \sum_{i=1}^n g_i^2(\boldsymbol{\varphi}) \le b_2\Big[h_1(\mathbf{x}) + \int_{-r}^0 h_1(\boldsymbol{\varphi}(s))\mu_1(ds)\Big]. \tag{2.4}$$

Remark 1.

Let us comment on the above assumptions.

- The above assumptions (and additional assumptions provided later) are not restrictive, and are easily verifiable. Such conditions are widely used in popular models in the literature; see Section 5.
- Parts (2) and (3) of Assumption 2.1 guarantee the existence and uniqueness of a strong solution to (2.1). We need part (1) of Assumption 2.1 to ensure that the solution to (2.1) is a non-degenerate diffusion. Moreover, as will be seen later that (3) implies the tightness of the family of transition probabilities associated with the solution to (2.1). One difficulty stems from the positive term $A_2 \int_{-r}^0 h\big(\boldsymbol{\varphi}(s)\big)\mu(ds)$ on the right-hand side of (2.2), which cannot be relaxed in practice.
- Assumption 2.2 plays an important role in guaranteeing the π-uniform integrability of the function $\sum_i \big(|f_i(\cdot)| + g_i^2(\cdot)\big)$, for any invariant measure π. It will become clear in Lemmas 3.4 and 3.5 as well as the remaining parts of the paper.

As was alluded to, persistence and extinction are concepts of vital importance in biology and ecology. It turns out that such concepts are features shared by all stochastic functional Kolmogorov systems. While the termination of a species in biology is referred to as extinction, the moment of extinction is generally considered to be the death of the last individual of the species. In contrast to extinction, we have the persistence of a species. To proceed, similar to [24,55,56], we define persistence and extinction as follows.

Definition 2.1. Let $\mathbf{X}(t) = (X_1(t), \ldots, X_n(t))^\top$ be the solution of (2.1). The process $\mathbf{X}$ is strongly stochastically persistent if for any $\varepsilon > 0$, there exists an $R = R(\varepsilon) > 0$ such that for any $\boldsymbol{\phi} \in \mathcal{C}_+^\circ$

$$\liminf_{t\to\infty} \mathbb{P}_{\boldsymbol{\phi}} \left\{ R^{-1} \le |X_i(t)| \le R \right\} \ge 1 - \varepsilon \text{ for all } i = 1, \dots, n. \tag{2.5}$$

Definition 2.2. With $\mathbf{X}(t)$ given in Definition 2.1, for $\boldsymbol{\phi} \in \mathcal{C}_+^\circ$ and some $i \in \{1, \dots, n\}$, we say X_i goes extinct with probability $p_{\boldsymbol{\phi}} > 0$ if

$$\mathbb{P}_{\boldsymbol{\phi}} \left\{ \lim_{t\to\infty} X_i(t) = 0 \right\} = p_{\boldsymbol{\phi}}.$$

In what follows, in the main theorems, we use the terminology "strongly stochastic persistence" for clarity. However, we will use "persistent" for "strongly stochastically persistent" and "persistence" for "strongly stochastic persistence" interchangeably in the rest of the paper for simplicity.

Let $\mathcal{M}$ be the set of ergodic invariant probability measures of $\mathbf{X}_t$ supported on the boundary $\partial\mathcal{C}_+$. Note that if we let $\boldsymbol{\delta}^*$ be the Dirac measure concentrated at $\mathbf{0}$, then $\boldsymbol{\delta}^* \in \mathcal{M}$ so that $\mathcal{M} \neq \emptyset$. 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 $\pi(\cdot) = \sum_{\nu\in\widetilde{\mathcal{M}}} p_\nu \nu(\cdot)$ with $p_\nu \ge 0$ and $\sum_{\nu\in\widetilde{\mathcal{M}}} p_\nu = 1$.

Assumption 2.3. For any $\pi \in \mathrm{Conv}(\mathcal{M})$, we have

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

where

$$\lambda_i(\pi) := \int_{\partial\mathcal{C}_+} \left(f_i(\boldsymbol{\varphi}) - \frac{\sigma_{ii} g_i^2(\boldsymbol{\varphi})}{2} \right) \pi(d\boldsymbol{\varphi}). \tag{2.6}$$

Theorem 2.1. *Assume that Assumptions 2.1–2.3 hold. The solution* $\mathbf{X}$ *of (2.1) is strongly stochastically persistent.*

It is well-recognized that the nondegeneracy of the diffusion is not sufficient to imply the strong Feller property as well as the uniqueness of an invariant probability measure of stochastic delay systems. The following assumption is needed to obtain the uniqueness of an invariant probability measure.

Assumption 2.4. The following conditions hold:

(i) There are some constants $D_0, d_0 > 0$ such that for any $\boldsymbol{\varphi}^{(1)}, \boldsymbol{\varphi}^{(2)} \in \mathcal{C}_+^\circ$, $i \in \{1, \dots, n\}$,

$$\begin{aligned} |f_i(\boldsymbol{\varphi}^{(1)}) - f_i(\boldsymbol{\varphi}^{(2)})| &\le D_0 \left|\mathbf{x}^{(1)} - \mathbf{x}^{(2)}\right| \left|1 + \mathbf{x}^{(1)} + \mathbf{x}^{(2)}\right|^{d_0} \\ &\quad + D_0 \int_{-r}^0 \left|\boldsymbol{\varphi}^{(1)}(s) - \boldsymbol{\varphi}^{(2)}(s)\right| \left|1 + \boldsymbol{\varphi}^{(1)}(s) + \boldsymbol{\varphi}^{(2)}(s)\right|^{d_0} \mu(ds), \end{aligned} \tag{2.7}$$

where $\mathbf{x}^{(1)} := \boldsymbol{\varphi}^{(1)}(0)$, $\mathbf{x}^{(2)} := \boldsymbol{\varphi}^{(2)}(0)$.

(ii) The conditions in (i) above hold with $f_i(\cdot)$ replaced by $g_i(\cdot)$ and $g_i^2(\cdot)$.

(iii) The inverse of matrix $(g_i(\boldsymbol{\varphi}) g_j(\boldsymbol{\varphi}) \sigma_{ij})_{n\times n}$ is uniformly bounded in $\mathcal{C}_+^\circ$.

Proposition 2.1. *Under Assumptions 2.1 and 2.4, the solution of Eq. (2.1) has at most one invariant probability measure on* $\mathcal{C}_+^\circ$.

Theorem 2.2. *Under Assumptions 2.1–2.4, system (2.1) has a unique invariant probability measure concentrated on $\mathcal{C}_+^\circ$.*

Remark 2. Assumption 2.3 means that all invariant measures on the boundary are repellers (because the maximum Lyapunov exponent of an invariant measure is positive), which guarantees that the solution in the interior cannot stay long near the boundary. As a result, the species coexist.

3. Preliminaries and key technical results

3.1. Existence, uniqueness, positivity, and key estimates of the solutions

To begin, we state the functional Itô formula for our processes; see [10] for more details. Let $\mathbb{D}$ be the space of càdlàg functions $\boldsymbol{\varphi} : [-r, 0] \mapsto \mathbb{R}^n$. For $\boldsymbol{\varphi} \in \mathbb{D}$, with $s \geq 0$ and $y \in \mathbb{R}^n$, we define horizontal and vertical perturbations as

$$\boldsymbol{\varphi}_s(t) = \begin{cases} \boldsymbol{\varphi}(t+s) \ \text{ if } \ t \in [-r, -s], \\ \boldsymbol{\varphi}(0) \ \text{ if } \ t \in [-s, 0], \end{cases}$$

and

$$\boldsymbol{\varphi}^y(t) = \begin{cases} \boldsymbol{\varphi}(t) \ \text{ if } \ t \in [-r, 0), \\ \boldsymbol{\varphi}(0) + y \ \text{if } t = 0, \end{cases}$$

respectively. The horizontal and vertical partial derivatives of $V : \mathbb{D} \to \mathbb{R}$ at $\boldsymbol{\varphi}$, denoted by $\partial_t V(\boldsymbol{\varphi})$, $(\partial_i V(\boldsymbol{\varphi}))_{i=1}^n$, are defined as

$$\partial_t V(\boldsymbol{\varphi}) = \lim_{s\to 0} \frac{V(\boldsymbol{\varphi}_s) - V(\boldsymbol{\varphi})}{s},$$

and

$$\partial_i V(\boldsymbol{\varphi}) = \lim_{s\to 0} \frac{V(\boldsymbol{\varphi}^{se_i}) - V(\boldsymbol{\varphi})}{s}, i = 1, \dots, n, \tag{3.1}$$

respectively, if the limits exist. In (3.1), e_i is the standard unit vector in $\mathbb{R}^n$ whose ith component is 1 and all other components are 0. Let $\mathbb{F}$ be the family of functions $V(\cdot) : \mathbb{D} \mapsto \mathbb{R}$ satisfying that

- V is continuous, that is, for any $\varepsilon > 0$, $\boldsymbol{\varphi} \in \mathbb{D}$ there is a $\delta > 0$ such that $|V(\boldsymbol{\varphi}) - V(\boldsymbol{\varphi}')| < \varepsilon$ as long as $\|\boldsymbol{\varphi} - \boldsymbol{\varphi}'\| < \delta$;
- the derivatives V_t, $V_x := (\partial_i V)$, and $V_{xx} := (\partial_{ij} V)$ exist and are continuous;
- V, V_t, $V_x = (\partial_i V)$ and $V_{xx} = (\partial_{ij} V)$ are bounded in each set $\{\boldsymbol{\varphi} \in \mathbb{D} : \|\boldsymbol{\varphi}\| \leq R\}$, $R > 0$.

Let $V(\cdot) \in \mathbb{F}$, we define the operator

$$\begin{aligned} \mathcal{L}V(\boldsymbol{\varphi}) =& \partial_t V(\boldsymbol{\varphi}) + \sum_{i=1}^n \varphi_i(0) f_i(\boldsymbol{\varphi}) \partial_i V(\boldsymbol{\varphi}) \\ &+ \frac{1}{2} \sum_{i,j=1}^n \varphi_i(0)\varphi_j(0)\sigma_{ij} g_i(\boldsymbol{\varphi}) g_j(\boldsymbol{\varphi}) \partial_{ij} V(\boldsymbol{\varphi}). \end{aligned} \tag{3.2}$$

We have the functional Itô formula (see [9,10]) as follows

$$dV(\mathbf{X}_t) = \big(\mathcal{L}V(\mathbf{X}_t)\big)dt + \sum_{i=1}^n X_i(t) g_i(\mathbf{X}_t) \partial_i V(\mathbf{X}_t) dE_i(t). \tag{3.3}$$

Lemma 3.1. *For any $\gamma < \gamma_b$ and $p_0 > 0$, $\boldsymbol{\rho} = (\rho_1, \dots, \rho_n) \in \mathbb{R}^n$ satisfying*

$$|\boldsymbol{\rho}| < \min\left\{\frac{\gamma_b}{2}, \frac{1}{n}, \frac{\gamma_b}{4\sigma^*}\right\} \text{ and } p_0 < \min\left\{1, \frac{\gamma_b}{8n\sigma^*}\right\}, \tag{3.4}$$

where $\sigma^* := \max\{\sigma_{ij} : 1 \le i, j \le n\}$, let

$$V_{\boldsymbol{\rho}}(\boldsymbol{\varphi}) := \left(1 + \mathbf{c}^\top \mathbf{x}\right) \prod_{i=1}^n x_i^{\rho_i} \exp\left\{A_2 \int_{-r}^0 \mu(ds) \int_s^0 e^{\gamma(u-s)} h\big(\boldsymbol{\varphi}(u)\big) du\right\}.$$

Then, we have

$$\begin{aligned}\mathcal{L}V_{\boldsymbol{\rho}}^{p_0}(\boldsymbol{\varphi}) \le & p_0 V_{\boldsymbol{\rho}}^{p_0}(\boldsymbol{\varphi}) \Bigg[A_0 \mathbf{1}_{\{|\mathbf{x}|<M\}} - \gamma_0 - Ah(\mathbf{x}) \\ & - A_2 \gamma \int_{-r}^0 \mu(ds) \int_s^0 e^{\gamma(u-s)} h\big(\boldsymbol{\varphi}(u)\big) du - \frac{\gamma_b}{2} \sum_{i=1}^n \Big(|f_i(\boldsymbol{\varphi})| + g_i^2(\boldsymbol{\varphi})\Big) \Bigg],\end{aligned} \tag{3.5}$$

where $\mathbf{x} := \boldsymbol{\varphi}(0)$ and A is a positive number satisfying $A < A_1 - A_2 \int_{-r}^0 e^{-\gamma s} \mu(ds)$. Recall that $\mathbf{c}$, M, A_0, A_1, A_2, γ_0, γ_b, $h(\cdot)$, and $\mu(\cdot)$ are defined in Assumption 2.1(3).

Proof. Let

$$\begin{aligned} U_{\boldsymbol{\rho}}(\boldsymbol{\varphi}) &= \ln V_{\boldsymbol{\rho}}(\boldsymbol{\varphi}) \\ &= \ln\Big(1 + \mathbf{c}^\top \mathbf{x}\Big) + \sum_{i=1}^n \rho_i \ln x_i + A_2 \int_{-r}^0 \mu(ds) \int_s^0 e^{\gamma(u-s)} h\big(\boldsymbol{\varphi}(u)\big) du. \end{aligned}$$

By [46, Remark 2.2] and direct calculation, we have

$$\begin{aligned} \partial_t U_{\boldsymbol{\rho}}(\boldsymbol{\varphi}) =& A_2 h(\mathbf{x}) \int_{-r}^0 e^{-\gamma s} \mu(ds) \\ & - A_2 \int_{-r}^0 h\big(\boldsymbol{\varphi}(s)\big) \mu(ds) - A_2 \gamma \int_{-r}^0 \mu(ds) \int_s^0 e^{\gamma(u-s)} h\big(\boldsymbol{\varphi}(u)\big) du, \\ \partial_i U_{\boldsymbol{\rho}}(\boldsymbol{\varphi}) =& \frac{c_i}{1 + \mathbf{c}^\top \mathbf{x}} + \frac{\rho_i}{x_i} \; ; \; \partial_{ij} U_{\boldsymbol{\rho}}(\boldsymbol{\varphi}) = \frac{-c_i c_j}{\Big(1 + \mathbf{c}^\top \mathbf{x}\Big)^2} + \frac{-\delta_{ij} \rho_i}{x_i^2}, \end{aligned}$$

where

$$\delta_{ij} = \begin{cases} 1 & \text{if} \quad i = j, \\ 0 & \text{otherwise.} \end{cases}$$

As a consequence, we obtain from the functional Itô formula that

$$\begin{aligned} \mathcal{L}U_{\boldsymbol{\rho}}(\boldsymbol{\varphi}) =& A_2 h(\mathbf{x}) \int_{-r}^0 e^{-\gamma s} \mu(ds) - A_2 \int_{-r}^0 h\big(\boldsymbol{\varphi}(s)\big) \mu(ds) \\ & - A_2 \gamma \int_{-r}^0 \mu(ds) \int_s^0 e^{\gamma(u-s)} h\big(\boldsymbol{\varphi}(u)\big) du \\ & + \frac{\sum_{i=1}^n c_i x_i f_i(\boldsymbol{\varphi})}{1 + \mathbf{c}^\top \mathbf{x}} - \frac{1}{2} \sum_{i,j=1}^n \frac{c_i c_j \sigma_{ij} x_i x_j g_i(\boldsymbol{\varphi}) g_j(\boldsymbol{\varphi})}{\Big(1 + \mathbf{c}^\top \mathbf{x}\Big)^2} \\ & + \sum_{i=1}^n \rho_i \Big(f_i(\boldsymbol{\varphi}) - \sigma_{ii} g_i^2(\boldsymbol{\varphi})\Big). \end{aligned} \tag{3.6}$$

Therefore, by the fact $V_{\boldsymbol{\rho}}^{p_0}(\boldsymbol{\varphi}) = e^{p_0 U_\rho(\varphi)}$ and an application of the functional Itô formula, we get

$$\begin{aligned}\mathcal{L}V_{\boldsymbol{\rho}}^{p_0}(\boldsymbol{\varphi}) =& p_0 V_{\boldsymbol{\rho}}^{p_0}(\boldsymbol{\varphi}) \Bigg(A_2 h(\mathbf{x}) \int_{-r}^{0} e^{-\gamma s}\mu(ds) - A_2 \int_{-r}^{0} h\big(\boldsymbol{\varphi}(s)\big)\mu(ds) \\ & - A_2\gamma \int_{-r}^{0} \mu(ds) \int_{s}^{0} e^{\gamma(u-s)} h\big(\boldsymbol{\varphi}(u)\big)du \\ & + \frac{\sum_{i=1}^n c_i x_i f_i(\boldsymbol{\varphi})}{1+\mathbf{c}^\top \mathbf{x}} - \frac{1}{2}\sum_{i,j=1}^n \frac{c_i c_j \sigma_{ij} x_i x_j g_i(\boldsymbol{\varphi}) g_j(\boldsymbol{\varphi})}{\big(1+\mathbf{c}^\top\mathbf{x}\big)^2} \\ & + \sum_{i=1}^n \rho_i \Big(f_i(\boldsymbol{\varphi}) - \sigma_{ii} g_i^2(\boldsymbol{\varphi})\Big) \\ & + \frac{1}{2} p_0 \sum_{i,j=1}^n \Big(\frac{c_i x_i}{1+\mathbf{c}^\top\mathbf{x}} + \rho_i\Big)\Big(\frac{c_j x_j}{1+\mathbf{c}^\top\mathbf{x}} + \rho_j\Big)\sigma_{ij} g_i(\boldsymbol{\varphi}) g_j(\boldsymbol{\varphi}) \Bigg).\end{aligned}$$

Since

$$\begin{aligned}&\frac{1}{2} p_0 \sum_{i,j=1}^n \Big(\frac{c_i x_i}{1+\mathbf{c}^\top\mathbf{x}} + \rho_i\Big)\Big(\frac{c_j x_j}{1+\mathbf{c}^\top\mathbf{x}} + \rho_j\Big)\sigma_{ij} g_i(\boldsymbol{\varphi}) g_j(\boldsymbol{\varphi}) \\ &\qquad \le \frac{1}{4} p_0 \sum_{i,j=1}^n (1+\rho_i)(1+\rho_j)\sigma_{ij}\Big(g_i^2(\boldsymbol{\varphi}) + g_j^2(\boldsymbol{\varphi})\Big) \\ &\qquad \le 2p_0 n\sigma^* \sum_{i=1}^n g_i^2(\boldsymbol{\varphi}),\end{aligned}$$

and $|\rho_i| < \frac{\gamma_b}{2}$; $|\rho_i|\sigma^* + 2p_0 n\sigma^* < \frac{\gamma_b}{2}$ $\forall i = 1,\dots,n$, using Assumption 2.1, we have

$$\begin{aligned}\mathcal{L}V_{\boldsymbol{\rho}}^{p_0}(\boldsymbol{\varphi}) \le& p_0 V_{\boldsymbol{\rho}}^{p_0}(\boldsymbol{\varphi}) \Bigg(A_0 \mathbf{1}_{\{|\mathbf{x}|<M\}} - \gamma_0 - h(\mathbf{x})\Big(A_1 - A_2\int_{-r}^0 e^{-\gamma s}\mu(ds)\Big) \\ & - A_2\gamma \int_{-r}^{0} \mu(ds) \int_{s}^{0} e^{\gamma(u-s)} h\big(\boldsymbol{\varphi}(u)\big)du - \frac{\gamma_b}{2}\sum_{i=1}^n \Big(|f_i(\boldsymbol{\varphi})| + g_i^2(\boldsymbol{\varphi})\Big)\Bigg) \\ \le& p_0 V_{\boldsymbol{\rho}}^{p_0}(\boldsymbol{\varphi}) \Bigg(A_0 \mathbf{1}_{\{|\mathbf{x}|<M\}} - \gamma_0 - A h(\mathbf{x}) \\ & - A_2\gamma \int_{-r}^{0} \mu(ds) \int_{s}^{0} e^{\gamma(u-s)} h\big(\boldsymbol{\varphi}(u)\big)du - \frac{\gamma_b}{2}\sum_{i=1}^n \Big(|f_i(\boldsymbol{\varphi})| + g_i^2(\boldsymbol{\varphi})\Big)\Bigg).\end{aligned}$$

The proof is complete. □

Theorem 3.1. *For any initial condition $\boldsymbol{\phi} \in \mathcal{C}_+$, there exists a unique global solution of (2.1). It remains in $\mathcal{C}_+$ (resp., $\mathcal{C}_+^\circ$), provided $\boldsymbol{\phi} \in \mathcal{C}_+$ (resp., $\boldsymbol{\phi} \in \mathcal{C}_+^\circ$). Moreover, for any $p_0, \boldsymbol{\rho}$ satisfying condition (3.4), we have*

$$\mathbb{E}_{\boldsymbol{\phi}} V_{\boldsymbol{\rho}}^{p_0}(\mathbf{X}_t) \le V_{\boldsymbol{\rho}}^{p_0}(\boldsymbol{\phi}) e^{A_0 p_0 t}. \tag{3.7}$$

In addition, if $\rho_i \ge 0, \forall i$, then

$$\mathbb{E}_{\boldsymbol{\phi}} V_{\boldsymbol{\rho}}^{p_0}(\mathbf{X}_t) \le V_{\boldsymbol{\rho}}^{p_0}(\boldsymbol{\phi}) e^{-\gamma_0 p_0 t} + \overline{M}_{p_0,\boldsymbol{\rho}}, \tag{3.8}$$

where

$$\overline{M}_{p_0,\boldsymbol{\rho}} := \frac{A_0}{\gamma_0} \sup_{\boldsymbol{\varphi}\in\mathcal{C}_{V,M}} V_{\boldsymbol{\rho}}^{p_0}(\boldsymbol{\varphi}) < \infty \textit{ provided } \rho_i \geq 0\, \forall i,$$

and $\mathcal{C}_{V,M} = \{\boldsymbol{\varphi} \in \mathcal{C}_+ : A_2\gamma \int_{-r}^0 \mu(ds) \int_s^0 e^{\gamma(u-s)} h\big(\boldsymbol{\varphi}(u)\big)du \leq A_0 \textit{ and } |\mathbf{x}| \leq M\}$.

Proof. We prove the existence and uniqueness of the solution with initial value $\boldsymbol{\phi} \in \mathcal{C}_+^\circ$. The other cases can be handled similarly. Let $\boldsymbol{\rho}^{(1)} = (\rho_1^{(1)}, \dots, \rho_n^{(1)}) \in \mathbb{R}^n$ with $\rho_i^{(1)} < 0\ \forall i = 1, \dots, n$ satisfy the conditions (3.4). We define the following stopping times $\tau_k^{(1)} = \inf\{t \geq 0 : V_{\boldsymbol{\rho}^{(1)}}^{p_0}(\mathbf{X}_t) \geq k\}$ and $\tau_\infty^{(1)} = \lim_{k\to\infty} \tau_k^{(1)}$. It is easily seen that

$$\begin{aligned}\lim_{m\to\infty} \inf\Big\{V_{\boldsymbol{\rho}^{(1)}}^{p_0}(\boldsymbol{\varphi}) : x_i \vee x_i^{-1} > m \text{ for some } i \in \{1,\dots,n\}, \mathbf{x} := \boldsymbol{\varphi}(0), \boldsymbol{\varphi} \in \mathcal{C}_+^\circ\Big\} \\ = \infty.\end{aligned} \tag{3.9}$$

The existence and uniqueness of local solutions can be seen in [42] due to the local Lipschitz continuity of the coefficients. To prove the solution is global and remains in $\mathcal{C}_+^\circ$, because of (3.9), it is sufficient to prove that $\tau_\infty^{(1)} = \infty$ a.s. We obtain from (3.5) that

$$\mathcal{L}V_{\boldsymbol{\rho}^{(1)}}^{p_0}(\boldsymbol{\varphi}) \leq A_0 p_0 V_{\boldsymbol{\rho}^{(1)}}^{p_0}(\boldsymbol{\varphi}),\ \ \forall \boldsymbol{\varphi} \in \mathcal{C}_+^\circ.$$

Hence, by the functional Itô formula, we get

$$\begin{aligned}\mathbb{E}_{\boldsymbol{\phi}} V_{\boldsymbol{\rho}^{(1)}}^{p_0}(\mathbf{X}_{t\wedge\tau_k^{(1)}}) &= V_{\boldsymbol{\rho}^{(1)}}^{p_0}(\boldsymbol{\phi}) + \mathbb{E}_{\boldsymbol{\phi}} \int_0^{t\wedge\tau_k^{(1)}} \mathcal{L}V_{\boldsymbol{\rho}^{(1)}}^{p_0}(X_s)ds \\ &\leq V_{\boldsymbol{\rho}^{(1)}}^{p_0}(\boldsymbol{\phi}) + p_0 A_0 \int_0^t \mathbb{E}_{\boldsymbol{\phi}} V_{\boldsymbol{\rho}^{(1)}}^{p_0}(\mathbf{X}_{s\wedge\tau_k^{(1)}})ds.\end{aligned}$$

Combined with Gronwall's inequality yields that

$$\mathbb{E}_{\boldsymbol{\phi}} V_{\boldsymbol{\rho}^{(1)}}^{p_0}(\mathbf{X}_{t\wedge\tau_k^{(1)}}) \leq V_{\boldsymbol{\rho}^{(1)}}^{p_0}(\boldsymbol{\phi}) e^{p_0 A_0 t},\ \ \forall t \geq 0. \tag{3.10}$$

As a consequence,

$$\mathbb{P}_{\boldsymbol{\phi}}\Big\{V_{\boldsymbol{\rho}^{(1)}}^{p_0}(\mathbf{X}_{t\wedge\tau_k^{(1)}}) \geq k\Big\} \leq \frac{V_{\boldsymbol{\rho}^{(1)}}^{p_0}(\boldsymbol{\phi}) e^{p_0 A_0 t}}{k} \to 0 \text{ as } k \to \infty,$$

which forces $\tau_\infty^{(1)} > t$ a.s. for any $t > 0$ and hence, $\tau_\infty^{(1)} = \infty$ a.s.

Next, we consider the second part. For any p_0, $\boldsymbol{\rho}$ satisfying (3.4), by applying (3.5), one has

$$\mathcal{L}V_{\boldsymbol{\rho}}^{p_0}(\boldsymbol{\varphi}) \leq A_0 p_0 V_{\boldsymbol{\rho}}^{p_0}(\boldsymbol{\varphi}) \text{ for all } \boldsymbol{\varphi} \in \mathcal{C}_+^\circ.$$

Thus, from (3.10), we get

$$\mathbb{E}_{\boldsymbol{\phi}} V_{\boldsymbol{\rho}}^{p_0}(\mathbf{X}_t) \leq V_{\boldsymbol{\rho}}^{p_0}(\boldsymbol{\phi}) e^{A_0 p_0 t}.$$

If $\rho_i \geq 0\ \forall i$, a consequence of (3.5) is

$$\mathcal{L}V_{\boldsymbol{\rho}}^{p_0}(\boldsymbol{\varphi}) \leq \gamma_0 p_0 \overline{M}_{p_0,\boldsymbol{\rho}} - \gamma_0 p_0 V_{\boldsymbol{\rho}}^{p_0}(\boldsymbol{\varphi}). \tag{3.11}$$

In (3.11), we have used the fact

$$A_0 \mathbf{1}_{\{|\mathbf{x}|<M\}} - A_2\gamma \int_{-r}^0 \mu(ds) \int_s^0 e^{\gamma(u-s)} h\big(\boldsymbol{\varphi}(u)\big)du \leq 0,\ \text{ if } \boldsymbol{\varphi} \notin \mathcal{C}_{V,M}.$$

By a standard argument (see e.g., [41, Theorem 5.2, p. 157]), we can obtain (3.8) from (3.11). The proof is complete. □

Lemma 3.2. *For any $R_1 > 0$, $T > r$, and $\varepsilon > 0$, there exists an $R_2 > 0$ such that*

$$\mathbb{P}_{\boldsymbol{\phi}}\Big\{\|\mathbf{X}_t\| \le R_2,\ \forall t \in [r, T]\Big\} > 1 - \varepsilon,$$

for any initial data $\boldsymbol{\phi}$ satisfying $V_{\mathbf{0}}(\boldsymbol{\phi}) < R_1$, where $V_{\mathbf{0}}$ is defined as in Lemma 3.1 corresponding to $\boldsymbol{\rho} = \mathbf{0} = (0, \dots, 0)$.

Proof. As the proof of Theorem 3.1, we define the following stopping times $\tau_k^{(2)} = \inf\{t \ge 0 : V_{\mathbf{0}}^{p_0}(\mathbf{X}_t) \ge k\}$ and $\tau_\infty^{(2)} = \lim_{k\to\infty} \tau_k^{(2)}$. Analogous (3.10), we obtain

$$\mathbb{E}_{\boldsymbol{\phi}} V_{\mathbf{0}}^{p_0}(\mathbf{X}_{t\wedge\tau_k^{(2)}}) \le V_{\mathbf{0}}^{p_0}(\boldsymbol{\phi}) e^{p_0 A_0 t},\ \forall t \ge 0.$$

Therefore, for any $R_1, T, \varepsilon > 0$, and initial condition $\boldsymbol{\phi}$ satisfying $V_{\mathbf{0}}(\boldsymbol{\phi}) < R_1$, there exists a finite constant k_0 such that

$$\frac{V_{\mathbf{0}}^{p_0}(\boldsymbol{\phi}) e^{p_0 A_0 T}}{k_0} < \varepsilon,$$

and

$$\mathbb{P}_{\boldsymbol{\phi}}\Big\{V_{\mathbf{0}}^{p_0}(\mathbf{X}_{T\wedge\tau_{k_0}^{(2)}}) \ge k_0\Big\} \le \frac{V_{\mathbf{0}}^{p_0}(\boldsymbol{\phi}) e^{p_0 A_0 T}}{k_0} < \varepsilon.$$

That means $\mathbb{P}_{\boldsymbol{\phi}}\{\tau_{k_0}^{(2)} \ge T\} > 1 - \varepsilon$ or

$$\mathbb{P}_{\boldsymbol{\phi}}\big\{V_{\mathbf{0}}^{p_0}(\mathbf{X}_t) \le k_0\ \forall t \in [0, T]\big\} > 1 - \varepsilon.$$

Note that $V_{\mathbf{0}}^{p_0}(\mathbf{X}_t) \ge 1 + \sum_{i=1}^n c_i X_i(t)$ and $c_i > 0\ \forall i = 1, \dots, n$. Therefore, it is easily seen that there exists a finite constant R_2 satisfying

$$\mathbb{P}_{\boldsymbol{\phi}}\Big\{\|\mathbf{X}_t\| \le R_2,\ \forall t \in [r, T]\Big\} > 1 - \varepsilon. \quad \square$$

Lemma 3.3. *There is a sufficiently small $\alpha > 0$ such that for any $R > 0$ and $\varepsilon > 0$, there exists $R_3 = R_3(R, \varepsilon) > 0$ satisfying*

$$\text{if } \|\boldsymbol{\phi}\| \le R \text{ then } \mathbb{P}_{\boldsymbol{\phi}}\left\{\|\mathbf{X}_t\|_{2\alpha} \le R_3\ \forall t \in [r, 3r]\right\} \ge 1 - \frac{\varepsilon}{2}. \tag{3.12}$$

As a consequence, for any $R > 0$ and $\varepsilon > 0$, there exists an $R_4 = R_4(\varepsilon, R) > 0$ satisfying that

$$\text{if } V_{\mathbf{0}}(\boldsymbol{\phi}) \le R \text{ then } \mathbb{P}_{\boldsymbol{\phi}}\left\{\|\mathbf{X}_t\|_{2\alpha} \le R_4\ \forall t \in [2r, 3r]\right\} \ge 1 - \varepsilon. \tag{3.13}$$

Proof. For any R and $\varepsilon > 0$, by slightly modifying the proof of Lemma 3.2, there exists an $\tilde{R} > 0$ depending only on R such that

$$\mathbb{P}_{\boldsymbol{\phi}}\{\|\mathbf{X}_t\| \le \tilde{R},\ \text{for all } t \in [0, 3r]\} \ge 1 - \frac{\varepsilon}{4} \text{ if } \|\boldsymbol{\phi}\| \le R. \tag{3.14}$$

Denote by $f_i^{\tilde{R}}(\cdot)$ and $g_i^{\tilde{R}}(\cdot)$ the truncated functions, where

$$f_i^{\tilde{R}}(\boldsymbol{\varphi}) = \begin{cases} f_i(\boldsymbol{\varphi}) \text{ if } \|\boldsymbol{\varphi}\| < \tilde{R}, \\ f_i\left(\dfrac{R_1\boldsymbol{\varphi}}{\|\boldsymbol{\varphi}\|}\right) \text{ otherwise,} \end{cases}$$

and $g_i^{\tilde{R}}(\cdot)$ is defined similarly. Then $f_i^{\tilde{R}}(\cdot)$ and $g_i^{\tilde{R}}(\cdot)$ are globally Lipschitz and bounded. Let $\widetilde{\mathbf{X}}(t) = \big(\widetilde{X}_1(t), \dots, \widetilde{X}_n(t)\big)$ be the solution of (2.1) when we replace $f_i(\cdot)$ and $g_i(\cdot)$ by $f_i^{\tilde{R}}(\cdot)$

and $g_i^{\tilde{R}}(\cdot)$, respectively. By a standard argument, it is easy to obtain that

$$\mathbb{E}_{\boldsymbol{\phi}} \left|\widetilde{X}_i(t)\right|^4 \le \tilde{K} \ \forall 0 \le t \le 3r, \|\boldsymbol{\phi}\| \le R$$

where $\tilde{K}$ is a constant depending only on R and $\tilde{R}$. On the other hand, by Burkholder's inequality we have that $\forall 0 \le s \le t \le 3r$, $\|\boldsymbol{\phi}\| \le R$,

$$\mathbb{E}_{\boldsymbol{\phi}} \left|\widetilde{X}_i(t) - \widetilde{X}_i(s)\right|^4 \le C_1 \mathbb{E}_{\boldsymbol{\phi}} \left|\int_s^t \widetilde{X}_i(y)dy\right|^4 + C_1 \mathbb{E}_{\boldsymbol{\phi}} \left(\int_s^t |\widetilde{X}_i(y)|^2 dy\right)^2,$$

where C_1 depends only on T, R, and $\tilde{R}$. Hence, by Hölder's inequality, we obtain for $0 \le s \le t \le 3r$, $\|\boldsymbol{\phi}\| \le R$ that

$$\mathbb{E}_{\boldsymbol{\phi}} \left|\widetilde{X}_i(t) - \widetilde{X}_i(s)\right|^4 \le 2C_1(t-s)^2 \mathbb{E}_{\boldsymbol{\phi}} \int_0^s \left|\widetilde{X}_i(y)\right|^4 dy \le C_2(t-s)^2,$$

where C_2 is a constant depending only on R and $\tilde{R}$. As a consequence of the Kolmogorov–Chentsov theorem, $\{\widetilde{\mathbf{X}}(t) : 0 \le t \le 3r\}$ has Hölder-continuous sample paths with an exponent $2\alpha \in (0, \frac{1}{2})$. Moreover, there is a $R_3 = R_3(R, \varepsilon)$ satisfying

$$\mathbb{P}_{\boldsymbol{\phi}} \left\{ \sup_{0\le t\le 3r} |\widetilde{\mathbf{X}}(t)| + \sup_{0\le s\le t\le 3r} \frac{|\widetilde{\mathbf{X}}(t) - \widetilde{\mathbf{X}}(s)|}{(t-s)^{2\alpha}} \le R_3 \right\} \ge 1 - \frac{\varepsilon}{4}, \|\boldsymbol{\phi}\| \le R,$$

which implies

$$\mathbb{P}_{\boldsymbol{\phi}} \left\{ \|\widetilde{\mathbf{X}}_t\|_{2\alpha} \le R_3 \ \forall t \in [r, 3r] \right\} \ge 1 - \frac{\varepsilon}{4}, \|\boldsymbol{\phi}\| \le R. \tag{3.15}$$

Combining (3.14) and (3.15) implies that

$$\mathbb{P}_{\boldsymbol{\phi}} \{\|\mathbf{X}_t\|_{2\alpha} \le R_3 \ \forall t \in [r, 3r]\} \ge 1 - \frac{\varepsilon}{2}, \text{ provided } \|\boldsymbol{\phi}\| < R,$$

and the first part of the proposition is proved.

Now, we consider the second part. By Lemma 3.2, there is an $R_5 = R_5(\varepsilon, R)$ such that

$$\mathbb{P}_{\boldsymbol{\phi}}\{\|\mathbf{X}_t\| < R_5 \ \forall t \in [r, 3r]\} \ge 1 - \frac{\varepsilon}{2} \text{ if } V_{\mathbf{0}}(\boldsymbol{\phi}) < R. \tag{3.16}$$

Hence, the second conclusion follows from the first part, (3.16) and the Markov property of $(\mathbf{X}_t)$. □

Proposition 3.1. *The following results hold.*

(i) *Let $\rho_1^{(3)}$ be a fixed constant satisfying $0 < \rho_1^{(3)} < \min\left\{\frac{\gamma_b}{2}, \frac{1}{n}, \frac{\gamma_b}{4\sigma^*}\right\}$. For any $T > r$ and $m > 0$ there exists a finite constant $K_{m,T}$ such that*

$$\mathbb{E}_{\boldsymbol{\phi}} \left\|X_{i,t}\right\|^{p_0\rho_1^{(3)}} \le K_{m,T}\phi_i^{p_0\rho_1^{(3)}}(0), \ \forall t \in [r, T], i = 1, \dots, n,$$

given that

$$|\boldsymbol{\phi}(0)| + \int_{-r}^0 \mu(ds) \int_s^0 e^{\gamma(u-s)} h\big(\boldsymbol{\phi}(u)\big)du < m,$$

where $\mathbf{X}_t =: (X_{1,t}, \dots, X_{n,t})$ and $\boldsymbol{\phi} =: (\phi_1, \dots, \phi_n)$ is the initial value.

(ii) *For any $T > r$, $\varepsilon > 0$, $R > 0$, there exists an $\varepsilon_1 > 0$ such that*

$$\mathbb{P}\left\{ \left\|\mathbf{X}_T^{\boldsymbol{\phi}_1} - \mathbf{X}_T^{\boldsymbol{\phi}_2}\right\| \le \varepsilon \right\} \ge 1 - \varepsilon \text{ whenever } V_{\mathbf{0}}(\boldsymbol{\phi}_i) < R, \left\|\boldsymbol{\phi}_1 - \boldsymbol{\phi}_2\right\| \le \varepsilon_1. \tag{3.17}$$

Moreover, the solution $(\mathbf{X}_t)$ has the Feller property in $\mathcal{C}_+$.

Proof. Let $\boldsymbol{\rho}^{(3)} = (\rho_1^{(3)}, 0, \ldots, 0)$. Then $\boldsymbol{\rho}^{(3)}$ satisfies condition (3.4). By the functional Itô formula, we obtain

$$\begin{aligned} V_{\boldsymbol{\rho}^{(3)}}^{\frac{p_0}{2}}(\mathbf{X}_t) &= V_{\boldsymbol{\rho}^{(3)}}^{\frac{p_0}{2}}(\boldsymbol{\phi}) + \int_0^t \mathcal{L} V_{\boldsymbol{\rho}^{(3)}}^{\frac{p_0}{2}}(\mathbf{X}_s)ds \\ &+ \int_0^t \frac{p_0}{2} V_{\boldsymbol{\rho}^{(3)}}^{\frac{p_0}{2}}(\mathbf{X}_s) \sum_{i=1}^n \left(\frac{c_i X_i(s)}{1 + \sum_{i'=1}^n c_{i'} X_{i'}(s)} + \delta_{1i}\rho_1^{(3)} \right) g_i(\mathbf{X}_s) dE_i(s), \end{aligned} \tag{3.18}$$

where $\delta_{1i} = 1$ if $i = 1$ and otherwise, $\delta_{1i} = 0$. Therefore, combining with (3.5) leads to that

$$\begin{aligned} V_{\boldsymbol{\rho}^{(3)}}^{\frac{p_0}{2}}(\mathbf{X}_t) \le & V_{\boldsymbol{\rho}^{(3)}}^{\frac{p_0}{2}}(\boldsymbol{\phi}) + A_0 p_0 \int_0^t V_{\boldsymbol{\rho}^{(3)}}^{\frac{p_0}{2}}(\mathbf{X}_s)ds \\ &+ \int_0^t \frac{p_0}{2} V_{\boldsymbol{\rho}^{(3)}}^{\frac{p_0}{2}}(\mathbf{X}_s) \sum_{i=1}^n \left(\frac{c_i X_i(s)}{1 + \sum_{i'=1}^n c_{i'} X_{i'}(s)} + \delta_{1i}\rho_1^{(3)} \right) g_i(\mathbf{X}_s) dE_i(s). \end{aligned} \tag{3.19}$$

In the estimates to follow, in fact we need the terms in (3.20) to be finite, which can be done by first using estimates for the solution at stopping time $t \wedge \tau_k$ with τ_k being the first time such that $|g(\mathbf{X}_s)| \vee V_{\boldsymbol{\rho}^{(3)}}(\mathbf{X}_s) > k$, and letting $k \to \infty$. Since it is a standard argument, we omit it for brevity. We obtain from (3.19) that

$$\begin{aligned} &\mathbb{E}_{\boldsymbol{\phi}} \sup_{t\in[0,T]} \left[V_{\boldsymbol{\rho}^{(3)}}^{\frac{p_0}{2}}(\mathbf{X}_t) \right]^2 \le C^{(1)} V_{\boldsymbol{\rho}^{(3)}}^{p_0}(\boldsymbol{\phi}) + C^{(1)} \mathbb{E}_{\boldsymbol{\phi}} \int_0^T \sup_{s'\in[0,s]} \left[V_{\boldsymbol{\rho}^{(3)}}^{\frac{p_0}{2}}(\mathbf{X}_{s'}) \right]^2 ds \\ &+ C^{(1)} \mathbb{E}_{\boldsymbol{\phi}} \sup_{t\in[0,T]} \left| \int_0^t V_{\boldsymbol{\rho}^{(3)}}^{\frac{p_0}{2}}(\mathbf{X}_s) \sum_{i=1}^n \left(\frac{c_i X_i(s)}{1 + \sum_{i'=1}^n c_{i'} X_{i'}(s)} + \delta_{1i}\rho_1^{(3)} \right) g_i(\mathbf{X}_s) dE_i(s) \right|^2, \end{aligned} \tag{3.20}$$

where $C^{(1)}$ is a constant, independent of $\boldsymbol{\phi}$. The Burkholder–Davis–Gundy inequality and the Hölder inequality imply that

$$\begin{aligned} &\mathbb{E}_{\boldsymbol{\phi}} \sup_{t\in[0,T]} \left| \int_0^t V_{\boldsymbol{\rho}^{(3)}}^{\frac{p_0}{2}}(\mathbf{X}_s) \sum_{i=1}^n \left(\frac{c_i X_i(s)}{1 + \sum_{i'=1}^n c_{i'} X_{i'}(s)} + \delta_{1i}\rho_1^{(3)} \right) g_i(\mathbf{X}_s) dE_i(s) \right|^2 \\ &\qquad \le 16 n \sigma^* \mathbb{E}_{\boldsymbol{\phi}} \int_0^T V_{\boldsymbol{\rho}^{(3)}}^{p_0}(\mathbf{X}_s) \sum_{i=1}^n g_i^2(\mathbf{X}_s) ds, \end{aligned} \tag{3.21}$$

for a constant $C_p^{(2)}$, independent of $\boldsymbol{\phi}$. In the display above, we have used

$$\begin{aligned} &\sum_{i,j=1}^n \left(\frac{c_i X_i(s)}{1 + \sum_{i'=1}^n c_{i'} X_{i'}(s)} + 1 \right) \left(\frac{c_j X_j(s)}{1 + \sum_{i'=1}^n c_{i'} X_{i'}(s)} + 1 \right) \sigma_{ij} g_i(\mathbf{X}_s) g_j(\mathbf{X}_s) \\ &\qquad \le 4 n \sigma^* \sum_{i=1}^n g_i^2(\mathbf{X}_s). \end{aligned}$$

It follows from (3.20) and (3.21) that

$$\begin{aligned}\mathbb{E}_{\boldsymbol{\phi}} \sup_{t\in[0,T]} V^{p_0}_{\boldsymbol{\rho}^{(3)}}(\mathbf{X}_t) \leq & C^{(1)} V^{p_0}_{\boldsymbol{\rho}^{(3)}}(\boldsymbol{\phi}) + C^{(1)}\mathbb{E}_{\boldsymbol{\phi}} \int_0^T \sup_{s'\in[0,s]} V^{p_0}_{\boldsymbol{\rho}^{(3)}}(\mathbf{X}_{s'})ds \\ & + 16n\sigma^* C^{(1)}\mathbb{E}_{\boldsymbol{\phi}} \int_0^T V^{p_0}_{\boldsymbol{\rho}^{(3)}}(\mathbf{X}_s)\sum_{i=1}^n g_i^2(\mathbf{X}_s)ds.\end{aligned} \tag{3.22}$$

On the other hand, by (3.7), we get

$$\mathbb{E}_{\boldsymbol{\phi}} V^{p_0}_{\boldsymbol{\rho}^{(3)}}(\mathbf{X}_t) \leq V^{p_0}_{\boldsymbol{\rho}^{(3)}}(\boldsymbol{\phi})e^{p_0A_0t}, \ \forall t \geq 0.$$

Therefore, we obtain from the functional Itô formula and (3.5) that

$$\begin{aligned} 0 \leq \mathbb{E}_{\boldsymbol{\phi}} V^{p_0}_{\boldsymbol{\rho}^{(3)}}(\mathbf{X}_T) &= V^{p_0}_{\boldsymbol{\rho}^{(3)}}(\boldsymbol{\phi}) + \mathbb{E}_{\boldsymbol{\phi}} \int_0^T \mathcal{L} V^{p_0}_{\boldsymbol{\rho}^{(3)}}(\mathbf{X}_s)ds \\ &\leq V^{p_0}_{\boldsymbol{\rho}^{(3)}}(\boldsymbol{\phi}) + \mathbb{E}_{\boldsymbol{\phi}} \int_0^T \left(p_0A_0V^{p_0}_{\boldsymbol{\rho}}(\mathbf{X}_s) - \frac{\gamma_b}{2}V^{p_0}_{\boldsymbol{\rho}}(\mathbf{X}_s)\sum_{i=1}^n g_i^2(\mathbf{X}_s)\right)ds \\ &\leq K_T^{(1)} V^{p_0}_{\boldsymbol{\rho}^{(3)}}(\boldsymbol{\phi}) - \frac{\gamma_b}{2}\mathbb{E}_{\boldsymbol{\phi}} \int_0^T V^{p_0}_{\boldsymbol{\rho}^{(3)}}(\mathbf{X}_s)\sum_{i=1}^n g_i^2(\mathbf{X}_s)ds, \end{aligned}$$

where $K_T^{(1)}$ is a finite constant depending only on T. It follows that

$$\mathbb{E}_{\boldsymbol{\phi}} \int_0^T V^{p_0}_{\boldsymbol{\rho}^{(3)}}(\mathbf{X}_s)\sum_{i=1}^n g_i^2(\mathbf{X}_s)ds \leq K_T^{(2)} V^{p_0}_{\boldsymbol{\rho}^{(3)}}(\boldsymbol{\phi}), \text{ for some constant } K_T^{(2)}. \tag{3.23}$$

Combining (3.22) and (3.23) yields that

$$\mathbb{E}_{\boldsymbol{\phi}} \sup_{t\in[0,T]} V^{p_0}_{\boldsymbol{\rho}^{(3)}}(\mathbf{X}_t) \leq K_T^{(3)} V^{p_0}_{\boldsymbol{\rho}^{(3)}}(\boldsymbol{\phi}) + K_T^{(3)}\mathbb{E}_{\boldsymbol{\phi}} \int_0^T \sup_{s'\in[0,s]} V^{p_0}_{\boldsymbol{\rho}^{(3)}}(\mathbf{X}_{s'})ds, \tag{3.24}$$

for some constant $K_T^{(3)}$ independent of $\boldsymbol{\phi}$. It is clear that there exists $K_{m,T}^{(4)}$ such that

$$V^{p_0}_{\boldsymbol{\rho}^{(3)}}(\boldsymbol{\phi}) \leq K_{m,T}^{(4)} \phi_1^{p_0\rho_1^{(3)}}(0), \tag{3.25}$$

given that

$$|\boldsymbol{\phi}(0)| + \int_{-r}^0 \mu(ds)\int_s^0 e^{\gamma(u-s)}h\big(\boldsymbol{\phi}(u)\big)du < m.$$

Combining (3.24), (3.25), and Gronwall's inequality, we have that

$$\mathbb{E}_{\boldsymbol{\phi}} \sup_{t\in[0,T]} V^{p_0}_{\boldsymbol{\rho}^{(3)}}(\mathbf{X}_t) \leq K_{m,T}^{(5)} \phi_1^{p_0\rho_1^{(3)}}(0), \tag{3.26}$$

where $K_{m,T}^{(5)}$ is a finite constant independent of $\boldsymbol{\phi}$. Note that

$$V^{p_0}_{\boldsymbol{\rho}^{(3)}}(\mathbf{X}_t) \geq X_1^{p_0\rho_1^{(3)}}(t). \tag{3.27}$$

It follows from (3.26) and (3.27) that

$$\mathbb{E}_{\boldsymbol{\phi}} \|X_{1t}\|^{p_0\rho_1^{(3)}} \leq K_{m,T}^{(5)} \phi_1^{p_0\rho_1^{(3)}}(0), \ \forall t \in [r, T].$$

Hence, by a similar argument, we obtain

$$\mathbb{E}_{\boldsymbol{\phi}} \|X_{it}\|^{p_0\rho_1^{(3)}} \le K_{m,T}\phi_i^{p_0\rho_1^{(3)}}(0),\ \forall t \in [r,T], i = 1,\dots,n,$$

for some constant $K_{m,T}$ depending only on m, T. As a result, the first part of the Theorem is proved.

Because our coefficients are Lipschitz continuous in each bounded set of $\mathcal{C}_+$, by using (3.7) and the truncation argument, the second conclusion is obtained (it is similar to the proof of Lemma 3.3). In addition, the Feller property can be obtained by slightly modifying the proof in [41, Lemma 2.9.4 and Theorem 2.9.3]. □

3.2. Tightness, weak convergence of occupation measures, and uniform integrability

Let $\boldsymbol{\rho} = \mathbf{0}$. We obtain from (3.5) that for all $\boldsymbol{\varphi} \in \mathcal{C}_+$, $\mathbf{x} := \boldsymbol{\varphi}(0)$,

$$\mathcal{L}V_{\mathbf{0}}^{p_0}(\boldsymbol{\varphi}) \le \gamma_0 p_0 \overline{M}_{p_0,\mathbf{0}} - Ap_0V_{\mathbf{0}}^{p_0}(\boldsymbol{\varphi})h(\mathbf{x}),$$

where $\overline{M}_{p_0,\mathbf{0}}$ is defined as in Theorem 3.1. Hence, by the functional Itô formula, we have

$$\mathbb{E}_{\boldsymbol{\phi}} V_{\mathbf{0}}^{p_0}(\mathbf{X}_t) \le V_{\mathbf{0}}^{p_0}(\boldsymbol{\phi}) + \gamma_0 p_0 \overline{M}_{p_0,\mathbf{0}}t - \mathbb{E}_{\boldsymbol{\phi}}\int_0^t Ap_0V_{\mathbf{0}}^{p_0}(\mathbf{X}_s)h(\mathbf{X}(s))ds.$$

Since $V_{\mathbf{0}}(\boldsymbol{\varphi}) \ge 1 + \mathbf{c}^\top\mathbf{x}$, we get

$$\int_0^T \mathbb{E}_{\boldsymbol{\phi}}\left(1+\sum_{i=1}^n c_iX_i(t)\right)^{p_0} h(\mathbf{X}(t))dt \le \frac{1}{Ap_0}\left(V_{\mathbf{0}}^{p_0}(\boldsymbol{\phi}) + T\gamma_0p_0\overline{M}_{p_0,\mathbf{0}}\right),\ \forall T \ge 0. \tag{3.28}$$

A consequence of (3.28) is that there is a constant H_1 such that

$$\begin{aligned}\int_r^T \mathbb{E}_{\boldsymbol{\phi}}\Bigg(\Big(1+\sum_{i=1}^n c_iX_i(t)\Big)^{p_0} h(\mathbf{X}(t))\\ + \int_{-r}^0\Big(1+\sum_{i=1}^n c_iX_i(t+s)\Big)^{p_0} h(\mathbf{X}(t+s))\mu(ds)\Bigg)dt\\ \le H_1\left(T + V_{\mathbf{0}}^{p_0}(\boldsymbol{\phi})\right),\ \forall T \ge r.\end{aligned} \tag{3.29}$$

On the other hand, using (3.5) again, we have

$$\mathcal{L}V_{\mathbf{0}}^{p_0}(\boldsymbol{\varphi}) \le \gamma_0p_0\overline{M}_{\mathbf{0},p_0} - \gamma_0p_0V_{\mathbf{0}}^{p_0}(\boldsymbol{\varphi}) \text{ for all } \boldsymbol{\varphi} \in \mathcal{C}_+. \tag{3.30}$$

Therefore, similarly to the process of getting (3.28), we obtain

$$\int_0^T \mathbb{E}_{\boldsymbol{\phi}} V_{\mathbf{0}}^{p_0}(\mathbf{X}_t)dt \le \frac{1}{p_0\gamma_0}(T\gamma_0p_0\overline{M}_{p_0,\mathbf{0}} + V_{\mathbf{0}}^{p_0}(\boldsymbol{\phi})),\ \forall T \ge 0. \tag{3.31}$$

Combining (3.31) and the Markov inequality leads to that for any $\varepsilon, R > 0$ there exists a finite constant $R_1 = R_1(\varepsilon, R)$ such that

$$\frac{1}{t}\int_0^t \mathbb{E}_{\boldsymbol{\phi}}\mathbf{1}_{\{V_{\mathbf{0}}^{p_0}(\mathbf{X}_s)<R_1\}}ds \ge 1 - \frac{\varepsilon}{2},\ \text{provided } V_{\mathbf{0}}(\boldsymbol{\phi}) < R. \tag{3.32}$$

Because of (3.32), Lemma 3.3, and the Markov property of $\mathbf{X}_t$, for any $\varepsilon, R > 0$, there exists a compact subset $\mathcal{K} = \mathcal{K}(\varepsilon, R) := \{\boldsymbol{\varphi} : \|\boldsymbol{\varphi}\|_{2\alpha} \le R_4\}$ of $\mathcal{C}_+$ satisfying

$$\frac{1}{t}\int_{2r}^{t+2r} \mathbb{E}_{\boldsymbol{\phi}} \mathbf{1}_{\{\mathbf{X}_s\in\mathcal{K}\}} ds \ge 1-\varepsilon, \text{ provided } V_{\mathbf{0}}(\boldsymbol{\phi}) < R. \tag{3.33}$$

In the above, $R_4 = R_4(\varepsilon, R)$ is determined as in Lemma 3.3; the compactness of $\mathcal{K}$ in $\mathcal{C}$ follows the Sobolev embedding theorem.

For each $t > r$, define the following occupation measures

$$\Pi_t^{\boldsymbol{\phi}}(\cdot) := \frac{1}{t}\mathbb{E}_{\boldsymbol{\phi}} \int_r^t \mathbf{1}_{\{\mathbf{X}_s\in\cdot\}} ds. \tag{3.34}$$

Then it follows from (3.33) that for $V_{\mathbf{0}}(\boldsymbol{\phi}) < R$,

$$\left\{\Pi_t^{\boldsymbol{\phi}}(\cdot) : t \ge 2r\right\} \text{ is tight.} \tag{3.35}$$

Note that $\Pi_t^{\boldsymbol{\phi}}(\cdot)$ defined in (3.34) is a subprobability measure for each $t > 2r$. However, its weak*-limit is still a probability measure.

Lemma 3.4. *Under Assumption 2.2(b), there is a constant, still denoted by H_1 (for simplicity of notation) such that*

$$\begin{aligned}\int_r^T \mathbb{E}_{\boldsymbol{\phi}} &\left(\left(1+\sum_{i=1}^n c_iX_i(t)\right)^{p_0} h_1(\mathbf{X}(t))\right.\\ &\left.+\int_{-r}^0\left(1+\sum_{i=1}^n c_iX_i(t+s)\right)^{p_0} h_1(\mathbf{X}(t+s))\mu_1(ds)\right)dt\\ &\le H_1\left(T+V_{\mathbf{0}}^{p_0}(\boldsymbol{\phi})\right),\ \forall T\ge r.\end{aligned} \tag{3.36}$$

Proof. By (3.5), we have

$$\mathcal{L}V_{\mathbf{0}}^{p_0}(\boldsymbol{\varphi}) \le \gamma_0 p_0 \overline{M}_{p_0,\mathbf{0}} - \frac{p_0\gamma_b}{2} V_{\mathbf{0}}^{p_0}(\boldsymbol{\varphi}) \sum_{i=1}^n\Big(|f_i(\boldsymbol{\varphi})| + g_i^2(\boldsymbol{\varphi})\Big).$$

In view of the functional Itô formula,

$$\begin{aligned}\mathbb{E}_{\boldsymbol{\phi}} V_{\mathbf{0}}^{p_0}(\mathbf{X}_t) \le& V_{\mathbf{0}}^{p_0}(\boldsymbol{\phi}) + \gamma_0 p_0 \overline{M}_{p_0,\mathbf{0}} t\\ &- \frac{p_0\gamma_b}{2}\mathbb{E}_{\boldsymbol{\phi}}\int_0^t \left(1+\sum_{i=1}^n c_iX_i(s)\right)^{p_0}\sum_{i=1}^n\Big(|f_i(\mathbf{X}_s)|+g_i^2(\mathbf{X}_s)\Big)ds.\end{aligned} \tag{3.37}$$

Therefore, we get

$$\begin{aligned}\int_0^T \mathbb{E}_{\boldsymbol{\phi}} &\left(1+\sum_{i=1}^n c_iX_i(t)\right)^{p_0}\sum_{i=1}^n\Big(|f_i(\mathbf{X}_s)|+g_i^2(\mathbf{X}_s)\Big)ds\\ &\le \frac{2}{p_0\gamma_b}\left(V_{\mathbf{0}}^{p_0}(\boldsymbol{\phi}) + T\gamma_0p_0\overline{M}_{p_0,\mathbf{0}}\right) \text{ for all } T\ge 0.\end{aligned} \tag{3.38}$$

In view of (2.4) and (3.38), for all $T > 0$ one has

$$\int_0^T \mathbb{E}_{\boldsymbol{\phi}}\left(1+\sum_{i=1}^n c_iX_i(t)\right)^{p_0} h_1(\mathbf{X}_s)ds \le \frac{2}{b_1p_0\gamma_b}\left(V_{\mathbf{0}}^{p_0}(\boldsymbol{\phi}) + T\gamma_0p_0\overline{M}_{p_0,\mathbf{0}}\right). \tag{3.39}$$

Hence, we obtain (3.36). □

Remark 3. It is easily seen that $\sum_i |f_i(\boldsymbol{\varphi})| + g_i^2(\boldsymbol{\varphi})$ is uniformly integrable owing to either (3.29) and Assumption 2.2(a) or (3.36) and Assumption 2.2(b). Lemma 3.4 reveals that Assumption 2.2(b) can play the same role as Assumption 2.2(a) in guaranteeing the uniform integrability of $\sum_i |f_i(\boldsymbol{\varphi})| + g_i^2(\boldsymbol{\varphi})$. Hence, from now on, when we assume Assumption 2.2 holds, without loss of generality, we can assume that Assumption 2.2(a) holds.

Lemma 3.5. *Assume that $(\boldsymbol{\phi}_k)_{k\in\mathbb{N}} \subset \mathcal{C}_+$, $(T_k)_{k\in\mathbb{N}} \subset \mathbb{R}_+$ are such that $V_{\mathbf{0}}(\boldsymbol{\phi}_k) \le R$, $T_k > r$, $\lim_{k\to\infty} T_k = \infty$, and the sequence $(\Pi_{T_k}^{\boldsymbol{\phi}_k})_{k\in\mathbb{N}}$ converges weakly to a probability measure π. Then π is an invariant probability measure and moreover,*

$$\lim_{k\to\infty} \int_{\mathcal{C}} G(\boldsymbol{\varphi}) \Pi_{T_k}^{\boldsymbol{\phi}_k}(d\boldsymbol{\varphi}) = \int_{\mathcal{C}} G(\boldsymbol{\varphi}) \pi(d\boldsymbol{\varphi}), \tag{3.40}$$

for any function $G : \mathcal{C}_+ \to \mathbb{R}$ satisfying

$$|G(\boldsymbol{\varphi})| \le K_G \left((1 + \mathbf{c}^\top \mathbf{x})^p h(\mathbf{x}) + \int_{-r}^0 \left(1 + \sum_{i=1}^n c_i \varphi_i(s)\right)^p h(\boldsymbol{\varphi}(s))\mu(ds) \right), \tag{3.41}$$

for some $p < p_0$, where $\mathbf{x} := \boldsymbol{\varphi}(0)$. Likewise, if Assumption 2.2 (b) holds, we also have (3.40) for G satisfying

$$|G(\boldsymbol{\varphi})| \le K_G \left((1 + \mathbf{c}^\top \mathbf{x})^p h_1(\mathbf{x}) + \int_{-r}^0 \left(1 + \sum_{i=1}^n c_i \varphi_i(s)\right)^p h_1(\boldsymbol{\varphi}(s))\mu_1(ds) \right),$$

where $\mathbf{x} := \boldsymbol{\varphi}(0)$.

Proof. For the proof of π being an invariant probability measure, we refer to [18, Theorem 9.9], or [20, Proposition 6.4] with a slight modification. We proceed to prove the second assertion. For any $\varepsilon > 0$, let l_ε be sufficiently large such that for any $\boldsymbol{\varphi}$ satisfying $|\mathbf{x}| + \int_{-r}^0 |\boldsymbol{\varphi}(s)|\mu(ds) \ge 2l_\varepsilon$,

$$\frac{(1 + \mathbf{c}^\top \mathbf{x})^p h(\mathbf{x}) + \int_{-r}^0 \left(1 + \sum_{i=1}^n c_i \varphi_i(s)\right)^p h(\boldsymbol{\varphi}(s))\mu(ds)}{(1 + \mathbf{c}^\top \mathbf{x})^{p_0} h(\mathbf{x}) + \int_{-r}^0 \left(1 + \sum_{i=1}^n c_i \varphi_i(s)\right)^{p_0} h(\boldsymbol{\varphi}(s))\mu(ds)} \le \frac{\varepsilon}{K_G H_1 (1 + R^{p_0})}. \tag{3.42}$$

The above inequality follows from

$$\lim_{|\mathbf{x}|\to\infty} \frac{(1 + \mathbf{c}^\top \mathbf{x})^p h(\mathbf{x})}{(1 + \mathbf{c}^\top \mathbf{x})^{p_0} h(\mathbf{x})} = 0$$

and

$$\lim_{\int_{-r}^0 |\boldsymbol{\varphi}(s)|\mu(ds)\to\infty} \frac{\int_{-r}^0 \left(1 + \sum_{i=1}^n c_i \varphi_i(s)\right)^p h(\boldsymbol{\varphi}(s))\mu(ds)}{\int_{-r}^0 \left(1 + \sum_{i=1}^n c_i \varphi_i(s)\right)^{p_0} h(\boldsymbol{\varphi}(s))\mu(ds)} = 0 \text{ (because } h(\cdot) \ge 1).$$

Denote by $u_{l_\varepsilon} : \mathcal{C} \to [0, 1]$, a continuous function satisfying

$$u_{l_\varepsilon}(\boldsymbol{\varphi}) = \begin{cases} 1 \text{ if } |\mathbf{x}| + \int_{-r}^0 |\boldsymbol{\varphi}(s)|\mu(ds) \le 2l_\varepsilon, \\ 0 \text{ if } |\mathbf{x}| + \int_{-r}^0 |\boldsymbol{\varphi}(s)|\mu(ds) \ge 4l_\varepsilon. \end{cases}$$

By Tonelli's theorem, we get that

$$\begin{aligned}
&\int_{\mathcal{C}}\left((1+\mathbf{c}^\top\mathbf{x})^{p_0}h(\mathbf{x})+\int_{-r}^{0}\Big(1+\sum_{i=1}^{n}c_i\varphi_i(s)\Big)^{p_0}h(\boldsymbol{\varphi}(s))\mu(ds)\right)\Pi_{T_k}^{\boldsymbol{\phi}_k}(d\boldsymbol{\varphi})\\
=&\frac{1}{T_k}\int_r^{T_k}\mathbb{E}_{\boldsymbol{\phi}_k}\left(\Big(1+\sum_{i=1}^{n}c_iX_i(t)\Big)^{p_0}h(\mathbf{X}(t))\right.\\
&\qquad\left.+\int_{-r}^{0}\Big(1+\sum_{i=1}^{n}c_iX_i(t+s)\Big)h(\mathbf{X}(t+s))\mu(ds)\right)dt.
\end{aligned}\tag{3.43}$$

Because of (3.41)–(3.43), and (3.29), one gets

$$\begin{aligned}
&\int_{\mathcal{C}}\big(1-u_{l_\varepsilon}(\boldsymbol{\varphi})\big)\,|G(\boldsymbol{\varphi})|\Pi_{T_k}^{\boldsymbol{\phi}_k}(d\boldsymbol{\varphi})\\
\le&K_G\int_{\mathcal{C}}\big(1-u_{l_\varepsilon}(\boldsymbol{\varphi})\big)\left((1+\mathbf{c}^\top\mathbf{x})^{p}h(\mathbf{x})\right.\\
&\qquad\left.+\int_{-r}^{0}\Big(1+\sum_{i=1}^{n}c_i\varphi_i(s)\Big)^{p}h(\boldsymbol{\varphi}(s))\mu(ds)\right)\Pi_{T_k}^{\boldsymbol{\phi}_k}(d\boldsymbol{\varphi})\\
\le&\frac{\varepsilon}{H_1(1+R^{p_0})}\int_{\mathcal{C}}\big(1-u_{l_\varepsilon}(\boldsymbol{\varphi})\big)\left((1+\mathbf{c}^\top\mathbf{x})^{p_0}h(\mathbf{x})\right.\\
&\qquad\left.+\int_{-r}^{0}\Big(1+\sum_{i=1}^{n}c_i\varphi_i(s)\Big)^{p_0}h(\boldsymbol{\varphi}(s))\mu(ds)\right)\Pi_{T_k}^{\boldsymbol{\phi}_k}(d\boldsymbol{\varphi})\\
\le&\varepsilon.
\end{aligned}\tag{3.44}$$

Similarly, because of (3.29) and π being invariant, we have

$$\int_{\mathcal{C}}\big(1-u_{l_\varepsilon}(\boldsymbol{\varphi})\big)\,|G(\boldsymbol{\varphi})|\pi(d\boldsymbol{\varphi})\le\varepsilon.\tag{3.45}$$

The weak convergence of $\Pi_{T_k}^{\boldsymbol{\phi}_k}$ to π implies

$$\lim_{k\to\infty}\int_{\mathcal{C}}u_{l_\varepsilon}(\boldsymbol{\varphi})|G(\boldsymbol{\varphi})|\Pi_{T_k}^{\boldsymbol{\phi}_k}(d\boldsymbol{\varphi})=\int_{\mathcal{C}}u_{l_\varepsilon}(\boldsymbol{\varphi})|G(\boldsymbol{\varphi})|\pi(d\boldsymbol{\varphi}).\tag{3.46}$$

Combining (3.44), (3.45), and (3.46) yields that

$$\limsup_{k\to\infty}\left|\int_{\mathcal{C}}|G(\boldsymbol{\varphi})|\Pi_{T_k}^{\boldsymbol{\phi}_k}(d\boldsymbol{\varphi})-\int_{\mathcal{C}}|G(\boldsymbol{\varphi})|\pi(d\boldsymbol{\varphi})\right|\le 2\varepsilon.$$

Hence, the proof of the lemma is concluded by letting $\varepsilon\to 0$. □

Lemma 3.6. *Let Y be a random variable, $\theta_0>0$ be a constant, and suppose*

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

for some finite constant K_1. Then the log-Laplace transform $\eta(\theta)=\ln\mathbb{E}\exp(\theta Y)$ is twice *differentiable on* $\left[0, \frac{\theta_0}{2}\right)$ *and*

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

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

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

Proof. The proof of this lemma can be found in [24, Proof of Lemma 3.5]. □

4. Persistence

This section is devoted to proving Theorem 2.1 and Theorem 2.2. It is shown in [56, Lemma 4], by the min–max principle that Assumption 2.3 is equivalent to the existence of $\boldsymbol{\rho}^* = (\rho_1^*, \dots, \rho_n^*)$ with $\rho_i^* > 0$ such that

$$\inf_{\pi\in\mathcal{M}} \left\{ \sum_{i=1}^n \rho_i^* \lambda_i(\pi) \right\} := 2\kappa^* > 0. \tag{4.1}$$

By rescaling if necessary, we can assume that $|\boldsymbol{\rho}^*|$ is sufficiently small to satisfy condition (3.4).

Lemma 4.1. *Assume Assumptions 2.1 and 2.2 hold. For any invariant measure* π*, one has*

$$\int_{\mathcal{C}_+} Q_{\mathbf{0}}(\boldsymbol{\varphi})\pi(d\boldsymbol{\varphi}) = 0,$$

where

$$\begin{aligned} Q_{\mathbf{0}}(\boldsymbol{\varphi}) =& A_2 h(\mathbf{x}) \int_{-r}^0 e^{-\gamma s}\mu(ds) - A_2 \int_{-r}^0 h\big(\boldsymbol{\varphi}(s)\big)\mu(ds) \\ &- A_2\gamma \int_{-r}^0 \mu(ds) \int_s^0 e^{\gamma(u-s)} h\big(\boldsymbol{\varphi}(u)\big)du \\ &+ \frac{\sum_{i=1}^n c_i x_i f_i(\boldsymbol{\varphi})}{1+\mathbf{c}^\top\mathbf{x}} - \frac{1}{2}\sum_{i,j=1}^n \frac{c_i c_j \sigma_{ij} x_i x_j g_i(\boldsymbol{\varphi}) g_j(\boldsymbol{\varphi})}{\left(1+\mathbf{c}^\top\mathbf{x}\right)^2}. \end{aligned}$$

Proof. Because of (3.29), (3.36), Lemma 3.5, and Assumption 2.2, $Q_{\mathbf{0}}$ is π-integrable. By the strong law of large numbers (see, e.g., [31, Theorem 4.2]) we have

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t Q_{\mathbf{0}}(\mathbf{X}_s)ds = \int_{\mathcal{C}_+} Q_{\mathbf{0}}(\boldsymbol{\varphi})\pi(d\boldsymbol{\varphi}), \quad \mathbb{P}_\pi\text{- a.s. }, \tag{4.2}$$

and

$$\begin{aligned} &\lim_{t\to\infty} \frac{1}{t}\int_0^t \frac{\sum_{i,j}\sigma_{ij}c_i c_j X_i(s) X_j(s) g_i(\mathbf{X}_s) g_j(\mathbf{X}_s)}{(1+\sum_i c_i X_i(s))^2} ds \\ &= \int_{\mathcal{C}_+} \frac{\sum_{i,j}\sigma_{ij}c_i c_j x_i x_j g_i(\boldsymbol{\varphi}) g_j(\boldsymbol{\varphi})}{(1+\mathbf{c}^\top\mathbf{x})^2}\pi(d\boldsymbol{\varphi}) < \infty \qquad \mathbb{P}_\pi\text{- a.s. }, \text{ where } \mathbf{x} := \boldsymbol{\varphi}(0). \end{aligned}$$

The above limit tells us that if we let $\langle L_{\cdot}, L_{\cdot}\rangle_t$ be the quadratic variation of the local martingale

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

then

$$\limsup_{t\to\infty} \frac{\langle L_{\cdot}, L_{\cdot}\rangle_t}{t} = \int_{\mathcal{C}_+} \frac{\sum_{i,j} \sigma_{ij} c_i c_j x_i x_j g_i(\boldsymbol{\varphi}) g_j(\boldsymbol{\varphi})}{(1+\mathbf{c}^\top \mathbf{x})^2} \pi(d\boldsymbol{\varphi}) < \infty \qquad \mathbb{P}_\pi\text{-a.s.}$$

Applying the strong law of large numbers for local martingales (see [41, Theorem 1.3.4]),

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t \frac{\sum_i c_i X_i(s) g_i(\mathbf{X}_s) dE_i(s)}{1+\sum_i c_i X_i(s)} = 0 \;\; \mathbb{P}_\pi\text{-a.s.} \tag{4.3}$$

As in (3.6), we have $\mathcal{L}U_{\mathbf{0}}(\boldsymbol{\varphi}) = Q_{\mathbf{0}}(\boldsymbol{\varphi})$, where

$$U_{\mathbf{0}}(\boldsymbol{\varphi}) = \ln(1+\mathbf{c}^\top \mathbf{x}) + A_2 \int_{-r}^0 \mu(ds) \int_s^0 e^{\gamma(u-s)} h\big(\boldsymbol{\varphi}(u)\big) du, \mathbf{x} := \boldsymbol{\varphi}(0). \tag{4.4}$$

Combining (4.2)–(4.4), and the functional Itô formula yields that

$$0 \le \lim_{t\to\infty} \frac{U_{\mathbf{0}}(\mathbf{X}_t)}{t} = \int_{\mathcal{C}_+} Q_{\mathbf{0}}(\boldsymbol{\varphi}) \pi(d\boldsymbol{\varphi}) \;\; \mathbb{P}_\pi\text{-a.s.} \tag{4.5}$$

A simple contradiction argument coupled with (3.8) and (4.5) leads to that

$$\int_{\mathcal{C}_+} Q_{\mathbf{0}}(\boldsymbol{\varphi}) \pi(d\boldsymbol{\varphi}) = 0. \quad \square$$

Lemma 4.2. *Assume Assumptions 2.1–2.3 hold. Let $\boldsymbol{\rho}^*$ be as in (4.1). For any compact set $\mathcal{K}$ of $\mathcal{C}_+$, there exists a $T_{\mathcal{K}} > r$ such that for any $T \ge T_{\mathcal{K}}$ and $\boldsymbol{\phi} \in \partial\mathcal{C}_+ \cap \mathcal{K}$, we have*

$$\int_r^T \mathbb{E}_{\boldsymbol{\phi}} Q_{\boldsymbol{\rho}^*}(\mathbf{X}_t) dt \le -\kappa^* T, \tag{4.6}$$

where

$$Q_{\boldsymbol{\rho}^*}(\boldsymbol{\varphi}) := Q_{\mathbf{0}}(\boldsymbol{\varphi}) - \sum_{i=1}^n \rho_i^* \left(f_i(\boldsymbol{\varphi}) - \frac{\sigma_{ii} g_i^2(\boldsymbol{\varphi})}{2} \right).$$

Proof. We prove the lemma by using a contradiction argument. Suppose that we can find $\boldsymbol{\phi}_k \in \partial\mathcal{C}_+ \cap \mathcal{K}$ and $T_k > r$, $T_k \uparrow \infty$ such that

$$\int_r^{T_k} \mathbb{E}_{\boldsymbol{\phi}} Q_{\boldsymbol{\rho}^*}(\mathbf{X}_t) dt \ge -\kappa^* T_k. \tag{4.7}$$

Since $\mathbb{E}_{\boldsymbol{\phi}} |Q_{\boldsymbol{\rho}^*}(\mathbf{X}_t)| \le H_1 \widetilde{K}(t + V_{\mathbf{0}}^{p_0}(\boldsymbol{\phi}))$ due to (3.29), (3.36), and Assumption 2.2, we can apply Tonelli's theorem to obtain

$$\int_{\mathcal{C}_+} Q_{\boldsymbol{\rho}^*}(\boldsymbol{\varphi}) \Pi_{T_k}^{\boldsymbol{\phi}_k}(d\boldsymbol{\varphi}) = \frac{1}{T_k} \int_r^{T_k} \mathbb{E}_{\boldsymbol{\phi}_k} Q_{\boldsymbol{\rho}^*}(\mathbf{X}_t) dt.$$

Under Assumption 2.2, as a consequence of Lemma 3.5,

$$\lim_{k\to\infty} \frac{1}{T_k} \int_r^{T_k} \mathbb{E}_{\boldsymbol{\phi}_k} Q_{\boldsymbol{\rho}^*}(\mathbf{X}_t) dt = \int_{\mathcal{C}_+} Q_{\boldsymbol{\rho}^*}(\boldsymbol{\varphi}) \pi(d\boldsymbol{\varphi}), \tag{4.8}$$

where the invariant measure π is the weak limit of $\{\Pi_{T_k}^{\boldsymbol{\phi}_k}\}$. Since the initial values lie on the boundary, π is supported in $\partial\mathcal{C}_+$. This combined with Lemma 4.1 implies that

$$\int_{\mathcal{C}_+} Q_{\boldsymbol{\rho}^*}(\boldsymbol{\varphi})\pi(d\boldsymbol{\varphi}) = -\sum_{i=1}^{n} \rho_i^* \lambda_i(\pi). \tag{4.9}$$

Thus, we obtain from (4.1), (4.8), and (4.9) that

$$\lim_{k\to\infty} \frac{1}{T_k} \int_r^{T_k} \mathbb{E}_{\boldsymbol{\phi}_k} Q_{\boldsymbol{\rho}^*}(\mathbf{X}_t)dt \le -2\kappa^*. \tag{4.10}$$

Combining (4.7) and (4.10) leads to a contradiction. As a result, the Lemma is proved. □

Now, let n^* be sufficiently large to satisfy

$$\gamma_0(n^*-1) - A_0 > 0, \tag{4.11}$$

and $p_1 > p_0$ but (3.4) still holds. Under Assumption 2.2, a consequence of (3.29) is that there is an H_1^* satisfying

$$\int_r^T \mathbb{E}_{\boldsymbol{\phi}} Q_{\boldsymbol{\rho}^*}(\mathbf{X}_t)dt \le H_1^*(T + V_{\mathbf{0}}^{p_0}(\boldsymbol{\phi})) \text{ for all } T \ge r. \tag{4.12}$$

Because of (3.8), we have

$$\mathbb{E}_{\boldsymbol{\phi}} V_{\mathbf{0}}^{p_1}(\mathbf{X}_t) \le V_{\mathbf{0}}^{p_1}(\boldsymbol{\phi})e^{-\gamma_0 p_1 t} + \overline{M}_{p_1,\mathbf{0}}. \tag{4.13}$$

Note that

$$\text{if } \boldsymbol{\phi} \in \mathcal{C}_{V,M} \text{ then } V_{\mathbf{0}}(\boldsymbol{\phi}) \le (1 + M\,|\mathbf{c}|)e^{A_0}, \tag{4.14}$$

where $\mathcal{C}_{V,M}$ is defined as in Theorem 3.1. Eqs. (4.13) and (4.14) imply that

$$\mathbb{E}_{\boldsymbol{\phi}} V_{\mathbf{0}}^{p_1}(\mathbf{X}_t) \le R_{V,M} \text{ if } \boldsymbol{\phi} \in \mathcal{C}_{V,M}, t \ge 0 \tag{4.15}$$

for some $R_{V,M} > 0$. Let $\varepsilon^* \in (0, \frac{1}{9})$ be such that

$$R_{V,M}^{\frac{p_0}{p_1}} \left(\varepsilon^*\right)^{\frac{p_1-p_0}{p_1}} + \varepsilon^* H_1^* T \le \frac{\kappa^*}{10}(T-2r) \text{ for any } T \ge 3r. \tag{4.16}$$

To take care of the case when $X_i(t)$ is small but the norm of the segment function $\|X_t\|$ is not, we derive the following Lemma.

Lemma 4.3. *Assume Assumptions 2.1–2.3 hold. There exist a $\widehat{\delta} > 0$ and $T^* > 0$ such that for any $T \in [T^*, n^*T^*]$,*

$$\mathbb{E}_{\boldsymbol{\phi}} \int_0^T Q_{\boldsymbol{\rho}^*}(\mathbf{X}_t)dt \le -\frac{1}{2}\kappa^* T, \text{ if } \boldsymbol{\phi} \in \mathcal{C}_V(\widehat{\delta}), \tag{4.17}$$

where

$$\mathcal{C}_V(\widehat{\delta}) := \left\{\boldsymbol{\phi} \in \mathcal{C}_+^\circ \cap \mathcal{C}_{V,M} \text{ such that } |\phi_i(0)| \le \widehat{\delta} \text{ for some } i\right\}.$$

Proof. For any event $\mathcal{A}$ with $\mathbb{P}_{\boldsymbol{\phi}}(\mathcal{A}) \geq 1 - \varepsilon$, we obtain from (4.12), the Hölder inequality, and (4.15) that

$$\begin{aligned}\mathbb{E}_{\boldsymbol{\phi}} \mathbf{1}_{\mathcal{A}^c} \int_{3r}^{T} Q_{\boldsymbol{\rho}^*}(\mathbf{X}_t)dt \leq & \mathbb{E}\mathbf{1}_{\mathcal{A}^c} H_1^* \left(T + V_{\mathbf{0}}^{p_0}(\mathbf{X}_{2r})\right) \\ \leq & \left(\varepsilon^*\right)^{\frac{p_1 - p_0}{p_1}} \left(\mathbb{E}_{\boldsymbol{\phi}} V_{\mathbf{0}}^{p_1}(\mathbf{X}_{2r})\right)^{\frac{p_0}{p_1}} + \varepsilon^* H_1^* T \\ \leq & R_{V,M}^{\frac{p_0}{p_1}} \left(\varepsilon^*\right)^{\frac{p_1 - p_0}{p_1}} + \varepsilon^* H_1^* T \text{ if } \boldsymbol{\phi} \in \mathcal{C}_{V,M},\end{aligned} \tag{4.18}$$

where $\mathcal{A}^c = \Omega \setminus \mathcal{A}$. Applying Lemma 3.3 implies that there is a compact set $\mathcal{K}^* = \mathcal{K}^*(\varepsilon^*) := \{\boldsymbol{\varphi} \in \mathcal{C}_+ : \|\boldsymbol{\varphi}\| \leq R, \|\boldsymbol{\varphi}\|_{2\alpha} - \|\boldsymbol{\varphi}\| \leq R_4\}$ ($R_4 = R_4(\varepsilon^*)$ is as in Lemma 3.3) such that

$$\mathbb{P}_{\boldsymbol{\phi}} \left\{\mathbf{X}_t \in \mathcal{K}^* \text{ for all } t \in [2r, 3r]\right\} \geq 1 - \frac{\varepsilon^*}{2} \text{ if } \boldsymbol{\phi} \in \mathcal{C}_{V,M}. \tag{4.19}$$

In view of Lemma 4.2, there exists $T^* = T^*(\mathcal{K}^*) > 0$ such that for all $T \geq T^* - 3r$,

$$\int_{r}^{T} \mathbb{E}_{\boldsymbol{\phi}} Q_{\boldsymbol{\rho}^*}(\mathbf{X}_t)dt \leq -\kappa^* T \text{ if } \boldsymbol{\phi} \in \partial\mathcal{C}_+ \cap \mathcal{K}^*. \tag{4.20}$$

Without loss of generality, we can choose $T^* > 3r$ sufficiently large such that for all $T \geq T^*$,

$$-\frac{7}{10}\kappa^*(T - 2r) + 3A_0 r \leq -\frac{1}{2}\kappa^* T. \tag{4.21}$$

In view of the Feller property of $\mathbf{X}_t$ and (4.20), we obtain that there is a $\delta_1 > 0$ such that for all $T \in [T^* - 3r, n^* T^*]$

$$\int_{r}^{T} \mathbb{E}_{\boldsymbol{\phi}} Q_{\boldsymbol{\rho}^*}(\mathbf{X}_t)dt \leq -\frac{9}{10}\kappa^* T \text{ if } \boldsymbol{\phi} \in \mathcal{K}^*, \text{dist}(\boldsymbol{\phi}, \partial\mathcal{C}_+) < \delta_1. \tag{4.22}$$

By virtue of (4.19), part (i) of Proposition 3.1, and the structure of $\mathcal{K}^*$, there exists a $\widehat{\delta} > 0$ such that

$$\mathbb{P}_{\boldsymbol{\phi}}(\mathcal{A}) \geq 1 - \varepsilon^* \text{ if } \boldsymbol{\phi} \in \mathcal{C}_V(\widehat{\delta}), \tag{4.23}$$

where

$$\mathcal{A} = \left\{\text{dist}(\mathbf{X}_{2r}, \partial\mathcal{C}_+) < \delta_1, \mathbf{X}_{2r} \in \mathcal{K}^*\right\}.$$

Combining (4.22) and (4.23) leads to that for all $T \in [T^*, n^* T^*]$

$$\mathbb{E}_{\boldsymbol{\phi}} \mathbf{1}_{\mathcal{A}} \int_{3r}^{T} Q_{\boldsymbol{\rho}^*}(\mathbf{X}_t)dt \leq -\frac{9}{10}\kappa^*(T - 2r)(1 - \varepsilon^*) \leq -\frac{8}{10}\kappa^*(T - 2r) \text{ if } \boldsymbol{\phi} \in \mathcal{C}_V(\widehat{\delta}). \tag{4.24}$$

We obtain from (4.18), (4.16), and (4.24) that for all $T \in [T^*, n^* T^*]$

$$\mathbb{E}_{\boldsymbol{\phi}} \int_{3r}^{T} Q_{\boldsymbol{\rho}^*}(\mathbf{X}_t)dt \leq -\frac{7}{10}\kappa^*(T - 2r), \boldsymbol{\phi} \in \mathcal{C}_V(\widehat{\delta}). \tag{4.25}$$

Using the functional Itô formula, Jensen's inequality, and (3.7), we have

$$\begin{aligned}\int_{0}^{3r} \mathbb{E}_{\boldsymbol{\phi}} Q_{\boldsymbol{\rho}^*}(\mathbf{X}_s)ds = & \frac{1}{p_0}\mathbb{E}_{\boldsymbol{\phi}} \left(\ln V_{-\boldsymbol{\rho}^*}^{p_0}(\mathbf{X}_{3r}) - \ln V_{-\boldsymbol{\rho}^*}^{p_0}(\boldsymbol{\phi})\right) \\ \leq & \frac{1}{p_0} \ln \frac{\mathbb{E}_{\boldsymbol{\phi}} V_{-\boldsymbol{\rho}^*}^{p_0}(\mathbf{X}_{3r})}{V_{-\boldsymbol{\rho}^*}^{p_0}(\boldsymbol{\phi})} \leq \frac{\ln e^{3A_0 p_0 r}}{p_0} = 3A_0 r.\end{aligned} \tag{4.26}$$

Therefore, we obtain from (4.25), (4.26), and (4.21) that if $\boldsymbol{\phi} \in \mathcal{C}_V(\widehat{\delta})$,

$$\mathbb{E}_{\boldsymbol{\phi}} \int_0^T Q_{\boldsymbol{\rho}^*}(\mathbf{X}_t)dt \leq -\frac{7}{10}\kappa^*(T-2r) + 3A_0 r \leq -\frac{1}{2}\kappa^* T, \text{ for all } T \in [T^*, n^*T^*].$$

The lemma is proved. □

Proposition 4.1. *Assume that Assumptions 2.1–2.3 hold. Then there are* $\theta \in \left(0, \frac{p_0}{2}\right)$ *and* $\widetilde{K}_\theta > 0$ *such that for any* $T \in [T^*, n^*T^*]$ *and* $\boldsymbol{\phi} \in \mathcal{C}_+^\circ \cap \mathcal{C}_{V,M}$,

$$\mathbb{E}_{\boldsymbol{\phi}} V^\theta_{-\boldsymbol{\rho}^*}(\mathbf{X}_T) \leq V^\theta_{-\boldsymbol{\rho}^*}(\boldsymbol{\phi}) \exp\Big(-\frac{1}{4}\theta\kappa^* T\Big) + \widetilde{K}_\theta, \tag{4.27}$$

where $-\boldsymbol{\rho}^* = \left(-\rho_1^*, \ldots, -\rho_n^*\right)$.

Proof. By the functional Itô formula, we obtain that

$$\begin{aligned} \ln V_{-\boldsymbol{\rho}^*}(\mathbf{X}_T) &= \ln V_{-\boldsymbol{\rho}^*}(\boldsymbol{\phi}) + \int_0^T Q_{\boldsymbol{\rho}^*}(\mathbf{X}_t)dt \\ &\quad + \int_0^T \left(\frac{\sum_i c_i X_i(t) g(\mathbf{X}_t) dE_i(t)}{1+\sum_i c_i X_i(t)} - \sum_i \rho_i^* g_i(\mathbf{X}_t) dE_i(t)\right) \\ &=: \ln V_{-\boldsymbol{\rho}^*}(\boldsymbol{\phi}) + z(T). \end{aligned} \tag{4.28}$$

Because of (3.7) and (4.28), we have

$$\mathbb{E}_{\boldsymbol{\phi}} \exp\left(p_0 z(T)\right) = \frac{\mathbb{E}_{\boldsymbol{\phi}} V^{p_0}_{-\boldsymbol{\rho}^*}(\mathbf{X}_T)}{V^{p_0}_{-\boldsymbol{\rho}^*}(\boldsymbol{\phi})} \leq e^{A_0 p_0 T}. \tag{4.29}$$

Another consequence of (3.7) is that

$$\frac{\mathbb{E}_{\boldsymbol{\phi}} V^{p_0}_{\boldsymbol{\rho}^*}(\mathbf{X}_T)}{V^{p_0}_{\boldsymbol{\rho}^*}(\boldsymbol{\phi})} \leq e^{A_0 p_0 T}. \tag{4.30}$$

We obtain from the definition of $V^{p_0}_{\boldsymbol{\rho}}(\boldsymbol{\varphi})$ that

$$\begin{aligned} V^{-p_0}_{-\boldsymbol{\rho}^*}(\boldsymbol{\varphi}) &= \left(1+\mathbf{c}^\top \mathbf{x}\right)^{-2p_0} \exp\Big\{-2p_0 A_2 \int_{-r}^0 \mu(ds) \int_s^0 e^{\gamma(u-s)} h\big(\boldsymbol{\varphi}(u)\big)du\Big\} V^{p_0}_{\boldsymbol{\rho}^*}(\boldsymbol{\varphi}) \\ &\leq V^{p_0}_{\boldsymbol{\rho}^*}(\boldsymbol{\varphi}), \text{ where } \mathbf{x} := \boldsymbol{\varphi}(0). \end{aligned} \tag{4.31}$$

Applying (4.30) and (4.31) to (4.28) yields

$$\begin{aligned} &\mathbb{E}_{\boldsymbol{\phi}} \exp(-p_0 z(T)) = \frac{\mathbb{E}_{\boldsymbol{\phi}} V^{-p_0}_{-\boldsymbol{\rho}^*}(\mathbf{X}_T)}{V^{-p_0}_{-\boldsymbol{\rho}^*}(\boldsymbol{\phi})} \\ &\leq \frac{\mathbb{E}_{\boldsymbol{\phi}} V^{p_0}_{\boldsymbol{\rho}^*}(\mathbf{X}_T)}{V^{\delta_0}_{\boldsymbol{\rho}^*}(\boldsymbol{\phi})} \left(1+\mathbf{c}^\top \boldsymbol{\phi}(0)\right)^{-2p_0} \exp\Big\{-2p_0 A_2 \int_{-r}^0 \mu(ds) \int_s^0 e^{\gamma(u-s)} h\big(\boldsymbol{\phi}(u)\big)du\Big\} \\ &\leq \left(1+\mathbf{c}^\top \boldsymbol{\phi}(0)\right)^{-2p_0} \exp\Big\{-2p_0 A_2 \int_{-r}^0 \mu(ds) \int_s^0 e^{\gamma(u-s)} h\big(\boldsymbol{\phi}(u)\big)du\Big\} \exp(A_0 p_0 T). \end{aligned} \tag{4.32}$$

In view of (4.29) and (4.32), an application of Lemma 3.6 for $z(T)$ implies that there is $\tilde{K}_2 \geq 0$ such that

$$0 \leq \frac{d^2\tilde{\eta}_{\boldsymbol{\phi},T}}{d\theta^2}(\theta) \leq \tilde{K}_2 \text{ for all } \theta \in \left[0, \frac{p_0}{2}\right), \ \boldsymbol{\phi} \in \mathcal{C}_V(\widehat{\delta}), T \in [T^*, n^*T^*],$$

where

$$\tilde{\eta}_{\boldsymbol{\phi},T}(\theta) = \ln \mathbb{E}_{\boldsymbol{\phi}} \exp(\theta z(T)).$$

Hence, using Lemma 3.6 and (4.17) yields

$$\frac{d\tilde{\eta}_{\boldsymbol{\phi},T}}{d\theta}(0) = \mathbb{E}_{\boldsymbol{\phi}} z(T) \le -\frac{1}{2}\kappa^* T \text{ for } \boldsymbol{\phi} \in \mathcal{C}_V(\widehat{\delta}), T \in [T^*, n^*T^*].$$

By a Taylor expansion around $\theta = 0$, for $\boldsymbol{\phi} \in \mathcal{C}_V(\widehat{\delta})$, $T \in [T^*, n^*T^*]$, and $\theta \in \left[0, \frac{p_0}{2}\right)$, we have

$$\tilde{\eta}_{\boldsymbol{\phi},T}(\theta) \le -\frac{1}{2}\kappa^* T\theta + \theta^2 \tilde{K}_2.$$

Now, if we choose $\theta \in \left(0, \frac{p_0}{2}\right)$ satisfying $\theta < \frac{\kappa^* T^*}{4\tilde{K}_2}$, we get

$$\tilde{\eta}_{\boldsymbol{\phi},T}(\theta) \le -\frac{1}{4}\kappa^* T\theta \quad \text{for all } \boldsymbol{\phi} \in \mathcal{C}_V(\widehat{\delta}), T \in [T^*, n^*T^*]. \tag{4.33}$$

In light of (4.33), we have for such θ, $\boldsymbol{\phi} \in \mathcal{C}_V(\widehat{\delta})$, and $T \in [T^*, n^*T^*]$ that

$$\frac{\mathbb{E}_{\boldsymbol{\phi}} V^{\theta}_{-\boldsymbol{\rho}^*}(\mathbf{X}_T)}{V^{\theta}_{-\boldsymbol{\rho}^*}(\boldsymbol{\phi})} = \exp \tilde{\eta}_{\boldsymbol{\phi},T}(\theta) \le \exp\left(-\frac{1}{4}\kappa^* T\theta\right). \tag{4.34}$$

On the other hand, because of (3.7), we have for $\boldsymbol{\phi} \notin \mathcal{C}_V(\widehat{\delta})$ but satisfying $\boldsymbol{\phi} \in \mathcal{C}_+^\circ \cap \mathcal{C}_{V,M}$ and $T \in [T^*, n^*T^*]$ that

$$\mathbb{E}_{\boldsymbol{\phi}} V^{\theta}_{-\boldsymbol{\rho}^*}(\mathbf{X}_T) \le \exp(\theta n^* T^* H) \sup_{\boldsymbol{\phi} \notin \mathcal{C}_V(\widehat{\delta}), \boldsymbol{\phi} \in \mathcal{C}_{V,M}} \{V^{\theta}_{-\boldsymbol{\rho}^*}(\boldsymbol{\phi})\} =: \widetilde{K}_\theta < \infty. \tag{4.35}$$

Combining (4.34) and (4.35) completes the proof. □

Theorem 4.1. *Assume that Assumptions 2.1, 2.2, and 2.3 hold. There is a finite constant K^* such that*

$$\limsup_{t\to\infty} \mathbb{E}_{\boldsymbol{\phi}} V^{\theta}_{-\boldsymbol{\rho}^*}(\mathbf{X}_t) \le K^* \textit{ for all } \boldsymbol{\phi} \in \mathcal{C}_+^\circ.$$

As a result, the solution $\mathbf{X}$ of (2.1) is strongly stochastically persistent.

Proof. Once we obtained Proposition 4.1, the proof is similar to [24, Theorem 4.1]. By virtue of (3.5), we have

$$\mathcal{L} V^{\theta}_{-\boldsymbol{\rho}^*}(\boldsymbol{\varphi}) \le -\theta\gamma_0 V^{\theta}_{-\boldsymbol{\rho}^*}(\boldsymbol{\varphi}) \text{ if } \boldsymbol{\varphi} \notin \mathcal{C}_{V,M}, \tag{4.36}$$

where $\mathcal{C}_{V,M}$ is as in Theorem 3.1. Define the stopping time

$$\tau = \inf\{t \ge 0 : \mathbf{X}_t \in \mathcal{C}_{V,M}\}. \tag{4.37}$$

We obtain from Dynkin's formula and (4.36) that

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{\phi}} \left[\exp\{\theta\gamma_0(\tau \wedge n^*T^*)\} V^{\theta}_{-\boldsymbol{\rho}^*}(\mathbf{X}_{\tau\wedge n^*T^*})\right] \\
&\quad \le V^{\theta}_{-\boldsymbol{\rho}^*}(\boldsymbol{\phi}) + \mathbb{E}_{\boldsymbol{\phi}} \int_0^{\tau\wedge n^*T^*} \exp\{\theta\gamma_0 s\} \left[\mathcal{L} V^{\theta}_{-\boldsymbol{\rho}^*}(\mathbf{X}_s) + \theta\gamma_0 V^{\theta}_{-\boldsymbol{\rho}^*}(\mathbf{X}_s)\right] ds \\
&\quad \le V^{\theta}_{-\boldsymbol{\rho}^*}(\boldsymbol{\phi}).
\end{aligned}$$

As a consequence,

$$\begin{aligned}
V^{\theta}_{-\boldsymbol{\rho}^*}(\boldsymbol{\phi}) \geq & \mathbb{E}_{\boldsymbol{\phi}}\left[\exp\big\{\theta\gamma_0\big(\tau\wedge n^*T^*\big)\big\}V^{\theta}_{-\boldsymbol{\rho}^*}\big(\mathbf{X}_{\tau\wedge n^*T^*}\big)\right]\\
= & \mathbb{E}_{\boldsymbol{\phi}}\left[\mathbf{1}_{\{\tau\leq(n^*-1)T^*\}}\exp\big\{\theta\gamma_0\big(\tau\wedge n^*T^*\big)\big\}V^{\theta}_{-\boldsymbol{\rho}^*}\big(\mathbf{X}_{\tau\wedge n^*T^*}\big)\right]\\
& +\ \mathbb{E}_{\boldsymbol{\phi}}\left[\mathbf{1}_{\{(n^*-1)T^*<\tau<n^*T^*\}}\exp\big\{\theta\gamma_0\big(\tau\wedge n^*T^*\big)\big\}V^{\theta}_{-\boldsymbol{\rho}^*}\big(\mathbf{X}_{\tau\wedge n^*T^*}\big)\right]\\
& +\ \mathbb{E}_{\boldsymbol{\phi}}\left[\mathbf{1}_{\{\tau\geq n^*T^*\}}\exp\big\{\theta\gamma_0\big(\tau\wedge n^*T^*\big)\big\}V^{\theta}_{-\boldsymbol{\rho}^*}\big(\mathbf{X}_{\tau\wedge n^*T^*}\big)\right]\\
\geq & \mathbb{E}_{\boldsymbol{\phi}}\left[\mathbf{1}_{\{\tau\leq(n^*-1)T^*\}}V^{\theta}_{-\boldsymbol{\rho}^*}(\mathbf{X}_\tau)\right]\\
& +\ \exp\left\{\theta\gamma_0((n^*-1)T^*)\right\}\mathbb{E}_{\boldsymbol{\phi}}\left[\mathbf{1}_{\{(n^*-1)T^*<\tau<n^*T^*\}}V^{\theta}_{-\boldsymbol{\rho}^*}(\mathbf{X}_\tau)\right]\\
& +\ \exp\left\{\theta\gamma_0 n^*T^*\right\}\mathbb{E}_{\boldsymbol{\phi}}\left[\mathbf{1}_{\{\tau\geq n^*T^*\}}V^{\theta}_{-\boldsymbol{\rho}^*}(\mathbf{X}_{n^*T^*})\right].
\end{aligned}\tag{4.38}$$

Combining the Markov property of $(\mathbf{X}_t)$ and Proposition 4.1 yields

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{\phi}}\left[\mathbf{1}_{\{\tau\leq(n^*-1)T^*\}}V^{\theta}_{-\boldsymbol{\rho}^*}(\mathbf{X}_{n^*T^*})\right]\\
&\quad\leq \mathbb{E}_{\boldsymbol{\phi}}\left[\mathbf{1}_{\{\tau\leq(n^*-1)T^*\}}\left[\widetilde{K}_\theta+e^{-\frac14\theta\kappa^*(n^*T^*-\tau)}V^{\theta}_{-\boldsymbol{\rho}^*}(\mathbf{X}_\tau)\right]\right]\\
&\quad\leq \widetilde{K}_\theta+\exp\left(-\frac14\theta\kappa^*T^*\right)\mathbb{E}_{\boldsymbol{\phi}}\left[\mathbf{1}_{\{\tau\leq(n^*-1)T^*\}}V^{\theta}_{-\boldsymbol{\rho}^*}(\mathbf{X}_\tau)\right].
\end{aligned}\tag{4.39}$$

Using again the strong Markov property of $(\mathbf{X}_t)$ and (3.7), we obtain that

$$\begin{aligned}
&\mathbb{E}_{\boldsymbol{\phi}}\left[\mathbf{1}_{\{(n^*-1)T^*<\tau<n^*T^*\}}V^{\theta}_{-\boldsymbol{\rho}^*}(\mathbf{X}_{n^*T^*})\right]\\
&\quad\leq \mathbb{E}_{\boldsymbol{\phi}}\left[\mathbf{1}_{\{(n^*-1)T^*<\tau<n^*T^*\}}e^{\theta A_0(n^*T^*-\tau)}V^{\theta}_{-\boldsymbol{\rho}^*}(\mathbf{X}_\tau)\right]\\
&\quad\leq \exp(\theta A_0T^*)\mathbb{E}_{\boldsymbol{\phi}}\left[\mathbf{1}_{\{(n^*-1)T^*<\tau<n^*T^*\}}V^{\theta}_{-\boldsymbol{\rho}^*}(\mathbf{X}_\tau)\right].
\end{aligned}\tag{4.40}$$

Applying (4.39) and (4.40) to (4.38) leads to

$$\begin{aligned}
&V^{\theta}_{-\boldsymbol{\rho}^*}(\boldsymbol{\phi})\\
&\geq\exp\left(\frac14\theta\kappa^*T^*\right)\mathbb{E}_{\boldsymbol{\phi}}\left[\mathbf{1}_{\{\tau\leq(n^*-1)T^*\}}V^{\theta}_{-\boldsymbol{\rho}^*}(\mathbf{X}_{n^*T^*})\right]-\exp\left(\frac14\theta\kappa^*T^*\right)\widetilde{K}_\theta\\
&\quad+\ \exp(-\theta A_0T^*)\exp\left(\theta\gamma_0(n^*-1)T^*\right)\mathbb{E}_{\boldsymbol{\phi}}\left[\mathbf{1}_{\{(n^*-1)T^*<\tau<n^*T^*\}}V^{\theta}_{-\boldsymbol{\rho}^*}(\mathbf{X}_{n^*T^*})\right]\\
&\quad+\ \exp\left(\theta\gamma_0n^*T^*\right)\mathbb{E}_{\boldsymbol{\phi}}\left[\mathbf{1}_{\{\tau\geq n^*T^*\}}V^{\theta}_{-\boldsymbol{\rho}^*}(\mathbf{X}_{n^*T^*})\right]\\
&\geq\exp(m\theta T^*)\mathbb{E}_{\boldsymbol{\phi}}V^{\theta}_{-\boldsymbol{\rho}^*}(\mathbf{X}_{n^*T^*})-\widetilde{K}_\theta\exp\left(\frac14\theta\kappa^*T^*\right),
\end{aligned}\tag{4.41}$$

where $m=\min\left\{\frac14\kappa^*,\gamma_0n^*,\gamma_0(n^*-1)-A_0\right\}>0$ by (4.11). As a result,

$$\mathbb{E}_{\boldsymbol{\phi}}V^{\theta}_{-\boldsymbol{\rho}^*}(\mathbf{X}_{n^*T^*})\leq\widehat{q}_1V^{\theta}_{-\boldsymbol{\rho}^*}(\boldsymbol{\phi})+q_1^*\ \text{ for all }\ \boldsymbol{\phi}\in\mathcal{C}_+^\circ,$$

for some $0<\widehat{q}_1<1$, $0<q_1^*<\infty$. Therefore, by the Markov property of $\mathbf{X}_t$, we have

$$\mathbb{E}_{\boldsymbol{\phi}}V^{\theta}_{-\boldsymbol{\rho}^*}(\mathbf{X}_{(k+1)n^*T^*})\leq\widehat{q}_1\mathbb{E}_{\boldsymbol{\phi}}V^{\theta}_{-\boldsymbol{\rho}^*}(\mathbf{X}_{kn^*T^*})+q_1^*\ \text{ for all }\ \boldsymbol{\phi}\in\mathcal{C}_+^\circ,$$

Using this recursively, we obtain

$$\mathbb{E}_{\boldsymbol{\phi}}V^{\theta}_{-\boldsymbol{\rho}^*}(\mathbf{X}_{kn^*T^*})\leq\widehat{q}_1^kV^{\theta}_{-\boldsymbol{\rho}^*}(\boldsymbol{\phi})+\frac{q_1^*(1-\widehat{q}_1^k)}{1-\widehat{q}_1}.\tag{4.42}$$

We obtain from (3.7) and (4.42) that

$$\mathbb{E}_{\boldsymbol{\phi}}V^{\theta}_{-\boldsymbol{\rho}^*}(\mathbf{X}_T)\leq\left[\widehat{q}_1^kV^{\theta}_{-\boldsymbol{\rho}^*}(\boldsymbol{\phi})+\frac{q_1^*(1-\widehat{q}_1^k)}{1-\widehat{q}_1}\right]e^{A_0\theta T^*}\ \text{ for all }\ T\in[kn^*T^*,kn^*T^*+T^*].$$

Hence, by letting $k \to \infty$, we obtain the existence of a finite constant K^* such that

$$\limsup_{t\to\infty} \mathbb{E}_{\boldsymbol{\phi}} V^{\theta}_{-\boldsymbol{\rho}^*}(\mathbf{X}_t) \le K^* \text{ for all } \boldsymbol{\phi} \in \mathcal{C}^{\circ}_{+}. \tag{4.43}$$

Finally, the strongly stochastic persistence of $\mathbf{X}$ is obtained by applying Markov's inequality to (4.43), and using $V_{-\boldsymbol{\rho}^*}(\boldsymbol{\varphi}) \ge \frac{1+\mathbf{c}^\top \mathbf{x}}{\prod_{i=1}^n x_i^{\rho_i^*}}$ and $\sum_{i=1}^n \rho_i^* < 1$, $\rho_i^* > 0$. $\square$

To proceed, we prove the uniqueness of the invariant probability measure under suitable assumptions. For $\mathbf{x} \in \mathbb{R}^{n,\circ}_{+}$, $\ln \mathbf{x}$ is understood as the component-wise logarithm of $\mathbf{x}$. By using the fact

$$\left|\mathbf{x}^{(1)} - \mathbf{x}^{(2)}\right| \le \left|\ln \mathbf{x}^{(1)} - \ln \mathbf{x}^{(2)}\right| \left|1 + \mathbf{x}^{(1)} + \mathbf{x}^{(2)}\right|$$

and basic inequalities (Young's inequality and the Cauchy–Schwarz inequality), we obtain from (2.7) and Assumption 2.4 (ii) that there is some constant D_1 depending only D_0, d_0 satisfying

$$\begin{aligned}
&\sum_i \Big(\left| f_i(\boldsymbol{\varphi}^{(1)}) - f_i(\boldsymbol{\varphi}^{(2)})\right| + \left| g_i(\boldsymbol{\varphi}^{(1)}) - g_i(\boldsymbol{\varphi}^{(2)})\right| + \left| g_i^2(\boldsymbol{\varphi}^{(1)}) - g_i^2(\boldsymbol{\varphi}^{(2)})\right| \Big) \\
&\quad \le D_1 \left|\ln \mathbf{x}^{(1)} - \ln \mathbf{x}^{(2)}\right| \left|1 + \mathbf{x}^{(1)} + \mathbf{x}^{(2)}\right|^{d_0+1} \\
&\qquad + D_1 \int_{-r}^{0} \left|\ln \boldsymbol{\varphi}^{(1)}(s) - \ln \boldsymbol{\varphi}^{(2)}(s)\right| \left|1 + \boldsymbol{\varphi}^{(1)}(s) + \boldsymbol{\varphi}^{(2)}(s)\right|^{d_0+1} \mu(ds),
\end{aligned} \tag{4.44}$$

and

$$\begin{aligned}
&\sum_i \Big(\left| f_i(\boldsymbol{\varphi}^{(1)}) - f_i(\boldsymbol{\varphi}^{(2)})\right| + \left| g_i(\boldsymbol{\varphi}^{(1)}) - g_i(\boldsymbol{\varphi}^{(2)})\right| \\
&\qquad + \left| g_i^2(\boldsymbol{\varphi}^{(1)}) - g_i^2(\boldsymbol{\varphi}^{(2)})\right| \Big) \left|\ln \mathbf{x}^{(1)} - \ln \mathbf{x}^{(2)}\right| \\
&\le D_1 \left|\ln \mathbf{x}^{(1)} - \ln \mathbf{x}^{(2)}\right|^2 \left|1 + \mathbf{x}^{(1)} + \mathbf{x}^{(2)}\right|^{d_0+1} \\
&\qquad + D_1 \int_{-r}^{0} \left|\ln \boldsymbol{\varphi}^{(1)}(s) - \ln \boldsymbol{\varphi}^{(2)}(s)\right|^2 \left|1 + \boldsymbol{\varphi}^{(1)}(s) + \boldsymbol{\varphi}^{(2)}(s)\right|^{2d_0+2} \mu(ds),
\end{aligned} \tag{4.45}$$

and

$$\begin{aligned}
&\sum_i \Big(\left| f_i(\boldsymbol{\varphi}^{(1)}) - f_i(\boldsymbol{\varphi}^{(2)})\right| + \left| g_i(\boldsymbol{\varphi}^{(1)}) - g_i(\boldsymbol{\varphi}^{(2)})\right| \\
&\qquad + \left| g_i^2(\boldsymbol{\varphi}^{(1)}) - g_i^2(\boldsymbol{\varphi}^{(2)})\right| \Big) \left|\ln \mathbf{x}^{(1)} - \ln \mathbf{x}^{(2)}\right|^3 \\
&\le D_1 \left|\ln \mathbf{x}^{(1)} - \ln \mathbf{x}^{(2)}\right|^4 \left|1 + \mathbf{x}^{(1)} + \mathbf{x}^{(2)}\right|^{4d_0+4} \\
&\qquad + D_1 \int_{-r}^{0} \left|\ln \boldsymbol{\varphi}^{(1)}(s) - \ln \boldsymbol{\varphi}^{(2)}(s)\right|^4 \left|1 + \boldsymbol{\varphi}^{(1)}(s) + \boldsymbol{\varphi}^{(2)}(s)\right|^{4d_0+4} \mu(ds).
\end{aligned} \tag{4.46}$$

To apply asymptotic couplings method and results in [23], we let $Y_i(t) = \ln X_i(t)$ and consider the following equations

$$\begin{cases} dY_i(t) = \left[f_i(\mathbf{X}_t) - \frac{g_i^2(\mathbf{X}_t)\sigma_{ii}^2}{2} \right] dt + g_i(\mathbf{X}_t) dE_i(t), \ i = 1, \dots, n, \\ \mathbf{Y}_0 = \ln \boldsymbol{\phi}, \ \boldsymbol{\phi} \in \mathcal{C}^{\circ}_{+}, \end{cases} \tag{4.47}$$

and

$$\begin{cases} d\widetilde{Y}_i(t) = \left[f_i(\widetilde{\mathbf{X}}_t) - \frac{g_i^2(\widetilde{\mathbf{X}}_t)\sigma_{ii}^2}{2} \right] dt + \widetilde{\lambda} \left[1 + X_i(t) + \widetilde{X}_i(t) \right]^{4d_0+4} \left[Y_i(t) - \widetilde{Y}_i(t) \right] dt \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad + g_i(\widetilde{\mathbf{X}}_t) dE_i(t), \ i = 1, \dots, n \\ \widetilde{\mathbf{Y}}_0 = \ln \widetilde{\boldsymbol{\phi}}, \ \widetilde{\boldsymbol{\phi}} \neq \boldsymbol{\phi} \in \mathcal{C}_+^\circ, \end{cases} \tag{4.48}$$

where $\widetilde{\lambda}$ is sufficiently large to be determined later; and $\widetilde{\mathbf{X}}(t) := \left(e^{\widetilde{Y}_1(t)}, \dots, e^{\widetilde{Y}_n(t)} \right)$. Put $\mathbf{Z} := \mathbf{Y} - \widetilde{\mathbf{Y}}$. Combining the functional Itô formula and (4.44)–(4.46), one has

$$\begin{aligned} d\left|\mathbf{Z}(t)\right|^2 \le & \left[D_2 \int_{-r}^0 \left| 1 + \widetilde{\mathbf{X}}(t+s) + \mathbf{X}(t+s) \right|^{4d_0+4} \left|\mathbf{Z}(t+s)\right|^2 \mu(ds) \right] dt \\ & - \left(\widetilde{\lambda} - D_2 \right) \left| 1 + \widetilde{\mathbf{X}}(t) + \mathbf{X}(t) \right|^{4d_0+4} \left|\mathbf{Z}(t)\right|^2 dt \\ & + 2 \sum_i \left(Y_i(t) - \widetilde{Y}_i(t) \right) \left(g_i(\mathbf{X}_t) - g_i(\widetilde{\mathbf{X}}_t) \right) dE_i(t), \end{aligned} \tag{4.49}$$

and

$$\begin{aligned} d\left|\mathbf{Z}(t)\right|^4 \le & \left[D_2 \int_{-r}^0 \left| 1 + \widetilde{\mathbf{X}}(t+s) + \mathbf{X}(t+s) \right|^{4d_0+4} \left|\mathbf{Z}(t+s)\right|^4 \mu(ds) \right] dt \\ & - \left(\widetilde{\lambda} - D_2 \right) \left| 1 + \widetilde{\mathbf{X}}(t) + \mathbf{X}(t) \right|^{4d_0+4} \left|\mathbf{Z}(t)\right|^4 dt \\ & + 4 \sum_i \left(Y_i(t) - \widetilde{Y}_i(t) \right)^3 \left(g_i(\mathbf{X}_t) - g_i(\widetilde{\mathbf{X}}_t) \right) dE_i(t), \end{aligned} \tag{4.50}$$

for some constant D_2 depending only D_0, d_0, σ_{ij} and independent of $\mathbf{X}_t$, $\widetilde{\mathbf{X}}_t$. For $\boldsymbol{\varphi}^{(1)}, \boldsymbol{\varphi}^{(2)} \in \mathcal{C}_+^\circ$, define

$$\begin{aligned} \widetilde{U}(\boldsymbol{\varphi}^{(1)}, \boldsymbol{\varphi}^{(2)}) := & \left| \ln \mathbf{x}^{(1)} - \ln \mathbf{x}^{(2)} \right|^4 \\ & + \frac{D_2 + 9n\sigma^* D_1}{\gamma} \int_{-r}^0 \mu(ds) \int_s^0 e^{\gamma(u-s)} \left| 1 + \boldsymbol{\varphi}^{(1)}(s) + \boldsymbol{\varphi}^{(2)}(s) \right|^{4d_0+4} \\ & \qquad \times \left| \boldsymbol{\varphi}^{(1)}(s) - \boldsymbol{\varphi}^{(2)}(s) \right|^4 du, \end{aligned}$$

and

$$\begin{aligned} U(\boldsymbol{\varphi}^{(1)}, \boldsymbol{\varphi}^{(2)}) := & \left| \ln \mathbf{x}^{(1)} - \ln \mathbf{x}^{(2)} \right|^2 \\ & + \frac{D_2 + 9n\sigma^* D_1}{\gamma} \int_{-r}^0 \mu(ds) \int_s^0 e^{\gamma(u-s)} \left| 1 + \boldsymbol{\varphi}^{(1)}(s) + \boldsymbol{\varphi}^{(2)}(s) \right|^{4d_0+4} \\ & \qquad \times \left| \boldsymbol{\varphi}^{(1)}(s) - \boldsymbol{\varphi}^{(2)}(s) \right|^2 du, \end{aligned}$$

where $\sigma^* := \max\{\sigma_{ij} : 1 \le i, j \le n\}$. Hence, by direct calculations using the functional Itô formula, [46, Remark 2.2] and then applying (4.49), (4.50), it is easily seen that we can choose

$\widetilde{\lambda}$ being sufficiently large such that

$$\begin{aligned} d\widetilde{U}(\mathbf{X}_t,\widetilde{\mathbf{X}}_t) \le &-2D_3\Bigg(\widetilde{U}(\mathbf{X}_t,\widetilde{\mathbf{X}}_t)+|1+\mathbf{X}(t)+\widetilde{\mathbf{X}}(t)|^{4d_0+4}|\mathbf{Z}(t)|^4 \\ &+\int_{-r}^0 |1+\mathbf{X}(t+s)+\widetilde{\mathbf{X}}(t+s)|^{4d_0+4}|\mathbf{Z}(t+s)|^4\mu(ds)\Bigg)dt \\ &+4\sum_i \big(Y_i(t)-\widetilde{Y}_i(t)\big)^3\big(g_i(\mathbf{X}_t)-g_i(\widetilde{\mathbf{X}}_t)\big)dE_i(t), \end{aligned} \tag{4.51}$$

and

$$\begin{aligned} dU(\mathbf{X}_t,\widetilde{\mathbf{X}}_t) \le &-2D_3\Bigg(U(\mathbf{X}_t,\widetilde{\mathbf{X}}_t)+|1+\mathbf{X}(t)+\widetilde{\mathbf{X}}(t)|^{4d_0+4}|\mathbf{Z}(t)|^2 \\ &+\int_{-r}^0 |1+\mathbf{X}(t+s)+\widetilde{\mathbf{X}}(t+s)|^{4d_0+4}|\mathbf{Z}(t+s)|^2\mu(ds)\Bigg)dt \\ &+2\sum_i \big(Y_i(t)-\widetilde{Y}_i(t)\big)\big(g_i(\mathbf{X}_t)-g_i(\widetilde{\mathbf{X}}_t)\big)dE_i(t), \end{aligned} \tag{4.52}$$

for some positive constant $D_3 > 8n\sigma^* D_1$.

Similar to [23, Theorem 3.1], let

$$\mathbf{v}(t)=\widetilde{\lambda}\left[(g_i(\widetilde{\mathbf{X}}_t)g_j(\widetilde{\mathbf{X}}_t)\sigma_{ij})_{n\times n}\right]^{-1}\left[1+|\mathbf{X}(t)|+|\widetilde{\mathbf{X}}(t)|\right]^{4d_0+4}\big(\mathbf{Y}(t)-\widetilde{\mathbf{Y}}(t)\big),$$

where $\left[(g_i(\boldsymbol{\varphi})g_j(\boldsymbol{\varphi})\sigma_{ij})_{n\times n}\right]^{-1}$ is the inverse matrix of matrix $\left[(g_i(\boldsymbol{\varphi})g_j(\boldsymbol{\varphi})\sigma_{ij})_{n\times n}\right]$ and for each $\varepsilon > 0$

$$\widetilde{\tau}_\varepsilon := \inf\left\{t\ge 0: \int_0^t |\mathbf{v}(s)|^2 ds \ge \varepsilon^{-1}\|\mathbf{Y}_0-\widetilde{\mathbf{Y}}_0\|^2\right\}.$$

Lemma 4.4. *The following assertions hold:*

$$\lim_{\varepsilon\to 0}\mathbb{P}\{\widetilde{\tau}_\varepsilon=\infty\}=1, \tag{4.53}$$

and

$$\lim_{t\to\infty}|\mathbf{Y}(t)-\widetilde{\mathbf{Y}}(t)|=0 \text{ a.s.} \tag{4.54}$$

Remark 4. In the proof of the Lemma, our purpose is to prove $\lim_{t\to\infty} U(\mathbf{X}_t,\widetilde{\mathbf{X}}_t)=0$ a.s. To handle the diffusion part of $U(\mathbf{X}_t,\widetilde{\mathbf{X}}_t)$, we need some helps from $\widetilde{U}(\mathbf{X}_t,\widetilde{\mathbf{X}}_t)$. That is why we introduced both $U(\mathbf{X}_t,\widetilde{\mathbf{X}}_t)$ and $\widetilde{U}(\mathbf{X}_t,\widetilde{\mathbf{X}}_t)$ in the above.

Proof. A consequence of (4.51) and (4.52) is that

$$d\left[e^{D_3t}\mathbb{E}U(\mathbf{X}_t,\widetilde{\mathbf{X}}_t)\right]\le -D_3e^{D_3t}\mathbb{E}U(\mathbf{X}_t,\widetilde{\mathbf{X}}_t)dt,$$

and

$$d\left[e^{D_3t}\mathbb{E}\widetilde{U}(\mathbf{X}_t,\widetilde{\mathbf{X}}_t)\right]\le -D_3e^{D_3t}\mathbb{E}\widetilde{U}(\mathbf{X}_t,\widetilde{\mathbf{X}}_t)dt,$$

which implies that

$$\lim_{t\to\infty}\mathbb{E}e^{D_3t}U(\mathbf{X}_t,\widetilde{\mathbf{X}}_t)=\lim_{t\to\infty}\mathbb{E}e^{D_3t}\widetilde{U}(\mathbf{X}_t,\widetilde{\mathbf{X}}_t)=0, \tag{4.55}$$

and

$$\mathbb{E}\int_0^\infty e^{D_3 t} U(\mathbf{X}_t, \widetilde{\mathbf{X}}_t)dt \le \frac{U(\boldsymbol{\phi}, \widetilde{\boldsymbol{\phi}})}{D_3}. \tag{4.56}$$

An application of Markov's inequality and (4.56) imply that

$$\mathbb{P}\left(\int_0^\infty e^{D_3 t} U(\mathbf{X}_t, \widetilde{\mathbf{X}}_t)dt \le \frac{U(\boldsymbol{\phi}, \widetilde{\boldsymbol{\phi}})}{D_3\sqrt{\varepsilon}}\right) \ge 1 - \sqrt{\varepsilon}. \tag{4.57}$$

To proceed, similar to the proof of Proposition 5.1 part (i) (in particular, the process of getting (3.24)), we can obtain that for $p < p_0$,

$$\mathbb{E}_{\boldsymbol{\phi}} \sup_{t\in[0,1]} V_{\mathbf{0}}^p(\mathbf{X}_t) \le C_1 V_{\mathbf{0}}^p(\mathbf{X}_0),$$

for some constant C_1. Then, combining with (3.8), we have

$$\mathbb{E}_{\boldsymbol{\phi}} \sup_{t\in[k,k+1]} V_{\mathbf{0}}^p(\mathbf{X}_t) \le C_2(1 + V_{\mathbf{0}}^p(\mathbf{X}_0)),$$

for some constant C_2. We have for any $C > 0$,

$$\begin{aligned}
&\mathbb{P}\left\{e^{\frac{-D_3 t}{2}} V_{\mathbf{0}}^{4d_0+4}(\mathbf{X}_t) \ge \frac{C}{\sqrt{\varepsilon}}, \text{ for some } t \in [k, k+1]\right\}\\
\le&\mathbb{P}\left\{V_{\mathbf{0}}^p(\mathbf{X}_t) \ge \frac{Ce^{D_3 pk/(8d_0+8)}}{\varepsilon^{p/(8d_0+8)}}, \text{ for some } t \in [k, k+1]\right\}\\
\le&C_2(1 + V_{\mathbf{0}}^p(\mathbf{X}_0))\frac{\varepsilon^{p/(8d_0+8)}}{Ce^{D_3 pk/(8d_0+8)}}.
\end{aligned}$$

Thus, one has

$$\begin{aligned}
\mathbb{P}\left\{e^{-\frac{D_3}{2}t} V_{\mathbf{0}}^{4d_0+4}(\mathbf{X}_t) \le \frac{C}{\sqrt{\varepsilon}}\forall t \ge 0\right\} \ge&1 - \sum_{k=0}^\infty C_2(1 + V_{\mathbf{0}}^p(\mathbf{X}_0))\frac{\varepsilon^{p/(8d_0+8)}}{Ce^{D_3 pk/(8d_0+8)}}\\
\ge&1 - K_{C,V_{\mathbf{0}}^p(\mathbf{X}_0)}\varepsilon^{p/(8d_0+8)},
\end{aligned} \tag{4.58}$$

for some finite constant $K_{C,V_{\mathbf{0}}^p(\mathbf{X}_0)}$ depending on C, $V_{\mathbf{0}}^p(\mathbf{X}_0)$. Combining (4.57) and (4.58), we can obtain

$$\begin{aligned}
&\mathbb{P}\left(\int_0^\infty \left[1 + |\mathbf{X}(t)| + |\widetilde{\mathbf{X}}(t)|\right]^{4d_0+4} \left|\mathbf{Y}(t) - \widetilde{\mathbf{Y}}(t)\right| dt \le \frac{C}{\varepsilon}\right)\\
&\ge 1 - \sqrt{\varepsilon} - K_{C,V_{\mathbf{0}}^p(\mathbf{X}_0)}\varepsilon^{\frac{p}{8d_0+8}}.
\end{aligned} \tag{4.59}$$

We obtain from the definition of $\widetilde{\tau}_\varepsilon$ that

$$\mathbb{P}\{\widetilde{\tau}_\varepsilon = \infty\} \ge \mathbb{P}\left\{\int_0^\infty |\mathbf{v}(s)|^2\, ds < \varepsilon^{-1}\|\boldsymbol{\phi} - \widetilde{\boldsymbol{\phi}}\|^2\right\}. \tag{4.60}$$

Since (4.59), (4.60), definition of $\mathbf{v}(\cdot)$, and Assumption 2.4(iii), we obtain (4.53) by letting $\varepsilon \to 0$.

On the other hand, applying the Burkholder–Davis–Gundy inequality and (4.44), we have

$$\begin{aligned}
&\mathbb{E}\sup_{t\in[0,1]}\sum_i \big(Y_i(t)-\widetilde{Y}_i(t)\big)\big(g_i(\mathbf{X}_t)-g_i(\widetilde{\mathbf{X}}_t)\big)\,dE_i(t)\\
&\le 4\mathbb{E}\left(n\sigma^*\sum_i^n\int_0^1 \big(Y_i(t)-\widetilde{Y}_i(t)\big)^2\big(g_i(\mathbf{X}_t)-g_i(\widetilde{\mathbf{X}}_t)\big)^2\,dt\right)^{\frac12}\\
&\le 4n\sqrt{\sigma^* D_1}\left(\mathbb{E}\int_0^1\left(|1+\mathbf{X}(t)+\widetilde{\mathbf{X}}(t)|^{4d_0+4}|\mathbf{Z}(t)|^4\right.\right.\\
&\qquad\left.\left.+\int_{-r}^0 |1+\mathbf{X}(t+s)+\widetilde{\mathbf{X}}(t+s)|^{4d_0+4}|\mathbf{Z}(t+s)|^4\mu(ds)\right)dt\right)^{\frac12}.
\end{aligned}\tag{4.61}$$

We obtain from (4.51) and the functional Itô formula that

$$\begin{aligned}
&\mathbb{E}\int_0^1\left(|1+\mathbf{X}(t)+\widetilde{\mathbf{X}}(t)|^{4d_0+4}|\mathbf{Z}(t)|^4\right.\\
&\qquad\left.+\int_{-r}^0 |1+\mathbf{X}(t+s)+\widetilde{\mathbf{X}}(t+s)|^{4d_0+4}|\mathbf{Z}(t+s)|^4\mu(ds)\right)dt\\
&\quad\le \frac{1}{2D_3}\mathbb{E}\widetilde{U}(\boldsymbol{\phi},\widetilde{\boldsymbol{\phi}}).
\end{aligned}\tag{4.62}$$

Applying (4.62) to (4.61) yields that

$$\mathbb{E}\sup_{t\in[0,1]}\sum_i\big(g_i(\mathbf{X}_t)-g_i(\widetilde{\mathbf{X}}_t)\big)\,dE_i(t)\le 4n\sqrt{\sigma^*D_1}\left(\frac{1}{2D_3}\mathbb{E}\widetilde{U}(\boldsymbol{\phi},\widetilde{\boldsymbol{\phi}})\right)^{\frac12}.\tag{4.63}$$

Hence, combining (4.52) and (4.63), by a standard argument, we conclude that

$$\mathbb{E}\sup_{t\in[0,1]}U(\mathbf{X}_{t+t_0},\widetilde{\mathbf{X}}_{t+t_0})\le D_4\left(\mathbb{E}U(\mathbf{X}_{t_0},\widetilde{\mathbf{X}}_{t_0})+\big(\mathbb{E}\widetilde{U}(\mathbf{X}_{t_0},\widetilde{\mathbf{X}}_{t_0})\big)^{\frac12}\right),\ \forall t_0>0,\tag{4.64}$$

for some constant D_4, independent of $\mathbf{X}_t,\widetilde{\mathbf{X}}_t$. A consequence of Markov's inequality and (4.64) is that

$$\begin{aligned}
&\mathbb{P}\left\{\sup_{t\in[n-1,n]}U(\mathbf{X}_t,\widetilde{\mathbf{X}}_t)\ge e^{-\frac{D_3 n}{4}}\right\}\\
&\le D_4 e^{\frac{D_3 n}{4}}\left(\mathbb{E}U(\mathbf{X}_{n-1},\widetilde{\mathbf{X}}_{n-1})+\big(\mathbb{E}\widetilde{U}(\mathbf{X}_{n-1},\widetilde{\mathbf{X}}_{n-1})\big)^{\frac12}\right).
\end{aligned}\tag{4.65}$$

We obtain from (4.55) and (4.65) that

$$\sum_{n=1}^\infty\mathbb{P}\left\{\sup_{t\in[n-1,n]}U(\mathbf{X}_t,\widetilde{\mathbf{X}}_t)\ge e^{-\frac{D_3 n}{4}}\right\}<\infty.\tag{4.66}$$

It follows from the Borel–Cantelli lemma and (4.66) that $\lim_{t\to\infty}U(\mathbf{X}_t,\widetilde{\mathbf{X}}_t)=0$ a.s. and thus we get (4.54). $\square$

Once we have Lemma 4.4, we can mimic the proof of [23, Theorem 3.1] to obtain the uniqueness of the invariant probability measure of (4.47), which is stated as in the following Proposition.

Proposition 4.2. *Under Assumptions 2.1 and 2.4, the solution process of (4.47) has at most one invariant probability measure, and moreover (2.1) has at most one invariant probability measure concentrated on $\mathcal{C}_+^\circ$.*

By the tightness (3.35) of the occupation measures and Theorem 4.1, the existence of invariant probability measure of (2.1) concentrated on $\mathcal{C}_+^\circ$ is guaranteed. Combined with Proposition 4.2, we have the following Theorem to end this section.

Theorem 4.2. *Under Assumptions 2.1–2.4, system (2.1) has a unique invariant probability measure concentrated on $\mathcal{C}_+^\circ$.*

5. Applications

This section presents a number of applications of our main results to different models. We make use of Theorems 2.2 together with the following lemma whose proof can be found in [24], to characterize the persistence.

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

Moreover, it is worth noting that these sufficient conditions for persistence are sharp and are almost necessary in the sense that if they are not satisfied and critical cases are excluded, the extinction will take place, which will be seen in part (II) [45].

5.1. Stochastic delay Lotka–Volterra competitive models

The Lotka–Volterra model, introduced in [39,63], is one of the most popular models in mathematical biology and has been studied extensively in the literature. When two or more species live in proximity and share the same basic resources, they usually compete for food, habitat, territory, etc., we therefore have the Lotka–Volterra competitive model. To capture many complex properties in real life, other terms (white noises, Markov switching, delayed time, etc.) are added to the original system. Stochastic delay Lotka–Volterra competitive models have also been widely studied; see, for example, [1,32] and references therein. This kind model for two species has the form

$$\begin{cases} dX_1(t) = X_1(t)\left(a_1 - b_{11}X_1(t) - b_{12}X_2(t) - \widehat{b}_{11}X_1(t-r) - \widehat{b}_{12}X_2(t-r)\right)dt \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad + X_1(t)dE_1(t), \\ dX_2(t) = X_2(t)\left(a_2 - b_{21}X_1(t) - b_{22}X_2(t) - \widehat{b}_{21}X_1(t-r) - \widehat{b}_{22}X_2(t-r)\right)dt \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad + X_2(t)dE_2(t). \end{cases} \tag{5.1}$$

Note that in the above $X_i(t)$ is the size of the species i at time t; $a_i > 0$ represents the growth rate of the species i; $b_{ii} > 0$ is the intra-specific competition of the ith species; $b_{ij} \geq 0$, $(i \neq j)$ stands for the inter-specific competition; $\widehat{b}_{ij} > -b_{ii}$ $(i, j = 1, 2)$ (i.e., $\widehat{b}_{ij}$ can be negative); r is the delay time; $(E_1(t), E_2(t))^\top = \Gamma^\top \mathbf{B}(t)$ with $\mathbf{B}(t) = (B_1(t), B_2(t))^\top$ being a vector of independent standard Brownian motions and Γ being a 2×2 matrix such that $\Gamma^\top \Gamma = (\sigma_{ij})_{2\times 2}$ is a positive definite matrix.

Before applying our Theorems, let us verify our Assumptions. First, it is easy to see that there is a sufficiently large M_1 such that

$$\frac{\sum_{i,j=1}^{2}\sigma_{ij}x_ix_j}{(1+x_1+x_2)^2}\geq 2\sigma_* \text{ if } |\mathbf{x}|>M_1, \mathbf{x}:=(x_1,x_2), \tag{5.2}$$

for some $\sigma_*>0$. There exist $0<b_2^*<b_1^*$ and $M_2>0$ satisfying

$$\begin{aligned}&\frac{\sum_{i=1}^{2}x_i\left(a_i-b_{i1}x_1-b_{i2}x_2-\widehat{b}_{i1}\varphi_1(-r)-\widehat{b}_{i2}\varphi_2(-r)\right)}{1+x_1+x_2}\\&<-b_1^*(1+|\mathbf{x}|)+b_2^*\left|\boldsymbol{\varphi}(-r)\right|,\end{aligned} \tag{5.3}$$

for all $\boldsymbol{\varphi}\in\mathcal{C}_+$ satisfying $|\mathbf{x}|:=|\boldsymbol{\varphi}(0)|>M_2$, and

$$\begin{aligned}&\frac{\sum_{i=1}^{2}x_i\left(a_i-b_{i1}x_1-b_{i2}x_2-\widehat{b}_{i1}\varphi_1(-r)-\widehat{b}_{i2}\varphi_2(-r)\right)}{1+x_1+x_2}\\&<|\mathbf{x}|\sum_i a_i+b_2^*\left|\boldsymbol{\varphi}(-r)\right|, \ \forall\boldsymbol{\varphi}\in\mathcal{C}_+.\end{aligned} \tag{5.4}$$

Let $M>\max\{M_1,M_2\}$, $\mathbf{c}=(1,1)$,

$$0<\gamma_b<\min\left\{\frac{b_1^*}{2\sum_i a_i},\frac{\sigma_*}{2},\frac{b_1^*-b_2^*}{\sum_{i,j}\left(b_{ij}+|\widehat{b}_{ij}|\right)}\right\},\quad 0<\gamma_0<\frac{b_1^*}{2}-\gamma_b\sum_i a_i,$$

A_1,A_2 be such that

$$0<b_2^*+\gamma_b\sum_{i,j}|\widehat{b}_{ij}|<A_2<A_1<b_1^*-\gamma_b\sum_{i,j}b_{ij} \text{ and } A_1-A_2<\frac{b_1^*}{2},$$

and $h(\mathbf{x}):=1+|\mathbf{x}|$, μ is be the Dirac delta measure (concentrated) at $\{-r\}$, and

$$\begin{aligned}A_0:=&\gamma_0+A_1(1+M)+\gamma_b\left(\sum_i a_i+M\sum_{ij}b_{ij}+2\right)+M\sum_i a_i\\&+\sup_{|\mathbf{x}|<M}\left\{\frac{\sum_{i,j=1}^{2}\sigma_{ij}x_ix_j}{(1+x_1+x_2)^2}+\frac{\sum_{i=1}^{2}x_i\left(a_i-b_{i1}x_1-b_{i2}x_2\right)}{1+x_1+x_2}\right\}.\end{aligned}$$

Combined with (5.2)–(5.4), direct calculations lead to that (2.2) is satisfied and that Assumption 2.1 holds. Moreover, it is easy to confirm that Assumptions 2.2 and 2.4 also hold.

Applying our Theorems in Section 2, we have that $\lambda_i(\boldsymbol{\delta}^*)=a_i-\frac{\sigma_{ii}}{2}, i=1,2$. Let $\mathcal{C}_{1+}^\circ:=\{(\varphi_1,0)\in\mathcal{C}_+:\varphi_1(s)>0\ \forall s\in[-r,0]\}$ and $\mathcal{C}_{2+}^\circ:=\{(0,\varphi_2)\in\mathcal{C}_+:\varphi_2(s)>0\ \forall s\in[-r,0]\}$. In view of Theorem 2.2, if $\lambda_i(\boldsymbol{\delta}^*)>0$, there is a unique invariant probability measure π_i on $\mathcal{C}_{i+}^\circ$, $i=1,2$. By Lemma 5.1, we have

$$\lambda_i(\pi_i)=a_i-\frac{\sigma_{ii}}{2}-\int_{\mathcal{C}_{i+}^\circ}\left(b_{ii}\varphi_i(0)+\widehat{b}_{ii}\varphi_i(-r)\right)\pi_i(d\boldsymbol{\varphi})=0, \text{ where } \boldsymbol{\varphi}=(\varphi_1,\varphi_2),$$

which implies

$$\int_{\mathcal{C}_{i+}^\circ}\left(b_{ii}\varphi_i(0)+\widehat{b}_{ii}\varphi_i(-r)\right)\pi_i(d\boldsymbol{\varphi})=a_i-\frac{\sigma_{ii}}{2}. \tag{5.5}$$

Since π_i is an invariant probability measure of $\{\mathbf{X}_t\}$, it is easy to see that

$$\int_{\mathcal{C}_{i+}^\circ} \varphi_i(0)\pi_i(d\boldsymbol{\varphi}) = \lim_{T\to\infty}\frac{1}{T}\int_0^T X_{i,t}(0)dt = \lim_{T\to\infty}\frac{1}{T}\int_0^T X_i(t)dt, \tag{5.6}$$

where $(X_{1,t}, X_{2,t}) = \mathbf{X}_t$. Similarly,

$$\int_{\mathcal{C}_{i+}^\circ} \varphi_i(-r)\pi_i(d\boldsymbol{\varphi}) = \lim_{T\to\infty}\frac{1}{T}\int_0^T X_i(t-r)dt. \tag{5.7}$$

By virtue of (5.6) and (5.7), we can prove that

$$\int_{\mathcal{C}_{i+}^\circ} \varphi_i(0)\pi_i(d\boldsymbol{\varphi}) = \int_{\mathcal{C}_{i+}^\circ} \varphi_i(-r)\pi_i(d\boldsymbol{\varphi}). \tag{5.8}$$

Combining (5.5) and (5.8) yields that

$$\int_{\mathcal{C}_{i+}^\circ} \varphi_i(0)\pi_i(d\boldsymbol{\varphi}) = \int_{\mathcal{C}_{i+}^\circ} \varphi_i(-r)\pi_i(d\boldsymbol{\varphi}) = \frac{a_i - \frac{\sigma_{ii}}{2}}{b_{ii}+\widehat{b}_{ii}}.$$

Therefore, we have

$$\begin{aligned}\lambda_2(\pi_1) &= \int_{\mathcal{C}_{1+}^\circ}\left[a_2 - \frac{\sigma_{22}}{2} - b_{21}\varphi_1(0) - \widehat{b}_{21}\varphi_1(-r)\right]\pi_1(d\boldsymbol{\varphi})\\ &= a_2 - \frac{\sigma_{22}}{2} - \left(a_1 - \frac{\sigma_{11}}{2}\right)\cdot\frac{b_{21}+\widehat{b}_{21}}{b_{11}+\widehat{b}_{11}},\end{aligned}$$

and

$$\begin{aligned}\lambda_1(\pi_2) &= \int_{\mathcal{C}_{2+}^\circ}\left[a_1 - \frac{\sigma_{11}}{2} - b_{12}\varphi_2(0) - \widehat{b}_{12}\varphi_2(-r)\right]\pi_2(d\boldsymbol{\varphi})\\ &= a_1 - \frac{\sigma_{11}}{2} - \left(a_2 - \frac{\sigma_{22}}{2}\right)\cdot\frac{b_{12}+\widehat{b}_{12}}{b_{22}+\widehat{b}_{22}}.\end{aligned}$$

If $\lambda_1(\boldsymbol{\delta}^*) > 0, \lambda_2(\boldsymbol{\delta}^*) > 0$ and $\lambda_1(\pi_2) > 0, \lambda_2(\pi_1) > 0$, any invariant probability measure in $\partial\mathcal{C}_+$ has the form $\pi = q_0\boldsymbol{\delta}^* + q_1\pi_1 + q_2\pi_2$ with $0 \le q_0, q_1, q_2$ and $q_0 + q_1 + q_2 = 1$. Then, one has $\max_{i=1,2}\{\lambda_i(\pi)\} > 0$ for any π having the form as above. As a consequence of Theorem 2.2, there is a unique invariant probability measure π^* on $\mathcal{C}_+^\circ$. This result generalizes the results of long-term properties in [32].

In the above, we considered a 2-dimension case to illustrate the idea as well as to simplify the explicit computation. For the stochastic delay Lotka–Volterra competitive model with n-species, our results can still be applied to characterize the long-term behavior of the solution.

5.2. Stochastic delay Lotka–Volterra predator–prey models

To continue our study of Lotka–Volterra competitive models, this section is devoted to applying our results to stochastic Lotka–Volterra predator–prey models with time delay. Such models are frequently used to describe the dynamics of biological systems in which two species interact, one as a predator and the other one as prey. In this section, we consider Lotka–Volterra predator–prey system with one prey and two competing predators as follows

$$\begin{cases} dX_1(t) = X_1(t)\Big\{ a_1 - b_{11}X_1(t) - b_{12}X_2(t) - b_{13}X_3(t) \\ \qquad\qquad -\widehat{b}_{11}X_1(t-r) - \widehat{b}_{12}X_2(t-r) - \widehat{b}_{13}X_3(t-r)\Big\}dt + X_1(t)dE_1(t), \\ dX_2(t) = X_2(t)\Big\{ -a_2 + b_{21}X_1(t) - b_{22}X_2(t) - b_{23}X_3(t) \\ \qquad\qquad -\widehat{b}_{21}X_1(t-r) - \widehat{b}_{22}X_2(t-r) - \widehat{b}_{23}X_3(t-r)\Big\}dt + X_2(t)dE_2(t), \\ dX_3(t) = X_3(t)\Big\{ -a_3 + b_{31}X_1(t) - b_{32}X_2(t) - b_{33}X_3(t) \\ \qquad\qquad -\widehat{b}_{31}X_1(t-r) - \widehat{b}_{32}X_2(t-r) - \widehat{b}_{33}X_3(t-r)\Big\}dt + X_3(t)dE_3(t), \end{cases} \tag{5.9}$$

where $X_1(t)$, $X_2(t)$, and $X_3(t)$ are the densities at time t of the prey, and two predators, respectively; $a_1 > 0$ is the growth rate; $a_2, a_3 > 0$ are the death rate of X_2, X_3; $b_{ii} > 0, i = 1, 2, 3$ denote the intra-specific competition coefficient of X_i; $b_{ij} \geq 0, i \neq j = 1, 2, 3$, in which b_{12}, b_{13} represent the capture rates, b_{21}, b_{31} represent the growth from food, and b_{23} and b_{32} signify the competitions between predators (species 2 and 3); for each $i, j \in \{1, 2, 3\}$, $\widehat{b}_{ij}$ is either positive or in $(-b_{ii}, 0]$; r is the time delay; $(E_1(t), E_2(t), E_3(t))^\top = \Gamma^\top \mathbf{B}(t)$ with $\mathbf{B}(t) = (B_1(t), B_2(t), B_3(t))^\top$ being a vector of independent standard Brownian motions and Γ being a 3×3 matrix such that $\Gamma^\top \Gamma = (\sigma_{ij})_{3\times 3}$ is a positive definite matrix.

The model in the current setup, was considered in [33]. However, by switching the sign of a_i or $b_{ij}, i \neq j$, we can obtain a stochastic time-delay Lotka–Volterra system with the prey and the mesopredator or intermediate predator. Note that the case involving a superpredator or top predator, was studied in [35,66], and the stochastic time-delay Lotka–Volterra system with one predator and two preys was investigated in [22].

By a similar calculation as in Section 5.1, we can check that (2.2) is satisfied if we let $\mathbf{c} = \left(1, \dfrac{b_{12}}{b_{21}}, \dfrac{b_{13}}{b_{31}}\right)$ and other parameters be similarly determined as in Section 5.1. Moreover, other assumptions also hold.

We consider the equation on the boundaries $\mathcal{C}_{12+} := \{(\varphi_1, \varphi_2, 0) \in \mathcal{C}_+ : \varphi_1(s), \varphi_2(s) \geq 0\ \forall s \in [-r, 0]\}$ and $\mathcal{C}_{13+} := \{(\varphi_1, 0, \varphi_3) \in \mathcal{C}_+ : \varphi_1(s), \varphi_3(s) \geq 0,\ \forall s \in [-r, 0]\}$. If $\lambda_1(\boldsymbol{\delta}^*) > 0$, there is an invariant probability measure π_1 on $\mathcal{C}^\circ_{1+} := \{(\varphi_1, 0, 0) \in \mathcal{C}_+ : \varphi_1(s) > 0\ \forall s \in [-r, 0]\}$.

In view of Lemma 5.1, we obtain

$$\int_{\mathcal{C}^\circ_{1+}} \big(b_{11}\varphi_1(0) + \widehat{b}_{11}\varphi_1(-r)\big)\,\pi_1(d\boldsymbol{\varphi}) = a_1 - \frac{\sigma_{11}}{2}. \tag{5.10}$$

Similar to the process of getting (5.8), we obtain from (5.10) that

$$\int_{\mathcal{C}^\circ_{1+}} \varphi_1(0)\pi_1(d\boldsymbol{\varphi}) = \int_{\mathcal{C}^\circ_{1+}} \varphi_1(-r)\pi_1(d\boldsymbol{\varphi}) = \frac{a_1 - \frac{\sigma_{11}}{2}}{b_{11} + \widehat{b}_{11}}.$$

Therefore,

$$\begin{aligned} \lambda_i(\pi_1) &= \int_{\mathcal{C}^\circ_{1+}} \Big[-a_i - \frac{\sigma_{ii}}{2} + b_{i1}\varphi_1(0) - \widehat{b}_{i1}\varphi_1(-r)\Big]\pi_1(d\boldsymbol{\varphi}) \\ &= -a_i - \frac{\sigma_{ii}}{2} + \Big(a_1 - \frac{\sigma_{11}}{2}\Big)\cdot \frac{b_{i1} - \widehat{b}_{i1}}{b_{11} + \widehat{b}_{11}},\ i = 2, 3. \end{aligned}$$

In case of $\lambda_1(\boldsymbol{\delta}^*) > 0$ and $\lambda_2(\pi_1) > 0$, Theorem 2.2 implies that there is an invariant probability measure π_{12} on $\mathcal{C}^\circ_{12+}$. In view of Lemma 5.1 and (5.8), we obtain

$$\int_{\mathcal{C}^\circ_{12+}} \varphi_1(0)\pi_{12}(d\boldsymbol{\varphi}) = \int_{\mathcal{C}^\circ_{12+}} \varphi_1(-r)\pi_{12}(d\boldsymbol{\varphi}) = A_1,$$
$$\int_{\mathcal{C}^\circ_{12+}} \varphi_2(0)\pi_{12}(d\boldsymbol{\varphi}) = \int_{\mathcal{C}^\circ_{12+}} \varphi_2(-r)\pi_{12}(d\boldsymbol{\varphi}) = A_2,$$

where the pair (A_1, A_2) is the unique solution to

$$\begin{cases} a_1 - \frac{\sigma_{11}}{2} - \left(b_{11} + \widehat{b}_{11}\right) A_1 - \left(b_{12} + \widehat{b}_{12}\right) A_2 = 0, \\ -a_2 - \frac{\sigma_{22}}{2} + \left(b_{21} - \widehat{b}_{21}\right) A_1 - \left(b_{22} + \widehat{b}_{22}\right) A_2 = 0. \end{cases}$$

In this case,

$$\begin{aligned} \lambda_3(\pi_{12}) &= \int_{\mathcal{C}^\circ_{12+}} \Big[-a_3 - \frac{\sigma_{33}}{2} + \left(b_{31}\varphi_1(0) - \widehat{b}_{31}\varphi_1(-r)\right) \\ &\qquad - \left(b_{32}\varphi_2(0) + \widehat{b}_{32}\varphi_2(-r)\right)\Big]\pi_{12}(d\boldsymbol{\varphi}) \\ &= -a_3 - \frac{\sigma_{33}}{2} + \left(b_{31} - \widehat{b}_{31}\right) A_1 - \left(b_{32} + \widehat{b}_{32}\right) A_2. \end{aligned}$$

Similarly, if $\lambda_1(\boldsymbol{\delta}^*) > 0$ and $\lambda_3(\pi_1) > 0$, by Theorem 2.2, there is an invariant probability measure π_{13} on $\mathcal{C}^\circ_{13+}$ and

$$\begin{aligned} \lambda_2(\pi_{13}) &= \int_{\mathcal{C}^\circ_{13+}} \Big[-a_2 - \frac{\sigma_{22}}{2} + \left(b_{21}\varphi_1(0) - \widehat{b}_{21}\varphi_1(-r)\right) \\ &\qquad - \left(b_{23}\varphi_3(0) + \widehat{b}_{23}\varphi_3(-r)\right)\Big]\pi_{13}(d\boldsymbol{\varphi}) \\ &= -a_2 - \frac{\sigma_{22}}{2} + \left(b_{21} - \widehat{b}_{21}\right) \widehat{A}_1 - \left(b_{32} + \widehat{b}_{23}\right) \widehat{A}_3, \end{aligned}$$

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

$$\begin{cases} a_1 - \frac{\sigma_{11}}{2} - \left(b_{11} + \widehat{b}_{11}\right) \widehat{A}_1 - \left(b_{13} + \widehat{b}_{13}\right) \widehat{A}_3 = 0, \\ -a_3 - \frac{\sigma_{33}}{2} + \left(b_{31} - \widehat{b}_{31}\right) \widehat{A}_1 - \left(b_{33} + \widehat{b}_{33}\right) \widehat{A}_3 = 0. \end{cases}$$

Because of the ergodic decomposition theorem, every invariant probability measure on $\partial\mathcal{C}_+$ is a convex combination of $\delta^*, \pi_1, \pi_{12}, \pi_{13}$ (when these measures exist). As a consequence, some computations for the Lyapunov exponents with respect to a convex combination of these ergodic measures together with an application of Theorem 2.2 yield that there exists a unique invariant probability measure in $\mathcal{C}^\circ_+$ if one of the following conditions is satisfied:

- $\lambda_1(\boldsymbol{\delta}^*) > 0$, $\lambda_2(\pi_1) > 0$, $\lambda_3(\pi_1) < 0$ and $\lambda_3(\pi_{12}) > 0$.
- $\lambda_1(\boldsymbol{\delta}^*) > 0$, $\lambda_2(\pi_1) < 0$, $\lambda_3(\pi_1) > 0$ and $\lambda_2(\pi_{13}) > 0$.
- $\lambda_1(\boldsymbol{\delta}^*) > 0$, $\lambda_2(\pi_1) > 0$, $\lambda_3(\pi_1) > 0$, $\lambda_3(\pi_{12}) > 0$, and $\lambda_2(\pi_{13}) > 0$.

The above assertions generalize the results in [33]. Moreover, if we switch the sign of a_i or $b_{ij}, i \neq j$, we obtain another modifications of Lotka–Volterra prey–predator equation as we mentioned at the beginning of this section with modification of the above characterization, which improve the results in [22,35,66].

Confining our analysis to $\mathcal{C}_{12+}$ (this describes the evolution of one predator and its prey), we get

$$\begin{cases} dX_1(t) = X_1(t)\Big\{ a_1 - b_{11}X_1(t) - \widehat{b}_{11}X_1(t-r) - b_{12}X_2(t) \\ \qquad\qquad\qquad\qquad\qquad -\widehat{b}_{12}X_2(t-r) \Big\}dt + X_1(t)dE_1(t), \\ dX_2(t) = X_2(t)\Big\{ -a_2 + b_{21}X_1(t) + \widehat{b}_{21}X_1(t-r) - b_{22}X_2(t) \\ \qquad\qquad\qquad\qquad\qquad -\widehat{b}_{22}X_2(t-r) \Big\}dt + X_2(t)dE_2(t). \end{cases} \tag{5.11}$$

This further leads to that if $\lambda_1(\boldsymbol{\delta}^*) > 0, \lambda_2(\pi_1) > 0$, there exists a unique invariant probability measure of (5.11) on $\mathcal{C}^\circ_{12+}$, which improves the results in [38].

5.3. Stochastic delay replicator equations

In evolutionary game theory, originally, a replicator equation is a deterministic monotone, nonlinear, and non-innovative game dynamic system. Such a deterministic system has been expanded to systems with stochastic perturbations. In this section, we consider the replicator dynamics for a game with n strategies, involving social-type time delay (see, e.g., [28] for details of such delays) and white noise perturbation. The system of interest can be expressed as

$$\begin{cases} dx_i(t) = x_i(t)\left(f_i(\mathbf{x}(t-r)) - \dfrac{1}{X}\displaystyle\sum_{j=1}^n x_j(t)f_j(\mathbf{x}(t-r)) \right)dt \\ \qquad\qquad +x_i(t)\left(\sigma_i dB_i(t) - \dfrac{1}{X}\displaystyle\sum_{j=1}^n \sigma_j x_j dB_j(t) \right);\ i = 1,\dots,n, \\ \mathbf{x}(s) = \mathbf{x}_0(s);\ t \in [-r,0], \end{cases} \tag{5.12}$$

where X is the size of the populations; $x_i(t)$ is the portion of population that has selected the ith strategy and the distribution of the whole population among the strategy; the fitness functions $f_i(\cdot) : \mathbb{R}^n_+ \to \mathbb{R}, i = 1,\dots,n$ are the payoffs obtained by the individuals playing the ith strategy; r is the time delay; and $\mathbf{x}_0(s) \in \Delta_X := \{\mathbf{x} \in \mathbb{R}^n_+ : \sum_{i=1}^n x_i = X\}$ for all $s \in [-r,0]$ is the initial value.

The replicator equation was introduced in 1978 by Taylor and Jonker in [59]. Since then significant contributions have been made in biology [26,49], economics [64], and optimization and control for a variety of systems [6,50,52,60]. Much attention has been devoted to studying their properties. For instance, when $f_i(\cdot) : \mathbb{R}^n_+ \to \mathbb{R}, i = 1,\dots,n$ are linear mappings, Eq. (5.12) without time delay was studied in [25,27]. Moreover, the deterministic version of Eq. (5.12) was studied in [28,51].

By a similar argument as in [51,64], we can show that Δ_X remains invariant a.s. As a consequence, our assumptions are verified. Hence, our results in Theorem 2.2 hold for (5.12). To demonstrate, for better visualization, we apply our results to some low-dimensional systems.

First, we consider Eq. (5.12) in case of two dimensions. Define

$$\begin{aligned}
\mathcal{C}_+^X &:= \{(\varphi_1, \varphi_2) : \varphi_1(s) + \varphi_2(s) = X \text{ and } \varphi_1(s), \varphi_2(s) \geq 0 \text{ for all } s \in [-r, 0]\},\\
\partial\mathcal{C}_+^X &:= \{(\varphi_1, \varphi_2) \in \mathcal{C}_+^X : \|\varphi_1\| = 0 \text{ or } \|\varphi_2\| = 0\},\\
\mathcal{C}_+^{X,\circ} &:= \{(\varphi_1, \varphi_2) \in \mathcal{C}_+^X : \varphi_1(s), \varphi_2(s) > 0 \text{ for all } s \in [-r, 0]\}.
\end{aligned}$$

In this case, it is clear that there are two invariant probability measures on the boundary $\partial\mathcal{C}_+^X$, which are $\boldsymbol{\delta}_1$ and $\boldsymbol{\delta}_2$ concentrating on $(X, 0)$ and $(0, X)$, respectively, where 0, X are understood to be constant functions. We have

$$\lambda_1(\boldsymbol{\delta}_2) = f_1((0, X)) - f_2((0, X)) - \frac{\sigma_1^2 + \sigma_2^2}{2}, \tag{5.13}$$

$$\lambda_2(\boldsymbol{\delta}_1) = f_2((X, 0)) - f_1((X, 0)) - \frac{\sigma_1^2 + \sigma_2^2}{2}. \tag{5.14}$$

By Theorem 2.2, in case of (5.12) of 2-dimensional systems, if $\lambda_1(\boldsymbol{\delta}_2) > 0$ and $\lambda_2(\boldsymbol{\delta}_1) > 0$, there is a unique invariant probability measure of (5.12) on $\mathcal{C}_+^{X,\circ}$.

Next, we consider (5.12) in three dimensions. Similarly, we also define the following set

$$\begin{aligned}
\mathcal{C}_+^X := \{(\varphi_1, \varphi_2, \varphi_3) :&\varphi_1(s) + \varphi_2(s) + \varphi_3(s) = X\\
&\text{and } \varphi_1(s), \varphi_2(s), \varphi_3(s) \geq 0 \text{ for all } s \in [-r, 0]\},
\end{aligned}$$

$$\begin{aligned}
\partial\mathcal{C}_+^X &:= \mathcal{C}_{12+}^X \cup \mathcal{C}_{23+}^X \cup \mathcal{C}_{13+}^X,\\
\mathcal{C}_{ij+}^X &:= \{(\varphi_1, \varphi_2, \varphi_3) \in \mathcal{C}_+^X : \|\varphi_k\| = 0, k \neq i, j\}, \text{ for } i \neq j \in \{1, 2, 3\},\\
\mathcal{C}_+^{X,\circ} &:= \{(\varphi_1, \varphi_2, \varphi_3) \in \mathcal{C}_+^X : \varphi_1(s), \varphi_2(s), \varphi_3(s) > 0 \text{ for all } s \in [-r, 0]\}.
\end{aligned}$$

Denote by $\boldsymbol{\delta}_1, \boldsymbol{\delta}_2, \boldsymbol{\delta}_3$ the invariant probability measures on the boundary $\partial\mathcal{C}_+^X$ of (5.12), concentrating on $(X, 0, 0)$, $(0, X, 0)$, and $(0, 0, X)$, respectively. We have

$$\begin{aligned}
\lambda_i(\boldsymbol{\delta}_1) &= f_i((X, 0, 0)) - f_1((X, 0, 0)) - \frac{\sigma_1^2 + \sigma_i^2}{2}, i = 2, 3,\\
\lambda_j(\boldsymbol{\delta}_2) &= f_j((0, X, 0)) - f_2((0, X, 0)) - \frac{\sigma_2^2 + \sigma_j^2}{2}, j = 1, 3,
\end{aligned}$$

and

$$\lambda_k(\boldsymbol{\delta}_3) = f_k((0, 0, X)) - f_3((0, 0, X)) - \frac{\sigma_3^2 + \sigma_k^2}{2}, k = 1, 2.$$

If $\max_{j=1,3} \lambda_j(\boldsymbol{\delta}_2) > 0$ and $\max_{k=1,2} \lambda_k(\boldsymbol{\delta}_3) > 0$, there is a unique invariant probability measure on $\mathcal{C}_{23+}^X$, denoted by π_{23}. When π_{23} exists, we have

$$\begin{aligned}
\lambda_1(\pi_{23}) = -\frac{\sigma_1^2}{2} + \int_{\mathcal{C}_{23+}^X} \Bigg(f_1(\boldsymbol{\varphi}) &- \frac{2X\varphi_2(0) f_2(\boldsymbol{\varphi}) + \sigma_2^2\varphi_2^2(0)}{X^2}\\
&- \frac{2X\varphi_3(0) f_3(\boldsymbol{\varphi}) + \sigma_3^2\varphi_3^2(0)}{X^2} \Bigg) \pi_{23}(d\boldsymbol{\varphi}).
\end{aligned}$$

By Lemma 5.1 and $\lambda_2(\pi_{23}) = \lambda_3(\pi_{23}) = 0$, we have

$$\begin{aligned}\int_{\mathcal{C}^X_{23+}} &\left(\frac{2X\varphi_2(0)f_2(\boldsymbol{\varphi}) + \sigma_2^2\varphi_2^2(0)}{2X^2} + \frac{2X\varphi_3(0)f_3(\boldsymbol{\varphi}) + \sigma_3^2\varphi_3^2(0)}{2X^2}\right)\pi_{23}(d\boldsymbol{\varphi})\\ &= \frac{\sigma_2^2}{2} + \int_{\mathcal{C}^X_{23+}} f_2(\boldsymbol{\varphi})\pi_{23}(d\boldsymbol{\varphi})\\ &= \frac{\sigma_3^2}{2} + \int_{\mathcal{C}^X_{23+}} f_3(\boldsymbol{\varphi})\pi_{23}(d\boldsymbol{\varphi}).\end{aligned}$$

As a result,

$$\begin{aligned}\lambda_1(\pi_{23}) &= -\frac{\sigma_1^2 + \sigma_2^2}{2} + \int_{\mathcal{C}^X_{23+}} (f_1(\boldsymbol{\varphi}) - f_2(\boldsymbol{\varphi}))\,\pi_{23}(d\boldsymbol{\varphi})\\ &= -\frac{\sigma_1^2 + \sigma_3^2}{2} + \int_{\mathcal{C}^X_{23+}} (f_1(\boldsymbol{\varphi}) - f_3(\boldsymbol{\varphi}))\,\pi_{23}(d\boldsymbol{\varphi}).\end{aligned}$$

The conditions to guarantee the existence of the unique invariant probability measure π_{12}, π_{13} on the boundary $\mathcal{C}^X_{12+}, \mathcal{C}^X_{13+}$ are similarly obtained and $\lambda_2(\pi_{13}), \lambda_3(\pi_{12})$ can be computed similar to $\lambda_1(\pi_{23})$. Therefore, we have the following classification for the long-run solution of (5.12) in three dimensions as follows. System (5.12) admits a unique invariant probability measure on $\mathcal{C}^{X,\circ}_+$ if

- $\max_{i=2,3}\lambda_i(\boldsymbol{\delta}_1) > 0$, $\max_{j=1,3}\lambda_j(\boldsymbol{\delta}_2) > 0$, $\max_{k=1,2}\lambda_k(\boldsymbol{\delta}_3) > 0$ and $\lambda_1(\pi_{23}) > 0$, $\lambda_2(\pi_{13}) > 0$, $\lambda_3(\pi_{12}) > 0$.

The (explicit) condition for persistence of (5.12) in n-dimensions is more complex. However, our results (Theorem 2.2) still hold and will be computable in practice under suitable conditions. Moreover, if $r = 0$ (i.e., there is no time delay) and $f_i(\cdot)$, $i = 1, \dots, n$ are linear, the condition of the persistence of (5.12) in this section is equivalent to results in [25,27].

5.4. Stochastic delay epidemic SIR models

The SIR model is one of the basic building blocks of compartmental models, from which many infectious disease models are derived. The model consists of three compartments, S for the number of susceptible, I for the number of infectious, and R for the number of recovered (or immune). First introduced by Kermack and McKendrick in [29,30], the models are deemed effective to depict the spread of many common diseases with permanent immunity such as rubella, whooping cough, measles, and smallpox. A variety of modifications of original equation are introduced due to the complexity of environment. Much attention has been devoted to analyzing the behavior of these systems; for example, see [12,15] and the references therein. In this section, we study the stochastic epidemic SIR models with time delay.

To start, we consider the equation with linear incidence rate of the following form

$$\begin{cases} dS(t) = (a - b_1 S(t) - c_1 I(t)S(t) - c_2 I(t)S(t-r))\,dt + S(t)dE_1(t),\\ dI(t) = (-b_2 I(t) + c_1 I(t)S(t) + c_2 I(t)S(t-r))\,dt + I(t)dE_2(t),\end{cases} \tag{5.15}$$

where $S(t)$ is the density of susceptible individuals, $I(t)$ is the density of infected individuals, $a > 0$ is the recruitment rate of the population, $b_i > 0$, $i = 1, 2$ are the death rates, $c_i > 0$, $i = 1, 2$ are the incidence rates, r is the delayed time, $(E_1(t), E_2(t))^\top = \Gamma^\top \mathbf{B}(t)$

with $\mathbf{B}(t) = (B_1(t), B_2(t))^\top$ being a vector of independent standard Brownian motions, and Γ being a 2×2 matrix such that $\Gamma^\top \Gamma = (\sigma_{ij})_{2\times 2}$ is a positive definite matrix. It is well-known that the dynamics of recovered individuals have no effect on the disease transmission dynamics and that is why we only consider the dynamics of $S(t), I(t)$ in (5.15).

Although Eq. (5.15) does not have the exact form as in (2.1), we can use the same idea and the same method to obtain similar results. First, we consider the equation on the boundary $\{(\varphi_1, 0) : \varphi_1(s) \geq 0\ \forall s \in [-r, 0]\}$ and let $\widehat{S}(t)$ be the solution of the equation on this boundary as follows

$$d\widehat{S}(t) = \big(a - b_1\widehat{S}(t)\big)\, dt + \widehat{S}(t) dE_1(t). \tag{5.16}$$

Since the drift coefficient of this equation is negative if $\widehat{S}(t)$ is sufficiently large and positive, if $\widehat{S}(t)$ is sufficiently small, we can show that there is a unique invariant probability measure π of (5.15) on $\mathcal{C}^\circ_{1+} := \{(\varphi_1, 0) : \varphi_1(s) > 0\ \forall s \in [-r, 0]\}$. On the other hand, since $\lambda_2(\boldsymbol{\delta}^*) = -b_2 - \dfrac{\sigma_{22}}{2} < 0$, there is no invariant probability measure in $\mathcal{C}^\circ_{2+} := \{(0, \varphi_2) : \varphi_2(s) > 0; \forall s \in [-r, 0]\}$.

Hence, we define the following threshold

$$\lambda(\pi) = -b_2 - \frac{\sigma_{22}}{2} + \int_{\mathcal{C}^\circ_{1+}} (c_1\varphi_1(0) + c_2\varphi_1(-r))\, \pi(d\boldsymbol{\varphi}), \tag{5.17}$$

whose sign will be able to characterize the permanence and extinction. As an application of Lemma 5.1, we get

$$\int_{\mathcal{C}^\circ_{1+}} \varphi_1(0)\pi(d\boldsymbol{\varphi}) = \frac{a}{b_1}. \tag{5.18}$$

By (5.8), we have that

$$\int_{\mathcal{C}^\circ_{1+}} \varphi_1(-r)\pi(d\boldsymbol{\varphi}) = \int_{\mathcal{C}^\circ_{1+}} \varphi_1(0)\pi(d\boldsymbol{\varphi}) = \frac{a}{b_1}.$$

Therefore, under this condition, we obtain from (5.17) and (5.18) that

$$\lambda(\pi) = -b_2 - \frac{\sigma_{22}}{2} + \frac{a(c_1 + c_2)}{b_1}.$$

Using the same idea and techniques, it is possible to obtain similar results to Theorem 2.2 for Eq. (5.15). We have that if $\lambda(\pi) > 0$, (5.15) has a unique invariant probability measure in $\mathcal{C}^\circ_+$. This characterization is equivalent to the result in [34,36].

In the above, we consider the linear incidence to make our computations be more explicit. The characterizations still hold for the following stochastic delay SIR epidemic model with general incidence rate

$$\begin{cases} dS(t) = \big(a - b_1 S(t) - I(t) f_1\big(S(t), S(t-r), I(t), I(t-r)\big)\big)\, dt + S(t) dE_1(t), \\ dI(t) = \big(-b_2 I(t) + I(t) f_2\big(S(t), S(t-r), I(t), I(t-r)\big)\big)\, dt + I(t) dE_2(t), \end{cases} \tag{5.19}$$

where $f_i : \mathbb{R}^4 \to \mathbb{R}, i = 1, 2$ are the incidence functions satisfying

- $f_1(0,0,i_1,i_2) = f_2(0,0,i_1,i_2) = 0$.
- there exists some $\kappa \in (0,\infty)$ such that for all $\boldsymbol{\varphi} \in \mathcal{C}_+$

$$\begin{aligned} f_2\left(\varphi_1(0), \varphi_1(-r), \varphi_2(0), \varphi_2(-r)\right) &\le \kappa f_1\big(\varphi_1(0), \varphi_1(-r), \varphi_2(0), \varphi_2(-r)\big) \\ &\le \kappa^2\left(1 + |\boldsymbol{\varphi}(0)| + |\boldsymbol{\varphi}(-r)|\right). \end{aligned}$$

- $f_2(s_1, s_2, i_1, i_2)$ is non-decreasing in s_1, s_2 and is non-increasing in i_1, i_2.

Almost all incidence functions used in the literature, e.g., linear functional response, Holling type II functional response, Beddington–DeAngelis functional response, etc., satisfy the above conditions. In the general case, the system has a unique invariant probability measure in $\mathcal{C}_+^\circ$ if $\lambda(\pi) > 0$, where $\lambda(\pi)$ is defined as follows

$$\lambda(\pi) = -b_2 - \frac{\sigma_{22}}{2} + \int_{\mathcal{C}_{1+}^\circ} f_2\big(\varphi_1(0), \varphi_1(-r), \varphi_2(0), \varphi_2(-r)\big)\pi(d\boldsymbol{\varphi}),$$

where $\boldsymbol{\varphi} = (\varphi_1, \varphi_2)$ and π is the invariant probability measure of (5.16). These results significantly generalize and improve that of [7,14,37,40].

5.5. Stochastic delay chemostat models

A chemostat is a bio-reactor. In a chemostat, fresh medium is continuously added, and culture liquid containing left-over nutrients, metabolic end products, and microorganisms are continuously removed at the same rate to keep a constant culture volume. The chemostat model is based on a technique introduced by Novick and Szilard in [48] and plays an important role in microbiology, biotechnology, and population biology. This section is devoted to studying a model of n-microbial populations competing for a single nutrient in a chemostat with delay in uptake conversion and under effects of white noises. Precisely, the model is described by the following system of stochastic functional differential equations

$$\begin{cases} dS(t) = \left(1 - S(t) + aS(t-r) - \displaystyle\sum_{i=1}^{n} x_i(t)p_i(S(t))\right)dt + S(t)dE_0(t), \\ dx_i(t) = x_i(t)\left(p_i(S(t-r)) - 1\right)dt + x_i(t)dE_i(t),\ i = 1,\dots,n, \end{cases} \tag{5.20}$$

where $S(t)$ is the concentration of nutrient at time t; $0 \le a < 1$ is a constant; $x_i(t), i = 1,\dots,n$ are the concentrations of the competing microbial populations; $p_i(S), i = 1,\dots,n$ are the density-dependent uptakes of nutrient by population x_i; r is the delayed time; and $(E_0(t),\dots,E_n(t))^\top = \Gamma^\top \mathbf{B}(t)$ with $\mathbf{B}(t) = (B_0(t),\dots,B_n(t))^\top$ being a vector of independent standard Brownian motions and Γ being a $(n+1)\times(n+1)$ matrix such that $\Gamma^\top\Gamma = (\sigma_{ij})_{(n+1)\times(n+1)}$ is a positive definite matrix. Moreover, in this section, $\mathcal{C} := \mathcal{C}([-r,0],\mathbb{R}^{n+1})$ instead of $\mathcal{C}([-r,0],\mathbb{R}^n)$. The deterministic version of (5.20) is studied and the long-time behavior is characterized in [19,21,65]. Recently, much attention is devoted to studying the related stochastic systems; see [57,58,67].

It is similar to Section 5.4, if we assume that $p_i : \mathbb{R} \to \mathbb{R}, i = 1,\dots,n$ satisfying non-decreasing and bounded properties and $p_i(0) = 0$, then our Assumptions hold. Therefore, our results in this paper can be applied to (5.20).

Before obtaining the results in multi-dimensional systems, we consider $n = 1, 2$. If $n = 1$, there is only one population x_1 together with the nutrient $S(t)$. Similar to Section 5.4, there is no invariant probability measure of (S_t, x_{1t}) in $\mathcal{C}_{1+}^\circ := \{(0,\varphi_1) \in \mathcal{C}_+ : \varphi_1(s) > 0, \forall s \in [-r,0]\}$, where x_{1t} is the memory segment function of $x_1(t)$. Moreover, there is a unique invariant probability measure π_0 in $\mathcal{C}^\circ_{0+} := \{(\varphi_0, 0) \in \mathcal{C}_+ : \varphi_0(s) > 0, \forall s \in [-r, 0]\}$. Hence, it is easy to see that for any invariant probability measure π in $\partial\mathcal{C}_+$, we have

$$\lambda_1(\pi) = \lambda_1(\pi_0) = -1 - \frac{\sigma_{11}}{2} + \int_{\mathcal{C}^\circ_{0+}} p_1(\varphi_0(-r))\pi_0(d\boldsymbol{\varphi}).$$

By applying our result, if $\lambda_1(\pi_0) > 0$ then (S_t, x_{1t}) admits a unique invariant probability measure in $\mathcal{C}^\circ_+$.

We next reveal the characterization of the longtime behavior in the case $n = 2$, which is similar to the case of $n = 1$. There is no invariant probability measure in $\mathcal{C}^\circ_{i+} := \{(0, \varphi_1, \varphi_2) \in \mathcal{C}_+ : \|\varphi_j\| = 0, j \neq i$ and $\varphi_i(s) > 0, \forall s \in [-r, 0]\}$, and there is a unique measure π_0 in $\mathcal{C}^\circ_{0+} := \{(\varphi_0, 0, 0) \in \mathcal{C}_+ : \varphi_0(s) > 0, \forall s \in [-r, 0]\}$. As characterized in the case $n = 1$, if $\lambda_i(\pi_0) > 0$, where

$$\lambda_i(\pi_0) = -1 - \frac{\sigma_{ii}}{2} + \int_{\mathcal{C}^\circ_{0+}} p_i(\varphi_0(-r))\pi_0(d\boldsymbol{\varphi}),\ i = 1, 2,$$

then there is a unique invariant probability measure π_{0i} in $\mathcal{C}^\circ_{0i+} := \{(\varphi_0, \varphi_1, \varphi_2) \in \mathcal{C}_+ : \|\varphi_j\| = 0, j \neq i$ and $\varphi_0(s), \varphi_i(s) > 0, \forall s \in [-r, 0]\}$. Hence, let

$$\lambda_j(\pi_{0i}) = -1 - \frac{\sigma_{jj}}{2} + \int_{\mathcal{C}^\circ_{0+}} p_j(\varphi_0(-r))\pi_{0i}(d\boldsymbol{\varphi}),\ j \neq i.$$

The persistence is classified as follows. The (S_t, x_{1t}, x_{2t}) admits a unique invariant probability measure in $\mathcal{C}^\circ_+$ if $\lambda_1(\pi_0) > 0$, $\lambda_2(\pi_0) > 0$, $\lambda_1(\pi_{02}) > 0$, and $\lambda_2(\pi_{01}) > 0$.

The two examples in low dimension ($n = 1, 2$) provide a scheme to construct recursively the characterization of the longtime behavior of (5.20) in higher dimensions. It is difficult to show concretely in case of general functions $p_i(\cdot)$, but it is computable in certain examples. These classifications improve the results in [58,67].

Remark 5. In fact, in all the examples in Sections 5.1–5.5, similar results can be obtained for multi-delays or distributed delays. We used a single delay in this Section for simplifying the notation and calculations so as to present the main ideas without notation complication. On the other hand, if $r = 0$, i.e., there is no time delay, the above results are consistent with and/or improve the existing results in the literature.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] A. Bahar, X. Mao, Stochastic delay Lotka–Volterra model, J. Math. Anal. Appl. 292 (2004) 364–380.
[2] J. Bao, G. Yin, C. Yuan, Asymptotic Analysis for Functional Stochastic Differential Equations, SpringerBriefs Math. Springer, 2016.
[3] M. Benaïm, Stochastic persistence, 2018, preprint.
[4] M. Benaïm, C. Lobry, Lotka–Volterra with randomly fluctuating environments or how switching between beneficial environments can make survival harder, Ann. Appl. Probab 26 (2016) 3754–3785.
[5] M. Benaïm, E. Strickler, Random switching between vector fields having a common zero, Ann. Appl. Probab 29 (2019) 326–375.
[6] I. Bomze, M. Pelillo, V. Stix, Approximating the maximum weight clique using replicator dynamics, IEEE Trans. Neural Network 11 (2000) 1228–1241.

[7] G. Chen, T. Li, Stability of stochastic delayed SIR model, Stoch. Dyn 9 (2009) 231–252.
[8] P.L. Chesson, S. Ellner, Invasibility and stochastic boundedness in monotonic competition models, J. Math. Biol. 27 (1989) 117–138.
[9] R. Cont, D.-A. Fournié, Change of variable formulas for non-anticipative functionals on path space, J. Funct. Anal 259 (2010) 1043–1072.
[10] R. Cont, D.-A. Fournié, Functional Itô calculus and stochastic integral representation of martingales, Ann. Probab 41 (2013) 109–133.
[11] Cushing, J.M, Integrodifferential Equations and Delay Models in Population Dynamics, Vol. 20, Springer Science & Business Media, 2013.
[12] N.T. Dieu, D.H. Nguyen, N.H. Du, G. Yin, Classification of asymptotic behavior in a stochastic SIR model, SIAM J. Appl. Dyn. Syst 15 (2016) 1062–1084.
[13] N.H. Du, N.T. Dieu, N.N. Nhu, Conditions for permanence and ergodicity of certain SIR epidemic models, Acta Appl. Math 160 (2019) (2019) 81–99.
[14] N.H. Du, N.N. Nguyen, Permanence and extinction of certain stochastic SIR models perturbed by a complex type of noises, Appl. Math. Lett. 64 (2017) 223–230.
[15] N.H. Du, N.N. Nguyen, Permanence and extinction for the stochastic SIR epidemic model, J. Differential Equations 269 (2020) 9619–9652.
[16] N.H. Du, D.H. Nguyen, G. Yin, Conditions for permanence and ergodicity of certain stochastic predator–prey models, J. Appl. Probab 53 (2016) 187–202.
[17] B. Dupire, Functional Itô Calculus, Bloomberg Portfolio Research Paper No. 2009-04-FRONTIERS Available at SSRN: http://ssrn.com/abstract=1435551 or http://dx.doi.org/10.2139/ssrn.1435551.
[18] S.N. Eithier, T.G. Kurtz, Markov Processes: Characterization and Convergence, vol. 282, Wiley, New York, 2009.
[19] S.F. Ellermeyer, Competition in the chemostat: global asymptotic behavior of a model with delayed response in growth, SIAM J. Appl. Math 54 (1994) 456–465.
[20] S.N. Evans, A. Hening, S.J. Schreiber, Protected polymorphisms and evolutionary stability of patch-selection strategies in stochastic environments, J. Math. Biol. 71 (2015) 325–359.
[21] H.I. Freedman, J.W.-H. So, P. Waltman, Coexistence in a model of competition in the chemostat incorporating discrete delays, SIAM J. Appl. Math 49 (1989) 859–870.
[22] J. Geng, M. Liu, Y. Zhang, Stability of a stochastic one-predator-two-prey population model with time delays, Commun. Nonlinear Sci. Numer. Simul. 53 (2017) 65–82.
[23] M. Hairer, J.C. Mattingly, M. Scheutzow, Asymptotic coupling and a general form of harris' theorem with applications to stochastic delay equations, Probab. Theory Related Fields 149 (2011) 223–259.
[24] A. Hening, D.H. Nguyen, Coexistence and extinction for stochastic Kolmogorov systems, Ann. Appl. Probab 28 (2018) 1893–1942.
[25] J. Hofbauer, L.A. Imhof, Time averages, recurrence and transience in the stochastic replicator dynamics, Ann. Appl. Probab 19 (2009) 1347–1368.
[26] J. Hofbauer, K. Sigmund, Evolutionary Games and Population Dynamics, Cambridge Univ. Press, 1998.
[27] L.A. Imhof, The long-run behavior of the stochastic replicator dynamics, Ann. Appl. Probab 15 (2005) 1019–1045.
[28] Alboszta J., Miekisz J., Stability of evolutionarily stable strategies in discrete replicator dynamics with time delay, J. Theoret. Biol. 231 (2004) 175–179.
[29] W.O. Kermack, A.G. McKendrick, Contributions to the mathematical theory of epidemics–I, Proc. Royal Soc. London Ser. A 115 (1927) 700–721.
[30] W.O. Kermack, A.G. McKendrick, Contributions to the mathematical theory of epidemics–II, Proc. Royal Soc. London Ser. A 138 (1932) 55–83.
[31] R. Khasminskii, Stochastic Stability of Differential Equations, second ed., Springer, Heidelberg, 2012.
[32] M. Liu, C. Bai, Optimal harvesting of a stochastic delay competitive model, Disc. Contin. Dyn. Syst. Ser. B 22 (2017) 1493–1508.
[33] M. Liu, C. Bai, Y. Jin, Population dynamical behavior of a two-predator one-prey stochastic model with time delay, Discrete Contin. Dyn. Syst 37 (2017) 2513–2538.
[34] Q. Liu, Q. Chen, D. Jiang, The threshold of a stochastic delayed SIR epidemic model with temporary immunity, Phys. A 450 (2016) 115–125.
[35] M. Liu, M. Fan, Stability in distribution of a three-species stochastic cascade predator–prey system with time delays, IMA J. Appl. Math 82 (2017) 396–423.

[36] Q. Liu, D. Jiang, The threshold of a stochastic delayed SIR epidemic model with vaccination, Phys. A 461 (2016) 140–147.
[37] Q. Liu, D. Jiang, N. Shi, T. Hayat, A. Alsaedi, Asymptotic behaviors of a stochastic delayed SIR epidemic model with nonlinear incidence, Commun. Nonlinear Sci. Numer. Simul. 40 (2016) 89–99.
[38] M. Liu, H. Qiu, K. Wang, A remark on a stochastic predator–prey system with time delays, Appl. Math. Lett. 26 (2013) 318–323.
[39] A.J. Lotka, Elements of Physical Biology, William and Wilkins, Baltimore, 1925.
[40] M. Mahrouf, J. Adnani, N. Yousfi, Stability analysis of a stochastic delayed SIR epidemic model with general incidence rate, Appl. Anal 97 (2018) 2113–2121.
[41] X. Mao, Stochastic Differential Equations and their Applications, Horwood Publishing Limited, Chichester, 1997.
[42] X. Mao, C. Yuan, Stochastic Differential Equations with Markovian Switching, Imperial College Press, London, 2006.
[43] S.-E.A. Mohammed, Stochastic Functional Differential Equations, Longman Scientific and Technical, 1986.
[44] D.H. Nguyen, N.N. Nguyen, G. Yin, General nonlinear stochastic systems motivated by chemostat models: Complete characterization of long-time behavior, optimal controls, and applications to wastewater treatment, Stochastic Process. Appl 130 (2020) 4608–4642.
[45] D.H. Nguyen, N.N. Nguyen, G. Yin, Stochastic functional Kolmogorov equations II: Extinction, J. Diff. Equ. 294 (2021) 1–39.
[46] D.H. Nguyen, G. Yin, Stability of stochastic functional differential equations with regime switching: Analysis using Dupire's functional Itô Formula, Potential Anal. 53 (2020) 247–265.
[47] D.H. Nguyen, G. Yin, C. Zhu, Long-term analysis of a stochastic SIRS model with general incidence rates, SIAM J. Appl. Math 80 (2020) (2020) 814–838.
[48] A. Novick, L. Szilard, Description of the chemostat, Science 112 (1950) 715–716.
[49] M.A. Nowak, K. Sigmund, Evolutionary dynamics of biological games, Science 303 (2004) 793–799.
[50] G. Obando, A. Pantoja, N. Quijano, Building temperature control based on population dynamics, IEEE Trans. Control Syst. Tech 22 (2014) 404–412.
[51] G. Obando, J.I. Poveda, N. Quijano, Replicator dynamics under perturbations and time delays, Math. Control Sig. Syst 28 (2016) 32.
[52] E. Ramírrez-Llanos, N. Quijano, A population dynamics approach for the water distribution problem, Internat. J. Control 83 (2010) 1947–1964.
[53] M. Reiß, M. Riedle, O. van Gaans, Delay differential equations driven by levy processes: stationarity and Feller properties, Stochastic Process. Appl 116 (2006) 1409–1432.
[54] G. Roth, S.J. Schreiber, Persistence in fluctuating environments for interacting structured populations, J. Math. Biol. 69 (2014) 1267–1317.
[55] S.J. Schreiber, Persistence for stochastic difference equations: a mini-review, J. Difference Equ. Appl 18 (2012) 1381–1403.
[56] S.J. Schreiber, M. Benaïm, K.A.S. Atchadé, Persistence in fluctuating environments, J. Math. Biol. 62 (2011) 655–683.
[57] S. Sun, X. Zhang, Asymptotic behavior of a stochastic delayed chemostat model with nonmonotone uptake function, Phys. A 512 (2018) 38–56.
[58] S. Sun, X. Zhang, Asymptotic behavior of a stochastic delayed chemostat model with nutrient storage, J. Biol. Syst 26 (2018) 225–246.
[59] P.D. Taylor, L.B. Jonker, Evolutionary stable strategies and game dynamics, Math Bio 40 (1978) 145–156.
[60] H. Tembine, E. Altman, R. El-Azouzi, Y. Hayel, Evolutionary games in wireless networks, IEEE Trans Syst Man Cybern. Part B 40 (2010) 634–646.
[61] T.D. Tuong, N.N. Nguyen, G. Yin, Longtime behavior of a class of stochastic tumor-immune systems, Systems Control Lett 146 (2020) 104806.
[62] P.F. Verhulst, Notice sur la loi que la population poursuit dans son accroissement, Corresp. Math. Phys. 10 (1838) 113–121.
[63] V. Volterra, Variazioni e fluttuazioni del numero d'individui in specie d'animali conviventi, Mem. Acad. Lincei 2 (1926) 31–113.
[64] J. Weibull, Evolutionary Game Theory, MIT Press, Cambridge, MA, USA, 1997.

[65] G.S.K. Wolkowicz, H. Xia, Global asymptotic behavior of a chemostat model with discrete delays, SIAM J. Appl. Math 57 (1997) 1019–1043.
[66] J. Wu, Analysis of a three-species stochastic delay predator–prey system with imprecise parameters, Methodol. Comput. Appl. Probab 21 (2019) 43–67.
[67] X. Zhang, R. Yuan, The existence of stationary distribution of a stochastic delayed chemostat model, Appl. Math. Lett. 93 (2019) 15–21.