



Research article

Resilience of a stochastic generalized Lotka–Volterra model for microbiome studies

Tuan A. Phan^{1,*}, Benjamin J. Ridenhour² and Christopher H. Remien²

¹ Institute for Modeling Collaboration and Innovation, University of Idaho, Moscow, ID 83844, USA

² Department of Mathematics and Statistical Science, University of Idaho, Moscow, ID 83844, USA

* **Correspondence:** Email: tphan@uidaho.edu; tuanpa141186@gmail.com.

Abstract: Microbial communities are constantly challenged by environmental stochasticity, rendering time-series data obtained from these communities inherently noisy. Traditional mathematical models, such as the first-order multivariate autoregressive (MAR) model and the deterministic generalized Lotka–Volterra model, are no longer suitable for predicting the stability of a microbiome from its time-series data, as they fail to capture volatility in the environment. To accurately measure microbiome stability, it is imperative to incorporate stochasticity into the existing mathematical models in microbiome research. In this paper, we introduce a stochastic generalized Lotka–Volterra (SgLV) system that characterizes the temporal dynamics of a microbial community. To study this system, we developed a comprehensive theoretical framework for calculating four resilience measures based on the SgLV model. These resilience metrics effectively capture the short- and long-term behaviors of the resilience of the microbiome. To illustrate the practical application of our approach, we demonstrate the procedure for calculating the four resilience measures using simulated microbial abundance datasets. The procedural simplicity enhances its utility as a valuable tool for application in various microbial and ecological communities.

Keywords: stochastic gLV; resilience measures; instantaneous return rate; average return rate; convergence rate; asymptotic resilience

1. Introduction

Microbiome dynamics are inherently stochastic. They are not just the assemblage of living microorganisms present in a defined environment, but also encompass the entire spectrum of molecules, structural elements, metabolites, and mobile genetic elements, produced by these microorganisms, known as their theater of activity [1]. Consequently, interpreting the causal link between microbiomes and human health, a critical aspect of microbiome research [2, 3], remains challenging.

Environmental fluctuations can induce variability in interactions within a microbiome and between microbes and their environment or host. For instance, some microbial species may exhibit varying levels of sensitivity to temperature, impacting their interactions with other microbes or with the environment/host depending on the ambient temperatures. This variability introduces errors in the estimation of microbial interactions, which, in turn, affect our ability to understand the drivers of dynamics. It becomes imperative to translate microbial interactions, which govern the microbiome's population dynamics, into a high-level property of a microbiome, such as its resilience.

Theories from dynamical systems, adapted from macroecology, offer insight into understanding the dynamics of the microbiome in different environments [4]. The resilience of a microbiome, i.e., its ability to maintain and recover function in the face of perturbations, such as those caused by antibiotics or opportunistic pathogens, can be studied by employing the concept of stability in dynamic systems. Microbiomes that more readily return to stationarity (i.e., when the dynamics are captured by white noise about an equilibrium) when faced with perturbations are considered to be more resilient. Figure 1 illustrates the concept of the resilience of the microbiome. The impact of a perturbation, classified as a “pulse” or a “press,” on a microbial community is determined by the duration and magnitude of the perturbation [5]. Following a perturbation, a microbial community in a healthy state (eubiosis) may either transition to an unstable and transient state or recover to its initial state. In some instances, the perturbation might drive the community towards the formation of an alternative healthy state or an unhealthy stable state (dysbiosis) associated with disease. Recent empirical evidence supports these theoretical concepts. Studies of gut microbiota dynamics have revealed intrinsic stability patterns and catastrophic shifts following antibiotic perturbations [6, 7]. Moreover, mechanistic research has elucidated how commensals regulate host immunity, such as the suppression of retinoic acid synthesis to control IL-2 activity and prevent dysbiosis [8]. Beyond the gut, investigations across different body sites demonstrate community-specific resilience patterns, from vaginal microbiome's stability [9] to oral ecosystems [10]. Building on these findings, recent comprehensive reviews have established resilience as the microbiota's universal capacity to resist and recover from perturbations [11, 12], with longitudinal studies revealing personalized resilience features, such as the individual-specific stability observed during the Mars500 mission [13]. Therefore, understanding how a microbiome responds to perturbations is crucial to preventing disease states and promoting health.

The concept of resilience has been intensively studied in macroecological systems. In fact, the ecological term “resilience” was originally coined by Holling 1973 [14]. In this seminal article, the author defined resilience as a “measure of the persistence of ecosystems and their ability to absorb change and disturbance and still maintain the same relationships between populations or state variables.” This concept was subjected to a comprehensive examination by Pimm and Lawton 1977 [15], DeAngelis 1980 [16], and Cottingham and Carpenter 1994 [17]. The index most frequently used to calculate resilience in ecosystems is based on the eigenvalues of an ordinary differential equation system near its equilibrium. Resilience is computed as the negative reciprocal of the maximum eigenvalue, providing a measure of the time it takes the ecosystem to recover from a perturbation. Subsequently, Neubert and Caswell 1997 [18] introduced four quantities (asymptotic resilience, reactivity, and two quantities of the amplification envelope) to measure the long- and short-term responses of an ecosystem to pulse perturbations. More recently, Arnoldi et al., [19, 20] attempted to unify all previous resilience measures for a community, making them more practical by extending the analysis of ecosystem responses to perturbations on all time scales.

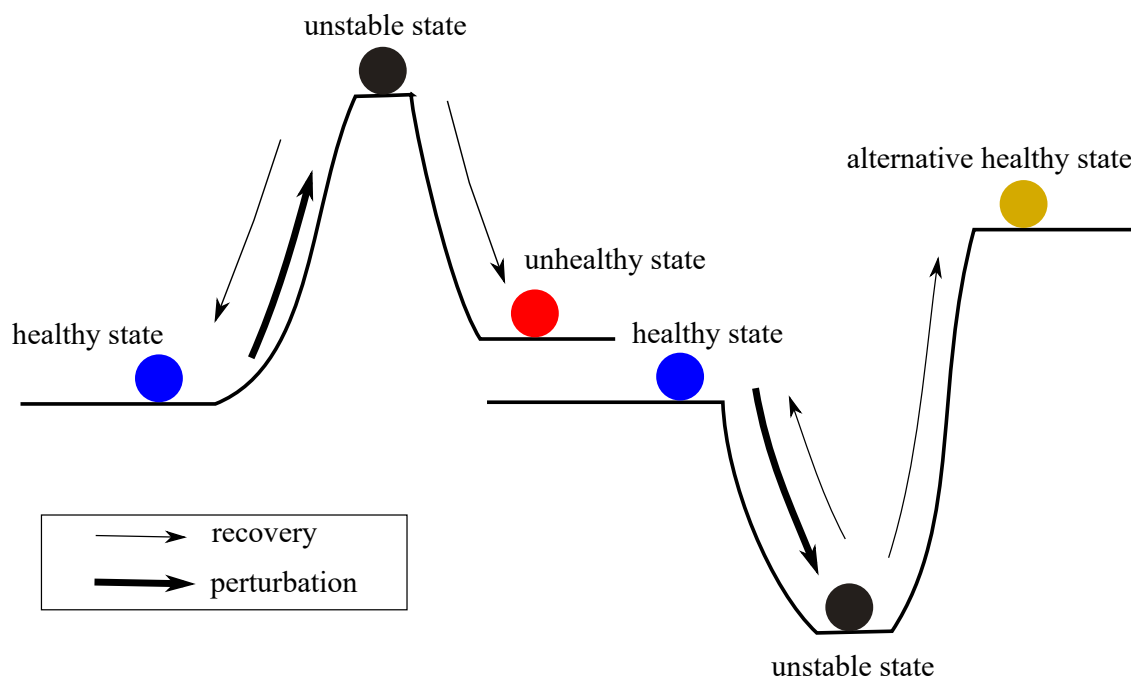


Figure 1. Conceptual illustration of the resilience of a microbiome.

All the previously mentioned definitions of resilience are based on the theory of ordinary differential equations, in which an equilibrium state is assumed to be a fixed point of a deterministic dynamic system. However, due to continuously fluctuating environments, equilibrium states may not be fixed points but fluctuating stationary distributions. To address this issue, Ives 1995 [21] proposed a definition of stochastic resilience based on a discrete-time linear stochastic model. This article used the process error, modeled as a sequence of normal random variables with a mean of 0 and a constant variance, to account for the variability in interactions between species. Subsequently, Ives et al., 2003 [22] extended this definition to three additional stability measures using a first-order discrete-time multivariate autoregressive model (MAR). Buckwar and Kelly 2014 [23] redefined the notions of asymptotic resilience, reactivity, and the amplification envelope, proposed by Neubert and Caswell 1997 [18], in terms of the second moment of a linear Ito stochastic differential equation system, with the goal of investigating the effect of persistent stochastic perturbations on a deterministic system.

Over the past decade, the conceptualization of resilience within microbiome research has evolved significantly through the integration of stochastic modeling approaches. Zaoli and Grilli 2021 [26] analyzed longitudinal human gut microbiome data, revealing that while most operational taxonomic units (OTUs) exhibit stable dynamics with fluctuations around a constant average, a minority display nonstationary behavior with abrupt transitions in abundance, potentially indicating shifts between alternative stable states. This study underscored the importance of considering stochastic fluctuations and the potential for multiple stable states in microbial communities.

Building on this foundation, Wolff et al., 2023 [27] investigated the daily dynamics of intraspecific

genetic variation in the human gut microbiome. They discovered that approximately 80% of the strains analyzed exhibited abundance fluctuations that could be predicted using a stochastic logistic model, suggesting that the stability and macroecological properties of the human gut microbiome emerge at the strain level. Hayashi et al., 2024 [24] conducted time-series analyses of experimental microbiomes, demonstrating that both deterministic and stochastic ecological processes drive the divergence of alternative states. This work highlighted the role of stochasticity in generating multiple stable states within microbial communities. Concurrently, Ponciano et al., 2024 [25] emphasized the necessity of integrating stochastic models to predict a population's persistence probabilities within the vaginal microbiome, advocating for a shift from static snapshots to dynamic modeling approaches. Collectively, these studies illustrate a progressive shift towards embracing stochastic modeling to more accurately capture the resilience and dynamic behavior of microbiome communities.

Although well studied, existing resilience measures are inadequate for microbiome studies because stochasticity, known to affect the stability of a dynamical system, does not influence these measures. Therefore, there is a pressing need to incorporate stochasticity, such as environmental noise and uncertainty, into the mechanistic models used to identify causal mechanisms using time-series microbiome data. In nature, the environment often fluctuates randomly and without memory, making white noise an excellent candidate to account for environmentally driven stochastic effects. Given the apparent need for such a model, we propose a stochastic generalized Lotka–Volterra (SgLV) model to describe the temporal dynamics of a community of microbial species. Stochasticity is induced in the model by adding white noise to the model parameters. Then, based on the linearization of our proposed SgLV model, we develop resilience measures to quantify resilience that can be applied to time-series microbiome data.

In our modeling framework, we represent environmental stochasticity using white noise. Specifically, we modify a constant parameter, such as α , to include stochastic fluctuations by redefining it as $\tilde{\alpha} = \alpha + \tau \frac{dW}{dt}$, where $W(t)$ is standard Brownian motion, and τ represents the intensity of noise. Consequently, $\tilde{\alpha}$ transforms from a fixed parameter into a stochastic process. Within a sufficiently small time interval $[0, t]$, we approximate the term $\frac{dW}{dt}$ by $\frac{W(t)}{t}$, which follows a normal distribution $N(0, \frac{1}{t})$. Thus, each time we simulate the model, the value of $\tilde{\alpha}$ will be chosen from the normal distribution $N(\alpha, \alpha + \tau^2/t)$. This way of incorporating noise is more realistic than that in Ives et al., 2003 [22], since the variance of the noise is not constant but changing over time.

The foundational work by Ives et al., 2003 [22] established critical methodologies for analyzing ecological community stability using discrete-time multivariate autoregressive (MAR) models, providing robust methods for estimating stability and ecological interactions from empirical time-series data. Their pioneering approach revealed important relationships between species interactions and environmental variability, significantly advancing the quantitative study of ecological dynamics. In our manuscript, we extend this foundational framework from discrete time to continuous time by developing stochastic generalized Lotka–Volterra models, which are particularly suited to capturing microbiome dynamics characterized by continuous growth and interaction processes. Moreover, we propose novel resilience metrics specifically tailored to microbiome community dynamics, including measures capable of quantifying recovery following perturbations and assessing how stochasticity influences the pace of community restoration. Thus, while conceptually rooted in the seminal work of Ives et al., 2003 [22], our model provides significant advancements optimized for addressing the unique complexities inherent in microbiome research.

The remainder of the paper is organized into three sections. First, we develop a general theoretical framework on how to calculate four resilience measures for the proposed SgLV model. Second, we demonstrate the practicality of our framework by applying it to a stochastic two-species competitive model to illustrate the effects of environmental noise on the resilience of the stochastic model. Lastly, we have a brief discussion on the possibilities of applying our framework to real microbiome data.

2. General theoretical framework

In this section, we develop a general theoretical framework to calculate four resilience measures that quantify the recovery of a microbial community from pulse perturbations (single shocks or perturbations of relatively short duration). In contrast to previous works that used ordinary differential equation (ODE) systems [18–20] and/or stochastic discrete-time systems [21,22], we use a stochastic differential equation (SDE) system to describe the temporal dynamics of species within a microbiome community and then study its transient dynamics under different pulse perturbations. We utilize white

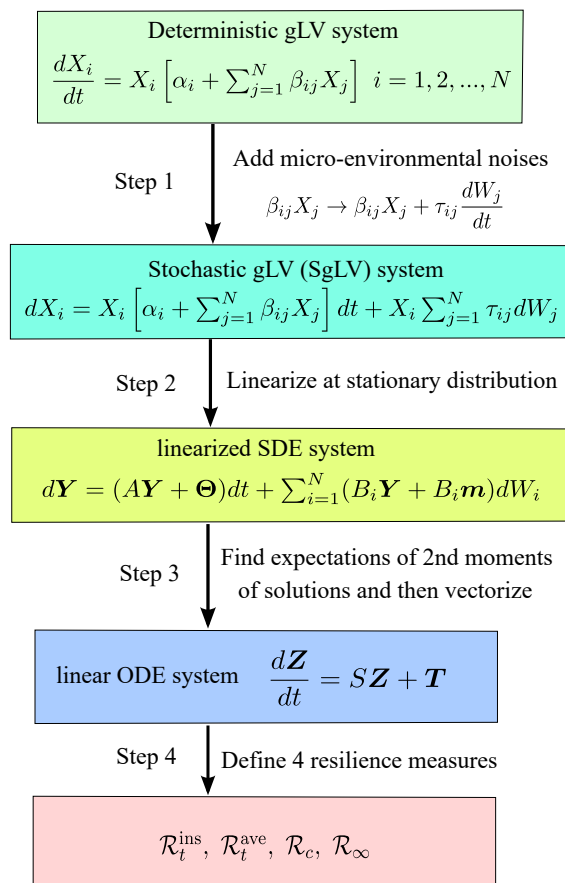


Figure 2. Four-step scheme of the method of calculating the resilience of a SgLV.

noise to represent random environmental fluctuations without memory to capture the fluctuations observed in microbiomic time-series data. We consider them to be the effects of persistent environmental perturbations. Mathematically, we add white noise, described as the “derivative of Brownian Mo-

tion,” to a deterministic gLV model. We then obtain a non-linear stochastic gLV model. Next, we linearize this nonlinear SDE system at its positive stationary distribution to get a linear SDE system. On the basis of two works by Arnoldi et al., [19,20], we indirectly establish four resilience measures, which are the *instantaneous return rate* (\mathcal{R}_t^{ins}), *average return rate* (\mathcal{R}_t^{ave}), *asymptotic resilience* (\mathcal{R}_∞), and the *convergence rate* (\mathcal{R}_c), of our stochastic gLV model, through the second moment of solutions of the corresponding linearized SDE system. These four resilience quantities, in which the first two are functions of time and the last two are just numbers, measure how strong the recovery dynamics of our original stochastic system is after a pulse perturbation is applied. Figure 2 shows the scheme of the theoretical method used to calculate the resilience of a stochastic generalized Lotka–Volterra system of a microbial community of N species. The details of this method will be presented in the following four subsections.

2.1. Model setup and notation

The widely used mechanistic model for microbiome studies is the deterministic generalized Lotka–Volterra model, which is given by

$$\frac{dX_i}{dt} = X_i \left[\alpha_i + \sum_{j=1}^N \beta_{ij} X_j \right], \quad i = 1, \dots, N, \quad (2.1)$$

in which X_i is the abundance of species i in a microbiome community. The parameter α_i represents the intrinsic growth rate of species i . The interaction between species i and species j ($j \neq i$) is captured by the parameter β_{ij} . When $i = j$, the parameter β_{ij} takes the form $\beta_{ii} = -\alpha_i c_i < 0$ where c_i is the coefficient of negative intraspecific interaction representing the inverse of the carrying capacity of the species i in isolation.

To characterize the microenvironmental fluctuations and noises in pairwise interactions between species within a microbiome community, we propose a general *stochastic generalized Lotka–Volterra* (SgLV) model that takes the following form

$$dX_i = X_i \left[\alpha_i + \sum_{j=1}^N \beta_{ij} X_j \right] dt + X_i \sum_{j=1}^N \tau_{ij} dW_j, \quad i = 1, \dots, N, \quad (2.2)$$

Notice that we perturbed each interaction coefficient β_{ij} in the system (2.1) by replacing each term $\beta_{ij} X_j$ with the term $\beta_{ij} X_j + \tau_{ij} \frac{dW_j}{dt}$ ($i, j = 1, \dots, N$) where $(W_1, \dots, W_N)^T$ is a standard Brownian motion of N dimensions.

For convenience, we introduce our notation that will be used throughout this paper. First, we need to specify an appropriate completed filtered probability space. Let $\Omega = \{\omega \in C(\mathbb{R}, \mathbb{R}^N), \omega(0) = 0\}$, let \mathcal{F} be the Borel σ -algebra on Ω , and let \mathbb{P} be the measure induced by $\{W = W(t)\}_{t \in \mathbb{R}}$, a two-sided N -dimensional Brownian motion. Then the elements of Ω will be identified with paths of the N -dimensional Brownian process $\omega(t) = W(t, \omega)$. Now we consider the \mathbb{P} -completion of \mathcal{F} , also denoted by \mathcal{F} , that is, \mathcal{F} contains all \mathbb{P} -null sets. The filtration \mathcal{F}_t is given by the canonical filtration generated by the Brownian motion $\{W(t)\}_{t \geq 0}$ completed by all \mathbb{P} -null sets of \mathcal{F} . Denote the probability measure given by the extension of \mathbb{P} to the completed \mathcal{F} again by \mathbb{P} . Therefore, a completed filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$ is obtained. Next, consider the SgLV (2.2) whose solutions take values in

$[0, \infty)^n$ and describe the abundance dynamics of N interacting species $\mathbf{X}(t) := (X_1(t), \dots, X_N(t))^T$, $t \geq 0$, in a microbial community. We write

$$\begin{aligned}\mathbb{R}^N &:= \{(x_1, \dots, x_N)^T : x_1 \in \mathbb{R}, \dots, x_N \in \mathbb{R}\}, \\ \mathbb{R}_+^N &:= \{(x_1, \dots, x_N)^T \in \mathbb{R}^N : x_1 \geq 0, \dots, x_N \geq 0\}, \\ \mathbb{R}_+^{N,\circ} &:= \{(x_1, \dots, x_N)^T \in \mathbb{R}^N : x_1 > 0, \dots, x_N > 0\}, \\ f_i = f_i(\mathbf{X}(t)) &:= \alpha_i + \sum_{j=1}^N \beta_{ij} X_j \text{ (the per-capita growth rate of the species } i), \\ \Gamma &:= (\tau_{ij})_{N \times N}, \text{ and } \Sigma := \Gamma \Gamma^T = (\sigma_{ij})_{N \times N} \text{ where } \sigma_{ij} = \sum_{k=1}^N \tau_{ik} \tau_{jk}.\end{aligned}$$

From now on, the stochastic process given by the solution of the system (2.2) will be denoted by \mathbf{X} or $(\mathbf{X}(t))_{t \geq 0}$. We use $\mathbb{P}_{\mathbf{x}}$ to denote the probability law on Ω when the solution path starts at $\mathbf{x} = (x_1, \dots, x_N)^T$ and $\mathbb{E}_{\mathbf{x}}$ is the expectation corresponding to $\mathbb{P}_{\mathbf{x}}$.

We use the norm $\|\mathbf{x}\| := \sum_{i=1}^N |x_i|$ in \mathbb{R}^N where $\mathbf{x} = (x_1, \dots, x_N)^T$. For $(u_1, \dots, u_N)^T \in \mathbb{R}^N$, we let $\bigwedge_{i=1}^N u_i := \min\{u_1, \dots, u_N\}$ and $\bigvee_{i=1}^N u_i := \max\{u_1, \dots, u_N\}$.

2.2. Persistence and extinction

Based on the excellent work of Hening and Nguyen 2018 [32–34], we briefly present, without proofs, sharp sufficient conditions for both persistence and extinction for SgLV (2.2) and its corresponding deterministic counterpart.

First, we consider the corresponding deterministic gLV system of the stochastic system (2.2) given by

$$\frac{dX}{dt} = h(X), \quad (2.3)$$

where $h(X) = (h_1(X), \dots, h_N(X))^T$, and $h_i(X) := X_i \left[\alpha_i + \sum_{j=1}^N \beta_{ij} X_j \right]$, $i = 1, \dots, N$. We define persistence and extinction for the deterministic system (2.3) as follows.

Definition 2.1. The system (2.3) is persistent if, for any initial value $\mathbf{X}(0)$ in $\mathbb{R}_+^{N,\circ}$, each component of the solution $\mathbf{X}(t)$ satisfies $\liminf_{t \rightarrow \infty} X_i(t) > 0$ for all $i = 1, \dots, N$.

Definition 2.2. For any initial value $\mathbf{X}(0) \in \mathbb{R}_+^{N,\circ}$, the species X_i is said to go extinct if $\limsup_{t \rightarrow \infty} X_i(t) = 0$.

In a deterministic setting, in order to investigate the persistence and extinction of the system (2.3) (i.e., we wish to know which species persist or which species go extinct in this system), we look at the set \mathcal{N} , which is the set of all the equilibria of the deterministic system (2.3) on the boundary $\partial \mathbb{R}_+^N := \mathbb{R}_+^N \setminus \mathbb{R}_+^{N,\circ}$, as well as the unique positive equilibrium E^* in $\mathbb{R}_+^{N,\circ}$. Clearly, $\mathcal{N} \neq \emptyset$ since $\mathbf{0} = (0, \dots, 0)^T \in \mathcal{N}$. Then, for each $E \in \mathcal{N} \cup \{E^*\}$, we compute the eigenvalues $\lambda_1(E), \dots, \lambda_N(E)$ of the variational matrix $\frac{dh}{dX} \Big|_{X=E}$. Hence, we have the following theorem about the persistence of the system (2.3).

Theorem 2.1. The microbial community described by the system (2.3) becomes persistent, i.e., the abundance of each species is positive, if all the equilibria on the boundary $\partial \mathbb{R}_+^{N,\circ}$ are unstable, which means that

$$\max_{i=1, \dots, N} \operatorname{Re} \lambda_i(E) > 0 \text{ for each } E \in \mathcal{N}, \quad (2.4)$$

This is equivalent to the fact that the unique positive equilibrium E^* is locally stable (i.e., all the eigenvalues of the linearized system at the positive equilibrium have non-positive real parts). If the deterministic system (2.3) is not persistent, we would like to track which species go extinct and which species survive. Indeed, for $\mathbf{0} \neq E = (x_{1,E}, \dots, x_{N,E})^T \in \mathcal{N}$, $1 \leq n_1 < \dots < n_k \leq N$ exist such that $x_{n_i,E} > 0$ for all $i = 1, \dots, k$ and $x_{j,E} = 0$ for all $j \in \{1, \dots, N\} \setminus \{n_1, \dots, n_k\}$. Let

$$\mathbb{R}_+^E := \{(x_1, \dots, x_N) \in \mathbb{R}_+^N : x_i = 0 \text{ for } i \in I_E^c\},$$

where $I_E := \{n_1, \dots, n_k\}$ and $I_E^c = \{1, \dots, N\} \setminus \{n_1, \dots, n_k\}$. If $E = \mathbf{0}$, then $\mathbb{R}_+^{\mathbf{0}} = \{\mathbf{0}\}$. Let

$$\mathbb{R}_+^{E,\circ} := \{(x_1, \dots, x_N) \in \mathbb{R}_+^N : x_i = 0 \text{ for } i \in I_E^c \text{ and } x_i > 0 \text{ for } i \in I_E\}$$

and $\partial R_+^E = \mathbb{R}_+^E \setminus \mathbb{R}_+^{E,\circ}$. Now we make the following assumption.

Assumption 1. *If a $E \in \mathcal{N}$ exists such that*

$$\max_{i \in I_E^c} \operatorname{Re} \lambda_i(E) < 0,$$

If $\mathbb{R}_+^E \neq \{\mathbf{0}\}$, then suppose further that for any $E' \in \mathcal{N}_E$, we have

$$\max_{i \in I_E} \operatorname{Re} \lambda_i(E') > 0,$$

where $\mathcal{N}_E := \{E' \in \mathcal{N} : \mathbb{R}_+^{E'} \subseteq \partial R_+^E\}$.

Define

$$\mathcal{N}^1 := \{E \in \mathcal{N} : E \text{ satisfies Assumption 1}\} \quad \text{and} \quad \mathcal{N}^2 := \mathcal{N} \setminus \mathcal{N}^1.$$

To determine exactly which species go extinct, we need an additional assumption ensuring that equilibria that are not in \mathcal{N}^1 are unstable.

Assumption 2. *Either $\mathcal{N}^2 = \emptyset$ or, for any $E' \in \mathcal{N}^2$, $\max_{i=1, \dots, N} \operatorname{Re} \lambda_i(E') > 0$.*

Theorem 2.2. *If Assumptions 1 and 2 are satisfied and $\mathcal{N}^1 \neq \emptyset$, then for any $E \in \mathcal{N}^1$, the species $X_i(t)$ converges to 0 with the convergence rate $\lambda_i(E)$ for any $i \in I_E^c$.*

Next, we consider the stochastic system (2.2). We make a standing assumption that will be used throughout this section.

Assumption 3. *The noise intensity matrix Γ of the system (2.2) satisfies the condition that the matrix $\Sigma := \Gamma \Gamma^T = (\sigma_{ij})_{N \times N}$ is a positive-definite matrix. Furthermore, $\mathbf{c} = (c_1, \dots, c_N)^T \in \mathbb{R}_+^{N,\circ}$ and $\delta > 0$ exist such that for sufficiently large $\|\mathbf{X}\|$, we have*

$$\frac{\sum_{i=1}^N c_i X_i f_i(\mathbf{X})}{1 + \mathbf{c}^T \mathbf{X}} - \frac{\sum_{i,j=1}^N \sigma_{ij} c_i c_j X_i X_j}{2(1 + \mathbf{c}^T \mathbf{X})^2} + \delta \left(1 + N + \sum_{i=1}^N |f_i(\mathbf{X})| \right) < 0. \quad (2.5)$$

Since the drift functions $X_i f_i(\mathbf{X}(t))$ are locally Lipschitz functions on \mathbb{R}^N for any $i = 1, \dots, N$ and the diffusion functions in the system (2.2) are linear, for each initial value in \mathbb{R}^N , the system (2.2) has a unique almost surely (a.s.) continuous solution $\mathbf{X} = \mathbf{X}(t)$ up to the explosion time

$$\tau_e = \inf \left\{ t > 0 : \bigwedge_{i=1}^N X_i(t) = -\infty \quad \text{or} \quad \bigvee_{i=1}^N X_i(t) = \infty \right\}.$$

Using the Ito's formula for $\log X_i(t)$, the system (2.2) can be written as

$$X_i(t) = X_i(0) \exp \left\{ \int_0^t f_i(s) ds - \left(\frac{1}{2} \sum_{j=1}^N \tau_{ij}^2 \right) t + \sum_{j=1}^N \tau_{ij} W_j(t) \right\} \text{ a.s.,}$$

which means that if $\mathbf{X}(0) \in \mathbb{R}_+^N$, then $\mathbf{X}(t) \in \mathbb{R}_+^N$ a.s. for all $t \in [0, \tau_e)$. Due to (2.5), we can obtain the tightness of the family of transition probabilities of the solution \mathbf{X} to the system (2.2). This means that $\tau_e = \infty$ a.s. Thus, for any initial value $\mathbf{X}(0)$ in \mathbb{R}_+^N , the system (2.2) has a unique a.s. continuous strong solution $\mathbf{X}(t)$ in \mathbb{R}_+^N for all $t \in [0, \infty)$.

The positive definiteness of the matrix Σ in Assumption 3 ensures that the solution to the system (2.2) is nondegenerate. This is an important assumption because it makes sure that we have enough noise that can locally push the dynamics of the solution in all directions, and hence any positive solution state can move close to any other positive solution state in a finite time. In other words, this condition guarantees that the family of transition probabilities of the solution has a smooth density, which will be used for the proof of the persistence and extinction of the system (2.2).

Now, we define what we mean by persistence and extinction in our stochastic model.

Definition 2.3. The process \mathbf{X} is strongly stochastically persistent if it possesses a unique invariant probability measure π^* in $\mathbb{R}_+^{N,0}$ and

$$\lim_{t \rightarrow \infty} \|P(t, \mathbf{x}, \cdot) - \pi^*(\cdot)\|_{TV} = 0, \quad \mathbf{x} \in \mathbb{R}_+^{N,0},$$

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

Definition 2.4. If $\mathbf{X}(0) = \mathbf{x} \in \mathbb{R}_+^{N,0}$, we say that the species X_i goes extinct with probability $p_{\mathbf{x}} > 0$ if

$$\mathbb{P}_{\mathbf{x}} \left\{ \lim_{t \rightarrow \infty} X_i(t) = 0 \right\} = p_{\mathbf{x}}.$$

We say that the species X_i goes extinct if, for all $\mathbf{x} \in \mathbb{R}_+^{N,0}$,

$$\mathbb{P}_{\mathbf{x}} \left\{ \lim_{t \rightarrow \infty} X_i(t) = 0 \right\} = 1.$$

In a stochastic setting, to look into the persistence and extinction, we consider the set \mathcal{M} as the collection of all ergodic invariant probability measures of the solution $\mathbf{X}(t)$ on the boundary $\partial \mathbb{R}_+^N$. Clearly, $\mathcal{M} \neq \emptyset$, since $\delta^* \in \mathcal{M}$ is the Dirac measure concentrated at $\mathbf{0} = (0, \dots, 0)^T \in \mathbb{R}_+^N$. For $\widetilde{\mathcal{M}} \subseteq \mathcal{M}$, we use $\text{Conv}(\widetilde{\mathcal{M}})$ to denote the convex hull of $\widetilde{\mathcal{M}}$, which is the collection of probability measures of the form

$$\pi(\cdot) = \sum_{\mu \in \widetilde{\mathcal{M}}} p_{\mu} \mu(\cdot)$$

with $p_\mu > 0$ and $\sum_{\mu \in \widetilde{\mathcal{M}}} p_\mu = 1$. For $\delta^* \neq \mu \in \mathcal{M}$, since the stochastic system (2.2) is nondegenerate by Assumption 3, $1 \leq n_1 < \dots < n_k \leq N$ exist such that the support of the ergodic invariant probability measure μ , denoted by $\text{supp}(\mu)$, is exactly \mathbb{R}_+^μ , in which

$$\mathbb{R}_+^\mu := \{(x_1, \dots, x_N) \in \mathbb{R}_+^N : x_i = 0 \text{ for } i \in I_\mu^c\}$$

for $I_\mu := \{n_1, \dots, n_k\}$ and $I_\mu^c := \{1, \dots, N\} \setminus I_\mu$. Let

$$\mathbb{R}_+^{\mu, \circ} := \{(x_1, \dots, x_N) \in \mathbb{R}_+^N : x_i = 0 \text{ for } i \in I_\mu^c \text{ and } x_i > 0 \text{ for } i \in I_\mu\}$$

and $\partial \mathbb{R}_+^\mu := \mathbb{R}_+^\mu \setminus \mathbb{R}_+^{\mu, \circ}$. Similarly to a deterministic setting, persistence in a stochastic setting means that every ergodic invariant probability measure on the boundary is a *repeller*. Therefore, ergodic invariant probability measures play a similar role as equilibria in a deterministic setting. To determine an ergodic invariant probability measure of the stochastic system (2.3) to be a repeller, we compute its Lyapunov exponents by looking at the equations for $\ln X_i(t)$, $i = 1, \dots, N$. Applying Ito's formula for $\ln X_i(t)$ gives

$$\frac{\ln X_i(t)}{t} = \frac{\ln X_i(0)}{t} + \frac{1}{t} \int_0^t f_i(\mathbf{X}(s)) ds - \frac{1}{2} \sum_{j=1}^N \tau_{ij}^2 + \sum_{j=1}^N \tau_{ij} \frac{W_j(t)}{t}.$$

When the solution $\mathbf{X}(t)$ to the system (2.3) is close to the support of an ergodic invariant probability measure μ , $\text{supp}(\mu)$, for a long time, then the Lyapunov exponent of the solution component X_i with respect to μ (which is also called the *invasion rate* of species X_i with respect to μ)

$$\lambda_i(\mu) := \lim_{t \rightarrow \infty} \frac{\ln X_i(t)}{t} = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t f_i(\mathbf{X}(s)) ds - \frac{1}{2} \sum_{j=1}^N \tau_{ij}^2,$$

can be approximated by the average with respect to μ (using the strong law of large numbers)

$$\int_{\partial \mathbb{R}_+^N} f_i(\mathbf{x}) \mu(d\mathbf{x}) - \frac{1}{2} \sum_{j=1}^N \tau_{ij}^2,$$

while the term

$$\frac{\ln X_i(0)}{t} + \sum_{j=1}^N \tau_{ij} \frac{W_j(t)}{t},$$

is negligible. This implies that the Lyapunov exponents $\lambda_i(\mu)$, $i = 1, \dots, N$, give the long-term growth rate of species X_i if $\mathbf{X}(t)$ is close to $\text{supp}(\mu)$. Therefore, an ergodic invariant probability measure μ is a repeller if $\max_{i=1, \dots, N} \lambda_i(\mu) > 0$. Now we impose the following assumption that ensures the strong stochastic persistence.

Assumption 4. For any μ in $\text{Conv}(\mathcal{M})$,

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

where

$$\lambda_i(\mu) = \int_{\partial \mathbb{R}_+^N} f_i(\mathbf{x}) \mu(d\mathbf{x}) - \frac{1}{2} \sum_{j=1}^N \tau_{ij}^2,$$

Theorem 2.3. Suppose that Assumptions 3 and 4 are true. Then the solution $\mathbf{X}(t)$ to the system (2.2) is strongly stochastically persistent, and its transition probabilities converge weakly to its unique invariant probability measure π^* in $\mathbb{R}_+^{N,\circ}$ exponentially fast in the total variation norm.

Next, we make an assumption that will imply the extinction of the stochastic system.

Assumption 5. There is a $\mu \in \mathcal{M}$ such that

$$\max_{i \in I_\mu^c} \lambda_i(\mu) < 0, \quad (2.6)$$

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

$$\max_{i \in I_\mu} \lambda_i(\nu) > 0, \quad (2.7)$$

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

Theorem 2.4. Under Assumptions 3 and 5, for any $\delta > 0$ sufficiently small and for any initial value $\mathbf{x} \in \mathbb{R}_+^{N,\circ}$, we obtain

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

We have several remarks for Assumption 5. First, if $\mu \in \mathcal{M}$ and $i \in I_\mu$, then $\lambda_i(\mu) = 0$. Intuitively, this is because if the solution is inside the support of an ergodic invariant probability measure μ , then it is at an “equilibrium” and it does not tend to grow or decay. If $\mu \in \mathcal{M}$ and $\max_{i \in I_\mu^c} \lambda_i(\mu) < 0$, then μ is regarded as an “attractor”, which attracts solutions starting nearby. So we require the condition (2.6) to make sure that the solution component $X_i(t)$, $i \in I_\mu^c$, will get close to 0 if it starts near $\mathbb{R}_+^{\mu,\circ}$. Second, the condition (2.7) is needed to guarantee that μ is really a “sink” in $\mathbb{R}_+^{\mu,\circ}$ but not the other invariant probability measures on $\partial \mathbb{R}_+^\mu$, that is, if \mathbf{X} is close to $\mathbb{R}_+^{\mu,\circ}$, then it is not pulled away to the boundary $\partial \mathbb{R}_+^\mu$.

Finally, let

$$\mathcal{M}^1 := \{\mu \in \mathcal{M} : \mu \text{ satisfies Assumption 5}\}$$

and $\mathcal{M}^2 := \mathcal{M} \setminus \mathcal{M}^1$. If there is no persistence, we would like to know exactly which species go extinct and which species survive. On the basis of the repulsion of the invariant probability measures in \mathcal{M}^2 , we can derive the “attracting” region of invariant probability measures in \mathcal{M}^1 for the solution \mathbf{X} . We make an additional assumption that ensures all the invariant probability measures outside $\text{Conv}(\mathcal{M}^1)$ are repellers.

Assumption 6. Suppose that one of the following holds true:

- $\mathcal{M}^2 = \emptyset$;
- For any $\nu \in \text{Conv}(\mathcal{M}^2)$, $\max_{i=1,\dots,N} \lambda_i(\nu) > 0$.

For any initial value $\mathbf{X}(0) = \mathbf{x} \in \mathbb{R}_+^{N,\circ}$, let $\mathcal{U} = \mathcal{U}(\omega)$ denote the weak*-limit set of the family of random occupation measures $\{\tilde{\Pi}_t(\cdot), t \geq 0\}$ where

$$\tilde{\Pi}_t(A) = \frac{1}{t} \int_0^t \mathbb{1}_{\{\mathbf{X}(s) \in A\}} ds, \quad t > 0, \quad A \in \mathcal{B}(\mathbb{R}_+^{N,\circ}),$$

which is the proportion of time the solution spends in a set A up to time t .

Theorem 2.5. Suppose that Assumptions 1, 5, and 6 hold true and $\mathcal{M}^1 \neq \emptyset$. Then for any $\mathbf{x} \in \mathbb{R}_+^{N,\circ}$, we obtain

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

in which

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

2.3. Linearization method

Let us get started with the linearization of the deterministic system (2.3). We assume that all the equilibria of this system on the boundary $\partial \mathbb{R}_+^{N,\circ}$ are unstable. Therefore, by Theorem 2.1, the unique positive equilibrium $E^* = (X_1^*, \dots, X_N^*)^T$ is locally stable. To understand the behavior of the deterministic system near E^* , we approximate the system (2.3) by the following linear system, as the solution is close to E^* :

$$\frac{dX_i}{dt} = \phi_0^i + \sum_{j=1}^N \theta_{ij}(X_j - X_j^*), \ i = 1, \dots, N, \quad (2.8)$$

where the values of ϕ_0^i and the θ_{ij} are constants to be determined. In fact, we would like to replace each right-hand side of the system (2.3), $h_i(\mathbf{X})$, by its linearized form $\phi_0^i + \sum_{j=1}^N \theta_{ij}(X_j - X_j^*)$ for $i = 1, \dots, N$ when $X_1 - X_1^*, \dots, X_N - X_N^*$ are close to 0. The constant coefficients ϕ_0^i and θ_{ij} are determined so that all the differences

$$\left| h_i(\mathbf{X}) - \phi_0^i - \sum_{j=1}^N \theta_{ij}(X_j - X_j^*) \right|,$$

for $i = 1, \dots, N$ approximate 0 as \mathbf{X} is close to E^* . If we use the norm $\|\mathbf{x}\|_2 = \sqrt{x_1^2 + \dots + x_N^2}$ in \mathbb{R}^N to identify how close two vectors are, then the linearization of the system (2.3) at the unique positive equilibrium E^* is equivalent to finding the vector $\boldsymbol{\phi}_0 = (\phi_0^1, \dots, \phi_0^N)^T$ and the matrix $\boldsymbol{\Theta} = (\theta_{ij})_{N \times N}$ so that the difference

$$\|h(\mathbf{X}) - \boldsymbol{\phi}_0 - \boldsymbol{\Theta}(\mathbf{X} - E^*)\|_2,$$

approximates 0 as $\|\mathbf{X} - E^*\|_2 \rightarrow 0$. For each $i = 1, \dots, N$, applying the Taylor expansion for the function $h_i(\mathbf{X})$ at E^* gives

$$\begin{aligned} h_i(\mathbf{X}) &= h_i(E^*) + \frac{\partial h_i(E^*)}{\partial \mathbf{X}} \cdot (\mathbf{X} - E^*) \\ &\quad + \frac{1}{2} (\mathbf{X} - E^*)^T \frac{\partial^2 h_i(E^*)}{\partial \mathbf{X}^2} (\mathbf{X} - E^*) + o(\|\mathbf{X} - E^*\|_2^2), \end{aligned}$$

in which

$$\frac{\partial h_i(E^*)}{\partial \mathbf{X}} = \left(\frac{\partial h_i(E^*)}{\partial X_1}, \dots, \frac{\partial h_i(E^*)}{\partial X_N} \right)^T$$

and

$$\frac{\partial^2 h_i(E^*)}{\partial \mathbf{X}^2} = \left(\frac{\partial^2 h_i(E^*)}{\partial X_j \partial X_k} \right)_{N \times N}.$$

This implies that as $\|\mathbf{X} - E^*\|_2 \rightarrow 0$, we get

$$\left| h_i(\mathbf{X}) - h_i(E^*) - \frac{\partial h_i(E^*)}{\partial \mathbf{X}} \cdot (\mathbf{X} - E^*) \right| \rightarrow 0,$$

Notice that $h_i(E^*) = 0$ for each $i = 1, \dots, N$. Therefore, when we choose $\phi_0 = \mathbf{0}$ and $\Theta = \frac{\partial h(E^*)}{\partial \mathbf{X}} := \left(\frac{\partial h_i(E^*)}{\partial X_j} \right)_{N \times N}$, then

$$h_i(\mathbf{X}) \approx \sum_{j=1}^N \theta_{ij}(X_j - X_j^*),$$

as \mathbf{X} is close to E^* . Then the linearization of the deterministic system (2.3) at E^* takes the form

$$\frac{dX_i}{dt} = \sum_{j=1}^N \frac{\partial h_i(E^*)}{\partial X_j} (X_j - X_j^*), \quad i = 1, \dots, N, \quad (2.9)$$

Next, we continue with the linearization of the stochastic system (2.2). We assume that the sufficient conditions (i.e., Assumption 3 and Assumption 4) for the strong persistence of the stochastic system are satisfied. By Theorem 2.3, there is a unique exponentially ergodic invariant probability measure π^* in $\mathbb{R}_+^{N,\circ}$. So, for any $\mathbf{x} \in \mathbb{R}_+^{N,\circ}$, all transition probabilities $P(t, \mathbf{x}, \cdot) = \mathbb{P}_{\mathbf{x}}\{\mathbf{X}(t) \in \cdot\}$ of the solution $\mathbf{X}(t)$ starting at \mathbf{x} become the same and equal to π^* as the time t becomes large enough. We call π^* the *unique positive stationary solution* of the stochastic system (2.2). We say that the solution $\mathbf{X}(t)$ is close to π^* if $\|P(t, \mathbf{x}, \cdot) - \pi^*(\cdot)\|_{TV}$ is close to 0. Note that $\mathbb{E}_{\mathbf{x}}\pi^*$ approaches E^* for any $\mathbf{x} \in \mathbb{R}_+^{N,\circ}$ when all entries in the noise intensity matrix Γ approach 0. Since, as time passes, all solutions $\mathbf{X}(t)$ will eventually end up in the unique positive stationary distribution no matter where it starts, we will drop the subscript \mathbf{x} in the notation of $\mathbb{E}_{\mathbf{x}}$ and $\mathbb{P}_{\mathbf{x}}$ in this subsection.

The linearization of the stochastic system (2.2) at the stationary solution π^* is to approximate this system by the following linear SDE system when the solution is close to π^* :

$$dX_i = \left[\Phi_i + \sum_{j=1}^N \gamma_{ij}(X_j - m_j) \right] dt + X_i \sum_{j=1}^N \tau_{ij} dW_j, \quad i = 1, \dots, N, \quad (2.10)$$

where $(m_1, \dots, m_N) = \mathbb{E} \pi^*$. In fact, our statistical linearization method is the replacement of each drift term $h_i(\mathbf{X}) = X_i \left[\alpha_i + \sum_{j=1}^N \beta_{ij} X_j \right]$ in (2.2) by its linearized form $\Phi_i + \sum_{j=1}^N \gamma_{ij}(X_j - m_j)$ for $i = 1, \dots, N$, in which the coefficients Φ_i 's and γ_{ij} 's are determined by using the Taylor expansion of each $h_i(\mathbf{X})$ at $\mathbf{m} = (m_1, \dots, m_N)$. In other words, we would like to find these coefficients to minimize

$$\left| h_i(\mathbf{X}) - \Phi_i - \sum_{j=1}^N \gamma_{ij}(X_j - m_j) \right|, \quad (2.11)$$

as the solution \mathbf{X} of the stochastic system (2.2) is close to π^* . As in the deterministic setting, we obtain

$$\Phi_i = h_i(\mathbf{m}) = m_i \left[\alpha_i + \sum_{j=1}^N \beta_{ij} m_j \right] \quad i = 1, \dots, N, \quad (2.12)$$

and for $i, j = 1, \dots, N$

$$\gamma_{ij} = \frac{\partial h_i}{\partial x_j}(\mathbf{m}) = \begin{cases} \alpha_i + \sum_{k=1}^N \beta_{ik} m_k & \text{if } i = j, \\ \beta_{ij} m_i & \text{if } i \neq j, \end{cases} \quad (2.13)$$

in which each m_i can be computed using the strong law of large numbers

$$m_i = \int_{\partial \mathbb{R}_+^N} X_i \pi^*(d\mathbf{X}) = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t X_i(s) ds, \quad (2.14)$$

Therefore the system (2.2) can be approximated by the linear SDE system at the positive stationary solution π^*

$$\begin{aligned} d \begin{bmatrix} \widetilde{X}_1 \\ \cdots \\ \widetilde{X}_N \end{bmatrix} &= \left(\begin{bmatrix} \Phi_1 \\ \cdots \\ \Phi_N \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \cdots & \gamma_{N1} \\ \vdots & \ddots & \vdots \\ \gamma_{1N} & \cdots & \gamma_{NN} \end{bmatrix} \begin{bmatrix} \widetilde{X}_1 - m_1 \\ \cdots \\ \widetilde{X}_N - m_N \end{bmatrix} \right) dt \\ &\quad + \sum_{i=1}^N \begin{bmatrix} \tau_{1i} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \tau_{Ni} \end{bmatrix} \begin{bmatrix} \widetilde{X}_1 \\ \cdots \\ \widetilde{X}_N \end{bmatrix} dW_i. \end{aligned} \quad (2.15)$$

Finally, we comment on our linearization method. We know that the solution of the nonlinear stochastic system (2.2) is a Markov process $\mathbf{X} = \mathbf{X}(t)$. In fact, our linearization is to find another Markov process, say $\widetilde{\mathbf{X}} = \widetilde{\mathbf{X}}(t)$, which represents the solution of the linear SDE system (2.15) such that both \mathbf{X} and $\widetilde{\mathbf{X}}$ have the same mean, and the variance $\mathbb{E} \|\mathbf{X} - \widetilde{\mathbf{X}}\|_2^2$ is to be minimized as the time becomes large enough. In other words, when time goes by and becomes sufficiently large, the long-term behaviors of the two Markov processes \mathbf{X} and $\widetilde{\mathbf{X}}$ are approximately the same. Because of that, a $\delta > 0$ exists such that the linear SDE system (2.15) is considered as a “ δ -perturbation” of the non-linear SDE system (2.2), and thus strong persistence of the system (2.2) implies that of the system (2.15) (see the definition of “ δ -perturbation” and Proposition 3 in [28] for more details).

2.4. Four resilience measures

On the basis of two excellent works by Arnoldi [19, 20], we indirectly establish four resilience measures, which are the *instantaneous return rate*, *average return rate*, *asymptotic resilience*, and *convergence rate* of our stochastic gLV model, through the second moment of the solutions of the corresponding linearized SDE system. These four resilience quantities, in which the first two are functions of time and the last two are just numbers, measure how strong the recovery dynamics of our original stochastic system is after a pulse perturbation is applied.

Now we start building up these four stability concepts by letting

$$\begin{aligned} A &= \begin{bmatrix} \gamma_{11} & \cdots & \gamma_{N1} \\ \vdots & \ddots & \vdots \\ \gamma_{1N} & \cdots & \gamma_{NN} \end{bmatrix}, \quad B_i = \begin{bmatrix} \tau_{1i} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \tau_{Ni} \end{bmatrix}, \quad i = 1, \dots, N, \\ \Theta &= (\Phi_1, \dots, \Phi_N)^T, \quad \mathbf{m} = (m_1, \dots, m_N)^T, \quad \text{and } \mathbf{Y} = \widetilde{\mathbf{X}} - \mathbf{m}. \end{aligned}$$

Then (2.15) can be written in matrix form:

$$d\mathbf{Y} = (A\mathbf{Y} + \Theta)dt + \sum_{i=1}^N (B_i\mathbf{Y} + B_i\mathbf{m})dW_i. \quad (2.16)$$

We assume that the linearized system (2.16) is already at its equilibrium state. A pulse perturbation at time $t = 0$ applied to this linear SDE system is characterized by a vector $\mathbf{u} = \mathbf{Y}(0)$, which describes the state of the system right after the perturbation. Then the recovery trajectory $\mathbf{Y}(t)$ after this pulse perturbation is given by

$$\begin{aligned} \mathbf{Y}(t) = & \Phi(t)\mathbf{u} + \Phi(t) \int_0^t \Phi^{-1}(s) \left[\Theta - \sum_{i=1}^N B_i^2 \mathbf{m} \right] ds \\ & + \sum_{i=1}^N \Phi(t) \int_0^t \Phi^{-1}(s) B_i \mathbf{m} dW_i(s), \end{aligned} \quad (2.17)$$

in which $\Phi(t)$ is the solution matrix of the homogeneous system

$$d\mathbf{Y} = A\mathbf{Y}dt + \sum_{i=1}^N B_i \mathbf{Y} dW_i. \quad (2.18)$$

with the initial condition $\Phi(0) = I$ (the identity matrix of size N). In general, we cannot explicitly compute the solution $\mathbf{Y}(t)$ in terms of A, B_1, \dots, B_N . Notice that the solution $\mathbf{Y}(t)$ is nowhere differentiable. Without the smoothness of the solution, it is difficult to build up stability concepts of a dynamic system that are able to be tractable. To avoid this difficulty, we can smooth out the solutions of the linearized system (2.16) by looking at the dynamics of their second moments. To proceed, let $P(t) = \mathbb{E}(\mathbf{Y}(t)\mathbf{Y}(t)^T)$ be the expectation of the matrix-valued process $\mathbf{Y}(t)\mathbf{Y}(t)^T$, which is given by the unique solution of the matrix ODE (see [23])

$$\begin{aligned} \frac{dP}{dt} = & AP + PA^T + \sum_{i=1}^N B_i P B_i + \Theta M^T + M \Theta^T \\ & + \sum_{i=1}^N B_i (M \mathbf{m}^T + \mathbf{m} M^T + \mathbf{m} \mathbf{m}^T) B_i, \end{aligned} \quad (2.19)$$

with the initial value $P(0) = \mathbf{Y}(0)\mathbf{Y}(0)^T = \mathbf{u}\mathbf{u}^T$, and $M := M(t) = \mathbb{E}\mathbf{Y}(t)$ is the solution to the linear ODE system, with the initial value $M(0) = \mathbf{u}$,

$$\dot{M} = AM + \Theta. \quad (2.20)$$

We can treat the matrix $P(t)$ as a N^2 -dimensional vector

$$\begin{aligned} \mathbf{Z}(t) := \text{vec}(P(t)) &= (Z_1(t), \dots, Z_{N^2}(t))^T \\ &= (\mathbb{E}(Y_1^2(t)), \mathbb{E}(Y_2(t)Y_1(t)), \dots, \mathbb{E}(Y_N(t)Y_1(t)), \\ &\quad \mathbb{E}(Y_1(t)Y_2(t)), \mathbb{E}(Y_2(t)^2), \dots, \mathbb{E}(Y_N(t)Y_2(t)), \dots, \mathbb{E}(Y_N(t)^2))^T. \end{aligned}$$

Applying the vectorization operation on both sides of the system (2.19) and then using the properties of Kronecker products and matrix vectorization in linear algebra [23], we obtain the equivalent N^2 -dimensional ODE system

$$\frac{d\mathbf{Z}}{dt} = S\mathbf{Z} + T, \quad (2.21)$$

in which

$$S = I \otimes A + A \otimes I + \sum_{i=1}^N B_i \otimes B_i,$$

and

$$T = \text{vec}(\Theta M^T + M \Theta^T) + \sum_{i=1}^N (B_i \otimes B_i) \text{vec}(M m^T + m M^T + m m^T),$$

The matrix S is called the *mean-square stability matrix* of the linear SDE system (2.16). The stationary solution to (2.16) is globally mean-square asymptotically stable if $\alpha(S) = \max_{i=1, \dots, N^2} \text{Re}(\lambda_i) < 0$, where $\text{Re}(\lambda_i)$ is the real part of the eigenvalue λ_i of the matrix S . Now let

$$W(t) := \sum_{i=1}^N \mathbb{E} Y_i^2(t) = \mathbb{E} \left(\sum_{i=1}^N Y_i^2(t) \right) = \mathbb{E}(\mathbf{Y}^T(t) \mathbf{Y}(t)),$$

then $W(t)$ measures the average squared distance of the solution $\mathbf{X}(t)$ of the system (2.2) to the expectation of its equilibrium state π^* . We define four resilience quantities based on $W(t)$ as follows:

$$\begin{aligned} \mathcal{R}_t^{\text{ins}} &= -\frac{d}{dt} \ln W(t)^{1/2} \text{ is the instantaneous return rate at time } t > 0, \\ \mathcal{R}_t^{\text{ave}} &= -\frac{\ln W(t)^{1/2} - \ln W(0)^{1/2}}{t} \text{ is the average return rate over the time interval } [0, t], \\ \mathcal{R}_c &= -\frac{1}{2} \alpha(S) \text{ is the rate of convergence of the ODE system (2.21), and} \\ \mathcal{R}_\infty &= \lim_{t \rightarrow \infty} \mathcal{R}_t^{\text{ave}} \text{ is the asymptotic resilience of the ODE system (2.21).} \end{aligned}$$

In these definitions, $\mathcal{R}_t^{\text{ins}}$ and $\mathcal{R}_t^{\text{ave}}$ are deterministic functions of time t , which capture the short-term behavior of the recovery trajectory $\mathbf{Y}(t)$, while \mathcal{R}_c and \mathcal{R}_∞ are only numbers, which represent the long-term dynamics of the solution's recovery. These two functions and two numbers give us more tractable computations that measure the recovery dynamics of the stochastic system (2.16) since we are able to explicitly compute the quantity $\ln W(t)^{1/2}$ and its derivative in terms of the matrices S and T . In fact, two quantities $\mathcal{R}_t^{\text{ins}}$ and $\mathcal{R}_t^{\text{ave}}$ can be computed on the basis of the derivative of $\ln W(t)$

$$\begin{aligned} \frac{d}{dt} \ln W &= \frac{1}{W} \frac{dW}{dt} = \frac{1}{W} \left\{ \mathbb{E}(\mathbf{Y}^T (A^T + A) \mathbf{Y}) + \Theta^T M + M^T \Theta \right. \\ &\quad \left. + \sum_{i=1}^N \left[\mathbb{E}(\mathbf{Y}^T B_i^2 \mathbf{Y}) + \mathbf{m}^T B_i^2 M + M^T B_i^2 \mathbf{m} + \mathbf{m}^T B_i^2 \mathbf{m} \right] \right\} \\ &= \frac{1}{W} \left\{ \left[\text{vec}(A^T + A) + \sum_{i=1}^N \text{vec}(B_i^2) \right]^T \mathbf{Z} + \Theta^T M + M^T \Theta \right. \\ &\quad \left. + \sum_{i=1}^N (\mathbf{m}^T B_i^2 M + M^T B_i^2 \mathbf{m} + \mathbf{m}^T B_i^2 \mathbf{m}) \right\}, \end{aligned}$$

where W , M , and \mathbf{Z} can be computed from two ODE systems (2.20) and (2.21). Lastly, when all noise intensities are equal to 0 in the system (2.2) (that is, $B_i = 0$ for all $i = 1, \dots, N$), all our definitions of the four resilience measures above correspond to those defined in Arnoldi [20] in which the rate of convergence and the asymptotic resilience coincide.

3. Example of a stochastic 2-species model

3.1. Analysis of the model

To demonstrate the method we proposed in previous section, let us look at the stochastic gLV model of two species

$$\begin{aligned}dX_1 &= X_1(\alpha_1 - \beta_{11}X_1 + \beta_{12}X_2)dt + X_1(\tau_{11}dW_1 + \tau_{12}dW_2), \\dX_2 &= X_2(\alpha_2 + \beta_{21}X_1 - \beta_{22}X_2)dt + X_2(\tau_{21}dW_1 + \tau_{22}dW_2),\end{aligned}\quad (3.1)$$

where $\beta_{11} > 0$ and $\beta_{22} > 0$. There are three cases:

- If $\beta_{12} < 0$ and $\beta_{21} < 0$, then the system (3.1) is a competitive model.
- If $\beta_{12} > 0$ and $\beta_{21} > 0$, then the system (3.1) is a mutualistic model.
- If $(\beta_{12} < 0, \beta_{21} > 0)$ or $(\beta_{12} > 0, \beta_{21} < 0)$, then the system (3.1) is a predator–prey model (mixed competition and mutualism).

Next, we analyze the dynamics of the system (3.1) on the boundary of \mathbb{R}_+^2 , which is $\partial\mathbb{R}_+^2 := \mathbb{R}_{1+}^\circ \cup \mathbb{R}_{2+}^\circ \cup \{(0, 0)\}$, where $\mathbb{R}_{1+}^\circ := \{(x_1, 0) : x_1 > 0\}$ and $\mathbb{R}_{2+}^\circ := \{(0, x_2) : x_2 > 0\}$.

Case 1. If $X_1(0) = 0$ and $X_2(0) = 0$, then, by (3.1), we get, for a.s.

$$\begin{aligned}X_1(t) &= X_1(0) \exp \left\{ \int_0^t \left[\alpha_1 - \beta_{11}X_1(s) + \beta_{12}X_2(s) - \frac{\tau_{11}^2}{2} - \frac{\tau_{12}^2}{2} \right] ds \right. \\&\quad \left. + \tau_{11}W_1(t) + \tau_{12}W_2(t) \right\}, \\X_2(t) &= X_2(0) \exp \left\{ \int_0^t \left[\alpha_2 + \beta_{21}X_1(s) - \beta_{22}X_2(s) - \frac{\tau_{21}^2}{2} - \frac{\tau_{22}^2}{2} \right] ds \right. \\&\quad \left. + \tau_{21}W_1(t) + \tau_{22}W_2(t) \right\}.\end{aligned}$$

This implies that $X_1(t) = X_2(t) = 0$ for all $t > 0$ a.s. So, the Dirac delta measure at $(0, 0)$, μ_0 , is an ergodic invariant probability measure for the system (3.1) on the boundary $\partial\mathbb{R}_+^2$.

Case 2. If $X_1(0) = 0$ and $X_2(0) > 0$, then, as in Case 1, $X_1(t) \equiv 0$ a.s. and $X_2(t) > 0$ for all $t > 0$ a.s. Plugging in 0 into X_1 in the second equation of (3.1) gives

$$dX_2 = X_2(\alpha_2 - \beta_{22}X_2)dt + X_2(\tau_{21}dW_1 + \tau_{22}dW_2). \quad (3.2)$$

For a fixed $\alpha_2 > 0$, consider

$$\begin{aligned}s(X_2) &= \int_{\alpha_2}^{X_2} \exp \left\{ - \int_{\alpha_2}^y \frac{(2\alpha_2 - 2\beta_{22}u)u}{(\tau_{21}^2 + \tau_{22}^2)u^2} du \right\} dy \\&= C_2 \int_{\alpha_2}^{X_2} y^{-\frac{2\alpha_2}{\tau_{21}^2 + \tau_{22}^2}} \exp \left\{ \frac{2\beta_{22}}{\tau_{21}^2 + \tau_{22}^2} y \right\} dy.\end{aligned}$$

Since the integrand in the last integral can be written as

$$y^{-\frac{2\alpha_2}{\tau_{21}^2 + \tau_{22}^2}} \left[1 + \frac{2\beta_{22}}{\tau_{21}^2 + \tau_{22}^2} y + \frac{1}{2!} \frac{4\beta_{22}^2}{(\tau_{21}^2 + \tau_{22}^2)^2} y^2 + \cdots \right],$$

a $k \in \mathbb{Z}_+$ exists such that $\frac{-2\alpha_2}{\tau_{21}^2 + \tau_{22}^2} + k > -1$, $s(\infty) = \infty$. If $\frac{-2\alpha_2}{\tau_{21}^2 + \tau_{22}^2} + 1 > 0$, which is equivalent to $\alpha_2 - \frac{\tau_{21}^2}{2} - \frac{\tau_{22}^2}{2} < 0$, then $s(0+) > -\infty$. This means that $\lim_{t \rightarrow \infty} X_2(t) = 0$ a.s. and so the equation (3.2) does not have any invariant probability measure on \mathbb{R}_+° . If $\frac{-2\alpha_2}{\tau_{21}^2 + \tau_{22}^2} + 1 < 0$, which is equivalent to $\alpha_2 - \frac{\tau_{21}^2}{2} - \frac{\tau_{22}^2}{2} > 0$, then $s(0+) = -\infty$. Therefore, $X_2(t)$ oscillates between 0 and ∞ . Thus the equation (3.2) has a unique invariant probability measure π_2 on \mathbb{R}_+° , which is the Gamma distribution

$$\pi_2 \sim \text{Gamma}\left(\frac{2\alpha_2}{\tau_{21}^2 + \tau_{22}^2} - 1, \frac{\tau_{21}^2 + \tau_{22}^2}{2\beta_{22}}\right).$$

So $\mu_2 := \delta_0^* \times \pi_2$ is an ergodic invariant probability measure for the system (3.1) on the boundary \mathbb{R}_{2+}° , in which δ_0^* is the Dirac Delta measure with the mass at 0.

Case 3. If $X_1(0) > 0$ and $X_2(0) = 0$ then $X_1(t) > 0$ for all $t > 0$ a.s. and $X_2(t) \equiv 0$ a.s. But then the first equation of (3.1) becomes

$$dX_1 = X_1(\alpha_1 - \beta_{11}X_1)dt + X_1(\tau_{11}dW_1 + \tau_{12}dW_2). \quad (3.3)$$

By the same argument as above, the equation (3.3) has a unique invariant probability measure π_1 on \mathbb{R}_+° , which is the Gamma distribution

$$\pi_1 \sim \text{Gamma}\left(\frac{2\alpha_1}{\tau_{11}^2 + \tau_{12}^2} - 1, \frac{\tau_{11}^2 + \tau_{12}^2}{2\beta_{11}}\right),$$

provided that $\alpha_1 - \frac{\tau_{11}^2}{2} - \frac{\tau_{12}^2}{2} > 0$. Otherwise, (3.3) does not have any invariant probability measure on \mathbb{R}_+° . Therefore, $\mu_1 = \pi_1 \times \delta_0^*$ is an ergodic invariant probability measure for the system (3.1) on the boundary \mathbb{R}_{1+}° .

Now we compute the Lyapunov exponents of $\mu_0 = \delta_0^* \times \delta_0^*$, μ_1 , and μ_2 . By Ito's formula, the system (3.1) is equivalent to

$$\begin{aligned} \frac{\ln X_1(t)}{t} &= \frac{\ln X_1(0)}{t} + \frac{1}{t} \int_0^t \left[\alpha_1 - \beta_{11}X_1(s) + \beta_{12}X_2(s) - \frac{\tau_{11}^2}{2} - \frac{\tau_{12}^2}{2} \right] ds \\ &\quad + \tau_{11} \frac{W_1(t)}{t} + \tau_{12} \frac{W_2(t)}{t}, \\ \frac{\ln X_2(t)}{t} &= \frac{\ln X_2(0)}{t} + \frac{1}{t} \int_0^t \left[\alpha_2 + \beta_{21}X_1(s) - \beta_{22}X_2(s) - \frac{\tau_{21}^2}{2} - \frac{\tau_{22}^2}{2} \right] ds \\ &\quad + \tau_{21} \frac{W_1(t)}{t} + \tau_{22} \frac{W_2(t)}{t}. \end{aligned}$$

By the strong law of large numbers, when the solution $\mathbf{X}(t) = (X_1(t), X_2(t))^T$ of (3.1) is close to $\text{supp}(\delta^*) = \{(0, 0)\}$ for a long time (meaning that the solution $\mathbf{X}(t)$ is within some small neighborhood of $(0, 0)$ for a long time), the Lyapunov exponent for the first component (the invasion rate of Species 1) with respect to μ_0 is

$$\lambda_1(\mu_0) = \lim_{t \rightarrow \infty} \frac{\ln X_1(t)}{t}$$

$$\begin{aligned}
&= \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t \left[\alpha_1 - \beta_{11}X_1(s) + \beta_{12}X_2(s) - \frac{\tau_{11}^2}{2} - \frac{\tau_{12}^2}{2} \right] ds \\
&= \int_{\partial \mathbb{R}_+^2} \left[\alpha_1 - \beta_{11}X_1 + \beta_{12}X_2 - \frac{\tau_{11}^2}{2} - \frac{\tau_{12}^2}{2} \right] \mu_0(d\mathbf{X}) \\
&= \alpha_1 - \frac{\tau_{11}^2}{2} - \frac{\tau_{12}^2}{2}.
\end{aligned}$$

Similarly,

$$\begin{aligned}
\lambda_2(\mu_0) &= \int_{\partial \mathbb{R}_+^2} \left[\alpha_2 + \beta_{21}X_1 + \beta_{22}X_2 - \frac{\tau_{21}^2}{2} - \frac{\tau_{22}^2}{2} \right] \mu_0(d\mathbf{X}) \\
&= \alpha_2 - \frac{\tau_{21}^2}{2} - \frac{\tau_{22}^2}{2}.
\end{aligned}$$

By the similar computations, the Lyapunov exponents of μ_1 and μ_2 can be computed as

$$\begin{aligned}
\lambda_1(\mu_1) &= 0, \\
\lambda_2(\mu_1) &= \alpha_2 - \frac{\tau_{21}^2}{2} - \frac{\tau_{22}^2}{2} + \frac{\beta_{21}}{\beta_{11}}\alpha_1 - \frac{\beta_{21}}{2\beta_{11}}(\tau_{11}^2 + \tau_{12}^2), \\
\lambda_1(\mu_2) &= \alpha_1 - \frac{\tau_{11}^2}{2} - \frac{\tau_{12}^2}{2} + \frac{\beta_{12}}{\beta_{22}}\alpha_2 - \frac{\beta_{12}}{2\beta_{22}}(\tau_{21}^2 + \tau_{22}^2), \\
\lambda_2(\mu_2) &= 0.
\end{aligned}$$

On the basis of the Lyapunov exponents of the set of all ergodic invariant probability measures of the system (3.1) on the boundary $\partial \mathbb{R}_+^2$, denoted $\mathcal{M} = \{\mu_0, \mu_1, \mu_2\}$, we can give the complete classification of the dynamics of the system (3.1). First of all, we assume that the matrix

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix},$$

where $\sigma_{ij} = \tau_{i1}\tau_{j1} + \tau_{i2}\tau_{j2}$ ($i, j = 1, 2$), is positive definite. This is equivalent to requiring that

$$\tau_{11}^2 + \tau_{12}^2 > 0, \quad \tau_{21}^2 + \tau_{22}^2 > 0, \quad \text{and} \quad \tau_{11}\tau_{22} - \tau_{12}\tau_{21} \neq 0.$$

Next, we claim that the family of transition probability functions of the solution to the system (3.1) is tight when the system (3.1) is either competitive or predator-prey. Furthermore, when the system (3.1) is mutualistic, the family of transition probability functions of its solution might not be tight. In this case, we show that the solution blows up in finite time almost surely or that there is no invariant measure on $\mathbb{R}_+^{2,\circ}$. Indeed, we first assume $\beta_{12} < 0$ and $\beta_{21} < 0$. In order to prove the tightness of the solution of (3.1), it suffices to show that there is a $\delta > 0$ such that

$$\frac{X_1 f_1(\mathbf{X}) + X_2 f_2(\mathbf{X})}{1 + X_1 + X_2} - \frac{\sum_{i,j=1}^2 \sigma_{ij} X_i X_j}{2(1 + X_1 + X_2)^2} + \delta(3 + |f_1(\mathbf{X})| + |f_2(\mathbf{X})|) < 0 \quad (3.4)$$

for any $\|\mathbf{X}\| = |X_1| + |X_2|$ that is sufficiently large, where $f_1(\mathbf{X}) = X_1(\alpha_1 - \beta_{11}X_1 + \beta_{12}X_2)$ and $f_2(\mathbf{X}) = X_2(\alpha_2 + \beta_{21}X_1 - \beta_{22}X_2)$. Since β_{12} and β_{21} are both negative, as $\|\mathbf{X}\| \rightarrow \infty$, we get

$$\frac{X_1 f_1(\mathbf{X}) + X_2 f_2(\mathbf{X})}{(1 + X_1 + X_2)^2} = \frac{\alpha_1 X_1 + \alpha_2 X_2 - (\beta_{11}X_1^2 + \beta_{22}X_2^2) + (\beta_{12} + \beta_{21})X_1 X_2}{(1 + X_1 + X_2)^2}$$

which approaches $-\infty$. So there is a $\tilde{\beta} > 0$ such that for any $\|\mathbf{X}\|$ that is sufficiently large

$$\frac{X_1 f_1(\mathbf{X}) + X_2 f_2(\mathbf{X})}{1 + X_1 + X_2} < -\tilde{\beta}(1 + X_1 + X_2).$$

On the other hand, using the Cauchy–Schwarz’s inequality, a $\tilde{\sigma} > 0$ exists such that

$$-\frac{\sum_{i,j=1}^2 \sigma_{ij} X_i X_j}{2(1 + X_1 + X_2)^2} \leq -\tilde{\sigma} \frac{X_1^2 + X_2^2}{(1 + X_1 + X_2)^2} \leq -\tilde{\sigma}$$

when $\|\mathbf{X}\|$ is sufficiently large. Then we can choose a $\delta > 0$ that is small enough so that we can obtain the inequality (3.4) as $\|\mathbf{X}\|$ is large enough. Second, we suppose either $(\beta_{12} < 0, \beta_{21} > 0)$ or $(\beta_{12} > 0, \beta_{21} < 0)$. Then $c_1 > 0$ and $c_2 > 0$ exist such that $c_1 \beta_{12} + c_2 \beta_{21} = 0$. This implies that

$$\lim_{\|\mathbf{X}\| \rightarrow \infty} \frac{c_1 X_1 f_1(\mathbf{X}) + c_2 X_2 f_2(\mathbf{X})}{(1 + c_1 X_1 + c_2 X_2)^2} = -\infty.$$

By the same reasoning as in the competitive case, we can easily find a $\delta' > 0$ that is sufficiently small such that, for any $\|\mathbf{X}\|$ large enough, we get the inequality (3.4). Lastly, for the case $\beta_{12} > 0$ and $\beta_{21} > 0$, we show that the solution of (3.1) blows up in finite time almost surely under the assumptions that

$$\alpha_1 - \frac{\tau_{11}^2}{2} - \frac{\tau_{12}^2}{2} > 0, \quad \alpha_2 - \frac{\tau_{21}^2}{2} - \frac{\tau_{22}^2}{2} > 0, \quad \text{and} \quad \beta_{11}\beta_{22} - \beta_{12}\beta_{21} \leq 0.$$

By way of contradiction, assume that $\mathbf{X}(t)$ does not blow up in finite time and has an invariant probability measure on $\mathbb{R}_+^{2,0}$. As a result, $\mathbf{X}(t)$ is a recurrent process. But then, by Ito’s formula,

$$\begin{aligned} \frac{\beta_{22} \ln X_1(t) + \beta_{12} \ln X_2(t)}{t} &= \frac{\beta_{22} \ln X_1(0) + \beta_{12} \ln X_2(0)}{t} \\ &+ \beta_{22} \left(\alpha_1 - \frac{\tau_{11}^2}{2} - \frac{\tau_{12}^2}{2} \right) + \beta_{12} \left(\alpha_2 - \frac{\tau_{21}^2}{2} - \frac{\tau_{22}^2}{2} \right) \\ &+ (\beta_{12}\beta_{21} - \beta_{11}\beta_{22}) \frac{1}{t} \int_0^t X_1(s) ds \\ &+ (\tau_{11}\beta_{22} + \tau_{21}\beta_{12}) \frac{W_1(t)}{t} \\ &+ (\tau_{12}\beta_{22} + \tau_{22}\beta_{12}) \frac{W_2(t)}{t}. \end{aligned}$$

It follows that, by the above assumptions,

$$\limsup_{t \rightarrow \infty} \frac{\beta_{22} \ln X_1(t) + \beta_{12} \ln X_2(t)}{t} > 0 \text{ a.s.}$$

which is a contradiction. Therefore, the system (3.1), which is competitive or predator–prey, satisfies Assumption 3 in Subsection 2.2.

Now we focus on the stochastic model (3.1) of the competitive type and predator–prey type. On the basis of Theorems 2.3, 2.4, and 2.5 in Subsection 2.2, we are able to describe the complete dynamic picture of the system (3.1) as follows.

C1. Suppose $\lambda_1(\mu_0) < 0$ and $\lambda_2(\mu_0) < 0$. This means that there are no other ergodic invariant probability measures in $\partial\mathbb{R}_+^2$ in addition to μ_0 . Then $\mathcal{M} = \{\mu_0\}$, which implies that Assumptions 3 and 5 hold for the system (3.1). By Theorem 2.4, the solution $\mathbf{X}(t) = (X_1(t), X_2(t))^T$ converges a.s. to $(0, 0)^T$ for any initial condition $\mathbf{x} = (x_1, x_2) \in \mathbb{R}_+^{2,\circ}$.

C2. Suppose $\lambda_1(\mu_0) > 0$ and $\lambda_2(\mu_0) < 0$. Then $\mathcal{M} = \{\mu_0, \mu_1\}$. There are two cases.

- (i) If $\lambda_2(\mu_1) > 0$ then Assumption 4 is fulfilled and hence transition probability function of the solution to the system (3.1) converges to an invariant probability measure $\bar{\mu}_1$ in $\mathbb{R}_+^{2,\circ}$ in total variation norm.
- (ii) If $\lambda_2(\mu_1) < 0$ then $\mathcal{M}^1 = \{\mu_1\}$ and so $\mathcal{M}^2 = \mathcal{M} \setminus \mathcal{M}^1 = \{\mu_0\}$. This means that Assumptions 3 and 6 are satisfied. By Theorem 2.5, $X_2(t)$ converges to 0 with the exponential rate $\lambda_2(\mu_1)$ and the family of random occupation measures of the solution converges to μ_1 .

C3. Suppose $\lambda_1(\mu_0) < 0$ and $\lambda_2(\mu_0) > 0$. Then $\mathcal{M} = \{\mu_0, \mu_2\}$. There are 2 cases:

- (i) If $\lambda_1(\mu_2) > 0$, then Assumption 4 is fulfilled, and hence the transition probability function of the solution to the system (3.1) converges to an invariant probability measure $\bar{\mu}_2$ in $\mathbb{R}_+^{2,\circ}$ in the total variation norm.
- (ii) If $\lambda_1(\mu_2) < 0$, then $\mathcal{M}^1 = \{\mu_2\}$ and so $\mathcal{M}^2 = \mathcal{M} \setminus \mathcal{M}^1 = \{\mu_0\}$. This means that Assumptions 3 and 6 are satisfied. By Theorem 2.5, $X_1(t)$ converges to 0 with the exponential rate $\lambda_1(\mu_2)$, and the family of random occupation measures of the solution converges to μ_2 .

C4. Suppose $\lambda_1(\mu_0) > 0$ and $\lambda_2(\mu_0) > 0$. Then $\mathcal{M} = \{\mu_0, \mu_1, \mu_2\}$. There are four cases.

- (i) If $\lambda_1(\mu_2) > 0$ and $\lambda_2(\mu_1) > 0$, then any invariant probability measure on $\partial\mathbb{R}_+^2$ has the form $\mu = p_0\mu_0 + p_1\mu_1 + p_2\mu_2$, where p_0, p_1 , and p_2 are positive numbers such that $p_0 + p_1 + p_2 = 1$. It is straightforward that $\max\{\lambda_1(\mu), \lambda_2(\mu)\} > 0$. Then Assumptions 3 and 4 are satisfied. By Theorem 2.3, there is a unique invariant probability measure μ^* on $\mathbb{R}_+^{2,\circ}$, and the transition probability function $P(t, \mathbf{x}, \cdot)$, $\mathbf{x} \in \mathbb{R}_+^{2,\circ}$, converges to μ^* in the total variation norm.
- (ii) If $\lambda_1(\mu_2) < 0$ and $\lambda_2(\mu_1) < 0$, then $\mathcal{M}^1 = \{\mu_1, \mu_2\}$ and so $\mathcal{M}^2 = \mathcal{M} \setminus \mathcal{M}^1 = \{\mu_0\}$. Hence, Assumptions 3 and 6 hold true. By Theorem 2.5, $p_1^x > 0$, $p_2^x > 0$, and $p_1^x + p_2^x = 1$, where

$$p_1^x = \mathbb{P}_x \left\{ \mathcal{U}(\omega) = \{\mu_1\} \text{ and } \lim_{t \rightarrow \infty} \frac{\ln X_2(t)}{t} = \lambda_2(\mu_1) < 0 \right\},$$

$$p_2^x = \mathbb{P}_x \left\{ \mathcal{U}(\omega) = \{\mu_2\} \text{ and } \lim_{t \rightarrow \infty} \frac{\ln X_1(t)}{t} = \lambda_1(\mu_2) < 0 \right\}.$$

(iii) If $\lambda_1(\mu_2) > 0$ and $\lambda_2(\mu_1) < 0$, then the conclusion is the same as in C2(ii).

(iv) If $\lambda_1(\mu_2) < 0$ and $\lambda_2(\mu_1) > 0$, then the conclusion is the same as in C3(ii).

Finally, to study the resilience of the two-dimensional stochastic gLV model (3.1), we assume that the system (3.1) is competitive or predator–prey type with the matrix Σ being positive-definite and satisfies the inequality (3.4) for some $\delta > 0$. Furthermore, we assume that the system (3.1) satisfies the assumptions as in C2(i) or C3(i) or C4(i).

Then the stochastic gLV model of two species is stochastically strongly persistent, that is, it has a unique positive stationary distribution π^* on $\mathbb{R}_+^{2,\circ}$ such that the transition probability function of its solution converges to π^* in the total variation norm. The linearized system of (3.1) at π^* is given by

$$\begin{aligned} d\tilde{X}_1 &= [\Phi_1 + \gamma_{11}(\tilde{X}_1 - m_1) + \gamma_{12}(\tilde{X}_2 - m_2)]dt + \tilde{X}_1(\tau_{11}dW_1 + \tau_{12}dW_2), \\ d\tilde{X}_2 &= [\Phi_2 + \gamma_{21}(\tilde{X}_1 - m_1) + \gamma_{22}(\tilde{X}_2 - m_2)]dt + \tilde{X}_2(\tau_{21}dW_1 + \tau_{22}dW_2), \end{aligned} \quad (3.5)$$

where the values of θ_i and the γ_{ij} are determined by

$$\begin{aligned} \Phi_1 &= \alpha_1 m_1 - \beta_{11} m_1^2 + \beta_{12} m_1 m_2, \\ \Phi_2 &= \alpha_2 m_2 + \beta_{21} m_2 m_1 - \beta_{22} m_2^2, \\ \gamma_{11} &= \alpha_1 + 2\beta_{11} m_1 + \beta_{12} m_2, \\ \gamma_{12} &= \beta_{12} m_1, \\ \gamma_{21} &= \beta_{21} m_2, \\ \gamma_{22} &= \alpha_2 + \beta_{21} m_1 + 2\beta_{22} m_2, \end{aligned}$$

with

$$m_i = \int_{\partial\mathbb{R}_+^{2,\circ}} X_i \pi^*(dX_1 dX_2) = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t X_i(s) ds \quad (i = 1, 2).$$

Let $A = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix}$, $\Theta = (\Phi_1, \Phi_2)^T$, $B_1 = \begin{bmatrix} \tau_{11} & 0 \\ 0 & \tau_{21} \end{bmatrix}$, $B_2 = \begin{bmatrix} \tau_{12} & 0 \\ 0 & \tau_{22} \end{bmatrix}$, $\sigma_{ij} = \tau_{i1}\tau_{j1} + \tau_{i2}\tau_{j2}$ ($i, j = 1, 2$), and $Y_i = \tilde{X}_i - m_i$ ($i = 1, 2$). Then the system (3.5) becomes

$$d\mathbf{Y} = (A\mathbf{Y} + \Theta)dt + (B_1\mathbf{Y} + B_1\mathbf{m})dW_1 + (B_2\mathbf{Y} + B_2\mathbf{m})dW_2 \quad (3.6)$$

where $\mathbf{Y} = (Y_1, Y_2)^T$ and $\mathbf{m} = (m_1, m_2)^T$. Next, by computation, we have

$$\begin{aligned} I_2 \otimes A &= \begin{bmatrix} \gamma_{11} & \gamma_{12} & 0 & 0 \\ \gamma_{21} & \gamma_{22} & 0 & 0 \\ 0 & 0 & \gamma_{11} & \gamma_{12} \\ 0 & 0 & \gamma_{21} & \gamma_{22} \end{bmatrix}, A \otimes I_2 = \begin{bmatrix} \gamma_{11} & 0 & \gamma_{12} & 0 \\ 0 & \gamma_{11} & 0 & \gamma_{12} \\ \gamma_{21} & 0 & \gamma_{22} & 0 \\ 0 & \gamma_{21} & 0 & \gamma_{22} \end{bmatrix}, \\ B_1 \otimes B_1 &= \begin{bmatrix} \tau_{11}^2 & 0 & 0 & 0 \\ 0 & \tau_{11}\tau_{21} & 0 & 0 \\ 0 & 0 & \tau_{21}\tau_{11} & 0 \\ 0 & 0 & 0 & \tau_{21}^2 \end{bmatrix}, B_2 \otimes B_2 = \begin{bmatrix} \tau_{12}^2 & 0 & 0 & 0 \\ 0 & \tau_{12}\tau_{22} & 0 & 0 \\ 0 & 0 & \tau_{22}\tau_{12} & 0 \\ 0 & 0 & 0 & \tau_{22}^2 \end{bmatrix}. \end{aligned}$$

Thus $S = I_2 \otimes A + A \otimes I_2 + B_1 \otimes B_1 + B_2 \otimes B_2$, which is equal to

$$S = \begin{bmatrix} 2\gamma_{11} + \sigma_{11} & \gamma_{12} & \gamma_{12} & 0 \\ \gamma_{21} & \gamma_{11} + \gamma_{22} + \sigma_{12} & 0 & \gamma_{12} \\ \gamma_{21} & 0 & \gamma_{11} + \gamma_{22} + \sigma_{21} & \gamma_{12} \\ 0 & \gamma_{21} & \gamma_{21} & 2\gamma_{22} + \sigma_{22} \end{bmatrix},$$

and $T = \text{vec}(\Theta M^T + M \Theta^T) + [B_1 \otimes B_1 + B_2 \otimes B_2] \text{vec}(M m^T + m M^T + m m^T)$, which is equal to

$$T = \begin{bmatrix} 2m_1\theta_1 + \sigma_{11}(2M_1m_1 + m_1^2) \\ m_1\theta_2 + m_2\theta_1 + \sigma_{12}(M_2m_1 + m_2M_1 + m_2m_1) \\ m_1\theta_2 + m_2\theta_1 + \sigma_{21}(M_1m_2 + m_1M_2 + m_1m_2) \\ 2m_2\theta_2 + \sigma_{22}(2M_2m_2 + m_2^2) \end{bmatrix},$$

in which $M = (M_1, M_2)^T = (\mathbb{E}Y_1, \mathbb{E}Y_2)^T$ is the solution to the ODE system

$$\dot{M} = AM + \Theta, \quad (3.7)$$

with the initial value $M(0) = \mathbf{u}$. If we let $\mathbf{Z} = (\mathbb{E}Y_1^2, \mathbb{E}Y_2Y_1, \mathbb{E}Y_1Y_2, \mathbb{E}Y_2^2)^T$, then we can smooth out the solutions of (3.6) by looking at the solutions of the following ODE system

$$\frac{d\mathbf{Z}}{dt} = S\mathbf{Z} + T, \quad (3.8)$$

By using two ODE systems (3.7) and (3.8), we will develop graphical representations of the four resilience measures, which are the instantaneous return rate, the average return rate, the asymptotic resilience, and the convergence rate, for the stochastic system (3.1) of the competitive type over the relevant parameter regimes in the next subsection.

3.2. Recovery after pulse perturbation

As a simple illustration of how environmental noise and microbial interactions combine to drive the rate of return of a microbial community to the equilibrium state after a pulse perturbation, we look at a SgLV model of two competitive species (3.1). Here, we assume that two species are in direct competition with each other, so that their interaction strengths β_{12} and β_{21} are negative.

In each panel of Figure 3, two trajectories of the system (3.1) are produced, in which the trajectories before time zero represent the equilibrium states of the two species and the trajectories after time zero are the recovery dynamics of the two species following a pulse perturbation at time zero. The difference between Figure 3A and Figure 3C is due to the difference in the interaction strength β_{12} , which measures the strength of interspecific competition of Species 2 on Species 1. In Figure 3A, β_{12} is twice as large as in Figure 3C, while the noise intensities are the same in both plots. In contrast, the difference between Figure 3A and Figure 3B is caused by the difference in the noise intensities of two white noises: In Figure 3B, the noise intensities of $\frac{dW_1}{dt}$ and $\frac{dW_2}{dt}$ are five times those in Figure 3A, while the interaction strengths between species are the same, leading to greater fluctuations in the species' abundances. Figure 3D differs from Figure 3A in that the noise intensities increase by a factor of 5 and the interaction strength β_{12} decreases by a factor of 2.

Resilience in a pulse-perturbed stochastic system can be illustrated by this example. When comparing Figure 3A and Figure 3C, we observe that decreasing the interaction strength β_{12} reduces the return time of the system to the equilibrium state after a pulse perturbation. That is, lower competition between two species intensifies the resilience of the system under the same environmental regime. When comparing Figure 3A and Figure 3B, increasing the noise intensities decreases the time required for recovery trajectories to return to an equilibrium state. In this example, it appears that increasing environmental noise actually increases stability, which contrasts with the intuition that increasing noise

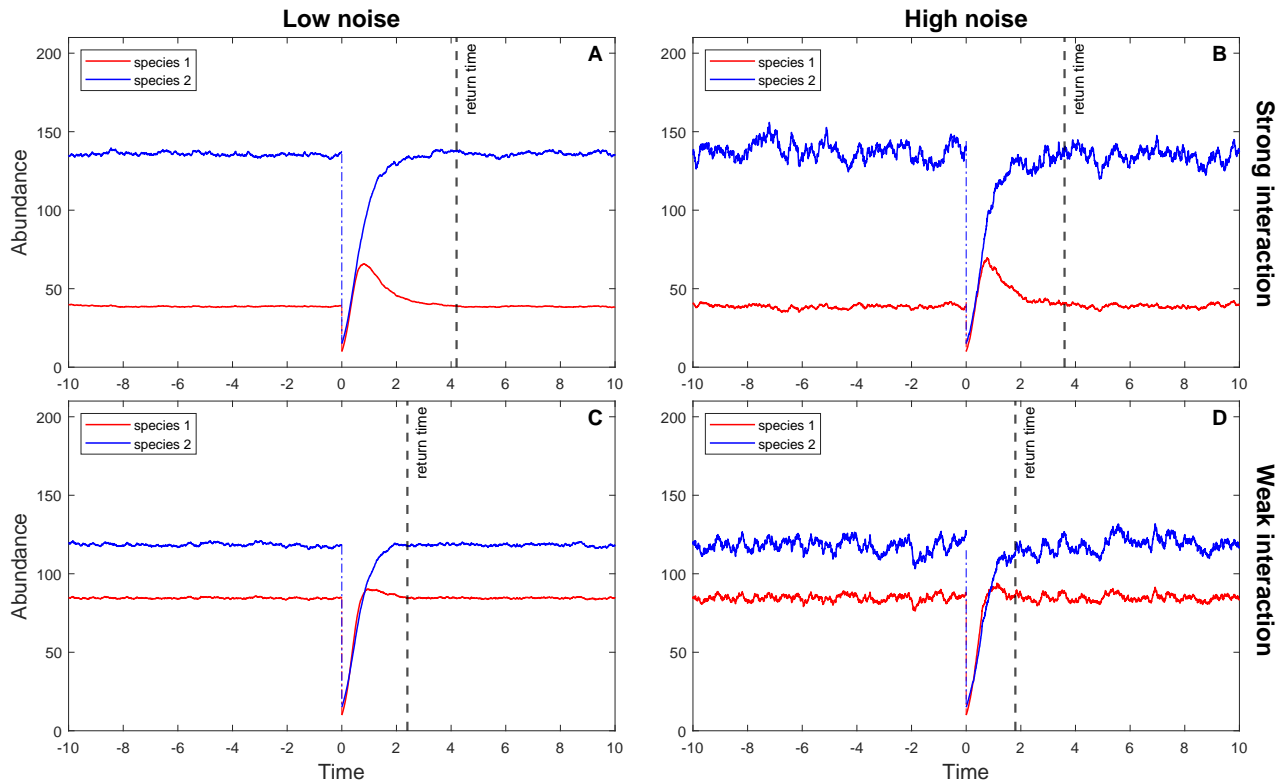


Figure 3. Dynamics of the stochastic system given by (3.1) with four parameter sets. Panel A corresponds to the parameter Set 1 (low noise, strong interaction): $(\alpha_1, \alpha_2) = (6, 4)$, $(\beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}) = (0.05, -0.03, -0.01, 0.027)$, and $(\tau_{11}, \tau_{12}, \tau_{21}, \tau_{22}) = (0.01, 0.015, 0.01, 0.02)$. Panel B corresponds to the parameter Set 2 (high noise, strong interaction): $(\alpha_1, \alpha_2) = (6, 4)$, $(\beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}) = (0.05, -0.03, -0.01, 0.027)$, and $(\tau_{11}, \tau_{12}, \tau_{21}, \tau_{22}) = (0.05, 0.075, 0.05, 0.1)$. Panel C corresponds to the parameter Set 3 (low noise, weak interaction): $(\alpha_1, \alpha_2) = (6, 4)$, $(\beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}) = (0.05, -0.015, -0.01, 0.027)$, and $(\tau_{11}, \tau_{12}, \tau_{21}, \tau_{22}) = (0.01, 0.015, 0.01, 0.02)$. Panel D corresponds to the parameter Set 4 (high noise, weak interaction): $(\alpha_1, \alpha_2) = (6, 4)$, $(\beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}) = (0.05, -0.03, -0.01, 0.027)$, and $(\tau_{11}, \tau_{12}, \tau_{21}, \tau_{22}) = (0.05, 0.075, 0.05, 0.1)$.

destabilizes the stability of the system [21–23]. Figure 3D shows the combined effect of environmental noise and interactions between species on the resilience of a stochastic system. In the next subsection, we will use our theoretical framework to test our observations in Figure 3.

3.3. Noise drives resilience

To demonstrate the usefulness of our theoretical framework developed above, we apply the procedure to calculate the resilience of a SgLV (described in Figure 2) to the simulated data in Figure 3. If we observe Figure 3 in Section 2, the recovery trajectories of two species return to an equilibrium state at time 4.2 after a pulse perturbation at time zero in Panel A. Reducing the interaction strength

β_{12} by 2-fold relative to Panel A, Panel C shows that the return time to the equilibrium state is reduced to approximately 2.4 after the same pulse perturbation as in Panel A at time zero. This means that the microbial community in Panel C is more resilient than that in Panel A. Increasing the noise intensities five times compared with Panel A, the return time to the equilibrium state in Panel B decreases to some time around 3.6 with the same perturbation as in Panel A. In other words, the microbial community in Panel C is more resilient than that in Panel A. This shows that environmental noise influences the resilience of the microbial community. Panel D shows the fastest return time, which is approximately 1.8, to the equilibrium state among communities when we decrease the interaction strength β_{12} two times and increase the noise intensities five times. Below, we calculate four resilience measures for each panel of Figure 3 to confirm our observation that the resilience of the microbial community is driven by environmental noise and to disentangle the sources of the effects of environmental fluctuations on the resilience of a microbiome system. To calculate the resilience measures, the first step is to lin-

Table 1. Coefficients of the linearized stochastic system (3.5) are computed in four scenarios that correspond to the four parameter sets in Figure 3.

	Parameter set 1	Parameter set 2	Parameter set 3	Parameter set 4
m_1	38.7470	38.8413	84.4804	84.4508
m_2	135.3518	134.9853	118.2069	117.8869
Φ_1	0.0813	0.3254	0.2431	0.7731
Φ_2	0.4263	1.6171	0.3562	1.3958
γ_{11}	-1.9353	-1.9337	-4.2211	-4.2134
γ_{12}	-1.1624	-1.1652	-1.2672	-1.2668
γ_{21}	-1.3535	-1.3499	-1.1821	-1.1789
γ_{22}	-3.6062	-3.5876	-3.1492	-3.1318

earize the stochastic system (3.1) at its stationary distribution π^* , and then we can obtain the linearized stochastic system (3.5), where all coefficients Φ_i , γ_{ij} , and m_i ($i, j = 1, 2$) are computed and given in Table 1. Note that $(m_1, m_2)^T$ is the expectation of the stationary distribution π^* . Figures 4A, 4B, 5A, and 5B show that the stochastic dynamics of both the original stochastic system (3.1) and the linearized stochastic system (3.5) converge to the equilibrium state characterized by the stationary distribution π^* following the same pulse perturbation at time 0 in four scenarios corresponding to the four parameter sets in Table 1. In Figure 4A, we see that the trajectories of the linearized system approach and merge into those of the original system between time 4 and 4.5, while the time in Figure 4B is between 3 and 3.5. In Figure 5A, the merging time of the trajectories of both systems is between 2.5 and 3. The fastest merging time between 1.5 and 2 is shown in Figure 5B. Thus, the patterns of the return time to the equilibrium state in both the original system and the linearized system when applied to the same pulse perturbation are the same, even though the dynamics of the trajectories of both systems are quite different before the merging time. Therefore, we are able to utilize a much simpler linearized system (3.5) to investigate the resilience of a microbial community whose temporal dynamics are captured by the original system (3.1).

Next, we transform the linearized system (3.5) into the two linear ODE systems (3.7) and (3.8) by, respectively, finding the first and second moment of solutions and then using vectorization in linear algebra. Then four resilience measures $\mathcal{R}_t^{\text{ins}}$, $\mathcal{R}_t^{\text{ave}}$, \mathcal{R}_c , and \mathcal{R}_∞ are calculated on the basis of $W(t)$,

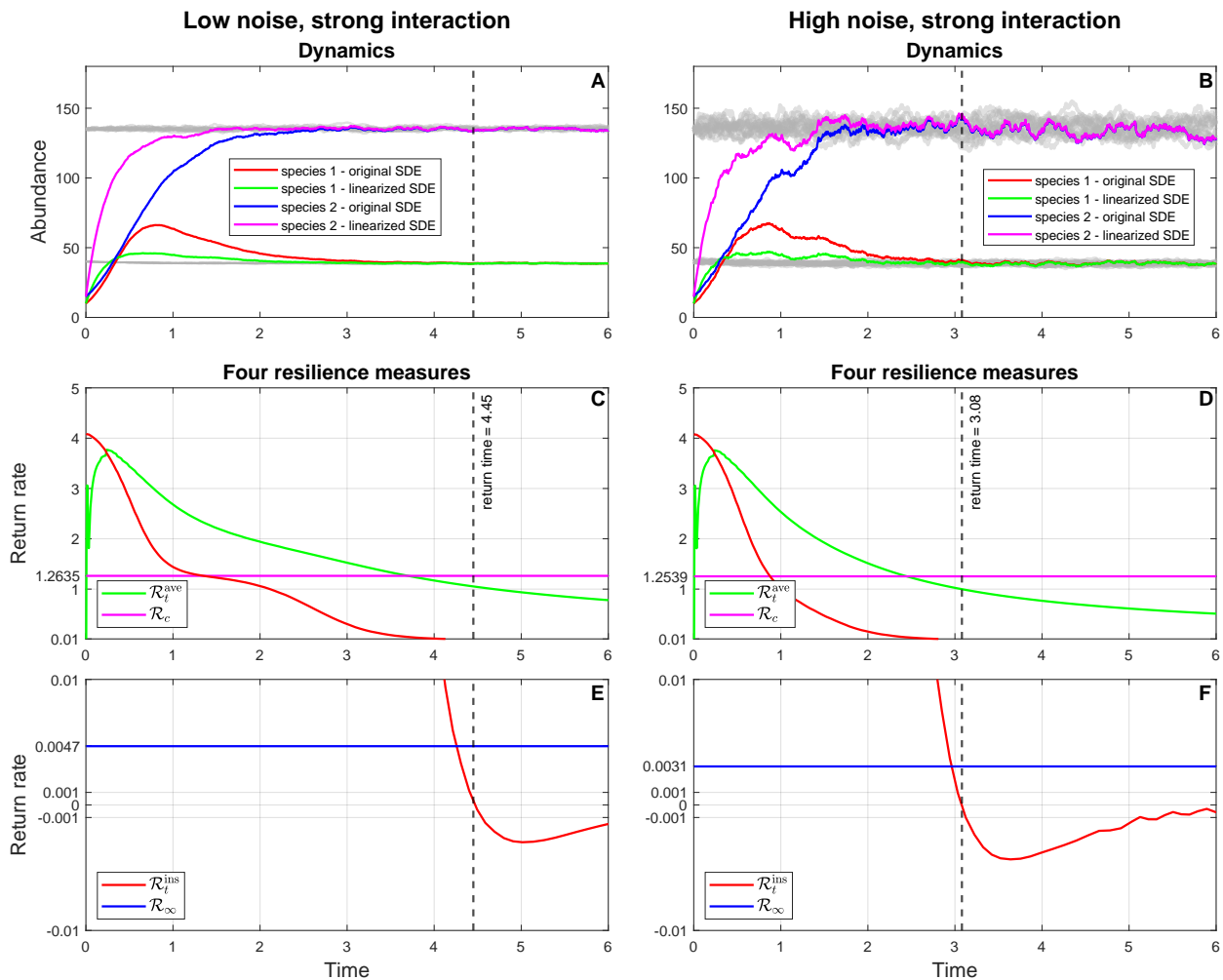


Figure 4. Recovery dynamics (Panels A and B) of both the linearized system (3.1) and the original system (3.5) following the same pulse perturbation (10, 15), and the dynamics of four resilience measures (average return rate, rate of convergence (Panels C and D) and instantaneous return rate, asymptotic resilience (Panels E and F)) in two scenarios of parameter sets 1 and 2.

the mean square difference between the solution of the linearized system (3.5), the expectation of the stationary distribution π^* , and the mean square stability matrix S . Note that the mean square stability matrix S and $W(t)$ can be calculated from the linearized coefficients shown in Table 1. The last four panels of Figure 4 and Figure 5 show the dynamics of four resilience measures over time in the four scenarios above.

Two resilience measures $\mathcal{R}_t^{\text{ins}}$ and $\mathcal{R}_t^{\text{ave}}$, which are deterministic functions of time, can help us track the short-term behavior of $\ln W(t)$. When $\mathcal{R}_t^{\text{ins}} > 0$, $\frac{d}{dt} \ln W(t)^{1/2} < 0$, which implies that $\ln W(t)$ is decreasing to its equilibrium point at time t . In other words, when $\mathcal{R}_t^{\text{ins}} > 0$, the solution $\tilde{X}(t)$ of the linearized system (3.5) moves closer to the equilibrium state π^* at time t . By the same argument, when

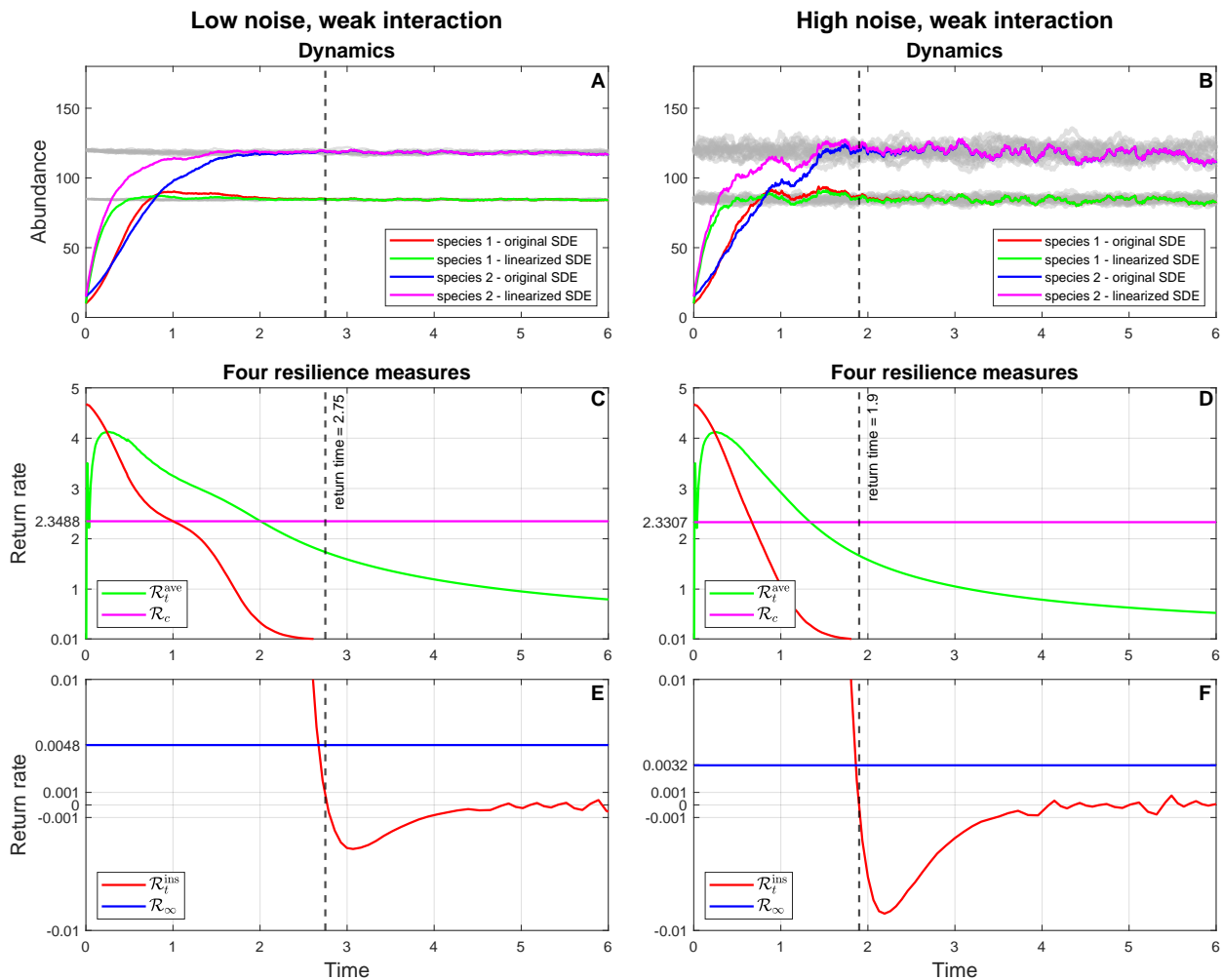


Figure 5. Recovery dynamics (Panels A and B) of both the linearized system (3.1) and the original system (3.5) following the same pulse perturbation (10, 15), and the dynamics of four resilience measures (average return rate, rate of convergence (Panels C and D) and instantaneous return rate, asymptotic resilience (Panels E and F)) in two scenarios of parameter sets 3 and 4.

$\mathcal{R}_t^{\text{ins}} < 0$, $\ln W(t)$ is increasing, and hence the solution $\tilde{X}(t)$ moves away from the equilibrium state π^* at time t . In terms of $\mathcal{R}_t^{\text{ave}}$, if $\mathcal{R}_t^{\text{ave}} > 0$, then $\ln W(t) < \ln W(0)$, which means that the solution $\tilde{X}(t)$ comes closer to the equilibrium state at time t , since the end of the perturbation at time 0. If $\mathcal{R}_t^{\text{ave}} < 0$, then the solution $\tilde{X}(t)$ moves far away from the equilibrium state since the end of the perturbation. The deterministic function $\mathcal{R}_t^{\text{ins}}$ measures the rate of change in $\ln W(t)$. So we can use it to track when $\ln W(t)$ comes back to its equilibrium point after the perturbation is applied. It is straightforward to see that the time when $\ln W(t)$ reaches its equilibrium point (which is a constant) is the time when $\mathcal{R}_t^{\text{ins}}$ crosses 0. Therefore, we can estimate the return time of the solution $\tilde{X}(t)$ to the equilibrium state π^* when the first time $\mathcal{R}_t^{\text{ins}}$ crosses 0. In the last four panels of Figure 4, the return time to the equilibrium

Table 2. The difference $\mathcal{R}_c - \mathcal{R}_\infty$ computed from the linearized coefficients in Table 1 in four scenarios corresponding to the four parameter sets in Figure 3.

Low noise, strong interaction	High noise, strong interaction
$\mathcal{R}_c^1 = 1.2635$	$\mathcal{R}_c^2 = 1.2539$
$\mathcal{R}_\infty^1 = 0.0047$	$\mathcal{R}_\infty^2 = 0.0031$
$\mathcal{R}_c^1 - \mathcal{R}_\infty^1 = 1.2588$	$\mathcal{R}_c^2 - \mathcal{R}_\infty^2 = 1.2508$
Low noise, weak interaction	High noise, weak interaction
$\mathcal{R}_c^3 = 2.3488$	$\mathcal{R}_c^4 = 2.3307$
$\mathcal{R}_\infty^3 = 0.0048$	$\mathcal{R}_\infty^4 = 0.0032$
$\mathcal{R}_c^3 - \mathcal{R}_\infty^3 = 2.3440$	$\mathcal{R}_c^4 - \mathcal{R}_\infty^4 = 2.3275$

state decreases from 4.45 to 3.08 when the noise intensities are increased by five times. The same decreasing pattern in the return time occurs in the last four panels of Figure 5, which is that the return times go down from 2.75 to 1.9. This pattern can be explained by the physical meaning of $W(t)$. By our definition in Section 2.4, $W(t)$ is the sum of all the variances of the solution components of the linearized system (3.5), which measures how far the solution of the linearized system (3.5) is from the equilibrium state π^* . Then high noise makes the equilibrium state π^* have higher variance so that there is less recovery required following perturbation to return to π^* . So, the return to the equilibrium state π^* is actually faster than in the system with high noise than with low noise.

Two resilience measures \mathcal{R}_c and \mathcal{R}_∞ , which are constants, can be used to predict the resilience of the stochastic system (3.1) by looking at the gap $\mathcal{R}_c - \mathcal{R}_\infty$ and how fast the average return rate $\mathcal{R}_t^{\text{ave}}$ enters this gap. In Table 2, when we increase the noise five times, the gap $\mathcal{R}_c - \mathcal{R}_\infty$ decreases from 1.2588 (parameter Set 1: low noise, strong interaction) to 1.2508 (parameter Set 2: high noise, strong interaction). From Panels C and D in Figure 4, we also observe that the time when the average return rate $\mathcal{R}_t^{\text{ave}}$ moves into the gap between \mathcal{R}_c and \mathcal{R}_∞ is shorter when the noise increases. This pattern is the same as we move from parameter Set 3 (low noise, weak interaction) to parameter Set 4 (high noise, weak interaction). The pattern of a decreased gap implies shorter for moving into the time gap, caused by increasing noise, which can be elucidated by looking at the mathematical meanings of the two resilience measures \mathcal{R}_c and \mathcal{R}_∞ . First, the quantity \mathcal{R}_c measures how fast the covariant matrix $P(t)$, which is the solution of the system (2.19), converges to its equilibrium state. Basically, \mathcal{R}_c equals the negative half of the real part of the largest eigenvalue of the mean square stability matrix S , which means that \mathcal{R}_c tells us about the smallest convergence rate among the convergence rates of the elements of the covariant matrix $P(t)$. While \mathcal{R}_∞ can be considered as an equilibrium state of the average return rate $\mathcal{R}_t^{\text{ave}}$, and so \mathcal{R}_∞ approximately equals the smallest convergence rate among the convergence rates of the diagonal elements of the covariance matrix $P(t)$. Therefore, the gap $\mathcal{R}_c - \mathcal{R}_\infty$ is always positive when noise appears in the original system (3.1). Without noise, this gap will be zero, and hence these two resilience measures coincide. Our simulations (Figures 4 and 5) and calculations (Table 2) demonstrate that increasing noise will decrease the gap $\mathcal{R}_c - \mathcal{R}_\infty$, which, in turn, decreases the time when $\mathcal{R}_t^{\text{ave}}$ enters the gap. This implies that increasing the level of noise increases the resilience of the system.

4. Discussion

The concept of resilience has been used for a long time in community ecology, although it is much less developed for microbial systems, even though resilience is at the center of many pressing questions in microbiome research. Microbiome communities differ from macroecological communities in at least two important ways that impact the modeling of ecological stability. First, while stability metrics in macroecological systems have been developed for first-order multivariate autoregressive models [21, 22], the nonlinear generalized Lotka–Volterra model is more commonly used in microbiome studies. This kind of model has been applied to study time-series inference and stability in intestinal microbiota [7, 38], to model primary succession of murine gut communities [35], and to capture dynamics of complex microbial metapopulations in aquatic environments [36, 37]. Second, and perhaps more importantly, microbial time-series data are incredibly noisy. Some of this noise is measurement error due to the sampling and sequencing processes involved in data collection. These challenges have been documented in various microbial systems, from cheese communities [40] to human microbiome studies with dense temporal sampling [41, 42]. Furthermore, noise arises in experimental studies of gut microbiota reconstitution and synthetic community design [43, 44], as well as in investigations of microbiome dynamics at different timescales [45, 46]. But there is also intrinsic noise in the system, with sources including fluctuations in temperature, different local concentrations of species, or slight variations among individuals.

Currently, existing definitions of resilience are based on the theory of ordinary differential equations or difference equations, so all existing stability metrics to measure resilience are not functions of noise. In this paper, we have built a general framework to allow stochasticity to be deeply involved in the dynamics of microbial communities. Specifically, we have proposed an SgLV model to describe the temporal dynamics of a microbial community with white noise added to each interaction. We then developed a general theoretical framework for how to calculate four resilience measures based on the SgLV model. Each measure captures a different facet of stability and, together, they give a picture of resilience.

All four resilience measures are inherently functions of noise, since they are defined in terms of the ODE system established by the variance matrix of the linearized SDE system. Therefore, all four measures can be manipulated by the magnitude of the noise. One surprising result to emerge from our framework is that the return to the original stationary state is actually faster in the system with high noise than with low noise. This is primarily driven by the fact that the stationary distribution has higher variance in the system with high noise, so that there is less recovery required following a perturbation to return to the stationary distribution.

The next step of this work is the possibility of applying this general framework to time-series data from real-world microbiomes. This application involves a fundamental question regarding the identifiability and estimability of the parameters within our proposed stochastic generalized Lotka–Volterra (SgLV) model. Parameter identifiability remains a significant challenge, particularly for complex hierarchical models involving microbiome dynamics. While traditional approaches to identifiability, such as those described in Browning 2020 [47] and in Remien 2021 [48], focus primarily on deterministic systems, fewer methodologies explicitly address stochastic models. To address this critical gap, methodologies such as data cloning, as described by Lele et al., 2010 [49] may be particularly valuable. Although originally designed for generalized linear mixed models, data cloning could be adapted

to our SgLV framework to provide a practical graphical test for assessing parameter estimability. This approach involves generating multiple replicated datasets (clones) from microbiome time-series data and using Markov Chain Monte Carlo (MCMC) methods to sample from the resulting posterior distributions. By assessing whether the parameters have degenerate posterior distributions as the number of clones increases, we can determine whether each parameter—or specific functions of parameters, such as the resilience metrics proposed in our study—is estimable. Such a rigorous assessment of estimability would significantly enhance the reliability of the scientific conclusions drawn from applying our framework to real-world microbiome data.

An important direction of future research is the application of our proposed stochastic generalized Lotka–Volterra (SgLV) framework and resilience measures to practical scenarios in microbiome studies. For example, predicting the recovery dynamics of microbial communities after perturbations such as single-dose antibiotic treatments represents a significant area of potential application. Recent studies, including those by Hayashi et al., 2024 [24] and Ponciano et al., 2024 [25], have demonstrated that microbial community compositions can shift between alternative stable states under environmental perturbations, influenced by both deterministic selection and stochastic drift. Using our developed resilience metrics, researchers could quantify the microbiome’s rate and likelihood of returning to a “healthy” or pre-perturbation state, providing a valuable predictive tool for microbiome resilience. This approach could significantly aid experimental design, inform intervention strategies, and guide the interpretation of the microbiome’s recovery patterns following ecological disturbances.

Nevertheless, applying our SgLV framework and resilience metrics to real-world microbiome data requires careful consideration of several microbial-data-specific challenges. Microbiome datasets frequently suffer from the undersampling of rare taxa, which can lead to inaccurate abundance estimates and consequently affect the reliability of the calculated resilience measures. Additionally, relative abundance data, as commonly obtained through sequencing technologies, complicate the accurate estimation of interaction parameters because these data lack information on the absolute abundances. As emphasized by Ponciano et al., 2024 [25], addressing these challenges might involve combining sequencing-based relative abundances with absolute abundance quantification methods, such as quantitative polymerase chain reaction (PCR), to improve the estimation accuracy of interaction coefficients. Additionally, sparse temporal sampling can obscure rapid microbial dynamics, potentially reducing the reliability of resilience predictions. Future studies might employ robust parameter estimation methods such as data cloning (Lele et al., 2010 [49]) to assess parameter estimability rigorously and thus enhance confidence in resilience measures calculated from real-world microbiome time-series data.

Recently, Flamant et al., 2020 [50] introduced an innovative approach using neural network solution bundles to approximate the solutions of ordinary differential equations, which could be extended effectively to stochastic generalized Lotka–Volterra (SgLV) models for microbiome studies. In this approach, neural networks are trained to approximate solution bundles of differential equations over a broad parameter space, enabling efficient exploration of model dynamics and rapid fitting to empirical data. Specifically, once trained, these neural network solution bundles could quickly approximate the trajectories predicted by the SgLV model, significantly reducing the computational overhead compared with traditional numerical solvers. When applied to real-world microbiome time-series data, this methodology could effectively address common data challenges. For instance, by efficiently handling sparse temporal sampling, neural network approximations could infer microbiome dynamics even when the data resolution is limited. Furthermore, the flexibility of neural networks might better

accommodate biases arising from relative abundance data, potentially integrating with methods used to infer absolute abundances. Lastly, the approach could also help mitigate the challenges arising from undersampling rare taxa by efficiently exploring parameter spaces to identify robust solutions despite incomplete information on abundance. Thus, adopting neural network solution bundles presents a promising strategy for reliably applying the SgLV framework to complex microbiome datasets.

Use of AI tools declaration

All authors declare that they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

Research reported in this publication was supported by the National Institute of General Medical Sciences of the National Institutes of Health under Award Number P20GM104420.

Conflict of interest

All authors declare that there are no conflicts of interest regarding the publication of this manuscript.

References

1. G. Berg, D. Rybakova, D. Fischer, T. Cernava, M. C. C. Vergès, T. Charles, et al., Microbiome definition re-visited: Old concepts and new challenges, *Microbiome*, **8** (2020), 103. <https://doi.org/10.1186/s40168-020-00875-0>
2. I. Cho, M. Blaser, The human microbiome: At the interface of health and disease, *Nat. Rev. Genet.*, **13** (2012), 260–270. <https://doi.org/10.1038/nrg3182>
3. C. M. Cullen, K. K. Aneja, S. Beyhan, C. E. Cho, S. Woloszynek, M. Convertino, et al., Emerging priorities for microbiome research, *Front. Microbiol.*, **11** (2020), 136. <https://doi.org/10.3389/fmicb.2020.00136>
4. E. K. Costello, K. Stagaman, L. Dethlefsen, B. J. Bohannan, D. A. Relman, The application of ecological theory toward an understanding of the human microbiome, *Science*, **336** (2012), 1255–1262. <https://doi.org/10.1126/science.1224203>
5. D. A. Relman, The human microbiome: Ecosystem resilience and health, *Nutr. Rev.*, **70** (2012), S2–S9. <https://doi.org/10.1111/j.1753-4887.2012.00489.x>
6. V. Bucci, J. B. Xavier, Towards predictive models of the human gut microbiome, *J. Mol. Biol.*, **426** (2014), 3907–3916. <https://doi.org/10.1016/j.jmb.2014.03.017>
7. R. R. Stein, V. Bucci, N. C. Toussaint, C. G. Buffie, G. Räscher, E. G. Pamer, et al., Ecological modeling from time-series inference: insight into dynamics and stability of intestinal microbiota, *PLoS Comput. Biol.*, **9** (2013), e1003388. <https://doi.org/10.1371/journal.pcbi.1003388>
8. M. Grizotte-Lake, G. Zhong, K. Duncan, J. Kirkwood, N. Iyer, I. Smolenski, et al., Commensals suppress epithelial cell retinoic acid synthesis to regulate intestinal

- interleukin-22 activity and prevent microbial dysbiosis, *Immunity*, **49** (2018), 1103–1115. e6. <https://doi.org/10.1016/j.immuni.2018.11.018>
9. J. Ravel, P. Gajer, Z. Abdo, G. M. Schneider, S. S. K. Koenig, S. L. McCulle, et al., Vaginal microbiome of reproductive-age women, *Proc. Natl. Acad. Sci.*, **108** (2011), 4680–4687. <https://doi.org/10.1073/pnas.1002611107>
 10. W. G. Wade, Resilience of the oral microbiome, *Periodontology 2000*, **86** (2021), 113–122. <https://doi.org/10.1111/prd.12365>
 11. S. K. Dogra, J. Doré, S. Damak, Gut microbiota resilience: Definition, link to health and strategies for intervention, *Front. Microbiol.*, **11** (2020), 572921. <https://doi.org/10.3389/fmicb.2020.572921>
 12. M. Fassarella, E. E. Blaak, J. Penders, A. Nauta, H. Smidt, E. G. Zoetendal, Gut microbiome stability and resilience: elucidating the response to perturbations in order to modulate gut health, *Gut*, **70** (2021), 595–605. <https://doi.org/10.1136/gutjnl-2020-321747>
 13. G. Bacci, A. Mengoni, G. Emiliani, C. Chiellini, E. G. Cipriani, G. Bianconi, et al., Defining the resilience of the human salivary microbiota by a 520-day longitudinal study in a confined environment: The Mars500 mission, *Microbiome*, **9** (2021), 152. <https://doi.org/10.1186/s40168-021-01070-5>
 14. C. S. Holling, Resilience and stability of ecological systems, *Annual Review of Ecology and Systematics*, **4** (1973), 1–23. <https://doi.org/10.1146/annurev.es.04.110173.000245>
 15. S. L. Pimm, J. H. Lawton, Number of trophic levels in ecological communities, *Nature*, **268** (1977), 329–331. <https://doi.org/10.1038/268329a0>
 16. D. L. DeAngelis, Energy flow, nutrient cycling, and ecosystem resilience, *Ecology*, **61** (1980), 764–771. <https://doi.org/10.2307/1936746>
 17. K. L. Cottingham, S. R. Carpenter, Predictive indices of ecosystem resilience in models of north temperate lakes, *Ecology*, **75** (1994), 2127–2138. <https://doi.org/10.2307/1941616>
 18. M. G. Neubert, H. Caswell, Alternatives to resilience for measuring the response of ecological systems to perturbations, *Ecology*, **78** (1997), 653–665. [https://doi.org/10.1890/0012-9658\(1997\)078\[0653:ATRFMT\]2.0.CO;2](https://doi.org/10.1890/0012-9658(1997)078[0653:ATRFMT]2.0.CO;2)
 19. J-F. Arnoldi, M. Loreau, B. Haegeman, Resilience, reactivity and variability: A mathematical comparison of ecological stability measures, *J. Theor. Biol.*, **389** (2016), 47–59. <https://doi.org/10.1016/j.jtbi.2015.10.012>
 20. J-F. Arnoldi, M. Loreau, B. Haegeman, How ecosystems recover from pulse perturbations: A theory of short- to long-term responses, *J. Theor. Biol.*, **436** (2018), 79–92. <https://doi.org/10.1016/j.jtbi.2017.10.003>
 21. A. R. Ives, Measuring resilience in stochastic systems, *Ecol. Monogr.*, **65** (1995), 217–233. <https://doi.org/10.2307/2937138>
 22. A. R. Ives, B. Dennis, K. L. Cottingham, S. R. Carpenter, Estimating community stability and ecological interactions from time-series data, *Ecol. Monogr.*, **73** (2003), 301–330. [https://doi.org/10.1890/0012-9615\(2003\)073\[0301:ECSAEI\]2.0.CO;2](https://doi.org/10.1890/0012-9615(2003)073[0301:ECSAEI]2.0.CO;2)

23. E. Buckwar, C. Kelly, Asymptotic and transient mean-square properties of stochastic systems arising in ecology, fluid dynamics, and system control, *SIAM J. Appl. Math.*, **74** (2014), 411–433. <https://doi.org/10.1137/120893859>
24. I. Hayashi, H. Fujita, H. Toju, Deterministic and stochastic processes generating alternative states of microbiomes, *ISME Commun.*, **4** (2024), ycae007. <https://doi.org/10.1093/ismeco/ycae007>
25. J. M. Ponciano, J. P. Gómez, J. Ravel, L. J. Forney, Inferring stability and persistence in the vaginal microbiome: A stochastic model of ecological dynamics, preprint, bioRxiv 2024.03.02.581600 <https://doi.org/10.1101/2024.03.02.581600>
26. S. Zaoli, J. Grilli, A macroecological description of alternative stable states reproduces intra- and inter-host variability of gut microbiome, *Sci. Adv.*, **7** (2021), eabj2882. <https://doi.org/10.1126/sciadv.abj2882>
27. R. Wolff, W. Shoemaker, N. Garud, Ecological stability emerges at the level of strains in the human gut microbiome, *mBio*, **14** (2023), e02502-22. <https://doi.org/10.1128/mbio.02502-22>
28. S. J. Schreiber, M. Benaïm, K. A. S. Atchade, Persistence in fluctuating environments, *J. Math. Biol.*, **62** (2011), 655–683. <https://doi.org/10.1007/s00285-010-0349-5>
29. R. Rudnicki, Long-time behaviour of a stochastic prey–predator model, *Stoch. Process. Their Appl.*, **108** (2003), 93–107. [https://doi.org/10.1016/S0304-4149\(03\)00090-5](https://doi.org/10.1016/S0304-4149(03)00090-5)
30. D. H. Nguyen, G. Yin, Coexistence and exclusion of stochastic competitive Lotka–Volterra models, *J. Differ. Equat.*, **262** (2017), 1192–1225. <https://doi.org/10.1016/j.jde.2016.10.005>
31. M. Benaïm, Stochastic persistence, arXiv:1806.08450.
32. A. Hening, H. D. Nguyen, Coexistence and extinction for stochastic Kolmogorov systems, *Ann. Appl. Probab.*, **28** (2018), 1893–1942. <https://doi.org/10.1214/17-AAP1347>
33. A. Hening, H. D. Nguyen, Persistence in stochastic Lotka–Volterra food chains with intraspecific competition, *Bull. Math. Biol.*, **80** (2018), 2527–2560. <https://doi.org/10.1007/s11538-018-0468-5>
34. A. Hening, H. D. Nguyen, Stochastic Lotka–Volterra food chains, *J. Math. Biol.*, **77** (2018), 135–163. <https://doi.org/10.1007/s00285-017-1192-8>
35. S. Marino, N. T. Baxter, G. B. Huffnagle, J. F. Petrosino, P. D. Schloss, Mathematical modeling of primary succession of murine intestinal microbiota, *Proc. Natl. Acad. Sci.*, **111** (2014), 439–444. <https://doi.org/10.1073/pnas.1311322111>
36. P. Dam, L. L. Fonseca, K. T. Konstantinidis, E. O. Voit, Dynamic models of the complex microbial metapopulation of lake mendota, *npj Syst. Biol. Appl.*, **2** (2016), 16007. <https://doi.org/10.1038/npjsba.2016.7>
37. P. Dam, L. M. Rodriguez-R, C. Luo, J. Hatt, D. Tsementzi, K. T. Konstantinidis, et al., Model-based comparisons of the abundance dynamics of bacterial communities in two lakes, *Sci. Rep.*, **10** (2020), 2423. <https://doi.org/10.1038/s41598-020-58769-y>
38. V. Bucci, B. Tzen, N. Li, M. Simmons, T. Tanoue, E. Bogart, et al., MDSINE: Microbial dynamical systems inference engine for microbiome time-series analyses, *Genome. Biol.*, **17** (2016), 121. <https://doi.org/10.1186/s13059-016-0980-6>

39. B. J. Ridenhour, S. Brooker, J. Williams, J. T. V. Leuven, A. W. Miller, M. D. Dearing, et al., Modeling time-series data from microbial communities, *ISME J.*, **11** (2017), 2526–2537. <https://doi.org/10.1038/ismej.2017.107>
40. J. Mounier, C. Monnet, T. Vallaëys, R. Arditi, A. S. Sarthou, A. Hélias, et al., Microbial interactions within a cheese microbial community, *Appl. Environ. Microbiol.*, **74** (2008), 172–181. <https://doi.org/10.1128/AEM.01338-07>
41. J. G. Caporaso, C. L. Lauber, E. K. Costello, D. Berg-Lyons, A. Gonzalez, J. Stombaugh, et al., Moving pictures of the human microbiome, *Genome. Biol.*, **12** (2011), R50. <https://doi.org/10.1186/gb-2011-12-5-r50>
42. L. A. David, A. C. Materna, J. Friedman, M. I. Campos-Baptista, M. C. Blackburn, A. Perrotta, et al., Host lifestyle affects human microbiota on daily timescales, *Genome. Biol.*, **15** (2014), R89. <https://doi.org/10.1186/gb-2014-15-7-r89>
43. C. Buffie, V. Bucci, R. Stein, P. T. McKenney, L. Ling, A. Gobourne, et al., Precision microbiome reconstitution restores bile acid mediated resistance to *Clostridium difficile*, *Nature*, **517** (2014), 205–208. <https://doi.org/10.1038/nature13828>
44. O. S. Venturelli, A. C. Carr, G. Fisher, R. H. Hsu, R. Lau, B. P. Bowen, et al., Deciphering microbial interactions in synthetic human gut microbiome communities, *Mol. Syst. Biol.*, **14** (2018), e8157. <https://doi.org/10.15252/msb.20178157>
45. B. H. Schlomann, R. Parthasarathy, Timescales of gut microbiome dynamics, *Current Opin. Microbiol.*, **50** (2019), 56–63. <https://doi.org/10.1016/j.mib.2019.09.011>
46. T. J. Wiles, B. H. Schlomann, E. S. Wall, R. Betancourt, R. Parthasarathy, K. Guillemin, Swimming motility of a gut bacterial symbiont promotes resistance to intestinal expulsion and enhances inflammation, *PLoS Biol.*, **18** (2020), e3000661. <https://doi.org/10.1371/journal.pbio.3000661>
47. A. P. Browning, D. J. Warne, K. Burrage, R. E. Baker, M. J. Simpson, Identifiability analysis for stochastic differential equation models in systems biology, *J. R. Soc. Interface*, **17** (2020), 20200652. <https://doi.org/10.1098/rsif.2020.0652>
48. C. H. Remien, M. J. Eckwright, B. J. Ridenhour, Structural identifiability of the generalized Lotka-Volterra model for microbiome studies, *R. Soc. Open Sci.*, **8** (2021), 201378. <https://doi.org/10.1098/rsos.201378>
49. S. R. Lele, K. Nadeem, B. Schmuland, Estimability and likelihood inference for generalized linear mixed models using data cloning, *J. Am. Statist. Assoc.*, **105** (2010), 1617–1625. <https://doi.org/10.1198/jasa.2010.tm09757>
50. C. Flamant, P. Protopapas, D. Sondak, Solving differential equations using neural network solution bundles, *arXiv:2006.14372*.



AIMS Press

©2025 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0>)