



# Impact of a block structure on the Lotka-Volterra model

Maxime Clenet, François Massol, Jamal Najim

## ► To cite this version:

Maxime Clenet, François Massol, Jamal Najim. Impact of a block structure on the Lotka-Volterra model. 2023. [hal-04298668](#)

**HAL Id: hal-04298668**

**<https://hal.science/hal-04298668v1>**

Preprint submitted on 21 Nov 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

# Impact of a block structure on the Lotka-Volterra model

Maxime Clenet<sup>(1,2)</sup>, François Massol<sup>(3)</sup>, Jamal Najim<sup>(1)\*</sup>

(1) CNRS and Université Gustave Eiffel, France

(2) Université de Sherbrooke, Canada

(3) CNRS, Université de Lille, INSERM, CHU, Institut Pasteur Lille

November 21, 2023

## Abstract

The Lotka-Volterra (LV) model is a simple, robust, and versatile model used to describe large interacting systems such as food webs or microbiomes. The model consists of  $n$  coupled differential equations linking the abundances of  $n$  different species. We consider a large random interaction matrix with independent entries and a block variance profile. The  $i$ th diagonal block represents the intra-community interaction in community  $i$ , while the off-diagonal blocks represent the inter-community interactions. The variance remains constant within each block, but may vary across blocks.

We investigate the important case of two communities of interacting species, study how interactions affect their respective equilibrium. We also describe equilibrium with feasibility (i.e., whether there exists an equilibrium with all species at non-zero abundances) and the existence of an attrition phenomenon (some species may vanish) within each community.

Information about the general case of  $b$  communities ( $b > 2$ ) is provided in the appendix.

**Keywords:** Lotka-Volterra model, Block structure, Linear Complementarity Problems, Large Random Matrices, Stability of food webs.

---

\*Supported by CNRS Project 80 Prime — KARATE.

# 1 Introduction

Understanding large ecosystems and the underlying mechanisms that support high species diversity is a major challenge in theoretical ecology. Motivated by the seminal work of May [May72], the introduction of random matrices has been a key mathematical step in modeling high-dimensional ecosystems. These tools have expanded our ability to understand the nature of interactions and how food webs can resist to small perturbations (stability) [AT12, TPA14].

Differential equations are frequently used in ecology to describe a system of interacting species. One of the most common models is the Lotka-Volterra (LV) model [Lot25, Vol26], which has been the subject of research in both ecology [Wan78, Jan87, LB92] and mathematics [GJ77, Goh77, Tay88, HS98, Tak96]. Certain properties of this model, such as its stability [GGRA18], have raised much interest. The conditions under which all species survive, referred to as feasibility, have also motivated many works [BN21, GAS<sup>+</sup>17, Sto18].

In nature, ecological networks are rather structured, and many studies have investigated the network structures that contribute to the stability of a given community [TF10, AGB<sup>+</sup>15]. One common network structure is food web compartmentalization, also known as modularity. The underlying concept is that the network is structured in the form of groups of nodes that interact more strongly within their group and more weakly between groups. A mathematical formulation of modularity was defined by Newman [New06]. Subsequently, modularity has been of great importance in ecology [GSSP<sup>+</sup>10], in complex networks [VML04], and in community detection (for a complete review, see Fortunato [For10]).

May had already mentioned that a multi-community structure should improve stability [May72], a hypothesis later investigated by Pimm [Pim79]. In the same framework as May, Grilli *et al.* [GRA16] studied the effect of modularity on the stability of the Jacobian of a system, the so-called “community matrix”. However, studies show that modularity improves the persistence ( $:=$  non-extinction of species, generally related to their resistance to external perturbations) of species in the dynamical system [SB11].

In this article, we study the Lotka-Volterra model where we consider a block structure network representing the inter- and intra-community interactions. Of particular interest are the interactions between the communities that affect their respective equilibrium and stability.

Each block is identified by its interaction strength, which is the standard deviation of the random part of the interactions. The idea that interaction strength plays a key role in the stability of ecosystems was introduced by May [May72]. For the sake of mathematical simplicity, we limit our model to two communities, although we can extend the model to more complex food webs and multi-community frameworks.

We study the existence and stability of an equilibrium, together with its properties. When an attrition phenomenon occurs (some species may vanish), we describe the proportion of surviving species and their distribution. We also provide conditions for which the equilibrium is feasible (i.e., whether there exists an equilibrium with all species at non-zero abundances).

**Model and assumptions.** The LV model is a standard model in ecology to study the dynamics of a community of species over time. It is defined by a system of  $n$  differential equations

$$\frac{dx_k}{dt}(t) = x_k(t) \left( r_k - x_k(t) + \sum_{\ell \in [n]} B_{k\ell} x_\ell(t) \right), \quad k \in [n] = \{1, \dots, n\}. \quad (1)$$

The abundance of species  $k$  at time  $t$  is represented by  $x_k(t)$  with  $\mathbf{x} = (x_1, \dots, x_n)$  the vector of abundances. Parameter  $r_k$  corresponds to the growth rate of species  $k$ . The coefficient  $B_{k\ell}$  represents the impact of species  $\ell$  on species  $k$ . The  $n \times n$  matrix  $B = (B_{k\ell})$ , which represents the interaction network, is decomposed into a block structure. This structure differentiates various groups of species in the form of communities that interact with each other. On the one hand, the diagonal blocks of  $B$  correspond to interactions within each community, each with its own interaction strength. On the other hand, the off-diagonal blocks correspond to the impact of the communities on each other. Analytically and within the framework of two communities, the matrix  $B = (B_{k\ell})_{n,n}$  is defined by blocks using random matrices  $A_{ij}$  and interaction strengths  $s_{ij}$ :

$$B = \frac{1}{\sqrt{n}} \begin{pmatrix} s_{11}A_{11} & s_{12}A_{12} \\ s_{21}A_{21} & s_{22}A_{22} \end{pmatrix}, \quad (2)$$

where  $\mathcal{I}_1 = [n_1]$  (resp  $\mathcal{I}_2 = \{n_1 + 1, \dots, n\}$ ), the subset of  $[n]$  of size  $|\mathcal{I}_1| = n_1$  (resp  $|\mathcal{I}_2| = n_2$  - here and below  $n = n_1 + n_2$ ) matching the index of species belonging to community 1 (resp community 2) and  $\beta = (\beta_1, \beta_2)$  with  $\beta_1 = n_1/n$ ,  $\beta_2 = 1 - \beta_1 = n_2/n$ . The random matrix  $A_{ij}$  is non-Hermitian of size  $|\mathcal{I}_i| \times |\mathcal{I}_j|$  with standard Gaussian entries i.e.  $\mathcal{N}(0, 1)$ . The Gaussianity assumption clarifies the explanations, but can be relaxed under certain circumstances (see the corresponding sections for details).

Notice a normalization parameter  $1/\sqrt{n}$  in the matrix  $B$ . This enables the interaction matrix  $B$  to have a macroscopic effect on system (1):

$$\mathbb{E} \left( \sum_{\ell \in [n]} B_{k\ell} x_\ell(t) \right) \sim O(1) ; \quad \text{Var} \left( \sum_{\ell \in [n]} B_{k\ell} x_\ell(t) \right) \sim O(1).$$

From an ecological perspective, an increase in the number of species may not necessarily lead to a corresponding increase in the overall strength of interactions between one species and all others.

The relative strength of interactions within and between blocks is controlled by the four  $s_{ij}$  coefficients, which can be grouped together in a matrix  $\mathbf{s}$ :

$$\mathbf{s} = \begin{pmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{pmatrix}.$$

The diagonal terms ( $s_{11}, s_{22}$ ) represent the interaction strength in each community. The off-diagonal term  $s_{12}$  (resp.  $s_{21}$ ) represents the interaction strength of the impact of community 2 on community 1 (resp. community 1 on community 2). The lower the value of  $s_{ij}$ , the lower the rates of interaction between species. Note that in the case of a unique community,  $s$  is the interaction strength coefficient, i.e. the standard deviation of the interspecific coefficients of the LV model.

**Remark 1.** For the sake of simplicity, the results are presented in the case of two interacting communities but can be extended to the case of  $b$  communities, see Appendix E.

There are two scenarios of interest: Let us consider two separate groups of species that follow the dynamics described in model (1). The matrices  $A_{ij}$  are each sampled once. In the first scenario, we consider a very weak interaction between the two communities

$$\mathbf{s} = \begin{pmatrix} 1/2 & \varepsilon \\ \varepsilon & 1/2 \end{pmatrix}, \varepsilon > 0,$$

and the intra-communities interactions are small enough (Fig 1a). We observe that both communities dynamics converge to a feasible equilibrium in the sense that all species survive (Fig. 1b). In the second scenario, we increase the interactions between the communities, i.e. the standard deviation matrix is defined by

$$\mathbf{s} = \begin{pmatrix} 1/2 & 1 \\ 1 & 1/2 \end{pmatrix}.$$

It is no longer possible for both communities to maintain the feasibility of all species. Some species are likely to disappear (Fig. 2).

**Properties of the dynamical system.** We are interested in the effect of a block structure on the food web, limit our study to the 2-blocks case (2) and focus on the model with constant growth rate<sup>1</sup>  $r_k = 1$ :

$$\frac{dx_k}{dt} = x_k (1 - x_k + (B\mathbf{x})_k), \quad k \in [n]. \quad (3)$$

Of major interest is the existence and uniqueness of an equilibrium  $\mathbf{x}^* = (x_k^*)_{k \in [n]}$ . The LV system is an autonomous differential system. If the initial conditions are positive i.e.  $\mathbf{x}(0) > 0$  (componentwise), it implies

<sup>1</sup>The simplifying assumption  $r_k = 1$  allows tractable computations and could be extended to  $r_k = c$  with  $c > 0$ . However, if the growth rate is different for each species, the mathematical development and result may be strongly affected and will be discussed in each section.

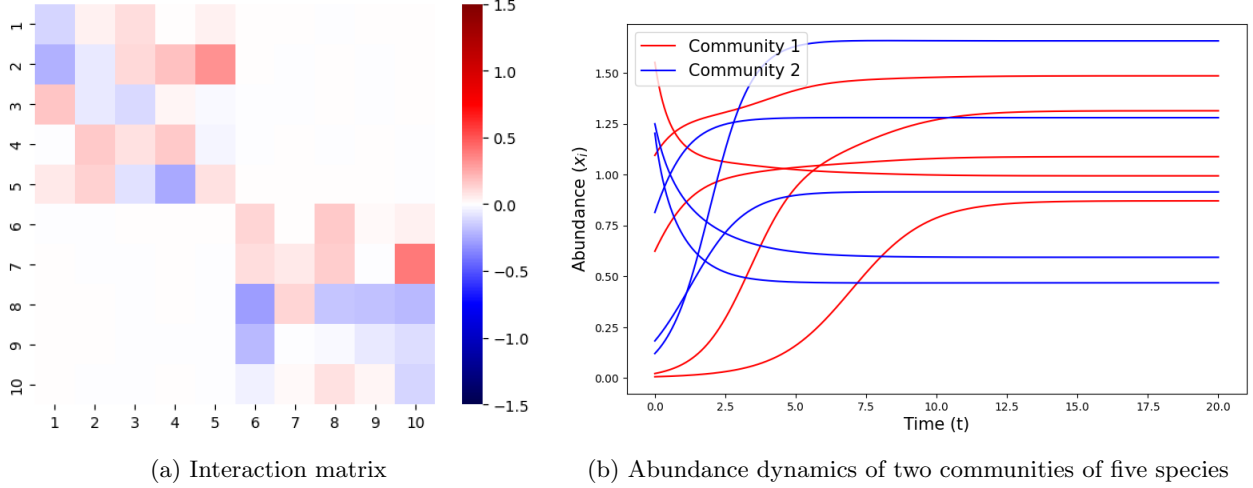


Figure 1: Dynamics of the model (1) with 2 distinct communities of 5 species each and interaction matrix given by (2). The two communities converge to their feasible equilibrium point and do not interact. In Fig. (a), a heat map illustrates the interaction matrix (2). Figure (b) shows the dynamics where the species of each community reach a feasible equilibrium.

$\mathbf{x}(t) > 0$  for every  $t > 0$ . However, some of the components  $x_k(t)$  may converge to zero if the equilibrium  $\mathbf{x}^*$  has components equal to zero. An equilibrium to the LV system should hence satisfy the following set of constraints:

$$\begin{cases} x_k^* (1 - x_k^* + (B\mathbf{x}^*)_k) = 0, & \forall k \in [n], \\ x_k^* \geq 0. \end{cases} \quad (4)$$

Two substantially different situations arise, that we will study hereafter.

First, if  $\mathbf{x}^*$  has vanishing components, the equilibrium equations are cast into a nonlinear optimization problem, which has been studied by Clenet *et al.* [CMN23] in the case of a single community.

If the equilibrium is feasible, that is  $\mathbf{x}^* > 0$ , then the equilibrium set of equations becomes a linear equation:

$$\mathbf{x}^* = \mathbf{1} + B\mathbf{x}^*. \quad (5)$$

In the context of a single community, the existence of a positive solution has been studied by Bizeul and Najim [BN21] and extended for more complex food webs in [AN22, CEFN22, LCP23].

A further consideration which will be addressed is whether the equilibrium  $\mathbf{x}^*$  is asymptotically globally stable, i.e. if for every initial vector  $\mathbf{x}_0 > 0$  the solution of (3), which starts at  $\mathbf{x}(0) = \mathbf{x}_0$ , satisfies

$$\mathbf{x}(t) \xrightarrow[t \rightarrow \infty]{} \mathbf{x}^*.$$

In the sequel, the term “stability” will refer to “asymptotic stability”.

**Outline of the article.** In Section 2, we describe sufficient conditions for the existence and uniqueness of a stable equilibrium in the model (3), see Theorem 3. Section 3 is devoted to the study of the properties of the species that survive in each of the communities. Finally, in Section 4 we provide conditions under which the equilibrium is feasible, see Theorem 4.

## 2 Existence of a unique equilibrium

In Figures 1 and 2, we notice that for different interaction coefficients  $\mathbf{s}$ , the system converges to an equilibrium (with or without vanishing species). Theorem 3 below will provide the adequate theoretical framework.

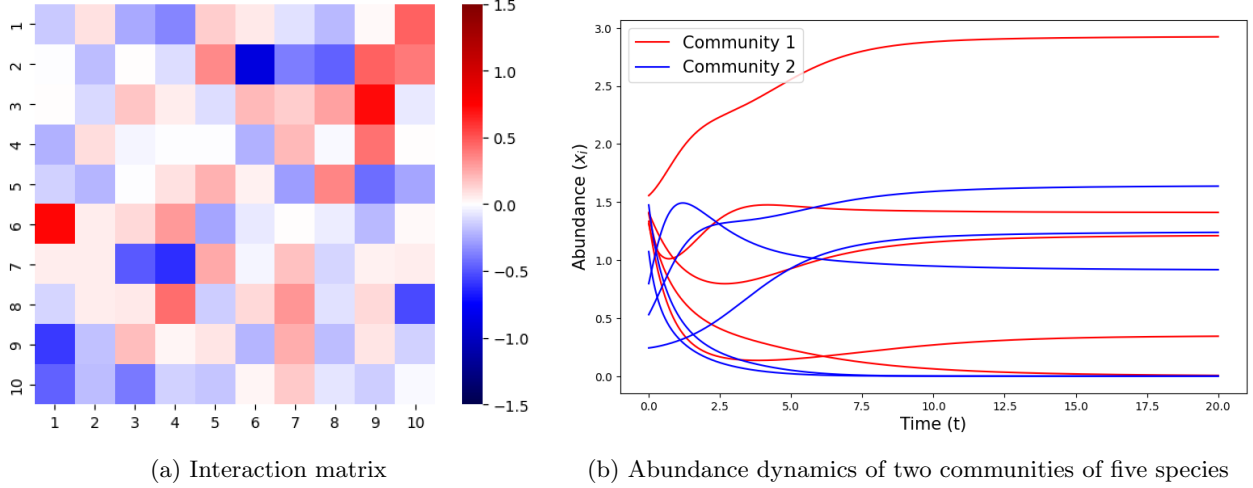


Figure 2: Dynamics of Model (1) with 2 distinct communities of 5 species each with interaction matrix given by (2). In Fig. (a), representation of the interaction matrix (2) when the interactions between the communities are strong. Fig. (b) shows the dynamics where the species of each community reach an equilibrium. Notice that there are vanishing species in each community.

## 2.1 Theoretical background

**Non-invadability condition.** The research of equilibrium points of (3) is equivalent to the identification of solutions of system (4). However, the number of potential solutions can be extremely large. In order for the equilibrium  $\mathbf{x}^*$  to be stable, there exists a necessary condition, known in ecology as the non-invadability condition [LM96], namely that

$$1 - x_k^* + (B\mathbf{x}^*)_k \leq 0. \quad (6)$$

In model (3), the non-invadability condition for a given species  $k \in [n]$  is equivalent to

$$\left( \frac{1}{x_k} \frac{dx_k}{dt} \right)_{x_k \rightarrow 0^+} \leq 0. \quad (7)$$

Condition (7) describes the fact that if we add a species to the system at a very low abundance, it will not be able to invade the system. As a consequence, the number of possible solutions should solve the following set of constraints:

$$\begin{cases} x_k^* (1 - x_k^* + (B\mathbf{x}^*)_k) &= 0 & \text{for } k \in [n], \\ 1 - x_k^* + (B\mathbf{x}^*)_k &\leq 0 & \text{for } k \in [n], \\ \mathbf{x}^* &\geq 0 & \text{componentwise.} \end{cases} \quad (8)$$

This casts the search of a nonnegative equilibrium problem into the class of linear complementarity problems (LCP). For a reminder of the definition of an LCP problem, see for instance [CMN23]. In the following, we recall the main Theorem for proving the existence and uniqueness of a single equilibrium.

**The equilibrium  $\mathbf{x}^*$  and its stability.** Let  $X^\top$  be the transpose of the matrix  $X$ .

**Definition 1** (Lyapunov diagonal stability). A matrix  $M$  is called Lyapunov diagonally stable, denoted by  $M \in \mathcal{D}$ , if and only if there exists a diagonal matrix  $D$  with positive diagonal elements such that  $DM + M^\top D$  is negative definite, i.e. all eigenvalues are negative.

This class of matrix was already mentioned in Volterra's historical paper [Vol31] and in Logofet's book [Log93, Chap. 4], in relation with the stability of LV models.

**Proposition 1** (Takeuchi *et al.* [TAT78]). *If  $M \in \mathcal{D}$  then  $-M$  is a  $P$ -matrix i.e. all its principal minors (sub-determinants) are strictly positive:*

$$\det(M_{\mathcal{I}}) > 0, \quad \forall \mathcal{I} \subset [n], \quad M_{\mathcal{I}} = (M_{k\ell})_{k,\ell \in \mathcal{I}}.$$

Recall System (1) with different growth rates for each species and consider matrix  $B$  is arbitrary,

$$\frac{dy_k}{dt} = y_k(r_k + ((-I + B)\mathbf{y})_k), \quad k \in [n]. \quad (9)$$

The LCP associated with (9) is as follows

$$\begin{cases} y_k^*(r_k - y_k^* + (B\mathbf{y}^*)_k) = 0 & \text{for } k \in [n], \\ r_k - y_k^* + (B\mathbf{y}^*)_k \leq 0 & \text{for } k \in [n], \\ \mathbf{y}^* \geq 0 & \text{componentwise.} \end{cases} \quad (10)$$

Takeuchi and Adachi [Tak96, Th. 3.2.1] establish a criterion for the existence of a unique globally stable equilibrium  $\mathbf{y}^*$  of (9).

**Theorem 2** (Takeuchi and Adachi [TA80]). *If  $-I + B \in \mathcal{D}$ , then System (10) admits a unique solution. In particular, for every  $\mathbf{r} \in \mathbb{R}^n$ , there is a unique equilibrium  $\mathbf{y}^*$  to (9), which is globally stable in the sense that for every  $\mathbf{y}_0 > 0$ , the solution to (9) which starts at  $\mathbf{y}(0) = \mathbf{y}_0$  satisfies:  $\mathbf{y}(t) \xrightarrow[t \rightarrow \infty]{} \mathbf{y}^*$ .*

## 2.2 Sufficient condition in the block model (3)

We recall the definition of the Stieltjes transform and some of its properties in Appendix A (for more details, see [BS10]). For a wide range of parameters  $(\beta, \mathbf{s})$ , we aim to ensure the existence of a globally stable equilibrium  $\mathbf{x}^*$  of (3) associated to LCP (8). Denote by  $\|\mathbf{x}\|_\infty$  the sup norm of a vector and by  $\|X\|_\infty$  its induced operator norm, i.e.

$$\|\mathbf{x}\|_\infty = \max_{k \in [n]} |x_k| \quad \text{and} \quad \|X\|_\infty = \max_{k \in [n]} \sum_{\ell=1}^n |X_{k\ell}|.$$

Let  $X, Y$  be matrices of the same size, then  $X \circ Y$  is their Hadamard product i.e.  $(X \circ Y)_{ij} = X_{ij}Y_{ij}$  and consider

$$\mathbf{s} \circ \mathbf{s} = \begin{pmatrix} s_{11}^2 & s_{12}^2 \\ s_{21}^2 & s_{22}^2 \end{pmatrix}.$$

**Theorem 3.** *Assume that*

$$\|\text{diag}(\beta) ((\mathbf{s} \circ \mathbf{s}) + (\mathbf{s} \circ \mathbf{s})^\top)\|_\infty < 1,$$

*then a.s. matrix  $(I - B) + (I - B)^\top$  is eventually positive definite: with probability one, there exists  $N$  depending on matrix  $B$ 's realization such that for  $n \geq N$ ,  $(I - B) + (I - B)^\top$  is positive definite. In particular,  $I - B \in \mathcal{D}$ . There exists a unique vector solution to the LCP (8). This vector  $\mathbf{x}^*$  is the unique (random) globally stable equilibrium of (3).*

**Remark 2.** Theorem 3 can be extended in two directions. The Gaussianity assumption can be relaxed to any reasonable distribution with finite second moment, and growth rates different from one i.e.  $r_k \neq 1$  can be considered (see for instance [BN21, Sections 4.2 and 4.3] in the context of a single community).

*Sketch of proof.* From Theorem 2, we need to verify the Lyapunov diagonally stable condition of the matrix  $(-I + B) \in \mathcal{D}$  by analyzing its largest eigenvalue

$$(-I + B) + (-I + B)^\top = -2I + \frac{1}{\sqrt{n}} \begin{pmatrix} s_{11}(A_{11} + A_{11}^\top) & s_{12}A_{12} + s_{21}A_{21}^\top \\ s_{21}A_{21} + s_{12}A_{12}^\top & s_{22}(A_{22} + A_{22}^\top) \end{pmatrix}.$$

Denote by  $H$  the symmetric matrix

$$H = \frac{1}{\sqrt{n}} \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} = \frac{1}{\sqrt{n}} \begin{pmatrix} s_{11}(A_{11} + A_{11}^\top) & s_{12}A_{12} + s_{21}A_{21}^\top \\ s_{21}A_{21} + s_{12}A_{12}^\top & s_{22}(A_{22} + A_{22}^\top) \end{pmatrix},$$

where  $H_{ij}$  is a matrix of size  $|\mathcal{I}_i| \times |\mathcal{I}_j|$  and each off-diagonal entries follow a Gaussian distribution  $\mathcal{N}(0, s_{ij}^2 + s_{ji}^2)$  for all  $i, j \in \{1, 2\}$ .

A matrix is negative definite if and only if all its eigenvalues are negative. Note here that  $-2I + H$  is negative definite if and only if the upper eigenvalue of  $H$  is less than 2. The goal of the proof is to give a condition on the parameter  $\mathbf{s}$  such that

$$\lambda_{\max}(H) < 2.$$

The matrix  $H$  has a variance profile, such a model has been studied in great details by Erdős *et al.* and is linked to the theory of the Quadratic Vector Equation (QVE, see [AEK17, AEK19] for more technical information). Given  $\mathbf{m}(z) = (m_1(z), \dots, m_n(z))$ , the QVE associated to the matrix  $H$  is decomposed as

$$\begin{aligned} k \in \mathcal{I}_1 \quad , \quad -\frac{1}{m_k(z)} &= z + \sum_{\ell \in \mathcal{I}_1} \frac{2s_{11}^2}{n} m_\ell(z) + \sum_{\ell \in \mathcal{I}_2} \frac{1}{n} (s_{12}^2 + s_{21}^2) m_\ell(z), \\ k \in \mathcal{I}_2 \quad , \quad -\frac{1}{m_k(z)} &= z + \sum_{\ell \in \mathcal{I}_1} \frac{1}{n} (s_{12}^2 + s_{21}^2) m_\ell(z) + \sum_{\ell \in \mathcal{I}_2} \frac{2s_{22}^2}{n} m_\ell(z). \end{aligned}$$

Denote by  $1/\mathbf{m}(z) = (1/m_1(z), \dots, 1/m_n(z))^\top$ ,  $\mathbf{1}_{\mathcal{I}_i}$  a vector whose entries are 1's of size  $|\mathcal{I}_i|$  and

$$S = \frac{1}{n} \begin{pmatrix} 2s_{11}^2 \mathbf{1}_{\mathcal{I}_1} \mathbf{1}_{\mathcal{I}_1}^\top & (s_{12}^2 + s_{21}^2) \mathbf{1}_{\mathcal{I}_1} \mathbf{1}_{\mathcal{I}_2}^\top \\ (s_{12}^2 + s_{21}^2) \mathbf{1}_{\mathcal{I}_2} \mathbf{1}_{\mathcal{I}_1}^\top & 2s_{22}^2 \mathbf{1}_{\mathcal{I}_2} \mathbf{1}_{\mathcal{I}_2}^\top \end{pmatrix},$$

the QVE can be written in the standard form

$$-\frac{1}{\mathbf{m}(z)} = z + S\mathbf{m}(z). \tag{11}$$

Following Theorem 2.1 in Ajanki *et al.* [AEK19],  $\forall z \in \mathbb{C}^+$ , Equation (11) has a unique solution  $\mathbf{m} = \mathbf{m}(z)$  where  $n^{-1} \sum m_i$  is the Stieltjes transform of a probability measure and the support of the associated measure is included in  $[-\Sigma, \Sigma]$ , where  $\Sigma = 2\|S\|_\infty^{1/2}$ .

This information gives an asymptotic bound on the support of the matrix  $H$  associated with (11), i.e. asymptotically  $\forall \varepsilon > 0$  there exists  $N$  depending on matrix  $B$ 's realization, such that for  $n \geq N$

$$\lambda_{\max}(H) \leq 2\|S\|_\infty^{1/2} + \varepsilon.$$

Recall that  $-2I + H$  is negative definite iff  $\lambda_{\max}(H) < 2$ . This condition is fulfilled if

$$2\|S\|_\infty^{1/2} < 2,$$

or equivalently

$$\|S\|_\infty < 1.$$

Note that this condition is sufficient but not necessary. Given the particular shape of the matrix  $S$ , computing its norm is equivalent to computing the norm of a matrix of size  $2 \times 2$

$$\|S\|_\infty = \|\text{diag}(\beta) ((\mathbf{s} \circ \mathbf{s}) + (\mathbf{s} \circ \mathbf{s})^\top)\|_\infty = \left\| \begin{pmatrix} \beta_1 & 0 \\ 0 & \beta_2 \end{pmatrix} \begin{pmatrix} 2s_{11}^2 & s_{12}^2 + s_{21}^2 \\ s_{12}^2 + s_{21}^2 & 2s_{22}^2 \end{pmatrix} \right\|_\infty,$$

which completes the proof. We can then rely on Theorem 2 to conclude.  $\square$

**Remark 3.**



1. In the context of a unique community, suppose that  $\mathbf{s} = s\mathbf{1}\mathbf{1}^\top$ , then the previous condition takes the simpler form  $s < 1/\sqrt{2}$ , a condition already mentioned in [CMN23]. Indeed, starting from the condition of Theorem 3, the condition on the matrix is

$$\left\| \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}^{1/2} \begin{pmatrix} 2s^2 & 2s^2 \\ 2s^2 & 2s^2 \end{pmatrix} \begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}^{1/2} \right\|_\infty = \left\| \begin{pmatrix} 1s^2 & 1s^2 \\ 1s^2 & 1s^2 \end{pmatrix} \right\|_\infty = 2s^2,$$

The same sufficient condition is obtained  $2s^2 < 1 \Leftrightarrow s < 1/\sqrt{2}$ .

2. The condition given in Theorem 3 is sufficient to guarantee a stable unique solution to LCP (8) but not necessary. Even in the single community case, finding optimal thresholds remains an open question (see [ABC<sup>+</sup>22, Table 1]).

### 3 Surviving species

In Section 2, we have given conditions on matrix  $\mathbf{s}$  and on  $\boldsymbol{\beta} = (\beta_1, \beta_2)$  for the existence of a globally stable equilibrium  $\mathbf{x}^*$  to (3) under the non-invadability condition. The equilibrium vector  $\mathbf{x}^*$  is random and depends on the realization of matrix  $B$ . Moreover since  $\mathbf{s}$  has fixed components and does not depend on  $n$ , the equilibrium  $\mathbf{x}^*$  will feature vanishing components (see the original argument for a unique community in [DVR<sup>+</sup>18] and the discussion in [BN21]). In an ecological context, we differentiate two kind of components in vector  $\mathbf{x}^*$ , the non-vanishing components  $x_k^* > 0$  corresponding to surviving species and the vanishing ones  $x_k^* = 0$  corresponding to the species going to extinction:

$$x_k(t) \xrightarrow[t \rightarrow \infty]{} 0.$$

Hereafter, we describe statistical properties of  $\mathbf{x}^*$ : the proportion of surviving species in each community, the distribution of the corresponding abundances, which turns out to be a truncated Gaussian, etc.

#### 3.1 Heuristics for the properties of surviving species

Starting from the model (3), the set of surviving species in community  $i \in \{1, 2\}$  is defined as

$$\begin{aligned} \mathcal{S}_1 &= \{k \in \mathcal{I}_1, x_k^* > 0\}; & \mathcal{I}_1 &= [1, \beta_1 n], \\ \mathcal{S}_2 &= \{k \in \mathcal{I}_2, x_k^* > 0\}; & \mathcal{I}_2 &= [\beta_1 n + 1, n]. \end{aligned}$$

Given the random equilibrium  $\mathbf{x}^*$ , we introduce the following quantities for each community  $i \in \{1, 2\}$

$$\hat{p}_i = \frac{|\mathcal{S}_i|}{|\mathcal{I}_i|}, \quad \hat{m}_i = \frac{1}{|\mathcal{S}_i|} \sum_{k \in \mathcal{S}_i} x_k^*, \quad \hat{\sigma}_i^2 = \frac{1}{|\mathcal{S}_i|} \sum_{k \in \mathcal{S}_i} (x_k^*)^2.$$

Quantity  $\hat{p}_i$  represents the proportion of surviving species in community  $i$ ,  $\hat{m}_i$  the empirical mean of the abundances of the surviving species in community  $i$  and  $\hat{\sigma}_i^2$ , the empirical mean square of the surviving species in community  $i$ .

Denote by  $Z \sim \mathcal{N}(0, 1)$  a standard Gaussian random variable and by  $\Phi$  the cumulative Gaussian distribution function:

$$\Phi(x) = \int_{-\infty}^x \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}} du.$$

**Heuristics 1.** Let  $\mathbf{s}$  be the  $2 \times 2$  matrix of interaction strengths and assume that the condition of Theorem 3 holds, then the following system of four equations and four unknowns  $(p_1, p_2, \sigma_1, \sigma_2)$

$$p_1 = 1 - \Phi(\delta_1), \tag{12}$$

$$p_2 = 1 - \Phi(\delta_2), \tag{13}$$

$$(\sigma_1)^2 = 1 + 2\Delta_1 \mathbb{E}(Z|Z > \delta_1) + \Delta_1^2 \mathbb{E}(Z^2|Z > \delta_1), \tag{14}$$

$$(\sigma_2)^2 = 1 + 2\Delta_2 \mathbb{E}(Z|Z > \delta_2) + \Delta_2^2 \mathbb{E}(Z^2|Z > \delta_2), \tag{15}$$

where

$$\Delta_i = \sqrt{p_1(\sigma_1)^2 \beta_1 s_{i1}^2 + p_2(\sigma_2)^2 \beta_2 s_{i2}^2} \quad \text{and} \quad \delta_i = \frac{-1}{\Delta_i}, \quad (16)$$

admits a unique solution  $(p_1^*, p_2^*, \sigma_1^*, \sigma_2^*)$  and for  $i \in \{1, 2\}$

$$\hat{p}_i \xrightarrow[n \rightarrow \infty]{a.s.} p_i^* \quad \text{and} \quad \hat{\sigma}_i \xrightarrow[n \rightarrow \infty]{a.s.} \sigma_i^*.$$

In order to simplify the following calculations, we denote by

$$\Delta_i^* = \sqrt{p_1^*(\sigma_1^*)^2 \beta_1 s_{i1}^2 + p_2^*(\sigma_2^*)^2 \beta_2 s_{i2}^2} \quad \text{and} \quad \delta_i^* = \frac{-1}{\Delta_i^*}.$$

There is a strong matching between the solutions obtained by solving (12)-(15) and their empirical counterparts obtained by Monte-Carlo simulations. This is illustrated in Fig. 3.

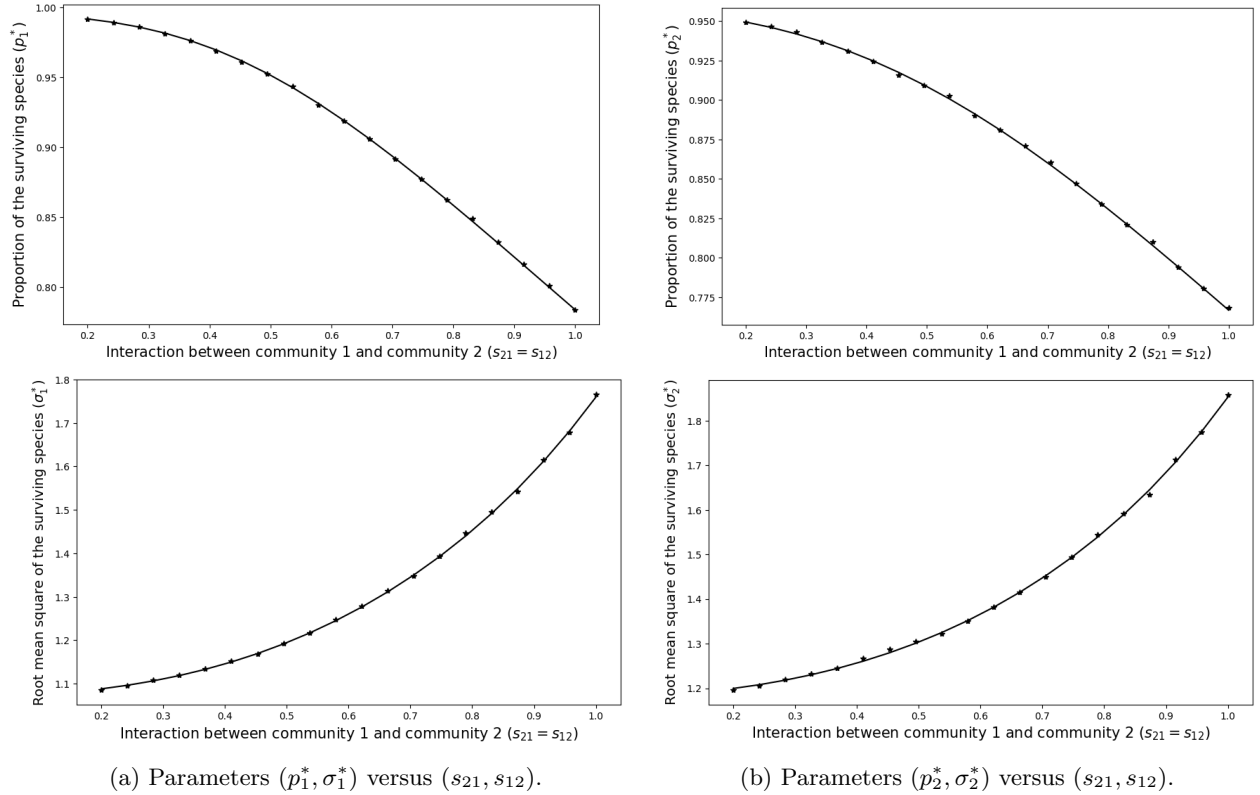


Figure 3: Comparison between the theoretical solutions  $(p_1^*, p_2^*, \sigma_1^*, \sigma_2^*)$  of (12)-(15) and their empirical Monte Carlo counterpart (the star marker) as functions of the off-diagonal block interaction strength  $(s_{12}, s_{21})$ . The left column is associated to the properties of community 1. The right column is associated to the properties of community 2. Matrix B has size  $n = 500$  and the number of Monte Carlo experiments is 500. The parameters are  $\mathbf{s} = \begin{pmatrix} 1/3 & s_{12} \\ s_{21} & 1/\sqrt{2} \end{pmatrix}$ ,  $\boldsymbol{\beta} = (\frac{1}{2}, \frac{1}{2})$ . When off-diagonal block interactions  $s_{12}, s_{21}$  increase, the proportion of surviving species  $p^*$  decreases but the root mean square of their equilibrium abundances  $\sigma^*$  increases.

### 3.2 Construction of the heuristics

Obtaining information about the fixed point is equivalent to solving the LCP problem

$$x_k^* \left( 1 - x_k^* + \sum_{\ell=1}^n B_{k\ell} x_\ell^* \right) = 0, \quad \forall k \in [n].$$

Consider the random variables:

$$\forall k \in [n], \quad \check{Z}_k = \sum_{\ell \in \mathcal{S}_1 \cup \mathcal{S}_2} B_{k\ell} x_\ell^*.$$

We assume that asymptotically the  $x_\ell^*$ 's are independent from the  $B_{k\ell}$ 's, an assumption supported by the chaos hypothesis, see for instance Geman and Hwang [GH82]. Denote by  $\mathbb{E}_{\mathbf{x}^*} = \mathbb{E}(\cdot \mid \mathbf{x}^*)$  the conditional expectation with respect to  $\mathbf{x}^*$ . Notice that conditionally to  $\mathbf{x}^*$ , the  $\check{Z}_k$ 's are independent Gaussian random variables, whose first two moments can easily be computed, see Appendix C for the details:

$$\forall k \in \mathcal{I}_i, \quad \text{Var}_{\mathbf{x}^*}(\check{Z}_k) = \hat{p}_1 \hat{\sigma}_1^2 \beta_1 s_{i1}^2 + \hat{p}_2 \hat{\sigma}_2^2 \beta_2 s_{i2}^2.$$

Notice that the fact that  $\text{Var}_{\mathbf{x}^*}(\check{Z}_k)$  only depends on  $\hat{p}_1, \hat{p}_2, \hat{\sigma}_1, \hat{\sigma}_2$  (which are converging quantities when  $n \rightarrow \infty$ ) supports the idea that  $\check{Z}_k$  is unconditionally a Gaussian random variable with second moment:

$$\text{Var}(\check{Z}_k) = p_1^* (\sigma_1^*)^2 \beta_1 s_{i1}^2 + p_2^* (\sigma_2^*)^2 \beta_2 s_{i2}^2,$$

where  $p_1^*, p_2^*, \sigma_1^*, \sigma_2^*$  are resp. the limits of  $\hat{p}_1, \hat{p}_2, \hat{\sigma}_1, \hat{\sigma}_2$ . We can introduce two families of standard Gaussian random variables  $(Z_k)_{k \in \mathcal{I}_1}$  and  $(Z_k)_{k \in \mathcal{I}_2}$ :

$$\forall k \in \mathcal{I}_i, \quad Z_k = \frac{\check{Z}_k}{\sqrt{\text{var}(\check{Z}_k)}} = \frac{\check{Z}_k}{\sqrt{p_1^* (\sigma_1^*)^2 \beta_1 s_{i1}^2 + p_2^* (\sigma_2^*)^2 \beta_2 s_{i2}^2}}.$$

Consider the equilibrium  $\mathbf{x}^* = (x_k^*)_{k \in [n]}$ , the definition of the LCP equilibrium implies if  $k \in \mathcal{S}_1 \cup \mathcal{S}_2$ :

$$x_k^* (1 - x_k^* + (B\mathbf{x}^*)_k) = 0 \quad \text{and} \quad 1 + (B\mathbf{x}^*)_k = 1 + \check{Z}_k > 0.$$

We finally obtain the following relationship for the surviving species:

$$x_k^* = 1 + \Delta_i^* Z_k \quad \text{for} \quad k \in \mathcal{S}_i. \quad (17)$$

Note that  $\Delta_i^*$  corresponds to the average variance of the interactions on community  $i$ .

**Heuristics (12)-(13).** We can write the first two equations:

$$\begin{aligned} \mathbb{P}(x_k^* \geq 0 \mid k \in \mathcal{S}_1) &= \mathbb{P}(Z_k < \delta_1^* \mid k \in \mathcal{S}_1) = 1 - \Phi(\delta_1^*), \\ \mathbb{P}(x_k^* \geq 0 \mid k \in \mathcal{S}_2) &= \mathbb{P}(Z_k < \delta_2^* \mid k \in \mathcal{S}_2) = 1 - \Phi(\delta_2^*), \end{aligned}$$

and finally obtain (12) and (13):

$$p_1^* = 1 - \Phi(\delta_1^*) \quad \text{and} \quad p_2^* = 1 - \Phi(\delta_2^*).$$

**Heuristics (14)-(15).** Our starting point is the following generic representation of an abundance at equilibrium (either of a surviving or vanishing species) in the case  $k \in \mathcal{S}_i$ :

$$x_k^* = (1 + \Delta_i^* Z_k) \mathbf{1}_{\{Z_k > \delta_i^*\}} = \mathbf{1}_{\{Z_k > \delta_i^*\}} + (\Delta_i^* Z_k) \mathbf{1}_{\{Z_k > \delta_i^*\}}.$$

Taking the square, we get:

$$(x_k^*)^2 = (1 + \Delta_i^* Z_k)^2 \mathbf{1}_{\{Z_k > \delta_i^*\}} = \mathbf{1}_{\{Z_k > \delta_i^*\}} + 2\Delta_i^* Z_k \mathbf{1}_{\{Z_k > \delta_i^*\}} + ((\Delta_i^*)^2 Z_k^2) \mathbf{1}_{\{Z_k > \delta_i^*\}}.$$

Summing over  $\mathcal{S}_i$  and normalizing, we get

$$\begin{aligned} \frac{1}{|\mathcal{S}_i|} \sum_{k \in \mathcal{S}_i} (x_k^*)^2 &= \frac{1}{|\mathcal{S}_i|} \sum_{k \in \mathcal{S}_i} \mathbf{1}_{\{Z_k > \delta_i^*\}} + 2\Delta_i^* \frac{1}{|\mathcal{S}_i|} \sum_{k \in \mathcal{S}_i} Z_k \mathbf{1}_{\{Z_k > \delta_i^*\}} + (\Delta_i^*)^2 \frac{1}{|\mathcal{S}_i|} \sum_{k \in \mathcal{S}_i} Z_k^2 \mathbf{1}_{\{Z_k > \delta_i^*\}}, \\ \hat{\sigma}_i^2 &\stackrel{(a)}{=} 1 + 2\Delta_i^* \frac{|\mathcal{I}_i|}{|\mathcal{S}_i|} \frac{1}{|\mathcal{I}_i|} \sum_{k \in \mathcal{I}_i} Z_k \mathbf{1}_{\{Z_k > \delta_i^*\}} + (\Delta_i^*)^2 \frac{|\mathcal{I}_i|}{|\mathcal{S}_i|} \frac{1}{|\mathcal{I}_i|} \sum_{k \in \mathcal{I}_i} Z_k^2 \mathbf{1}_{\{Z_k > \delta_i^*\}}, \\ \hat{\sigma}_i^2 &\stackrel{(b)}{\simeq} 1 + 2\Delta_i^* \frac{1}{\mathbb{P}(Z > \delta_i^*)} \mathbb{E}(Z \mathbf{1}_{\{Z > \delta_i^*\}}) + (\Delta_i^*)^2 \frac{1}{\mathbb{P}(Z > \delta_i^*)} \mathbb{E}(Z^2 \mathbf{1}_{\{Z > \delta_i^*\}}), \\ \hat{\sigma}_i^2 &\simeq 1 + 2\Delta_i^* \mathbb{E}(Z \mid Z > \delta_i^*) + (\Delta_i^*)^2 \mathbb{E}(Z^2 \mid Z > \delta_i^*). \end{aligned}$$

where (a) follows from the fact that  $|\mathcal{S}_i| = \sum_{k \in \mathcal{S}_i} \mathbf{1}_{\{Z_k > \delta_i^*\}}$  (by definition of  $\mathcal{S}_i$ ), (b) from the law of large numbers  $\frac{1}{|\mathcal{I}_i|} \sum_{k \in \mathcal{I}_i} Z_k^j \mathbf{1}_{\{Z_k > \delta_i^*\}} \xrightarrow[n \rightarrow \infty]{} \mathbb{E} Z^j \mathbf{1}_{\{Z > \delta_i^*\}}$ ,  $j = 1, 2$  and  $\frac{|\mathcal{S}_i|}{|\mathcal{I}_i|} \xrightarrow[n \rightarrow \infty]{} \mathbb{P}(Z > \delta_i^*)$  with  $Z \sim \mathcal{N}(0, 1)$ . It remains to replace  $\hat{\sigma}_i$  by its limit  $\sigma_i^*$  to obtain (14)-(15). We finally obtain the third and fourth equations:

$$\begin{aligned} (\sigma_1^*)^2 &= (1 + \lambda_1^*)^2 + 2(1 + \lambda_1^*)\Delta_1^* \mathbb{E}(Z | Z > \delta_1^*) + (\Delta_1^*)^2 \mathbb{E}(Z^2 | Z > \delta_1^*) , \\ (\sigma_2^*)^2 &= (1 + \lambda_2^*)^2 + 2(1 + \lambda_2^*)\Delta_2^* \mathbb{E}(Z | Z > \delta_2^*) + (\Delta_2^*)^2 \mathbb{E}(Z^2 | Z > \delta_2^*) . \end{aligned}$$

### 3.3 General properties of the ecosystem

The properties at equilibrium, such as the proportion and mean square of the abundance of surviving species, can be computed for each community by solving the system of equations in Heuristics 1. An additional property, the mean abundance of the surviving species at equilibrium for each community  $(m_1^*, m_2^*)$ , can be calculated using a method similar to the mean square of the abundances (see Appendix C.2 for the details of the computations).

$$m_1^* = 1 + \Delta_1^* \mathbb{E}(Z | Z > \delta_1^*) , \quad (18)$$

$$m_2^* = 1 + \Delta_2^* \mathbb{E}(Z | Z > \delta_2^*) . \quad (19)$$

The two equations are not necessary for solving Heuristics 1, but they provide new information. In particular, strong inter- or intra-community interactions increase the mean abundance of the surviving species (see Fig. 4).

Conditional on each community, one can easily extend the properties of each community to the whole ecosystem. We denote by  $p^*$  the proportion,  $(\sigma^*)^2$  the mean square and  $m^*$  the mean of surviving species. We observe the linear effect of community size  $\beta$  on general properties:

1. Proportion of surviving species.

$$\begin{aligned} \mathbb{P}(x_k^* \geq 0) &= \mathbb{P}(x_k^* \geq 0 | k \in \mathcal{I}_1) \mathbb{P}(k \in \mathcal{I}_1) + \mathbb{P}(x_k^* \geq 0 | k \in \mathcal{I}_2) \mathbb{P}(k \in \mathcal{I}_2) , \\ p^* &= p_1^* \beta_1 + p_2^* \beta_2 . \end{aligned}$$

2. Mean square of the abundance of the surviving species.

$$\begin{aligned} \mathbb{E}((x_k^*)^2) &= \mathbb{E}((x_k^*)^2 | k \in \mathcal{I}_1) \mathbb{P}(k \in \mathcal{I}_1) + \mathbb{E}((x_k^*)^2 | k \in \mathcal{I}_2) \mathbb{P}(k \in \mathcal{I}_2) , \\ (\sigma^*)^2 &= (\sigma_1^*)^2 \beta_1 + (\sigma_2^*)^2 \beta_2 . \end{aligned}$$

3. Mean of the abundance of the surviving species

$$\begin{aligned} \mathbb{E}(x_k^*) &= \mathbb{E}(x_k^* | k \in \mathcal{I}_1) \mathbb{P}(k \in \mathcal{I}_1) + \mathbb{E}(x_k^* | k \in \mathcal{I}_2) \mathbb{P}(k \in \mathcal{I}_2) , \\ m^* &= m_1^* \beta_1 + m_2^* \beta_2 . \end{aligned}$$

### 3.4 Distribution of the surviving species

We may recall the following representation of the abundance  $x_k^*$  of a surviving species when  $k \in \mathcal{S}_i$ :

$$x_k^* = 1 + \Delta_i^* Z_k \text{ if } k \in \mathcal{S}_i ,$$

where  $Z_k \sim \mathcal{N}(0, 1)$  and  $Z_k > \delta_i^* = \delta_i(p_i^*, \sigma_i^*)$  defined in (17). This representation allows to characterize the distribution of  $x_k^*$  of each community. It turns out that the surviving species of each community follow a truncated Gaussian distribution.

**Heuristics 2.** Let  $\mathbf{s}$  be the  $2 \times 2$  matrix of interaction strengths and assume that the condition of Theorem 3 holds and let  $(p_1^*, p_2^*, \sigma_1^*, \sigma_2^*)$  be the solution of the system (12)-(15). Recall the definition (16) of  $\Delta_i$  and  $\delta_i$

and denote by  $\delta_i^* = \delta_i(p_i^*, \sigma_i^*)$ . Let  $x_k^* > 0$  be a positive component of  $\mathbf{x}^*$  belonging to the community  $i$ , then the law of  $x_k^*$  is

$$\mathcal{L}(x_k^*) \xrightarrow{n \rightarrow \infty} \mathcal{L}(1 + \Delta_i^* Z \mid Z > \delta_i^*) ,$$

where  $Z \sim \mathcal{N}(0, 1)$ . Otherwise stated, asymptotically for  $k \in \mathcal{S}_i$ ,  $x_k^*$  admits the following density

$$f_k(y) = \frac{\mathbf{1}_{\{y > 0\}}}{\Phi(-\delta_i^*)} \frac{1}{\Delta_i^* \sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left( \frac{y}{\Delta_i^*} + \delta_i^* \right)^2 \right\} . \quad (20)$$

The heuristics simply follows from the fact that if  $x_k^*$  is a surviving species and  $k \in \mathcal{S}_i$  then

$$x_k^* = 1 + \Delta_i^* Z_k ,$$

conditionally on the fact that the right-hand side of the equation is positive, that is  $Z_k > \delta_i^*$ . A simple change of variable yields the density - details are provided in Appendix C.

Fig. 4 illustrates the matching between the theoretical distribution obtained by equation (20) and a histogram obtained by generating the interaction matrix for 2 communities. In Fig. 5, the validity of heuristics in the case of non-Gaussian entries is illustrated.

**Remark 4.** The proof relies on the Gaussianity assumption, but we are convinced that it could be extended beyond. In particular, in Figure 5, non-Gaussian entries centered  $\mathbb{E}(B_{k\ell}) = 0$  with variance one  $\mathbb{E}(|B_{k\ell}|^2) = 1$  are considered. The distribution of surviving species still fits the truncated Gaussian in this case.

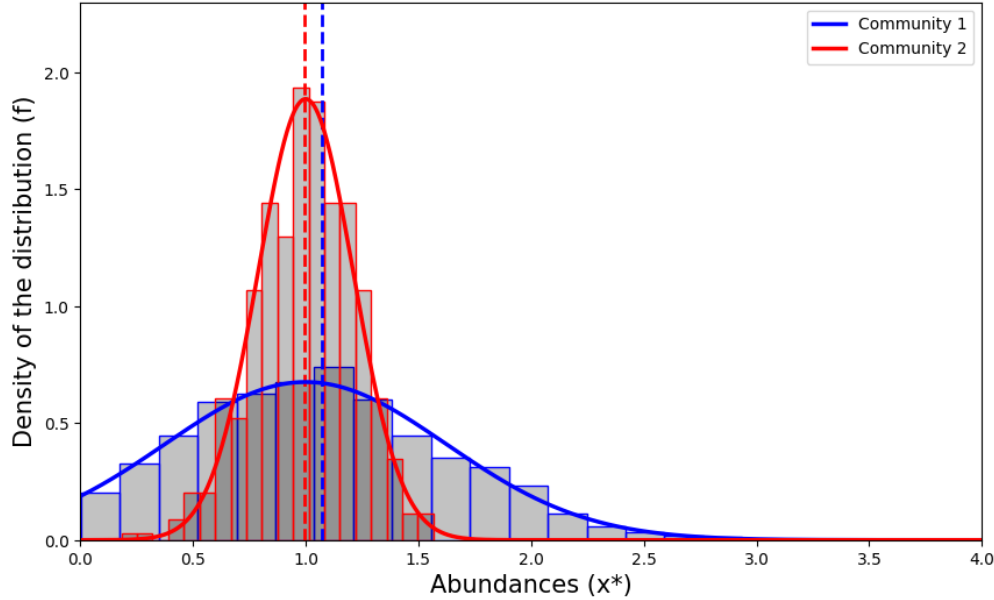


Figure 4: Distribution of surviving species in each community. The  $x$ -axis represents the value of the abundances and the histogram is built upon the positive components of equilibrium  $\mathbf{x}^*$  associated to each community. The blue-solid line (resp. red-solid line) represents the theoretical distribution of community 1 (resp. community 2) for parameters  $\mathbf{s}$  as given by Heuristics 2. The blue-dashed vertical line (resp. red-dashed vertical line) corresponds to the mean abundance of community 1 (resp. community 2) given by equations (18)-(19). All the matrix entries  $A_{ij}$ 's are independent Gaussian  $\mathcal{N}(0, 1)$ ; the parameters are set to  $n = 2000$ ,  $\beta = (0.75, 0.25)$  and  $\mathbf{s} = \begin{pmatrix} 1/2 & 1/\sqrt{2} \\ 1/5 & 1/9 \end{pmatrix}$ .

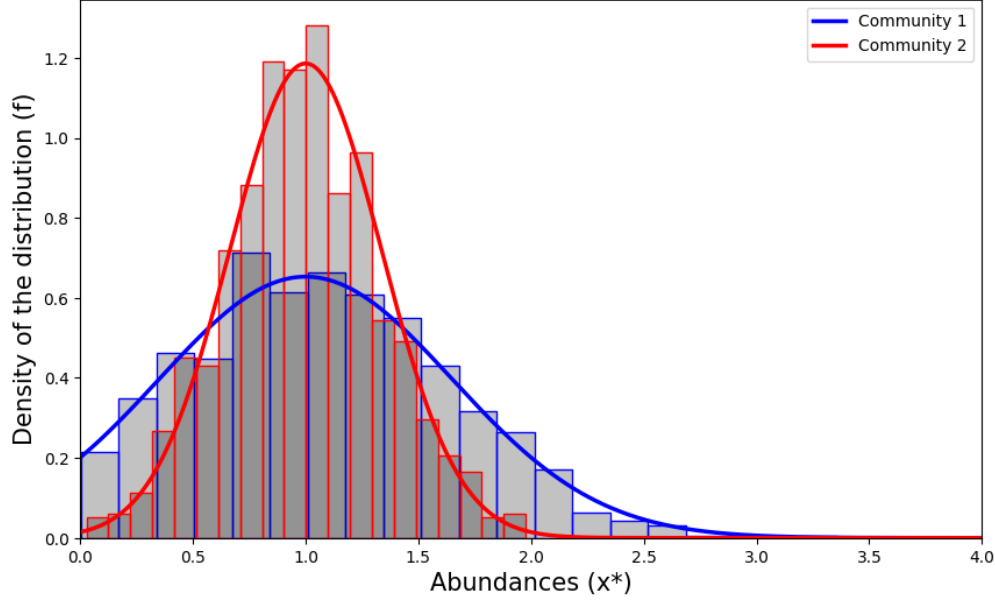


Figure 5: Distribution of surviving species in each community with non-Gaussian entries. The  $x$ -axis represents the value of the abundances. The histogram is built upon the positive components of equilibrium  $\mathbf{x}^*$  associated to each community. (blue for community 1, red for community 2). The solid lines represent the theoretical distributions associated to parameters  $\mathbf{s}$  as given by Heuristics 2. The entries of the  $A_{ij}$  matrices are uniform  $\mathcal{U}(-\sqrt{3}, \sqrt{3})$  with variance 1 and the parameters are set to  $n = 2000$ ,  $\beta = (0.5, 0.5)$  and  $\mathbf{s} = \begin{pmatrix} 1/2 & 2/3 \\ 1/3 & 1/4 \end{pmatrix}$ . Notice in particular that the histogram is well predicted by the theoretical distributions even if the entries are non-Gaussian.

## 4 Feasibility

Recall  $s$  the interaction parameter in the case of a unique community. According to the work of Dougoud *et al.* [DVR<sup>+</sup>18], if  $s$  is fixed (i.e. does not depend on  $n$ ) then there can be no feasible equilibrium at large  $n$ . Following this work, Bizeul and Najim [BN21] provided the appropriate normalization of  $s$  to have a feasible equilibrium. The threshold corresponds to  $s \sim 1/\sqrt{2\log(n)}$ . The equilibrium is feasible almost surely when  $s$  is less than this threshold value, i.e. when elements of random matrix  $B$  are divided by  $\sqrt{2n\log(n)}$  or a larger factor. Some extensions of these results have been made in the sparse case [AN22] and with a mean and pairwise correlated entries [CEFN22]. In this section, conditions are given on the matrices  $\mathbf{s}$  to get a feasible equilibrium in each community, called co-feasibility. We then provide some ecological interpretations.

### 4.1 Theoretical analysis of the threshold

Recall the notation  $\mathbf{x} = (x_k)_{k \in [n]}$  and denote by  $\|\mathbf{x}\|_\infty = \max_{k \in [n]} |x_k|$ . We are interested in the existence of a feasible solution of the fixed point problem associated with the model (3). To consider this problem, we extend the computations of Bizeul and Najim in the framework of a block structure network. Consider  $\mathbf{s}$  such that  $I - B$  is invertible. The problem is defined by

$$\mathbf{x}^* = \mathbf{1} + B\mathbf{x}^* \Leftrightarrow \mathbf{x}^* = (I - B)^{-1}\mathbf{1}, \quad (21)$$

The problem (21) admits a unique solution. We consider a matrix  $\mathbf{s}$  which depends on  $n$ , i.e.  $\mathbf{s} = \mathbf{s}_n$  such that:

$$\mathbf{s}_n \xrightarrow{n \rightarrow \infty} \mathbf{0} \Leftrightarrow \forall i, j \in \{1, 2\}, s_{ij} \xrightarrow{n \rightarrow \infty} 0.$$

Note that for sufficiently large  $n$ , the problem satisfies the sufficient condition of Theorem 3 to have a unique globally stable equilibrium.

Let matrix  $B_n$  depending on the interaction matrix  $\mathbf{s}_n$  defined by

$$B_n = V \mathbf{s}_n V^\top \circ \frac{1}{\sqrt{n}} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}, \quad (22)$$

where

$$V \in \mathcal{M}_{n \times 2}, \quad V = \begin{pmatrix} \mathbf{1}_{\mathcal{I}_1} & 0 \\ 0 & \mathbf{1}_{\mathcal{I}_2} \end{pmatrix}.$$

The spectral radius of  $\frac{1}{\sqrt{n}} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$  a.s. converges to 1 due to the circular law [TVK10]. So as long as  $\mathbf{s}_n$  is close to zero, the matrix  $I - B_n$  is eventually (for large enough values of  $n$ ) invertible.

**Theorem 4** (Co-feasibility for the 2-blocks model). *Assume that matrix  $B_n$  is defined by the 2-blocks model (22). Let  $\boldsymbol{\beta} = (\beta_1, \beta_2)$ ,  $\beta_1 + \beta_2 = 1$  represents the proportion of each community. Let  $\mathbf{s}_n \xrightarrow[n \rightarrow \infty]{} 0$  and denote by  $s_n^* = 1/\sqrt{2 \log n}$  the critical threshold. Let  $x_n = (x_k)_{k \in [n]}$  be the solution of (21).*

1. *If there exists  $\varepsilon > 0$  such that eventually  $\left\| (\mathbf{s}_n \circ \mathbf{s}_n) \boldsymbol{\beta}^\top \right\|_\infty \geq (1 + \varepsilon)(s_n^*)^2$  then*

$$\mathbb{P} \left\{ \min_{k \in [n]} x_k > 0 \right\} \xrightarrow[n \rightarrow \infty]{} 0.$$

2. *If there exists  $\varepsilon > 0$  such that eventually  $\left\| (\mathbf{s}_n \circ \mathbf{s}_n) \boldsymbol{\beta}^\top \right\|_\infty \leq (1 - \varepsilon)(s_n^*)^2$  then*

$$\mathbb{P} \left\{ \min_{k \in [n]} x_k > 0 \right\} \xrightarrow[n \rightarrow \infty]{} 1.$$

A sketch of proof is postponed to Appendix D, the extension to the  $b$ -blocks case can be found in Appendix E.

Proof of Theorem 4 strongly depends on the assumption of Gaussianity and equal growth rates of each species. However, according to the approach of Bizeul *et al.* [BN21], these assumptions could be relaxed. In particular, the phenomenon seems to be universal, i.e. the feasibility threshold works for a wide range of distribution choices.

in the critical regime  $s \propto 1/\sqrt{\log(n)}$  or equivalently  $s^{-1} \propto \sqrt{\log(n)}$ . We thus introduce matrix  $\boldsymbol{\kappa}$  defined by

$$\boldsymbol{\kappa} = \frac{1}{\sqrt{\log(n)}} \begin{pmatrix} s_{11}^{-1} & s_{12}^{-1} \\ s_{21}^{-1} & s_{22}^{-1} \end{pmatrix}.$$

Notice that at criticality  $\boldsymbol{\kappa}$  will be of order  $O(1)$ . This will be convenient for ecological interpretations. Using the inequality of Theorem 4, the co-feasibility condition on  $\boldsymbol{\kappa}$  writes

$$\left\| (\mathbf{s}_n \circ \mathbf{s}_n) \boldsymbol{\beta}^\top \right\|_\infty < (s_n^*)^2 \quad \Leftrightarrow \quad \max \left( \frac{2\beta_1}{\kappa_{11}^2} + \frac{2\beta_2}{\kappa_{12}^2}, \frac{2\beta_1}{\kappa_{21}^2} + \frac{2\beta_2}{\kappa_{22}^2} \right) < 1. \quad (23)$$

If for  $i = 1, 2$ ,  $\beta_i = \frac{1}{2}$  and the entry of the matrix  $\boldsymbol{\kappa}$  are equal, then condition (23) gives the threshold  $\kappa_{ij} > \sqrt{2}$ , and we recover the same critical threshold  $\sqrt{2 \log(n)}$  as in [BN21].

**Remark 5.** Assume  $\kappa_{11} = \kappa_{22} = \nu_1$  and  $\kappa_{12} = \kappa_{21} = \nu_2$ , condition (23) is reformulated as:

$$\max \left( \frac{2\beta_1}{\nu_1^2} + \frac{2\beta_2}{\nu_2^2}, \frac{2\beta_1}{\nu_2^2} + \frac{2\beta_2}{\nu_1^2} \right) < 1.$$

If  $\beta_1, \beta_2$  and  $\nu_2$  are fixed, then the phase transition on the intra-community interactions occurs at

$$\nu_1 > \min \left( \sqrt{\frac{\beta_1}{\frac{1}{2} - \frac{\beta_2}{\nu_2^2}}}, \sqrt{\frac{\beta_2}{\frac{1}{2} - \frac{\beta_1}{\nu_2^2}}} \right).$$

In Fig. 6, the phase transition is represented for a selected set of parameters. Note that the transition is rather smooth. The threshold depends on  $\nu_2$ . Increasing  $\nu_2$  (decreasing the inter-block interactions) lowers the co-feasibility threshold to at least 1 (for communities of the same size).

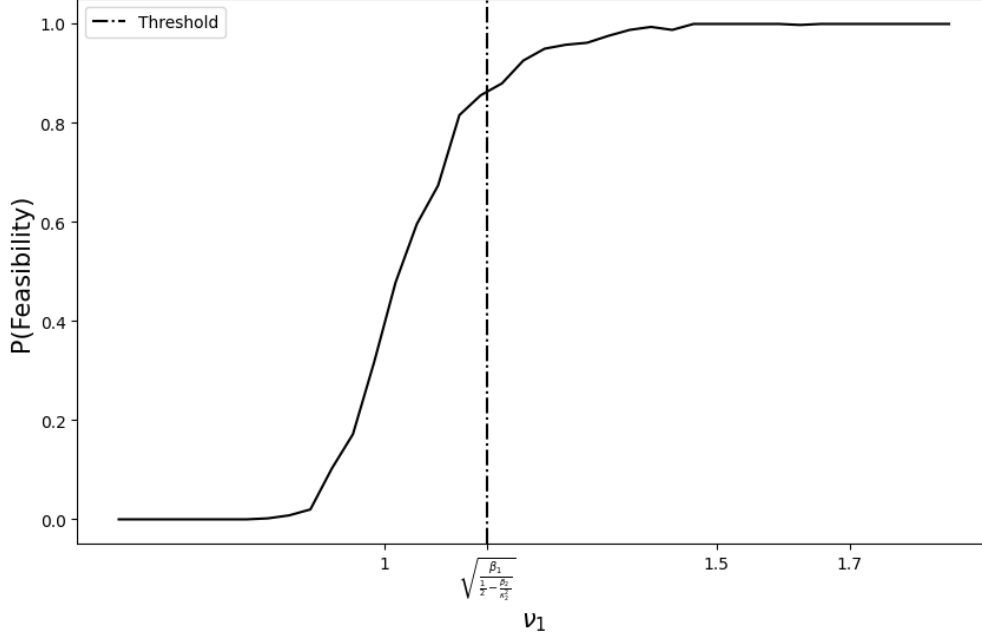


Figure 6: Transition towards co-feasibility for the 2-blocks model (2). For each value  $\nu_1$  on the x-axis, we simulate 500 matrices  $B$  of size  $n = 5000$  with two communities of the same size ( $\beta_1 = \beta_2 = 0.5$ ) with the inter-block interactions fixed at  $s_{21}(\nu_2) = s_{12}(\nu_2) = 1/2\sqrt{\log(n)}$  and compute the solution  $\mathbf{x}$  of Theorem 4 at the scaling for the intra-block interactions  $s_{11}(\nu_1) = s_{22}(\nu_1) = 1/\nu_1\sqrt{\log(n)}$ . The curve represents the proportion of feasible solutions  $\mathbf{x}$  obtained for the 500 simulations. The dotdashed vertical line corresponds to  $\nu_1 = \sqrt{\frac{\beta_1}{\frac{1}{2} - \frac{\beta_2}{\nu_2^2}}} = 2/\sqrt{3}$ .

## 4.2 Preservation of co-feasibility

Equation (23) defines a “co-feasibility domain” and gives a constraint in five dimensions. The two communities of species can be studied independently i.e. the two components of equation (23) respectively give the feasibility condition for each community:

$$\begin{cases} \text{If } \frac{2\beta_1}{\kappa_{11}^2} + \frac{2\beta_2}{\kappa_{12}^2} < 1, \text{ then community 1 is feasible.} \\ \text{If } \frac{2\beta_1}{\kappa_{21}^2} + \frac{2\beta_2}{\kappa_{22}^2} < 1, \text{ then community 2 is feasible.} \end{cases}$$

The first community (resp. the second one) will be affected by changing  $\kappa_{11}$ ,  $\kappa_{12}$  (resp.  $\kappa_{21}$ ,  $\kappa_{22}$ ). In general, increasing the inter- or intra- interaction strength will decrease the probability of having a co-feasible equilibrium.



If  $\kappa_{12} = \kappa_{21} = \infty$ , then condition (23) gives the co-feasibility conditions for each community:

$$s_{11} < \frac{1}{\sqrt{2\beta_1 \log(n)}} \quad \text{and} \quad s_{22} < \frac{1}{\sqrt{2\beta_2 \log(n)}}.$$

For the same  $s$ , it means

$$s < \frac{1}{\sqrt{2 \log(n) \max(\beta_1, \beta_2)}}.$$

As an example of application, suppose we start with co-feasible communities of equal size ( $\beta_1 = \beta_2 = 0.5$ ) and add interactions between these two groups, co-feasibility may be dropped (see Fig. 2). The co-feasibility domain is illustrated in Fig. 7. It shows a threshold where the co-feasibility property is satisfied above the curve. This means that the lower the values of  $\kappa_{11}$  and  $\kappa_{22}$ , i.e. the stronger the interactions within the groups, the more likely the co-feasibility property is lost. We can conclude that an independent group structure is more likely to be co-feasible and therefore stable, which supports previous work on compartmentalization models [SB11].

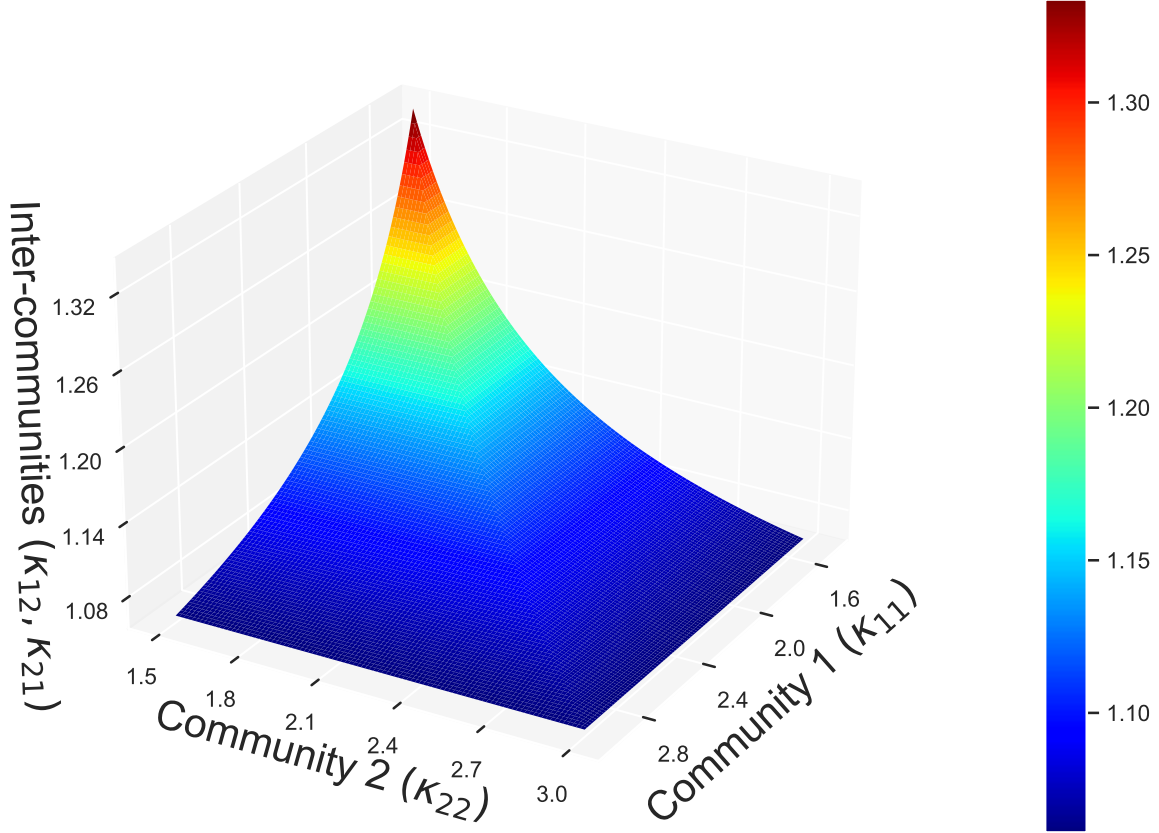


Figure 7: Representation of the co-feasibility phase diagram. The co-feasible domain is above the surface. The z-axis (resp x-axis) is the strength of interaction within community 1 -  $\kappa_{11}$  (resp community 2 -  $\kappa_{22}$ ). The y-axis is the inter-community interactions  $\kappa_{12} = \kappa_{21}$ . The colored area illustrates the threshold between the co-feasible and non-co-feasible domains in the system (3).

### 4.3 Impact of the community size

For a fixed matrix  $\kappa$ , the condition to have a co-feasible fixed point can be computed as a function of the size of each community i.e. the pair  $\beta = (\beta_1, \beta_2)$ . Starting from the co-feasibility inequality (23):

$$\max \left( \frac{2\beta_1}{\kappa_{11}^2} + \frac{2\beta_2}{\kappa_{12}^2}, \frac{2\beta_1}{\kappa_{21}^2} + \frac{2\beta_2}{\kappa_{22}^2} \right) < 1,$$

the two components are studied independently,

$$\frac{2\beta_1}{\kappa_{11}^2} + \frac{2(1-\beta_1)}{\kappa_{12}^2} < 1 \quad \Leftrightarrow \quad \beta_1 \left( \frac{2}{\kappa_{11}^2} - \frac{2}{\kappa_{12}^2} \right) < 1 - \frac{2}{\kappa_{12}^2} \quad \Rightarrow \quad \beta_1 < \frac{1 - \frac{2}{\kappa_{12}^2}}{\left( \frac{2}{\kappa_{11}^2} - \frac{2}{\kappa_{12}^2} \right)} \quad \text{if } \kappa_{11} < \kappa_{12}.$$

Similarly, one has

$$\sqrt{\frac{2\beta_1}{\kappa_{21}^2} + \frac{2(1-\beta_1)}{\kappa_{22}^2}} < 1 \quad \Rightarrow \quad \beta_1 > \frac{1 - \frac{2}{\kappa_{22}^2}}{\left( \frac{2}{\kappa_{21}^2} - \frac{2}{\kappa_{22}^2} \right)} \quad \text{if } \kappa_{22} < \kappa_{21}.$$

In the case where the intra-community interactions ( $\kappa_{11}, \kappa_{22}$ ) are smaller than the inter-community interactions ( $\kappa_{12}, \kappa_{21}$ ), we obtain an upper and a lower bound for the admissible size of each community  $\beta_1, \beta_2$  to have a co-feasible equilibrium. In Fig. 8, different cases of the co-feasibility zone are represented according to the inter-community interactions ( $\kappa_{12}, \kappa_{21}$ ). If the intra-community interactions are different, the community with the lowest interaction  $\kappa_{ii}$  is advantaged i.e. the size of the community can be larger.

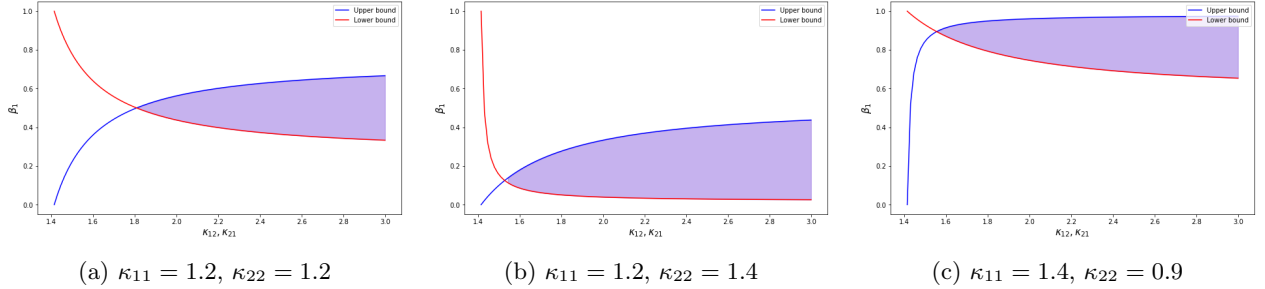


Figure 8: Representation of the co-feasibility domain depending on the fixed intra-community interaction. In (a), (b), (c), a different scenario of intra-community interaction is presented. Each panel represents the upper-bound (blue curve) and the lower-bound (red curve) of the size of community 1 as a function of the interaction between two communities ( $\kappa_{12}, \kappa_{21}$ ). The blue area is the admissible zone to have a co-feasible fixed point in (3). The size of community 2 is equal to  $\beta_2 = 1 - \beta_1$ .

### 4.4 Connection increases co-feasibility

In Section 4.2, we analyzed the co-feasibility condition for a scenario involving two communities. We presented a co-feasibility domain defined by

$$\begin{aligned} & \left\| (s_n \circ s_n) \beta^\top \right\|_\infty < \frac{1}{2 \log(n)} := (s_n^*)^2, \\ \Leftrightarrow & \max (\beta_1 s_{11}^2 + \beta_2 s_{12}^2, \beta_1 s_{21}^2 + \beta_2 s_{22}^2) < \frac{1}{2 \log(n)}, \\ \Leftrightarrow & \beta_1 s_{11}^2 + \beta_2 s_{12}^2 < \frac{1}{2 \log(n)} \quad \text{and} \quad \beta_1 s_{21}^2 + \beta_2 s_{22}^2 < \frac{1}{2 \log(n)}. \end{aligned} \quad (24)$$

These two distinct conditions within the communities have led us to the conclusion that community isolation is beneficial for coexistence. However, general constraints that affect all interactions could be considered

at the ecosystem scale. To this end, we introduce a complementary condition: the global variance of the interaction coefficients in the ecosystem remains invariant, i.e.

$$\beta_1 s_{11}^2 + \beta_2 s_{12}^2 + \beta_1 s_{21}^2 + \beta_2 s_{22}^2 = \Gamma,$$

with  $\Gamma > 0$ .

In order to remove the  $n$  dependency, we define  $\gamma_{ij}^2 = s_{ij}^2 \log(n)$  and  $\Gamma_n = \Gamma \log(n)$  (note that the  $\gamma$  coefficients are defined differently from the  $\kappa$  coefficients). Combined with condition (24), we get the following system of equations:

$$\begin{cases} \beta_1 \gamma_{11}^2 + \beta_2 \gamma_{12}^2 + \beta_1 \gamma_{21}^2 + \beta_2 \gamma_{22}^2 = \Gamma_n, \\ 2\beta_1 \gamma_{11}^2 + 2\beta_2 \gamma_{12}^2 < 1, \quad (\text{feasibility condition for community 1}), \\ 2\beta_1 \gamma_{21}^2 + 2\beta_2 \gamma_{22}^2 < 1, \quad (\text{feasibility condition for community 2}). \end{cases} \quad (25)$$

Assuming the fixed intra-community variances  $\gamma_{11}$  and  $\gamma_{22}$ , we seek to determine the co-feasibility conditions for the inter-community variances  $\gamma_{12}$  and  $\gamma_{21}$ . In this case, the constraint on the total variance corresponds to the equation of an ellipse in the  $(\gamma_{12}, \gamma_{21})$  plane:

$$\begin{aligned} & \beta_1 \gamma_{11}^2 + \beta_2 \gamma_{12}^2 + \beta_1 \gamma_{21}^2 + \beta_2 \gamma_{22}^2 = \Gamma_n, \\ \Leftrightarrow & \left( \frac{\beta_2}{\Gamma_n - \beta_1 \gamma_{11}^2 - \beta_2 \gamma_{22}^2} \right) \gamma_{12}^2 + \left( \frac{\beta_1}{\Gamma_n - \beta_1 \gamma_{11}^2 - \beta_2 \gamma_{22}^2} \right) \gamma_{21}^2 = 1, \end{aligned}$$

The equations below provide the values for the semi-major axis  $a$  and the semi-minor axis  $b$  of the ellipse:

$$a = \sqrt{\frac{\Gamma_n - \beta_1 \gamma_{11}^2 - \beta_2 \gamma_{22}^2}{\beta_2}}, \quad b = \sqrt{\frac{\Gamma_n - \beta_1 \gamma_{11}^2 - \beta_2 \gamma_{22}^2}{\beta_1}}.$$

Note that if both communities are of equal size ( $\beta_1 = \beta_2$ ), a circle with radius  $a$  is obtained.

From a visual standpoint, the conditions (25) are depicted in Figure 9. Since the coefficients  $(\gamma_{12}, \gamma_{21})$  are non-negative, we are only interested in the positive orthant. The feasibility condition for community 1 is given by the horizontal axis defined by

$$2\beta_1 \gamma_{11}^2 + 2\beta_2 \gamma_{12}^2 < 1 \quad \Leftrightarrow \quad \gamma_{12} < \sqrt{\frac{1 - 2\beta_1 \gamma_{11}^2}{2\beta_2}},$$

and the one of community 2 is given by the vertical axis defined by

$$2\beta_1 \gamma_{21}^2 + 2\beta_2 \gamma_{22}^2 < 1 \quad \Leftrightarrow \quad \gamma_{21} < \sqrt{\frac{1 - 2\beta_2 \gamma_{22}^2}{2\beta_1}}.$$

The intersection between the vertical (resp. horizontal) line and the ellipse occurs when the semi-major axis (resp. semi-minor axis) exceeds the vertical condition  $a > \sqrt{(1 - 2\beta_1 \gamma_{11}^2)/(2\beta_2)}$  (resp. horizontal condition  $b > \sqrt{(1 - 2\beta_2 \gamma_{22}^2)/(2\beta_1)}$ ).

**Remark 6.** By replacing the feasibility condition of  $\gamma_{21}$  in the equation of the ellipse, we can derive the intersection between the vertical axis and the ellipse as follows:

$$\gamma_{21}^2 = \frac{\Gamma_n - \beta_2 \gamma_{22}^2 - 1/2}{\beta_1},$$

equivalent to the feasibility condition of  $\gamma_{12}$  (by replacing  $\Gamma_n$ ):

$$\gamma_{12} < \sqrt{\frac{1 - 2\beta_1 \gamma_{11}^2}{2\beta_2}}.$$

We identify the range of co-feasibility between the two groups for  $\gamma_{21}$

$$\frac{\Gamma_n - \beta_2 \gamma_{22}^2 - 1/2}{\beta_1} < \gamma_{21}^2 < \frac{1 - 2\beta_2 \gamma_{22}^2}{2\beta_1}.$$

This simple framework allows for testing different scenarios. Figure 9a is the reference figure. It represents a situation where all the potential interactions between communities lead to co-feasibility. In Figure 9b, the total system variance is increased, which results in a reduction of co-feasibility options: the interactions between the two communities must be high  $\gamma_{12}, \gamma_{21} \gg 0$ . When the intra-community variances are increased as shown in Figure 9c, the ellipse shrinks, and the set of inter-community interaction variances is reduced. This observation reinforces the findings of Section 4.2 where large interaction between communities enables co-feasibility through community isolation. In the concluding example in Figure 9d, we reduce only the  $\gamma_{11}$  interaction in community 1. We observe that the impact of community 1 on 2  $\gamma_{21}$  must be weaker, but the impact of community 2 on 1  $\gamma_{12}$  can no longer be weak. Weaker interactions within community 1 imply a stronger connection between the communities for co-feasibility.

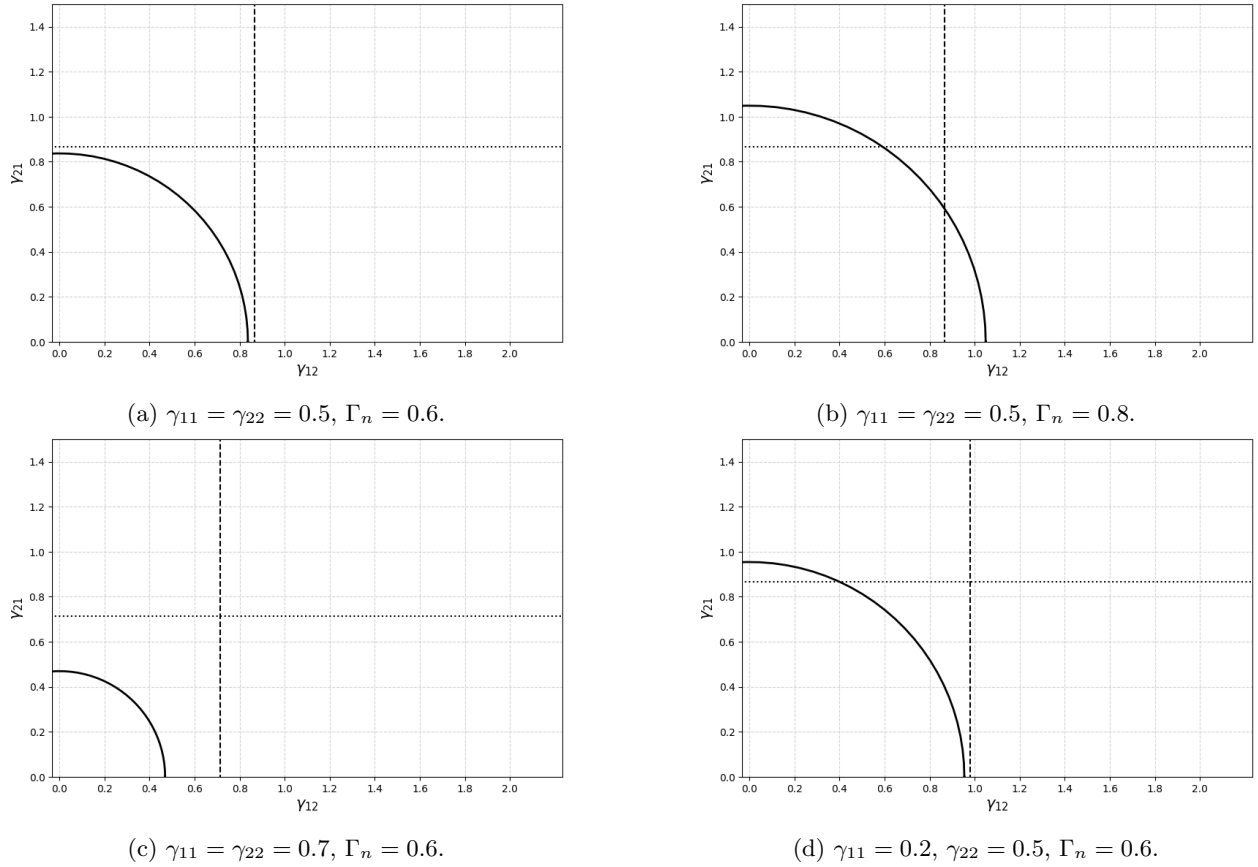


Figure 9: Graphical representation of equation system (25). The solid line circle constrains the total variance of the system (equation 1 in (25)), while the dashed and dotted lines correspond to feasibility conditions for community 1 (equation 2 in (25)) and 2 (equation 3 in (25)), respectively. In each figure, two communities of equal size are considered with  $\beta_1 = \beta_2 = 1/2$ , creating the solid line circle (rather than an ellipse) in this particular case. Each figure illustrates a distinct interaction situation between two communities outlined in the caption. Figure (a) serves as the reference against which the other figures (b)-(c)-(d) are compared. The co-feasibility arises only for the values of  $(\gamma_{12}, \gamma_{21})$  on the line found within the inner square at the bottom left-hand corner of the figure. When values of  $(\gamma_{12}, \gamma_{21})$  on the line are outside this inner square, then one or the other community is not feasible anymore.

## 5 Discussion

In this paper, we described a model of the dynamics of species abundances when the interaction among species is structured in multiple communities. The main interest is to outline the effect of a block structure on the stability and persistence of species. We defined an interaction matrix per block which has several characteristics such as the strength of the interactions  $\mathbf{s}$  and the size of the community  $\beta$ . Specifically, we described the dynamics and properties of each community in the system (feasibility, proportion of surviving species, mean and root mean square of the abundances of surviving species) and their effect on each other. In this context, we focused most of our analysis on the case of two interacting communities. However, our results can be extended to more than 2 communities (see Appendix E).

First, theoretical conditions were given for a unique globally stable equilibrium in the model (3) with surviving and vanishing species. This follows from Lyapunov conditions related to a result of Takeuchi and Adachi [TA80] and random matrix theory. These stability results had been found in the case of a single community by Clenet *et al.* [CMN23]. This complements the stability properties in the Lotka-Volterra system studied by Stone [Sto18] and Gibbs *et al.* [GGRA18]. Recent random matrix methods allow us to describe the spectrum of a block matrix and plot it numerically. For a detailed discussion of random matrices in the Lotka-Volterra model, see Akjouj *et al.* [ABC+22].

Subsequently, we gave heuristics on the surviving species (proportion, mean and root mean square of their abundances). These heuristics have also been found in the case of a single community by Clenet *et al.* [CMN23]. From a physicist's point of view and using the methods of Bunin [Bun17] and Galla [Gal18], Barbier *et al.* [BABL18] and Poley *et al.* [PBG23] have extended the heuristics in the block and cascade model. Previously, obtaining properties on surviving species in the LV model (not normalized by  $\sqrt{n}$ ) was already done by Servan *et al.* [SCG+18] where they consider a different growth rate for each species. The study of the stability and properties of surviving species in the LV system has also been carried out by Pettersson *et al.* [PSJ20a, PSJ20b]. From an ecological point of view, heuristics are derived from the properties of interactions between multiple communities.

In a third part, we studied the condition under which the feasibility threshold exists where all species co-exist. We extend the feasibility results found by Bizeul and Najim [BN21] in the case of a block structure. A co-feasibility threshold was found in the form of an inequality that must be verified to have a feasible community set. This complements the recent results on interactions with a sparse structure [AN22] and interactions with a correlation profile [CEFN22]. We notice that to maximize the probability of co-feasibility, we need to minimize the interactions between the communities. Additionally, a community with weaker interactions can exhibit a larger total abundance in the ecosystem while maintaining the co-feasibility threshold. At the ecosystem level, when a generic constraint that affects all interactions is added, weaker interactions within one of the communities suggest a stronger connection between the communities for co-feasibility.

There are still many mathematical and ecological questions that remain unanswered in this type of model.

First, a rigorous mathematical proof of the heuristics presented here would be of interest, although the LCP procedure induces an a priori statistical bias that is difficult to handle. This issue is still pending in the single community case [CMN23] and appears to be challenging to address. Recently, Akjouj *et al.* [AHMN23] provided a rigorous proof using an approximate message passing (AMP) approach in the single community model with an interaction matrix taken from the Gaussian Orthogonal Ensemble (GOE). Their approach was based on work by Hachem [Hac23].

Second, we could extend the heuristics for two different scenarios. On the one hand, it would be interesting to add pairwise correlations between species coefficients  $A_{ij}$ . This has already been done by physicists, see [BABL18, PBG23]. In the study of feasibility, it was shown that a correlation profile does not change the feasibility threshold [CEFN22]. On the other hand, for the sake of simplicity, we have chosen to set the growth rates equal to the same value  $r_k = 1$  for  $k \in [n]$ . It would be relevant to control the distribution of the growth rate as in [SCG+18] or to consider structural stability as in Saavedra *et al.* [SRB+17], i.e. how much can the growth rates be perturbed (initially all equal to 1) without changing the type of equilibrium  $\mathbf{x}^*$  obtained.

There are many applications of this kind of models in ecology. We could consider a spatial structure that accounts for spatial proximity in the sense that two nearby communities tend to be more strongly connected. For example, in an aquatic environment, we could imagine the existence of an up/down gradient in a water

column. Fig. 10 illustrates a situation where three communities are involved.

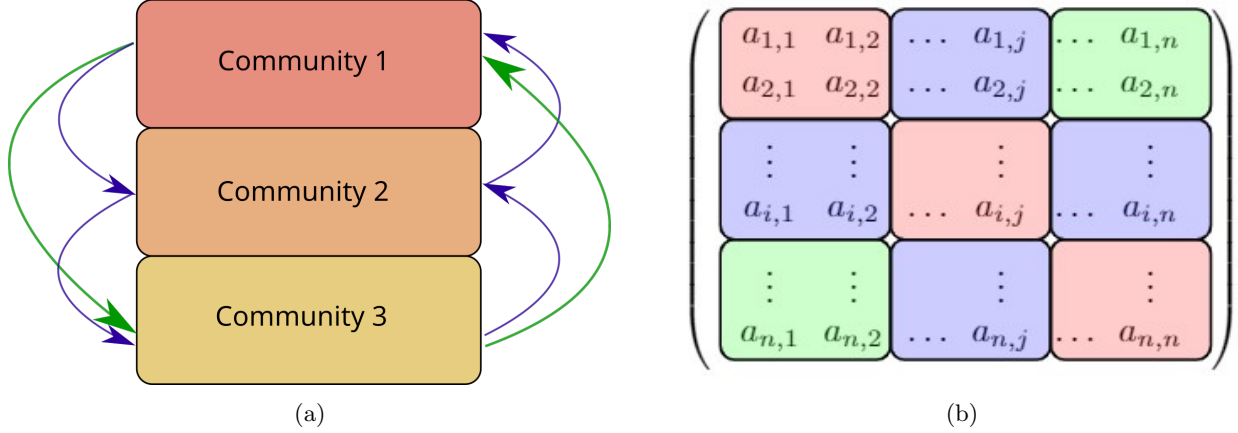


Figure 10: In (a), a representation of the gradient of interaction between three communities in a water column is represented. The blue arrows correspond to strong interaction strength due to their spatial proximity. On the opposite, the communities 1 and 3 are separated, the green arrow represents a weaker interaction. In (b), the block matrix associated with this type of model is displayed. The colors of the blocks corresponds to the colors of the arrows. The red colored block corresponds to intra-community interactions.

Originally introduced by R.T. Paine [Pai66, Pai69], the concept of keystone species is widely used in ecology i.e. one species controls the coexistence of the others and species are lost after the removal of this keystone species. Mouquet *et al.* [MGMC13] suggested extending the concept of keystone species to communities. In the block system, one could analyze the existence of a keystone community that would have disproportionately large effect on other communities. In a metacommunity dynamic, Resetarits *et al.* [RCL18] have explored the concept of keystone communities, where some patches have stronger effect on others.

One could imagine that the same species is present several times in the system, but in different blocks, see Gravel *et al.* [GML16]. In this case, the inter-blocks represent interactions between spatially isolated communities (so should be less strong). If each diagonal or non-diagonal block is a copy of the same interaction pattern (possibly slightly perturbed) and we can add linear effects to the system to represent emigration and immigration, then we could study the feasibility properties of this system. In [GML16], they found that stability is most likely when dispersal (which controls off-diagonal blocks) is intermediate.

Last but not least, it would be relevant to compare the patterns obtained with data in ecology, as in the recent article by Hu *et al.* [HAB<sup>+</sup>22] in the case of a single community.

## References

- [ABC<sup>+</sup>22] I. Akjouj, M. Barbier, M. Clenet, W. Hachem, M. Maïda, F. Massol, J. Najim, and V. C. Tran. Complex systems in Ecology: a guided tour with large Lotka-Volterra models and random matrices, 2022. arXiv:2212.06136.
- [AEK17] O. H. Ajanki, L. Erdős, and T. Krüger. Universality for general Wigner-type matrices. *Probability Theory and Related Fields*, 169(3):667–727, 2017.
- [AEK19] O. Ajanki, L. Erdős, and T. Krüger. Quadratic vector equations on complex upper half-plane. *Memoirs of the American Mathematical Society*, 261(1261), 2019.
- [AGB<sup>+</sup>15] S. Allesina, J. Grilli, G. Barabás, S. Tang, J. Aljadeff, and A. Maritan. Predicting the stability of large structured food webs. *Nature Communications*, 6(1):7842, 2015.
- [AHMN23] I. Akjouj, W. Hachem, M. Maïda, and J. Najim. Equilibria of large random Lotka-Volterra systems with vanishing species: a mathematical approach, 2023. arXiv:2302.07820.
- [AN22] I. Akjouj and J. Najim. Feasibility of sparse large Lotka-Volterra ecosystems. *Journal of Mathematical Biology*, 85(6):66, 2022.
- [AT12] S. Allesina and S. Tang. Stability criteria for complex ecosystems. *Nature*, 483(7388):205–208, 2012.
- [BABL18] M. Barbier, J-F. Arnoldi, G. Bunin, and M. Loreau. Generic assembly patterns in complex ecological communities. *Proceedings of the National Academy of Sciences*, 115(9):2156–2161, 2018.
- [BN21] P. Bizeul and J. Najim. Positive solutions for large random linear systems. *Proceedings of the American Mathematical Society*, 149(6):2333–2348, 2021.
- [BS10] Z. Bai and J. W. Silverstein. *Spectral analysis of large dimensional random matrices*. Springer series in statistics. Springer, New York ; London, 2nd ed edition, 2010.
- [Bun17] G. Bunin. Ecological communities with Lotka-Volterra dynamics. *Physical Review E*, 95(4):042414, 2017.
- [CEFN22] M. Clenet, H. El Ferchichi, and J. Najim. Equilibrium in a large Lotka–Volterra system with pairwise correlated interactions. *Stochastic Processes and their Applications*, 153:423–444, 2022.
- [Cle23] M. Clenet. Impact of a block structure on the lotka-volterra model. <https://github.com/maxime-clenet/Impact-of-a-block-structure-on-the-Lotka-Volterra-model>, 2023.
- [CMN23] M. Clenet, F. Massol, and J. Najim. Equilibrium and surviving species in a large Lotka–Volterra system of differential equations. *Journal of Mathematical Biology*, 87(1):13, 2023.
- [DVR<sup>+</sup>18] M. Dougoud, L. Vinckenbosch, R. P. Rohr, L-F. Bersier, and C. Mazza. The feasibility of equilibria in large ecosystems: A primary but neglected concept in the complexity-stability debate. *PLOS Computational Biology*, 14(2):e1005988, 2018.
- [For10] S. Fortunato. Community detection in graphs. *Physics Reports*, 486(3):75–174, 2010.
- [Gal18] T. Galla. Dynamically evolved community size and stability of random Lotka-Volterra ecosystems. *EPL (Europhysics Letters)*, 123(4):48004, 2018.
- [GAS<sup>+</sup>17] J. Grilli, M. Adorisio, S. Suweis, G. Barabás, J. R. Banavar, S. Allesina, and A. Maritan. Feasibility and coexistence of large ecological communities. *Nature Communications*, 8(1):14389, 2017.



- [GGRA18] T. Gibbs, J. Grilli, T. Rogers, and S. Allesina. Effect of population abundances on the stability of large random ecosystems. *Physical Review E*, 98(2):022410, August 2018.
- [GH82] S. Geman and C. R. Hwang. A chaos hypothesis for some large systems of random equations. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 60(3):291–314, 1982.
- [GJ77] B. S. Goh and L. S. Jennings. Feasibility and stability in randomly assembled Lotka-Volterra models. *Ecological Modelling*, 3(1):63–71, 1977.
- [GML16] D. Gravel, F. Massol, and M. A. Leibold. Stability and complexity in model meta-ecosystems. *Nature Communications*, 7(1):12457, 2016.
- [Goh77] B. S. Goh. Global Stability in Many-Species Systems. *The American Naturalist*, 111(977):135–143, 1977.
- [GRA16] J. Grilli, T. Rogers, and S. Allesina. Modularity and stability in ecological communities. *Nature Communications*, 7(1):12031, 2016.
- [GSSP<sup>+</sup>10] R. Guimerà, D. B. Stouffer, M. Sales-Pardo, E. A. Leicht, M. E. J. Newman, and L. A. N. Amaral. Origin of compartmentalization in food webs. *Ecology*, 91(10):2941–2951, 2010.
- [HAB<sup>+</sup>22] J. Hu, D. R. Amor, M. Barbier, G. Bunin, and J. Gore. Emergent phases of ecological diversity and dynamics mapped in microcosms. *Science*, 378(6615):85–89, 2022.
- [Hac23] W. Hachem. Approximate Message Passing for sparse matrices with application to the equilibria of large ecological Lotka-Volterra systems, 2023. arXiv:2302.09847.
- [HS98] J. Hofbauer and K. Sigmund. *Evolutionary Games and Population Dynamics*. Cambridge University Press, 1998.
- [Jan87] W. Jansen. A permanence theorem for replicator and Lotka-Volterra systems. *Journal of Mathematical Biology*, 25(4):411–422, 1987.
- [Lam19] A. Lamperski. Lemke’s algorithm for linear complementarity problems. <https://github.com/AndyLamperski/lemkelcp>, 2019.
- [LB92] R. Law and J. C. Blackford. Self-Assembling Food Webs: A Global Viewpoint of Coexistence of Species in Lotka-Volterra Communities. *Ecology*, 73(2):567–578, 1992.
- [LCP23] X. Liu, G. W. A. Constable, and J. W. Pitchford. Feasibility and stability in large Lotka Volterra systems with interaction structure. *Physical Review E*, 107(5):054301, 2023.
- [LM96] R. Law and R. D. Morton. Permanence and the Assembly of Ecological Communities. *Ecology*, 77(3):762–775, 1996.
- [Log93] D. O. Logofet. *Matrices and graphs: stability problems in mathematical ecology*. CRC Press, Boca Raton, 1993.
- [Lot25] A. J. Lotka. *Elements of Physical Biology*. Williams and Wilkins Company, 1925.
- [May72] R. M. May. Will a Large Complex System be Stable? *Nature*, 238(5364):413–414, 1972.
- [MGMC13] N. Mouquet, D. Gravel, F. Massol, and V. Calcagno. Extending the concept of keystone species to communities and ecosystems. *Ecology Letters*, 16(1):1–8, 2013.
- [New06] M. E. J. Newman. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, 103(23):8577–8582, 2006.
- [Pai66] R. T. Paine. Food Web Complexity and Species Diversity. *The American Naturalist*, 100(910):65–75, 1966.



- [Pai69] R. T. Paine. The Pisaster-Tegula Interaction: Prey Patches, Predator Food Preference, and Intertidal Community Structure. *Ecology*, 50(6):950–961, 1969.
- [PBG23] L. Poley, J. W. Baron, and T. Galla. Generalized Lotka-Volterra model with hierarchical interactions. *Physical Review E*, 107(2):024313, 2023.
- [Pim79] S. L. Pimm. The structure of food webs. *Theoretical Population Biology*, 16(2):144–158, 1979.
- [PSJ20a] S. Pettersson, V. M. Savage, and M. N. Jacobi. Predicting collapse of complex ecological systems: quantifying the stability–complexity continuum. *Journal of The Royal Society Interface*, 17(166):20190391, 2020.
- [PSJ20b] S. Pettersson, V. M. Savage, and M. N. Jacobi. Stability of ecosystems enhanced by species-interaction constraints. *Physical Review E*, 102(6):062405, 2020.
- [RCL18] E. J. Resetarits, S. E. Cathey, and M. A. Leibold. Testing the keystone community concept: effects of landscape, patch removal, and environment on metacommunity structure. *Ecology*, 99(1):57–67, 2018.
- [SB11] D. B. Stouffer and J. Bascompte. Compartmentalization increases food-web persistence. *Proceedings of the National Academy of Sciences*, 108(9):3648–3652, 2011.
- [SCG<sup>+</sup>18] C. A. Serván, J. A. Capitán, J. Grilli, K. E. Morrison, and S. Allesina. Coexistence of many species in random ecosystems. *Nature Ecology & Evolution*, 2(8):1237–1242, 2018.
- [SRB<sup>+</sup>17] S. Saavedra, R. P. Rohr, J. Bascompte, O. Godoy, N. J. B. Kraft, and J. M. Levine. A structural approach for understanding multispecies coexistence. *Ecological Monographs*, 87(3):470–486, 2017.
- [Sto18] L. Stone. The feasibility and stability of large complex biological networks: a random matrix approach. *Scientific Reports*, 8(1):8246, 2018.
- [TA80] Y. Takeuchi and N. Adachi. The existence of globally stable equilibria of ecosystems of the generalized Volterra type. *Journal of Mathematical Biology*, 10(4):401–415, 1980.
- [Tak96] Y. Takeuchi. *Global dynamical properties of Lotka-Volterra systems*. World Scientific, Singapore ; River Edge, NJ, 1996.
- [TAT78] Y. Takeuchi, N. Adachi, and H. Tokumaru. Global stability of ecosystems of the generalized volterra type. *Mathematical Biosciences*, 42(1):119–136, 1978.
- [Tay88] P. J. Taylor. Consistent scaling and parameter choice for linear and Generalized Lotka-Volterra models used in community ecology. *Journal of Theoretical Biology*, 135(4):543–568, 1988.
- [TF10] E. Thébault and C. Fontaine. Stability of ecological communities and the architecture of mutualistic and trophic networks. *Science*, 329(5993):853–856, 2010.
- [TPA14] S. Tang, S. Pawar, and S. Allesina. Correlation between interaction strengths drives stability in large ecological networks. *Ecology Letters*, 17(9):1094–1100, 2014.
- [TVK10] T. Tao, V. Vu, and M. Krishnapur. Random matrices: Universality of ESDs and the circular law. *The Annals of Probability*, 38(5):2023–2065, 2010.
- [VML04] E. A. Variano, J. H. McCoy, and H. Lipson. Networks, Dynamics, and Modularity. *Physical Review Letters*, 92(18):188701, 2004.
- [Vol26] V. Volterra. Fluctuations in the Abundance of a Species considered Mathematically. *Nature*, 118(2972):558–560, 1926.
- [Vol31] V. Volterra. *Leçons sur la théorie mathématique de la lutte pour la vie*. Gauthier-Villars, 1931.

- [Wan78] P. J. Wangersky. Lotka-Volterra Population Models. *Annual Review of Ecology and Systematics*, 9:189–218, 1978.

## A Stieltjes transform

We provide some reminders regarding Stieltjes transforms, a central element of proofs in random matrix theory. We denote by

$$\mathbb{C}^+ := \{z \in \mathbb{C} : \text{Im}(z) > 0\}$$

the upper half of the complex plane.

**Definition 2** (Stieltjes transform). Let  $\nu \in \mathcal{P}(\mathbb{R})$  be a probability measure. The Stieltjes transform of  $\nu$ , denoted by  $g_\nu : \mathbb{C}^+ \rightarrow \mathbb{C}$ , is defined by

$$g_\nu(z) = \int \frac{1}{\lambda - z} \nu(d\lambda), z \in \mathbb{C}^+.$$

**Remark 7.** Let  $\nu_X$  be the empirical measure of the eigenvalues  $\lambda_1(X), \dots, \lambda_n(X)$  of the symmetric matrix  $X \in \mathcal{M}_n(\mathbb{C})$  define by

$$\nu_X := \frac{1}{n} \sum_{k=1}^n \delta_{\lambda_k(X)}.$$

then the associated Stieltjes transform is given by

$$g_{\nu_X}(z) = \int \frac{1}{\lambda - z} \nu_X(d\lambda) = \frac{1}{n} \sum_{i=1}^n \frac{1}{\lambda_i - z} = \frac{1}{n} \text{Tr}((X - zI)^{-1}),$$

where  $Q = (X - zI)^{-1}$  is the resolvent of the matrix  $X$  and  $\text{Tr}(Q)$  is the trace of matrix  $Q$ .

**Proposition 5** (Stieltjes inversion). *Let  $g_\nu$  the Stieltjes transform of the measure  $\nu$  of finite mass  $\nu(\mathbb{R})$ . If  $a, b \in \mathbb{R}$  and  $\nu(\{a\}) = \nu(\{b\}) = 0$ , then*

$$\nu(a, b) = \frac{1}{\pi} \lim_{y \rightarrow 0^+} \text{Im} \int_a^b g_\nu(x + iy) dx,$$

and

$$\forall x \in \mathbb{R}, \nu(\{x\}) = \frac{1}{\pi} \lim_{y \rightarrow 0^+} \text{Im}(g_\nu(x + iy)).$$

## B Numerical methods

All figures and code are available on Github [Cle23]. The code is written in Python.

To verify the system of equations of heuristics 1, the simulations on the properties of surviving species are performed in two distinct methods (see Fig. 3). On the one hand, we use a standard solver (cf. `scipy.optimize`) to find the theoretical solutions by finding a local minimum of the system of equations (a modification of the Powell hybrid method). On the other hand, we simulate a large number of matrix  $B$ , each corresponding to an experiment, and we resolve the associated LCP problem using the Lemke's algorithm (see the `lemkelcp` package [Lam19]). The empirical solutions are computed using a Monte Carlo experiment, i.e. we use the LCP solution to compute the properties of the surviving species and we make an average over the ensemble of experiments. As a baseline, the dynamics of Lotka-Volterra are achieved by a Runge-Kutta method of order 4 (RK4) implemented in the code.

**Spectrum: a computer based approach.** Theorem 3 only provides sufficient conditions for the existence of a unique stable equilibrium and is based on the rough asymptotic upper bound estimation  $\Sigma = 2 \|S\|_\infty^{1/2}$ . We can assess the sharpness of this bound by comparing it to the limiting spectrum of matrix  $H$ , which can be plotted via numerical simulations. An efficient way to compute numerically the spectrum of the matrix  $H$  comes from the system of non linear equations (11).

Starting from the QVE (11) associated to the matrix  $H$ , the system takes the simpler form

$$\begin{cases} -\frac{1}{\mu(z)} &= z + 2\beta_1 s_{11}^2 \mu(z) + \beta_2 (s_{12}^2 + s_{21}^2) \nu(z) \\ -\frac{1}{\nu(z)} &= z + \beta_1 (s_{12}^2 + s_{21}^2) \mu(z) + 2\beta_2 s_{22}^2 \nu(z) \end{cases},$$

where  $\forall k \in \mathcal{I}_1$ ,  $m_k(z) = \mu(z)$  and  $\forall k \in \mathcal{I}_2$ ,  $m_k(z) = \nu(z)$ . All the knowledge of equation (11) relies on the functions  $\mu(z)$  and  $\nu(z)$ . Then, using RMT theory, the resolvent  $G$  of the symmetric matrix  $H$  can be approximated by

$$G(z) = (H - zI)^{-1} \simeq \text{diag}(\mu(z)\mathbf{1}_{\mathcal{I}_1}^\top, \nu(z)\mathbf{1}_{\mathcal{I}_2}^\top).$$

From Remark 7, the trace of the resolvent is equal to the Stieltjes transform

$$g(z) = \frac{1}{n} \text{Tr}(G) \simeq \beta_1 \mu(z) + \beta_2 \nu(z)$$

of the spectral measure. Finally, the spectral density can be obtained using the Stieltjes inversion (Prop. 5). The spectral density of the matrix  $H$  can be computed numerically by an iterative scheme. The initial condition of the two measurements  $(\mu, \nu)$  is  $\mu_0 = \nu_0 = -\frac{1}{z}$ . Then, the iterative scheme

$$\begin{cases} -\frac{1}{\mu_p} &= z + 2\beta_1 s_{11}^2 \mu_{p-1} + \beta_2 (s_{12}^2 + s_{21}^2) \nu_{p-1} \\ -\frac{1}{\nu_p} &= z + \beta_1 (s_{12}^2 + s_{21}^2) \mu_{p-1} + 2\beta_2 s_{22}^2 \nu_{p-1} \end{cases},$$

converge to  $\mu_\infty = \lim_{p \rightarrow +\infty} \mu_p$  and  $\nu_\infty = \lim_{p \rightarrow +\infty} \nu_p$ . The last step consist of using the property of the Stieltjes inversion (Prop. 5).

**Remark 8.** To handle the Stieltjes inversion (Prop. 5) numerically, it is similar as starting with  $z = x + \epsilon i$ ,  $\epsilon \approx 10^{-3}$ .

In Fig. 11, we present the numerical estimation of the spectral density for different types of interactions of the matrix  $H$ .

## C Remaining computations

### C.1 Moments of $\check{Z}_k$

We compute hereafter the conditional variance of  $\check{Z}_k = (B\mathbf{x}^*)_k$  with respect to  $\mathbf{x}^*$ . We rely on the following identities  $\forall k \in \mathcal{I}_i$ ,  $\forall \ell \in \mathcal{I}_j$  and  $\forall o \in \mathcal{I}_q$ :

$$\mathbb{E}B_{k\ell} = 0 \quad , \quad \mathbb{E}(B_{k\ell}^2) = \frac{s_{ij}^2}{n} \quad , \quad \mathbb{E}B_{k\ell}B_{ko} = 0 \quad (\ell \neq o).$$

We first compute the conditional mean:

$$\forall k \in \mathcal{I}_i, \quad \mathbb{E}_{\mathbf{x}^*}(\check{Z}_k) = \sum_{\ell \in [n]} \mathbb{E}(B_{k\ell})x_\ell^* = \sum_{\ell \in \mathcal{S}_1} \mathbb{E}(B_{k\ell})x_\ell^* + \sum_{\ell \in \mathcal{S}_2} \mathbb{E}(B_{k\ell})x_\ell^* = 0.$$

We now compute the second moment:

$$\begin{aligned} \forall k \in \mathcal{I}_i, \quad \mathbb{E}_{\mathbf{x}^*}(\check{Z}_k^2) &= \mathbb{E}_{\mathbf{x}^*} \left( \sum_{\ell \in [n]} B_{k\ell} x_\ell^* \right)^2 = \mathbb{E}_{\mathbf{x}^*} \sum_{\ell, o \in [n]} B_{k\ell} B_{ko} x_\ell^* x_o^*, \\ &= \sum_{\ell \in \mathcal{S}_1 \cup \mathcal{S}_2} \mathbb{E}(B_{k\ell}^2) x_\ell^{*2} + \sum_{\ell \neq o} \mathbb{E}(B_{k\ell} B_{ko}) x_\ell^* x_o^*, \\ &= \frac{s_{i1}^2}{n} \sum_{\ell \in \mathcal{S}_1} x_\ell^{*2} + \frac{s_{i2}^2}{n} \sum_{\ell \in \mathcal{S}_2} x_\ell^{*2}, \\ &\stackrel{(a)}{\simeq} \beta_1 \hat{p}_1 \hat{\sigma}_1^2 s_{i1}^2 + \beta_2 \hat{p}_2 \hat{\sigma}_2^2 s_{i2}^2, \end{aligned}$$

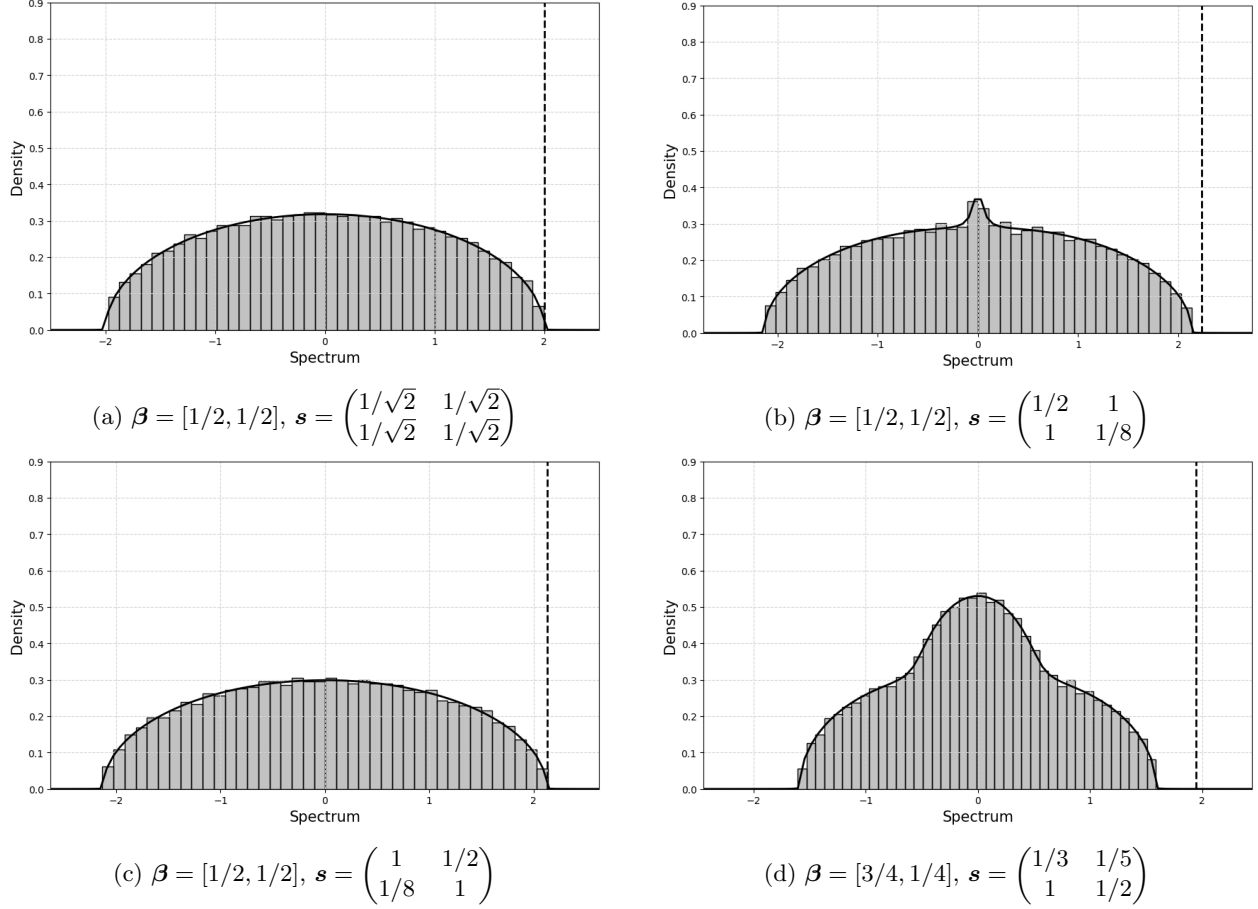


Figure 11: Spectrum (histogram) of the Hermitian random matrix  $H$  ( $n = 1000$ ), conditions on  $(\beta, \mathbf{s})$  are given in each panel. The numerical approach is used to compute the solid line spectrum distribution. An upper bound for the largest eigenvalue of  $H$ , given by  $2\|S\|_\infty^{1/2}$ , is denoted by the dashed vertical line.

where the approximation in (a) follows from the fact that

$$\frac{1}{n^2} \sum_{\ell, o \in [n]} x_\ell^* x_o^* = \frac{1}{n^2} \sum_{\ell \neq o} x_\ell^* x_o^* + \mathcal{O}\left(\frac{1}{n}\right).$$

We can now compute the variance:

$$\forall k \in \mathcal{I}_i, \text{Var}_{\mathbf{x}^*}(\check{Z}_k) = \mathbb{E}_{\mathbf{x}^*}(\check{Z}_k^2) - (\mathbb{E}_{\mathbf{x}^*}\check{Z}_k)^2 = \beta_1 \hat{p}_1 \hat{\sigma}_1^2 s_{i1}^2 + \beta_2 \hat{p}_2 \hat{\sigma}_2^2 s_{i2}^2.$$

## C.2 Details of heuristics of the mean

Our starting point is the following generic representation of an abundance at equilibrium (either of a surviving or vanishing species) in the case  $k \in \mathcal{S}_i$ :

$$x_k^* = (1 + \Delta_i^* Z_k) \mathbf{1}_{\{Z_k > \delta_i^*\}} = \mathbf{1}_{\{Z_k > \delta_i^*\}} + (\Delta_i^* Z_k) \mathbf{1}_{\{Z_k > \delta_i^*\}}.$$

Summing over  $\mathcal{S}_i$  and normalizing,

$$\begin{aligned} \frac{1}{|\mathcal{S}_i|} \sum_{k \in \mathcal{S}_i} x_k^* &= \frac{1}{|\mathcal{S}_i|} \sum_{k \in \mathcal{S}_i} \mathbf{1}_{\{Z_k > \delta_i^*\}} + \Delta_i^* \frac{1}{|\mathcal{S}_i|} \sum_{k \in \mathcal{S}_i} Z_k \mathbf{1}_{\{Z_k > \delta_i^*\}}, \\ \hat{m}_i &\stackrel{(a)}{=} 1 + \Delta_i^* \frac{|\mathcal{I}_i|}{|\mathcal{S}_i|} \frac{1}{|\mathcal{I}_i|} \sum_{k \in \mathcal{I}_i} Z_k \mathbf{1}_{\{Z_k > \delta_i^*\}}, \\ \hat{m}_i &\stackrel{(b)}{\simeq} 1 + \Delta_i^* \frac{1}{\mathbb{P}(Z > \delta_i^*)} \mathbb{E}(Z \mathbf{1}_{\{Z > \delta_i^*\}}), \\ \hat{m}_i &\simeq 1 + \Delta_i^* \mathbb{E}(Z \mid Z > \delta_i^*). \end{aligned}$$

where (a) follows from the fact that  $|\mathcal{S}_i| = \sum_{k \in \mathcal{S}_i} \mathbf{1}_{\{Z_k > \delta_i^*\}}$  (by definition of  $\mathcal{S}_i$ ), (b) from the law of large numbers  $\frac{1}{|\mathcal{I}_i|} \sum_{k \in \mathcal{I}_i} Z_k \mathbf{1}_{\{Z_k > \delta_i^*\}} \xrightarrow[n \rightarrow \infty]{} \mathbb{E} Z \mathbf{1}_{\{Z > \delta_i^*\}}$  and  $\frac{|\mathcal{S}_i|}{|\mathcal{I}_i|} \xrightarrow[n \rightarrow \infty]{} \mathbb{P}(Z > \delta_i^*)$  with  $Z \sim \mathcal{N}(0, 1)$ . It remains to replace  $\hat{m}_i$  by its limit  $m_i^*$  to obtain the heuristics of the mean:

$$\begin{aligned} m_1^* &= 1 + \Delta_1^* \mathbb{E}(Z \mid Z > \delta_1^*), \\ m_2^* &= 1 + \Delta_2^* \mathbb{E}(Z \mid Z > \delta_2^*). \end{aligned}$$

### C.3 Density of the distribution of the surviving species.

Assume that  $x^* > 0$ , and let  $f = \mathbb{R} \rightarrow \mathbb{R}$  be a bounded continuous test function, then  $\forall k \in \mathcal{S}_i$

$$\begin{aligned} \mathbb{E} f(x_k^*) &= \mathbb{E} \left[ f(1 + \lambda_i^* + \Delta_i^* Z_k) \mid Z_k > \delta_i^* \right], \\ &= \int_{-\infty}^{\infty} f(1 + \lambda_i^* + \Delta_i^* u) \frac{\mathbf{1}_{\{u > \delta_i^*\}}}{1 - \Phi(\delta_i^*)} \frac{e^{-\frac{u^2}{2}}}{\sqrt{2\pi}} du, \\ &= \int_0^{\infty} f(y) e^{-\frac{1}{2} \left( \frac{y}{\Delta_i^*} + \delta_i^* \right)^2} \frac{1}{\sqrt{2\pi} \Phi(-\delta_i^*) \Delta_i^*} dy, \end{aligned}$$

hence the density of  $x_k^*$ ,  $\forall k \in \mathcal{S}_i$ .

## D Sketch of proof of Theorem 4

The first step consists in decomposing the equilibrium  $\mathbf{x}^*$ :

$$\begin{aligned} x_k^* &= e_k^\top \mathbf{x}^* = e_k^\top (I - B)^{-1} \mathbf{1} = \sum_{\ell=0}^{\infty} e_k^\top B^\ell \mathbf{1} = 1 + e_k^\top B \mathbf{1} + e_k^\top B^2 (I - B)^{-1} \mathbf{1}, \\ &= 1 + Z_k + R_k, \end{aligned}$$

where  $Z_k = \sum_{\ell=1}^n B_{k\ell}$ ,  $\forall k \in [n]$ .

One can prove that  $\forall k \in [n]$ ,  $R_k$  is a negligible term if  $n$  is sufficiently large. From a technical point, it relies on Gaussian concentration of Lipschitz functionals and we are confident that the techniques applied in [BN21] will succeed in handling  $R_k$ . However, this part of the proof is not been treated here since we want to stick a concise argumentation of the proof which gives the reader information about the critical bound of the feasibility threshold.

The feasibility of the two communities is studied independently. Using Gaussian addition properties, a simpler form of  $Z_k$  is first deduced. Consider a family  $(\tilde{Z}_k)_{k \in [n]}$  of i.i.d. random variables  $\mathcal{N}(0, 1)$ .

$$\begin{aligned} \text{If } k \in \mathcal{I}_1, \quad Z_k &= \sum_{\ell \in \mathcal{I}_1} B_{k\ell} + \sum_{\ell \in \mathcal{I}_2} B_{k\ell}, \\ &\sim \mathcal{N}(0, \beta_1 s_{11}^2) + \mathcal{N}(0, \beta_2 s_{12}^2), \\ &\sim \sqrt{\beta_1 s_{11}^2 + \beta_2 s_{12}^2} \tilde{Z}_k. \end{aligned}$$

Similarly

$$\text{if } k \in \mathcal{I}_2, Z_k \sim \sqrt{\beta_1 s_{21}^2 + \beta_2 s_{22}^2} \check{Z}_k.$$

Given  $\beta = (\beta_1, \beta_2)$ , conditions on the matrix  $\mathbf{s}$  are inferred to have:

$$\mathbb{P}\left(\min_{k \in [n]} x_k > 0\right) = 1 \quad \Leftrightarrow \quad \mathbb{P}\left(\min_{k \in [n]} Z_k > -1\right) = 1. \quad (26)$$

In order to compute a tractable form of  $\min_{k \in [n]} Z_k$ , an additional approximation is made, if  $n$  is large enough

$$\forall i \in \{1, 2\}, \quad \min_{k \in \mathcal{I}_i} \check{Z}_k \sim -\sqrt{2 \log(\beta_i n)} \approx -\sqrt{2 \log(n)} \quad (27)$$

$$\begin{aligned} \min_{k \in [n]} Z_k &= \min \left( \sqrt{\beta_1 s_{11}^2 + \beta_2 s_{12}^2} \min_{k \in \mathcal{I}_1} \check{Z}_k, \sqrt{\beta_1 s_{21}^2 + \beta_2 s_{22}^2} \min_{k \in \mathcal{I}_2} \check{Z}_k \right), \\ &\stackrel{(a)}{\simeq} \min \left( \sqrt{\beta_1 s_{11}^2 + \beta_2 s_{12}^2} \left( -\sqrt{2 \log(n)} \right), \sqrt{\beta_1 s_{21}^2 + \beta_2 s_{22}^2} \left( -\sqrt{2 \log(n)} \right) \right), \\ &= \min \left( -\sqrt{2\beta_1 s_{11}^2 \log(n) + 2\beta_2 s_{12}^2 \log(n)}, -\sqrt{2\beta_1 s_{21}^2 \log(n) + 2\beta_2 s_{22}^2 \log(n)} \right), \\ &= -\max \left( \sqrt{2\beta_1 s_{11}^2 \log(n) + 2\beta_2 s_{12}^2 \log(n)}, \sqrt{2\beta_1 s_{21}^2 \log(n) + 2\beta_2 s_{22}^2 \log(n)} \right). \end{aligned}$$

where (a) comes from the approximation (27). The condition  $\min_{k \in [n]} Z_k > -1$  asymptotically boils down to

$$\begin{aligned} &\max \left( \sqrt{2\beta_1 s_{11}^2 \log(n) + 2\beta_2 s_{12}^2 \log(n)}, \sqrt{2\beta_1 s_{21}^2 \log(n) + 2\beta_2 s_{22}^2 \log(n)} \right) < 1, \\ \Leftrightarrow &\max \left( 2\beta_1 s_{11}^2 \log(n) + 2\beta_2 s_{12}^2 \log(n), 2\beta_1 s_{21}^2 \log(n) + 2\beta_2 s_{22}^2 \log(n) \right) < 1, \\ \Leftrightarrow &\max \left( \beta_1 s_{11}^2 + \beta_2 s_{12}^2, \beta_1 s_{21}^2 + \beta_2 s_{22}^2 \right) < \frac{1}{2 \log(n)}, \\ \Leftrightarrow &\left\| (\mathbf{s}_n \circ \mathbf{s}_n) \boldsymbol{\beta}^\top \right\|_\infty < \frac{1}{2 \log(n)} := (s_n^*)^2. \end{aligned}$$

## E Extension to the $b$ -blocks model

### E.1 Interaction matrix with $b$ communities

Within the framework of  $b$  communities, the matrix  $B = (B_{k\ell})_{n,n}$  is defined as

$$B = V \mathbf{s} V^\top \circ \frac{1}{\sqrt{n}} A, \quad (28)$$

where

$$V \in \mathcal{M}_{n \times b}, V = \begin{pmatrix} \mathbf{1}_{\mathcal{I}_1} & 0 & \cdots & 0 \\ 0 & \mathbf{1}_{\mathcal{I}_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{1}_{\mathcal{I}_b} \end{pmatrix}, A = \begin{pmatrix} A_{11} & \cdots & A_{1b} \\ \vdots & \ddots & \vdots \\ A_{b1} & \cdots & A_{bb} \end{pmatrix}, \mathbf{s} = \begin{pmatrix} s_{11} & \cdots & s_{1b} \\ \vdots & \ddots & \vdots \\ s_{b1} & \cdots & s_{bb} \end{pmatrix},$$

where  $\mathcal{I}_1 = [n_1]$ ,  $\mathcal{I}_2 = \{n_1 + 1, \dots, n_1 + n_2\}, \dots, \mathcal{I}_b = \{n_1 + \dots + n_{b-1} + 1, \dots, n\}$  the subset of  $[n]$  of size  $|\mathcal{I}_i| := n_i$  matching the index of species belonging to community  $i$  and  $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_b)$  with  $\forall i \in [b]$ ,  $\beta_i = n_i/n$  and  $\sum_{i=1}^b \beta_i = 1$ . The random matrix  $A_{ij}$  is non-Hermitian of size  $|\mathcal{I}_i| \times |\mathcal{I}_j|$  with standard Gaussian entries i.e.  $\mathcal{N}(0, 1)$ . Recall that  $\mathbf{1}_{\mathcal{I}_i}$  be the element-wise vector of 1 with size  $|\mathcal{I}_i|$ .

## E.2 Existence of a unique equilibrium

Let  $H$  be the symmetric matrix

$$H = B + B^\top = \frac{1}{\sqrt{n}} \begin{pmatrix} H_{11} & \cdots & H_{1b} \\ \vdots & \ddots & \vdots \\ H_{b1} & \cdots & H_{bb} \end{pmatrix},$$

where  $\forall i, j \in [b]$ ,  $H_{ij}$  is a matrix of size  $|\mathcal{I}_i| \times |\mathcal{I}_j|$  and each off-diagonal entries follow a Gaussian distribution  $\mathcal{N}(0, s_{ij}^2 + s_{ji}^2)$ . The QVE associated to the matrix  $H$  is decomposed as

$$k \in \mathcal{I}_i, -\frac{1}{m_k(z)} = z + \sum_{j=1}^b \sum_{\ell \in \mathcal{I}_j} \frac{1}{n} (s_{ij}^2 + s_{ji}^2) m_\ell(z).$$

Given  $\mathbf{m}(z) = (m_1(z), \dots, m_n(z))$ , denote by  $1/\mathbf{m}(z) = (1/m_1(z), \dots, 1/m_n(z))$  and  $S = \frac{1}{n} V(\mathbf{s} + \mathbf{s}^\top) V^\top$  the QVE can be written in the standard form

$$-\frac{1}{\mathbf{m}(z)} = z + S\mathbf{m}(z). \quad (29)$$

Following the same arguments as in Theorem 3 and relying on Theorem 2, given the particular shape of the matrix  $S$ , computing its norm is equivalent to computing the norm of a matrix of size  $b$

$$\|S\|_\infty = \|\text{diag}(\boldsymbol{\beta}) ((\mathbf{s} \circ \mathbf{s}) + (\mathbf{s} \circ \mathbf{s})^\top)\|_\infty.$$

## E.3 Surviving species

**Heuristics 3.** Let  $\mathbf{s}$  be the  $b \times b$  matrix of interaction strengths and assume that the condition of Theorem 3 holds, then the following system of  $2b$  equations and  $2b$  unknowns  $\mathbf{p} = (p_1, p_2, \dots, p_b)$ ,  $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \dots, \sigma_b)$

$$\begin{aligned} \forall i \in [b], p_i &= 1 - \Phi(\delta_i), \\ \forall i \in [b], (\sigma_i)^2 &= 1 + 2\Delta_i \mathbb{E}(Z|Z > \delta_i) + \Delta_i^2 \mathbb{E}(Z^2|Z > \delta_i), \end{aligned}$$

where

$$\Delta_i = \sqrt{\sum_{j=1}^b p_j (\sigma_j)^2 \beta_j s_{ij}^2}; \delta_i = -\frac{1}{\Delta_i},$$

admits a unique solution  $(\mathbf{p}^*, \boldsymbol{\sigma}^*)$  and  $\forall i \in [b]$

$$\hat{p}_i \xrightarrow[n \rightarrow \infty]{a.s.} p_i^* \quad \text{and} \quad \hat{\sigma}_i \xrightarrow[n \rightarrow \infty]{a.s.} \sigma_i^*.$$

## E.4 Distribution of the surviving species

Let  $\mathbf{s}$  be the  $b \times b$  matrix of interaction strengths and assume that the condition of Theorem 3 holds. Let  $\mathbf{x}^*$  the solution of (8) and  $(\mathbf{p}^*, \boldsymbol{\sigma}^*)$  the solution of the heuristic 3. Recall the definition of  $\Delta_i$ ,  $\delta_i$  and denote by  $\delta_i^* = \delta_i(p_i^*, \sigma_i^*)$ . Let  $x_k^* > 0$  a positive component of  $\mathbf{x}^*$  belonging to the community  $i$ , then:

$$\mathcal{L}(x_k^*) \xrightarrow[n \rightarrow \infty]{} \mathcal{L} \left( 1 + \Delta_i^* Z \mid Z > \delta_i^* \right),$$

where  $Z \sim \mathcal{N}(0, 1)$ . Otherwise stated, asymptotically  $\forall k \in \mathcal{S}_i$ ,  $x_k^*$  admits the following density

$$f_k(y) = \frac{\mathbf{1}_{\{y>0\}}}{\Phi(-\delta_i^*)} \frac{1}{\Delta_i^* \sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left( \frac{y}{\Delta_i^*} + \delta_i^* \right)^2 \right\}. \quad (30)$$



## E.5 Feasibility

We consider a growing scaling matrix

$$\mathbf{s}_n \xrightarrow[n \rightarrow \infty]{} \mathbf{0} \quad \Leftrightarrow \quad \forall i, j \in \{1, b\}, s_{ij} \xrightarrow[n \rightarrow \infty]{} 0.$$

Let  $B_n$  a matrix defined by

$$B_n = V \mathbf{s}_n V^\top \circ \frac{1}{\sqrt{n}} A. \quad (31)$$

The spectral radius of  $\frac{1}{\sqrt{n}} A$  a.s. converges to 1 (circular law). So as long as  $\mathbf{s}_n$  is close to zero, the matrix  $I - B_n$  is eventually invertible.

Recall the problem which admits a unique solution defined by

$$\mathbf{x}^* = \mathbf{1} + B \mathbf{x}^* \Leftrightarrow \mathbf{x}^* = (I - B)^{-1} \mathbf{1}, \quad (32)$$

**Theorem 6** (Co-feasibility for the  $b$ -blocks model). *Assume that matrix  $B_n$  is defined by the  $b$ -blocks model (31). Let  $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_b)$ ,  $\sum_{i=1}^b \beta_i = 1$  represents the proportion of each community. Let  $\mathbf{s}_n \xrightarrow[n \rightarrow \infty]{} \mathbf{0}$  and denote by  $s_n^* = 1/\sqrt{2 \log n}$ . Let  $x_n = (x_k)_{k \in [n]}$  be the solution of (32).*

1. *If there exists  $\varepsilon > 0$  such that eventually  $\left\| (\mathbf{s}_n \circ \mathbf{s}_n) \boldsymbol{\beta}^\top \right\|_\infty \geq (1 + \varepsilon) (s_n^*)^2$  then*

$$\mathbb{P} \left\{ \min_{k \in [n]} x_k > 0 \right\} \xrightarrow[n \rightarrow \infty]{} 0.$$

2. *If there exists  $\varepsilon > 0$  such that eventually  $\left\| (\mathbf{s}_n \circ \mathbf{s}_n) \boldsymbol{\beta}^\top \right\|_\infty \leq (1 - \varepsilon) (s_n^*)^2$  then*

$$\mathbb{P} \left\{ \min_{k \in [n]} x_k > 0 \right\} \xrightarrow[n \rightarrow \infty]{} 1.$$

*Sketch of proof.* Starting from the decomposition, the equilibrium  $\mathbf{x}^*$ :

$$x_k^* = 1 + Z_k + R_k,$$

where  $Z_k = \sum_{\ell=1}^n B_{k\ell}$ ,  $\forall k \in [n]$  and we assume that  $\forall k \in [n]$ ,  $R_k$  is a negligible term if  $n$  is sufficiently large.

The feasibility of the  $b$  communities is studied independently. Using Gaussian addition properties, a simpler form of  $Z_k$  is derived. Consider a family  $(\check{Z}_k)_{k \in [n]}$  of i.i.d. random variables  $\mathcal{N}(0, 1)$ .

$$\begin{aligned} \text{If } k \in \mathcal{I}_i, \quad Z_k &= \sum_{j=1}^b \sum_{\ell \in \mathcal{I}_j} B_{k\ell}, \\ &\sim \sum_{j=1}^b \mathcal{N}(0, \beta_j s_{ij}^2), \\ &\sim \sqrt{\sum_{j=1}^b \beta_j s_{ij}^2} \check{Z}_k. \end{aligned}$$

Given  $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_b)$ , conditions on the matrix  $\mathbf{s}$  are inferred to have

$$\mathbb{P}(\min_{k \in [n]} x_k > 0) = 1 \Leftrightarrow \mathbb{P}(\min_{k \in [n]} Z_k > -1) = 1.$$

In order to compute a tractable form of  $\min_{k \in [n]} Z_k$ , an additional approximation is made, if  $n$  is large enough

$$\min_{k \in \mathcal{I}_i} \check{Z}_k \sim -\sqrt{2 \log(\beta_i n)} \simeq -\sqrt{2 \log(n)}. \quad (33)$$

$$\begin{aligned} \min_{k \in [n]} Z_k &= \min_{i \in [b]} \left( \sqrt{\sum_{j=1}^b \beta_j s_{ij}^2} \min_{k \in \mathcal{I}_i} \check{Z}_k \right), \\ &\simeq \min_{i \in [b]} \left( \sqrt{\sum_{j=1}^b \beta_j s_{ij}^2} \left( -\sqrt{2 \log(n)} \right) \right), \\ &= \min_{i \in [b]} \left( -\sqrt{\sum_{j=1}^b 2\beta_j s_{ij}^2 \log(n)} \right), \\ &= -\max_{i \in [b]} \left( \sqrt{\sum_{j=1}^b 2\beta_j s_{ij}^2 \log(n)} \right). \end{aligned}$$

Following the approximation (33), the condition  $\min_{k \in [n]} Z_k > -1$  asymptotically boils down to

$$\begin{aligned} &\max_{i \in [b]} \left( \sqrt{\sum_{j=1}^b 2\beta_j s_{ij}^2 \log(n)} \right) < 1, \\ \Leftrightarrow &\max_{i \in [b]} \left( \sum_{j=1}^b 2\beta_j s_{ij}^2 \log(n) \right) < 1, \\ \Leftrightarrow &\max_{i \in [b]} \left( \sum_{j=1}^b \beta_j s_{ij}^2 \right) < \frac{1}{2 \log(n)}, \\ \Leftrightarrow &\left\| (\mathbf{s}_n \circ \mathbf{s}_n)^2 \boldsymbol{\beta}^\top \right\|_\infty < \frac{1}{2 \log(n)} := (s_n^*)^2. \end{aligned}$$

□