



FEDERAL UNIVERSITY OF BAHIA

PHYSICS INSTITUTE

PHYSICS POST-GRADUATE PROGRAM

Exploring Ecological Interactions Using the Generalized Lotka-Volterra Model: Coexistence and Resilience of Populations

Supervisor:

Prof. Dra. Suanie Tavares Rubim de Pinho

Author:

Rafael Menezes dos Santos

Co-Supervisors:

Prof. Dra. Flora Souza Bacelar

Prof. Dr. Pedro Milet Meirelles

A dissertation submitted for the degree of

MSc Physics

January 21, 2021

*Para (mãe) Natureza,
que o conhecimento, fala e ação, sejam amor.*

*Para o meu povo,
que o conhecimento, fala e ação, se transformem em profunda liberdade.*

Agradecimentos / Acknowledgements

Agradeço imensamente às milhões de pessoas que, invisíveis, permitiram que me dedicasse à ciência através de seus trabalhos, sonhos e lutas.

Agradeço o apoio incondicional da família que me fez: A Shini, pelo carinho, cuidado e parceria; A Joseh, pelo suporte e os horizontes vastíssimos; A Dani, guerreiro sensível, pelos aprendizados e imenso companherismo em mundos reais e imaginados; A Tobe, Jaben, Zelda e Gaia, que insistem em um dia me ensinar que, na verdade, viver e amar são coisas simples.

Agradeço a vó Marlene, pela eterna doçura. A meus tios e tia, primas e primos por muitas vezes também sermos a família que escolhemos, na luta e na paz, com amor. Agradeço *in memoriam* a meus avôs Rona e Sevê por deixarem os tesouros de tantas viagens comigo. Agradeço *in memoriam* a vó Margô, que vai sorrir ao lado de Santo Antônio ao saber de mais um passo dado (apesar do cabelo).

Agradeço a amizade e presença de amigues. Vocês, essas pessoas maravilhosas, fazem a vida ter mais gosto e alegrias. Obrigado pela paciência com um semi-eremita e (pós pandemia) vamos festejar minha (nossa) volta ao mundo dos que vivem! Às amizades multiétnicas, ao Breakdown (tô muito feliz), ao almoço de Cacha Pregos, ao Delta Force.

Agradeço o apoio financeiro da CAPES através da bolsa de mestrado recebida. Agradeço à comunidade do Instituto de Física da UFBA pela acolhida e convivência maravilhosa. Espero ter podido contribuir com o sonho de um instituto inclusivo, politicamente ativo e de excelência. Aos amigos e colegas (discentes ou não), muito obrigado pelos papos, trocas e festas. Ao corpo docente, meu profundo agradecimento por pavimentarem o (não trivial) caminho das abstrações, conceitos, práticas e técnicas - saibam de minha grande admiração. Aos terceirizados e servidores, também parceiros, obrigado pelos ensinamentos e apoio diurno. Agradeço em particular às queridas pessoas parceiras de DA e da vida. A Luan Orion (Magistralis), pela amizade, convivência suave durante dois anos de apartamento 103 e pela parceria de Kerbal à Seabra!

Agradeço às queridas pessoas que são, para mim, Seabra. Pelos aprendizados intensos com cada estudante (e que os sonhos contagiem!). Aos colegas que compartilham o sonho possível de se fazer educação, em particular João Bina, Marcão, Nivaldo, Ronízia, Janaína, Thiago, Adelson (1.e4 c5), Teresa, Drica e André.

Às pessoas que fizeram parte da minha vida e que a vida levou, obrigado por tudo que ficou.

Agradeço todas as professoras e professores que me ajudaram a aprender (e aprender a aprender). I thank everyone that is dedicated to removing all the barriers in the way of science. I am profoundly grateful to prof. Jan Hendriks for the keen guidance and kind encouragement - the interdisciplinarity and overarching principles he showed made Ecology much closer to my mind and heart. Aos professores Roberto Kraenkel e Paulo Inácio, agradeço a oportunidade ímpar de aprendizado e vivência que é a escola de matemática biológica. Ao prof. José Garcia, pelo entusiasmo cativante em revelar um mundo complexo. À profa. Flora Bacelar, pelo exemplo, guiança e apoio em tantas pesquisas dinâmicas. Ao prof. Pedro Meirelles pela parceria, e maravilhosa acolhida em seu laboratório. À profa. Suani Pinho, pela enorme paciência em construir junto, passo a passo, uma dissertação e um mestre em física (mesmo diante de seu compromisso pétreo em evitar o fim termodinâmico da Universidade).

Agradeço a Amanda, pelo carinho, pela parceria, pela convivência. Por nosso encontro multidimensional. Sem seu sonhar junto, a existência dessa dissertação seria só um sonho que se sonha só.

Agradeço ao silêncio imóvel dentro de mim, pela paciência, pelo acolhimento, pela inspiração.

*Politics without principles,
Education without character,
Science without humanity,
Commerce without morality —
Are not only useless but positively dangerous*
— Sri Sathya Sai Baba

Abstract

To understand ecological communities, we must understand the complex dynamical patterns that emerge from population interactions. Indeed, ecological communities are non-randomly assembled, and the interactions between populations have distinct architectures or topologies. Through this structure of interactions, populations and species influence each other's growth in an intricate manner. These dynamics unveil themselves in community-wide emergent properties, such as community stability. During the last century, many central advancements in ecology stemmed from mathematical methods based on dynamical systems theory. More recently, the representation of ecological communities as complex networks, such as food webs and competitive or mutualistic networks, provided additional insights into the ecology of communities. Several authors contributed to our current knowledge on the interplay between the structure and dynamics of complex ecosystems. Despite the extensive developments in this area, many questions are not yet settled. Notably, the distribution of each type of dynamics (attractors) in ecological models was seldomly studied in favor of the assumption of equilibrium. Additionally, an approach exploring the ubiquity of known results linking community stability to structure-related parameters in the parameter space is of great importance. In this study, we aim at contributing to this endeavor by performing simulations based on the generalized Lotka-Volterra model (gLV) across a wide parameter space range. In accordance, our research objective is to present an overview of the gLV asymptotic dynamics, coexistence, and resilience by performing a comprehensive computational exploration integrating variability in type, intensity, and distribution of interspecific ecological interactions. We found that different dynamics of ecological communities can be described by specific parametrizations of the gLV, despite a large prevalence of parameter combinations leading to unlimited growth of some populations in the community. We were also able to detect known patterns, such as the negative impact of cooperation upon the resilience of communities and the occurrence of competitive exclusion in speciose communities. Significantly, our research contributes to the longstanding question of the interconnections between the structure and stability of ecological communities.

Resumo

Para entender as comunidades ecológicas, é essencial que entendamos os padrões dinâmicos complexos que emergem das interações entre as populações. De fato, as interações entre as populações em uma comunidade ecológica possuem arquiteturas ou topologias singulares. Por meio dessa estrutura de interações, as populações e as espécies influenciam o crescimento umas das outras de maneira intrincada. Essa dinâmica se revela em propriedades emergentes ao nível de comunidade, como a estabilidade. Durante o último século, muitos avanços centrais na ecologia resultaram de métodos matemáticos baseados na teoria dos sistemas dinâmicos. Mais recentemente, a representação de comunidades ecológicas como redes complexas, a exemplo das teias alimentares e redes competitivas ou mutualísticas, permitiu novas descobertas acerca da ecologia das comunidades. Vários autores contribuíram para o nosso conhecimento atual sobre a interação entre a estrutura e a dinâmica de ecossistemas complexos. Apesar dos extensos desenvolvimentos nesta área, muitas questões ainda não foram resolvidas. Notavelmente, a distribuição de cada tipo de dinâmica (atratores) em modelos ecológicos raramente foi estudada em favor da suposição de que comunidades ecológicas estariam em equilíbrio. Além disso, uma abordagem explorando a ubiquidade dos resultados conhecidos relacionando estabilidade e estrutura de comunidades no espaço de parâmetros é fundamental. Neste estudo, almejamos contribuir para este esforço realizando simulações baseadas no modelo generalizado de Lotka-Volterra (gLV) em uma ampla faixa de espaço de parâmetros. Para tanto, nosso objetivo de pesquisa é apresentar uma visão geral da dinâmica assintótica, coexistência e resiliência de comunidades ecológicas representadas através do modelo gLV, realizando uma exploração computacional abrangente que integra variabilidade no tipo, intensidade e distribuição de interações ecológicas. Descobrimos que diferentes dinâmicas de comunidades ecológicas podem ser descritas por parametrizações específicas do modelo gLV, apesar de uma grande prevalência de combinações de parâmetros levando ao crescimento ilimitado de algumas populações na comunidade. Também fomos capazes de detectar padrões conhecidos, como o impacto negativo da cooperação sobre a resiliência das comunidades e a ocorrência de exclusão competitiva em comunidades específicas. Assim, nossa pesquisa contribui significativamente para o avanço científico no entendimento das interconexões entre a estrutura e a estabilidade das comunidades ecológicas.

Contents

1	Introduction	1
2	Fundamental Concepts	4
2.1	Dynamical Systems Foundations	4
2.1.1	Basic Concepts	4
2.1.2	Dynamical Systems	4
2.1.3	State Space	5
2.1.4	Forward Invariance	7
2.1.5	Stability	7
2.1.6	Parameter Space	9
2.2	Network Theory Foundations	11
2.2.1	Basic concepts	11
2.2.2	Topological Properties	12
2.2.3	Models of Random Graphs	17
2.2.4	Random Graphs' Topological Parameters	18
2.3	Ecological Foundations	22
2.3.1	Levels of Biological Organization: Individual, Population, Community	22
2.3.2	Ecological Interactions	22
2.3.3	Ecological Community	23
2.3.4	Stability of Ecological Communities	23
3	Application of Fundamental Concepts to Ecological Communities	25
3.1	Generalized Lotka-Volterra Model Analytical Properties	25
3.1.1	The Generalized Lotka-Volterra Model (gLV)	25
3.1.2	Single Population Dynamics	26
3.1.3	Two interacting populations	26
3.1.4	Ecological Community with N Interacting Populations	38
3.2	Interplay between Structure and Dynamics in Ecological Communities	52
3.2.1	Ecological Communities as Dynamical Systems	52
3.2.2	Ecological Communities as Complex Networks	53
3.2.3	The Topology of Ecological Community Networks	53
3.2.4	Structure and Stability of Ecological Networks	55
4	Methodology	58
4.1	Input Parameters	60
4.1.1	Network Topologies, Number of Nodes, and Densities	60
4.1.2	Types of Interaction	61
4.1.3	Intensities of Interaction	61
4.1.4	Statistical Parametrization	62
4.1.5	Replicates	63
4.2	Temporal Evolution of Ecological Communities	64
4.2.1	Asymptotic Dynamics detected as function of Computational Effort	65
4.2.2	Final Community	65
4.3	Measurements	66
4.3.1	Topological Properties	66
4.3.2	Interaction Properties	66
4.3.3	Dynamical Properties	67

4.4	Techniques of Analysis	67
4.4.1	Measures of Variability	67
4.4.2	Stacked Area Plot	68
4.4.3	Histograms	68
4.4.4	Scatter Plots and Regressions	68
5	Results	69
5.1	Measured Variables Profile	69
5.1.1	Descriptive Statistics	69
5.1.2	Correlations	74
5.2	Variability in Experiments	74
5.2.1	Variability of Measured Quantities	74
5.2.2	Variability in Asymptotic Dynamics	77
5.3	Distribution of Asymptotic Dynamics	77
5.3.1	Overall	78
5.3.2	Distribution of Attractors Across Density and Intensity of Interactions	78
5.3.3	Dependence of Attractors on Interaction Types	79
5.3.4	Interaction Topology and Asymptotic Dynamics	80
5.4	Effects of Structure on Coexistence of Populations	81
5.4.1	Relation between Coexistence and Interaction Types	81
5.4.2	Relation between Coexistence and Density of Interactions	82
5.4.3	Relation between Coexistence and Topologies	83
5.5	Resilience of Communities	83
5.5.1	Resilience and Density of Interactions	83
5.5.2	Resilience and Intensity of Interactions	85
5.5.3	Resilience and Types of Interaction	85
5.5.4	Resilience and Attractors / Topologies	86
5.5.5	Distribution of Eigenvalues Before and After Stabilization	87
6	Discussion	90
7	Conclusion	93
7.1	Caveats and Perspectives	93
7.2	Take-home message	94
A	Detection of Attractors' Accuracy	95
A.1	Systems from the Literature	95
A.1.1	Equilibrium Points	95
A.1.2	Cycles	96
A.2	Temporal Evolution	97
B	Supplementary Figures and Tables	108
Bibliography		119

List of Figures

2.1	Topological Properties of Random Graphs	20
2.2	Degree Distribution of $G_{N,M}$ and Small World Models	21
2.3	Preferential Attachment Model Degree Distribution	21
3.1	Diagram of the gLV-2D Parameter Space	30
3.2	State Space Portrait for All gLV-2D Parameter Space Regions	36
3.3	Ecological Interpretation of Two-Dimensional Generalized Lotka Volterra Model Parameter Space	37
3.4	3D Examples of Type of Dynamics in gLV Model	51
4.1	Schematic Representation of the Methodology	60
4.2	Ternary Plot Schematic Diagram	62
4.3	Distributions of Interactions Intensities	63
4.4	Asymptotic Dynamics Detection	66
5.1	Coefficient of Variation of Replicates Generated with Same Parameters	76
5.2	Proportion of Bounded Scenarios in Each Experiment	77
5.3	Overall Distribution of Detected Attractors	78
5.4	Distribution of Types of Dynamics across differences in Network Density	79
5.5	Distribution of Types of Dynamics across differences in Interaction Intensity	80
5.6	Relationship between Interaction Intensity and Density on the proportion of bounded scenarios:	81
5.7	Proportion of Bounded Attractors across Interaction Types and Intensities	82
5.8	Distribution of Attractors across Topologies	83
5.9	Percentage of Surviving populations Decreases with Increasing Competition	84
5.10	Percentage of Surviving populations decreases with Increasing Competition	84
5.11	Proportion of surviving populations across topologies	85
5.12	Effect of network Density on network Resilience:	86
5.13	Effect of Interaction Intensity on network Resilience:	87
5.14	The proportion of Interaction Types influences the Resilience across different network Densities	88
5.15	Resilience is similar for different network Topologies	88
5.16	Standard case depicting the Resilience variation between Initial and Final Network	89
A.1	Example of Numerical Simulation	98
B.1	Correlation of All Variables	110
B.2	Proportion of Bounded Attractors across Interaction Types and Densities	111
B.3	Order of the Final Network Decreases with Increasing EXPLOITATION	112
B.4	Percentage of surviving population increases as proportion of Cooperation interaction increases	112
B.5	Order of the Final Network Decreases with Increasing Competition	113
B.6	Distribution of Input Density and Proportion of Competition	113
B.7	Order of the Final Network Decreases with Increasing Competition	114
B.8	Distribution of the Rightmost Eigenvalue Across Attractors	119
B.9	Median Rightmost Eigenvalue Ternary Plot	119

List of Tables

2.2	Classification of fixed points in 2D	9
3.1	Dynamical Behavior at gLV-2D Parameter Space Regions	33
4.1	Characteristics of Latin Hypercube Sampling (LHS) methods	64
4.2	List of Dynamical Variables	67
5.1	Descriptive Statistics of Dynamical Variables - I	70
5.2	Descriptive Statistics of Dynamical Variables - II	71
5.3	Descriptive Statistics of Interaction Variables	71
5.4	Descriptive Statistics of Topological Variables - I	72
5.5	Descriptive Statistics of Topological Variables - II	72
5.6	Descriptive Statistics of Topological Variables - III	73
5.7	Descriptive Statistics of Topological Variables - IV	73
5.8	Descriptive Statistics of the Distributions of Rightmost Eigenvalue for Different Subsets	86
B.1	List of Symbols and Variables used in Numerical Results	108
B.2	Regressions Between the Proportion of Survival and Input Variables	115
B.3	Regressions Between Rightmost Eigenvalue and Input Variables	117

Chapter 1

Introduction

Traditional indigenous wisdom teaches that "The idea of us humans being detached from the Earth [...] is absurd"¹(Krenak 2019). Indeed, humanity faces inevitable upcoming challenges to its very existence stemming from its profound impacts on Earth's ecosystems, such as climate change and the Anthropocene mass extinction (Dyurgerov and Meier 2000; Dirzo et al. 2014). These global changes are bound to have profound effects on natural systems and modern civilization alike. In this context, it is necessary to understand how the interactions among various species across different ecosystems shape ecological communities and how they are related to ecosystem stability. Despite the many possible definitions of ecological stability, knowing if a community existence, composition, and status are likely to change is essential. Even more important is to understand why that is the case.

Remarkably, early ecologists suggested that rich and complex ecosystems were associated with increased stability (Dunne 2006). However, this idea was challenged in 1972 when May put forward evidence that richer and more complex systems should be less stable (May 1972b). Despite the extensive research dedicated to elucidating the interplay between structure and dynamics in ecological networks, several inconsistencies are still present in the field.

In particular, it is mostly unknown how different aspects of structure influence the temporal evolution of the species' densities (Serván et al. 2018). Highlighting the importance of taking dynamics into account, Fowler (2009) found the relationship between structure and stability to be different when species' densities oscillated. Furthermore, due to the interconnections of the many aspects of structure (topology, type, and intensity of interactions), it is necessary to address multiple ecological interaction types while also allowing variability in how they are distributed and how intense they are.

Mathematical models are powerful tools in answering such questions. Accordingly, the use of mathematical models to study interactions between populations has at least one hundred years of history. In the foundational works of Lotka (Lotka 1920) and Volterra (Volterra 1926), fluctuations in populations' densities were mathematically considered through a set of coupled differential equations in time that takes into account both the intrinsic growth rate of the populations and the interactions between them. Even though these early investigations did not preclude the study of systems composed of many populations, most studies focused on the dynamics of few interacting populations. When many populations are considered, these interactions' patterns can be represented as a complex network, which embeds the structure of the interspecific interactions in an ecological community.

The framework most commonly used to represent systems of multiple interacting populations is the generalized Lotka-Volterra (gLV) model. Such a relatively simple model can encompass all possible pairwise interactions between an arbitrary number of populations. In this model, each interaction is represented by a law of masses – the magnitude of the interaction is directly proportional to the product of the densities of the two interacting populations. Despite its seeming simplicity, this model can display any dynamics (Smale 1976). Additionally, it is convenient that there is a direct connection between the interaction matrix of the gLV and the adjacency matrix of the network representation of the system.

¹Original phrase: "A ideia de nós, os humanos, nos descolarmos da terra, vivendo numa abstração civilizatória, é absurda. Ela suprime a diversidade, nega a pluralidade das formas de vida, de existência e de hábitos". Our translation.

Remarkably, one aspect that hindered the development of research on alternative asymptotic dynamics is the difficulty of addressing high-dimensional systems analytically. The dimensionality of the problem hinders the usefulness of traditional analytical tools developed for systems with few interacting populations. Moreover, the overall relation between structure and dynamics can be concealed by how structure and stability are defined or measured. Positively, many aspects of this relation are currently known, and some integrative approaches are starting to present a broader perspective. However, there is a need to join this complex and diverse knowledge corpus into an overall coherent framework. Undoubtedly, it is a substantial challenge to unravel the complex interplay between structure and dynamics in ecological communities.

In this study, we aim at contributing to this endeavor by performing simulations based on the generalized Lotka-Volterra model across a wide parameter space range. In accordance, our research objective is to present an overview of the generalized Lotka-Volterra model asymptotic dynamics, coexistence, and resilience by performing a comprehensive computational exploration integrating variability in type, intensity, and distribution of interspecific ecological interactions.

To fulfill our objectives, we generate random communities and simulate their temporal evolution with different structures. This procedure allows us to study populations' resilience and coexistence in each generated network. Our method can be broadly described as a series of three steps:

1. Creating a putative structure for the community with given topology and distribution of interspecific interactions types and intensities, the initial network;
2. Simulating community dynamical evolution computationally using the generalized Lotka-Volterra model and updating the initial network;
3. Characterizing initial and final communities' dynamical and topological properties.

This procedure allows us to investigate the effect of different structures upon communities' stability. Throughout our research, we focus on the resilience and coexistence aspects of ecological stability.

The dissertation is organized into seven chapters. This first chapter puts forward the setting of our research, alongside its motivation and core concepts.

In the second chapter, we present an overview of the fundamental concepts that form the technical and theoretical basis of the dissertation. Central ideas on dynamical systems, complex networks, and ecology are briefly discussed. The topics we addressed correspond to concepts and techniques which were used in the development of the dissertation.

Namely, some possible classes of asymptotic dynamics, the state space and parameter space, and local asymptotic stability are addressed in the dynamical systems section. A fair few topological measures and random network models are presented alongside canonical definitions of graph theory in the complex network theory section. In the section related to ecological theory, definitions of levels of biological organization, ecological interaction and stability are presented.

In chapter three, we apply the theory and concepts presented in the introduction to ecological communities. We study the analytical properties of the generalized Lotka-Volterra model for one, two, and N dimensions. For the bidimensional case, we present a general picture of the model parameter space. We further highlight the correspondence between the type of ecological interactions and attractors. For the N -dimensional case, we present classical results on the existence, feasibility, local and global stability of equilibrium points. Additionally, we present examples of other possible dynamics.

We supplement this chapter with a brief revision of the literature concerning the interplay between structure and dynamics in ecological communities. This revision addresses an ecological community's representation as a dynamical system and as a complex network alongside its topological fingerprints. An overview of current knowledge on the mutual influences of structure and stability in ecological communities is also presented.

In the fourth chapter, we present the methodology employed in this research. Technical aspects of the use of Latin Hypercube Sampling (LHS) for the generation of putative ecological networks and parametrization of the generalized Lotka-Volterra model are addressed. We provide a summary of the program developed to simulate the construction and dynamics of each ecological community. We also indicate parameters used in the computations and define the measures performed upon the generated data.

In chapter five, we present the results of our numerical simulations. Concerning the distribution of asymptotic dynamics, our results indicate the ubiquity of unbounded growth of some populations' densities in the model for the parametrization procedure adopted. The type of dynamics was influenced by intensity and type of interactions. We also show that populations' coexistence in a community is affected negatively by the proportion of competitive interactions and the network density. Finally, we provide evidence that community resilience is improved with increasing competition and decreasing interaction strength.

We discuss our findings in chapter six, relating them to current knowledge present in the literature. In chapter seven, the conclusion, we consider the take-home message of this dissertation. Additionally, we also point out the caveats of the current research that should be addressed by future studies.

Lastly, we present additional supporting material in the appendix. The additional resources encompass tests of the algorithm developed to detect attractors, supplementary analysis, and results.

Chapter 2

Fundamental Concepts

2.1 Dynamical Systems Foundations

In the following sections, some of the fundamental elements of dynamical systems theory are presented. The concepts and methods presented constitute an essential basis for the dissertation.

2.1.1 Basic Concepts

In order to study and understand the natural processes in the world, it is customary and convenient to focus on a particular set of objects and phenomena that constitute a somewhat cohesive whole. In practice, this can be anything from a single atom or cell to the entire universe. The central aspect of this process is identifying this "cohesive whole," i.e., the *system*, which is conveniently separated from the rest of the world as being the object of study.

Etymologically, the root of the contemporary word "system" is the Greek word *συστῆμα*, "systema", meaning "what stands together, a body of united elements" (Wedgwood 2017). By its definition, it is usual that the system itself comprises other interacting constituents or elements. It is the interaction among these elements, entities, or constituents that often govern eventual temporal changes on the system and, ultimately, its evolution in time. This process of abstracting the real system is usually referred to as *modeling*, and the identification of the system itself can arguably be seen as a kind of modeling.

It is convenient to use numerical quantities, *state variables*, that can describe the system (given the meaning associated with such numbers) to identify and characterize it. The system evolution in time is then studied by assessing the temporal change of the state variables. When it is possible to write down an equation describing precisely how we expect these variables to change, we can build a *mathematical model* representing the system.¹

The mathematical model should not be seen as a result of empirical observation alone, nor can it be independent. Therefore, the mathematical model is derived from theoretical expectations for the system's phenomenology. In a sense, the model becomes a tool, aiding the scientist to explore the consequences of assumptions concerning the system. Once a model is constructed, it is possible to manipulate this representation of the system to derive new results and predictions, with caution not to expect the model to be "more real than the system itself".

2.1.2 Dynamical Systems

In this work, we use a generalized mathematical model based on the dynamical systems theory to represent the real system. In the following, we focus on the techniques suitable for the study of *dynamical systems*. A dynamical system is described by the state variables $\{X_i\}$ for which their functions $\{F_i\}$ give the system's temporal evolution.

If the state variables ($\{X_i\}$) that characterize the system vary continually in time following a vectorial function that is composed of the functions $\{F_i\}$ of the state variables, parameters, and potentially the time, it is possible to write

¹Note that other frameworks are possible for computational and mathematical modeling, such as agent-based modeling.

$$\begin{aligned}\frac{dX_1}{dt} &= F_1(X_1, X_2, \dots, X_n, \vec{\beta}_1, \vec{\beta}_2, \dots, \vec{\beta}_n, t) \\ \frac{dX_2}{dt} &= F_2(X_1, X_2, \dots, X_n, \vec{\beta}_1, \vec{\beta}_2, \dots, \vec{\beta}_n, t) \\ &\vdots \\ \frac{dX_n}{dt} &= F_n(X_1, X_2, \dots, X_n, \vec{\beta}_1, \vec{\beta}_2, \dots, \vec{\beta}_n, t),\end{aligned}$$

This expression can be written in a more succinct way as

$$\frac{d\vec{X}}{dt} = \vec{F}(\vec{X}, \vec{\beta}, t), \quad (2.1)$$

in which $\vec{X}, \vec{\beta}$ represent respectively the vector of state variables and control parameters of the system. The vectorial function $\vec{F}(\vec{X}, \vec{\beta}, t)$ denote system's *flow*, i.e., the phenomenologically or mechanistically-derived function describing how each state variable changes in time.

The flow dependency on the state variables indicates their mutual interaction. In contrast, the control parameters convey information on particular values for the system's properties (e.g., rates of interaction, intrinsic birth and death rates for an ecological system). Parameters establish a connection between the abstract mathematical representation of a function and the real system's phenomenology as each choice of parameters' values lead to particular system behavior, albeit a different set of parameters may lead to equivalent behavior (see section 2.1.6).

The flow dependency on the state variables indicates the mutual interaction between state variables. In contrast, the control parameters convey information on particular values for the system's properties (e.g., rates of interaction, intrinsic birth, and death rates for an ecological system).

Usually, the parameters are not time-dependent, i.e., they are constant $\frac{d\vec{\beta}}{dt} = 0$. A common justification for this assumption is if the rates at which the parameters change are not at the same time scale as the rates at which the state variables change. Also, from a mathematical standpoint, the parameters of a model can unveil structural similarities of very different phenomena if the only difference between them is the values they can assume.

The state variables do not need to change in time, as they could change in response to changes in other *independent variables*, like space or temperature, for instance. Nevertheless, the dynamical systems framework presented thus far still holds if we study only one independent variable. In the case that the vector function $\vec{F}(\vec{X}, \vec{\beta})$ does not depend on time (independent variable), the system is said to be an *autonomous dynamical system*.

2.1.3 State Space

As seen in subsection 2.1.1, the configuration of a dynamical system is determined by the specific values of the set of state variables. It is then possible to abstract the individual values of each state variable and consider the set of all possible configurations of a given system, the *state space* or *phase space*. In the state space, each state of the system corresponds to a single point, which allows the use of geometrical methods in the analysis of the dynamical system.

The temporal evolution of a particular state in a dynamical system is represented by some curve in the state space, a *trajectory*, as the flow direction at each point indicates the direction of evolution of this curve. In other words, the trajectory is always tangent to the flow, and each point on the phase space corresponds to one unique curve, as no two distinct trajectories may intersect each other. This property indicates that the state space presents a complete picture of the system's evolution for a particular set of parameters.

The dynamical systems theory shifted from trying to completely solve the mathematical system, which is always possible for the linear systems, to gaining insights into its state space's main properties. The main idea of this "graphical" approach is to describe the flow's qualitative properties based on its state space, studying typical invariant sets in such space. Importantly, if some trajectory starts at an invariant set, it never leaves the invariant set.

Some examples of typical invariant sets are points that correspond to a stationary flow (*equilibria*) and closed orbits (*cycles*). Other invariant sets are toroidal regions formed by two or more incommensurable periodic motions (*torus*) and even *strange attractors*, complex topological objects which are the fingerprint of a chaotic dynamical system. Generically, a topological object on the state space for which neighboring trajectories converge to is an *attractor*.

2.1.3.1 Equilibrium Points

Equilibrium points are points in the state space for which the flow is stationary (Monteiro 2011). In other words, if the system is at an equilibrium point, no temporal change occurs, and the values of its state variables do not change. An equilibrium point X^* can be characterized mathematically as

$$\vec{X}^*, \text{ such that } \vec{F}\left(\vec{X}^*, \vec{\beta}\right) = 0. \quad (2.2)$$

In order to find such points, it is possible to solve the set of equations

$$\frac{d\vec{X}}{dt} = 0. \quad (2.3)$$

It is possible to obtain the solution for the system (2.3) both in terms of an analytic expression² or numerically. Since, in general, the flow is non-linear, solving the system of equations analytically may be challenging.

Temporal changes of variables are smaller on the neighborhood of equilibrium points, indicating its study is essential when considering the correspondence between the model and the real system.

2.1.3.2 Limiting Cycles

Limit cycles are closed isolated trajectories on the state space for which near trajectories converge to or diverge from (Monteiro 2011). The requirement that these trajectories are closed implies that in the system's temporal evolution, the same point in the phase space will be revisited at a later time, i.e.,

$$\vec{X}(t_0) = \vec{X}(t_0 + T), \quad (2.4)$$

in which T denotes the period of such cycle. The cycles must be isolated so that there are no other closed trajectories infinitesimally near the limit cycle. If all trajectories on a cycle's neighborhood converge to it, the cycle is classified as stable, and when some trajectories diverge from it is correspondingly classified as unstable.

2.1.3.3 Strange Attractors (Chaos)

Strange attractors are a typical fingerprint of chaotic systems. Although there are different ways to identify chaotic behavior, it is possible to characterize a dynamical system as chaotic if it has topological transitivity and adjacent trajectories diverge from one another (Thompson and Stewart 2002). By transitivity, it is implied that each trajectory eventually goes arbitrarily close to any point in the attractor, i.e., the trajectory is dense. The divergence of adjacent trajectories indicates sensitivity to initial conditions, i.e., two very close initial conditions will lead to very distinct trajectories.

²An analytical expression is a mathematical expression composed of well known operations and functions. This is, in some sense, a subjective statement as different set of mathematical operations and functions can be classified as "well known" depending on the context. Nevertheless, the important aspect is that calculating a numeric value for an analytic expression should be relatively straightforward once the value of its parameters are available.

2.1.4 Forward Invariance

If some region \mathbb{S}' of the state space \mathbb{S} , $\mathbb{S}' \subseteq \mathbb{S}$, is such that no trajectory starting in \mathbb{S}' ever leaves this region, then this region is said to be forward invariant (Hofbauer and Sigmund 1998). This property is particularly relevant if there is a phenomenological reason why only a subregion of this state space is meaningful or has essential properties for the modeled system. For instance, this is the case when considering densities or abundances of populations as state variables. Although negative values for the densities are, in principle, part of the state space, there is no biological justification for these values.

Consider, for example, an ecological model describing the dynamics of wolves and sheep in which the state variables are the densities of wolves and sheep. It makes no sense to talk about negative densities, so the forward invariance guarantees that no sane, positive densities of wolves and sheep will evolve into unmeaningful negative densities.

Mathematically, forward invariance can be defined as

$$\vec{X}(t_0) \in \mathbb{S} \implies \vec{X}(t_0 + t) \in \mathbb{S}. \quad (2.5)$$

2.1.5 Stability

Some theoretical results guarantee that the change of behavior of the dynamical systems is associated with the attracting structures (Fiedler-Ferrara and Prado 1994). Due to that, it is essential to identify and analyze the stability of those structures. For instance, we say that an equilibrium point is globally stable if the trajectories converge to it for any initial conditions. Otherwise, it may be just locally stable but not globally stable.

2.1.5.1 Local Stability of Equilibria

In order to analyze the local stability of the equilibria of a dynamical system, we search for a linearized version of the system in the neighborhood of such equilibria. It is then necessary to make a translational transformation in the state space to bring the equilibrium point (\vec{X}^*) to the state space's origin. This transformation is $\vec{X} - \vec{X}^* \rightarrow \vec{X}'$ so that $\vec{X}'^* = \vec{0}$. We can then expand (2.1) around the equilibrium point as

$$\begin{aligned} \frac{dX_i}{dt} &= F_i(\vec{X}, \vec{\beta}) \\ \frac{dX_i}{dt} &= \sum_{k=0}^{\infty} \left[\frac{1}{k!} \left[\sum_{j=1}^n (X_j - X_j^*) \frac{\partial}{\partial X_j} \right]^k \vec{F}_i(\vec{X}, \vec{\beta}) \right] \Bigg|_{\vec{X}=\vec{X}^*} \end{aligned}$$

It is not possible to write this expansion in a compact matrix form for its higher terms include tensors of higher orders. However, it is convenient to write the expansion in the full vector-matrix form up to the linear term in $(\vec{X} - \vec{X}^*)$ as

$$\begin{aligned} \frac{d\vec{X}}{dt} &= \frac{1}{0!} (\vec{X} - \vec{X}^*)^0 \vec{F}(\vec{X}^*, \vec{\beta}) + \frac{1}{1!} (\vec{X} - \vec{X}^*)^1 \mathbf{J} + \mathcal{O}\left((\vec{X} - \vec{X}^*)^2\right) \\ \frac{d\vec{X}}{dt} &= (\vec{X} - \vec{X}^*) \mathbf{J} \Big|_{\vec{X}=\vec{X}^*} + \mathcal{O}\left((\vec{X} - \vec{X}^*)^2\right) \end{aligned}$$

where we employed a Jacobian application, resulting in a *Jacobian matrix* (\mathbf{J})

$$J_{ij} \equiv \frac{\partial F_i(\vec{X}, \vec{\beta})}{\partial X_j} \quad (2.6)$$

If we consider the behavior of the system on the neighborhood of \vec{X}^* , we can assume $(\vec{X} - \vec{X}^*) \approx 0$ and disregard higher-order terms so that

$$\begin{aligned}\frac{d\vec{X}}{dt} &\approx \left(\vec{X} - \vec{X}^* \right) \mathbf{J} \Big|_{\vec{X} = \vec{X}^*} & \vec{X} - \vec{X}^* \rightarrow \vec{X}' \\ \frac{d\vec{X}'}{dt} &\approx \vec{X}' \mathbf{J} \Big|_{\vec{X}' = 0}\end{aligned}\tag{2.7}$$

With this linearized system, it is possible to study flow behavior in the neighborhood of equilibria. Dropping the notation for the transformed vectors, we have (note that $\dot{\vec{X}} = \frac{d\vec{X}}{dt}$)

$$\dot{\vec{X}}' \approx \mathbf{J} \Big|_{\vec{X}' = 0} \vec{X}' \tag{2.8}$$

$$\dot{\vec{X}}' \approx \mathbf{J} \Big|_{\vec{X}' = 0} (k_1 \vec{\nu}_1 + k_2 \vec{\nu}_2 + \dots + k_n \vec{\nu}_n) \tag{2.9}$$

$$\dot{\vec{X}}' \approx \left(k_1 \mathbf{J} \Big|_{\vec{X}' = 0} \vec{\nu}_1 + k_2 \mathbf{J} \Big|_{\vec{X}' = 0} \vec{\nu}_2 + \dots + k_n \mathbf{J} \Big|_{\vec{X}' = 0} \vec{\nu}_n \right) \tag{2.10}$$

$$\dot{\vec{X}}' \approx (k_1 \lambda_1 \vec{\nu}_1 + k_2 \lambda_2 \vec{\nu}_2 + \dots + k_n \lambda_n \vec{\nu}_n) \tag{2.11}$$

$$(k_1 \vec{\nu}_1 + k_2 \vec{\nu}_2 + \dots + k_n \vec{\nu}_n) \approx (k_1 \lambda_1 \vec{\nu}_1 + k_2 \lambda_2 \vec{\nu}_2 + \dots + k_n \lambda_n \vec{\nu}_n) \tag{2.12}$$

$$\dot{k}_i \approx k_i \lambda_i \tag{2.13}$$

$$\Rightarrow k_i(t) \approx e^{\lambda_i t} k_0 \tag{2.14}$$

$$\Rightarrow \vec{X}'(t) \approx \sum_{i=1}^n k_{0i} e^{\lambda_i t} \vec{\nu}_i. \tag{2.15}$$

It is convenient to express $\vec{X}' = \sum_{i=1}^n k_i \vec{\nu}_i$, as the expansion of an arbitrary vector in the eigenbasis of \mathbf{J} assuming the eigenvector equation $\mathbf{J} \vec{\nu}_i = \lambda_i \vec{\nu}_i$

Classification of the Stability of Equilibria As it is possible to see through the examination of (2.15), if all eigenvalues of the Jacobian matrix have negative real parts, the temporal evolution of the system, on the neighborhood of an equilibrium point, converges to the equilibrium point itself, and it is said to be *asymptotically stable* in Lyapunov's sense. Since we have carried out a coordinate transformation to bring the equilibrium point to the origin of the state space, it is possible to see that

$$\operatorname{Re}(\lambda_i) < 0 \forall i \Rightarrow \lim_{t \rightarrow \infty} e^{\lambda_i t} = 0 \Rightarrow \lim_{t \rightarrow \infty} \vec{X}'(t) = \lim_{t \rightarrow \infty} \sum_{i=1}^n k_{0i} e^{\lambda_i t} \vec{\nu}_i = \vec{0} \Rightarrow \lim_{t \rightarrow \infty} \vec{X}(t) = \vec{X}^*.$$

If there is at least one eigenvalue with a positive real part, then the limit above does not hold, and the flux will deviate from the equilibrium point in at least one direction. In this case, the equilibrium point is said to be *unstable* in Lyapunov's sense. Since the approximation (2.15) only holds on the neighborhood of the equilibrium point, it is not possible to study its behavior on regions far from equilibrium.

In the intermediate situation, namely, if at least the real part of one eigenvalue is zero and the real part of others is either zero or negative, the system on equilibrium point's neighborhood is *stable*, but not asymptotically stable, in Lyapunov's sense. In this case, the system does not approach the equilibrium point in every direction. However, it does not leave its neighborhood either, in a way that all trajectories beginning in some arbitrary close region near to the equilibrium point will remain inside such region.

Concerning the imaginary part of the eigenvalues, if these are nonzero, then the flux has some oscillatory movement. If the real parts of all eigenvalues are negative (one of the real parts is positive), then the flow oscillates, converging into the equilibrium point ('escaping' from the equilibrium point). If the real parts of all eigenvalues are zero, and the imaginary part is nonzero, the system oscillates around the equilibrium point without approaching or leaving its neighborhood.

Classification of Equilibria in 2 dimensions In two dimensions, we have the generic system

$$\dot{X}_1 = f(\vec{X}, \vec{\beta}) \quad (2.16)$$

$$\dot{X}_2 = g(\vec{X}, \vec{\beta}) \quad (2.17)$$

which can be linearized by means of a Jacobian matrix, yielding:

$$\mathbf{J} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad (2.18)$$

In which the explicit dependence on $\vec{X}, \vec{\beta}$ was suppressed. We can calculate the eigenvalues of this Jacobian matrix as the roots of the characteristic equation. Therefore, by calculating $\det(\mathbf{J} - \lambda \mathbf{I})$ we have

$$\lambda_{\pm} = \frac{1}{2} \left[b \pm \sqrt{b^2 - 4c} \right] \Leftrightarrow \lambda_{\pm} = \frac{1}{2} \left[\text{tr}(\mathbf{J}) \pm \sqrt{(\text{tr}(\mathbf{J}))^2 - 4 \det(\mathbf{J})} \right]$$

$$\begin{aligned} b &= \text{tr} \mathbf{J} \\ c &= \det \mathbf{J} \end{aligned}$$

It is then possible to classify the different types of equilibria presented in a 2D system, as in the table 2.1.5.1.

$\det \mathbf{J}$	$(\text{tr} \mathbf{J})^2 - 4 \det \mathbf{J}$	$\text{tr} \mathbf{J}$	$\text{Re}(\lambda)$	$\text{Im}(\lambda)$	Classification
-	$\Rightarrow +$	\pm	$\{+, -\}$	0	Saddle point
+	+	-	-	0	Stable node
+	+	+	+	0	Unstable node
+	0	-	$-(\times 2)$	0	Stable degenerate node
+	0	+	$+(\times 2)$	0	Unstable degenerate node
+	-	-	-	\pm	Stable Spiral
+	-	+	+	\pm	Unstable Spiral
+	-	0	0	\pm	Center (neutral)
0	$\Rightarrow +$	\pm	$\{0, \pm\}$	0	Line or plane of fixed points

Table 2.2: Classification of fixed points in 2D

2.1.6 Parameter Space

In an analogous manner to the state space, it is possible to think about a system's parameters as being the coordinates of a point in an abstract space. The dimensions of this space are the parameters. Accordingly, the limits of such space are given by the set of allowed values of the corresponding parameters. As the parameter values are not independent of the real system being modeled, defining the parameter space is also a part of the modeling effort. In fact, investigating how different parameter configurations might induce changes in the state space is crucial in dynamical systems theory.

As each point in parameter space corresponds to a particular set of model parameters, it also corresponds to a particular state space configuration. In contrast to the state space, there is no dynamical concept of trajectory in the parameter space for a system with fixed or constant parameters. Nevertheless, the parameter space study might provide many important insights into the system's behavior. Significantly, two neighboring points on the parameter space may lead to very different state space portraits, i.e., not *topologically equivalent*. In the case that two neighboring points in the parameter space do not lead to topologically equivalent state-space portraits, it is said that there is a *bifurcation* at the *bifurcation point* which separates the two distinct regions.

In large systems, the components might display an intricate pattern of interactions, with each component interacting with a potentially distinct subset of the whole system. In such cases, the system can be represented as a network, with components and their interactions being the nodes and edges. The parameters in the dynamical systems are then identified with network quantities.

2.2 Network Theory Foundations

Networks can be understood intuitively as a set of elements that have some kind of pairwise interaction. These elements can represent neurons, galaxies, cities, people, populations, species, genes, molecules, among many more. For instance, if there is an interest in studying interactions among people, each element might be defined as a person. However, the interaction can still be defined in many different ways, such as friendship, family, work, and collaboration. Remarkably, the concept of network is part of the current globalized culture as exemplified by the internet, traffic, plumbing, electrical, and social networks.

Since networks can be constructed to represent all sorts of interacting elements, such representation might provide valuable insights into various complex systems and highlight similarities between seemingly unrelated phenomena. Among the advantages of representing a system as a network is making the structural features of its interactions evident. In the following, we define basic concepts in graph and network theory, present some random network models, and provide an overview of some of the properties/quantities used to classify and study networks. We refer to graphs when explicitly addressing the mathematical object with which networks can be represented. The definitions put forward in this section are based on Monteiro (2011) and M. Newman (2018a).

2.2.1 Basic concepts

A generic graph G is defined as $G = (V, E)$ in which V is the set of all the elements present in the network, denominated nodes or vertices, and E is the set of all the interactions between elements of V , the edges. Each edge $e_k \in E$ denotes an interaction between two elements of V . If the interactions are symmetrical, the graph is said to be *undirected*, and an edge can be represented by a set of two elements $e_k = \{v_i, v_j\}$ as v_i is connected to v_j if, and only if, v_j is connected to v_i . Conversely, if the interactions are not symmetric, the graph is said to be *directed* and the edge is represented by an ordered pair $e_k = (v_i, v_j)$, in which $v_i, v_j \in V$.

Two nodes $v_i, v_j \in V$ are said to be *adjacent* if, and only if, there is an edge $e_k \in E$ such that $e_k = (v_i, v_j)$. The set of all nodes that are adjacent to a given node v_i is the *neighborhood* of v_i and the *degree of the node* $\deg v_i$ is the number of neighbors it has. If all nodes in a graph have the same degree, this graph is said to be a *regular graph* or *k-regular graph* where k indicates its degree. A subgraph of a given graph $G = (V, E)$ is a graph $G' = (V', E')$ such that $V' \subseteq V \wedge E' \subseteq E$.

Interestingly, nodes that are not adjacent may also be able to exchange information. This process is essential in many applications of graphs; thus, it is useful to develop some nomenclature in this sense. In the graph, a *walk* is a sequence of edges $S = (e_1, e_2, \dots, e_n)$ such that each edge starts in the node its predecessor ends, i.e. $e_k = (v_i, v_j) \iff e_{k+1} = (v_j, v_l) \forall k \in (1, \dots, n-1)$. The length of a walk is then defined as the number of edges in such walk, $|S|$. A walk might go over the same edge multiple times, but this is not the case for a *path* in the graph, which is a walk with the additional requirement that no two edges in the sequence S are equal, $e_i \neq e_j \forall e_i, e_j \in S$ if $i \neq j$. It is often convenient to denote a walk or path as a sequence of nodes instead so that all adjacent nodes are linked by an edge, i.e. $S = (v_1, v_2, \dots, v_n)$ if $\exists e_k = (v_i, v_{i+1}) \in V \forall 1 \leq i < n$. A *simple path* is then defined as a walk in which no vertex is visited more than once, i.e. $S = (v_1, v_2, \dots, v_n), v_i \in V \forall v_i$ such that $v_i \neq v_j \forall v_i, v_j$. The notion of paths can be used to conceptualize distances between nodes so that the distance between two given nodes $l_{v_i, v_j} = l_{ij}$ is the length of the shortest path between the two nodes. If there are no paths that connect the nodes v_i and v_j , the distance is then undefined or infinite, $l_{ij} = \infty$. Conversely, the distance from a node to itself is defined to be 0.

If a path is such that its first and last nodes are the same, this path is classified as a *cycle*. In accordance, if all nodes in a cycle except the first/last one are distinct, the cycle is a *simple cycle*. If some node in the graph has no neighbors, it is said to be an *isolated node*. Conversely, if there is a path between every pair of a given graph's vertices, it is said to be *connected*. For directed graphs, a graph in which there is a path in both directions between each pair of vertices is said to be *strongly connected*. In contrast, it is said to be *weakly connected* if the undirected graph obtained by replacing all its directed edges with undirected ones is connected. In this sense, a *connected component* is the maximally connected subgraph of a given graph; no other node of the graph can be added to the connected component so that it remains connected. There is always only one connected component in a connected graph, and every isolated node is also defined as a

connected component of its respective graph. A *complete graph* is a graph in which all potential edges exist.

Networks and graphs can be represented in a multitude of ways. While it is possible to simply list the sets V and E , some applications benefit of alternative representations such as adjacency and distance matrix. The entries of adjacency matrix $\mathbf{A} = (a_{ij})$ are 1 if there is an edge connecting v_i to v_j and 0 otherwise. Therefore

$$a_{ij} = \begin{cases} 1 & (v_i, v_j) \in E \\ 0 & (v_i, v_j) \notin E \end{cases}.$$

On the other hand, each entry of the distance matrix $\mathbf{B} = (b_{ij})$ is given by the distance between nodes, i.e. $b_{ij} = l_{ij}$.

The definitions and nomenclature put forward in this section allow for a common language to address network properties and are used throughout the text.

2.2.2 Topological Properties

In order to numerically characterize the topology of both observed and theoretical networks, it is necessary to define some metrics that reflect the topological information. In the following, different numerical topological properties are defined and exemplified by its characteristic values.

2.2.2.1 Number of Nodes (Order) - N

Order indicates the number of nodes in a given network. For a graph $G = (V, E)$, the order is the number of elements in the V set of nodes, $N = |V|$.

2.2.2.2 Number of Edges (Size) - M

Size indicates the number of edges in a given network. For a graph $G = (V, E)$, the size is the number of elements in the E set of edges, $M = |E|$.

2.2.2.3 Density - D

The density of a network, also referred to as *connectance*, measures the ratio between the size of the graph and the maximum size it could have, given its order. The minimum density for a network is $D = 0$ for a network composed entirely of isolated nodes and $D = 1$ for the complete graph.

For an undirected graph, the network density is defined as

$$D = \frac{2M}{N(N-1)}, \quad (2.19)$$

while for directed graphs, the network density is

$$D = \frac{M}{N(N-1)}. \quad (2.20)$$

The difference between both definitions is a consequence that for a given pair of nodes v_i, v_j , there is only one possible edge $e_k = \{v_i, v_j\}$ in an undirected network, but there are two possible edges in a directed network, namely $e_l = (v_i, v_j)$ and $e_m = (v_j, v_i)$.

2.2.2.4 Degree Distribution - $P(k)$

The degree distribution of all nodes in a network is a very representative measure of its topology. Such distribution provides the probability that a randomly selected node has degree k . It is then defined as

$$P(k) = \frac{1}{N} |\{v_i \in V \text{ if } \deg v_i = k\}|, \quad (2.21)$$

in which $|X|$ indicates the number of elements in the set X . As a probability, the sum of $P(k)$ for all possible degrees k in the network is one:

$$\sum_{k=k_{min}}^{k_{max}} P(k) = 1$$

2.2.2.5 Assortativity - r

Assortativity is a topological measure that assesses nodes' tendency to be connected to other nodes with similar properties: networks in which nodes are connected with others of similar (dissimilar) properties are classified as assortative (disassortative) networks (M. E. J. Newman 2003). This measure is also referred to as *mixing pattern* in the literature. Assortativity can be measured both for categorical and scalar properties, and it is commonly applied to node degree, in which case it is named *degree assortativity*. In its basics, this metric measures the correlation of pair of nodes in relation to some property attributed to the nodes. It can be assessed with many of the tools available to measure correlations.

Categorical Node Characteristics Following the foundational definitions by M. E. J. Newman (2003) we have:

$$r = \frac{\sum_i e_{ii} - \sum_i a_i b_i}{1 - \sum_i a_i b_i} \quad (2.22)$$

in which the following definitions hold:

$(\mathbf{e})_{ij} = e_{ij}$ fraction edges that connect nodes with characteristic i and j ;

a_i, b_i fraction of each type of edge end that is attached to a node with characteristic i .

These variables satisfy the sum rules:

$$\sum_{ij} e_{ij} = 1, \quad \sum_j e_{ij} = a_i, \quad \sum_i e_{ij} = b_j \quad (2.23)$$

It is also possible to rewrite expression (2.22) in matricial form noting that the trace of a matrix can be given as $\text{Tr } \mathbf{A} = \sum_i A_{ii}$ and for a given matrix product $\mathbf{C} = \mathbf{AB} \Leftrightarrow c_{ij} = \sum_k a_{ik} b_{kj}$. Therefore we have

$$\begin{aligned} \sum_k a_k b_k &= \sum_k (\sum_j e_{kj})(\sum_i e_{ik}) \\ &= \sum_i \sum_j \sum_k e_{ik} e_{kj} \\ &= \sum_i \sum_j (\mathbf{e}^2)_{ij} \\ &= \|\mathbf{e}^2\|, \end{aligned}$$

In which the norm $\|\mathbf{A}\|$ indicates the sum of all the elements in matrix \mathbf{A} . From the above, it is possible to rewrite expression (2.22) as

$$r = \frac{\text{Tr } \mathbf{e} - \|\mathbf{e}^2\|}{1 - \|\mathbf{e}^2\|}. \quad (2.24)$$

The assortativity is bounded in the range $-1 \leq r \leq 1$. Negative values indicate a disassortative network in which nodes of distinct characteristics are systematically connected. In contrast, positive values indicate an assortative network in which nodes with similar characteristics are connected among themselves.

Scalar Node Characteristics The general definition of linear correlation between two random variables X, Y is

$$\rho_{XY} = \frac{E(XY) - E(X)E(Y)}{\sqrt{E(X^2) - E(X)^2}\sqrt{E(Y^2) - E(Y)^2}} \quad (2.25)$$

For node characteristics that are described by integers, the assortativity index presented thus far is not adequate. It is then necessary to lay a methodological basis to answer questions concerning tendencies of these connections, e.g., "are richer individuals more likely to have other rich individuals as their friends?". In order to do so, it is convenient to make use of the mixing matrix \mathbf{e} once again, in which the element e_{ij} presents the fraction of edges that connect a node with characteristic value i to other with characteristic value j . Under these conditions, the assortativity is

$$r = \frac{\sum_{xy} xy(e_{xy} - a_x b_y)}{\sigma_a \sigma_b}, \quad (2.26)$$

in which σ_a, σ_b are the standard deviations of a and b . Expression (2.26) follows from (2.25), by observing that

$$\begin{aligned} E(X) &= xa_x = \sum_y xe_{xy} \\ E(Y) &= yb_y = \sum_x ye_{xy} \\ E(XY) &= xye_{xy} \\ \sigma_a &= \sqrt{E(X^2) - E(X)^2} \\ \sigma_b &= \sqrt{E(Y^2) - E(Y)^2}. \end{aligned}$$

Equivalently, for continuous variables we define the joint probability density $e(x, y)$ and

$$\begin{aligned} E(X) &= xa(x) = \int xe(x, y) dy \\ E(Y) &= yb(y) = \int ye(x, y) dx \\ E(XY) &= \iint xye(x, y) dx dy. \end{aligned}$$

For a continuous variable, the assortativity can be calculated as

$$r = \frac{\iint xy(e(x, y) - a(x)b(y)) dx dy}{\sigma_a \sigma_b}, \quad (2.27)$$

Once again, the value of the assortativity is bounded ($-1 \leq r \leq 1$).

Degree Assortativity categorical or numerical characteristics. Examples of such characteristics are gender (categorical) and age (numerical) in social networks or developmental stage (categorical) and abundance (numerical) in ecological networks. Nevertheless, when solely considering the topological structure of graphs, it is still possible to calculate assortativity based on topological characteristics, as node degree. Indeed, node degree assortativity has been extensively studied (Noldus and Van Mieghem 2015). In order to avoid redundancy when considering a pair of nodes connected by an edge, the remaining degree is used (one less than the node's degree).

To calculate the probability that a randomly selected edge is incident on a vertex with remaining degree k , consider that the total number of edges M in the network is

$$M = N\bar{k} \\ = N \sum_k kP(k),$$

in which \bar{k} is the average degree of the network. Additionally, the number of edges linked to a node with degree k is $NkP(k)$. Thus, the remaining degree distribution is

$$q_k = \frac{(k+1)p_{k+1}}{\sum_k kP(k)}. \quad (2.28)$$

Defining e_{jk} to be the fraction of edges connecting nodes with remaining degrees j and k , it follows that e_{jk} and q_k are related as

$$\sum_j e_{jk} = q_k. \quad (2.29)$$

The degree assortativity can then be calculated as

$$r_D = \frac{\sum_{jk} jk(e_{jk} - q_j q_k)}{\sigma_q^2}, \quad (2.30)$$

in which σ_q is the standard deviation of the remaining degree distribution q_k . Alternatively, for a directed network

$$r_D = \frac{\sum_{jk} jk(e_{jk} - q_j^{\text{in}} q_k^{\text{out}})}{\sigma_q^{\text{in}} \sigma_q^{\text{out}}}, \quad (2.31)$$

in which superscripts $\square^{\text{in}}, \square^{\text{out}}$ denote variables related to remaining in-degree and out-degree, respectively.

2.2.2.6 Average Shortest Path Length - $\langle l \rangle$

The average shortest path length, also referred to as average distance, captures the central tendency of distances between pairs of nodes in a network. Since the distance between two nodes that are not connected diverges, this measure is undefined for unconnected networks. The definition of the average shortest path length is similar for directed and undirected networks, noting that $l_{ij} = l_{ji}$ for undirected networks. Thus

$$\langle l \rangle = \frac{1}{N(N-1)} \sum_{i=1}^N \sum_{j=1}^N l_{ij}. \quad (2.32)$$

2.2.2.7 Diameter - d

While the average shortest path length accounts for the average distance between the nodes of a network, the diameter measures the maximum shortest path length among the shortest paths between all pairs of nodes in the graph. Therefore, the diameter is an answer to the question "how far from each other can a pair of nodes be in a given graph?". Mathematically, it is possible to define it as

$$d = \max \{l_{ij} \forall i, j\}. \quad (2.33)$$

2.2.2.8 Clustering coefficients - C and $\langle C \rangle$

Clustering measures the tendency of nodes to cluster together. It provides insight on whether the neighbors of an arbitrary node are also neighbors among themselves, both locally and globally.

The local cluster coefficient, C_i , proposed by Watts and Strogatz (1998), measures the proportion of edges that exist between the neighbors of a given node in relation to the maximum number of edges that could be expected from the nodes' degree. In other words, it is the density of the subgraph constructed with the neighbors of a given node. If the given node has degree 0 or 1, then there are no possible links between its neighbors, and the local clustering is 0. It can thus be defined as

$$C_i = \frac{2c_i}{k_i(k_i - 1)}, \quad (2.34)$$

(2.35)

for undirected networks, and

$$C_i = \frac{c_i}{k_i(k_i - 1)} \quad (2.36)$$

for directed ones. The coefficient c_i , the number of links between neighbors, can be calculated employing the adjacency matrix as

$$c_i = \frac{1}{2} \sum_{j=1}^N \sum_{l=1}^N a_{ij} a_{jl} a_{li}.$$

In order to characterize network-wide properties, it is possible to define the average local clustering coefficient as

$$\langle C \rangle = \frac{1}{N} \sum_{i=1}^N C_i. \quad (2.37)$$

It is also possible to estimate the clustering tendency of the entire network. The global clustering coefficient is then defined as the ratio of closed triplets in relation to all triplets in the network. A triplet is defined as a simple path composed of three nodes connected by two edges. Alternatively, as a triangle is composed of three closed triplets (centered in each one of its vertices), the global clustering is three times the ratio of the numbers of triangles over the total number of triplets in the network:

$$\begin{aligned} C &= \frac{\sum_i \text{number of triplets centered in } i}{\text{total number of possible triplets}} \\ &= \frac{3 \times \text{number of triangles}}{\text{total number of possible triplets}}. \end{aligned}$$

This index can be calculated using the adjacency matrix as

$$C = \frac{\sum_{i=1}^N \sum_{j=1}^N \sum_{l=1}^N a_{ij} a_{jl} a_{li}}{\sum_{i=1}^N k_i(k_i - 1)} \quad (2.38)$$

From the definitions of both the local and global clustering coefficients, it is possible to show that global clustering is a weighted average of local clusterings with coefficients $k_i(k_i - 1)$. As a result, the global clustering coefficient is more sensitive to the clustering of more connected nodes than the average of the local clustering coefficient for the entire network.

2.2.3 Models of Random Graphs

When studying systems described as networks, it is relevant to understand which processes generate that particular network topology. In this sense, it is crucial to develop algorithms with explicit assumptions to generate networks. The relation between graph properties and graph generation assumptions can then be studied in detail.

In the following, we present some canonical models to generate random graphs.

2.2.3.1 $G_{N,M}$ Random Graph (Erdős and Rényi 1959)

Erdős and Rényi proposed an algorithm to sample the space of graphs with N nodes and M edges uniformly (Erdős and Rényi 1959). Given N labelled nodes, $V = \{v_1, n_2, \dots, v_N\}$, the algorithm consist in constructing the set of edges E by choosing at random one of the $N(N - 1)/2$ possible edges so that every edge has a probability $2/N(N - 1)$ of being chosen. The chosen edge is now part of the set E of the network and is removed from the available choices. A new edge is selected at random, from the $N(N - 1)/2 - 1$ edges and this process is repeated until the number of chosen edges is equal to M .

The resulting network is a random element of the ensemble of networks with N labeled nodes and M edges. For fixed N and M , the number of distinct graphs that can be constructed is $\binom{\binom{N}{2}}{M}$.

The parameters used as input in this algorithm are the total number of nodes N and edges M . It is also possible to construct a similar random graph with a different algorithm. The $G_{N,\bar{p}}$ model associates a probability \bar{p} to each possible edge in a graph with N nodes. The expected number of edges in this model is $M \approx pN(N - 1)/2$ and for a large number of nodes, these models of Random Graphs are equivalent.

2.2.3.2 Random Regular Graph

In opposition of the $G_{N,M}$ random graph described above, the random regular graph is k -regular. One possible algorithm to construct this graph is associating each node with k "half-edges". Then, a pair of "half-edges" is chosen at random and connected to form an edge. The process is repeated until no "half-edges" remain. Extra care must be taken to avoid self-loops, i.e., edges having the same node at both of its endpoints. This model corresponds to a particular case of the *configuration model*, which generates random graphs with a prescribed sequence of degrees (M. Newman 2018b).

For this category of graphs, the algorithm samples networks uniformly from the ensemble of k -regular networks. This model's input parameters are the number of nodes N and the degree k .

2.2.3.3 Preferential Attachment Graph (Barabási and Albert 1999)

The model proposed by Barabási and Albert (1999) is based on two main processes: growth and preferential attachment. These two processes combined lead to a network topology that displays distinct node degrees distributions in the form $P(k) = ak^{-\gamma}$.

The model can also be described as an algorithm: The algorithm starts from a set of k_0 isolated nodes. A new node is then added to the network at each interaction and connects with k nodes. The probability p of connection with a given node v_i , that is already in the network, is proportional to the vertex degree k_i so that $p(k_i) = k_i / \sum_j k_j$. After n iterations, the network will be composed of $k_0 + n$ vertices and nk edges. Making $k_0 = k$, the two parameters used in the definition of the model are N and k .

2.2.3.4 Small World Graph (Watts and Strogatz 1998)

The model proposed by Duncan Watts and Steven Strogatz in 1998 reproduces the effect of having a quasi-regular network with few edges between random nodes (Watts and Strogatz 1998). This small modification in the network can bring dramatic changes in the topological features of the network. The underlying algorithm starts from a k -regular lattice with N nodes. Each of the Nk edges is then rewired to other randomly selected nodes with probability p . For large values of $p \approx 1$, the network is indistinguishable from a $G_{N,M}(N, Nk)$ random network.

This model is termed "small-world" in reference to the topological property of the same name. A network is said to be a small-world network (or to satisfy the small-world property) if its average shortest path is small relative to the number of nodes in the network. More specifically, the average

shortest path length or the diameter of the network grows logarithmically (or sub-logarithmically) as the number of nodes tends to infinity.

The parameters used to construct a random network through this model are the number of nodes N , degree of initial regular network k and rewiring probability p .

2.2.4 Random Graphs' Topological Parameters

As mentioned, much of the interest in using random algorithms (or models) of networks is due to the possibility of addressing typical values for different topological properties. In the following sections, numerical simulations are performed in order to display the distribution of typical values. The expected value of density is analytically derived for previously defined topologies.

2.2.4.1 Density

For the random undirected graph, the density can be directly calculated from its definition (2.19).

However, for other network models, it is necessary to derive the number of edges in the final network. For the random regular and small-world model, this number is equivalent to that of a regular network with N nodes with degree k , therefore the number of edges are $Nk/2$ and the final density is $D = k/(N - 1)$.

In the scale free model the algorithm starts from a network with k isolated nodes, therefore the total number of edges in the final network is $(N - k) * k$ and thus the expected density is $D = 2k(N - k)/N(N - 1)$.

2.2.4.2 Estimating topological indices for network models

In figure 2.1, the distribution of the average shortest path length, diameter, global clustering, and degree assortativity for different topologies are shown. The topologies analyzed were used to construct the ecological communities (see chapter 4). Replicates of each network topology were simulated, then the topological properties were measured, and its distribution was represented as violin plots. We can observe that most of the models are associated with small values of the average shortest path length and diameter, which is a consequence of the edge configuration of the network. Small-world topologies with a low probability of reconnection display a comparatively larger value for all measured parameters. However, the networks constructed with $p = 0.1$ have the diameter and average shortest path length values equivalent to those with random and random-regular topologies. Concerning the diameter, the topologies do not present variability, as indicated by the lack of spread in the violin plot's kernel density function, except for the small-world model with a low probability of reconnection, which presents variability in network diameter.

Small-world topologies are also more clustered in comparison to the other ones. This effect is a consequence of its similarity with a perfectly regular cyclic network. Note that the clustering tends to be higher since most neighbors are connected among themselves in cyclic graphs. All topologies display degree assortativity close to zero. The degree assortativity of random regular networks is not defined since its calculation requires the degree distribution's standard deviation and it is not defined for regular networks.

2.2.4.3 Estimating Degree Distribution for Network models

As mentioned, the degree distribution is a distinctive feature of network topology. The degree distribution can serve as a basis for calculating many other properties such as degree assortativity, density, and diameter. In the figures 2.2 and 2.3 the degree distribution of the networks models presented in section 2.2.3 are shown. Note that the degree distribution of a k -regular network is

$$P(k') = \begin{cases} 1 & k' = k \\ 0 & k' \neq k \end{cases}, \quad (2.39)$$

i.e., the probability of a node having degree k is 100%, while the probability of having a different degree is 0. In accordance, the degree distribution of the random regular model is not shown.

Since the small-world algorithm starts from a k -regular network and rewrites them with individual probability p , it is expected that for small values of p this algorithm produces a network in

which most nodes have degrees close to k . The distribution of degrees for small-world networks is represented in figure 2.2, through the Complementary Cumulative Density Function (CCDF). The abrupt change of the CCDF is also indicative that most nodes have degrees close to k . For small values of p , the small-world networks had an exponential degree distribution with a small standard deviation. Considering larger values of p also gives the expected result of an increase in the distribution's standard deviation while preserving its mean. For the largest value of $p = 0.1$, it is possible to perceive the tendency of the small-world topology to produce networks that are increasingly similar to the random topology.

For the $G_{N,M}$ random graph, M edges are placed at random among N nodes. As the chance of a given node being an endpoint of a given edge is the same for all nodes, it is reasonable to expect that the average degree of a node is $\langle k \rangle \approx 2M/N$, number of endpoints divided by number of nodes. For $N = 400$ and $D = 2M/N(N - 1) = 0.1$ we have $M = 7980$ and thus $\langle k \rangle \approx 39.9$, which is precisely the mean value of the normal distribution fitted for the $G_{N,M}$ random graph simulation data.

The degree distribution of the preferential attachment network topology is remarkably distinct from the other degree distributions presented thus far. This difference is a consequence of this topology's underlying algorithm, i.e., nodes with different degrees have different probabilities of receiving further connections. According to M. Newman (2018a), if at each evolution step a node with degree c is added to the network, the probability distribution of this model is

$$P(k) = \frac{2c(c+1)}{k(k+1)(k+2)} \quad (2.40)$$

and it follows a power law as it is possible to see that $P(k') \propto 1/k^3$. In figure 2.3, the theoretical probability density function is plotted alongside data from numerical simulations. It is possible to see a remarkable agreement between predictions in data despite the effects of the limited number of nodes in the network.

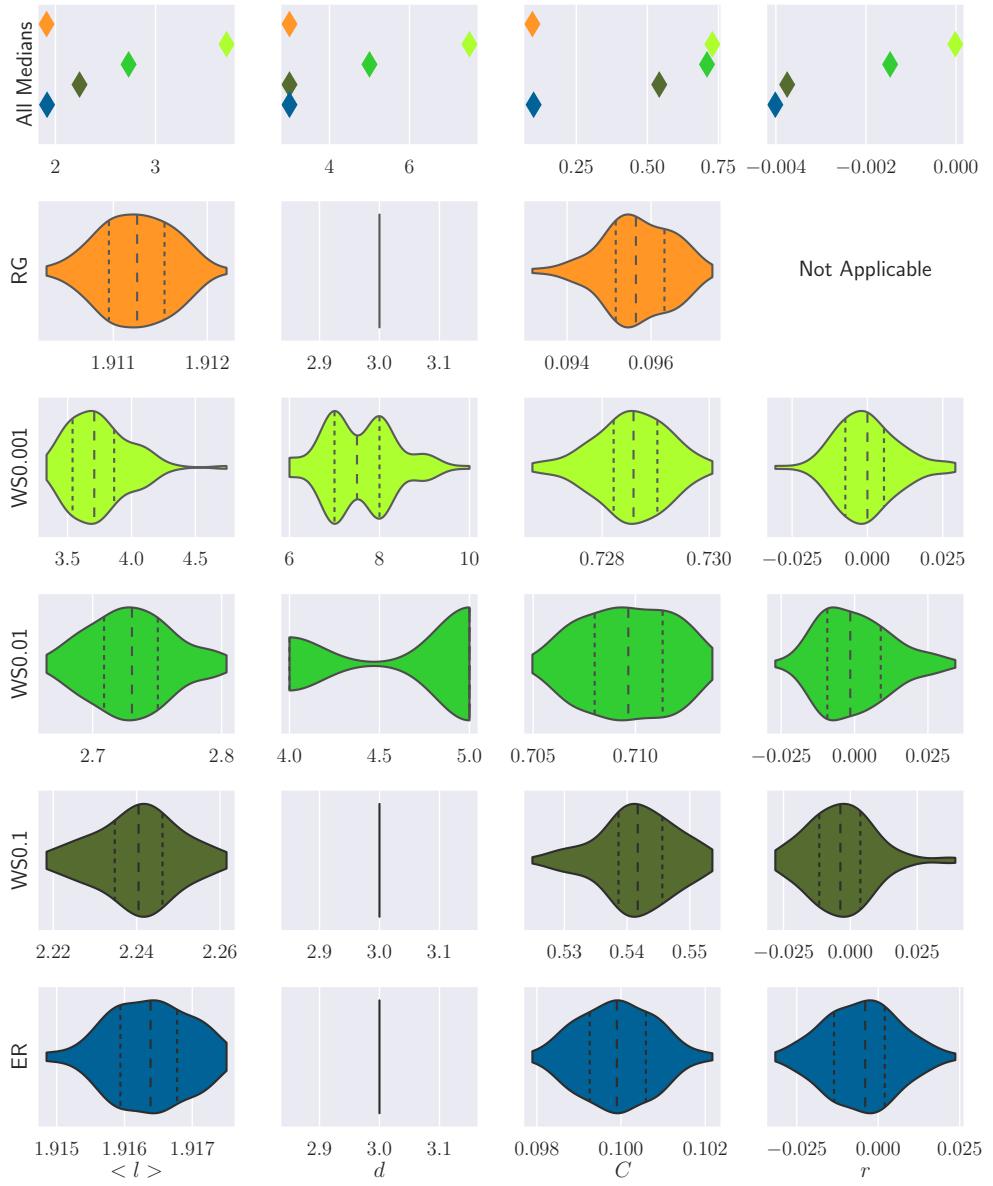


Figure 2.1: Distribution of Topological Properties for in Random Graphs: the distribution of shortest pathlength ($\langle l \rangle$), diameter (d), clustering coefficient (C), degree assortativity (r) for some of the random network models presented in section 2.2.3 is shown. The simulated topologies were $G_{N,M}$ random graph – ER (blue), random regular graph – RG (orange), and small-world model with different probabilities of reconnection: $p = 0.1$ – WS0.1 (dark green), $p = 0.01$ – WS-0.01 (green), $p = 0.001$ – WS-0.001 (light green). 100 networks of each model with 400 nodes and density $D = 0.1$ were simulated. The medians of measures for each topology are indicated by a diamond-shaped marker presented in the top row. Results aggregated by network topology are shown in the violin plots. The "violins" contour is a continuous generalization of a histogram, i.e., each distribution's kernel density estimation. The first and third quartiles are indicated by dotted vertical lines inside each distribution and dashed lines indicating the median. Note that individual plots, in general, do not share the x-axis.

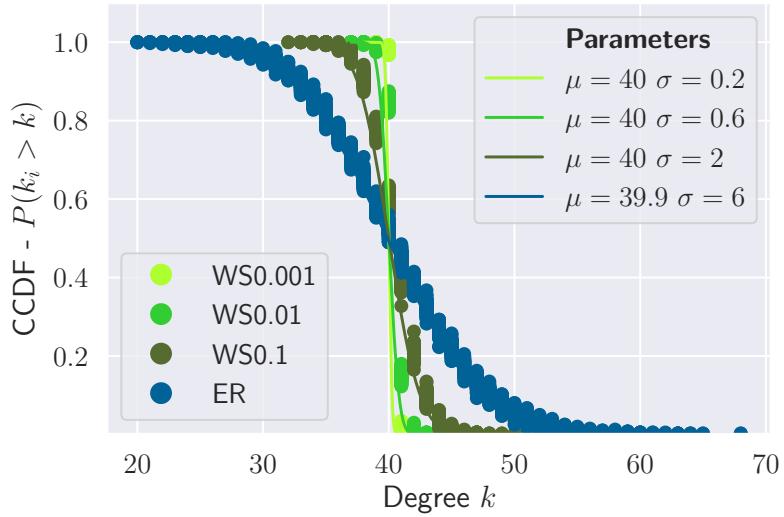


Figure 2.2: **Complementary Cumulative Density Function (CCDF) of Degrees for Watts-Strogatz (WS) and Erdős-Rényi (ER) models:** For this analysis, 100 networks of each model with 400 nodes and density $D = 0.1$ were simulated. The simulated topologies were $G_{N,M}$ random graph – ER (blue), and small-world model with different probabilities of reconnection: $p = 0.1$ – WS0.1 (dark green), $p = 0.01$ – WS-0.01 (green), $p = 0.001$ – WS-0.001 (light green). The different rewiring probabilities p in the small world model are indicated by the suffixes. Full lines indicate fitted normal distribution and points are derived from numeric simulations' data. The CCDF for all models fits normal distribution with similar mean ($\mu \approx 40$). The values of Standard deviation (σ) of nodes' degree are $\sigma = 0.2$, $\sigma = 0.6$, $\sigma = 2$ and $\sigma = 6$ for WS0.001, WS0.01, WS0.1 and ER respectively.

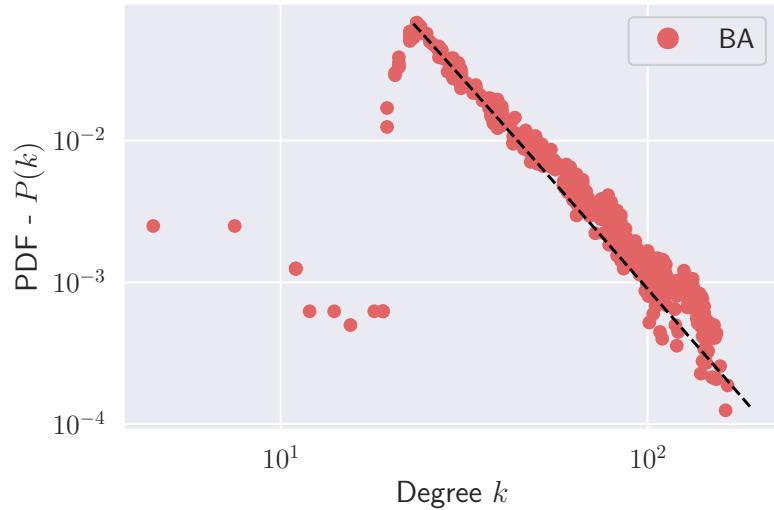


Figure 2.3: **Degree distribution of preferential attachment network topology (BA) follows a power law:** Simulated data for the BA topology are plotted alongside the theoretical expectation given by equation (2.40). The data was composed by 100 simulated networks with $N = 400$ nodes and $c = 21$ edges added at each iteration. The black dashed line represents the theoretical prediction of the probability density function , equation (2.40), with $c = 21$.

2.3 Ecological Foundations

In the following sections, elements of the ecological theory upon which this research is constructed are presented to introduce crucial concepts and nomenclature used in this work.

2.3.1 Levels of Biological Organization: Individual, Population, Community

In a nutshell, ecology is dedicated to understanding the spatial-temporal interactions of living matter with its surrounding biotic and abiotic environment. Since organisms interact with each other and with their surrounding environment across multiple spatial-temporal scales, and according to a plethora of mechanisms, investigations in ecology can be partitioned in the sense of levels of biological organization: individual organisms, the population of organisms, or community of populations.

Individual organisms are the building blocks of the population and community levels. It is related to the specific characteristics of a living being, such as its lifespan and morphology. Populations are groups of individual organisms of the same species that interact due to an interest in the same resources, usually food, shelter, and a mating partner. These interactions are denominated intra-specific since they represent processes occurring among individuals of the same population. Finally, biological communities are a group of populations that interact either positively or negatively, a process denominated inter-specific interaction since it occurs among different populations.

It is essential to acknowledge the changes in perspective when navigating across organizational levels. For instance, the lifespan of an individual organism is well defined, but it only contributes to the average lifespan at the population level. In this sense, the lifespan of a particular organism in a population may not be of much relevance, as long as the population is sufficiently abundant. Inversely, since the community comprises various interacting populations, its composition can slowly drift as the populations' densities fluctuate, new populations invade, and others go extinct at the expense of individual life histories. In the present work, we are particularly interested in understanding how interacting populations can shape ecological communities. We use populations, individuals that interact in a similar manner, as a unit of interaction. Therefore we focus on the community level of biological organization.

2.3.2 Ecological Interactions

Ecological populations and communities emerge from interactions. Precisely, ecological populations are built upon interactions between individuals of the same species, denominated as *intra-specific interactions*. In contrast, ecological communities are built from the interactions between different populations, denominated *inter-specific interactions*.

Intra-specific interactions occur due to the mutual interest of individuals of the same species in the same resources, usually related to their survival and reproduction. While individuals' reproduction fosters the population's density, the increase in density relates to an increased demand for crucial resources. Notably, resources are limited and limiting. Therefore it generates competition between the individuals in a population and ultimately sets the maximal number of individuals supported in a given environment. This maximum number is usually referred to as *carrying capacity*, usually denoted by K in mathematical models.

In an environment with abundant resources, population's growth can be estimated by the average number of offspring the species produces in a given time frame, referred to as the average growth rate (b), minus the average number of individual deaths in the same time frame, referred to as the average death rate (d). This average of births and deaths in a population is also called the population's intrinsic growth rate, usually represented as a r . As the density of the population increases, the population's growth rate decreases, which can be ascribed to a density-dependent increase in the death rate, d . The dynamics of a single population in an environment with finite resources can then be modeled by the logistic equation below (Hofbauer and Sigmund 1998)

$$\frac{dN}{dt} = rN \left(1 - \frac{N}{K}\right). \quad (2.41)$$

On the other hand, interactions are not necessarily competitive. Interspecific interactions refer to the interactions occurring between different populations. While many distinct processes mediate these interactions, their results can usually be observed in changes in the attributes of both populations, usually their abundance, density or biomass. Here we will focus on the density of populations, but abundance is at times used interchangeably.

There are three general types of interspecific interaction: competition (-/-) – when both populations suffer a decrease in their density due to the interaction; exploitation (+/-) – when one population suffer an increase in their density at the expense of the other; and mutualism (+/+) – when the interaction affects the densities of both populations positively. Notably, there are particular types of interactions in which only one population suffers the impact of the interaction: amensalism (0/-), which is considered a particular case of competition; and commensalism (0/+), which is considered a particular case of mutualism. In the present work, we are particularly interested in inter-specific pairwise interactions in which both populations are affected by their interaction, namely competition, exploitation, and mutualism.

2.3.3 Ecological Community

As an emergent property of interacting populations, the structure of an ecological community can be studied through the lenses of its compositional, relational, distribution, and functional traits.

Compositional traits are related to the actual populations that together build the community. We can study this aspect by estimating the *richness* and *diversity* of the ecological community, which refers, respectively, to the number of populations composing a community and the equitability of the abundance distribution of such populations. Relational traits refer to the interactions among the populations, specifically in the types of interactions, their strengths, and their distribution among pairs of populations. Functional traits are measures of the overall role of the community. Questions like "What is the production of a forest?" are related to the functional perspective of the community. Finally, the distribution traits are considered when addressing the characteristics of a community over a gradient, which is commonly the physical space, but can also be other abiotic variables such as temperature.

These perspectives are firmly linked, as composition changes might affect the community relations, function, and ultimately distribution. In the present work, we give special attention to the compositional and relational aspects of ecological communities. Namely, we investigate the community richness, types of interaction, the strength of interactions, and their distribution among pairs of interacting populations. These community attributes can be abstracted in what is denominated ecological community topology, encapsulating information on which and how populations interact in a given community.

2.3.4 Stability of Ecological Communities

The notion of the *ecological stability* of a community can be approached in many ways. Different aspects of stability can be investigated, such as invasibility, resistance, resilience, permanence, and robustness, which are not mutually exclusive (Pascual and Dunne 2006). The invasibility, also referred to as external stability, is related to how easily and frequently new species could invade it (Pascual and Dunne 2006; Qian and Akçay 2020).

The persistence of a population in a given region is related to its interactions with the surrounding environment and other populations. In constant biotic and abiotic conditions, a population might persist in constant density. In other scenarios, spatial-temporal environmental variability and complex biotic interactions might lead to long term cyclic fluctuations of densities. Acyclic fluctuations are also possible, e.g., the chaotic temporal evolution of densities.

As populations are the building blocks of ecological communities, the persistence of populations directly impacts the properties of ecological communities. If all the populations that compose a community persist, this is translated into *community persistence*. It is then relevant to study whether this persistence is *stable* or if any minor disturbance or variation in the conditions might disrupt the community.

Resistance is related to how much perturbation a community can endure while staying relatively constant, whereas *resilience* is related to how quickly a community returns to an original equilibrium state after a perturbation. *Permanence* requires community persistence even if a perturbation drives one of its population very close to extinction. Community *robustness* accounts for the variability of the populations in the face of environmental or populational variations, with more

robust communities displaying overall less variability in the face of environmental or populational variations.

Notably, most of the above-mentioned aspects of community stability assume the existence of an equilibrium point where all populations co-exist, and the stability of this point of equilibrium is tested by perturbing the community. In the present work, we focus on the resilience aspect of community stability.

Chapter 3

Application of Fundamental Concepts to Ecological Communities

3.1 Generalized Lotka-Volterra Model Analytical Properties

In this chapter, some analytical properties of the generalized Lotka-Volterra model (gLV) are studied and derived. Many of the techniques used are explained or based on the concepts presented in the section 2.1. While there is an attempt to rediscover many of the canonical results based on first principles, many of the developments presented here are extensively based on the literature (Hofbauer and Sigmund 1998; Takeuchi 1996; Allesina 2020). Nevertheless, the presentation of the complete map of the generalized Lotka-Volterra for two interacting populations (section 3.1.3) and the algorithm for boundedness verification of a given N -dimensional interacting matrix (3.1.4.7) are original results. In order to maintain generality, an attempt is made to use analytical developments as much as possible. However, some sections also present numerical results to illustrate or investigate properties for which analytical treatment was insufficient or impractical.

3.1.1 The Generalized Lotka-Volterra Model (gLV)

The gLV model can be formulated as

$$\frac{dX_i}{dt} = r_i X_i \left(1 + \sum_{j=1}^N a_{ij} X_j \right) \quad i = 1, 2, \dots, N, \quad (3.1)$$

in which r_i is the intrinsic growth rate of population i and the matrix of interaction is denoted by \mathbf{A} with elements a_{ij} . The elements a_{ij} of this matrix indicate the influence of population j on population's i growth rate stemming from the interaction between them. In particular, the element a_{ii} supposed negative indicates the self-limiting intraspecific competition and it is related to the population's carrying capacity as $a_{ii} = -1/K_i$. That is, the value of a_{ii} is the inverse of the density of X_i necessary to change its growth rate by 100%. Whether this change increases or decreases the effective growth rate depends on the sign of a_{ii} . The values of all elements a_{ij} can be understood similarly: a_{ij} is the inverse of the density of X_j that is necessary to change the growth rate of X_i by 100%.

Small values for a_{ij} indicate weak interaction between populations so that variances on X_j 's density have little effect on X_i 's growth rate, and vice versa. A value zero ($a_{ij} = 0$) indicate no influence of X_j upon X_i .

In this framework, the ecological interactions between two populations i, j will be given by the values of the mutual interaction coefficients a_{ij}, a_{ji} . If these coefficients are both positive, there is a mutualistic interaction between i, j while negative values indicate competitive interactions. If the coefficients have opposite signs, it indicates an exploitative interaction. For instance, if $a_{ij} > 0$ and $a_{ji} < 0$ the interaction benefits population X_i while it harms population X_j . This scenario corresponds to the case that the population X_i predares the population X_j .

In the following, we study the model for different total number of populations in the community. Starting from the simplest case of an isolated population and up to general case with an arbitrary number of N populations interacting.

3.1.2 Single Population Dynamics

For a single population, the model (3.1) can be written as

$$\frac{dX}{dt} = rX(1 + aX), \quad (3.2)$$

which is precisely the logistic model (2.41), with $N \rightarrow X$ and $K \rightarrow -1/a$.

3.1.2.1 Equilibrium Points

Setting $\frac{dX}{dt} = 0$ in equation (3.2) leads to:

$$\begin{cases} X' = 0 \\ X'' = -1/a \end{cases}. \quad (3.3)$$

Therefore there are two possible equilibrium points: one at which the density of the population is null (X') and another for which the population is at its carrying capacity (X''). X'' indicates the population density at which intraspecific competition prevents further increases in this density. It is then convenient to study what is the stability of these equilibrium points.

3.1.2.2 Stability of Equilibrium Points

For one-dimensional systems, the Jacobian is a 1–matrix and equivalent to a scalar. For the model (3.2), it can be simply calculated deriving its right-hand side by X , therefore

$$\mathbf{J} = r(1 - 2aX), \quad (3.4)$$

which is also its eigenvalue, λ . Evaluating these eigenvalues at the equilibrium points, it follows that

$$\lambda = r(1 - 2aX) \quad (3.5)$$

$$\begin{cases} \lambda|_{X=0} = r \\ \lambda|_{X=X''} = -r \end{cases}. \quad (3.6)$$

Therefore, it is possible to see that for a single population dynamics with positive growth rate $r > 0$ the stable equilibrium point is the population having a constant density $K = -1/a$, while for a population with negative growth rate $r < 0$, the equilibrium point with no density is stable. This transition occurs through a transcritical bifurcation at $r = 0$. With no intrinsic growth, the model (3.2) does not capture the dynamics of an evolving population as it predicts no change in density irrespective of the initial density of the population.

3.1.3 Two interacting populations

For a situation in which only two populations are interacting, the gLV model is

$$\begin{cases} \frac{dX_1}{dT} = r_1 X_1 (1 + a_{11} X_1 + a_{12} X_2) \\ \frac{dX_2}{dT} = r_2 X_2 (1 + a_{22} X_2 + a_{21} X_1) \end{cases}. \quad (3.7)$$

3.1.3.1 Nondimensionalization

For the model (3.7) there are now six parameters and two variables for describing the densities of the populations besides the independent variable, time. The current formulation of the model is agnostic concerning the units used to measure each of these variables. However, it is possible to simplify the mathematical formulation of the model by appropriately rescaling the variables. A useful technique to accomplish such a goal is the model's nondimensionalization (Witelski and Bowen 2015). In order to do so, we note that parameters r_i have dimensions 1/time and parameters a_{ij} have dimensions 1/density. By rescaling the populations density $X_i = K_i x_i$ and time $T = \tau t$ so that x_i, t are dimensionless variables, the model can be rewritten as

$$\begin{cases} \frac{dx_1}{dt} = \tau r_1 x_1 (1 + a_{11} K_1 x_1 + a_{12} K_2 x_2) \\ \frac{dx_2}{dt} = \tau r_2 x_2 (1 + a_{22} K_2 x_2 + a_{21} K_1 x_1) \end{cases}.$$

Defining $\tau = 1/r_1$, $K_i = -1/a_{ii}$, $a'_{ij} = a_{ij} K_j$ and $r' = \tau r_2 = r_2/r_1$ it is possible to simplify the model as

$$\begin{cases} \frac{dx_1}{dt} = x_1 (1 - x_1 + a'_{12} x_2) \\ \frac{dx_2}{dt} = r' x_2 (1 - x_2 + a'_{21} x_1) \end{cases}. \quad (3.8)$$

The model presented in the system of equations (3.8) is equivalent to model (3.7), although it has only three parameters. Note however that this parametrization does not hold for $r_1 = 0$ or $a_{ii} = 0$. In all the analytical developments we assume positive growth rates for all populations considered, i.e. $r_i > 0$ and therefore $r_1 \neq 0$. Additionally, it is ecologically reasonable that $a_{ii} \neq 0$ for all populations as intraspecific competition implies $a_{ii} < 0$.

3.1.3.2 Equilibrium Points

Setting $\frac{dx_i}{dt} = 0$ in the simplified model (3.8) results in the algebraic system of equations

$$\begin{cases} x_1 (1 - x_1 + a'_{12} x_2) = 0 \\ r' x_2 (1 - x_2 + a'_{21} x_1) = 0 \end{cases}, \quad (3.9)$$

for which the solutions are

$$E_0 = \{0, 0\} \quad (3.10)$$

$$E_1 = \{1, 0\} \quad (3.11)$$

$$E_2 = \{0, 1\} \quad (3.12)$$

$$E_3 = \left\{ \frac{1 + a'_{12}}{1 - a'_{12} a'_{21}}, \frac{1 + a'_{21}}{1 - a'_{12} a'_{21}} \right\}. \quad (3.13)$$

The last solution, E_3 , can be found under the assumption that $x_i \neq 0$, so that

$$\begin{aligned} & \begin{cases} 1 - x_1 + a'_{12} x_2 = 0 \\ 1 - x_2 + a'_{21} x_1 = 0 \end{cases} \\ \Rightarrow & \begin{cases} 1 + a'_{12} + a'_{12} a'_{21} x_1 = x_1 \\ 1 + a'_{21} + a'_{12} a'_{21} x_2 = x_2 \end{cases} \\ \Rightarrow & \begin{cases} x_1 = (1 + a'_{12}) / (1 - a'_{12} a'_{21}) \\ x_2 = (1 + a'_{21}) / (1 - a'_{12} a'_{21}) \end{cases}. \end{aligned}$$

Therefore, there are four possible equilibrium points: one solution in which all the populations are extinct; Two solutions in which one of the population is extinct and the other is at its carrying capacity density $x_i = 1 \implies X_i = K_i$; A final solution in which both populations coexist at some non-zero density.

3.1.3.3 Local Stability of Equilibria

In order to study stability of the equilibrium points, the Jacobian matrix of the model (3.8) can be calculated as

$$\mathbf{J} = \begin{pmatrix} 1 - 2x_1 + a'_{12}x_2 & a'_{12}x_1 \\ r'a'_{21}x_2 & r'(1 - 2x_2 + a'_{21}x_1) \end{pmatrix} \quad (3.14)$$

It is then possible to study the stability of each solution by calculating the eigenvalues of the Jacobian matrix evaluated at each of the equilibrium points. In the following, this analysis is presented for each solution, results are synthesized, and limit cases are discussed.

Stability of E_0 The Jacobian matrix (3.14) evaluated at the equilibrium point is

$$\mathbf{J}\Big|_{\vec{x}=E_0} = \begin{pmatrix} 1 & 0 \\ 0 & r' \end{pmatrix} \quad (3.15)$$

and its eigenvalues are $\lambda' = 1$ and $\lambda'' = r'$. Since $\lambda' > 0$ and $\lambda'' > 0$, this equilibrium point is unstable.

Stability of E_1 The Jacobian matrix (3.14) evaluated at the equilibrium point is

$$\mathbf{J}\Big|_{\vec{x}=E_1} = \begin{pmatrix} -1 & a'_{12} \\ 0 & r'(1 + a'_{21}) \end{pmatrix} \quad (3.16)$$

which has eigenvalues $\lambda' = -1$ and $\lambda'' = r'(1 + a'_{21})$. This equilibrium point is thus stable if, and only if, $a'_{21} < -1$ since $r' > 0$ by assumption. Therefore, this solution can only be stable if there is a strong negative influence of x_1 upon x_2 . A value $a'_{21} = -1$ would imply a 100% decrease on x_2 's growth rate and thus $a'_{21} < -1$ indicates that the density of population x_2 decreases over time whenever population x_1 is at its carrying capacity.

Stability of E_2 The Jacobian matrix (3.14) evaluated at the equilibrium point is

$$\mathbf{J}\Big|_{\vec{x}=E_2} = \begin{pmatrix} 1 + a'_{12} & 0 \\ r'a'_{21} & -r' \end{pmatrix} \quad (3.17)$$

which has eigenvalues $\lambda' = -r'$ and $\lambda'' = (1 + a'_{12})$.

In a similar analysis as E_1 , this solution can only be stable if $a'_{12} < -1$ which corresponds to a very intense negative influence of population x_2 upon population x_1 .

Stability of E_3 The Jacobian matrix (3.14) evaluated at the equilibrium point is

$$\mathbf{J}\Big|_{\vec{x}=E_3} = \begin{pmatrix} -x_1^* & a'_{12}x_1^* \\ r'a'_{21}x_2^* & -r'x_2^* \end{pmatrix}, \quad (3.18)$$

in which $E_3 = (x_1^* \ x_2^*)^T$. By substituting the values from equality (3.13), it follows that

$$\mathbf{J}\Big|_{\vec{x}=E_3} = \frac{1}{1 - a'_{12}a'_{21}} \begin{pmatrix} -1 - a'_{12} & a'_{12}(1 + a'_{12}) \\ r'a'_{21}(1 + a'_{21}) & -r'(1 + a'_{21}) \end{pmatrix}, \quad (3.19)$$

As seen in section 2.1.5, for a two-dimensional system, the classification of fixed point stabilities can be determined by the Jacobian matrix trace and determinant. For the Jacobian (3.19), these are

$$\text{tr } \mathbf{J} \Big|_{\vec{x}=E_3} = \frac{-1}{1 - a'_{12}a'_{21}}(1 + r' + a'_{12} + r'a'_{21}) \quad (3.20)$$

$$\det \mathbf{J} \Big|_{\vec{x}=E_3} = \frac{r'}{1 - a'_{12}a'_{21}}(1 + a'_{12})(1 + a'_{21}). \quad (3.21)$$

The characteristic polynomial of $\mathbf{J}|_{E_3}$ is $P_{E_3}(\lambda) = \det(\mathbf{J}|_{E_3} - \lambda \mathbf{I})$ so that

$$P_{E_3}(\lambda) = \det \begin{pmatrix} -1 - a'_{12} - \lambda & a'_{12}(1 + a'_{12}) \\ r'a'_{21}(1 + a'_{21}) & -r' - r'a'_{21} - \lambda \end{pmatrix} \quad (3.22)$$

$$= \lambda^2 - \underbrace{\frac{1 + r' + a'_{12} + r'a'_{21}}{1 - a'_{12}a'_{21}} \lambda}_{\text{tr } \mathbf{J}|_{E_3}} + \underbrace{\frac{r'(1 + a'_{12})(1 + a'_{21})}{1 - a'_{12}a'_{21}}}_{\det \mathbf{J}|_{E_3}}. \quad (3.23)$$

The coexistence equilibrium point is then stable if $\text{tr } \mathbf{J}|_{E_3} < 0$ and $\det \mathbf{J}|_{E_3} > 0$.

3.1.3.4 A Parameter Space Map

In order to thoroughly study the dynamics of the system at each point on the state space, it is convenient to divide the parameter space into regions with well-defined behavior. The following curves are selected for their relevance to system properties as indicated:

Line a: $a'_{12} + 1 = 0$, determines the stability of E_2 , is one of the $\det \mathbf{J}|_{\vec{x}=E_3}$ factors and numerator of x_1^* ;

Line b: $a'_{21} + 1 = 0$, determines the stability of E_1 , is one of the $\det \mathbf{J}|_{\vec{x}=E_3}$ factors, and numerator of x_2^* ;

Line c: $1 + a'_{12} + r'(1 + a'_{21}) = 0$, numerator of $\text{tr } \mathbf{J}|_{\vec{x}=E_3}$;

Curve d: $a'_{12}a'_{21} - 1 = 0$, present in the equilibrium point E_3 and denominator of $\det \mathbf{J}|_{\vec{x}=E_3}$ and $\text{tr } \mathbf{J}|_{\vec{x}=E_3}$.

Particular cases are the intersection of such curves,

Point A: $(-1, -1)$, $A = \mathbf{a} \cap \mathbf{b} \cap \mathbf{c} \cap \mathbf{d}$;

Point B: $(-r, -1/r)$, $B = \mathbf{c} \cap \mathbf{d}$.

The curves defined above divide the parameter space into up to eleven regions. In the interior of each of these regions, the determinant and trace of the Jacobian of the system evaluated at the equilibrium E_3 have a well-defined sign, i.e., their sign does not change. Each of these regions is defined and studied in table 3.1.

The relevant curves, points, and regions limited by them are represented in figure 3.1. The effect of changing r' is simply a change on line c's angular coefficient, with $r' \rightarrow 0 \implies \mathbf{c} \rightarrow \mathbf{a}$ and $r' \rightarrow \infty \implies \mathbf{c} \rightarrow \mathbf{b}$. Note that by the conditions presented on table 3.1 regions 6 and 7 cannot both exist for the same value of r' , i.e. region 6 has non-zero area if $r' > 1$, while region 7 requires $r' < 1$.

In the following, the behavior of the system at each of the lines and points of interest is individually studied.

Line a: At this line, $a'_{12} = -1$ and thus $E_3 = ((1 - 1)/(1 + a'_{21}), (1 + a'_{21})/(1 + a'_{21})) = (0, 1) = E_2$, i.e. there is a convergence of equilibrium points. The eigenvalues for this equilibrium point are $\lambda' = 0$ and $\lambda'' = -r' < 0$. The null eigenvalue corresponds to a central manifold, while the negative eigenvalue indicates that the flow is stable on the direction associated with λ'' eigenvalue.

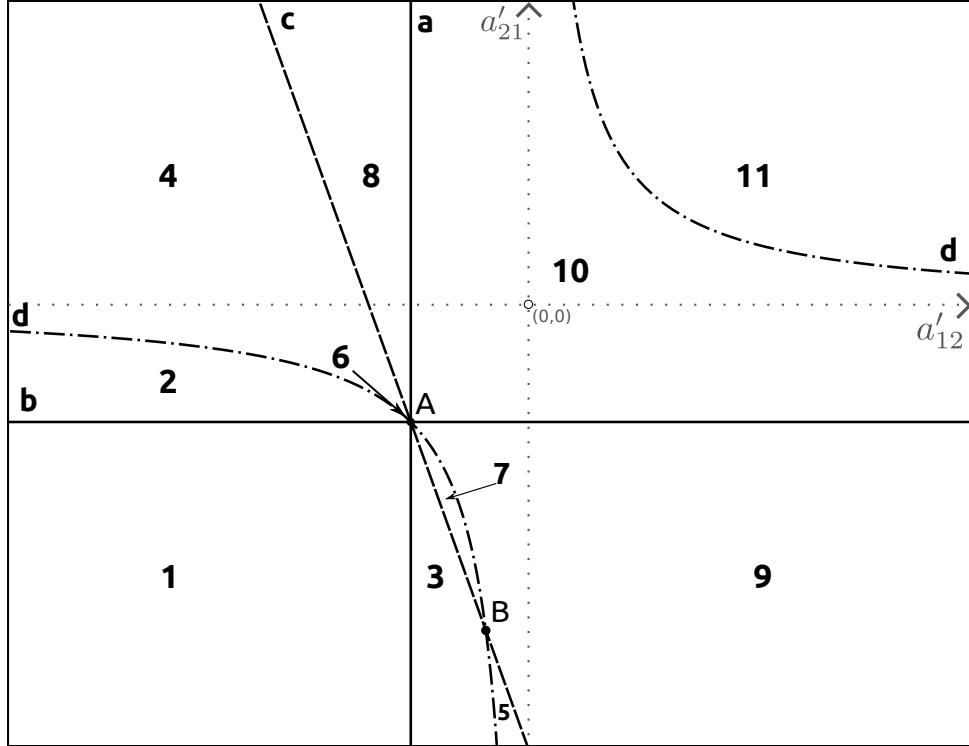


Figure 3.1: **Distinct regions and curves on the two-dimensional generalized Lotka-Volterra model (gLV-2D) $a'_{12} \times a'_{21}$ parameter space:** This diagram displays the curves (bold lowercase letters), points (uppercase letters), and regions (numbers) associated with the dynamical behavior of the system (3.7). The determinant and trace of the Jacobian of the referred system do not change sign in the interior of any of the specified regions. The curves are relations obtained when equating factors present in the determinant and trace of the Jacobian to zero. The schematic diagram was constructed for $r' = 0.2$. The axes are presented with dotted lines for reference and are not borders of any region. All elements in the figure are defined in the text.

Line b: In a analogous way to line **a**, at this line $a'_{21} = -1$ and thus $E_3 = (1, 0) = E_1$. The eigenvalues associated with the Jacobian evaluated at this equilibrium point are $\lambda' = 0$ and $\lambda'' = -1$, indicating the presence of a central manifold and a stable hyperbolic manifold.

Line c: at this line, $a'_{12} + 1 + r'(a'_{21} + 1) = \text{tr } \mathbf{J}|_{\vec{x}=E_3} = 0$ and thus the equilibrium point E_3 can only be neutrally stable or a saddle point. With the additional condition that $\det \mathbf{J}|_{\vec{x}=E_3} > 0$, the equilibrium point goes through a Hopf bifurcation and the eigenvalues become $\vec{\lambda} = (a + bi, a - bi)$. Nevertheless, this bifurcation can only affect the dynamics of the populations if it occurs in a region of the parameter space for which $E_3 > \vec{0}$ since negative densities are unmeaningful.

Observing that at this line $a'_{12} + 1 = -r'(a'_{21} + 1)$, it is possible to calculate the Jacobian's determinant evaluated at the coexistence equilibrium point as

$$\begin{aligned} \det \mathbf{J} \Big|_{\vec{x}=E_3} &= \frac{-r'^2(a'_{21} + 1)^2}{(a'_{21} + 1)(r'a'_{21} + 1)} \\ &= \frac{-r'^2(a'_{21} + 1)}{(r'a'_{21} + 1)}. \end{aligned} \quad (3.24)$$

The determinant is in the region $(a'_{21} + 1)/(r'a'_{21} + 1) < 0$, i.e. when the numerator and denominator have opposite signs. This occurs for the region $-1 < a'_{21} < -1/r'$ if $r' < 1$ or $-1 > a'_{21} > -1/r'$ if $r' > 1$. In the figure 3.1, this corresponds to the line segment \overline{AB} . In either situation, $1 - a'_{12}a'_{21} = (1 + a'_{21})(1 + r'a'_{21}) < 0$, and thus

$$\begin{aligned}
x_1^* &= \frac{1 + a'_{12}}{1 - a'_{12}a'_{21}} \\
&= -\frac{r'}{1 + r'a'_{21}} < 0 \quad \text{if} \quad -\frac{1}{r'} < a'_{21} \\
x_2^* &= \frac{1 + a'_{21}}{1 - a'_{12}a'_{21}} \\
&= \frac{1}{(1 + r'a'_{21})} < 0 \quad \text{if} \quad d'_{21} < -\frac{1}{r'},
\end{aligned}$$

and therefore at least one of the components of E_3 is negative. This equilibrium point is not inside an ecologically meaningful region when the conditions for a Hopf bifurcation are satisfied.

Curve d: At curve **c**, $a'_{12}d'_{21} = 1$, the value of E_3 as well as the trace and determinant of the Jacobian evaluated at this point diverge. In order to study the equilibrium of the system in this curve, it is possible to solve (3.9) with the additional assumption that $a'_{12} = 1/a'_{21} = a$. Therefore

$$\begin{aligned}
&\begin{cases} x_1(1 - x_1 + ax_2) = 0 \\ r'x_2\left(1 - x_2 + \frac{1}{a}x_1\right) = 0 \end{cases} \\
x_1, x_2 \neq 0 \implies &\begin{cases} 1 - x_1 + ax_2 = 0 \\ 1 - x_2 + \frac{1}{a}x_1 = 0 \end{cases} \\
\implies &\begin{cases} x_1 = 1 + ax_2 \\ x_2 = 1 + \frac{1}{a} + \frac{a}{a}x_2 \end{cases} \\
\implies &\frac{1}{a} = a = -1 \\
\implies &x_1 + x_2 = 1.
\end{aligned}$$

It is possible to see that there is a line $x_1 + x_2 = 1$ at which the flow of the system is stationary. The Jacobian (3.14), evaluated at this line then becomes

$$\mathbf{J}\Big|_{x_2=1-x_1} = \begin{pmatrix} -x_1 & -x_1 \\ -r'(1-x_1) & -r'(1-x_1) \end{pmatrix}, \quad (3.25)$$

and its eigenvalues are $\lambda' = 0$ and $\lambda'' = -x_1 - r'x_2 < 0$. The eigenvector $\vec{e}_0 = (a, b)^T$ associated with λ' can be found by solving the eigenvector equation $\mathbf{J}\Big|_{x_2=1-x_1} \vec{e}_0 = \lambda' \vec{e}_0$, therefore

$$\begin{aligned}
&\mathbf{J}\Big|_{x_2=1-x_1} \vec{e}_0 = \vec{0} \\
\begin{pmatrix} -x_1 & -x_1 \\ -r'x_2 & -r'x_2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\
\begin{cases} -x_1(a+b) = 0 \\ -r'x_2(a+b) = 0 \end{cases} &= 0 \\
\implies a &= -b \\
\implies \vec{e}_0 &= \begin{pmatrix} 1 \\ -1 \end{pmatrix}.
\end{aligned}$$

This vector indicates that the line of neutral stability has an angular coefficient -1 , and for this particular set of parameters, the line $x_1 + x_2 = 1$ corresponds to a line of asymptotically stable equilibria.

The analysis of the flux behavior on the state space for the particular choice of parameters in line **c** lead to the conclusion that the equilibrium $E_3 = (x_1^*, x_2^*) \neq (0, 0)$ is only possible under the additional assumptions that $a'_{12} = a'_{21} = -1$. All the other points in this curve on the parameter space are associated with a state space flux for which there is no equilibrium point with coexistence.

Point A As indicated in the study of curve **d**, the point $(a'_{12}, a'_{21}) = (-1, -1)$ in the parameter space is associated with a particular dynamics. In this configuration, interspecific competition is symmetric and numerically identical to the intraspecific competition, suggesting this could be the case of two populations that are ecologically similar or the same population being artificially subdivided in two populations with potentially different basal growth rates. If we assume the growth rates of the two populations to be equal, i.e. $r_1 = r_2 \implies r' = 1$, both equations are completely equivalent and by taking $x = x_1 + x_2$ it follows that

$$\begin{aligned}\frac{dx}{dt} &= \frac{dx_1}{dt} + \frac{dx_2}{dt} \\ &= x_1(1 - x_1 - x_2) + x_2(1 - x_1 - x_2) \\ &= (x_1 + x_2)(1 - (x_1 + x_2)) \\ \frac{dx}{dt} &= x(1 - x)\end{aligned}$$

Regions 1 - 11 The regions limited by the relevant lines and points highlighted have well defined dynamical behavior. In order to characterize these dynamics, the stability of all equilibrium points and the feasibility¹ of the coexistence solution E_3 for each region are presented in table 3.1. These regions are indicated by the numbers in figure 3.1, which depicts a section of the parameter space for constant r' .

Indeed, considering the conditions for the local stability of all equilibrium points, it is possible to investigate which ones are going to be stable at some values of the parameters a'_{12} , a'_{21} and r' . The equilibrium points E_0 , E_1 and E_2 do not depend on any parameters, and are always in region $\mathbb{R}_{\geq 0}^2$ for which the densities are ecologically reasonable. For equilibrium point E_3 the densities depend on the parameters and therefore it is important to determine the sign of populations' densities at each region. Indeed, the sign of the densities is indicated on table 3.1 under column $\text{sgn}(E_3)^T$.

All stable equilibria for parameters inside a given parameter space region are listed under column "Stable". Note that the equilibrium point E_3 might be associated with negative densities and thus be in an unreasonable region of the state space. In particular, for regions 2 and 3, the equilibrium E_3 has a negative component.

In order to illustrate the dynamics in each region of the parameter space, it is convenient to study the state space portraits associated with each set of parameters. Figure 3.2 displays the state space portrait for every region illustrated in figure 3.1.

In this figure, the lines indicating null net growth for each population are presented. These lines are named isoclines, and for the model (3.8) it is possible to show that each population has two isoclines, one of which is the line $x_i = 0$. Under the assumption that $x_i \neq 0$, it is possible to find the other isoclines as

$$x_i = 1 + a'_{ij}x_j \quad i, j \in \{1, 2\}. \quad (3.26)$$

As also depicted in figure 3.2, there are five distinct possibilities for the state space diagram in the two-dimensional generalized Lotka-Volterra system. The system might have one stable equilibrium point with only of the populations persisting at its carrying capacity as in regions 2 - 9 of the parameter space and depicted in subfigures **A** - **D**. In this scenario, the system converges for the state with only one of the populations at its carrying capacity for any initial condition. A scenario in which the growth of both populations in isolation is stable but their coexistence is not, corresponding to region 1 of the parameter space, is represented in subfigure **E**. In this scenario, the isoclines intercept in an unstable node with positive coordinates for both populations, which

¹A solution of an ecologically-motivated dynamical system is said to be feasible if all the densities on this solution are positive.

n^0	Region	$(\text{tr } \mathbf{J}, \det \mathbf{J}, \Delta) _{\vec{x}=E_3}$	$\text{sgn}(E_3)^T$	E_3 Stab.	Stable
1	$a'_{12} < -1$ and $a'_{21} < -1$	$(-, -, +)$	$(+, +)$	Saddle Point	$[E_1, E_2]$
2	$a'_{12} < -1$ and $-1 < a'_{21} < \min\left(\frac{1}{a'_{12}}, -\frac{1+r'+a'_{12}}{r'}\right)$	$(-, +, \pm)$	$(+, -)$	Stable	$[E_2, E_3]$
3	$a'_{12} <$ $\min(1/a'_{21}, -(1+r'+r'a'_{21}))$ and $-1 < a'_{12}$ and $a'_{21} < -1$	$(-, +, \pm)$	$(-, +)$	Stable	$[E_1, E_3]$
4	$a'_{12} < -1$ and $1/a'_{12} < a'_{21} < -(1+r'+a'_{12})/(r')$	$(+, -, +)$	$(-, +)$	Saddle Point	$[E_2]$
5	$1/a'_{21} < a'_{12} < -(1+r'+r'a'_{21})$ and $a'_{21} < -1$	$(+, -, +)$	$(+, -)$	Saddle Point	$[E_1]$
6	$a'_{12} < -1$ and $-(1+r'+a'_{12})/(r') < a'_{21} < 1/a'_{12}$	$(+, +, \pm)$	$(+, -)$	Unstable	$[E_2]$
7	$-(1+r'+r'a'_{21}) < a'_{12} < 1/a'_{21}$ and $a'_{21} < -1$	$(+, +, \pm)$	$(-, +)$	Unstable	$[E_1]$
8	$a'_{12} < -1$ and $\max(1/a'_{12}, -(1+r'+a'_{12})/(r')) <$ a'_{21}	$(-, -, +)$	$(-, +)$	Saddle Point	$[E_2]$
9	$\max(1/a'_{21}, -(1+r'+r'a'_{21})) <$ a'_{12} and $a'_{21} < -1$	$(-, -, +)$	$(+, -)$	Saddle Point	$[E_1]$
10	$-1 < a'_{12}$ and $-1 < a'_{21} < 1/a'_{12}$	$(-, +, \pm)$	$(+, +)$	Stable	$[E_3]$
11	$0 < a'_{12}$ and $1/a'_{12} < a'_{21}$	$(+, -, +)$	$(-, -)$	Saddle Point	$[\emptyset]$

Table 3.1: **Dynamical behavior at parameter space regions:** The behavior of each region highlighted on diagram 3.1 is presented and studied. Each region is defined by a set of inequalities on a'_{12} and a'_{21} , and the trace, determinant and discriminant $\Delta = (\text{tr } \mathbf{J})^2 - 4 \det(\mathbf{J})$ of the Jacobian evaluated at the coexistence equilibrium are presented. This equilibrium point's feasibility is presented using the sign ($\text{sgn}(E_3)^T$) of its components, and the stability is classified. The stability of E_3 is presented in the column $E_3\text{Stab.}$. As presented in the text, solution E_0 is always an unstable node and solutions E_1, E_2 are always nodes for which the stability is indicated at the last column of the table.

lie on the limiting border between the basins of attraction of the two stable equilibria. It is also possible that the coexistence of both populations is stable so that all trajectories passing through a point with non-zero densities for both populations converges to the coexistence equilibrium point as seen in subfigure F. Subfigure H corresponds to the particular case in which $a'_{12} = a'_{21} = -1$ and displays a line of asymptotically stable equilibria.

In subfigure G, the populations display a tendency to grow indefinitely. This is a consequence of the fact that mutual influences between populations are positive $a'_{12}, a'_{21} > 0$ and that isoclines do not intercept each other at the ecologically meaningful subspace \mathbb{R}_+^2 . Consider that the two isoclines are defined by equation (3.26) so that

$$\begin{aligned} & \begin{cases} l': & x_1 = a'_{12}x_2 + 1 \\ l'': & x_2 = a'_{21}x_1 + 1 \end{cases} \\ \Rightarrow & \begin{cases} l': & x_1 = a'_{12}x_2 + 1 \\ l'': & x_1 = x_2/a'_{21} - 1/a'_{21} \end{cases}. \end{aligned}$$

Therefore, as $a'_{12}, a'_{21} > 0$ the condition for existence of an unbounded growth region is that the angular coefficient of line l' is larger or equal to the angular coefficient of line l'' , i.e.,

$$a'_{12} \geq \frac{1}{a'_{21}} \\ \implies a'_{12}a'_{21} \geq 1,$$

which is precisely the union of the conditions defining the region 11 and its border, the positive portion of curve **d**.

3.1.3.5 Ecological Interpretation

The dynamical properties of the generalized two-dimensional Lotka-Volterra model presented thus far are intrinsically related to the types of ecological interaction in the system. It is then possible to provide an ecological interpretation of the dynamical results. Indeed, figure 3.3 illustrates this interplay by presenting the ecological regions corresponding to cooperative, exploitative, and competitive interactions in a green, yellow, and red overlay. In the same figure, the regions of distinct dynamical behaviors are also indicated, i.e., stable coexistence, competitive exclusion, bistability, and unlimited (infinite) growth by a dotted, striped, checkboard, and wavy overlay, respectively.

There is never the extinction of any population for cooperative dynamics, but the model predicts unlimited growth for intense cooperation between both populations (region 11 in figure 3.1). The unlimited growth is unreasonable from an ecological standpoint and is a consequence of the modeling procedure as it signals that interactions between populations suppressed the intraspecific competition. This type of dynamics occurs for systems in which the per capita growth benefit stemming from cooperation has no upper bound and depends linearly on the cooperative population density. As some authors have pointed out, this unlimited growth can be avoided by incorporating saturation on cooperation benefits (Qian and Akçay 2020).

The generalized Lotka-Volterra model also indicates that strong exploitative interactions might drive the exploited population to extinction. The extinction of the exploited population occurs whenever the exploitation interaction can reduce the net per capita growth rate of the exploited population to zero when the exploiting population is at its carrying capacity, i.e., $a'_{ij} < -1$ for exploited population i and exploiting population j . The conditions for the extinction of one of the populations interacting through exploitation are satisfied in the intersection of regions 4, 5, 8 and 9 in figure 3.1 in the second and fourth quadrants of the $a'_{ij} \times a'_{ji}$ plane.

Suppose the interaction is not as intense, so that $a'_{ij} > -1$. In that case, the exploited population has a positive net per capita growth rate if its density is small, and the exploiting population is at its carrying capacity. In other words, the exploited population can invade a system with population j at its carrying capacity. As this is symmetrical for both populations, the condition for coexistence is that both populations are able to invade a system in which just the other population is at its carrying capacity, $a'_{12} > -1$ and $a'_{21} > -1$ (intersection of the second and fourth quadrants and region 10 in figure 3.1).

For two competing populations with intense competitive interaction $a'_{12} < -1$ and $a'_{21} < -1$, there is mutual competitive exclusion. Under these circumstances, no population can invade a system in which the other is at its carrying capacity. The final state of the community will then be dependent on the initial densities of both populations. The mutual competitive exclusion scenario corresponds to region 1 in figure 3.1.

If one population i is a strong competitor as indicated by an intense negative influence upon the competing population j , $a_{ji} < -1$ but the other population does not exert such strong competitive regulation, the situation is analogous to the intense exploitation scenario. The strong competitor will be dominant and drive the weak competitor to extinction in all scenarios for which both populations are initially present. This scenario corresponds to the intersection of regions 2 to 9 with the third quadrant of the $a'_{ij} \times a'_{ji}$ plane, i.e., with $a'_{ij} < 0$ and $a'_{ji} < 0$.

Nevertheless, both competing populations can also coexist if the interaction is relatively weak. Specifically, both populations should be able to invade a system in which the other is at the carrying capacity ($a'_{ij} > -1$ and $a'_{ji} > -1$). This condition is satisfied in the portion of region 10 that lies in the third quadrant. In general, many outcomes are possible in a competitive scenario, and the particular result depends on the intensities of the interspecific interactions between both populations.

Note that the definition of the gLV model (3.7) is best suited for positive intrinsic growth rates, $\vec{r} > \vec{0}$. The model still holds for negative growth rates, but it loses much of the ecological interpretability in such scenario, as the meaning of interaction variables a_{ij} would depend on r_i : For positive intrinsic growth rates, a_{ij} is the inverse of j 's density necessary to change i 's intrinsic growth rate by 100% positively or negatively in accordance to a_{ij} sign; For negative growth rates the sign is changed and thus if $r_i < 0$ a positive value of a_{ij} is associated with a decrease in i 's population and a negative value of a_{ij} is beneficial to i 's population. This mismatch in the meaning of different interaction coefficients hinders interpretation and parametrization of this model for mixed positive and negative growth rates.

The specific growth rates of the populations are not determinant for the overall dynamics of the system. For instance, consider the situation in which one population has an intrinsic growth much larger than the other. Under this assumption, two limiting cases are possible. Considering $r_1 \gg r_2 \implies r' \rightarrow 0$, the line **c**, defined as $a'_{12} + 1 + r'(a'_{21} + 1) = 0$ coincides with line **a**, and the parameter space regions 3, 5, 6 and 8 cease to exist. Conversely, if $r_1 \ll r_2 \implies r' \rightarrow +\infty$, than line **c** coincides with line **b**, as $\lim_{r' \rightarrow \infty} ((a'_{12} + 1)/r' + a'_{21} + 1) = a'_{21} + 1$ and under these assumption parameter space regions 2, 4, 7 and 9 cease to exist. Despite these changes, regions 2, 4, 6, and 8 share the same properties of the state space portrait, and the same occurs in regions 3, 5, 7, and 9 as it can be seen on table 3.1. Therefore even in these limiting cases, the intrinsic growth rates do not have a sizeable structural impact.

Nevertheless, it is relevant to state that this is also a consequence of model parameterization and the assumption that coefficients a'_{ij} indicate the proportion of change on the intrinsic growth rate of population i due to the presence of populations j when it is at its carrying capacity.

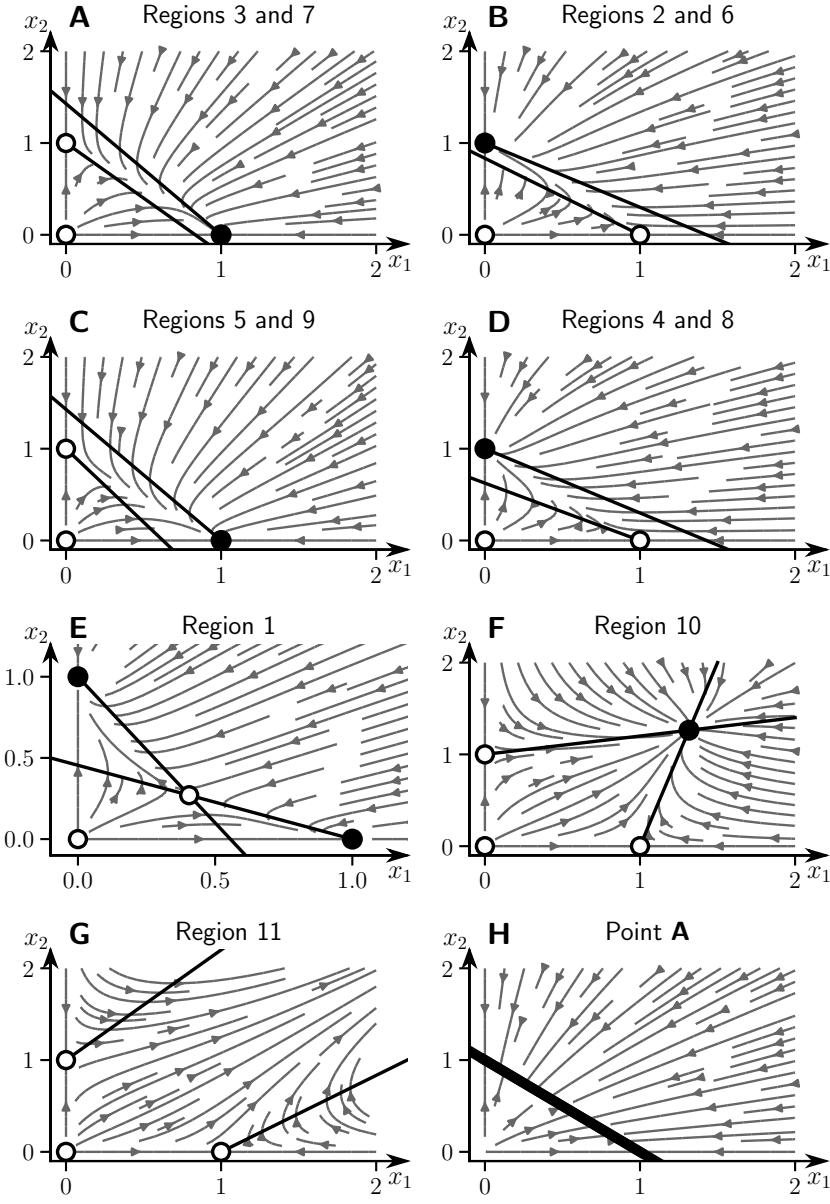


Figure 3.2: State space portrait for all regions of the two dimensional gLV parameter space: A typical state space portrait for all regions of the gLV-2D defined in table 3.1 and pictured in diagram 3.1 is shown, with the equilibrium points indicated by white and dark dots for unstable and stable solutions, respectively. Oriented grey curves correspond to subsets of possible trajectories of the system for different initial conditions. The bold lines are isoclines, curves indicating a region in which one of the populations has null net growth. In each plot, isoclines $x_i = 0$ were omitted, and it is possible to identify which isocline corresponds to each population by observing to which fixed-point it is connected. **A, B, C, D** display a state space typical of competitive exclusion, population x_1 is the better competitor in diagrams **A** and **B** while population x_2 drive its competitor to extinction in diagrams **C** and **D**. **E** indicates a bi-stable situation, in which both populations can exclude the other competitively, depending on the initial densities of both populations. **F** indicates a coexistence scenario in which both populations stabilize at a feasible equilibrium. **G** exemplifies a region of the parameter space for which no equilibrium point is stable, and there is an infinite growth of both populations. **H** represents the state space for the particular choice of parameters corresponding to point A of the parameter space, and a bold line indicates both isoclines and the set of equilibrium points for which $x_1 + x_2 = 1$. Interaction parameters (a'_{12}, a'_{21}) for each subplot were: **A** $(-0.7, -1.2)$, **B** $(-1.2, -0.7)$, **C** $(-0.7, -1.6)$, **D** $(-1.6, -0.7)$, **E** $(-2.2, -1.8)$, **F** $(0.25, 0.2)$, **G** $(1.2, 1.2)$ and **H** $(-1, -1)$. All graphs were constructed with a relative growth rate $r' = 0.8$.

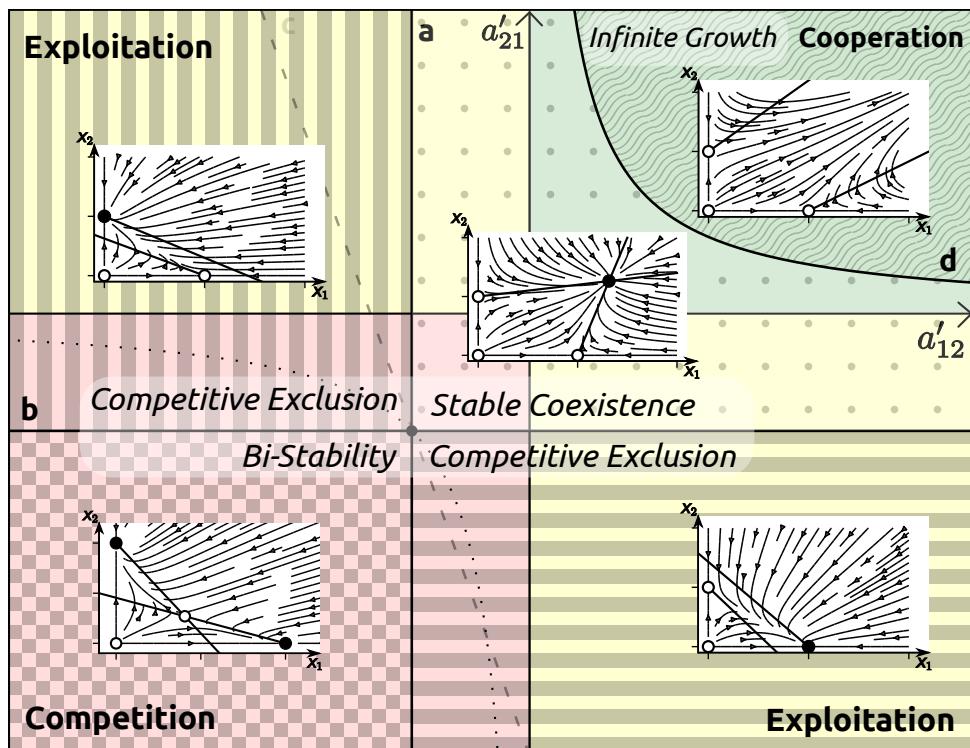


Figure 3.3: **Ecological interpretation of two-dimensional generalized Lotka-Volterra model parameter space:** The ecological regions corresponding to cooperative, exploitative and competitive interactions are respectively shown in green, yellow, and red. The parameter space regions corresponding to stable coexistence, competitive exclusion, bi-stability, and unlimited (infinite) growth are respectively indicated by a dotted, striped, check-board, and wavy overlay. In each of these regions, an inset with a corresponding portrait of the state space is shown. The state space portraits were taken from figure 3.2. Note that different overlays can exist for the same background color, indicating that the same ecological interaction might correspond to distinct topological properties of the state space and vice-versa.

3.1.4 Ecological Community with N Interacting Populations

Despite the completeness possible in studies with one and two interacting populations, real ecological communities are seldom straightforward. The richness of many ecosystems implies that the modeling of ecological communities is a high dimensional endeavor, thus requiring caution in generalizing results and principles derived from models with low dimensionality. Indeed, it is well known that dynamical systems with many variables can display chaos and other types of dynamics that are not possible in two dimensions (Fiedler-Ferrara and Prado 1994).

Due to the intrinsic complexity of studying dynamical systems in many dimensions, no attempt is made to present a complete picture of the parameter space of the generalized Lotka-Volterra model (Takeuchi 1996). In the following, some essential properties are shown, such as forward invariance, the existence of equilibrium points and their local and global stability, the possible asymptotic dynamics in the model, and conditions for boundedness of the solutions. Additionally, the time average of population densities in a solution of the model is shown to be equal to the densities of the equilibrium point, even if this solution is not stable.

For many dimensions, it is often convenient to make use of a vectorial formulation of the model (3.1) as:

$$\frac{d\vec{X}}{dt} = \text{diag}(\vec{X} \circ \vec{r}) (\vec{1} + \mathbf{A}\vec{X}) \quad (3.27)$$

in which $(\vec{A} \circ \vec{B})_i = A_i B_i$ is the Hadamard product, i.e. indicate multiplication of the elements with same index.

3.1.4.1 Forward Invariance

The generalized Lotka-Volterra model, as defined by the system of equations (3.1), assumes non-negative values for the X_i variables. It is then worth to show that no trajectory starting in the ecologically meaningful region $\mathbb{R}_{\geq 0}^N$ entails points in non-meaningful region $\mathbb{R}^N \setminus \mathbb{R}_{\geq 0}^N$ and vice versa. In fact, this is guaranteed if no trajectory can ever cross the boundaries between regions, i.e. if and only if $\frac{dX_i}{dt}|_{X_i=0} = 0 \quad \forall i$. Therefore

$$\begin{aligned} \left. \frac{dX_i}{dt} \right|_{X_i=0} &= 0 \\ \implies 0 \cdot r_i (\vec{1} + \mathbf{A}\vec{X}) &= 0 \quad \forall i \\ 0 &= 0 \quad \square \end{aligned} \quad (3.28)$$

Thus, the forward invariance of the gLV model for N dimensions is guaranteed.

3.1.4.2 Equilibrium Points

As indicated in subsection 3.1.4.1, $\vec{X} = \vec{0}$ is always a equilibrium point. Under the assumptions $X_i \neq 0 \quad \forall i$ and $r_i \neq 0 \quad \forall i$ it is possible to calculate the equilibrium points, denoted by \vec{X}^* , for model (3.27) as

$$\begin{aligned} \text{diag}(\vec{X}^*) \text{diag}(\vec{r}) (\vec{1} + \mathbf{A}\vec{X}^*) &= \vec{0} \\ \implies \vec{1} + \mathbf{A}\vec{X}^* &= \vec{0} \\ \mathbf{A}\vec{X}^* &= -\vec{1} \\ \vec{X}^* &= -\mathbf{A}^{-1}\vec{1}. \end{aligned} \quad (3.29)$$

Therefore, this solution is possible if the interaction matrix \mathbf{A} is non-singular and if $-\mathbf{A}^{-1}\vec{1} > \vec{0}$, the equilibrium point is said to be feasible. In general, in terms of the index-based formulation of the model (3.1), there are two possible solutions for the equation $\frac{dX_i}{dt} = 0$ for a given population i

$$\frac{dX_i}{dt} = 0 \implies \begin{cases} X_i^* = 0 & \text{or} \\ \sum_j a_{ij} X_j^* = -1 \end{cases}. \quad (3.30)$$

Consider that for a given equilibrium point $X_i > 0$ for $i \in \{1, \dots, l\}$ and $X_j = 0$ for $j \in \{l+1, \dots, N\}$. Under these assumptions, it is possible to construct a reduced interaction matrix

$$(\mathbf{A}')_{ij} = a_{ij} \quad i, j \in \{1, \dots, l\} \quad (3.31)$$

and denoting $(X_1, \dots, X_l)^T = \vec{X}'$ it follows that

$$\vec{X}'^* = -\mathbf{A}'^{-1} \vec{1}. \quad (3.32)$$

Therefore, all the equilibrium points of the N dimensional gLV model can be found by inverting a reduced matrix obtained from \mathbf{A} removing the columns and rows correspondent to variables that are zero in such equilibrium point. Since each of the N variables can have a zero and a non-zero solution for every combination of all other variables, the model displays at most 2^N distinct solutions. It is expected that some of them have at least one variable $X_j^* < 0$ and thus not all the solutions are in the ecologically relevant subspace $\mathbb{R}_{\geq 0}^N$.

3.1.4.3 Linear Approximation of the Flow - Jacobian Matrix

In order to investigate the stability of equilibrium points of system (3.27) for N populations, it is necessary to perform the linearization of the system around this equilibrium. As indicated in equation (2.6), it is possible to find a set of linear differential equations

$$\frac{d}{dt} \vec{X}' = \mathbb{J} \vec{X}', \quad (3.33)$$

which are locally equivalent to the general nonlinear system. For the generalized Lotka-Volterra model (3.1), the Jacobian matrix can be calculated as

$$\begin{aligned} J_{ij} &= \frac{\partial}{\partial X_j} \frac{dX_i}{dt} \\ \implies J_{ij} &= \begin{cases} r_i \left(1 + a_{ii} X_i + \sum_j a_{ij} X_j \right) & i = j \\ r_i a_{ij} X_i & i \neq j \end{cases}. \end{aligned} \quad (3.34)$$

If the Jacobian matrix is being evaluated at the neighborhood of the internal equilibrium (3.29), it follows that $\sum_j a_{ij} X_j^* + 1 = 0$, thus the Jacobian for this equilibrium point can be defined as

$$J_{ij} \Big|_{\vec{X}^*} = r_i a_{ij} X_i^* \quad \forall i, j. \quad (3.35)$$

Instead, if the Jacobian is being evaluated at the equilibrium point (3.32), with some populations virtually extinct, it follows that

$$J_{ij} \Big|_{\vec{X}'^*} = \begin{pmatrix} \mathbf{J}' & \mathbf{K} \\ 0 & \text{diag} \left(r_i \left(\sum_j a_{ij} X_j^* + 1 \right) \right) \end{pmatrix}, \quad (3.36)$$

in which \mathbf{J}' is the $l \times l$ matrix $(\mathbf{J}')_{ij} = \frac{\partial}{\partial X_j} X_i \quad i, j \in \{1, \dots, l\}$, \mathbf{K} is the $l \times (N-l)$ matrix $(\mathbf{K})_{ij} = a_{ij} X_i \quad i \in \{1, \dots, l\}$ and $j \in \{l+1, \dots, N\}$.

3.1.4.4 Local Stability of the equilibrium points

As presented in section 2.1.5.1, the local stability of a equilibrium point can be studied by means of the linearized system (3.33). The condition for the local stability of a given equilibrium point is that all eigenvalues of the Jacobian matrix evaluated at this equilibrium point should have negative real parts. For the internal equilibrium point (3.29), it is thus necessary to find the eigenvalues of the matrix

$$\mathbf{M} = \mathbf{J} \Big|_{\vec{X}=\vec{X}^*} = \text{diag}(\vec{r} \circ \vec{X}^*) \mathbb{A}, \quad (3.37)$$

in which \mathbf{M} is the community matrix, the coefficients m_{ij} of which indicate the net effect of the interaction between populations i and j upon the growth rate of population i near the equilibrium point. The distribution of eigenvalues is then dependent on \vec{X}^* , as seen in Gibbs et al. (2018).

For solutions with some subset of the populations extinct (3.32), the linearized system gets decoupled for a subset of the populations, and their eigenvalues can then be calculated directly. Indeed, considering the eigenvalues of matrix $\mathbf{M}' = J_{ij}|_{\vec{X}^*,*}$ yields

$$\det(\mathbf{M}' - \lambda \mathbf{I}) = \det \begin{pmatrix} \mathbf{J}' - \lambda \mathbf{I}_l & \mathbb{K} \\ 0 & \text{diag}\left(r_i \left[\sum_j a_{ij} X_j^* + 1\right] - \lambda\right) \end{pmatrix},$$

and since the determinant of a block upper triangular matrix can be calculated as the product of the determinants of its diagonal blocks, it follows that

$$\begin{aligned} \det(\mathbf{M}' - \lambda \mathbf{I}) &= \det(\mathbf{J}' - \lambda \mathbf{I}_l) \det \left[\text{diag}\left(r_i \left(\sum_j a_{ij} X_j^* + 1\right) - \lambda\right) \right] \\ \det(\mathbf{M}' - \lambda \mathbf{I}) &= \det(\mathbf{J}' - \lambda \mathbf{I}_l) \prod_{i=l+1}^N \left(r_i \left(\sum_j a_{ij} X_j^* + 1\right) - \lambda \right). \end{aligned} \quad (3.38)$$

Therefore, a necessary condition for local stability of the equilibrium for populations X_{l+1}, \dots, X_N are extinct is

$$\sum_j a_{ij} X_j^* + 1 < 0 \quad \forall i \in \{l+1, \dots, N\}. \quad (3.39)$$

The set of conditions (3.39) indicates that the extinct populations cannot invade the community. In other words, if any of the populations that have zero density in the equilibrium point is reintroduced into the community at a low density, the per capita growth rate of such population will be negative. Also, the coexistence of all populations with positive densities in this equilibrium point must be stable since all eigenvalues of \mathbf{J}' are also eigenvalues of \mathbf{M}' .

3.1.4.5 Global Stability of Equilibrium Points

In opposition to local stability, the global stability of a dynamical system ensures that all trajectories that do not start in some (unstable) equilibrium point will converge to a given (globally stable) equilibrium point. In other words, the basin of attraction of the globally stable equilibrium point is the entire state space, except for other equilibrium points. The techniques used to prove global stability for a given solution involve the construction of a Lyapunov function $V(\vec{X})$, $V : \mathbb{R}_{\geq 0}^N \rightarrow \mathbb{R}$ which has the properties.

$$\left\{ \begin{array}{l} V(\vec{X}) \geq 0 \quad \forall \vec{X} \in \mathbb{R}_{\geq 0}^N \\ V(\vec{X}) = 0 \iff \vec{X} = \vec{X}^* \\ \frac{dV(\vec{X}')}{dt} = \left(\sum_i \frac{\partial V(\vec{X})}{\partial X_i} \frac{dX_i}{dt} \right) \Big|_{\vec{X}=\vec{X}'} \leq 0 \quad \forall \vec{X}' \in \mathbb{R}_{\geq 0}^N \\ \frac{dV(\vec{X}')}{dt} = 0 \iff \vec{X}' = \vec{X}^*. \end{array} \right. \quad (3.40)$$

In order to construct a Lyapunov function for the generalized Lotka-Volterra model, it will be necessary to define some matrix properties.

- **Negative Definite Matrix:** A matrix \mathbf{B} is said to be negative definite if and only if $\vec{X}^T \mathbf{B} \vec{X} < 0 \quad \forall \vec{X} \neq \vec{0}$. In addition, if \mathbf{B} is also symmetric such that $\mathbf{B} = \mathbf{B}^T$, then all of its eigenvalues are real and negative.
- **Volterra-Lyapunov Stability:** A matrix \mathbf{B} is said to be Volterra-Lyapunov stable, Lyapunov diagonally stable or VL-stable, if there exists a positive diagonal matrix $\mathbf{D} > 0$ such that the symmetric matrix $\mathbf{DB} + \mathbf{B}^T \mathbf{D}$ is negative definite.
- **P-matrix:** A matrix \mathbf{B} is said to be a *P*-matrix if, and only if, all the principal minors of \mathbf{B} are positive.

More explicitly, if \mathbf{B} is VL-stable, then

$$\vec{X}^T (\mathbf{DB} + \mathbf{B}^T \mathbf{D}) \vec{X} < 0, \quad (3.41)$$

and since $(\mathbf{DB} + \mathbf{B}^T \mathbf{D})^T = (\mathbf{DB})^T + (\mathbf{B}^T \mathbf{D})^T = \mathbf{B}^T \mathbf{D} + \mathbf{DB} = \mathbf{DB} + \mathbf{B}^T \mathbf{D}$, i.e. the matrix $\mathbf{DB} + \mathbf{B}^T \mathbf{D}$ is symmetric, then all its eigenvalues are negative.

Murty (1972) has shown that if a matrix \vec{B} is a *P*-matrix, there is an unique solution for the linear complementarity problem

$$\begin{cases} \vec{y} = \mathbf{B} \vec{x} + \vec{c} \geq \vec{0} \\ \vec{x} \geq \vec{0} \\ y_i x_i = 0 \end{cases},$$

in which \mathbf{B} is a $N \times N$ real matrix and $\vec{x}, \vec{y} \in \mathbb{R}_{\geq 0}^N$. However, by allowing $\mathbf{B} = -\mathbf{A}$ and $\vec{c} = \vec{1}$, this linear complementarity problem is equivalent to finding a equilibrium point for the generalized Lotka-Volterra model (Takeuchi 1996), with the additional assumption that

$$1 + \sum_{j=1}^N a_{ij} X_j^* \leq 0 \quad i = 1, \dots, N. \quad (3.42)$$

Thus, if it is possible to show that $-\mathbf{A}$ is a *P*-matrix, then the underlying system is guaranteed to have a unique solution in $\mathbb{R}_{\geq 0}^N$. Indeed, (Cross 1978) proved that if a matrix \mathbf{B} is VL-stable than $-\mathbf{B}$ is a *P*-matrix, and therefore the associated gLV model has a unique solution. This solution does not need to be internal, i.e. there might be some populations X_i for which $X_i^* = 0$, but if that is the case it is always possible to consider the subsystem for which $X_i^* > 0 \forall i$ and thus it is guaranteed that some subsystem has a internal equilibrium if the interaction matrix \mathbf{A} is VL-stable.

However, suppose that system (3.27) has a non-negative equilibrium $\vec{X}^* = (X_1^*, \dots, X_l^*, 0, \dots, 0)^T$. It is then possible to construct a Lyapunov function to evaluate the global stability of this equilibrium. Following (Goh 1977; Takeuchi 1996), it is possible to write

$$V(\vec{X}) = \overbrace{\sum_{i=1}^l d_i \left(X_i - X_i^* - X_i^* \log \left(\frac{X_i}{X_i^*} \right) \right)}^{S1} + \overbrace{\sum_{i=l+1}^N d_i X_i}^{S2}, \quad (3.43)$$

in which $\vec{d} \in \mathbb{R}^N > \vec{0}$ is yet unspecified.

In order to show that the putative Lyapunov function satisfies the conditions (3.40), let us first define the family of functions f_{x^*} , parametrized by $x^* \in \mathbb{R}_{>0}$:

$$\begin{aligned} f_{x^*} : \quad \mathbb{R}_{>0} &\rightarrow \mathbb{R} \\ x &\mapsto x - x^* - x^* \ln(x/x^*) \end{aligned}$$

The derivative of this function can be calculated as

$$\begin{aligned} f'_{x^*}(x) &= \frac{df_{x^*}(x)}{dx} = 1 - \frac{x^*}{x} \\ \implies f'_{x^*}(x^*) &= 0 \\ \implies f''_{x^*}(x) \Big|_{x=x^*} &= \frac{x^*}{x^2} \Big|_{x=x^*} = \frac{1}{x^*} > 0. \end{aligned}$$

From the development above, it is possible to see that $f_{x^*}(x)$ has a unique global minimum at $x = x^*$. Since $S1 = \sum_{i=1}^l d_i f_{X_i^*}(X_i)$, it follows that $S1$ has a global minimum at $X_i = X_i^*$, $i = 1, \dots, l$. Additionally since $X_i \geq 0$, the global minimum for $S2$ is $X_i = 0$, $i = l+1, \dots, N$. Therefore $\min V(\vec{X}) = V(\vec{X}^*) = 0$.

$$\begin{aligned} \frac{d}{dt} V(\vec{X}) &= \sum_{i=1}^N \frac{\partial V(X)}{\partial X_i} \frac{dX_i}{dt} \\ &= \sum_{i=1}^l d_i \left(1 - \frac{X_i^*}{X_i}\right) \frac{dX_i}{dt} + \sum_{i=l+1}^N d_i \frac{dX_i}{dt} \\ &= \sum_{i=1}^l d_i (X_i - X_i^*) \frac{1}{X_i} \frac{dX_i}{dt} + \sum_{i=l+1}^N d_i \frac{dX_i}{dt} \\ &= \sum_{i=1}^l d_i (X_i - X_i^*) r_i \left(1 + \sum_{j=1}^N a_{ij} X_j\right) + \sum_{i=l+1}^N d_i r_i X_i \left(1 + \sum_{j=1}^N a_{ij} X_j\right) \end{aligned}$$

But $\sum_{j=1}^N a_{ij} X_j^* = -1$ for $i = 1, \dots, l$. Conversely, for $i = l+1, \dots, N$ it holds that $X_i^* = 0$, so it the followings relations always hold

$$\begin{aligned} \sum_{j=1}^N a_{ij} X_j &= \sum_{j=1}^N a_{ij} X_j - \sum_{j=1}^N a_{ij} X_j^* + \sum_{j=1}^N a_{ij} X_j^* \\ &= \sum_{j=1}^N a_{ij} (X_j - X_j^*) + \sum_{j=1}^N a_{ij} X_j^* \\ 1 + \sum_{j=1}^N a_{ij} X_j &= \begin{cases} \sum_{j=1}^N a_{ij} (X_j - X_j^*) & i = 1, \dots, l \\ \sum_{j=1}^N a_{ij} (X_j - X_j^*) + \left(1 + \sum_{j=1}^N a_{ij} X_j^*\right) & i = l+1, \dots, N \end{cases} \end{aligned}$$

Therefore,

$$\begin{aligned}
\frac{d}{dt} V(\vec{X}) &= \sum_{i=1}^l (X_i - X_i^*) \sum_{j=1}^N d_i r_i a_{ij} (X_j - X_j^*) + \sum_{i=l+1}^N (X_i - X_i^*) \sum_{j=1}^N d_i r_i a_{ij} (X_j - X_j^*) + \\
&\quad + \sum_{i=l+1}^N d_i r_i X_i \left(1 + \sum_{j=1}^N a_{ij} X_j^* \right) \\
&= \sum_{i=1}^N \sum_{j=1}^N (X_i - X_i^*) d_i r_i a_{ij} (X_j - X_j^*) + \sum_{i=l+1}^N d_i r_i X_i \left(1 + \sum_{j=1}^N a_{ij} X_j^* \right) \\
\therefore \frac{d}{dt} V(\vec{X}) &= (1/2)(\vec{X} - \vec{X}^*)^T (\mathbf{CA} + \mathbf{A}^T \mathbf{C})(\vec{X} - \vec{X}^*) + \sum_{i=l+1}^N d_i r_i X_i \left(1 + \sum_{j=1}^N a_{ij} X_j^* \right) \quad (3.44)
\end{aligned}$$

in which $\mathbf{C} = \text{diag}(d_i r_i)$. Since \mathbf{A} is VL-stable hypothesis, the matrix $(\mathbf{CA} + \mathbf{A}^T \mathbf{C})$ is negative definite and thus the first term of the equation is negative in the domain of $V(\vec{X})$. The second term in the equation (3.44) is always negative due to the requirement for the existence of a unique equilibrium (3.42). Therefore it is possible to conclude that, if it exists, this unique equilibrium is globally stable for a Lotka-Volterra system with VL-stable interaction matrix.

A significant result is that all principal submatrices of a VL-stable matrix are also VL-stable. A possible ecological interpretation of this fact is that if the equilibrium of populations in a community is globally stable, all possible assemblies of the same populations will also display globally stable equilibria (Cross 1978).

3.1.4.6 Unbounded, Cyclic and Chaotic Dynamics

In order to characterize the solutions of system (3.27), it is necessary to define what types of asymptotic dynamics can be expected. As asymptotic dynamics, we refer to the topological properties of the solution of the system after an arbitrary initial temporal evolution, in which the system might display some undefined dynamics. This initial temporal evolution is referred to as *transient*.

It is known that generalized Lotka-Volterra systems might display any dynamics (Smale 1976). In what follows, an attempt is made to identify or present conditions for particular dynamics from the parameters of the model (3.27).

3.1.4.7 Unbounded Dynamics

For the applicability of the gLV model to real problems, it is relevant to investigate if the model may predict unreasonable dynamics for some parameters. Indeed, a very unreasonable scenario when considering N interacting populations is that some subset with M populations grow in density indefinitely. In order to investigate the existence of such scenario, consider the interaction matrix \mathbb{A} and suppose it has an eigenvector $\vec{c} \geq \vec{0}$ such that $\mathbf{A}\vec{c} = \lambda\vec{c}$, $\lambda > 0$. If we suppose the vector of populations densities is linear dependent on the eigenvector \vec{c} such that $\vec{X} = k\vec{c}$, $k \geq 0$, it follows that:

$$\begin{aligned}
\frac{d\vec{X}}{dt} \Bigg|_{\vec{X}=k\vec{c}} &= \text{diag}(k\vec{c} \circ \vec{r}) (\vec{1} + k\mathbf{A}\vec{c}) \\
\therefore \frac{d\vec{X}}{dt} \Bigg|_{\vec{X}=k\vec{c}} &= k \text{diag}(\vec{c} \circ \vec{r}) (\vec{1} + k\lambda\vec{c}) \geq 0, \quad (3.45)
\end{aligned}$$

If the interaction matrix satisfies the previously stated conditions, there is always a line in $\mathbb{R}_{\geq 0}^N$ for which the growth of all populations is always positive (or null). However, it is not guaranteed that the temporal evolution will drive all the populations to infinite growth for a given initial condition on this line, as the flow might not be parallel to such a line at all points.

The results for the boundedness of gLV systems are developed following Hofbauer and Sigmund (1998). A matrix \mathbf{A} is said to be a *B*-matrix if solutions of the associated Lotka-Volterra model (3.1) are bounded:

Definition 3.1. Let \mathbf{A} be an $n \times n$ interaction matrix with real entries $a_{ij} \in \mathbb{R}$. \mathbf{A} is said to be a B -matrix if solutions of the associated Lotka-Volterra model (3.1), $X_i(t)$, are such that $\exists M > 0$ such that $X_i(t) < M$ for $t \rightarrow \infty \forall i$.

Theorem 3.1. [adapted from (Hofbauer and Sigmund 1998)] Suppose that $\vec{X} \geq \vec{0}$ with $\vec{X} \neq \vec{0}$ and $\vec{X} \in \mathbb{R}^n$, if there is an i such that $X_i > 0$ and $(\mathbf{A}\vec{X})_i < 0$ for any vector \vec{X} satisfying the condition above, then the matrix \mathbf{A} is a B -matrix.

Many of the properties equivalent to \mathbf{A} being a B -matrix were not presented, for an extensive list see Hofbauer and Sigmund (1998). It is possible to provide an ecological interpretation for the theorem 3.1 by noting that a given community is guaranteed to have bounded dynamics if, for any combinations of populations densities in a community, at least one of the populations with non-zero density has its growth rate decreased due to the effect of community interactions, i.e., at least one population should have a growth rate lower than it would have if it were isolated and at low densities.² In this condition, intraspecific competition is also included. Note that it is not required that the density of at least one population to be decreasing, since the condition would always be violated by considering a very small \vec{X} .

Indeed, the theorem requires that for all combinations of populations densities, there must be at least one population for which the community interaction effect is negative, provided $\vec{X} \neq \vec{0}$. It is, therefore, possible to investigate conditions that individual and pairs of interacting populations must satisfy so to guarantee bounded dynamics.

For isolated populations, consider the circumstance in which just one of the populations has non-zero density, for instance $X'_i = 0 \forall i \neq j$ and $X'_i = 1 i = j$. The requirement for boundedness then states

$$\begin{aligned} (\mathbf{A}\vec{X}')_i &< 0 \\ \implies a_{jj} &< 0, \end{aligned} \tag{3.46}$$

which is the natural requirement that the intraspecific competition should have a negative influence on population growth.

For two interacting populations, say $X_j = x_j$ and $X_k = x_k$, it follows that

$$\begin{aligned} &\begin{cases} (\mathbf{A}\vec{X}')_j < 0 \\ (\mathbf{A}\vec{X}')_k < 0 \end{cases} \\ \implies &\begin{cases} a_{jj}x_j + a_{jk}x_k < 0 \\ a_{kj}x_j + a_{kk}x_k < 0 \end{cases} \end{aligned} \tag{3.47}$$

Suppose that $a_{jk} > 0$ and $a_{kk}, a_{jj} < 0$. It thus follows that

$$\begin{aligned} \implies &\begin{cases} x_k < -(a_{jj}x_j)/(a_{jk}) \\ x_k > -(a_{kj}x_j)/(a_{kk}) \end{cases} \\ \implies &\frac{a_{jj}x_j}{a_{jk}} < \frac{a_{kj}x_j}{a_{kk}} \end{aligned} \tag{3.48}$$

$$\therefore a_{kk}a_{jj} - a_{jk}a_{kj} > 0 \tag{3.49}$$

Conversely, for $a_{jk} \leq 0$ the first inequality is always satisfied:

$$\implies a_{jj}x_j + a_{jk}x_k < 0 \quad \forall (x_j, x_k) \in \mathbb{R}_{>0}^2.$$

From the symmetry of conditions (3.47), the same considerations drawn for a_{jk} hold for a_{kj} . Observing that inequality (3.49) also holds for $a_{jk} \leq 0$ or $a_{kj} \leq 0$, the condition for a system of two populations growing in accordance to the generalized Lotka-Volterra model (3.1) to be bounded is

²There is no free dinner.

the inequality (3.47). Therefore, each pair j and k must have mutual interaction coefficients that satisfy this inequality for a community with many populations.

In general, considering a bounded N -populations community, the inequality $\sum_j a_{ij} < 0$ is true for at least one of the populations. To see that this is the case, consider the vector of population densities $\vec{X} = \vec{1}$, the result is then a direct consequence of the theorem 3.1.

To verify the condition that an N -dimensional system is bounded, the theorem is not convenient, as it requires testing an inequality for all points in $\mathbb{R}_{\geq 0}^N$. The following theorem presented by Hofbauer and Sigmund allows for algorithmic verification of the boundedness of the system.

Theorem 3.2 ((Hofbauer and Sigmund 1998)). *If all proper principal submatrices of A are B -matrices and $\det(-A) > 0$ then A itself is a B -matrix.*

Note that the theorem is recursive: Starting from a given $N \times N$ matrix \mathbf{A} , we first check if the 1×1 principal submatrices are B -matrices. Based on the development of the boundedness of an isolated population, this is equivalent to

$$a_{ii} < 0 \quad i = 1, \dots, N \quad (3.50)$$

If this is the case, then we proceed to verify the determinants of 2×2 principal submatrices of A . If \mathbf{A}'_2 is a 2×2 principal submatrix of \mathbf{A} and the condition (3.50) is satisfied for \mathbf{A} , then if $\det \mathbf{A}'_2 < 0$, \mathbf{A}'_2 is also a B -matrix. The condition $\det \mathbf{A}'_2 < 0$ is precisely (3.47). The process then continues up to $N - 1$, testing all the determinants of the principal submatrices with order $3, \dots, N - 1$ so that if $\det -\mathbf{A}'_n < 0 \quad n = 3, \dots, N - 1$ for all the submatrices, then \mathbf{A} is a B -matrix and the system is bounded. An additional advantage of an algorithmic approach is to be able to conclude that \mathbf{A} is not a B -matrix at any intermediate step and thus conclude that the gLV system described by \mathbf{A} is not guaranteed to be bounded.

The algorithm 1 indicates a possible way to iteratively verify if a matrix \mathbf{A} is indeed a B -matrix.

Algorithm 1 An algorithm for testing if some matrix is a B -matrix

```

1: function ISBMATRIX( $N, \mathbf{A}$ ) ▷ Inputs are matrix  $\mathbf{A}$  and its order  $N$ 
2:    $n \leftarrow 1$ 
3:   while  $n \leq N$  do
4:     for all  $\mathbf{C}, n \times n$  principal submatrix of  $\mathbf{A}$  do
5:       if  $\det(-\mathbf{C}) < 0$  then
6:         return False ▷  $\mathbf{A}$  is not a  $B$ -matrix
7:       end if
8:     end for
9:     if  $n = N$  then
10:      return True ▷  $\mathbf{A}$  is a  $B$ -matrix
11:    else
12:       $n \leftarrow n + 1$ 
13:    end if
14:   end while
15: end function

```

3.1.4.8 Limit Cycles

Limit cycles are known to happen in systems similar to generalized Lotka-Volterra systems since 1972 (May 1972a). In a limit cycle, the system solution is periodic in the sense that $\vec{X}(t) = \vec{X}(t + T)$, additionally, if such cycle is stable neighbouring trajectories on the state space tend to converge to or diverge from it.

Limit cycles are possible in the generalized Lotka-Volterra model (3.1) for three or more interacting populations (Takeuchi 1996). For instance, consider the particular case of three competing populations, such that population A outcompetes B , B outcompetes C , and C outcompetes A in an ecological analog of the classical rock-paper-scissors game. The analysis presented here is developed based on the results presented by May and Leonard (1975). Let α be the coefficient related to the competitive regulation of the best competitor over the weaker competitor and β be

the regulation of the weaker competitor over the stronger one. This non-transitive system can be parametrized as

$$\frac{d\vec{X}}{dt} = \text{diag}(\vec{X} \circ \vec{r}) (\vec{1} - \mathbf{A}\vec{X}) \quad (3.51)$$

$$\mathbf{A} = \begin{pmatrix} 1 & \alpha & \beta \\ \beta & 1 & \alpha \\ \alpha & \beta & 1 \end{pmatrix},$$

with $\vec{r} = \vec{1}$, $0 < \beta < 1 < \alpha$ and $\alpha + \beta > 2$. Remarkably, any two species from these three competitors satisfy the conditions for a competitive exclusion scenario presented in section 3.1.3.5.

The equilibrium points for the system can be calculated by solving the algebraic system of equations $\frac{d\vec{X}}{dt} = 0$, which yields

$$\begin{aligned} \vec{X}_0^* &= (0, 0, 0)^T \\ \vec{X}_1^* &= (1, 0, 0)^T \\ \vec{X}_2^* &= (0, 1, 0)^T \\ \vec{X}_3^* &= (0, 0, 1)^T. \end{aligned}$$

Additionally, there are three solutions in which two populations coexist that can be found as

$$\begin{cases} x_< + \alpha x_> = 1 \\ \beta x_< + x_> = 1 \end{cases} \implies (x_<, x_>) = \frac{1}{1 - \alpha\beta}(1 - \alpha, 1 - \beta),$$

in which we denote the better competitor by $x_>$ and the worst competitor by $x_<$. However, since $\alpha > 1$, these solutions lie outside of the region $\mathbb{R}_{\geq 0}^3$ and thus are not relevant for further analysis.

The interior solution is $\vec{x}^* = \mathbf{A}^{-1}\vec{1}$ which can be evaluated using the cofactor matrix $\bar{\mathbf{A}}$ as

$$\begin{aligned} \mathbf{A}^{-1} &= \frac{1}{\det(\mathbf{A})} \bar{\mathbf{A}}^T \\ \mathbf{A}^{-1} &= \frac{-1}{1 + \alpha^3 + \beta^3 - 3\alpha\beta} \begin{pmatrix} 1 - \alpha\beta & \alpha^2 - \beta & \beta^2 - \alpha \\ \beta^2 - \alpha & 1 - \alpha\beta & \alpha^2 - \beta \\ \alpha^2 - \beta & \beta^2 - \alpha & 1 - \alpha\beta \end{pmatrix} \\ \implies -\mathbf{A}^{-1}\vec{1} &= \frac{1}{1 + \alpha^3 + \beta^3 - 3\alpha\beta} \begin{pmatrix} 1 + \alpha^2 + \beta^2 - \alpha\beta - \beta - \alpha \\ 1 + \alpha^2 + \beta^2 - \alpha\beta - \beta - \alpha \\ 1 + \alpha^2 + \beta^2 - \alpha\beta - \beta - \alpha \end{pmatrix}. \end{aligned}$$

Thus the flow is stationary if all populations have the densities given by

$$\begin{aligned} X_{4i}^* &= \frac{1 + \alpha^2 + \beta^2 - \alpha\beta - \beta - \alpha}{1 + \alpha^3 + \beta^3 - 3\alpha\beta} \\ &= \left(\frac{1 + \alpha + \beta}{1 + \alpha + \beta} \right) \left(\frac{1 + \alpha^2 + \beta^2 - \alpha\beta - \beta - \alpha}{1 + \alpha^3 + \beta^3 - 3\alpha\beta} \right) \\ &= \frac{1}{1 + \alpha + \beta} \left(\frac{1 + \alpha^3 + \beta^3 - 3\alpha\beta}{1 + \alpha^3 + \beta^3 - 3\alpha\beta} \right) \\ \therefore X_{4i}^* &= \frac{1}{1 + \alpha + \beta} \quad \forall i. \end{aligned}$$

In order to study the stability of these equilibrium points, we proceed with the calculation of the eigenvalues associated with each of them. For X_0^* , all the eigenvalues equal and can be calculated as

$$\lambda_0 = 1 - 1 \times 0 - \alpha \times 0 - \beta \times 0 = 1,$$

and this solution is a source.

By symmetry, the distribution of eigenvalues associated with the solutions $\vec{X}_1^*, \dots, \vec{X}_3^*$ is the same. Calculating the eigenvalues of $\text{vec}X_1^*$, yields

$$\begin{aligned}\lambda_1 &= 1 - 1 \times 1 - \alpha \times 0 - \beta \times 0 = 0 \\ \lambda_2 &= 1 - 1 \times 0 - \alpha \times 0 - \beta \times 1 = 1 - \beta \\ \lambda_3 &= 1 - 1 \times 0 - \alpha \times 1 - \beta \times 0 = 1 - \alpha,\end{aligned}$$

which indicates that all the points with the prevalence of just one population are saddles, since $\beta < 1$.

In order to evaluate the stability of the internal equilibrium point, it is possible to evaluate the Jacobian matrix \mathbf{J} of the system (3.51), by means of (3.34) as

$$\mathbf{J} = \frac{-\mathbf{A}}{1 + \alpha + \beta}. \quad (3.52)$$

The characteristic polynomial associated with the Jacobian matrix is

$$\begin{aligned}P(\lambda) &= \det(\mathbf{J} - \lambda\mathbf{I}) \\ &= \lambda^3 + \frac{3}{1 + \alpha + \beta}\lambda^2 + \frac{3(1 - \alpha\beta)}{(1 + \alpha + \beta)^2}\lambda + \frac{1 + \alpha^3 + \beta^3 - 3\alpha\beta}{(1 + \alpha + \beta)^3}.\end{aligned} \quad (3.53)$$

In order to find all the roots of the characteristic polynomial, we first note that $\lambda_1 = -1$ is a solution:

$$\begin{aligned}P(-1) &= -\frac{(1 + \alpha + \beta)^3}{(1 + \alpha + \beta)^3} + \frac{3(1 + \alpha + \beta)^2}{(1 + \alpha + \beta)^3} - \frac{3(1 - \alpha\beta)(1 + \alpha + \beta)}{(1 + \alpha + \beta)^3} + \frac{1 + \alpha^3 + \beta^3 - 3\alpha\beta}{(1 + \alpha + \beta)^3} \\ &= \frac{1}{(1 + \alpha + \beta)^3} (-\alpha^3 - 3\alpha^2\beta - 3\alpha^2 - 3\alpha\beta^2 - 6\alpha\beta - 3\alpha - \beta^3 - 3\beta^2 - 3\beta - 1 + \\ &\quad + 3\alpha^2 + 6\alpha\beta + 6\alpha + 3\beta^2 + 6\beta + 3 + \\ &\quad + 3\alpha^2\beta + 3\alpha\beta^2 + 3\alpha\beta - 3\alpha - 3\beta - 3 + \\ &\quad \alpha^3 - 3\alpha\beta + \beta^3 + 1) \\ &= 0.\end{aligned}$$

Rewriting the characteristic polynomial as $P(\lambda) = p(\lambda)(\lambda + 1)$ it follows that

$$p(\lambda) = \lambda^2 + \overbrace{\frac{2 - \alpha - \beta}{1 + \alpha + \beta}\lambda}^b + \overbrace{\frac{(\alpha - \beta)^2 + (\alpha - 1)(\beta - 1)}{(1 + \alpha + \beta)^2}}^c,$$

and the other eigenvalues are roots of this expression roots, so that

$$\begin{aligned}
\Delta &= b^2 - 4c \\
&= \frac{1}{(1+\alpha+\beta)^2} (\alpha^2 + 2\alpha\beta + \beta^2 - 4\alpha - 4\beta + 4 - 4(\alpha^2 - \alpha\beta + \beta^2 - \alpha - \beta + 1)) \\
&= -3 \frac{\alpha^2 - 2\alpha\beta + \beta^2}{(1+\alpha+\beta)^2} \\
&= -3 \frac{(\alpha - \beta)^2}{(1+\alpha+\beta)^2} \\
\implies \lambda_2 &= \frac{-b + \sqrt{\Delta}}{2} = \frac{\alpha + \beta - 2 + \sqrt{3}(\alpha - \beta)i}{2(1+\alpha+\beta)} \\
\implies \lambda_3 &= \frac{-b - \sqrt{\Delta}}{2} = \frac{\alpha + \beta - 2 - \sqrt{3}(\alpha - \beta)i}{2(1+\alpha+\beta)}.
\end{aligned}$$

The real part of these eigenvalues is

$$\operatorname{Re}(\lambda_2) = \operatorname{Re}(\lambda_3) = \frac{\alpha + \beta - 2}{2(1+\alpha+\beta)} > 0, \quad (3.54)$$

since $\alpha + \beta > 2$ by hypothesis. The interior solution is also a saddle since it has eigenvalues with both positive and negative real parts. Therefore, this system does not present stable equilibrium points.

The system could display an unbounded dynamics, and in order to investigate whether this is the case, it is possible to investigate the total density $S = X_1 + X_2 + X_3$, so that

$$\begin{aligned}
\frac{dS}{dt} &= \frac{dX_1}{dt} + \frac{dX_2}{dt} + \frac{dX_3}{dt} \\
\frac{dS}{dt} &= X_1 + X_2 + X_3 - (X_1^2 + X_2^2 + X_3^2 + (\alpha + \beta)(X_1X_2 + X_1X_3 + X_2X_3)).
\end{aligned} \quad (3.55)$$

In order to compare $\frac{dS}{dt}$ to the logistic growth, let us define a function $L(S) = S(1 - S)$, which can be expanded in terms of X_1, X_2 and X_3 as

$$\begin{aligned}
L &= S(1 - S) \\
L &= (X_1 + X_2 + X_3)[1 - (X_1 + X_2 + X_3)] \\
L &= X_1 + X_2 + X_3 - [X_1^2 + X_2^2 + X_3^2 + 2(X_1X_2 + X_1X_3 + X_2X_3)].
\end{aligned}$$

Therefore,

$$\begin{aligned}
\frac{dS}{dt} - L &= [2 - (\alpha + \beta)](X_1X_2 + X_1X_3 + X_2X_3) \leq 0 \\
\implies \frac{dS}{dt} &\leq S(1 - S).
\end{aligned}$$

The total density will always increase at a lower rate than a corresponding logistic model. Since we have shown in section 3.1.2 that the logistic model is bounded, the model (3.51) is also bounded under the condition that $\alpha + \beta > 2$. In order to investigate what is the dynamics of this system, it is useful to define the function $P(\vec{X}) = X_1X_2X_3$, for which

$$\begin{aligned}
\frac{dP}{dt} &= \frac{dX_1}{dt}X_2X_3 + X_1\frac{dX_2}{dt}X_3 + X_1X_2\frac{dX_3}{dt} \\
&= X_1X_2X_3(3 - X_1 - \alpha X_2 - \beta X_3 - \beta X_1 - X_2 - \alpha X_3 - \beta X_1 - \alpha X_2 - X_3) \\
&= P(3 - (1 + \alpha + \beta)S).
\end{aligned} \quad (3.56)$$

It is then possible to define the Lyapunov function $V(\vec{X}) = (P(\vec{X}))/((S(\vec{X}))^3 + \epsilon)$ in which ϵ is an arbitrarily small number. This Lyapunov function is positive semidefinite, since $\vec{X} \in \mathbb{R}_{\geq 0}^3$ with the function vanishing on the boundary of $\mathbb{R}_{\geq 0}^3 : V(X_1, 0, 0) = V(0, X_2, 0) = V(0, 0, X_3) = 0$. Observe that

$$\begin{aligned}\frac{dV}{dt} &= \frac{d}{dt} \left(\frac{P}{S^3 + \epsilon} \right) \\ &= \frac{\frac{dP}{dt}(S^3 + \epsilon) - P \frac{dS^3}{dt}}{S^6 + S^3\epsilon + \epsilon^2} \\ \implies \lim_{\epsilon \rightarrow 0} \frac{dV}{dt} &= \frac{\frac{dP}{dt}S^3 - 3PS^2 \frac{dS}{dt}}{S^6},\end{aligned}$$

and thus by making ϵ sufficiently small, the derivative of the Lyapunov function will be completely determined by $\frac{d}{dt} \left(\frac{P}{S^3} \right)$. Therefore

$$\begin{aligned}\frac{d}{dt} \left(\frac{P}{S^3} \right) &= \frac{\frac{dP}{dt}S^3 - 3PS^2 \frac{dS}{dt}}{S^6} \\ &= \frac{P}{S^4} [S(3 - (1 + \alpha + \beta)S) - 3S + 3[X_1^2 + X_2^2 + X_3^2 + (\alpha + \beta)(X_1X_2 + X_1X_3 + X_2X_3)]] \\ &= \frac{P}{S^4} [- (1 + \alpha + \beta)(X_1^2 + X_2^2 + X_3^2 + 2X_1X_2 + 2X_1X_3 + 2X_2X_3) + \\ &\quad + 3[X_1^2 + X_2^2 + X_3^2 + (\alpha + \beta)(X_1X_2 + X_1X_3 + X_2X_3)]] \\ &= \frac{P}{S^4} [(X_1 - X_2)^2 + (X_1 - X_3)^2 + (X_2 - X_3)^2 + (\alpha + \beta)(X_1X_2 + X_1X_3 + X_2X_3 - X_1^2 - X_2^2 - X_3^2)] \\ &= \frac{P}{S^4} [(X_1 - X_2)^2 + (X_1 - X_3)^2 + (X_2 - X_3)^2 - \left(\frac{\alpha + \beta}{2} \right) [(X_1 - X_2)^2 + (X_1 - X_3)^2 + (X_2 - X_3)^2]] \\ &= \frac{P}{S^4} \left(1 - \frac{\alpha + \beta}{2} \right) [(X_1 - X_2)^2 + (X_1 - X_3)^2 + (X_2 - X_3)^2],\end{aligned}$$

and thus $V(\vec{X})$ is negative semi-definite as $\frac{d}{dt} \left(\frac{P}{S^3} \right) \leq 0$, with $\frac{d}{dt} \left(\frac{P}{S^3} \right) = 0$ being true on the boundaries of $\mathbb{R}_{\geq 0}^3$ and on the diagonal $X_1 = X_2 = X_3$. If $X_1 = X_2 = X_3$, the system can be written as $\frac{dX_i}{dt} = X_i(1 - (1 + \alpha + \beta)X_i) \forall i$ which is a set of three uncoupled logistic equations with carrying capacity $1/(1 + \alpha + \beta)$. This diagonal line is thus the α -limit set of the interior equilibrium point.

It is also possible to conclude that the boundary of $\mathbb{R}_{\geq 0}^3$ is a ω -limit set for the region $\mathbb{R}_{\geq 0}^3 \setminus \{(x, x, x) : x \in \mathbb{R}\}$. Therefore to understand the dynamics of the system, it is necessary to study the behavior of the flow at the boundary. Since the system is symmetric by a cyclic replacement of variables ($X_1 \rightarrow X_2, X_2 \rightarrow X_3, X_3 \rightarrow X_1$ for instance), it is necessary to study only one of the boundaries.

Taking $X_3 = 0$, the model is then a two dimensional generalized Lotka-Volterra system, the behavior of which has been thoroughly studied in section 3.1.3. For this particular system, since $\alpha > 1$, the population 1 outcompetes population 2, which go extinct for any initial condition in the (restricted) state-space with non-zero densities for both populations. Therefore, the entire region $\{X_1, X_2, 0\} \setminus \{0, X_2, 0\} \subset \mathbb{R}_{\geq 0}^3$ is the ω -limit of the equilibrium point \vec{X}_1^* and there is one orbit \mathbf{O}_1 going from \vec{X}_2^* to \vec{X}_1^* .

By the symmetry of the system, there are two other orbits, \mathbf{O}_2 and \mathbf{O}_3 which join the fixed points \vec{X}_3^* to \vec{X}_2^* and \vec{X}_1^* to \vec{X}_3^* , respectively. Thus it is possible to conclude that the ω -set of the whole system is the union of the three orbits $\mathbf{O}_1, \mathbf{O}_2, \mathbf{O}_3$ and the saddles $\vec{X}_1^*, \vec{X}_2^*, \vec{X}_3^*$. This set is called an *heteroclinic cycle*, which joins different equilibrium points with closed orbits so that the whole system is periodic.

Even though this is a very particular example, it presents enough evidence that limit cycles are possible in the generalized Lotka-Volterra model with more than two dimensions.

3.1.4.9 Chaos

Since all dynamics are possible for a generalized Lotka-Volterra model (3.27) with a sufficient number of populations (Smale 1976), it is clear that chaotic dynamics can be found for these systems. Indeed most of the evidence for the existence of chaos in this model is based on examples (Hofbauer and Sigmund 1998). Nevertheless, analytical investigations also corroborate the existence of chaotic dynamics (Kozlov and Vakulenko 2013).

Following (Arneodo, Coullet, and Tresser 1980), a slightly different parametrization of the gLV model with three interacting populations displays chaotic behavior. Consider the model

$$\frac{dN_i}{dt} = N_i \left(\sum_{j=1}^3 \alpha_{ij} (1 - N_j) \right) \quad (3.57)$$

with the parametrization

$$(\alpha_{ij}) = \begin{pmatrix} 0.5 & 0.4 & 0.1 \\ -0.5 & -0.1 & 0.1 \\ 1.43 & 0.1 & 0.1 \end{pmatrix}.$$

For any initial condition, this system evolves into a strange attractor. This highlights the possibility of chaotic behavior in Lotka-Volterra models with $N \geq 3$.

Figure 3.4 shows the three-dimensional state space of the four possible dynamics. It is possible to observe the strange attractor of the chaotic system presented with its typical turns and twists.

3.1.4.10 The Densities at The Equilibrium Point are the Time Average

Even when the equilibrium of some particular gLV system is not stable, it still carries valuable information. In particular, if some system has an equilibrium point \vec{X}^* , $\vec{X}^* \in \mathbb{R}_{\geq 0}^N$, it is possible to show that the time average of the populations densities are equal to \vec{X}^* . The time average is defined as

$$\langle \vec{X} \rangle = \frac{1}{T} \int_{\tau}^{\tau+T} \vec{X}(t) dt, \quad (3.58)$$

in which $\vec{X}(t)$ is the vector of populations' densities at time t . By performing a change of variables in the gLV model (3.1), such that $\vec{y} = \log \vec{X}$, it follows that

$$\begin{aligned} \frac{dX_i}{dt} &= \frac{dX_i}{d \log X_i} \frac{d \log X_i}{dt} \\ &= X_i \frac{d \log X_i}{dt} \\ \implies X_i \frac{d \log X_i}{dt} &= X_i r_i \left(1 + \sum a_{ij} X_j \right) \\ \frac{d \log X_i}{dt} &= r_i \left(1 + \sum a_{ij} X_j \right). \end{aligned}$$

By converting the expression above into vectorial notation and computing the time average on both sides of the equation using the fundamental theorem of calculus, it follows that

$$\begin{aligned} \frac{1}{T} \int_{\tau}^{\tau+T} \frac{d \log \vec{X}}{dt} dt &= \frac{\text{diag}(\vec{r})}{T} \int_{\tau}^{\tau+T} (1 + \mathbf{A} \vec{X}) dt \\ \frac{\text{diag}(\vec{r})^{-1}}{T} \log \vec{X} \Big|_{\tau}^{\tau+T} &= \frac{1}{T} \int_{\tau}^{\tau+T} (1 + \mathbf{A} \vec{X}) dt. \end{aligned} \quad (3.59)$$

The system could display three distinct dynamical behaviors. If the system is evolving towards a stable equilibrium point, it follows that $\vec{X}(\tau) \approx \vec{X}(\tau+T)$ for sufficient large τ so that the system

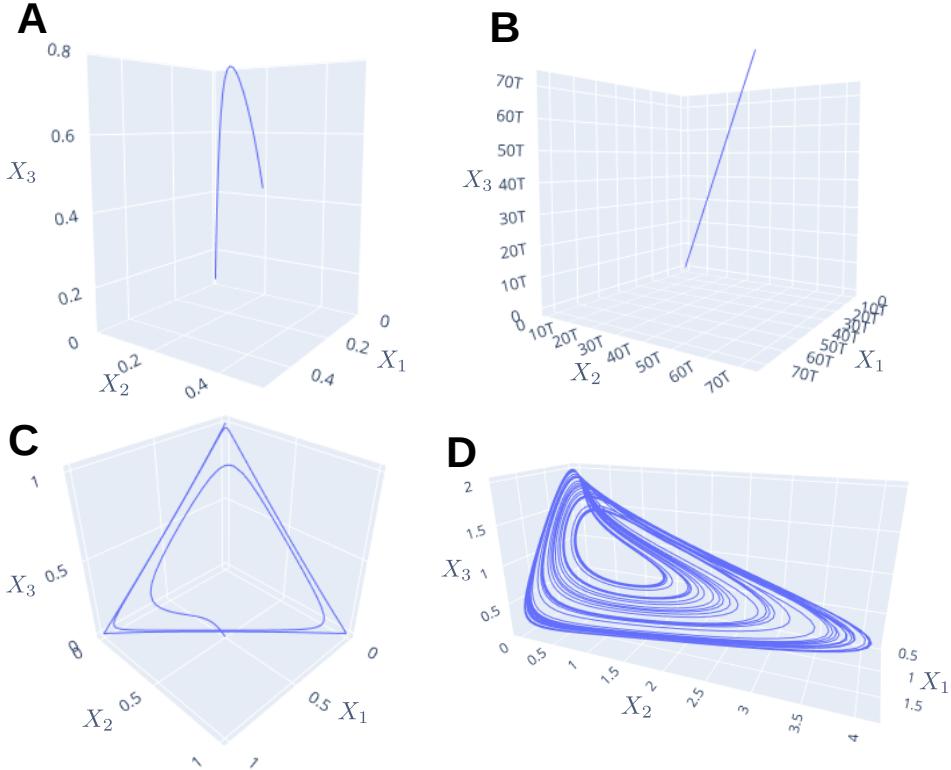


Figure 3.4: **Examples of Dynamics in gLV-3D:** Examples of the state space curves displaying the evolution of three dimensional generalized Lotka-Volterra systems with stable equilibrium points, unbounded dynamics, limit cycles, and chaotic dynamics are shown. **A:** The system (3.51) with $\alpha = \beta = 0.5$ presents a stable equilibrium point in which $X_1^* = X_2^* = X_3^* = 0.5$, the stability of this solution can be verified through the equation (3.54) (May and Leonard 1975). **B:** The system (3.51) with $\alpha = -0.8$, $\beta = -1.3$ presents unlimited growth. Note that the tick labels indicate values in 'tera' = $T = 1 \times 10^{12}$, highlighting the extreme "densities" that the variables can achieve in limited time. **C:** The system (3.51) with $\alpha = 0.8$, $\beta = 1.3$ presents an heteroclinic orbit which is a limit cycle of the system (May and Leonard 1975). **D:** The model (3.57) with parametrization (3.58) shows a distinct strange attractor, which is a fingerprint of chaos (Arneodo, Coullet, and Tresser 1980).

is already on the vicinity of the equilibrium point and any T . If on the vicinity of a limit cycle, there is a well defined period T such that $\vec{X}(\tau) \approx \vec{X}(\tau+T)$ for sufficient large τ . Lastly, in a chaotic dynamics the system display dense periodic orbits which are limited in some region $S \subset \mathbb{R}_{\geq 0}^N$, in this circumstances, the largest difference of to elements in this set is finite, say $\vec{X}(\tau) - \vec{X}(\tau + T) \leq \vec{M}$ for any T .

For the equilibria and periodic solutions, the equation (3.59) reads

$$\begin{aligned}
 & \frac{1}{T} \int_{\tau}^{\tau+T} \mathbf{A} \vec{X} dt + \frac{T}{T} \vec{1} = \frac{\text{diag}(\vec{r})^{-1}}{T} \times 0 \\
 & \mathbf{A} \frac{1}{T} \int_{\tau}^{\tau+T} \vec{X} dt = -\vec{1} \\
 \implies & \left\langle \vec{X} \right\rangle = \frac{1}{T} \int_{\tau}^{\tau+T} \vec{X} dt = -\mathbf{A}^{-1} \vec{1} = \vec{X}^* \quad .
 \end{aligned} \tag{3.60}$$

For the chaotic solution it is possible to calculate

$$\begin{aligned}
& \frac{1}{T} \int_{\tau}^{\tau+T} \mathbf{A} \vec{X} dt + \frac{T}{T} \vec{1} = \text{diag}(\vec{r})^{-1} \lim_{T \rightarrow \infty} \frac{\vec{X}(\tau) - \vec{X}(\tau+T)}{T} = 0 \\
& \mathbf{A} \frac{1}{T} \int_{\tau}^{\tau+T} \vec{X} dt = -\vec{1} \\
\implies & \left\langle \vec{X} \right\rangle = \frac{1}{T} \int_{\tau}^{\tau+T} \vec{X} dt = -\mathbf{A}^{-1} \vec{1} = \vec{X}^* \quad \square
\end{aligned} \tag{3.61}$$

Thus, we conclude that the time average for sufficiently large T is equal to the equilibrium point for all bounded attractors.

3.2 Interplay between Structure and Dynamics in Ecological Communities

This section aims to present the state of the art of research focused on integrating ecological structure and dynamics of networks. Firstly we briefly address why and how large ecological communities are modeled by dynamical systems (section 3.2.1) and complex networks (section 3.2.2), and then review the literature focusing on the topological classifications of ecological communities (section 3.2.3) and the many facets of the interplay between the stability and structure of these ecological communities (3.2.4).

3.2.1 Ecological Communities as Dynamical Systems

The innovative idea of studying ecological systems through mathematical models put forward by Lotka (1920) and Volterra (1926) opened a tradition of applying techniques derived from dynamical systems theory into ecological questions. Although several systems with few interacting species are studied through modeling, speciose ecological communities can also benefit from a dynamical system description (May 1972b).

In ecological models, the abundance, density, or biomass of the populations often play the role of state variables, and the interactions between populations are the main drivers of changes in these variables.

Indeed, Lotka indicated that the per-capita growth rate of a given population should be a function of the densities of all populations in the community (Lotka 1920). This unspecified function was then considered a multilinear function representing the effect of each species' density on the growth rate, also known as the law of masses. Consequently, a community model in which populations' per capita growth rate is a multilinear function considering populations' densities as variables is a generalized Lotka-Volterra model (gLV) (Hofbauer and Sigmund 1998). If $X_i, i \in 1, 2, \dots, N$ represent the density of each of the N populations in the community, in the gLV framework, its per capita growth rate is given by the equation (3.1). Although this model does not incorporate any functional response besides the law of masses, its impact in theoretical ecology is vast.

The advantages of such a "simple" model reside both in its mathematical tractability and the relative ease of interpreting its features and results in the light of ecological theory. Overall, since the linear function is a truncated Taylor expansion of an arbitrary derivable function, all possible ecological models resemble the gLV if analyzed in the vicinity of their respective equilibria (as seen in Rohr, Saavedra, and Bascompte 2014, for example). Indeed, this model can display any asymptotic dynamics for the densities of the populations and thus could, in principle, represent any real ecological system with the appropriate parametrization, as discussed in the section 3.1. For instance, May 1972a demonstrated the importance of limit-cycles in predator-prey systems, while Gonze et al. 2017 discussed that multistability could mechanistically explain the alternative states in ecological communities using the gLV. Nonetheless, the effect of the density of other populations on the growth rate of a given population can be represented as other functional responses when particular features of this interaction, such as predator satiation, are considered (Holling 1959). Certainly, non-linear functional responses can be an important consideration in ecological models, impacting the stability of ecological communities (Qian and Akçay 2020). Altogether, the generalized Lotka-Volterra can be considered a standard model for multispecies communities.

Unsurprisingly, the gLV model has been extensively used to investigate ecological communities in many circumstances: Inference of interactions among taxa from data (Xiao et al. 2017; Maynard, Miller, and Allesina 2020); Incorporation of microbial interactions for studies on the microbiome (Succurro and Ebenhöh 2018; Faust and Raes 2012; Gonze et al. 2018); Evaluation of the accuracy of computational techniques in microbiology and theoretical exploration of the structural role of keystone taxa (Berry and Widder 2014); Investigation on the stability and coexistence of generic communities (Gibbs et al. 2018; Barabás, J. Michalska-Smith, and Allesina 2016).

3.2.2 Ecological Communities as Complex Networks

In a community, each population might interact with all remaining populations leading to several possible interactions among pairs that grow with the square of community richness. For example, in a community composed of 300 populations, there could be up to almost one hundred thousand possible pairwise interactions. However, in speciose communities, the interactions between populations have non-trivial distributions, and many population pairs do not interact in an effective manner (Pinheiro et al. 2019; Paine 1992; Dunne 2006). As a result, it is reasonable that the pattern of ecological interactions among populations in a community can be studied employing a complex network representation (Guimarães 2020).

Note that community-level interactions are a consequence of interactions between organisms, and thus the interactions between populations are an abstraction of interactions between individuals (Guimarães 2020). Importantly, ecological interactions can have positive or negative influences of varying intensities on both populations (Paine 1992). Therefore, the effects of all interactions among individuals from two populations are summarized into a population-level coefficient that encapsulates information on frequency, type, and intensity of pairwise interactions. The intensity and direction of the interaction's influence are then denoted by the absolute value and signal of the interaction coefficient, respectively. Suppose this coefficient is defined as the relative effect of population j on population i growth rate compared to the law of masses. In that case, this is identical to the a_{ij} coefficient of the gLV model (Staniczenko, Kopp, and Allesina 2013; Dormann, Fründ, and Schaefer 2017).

In order to represent the structure of interactions at this level of biological organization, ecological networks are constructed considering populations and interactions as nodes and edges/links, respectively. This network of ecological interactions can be *undirected* so that two nodes A and B are symmetrically connected if there is any ecological relation involving both populations. Alternatively, the network can be *directed*, so that there is a link from A to B if, and only if, A has relevant influence on population B (Dunne 2006).

Therefore, the non-trivial pattern of mutual pairwise interactions among populations can be fully represented by directed and weighted edges in a generic ecological network encompassing all possible interaction types. Besides, there is a direct connection between the ecological community representation as a complex network and as a dynamical system: The off-diagonal elements of the interaction matrix $\mathbf{A} = (a_{ij})$ which describe interspecific interaction coefficients between state variables in the gLV represent the weights of links between populations in the directed complex network (Succurro and Ebenhöh 2018). Indeed, the interaction matrix of the dynamical model is the adjacency matrix of the directed complex network if the elements on the main diagonal are not considered. For an ecological community, these distinct mathematical objects are equivalent: the matrix of interactions determining the community's evolution in time is the adjacency matrix of the weighted, directed ecological network that conveys the community structure.

3.2.3 The Topology of Ecological Community Networks

One of the main advantages of network representation is the possibility to highlight patterns of interactions' distribution in a network, i.e., its interaction structure, architecture, or topology (Guimarães 2020). In this sense, much effort has been made to investigate the topological properties of ecological networks. Due to the challenge of constructing well-resolved networks encompassing an entire community and all possible interactions therein, much of the scientific effort focused on particular types of interactions (food-webs, competitive and mutualistic networks) (Dormann, Fründ, and Schaefer 2017; Dunne 2006). Despite intrinsic difficulties, it is a long-standing consensus that ecological communities have non-random topology (Dunne 2006; Guimarães 2020; Dormann, Fründ, and Schaefer 2017; Takemoto and Iida 2019).

Food-web topology is known to present hierarchical pattern (Mougi and Kondoh 2014) and nested structure (Kondoh, Kato, and Sakato 2010). In addition, some food-webs also present small world properties (Watts and Strogatz 1998; Dunne 2006). Small world property is characterized by the short path lengths between nodes (Watts and Strogatz 1998). It can be understood ecologically as indicating that few intermediary populations separate two populations selected at random in the network representation of the community. Some food-webs also display scale-free distribution of interactions (Solé and Montoya 2001; Dunne, Williams, and Martinez 2002a) while most have exponential degree distribution compatible with random models (Dunne, Williams, and Martinez 2002a). In the scale-free structure, the distribution of interactions between populations follows a power-law, with few populations concentrating most of the network's ecological interactions. On the other hand, the structural organization analysis of 52 real species-rich plant-animal mutualistic networks found that they are highly nested, indicating that in these communities, specialist species interacted with subsets of the interaction partners of generalist species (Bascompte et al. 2003; Almeida-Neto et al. 2008).

While most theories and techniques used to study ecological networks were developed in the context of macro-organisms, microbial communities offer a unique opportunity to extend this area. Indeed, studies on the microbial interaction network can encompass all possible types of interactions and incorporate the entire community. In particular, high throughput sequencing techniques allowed profiling the relative densities of different taxa in microbial communities through time, but retrieving interaction information from this data is a scientific challenge (Berry and Widder 2014). Traditional techniques rely on co-occurrence networks' constructed through correlations among taxa densities from multiple samples of taxonomically annotated metagenomes (Faust and Raes 2012; Li et al. 2016). However, the recovery of interactions from correlations was shown to yield a large proportion of false positives and/or negatives on the resulting network depending on the properties of underlying interactions and methodological procedures employed (Berry and Widder 2014). These caveats hinder the interpretability of any topological features based on such results. On the bright side, new dynamical-systems derived techniques could represent a significant advancement in this field (Xiao et al. 2017; Shang et al. 2017; Maynard, Miller, and Allesina 2020). Altogether, current knowledge indicates that microbial ecological networks have non-random scale-free and clustered topologies (Faust and Raes 2012) and that competition is the most common observed interaction among culturable microbial species (Foster and Bell 2012).

In order to characterize the topological features of ecological networks, one of the most useful measures is the proportion of existent interactions among populations in relation to the maximum number of interactions that could be expected from the number of nodes in the network, i.e. $\text{Con} = E/N(N - 1)$ or $\text{Con}' = E/N^2$. Literature on network theory associates this measure to the *density of the network* (see 2.2.2.3), but in ecology it is often referred as *connectance* or *interactive connectance* (Dunne 2006). Even if definitions of connectance and density differ slightly, this difference tends to zero when considering the limit of large number of populations (nodes) in the community since $\lim_{N \rightarrow \infty} \text{Con}' = \lim_{N \rightarrow \infty} \text{Con} = D$. Notably, ecological networks are characterized by small connectance values, in the range of 3% to 30% but generally near 10% (Dunne 2006). Albeit the connectance indicates how densely connected a given network is, the distribution of connections among nodes might be very uneven (Guimarães 2020).

The network connectance also affects the interplay between other topological measures. Fortuna et al. (2010) found a positive correlation between modularity and nestedness at low connectance, whereas these measures are negatively correlated at high connectance. Indeed, for a fixed number of species, an increase in the connectance is associated with increased nestedness values in mutualistic networks (Bascompte et al. 2003). Additionally, correlations between nestedness and dissimilarity have also been reported in the literature (Johnson and Omland 2004).

Since the interconnections between populations in an ecological network reflect ecological interactions between individuals from these populations, it is natural to ask whether different interaction types would originate topologically distinct networks. Indeed, this question was posed as a challenge to the scientific community by Michalska-Smith and Allesina (2019). The challenge is that despite the expectation that different types of interactions would originate distinct topologies, it is difficult to identify correctly which type(s) of interaction(s) originated a given network having only its topology as information. An early answer to this question points out that it is possible to effectively distinguish between antagonistic and mutualistic ecological networks by incorporating environmental information (Song and Saavedra 2020).

3.2.4 Structure and Stability of Ecological Networks

The fact that ecological communities have a non-random structure and are ecologically stable to some extent, promoted the question of whether there is a relation between topological properties and stability (Dunne 2006). The expectation is that the topological features of the communities could ultimately be related to evolutionary or aggregation processes.

Indeed, the factors that determine the stability of ecological communities is a central question in ecology. A seminal work from May (1972b) started a venue of investigation which was extended for almost half-century through the works of many scientists. Nevertheless, even today, this question is still open to new contributions.

In order to better address the interplay between structure and stability, it is relevant to highlight some aspects of these two essential concepts. First, community structure can be related to the pattern of presence/absence of interactions, the types of interaction, or, ultimately, the intensity of such interactions (Guimaraes 2020). Even though these aspects of the community structure are interdependent, this division aims to identify distinct features contributing to ecological communities' overall complexity.

On the other hand, ecological stability can have many definitions, as discussed in section 2.3.4. Although all definitions of stability are related to some extent, each contribution to this long-standing debate focused on one or a few of them. Accordingly, differences in findings concerning how structure and stability are related can in part be ascribed to the different aspects of structure and stability concepts studied.

An overview of the many studies comparing structural features of communities and their stability is presented in the following. While this review is not intended to be complete, an effort is made to reflect current knowledge.

Richness and Stability Perhaps one of the most striking aspects of natural environments is its richness. In accordance, how can so many populations coexist in a (rich) community is a core question in community ecology. The mainstream answer to this seminal question before 1970 advocated that richer and more connected communities were more stable (Dunne 2006). May (1972b) challenged this view by arguing that rich communities tend to be less stable, considering local asymptotic stability or resilience. The criterion³ derived by May (1972b) for the stability of communities with random interactions was

$$\mu\sqrt{SD} < 1, \quad (3.62)$$

with mean intensity of interaction, richness and the density of interactions denoted by μ , S and D , respectively.

This criterion indicates that rich communities would inevitably be unstable, thus contradicting the observed rich and complex communities. Nonetheless, recent theoretical and computational results indicate that May's criterion does not preclude rich communities. In these frameworks, speciose communities would arise from even richer putative communities (Serván et al. 2018) or iterative assembly processes (Maynard, Serván, and Allesina 2018; Qian and Akçay 2020). Specifically, Serván et al. (2018) showed that from an initial putative community composed of S populations assembled randomly, approximately $S/2$ populations would be able to thrive if the non-invasive solution is assumed to be globally stable. Interestingly, the authors point out that this result would also be found if there were no interspecific interactions in the community.

Connectance and Stability Several works investigated which properties of rich communities allow them to be stable, with some early studies in empirical food webs indicating an inverse scaling between richness and density, i.e., $S \propto D^{-1}$ (Dunne 2006). Nevertheless, this correlation was a consequence of methodological inconsistencies in the network data available at that time (Dunne 2006)

In fact, higher connectance appears to increase the robustness of food-webs (Dunne, Williams, and Martinez 2002b), while having the opposite effect in coexistence and resilience (Thébaud and Fontaine 2010). It is argued that increasing connectance increases the tolerance to the removal of most highly connected nodes due to an increase in the number of paths in the network, therefore

³May's work was based on random matrix techniques, initially developed by Wigner and others as a tool for solving the problem of excited states in the atomic nucleus (May 1972b; Wigner 1967). An excellent example of fruitful interdisciplinarity between physics and ecology, upon which this work is constructed.

diminishing the probability that any single predator would be left without prey after such disruption event (Dunne, Williams, and Martinez 2002b).

For competitive communities in which some individual populations had oscillatory dynamics when isolated, the increase in community size and connectance was also found to positively affect the stability (Fowler 2009). Based on a discrete generalized Lotka-Volterra map, Fowler (2009) argued that by increasing the number of links in these communities, the per capita growth rate of species reduces due to an increase in competitive feed-backs. Ultimately, the increase in density stabilized and reduced the probability of out-of-equilibrium dynamics for the whole community. However, the stabilizing effect of size and connectance was limited for large densely connected competitive communities with oscillatory dynamics and all communities without oscillatory dynamics, possibly due to large competition leading to competitive exclusion. On the other hand, connectance was shown to affect the coexistence and resilience of mutualistic communities (Thébault and Fontaine 2010) positively.

Effects of Network Topology in Stability One of the main criticism of May's findings on the instability of large communities is his assumption of random networks (May 1972b) since the topology of ecological networks is not random (see section 3.2.3). As a result, several studies were conducted to elucidate topology's influence upon stability, either using a network model with distinctive topological features or directly assessing the effect of a particular topological measure upon stability (see section 2.2 for definitions on networks measures and models).

In particular, it was shown that the effect of removal of highly connected populations cascade across the food-webs, leading to topologically determined secondary extinctions (Solé and Montoya 2001; Dunne, Williams, and Martinez 2002b). The result that food-webs are more vulnerable to selective removal of highly connected nodes is similar to the ones found for scale-free networks and distinct from random topology features. Watts and Strogatz (1998) also hypothesized that the small-world feature of some ecological networks would lead to susceptibility to perturbation, as more populations would be in the neighborhood of any perturbed population due to small average path lengths. Indeed, modeling also indicated that network structure plays a crucial role in maintaining community diversity, with food-web like networks being more stable than its random counterparts (Mougi and Kondoh 2014). For instance, it was found that topologies with increased modularity can have modest positive effects on ecological networks' stability, while antimodularity has powerful destabilizing effects (Thébault and Fontaine 2010; Grilli, Rogers, and Allesina 2016).

Nestedness' Effects on Stability Among all possible topological measures in ecological networks, nestedness perhaps received the most attention. It was pointed out that nested community structures are associated to increased (e.g. Thébault and Fontaine 2010) and decreased (e.g Allesina and Tang 2012) mutualistic community stability. Burgos et al. (2007) found that the nested pattern of contacts is the best possible one as far as robustness is concerned, but only when the least linked species have a greater probability of becoming extinct. Additionally, Rohr, Saavedra, and Bascompte (2014) studied structural stability and its relation to robustness in mutualistic systems and determined that the nested network topology maximizes structural stability. Indeed, the nested configuration was also found to benefit coexistence in mutualistic communities (Bastolla et al. 2009). However, Barabás, J. Michalska-Smith, and Allesina (2016) also showed that nestedness structure appears in most (and least) resilient, competitive networks.

Despite significant advances in our understanding of the effects of nested network structure on different aspects of stability, it is essential to note that the principal measure of nestedness is based on unweighted bipartite networks (Almeida-Neto et al. 2008). A robust generalization of this concept to weighted networks demonstrated that many unweighted nested networks had non-nested weighted counterparts when quantitative data concerning interactions were incorporated (Staniczenko, Kopp, and Allesina 2013). Furthermore, the nested distribution of interactions is also guaranteed to be associated with low community resilience (Staniczenko, Kopp, and Allesina 2013).

Note that the general underlying assumption linking the structure of networks and stability is that evolution shapes the structure of ecological interactions (Segar et al. 2020). Therefore, if some pattern is found in ecological communities, this pattern would probably have been selected to improve the stability of the community. There is, however, the possibility that these patterns are byproducts of other processes and not selected *per se*. For instance, Maynard, Serván, and Allesina (2018) showed that different assembly processes could lead to different stable community

structures, further suggesting that these structures could be "completely unrelated to stability or selection".

Types of Interaction Another critical assumption May (1972b) implicitly used when assessing the stability of communities with random interactions was the proportions of each type of ecological interaction in the network. The assumption of random interactions results in 50% of exploitative interactions and 25% of mutualistic and competitive interactions when all possible pairs of populations interact, i.e., in a complete network.

Despite this early focus on communities with multiple types of ecological interactions, later developments focused on networks constructed for a single type of interaction (Dunne 2006). Indeed, food webs are mainly composed of predator-prey interactions, while mutualistic interactions are fingerprints of plant-pollinator networks. Despite this appealing simplicity, natural communities are often composed of a diversity of interactions, which are thought to have specific impacts on community stability and richness.

The general aim of studies focusing on the effect of interaction types upon networks' stability is to indicate which interactions improve ecological communities' stability. Indeed, mutualism was found to have a destabilizing effect on communities (Coyte, Schluter, and Foster 2015; Allesina and Tang 2012). Specifically, mutualistic interactions might generate unbounded positive feedbacks if the per capita effect of the mutualistic taxa on another is independent of their population density (Qian and Akçay 2020), see section 3.1.4.5 for analytical results on this. Nevertheless, if per capita mutualistic benefits are density-dependent and limited, mutualism can promote community richness and resistance to invasions (external stability) (Qian and Akçay 2020).

In other direction, it was pointed out that diversity of interactions could improve stability so that an intermediate proportion of each interaction type in the community was able to stabilize community dynamics (Mougi and Kondoh 2012; Mougi and Kondoh 2014).

In accordance, Qian and Akçay (2020) used sequentially assembled communities and non-linear functional responses to show that population dynamics converges to equilibrium for any combination of types of interactions. Furthermore, the authors demonstrated that communities with many exploitative interactions were less prone to invasion. This result is in accordance with previous results that found that exploitative interactions had the most significant contribution to community stability (Allesina and Tang 2012).

Intensity of Interaction and Stability Apart from interaction topology and types, the intensity of interaction also impacts community dynamics. Once again, this result was firstly found by May (1972b), which indicated that an increased average interaction intensity would lead to unstable communities (see equation (3.62)). Since then, it was shown that most interactions in an ecological network are indeed weak (Paine 1992; Coyte, Schluter, and Foster 2015; Guimarães 2020). Additionally, further theoretical support to the result that "less intense interactions increase stability" was put forward (e.g. Rohr, Saavedra, and Bascompte 2014; Coyte, Schluter, and Foster 2015), but some authors challenged the accuracy of this result (Fowler 2009; Allesina and Tang 2012). Indeed, Serván et al. (2018) pointed out that increased intraspecific competition is beneficial for stability in rich random communities. It was also found that the effect of an increase in the mean of the distribution of coefficients related to interspecific interactions is not as relevant for community resilience as its variance (Allesina et al. 2015). This result is also corroborated by Fowler (2009), who found that increasing variance of interactions diminished the stability of communities if the mean was kept constant.

Throughout our research, we focus on the resilience and coexistence aspects of the community. Following the literature, resilience is measured by considering the rightmost eigenvalues of the interaction matrix, and coexistence is defined as the number of extant populations after system stabilization (e.g. Coyte, Schluter, and Foster 2015; Barabás, J. Michalska-Smith, and Allesina 2016). We also address the community's asymptotic dynamical behavior, identifying equilibrium points, cyclic solutions, and unbounded dynamics.

Chapter 4

Methodology

In this chapter, the techniques employed to generate and analyze the numerical data are presented. Our methodology is centered on the generation of artificial ecological communities. The artificial communities are represented by random networks constructed through a random graph generator with a predefined number of populations, densities of interaction, distributions of intensities, and distributions of types of interactions.

The code used to simulate all data is available through GitHub at <https://github.com/r-menezes/simulation-econetworks>. All the code is based on Python (version 3.6) and Cython (version 0.29) programming languages. Cython is a Python and C/C++ static compiler that allows writing C code with Python syntax and integration of C and Python environments.

In order to summarize the steps of the methodology, we constructed figure 4.1, divided into three parts. Figure 4.1 presents a schematic representation of these stages, with each stage corresponding to a panel. It is possible to see the simulations and analysis workflow and the main elements used to construct and characterize the simulated communities. In the following, we will provide further technical information on each of these stages.

1. **Input Parameters:** the generation of the design of the experiment and construction of the initial network, represented in panel 1 of figure 4.1.

- (a) We constructed an experimental design with 1000 distinct sets of input parameters homogeneously distributed across a preset range satisfying a Latin Hypercube design.
- (b) We performed a combination of the distinct sets of input parameters generated and each of the five different random graph algorithms defined in section 2.2.3. Each graph was associated with each set of input parameters generated by the Latin Hypercube design, amounting to 5000 distinct combinations of algorithms and parameters in total, referred to as *experiments*.
- (c) We generated 10 random seeds for each experiment. Each random seed, associated with the experiment's input parameters, is a *replicate*.
- (d) For each replicate, we used the specified algorithm and input density to construct an undirected random graph. All generated graphs had 400 nodes. In total, 50 000 distinct graphs were constructed, in accordance with the total number of replicates.
- (e) For each edge on the undirected random graph, we associated one type of ecological interaction, the distribution of which was determined by experiments' input parameters. The mutual influence of each adjacent node upon the other was determined as the modulus of a number taken from a normal distribution centered in zero and with standard deviation set by the input parameters. By the end of this process, the undirected network was effectively transformed into a weighted, directed random network, hereafter referred to as *initial network*.

2. **Numerical Integration:** the parametrization of the generalized Lotka-Volterra model from the initial network and the simulation of the temporal evolution of the system, represented in panel 2 of figure 4.1.

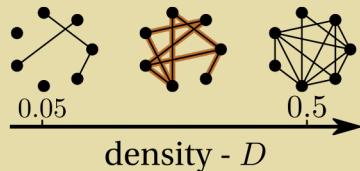
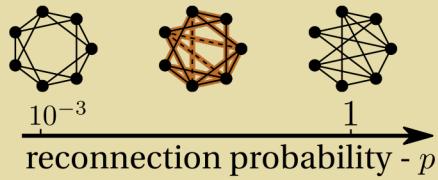
- (a) The matrix of interactions of the generalized Lotka-Volterra model (3.1) was constructed from the initial network's adjacency matrix by assuming all populations have intraspecific competition and setting all elements of the main diagonal to minus one ($a_{ii} = -1$).

Schematic Representation of the Methodology

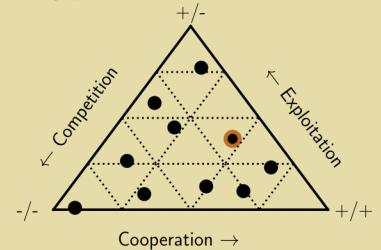
1 Input Parameters

5000 unique parameter combinations: *Experiments*

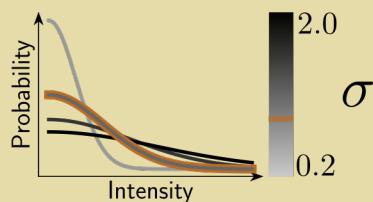
Topology - Random Graph Generator Algorithm



Type of Interaction



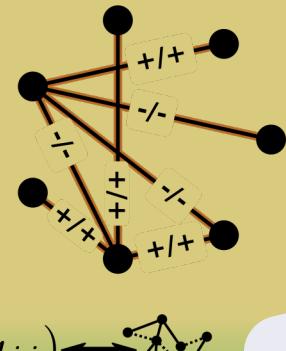
Intensity of Interaction



Random Seed



Initial Network



10 unique random seeds for each *Experiment*: *Replicates*

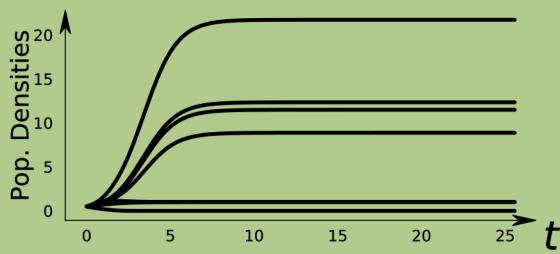
Dynamical System and Complex Network Correspondence

$$(a_{ij}) \longleftrightarrow \text{Network Graph}$$

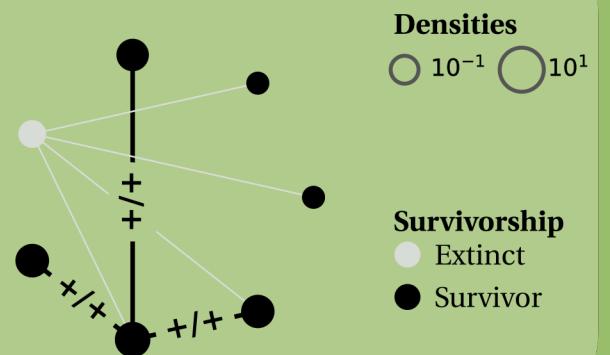
2 Numerical Integration

$$\frac{dX_i}{dt} = r_i X_i \left(1 + \sum_j a_{ij} X_j \right)$$

$r_i = 1$
 $a_{ii} = -1$
 $X_i(t=0) = 1$



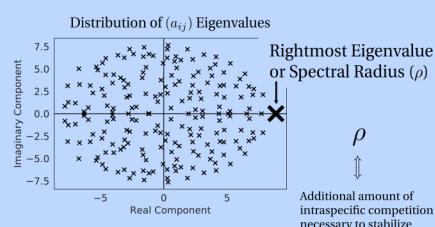
Final Network



3 Measures (Outputs)

Network Measures

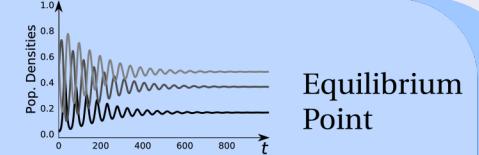
Resilience



Interaction Properties

Coexistence

$$\text{Survival Proportion} = \frac{\# \bullet (\text{Survivors})}{\# \bullet + \# \circ}$$



Equilibrium Point

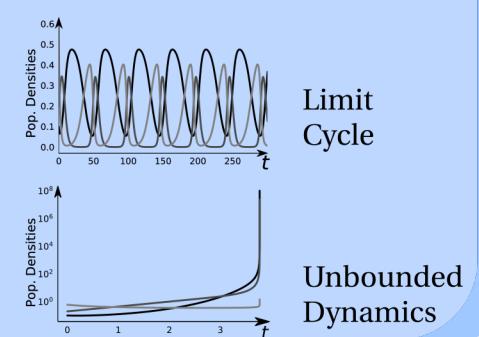


Figure 4.1: (*previous page*) **Schematic Representation of the Methodology:** The colors represent each stage of the methodological pipeline used in this work. A detailed description of each step is provided in the text.

- (b) We simulated the temporal evolution of the community by setting the intrinsic growth rate of all populations to one ($r_i = 1$) and starting the simulation with all populations at a density equal to its carrying capacity ($X_i(t = 0) = 1$). During the temporal evolution, populations that attained densities lower than a millionth (10^{-6}) of the carrying capacity were considered extinct. The simulation was finished when: a) a cycle or equilibrium point was detected; b) any population's density surpassed a million (10^6) times its carrying capacity; c) the time $t = 2000$ was reached.
3. **Measures (Outputs):** measurement of topological and dynamical properties of each replicate, represented in panel 3 of figure 4.1.
- (a) We recorded the attractor detected at the end of the numerical integration of the dynamical model. The algorithmic implementation of the detection of each type of attractors was based on the analytical properties studied in section 3.1.4.
 - (b) The rightmost eigenvalue and other topological measures were logged for the initial network.
 - (c) The final network was constructed by removing nodes and edges associated with extinct populations from the initial network.
 - (d) The dynamical and topological properties of the final network were recorded.

Importantly, we generated networks for 5000 unique combinations of input parameters, considering the topologies, densities, distribution of types of interaction, and standard deviation of interaction intensities. Each unique combination of input parameters will be referred to as *an experiment*. Each experiment was repeated ten times, with different random seeds. The random seed set the initial state of the pseudorandom number generator, leading to distinct topologies due to the intrinsic randomness in the generation of the unweighted network and the distribution of types and intensities of interactions. After the random seed has been determined, the numerical simulation was deterministic. Each repetition of the experiment with a unique random seed is referred as a *replicate* or *scenario*.

4.1 Input Parameters

The first stage in the simulation and analysis of the ecological networks is the parametrization and construction of the initial network. To systematically perform the analysis, it is then necessary to use a method to choose what sets of input parameters will be used in simulations from the unlimited range of possibilities. First, the input variables used to construct the networks are described, and then the Latin Hypercube Sampling (LHS) method is presented.

4.1.1 Network Topologies, Number of Nodes, and Densities

The algorithms to generate random graphs were: a version of Erdős-Renyi random graph; a random regular graph obtained through a configurational algorithm; the algorithm of Watts-Strogatz with probabilities of reconnection being one in one thousand, one in one-hundred and one in ten.

To generate random undirected graphs, we used the random graph generators made available by the NetworkX Python package (Hagberg, Schult, and Swart 2008). The Erdős-Renyi random graph was constructed with the `gnm_random_graph` function. The Watts-Strogatz random and random regular graphs were constructed with `connected_watts_strogatz_graph` and `random_regular_graph` functions, respectively. All these functions have the number of nodes (N) as input, and all networks were constructed with 400 nodes ($N = 400$).

Additionally, all generator functions have a parameter associated with the number of edges in the network (see section 2.2.3). We chose to use network density as the input parameter. Therefore, each function was called with connection-related parameters that allowed for the lowest relative error between the generated network density and the input density. The absolute error between the

input density and density of the generated network was always smaller than 3.7×10^{-3} (median 7.5×10^{-4}).

The Watts-Strogatz algorithm has an additional parameter, which is the rewiring probability (p). To capture distinct regions on the regular-random gradient, we simulated the Watts-Strogatz topology with three distinct rewiring probabilities. Namely $p = 0.001$, $p = 0.01$, and $p = 0.1$, with the $p = 0.001$ ($p = 0.1$) network being more akin to a regular (random) network (Watts and Strogatz 1998).

4.1.2 Types of Interaction

A central aspect of our methodology is the possibility of considering multiple types of ecological interactions in the same community. In accordance, each experiment was associated with a particular proportion of competitive, cooperative, and exploitative interactions.

Through the text, we represent the competitive, cooperative, and exploitative interactions by $\% -/-$, $\% +/+$, and $\% +/-$, respectively. Since this is not a useful notation for algebraic manipulation, consider p_C , p_M , p_E to represent the proportion of Competition, Mutualism (cooperation), and Exploitation. Then, the sum of the proportions of each type of interaction must be equal to one:

$$p_C + p_M + p_E = 1 \quad (4.1)$$

Note that the equation (4.1) implies that it is always sufficient to specify two of the proportions. Therefore, we chose to parametrize the networks using the proportion of cooperative and competitive interactions and derive the proportion of exploitative interactions from (4.1).

A convenient way of representing the proportions of competition, cooperation, and exploitative interaction is through a ternary plot. In a ternary plot, three variables can be represented in a 2-simplex (a triangle) because they must satisfy the condition of summing to a constant total value. Figure 4.2 presents the ternary plot for the types of interactions, indicating the plot's essential elements.

To associate a type of ecological interaction with each pair of interacting populations, we first calculated the number of interactions of each type based on the total number of edges in the (undirected) network, E , and p_C and p_M . The number of competitive and cooperative edges were then $E_C = p_C E$ and $E_M = p_M E$ rounded to the closest integer. The number of edges associated with exploitative interactions was simply $E_E = E - E_C - E_M$.

Second, we performed a random shuffle of the list of edges in the network. We then iterated through this shuffled list, and the first E_C edges were associated with competitive interactions, the next E_M were associated with cooperation. The remaining edges were assigned exploitative interactions. Importantly, if an edge (u, v) was associated with exploitative interactions, there was an equal probability that the interaction would benefit u or v .

After this second step, the final network was directed, as for each pair of connected nodes u and v both the edges (u, v) and (v, u) existed.

4.1.3 Intensities of Interaction

The distribution of interaction strengths in the community was given by the absolute value of a normal distribution centered at zero. In this framework, a larger standard deviation of the normal distribution is associated with a higher proportion of strong interactions. By imposing that the distribution is centered at zero, we ensured that for all distributions considered, there is always a higher prevalence of weaker interactions, as observed in natural systems (Coyte, Schluter, and Foster 2015). As the signal of interactions was predetermined by the type of interactions between the two populations (previous section), only the modulus of each value drawn from this distribution was relevant. In figure 4.3, it is possible to see the probability density function of the interaction intensities for different values of its standard deviation, which serves as an input parameter.

To associate distinct interaction strengths to each interaction in the network, we iterated through each edge in the *directed* network. For each edge, we set its weight to a random number drawn from the normal distribution. Even though each weight could be arbitrarily close to zero, zeros are not allowed, as this would fundamentally change the type of interaction. After allocating strengths to all interactions in the network, each element a_{ij} of the adjacency matrix indicated the

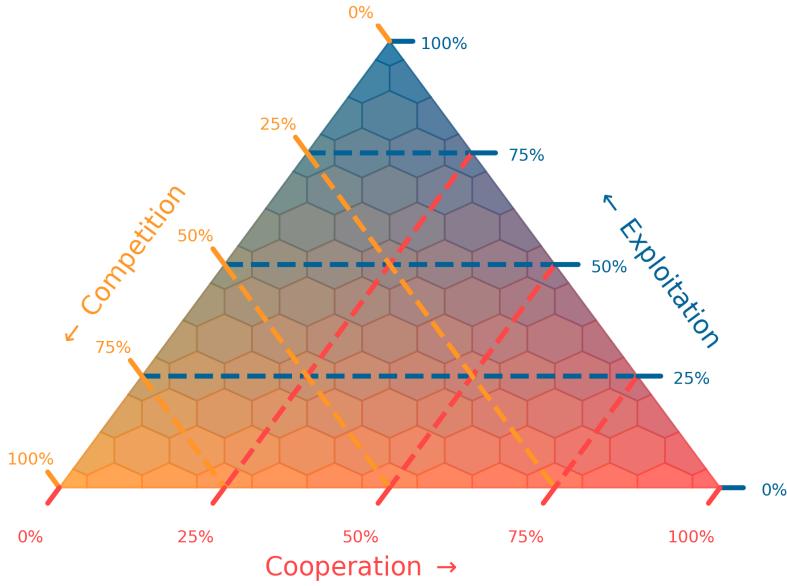


Figure 4.2: **The ternary diagram depicts the proportion of interaction types:** Each of the triangle sides indicates one type of interaction, its color corresponds to legend color, and the dashed lines cross the triangle indicating a constant proportion of the type of interaction with the corresponding color. For instance, blue dashed lines indicate a constant exploitation value, which can be read at the triangle’s left side. The triangle is covered in a hexagonal grid overlay with a color gradient corresponding to the proportions of each interaction type represented by each hexagon. Extreme values of each type of interaction are at the triangle corners and are colored in correspondence to the interaction type: yellow left, red right, and blue top corners corresponding to 100 % of the competition, cooperation, and exploitation, respectively.

influence of population j upon the growth rate of population i . The magnitude of the interaction coefficient a_{ij} is the proportion of change in i ’s growth rate due to j ’s influence when the population j is at its carrying capacity (see section 3.1.1).

Thus, the signal of a_{ij} and a_{ji} was determined by the type of interaction between both populations. The magnitude of both influences was randomly and independently sampled from a zero-centred normal distribution.

4.1.4 Statistical Parametrization

All sets of input parameters were generated before the simulation of the ecological communities. The parameters were generated using a Latin Hypercube Sampling (LHS) method. The LHS method allows for reduced variance when performing Monte Carlo (MC) simulations compared to traditional random sampling (Stein 1987).

The LHS workflow involves determining the number of numerical input variables and generating LHS samples on an n -dimensional hypercube. The input distribution of each variable is then determined, and the resulting samples are then transformed into the input distribution through quantile or inverse distribution functions. Since the Latin Hypercube samples are always uniformly distributed in the range $[0, 1]$, they can be used as quantile values of any other distribution. Therefore, the LHS can provide samples from any statistical distribution, provided this distribution has a well-defined inverse distribution function.

In our study, the standard deviations inputs were uniformly distributed from 0.2 to 2.0. The density was also uniformly distributed from 0.05 to 0.5. However, the proportions of interactions are not as simple since the proportion of cooperation and competition are not entirely independent. In order to guarantee that input variables satisfy the relation (4.1), it was necessary to jointly transform the variables (Frigyik, Kapila, and Gupta 2010).

We considered that the type of interactions follows a Dirichlet distribution with a shape parameter equal to one for the three dimensions. This assumption is equivalent to considering that

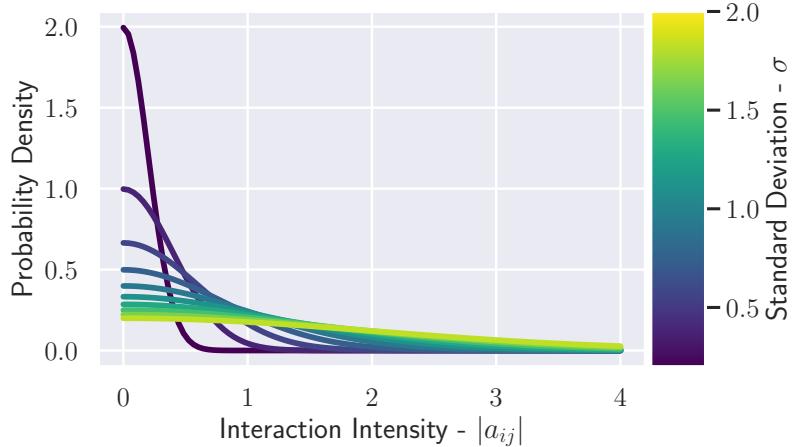


Figure 4.3: **Distributions of Interactions Intensities:** The probability densities of the distributions of the absolute value of the coefficients of interaction for different standard deviations are shown. The colors of the curves correspond to different standard deviations, with brighter indicating larger standard deviations. It is possible to see that the values are more concentrated around the origin for smaller values of the standard deviation. Correspondingly, larger standard deviations are related to flatter distributions, with comparatively more intense interactions.

each possible combination of proportions of competitive, cooperative, and exploitative interactions is equally likely. Under these assumptions, it was then possible to sample from the Dirichlet distribution by transforming three variables from the LHS using a quantile gamma distribution with shape parameter also equal to the unity. When the variables' values were normalized by their sum, the result satisfied relation (4.1) and was sampled from the goal Dirichlet distribution (Frifyik, Kapila, and Gupta 2010; Carnell 2020). Thus, this method allows for the uniform sampling of all possible combinations of proportions of types of interactions while preserving the LHS properties.

Since only two variables are needed to completely determine the types of interaction, we only used the proportion of competition and cooperation as input variables. The python function developed to make this transformation was based on a solution presented by Carnell (2020). Note that the mean values of the proportion of competition and cooperation were $1/3$ since all combinations of types of interactions in the 2-simplex were equally probable.

The LHS was designed with Python package pyDOE2 (Sjögren and Svensson 2020). All variable transformations were performed with our code and SciPy's functions (Virtanen et al. 2020). Among the different possible criteria for the Latin Hypercube Sampling provided by the function `lhs` from the pyDOE2 package, we used the `centermaxmin` criterion. This criterion provided considerable average distance and low correlations samples, as it can be seen on table 4.1. Generally, it is desirable to have larger distances and smaller correlations between samples.

4.1.5 Replicates

For the simulation of the ecological communities, a random seed is a necessary additional input. This random seed guarantees the reproducibility of the simulations since it determines all random processes. For a given seed, the simulation is entirely deterministic.

For each set of input parameters (without the seed), ten distinct seeds were generated through the `SeedSequence` function from NumPy's random module (Harris et al. 2020). The random generator used (PCG64) has a period of 2^{128} .

In total, one thousand distinct sets of input parameters (proportion of competition and cooperation, density, and standard deviation of intensities) were generated using the LHS method. Every distinct set of input parameters was associated with each of the five topologies considered, amounting to 5000 unique *experiments*. Ten unique seeds were associated with each experiment. Overall, 50 000 unique *replicates* or *scenarios* were simulated.

method	distance				correlation		
	max	min	ave	median	max	min	ave
center	1.96	0.02	0.88	0.88	0.01	-0.01	7.0×10^{-3}
maximin	1.98	0.02	0.88	0.88	0.02	-0.02	8.1×10^{-3}
centermaximin	2.01	0.03	0.88	0.88	7.6×10^{-3}	-0.01	5.8×10^{-3}
correlation	2.00	0.02	0.88	0.88	6.6×10^{-3}	-0.02	7.5×10^{-3}
None	2.01	0.02	0.88	0.88	0.01	-0.02	9.6×10^{-3}

Table 4.1: **The characteristics of different LHS generation methods.** For each method, a Latin Hypercube design of experiment was constructed with 10 000 samples across 5 dimensions, with samples uniformly distributed between 0 and 1 in each dimension. Each sample is thus a five-dimensional vector with entries as the outputs of the LHS algorithm. The Euclidean distance and correlations between samples was calculated. The statistical descriptives for distance and correlation are presented.

4.2 Temporal Evolution of Ecological Communities

The dynamical generalized Lotka-Volterra (gLV) model was parameterized from the ecological network by setting the off-diagonal elements of the interaction matrix \mathbf{A} to the values present of the directed network's adjacency matrix $(\mathbb{M})_{ij} = m_{ij}$, i.e. $a_{ij} = m_{ij} \forall i \neq j$. The diagonal elements of the interaction matrix were chosen to be minus one to account for intraspecific competition $a_{ii} = -1 \forall i$. Additionally, in order to focus on the interaction matrix, the growth rate of all populations was set to one ($r_i = 1 \forall i$). The gLV and its Jacobian matrix were implemented computationally using matrix-multiplication algorithms from the Basic Linear Algebra Subprograms (BLAS) library made available through SciPy for improved performance (Virtanen et al. 2020). Although we integrated the parametrized Lotka-Volterra model (3.1) numerically, the analytical background that guarantees the asymptotic behavior was presented in section 3.1.4.

In order to integrate numerically the parametrized lotka-volterra model we set all the populations initial value to half of their carrying capacity ($K_i/2 = -2/a_{ii} = 1/2$). The numerical integration was performed through the LSODA algorithm, available through SciPy (Virtanen et al. 2020), with prescribed absolute and relative tolerances being 1×10^{-7} and 1×10^{-6} , respectively. This initial value problem solver is convenient for its varying time-steps and automatic detection of stiffness. Additionally, SciPy provides local interpolants of the numerical solution so that it is possible to get values for the densities of the populations at times that did not correspond to any particular iteration of the LSODA integrator.

After initialization, the integrator was iterated repeatedly until a stopping condition was found. At each iteration, the density of each variable was checked. If any population had a density above a million times its carrying capacity ($X_i \leq 1 \times 10^6$), the replicate would be associated with an unbounded attractor. If the absolute value of the time derivative of each population's densities were smaller than ten times the prescribed absolute tolerance of the integrator, the replicate would be associated with an equilibrium attractor.

The detection of cycles occurred after a transient time ($t = 100$) had elapsed and then once every 100 integrator iterations. Then, interpolators spanning a time interval of 256 time units were sampled at 1024 equally spaced intervals for all populations to detect the cycles. Note that each iteration can correspond to different time increments due to varying time-steps of the algorithm. In sequence, populations that eventually reached densities lower than 1×10^{-6} were not taken into account, as they were virtually extinct. We subtracted the population-level density mean from the vector of each population's densities to prepare this vector for a Fourier transformation. We then performed a fast Fourier transformation using the `rfftn` function provided by SciPy. Subsequently, we used SciPy's function `find_peaks` to find the peaks on the Fourier transformed densities of each population individually. The peaks were then compared to check for common peaks for all populations, which were then inversely Fourier transformed and used as putative periods. Since the putative period could be slightly off the system's actual period, we did a numerical optimization of the period. This optimization was performed by choosing the period that minimizes the maximum absolute distance between the vector of densities $\vec{X}(t)$ and $\vec{X}(t-T')$, in which T' was

the putative period. This optimization was performed using SciPy's function `minimize_scalar`. Finally, it was checked if the maximum absolute distance between the vectors $\vec{X}(t)$ and $\vec{X}(t - T')$, and $\vec{X}(t)$ and $\vec{X}(t - 2T')$ was within ten times the absolute tolerance of the integrator. If the previous condition was true, the period was recorded, and the replicate would be associated with a cyclic attractor.

The numerical integration was terminated whenever an attractor was detected. If no attractor was detected, the integration was finished at $t_f = 2000$, and the replicate was associated with transient dynamics. After the simulation of all 50 000 replicates, we investigated transient attractors further, as indicated in section 4.2.1. We tested the detection of attractors by using parametrizations with known dynamics reported in the literature as described in the appendix A. We also plotted the temporal evolution of several simulations and investigated the attractor detection accuracy by visual inspection (see the appendix A).

Numerical errors occurred for some replicates. When this was the case, we drew a distinct random seed and repeated the simulation. That is, the replicate was changed, but the experiment was maintained.

4.2.1 Asymptotic Dynamics detected as function of Computational Effort

After the initial simulation of all replicates, a portion of these replicates could have been associated with "transient" attractors, i.e., no equilibrium, cycle, or unbounded dynamics was detected. To further study these replicates, we simulated the communities again, adjusting the numerical tolerances of the integrator. We diminished the numerical tolerances, with absolute and relative tolerances being 1×10^{-9} and 1×10^{-8} , respectively. Additionally, we increased the temporal range of the simulation by setting $t_f = 6000$. We also placed a relatively less restrictive condition on the detection of cycles by requiring the population densities to be separated by one and two periods to be equal within 50 times the prescribed tolerances.

For instance, in our simulations, a total of 225 replicates were initially associated with "transient" attractors. Following the more precise numerical integration, we found that 11 of the 225 replicates were associated with cycles. In accordance, 16 replicates were associated with equilibrium points, and 69 displayed unbounded dynamics. This information is represented in figure 4.4.

There could be replicates still being classified as possessing transient dynamics after simulating such replicates with improved numerical precision. We performed a further reanalysis by drawing new random seeds and effectively substituting these replicates to remove these transients. In this final step, whenever a simulation resulted in transient dynamics, we drew another random seed and repeated the simulation with this updated replicate.

In our simulations, a total of 129 scenarios remained unspecified after the simulation with improved numerical precision. After new seeds were introduced, three additional replicates were associated with cyclic attractors, totaling 14 cyclic scenarios. Another 30 replicates were associated with equilibrium points, with a total of 46 replicates. Unbounded dynamics were detected in 165 scenarios overall, representing an increase of 96 scenarios. The distribution of attractors detected at each stage is pictured in figure 4.4. The final distribution of attractors is presented in the results section 5.3.

4.2.2 Final Community

We considered that the populations which reached less than a millionth of its carrying capacity were virtually extinct, i.e., whenever $X_i < 1 \times 10^{-6}$ we considered the i -th population to be extinct. To construct the final community, we removed all virtually extinct populations and their connections from the initial community network. As the removal of nodes and edges could result in a disconnected graph, we then considered the largest connected subgraph to be the final community. This procedure is in accordance with the definition of a community as a set of interacting populations in an ecosystem (see section 2.3.2). All topological measures from the final network were based on this largest connected subgraph. Other measures, however, were based on all nodes, connected or not. For instance, the proportion of survival was taken by dividing the number of populations with final densities larger than 1×10^{-6} by the number of initial populations (400 in our simulations).

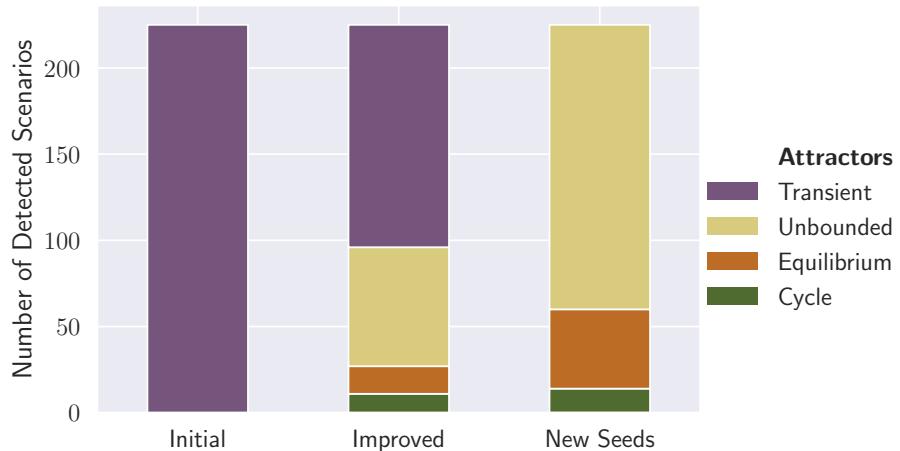


Figure 4.4: Investigation of Transient Dynamics: The removal of transient dynamics was made in two stages. First, the temporal evolution was extended and numerical precision was increased to allow for convergence ("Improved" column). Second, the remaining transient dynamics were allowed to get a new random seed, which slightly changed the experiment replicates ("New Seeds" column). For both "Improved" and "New Seeds" updated simulations, 20% of the scenarios corresponded to equilibrium points, near 74% corresponded to unbounded dynamics, and 6% and 2% respectively were detected to be cycle dynamics.

4.3 Measurements

Aiming at characterizing the ecological communities, we measured different aspects of the community. The measures were taken from the initial, putative, community or the final, extant, community. Some measures were based both on the initial and final communities, as it is the case of topological properties. We divide the measures into three categories, corresponding to which aspect of the ecological community they represent: Topology, Interaction, and Dynamics. The measures are discussed in more detail in the following. A complete list of names, symbols and meanings of the variables is presented in table B.1

4.3.1 Topological Properties

The topological variables measured were average shortest path length (l_G), degree assortativity (r_D), density (D), diameter (d), global clustering (C), average local clustering (\bar{C}), number of modules, modularity (Q), order (N) and size (E). All these variables were calculated using algorithms provided by the networkx and networkkit Python packages (Hagberg, Schult, and Swart 2008; Staudt, Sazonovs, and Meyerhenke 2016). The definitions of these variables are presented and discussed in the network background section (2.1). All topological variables are measured for undirected networks in which two nodes (populations) are connected if and only if there exists an ecological interaction between them.

The input variables related to these measurements were the density and the particular random graph algorithm used to construct the networks.

4.3.2 Interaction Properties

We characterized the communities' interaction properties by measuring the mean and variance of the set $\{a_{ij}\}$. Additionally, we measured the correlation of the coefficients indicating the influence of species j upon species i and species i upon species j , i.e., a_{ij} and a_{ji} .

The input variable most related to the interaction was the standard deviation associated with the interaction intensities' normal distributions.

4.3.3 Dynamical Properties

The most fundamental aspect of the dynamics is to characterize the type of dynamics, or attractors, of each simulated community. Depending on the attractor, other dynamical variables are then measured.

In particular, we recorded the proportion of survival (% surv.), the populations' abundances at the end, and the abundance assortativity in the final community for communities associated with cyclic (CY) and equilibrium attractors (EP). Additionally, we recorded the number of subcommunities (n. comm.) at the end, i.e., the number of connected components after removing virtually extinct populations from the network. We also recorded the average amplitude of oscillation and period (T) for communities with cyclic attractors. Exclusively for communities with equilibria, we recorded the eigenvalues at the final time-step ($\{\lambda_i\}$) and the real part of the rightmost eigenvalue (λ_{max}). Finally, we recorded the final time for all attractors (t_f) and individually for communities at equilibrium (EP t_f). The complete list of dynamical variables measured is presented on table 4.2.

Measurement	Abbreviation/Symbol	Type of Dynamics
Abundance	abund.	EP, CY
Abundance Assortativity	abund. assort.	EP, CY,
Average Amplitude	ave. amp.	CY
Eigenvalues at EP	λ_i	EP
Final Time	t_f	EP, CY, , UN
Number of Populations Coexisting	N populations	EP, CY
Period	T	CY
Proportion of Survival	% surv.	EP, CY
Real Part Rightmost Eigenvalue	λ_{max}	EP
Time to EP	EP t_f	EP

Table 4.2: **List of dynamical variables:** The type of dynamics for each variable measured is presented. Equilibrium Point, Cycle and Unbounded dynamics are denoted by EP, CY and UN, respectively.

4.4 Techniques of Analysis

In this section, the methodology used to analyze the data recovered from the simulations is presented. We focus our efforts on describing the techniques employed in constructing non-standard figures and analysis with more detail.

Most data handling was performed through Python's library pandas (McKinney 2010). Additional functions from the NumPy library were also used (Harris et al. 2020). All plots were made using Python's libraries matplotlib, seaborn, and python-ternary (Hunter 2007; Waskom and team 2020; Harper et al. 2019). Networks were plotted with networkx and matplotlib (Hagberg, Schult, and Swart 2008; Hunter 2007).

4.4.1 Measures of Variability

We measured variability in the outcome of each experiment through the coefficient of variation. The coefficient of variation is defined as the ratio between the standard deviation and the mean's absolute value for a set of observations. Mathematically, this measure is defined as

$$CF = \sigma/\mu \tag{4.2}$$

$$= \sqrt{E[X^2] - (E[X])^2}/E[X] \tag{4.3}$$

This quantity is most suited for variables whose mean is not close to zero, as low values for the mean could lead to high values of the coefficient without this being related to large standard deviations. The coefficient of variation is useful as it allows for comparisons of the amount of variance in different measures.

4.4.2 Stacked Area Plot

To create a stacked area plot for the types of attractors, we first constructed a kernel density estimator of each attractor's occurrence along a continuous axis. This kernel density estimator was normalized to one, indicating the probability that a randomly selected replicate of a given attractor would have each value of the continuous variable. For instance, when considering input density as the continuous variable, the cyclic attractors' kernel density estimator provided the probability density function related to the densities of all replicates associated with cyclic attractors.

When all the appropriate kernel densities had been measured, we normalized the values once more, considering that, at each level of the continuous variable, the estimates of the probability densities of all attractors must add up to one. After this process, the normalized kernel densities associated with each attractor could be plotted and provided a reliable estimative of each attractor's relative occurrence for a given level of the continuous variable.

4.4.3 Heatmaps

To construct the heatmaps, we defined an aggregation function and then applied this aggregation function to all replicates whose variables occurred inside a unit area of the heatmap. The aggregated value was then associated with that area and plotted with a color according to its value. We used two distinct types of aggregation functions: the proportion of bounded attractors and the median of the rightmost eigenvalue.

When assessing the proportion of bounded attractors, we counted the number of bounded scenarios inside a given area of the heatmap. We then divided the number of bounded scenarios by the number of all scenarios occurring inside that area. This procedure was then repeated for all areas in the heatmap. When assessing the median of the rightmost eigenvalue, we considered the rightmost eigenvalue of all replicates that occurred inside a given region of the heatmap and recorded their median.

Many of the heatmaps were constructed on the ternary diagrams. The principles of constructing heatmaps on the ternary diagrams are the same as when constructing them on regular square grids. The only difference is that the ternary diagram's hexagonal regions require that the process of filtering which replicates fall into each hexagon to be performed with additional functions. Therefore, we have developed a filtering function that identifies the linear relationships between the proportion of cooperation and the proportion of competition of the replicate that defines each hexagonal region's borders. Since these six identified lines define the plane's hexagonal region, we then filtered all the replicates with a proportion of cooperation and proportion of competition that fell into the given hexagonal region. Once the aggregated value of each hexagon was obtained, we plotted the ternary diagram using python-ternary.

4.4.4 Scatter Plots and Regressions

The scatter plots were plotted with the aid of the plotting libraries seaborn and matplotlib (Waskom and team 2020; Hunter 2007). The trends are automatically calculated and plotted by seaborn, but in order to display its equation, we recalculated the regressions using the Python library statmodels (Seabold and Perktold 2010). The same library was also used to calculate all the linear regressions presented throughout this work.

Chapter 5

Results

In this chapter, the results of this research are presented. The results are divided in six sections. The first section (section 5.1) presents summary statistics of the measured variables and the correlations among them. These correlations show that it is possible to reduce the number of variables analyzed. The second section (section 5.2) addresses the variability present in each experiment, how much each property varied among replicates in the same experiment and overall. The third section (section 5.3) show the distribution of attractors in the simulated communities, and its dependence on input parameters.

Section four (section 5.4) presents our results concerning how coexistence of populations in a community is related to density, intensity and types of interaction. Lastly, in section five (section 5.5), we discuss the resilience of communities, and how are they related to parameters used to construct the community.

5.1 Measured Variables Profile

In the following section, we present the statistical profile of the variables measured for each replicate (see section 4.3). We characterize the measurements of each variable by presenting a set of summary statistics. Specifically, we present the total count, mean, standard deviation, minimum, maximum, and first through third quartiles. While these values do not fully determine the distribution, they provide valuable insights corresponding to the range, spread, and central tendency for each variable. Additionally, we perform a correlation clusterization of variables and identify several groups of highly correlated measured variables. These sets of variables provide redundant information concerning the structure or dynamics of the simulated ecological networks. Thus, we choose one representative of each group to study in subsequent analysis.

5.1.1 Descriptive Statistics

In the following, variables are grouped according to the aspect of structure/stability it characterizes. Therefore, we present the descriptive properties of variables related to dynamics, interaction, and topology. Several of these variables are measured once for the initial network and then for the final network. In such cases, the variable name is suffixed by 0 or f , indicating a measure performed for the initial or final network, respectively.

The results put forward in this section provide an initial overview of the data gathered from numerical simulations and serve as preparation for more complex analysis.

Dynamical Properties/Variables The dynamical properties entail information concerning densities, state space dynamics, number of populations, and connected components of the networks. In particular, these variables provide information concerning communities' resilience and coexistence through the rightmost eigenvalue and number of surviving populations.

In table 5.1 we present the summary statistics for abundance assortativity (abun. assort.), average amplitude (ave. amp.), the number of populations in the final community (N populations), and the proportion of population survival (% surv.).

	abun.	assort.	ave.	amp.	N	populations	% surv.
n	288		19		288	288	
mean	0.11		8.19		134.9	33.74	
std	0.15		14.75		83.30	20.82	
min	-0.12		1.13		13	3.25	
25%	6.2×10^{-3}		3.48		51	12.75	
50%	0.07		4.65		148	37	
75%	0.19		5.79		175	43.75	
max	0.81		68.31		295	73.75	

Table 5.1: **The descriptive statistics of dynamical variables (I):** We present the number of observations (n), mean, standard deviation (std), minimum and maximum values (min, max) and the 25%, 50% and 75% quartiles that characterize the distribution of the abundance assortativity, average amplitude, number of extant populations and proportion of survival. The 50% quartile is also referred to as the median of the data.

The number of observations of each variable allows us to conclude that from the 50 000 scenarios considered, only 288 did not lead to unbounded dynamics. Therefore, almost all sets of input variables led to the unbounded growth of at least one population 4.2. By analogous reasoning, equilibrium point and cyclic attractors were observed in 269 and 19 scenarios, respectively.

It is possible to see that most bounded scenarios had weakly assortative abundance distribution across nodes as indicated by the mean and 50%, and 75% percentiles. Nevertheless, the standard deviation and distance between median and mean indicate high variance and the presence of assortative outliers. Similarly, the average amplitude of the detected cycles had a large standard deviation compared to its mean value. Average amplitudes were always greater than one, indicating that the oscillations in populations' densities were larger than populations' carrying capacity in all cases.

The average state space speed is dominated by many small values, as indicated by the quartiles' relatively low values. However, the maximum and the standard deviation indicate a few large values that may be outliers. Accordingly, the distance to equilibrium point displays the same pattern, with many small values and few outliers indicated by maximum, standard deviation, and distance between mean and the quartiles.

The number of surviving populations indicates that at least 13 and at most 295 populations coexisted in the final community. In accordance, the proportion of survival indicates that a third of the total number of populations survive, on average. Overall, 135 ± 83 (mean \pm standard deviation) populations coexisted in the final community.

In the table 5.2, we present summary statistics of the number of weakly connected components in the initial and final networks ((0) n comm. / (f) n comm.), the period of cyclic solutions (T), initial and final rightmost eigenvalue (λ_{max0} , λ_{max}), final time (t_f), and final time for equilibrium points (EP t_f).

As a consequence of our methodology, all initial networks had a single connected component. The quartiles of the number of connected components in the final network also indicate that most networks comprised a single connected component. Few scenarios displayed some level of disaggregation and were composed of up to 22 connected components. Therefore, it is possible to deduce that despite the extinction of an average of two-thirds of interacting populations, the community generally remained connected. The periods of scenarios with cyclic dynamics ranged from 1 up to 120 units of time, indicating a large variance on this measure.

The initial rightmost eigenvalue (λ_{max0}) of all communities had positive real components, indicating that no community would be stable in the absence of intraspecific competition. Only a small proportion of communities satisfied the condition $\lambda_{max0} < 1$, which would guarantee the stability of an internal equilibrium point. Upon inspection, it was verified that all the 19 initial communities satisfying such condition were associated with equilibrium points. Generally, values of λ_{max0} occurred in the range [0.9, 196], with mean 24 and standard deviation 27, indicating that the parametrization typically led to nonresilient communities. Note that the initial eigenvalues are the eigenvalues of the interaction matrix with intraspecific competition set to zero, i.e., $\mathbf{A} + \mathbf{I}$.

	(0) n comm.	(f) n comm.	T	λ_{max0}	λ_{max}	t_f	EP t_f
n	5.0×10^4	288	19	5.0×10^4	269	5.0×10^4	269
mean	1	1.11	36.29	24.10	0.04	26.18	1219.6
std	0.0	1.28	40.31	27.07	0.13	231.5	1204.2
min	1	1	1.05	0.91	-0.29	0.01	63.86
25%	1	1	10.22	8.26	-7.2×10^{-3}	0.08	509.6
50%	1	1	16.81	14.94	-1.4×10^{-3}	0.23	799.0
75%	1	1	44.47	27.44	0.04	1.17	1481.9
max	1	22	122.2	196.0	0.69	5961.2	5961.2

Table 5.2: **The descriptive statistics of dynamical variables (II):** We present the number of observations (n), mean, standard deviation (std), minimum and maximum values (min, max) and the 25%, 50% and 75% quartiles that characterize the distribution of the number of communities in the initial and final networks (n. comm.), the period (T), real part of the rightmost eigenvalue of the initial community (λ_{max0}), rightmost eigenvalue at the equilibrium point (λ_{max}), final time (t_f) and time at the detection of equilibrium points (EP t_f). The 50% quartile is also referred to as the median of the data.

However, after communities' temporal evolution, the distribution of rightmost eigenvalues changed. Most final rightmost eigenvalues were close to zero, as indicated by the mean and quartiles. The 75% quartile indicates that some communities with equilibrium points were associated with positive, albeit small, values of λ_{max} . The final eigenvalues are calculated for the Jacobian matrix evaluated at the final densities.

The final time of most scenarios was considerably small, with the 75% percentile of the data being less than 2 units of time. On the other hand, when considering only equilibrium points final times, more than 50% of scenarios presented final times in the range of 500 to 1500.

Interaction Variables Interaction variables convey much of the information concerning the intensity, distribution, and types of ecological interactions among populations. The summary statistics of the mean, variance, and correlation of ecological interaction coefficients are presented in table 5.3.

	a_{ij} mean (μ)	a_{ij} variance	a_{ij} correlation
n	5.0×10^4	5.0×10^4	5.0×10^4
mean	4.6×10^{-3}	0.40	0.20
std	0.12	0.39	0.30
min	-0.54	2.5×10^{-3}	-0.60
25%	-0.04	0.09	-6.3×10^{-3}
50%	1.5×10^{-3}	0.26	0.23
75%	0.05	0.61	0.46
max	0.49	1.96	1

Table 5.3: **The descriptive statistics of interaction variables:** We present the number of observations (n), mean, standard deviation (std), minimum and maximum values (min, max) and the 25%, 50% and 75% quartiles that characterize the distribution of the interactions mean value (a_{ij} mean), variance (a_{ij} variance) and correlation (a_{ij} correlation). The 50% quartile is also referred to as the median of the data.

The mean of interaction coefficients in each community was distributed roughly symmetrically around zero. Therefore, half of the simulated scenarios had interaction coefficients with positive mean, and the other half was associated with negative mean. The interaction coefficients variance covered a range of values between 1×10^{-3} to 2, with mean 0.4. The correlation among coefficients in a community was, on average, slightly positive. Nevertheless, anticorrelated and highly

correlated distributions of interaction coefficients were registered, ranging from -0.6 to 1 .

Topology Variables The measured topological variables characterize the network structure of putative and final communities. In table 5.4, summary statistics for initial and final average minimum path length (l_G) and degree assortativity (r_D) are presented.

	l_{G0}	l_{Gf}	r_{D0}	r_{Df}
n	5.0×10^4	288	4.0×10^4	288
mean	1.98	2.90	-1.9×10^{-3}	0.28
std	0.67	1.58	7.9×10^{-3}	0.29
min	1.50	1.55	-0.05	-0.27
25%	1.63	1.85	-6.4×10^{-3}	5.1×10^{-3}
50%	1.78	2.37	-2.6×10^{-3}	0.26
75%	2.05	3.11	1.9×10^{-3}	0.55
max	10.04	9.23	0.07	0.81

Table 5.4: **The descriptive statistics of topological variable (I)s:** We present the number of observations (n), mean, standard deviation (std), minimum and maximum values (min, max) and the 25%, 50% and 75% quartiles that characterize the distribution of the average minimum path length (l_G) and degree assortativity (r_D). The measures taken from the initial network are indicated by a 0 subscript, while the measures taken from the final network are indicated by a f subscript. The 50% quartile is also referred to as the median of the data.

The initial networks typically had a small average shortest path length with mean and median below 2, despite the relatively broad range (1.5 to 10). The average shortest path length was higher in the final community, with a mean close to 3. This increase in path lengths occurs independently of the reduced number of nodes in the final network compared to the initial network. The final network also present increased standard deviations in comparison with its initial counterpart. Importantly, the average shortest path length was always defined since all networks were connected.

Initial networks were consistently disassortative, as indicated by minimum and maximum values close to zero. In opposition, final networks were, on average, assortative - albeit with coefficients close to zero. Overall, these networks had a broad range of assortativity values, ranging from disassortative to highly assortative (-0.27 to 0.81).

The table 5.5 displays descriptives for initial and final density (D), diameter (d), and global clustering (C).

	D_0	D_f	d_0	d_f	C_0	C_f
n	5.0×10^4	288	5.0×10^4	288	5.0×10^4	288
mean	0.27	0.18	3.06	5.48	0.52	0.50
std	0.13	0.12	1.66	3.74	0.22	0.26
min	0.05	0.05	2	2	0.04	0.04
25%	0.16	0.06	2	3	0.33	0.24
50%	0.28	0.14	3	4	0.58	0.56
75%	0.39	0.23	3	6	0.73	0.72
max	0.50	0.47	20	19	0.75	0.85

Table 5.5: **The descriptive statistics of topological variables (II):** We present the number of observations (n), mean, standard deviation (std), minimum and maximum values (min, max) and the 25%, 50% and 75% quartiles that characterize the distribution of the density (D), diameter (d) and global clustering (C). The measures taken from the initial network are indicated by a 0 subscript, while the measures taken from the final network are indicated by a f subscript. The 50% quartile is also referred to as the median of the data.

The density of initial communities reflects the uniform distribution of input densities from 0.05 to 0.5. However, the final densities of final communities were lower, as indicated by the quartiles and mean. Mean and standard deviations of diameter for final graphs were also higher than initial ones. This increase in diameter occurs independently of the decrease in the order and size of the final network. On the other hand, the distribution of global clustering was roughly similar across initial and final networks. On average, the networks had a clustering coefficient around 0.5 and ranging from 0.2 to 0.8.

Next, the summary statistics for initial and final local clustering (\bar{C}), modularity (Q), and the number of modules (n modul.) are presented in table 5.6.

	\bar{C}_0	\bar{C}_f	Q_0	Q_f	(0) n modul.	(f) n modul.
n	5.0×10^4	288	5.0×10^4	288	5.0×10^4	288
mean	0.52	0.51	0.29	0.45	4.89	5.09
std	0.22	0.26	0.20	0.20	2.38	2.12
min	0.04	0.02	0.04	0.09	2	2
25%	0.33	0.23	0.08	0.21	3	3
50%	0.58	0.58	0.32	0.49	3	4
75%	0.73	0.72	0.46	0.62	7	7
max	0.75	0.87	0.76	0.81	13	12

Table 5.6: **The descriptive statistics of topological variables (III):** We present the number of observations (n), mean, standard deviation (std), minimum and maximum values (min, max) and the 25%, 50% and 75% quartiles that characterize the distribution of the average local clustering coefficient (\bar{C}), modularity (Q), and number of modules in the network. The measures taken from the initial network are indicated by a 0 subscript, while the measures taken from the final network are indicated by a f subscript. The 50% quartile is also referred to as the median of the data.

Despite having different definitions, the global average of local clustering and global clustering had very similar distributions both in initial and final networks. In contrast, the modularity was higher for final communities than for initial ones. On average, typical values for the modularity of initial and final networks were $Q_0 \approx 0.2 \pm 0.2$, and $Q_f \approx 0.45$. Despite the difference in modularity, the number of modules was similar between the two groups. Both initial and final networks had around 5 modules on average.

Lastly, table 5.7 contains summary statistics for the number of nodes (N) and edges (E).

	N_0	N_f	E_0	E_f
n	5.0×10^4	288	5.0×10^4	288
mean	400	134.8	2.2×10^4	1211.9
std	0.0	83.44	1.0×10^4	917.8
min	400	13	4000	26
25%	400	51	1.3×10^4	267.2
50%	400	148	2.2×10^4	1063
75%	400	175	3.1×10^4	2182.8
max	400	295	4.0×10^4	3345

Table 5.7: **The descriptive statistics of topological variables (IV):** We present the number of observations (n), mean, standard deviation (std), minimum and maximum values (min, max) and the 25%, 50% and 75% quartiles that characterize the distribution of the number of nodes(N) and edges (E). The measures taken from the initial network are indicated by a 0 subscript, while the measures taken from the final network are indicated by a f subscript. The 50% quartile is also referred to as the median of the data.

As expected by the employed methodology, the number of nodes and edges decreased when comparing initial and final networks. Initial communities always had the same number of populations (400) as indicated by N_0 . Initial number of edges followed the input distribution of densities, thus being fairly uniform between $N_0 \times (N_0 - 1) \times d_{0\min}/2 \approx 4000$ and $N_0 \times (N_0 - 1) \times d_{0\max}/2 \approx 40\,000$. The number of edges in the final network is much smaller than the number of edges in the initial network. While most initial networks had more than 2×10^4 edges, final networks have less than 4×10^3 in all cases. This reduction decreases the number of edges by at least one order of magnitude. It is relevant to note that number of nodes in the final graph is always equal or larger than the number of populations in the final community.

5.1.2 Correlations

The study of correlations among variables contributes to the overall analysis in different ways. It is possible to use correlations as an indication of results that should be further investigated. Also, since highly correlated variables convey similar information, we reduce the parameters to be analyzed by avoiding redundancy on the investigated variables. In itself, the correlation provides valuable insight on the interplay of multiple variables, albeit *correlation does not imply causation*.

The correlations among all variables are presented in figure B.1. From all the correlations, we highlight several sets of variables that presented large Pearson correlation (ρ) values with each other. The list of all highly correlated variables, $\rho > 0.95$ is

- **Final time** (t_f) and final time for equilibrium point attractors (EP t_f) [$\rho = 1$.];
- Average state space speed (ave. SS-speed) and **distance to equilibrium point** (EP dist.) [$\rho = 1$.];
- Initial/final **global clustering** (C) initial/final average local clustering (\bar{C}) [$\rho \geq 0.99$ for all pairs];
- Initial **average shortest path** (l_{G_0}) and initial diameter (d_0) [$\rho = 0.96$.];
- Final **average shortest path** (l_{G_f}) and final diameter (d_f) [$\rho = 0.98$.];
- Number of populations in the final community (N populations) and **proportion of survival** (% surv.) [$\rho = 1.0$.];
- Initial density/final density (D_0 / D_f), **input density**(D_{in}) and initial number of edges (E_0) [$\rho > 0.98$ for all pairs].

To avoid the repetition of analysis, we conduct further analysis solely the variable highlighted in bold for each set. Additionally, strong negative correlation between the period and initial degree assortativity ($\rho = -0.83$) and strong positive correlation between initial and final average shortest path length ($\rho = 0.88$) were detected.

5.2 Variability in Experiments

5.2.1 Variability of Measured Quantities

For each unique set of input parameters, the community simulation was repeated ten times, with different random seeds (see section 4.1.4). Since the initial networks are randomly generated, the output from each seed is unique. Therefore, we computed the coefficient of variation (CV) to address how similar the networks generated with the same set of input parameters were. The CV is the ratio of the standard deviation to the mean. Since this measure is normalized, it is possible to compare its values for different data distributions: Smaller values indicate less variance around the mean, and larger values are associated with increased variance. The coefficients of variation calculated for each unique experiment individually are shown in figure 5.1. For reference, we also present the CV calculated across all generated networks.

In general, the coefficient of variation calculated for each experiment's replicates was almost one order of magnitude smaller than its counterpart calculated for the entire dataset. Some few exceptions to this rule were measures related to assortativity, the real part of the rightmost eigenvalue, and the average state space speed. However, as indicated in the previous section (5.1.1),

these measures' means were generally very close to zero, which hinders the interpretation of the coefficient of variation. As a reference, observe that the seed displayed a global and experiment-based coefficient of variation close to one (10^0). Indeed, since the seed is sourced from a random number generator, its distribution is independent of the experiment.

Most of the measured variables had coefficients of variation around or below 10% (dotted white line). Therefore, despite the variability inherent to the employed methodology, measures taken on distinct replicates of the same experiment were relatively similar.

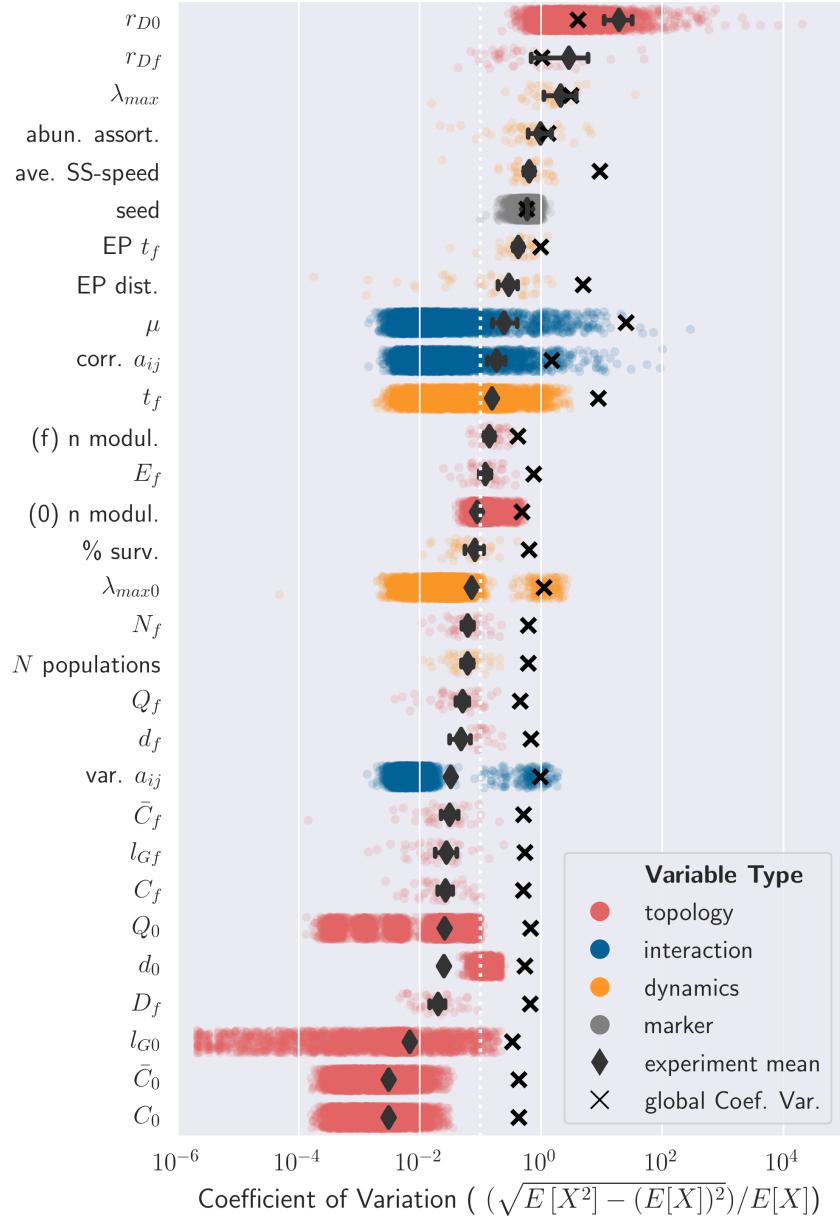


Figure 5.1: Global and Experiment variability: The experiments are composed of a unique set of input parameters. Each experiment has ten replicates, each performed with different seeds (see section 4.1.4). The normalized coefficient of variation (CV) was calculated by dividing the standard deviation by the mean, calculated over the same data. We computed the CV across the ten replicates of each experiment to account for variability in each experiment (colored circles) and across the whole dataset (x marker). For each measure, we calculated the mean of all the CVs determined over the replicates of each experiment. This mean is represented by grey diamond and the whiskers indicate the 95% confidence interval for the mean, obtained through bootstrapping. Note that smaller values indicate less variance around the mean, and larger values are associated with increased variance. In general, the CV calculated within experiments was almost one order of magnitude smaller than the CV calculated for the entire dataset. The colors indicate variable types. The list of variables' names and symbols used in the y axis is presented in table B.1.

5.2.2 Variability in Asymptotic Dynamics

Apart from presenting potentially distinct values for the measured variables, replicates can also display distinct dynamics. To study the variability of attractors in each experiment, we aggregated the detected attractors by experiment. Since almost all simulated communities displayed unbounded dynamics, we focus on the variability of bounded attractors. Figure 5.2 display the distribution of experiments with respect to the proportion of bounded scenarios. For convenience, experiments with no bounded scenarios were suppressed.

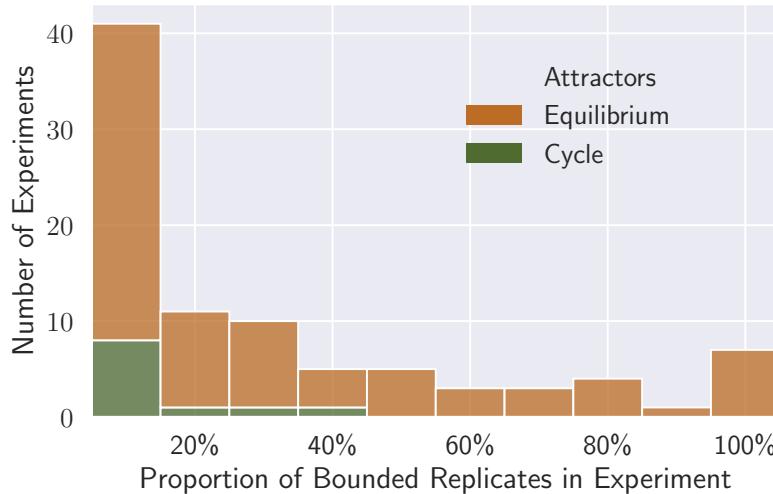


Figure 5.2: **Variability of dynamics type within replicates:** To observe the variability between cycles and equilibrium dynamics, unbounded scenarios were omitted. The proportion of replicates with bounded dynamics was calculated within the ten replicates of each experiment. The number of experiments with a given proportion of bounded dynamics was estimated. The majority of experiments presented bounded dynamics in only 1 out of the 10 replicates. Cycle dynamics were only detected in experiments in which bounded attractors were detected less than 5 out of 10 times. Equilibrium dynamics were present in experiments displaying any proportions of bounded dynamics among replicates. Nevertheless, experiments displaying more bounded than unbounded scenarios were the exception.

The variability in the type of dynamics displayed in an experiment's replicates is directly related to the proportion of bounded attractors. Experiments with 100% of bounded attractors showed no variability, and all replicates were similar. On the other hand, experiments with 10% of bounded attractors showed a bounded attractor in only 1 out of 10 replicates.

In our simulations, most experiments with bounded attractors detected them in just one out of ten replicates. The result is similar for experiments that detected cyclic dynamics or equilibrium points. However, cycles were not detected in experiments with more than 40% of bounded replicates, possibly due to their rarity. Some experiments presented a large proportion of bounded communities, with six experiments detecting equilibria in all replicates. Irrespective of a few experiments with a majority of bounded attractors, unbounded attractors were most common even for experiments that detected bounded replicates.

In the following analysis, each replicate is considered on its own, and experiments are not aggregated. With this procedure, the variability in measures and the detection of attractors is preserved.

5.3 Distribution of Asymptotic Dynamics

The asymptotic dynamics displayed by random putative communities represents a crucial aspect of community dynamics. As indicated in section 3.1.4, an ecological community whose dynamics can be correctly described by the gLV model can display many different asymptotic behaviors. From a

state space standpoint, these asymptotic dynamics represent an invariant set of points (see section 2.1.3).

In the following, we present the distribution of attractors found in the computational simulations. Note that the distribution displayed was obtained after all replicates with transient attractors were reanalyzed or reseeded (see section 4.2.1).

5.3.1 Overall

As mentioned in section 5.1.1, most considered replicates led to unbounded dynamics. The unbounded dynamics is characterized by the presence of one or more populations that grow exponentially in density and crosses the barrier of a million times its carrying capacity ($X_j > 10^6 K_j$ for some j). In the performed experiments, this dynamics was found in 99.4 % of simulated communities.

The number of occurrences of each attractor is presented in figure 5.3. Apart from the evident dominance of unbounded dynamics, it is also possible to see that most bounded communities achieved an equilibrium point. Equilibria were detected in 269 replicates in total, which represent 0.54% of the total amount of simulations and 93.4% of simulations with bounded attractors. Cycles were detected in 19 communities, which represented 6.6% of bounded communities and 0.038% of all simulated communities.

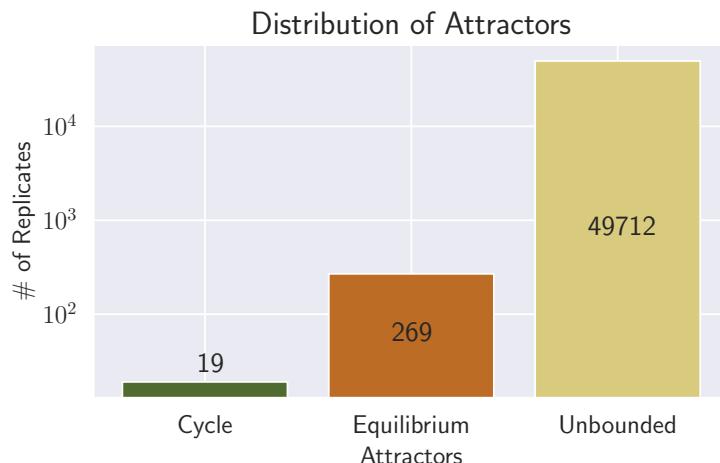


Figure 5.3: **Occurrences of each attractor in the computational experiments:** Most replicates considered led to unbounded dynamics. The equilibrium attractor is more observed than cycles in bounded dynamics. The logarithmic scale on the y-axis highlights the very uneven distribution of types of attractors.

5.3.2 Distribution of Attractors Across Density and Intensity of Interactions

The type of asymptotic dynamics associated with each scenario depended on input parameters. Figure 5.4 indicates how the proportion of asymptotic dynamics types varied with the density.

Note that in figure 5.4, the y-axis is log-transformed to allow for the visualization of the proportions of all three types of dynamics. Therefore, it does not follow that the area or height of different colors is proportional to the number of scenarios associated with a given attractor. As an example, for $D_{in} = 0.1$, the height of the color associated with cyclic dynamics is roughly equal to the color representing equilibria. However, equilibria are ten times more likely than cyclic dynamics in this region.

The proportion of equilibria dynamics varied with the input density. Equilibria represented around 1% of scenarios for small values of density and around 0.1% of scenarios for intermediate density levels. Overall, bounded attractors represented at most 1% of the scenarios for any input

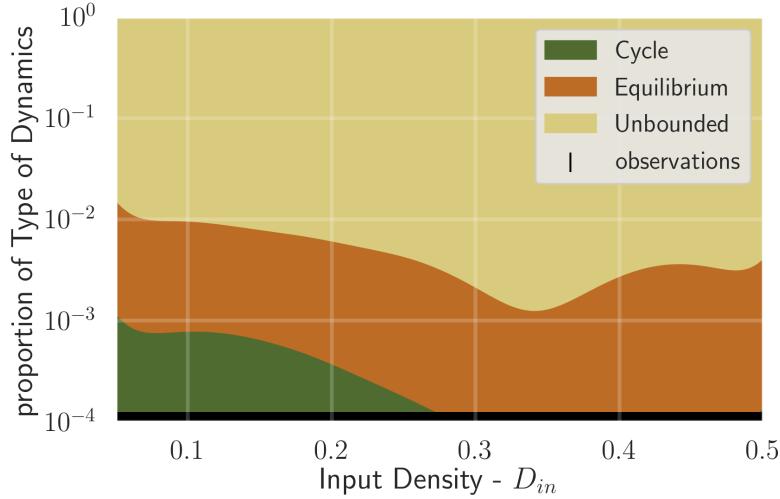


Figure 5.4: Distribution of Types of Dynamics across differences in Network Density: The proportion of scenarios detecting cycles, equilibria or unbounded dynamics is depicted in green, orange and beige, respectively. The input density of interactions of the network is measured as the ratio of the number of pairwise interactions in the network divided by the maximum number of pairwise interactions that could occur with the same number of populations. Each simulation (observation) is depicted as an individual tick in the x-axis and the large amount of simulations hinder identification of individual ticks. For all densities, unbounded scenarios are majorly observed, although equilibrium dynamics are present for all densities. Cycle dynamics are only observed for densities lower than 0.3. Note that the proportion of dynamics type is log-transformed.

density value. Cyclic attractors were only detected for small values of density and did not represent more than 1 in every 1000 scenarios.

Considering other measurements, figure 5.5 indicates how the proportion of asymptotic dynamics types varied with the interaction intensity. Interaction intensity had a very pronounced effect upon the distribution of attractors. All bounded scenarios occurred for relatively small values of the standard deviation of the intensity distribution. Indeed, only one community presented equilibrium for $\sigma_{in} > 1$. Conversely, bounded attractors occurred in up to 10% of the replicates for communities with weaker interactions ($\sigma_{in} \approx 0.2$). Cyclic attractors were also concentrated for small values of σ_{in} , representing more than 0.1% of replicates for some range of the intensity parameter.

To consider how bounded attractors are distributed across both density and intensity of interactions, we plotted a heatmap (figure 5.6). The effect of the intensity of interactions was more pronounced than the effect of the density of interactions as it was possible to observe bounded interactions only for small intensities standard deviation. A combination of small intensity and density was optimal for bounded attractors, since ecological communities presented bounded dynamics in more than one out of every five replicates in this situation.

5.3.3 Dependence of Attractors on Interaction Types

The distribution of attractors was also related to the distribution of interaction types. In figure 5.7, the proportion of bounded attractors for different combinations of interaction density and types is shown. Overall, bounded attractors were more frequent in regions with lower proportions of cooperative interactions (figure 5.7-A). Accordingly, only ten replicates with more than 25% of interactions being cooperative displayed bounded attractors.

For small proportions of cooperative interactions, the distribution of attractors does not clearly depend on the percentage of exploitative or cooperative interactions. Nevertheless, when considering communities with intermediate or large density values, bounded attractors were only registered for simulations with more than 70% of interactions being competitive (figures 5.7-C and 5.7-D). Bounded attractors were also much more common in more competitive communities when only

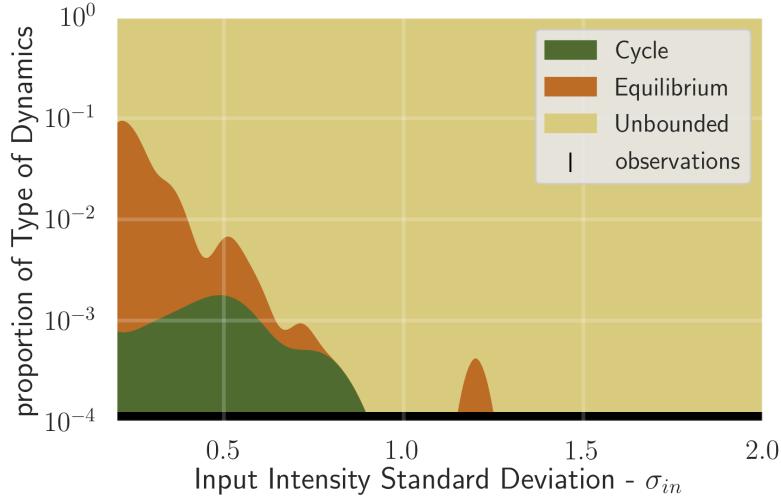


Figure 5.5: Distribution of Types of Dynamics across differences in Interaction Intensity: The proportion of scenarios detecting cycles, equilibria or unbounded dynamics is depicted in green, orange and beige, respectively. The standard deviation indicates the distribution of interaction intensity; larger standard deviation values indicate a higher proportion of strong interactions. Each simulation (observation) is depicted as an individual tick in the x-axis and the large amount of simulations hinder identification of individual ticks. The interaction intensity had a pronounced effect upon attractors' distribution, where bounded dynamics occurred only for relatively small values of standard deviation.

communities with a weak standard deviation of interactions are taken into account (figure B.2-B).

5.3.4 Interaction Topology and Asymptotic Dynamics

The proportion of attractors as a function of input density for each considered topology is shown in figure 5.8. The distribution of bounded and unbounded attractors was roughly similar for all considered topologies. The maximum proportion bounded attractors of roughly 1% of replicates was attained for small densities ($D_i n \approx 0.05$) for all topologies considered. For networks following the Watts-Strogatz topology, the minimum proportion of bounded attractors decreased with increasing probability of reconnection p . Indeed, the topologies constructed with $p = 1 \times 10^{-3}$ consistently had a larger proportion of bounded scenarios than the ones constructed with $p = 1 \times 10^{-1}$ or $p = 1 \times 10^{-2}$ for the same input density.

The number of cyclic attractors detected was negatively related to the reconnection probability in Watts-Strogatz (WS) topology. Most cyclic attractors were detected for the Watts-Strogatz topology with reconnection probability $p = 1 \times 10^{-3}$, WS0.001, which registered nine cycles. The networks constructed with reconnection probabilities equal to 1×10^{-2} (WS0.01) and 1×10^{-1} (WS0.1) were associated with four and one cyclic scenarios, respectively. The networks constructed with random (ER) topology registered three cyclic scenarios, while networks constructed with random-regular (RG) topology registered two.

The number of detected equilibrium points also differed among topologies. Once again, the topology WS0.001 was associated with the largest number of detected equilibria (68). The topology WS0.1 registered 57 replicates with equilibrium points. Both the topologies WS0.01 and ER registered 51 equilibrium points. The random-regular topology was associated with the smallest number of equilibria in total (42).

Since some measures are only defined for bounded attractors, some analysis do not cover the entire parameter space evenly. There is an under-representation of scenarios with more cooperation,

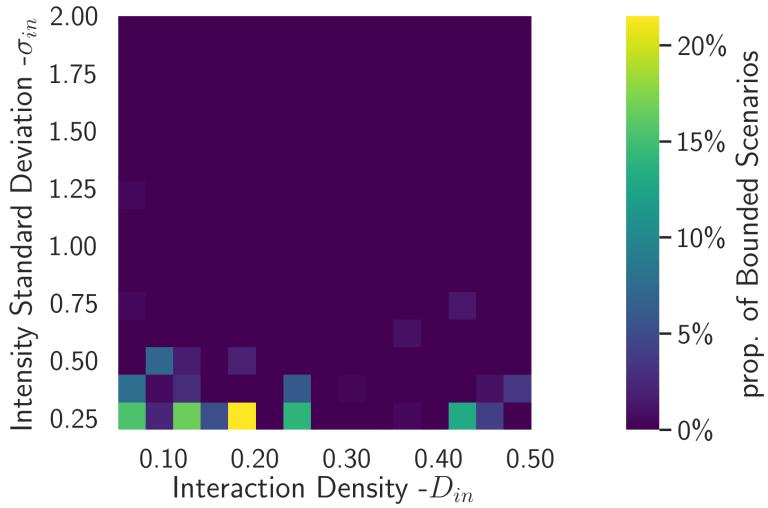


Figure 5.6: Relationship between Interaction Intensity and Density on the proportion of bounded scenarios: The proportion of scenarios displaying bounded dynamics is presented in a heatmap. Brighter colors indicate a larger proportion of bounded scenarios. Scenarios displaying equilibria or cyclic dynamics are classified as bounded. The input density of interactions of the network is measured as the ratio of the number of pairwise interactions in the network divided by the maximum number of pairwise interactions that could occur with the same number of populations. The standard deviation indicates the distribution of interaction intensity; larger standard deviation values indicate a higher proportion of strong interactions. The intensity of interactions was more limiting to the appearance of bounded scenarios than the density of interactions. A combination of small interaction intensity and density was optimal for bounded attractors.

larger densities, or more intense interactions' coefficients among scenarios with bounded dynamics.

5.4 Effects of Structure on Coexistence of Populations

5.4.1 Relation between Coexistence and Interaction Types

The proportion of surviving populations exhibited a strong dependence on the types of interaction in the community. The proportion of competition was strongly negatively correlated with the proportion of surviving populations (figure B.1). Figure 5.9 displays a scatter plot of the proportion of survival (y-axis) and proportion of competition (x-axis). A linear regression between proportion of survival and proportion of competitive interactions indicated 68% of expected decrease in coexistence when the community shifted from 0 to 100% of competitive interactions (figure 5.9-A, black dashed line; Angular coefficient $a = -0.68$, 95% CI $[-0.73, -0.63]$). The line intercept indicated an expected survival of 70% strictly non-competitive communities (intercept $b = 0.70$, 95% CI $[0.67, 0.73]$). Interestingly, the proportion of survival for entirely competitive communities was close to zero for the simulation data. The linear regression between proportion of survival and competition was a good fit of the data ($R^2 = 0.72$; $p \approx 3 \times 10^{-80}$; AIC for linear regression: -438 ; AIC for just the intercept: -77).

The decrease in the proportion of survival with increasing competition was similar for all topologies (figure 5.9-A, colored lines). However, the linear regression for the subset of communities with cyclic attractors was associated with a shallower slope (figure 5.9-B, brown line; angular coefficient $a = -0.68$, 95% CI $[-0.73, -0.63]$).

Apart from the negative relationship with the proportion of competition, the proportion of survival positively depended on the proportion of exploitation (figure B.3 and model parameters in supplementary table B.2). On the other hand, the effect of cooperation upon the proportion of survival was unclear (figure B.4 and model parameters in supplementary table B.2). It is possible to represent the relationship between coexistence and type of interaction without focusing

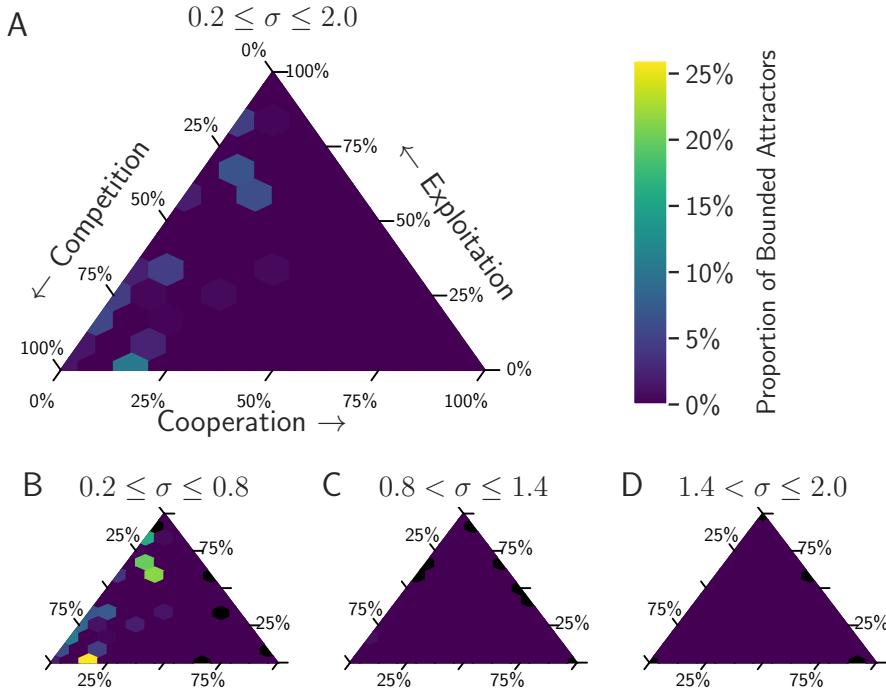


Figure 5.7: The proportion of scenarios displaying bounded dynamics distributed across interaction type and intensity: The proportion of scenarios displaying bounded dynamics is presented in a ternary heatmap. Brighter colors indicate a larger proportion of bounded scenarios. This proportion was measured by dividing the number of scenarios displaying bounded dynamics (equilibrium points and cycles) by the total number of scenarios inside each hexagonal grid cell. Regions in which no scenario was simulated are colored black. The input density of interactions of the network is measured as the ratio of the number of pairwise interactions in the network divided by the maximum number of pairwise interactions that could occur with the same number of populations. The vast regions colored purple are related to the fact that 98.97 % of the simulated scenarios were unbounded. (A) Overall, most of the bounded scenarios were found for a small portion of cooperation and varying amounts of asymmetric and competitive interactions, with a larger proportion of competitive interactions related to more bounded scenarios. (B) The largest proportion of bounded dynamics was found for scenarios with few cooperative interactions ($< 25\%$) and less intense coefficients ($0.2 \leq \sigma \leq 0.8$). (C) and (D) For moderate and intense interaction's coefficients, only one bounded scenario was found.

on a particular type of interaction through a heatmap overlaid on a ternary plot of the types of interaction (see supplementary figure B.5).

5.4.2 Relation between Coexistence and Density of Interactions

In our simulations, there is a negative correlation between the proportion of survival and input communities densities. This dependence is depicted in figure 5.10, which shows a typical decrease of roughly 60% in the proportion of surviving populations as density increases from 5% to 50%. Similarly to the dependence of the proportion of survival with input competition, there was no clear difference in the signal between topologies. The communities with cyclic and equilibrium attractors also showed a similar decrease in the proportion of survival with increasing density. All information on linear regressions constructed for the data can be found in the supplementary table B.2.

Even though all scenarios were generated with uncorrelated input parameters, the data displayed a strong correlation between input density and proportion of competition when considering only communities with bounded dynamics (see supplementary figure B.6). The coexistence did not depend strongly on the intensity of interactions (see supplementary figure B.7 and table B.2).

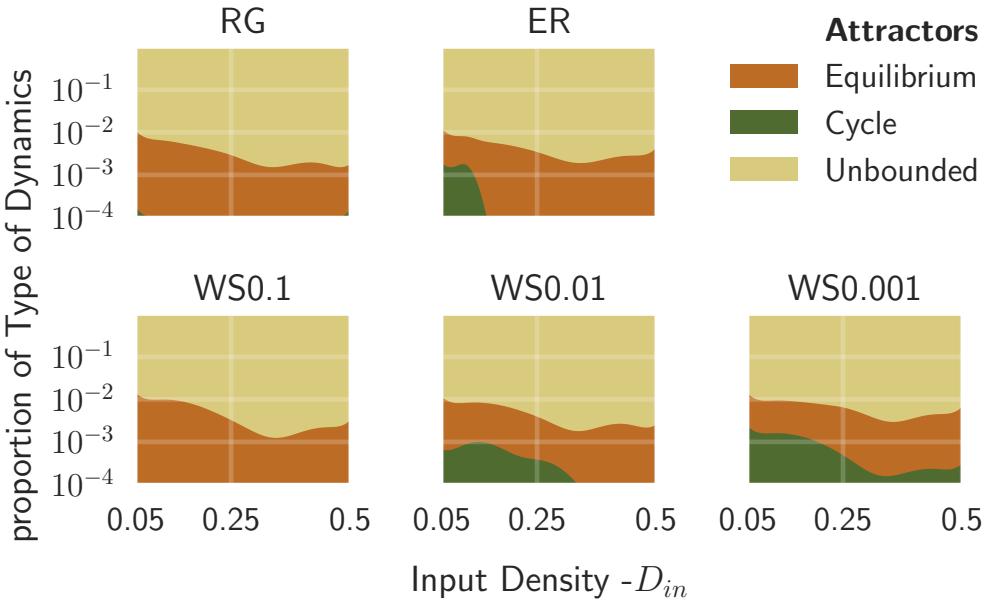


Figure 5.8: The effect of Density on Distribution of Attractors accross different Network Topologies The density indicates how heavily taxa are connected in a given network. The number of cyclic attractors detected was negatively related to the reconnection probability (p) in Watts-Strogatz (WS) topology: when $p = 0.1$, only one cycle dynamics is observed. Similarly, the random-regular (RG) topology registered only two cycle dynamics. The number of equilibrium points detected also differed among topologies. For $p = 0.001$, the topology WS was associated with the largest number of equilibria dynamics detected (68).

5.4.3 Relation between Coexistence and Topologies

In accordance with the result that regressions between the proportion of survival and interaction parameters, the distributions of the proportion of survival were similar across topologies 5.11. The topologies resembled each other in the violin plot overlaid with a swarm plot identifying individual values.

5.5 Resilience of Communities

5.5.1 Resilience and Density of Interactions

In our simulations, the resilience tended to decrease with the increasing density of interactions. The rightmost eigenvalue increased with input density across all topologies, as depicted in figure 5.12-A. This trend was also consistent between unbounded and equilibrium types of attractors, although regressions performed for these subsets of data presented different values for coefficients (figure 5.12-B, parameters reported in the supplementary table B.3). Linear regression for the whole data indicates that if input density increases by 0.1, the rightmost eigenvalue increases by 7, on average. If the same regression is performed by considering only communities with equilibrium points, the rightmost eigenvalue is expected to increase by 20. On the other hand, data on cycles do not support a strong influence of density upon the rightmost eigenvalue.

The data presented a large spread as highlighted by the small coefficients of determination, with $r^2 \approx 0.12$ for all regressions except the one based on the subset of data exhibiting equilibrium point attractors ($r^2 = 0.27$) and cycles ($r^2 \approx 0$). Remarkably, the rightmost eigenvalue distribution presented an upper bound, which seemed to depend linearly on the input density. The parameters of all linear regressions addressed are available in supplementary table B.3.

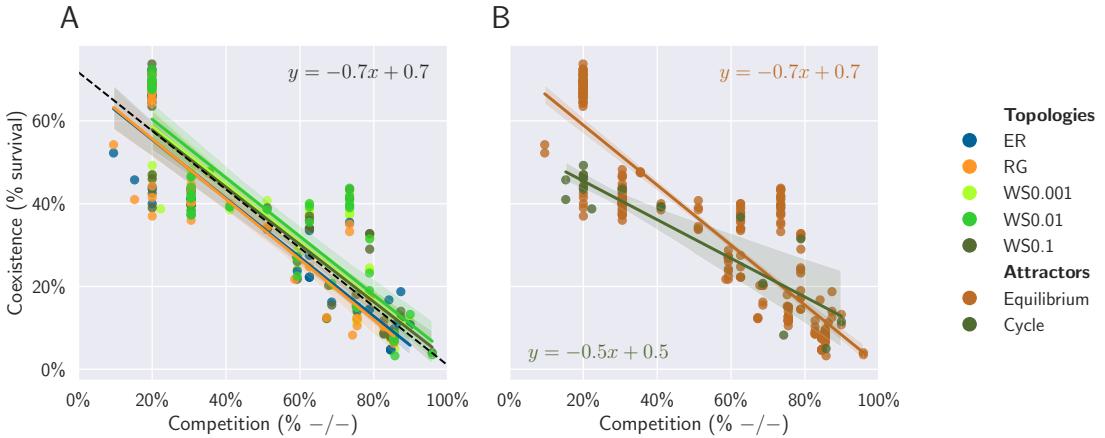


Figure 5.9: More competition results in a lower percentage of surviving populations: The percentage of survivors is calculated as the number of nodes in the final network divided by the number of nodes in the initial network. The percentage of survivors in the final network is affected by the proportion of competition in the initial network, where more competition results in a lower percentage of surviving populations. (A) A comparable result was found for all topologies, as it is possible to see the similarity between colored full lines and black dashed line. The black line intercept and angular coefficient were calculated as the mean of these parameters for the colored lines. (B) Equilibria and cyclic dynamics were associated with distinct trends in reducing the percentage of survivors with increasing competition, as indicated by the angular coefficient. However, scenarios with cycles had a larger decrease in the percentage of survivors for the no competition best-case scenario as revealed by the lower intercept.

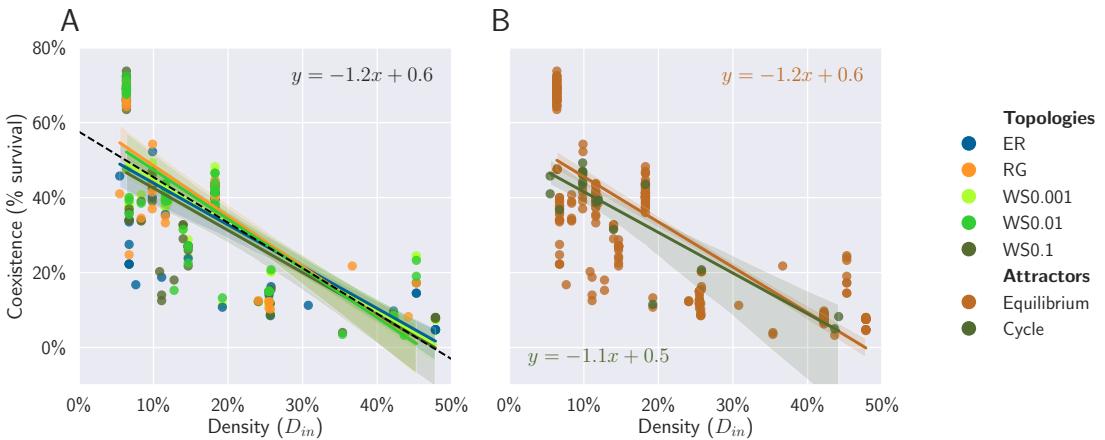


Figure 5.10: Highly connected networks result in a lower percentage of surviving populations: The percentage of survivors is estimated by the number of nodes in the final network divided by the number of nodes in the initial network. The percentage of survivors in the final network is affected by interaction density in the initial network, where higher densities result in a lower percentage of surviving populations. The density indicates how heavily taxa are connected in a given network. (A) A comparable result was found for all topologies, as it is possible to see the similarity between colored full lines and black dashed line. The black line intercept and angular coefficient were calculated as the mean of these parameters for the colored lines. (B) Equilibria and cyclic dynamics were associated with slightly distinct trends in reducing the percentage of surviving populations. Equilibria scenarios had a slightly steeper decrease in the percentage of survivors with increasing competition, as indicated by the angular coefficient.

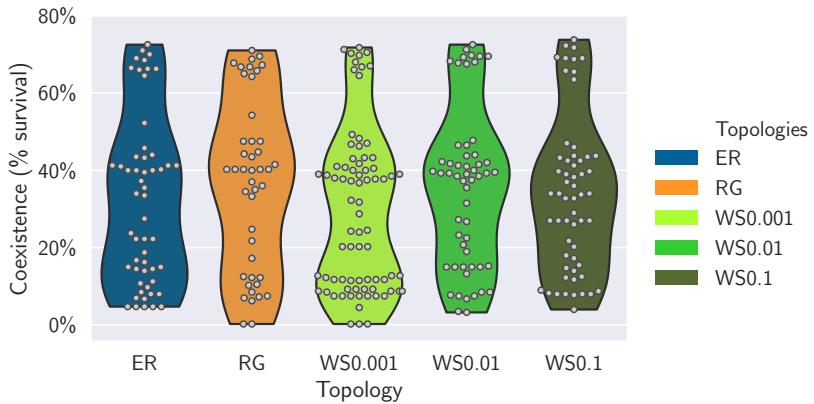


Figure 5.11: Proportion of surviving populations across topologies The percentage of survivors is estimated by the number of nodes in the final network divided by the number of nodes in the initial network. The distributions of the proportion of survival were similar across topologies. The 'violin' shape indicates the distribution of values.

5.5.2 Resilience and Intensity of Interactions

In a similar fashion than the density of interactions, the rightmost eigenvalue depended positively on the standard deviation of interactions intensities (figure 5.13). Despite this overall indication that the resilience diminished with increasing interactions' standard deviation, the data displayed a large variance ($r^2 = 0.11$). The linear regression based on communities with equilibria attractors showed a negligible coefficient of determination ($r^2 \approx 0$, $p = 0.72$). Conversely, the intensity of interaction explained much of the variance on the rightmost eigenvalue of cycle-related data ($r^2 = 0.79$, $p = 0.04$). The different network topologies showed similar results for the resilience vs. intensity trend.

5.5.3 Resilience and Types of Interaction

Resilience showed a strong dependence on the distribution of types of interaction (figure 5.14). Communities were more resilient for larger proportions of competitive and exploitative interactions and less resilient for increasing proportions of cooperative interactions. Indeed, linear regressions between the rightmost eigenvalue and proportion of competition or exploitation had consistently negative angular coefficients overall and for each topology considered ($-27.4 \leq a \leq -24.1$). Linear regressions also indicated the positive correlation between the proportion of cooperation and rightmost eigenvalue (angular coefficients $52.3 \leq a \leq 54.9$). This tendency was also consistent over different network density levels (figures 5.14-B, 5.14-C, and 5.14-D), with the least resilient communities being found for more cooperative communities.

The combined effects of the standard deviation of intensities and types of interaction were similar to the combined effect of types and density of interactions. Furthermore, the effects of density, intensity and types of interactions were cumulative. Thus, the least stable communities were dense cooperative communities with a higher proportion of intense interactions (figures 5.14 and B.9).

For the whole data, competition and exploitation did not explain much of the resilience variance ($r^2 = 0.05$ for both regressions). On the other hand, the effect of cooperative interactions was much more pronounced, explaining up to 21% of the overall variance of the rightmost eigenvalue ($r^2 = 0.21$).

Interestingly when only bounded attractors were considered, the competition and exploitation effects upon resilience were distinct. For bounded attractors, competition led to decreased resilience ($a = 76$, $p = 1.2 \times 10^{-9}$, $r^2 = 0.15$) while exploitation had the opposite effect, boosting resilience by decreasing the expected rightmost eigenvalue ($a = -75$, $p = 5.1 \times 10^{-9}$, $r^2 = 0.14$). Also, cooperation did not appear to have substantial effects on resilience for this subset of the data ($r^2 = 0.1$, $p = 0.13$).

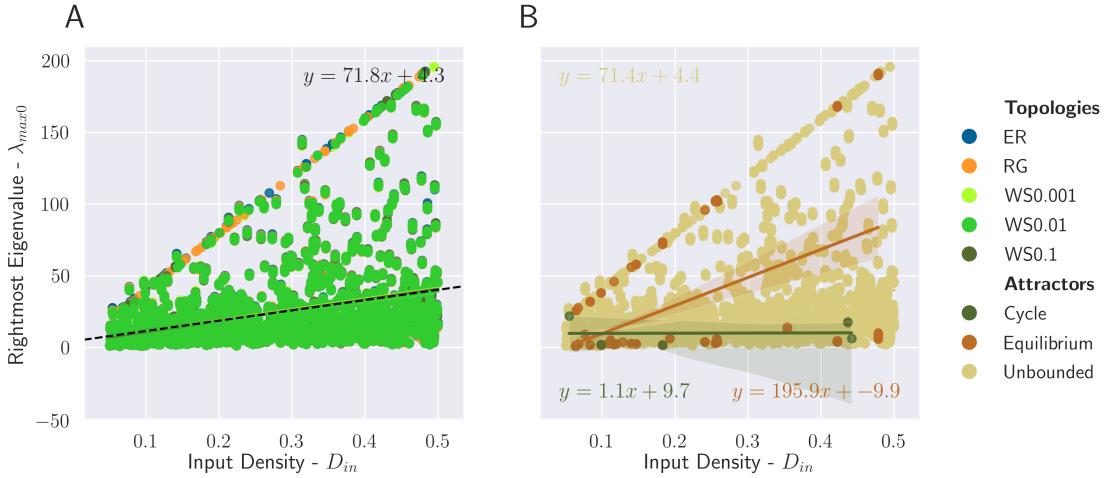


Figure 5.12: Effect of network Density on network Resilience: The density indicates how heavily taxa are connected in a given network. The rightmost eigenvalue corresponds to a measure of resilience in the initial network; as the rightmost eigenvalue gets closer to the origin, the network becomes more resilient. (A) Considering the different network topologies, as density increases, the resilience decreases, and appears to be no difference in this tendency across network topologies. (B) Considering different dynamics types, the relationship between network density and network resilience appear to differ between attractor types. For equilibrium dynamics, as network density increases, resilience decreases, whereas there is an overall small impact of density in resilience for the cycle dynamics.

5.5.4 Resilience and Attractors / Topologies

The distribution of the rightmost eigenvalue of networks was similar across topologies, as it can be seen in figure 5.15. Accordingly, summary values for these distributions were comparable regardless of the particular distribution used to construct the networks, as seen in table 5.8. Overall, the rightmost eigenvalue was distributed between 0.9 and 196, but heavily concentrated for small values (75% percentile being $\lambda_{max0} = 27.4$).

In contrast, the distributions of the rightmost eigenvalue are distinct when different types of dynamics are considered, as seen in figure B.8. There were more small eigenvalues in communities that displayed equilibrium points in contrast to unbounded dynamics (see table 5.8). There were also differences in the standard deviations, driven by each the number of occurrences of each attractor.

		n	mean	std	min	25%	50%	75%	max
1	<i>Overall</i>	5.0×10^4	24.10	27.07	0.91	8.26	14.94	27.44	196.0
2	<i>Topology</i>	RG	9970	24.15	27.39	0.91	8.23	14.86	27.28
3		ER	9959	24.08	27.04	0.93	8.29	14.91	27.32
4		WS0.001	9943	24.49	27.58	0.94	8.35	15.19	27.78
5		WS0.01	9944	23.75	25.99	0.95	8.30	14.99	27.38
6		WS0.1	9959	24.04	27.30	0.94	8.14	14.83	27.30
7	<i>Attractor</i>	Unbounded	5.0×10^4	24.08	26.90	1.03	8.31	14.97	27.44
8		Equilibrium	223	28.94	52.36	0.91	1.63	2.75	26
9		Cycle	5	10.00	9.37	1.70	2.15	6.22	17.90
									22

Table 5.8: Descriptive Statistics of the Distributions of Rightmost Eigenvalue for Different Subsets: We present the number of observations (n), mean, standard deviation (std), minimum and maximum values (min, max) and the 25%, 50% and 75% quartiles that characterize the distribution of the rightmost eigenvalue unfiltered, filtered by topology, and filtered by type of attractor. The 50% quartile is also referred to as the median of the data.

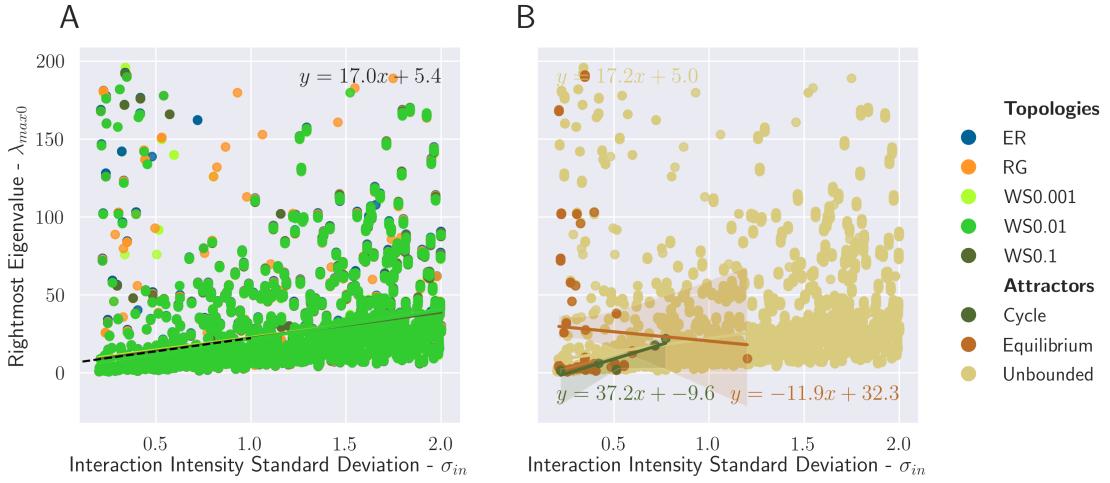


Figure 5.13: Effect of Interaction Intensity on network Resilience: The standard deviation indicates the distribution of interaction intensity; larger standard deviation values indicate a higher proportion of strong interactions. The rightmost eigenvalue corresponds to a measure of resilience in the initial network; as the rightmost eigenvalue gets closer to the origin, the network becomes more resilient. (A) Considering the different network topologies, although there appears to be no clear signal, in general, as the intensity of interaction increases, the resilience decreases. (B) Considering different dynamics types, the relationship between interaction intensity and network resilience appear to differ between attractor types. For equilibrium dynamics, as interaction intensity increases, resilience decreases, which is the opposite observed for the cycle dynamics.

5.5.5 Distribution of Eigenvalues Before and After Stabilization

The distributions of eigenvalues for the initial and final networks were consistently different. The initial eigenvalues were calculated for the matrix $\mathbb{A}' = \mathbb{A} + \mathbb{I}$, which is the interaction matrix but with no intraspecific competition, i.e. $a_{ii} = 0$ $i = 1, \dots, N$. The initial distribution was generally centered at the complex plane's origin, with most eigenvalues being scattered across an ellipsoidal region. After the community's temporal evolution, the eigenvalues were calculated for the Jacobian matrix of the gLV model, equation (5.16). The final networks showed a distinctive distribution of eigenvalues with the eigenvalue with the largest real part being very close to zero.

A significant part of the remaining eigenvalues was distributed in a "tail" with small or null imaginary parts and negative real parts. Some eigenvalues also displayed larger imaginary parts and were distributed with negative real parts with larger moduli associated with the imaginary part of the complex number's larger absolute values. In general, the final eigenvalues typically assumed a spear-point or arrow-point shape oriented from left to right with the tip very close to the complex plane origin. In order to illustrate the distinct typical distributions of eigenvalues, we present an example in figure 5.16. In figure 5.16, the orange ellipsoidal distribution presents the eigenvalues for the interaction matrix without intraspecific competition. Conversely, the blue distribution is constructed for eigenvalues of the Jacobian matrix of the underlying gLV model evaluated at the final densities (\bar{X}^*).

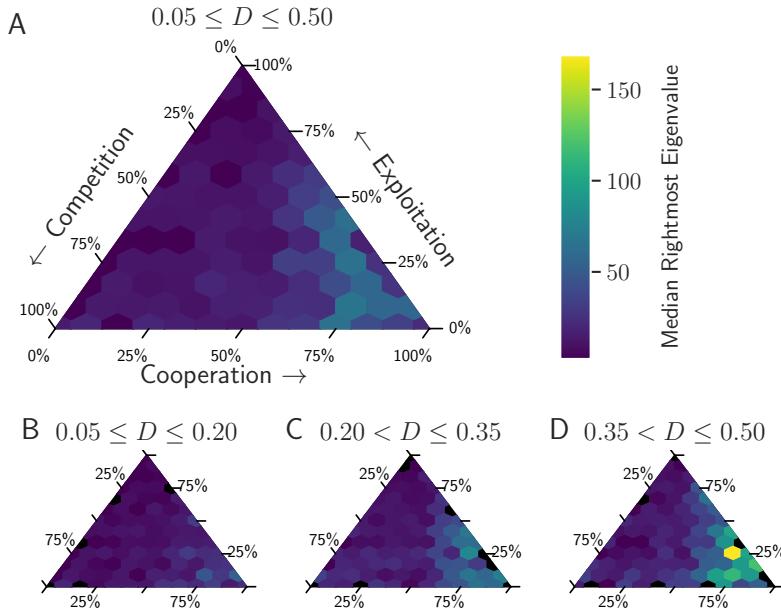


Figure 5.14: The proportion of Interaction Types influences the Resilience across different network Densities The density indicates how heavily taxa are connected in a given network. The rightmost eigenvalue corresponds to a measure of resilience in the initial network; as the rightmost eigenvalue gets closer to the origin the network becomes more resilient. (A) In general, for a large variability of densities, we observe that resilience increase as the proportion of cooperation decreases. (B-D) This pattern is maintained across different density intervals. Note that as density increases, the destabilizing effect of large proportions of cooperation increases, and competition and exploitation decreases (C).

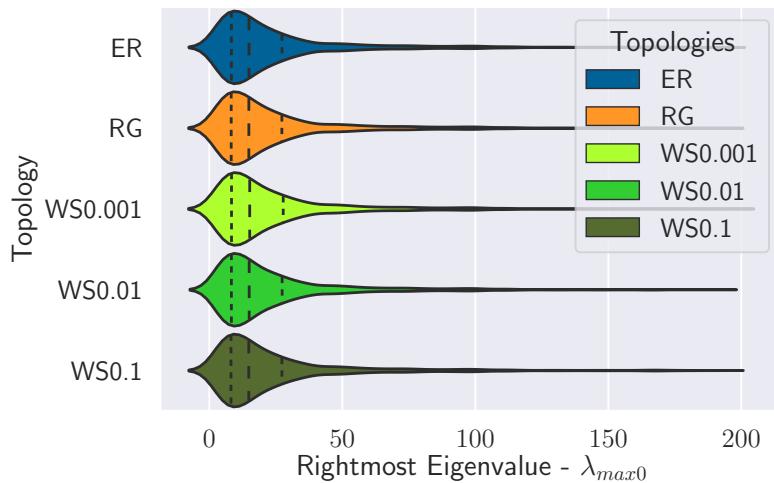


Figure 5.15: Resilience is similar for different network Topologies The rightmost eigenvalue corresponds to a measure of resilience in the initial network; as the rightmost eigenvalue gets closer to the origin the network becomes more resilient. The left dashed line indicates the 25% quartile; the central line indicates the median; the right dashed line indicates the 75% quartile. The resilience pattern is similar for all topologies. The majority of rightmost eigenvalues are smaller than 50.

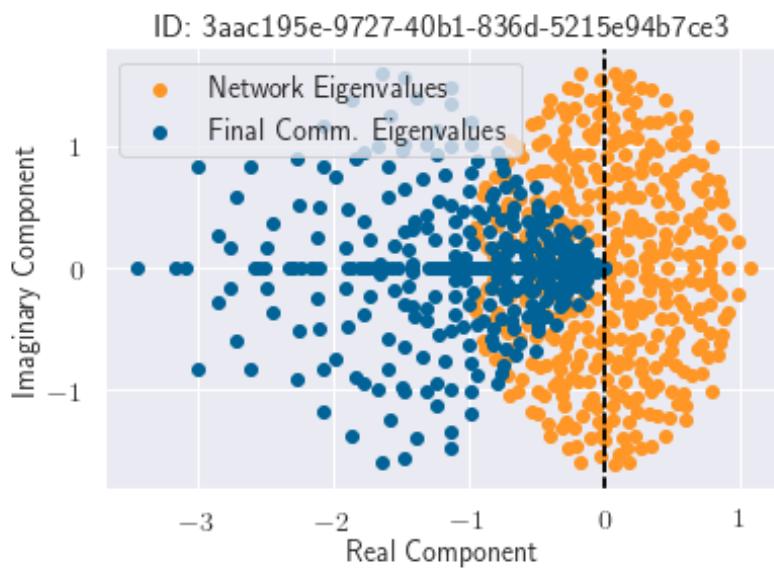


Figure 5.16: Standard case depicting the Resilience variation between Initial and Final Network: Orange depicts the eigenvalues from the $A+I$ matrix, which corresponds to the initial network. We can observe that the real and the imaginary component are centered around zero. Blue depicts the eigenvalues for the Jacobian matrix evaluated around Equilibrium Point, which corresponds to the final network. We can observe that the eigenvalues tend to concentrate on the negative real component, which corresponding to a stabilization of the network after temporal evolution.

Chapter 6

Discussion

Links between different aspects of community structure and stability abound. Nevertheless, it is a longstanding question whether particular trends are restricted to small regions of the parameter space or are ubiquitous "law-like" relations. We approached this problem by exploring the parameter space of the generalized Lotka-Volterra model allowing for simultaneous changes in all input variables. In this sense, the regularities we find are related to the entire region of the parameter space we covered with the ecologically-motivated parametrization.

Indeed, the small number of input parameters we used to construct the ecological communities is at odds with the N^2 parameters needed to completely determine the generalized Lotka-Volterra model. Namely, the parametrization of the gLV model was based on: i) interaction density; ii) The standard deviation of interaction intensities; iii-iv) The proportions of cooperative and competitive interactions; v) The number of populations in the community; vi) The choice of the random graph model among the total of five possibilities. Therefore, it is natural that we expect our parametrization to display some level of variability when measures taken from the input or output communities are concerned. Nevertheless, it was possible to show that our inputs indeed determined much of the studied variables to a high degree.

We measured smaller coefficients of variation when considering all the replicates that shared the same input parameters than when all simulations (across distinct inputs) were considered. In this sense, a large (small) coefficient of variation for some measure indicates that this measure is heterogeneous (homogeneous) across the parameter space (see figure 5.1). That is, if the coefficient of variation measured from one subset of the data presents smaller values than when measured for all datasets, it indicates that the measure is more homogeneous in the subset. Overall, this indicates that the set of all possible $N^2 a_{ij}$ parameters that satisfy the restrictions imposed by the few input parameters presented characteristic values for most of the measured variables. Thus, we confirm that density, intensity (standard deviation), and types of interactions determine the values of many other properties and quantities of interest to a large extent.

Notably, topological variables displayed a minimal coefficient of variation overall, which is a consequence of the fact that input variables directly determine the network's topology. The random network algorithm and the density determine much of the properties of the undirected network. Other variables, such as the maximum rightmost eigenvalue, also display smaller than random variabilities, supporting the linkage between ecological communities in their mathematical model representations.

Overall there was a vast prevalence of unbounded attractors on simulated communities (see figure 5.3). This high prevalence of unbounded attractors indicates that the generalized Lotka-Volterra model's parameter space is dominated by parameter sets that do not correspond to biologically accurate depictions of ecological communities. This lack of correspondence between the generalized Lotka-Volterra and the temporal evolution of ecological communities was suggested before, mainly due to the unlimited benefits of mutualism assumed in the model (Qian and Akçay 2020). A complete analysis of the two-dimensional gLV indicates that when the parameters of the mutual influence are such that $a_{12}a_{21} > 1$, the time evolution of the system leads to unbounded growth of both populations (see section 3.1.3). Possible solutions to this mismatch between model and real system are modifying the functional response of the cooperation inside the model and performing an informed parametrization of the model. In this sense, the algorithm 1 could be used as a tool for parametrization so that regions associated with unbounded attractors are actively

avoided.

Our simulations also support that equilibrium points are much more common than cycles throughout the parameter space (see figure 5.3) since equilibria were associated with a substantially higher number of replicates. Thus, the generalized Lotka-Volterra model predicts that ecological communities should more commonly display coexistence dynamics for which the populations are approximately fixed.

Additionally, bounded attractors were much more common for less intense interaction coefficient values, i.e., the standard deviation of interaction intensities was smaller (see figure 5.5). Presumably, larger proportions of intense interactions are associated with a higher probability that at least two populations with cooperative interaction would overcome each other intraspecific competition, leading to sustained positive growth rates for arbitrarily large densities. However, additional investigation is needed on why unbounded attractors were ubiquitous even for less dense and more competitive communities.

There was no striking difference in distributions of attractors across densities and topologies covering the range of regular to random networks (see figure 5.8). Remarkably, the fact that proportions of attractors did not depend strongly on the densities of communities overall might result from the confounding effects stemming from considering all possible combinations of ecological interactions (see figure 5.4). Indeed, cooperative and competitive scenarios are associated with conflicting expectations for increased density levels. Cooperative scenarios would allow for more unbounded attractors, as more pairs of populations could potentially achieve unlimited growth. Conversely, competitive scenarios would have a decreased proportion of unbounded attractors due to increased competitive regulations in the community. Thus, our results were in the opposite direction of the strong link between higher connectance and equilibrium attractors reported by Fowler (2009).

We provide additional support for the link between cooperation and decreased resilience (Coyte, Schluter, and Foster 2015; Allesina and Tang 2012). Indeed, the strongest predictor of a community's resilience was the proportion of cooperative interactions, with more cooperative communities being much less resilient overall (see figure 5.14). Since we did not measure the community's invasibility, it is not possible to indicate whether data gathered using our methodology would support recent findings linking increased proportions of cooperation and community-level resistance to invasions (Qian and Akçay 2020). Also, it is not possible to confirm that random communities with diverse types of interaction would be locally stable most often (Mougi and Kondoh 2012; Mougi and Kondoh 2014).

In our simulations, increased densities generally led to less stable communities, which is in agreement with known results (e.g. Thébault and Fontaine 2010). The coefficient of determination of the subjacent linear regression was relatively small overall ($R^2 \approx 0.12$) but presented higher levels when only scenarios leading to equilibrium were considered ($R^2 \approx 0.27$; see figure 5.14 and table B.3). The difference in the coefficients of determination either suggests that densities might become more relevant in determining resilience for communities that attain equilibrium or that the requirements for equilibrium impose correlations that lead to increased coefficients of determination. Independently of the specific attractor considered, higher densities were always associated with rightmost eigenvalues with larger real parts.

The intensity of interactions was positively associated with the rightmost eigenvalue, so that communities with more (less) intense interactions had larger (smaller) rightmost eigenvalues and lower (higher) resilience (see figure B.9 and table B.3). Thus, we join other authors in predicting that communities with lower densities and intensities of interactions and more competition/exploitation would be more resilient (Rohr, Saavedra, and Bascompte 2014; Coyte, Schluter, and Foster 2015).

However, the signal linking resilience to other parameters may depend upon the type of dynamics in the community. In another way, the properties of interactions among populations jointly determine the type of dynamics and resilience parameters. Nevertheless, the trends observed for the regions in which equilibrium communities were detected could better represent real communities and thus be more adequate.

Additionally, we did not find differences in the distribution of the rightmost eigenvalue between topologies (see figure 5.15). This result accumulates to a body of literature that found out the network topology is not the most crucial predictor of stability in ecological communities (Allesina et al. 2015; Serván et al. 2018). Therefore, our results allow us to conclude that the type, density, and intensities of interactions are more critical to the distribution of eigenvalues than the topology itself.

In predicting ecological communities' dynamical properties, interaction properties are more relevant than the network's specific topology. The similarity between rightmost eigenvalues distributions for different network topologies further indicates that the variance detected on these networks was not a direct consequence of topological differences between random network models considered.

Even though the input parameters were evenly distributed across the parameter space's predefined range, communities that showed bounded attractors were concentrated for particular values and ranges of the parameters. Indeed, the set of communities with bounded attractors displayed a high correlation of the input variables as it was the case of competition/exploitation and competition/density. These relations indicate trade-offs that must be satisfied in communities that present bounded dynamics. For instance, denser networks should be associated with a higher proportion of competitive interactions to present equilibrium points (see figure B.6).

Our results support the idea that the exclusive competition principle is at play for larger communities, even though its traditional formulation is based on reasoning ecological interactions between populations pairs (see figure 5.9). Indeed we found that communities with larger proportions of competitive interactions were associated with smaller proportions of survival. This competitive exclusion principle was already discussed for larger random communities by other authors, and the evidence in this work provides further confirmation of this trend (Serván et al. 2018).

The proportion of survival also depends on other input variables such as proportion of exploitation and density of interactions (see figures B.3 and B.10, and table B.2). However, it is unclear whether these patterns represent distinct effects or whether they are consequences of the correlations between the proportion of competition and other parameters. To elucidate the effects of each input parameter upon coexistence, a methodology preventing such correlations will be necessary. Once again, the particular topology of interactions was not determinant of coexistence levels in the community.

Chapter 7

Conclusion

In this section, summarize the findings and the developments of this work. The perspectives and open questions are presented at the same time that limitations are discussed. In the end, we provide some concluding remarks on what we think is the central message of this work.

7.1 Caveats and Perspectives

Perhaps most importantly, a relatively large proportion of simulated communities detected presented unbounded dynamics (see section 5.3). The proportion of unbounded dynamics did not interfere with the variability of parameters and measures or the links between input parameters and resilience. However, since coexistence could only be determined for communities with bounded dynamics, correlations among the input variables hinder conclusions concerning coexistence. Further studies should address this difficulty by implementing statistical parameterizations that avoid such correlations of input parameters on bounded communities. On a related issue, it is relevant to study our conclusions in the face of more complex functional responses for the benefit of cooperative interactions among populations.

The generation of the initial communities followed canonical models of random graphs in network science (see section 2.2.3). However, these random graphs do not incorporate many structural details of real ecological communities. Therefore, it would be very beneficial to consider other algorithms that generate networks more akin to real ecological communities as measured in experimental and laboratory studies. Ultimately, it is essential to consider real network topologies and verify whether modeling predictions are closely related to natural systems.

More realistic networks could be constructed by lifting the hypothesis that quantity, intensity, and the type of interactions are independent of node identity. For instance, topology could be modified by considering all the populations have limited available time for interactions so that they should prioritize which populations they interact with, thus imposing restrictions on the amount and intensities of interactions for each population. Additionally, the usage of algorithms generating random graphs with other topological features such as assortativity or power-law degree distribution would provide further insights on whether such features impact resilience and coexistence.

By incorporating new measures to characterize topology and interaction properties, it would be possible to compare predictions for different stability types. Indeed, the measures of stability could be expanded to include other aspects beyond resilience and coexistence. This methodology could provide further insights into the conflicting results reported in the literature by investigating different stability concepts.

As the number of theoretical predictions of the rightmost eigenvalue and eigenvector distributions in ecological communities abounds, it would be possible to use the methods employed in this work to verify the accuracy of those predictions when considering large regions of the parameter space in a comparative manner. Additionally, lifting the assumption of an equal intrinsic growth rate for all populations is a significant step for future investigations. We expect the picture presented here to be improved by including the possibility of different and potentially negative intrinsic growth rates. Indeed, by considering specific growth rates for each population, it will be possible to study the feasibility regions for each parameter set.

In future developments, computational program enhancements should focus on implementing algorithms for the detection of chaotic dynamics. Additionally, the detection of (stable) equilibrium points can be improved by checking if the eigenvalues at the equilibrium point are all negative. This update will also allow for the detection of homoclinic and heteroclinic cycles, as the numerical integration would no longer terminate in the vicinity of unstable equilibrium points.

We have focused on the properties of the ecological systems at the community level. Additional studies can also adapt this methodology to focus on particular populations' properties, answering questions such as what features allow a particular population to thrive or not in a complex community.

Finally, our work is prone to the standard limitations of mathematical models in ecology. Some assumptions are implicitly made due to the ordinary differential equations formulation. For instance, we assume population density is continuous and that it changes smoothly in time. We assume ecological interactions directly impact the growth rate of each population and that this impact is proportional to the product of the densities of the interacting populations. The effects of the interactions with many populations upon the growth rate of each population are additive. We also assume that the coefficients of interaction do not change substantially in the time-scale of the temporal evolution of the system. Indeed, such limitations are part of the generalized Lotka-Volterra modeling framework and reinforce the importance of avoiding hasty generalizations and inferences. Nevertheless, the usage of "toy models" in general, and the gLV in particular, have an invaluable role in constructing solid ecological theories, as seen in the last century.

7.2 Take-home message

In summary, we have investigated the generalized Lotka-Volterra model through an ecologically informed parametrization that allowed us to verify global patterns linking structure and dynamics in ecological communities. We recovered stable speciose communities that displayed coexistence in equilibrium points and cycles, reinforcing the suitability of the generalized Lotka-Volterra model to grasp the essential properties of such communities. However, we conclude that a large portion of the parameter space is dominated by ecologically inconsistent dynamics, indicating care should be taken when linking model and ecological communities. Furthermore, we have reproduced many of the previous results from mathematical and computational studies linking communities' topological and interaction properties to resilience and coexistence of populations. We conclude by pointing out that it is essential to develop theories and computational models that entail large proportions of the parameter space.

Appendix A

Detection of Attractors' Accuracy

The correct detection of the attractors is crucial to our methodology. In this appendix, we present tests on the accuracy of the detection of the attractors. The first set of tests, presented in section [A.1](#), was based on generalized Lotka-Volterra systems with parametrizations whose attractors are known. The parametrization source is indicated for each attractor and the attractor detected by the program developed for this research.

Overall, the detected and reported attractors are equal. The only exception is a heteroclinic cycle, which was detected as an equilibrium. This error is a consequence of the fact that if the system evolution presents very slow dynamics, it is detected as an equilibrium point. In the referred case, the system's evolution goes very close to an equilibrium point, therefore being detected as one. Additionally, parametrizations leading to cycles with no determined period are not detectable through our algorithm since it assumes a constant period.

The system being integrated is the generalized Lotka-Volterra as presented in equation [\(3.27\)](#), which we present below. We therefore present only the parametrizations and the initial values of the populations' densities. Most tests were performed in three dimensional systems, as these were more abundant in the consulted literature.

$$\frac{d\vec{X}}{dt} = \text{diag}(\vec{X} \circ \vec{r}) (\vec{1} + \mathbf{A}\vec{X}) \quad (\text{A.1})$$

The second set of tests (section [A.2](#)) was performed by simulating 50 communities for which we show the initial and final networks alongside the temporal evolution of the populations.

A.1 Systems from the Literature

In the following, we present the systems used to test attractors' identification, reference, and algorithm's outcome.

A.1.1 Equilibrium Points

A.1.1.1 Takeuchi ([1996](#), p.70)

$$\mathbf{A} = \begin{pmatrix} -1 & -1 & -4 \\ -1.5 & -1 & -1 \\ 2 & 0.5 & 0 \end{pmatrix} \quad (\text{A.2a})$$

$$\vec{r} = \vec{1} \quad (\text{A.2b})$$

$$\vec{X}(t=0) = \text{sampled randomly from uniform distribution } [0, 1] \quad (\text{A.2c})$$

The equilibrium point was correctly identified.

A.1.1.2 Allesina (2020)

$$\mathbf{A} = - \begin{pmatrix} 10 & 7 & 12 \\ 15 & 10 & 8 \\ 7 & 11 & 10 \end{pmatrix} \quad (\text{A.3a})$$

$$\vec{r} = \vec{1} \quad (\text{A.3b})$$

$$\vec{X}(t=0) = \text{sampled randomly from uniform distribution } [0, 1] \quad (\text{A.3c})$$

The equilibrium point was correctly identified.

A.1.1.3 Many Isolated Populations

$$\mathbf{A} = -\text{diag}(a_i) \quad (\text{A.4a})$$

$$\vec{r} = \vec{1} \quad (\text{A.4b})$$

$$\vec{X}(t=0) = \text{sampled randomly from uniform distribution } [0, 1] \quad (\text{A.4c})$$

The equilibrium point was correctly identified for all tested values of a_i . We tested up to 500 populations, randomly sampling the parameters a_i from a uniform distribution between 0 and 1. The rationale behind this example is that when populations are not interacting, each one follows a logistic growth, independently from the others.

A.1.2 Cycles

A.1.2.1 Allesina (2020)

$$\mathbf{A} = - \begin{pmatrix} 10 & 6 & 12 \\ 14 & 10 & 2 \\ 8 & 18 & 10 \end{pmatrix} \quad (\text{A.5a})$$

$$\vec{r} = \vec{1} \quad (\text{A.5b})$$

$$\vec{X}(t=0) = (0.1 \ 0.2 \ 0.6)^T \quad (\text{A.5c})$$

The cycle was correctly identified.

A.1.2.2 Takeuchi (1996, p.71)

$$\mathbf{A} = \begin{pmatrix} -1 & -1 & -7 \\ -1.5 & -1 & -1 \\ 3.5 & 0.5 & 0 \end{pmatrix} \quad (\text{A.6a})$$

$$\vec{r} = \vec{1} \quad (\text{A.6b})$$

$$\vec{X}(t=0) = (0.1 \ 0.2 \ 0.6)^T \quad (\text{A.6c})$$

The cycle was correctly identified.

A.1.2.3 Takeuchi (1996, p.76)

$$\mathbf{A} = - \begin{pmatrix} 1 & 1 & 0 & 6 \\ 1.5 & 1 & 1.5 & 1 \\ 0 & 1 & 1 & 6 \\ -3 & -0.5 & -3 & 0 \end{pmatrix} \quad (\text{A.7a})$$

$$\vec{r} = (1 \ 1 \ 1 \ -1)^T \quad (\text{A.7b})$$

$$\vec{X}(t=0) = (0.1 \ 0.7 \ 0.1 \ 0.08)^T \quad (\text{A.7c})$$

The cycle was correctly identified.

A.1.2.4 Takeuchi (1996, p.82)

$$\mathbf{A} = - \begin{pmatrix} 1 & 1.5 & 4 & 2 \\ 0.6 & 1 & 2 & 4 \\ -4 & -2 & 0 & 0 \\ -2 & -4 & 0 & 0 \end{pmatrix} \quad (\text{A.8a})$$

$$\vec{r} = (1 \ 1 \ -1 \ -1)^T \quad (\text{A.8b})$$

$$\vec{X}(t=0) = (0.2 \ 0.18 \ 0.05 \ 0.2)^T \quad (\text{A.8c})$$

The cycle was correctly identified.

A.1.2.5 Takeuchi (1996, p.85-86)

$$\mathbf{A} = - \begin{pmatrix} 1 & 1.12 & 2 & 0.6 \\ 1.12 & 1 & 0.6 & 2 \\ -2 & -0.6 & 0 & 0 \\ -0.6 & -2 & 0 & 0 \end{pmatrix} \quad (\text{A.9a})$$

$$\vec{r} = (1 \ 1 \ -1 \ -1)^T \quad (\text{A.9b})$$

$$\vec{X}(t=0) = (0.2 \ 0.18 \ 0.05 \ 0.2)^T \quad (\text{A.9c})$$

The cycle was correctly identified.

A.1.2.6 Takeuchi (1996, p.85-86)

$$\mathbf{A} = - \begin{pmatrix} 1 & 1.18 & 2 & 0.6 \\ 1.18 & 1 & 0.6 & 2 \\ -2 & -0.6 & 0 & 0 \\ -0.6 & -2 & 0 & 0 \end{pmatrix} \quad (\text{A.10a})$$

$$\vec{r} = (1 \ 1 \ -1 \ -1)^T \quad (\text{A.10b})$$

$$\vec{X}(t=0) = (0.2 \ 0.18 \ 0.05 \ 0.2)^T \quad (\text{A.10c})$$

The cycle was correctly identified.

A.1.2.7 Takeuchi (1996, p.85-86) - Heteroclinic Cycle

$$\mathbf{A} = - \begin{pmatrix} 1 & 1.24 & 2 & 0.6 \\ 1.24 & 1 & 0.6 & 2 \\ -2 & -0.6 & 0 & 0 \\ -0.6 & -2 & 0 & 0 \end{pmatrix} \quad (\text{A.11a})$$

$$\vec{r} = (1 \ 1 \ -1 \ -1)^T \quad (\text{A.11b})$$

$$\vec{X}(t=0) = (0.2 \ 0.18 \ 0.05 \ 0.2)^T \quad (\text{A.11c})$$

The cycle was not identified. The attractor detected was an equilibrium point.

A.2 Temporal Evolution

In the following, we present 50 examples parametrized according to a simplified experiment design that followed the same procedure as the main results (see section 4.1). In this simplified version, each experiment was associated with a single replicate. The Latin hypercube sampling presented only ten distinct parameter combinations, and the number of initial populations was set to 20. The outcomes of all replicates are shown, together with additional information on their input parameters, initial rightmost eigenvalue, and the proportion of survival.

44

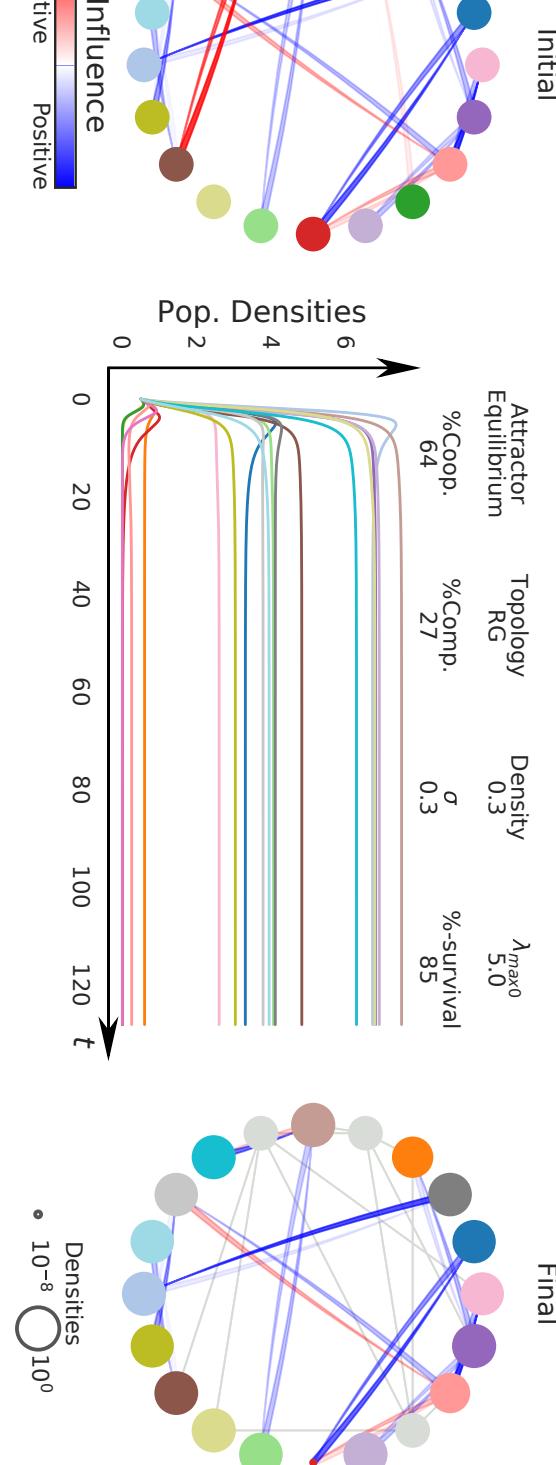
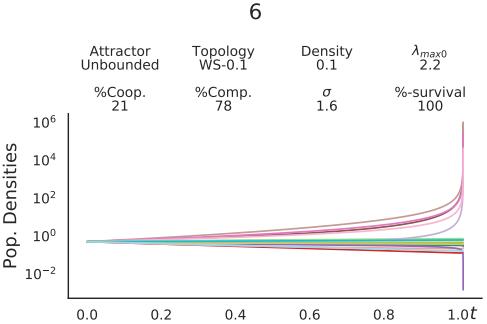
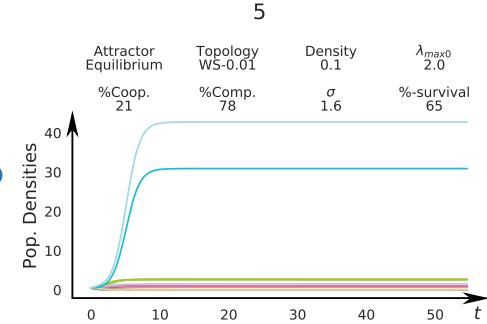
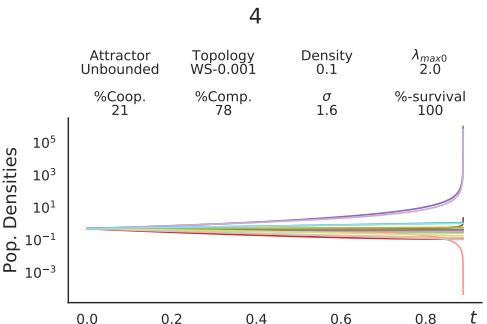
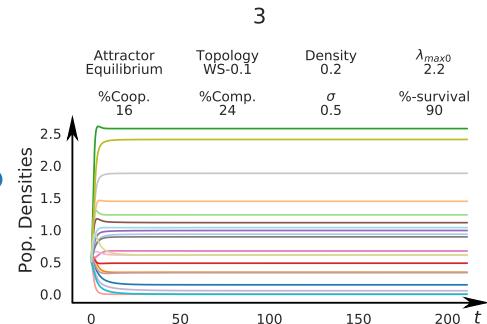
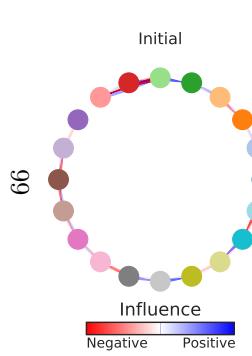
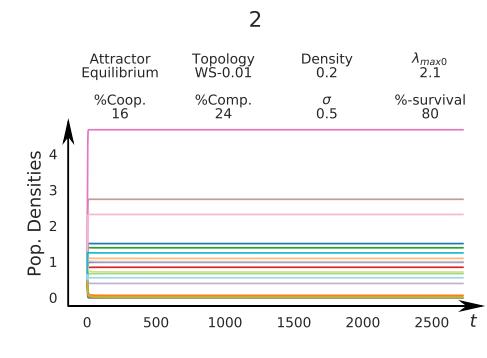
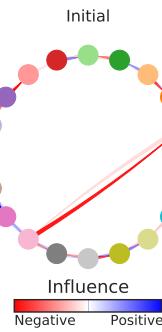
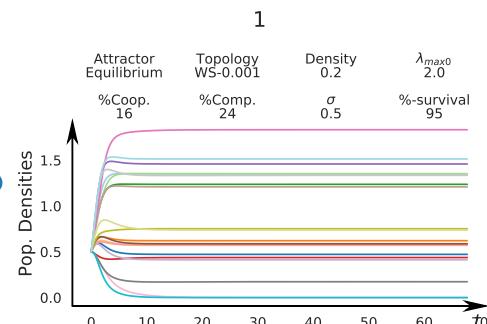
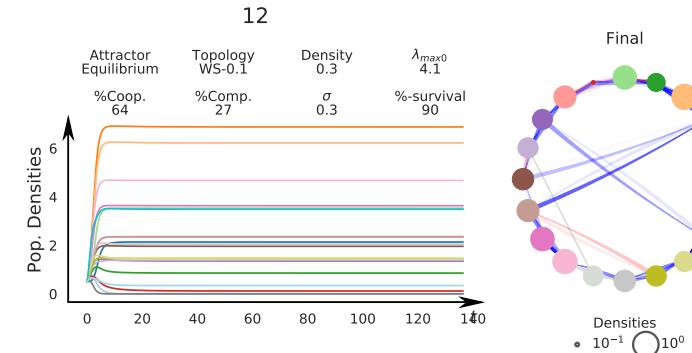
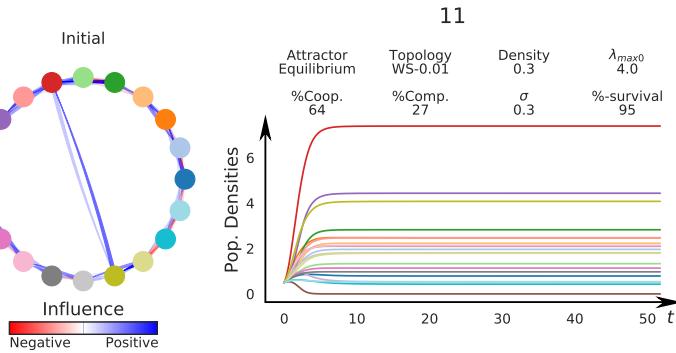
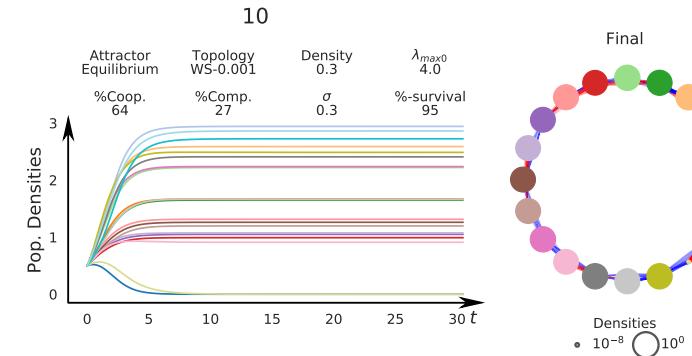
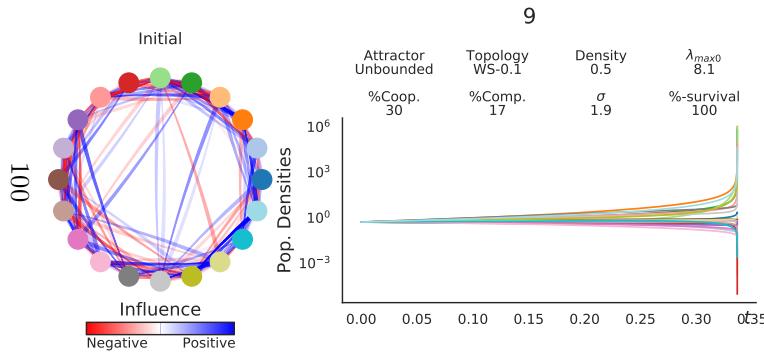
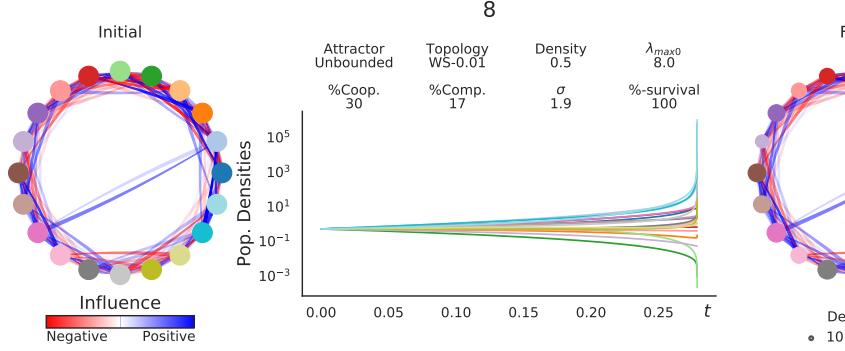
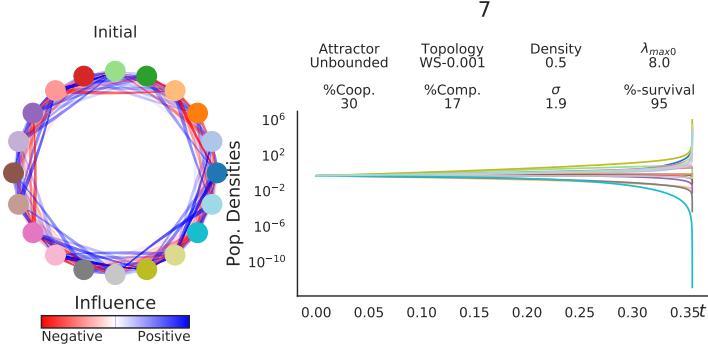
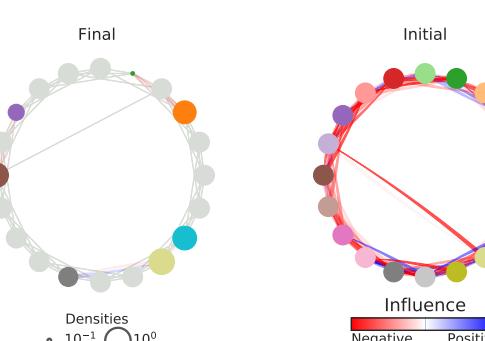
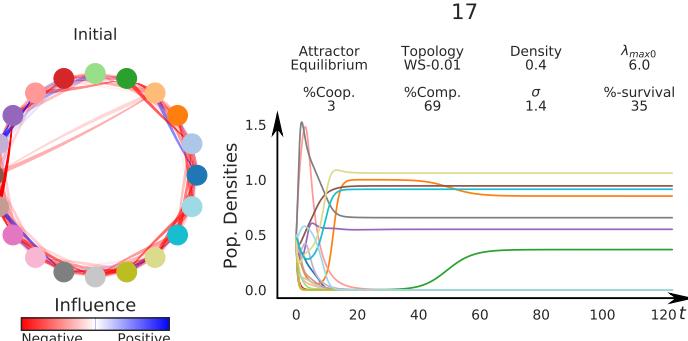
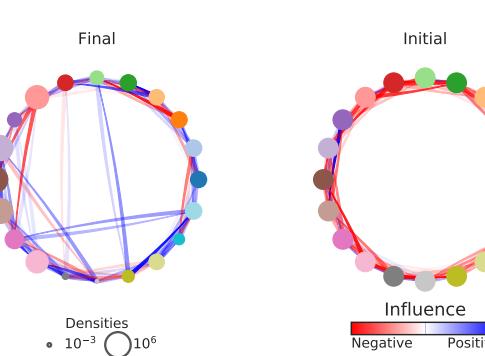
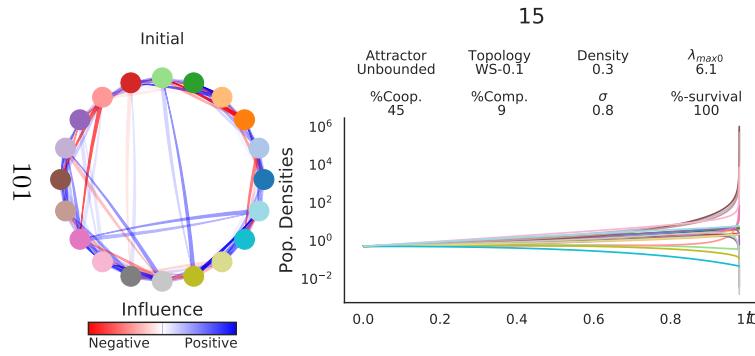
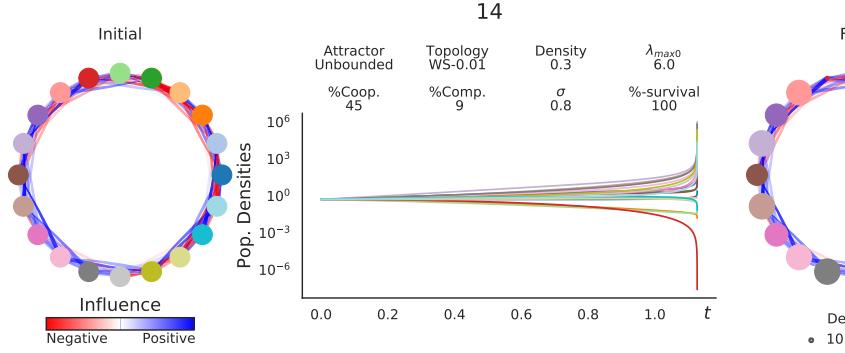
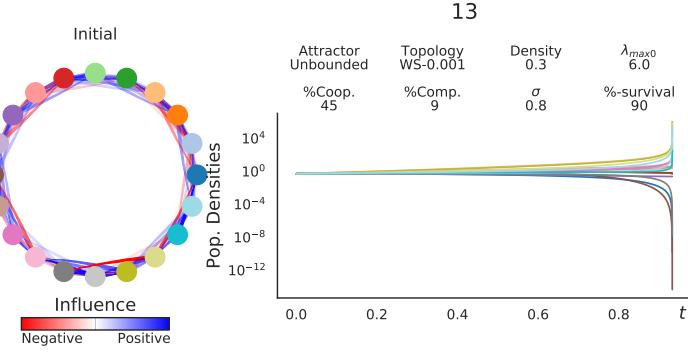
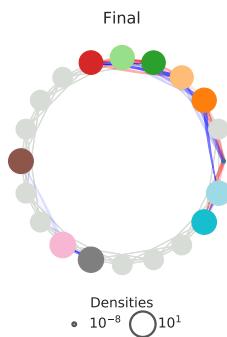
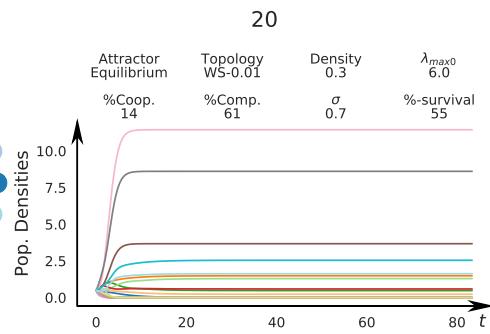
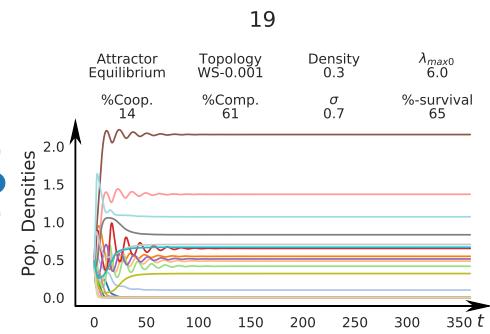


Figure A.1: **Example of Numerical Simulation:** In this and the figures below, the initial network is represented in the left. Each node corresponds to a different color and position, which is the same in the final network. The temporal evolution of the density of each population is pictured in the graph in the middle, with the color of the line being equal to the color of the corresponding node. Virtually extinct populations are indicated in the final network by light gray nodes adjacent to gray edges. The color of the edges indicates whether the influence of one population upon the other is positive (blue) or negative (red). The brightness of the edges indicates how large is the absolute value of the coefficient related to that interaction, more vivid colors represent more intense interactions. Edges are directed from the node adjacent to the side with larger width to the node adjacent to the thin side of the edge. The size of surviving nodes in the final community indicates the final density of the corresponding population. Input variables, detected attractors, and values of the rightmost eigenvalue and proportion of survival are presented in a table above the temporal evolution of the populations.

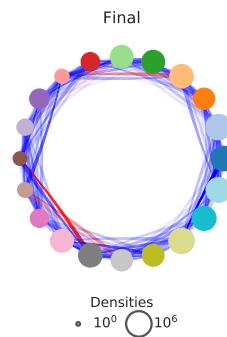
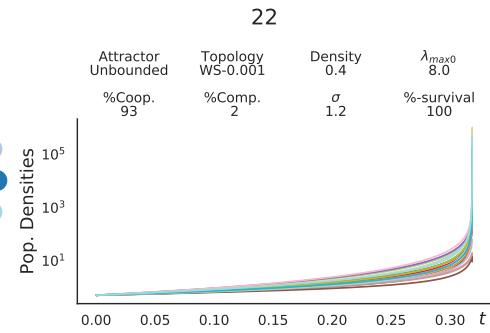
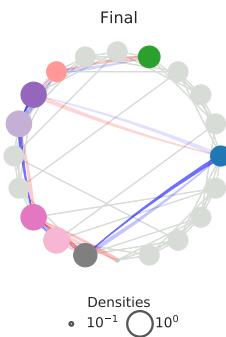
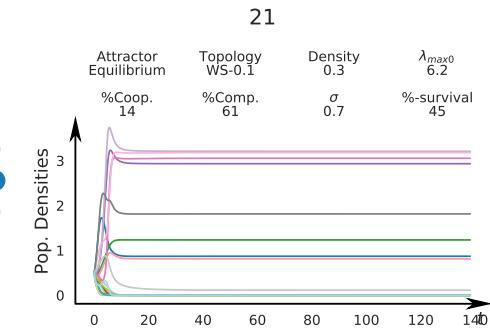




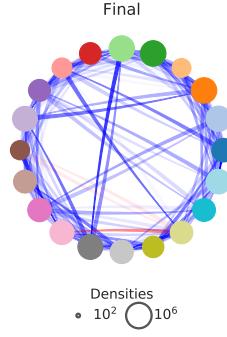
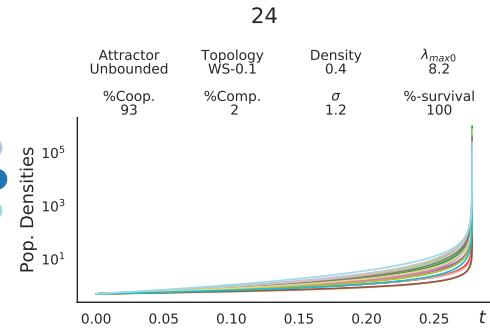
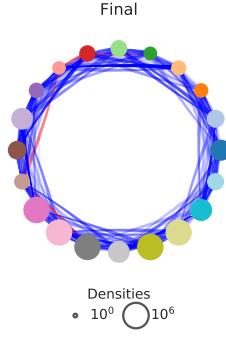
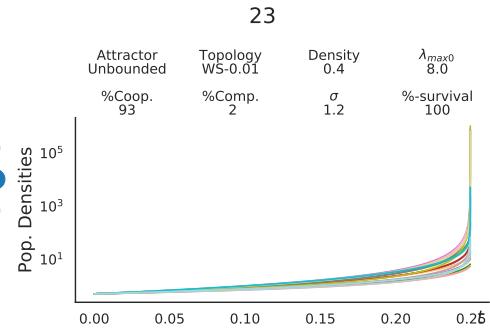


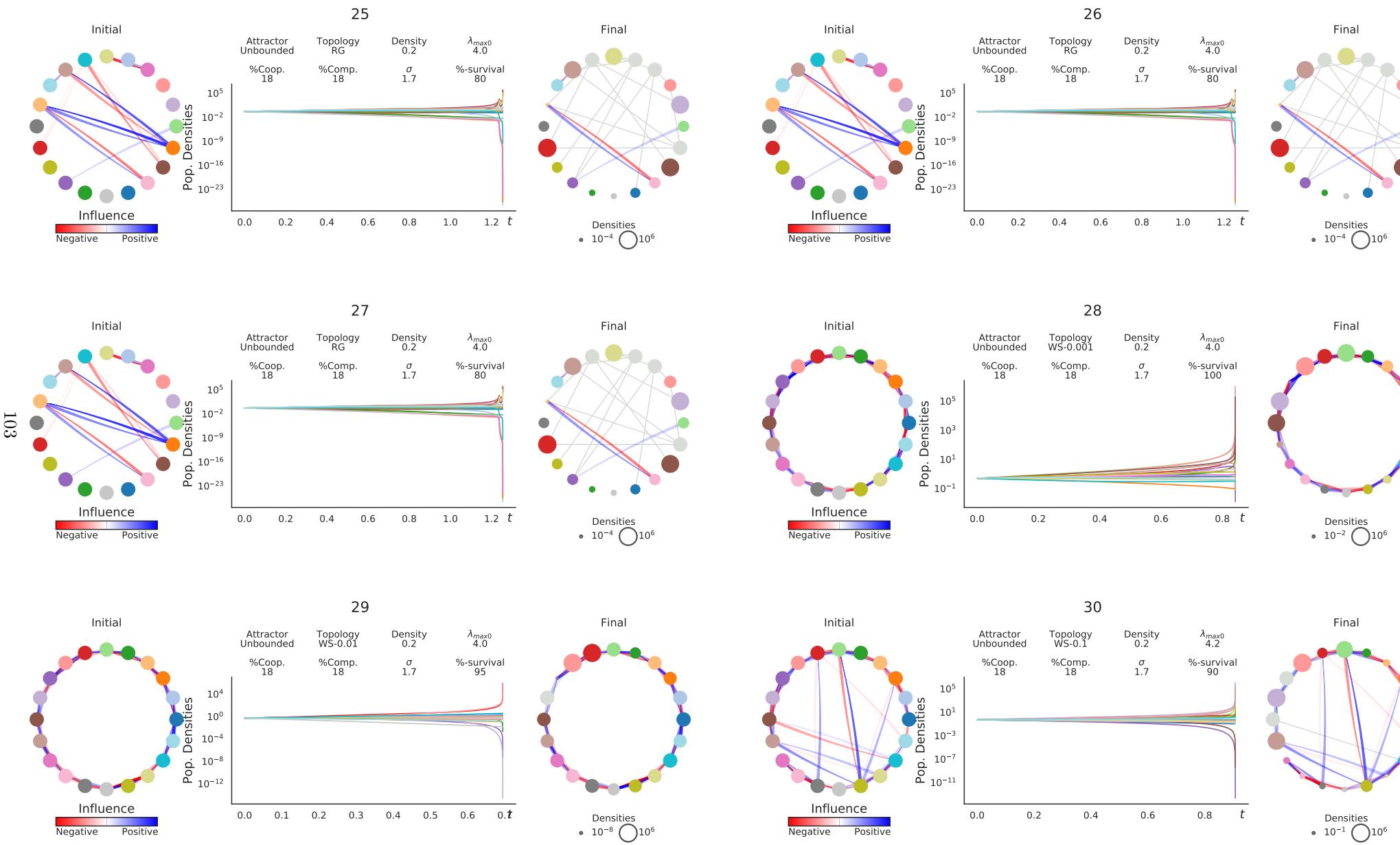


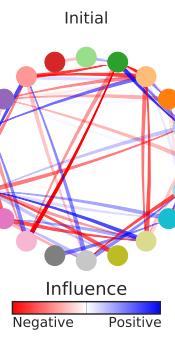
102



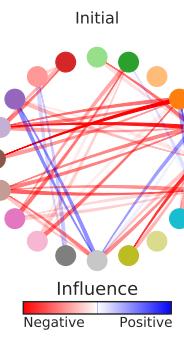
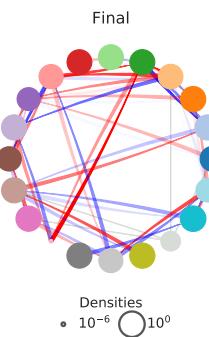
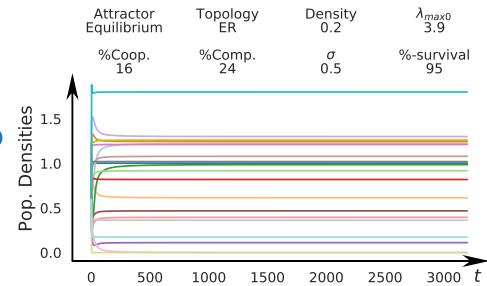
23



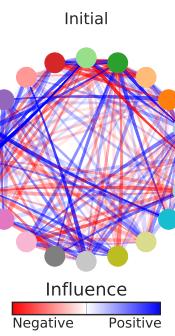
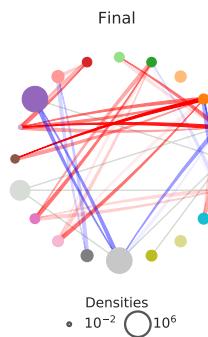
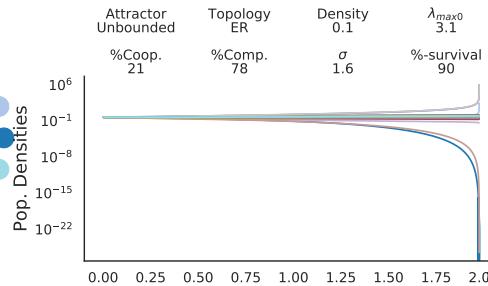




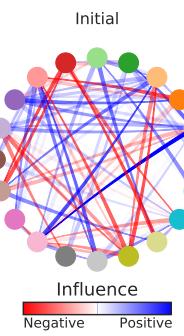
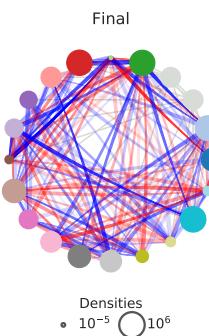
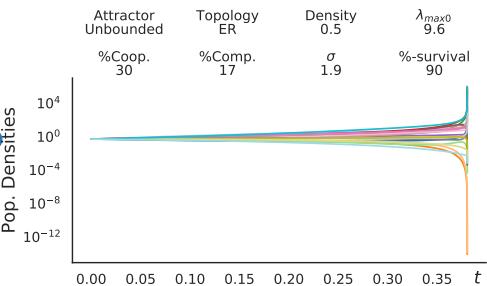
31



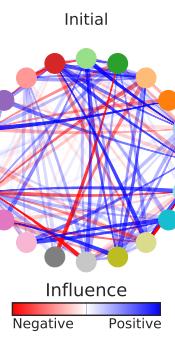
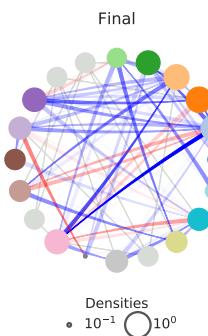
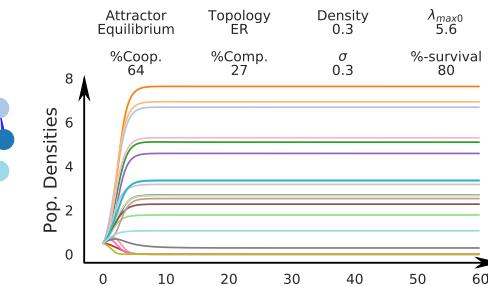
32



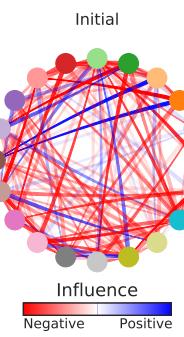
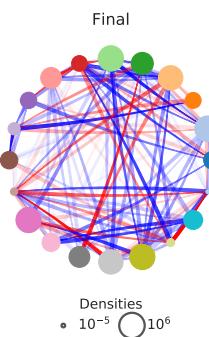
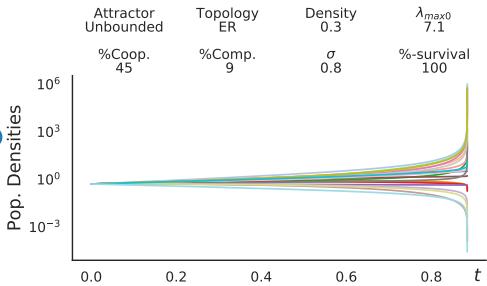
33



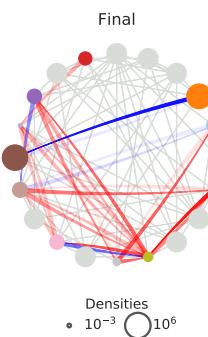
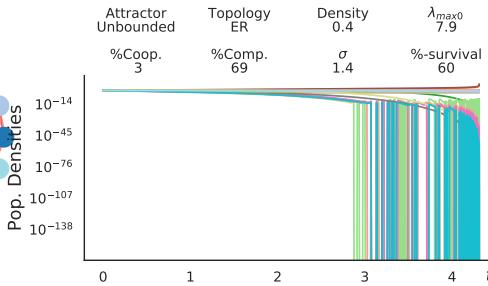
34



35

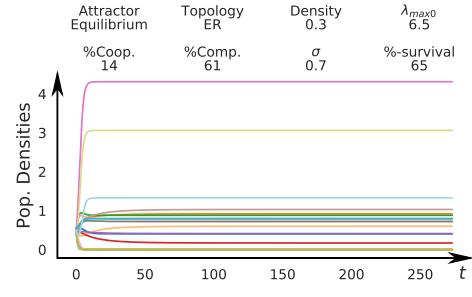
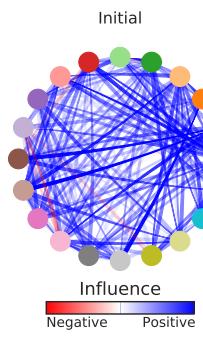


36

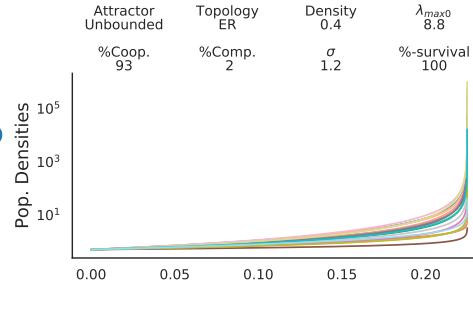
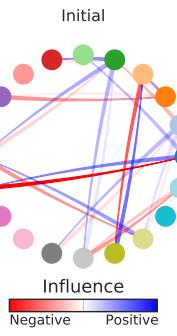




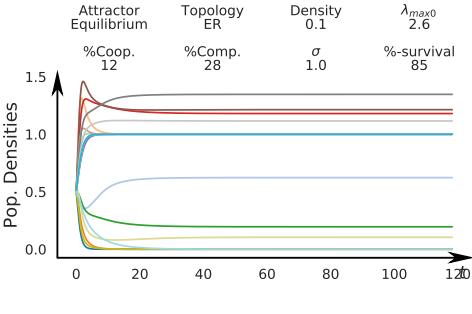
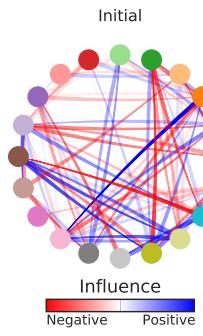
37

**Final**

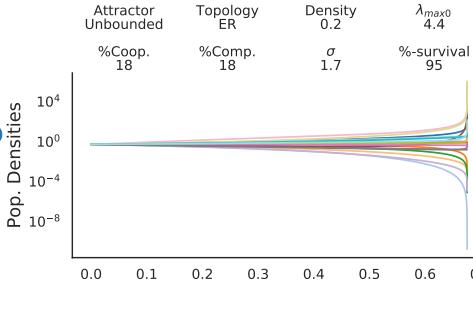
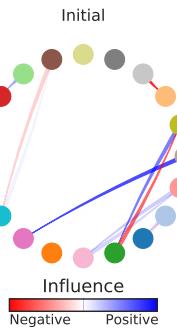
38

**Final**

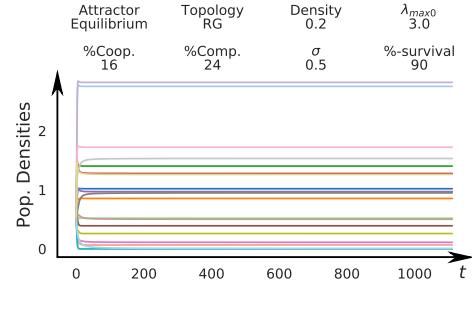
39

**Final**

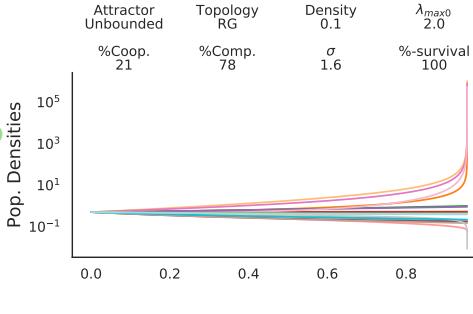
40

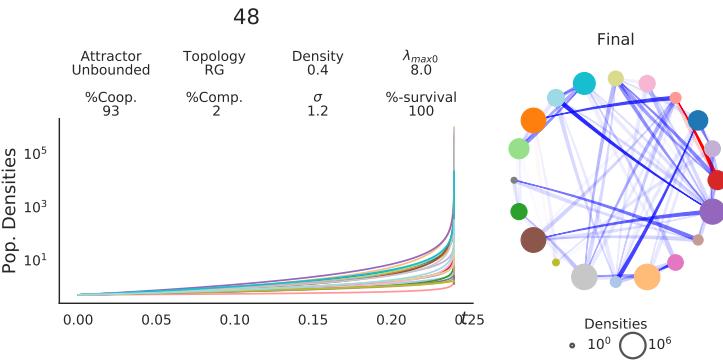
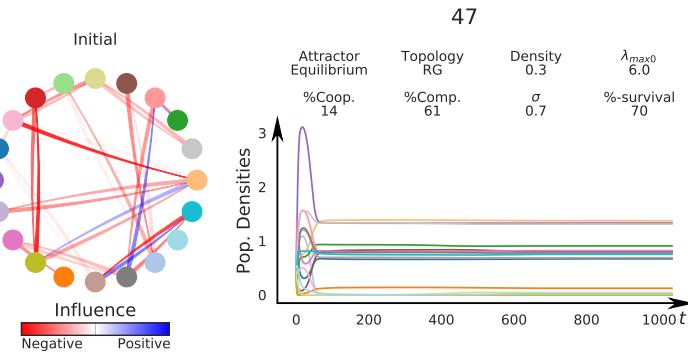
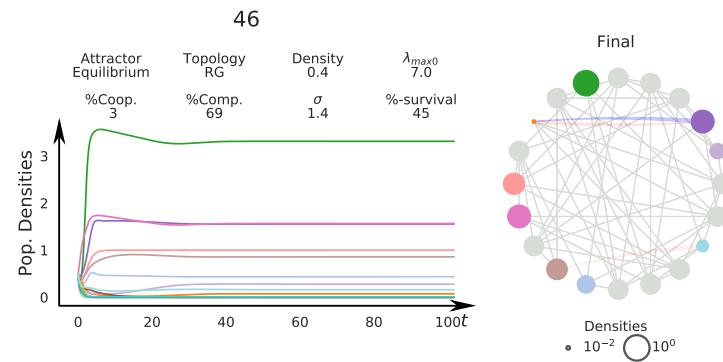
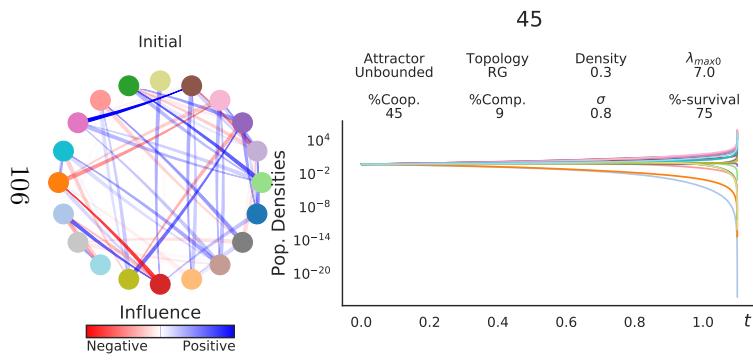
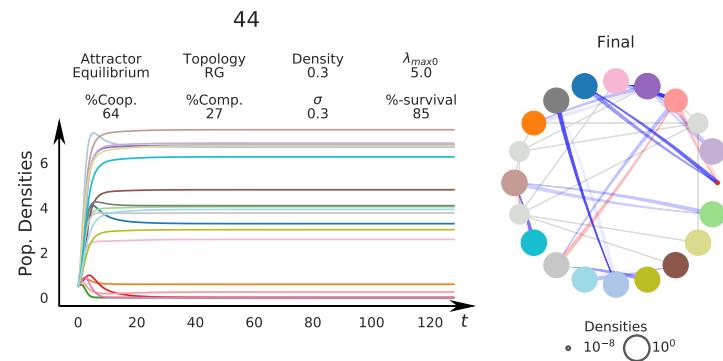
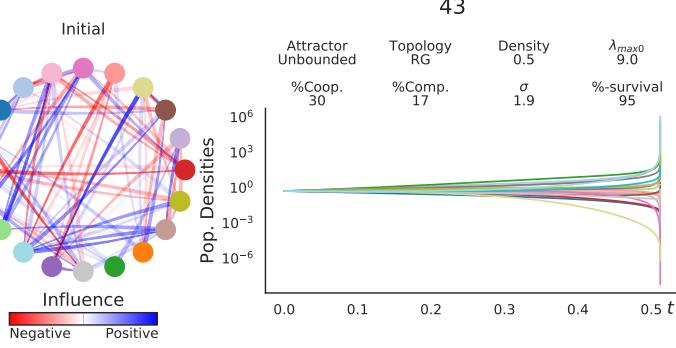
**Final**

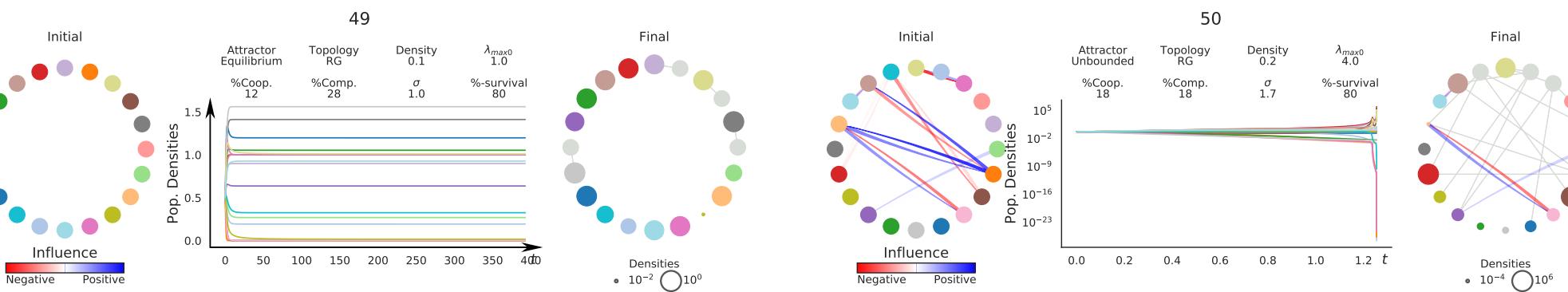
41

**Final**

42

**Final**





Appendix B

Supplementary Figures and Tables

Table B.1: List of Symbols and Variables used in Numerical Results

Variable	Representation	
Abundance Assortativity	abun. assort.	see section 2.2.1
Abundances	abund.	final abundances of populations
Average Amplitude	ave. amp.	average amplitude of oscillating populations
Eigenvalues	λ_i	eigenvalues calculated at the point in which equilibrium was detected
Initial/Final Average Shortest Path Lengths	l_{G0} / l_{Gf}	see section 2.2.1
Initial/Final Connected Networks	(0) / (f) connect.?	1 if network was connected, 0 otherwise
Initial/Final Degree Assortativities	r_{D0} / r_{Df}	see section 2.2.1
Initial/Final Degree Distributions	$P(k)_0 / P(k)_f$	see section 2.2.1
Initial/Final Densities	D_0 / D_f	see section 2.2.1
Initial/Final Diameter	d_0 / d_f	see section 2.2.1
Initial/Final Global Clustering coefficients	C_0 / C_f	see section 2.2.1
Initial/Final Local Clustering coefficients	\bar{C}_0 / \bar{C}_f	see section 2.2.1
Initial/Final Modularities	Q_0 / Q_f	see section 2.2.1
Initial/Final Numbers of Communities	(0) / (f) n comm.	number of connected components in the network
Initial/Final Numbers of Modules	(0) / (f) n modul.	number of modules in the network
Initial/Final Orders	N_0 / N_f	see section 2.2.1
Initial/Final Sizes	E_0 / E_f	see section 2.2.1
Proportion of competitive interactions (input)	% -/-	proportion of competitive interactions on the community
Proportion of cooperative interactions (input)	% +/+	proportion of cooperative interactions on the community
Proportion of exploitative interactions	% +/-	calculated from proportions of competition and cooperation, based on equation (4.1)
Density (input)	D_{in}	input density of interactions
Order (input)	N_{in}	number of initial populations in the community (set to 400)
WS topology Reconnection Probability (input)	p_{in}	probability of rewiring; see section 2.2.1

Continued on next page

Variable	Representation	
Interaction Intensity	σ_{in}	see section 4.1.3
Standard Deviation		
Number of Coexisting Populations	N populations	number of extant populations in the community
Oscillation Period	T	oscillation period in communities with cyclic attractors
Proportion of Survival	% surv.	number of coexisting populations divided by number of initial populations
Rightmost Eigenvalue	λ_{max}	real part of the eigenvalue with largest real part
seed (input)	seed	random seed used as input; see section 4.1.4
Final Time for Equilibrium Attractors	EP t_f	see section 4.3.1
Type of Dynamics / Attractor	Type of Dynamics	see section 4.2

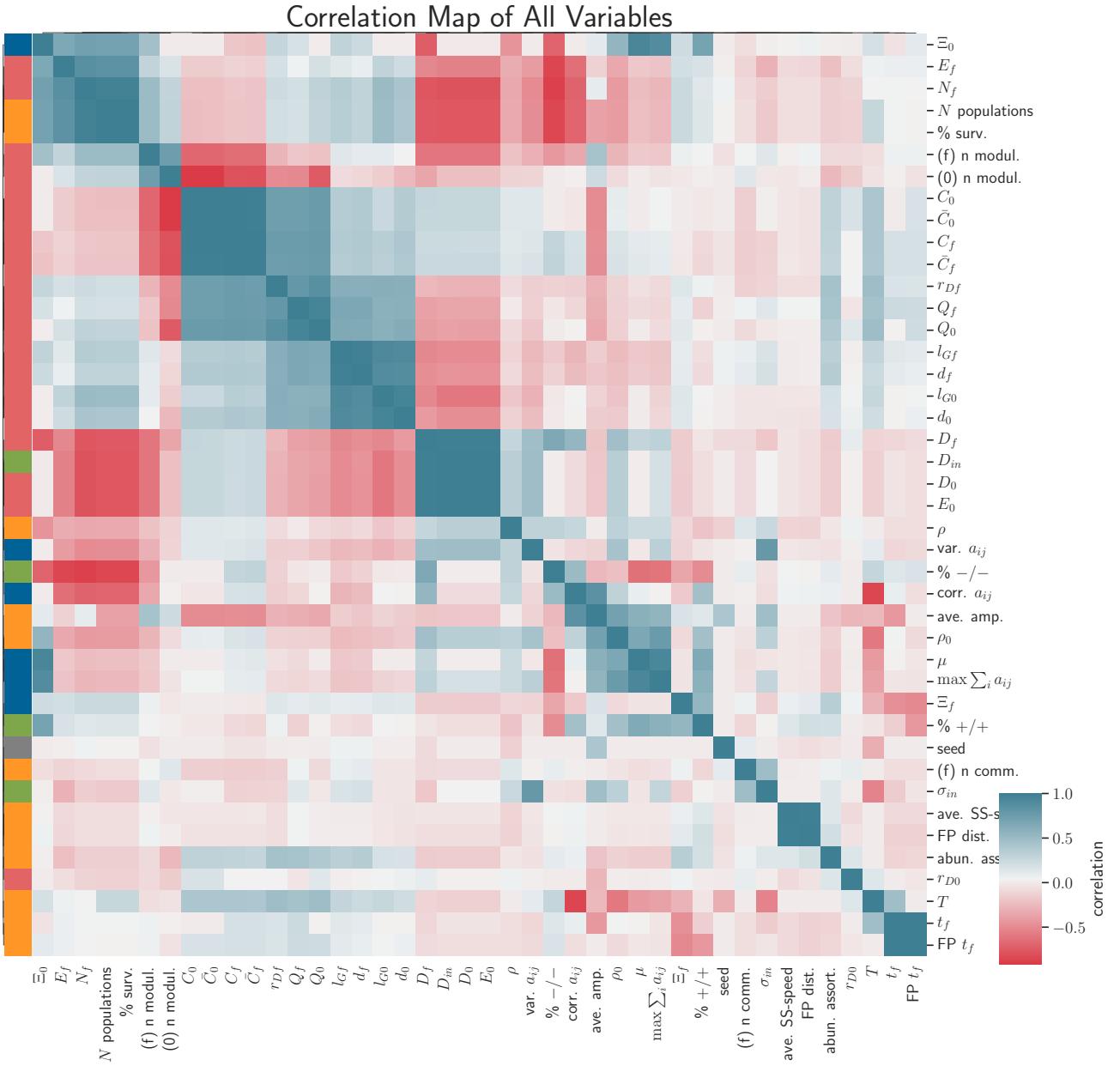


Figure B.1: **The correlation between all pairs of measured variables:** The Pearson correlation of each pair of variables is plotted as a symmetrical heatmap with correlated/uncorrelated pairs represented by blue/red color. Variables are clustered together and sorted according to the euclidean distance between the vector of correlations. At the left, a color code indicates the variable type: the dynamic variables are represented in yellow, the topological variables are in brown, the interaction variables are in blue, the input variables are in green, and identification variables are in gray. Some groups of variables display similar properties, as seen by adjacent analogous colors. The variables in each group correlate with all measured variables similarly. For reference on the symbols used, see table B.1.

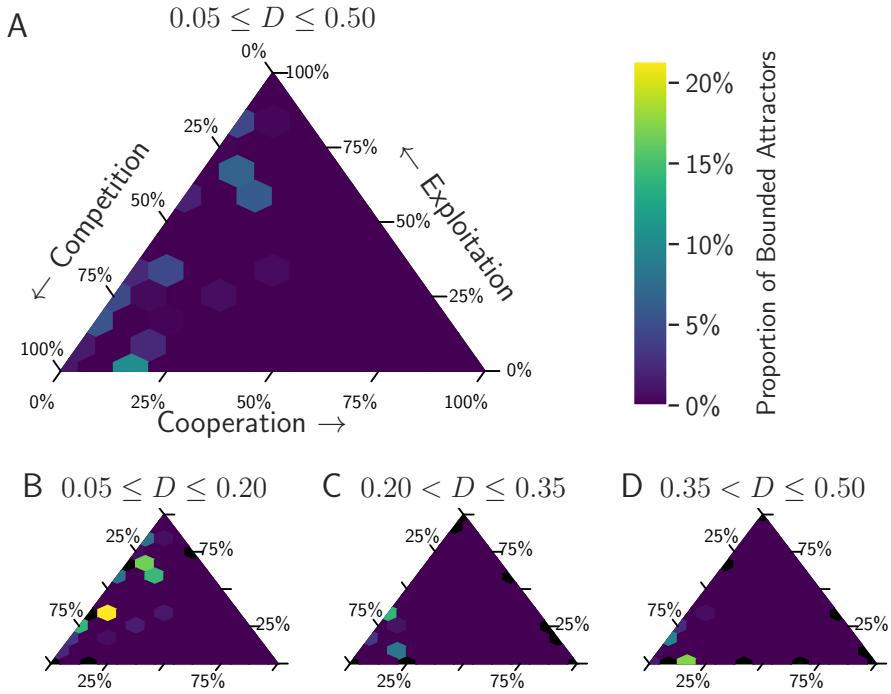


Figure B.2: Proportion of scenarios displaying bounded dynamics distributed across interaction type and density: The proportion of scenarios displaying bounded dynamics is presented in a ternary heatmap. Brighter colors indicate a larger proportion of bounded scenarios. This proportion was measured by dividing the number of scenarios displaying bounded dynamics (equilibrium points and cycles) by the total number of scenarios inside each hexagonal grid cell. The input density of interactions of the network is measured as the ratio of the number of pairwise interactions in the network divided by the maximum number of pairwise interactions that could occur with the same number of populations. Regions in which no scenario was simulated are colored black. The vast regions colored purple are related to the fact that 98.97 % of the simulated scenarios were unbounded, in total. (A) Overall, most of the bounded scenarios were found for a small portion of cooperation and varying amounts of asymmetric and competitive interactions, with a larger proportion of competitive interactions related to more bounded scenarios. (B) The largest proportion of bounded dynamics was found for scenarios with few cooperative interactions ($< 25\%$) and less dense interactions ($0.05 \leq D \leq 0.20$). (C) and (D) Albeit bounded scenarios were much rarer in communities with larger interaction densities, we found bounded dynamics associated with highly competitive communities.

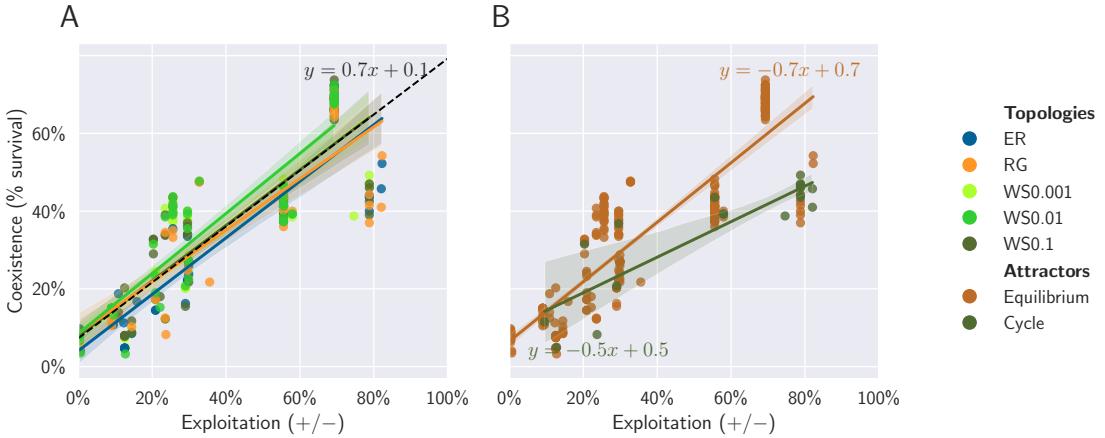


Figure B.3: Percentage of surviving population increases as proportion of Exploitation interaction increases The percentage of survivors is estimated by the number of nodes in the final network divided by the number of nodes in the initial network. The percentage of survivors in the final network is affected by the proportion of exploitative interactions; as the proportion of exploitation increases, the percentage of surviving population increases. (A) A comparable result was found for all topologies, as it is possible to see the similarity between colored full lines and black dashed line. The black line intercept and angular coefficient were calculated as the mean of these parameters for the colored lines. (B) Equilibrium point and cycle dynamics show a similar overall trend. Equilibrium point scenarios had a more steep increase in the percentage of surviving populations with increasing proportions of exploitative interactions, as indicated by the angular coefficient than cycle dynamics.

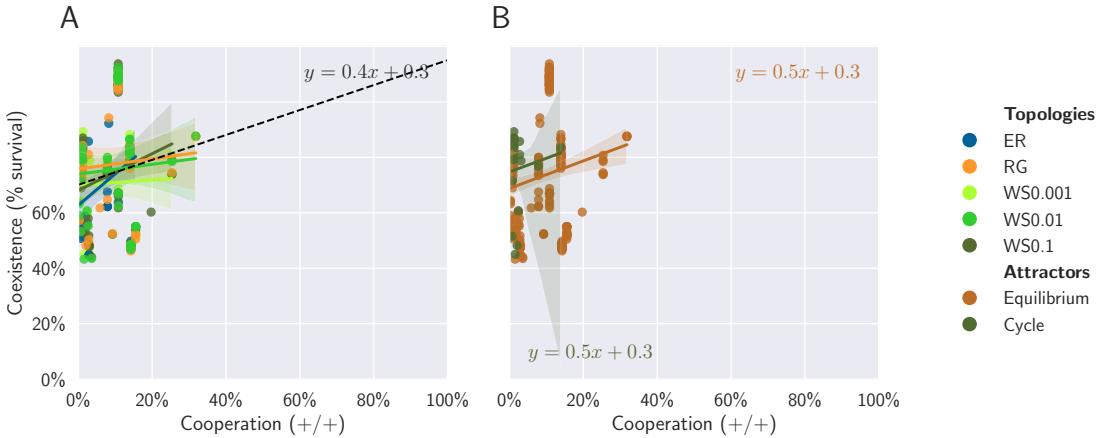


Figure B.4: Percentage of surviving population increases as proportion of Cooperation interaction increases: The percentage of survivors is estimated by the number of nodes in the final network divided by the number of nodes in the initial network. The percentage of survivors in the final network is affected by the proportion of exploitative interactions; as the proportion of exploitation increases, the percentage of surviving population increases. (A) A comparable result was found for all topologies, as it is possible to see the similarity between colored full lines and black dashed line. The black line intercept and angular coefficient were calculated as the mean of these parameters for the colored lines. (B) Equilibrium point and cycle dynamics show a similar overall trend. Equilibrium point scenarios had a more steep increase in the percentage of surviving populations with increasing proportions of exploitative interactions, as indicated by the angular coefficient than cycle dynamics.

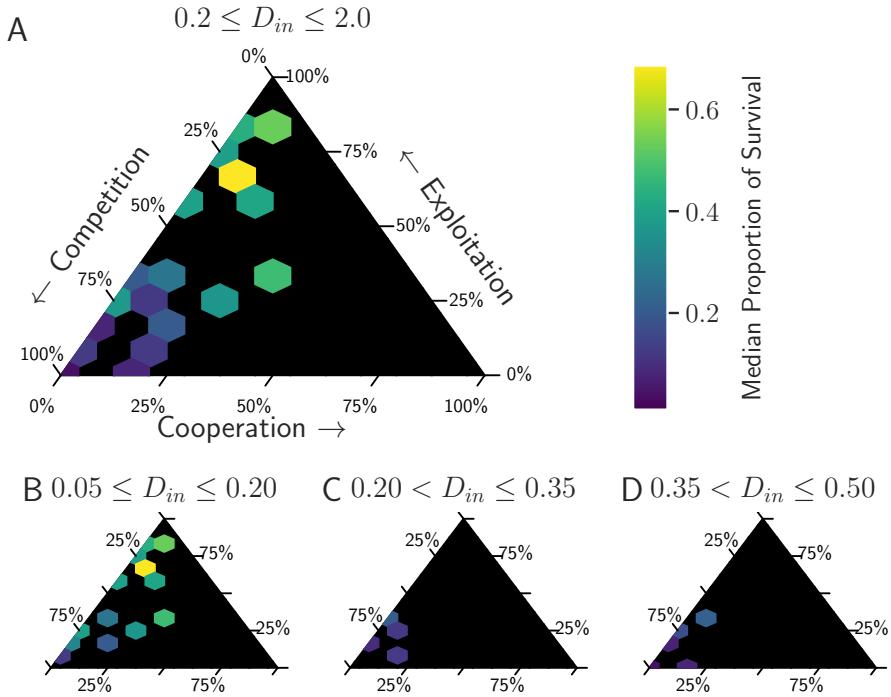


Figure B.5: The proportion of scenarios displaying bounded dynamics distributed across interaction type and intensity The percentage of survivors is estimated by the number of nodes in the final network divided by the number of nodes in the initial network. The density indicates how heavily taxa are connected in a given network. The proportion was measured by dividing the number of scenarios displaying bounded dynamics (equilibrium points, and cycles) by the total number of scenarios inside each hexagonal grid cell. Regions in which no scenario was simulated are colored black.

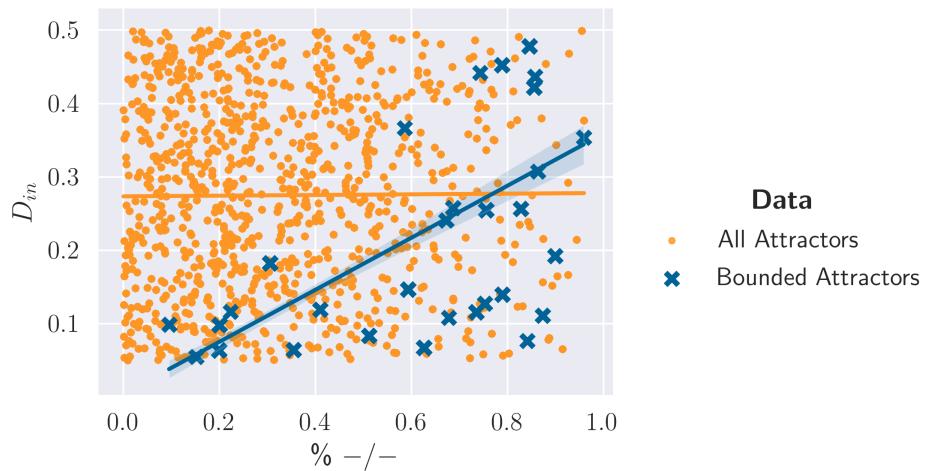


Figure B.6: Distribution of Input Density and Proportion of Competition: The percentage of survivors is estimated by the number of nodes in the final network divided by the number of nodes in the initial network. The overall distribution of input density and proportion of competition is shown in orange dots. The subset of the distribution of input density and proportion of competition leading to bounded attractors is highlighted in blue crosses.

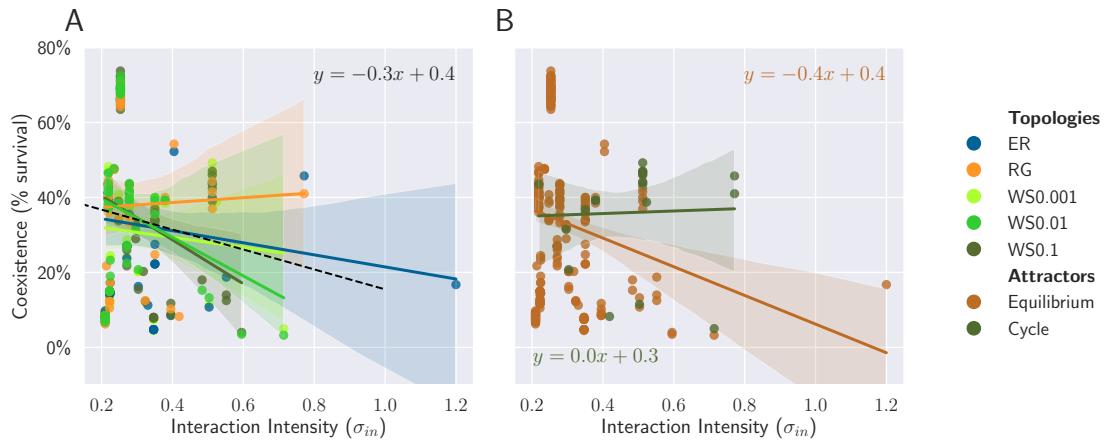


Figure B.7: Order of the Final Network Decreases with Increasing Competition: Order of the final network is affected by proportion of competition in the network, more competition being related smaller networks. (A) A comparable result was found for all topologies, as it is possible to see the similarity between colored full lines and black dashed line. The black line intercept and angular coefficient were calculated as the mean of these parameters for the colored lines. (B) Equilibria and cyclic were associated to distinct trends on the reduction of final network order. Cyclic scenarios had a more steep decrease in order of final network with increasing competition as indicated by the angular coefficient. Scenarios with cycles, however, had a larger decrease in order for the no competition best case scenario as revealed by lower intercept.

Table B.2: Regressions Between the Proportion of Survival and Input Variables: A summary of statistical properties of regressions between the proportion of survival and proportion of competition, exploitation and cooperation, density, and standard deviation of the distribution of interaction intensities.

x-variable			r^2	AIC	n	a	95% CI	p-value	b	95% CI	p-value		
0	Prop. Competition (-/-)	<i>Overall</i>	0.72	-437.6	288	-0.68	(-0.73, -0.63)	3.5×10^{-80}	0.70	(0.67, 0.73)	2.6×10^{-134}		
1		<i>Intercept</i>	0.0	-76.96	288				0.33	(0.31, 0.36)	9.9×10^{-80}		
2		<i>Topology</i>	WS0.001	0.60	-88.95	77	-0.64	(-0.76, -0.52)	2.1×10^{-16}	0.68	(0.60, 0.76)	1.1×10^{-27}	
3		ER		0.85	-109.1	54	-0.71	(-0.80, -0.63)	8.3×10^{-23}	0.70	(0.65, 0.75)	5.7×10^{-33}	
4		RG		0.67	-52.99	44	-0.70	(-0.85, -0.54)	1.3×10^{-11}	0.68	(0.60, 0.75)	8.1×10^{-21}	
5		WS0.1		0.78	-105.1	58	-0.69	(-0.79, -0.59)	6.7×10^{-20}	0.72	(0.66, 0.78)	3.4×10^{-31}	
6		WS0.01		0.76	-93.75	55	-0.71	(-0.82, -0.60)	3.2×10^{-18}	0.75	(0.68, 0.81)	4.7×10^{-29}	
7		<i>Attractor</i>	Equilibrium	0.80	-494.3	269	-0.72	(-0.77, -0.68)	5.1×10^{-94}	0.73	(0.71, 0.76)	5.2×10^{-145}	
8			Cycle	0.19	-9.17	19	-0.32	(-0.66, 0.01)		0.41	(0.25, 0.57)	4.5×10^{-5}	
9	Prop. Exploitation (+/-)	<i>Overall</i>	0.70	-420.6	288	0.69	(0.64, 0.74)	1.6×10^{-76}	0.08	(0.05, 0.10)	2.9×10^{-10}		
10		<i>Intercept</i>	0.0	-76.96	288				0.33	(0.31, 0.36)	9.9×10^{-80}		
11		<i>Topology</i>	WS0.001	0.60	-90.26	77	0.65	(0.53, 0.78)	1.1×10^{-16}	0.08	(0.03, 0.13)	1.1×10^{-3}	
12		ER		0.82	-99.51	54	0.73	(0.63, 0.82)	8.6×10^{-21}	0.04	(-5.0 $\times 10^{-3}$, 0.09)	0.08	
13		RG		0.57	-41.48	44	0.62	(0.45, 0.79)	3.5×10^{-9}	0.09	(7.2 $\times 10^{-3}$, 0.18)	0.03	
14		WS0.1		0.75	-99.35	58	0.71	(0.61, 0.82)	1.1×10^{-18}	0.08	(0.03, 0.13)	2.1×10^{-3}	
15		WS0.01		0.82	-108.8	55	0.77	(0.67, 0.87)	2.2×10^{-21}	0.08	(0.04, 0.13)	1.9×10^{-4}	
16		<i>Attractor</i>	Equilibrium	0.81	-507.6	269	0.76	(0.72, 0.81)	6.8×10^{-97}	0.07	(0.05, 0.09)	2.4×10^{-10}	
17			Cycle	0.23	-9.92	19	0.34	(0.02, 0.67)		0.04	0.09	(-0.12, 0.29)	0.38
18	Prop. Cooperation (+/+)	<i>Overall</i>	0.02	-81.20	288	0.48	(0.10, 0.85)		0.01	0.29	(0.25, 0.33)	1.6×10^{-32}	
19		<i>Intercept</i>	0.0	-76.96	288				0.33	(0.31, 0.36)	9.9×10^{-80}		
20		<i>Topology</i>	WS0.001	1.4×10^{-3}	-19.37	77	0.13	(-0.66, 0.92)		0.75	0.28	(0.21, 0.36)	4.5×10^{-10}
21		ER		0.08	-12.19	54	1.13	(0.03, 2.23)		0.04	0.23	(0.12, 0.34)	1.3×10^{-4}
22		RG		0.01	-5.19	44	0.35	(-0.53, 1.22)		0.43	0.33	(0.21, 0.45)	2.2×10^{-6}
23		WS0.1		0.04	-20.65	58	0.66	(-0.16, 1.48)		0.11	0.28	(0.19, 0.37)	5.5×10^{-8}
24		WS0.01		3.2×10^{-3}	-14.65	55	0.18	(-0.68, 1.03)		0.68	0.34	(0.24, 0.44)	1.3×10^{-8}
25		<i>Attractor</i>	Equilibrium	0.02	-73.40	269	0.50	(0.10, 0.89)		0.01	0.29	(0.24, 0.33)	2.5×10^{-28}

Continued on next page

	x-variable		r^2	AIC	n	a	95% CI	p-value	b	95% CI	p-value	
26		Cycle	0.09	-6.88	19	-1.85	(-4.85, 1.15)	0.21	0.33	(0.21, 0.45)	2.1×10^{-5}	
27	Input Density (D_{in})	<i>Overall</i>	0.59	-330.6	288	-1.18	(-1.30, -1.07)	4.6×10^{-57}	0.56	(0.54, 0.59)	4.6×10^{-120}	
28		<i>Intercept</i>	0.0	-76.96	288				0.33	(0.31, 0.36)	9.9×10^{-80}	
29		<i>Topology</i>	WS0.001	0.60	-89.63	77	-1.13	(-1.34, -0.91)	1.5×10^{-16}	0.54	(0.49, 0.60)	5.7×10^{-31}
30		ER	0.58	-54.53	54	-1.12	(-1.38, -0.85)	2.6×10^{-11}	0.55	(0.48, 0.62)	2.0×10^{-22}	
31		RG	0.60	-45.02	44	-1.31	(-1.64, -0.98)	6.3×10^{-10}	0.60	(0.53, 0.67)	5.0×10^{-20}	
32		WS0.1	0.51	-59.93	58	-1.13	(-1.42, -0.83)	2.4×10^{-10}	0.54	(0.47, 0.60)	1.1×10^{-23}	
33		WS0.01	0.65	-71.72	55	-1.31	(-1.58, -1.05)	1.4×10^{-13}	0.61	(0.54, 0.67)	2.3×10^{-26}	
34		<i>Attractor</i>	Equilibrium	0.62	-326.5	269	-1.21	(-1.32, -1.09)	8.5×10^{-58}	0.58	(0.55, 0.60)	4.0×10^{-117}
35		Cycle	0.35	-13.21	19	-1.03	(-1.74, -0.31)	7.8×10^{-3}	0.44	(0.31, 0.57)	2.5×10^{-6}	
36	Interaction Intensity (σ_{in})	<i>Overall</i>	0.03	-84.04	288	-0.32	(-0.53, -0.11)	2.7×10^{-3}	0.43	(0.36, 0.49)	2.2×10^{-29}	
37		<i>Intercept</i>	0.0	-76.96	288				0.33	(0.31, 0.36)	9.9×10^{-80}	
38		<i>Topology</i>	WS0.001	0.01	-20.34	77	-0.26	(-0.75, 0.24)	0.31	0.37	(0.22, 0.53)	9.0×10^{-6}
39		ER	0.01	-8.74	54	-0.16	(-0.53, 0.20)	0.38	0.38	(0.24, 0.51)	4.7×10^{-7}	
40		RG	0.03	-5.67	44	-0.31	(-0.92, 0.29)	0.30	0.46	(0.27, 0.64)	1.0×10^{-5}	
41		WS0.1	0.09	-23.20	58	-0.59	(-1.10, -0.07)	0.03	0.52	(0.35, 0.69)	5.8×10^{-8}	
42		WS0.01	0.06	-17.80	55	-0.51	(-1.08, 0.05)	0.07	0.50	(0.33, 0.66)	1.6×10^{-7}	
43		<i>Attractor</i>	Equilibrium	0.03	-76.43	269	-0.38	(-0.63, -0.13)	2.6×10^{-3}	0.44	(0.37, 0.52)	1.0×10^{-25}
44		Cycle	1.3×10^{-4}	-5.08	19	-0.01	(-0.68, 0.65)	0.96	0.29	(-0.05, 0.63)	0.09	

Table B.3: Regressions Between Rightmost Eigenvalue and Input Variables: A summary of statistical properties of regressions between the rightmost eigenvalue and proportion of competition, exploitation and cooperation, density, and standard deviation of the distribution of interaction intensities.

	x-variable			r^2	AIC	n	a	95% CI	p-value	b	95% CI	p-value
0	Prop. Competition ($-/-$)	<i>Overall</i>	—	0.05	4.7×10^5	5.0×10^4	-25.95	(-26.96, -24.94)	0.0	32.68	(32.27, 33.08)	0.0
1		<i>Intercept</i>	—	0.0	4.7×10^5	5.0×10^4	—	—	—	24.10	(23.86, 24.34)	0.0
2		<i>Topology</i>	WS0.001	0.04	9.4×10^4	9943	-24.06	(-26.37, -21.75)	1.1×10^{-90}	32.44	(31.51, 33.37)	0.0
3		ER	0.05	9.3×10^4	9959	-26.86	(-29.10, -24.61)	1.5×10^{-118}	32.97	(32.06, 33.87)	0.0	
4		RG	0.05	9.4×10^4	9970	-25.42	(-27.70, -23.14)	2.1×10^{-103}	32.56	(31.64, 33.48)	0.0	
5		WS0.1	0.05	9.4×10^4	9959	-26.05	(-28.32, -23.78)	3.8×10^{-109}	32.65	(31.73, 33.56)	0.0	
6		WS0.01	0.06	9.2×10^4	9944	-27.36	(-29.52, -25.21)	1.1×10^{-132}	32.78	(31.91, 33.65)	0.0	
7		<i>Attractor</i>	Unbounded	0.05	4.6×10^5	5.0×10^4	-26.76	(-27.76, -25.75)	0.0	32.90	(32.50, 33.30)	0.0
8			Equilibrium	0.15	2363.8	223	75.97	(52.40, 99.54)	1.2×10^{-9}	-11.28	(-25.29, 2.73)	0.11
9			Cycle	0.02	39.35	5	4.02	(-48.37, 56.41)	0.82	8.18	(-19.96, 36.32)	0.42
10	Prop. Exploitation (+/ $-$)	<i>Overall</i>	—	0.05	4.7×10^5	5.0×10^4	-26.90	(-27.89, -25.91)	0.0	33.10	(32.70, 33.51)	0.0
11		<i>Intercept</i>	—	0.0	4.7×10^5	5.0×10^4	—	—	—	24.10	(23.86, 24.34)	0.0
12		<i>Topology</i>	WS0.001	0.05	9.4×10^4	9943	-27.04	(-29.30, -24.78)	2.2×10^{-118}	33.55	(32.62, 34.47)	0.0
13		ER	0.05	9.3×10^4	9959	-26.80	(-29.02, -24.59)	3.9×10^{-121}	33.04	(32.14, 33.94)	0.0	
14		RG	0.05	9.4×10^4	9970	-27.14	(-29.38, -24.90)	2.8×10^{-121}	33.23	(32.32, 34.15)	0.0	
15		WS0.1	0.05	9.4×10^4	9959	-26.47	(-28.70, -24.23)	4.5×10^{-116}	32.90	(31.98, 33.81)	0.0	
16		WS0.01	0.06	9.2×10^4	9944	-27.05	(-29.17, -24.93)	7.0×10^{-134}	32.80	(31.94, 33.67)	0.0	
17		<i>Attractor</i>	Unbounded	0.05	4.6×10^5	5.0×10^4	-26.64	(-27.63, -25.65)	0.0	32.99	(32.59, 33.39)	0.0
18			Equilibrium	0.14	2366.6	223	-74.99	(-99.28, -50.70)	5.1×10^{-9}	56.29	(45.35, 67.22)	4.2×10^{-20}
19			Cycle	4.9×10^{-3}	39.42	5	-2.09	(-56.64, 52.46)	0.91	11.05	(-20.52, 42.62)	0.35
20	Prop. Cooperation (+/+)	<i>Overall</i>	—	0.21	4.6×10^5	5.0×10^4	54.14	(53.21, 55.06)	0.0	5.98	(5.61, 6.35)	2.5×10^{-213}
21		<i>Intercept</i>	—	0.0	4.7×10^5	5.0×10^4	—	—	—	24.10	(23.86, 24.34)	0.0
22		<i>Topology</i>	WS0.001	0.19	9.2×10^4	9943	52.33	(50.19, 54.46)	0.0	6.96	(6.09, 7.83)	2.5×10^{-55}

Continued on next page

	x-variable		r^2	AIC	n	a	95% CI	p-value	b	95% CI	p-value			
23		ER	0.22	9.2×10^4	9959	54.94	(52.89, 56.99)	0.0	5.69	(4.86, 6.53)	1.0×10^{-40}			
24		RG	0.20	9.2×10^4	9970	53.82	(51.73, 55.92)	0.0	6.16	(5.31, 7.01)	2.0×10^{-45}			
25		WS0.1	0.20	9.2×10^4	9959	53.82	(51.73, 55.90)	0.0	6.03	(5.18, 6.88)	7.3×10^{-44}			
26		WS0.01	0.24	9.0×10^4	9944	55.77	(53.83, 57.71)	0.0	5.06	(4.27, 5.85)	5.4×10^{-36}			
27	<i>Attractor</i>	Unbounded	0.22	4.5×10^5	5.0×10^4	54.52	(53.61, 55.44)	0.0	5.77	(5.40, 6.14)	4.4×10^{-201}			
28		Equilibrium	0.01	2398.9	223	-85.70	(-197.9, 26.46)	0.13	38.02	(24.28, 51.75)	1.3×10^{-7}			
29		Cycle	0.19	38.42	5	-74.78	(-361.9, 212.3)	0.47	13.20	(-5.35, 31.76)	0.11			
30	Input	Density	<i>Overall</i>			0.12	4.6×10^5	5.0×10^4	71.77	(70.05, 73.49)	0.0	4.35	(3.83, 4.87)	1.6×10^{-59}
31			<i>Intercept</i>			0.0	4.7×10^5	5.0×10^4	—	—	—	24.10	(23.86, 24.34)	0.0
32		<i>Topology</i>	WS0.001			0.12	9.3×10^4	9943	74.50	(70.58, 78.41)	2.6×10^{-285}	3.99	(2.80, 5.18)	5.6×10^{-11}
33			ER			0.11	9.3×10^4	9959	69.95	(66.10, 73.81)	3.6×10^{-261}	4.82	(3.65, 5.99)	8.8×10^{-16}
34			RG			0.11	9.3×10^4	9970	71.38	(67.48, 75.28)	1.6×10^{-265}	4.50	(3.32, 5.69)	1.0×10^{-13}
35			WS0.1			0.12	9.3×10^4	9959	73.92	(70.05, 77.79)	1.8×10^{-287}	3.71	(2.53, 4.88)	6.8×10^{-10}
36			WS0.01			0.12	9.2×10^4	9944	69.09	(65.39, 72.78)	5.3×10^{-276}	4.73	(3.61, 5.86)	1.8×10^{-16}
37	<i>Attractor</i>	Unbounded				0.12	4.6×10^5	5.0×10^4	71.36	(69.64, 73.07)	0.0	4.42	(3.89, 4.94)	1.6×10^{-61}
38		Equilibrium				0.27	2331.8	223	195.9	(152.9, 238.9)	1.2×10^{-16}	-9.89	(-20.28, 0.49)	0.06
39		Cycle				5.0×10^{-4}	39.45	5	1.13	(-91.91, 94.17)	0.97	9.72	(-17.61, 37.05)	0.34
40	Interaction	In-	<i>Overall</i>			0.11	4.6×10^5	5.0×10^4	16.98	(16.54, 17.41)	0.0	5.36	(4.84, 5.89)	1.8×10^{-87}
41			<i>Intercept</i>			0.0	4.7×10^5	5.0×10^4	—	—	—	24.10	(23.86, 24.34)	0.0
42		<i>Topology</i>	WS0.001			0.09	9.3×10^4	9943	16.28	(15.28, 17.28)	7.1×10^{-214}	6.51	(5.30, 7.73)	1.2×10^{-25}
43			ER			0.11	9.3×10^4	9959	17.00	(16.03, 17.97)	5.7×10^{-245}	5.32	(4.14, 6.50)	1.3×10^{-18}
44			RG			0.10	9.3×10^4	9970	16.65	(15.67, 17.64)	7.3×10^{-229}	5.79	(4.59, 6.99)	3.6×10^{-21}
45			WS0.1			0.10	9.3×10^4	9959	16.68	(15.70, 17.66)	1.9×10^{-230}	5.63	(4.43, 6.83)	3.8×10^{-20}
46			WS0.01			0.13	9.2×10^4	9944	18.27	(17.35, 19.19)	6.5×10^{-309}	3.57	(2.45, 4.69)	4.7×10^{-10}
47	<i>Attractor</i>	Unbounded				0.11	4.6×10^5	5.0×10^4	17.24	(16.81, 17.67)	0.0	4.99	(4.46, 5.52)	5.1×10^{-76}
48		Equilibrium				6.0×10^{-4}	2401.1	223	-11.86	(-76.06, 52.34)	0.72	32.35	(12.64, 52.05)	1.4×10^{-3}
49		Cycle				0.79	31.57	5	37.17	(2.27, 72.06)	0.04	-9.59	(-29.27, 10.09)	0.22

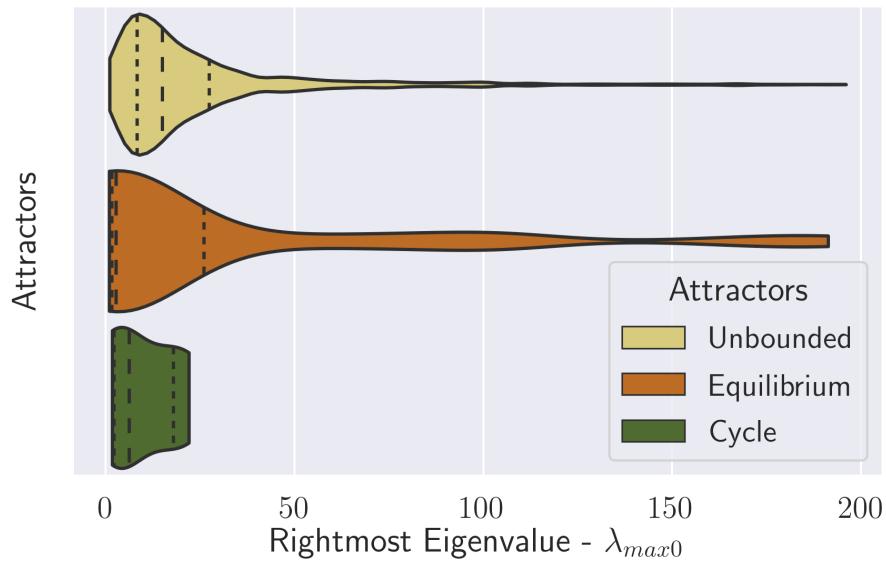


Figure B.8: Distribution of rightmost eigenvalues across topologies: The rightmost eigenvalue is measured for the initial community. The distributions of the proportion of survival were similar across topologies. The 'violin' shape indicates the distribution of values.

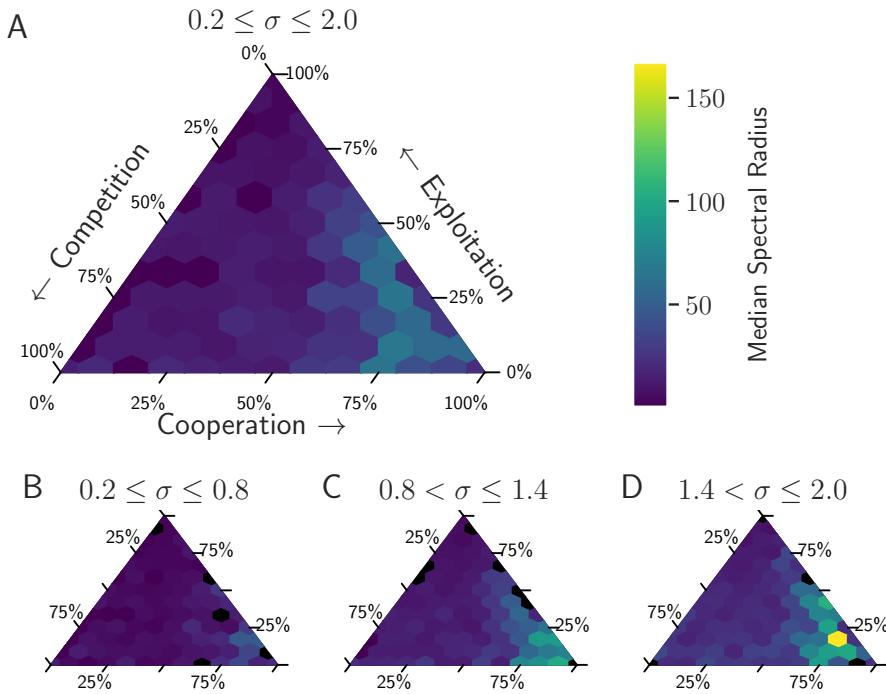


Figure B.9: The rightmost eigenvalue distributed across interaction type and intensity of interaction The rightmost eigenvalue is measured for the initial community. The median is taken from all values inside each hexagonal area, and brighter values indicate larger rightmost eigenvalue and lower resilience. The intensity of interaction is measured through the standard deviation of the normal distribution of interaction intensities. Regions in which no scenario was simulated are colored black.

Bibliography

- Allesina, Stefano (2020). *A Tour of the Generalized Lotka-Volterra Model*.
- Allesina, Stefano and Si Tang (Mar. 2012). “Stability Criteria for Complex Ecosystems”. en. In: *Nature* 483.7388, pp. 205–208. ISSN: 1476-4687. DOI: [10.1038/nature10832](https://doi.org/10.1038/nature10832).
- Allesina, Stefano et al. (July 2015). “Predicting the Stability of Large Structured Food Webs”. en. In: *Nature Communications* 6.1, pp. 1–6. ISSN: 2041-1723. DOI: [10.1038/ncomms8842](https://doi.org/10.1038/ncomms8842).
- Almeida-Neto, Mário et al. (2008). “A Consistent Metric for Nestedness Analysis in Ecological Systems: Reconciling Concept and Measurement”. en. In: *Oikos* 117.8, pp. 1227–1239. ISSN: 1600-0706. DOI: [10.1111/j.0030-1299.2008.16644.x](https://doi.org/10.1111/j.0030-1299.2008.16644.x).
- Arneodo, A, P Coullet, and C Tresser (Oct. 1980). “Occurrence of Strange Attractors in Three-Dimensional Volterra Equations”. en. In: *Physics Letters A* 79.4, pp. 259–263. ISSN: 0375-9601. DOI: [10.1016/0375-9601\(80\)90342-4](https://doi.org/10.1016/0375-9601(80)90342-4).
- Barabás, György, Matthew J. Michalska-Smith, and Stefano Allesina (July 2016). “The Effect of Intra- and Interspecific Competition on Coexistence in Multispecies Communities”. In: *The American Naturalist* 188.1, E1–E12. ISSN: 0003-0147. DOI: [10.1086/686901](https://doi.org/10.1086/686901).
- Barabási, Albert-László and Réka Albert (Oct. 1999). “Emergence of Scaling in Random Networks”. en. In: *Science* 286.5439, pp. 509–512. ISSN: 0036-8075, 1095-9203. DOI: [10.1126/science.286.5439.509](https://doi.org/10.1126/science.286.5439.509).
- Bascompte, Jordi et al. (Aug. 2003). “The Nested Assembly of Plant–Animal Mutualistic Networks”. en. In: *Proceedings of the National Academy of Sciences* 100.16, pp. 9383–9387. ISSN: 0027-8424, 1091-6490. DOI: [10.1073/pnas.1633576100](https://doi.org/10.1073/pnas.1633576100).
- Bastolla, Ugo et al. (Apr. 2009). “The Architecture of Mutualistic Networks Minimizes Competition and Increases Biodiversity”. en. In: *Nature* 458.7241, pp. 1018–1020. ISSN: 1476-4687. DOI: [10.1038/nature07950](https://doi.org/10.1038/nature07950).
- Berry, David and Stefanie Widder (2014). “Deciphering Microbial Interactions and Detecting Keystone Species with Co-Occurrence Networks”. English. In: *Frontiers in Microbiology* 5. ISSN: 1664-302X. DOI: [10.3389/fmicb.2014.00219](https://doi.org/10.3389/fmicb.2014.00219).
- Burgos, Enrique et al. (Nov. 2007). “Why Nestedness in Mutualistic Networks?” en. In: *Journal of Theoretical Biology* 249.2, pp. 307–313. ISSN: 0022-5193. DOI: [10.1016/j.jtbi.2007.07.030](https://doi.org/10.1016/j.jtbi.2007.07.030).
- Carnell, Robert (July 2020). *Using Latin Hypercube Sampling with a Condition That the Sum of Two Variables Should Be Less than One*. English. Cross Validated. Answer. eprint: <https://stats.stackexchange.com/q/476433>.
- Coyte, Katharine Z., Jonas Schlüter, and Kevin R. Foster (Nov. 2015). “The Ecology of the Microbiome: Networks, Competition, and Stability”. en. In: *Science* 350.6261, pp. 663–666. ISSN: 0036-8075, 1095-9203. DOI: [10.1126/science.aad2602](https://doi.org/10.1126/science.aad2602).
- Cross, G. W. (June 1978). “Three Types of Matrix Stability”. In: *Linear Algebra and its Applications* 20.3, pp. 253–263. ISSN: 0024-3795. DOI: [10.1016/0024-3795\(78\)90021-6](https://doi.org/10.1016/0024-3795(78)90021-6).
- Dirzo, Rodolfo et al. (July 2014). “Defaunation in the Anthropocene”. en. In: *Science* 345.6195, pp. 401–406. ISSN: 0036-8075, 1095-9203. DOI: [10.1126/science.1251817](https://doi.org/10.1126/science.1251817).
- Dormann, Carsten F., Jochen Fründ, and H. Martin Schaefer (2017). “Identifying Causes of Patterns in Ecological Networks: Opportunities and Limitations”. In: *Annual Review of Ecology, Evolution, and Systematics* 48.1, pp. 559–584. DOI: [10.1146/annurev-ecolsys-110316-022928](https://doi.org/10.1146/annurev-ecolsys-110316-022928).
- Dunne, Jennifer A. (2006). “The Network Structure of Food Webs”. In: *Ecological Networks: Linking Structure to Dynamics in Food Webs*. Ed. by Mercedes Pascual and Jennifer A. Dunne. Santa Fe Institute Studies in the Sciences of Complexity. New York: Oxford University Press, pp. 27–86. ISBN: 978-0-19-518816-5.

- Dunne, Jennifer A., Richard J. Williams, and Neo D. Martinez (Oct. 2002a). "Food-Web Structure and Network Theory: The Role of Connectance and Size". en. In: *Proceedings of the National Academy of Sciences* 99.20, pp. 12917–12922. ISSN: 0027-8424, 1091-6490. DOI: [10.1073/pnas.192407699](https://doi.org/10.1073/pnas.192407699).
- (July 2002b). "Network Structure and Biodiversity Loss in Food Webs: Robustness Increases with Connectance". In: *Ecology Letters* 5.4, pp. 558–567. ISSN: 1461-023X. DOI: [10.1046/j.1461-0248.2002.00354.x](https://doi.org/10.1046/j.1461-0248.2002.00354.x).
- Dyurgerov, Mark B. and Mark F. Meier (Feb. 2000). "Twentieth Century Climate Change: Evidence from Small Glaciers". en. In: *Proceedings of the National Academy of Sciences* 97.4, pp. 1406–1411. ISSN: 0027-8424, 1091-6490. DOI: [10.1073/pnas.97.4.1406](https://doi.org/10.1073/pnas.97.4.1406).
- Erdős, P. and A. Rényi (1959). "On Random Graphs I". In: *Publ. math. debrecen* 6.290-297, p. 18.
- Faust, Karoline and Jeroen Raes (Aug. 2012). "Microbial Interactions: From Networks to Models". en. In: *Nature Reviews Microbiology* 10.8, pp. 538–550. ISSN: 1740-1526, 1740-1534. DOI: [10.1038/nrmicro2832](https://doi.org/10.1038/nrmicro2832).
- Fiedler-Ferrara, Nelson and Carmen P. Cintra do Prado (Jan. 1994). *Caos: uma Introdução*. Português. 1ª Edição. Blucher. ISBN: 978-85-212-0058-1.
- Fortuna, Miguel A. et al. (2010). "Nestedness versus Modularity in Ecological Networks: Two Sides of the Same Coin?" en. In: *Journal of Animal Ecology* 79.4, pp. 811–817. ISSN: 1365-2656. DOI: [10.1111/j.1365-2656.2010.01688.x](https://doi.org/10.1111/j.1365-2656.2010.01688.x).
- Foster, Kevin R. and Thomas Bell (Oct. 2012). "Competition, Not Cooperation, Dominates Interactions among Culturable Microbial Species". In: *Current Biology* 22.19, pp. 1845–1850. ISSN: 0960-9822. DOI: [10.1016/j.cub.2012.08.005](https://doi.org/10.1016/j.cub.2012.08.005).
- Fowler, Mike S. (May 2009). "Increasing Community Size and Connectance Can Increase Stability in Competitive Communities". en. In: *Journal of Theoretical Biology* 258.2, pp. 179–188. ISSN: 0022-5193. DOI: [10.1016/j.jtbi.2009.01.010](https://doi.org/10.1016/j.jtbi.2009.01.010).
- Frigyik, Bela A., Amol Kapila, and Maya R. Gupta (2010). *Introduction to the Dirichlet Distribution and Related Processes*. English. Tech. rep. UWEETR-2010-0006. Washington: Department of Electrical Engineering, University of Washignton, pp. 1–27.
- Gibbs, Theo et al. (Aug. 2018). "Effect of Population Abundances on the Stability of Large Random Ecosystems". In: *Physical Review E* 98.2, p. 022410. DOI: [10.1103/PhysRevE.98.022410](https://doi.org/10.1103/PhysRevE.98.022410).
- Goh, B. S. (Jan. 1977). "Global Stability in Many-Species Systems". In: *The American Naturalist* 111.977, pp. 135–143. ISSN: 0003-0147. DOI: [10.1086/283144](https://doi.org/10.1086/283144).
- Gonze, Didier et al. (Oct. 2017). "Multi-Stability and the Origin of Microbial Community Types". en. In: *The ISME Journal* 11.10, pp. 2159–2166. ISSN: 1751-7370. DOI: [10.1038/ismej.2017.60](https://doi.org/10.1038/ismej.2017.60).
- Gonze, Didier et al. (Aug. 2018). "Microbial Communities as Dynamical Systems". In: *Current Opinion in Microbiology*. Microbiota 44, pp. 41–49. ISSN: 1369-5274. DOI: [10.1016/j.mib.2018.07.004](https://doi.org/10.1016/j.mib.2018.07.004).
- Grilli, Jacopo, Tim Rogers, and Stefano Allesina (June 2016). "Modularity and Stability in Ecological Communities". en. In: *Nature Communications* 7.1, pp. 1–10. ISSN: 2041-1723. DOI: [10.1038/ncomms12031](https://doi.org/10.1038/ncomms12031).
- Guimarães, Paulo R. (2020). "The Structure of Ecological Networks Across Levels of Organization". In: *Annual Review of Ecology, Evolution, and Systematics* 51.1, null. DOI: [10.1146/annurev-ecolsys-012220-120819](https://doi.org/10.1146/annurev-ecolsys-012220-120819).
- Hagberg, Aric A., Daniel A. Schult, and Pieter J. Swart (2008). "Exploring Network Structure, Dynamics, and Function Using NetworkX". In: *Proceedings of the 7th Python in Science Conference*. Ed. by Gaël Varoquaux, Travis Vaught, and Jarrod Millman. Pasadena, CA USA, pp. 11–15.
- Harper, Marc et al. (Apr. 2019). *Marcharper/Python-Ternary: Version 1.0.6*. Zenodo. DOI: [10.5281/zenodo.2628066](https://doi.org/10.5281/zenodo.2628066).
- Harris, Charles R. et al. (Sept. 2020). "Array Programming with NumPy". en. In: *Nature* 585.7825, pp. 357–362. ISSN: 1476-4687. DOI: [10.1038/s41586-020-2649-2](https://doi.org/10.1038/s41586-020-2649-2).
- Hofbauer, Josef and Karl Sigmund (1998). *Evolutionary Games and Population Dynamics*. English. ISBN: 978-1-139-17317-9 978-0-521-62365-0 978-0-521-62570-8.
- Holling, C. S. (1959). "Some Characteristics of Simple Types of Predation and Parasitism". In: *The Canadian Entomologist* 91.7, pp. 385–398. ISSN: 0008-347X. DOI: [10.4039/Ent91385-7](https://doi.org/10.4039/Ent91385-7).
- Hunter, J. D. (2007). "Matplotlib: A 2D Graphics Environment". In: *Computing in Science & Engineering* 9.3, pp. 90–95. DOI: [10.1109/MCSE.2007.55](https://doi.org/10.1109/MCSE.2007.55).

- Johnson, Jerald B. and Kristian S. Omland (2004). "Model Selection in Ecology and Evolution". In: *Trends in Ecology and Evolution* 19.2, pp. 101–108. ISSN: 01695347. DOI: [10.1016/j.tree.2003.10.013](https://doi.org/10.1016/j.tree.2003.10.013).
- Kondoh, Michio, Satoshi Kato, and Yoshikuni Sakato (2010). "Food Webs Are Built up with Nested Subwebs". en. In: *Ecology* 91.11, pp. 3123–3130. ISSN: 1939-9170. DOI: [10.1890/09-2219.1](https://doi.org/10.1890/09-2219.1).
- Kozlov, Vladimir and Sergey Vakulenko (July 2013). "On Chaos in Lotka–Volterra Systems: An Analytical Approach". en. In: *Nonlinearity* 26.8, pp. 2299–2314. ISSN: 0951-7715. DOI: [10.1088/0951-7715/26/8/2299](https://doi.org/10.1088/0951-7715/26/8/2299).
- Krenak, Ailton (2019). *Ideias Para Adiar o Fim Do Mundo*. Editora Companhia das Letras.
- Li, Chenhao et al. (June 2016). "Predicting Microbial Interactions through Computational Approaches". In: *Methods. Pan-Omics Analysis of Biological Data* 102, pp. 12–19. ISSN: 1046-2023. DOI: [10.1016/j.ymeth.2016.02.019](https://doi.org/10.1016/j.ymeth.2016.02.019).
- Lotka, Alfred J. (July 1920). "Analytical Note on Certain Rhythmic Relations in Organic Systems". en. In: *Proceedings of the National Academy of Sciences* 6.7, pp. 410–415. ISSN: 0027-8424, 1091-6490. DOI: [10.1073/pnas.6.7.410](https://doi.org/10.1073/pnas.6.7.410).
- May, Robert M. (1972a). "Limit Cycles in Predator-Prey Communities". In: *Science* 177.4052, pp. 900–902. ISSN: 0036-8075.
- (Aug. 1972b). "Will a Large Complex System Be Stable?" en. In: *Nature* 238.5364, pp. 413–414. ISSN: 1476-4687. DOI: [10.1038/238413a0](https://doi.org/10.1038/238413a0).
- May, Robert M. and Warren J. Leonard (Sept. 1975). "Nonlinear Aspects of Competition Between Three Species". In: *SIAM Journal on Applied Mathematics* 29.2, pp. 243–253. ISSN: 0036-1399. DOI: [10.1137/0129022](https://doi.org/10.1137/0129022).
- Maynard, Daniel S., Zachary R. Miller, and Stefano Allesina (Jan. 2020). "Predicting Coexistence in Experimental Ecological Communities". en. In: *Nature Ecology & Evolution* 4.1, pp. 91–100. ISSN: 2397-334X. DOI: [10.1038/s41559-019-1059-z](https://doi.org/10.1038/s41559-019-1059-z).
- Maynard, Daniel S., Carlos A. Serván, and Stefano Allesina (2018). "Network Spandrels Reflect Ecological Assembly". en. In: *Ecology Letters* 21.3, pp. 324–334. ISSN: 1461-0248. DOI: [10.1111/ele.12912](https://doi.org/10.1111/ele.12912).
- McKinney, Wes (2010). "Data Structures for Statistical Computing in Python". In: *Proceedings of the 9th Python in Science Conference*. Ed. by Stéfan van der Walt and Jarrod Millman, pp. 56–61. DOI: [10.25080/Majora-92bf1922-00a](https://doi.org/10.25080/Majora-92bf1922-00a).
- Michalska-Smith, Matthew J. and Stefano Allesina (June 2019). "Telling Ecological Networks Apart by Their Structure: A Computational Challenge". en. In: *PLOS Computational Biology* 15.6, e1007076. ISSN: 1553-7358. DOI: [10.1371/journal.pcbi.1007076](https://doi.org/10.1371/journal.pcbi.1007076).
- Monteiro, Luiz Henrique Alves (Jan. 2011). *Sistemas Dinâmicos*. Português. 3^a Edição. São Paulo (SP): Livraria da Física. ISBN: 978-85-7861-102-6.
- Mougi, Akihiko and Michio Kondoh (July 2012). "Diversity of Interaction Types and Ecological Community Stability". en. In: *Science* 337.6092, pp. 349–351. ISSN: 0036-8075, 1095-9203. DOI: [10.1126/science.1220529](https://doi.org/10.1126/science.1220529).
- (Nov. 2014). "Stability of Competition–Antagonism–Mutualism Hybrid Community and the Role of Community Network Structure". en. In: *Journal of Theoretical Biology* 360, pp. 54–58. ISSN: 0022-5193. DOI: [10.1016/j.jtbi.2014.06.030](https://doi.org/10.1016/j.jtbi.2014.06.030).
- Murty, Katta G. (Jan. 1972). "On the Number of Solutions to the Complementarity Problem and Spanning Properties of Complementary Cones". en. In: *Linear Algebra and its Applications* 5.1, pp. 65–108. ISSN: 0024-3795. DOI: [10.1016/0024-3795\(72\)90019-5](https://doi.org/10.1016/0024-3795(72)90019-5).
- Newman, M. E. J. (Feb. 2003). "Mixing Patterns in Networks". In: *Physical Review E* 67.2, p. 026126. ISSN: 1063-651X, 1095-3787. DOI: [10.1103/PhysRevE.67.026126](https://doi.org/10.1103/PhysRevE.67.026126). arXiv: [cond-mat/0209450](https://arxiv.org/abs/cond-mat/0209450).
- Newman, Mark (July 2018a). *Networks*. Inglês. 2^a Edição. Oxford, United Kingdom ; New York, NY, United States of America: OUP Oxford. ISBN: 978-0-19-880509-0.
- (July 2018b). *Networks*. en. Oxford University Press. ISBN: 978-0-19-252749-3.
- Noldus, Rogier and Piet Van Mieghem (Dec. 2015). "Assortativity in Complex Networks". en. In: *Journal of Complex Networks* 3.4, pp. 507–542. ISSN: 2051-1310, 2051-1329. DOI: [10.1093/comnet/cnv005](https://doi.org/10.1093/comnet/cnv005).
- Paine, R. T. (Jan. 1992). "Food-Web Analysis through Field Measurement of per Capita Interaction Strength". en. In: *Nature* 355.6355, pp. 73–75. ISSN: 1476-4687. DOI: [10.1038/355073a0](https://doi.org/10.1038/355073a0).

- Pascual, Mercedes and Jennifer A. Dunne (2006). "From Small to Large Ecological Networks in a Dynamic World". In: *Ecological Networks: Linking Structure to Dynamics in Food Webs*, pp. 3–24.
- Pinheiro, Rafael B. P. et al. (2019). "A New Model Explaining the Origin of Different Topologies in Interaction Networks". en. In: *Ecology* 100.9, e02796. ISSN: 1939-9170. DOI: [10.1002/ecy.2796](https://doi.org/10.1002/ecy.2796).
- Qian, Jimmy J. and Erol Akçay (Mar. 2020). "The Balance of Interaction Types Determines the Assembly and Stability of Ecological Communities". en. In: *Nature Ecology & Evolution* 4.3, pp. 356–365. ISSN: 2397-334X. DOI: [10.1038/s41559-020-1121-x](https://doi.org/10.1038/s41559-020-1121-x).
- Rohr, Rudolf P., Serguei Saavedra, and Jordi Bascompte (July 2014). "On the Structural Stability of Mutualistic Systems". en. In: *Science* 345.6195. ISSN: 0036-8075, 1095-9203. DOI: [10.1126/science.1253497](https://doi.org/10.1126/science.1253497).
- Seabold, Skipper and Josef Perktold (2010). "Statsmodels: Econometric and Statistical Modeling with Python". In: *9th Python in Science Conference*.
- Segar, Simon T. et al. (May 2020). "The Role of Evolution in Shaping Ecological Networks". en. In: *Trends in Ecology & Evolution* 35.5, pp. 454–466. ISSN: 0169-5347. DOI: [10.1016/j.tree.2020.01.004](https://doi.org/10.1016/j.tree.2020.01.004).
- Serván, Carlos A. et al. (Aug. 2018). "Coexistence of Many Species in Random Ecosystems". en. In: *Nature Ecology & Evolution* 2.8, pp. 1237–1242. ISSN: 2397-334X. DOI: [10.1038/s41559-018-0603-6](https://doi.org/10.1038/s41559-018-0603-6).
- Shang, Yu et al. (Mar. 2017). "Inferring Interactions in Complex Microbial Communities from Nucleotide Sequence Data and Environmental Parameters". en. In: *PLOS ONE* 12.3, e0173765. ISSN: 1932-6203. DOI: [10.1371/journal.pone.0173765](https://doi.org/10.1371/journal.pone.0173765).
- Sjögren, Rickard and Daniel Svensson (Jan. 2020). *pyDOE2*.
- Smale, S. (Mar. 1976). "On the Differential Equations of Species in Competition". en. In: *Journal of Mathematical Biology* 3.1, pp. 5–7. ISSN: 1432-1416. DOI: [10.1007/BF00307854](https://doi.org/10.1007/BF00307854).
- Solé, Ricard V. and M. Montoya (Oct. 2001). "Complexity and Fragility in Ecological Networks". In: *Proceedings of the Royal Society of London. Series B: Biological Sciences* 268.1480, pp. 2039–2045. DOI: [10.1098/rspb.2001.1767](https://doi.org/10.1098/rspb.2001.1767).
- Song, Chuliang and Serguei Saavedra (Apr. 2020). "Telling Ecological Networks Apart by Their Structure: An Environment-Dependent Approach". en. In: *PLOS Computational Biology* 16.4, e1007787. ISSN: 1553-7358. DOI: [10.1371/journal.pcbi.1007787](https://doi.org/10.1371/journal.pcbi.1007787).
- Staniczenko, Phillip P. A., Jason C. Kopp, and Stefano Allesina (Jan. 2013). "The Ghost of Nest-edness in Ecological Networks". en. In: *Nature Communications* 4.1, pp. 1–6. ISSN: 2041-1723. DOI: [10.1038/ncomms2422](https://doi.org/10.1038/ncomms2422).
- Staudt, Christian L., Aleksejs Sazonovs, and Henning Meyerhenke (Dec. 2016). "NetworKit: A Tool Suite for Large-Scale Complex Network Analysis". en. In: *Network Science* 4.4, pp. 508–530. ISSN: 2050-1242, 2050-1250. DOI: [10.1017/nws.2016.20](https://doi.org/10.1017/nws.2016.20).
- Stein, Michael (May 1987). "Large Sample Properties of Simulations Using Latin Hypercube Sampling". In: *Technometrics* 29.2, pp. 143–151. ISSN: 0040-1706. DOI: [10.1080/00401706.1987.10488205](https://doi.org/10.1080/00401706.1987.10488205).
- Succurro, Antonella and Oliver Ebenhöh (Apr. 2018). "Review and Perspective on Mathematical Modeling of Microbial Ecosystems". en. In: *Biochemical Society Transactions* 46.2, pp. 403–412. ISSN: 0300-5127, 1470-8752. DOI: [10.1042/BST20170265](https://doi.org/10.1042/BST20170265).
- Takemoto, Kazuhiro and Midori Iida (Jan. 2019). "Ecological Networks". en. In: *Encyclopedia of Bioinformatics and Computational Biology*. Ed. by Shoba Ranganathan et al. Oxford: Academic Press, pp. 1131–1141. ISBN: 978-0-12-811432-2. DOI: [10.1016/B978-0-12-809633-8.20203-3](https://doi.org/10.1016/B978-0-12-809633-8.20203-3).
- Takeuchi, Yoshito (1996). *Global Dynamical Properties of Lotka-Volterra Systems*. English. Singapore: World scientific. ISBN: 978-981-02-2471-4.
- Thébault, Elisa and Colin Fontaine (Aug. 2010). "Stability of Ecological Communities and the Architecture of Mutualistic and Trophic Networks". en. In: *Science* 329.5993, pp. 853–856. ISSN: 0036-8075, 1095-9203. DOI: [10.1126/science.1188321](https://doi.org/10.1126/science.1188321).
- Thompson, J. M. T. and H. B. Stewart (Apr. 2002). *Nonlinear Dynamics and Chaos*. eng. Wiley. ISBN: 978-0-471-87645-8.
- Virtanen, Pauli et al. (Mar. 2020). "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python". en. In: *Nature Methods* 17.3, pp. 261–272. ISSN: 1548-7105. DOI: [10.1038/s41592-019-0686-2](https://doi.org/10.1038/s41592-019-0686-2).
- Volterra, Vito (Oct. 1926). "Fluctuations in the Abundance of a Species Considered Mathematically¹". en. In: *Nature* 118, pp. 558–560. ISSN: 1476-4687. DOI: [10.1038/118558a0](https://doi.org/10.1038/118558a0).

- Waskom, Michael and the seaborn development team (Sept. 2020). *Mwaskom/Seaborn*. Zenodo. DOI: [10.5281/zenodo.592845](https://doi.org/10.5281/zenodo.592845).
- Watts, Duncan J. and Steven H. Strogatz (June 1998). “Collective Dynamics of ‘Small-World’ Networks”. en. In: *Nature* 393.6684, pp. 440–442. ISSN: 1476-4687. DOI: [10.1038/30918](https://doi.org/10.1038/30918).
- Wedgwood, Hensleigh (Aug. 2017). *Dictionary of English Etymology*. English. Andesite Press. ISBN: 978-1-375-61565-5.
- Wigner, Eugene P. (Jan. 1967). “Random Matrices in Physics”. In: *SIAM Review* 9.1, pp. 1–23. ISSN: 0036-1445. DOI: [10.1137/1009001](https://doi.org/10.1137/1009001).
- Witelski, Thomas and Mark Bowen (2015). *Methods of Mathematical Modelling: Continuous Systems and Differential Equations*. eng. Springer Undergraduate Mathematics Series. Cham Heidelberg New York Dordrecht London: Springer. ISBN: 978-3-319-23042-9 978-3-319-23041-2.
- Xiao, Yandong et al. (Dec. 2017). “Mapping the Ecological Networks of Microbial Communities”. en. In: *Nature Communications* 8.1, pp. 1–12. ISSN: 2041-1723. DOI: [10.1038/s41467-017-02090-2](https://doi.org/10.1038/s41467-017-02090-2).