Inequalities in Cooperative Generalized Lotka-Volterra model with Random Regular Graph Structure

by

Alessandro Zuliani

to obtain the degree of Master of Science in Applied Mathematics
at the Delft University of Technology,

Student number:	5852870	
Project duration:	February 1, 2024 – September 18, 2024	
Thesis committee:	Dr. A. Barbaro,	TU Delft, responsible supervisor
	Dr. F. A. Lopez,	EconophysiX Chair, external daily supervisor
	Dr. C. A. Urzúa-Torres,	TU Delft, responsible professor

Abstract

The Generalized Lotka-Volterra (GLV) model, widely employed in population dynamics, serves to characterize how interactions between species influence system equilibria. This model can be extended to the economic domain to quantify wealth inequality among individuals. In this thesis, we explore the impact of the underlying random network structure on the final distribution of abundances, with a focus on homogeneous degree settings, particularly regular random graphs. While the GLV model has been extensively studied in competitive and predator-prey scenarios, the mutualistic (or cooperative) case has received less attention due to the divergence of species abundances. In the context of this divergence, we investigate the hierarchy of infinities by revisiting the GLV model and introducing a framework wherein species undergo relative extinction when compared to others. For large interaction strengths, we observe the emergence of effective competition within a cooperative system, resulting in localization, a phenomenon in which a few highly connected species dominate and accumulate most of the wealth. As interaction strength decreases, the surviving component follows a geometrical evolution, which we analytically quantify in this work. Finally, we discuss the implications of this framework for wealth inequality through the Gini index and extend the analysis to heterogeneous degree settings using the Erdős-Rényi random graph model.

Keywords: Generalized Lotka-Volterra, mutualistic divergence, regular random graph, Erdos-Renyi random graph, replicator equations, wealth condensation, Anderson localization.

Acknowledgements

I would like to express my deep gratitude to the EconophysiX chair and Capital Fund Management (CFM) for providing me with the opportunity to conduct my thesis research within their institution. Specifically, I would like to thank J.P. Bouchaud, M. Benzaquen, F.A. Lopez and R. Zakine for fostering an intellectually stimulating and high-quality research environment characterized by a collaborative exchange of knowledge. I feel that this experience has likely offered me a valuable chance to broaden my knowledge in different directions of macro-economy, quantitative finance and statistical physics. This is due to them and due to the guest speakers who delivered seminars at CFM. Their insights proved to be highly enriching. G. Bunin's inspiring motivation for pursuing research particularly impressed me: he introduced his work by saying "I am only interested in understanding why life exists where it shouldn't" and that added to those topics an underlying beauty that I have cherished during all my thesis period.
Additionally, I consider myself fortunate to have shared the laboratory with such skilled and passionate researchers as Alessandro, Jutta, Elia, Cécilia, Salma, Victor, Patil, Natasha, Max, Max, Tomas, Guillaume, Antoine-Cyrus, Swann, Thomas, and Pierre. My deepest gratitude goes to Fabian, my supervisor, whose professionalism and availability have been invaluable. I especially admire his foresight when approaching new problems, which stems from a profound understanding of the subjects we explored. Beyond the theoretical insights, Fabian taught me a solid research methodology that starts from the attention to posing the right research questions. I hope we will have new discussions about random graphs and population dynamics at Le Temps des Cerises.

I would also like to express my gratitude to Alethea Barbaro, my supervisor at TU Delft, who regularly provided me with valuable advice during each of our meetings. She has the ability to make the most challenging moments seem manageable and I hope our collaboration will continue to be just as harmonious in the future. I finally thank Carolina Urzúa Torres for her time and patience in reviewing this thesis.

I conclude here the official acknowledgements and continue in Italian to express a few thoughts on the people in Italy who have been supporting and inspiring me in these uni-

versity years. Vorrei iniziare con Alberto, a cui voglio un bene indescrivibile, che mi ha accompagnato durante tutte le fasi della vita, con i pranzi dopo scuola insieme a Franca e le versioni di latino. E' sempre stato un esempio da seguire e nonostante mi senta suo studente da quando sono nato, percepisco che c'è ancora moltissimo da imparare da lui. Sono sicuro che condivida questo pensiero perchè la qualità che più ammiro di lui è il continuo desiderio di approfondire nuovi argomenti. Mi auguro di vivere sempre secondo il suo ideale di capire per condividere. Grazie a Stefano, che è la prima persona di cui chiedo il consiglio e che mi ha insegnato il valore di una vita alla ricerca della bellezza e della diversità. Grazie a Marilisa, alla quale forse dovrei dire queste cose più spesso, ma vorrei cogliere questa occasione per esprimere tutta la stima che ho nei suoi confronti. Mi auguro in futuro di avere la capacità che ha lei di gestire situazioni difficili facendolo sembrare così facile. Entrambi mi hanno sempre fatto percepire un raro senso di leggerezza nella vita, soprattutto nei confronti della libertà di scelta, e mi riprometto di ripagare negli anni che verranno questa loro fiducia. Continuo pensando a Silvia, che sono sicuro farà meglio di me all'università, con la quale in questi anni vorrei creare un rapporto di spontanea complicità. Ringrazio Nice per essere stata la dimostrazione di cosa sia l'amore incondizionato per la famiglia, che in questi anni universitari ho un po' trascurato, ma vorrei relativamente restaurare. Mi rivolgo a Francesco, con cui mi sento estremamente affine e legato. Mi fa sorridere pensare alla pazienza che avevi quando mi spiegavi la fisica al liceo. E' bello stare accanto ad una persona altruista e appassionata come te. Infine vorrei rivolgermi ad Ernesto e Paolo per augurar loro il meglio e dir loro di vivere ogni momento con curiosità ed entusiasmo.

Concludo con un ringraziamento speciale agli amici. Ne nominerò solo alcuni ma ognuno di coloro che non è menzionato sa del bene che gli riservo. Lorenzo e Alessio, è difficile far meglio di Ale a riguardo, ma sono veramente felice che si sia creato questo gruppo anni fa e penso che non ci sia distanza geografica che possa farci passare la voglia di prendere treni, bus e aerei per vederci. Grandi cose aspettano entrambi e proprio non saprei predire il futuro per noi tre, ma la certezza è che non sarà una novità stravolgente perchè so che il primo pensiero di ogni pausa sarà sempre quello di chiamarvi e sentire che avete da raccontare. Grazie a Luca per essere stato il più grande amico, compagno di progetti e ultimamente anche coinquilino che potessi desiderare qui a Delft. Grazie a Fernando con cui ho condiviso Menno Ter Braaklaan, belle serate in Andalusia e qualche impulsività del lunedì. Vorrei menzionare anche i collegiali Fra, Nico, Chiara e Paolo, che mi sembra corrispondano ad un'altra vita ma con cui ho transcorso uno dei miei periodi più belli. Infine un pensiero per Saso con il quale ci sentiamo di rado, ma ci unisce un'amicizia indissolubile.

Delft, 18 September 2024
Alessandro Zuliani

Contents

List of Figures

Acronyms

GLV Generalized Lotka Volterra 2

DMFT Dynamic Mean Field Theory 2

ABM Agent Based Model 5

IC Initial Conditions 5

RMSE Root Mean Squared Error 58

MPA Message Passing Algorithm 3

ODE Ordinary Differential Equation 5

MIS Maximal Independent Set 9

1 Introduction

1.1 Background and context of the research

To explain the motivations behind this thesis, we will initially explore the roles of statistical physics and populations biology in Econophysics: the study of complex systems central to understanding the multifaceted nature of economic dynamics. A complex system, characterized by numerous interacting components that exhibit collective behavior, cannot be comprehensively described by simple, linear models. This complexity requires the use of statistical physics, a field adept at modeling systems composed of large numbers of particles, each obeying specific rules governing their dynamics and interactions. The application of these principles to economics, particularly after the financial crisis, highlighted the inadequacies of classical economic theories, which often failed to account for the emergent and nonlinear behaviors observed in real-world markets. *J.P. Bouchaud* in "The (Unfortunate) Complexity of the Economy" [11] underscores this shift, emphasizing how traditional models struggled to predict and manage the complexities revealed during economic turmoil. Econophysics treats wealth, income, and money distributions akin to particle distributions in physical systems. In [29] and [52], *T. Lux* and *V. M. Yovanenko* merged statistical mechanics principles with economic measures like the Gini index and the Lorenz curve (we will use them later) to model the heavy-tailed distributions observed in real economies, where wealth and income are unevenly distributed across the population. In the book *The economy as an evolving complex system* by *P. W. Anderson* [4], multiple physics phenomena that arise in economics are explained, such as self-reinforcing mechanisms that describe processes where initial advantages tend to accumulate over time, leading to increasing returns and potential market dominance by particular firms. Or the influence of nonlinear dynamics that occur when changes in input do not result in proportional changes in output, leading to phenomena such as bifurcations, where a small change in initial conditions can cause a sudden shift from one equilibrium to another [14]. In the context of economic development, dynamics away from attractors provide a framework for understanding how economies transition between different states of equilibrium. Traditional economic models often assume that economies tend toward a single, stable equilibrium. However, *M. Kauffman* and *D. Farmer*'s work suggests that economies can exhibit complex adaptive behaviors, moving away from traditional

attractors toward new configurations in response to shocks or innovations [22]. In this last paper the interactions between different market participants can be modeled as ecological interactions, where profits and losses correspond to the survival and reproduction of species. There, *D. Farmer* explores how reinvestment of profits within markets leads to a capital allocation model analogous to population dynamics, where trend followers and value investors play roles similar to different species within an ecosystem. This approach also demonstrates how market efficiency is influenced by the behaviors of these participants, with trend followers often impeding the progress toward market efficiency, much like invasive species might disrupt ecological balance. Among the models from population biology, which describe the interactions between different economic agents as analogous to predator-prey dynamics, the Generalized Lotka Volterra (GLV) is the most versatile. For instance, *R. Goodwin* applied it to provide insights into the cyclical nature of economic growth and income distribution, highlighting how competition and cooperation between different economic actors can lead to emergent behaviors in markets. This approach, combined with the insights in *J. Moran* and *J.P. Bouchaud*'s work on instability in large economies, underscores the inherent fragility of economic systems as they become larger and more interconnected [39].

Motivated by these considerations, we will investigate the GLV, focusing on the influence of the underlying network to the final state of the system. To do that we will introduce some concept from random graph theory for which the reader can refer to the book by *R. V. Hofstad* [51]. The role of the underlying graph was studied both from an economical point of view [7] and, more extensively, from an ecological point of view [2], [30], [15]. Recently, three notable studies were published almost simultaneously, focusing on the effects of heterogeneous interactions on the equilibria of systems governed by GLV dynamics [47], [42], [1]. Of particular relevance is the study by *F. A. Lopez*, which serves as a primary inspiration for our analysis. A common thread among these studies is the assumption of the *high connectivity limit*, denoted as $1 \ll C \ll N$, where N represents the population size and C the average degree of connectivity, leading to dense adjacency matrices. This assumption is crucial as it facilitates the application of Dynamic Mean Field Theory (DMFT), a powerful tool that simplifies the analysis of such complex systems. DMFT enables the reduction of a high-dimensional problem, wherein the dynamics of each species are influenced by interactions with many others, to a more tractable one-dimensional stochastic process. In essence, the mean field approach approximates the intricate, individual interactions within the system by an average effect that each species experiences from the collective influence of the entire population.

An alternative approach to studying the stability of equilibria in the context of underlying random networks was pioneered by *R. M. May* [31]. This approach involves determining the spectral boundary of the Jacobian matrix, heavily relying on the principles of random matrix theory. Specifically, the stability of an equilibrium is assessed by evaluating the eigenvalues of the Jacobian at that equilibrium; if all eigenvalues are negative, the equilibrium is considered

a stable attractor. In its most basic form, this approach reduces to analyzing the maximum eigenvalue. However, the challenge arises in defining the spectrum of an interaction matrix, particularly when its entries are random variables following a specific distribution [43]. While this spectral approach has the advantage of being applicable to systems with finite N and sparse interaction matrices with low connectivity, it is often complex to manage and generalize due to its case dependent nature.

In our case we will mainly deal with homogeneous degree random graphs which lead to deterministic equilibria. The goal is then to find those equilibria (or fixed points) and find the Jacobian spectrum boundary. In our framework, given the regularity of the graph, we could always expect the same final structures. Therefore, what we did was to run the simulations, observing the structure of the equilibria and building the theory on top of it. In general, for more complex networks, Message Passing Algorithm (MPA) can be used to find (or approximate) the fixed points [35], [20]. It works by iteratively updating the messages based on information from neighboring nodes which are propagated throughout the network. This propagation continues until the messages converge to stable values, representing the fixed points of the system. It is particularly well-suited to sparse matrices where the number of non-zero elements is relatively small compared to the size of the matrix, and the algorithm efficiently handles the propagation of information through the network. Moreover, thanks to the Schur Complement formula which simplifies the matrix by iteratively eliminating rows, the determinant of the Jacobian evaluated in the fixed point found with the MPA can be factorized. In most cases it is an approximation, but in the case of trees (regular and non-regular) the factorization of the characteristic equation is exact as there are no cycles or multi-edges. This algorithm proved to be a very useful tool also in information theory and compressed sensing because it drastically reduces the dimension of a complex system. For what concerns the topics of information theory that we will use in this research, like the Shannon Entropy and the Fractal Exponent, the reader can refer to the book *Information, Physics and Computation* by *A. Montanari* and *M. Mezard* [35].

I will now anticipate the explanations of two results that we will observe later. In particular there are two phenomena that belong to information theory, whose origins come from statistical mechanics. In order to achieve that, I have to introduce the underlying process of my research. We will observe the dynamics of a system of N species that are connected via an underlying random graph. It is a one dimensional problem because the only variable that changes in time is the abundance of each species. Finally the dynamics converge to a stable attractor (or fixed point or equilibrium) and I have a final state given by the N-dimensional vector with the final abundances of each species. In general, a state at time t is the vector of abundances in that instant. After this incipit, we can delve into the description of the two phenomena.

In statistical mechanics, *condensation* describes the phenomenon where a large fraction of particles occupies a single quantum state, as in Bose-Einstein condensation [23]. In the context of the GLV dynamics, in a condensed final state, the total abundance of the system is dominated by a few species whose abundances are large (of $O(N)$) and whose normalization against the total sum of abundances does not drop to zero when the total number of species N is brought to infinity. We then have a highly skewed abundances distribution where the majority of species might have negligible or zero abundance (we say that they go extinct), analogous to the way particles in a Bose-Einstein condensate occupy a single quantum state. This phenomenon is studied also with respect to the wealth of individuals in [12], [52] where it is measured with the *inverse participation ratio*.

Anderson localization refers to the phenomenon where electrons become localized in space due to disorder in the system, hindering their ability to move across a lattice. When the disorder in the system is sufficiently strong, the wave functions of the particles become exponentially localized, preventing them from contributing to transport phenomena like electrical conduction. In the paper by *V. E. Kravtsov* [28], which we will take as a term of comparison for our results, the Anderson localization is detected by showing the fractal dimension D of wave functions with respect to the disorder strength W (see Figure 4.24). If $D = 1$ corresponds to the regime of ergodic extended states and $D = 0$ is the one for non-ergodic localized states where particles are confined to specific regions, then D undergoes a smooth transition from 1 to 0. Simply think of an ergodic regime as one where no single species or small group of species dominates. While in the non-ergodic regime, the system's state is "localized" in a subset of the possible configurations. In this case, the Anderson localization is anticipated by the non-ergodic extended phase with $0 < D < 1$ where the wave functions are extended but still localized in a fractal-like manner. In the context of the final state in the GLV, the fractal exponent (or fractal dimension) describes the spatial distribution of species abundances at equilibrium. With spatial distribution I mean the clusters of the distribution, i.e. regions where the abundance values are dense, separated by regions where abundances are sparse or zero. For example, if there are M species with final abundances all at $y = 100$, and N–M species that go extinct with abundances around 0, then I have only two clusters and no self-similarity in the first one. Let us see what it means. If you zoom in on a portion of this distribution, you might find that within a cluster, the distribution still has irregularities, or if you look at the distribution on different scales, similar patterns might appear. Fractal dimensions are indeed a mathematical tool used to describe the complexity of structures that exhibit patterns repeating at different scales. This is *self similarity* and it means that the structure looks similar at different scales. In the GLV framework, this could manifest as similar patterns of clustering at different levels of abundance. In the GLV, in a localized regime, many species might go extinct because the interaction strengths lead to highly competitive dynamics where only a few species can sustain themselves. Indeed the disorder strength W of *Kravtsov*'s paper corresponds to our interaction strength. This is

analogous to the wave function in a quantum system being concentrated in a small region, with most of the space having near-zero amplitude [8], [16]. We will observe this localization in the case of a C-regular trees which is a finite size simplification of an infinite size *Bethe lattice* where the Anderson localization was observed in [41], [46].

One last consideration before delving into our research and the first description of the model we are going to use. In this study we cite many research projects on Agent Based Model (ABM), it is then worth mentioning whether the GLV is an ABM and whether we can get some insights that could be useful for the ABM scope [27], [44], [12].
A GLV model with N species and an ABM with N agents share the common goal of modeling complex dynamics resulting from interactions within a system. However, the GLV model represents these interactions through continuous, deterministic differential equations that describe the overall population dynamics in a more aggregated form, focusing on species-level behaviors. In contrast, an ABM captures the individual actions and interactions of discrete agents, often incorporating stochastic elements that reflect variability and randomness in agent behaviors. The former provides a macroscopic view while the latter offers a microscopic perspective. We know from [40] that the ABMs are very sensible to the choice of the initial configuration, hence understanding how the choice of Initial Conditions (IC) influences the equilibria of the GLV could be an insightful starting point for a better comprehension of ABMs.

1.2 Generalized Lotka-Volterra

Before we state the problem and we delve into the results, we present the dynamic model with a detailed analysis of the roles of the different terms.
The Generalized Lotka-Volterra model consists of a system of N Ordinary Differential Equation (ODE) expressed as:

$$\dot{x}_i(t) = x_i(t)\left(1 - x_i(t) + \sum_j A_{ij}\alpha_{ij}x_j(t)\right), \quad i = 1, \dots, N. \tag{1.1}$$

where each term has a different role,

- $x_i(t)$: represents the abundance (or wealth when we consider the economical implications of the study) of species i at time t.
- $\dot{x}_i(t)$: denotes the rate of change of the abundance of species i with respect to time. It tells us how the species size is evolving at any given moment.
- 1: represents the intrinsic growth rate of species i in the absence of any competition or interaction.

- $-x_i(t)$: represents the self-limitation effect due to intraspecific competition. As the population $x_i(t)$ increases, this negative term becomes more significant, reflecting the carrying capacity limitation and the ageing of a certain species. It models the logistic growth of species i.

- Interaction strength parameter α: modulates the influence of the connections on the population dynamics. A larger $|\alpha|$ indicates stronger interactions among species.

- Adjacency matrix A: Matrix A quantifies the connection of species. In particular, $A_{ij} \in \{0, 1\}$ is 1 if the two nodes are connected. The matrix $A_{ij}\alpha_{ij}$ is called Interaction matrix instead.

For the rest of the study the terms *species, individual, node* will be interchanged but they all refer to i. We take the adjacency matrix A to be that of a random network sampled from a configuration model (see A), where the degree of each node, $k_i(A) = \sum_j A_{ij}$, is fixed to value k_i. The vector of degrees $\mathbf{k} = (k_1, ..., k_N)$ is sampled from a distribution $p(k)$.
We consider A to be a non-directed symmetric matrix without self-loops, which means $A_{ii} = 0 \quad \forall i \in \{1, ..., N\}$, while the number of ones on the row i is given by k_i.

Then we call C the average degree given by $C = \sum_k p(k)k$.

The other source of randomness is given by the interaction strength, where the asymmetry of interaction is introduced by taking the pairs $(\alpha_{ij}, \alpha_{ji})$ as i.i.d. random variables with

$$\alpha_{ij} = \frac{\mu}{C} + \frac{\sigma}{\sqrt{C}} z_{ij}.$$

where $z_{ij} \sim \mathcal{N}(0, 1)$, and $\text{cov}(z_{ij}, z_{ji}) = \gamma$. In this way the correlation between the weights is controlled by γ; this changes the proportion of the different type of interactions that can occur between species (cooperation, competition) [1]. We will always consider that the model starts from an initial condition $\mathbf{x}_0$ whose entries are random i.i.d. numbers sampled from a distribution $x_0 \sim \mathcal{U}(0, 1)$.

In this research, we are interested in analyzing the effect of the network structure on the equilibria of the system, so we get rid of the randomness generated by the interaction strength. This corresponds to setting $\sigma = 0$ and it means all interactions have the same value. Considering $\alpha = \frac{\mu}{C}$, we analyze the homogeneous interaction version of Equation 1.2 that is,

$$\dot{x}_i(t) = x_i(t) \left(1 - x_i(t) + \alpha \sum_j A_{ij} x_j(t) \right), \quad i = 1, \ldots, N. \tag{1.2}$$

The only disorder to consider is that of the random matrix A. In this case if we have $\alpha > 0$, it corresponds to cooperation between all species and if we have $\alpha < 0$ it corresponds to pure competition between everyone.

Vectorially Equation 1.2 can be expressed as,

$$\frac{d\mathbf{x}(t)}{dt} = D(\mathbf{x}(t))(\mathbf{1} - \mathbf{x}(t) + \alpha A\mathbf{x}(t)). \tag{1.3}$$

Where $D(\mathbf{x}(t))$ is the diagonal matrix with the vector of abundances in the diagonal. The system converges to one of the 2^N equilibria (or fixed points) that are obtained by putting

$$\frac{d\mathbf{x}}{dt} = \mathbf{0}. \tag{1.4}$$

Since we are dealing with species abundances (or firms wealth), the equilibrium must have N non-negative values. This condition is called *feasibility* of the equilibrium. We introduce also the other condition of *uninvadability* which refers to the impossibility of species' abundances to increase after they get to zero. This means that there is no resurrection possibility. If all species are surviving (meaning that they have positive final abundance) the second term must be zero and the equilibrium is

$$\mathbf{x}^\star = (-I + \alpha A)^{-1}\mathbf{1}. \tag{1.5}$$

Where we can operate the inversion only if the matrix is non-singular, so we need it to be full rank. As we explain in 2.1, finding the stable equilibria is reduced to a problem of finding the maximal eigenvalue of the Jacobian calculated in the fixed point. In the standard GLV case it is

$$J_{ij} = \begin{cases} 1 - 2x_i + \alpha \sum_k A_{ik}x_k & \text{if } i = j \\ \alpha x_i A_{ij} & \text{if } i \neq j \end{cases} \tag{1.6}$$

Then it has to be evaluated in $\mathbf{x}^\star = (x_1^*, ..., x_N^*)$ to find the maximal eigenvalue and check if $\mathbf{x}^\star$ is stable. With this analysis we can study which is the structure of the stable equilibrium for each value of α. This is what *G. Bunin et al.* did in [30], and the results obtained in that paper serve as the starting point for our investigation.

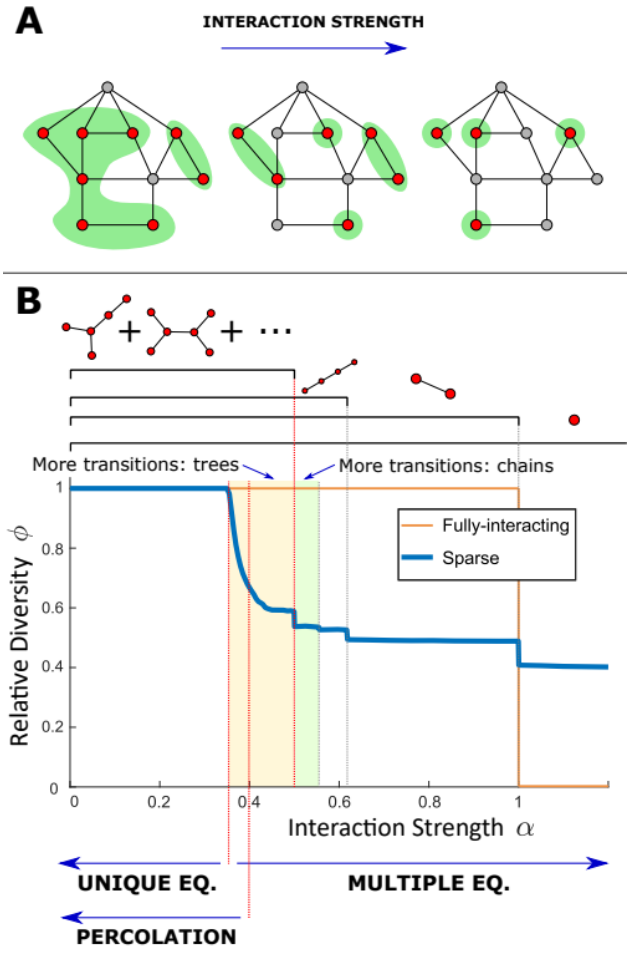


Figure 1.1: A) Evolution of the surviving components. Surviving nodes are highlighted in red. B) ϕ is the relative diversity. ϕ against α in the competitive case; differently from ours, in that paper competition corresponds to $\alpha > 0$. This curve is obtained by running the dynamics on a regular random graph with $N = 400$ and $C = 3$ and checking how many nodes survive at equilibrium. The transitions are of two genres: structural (above) and in the number of surviving species (below).

Here the *relative diversity* is introduced, and it is the measure that we will use the most during this research. It is the ratio of surviving species over the total amount of species. Looking at Figure 1.1.B, we notice that the curve undergoes multiple transitions of different genres. For the initial values of α the relative diversity is 1, meaning that all species survive with the same abundances expressed by equation 1.5. We call it uniform (or mean field) equilibrium. As α is increased and species start dying, multiple equilibria become stable and their structure goes from trees to chains to end up with size one disconnected components

(singlets). At $\alpha = 1$ the final plateau starts, where almost 40% survive as singlets. Figure 1.1.A shows the structure of the surviving components (green sets of red nodes) in relation with the initial network that includes also the grey dying nodes.

1.3 Statement of the problem

The first goal of this research is to characterize the final structures geometrically and to find what are the analytical values of α (or μ) where the transitions occur. We will do that by looking at the edges of the distribution of the eigenvalues of the Jacobian. We want to create a universal theory for C-regular random graphs for $\alpha \in \mathbb{R}$, including also the cooperative case with $\alpha > 0$. The extension on the whole real set for the interaction strength will not be perfectly symmetric with respect to the model used because in cooperation we will adopt a modified version of the standard GLV. Therefore it is not a perfect universal theory but it is a useful unified theory where the two cases can be critically compared. In the literature the purely competitive and the negatively correlated predator-pray dynamics scenarios are vastly studied [49][30], but the purely mutualistic case is not, because species diverge and there is only one possible state with all infinities.
Using the homogeneous degree case, for a given interaction strength and a given average degree we will be able to immediately predict to which stable equilibrium the system is going to relax and which is the final abundances distribution. That enables us to quantify the wealth inequality of the network and to see how that changes when we modify either the number or the strength of the connections.
Keep in mind that until we do not totally understand how the initial conditions influence the equilibria, telling us exactly which species are going to survive, we can make our evaluations only in terms of what final structure we can expect. Then the next work consists on counting the number of those structures on a given random graph. If, for instance, we know that our random graph allows only one tree and the stability analysis tells us that the attractor should have exactly that geometry, then we are sure that such single tree is the component that will survive. But from the theory we know that a C-regular graph is locally tree like [9], so if we discover that the final surviving component is a height-L tree (L low enough not to encounter a long cycle), then it is a matter of counting the number of trees in such regular graph. In particular, if N is large, using the cofactor of the Laplacian matrix [17], we know that the number of spanning trees scales with N^{C-1}.
Within the competitive case, *M. N. Mooij* understood that, for a high enough interaction strength, the GLV was converging to a Maximal Independent Set (MIS) and used this realization to reverse engineer the process and build an algorithm that finds the MIS in a regular graph just by running the proper GLV dynamics on it [38]. However, once again, the MIS is not unique, so the algorithm gives you one of the many MIS.

1.4 Research objectives and questions

As it was discussed in 1.1, this problem was largely investigated in the thermodynamic limit with $N \to \infty$ in the setting of the DMFT. But that requires the *high connectivity limit* with $C >> 1$, $C < N$. The underlying mean field hypothesis is that each species experiences the average effect of the interactions with all others and the single average behavior represents the collective behavior of the entire system. In this study we want to relax that hypothesis, taking the *low connectivity limit* where the interaction matrix is sparse and each species is studied separately, increasing the complexity in the sense of rising behaviours. We can get a first grasp of what this means by looking at the Chapter *Beyond mean-field* of the paper by *F. A. Lopez* [1].

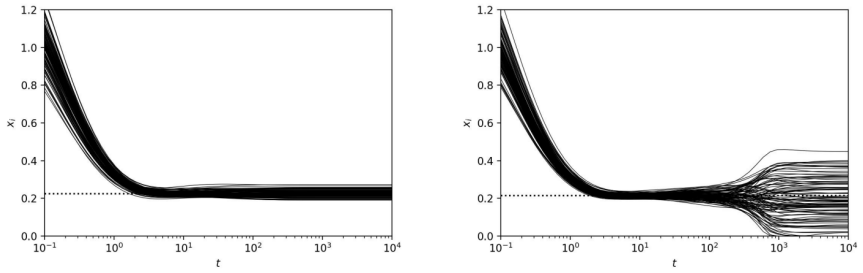


Figure 1.2: Figure from Chapter 7 of [1]. Random regular graph with no degree heterogeneity with a population of 30000 species (50 shown) and average connectivity $C = 50$. There is a value of $\mu = \mu_u \approx -3.57$ such that for $\mu < \mu_u$ there are deviations from the uniform solution. On the left $\mu = \mu_u + \epsilon = -3.56$, on the right $\mu = \mu_u - \epsilon = -3.58$.

In the competitive case of Figure 1.2, if $\mu < \mu_u$ the DMFT does not reflect the system behavior anymore and a new case dependent theory like ours is necessary. But if $\mu = \mu_u - \epsilon$, for ϵ low enough, the DMFT is still useful to describe the transient regime as the unstable directions have positive but small eigenvalues associated with them, making the instability taking very long to develop (the transient time goes like $\frac{1}{|\lambda_{max}|}$, see 2.1). Moreover, in terms of the wealth inequality, we see that beyond the mean field there are many extinctions also when there is no degree heterogeneity and that is what our research is going to focus on.

In our context of the low connectivity case, we want to establish whether there is some kind of mirroring around $\mu = \alpha = 0$ between the competitive and cooperative scenarios, comparing our results on cooperation with the results of [30] (see Figure 1.1). Initially, we want to answer the following questions:

Do we encounter the same steep transitions at the level of the relative diversity pitfalls also in cooperation?

What is the minimal structure of the surviving components in the last plateau? Do we have singlets as in the competitive case?

To what extent are the two cases mirrored?

As anticipated, we will use a modified re-scaled model based on the replicator equations to cope with the divergence of the species abundances in cooperation. This prevents an eventual perfect mirroring but the results are still very interesting to compare. For example, in the Section 4.2 we see that, even if we use the different model, the uniform critical value of the interaction strength is the opposite of the one obtained in competition: $\alpha_u = -\frac{1}{2\sqrt{C-1}}$. That stems an interest to check whether also the other critical values are preserved. We will discover that they are not but the new re-scaled model introduces an *effective competition* within cooperation that leads to interesting results on the final geometries of the surviving components. After the study of the homogeneous case we will introduce the degree heterogeneity with the introduction of the Erdos-Renyi random graph, which in ecological literature was deeply studied in [3], [2].

Moreover, we can give an economical interpretation of our results in terms of the wealth inequality of the final states, where we can consider the species i as individuals or firms. If we consider the final abundances as wealth of individuals or firms for example, it is interesting to study the wealth inequality in the whole population. An economical metric widely used in literature is the Gini coefficient. It ranges from 0 to 1, where 0 represents perfect equality (everyone has the same income or wealth) and 1 represents perfect inequality (one person has all the income or wealth). This is a widely used tool for analyzing wealth distribution in countries and regions [37], [6].

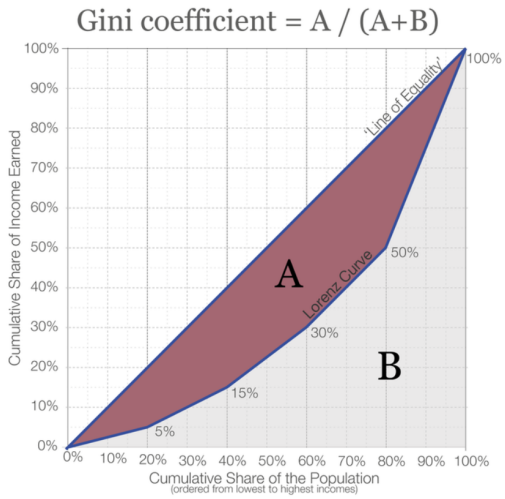


Figure 1.3: Definition of Gini coefficient as the ratio of the two areas [24].

It is defined as the ratio between the area above the Lorenz curve and the total area under the equality line where all individuals have the same wealth. The Lorenz curve $L(F(x))$ plots the cumulative share of income earned by the bottom x fraction of the population, where $F(x)$ is the Cumulative Distribution Function (CDF). In Figure 1.3, for example, the Lorenz curve shows that the poorest 20% of the population has 5% of the income, while the poorest 80% detains half of the total income.
Let y be the abundances (or income) and $F(y)$ be the CDF of the income distribution, the Gini coefficient can be expressed as:

$$G = 1 - \frac{1}{\langle y \rangle} \int_0^\infty (1 - F(y))^2 \, dy. \tag{1.7}$$

Where $\langle y \rangle$ is the mean of the income distribution and $F(y)$ is zero for all negative values.

This thesis will be then structured as follows. After a preparatory explanation of mathematical ecology in Chapter 2, in Chapter 3 we explain the methods we used to find our results, including the environments we worked in and the functions that were used, as well as their computational criticalities. In Chapter 4 we introduce the results, incorporating the conclusions associated to the phenomena we observe. We begin by introducing the arise of divergence in cooperation and how to deal with it with the new model in 4.1, then we study the homogeneous degree case starting with the rings in Section 4.3, to then generalize to any degree C in 4.4. We give an initial flare of how things would change taking an heterogeneous degree distribution in 4.5, focusing on the Erdos-Renyi random graph and finally we sum up our considerations, indicate the limits of the study and propose further research in Chapter 5. If the reader wants to get a grasp of the sequence of results or wants to orient

him/her-self in the structure of this research, it is suggested to look at the *List of Figures* which summarizes the results and the means to achieve them.

2 Mathematical Background

2.1 Mathematical Ecology and Stability Analysis

For this section the reader can refer to the book by *Boyce, Diprima* and *Meade* [13]. The Lotka-Volterra equations were originally formulated to model the interaction between two species, often referred to as predator and prey. Introduced by A. Lotka in 1910 and independently by V. Volterra in 1926, these equations were first applied to describe the dynamics of fish populations in the Adriatic Sea. The classical system is defined as follows:

$$\frac{dx_1}{dt} = ax_1 - \beta x_1 x_2,$$
$$\frac{dx_2}{dt} = \delta x_1 x_2 - \gamma x_2,$$

where x_1 represents the population of prey and x_2 the population of predators. The parameters a, β, δ, and γ are positive constants that describe the growth rates and interactions between the two species: a is the natural growth rate of the prey, β describes the rate at which predators consume prey, δ is the rate at which predators grow due to feeding, and γ is the natural death rate of the predator population.
This system is an example of an *autonomous system of non-linear differential equations*. An autonomous system is one where the independent variable, in this case, time t, does not explicitly appear in the equations, meaning the system's behavior only depends on the current values of the populations x_1 and x_2. The non-linearity arises from the interaction terms $\beta x_1 x_2$ and $\delta x_1 x_2$, which reflect the complex interdependencies between the predator and prey populations.
A *fixed point* (or equilibrium point) of the system occurs when the populations are constant, meaning both derivatives are zero. To find the fixed points, we solve $\frac{dx_1}{dt} = 0$ and $\frac{dx_2}{dt} = 0$. For the Lotka-Volterra equations, this leads to the following four fixed points:

$$(x_1, x_2) = (0, 0) \quad ; \quad (x_1, x_2) = \left(\frac{\gamma}{\delta}, 0\right) \quad ; \quad (x_1, x_2) = \left(0, \frac{a}{\beta}\right) \quad ; \quad (x_1, x_2) = \left(\frac{\gamma}{\delta}, \frac{a}{\beta}\right).$$

They correspond to the *trivial fixed point*, the *prey-only fixed point*, the *predator-only fixed point*, the *non-trivial coexistence fixed point*.
The *phase plane* is a graphical representation in which each axis corresponds to one of the state variables, typically x_1 (prey) on the horizontal axis and x_2 (predators) on the vertical axis. A *phase portrait* consists of the trajectories that represent the evolution of the system over time for different initial conditions. These trajectories show how the populations of both species evolve, and their direction reflects the dynamics dictated by the system of equations. The *trajectories* in the phase plane are curves that indicate the population levels at any given time, starting from different initial conditions. For the Lotka-Volterra system, trajectories typically exhibit cyclical patterns, reflecting the natural oscillations between predator and prey populations. *Nullclines* are important geometric tools for analyzing the behavior of the system in the phase plane. A nullcline is a curve where the derivative of one of the variables is zero, meaning the population of that species remains constant along that curve. For x_1 (prey), the nullcline is found by setting $\frac{dx_1}{dt} = 0$, which gives:

$$ax_1 - \beta x_1 x_2 = 0,$$

or equivalently:

$$x_1 = 0 \quad \text{or} \quad x_2 = \frac{a}{\beta}.$$

For x_2 (predators), the nullcline is obtained by setting $\frac{dx_2}{dt} = 0$, which yields:

$$\delta x_1 x_2 - \gamma x_2 = 0,$$

or equivalently:

$$x_2 = 0 \quad \text{or} \quad x_1 = \frac{\gamma}{\delta}.$$

The intersections of the nullclines correspond to the equilibrium points of the system. These intersections help to determine the system's behavior near the fixed points and give insight into the dynamics of the interacting populations.
Finally, *bifurcation* refers to a qualitative change in the system's behavior as a parameter is slightly varied. In the Lotka-Volterra model, bifurcations can occur when certain parameters, such as β (predation rate) or δ (predator growth rate), are altered. As these parameters increase or decrease, the system may transition between different types of behavior, such as moving from a stable equilibrium to oscillatory dynamics. Two common types of bifurcations relevant to this system are the *saddle-node bifurcation* and the *Hopf bifurcation*. A saddle-node bifurcation occurs when two equilibrium points, one stable and one unstable, collide and annihilate each other as a parameter is varied. This typically happens when one real eigenvalue passes through zero, causing the system to lose or gain an equilibrium point. On the other hand, a Hopf bifurcation occurs when a pair of complex conjugate eigenvalues crosses the imaginary axis, moving from having negative real parts to positive real parts. In

this case, the system transitions from a stable equilibrium to oscillatory behavior, potentially leading to the formation of a limit cycle.

The stability analysis tells us how ecological systems respond to perturbations and whether they return to equilibrium after disturbances. We have different kinds of stabilities.

Lyapunov stability refers to an equilibrium point x^* where, for any small perturbation around x^*, the system remains within a small neighborhood of x^*. Formally, for any neighborhood U around x^*, there exists a smaller neighborhood $W \subset U$ such that:

$$x(0) \in W \implies x(t) \in U \quad \forall t \geq 0.$$

This means that small disturbances do not cause the system to deviate significantly from x^*, but it does not necessarily return to x^*.

Asymptotic stability is a stronger condition. The equilibrium x^* is asymptotically stable if it is Lyapunov stable and if small perturbations cause the system to eventually return to x^* as $t \to \infty$. Mathematically, this implies that for small perturbations:

$$x(0) \in W \implies x(t) \to x^* \text{ as } t \to \infty.$$

In this case, the system not only remains close to x^* but also returns to it over time.

Global stability refers to an equilibrium that is stable for any initial condition in the entire state space. In other words, no matter where the system starts, it will eventually converge to the equilibrium x^*. This means that for any initial condition $x(0) \in (0, \infty)^N$, we have:

$$x(t) \to x^* \text{ as } t \to \infty.$$

Global stability implies that x^* is a global attractor for the entire state space.

Suppose that a feasible equilibrium x^* exists. Then we can ask whether it is *attractive*, i.e., if trajectories started at initial condition $x(0)$ will eventually reach x^*. This problem is in general difficult to solve and as an alternative, we can test for local asymptotic stability, i.e., ask whether the system will return to the equilibrium if perturbed infinitesimally away from it. In general, whenever we describe an ecological community as a system of nonlinear, autonomous ODEs:

$$\frac{dx_i(t)}{dt} = f_i(x(t)),$$

we define an equilibrium x^* as a vector of densities such that:

$$\left.\frac{dx_i}{dt}\right|_{x=x^*} = 0 \quad \forall i$$

A given system might have a multitude of equilibria. When the system is resting at an equilibrium point, it will remain there unless it is perturbed away from it.
Suppose that the system is resting at an equilibrium x^*, and that it is slightly perturbed away from it. $\delta x(t) = x(t) - x^*$ is the state of the system immediately after the perturbation. We Taylor-expand the first derivative around x^*:

$$f(x(t)) = f(x^*) + J_x \delta x(t) + \cdots$$

Where J is the Jacobian matrix of the system, whose elements are defined as:

$$J_{ij} = \frac{\partial f_i(x)}{\partial x_j}$$

Each element of this matrix is therefore a function, whose value depends on x. When we evaluate the Jacobian matrix at an equilibrium point x^*, we obtain the community matrix $J|_{x=x^*}$.

Note that, although each system has a unique Jacobian matrix, there are as many community matrices as there are equilibria. The community matrix details the effect of increasing the density of one species on any other species around the equilibrium point.
We can therefore write the differential equation:

$$\frac{d\delta x(t)}{dt} \approx J|_{x=x^*} \delta x(t)$$

with solution:

$$\delta x(t) = \delta x(0) e^{-tJ|_{x=x^*}} = \delta x(0) V e^{-\Lambda t} V^{-1}$$

Where V is the matrix containing the (unit) eigenvectors of $J|_{x=x^*}$, and Λ is a diagonal matrix containing the eigenvalues of $J|_{x=x^*}$. We can rewrite it as

$$\delta x(t) = \sum_{k=1}^{N} c_k v_k e^{\lambda_k t}$$

Where v_k are the eigenvectors of $J|_{x=x^*}$ and $c_k = \delta x_k(0)$ are constants determined by the initial conditions.

Calling *mode* the direction given by an eigenvector, each eigenvalue λ_k corresponds to the growth or decay rate of a particular mode v_k . So if λ_k is positive, the corresponding mode grows exponentially over time. As time progresses, the mode associated with the largest eigenvalue λ_{max} (in terms of real part) will dominate the system's behavior because it grows or decays faster than all other modes. When there is a transient regime where the species

seems to relax to a uniform equilibrium, the system exits the transient regime when the most dominant mode associated with λ_{max} begins to clearly dictate the behavior of the system. The time to exit the transient regime is roughly given by $t_{trans} \approx \frac{1}{|\lambda_{max}|}$. This is because exponential functions $e^{\lambda t}$ have characteristic time scales given by $\frac{1}{|\lambda|}$.

The eigenvalues of $J|_{x=x^*}$ determine the stability of the equilibrium x^*: if all the eigenvalues have negative real part, then the system will eventually return to the equilibrium after sufficiently small perturbations; conversely, if any of the eigenvalues have positive real part, the system will move away from the equilibrium whenever perturbed. Therefore, depending on the sign of the real part of the "rightmost" eigenvalue of $J|_{x=x^*}$, λ_1, we can determine the stability of x^*:

$$\mathrm{Re}(\lambda_1) \begin{cases} < 0 & \to x^* \text{ is asymptotically stable} \\ = 0 & \to x^* \text{ is Lyapunov stable} \\ > 0 & \to x^* \text{ is unstable} \end{cases}$$

While local stability analysis considers behavior near an equilibrium point, we may also be interested in global stability. Global stability implies that the equilibrium x^* is stable for any initial condition, meaning that no matter where the system starts, it will eventually return to x^*. After understanding the stability of equilibrium points through eigenvalues and their real parts, another important concept in the analysis of dynamical systems is the *basin of attraction*. The basin of attraction of an equilibrium point x^* refers to the set of initial conditions from which the system will eventually converge to that equilibrium point as $t \to \infty$.
Mathematically, the basin of attraction $B(x^*)$ for an equilibrium x^* can be defined as:

$$B(x^*) = \{x(0) \in \mathbb{R}^n \mid \lim_{t\to\infty} x(t) = x^*\}.$$

This means that if the system starts at any point $x(0) \in B(x^*)$, the trajectory $x(t)$ will approach the equilibrium x^* as time tends to infinity. In other words, the basin of attraction encompasses all initial conditions that lead to convergence to a particular stable equilibrium point. For systems with multiple stable equilibria, each equilibrium will have its own distinct basin of attraction. The boundary separating these basins is often where the system can exhibit different behavior, depending on small perturbations.

To study the basin of attraction and understand the influence of the initial conditions on the stable equilibria, we analyzed the competitive case on simple cyclic structures: the *triangle* and the *square*.

The triangle is characterized by the following adjacency matrix:

$$\mathrm{A} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

Figure 2.1: Triangle adjacency matrix, N = 3

and the following system of ODEs, with initial conditions $x(0) = x_0, y(0) = y_0, z(0) = z_0$:

$$\begin{aligned} \dot{x}(t) &= x(t)(1 - x(t) - \alpha y(t) - \alpha z(t)) \\ \dot{y}(t) &= y(t)(1 - y(t) - \alpha x(t) - \alpha z(t)) \\ \dot{z}(t) &= z(t)(1 - z(t) - \alpha x(t) - \alpha y(t)) \end{aligned}$$

We can study the stability of the system by starting from the stability of the locally linear system around the 8 fixed points of the triangle. The general procedure is to compute the Jacobian matrix, evaluate it in the fixed point i, compute its eigenvalues λ_j^i and eigenvectors $\boldsymbol{\xi_j^i}$ to get the following dynamics around the fixed point:

$$x_i(t) = \sum_{j=1}^{N} c_j^i \boldsymbol{\xi_j^i} e^{\lambda_j^i t}. \tag{2.1}$$

where the constants are found via the initial conditions. Then we say that a fixed point is an asymptotically stable attractor if all its eigenvalues are negative. If they are all positive it is an unstable node or spiral point (that repels away), while if some are positive and some are negative it is a saddle point.

For the triangle, the Jacobian is,

$$J = J(x, y, z) = \begin{bmatrix} 1 - 2x - \alpha y - \alpha z & -\alpha x & -\alpha x \\ -\alpha y & 1 - 2y - \alpha x - \alpha z & -\alpha y \\ -\alpha z & -\alpha z & 1 - 2z - \alpha x - \alpha y \end{bmatrix}$$

Now we can group the fixed points based on the number of surviving species (#SS) and find the eigenvalues associated to them.

From Table 2.1 we can study the α intervals for when the maximal eigenvalue of each group is negative. We obtain that:

- for $\alpha < 1$ only the uniform fixed point with 3 species surviving (3SS) is stable.
- for $\alpha > 1$ only the fixed points with 1 species surviving (1SS) are stable.

(#SS)	Fixed point	Eigenvalues
0SS	$(0,0,0)$	$(1,1,1)$
1SS	$(1,0,0),(0,1,0),(0,0,1)$	$(1-\alpha, 1-\alpha, -1)$
2SS	$(\frac{1}{\alpha+1}, \frac{1}{\alpha+1}, 0), (\frac{1}{\alpha+1}, 0, \frac{1}{\alpha+1}), (0, \frac{1}{\alpha+1}, \frac{1}{\alpha+1})$	$\left(-\frac{\alpha-1}{\alpha+1}, -1, \frac{\alpha-1}{\alpha+1}\right)$
3SS	$\left(\frac{1}{2\alpha+1}, \frac{1}{2\alpha+1}, \frac{1}{2\alpha+1}\right)$	$\left(-1, \frac{\alpha-1}{2\alpha+1}, \frac{\alpha-1}{2\alpha+1}\right)$

Table 2.1: Fixed points and eigenvalues for different numbers of surviving species.

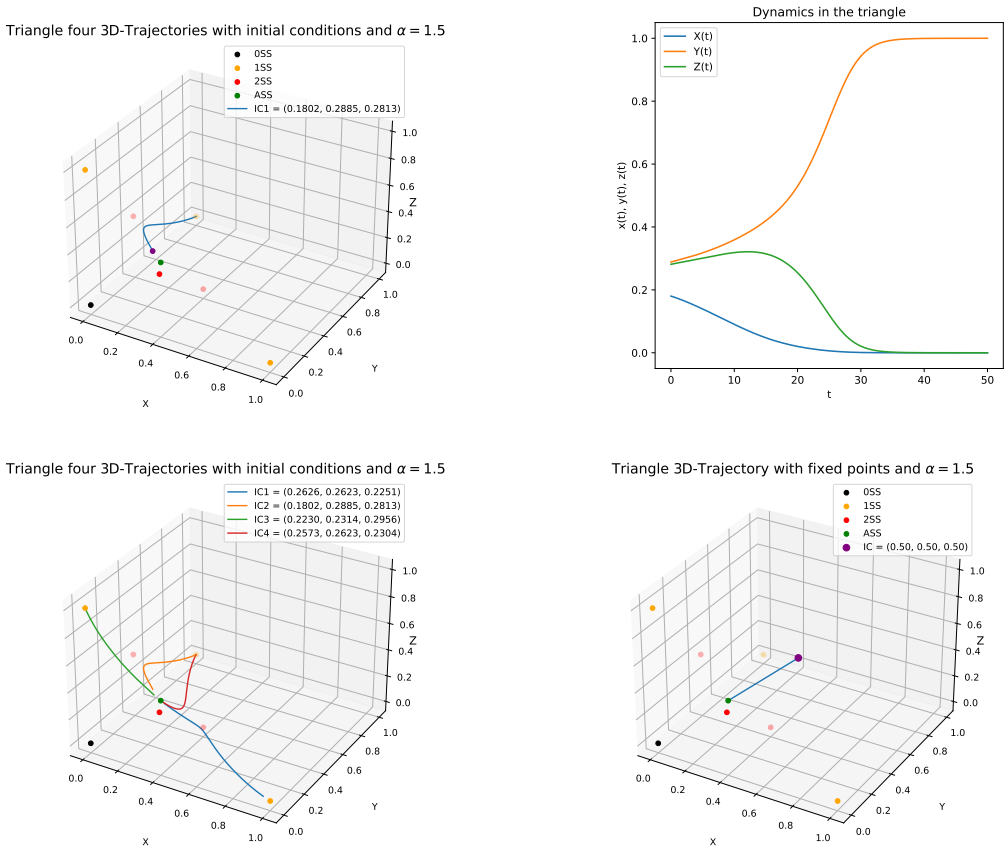
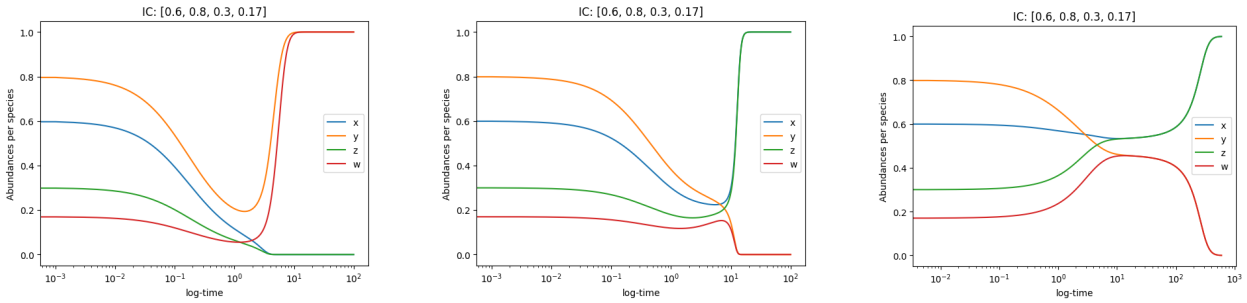


Figure 2.2: First Row) 3D trajectory and dynamics of the system for x, y, z starting from the same initial conditions with $\alpha = 1.5$. Bottom-Left) 3-D trajectories to show that among the 3 possible fixed points in group #1SS the chosen one is always the one associated to the biggest IC component. In $\{2, 4\}$ the biggest IC component is y, in $\{1\}$ is x, in $\{3\}$ is z. Bottom-Right) Stable manifold example. The initial condition are given by a perturbation of the uniform fixed point in the direction of the stable eigenvector $(1, 1, 1)$. The result is that the trajectory follows the stable manifold and converges to the uniform fixed point even if $\alpha > 1$. 'ASS' stands for 'all surviving species' and in this case is equivalent to 3SS.

In the first row of Figure 2.2 we notice how the dynamics initially approach the saddle point associated to a fixed point in #2SS and then, since we are in the α-range where the stable equilibrium belongs to #1SS, the z-component changes direction and converges to zero. The smaller the positive eigenvalue of the associated fixed point in #2SS, the longer it takes to change derivative sign and descend towards zero. In the third image we observe that when we are in the regime $\alpha > 1$ and it is always one species surviving, the choice of which of the three survives is given by a simple criterion: the species with highest IC survives. There can be exceptions from the rules indicated by the α regimes. In the last image of 2.2 we notice that the system converges to the uniform equilibrium even if $\alpha > 1$. That is because the initial conditions are set on the stable manifold of the uniform fixed point $\left(\frac{1}{2\alpha+1}, \frac{1}{2\alpha+1}, \frac{1}{2\alpha+1}\right)$. A stable manifold is the curve, passing through the fixed point, that is tangent to the eigenvector associated to the eigenvalue connected to that fixed point. During our research we will often give these one-to-one relations α-to-equilibrium, and when they are not verified in the simulations it is probably because the IC where chosen on the stable manifold of another equilibrium.

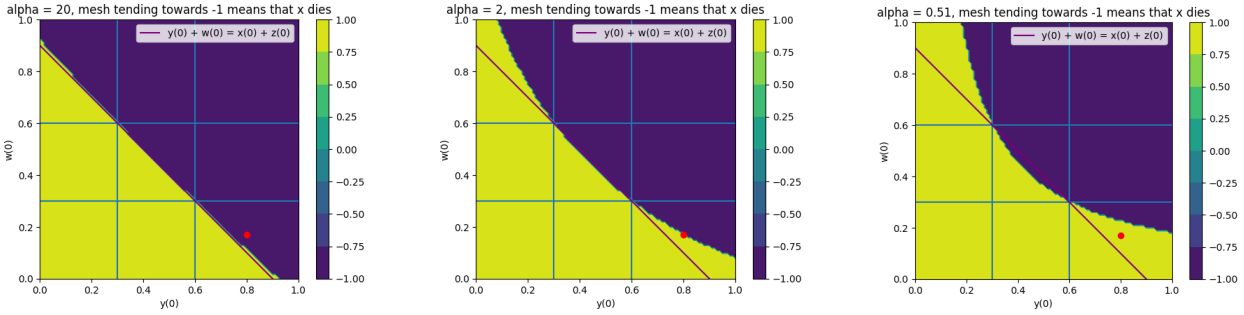
However, it is not always that easy to establish which fixed point is chosen within those of the same group #SS. In the case of the triangle it is simply a matter of taking the component with highest IC. In the *square*, on the other hand, we have that for $\alpha > 0.5$, the system converges to a fixed point where the surviving species are the two non adjacent nodes that lay on the diagonal of the square. The question we tried to answer is: *what is the criterion on the initial conditions that tells us which of the two pairs is going to survive?*

From the example $(x(0), y(0), z(0), w(0)) = (0.6, 0.8, 0.3, 0.17)$ and $\alpha = 2$ we can see that $y(0) + w(0) > x(0) + z(0)$ but y and w die. On the other hand, we see that if we take $\alpha = 6$, y and w are the ones that survive.



(a) $\alpha = 6$ (b) $\alpha = 2$ (c) $\alpha = 0.51$

Figure 2.3: Dynamics of the species x, y, z, w with a square-shaped network with the same initial conditions but different values of the interaction strength $\alpha = (6, 2, 0.5)$. According to the chosen value of α a different non-adjacent pair survives.



(a) $\alpha = 20$ (b) $\alpha = 2$ (c) $\alpha = 0.51$

Figure 2.4: We set $x(0) = 0.6$ and $z(0) = 0.3$. The purple dotted line is such that $y(0) + w(0) = 0.9$ The color spectrum is x that takes the value 1 when it survives and -1 when it dies. As α decreases the boundary takes a more parabolic shape that always seems to be passing via $(y(0), w(0)) = (0.3, 0.3)$ and $(y(0), w(0)) = (0.6, 0.6)$. The red dot is the point $(y(0), w(0)) = (0.8, 0.17)$ that we know is associated with the death of y and w for $\alpha = 2$. Indeed, in the middle figure the red dot is just outside the boundary.

We tried to find a rule that guarantees the possibility to predict the surviving pair from the information given by the initial conditions, but the non linearity of the system makes this task very complicated. From Figure 2.4, we see that for high values of α (here $\alpha = 20$) the pair that survives is the one with highest sum of initial conditions. But as we can see, as α decreases towards the uniform fixed point regime, the boundary of the mesh assumes a parabolic shape that is difficult to quantify analytically. This example shows how, even for a simple structure like the *square*, predicting the exact species that are going to survive just by looking at the initial conditions and the value of α is a very complicated problem. In this research we will limit our scope to the quest of the group of equilibria that should survive, relying only on the information given by the value of α and the structure of the network. In this case, that consists in understanding that for $\alpha > 0.5$ we expect the equilibrium to be a pair of disconnected nodes.

3 Methodology

The goal of the research is to find the structure and the final state distribution of the system of N species which all obey the ordinary differential equation of the Generalized Lotka-Volterra model. This research starts from numerical observations that hint the emergence of specific phenomena and then focuses on finding their analytical proof. The process was often to verify an intuition numerically, observe the phenomena and then prove them analytically using the *ansatz* extracted from the simulations; then finally verify, again numerically, the theoretical results. For instance, in the case of the result in Equation 4.11, first we observed the phenomenon in Figure 4.6 and understood that the surviving species where distributed as a symmetric chain, then we used that assumption to find the theoretical value that drives us from a length-$(L-1)$ chain to a length-L chain.

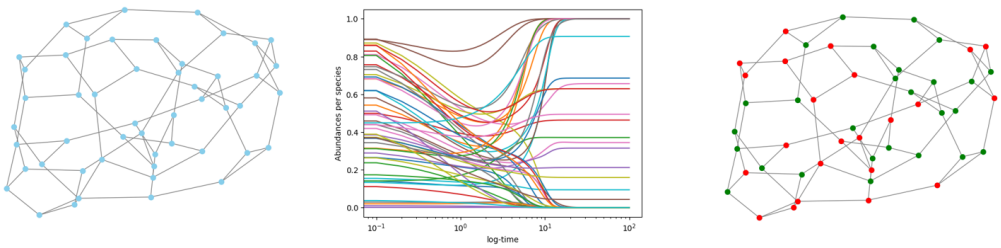


Figure 3.1: Simulations process: build the underlying random graph, then run the dynamics, then check how the final abundances reflect on the nodes of the graph. The x-axis is just time in a logarithmic rescaling.

The underlying process to run the simulations is the same for all the steps of the research. Initially, generate the initial conditions uniformly at random in $(0, 1)$ and the random graph with the *Networkx* library of *Python*. According to the number of iterations that have to be done with the same parameter, one or more seeds are used. For instance, in Figure 4.20 only one seed is used, while in Figure 4.31 twenty seeds are used. When the network is set, the dynamics are run with integrating functions from *Scipy.integrate*. In particular in the competitive case we use:

scipy.integrate.odeint(LV_ODEs, initial_conditions, t, args=(A_adjacency, alpha))

Where *LV_ODEs* is the function that we define to respect the equations in 1.2. This function is too slow for large values of N so, in particular in the cooperative case where we want to find the fractal exponent, we replace it with:

scipy.integrate.solve_ivp(Cooperative_LV_ODEs, [0, tmax], initial_conditions, args=(A_adjacency, alpha, C), method = 'DOP853', t_eval=t, rtol = 1e-8, atol = 1e-10)

Where *Cooperative_LV_ODEs* is the function we define to respect the equations in 4.3.

solve_ivp() allows us to arrive to N up to 50000, but it requires a very accurate choice of the last integrating time *tmax*. As we increase μ, we have to decrease *tmax*, but we need to find a very accurate upper bound of the real converging time because otherwise the server crushes. Moreover, to obtain the locally tree-like structure of C-regular graphs, when we increase C we have to super-linearly increase N. If we consider the limitations on the size we can use, this entails that the results for higher C are less accurate. This can be observed in Subsection 4.4.1, where for $C = 3$ the theoretical results are very accurate, while for $C = 5$ they are not as much. That has a theoretical explanation, that has to be carefully paired with this computational one.

Once we have run the dynamics, we can get the final state vector and establish that the surviving species are those with abundance greater than a certain tolerance (that we set to $1e - 6$). Then we can observe what nodes in the original graph they correspond to. That allows us to find the structure of the final component(s) and to compute metrics like the Gini index or the Shannon entropy.

4 Results

In the Introduction we introduced our will to extend to the cooperative case the analysis of the evolution of the equilibria with the interaction strength. This chapter will be dedicated to that.

4.1 Divergence

If we look at the phase portrait of a two species GLV (x and y) under cooperation, we notice that the species benefit from the others' presence reaching an equilibrium where either they both survive with the same abundance or they diverge.

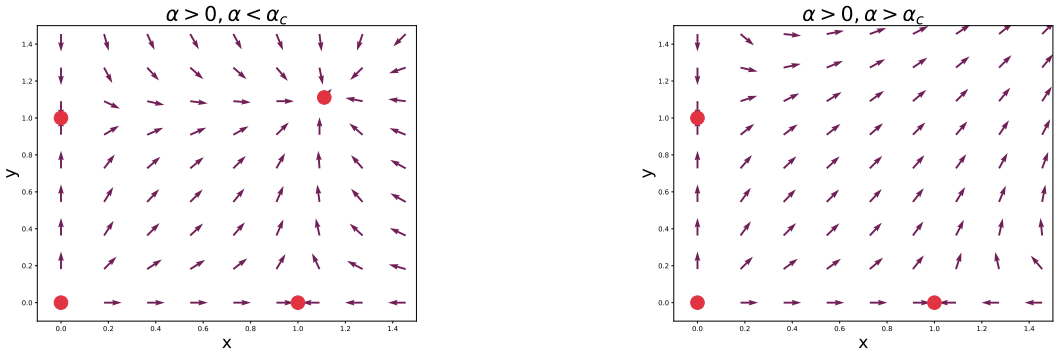


Figure 4.1: GLV with $N = 2$. Phase portrait with vector field and 2^N fixed points. According to whether the interaction strength is before or after the critical α_c (we will call it α_d to indicate divergence), the species converge or diverge.

From Equation 1.2, we can find the uniform equilibrium where all species survive with the same abundance $x^\star$. This corresponds to setting the second factor to zero for all species.

$$1 - x^\star + \alpha C x^\star = 0 \Leftrightarrow x^\star = \frac{1}{1 - \alpha C}.$$

If we call α_d the critical value of interaction strength after which abundances start diverging, then we can obtain $\alpha_d = \frac{1}{C}$ by setting the uniform solution to infinity. Those abundances diverge to different orders of infinity, so even if the lowest are not exactly extinct, relatively to the most wealthy species they are. It is of great interest to study this hierarchy because in many settings it is not important how much you have, but how much you have compared to the other agents in the environment. Talking about firms or national economies, even if two of them have positive liquidity but very uneven in terms of inequality, then the poorest one cannot benefit from its positive liquidity simply for the limited nature of goods or resources.

To cope with the divergence we introduced a re-scaled model that resembles the replicator equations in [25].

Let us start by recalling Equation 1.2,

$$\dot{x}_i(t) = x_i(t)(1 - x_i(t) + \alpha \sum_j A_{i,j} x_j(t)).$$

given that the degree of vertex i is $k_i = \sum_j A_{i,j}$, we introduce the normalizing factor

$$M(t) = \frac{1}{N} \sum_{i,j} A_{i,j} x_j(t) = \frac{1}{N} \sum_j k_j x_j(t).$$

whose time derivative is

$$\dot{M}(t) = \frac{1}{N} \sum_j k_j x_j(t)(1 - x_j(t) + \alpha \sum_l A_{j,l} x_l(t)).$$

Now we re-scale the abundances with that normalizing factor

$$y_i(t) = \frac{x_i(t)}{M(t)}. \tag{4.1}$$

and derive with respect to t,

$$\dot{y}_i(t) = \frac{\dot{x}_i(t)}{M(t)} - \frac{x_i(t)\dot{M}(t)}{M^2(t)}$$

$$= x_i \left(\frac{1}{M} - \frac{x_i}{M} + \alpha \sum_j A_{i,j} \frac{x_j}{M} - \sum_j \frac{k_j}{N} \frac{x_j}{M} \left(\frac{1}{M} - \frac{x_j}{M} + \alpha \sum_l A_{j,l} \frac{x_l}{M} \right) \right)$$

$$= M y_i \left(\frac{1}{M} - y_i + \alpha \sum_j A_{i,j} y_j - \frac{1}{N} \sum_j k_j y_j \left(\frac{1}{M} - y_j + \alpha \sum_l A_{j,l} y_l \right) \right).$$

where we stopped expliciting that x, y, M depend on t.
Now we define t_c as the time when the abundances start diverging, or $\lim_{t\to t_c} M(t) = \infty$.
Since we want to have the dynamics defined from zero to infinity, we re-scale time as

$$\tau = \int_0^t M(s)\, ds. \tag{4.2}$$

In this way time slows down as t_c is approached. Moreover, if we integrate 4.2 from zero to t_c we get that $\tau = \infty$.
Finally, we can neglect the $\frac{1}{M}$ because when the system stabilizes to an equilibrium where the abundances tend to infinity it is infinitesimal. Keeping implicit the dependence of y_i on τ, we get

$$\frac{\partial y_i}{\partial \tau} = y_i(-y_i + \alpha \sum_j A_{i,j} y_j - \frac{1}{N} \sum_j k_j y_j(-y_j + \alpha \sum_l A_{j,l} y_l)). \tag{4.3}$$

And when we are in the C-regular case ($k_i = C\ \forall i$), it becomes

$$\frac{\partial y_i}{\partial \tau} = y_i(-y_i + \alpha \sum_j A_{i,j} y_j - \frac{C}{N} \sum_j y_j(-y_j + \alpha \sum_l A_{j,l} y_l)). \tag{4.4}$$

Here we use a degree weighted normalizing factor, which is different than the $M = \frac{1}{N}\sum_j x_j(t)$ used in [25]. The reason why we include the degree is that the extension of this work would entail using an heterogeneous degree random graph and our normalization allows to give more importance to the highly connected nodes in terms of the influence they have on others [1]. Using $M = \frac{1}{N}\sum_j x_j(t)$ in our regular case would practically mean getting rid of the dependence on C in Equation 4.4. Normalizing by $M(t)$ has the same logic of using Lagrange multipliers to impose a global constraint on the system. In the case of $M(t)$, the normalization ensures that the total influence of the system (or total "mass") is accounted for, similarly to how Lagrange multipliers enforce a constraint on the solution space of an optimization problem. Another relevant observation about the new model is that the third term in the parenthesis is global as it does not depend on i. In particular $-\frac{1}{N}\sum_j k_j y_j \alpha \sum_l A_{j,l} y_l$ induces an *effective cooperation* that will be crucial to explain the difference with the competitive model.

Now we propose an example of the dynamics stopped at t_c with the original model and with the replicator model.

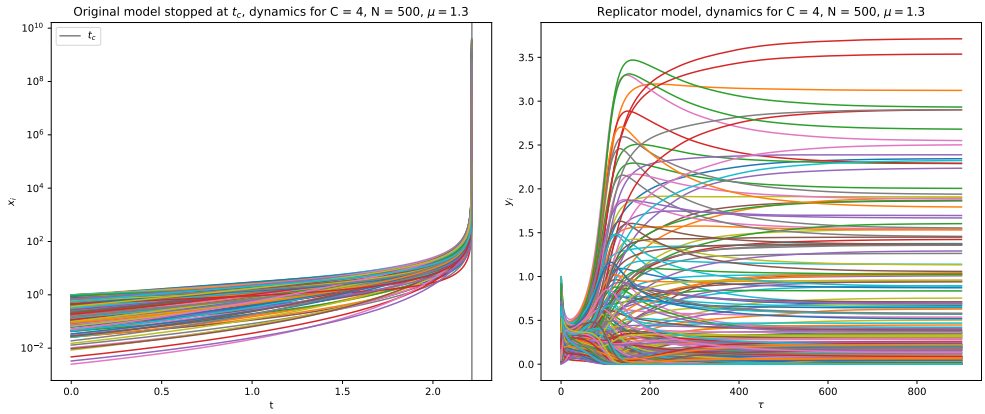


Figure 4.2: For $\mu = 1.3$ the cooperative original Lotka-Volterra model diverges at time t_c (Left). The replicator model is well defined on $\tau \in [0, \infty)$ and the hierarchy between species is more evident with many of them going extinct relatively to the others (Right).

4.2 Different regimes for C-regular graphs

To analyze the stability of an equilibrium and see how the regimes change with α, we need to write the analytical expression for the Jacobian. Let us start from the vectorial version of Equation 4.4.

$$\dot{\mathbf{y}} = F(\mathbf{y}) = D(\mathbf{y}) \left(-\mathbf{y} + \alpha A\mathbf{y} - \frac{C}{N}(\mathbf{y}^T(-\mathbf{y} + \alpha A\mathbf{y}))\mathbf{1} \right) = D(\mathbf{y})H(\mathbf{y}).$$

$$H(\mathbf{y}) = -\mathbf{y} + \alpha A\mathbf{y} - \frac{C}{N}(\mathbf{y}^T(-\mathbf{y} + \alpha A\mathbf{y}))\mathbf{1}.$$

where $\mathbf{y}$ is a column vector $\mathbf{y} = (y_1, \ldots, y_N)$, $D(\mathbf{y})$ is a $N \times N$ diagonal matrix with the elements of $\mathbf{y}$ on the diagonal, and $\mathbf{1}$ is a column vector of ones. Here we consider the C-regular graph case. To find the Jacobian matrix $J(\mathbf{y})$ of the function $F(\mathbf{y})$, we need to compute the partial derivative of each component $F_i(\mathbf{y})$ with respect to each y_j.
Using the product rule,

$$\frac{\partial F_i(\mathbf{y})}{\partial y_j} = \frac{\partial}{\partial y_j} \left(D(\mathbf{y})H(\mathbf{y}) \right)_i = \frac{\partial}{\partial y_j} \left(D_i(\mathbf{y})H_i(\mathbf{y}) \right) = \delta_{ij}H_i(\mathbf{y}) + y_i \frac{\partial H_i(\mathbf{y})}{\partial y_j}.$$

where δ_{ij} is the Kronecker delta.

$$H_i(\mathbf{y}) = -y_i + \alpha(A\mathbf{y})_i - \frac{C}{N}\left(\sum_{k=1}^{N} y_k(-y_k + \alpha(A\mathbf{y})_k)\right).$$

$$\frac{\partial H_i(\mathbf{y})}{\partial y_j} = -\delta_{ij} + \alpha A_{ij} - \frac{C}{N}\left(-2y_j + 2\alpha(A\mathbf{y})_j\right).$$

Where the third term comes from:

$$\frac{\partial}{\partial y_j}\left(\sum_{k=1}^{N} y_k(-y_k + \alpha(A\mathbf{y})_k)\right) = -2y_j + \alpha(A\mathbf{y})_j + \alpha(\mathbf{y}'A)_j.$$

but since A is symmetric, $(A\mathbf{y})_j = (\mathbf{y}'A)_j$.

Finally combine the results,

$$J_{ij}(\mathbf{y}) = \delta_{ij}H_i(\mathbf{y}) + y_i\left(-\delta_{ij} + \alpha A_{ij} - \frac{C}{N}\left(-2y_j + 2\alpha(A\mathbf{y})_j\right)\right). \tag{4.5}$$

Then we have to plug in the equilibrium we are interested to investigate. Let us start with the uniform fixed point where all the species survive with the same abundances $y^\star$. Similarly to the original model, we have to impose that the second factor in Equation 4.4 is zero.

$$-y^\star + \alpha C y^\star - \frac{C}{N} N y^\star(-y^\star + \alpha C y^\star) = 0 \Leftrightarrow y^\star = \frac{1}{C}.$$

Where we disregard the unfeasible solution $y^\star = 0$

If $\mathbf{y}^\star = (y^\star, \ldots, y^\star)^T)$ then $(A\mathbf{y}^\star)_j = 1$ and $H_i(\mathbf{y}^\star) = 0$ and the Jacobian evaluated in $\mathbf{y}^\star$ becomes:

$$J_{ij}(\mathbf{y}^\star) = \frac{1}{C}\left(-\delta_{ij} + \alpha A_{ij} + \frac{2}{N} - \frac{2\alpha C}{N}\right).$$

Now for a C-regular random graph with adjacency matrix A, the maximal eigenvalue is $\lambda^A_{max} = 2\sqrt{C-1}$ [32]. It follows that,

$$\lambda^J_{max} = \frac{1}{C}\left(-1 + \alpha 2\sqrt{C-1} + \frac{2}{N} - \frac{2\alpha C}{N}\right).$$

which needs to be negative to have stability, leading to a critical value

$$\alpha_u = \frac{1 - \frac{2}{N}}{2\sqrt{C-1} - \frac{2C}{N}} \approx \frac{1}{2\sqrt{C-1}} \qquad \text{for} \quad N >> C.$$

Which is in accordance with the value obtained in [30] for the competition but with opposite sign. This means that for small values of $|\alpha|$ we have symmetry between the two scenarios.

The question now is how is this symmetry broken as we increase $|\alpha|$. We start answering this by showing how the choice of the interaction strength leads to different regimes in cooperation.

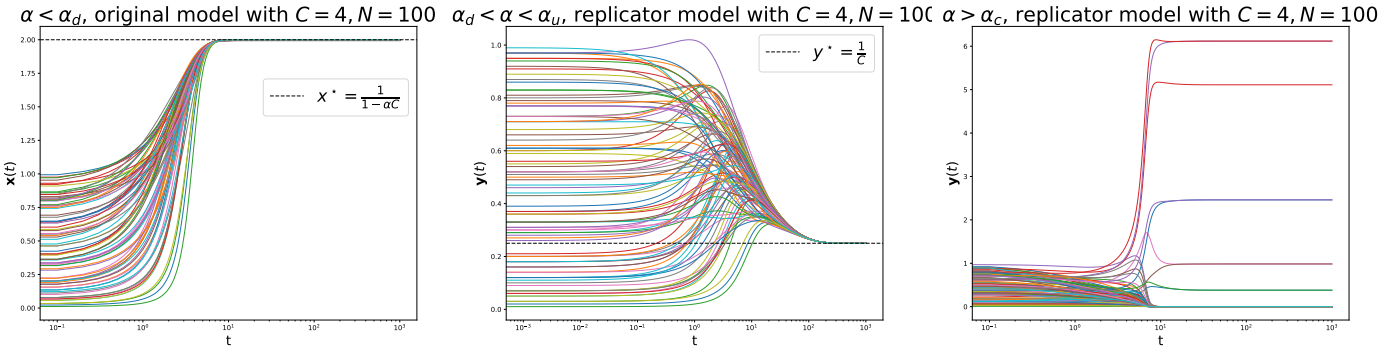


Figure 4.3: Left) $\alpha < \alpha_d = \frac{1}{C}$: uniform solution with the original model. Middle) $\alpha_d < \alpha < \alpha_u \approx \frac{1}{2\sqrt{C-1}}$: uniform solution with the replicator model. Right) $\alpha > \alpha_u$: solution away from the mean field where most of the species go extinct.

Keep in mind that if the dynamics are converging and we have $\alpha < \alpha_d$ we cannot use the replicator model because that neglects the factor $\frac{1}{M}$, that for convergence would not be infinitesimal. For this reason in the first image I run the dynamics in x and not in y. From the third image we can already notice that the number of surviving species is very low, much lower then the $\phi = 0.4$ that we encounter in 1.1. Now we want to characterize those equilibria for high α and we want to see how ϕ scales with N.
In the following section we study the stability of the final structures by analyzing how the maximal eigenvalue changes with the interaction strength, and for which values of α we encounter the transitions. To do that analytically, we need to do use an *ansatz* for the structure of the final equilibrium, because we need to evaluate the Jacobian in the solution $\mathbf{y}$. Hence the strategy is to run numerical simulations to understand what are these structures, then impose analytically that the solution is a chain (or star or tree or cyclic graph) and finally study the spectrum of the eigenvalues of the Jacobian evaluated in that solution.

4.3 Rings

We start from the analysis of a population connected with a ring shaped network, that is equivalent to setting $C = 2$. The handy property of this structure is that the adjacency

matrix is fixed, up to a labels permutation:

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix}$$

We run the dynamics of a system with such adjacency matrix, for a high value of the interaction strength, to check what is the minimal solution (or minimal geometry).

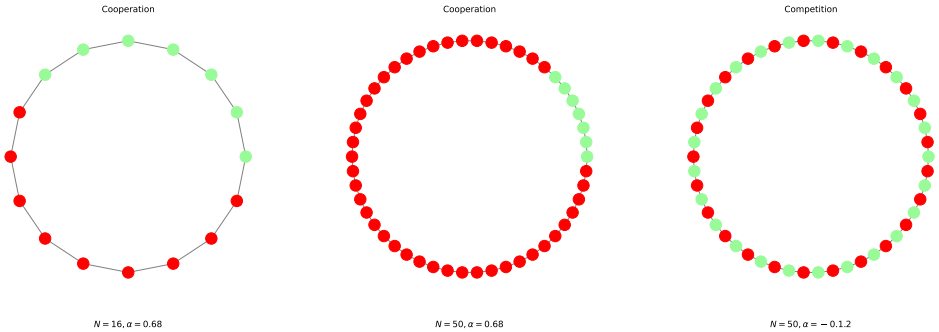


Figure 4.4: Graph visualization of the surviving species after running the dynamics in a ring. A species is considered to be extinct if its abundance is below a certain tolerance (here tol = 1e-3). Left) Cooperation with $\alpha = 0.68$ and $N = 16$. Middle) Cooperation with $\alpha = 0.68$ and $N = 50$. Right) Competition with $\alpha = -0.68$ and $N = 50$.

The simulations suggest that the surviving species are grouped into length-L chains (in the example, L = 7). In the competitive case, the dynamics produced the maximal independent sets, which in the ring are non-connected alternating singlets. That is in accordance with [30]. It is remarkable that the length of the chain in the cooperative case does not depend on N, but only on α. Within competition, what is conserved is the surviving fraction ϕ, meaning that if we increase N, ϕ will remain 50%. On the other hand, within cooperation the conservation is on the number of survivors. This phenomenon leads to what is called *condensation* (see 1.1), meaning that a significant portion of the total wealth is "condensed" or accumulated by a few agents in a way that is independent on N. We will see that this will hold for all the C-regular cooperative analysis. It is initially counter intuitive to think that a concentration of wealth can happen within cooperation, where all individuals should benefit from the others' presence, but it is not if you consider that it is relative to the wealth of other species. What is interesting is that the number of species that detains almost all

the wealth does not increase with the total population size.

In [12] *J.P.Bouchaud* and *M. Mezard* introduce the idea of condensation in wealth inequality starting from a model that is very similar to the GLV; the difference is that here such phenomenon is produced by a self-generated noise, while there it is produced by the explicit volatility σ of the spontaneous growth. So there the noise is imposed, here it is not.

Within the ring, it is interesting to find the analytical values of the critical interaction strengths where we pass from having a chain of length $L-1$ to a chain of length L. First we need to find the analytical solution $\mathbf{y}^\star$.

Since the ring is invariant to combinations of length L chains, we can arbitrarily take

$$\mathbf{y}^\star = (y_1^\star, ..., y_L^\star)^T. \tag{4.6}$$

$$A = A_{1:L,1:L}. \tag{4.7}$$

That is because all the other $N-L$ entrances of $\mathbf{y}^\star$ and of A are zero and they do not affect the sign of the maximal eigenvalue.

From 4.4 we have,

$$y_i^\star = \alpha \sum_j A_{ij} y_j^\star - \rho \beta'(\mathbf{y}^\star)$$

$$\rho = \frac{C}{N}$$

$$\beta'(\mathbf{y}^\star) = \sum_j y_j^\star (-y_j^\star + \alpha \sum_l A_{jl} y_l^\star) \in \mathbb{R}.$$

Which in vectorial form is,

$$\mathbf{y}^\star = \alpha A \mathbf{y}^\star - \rho \beta'(\mathbf{y}^\star)\mathbf{1}$$

$$\beta'(\mathbf{y}^\star) = \mathbf{y}^{\star T}(-I + \alpha A)\mathbf{y}^\star.$$

Which replacing $\mathbf{y}^\star$ into $\beta'(\mathbf{y}^\star)$ leads to,

$$\mathbf{y}^\star = \frac{\beta}{\rho}(-I + \alpha A)^{-1}\mathbf{1} \tag{4.8}$$

$$\beta = (\mathbf{1}^T(-I + \alpha A)^{-1}\mathbf{1})^{-1}.$$

Where we used,

$$A = A^T \Rightarrow (-I + \alpha A)^T = (-I + \alpha A) \Rightarrow (-I + \alpha A)^{-T} = (-I + \alpha A)^{-1}.$$

The analytical solution for the final abundances is in accordance with the simulations as we can see from the following figure,

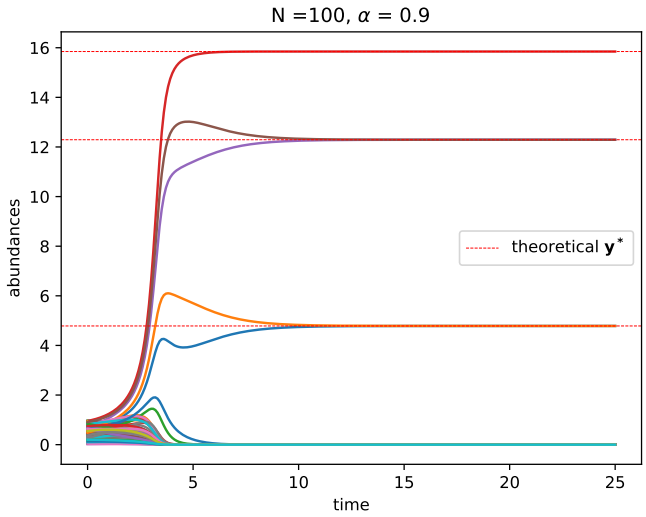


Figure 4.5: Subset of the final state with only the surviving species that are grouped in a length-5 chain. $\mathbf{y}^\star = [4.78, 12.29, 15.84, 12.29, 4.78]$, the abundances are symmetric with respect to the central node.

Now, if we take only the non-zero entrances for $\mathbf{y}^\star$ and A as in Equation 4.6, then $H_i(\mathbf{y}^\star) = 0$ for all i, and we can write the Jacobian as

$$\begin{aligned} J(\mathbf{y}^*) &= D(\mathbf{y}^*)(-I + \alpha A - 2\rho((-I + \alpha A)\mathbf{y}^*\mathbf{1}^T)^T) \\ &= D(\mathbf{y}^*)(-I + \alpha A - 2\rho\mathbf{1}\mathbf{y}^{*T}((-I + \alpha A))^T) \\ &= D(\mathbf{y}^*)(I - 2\rho\mathbf{1}\mathbf{y}^{*T})(-I + \alpha A) \\ &= D(\mathbf{y}^*)(-I + \alpha A - 2\beta\mathbf{1}\mathbf{1}^T(-I + \alpha A)^{-T}(-I + \alpha A)) \\ &= D(\mathbf{y}^*)(-I + \alpha A - 2\beta\mathbf{1}\mathbf{1}^T) \\ &= D(\mathbf{y}^*)M. \end{aligned} \tag{4.9}$$

where in the fourth equation we substituted the solution found in 4.8, and in the following equation we used the symmetry of $(-I + \alpha A)$. Remember that $\beta = (\mathbf{1}^T(-I + \alpha A)^{-1}\mathbf{1})^{-1}$.

Definition 4.3.1 (Inertia of a Matrix)**.** The inertia of a symmetric matrix A is defined as the triplet (n_+, n_-, n_0), where:

- n_+ is the number of positive eigenvalues of A,

- n_- is the number of negative eigenvalues of A, and
- n_0 is the number of zero eigenvalues of A.

Definition 4.3.2 (Congruence of two Matrices)**.** A and B are congruent if $C^T AC = B$ for some non-regular matrix C.

Theorem 4.3.0.1 (Sylvester's Law of Inertia)**.** *A and B are two real symmetric matrices, then*
A and B are congruent $\Leftrightarrow$ A and B have the same inertia.

Corollary 4.3.0.1.1. *If A is positive definite and symmetric and B is symmetric, then AB and B have the same inertia.*

Proof. Since A is positive definite and symmetric, $A = Q\Lambda Q^T$ where where Q is an orthogonal matrix and Λ is a diagonal matrix with positive entries. Then the square root is unique and positive definite as well, $A^{1/2} = Q\Lambda^{1/2}Q^T$. Consider the matrix $C = A^{1/2}BA^{1/2}$. The matrix C is symmetric because $A^{1/2}$ and B are symmetric. By Sylvester's law of inertia, B and C have the same inertia. Since AB and C are similar matrices (because $AB = A^{1/2}CA^{-1/2}$), they have the same eigenvalues and hence the same inertia. Therefore, AB and B have the same inertia [34]. □

Since for the stability we want to guarantee that there are not any positive eigenvalues of the Jacobian, for the Corollary 4.3.0.1.1 we can limit our research to the eigenvalues of $M = (-I + \alpha A - 2\beta \mathbf{1}\mathbf{1}^T)$. And we can do that because A is symmetric and positive definite and M is symmetric and the corollary hypothesis are respected.

If $\{\mu_i\}_{i\in(1,\dots,L)}$ are the eigenvalues of $-I + \alpha A$ and $\{\lambda_i\}_{i\in(1,\dots,L)}$ the eigenvalues of M, then

$$\begin{cases} \lambda_1 = \mu_1 - 2\beta L \\ \lambda_i = \mu_i \qquad\qquad i = 2, \dots, L \end{cases} \tag{4.10}$$

Because $2\beta\mathbf{1}\mathbf{1}^T$ ensures that M is a *rank-1 perturbation* of $-I + \alpha A$, which affects only the first eigenvalue [19].

Now since A is a *tridiagonal Toeplitz matrix* [18], [48] its eigenvalues can be written as

$$\gamma_i = 2\cos\left(\frac{i\pi}{L+1}\right) \quad i = 1, \dots, L.$$

which leads to

$$\mu_i = -1 + 2\alpha\cos\left(\frac{i\pi}{L+1}\right) \quad i = 1, \dots, L.$$

Given the rank-1 perturbation, the highest eigenvalue of M is

$$\lambda_2 = -1 + 2\alpha \cos\left(\frac{2\pi}{L+1}\right).$$

Which finally leads to the critical value of the interaction strength that tells us when a chain of length L becomes stable:

$$\alpha_L = \frac{1}{2\cos\left(\frac{2\pi}{L+1}\right)}. \tag{4.11}$$

We can use such result only for chains of $L > 3$ because otherwise the critical value diverges. That is in accordance with the simulations that show that even if we impose $\alpha \to \infty$ we get a chain of length $L = 3$. We will observe in this thesis that in cooperation we never have final singlets and that applies also to $L = 2$ because the symmetry of chains makes sure that if there are two species surviving, they must have the same abundance. The effective competition we discussed before reduces the final component to a minimal connected geometry, but always requires some kind of abundances inequality when the interaction strength is high, and that cuts off singlets and length-2 chains. We now show that the analytical result in 4.11 matches the numerical simulations.

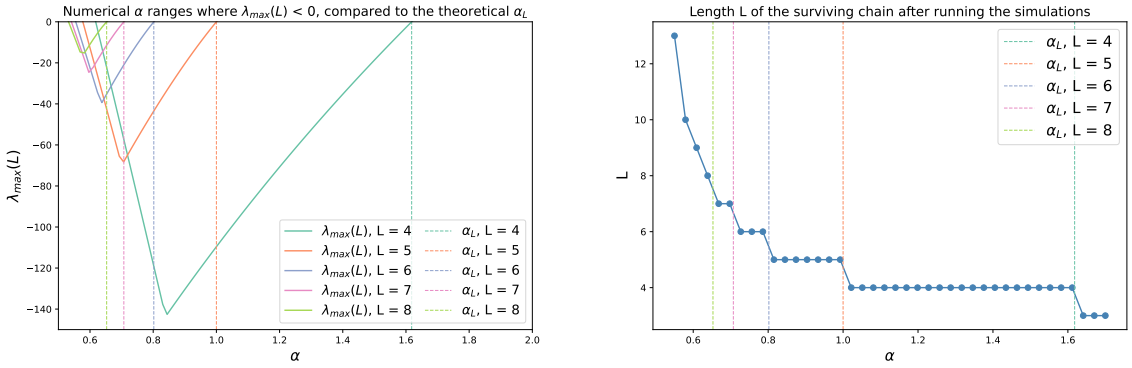


Figure 4.6: Left) Interval of stability for a length-L chain, showing how its maximal eigenvalue changes with α. This is done for $L \in \{4, ..., 8\}$. The maximal eigenvalue is computed numerically by plugging the fixed point of 4.8 in the Jacobian of 4.9 and finding the spectrum boundary. Only the range where $\lambda_{max} < 0$ is shown. Then we take the analytical α_L of 4.11 and we see that it coincides with the right-most α of such range. Right) We run the dynamics for many α and plot the length of the surviving chain L against α. The jumps coincide with the associated theoretical α_L.

We see from the second plot of Figure 4.6, that as the interaction strength approaches $\alpha_u = \alpha_d = 0.5$, the succession $\{\alpha_L\}_{L>3}$ gets denser and denser resembling a continuous

regime. Another interesting phenomenon is that the first plot suggests that the intervals where a chain of length L is stable are not disjoint for the different L. However, from the simulations we see that it is always the longest stable chain that survives. For instance, for $\alpha \in (0.8, 1)$, $L = 5$ and $L = 4$ are both stable but we observe that for different IC it is always the chain of length 5 that survives.

Moreover, we can look at the differences between competition and cooperation in the ring also from the economical perspective of the Gini index.

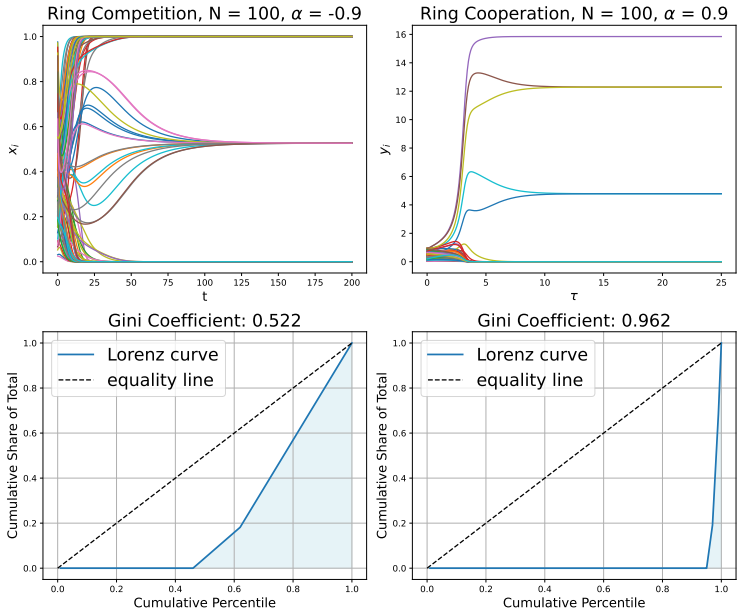


Figure 4.7: Cooperation (left column) and competition (right column) for the same absolute value of interaction strength $|\alpha| = 0.9$. On the second row the Gini index is computed from the Lorenz curve obtained with 100 species. This was done only to show that the CDF jump are grouped in pairs by amplitude, except for the highest-one that is alone because the chain has odd length.

From Figure 4.7 we can observe that in the cooperation case where only a length 5 chain survives, the wealth is concentrated only on a few species and the associated Gini index is proximate to 1. On the competitive case instead there are three possible values for the final wealth but 54% of the species survive and this is reflected on a much smaller Gini index.

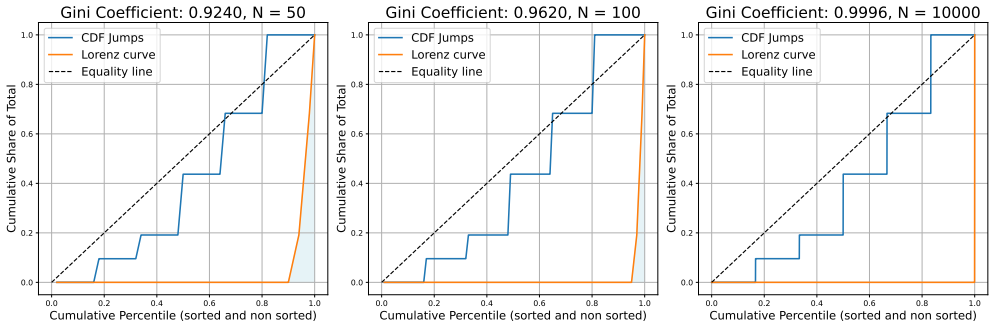


Figure 4.8: Lorenz curve and Gini coefficient for $\alpha = 0.9$ in the cooperative condensed case with increasing N = (50, 100, 10000). In blue, the CDF jumps where the percentile on the x-axis is not sorted (it is only equally spaced) to show the jumps in increasing order.

Figure 4.8 suggests that, since the number of surviving species is always the same, if we increase N the Lorenz curve is pushed to the right and the Gini coefficients approaches 1. The jumps in the CDF were included to show that the ratio $\frac{y_i}{\sum_i y_i}$ for the surviving species stays the same and it is given by the fact that the jump amplitudes are constant for increasing N. The location on the x-axis of those jumps is artificially made just to have them in increasing order. For the ring the Gini index can also be computed analytically because we know that the final component evolves as a chain that gets shorter as α is decreased. Moreover, the abundances of $\mathbf{y}^\star$ in Equation 4.8 are symmetric so the computation can be further simplified.

$$\hat{\mathbf{y}}^\star = (y_1^\star, \dots, y^\star_{\lceil \frac{L}{2} \rceil})^T.$$

$$\Delta^\star = \left(\hat{y}_1^\star, \quad \hat{y}_2^\star - \hat{y}_1^\star, \quad \dots \quad , \quad \hat{y}^\star_{\lceil \frac{L}{2} \rceil} - \hat{y}^\star_{\lceil \frac{L}{2} \rceil - 1}\right)^T.$$

$$G(\alpha) = 1 - \frac{1}{\langle y \rangle} \int_0^\infty (1 - F(y))^2 \, dy = 1 - \frac{N}{\langle \mathbf{y}^\star, \mathbf{1} \rangle} \sum_{i=0}^{\lceil \frac{L}{2} \rceil - 1} \left(\frac{L - 2i}{N}\right)^2 \Delta^\star_{i+1}. \tag{4.12}$$

Where G is the Gini index and $\langle \mathbf{y}^\star, \mathbf{1} \rangle$ is the dot product between the L-dimensional vector $\mathbf{y}^\star$ and the vector of ones with same length. Inside the sum I use only the first $\lceil \frac{L}{2} \rceil$ elements of $\mathbf{y}^\star$ (via $\Delta^\star$) because of the symmetry of the abundances in $\mathbf{y}^\star$. The dependence on α of the Gini index is given by $\mathbf{y}^\star = \mathbf{y}^\star(\alpha)$. Now let us compare this analytical result with the simulations.

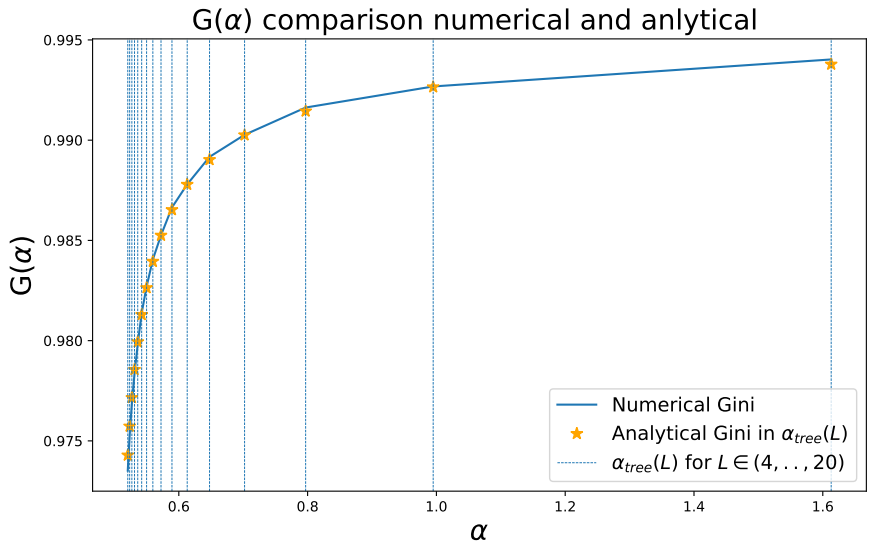


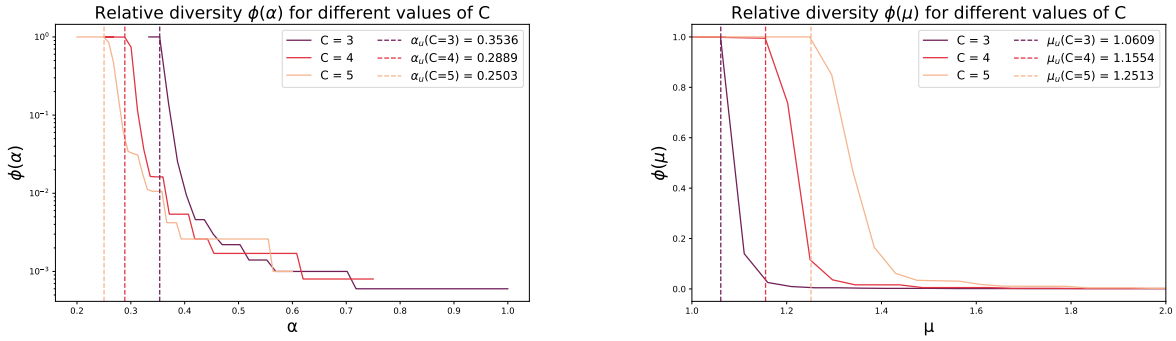
Figure 4.9: Gini coefficient for the critical values of α found in 4.11. In blue numerical results, in yellow analytical results from 4.12.

Equation 4.11 requires to know what is the length of the chain for any given α. Theoretically we do know it because we can partition the α support into the intervals dictated by Equation 4.11, but the Equation 4.12 does not assume that embedded information so it is more correct to define it as $G(\alpha) = G(\alpha, L)$. Other than a good matching between theoretical and numerical values, the figure confirms that as the interaction strength increases the wealth is localized on fewer and fewer species and the Gini coefficient approaches 1 when we take a fixed population size.

4.4 General degree C

In this section we extend the case of the ring to a C-regular random graph. As it was done for the competitive scenario, we start the investigation on the transitions by plotting the relative diversity against the interaction strength. Here we use $\mu = \alpha C$ instead of α to be able to compare in the same figure the curves for different values of C. Indeed in this way they all begin from $\mu = 1$, because, before it, species do not diverge. μ should not be carelessly interchanged with the interaction strength because it takes into account also the number of connections of each node. Hence it is better to refer to it as *social strength*. Between two graphs with the same interaction strengths but different average degrees C and $C' > C$, the social strength will be larger in the network with C'. Except for a few paragraphs where we consider different connectivities, in this chapter we will mostly consider a single C, and in

that case the two terms can be interchanged. The differences can be spotted from the figure below.



(a) (b)

Figure 4.10: a) Relative diversity against α for average degree $C = 3,4,5$ and $N = 10000$ in a plot where the y-axis was re-scaled with the logarithm to highlight the jumps. b) Relative diversity against μ, with the same choice of parameters without re-scaling of the y-axis. We do not re-scale here because we are interested to show the transition from 1 to *almost* 0.

Even if the comparison was made on two different scales of the y-axis, we can see that using the social strength is more intuitive in terms of the starting point and the overlapping of the curves. In Figure 4.10a the three relative diversity curves end in different x-points because we equally spaced the inputs in μ and not in α. We observe from Figure 4.10b that the single relative diversity curve presents three regimes:

- *mean field regime*: for $\alpha < \alpha_u$ (or $\mu < \mu_u$) the species all survive with the same abundance.
- *intermediate regime*: where the dimension of the surviving component decays. In this regime that size is dependent on the total size of the system N.
- *localization regime*: where the surviving species are localized in a connected component predictable via μ whose size is independent of N.

The first difference with the competitive case is that the last plateau is almost at $\phi = 0$. In particular, from the simulations with high values of social strength, we see that the final structures are *stars* with symmetric abundances around the central node. This follows the same property of *chains*, where the abundances where gradually symmetric around the central (or centrals, because in case of even length chains there are two) nodes.
The survival of a single minimal component is explained by the *effective competition* we introduced before. In the competitive case we have many disconnected components surviving because every node affects only the ones it is connected to. In particular, there we observed

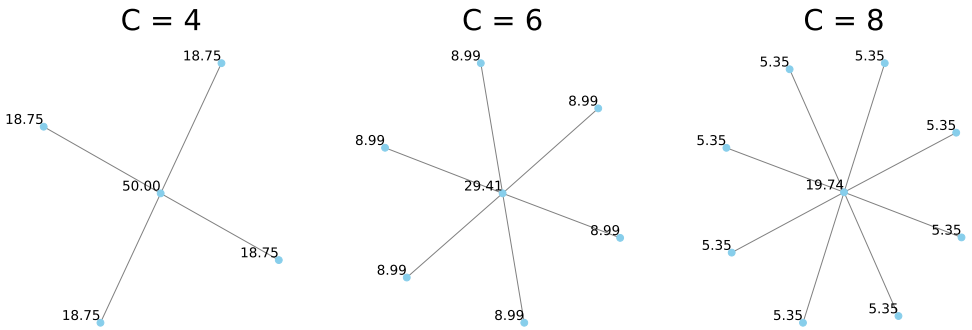
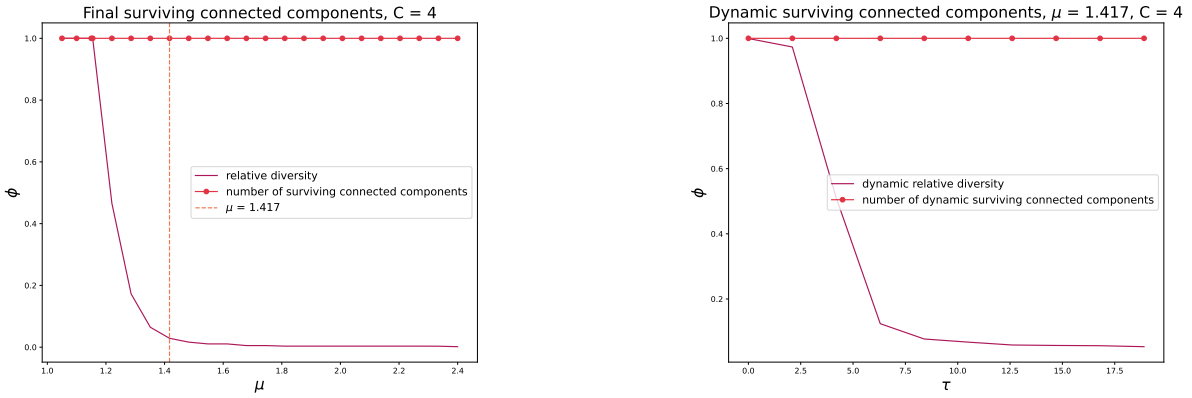


Figure 4.11: Star structures with abundances obtained by running the dynamics for $\mu = 3$. Increasing N the structure stays the same but the abundances are scaled.

a final plateau that suggested that 40% of the species were surviving. Here, that final global term from Equation 4.3 makes sure that every node influences all the other nodes, regardless of their connection. This global influence leads to another interesting phenomenon: even when we decrease the social strength and more species survive, the number of connected components for the survivors is always one. Also dynamically the extinction works by removal (subtraction); indeed for a fixed μ at every time from 0 to the t_{eq}, the surviving species are always connected. We will later investigate what are the properties of the network of the extinct species.



(a) (b)

Figure 4.12: (a) For all μ, also the small ones with a lot of surviving species, the number of connected components among the surviving species is always one. (b) Fixing $\mu = 1.417$, dynamically the number of surviving species decreases with time, but they are always grouped into a single connected component

Now that we know what the final structure is, we want to see how that evolves as we decrease the social strength.

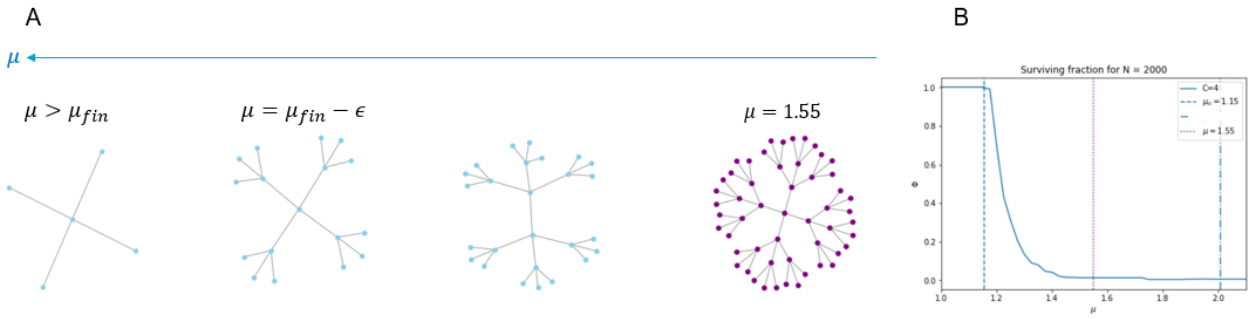


Figure 4.13: A) from left to right μ decreases and the star grows as a tree with C=4. B) In purple, relative diversity associated to $\mu = 1.55$ for $N = 2000$

From Figure 4.13.B we can observe that the tree associated with $\mu = 1.55$ corresponds to a level of relative diversity proximate to zero (it is 53/2000). Hence, we do not really know if this tree behaviour is maintained also for μ in the part of the *intermediate regime* that is closer to μ_u.
Let us start by characterizing these trees.

Definition 4.4.1 (C-regular tree)**.** A C-regular tree is an acyclic graph where all nodes have degree C. It is defined by the adjacency matrix A_{reg} and it has height L, where the root node corresponds to $L = 0$.

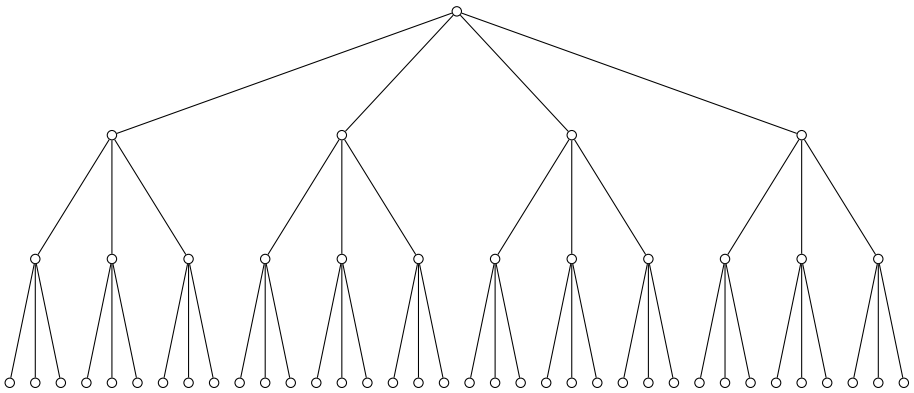


Figure 4.14: 4-regular tree with $L = 3$

If l is the index of the layer, $S_C(l)$ is the size of the tree until a total completion of layer l, therefore $S_C(L)$ is the total size of the tree when the last layer is complete. This implies that the size of the layer l is $S_C(l) - S_C(l-1)$. $S_C(\mu)$ is the size of the tree associated with social strength μ which can be $S_C(L-1) < S_C(\mu) < S_C(L)$ if the layer is incomplete.

The formula to compute the size of a tree with complete final layer L is:

$$S_C(L) = 1 + C\sum_{i=0}^{L-1}(C-1)^i = 1 + C\frac{(C-1)^L - 1}{C-2}. \tag{4.13}$$

In the previous chapter we saw that as we decrease μ, the length of the stable chain increases and we could analytically find some critical α_L to identify this transition. Similarly we can find α_L for trees to understand when we switch from a tree with height $L-1$ to one with height L. To do that we need another graph theory definition.

Definition 4.4.2 (Q-ary tree)**.** A Q-ary tree is a tree where all nodes have Q children. It is characterized by uniform first degree progeny. It is defined by the adjacency matrix A_{ary} and it has height L', where the root node corresponds to $L' = 0$.

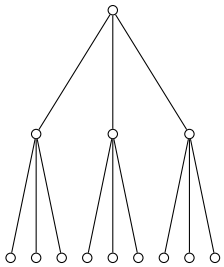


Figure 4.15: 3-ary (ternary) tree with $L' = 2$

Definition 4.4.3 (Second type Chebyshev Polynomials)**.** The second type Chebyshev polynomials $U_n(x)$ are defined by the recurrence relation:

$$U_0(x) = 1,$$
$$U_1(x) = 2x,$$
$$U_n(x) = 2xU_{n-1}(x) - U_{n-2}(x) \quad \text{for} \quad n \geq 2.$$

The roots of the Chebyshev polynomial $U_n(x)$ are given by:

$$x_k = \cos\left(\frac{k\pi}{n+1}\right) \quad \text{for} \quad k = 1, 2, \ldots, n.$$

The non-zero eigenvalues of the adjacency matrix A_{ary} of a Q-ary tree can be determined from the roots of the second type Chebyshev Polynomials $U_{L'+1}(x)$ [50], [21]. Those eigenvalues are $\{\gamma_k\}_{k\in(1,\ldots,L'+1)}$ and are obtained as

$$U_{L'+1}\left(\frac{\gamma}{2\sqrt{Q}}\right) = 0.$$

Then we have

$$\frac{\gamma_k}{2\sqrt{Q}} = \cos\left(\frac{k\pi}{L'+2}\right) \quad \text{for} \quad k = 1, 2, \ldots, L'+1.$$
$$\gamma_k = 2\sqrt{Q}\cos\left(\frac{k\pi}{L'+2}\right) \quad \text{for} \quad k = 1, 2, \ldots, L'+1.$$

We can observe from Figure 4.14 that a C-regular tree is composed of a number C of $(C-1)$-ary identical trees attached to a root node. We then analyze the case where $Q = C - 1$ and

$L = L' + 1$. Also here we start counting from zero, so there are $L + 1$ layers in total. This means that its adjacency matrix takes the form

$$A_{\text{reg}} = \begin{pmatrix} 0 & & \mathbf{u^T} & \\ & \begin{pmatrix} A_{\text{ary}} & 0 & 0 & 0 \\ 0 & A_{\text{ary}} & 0 & 0 \\ 0 & 0 & A_{\text{ary}} & 0 \\ 0 & 0 & 0 & A_{\text{ary}} \end{pmatrix} & & \\ \mathbf{u} & & & \end{pmatrix} = \begin{pmatrix} 0 & \mathbf{u^T} \\ \mathbf{u} & B_{\text{ary}} \end{pmatrix} \tag{4.14}$$

Where we took the example of Figures 4.14, 4.15 with $C = 4$. Then $\mathbf{u}$ is a vector defined as $\mathbf{u} = (1, 0, \ldots, 0, 1, 0, \ldots, 0, 1, 0, \ldots, 0, 1, 0, \ldots, 0)^T$, with ones only at the level of the first row of each C-ary matrix (here at positions $1, L'+2, 2L'+2, 3L'+2$); everything else is zero. For the sake of the theory, the following results are not restricted to those specific positions of the 1-entries of $\mathbf{u}$, what counts is only the number and that is always C. It is entailed that the matrix A_{reg} is formed by adding a rank-1 perturbation to B_{ary}.

The spectrum of B_{ary} is given by $\{\gamma_k\}_{k\in(1,\ldots,L'+1)}$ with multiplicity C (the other eigenvalue is zero with multiplicity $C(S_C(L') - L' + 1)$). Then the rank-1 perturbation adds an extra eigenvalue.

To find the new eigenvalue, we analyze the characteristic equation of A_{reg}. Its characteristic polynomial for A_{reg} is given by:

$$\det(A_{\text{reg}} - \theta I) = \det \begin{pmatrix} -\theta & \mathbf{u}^T \\ \mathbf{u} & B_{ary} - \theta I \end{pmatrix}$$

Using the formula for the determinant of a block matrix, we get:

$$\det(A_{\text{reg}} - \theta I) = -\theta \cdot \det(B_{ary} - \theta I) - \mathbf{u}^T B_{ary}^{-1} \mathbf{u} \cdot \det(B_{ary} - \theta I).$$

Theorem 4.4.0.1 (Block Matrix Determinant Formula)**.** *Let M be a block matrix of the form:*

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

where A and D are square matrices. If A is invertible, then the determinant of M is given by:

$$det(M) = det(A) \cdot det(D - CA^{-1}B).$$

If D is invertible, the determinant can also be expressed as:

$$det(M) = det(D) \cdot det(A - BD^{-1}C).$$

Where we used the second form, assuming that $B_{ary} - \theta I$ is invertible (only B_{ary} is singular). Simplifying the characteristic equation, we obtain:

$$\det(A_{\text{reg}} - \theta I) = \det(B_{ary} - \theta I) \left(-\theta - \mathbf{u}^T \left(B_{ary} - \theta I\right)^{-1} \mathbf{u}\right).$$

The eigenvalues are solutions to $\det(A_{\text{reg}} - \theta I) = 0$. Thus, either:

- $\det(B_{ary} - \theta I) = 0$ (which gives the original eigenvalues γ_k with multiplicity C), or
- $-\theta - \mathbf{u}^T \left(B_{ary} - \theta I\right)^{-1} \mathbf{u} = 0$, which gives the new eigenvalue θ_{new}.

The new eigenvalue θ_{new} satisfies:

$$\theta_{\text{new}} = -\mathbf{u}^T \left(B_{ary} - \theta_{\text{new}} I\right)^{-1} \mathbf{u}$$

Since $\mathbf{u}$ has only C (in the example $C = 4$) non-zero entries, this reduces to:

$$\theta_{\text{new}} = -\sum_{i=1}^{C} \frac{1}{\gamma_1 - \theta_{\text{new}}} = -\frac{C}{\gamma_1 - \theta_{\text{new}}}.$$

Therefore, we solve:

$$\theta_{\text{new}}(\gamma_1 - \theta_{\text{new}}) = -C$$

The quadratic equation for θ_{new} is:

$$\theta_{\text{new}} = \frac{\gamma_1 \pm \sqrt{\gamma_1^2 + 4C}}{2}$$

The perturbation introduced by the rank-1 update typically adds a new eigenvalue that lies outside the original spectrum of B_{ary}, hence we choose the positive sign. Therefore, the new eigenvalue is:

$$\theta_{\text{new}} = \frac{\gamma_1 + \sqrt{\gamma_1^2 + 4C}}{2} > \gamma_1.$$

Finally, remembering $L = L' + 1$, the eigenvalues of A_{reg} are:

$$\begin{cases} \theta_0 = \frac{\gamma_1 + \sqrt{\gamma_1^2 + 4C}}{2} = \frac{\theta_1 + \sqrt{\theta_1^2 + 4C}}{2} \\ \theta_k = \gamma_k = 2\sqrt{C-1}\cos\left(\frac{k\pi}{L+1}\right) \quad \text{for} \quad k = 1, \ldots, L \quad \text{with multiplicity C} \end{cases}$$

From which we can observe that $\theta_0 > \theta_1 > ... > \theta_L$.
And the eigenvalues of $-I + \alpha A_{reg}$ are:

$$\begin{cases} \lambda_0 = -1 + \alpha \frac{\lambda_1 + \sqrt{\lambda_1^2 + 4C}}{2} \\ \lambda_k = -1 + 2\alpha\sqrt{C-1}\cos\left(\frac{k\pi}{L+1}\right) \quad \text{for} \quad k = 1, \dots, L \quad \text{with multiplicity C} \end{cases} \tag{4.15}$$

Once again $\lambda_0 > \lambda_1 > ... > \lambda_L$.

Moreover, the Equations 4.8 and 4.9 are true for every C-regular graph and in $M = -I + \alpha A - 2\beta \mathbf{1}\mathbf{1}^T$ the perturbation affects only the first eigenvalue ensuring that $\lambda_1 > \lambda_0$ where $\{\lambda_l\}_{l \in (0,\dots,L)}$ are the eigenvalues of M.

Studying the stability of a height-L tree with complete last layer is then reduced to studying the sign of the maximal eigenvalue

$$\lambda_1(L) = -1 + 2\alpha\sqrt{C-1}\cos\left(\frac{\pi}{L+1}\right). \tag{4.16}$$

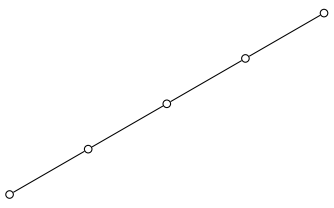
which finally leads to the critical social strength

$$\begin{aligned} \alpha_{tree}(L) &= \frac{1}{2\sqrt{C-1}\cos\left(\frac{\pi}{L+1}\right)}. \\ \mu_{tree}(L) &= \frac{C}{2\sqrt{C-1}\cos\left(\frac{\pi}{L+1}\right)}. \end{aligned} \tag{4.17}$$

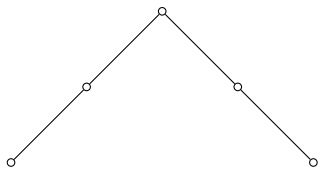
Before checking the simulations we can make some observations that confirm this result. In particular, in the case of the ring condensing into a chain we have,

$$\alpha_{chain}(L) = \frac{1}{2\cos\left(\frac{2\pi}{L_{chain}+1}\right)} = \frac{1}{2\cos\left(\frac{2\pi}{(2L_{tree}+1)+1}\right)} = \frac{1}{2\sqrt{2-1}\cos\left(\frac{\pi}{L_{tree}+1}\right)}$$

because $C = 2$ and $L_{chain} = 2L_{tree} + 1$.



(a) Chain with $L_{chain} = 5$



(b) 2-regular tree with $L_{tree} = 2$

Moreover,

$$\alpha_{tree}(L) \xrightarrow{L\to\infty} \frac{1}{2\sqrt{C-1}} = \alpha_u$$

That is hinting that the only way to continuously arrive to the mean field regime, where everyone survives, only via trees growth, is having $N \to \infty$ ($L < N$). It is already suggesting that, if we consider the finite size case, this tree growth must be broken for a certain μ. For the moment let us focus on high values of the social strength where we have trees.

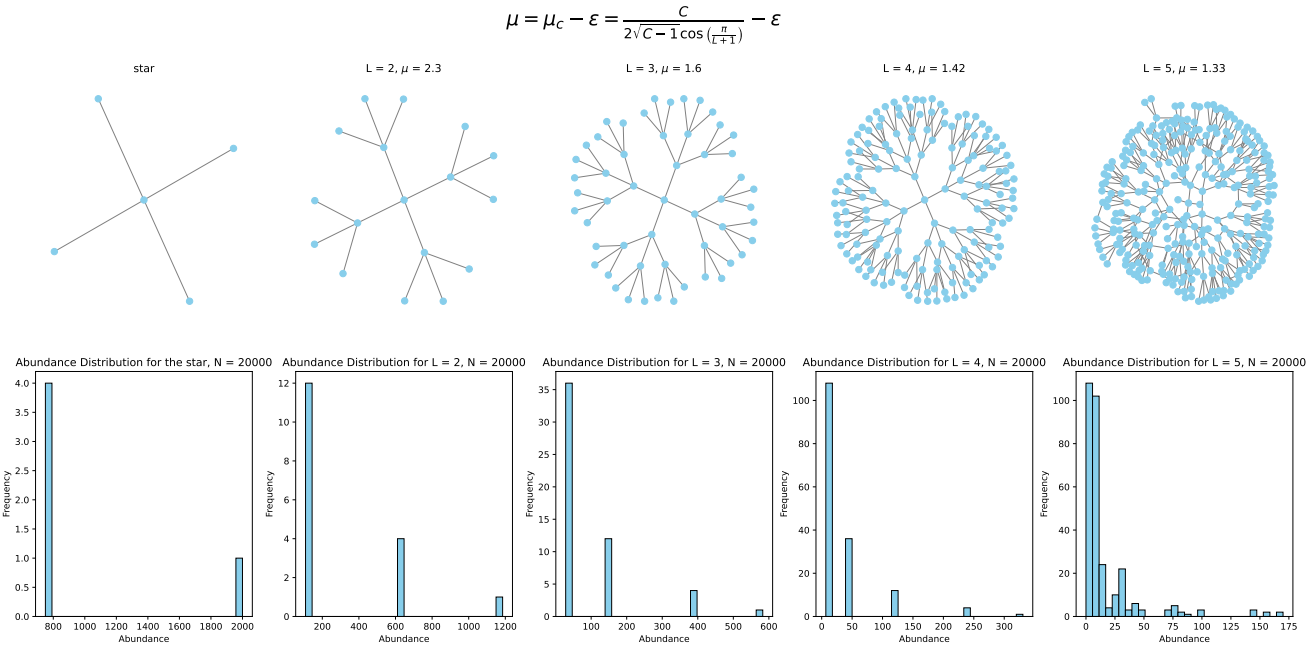


Figure 4.17: Tree structure for $\mu = \mu_c(L) - \epsilon$ with increasing L. Below we include the abundances distribution. Until $L = 4$ it follows the analytical prediction, then from $L = 5$ loops are generated and the tree structure with symmetric abundances is abandoned.

We observe that if we take $\mu = \mu_{tree}(L) - \epsilon$, the tree with height L is stable. Then as soon as we surpass $\mu_{tree}(L)$, it becomes unstable and a smaller tree appears for $L < 5$. Now we introduce another graph theory metric that is useful to characterize the length or the extension of a graph.

Definition 4.4.4 (Graph Diameter)**.** The diameter of a graph $G(V, E)$, denoted as $\text{diam}(G)$, is defined as the longest distance between any two vertices in G. It can be expressed as:

$$\text{diam}(G) = \max\{d(u, v) \mid u, v \in V(G)\},$$

where $d(u, v)$ represents the distance (i.e., the minimal number of edges) between vertices u and v in the graph G. If u and v are not connected then $d(u, v) = \infty$.

In case G is a regular tree with height L, then $\text{diam}(G) = 2L$ or $\text{diam}(G) = 2L - 1$ if all the last layer nodes belong to the progeny of the same child of the root node. To prove the truthfulness of the analytical result we also compared it to the even jumps in the diameter of the surviving tree.

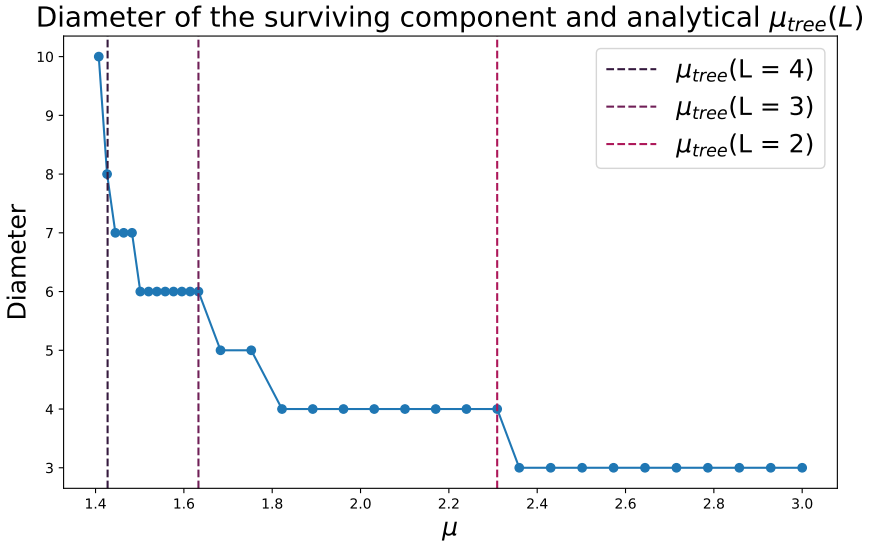


Figure 4.18: Diameter of the surviving component over different values of the social strength μ, for $C = 4$ and $N = 10000$. The dotted lines are the analytical values of $\mu_{tree}(L)$ found in 4.17.

From Figure 4.18 we see that the analytical jumps are correctly avoiding the odd diameter values because in a tree a diameter is odd only when the last layer is not complete. Moreover, this check was conducted only on small values of L because for larger heights the distance between $\mu_c(L)$ and $\mu_c(L+1)$ gets smaller and smaller and the noise in the simulations cannot completely capture it. However, like in the case $L = 5$ from Figure 4.17, there is a value of μ where we must stop considering the surviving component as a tree; we will soon delve into it. The intuition tells us that if the tree continuously grows it arrives to a perfect C-regular tree with $S_C(\mu_u) = N$ and this cannot be the case because we initially generate a generic C-regular graph, which is locally tree like (finite number of short cycles even when $N \to \infty$ [9]) but is very likely to present long cycles.
With the critical $\mu_c = \mu_{tree}(L)$ found in Equation 4.17 and from the Equation 4.13 that tells us the size of the tree for a given height L, the theoretical relative diversity for these μ_c

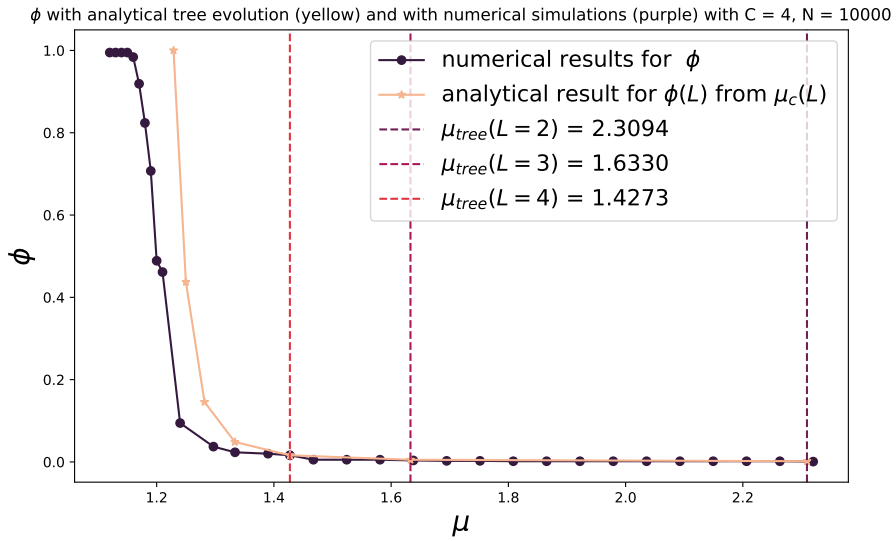


Figure 4.19: Relative diversity against social strength. The numerical simulations are run for 30 values of μ. The theoretical values of $\phi(L)$ are for $L \in (2, 3, 4, 5, 6, 7, 8)$ because afterwards larger L would produce $\phi(L) > 1$. We see that from $L = 5$ the tree does not reflect the nature of the surviving component.

can be calculated for $L \geq 2$. From Figure 4.19 we can notice that the tree evolution stops mirroring the dynamics of the system for $L \geq 5$ where the two curves stop overlapping. The theoretical relative diversity grows to infinity as the size of the tree becomes larger than N (4.13 does not depend on N). Indeed, for $L \geq 9$, the theoretical value of Φ is also larger than one and that is because $S_4(9) > N = 10000$. In particular, the *localization regime* continues until long cycles appear and we start having this non-ergodic expanded states. To prove the occurrence of these long cycles we can look at the diameter as well.

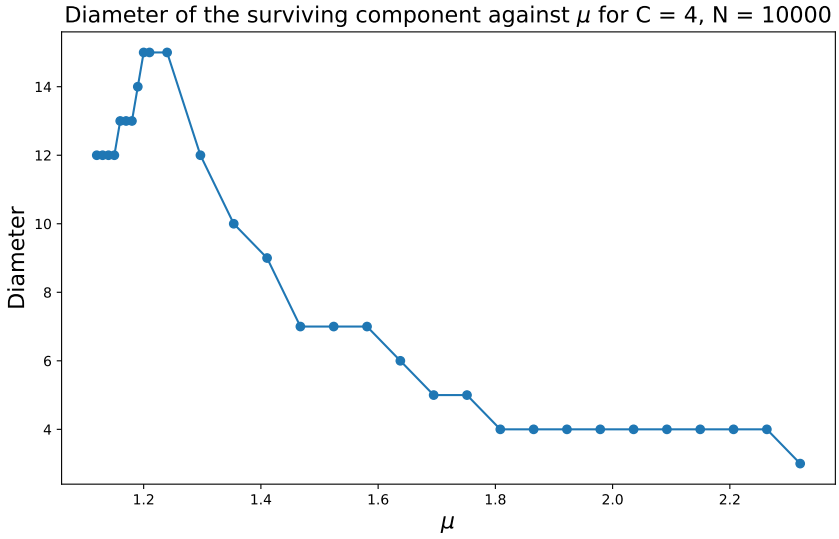
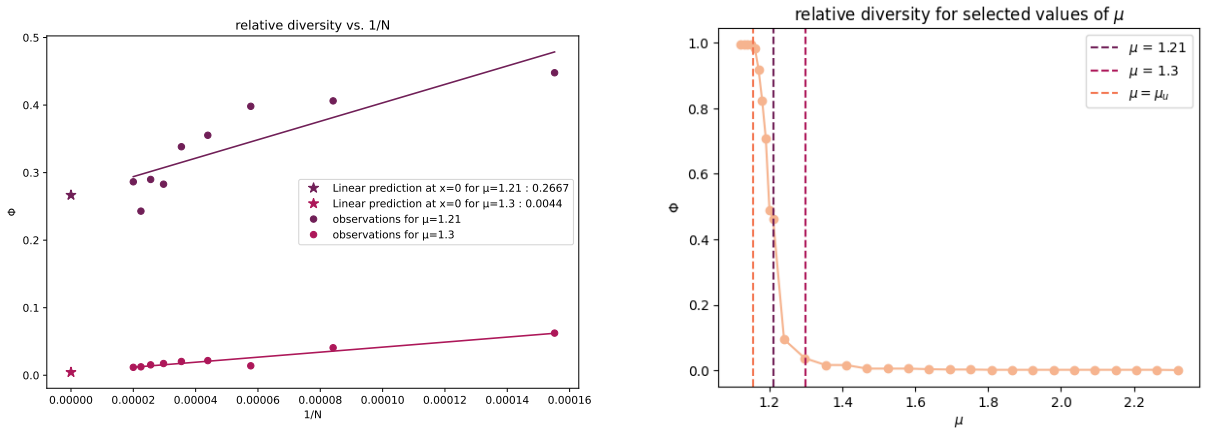


Figure 4.20: Diameter of the surviving component against μ after running the dynamics with $C = 4$, $N = 10000$. We want to show that, from right to left, the diameter does not monotonously increase, therefore the tree structure must be broken at a certain point. Note that the point where it changes direction does not tell us when that happens. It is just a proof that it must happen.

We can observe that the diameter is not monotonously decreasing with μ, and this is explained by the appearance of nodes which are connected to the tree at a certain level, such that the new diameter is lower than $2L$. Keep in mind that this inflection point does not occur exactly at the μ where the tree evolution stops. And the reason is that even if long cycles appear, the number of surviving species significantly increases and hence the longest distance between two nodes too. After observing these results we can re-name the *intermediate regime* in a way that stresses what breaks this tree growth: *long cycles regime*.

Until now we have run the simulations with finite N, now we want to investigate how this regimes characterization changes if we take $N \to \infty$. In particular we want to see if by increasing N the *intermediate regime* disappears and the relative diversity resembles an Heaviside step function. Since for computational limitations we cannot run the dynamics for extremely high N, we plot ϕ against $1/N$ and we regress the line until $1/N = 0$.

It is enough to take two values of μ in the transitory regime, one closer to the *localization regime* ($\mu = 1.3$) and one closer to the *mean field regime* ($\mu = 1.21$). In both cases we expect long cycles to arise breaking the tree property.



(a) (b)

Figure 4.21: (a) Linear regression for ϕ against $1/N$ for $\mu = 1.3$ and $\mu = 1.21$ with N equally spaced from 1000 to 50000. (b) ϕ for $\mu = 1.3$ and $\mu = 1.21$ for fixed $N = 10000$ to show where the values are located in the *long cycles regime*.

We observe that the projection at zero is positive for both values of μ, excluding the possibility of a step relative diversity function. Moreover, we also observe that as μ approaches the mean-field, $\phi_{\frac{1}{N}=0}(\mu)$ is higher because we get closer to the scenario where they all survive.

We then question whether there is an exponent $\xi(\mu)$ such that,

$$\Phi(N, \mu) = A\frac{1}{N^{\xi(\mu)}} + \Phi_\infty(\mu) \tag{4.18}$$

We introduced the extra term Φ_∞ to make a generalization and comment the asymptotic fitted value. We then conduct a non-linear regression to find the $\xi(\mu)$ such that 4.18 is verified. The goal is to make a full characterization of $\xi = \xi(\mu)$.

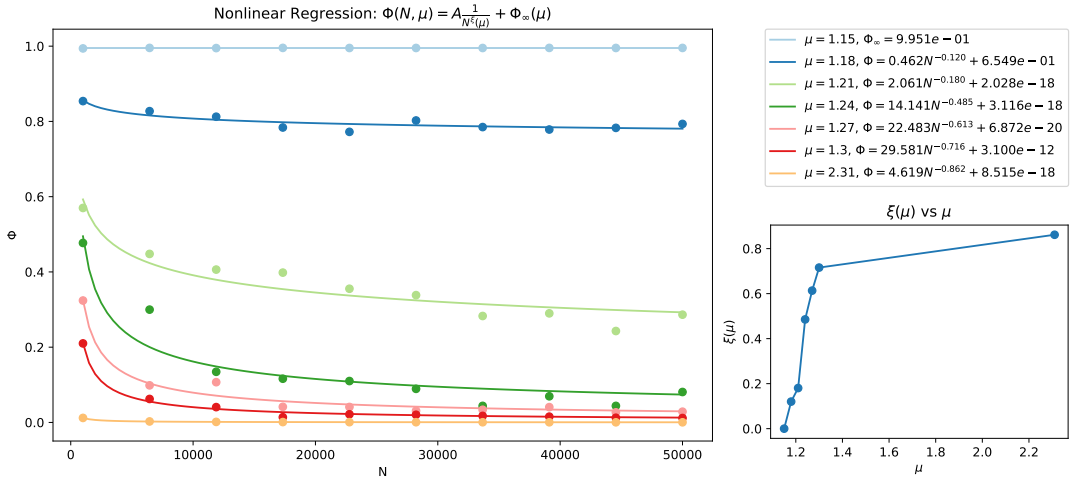


Figure 4.22: Non-linear regression on the relative diversity for different values of μ. From that the coefficient $\xi(\mu)$ is extracted and plotted against μ.

From the Figure we can observe that $\xi \to 0$ as μ gets smaller and approaches the *mean field regime* , hence for small μ, $\frac{1}{N^{\xi}}$ drops to zero very steeply, while for μ closer to the *localization regime* $\xi \to 1$. Moreover, for the regime we are interested in, $\phi_\infty(\mu) \approx 0$, which suggests that, for $N \to \infty$, ϕ completely depends on N. We could have then used $\phi(N,\mu) = AN^{-\xi(\mu)}$.

Now let us introduce a measure of concentration and quantify the condensation of the states. Remember that a condensed state does not depend on N so if we take the abundances normalized by the system size and we take the large N limit, they do not collapse to zero.

Definition 4.4.5 (Herfindahl–Hirschman index)**.** It is defined as,

$$HHI = \sum_{i=1}^{N} MS_i^2 = \sum_{i=1}^{N} (\frac{y_i}{Y})^2$$

$$Y = \sum_{i=1}^{N} y_i$$

The HHI tends to zero when wealth is distributed between many species and to one when all the wealth belongs to a single species. It is an economical index that takes into account the *market share* MS_i of species i. If all the MS_i are of order $1/N$ then $HHI \sim 1/N$ and tends to zero for large N. On the other hand, if at least one MS_i remains finite when

$N \to \infty$, then HHI will also be finite and we have condensation. It can be associated to the *inverse participation ratio* used to measure condensation in [36], [12]. In an extended state, the MS_i are roughly of the order $\frac{1}{N}$, so the HHI becomes small, of the order $\frac{1}{N}$. In a localized state, where the wealth is concentrated on a small number of sites, the HHI becomes larger. If the state is entirely localized on a single site, the HHI equals 1. We can observe these properties in Figure 4.23.

We then want to measure how HHI scales with N, in particular we want to see what is the $\psi = \psi(\mu)$ such that,

$$HHI(N, \mu) = A \frac{1}{N^{\psi(\mu)}}$$

$$\log(HHI(N, \mu)) = \log(A) - \psi(\mu) \log(N) \tag{4.19}$$

As for the relative diversity, also here we run the non-linear regression and observe that the term $HHI_\infty \approx 0$, hence we directly provide the linear regression on $log - log$ as in 4.19.

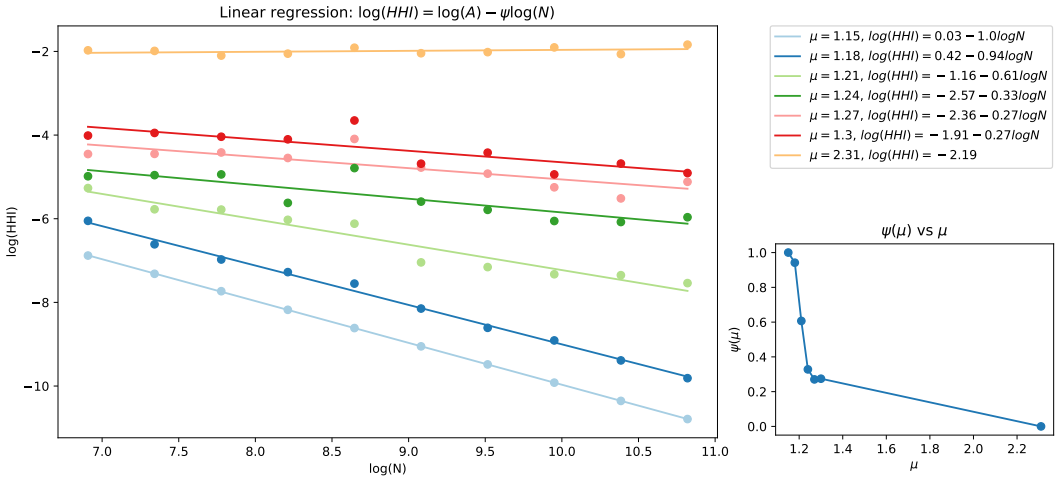


Figure 4.23: Linear regression on the log-transformation of the Herfindahl–Hirschman index. From that the coefficient $\psi(\mu)$ is extracted and plotted against μ.

Looking at Figure 4.23, we notice how $\psi \to 0$ as μ increases, suggesting that the HHI index does not depend on N in the *localization regime* and all the wealth (or abundances) is concentrated in just a few species. This entails that in such regime there is an oligarchy where only a few species detain most of the wealth. On the other hand, as μ approaches the *mean field regime*, $\psi \to 1$ and the wealth is heterogeneously distributed among many species because $HHI \to 0$.

All these results suggest the occurrence of the last phenomenon that we want to investigate for the homogeneous degree cooperative case: the *Anderson Localization*. It describes the regime where the abundances distribution does not scale linearly with N and it is localized in just a few elements (see 1.1). To do that we have to introduce some measures.

Definition 4.4.6 (Shannon Entropy)**.** It is defined as,

$$H(N) = \mathbb{E}[log(\frac{1}{MS})] = -\sum_{i=1}^{N} MS_i \log(MS_i)$$

where MS_i stands for market share of species i.

In particular, calling *state* the vector of final abundances $(y_1, ..., y_N)$, the Shannon Entropy gives information on the states distribution. If H is high, it indicates that the wealth is more evenly distributed among the nodes (mean field). Each node has a relatively similar share of the total wealth. On the contrary if H is low, it suggests that the wealth distribution is more uneven, with some nodes holding significantly more wealth than others. This is the case as we move towards the high values of μ.

Definition 4.4.7 (Fractal Exponent)**.** For a fixed μ the fractal exponent D is defined via the asymptotic Shannon entropy as,

$$H(N) \sim D \log(N), \quad \text{for} \quad N \to \infty$$

In presence of this localization as $\log N$ increases, you would expect D to decrease, indicating that wealth is becoming increasingly localized.

If we are in the mean field regime where the abundances are uniform then we have $H(N) = logN$ and $D = 1$. In general, high D suggests that the state is distributed over a large number of sites, characteristic of an extended state. The distribution has a fractal-like structure that spans the entire system.

But also in the competitive case where we have that the number of surviving states is a fraction of N, $D = 1$. For example, take a system like the ring where 50% of the species are surviving. There we have $H(N) = O(log(\frac{N}{2})) = O(logN - log2)) = O(log(N))$ when N is big. So also here $D = 1$. Every time the fraction and not the number of surviving states is conserved, then $D = 1$.
Now let us relate it to the behaviour showed in 4.20 where there are long cycles and the diameter decreases (before entering the tree regime). It is a known result that the average length of a long cycle in a C-regular graph is of $O(logN)$ [33], then in the part of the transitory regime where we do not have trees we have $D = 1$ or at least $D > 0$. On the

other hand, when trees appear for larger values of μ the number of states is of $O(S(L))$ and $H = O(log(S(L)))$ where $S(L)$ is the size of the tree. That does not depend on N and hence $D = 0$ there. Localization indeed means that the states distribution $\{MS_i\}_{i=1,...,N}$ is concentrated only on a few number of states, mainly $S(L)$. So if I find the μ such that $D = 0$, then we can say when we transition from a cyclic surviving graph to a surviving tree.

This is what was done by *Kravstov et al.* in [28] to prove the presence of the Anderson localization on the *Bethe-lattice*, where their *disorder strength* W, corresponds to our social strength. A Bethe lattice is an infinite, loopless, connected graph where each node is connected to exactly C neighbors, and the number of nodes increases exponentially with the distance from any given node.

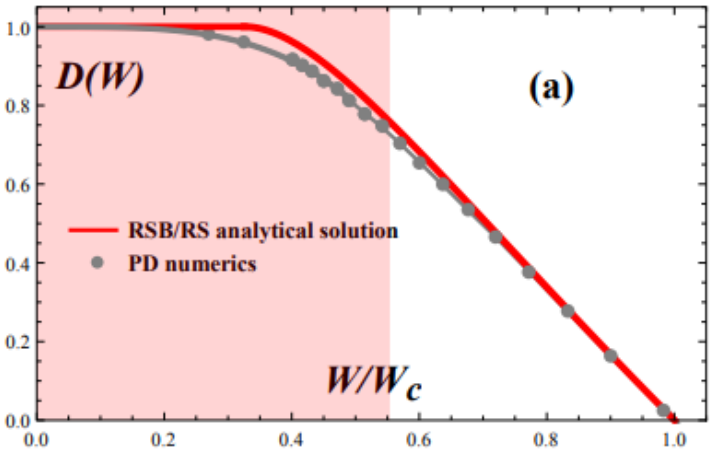


Figure 4.24: On the x-axis W/W_c is taken to have it between zero and one, as W_c is the maximal disorder strength. The red area represents a one-step replica symmetry breaking that predicts a continuous ergodic transition, but it is not in the scope of our research

What is important for us about Figure 4.24 is the monotonous decay of the fractal exponent from one to zero. Our goal now is to replicate that $D(\mu)$ against μ plot to check if we have monotonous decay as well and finally prove that we have Anderson Localization for the critical $\mu_{AL} : D(\mu) = 0 \quad \forall \mu > \mu_{AL}$
Therefore, for each μ we found the fractal exponent as the coefficient of the regression of $H(N)$ against $log(N)$ with H_0 that is the additional term.

$$H(N, \mu) = D_{reg}(\mu) \log(N) + H_0.$$

Where $D_{reg}(\mu)$ is the coefficient obtained from the regression. Then plot $D_{reg}(\mu)$ against μ

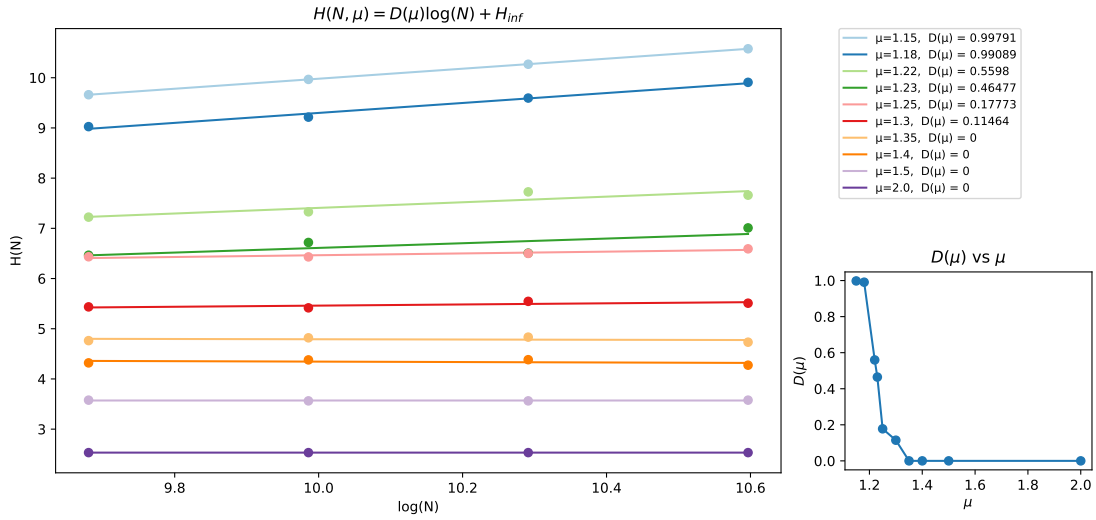


Figure 4.25: Linear regression on the Shannon Entropy against $log(N)$. From that the coefficient $D(\mu)$ is extracted and plotted against μ. It appears to be monotonously decreasing and between 1 and 0.

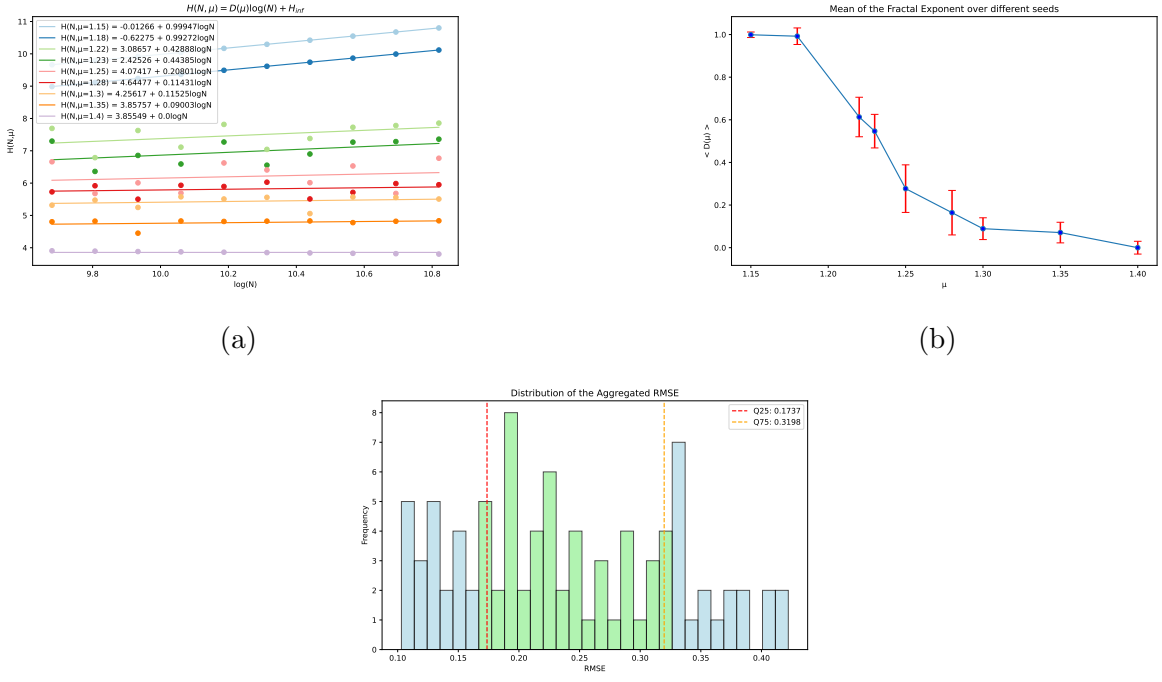
Figure 4.25 confirms the *Anderson localization* and suggests that $1.3 < \mu_{AL} < 1.35$ that is in accordance also with the previous results that showed that the last tree is the one associated to height $L = 4$.

So far, via the observations obtained with one simulation, we have shown that as μ increases $\psi(\mu)$ and $D(\mu)$ tend to zero while ξ tends to one. Another property of interest is the monotonicity of the curves. To validate our results we conducted a statistical analysis over the regressions made on many simulations. In this study we are not dealing with a real dataset where the observations are subject to endogenous or exogenous noise, like the evolution of two stocks on the financial exchange for example. We are studying a synthetic environment over which we have total control and the noise is only self-generated. Hence we avoided statistical tests like the *Spearman's Test* to validate the monotonicity hypothesis. Here we are dealing with linear and non-linear regressions with limited observations, hence the most useful metric is the Root Mean Squared Error (RMSE).

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{4.20}$$

where n represents the number of observations, y_i denotes the actual value of the i-th observation, while $\hat{y}_i$ represents the predicted value obtained by fitting the regression. The

difference $y_i - \hat{y}_i$ is the error or residual for the i-th observation. We take the square root of the standard MSE to have the same units as the target variable y. The best characteristic of this measure is its interpretability; indeed we can see it as the average magnitude of the prediction errors.



(a) (b)

(c)

Figure 4.26: a) Linear regression of the Shannon Entropy with a single seed. 10 observations for each of the 9 values of μ. b) Mean and standard deviation of the fractal exponent. Simulations on 10 different seeds were run, obtaining 10 different values of $D(\mu)$ for each μ. From them, the mean is shown and the standard deviation is used to obtain red error bars. c) Distribution of the 90 RMSE points, with 25th and 75th quantiles highlighted in green

N was equally log-spaced from 16000 to 50000 to obtain 10 observations every time. We then used 10 seeds and 9 values of μ because of the computational costs of running the dynamics for such a big size of the system. Instead of doing the regression on the means of the observations over the seeds, we take each regression separately, take the 9 fractal exponents and then average them. Moreover, we did not show the mean of the RMSE distribution but we showed the 25th and 75th quantiles to give more information. The mean would have been misleading for a bimodal distribution for example. In order to have a larger dataset we conducted a non parametric bootstrapping with 100 samples taken from resampling our data

and that led us to the following fractal exponent curve that resembles 4.24 in its monotonous direction from 1 to 0.

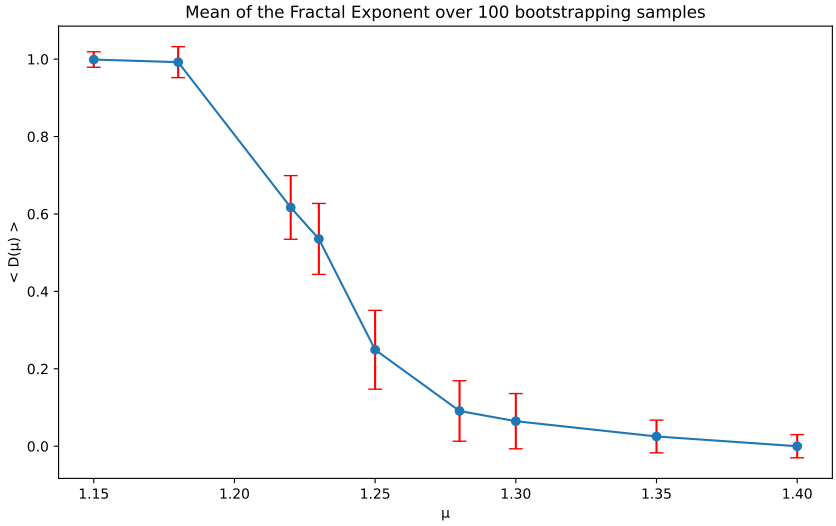


Figure 4.27: Mean and standard deviation of the fractal exponent for each of the 100 samples obtained via non-parametric bootstrapping.

The non-parametric bootstrapping is not as informative as 100 independent samples (which would have been too computationally expensive) because multiple outcomes rely on partially equal data, but nevertheless provides an additional guarantee of the goodness of the regressions. Moreover, the repeated resampling process allows for a more thorough exploration of possible outcomes, ensuring that the results are not overly dependent on a single dataset configuration. By generating multiple resampled datasets, bootstrapping captures the natural variability within the original data, making it possible to assess the stability and consistency of the regression outcomes. This process adds robustness by reducing the risk of overfitting to a particular dataset and providing more reliable estimates of uncertainty, which strengthens confidence in the validity of the conclusions drawn from the analysis.

4.4.1 Effects of increasing C

The reason why all the simulations of this chapter were conducted with $C = 4$ is that last time of the dynamics had to be manually selected for every value of social strength. The cost of using a integrating function that does not scale exponentially with the population size is that the last time has to be carefully chosen and it crashes if you choose an upper bound that is distant from the true convergence time. We applied a non linear interpolation

but the integrating function is so sensitive that also that has to be carefully tuned.

As we increase C, the critical μ_u shifts to the right, as well as the beginning of the localization regime. Figure 4.21a suggests that the *intermediate regime* is expanded. It is left to see whether that is because the *localization regime* is shrinked. In particular, we can replicate Figure 4.19 with the new $\mu_c(L) = \mu_c(C, L)$ and see if the last admissible height of the tree is still $L = 4$. The intuition is that for higher C it is more difficult to reach the completion of the same height tree without a defected node creating a loop. That is because, for the same L, the corresponding $S_C(L)$ drastically increases if we increase C.

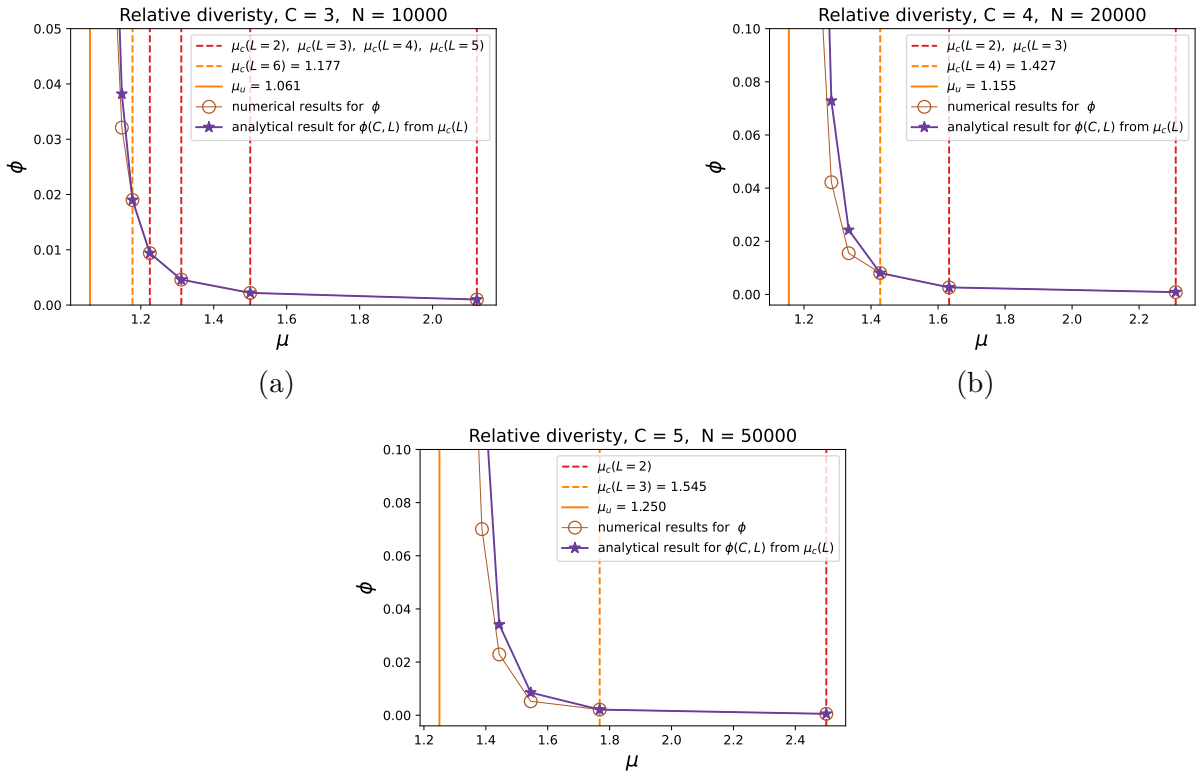


Figure 4.28: Relative diversity against social strength with numerical simulations and with the analytical values given by 4.13 for C = (3,4,5). As C increases also S(C) does and hence for a good comparison N must be increased as well. For C = 5 we used N = 500000.

As C gets larger the *long cycles regime* gets larger too. The motivation is that fixing μ, I am expecting a tree of height L, and a regular tree of height L with degree 5 has exponentially many more nodes than a regular tree of height L with degree 3. Hence higher probability

of forming a cycle. That is why in the three figures the maximal L such that the analytical relative diversity matches the simulations are $(6, 4, 3)$ for $C = (3, 4, 5)$. There is also a computational motivation behind this expansion of the *long cycles regime* as C is increased. Indeed, to obtain a locally tree like graph we have to impose a very large N, the higher the C the higher the N, so Figure 4.28 is very accurate for $C = 3$ but less accurate for $C = 5$. If we want to assess which is the real boundary of the *localization regime* without being excessively affected by the finite size curse, then we must study where the fractal exponent approaches zero, as shown in Figure 4.25.

4.4.2 Wealth Inequality

We can look into this regimes difference also from a wealth inequality perspective, using the Gini index. Looking at Figure 4.17 and the symmetric abundances of the trees we notice the break of symmetry given by the rise of long loops. The Lorenz curve shows this phenomenon too. However, to obtain the trees from the simulations, it is necessary to select a very large population size and in the case of states localization that means having a Lorenz curve that resembles a Kronecker delta $\delta_{x=1}$. Hence, these infinitesimal orders of magnitude make it difficult to distinguish a localized tree of $L = 4$ from one of $L = 2$ from the Gini coefficient. However, the three regimes can be easily distinguished.

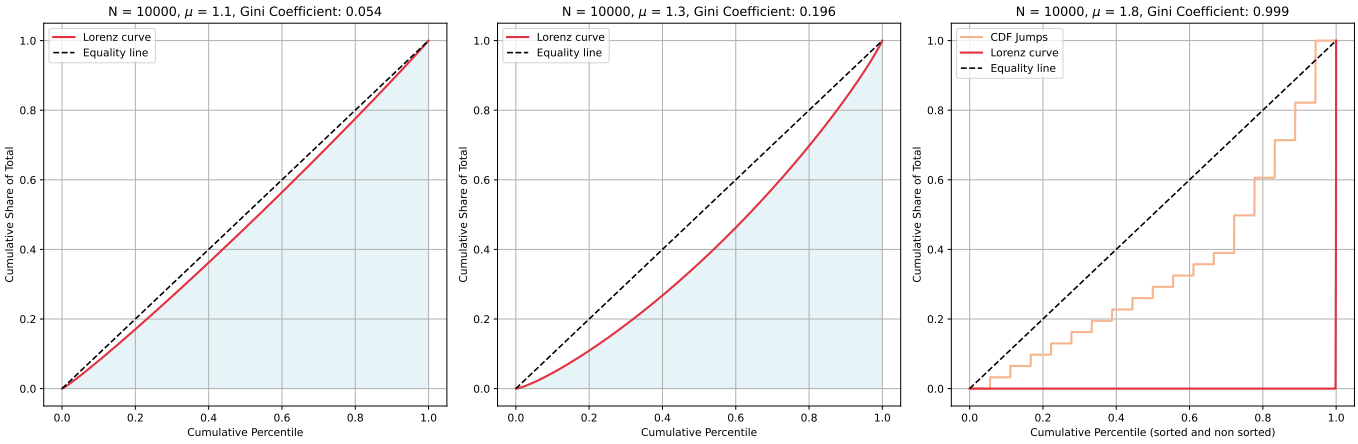


Figure 4.29: From left to right: *mean field regime, intermediate regime, localization regime*, with the Gini coefficient increasing with μ. For $\mu = 1.8$ in the localization regime, the jumps in CDF are included as for Figure 4.8.

In the *mean field regime* all species survive with the same abundance so the Lorenz curve overlaps with the equality line. Here that is not exactly the case because even if they all survive, the numerical noise produces some wealth inequality that generates a positive Gini

coefficient. The jumps of the CDF in the third image for μ in the *localization regime* agree with a height-2 surviving tree and that is in accordance with the theoretical values of Equation 4.17. Considering that before a certain μ_c the tree growth is broken, we cannot find an analytical value for Gini(μ). The reason is that we know the abundances distribution only for surviving trees but when the surviving component has long cycles we cannot say anything about the CDF used to compute Gini(μ). In the chapter where we analyze the heterogeneous degree with the Erdos-Renyi random graph, we show the numerical evolution of the Gini index against the social strength in the two cases.

4.5 Erdos-Renyi

So far we have dealt with *degree homogeneity*, which was convenient because having $\sum_j A_{ij} = C$ fixed guaranteed a non randomic structure of the final surviving component, which could be exactly predicted knowing the social strength.
A natural extension of the work would be to consider *degree heterogeneity* in a way that is relatable to the C-regular case. With this goal in mind, in this section we introduce an Erdos-Renyi (ER) graph $G_{N,p}$ with N nodes, where the connection between node i and node j is a *i.i.d.* random variable with the Bernoulli distribution $\mathcal{B}(p)$. If K indicates the random variable that gives the degree of each node, we can write

$$K \sim \mathcal{B}(N-1,p) \xrightarrow[\substack{N\to\infty \\ p\to 0}]{d} \mathcal{P}(\lambda),\ \lambda = (N-1)p$$
$$\mathbb{E}(K) = \lambda = C, \quad \mathrm{Var}(K) = \lambda = C \quad for \quad N \to \infty. \tag{4.21}$$

Where there are $N-1$ possible connections because a node cannot have a self-loop. $p = \frac{C}{N-1}$ is chosen such that $\mathbb{E}(K) = C$.

In the paper by *G. Bunin et al.* [30], it is shown that, after a transitory regime, the jumps happen for the same critical values α_c in the C-regular graph and in the $G_{N,p}$ graph with $p = \frac{C}{N-1}$; The relative diversity curves are different but those values are conserved. As we see from the asymptotic distribution in 4.21, the degree distribution does not converge to that of a C-regular, which is a delta centered in C. For what concerns the spectrum, in the regular case there is uniform interval with right edge of the eigenvalues distribution in C. For the ER with average degree C, the maximal eigenvalue is C as well but that point is an outlier from the semi-circle where almost all the eigenvalues are concentrated (see Appendix A).
From Figure 4.30, we observe a first difference between the competitive ER and the cooperative ER. In the former case, for a social strength larger than 0.5 there is a perfect matching

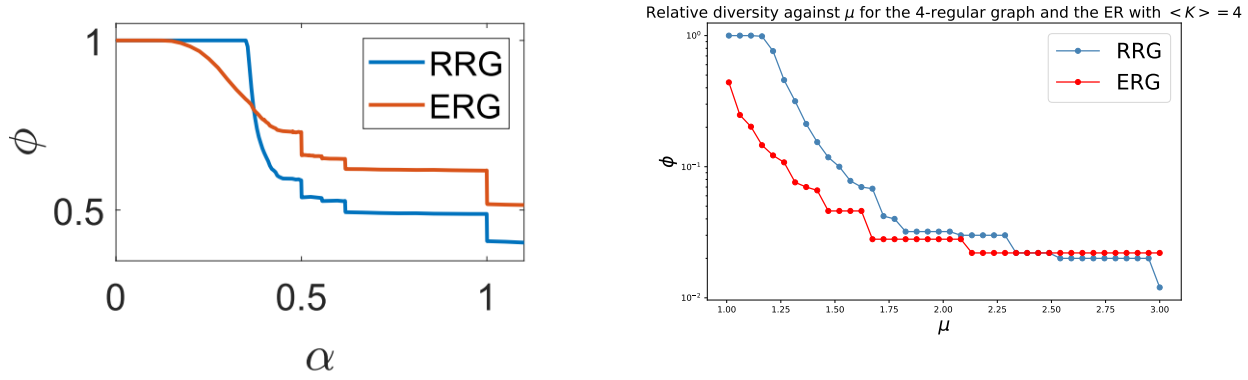


Figure 4.30: Left) From the paper [30]. ER with average degree 3 in red and 3-regular graph in blue; both taken from a system size $N = 400$. They present the same jumps for $\alpha > 0.5$. Right) Comparison of the relative diversity for the ER and the C-regular with C = 4 and N = 500 in our case of cooperation. The y-axis is log-scaled to highlight the difference at the last $\mu = 3$ where the ER has a larger number of species surviving.

between the points where the jumps occur. On the other hand, in our replicator model those jumps do not match. And it is not just a matter of averaging over many simulations; in the ER the last plateau always starts much earlier. In the C-regular case for $\mu < 1$, there is convergence, making the replicator model inapplicable. While in both the ER cases $\phi < 1$ for $\mu = 1$. That is in accordance with [1] where it is shown that $\mu_u = \frac{C^2}{\langle k^2 \rangle}$, that for $C = 4$ corresponds to $\mu_u = 0.8$. To verify this result we observe the relative diversity of 20 simulations and we notice that they all abandon the *mean field regime* in that specific value of μ.

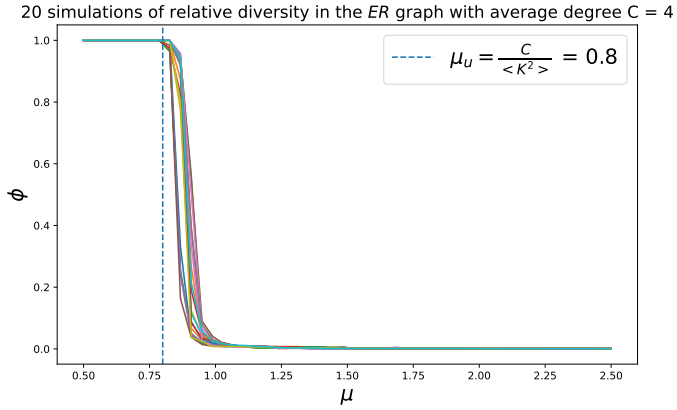


Figure 4.31: 20 simulations to plot the relative diversity and show that it abandons the *mean field* at $\mu_u = \frac{C}{\langle k^2 \rangle}$.

To show that there is no matching between the jumps in the C-regular and in the associated ER, we can look at the $\mu_{tree}(L)$ values of Equation 4.17. We see, from the simulation shown in Figure 4.32, that $\mu_{tree}(L = 2)$ and $\mu_{tree}(L = 3)$ are located where there is the last plateau (while in the regular case they corresponded to the first two jumps).

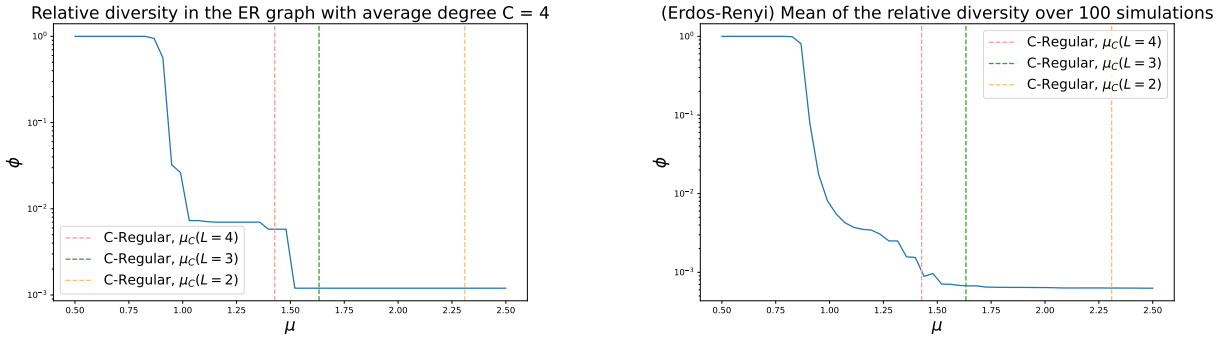


Figure 4.32: Relative diversity over different values of social strength in the Erdos-Renyi graph with average degree 4 and $N = 10000$. The results are compared to the jumps $\mu_C(L) = \mu_{tree}(L)$ in relative diversity of the associated regular graph to show that there is no matching. Left) Relative diversity for the single simulation. Right) Mean of the relative diversity over 100 simulations.

Now we want to understand what are the final structures in the right-most jumps in the case of heterogeneous degree and to do that we will analyze the final structures associated to the social strength values of the jumps in Figure 4.32.

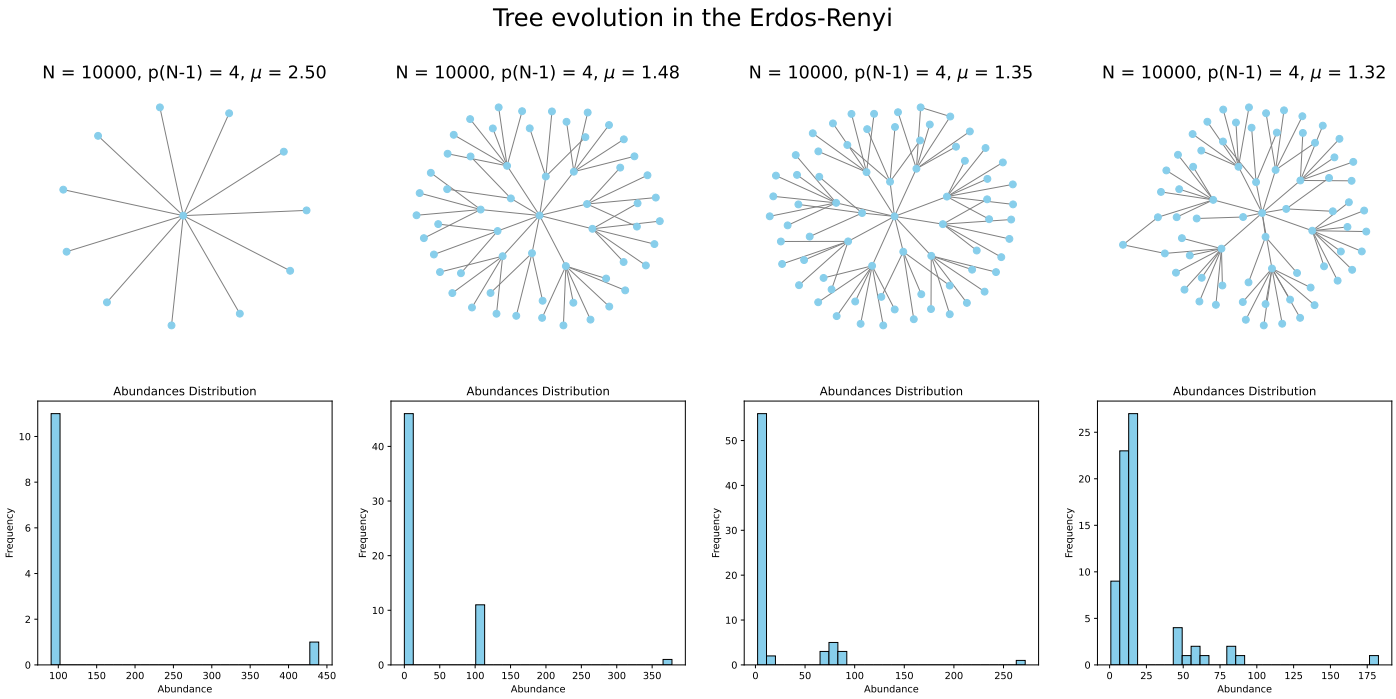


Figure 4.33: Final structures of the ER graph with C = 4 for the values μ where we have the last four jumps in the Figure 4.32. We have a tree evolution as μ decreases and from the abundances distribution we can see that we have symmetry around the central node.

We notice that here as well the evolution follows a tree-like growth where the tree now is not regular anymore. Indeed we notice, looking at the third height-2 tree, that in the first layer there are a node with degree 7 and a node with degree 3. This phenomenon allows us to find the the critical values where the transitions happen by taking the adjacency matrix of a branching process with an underlying Poisson distribution for the number of children. In particular the MPA, that we referred to in the introduction, can be used to implement a factorization of the characteristic equations and find the exact eigenvalues. Therefore, we would have the exact solution of the critical μ_L also in the case where the degree is heterogeneous. But that would be different every time because the branching process generates different outcomes every time.

Looking at Figure 4.33, the first question that comes up is whether the central node of the star (that has degree 11) is the one with highest degree in the whole graph. We supposed that the surviving component should have concentrated around that. But it is not; there is indeed a node with degree 13 which unexpectedly dies. In particular, we run 20 different simulations with 20 different seeds for the initial conditions on the same graph with max degree 13. We want to shuffle various IC combinations to see if this was an exception and the star that survives is usually the one with the highest degree central node. That would be explained by the bias affecting the surviving component influenced by the second level of randomness that we have in the heterogeneous degree case. But Figure 4.34 shows us that, over 20 simulations, only 4 times the surviving component expands around the node with maximum degree in the graph. The mode is 11 as in the example of Figure 4.33.

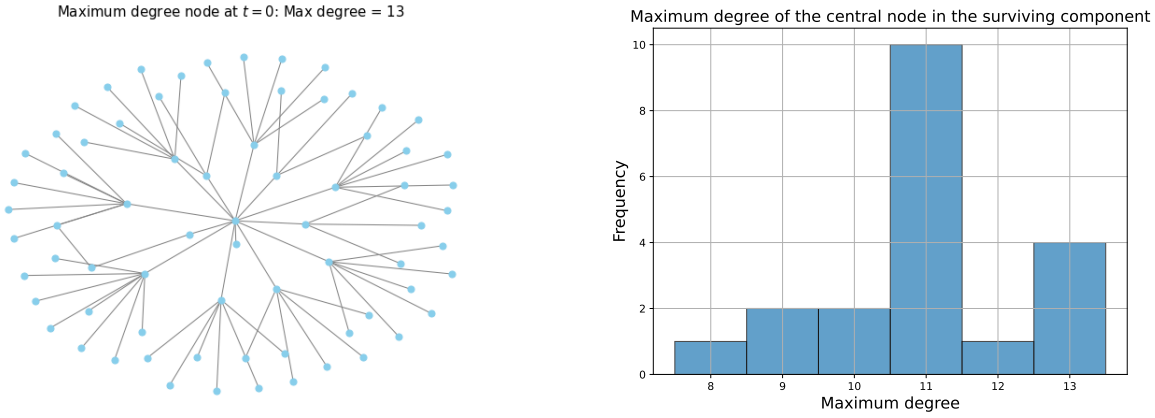


Figure 4.34: Left) Before running the dynamics the node with max degree has 13 connections and this is the component with first and second neighbours. Right) distributions of the final degrees of central nodes

Finally, we compare the homogeneous and the heterogeneous degree cases from the point of view of the Gini index to assess what is the influence of a degree regularity on the wealth inequality.

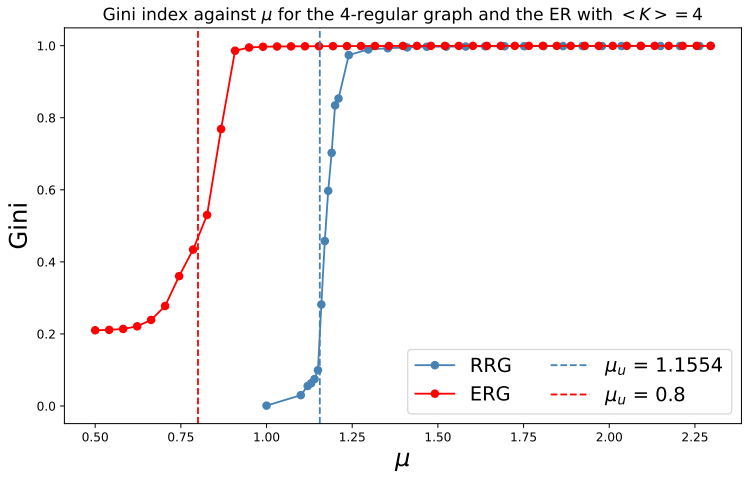


Figure 4.35: Gini index against the social strength in the 4-regular random graph and in the Erdos-Renyi with average degree 4. It was computed over a population of size $N = 10000$. In blue the observations for the regular case over a single simulation for the graph generation. In red the mean of the Gini index over 100 simulations for the graph generation. The dotted lines indicate when we abandon the regime where all species survive (the fist μ such that $\phi(\mu) < 1$).

From the last Figure 4.35 we can observe that in both cases we converge to a value very proximate to one because we arrive to the final localized star shaped structure. After the mean field critical value μ_u, we have a transition towards the extreme inequality, but it is steeper in the regular case because the *long cycles regime* is shorter. Finally, we can observe that in the regular graph case for $\mu < \mu_u$ the Gini coefficient is around zero because all the species have the same uniform abundance, while in the heterogeneous case, even if all species survive, they do that with different levels of wealth, producing an un-evenly distributed state.

5 Discussion

The main goal of this thesis is to investigate the equilibria of the cooperative Generalized Lotka-Volterra model on a random network. To cope with the divergence of the abundances of the species, we revisited the model using an alternative version of the replicator equations and that allows us to find inequality within a setting where species mutualistically help each other. In particular, effective competition arises and that has the consequence of an uneven wealth distribution where a few species get much richer than others. We study the two cases of homogeneous and heterogeneous degree distributions with the C-regular random graph and the Erdos-Renyi random graph with average degree C. In both of them we observe a localization of wealth that we traced to the phenomena of condensation and Anderson localization, typically found in statistical mechanics. To do that we have to use some notions of information theory, mainly the fractal exponent and the Shannon entropy, which we use to measure the inequality in the final states distributions. We observe that the stronger the communication between species, the stronger the effect of the effective competition. We quantify that from the relative diversity, a measure that indicates the fraction of surviving species over all the population size, and the abundances final distribution. It is extremely interesting to observe the geometrical evolution of the surviving components as the interaction strength decreases. In their minimal shape (high strength), the species are grouped into stars that gradually expand as trees, until they enter a regime where long cycles appear in the surviving component and the regularity is broken. From an economical point of view our study can be reflected in a scenario where different agents have all positive wealth or liquidity and they all benefit from each other's presence. However, within a positive wealth for everyone, hierarchies arise and a few agents get significantly richer than others, leading to a relative oligarchy. This induces an inequality that may have impacting consequences when for example there is a limited amount of available resources, which can be assigned only to those who are willing to pay more for them. This is not the scope of the research, but we could make a comparison with the national economies. In a setting of global cooperation, there is still a marked discrepancy between the developed and undeveloped ones, that gets higher with rising intra-economies communication. In this context, it would be interesting to extend this work by investigating analytically the consequences of freezing the dynamics of the Lotka-Volterra model to redistribute the wealth halfway through the process. Something

similar was done in [45] in the context of a good-exchange market with the introduction of helicopter money that were assigned at $0 < t < t_{final}$ to specific agents according to their rank. We limited our research on wealth inequality to the Gini coefficient, a metric widely used in literature that is based on the Lorenz curve. The further research that arises from this thesis spans over many field. From a random graph and random matrices point of view, we could extend this by studying the effects of different degree distributions or a preferential attachment model [51] where nodes are added also based on the current abundances. Finally, we could find a computational application to this study. In the introduction we mention that the competitive Generalized Lotka Volterra with high enough interaction strength is used as an algorithm to find the maximal independent sets in a graph [38]. Similarly, we can use our cooperative version of the model to find the long loops in the graph, without going through every possible path, with the only foresight of choosing the right interaction strength in the *long cycles regime*.

Bibliography

[1] Fabián Aguirre-López. "Heterogeneous mean-field analysis of the generalized Lotka-Volterra model on a network". In: *Journal of Physics A: Mathematical and Theoretical* (2024).

[2] Imane Akjouj et al. "Complex systems in ecology: a guided tour with large Lotka–Volterra models and random matrices". In: *Proceedings of the Royal Society A* 480.2285 (2024), p. 20230284.

[3] Stefano Allesina and Si Tang. "Stability criteria for complex ecosystems". In: *Nature* 483.7388 (2012), pp. 205–208.

[4] Philip W Anderson. *The economy as an evolving complex system*. CRC Press, 2018.

[5] Krishna B Athreya, Peter E Ney, and PE Ney. *Branching processes*. Courier Corporation, 2004.

[6] Anthony B Atkinson and Andrea Brandolini. "Promise and pitfalls in the use of "secondary" data-sets: Income inequality in OECD countries as a case study". In: *Journal of economic literature* 39.3 (2001), pp. 771–799.

[7] Andrea Bacilieri et al. "Firm-level production networks: what do we (really) know". In: *Unpublished results, University of Oxford* (2023).

[8] J Bauer, T-M Chang, and JL Skinner. "Correlation length and inverse-participation-ratio exponents and multifractal structure for Anderson localization". In: *Physical Review B* 42.13 (1990), p. 8121.

[9] Béla Bollobás. "A probabilistic proof of an asymptotic formula for the number of labelled regular graphs". In: *European Journal of Combinatorics* 1.4 (1980), pp. 311–316.

[10] Béla Bollobás. "Random graphs. 2001". In: *Cambridge Stud. Adv. Math* (2001).

[11] Jean-Philippe Bouchaud. "The (unfortunate) complexity of the economy". In: *Physics World* 22.04 (2009), p. 28.

[12] Jean-Philippe Bouchaud and Marc Mézard. "Wealth condensation in a simple model of economy". In: *Physica A: Statistical Mechanics and its Applications* 282.3-4 (2000), pp. 536–545.

[13] William E Boyce, Richard C DiPrima, and Douglas B Meade. *Elementary differential equations and boundary value problems*. John Wiley & Sons, 2021.

[14] William A Brock. "Pathways to randomness in the economy: emergent nonlinearity and chaos in economics and finance". In: *Estudios Economicos* (1993), pp. 3–55.

[15] Guy Bunin. "Ecological communities with Lotka-Volterra dynamics". In: *Physical Review E* 95.4 (2017), p. 042414.

[16] T-M Chang, JD Bauer, and JL Skinner. "Critical exponents for Anderson localization". In: *The Journal of chemical physics* 93.12 (1990), pp. 8973–8982.

[17] "Counting the number of spanning trees in a graph - A spectral approach". In: 1 (1966), pp. 34–54.

[18] C.M. da Fonseca and J. Petronilho. "Explicit inverses of some tridiagonal matrices". In: *Linear Algebra and its Applications* 325.1 (2001), pp. 7–21. ISSN: 0024-3795. DOI: https://doi.org/10.1016/S0024-3795(00)00289-5. URL: https://www.sciencedirect.com/science/article/pii/S0024379500002895.

[19] Jiu Ding and Aihui Zhou. "Eigenvalues of rank-one updated matrices with some applications". In: *Applied Mathematics Letters* 20.12 (2007), pp. 1223–1226. ISSN: 0893-9659. DOI: https://doi.org/10.1016/j.aml.2006.11.016. URL: https://www.sciencedirect.com/science/article/pii/S0893965907000614.

[20] David L Donoho, Arian Maleki, and Andrea Montanari. "Message-passing algorithms for compressed sensing". In: *Proceedings of the National Academy of Sciences* 106.45 (2009), pp. 18914–18919.

[21] Ioana Dumitriu and Soumik Pal. "Sparse regular random graphs: spectral density and eigenvectors". In: (2012).

[22] J Doyne Farmer. "Market force, ecology and evolution". In: *Industrial and Corporate Change* 11.5 (2002), pp. 895–953.

[23] Allan Griffin, David W Snoke, and Sandro Stringari. *Bose-einstein condensation*. Cambridge University Press, 1996.

[24] Joe Hasell. "Measuring inequality: what is the Gini coefficient?" In: *Our World in Data* (2023). https://ourworldindata.org/what-is-the-gini-coefficient.

[25] Josef Hofbauer. "On the occurrence of limit cycles in the Volterra-Lotka equation". In: *Nonlinear Analysis: Theory, Methods & Applications* 5.9 (1981), pp. 1003–1007.

[26] Samuel Karlin and James McGregor. "Spectral representation of branching processes. II: Case of continuous spectrum". In: *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 5.1 (1966), pp. 34–54.

[27] Max Sina Knicker et al. "Post-COVID inflation and the monetary policy dilemma: An agent-based scenario analysis". In: *Journal of Economic Interaction and Coordination* (2024), pp. 1–55.

[28] VE Kravtsov, BL Altshuler, and LB3762015 Ioffe. "Non-ergodic delocalized phase in Anderson model on Bethe lattice and regular graph". In: *Annals of Physics* 389 (2018), pp. 148–191.

[29] Thomas Lux. "Applications of statistical physics in finance and economics". In: *Handbook of research on complexity*. Edward Elgar Publishing, 2009.

[30] Stav Marcus, Ari M Turner, and Guy Bunin. "Local and collective transitions in sparsely-interacting ecological communities". In: *PLoS computational biology* 18.7 (2022), e1010274.

[31] Robert M May. *Stability and Complexity in Model Ecosystems*. Princeton: Princeton University Press, 2001. ISBN: 9780691206912. DOI: doi:10.1515/9780691206912. URL: https://doi.org/10.1515/9780691206912.

[32] Brendan D McKay. "The expected eigenvalue distribution of a large regular graph". In: *Linear Algebra and its applications* 40 (1981), pp. 203–216.

[33] Brendan D McKay, Nicholas C Wormald, and Beata Wysocka. "Short cycles in random regular graphs". In: *the electronic journal of combinatorics* (2004), R66–R66.

[34] Carl D Meyer. *Matrix analysis and applied linear algebra*. SIAM, 2023.

[35] Marc Mézard and Andrea Montanari. "Information, Physics, and Computation". In: (2009). URL: https://web.stanford.edu/~montanar/RESEARCH/book.html.

[36] Marc Mézard et al. "Replica symmetry breaking and the nature of the spin glass phase". In: *Journal de Physique* 45.5 (1984), pp. 843–854.

[37] Branko Milanovic. *Global inequality: A new approach for the age of globalization*. Harvard University Press, 2016.

[38] M. N. Mooij. "Generating Maximal Independent Sets Using Lotka-Volterra Dynamics". In: *Utrecht University Master Thesis*. Utrecht University Thesis Repository, 2022.

[39] José Moran and Jean-Philippe Bouchaud. "May's instability in large economies". In: *Physical Review E* 100.3 (2019), p. 032307.

[40] Karl Naumann-Woleske et al. "Exploration of the Parameter Space in Macroeconomic Agent-Based Models". In: *arXiv preprint arXiv:2111.08654* (2021).

[41] Giorgio Parisi et al. "Anderson transition on the Bethe lattice: an approach with real energies". In: *Journal of Physics A: Mathematical and Theoretical* 53.1 (2019), p. 014003.

[42] Jong Il Park et al. "Incorporating Heterogeneous Interactions for Ecological Biodiversity". In: *arXiv preprint arXiv:2403.15730* (2024).

[43] Nirbhay Patil, Fabián Aguirre-López, and Jean-Philippe Bouchaud. "The spectral boundary of block structured random matrices". In: *Journal of Physics: Complexity* 5.3 (2024), p. 035001.

[44] Nirbhay Patil and Jean-Philippe Bouchaud. "Emergent Inequalities in a Primitive Agent-Based Good-Exchange Model". In: *arXiv preprint arXiv:2405.18116* (2024).

[45] Nirbhay Patil and Jean-Philippe Bouchaud. "Emergent Inequalities in a Primitive Agent-Based Good-Exchange Model". In: *arXiv preprint arXiv:2405.18116* (2024).

[46] Francesca Pietracaprina, Valentina Ros, and Antonello Scardicchio. "Forward approximation as a mean-field approximation for the Anderson and many-body localization transitions". In: *Phys. Rev. B* 93 (5 Feb. 2016), p. 054201. DOI: 10.1103/PhysRevB.93.054201. URL: https://link.aps.org/doi/10.1103/PhysRevB.93.054201.

[47] Lyle Poley, Tobias Galla, and Joseph W Baron. "Interaction networks in persistent Lotka-Volterra communities". In: *arXiv preprint arXiv:2404.08600* (2024).

[48] Lothar Reichel and Lloyd N Trefethen. "Eigenvalues and pseudo-eigenvalues of Toeplitz matrices". In: *Linear algebra and its applications* 162 (1992), pp. 153–185.

[49] Valentina Ros et al. "Generalized Lotka-Volterra Equations with Random, Nonreciprocal Interactions: The Typical Number of Equilibria". In: *Physical Review Letters* 130.25 (June 2023). v2 with minor changes, p. 257401. DOI: 10.1103/PhysRevLett.130.257401. URL: https://hal.science/hal-04172871.

[50] Patrick Solé et al. "Spectra of regular graphs and hypergraphs and orthogonal polynomials". In: *European Journal of Combinatorics* 17.5 (1996), pp. 461–477.

[51] Remco Van Der Hofstad. *Random graphs and complex networks*. Vol. 54. Cambridge university press, 2024.

[52] Victor M Yakovenko. "Econophysics, statistical mechanics approach to". In: *arXiv preprint arXiv:0709.3662* (2007).

A Random Graphs Theory

In this appendix we provide some basic notions of random graphs theory that can be useful for a better comprehension of some concepts found in the results. Most of these definitions and theoretical results are taken from [51], [3], [38].

A.0.1 Definitions

Let us start with some fundamental definitions.

Graph: A graph $G = (V, E)$ consists of a set of vertices $V = \{v_i : i \in \{1, 2, \dots, n\}\}$ and a set of edges $E \subseteq V \times V$. With the size of the graph $|G|$, we denote the number of vertices $|V|$ in the graph.

Adjacency matrix: We define the adjacency matrix of a graph as

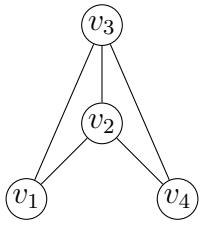
$$A_{ij} = \begin{cases} 1 & \text{if there is an edge between vertex } v_i \text{ and vertex } v_j, \\ 0 & \text{otherwise.} \end{cases}$$

Undirected: We call a graph undirected if $A_{ij} = A_{ji}$ for every $i, j \in \{1, 2, \dots, n\}$. If not, we call the graph *directed*.

Neighborhood: Let $G = (V, E)$ be a graph. For any vertex $v_i \in V$, we denote the neighborhood of v_i by $N(v_i) := \{v_j \in V \setminus \{v_i\} : A_{ij} = 1\}$.

Degree: The *degree* of a vertex $v_i \in V$ is the number of neighbors of v_i. We denote the degree by $d_i = |N(v_i)|$.

Vertex	**Degree**
v_1	2
v_2	3
v_3	3
v_4	2



$$A = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$$

Connectance: The connectance of a graph is defined as the ratio of the number of edges $|E|$ in the graph to the total number of possible edges between all pairs of vertices. For an undirected graph, the connectance is given by:

$$C = \frac{|E|}{\frac{|V|(|V|-1)}{2}},$$

where $|V|$ is the number of vertices, and $|E|$ is the number of edges in the graph.
Walk: A walk on a graph is a sequence of vertices, such that every pair of consecutive vertices is connected by an edge.

Path: A path in a graph is a walk where every vertex occurs only once.

Cycle: A walk in which the first vertex is the same as the last vertex and every other vertex occurs at most once, we call a cycle.

Connected: A graph is connected if there exists a path between any two distinct vertices in the graph.

Diameter: The *diameter* of a graph is defined as the longest shortest path between its vertices. Given the relevance of the definition in this study, it was explicited in 4.4.4 in the Results.

Typical Distance: It is the expected value of the distance between two randomly chosen vertices in the graph. It is computed as the average of the distances between all pairs of vertices. Let $X_N = d_G(u_N, v_N)$ represent the distance between two randomly chosen vertices u_N and v_N in a graph G with N vertices. The order of X_N provides insight into the graph's nature: if $X_N = O(\log \log N)$, the graph exhibits a small-small-world property, characterized by extremely short paths between vertices. In networks that follow a power-law degree distribution (also known as scale-free networks), the typical distance between nodes grows as $log - log(N)$. If $X_N = O(\log N)$, the graph exhibits the small-world property, where most pairs of vertices are connected by short paths. Conversely, if $X_N = O(N)$, the graph is considered to exhibit a large-world property, where distances scale linearly with the number of vertices, typical of sparse graphs with low connectivity.

A.0.2 Types of graphs

We call a graph simple if it is undirected and contains no self-loops or parallel edges. In this thesis, we only consider simple graphs. Now we give some examples of simple graphs that will recur.

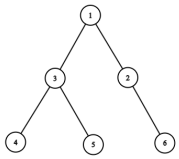


Figure A.1

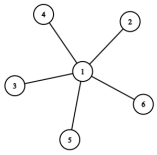


Figure A.2

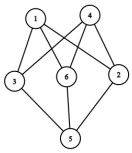


Figure A.3

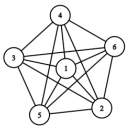


Figure A.4

Tree graph: We call a graph a tree if every pair of vertices is connected by exactly one path. A tree is a connected acyclic graph (A.1). Within the tree we have the following notions:

- The *root node* serves as the starting point from which all other nodes are explored.
- Each node in a tree, except the root, has exactly one parent, and the nodes directly connected to a given node are called its *children*.
- The *progeny* of a node v is the set of all its descendants, meaning all nodes that can be reached from v by following edges downward in the tree.
- The *offspring* of a node refers to the direct children of that node. If a node v has c children, then it is said to have c offspring.
- Nodes with no children are called *leaves*. The root node and the leaves form important parts of the structure of a tree, with the leaves representing the "endpoints" of the tree.
- The *height* of a tree is defined as the maximum distance from the root node to any leaf. In other words, if the tree is rooted at a node r, the height is the length of the longest path from r to any other node in the tree. The height gives a measure of the "depth" of the tree, indicating how many levels of branching occur.
- The *k-th degree neighborhood* of a node v in a tree is the set of all nodes that are exactly k edges away from v. This neighborhood encapsulates the nodes that can be reached from v by traversing exactly k edges, and it provides a way to study the local structure around any given node within the tree.

Star graph: We call a graph G a *star* if it is a tree with one vertex having degree $|V| - 1$ and the other vertices having degree 1 (A.2).

Regular graph: We call a graph *regular* if the degree of all vertices is the same (A.3).

Complete graph: We call a graph *complete* if there is an edge between any two distinct vertices in the graph(A.4).

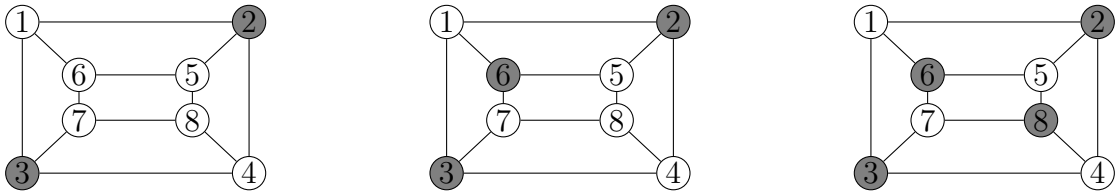


Figure A.5: Left) Maximal IS. Middle) Maximal IS. Right) Maximum IS.

Maximal Independent Set: An *independent set* (IS) in a graph is a set of vertices such that no two vertices in the set are adjacent. A *maximal independent set* (MIS) is an independent set that cannot be extended by including one more vertex from the graph, meaning adding any other vertex to the set would violate its independence. Within this definition, we also distinguish a *maximum independent set*, which is an independent set that contains the largest possible number of vertices for a given graph.

A.0.3 Random Graph Generation Models

Erdős–Rényi graph

Let $n \in \mathbb{N}$ and $p \in [0, 1]$. An Erdős–Rényi graph is constructed in the following way:

1. Initialize n vertices $\{v_i\}_{i=1}^{n}$.
2. Add each edge independently with probability p.
3. Return the obtained graph.

There exists a critical threshold for the connection probability p at which the graph undergoes a phase transition from being almost surely disconnected to being almost surely connected. This transition is characterized by the following critical values of p:

- If $p < \frac{(1-\epsilon)\log n}{n}$, for any $\epsilon > 0$, the graph is almost surely disconnected.

- If $p > \frac{(1+\epsilon)\log n}{n}$, for any $\epsilon > 0$, the graph is almost surely connected.

For connection probabilities p below this threshold, the graph is highly likely to remain disconnected, often consisting of small isolated components. However, as p increases and surpasses the threshold, the graph rapidly becomes connected, meaning that a giant component emerges and spans almost all vertices in the graph [38].

The Configuration Model

The Configuration Model is a random graph model that allows the creation of graphs with a specified degree sequence. Unlike the Erdős–Rényi model, which generates random graphs with a given edge probability, the Configuration Model constructs a graph where the degrees of the vertices are pre-specified.

Given a degree sequence $\{d_1, d_2, \ldots, d_n\}$ for a graph with n vertices, the Configuration Model is obtained by following these steps:

1. For each vertex v_i, create d_i "half-edges". The number of half-edges corresponds to the degree of each vertex.
2. Pair the half-edges uniformly at random to create edges between the vertices. Each pair of half-edges forms an edge.
3. If there are any leftover unpaired half-edges or self-loops (edges where both ends are connected to the same vertex) and multiple edges (edges connecting the same pair of vertices more than once), retry the pairing step.
4. The resulting graph has the degree sequence $\{d_1, d_2, \ldots, d_n\}$.

To build a C-regular random graph, where each vertex has degree C, one can apply the Configuration Model with a degree sequence $\mathbf{d} = \{C, C, \ldots, C\}$.

The Branching Process

A branching process is a stochastic model used to describe the growth of populations across generations. Let $X_{i,j}$ represent the number of children of the j-th individual in generation i. Each $X_{i,j}$ is an independent and identically distributed (i.i.d.) random variable following a Poisson (or Binomial) distribution with parameter λ, which governs the average number of offspring. The process starts with a single individual at generation 0, denoted by $Z_0 = 1$, and evolves over generations as follows: for each individual in generation i, the number of

children $X_{i,j} \sim \text{Poisson}(\lambda)$ is generated. The total number of individuals in generation $i+1$ is given by the sum of all offspring from generation i:

$$Z_{i+1} = \sum_{j=1}^{Z_i} X_{i,j}.$$

The expected population size at generation i is $\mathbb{E}[Z_i] = \lambda^i$, meaning that the population grows exponentially if $\lambda > 1$, remains constant if $\lambda = 1$, and faces extinction if $\lambda < 1$. Additionally, the variance of the population size grows linearly with the current population, as $\text{Var}(Z_{i+1}) = Z_i \cdot \lambda$. An important feature of the branching process is the probability of eventual extinction, denoted by q, which satisfies the equation

$$q = e^{-\lambda(1-q)}.$$

For $\lambda \leq 1$, extinction is certain ($q = 1$), whereas for $\lambda > 1$, there is a positive probability that the population survives indefinitely ($0 < q < 1$).
The total population size over all generations, denoted by T, is the sum of individuals in each generation:

$$T = Z_0 + Z_1 + Z_2 + \dots.$$

The total population size is finite if the process becomes extinct, but for $\lambda > 1$, there is a positive chance that the population grows without bound, depending on whether the process escapes extinction.

A.0.4 Random Graphs Spectrum

In graph theory, the spectrum of a graph is the set of eigenvalues of its corresponding adjacency matrix. The spectrum of G is often referred to as $\lambda_1, \lambda_2, \dots, \lambda_n$, where n is the number of vertices of the graph.
If the adjacency matrix A of a graph G is *Hermitian* (meaning that $\bar{A^T} = A$), then the spectrum is real. The adjacency matrix of a non directed graph is symmetric and therefore its eigenvalues are real numbers and can be ordered as $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$.
In the case of C-regular graphs it is a known result that the largest eigenvalue λ_1 is C and it corresponds to the eigenvector of ones. As we saw, random regular graphs can be generated using the configuration model with $\delta_{x=C}$ degree distribution. *B. McKay* studied the spectral density from the adjacency matrix in such random graphs in [33] and proved that the spectral density for a regular graph X with $C \geq 2$, where the size goes to infinity and the number of short loops normalized by such size goes to zero is:

$$f(x) = \begin{cases} \frac{C\sqrt{4(C-1)-x^2}}{2\pi(C^2-x^2)} & \text{for } |x| \leq 2\sqrt{C-1}, \\ 0 & \text{otherwise.} \end{cases}$$

This is the limiting density for $n \to \infty$, therefore even if the largest eigenvalue of a regular graph is C, the limiting portion of eigenvalues bigger than $2\sqrt{C-1}$ is zero. In particular, given the randomness in the choice of vertices and edges, the eigenvalues of most vertices are influenced by local fluctuations in the graph's structure and that is captured by *McKay*'s distribution. $[-2\sqrt{C-1}, 2\sqrt{C-1}]$ is then the *bulk of the eigenvalues* and as the size of the graph goes to infinity the fraction of eigenvalues outside that bulk goes to zero. A clarifying example about the difference of a regular random graph and a regular graph is the *ring*, a deterministic regular graph where eigenvalues are structured and deterministic.

In the case or Erdos-Renyi random graph we know the geometric and spectral properties of the adjacency matrix [10], [51]. In the low connectivity regime ($C = O(1)$), the ER adjacency matrix is a rank one deformation of a matrix with centred i.i.d. entries, so that we observe a circular law and one outlier.

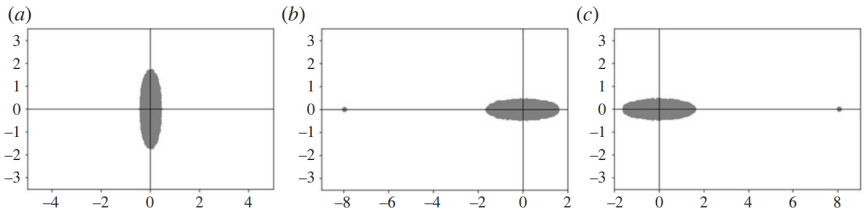


Figure A.6: Case of the non-Hermitian adjacency matrix from [2]. We have complex eigenvalues because graph is directed and the matrix is not symmetric. Left) Predator-Prey case with centered elliptic spectrum. Middle) Competitive case with elliptic spectrum and negative outlier. Right) Competitive case with elliptic spectrum and positive outlier equal to $C =< K >$.

The spectral density of Poisson branching processes has been studied extensively in [26], [5]. The process is governed by the generating function $f(s) = e^{-\lambda(1-s)}$. In the case of a process stopped at height L, which creates a tree of depth L, the corresponding transition matrix becomes finite-dimensional, even in the supercritical regime where $\lambda > 1$. This truncation results in a discrete spectrum. For large L, the eigenvalue distribution begins to approximate that of the infinite process, with smaller eigenvalues decaying exponentially, such that $\lambda_r \sim e^{-\lambda r}$. As discussed in [5], the spectral density $\rho(\lambda)$ for large L can be approximated by a continuous function that decreases exponentially with L: $\rho(\lambda) \sim \frac{1}{L} e^{-\lambda L}$.