GRAN SASSO SCIENCE INSTITUTE
SCHOOL OF ADVANCED STUDIES
Scuola Universitaria Superiore

Natural communities in sparse networks: a statistical mechanics approach

Submitted by

Tommaso Tonolo

Supervisors

Dr. Maria Chiara Angelini
Prof. Giacomo Gradenigo

Cycle XXXVII

Degree of Doctor of Philosophy in
Mathematics in Natural, Social and Life Sciences

Gran Sasso Science Institute
L'Aquila, Italy
2026

Abstract

Ecological communities are high-dimensional interacting systems whose stability emerges from the structure of their interaction networks. This thesis applies tools from the statistical mechanics of disordered systems to study the equilibrium properties of large ecosystems modeled by generalized Lotka–Volterra (gLV) equations with sparse and symmetric interactions. Moving beyond the traditional fully-connected (mean-field) setting, we analyze systems on locally tree-like random graphs using the Cavity Method and Belief Propagation (BP). In the single-equilibrium regime where BP converges, we compute species abundance marginals and show that increasing interaction disorder leads to a robust crossover from Gaussian to Gamma-like distributions. This behaviour is consistent with empirical observations and differs from the results obtained in fully-connected models. Remarkably, unlike dense systems, sparse ecosystems do not develop a glassy phase at large interaction disorder. At zero disorder, however, a peculiar phenomenon emerges, specific to sparse topologies: increasing competition among species induces a topological multiple-equilibria phase, signalled by the lack of convergence of the BP equations. To characterize this regime, we develop a one-step replica symmetry breaking (1RSB) cavity framework. The persistent non-convergence of both RS and 1RSB iterative schemes across the multiple-equilibria regime suggests a more intricate organization of equilibria, compatible with a full replica symmetry breaking (FRSB) phase with hierarchically structured basins. Altogether, these results demonstrate how sparsity reshapes stability, abundance distributions, and phase structure in large ecological communities, showing distinctive results compared to fully-connected models.

Acknowledgments

These PhD years have been full of people I would like to thank.

First, I would like to thank my supervisors, who guided me throughout these years. Grazie a Giacomo, che in questi anni mi ha riempito di consigli, con sincerità e affetto. Grazie a Maria Chiara, che mi ha insegnato la cura, la passione e il divertimento nella ricerca.

Thanks to all the *Angry Laundry People* for having shared life for one year at the Grand Hotel (e Del Parco). Thank you to the entire GSSI community, a unique example of a big family full of surprising people. Thanks in particular to all my companions of *Math37*, with whom I have shared these years from the very first moment. Thanks to all the associations and social spaces, *R4GE, CaseMatte, United L'Aquila, Radici, Polarville, Ju Boss* and many more, which contribute every day to making L'Aquila a better city.

Grazie ai miei coinquilini. A Fra, grazie per essere stato un compagno di tardi risvegli. A Fede, grazie per la pizza. A Peppe, senza il quale questi anni non sarebbero stati gli stessi. Ciao Rick anche tu sei un mio coinquilino adesso, ti voglio bene. Grazie a Davide, sei devastante. Grazie ad Anjali, per non aver mai smesso di darmi da mangiare.

Grazie a Tommy, Fede, Matte e Gio, ci conosciamo da un'eternità. È incredibile la semplicità con cui tutto resta come sempre.

Grazie agli amici di Mauro (MostiBondielliscolpitoremarmo), Ale, Checco, Ele, Eusebio, Fra, Gio, Gabri, Marghe, Nenni, all'entusiasmo e alla follia con cui ci ritroviamo.

Grazie a Margherita, alle marmitte, alle botti, alle albe al mare, ai trattori e ai prossimi mammut, bunker e forum. Sempre insieme.

Grazie ai miei genitori, che mi sono sempre vicini e mi supportano. Che sia dandomi un consiglio, chiamandomi per un saluto o mettendomi in imbarazzo, mi dimostrano sempre un sacco di affetto, con un tocco di pazzia.

Grazie a Francesca. È da più di 28 anni che mi lascio guidare da lei. Inizialmente senza scelta e poi, quasi subito, scegliendo di farlo. La seguo e la ammiro da sempre. Giocavamo a *noi eravamo* e ora giochiamo insieme a *noi saremo*, con convinzione ma anche con la spensieratezza che lei per prima mi ha insegnato.

Infine, grazie alle mie nonne Bruna e Graziella, alle quali dedico questa tesi, che mi hanno insegnato ad essere felice e spettinato.

Contents

List of Publications

Jacopo Niedda, Tommaso Tonolo, Giacomo Gradenigo (2024). **Probing marginal stability in the spherical p = 2 model**. Published in *Journal of Statistical Mechanics: Theory and Experiment*.

Tommaso Tonolo, Maria Chiara Angelini, Sandro Azaele, Amos Maritan, Giacomo Gradenigo (2025). **Generalized Lotka-Volterra model with sparse interactions: non-Gaussian effects and topological multiple-equilibria phase**. Published in *PRX Life*.

David Machado, Pietro Valigi, Tommaso Tonolo, Maria Chiara Angelini (2025). **Local equations for the generalized Lotka-Volterra model on sparse asymmetric graphs**, arXiv:2511.17499.

Introduction

A forest, a lake, a microbiome, a field of grass with flowers and bees. These are not merely words. They describe systems composed of many interacting species and their environment. The aim of this thesis is to characterize such systems, uncover connections between them, and provide insight into their structure and stability.

Understanding ecological communities means facing an extraordinary level of complexity. Species interact through competition, mutualism, predation and cooperation, forming networks that are heterogeneous, non-linear, and adaptive. Describing their coexistence and collective behaviour is an ambitious goal, one that calls for a synthesis of ideas from different fields. In this context, Statistical Mechanics can play an important role: it provides a natural framework to understand how macroscopic emergent properties arise from the interactions of many microscopic components.

Ecosystems are very delicate and their stability, defined by the existence of stationary states, their response to perturbations, and the occurrence of extinctions, is crucial, especially in the face of rapid human-induced change.

A central question concerns the role of biodiversity in determining ecosystems stability. Early empirical studies suggested that simplified ecosystems are more vulnerable, while more diverse ones are less sensitive to perturbations [1, 2]. However, this has never been a general rule. Also mathematical models have always offered contrasting perspectives. Robert MacArthur in 1955 [3] showed that a large number of species in an ecosystem results in avoiding dramatic changes in the population abundances, leading to a stable community. In contrast, Robert May's work in the 1970s [4, 5] demonstrated that ecosystems with many interacting species and a large variety of interactions can become unstable beyond a critical level of complexity. The relationship between diversity and stability thus remains a central and open problem in theoretical ecology.

Statistical Mechanics offers essential tools to disentangle the mechanisms that drive stability or instability in ecosystems. It helps identify structural and dynamical factors that determine how communities evolve and respond to perturbations. This is particularly relevant today, as ecosystems are increasingly affected by rapid, human-induced changes. While some systems are capable of recovering, such as the recent reappearance of the Galápagos rail on Floreana Island after the removal of invasive species, many others, like the Rapa Nui subtropical forest on Easter Island, are not. Understanding the principles governing stability is therefore not only a theoretical challenge but also a practical necessity.

The introduction of statistical-mechanical ideas marked a turning point in ecology,

considering ecosystems as complex systems whose collective behaviour emerges from disordered interaction networks, much like spins in glassy systems. The structure of these networks proves fundamental in shaping the resulting phenomenology of the system. Building on this perspective, the present thesis applies concepts and methods from the statistical mechanics of disordered systems to study the equilibrium structure and stability of large ecological communities described by sparse and symmetric interaction networks. The model that we employ is the generalized Lotka-Volterra (gLV) one, widely used in the literature, typically with dense interaction networks. Our approach based on sparse topologies is novel in this context and reveals distinctive phenomenology, including realistic non-Gaussian species distributions and transitions to regimes highly destabilized by species extinctions. These results differ from those obtained in fully-connected models and highlight the crucial role of network structure in shaping ecosystem behaviour.

The present thesis is organized as follows. In Chapter 1, we introduce the topic of Theoretical Ecology, starting from simple mathematical models and progressing to more complex ones such as the Consumer-Resource and generalized Lotka-Volterra (gLV) models. Within the framework of the gLV model, we also discuss the application of statistical physics methods to ecology and review the first results for ecosystems with symmetric and dense interaction networks.

Having established that Statistical Mechanics of Disordered Systems provides a natural approach to studying ecosystems, Chapter 2 presents the phenomenology of fully-connected disordered spin models in order to overview the techniques used in the field. We derive exact solutions for a prototypical model of glasses, focusing on the mechanisms which drive the emergence of multiple-equilibria in the low temperature phase. The chapter concludes by highlighting the properties and limitations of dense interaction networks and the need to extend the analysis to sparse systems, in the perspective of ecological applications.

The study of gLV ecosystems on sparse random graphs (Bethe lattices) with symmetric interactions is the central topic of Chapter 3. We first introduce the tools used to study the thermodynamics of such systems, namely, the Cavity Method and the Belief Propagation (BP) algorithm. We then present our results for the Generalized Lotka-Volterra model on sparse graphs, which differs significantly from that of fully-connected models. First, we find that for large values of the interaction disorder σ, the system remains in a single equilibrium where abundance distributions show a crossover, upon increasing the disorder σ, from a Gaussian to a Gamma-like distribution shape, the latter in good agreement with empirical data. This finding, together with the absence of a multiple-equilibria phase at large values of disorder, where it is only found a transition from single equilibrium to unbounded growth of species, is a novel and distinctive result compared to the fully-connected case. At the end of Chapter 3, we illustrate a mechanism that is specific to sparse topologies and has no analogue in dense networks. This mechanism drives the emergence of a multiple-equilibria phase in the gLV model on the Bethe lattice in the absence of disorder, triggered by an increase in competition among interacting species.

This new zero-disorder *topological* multiple-equilibria phase is further investigated in Chapter 4 using a more advanced algorithmic framework suited to regimes where

a single equilibrium no longer captures the system's thermodynamics. This multiple-equilibria phase is found to exhibit a rich, possibly fractal structure of coexisting equilibria.

Finally, in Chapter 5 we present our conclusions, highlight open questions and propose new research directions.

1

Statistical Mechanics and Ecology

The purpose of this thesis is to show how tools and ideas from the statistical mechanics of disordered systems can be successfully used to model real ecosystems and to investigate their emergent properties in a simplified yet insightful framework. In particular, among the open questions in theoretical ecology, the problem of defining and assessing the stability of large ecosystems is particularly well suited to a statistical mechanics approach. The goal of the present work is to show how the statistical mechanics modelling of ecological systems can unravel and frame non-trivial aspects of their phenomenology, therefore actively contributing to the advancement of theoretical ecology. This thesis thus offers a paradigmatic example of the interplay between theoretical ecology and theoretical physics, namely of the constant struggle to match empirical observations with mathematical models. The ultimate goal of applying statistical mechanics to the study of ecosystems is to achieve a deeper understanding of the underlying mechanisms governing their behaviour, ideally moving beyond the standard trial-and-error approach, which usually works only for few paradigmatic and simple situations, as for instance the famous case of Lake Victoria [6]. In the late 1950s, the Nile perch was introduced in the lake, with the aim of increasing the fish biomass. The result was the opposite: the Nile perch, a top predator, had a strong impact on the ecosystem, leading to the extinction of many species and to a decrease in the overall biomass. This is the typical case where the decrease of biodiversity leads to the increase of the instability of the ecosystem. But in many cases the mechanism behind instabilities is not that simple.

It was during the 1970s that a statistical mechanics perspective was introduced as a working tool to model a problem in Theoretical Ecology, particularly in addressing the problem of stability in large ecosystems. Let us therefore dedicate the next section to a brief historical account on the role played by mathematical models in the study of population growth and ecosystem stability.

1.1 Ecological Models

Attempts to model the growth of animal/human populations using mathematics date back several centuries. One of the earliest known examples is Fibonacci's 13th-century model describing the growth of a population of rabbits. The model is based on the assumption that in an idealized colony each pair of rabbits produces a new pair every month, starting from their second month of life. Let us assume now that $\tau \in \mathbb{N}$ is the positive integer labeling the months: it turns out that the total number of rabbit pairs present each month, which we denote as N_τ, can be determined recursively as

$$N_\tau = N_{\tau-1} + N_{\tau-2}, \tag{1.1}$$

and the sequence of rabbit pairs present at each month yields the well known Fibonacci sequence:

$$0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, ... \tag{1.2}$$

What matters for the purposes of the present discussion is that asymptotically the $\tau-$th element of the Fibonacci sequence increases exponentially with τ:

$$N_\tau \approx \varphi^\tau = e^{\tau\kappa}, \tag{1.3}$$

where $\kappa = \log(\varphi)$ and $\varphi = \frac{1+\sqrt{5}}{2}$ is the so-called *"golden ratio"*.

The result from Fibonacci on the population of rabbits is a particular case of the more general exponential law for the growth of populations later discovered by Malthus, which he proposed for the first time in 1798. Malthus not only realized that a population breeding at constant rate undergoes exponential growth, but also pointed out that this can happen only in the presence of unlimited resources – "the power of population is indefinitely greater than the power in the earth to produce subsistence for men". His work motivated the *Census Act* of 1800, which established the first census in England and Wales. The first mathematical model also accounting for the effects of limited resource availability was proposed by Verhulst in 1838 [7]: it was the first example of a *logistic model*. Such a model describes the evolution of a population $N(t)$ as:

$$\frac{dN}{dt} = rN\left(1 - \frac{N}{K}\right), \tag{1.4}$$

where the intrinsic growth rate r sets the exponential growth of the population, while the parameter K is the carrying capacity of the environment, which limits the growth of the population.

Another source of limitation in the unbounded exponential growth of a population is the interaction among different species, for instance on the form of a prey-predator interaction. The first mathematical model to account for this sort of interactions has been the Lotka-Volterra model in the 1920s [8, 9]. If we denote with $x(t)$ the density of preys population, e.g., rabbits, at time t and with $y(t)$ the density of predators, e.g.,

foxes, the Lotka-Volterra model consists in the following set of two coupled non-linear differential equations:

$$\begin{aligned} \frac{dx}{dt} &= \alpha x - \beta xy \\ \frac{dy}{dt} &= -\gamma y + \delta xy, \end{aligned} \tag{1.5}$$

where $\alpha, \beta, \gamma, \delta$ are all assumed to be positive real numbers. From Eq. (1.5) we see that, in absence of the mutual interaction terms, i.e., those proportional to the product of the two densities xy, the density of rabbits would experience unbounded exponential growth, due to the lack of predators, while the density of foxes would decay exponentially, due to the lack of preys, i.e., food resources. Instead, what has been discovered by Lotka and Volterra, is that by limiting the growth of rabbits with a negative contribution to the rate of growth of the kind $-\beta xy$ and by enhancing the rate of growth of foxes with a positive term δxy allows for stable equilibria where the density of both rabbits and foxes oscillates around finite values.

The generalization of the Lotka-Volterra equations to higher dimensions and so to a higher number of species gives rise to the so called *generalized Lotka-Volterra model* (gLV), which is a system of N differential equations (one for each species):

$$\frac{dn_i}{dt} = \frac{r_i}{K_i} n_i(t) \left[K_i - n_i(t) - \sum_{j \neq i} \alpha_{ij} n_j(t) \right], \tag{1.6}$$

where $i \in \{1, \ldots, N\}$ indexes the different species in the ecosystem. The gLV model is the most widely used in statistical physics to describe ecosystems: let us notice that it takes into account both the individual species intrinsic limitation of growth, as suggested by Verhlust, through the carrying capacities K_i in the terms $-r_i n_i^2/K_i$, and the mutual effect of species interactions $-\alpha_{ij} n_i n_j$, where the coefficients α_{ij} depend on the species.
When the number of species N is already larger than or equal to two, it is not possible anymore to provide an exact analytical solution to the differential equations (1.6). Here is where the average properties of the solutions start to be important and the coarse-grained approach of statistical mechanics becomes central.

When modelling the dynamics of large ecosystems, different approaches arise depending on the assumptions made about species diversity and interactions. A central question is whether species within an ecological community can be considered indistinguishable, or whether accounting for their differences in the model has important consequences. The gLV model in Eq. (1.6) for example explicitly assumes that species are not indistinguishable: each species is characterized by its own parameters r_i, K_i, and interaction coefficients α_{ij}. Species are thus fundamentally distinct, and their individual traits matter. However, this is not the only approach used in the study of complex ecosystems. Other approaches focus on the importance of stochasticity, assuming that random processes play a major role in shaping ecosystems, rather than niche differences. In the following, we introduce two broad classes of models, neutral models and niche models, which embody different answers to the question of species

equivalence and adopt different assumptions about the drivers of biodiversity and stability.

Neutral models

The central assumption of neutral models is that some key features of real ecosystems can be recovered even by assuming that all species play an identical role in the model. This is the core idea behind Stephen Hubbell's *Unified neutral theory of Biodiversity and Biogeography* (2001) [10], which proposes that species occupying the same trophic level can be treated as functionally interchangeable. The term "neutral" refers to the assumption that any differences among different species are neutral with respect to the success of individuals. Species are assumed to be indistinguishable and thus demographic stochasticity is particularly relevant in driving fluctuations in population sizes over time.

Hubbell's framework builds on earlier theories, such as MacArthur and Wilson's Theory of Island Biogeography [11], and has since been widely applied to biodiversity studies [12, 13]. These models are inherently minimalistic and are especially useful for describing species distributions in systems where species occupy the same position in the food chain, for instance trees in a forest. However, they are not well-suited to describe more complex ecosystems with significant trophic or functional differentiation.

A simple example of neutral dynamics is presented in [14], where the population size $x(t)$ of a species evolves according to the Langevin equation:

$$\dot{x}(t) = b - \frac{x(t)}{\tau} + \sqrt{Dx(t)}\,\xi(t). \tag{1.7}$$

Within a neutral framework, all species are assumed to be ecologically equivalent and therefore obey the same stochastic dynamics. The full community can thus be viewed as an ensemble of independent replicas of Eq. 1.7, differing only by their realizations of the demographic noise. Here, x is a continuous random variable representing species abundance, b is an immigration rate, τ is the characteristic relaxation time of the system, and D controls the amplitude of the demographic noise $\xi(t)$, modeled as Gaussian white noise with $\langle \xi(t) \rangle = 0$ and $\langle \xi(t)\xi(t') \rangle = 2\delta(t - t')$. The noise term captures ecological drift, reflecting the randomness in birth, death, and immigration events.

The Fokker-Planck equation associated with Eq. (1.7) governs the evolution of the probability distribution $p(x, t)$ of observing a species with abundance x at time t. Its stationary solution $p_0(x)$ reads:

$$p_0(x) = \frac{(D\tau)^{-b/D}}{\Gamma(b/D)} x^{\frac{b}{D}-1} e^{-\frac{x}{D\tau}}, \tag{1.8}$$

where $\Gamma(y)$ is the Gamma function. This stationary distribution has been shown to accurately fit empirical species abundance data, particularly those from forest ecosystems, like for example the one in Barro Colorado Island, Panama, where plenty of data were collected in the '90s (Center for Tropical Forest Science website, http://ctfs.si.edu).

Despite their empirical success in capturing some macroecological patterns, such as species abundance distributions, for species at the same trophic level, neutral models

are not microscopic models. They offer effective descriptions but do not account for the detailed structure of ecological interaction networks. In reality, species modify their environments, occupy specific niches, and interact in highly non-neutral ways [15, 16].

An alternative to the neutral approach is to explicitly model these microscopic interactions and investigate whether interesting, data-compatible patterns can emerge from them. This is the idea behind niche models.

Niche models

As discussed earlier, models that take into account differences between species are generally referred to as niche models. These frameworks rely on the idea that each species occupies a distinct ecological niche, consistent with the competitive exclusion principle, which states that two species competing for the same limiting resource cannot stably coexist. Over time, the one better adapted to the environment tends to outcompete the other, potentially leading to its extinction. Among niche-based approaches, the Lotka–Volterra equations, introduced in Eq. (1.5), play a central role. They describe how a species' growth rate depends not only on its own abundance but also on the presence and abundance of other species in the system. Species, in LV equations, are clearly non equivalent: preys and predators have different kind of dynamical equations, with different growth/death rates and interactions.

By moving away from the assumptions of neutrality, niche models allow us to investigate how specific interspecies interactions shape coexistence, competition, and the stability of ecological communities.

In this perspective, the two most important niche models which allow us to describe large, diverse and complex ecosystems are the MacArthur Consumer–Resource model and the generalized Lotka–Volterra model.

1.2 Competition for resources: MacArthur model

Let us begin this discussion of the competitive Consumer-Resource model by highlighting the role of competition in shaping ecosystems. Competition is a fundamental ecological force, deeply influencing the population dynamics of species. Even at the single-species level, competition for limited resources gives rise to density-dependent regulation, commonly modeled by the logistic growth term. When multiple species are present, competition becomes interspecific and can profoundly shape community structure. In the classical Lotka-Volterra framework, predators and prey are interdependent: an abundance of predators limits the growth of the prey population. The notion becomes more tangible when we consider species as organisms that exploit common resources, which are in limited supply, as the Consumer-Resource model describes. In consumer-resource models, a solitary species would face little difficulty in sustaining itself, but when others relying on the same resources are present, survival becomes uncertain for all.

This is the basis of what is known as the *competitive exclusion principle*: in a system with M available resources, the number of coexisting species N must satisfy $N \leq M$.

Although supported by consumer-resource models (and we will see it quantitatively at the end of this paragraph), this principle does not always match ecological observations. For reasons still not fully understood, natural ecosystems frequently display a higher level of coexistence than predicted. A well-known example is the *paradox of the plankton* [17]. Numerous plankton species live on a very limited number of shared resources like sunlight and dissolved minerals. Nevertheless, they manage to coexist in large numbers in relatively small regions of open ocean, in disagreement with the competitive exclusion principle.

To move from intuitive reasoning to a more quantitative description of resource-driven competitive dynamics, we now introduce the mathematical framework of the MacArthur Consumer–Resource model [18]. This model describes the dynamics of a community composed of consumers (species that feed on resources) and resources (the food they consume), and it consists of the following set of $N+M$ differential equations:

$$\begin{aligned} \frac{dn_\sigma}{dt} &= n_\sigma \left(\sum_{i=1}^{M} \alpha_{\sigma i} c_i - d_\sigma \right), \\ \frac{dc_i}{dt} &= c_i \left(f(c_i) - \sum_{\sigma=1}^{N} \alpha_{i\sigma} n_\sigma \right), \end{aligned} \tag{1.9}$$

where n_σ, with $\sigma = 1, ..., N$, are the abundances of the N consumer species, while c_i, with $i = 1, ..., M$, are the abundances of the M resources. The parameters $\alpha_{\sigma i}$ are the interaction coefficients which represent the preference of consumer σ towards resource i. The parameters d_σ are the death rates of the consumers. The function $f(c_i)$ represents the resource production rate. Usually $f(c_i) = \frac{r_i}{K_i}(K_i - c_i)$, with r_i the intrinsic growth rate of resource i and K_i its carrying capacity.

The MacArthur model is a niche model because it explicitly incorporates differences among consumers and resources through separate dynamical equations, distinct parameters d_σ, c_i and consumer–resource preferences $\alpha_{\sigma i}$.

While we will analyze the generalized Lotka–Volterra model in more detail later, we show here how the Consumer–Resource equations, under reasonable assumptions, naturally give rise to the gLV ones. Population dynamics of resources is assumed to be faster than that of consumers, so we can assume that resources are always at equilibrium. Assuming equilibrium for resources and considering for simplicity $f(c_i) = 1 - c_i$, from the second of Eqs. (1.9) we get that for any $i \in \{1, ..., M\}$:

$$c_i = 1 - \sum_{\sigma=1}^{N} \alpha_{i\sigma} n_\sigma. \tag{1.10}$$

Inserting Eq.(1.10) in the first of Eqs. (1.9), we get:

$$\frac{dn_\sigma}{dt} = n_\sigma \left(\sum_{i=1}^{M} \alpha_{\sigma i} c_i - d_\sigma \right) = n_\sigma \left(\sum_{i=1}^{M} \alpha_{\sigma i} \left(1 - \sum_{\sigma'=1}^{N} \alpha_{i\sigma'} n_{\sigma'} \right) - d_\sigma \right). \tag{1.11}$$

This corresponds to having:

$$\frac{dn_\sigma}{dt} = n_\sigma \left(a_\sigma - \sum_{\sigma'=1}^{N} A_{\sigma\sigma'} n_{\sigma'} - d_\sigma \right), \tag{1.12}$$

where $a_\sigma = \sum_{i=1}^{M} \alpha_{\sigma i}$ and $A_{\sigma\sigma'} = \sum_{i=1}^{M} \alpha_{\sigma i}\alpha_{i\sigma'}$. In particular we can define $r_\sigma = a_\sigma - d_\sigma$ as the intrinsic growth rate of species σ. Then, separating from the sum the term with $\sigma' = \sigma$, we get:

$$\frac{dn_\sigma}{dt} = n_\sigma \left(r_\sigma - A_{\sigma\sigma} n_\sigma - \sum_{\sigma' \neq \sigma} A_{\sigma\sigma'} n_{\sigma'} \right). \tag{1.13}$$

We thus recover an equation of the same form as Eq. (1.6): the generalized Lotka–Volterra dynamics emerge naturally from the Consumer–Resource model.

Let us now highlight how the MacArthur Consumer–Resource equations are aligned with the competitive exclusion principle. When resource dynamics are fast, the system admits a Lyapunov function, originally introduced by MacArthur [19], which guarantees convergence toward a stable equilibrium. After eliminating the resource variables through their equilibrium values, this Lyapunov function can be expressed solely in terms of the consumer abundances, and its curvature is determined by the consumer–resource preference matrix $\boldsymbol{\alpha}$.

A crucial property of this Lyapunov function is that it is strictly convex in the space of consumer abundances only if the number of resources is greater than or equal to the number of consumers, $M \geq N$. Strict convexity ensures the existence of a unique stable equilibrium in which all consumer species persist with strictly positive abundances. When $M < N$, the Lyapunov function loses strict convexity, and its minimization drives the system toward states where one or more consumer abundances vanish. This provides the formal dynamical basis of the competitive exclusion principle within the MacArthur Consumer–Resource framework.

1.3 Generalized Lotka-Volterra model

We have already introduced the generalized Lotka-Volterra (gLV) model, first in Eq. (1.6) as a generalization of the 2-species Lotka-Volterra equations and then in Eq. (1.13) as an effective description emerging from the Consumer-Resource model.

At its core, the gLV model is a niche model: species' niches are defined by their interactions, both within their own population and across different species. The deterministic version of the model combines logistic growth with pairwise interactions, as discussed in the previous sections. This description can be refined by including stochastic fluctuations, which account for random effects not captured by deterministic dynamics.

In this way, the general model consists of a set of N coupled stochastic differential equations describing the time evolution of the abundances of N species. The dynamics is given by:

$$\frac{dn_i(t)}{dt} = \frac{r_i}{K_i} n_i(t) \left[K_i - n_i(t) - \sum_{j \neq i} \alpha_{ij} n_j(t) \right] + \sqrt{n_i(t)}\eta_i(t) + \lambda, \tag{1.14}$$

where, with respect to Eqs. (1.6) and (1.13), we have also the stochastic noise $\sqrt{n_i(t)}\eta_i(t)$ and the immigration rate λ. Each species is labeled by an index $i = 1, \ldots, N$, and its

abundance at time t is denoted by $n_i(t)$. The *ecological* parameters r_i and K_i represent, respectively, the intrinsic growth rate and the carrying capacity of species i. Stochasticity is introduced through *demographic noise*, a multiplicative noise term proportional to $\sqrt{n_i(t)}$ that accounts for random birth and death events. The term $\eta_i(t)$ in the multiplicative demographic noise is a Gaussian white noise with zero mean and covariance $\langle \eta_i(t)\eta_j(t')\rangle \propto \delta_{ij}\delta(t-t')$. The demographic noise $\sqrt{n_i(t)}\eta_i(t)$ differs from another kind of multiplicative noise, the *environmental* noise, which scales with the abundance $n_i(t)$ and thus affects the per-capita demographic rates (e.g. r_i and K_i) independently of the value of $n_i(t)$. In this thesis, we will consider the case of demographic noise because, in the case of symmetric interactions, it satisfies detailed balance and allows the derivation of a closed-form Hamiltonian, as we show in Sec. 1.5. The last term in Eq. (1.14), λ, represents a small immigration rate that is often included to prevent complete extinction due to stochastic fluctuations.

Let us now focus on the deterministic part of Eq. (1.14), neglecting immigration. In the absence of interactions, the system admits two fixed points for each species: $n_i = 0$ (extinction) and $n_i = K_i$. For $r_i > 0$, the point $n_i = K_i$ is stable while $n_i = 0$ is unstable. Once interactions are included, however, the dynamics becomes more interesting and more complicated. The interaction coefficients α_{ij} encode the effect of species j on species i, which can be either beneficial or harmful depending on the sign of α_{ij}.

Following Robert May's influential 1973 work [5], we can classify the type of interaction between a pair of species (i, j) by looking at the signs of α_{ij} and α_{ji}. This classification is summarized in Table 1.1.

α_{ij}	α_{ji}	Interaction type
< 0	< 0	Mutualism
> 0	> 0	Competition
> 0	< 0	Predator-prey (prey i, predator j)
< 0	> 0	Predator-prey (predator i, prey j)
$= 0$	$\neq 0$	Commensalism ($\alpha_{ji} < 0$) / Amensalism ($\alpha_{ji} > 0$)
$\neq 0$	$= 0$	Commensalism ($\alpha_{ij} < 0$) / Amensalism ($\alpha_{ij} > 0$)
$= 0$	$= 0$	No interaction

Table 1.1: Type of interaction between two species i and j given the sign of the interaction coefficients α_{ij} and α_{ji}.

These interaction types (mutualism, competition, predation, commensalism and amensalism) are commonly recognized by ecologists. For instance, mutualistic relationships are often observed in pollination networks, competition shows up in resource-limited communities, and predator–prey interactions are, well, everywhere. That said, the real challenge comes when trying to measure how strong these interactions are.

Quantifying the strength of interactions, especially in large ecosystems, remains an open problem. Often they're not directly observable and are context-dependent. Inferring them from abundance data is far from straightforward. This is where the statistical mechanics perspective, on which the next section is focused, becomes fundamental.

1.4 Statistical Mechanics approach in ecological models

In the models introduced in sections 1.2 and 1.3, the ecological dynamics is governed by a number of parameters which grows with the number of species. Specifically, the generalized Lotka-Volterra model involves an interaction matrix α_{ij} of size $N \times N$ for a community of N species, while in the MacArthur model, the consumer-resource preferences are encoded in a matrix $\alpha_{\sigma i}$ with $N \times M$ entries for N consumers and M resources.

Given that ecological communities are typically very large and that it is practically impossible to measure all pairwise interactions in detail, we adopt a modelling approach in which the interaction coefficients are treated as independent and identically distributed (i.i.d.) random variables, which provides a controlled mean-field description of heterogeneous interactions in large communities. We stress that this assumption is a simplifying idealization rather than a claim of ecological realism, as correlations can lead to different behaviour in the large N limit. Its usefulness lies instead in the fact that, under appropriate scaling with system size, macroscopic properties become insensitive to the detailed form of the distribution and depend only on a small number of statistical parameters.

Let us recall the deterministic form of generalized Lotka-Volterra equations for species i:

$$\begin{aligned} \frac{dn_i}{dt} &= F_i(\mathbf{n}) \\ F_i(\mathbf{n}) &= \frac{r_i}{K_i} n_i(t) \left[K_i - n_i(t) - \sum_{j \neq i} \alpha_{ij} n_j(t) \right]. \end{aligned} \tag{1.15}$$

Imagine that the number of species N in the model is very large, $N \gg 1$, so that it is not possible to provide a detailed estimate of the coefficients α_{ij} weighting the interactions between species. In this regime, we model the α_{ij} as i.i.d. random variables. Let us now assume that the non-linear system of differential equations in Eq. (1.15) admits a fixed point $\mathbf{n}^* = \{n_1^*, \ldots, n_N^*\}$. It is then possible to linearize the gLV equations around these stationary points for small deviations $\delta n_i = n_i - n_i^*$ from the stationary solution:

$$\frac{d}{dt} \delta n_i(t) = [\mathbf{M}\, \delta \mathbf{n}(t)]_i, \tag{1.16}$$

where the matrix $\mathbf{M}$ has elements

$$M_{ij} = \left. \frac{\partial F_i}{\partial n_j} \right|_{\mathbf{n}=\mathbf{n}^*}. \tag{1.17}$$

In this context, the core idea proposed by Robert May in the work published in 1972, which can be considered as the milestone of the modern quantitative approach to theoretical ecology [4], has been to assume that $\mathbf{M}$ is itself a random matrix and to deduce the stability of the stationary point of the ecosystem dynamics from the properties of the random matrix eigenvalues spectrum.

May was the first one to relate the problem of ecosystems stability with the properties of a random matrix eigenvalues spectrum. If all eigenvalues have negative real parts, the

equilibrium is stable; if at least one has a positive real part, the equilibrium is unstable. The assumptions in May's work are that the elements of the interaction matrix are zero if the species do not interact, positive if they are beneficial, and negative if one species harms another. In particular, we can define as the *connectance* C the probability that two species interact. Another assumption of [4] is that each non-zero interaction element, except the diagonal ones which are all fixed to be -1 (each species is stable when isolated), is a random variable drawn from a Gaussian distribution with mean 0 and variance σ^2. From random matrix theory [20] it is known that, for large values of the number of species N the system is stable only for

$$\sigma < \frac{1}{\sqrt{CN}}. \tag{1.18}$$

This leads to the well known paradox of the May's bound: in the large-N limit the ecosystem is always unstable, in clear contrast with the empirical evidence that even very large and diverse ecosystems are stable.

The statistical approach introduced by May has strongly influenced Theoretical Ecology in the last 50 years. Consider for instance the case of the generalized Lotka-Volterra model, in which the interaction coefficients are modeled as normally distributed random variables, characterized by the following moments:

$$\begin{aligned} \langle \alpha_{ij} \rangle &= \mu/N, \\ \langle \alpha_{ij}^2 \rangle - \langle \alpha_{ij} \rangle^2 &= \sigma^2/N, \\ \langle \alpha_{ij}\alpha_{ji} \rangle - \langle \alpha_{ij} \rangle \langle \alpha_{ji} \rangle &= \gamma. \end{aligned} \tag{1.19}$$

The mean and variance parameters, μ and σ^2, are rescaled by the system size N so that the total interaction strength remains finite, i.e. $\sum_{i \neq j} \alpha_{ij} n_j \sim 1$ even in the limit $N \to \infty$. The parameter μ controls the average sign of the interactions coefficients: positive values of μ bias interactions towards competition, while negative values bias them towards mutualism on average. The parameter σ sets the width of the interaction distribution and thus controls the degree of heterogeneity in species interactions.

The symmetry parameter γ measures the correlation between α_{ij} and α_{ji}. The case $\gamma = 1$ corresponds to fully symmetric interactions, $\gamma = -1$ to fully antisymmetric ones, and $\gamma = 0$ to statistically independent pairs $\{\alpha_{ij}, \alpha_{ji}\}$. Intermediate values interpolate between these extremes, allowing partial symmetry or antisymmetry. Ecologically, values of γ close to -1 represent systems dominated by predator–prey dynamics, while values near 1 describe communities with mostly symmetric interactions, where predator–prey interactions are absent.

By varying μ, σ, and γ, one can explore a wide variety of ecological scenarios and study how different interaction structures give rise to different dynamics and emergent phenomena.

1.5 Ecosystem stability and multiple-equilibria phases

The result obtained by May in Eq. (1.18) is quite strong: sufficiently large ecosystems are always unstable. Due to the paradoxical nature of this conclusion one is forced to a

careful analysis of the underlying hypothesis and approximations. The debate around this problem is precisely the *diversity–stability* debate which has animated theoretical ecology over the past decades [21].

For instance, just a few years before May's work, Robert MacArthur, in [18, 19], showed that a very large number of competing species can, in fact, stably coexist. Starting from the Consumer–Resource equations given in Eq. (1.9), MacArthur introduced two key assumptions: first, that resources equilibrate much faster than consumers; second, that the effective interaction α_{ij} between consumer species i and j is symmetric, i.e., $\alpha_{ij} = \alpha_{ji}$. Under these conditions, he demonstrated the existence of a stable equilibrium for the ecosystem, which is robust in the $N, M \to \infty$ limit.

Such a stability can be interpreted as a form of *uninvadability*, in the following sense: once the ecosystem settles into its unique steady state, any attempt to reintroduce species or resources will result in the dying out of the reintroduced element. The equilibrium is therefore robust against small perturbations of the system.

Yet, this leads to a further question: is such an *uninvadable* unique equilibrium a generic feature of ecosystems? Or is it possible to have a less trivial equilibrium landscape, perhaps characterized by multiple equilibria?

Questions of this kind have long been a natural domain for statistical physics. For this reason, several ecological models were subsequently investigated using tools from statistical mechanics, with the aim of understanding the conditions under which intricate equilibrium landscapes can arise. A notable early example is the random replicator model, which describes the co-evolution of species interacting through random couplings. This model can be mapped onto a gLV one and was analyzed using replica theory in [22] and [23]. These works laid the first methodological foundations for the statistical-mechanical analysis of equilibrium properties in large ecological systems.

Building more directly on MacArthur's framework and its notion of an uninvadable equilibrium, the authors of [24] analytically solved the MacArthur's model in the limit of large N, using methods of statistical physics. They showed that an *uninvadable* phase, *shielded* from external conditions, exists, but that it can undergo a transition to a *vulnerable* phase in which external perturbations induce a reorganization of the community.

In this perspective, also the authors of [25] analyze a resource-competition model with tools from the statistical mechanics of disordered systems. Their results again reveal the presence of two distinct phases: alongside the uninvadable phase, a *vulnerable* phase emerges. These two regimes are characterized in [25] by a stability analysis: in the uninvadable phase, the system reaches a single equilibrium that remains completely unaffected by external perturbations, whereas in the vulnerable phase no such unique stable equilibrium exists. An interesting feature arises at the transition between the two phases, where the system is found to be marginally stable: a unique stable equilibrium still exists, but it becomes highly sensitive to even the smallest perturbations.

The analysis of [24] and [25] naturally leads to a further question: in the vulnerable phase, besides the absence of a stable fixed point, can we say more about the structure of equilibrium configurations in general ecological communities?

That is precisely the line of study pursued in works such as [26],[27],[28], and [29],

where statistical mechanics tools are employed to detect and characterize such vulnerable phase. In those works, following the ideas of [23], it has been found that the vulnerable phase of [25] corresponds to a *multiple-equilibria phase* which differs from the *single-equilibrium* regime which characterizes the uninvadable phase.

Following the approach of [26], let us rewrite for clarity the gLV equations governing the dynamics of a community of N species in the presence of demographic noise and immigration:

$$\frac{dn_i(t)}{dt} = n_i(t)\left[1 - n_i(t) - \sum_{j\neq i}\alpha_{ij}n_j(t)\right] + \sqrt{n_i(t)}\eta_i(t) + \lambda, \qquad i \in \{1, ..., N\}. \quad (1.20)$$

Here, for simplicity, all intrinsic growth rates r_i and carrying capacities K_i in Eq. (1.14) are set to 1. The Gaussian noise $\eta_i(t)$ has zero mean and covariance $\langle \eta_i(t)\eta_j(t')\rangle = 2T\delta_{ij}\delta(t-t')$, where the parameter T controls the amplitude of noise fluctuations. It is conventional to denote this noise amplitude by the temperature T, even though the system is not coupled to any thermostat and, in general, does not admit a thermodynamic equilibrium. Only in the special case where the interaction matrix α_{ij} is symmetric and the noise is demographic one can prove that the Fokker–Planck equation associated with Eq. (1.20) admits a stationary Boltzmann-like solution, with T effectively playing the role of an equilibrium temperature.

More explicitly, the Fokker–Planck equation governing the evolution of $P(\mathbf{n}, t)$, the probability distribution of species abundances $\mathbf{n} = \{n_1, \ldots, n_N\}$ at time t, admits as stationary solution, in the case of symmetric interactions and demographic noise, a distribution of the form:

$$P(\mathbf{n}) = \frac{e^{-H(\mathbf{n})/T}}{Z}, \quad (1.21)$$

where Z is the normalization constant and the effective Hamiltonian $H(\mathbf{n})$ is given by:

$$H(\mathbf{n}) = -\sum_{i=1}^{N}\left(n_i - \frac{n_i^2}{2}\right) + \sum_{i<j}\alpha_{ij}n_in_j + \sum_{i=1}^{N}\left[T\ln\, n_i - \ln\theta(n_i - \lambda)\right], \quad (1.22)$$

with $\theta(x)$ denoting the Heaviside function. This mapping to an equilibrium statistical mechanics problem allows us to apply tools from the theory of disordered systems to analyze the equilibrium phases of the gLV model and their dependence on the system's parameters. In particular, for dense interaction networks, the model's thermodynamics, which comprehends the analysis of the number and stability of equilibria, can be solved exactly.

The authors of [26] showed that, for fixed $\mu > 0$ and for sufficiently large T (or equivalently small σ), the system lies in a single-equilibrium phase, relaxing toward a unique stationary state. As σ increases (or T decreases), the system undergoes two transitions into multiple-equilibria regimes. The first transition leads to a phase in which the number of equilibria grows exponentially with the number of species N. These equilibria are separated by extensive free-energy barriers, so that in the

thermodynamic limit $N \to \infty$ they remain well-isolated and robust. In this regime, each equilibrium configuration is highly stable and the system is essentially trapped within one of them once reached.

By further increasing σ, a second transition takes place toward a different kind of multiple-equilibria phase. In this regime the number of equilibria is sub-exponential in N, and the basins of attraction are separated by non-extensive barriers. As a consequence, such equilibria are *marginally stable*: the system can move from one equilibrium to another under small perturbations.

From an ecological perspective, these results suggest that communities can exhibit very different equilibrium structures depending on the level of heterogeneity. On the one hand, ecosystems may settle into stable configurations that persist even in the presence of perturbations; on the other hand, they may display marginally stable equilibria where even small external perturbations can drive the community from one equilibrium to another, triggering a reorganization of the entire ecosystem.

In Chapter 2, we will show how these phases are described in replica theory language: the single-equilibrium phase corresponds to the *replica-symmetric* (RS) solution, the exponentially many robust equilibria to the *one-step replica symmetry breaking* (1RSB) phase, and the marginally stable equilibria to the *full replica symmetry breaking* (FRSB) phase.

Let us emphasize that all the results discussed so far rely on two key assumptions: interactions are symmetric, and the interaction network is fully connected. In real ecological systems, however, neither of these conditions is typically satisfied: interaction networks are sparse rather than complete, and interactions are often asymmetric. This raises a central question: how robust is the multiple-equilibria scenario once these assumptions are relaxed? In other words, do multiple stable stationary states persist when interactions are asymmetric and/or the network is sparse?

Another crucial issue concerns species abundance distributions which represent the most empirically accessible observables in ecology and, for this reason, are among the most widely studied in theoretical models. In fully connected models, at least when the interactions' distribution is Gaussian, species abundance distributions are always found to be Gaussian, in contrast with real-world data, which show highly non-Gaussian features [30]. Empirical distributions, in some cases well captured by neutral models, see Eq. (1.8), are typically log-normal or Gamma-like, with heavier tails than a Gaussian. Fully connected gLV models in general fail to reproduce these features.

Since dense networks are not realistic representations of ecosystems, and moreover lead to species abundance distributions which are inconsistent with empirical ecological data, it is essential to investigate what happens when we move to sparse interaction networks. This thesis is dedicated precisely to this goal: understanding the role of network connectivity by analyzing the behaviour of the model when interactions are sparse rather than dense.

For certain classes of sparse networks, exact methods do exist to analyze disordered systems. However, before introducing these techniques, it is necessary to first develop the theory of disordered systems on dense networks. Indeed, methods from the statistical mechanics of disordered systems are exact in the fully connected case and only

in very specific sparse cases. This motivates the structure of the thesis: in Chapter 2 we will analyze dense disordered systems, and in Chapter 3 we will then relax the assumption of denseness to study the sparse case. This approach mirrors the common methodology of statistical mechanics, where phase transitions are typically understood first in mean-field, fully connected models and then re-examined in finite dimensions or low-connectivity settings to test the robustness of those transitions.

The next chapter provides a detailed overview of disordered systems with dense interaction networks. We will introduce the replica method and show how multiple-equilibria (RSB) phases arise in such models. This will lay the groundwork for the central part of the thesis, where we investigate how these theoretical tools can be extended to sparse ecological networks, and which key features of the fully connected theory remain valid or fail in more realistic settings.

2

Mean Field Disordered Systems: dense interaction networks

This chapter is dedicated to introduce and discuss the exact solutions which, in the thermodynamic limit, can be computed for a continuous spin model with quenched random interactions. We will also present some new original results, but the main purpose of this chapter will be to introduce the jargon of replica theory and describe in some detail the properties of the most important multiple-equilibria phases which, at least in the case of dense interaction networks, represent exact solutions of the thermodynamics. In particular, we will point out the features that might carry over to the modelling of ecosystems.

2.1 Introduction

Statistical mechanics of disordered systems is a great framework for ecosystems modelling. 'Statistical Mechanics' enter the game because ecosystems are large communities of species interacting among themselves, and the main goal of Statistical Mechanics is exactly to study large collections of entities by focusing on the overall, emergent behaviour. This makes it a natural choice for studying ecosystems. After that, ecosystems are 'disordered systems' and it is worth going deeper into what this means, on how "disorder" enters the game. In particular, disorder is a form of randomness embedded in the structural elements of a system, a randomness that does not vary or varies slowly over time and differs from fast stochastic fluctuations.

In order to analyze this in the context of ecosystems modelling, let us recall for a moment the gLV Eq. (1.20) from Chapter 1, which governs the dynamics of a species i of the ecosystem:

$$\frac{dn_i(t)}{dt} = n_i(t)\left[1 - n_i(t)\right] - n_i(t)\sum_{j\neq i} \boldsymbol{\alpha_{ij}} n_j(t) + \sqrt{n_i(t)}\boldsymbol{\eta_i(t)} + \lambda. \tag{2.1}$$

Let us focus on the two sources of randomness in this problem, which are in bold in Eq. (2.1): we have the random interaction coefficients $\{\alpha_{ij}\}$ and the noise $\eta_i(t)$. Their main difference is that while $\{\alpha_{ij}\}$ are fixed once and for all, the noise $\eta_i(t)$ fluctuates along the dynamics. According to the jargon of statistical mechanics we say that the randomness of $\{\alpha_{ij}\}$ is "quenched" with respect to the dynamical fluctuations, and the random choice on the set of $\{\alpha_{ij}\}$ is the source of disorder in the system. In [26, 27] it is explained how the dynamics of Eq. (2.1) evolves to a stationary state where the dynamical fluctuation follows the equilibrium distribution $P_\alpha[\boldsymbol{n}]$. Let us remark that such evolution of the dynamics can be rapid or instead it can display long transients and aging, as also suggested by results on mean-field spin glasses [31, 32]. The latter is for example the case of the small demographic noise regime.

Anyway, the equilibrium distribution $P_\alpha[\boldsymbol{n}]$ which governs the dynamical fluctuations depends on the fixed set of coefficients $\{\alpha_{ij}\}$, which therefore represent a "quenched randomness" also with respect to thermal equilibrium fluctuations. In order to assess the properties of a given large ecosystem we need to know the typical properties of the equilibrium fluctuations in presence of quenched disorder.

Another important point that has to be highlighted is that, in order to tackle disordered systems analytically, we need to assume dense interactions, often implemented through fully-connected models, where each element interacts with all the others. It is a strong assumption, clearly not always realistic, but it allows for powerful analytical tools. In particular, in dense models, *mean-field* computations, where one considers the average collective behaviour and smooths out local fluctuations, become exact, greatly simplifying analytical treatment.

But in many real-world systems, including ecosystems, local effects matter, and interactions are sparse: species do not all interact with one another, but instead form a network with relatively few, structured connections [33–35].

In this chapter, we will focus on fully connected disordered models. This will help us build up the main ideas and techniques used in the field, and provide a reference point to see what changes when we drop full connectivity and move beyond mean-field calculations. In Chapter 3 we will then move on to studying ecological systems in the more realistic sparse case. This will allow us to identify both the similarities and the key differences that arise due to the network structure which, as we will see, plays a fundamental role in shaping the dynamics and stability of ecosystems.

2.2 Sherrington-Kirkpatrick model: fractal free-energy landscape

It is worth at this point to briefly introduce and comment the equilibrium thermodynamic properties of some disordered spin systems, with the purpose of gaining some familiarity on the methodology similar to the one used to study the Generalized Lotka-

Volterra systems in [26] and learn about the nature of the thermodynamic phases described therein by considering prototypical models. We begin with the Sherrington–Kirkpatrick (SK) model [36], which is a disordered Ising model on a complete graph (each spin interacting with any other spin):

$$H_J[\mathbf{S}] = -\sum_{i<j} J_{ij} S_i S_j, \tag{2.2}$$

where the couplings J_{ij} are i.i.d. Gaussian random variables with zero mean and variance $\sigma_J^2 = J^2/N$, the spins are dichotomic variables, $S_i = \pm 1$, and the subscript J in H_J denotes the dependence of the energy on a specific instance of the disorder. The exact solution of this model thermodynamics, revealing the existence of a transition between a high-temperature paramagnetic phase and a low-temperature spin-glass phase, was first provided by de Almeida and Thouless in [37], which also showed that their method was not able to provide a physically consistent description of the low temperature spin glass phase. The first analytical solution able to correctly describe this phase was the one provided by Parisi in 1979 [38].

We will discuss here few key steps in the derivation of the thermodynamic properties of the model, referring the reader interested in a complete calculation to [39]. In particular, we will see how it is only the introduction of a matrix order parameter that allows to correctly characterize the multiple-equilibria nature of the low temperature spin-glass phase, where spins are frozen in random configurations with zero magnetization. According to the standard statistical mechanics approach, the equilibrium thermodynamic phases of the Sherrington-Kirkpatrick model can be characterized by studying the free energy

$$F_J = -\frac{1}{\beta} \log \mathcal{Z}_J, \tag{2.3}$$

where β is the inverse temperature while $\mathcal{Z}_J$ is the canonical partition function

$$\mathcal{Z}_J = \sum_{\{S_1,\dots,S_N\}} e^{-\beta H_J[\mathbf{S}]}. \tag{2.4}$$

Since the free energy depends on the specific instance of disorder, it is itself a random variable. It is therefore natural to wonder which instance of the random couplings $\{J_{ij}\}_{i<j}$ is more representative of the system behaviour in the large-N limit. Luckily, it can be shown that, when the J_{ij}'s are Gaussian random variables, the free energy is *self-averaging*, namely the probability distribution of the free energy, $P(F) = \int \prod_{i<j} dJ_{ij} p(J_{ij}) \delta(F - F_J)$, is peaked around its average value $\overline{F_J}$ with a width decreasing with system size. More precisely, the self averaging property corresponds to have, in the large-N limit

$$\frac{\sqrt{\overline{F_J^2} - \overline{F_J}^2}}{\overline{F_J}} \sim \frac{1}{\sqrt{N}}. \tag{2.5}$$

This means two things: physically, any realization of the disorder yields a *typical* behaviour in the large-size limit; mathematically, to get the typical behaviour we can average the free energy over all random coupling instances.

$$\beta f \equiv \beta \overline{f_J} = \lim_{N\to\infty} -\frac{1}{N}\,\overline{\ln \mathcal{Z}_J}. \tag{2.6}$$

The standard procedure to average the logarithm of the partition function is to use the so-called replica trick, which amounts to the mathematical identity:

$$\overline{\ln \mathcal{Z}} = \lim_{n\to 0} \frac{\overline{\mathcal{Z}^n} - 1}{n}. \tag{2.7}$$

The subtlety consists in the fact that within the identity of Eq. (2.7) the parameter n is in principle a real variable, $n \in \mathbb{R}$, but all formal manipulations from here on are done by promoting n to an integer number, $n \in \mathbb{N}$, namely the *"number of replicas"*, and then taking the analytic continuation to real values $n \to 0$ (after taking the limit $N \to \infty$) only at the end of the calculation. These formal manipulations, described in all the original references which can be found in [40] and summarized in a synthetic but complete account in [39], allow to replace in the expression of the replicated partition function the trace over spin variables with the integration over the elements of a matricial order parameter, the *overlap* matrix q_{ab}, with the indices a and b running from 1 to n, the number of replicas, so to have:

$$\begin{aligned}
\overline{\mathcal{Z}_J^n} &= \int \prod_{i<j} dJ_{ij}\, p(J_{ij})\, \mathcal{Z}_J^n \\
&= \int \prod_{i<j} dJ_{ij}\, p(J_{ij}) \sum_{\{S_1^a,\ldots,S_N^a\}} \exp\left\{\beta \sum_{a=1}^{n} \sum_{i<j} J_{ij} S_i^a S_j^a\right\} \\
&= e^{\frac{nN\beta^2 J^2}{4}} \sum_{\{S_1^a,\ldots,S_N^a\}} \exp\left\{\frac{\beta^2 J^2}{2N} \sum_{a<b}^{n} \left[\left(\sum_{i=1}^{N} S_i^a S_i^b\right)^2 + O(N)\right]\right\} \\
&\simeq e^{\frac{nN\beta^2 J^2}{4}} \sum_{\{S_1^a,\ldots,S_N^a\}} \int \prod_{a<b} dq_{ab} \exp\left\{-\frac{N\beta^2 J^2}{2} \sum_{a<b} q_{ab}^2\right\} \exp\left\{\beta^2 J^2 \sum_{a<b} q_{ab} \sum_{i=1}^{N} S_i^a S_i^b\right\} \\
&= \int \prod_{a<b} dq_{ab}\; e^{-nNG[q_{ab}]},
\end{aligned} \tag{2.8}$$

where

$$G[q_{ab}] = -\frac{\beta^2 J^2}{4} + \frac{\beta^2 J^2}{2n} \sum_{a<b} q_{ab}^2 - \frac{1}{n} \ln[\mathrm{Tr}\, e^{L(q_{ab})}], \tag{2.9}$$

and where we used the compact notation

$$\mathrm{Tr}\, e^{L(q_{ab})} = \sum_{\{S^1,\ldots,S^n\}} e^{\beta^2 J^2 \sum_{a<b} q_{ab} S^a S^b}. \tag{2.10}$$

The way of writing $G[q_{ab}]$ in Eq. (2.9) emphasizes that it depends only on the order parameter q_{ab}, being the trace over the spins carried over. Let us summarize in words the main steps of the calculation in Eq. (2.8). From the first to second line of Eq. (2.8) we have substituted the expression for $\mathcal{Z}_J^n$. From the second to the third, one performs the integration over the Gaussian distributed random couplings and separates the $O(N)$ terms, subleading in N, from the $O(N^2)$ ones. From the third to fourth line of Eq. (2.8) we have dropped the $O(N)$ terms and we have introduced the order parameter q_{ab} by means of a Hubbard-Stratonovich transformation, while from the fourth to fifth line of the same equation we have just rearranged terms in order to make explicit that the expression depends only on the order parameter.

Let us emphasize that it is only because the Hamiltonian (2.2) involves a summation $\sum_{i<j}$ over all pairs of indices that, in the second line of Eq. (2.8), we were able to rewrite the double sum corresponding to a two-body interaction as the square of a single summation. And so at the end, only thanks to the all-to-all interaction between the spins, we obtained a factorization of the partition function with respect to the latin index i, which in particular allows us to write

$$\sum_{\{S_1^a,\dots,S_N^a\}} e^{\beta^2 J^2 \sum_{a<b} q_{ab} \sum_{i=1}^N S_i^a S_i^b} = \left[\sum_{\{S^a,\dots,S^a\}} e^{\beta^2 J^2 \sum_{a<b} q_{ab} S^a S^b} \right]^N . \tag{2.11}$$

The physical interpretation of the order parameter is obtained by making an intermediate step at the fourth line of Eq. (2.8): if one tries to solve the integration over q_{ab} by means of a saddle-point approximation in the large-N limit before taking the trace over the spins, she finds out that the saddle point values q_{ab}^* of the overlap matrix are

$$q_{ab}^* = \frac{1}{N} \sum_{i=1}^N S_i^a S_i^b . \tag{2.12}$$

The above expression tells us that, precisely as suggested by its name, in the large-N limit the matrix element q_{ab}^* represents indeed the scalar product, *overlap*, between the spin configurations in two different replicas. By then finally solving the integral in the last line of Eq. (2.8), again by means of a saddle-point approximation, and then expanding the result for small values of n, one is able to show that the whole computation of the free energy boils down to the solution of a variational problem with respect to a convenient parametrization of the matrix q_{ab}:

$$\overline{f_J} = \lim_{\substack{N\to\infty \\ n\to 0}} -\frac{1}{N\beta} \frac{\overline{Z^n} - 1}{n} = \frac{1}{\beta} \max_{q_{ab}} G[q_{ab}]. \tag{2.13}$$

There are two steps in the extremization of the function $G[q_{ab}]$ with respect to the overlap matrix elements q_{ab}: first, one has to make an ansatz on the matrix structure, which defines how many variational parameters one has, and then extremize the expression of $G[q_{ab}]$ analytically continued to $n = 0$ with respect to the value of these parameters. The fact that in Eq. (2.13) there is a maximization rather than a minimization is a peculiarity related to the analytic continuation at $n = 0$, whose mathematical

consistency has been proven with a probabilistic approach which does not make use of replicas in the famous Guerra-Talagrand theorem [41, 42]. The choice of this parametrization, together with the extremal values of the variational parameters, fully encodes the nature of the low-temperature spin glass phase.

The simplest possible choice for the non-diagonal matrix elements is to choose all of them identical, $q_{ab} = q$ for every couple of indices $a \neq b$. This is called the *replica symmetric ansatz*, and it was the first one considered in the first solution of the problem considered by de Almeida and Thouless. The first indication was of a spin-glass transition at T_c where in the low temperature phase the spin system does not magnetize but ergodicity is nevertheless broken, since at thermal equilibrium the typical value of overlap becomes positive, $q > 0$. But this new low temperature phase was also presenting non-physical features, as a negative entropy in the zero temperature limit, which is not possible for a system of discrete variables. It was then realized that the nature of the low temperature phase cannot be described by a single variational parameter [38], a more rich block diagonal structure must be assumed for the overlap matrix q_{ab}: the symmetry between replicas must be broken. To break the replica symmetry means, in jargon, to assume a block diagonal structure with nested blocks where many possible different values of the matrix elements are considered as variational parameters, see Fig. 2.1. For instance, a k-step breaking of the replica symmetry is an ansatz where one consider $q_0, \ldots, q_{k-1}$ possible values for each row of the matrix. The necessity of many possible equilibrium values of the overlap is what is needed to represent a free-energy landscape with many different minima, which is usually known as a *complex* free energy landscape, such that configurations of the system which remain trapped in different minima are characterized by different degrees of similarity, i.e., a different value of the overlap.

The key intuition was to understand that, below the spin glass critical temperature T_c, the system is always frozen in a given disordered configuration, but the degree of similarity between two different configurations cannot be captured by a single overlap parameter. In practice, the disordered configurations in the spin-glass phase can exhibit different degrees of similarity/overlap. In fact, the simplest non trivial extension of the RS ansatz is the one-step replica symmetry breaking (1RSB), in which one introduces a positive integer $m_1 (\leq n)$ and divides the n replicas into n/m_1 blocks, see Fig. 2.1. Elements q_{ab} in off-diagonal blocks are set to q_0, while those within the diagonal blocks are set to q_1. This procedure can be iterated, leading to two-step (2RSB), three-step (3RSB), and higher levels of symmetry breaking.

What Parisi showed in [38] is that the correct stable solution for the SK model involves an *infinite* number of such steps: in the so-called *full replica symmetry breaking* (full-RSB) solution the overlap matrix is parametrized by an infinite set $\{q_0, \ldots, q_\infty\}$ of order parameters. In fact, in the spin-glass phase the overlap between two equilibrium configurations at a given temperature can take infinitely many values according to a non-trivial distribution $P(q)$, with $q \in [0, 1]$, where it is actually the probability distribution $P(q)$ itself playing the role of a generalized order parameter [43].

This infinite number of order parameters, and the corresponding hierarchical structure of the q_{ab} matrix, captures the hierarchical nature of the free-energy landscape.

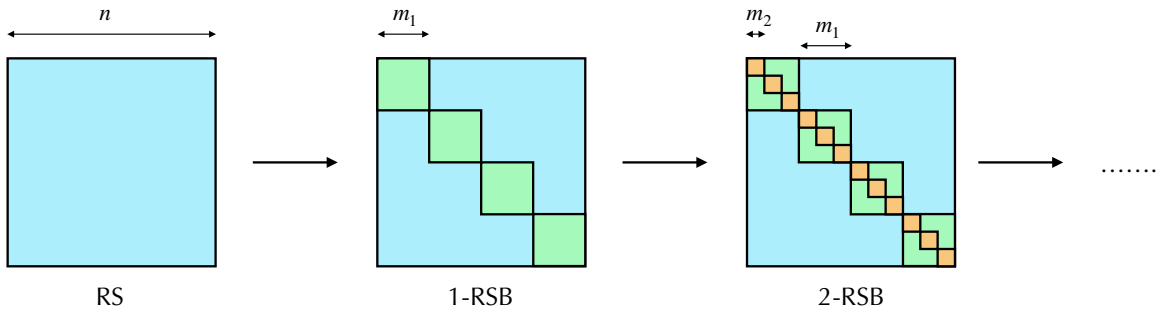

Figure 2.1: Block diagonal structure of the overlap matrix in the cases of RS, 1-RSB and 2-RSB ansatz. It is then clear how to further proceed with a generic k-RSB ansatz. In particular, each colour of the blocks corresponds to different values of the off-diagonal terms inside that block.

The low-temperature phase of the SK model is characterized by a landscape filled with nested minima (a multi-valley structure) which is typical of systems with quenched disorder, such as spin glasses.

This strategy of analysis was the one which revealed the existence of a multi-minima scenario in the equilibrium thermodynamics of the generalized Lotka-Volterra model where the coupling between species is represented by a symmetric random matrix [26]. As explained in Chapter 1, in this case the stationary dynamical fluctuations of the system are well represented by the Boltzmann weight $\exp[-\beta H_{\text{eff}}]$, where H_{eff} is an effective Hamiltonian which contains all the information on the ecosystem parameters and interactions and can be used to study the equilibrium properties of the gLV system:

$$H_{\text{eff}}(\mathbf{n}) = -\sum_{i=1}^{N}\left(n_i - \frac{n_i^2}{2}\right) + \sum_{i<j}\alpha_{ij}n_i n_j + \sum_{i=1}^{N}\left[T\ln\, n_i - \ln\theta(n_i-\lambda)\right]. \tag{2.14}$$

From the disordered partition function $\mathcal{Z} = \text{Tr}[e^{-\beta H_{\text{eff}}}]$ one then computes the typical free energy averaging over disorder and carries on an analysis analogous to the one of the Sherrington-Kirkpatrick model. At the core of this analysis is the study of the replica symmetric ansatz stability, which is carried on by computing the eigenvalues of the stability matrix

$$M_{ab,cd}[q^*_{ab}] = \left.\frac{\delta^2 G}{\delta q_{ab}\delta q_{cd}}\right|_{q^*_{ab}}, \tag{2.15}$$

which comes from an expansion in small deviations of matrix elements, $\delta\phi_{ab} = q_{ab} - q^*_{ab}$, from the chosen replica symmetry ansatz:

$$\overline{\mathcal{Z}_J^n} = e^{-nNG[q^*_{ab}]}\int\prod_{a<b}d\phi_{ab}\; e^{-\frac{1}{2}\sum_{ab}M_{ab,cd}\delta\phi_{ab}\delta\phi_{cd}} \tag{2.16}$$

At the heart of the analysis carried on in [26] is the calculation of the lowest eigenvalue of $M_{ab,cd}[q^*_{ab}]$, known in jargon as *replicon*: the transition from the high temperature

ergodic phase to the low temperature multi-minima phase can be typically recognized from the transition of the replicon, computed with respect to the replica-symmetric phase, from a positive to a negative value. In particular, in this case the expression of the RS replicon is

$$\lambda_{\mathrm{R}} = (\beta\sigma)^2 \left[1 - (\beta\sigma)^2 \overline{(\langle N_i^2 \rangle - \langle N_i \rangle^2)^2}\right], \tag{2.17}$$

where the angular brackets $\langle \cdot \rangle$ denote the thermal average with respect to the Boltzmann weight $\exp(-\beta H_{\mathrm{eff}})$, where H_{eff} is the effective Hamiltonian of the problem, Eq. (2.14), and where the overline denotes the average over the disorder. We have reported the explicit expression of the replicon, computed in the replica-symmetric phase, to show that it is made of physically transparent quantities: one has to compute the width of a single population fluctuations for a given instance of disordered couplings and then average it over disorder. When the extent of this fluctuations grows above a certain threshold limit corresponding to $1/(\beta\sigma)^2$ then a multi-minima phase arises. The authors of [26] in particular show that two continuous transitions are encountered by increasing the standard deviation σ of the disordered couplings distribution at fixed temperature, see Fig. 2.2, in both cases studying the behaviour of the replicon with respect to the appropriate phase.

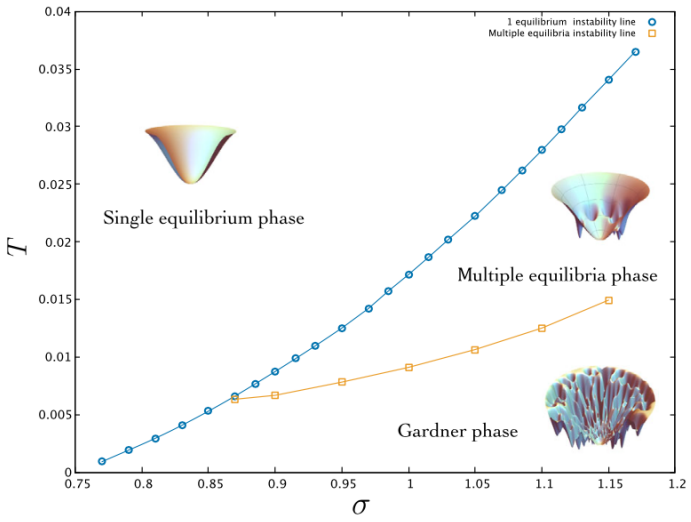

Figure 2.2: Phase diagram of the fully-connected gLV model studied in [26]. Two transitions to multiple-equilibria phases appear increasing the interaction disorder σ.

The first transition is from an ergodic phase to a 1-step replica symmetry breaking phase, namely a phase with many basins divided by extensive free-energy barriers and without any internal structure. Ecologically, this scenario would correspond to many very distinct and very stable equilibria: small perturbations are not able to trigger any relevant change in the ecosystem. But if one increases further σ there is another transition to a different sort of multivalley scenario: many more equilibria appears, which this time might have a very broad degree of similarity and where the transition to an equilibrium from slightly different ones can take place even upon the action of a

rather small perturbation. This second scenario clearly describes a much more versatile situation from the ecological point of view. Nevertheless, although these scenarios may appear suggestive, their existence could strongly depend on specific and ecologically arbitrary features of the model, such as the density of interactions. As we will discuss in Sec. 2.5, in some models having a sparse interaction network rather than a dense one could lead to the disappearance of such multiple-equilibria scenario [44].

2.3 The $p = 2$ spherical model: marginal stability

It is both interesting and instructive to examine how the nature of the low-temperature phase changes when an apparently small modification of the Sherrington–Kirkpatrick model, discussed in the previous section, is introduced. In particular, let us consider the so called *spherical model*, exactly solved for the first time by Michael Kosterlitz [45], which has the same interaction energy of the Sherrington-Kirkpatrick model, namely

$$\mathcal{H}_J(\boldsymbol{\sigma}) = -\sum_{i<j}^{1,N} J_{ij}\ \sigma_i\sigma_j. \tag{2.18}$$

In this case, differently from Sherrington-Kirkpatrick, we have *soft* spins, i.e., $\boldsymbol{\sigma} \in \mathbb{R}^N$ is an array of N, locally unbounded, real variables subjected to a global spherical constraint:

$$\boldsymbol{\sigma}\cdot\boldsymbol{\sigma} = \sum_{i=1}^{N}\sigma_i^2 = N. \tag{2.19}$$

The couplings J_{ij} are i.i.d. random variables taken from the Gaussian distribution

$$p(J_{ij}) = \frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{(J_{ij}-J_0)^2}{2\sigma_2^2}}, \tag{2.20}$$

where $\sigma_2^2 = J^2/N$ and J is a free parameter. The simplest version of the model is the one with $J_0 = 0$. The scaling of the variance ensures that the average internal energy is extensive in the large-N limit. For any given instance of the random coefficients J_{ij}, the probability distribution of the spin configurations is given by

$$P_{\beta,J}(\boldsymbol{\sigma}) = \frac{1}{\mathcal{Z}_J(\beta)} e^{-\beta\mathcal{H}_J(\boldsymbol{\sigma})}\delta\left(N - \boldsymbol{\sigma}\cdot\boldsymbol{\sigma}\right), \tag{2.21}$$

where the partition function $\mathcal{Z}_J(\beta)$ reads

$$\mathcal{Z}_J(\beta) = \int\left(\prod_{i=1}^{N} d\sigma_i\right)\ e^{-\beta\mathcal{H}_J(\boldsymbol{\sigma})}\delta\left(N - \boldsymbol{\sigma}\cdot\boldsymbol{\sigma}\right). \tag{2.22}$$

Quite remarkably, what has been found already in [45], is that the *softness* of the spins, which concur to reduce the frustration in the model, remove the presence of a

multi-minima complex landscape at small temperatures. At the same critical temperature T_c of the Sherrington–Kirkpatrick model, a transition occurs, where the average overlap between different configurations, which is $q = 0$ for $T > T_c$, becomes positive below the critical temperature. But this time the replica-symmetric ansatz is stable! This means that, if one computes the lowest eigenvalue of the stability matrix, the replicon, it vanishes at T_c but then remains zero all the way down to $T = 0$. This phenomenon is called marginal stability and the replica-symmetric phase with $q > 0$ is called *trivial spin-glass* phase. It turned out that the replacement of $S_i \in \{-1, +1\}$ variables with $\sigma_i \in \mathbb{R}$ variables completely changed the low temperature phase: the possibility to better accommodate the variables in the interactions completely removed the low-temperature ruggedness of the landscape. On the other hand the low temperature phase is still quite peculiar. The marginal stability expressed by the zero replicon encodes an interesting phenomenology: there is a vast basin of configurations which are not identical but have a finite degree of similarity, among which it is possible to move at zero energetic cost. This situation, as just mentioned, is called marginal stability, and is an interesting scenario also for ecological systems.

The exact solution of the model shows that different kind of perturbations have very different effects.

If a linear coupling with an external magnetic field is introduced, by considering a total Hamiltonian of the kind $\mathcal{H}_{\text{tot}}(\boldsymbol{\sigma}) = \mathcal{H}_h(\boldsymbol{\sigma}) + \mathcal{H}_J(\boldsymbol{\sigma})$, where $\mathcal{H}_J(\boldsymbol{\sigma})$ is the one of Eq. (2.18) and $\mathcal{H}_h(\boldsymbol{\sigma}) = h \sum_{i=1}^{N} \sigma_i$, one finds that the magnetic field completely washes the transition away. This means that drawing the phase diagram of the model in the plane (h, T) the only critical point is at $(h = 0, T = J)$ and for all values $h \neq 0$ there are no phase transitions. The behaviour is different if we take into account ferromagnetic couplings in addition to random couplings, i.e., if we allow the parameter J_0 to be different from zero in the distribution of Eq. (2.20). In this case, taking $J_0 = \epsilon/N$, one finds that the low-temperature trivial spin-glass phase of the spherical model is in fact stable with respect to an increase of ϵ, until a transition to a ferromagnetic phase occurs at large enough ϵ. This is shown in Fig. 2.3, which simply reproduces the phase diagram found by Kosterlitz, Thouless and Jones in Ref. [45].

In the following sections we will show how, by adding also different kind of non-linear interaction terms to the two-body interactions, it is then possible to modify the nature of the low temperature phase of the spherical model, triggering the formation of a multi-minima with a rugged free-energy landscape similar to the one of the Sherrington-Kirkpatrick model. Such analysis was done in Ref. [46]. This exercise is particularly instructive to understand the phenomenology of the generalized Lotka-Volterra model with symmetric random interactions studied in [26]. In this case we also have a two body interaction with a continuous variable, N_i, from which in principle we could expect, for the nature of the variable, a behaviour more similar to the spherical model rather than the Sherrington-Kirkpatrick model, therefore a single marginally stable trivial spin glass phase. Instead it has been shown that for the gLV with dense interactions there are transitions to a more structured low-temperature landscape described by non-trivial patterns of replica symmetry breakings. It is this sort of analogies between the behaviour of the gLV and prototypical models of spin glasses which makes the study of the latter a necessary preliminary step to grasp some intu-

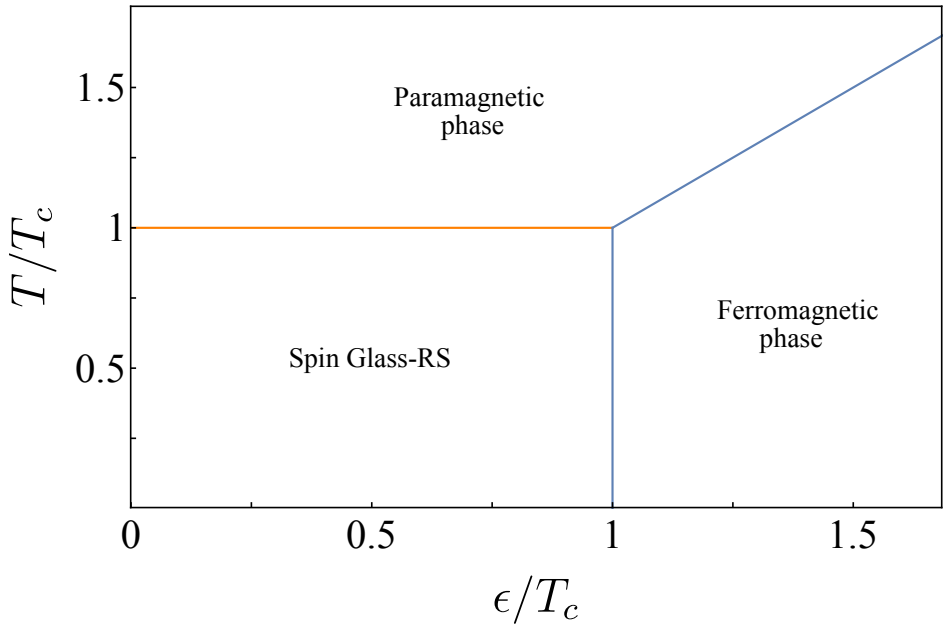

Figure 2.3: Phase diagram in the (ϵ, T) plane for the $p = 2$ spherical spin glass, where ϵ represents the strength of two-body ferromagnetic interactions.

ition on the role played by the different terms in the effective Hamiltonian describing the interactions among species. As already discussed in Chapter 1, the stability of stationary states is a crucial property for theoretical ecology: it helps us identify which aspects of an ecosystem make it fragile and what types of perturbations might lead to drastic changes in an ecological community.

2.3.1 Ordered non-linearity

As anticipated in the previous section the most interesting peculiarity of the $p = 2$ spherical model is to have a low-temperature trivial spin-glass phase, characterized at all temperatures below the critical one T_c by a zero smallest eigenvalue of the stability matrix $M_{ab,cd}$ [47]. It is therefore interesting to wonder if the addition to the Hamiltonian of some kind of terms could drive, from this marginally stable phase, an instability to a more structured low temperature phase.

To this purpose, we studied the effect of different kinds of ordered and disordered non-linear perturbations to the spherical $p = 2$ model.

The first perturbation to the spherical $p = 2$ model that we study is represented by an ordered 4-body interaction term. The model is defined by the Hamiltonian

$$\mathcal{H}_{\text{ord}}(\boldsymbol{\sigma}) = \mathcal{H}_J(\boldsymbol{\sigma}) - \frac{4!\epsilon}{N^3} \sum_{i<j<k<l}^{1,N} \sigma_i \sigma_j \sigma_k \sigma_l, \tag{2.23}$$

where $\mathcal{H}_J$ is the Hamiltonian (2.18) of the spherical $p = 2$ model with $J_0 = 0$. The scaling of the ordered coupling magnitude is chosen in order for the non-linear term to be extensive. The parameter ϵ can be tuned to probe different regimes according to the

strength of the non-linearity. The free energy of the model is obtained by means of a standard replica approach, which provides the exact mean-field solution in the large-N limit.

In this case, differently from the SK model analyzed in Sec. 2.2, we will have two order parameters. This is due to the presence of the ferromagnetic term (the ordered non-linearity) in the Hamiltonian (2.23). The two global order parameters that one has to consider are:

$$\begin{aligned} m_a &= \frac{1}{N}\sum_{i=1}^{N}\sigma_i^a \\ q_{ab} &= \frac{1}{N}\sum_{i=1}^{N}\sigma_i^a\sigma_i^b, \end{aligned} \tag{2.24}$$

where m_a and q_{ab} are respectively the magnetization elements and the overlap matrix elements. From here on, we will use the following symbols to denote the full overlap matrix and magnetization vector in replica space: $\mathbb{Q} = \{q_{ab}\}_{a,b=1,\dots,n}$ and $\mathbf{m} = \{m_a\}_{a=1,\dots,n}$. In terms of these order parameters, following similar passages to what we showed in Sec. 2.2, the replicated partition function averaged over the disorder reads as

$$\overline{\mathcal{Z}^n} = \int \prod_{a<b}^{1,n} \mathrm{d}q_{ab} \int \prod_{a=1}^{n} \mathrm{d}m_a \; e^{-NG[\mathbb{Q},\mathbf{m}]}, \tag{2.25}$$

where the action functional G is defined as

$$G[\mathbb{Q},\mathbf{m}] = -\frac{(\beta J)^2}{4}\sum_{ab}^{1,n} q_{ab}^2 - \beta\epsilon\sum_{a=1}^{n} m_a^4 - \frac{1}{2}\ln\det\Big(\mathbb{Q} - \mathbf{m}\otimes\mathbf{m}^T\Big), \tag{2.26}$$

with

$$\beta\overline{f_J} = \lim_{n\to 0}\frac{1}{n}\min_{m_a}\max_{q_{ab}} G[q_{ab}, m_a]. \tag{2.27}$$

In Eq. (2.26) the symbol $\mathbf{m}\otimes\mathbf{m}^T$ denotes a matrix with elements $(\mathbf{m}\otimes\mathbf{m}^T)_{ab} = m_a m_b$. Details on the derivation of $G[\mathbb{Q},\mathbf{m}]$ are given in App. A.1. As in standard replica approach, in order to find the physical value of the free energy we have to maximize with respect to the matricial order parameters and minimize with respect to the vectorial ones [40]. Stationary points are found by solving the following saddle-point equations:

$$\frac{\delta G}{\delta q_{ab}} = (\beta J)^2 q_{ab} + (\mathbb{Q} - \mathbf{m}\otimes\mathbf{m}^T)^{-1}_{ab} = 0, \tag{2.28}$$

$$\frac{\delta G}{\delta m_a} = -4\beta\epsilon\; m_a^3 + [(\mathbb{Q} - \mathbf{m}\otimes\mathbf{m}^T)^{-1}\cdot\mathbf{m}]_a = 0. \tag{2.29}$$

As pointed out in Sec. 2.2 we now have to choose an ansatz for the matrix q_{ab}. In particular, the RS ansatz turns out to be stable for all values of temperature T and strength ϵ of the non-linear couplings.

Replica Symmetric Ansatz

Let us then assume for the matrix $\mathbb{Q}$ a RS ansatz:

$$q_{ab} = \delta_{ab} + (1 - \delta_{ab})\ q_0. \tag{2.30}$$

For what concerns the vector $\mathbf{m}$ we assume that all elements are identical, since physical quantities cannot depend on the replicas. By doing so we get:

$$\left[(\mathbb{Q} - \mathbf{m} \otimes \mathbf{m}^T)^{-1}\right]_{ab} = \begin{cases} \dfrac{1 - 2q_0 + m^2}{(1 - q_0)^2} & a = b \\ -\dfrac{q_0 - m^2}{(1 - q_0)^2} & a \neq b. \end{cases} \tag{2.31}$$

As detailed in App. A.2, the RS effective action reads, in the limit $n \to 0$, as

$$\lim_{n\to 0} \frac{1}{n} G[q_0, m] = -\frac{(\beta J)^2}{4}(1 - q_0^2) - \beta\epsilon m^4 - \frac{1}{2}\ln(1 - q_0) - \frac{1}{2}\frac{q_0 - m^2}{1 - q_0}. \tag{2.32}$$

The RS saddle-point equations are obtained by plugging the ansatz (2.30) in Eqs. (2.28),(2.29) or by deriving Eq. (2.32) with respect to q_0 and m:

$$\frac{\partial G}{\partial q_0} = (\beta J)^2 q_0 - \frac{q_0 - m^2}{(1 - q_0)^2} = 0 \tag{2.33}$$

$$\frac{\partial G}{\partial m} = -4\beta\epsilon m^3 + \frac{m}{1 - q_0} = 0. \tag{2.34}$$

A trivial solution for Eq. (2.34) is $m = 0$: in this case Eq. (2.33) simply becomes identical to the RS saddle-point equation for the original $p = 2$ spherical model:

$$[1 - (\beta J)^2(1 - q_0)^2]q_0 = 0. \tag{2.35}$$

The above equation has two solutions, corresponding to the two possible phases of the model: the paramagnetic phase with $q_0 = 0$, and the trivial spin-glass phase with $q_0 = 1 - \frac{1}{\beta J}$. Eq. (2.34) admits also a $m \neq 0$ solution, which reads:

$$m^2 = \frac{1}{4\beta\epsilon}\frac{1}{1 - q_0}, \tag{2.36}$$

that plugged into the equation for q_0 gives

$$q_0(1 - q_0)^3 \left[\frac{1}{1 - q_0} - (\beta J)^2\right] = \frac{1}{4\beta\epsilon}. \tag{2.37}$$

Summarizing, the analysis of the saddle-point equations obtained by assuming a RS ansatz yields overall three kind of solutions: the paramagnetic one with both $m = 0$ and $q_0 = 0$; the solution corresponding to the trivial spin glass with $m = 0$ and

$q_0 = 1 - 1/(\beta J)$ and a ferromagnetic solution where both m and q_0 are different from zero and are given by the solutions of Eqs. (2.36) and (2.37). The first two solutions are the only possible ones with $\epsilon = 0$, that is the unperturbed case. The third solution is the interesting one in presence of non-linearity and can be studied numerically.

The following step of our analysis consist in studying whether the *ordered* quartic non-linearity may trigger or not an instability of the RS ansatz, in particular whether or not it drives, in the low temperature regime, a transition from a trivial spin-glass phase to a true spin-glass phase with a non-trivial distribution of the overlap. In order to do that we need to study the stability of the RS solutions.

Stability of the RS solution and phase diagram

The nature of the fluctuations around the saddle point determines the stability of the solutions. By following [37], it is convenient to define a single array containing all variational parameters:

$$\boldsymbol{\eta} = \begin{pmatrix} \mathbf{m} \\ [\mathbb{Q}]_{a \neq b} \end{pmatrix}. \tag{2.38}$$

The object defined in Eq. (2.38) is a vector in an $n(n+1)/2$-dimensional space, with element η_A where the subscript index takes values in the range $A = 1, \ldots, n(n+1)/2$. The number of variational parameters, with respect to which fluctuations must be taken into account, is $n(n+1)/2$ because we have to consider the n components of $\mathbf{m}$ and $n(n-1)/2$ components for $\mathbb{Q}$, since for the overlap matrix the diagonal elements are fixed to $q_{aa} = 1$ by the spherical constraint. The effective action (2.26) can be expanded around the saddle point $\boldsymbol{\eta}^* = (\boldsymbol{m}^*, \mathbb{Q}^*)$ up to the second order in the deviations from saddle-point solutions:

$$G[\boldsymbol{\eta}^* + \delta\boldsymbol{\eta}] = G[\boldsymbol{\eta}^*] + \frac{1}{2} \sum_{AB} \frac{\partial^2 G[\boldsymbol{\eta}^*]}{\partial \eta_A \partial \eta_B} \, \delta\eta_A \delta\eta_B + \ldots, \tag{2.39}$$

In Eq. (2.39) there are no linear terms because when $\boldsymbol{\eta}^*$ is the solution of the saddle-point equations and so we have by definition:

$$\frac{\partial G[\boldsymbol{\eta}^*]}{\partial \eta_A} = 0 \qquad \forall \; A. \tag{2.40}$$

In order to assess the stability of the solution $\boldsymbol{\eta}^*$ one has to study the spectrum of the Hessian $\mathbb{H}_{AB}(\boldsymbol{\eta}^*) = \partial^2 G[\boldsymbol{\eta}^*]/\partial\eta_A \partial\eta_B$. By explicitating the dependence on the overlap and magnetization fluctuations, the expansion in Eq. (2.39) can be rewritten as:

$$G[\mathbb{Q}, \mathbf{m}] = G[\mathbb{Q}^*, \mathbf{m}^*] + \frac{1}{2} \sum_{ab} \frac{\partial^2 G[\mathbb{Q}^*, \mathbf{m}^*]}{\partial m_a \partial m_b} \delta m_a \delta m_b + \sum_{ab,c} \frac{\partial^2 G[\mathbb{Q}^*, \mathbf{m}^*]}{q_{ab} \partial m_c} \delta q_{ab} \delta m_c + \\ + \frac{1}{2} \sum_{ab,cd} \frac{\partial^2 G[\mathbb{Q}^*, \mathbf{m}^*]}{\partial q_{ab} \partial q_{cd}} \delta q_{ab} \delta q_{cd} + \ldots \tag{2.41}$$

It is well known that, in the limit $n \to 0$, the smallest eigenvalue of the Hessian $\mathbb{H}[\boldsymbol{\eta}^*]$ is the replicon [37], which is related only to overlap fluctuations and can be obtained from the diagonalization of the submatrix

$$G_{(ab),(cd)} = \frac{\partial^2 G[\mathbb{Q}^*, \mathbf{m}^*]}{\partial q_{ab} \partial q_{cd}}. \tag{2.42}$$

In general, i.e. without specifying the solution ansatz for the saddle point, this matrix has three different kinds of elements, defined by taking: (*i*) $a = c$ and $b = d$; (*ii*) either $a \neq c$ and $b = d$ or $a = c$ and $b \neq d$; (*iii*) $a \neq c$ and $b \neq d$. These elements are related to different correlations of the replicated local variables. In the case of spherical variables, we can rewrite the action functional (2.26) in the following way

$$G[\mathbb{Q}, \mathbf{m}] = -\frac{(\beta J)^2}{2} \sum_{a<b}^{1,n} q_{ab}^2 - \beta\epsilon \sum_{a=1}^{n} m_a^4 + \ln \int \prod_{a=1}^{n} \mathrm{d}x_a e^{-\sum_{a\leq b}^{n} [(\mathbb{Q} - \mathbf{m}\otimes\mathbf{m}^T)^{-1}]_{ab} x_a x_b}, \tag{2.43}$$

where we have introduced back the local variables as correlated Gaussian auxiliary variables to represent the entropic term $\ln\det\left(\mathbb{Q} - \mathbf{m}\otimes\mathbf{m}^T\right)$. With this formalism, we have

$$\frac{\partial G[\mathbb{Q}, \mathbf{m}]}{\partial q_{ab}} = -(\beta J)^2 q_{ab} - \langle x_a x_b \rangle, \tag{2.44}$$

where the average is computed over the Gaussian distribution of the x_a variables:

$$\langle \cdots \rangle = \frac{\int \prod_{a=1}^{n} \mathrm{d}x_a e^{-\sum_{a\leq b}^{n} (\mathbb{Q} - \mathbf{m}\otimes\mathbf{m}^T)_{ab} x_a x_b} (\cdots)}{\int \prod_{a=1}^{n} \mathrm{d}x_a e^{-\sum_{a\leq b}^{n} (\mathbb{Q} - \mathbf{m}\otimes\mathbf{m}^T)_{ab} x_a x_b}}. \tag{2.45}$$

Therefore, by using the fact that

$$\frac{\partial}{\partial q_{cd}} \langle x_a x_b \rangle = -\langle x_a x_b x_c x_d \rangle + \langle x_a x_b \rangle \langle x_c x_d \rangle, \tag{2.46}$$

we have

$$G_{(ab),(ab)} \equiv \frac{\partial^2 G}{\partial q_{ab}^2} = -(\beta J)^2 + \langle x_a^2 x_b^2 \rangle - \langle x_a x_b \rangle^2 \tag{2.47a}$$

$$G_{(ab),(ac)} \equiv \frac{\partial^2 G}{\partial q_{ac} \partial q_{ab}} = \langle x_a^2 x_b x_c \rangle - \langle x_a x_b \rangle \langle x_a x_c \rangle \tag{2.47b}$$

$$G_{(ab),(cd)} \equiv \frac{\partial^2 G}{\partial q_{cd} \partial q_{ab}} = \langle x_a x_b x_c x_d \rangle - \langle x_a x_b \rangle \langle x_c x_d \rangle. \tag{2.47c}$$

We can now use Wick's theorem to compute the contractions of the 4-point Gaussian correlators and express them in terms of the second moment of the Gaussian distribution [47]. We have

$$G_{(ab),(ab)} = -(\beta J)^2 + \langle x_a^2 \rangle \langle x_b^2 \rangle + \langle x_a x_b \rangle^2 \tag{2.48a}$$

$$G_{(ab),(ac)} = \langle x_a^2 \rangle \langle x_b x_c \rangle + \langle x_a x_b \rangle \langle x_a x_c \rangle \tag{2.48b}$$

$$G_{(ab),(cd)} = \langle x_a x_c \rangle \langle x_b x_d \rangle + \langle x_a x_d \rangle \langle x_b x_c \rangle, \tag{2.48c}$$

where

$$\langle x_a x_b \rangle = \left[(\mathbb{Q} - \mathbf{m} \otimes \mathbf{m}^T)^{-1} \right]_{ab}. \tag{2.49}$$

Replicon

The diagonalization of $G_{(ab),(cd)}$ has to be performed case by case depending on the solution ansatz for the saddle point problem. If $(\mathbb{Q}^*, \mathbf{m}^*)$ is a RS saddle point, due to the symmetry of the overlap matrix $\mathbb{Q}^*$ under the permutation of n replicas, $G_{(ab),(cd)}$ depends only on three numbers:

$$P = -(\beta J)^2 + \langle x_a^2 \rangle^2 + \langle x_a x_b \rangle^2 \tag{2.50a}$$
$$Q = \langle x_a^2 \rangle \langle x_a x_b \rangle + \langle x_a x_b \rangle^2 \tag{2.50b}$$
$$R = 2\langle x_a x_b \rangle^2. \tag{2.50c}$$

Therefore, the eigenvalues of $G_{(ab),(cd)}$ depend on combinations of P, Q and R and can be computed in a straightforward manner by following the procedure of Ref. [37] or by using the Replica Fourier Transform [47, 48]. There are three types of eigenvalues at finite n: the *longitudinal* one, corresponding to fluctuations of the overlap which do not break the replica symmetry, the *anomalous* one, corresponding to fluctuations which break the replica symmetry, but violate the property of the overlap matrix of having the same sum over each row, and the *replicon*, connected to fluctuations which break the replica symmetry preserving the properties of the overlap matrix. In the limit $n \to 0$, the longitudinal and anomalous eigenvalues degenerate in the same one. The replicon has the following expression:

$$\begin{aligned} \lambda_R &= P - 2Q + R \\ &= -(\beta J)^2 + \left(\langle x_a^2 \rangle - \langle x_a x_b \rangle \right)^2 . \end{aligned} \tag{2.51}$$

The expression in Eq. (2.31) for $(\mathbb{Q} - \mathbf{m} \otimes \mathbf{m}^T)^{-1}$ in the RS case implies that the explicit expression of the replicon is

$$\lambda_R = -(\beta J)^2 + \frac{1}{(1 - q_0)^2}. \tag{2.52}$$

Let us specify that for RSB saddle points, the diagonalization of the Hessian matrix is more difficult, since the three different kinds of elements in Eq. (2.47) have more than just three possible values. The computation for the fluctuations around a 1RSB saddle point can be found in Ref. [49], while the generalization to generic k-RSB saddle points can be found in Ref. [50]. But for the present case of an ordered non-linearity this further step of the calculation is not necessary, since the RS ansatz turns out to be always stable.

Phase diagram

The phase diagram of the model is then obtained by looking for the solutions of the RS saddle-point equations (2.33)(2.34) and studying their stability. The result of this analysis is presented in Fig. 2.4. For $T > T_c$ and small values of ϵ the system is in a paramagnetic phase, which is stable because the replicon is always positive, as can be checked by plugging into Eq. (2.52) the values $q_0 = 0$ and $T > T_c = J$.

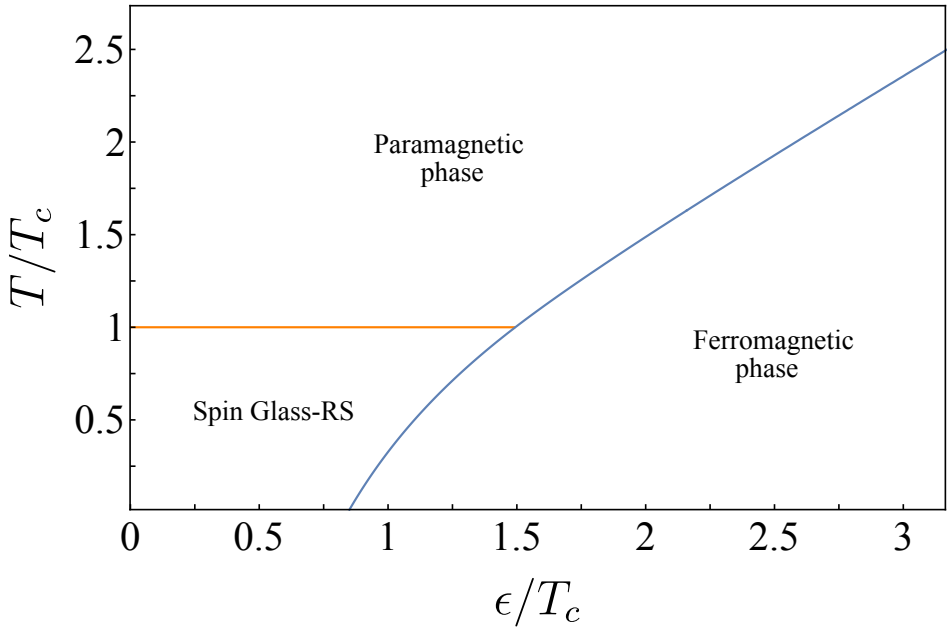

Figure 2.4: Phase diagram in the (ϵ, T) plane for the $p = 2$ spherical spin glass plus four-body ferromagnetic interactions, where ϵ represents the strength of 4-body ferromagnetic interactions.

Then, for small values of ϵ by lowering the temperature below T_c the system passes to a trivial spin-glass phase, which is marginally stable, since $\lambda_R = 0$ in all this region of the parameters as can be checked by using the expression $q_0 = 1 - \frac{1}{\beta J}$ in Eq. (2.52). On the contrary, if ϵ is large enough and the non-linear ferromagnetic interaction prevails on the disordered two-body interaction, that is for $\epsilon \gtrsim 1.5$, upon lowering the temperature there is no spin-glass transition and the system moves directly to the ferromagnetic phase. By using Eqs. (2.37) and (2.52) one finds the expression

$$\lambda_R = \frac{1}{4\beta\epsilon q_0(1-q_0)^3} \tag{2.53}$$

from which one concludes that $\lambda_R > 0$ for all values of $q_0 \in [0, 1]$. We also find a small interval of ϵ values, approximately $\epsilon \in [0.85 : 1.5]$ where at $T = T_c$ the system encounters first a transition to the (trivial) spin-glass phase and then, upon lowering further the temperature, a transition from (trivial) spin-glass to the ferromagnetic phase. Let us stress that in Fig. 2.4 the transition line which separates the ferromagnetic phase from the other phases is a line of first order transitions: the paramagnetic phase and the trivial spin glass phase, respectively at temperatures above and below the critical one, are always stable upon increasing ϵ and the transition to the ferromagnetic phase is controlled by the free-energy balance.

We can therefore conclude that the addition of ordered non-linear interactions does not produce any sort of RSB multiple-equilibria phases in the trivial spin glass phase of the spherical model.

Notice also that this scenario holds for arbitrary ordered non-linearity. If we denote by p the power of the non-linearity, the only change in the action (2.26) is that

the magnetization term is raised to the power p. As a consequence the RS saddle point equation obtained deriving the action w.r.t q_0 (2.33) remains the same and that obtained deriving w.r.t. m reads

$$-p\beta\epsilon m^{p-1} + \frac{m}{1-q_0} = 0. \tag{2.54}$$

Therefore, the only change is in the equation for the overlap in the ferromagnetic phase, which reads

$$q_0(1-q_0)^{\frac{2(p-1)}{p-2}} \left[\frac{1}{1-q_0} - (\beta J)^2\right] = \frac{1}{(p\beta\epsilon)^{\frac{2}{p-2}}}, \tag{2.55}$$

from which one immediately recognize that

$$\lambda_R = \left(\frac{1}{p\beta\epsilon q_0(1-q_0)^{p-1}}\right)^{\frac{2}{p-2}} > 0 \tag{2.56}$$

for all values of p, β, ϵ and $q_0 \in [0,1]$.

2.3.2 Disordered nonlinearity

We now consider a different kind of perturbation to the spherical $p = 2$ model, namely we add a 4-body interaction term with quenched random couplings. The model is defined by the Hamiltonian

$$\mathcal{H}_{\rm dis}(\boldsymbol{\sigma}) = \mathcal{H}_J(\boldsymbol{\sigma}) - \sum_{i<j<k<l}^{1,N} J^{(4)}_{ijkl}\ \sigma_i\sigma_j\sigma_k\sigma_l, \tag{2.57}$$

where the couplings J_{ijkl} are i.i.d. random variables following the Gaussian distribution

$$p(J_{ijkl}) = \frac{1}{\sqrt{2\pi\sigma_4^2}} e^{-\frac{J^2_{ijkl}}{2\sigma_4^2}}, \tag{2.58}$$

with variance

$$\sigma_4^2 = \frac{\epsilon^2\ 4!}{2N^3}. \tag{2.59}$$

By averaging over the disorder, one reaches the following form for the replicated partition function:

$$\overline{\mathcal{Z}^n} = \int \prod_{a<b}^{1,n} \mathrm{d}q_{ab}\ e^{-NG[\mathbb{Q}]}, \tag{2.60}$$

where the action $G[\mathbb{Q}]$ is defined as

$$G[\mathbb{Q}] = -\frac{(\beta J_2)^2}{2}\sum_{a\le b}^{1,n} q_{ab}^2 - \frac{(\beta\epsilon)^2}{2}\sum_{a\le b}^{1,n} q_{ab}^4 - \frac{1}{2}\ln\det\mathbb{Q}. \tag{2.61}$$

Let us notice that here, at variance with the model studied in Sec. 2.3.1, due to the absence of any ferromagnetic term in the Hamiltonian there is only one ordered parameter, the overlap matrix $\mathbb{Q}$.

RS instability and RSB equations

It can be found that in the present case the RS saddle point equation is

$$q_0 \left[(\beta J_2)^2 + 2(\beta\epsilon)^2 q_0^2 - \frac{1}{(1-q_0^2)} \right] = 0, \tag{2.62}$$

and the replicon is

$$\lambda_R = -(\beta J_2)^2 - 6(\beta\epsilon)^2 q_0^2 + \frac{1}{(1-q_0)^2}. \tag{2.63}$$

By plugging the expression of Eq. (2.62) into Eq. (2.63) it is easy to realize that λ_R is negative for any value of the parameter q_0 different from zero, so that in the RS case only the solution $q_0 = 0$ can be stable. By then plugging the latter into the expression of the replicon we get $\lambda_R = 1 - (\beta J_2)^2$, which for temperatures $T < T_c = J_2$ becomes negative, signalling the instability of the RS ansatz and therefore forcing us to consider RSB solutions.

The first attempt to go beyond a RS ansatz it is always represented by considering one step of replica symmetry breaking (1RSB), corresponding to an overlap matrix, which we may refer to as $\mathbb{Q}^{\text{1step}}$, with the following structure:

$$[\mathbb{Q}^{\text{1step}}]_{ab} = (1-q_1)\delta_{ab} + (q_1 - q_0)\gamma_{ab} + q_0 \mathbb{I}_{ab}. \tag{2.64}$$

$\mathbb{I} = \mathbf{1} \otimes \mathbf{1}^T$ is a matrix whose elements are all identically equal to 1, while γ is a block diagonal matrix, with diagonal blocks all equal to $\mathbb{I}_x = \mathbf{1}_x \otimes \mathbf{1}_x^T$, namely $x \times x$ square matrices with all elements identically equal to one. For the 1RSB ansatz we have therefore three variational parameters: q_0, q_1 and x. In the present case the inverse matrix of matrix q_{ab} reads as:

$$[(\mathbb{Q}^{\text{1step}})^{-1}]_{ab} = A\delta_{ab} + B\gamma_{ab} + C\mathbb{I}_{ab}, \tag{2.65}$$

with

$$A = \frac{1}{1-q_1} \tag{2.66}$$

$$B = -\frac{(q_1 - q_0)}{(1-q_1)(1-q_1+x(q_1-q_0))} \tag{2.67}$$

$$C = -\frac{q_0}{[1-q_1+x(q_1-q_0)]^2}. \tag{2.68}$$

Given this structure of the overlap matrix, with very similar calculations to the RS case and defining the function

$$\chi_p = \chi_p(q_0, q_1, x) \equiv 1 - q_1^p + x(q_1^p - q_0^p), \tag{2.69}$$

we get as 1RSB effective action:

$$\lim_{n\to 0} \frac{1}{n} G[q_0, q_1, x] = -\frac{(\beta J_2)^2}{4}\chi_2 - \frac{(\beta\epsilon)^2}{4}\chi_4 - \frac{x-1}{2x}\log(1-q_1) - \frac{1}{2x}\log(\chi_1) - \frac{1}{2}\frac{q_0}{\chi_1}. \tag{2.70}$$

The 1RSB saddle point equations read as:

$$\frac{\partial G}{\partial q_0} = \frac{x}{2} q_0 \left[(\beta J_2)^2 + 2(\beta)^2 q_0^2 - \frac{1}{\chi_1^2} \right] = 0, \tag{2.71a}$$

$$\frac{\partial G}{\partial q_1} = (x-1) \left[-q_1 (\beta J_2)^2 - 2q_1^3 (\beta\epsilon)^2 + \frac{q_0}{\chi_1^2} + \frac{q_1 - q_0}{(1-q_1)\chi_1} \right] = 0, \tag{2.71b}$$

$$\frac{\partial G}{\partial x} = -\frac{(\beta J_2)^2}{2}[q_1^2 - q_0^2] - \frac{(\beta\epsilon)^2}{2}[q_1^4 - q_0^4] + \frac{1}{x^2}\log\frac{\chi_1}{1-q_1} + (q_1 - q_0)\left[\frac{q_0}{\chi_1^2} - \frac{1}{x \cdot \chi_1}\right] = 0. \tag{2.71c}$$

Replicon and 1RSB stability

For the diagonalization of the stability matrix of the 1RSB ansatz we follow the procedure of [49].

There are three different classes of fluctuations around the 1RSB saddle point, leading to nine different eigenvalues, which reduce to seven in the limit $n \to 0$. Following [51, 52], the two relevant eigenvalues for the study of the 1RSB stability are

$$\Lambda^{(1)} = -(\beta J_2)^2 - 6(\beta\epsilon)^2 q_1^2 + \frac{1}{(1-q_1)^2} \tag{2.72}$$

$$\Lambda^{(2)} = -(\beta J_2)^2 - 6(\beta\epsilon)^2 q_0^2 + \frac{1}{(1 - q_1 + x(q_1 - q_0))^2}. \tag{2.73}$$

Fluctuations with respect to a given ansatz for the breaking of replica symmetry means fluctuations which alter the structure of the matrix $\mathbb{Q}^{\text{1step}}$. In general for these instabilities it is also possible to give a physical interpretation connected to the modification of the matrix structure. In particular, in the case of a block diagonal matrix $\mathbb{Q}^{\text{1step}}$, each row of the matrix has the same structure: one element is unity, there are $x - 1$ elements equal to q_1, those belonging to the same block $\mathbb{I}_x$, and $n - x$ elements equal to q_0. This structure of the matrix corresponds to a *"one-step"* fragmentation of phase space in disjoint ergodic components, which we may refer to as *states* or *clusters*. Configurations belonging to the same state have typical overlap q_1, while configurations belonging to different states have overlap q_0. By considering the structure of the matrix $\mathbb{Q}^{\text{1step}}$, a consistent interpretation is to regard the elements of blocks $\mathbb{I}_x$ as the typical overlap between configurations belonging to the same state and the elements outside blocks $\mathbb{I}_x$ as the typical overlap between configurations belonging to different states.

Given this scenario, we consider two kind of instabilities: 1) a single state can undergo a further fragmentation process, which corresponds to an emerging block structure *inside* $\mathbb{I}_x$, with new elements q_2 appearing within $\mathbb{I}_x$ (again in a block-diagonal structure); 2) different states can merge into a single one, which corresponds to a rearrangement of the structure of the whole $\mathbb{Q}^{\text{1step}}$, with *all* inner elements of different blocks $\mathbb{I}_x$ changing value from q_1 to q_2 and *some* of the extra-block elements rising from q_0 to q_1. These two patterns to alter the 1RSB structure of $\mathbb{Q}^{\text{1step}}$ are connected respectively to the eigenvalues $\Lambda^{(1)}$ and $\Lambda^{(2)}$. In particular, the fluctuations inside the same cluster represented by $\Lambda^{(1)}$ are the ones which destabilize the 1RSB phase in the SK model and that

lead, across an infinite sequences of further other breakings, to the well known fractal free-energy structure [38], that we described in Sec. 2.2. This instability pattern for the 1RSB phase is the most common. In particular it is also found in ecological models studied through replica theory [26], as we have discussed in Sec. 1.5.

On the other hand, for mixed p-spin models, i.e., models with spherical variables and competition between linear and non-linear interactions, it turns out the dominant instabilities are those related to $\Lambda^{(2)}$, as demonstrated in [52] by showing that $\Lambda^{(1)} > \Lambda^{(2)}$, so that the latter is the dominant eigenvalue. This means that the *relevant* replicon is the one related to fluctuations between clusters. In this situation it was pointed out in [52] that an intermediate phase also emerges, the so-called *1-full-RSB* phase, characterized by the coexistence between a one-step and full replica symmetry breaking.

Phase diagram

In the mixed p-spin with purely disordered nonlinearity we keep fixed the variance of the two-body disordered interaction, as we did in Sec. 2.3.1, and study the phase diagram in the plane (ϵ, T), where in the present case ϵ represents the standard deviation of the four-body couplings disorder. We show in Fig. 2.5 that the lower limit of stability of the paramagnetic phase is still $T_c = J_2$. For ϵ high enough one finds at all temperatures a transition to a 1RSB glassy phase. Then, depending on the temperature, upon lowering ϵ we can either find a transition to the paramagnetic phase, for $T > T_c$, or a transition to a full-RSB phase, for $T < T_c$, (either *1-full-RSB* or standard full-RSB).

Let us remark the difference between the transition taking place from the 1-RSB phase to the paramagnetic one above T_c and the one taking place from the 1-RSB to the full-RSB below T_c. Considering the probability distribution of the overlap between replicas, which corresponds to the parametrization of the matrix $\mathbb{Q}^{\text{1step}}$, the transition at temperatures $T > T_c$ has the features of a first-order phase transition, while the transition at temperatures $T < T_c$ has the features of a second order phase transition. This means that for $T > T_c$, upon reducing ϵ, the 1RSB ansatz for $\mathbb{Q}^{\text{1step}}$ is always a *locally stable* solution, even when it becomes less convenient than the RS ansatz from the point of view of free energy, until when it completely disappears as a solution at the line marked in the phase diagram 2.5 as *1RSB spinodal*. This is the transition mechanism typical of first-order transitions.

On the other hand, for $T < T_c$ and sufficiently small ϵ, the phase described by the 1RSB solution makes way for a 1-full-RSB phase, with the typical mechanism of second order phase transitions, i.e the former loses stability in favour of the latter. Eventually, the 1-full-RSB will become itself unstable on the transition line towards the full-RSB phase at even smaller values of ϵ [52]. The transition line in Fig. 2.5 has been obtained by finding the values of the parameters where $\Lambda^{(2)} = 0$.

2.3.3 Discussion

In Secs. 2.3.1 and 2.3.2 we have probed the stability of a specific disorder system, i.e. the spherical model, by studying its response to both ordered and disordered

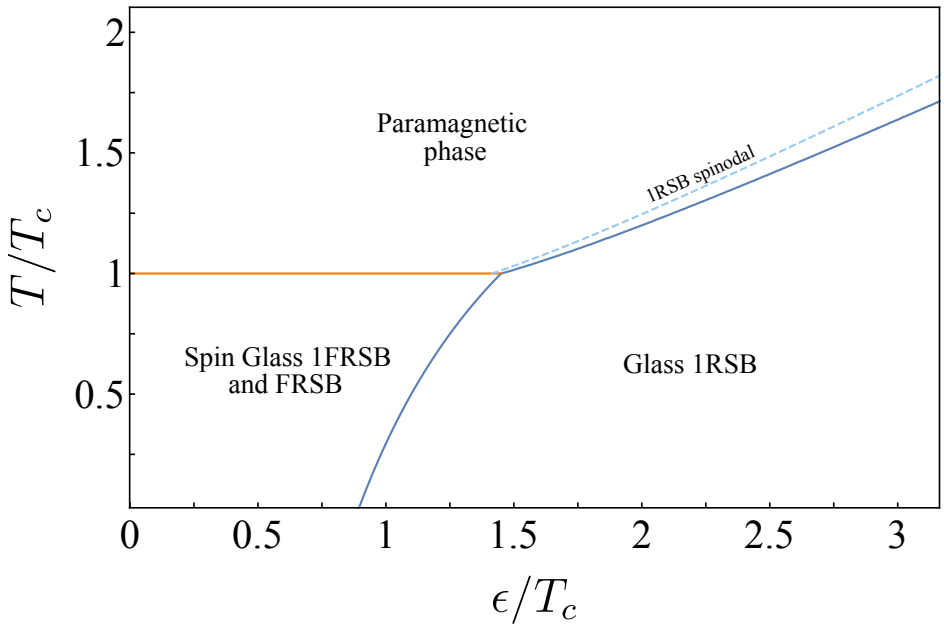

Figure 2.5: Phase diagram in the (ϵ, T) plane for the $p = 2$ spherical spin glass plus four-body *purely disordered* interactions, where ϵ represents the strength of non-linear interactions.

perturbations. Our focus was on identifying which types of perturbations can induce a non-trivial replica symmetry breaking (RSB) pattern within the low-temperature spin-glass phase of the model, which is known to be marginally stable.

We found that ordered non-linear interactions are ineffective in destabilizing the marginal spin-glass phase: no RSB pattern emerges. In this sense, the phase remains effectively stable with respect to such ordered perturbations. In contrast, when we introduce disordered four-body couplings, even with infinitesimal strength, the system reacts differently. The trivial spin-glass phase then truly shows its marginal character and any finite amount of disordered non-linearity is sufficient to trigger a non-trivial RSB pattern.

In [46], we extended this analysis to the case where the disordered four-body interactions are drawn from a Gaussian distribution with a non-zero mean. This setup leads to an interesting competition between the ordered and disordered components of the perturbation. The resulting phase diagram is even richer, displaying multiple distinct phases, among which an RSB phase emerges once again. Details on this scenario are shown in App. B.

It is worth emphasizing that in the $2+4$ spherical model, the instability of the 1RSB phase unfolds through the merging of 1RSB states into broader metabasins. This contrasts with the behaviour observed in the SK model and in many other systems, including ecological models, where instability tends to arise from the fragmentation of states into smaller and smaller clusters. This distinction reveals different underlying mechanisms of instability in disordered systems, and may help us understand which kinds of losses of stability are relevant in large ecological communities.

2.4 Mean-field ecological models

The disordered systems techniques introduced in Sec. 2.3 have been widely applied to the analysis of mean-field ecological models.

As discussed in Sec. 1.5, generalized Lotka–Volterra models for large ecological communities with dense interaction matrices display transitions to RSB multiple-equilibria phases. As ecosystem diversity increases, glassy multi-attractor regimes emerge. These phases arise from increasing heterogeneity in species interactions, which increases the frustration in the system and leads to the emergence of many distinct equilibrium states that the ecological community can relax to.

Such multiple-equilibria regimes have been shown to be analogous to the spin-glass phases encountered in disordered systems in condensed matter physics [27]. In particular, the analytical characterization of multiple-equilibria phases has been made possible not only through replica theory [26], as discussed in Sec. 2.3, but also via other mean-field approaches like Dynamical Mean-Field theory (DMFT) [53] and the Dynamical Cavity Method [28].

Without entering into the technical details of the latter two, thoroughly reviewed in [54], it is enough to say that both rely on deriving the effective dynamics of a *typical* species, where interactions with all other species are treated as a source of stochastic noise. In this framework, the quenched disorder in the interaction matrix of the original gLV equations is effectively encoded as noise in the species' abundance trajectories.

Similarly, replica theory provides tools to describe only the typical behaviour of a species, rather than the full set of species-specific dynamics, for instance by formulating an effective Hamiltonian for a representative species abundance [27, 55].

Besides the identification of multiple-equilibria phases, which offer fundamental insights into ecosystems stability but are not directly observable in real ecological systems, these theoretical models can also be used to derive more empirical ecological patterns.

Among the most commonly studied observables in ecological communities are species abundance distributions. These can be computed within theoretical frameworks and directly compared with empirical distributions. What emerges from replica computations [55], dynamical mean-field theory [53], and the dynamical cavity method [28] is that species abundance distributions are typically Gaussian. More precisely, since species abundances cannot be negative, the resulting distributions are truncated Gaussians. This feature remains robust across different parameter regimes: for instance, increasing interaction diversity eventually leads to extinctions, which result as a peak at zero abundance in the distribution. However, the shape of the bulk distribution remains unchanged, consistently maintaining its truncated Gaussian form.

These results disagree with real ecological data, where species abundance distributions are highly non-Gaussian, typically following log-normal or Gamma-like distributions [13, 30, 56], and exhibit much heavier tails than Gaussian ones. This discrepancy highlights the need to go beyond these mean-field techniques in order to accurately capture the phenomenology of ecosystems.

2.5 Disordered systems beyond mean-field

All the ecosystems properties discussed in the previous literature, based on mean-field statistical mechanics techniques and introduced in Sec. 2.4, have been derived under the assumption of a dense interaction matrix. In other words, they rely on the simplifying hypothesis that large ecosystems are fully connected, meaning that each species interacts with all the others in the community. This assumption, while mathematically convenient and essential for mean-field computations, is ecologically unrealistic. In real ecological communities, species typically interact with only a limited subset of other species. Thus, the fully-connected hypothesis, although central to the analytic tractability of previous models, is definitely one that must be challenged.

What happens when we relax this assumption? Do the same results and phenomenology hold? In particular, regarding ecosystem stability, do we still observe a transition to complex multiple-equilibria phases once the single-equilibrium regime becomes unstable? Once again, disordered systems and spin glass theory can offer valuable insights, or at least provide useful analogies, for our understanding of ecosystems.

In particular, as shown in [44], the glass transitions, and thus the multiple-equilibria phases, found in some mean-field spin systems can actually disappear in systems with low connectivity. The authors argue that the mean-field scenario in certain disordered spin models is rather fragile: as the connectivity, and hence the effective dimensionality, of the system decreases, the existence of glassy phases is no longer guaranteed. Their analysis focuses on Potts models, and while they do not identify a general mechanism that forbids the presence of multiple equilibria, they do find that in these models glassy transitions vanish when the connectivity becomes sufficiently low.

This phenomenon can be qualitatively understood using the coloring model, which at zero temperature corresponds to a disordered antiferromagnetic Potts model. In the fully-connected case, the model exhibits a large degree of frustration: each node prefers to adopt a different "color" or state from the nodes it is connected to, and since each node is connected to all others, the system faces many competing constraints. Even when the number of available colors is large, the high connectivity ensures a high level of frustration, which is the main driver behind the emergence of a multiple-equilibria phase. In a finite-connectivity regime instead, the number of such constraints is much smaller. With few connections and a sufficiently large number of colors, frustration can vanish entirely, along with the associated glassy behaviour.

This raises an important question for ecosystems: could the same mechanism apply here? Do the glassy phases observed in fully-connected models survive when we consider more realistic ecological networks with low connectivity?

A contrasting result is presented in [57], where the authors study a ferromagnetic multi-spin model with finite connectivity and discrete spins. Their analysis shows that a glassy, multiple-equilibria phase can still emerge even in the sparse (finite-connectivity) regime. This result could suggest the opposite conclusion to that of [44], namely that glassy phases might persist in sparse ecological networks as well. However, the model analyzed in [57] involves Ising-like variables, which can only take two values, $+1$ and -1. In contrast, species abundances, while discrete, typically span a much broader

range of possible values. This feature makes ecological systems more comparable to the setting studied in [44].

All these considerations remain hypothetical, but they clearly motivate our interest in studying sparse ecosystems. In addition to being a more realistic scenario, several indications suggest that their phenomenology could differ significantly from that of fully-connected systems. For this reason, the next chapter, the core of this thesis, is devoted to an exhaustive study of the equilibrium properties of the generalized Lotka–Volterra equations on sparse interaction networks, highlighting the surprising differences that emerge when compared to fully-connected models.

3

Generalized Lotka-Volterra model on sparse interaction networks

3.1 Introduction

Since the pioneering works of May [4, 5], recent advances in statistical mechanics have opened new avenues for understanding the stability and phase behaviour of large ecosystems. In particular, the equilibrium statistical mechanics approach to the generalized Lotka-Volterra (gLV) model has emerged as a powerful framework for studying species coexistence in complex ecological networks.

When interactions are both symmetric and dense, the equilibrium phase diagram can be determined exactly. As discussed in Chapter 1, the equilibrium phases of the gLV model with disordered interactions are controlled by three parameters: the average interaction strength μ, the width of the disordered interactions distribution σ and the thermal fluctuations amplitude T. In fully connected systems, a glassy multiple-equilibria phase typically emerges at large values of σ, for finite μ and T [26]. However, real ecological networks are neither dense nor characterized by symmetric interactions, so it is not clear whether the appearance of these multiple-equilibria phases is a true relevant phenomenon for real ecosystems or is a model artifact.

Another severe limitation of fully connected models lies in their predictions for species abundance distributions. In all such cases with normally distributed interactions, abundances follow truncated Gaussian distributions [28, 53, 55], in contrast to empirical observations, which consistently exhibit log-normal or Gamma-like distributions [13, 30, 56, 58–60]. For these reasons it is worth to investigate which aspects of the phenomenology of the gLV model are robust with respect to the dilution of interactions.

While several works have extended the gLV framework to asymmetric interactions [28, 29, 53, 61, 62], in this chapter we restrict our attention to the symmetric case, which

still allows to exploit the tools of equilibrium statistical mechanics. The key novelty of our approach lies in moving beyond dense networks: we study sparse ecological interactions which, as also pointed out in [63–65], are more representative of real-world ecosystems. This choice is further supported by empirical evidence showing that ecological networks are typically sparse [33–35, 66–68].

We focus on the case of quenched disorder, where interaction coefficients are fixed in time. The case of annealed disorder, involving time-varying interactions, is discussed, for fully-connected systems elsewhere [69].

The mean-field techniques introduced in Chapter 2, such as the replica method, are formulated for fully connected systems. These methods describe the macroscopic behaviour of the system in terms of global order parameters, averaging over all microscopic degrees of freedom. While this framework is extremely powerful and provides exact results in the thermodynamic limit for dense networks, it inherently neglects local structural details. When the interaction network is sparse, a global order parameter cannot be defined and the replica approach cannot be applied. In such cases, one must turn to a different set of tools: the *Cavity Method* and the associated message-passing *Belief Propagation* (BP) algorithm. These approaches preserve local heterogeneities and are well suited to the analysis of sparse disordered systems.

The central focus of this chapter, based on the work presented in [70], is the behaviour of the gLV model on a Random Regular Graph (RRG), which we will also refer to as *Bethe lattice*. In an RRG, each node (representing a species) interacts with exactly k others, where k is the connectivity of the graph (in the following, we will also refer to the connectivity of a species as its *degree*). We concentrate on graphs with small connectivity, reflecting the sparsity typical of ecological networks. The Random Regular Graph is a paradigmatic example of a *locally tree-like* random graph: although loops are present, their typical length scales as $\log N$, where N is the number of nodes, and therefore diverges in the thermodynamic limit. This property makes RRGs particularly suitable for analytical approaches such as the Cavity Method, which rely on local tree-like structure.

A crucial difference with respect to fully connected models is that sparsity may lead to a partial loss of universality: equilibrium properties can depend on local structural details and on the specific choice of interaction network or coupling statistics. While we focus on fixed connectivity graphs for simplicity, similar results are expected to hold for other sparse random tree-like topologies such as Erdős–Rényi graphs, where the degree distribution is Poissonian and the connectivity varies from node to node. We expect the main features discussed below to be robust within this class, while a systematic comparison with other sparse graph ensembles is left for future work.

We will study the properties of the gLV model on a sparse network varying the three parameters μ, σ and T already introduced before. The variables on the network are species abundances which in our model are discrete, corresponding to integer counts of individuals. We find two main results: strong deviation from Gaussianity in abundance distributions already in the single equilibrium phase and, for large enough values of the parameter μ, a topological glass phase at finite temperature. This phase looks very different from the glass phase of fully connected models; in particular, as first noticed in [71] in the $T = 0$ case, in this phase glassiness is strongly correlated to the

presence of extinctions. The gLV model on a random regular graph was in fact already considered in [71], where a phase transition to multiple equilibria was studied at zero temperature, $T = 0$, zero disorder, $\sigma = 0$, and finite interaction strength μ. Here we extend a similar analysis to finite temperatures.

In Section 3.2 we will introduce the Generalized Lotka-Volterra model on a sparse interaction network and show which is the energy functional (Hamiltonian) which governs its equilibrium properties; in Section 3.3 we will present the framework of Cavity method and Belief Propagation (BP) equations and show how to apply these tools to the case of sparse gLV models. In Section 3.4 we will discuss in detail the deviations from Gaussianity which can be found in the abundance distributions on a sparse ecological network. We will then show that no glass phase emerges at fixed μ increasing σ, at variance with fully-connected models. Sec. 3.5 will be then devoted to study the equilibrium phase diagram of the model at zero disorder ($\sigma = 0$), varying only μ and T. We will show how, at large values of competition μ, a peculiar multiple-equilibria phase emerges. The equilibrium properties in this regime have been investigated using both the BP equations and the Langevin dynamics, finding consistency between the two methods. In Sec. 3.6 we show that also considering species abundances as continuous variables in the interval $[0, 1]$ does not affect the phenomenology shown in Sec. 3.5.

3.2 Generalized Lotka-Volterra model on the Bethe lattice

The dynamical equations of the generalized Lotka-Volterra model are:

$$\frac{dn_i(t)}{dt} = \frac{r_i}{K_i} n_i(t)\Big[K_i - n_i(t) - \sum_{j\in\partial i} \alpha_{ij} n_j(t)\Big] + \sqrt{n_i(t)}\xi_i(t), \tag{3.1}$$

where $n_i(t)$ is the abundance of species ($i = 1, ..., N$) at time t and ∂i represents the set of species which interact with species i. We will focus on random locally tree-like graphs, also named Bethe lattices, for which the cardinality of ∂i is fixed and equal to $k_i = 3$ for all i's (Random Regular Graphs), even if the methods are exact in the $N \to \infty$ limit also for a Poissonian distribution of the connectivities k_i's (Erdős-Rényi Graphs). The factors r_i and K_i are respectively the intrinsic growth rate and the carrying capacity of species i, while $\xi_i(t)$ represents a Gaussian noise with zero mean and covariance

$$\langle \xi_i(t)\xi_j(t')\rangle = 2T\delta_{ij}\delta(t - t'). \tag{3.2}$$

We consider a symmetric interaction matrix with elements $\alpha_{ij} = \alpha_{ji}$. Because of the symmetric interactions we know that the Langevin dynamics admits an equilibrium distribution of the form $P(\mathbf{n}) = \exp(-H_{\text{eff}}(\mathbf{n})/T)$, with $\mathbf{n} = \{n_1, \ldots, n_N\}$ [26, 27]. When the stochastic dynamics of Eq. (3.1) is complemented with a reflecting wall condition $n_i^{\min} = \lambda$ with $\lambda \ll 1$ for every species i, as used in numerical simulations to prevent extensive extinction due to demographic fluctuations, the effective Hamiltonian

$H_{\text{eff}}(\mathbf{n})$ reads as [26]:

$$H_{\text{eff}}(\mathbf{n}) = -\sum_{i=1}^{N} r_i \left(n_i - \frac{n_i^2}{2K_i} \right) + \sum_{(ij)\in E} \frac{\alpha_{ij}}{2} \left(\frac{r_i}{K_i} + \frac{r_j}{K_j} \right) n_i n_j + \sum_{i=1}^{N} \left[T \ln (n_i) - \ln \theta(n_i - \lambda) \right], \tag{3.3}$$

where the term with the Heaviside function $\theta(n_i - \lambda)$ accounts for the presence of the reflecting wall. The symbol E in the double sum of Eq. (3.3) denotes the set of edges connecting interacting species, which are arranged in a locally tree-like sparse random graph of fixed connectivity $k = 3$. Without any lack of generality, we will consider identical carrying capacities, $K_i = K$, and identical intrinsic growth rates, $r_i = r$, for all species. In particular, we set $r = 1$ and $K = 280$, with K chosen to be much larger than the discretization step, which is $dn = 1$ since species abundances are integers. Details on the derivation of Eq. (3.3) are given in [27] and in App. C.

Differently from previous literature [26, 28], we consider the variables n_i to take only discrete non-negative values, consistently with the interpretation of n_i as the number of individuals for the species i. In this particular case, the choice of discrete variables is also an algorithmic necessity to solve numerically the Belief Propagation equations, which will be discussed in detail in Sec. 3.3. For the present purpose we just need to know that the BP equations represent an algorithmic strategy to evaluate the equilibrium Boltzmann distribution $P(\mathbf{n}) = \exp\{-\beta H_{\text{eff}}(\mathbf{n})\}$. Due to the discreteness of n_i values, we need to modify the effective Hamiltonian of Eq. (3.3): the term $\ln \theta(n_i - \lambda)$ is dropped and the term $T \ln(n_i)$ is replaced by $T \ln(n_i + \epsilon)$, with ϵ small but finite ($\epsilon = 0.0001$) in order to regularize the distribution $P(\mathbf{n})$ at $n_i = 0$. The parameter ϵ plays the same role of λ in previous literature, avoiding the extinction of all the species of the system.

The elements of the symmetric matrix α_{ij} are taken from a Gaussian distribution with mean and variance which, consistently with the literature on disordered systems [26, 72], are defined as:

$$\text{mean}[\alpha_{ij}] = \frac{\mu}{k} = \hat{\mu}, \quad \text{var}[\alpha_{ij}] = \frac{\sigma^2}{k} = \hat{\sigma}^2. \tag{3.4}$$

We thus rewrite the version of Eq. (3.3) adapted for discrete variables as:

$$H_\alpha(\mathbf{n}) = \sum_{i=1}^{N} h_i(n_i) + \sum_{(ij)\in E} h_{ij}^{\alpha}(n_i, n_j), \tag{3.5}$$

where the label α in $H_\alpha(\mathbf{n})$ denotes a given instance of the quenched disordered couplings α_{ij}. In particular we have:

$$\begin{aligned} h_i(n_i) &= -r \left(n_i - \frac{n_i^2}{2K} \right) + T \ln (n_i + \epsilon), \\ h_{ij}^{\alpha}(n_i, n_j) &= \frac{r}{K} \alpha_{ij} n_i n_j. \end{aligned} \tag{3.6}$$

The following section is dedicated to the general introduction of the Cavity Method and Belief Propagation for sparse locally tree-like random graphs and to the applications of such tools to the case of gLV sparse systems. These methods will allow us to

sample the equilibrium Boltzmann distribution $P(\mathbf{n})$ of the discrete variables n_i and to compute their marginal distributions.

3.3 Cavity Method and Belief Propagation for sparse graphs

The interaction network we want to study is a Random Regular Graph (RRG), a locally tree-like random graph with the specific property of fixed connectivity, meaning that each variable node is connected to exactly k other nodes, where k is the connectivity of the graph. In the context of gLV systems, where interactions involve only pairs of species, each interaction corresponds to an edge that connects two variable nodes.

One of the key properties of RRGs (and in general of locally tree-like random graphs), crucial to the validity of the cavity method, is related to their loop structure. Although RRGs do contain loops, their typical size is of order $\log(N)$, where N is the number of nodes in the graph. Consequently, in the thermodynamic limit, the length of loops tends to infinity. This property means that the local neighbourhood of any given node, up to a distance smaller than the typical loop size, has the same topology of a tree, as illustrated in Fig. 3.1. We will also refer to such kind of graphs as *Bethe lattices*.

Let us now write the partition function of the model in a way which is convenient to introduce the cavity method, following the approach of [43, 73]. A system with two-body interactions between variables n_i can be represented as a set of nodes i connected by edges (ij), so that the partition function can be written as:

$$Z = \sum_{\{n_i\}} \prod_i \psi_i(n_i) \prod_{(ij)\in E} \psi_{ij}(n_i, n_j), \tag{3.7}$$

where E is the set of edges in the graph, $\psi_i(n_i)$ takes into account the local field contribution and $\psi_{ij}(n_i, n_j)$ accounts for the pairwise interaction between nodes i and j. In particular, for a model like the generalized Lotka-Volterra which is described by a Hamiltonian of the form:

$$H(\mathbf{n}) = \sum_{i=1}^{N} h_i(n_i) + \sum_{(ij)\in E} h_{ij}(n_i, n_j), \tag{3.8}$$

where $h_i(n_i)$ represents a local field term and $h_{ij}(n_i, n_j)$ the interaction term, the factors ψ_i and ψ_{ij} read as:

$$\psi_i(n_i) = \exp\left(-\beta h_i(n_i)\right) \qquad \psi_{ij}(n_i, n_j) = \exp\left(-\beta h_{ij}(n_i, n_j)\right). \tag{3.9}$$

The *Cavity method* for RRGs represents an approximation which becomes exact in the $N \to \infty$ limit. The approximation is based on the tree-like structure of the graph which is shown in Fig. 3.1: for a tree, if one removes the node i (creating a *cavity*), the marginal distributions of n_j, n_p, n_l are independent due to the absence of loops in a tree. In particular, such marginal distributions can be computed by tracing out all the remaining degrees of freedom of the subtrees rooted at j, p, l. Let us refer for example

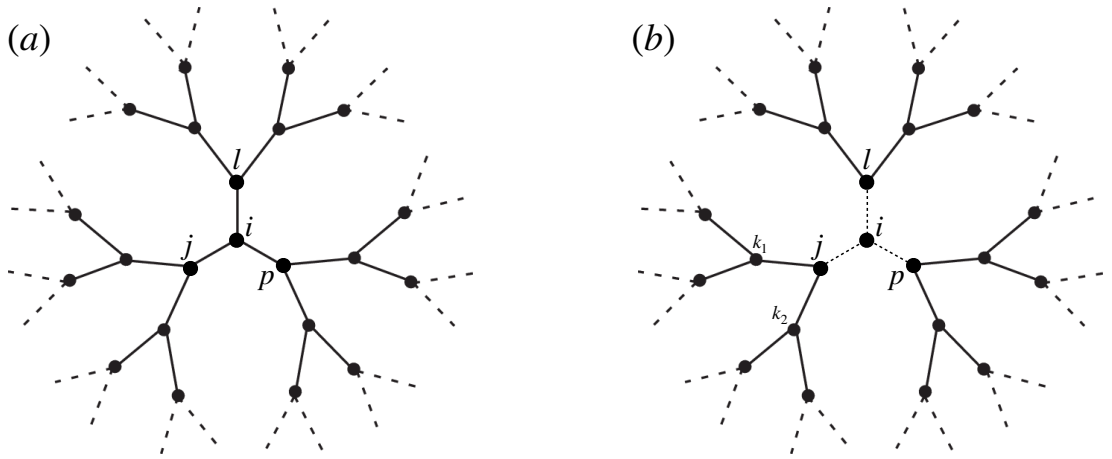

Figure 3.1: Local tree-like structure of a Random Regular Graph with 2-body interactions and connectivity $k = 3$. **(*a*)**: full graph. **(*b*)**: graph with a cavity in node i, all the edges connecting node i are removed.

to node j in Fig. 3.1(b): if the interaction graph is a tree, once we remove the edge with node i, we can define the *cavity marginal* distribution $\eta_{j\to i}(n_j)$ as

$$\eta_{j\to i}(n_j) = \frac{\psi_j(n_j)}{z_{j\to i}} \prod_{k\in\partial j\backslash i} \left(\sum_{\{n_k\}} \eta_{k\to j}(n_k)\psi_{jk}(n_j, n_k) \right), \tag{3.10}$$

where $z_{j\to i}$ is the normalization and $\partial j\backslash i$ denotes the set of neighbours of j excluding node i. Note that in Eq. (3.10) the contributions of the neighbouring cavity marginals $\eta_{k\to j}$ are multiplied because, in absence of loops, they are independent. Referring to Fig. 3.1(b), the set $\partial j\backslash i$ consists of the nodes k_1, k_2 and so, following Eq. (3.10), $\eta_{j\to i}$ will be written in function of $\eta_{k_1\to j}$ and $\eta_{k_2\to j}$. At this point, also for both the latter messages ($\eta_{k_1\to j}$ and $\eta_{k_2\to j}$) we can write an expression like Eq. (3.10), where this time the cavity marginals in the product will be the ones coming from the neighbours of k_1 and k_2 except node j. On a tree, this process can be reiterated many times till we reach the leaves (the border) of the tree. This means that, starting from the leaves of the graph and propagating the recursion, we will find an exact solution for all the cavity marginals.

In conclusion, on a tree, given a node i, we can find an exact solution for all the cavity marginals entering node i. In particular, considering Fig. 3.1, we can compute the cavity marginals $\eta_{j\to i}(n_j)$, $\eta_{l\to i}(n_l)$ and $\eta_{p\to i}(n_p)$. At this point, we can define as the *marginal* distribution of the abundance n_i of node i:

$$\eta_i(n_i) = \frac{\psi_i(n_i)}{z_i} \prod_{j\in\partial i} \left(\sum_{\{n_j\}} \eta_{j\to i}(n_j)\psi_{ij}(n_i, n_j) \right), \tag{3.11}$$

where z_i is the normalization and j indexes the neighbours of i (which for the case of Fig. 3.1 are j, l, p). While the cavity marginal $\eta_{i\to j}(n_i)$ represents the probability

distribution of the variable n_i in a modified system where the link (i, j) has been removed, $\eta_i(n_i)$ is the marginal distribution of n_i in the full system.

Finally, the free energy of the system can be determined from the collection of marginals η_i and cavity messages $\eta_{i\to j}$, together with ψ_i and ψ_{ij}. This provides a complete characterization of the thermodynamic properties of the system on a tree.

Let us stress that Equations (3.10) and (3.11) are exact on a tree, because for any given node i, the incoming cavity messages $\eta_{j\to i}(n_j)$ from its neighbours become statistically independent once the corresponding edges (i, j) are removed. This independence follows directly from the absence of loops in trees, which ensures that the subtrees rooted at different neighbours of i do not interact. As a result, the contributions of the cavity marginals from the neighbours can be simply multiplied, as it is shown in Eqs. (3.10) and (3.11). This property does not hold exactly in the case of Random Regular Graphs which, while being locally tree-like, do not have leaves and do contain loops. In a finite-size RRG, the messages $\eta_{j\to i}(n_j)$ coming from different neighbours of node i are not guaranteed to be independent and their contributions cannot be simply multiplied. In such cases, the exact expression for the cavity messages becomes:

$$\eta_{i\to j}(n_i) = \frac{\psi_i(n_i)}{z_{i\to j}} \sum_{\{n_k\},\, k\in\partial i\backslash j} \eta_{(\partial i\backslash j)\to i}(\{n_k\}, k \in \partial i\backslash j) \prod_{k\in\partial i\backslash j} \psi_{ik}(n_i, n_k), \tag{3.12}$$

where $\eta_{(\partial i\backslash j)\to i}$ is the joint distribution of the variables in $\partial i\backslash j$ in absence of the links that connect them with i. Equations (3.12), in a graph with loops, are not closed and cannot be solved. To make progress, we can assume that the cavity variables are uncorrelated, factorizing $\eta_{(\partial i\backslash j)\to i}$ and getting again equations (3.10) (and (3.11)). This factorization becomes exact in the thermodynamic limit, $N \to \infty$, due to the diverging length of loops in RRGs. As the loops become long, correlations between cavity variables decay, and the independence assumption becomes valid.

Even for finite N, one can still look for a solution of equations (3.10) on a given locally tree-like graph. While the method is no longer exact, it provides a good approximation, which improves as N increases. This procedure is known as *Belief Propagation* (BP) algorithm, and it is the approach we will use in this chapter.

The replica-symmetric nature of Belief Propagation

The effectiveness of the Belief Propagation method depends on the assumption that cavity marginals pointing to the same node are uncorrelated. This assumption can fail for two main reasons: the presence of short loops in the graph, or the emergence of long-range correlations among variables. In the Bethe lattice, because of its locally tree-like structure, short loops are avoided but long-range correlations might still be found in certain phases. This is for instance the case in disordered systems when a replica-symmetry-breaking transition takes place from an ergodic phase to a multi-state phase where the configuration space splits in "almost disjoint" ergodic components, see Chapter 2. In this case, the thermodynamic weight of the different states, together with the existence of long-range correlations, must be accounted for by the algorithm.

For this reason, the application of Eqs. (3.10) and (3.11) to sparse RRGs is commonly referred to as the *replica-symmetric cavity method*. While effective in the single-

equilibrium regime, this method breaks down in the presence of multiple-equilibria phases. In such cases, more sophisticated approaches, namely *replica-symmetry-breaking* cavity methods, are required to correctly describe the system's behaviour. We will analyze such methods in Chapter 4.

3.3.1 Belief Propagation for Generalized Lotka-Volterra models

Inserting the expressions for $\psi_i(n_i)$ and $\psi_{ij}(n_i, n_j)$ from Eq. (3.9) into the cavity equations (3.10), we find that for systems whose equilibrium properties are described by a Hamiltonian of the form:

$$H(\mathbf{n}) = \sum_{i=1}^{N} h_i(n_i) + \sum_{(ij)\in E} h_{ij}(n_i, n_j), \tag{3.13}$$

the cavity marginals $\eta_{i\to j}(n_i)$ satisfy the self-consistent recursive relation:

$$\eta_{i\to j}(n_i) = \frac{e^{-\beta h_i(n_i)}}{z_{i\to j}} \prod_{k\in\partial i\setminus j} \left(\sum_{\{n_k\}} \eta_{k\to i}(n_k) e^{-\beta h_{ik}(n_i,n_k)} \right). \tag{3.14}$$

As discussed in Sec. 3.2, for the disordered Generalized Lotka-Volterra model, the energy terms take the explicit forms:

$$h_i(n_i) = -r\left(n_i - \frac{n_i^2}{2K}\right) + T\ln(n_i + \epsilon), \qquad h_{ij}^{\alpha}(n_i, n_j) = \frac{r}{K}\alpha_{ij} n_i n_j, \tag{3.15}$$

where the superscript α in h_{ij}^{α} denotes the dependence of the interaction term on a particular realization of the quenched disorder α_{ij}. We can then write the Belief Propagation equations for the cavity marginals $\eta_{i\to j}^{\alpha}(n_i)$, which now explicitly depend on the disorder realization:

$$\eta_{i\to j}^{\alpha}(n_i) = \frac{e^{-\beta h_i(n_i)}}{z_{i\to j}^{\alpha}} \prod_{k\in\partial i\setminus j} \left[\sum_{\{n_k\}} \eta_{k\to i}^{\alpha}(n_k) e^{-\beta h_{ik}^{\alpha}(n_i,n_k)} \right], \tag{3.16}$$

with the normalization factor

$$z_{i\to j}^{\alpha} = \sum_{\{n_i\}} e^{-\beta h_i(n_i)} \prod_{k\in\partial i\setminus j} \left[\sum_{\{n_k\}} \eta_{k\to i}^{\alpha}(n_k) e^{-\beta h_{ik}^{\alpha}(n_i,n_k)} \right]. \tag{3.17}$$

Iteration procedure

To solve the self-consistent equations (3.16), we use an iterative procedure. While the cavity marginals $\eta_{i\to j}(n_i)$ are, in principle, continuous distributions, in practice we discretize the values of the variables n_i. Specifically, we assume:

$$n_i \in \{0, 1, \ldots, n_{\max} - 1\}, \tag{3.18}$$

where $n_{\max}$ is chosen large enough such that $\eta_i(n_{\max}-1) = 0$ within numerical precision for all i.

Notice that the choice to discretize the abundance distributions into non-negative integers is not inevitable. However, we do not expect the integer discretization to affect the results, since the discretization step should be compared with the natural scale of n_i, given by K (the fixed point of the single-species dynamics), which is chosen to be large for this purpose ($K = 280$).

For each oriented edge $i \to j$ in the interaction graph, a normalized initial condition is taken for the cavity marginal $\eta_{i\to j}(n_i)$, for instance, the uniform distribution $\eta^{(0)}_{i\to j}(n_i) = \frac{1}{n_{\max}}, \forall\, n_i \in [0, 1, ..., n_{\max}-1]$. Starting from $\eta^{(0)}_{i\to j}$, the cavity marginals are then updated iteratively. At each iteration step t, we compute the updated normalization constants $z^{(t+1)}_{i\to j}$ and the cavity marginals $\eta^{(t+1)}_{i\to j}$ as follows:

$$z^{(t+1)}_{i\to j} = \sum_{\{n_i\}} \exp\left\{\left(\beta r\left(n_i - \frac{n_i^2}{2K}\right) - \ln\left(n_i + \epsilon\right)\right)\right\} \prod_{k\in\partial i\setminus j}\left(\sum_{\{n_k\}} \eta^{(t)}_{k\to i}(n_k) \exp\left\{-\beta\alpha_{ik}\frac{r}{K} n_i n_k\right\}\right), \tag{3.19}$$

$$\eta^{(t+1)}_{i\to j}(n_i) = \frac{1}{z^{(t+1)}_{i\to j}} \exp\left\{\left(\beta r\left(n_i - \frac{n_i^2}{2K}\right) - \ln\left(n_i + \epsilon\right)\right)\right\} \prod_{k\in\partial i\setminus j}\left(\sum_{\{n_k\}} \eta^{(t)}_{k\to i}(n_k) \exp\left\{-\beta\alpha_{ik}\frac{r}{K} n_i n_k\right\}\right), \tag{3.20}$$

where we have explicitly written the hamiltonian terms but omitted the label α for the ease of notation. At each iteration, we also compute the marginal distributions $\eta^{(t)}_i(n_i)$ from the updated cavity marginals using:

$$\eta^{(t)}_i(n_i) = \frac{1}{z^{(t)}_i} \exp\left\{\left(\beta r\left(n_i - \frac{n_i^2}{2K}\right) - \ln\left(n_i + \epsilon\right)\right)\right\} \prod_{j\in\partial i}\left(\sum_{\{n_j\}} \eta^{(t)}_{j\to i}(n_j) \exp\left\{-\beta\alpha_{ij}\frac{r}{K} n_i n_j\right\}\right), \tag{3.21}$$

where $z^{(t)}_i$ is the normalization factor. Iteration is repeated until convergence, which is assessed as follows. At each iteration t, we compute the mean, variance, and kurtosis of each marginal distribution $\eta^{(t)}_i(n_i)$ for all nodes i. We then compute the relative change of these quantities compared to the previous iteration step. The algorithm is iterated until when, for each marginal $\eta^{(t)}_i(n_i)$, the relative change in all three moments is below 10^{-6} % compared to the previous step and this condition is maintained for ten consecutive iterations. Once these conditions are met, we assume that the algorithm has reached convergence.

3.4 The single-equilibrium phase for sparse gLV models

In what follows we present the two main results found studying the gLV on a sparse locally tree-like graph, which show a remarkable difference with the phenomenology of the same model on a dense interaction network. In particular, we find non-Gaussian, Gamma-like species abundance distributions and we state that, in sparse networks, multiple-equilibria phases driven by interaction disorder disappear.

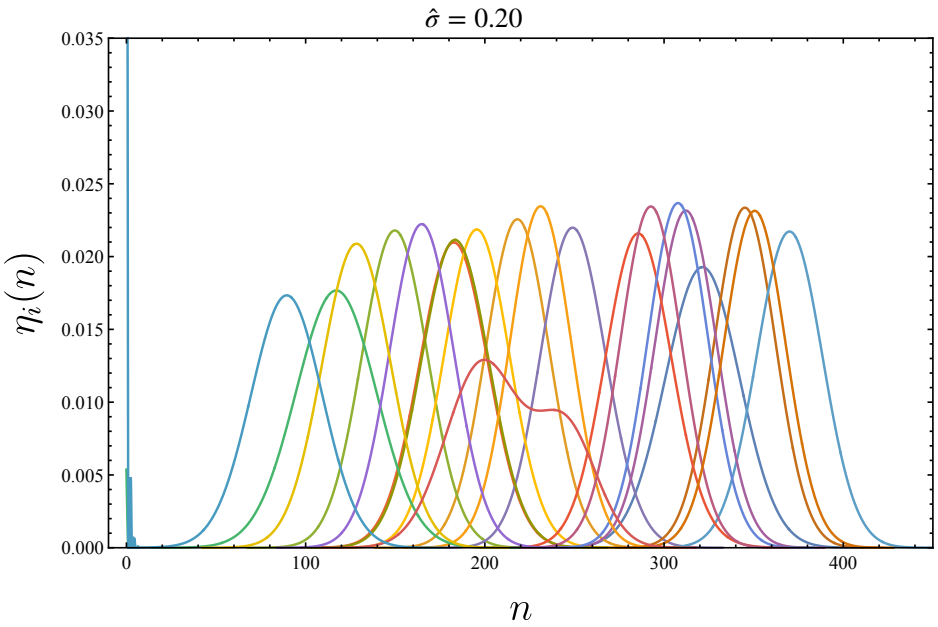

Figure 3.2: Marginal distributions for different species in a sparse system with parameters $T = 1$, $\hat{\mu} = 0.1$ and $\hat{\sigma} = 0.2$, for a particular choice of disordered couplings. The system size is $N = 256$.

3.4.1 Non-Gaussianity in the single-equilibrium phase

As thoroughly discussed in Chap. 2, the thermodynamic properties of a disordered system in the thermodynamic limit can be investigated by studying the free energy of the system averaged over all possible disorder instances. What is done numerically is to consider the iteration of the BP algorithm on the largest possible number of quenched disorder instances, where each instance corresponds to a different graph, compute the physical observables of interest for each graph once BP has converged, and then take the average over the disorder, i.e., over the different graphs. For each graph, we have a number $2|E|$ of cavity marginals $\eta^{\alpha}_{i\to j}$, where $|E|$ is the number of non-oriented edges, and a number N of marginal distributions $\eta^{\alpha}_i(n_i)$, one for each species i, given by:

$$\eta^{\alpha}_i(n_i) = \frac{1}{z^{\alpha}_i} e^{-\beta h_i(n_i)} \prod_{j\in\partial i} \left[\sum_{\{n_j\}} \eta^{\alpha}_{j\to i}(n_j) e^{-\beta h^{\alpha}_{ij}(n_i,n_j)} \right] . \tag{3.22}$$

The marginal distributions $\eta^{\alpha}_i(n_i)$ thus represent the probability for a species i to have a number n_i of individuals for a given instance of the disorder. Let us emphasize a first crucial difference compared to the fully-connected case where the marginals of all sites of a given graph are identical by construction: the iteration of the BP algorithm on a sparse graph shows that, at convergence, marginals are different site by site. This is clearly shown in Fig. 3.2 for a system of parameters $T = 1$, $\hat{\mu} = 0.1$ and $\hat{\sigma} = 0.2$. It is clear from Fig. 3.2 that marginal distributions for different species have different shapes, different peak position, some of them display extinctions (shown by the peaks in $n = 0$), some of them not.

Taking into account that on sparse graphs the marginals may have fluctuations not only with respect to the graph realization but also from species to species, we consider two kinds of average for the distributions $\eta_i^\alpha(n_i)$: the *sample* average,

$$\eta^\alpha(n) \equiv \frac{1}{N}\sum_{i=1}^{N}\eta_i^\alpha(n), \tag{3.23}$$

and the *disorder* average,

$$\eta(n) \equiv \overline{\eta^\alpha(n)}, \tag{3.24}$$

which we indicate, as customary, with an overbar $\bar{\cdot}$.

In the following we will always present marginal distributions averaged both over the sample and the disorder.

Let us stress again that on fully-connected networks, with gaussian random couplings, the local marginal $\eta(n_i)$ has, by construction, a truncated Gaussian form, with the extinction peak at $n_i = 0$. The main advantage of studying the generalized Lotka-Volterra model on a sparse topology is that the shape of local marginals $\eta_i(n_i)$ is not constrained to be a truncated Gaussian, even though the interactions are drawn from a Gaussian distribution. The remarkable finding of our study is that the average $\eta(n)$ of single-site marginals $\eta_i^\alpha(n)$ shows a remarkable deviation from gaussianity, exhibiting a Gamma-like distribution shape for a large enough value of $\hat{\sigma}$. In particular, as shown in Fig. 3.3, we find a crossover from a Gaussian to a Gamma-like shape of $\eta(n)$ upon increasing $\hat{\sigma}$. Quite remarkably this transition takes place in the single-equilibrium phase as we know from the convergence of the algorithm.

In Fig. 3.3 we show the averaged species abundance marginal distribution $\eta(n)$, Eq. (3.24), at $T = 1$ and $\hat{\mu} = 0.1$, for two different values of the standard deviation of the disordered couplings: $\hat{\sigma} = 0.02$ and $\hat{\sigma} = 0.20$. We find a crossover from a *low disorder* phase, where the mean marginal $\eta(n)$ is Gaussian, to a *high disorder* phase where $\eta(n)$ is highly non-Gaussian and is well described by a Gamma distribution. In particular, for $\hat{\sigma} = 0.20$, the distribution $\eta(n)$ is well fitted by the red curve in Fig. 3.3, given by:

$$\gamma(n;\alpha,\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)}n^{\alpha-1}e^{-\beta n}, \tag{3.25}$$

where $\Gamma(\alpha)$ is the Gamma function and the fitted parameters are $\alpha = 7.5$ and $\beta = 0.029$. Let us specify that the red line fit in Fig. 3.3 is a Gamma function plus a gaussian peak, centered in $n = 0$, which represents the extinctions. This is a very significant result as real ecosystems data exhibit species abundance distributions that follow a Gamma distribution [30, 56]. Moreover, we show in App. D that in this region of the parameters where the marginals exhibit non-gaussianity, variances of the single species abundance distributions follow a power-law dependence on the means of the distributions, well known as Taylor's law, documented extensively in real ecosystems (see [74]). In this region of parameters, what our sparse gLV model predicts is thus in perfect agreement with what is observed in real ecosystems.

Since the finding of a crossover from Gaussian to non-Gaussian within the single-equilibrium phase is an unprecedented result for this kind of model, we needed to

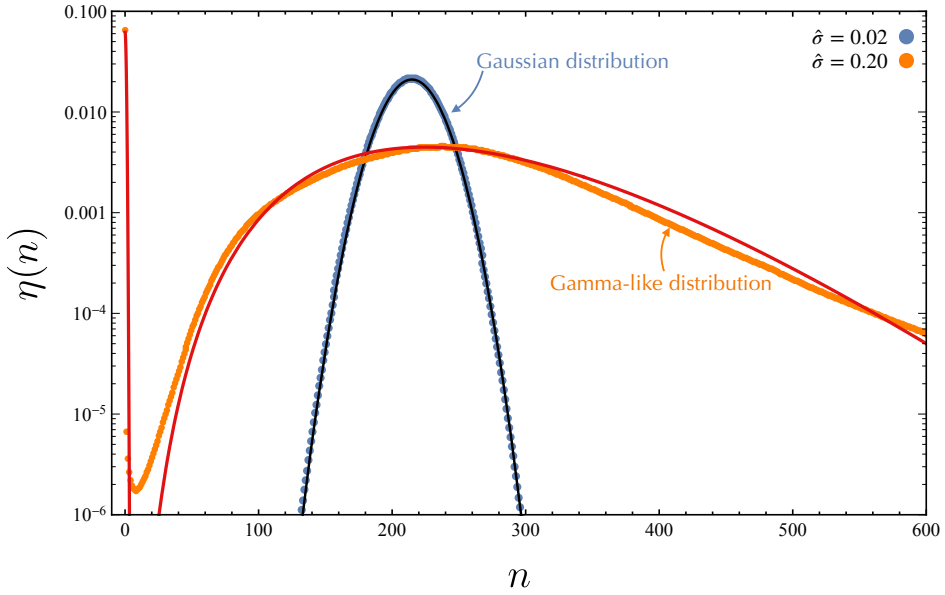

Figure 3.3: Species abundance average marginal distribution $\eta(n)$ for $T = 1$ and $\hat{\mu} = 0.1$, for $\hat{\sigma} = 0.02$ (blue) and $\hat{\sigma} = 0.20$ (orange) (Both in the single equilibrium phase). Solid lines are fits for the data: the black line is a Gaussian fit for $\hat{\sigma} = 0.02$, the red line is a fit with a Gamma distribution as in (3.25), with parameters $\alpha = 7.5$ and $\beta = 0.029$ plus a gaussian peaked at $n = 0$ for $\hat{\sigma} = 0.20$. The system size is $N = 256$ and we averaged over 100 disorder realizations.

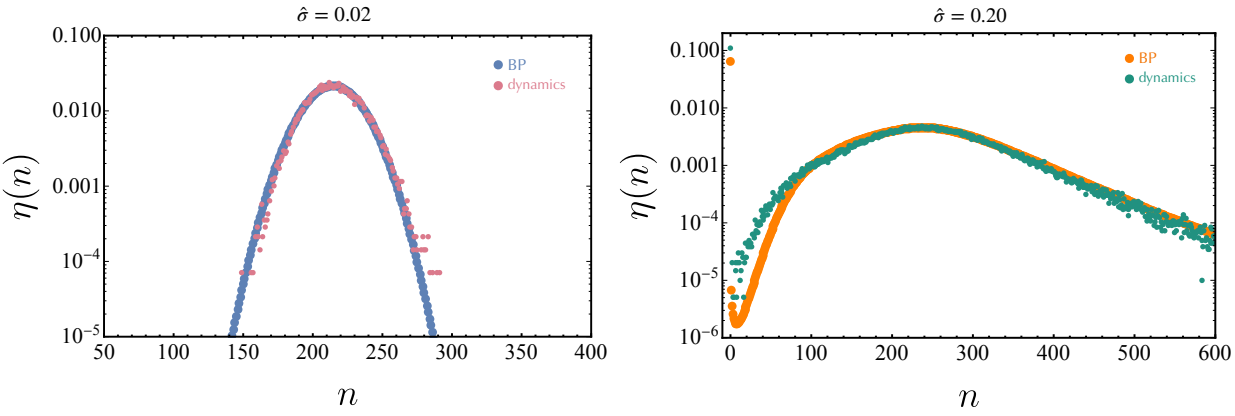

Figure 3.4: Species abundance average marginal distribution $\eta(n)$ for $T = 1$ and $\hat{\mu} = 0.1$, for $\hat{\sigma} = 0.02$ and $\hat{\sigma} = 0.20$ computed from BP and from the dynamics. The dynamics confirm the results of BP for both values of $\hat{\sigma}$. The number of species considered is $N = 256$.

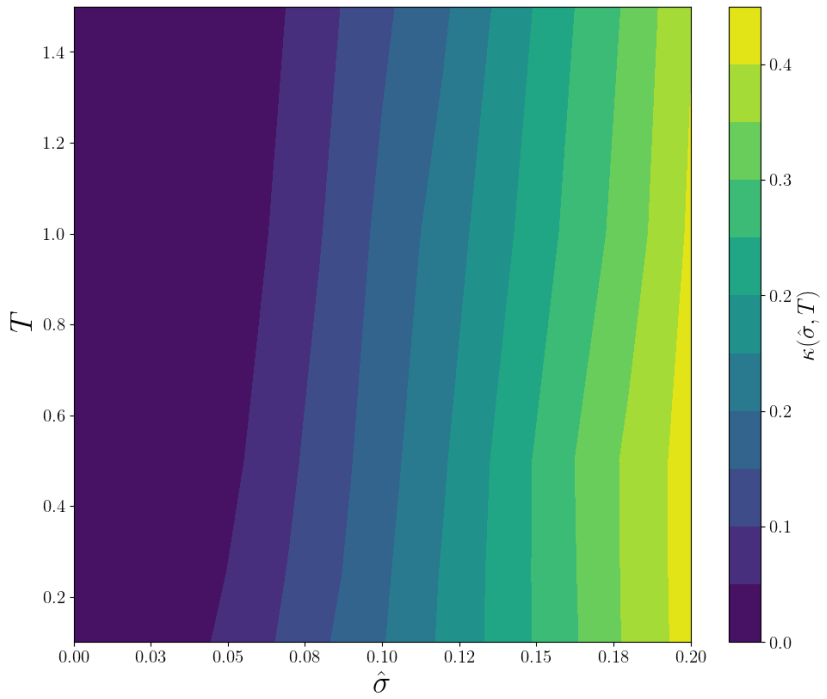

Figure 3.5: Kurtosis $\kappa(\hat{\sigma}, T)$ as a function of standard deviation $\hat{\sigma}$ and temperature T for fixed $\hat{\mu} = 0.1$ and $N = 256$. The kurtosis was computed on marginal distributions averaged over 50 disorder realizations. For high $\hat{\sigma}$ it is evident that $\kappa > 0$ implying that the marginals distributions are developing a non-Gaussian tail.

ascertain it was not an artifact of the approximations at the basis of the BP algorithm. For this reason, we also studied the exact Langevin dynamics on the Bethe lattice. This consists in fixing a disordered graph, running the set of equations given by Eq. 3.1 and stopping the simulation at a sufficiently large time $t_{\max}$ to collect all the $n_i(t_{\max})$ for every i. By repeating this process for multiple disorder realizations and combining all $n_i(t_{\max})$ from each realization, we construct the normalized histogram of species abundances, shown with pink and green points in Fig. 3.4. As we show in Fig. 3.4, the shape of both the Gaussian, for $\hat{\sigma} = 0.02$, and the non-Gaussian distribution, for $\hat{\sigma} = 0.20$, are accurately reproduced by running the Langevin dynamics. Further details on the implementation of the dynamics are provided in Sec. 3.5.2.

The crossover from Gaussian to non-Gaussian behaviour varying both the disorder strength $\hat{\sigma}$ and the temperature T is represented in the phase diagram shown in Fig. 3.5, where color code represents the kurtosis $\kappa(\hat{\sigma}, T)$ of the marginal distribution $\eta(n)$:

$$\kappa(\hat{\sigma}, T) \equiv \frac{< (n - < n >)^4 >}{< (n - < n >)^2 >^2} - 3. \tag{3.26}$$

Here the averages $< \cdot >$ are taken with respect to $\eta(n)$. For high values of $\hat{\sigma}$ it can be easily recognized the strong deviation from Gaussianity due to the non-gaussian tail of $\eta(n)$ shown in Fig. 3.3.

Let us now show that, while the correlations between species related to a transition from a single to a multi-state phase cannot be captured simply by the BP algorithm,

there are other kind of correlations which are well-reproduced by this algorithm. In particular, the transition from Gaussian to non-Gaussian behaviour is quite robust: it does not only affect the single variable marginal distribution but has also an influence on correlations. This fact is evident when studying the joint probability distribution $P_{ij}^{\alpha}(n_i, n_j)$ of a pair of nearest neighbour species on a given graph, which is defined as:

$$P_{ij}^{\alpha}(n_i, n_j) = \frac{1}{\mathcal{Z}^{\alpha}} \sum_{\{n_1,\ldots,n_N\}\setminus(n_i,n_j)} e^{-\beta H_{\alpha}(\mathbf{n})}, \tag{3.27}$$

where $\mathcal{Z}^{\alpha}$ is a normalization factor. Contrary to what one finds in fully-connected networks, $P_{ij}^{\alpha}(n_i, n_j)$ does not factorize in a sparse graph. Indeed, on a dense network, where the connectivity grows with N, the joint distribution $P_{ij}^{\alpha}(n_i, n_j)$ should simply obey a factorization property of the kind:

$$P_{ij}^{\alpha}(n_i, n_j) \xrightarrow[N\to\infty]{} \eta^{\alpha}(n_i)\eta^{\alpha}(n_j). \tag{3.28}$$

On the Bethe lattice the *two-point joint probability* on a given graph can be exactly computed as:

$$P_{ij}^{\alpha}(n_i, n_j) = \frac{1}{\Omega_{ij}^{\alpha}} \eta_{i\to j}^{\alpha}(n_i)\eta_{j\to i}^{\alpha}(n_j) e^{-\beta \frac{r}{K}\alpha_{ij} n_i n_j}, \tag{3.29}$$

where Ω_{ij}^{α} is the normalization factor. Similarly to what we have done for the single species marginal, we can define also for this two-species joint probability the *sample* average as

$$P^{\alpha}(n, m) \equiv \frac{1}{|E|} \sum_{(ij)\in E} P_{ij}^{\alpha}(n, m), \tag{3.30}$$

where both n and m represent species abundances and the summation runs over all edges in a given graph. The *disorder* average is defined as

$$P(n, m) \equiv \overline{P^{\alpha}(n, m)}. \tag{3.31}$$

The contour plots showing the behaviour of $P(n, m)$ in the plane (n, m) are reported on the two left panels of Fig. 3.6 for two different values of the standard deviation: $\hat{\sigma} = 0.02$, in the Gaussian regime, and $\hat{\sigma} = 0.20$, in the non-Gaussian regime.

On the right panels of Fig. 3.6 we show the contour plot of $P(n, m) - \eta(n)\eta(m)$ for the same two different values of $\hat{\sigma}$. It is clear that $P(n, m) \neq \eta(n)\eta(m)$ and the presence of either correlations or anticorrelations can be fully appreciated. This is an important difference compared to the fully-connected case where there are no correlations between different species. In Fig. 3.7 we compare $P(n, m)$ for $n = m$ with $\eta(n)\eta(n)$ in order to highlight again the relevance of correlations. In particular, looking at the case $\hat{\sigma} = 0.20$ in Fig. 3.7, we see that $P(n, n) < \eta(n)\eta(n)$ at small values of n, which means that positive (competitive) interactions dominate the average correlation. On the contrary we find that for large values of the abundances one has $P(n, n) > \eta(n)\eta(n)$: in this regime negative (mutualistic) interactions are the dominant ones. From the study of

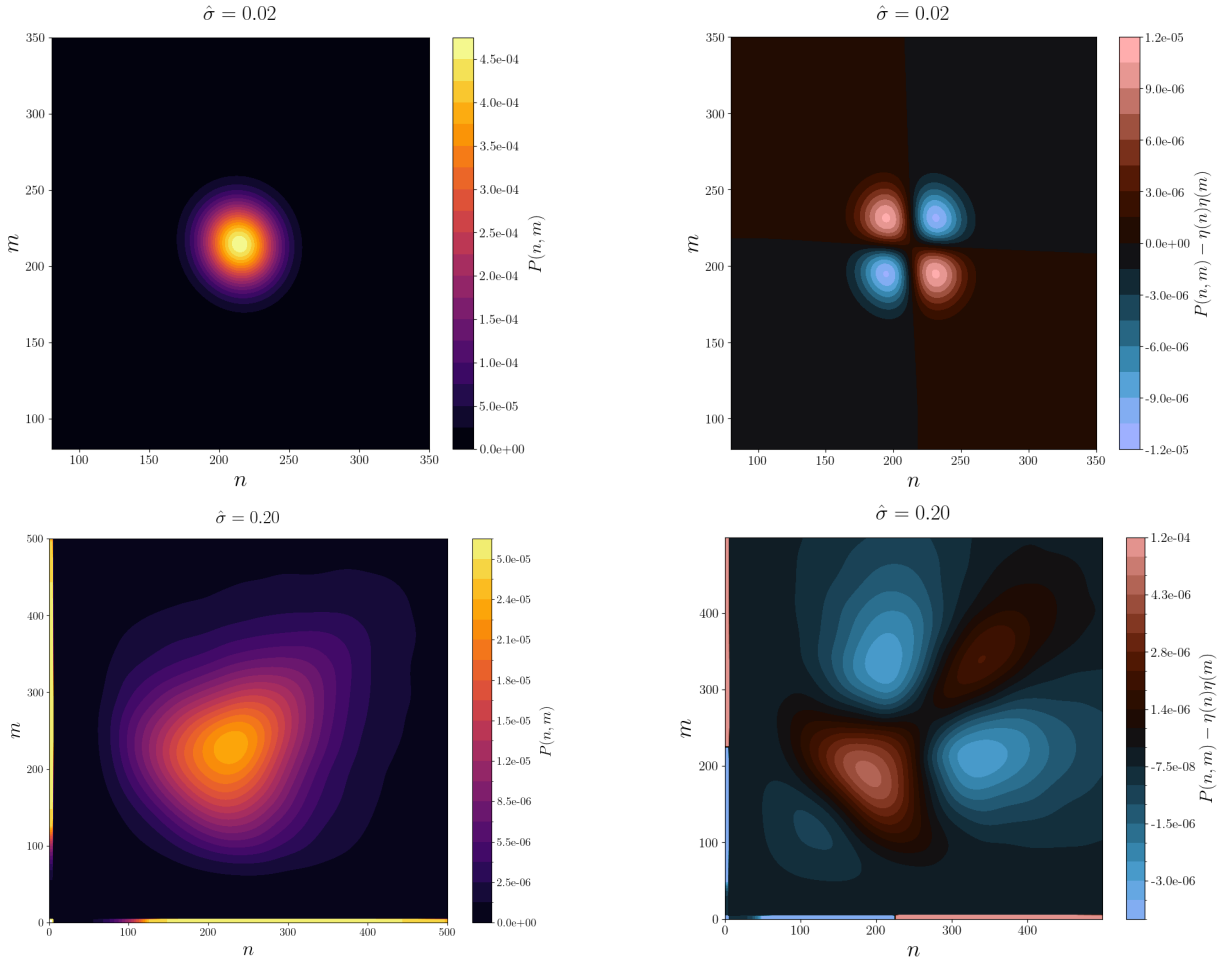

Figure 3.6: **Left:** Contour plots for the joint distribution $P(n, m)$, see Eq.(3.31), for $\hat{\sigma} = 0.02$, in the Gaussian regime (top), and $\hat{\sigma} = 0.20$, in the non-Gaussian regime (bottom). **Right:** connected correlation function $P(n, m) - \eta(n)\eta(m)$ for $\hat{\sigma} = 0.02$ (top), and $\hat{\sigma} = 0.20$ (bottom). Correlations are different from zero, at variance to what happens in the fully connected case. For all the panels $\hat{\mu} = 0.1$, $T = 1$ and $N = 256$. We averaged over 30 disorder realizations.

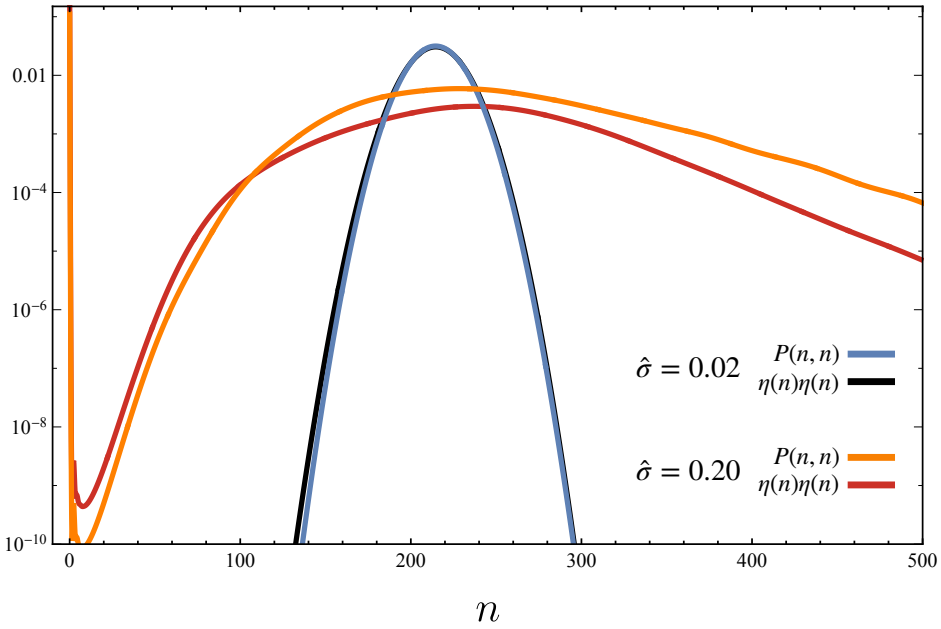

Figure 3.7: Normalized sections of $P(n,n)$ and $\eta(n)\eta(n)$ for $\hat{\sigma} = 0.02$, Gaussian regime, and $\hat{\sigma} = 0.20$, non-Gaussian regime. We have non trivial correlation effects differently from the fully connected case for which $P(n,n) = \eta(n)\eta(n)$. Also here $\hat{\mu} = 0.1$, $T = 1$, $N = 256$ and we averaged over 30 disorder realizations.

the connected probability $P(n,n) - \eta(n)\eta(n)$ it is therefore clear that non-Gaussianity affects (in a non trivial way) also the correlations when the disorder parameter $\hat{\sigma}$ is increased.

In order to separate the effect on correlations of competitive and mutualistic interactions, we have then studied the function $P(n,m) - \eta(n)\eta(m)$ in the specific cases of either exclusively positive ($\alpha_{ij} > 0$) or exclusively negative ($\alpha_{ij} < 0$) interactions. Namely, we have specialized the average of $P^{\alpha}(n,m)$ with respect to the sign of the interaction:

$$\begin{aligned} P^{(+)}(n,m) &= \overline{P^{\alpha}(n,m)}\Big|_{\{\alpha_{ij}>0\}}, \\ P^{(-)}(n,m) &= \overline{P^{\alpha}(n,m)}\Big|_{\{\alpha_{ij}<0\}}. \end{aligned} \tag{3.32}$$

The results for $P^{(+)}(n,m)$ and $P^{(-)}(n,m)$ in the case $\hat{\sigma} = 0.20$ are shown in Figures 3.8, 3.9: it is clear that competitive (positive sign) interactions lead to anti-correlations, while mutualistic (negative sign) interactions induce positive correlations.

3.4.2 Absence of glassiness for increasing disorder $\hat{\sigma}$

Let us recall that on fully-connected networks with symmetric interactions [26] a multiple-equilibria phase is found at $T > 0$ for any positive μ, provided that σ is large enough, before the transition to an unbounded growth phase, which takes place at even larger values of $\hat{\sigma}$. On the contrary, what we find here is that, on sparse networks, for finite temperatures and for $\hat{\mu} = 0.1$, by increasing $\hat{\sigma}$ the system passes

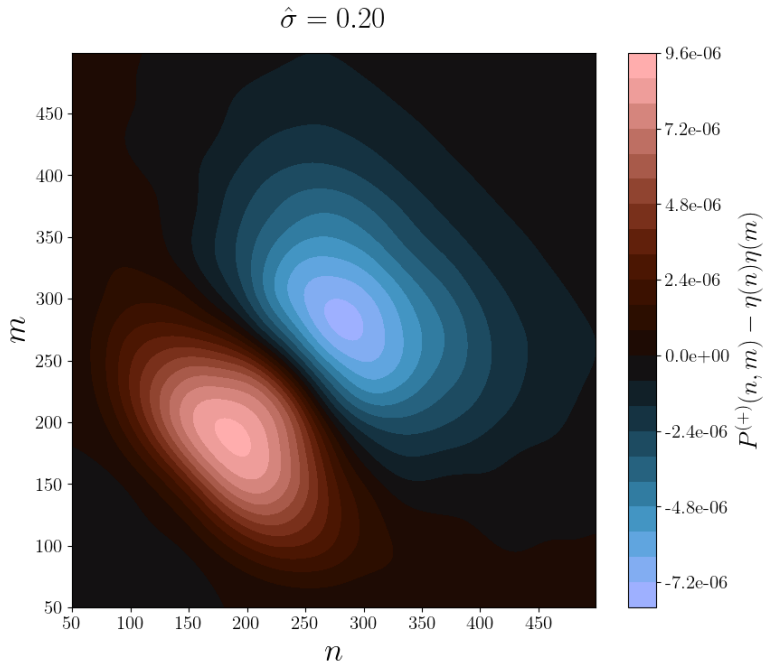

Figure 3.8: Connected correlation function for exclusively competitive $P^{(+)}(n,m) - \eta(n)\eta(m)$ interactions, for $\hat{\sigma} = 0.20$. The system size is $N = 256$ and we averaged over 30 disorder realizations.

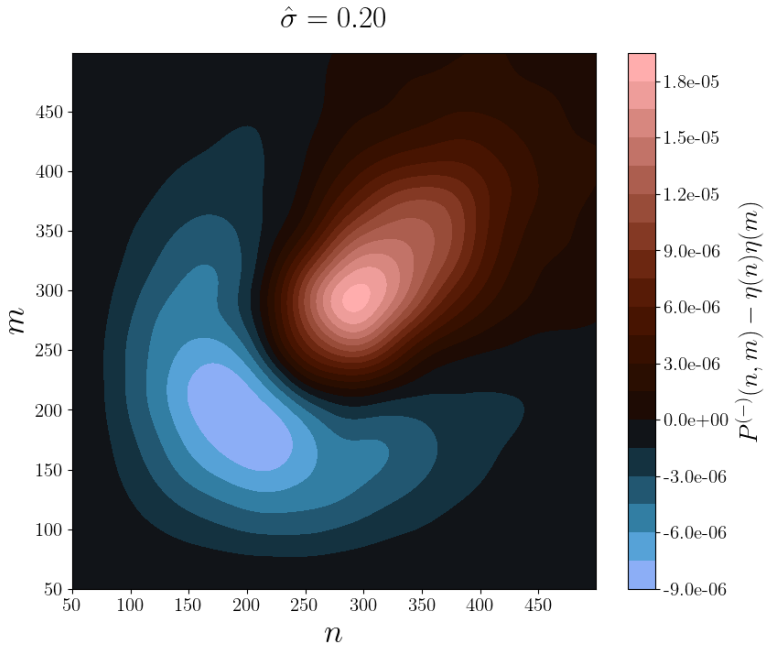

Figure 3.9: Connected correlation functions for exclusively mutualistic $P^{(-)}(n,m) - \eta(n)\eta(m)$ interactions, for $\hat{\sigma} = 0.20$. The system size is $N = 256$ and we averaged over 30 disorder realizations.

directly from the single equilibrium phase to the unbounded growth one, without any signature of an intermediate glassy phase.

To assess this result, we have scanned, for 3 fixed temperatures, a much wider interval of $\hat{\sigma}$ than the one shown in Fig. 3.5. In particular, our results show that for $\hat{\mu} = 0.1$ and temperatures $T = 0.5, 0.75, 1.0$, the BP equations always converge, and the system transitions directly from a single-equilibrium phase to an unbounded growth regime as $\hat{\sigma}$ increases, without any intermediate multiple-equilibria phase, whose presence would instead be signalled by a lack of convergence of the algorithm. This behaviour is illustrated in Fig. 3.10, where we plot the fraction N_{UG}/N as a function of $\hat{\sigma}$, with N_{UG} denoting the number of species exhibiting unbounded growth and N the total number of species. What we see in Fig. 3.10 is that the transition from the single-equilibrium phase to the unbounded-growth one arises around $\hat{\sigma} = 0.35$. Since the assessment of a transition from a single-equilibrium phase to an unbounded growth one, without passing through a glassy phase, is crucial, we need to clarify what is the conventional definition of unbounded growth (UG) on a sparse topology.

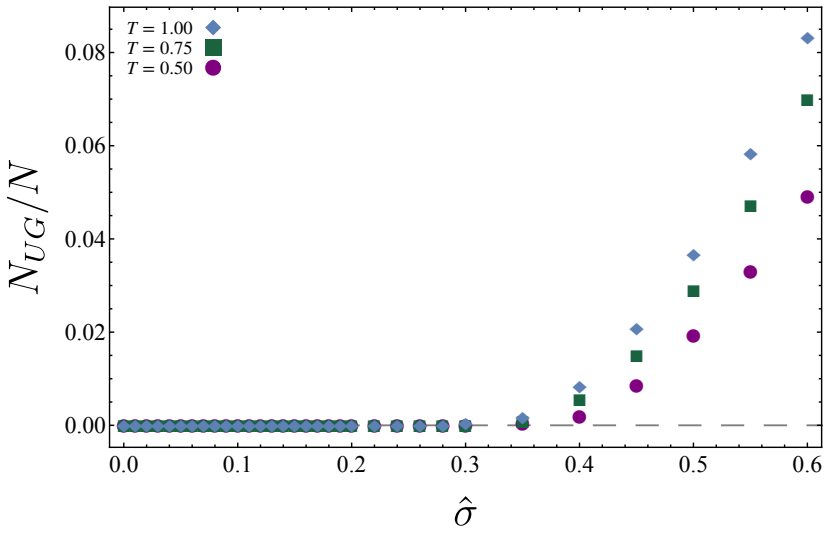

Figure 3.10: Fraction of species that grow indefinitely, for three different temperatures, for $\hat{\mu} = 0.1$ and $N = 128$. Around $\hat{\sigma} \sim 0.35$ the fraction becomes finite (it deviates from the dashed grey line that represents $N_{UG} = 0$) and the unbounded growth phase appears. For each point in the plot we considered 30 disorder realizations.

Detecting the Unbounded Growth phase

In fully connected models [28], all species behave identically, so that the UG phase is characterized by the unbounded growth of all species. On the contrary, in sparse networks species have an heterogeneous behaviour: some species may exhibit unbounded growth while others remain bounded. We can therefore explicitly identify which and how many species diverge. In this context, we have defined the UG phase as the regime in which at least one species undergoes unbounded growth. In order to state that

a certain species has undergone unbounded growth we analyze its equilibrium marginal which, we recall, is studied in the finite interval $[0, n_{\max}]$. If, after convergence, the marginal distribution for a given species j accumulates at the right boundary, i.e. $\eta_j(n_{\max} - 1) > 0$, we increase $n_{\max}$ and repeat the iteration procedure. If this accumulation persists for any value of $n_{\max}$, the species j is considered to be in the *unbounded growth phase*.

Let us stress that the transition to the UG phase that we identified in Fig. 3.10 does not depend on the choice of $n_{\max}$. In order to highlight that, here we show the differences on the final marginals in the single fixed point phase and in the Unbounded Growth phase. In Fig. 3.11 we show the marginal probabilities on a given network for three different species in the unique fixed point phase. If n_{max} is too small, the marginals for some species will be wrong: we detect the error looking if there are some species for which $\eta_i(n_{max} - 1) \neq 0$. However, enlarging n_{max} will then correct the marginal. Once we identify a value of n_{max} such that $\eta_i(n_{\max}) = 0$ inside computer precision, enlarging n_{max} further will have no effect on the marginals: this is a clear indication of the fact that the solution found for the BP equations is not influenced by our numerical implementation, as long as n_{max} is large enough.

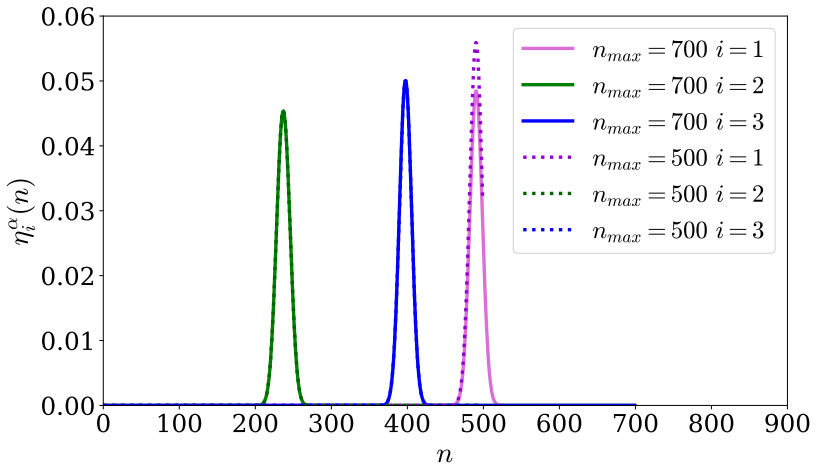

Figure 3.11: Marginals $\eta_i^\alpha(n)$ for three species changing n_{max} in the unique fixed point phase. In the unique fixed point phase, if n_{max} is too small, the marginals for some species will be wrong: we detect the error looking if there are some species for which $\eta_i(n_{max} - 1) \neq 0$, as happens for species 1 in the figure. Enlarging n_{max} will then correct the marginal. The interaction matrix here has parameters $\hat{\mu} = 0.1$, $T = 0.2$, $\hat{\sigma} = 0.20$, $N = 200$.

The behaviour in the unbounded growth phase is instead different. The numerical implementation of the BP equations with finite n_{max} is converging even in the unbounded growth phase. However, once implemented with a finite n_{max}, the BP equations are searching a solution for the marginal distributions generated not by the original Hamiltonian in eq. (3.5), but by a modified Hamiltonian in which we are manually injecting a reflecting wall at $n = n_{max}$. While in the unique fixed point phase the final marginals are independent on the value of n_{max} as long as it is large enough, in the Unbounded Growth phase some marginals will always depend on n_{max}, irrespectively on how large n_{max} is, and in particular they will always have $\eta_i(n_{max} - 1) \simeq 1$, as shown

in Fig. 3.12: this is a clear evidence that the abundance distribution for $n_{max} \to \infty$ would run off to infinity.

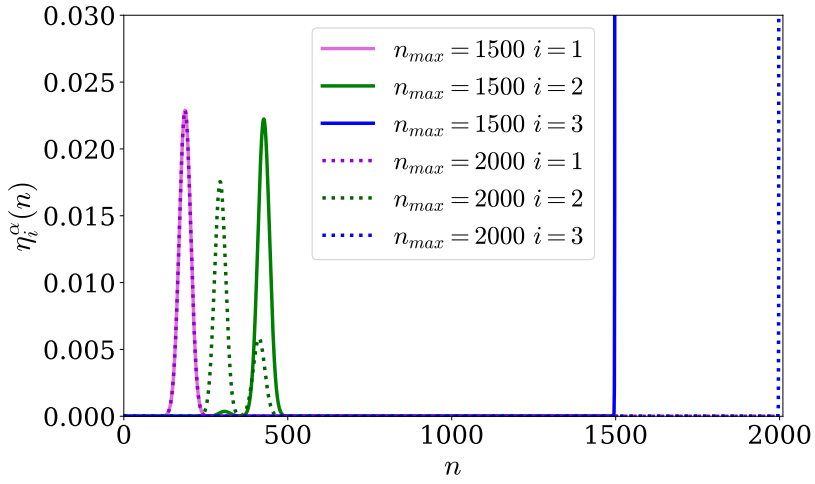

Figure 3.12: Marginals $\eta_i^\alpha(n)$ for three species changing n_{max} in the Unbounded Growth phase. While for some species changing n_{max} does not change their marginals, like species 1 in the figure, some other species have $\eta_i(n_{max} - 1) \simeq 1$ irrespectively on the value of n_{max}, like species 3. The interaction matrix here has parameters $\hat{\mu} = 0.1$, $T = 1$, $\hat{\sigma} = 0.5$, $N = 200$.

In order to provide further evidence of the consistency between our results and others from the literature, we have also computed, in the single equilibrium phase, the survival probability Φ, i.e., the ratio between the number of surviving species and the total number of species, and the average population abundance $M = \sum_{i=1}^{N} M_i/N$, as a function of the disorder strength $\hat{\sigma}$. Data for Φ and M are shown respectively in the upper and the lower panel of Fig. 3.13. In both panels the behaviour of the corresponding observable is represented up to the value of $\hat{\sigma}$ at which unbounded growth is found (see Fig. 3.10). These results are consistent with those presented in other works [28, 72]: increasing the interactions heterogeneity $\hat{\sigma}$ both the fraction of extinctions $1 - \Phi$ and the mean species abundance M of the ecosystem increase.

3.5 Multiple equilibria at zero disorder: topological glass phase

In Sec. 3.4, we showed that, in sparse ecological networks, there is no multiple-equilibria phase emerging upon the increase of $\hat{\sigma}$ for $\hat{\mu} = 0.1$. However, for a sparse interaction network, it was recently shown in [71] that, at zero temperature ($T = 0$) and zero interaction disorder ($\sigma = 0$), there is a transition from single to multiple equilibria at the critical value:

$$\hat{\mu}_c = \frac{1}{2\sqrt{k-1}}, \tag{3.33}$$

where k is the connectivity of the graph. This is a quite peculiar transition to a multiple-equilibria phase, since it is not due to disordered interactions (as we are considering the

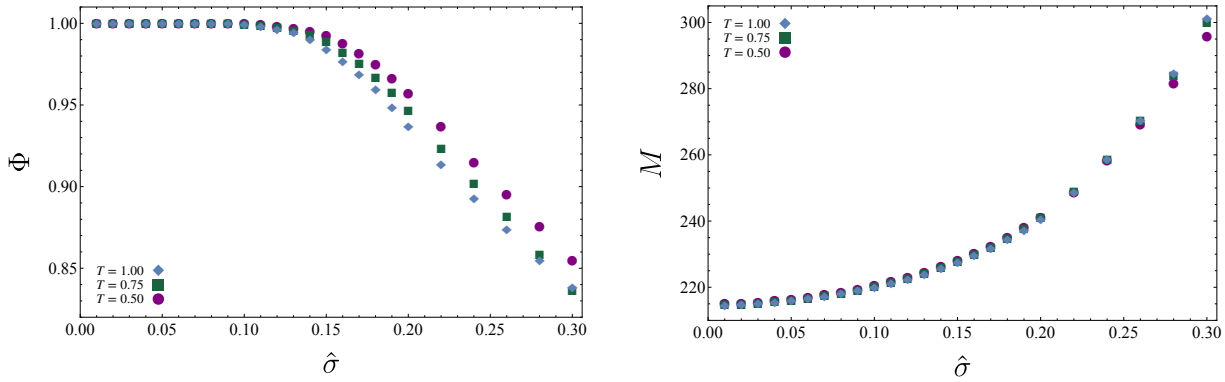

Figure 3.13: We show, for $\hat{\mu} = 0.1$ and $N = 128$, and three different temperatures, the survival probability Φ and the mean abundance M, varying $\hat{\sigma}$ up to the appearance of the unbounded growth phase. For each point in the plots we considered 30 disorder realizations.

$\sigma = 0$ case) but rather to high competition between species. For $\hat{\mu} > \hat{\mu}_c$, the increased competitive interactions cause the phase space to fragment into a large set of different equilibrium states, each characterized by the extinction of a different set of species and different pools of surviving ones. This kind of transition is absent in fully-connected models, it is peculiar of sparse topologies.

In ref. [71], the transition point $\hat{\mu}_c$ is exactly located looking at the local stability of the single equilibrium phase. Very briefly and referring to [71] for further details, the authors study the stability of the single-equilibrium by analyzing the spectrum of the stability matrix $\alpha_{ij} = \delta_{ij} + \hat{\mu} A_{ij}$, where A_{ij} is the adjacency matrix of a random regular graph with size N and connectivity k. At large N, the minimal eigenvalue of A_{ij} is $\lambda^A_{\min} = -2\sqrt{k-1}$. Therefore the minimal eigenvalue of α_{ij} is $\lambda^\alpha_{\min} = 1 - 2\hat{\mu}\sqrt{k-1}$ and the critical value of $\hat{\mu}$ which determines the transition from a stable to an unstable single-equilibrium is $\hat{\mu}_c = 1/2\sqrt{k-1}$, which corresponds to Eq. 3.33.

Motivated by the results of [71], we analyze the same transition, adding a finite temperature $T > 0$, in order to see how the presence of demographic noise influences such a transition.

First, in Sec. 3.5.1, we study the transition through the equilibrium solutions of BP and then, in Sec. 3.5.2, we simulate the Langevin dynamics.

3.5.1 Zero-disorder phase diagram from Belief Propagation equations

As already anticipated in Sec. 3.4.2, the standard criterion to assess the transition from a single-equilibrium to a multiple-equilibria phase is the lack of convergence of the BP algorithm (Sec. 3.3.1). This failure to converge means that, as we discussed in Sec. 3.3, the equations derived by assuming the existence of a single equilibrium state are not appropriate to describe the equilibrium thermodynamics of the system and more complicated replica-symmetry-breaking ansatzs (*multiple-equilibria adapted* assumptions on the equilibrium distribution) are needed to study the cavity marginals [43]. On the basis of this criterion we have computed, for the zero disorder case, $\hat{\sigma} = 0$, the

critical line in the plane $(\hat{\mu}, T)$ which separates the single equilibrium phase from a multiple-equilibria one: data are shown in Fig. 3.14. We highlight that a linear extrapolation of the transition line down to $T = 0$ nicely matches the critical value $\hat{\mu}_c$ found in [71]. The extrapolation is needed because the BP equations at finite temperatures have notoriously a pathological behaviour in the $T \to 0$ limit.

The multiple-equilibria phase found upon increasing $\hat{\mu}$ at zero disorder is quite different from the multiple equilibria phase found upon increasing σ at finite T in the fully-connected case [26]. In the sparse network the presence of many equilibria is determined by the following mechanism: when $\hat{\mu} > \hat{\mu}_c(T)$ the interactions between species become *too competitive* and this causes the extinction of some species. For even higher values of $\hat{\mu}$ the graph breaks into disconnected groups of surviving species. Since all species are equivalent, because they have the same number of neighbours and identical interactions, which species go extinct and which survive solely depend on the initial values of the marginals and on thermal fluctuations. In particular, different instances of the thermal noise can lead to the extinction of different subsets of species in the sparse network, giving rise to a glass phase which we have termed *topological*, due to its heterogeneous nature specifically tied to the sparseness of the network and which has no counterpart in dense networks. In Fig. 3.14 the critical line for the transition from single to multiple equilibria is plotted in the $(\hat{\mu}, T)$ plane: the critical value $\mu_c(T)$ decreases by increasing the temperature. This result can be easily understood as thermal fluctuations enhance the extinction of some species, shifting $\hat{\mu}_c$ further to the left as temperature increases.

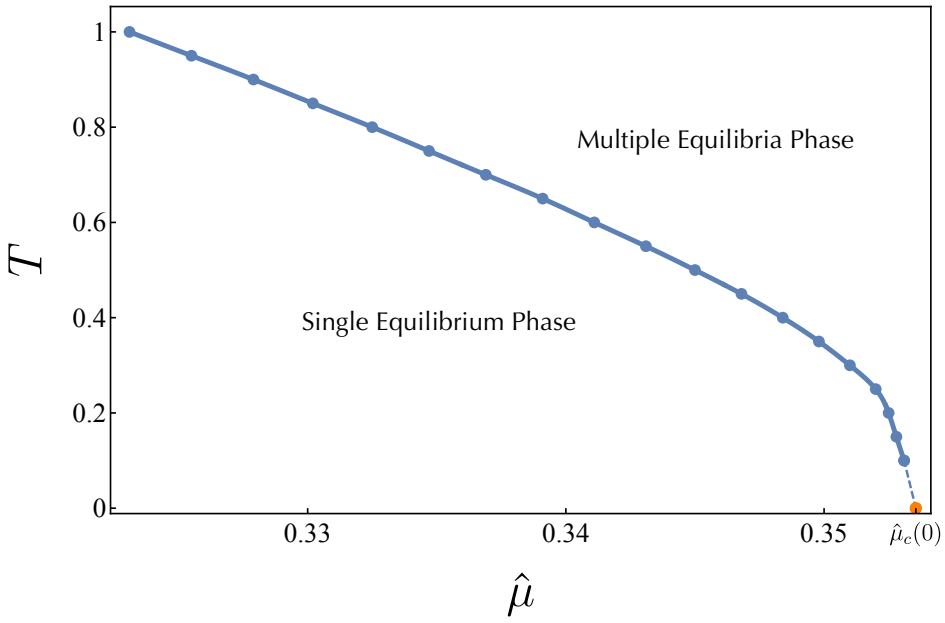

Figure 3.14: Transition line from single to multiple equilibria in the $(\hat{\mu}, T)$ space, with $\hat{\sigma} = 0$. The transition line has been computed for $N = 256$.

3.5.2 Zero-disorder phase diagram from Langevin Dynamics

In order to validate the equilibrium results obtained for the zero disordered case by means of the Belief Propagation algorithm, we simulated the Langevin dynamics at finite temperature. In this context, the Langevin dynamics is employed as a tool to sample the exact equilibrium distribution of the system without any of the assumptions at the basis of the cavity method, which is exact only in the $N \to \infty$ limit. More precisely, we used a generalization of the Runge-Kutta method for stochastic dynamics to integrate the following set of Langevin equations:

$$\frac{dn_i(t)}{dt} = \frac{r}{K} n_i(t)\Big[K - n_i(t) - \hat{\mu} \sum_{j \in \partial i} n_j(t)\Big] + \xi_i(t), \tag{3.34}$$

where the noise variance is

$$\langle \xi_i(t)\xi_j(t')\rangle = 2T\delta_{ij} n_i(t)\delta(t-t'). \tag{3.35}$$

Following [26] we have introduced a reflecting wall condition at $n_i = \lambda$ for each species.

We have found that the results from Langevin dynamics are in agreement with the results obtained from the Belief Propagation equations. In the phase which, according to BP, is of single-equilibrium, the Langevin dynamics produces trajectories with small fluctuations and no extinctions. On the contrary, as soon as one tries to run the Langevin dynamics in the multiple-equilibria phase, an abrupt spreading of $n_i(t)$ trajectories is found, accompanied by the immediate appearing of extinctions. This peculiar behaviour is shown in the panels of Fig. 3.15: data are taken from the simulations of a random regular graph with $N = 256$ species and the value of the parameters $\hat{\mu}$ and T corresponding to each panel is indicated with a symbol (different in shape and color) in the ($\hat{\mu}$,T) phase diagram in the center.

Let us comment the simulations of dynamics in Eq. (3.34) by starting from the $T = 0$ case. By looking at the panels at the bottom of Fig. 3.15 corresponding to $T = 0$ and $\hat{\mu} = 0.35, 0.353$, it is clear that Langevin dynamics converges to the same abundance value for every species. On the contrary, the bottom-right panel of Fig. 3.15, corresponding to $T = 0$ and $\hat{\mu} = 0.358$, clearly shows the appearance of extinctions and the spreading of $n_i(t)$ trajectories, consistently with the nature of the topological glass phase discussed above. The same transition from a single-equilibrium to a topological multiple-equilibria phase is found by increasing the temperature at fixed $\hat{\mu}$, provided we choose a value $\hat{\mu} < \hat{\mu}_c(T = 0)$, i.e., smaller than the critical one at zero temperature. This is clearly shown by the two top-right panels, corresponding to ($\hat{\mu} = 0.35$, $T = 0.6$) and to ($\hat{\mu} = 0.353$, $T = 0.2$), which show a spreading of $n_i(t)$ and the appearance of extinctions, at variance with the behaviour of trajectories at the same values of $\hat{\mu}$ and $T = 0$. All panels in Fig. 3.15 represent trajectories of abundances $n_i(t)$ obtained at different values of $\hat{\mu}$ and temperature T by exploiting the same initial conditions and stochastic noise realization. Clearly, different initial conditions and different realization of the noise would have led to the extinction of different species. Finally, by comparing the top-left panel ($\hat{\mu} = 0.33$, $T = 0.6$) to the top-central one ($\hat{\mu} = 0.35$, $T = 0.6$) of

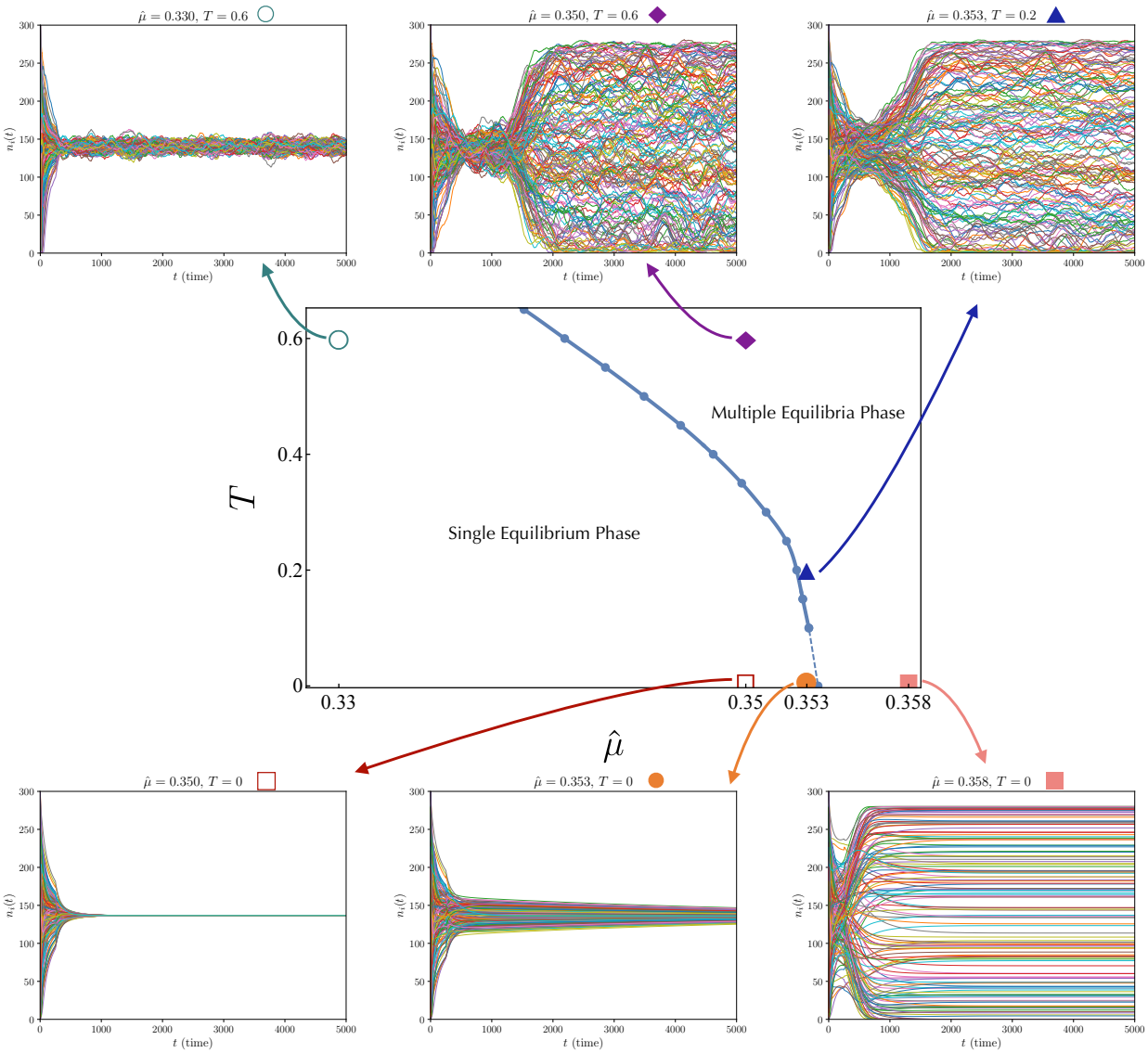

Figure 3.15: In the center, the same $(\hat{\mu}, T)$ diagram of Fig. 3.14 is reported. In this plot we identify six specific values of $(\hat{\mu}, T)$ (identified by different markers), for which we show a single dynamics for the single-species abundance populations of a single ecosystem. The multiple equilibria phase appears when extinctions appear. In the plots above, for each species i, at each time t, $n_i(t)$ is the average over the previous 250 time-steps. Here $N = 256$.

Fig. 3.15, it is clear how the same transition from single to multiple equilibria can be found by increasing $\hat{\mu}$ at fixed temperature.

While the results of Langevin dynamics are perfectly compatible with the results obtained from the Belief Propagation equations, we want to stress that from the Langevin approach the transition line between single and multiple equilibria cannot be defined as sharply as from the lack of convergence of BP, which turns out to be a much more powerful tool to study this phase diagram.

We also want to point out that it is not always true that as long as extinctions appear, the single equilibria breaks into multiple equilibria states. In fact, we showed that at low $\hat{\mu}$ and high enough $\hat{\sigma}$, there are extinctions but the single equilibrium phase is the only stable phase, see Fig. 3.13.

3.5.3 A re-entrant transition

We have shown that in the case of a sparse interaction network it is possible to find a *topological* glass phase, characterized by the presence of extinctions. In this phase, at variance with the glass phase typically found in fully-connected models, frustration is due to highly competitive interactions rather than disordered ones. To fully characterize the transition line in the plane $(\hat{\mu}, T)$ at $\sigma = 0$, we studied the convergence of the BP algorithm varying the temperature for four fixed values of the (ordered) interaction strength: $\hat{\mu} = 0.1, 0.15, 0.2, 0.3$. At $\hat{\mu} = 0.1$ we are able to heat up the system without ever losing convergence of the algorithm, a signal that no glassy phase is encountered. Instead, at $\hat{\mu} = 0.15, 0.2, 0.3$, we find a critical temperature that divides a single-equilibria phase from a multiple equilibria one increasing T from low temperature, $T_c^{\text{down}}(\hat{\mu})$ (black dots in Fig. 3.16), and another critical temperature that divides a single-equilibria phase from a multiple equilibria one decreasing T from high temperature, $T_c^{\text{up}}(\hat{\mu})$ (red dots in Fig. 3.16). These findings suggest the following scenario: there is a critical line enclosing the topological glass phase which starts at $\hat{\mu}_c$ at $T = 0$, moves at smaller $\hat{\mu}$ for increasing T and then bends to the right at a critical value of the interactions strength μ^* which, according to our study, must be located in the interval $0.1 < \mu^* < 0.15$ and goes to some horizontal asymptote at a certain critical temperature that we have not precisely located. We have just assessed that this temperature must be smaller than $T = 15$, since we have checked that at this temperature the system is in the replica symmetric phase for any $\hat{\mu}$. The re-entrant transition in the $(\hat{\mu}, T)$ plane is pictured as a gray dashed line in Fig. 3.16.

From a general perspective, the critical line sketched in Fig. 3.14 represents an ergodicity-breaking transition remarkably different from the standard replica-symmetry-breaking transition found in fully-connected graphs. In the latter case the transition from single to multiple equilibria is always triggered by the decrease of thermal fluctuations. In contrast, the sparse gLV model exhibits the opposite behaviour. The reentrant transition of Fig. 3.16 is very similar to the inverse freezing transition, present in the Random Blume-Capel model [75, 76]. It is known that, in order to have an inverse-freezing transition, it is essential the presence of active and inactive nodes,

that are modeled as bosonic spins that can take values $(+1, -1, 0)$ in the Blume-Capel model, while in the sparse gLV the active and inactive nodes correspond to surviving and extincted species.

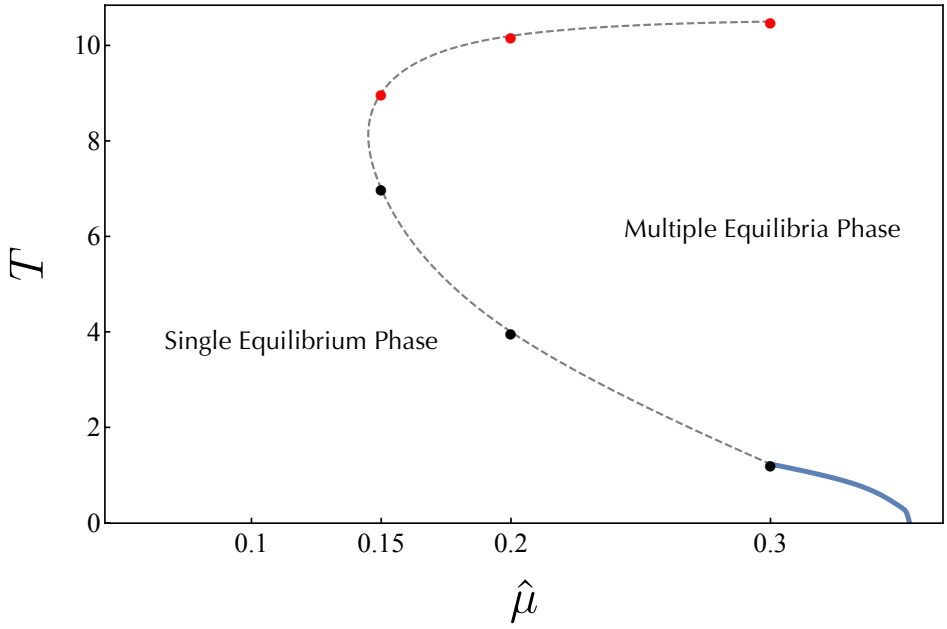

Figure 3.16: Transition line, at $\hat{\sigma} = 0$, between the single equilibrium and the multiple equilibria phase. The blue line reproduces the transition line from Fig. 3.14. The black dots are the transition points T_c^{down} found by increasing the temperature from below, while the red dots correspond the transition points T_c^{up} found by decreasing the temperature from above. Based on these points, we suggest that the transition line eventually bends to the right, forming a nose-like shape, which we represent as the grey dashed line in the figure.

3.6 Continuous gLV model at zero disorder

Let us emphasize that all the calculations in the previous sections were carried out for discrete variables, i.e. discrete species abundances. In this setting, one considers n_i as the actual number of individuals of a species: extinction ($n_i = 0$), a single individual ($n_i = 1$), or a generic population size $n_i = \overline{n}$. This approach is motivated by the very meaning of abundance as the number of individuals in a population. However, it is not the most common perspective in the literature, where species abundances are often treated as continuous variables taking values in the interval $[0, 1]$, interpreted as relative abundances. The latter choice is natural in several contexts, for instance in microbial ecology, where sequencing techniques measure gene frequencies across samples. In such cases, abundances are expressed as relative frequencies, naturally taking continuous values. Assuming that n_i is a continuous variable is also consistent with the mathematical structure of the gLV equations, which are properly defined for continuous variables. It is therefore worth to study whether and to which extent the

choice of a continuous rather than discrete n_i affects the results so far presented for the equilibrium phases of the model. In this section we show that adopting a continuous formulation of BP, as done in [77], leads to equivalent results.

Recall that the BP update rule for cavity messages $\eta_{i\to j}(n_i)$, as introduced in Ref. [70] and Sec. 3.3.1, reads:

$$\eta_{i\to j}(n_i) \propto \frac{1}{n_i+\epsilon}\, \exp\left\{-\beta r\left(\frac{n_i^2}{2K}-n_i\right)\right\} \prod_{k\in\partial i\setminus j}\left(\sum_{\{n_k\}} \eta_{k\to i}(n_k)\exp\left\{-\beta\frac{r}{K}\alpha_{ik}n_i\, n_k\right\}\right), \tag{3.36}$$

where $\epsilon>0$ prevents divergences at $n_i=0$ that would otherwise make the distribution non-normalizable. Notice that Eq. (3.36) is written in a discretized space, which is why the expression involves sums ($\sum_{\{n_k\}}...$) rather than integrals. In previous sections, we used $r=1$ and $K=280$, with K representing the carrying capacity, i.e. the stationary abundance of a species in isolation.

For continuous variables, Eq. (3.36) must be rewritten replacing sums with integrals. In this case one adds an immigration term λ to the dynamics to regularize behaviour at $n_i=0$, replacing the term $(n_i+\epsilon)^{-1}$ of Eq. (3.36) with $n_i^{\beta\lambda-1}$. The BP recursion then, with $K=1$ and $r=1$, takes the form:

$$\eta_{i\to j}(n_i) = \frac{1}{z_{i\to j}}\, n_i^{\beta\lambda-1}\, \exp\left\{-\frac{\beta}{2}(n_i^2-2n_i)\right\} \prod_{k\in\partial i\setminus j}\left(\int_0^\infty dn_k\, \eta_{k\to i}(n_k)\exp\left\{-\beta\alpha_{ik}n_i\, n_k\right\}\right). \tag{3.37}$$

So far we have considered the cavity marginals $\eta_{i\to j}(n_i)$, which are the probability distributions for the abundance of species i once we cut the edge with species j. Now, we introduce for convenience another cavity marginal, which we denote $\hat\eta_{i\to j}(n_j)$. This cavity marginal is the distribution of the abundance of species j once all its edges, except the one with i, are cut. This new cavity marginal is defined as:

$$\hat\eta_{i\to j}(n_j) \propto \int_0^\infty dn_i\, \eta_{i\to j}(n_i)\exp\left\{-\beta\alpha_{ij}n_i\, n_j\right\}. \tag{3.38}$$

In particular, inserting Eq. (3.37) into Eq. (3.38), we get:

$$\hat\eta_{i\to j}(n_j) \propto \int_0^\infty dn_i\, n_i^{\beta\lambda-1}\, \exp\left\{-\frac{\beta}{2}(n_i^2-2n_i)\right\}\exp\left\{-\beta\alpha_{ij}n_i\, n_j\right\} \prod_{k\in\partial i\setminus j}\hat\eta_{k\to i}(n_i), \tag{3.39}$$

which can be rewritten as

$$\hat\eta_{i\to j}(n_j) \propto \exp\left\{\frac{\beta}{2}(1-\alpha_{ij}n_j)^2\right\}\int_0^\infty dn_i\, n_i^{\beta\lambda-1}\, \exp\left\{-\frac{\beta}{2}(n_i-1+\alpha_{ij}n_j)^2\right\} \prod_{k\in\partial i\setminus j}\hat\eta_{k\to i}(n_i). \tag{3.40}$$

In order to analyze the integral part of Eq. (3.40), let us split it at some $\delta>0$:

$$I_{i\to j}(n_j) = \int_0^\delta dn_i\, n_i^{\beta\lambda-1}\, \exp\left\{-\frac{\beta}{2}(n_i-1+\alpha_{ij}n_j)^2\right\} \prod_{k\in\partial i\setminus j}\hat\eta_{k\to i}(n_i)+$$

$$+\int_{\delta}^{\infty} dn_i\, n_i^{\beta\lambda-1}\, \exp\Big\{-\frac{\beta}{2}(n_i-1+\alpha_{ij}n_j)^2\Big\} \prod_{k\in\partial i\setminus j}\hat{\eta}_{k\to i}(n_i). \tag{3.41}$$

For $\delta \ll 1$, the first contribution, i.e. the integral between 0 and δ, can be approximated as:

$$I^{\delta}_{i\to j}(n_j) = \exp\Big\{-\frac{\beta}{2}(1-\alpha_{ij}n_j)^2\Big\} \prod_{k\in\partial i\setminus j}\hat{\eta}_{k\to i}(0)\frac{\delta^{\beta\lambda}}{\beta\lambda}, \tag{3.42}$$

where we performed the integration of the factor $n_i^{\beta\lambda-1}$ over the interval $[0,\delta]$, while approximating the remaining part of the integrand by its value at $n_i = 0$.

Substituting Eqs. (3.41) and (3.42) into Eq. (3.40), we obtain the BP update rule for the cavity messages $\hat{\eta}_{i\to j}$:

$$\begin{aligned}\hat{\eta}_{i\to j}(n_j) = \frac{1}{\hat{z}_{i\to j}}\Big[&\frac{\delta^{\beta\lambda}}{\beta\lambda}\prod_{k\in\partial i\setminus j}\hat{\eta}_{k\to i}(0)+\\ &+\int_{\delta}^{\infty} dn_i\, n_i^{\beta\lambda-1}\exp\Big\{-\frac{\beta}{2}(n_i^2-2n_i)\Big\}\exp\Big\{-\beta\alpha_{ij}n_in_j\Big\}\prod_{k\in\partial i\setminus j}\hat{\eta}_{k\to i}(n_i)\Big].\end{aligned} \tag{3.43}$$

In Eq. (3.43) we have thus managed to compute the contribution in the surrounding of $n_i = 0$ considering continuous variables and performing an integral, see Eq. (3.42). However, for the remaining part of the domain of integration $[\delta,\infty]$, we need to compute the integral by discretizing.

Importantly, once the messages $\hat{\eta}_{i\to j}$ are determined, the cavity marginals $\eta_{i\to j}$ follow directly, as we can see by rewriting Eq. (3.37) as:

$$\eta_{i\to j}(n_i) \propto n_i^{\beta\lambda-1}\exp\Big\{-\frac{\beta}{2}(n_i^2-2n_i)\Big\}\prod_{k\in\partial i\setminus j}\hat{\eta}_{k\to i}(n_i). \tag{3.44}$$

In order to do an analogous analysis of Sec. 3.5, we can then analyze convergence directly at the level of $\hat{\eta}_{i\to j}$. In particular, for $\hat{\sigma} = 0$ and competitive interactions $\hat{\mu} > 0$, species abundances will never be larger than the carrying capacities, which in this case are all set to be 1. This means that the integral $\int_{\delta}^{\infty}$ in Eq. (3.43) can be restricted to the interval $[\delta, 1]$. Discretizing this interval we can compute the integral numerically.

In Fig. 3.17 we show the transition from the single-equilibrium to the multiple-equilibria phase in the $(\hat{\mu}, T)$ plane, for $\hat{\sigma} = 0$. As in Sec. 3.5.1, the transition to the multiple-equilibria phase is signalled by the non convergence of the iteration algorithm for which the update rule is given by Eq. (3.43).

The shape of the phase diagram of Fig. 3.17 is identical to the case with discrete variables. In both cases, the critical value of $\hat{\mu}$ at which the transition appears approaches $\hat{\mu} \simeq 0.353$ when $T \to 0$, consistently with the result of [28] (see Eq. (3.33)). The only difference between the discrete (Fig. 3.16) and the continuous case (Fig. 3.17) lies in the values of the temperature T. This follows from the value of the carrying

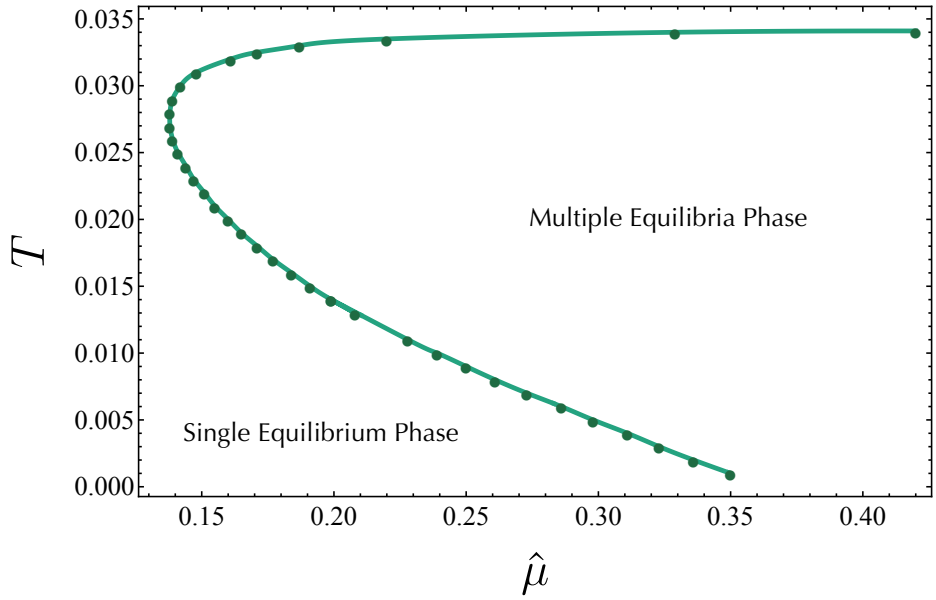

Figure 3.17: Transition line, at $\hat{\sigma} = 0$, between the single equilibrium and the multiple equilibria phase. We see that, as in Fig. 3.16, the transition line has a nose-like shape.

capacity K, which rescales the abundances. We can understand this by looking at the stochastic gLV equations. In fact, the gLV dynamics reads

$$\frac{dn_i(t)}{dt} = \frac{n_i(t)}{K}\Big[K - n_i(t) - \sum_{j\in\partial i} \alpha_{ij} n_j(t)\Big] + \sqrt{2Tn_i(t)}\xi_i(t), \tag{3.45}$$

with $\xi_i(t)$ Gaussian white noise. Rescaling abundances as $\tilde{n}_i = n_i/K$ yields

$$\frac{d\tilde{n}_i(t)}{dt} = \tilde{n}_i(t)\Big[1 - \tilde{n}_i(t) - \sum_{j\in\partial i} \alpha_{ij} \tilde{n}_j(t)\Big] + \sqrt{2\frac{T}{K}\tilde{n}_i(t)}\xi_i(t). \tag{3.46}$$

Thus, the effect of K is simply to rescale the temperature by a factor $1/K$. Accordingly, the "nose-tip" of Fig. 3.16 appears around $T = 8$, while in Fig. 3.17 it is located at $T = 0.028 = 8/280$.

3.7 Conclusions

In this chapter we have studied the equilibrium thermodynamics of the generalized Lotka-Volterra (gLV) model on a sparse random graph with quenched disordered interactions, with particular attention to the behaviour of species abundance distributions. By exploiting the cavity method and the Belief Propagation equations, we have studied the properties of the gLV model on a locally tree-like random graph with small and fixed connectivity k, varying three parameters: the average interaction strength $\hat{\mu}$, the width $\hat{\sigma}$ of the randomly distributed disordered couplings and the extent of thermal fluctuations, parametrized by the temperature T. Our first important observation has been

that, by increasing $\hat{\sigma}$ at finite and small $\hat{\mu}$ and finite temperature, the abundance marginal distributions become strongly non-Gaussian while staying in a single-equilibrium phase. In models with the same sort of random couplings but living on a dense network, there is by construction no room for non-Gaussian effects. We have thus argued that this result is purely due to the sparsity of the network, therefore providing a new perspective on the emergence of non-Gaussian and Gamma-like distributions in ecological systems. In particular, the finding of Gamma-like marginals is strongly consistent with the empirical observations from real ecosystems [30, 60]. We have observed non-Gaussian effects also in two-species joint probabilities, due to the emergence of non trivial correlations. In addition to the observation of these non-Gaussian effects in the single equilibrium phase, we have ascertained that by further increasing $\hat{\sigma}$ for the same choice of $\hat{\mu}$ and T the glass phase usually found in fully-connected networks is never met: on the Bethe lattice with fixed and small connectivity, at small values of $\hat{\mu}$, upon increasing $\hat{\sigma}$ the system has a transition from the single-equilibrium phase directly to the unbounded growth one.

After having studied the effect of varying $\hat{\sigma}$ at fixed $\hat{\mu}$, we focused on the model behaviour at zero disorder, $\hat{\sigma} = 0$, varying both the temperature T and the average strength $\hat{\mu}$ of the ordered interactions. Doing so, we have extended the zero temperature results of [71], showing that the formation of the topological glass phase pointed out in [71] at $T = 0$ for high enough interaction strength, $\hat{\mu} > \hat{\mu}_c$, is robust against the introduction of temperature. This phase emerges due to the occurrence of extinctions, resulting from high competition within the system (large $\hat{\mu}$). The higher is $\hat{\mu}$, the larger is the amount of extinctions, which eventually determines the formation of disconnected groups of surviving species and the *freezing* of the system into a configuration where a certain number of species have gone extinct, forming non-communicating islands of surviving ones. The pattern of surviving/disappearing species strongly depends on the initial conditions and the sequence of thermal fluctuations, so that there is a huge multiplicity of possible target stationary states: that is why it makes sense to talk about a glass phase, but of very different nature compared to the one induced by the disordered couplings in fully connected networks [26].

The main outcomes of our investigation have been the highlight of realistic non-Gaussian effects in the single-equilibrium phase and the uncovering of a non-trivial glass phase at finite temperature and in the absence of disorder, therefore questioning two of the main ingredients of the models so-far used to mimic the behaviour of large ecosystems: dense networks and large fluctuations in the disordered couplings. From a model-building perspective, we provided the evidence that the study of sparse networks offers a new avenue for understanding the behaviour of large ecosystems within a more realistic framework.

The transition to the *topological* multiple-equilibria phase observed in the sparse system requires deeper investigation. We previously argued that, for sufficiently large values of the competition strength $\hat{\mu}$, a glassy phase emerges. Up to this point, our analysis, based on the *replica-symmetric* (RS) formulation of the Belief Propagation algorithm, has only indicated that beyond the transition, the system is no longer in a single-equilibrium phase. In the language of replica theory, this means the system

leaves the RS phase and begins to exhibit features characteristic of *replica symmetry breaking* (RSB). But at this point other questions arise: which type of RSB phases arise? Can we characterize them? Is the transition towards a 1RSB phase or the system passes directly to a full replica symmetry breaking phase?

The next chapter will be devoted to investigating the nature of the RSB phase that emerges beyond the transition line in Fig. 3.14. In particular, the standard BP equations become inadequate for analyzing the multiple-equilibria regime, and we will introduce their replica-symmetry-breaking (RSB) extension.

4

Exploring the gLV multiple-equilibria phase

In Chapter 3, we investigated the generalized Lotka-Volterra model on a locally tree-like random graph with finite connectivity, using the Cavity method and Belief Propagation (BP) equations. Within the region of the phase diagram where BP converges, we were able to study the marginal distributions of each species of the ecosystem and to characterize the phenomenology of sparse ecological communities. However, as shown in Sections 3.5 and 3.6, BP does not always converge. Its failure to converge indicates that, in those regions of parameter space, the thermodynamics cannot be captured by a single equilibrium state. To explore these regimes, where multiple equilibria emerge, in this chapter we introduce the one-step replica symmetry breaking (1RSB) Cavity method.

As discussed in Secs. 3.5, 3.6, the lack of convergence of the Belief Propagation equations typically signals the formation of many distinct equilibrium states, each associated with its own Boltzmann weight. In such a scenario, the assumption of a single Gibbs measure, at the basis of the *replica-symmetric* (RS) BP algorithm, becomes inconsistent. In presence of multiple equilibria, two main situations are commonly encountered: a fragmentation of configuration space following a *one-step replica symmetry breaking* (1RSB) pattern, characterized by many disjoint and well distinguished basins, and a *full replica symmetry breaking* (FRSB) structure, where equilibria form a hierarchical landscape. While the Belief Propagation algorithm cannot be easily adapted to the latter case, a modified version exists for the 1RSB scenario.

In Chapter 3 (see Fig. 3.14), we identified the transition line separating the single-equilibrium phase, where BP converges to the unique fixed-point of the system, from the multiple-equilibria phase, where BP fails to converge and no longer provides a valid description. This transition arises in sparse ecosystems with homogeneous interactions ($\hat{\sigma} = 0$) under strong competition (large $\hat{\mu}$). In this chapter, we therefore introduce and apply the *one-step replica symmetry breaking* (1RSB) Cavity method to study the nature of this peculiar multiple-equilibria phase.

4.1 1RSB Cavity Method

In order to introduce the 1RSB cavity method, let us first recall that the Hamiltonian which governs the thermodynamics of the gLV model is of the form:

$$H(\mathbf{n}) = \sum_{i=1}^{N} h_i(n_i) + \sum_{(ij)\in E} h_{ij}(n_i, n_j), \tag{4.1}$$

where $\mathbf{n}$ denotes the vector of species abundances, $h_i(n_i)$ and $h_{ij}(n_i, n_j)$ are the field and interaction terms and E represents the set of edges in the interaction network. Defining

$$\psi_i(n_i) = \exp\left(-\beta h_i(n_i)\right) \qquad \psi_{ij}(n_i, n_j) = \exp\left(-\beta h_{ij}(n_i, n_j)\right), \tag{4.2}$$

the RS equations for the cavity marginals read:

$$\eta_{i\to j}(n_i) = \frac{\psi_i(n_i)}{z_{i\to j}} \prod_{k\in\partial i\setminus j} \left[\sum_{\{n_k\}} \eta_{k\to i}(n_k)\psi_{ik}(n_i, n_k)\right] = \mathcal{F}_{i\to j}[\eta_{k\to i}, k \in \partial i\setminus j], \tag{4.3}$$

where $z_{i\to j}$ is the normalization and $\mathcal{F}_{i\to j}$ denotes the function which implements the BP equations for the messages $\eta_{i\to j}$.

In the presence of multiple equilibria, each characterized by its own equilibrium measure, the RS BP equations (4.3) may either converge to different stationary points depending on initial conditions or wander around these fixed points failing to ever converge to one. The latter is the case of gLV sparse random networks in the topological multiple-equilibria phase described in Sec. 3.5. We therefore need to have an algorithm adapted to the situation.

The goal is actually to constrain the equations (4.3) to enter a single state and, more generally, to sample all possible states. To achieve this, we employ the "coupled-replicas" method introduced by Monasson [78]. The idea is to consider m copies (replicas) of the system and introduce a weak coupling among them, so that all replicas fall into the same state while remaining statistically independent within it. The coupling is then sent to zero, restoring the original, unperturbed system but with all replicas aligned in one particular state.

This procedure is analogous to selecting a magnetized phase in a ferromagnet: a small positive or negative external field biases the system toward one of the two magnetization states, and removing the field afterwards leaves the system in the chosen phase. However, unlike the ferromagnetic case where the two possible states are characterized by their magnetization (+ or −), here the "label" of each state is unknown. The infinitesimal coupling between replicas thus provides a way to align them within a single state without specifying in advance which one.

For such a replicated system, since the m replicas are independent inside a given state a, their combined contribution to the free-energy is simply m times the free-energy of a single replica. This means that the *replicated* partition function of the system reads as:

$$Z(m) = \sum_a e^{-m\beta F^a}, \tag{4.4}$$

where the summation is over all the states and F^a represents the free-energy of state a. The weight of each state, i.e. the probability that such state will be visited, is $w_a \propto e^{-m\beta F^a}$.

Within a single state a, BP equations (4.3) still hold, meaning that the equations for the cavity messages $\eta^a_{i\to j}$ inside that state read

$$\eta^a_{i\to j} = \mathcal{F}_{i\to j}[\eta^a_{k\to i}, k \in \partial i \backslash j], \tag{4.5}$$

with fixed-point solutions $\eta^a_{i\to j}$ that vary from state to state. Our goal is to study the distribution of such RS fixed points across the ensemble of states. This motivates the introduction of a new object,

$$Q_{i\to j}(\eta_{i\to j}),$$

the probability distribution of the cavity marginal distributions $\eta_{i\to j}$. In this sense, the 1RSB equations go one step beyond the RS ones: instead of describing distributions of species abundances ($\eta_{i\to j}$), they describe distributions of distributions of species abundances ($Q_{i\to j}(\eta_{i\to j})$). Higher levels of replica-symmetry-breaking follow the same logic, with the 2RSB algorithm involving distributions of distributions of distributions, and so on.

The 1RSB cavity equations for the distributions $Q_{i\to j}$, which are derived in detail in App. E, are found to be

$$Q_{i\to j}(\eta_{i\to j}) \propto \sum_{\{\eta_{k\to i}\},\, k\in\partial i\backslash j} z^m_{i\to j}(\{\eta_{k\to i}\})\delta\left(\eta_{i\to j} - \mathcal{F}_{i\to j}\left[\{\eta_{k\to i}\}\right]\right) \prod_{k\in\partial i\backslash j} Q_{k\to i}(\eta_{k\to i}). \tag{4.6}$$

In Eq. (4.6), the deltas $\delta\left(\eta_{i\to j} - \mathcal{F}_{i\to j}\left[\{\eta_{k\to i}\}\right]\right)$ impose that the cavity messages $\eta_{i\to j}$ are solutions of the RS BP equations (4.3) and the term $z^m_{i\to j}$ correctly weights each solution. Let us recall that $z_{i\to j}$ is the normalization of the cavity marginals, as it is shown in Eq. (4.3).

From Eq. (4.6), we see that the parameter m, introduced in Eq. (4.4) as the number of coupled replicas, now appears as an additional variable of the system. Although m is initially defined as an integer, an analytic continuation to real values is then performed. After this analytic continuation, one finds that values $m > 1$ are unphysical, restricting the parameter to the interval $m \in [0, 1]$. The details of this procedure are discussed in [43].

Eq. (4.6) tells us that, for each directed edge $i \to j$, we have a different distribution $Q_{i\to j}$ of cavity messages. In particular, $Q_{i\to j}$ is the probability distribution of the cavity message $\eta_{i\to j}$ over the pure states, where each state has weight $e^{-m\beta F^a}$, as shown in Eq. (4.4). This means that each $Q_{i\to j}$ can be also written as a sum of deltas over all the states:

$$Q_{i\to j}(\eta_{i\to j}) = \sum_a e^{-m\beta F^a}\delta\left[\eta_{i\to j} - \eta^a_{i\to j}\right]. \tag{4.7}$$

4.1.1 Homogeneous 1RSB equations

The aim of this chapter is to characterize the multiple-equilibria phase that arises in sparse gLV ecosystems under strong competition between species, in absence of interaction disorder ($\hat{\sigma} = 0$). We focus on systems whose interaction network is a Random Regular Graph, a locally tree-like random graph with fixed connectivity $k = 3$. In such a graph, in absence of interaction disorder, all species are equivalent: each species interacts with the same number of neighbours and through identical interaction coefficients. For such a homogeneous system, where no local disorder is present, it is then natural to assume that the distribution $Q_{i\to j}$ is independent of the specific edge $i \to j$. We can therefore replace $Q_{i\to j}(\eta_{i\to j})$ with a single distribution $Q_{\to}(\eta_{\to})$, which no longer carries explicit i, j indices. Rewriting Eq. (4.6) under this assumption, we obtain:

$$Q_{\to}(\eta_{\to}) \propto \sum_{\{\eta_{\to}^{a_1},\dots,\eta_{\to}^{a_{k-1}}\}} z_{\to}(\{\eta_{\to}^{a_1},\dots,\eta_{\to}^{a_{k-1}}\})^m \delta\left[\eta_{\to} - \mathcal{F}(\{\eta_{\to}^{a_1},\dots,\eta_{\to}^{a_{k-1}}\})\right] \prod_{i=1}^{k-1} Q_{\to}(\eta_{\to}^{a_i}), \tag{4.8}$$

where k denotes the graph connectivity and $z_{\to}$ is the normalization of the cavity marginals, which also does not depend on indices i, j. The term $z_{\to}^m$ is a reweighting factor, which assigns the proper statistical weight to each solution. Equation (4.8) thus describes the distribution of all cavity messages. Rewriting also Eq. (4.7) for the homogeneous case we get:

$$Q_{\to}(\eta_{\to}) = \sum_a w^a \delta\left[\eta_{\to} - \eta_{\to}^a\right], \tag{4.9}$$

meaning again that $Q_{\to}(\eta_{\to})$ is a sum of deltas over the states.

4.1.2 Iteration procedure for the homogeneous 1RSB cavity equations

Let us now introduce the iteration procedure for the self-consistent homogeneous 1RSB cavity equations (4.8). Following Eq. (4.9), the distribution $Q_{\to}(\eta_{\to})$ is defined as a sum of delta functions over M possible states of the cavity messages $\eta_{\to}$. A message labeled by a, i.e. $\eta_{\to}^a$, corresponds to the cavity message in state a, where a runs from 1 to M. During each iteration step, the set of messages $\eta_{\to}^a$ is updated, a corresponding weight w^a is assigned to each state and the cumulative distribution $C(w)$ of the weights is constructed. The introduction of the cumulative distribution ensures that cavity messages with very small weights are not automatically discarded, even if their contribution is small.

At the end of each iteration step the distribution $Q_{\to}(\eta_{\to})$ is completely defined by two sets of variables: the set $\{\eta_{\to}^a\}$ of cavity messages and the set $\{w^a\}$ of corresponding weights.

In practice, once the parameters of the gLV system ($\hat{\mu}$ and T, with $\hat{\sigma} = 0$), the number M of messages and the replica parameter $m \in [0, 1]$ are fixed, the iterative scheme proceeds as follows:

Algorithm 1

1: Initialize $Q_{\rightarrow}(\eta_{\rightarrow})$ with a random set of M messages $\eta^a_{\rightarrow}$ and random normalized weights w^a.

2: Construct the cumulative distribution $C(w)$ for the weights.

3: Extract one index $a \in \{1, ..., M\}$, corresponding to the state of the message to be updated.

4: Using the cumulative distribution $C(w)$, draw $k-1$ indices $\{a_1, ..., a_{k-1}\} \subset \{1, ..., M\}$, representing the states of the $k-1$ neighbouring cavity messages required for the update.

5: Update the message according to $\eta^a_{\rightarrow} = \mathcal{F}(\{\eta^{a_1}_{\rightarrow}, ..., \eta^{a_{k-1}}_{\rightarrow}\})$, following Eq. (4.3).

6: Compute the non-normalized reweighting factor $\tilde{w}^a = z^m(\{\eta^{a_1}_{\rightarrow}, ..., \eta^{a_{k-1}}_{\rightarrow}\})$ associated to $\eta^a_{\rightarrow}$.

7: Repeat steps 3-6 M times.

8: Normalize the weights as $w^a = \tilde{w}^a / \sum_{a=1}^{M} \tilde{w}^a$.

9: Update the weights cumulative distribution $C(w)$ accordingly and return to step 3.

After each iteration step, once the distribution of cavity messages $Q_{\rightarrow}(\eta_{\rightarrow})$ has been updated, we also compute the corresponding distribution $Q(\eta)$ of marginal distributions. Let us recall that the RS equations for the marginal distributions, introduced in Sec. 3.3, are given by

$$\eta_i(n_i) = \frac{\psi_i(n_i)}{z_i} \prod_{k \in \partial i} \left[\sum_{\{n_k\}} \eta_{k \rightarrow i}(n_k) \psi_{ik}(n_i, n_k) \right] = \mathcal{G}_i[\eta_{k \rightarrow i}, k \in \partial i], \tag{4.10}$$

where z_i is the normalization and $\mathcal{G}_i$ denotes the functional form of the BP update for the marginal distributions η_i.

In practice, to compute $Q(\eta)$, at each iteration step we proceed as follows. For M times, where M is the fixed number of possible states of the system, we sample k cavity messages $\{\eta^{a_1}_{\rightarrow}, ..., \eta^{a_k}_{\rightarrow}\}$ from the set of cavity messages in $Q_{\rightarrow}(\eta_{\rightarrow})$, using the cumulative distribution $C(w)$. This means that each message $\eta^{a^*}_{\rightarrow}$ is selected with probability equal to its weight w^{a^*}. The updated marginal is then computed as $\eta^b(n) = \mathcal{G}(\{\eta^{a_1}_{\rightarrow}, ..., \eta^{a_k}_{\rightarrow}\})$, where $b \in \{1, ..., M\}$ labels the M possible states of the marginals. After performing this procedure M times, we obtain M marginal distributions $\eta^{b_i}(n)$ corresponding to the M states b_i $(i = 1, ..., M)$. The distribution $Q(\eta)$ is then given by the average over all these marginals:

$$Q(\eta) = Q_\eta(n) = \frac{1}{M} \sum_{b_i, i \in \{1, ..., M\}} \eta^{b_i}(n). \tag{4.11}$$

At the end of each iteration step of Algorithm 1, we compute both the set of marginals $\{\eta^{b_i}(n)\}$ and their averaged distribution $Q_\eta(n)$. Convergence is reached when the distribution $Q_\eta(n)$ converges.

Up to this point, we have assumed a fixed value of the parameter m. However, m is a variable of the system and is not determined a priori. Consequently, in order to identify a stable 1RSB phase, the algorithm must be executed for different values of the variational parameter $m \in [0, 1]$. For each value of m, once convergence is reached, we compute the corresponding free energy (the explicit expression for its computation is provided in App. E). By plotting the free energy as a function of m, we can then determine its equilibrium value by selecting the minimum of the curve and identify the corresponding value $m = m^*$ at which this minimum occurs.

4.1.3 Absence of a 1RSB phase for sparse and ordered gLV systems

Let us now apply all the 1RSB machinery introduced in previous sections to study the *topological* multiple-equilibria phase that appears in sparse ecological networks as the competition parameter $\hat{\mu}$ increases, at fixed temperature T and in the absence of interaction disorder ($\hat{\sigma} = 0$). For clarity, we reproduce in Fig. 4.1 the transition line, that was already shown in previous chapter in Fig. 3.14, separating the single-equilibrium phase from the topological multiple-equilibria phase.

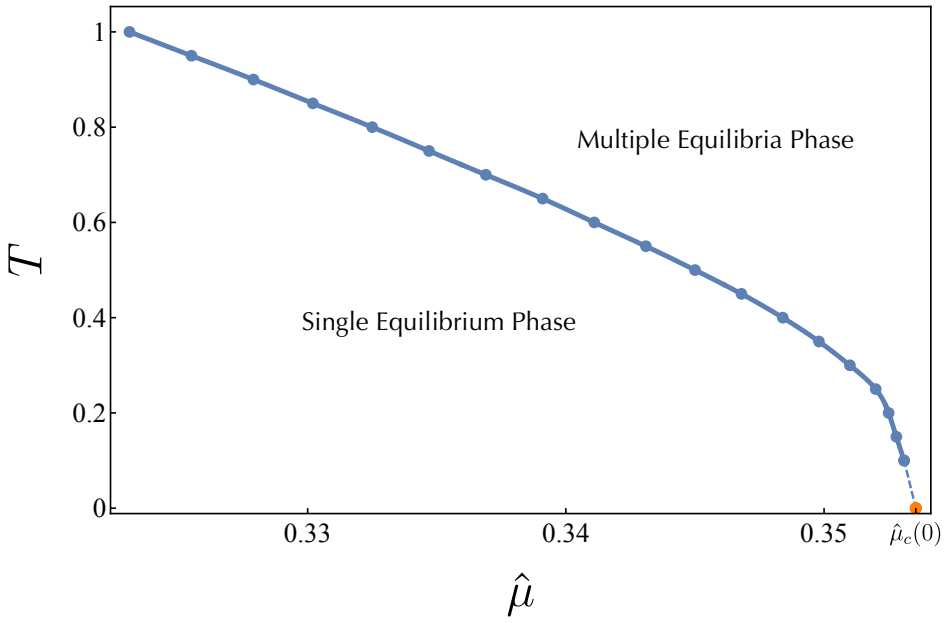

Figure 4.1: Transition line from single to multiple equilibria in the $(\hat{\mu}, T)$ space, for $\hat{\sigma} = 0$.

Let us focus on the transition in Fig. 4.1 at temperature $T = 1$. For $T = 1$ the critical value of the competition parameter $\hat{\mu}$ is found to be $\hat{\mu}_c = 0.3231$. This means that for $\hat{\mu} < 0.3231$ the system remains in a single-equilibrium phase, while for $\hat{\mu} \geq 0.3231$ it enters a multiple-equilibria regime. We performed the 1RSB iteration procedure of Sec. 4.1.2 for $T = 1$, considering two representative values of the competition parameter, $\hat{\mu} = 0.3230$ and $\hat{\mu} = 0.3231$. Figures 4.2 and 4.3 show the sets of marginal distributions $\eta^b(n)$, with $b \in \{1, ..., M\}$, obtained after 2000 iterations for two different values of the parameter m, $m = 1$ and $m = 0.1$. The number of sampled states was

set to $M = 1000$.

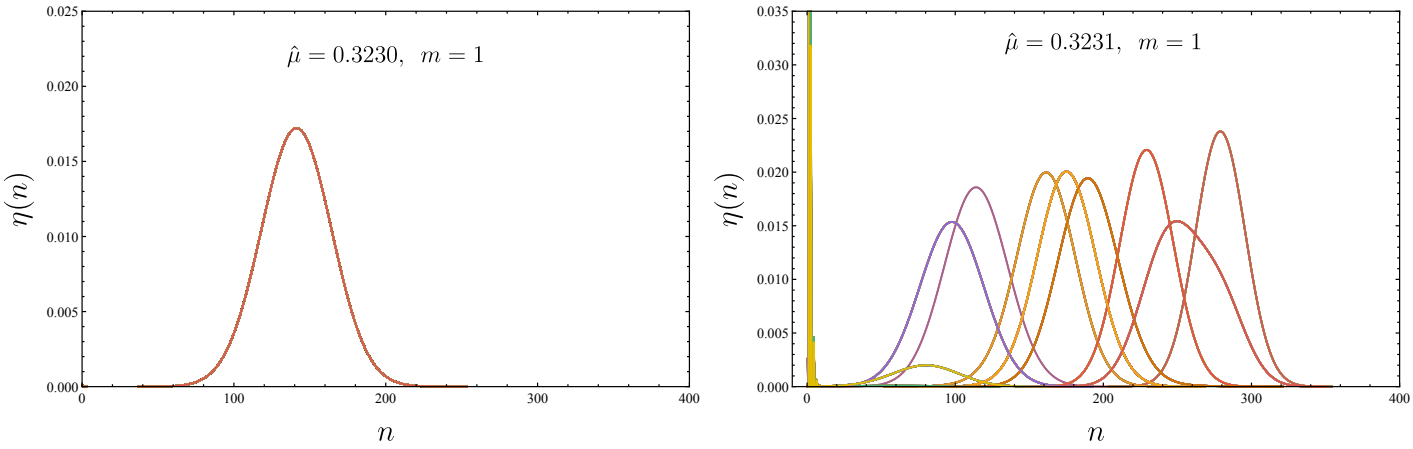

Figure 4.2: Set of $M = 1000$ marginal distributions $\eta(n)$ corresponding to the M possible states of the system, computed after 2000 iterations for $T = 1$ and $m = 1$. **Left**: for $\hat{\mu} = 0.3230$, all marginals collapse onto a single curve. **Right**: for $\hat{\mu} = 0.3231$, marginals differ across states, indicating the emergence of multiple equilibria.

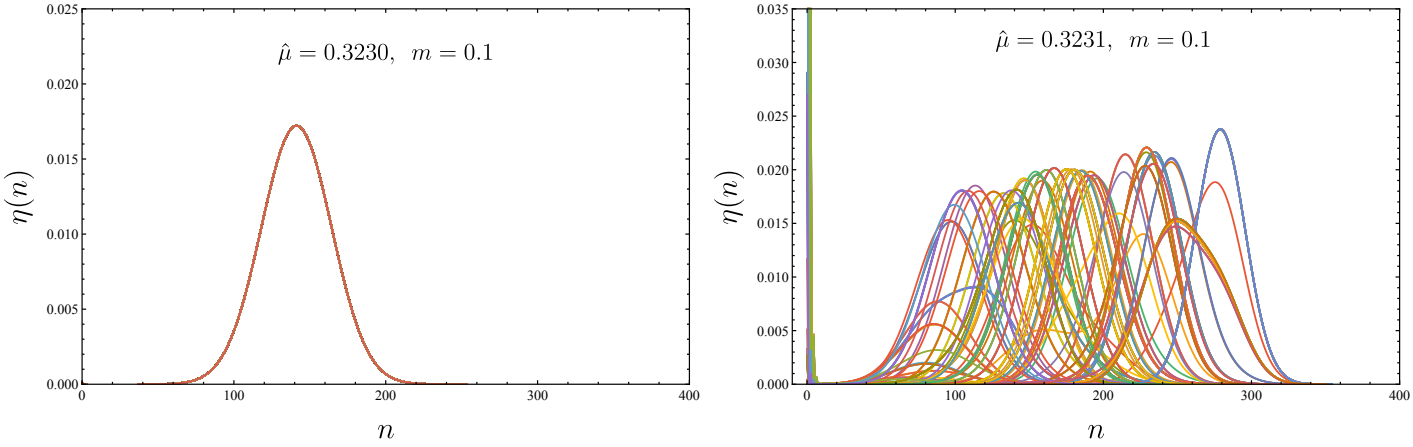

Figure 4.3: Same as Fig. 4.2, but for $m = 0.1$. Set of marginal distributions $\eta(n)$ for $T = 1$. **Left**: for $\hat{\mu} = 0.3230$, all marginals collapse onto a single curve. **Right**: for $\hat{\mu} = 0.3231$, marginals differ across states.

The difference between the two values of $\hat{\mu}$ is evident in both Figures 4.2 and 4.3. For $\hat{\mu} = 0.3230$ all marginal distributions corresponding to different states converge collapsing to the same one, confirming that the system is in a single-equilibrium phase. The situation changes at $\hat{\mu} = 0.3231$. Here, the marginals associated with different states are no longer identical: they display distinct shapes and ranges; some exhibit extinction peaks at $n = 0$, while others do not. This clearly confirms that, for $\hat{\mu} \geq \hat{\mu}_c$, the system is no longer in the single-equilibrium phase.

However, the distribution $Q_\eta(n)$, defined in Eq. (4.11), fails to converge for $\hat{\mu} = 0.3231$, for both $m = 1$ and $m = 0.1$. In practical terms, continuing the iteration

causes the distribution $Q_\eta(n)$ to fluctuate indefinitely without ever converging. This fact indicates that the 1RSB phase is not stable. We tested also intermediate values of m and larger values of $\hat{\mu}$, obtaining the same behaviour: the distribution $Q_\eta(n)$ never stabilizes. The analysis presented here focuses on $T = 1$, but equivalent results are obtained for other temperatures. This lack of convergence signals the absence of any stable 1RSB phase and suggests that the system instead enters a more complex replica-symmetry-breaking regime.

4.2 Conclusions

In this chapter, we have investigated the topological multiple-equilibria phase, previously introduced in Chapter 3, which appears in sparse ecological networks in absence of interaction disorder ($\hat{\sigma} = 0$) as the interaction strength $\hat{\mu}$ increases. An increase in $\hat{\mu}$ corresponds to stronger competition among species and, beyond a critical value $\hat{\mu}_c$, which is plotted in function of temperature T in Fig. 4.1, extinctions start to occur. In this regime, the system is fully ordered: the interaction network is a Random Regular Graph with fixed connectivity and, since $\hat{\sigma} = 0$, all the interactions are equivalent. Consequently, the particular set of extinct species depends only on the initial conditions: different initial conditions lead to different equilibria, each characterized by different sets of extinct and surviving species. For $\hat{\mu} \geq \hat{\mu}_c$, the system thus passes from a single-equilibrium regime to a multiple-equilibria one.

The multiple-equilibria phase found on the right of the transition line in Fig. 4.1 requires tools beyond those employed for the single-equilibrium regime in Chapter 3. For this reason, we introduced and applied the *one-step replica symmetry breaking* (1RSB) cavity method to analyze its structure and stability. Our results, based on the absence of convergence of the 1RSB algorithm, indicate that the multiple-equilibria phase is not of the 1RSB type. This suggests that additional levels of replica-symmetry breaking are necessary, possibly infinite ones implying the emergence of a *full replica symmetry breaking* (FRSB) phase.

The computational complexity of k-RSB cavity methods, however, becomes prohibitive for $k \geq 2$. Nevertheless, as we already discussed at the end of Chapter 3, the fact that both the replica-symmetric BP and the 1RSB equations never converge beyond the critical line provides evidence that the multiple-equilibria phase observed here could be a *full replica symmetry breaking* one, characterized by a fractal structure of equilibria with small basins of attraction, which causes the RS and the 1RSB cavity equations to wander indefinitely among such fixed points without ever converging [43].

5
Conclusions

The work presented in this thesis has explored the statistical mechanics of complex ecological systems through the lens of disordered systems theory. Using the generalized Lotka–Volterra (gLV) framework, we have studied the collective behaviour and stability of large ecological communities in which species interact through sparse networks. Our approach combined analytical and numerical techniques developed in the study of disordered systems, adapting them to the description of ecological communities.

In Chapter 1, we introduced the general context and motivation of this work. We discussed how ecosystems, composed of many interacting species and environmental factors, display a level of complexity that challenges traditional approaches in theoretical ecology. Within this perspective, statistical mechanics provides a natural language to describe macroscopic observables such as abundance distributions and stability, starting from microscopic interaction parameters. We reviewed the main models used in Theoretical Ecology and introduced the generalized Lotka–Volterra (gLV) model as the central theoretical framework. This model describes large ecological communities through pairwise interactions encoded in a random interaction matrix whose elements are Gaussian variables. Within this context, we presented the statistical-mechanical approach to ecosystems and reviewed existing results for fully connected systems where each species interacts with all others. In such dense networks, increasing interaction disorder leads to the emergence of multiple-equilibria phases, where the configuration landscape fragments into many distinct basins of attraction.

In Chapter 2, we developed the theoretical tools necessary to study fully connected disordered systems. We introduced key concepts from mean-field disordered systems theory, in particular the replica method, which provides a framework to analyze multiple-equilibria phases in physics. We then focused on the spherical 2-spin model perturbed by different types of nonlinearities, showing that only in the presence of disordered nonlinear couplings the model develops a replica-symmetry-breaking equilibrium configuration. We concluded the chapter by emphasizing that, while these

techniques are fundamental to the study of ecosystems with dense interaction networks, they must be extended or replaced to capture the properties of more realistic ecological systems characterized by finite connectivity.

In Chapter 3, we introduced the Cavity Method and the Belief Propagation (BP) algorithm, tools specifically suited to study the stationary states of the gLV model on sparse random networks. In this analysis, we considered species abundances as discrete variables, corresponding to integer counts of individuals, and interactions mediated by a locally tree-like random graph with fixed connectivity. Such a topology captures the local structure of ecological interactions more realistically than the fully connected case. We were able to characterize the marginal distributions of species abundances in the region of the phase diagram where BP converges. This study revealed that, depending on the parameters controlling stochastic fluctuations (T) and the mean ($\hat{\mu}$) and variance ($\hat{\sigma}$) of the interaction coefficients, the system exhibits distinct ecological regimes. In particular, we identified a region at large interaction disorder $\hat{\sigma}$ where the abundance distributions are Gamma-like. This is a result consistent with empirical data and different from the Gaussian-like distributions of fully connected models. Moreover, we showed that, in sparse ecological networks, no multiple-equilibria phase appears upon increasing $\hat{\sigma}$, in contrast to dense systems. However, for $\hat{\sigma} = 0$, increasing the interaction strength $\hat{\mu}$, which controls the level of competition between species, leads to non-convergence of the BP equations, signalling the emergence of a replica-symmetry-breaking phase. This multiple-equilibria regime, driven by competition rather than disorder, is qualitatively different from the one found in fully connected models. The transition between the single-equilibrium and multiple-equilibria phases was also analyzed through direct numerical simulations of the stochastic gLV dynamics, confirming our findings.

In Chapter 4, we extended our analysis to characterize this peculiar multiple-equilibria phase emerging in highly competitive sparse ecosystems. We introduced the one-step replica symmetry breaking (1RSB) Cavity Method, which is suitable for studying equilibrium phases composed of many well-separated basins of attraction. The absence of convergence of the 1RSB algorithm, together with that of the replica-symmetric BP equations, in the parameter region corresponding to the multiple-equilibria phase suggests that this regime is instead characterized by a more complex, possibly full replica symmetry breaking (FRSB) structure, with a hierarchical organization of equilibria and small basins of attraction.

The work presented in this thesis opens several directions for future research. From an ecological perspective, a natural extension would be to study interaction matrices that are not only sparse but also asymmetric, incorporating predator–prey interactions. In this case, equilibrium statistical mechanics, and consequently the Cavity Method, can no longer be applied, as the system is expected to evolve toward non-equilibrium stationary states. A first attempt in this direction, focusing on sparse and asymmetric interaction networks, has been carried out in [77]. Since the Cavity Method framework is particularly well suited to the study of sparse networks, it would also be interesting to extend it to spatially structured ecosystems, where nodes represent not different species but different spatial patches or locations. Such an approach would make it possible to capture the interplay between spatial heterogeneity and ecological stability.

From the perspective of empirical ecology, the theoretical results presented in this thesis provide a conceptual framework to interpret observed patterns in large ecological communities. In particular, the emergence of non-Gaussian, Gamma-like abundance distributions in sparse interaction networks offers a possible explanation for a widely reported empirical feature of ecological data, which is not captured by fully connected models. Moreover, the absence of disorder-driven multiple-equilibria phases in sparse systems suggests that high variability in interaction strengths alone may not be sufficient to induce instability in real ecosystems, where interactions are typically limited and structured. Instead our results indicate that strong competitive interactions, combined with network sparsity, can generate regimes characterized by many coexisting equilibria, which may be relevant for interpreting variability and non-unique outcomes observed in ecological data.

Altogether, the results presented here contribute to clarifying how sparsity shapes the stability and phenomenology of large ecological systems. However, real ecosystems are not only sparse: their interactions are asymmetric, adaptive, and embedded in complex, dynamic networks. Understanding these features represents the next challenge, one that this work has only begun to explore.

A

Details of computation for the 2-spin spherical model with an ordered non-linearity

In this Appendix we present in detail the computation of the quenched free energy of the spherical 2-spin model with ordered 4-body non-linearity and we show its explicit form in the replica-symmetric case.

A.1 Computation of the quenched free energy

The computation of the quenched free energy with the replica method requires the knowledge of the replicated partition function, which is given by

$$\mathcal{Z}_J^n = \int \mathcal{D}\sigma \exp\left[\beta \sum_{i<j}^{1,N} \sum_{a=1}^{n} J_{ij}\sigma_i^a\sigma_j^a\right] \exp\left[\frac{4!\beta\epsilon}{N^3} \sum_{i<j<k<l}^{1,N} \sum_{a=1}^{n} \sigma_i^a\sigma_j^a\sigma_k^a\sigma_l^a\right], \tag{A.1}$$

where we have introduced the shorthand notation

$$\mathcal{D}\sigma = \prod_{i=1}^{N}\prod_{a=1}^{n} \mathrm{d}\sigma_i^a \delta\left(N - \sum_{i=1}^{N}(\sigma_i^a)^2\right) \tag{A.2}$$

for the integration element over the replicated hypersphere.

By considering the J-dependent part of the expression (A.1), we perform the average

over the coupling distribution as follows

$$\overline{e^{\beta\sum_{i<j}^{N}\sum_{a}^{n}J_{ij}\sigma_i^a\sigma_j^a}} = \prod_{i<j}^{1,N}\int\frac{\mathrm{d}J_{ij}}{\sqrt{2\pi\sigma_J^2}}e^{-\frac{J_{ij}^2}{2\sigma_J^2}}e^{\beta\sum_{a}^{n}\sigma_i^a\sigma_j^aJ_{ij}}$$
$$= \exp\left[\frac{(\beta J)^2}{2N}\sum_{i<j}^{1,N}\sum_{ab}^{1,n}\sigma_i^a\sigma_i^b\sigma_j^a\sigma_j^b\right].$$

We now retain only the leading contributions to the free energy density in the large N limit, i.e. only terms of order $O(1)$. In the expression

$$2\sum_{i<j}^{1,N}\sum_{ab}^{1,n}\sigma_i^a\sigma_i^b\sigma_j^a\sigma_j^b = \sum_{ij}^{1,N}\sum_{ab}^{1,n}\sigma_i^a\sigma_i^b\sigma_j^a\sigma_j^b - \sum_{i=1}^{N}\sum_{ab}^{1,n}(\sigma_i^a)^2(\sigma_i^b)^2$$

we neglect the second sum, which leads to a $O(1/N)$ correction to the free energy density. Similarly in the following expression appearing in the ordered part of the partition function (A.1)

$$4!\sum_{i<j<k<l}^{1,N}\sum_{a=1}^{n}\sigma_i^a\sigma_j^a\sigma_k^a\sigma_l^a = \sum_{ijkl}^{1,N}\sum_{a=1}^{n}\sigma_i^a\sigma_j^a\sigma_k^a\sigma_l^a - 6\sum_{ijk}^{1,N}\sum_{a=1}^{n}(\cdots) - 4\sum_{ij}^{1,N}\sum_{a=1}^{n}(\cdots) - \sum_{i=1}^{N}\sum_{a=1}^{n}(\cdots)$$

we neglect all the sums apart from the first one, since they lead to corrections of order $O(1/N)$, $O(1/N^2)$ and $O(1/N^3)$ respectively to the free energy density. Therefore, the averaged expression of the replicated partition function at the leading order in N reads

$$\overline{\mathcal{Z}^n} = \int\mathcal{D}\sigma\exp\left[\frac{(\beta J)^2}{4N}\sum_{ab}^{n,1}\left(\sum_{i=1}^{N}\sigma_i^a\sigma_i^b\right)^2\right]\times\exp\left[\frac{\beta\epsilon}{N^3}\sum_{a=1}^{n}\left(\sum_{i=1}^{N}\sigma_i^a\right)^4\right]. \tag{A.3}$$

At this point of the calculation the partition function depends only on two global parameters: the magnetization vector

$$m_a = \frac{1}{N}\sum_{i=1}^{N}\sigma_i^a \tag{A.4}$$

and the overlap matrix with elements

$$q_{ab} = \frac{1}{N}\sum_{i=1}^{N}\sigma_i^a\sigma_i^b. \tag{A.5}$$

By definition the matrix $\mathbb{Q}$ is symmetric and positive semidefinite and its diagonal elements are $q_{aa} = 1$, due to the spherical constraint. These global parameters can be introduced in the computation of the partition function by exploiting the following identities

$$1 = \prod_{a=1}^{n}\int\mathrm{d}m_a\delta\left(Nm_a - \sum_{i}^{N}\sigma_i^a\right)$$

$$= \prod_{a=1}^{n} \int \mathrm{d}m_a \int_{-i\infty}^{+i\infty} \frac{N}{2\pi i} \mathrm{d}\rho_a e^{-\rho_a (N m_a - \sum_i^N \sigma_i^a)}$$
$$= \int \mathcal{D}m \int_{-i\infty}^{+i\infty} \mathcal{D}\rho \; e^{-N \sum_a^n \rho_a m_a + \sum_a^n \sum_i^N \rho_a \sigma_i^a} \tag{A.6}$$

with

$$\mathcal{D}m = \prod_{a=1}^{n} \mathrm{d}m_a \; , \qquad \mathcal{D}\rho = \prod_{a}^{n} \frac{N}{2\pi i} \mathrm{d}\rho_a \tag{A.7}$$

and

$$1 = \prod_{a<b}^{1,n} \int \mathrm{d}q_{ab} \delta \left(N q_{ab} - \sum_i^N \sigma_i^a \sigma_i^b \right)$$
$$= \prod_{a<b}^{1,n} \int \mathrm{d}q_{ab} \int_{-i\infty}^{+i\infty} \frac{N}{2\pi i} \mathrm{d}\lambda_{ab} e^{-\lambda_{ab} (N q_{ab} - \sum_i^N \sigma_i^a \sigma_i^b)}$$
$$= \int \mathcal{D}q \int_{-i\infty}^{+i\infty} \mathcal{D}\lambda \; e^{-\frac{N}{2} \sum_{a\neq b}^n \lambda_{ab} q_{ab} + \frac{1}{2} \sum_{a\neq b}^n \sum_i^N \sigma_i^a \lambda_{ab} \sigma_i^b}, \tag{A.8}$$

with

$$\mathcal{D}q = \prod_{a<b}^{1,n} \mathrm{d}q_{ab} \; , \qquad \mathcal{D}\lambda = \prod_{a<b}^{1,n} \frac{N}{2\pi i} \mathrm{d}\lambda_{ab}. \tag{A.9}$$

where the delta functions have been opened through a Laplace transformation and the conjugate variables of m_a and q_{ab} have been introduced, respectively ρ_a and λ_{ab}. Due to the fact that $\mathbb{Q}$ is symmetric, the matrix Λ (with elements λ_{ab}) is symmetric as well. Note that in the last line of Eq. (A.8) the sums in the exponential have been symmetrized .

By exploiting a similar relation, the spherical constraint, which until now has been hidden in the definition (A.2), can be opened as follows

$$\prod_{a=1}^{n} \delta \left(N - \sum_{i=1}^{N} (\sigma_i^a)^2 \right) = \prod_{a}^{n} \int_{-i\infty}^{+i\infty} \frac{N}{4\pi i} \mathrm{d}\lambda_{aa} e^{-\frac{\lambda_{aa}}{2} (N - \sum_i^N (\sigma_i^a)^2)}$$
$$= \int_{-i\infty}^{+i\infty} \prod_{a}^{n} \frac{N}{4\pi i} \mathrm{d}\lambda_{aa} e^{-\frac{N}{2} \sum_a^n \lambda_{aa} q_{aa} + \frac{1}{2} \sum_a^n \sum_i^N \lambda_{aa} (\sigma_i^a)^2}. \tag{A.10}$$

Note that the previous expression for the spherical constraint perfectly matches with Eq. (A.8). Hence, by neglecting constant prefactors and subleading contributions of order $O(\ln N/N)$, the expression of the replicated partition function reads

$$\overline{\mathcal{Z}^n} = \int \mathcal{D}\sigma \int \mathcal{D}q \int_{-i\infty}^{+i\infty} \mathcal{D}\lambda \int \mathcal{D}m \int_{-i\infty}^{+i\infty} \mathcal{D}\rho \; \times$$
$$\times \exp \left[\frac{(\beta J)^2 N}{4} \sum_{ab}^{1,n} q_{ab}^2 + \beta \epsilon N \sum_{a=1}^{n} m_a^4 - \frac{N}{2} \sum_{ab}^{1,n} \lambda_{ab} q_{ab} \right.$$
$$\left. + \frac{1}{2} \sum_{ab}^{1,n} \sum_{i=1}^{N} \sigma_i^a \lambda_{ab} \sigma_i^b - N \sum_{a=1}^{n} \rho_a m_a + \sum_{i=1}^{N} \sum_{a=1}^{n} \rho_a \sigma_i^a \right], \tag{A.11}$$

where now

$$\mathcal{D}\sigma = \prod_{i=1}^{N}\prod_{a=1}^{n} \mathrm{d}\sigma_i^a \qquad \mathcal{D}\lambda = \prod_{a\leq b}^{1,n} \frac{N}{2\pi i}\mathrm{d}\lambda_{ab} \tag{A.12}$$

We notice that the spin dependent part of the action can be factorized with respect to the site indices, yielding the following expression

$$\exp\left[N \ln \int \prod_{a}^{n} d\sigma^a e^{\frac{1}{2}\sum_{ab}^{n}\sigma^a\lambda_{ab}\sigma^b + \sum_a^n \rho_a\sigma^a}\right].$$

The Gaussian integral in the spin variables can be easily performed

$$\int \prod_{a}^{n} \mathrm{d}\sigma^a e^{\frac{1}{2}\sum_{ab}^{n}\sigma^a\lambda_{ab}\sigma^b + \sum_a^n \rho_a\sigma^a} = \sqrt{\frac{(2\pi)^n}{\det(-\Lambda)}} \cdot e^{-\frac{1}{2}\sum_{ab}^{n}\rho_a(\Lambda^{-1})_{ab}\rho_b}.$$

Note that the previous integral is well defined only after shifting the integration of the Λ elements in order for them to have a negative real part. This shift does not affect significantly the partition function. Eventually, by neglecting a constant, the partition function can be written as

$$\overline{\mathcal{Z}^n} = \int \mathcal{D}q \int_{-i\infty}^{+i\infty} \mathcal{D}\lambda \int \mathcal{D}m \int_{-i\infty}^{+i\infty} \mathcal{D}\rho \; e^{-NG[\mathbb{Q},\Lambda,\mathbf{m},\boldsymbol{\rho}]}, \tag{A.13}$$

where the effective action G has been defined as

$$\begin{aligned} G[\mathbb{Q},\Lambda,\mathbf{m},\boldsymbol{\rho}] = -\frac{(\beta J)^2}{4}\sum_{ab}^{1,n} q_{ab}^2 - \beta\epsilon\sum_{a=1}^{n} m_a^4 + \frac{1}{2}\sum_{ab}^{1,n}\lambda_{ab}q_{ab} + \sum_{a=1}^{n}\rho_a m_a + \\ + \frac{1}{2}\sum_{ab}^{1,n}\rho_a(\Lambda^{-1})_{ab}\rho_b + \frac{1}{2}\ln\det(-\Lambda). \end{aligned} \tag{A.14}$$

A.1.1 The Λ integration

The integrals in Eq. (A.13) can be performed through the saddle point method, which is why we only retained terms of order $O(1)$ in the large N limit. The Λ-dependent part of eq. (A.14) is

$$G_1[\Lambda] = \frac{1}{2}\sum_{ab}^{1,n}\lambda_{ab}q_{ab} + \frac{1}{2}\sum_{ab}^{1,n}\rho_a(\Lambda^{-1})_{ab}\rho_b + \frac{1}{2}\ln\det(-\Lambda). \tag{A.15}$$

The following relation holds:

$$\begin{aligned} \ln\det\left(-\Lambda - \boldsymbol{\rho}\otimes\boldsymbol{\rho}^T\right) &= \ln\left[(1+\boldsymbol{\rho}^T\Lambda^{-1}\boldsymbol{\rho})\det(-\Lambda)\right] \\ &= \ln\left(1+\boldsymbol{\rho}^T\Lambda^{-1}\boldsymbol{\rho}\right) + \ln\det(-\Lambda) \\ &= \ln\det(-\Lambda) + \boldsymbol{\rho}^T\Lambda^{-1}\boldsymbol{\rho} - \frac{1}{2}(\boldsymbol{\rho}^T\Lambda^{-1}\boldsymbol{\rho})^2 + O(n^3), \end{aligned} \tag{A.16}$$

where we have used a general matrix relation for the determinant of the sum of a matrix $\boldsymbol{A}$ and the external product of two vectors u and v, i.e.

$$\det\Big(\boldsymbol{A} + uv^T\Big) = (1 + v^T \boldsymbol{A}^{-1} u) \det \boldsymbol{A},$$

and then we have expanded the logarithm. We also recall that $\boldsymbol{\rho} \otimes \boldsymbol{\rho}^T$ denotes a matrix with elements $(\boldsymbol{\rho} \otimes \boldsymbol{\rho}^T)_{ab} = \rho_a \rho_b$. Thus, from the relation (A.16) one gets

$$\ln\det(-\Lambda) + \boldsymbol{\rho}^T \Lambda^{-1} \boldsymbol{\rho} = \ln\det\Big(-\Lambda - \boldsymbol{\rho} \otimes \boldsymbol{\rho}^T\Big) + \frac{1}{2}(\boldsymbol{\rho}^T \Lambda^{-1} \boldsymbol{\rho})^2 + O(n^3).$$

By using the previous relation, eq. (A.15) can be written as

$$\begin{aligned} G_1[\Lambda] = \frac{1}{2}\sum_{ab}^{1,n} \lambda_{ab} q_{ab} + \frac{1}{2}\ln\det\Big(-\Lambda - \boldsymbol{\rho} \otimes \boldsymbol{\rho}^T\Big) \\ + \frac{1}{4}\left(\sum_{ab}^{1,n} \rho_a (\Lambda^{-1})_{ab} \rho_b\right)^2 + O(n^3). \end{aligned} \tag{A.17}$$

However, the relevant contribution to the free energy is of order $O(n)$, due to Eq. (2.27). Thus the only part of $G_1[\Lambda]$ that we have to consider is

$$G_1[\Lambda] = \frac{1}{2}\sum_{ab}^{1,n} \lambda_{ab} q_{ab} + \frac{1}{2}\ln\det\Big(-\Lambda - \boldsymbol{\rho} \otimes \boldsymbol{\rho}^T\Big). \tag{A.18}$$

The stationary point of $G_1[\Lambda]$ is, thus, given by

$$\lambda^*_{ab} = -\rho_a \rho_b - (\mathbb{Q}^{-1})_{ab} \tag{A.19}$$

leading to

$$G_1[\Lambda^*] = -\frac{1}{2}\sum_{ab}^{1,n} \rho_a q_{ab} \rho_b - \frac{1}{2}\ln\det\mathbb{Q}, \tag{A.20}$$

where a constant has been neglected. Therefore, the complete effective action (A.14) reads as

$$\begin{aligned} G[\mathbb{Q}, \mathbf{m}, \boldsymbol{\rho}] = -\frac{(\beta J)^2}{4}\sum_{ab}^{1,n} q_{ab}^2 - \beta\epsilon \sum_{a=1}^{n} m_a^4 + \sum_{a=1}^{n} \rho_a m_a \\ - \frac{1}{2}\sum_{ab}^{1,n} \rho_a q_{ab} \rho_b - \frac{1}{2}\ln\det\mathbb{Q}. \end{aligned} \tag{A.21}$$

A.1.2 The $\boldsymbol{\rho}$ integration

Let us now perform the integration over $\boldsymbol{\rho}$. For this purpose the relevant part of the effective action is

$$G_2[\boldsymbol{\rho}] = -\frac{1}{2}\sum_{ab}^{1,n} \rho_a q_{ab} \rho_b + \sum_{a=1}^{n} \rho_a m_a. \tag{A.22}$$

The stationary point of $G_2[\boldsymbol{\rho}]$ is given by the relation

$$\rho_a^* = \sum_{b=1}^{n} (\mathbb{Q}^{-1})_{ab} m_b, \tag{A.23}$$

that leads to

$$G_2[\boldsymbol{\rho}^*] = \frac{1}{2} \sum_{ab}^{1,n} m_a (\mathbb{Q}^{-1})_{ab} m_b. \tag{A.24}$$

We have finally arrived to an expression of the effective action only in terms of the magnetization vector and the overlap matrix:

$$\begin{aligned} G[\mathbb{Q}, \mathbf{m}] = &-\frac{(\beta J)^2}{4} \sum_{ab}^{1,n} q_{ab}^2 - \beta\epsilon \sum_{a=1}^{n} m_a^4 \\ &- \frac{1}{2} \ln\det \mathbb{Q} + \frac{1}{2} \sum_{ab}^{1,n} m_a (\mathbb{Q}^{-1})_{ab} m_b. \end{aligned} \tag{A.25}$$

We can now use the same trick used before to retain only the relevant terms for the saddle point. By exploiting the following relation

$$\begin{aligned} \ln\det\Big(\mathbb{Q} - \mathbf{m} \otimes \mathbf{m}^T\Big) = &\ln\det \mathbb{Q} - \mathbf{m}^T \mathbb{Q}^{-1} \mathbf{m} \\ &- \frac{1}{2} (\mathbf{m}^T \mathbb{Q}^{-1} \mathbf{m})^2 + O(n^3), \end{aligned} \tag{A.26}$$

one gets

$$\begin{aligned} G[\mathbb{Q}, \mathbf{m}] = &-\frac{(\beta J)^2}{4} \sum_{ab}^{n} q_{ab}^2 - \beta\epsilon \sum_{a}^{n} m_a^4 - \frac{1}{2} \ln\det\Big(\mathbb{Q} - \mathbf{m} \otimes \mathbf{m}^T\Big) \\ &- \frac{1}{4} \left(\sum_{ab}^{n} m_a (\mathbb{Q}^{-1})_{ab} m_b \right)^2 + O(n^3). \end{aligned} \tag{A.27}$$

However the only relevant contribution to the saddle point is given by terms of order $O(n)$, so that we have

$$G[\mathbb{Q}, \mathbf{m}] = -\frac{(\beta J)^2}{4} \sum_{ab}^{n} q_{ab}^2 - \beta\epsilon \sum_{a}^{n} m_a^4 - \frac{1}{2} \ln\det\Big(\mathbb{Q} - \mathbf{m} \otimes \mathbf{m}^T\Big). \tag{A.28}$$

A.2 Replica symmetric free energy

In order to solve the saddle point problem given by Eqs. (2.28), (2.29) we take the simplest possible ansatz on the structure of the matrix $\mathbb{Q}$, i.e. the Replica Symmetric (RS) ansatz. We assume the overlap to be parametrized by only one variable

$$q_{ab} = (1 - q_0)\delta_{ab} + q_0 I_{ab}, \tag{A.29}$$

since the diagonal elements are $q_{aa} = 1$ due to the spherical constraint. In the previous expression I is a matrix whose elements are all ones. Moreover, we assume the magnetization vector to have all its components $m_a = m$. The energetic part of the action depending on the overlap can be then written as

$$\sum_{ab}^{n} q_{ab}^2 = n + n(n-1)q_0^2 = n(1-q_0^2) + O(n^2)$$

and the entropic term as

$$\begin{aligned}
\ln\det\Big(\mathbb{Q} - \mathbf{m}\otimes\mathbf{m}^T\Big) &= \ln\Big[(1-q_0)^{n-1}(1-q_0+n(q_0-m^2))\Big] \\
&= \ln\left[(1-q_0)^n\left(1+n\frac{q_0-m^2}{1-q_0}\right)\right] \\
&= n\ln(1-q_0) + \ln\left(1+n\frac{q_0-m^2}{1-q_0}\right) \\
&= n\ln(1-q_0) + n\frac{q_0-m^2}{1-q_0} + O(n^2),
\end{aligned} \tag{A.30}$$

since the RS matrix $\mathbb{Q} - \mathbf{m}\otimes\mathbf{m}^T$ has only two kind of eigenvalues $\lambda_1 = 1-q_0$, with degeneracy $n-1$, and $\lambda_2 = 1-q_0+n(q_0-m^2)$. Hence, in the $n\to 0$ limit the RS effective action reads as

$$\lim_{n\to 0}\frac{1}{n}G[q_0,m] = -\frac{(\beta J)^2}{4}(1-q_0^2) - \beta\epsilon m^4 - \frac{1}{2}\ln(1-q_0) - \frac{1}{2}\frac{q_0-m^2}{1-q_0}. \tag{A.31}$$

The RS free energy is then given by

$$f_{\mathrm{RS}} = \frac{1}{\beta}\lim_{n\to 0}\frac{1}{n}G[q_0,m]. \tag{A.32}$$

B

The 2-spin spherical model with an ordered plus a disordered non-linearity

Having assessed in Secs. 2.3.1 and 2.3.2 the effect of adding respectively purely ordered and purely disordered non-linear terms to the $p = 2$ spherical model, we study here the effect of non-linear terms with both competing disordered and ferromagnetic interactions. This can be done by extracting the couplings of the 4-body interaction term from a non-zero average probability distribution, namely we consider

$$p(J_{ijkl}) = \frac{1}{\sqrt{2\pi\sigma_4^2}} e^{-\frac{(J_{ijkl}-J_0)^2}{2\sigma_4^2}}, \tag{B.1}$$

with

$$\sigma_4^2 = \frac{\epsilon^2\ 4!}{2N^3} \qquad J_0 = \frac{\epsilon^2\ 4!}{N^3}. \tag{B.2}$$

The variable ϵ parametrizes both the mean and the variance of the 4-body couplings distribution. This choice allows to have just one free parameter for the whole Gaussian distribution. With a similar computation with respect to the previous cases we find the effective action

$$\begin{aligned} G[\mathbb{Q}, \mathbf{m}] = &- \frac{(\beta J)^2}{2} \sum_{a<b}^{1,n} q_{ab}^2 - \frac{(\beta\epsilon)^2}{2} \sum_{a<b}^{1,n} q_{ab}^4 \\ &- \beta\epsilon^2 \sum_{a=1}^{n} m_a^4 - \frac{1}{2} \ln\det\left(\mathbb{Q} - \mathbf{m}\otimes\mathbf{m}^T\right). \end{aligned} \tag{B.3}$$

B.1 RS solutions

In the same way of previous sections we begin by considering a RS ansatz for the matrix $\mathbb{Q}$. This leads to a RS action reading as:

$$\lim_{n\to 0}\frac{1}{n}G[q_0,m] = -\frac{(\beta J_2)^2}{4}(1-q_0^2) - \frac{(\beta\epsilon)^2}{4}(1-q_0^4) - \beta\epsilon^2 m^4 - \frac{1}{2}\log(1-q_0) - \frac{1}{2}\frac{q_0-m^2}{1-q_0}. \tag{B.4}$$

The RS saddle-point equations read as:

$$\begin{aligned}\frac{\partial G}{\partial q_0} &= (\beta J_2)^2 q_0 + 2\beta^2\epsilon^2 q_0^3 - \frac{q_0-m^2}{(1-q_0)^2} = 0\\ \frac{\partial G}{\partial m} &= -4\beta\epsilon^2 m^3 + \frac{m}{1-q_0} = 0\end{aligned} \tag{B.5}$$

There are three possible solutions for Eq. (B.5):

1. $m = 0$;
   $q_0 = 0$.

2. $m = 0$;
   $(\beta J_2)^2 + 2\beta^2\epsilon^2 q_0^2 = 1/(1-q_0)^2$.

3. $m = 1/\sqrt{4\beta\epsilon^2(1-q_0)}$;
   $q_0(1-q_0)[1-(1-q_0)^2((\beta J_2)^2 + 2\beta^2\epsilon^2 q_0^2)] = 1/(4\beta\epsilon^2)$.

In order to study the stability of the solution we have computed the replicon, which for the present case reads as:

$$\lambda_R = \frac{1}{(1-q_0)^2} - (\beta J_2)^2 - 6\beta^2\epsilon^2 q_0^2. \tag{B.6}$$

The paramagnetic solution (PM, n.1) with $m = 0$ and $q_0 = 0$ is unstable everywhere below the critical line $T_c = J_2$, since $\lambda_R > 0$. Therefore, below this line, which is the horizontal orange one in the phase diagram of Fig. B.1, we can only have phases with either finite magnetization or replica symmetry breaking or both. The trivial spin glass solution (n.2), is never stable for this model because one always has $\lambda_R < 0$, therefore it is always excluded from the phase diagram. For what concerns the ferromagnetic phase (FM, n.3) in the list above, the numerical solution of the corresponding saddle-point equations shows that in the range of parameters (ϵ, T) where such a solution exists and is nontrivial, it is stable, i.e. $\lambda_R > 0$.

Regarding the PM-FM transition, since both solutions are always stable, it is only by comparing their free energies that we can determine the transition line. The coordinates

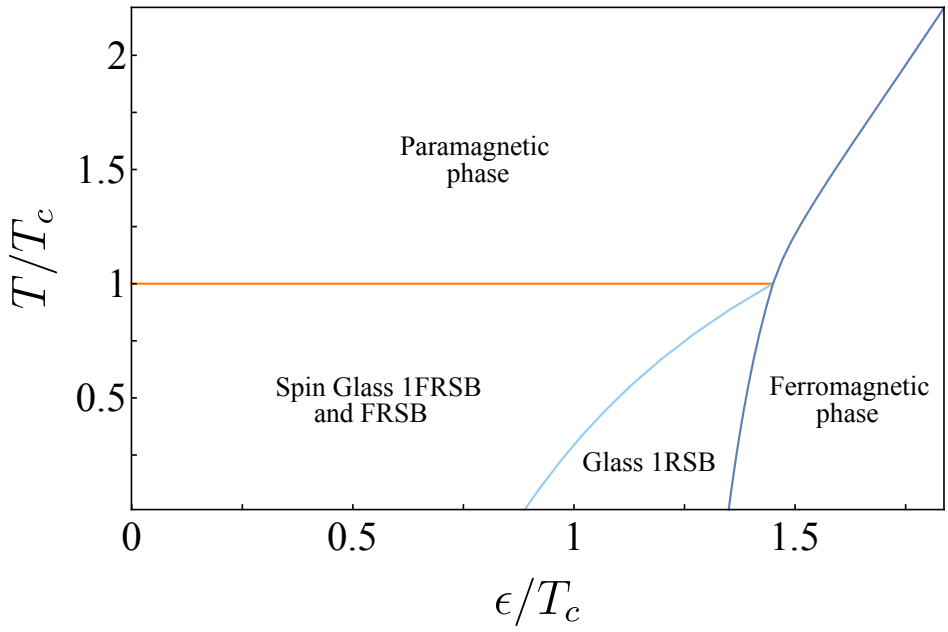

Figure B.1: Phase diagram in the (ϵ, T) plane for the $p = 2$ spherical spin glass plus four-body interactions, with competing disordered and ferromagnetic terms, where the variance of the disorder and the strength of ferromagnetic couplings are both controlled by ϵ, see Eq. (B.2) in the text.

of this first-order transition line in the phase diagram in Fig. B.1 are obtained by the implicit equation:

$$f^{\text{para}}(\epsilon_c^{\text{ferro}}, T_c^{\text{ferro}}) = f^{\text{ferro}}(\epsilon_c^{\text{ferro}}, T_c^{\text{ferro}}), \tag{B.7}$$

where:

$$\begin{aligned} \text{sol n.1} \quad &\Longrightarrow \quad f^{\text{para}}(\epsilon, T) = \frac{1}{\beta} G(q_0 = 0, m = 0) \\ &\qquad\quad = -\frac{\beta}{4}(J_2^2 + \epsilon^2) \\ \text{sol n.3} \quad &\Longrightarrow \quad f^{\text{ferro}} = G(q_0 > 0, m > 0). \end{aligned} \tag{B.8}$$

We will comment the remaining part of the phase diagram, namely what happens at temperatures lower than the critical one for the stability of the paramagnetic phase, $T < T_c$, in the following section after having introduced the 1RSB ansatz for the replica matrix.

B.2 1RSB ansatz

In order to study the phase diagram of the model for $T < T_c$ we need to consider the possibility of replica symmetry breaking. By recalling the definition

$$\chi_p(q_0, q_1, x) = 1 - q_1^p + x(q_1^p - q_0^p) \tag{B.9}$$

and by assuming the 1RSB structure for the matrix $\mathbb{Q}$ we obtain the following form for the effective action in the limit $n \to 0$:

$$\lim_{n\to 0} \frac{1}{n} G[q_0, q_1, x, m] = -\frac{(\beta J_2)^2}{4}\chi_2(q_0, q_1, x) - \frac{\beta^2\epsilon^2}{4}\chi_4(q_0, q_1, x) - \beta\epsilon^2 m^4 - \frac{x-1}{2x}\ln(1-q_1) - \frac{\ln(\chi_1)}{2x} - \frac{q_0 - m^2}{2\chi_1(q_0, q_1, x)},$$

from which the four 1RSB saddle point equations read as:

$$\begin{aligned}
\frac{\partial G}{\partial q_0} &= \frac{(\beta J_2)^2}{2} x q_0 + \beta^2\epsilon^2 x q_0^3 - \frac{1}{2}\frac{(q_0 - m^2)x}{\chi_1^2(q_0, q_1, x)} = 0 \\
\frac{\partial G}{\partial q_1} &= (x-1)\left[-\frac{(\beta J_2)^2}{2} q_1 - \beta^2\epsilon^2 q_1^3 + \frac{1}{2x(1-q_1)} - \frac{1}{2x\chi_1(q_0, q_1, x)} + \frac{q_0 - m^2}{2\chi_1^2(q_0, q_1, x)}\right] = 0 \\
\frac{\partial G}{\partial x} &= -\frac{(\beta J_2)^2}{4}(q_1^2 - q_0^2) - \frac{\beta^2\epsilon^2}{4}(q_1^4 - q_0^4) - \frac{1}{2}\frac{q_1 - q_0}{x\chi_1(q_0, q_1, x)} + \frac{1}{2x^2}\ln\left(1 + \frac{x(q_1 - q_0)}{1 - q_1}\right) + \\
&\quad + \frac{(q_0 - m^2)(q_1 - q_0)}{2\chi_1^2(q_0, q_1, x)} = 0 \\
\frac{\partial G}{\partial m} &= m\left(-4\beta\epsilon^2 m^2 + \frac{1}{\chi_1(q_0, q_1, x)}\right) = 0.
\end{aligned} \tag{B.10}$$

The expression of the two replicon eigenvalues which are relevant in the limit $n \to 0$, which can be obtained following the lines of [52], reads as

$$\Lambda^{(1)}[q_1] = -(\beta J_2)^2 - 6\beta^2\epsilon^2 q_1^2 + \frac{1}{(1-q_1)^2}, \tag{B.11}$$

$$\Lambda^{(2)}[q_0, q_1, x] = -(\beta J_2)^2 - 6\beta^2\epsilon^2 q_0^2 + \frac{1}{(1 - q_1 + x(q_1 - q_0))^2}. \tag{B.12}$$

As already mentioned, for this model the most relevant between the two is $\Lambda^{(2)}$, signalling a pattern of instability of the 1RSB ansatz characterized by the merging of 1RSB clusters.

B.2.1 Phase Diagram

Let us here summarize how all the transition lines for $T < T_c$ in Fig. B.1 have been obtained. We find, for ϵ large enough, a stable 1RSB phase with zero magnetization, $m = 0$. The transition line which separates this 1RSB phase, which we refer to simply as *glass* phase, from the spin-glass phase (with full-RSB) at smaller value of ϵ has been determined by simply tracking the value of the control parameters where the 1RSB phase loses stability. That is, the intrinsic equation of this line is determined as

$$\Lambda^{(2)}[q_0(\epsilon_c^{sg}, T_c^{sg}), q_1(\epsilon_c^{sg}, T_c^{sg}), x(\epsilon_c^{sg}, T_c^{sg})] = 0. \tag{B.13}$$

Inside the region of the phase diagram delimited by the lines $T = T_c$ above and $(\epsilon_c^{sg}, T_c^{sg})$ to the right, both the RS and the 1RSB solutions are unstable. This is therefore

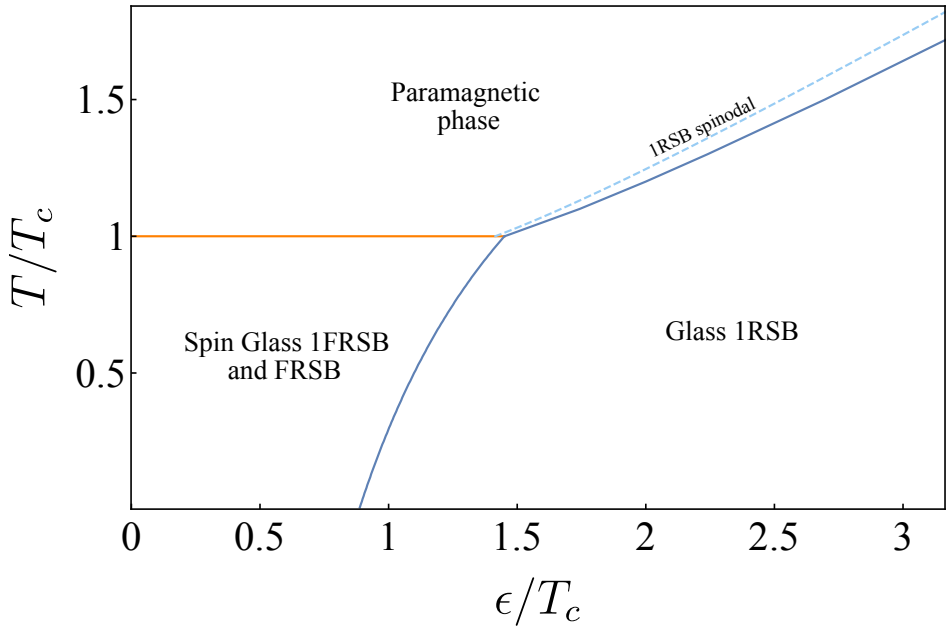

Figure B.2: Phase diagram in the (ϵ, T) plane for the $p = 2$ spherical spin glass plus four-body interactions: Gaussian distribution of non-linear couplings with standard deviation $\sigma_4 \sim \epsilon$ and mean $J_0 \sim \epsilon$ ($R = \text{const}$).

necessarily a region where further breakings of the symmetry between replica occurs. As already mentioned, we know from the results of [51] that different kinds of full-RSB occur here. Both the transitions occurring along these lines are continuous.

Finally, the transition line at $T < T_c$ between the RS ferromagnetic phase at large ϵ and the intermediate zero magnetization 1RSB phase at smaller values of ϵ is determined by simply comparing their free energies and has the implicit equation:

$$f^{\text{ferro}}(\epsilon_c^{\text{ferro/1rsb}}, T_c^{\text{ferro/1rsb}}) = f^{\text{1rsb}}(\epsilon_c^{\text{ferro/1rsb}}, T_c^{\text{ferro/1rsb}}). \tag{B.14}$$

This is clearly a first order transition.

B.3 Other choices of the ratio σ_4/J_0

It is legitimate to wonder whether the choice

$$\sigma_4/J_0 = R(\epsilon) \sim \epsilon^{-1} \tag{B.15}$$

considered so far for the Gaussian distribution of non-linear couplings is fully representative of the case where for the non-linear couplings we have a competition between frustrated and ferromagnetic interactions. In general for the ratio $R(\epsilon)$ we have three possibilities: $R(\epsilon)$ is constant with respect to the parameter ϵ, which means that ferromagnetic and frustrated interactions are always competing; $R(\epsilon) \to \infty$ when $\epsilon \to \infty$,

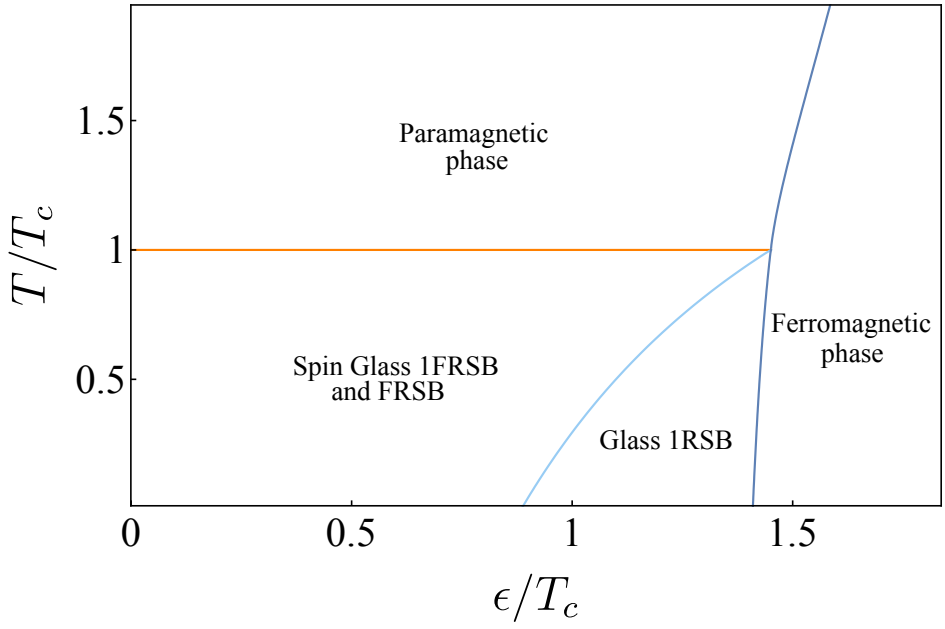

Figure B.3: Phase diagram in the (ϵ, T) plane for the $p = 2$ spherical spin glass plus four-body interactions: Gaussian distribution of non-linear couplings with standard deviation $\sigma_4 \sim \epsilon$ and mean $J_0 \sim \epsilon^4$ ($R \sim 1/\epsilon^3$).

which means that in the large-ϵ limit frustration is dominating; $R(\epsilon) \to 0$ when $\epsilon \to 0$, which means that in the large-ϵ limit the ferromagnetic ordering effect is dominating.

In order to have a clearer picture on how the phenomenology of the model depends on the choice of $R(\epsilon)$ we have studied the phase diagram, repeating the analysis already discussed above, also in the case where $\sigma_4 \propto \epsilon$ and $J_0 \propto \epsilon^4$, i.e., $R \propto \varepsilon^{-3}$, to check whether the increase of ferromagnetic terms relative importance introduces new glass phases (for instance the ferro-1RSB phase found in [79]), and in the case $R(\epsilon) = \text{const}$, in particular assuming $\sigma_4 \propto \epsilon$ and $J_0 \propto \epsilon$.

We have obtained phase from the case $R(\epsilon) = \text{const}$, for which the phase diagram is shown in Fig. B.2: it can be easily seen that it has the same topology of the phase diagram of the purely frustrated case in Fig. 2.5, which can be formally identified with the case $R = \infty$, since there we have $J_0 = 0$. In both cases there is no ferromagnetic phase! We have therefore checked, even without a full detailed analysis of the three-dimensional (T, σ_4, J_0) phase diagram, that already a constant ratio between σ_4 and J_0 is enough to produce the phenomenology that we find also in the limit $R = \infty$: the complete disappearance of the ordered ferromagnetic phase. On the other hand we have found that the phase diagram of the case $R \sim \epsilon^{-3}$, which is shown in Fig. B.3, is qualitatively identical to the case $R \sim \epsilon^{-1}$ of Fig. B.1, i.e., we do find ferromagnetic ordering at large ϵ and the same glassy phases at smaller values of the order parameter. We can therefore conclude that, despite the choice $R \sim \epsilon^{-1}$ of the present chapter looks quite arbitrary, is indeed representative of the model phenomenology when the importance of the ferromagnetic terms grows with ϵ.

C

Derivation of the effective Hamiltonian for the generalized Lotka-Volterra model

Let us now show that the effective Hamiltonian

$$H_{\text{eff}}(\mathbf{n}) = -\sum_{i=1}^{N} r_i \left(n_i - \frac{n_i^2}{2K_i} \right) + \sum_{(ij)\in E} \frac{\alpha_{ij}}{2} \left(\frac{r_i}{K_i} + \frac{r_j}{K_j} \right) n_i n_j + \sum_{i=1}^{N} \left[T \ln(n_i) - \ln \theta(n_i - \lambda) \right] \tag{C.1}$$

is the one that allows to describe the equilibrium thorugh a Gibbs–Boltzmann distribution $P(\mathbf{n}) \propto \exp(-\beta H_{\text{eff}}(\mathbf{n}))$. In particular, we outline why the symmetry of interactions is essential for defining such an equilibrium distribution.

It is known that a stochastic equation of the form

$$\dot{n}_i = f_i(n) + \sqrt{\Phi_i}\, \xi_i(t) \tag{C.2}$$

is associated, following Ito's prescription [80], with the Fokker-Planck equation for the distribution $q(\mathbf{n}, t)$:

$$\dot{q} = \sum_i \partial_i \Big[-f_i\, q + \frac{1}{2} \partial_i (\Phi_i\, q) \Big] . \tag{C.3}$$

For the gLV model, described by the dynamics:

$$\frac{dn_i(t)}{dt} = \frac{r}{K} n_i(t) \Big[K - n_i(t) - \sum_{j\in\partial i} \alpha_{ij} n_j(t) \Big] + \sqrt{n_i(t)} \xi_i(t), \tag{C.4}$$

the drift and noise terms read

$$f_i = \frac{r}{K}\, n_i \left[K - n_i - \sum_{j\in\partial i} \alpha_{ij} n_j \right], \tag{C.5}$$

$$\Phi_i = 2T\, n_i. \tag{C.6}$$

Substituting Eq. (C.6) into Eq. (C.3) and computing the derivative inside the square brackets, the Fokker-Planck equation becomes

$$\dot{q} = \sum_i \partial_i \left[q(-f_i + T) + T n_i \partial_i q \right]. \tag{C.7}$$

Let us now consider the Hamiltonian of Eq. (3.3), in the case of $r_i = r$, $K_i = K$, and without the immigration term (which plays only the role of a reflecting boundary):

$$H = -\sum_{i=1}^{N} r \left(n_i - \frac{n_i^2}{2K} \right) + \sum_{(ij)\in E} \alpha_{ij} \frac{r}{K}\, n_i n_j + \sum_{i=1}^{N} T \ln(n_i). \tag{C.8}$$

A straightforward computation yields

$$-n_i \partial_i H = f_i - T. \tag{C.9}$$

This allows Eq. (C.7) to be rewritten as

$$\dot{q} = \sum_i \partial_i \left[n_i q\, \partial_i H + T n_i \partial_i q \right]. \tag{C.10}$$

If we impose stationarity ($\dot{q} = 0$) and detailed balance, we obtain

$$q\, \partial_i H = -T\, \partial_i q, \tag{C.11}$$

which integrates to the stationary solution

$$q \propto \exp(-\beta H). \tag{C.12}$$

At this point, we have to emphasize an important subtlety. The above integration step is only consistent if H is a *true scalar potential*, which requires that the interaction matrix α_{ij} is symmetric. In that case, the drift field is conservative, detailed balance holds, and the stationary solution is indeed a Gibbs-Boltzmann equilibrium distribution.

If the interactions are not symmetric, the relation (C.9) still holds, but the function H does not define a proper potential (the mixed derivatives $\partial_i \partial_j H$ and $\partial_j \partial_i H$ do not coincide).

Thus, the Hamiltonian in Eq. (C.8) governs the equilibrium properties of the gLV system only when the interactions are symmetric, while in the general asymmetric case one must instead deal with non-equilibrium stationary states.

D

Universality and Scaling Law for the Sparse Ecosystems Dynamics

Here we focus on an ecologically relevant aspect of our system. In real ecosystems, extensive observations have documented the presence of Taylor's law (see [74]), which describes the relationship between the variance and mean of species abundances. Specifically, if one measures the variance V_i and the mean M_i for the abundance trajectory $n_i(t)$ of each species in the ecosystem, Taylor's law states that the variances follow a power-law dependence on the means:

$$V = aM^b, \tag{D.1}$$

where, in ecological systems, the exponent b has widely been reported to be $b \simeq 2$ [74, 81].

To investigate this in our sparse model, we set $\hat{\mu} = 0.1$ and $\hat{\sigma} = 0.25$, ensuring that the system is in the non-Gaussian (ecologically realistic) regime, see Sec. 3.4.1. We then run Langevin dynamics (see Fig. D.1) for the generalized Lotka-Volterra equations (3.1), computing the mean M_i and variance V_i for each species i:

$$M_i \equiv \frac{1}{t_{\max} - t_0} \sum_{t=t_0}^{t_{\max}} n_i(t), \quad V_i \equiv \frac{1}{t_{\max} - t_0} \sum_{t=t_0}^{t_{\max}} \left(n_i(t) - M_i\right)^2, \tag{D.2}$$

where t_0 is the initial transient time beyond which the distributions no longer depend on initial conditions and $t_{\max}$ is the final simulation time. We performed these simulations for $N = 400$ species and various temperatures T. In the left panel of Fig. D.2, we plot all pairs (M_i, V_i), with different colors representing different temperatures.

Regardless of the temperature, the data always exhibit a *power law* behaviour for small M_i, followed by a *saturation* region for large M_i. As shown in the left plot of Fig. D.2, the power law remains the same for all T, with the exponent b of Eq. (D.1)

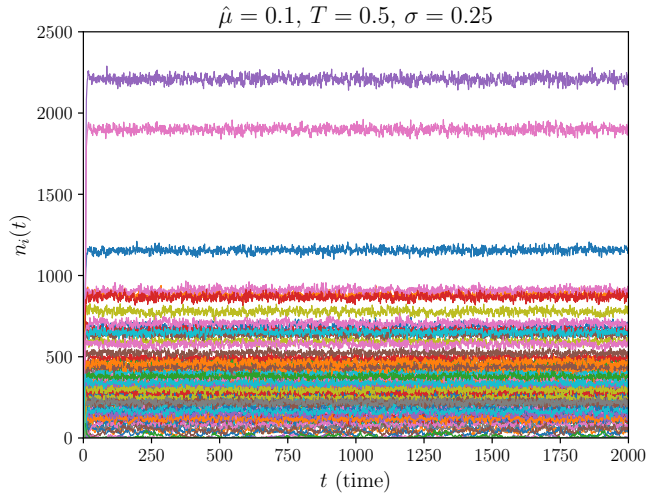

Figure D.1: Species abundance dynamics for $\hat{\mu} = 0.1, \hat{\sigma} = 0.25$ and $T = 0.5$. The number of species here is $N = 400$.

being $b \simeq 2$. In particular the black solid line represents the fit of the data for small M_i and is given by $V = aM^b$, with $a = 3.4$ and $b = 2.05$. This is a remarkable result, given the widespread occurrence of this power law in nature. Moreover, the same power law is found for different values of disorder $\hat{\sigma}$, provided that they are sufficiently large ($\hat{\sigma} \gtrsim 0.15$), corresponding to the non-Gaussian regime.

In the saturation regime instead, the variance is more or less constant ($V(M,T) = V_S(T)$), no matter which is the value of M. In particular, the higher is the temperature, the higher is the saturation value $V_S(T)$.

Given these observations on the power-law and saturation regimes, we propose a scaling ansatz for the variance $V(M,T)$:

$$V(M,T) = M^{\delta} F\left(\frac{M}{T^{\psi}}\right), \tag{D.3}$$

where V is the variance, M is the mean, T is the temperature and δ and ψ are critical exponents that have to be computed. Looking at our data we state that the function F is such that:

- $F(x) \sim$ constant for $x \to 0$,
- $F(x) \sim x^{-\delta}$ for $x \to \infty$.

The exponent δ is found by fitting the low-M regime data. In particular, as we already said, we get:

$$\delta \simeq 2.$$

Regarding the exponent ψ, we can compute it by noticing that, for each temperature T, the variance saturation value is:

$$V_S(T) = \lim_{M \to \infty} V(M,T).$$

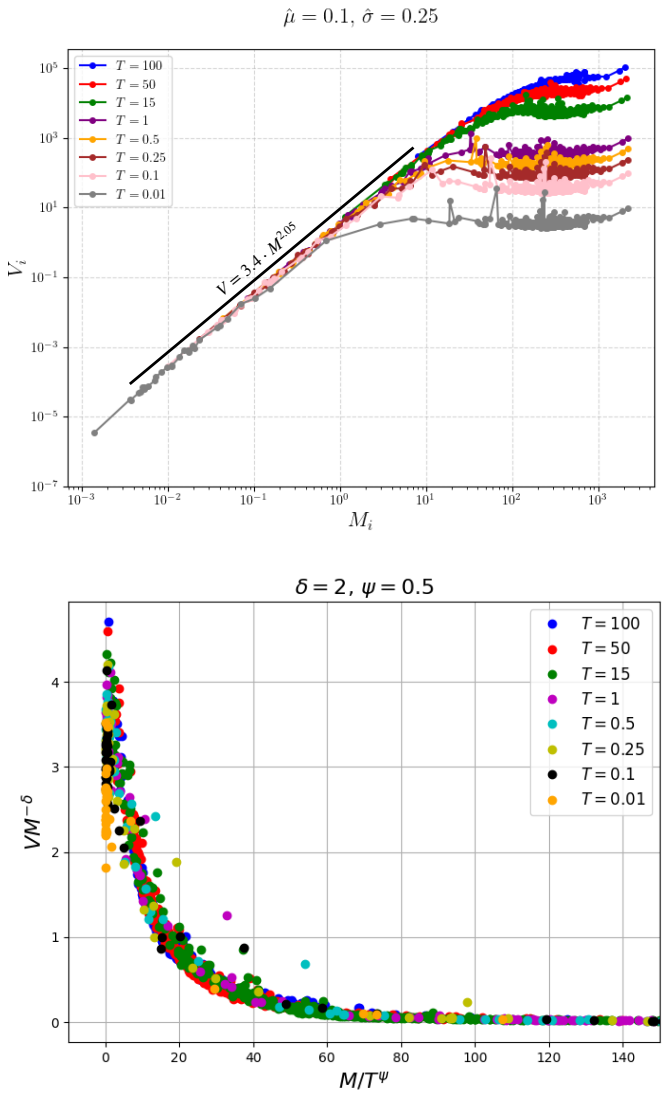

Figure D.2: For this plots $\hat{\mu} = 0.1, \hat{\sigma} = 0.25$ and $N = 400$. Left panel: Plot of the pairs of mean and variance (M_i, V_i) for different values of the temperature. For all the temperatures there is a power-law region for low values of M_i and a saturation region for large values of M_i. Right panel: Representation of the scaling law which allows us to make all the data collapse into a unique curve, see eq. (D.3).

Recalling the behaviour of $F(x)$ for $x \to \infty$, from Eq. (D.3) we get:

$$V_S(T) \sim M^{\delta} \frac{M^{-\delta}}{T^{-\psi\delta}} = T^{\psi\delta}.$$

So, once we have the value of δ, we can compute ψ by fitting all the values $V_S(T)$ with the function $V_S(T) = T^{\psi\delta}$. In this way we get:

$$\psi \simeq 0.5.$$

To conclude, on the right panel of Fig. D.2 we plotted, for all the data of the left panel, $VM^{-\delta}$ vs M/T^{ψ}, with δ and ψ set to the values that we found. In this way we can really see that the scaling of Eq. (D.3) works and all the data from different ecosystems at different temperatures collapse onto the same curve. These results highlight the emergence of a universality mechanism.

E

Full derivation of the 1RSB cavity equations

In Chapter 4, in ordered to study the 1RSB multiple-equilibria phases, we discussed the need to create m different replicas of the system with a weak coupling between them. Since the m replicas are independent inside a given state a, the partition function of the replicated system is

$$Z(m) = \sum_a e^{-m\beta F^a}. \tag{E.1}$$

By rewriting Eq. (E.1) as an integral over all the possible cavity messages on the interaction graph we get:

$$Z(m) = \int \mathbf{D}\eta \, e^{-m\beta F(\eta)} \prod_{(ij)\in E} \left[\delta(\eta_{i\to j} - \mathcal{F}_{i\to j}(\{\eta_{k\to i}\}))\delta(\eta_{j\to i} - \mathcal{F}_{j\to i}(\{\eta_{l\to j}\}))\right], \tag{E.2}$$

where $\mathbf{D}\eta = \prod_{i,j\in E} d\eta_{i\to j}$ and the deltas impose the BP equations on each edge (i,j).

Let us specify, and it is detailed in [73], that the free energy of each state is given by:

$$-\beta F\left[\{\eta_{i\to j}\}\right] = \sum_i \log z_i - \sum_{\langle i,j\rangle} \log z_{ij}. \tag{E.3}$$

In particular, let us recall the expression of the normalizations $z_{i\to j}$ and z_i

$$z_{i\to j} = \sum_{n_i} \psi_i(n_i) \prod_{k\in\partial i\setminus j} \left(\sum_{n_k} \eta_{k\to i}(n_k)\psi_{ik}(n_i,n_k)\right), \tag{E.4}$$

$$z_i = \sum_{n_i} \psi_i(n_i) \prod_{j\in\partial i} \left(\sum_{n_j} \eta_{j\to i}(n_j)\psi_{ij}(n_i,n_j)\right). \tag{E.5}$$

The term z_{ij} of Eq. (E.3) can be expressed as a function of $z_{i\to j}$ and z_i in the following way:

$$z_{ij} = \sum_{n_i,n_j} \eta_{j\to i}(n_j)\eta_{i\to j}(n_i)\psi_{ij}(n_i,n_j) = \frac{z_j}{z_{j\to i}} = \frac{z_i}{z_{i\to j}}. \tag{E.6}$$

This means that, from Eq. (E.3), we get:

$$e^{-\beta m F(\eta)} = \frac{\prod_i z_i^m}{\prod_{ij} z_{ij}^m}. \tag{E.7}$$

Inserting Eq. (E.7) in Eq. (E.2) we get:

$$Z_m = \int \mathcal{D}\eta_{i\to j} \prod_i z_i^m \prod_{\langle i,j\rangle} z_{ij}^{-m} \delta[\eta_{i\to j} - \mathcal{F}_{i\to j}]\delta[\eta_{j\to i} - \mathcal{F}_{j\to i}] \tag{E.8}$$

$$= \int \mathcal{D}\eta_{i\to j} \prod_i \left(z_i^m \prod_{j\in\partial i} \delta[\eta_{i\to j} - \mathcal{F}_{i\to j}] \right) \prod_{\langle i,j\rangle} z_{ij}^{-m} \tag{E.9}$$

$$= \int \mathcal{D}\eta_{i\to j} \prod_i \Psi_i(\{\eta_{i\to j}, \eta_{j\to i}\}_{j\in\partial i}) \prod_{\langle i,j\rangle} \Psi_{ij}(\eta_{i\to j}, \eta_{j\to i}) \tag{E.10}$$

In Eq. (E.10) there are two interaction terms. The first one is a product over all the sites of Ψ_i which depends on all the cavity messages which involve site i. The second term is a product over all the links of Ψ_{ij} that depends only on the cavity messages on that link.

Let us then consider the cavity messages as variables. From Eq. (E.9), we can say that on each link (ij) of the original graph we have a variable made by the two messages $\{\eta_{i\to j}, \eta_{j\to i}\}$ with a local field $z_{ij}^{-m} = \Psi_{ij}$. Instead on each variable-node of the original graph, we have a *many-body* interaction $z_i^m \prod_{j\in\partial i} \delta[\eta_{i\to j} - \mathcal{F}_{i\to j}] = \Psi_i$. Therefore we can re-interpret the original graph, made of only variable nodes, as a factor graph, where on each link there is a variable node, while the original variable nodes become factor/interaction nodes.

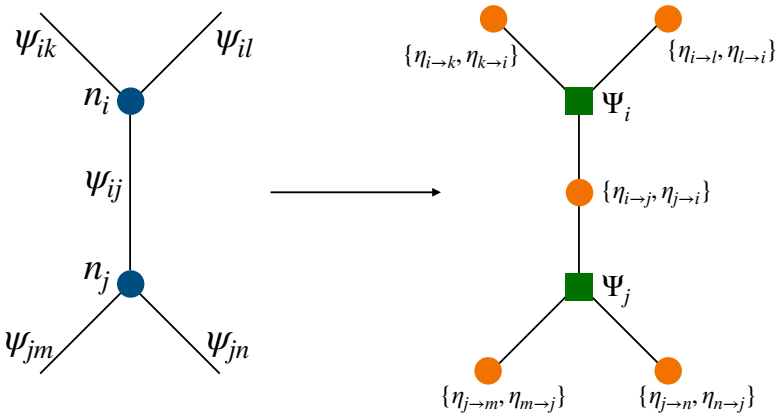

Figure E.1: **Left**: original graph made of only variable nodes n_i. **Right**: associated factor graph for the construction of the 1RSB model. On each link (ij) of the original graph we have the variable $\{\eta_{i\to j}, \eta_{j\to i}\}$, on each variable node of the original graph we have a many-body interaction term Ψ_i.

For this auxiliary model then, we can write the BP cavity equations for the cavity distribution $Q_{i\to j}$ of node variables $\{\eta_{i\to j}, \eta_{j\to i}\}$. As the RS cavity messages $\eta_{i\to j}$ were

written in Eq. (3.10), in an equivalent way we can write the BP equatons for the cavity distributions $Q_{i\to j}$:

$$Q_{i\to j}(\eta_{i\to j},\eta_{j\to i}) = \frac{\Psi_{ij}(\eta_{i\to j},\eta_{j\to i})}{\mathcal{Z}_{i\to j}} \sum_{\{\eta_{i\to k},\eta_{k\to i}\}_{k\in\partial i\setminus j}} \Psi_i(\{\eta_{i\to l},\eta_{l\to i}\}_{l\in\partial i}) \prod_{k\in\partial i\setminus j} Q_{k\to i}(\eta_{k\to i},\eta_{i\to k}). \tag{E.11}$$

Using the explicit expressions of Ψ_i and Ψ_{ij} we get:

$$\begin{aligned}
&Q_{i\to j}(\eta_{i\to j},\eta_{j\to i}) = \\
&= \frac{z_{ij}\left(\eta_{i\to j},\eta_{j\to i}\right)^{-m}}{\mathcal{Z}_{i\to j}} \sum_{\{\eta_{i\to k},\eta_{k\to i}\}_{k\in\partial i\setminus j}} z_i\left(\{\eta_{k\to i}\}_{k\in\partial i}\right)^m \prod_{l\in\partial i} \delta\left[\eta_{i\to l}-\mathcal{F}_{i\to l}\right] \prod_{k\in\partial i\setminus j} Q_{k\to i}\left(\eta_{k\to i},\eta_{i\to k}\right) \\
&= \frac{z_{ij}\left(\eta_{i\to j},\eta_{j\to i}\right)^{-m}}{\mathcal{Z}_{i\to j}} \sum_{\{\eta_{k\to i}\}_{k\in\partial i\setminus j}} z_i\left(\{\eta_{k\to i}\}_{k\in\partial i}\right)^m \delta\left[\eta_{i\to j}-\mathcal{F}_{i\to j}\right] \prod_{k\in\partial i\setminus j} Q_{k\to i}\left(\eta_{k\to i},\mathcal{F}_{i\to k}\right) \\
&= \frac{1}{\mathcal{Z}_{i\to j}} \sum_{\{\eta_{k\to i}\}_{k\in\partial i\setminus j}} z_{i\to j}\left(\{\eta_{k\to i}\}_{k\in\partial i\setminus j}\right)^m \delta\left[\eta_{i\to j}-\mathcal{F}_{i\to j}\right] \prod_{k\in\partial i\setminus j} Q_{k\to i}\left(\eta_{k\to i},\mathcal{F}_{i\to k}\right),
\end{aligned} \tag{E.12}$$

where in the last step we used Eq. (E.6). In particular, looking at the last line of Eq. (E.12), we notice that $\eta_{j\to i}$ only appears in the second argument of $Q_{k\to i}$, since it is in the argument of $\mathcal{F}_{i\to k}$. This means that, if the $Q_{i\to j}(\eta_{i\to j},\eta_{j\to i})$ are initialized as independent of their second argument $\eta_{j\to i}$, they will remain independent of their second argument also after the iteration. So, we can write:

$$Q_{i\to j}(\eta_{i\to j}) = \frac{1}{Z_{i\to j}} \sum_{\{\eta_{k\to i}\}_{k\in\partial i\setminus j}} z_{i\to j}(\{\eta_{k\to i}\}_{k\in\partial i\setminus j})^m \delta\left[\eta_{i\to j}-\mathcal{F}_{i\to j}\left[\{\eta_{k\to i}\}_{k\in\partial i\setminus j}\right]\right] \prod_{k\in\partial i\setminus j} Q_{k\to i}(\eta_{k\to i}), \tag{E.13}$$

which is Eq. (4.6) in Chapter 4.

In particular, we can also write the replicated free energy as:

$$-\beta N\Phi(m,T) = \log Z_m = \sum_i \log \mathcal{Z}_i - \sum_{ij} \log \mathcal{Z}_{ij}, \tag{E.14}$$

where

$$\mathcal{Z}_i = \sum_{\{\eta_{k\to i}\}_{k\in\partial i}} z_i(\{\eta_{k\to i}\}_{k\in\partial i})^m \prod_{k\in\partial i} Q_{k\to i}(\eta_{k\to i}) = \langle z_i^m\rangle_Q, \tag{E.15}$$

$$\mathcal{Z}_{ij} = \sum_{\eta_{i\to j},\eta_{j\to i}} z_{ij}(\eta_{i\to j},\eta_{j\to i})^m Q_{i\to j}(\eta_{i\to j}) Q_{j\to i}(\eta_{j\to i}) = \langle z_{ij}^m\rangle_Q. \tag{E.16}$$

Bibliography

[1] E. Odum. *Fundamentals of Ecology*. Saunders, 1953.

[2] C. Elton. *Ecology of invasions by animals and plants*. Chapman and Hall, 1958.

[3] R. MacArthur.
"Fluctuations of Animal Populations and a Measure of Community Stability".
In: *Ecology* 36.3 (1955), pp. 533–536.

[4] R. M. May. "Will a Large Complex System be Stable?"
In: *Nature* 238 (1972), pp. 413–414.

[5] R. M. May. "Qualitative stability in model ecosystems".
In: *Ecology* 54.3 (1973), pp. 638–641.

[6] B. E. Marshall. "Guilty as charged: Nile perch was the cause of the haplochromine decline in Lake Victoria". In: *Canadian Journal of Fisheries and Aquatic Sciences* 75.9 (2018), pp. 1542–1559.

[7] P. Verhulst. "Notice sur la loi que la population suit dans son accroissement".
In: *Correspondance mathématique et physique* 10 (1838), pp. 113–121.

[8] A. Lotka.
"Analytical Note on Certain Rhythmic Relations in Organic Systems".
In: *Proceedings of the National Academy of Sciences* 6.7 (1920), pp. 410–415.

[9] V. Volterra.
"Variazioni e fluttuazioni del numero d'individui in specie animali conviventi".
In: *Mem. Acad. Lincei Roma* (1926).

[10] S. Hubbell. *The Unified Neutral Theory of Biodiversity and Biogeography*.
Princeton University Press, 2001.

[11] R. MacArthur and E. O. Wilson. *The Theory of Island Biogeography*.
Princeton University Press, 1967.

[12] I. Volkov et al. "Neutral theory and relative species abundance in ecology".
In: *Nature* 424.6952 (2003), pp. 1035–1037.

[13] S. Azaele et al.
"Statistical mechanics of ecological systems: Neutral theory and beyond".
In: *Reviews of Modern Physics* 88.3 (2016), p. 035003.

[14] S. Azaele et al. "Dynamical evolution of ecosystems".
In: *Nature* 444 (2006), pp. 926–928.

[15] S. Nee. “The neutral theory of biodiversity: do the numbers add up?” In: *Functional Ecology* 19.1 (2005), pp. 173–176.

[16] J. S. Clark. “Beyond neutral science”. In: *Trends in ecology and evolution* 24 (2009), pp. 8–15.

[17] G. E. Hutchinson. “The Paradox of the Plankton”. In: *The American Naturalist* 95.882 (1961), pp. 137–145.

[18] R. MacArthur. “Species Packing, and What Competition Minimizes”. In: *PNAS* 64 (1969), pp. 1369–1371.

[19] R. MacArthur. “Species packing and competitive equilibrium for many species”. In: *Theoretical Population Biology* 1 (1970), pp. 1–11.

[20] J. Ginibre. “Statistical Ensembles of Complex, Quaternion, and Real Matrices”. In: *Journal of Mathematical Physics* 6.3 (Mar. 1965), pp. 440–449.

[21] K. S. McCann. “The diversity–stability debate”. In: *Nature* 405 (2000), pp. 228–233.

[22] S. Diederich and M. Opper. “Replicators with random interactions: A solvable model”. In: *Physical Review A* 39.8 (1989), p. 4333.

[23] P. Biscari and G. Parisi. “Replica symmetry breaking in the random replicant model”. In: *Journal of Physics A: Mathematical and General* 28.17 (Sept. 1995).

[24] M. Tikhonov and R. Monasson. “Collective phase in resource competition in a highly diverse ecosystem”. In: *Phys. Rev. Lett.* **118** (4 2017), p. 048103.

[25] A. Altieri and S. Franz. “Constraint satisfaction mechanisms for marginal stability and criticality in large ecosystems”. In: *Phys. Rev. E* **99** (1 2019), p. 010401.

[26] A. Altieri et al. “Properties of equilibria and glassy phases of the random Lotka-Volterra model with demographic noise”. In: *Phys. Rev. Lett.* **126** (25 2021), p. 258301.

[27] G. Biroli, G. Bunin and C. Cammarota. “Marginally stable equilibria in critical ecosystems”. In: *New Journal of Physics* 20.8 (Aug. 2018), p. 083051.

[28] G. Bunin. “Ecological communities with Lotka-Volterra dynamics”. In: *Physical Review E* 95.4 (2017), p. 042414.

[29] V. Ros et al. “Generalized lotka-volterra equations with random, nonreciprocal interactions: The typical number of equilibria”. In: *Physical Review Letters* 130.25 (2023), p. 257401.

[30] J. Grilli. “Macroecological laws describe variation and diversity in microbial communities”. In: *Nature communications* 11.1 (2020), p. 4743.

[31] H. Sompolinsky and A. Zippelius. “Relaxational dynamics of the Edwards-Anderson model and the mean-field theory of spin-glasses”. In: *Physical Review B* 25.11 (1982), p. 6860.

[32] L. F. Cugliandolo and J. Kurchan. “Analytical solution of the off-equilibrium dynamics of a long-range spin-glass model”. In: *Phys. Rev. Lett.* 71 (1 July 1993), pp. 173–176.

[33] J. A. Dunne, R. J. Williams and N. D. Martinez. “Food-web structure and network theory: The role of connectance and size”. In: *Proceedings of the National Academy of Sciences* 99.20 (2002), pp. 12917–12922.

[34] J. E. Cohen, F. Briand and C. M. Newman. *Community Food Webs: Data and Theory*. Springer, 2011.

[35] *Web of Life: Ecological Networks Database*. `https://www.web-of-life.es/`.

[36] D. Sherrington and S. Kirkpatrick. “Solvable model of a spin-glass”. In: *Phys. Rev. Lett.* **35** (1975), p. 1792.

[37] J. R. L. de Almeida and D. J. Thouless. “Stability of the Sherrington-Kirkpatrick solution of a spin glass model”. In: *J. Phys. A Math. Gen.* **11**.5 (1978), p. 983.

[38] G. Parisi. “Infinite number of order parameters for spin-glasses”. In: *Phys. Rev. Lett.* **43** (1979), pp. 1754–1756.

[39] H. Nishimori. *Statistical Physics of Spin Glasses and Information Processing: An Introduction*. Oxford University Press, July 2001.

[40] M. Mezard, G. Parisi and M. Virasoro. *Spin Glass Theory and Beyond*. World Scientific, 1986.

[41] F. Guerra. “Broken Replica Symmetry Bounds in the Mean Field Spin Glass Model”. In: *Communications in Mathematical Physics* 233 (June 2002).

[42] M. Talagrand. “The Parisi formula”. In: *Ann. Math. (2)* 163.1 (2006), pp. 221–263.

[43] M. Mezard and A. Montanari. *Information, physics, and computation*. Oxford University Press, 2009.

[44] C. Cammarota et al. “Fragility of the mean-field scenario of structural glasses for disordered spin models in finite dimensions”. In: *Phys. Rev. B* 87 (6 Feb. 2013), p. 064202.

[45] J. M. Kosterlitz, D. J. Thouless and R. C. Jones. “Spherical model of a spin-glass”. In: *Phys. Rev. Lett.* **36** (20 1976), pp. 1217–1220.

[46] J. Niedda, T. Tonolo and G. Gradenigo. “Probing marginal stability in the spherical p = 2 model”. In: *Journal of Statistical Mechanics: Theory and Experiment* 2024.11 (2024), p. 113301.

[47] C. D. Dominicis and I. Giardina.
Random Fields and Spin Glasses: A Field Theory Approach.
Cambridge University Press, 2006.

[48] A. Crisanti and C. De Dominicis.
"Replica Fourier Transform: properties and applications".
In: *Nucl. Phys. B* **891** (2015), pp. 73–105.

[49] A. Crisanti and H. Sommers.
"The spherical p-spin interaction spin-glass model - the statics".
In: *Z. Phys. B* **87** (1992), p. 341.

[50] T. Temesvari, C. D. Dominicis and I. Kondor.
"Block diagonalizing ultrametric matrices".
In: *Jour. Phys. A: Math. Gen.* **27** (1994), p. 7569.

[51] A. Crisanti and L. Leuzzi. "Spherical $2 + p$ spin-glass model: an exactly solvable model for glass to spin-glass transition".
In: *Phys. Rev. Lett.* **93** (21 2004), p. 217203.

[52] A. Crisanti and L. Leuzzi. "Spherical $2 + p$ spin-glass model: an analytically solvable model with a glass-to-glass transition".
In: *Phys. Rev. B* **73** (1 2006), p. 014412.

[53] T. Galla. "Dynamically evolved community size and stability of random Lotka-Volterra ecosystems (a)". In: *Europhysics Letters* 123.4 (2018), p. 48004.

[54] T. Galla. *Generating-functional analysis of random Lotka-Volterra systems: A step-by-step guide*. 2024. arXiv: 2405.14289.

[55] G. Garcia Lorenzana and A. Altieri. "Well-mixed Lotka-Volterra model with random strongly competitive interactions".
In: *Physical Review E* 105.2 (2022), p. 024307.

[56] I. Volkov et al.
"Patterns of relative species abundance in rainforests and coral reefs".
In: *Nature* 450.7166 (2007), pp. 45–49.

[57] S. Franz et al. "A ferromagnet with a glass transition".
In: *Europhysics Letters* 55.4 (2001), p. 465.

[58] R. May and A. R. McLean. *Theoretical ecology: principles and applications*.
Oxford University Press, 2007.

[59] R. V. Solé, D. Alonso and A. McKane.
"Self–organized instability in complex ecosystems".
In: *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 357.1421 (2002), pp. 667–681.

[60] B. J. McGill et al. "Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework".
In: *Ecology Letters* 10.10 (2007), pp. 995–1015.

[61] V. Ros et al. “Quenched complexity of equilibria for asymmetric generalized lotka–volterra equations”.
In: *Journal of Physics A: Mathematical and Theoretical* 56.30 (2023), p. 305003.

[62] G. G. Lorenzana et al. “Non-reciprocal spin-glass transition and aging”.
In: *arXiv preprint arXiv:2408.17360* (2024).

[63] P. Valigi, I. Neri and C. Cammarota. “Local sign stability and its implications for spectra of sparse random graphs and stability of ecosystems”.
In: *Journal of Physics: Complexity* 5.1 (2024), p. 015017.

[64] J. Camacho-Mateu et al. “Sparse species interactions reproduce abundance correlation patterns in microbial communities”.
In: *Proceedings of the National Academy of Sciences* 121.5 (2024).

[65] J. I. Park et al.
“Incorporating Heterogeneous Interactions for Ecological Biodiversity”.
In: *Phys. Rev. Lett.* 133 (19 2024).

[66] P. Jordano. “Sampling networks of ecological interactions”.
In: *Functional Ecology* 30.12 (2016), pp. 1883–1893.

[67] T. Strydom et al. “A roadmap towards predicting species interaction networks (across space and time)”. In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 376.1837 (2021).

[68] A. A. M. MacDonald, F. Banville and T. Poisot.
“Revisiting the Links-Species Scaling Relationship in Food Webs”.
In: *Patterns* 1.7 (2020), p. 100079.

[69] S. Suweis et al. “Generalized lotka-volterra systems with time correlated stochastic interactions”. In: *Physical Review Letters* 133.16 (2024), p. 167101.

[70] T. Tonolo et al. “Generalized Lotka-Volterra model with sparse interactions: non-Gaussian effects and topological multiple-equilibria phase”.
In: *PRX Life* (2025).

[71] S. Marcus, A. M. Turner and G. Bunin.
“Local and collective transitions in sparsely-interacting ecological communities”.
In: *PLoS computational biology* 18.7 (2022), e1010274.

[72] L. Poley, T. Galla and J. W. Baron.
“Interaction networks in persistent Lotka-Volterra communities”.
In: *arXiv preprint arXiv:2404.08600* (2024).

[73] F. Zamponi. *Mean field theory of spin glasses*. 2014.
arXiv: `1008.4844 [cond-mat.stat-mech]`.

[74] A. Giometto et al.
“Sample and population exponents of generalized Taylor’s law”. In: *Proceedings of the National Academy of Sciences* 112.25 (2015), pp. 7755–7760.

[75] A. Crisanti and L. Leuzzi.
“Stable Solution of the Simplest Spin Model for Inverse Freezing”.
In: *Phys. Rev. Lett.* 95 (8 2005).

[76] L. Leuzzi, M. Paoluzzi and A. Crisanti.
"Random Blume-Capel model on a cubic lattice: First-order inverse freezing in a three-dimensional spin-glass system". In: *Physical Review B—Condensed Matter and Materials Physics* 83.1 (2011), p. 014107.

[77] D. Machado et al. *Local equations for the generalized Lotka-Volterra model on sparse asymmetric graphs*. 2025. arXiv: 2511.17499.

[78] R. Monasson.
"Structural Glass Transition and the Entropy of the Metastable States".
In: *Phys. Rev. Lett.* 75 (15 1995), pp. 2847–2850.

[79] A. Crisanti and L. Leuzzi.
"Exactly solvable spin–glass models with ferromagnetic couplings: The spherical multi-p-spin model in a self-induced field".
In: *Nuclear Physics B* 870.1 (2013), pp. 176–204.

[80] C. Gardiner. *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences*. Springer, 2004.

[81] R. Anderson, D. Gordon and M. e. a. Crawley.
"Variability in the abundance of animal and plant species".
In: *Nature* 296 (1982), pp. 245–248.

List of Figures