Universidade de São Paulo
Instituto de Física

Informação Integrada e Medidas de Complexidade para Sistemas Desordenados

Otávio Cístolo Citton

Orientador: Prof. Dr. Nestor Caticha

Dissertação de mestrado apresentada ao Instituto de Física da Universidade de São Paulo, como requisito parcial para a obtenção do título de Mestre em Ciências.

Banca Examinadora:
Prof. Dr. Nestor Felipe Caticha Alfonso - Orientador (IFUSP)
Prof. Dr. Renato Vicente (IME-USP)
Prof. Dr. Juan Pablo Neirotti (Aston)

São Paulo
2023

FICHA CATALOGRÁFICA
Preparada pelo Serviço de Biblioteca e Informação
do Instituto de Física da Universidade de São Paulo

Citton, Otávio Cistolo

Informação integrada e medidas de complexidade para sistemas desordenados / Integrated Information and Complexity Measures for Disordered Systems. São Paulo, 2023.

Dissertação (Mestrado) – Universidade de São Paulo. Instituto de Física. Departamento de Física Geral.
Orientador(a): Profº Drº Nestor Felipe Caticha Alfonso

Área de Concentração: Física Geral

Unitermos: 1. Mecânica estatística; 2. Sistemas complexos; 3. Índices de complexidade; 4. Informação integrada. 5. Sistemas desordenados.

USP/IF/SBI-001/2023

University of São Paulo
Physics Institute

Integrated Information and Complexity Measures for Disordered Systems

Otávio Cístolo Citton

Supervisor: Prof. Dr. Nestor Caticha

Dissertation submitted to the Physics Institute of the University of São Paulo in partial fulfillment of the requirements for the degree of Master of Science.

Examining Committee:
Prof. Dr. Nestor Felipe Caticha Alfonso - Supervisor (IFUSP)
Prof. Dr. Renato Vicente (IME-USP)
Prof. Dr. Juan Pablo Neirotti (Aston)

São Paulo
2023

Acknowledgements

First of all I would like to thanks all my family, Adilson Ferreira Citton, Geovana Márcia de Cístolo Citton and Luana Cístolo Citton, without whom this work would not have been possible. They have given me all the support throughout my whole academic and personal life. This achievement is as much theirs as it is mine.

I would like to give a special thanks to my advisor prof. Nestor Caticha. It was a pleasure having him as a professor and working with him. I always enjoyed our conversations about physics, mathematics, philosophy, neuroscience, literature, coffee and all sort of other unrelated topics.

I also want to thank all my colleagues that accompanied me during my undergraduate studies and this master's project: Evanildo Lacerda, Felippe Pereira, Ian Pinto, José Arthur, Mariana Mercucci, Pietro Zanin and specially João Henrique de Sant'ana and Gustavo Eduardo Menesse with whom I had multiple interesting discussions. They always listened and helped me with problems related to this work, so they also share a portion of this accomplishment.

I am grateful also to all my friends that helped me distract and think about other things. You are all amazing and have made my life way better by joining me for movies, games, books, travels or even just chatting, specially during the pandemics.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001. I thank CAPES for the financial support that made possible the development of this project.

Abstract

Motivated by the possible applications that a better understanding of consciousness might bring, we follow Tononi's idea and calculate analytically a complexity index for a disordered system. Utilizing the information geometry formulation of integrated information theory, and by restricting our analysis to bipartitions of the system, we calculate the geometric integrated information index for the model we call Little-Sherrington-Kirkpatrick, a synchronous dynamics version of the SK spin-glass model with quenched Gaussian interactions. The effects of partitioning are taken into account by introducing site dilution. We show that this complexity index can be used to rank the three phases of the system in terms of its complexity, and make an analysis on how it changes when we vary the partitioning of the system. Finally, by approximating the dynamics near-equilibrium, we briefly analyze the behavior of the geometric integrated information index out of equilibrium.

Keywords: Statistical Mechanics; Complex Systems; Complexity Index; Integrated Information; Disordered Systems.

Resumo

Motivados pelas possiveis aplicações que um melhor entendimento da consciência pode trazer, seguimos a ideia de Tononi e calculamos analiticamente um índice de complexidade para um sistema desordenado. Utilizando a formulação de geometria da informação para a teoria de informação integrada, e restringindo nossa análise a bipartições do sistema, nós calculamos o índice de informação integrada geométrico para o modelo que chamamos de Little-Sherrington-Kirkpatrick, uma versão com dinâmica síncrona do modelo SK de vidro de spin com interações Gaussianas temperadas. Os efeitos de particionamento foram levados em conta introduzindo diluição de sítios. Mostramos que esse índice de complexidade pode ser usado para classificar as três fases do sistema em termos de sua complexidade e, fazemos uma análise de como ele muda quando variamos o particionamento do sistema. Finalmente, aproximando a dinâmica próxima do equilíbrio, analisamos brevemente o comportamento do índice de informação integrada fora do equilíbrio.

Palavras-chave: Mecânica Estatística; Sistemas Complexos; Índices de Compelxidade; Informação Integrada; Sistemas Desordenados.

Contents

1 Introduction **5**
- 1.1 Complex Systems . . . 5
- 1.2 Complexity Indices . . . 6
  - 1.2.1 Integrated Information Theory . . . 6
  - 1.2.2 Neuroscientific Applications . . . 7
  - 1.2.3 Information Geometric Framework . . . 9

2 Geometric Integrated Information **10**
- 2.1 Integrated Information for a Bipartition . . . 11

3 The Little-Sherrington-Kirkpatrick Model **13**
- 3.1 The Model Definition . . . 13
- 3.2 Free Energy . . . 14
- 3.3 Equations of State . . . 16
- 3.4 Near Equilibrium Dynamics . . . 17

4 Calculation of ϕ for the LSK Model **20**
- 4.1 Integrated Information for a Disordered System . . . 20
- 4.2 Implementing Disconnections by Dilutions . . . 21
- 4.3 Calculation of $\phi_{\boldsymbol{\eta}}$ . . . 22

5 Results Analysis **25**
- 5.1 The γ dependence . . . 25
- 5.2 Map of ϕ at equilibrium . . . 26
- 5.3 ϕ on its way to equilibrium . . . 29
- 5.4 Connection with Stochastic Interactions . . . 32

6 Conclusions **33**

A Free Energy for LSK Model **37**
- A.1 Replica Symmetric Ansatz . . . 40

B $\Phi_{\boldsymbol{\eta}}$ Calculation **41**
- B.1 Calculation of ϕ_A . . . 41
- B.2 Calculation of ϕ_B . . . 50
- B.3 Calculation of ϕ_C . . . 54

Chapter 1

Introduction

In this dissertation we investigate a model we have called the Little-Sherrington-Kirkpatrick spin glass model, with the lens of Integrated Information Theory.

This model belongs to a very broad class of problems called Complex Systems. The most prominent example of such systems in physics is probably the Sherrington-Kirkpatrick spin glass [1] a tractable extension of the seminal Edwards-Anderson spin glass model [2] that has been extensively studied since 1975. These studies yielded a very good understanding [3], [4] of the kind of problems such systems may exhibit, led to the development of useful tools and raised new questions that are yet to be answered. For technical reasons, we study a Sherrington-Kirkpatrick model with parallel or synchronous dynamics, and hence the addition of Little to the name of the model [5].

The second ingredient of this work is the concept from Integrated Information Theory, a theory first introduced in 2004 by Tononi [6] as an attempt to describe consciousness from a first-person perspective and, in principle, to provides us with tools to infer if a given system is conscious. This is a very ambitious idea, but the theory is still in its early days and this work is an attempt to shed some light on Tononi's theory applying its concepts to a complex physical model. The main contribution of Tononi to the discussion of consciousness is the attempt to introduce a quantitative measure Φ, a complexity index to characterize a conscious system. Despite the long time since its introduction, the calculation of Φ and its variants remains elusive and only very simple systems with a few degrees of freedom, which would hardly classify as conscious, have been analyzed.

1.1 Complex Systems

The studies of complex systems is a subject that extends across all areas of science and is used to model all sorts of problems. The understanding of such systems is crucial to gain insight and tackle very interesting problems that may seem hopeless at a first glance.

There is no consensus on the exact definition of a complex system, but there are some concepts that appear in most of them: complex systems are systems that present emergent properties that cannot be reduced to properties of the individual elements, they present multiscale phenomena and there is an interplay between order and disorder which make this kind of systems very difficult to predict.

While a precise definition of complex is not necessary, Giorgio Parisi (see [7]) states that a system is complex if its behavior is very dependent on its details, that is, to know the model approximately might not help predict the exact system behavior. This sensibility and uncertainty on the model is due to incomplete information and the formalism of statistical mechanics and information theory can be used to find the (potentially few) interesting quantities that systematically describe the complicated dynamics in a simpler manner and, with this, make predictions about the states of the system.

1.2 Complexity Indices

The study of complex systems is very challenging, and an interesting concept that tries to encapsulate its complex behavior are the so-called complexity indices. These quantities have found applications in many areas and they provide a way to compare different systems or different states of the same system, and equip us with tools to classify the behavior of complex systems. One such index that we focus here is the Integrated Information Index. It has been introduced by Giulio Tononi in 2004 [6] as part of its Integrated Information Theory and was firstly designed to measure the intrinsic irreducibility of a system, a fundamental aspect for the conscious experience, according to Tononi.

1.2.1 Integrated Information Theory

Tononi's Integrated Information Theory (IIT) can be viewed as a new approach to the problem of consciousness: instead of trying to give a physical description for the conscious experience, he takes as a given that the only certainty we have is our inner experiences and, from there, proposes a quantitative measure a system ought to have in order to present the properties he postulated for our conscious experience. A physical system that satisfy these properties would be called a Physical Substrate of Consciousness (PSC).

The main quantity proposed by IIT that encapsulates these properties was called the Integrated Information Index, denoted by Φ, and it is a measure of the causal influences of the different parts of the system, how much influence parts of the systems exert on each other that can not be reduced to small independent systems.

Since 2004 IIT has gone through many iterations and the properties postulated by Tononi and the definition of Φ have changed, and no consensus was achieved. But the main structure of IIT, where one assumes that the conscious experience is real and tries to describe the properties of a physical substrate of consciousness, remained.

The postulates are at least six and throughout the different iterates of the theory have been subject to changes, not always in the direction of clarification. On its current state (see [8] and [9]), the postulates which try to translate the properties of the conscious experience, can be summarized as follows:

- *Existence* - (or causality) Experience exists. It is an undeniable aspect of reality and can be traced back to Descartes. A PSC must present the cause-effect potential in order to change its own state. A measure that encapsulates this property must be equal to zero for a system composed by independent parts.

- *Intrinsicality* - Experience is subjective.

- *Composition* - Consciousness is compositional (structured), that is, each experience consists of multiple aspects in various combinations.

- *Information* - (or specificity) The perspective is specific and the measure of complexity should be calculated taking into account only one state of the system, the one that carries the most information.

- *Integration* - A PSC must be unified, irreducible to its independent systems.

- *Exclusion* - The causal structure defined by the PSC must be definite, defined with respect to one state, the one that maximizes the irreducibility of the system.

The reader should go to the original references in order to be fair with these authors. However, no logician would accept such set as the desiderata of a formal theory. It sounds more like a vague declaration of what elements an acceptable theory ought to present. It is not possible to extract a unique tractable mathematical theory and hence many possibilities arise when trying to translate this declaration of purpose into something that can be calculated, measured and permits comparison between theory and experiment. The merit lies in presenting the first attempt in history to measure and quantify consciousness. It is certainly an unfinished chapter in the history of science.

1.2.2 Neuroscientific Applications

In 2013 Casali et al. published a paper inspired by Tononi's idea where they tried to define and test a neural correlate of consciousness, a marker that could identify and measure the level of consciousness a patient presents [10]. This measure, namely the Perturbational Complexity Index (PCI), was defined in such a way to assess the presence of integration and differentiation, both concepts also present in the integrated information index definition.

Their method consists in measuring the neural response to a transcranial magnetic stimulation (TSM) and calculating PCI from this spatio-temporal data. In the TMS technique, a coil generates a strong magnetic field near the patient's head, inducing electric currents in the brain. The response of the neurons to this strong magnetic field is then measured in the form of a spatio-temporal matrix. This data matrix is compressed and by calculating the normalized algorithmic complexity (Lempel-Ziv complexity) we have the PCI for that subject.

In addition to defining the PCI, they have tested it in subjects under different conditions to see if this index could be used as a neural correlate of consciousness. Figure 1.1 shows their results for patients subjected to different levels of anesthesia and for different sleep cycles, and the comparison of their PCI with wakeful patients. There is a correlation between their level of consciousness assessed by conventional clinical methods and the PCI, which indicates that quantitative complexity indices can be used to discriminate those patients.

They then turned to the application of PCI for brain-injured individuals. They have compared patients with four types of brain injury diagnosis: vegetative state or unresponsive wakefulness syndrome (VS or UWS), locked-in syndrome (LIS) and two classes in the coma recovery spectrum, the minimally conscious state (MCS) and the emerging from the minimally conscious state (EMCS). The PCI comparison for these subjects can be found in figure 1.2. Again, we notice that there is a correlation between the brain-injured group and their PCI and, by comparing it with the PCI in wakefulness, it might be used to infer if a patient is conscious or not.

These results have motivated us to continue to investigate the complexity index and to better understand what exactly they are measuring. The PCI used by Casali et al. is defined operationally and is hard to use it analytically. For this reason, we turn to other measures defined to present the same interesting properties and that might help us with our goal.

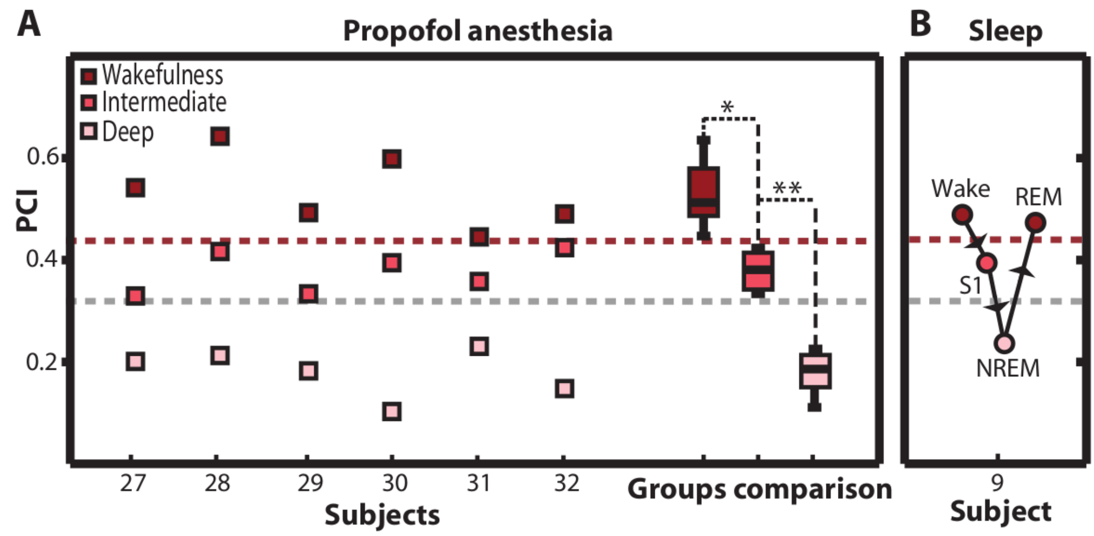


Figure 1.1: Comparison of PCI for subjects in different levels of anesthesia and sleep. Image taken from reference [10].

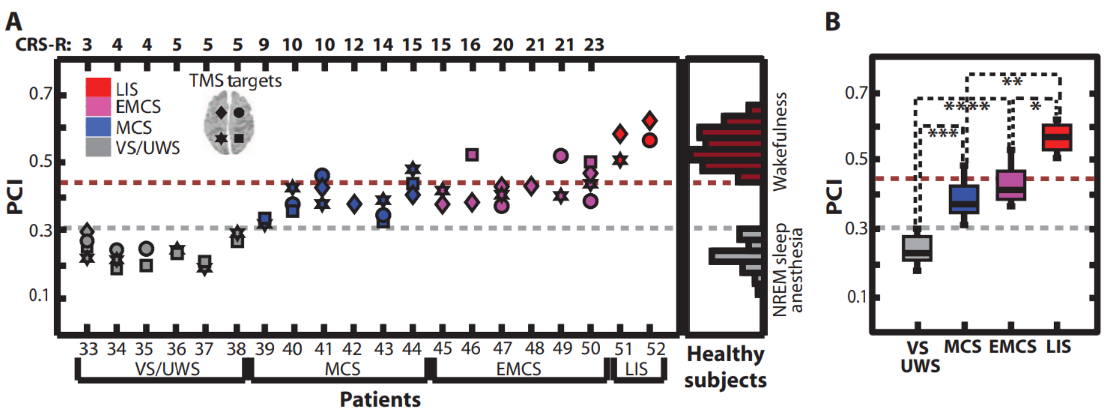


Figure 1.2: Comparison of PCI for different types of brain-injured patients. Image taken from reference [10].

1.2.3 Information Geometric Framework

Another interesting work that originated from Tononi's IIT is the 2016 M. Oizumi et al. paper, where they derived an information geometry based framework to study complexity measures as "distances" between probability manifolds. In all definitions of Φ, a concept that is always present is that it is a measure of how different a system is to a disconnected version of itself, and the information geometry notion of distance (or divergence) seems natural to describe such a measure.

This approach is very appealing because, besides the definition of Φ being conceptually simple, its generalization for other measures is straightforward and the tools of information geometry are developed enough to permit calculations. For those reasons, this framework was chosen for our analysis.

The details of this approach will be further developed in the following chapters.

Chapter 2

Geometric Integrated Information

Information geometry is a field that combines differential geometry with information and probability theory and presents a different way of studying relations between probability distributions. By parametrization of a family of distributions, it is possible to define a manifold and study their properties such as curvature and distance to other distributions [11].

In 2016, M. Oizumi, N. Tsuchyia and S. Amari used the tools available in information geometry to develop a framework to analyse complexity measures for stochastic systems [12]. Their approach is to measure how different the full system is to a disconnected one, thus quantifying the strength of the influences that the disconnected elements exerts on each other. Their calculations were performed on very small systems, with only a very small number of units.

Let's consider a system with N units, indexed on a set $\Lambda \subseteq \mathbb{Z}$, $|\Lambda| = N$. We denote by $X = \{x_i\}_{i\in\Lambda}$ and $Y = \{y_i\}_{i\in\Lambda}$ the system state at two consecutive times of a discrete dynamics. The spatio-temporal interactions between its elements are described by a distribution $P(X,Y)$[1], the full model. We then define $Q(X,Y)$ a distribution in which some of the influences of interest are disconnected. They postulated that the difference between the full model and the disconnected model would be measured by the minimum Kullback-Leibler divergence:

$$\min_Q D_{\mathrm{KL}}[P\|Q] = \min_Q \sum_{X,Y} P(X,Y) \log \frac{P(X,Y)}{Q(X,Y)}, \tag{2.1}$$

and this quantity is a measure of the strength of the influence between the disconnected elements.

By considering different disconnected models, Oizumi et al. were able to derive different information theoretical quantities that can be used as measures of complexity for the system. In figure 2.1 we can see what such disconnections look like and what quantity can be obtained calculating (2.1). The different disconnections arise from performing the variations on a particular manifold. Let Π be a partition of the set Λ into non-overlapping subsets and denote by I and J components of the partition Π. Examples of the probability manifolds that will be considered below are

$$\mathcal{M}_G := \Big\{Q(X,Y)\Big|Q(Y_I|X) = Q(Y_I|X_I), \forall I \in \Pi\Big\}, \tag{2.2}$$

$$\mathcal{M}_S := \left\{Q(X,Y)\middle|Q(Y|X) = \prod_{I\in\Pi} Q(Y_I|X_I)\right\}, \tag{2.3}$$

[1]Conditional information about the details of the model is not shown for notational simplicity.

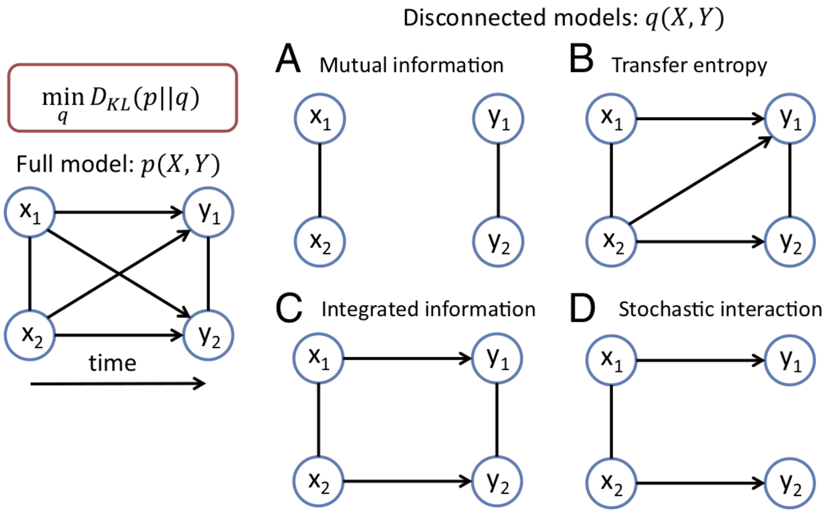


Figure 2.1: Different information theoretic quantities achieved by different disconnected models. X_1 and X_2 (or Y_1 and Y_2) are two disjoint subsets of X (or Y), and represents the disconnected elements. Image taken from reference [12]

where $X_I = \{x_i | i \in I\}$ and $Y_I = \{y_i | i \in I\}$. We do not write explicitly the dependence on the partition.

Integrated information aims to quantify the amount of "synergistic" influences the whole system exerts on its future in excess of what the independent parts of the system do [2]. In the information geometry framework, given a partitioning Π, this can be achieved by considering the following disconnection:

$$Q\left(Y_I|X\right) = Q\left(Y_I|X_I\right), \tag{2.4}$$

for every component $I \in \Pi$.

The geometric integrated information is then defined as

$$\phi = \min_{Q \in \mathcal{M}_G} \sum_{X,Y} P\left(X,Y\right) \log \frac{P\left(X,Y\right)}{Q\left(X,Y\right)}, \tag{2.5}$$

where $\mathcal{M}_G$ is the manifold (2.2) of distributions that satisfy (2.4). This is a complexity index that is relative to the particular partition under consideration.

2.1 Integrated Information for a Bipartition

Consider a partition of the elements of the system into two groups, denoted by I and J, the complement of $I \subset \Lambda$. We define $\mathcal{M}_{IJ}$ the probability manifold whose elements satisfy the integrated information disconnection constraint (2.4) for this bipartion $\Pi = (I, J)$.

Every distribution $Q_{IJ} \in \mathcal{M}_{IJ}$ must decompose as follows:

$$Q_{IJ}\left(X,Y\right) = Q_{IJ}\left(X\right) Q_{IJ}\left(Y_I|X_I\right) Q_{IJ}\left(Y_J|X_J Y_I\right). \tag{2.6}$$

[2]Synergy is an unusual word in the context of phase transitions, and it probably refers to interactions between the system's degrees of freedom and their effect on emergent or collective properties.

To calculate the minimum KL divergence we use the Lagrange multiplier method. Define

$$\mathcal{L} = D_{\mathrm{KL}}\left[P\|Q_{IJ}\right] + \lambda\left(\sum_X Q_{IJ}\left(X\right) - 1\right) + \sum_{X_I}\mu\left(X_I\right)\left(\sum_{Y_I} Q_{IJ}\left(Y_I|X_I\right) - 1\right) + \\ + \sum_{X_J,Y_I}\nu\left(X_J,Y_I\right)\left(\sum_{Y_J} Q_{IJ}\left(Y_J|X_JY_I\right) - 1\right), \quad (2.7)$$

where the Lagrange multipliers enforce the normalization of each factor in (2.6). To obtain the extreme,

$$Q^*_{IJ} = \operatorname{argmin}\mathcal{L}, \quad (2.8)$$

impose the variations

$$\frac{\delta\mathcal{L}}{\delta Q_{IJ}\left(X\right)} = 0, \qquad \frac{\delta\mathcal{L}}{\delta Q_{IJ}\left(Y_I|X_I\right)} = 0, \qquad \frac{\delta\mathcal{L}}{\delta Q_{IJ}\left(Y_J|X_JY_I\right)} = 0, \quad (2.9)$$

leading to conditions that Q^*_{IJ}, the closest one to P, must satisfy:

$$\begin{aligned} Q^*_{IJ}\left(X\right) &= P\left(X\right), & (2.10)\\ Q^*_{IJ}\left(Y_I|X_I\right) &= P\left(Y_I|X_I\right), & (2.11)\\ Q^*_{IJ}\left(Y_J|X_J,Y_I\right) &= P\left(Y_J|X_J,Y_I\right). & (2.12)\end{aligned}$$

Using these conditions, we finally derive an expression for the integrated information for a bipartition (I,J):

$$\phi_{IJ} = \sum_{X,Y} P\left(X,Y\right)\log\frac{P\left(Y|X\right)P\left(Y_I|X_J\right)}{P\left(Y|X_J\right)P\left(Y_I|X_I\right)}. \quad (2.13)$$

Suppose, as we will show for a specific model in the next chapters, that ϕ_{IJ} depends only on the fraction of units in one of the components; $\phi_{IJ} = \phi(\gamma)$, with $\gamma = |I|/|\Lambda|$, then a candidate for the integrated information may be the average value

$$\Phi = \langle\phi(\gamma)\rangle_\gamma, \quad (2.14)$$

where the probability of γ depends on the particular problem under consideration. We will mainly concentrate on the properties of $\phi(\gamma)$ in the next chapters. The problem of finding an appropriate partition, already present in [13] and central in [12], or the distribution of partitions, $P(\gamma)$, remains unsolved in general.

Chapter 3

The Little-Sherrington-Kirkpatrick Model

The model chosen for our analysis of the integrated information is what we called the Little-Sherrington-Kirkpatrick (LSK) model. The model is a version of W. Little's attractor synchronous neural network [5] with the gaussian quenched disorder introduced by D. Sherrington and S. Kirkpatrick on their study of the infinite range Edwards and Anderson spin glass model [14]. The relation of the LSK to the SK model is analogous to that of the Little model of an attractor neural network to the Hopfield model [15], studied using the tools of statistical mechanics by Peretto [16] and Fontanari and Köberle [17].

The main reason for the choice of this model is that, besides the fact it has some interesting features and complex behavior, it is formulated with an intrinsic dynamics and defined through the interaction of two consecutive times, which fits perfectly the information geometry description of the integrated information index.

3.1 The Model Definition

The system's state is described by a set of N binary variables that takes values in $\{-1, 1\}$. Lets denote by $X = \{x_i\}_{i=1}^N$ and $Y = \{y_i\}_{i=1}^N$ the state of the system at two consecutive times, then, the interaction between the variables is described by the Hamiltonian

$$\mathcal{H}\left(X, Y|\boldsymbol{J}\right) = -\sum_{i,j} J_{ij} x_j y_i, \tag{3.1}$$

where $\boldsymbol{J} = \{J_{ij}\}$ is a set of quenched numbers that can take positive or negative values, and each J_{ij} describes the interaction between the i-th and j-th elements: if it is positive, the energy with the pair will be lower when x_j and y_i have the same sign and greater if their signs are opposite.

Imposing the constraint that $\mathcal{H}\left(X, Y|\boldsymbol{J}\right)$ has a fixed mean value E (canonical ensemble), the maximum entropy principle (see [18]) gives as the equilibrium probability distribution the Gibbs-Boltzmann measure

$$P\left(X, Y|\boldsymbol{J}\right) = \frac{1}{Z} \exp\left\{\beta \sum_{i,j} J_{ij} x_j y_i\right\}, \tag{3.2}$$

where β is a Lagrange multiplier that ensures the above constraint and can be identified as the inverse temperature for the system $\beta = 1/T$, and Z is the partition function, that ensures normalization.

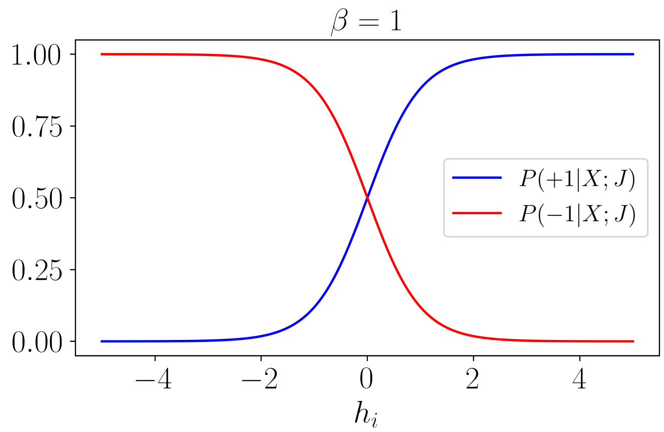


Figure 3.1: Transition probability function for the LSK model, with $\beta = 1$

Calculating the transition probability distribution

$$P\left(Y|X;\boldsymbol{J}\right) = \frac{P\left(X,Y|\boldsymbol{J}\right)}{P\left(X|\boldsymbol{J}\right)} = \frac{P\left(X,Y|\boldsymbol{J}\right)}{\sum_Y P\left(X,Y|\boldsymbol{J}\right)} = \frac{\exp\left\{\beta\sum_{i,j} J_{ij}x_jy_i\right\}}{\prod_i 2\cosh\left(\beta\sum_j J_{ij}x_j\right)}, \tag{3.3}$$

we recover the same distribution introduced by W. Little [5] and studied by P. Peretto [16].

Note that this probability is a product of the transition distributions for each element,

$$P\left(Y|X;\boldsymbol{J}\right) = \prod_{i=1}^{N} \frac{\exp\left\{\beta h_i y_i\right\}}{2\cosh\left(\beta h_i\right)} = \prod_{i=1}^{N} P\left(y_i|X;\boldsymbol{J}\right), \tag{3.4}$$

where we define the field generated by the system on the element i as

$$h_i = \sum_{j=1}^{N} J_{ij}x_j. \tag{3.5}$$

That is, conditioned on the past, the elements evolve independently of each other, each of then following a logistic function of h_i, represented on figure 3.1.

3.2 Free Energy

Until now, we have not said anything about the interactions $\boldsymbol{J}$. We are not interested in a specific configuration of $\boldsymbol{J}$, but in a disordered system, where the interactions are random variables drawn, independently from anything else, from a distribution $P\left(\boldsymbol{J}\right)$.

These new random variables may evolve in the same time scale as the state variables, or they can evolve on a very slow time scale, which can be taken to be infinitely slow, then these are fixed random variables during the dynamics of the system. Those two types of disorder are called annealed and quenched, respectively.

In this work, we focus on the system with quenched disorder. To find an expression for the free energy, we consider the following constraints for the joint probability distribution

$P(X, Y, \boldsymbol{J})$:

$$\sum_{X,Y} P(X, Y|\boldsymbol{J}) = 1, \tag{3.6}$$

$$\sum_{X,Y} P(X, Y, \boldsymbol{J}) = P(\boldsymbol{J}), \tag{3.7}$$

$$\sum_{X,Y} \mathcal{H}(X, Y|\boldsymbol{J}) P(X, Y|\boldsymbol{J}) = E. \tag{3.8}$$

With these constraints, the Shannon entropy for the joint probability distribution is

$$S[P(X, Y, \boldsymbol{J})] = \beta E + \langle \log Z \rangle_{\boldsymbol{J}} + S[P(\boldsymbol{J})], \tag{3.9}$$

so, as E and $S[P(\boldsymbol{J})]$ are constants, the maximum entropy is obtained minimizing the free energy

$$F = -\frac{1}{\beta} \langle \log Z \rangle_{\boldsymbol{J}}. \tag{3.10}$$

The form of the disorder we consider in this work is the one studied by D. Sherrington and S. Kirkpatrick [14], where all J_{ij} are independent of each other and all of them distributed accordingly to

$$P(J_{ij}) = \sqrt{\frac{N}{2\pi J^2}} \exp\left\{ -\frac{N}{2J^2} \left(J_{ij} - \frac{J_0}{N} \right)^2 \right\}, \tag{3.11}$$

a normal distribution with mean J_0/N and variance J^2/N. Such a disorder introduce positive and negative interactions between the elements of the system. This will lead to frustration and will allow the appearance of a spin glass phase, under certain circumstance.

To calculate $\langle \log Z \rangle_{\boldsymbol{J}}$ we use the replica method: considering the identity

$$\log Z = \lim_{\delta \to 0} \frac{Z^\delta - 1}{\delta}, \tag{3.12}$$

we calculate Z^δ for integer δ and assume that the result holds for real δ in the limit $\delta \to 0$. We can interpret each term on this product of partition functions as a replica of the system, thus the name of the method.

Our work will be to calculate

$$\left\langle Z^\delta \right\rangle_{\boldsymbol{J}} = \sum_{\{X^a Y^a\}} \left\langle \exp\left\{ \beta \sum_{a=1}^{\delta} \sum_{i,j} J_{ij} x_j^a y_i^a \right\} \right\rangle_{\boldsymbol{J}}, \tag{3.13}$$

with $P(\boldsymbol{J})$ given by equation (3.11).

Although we do not have a Hamiltonian with explicit replica interaction, when we perform the average over the disorder, terms involving $x_i^a x_i^b$ and $y_i^a y_i^b$ will appear. So we define the following quantities that will be used as order parameters for the system:

$$m_a = \langle x_i^a \rangle_{\text{rep}}, \tag{3.14}$$

$$n_a = \langle y_i^a \rangle_{\text{rep}}, \tag{3.15}$$

$$q_{ab} = \left\langle x_i^a x_i^b \right\rangle_{\text{rep}}, \tag{3.16}$$

$$r_{ab} = \left\langle y_i^a y_i^b \right\rangle_{\text{rep}}, \tag{3.17}$$

where $\langle\cdot\rangle_{\text{rep}}$ denotes the average over the canonical distribution associated with the replica interaction Hamiltonian

$$\mathcal{H}_{\text{rep}} = \beta J_0 \sum_a n_a x^a + \beta^2 J^2 \sum_{a<b} r_{ab} x^a x^b + \beta J_0 \sum_a m_a y^a + \beta^2 J^2 \sum_{a<b} q_{ab} y^a y^b. \tag{3.18}$$

The quantities m and n are the magnetization at the two consecutive times, and q and r the replica overlaps.

Considering the Replica Symmetric ansatz, where the above parameters are symmetric under replica index permutation, we calculate the free energy $f = F/N$ in the thermodynamic limit, $N \to \infty$, in terms of the parameters m, n, q and r:

$$\begin{aligned} f_{\text{RS}} = J_0 mn - \frac{\beta J^2}{2}(1-q)(1-r) - \frac{1}{\beta}\int Dz \log 2\cosh\left[\beta\left(J_0 n + J\sqrt{r}z\right)\right] + \\ -\frac{1}{\beta}\int Dz \log 2\cosh\left[\beta\left(J_0 m + J\sqrt{q}z\right)\right], \end{aligned} \tag{3.19}$$

where

$$Dz = \frac{e^{-\frac{1}{2}z^2}}{\sqrt{2\pi}} dz, \tag{3.20}$$

is the Gaussian measure with zero mean and unit variance.

Details of the calculation can be found in appendix A.

3.3 Equations of State

The equilibrium states for the model are those that minimize the free energy. Setting the derivative of f_{RS} with respect to m, n, q and r equal to zero, we find

$$n = \int Dz \tanh\left[\beta\left(J_0 m + J\sqrt{q}z\right)\right], \tag{3.21}$$

$$r = \int Dz \tanh^2\left[\beta\left(J_0 m + J\sqrt{q}z\right)\right], \tag{3.22}$$

$$m = \int Dz \tanh\left[\beta\left(J_0 n + J\sqrt{r}z\right)\right], \tag{3.23}$$

$$q = \int Dz \tanh^2\left[\beta\left(J_0 n + J\sqrt{r}z\right)\right]. \tag{3.24}$$

In equilibrium, as we expected, we have $m = n$ and $q = r$ (the magnetization and replica overlap do not change). In this case, the self-consistency equations above are exactly the same as those of the standard Sherrington-Kirkpatrick model.

Solving numerically the equations (3.21) to (3.24) we get the phase diagram on figure 3.2.

At high temperatures, with $T > J_0$ the solution to the state equations is $m = n = 0$ and $q = r = 0$, a paramagnetic phase. For lower temperatures with $T < J_0$ the state equations give us a solution with $m = n > 0$ and $q = r > 0$, a ferromagnetic phase. And finally, for low T and low J_0 we have a solution with $m = n = 0$ and $q = r > 0$, a spin glass phase where the elements are randomly frozen.

The validity of the solution we get in the replica symmetric ansatz was studied by J. de Almeida and D. Thouless [19]. They found that the result we get is unstable in the spin glass phase and the low temperature part of the ferromagnetic phase.

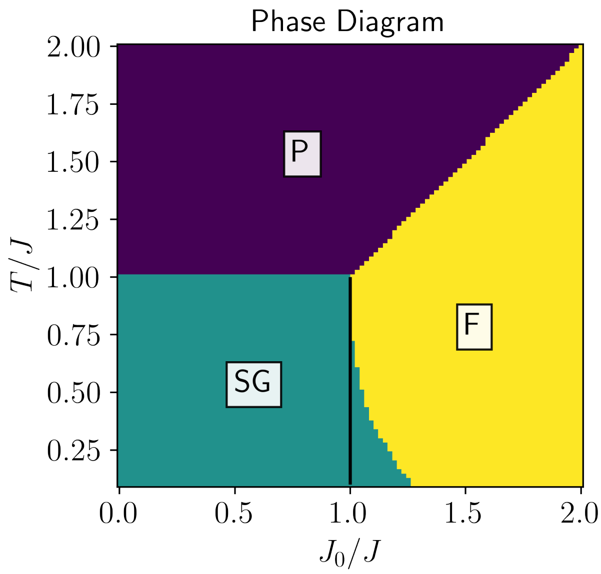


Figure 3.2: Phase diagram for the LSK model: P denotes the Paramagnetic phase, F denotes the Ferromagnetic phase and SG the Spin Glass phase.

An exact solution was found by G. Parisi in 1979 [3], where he considered a specific form for the matrix q_{ab} and find a phase transition boundary between the spin glass and ferromagnetic phase at $J_0 = 1$ (represented by the black line in figure 3.2)

We see that for most of the phase diagram, the replica symmetric approximation gives a good solution, so in this work, for a first analysis, we stick to the RS ansatz bearing in mind that for low temperatures and close to the SG-F phase boundary our solution is wrong.

3.4 Near Equilibrium Dynamics

The dynamics of such a system is very complex and involves too many variables to describe. However, in the equilibrium, the system is well described by the quantities defined in the equations (3.14) to (3.17) and, by the definition of the variables $\{x_i\}$ and $\{y_i\}$, we have a dynamic relation between the order parameters: n_a and r_{ab} are the future of the variables m_a and q_{ab}, respectively. With that in mind, we might look at the equations of state as a discrete dynamical system for the state vector (m, q). This system would be exact only in the equilibrium, but we expect that it may be a good approximation for its dynamics in the vicinity of the equilibrium manifold.

With this idea in mind we look at trajectories for a grid of initial conditions in the $m \times q$ plane for the three different phases. These images can be found in the figure 3.3. The right and left column represents the same dynamics at the same conditions, but with a different coloring: on the left column, the earlier in the iteration, the darker the arrows are, and on the right column, we have darker arrows for points later in the iteration, in such a way that we can better visualize the whole dynamics.

Although our approximation is only good near the equilibrium point, we look for trajectories with initial conditions all over the $m \times q$ plane. There might be some interesting behavior and we do not lose information doing this.

In the paramagnetic and in the spin glass phases, we see that the points quickly shrink

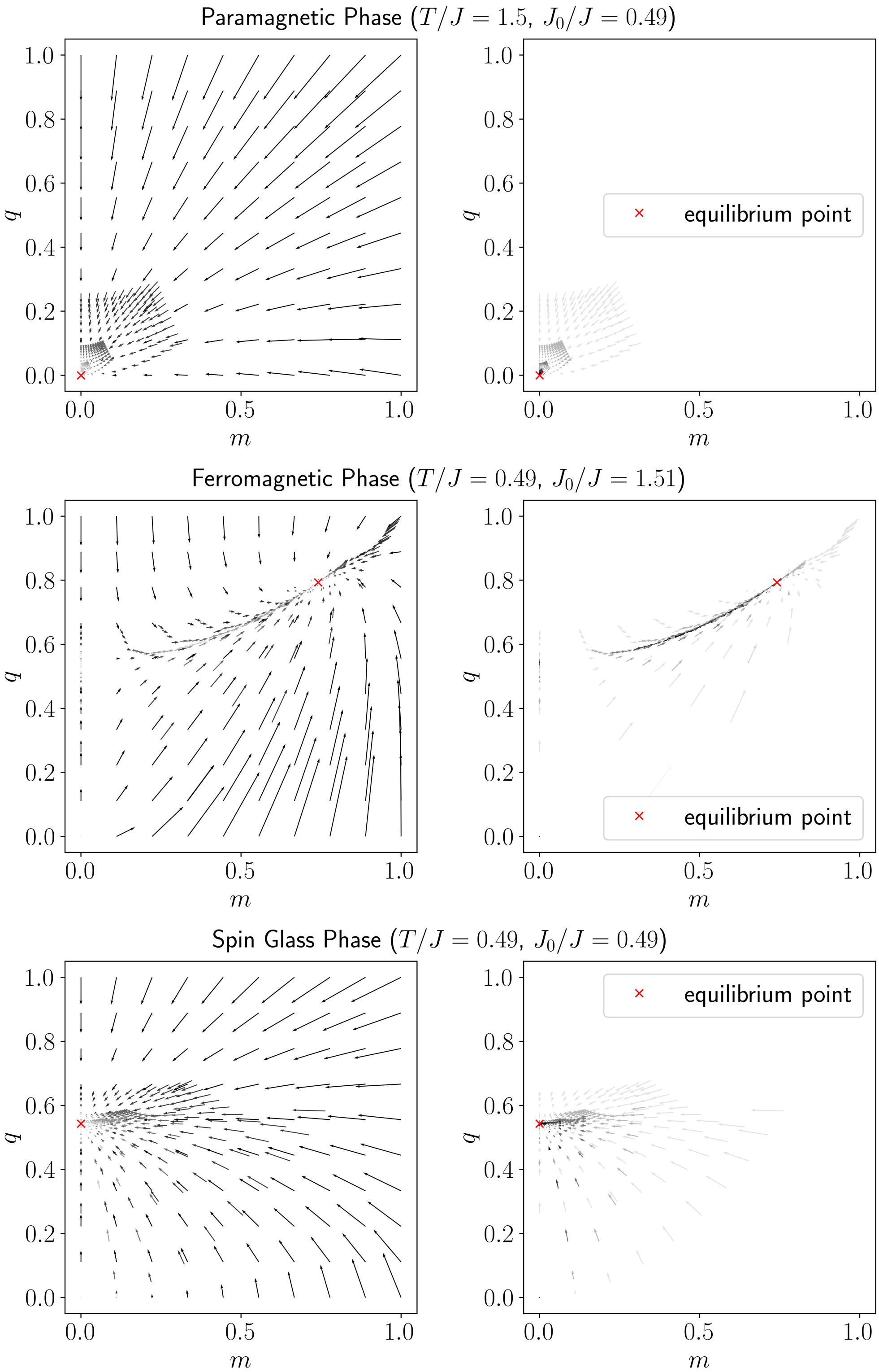


Figure 3.3: Comparison of the dynamics for a grid of initial conditions in the three phases of the LSK model. The two columns present the same data. On the left as the iteration index grows the gray level representation becomes lighter, and on the right, darker.

towards the equilibrium point (marked as a red cross in the graphics). However, in the ferromagnetic phase the points seems to quickly go to a slow manifold, where we have a dynamics that slowly carries them to the equilibrium. This qualitative discrepancy might yield interesting results when we look at the integrated information index for this approximated dynamical system.

Chapter 4

Calculation of ϕ for the LSK Model

Now we calculate the integrated information index for the Little-Sherrington-Kirkpatrick model. We note that our formula for ϕ_{IJ}, equation (2.13), does not take into account the disorder of the LSK model. So we first need a slight modification in our formula.

4.1 Integrated Information for a Disordered System

To take into account the disorder variables $\boldsymbol{J}$ we need to consider the joint probability, $P(X, Y, \boldsymbol{J})$. In this case the integrated information is defined as

$$\phi = \min_{Q \in \mathcal{M}_G} \sum_{X,Y} \int d\boldsymbol{J} P(X, Y, \boldsymbol{J}) \log \frac{P(X, Y, \boldsymbol{J})}{Q(X, Y, \boldsymbol{J})}, \tag{4.1}$$

where $\mathcal{M}_G$ is a manifold that satisfy the same disconnection constraint

$$Q(Y_I | X; \boldsymbol{J}) = Q(Y_I | X_I; \boldsymbol{J}), \tag{4.2}$$

Therefore, for a bipartition (I, J) of the system, a typical distribution $Q_{IJ} \in \mathcal{M}_G$ is decomposed as

$$Q_{IJ}(X, Y, \boldsymbol{J}) = Q_{IJ}(\boldsymbol{J}) Q_{IJ}(X|\boldsymbol{J}) Q_{IJ}(Y_I | X_I; \boldsymbol{J}) Q_{IJ}(Y_J | X_J, Y_I; \boldsymbol{J}). \tag{4.3}$$

Defining the Langrange function

$$\begin{aligned}\mathcal{L} = D_{\mathrm{KL}}[P \| Q_{IJ}] + \alpha \left(\int Q_{IJ}(\boldsymbol{J}) d\boldsymbol{J} - 1 \right) + \int \lambda(\boldsymbol{J}) \left(\sum_{X,Y} Q_{IJ}(X, Y | \boldsymbol{J}) - 1 \right) d\boldsymbol{J} + \\ + \sum_{X_I} \int \mu(X_I, \boldsymbol{J}) \left(\sum_{Y_I} Q_{IJ}(Y_I | X_I; \boldsymbol{J}) - 1 \right) d\boldsymbol{J} + \\ + \sum_{X_J, Y_I} \int \nu(X_J, Y_I, \boldsymbol{J}) \left(\sum_{Y_J} Q_{IJ}(Y_J | X_J, Y_I; \boldsymbol{J}) - 1 \right) d\boldsymbol{J}.\end{aligned} \tag{4.4}$$

Now, as before, we impose that the variations of $\mathcal{L}$ with respect to each factor of $Q_{IJ}(X, Y, \boldsymbol{J})$ are zero. The variations of $Q_{IJ}(X, Y|\boldsymbol{J})$, $Q_{IJ}(Y_I|X_I; \boldsymbol{J})$ and $Q_{IJ}(Y_J|X_J, Y_I; \boldsymbol{J})$

yield similar conditions as before:

$$
\begin{aligned}
Q^*_{IJ}\left(X|\boldsymbol{J}\right) &= P\left(X|\boldsymbol{J}\right); &(4.5)\\
Q^*_{IJ}\left(Y_I|X_I;\boldsymbol{J}\right) &= P\left(Y_I|X_I;\boldsymbol{J}\right); &(4.6)\\
Q^*_{IJ}\left(Y_J|X_J,Y_I;\boldsymbol{J}\right) &= P\left(Y_J|X_J,Y_I;\boldsymbol{J}\right). &(4.7)
\end{aligned}
$$

But, when we make the variations with respect to $Q_{IJ}\left(\boldsymbol{J}\right)$ we obtain a new condition that the closest one to P must satisfy:

$$Q^*_{IJ}\left(\boldsymbol{J}\right) = P\left(\boldsymbol{J}\right). \tag{4.8}$$

Now, using these conditions, we have a formula for the integrated information (with respect to a bipartition) for a disordered system

$$\phi_{IJ} = \left\langle \sum_{X,Y} P\left(X,Y|\boldsymbol{J}\right)\log\frac{P\left(Y|X;\boldsymbol{J}\right)P\left(Y_I|X_J;\boldsymbol{J}\right)}{P\left(Y|X_J;\boldsymbol{J}\right)P\left(Y_I|X_I;\boldsymbol{J}\right)}\right\rangle_{\boldsymbol{J}}. \tag{4.9}$$

Before we begin the calculation, there is only one detail left: we have only the transition probability for the full LSK model $P\left(Y|X;\boldsymbol{J}\right)$, but not disconnected transition probabilities distributions. So first we need a way to take into account the partitioning and write those probabilities.

4.2 Implementing Disconnections by Dilutions

To mimic the disconnections we consider a system with site dilution. We introduce a new set of variables $\boldsymbol{\eta} = \{\eta_i\}_{i=1}^N$, with each η_i taking values 1 or 0, representing that the i-th element belongs to the partition I or J, respectively.

Note that each configuration $\boldsymbol{\eta}$ corresponds to one and only one partitioning (I,J) (considering that $(I,J)\neq(J,I)$) and that the number of both configurations $\boldsymbol{\eta}$ and partitions is 2^N, implying a one to one correspondence. Additionally, by calculating expected values with respect to a probability measure, $P(\boldsymbol{\eta})$, we are able to average over all possible partitions, which can be interesting in our analysis.

Therefore, modifying the interaction matrix with the new variables, it is possible to select only the interactions we are interested in. For example, if we consider the new interaction matrix $J_{ij}\eta_i(1-\eta_j)$, we immediately see that only when $i\in I$ and $j\in J$ the interaction will be non zero. Thus, the Hamiltonian, in this case, is

$$\mathcal{H}\left(X,Y|\boldsymbol{J},\boldsymbol{\eta}\right) = -\sum_{i,j} J_{ij}\eta_i(1-\eta_j)x_jy_i = \mathcal{H}\left(X_J,Y_I|\boldsymbol{J},\boldsymbol{\eta}\right), \tag{4.10}$$

only the interaction between $x_j\in X_J$ and $y_i\in Y_I$ is been considered.

In general, we use the following prescription to account for any interaction between the elements of the two partitions we are interested in:

$$J_{ij} \to \begin{cases} J_{ij}\eta_i\eta_j & i,j \in I \\ J_{ij}\left(1-\eta_i\right)\left(1-\eta_j\right) & i,j \in J \\ J_{ij}\left(1-\eta_i\right)\eta_j & i \in J, \quad j \in I \\ J_{ij}\eta_i\left(1-\eta_j\right) & i \in I, \quad j \in J \\ J_{ij}\left(1-\eta_j\right) & i \in I \cup J, \quad j \in J \\ J_{ij}\eta_j & i \in I \cup J, \quad j \in I \\ J_{ij}\left(1-\eta_i\right) & i \in J, \quad j \in I \cup J \\ J_{ij}\eta_i & i \in I, \quad j \in I \cup J \end{cases} \tag{4.11}$$

Now, we write all the transition probabilities present in equation (4.9):

$$P\left(Y|X;\boldsymbol{J},\boldsymbol{\eta}\right) = \frac{\exp\left\{\beta\sum_{i,j}J_{ij}x_jy_i\right\}}{\prod_i 2\cosh\left(\beta\sum_j J_{ij}x_j\right)}, \tag{4.12}$$

$$P\left(Y_I|X_I;\boldsymbol{J},\boldsymbol{\eta}\right) = \frac{\exp\left\{\beta\sum_{i,j}J_{ij}\eta_i\eta_jx_jy_i\right\}}{\prod_i 2\cosh\left(\beta\sum_j J_{ij}\eta_i\eta_jx_j\right)}, \tag{4.13}$$

$$P\left(Y_I|X_J;\boldsymbol{J},\boldsymbol{\eta}\right) = \frac{\exp\left\{\beta\sum_{i,j}J_{ij}\eta_i\left(1-\eta_j\right)x_jy_i\right\}}{\prod_i 2\cosh\left(\beta\sum_j J_{ij}\eta_i\left(1-\eta_j\right)x_j\right)}, \tag{4.14}$$

$$P\left(Y|X_J;\boldsymbol{J},\boldsymbol{\eta}\right) = \frac{\exp\left\{\beta\sum_{i,j}J_{ij}\left(1-\eta_j\right)x_jy_i\right\}}{\prod_i 2\cosh\left(\beta\sum_j J_{ij}\left(1-\eta_j\right)x_j\right)}. \tag{4.15}$$

Finally we are ready for the calculation of $\phi_{IJ} = \phi_{\boldsymbol{\eta}}$.

4.3 Calculation of $\phi_{\boldsymbol{\eta}}$

With the distributions (4.12) to (4.15) the integrated information for the LSK model is

$$\begin{aligned}\phi_{\boldsymbol{\eta}} = \Bigg\langle \sum_{X,Y} P\left(X,Y|\boldsymbol{J}\right)\Bigg[\left(\beta\sum_{i=1}^{N}y_ih_i - \sum_{i=1}^{N}\log 2\cosh(\beta h_i)\right) + \\ - \left(\beta\sum_{i=1}^{N}y_ih_{i|S} - \sum_{i=1}^{N}\log 2\cosh(\beta h_{i|S})\right)\Bigg]\Bigg\rangle_{\boldsymbol{J}},\end{aligned} \tag{4.16}$$

where we introduce the quantity $h_{i|S}$ and its complementary $h_{i|D}$, defined as

$$h_{i|S} = \begin{cases} \sum_j J_{ij}\eta_jx_j, & \text{for} \quad i \in I \\ \sum_j J_{ij}\left(1-\eta_j\right)x_j, & \text{for} \quad i \in J \end{cases} \tag{4.17}$$

$$h_{i|D} = \begin{cases} \sum_j J_{ij} \left(1-\eta_j\right) x_j, & \text{for} \quad i \in I \\ \\ \sum_j J_{ij} \eta_j x_j, & \text{for} \quad i \in J \end{cases} \tag{4.18}$$

the fields generated on i by the elements in the *same* partition and in the *different* partition, respectively. We immediately note that the total field is $h_i = h_{i|S} + h_{i|D}$.

To proceed with the calculation, we rewrite this expression as

$$\phi_{\boldsymbol{\eta}} = \left\langle \sum_{X,Y} P\left(X,Y|\boldsymbol{J}\right) \left(\beta \sum_{i=1}^{N} y_i h_{i|D} - \sum_{i=1}^{N} \log 2\cosh(\beta h_i) + \sum_{i=1}^{N} \log 2\cosh(\beta h_{i|S}) \right) \right\rangle_{\boldsymbol{J}}, \tag{4.19}$$

and consider the three terms separately:

$$\begin{aligned} \phi_A &= \beta \left\langle \sum_{X,Y} P\left(X,Y|\boldsymbol{J}\right) \sum_{i=1}^{N} y_i h_{i|D} \right\rangle_{\boldsymbol{J}}, & (4.20) \\ \phi_B &= -\left\langle \sum_{X,Y} P\left(X,Y|\boldsymbol{J}\right) \sum_{i=1}^{N} \log 2\cosh(\beta h_i) \right\rangle_{\boldsymbol{J}}, & (4.21) \\ \phi_C &= \left\langle \sum_{X,Y} P\left(X,Y|\boldsymbol{J}\right) \sum_{i=1}^{N} \log 2\cosh(\beta h_{i|S}) \right\rangle_{\boldsymbol{J}}, & (4.22) \end{aligned}$$

such that $\phi_{\boldsymbol{\eta}} = \phi_A + \phi_B + \phi_C$.

To calculate each of these terms, we consider auxiliary distributions tailored in such a way that the above averages can be extracted from their partition functions (see appendix B for more details).

The result we got for each of those terms are:

$$\phi_A = 2\gamma\left(1-\gamma\right) N \left[\beta J_0 m n + \beta^2 J^2 (1-qr)\right], \tag{4.23}$$

$$\begin{aligned} \phi_B = &- N \int Dz \sum_{y\in\{-1,1\}} \frac{\exp\left\{\beta\left(J_0 m + J\sqrt{q}z\right) y\right\}}{2\cosh\left[\beta\left(J_0 m + J\sqrt{q}z\right)\right]} \times \\ &\times \int D\zeta \log 2\cosh\left[\beta\left(J_0 m + J\sqrt{q}z\right) + \beta^2 J^2 \left(1-q\right) y + \beta J \zeta \sqrt{1-q}\right], \end{aligned} \tag{4.24}$$

$$\begin{aligned} \phi_C = &\gamma N \int Dz_S \int Dz_D \sum_{y_i\in\{-1,1\}} \frac{\exp\left\{\beta\left(J_0 m + J z_D \sqrt{(1-\gamma)q} + J z_S \sqrt{\gamma q}\right) y_i\right\}}{2\cosh\left[\beta\left(J_0 m + J z_S \sqrt{\gamma q} + J z_D \sqrt{(1-\gamma)q}\right)\right]} \times \\ &\times \int D\zeta \log 2\cosh\left[\beta\left(J_0 \gamma m + J z_S \sqrt{\gamma q}\right) + \beta^2 J^2 \gamma (1-q) y_i + \beta J \sqrt{\gamma(1-q)}\zeta\right] + \\ &+ (1-\gamma) N \int Dz_S \int Dz_D \sum_{y_i\in\{-1,1\}} \frac{\exp\left\{\beta\left(J_0 m + J z_D \sqrt{\gamma q} + J z_S \sqrt{(1-\gamma)q}\right) y_i\right\}}{2\cosh\left[\beta\left(J_0 m + J z_S \sqrt{(1-\gamma)q} + J z_D \sqrt{\gamma q}\right)\right]} \times \\ &\times \int D\zeta \log 2\cosh\left[\beta\left(J_0 (1-\gamma) m + J z_S \sqrt{(1-\gamma)q}\right) + \beta^2 J^2 (1-\gamma)(1-q) y_i + \beta J \sqrt{(1-\gamma)(1-q)}\zeta\right]. \end{aligned} \tag{4.25}$$

In all of the above equation,

$$Dz = \frac{e^{-z^2/2}}{\sqrt{2\pi}} dz, \tag{4.26}$$

denotes the Gaussian measure with zero mean and unit variance, and $\gamma = \sum_i \eta_i / N$, the fraction of the elements that belongs to the component I.

Chapter 5

Results Analysis

In the last chapter we calculated the integrated information for the LSK model in terms of its order parameters m, q, n and r:

$$
\begin{aligned}
\frac{\phi_{\boldsymbol{\eta}}}{N} =& 2\gamma\left(1-\gamma\right)\left[\beta J_0 mn + \beta^2 J^2 (1-qr)\right] + \\
& - \int Dz \sum_{y\in\{-1,1\}} \frac{\exp\left\{\beta\left(J_0 m + J\sqrt{q}z\right)y\right\}}{2\cosh\left[\beta\left(J_0 m + J\sqrt{q}z\right)\right]} \times \\
& \times \int D\zeta \log 2\cosh\left[\beta\left(J_0 m + J\sqrt{q}z\right) + \beta^2 J^2\left(1-q\right)y + \beta J\zeta\sqrt{1-q}\right] \\
& + \gamma \int Dz_S \int Dz_D \sum_{y\in\{-1,1\}} \frac{\exp\left\{\beta\left(J_0 m + Jz_D\sqrt{(1-\gamma)q} + Jz_S\sqrt{\gamma q}\right)y\right\}}{2\cosh\left[\beta\left(J_0 m + Jz_S\sqrt{\gamma q} + Jz_D\sqrt{(1-\gamma)q}\right)\right]} \times \\
& \times \int D\zeta \log 2\cosh\left[\beta\left(J_0\gamma m + Jz_S\sqrt{\gamma q}\right) + \beta^2 J^2 \gamma(1-q)y + \beta J\sqrt{\gamma(1-q)}\zeta\right] + \\
& + (1-\gamma) \int Dz_S \int Dz_D \sum_{y\in\{-1,1\}} \frac{\exp\left\{\beta\left(J_0 m + Jz_D\sqrt{\gamma q} + Jz_S\sqrt{(1-\gamma)q}\right)y\right\}}{2\cosh\left[\beta\left(J_0 m + Jz_S\sqrt{(1-\gamma)q} + Jz_D\sqrt{\gamma q}\right)\right]} \times \\
& \times \int D\zeta \log 2\cosh\Big[\beta\left(J_0(1-\gamma)m + Jz_S\sqrt{(1-\gamma)q}\right) + \\
& + \beta^2 J^2 (1-\gamma)(1-q)y + \beta J\sqrt{(1-\gamma)(1-q)}\zeta\Big].
\end{aligned} \tag{5.1}
$$

5.1 The γ dependence

The first thing we notice is that the dependence in the partition is only through γ, the size of the partition: $\phi_{\boldsymbol{\eta}} = \phi(\gamma)$. This result was expected, since all the elements are symmetric and nothing differentiates between then. Another obvious fact that we confirm with our calculation is that $\phi(\gamma)$ is symmetric around $\gamma = 1/2$, this was also already expected because it should not matter which partition we name I or J.

The first term, ϕ_A, can be easily interpreted as it is proportional to the average energy for the LSK model

$$\frac{E}{N} = -J_0 mn - \beta J^2(1 - qr). \tag{5.2}$$

It is also proportional to $2\gamma(1-\gamma)N$ which is the number of interactions (per number of elements, N) between elements from the partition I with elements from the partition J.

The rest of our expression is not that clear, and to better understand it we calculate it numerically.

We start by choosing a typical point $(J_0/J, T/J)$ in each of the three phases and plotting a graphic of $\phi \times \gamma$, as seen in the left column of figure 5.1. We notice that the general behavior is very similar for the three phases: if we have a trivial partition ($\gamma = 0$ or $\gamma = 1$) all elements will belong to one partition and there will be no distinction between the full model and the disconnected one. Additionally, the maximum happens when $\gamma = 1/2$ and we have the maximum number of disconnected interactions.

In the previous paragraphs we also discussed the parabolic dependence on γ of the first term ϕ_A. We can extend this discussion asking how much the rest of our expression deviates from a parabola. This analysis can be found in the right column of figure 5.1 for each phase, where we compare our result (in blue) with a parabola (in orange) centered in $\gamma = 1/2$, with roots 0 and 1 and with the same maximum value.

We notice that in the paramagnetic phase our result matches very well with a parabola, while in the spin-glass phase we have a little difference and in the ferromagnetic phase we have an even larger discrepancy. This behavior can be explained when we consider the fact that in the paramagnetic phase the average energy for the system is larger than the other phases and the term that depends on the energy (and has a parabolic dependence on γ) dominates. At the same time, the ferromagnetic phase is the phase with less average energy and the term that depends on the energy is less dominant.

5.2 Map of ϕ at equilibrium

Now that we know the general behavior of ϕ when we change the size partition γ, we proceed to analyse how ϕ varies across the phase diagram for a given γ.

We begin choosing $\gamma = 1/2$ and calculating $\phi(\gamma = 1/2)$ for all points of our phase diagram. The result can be found in figure 5.2.

The range of values we used for the control parameters were chosen in order to avoid values of ϕ that does not make sense: We know that the replica symmetric solution is unstable below the de Almeida-Thouless line and yield nonsensical solutions, for example negative entropy for low temperatures. In the following sections we will see that ϕ is closely related with the entropy of the system and this negative entropy solution will also result in negative ϕ, which is by definition absurd (ϕ is a

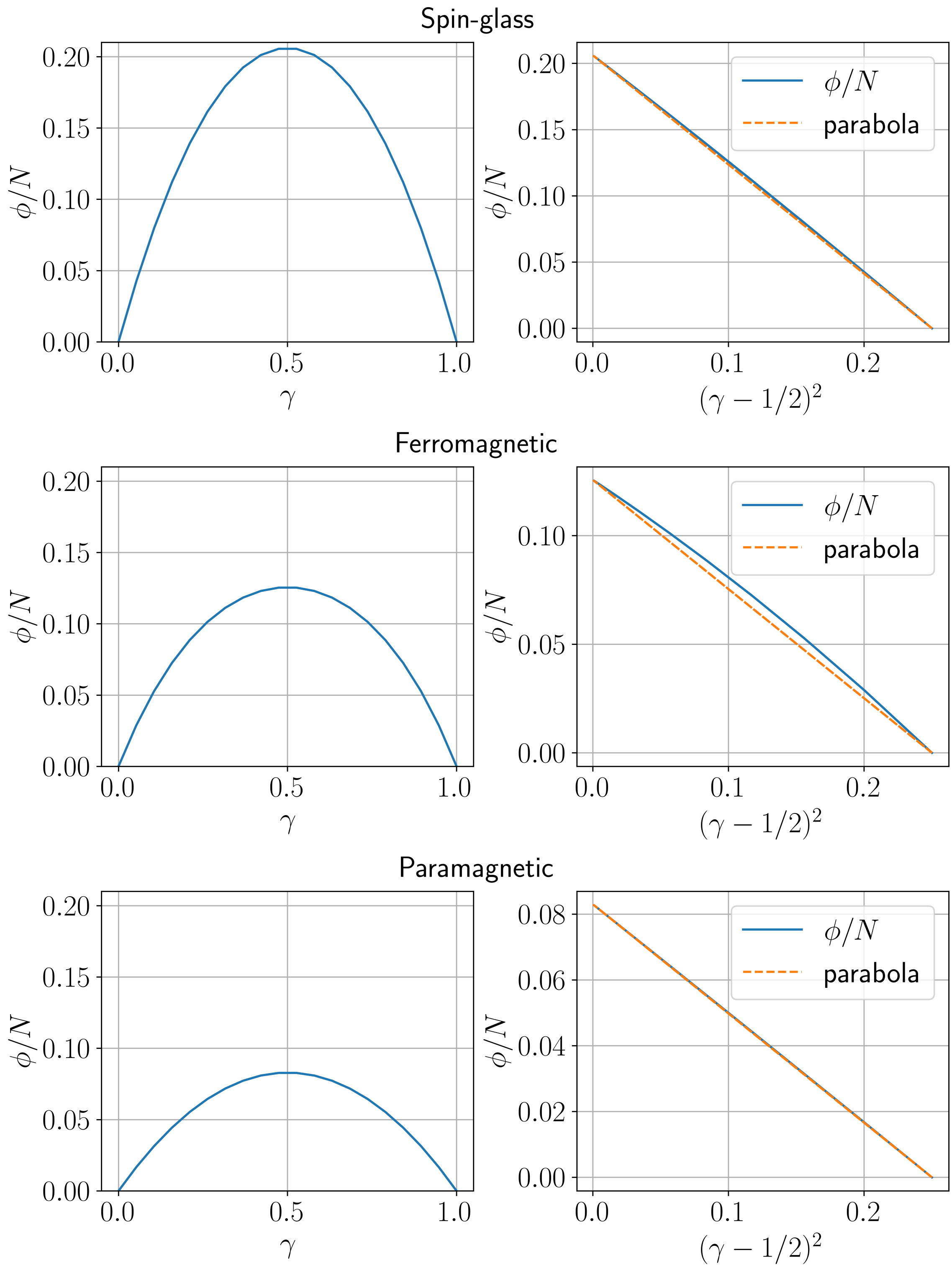


Figure 5.1: Comparison of the γ-dependence for each of the three phases for the LSK model in equilibrium. The control parameters used were: $T/J = 0.5$ and $J_0/J = 0.5$ for the Spin-glass phase, $T/J = 1.75$ and $J_0/J = 1$ for the Ferromagnetic phase and $T/J = 1.5$ and $J_0/J = 0.5$ for the Paramagnetic phase.

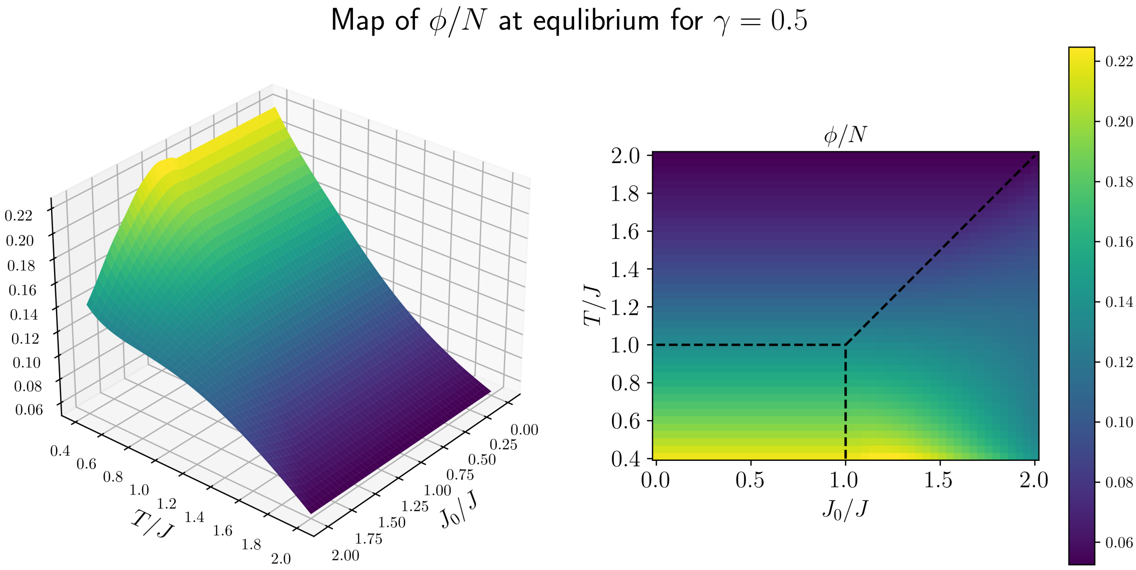


Figure 5.2: Map of ϕ/N across the whole phase diagram for $\gamma = 1/2$.

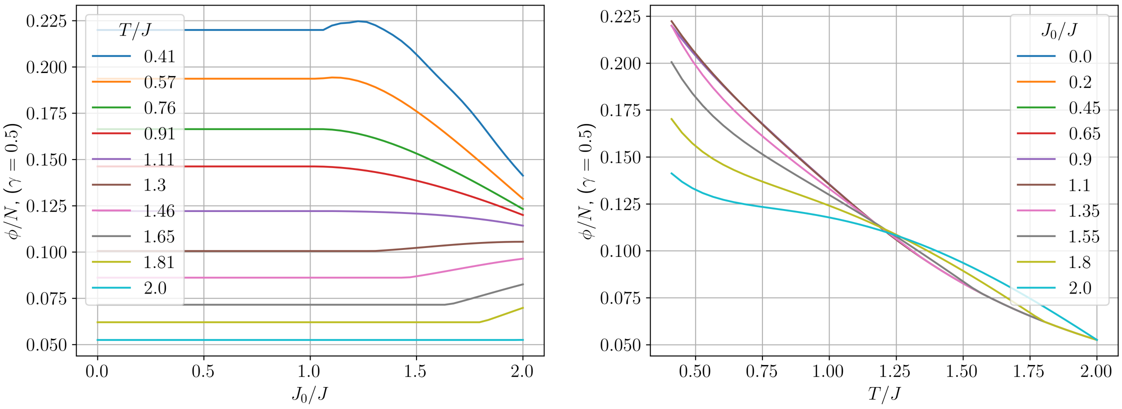
Kullback-Leibler divergence and by Jensen's inequality it is always positive). Thus we omitted the low temperature region for visualization sake, always remembering that our conclusions should be taken with a grain of salt for all the unstable region.

The first thing that is clear when we look at this map is that ϕ is typically lower in the paramagnetic phase and higher in the spin-glass phase. This is an interesting behavior for a complexity index, since the spin glass phase is known to have a complex structure with an interplay between order and disorder.

To better visualize the the behavior of ϕ we have made graphics of slices at constant J_0/J (figure 5.3a) and constant T/J (figure 5.3b). In both graphics we note a non-continuous derivative in the region of high T/J and high J_0/J indicating a second order phase transition.

Firstly, on the graphic of $\phi/N \times J_0/J$ we notice that for all temperatures ϕ is constant (with its values depending on the temperature) for low values of J_0/J and, at some point, they begin to change. For lower temperatures this change is smooth, but for high temperatures the change is not differentiable. These two transitions occurs when the system goes from the spin glass phase to the ferromagnetic phase and from the paramagnetic phase to the ferromagnetic phase, respectively. On the first transition (spin-glass to ferromagnetic) we see a drop on the value of ϕ and on the second one (paramagnetic to ferromagnetic) we have an increase on the value of ϕ. Another interesting feature is that for lower temperatures the value of ϕ is consistently higher than for lower temperatures, independently of the value of J_0/J, and in particular on the transition from paramagnetic to spin glass we have an increase on ϕ.

On the graphic of $\phi/N \times T/J$ the behavior is quite different, for lower values of J_0/J all the values of ϕ is the same monotonically decreasing function of T/J, and around $J_0/J = 1$ the curves present a different behavior, while also being monotonically decreasing. Again we notice that for lower temperatures ϕ is typically higher.



(a) Graphics of $\phi/N \times J_0/J$ for various values of T/J.

(b) Graphics of $\phi/N \times T/J$ for various values of J_0/J.

Figure 5.3: Slices from the surface plot in figure 5.2

5.3 ϕ on its way to equilibrium

Using the approximation described at the last section of chapter 3 we briefly studied how ϕ behaves over our approximate dynamics near equilibrium described earlier.

This analysis is an even bigger stretch on the validity of our results, just like the dynamical system, the formula for ϕ we calculated is only correct in the equilibrium. As this calculations are a novelty, we might as well overextend our result out of sheer curiosity.

With that in mind, we begin considering how ϕ evolves over trajectories beginning at a grid of initial conditions for the three different phases. The result can be found in figure 5.4.

Note that the behavior is very similar for all the phases. We again notice that the trajectories seems to go to the equilibrium very quickly in the paramagnetic phase and slower in the ferromagnetic phase, a fact that we have already noted when we looked at the trajectories alone (see figure 3.3). Besides this, there is no visible pattern that distinguishes them and, if it exists, our approximation probably does not captures it.

Another attempt to analyse this toy dynamics of ϕ was to consider a system in equilibrium at an initial T/J and J_0/J and change these values measuring how ϕ reacts to this. We have made such study for all sort of combinations of initial and final T/J and J_0/J, but, again, we have not been able to identify a pattern or gain information on the nature of ϕ. Some of those results can be found at figure 5.5. In the graphics in the left column, the green dashed line represents the value of ϕ the system had initially, and the red dashed line the value of ϕ for the new equilibrium point. The middle column diagrams shows how the initial and final conditions are located with respect to the phase boundaries, and on the left we have the trajectory the system describes in the $m \times q$ plane.

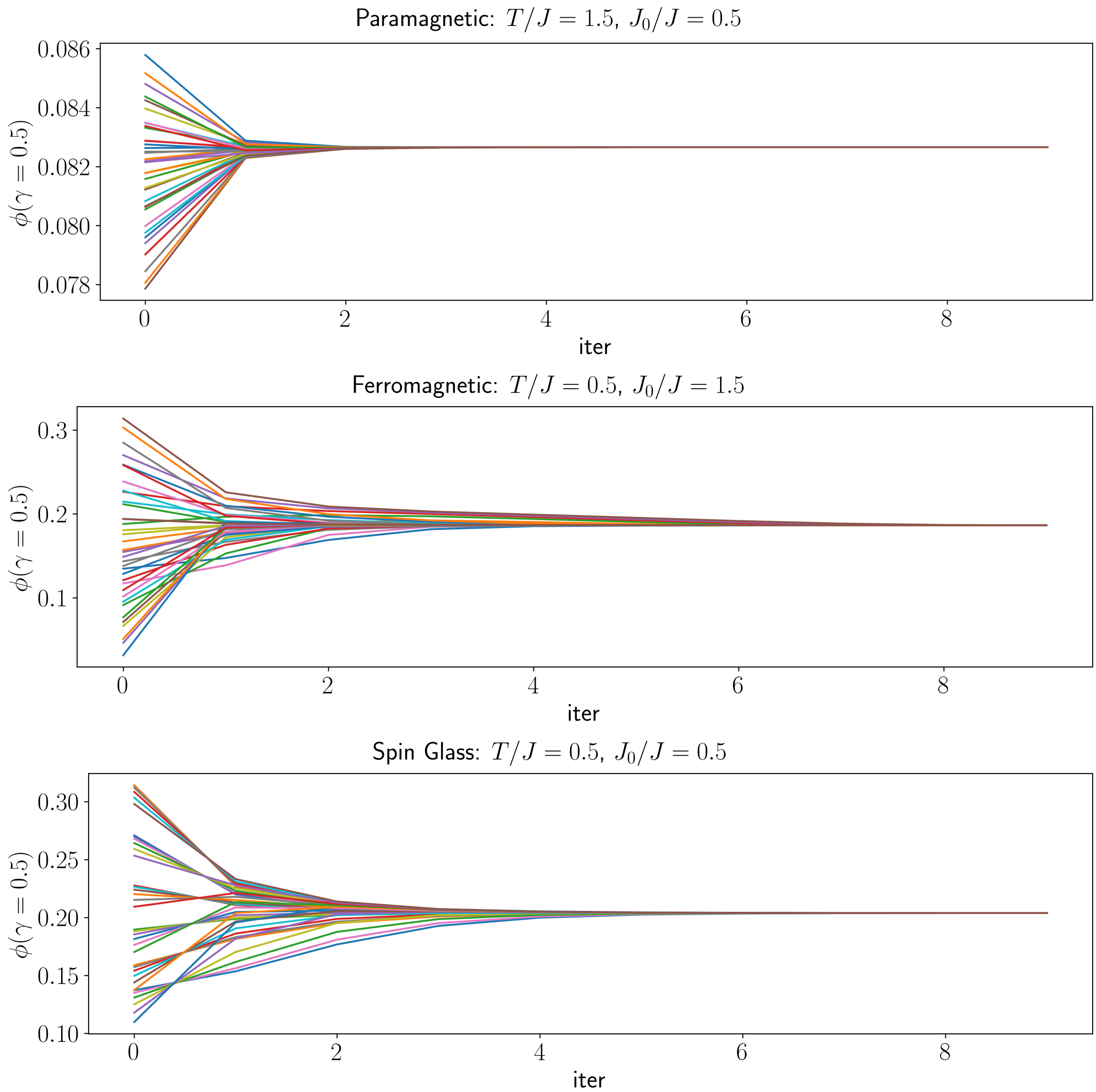


Figure 5.4: Comparison of the evolution of ϕ for a grid of initial conditions at the three phases.

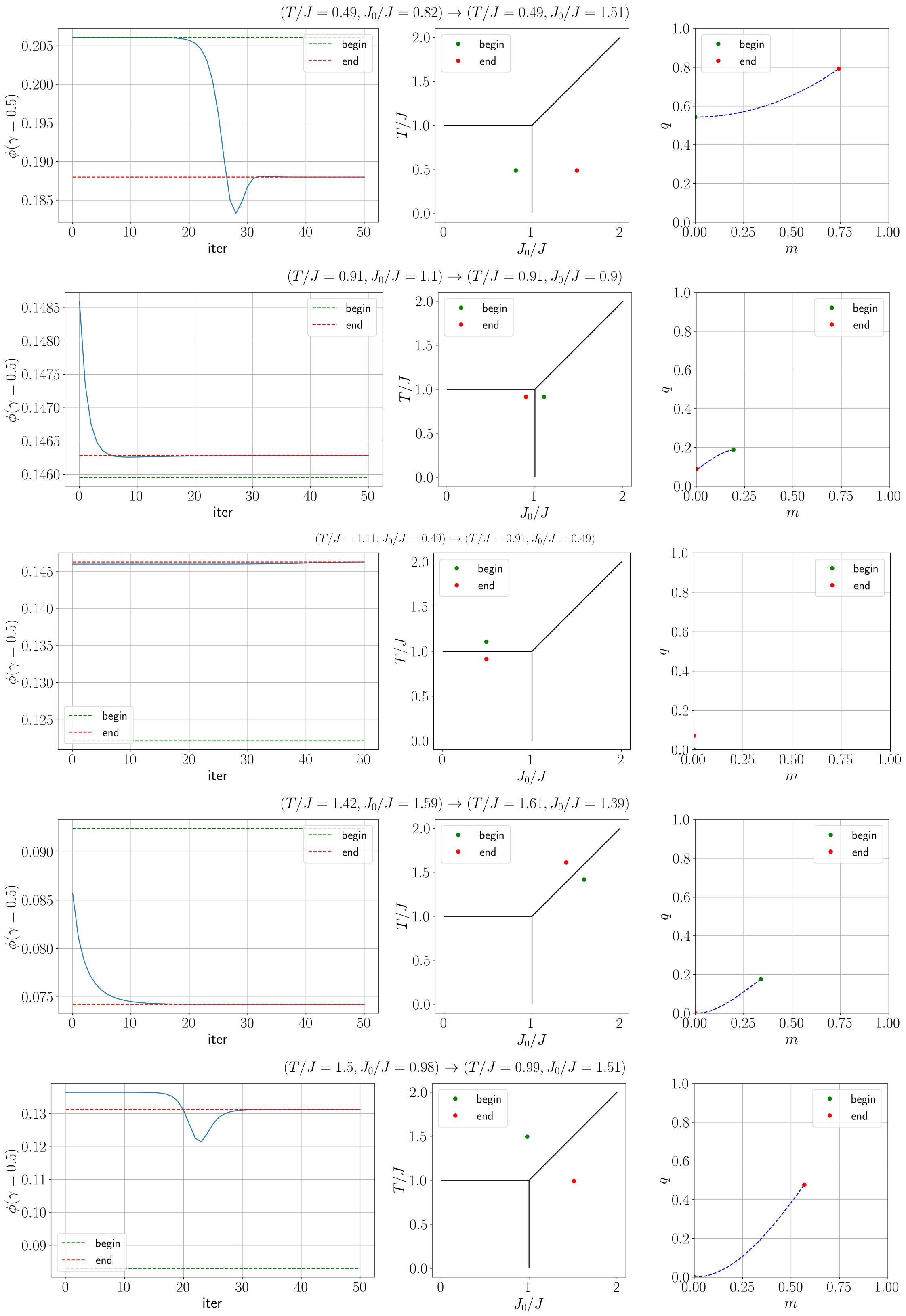


Figure 5.5: Examples of behavior of ϕ when the equilibrium condition T/J and J_0/J is suddenly changed.

5.4 Connection with Stochastic Interactions

In an attempt to interpret the meaning of our result, we take a step back and look at other complexity measures. In particular we look at what is called Stochastic Interactions (I_S), for it is known that this index have a close relationship with integrated information, as it is indicated by Ito et al. in [20].

In the geometric information framework, such complexity measure is obtained when we consider the following statistical manifold

$$\mathcal{M}_S = \left\{ Q(X,Y) \middle| Q(Y|X) = \prod_{I\in\Pi} Q(Y_I|X_I) \right\}, \tag{5.3}$$

that is, given a partitioning of the system Π (that we take here as a bipartition (I, J)), is the set of distributions in which each component of Π evolves independently. I_S is then defined as

$$I_S = \min_{Q\in\mathcal{M}_S} D_{KL}\left[P\|Q\right]. \tag{5.4}$$

Performing the minimization just like we did for integrated information we find

$$I_S = \left\langle \sum_{X,Y} P\left(X,Y|\boldsymbol{J}\right) \log \frac{P\left(Y|X;\boldsymbol{J}\right)}{P\left(Y_I|X_I;\boldsymbol{J}\right) P\left(Y_J|X_J;\boldsymbol{J}\right)} \right\rangle_{\boldsymbol{J}}. \tag{5.5}$$

Now we use the prescription for implementing disconnections by site dilution (see equation (4.11)), to write

$$\begin{aligned} P\left(Y_I|X_I;\boldsymbol{J},\boldsymbol{\eta}\right) &= \frac{\exp\left\{\beta\sum_{i,j} J_{ij}\eta_i\eta_j x_j y_i\right\}}{\prod_i 2\cosh\left(\beta\sum_j J_{ij}\eta_i\eta_j x_j\right)}, & (5.6)\\ P\left(Y_J|X_J;\boldsymbol{J},\boldsymbol{\eta}\right) &= \frac{\exp\left\{\beta\sum_{i,j} J_{ij}\left(1-\eta_i\right)\left(1-\eta_j\right) x_j y_i\right\}}{\prod_i 2\cosh\left(\beta\sum_j J_{ij}\left(1-\eta_i\right)\left(1-\eta_j\right) x_j\right)}. & (5.7) \end{aligned}$$

Substituting in our expression for I_S we have

$$I_S = \left\langle \sum_{X,Y} P\left(X,Y|\boldsymbol{J}\right) \left(\beta\sum_{i=1}^{N} y_i h_{i|D} - \sum_{i=1}^{N} \log 2\cosh(\beta h_i) + \sum_{i=1}^{N} \log 2\cosh(\beta h_{i|S}) \right) \right\rangle_{\boldsymbol{J}}, \tag{5.8}$$

which can be identified as the same expression for the integrated information index.

Looking at the equation (5.5) we see that it can be written as a difference of transitional entropy:

$$I_S = S(Y_I|X_I) + S(Y_J|X_J) - S(Y|X) = \phi_{IJ}, \tag{5.9}$$

giving us, at least for this particular model, an interpretation for the geometric integrated information index: it is a measure of how much information is gained when we let parts of the system communicate with each other in order to predict the future state.

Chapter 6

Conclusions

In this work, we have proposed an analysis of complexity measures, in particular the integrated information index, for physical systems, as an attempt to better understand how it can be used to assess complexity in all sorts of systems. Our main motivation was the experimental application of such indices as markers of consciousness in patients whose brains are subject to different conditions.

Due to the large toolbox of information geometry and the very straightforward definition of complexity measures, the information geometric framework developed by Oizumi et al. have been chosen as our main framework. By restricting the definition to a bipartition, we have been able to calculate an exact formula for the integrated information index. We also made progress deriving a formula for the geometric integrated information for quenched disordered systems.

We investigated the Little-Sherrington-Kirkpatrick model, a model that has pretty much the same behavior as the standard Sherrington-Kirkpatrick model at the equilibrium, but has a dynamics built-in in its definition, which may be very useful depending on the type of problem we are trying to study. Although we have successfully calculated its free energy and fully described the system in equilibrium (at least in the replica symmetric ansatz), there is still much to be done when we talk about out-of-equilibrium statistical mechanics. The near equilibrium approximation dynamics analysis did not give insight on the behavior of ϕ and should be properly addressed in a follow-up project.

Our analysis for the LSK model has shown how ϕ changes as we go from one phase to another: going from the paramagnetic phase to any other phase ϕ increases, as well as going from the ferromagnetic to the spin glass phase. By looking at the geometric integrated information index as a complexity measure, we are able to rank the three phases of the LSK model in terms of its complexity, where the spin glass is the most complex phase and the paramagnetic the least complex one, as one would have expected.

We also have analyzed how ϕ depends on the partitioning. Our result shows that for this statistically uniform model, this dependence is only through the size of the partition and it is very similar in all three phases: ϕ goes to zero when the partition contains zero or all the elements, and it has a maximum when the system is cut in half.

Finally, we noticed that, for the LSK model, our formula for the geometric integrated information is equal to the stochastic interactions measure, a difference of the transitional entropy for the disconnected model and the full model, and we can give it a nice interpretation as the amount of information that is gained when we let parts of the system communicate with each other in order to predict the future state. This is a nice feature that our system has, but it would also be interesting to study under which conditions a system has this property and how the integrated information differs from stochastic interactions if this relation does not hold.

In conclusion, we show that the geometric integrated information index can be used as a measure that assess the complexity of a physical system. The symmetries present in our LSK model have facilitated our calculations, but have produced results that does not seem very interesting at a first glance. The definition of ϕ is dynamic and encompasses two consecutive states in time, so our equilibrium analysis only captures a small part of its phase space and a lot more might be concluded studying the system out of equilibrium.

A problem that was not addressed here, is the calculation of Φ, the average of ϕ over all possible partitioning. In our case, this average would be described by a distribution $P(\gamma)$, over the all possible sizes of partitions. We believe this prescription is related with the kind of system we are analyzing, and for the LSK model there is no natural distribution for the partitions, so the choice is arbitrary. Information processing systems, arising from evolution, have a modular architecture of specialized macroscopic units and may suggest a particular partition as natural.

Thus, for future projects we might try to change the disorder, for example a system with Hebbian interaction, that presents a different order parameter that combines states in two different times, make a more in-depth analysis for the system out of equilibrium, and, last but not least, investigate how the symmetries of a synchronous system affect the behavior of the integrated information index, and consider systems that have a natural way of partitioning.

Bibliography

[1] David Sherrington and Scott Kirkpatrick. Solvable model of a spin-glass. *Phys. Rev. Lett.*, 35:1792–1796, Dec 1975.

[2] S F Edwards and P W Anderson. Theory of spin glasses. *Journal of Physics F: Metal Physics*, 5(5):965, may 1975.

[3] G. Parisi. The order parameter for spin glasses: a function on the interval 0-1. *Journal of Physics A: Mathematical and General*, 13(3):1101, mar 1980.

[4] G. Parisi. Infinite number of order parameters for spin-glasses. *Phys. Rev. Lett.*, 43:1754–1756, Dec 1979.

[5] W. A. Little. *The Existence of Persistent States in the Brain*, pages 145–164. Springer US, Boston, MA, 1974.

[6] Giulio Tononi. An information integration theory of consciousness. *BMC Neuroscience*, 5(1):42, November 2004.

[7] Giorgio Parisi. Complex systems: a physicist's viewpoint. *Physica A: Statistical Mechanics and its Applications*, 263(1):557–564, 1999. Proceedings of the 20th IUPAP International Conference on Statistical Physics.

[8] Leonardo S. Barbosa, William Marshall, Sabrina Streipert, Larissa Albantakis, and Giulio Tononi. A measure for intrinsic information. *Scientific Reports*, 10(1), 2020.

[9] Leonardo S. Barbosa, William Marshall, Larissa Albantakis, and Giulio Tononi. Mechanism integrated information. *Entropy*, 23(3), 2021.

[10] Adenauer G Casali, Olivia Gosseries, Mario Rosanova, Mélanie Boly, Simone Sarasso, Karina R Casali, Silvia Casarotto, Marie-Aurélie Bruno, Steven Laureys, Giulio Tononi, and Marcello Massimini. A theoretically based index of consciousness independent of sensory processing and behavior. *Sci Transl Med*, 5(198):198ra105, August 2013.

[11] Shun-ichi Amari. *Information geometry and its applications*, volume 194. Springer, 2016.

[12] Masafumi Oizumi, Naotsugu Tsuchiya, and Shun-ichi Amari. Unified framework for information integration based on information geometry. *Proceedings of the National Academy of Sciences*, 113(51):14817–14822, 2016.

[13] Tononi G. Consciousness as integrated information: A provisional manifesto. *Biological Bulletin*, 215(3):216–242, 2008.

[14] Scott Kirkpatrick and David Sherrington. Infinite-ranged models of spin-glasses. *Phys. Rev. B*, 17:4384–4403, Jun 1978.

[15] Hopfield JJ. Neural networks and physical systems with emergent collective computational abilities. *Proc Natl Acad Sci U S A.*, 79(8):2554–8, 1982.

[16] P. Peretto. Collective properties of neural networks: A statistical physics approach. *Biological Cybernetics*, 50:51–62, 1984.

[17] Fontanari JF and Köberle R. Information storage and retrieval in synchronous neural networks. *Phys Rev A Gen Phys.*, 36(5):2475–2477, 1987.

[18] E. T. Jaynes. *Probability Theory: The Logic of Science*. Cambridge University Press, 2003.

[19] J. R. L. de Almeida and D. J. Thouless. Stability of the sherrington-kirkpatrick solution of a spin glass model. *Journal of Physics A: Mathematical and General*, 11(5):983, may 1978.

[20] Sosuke Ito, Masafumi Oizumi, and Shun-ichi Amari. Unified framework for the entropy production and the stochastic interaction based on information geometry. *Phys. Rev. Res.*, 2:033048, Jul 2020.

Appendix A

Free Energy for LSK Model

Here we perform the calculation of the free energy for the LSK model

$$f = -\lim_{N\to\infty} \frac{1}{\beta N} \langle \log Z \rangle_{\boldsymbol{J}} , \tag{A.1}$$

where the partition function Z is given by

$$Z = \sum_{X,Y} \exp\left\{ \beta \sum_{i,j} J_{ij} x_j y_i \right\} , \tag{A.2}$$

and the distribution $P(\boldsymbol{J}) = \prod_{i,j} P(J_{ij})$ with

$$P(J_{ij}) = \sqrt{\frac{N}{2\pi J^2}} \exp\left\{ -\frac{N}{2J^2} \left(J_{ij} - \frac{J_0}{N} \right)^2 \right\} . \tag{A.3}$$

Using the replica method, our work will resume to the calculation of

$$\left\langle Z^\delta \right\rangle_{\boldsymbol{J}} = \sum_{\{X^a Y^a\}} \left\langle \exp\left\{ \beta \sum_a \sum_{i,j} J_{ij} x_j^a y_i^a \right\} \right\rangle_{\boldsymbol{J}} \tag{A.4}$$

Performing the average of the disorder $\boldsymbol{J}$:

$$\left\langle Z^\delta \right\rangle_{\boldsymbol{J}} = \sum_{\{X^a Y^a\}} \exp\left\{ \frac{\beta J_0}{N} \sum_{i,j} \sum_a x_j^a y_i^a + \frac{\beta^2 J^2}{2N} \sum_{i,j} \left(\sum_a x_j^a y_i^a \right)^2 \right\} \tag{A.5}$$

$$\begin{aligned} \left\langle Z^\delta \right\rangle_{\boldsymbol{J}} = \exp\left(\frac{\beta^2 J^2 \delta N}{2} \right) \sum_{\{X^a Y^a\}} \exp\Bigg\{ & \frac{\beta J_0}{N} \sum_a \left(\sum_i x_i^a \right) \left(\sum_i y_i^a \right) + \\ & + \frac{\beta^2 J^2}{N} \sum_{a<b} \left(\sum_i x_i^a x_i^b \right) \left(\sum_i y_i^a y_i^b \right) \Bigg\} \end{aligned} \tag{A.6}$$

Introducing integrals over Dirac delta distributions:

$$
\begin{aligned}
\left\langle Z^{\delta}\right\rangle_{\boldsymbol{J}} =& \exp\left(\frac{\beta^2 J^2 \delta N}{2}\right) \sum_{\{X^a Y^a\}} \int \prod_a N dm_a N dn_a \int \prod_{a<b} N dq_{ab} N dr_{ab} \times \\
& \times \exp\left\{\beta J_0 N \sum_a m_a n_a + \beta^2 J^2 N \sum_{a<b} q_{ab} r_{ab}\right\} \times \\
& \times \prod_a \delta\left(N m_a - \sum_i x_i^a\right) \delta\left(N n_a - \sum_i y_i^a\right) \times \\
& \times \prod_{a<b} \delta\left(N q_{ab} - \sum_i x_i^a x_i^b\right) \delta\left(N r_{ab} - \sum_i y_i^a y_i^b\right)
\end{aligned} \tag{A.7}
$$

Using the the Fourier integral representation of the delta distributions:

$$
\begin{aligned}
\left\langle Z^{\delta}\right\rangle_{\boldsymbol{J}} =& \exp\left(\frac{\beta^2 J^2 \delta N}{2}\right) \sum_{\{X^a Y^a\}} \int \prod_a \frac{dm_a d\hat{m}_a}{2\pi/N} \frac{dn_a d\hat{n}_a}{2\pi/N} \int \prod_{a<b} \frac{dq_{ab} d\hat{q}_{ab}}{2\pi/N} \frac{dr_{ab} d\hat{r}_{ab}}{2\pi/N} \times \\
& \times \exp\left\{\beta J_0 N \sum_a m_a n_a + \beta^2 J^2 N \sum_{a<b} q_{ab} r_{ab} + iN \sum_a (m_a \hat{m}_a + n_a \hat{n}_a) + \right. \\
& + iN \sum_{a<b} (q_{ab}\hat{q}_{ab} + r_{ab}\hat{r}_{ab}) - i \sum_a \hat{m}_a \sum_i x_i^a - i \sum_a \hat{n}_a \sum_i y_i^a + \\
& \left. - i \sum_{a<b} \hat{q}_{ab} \sum_i x_i^a x_i^b - i \sum_{a<b} \hat{r}_{ab} \sum_i y_i^a y_i^b \right\}.
\end{aligned} \tag{A.8}
$$

Grouping the terms that depend on the variables $\{X^a\}$ and $\{Y^a\}$

$$
\begin{aligned}
\left\langle Z^{\delta}\right\rangle_{\boldsymbol{J}} =& \exp\left(\frac{\beta^2 J^2 \delta N}{2}\right) \int \prod_a \frac{dm_a d\hat{m}_a}{2\pi/N} \int \prod_a \frac{dn_a d\hat{n}_a}{2\pi/N} \int \prod_{a<b} \frac{dq_{ab} d\hat{q}_{ab}}{2\pi/N} \int \prod_{a<b} \frac{dr_{ab} d\hat{r}_{ab}}{2\pi/N} \times \\
& \times \exp\left\{\beta J_0 N \sum_a m_a n_a + \beta^2 J^2 N \sum_{a<b} q_{ab} r_{ab} + iN \sum_a (m_a \hat{m}_a + n_a \hat{n}_a) + \right. \\
& \left. + iN \sum_{a<b} (q_{ab}\hat{q}_{ab} + r_{ab}\hat{r}_{ab}) + N \log \mathcal{Z}_{\text{rep}} \right\},
\end{aligned} \tag{A.9}
$$

where

$$
\mathcal{Z}_{\text{rep}} = \sum_{\{x^a y^a\}} \exp\left\{-\mathcal{H}_{\text{rep}}\right\}, \tag{A.10}
$$

is the canonical partition function associated with the replica interaction Hamiltonian $\mathcal{H}_{\text{rep}}$, defined as:

$$
\mathcal{H}_{\text{rep}} = i \sum_a \hat{m}_a x^a + i \sum_{a<b} \hat{q}_{ab} x^a x^b + i \sum_a \hat{n}_a y^a + i \sum_{a<b} \hat{r}_{ab} y^a y^b. \tag{A.11}
$$

Now, the integrand in the expression for $\left\langle Z^{\delta}\right\rangle_{\boldsymbol{J}}$ is an exponential with exponent proportional to N and as we are interested in the study of the thermodynamic limit,

$N \to \infty$, we can apply the saddle point method to approximate the integral as

$$\left\langle Z^\delta \right\rangle_{\boldsymbol{J}} = \exp\left\{ \frac{\beta^2 J^2 \delta N}{2} + \beta J_0 N \sum_a m_a n_a + iN \sum_a (m_a \hat{m}_a + n_a \hat{n}_a) + \right.$$
$$\left. + \beta^2 J^2 N \sum_{a<b} q_{ab} r_{ab} + iN \sum_{a<b} (q_{ab}\hat{q}_{ab} + r_{ab}\hat{r}_{ab}) + N \log \mathcal{Z}_{\text{rep}} \right\}, \qquad \text{(A.12)}$$

where, now, the parameters $\{m_a, \hat{m}_a\}$, $\{n_a, \hat{n}_a\}$, $\{q_{ab}, \hat{q}_{ab}\}$ and $\{r_{ab}, \hat{r}_{ab}\}$ are the saddle points of the exponent.

By imposing that the derivative of the exponent with respect to each one of them is equal to zero, we find:

$$m_a = \langle x_i^a \rangle_{\text{rep}} \qquad \hat{m}_a = i\beta J_0 n_a \qquad \text{(A.13)}$$
$$n_a = \langle y_i^a \rangle_{\text{rep}} \qquad \hat{n}_a = i\beta J_0 m_a \qquad \text{(A.14)}$$
$$q_{ab} = \left\langle x_i^a x_i^b \right\rangle_{\text{rep}} \qquad \hat{q}_{ab} = i\beta^2 J^2 r_{ab} \qquad \text{(A.15)}$$
$$r_{ab} = \left\langle y_i^a y_i^b \right\rangle_{\text{rep}} \qquad \hat{r}_{ab} = i\beta^2 J^2 q_{ab} \qquad \text{(A.16)}$$

We immediately see that, on the saddle points, the Fourier variables are not independent from the other parameters. Substituting this relationship into the former equations:

$$\left\langle Z^\delta \right\rangle_{\boldsymbol{J}} = \exp\left\{ \frac{\beta^2 J^2 \delta N}{2} - \beta J_0 N \sum_a m_a n_a - \beta^2 J^2 N \sum_{a<b} q_{ab} r_{ab} + N \log \mathcal{Z}_{\text{rep}} \right\}, \qquad \text{(A.17)}$$

$$\mathcal{Z}_{\text{rep}} = \sum_{\{x^a y^a\}} \exp\left\{ \beta J_0 \sum_a n_a x^a + \beta^2 J^2 \sum_{a<b} r_{ab} x^a x^b + \beta J_0 \sum_a m_a y^a + \beta^2 J^2 \sum_{a<b} q_{ab} y^a y^b \right\}. \qquad \text{(A.18)}$$

From the equation (3.13) we see that we are interested in the limit $\delta \to 0$, so we can linearize the expression for $\left\langle Z^\delta \right\rangle_{\boldsymbol{J}}$ around $\delta = 0$

$$\left\langle Z^\delta \right\rangle_{\boldsymbol{J}} \approx 1 - \beta N \delta \left(-\frac{\beta J^2}{2} + \frac{J_0}{\delta} \sum_a m_a n_a + \frac{\beta J^2}{\delta} \sum_{a<b} q_{ab} r_{ab} - \frac{1}{\beta\delta} \log \mathcal{Z}_{\text{rep}} \right), \qquad \text{(A.19)}$$

and rearrange the terms

$$-\frac{1}{\beta N} \frac{\left\langle Z^\delta \right\rangle_{\boldsymbol{J}} - 1}{\delta} \approx -\frac{\beta J^2}{2} + \frac{J_0}{\delta} \sum_a m_a n_a + \frac{\beta J^2}{\delta} \sum_{a<b} q_{ab} r_{ab} - \frac{1}{\beta\delta} \log \mathcal{Z}_{\text{rep}}. \qquad \text{(A.20)}$$

Now, taking the limits $\delta \to 0$ and $N \to \infty$, we have

$$f = \lim_{\delta \to 0} \left\{ -\frac{\beta J^2}{2} + \frac{J_0}{\delta} \sum_a m_a n_a + \frac{\beta J^2}{\delta} \sum_{a<b} q_{ab} r_{ab} - \frac{1}{\beta\delta} \log \mathcal{Z}_{\text{rep}} \right\}. \qquad \text{(A.21)}$$

A.1 Replica Symmetric Ansatz

To proceed, we assume the ansatz that the saddle points $\{m_a\}$, $\{n_a\}$, $\{q_{ab}\}$ and $\{r_{ab}\}$ are symmetric under replica index permutation. In this case we can drop off the replica index

$$m_a = m \qquad \hat{q}_{ab} = q \tag{A.22}$$

$$n_a = n \qquad \hat{r}_{ab} = r \tag{A.23}$$

and the expression for $\mathcal{Z}_{\text{rep}}$ becomes

$$\mathcal{Z}_{\text{RS}} = \sum_{\{x^a y^a\}} \exp\left\{\beta J_0 n \sum_a x^a + \beta^2 J^2 r \sum_{a<b} x^a x^b + \beta J_0 m \sum_a y^a + \beta^2 J^2 q \sum_{a<b} y^a y^b\right\}. \tag{A.24}$$

Introducing gaussian integrals to linearize the quadratic terms in $\{x^a\}$ and $\{y^a\}$

$$\mathcal{Z}_{\text{RS}} = \exp\left\{-\frac{\beta^2 J^2 \delta\,(q+r)}{2}\right\} \int Dz \left(2\cosh\left[\beta\left(J_0 m + J\sqrt{q}z\right)\right]\right)^\delta \times \\ \times \int Dz' \left(2\cosh\left[\beta\left(J_0 n + J\sqrt{r}z'\right)\right]\right)^\delta \tag{A.25}$$

Taking the logarithm and linearizing around $\delta = 0$

$$\log \mathcal{Z}_{\text{RS}} = -\frac{\beta^2 J^2 \delta\,(q+r)}{2} + \delta \int Dz \log 2\cosh\left[\beta\left(J_0 m + J\sqrt{q}z\right)\right] + \\ + \delta \int Dz' \log 2\cosh\left[\beta\left(J_0 n + J\sqrt{r}z'\right)\right]. \tag{A.26}$$

Then, in this ansatz, the free energy becomes

$$f_{\text{RS}} = J_0 m n - \frac{\beta J^2}{2}(1-q)(1-r) - \frac{1}{\beta}\int Dz \log 2\cosh\left[\beta\left(J_0 n + J\sqrt{r}z\right)\right] + \\ - \frac{1}{\beta}\int Dz \log 2\cosh\left[\beta\left(J_0 m + J\sqrt{q}z\right)\right], \tag{A.27}$$

where we already took the limit $\delta \to 0$.

Appendix B

$\Phi_{\boldsymbol{\eta}}$ Calculation

Our goal is to calculate

$$\phi_{\boldsymbol{\eta}} = \left\langle \sum_{X,Y} P\left(X,Y|\boldsymbol{J}\right) \left(\beta \sum_{i=1}^{N} y_i h_{i|D} - \sum_{i=1}^{N} \log 2\cosh(\beta h_i) + \sum_{i=1}^{N} \log 2\cosh(\beta h_{i|S}) \right) \right\rangle_{\boldsymbol{J}}, \tag{B.1}$$

and to do so, we consider the following terms separately:

$$\phi_A = \beta \left\langle \sum_{X,Y} P\left(X,Y|\boldsymbol{J}\right) \sum_{i=1}^{N} y_i h_{i|D} \right\rangle_{\boldsymbol{J}}, \tag{B.2}$$

$$\phi_B = -\left\langle \sum_{X,Y} P\left(X,Y|\boldsymbol{J}\right) \sum_{i=1}^{N} \log 2\cosh(\beta h_i) \right\rangle_{\boldsymbol{J}}, \tag{B.3}$$

$$\phi_C = \left\langle \sum_{X,Y} P\left(X,Y|\boldsymbol{J}\right) \sum_{i=1}^{N} \log 2\cosh(\beta h_{i|S}) \right\rangle_{\boldsymbol{J}}, \tag{B.4}$$

such that $\phi_{\boldsymbol{\eta}} = \phi_A + \phi_B + \phi_C$.

B.1 Calculation of ϕ_A

To calculate ϕ_A we define an auxiliary system with equilibrium distribution given by

$$P_A\left(X,Y|\boldsymbol{J},\boldsymbol{\eta},\beta_S,\beta_D\right) = \frac{1}{Z_A} \exp\left\{ \sum_i \left(\beta_S h_{i|S} + \beta_D h_{i|D}\right) y_i \right\}, \tag{B.5}$$

with disorder $\boldsymbol{J}$ described by the same prescription as the standard SK model (equation (3.11)). That is, a LSK model with elements from the same partition interacting with temperature β_S and elements from different partitions interacting with temperature β_D. Note that when we take $\beta_S = \beta_D = \beta$ we recover the standard LSK model.

Taking the derivative of $\langle \log Z_A \rangle_{\boldsymbol{J}}$ with respect to β_D and taking $\beta_S = \beta_D = \beta$,

$$\frac{\partial}{\partial \beta_D} \left\langle \log Z_A \right\rangle_{\boldsymbol{J}} \Big|_{\beta_S=\beta_D=\beta} = \left\langle \sum_{X,Y} P\left(X,Y|\boldsymbol{J}\right) \left(\sum_i h_{i|D} y_i \right) \right\rangle_{\boldsymbol{J}}. \tag{B.6}$$

We note that the RHS is, except for a factor of β, equal to ϕ_A,

$$\phi_A = \beta \, \frac{\partial}{\partial \beta_D} \left\langle \log Z_A \right\rangle_{\boldsymbol{J}} \Big|_{\beta_S=\beta_D=\beta} . \tag{B.7}$$

Our work then is to calculate $\left\langle \log Z_A \right\rangle_{\boldsymbol{J}}$. To do this we use the replica method. We calculate $\left\langle Z_A^\delta \right\rangle_{\boldsymbol{J}}$ for integer δ:

$$\left\langle Z_A^\delta \right\rangle_{\boldsymbol{J}} = \left\langle \sum_{\{X^a,Y^a\}} \exp\left\{ \sum_a \sum_i y_i^a \left(\beta_S h_{i|S}^a + \beta_D h_{i|D}^a \right) \right\} \right\rangle_{\boldsymbol{J}} , \tag{B.8}$$

$$= \left\langle \sum_{\{X^a,Y^a\}} \exp\left\{ \sum_a \sum_{i,j} J_{ij} x_j^a y_i^a \left(\beta_S \theta_{ij}^S + \beta_D \theta_{ij}^D \right) \right\} \right\rangle_{\boldsymbol{J}} , \tag{B.9}$$

where we introduced two new quantities, the projectors $\theta_{ij}^S = \eta_i \eta_j + (1-\eta_i)(1-\eta_j)$ and $\theta_{ij}^D = \eta_i(1-\eta_j) + (1-\eta_i)\eta_j$, that take into account the cases where i and j belongs to the same partition and to different partition, respectively.

Those projectors have the following interesting properties:

$$\left(\theta_{ij}^S\right)^2 = \theta_{ij}^S, \tag{B.10}$$

$$\left(\theta_{ij}^D\right)^2 = \theta_{ij}^D, \tag{B.11}$$

$$\theta_{ij}^S \theta_{ij}^D = 0, \tag{B.12}$$

$$\theta_{ij}^S + \theta_{ij}^D = 1. \tag{B.13}$$

Calculating the average over the disorder $\boldsymbol{J}$,

$$\left\langle Z_A^\delta \right\rangle_{\boldsymbol{J}} = \sum_{\{X^a,Y^a\}} \exp\left\{ \frac{J_0}{N} \sum_{i,j} \sum_a \left(\beta_S \theta_{ij}^S + \beta_D \theta_{ij}^D \right) x_j^a y_i^a + \frac{J^2}{2N} \sum_{i,j} \left(\sum_a \left(\beta_S \theta_{ij}^S + \beta_D \theta_{ij}^D \right) x_j^a y_i^a \right)^2 \right\} , \tag{B.14}$$

and using the properties of θ_{ij}^S and θ_{ij}^D we can write

$$\left\langle Z_A^\delta \right\rangle_{\boldsymbol{J}} = \sum_{\{X^a,Y^a\}} \exp\left\{ \frac{J_0}{N} \sum_{i,j} \sum_a \left(\beta_S \theta_{ij}^S + \beta_D \theta_{ij}^D \right) x_j^a y_i^a + \frac{J^2}{2N} \sum_{i,j} \left(\beta_S^2 \theta_{ij}^S + \beta_D^2 \theta_{ij}^D \right) \left(\sum_a x_j^a y_i^a \right)^2 \right\} . \tag{B.15}$$

Rearranging the terms and defining $\gamma = \sum_i \eta_i / N$, such that γN is the number of elements in the partition I (and $(1-\gamma)N$ the number of elements in partition J), we have:

$$
\begin{aligned}
\left\langle Z_A^\delta \right\rangle_{\boldsymbol{J}} = & \exp\left\{ \frac{1}{2} J^2 N \delta \left[\beta_S^2 \left(\gamma^2 + (1-\gamma)^2 \right) + \beta_D^2 2\gamma \left(1-\gamma\right) \right] \right\} \times \\
& \times \sum_{\{X^a, Y^a\}} \exp\left\{ \frac{\beta_S J_0}{N} \sum_a \left(\sum_i \eta_i x_i^a \right) \left(\sum_i \eta_i y_i^a \right) + \right. \\
& + \frac{\beta_S J_0}{N} \sum_a \left(\sum_i \left(1-\eta_i\right) x_i^a \right) \left(\sum_i \left(1-\eta_i\right) y_i^a \right) + \\
& + \frac{\beta_D J_0}{N} \sum_a \left(\sum_i \eta_i x_i^a \right) \left(\sum_i \left(1-\eta_i\right) y_i^a \right) + \\
& + \frac{\beta_D J_0}{N} \sum_a \left(\sum_i \left(1-\eta_i\right) x_i^a \right) \left(\sum_i \eta_i y_i^a \right) + \\
& + \frac{\beta_S^2 J^2}{N} \sum_{a<b} \left(\sum_i \eta_i x_i^a x_i^b \right) \left(\sum_i \eta_i y_i^a y_i^b \right) + \\
& + \frac{\beta_S^2 J^2}{N} \sum_{a<b} \left(\sum_i \left(1-\eta_i\right) x_i^a x_i^b \right) \left(\sum_i \left(1-\eta_i\right) y_i^a y_i^b \right) + \\
& + \frac{\beta_D^2 J^2}{N} \sum_{a<b} \left(\sum_i \eta_i x_i^a x_i^b \right) \left(\sum_i \left(1-\eta_i\right) y_i^a y_i^b \right) + \\
& \left. + \frac{\beta_D^2 J^2}{N} \sum_{a<b} \left(\sum_i \left(1-\eta_i\right) x_i^a x_i^b \right) \left(\sum_i \eta_i y_i^a y_i^b \right) \right\}.
\end{aligned}
\tag{B.16}
$$

Completing the square in each line of the above equation,

$$
\begin{aligned}
\left\langle Z_A^\delta \right\rangle_{\boldsymbol{J}} = & \exp\left\{\frac{1}{2}J^2 N\delta\left[\beta_S^2\left(\gamma^2+(1-\gamma)^2\right)+\beta_D^2 2\gamma(1-\gamma)\right]\right\}\times \\
& \times \sum_{\{X^a,Y^a\}} \exp\left\{\beta_S J_0 N\sum_a\left[\left(\frac{1}{N}\sum_i \eta_i\left(x_i^a+y_i^a\right)\right)^2-\left(\frac{1}{N}\sum_i \eta_i x_i^a\right)^2-\left(\frac{1}{N}\sum_i \eta_i y_i^a\right)^2\right]+\right. \\
& + \beta_S J_0 N\sum_a\left[\left(\frac{1}{N}\sum_i (1-\eta_i)\left(x_i^a+y_i^a\right)\right)^2-\left(\frac{1}{N}\sum_i (1-\eta_i) x_i^a\right)^2-\left(\frac{1}{N}\sum_i (1-\eta_i) y_i^a\right)^2\right]+ \\
& + \beta_D J_0 N\sum_a\left[\left(\frac{1}{N}\sum_i \left(\eta_i x_i^a+(1-\eta_i) y_i^a\right)\right)^2-\left(\frac{1}{N}\sum_i \eta_i x_i^a\right)^2-\left(\frac{1}{N}\sum_i (1-\eta_i) y_i^a\right)^2\right]+ \\
& + \beta_D J_0 N\sum_a\left[\left(\frac{1}{N}\sum_i \left((1-\eta_i) x_i^a+\eta_i y_i^a\right)\right)^2-\left(\frac{1}{N}\sum_i (1-\eta_i) x_i^a\right)^2-\left(\frac{1}{N}\sum_i \eta_i y_i^a\right)^2\right]+ \\
& + \beta_S^2 J^2 N\sum_{a<b}\left[\left(\frac{1}{N}\sum_i \eta_i\left(x_i^a x_i^b+y_i^a y_i^b\right)\right)^2-\left(\frac{1}{N}\sum_i \eta_i x_i^a x_i^b\right)^2-\left(\frac{1}{N}\sum_i \eta_i y_i^a y_i^b\right)^2\right]+ \\
& + \beta_S^2 J^2 N\sum_{a<b}\left[\left(\frac{1}{N}\sum_i (1-\eta_i)\left(x_i^a x_i^b+y_i^a y_i^b\right)\right)^2-\left(\frac{1}{N}\sum_i (1-\eta_i) x_i^a x_i^b\right)^2-\left(\frac{1}{N}\sum_i (1-\eta_i) y_i^a y_i^b\right)^2\right]+ \\
& + \beta_D^2 J^2 N\sum_{a<b}\left[\left(\frac{1}{N}\sum_i \left(\eta_i x_i^a x_i^b+(1-\eta_i) y_i^a y_i^b\right)\right)^2-\left(\frac{1}{N}\sum_i \eta_i x_i^a x_i^b\right)^2-\left(\frac{1}{N}\sum_i (1-\eta_i) y_i^a y_i^b\right)^2\right]+ \\
& \left. + \beta_D^2 J^2 N\sum_{a<b}\left[\left(\frac{1}{N}\sum_i \left((1-\eta_i) x_i^a x_i^b+\eta_i y_i^a y_i^b\right)\right)^2-\left(\frac{1}{N}\sum_i (1-\eta_i) x_i^a x_i^b\right)^2-\left(\frac{1}{N}\sum_i \eta_i y_i^a y_i^b\right)^2\right]\right\}.
\end{aligned}
\tag{B.17}
$$

Now, for each of the square term in the exponent we introduce a gaussian integral using the identity:

$$
e^{ab^2}=\sqrt{\frac{a}{\pi}}\int e^{-ax^2+2abx}dx. \tag{B.18}
$$

Ignoring the multiplicative constants, we have

$$\begin{aligned}
\left\langle Z_A^\delta \right\rangle_{\boldsymbol{J}} =& \exp\left\{\frac{1}{2}J^2 N\delta\left[\beta_S^2\left(\gamma^2+(1-\gamma)^2\right)+\beta_D^2 2\gamma(1-\gamma)\right]\right\}\times\\
&\times\int\prod_a dl_a^I dl_a^J dl_a^{IJ} dl_a^{JI}\int\prod_a dm_a^I dm_a^J\int\prod_a dn_a^I dn_a^J\times\\
&\times\int\prod_{a<b} dp_{ab}^I dp_{ab}^J dp_{ab}^{IJ} dp_{ab}^{JI}\int\prod_{a<b} dq_{ab}^I dq_{ab}^J\int\prod_{a<b} dr_{ab}^I dr_{ab}^J\times\\
&\times\exp\left\{-\frac{\beta_S J_0 N}{2}\sum_a\left[\left(l_a^I\right)^2+\left(m_a^I\right)^2+\left(n_a^I\right)^2+\left(l_a^J\right)^2+\left(m_a^J\right)^2+\left(n_a^J\right)^2\right]+\right.\\
&-\frac{\beta_D J_0 N}{2}\sum_a\left[\left(l_a^{IJ}\right)^2+\left(m_a^I\right)^2+\left(n_a^J\right)^2+\left(l_a^{JI}\right)^2+\left(m_a^J\right)^2+\left(n_a^I\right)^2\right]+\\
&-\frac{\beta_S^2 J^2 N}{2}\sum_{a<b}\left[\left(p_{ab}^I\right)^2+\left(q_{ab}^I\right)^2+\left(r_{ab}^I\right)^2+\left(p_{ab}^J\right)^2+\left(q_{ab}^J\right)^2+\left(r_{ab}^J\right)^2\right]+\\
&\left.-\frac{\beta_D^2 J^2 N}{2}\sum_{a<b}\left[\left(p_{ab}^{IJ}\right)^2+\left(q_{ab}^I\right)^2+\left(r_{ab}^J\right)^2+\left(p_{ab}^{JI}\right)^2+\left(q_{ab}^J\right)^2+\left(r_{ab}^I\right)^2\right]+\sum_i\log\zeta_i\right\},
\end{aligned}\tag{B.19}$$

with

$$\begin{aligned}
\zeta_i =& \sum_{\{x_i^a,y_i^a\}}\exp\left\{\beta_S J_0\sum_a\left[\left(l_a^I+im_a^I\right)\eta_i x_i^a+\left(l_a^I+in_a^I\right)\eta_i y_i^a\right]+\right.\\
&+\beta_S J_0\sum_a\left[\left(l_a^J+im_a^J\right)(1-\eta_i)x_i^a+\left(l_a^J+in_a^J\right)(1-\eta_i)y_i^a\right]+\\
&+\beta_D J_0\sum_a\left[\left(l_a^{IJ}+im_a^I\right)\eta_i x_i^a+\left(l_a^{IJ}+in_a^J\right)(1-\eta_i)y_i^a\right]+\\
&+\beta_D J_0\sum_a\left[\left(l_a^{JI}+im_a^J\right)(1-\eta_i)x_i^a+\left(l_a^{JI}+in_a^I\right)\eta_i y_i^a\right]+\\
&+\beta_S^2 J^2\sum_{a<b}\left[\left(p_{ab}^I+iq_{ab}^I\right)\eta_i x_i^a x_i^b+\left(p_{ab}^I+ir_{ab}^I\right)\eta_i y_i^a y_i^b\right]+\\
&+\beta_S^2 J^2\sum_{a<b}\left[\left(p_{ab}^J+iq_{ab}^J\right)(1-\eta_i)x_i^a x_i^b+\left(p_{ab}^J+ir_{ab}^J\right)(1-\eta_i)y_i^a y_i^b\right]+\\
&+\beta_D^2 J^2\sum_{a<b}\left[\left(p_{ab}^{IJ}+iq_{ab}^I\right)\eta_i x_i^a x_i^b+\left(p_{ab}^{IJ}+ir_{ab}^J\right)(1-\eta_i)y_i^a y_i^b\right]+\\
&\left.+\beta_D^2 J^2\sum_{a<b}\left[\left(p_{ab}^{JI}+iq_{ab}^J\right)(1-\eta_i)x_i^a x_i^b+\left(p_{ab}^{JI}+ir_{ab}^I\right)\eta_i y_i^a y_i^b\right]\right\}.
\end{aligned}\tag{B.20}$$

Each ζ_i can be interpreted as the partition function associated with the interaction of the i-th component of the replicas.

Now, as we are interested in the thermodynamic limit, $N\to\infty$, we are able to calculate the integrals using the saddle point method,

$$\begin{aligned}\left\langle Z_A^\delta\right\rangle_{\boldsymbol{J}} =&\exp\left\{\frac{1}{2}J^2N\delta\left[\beta_S^2\left(\gamma^2+(1-\gamma)^2\right)+\beta_D^2 2\gamma\left(1-\gamma\right)\right]+\sum_i\log\zeta_i+\right.\\
&-\frac{\beta_S J_0 N}{2}\sum_a\left[\left(l_a^I\right)^2+\left(m_a^I\right)^2+\left(n_a^I\right)^2+\left(l_a^J\right)^2+\left(m_a^J\right)^2+\left(n_a^J\right)^2\right]+\\
&-\frac{\beta_D J_0 N}{2}\sum_a\left[\left(l_a^{IJ}\right)^2+\left(m_a^I\right)^2+\left(n_a^J\right)^2+\left(l_a^{JI}\right)^2+\left(m_a^J\right)^2+\left(n_a^I\right)^2\right]+\\
&-\frac{\beta_S^2 J^2 N}{2}\sum_{a<b}\left[\left(p_{ab}^I\right)^2+\left(q_{ab}^I\right)^2+\left(r_{ab}^I\right)^2+\left(p_{ab}^J\right)^2+\left(q_{ab}^J\right)^2+\left(r_{ab}^J\right)^2\right]+\\
&\left.-\frac{\beta_D^2 J^2 N}{2}\sum_{a<b}\left[\left(p_{ab}^{IJ}\right)^2+\left(q_{ab}^I\right)^2+\left(r_{ab}^J\right)^2+\left(p_{ab}^{JI}\right)^2+\left(q_{ab}^J\right)^2+\left(r_{ab}^I\right)^2\right]\right\},\end{aligned} \tag{B.21}$$

where the parameters now represent the saddle point of the exponent.

Now, as we are interested in the limit $\delta\to 0$, we linearize around $\delta=0$ and rearrange the terms in the form of (3.13),

$$\begin{aligned}\left\langle \log Z_A\right\rangle_{\boldsymbol{J}} =&\lim_{\delta\to 0}\left\{\frac{1}{2}J^2N\left[\beta_S^2\left(\gamma^2+(1-\gamma)^2\right)+\beta_D^2 2\gamma\left(1-\gamma\right)\right]+\frac{1}{\delta}\sum_i\log\zeta_i+\right.\\
&-\frac{\beta_S J_0 N}{2\delta}\sum_a\left[\left(l_a^I\right)^2+\left(m_a^I\right)^2+\left(n_a^I\right)^2+\left(l_a^J\right)^2+\left(m_a^J\right)^2+\left(n_a^J\right)^2\right]+\\
&-\frac{\beta_D J_0 N}{2\delta}\sum_a\left[\left(l_a^{IJ}\right)^2+\left(m_a^I\right)^2+\left(n_a^J\right)^2+\left(l_a^{JI}\right)^2+\left(m_a^J\right)^2+\left(n_a^I\right)^2\right]+\\
&-\frac{\beta_S^2 J^2 N}{2\delta}\sum_{a<b}\left[\left(p_{ab}^I\right)^2+\left(q_{ab}^I\right)^2+\left(r_{ab}^I\right)^2+\left(p_{ab}^J\right)^2+\left(q_{ab}^J\right)^2+\left(r_{ab}^J\right)^2\right]+\\
&\left.-\frac{\beta_D^2 J^2 N}{2\delta}\sum_{a<b}\left[\left(p_{ab}^{IJ}\right)^2+\left(q_{ab}^I\right)^2+\left(r_{ab}^J\right)^2+\left(p_{ab}^{JI}\right)^2+\left(q_{ab}^J\right)^2+\left(r_{ab}^I\right)^2\right]\right\}.\end{aligned} \tag{B.22}$$

Before we proceed, we perform the following change of variables:

$$m_a^{I(J)} \to im_a^{I(J)}, \tag{B.23}$$
$$n_a^{I(J)} \to in_a^{I(J)}, \tag{B.24}$$
$$q_{ab}^{I(J)} \to iq_{ab}^{I(J)}, \tag{B.25}$$
$$r_{ab}^{I(J)} \to ir_{ab}^{I(J)}. \tag{B.26}$$

This is a necessary step in order to ensure that the order parameters will be real.

With this, we can now set the derivative of $\left\langle\log Z_A\right\rangle_{\boldsymbol{J}}$ with respect to each parameter equal to zero,

$$l_a^I = \frac{1}{N}\sum_i \left(\eta_i \langle x_i^a\rangle_i^{\text{rep}} + \eta_i \langle y_i^a\rangle_i^{\text{rep}}\right) \qquad m_a^I = \frac{1}{N}\sum_i \eta_i \langle x_i^a\rangle_i^{\text{rep}} \tag{B.27}$$

$$l_a^J = \frac{1}{N}\sum_i \left((1-\eta_i) \langle x_i^a\rangle_i^{\text{rep}} + (1-\eta_i) \langle y_i^a\rangle_i^{\text{rep}}\right) \qquad m_a^J = \frac{1}{N}\sum_i (1-\eta_i) \langle x_i^a\rangle_i^{\text{rep}} \tag{B.28}$$

$$l_a^{IJ} = \frac{1}{N}\sum_i \left(\eta_i \langle x_i^a\rangle_i^{\text{rep}} + (1-\eta_i) \langle y_i^a\rangle_i^{\text{rep}}\right) \qquad n_a^I = \frac{1}{N}\sum_i \eta_i \langle y_i^a\rangle_i^{\text{rep}} \tag{B.29}$$

$$l_a^{JI} = \frac{1}{N}\sum_i \left((1-\eta_i) \langle x_i^a\rangle_i^{\text{rep}} + \eta_i \langle y_i^a\rangle_i^{\text{rep}}\right) \qquad n_a^J = \frac{1}{N}\sum_i (1-\eta_i) \langle y_i^a\rangle_i^{\text{rep}} \tag{B.30}$$

$$p_{ab}^I = \frac{1}{N}\sum_i \left(\eta_i \left\langle x_i^a x_i^b\right\rangle_i^{\text{rep}} + \eta_i \left\langle y_i^a y_i^b\right\rangle_i^{\text{rep}}\right) \qquad q_{ab}^I = \frac{1}{N}\sum_i \eta_i \left\langle x_i^a x_i^b\right\rangle_i^{\text{rep}} \tag{B.31}$$

$$p_{ab}^J = \frac{1}{N}\sum_i \left((1-\eta_i) \left\langle x_i^a x_i^b\right\rangle_i^{\text{rep}} + (1-\eta_i) \left\langle y_i^a y_i^b\right\rangle_i^{\text{rep}}\right) \qquad q_{ab}^J = \frac{1}{N}\sum_i (1-\eta_i) \left\langle x_i^a x_i^b\right\rangle_i^{\text{rep}} \tag{B.32}$$

$$p_{ab}^{IJ} = \frac{1}{N}\sum_i \left(\eta_i \left\langle x_i^a x_i^b\right\rangle_i^{\text{rep}} + (1-\eta_i) \left\langle y_i^a y_i^b\right\rangle_i^{\text{rep}}\right) \qquad r_{ab}^I = \frac{1}{N}\sum_i \eta_i \left\langle y_i^a y_i^b\right\rangle_i^{\text{rep}} \tag{B.33}$$

$$p_{ab}^{JI} = \frac{1}{N}\sum_i \left((1-\eta_i) \left\langle x_i^a x_i^b\right\rangle_i^{\text{rep}} + \eta_i \left\langle y_i^a y_i^b\right\rangle_i^{\text{rep}}\right) \qquad r_{ab}^J = \frac{1}{N}\sum_i (1-\eta_i) \left\langle y_i^a y_i^b\right\rangle_i^{\text{rep}} \tag{B.34}$$

In the above expressions, $\langle\cdot\rangle_i^{\text{rep}}$ denotes the average over the distribution associated with the i-th component replica interaction Hamiltonian, $\mathcal{H}_i^{\text{rep}}$, such that

$$\zeta_i = \sum_{\left\{x_i^a, y_i^a\right\}} e^{-\mathcal{H}_i^{\text{rep}}}. \tag{B.35}$$

Note that there are relations between the order parameters,

$$l_a^I = m_a^I + n_a^I \qquad l_a^J = m_a^J + n_a^J \tag{B.36}$$

$$l_a^{IJ} = m_a^I + n_a^J \qquad l_a^{JI} = m_a^J + n_a^I \tag{B.37}$$

$$p_{ab}^I = q_{ab}^I + r_{ab}^I \qquad p_{ab}^J = q_{ab}^J + r_{ab}^J \tag{B.38}$$

$$p_{ab}^{IJ} = q_{ab}^I + r_{ab}^J \qquad p_{ab}^{JI} = q_{ab}^J + r_{ab}^I \tag{B.39}$$

Substituting in our result,

$$\begin{aligned}\langle \log Z_A\rangle_{\boldsymbol{J}} = \lim_{\delta\to 0}\Bigg\{ & \frac{1}{2}J^2 N\left[\beta_S^2\left(\gamma^2 + (1-\gamma)^2\right) + \beta_D^2 2\gamma(1-\gamma)\right] + \frac{1}{\delta}\sum_i \log\zeta_i + \\ & - \frac{\beta_S J_0 N}{\delta}\sum_a \left(m_a^I n_a^I + m_a^J n_a^J\right) - \frac{\beta_D J_0 N}{\delta}\sum_a \left(m_a^I n_a^J + m_a^J n_a^I\right) + \\ & - \frac{\beta_S^2 J^2 N}{\delta}\sum_{a<b} \left(q_{ab}^I r_{ab}^I + q_{ab}^J r_{ab}^J\right) - \frac{\beta_D^2 J^2 N}{\delta}\sum_{a<b} \left(q_{ab}^I r_{ab}^J + q_{ab}^J r_{ab}^I\right)\Bigg\}\end{aligned} \tag{B.40}$$

$$\zeta_i = \sum_{\{x_i^a, y_i^a\}} \exp\left\{ \beta_S J_0 \sum_a \left[n_a^I \eta_i + n_a^J (1-\eta_i) \right] x_i^a + \beta_S J_0 \sum_a \left[m_a^I \eta_i + m_a^J (1-\eta_i) \right] y_i^a + \right.$$
$$+ \beta_D J_0 \sum_a \left[n_a^J \eta_i + n_a^I (1-\eta_i) \right] x_i^a + \beta_D J_0 \sum_a \left[m_a^J \eta_i + m_a^I (1-\eta_i) \right] y_i^a +$$
$$+ \beta_S^2 J^2 \sum_{a<b} \left[r_{ab}^I \eta_i + r_{ab}^J (1-\eta_i) \right] x_i^a x_i^b + \beta_S^2 J^2 \sum_{a<b} \left[q_{ab}^I \eta_i + q_{ab}^J (1-\eta_i) \right] y_i^a y_i^b +$$
$$\left. + \beta_D^2 J^2 \sum_{a<b} \left[r_{ab}^J \eta_i + r_{ab}^I (1-\eta_i) \right] x_i^a x_i^b + \beta_D^2 J^2 \sum_{a<b} \left[q_{ab}^J \eta_i + q_{ab}^I (1-\eta_i) \right] y_i^a y_i^b \right\} \tag{B.41}$$

We note that the exponent in the expression for ζ_i is the same for every i when we take $\beta_S = \beta_D$. In this case, the averages $\langle x_i^a \rangle_i^{\mathrm{rep}}$, $\langle y_i^a \rangle_i^{\mathrm{rep}}$, $\langle x_i^a x_i^b \rangle_i^{\mathrm{rep}}$ and $\langle y_i^a y_i^b \rangle_i^{\mathrm{rep}}$ are the same for all i, and if we call them m_a, n_a, q_{ab} and r_{ab}, respectively, we have:

$$m_a^I = \gamma m_a \qquad m_a^J = (1-\gamma)\, m_a \tag{B.42}$$
$$n_a^I = \gamma n_a \qquad n_a^J = (1-\gamma)\, n_a \tag{B.43}$$
$$q_{ab}^I = \gamma q_{ab} \qquad q_{ab}^J = (1-\gamma)\, q_{ab} \tag{B.44}$$
$$r_{ab}^I = \gamma r_{ab} \qquad r_{ab}^J = (1-\gamma)\, r_{ab} \tag{B.45}$$

Now, back to a general β_S and β_D, to proceed we consider the replica symmetric ansatz, just like we did with the LSK model:

$$m_a^I = m^I \qquad m_a^J = m^J \tag{B.46}$$
$$n_a^I = n^I \qquad n_a^J = n^J \tag{B.47}$$
$$q_{ab}^I = q^I \qquad q_{ab}^J = q^J \tag{B.48}$$
$$r_{ab}^I = r^I \qquad r_{ab}^J = r^J \tag{B.49}$$

In this ansatz we have

$$\zeta_i^{\mathrm{RS}} = \sum_{\{x_i^a, y_i^a\}} \exp\left\{ \beta_S J_0 \left[n^I \eta_i + n^J (1-\eta_i) \right] \left(\sum_a x_i^a \right) + \beta_S J_0 \left[m^I \eta_i + m^J (1-\eta_i) \right] \left(\sum_a y_i^a \right) + \right.$$
$$+ \beta_D J_0 \left[n^J \eta_i + n^I (1-\eta_i) \right] \left(\sum_a x_i^a \right) + \beta_D J_0 \left[m^J \eta_i + m^I (1-\eta_i) \right] \left(\sum_a y_i^a \right) +$$
$$+ \beta_S^2 J^2 \left[r^I \eta_i + r^J (1-\eta_i) \right] \left(\sum_{a<b} x_i^a x_i^b \right) + \beta_S^2 J^2 \left[q^I \eta_i + q^J (1-\eta_i) \right] \left(\sum_{a<b} y_i^a y_i^b \right) +$$
$$\left. + \beta_D^2 J^2 \left[r^J \eta_i + r^I (1-\eta_i) \right] \left(\sum_{a<b} x_i^a x_i^b \right) + \beta_D^2 J^2 \left[q^J \eta_i + q^I (1-\eta_i) \right] \left(\sum_{a<b} y_i^a y_i^b \right) \right\} \tag{B.50}$$

Lets define m_i^S and m_i^D as:

$$m_i^S = m^I \eta_i + m^J (1-\eta_i) = \begin{cases} m^I & \text{se } i \in I \\ m^J & \text{se } i \in J \end{cases} \tag{B.51}$$

$$m_i^D = m^J \eta_i + m^I (1-\eta_i) = \begin{cases} m^J & \text{se } i \in I \\ m^I & \text{se } i \in J \end{cases} \tag{B.52}$$

and with similar definitions for $n_i^{S(D)}$, $q_i^{S(D)}$ and $r_i^{S(D)}$, we write

$$\begin{aligned}\zeta_i^{\mathrm{RS}} = \sum_{\{x_i^a, y_i^a\}} \exp\Bigg\{ & J_0 \left(\beta_S n_i^S + \beta_D n_i^D\right) \left(\sum_a x_i^a\right) + J_0 \left(\beta_S m_i^S + \beta_D m_i^D\right) \left(\sum_a y_i^a\right) + \\ & + J^2 \left(\beta_S^2 r_i^S + \beta_D^2 r_i^D\right) \left(\sum_a x_i^a\right)^2 + J^2 \left(\beta_S^2 q_i^S + \beta_D^2 q_i^D\right) \left(\sum_a y_i^a\right)^2 + \\ & - \frac{J^2}{2} \left(\beta_S^2 r_i^S + \beta_D^2 r_i^D\right) \delta - \frac{J^2}{2} \left(\beta_S^2 q_i^S + \beta_D^2 q_i^D\right) \delta \Bigg\}\end{aligned} \tag{B.53}$$

which can be calculated by introducing gaussian integrals for the quadratic terms.

In the first order in δ this expression results in

$$\begin{aligned}\log \zeta_i^{\mathrm{RS}} = & -\frac{J^2 \delta}{2} \left[\beta_S^2 \left(q_i^S + r_i^S\right) + \beta_D^2 \left(q_i^D + r_i^D\right)\right] + \\ & + \delta \int Dz \log \left\{ 2\cosh \left[J_0 \left(\beta_S n_i^S + \beta_D n_i^D\right) + Jz\sqrt{\beta_S^2 r_i^S + \beta_D^2 r_i^D} \right] \right\} + \\ & + \delta \int Dz \log \left\{ 2\cosh \left[J_0 \left(\beta_S m_i^S + \beta_D m_i^D\right) + Jz\sqrt{\beta_S^2 q_i^S + \beta_D^2 q_i^D} \right] \right\}\end{aligned} \tag{B.54}$$

Substituting in our expression for $\langle \log Z_A \rangle_{\boldsymbol{J}}$, and taking the limit $\delta \to 0$:

$$\begin{aligned}\langle \log Z_A^{\mathrm{RS}} \rangle_{\boldsymbol{J}} = & \frac{1}{2} J^2 N \left[\beta_S^2 \left(\gamma^2 + (1-\gamma)^2\right) + \beta_D^2 2\gamma (1-\gamma) \right] + \\ & - \frac{1}{2} \beta_S^2 J^2 N \left[\gamma \left(q^I + r^I\right) + (1-\gamma)\left(q^J + r^J\right)\right] + \\ & - \frac{1}{2} \beta_D^2 J^2 N \left[(1-\gamma) \left(q^I + r^I\right) + \gamma \left(q^J + r^J\right)\right] + \\ & + \gamma N \int Dz \log \left\{ 2\cosh \left[J_0 \left(\beta_S n^I + \beta_D n^J\right) + Jz\sqrt{\beta_S^2 r^I + \beta_D^2 r^J} \right] \right\} + \\ & + \gamma N \int Dz \log \left\{ 2\cosh \left[J_0 \left(\beta_S m^I + \beta_D m^J\right) + Jz\sqrt{\beta_S^2 q^I + \beta_D^2 q^J} \right] \right\} + \\ & + (1-\gamma) N \int Dz \log \left\{ 2\cosh \left[J_0 \left(\beta_S n^J + \beta_D n^I\right) + Jz\sqrt{\beta_S^2 r^J + \beta_D^2 r^I} \right] \right\} + \\ & + (1-\gamma) N \int Dz \log \left\{ 2\cosh \left[J_0 \left(\beta_S m^J + \beta_D m^I\right) + Jz\sqrt{\beta_S^2 q^J + \beta_D^2 q^I} \right] \right\} + \\ & - \beta_S J_0 N \left(m^I n^I + m^J n^J\right) - \beta_D J_0 N \left(m^I n^J + m^J n^I\right) + \\ & + \frac{1}{2} \beta_S^2 J^2 N \left(q^I r^I + q^J r^J\right) + \frac{1}{2} \beta_D^2 J^2 N \left(q^I r^J + q^J r^I\right)\end{aligned} \tag{B.55}$$

Now we take the derivative with respect to β_D and calculate at $\beta_S = \beta_D = \beta$. After some straightforward calculations we finally get

$$\phi_A = \beta \frac{\partial}{\partial \beta_D} \langle \log Z_A^{\mathrm{RS}} \rangle_{\boldsymbol{J}} \Big|_{\beta_S = \beta_D = \beta} = 2\gamma (1-\gamma) N \left[\beta J_0 m n + \beta^2 J^2 (1 - qr)\right] \tag{B.56}$$

where we used the fact that for $\beta_S = \beta_D$ the equations (B.42) to (B.45) holds, and for the replica symmetric ansatz they don't depend on the replica index.

B.2 Calculation of ϕ_B

To calculate ϕ_B we define a auxiliary probability distribution given by

$$P_B\left(X|\boldsymbol{J},\lambda\right) = \frac{1}{Z_B}\exp\left\{\lambda\sum_i \log 2\cosh\left(\beta\sum_j J_{ij}x_j\right)\right\}, \tag{B.57}$$

with disorder variables $\boldsymbol{J}$ distributed accordingly to equation (3.11).

Taking $\lambda = 1$ the probability P_B becomes the Little probability distribution marginalized over the future, $P\left(X|\boldsymbol{J}\right) = \sum_Y P\left(X,Y|\boldsymbol{J}\right)$.

Taking the derivative of $\langle\log Z_B\rangle_{\boldsymbol{J}}$ with respect to λ and calculating at $\lambda = 1$,

$$\left.\frac{\partial}{\partial\lambda}\langle\log Z_B\rangle_{\boldsymbol{J}}\right|_{\lambda=1} = \left\langle\sum_{X,Y} P\left(X,Y|\boldsymbol{J}\right)\sum_i \log 2\cosh\left(\beta\sum_j J_{ij}x_j\right)\right\rangle_{\boldsymbol{J}}. \tag{B.58}$$

Note that the RHS is equal to $-\phi_B$,

$$\phi_B = -\left.\frac{\partial}{\partial\lambda}\langle\log Z_B\rangle_{\boldsymbol{J}}\right|_{\lambda=1}. \tag{B.59}$$

Our work then is to calculate $\langle\log Z_B\rangle_{\boldsymbol{J}}$. To do this we use the replica method. So first we calculate $\left\langle Z_B^\delta\right\rangle_{\boldsymbol{J}}$ for integer δ:

$$\left\langle Z_B^\delta\right\rangle_{\boldsymbol{J}} = \left\langle\sum_{\{X^a\}}\exp\left\{\lambda\sum_{i,a}\log 2\cosh\left(\beta\sum_j J_{ij}x_j^a\right)\right\}\right\rangle_{\boldsymbol{J}}. \tag{B.60}$$

Introducing an integral over a Dirac delta distribution,

$$\left\langle Z_B^\delta\right\rangle_{\boldsymbol{J}} = \left\langle\sum_{\{X^a\}}\int\prod_{i,a} dw_i^a\,\delta\left(w_i^a - i\beta\sum_j J_{ij}x_j^a\right)\exp\left\{\lambda\sum_{i,a}\log 2\cosh\left(-iw_i^a\right)\right\}\right\rangle_{\boldsymbol{J}}. \tag{B.61}$$

Using the Fourier representation of the delta distribution,

$$\left\langle Z_B^\delta\right\rangle_{\boldsymbol{J}} = \left\langle\sum_{\{X^a\}}\int\prod_{i,a}\frac{dw_i^a d\hat{w}_i^a}{2\pi}\exp\left\{i\sum_{i,a}\hat{w}_i^a w_i^a + \beta\sum_{i,a}\hat{w}_i^a\sum_j J_{ij}x_i^a + \lambda\sum_{i,a}\log 2\cosh\left(-iw_i^a\right)\right\}\right\rangle_{\boldsymbol{J}}. \tag{B.62}$$

Calculating the average over the disorder variables,

$$\begin{aligned}\left\langle Z_B^\delta\right\rangle_{\boldsymbol{J}} = \sum_{\{X^a\}}\int\prod_{i,a}\frac{dw_i^a d\hat{w}_i^a}{2\pi}\exp\Bigg\{&i\sum_{i,a}\hat{w}_i^a w_i^a + \lambda\sum_{i,a}\log 2\cosh\left(-iw_i^a\right) + \\ &+\frac{\beta J_0}{N}\sum_{i,j}\sum_a x_j^a\hat{w}_i^a + \frac{\beta^2 J^2}{2N}\sum_{i,j}\left(\sum_a x_j^a\hat{w}_i^a\right)^2\Bigg\}.\end{aligned} \tag{B.63}$$

Rearranging the terms and introducing new integrals over Dirac delta distributions for the order parameters,

$$\left\langle Z_B^\delta\right\rangle_{\boldsymbol{J}} = \sum_{\{X^a\}} \int \prod_{i,a} \frac{dw_i^a d\hat{w}_i^a}{2\pi} \int \prod_a dm_a du_a dv_a \int \prod_{a<b} dq_{ab} ds_{ab} \times$$
$$\times \prod_a \delta\left(Nm_a - \sum_i x_i^a\right) \delta\left(Nu_a - \sum_i \hat{w}_i^a\right) \delta\left(Nv_a - \sum_i \left(\hat{w}_i^a\right)^2\right) \times$$
$$\times \prod_{a<b} \delta\left(Nq_{ab} - \sum_i x_i^a x_i^b\right) \delta\left(Ns_{ab} - \sum_i \hat{w}_i^a \hat{w}_i^b\right) \times$$
$$\times \exp\left\{ i\sum_{i,a} \hat{w}_i^a w_i^a + \lambda \sum_{i,a} \log 2\cosh\left(-iw_i^a\right) + \beta J_0 N \sum_a m_a u_a + \right.$$
$$\left. + \beta^2 J^2 N \sum_{a<b} q_{ab} s_{ab} + \frac{1}{2}\beta^2 J^2 N \sum_a v_a \right\}. \quad \text{(B.64)}$$

Using the Fourier representations just like before and grouping the terms that depends on $\{x_i^a\}$ and $\{w_i^a, \hat{w_i^a}\}$,

$$\left\langle Z_B^\delta\right\rangle_{\boldsymbol{J}} = \int \prod_a \frac{dm_a d\hat{m}_a}{2\pi/N} \int \prod_a \frac{du_a d\hat{u}_a}{2\pi/N} \int \prod_a \frac{dv_a d\hat{v}_a}{2\pi/N} \int \prod_{a<b} \frac{dq_{ab} d\hat{q}_{ab}}{2\pi/N} \int \prod_{a<b} \frac{ds_{ab} d\hat{s}_{ab}}{2\pi/N}$$
$$\times \exp\left\{ iN\sum_a \hat{m}_a m_a + iN\sum_a \hat{u}_a u_a + iN\sum_a \hat{v}_a v_a + iN\sum_{a<b} \hat{q}_{ab} q_{ab} + iN\sum_{a<b} \hat{s}_{ab} s_{ab} + \right.$$
$$\left. + \beta J_0 N \sum_a m_a u_a + \beta^2 J^2 N \sum_{a<b} q_{ab} s_{ab} + \frac{1}{2}\beta^2 J^2 N \sum_a v_a + N \log \mathcal{Z}_{\text{rep}} \right\}. \quad \text{(B.65)}$$

where $\mathcal{Z}_{\text{rep}}$ is a the partition function associated with the replica interaction Hamiltonian $\mathcal{H}_{\text{rep}}$,

$$\mathcal{Z}_{\text{rep}} = \sum_{\{x^a\}} \int \prod_a \frac{dw^a d\hat{w}^a}{2\pi} \exp\left\{-\mathcal{H}_{\text{rep}}\right\}, \quad \text{(B.66)}$$

$$\mathcal{H}_{\text{rep}} = -i\sum_a \hat{w}^a w^a - \lambda \sum_{i,a} \log 2\cosh\left(-iw_i^a\right) + i\sum_a \hat{m}_a x^a + i\sum_{a<b} \hat{q}_{ab} x^a x^b +$$
$$+ i\sum_a \hat{u}_a \hat{w}^a + i\sum_{a<b} \hat{s}_{ab} \hat{w}^a \hat{w}^b + i\sum_a \hat{v}_a \left(\hat{w}^a\right)^2. \quad \text{(B.67)}$$

Using the saddle point method, we are able to calculate the integrals over the order parameters. Expanding the solution in first order in δ and rearranging the terms in the same way as the replica identity (equation (3.13)) we have:

$$\langle \log Z_B \rangle_{\boldsymbol{J}} = \lim_{\delta\to 0} \left\{ \frac{iN}{\delta} \sum_a \hat{m}_a m_a + \frac{iN}{\delta} \sum_a \hat{u}_a u_a + \frac{iN}{\delta} \sum_a \hat{v}_a v_a + \frac{iN}{\delta} \sum_{a<b} \hat{q}_{ab} q_{ab} + \frac{iN}{\delta} \sum_{a<b} \hat{s}_{ab} s_{ab} + \right.$$
$$\left. + \frac{\beta J_0 N}{\delta} \sum_a m_a u_a + \frac{\beta^2 J^2 N}{\delta} \sum_{a<b} q_{ab} s_{ab} - \frac{\beta J^2 N}{2\delta} \sum_a v_a + N \log \mathcal{Z}_{\text{rep}} \right\}, \quad \text{(B.68)}$$

where the order parameters are now the saddle points.

Setting the derivative of $\langle \log Z_B \rangle_{\boldsymbol{J}}$ with respect to to each parameter equal to zero, we have the equations for the saddle points:

$$\hat{m}_a = i\beta J_0 u_a \qquad m_a = \langle x^a \rangle_{\text{rep}} \tag{B.69}$$

$$\hat{u}_a = i\beta J_0 m_a \qquad u_a = \langle w^a \rangle_{\text{rep}} \tag{B.70}$$

$$\hat{q}_{ab} = i\beta^2 J^2 s_{ab} \qquad q_{ab} = \left\langle x^a x^b \right\rangle_{\text{rep}} \tag{B.71}$$

$$\hat{s}_{ab} = i\beta^2 J^2 q_{ab} \qquad s_{ab} = \left\langle w^a w^b \right\rangle_{\text{rep}} \tag{B.72}$$

$$\hat{v}_a = \frac{i}{2}\beta^2 J^2 \qquad v_a = \left\langle (w^a)^2 \right\rangle_{\text{rep}} \tag{B.73}$$

where, again, $\langle \cdot \rangle_{\text{rep}}$ denotes the average over the canonical distribution associated with the replica interaction Hamiltonian.

With these relations, we can write

$$\langle \log Z_B \rangle_{\boldsymbol{J}} = \lim_{\delta \to 0} \left\{ -\frac{\beta J_0 N}{\delta} \sum_a m_a u_a - \frac{\beta^2 J^2 N}{\delta} \sum_{a<b} q_{ab} s_{ab} + N \log \mathcal{Z}_{\text{rep}} \right\}, \tag{B.74}$$

and

$$\mathcal{Z}_{\text{rep}} = \sum_{\{x^a\}} \exp\left\{ \beta J_0 \sum_a u_a x^a + \beta^2 J^2 \sum_{a<b} s_{ab} x^a x^b \right\} \int \prod_a \frac{dw^a d\hat{w}^a}{2\pi} \exp\left\{ i \sum_a \hat{w}^a w^a + \right.$$
$$\left. + \lambda \sum_a \log 2\cosh\left(-i\hat{w}^a\right) + \beta J_0 \sum_a m_a w^a + \beta^2 J^2 \sum_{a<b} q_{ab} w^a w^b + \frac{1}{2}\beta^2 J^2 \sum_a (w^a)^2 \right\}. \tag{B.75}$$

Now we consider the replica symmetric ansatz, where the saddle points are symmetreic under replica index permutation:

$$m_a = m \tag{B.76}$$

$$u_a = u \tag{B.77}$$

$$q_{ab} = q \tag{B.78}$$

$$s_{ab} = s \tag{B.79}$$

In this ansatz the replica interaction pertition function becomes

$$\mathcal{Z}_{\text{RS}} = \sum_{\{x^a\}} \exp\left\{ \beta J_0 u \sum_a x^a + \beta^2 J^2 s \sum_{a<b} x^a x^b \right\} \int \prod_a \frac{dw^a d\hat{w}^a}{2\pi} \exp\left\{ i \sum_a \hat{w}^a w^a + \right.$$
$$\left. + \lambda \sum_a \log 2\cosh\left(-i\hat{w}^a\right) + \beta J_0 m \sum_a w^a + \beta^2 J^2 q \sum_{a<b} w^a w^b + \frac{1}{2}\beta^2 J^2 \sum_a (w^a)^2 \right\}. \tag{B.80}$$

The summation over $\{x^a\}$ can be easily performed, yielding

$$\exp\left\{ -\frac{1}{2}\beta^2 J^2 s \delta \right\} \int Dz \left(2\cosh\left[\beta \left(J_0 u + J\sqrt{s} z \right) \right] \right)^{\delta}, \tag{B.81}$$

where Dz is the normalized gaussian measure, with zero mean and unit variance.

Taking the logarithm:

$$\log \mathcal{Z}_{\mathrm{RS}} = -\frac{1}{2}\beta^2 J^2 s\delta + \log \int Dz \left(2\cosh\left[\beta\left(J_0 u + J\sqrt{s}z\right)\right]\right)^{\delta} + \\ + \log \int \prod_a \frac{dw^a d\hat{w}^a}{2\pi} \exp\left\{ i\sum_a \hat{w}^a w^a + \lambda \sum_a \log 2\cosh\left(i\hat{w}^a\right) + \right. \\ \left. + \beta J_0 m \sum_a w^a + \frac{1}{2}\beta^2 J^2 q \left(\sum_a w^a\right)^2 + \frac{1}{2}\beta^2 J^2 (1-q) \sum_a \left(w^a\right)^2 \right\}. \tag{B.82}$$

Introducing yet another gaussian integral for the remaining quadratic term in the exponent and factoring the integrals in the index a,

$$\log \mathcal{Z}_{\mathrm{RS}} = -\frac{1}{2}\beta^2 J^2 s\delta + \log \int Dz \left(2\cosh\left[\beta\left(J_0 u + J\sqrt{s}z\right)\right]\right)^{\delta} + \\ + \log \int Dz \left[\int \frac{dw d\hat{w}}{2\pi} \exp\left\{ i\hat{w}w + \lambda \log 2\cosh\left(i\hat{w}\right) + \right.\right. \\ \left.\left. + \beta\left(J_0 m + J\sqrt{q}z\right) w + \frac{1}{2}\beta^2 J^2 (1-q) w^2 \right\} \right]^{\delta}. \tag{B.83}$$

Now we linearize in δ around $\delta = 0$,

$$\log \mathcal{Z}^{\mathrm{RS}}_{\mathrm{rep}} = -\frac{1}{2}\beta^2 J^2 s\delta + \delta \int Dz \log 2\cosh\left[\beta\left(J_0 u + J\sqrt{s}z\right)\right] + \\ + \delta \int Dz \log \int \frac{dw d\hat{w}}{2\pi} \exp\left\{ i\hat{w}w + \lambda \log 2\cosh\left(i\hat{w}\right) + \right. \\ \left. + \beta\left(J_0 m + J\sqrt{q}z\right) w + \frac{1}{2}\beta^2 J^2 (1-q) w^2 \right\}, \tag{B.84}$$

and we finally have an expression for $\langle \log Z_B \rangle_{\boldsymbol{J}}$:

$$\left\langle \log Z_B^{\mathrm{RS}} \right\rangle_{\boldsymbol{J}} = -\beta J_0 N m u + \frac{1}{2}\beta^2 J^2 N q s - \frac{1}{2}\beta^2 J^2 N s + N \int Dz \log 2\cosh\left[\beta\left(J_0 u + J\sqrt{s}z\right)\right] + \\ + N \int Dz \log \int \frac{dw d\hat{w}}{2\pi} \exp\left\{ i\hat{w}w + \lambda \log 2\cosh\left(i\hat{w}\right) + \right. \\ \left. + \beta\left(J_0 m + J\sqrt{q}z\right) w + \frac{1}{2}\beta^2 J^2 (1-q) w^2 \right\} \tag{B.85}$$

Now we are able to take the derivative with repect to λ and set $\lambda = 1$,

$$\phi_B = -\frac{\partial}{\partial\lambda} \left\langle \log Z_B^{\mathrm{RS}} \right\rangle_{\boldsymbol{J}} \Big|_{\lambda=1} = -N \int Dz \frac{\exp\left\{-\frac{1}{2}\beta^2 J^2 (1-q)\right\}}{2\cosh\left[\beta\left(J_0 m + J\sqrt{q}z\right)\right]} \times \\ \times \int \frac{dw d\hat{w}}{2\pi} \left[\log 2\cosh\left(i\hat{w}\right)\right] \exp\left\{ i\hat{w}w + \log 2\cosh\left(i\hat{w}\right) + \right. \\ \left. + \beta\left(J_0 m + J\sqrt{q}z\right) w + \frac{1}{2}\beta^2 J^2 (1-q) w^2 \right\} \tag{B.86}$$

After some manipulation we find that this can be written as:

$$\phi_B = -N\int Dz \sum_{y\in\{-1,1\}} \frac{\exp\left\{\beta\left(J_0 m + J\sqrt{q}z\right)y\right\}}{2\cosh\left[\beta\left(J_0 m + J\sqrt{q}z\right)\right]}\times$$
$$\times\int D\zeta \log 2\cosh\left[\beta\left(J_0 m + J\sqrt{q}z\right) + \beta^2 J^2\left(1-q\right)y + \beta J\zeta\sqrt{1-q}\right] \quad \text{(B.87)}$$

where $D\zeta$ is also a gaussian integral with unit variance and vanishing mean.

B.3 Calculation of ϕ_C

The third and last term we need to calculate is

$$\phi_C = \left\langle \sum_{X,Y} P\left(X,Y|\boldsymbol{J}\right)\sum_i \log 2\cosh\left(\beta h_{i|S}\right)\right\rangle_{\boldsymbol{J}}, \quad \text{(B.88)}$$

where $h_{i|S}$ is, again, the field defined as

$$h_{i|S} = \begin{cases} \sum_j J_{ij}\eta_j x_j & \text{para} \quad i\in I \\ \\ \sum_j J_{ij}\left(1-\eta_j\right)x_j & \text{para} \quad i\in J \end{cases} \quad \text{(B.89)}$$

To do so, we consider the following probability distribution:

$$P_C\left(X,Y|\boldsymbol{J};\lambda_S\right) = \frac{1}{Z_C}\exp\left\{\beta\sum_{i,j}J_{ij}x_j y_i + \lambda_S\sum_i \log 2\cosh\left(\beta h_{i|S}\right)\right\}. \quad \text{(B.90)}$$

Taking the derivative of $\langle \log Z_C\rangle_{\boldsymbol{J}}$ with respect to λ_S and calculating at $\lambda_S = 0$

$$\frac{\partial}{\partial\lambda_S}\left\langle \log Z_C\right\rangle_{\boldsymbol{J}}\Big|_{\lambda_S=0} = \left\langle \frac{1}{Z}\sum_{X,Y}\left[\sum_i \log 2\cosh\left(\beta h_{i|S}\right)\right]\exp\left\{\beta\sum_{i,j}J_{ij}x_j y_i\right\}\right\rangle_{\boldsymbol{J}}, \quad \text{(B.91)}$$

we recover the average we are interested in.

Therefore we need to calculate the free energy for this auxiliary distribution. We again turn to the replica method, where we first calculate:

$$\left\langle Z_C^\delta\right\rangle_{\boldsymbol{J}} = \left\langle \sum_{\{X^a,Y^a\}}\exp\left\{\beta\sum_{i,j}\sum_a J_{ij}x_j^a y_i^a + \lambda_S\sum_{i,a}\log 2\cosh\left(\beta\sum_j \theta_{ij}^S J_{ij}x_j^a\right)\right\}\right\rangle_{\boldsymbol{J}} \quad \text{(B.92)}$$

where $\theta_{ij}^S = \eta_i\eta_j + \left(1-\eta_i\right)\left(1-\eta_j\right)$ is the same projector we used for the calculation of ϕ_A, and, together with $\theta_{ij}^D = \eta_i\left(1-\eta_j\right) + \left(1-\eta_i\right)\eta_j$, have the properties (B.10) to (B.13).

Introducing an integral over a Dirac delta distribution:

$$\left\langle Z_C^\delta\right\rangle_{\boldsymbol{J}} = \left\langle \sum_{\{X^a,Y^a\}}\int\prod_{i,a}\delta\left(w_i^a - i\beta\sum_j\theta_{ij}^S J_{ij}x_j^a\right)dw_i^a\right.$$
$$\left.\times\exp\left\{\beta\sum_{i,j}J_{ij}\sum_a x_j^a y_i^a + \lambda_S\sum_{i,a}\log 2\cosh\left(-iw_i^a\right)\right\}\right\rangle_{\boldsymbol{J}}. \quad \text{(B.93)}$$

Using the Fourier representation of the delta:

$$\left\langle Z_C^\delta \right\rangle_{\boldsymbol{J}} = \sum_{\{X^a,Y^a\}} \int \prod_{i,a} \frac{dw_i^a d\hat{w}_i^a}{2\pi} \left\langle \exp\left\{ \beta \sum_{i,j} J_{ij} \sum_a x_j^a y_i^a + \lambda_S \sum_{i,a} \log 2\cosh\left(-iw_i^a\right) + \right.\right.$$
$$\left.\left. + i\sum_{i,a} \hat{w}_i^a w_i^a + \beta \sum_{ij} \theta_{ij}^S J_{ij} \sum_a x_j^a \hat{w}_i^a \right\} \right\rangle_{\boldsymbol{J}}. \tag{B.94}$$

Taking the average over the disrder variables $\boldsymbol{J}$

$$\left\langle Z_C^\delta \right\rangle_{\boldsymbol{J}} = \sum_{\{X^a,Y^a\}} \int \prod_{i,a} \frac{dw_i^a d\hat{w}_i^a}{2\pi} \exp\left\{ i\sum_{i,a} \hat{w}_i^a w_i^a + \lambda_S \sum_{i,a} \log 2\cosh\left(-iw_i^a\right) + \right.$$
$$\left. + \frac{\beta J_0}{N} \sum_{ij} \sum_a x_j^a y_i^a + \frac{\beta J_0}{N} \sum_{ij} \theta_{ij}^S \sum_a x_j^a \hat{w}_i^a + \frac{\beta^2 J^2}{2N} \sum_{ij} \left[\sum_a x_j^a \left(y_i^a + \theta_{ij}^S \hat{w}_i^a \right) \right]^2 \right\}. \tag{B.95}$$

Rearranging the terms:

$$\begin{aligned}
\left\langle Z_C^\delta \right\rangle_{\boldsymbol{J}} = & \sum_{\{X^a,Y^a\}} \int \prod_{i,a} \frac{dw_i^a d\hat{w}_i^a}{2\pi} \exp\left\{ i\sum_{i,a} \hat{w}_i^a w_i^a + \lambda_S \sum_{i,a} \log 2\cosh\left(-iw_i^a\right) + \right. \\
& + \frac{\beta J_0}{N} \sum_a \left(\sum_i x_i^a \right) \left(\sum_i y_i^a \right) + \frac{\beta J_0}{N} \sum_a \left(\sum_i \eta_i x_i^a \right) \left(\sum_i \eta_i \hat{w}_i^a \right) + \\
& + \frac{\beta J_0}{N} \sum_a \left(\sum_i (1-\eta_i) x_i^a \right) \left(\sum_i (1-\eta_i) \hat{w}_i^a \right) + \\
& + \frac{\beta^2 J^2}{N} \sum_{a<b} \left(\sum_i x_i^a x_i^b \right) \left(\sum_i y_i^a y_i^b \right) + \frac{1}{2}\beta^2 J^2 N \delta + \\
& + \frac{2\beta^2 J^2}{N} \sum_{a<b} \left(\sum_i \eta_i x_i^a x_i^b \right) \left(\sum_i \eta_i y_i^a \hat{w}_i^b \right) + \beta^2 J^2 \gamma \sum_{i,a} \eta_i y_i^a \hat{w}_i^a + \\
& + \frac{2\beta^2 J^2}{N} \sum_{a<b} \left(\sum_i (1-\eta_i) x_i^a x_i^b \right) \left(\sum_i (1-\eta_i) y_i^a \hat{w}_i^b \right) + \\
& + \beta^2 J^2 (1-\gamma) \sum_{i,a} (1-\eta_i) y_i^a \hat{w}_i^a + \\
& + \frac{\beta^2 J^2}{N} \sum_{a<b} \left(\sum_i \eta_i x_i^a x_i^b \right) \left(\sum_i \eta_i \hat{w}_i^a \hat{w}_i^b \right) + \frac{1}{2}\beta^2 J^2 \gamma \sum_{i,a} \eta_i \left(\hat{w}_i^a\right)^2 + \\
& + \frac{\beta^2 J^2}{N} \sum_{a<b} \left(\sum_i (1-\eta_i) x_i^a x_i^b \right) \left(\sum_i (1-\eta_i) \hat{w}_i^a \hat{w}_i^b \right) + \\
& \left. + \frac{1}{2}\beta^2 J^2 (1-\gamma) \sum_{i,a} (1-\eta_i) \left(\hat{w}_i^a\right)^2 \right\}.
\end{aligned} \tag{B.96}$$

where we define $\gamma = \sum_i \eta_i$. That is, γ is the fraction of elements in the partition I (and $1-\gamma$ is the fraction of elements in the partition J).

Again we introduce integrals over delta distributions and use its Fourier representation:

$$\begin{aligned}\langle Z_C^\delta \rangle_{\boldsymbol{J}} =& \int \prod_a \frac{dm_a d\hat{m}_a}{2\pi/N} \int \prod_a \frac{dm_a^I d\hat{m}_a^I}{2\pi/N} \int \prod_a \frac{dm_a^J d\hat{m}_a^J}{2\pi/N} \int \prod_a \frac{dn_a d\hat{n}_a}{2\pi/N} \int \prod_a \frac{du_a^I d\hat{u}_a^I}{2\pi/N} \int \prod_a \frac{du_a^J d\hat{u}_a^J}{2\pi/N} \times \\ & \times \int \prod_{a<b} \frac{dq_{ab} d\hat{q}_{ab}}{2\pi/N} \int \prod_{a<b} \frac{dq_{ab}^I d\hat{q}_{ab}^I}{2\pi/N} \int \prod_{a<b} \frac{dq_{ab}^J d\hat{q}_{ab}^J}{2\pi/N} \int \prod_{a<b} \frac{dr_{ab} d\hat{r}_{ab}}{2\pi/N} \int \prod_{a<b} \frac{ds_{ab}^I d\hat{s}_{ab}^I}{2\pi/N} \int \prod_{a<b} \frac{ds_{ab}^J d\hat{s}_{ab}^J}{2\pi/N} \times \\ & \times \int \prod_{a<b} \frac{dt_{ab}^I d\hat{t}_{ab}^I}{2\pi/N} \int \prod_{a<b} \frac{dt_{ab}^J d\hat{t}_{ab}^J}{2\pi/N} \int \prod_a \frac{dv_a^I d\hat{v}_a^I}{2\pi/N} \int \prod_a \frac{dv_a^J d\hat{v}_a^J}{2\pi/N} \int \prod_a \frac{d\tau_a^I d\hat{\tau}_a^I}{2\pi/N} \int \prod_a \frac{d\tau_a^J d\hat{\tau}_a^J}{2\pi/N} \times \\ & \times \exp\left\{ \beta J_0 N \sum_a \left(m_a n_a + m_a^I u_a^I + m_a^J \right) + \beta^2 J^2 N \sum_{a<b} \left(q_{ab} r_{ab} + 2 q_{ab}^I t_{ab}^I + 2 q_{ab}^J t_{ab}^J + q_{ab}^I s_{ab}^I + q_{ab}^J s_{ab}^J \right) + \right. \\ & + \frac{1}{2} \beta^2 J^2 N \sum_a \left[2\gamma \tau_a^I + 2(1-\gamma) \tau_a^J + \gamma v_a^I + (1-\gamma) v_a^J \right] + \frac{1}{2} \beta^2 J^2 N \delta + \\ & + iN \sum_a \left(m_a \hat{m}_a + m_a^I \hat{m}_a^I + m_a^J \hat{m}_a^J + n_a \hat{n}_a + u_a^I \hat{u}_a^I + u_a^J \hat{u}_a^J \right) + \\ & + iN \sum_{a<b} \left(q_{ab} \hat{q}_{ab} + q_{ab}^I \hat{q}_{ab}^I + q_{ab}^J \hat{q}_{ab}^J + r_{ab} \hat{r}_{ab} + s_{ab}^I \hat{s}_{ab}^I + s_{ab}^J \hat{s}_{ab}^J + t_{ab}^I \hat{t}_{ab}^I + t_{ab}^J \hat{t}_{ab}^J \right) + \\ & \left. + iN \sum_a \left(v_a^I \hat{v}_a^I + v_a^J \hat{v}_a^J + \tau_a^I \hat{\tau}_a^I + \tau_a^J \hat{\tau}_a^J \right) + \sum_i \log \zeta_i \right\}, \end{aligned} \tag{B.97}$$

where ζ_i is the partition function associated with the interaction between the i-th component of the replicas, and it is given by:

$$\begin{aligned}\zeta_i =& \sum_{\{x_i^a, y_i^a\}} \int \prod_a \frac{dw_i^a d\hat{w}_i^a}{2\pi} \exp\left\{ i \sum_a \hat{w}_i^a w_i^a + \lambda_S \sum_a \log 2\cosh\left(-i w_i^a\right) + \right. \\ & + -i \sum_a \left[\hat{m}_a x_i^a + \hat{m}_a^I \eta_i x_i^a + \hat{m}_a^J (1-\eta_i) x_i^a + \hat{n}_a y_i^a + \hat{u}_a^I \eta_i \hat{w}_i^a + \hat{u}_a^J (1-\eta_i) \hat{w}_i^a \right] + \\ & + -i \sum_{a<b} \left[\hat{q}_{ab} x_i^a x_i^b + \hat{q}_{ab}^I \eta_i x_i^a x_i^b + \hat{q}_{ab}^J (1-\eta_i) x_i^a x_i^b + \hat{r}_{ab} y_i^a y_i^b + \right. \\ & + \left. \hat{s}_{ab}^I \eta_i \hat{w}_i^a \hat{w}_i^b + \hat{s}_{ab}^J (1-\eta_i) \hat{w}_i^a \hat{w}_i^b + \hat{t}_{ab}^I \eta_i t_i^a \hat{w}_i^b + \hat{t}_{ab}^J (1-\eta_i) y_i^a \hat{w}_i^b \right] + \\ & \left. + -i \sum_a \left[\hat{v}_a^I \eta_i \left(\hat{w}_i^a\right)^2 + \hat{v}_a^J (1-\eta_i) \left(\hat{w}_i^a\right)^2 + \hat{\tau}_a^I \eta_i y_i^a \hat{w}_i^a + \hat{\tau}_a^J (1-\eta_i) y_i^a \hat{w}_i^a \right] \right\}. \end{aligned} \tag{B.98}$$

Using the saddle point method we are able to calculate the integrals over the order parameters in the limit $N \to \infty$. Then, as we are interested in the limit $\delta \to 0$, we linearize the result around $\delta = 0$ and rearrange the terms in the form of (3.13):

$$\begin{aligned}
\langle \log Z_C \rangle_{\boldsymbol{J}} = \lim_{\delta \to 0} \Bigg\{ & \frac{\beta J_0 N}{\delta} \sum_a \left(m_a n_a + m_a^I u_a^I + m_a^J u_a^J \right) + \frac{1}{2} \beta^2 J^2 N + \\
& + \frac{\beta^2 J^2 N}{\delta} \sum_{a<b} \left(q_{ab} r_{ab} + 2 q_{ab}^I t_{ab}^I + 2 q_{ab}^J t_{ab}^J + q_{ab}^I s_{ab}^I + q_{ab}^J s_{ab}^J \right) + \\
& + \frac{\beta^2 J^2 N}{2\delta} \sum_a \left[2\gamma \tau_a^I + 2 (1-\gamma) \tau_a^J + \gamma v_a^I + (1-\gamma) v_a^J \right] + \\
& + i \frac{N}{\delta} \sum_a \left(m_a \hat{m}_a + m_a^I \hat{m}_a^I + m_a^J \hat{m}_a^J + n_a \hat{n}_a + u_a^I \hat{u}_a^I + u_a^J \hat{u}_a^J \right) + \\
& + i \frac{N}{\delta} \sum_{a<b} \left(q_{ab} \hat{q}_{ab} + q_{ab}^I \hat{q}_{ab}^I + q_{ab}^J \hat{q}_{ab}^J + r_{ab} \hat{r}_{ab} + s_{ab}^I \hat{s}_{ab}^I + s_{ab}^J \hat{s}_{ab}^J + t_{ab}^I \hat{t}_{ab}^I + t_{ab}^J \hat{t}_{ab}^J \right) + \\
& + i \frac{N}{\delta} \sum_a \left(v_a^I \hat{v}_a^I + v_a^J \hat{v}_a^J + \tau_a^I \hat{\tau}_a^I + \tau_a^J \hat{\tau}_a^J \right) + \frac{1}{\delta} \sum_i \log \zeta_i \Bigg\}
\end{aligned} \tag{B.99}$$

where the parameters represent the saddle points.

To find the saddle points, we take the derivative of $\langle \log Z \rangle_{\boldsymbol{J}}$ with respect to each parameter and set it equal to zero:

$$\hat{m}_a = i\beta J_0 n_a \qquad m_a = \frac{1}{N} \sum_i \langle x_i^a \rangle_i^{\text{rep}} \tag{B.100}$$

$$\hat{m}_a^I = i\beta J_0 u_a^I \qquad m_a^I = \frac{1}{N} \sum_i \eta_i \langle x_i^a \rangle_i^{\text{rep}} \tag{B.101}$$

$$\hat{m}_a^J = i\beta J_0 u_a^J \qquad m_a^J = \frac{1}{N} \sum_i (1-\eta_i) \langle x_i^a \rangle_i^{\text{rep}} \tag{B.102}$$

$$\hat{n}_a = i\beta J_0 m_a \qquad n_a = \frac{1}{N} \sum_i \langle y_i^a \rangle_i^{\text{rep}} \tag{B.103}$$

$$\hat{u}_a^I = i\beta J_0 m_a^I \qquad u_a^I = \frac{1}{N} \sum_i \eta_i \langle \hat{w}_i^a \rangle_i^{\text{rep}} \tag{B.104}$$

$$\hat{u}_a^J = i\beta J_0 m_a^J \qquad u_a^J = \frac{1}{N} \sum_i (1-\eta_i) \langle \hat{w}_i^a \rangle_i^{\text{rep}} \tag{B.105}$$

$$\hat{q}_{ab} = i\beta^2 J^2 r_{ab} \qquad q_{ab} = \frac{1}{N}\sum_i \left\langle x_i^a x_i^b \right\rangle_i^{\text{rep}} \tag{B.106}$$

$$\hat{q}_{ab}^I = i\beta^2 J^2 \left(s_{ab}^I + 2t_{ab}^I\right) \qquad q_{ab}^I = \frac{1}{N}\sum_i \eta_i \left\langle x_i^a x_i^b \right\rangle_i^{\text{rep}} \tag{B.107}$$

$$\hat{q}_{ab}^J = i\beta^2 J^2 \left(s_{ab}^J + 2t_{ab}^J\right) \qquad q_{ab}^J = \frac{1}{N}\sum_i (1-\eta_i) \left\langle x_i^a x_i^b \right\rangle_i^{\text{rep}} \tag{B.108}$$

$$\hat{r}_{ab} = i\beta^2 J^2 q_{ab} \qquad r_{ab} = \frac{1}{N}\sum_i \left\langle y_i^a y_i^b \right\rangle_i^{\text{rep}} \tag{B.109}$$

$$\hat{s}_{ab}^I = i\beta^2 J^2 q_{ab}^I \qquad s_{ab}^I = \frac{1}{N}\sum_i \eta_i \left\langle \hat{w}_i^a \hat{w}_i^b \right\rangle_i^{\text{rep}} \tag{B.110}$$

$$\hat{s}_{ab}^J = i\beta^2 J^2 q_{ab}^J \qquad s_{ab}^J = \frac{1}{N}\sum_i (1-\eta_i) \left\langle \hat{w}_i^a \hat{w}_i^b \right\rangle_i^{\text{rep}} \tag{B.111}$$

$$\hat{t}_{ab}^I = 2i\beta^2 J^2 q_{ab}^I \qquad t_{ab}^I = \frac{1}{N}\sum_i \eta_i \left\langle y_i^a \hat{w}_i^b \right\rangle_i^{\text{rep}} \tag{B.112}$$

$$\hat{t}_{ab}^J = 2i\beta^2 J^2 q_{ab}^J \qquad t_{ab}^J = \frac{1}{N}\sum_i (1-\eta_i) \left\langle y_i^a \hat{w}_i^b \right\rangle_i^{\text{rep}} \tag{B.113}$$

$$\hat{v}_a^I = \frac{i}{2}\beta^2 J^2 \gamma \qquad v_a^I = \frac{1}{N}\sum_i \eta_i \left\langle (\hat{w}_i^a)^2 \right\rangle_i^{\text{rep}} \tag{B.114}$$

$$\hat{v}_a^J = \frac{i}{2}\beta^2 J^2 (1-\gamma) \qquad v_a^J = \frac{1}{N}\sum_i (1-\eta_i) \left\langle (\hat{w}_i^a)^2 \right\rangle_i^{\text{rep}} \tag{B.115}$$

$$\hat{\tau}_{ab}^I = i\beta^2 J^2 \gamma \qquad \tau_a^I = \frac{1}{N}\sum_i \eta_i \left\langle y_i^a \hat{w}_i^a \right\rangle_i^{\text{rep}} \tag{B.116}$$

$$\hat{\tau}_a^J = i\beta^2 J^2 (1-\gamma) \qquad \tau_a^J = \frac{1}{N}\sum_i (1-\eta_i) \left\langle y_i^a \hat{w}_i^a \right\rangle_i^{\text{rep}} \tag{B.117}$$

where $\langle\cdot\rangle_i^{\text{rep}}$ denotes the average over the distribution associated with the i-th component replica interaction Hamiltonian, $\mathcal{H}_i^{\text{rep}}$, such that

$$\zeta_i = \sum_{\left\{x_i^a, y_i^a\right\}} \int \prod_a \frac{dw_i^a d\hat{w}_i^a}{2\pi} e^{-\mathcal{H}_i^{\text{rep}}}. \tag{B.118}$$

We note that some order parameters depends on each other. So we use these relationships to write:

$$\begin{aligned}\langle \log Z_C \rangle_{\boldsymbol{J}} =& \lim_{\delta\to 0}\Bigg\{ -\frac{\beta J_0 N}{\delta}\sum_a \left(m_a n_a + m_a^I u_a^I + m_a^J u_a^J\right) + \frac{1}{2}\beta^2 J^2 N + \frac{1}{\delta}\sum_i \log \zeta_i + \\ & - \frac{\beta^2 J^2 N}{\delta}\sum_{a<b}\left(q_{ab} r_{ab} + 2q_{ab}^I t_{ab}^I + 2q_{ab}^J t_{ab}^J + q_{ab}^I s_{ab}^I + q_{ab}^J s_{ab}^J\right)\Bigg\}\end{aligned} \tag{B.119}$$

and

$$\zeta_i = \sum_{\{x_i^a, y_i^a\}} \int \prod_a \frac{dw_i^a d\hat{w}_i^a}{2\pi} \exp\left\{ i \sum_a \hat{w}_i^a w_i^a + \lambda_S \sum_a \log 2\cosh\left(-i w_i^a\right) + \right.$$
$$+ \beta J_0 \sum_a \left[n_a + \eta_i u_a^I + (1-\eta_i) u_a^J\right] x_i^a + \beta J_0 \sum_a m_a y_i^a + \beta J_0 \sum_a \left[\eta_i m_a^I + (1-\eta_i) m_a^J\right] \hat{w}_i^a +$$
$$+ \beta^2 J^2 \sum_{a<b} \left[r_{ab} + \eta_i \left(s_{ab}^I + 2t_{ab}^I\right) + (1-\eta_i)\left(s_{ab}^J + 2t_{ab}^J\right)\right] x_i^a x_i^b + \beta^2 J^2 \sum_{a<b} q_{ab} y_i^a y_i^b +$$
$$+ \beta^2 J^2 \sum_{a<b} \left[\eta_i q_{ab}^I + (1-\eta_i) q_{ab}^J\right] \hat{w}_i^a \hat{w}_i^b + \beta^2 J^2 \sum_{a<b} \left[\eta_i q_{ab}^I + (1-\eta_i) q_{ab}^J\right] 2 y_i^a \hat{w}_i^b +$$
$$\left. + \frac{1}{2}\beta^2 J^2 \left[\eta_i \gamma - (1-\eta_i)(1-\gamma)\right] \sum_a \left(\hat{w}_i^a\right)^2 + \frac{1}{2}\beta^2 J^2 \left[\eta_i \gamma - (1-\eta_i)(1-\gamma)\right] \sum_a 2 y_i^a \hat{w}_i^a \right\} \tag{B.120}$$

For $\lambda_S = 0$, we calculate the integral over $\{w_i^a\}$, resulting in a delta distribution $\delta\left(\hat{w}_i^a\right)$. Thus, when we calculate the order parameters, the ones that involves averages of $\{\hat{w}_i^a\}$ will be zero. In fact, in this case, the only ones that will be non zero is:

$$m_a = \frac{1}{N} \sum_i \langle x_i^a \rangle_i^{\text{rep}} \qquad q_{ab} = \frac{1}{N} \sum_i \left\langle x_i^a x_i^b \right\rangle_i^{\text{rep}} \tag{B.121}$$
$$m_a^I = \frac{1}{N} \sum_i \eta_i \langle x_i^a \rangle_i^{\text{rep}} \qquad q_{ab}^I = \frac{1}{N} \sum_i \eta_i \left\langle x_i^a x_i^b \right\rangle_i^{\text{rep}} \tag{B.122}$$
$$m_a^J = \frac{1}{N} \sum_i (1-\eta_i) \langle x_i^a \rangle_i^{\text{rep}} \qquad q_{ab}^J = \frac{1}{N} \sum_i (1-\eta_i) \left\langle x_i^a x_i^b \right\rangle_i^{\text{rep}} \tag{B.123}$$
$$n_a = \frac{1}{N} \sum_i \langle y_i^a \rangle_i^{\text{rep}} \qquad r_{ab} = \frac{1}{N} \sum_i \left\langle y_i^a y_i^b \right\rangle_i^{\text{rep}} \tag{B.124}$$

Additionally, if we look at ζ_i for $\lambda_S = 0$,

$$\zeta_i = \sum_{\{x_i^a, y_i^a\}} \exp\left\{ \beta J_0 \sum_a n_a x_i^a + \beta^2 J^2 \sum_{a<b} r_{ab} x_i^a x_i^b + \beta J_0 \sum_a m_a y_i^a + \beta^2 J^2 \sum_{a<b} q_{ab} y_i^a y_i^b \right\} \tag{B.125}$$

we recover the partition function for the LSK model, as expected. In particular we notice that it is symmetric in the index i. Therefore the averages $\langle \cdot \rangle_i^{\text{rep}}$ present in the order parameters expressions will be equal for all i:

$$\langle x_i^a \rangle_i^{\text{rep}} = m_a \qquad \left\langle x_i^a x_i^b \right\rangle_i^{\text{rep}} = q_{ab} \tag{B.126}$$
$$\langle y_i^a \rangle_i^{\text{rep}} = n_a \qquad \left\langle y_i^a y_i^b \right\rangle_i^{\text{rep}} = r_{ab} \tag{B.127}$$

and we can write

$$m_a^I = \gamma m_a \qquad q_{ab}^I = \gamma q_{ab} \tag{B.128}$$
$$m_a^J = (1-\gamma) m_a \qquad q_{ab}^J = (1-\gamma) q_{ab} \tag{B.129}$$

Now, back to a general λ_S, we consider the replica symmetric ansatz. In this ansatz the order parameters no longer depends on the replica index, and we write ζ_i as:

$$\begin{aligned}\zeta_i^{\mathrm{RS}} =& \sum_{\{x_i^a, y_i^a\}} \int \prod_a \frac{dw_i^a d\hat{w}_i^a}{2\pi} \exp\Bigg\{ i \sum_a \hat{w}_i^a w_i^a + \lambda_S \sum_a \log 2\cosh\left(-iw_i^a\right) + \\ & + \beta J_0 \left[n + \eta_i u^I + (1-\eta_i) u^J\right] \sum_a x_i^a + \beta J_0 m \sum_a y_i^a + \beta J_0 \left[\eta_i m^I + (1-\eta_i) m^J\right] \sum_a \hat{w}_i^a + \\ & + \frac{1}{2}\beta^2 J^2 \left[r + \eta_i \left(s^I + 2t^I\right) + (1-\eta_i)\left(s^J + 2t^J\right)\right] \left(\sum_a x_i^a\right)^2 + \frac{1}{2}\beta^2 J^2 q \left(\sum_a y_i^a\right)^2 + \\ & + \frac{1}{2}\beta^2 J^2 \left[\eta_i q^I + (1-\eta_i) q^J\right] \left(\sum_a \hat{w}_i^a\right)^2 + \frac{1}{2}\beta^2 J^2 \left[\eta_i q^I + (1-\eta_i) q^J\right] 2 \left(\sum_a y_i^a\right)\left(\sum_a \hat{w}_i^a\right) + \\ & - \frac{1}{2}\beta^2 J^2 \left[r + \eta_i \left(s^I + 2t^I\right) + (1-\eta_i)\left(s^J + 2t^J\right)\right] \delta - \frac{1}{2}\beta^2 J^2 q \delta + \\ & + \frac{1}{2}\beta^2 J^2 \left[\eta_i \left(\gamma - q^I\right) + (1-\eta_i)\left((1-\gamma) - q^J\right)\right] \sum_a \left(\hat{w}_i^a\right)^2 + \\ & + \frac{1}{2}\beta^2 J^2 \left[\eta_i \left(\gamma - q^I\right) + (1-\eta_i)\left((1-\gamma) - q^J\right)\right] \sum_a 2 y_i^a \hat{w}_i^a \Bigg\} \end{aligned} \tag{B.130}$$

For the sake of clarity, we define the quantities

$$\sigma_{i|S}^2 = \eta_i \left(\gamma - q^I\right) + (1-\eta_i)\left(1 - \gamma - q^I\right) \tag{B.131}$$

and

$$\rho_i^S = \eta_i \rho^I + (1-\eta_i)\rho^J \tag{B.132}$$
$$\rho_i^D = (1-\eta_i)\rho^I + \eta_i \rho^J \tag{B.133}$$

for $\rho = m, q, u, s, t$.

In terms of these quantities, we have

$$\begin{aligned}\zeta_i^{\mathrm{RS}} =& \exp\left\{-\frac{1}{2}\beta^2 J^2 \left(r + 2t_i^S + s_i^S + q\right)\delta\right\} \sum_{\{x_i^a, y_i^a\}} \int \prod_a \frac{dw_i^a d\hat{w}_i^a}{2\pi} \exp\Bigg\{ i \sum_a \hat{w}_i^a w_i^a + \\ & + \lambda_S \sum_a \log 2\cosh\left(-iw_i^a\right) + \beta J_0 \left(n + u_i^S\right) \sum_a x_i^a + \beta J_0 m \sum_a y_i^a + \beta J_0 m_i^S \sum_a \hat{w}_i^a + \\ & + \frac{1}{2}\beta^2 J^2 \left(r + 2t_i^S + s_i^S\right)\left(\sum_a x_i^a\right)^2 + \frac{1}{2}\beta^2 J^2 q \left(\sum_a y_i^a\right)^2 + \frac{1}{2}\beta^2 J^2 q_i^S \left(\sum_a \hat{w}_i^a\right)^2 + \\ & + \frac{1}{2}\beta^2 J^2 q_i^S 2 \left(\sum_a y_i^a\right)\left(\sum_a \hat{w}_i^a\right) + \frac{1}{2}\beta^2 J^2 \sigma_{i|S}^2 \sum_a \left(\hat{w}_i^a\right)^2 + \frac{1}{2}\beta^2 J^2 \sigma_{i|S}^2 \sum_a 2 y_i^a \hat{w}_i^a \Bigg\}, \end{aligned} \tag{B.134}$$

and performing the average over $\{x_a\}$ and completing the squares,

$$
\begin{aligned}
\zeta_i^{\mathrm{RS}} = & \exp\left\{-\frac{1}{2}\beta^2 J^2 \left(r + 2t_i^S + s_i^S + q + \left(\sigma_i^S\right)^2\right)\delta\right\} \times \\
& \times \int Dz \left(2\cosh\left[\beta\left(J_0\left(n + u_i^S\right) + Jz\sqrt{r + 2t_i^S + s_i^S}\right)\right]\right)^\delta \times \\
& \times \sum_{\{y_i^a\}} \int \prod_a \frac{dw_i^a d\hat{w}_i^a}{2\pi} \exp\left\{ i\sum_a \hat{w}_i^a w_i^a + \lambda_S \sum_a \log 2\cosh\left(-iw_i^a\right) + \right. \\
& + \beta J_0 m_i^D \sum_a y_i^a + \frac{1}{2}\beta^2 J^2 q_i^D \left(\sum_a y_i^a\right)^2 + \beta J_0 m_i^S \sum_a \left(\hat{w}_i^a + y_i^a\right) + \\
& \left. + \frac{1}{2}\beta^2 J^2 q_i^S \left(\sum_a \left(\hat{w}_i^a + y_i^a\right)\right)^2 + \frac{1}{2}\beta^2 J^2 \left(\sigma_i^S\right)^2 \sum_a \left(\hat{w}_i^a + y_i^a\right)^2 \right\}
\end{aligned}
\tag{B.135}
$$

Introducing gaussian integrals for the quadratic terms we notice that the integrals factor in the replica index a

$$
\begin{aligned}
\zeta_i^{\mathrm{RS}} = & \exp\left\{-\frac{1}{2}\beta^2 J^2 \left(r + 2t_i^S + s_i^S + q + \left(\sigma_i^S\right)^2\right)\delta\right\} \times \\
& \times \int Dz \left(2\cosh\left[\beta\left(J_0\left(n + u_i^S\right) + Jz\sqrt{r + 2t_i^S + s_i^S}\right)\right]\right)^\delta \times \\
& \times \int Dz_S \int Dz_D \left(\sum_{y_i \in \{-1,1\}} \int \frac{dw_i d\hat{w}_i}{2\pi} \exp\left\{ i\hat{w}_i w_i + \lambda_S \log 2\cosh\left(-iw_i\right) + \right.\right. \\
& \left.\left. + \beta\left(J_0 m_i^D + Jz_D\sqrt{q_i^D}\right) y_i + \beta\left(J_0 m_i^S + Jz_S\sqrt{q_i^S}\right)\left(\hat{w}_i + y_i\right) + \frac{1}{2}\beta^2 J^2 \left(\sigma_i^S\right)^2 \left(\hat{w}_i + y_i\right)^2 \right\}\right)^\delta
\end{aligned}
\tag{B.136}
$$

Taking the logarithm and linearizing around $\delta = 0$

$$
\begin{aligned}
\log\zeta_i^{\mathrm{RS}} = & -\frac{1}{2}\beta^2 J^2 \left(r + 2t_i^S + s_i^S + q + \left(\sigma_i^S\right)^2\right)\delta + \\
& + \delta \int Dz \log 2\cosh\left[\beta\left(J_0\left(n + u_i^S\right) + Jz\sqrt{r + 2t_i^S + s_i^S}\right)\right] + \\
& + \delta \int Dz_S \int Dz_D \log \sum_{y_i \in \{-1,1\}} \int \frac{dw_i d\hat{w}_i}{2\pi} \exp\left\{ i\hat{w}_i w_i + \lambda_S \log 2\cosh\left(-iw_i\right) + \right. \\
& \left. + \beta\left(J_0 m_i^D + Jz_D\sqrt{q_i^D}\right) y_i + \beta\left(J_0 m_i^S + Jz_S\sqrt{q_i^S}\right)\left(\hat{w}_i + y_i\right) + \frac{1}{2}\beta^2 J^2 \left(\sigma_i^S\right)^2 \left(\hat{w}_i + y_i\right)^2 \right\}
\end{aligned}
\tag{B.137}
$$

Now we are finally able to substitute in the expression for $\langle \log Z_C \rangle_{\boldsymbol{J}}$ and take the limit $\delta \to 0$.

Taking the derivative with respect to λ_S and calculating at $\lambda_S = 0$ we get ϕ_C

$$\begin{aligned}\phi_C =& \sum_i \int Dz_S \int Dz_D \left(\sum_{y_i \in \{-1,1\}} \int \frac{dw_i d\hat{w}_i}{2\pi} \exp\left\{ i\hat{w}_i w_i + \beta\left(J_0 m_i^D + Jz_D\sqrt{q_i^D}\right) y_i + \right.\right.\\ & \left.\left. + \beta\left(J_0 m_i^S + Jz_S\sqrt{q_i^S}\right)(\hat{w}_i + y_i) + \frac{1}{2}\beta^2 J^2 \left(\sigma_i^S\right)^2 (\hat{w}_i + y_i)^2 \right\}\right)^{-1} \times \\ & \times \sum_{y_i \in \{-1,1\}} \int \frac{dw_i d\hat{w}_i}{2\pi} \left[\log 2\cosh(-iw_i)\right] \exp\left\{ i\hat{w}_i w_i + \beta\left(J_0 m_i^D + Jz_D\sqrt{q_i^D}\right) y_i + \right.\\ & \left. + \beta\left(J_0 m_i^S + Jz_S\sqrt{q_i^S}\right)(\hat{w}_i + y_i) + \frac{1}{2}\beta^2 J^2 \left(\sigma_i^S\right)^2 (\hat{w}_i + y_i)^2 \right\} \end{aligned} \tag{B.138}$$

After some manipulations we were able to get an expression for ϕ_C

$$\begin{aligned}\phi_C =& \sum_i \int Dz_S \int Dz_D \sum_{y_i \in \{-1,1\}} \frac{\exp\left\{\beta\left(J_0 m + Jz_D\sqrt{q_i^D} + Jz_S\sqrt{q_i^S}\right) y_i\right\}}{2\cosh\left[\beta\left(J_0 m_i^S + Jz_S\sqrt{q_i^S} + J_0 m_i^D + Jz_D\sqrt{q_i^D}\right)\right]} \times \\ & \times \int D\zeta \log 2\cosh\left[\beta\left(J_0 m_i^S + Jz_S\sqrt{q_i^S}\right) + \beta^2 J^2 \left(\sigma_i^S\right)^2 y_i + \beta J \sigma_i^S \zeta\right] \end{aligned} \tag{B.139}$$

This expression can be developed even further by separating the sum over i into two sums, one over $i \in I$ and another over $i \in J$, and by doing this we get a expression with the explicit dependence in γ

$$\begin{aligned}\phi_C =& \gamma N \int Dz_S \int Dz_D \sum_{y_i \in \{-1,1\}} \frac{\exp\left\{\beta\left(J_0 m + Jz_D\sqrt{(1-\gamma)q} + Jz_S\sqrt{\gamma q}\right) y_i\right\}}{2\cosh\left[\beta\left(J_0 m + Jz_S\sqrt{\gamma q} + Jz_D\sqrt{(1-\gamma)q}\right)\right]} \times \\ & \times \int D\zeta \log 2\cosh\left[\beta\left(J_0 \gamma m + Jz_S\sqrt{\gamma q}\right) + \beta^2 J^2 \gamma (1-q) y_i + \beta J\sqrt{\gamma(1-q)}\zeta\right] + \\ & + (1-\gamma) N \int Dz_S \int Dz_D \sum_{y_i \in \{-1,1\}} \frac{\exp\left\{\beta\left(J_0 m + Jz_D\sqrt{\gamma q} + Jz_S\sqrt{(1-\gamma)q}\right) y_i\right\}}{2\cosh\left[\beta\left(J_0 m + Jz_S\sqrt{(1-\gamma)q} + Jz_D\sqrt{\gamma q}\right)\right]} \times \\ & \times \int D\zeta \log 2\cosh\left[\beta\left(J_0(1-\gamma)m + Jz_S\sqrt{(1-\gamma)q}\right) + \beta^2 J^2 (1-\gamma)(1-q) y_i + \beta J\sqrt{(1-\gamma)(1-q)}\zeta\right] \end{aligned} \tag{B.140}$$