UNIVERSITÀ DEGLI STUDI DI MILANO

Facoltà di Scienze Matematiche, Fisiche e Naturali

Corso di laurea in Fisica

Cavity Method Analysis for Random Assignment Problems

Relatore: Prof. Sergio Caracciolo

Correlatore: Prof. Bruno Bassetti

PACS: 05.70.Fh, 75.10.Nr;
02.50.Tt, 89.20.Ff.

Tesi di laurea di
Caterina Grosso
matr. 604030

Anno Accademico 2002-2003

02.50.Tt Inference methods
05.70.Fh Phase transitions: general studies
75.10.Nr Spin-glass and other random models
89.20.Ff Computer science and technology

Contents

Introduction

In the 80's, the field of disordered systems received a boost by the solution of a prototype problem, the Sherrington-Kirkpatrick Model, achieved by Parisi and collaborators. New theoretical methods led to the understanding of new, interesting phenomena, typical of disordered frustrated systems in mean-field approximation. These features, like the existence of an extensive number of pure phases, not related by any symmetry, and their organization in an ultrametric structure, cannot be treated with standard tools of statistical mechanics.

The theory of disordered system receives interest also for its applications in Computer Science. A combinatiorial optimization problem can be restated in terms of a statistical mechanics problem with a Hamiltonian mimicking the cost function. Recently this research field is receiving considerable attention, since the development a new class of algorithms, named Survey Propagation algorithms. They are based on equations derived from an analysis of Onsager cavity fields, exploiting the physical knowledge of the phase space complexity structure, and they are capable of solving optimization problems even for instances of hardness and sizes where traditional algorithms fail.

Many questions are open in this field, both from a numerical and a theoretical point of view. It would be important to understand what kind of problems this way of proceding can be extended to, and to improve the performances of existing algorithms, with a better control of the approximations they rely on.

In the first chapter, we review some basic concepts of statistical mechanics, and we introduce some tools useful in the study of disordered materials. In particular, we shortly present the Parisi Solution to the Sherrington-Kirkpatrick Model, discussing the replica method, the phenomenon of Replica Symmetry Breaking, and the ultrametric structure of the phase space.

In the second chapter, another method is presented, the Cavity approach. The main idea dates back to the same years of the Replica method, and its main difference with it is that the cavity method can be applied also at given instance; only recently it has been applied to systems with Replica Symmetry Breaking, and it has been recognized to be particularly suitable for finite connectivity systems. Besides, Survey Propagation Algorithms are based on it. The derivation presented here is original. We describe a general class of models it applies to, and derive cavity

equations, emphasizing all approximations involved, both in the Replica Symmetric case, and in the case where many pure phases are present. We finally discuss how to extend these equations when averaging over disorder, a delicate topic for which a general discussion is lacking.

The third chapter is a short introduction to optimization theory, and describes qualitatively some of the most fruitful exchanges between statistical physics and computer science. The connection between optimization problems and physics is briefly exposed.

In this context, the study of a simple model, the *Random Assignment Problem*, is significant. This problem has considerable practical drawbacks in the world of information theory: it is possible to connect it to the traffic scheduling in routers for network applications. Besides, it shows many interesting features also from a theoretical point of view. The fourth chapter focuses on this aspect. We introduce the model and apply the general theory of the Cavity Method to derive cavity equations at finite temperature for this problem. The statistical investigation of these equations, when averaging over disorder, is possible in the context of distributional equations. Using general theorems on Independent Point Processes, Cavity Equations are translated into differential equations for the cavity-fields distributions, from which one can find several statistical properties of the model. The mathemathical tool of Independent Point Processes is shortly treated in the Appendix, where also an original theorem is presented. This tool has allowed us to solve the zero-temperature limit of the equations, reobtaining some known results; it chiefly has allowed us to extend the investigation to new quantities, like the finite-temperature distributions of the fields.

In the fifth chapter we again exploit the same technique to study the thermodynamics of an extension of the Random Assignment Problem, named the k-assignment problem, proposed in literature, but for which a statistical analysis was lacking. After deriving Cavity Equations, we promote them to distributional equations, and find a local differential equation for the cavity-fields distribution. Also an equation for the average ground state energy is obtained. Analytic expressions are derived in the large k limit, with finite-k corrections.

The sixth chapter comes back to the world of computer science: we present some methods for an algorithmic solution of the Assignment Problem and its variations. In particular, we describe in general how linear programming techniques work, presenting the Simplex Method. The Linear Programming formulation for Assignment problems is given. We then describe another exact algorithm to solve the Assignment Problem, named the Hungarian Algorithm, showing how it works. An original algorithmic implementation of Cavity Equations for the Assignment Problem is given, highlighting some difficulties. The hardness of the satisfiability version of the Assignment Problem, as a function of the threshold, is discussed both in the SAT and in the UNSAT regions. Finally, we briefly mention the scheduling problem in routers, one of the main application for the Assignment Problem.

This thesis has also raised many open questions, both from a theoretical and an algorithmical point of view. In the last chapter, a short presentation of future development is given, ranging from a possible treatment of other optimization problems with cavity-method based distributional equations, to a future sistematic investigation of algorithmic implementations for the Cavity Equations, applied to the Assignment problem.

1. Introduction to the Statistical Mechanics of random systems

This chapter presents some basic concepts, used extensively in this work. The first three sections deal a prototype model for ordered systems, the Ising Model: the concepts of phase transitions, pure phases, Cluster property are stressed. We then turn our attention to disordered systems: in the fourth section, we describe the relevant questions for these models, while in the fifth and sixth, a prototype model for disordered systems, the Sherrington-Kirkpatrick Model, is discussed, presenting the so-called Parisi Solution. The phenomenon of Replica Symmetry Breaking and the ultrametricity of the phase space are described, as well.

1.1 Spontaneous Symmetry Breaking in infinite systems

Equilibrium Statistical Mechanics deals with thermodynamical systems with a large number of degrees of freedom [1]. The thermodynamics is encoded in a small set of physical quantities (such as temperature, pressure, free energy density, ...), arising from average over the macroscopic number of variables (positions, momenta, spins, ...) which describe the microscopic components.

A mathematical definition of a system at equilibrium is the datum of a set of external parameters (temperature, pressure, etc.), a finite-dimensional space of configurations $X = X_0^N$, equipped with a reference measure $\mathrm{d}x$, and a Hamiltonian $\mathcal{H}(x) : X \to \mathbb{R}$. For each temperature $T = 1/\beta$, the Gibbs probability measure

$$\mathrm{d}\mu_N(x) = \frac{1}{Z_N} e^{-\beta \mathcal{H}_N(x)} \mathrm{d}x \tag{1.1}$$

is stable under any local dynamics preserving detailed balance at given β, which is a standard request for a first modeling of the microscopic dynamics for systems at equilibrium.[1]

The normalization factor Z_N is named *partition function*, and its logarithm is extensive in system size. This leads to the definition of the free energy F_N and the free energy density

[1] In formulas, the transition rates $W_{x \to x'}$ of the dynamics should satisfy $W_{x \to x'} \to 0$ for $|x - x'| \to \infty$ (locality) and $W_{x \to x'} / W_{x' \to x} = \exp[-\beta(\mathcal{H}(x') - \mathcal{H}(x))]$ (detailed balance).

$f = F_N/N$

$$Z_N = \int_X \mathrm{d}x\, e^{-\beta \mathcal{H}_N(x)}\,; \qquad\qquad F_N = -\frac{1}{\beta} \ln Z_N\,. \tag{1.2}$$

A physical state for a given system corresponds to a linear functional over the space of observables $\{A\}$. The linearity property is implicit in functionals Ω of the form

$$\Omega(A) = \langle A \rangle_\Omega = \frac{\int_X \mathrm{d}\mu_\Omega(x)\, A(x)}{\int_X \mathrm{d}\mu_\Omega(x)}\,. \tag{1.3}$$

Note that, if ergodicity fails, the Gibbs measure is not guaranteed to be the only equilibrium measure on the system. Indeed, the infinite-volume limit (or *thermodynamic limit*) can lead to new phenomena forbidden in finite systems, namely the ergodicity under local dynamics can be broken, and the free-energy density of the system can be a non-analytic function of the physical parameters. The points in the space of parameters in which f is singular are called *critical points*. At these points, a *phase transition* can occur in the system.

The study of Critical Phenomena receives a boost from *Universality Hypothesis*[2]: the kind of singularity at the critical points is determined only by general properties of the configuration space and of the Hamiltonian (dimensionality of the underlying space, range of interaction, symmetry property of the variables involved, ...). This fact justifies an abstract mathematical approach to critical phenomena: the study of idealized models reveals critical properties of complicated physical systems sharing the same universality characteristics of the model, i.e. which belong to the same *universality class*.

A prototype phase transition, the spontaneous magnetization of ferromagnetic materials below the Curie temperature, has provided a general lexicon in the field, for models describing the most diverse phenomenologies. The *Ising Model*, introduced to describe this kind of transition, has been a milestone in the theory of Critical Phenomena, since the acclaimed solution of the two-dimensional version of the problem by Onsager in 1944 [3].

We consider a regular lattice in d dimensions (to fix the ideas, think of a hypercubic lattice of side L), with ions on the $N = L^d$ vertices. Since we are interested in the magnetic behaviour, we only consider the spin quantum number of each ion, neglecting vibrational modes around the equilibrium positions. Anisotropy suggests us to deal with classical spins $\sigma_i \in \{\pm 1\}$. The configuration space is thus $\{\pm 1\}^N$, where N is the number of sites in the lattice. The Hamiltonian is

$$\mathcal{H}(\boldsymbol{\sigma}) = -J \sum_{\langle i,j \rangle} \sigma_i \sigma_j - h \sum_i \sigma_i\,. \tag{1.4}$$

where $\langle i, j \rangle$ denotes pairs of first-neighbouring sites, $J > 0$ is the parameter defining the intensity of the interaction, and h is an external magnetic field. Consider the magnetization of the system

[2]Nowadays, this hypothesis finds a justification in the context of Renormalization Group [2].

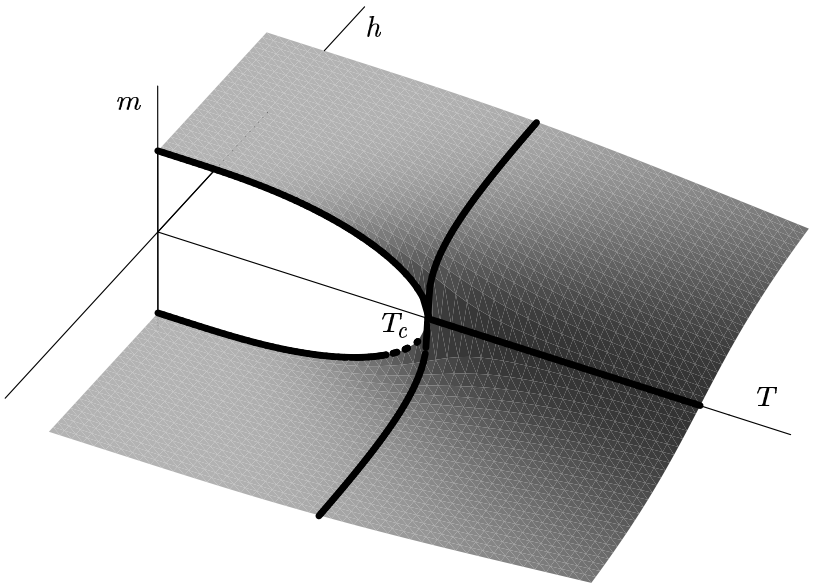


Figure 1.1 Qualitative shape of the average magnetization $m(T,h)$ for the Ising Model in $d > 1$.

as a function of temperature and magnetic field

$$m(T,h) = \langle m(\boldsymbol{\sigma}) \rangle \ , \qquad\qquad m_N(\boldsymbol{\sigma}) = \frac{1}{N} \sum_i \sigma_i \, . \tag{1.5}$$

In the thermodynamic limit, for $d > 1$ a critical point exists in the plane (T,h), for $h = 0$ and $T = T_c > 0$. The function $m(T,h)$ is analytic on the whole half-plane $T > 0$, with the exclusion of the line with zero external field and $T \in [0, T_c]$, and the original symmetry of the problem reflects into $m(T,h) = -m(T,-h)$. On the critical line there is a cut discontinuity: the limits $m^+(T) = \lim_{h \to 0^+} m(T,h)$ and $m^-(T) = \lim_{h \to 0^-} m(T,h)$ do not coincide. This fact has strong implications on the equilibrium measures of the system. Consider the two limit measures

$$\mu_+^{(T)}(\boldsymbol{\sigma}) = \lim_{h \to 0^+} \mu_{\text{Gibbs}}^{(T,h)}(\boldsymbol{\sigma}) \, ; \qquad\qquad \mu_-^{(T)}(\boldsymbol{\sigma}) = \lim_{h \to 0^-} \mu_{\text{Gibbs}}^{(T,h)}(\boldsymbol{\sigma}) \, . \tag{1.6}$$

A finite difference on the expectation value of an intensive quantity (as the average magnetization m) implies that the measures $\mu_+^{(T)}$ and $\mu_-^{(T)}$ do not coincide. This phenomenon is called *Spontaneous Symmetry Breaking*.

Indeed, on the critical line, both measures are equilibrium measures for the system: two *pure phases* coexist. The set of equilibrium measures is a convex: any equilibrium measure can be written in one and only one way as a convex combination of extremal measures. The extremal measures on this convex correspond to pure phases. In this case, both $\mu_+^{(T)}$ and $\mu_-^{(T)}$ correspond to restrictions of the Gibbs measure to a region of the phase space (space of configurations).

In our case we can state

$$\mu_+^{(T)}(\boldsymbol{\sigma}) \propto e^{-\frac{1}{T}\mathcal{H}_{h=0}(\boldsymbol{\sigma})}\theta(m(\boldsymbol{\sigma}))\,; \qquad \mu_-^{(T)}(\boldsymbol{\sigma}) \propto e^{-\frac{1}{T}\mathcal{H}_{h=0}(\boldsymbol{\sigma})}\theta(-m(\boldsymbol{\sigma}))\,. \tag{1.7}$$

Note that this definition makes sense also at finite size, although stability under time evolution is obtained only in the thermodynamic limit, and derives from two facts

- local stability far from the θ discontinuity is a consequence of proportionality with Gibbs measure, and locality of the dynamics;
- the lack of stability near the edge of the θ is thermodynamically irrelevant, as the measure of the region of phase space with $m = \mathcal{O}(1/N)$ goes to zero exponentially with the size: indeed, far from the critical point, thermal fluctuations of the magnetization are of order $\mathcal{O}(N^{-1/2})$, and $|m(T)|$ is of order 1.

We could say that the finite-size measures (1.7) correspond to *"quasi"-pure phases*: pictorially speaking, they correspond to "valleys" of the phase space, whose border has a measure which goes to zero exponentially with the size, and thus become a "separated world" in the thermodynamic limit.

This picture can be generalized to more complex situations, and in particular to disordered systems. On the other hand, some features are peculiar of Ising-like systems:

- The Hamiltonian at zero magnetic field is symmetric for flip of all spins: this implies that time evolution commutes with this operation. Thus, either we have only one equilibrium measure which preserves this symmetry, as in the high-temperature phase $T > T_c$, or the set of equilibrium measures is stable under symmetry operation, $\mu_+^{(T)}(\boldsymbol{\sigma}) = \mu_-^{(T)}(-\boldsymbol{\sigma})$, as in the low-temperature phase $T < T_c$. On the contrary, a typical Hamiltonian of a disordered system does not have any symmetry, and it is not possible to relate a pure phase to the others via the action of a symmetry group.
- The free energy of the two phases, related to normalizations of the two (1.7), and defined in analogy with (1.2), strictly coincide as a consequence of the original symmetry. In disordered systems, many pure phases can exist, with free energies differences of order 1.
- The two pure phases of the system are identified by the expectation value of magnetization. In this case we say that m is an *order parameter* for the transition. As shown above, we can select one pure phase via a limit process (Bogoliubov criterium [4]). In disordered systems, it is a difficult task to identify an order parameter, and to select a certain pure phase.

1.2 Characterization of pure phases and Cluster Property

A different approach to the definition of pure phases in a thermodynamical system exists. This approach starts from the physical assumption that the ground state of a system is perturbed only locally under local perturbations (say, measurements of a physical quantity), so measurements far away in space do not allow non-trivial correlations.

This property must be true only for a given ground state: when the system allows for different ground states (and thus several "valleys" for a perturbation-theory approach), if we perform averages with a measure which is not extremal, correlations can originate from different expectation values of physical observables in different pure phases. Indeed, if we have two distinct pure phases, at least one physical observable must exist, such that its expectation value in the two phases is different. Thus, a formalization of the physical picture above (which goes under the name of *Cluster Property*), under general physical hypothesis, can be reconducted to an equivalent definition of pure phase.

Theorem 1 ((naïve) Cluster Property) *A given equilibrium measure μ on a physical system describes a pure phase if and only if for each pair of local physical observables the connected correlation function vanishes in the large distance limit:*

$$\forall \quad A, B \qquad \lim_{|x-x_0|\to\infty} \Big(\langle A(x_0)B(x)\rangle - \langle A(x_0)\rangle \langle B(x)\rangle \Big) = 0\,. \tag{1.8}$$

In the previous section, we defined some quantities using large finite systems, in order to guarantee a more precise mathematical control. In particular, we gave a hint on the concept of finite-system "quasi"-pure phases (in the Ising Model at $h = 0$ and $T < T_c$, the two measures (1.7), when considered at finite size): although only in the thermodynamic limit they have all the properties of pure phases (for instance, stability under a local dynamic), quantitative corrections to these properties are typically bounded by functions of the size (for instance, metastability lifetimes are expected to scale as $\tau \sim \exp(aN)$).

In particular, we expect a version of the Cluster Property to hold in finite-size:

Theorem 2 ((naïve) finite-size Cluster Property) *Consider a family of equilibrium measures μ_N on finite-size physical systems, which converge to a measure μ in the large N limit. The limit measure μ corresponds to a pure phase if and only if for each pair of local physical observables the connected correlation function, calculated w.r.t. the finite-size measures, vanishes in the large size, large distance limit, provided that some typical diameter $d(N)$ of the system is kept large w.r.t. the distance $|x - x_0|$ between the supports of the operators:*

$$\forall \quad A, B \qquad \lim_{\substack{|x-x_0|\to\infty \\ N\to\infty \\ |x-x_0|\ll d(N)}} \Big(\langle A(x_0)B(x)\rangle_N - \langle A(x_0)\rangle_N \langle B(x)\rangle_N \Big) = 0\,. \tag{1.9}$$

An intuition on the validity of Cluster Property is given by the example of the Ising Model. Consider the model at $h = 0$ and $T < T_c$, on a finite periodic lattice of size $N = L^D$. In this case $d(N) \sim L/2$. As all sites are equivalent under translations, we have

$$\frac{1}{N}\sum_x \langle \sigma_x \rangle_N \equiv \overline{m}_N = \langle \sigma_{x_0} \rangle_N \qquad \forall \quad x_0 \in \{1, \dots N\}\,. \tag{1.10}$$

From the discussion presented in the previous section, it is clear that any stable measure can be reconducted to convex combinations of the two (1.7), in the termodinamic limit.[3] Thus, we have a one-parameter family of measures

$$\mu_t(\boldsymbol{\sigma}) = t\mu_-(\boldsymbol{\sigma}) + (1-t)\mu_+(\boldsymbol{\sigma})\,, \qquad t \in [0,1]\,; \tag{1.11}$$

and, if $m = m^+(T)$ is the average magnetization in the phase μ_+, the average magnetization in the phase μ_t is given by $m(1-2t)$.

What happens is that, when $T < T_c$, the correlation length of the system is finite, thus long-range thermal fluctuations are suppressed. In the correlation function $\langle \sigma_{x_0}\sigma_x \rangle$, at large distance, either both spins are on average m, with almost decorrelated fluctuations (this happens with probability $(1-t)$), or both spins are on average $-m$, with almost decorrelated fluctuations (this happens with probability t). So the expectation value of the correlation function in this measure is

$$\lim_{|x-x_0|\to\infty} \langle \sigma_{x_0}\sigma_x \rangle_t = m^2\,, \tag{1.12}$$

for all choices of t. On the other side, both the one-point function $\langle \sigma_{x_0} \rangle$ and $\langle \sigma_x \rangle$ are on average $m(1-2t)$, so we have

$$\lim_{|x-x_0|\to\infty} \Big(\langle \sigma_{x_0} \rangle_t \langle \sigma_x \rangle_t \Big) = m^2(1-2t)^2 = m^2(1-4t(1-t))\,, \tag{1.13}$$

which coincides with equation (1.12) only in the two pure phase measures $t = \{0, 1\}$.

To summarize, this section presents some simple non-rigorous hint on a mathematically delicate task, that is, characterizing pure phases of a thermodynamical system via the cluster property of correlation functions. The proof of the equivalence of this characterization with the more intuitive "valley" picture, and with stability under local dynamic, is out of our purposes, and is done in literature only for traditional ordered systems, with particular emphasis over S-matrix theory in Quantum Field Theory [5], [6], [4].

The lack of an equivalent of Cluster Property for disordered systems, with a sound mathematical formulation, is an important subject of research. Indeed, as we will see in section 2.3, Cluster

[3] At finite size, "quasi"-stable measures can be reconducted to convex combinations of the two (1.7), up to discrepancies in a part of the phase space, whose measure in any equilibrium distribution vanishes in the thermodynamic limit.

Property is a fundamental concept in the derivation of Cavity Equations for weak-inference disordered systems. Attention of the mathematical-physics community over this delicate point is welcome.

1.3 Statistical Mechanics of Disordered Systems

Random models arose from the study of disordered materials. We have seen in section 1.1 that the Ising Model is a prototype model for the realistic interaction of ferromagnetic ions on a regular lattice. We now turn our attention to new features arising from the physics of systems in which variables are located on a disordered lattice, like the one of a glassy material.

The Hamiltonian of this kind of system is described by an extensive number of external parameters (e.g. the exact position of the atoms in a glass sample). A physicist is *not* interested in the exact thermodynamics of a given sample, but rather to average behaviour over all possible samples.

For this reason, a powerful modelization of these systems consists of an ensemble for the parameters J describing the specific instance, equipped with an *a priori* measure $\mathrm{d}\mu(J)$, and an ensemble for the configurations $\boldsymbol{\sigma}$. The partition function will depend on J, and the measure over configurations will be given by

$$\mu^{(J)}_{\text{Gibbs}}(\boldsymbol{\sigma}) = \frac{1}{Z_J} e^{-\beta \mathcal{H}_J(\boldsymbol{\sigma})}\,, \qquad\qquad Z_J = \int_X \mathrm{d}\boldsymbol{\sigma}\, e^{-\beta \mathcal{H}_J(\boldsymbol{\sigma})}\,. \tag{1.14}$$

As measurements are performed on a given sample, the calculation of a certain expectation value should be the result of a *quenched average*

$$\overline{A} = \int \mathrm{d}\mu(J)\ \langle A \rangle_J\ , \qquad\qquad \langle A \rangle_J = \frac{1}{Z_J} \int_X \mathrm{d}\boldsymbol{\sigma}\, A(\boldsymbol{\sigma}) e^{-\beta \mathcal{H}_J(\boldsymbol{\sigma})}\,. \tag{1.15}$$

Remark that if the disordered degrees of freedom J were at thermal equilibrium as in standard Statistical Mechanics, the result of the measurement would have been given instead by the so-called *annealed average*

$$\langle A \rangle_{\text{ann.}} = \frac{1}{Z} \int \mathrm{d}\mu(J) \int_X \mathrm{d}\boldsymbol{\sigma}\, A(\boldsymbol{\sigma}) e^{-\beta \mathcal{H}_J(\boldsymbol{\sigma})}\,. \tag{1.16}$$

It is interesting to know whether the result of a measurement depends on the sample or not. Observables whose average does not fluctuate with the sample in the thermodynamic limit are called *self averaging*, i.e. if A is such an observable

$$\lim_{N\to\infty} \frac{\overline{(\overline{A} - \langle A \rangle_J)^2}}{\overline{A}^2} = 0\,, \tag{1.17}$$

where the bar denotes an average over disorder. In particular, for most physical systems the free energy density is a self-averaging quantity.

Quenched averages are difficult. Many standard tools of statistical mechanics cannot be used any more. The *replica trick* allows to compute the quenched average of any observable A, as defined in equation (1.15), as an analytic continuation over a set of more controlled quantities:

$$\langle A\rangle^{(n)} = \frac{\int \mathrm{d}\mu(J)\, Z_J^{n-1} \int \mathrm{d}\boldsymbol{\sigma}\, A(\boldsymbol{\sigma}) e^{-\beta\mathcal{H}_J(\boldsymbol{\sigma})}}{\int \mathrm{d}\mu(J)\, Z_J^n}; \qquad \overline{A} = \lim_{n\to 0} \langle A\rangle^{(n)} \,. \tag{1.18}$$

At each integer positive n, the quantity $\langle A\rangle^{(n)}$ corresponds to a traditional annealed average on a system in which n copies of the original one (*replicas*) are coupled via the same realization of disorder J. Indeed, define a n-replica Hamiltonian as

$$\mathcal{H}_J(\boldsymbol{\sigma}_1, \cdots, \boldsymbol{\sigma}_n) = \sum_{a=1}^{n} \mathcal{H}_J(\boldsymbol{\sigma}_a)\,, \tag{1.19}$$

we recognize that

$$\langle A\rangle^{(n)} = \frac{\int \mathrm{d}\mu(J) \int \mathrm{d}\boldsymbol{\sigma}_1 \cdots \mathrm{d}\boldsymbol{\sigma}_n\, A(\boldsymbol{\sigma}_1) e^{-\beta\mathcal{H}_J(\boldsymbol{\sigma}_1,\cdots,\boldsymbol{\sigma}_n)}}{\int \mathrm{d}\mu(J) \int \mathrm{d}\boldsymbol{\sigma}_1 \cdots \mathrm{d}\boldsymbol{\sigma}_n\, e^{-\beta\mathcal{H}_J(\boldsymbol{\sigma}_1,\cdots,\boldsymbol{\sigma}_n)}}\,. \tag{1.20}$$

Note that the partition function of the n-replica system corresponds to the average n-th moment of the partition function in the original system, $\overline{Z^n}$.

In particular, this technique is useful for finding the quenched average of the free energy density; since

$$\ln Z = \lim_{n\to 0} \frac{Z^n - 1}{n}\,, \tag{1.21}$$

the free energy density can be calculated as an analytic continuation from the formula:

$$\overline{f} = -\frac{1}{\beta N}\overline{\ln Z} = \lim_{n\to 0} -\frac{1}{\beta n N}(\overline{Z^n} - 1)\,. \tag{1.22}$$

So we must find the expectation value over the disorder of the n-th moment of the partition function. This is an annealed average, thus for each finite n it is easier to calculate, in general, than a quenched average. Then we have to perform an analitic continuation on the parameter n, which has a physical interpretation only when it is a positive integer. Indeed, the analitic continuation will turn out to be the difficult point of the whole procedure. This task has been fulfilled for the first time by Parisi and collaborators in the 80's for a specific model, the *Sherrington-Kirkpatrick Model*, which since then has become a prototype for mean-field disordered systems.

1.4 The Sherrington-Kirkpatrick model

In this section we will consider the Sherrington Kirkpatrick model at zero external magnetic field. For a more extensive presentation, we refer the reader to [7, 8]. The Hamiltonian is:

$$\mathcal{H}_J(\boldsymbol{\sigma}) = -\sum_{1\le i<j\le N} J_{ij}\,\sigma_i\sigma_j\,. \tag{1.23}$$

It depends on two sets of variables: a set of N spins $\boldsymbol{\sigma}$, taking value in $\{\pm 1\}^N$, and a set of couplings J. For each unordered pair (i,j), the value of $J_{ij} = J_{ji}$ is a random number, drawn from a probability distribution $\mu_0(J)$; we will assume that the distribution is a Gaussian with zero mean

$$\mu_0(J) = \sqrt{\frac{N}{2\pi}} e^{-\frac{N}{2}J^2}\,. \tag{1.24}$$

Notice that the interactions are infinite range: each spin interacts with all others and there is no notion of 'neighbouring' or geometric distance between pairs of spins.

We want to use formula (1.22) to find the free energy density. We start with the calculation of average moments of the partition function:

$$\overline{Z_N^n} = \sum_{\boldsymbol{\sigma}} \int \prod_{i<j} \left(\mathrm{d}J_{ij} \sqrt{\frac{N}{2\pi}} e^{-\frac{N}{2}J_{ij}^2 + \sum_{a=1}^n \beta J_{ij}\sigma_i^a\sigma_j^a} \right)\,; \tag{1.25}$$

note that $\boldsymbol{\sigma}$ is now an element of the configuration space $\{\pm 1\}^{nN}$. Integrating over J we get

$$\overline{Z_N^n} = \sum_{\boldsymbol{\sigma}} e^{\frac{\beta^2}{2N}\sum_{i<j}\left(\sum_{a=1}^n \sigma_i^a\sigma_j^a\right)^2} = e^{\frac{(N-n)n\beta^2}{4}} \sum_{\boldsymbol{\sigma}} e^{\frac{\beta^2}{2N}\sum_{a<b}\left(\sum_{i=1}^N \sigma_i^a\sigma_i^b\right)^2}\,. \tag{1.26}$$

Note that $\overline{Z_N^n}$ is of the same form if we exchange N with n, up to a constant factor. It is thus natural to repeat backwards all the passages we have done so far, where now n is the system size, and N is the number of replicas; after Gaussian integration, all the constants conspire to highlight the N dependence of this quantity:

$$\overline{Z_N^n} = e^{\frac{(N-n)n\beta^2}{4}} \int \prod_{a<b} \frac{\mathrm{d}Q_{ab}}{\sqrt{2\pi/(N\beta^2)}} e^{-NS[Q]}\,, \tag{1.27}$$

where Q_{ab} is a $n\times n$ symmetric matrix, zero on diagonal, which is the counterpart of J_{ij} in equation (1.25), and describes the couplings of the n-site system. The action S is a sum of two parts

$$S[Q] = -\ln M[Q] - \ln Z[Q]\,; \tag{1.28}$$

$M[Q]$ is the Gaussian measure on Q_{ab}

$$M[Q] = e^{-\frac{\beta^2}{2}\sum_{a<b} Q_{ab}^2}\,, \tag{1.29}$$

and Z[Q] is the partition function of the n-site system,

$$Z[Q] = \sum_{\boldsymbol{\sigma}\in\{\pm 1\}^n} e^{-\beta^2 \sum_{a<b} Q_{ab}\sigma_a\sigma_b} \,. \tag{1.30}$$

It is now possible to apply the saddle point method to evaluate the integral. The only relevant part of the action is $S[Q^*]$, where Q^* solves the saddle point equations

$$\left.\frac{\partial S[Q]}{\partial Q_{ab}}\right|_{Q^*} = 0\,, \qquad \text{for all } a,\, b\,, \tag{1.31}$$

and corresponds to a minimum of the action.

Note that $S[Q]$ is invariant under the action of the group of permutations of n elements, $\mathfrak{S}_n$, so either Q^* is symmetric, or there exist several solutions of these equations, which define an orbit of a non-trivial subgroup of $\mathfrak{S}_n$.

Once we have found the solution to these equations for n positive integer, the free energy, up to trivial contributions coming from the constant multiplying factor, is given by the analytic continuation

$$\overline{f} = \lim_{n\to 0} \frac{1}{\beta n} S[Q^*]\,. \tag{1.32}$$

Equations (1.31) read

$$Q^*_{ab} = \langle \sigma_a \sigma_b \rangle \,, \tag{1.33}$$

where the average is evaluated on the n-site replica system; also

$$Q^*_{ab} = \overline{\langle \sigma_i^a \sigma_i^b \rangle} = \frac{1}{N}\sum_i \overline{\langle \sigma_i^a \sigma_i^b \rangle}\,, \tag{1.34}$$

in the original n-replica, N-site system.[4] This equation allows us to give a physical meaning to the saddle-point matrix Q^*, even for its analytic continuation in n.

Given two configurations, we define their overlap as

$$q(\boldsymbol{\sigma}^a, \boldsymbol{\sigma}^b) = \frac{1}{N}\sum_i \sigma_i^a \sigma_i^b\,, \tag{1.35}$$

and their distance as

$$d(\sigma_i^a, \sigma_i^b) = 1 - q(\sigma_i^a, \sigma_i^b) = \frac{1}{2N}\sum_i (\sigma_i^a - \sigma_i^b)^2\,. \tag{1.36}$$

In replicated systems, these quantities give a measure of the correlation and the distance between configurations in different replicas. The distribution at fixed instance

$$p_J(q) = \Big\langle \delta(q(\sigma^a, \sigma^b) - q) \Big\rangle_J \tag{1.37}$$

[4] This can be proved for example introducing a source term $\Delta\mathcal{H} = \sigma_i^a h_i^a + \sigma_i^b h_i^b$ in formula (1.25).

does not depend on n since the contribution of the other $n-2$ replicas is factorized. It can be shown that the overlap distribution is *not* self-averaging. Consider its quenched average

$$\overline{p(q)} = \int \mathrm{d}\mu(J)\, p_J(q)\,, \tag{1.38}$$

and the distribution of Q_{ab} entries on the saddle point

$$p(Q) = \lim_{n\to 0} \frac{2!}{n(n-1)} \sum_{a<b} \delta(Q^*_{ab} - Q)\,; \tag{1.39}$$

we find that equation (1.34) implies

$$p(Q) = \overline{p(q)}\,. \tag{1.40}$$

1.5 Replica Symmetry Breaking and the Parisi order parameter

We seek for a solution of the saddle point equations (1.31). Proposing an Ansatz for the matrix Q, it is natural to use a form which is symmetric under permutation of rows and columns, since the action has this property. We thus try

Ansatz 1.5.1 (Replica Symmetric Ansatz) *For all $a < b$, $Q_{ab} = q$.*

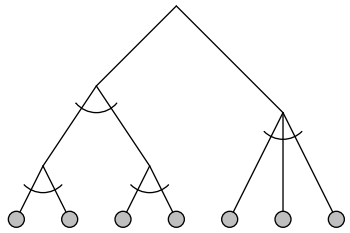
Plugging this form in equation (1.31) and performing an analytic continuation of the result for $n \to 0$, one finds that the only solution at $T > T_c = 1$ is $q = 0$, while below the critical temperature there is another non-null solution of the saddle point equation.

Starting from the result we have found this way, it is possible to calculate the free energy density and the entropy. Unfortunately, it turns out that at low temperature not only the free energy density does not agree with numerical simulations, but also, the entropy is negative. This is a clear sign that the RS Ansatz is not stable at all temperatures, and the $n \to 0$ limit hides some difficulties. In particular, it is possible to show that the proposed solution, who was a minimum for the action at $n > 1$, becomes a maximum for $0 < n < 1$. We are thus forced to propose an Ansatz breaking the symmetry of the action.

We can classify all possible matrices Q_{ab} with respect to their symmetry group G_Q of permutations. For each subgroup $G_Q \subseteq \mathfrak{S}_N$, a set of generators exists, containing only simple transpositions of index pairs, or of pairs of subsets of indices with equal cardinality. For example, the matrix

$$Q = \left(\begin{array}{ccc|cc|ccc}
0 & a & & b & b & d & d & d \\
 & 0 & & b & b & d & d & d \\
\hline
 & & 0 & a & & d & d & d \\
 & & & 0 & & d & d & d \\
\hline
 & & & & 0 & c & c & \\
 & * & & & & 0 & c & \\
 & & & & & & 0 &
\end{array}\right)$$

has corresponding group of invariance generated by the permutations (12), (34), (12)(34), (56) and (57). Each group G_Q can be graphically represented as a rooted tree with n leaves: at each node, there is an arc collecting c branches iff the permutations which exchange rigidly the sets of leaves in two of these branches are in G_Q. For example, for the matrix Q we would have



A special class of matrices is the one corresponding to fully symmetric trees, i.e. trees in which each node has a unique arc connecting all its branches. This implies that the tree is graded[5], with a certain height $k+1$, and that nodes at distance k' from the root have all the same number of leaves $m_{k'}$ below them. Of course, these numbers are integers and decreasing, $n \equiv m_0 > m_1 > \ldots > m_k > m_{k+1} \equiv 1$, and each m_h is a divisor of m_{h-1}.

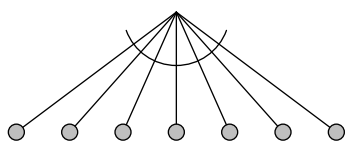
Define the *distance* between leaves on the tree, $d(a,b)$, as the height of the node at which the paths from the root to a and b divide. The symmetry described by a fully symmetric tree implies that an element Q_{ab} in the corresponding matrix only depends on $d(a,b)$, and will be denoted as q_d. From simple combinatorics, given the set of parameters $m_{k'}$, the number of pairs (a,b) at distance d is given by

$$C_d = \frac{n(m_{d+1} - m_d)}{2},$$

from which we deduce that the function $p(Q)$ corresponding to a fully symmetric tree with height $k+1$ is the sum of k deltas, with coefficients

$$p(Q) = \frac{1}{n(n-1)/2} \sum_{d+0}^{k} C_d \delta(Q - q_d) = \frac{1}{1-n} \sum_{d=0}^{k} (m_d - m_{d+1}) \delta(Q - q_d)\,. \qquad (1.41)$$

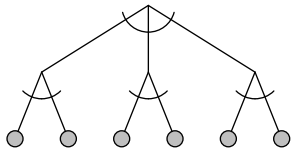
The tree describing the trivial group $G_Q = \mathfrak{S}_n$, corresponding to a matrix of the form of RS Ansatz (1.5.1), is the tree



[5] A tree T, with vertices $V(T)$ and edges $E(T)$, is graded if an height function $h : V(T) \longrightarrow \mathbb{N}$ exists such that $h(v) = 0$ if and only if $v \in V(T)$ is a leaf, and $(v_1, v_2) \in E(T) \Rightarrow |h(v_1) - h(v_2)| = 0$.

As $d(a,b) = 1$ for each pair $a \neq b$, we recover the fact that we have only one free parameter, q_1.

Fully symmetric trees with height 3 describe to the so-called *one-step Replica Symmetry Breaking* (1-RSB) pattern: this ansatz is fully given by a positive integer m_1, which divides n, and two real parameters q_1, q_2, and corresponds to the tree



In the analytic continuation, the constraint on m_1 to be a divisor of n should be discarded. Indeed, performing all calculations, the resulting saddle-point equations inside this pattern yield $m_1 \in [0,1]$, and $q_1 > q_2$. It can be shown that, from a variational point of view, this choice is better than the RS one, and the agreement with numerical simulations is improved. Nevertheless, it still presents the paradox of a negative entropy at zero temperature, so a better ansatz is welcome.

If we considered fully symmetric trees with larger depth (*k-step Replica Symmetry Breaking*, or k-RSB, pattern) we would have a larger number of integer parameters $m_1, \ldots, m_k$, such that m_h divides m_{h+1} for each $h \leq k$, and of real parameters $q_1, \ldots, q_{k+1}$.

Again, in the analytic continuation, we could treat the variables m_h as real numbers. The resulting saddle-point inside this pattern would have $0 \equiv m_0 < m_1 < \ldots < m_k < m_{k+1} \equiv 1$, and $q_h > q_{h+1}$ for each $h \leq k$.

This ideal procedure should suggest that, as we treat the variables m_h as real variables in the analytic continuation, we allow for a limit process in which both q_h and m_h vary in a continuous way, and h is promoted from an integer index to a real continuous variable. This limit process is called *full Replica Symmetry Breaking* (∞-RSB). Define

$$x(q) = \int_0^q \mathrm{d}Q\, p(Q)\,. \tag{1.42}$$

At each finite depth k, the set of allowed $x(q)$ is the set of monotonic k-step functions which map $\mathbb{R}$ into $[0,1]$. The limit $k \to \infty$ allows to deal with generic monotonic functions, and also the distribution $p(q) = \frac{\mathrm{d}}{\mathrm{d}q} x(q)$ admits a continuos part, and is not constrained to be a finite sum of delta functions. If one solves the saddle-point equations, its form turns out to be

$$p(q) = x_{\min}\delta(q - q_{\min}) + x_{\max}\delta(q - q_{\max}) + \widetilde{p}(q)\,, \tag{1.43}$$

where the only singular parts correspond to a maximum and a minimum allowed value of q, while $\widetilde{p}(q)$ is a smooth convex function on the interval $[q_{\min}, q_{\max}]$. It is possible to show that it gives

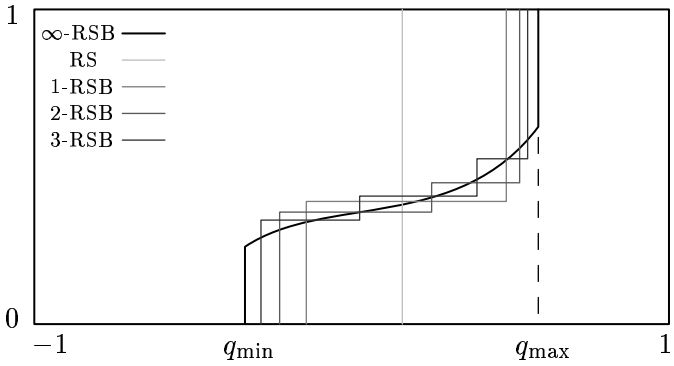


Figure 1.2 Qualitative plot of the function $x(q)$, and of the functions $x_k(q)$ found in the RS pattern, and in the k-RSB pattern, for $k = 1, 2, 3$.

the correct value of free energy and entropy at low temperatures. In k-RSB patterns, with finite k, not only one finds good numerical convergence to the full-RSB results for thermodynamic quantities (free energy, entropy, ...), but also the function $x_k(q)$ is the "best" approximant of the full-RSB limit $x(q)$ in the space of k-step functions (i.e., $x(q)$ is the maximum of a certain functional, and $x_k(q)$ is the maximum of the same functional, restricted to the space of k-step functions, see fig. 1.5). The physical content of this solution, also called the *Parisi solution*, turned out to be quite surprising. We have seen in the previous paragraph that the function $p(q)$ solving saddle point equations was related by formula (1.40) to a physical quantity, the quenched average of the overlap between two configurations. In a similar way, it is possible to find the average values of joint probability that three configurations have a given set of mutual overlaps, i.e. we ask for the probability that the overlap between the first and the second replica is q, between the second and the third is q', between the first and the third is q''; from saddle point equations, with reasonings similar to the ones leading to equation (1.40), one finds:

$$\overline{p_J(q, q', q'')} = p(Q, Q', Q'') \,, \tag{1.44}$$

where $p(Q, Q', Q'')$ is the distribution of the triplets of values which determine diagonal 3×3 submatrices of Q_{ab}:

$$p(Q, Q', Q'') = \lim_{n \to 0} \frac{3!}{n(n-1)(n-2)} \sum_{a<b<c} \delta(Q^*_{ab} - Q)\delta(Q^*_{bc} - Q')\delta(Q^*_{ac} - Q'') \,; \tag{1.45}$$

this quantity can be calculated for a generic $p(q)$ of the form (1.43), not necessarily solving saddle point equations; only at saddle point, instead, one gets the formula:

$$\begin{aligned} \overline{p_J(q, q', q'')} = & \frac{1}{2} p(q) x(q) \delta(q - q') \delta(q - q'') \\ & + \frac{1}{2} \Big(p(q) p(q') \theta(q - q') \delta(q' - q'') \quad + \; (q \leftrightarrow q') \; + \; (q \leftrightarrow q'') \Big) \,, \end{aligned} \tag{1.46}$$

where $x'(q) \equiv p(q)$. This equation states that, given three replicas, and three possible overlaps between pairs of them, either they all have the same overlap, or two overlaps have the same value, and the third one is larger. This implies that the phase space of the SK Model is *ultrametric* w.r.t. the distance defined in (1.36). An ultrametric space is a metric space where the standard triangular inequality is replaced by a stronger one: the distances between any three points a, b, c satisfy

$$d_{ab} \leq \max(d_{bc}, d_{ac}) \,, \tag{1.47}$$

i.e. all triangles are either equilateral, or isosceles with a side smaller than the other two. A simple example of an ultrametric space is, for a given graded tree, the set of leaves, with the distance defined previously. The fully symmetric trees are the graded trees in which for each pair of leaves (a, b) an automorphism ϕ of the tree exists such that $b = \phi(a)$. Ultrametric spaces described by trees only allows for finite sets of distance values. The extrapolation of the structure of fully symmetric trees to the $n \to 0$ limit mathematically corresponds to the natural extension to the whole set of fully symmetric ultrametric structures. Similar calculations can be performed for an arbitrary number of replicas.

Parisi solution turns out to imply a precise picture of configuration space: there are many pure phases, organized in a hierarchical way with a branching structure described by the Parisi order parameter $p(q)$. Indeed, define the thickness of a branch as the total probability of all configurations steming from it. At any level of the hierarchical process, the joint distributions for an arbitrary number of thickness values are known. The functional dependence of the weights from the order parameter turns out to have an universal form, not related to the detailed form of the Hamiltonian, or to the thermodynamic parameters.

As the distibution of the thickness is related to the thermodynamic weight of a certain pure phase, this quantity, too, at any level of the tree (where a continuous branching process is present), can be found from the order parameter. The distibution of the thickness is related to the thermodynamic weight of a certain pure phase, thus for any model with an ultrametric structure of the configuration space, all information on pure phases is contained in the Parisi order parameter.

This surprising form of the phase space is not a peculiarity of the SK model. The whole replica machinery can be applied to different disordered systems. For some of them, the RS ansatz turns out to be the correct one for all temperatures. To this group belongs the problem we will focus on, the Assignment problem [9], and also the famous Travelling Salesman Problem [10], both deriving from the world of Computer Science. For other systems, like the K-SAT problem [11], [12], or the Viana-Bray Model [13] (a finite-connectivity version of SK Model), the correct solution is given by the 1-RSB Ansatz in a low-temperature region, and by the RS Ansatz in a high temperature region. There exist systems, like the p-spin Model [14], where

the structure of the phase space is described by a RS Ansatz in a high temperature region, by an 1-RSB Ansatz at an intermediate range of temperature, and by full RSB Ansatz, in a low temperature region.

2. General theory of the Cavity Method

In the previous chapter we have described the replica approach to disordered systems. The replica method relies on some hypotheses (e.g. the validity of the $n \to 0$ limit procedure) whose physical meaning is not completely clear. Besides, since the very first step is averaging over disorder, it doesn't provide a powerful tool for the problem at given instance.

However, different methods are available for dealing with disordered systems. One of them is the Cavity approach. Even though it has some drawbacks (generally, a heavier notation) this approach has several advantages. First of all, approximations have a clear physical meaning and can be stated in a precise and controlled way. The Guerra-Toninelli approach [15], inspired by the cavity method, have recently led to the result that in SK Model the free energy of the Parisi solution is a rigorous upper bound for the free energy of the model. This approach generalizes to other mean-field disordered models ([16]). Another fruitful application is in finite-connectivity systems [17] Besides, Cavity Equations make sense also at fixed instance of disorder, and this quality makes them a good starting point for algorithm design. This procedure has shown to be surprisingly fruitful in many practical applications as, for example, a powerful algorithm for solving 3-SAT problems or q-Colouring problems [18], [19], [20].

In this chapter we will give a derivation of Cavity Equations valid for a general class of models. We start with a description of the models our derivation applies to, and their factor graph representation; we introduce some hypotheses based on Cluster Property inside each pure state and find the corresponding Cavity Equations, valid in the Replica-symmetric approximation. Then, we discuss a generalization to systems where replica symmetry is spontaneously broken, and we derive the Cavity Equations valid in 1-RSB approximation. In the last section we explain what happens to Cavity Equations when an average over disorder is performed.

2.1 A general class of Random Mean-Field Models

In this chapter we want to deal with a class of statistical mechanics models sufficiently large to include problems like the Sherrington-Kirkpatrick spin glass or the Random K-SAT as prototype

cases.

Variables are defined on a certain (one-variable) state space S, e.g. $S = \{\pm 1\}$ for spin systems, or $S = \{\text{true, false}\}$ for boolean problems. A configuration $\boldsymbol{\sigma} = \{\sigma_i\}_{i=1,\dots,N}$ is an element of S^N. Assume one-variable normalization $\int_S \mathrm{d}\sigma = 1$ so that comparing the free energies of systems with a different number of variables will not lead to irrelevant shifts due to a different number of variables. Integrals should be intended as sums for S discrete. Each interaction a, involving k_a variables, corresponds to a real function $E_a : S^{k_a} \to \mathbb{R}$. One-body terms of the Hamiltonian, corresponding to real functions $W_i : S \to \mathbb{R}$, are understood to sit on each variable, and not labeled by interaction indices. The parameters defining these functions are quenched, i.e., in the spirit of chapter 1, even when the set of parameters is macroscopically large, we will perform statistical averages only on the variables. The generic Hamiltonian is

$$\mathcal{H}(\boldsymbol{\sigma}) = \sum_{a=1}^{M} E_a(\{\sigma_{i_1}, \dots, \sigma_{i_{k_a}}\}) + \sum_{i=1}^{N} W_i(\sigma_i)\,. \tag{2.1}$$

For Statistical Mechanics models on a lattice, the typical picture we have in mind is a regular lattice, in which variables are sitting on the vertices, and the pattern of interaction is quite simple: usually each interaction involves a fixed number of spins, and they relate spins situated nearby on the lattice. For instance, often interactions are sitting on each link (i.e. only first neighbours interact), or on each plaquette, so the lattice picture is sufficient to describe in a complete and intuitive way the whole system. When dealing with disordered systems and models arising in other contexts, such as coding or optimization theory, often variables interact in a complicated and heterogeneous way, thus a more flexible graphical representation is advisable. The common trend in Information Science is on structures named *factor graphs*: given a system with N variables and M interactions, we can represent it via a bipartite graph G, i.e. a graph in which we have vertices $V(G)$ of two species $V_1(G)$ and $V_2(G)$ (say, circles and squares), and edges $E(G)$ only between vertices of different species

$$V(G) = V_1(G) \cup V_2(G)\,; \qquad\qquad E(G) \subseteq V_1(G) \times V_2(G)\,. \tag{2.2}$$

Associate to each variable a circle-vertex i (variable node) and to each interaction a square-vertex a (function node). Draw an edge between a circle and a square if the variable associated with the circle is involved in the interaction associated with the square. With our choice of graphical representation, there are no square vertices corresponding to the one-body interactions $W_i(\sigma_i)$, which are implicitly represented in circle vertices. Call $(\!(i)\!)$ the neighbours of variable node i and $(\!(a)\!)$ the neighbours of function node a. We denote with $c(a)$, $c(i)$ the cardinality of these sets, respectively. A factor graph inherits the natural notion of distance over graphs, i.e., for i, j sites over G, the distance $d(i,j)$ (or $|i-j|$) is defined as the length of the shortest path on the graph going from i to j.

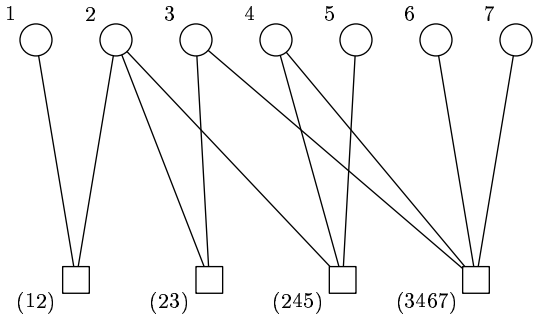


Figure 2.1 Factor graph corresponding to the Hamiltonian of equation (2.3).

As an example, we show in figure 2.1 the factor graph associated to the Hamiltonian for the small spin system

$$\mathcal{H}(\boldsymbol{\sigma}) = \sum_{i=1}^{7} h_i \sigma_i + J_{12}\sigma_1\sigma_2 + J_{23}\sigma_2\sigma_3 + J_{245}\sigma_2\sigma_4\sigma_5 + J_{3467}\sigma_3\sigma_4\sigma_6\sigma_7 \,. \tag{2.3}$$

2.2 The Cavity approach

The basic idea of Cavity Method is that we can understand the properties of a large system of the kind defined in section 2.1 by comparison with systems which only locally differ from the original one (cavity systems). These systems are chosen in order to almost decorrelate certain highly correlated variables close to the deformation, while minimally perturbing other observables, with support far away from the deformation (Cavity assumption).

We denote the set of neighbours of a given interaction a, but one given neighbour i, with the symbol $(\!(a)\!)_i$, and similarly, the set of neighbours of a given variable i, but one given neighbour a, with the symbol $(\!(i)\!)_a$.

We define some Hamiltonians connected to different cavity systems. Given a pair of neighbouring vertex-interaction (i, a), we introduce a further spin variable i_a and consider the Hamiltonian

$$\mathcal{H}_{a \nleftrightarrow i}(\boldsymbol{\sigma}) = \sum_{i} W_i(\sigma_i) + \sum_{a' \neq a} E_{a'}(\sigma|_{(\!(a')\!)}) + E_a(\sigma_{i_a}, \sigma|_{(\!(a)\!)_i}) \,, \tag{2.4}$$

where

$$\boldsymbol{\sigma} = \{\sigma_i\}_{i=1,\ldots,N} \cup \{\sigma_{i_a}\} \,.$$

This Hamiltonian corresponds to the cavity system in which we remove the effect of the interaction a on the spin i. We stress the fact that there is no one-body contribution associated to the

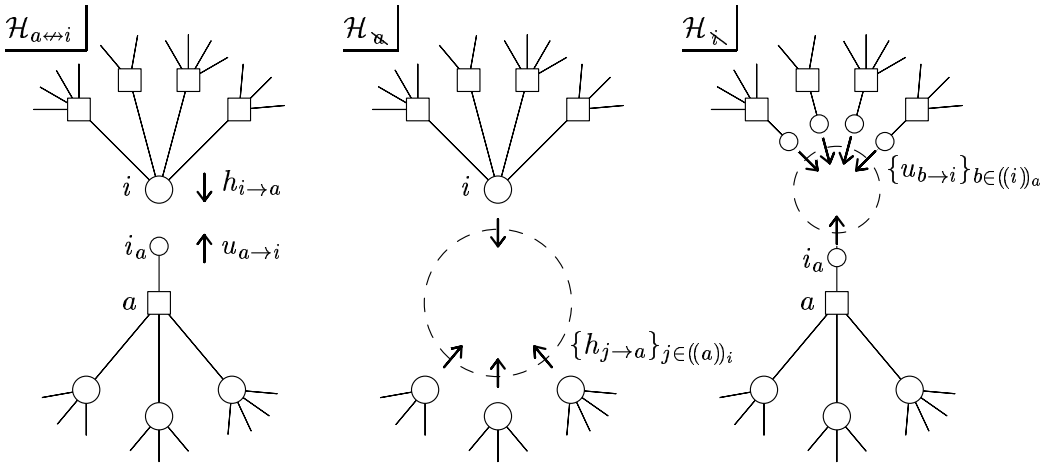


Figure 2.2 A portion of the factor graphs associated to the cavity Hamiltonians of equations (2.4), on the left, (2.5), in the middle, and (2.6), on the right.

new variable i_a. This fact is visualized on factor-graph representation using a dashed circle.[1] A sketch of the factor graph associated to this system, in a neighbourhood of the bond (i, a), compared to the original factor graph, is shown in figure 2.2 (left).

Systems in which a whole variable or a whole interaction are removed can be seen as reiterated applications of this procedure. If, for a given interaction a, we remove its effects on all the neighbouring variables $i \in (\!(a)\!)$, the small system containing interaction a and all the auxiliary variables $\{i_a\}_{i\in(\!(a)\!)}$ is disconnected from the rest of the system, and its contribution to the partition function is trivially factorized. This provides us the recipe to obtain a system with $M-1$ interactions modifying as little as possible the system with M interactions. The corresponding Hamiltonian in which we removed the interaction a is

$$\mathcal{H}_{\not a}(\boldsymbol{\sigma}) = \sum_i W_i(\sigma_i) + \sum_{a'\neq a} E_{a'}(\sigma|_{(\!(a')\!)})\,; \tag{2.5}$$

note that $\boldsymbol{\sigma} = \{\sigma_i\}_{i=1,\dots,N}$. The factor graph associated to this system is shown in figure 2.2 (center).

[1] A hint of motivation for this choice is the following. In the spirit of cavity approximation, we want to consider the minimal deformation for the system in which the "inference channel" (i, a) has been switched off. As there is no natural restriction of the interaction functions E_a to a smaller space $S^{c(a)-1}$, we are forced to introduce the auxiliary variable i_a. The remaining part of the system should be left unperturbed, so on site i we should still have the corresponding one-body term and all the interactions $b \in (\!(i)\!)_a$), while on the auxiliary site i_a we should not have any new one-body term. In this way the cavity system still contains N one-body terms and M interactions.

If, for a given variable i, we remove the effect of all the neighbouring interactions $a \in (\!(i)\!)$ on it, the small system containing only the variable i with its one-body term is disconnected from the rest of the system, and its contribution to the partition function is trivially factorized. This provides us the recipe to obtain a system with $N-1$ variables modifying as little as possible the system with N variables. The corresponding Hamiltonian in which we removed the variable i is

$$\mathcal{H}_{\not i}(\boldsymbol{\sigma}) = \sum_{i' \neq i} W_{i'}(\sigma_{i'}) + \sum_{a \notin (\!(i)\!)} E_a(\sigma|_{(\!(a)\!)}) + \sum_{a \in (\!(i)\!)} E_a(\sigma_{i_a}, \sigma|_{(\!(a)\!)_i}) \,; \tag{2.6}$$

where

$$\boldsymbol{\sigma} = \{\sigma_{i'}\}_{i' \in \{1,\dots,N\} \setminus i} \cup \{\sigma_{i_a}\}_{a \in (\!(i)\!)} \,.$$

The factor graph associated to this system is shown in figure 2.2 (right).

Given an operator $\mathcal{O}$, as usual $\langle \mathcal{O} \rangle$ denotes its expectation value over the Hamiltonian of the system. We denote averages over cavity Hamiltonians with the subscript associated to the corresponding cavity system: $\langle \mathcal{O} \rangle_{a \leftrightarrows i}$ for the system in figure 2.2, $\langle \mathcal{O} \rangle_{\not i}$ for the one in figure 2.2 (right) and $\langle \mathcal{O} \rangle_{\not a}$ for the one in figure 2.2 (center). In particular, F denotes the free energy of the cavity systems, for instance

$$F_{a \leftrightarrows i} = -\frac{1}{\beta} \ln Z_{a \leftrightarrows i} \,, \qquad Z_{a \leftrightarrows i} = \int \mathrm{d}\boldsymbol{\sigma}\, e^{-\beta \mathcal{H}_{a \leftrightarrows i}(\boldsymbol{\sigma})} \,.$$

Consider the local operators $\theta_i(\sigma) = \delta(\sigma_i, \sigma)$, and the algebra of observables containing generic polynomials in these operators. A space $\mathbf{h}$ exists, parametrizing the normalized distributions over S, that is, for all $f(\sigma) : S \to \mathbb{R}^+$ such that $\int_S \mathrm{d}\sigma f(\sigma) = 1$, there exists one and only one $h \in \mathbf{h}$, such that $f(\sigma) = p(\sigma, h)$. For example, if S is discrete and finite, with q elements, the space $\mathbf{h}$ is isomorphic to $\mathbb{R}^{q-1}$, i.e. $q-1$ magnetic fields fully describe the marginal probability over one q-state variable. In particular, in many applications, such as spin models, boolean variables, covering problems etc., $|S| = 2$. In this case, adopting the terminology of magnetic systems, with $S = \{\pm 1\}$, we can write a generic one-body term in the form $W_i(\sigma_i) = h_i \sigma_i$, and a generic normalized distribution as

$$f_h(\sigma) = \frac{e^{-\beta h \sigma}}{2\cosh(\beta h)} \,.$$

Also in the general case, in analogy with magnetic-system jargon, we will deal with (generalized) magnetic fields. In particular, we introduce a set of

magnetic field h_i**:** the element $h_i \in \mathbf{h}$ such that $p(\sigma, h_i) = \langle \theta_i(\sigma) \rangle$, that is, the marginal probability distribution over site i in the original system.

cavity field $h_{i \to a}$**:** the element $h_{i \to a} \in \mathbf{h}$ such that $p(\sigma, h_{i \to a}) = \langle \theta_i(\sigma) \rangle_{a \leftrightarrows i}$, that is, the marginal probability distribution over site i in the cavity system in which interaction

a does not infer directly over site i, which reflects the direct influence on site i of all its neighbouring interactions but a, in the original system.

cavity bias $u_{a\to i}$**:** the element $u_{a\to i} \in \mathbf{h}$ such that $p(\sigma, u_{a\to i}) = \langle \theta_{i_a}(\sigma) \rangle_{a \nrightarrow i}$, that is, in the same cavity system as above, the marginal probability distribution over the site i_a, which reflects the direct influence of only interaction a over site i in the original system.

A pictorial description of cavity fields and cavity bias is given in figure 2.2.

2.3 Single-state Cavity Equations

Informations on the typical length scales in the system are contained in expectation values of higher-order monomials in the characteristic functions θ_i. For example, the connected two-point function

$$\langle \theta_i(\sigma_1)\theta_j(\sigma_2) \rangle^{\text{conn.}} = \langle \theta_i(\sigma_1)\theta_j(\sigma_2) \rangle - \langle \theta_i(\sigma_1) \rangle \langle \theta_j(\sigma_2) \rangle \ , \tag{2.7}$$

in the limit of large distances describes the correlation function of the system

$$G_{\sigma_1,\sigma_2}(x) = \frac{1}{\#\{\text{site pairs at dist. } x\}} \sum_{\substack{i,j \\ d(i,j)=x}} \langle \theta_i(\sigma_1)\theta_j(\sigma_2) \rangle^{\text{conn.}} \ . \tag{2.8}$$

We expect this quantity to decrease for i, j far away inside a pure phase, as a consequence of the Cluster Property discussed in section 1.2. Similar reasonings concern generic connected n-point functions.

The way in which correlation functions decrease with distance in different systems is a delicate topic. In systems whose factor graph representations are random bipartite Erdös-Renyi graphs, the typical size of loops is of order $\ln N$. Furthermore, connected correlation functions are expected to vanish exponentially with distance, and thus to decrease with a power law in N. Also in many fully connected systems (as the SK model, or the Assignment problem), short loops give contribution to the correlation functions negligible w.r.t. the one given by long loops (typically, of lenght order $\ln N$). This is a delicate consequence of the fact that interactions should be properly rescaled with the system size.

In the following, extremizing the Cluster Property statement, we will implicitly assume that for our system correlation functions of operators far away inside a pure phase can be neglected in large N limit. A systematic expansion can be the subject of future investigations (cfr. chapter 7).

Of course, we cannot in general expect that correlation functions of operators located nearby are in some sense small. Nevertheless there is a remarkable exception. Think of local operators, concentrated on a certain region, reproducing a n-point function weighted with the cavity

Hamiltonian in which we have removed that region. As the direct inference among neighbouring sites is switched off, Cluster Property reasonings applied to the cavity system allow to assume decorrelation for operators situated nearby in the original system.

For example, consider a pair of operators $\mathcal{O}_1$, $\mathcal{O}_2$ located nearby in the original system, with Hamiltonian $\mathcal{H}$, but far away in a certain cavity system with Hamiltonian $\mathcal{H}_{[\text{cav.}]} = \mathcal{H} + \Delta\mathcal{H}$. We expect

$$\begin{aligned} e^{-\beta(F-F_{[\text{cav.}]})} \left\langle \mathcal{O}_1 \mathcal{O}_2 e^{-\beta\Delta\mathcal{H}} \right\rangle - e^{-2\beta(F-F_{[\text{cav.}]})} \left\langle \mathcal{O}_1 e^{-\beta\Delta\mathcal{H}} \right\rangle \left\langle \mathcal{O}_2 e^{-\beta\Delta\mathcal{H}} \right\rangle \\ = \langle \mathcal{O}_1 \mathcal{O}_2 \rangle_{[\text{cav.}]} - \langle \mathcal{O}_1 \rangle_{[\text{cav.}]} \langle \mathcal{O}_2 \rangle_{[\text{cav.}]} \simeq 0 \qquad \text{(Cluster Property)} \end{aligned} \tag{2.9}$$

This property, applied to specific choices of local operators, will lead us to Cavity Equations. In the following statement we specify the factorization requirements we need in this procedure.

Ansatz 2.3.1 (Cavity Ansatz) *In the thermodynamic limit, inside a pure thermodynamic state, for the cavity system in which we removed site i we have*

$$\left\langle \prod_{b \in (\!(i)\!)_a} \theta_{i_b}(\sigma) \right\rangle_{\not i} \simeq \prod_{b \in (\!(i)\!)_a} \langle \theta_{i_b}(\sigma) \rangle_{\not i} ; \tag{2.10a}$$

for the system in which we removed interaction a we have

$$\left\langle \prod_{j \in (\!(a)\!)_i} \theta_j(\sigma) \right\rangle_{\not a} \simeq \prod_{i \in (\!(a)\!)_i} \langle \theta_j(\sigma) \rangle_{\not a} . \tag{2.10b}$$

Furthermore, the presence of a larger cavity in a region out of the inference neighbourhood of a given site does not affect the marginal probability on that site

$$\langle \theta_{i_a}(\sigma) \rangle_{\not i} \simeq \langle \theta_{i_a}(\sigma) \rangle_{a \nleftrightarrow i} ; \tag{2.10c}$$

$$\langle \theta_i(\sigma) \rangle_{\not a} \simeq \langle \theta_i(\sigma) \rangle_{a \nleftrightarrow i} . \tag{2.10d}$$

A consequence of the Cavity Ansatz 2.3.1 is the set of equations (Cavity Equations)

$$\begin{cases} p(\sigma, h_{i \to a}) \propto e^{-\beta W_i(\sigma)} \displaystyle\prod_{b \in (\!(i)\!)_a} p(\sigma, u_{b \to i}) , \\ p(\sigma, u_{a \to i}) \propto \displaystyle\int e^{-\beta E_a(\sigma_i = \sigma, \{\sigma_j\})} \prod_{j \in (\!(a)\!)_i} \Big(\mathrm{d}\sigma_j \, p(\sigma_j, h_{j \to a}) \Big) . \end{cases} \tag{2.11}$$

Indeed, the first equation represents the fact that

$$\langle \theta_i(\sigma) \rangle_{a \nleftrightarrow i} = e^{-\beta W_i(\sigma)} \left\langle \prod_{b \in (\!(i)\!)_a} \theta_{i_b}(\sigma) \right\rangle_{\not i} ,$$

while the second equation is derived from

$$\langle \theta_{i_a}(\sigma) \rangle_{a \nleftrightarrow i} = \left\langle e^{-\beta E_a(\sigma_{i_a}=\sigma,\{\sigma_j\})} \prod_{j \in (\!(a)\!)_i} \theta_j(\sigma_j) \right\rangle_{\not a} = \int e^{-\beta E_a(\sigma_{i_a}=\sigma,\{\sigma_j\})} \left\langle \prod_{j \in (\!(a)\!)_i} \theta_j(\sigma_j) \mathrm{d}\sigma_j \right\rangle_{\not a} .$$

The Cavity Equations (2.11) define a set of functionals on the space of cavity fields:

$$\Phi_{i \to a} : \mathbf{h}^{c(i)-1} \longrightarrow \mathbf{h} \qquad \Phi_{i \to a}(\{u_{b \to i}\}_{b \in (\!(i)\!)_a}) = h_{i \to a} ; \tag{2.12a}$$

$$\Psi_{a \to i} : \mathbf{h}^{c(a)-1} \longrightarrow \mathbf{h} \qquad \Psi_{a \to i}(\{h_{j \to a}\}_{j \in (\!(a)\!)_i}) = u_{a \to i} ; \tag{2.12b}$$

They also define two functions for the normalizations involved in equations (2.11)

$$\varphi_{i \to a}(\{u_{b \to i}\}) := -\frac{1}{\beta} \ln \int \mathrm{d}\sigma \, e^{-\beta W_i(\sigma)} \prod_{b \in (\!(i)\!)_a} p(\sigma, u_{b \to i}) ; \tag{2.13a}$$

$$\psi_{a \to i}(\{h_{j \to a}\}) := -\frac{1}{\beta} \ln \int \mathrm{d}\sigma_i \int \prod_{j \in (\!(a)\!)_i} (\mathrm{d}\sigma_j \, p(\sigma_j, h_{j \to a})) e^{-\beta E_a(\sigma_i, \{\sigma_j\})} . \tag{2.13b}$$

Consistently with Cavity Approximation, we can relate the cavity fields to the free energy shifts w.r.t. the free energy of the original system

$$F - F_{a \nleftrightarrow i} \simeq -\frac{1}{\beta} \ln \int \mathrm{d}\sigma \, p(\sigma, h_{i \to a}) p(\sigma, u_{a \to i}) ; \tag{2.14a}$$

$$F - F_{\not i} \simeq -\frac{1}{\beta} \ln \int \mathrm{d}\sigma \, e^{-\beta W_i(\sigma)} \prod_{b \in (\!(i)\!)} p(\sigma, u_{b \to i}) ; \tag{2.14b}$$

$$F - F_{\not a} \simeq -\frac{1}{\beta} \ln \int \prod_{j \in (\!(a)\!)} (\mathrm{d}\sigma_j \, p(\sigma_j, h_{j \to a})) e^{-\beta E_a(\{\sigma_j\})} . \tag{2.14c}$$

We can also interpret the normalizations (2.13) as free-energy shifts

$$\varphi_{i \to a}(\{u_{b \to i}\}) \simeq F_{a \nleftrightarrow i} - F_{\not i} ; \tag{2.15a}$$

$$\psi_{a \to i}(\{h_{j \to a}\}) \simeq F_{a \nleftrightarrow i} - F_{\not a} . \tag{2.15b}$$

In two-state variable applications, we can write the Cavity Equations (2.11) with the language of standard magnetic fields. Say that $W_i(\sigma) = w_i \sigma$, we have

$$\begin{cases} h_{i \to a} = w_i + \sum_b u_{b \to i} , \\ u_{a \to i} = -\dfrac{1}{2\beta} \ln \dfrac{\sum_{\sigma_j} \exp\left(-\beta E_a(+1, \{\sigma_j\}) - \beta \sum_j h_{j \to a} \sigma_j\right)}{\sum_{\sigma_j} \exp\left(-\beta E_a(-1, \{\sigma_j\}) - \beta \sum_j h_{j \to a} \sigma_j\right)} . \end{cases} \tag{2.16}$$

Analogously, when $S = \{0, 1\}$, we can parametrize marginalizations as (for notational convenience we have called variables n)

$$p(n, h) = \frac{e^{-\beta h} + (1 - e^{-\beta h}) \delta(n, 0)}{1 + e^{-\beta h}} , \tag{2.17}$$

and one-body terms as $W_i(n) = w_i n$, equations (2.11) become

$$\begin{cases} h_{i\to a} = w_i + \sum_b u_{b\to i}\,, \\ u_{a\to i} = -\dfrac{1}{\beta}\ln \dfrac{\sum_{n_j} \exp\left(-\beta E_a\left(1,\{n_j\}\right) - \beta\sum_j h_{j\to a} n_j\right)}{\sum_{n_j} \exp\left(-\beta E_a\left(0,\{n_j\}\right) - \beta\sum_j h_{j\to a} n_j\right)}\,. \end{cases} \tag{2.18}$$

2.4 Decomposition of Gibbs measure into pure phases

The cavity factorization hypotheses are strictly related to the Cluster Property of a statistical mechanics system. In general, these hypotheses are not valid for a system showing Spontaneus Symmetry Breaking, where many pure phases are present; nevertheless, they still hold inside each pure phase. Following the seminal idea of Mézard and Parisi [17] we want to extend equations (2.11) to systems with many states, promoting the cavity-field unknowns to *survey* unknowns (probability distributions over pure phases). This approach does not require any procedure for selecting one pure phase (the analogous of a Ruelle criterium, or a Boguliubov criterium, for finite dimensional regular systems), this task being of overwhelming difficulty in spin glasses.

This kind of approach must necessary make use of some properties of pure phases, but we have seen in the previous chapter that any formalization of these concepts for disordered systems is delicate, expecially if we deal with finite-size systems. So the whole discussion, even though it is reasonable and self-consistent from a physical point of view, is not a formal demonstration and is essentially non-rigorous.

The ergodicity breaking and the presence of many pure phases are mathematically depicted as a non-trivial decomposition of Gibbs measure into a sum of measures, each one being almost coincident with the Gibbs measure inside a certain "valley" of the phase space, and almost zero out of the valley, with the property that they are stationary under a local dynamic.

Say we have $\mathcal{N}$ pure phases, labeled with an index $\alpha = 1, \ldots, \mathcal{N}$. The pure-phase measures $\{\mu_\alpha(\boldsymbol{\sigma})\}$ are defined via a partition of the phase space $\{\chi^{(\alpha)}(\boldsymbol{\sigma})\}$, that is a set of positive functions such that $\sum_\alpha \chi^{(\alpha)}(\boldsymbol{\sigma}) = 1$. We define

$$\mu_\alpha(\boldsymbol{\sigma}) = \frac{1}{Z_\alpha}\chi^{(\alpha)}(\boldsymbol{\sigma}) e^{-\beta\mathcal{H}(\boldsymbol{\sigma})}\,; \qquad\qquad Z_\alpha = \int \mathrm{d}\boldsymbol{\sigma}\, \chi^{(\alpha)}(\boldsymbol{\sigma}) e^{-\beta\mathcal{H}(\boldsymbol{\sigma})}\,. \tag{2.19}$$

For each pure phase α, we have a free energy $F_\alpha = -(1/\beta)\ln Z_\alpha$ and, for each observable $\mathcal{O}$, the corresponding expectation value $\langle\mathcal{O}\rangle^{(\alpha)}$ is given by

$$\langle\mathcal{O}\rangle^{(\alpha)} = \int \mathrm{d}\mu_\alpha(\boldsymbol{\sigma})\, \mathcal{O}(\boldsymbol{\sigma})\,. \tag{2.20}$$

The expectation value of the operator in the full Gibbs measure (*symmetric phase*) is a baricentric combination of the expectation values in single pure phases, with weights related to the free energies

$$\langle \mathcal{O} \rangle = \sum_{\alpha} e^{-\beta(F-F^{(\alpha)})} \langle \mathcal{O} \rangle^{(\alpha)} \ . \tag{2.21}$$

We remark that, for a given pair of cavity systems, there is a natural identification of the pure phases in the two perturbed systems.[2] From now on, in the phase labeling for the various cavity systems, this identification is understood.

We expect that in each phase the Cluster Property holds. Of course, it does not hold in the Gibbs Measure, and connected correlation functions do not vanish, so that cavity hypotheses, in particular, are no longer valid.

As we have seen in section 1.1, for traditional lattice systems (whose prototype is the Ising Model), Spontaneus Symmetry Breaking leads to a phase space landscape with a finite number of pure phases, related by an exact symmetry of the Hamiltonian, with exactly the same energy (in Ising Model below the critical temperature, two pure phases, related by the symmetry $\boldsymbol{\sigma} \to -\boldsymbol{\sigma}$). In disordered frustrated systems, a new feature can appear: in a certain region of the parameters space, we can have a number of pure states exponentially large in the system size, with free energy differencies of order 1. For these systems, it is natural to approximate the discrete spectrum of free energies with a continuous distribution, whose logarithm is an extensive quantity, and introduce a complexity function $\Sigma(F)$

$$e^{\Sigma(F)} \mathrm{d}F \ = \#\{\text{pure phases } \alpha : F^{(\alpha)} \in [F, F + \mathrm{d}F\,]\} \, . \tag{2.22}$$

2.5 Survey of Cavity Equations over many states

The cavity fields $h_{i\to a}$, $u_{a\to i}$ of the section 2.2 were defined starting from averages of local operators. The set of hypotheses in Cavity Ansatz 2.3.1 are valid only for averages over the pure phase measures. Thus, for each bond (i, a), instead of $h_{i\to a}$, $u_{a\to i}$, we define a set of cavity fields $\{h^{(\alpha)}_{i\to a}, u^{(\alpha)}_{a\to i}\}_{\alpha=1,\dots,\mathcal{N}}$ starting from averages of the same operators not over the Gibbs measure but over the pure phase measures. Equations (2.11) are still valid for fields at fixed index α.

Consider a given instance of the general model defined in 2.1, with factor graph G. Define the space $\mathbf{H} = \mathbb{R} \times \mathbf{h}^{2|E(G)|}$: a point $\mathfrak{p} \in \mathbf{H}$ is the datum of a free energy and a set of cavity fields, $\mathfrak{p} = (F, \{h_{i\to a}, u_{a\to i}\}_{(i,a)})$. Different free energies and sets of cavity fields, corresponding

[2] This fact is pictorially clear: as pure phases correspond to "valleys" in the phase space, separated by macroscopically large free-energy barriers, a perturbation of order 1 of the Hamiltonian cannot change the pure-phase landscape of the system.

to different pure phases of the systems, can be seen as a set of $\mathcal{N}$ points $\{\mathfrak{p}^{(\alpha)}\}_{\alpha=1,\dots,\mathcal{N}}$ on space $\mathbf{H}$.

For each point $\mathfrak{p}^{(\alpha)}$, we can consider restrictions to subsets of cavity fields. In particular, for each cavity system, it is natural to turn the attention on the restriction of $\mathfrak{p}^{(\alpha)}$ to cavity fields on the boundary of the cavity. Using equations (2.14), we can also reparametrize the free energy of the original system into the free energy of the cavity system

$$\begin{aligned} F^{(\alpha)}_{a\leftrightarrow i} &= F_{a\leftrightarrow i}(F^{(\alpha)}, h^{(\alpha)}_{i\to a}, u^{(\alpha)}_{a\to i})\,, \\ F^{(\alpha)}_{\diagdown} &= F_{\diagdown}(F^{(\alpha)}, \{u^{(\alpha)}_{b\to i}\})\,, \\ F^{(\alpha)}_{\diagup} &= F_{\diagup}(F^{(\alpha)}, \{h^{(\alpha)}_{j\to a}\})\,. \end{aligned}$$

So, for a point $\mathfrak{p}$ in the original set, we are interested in the points

$$\mathfrak{p}_{a\leftrightarrow i} = (F_{a\leftrightarrow i}, h_{i\to a}, u_{a\to i})\,; \tag{2.23}$$

$$\mathfrak{p}_{\diagdown} = (F_{\diagdown}, \{u_{b\to i}\})\,; \tag{2.24}$$

$$\mathfrak{p}_{\diagup} = (F_{\diagup}, \{h_{j\to a}\})\,. \tag{2.25}$$

The sets of points $\{\mathfrak{p}^{(\alpha)}_{a\leftrightarrow i}\}$, $\{\mathfrak{p}^{(\alpha)}_{\diagdown}\}$ and $\{\mathfrak{p}^{(\alpha)}_{\diagup}\}$ must satisfy some set identities reflecting the pointwise (in $\mathbf{H}$) validity of Cavity Equations

$$\{(F^{(\alpha)}_{a\leftrightarrow i}, h^{(\alpha)}_{i\to a}, u^{(\alpha)}_{a\to i})\} \equiv \left\{\left(F^{(\alpha)}_{\diagdown} + \varphi^{(\alpha)}_{i\to a}, \Phi^{(\alpha)}_{i\to a}, u^{(\alpha)}_{a\to i}\right)\right\}\,, \tag{2.26a}$$

$$\{(F^{(\alpha)}_{a\leftrightarrow i}, h^{(\alpha)}_{i\to a}, u^{(\alpha)}_{a\to i})\} \equiv \left\{\left(F^{(\alpha)}_{\diagup} + \psi^{(\alpha)}_{a\to i}, h^{(\alpha)}_{i\to a}, \Psi^{(\alpha)}_{a\to i},\right)\right\}\,, \tag{2.26b}$$

where

$$\begin{aligned} \varphi^{(\alpha)}_{i\to a} &= \varphi_{i\to a}(\{u^{(\alpha)}_{b\to i}\}_{b\in(\!(i)\!)_a})\,, & \Phi^{(\alpha)}_{i\to a} &= \Phi_{i\to a}(\{u^{(\alpha)}_{b\to i}\}_{b\in(\!(i)\!)_a})\,, \\ \psi^{(\alpha)}_{a\to i} &= \psi_{a\to i}(\{h^{(\alpha)}_{j\to a}\}_{j\in(\!(a)\!)_i})\,, & \Psi^{(\alpha)}_{a\to i} &= \Psi_{a\to i}(\{h^{(\alpha)}_{j\to a}\}_{j\in(\!(a)\!)_i})\,. \end{aligned}$$

Note that equations (2.26) are set identities. Suppose to introduce a canonical ordering in each set, for example ordering the free energies: then, the identities (2.26) are not pointwise, since a reshuffling of free energies from a cavity system to another is possible. The comprehension of this mechanism is a crucial point in deriving Survey Propagation Equations.

The requirement 2.26 that two finite sets of points coincide is technically difficult to handle. If we deal with continuous distributions on the spaces $\mathbb{R}\times\mathbf{h}^c$, with c of order one, we can use the methods of differential calculus. This step is possible only in certain situations. In particular, we focus on a system with a macroscopic number of phases, where the spectrum of free energies is a continuous function in the large N limit, in the scheme depicted on section 2.4. If the points are canonically ordered w.r.t. the free energy parameter, they can be described by a point process

of exponentially large rate. We choose instead to describe the two sets $\{\mathfrak{p}_{i}^{(\alpha)}\}$ and $\{\mathfrak{p}_{a}^{(\alpha)}\}$ with a continuous density

$$\rho_a(F_a,\{h_{j\to a}\}) = e^{\Sigma_a(F_a)} Q_a(\{h_{j\to a}\}|F_a)\,, \qquad \int_{\mathbf{h}^{c(a)}} \mathrm{d}\vec{h}\, Q(\vec{h}|F) = 1\,; \tag{2.27a}$$

$$\rho_i(F_i,\{u_{b\to i}\}) = e^{\Sigma_i(F_i)} P_i(\{u_{b\to i}\}|F_i)\,, \qquad \int_{\mathbf{h}^{c(i)}} \mathrm{d}\vec{u}\, P(\vec{u}|F) = 1\,. \tag{2.27b}$$

The consistency requirement (2.26) implies, in particular, for each edge (i,a), a set identity on the Laplace Transforms w.r.t. the free energy:

$$\begin{aligned}
&\sum_{\mathfrak{p}_{a\rightarrow i}^{(\alpha)}} e^{-yF_{a\rightarrow i}^{(\alpha)}}\ \delta\Big(h, h_{i\to a}^{(\alpha)}\Big)\ \delta\Big(u, u_{a\to i}^{(\alpha)}\Big) \\
&\quad = \sum_{\mathfrak{p}_{i}^{(\alpha)}} e^{-y(F_i+\varphi_{i\to a}(\{u_{b\to i}^{(\alpha)}\}_{b\in(\!(i)\!)_a}))}\ \delta\Big(h, \Phi_{i\to a}(\{u_{b\to i}^{(\alpha)}\}_{b\in(\!(i)\!)_a})\Big)\ \delta\Big(u, u_{a\to i}^{(\alpha)}\Big) \\
&\quad = \sum_{\mathfrak{p}_{a}^{(\alpha)}} e^{-y(F_a+\psi_{a\to i}(\{h_{j\to a}^{(\alpha)}\}_{j\in(\!(a)\!)_i}))}\ \delta\Big(h, h_{i\to a}^{(\alpha)}\Big)\ \delta\Big(u, \Psi_{a\to i}(\{h_{j\to a}^{(\alpha)}\}_{j\in(\!(a)\!)_i})\Big)\,.
\end{aligned} \tag{2.28}$$

Plugging the expressions (2.27) for the densities we get

$$\begin{aligned}
&\int \mathrm{d}F_a \int \Big(\prod_{j\in(\!(a)\!)_i} \mathrm{d}h_{j\to a}\Big) \exp\Big(-y(F_a+\psi_{a\to i}(\{h_{j\to a}\})) + N\Sigma_a(F_a/N)\Big) \\
&\qquad\qquad \cdot\,\delta\Big(u_{a\to i}, \Psi_{a\to i}(\{h_{j\to a}\})\Big)\ Q_a(\{h_{j\to a}\}|F_a) \\
&= \int \mathrm{d}F_i \int \Big(\prod_{b\in(\!(i)\!)_a} \mathrm{d}u_{b\to i}\Big) \exp\Big(-y(F_i+\varphi_{i\to a}(\{u_{b\to i}\})) + N\Sigma_i(F_i/N)\Big) \\
&\qquad\qquad \cdot\,\delta\Big(h_{i\to a}, \Phi_{i\to a}(\{u_{b\to i}\})\Big)\ P_i(\{u_{b\to i}\}|F_i)
\end{aligned} \tag{2.29}$$

We now introduce some hypotheses in order to make equation (2.29) more tractable. These hypotheses are related to the assumption that the phase space is of the 1-RSB kind.

- Because of the concavity of the complexity function, the parameter y (named the *reweighting parameter*) selects a narrow interval of free energies relevant for the integrals in (2.29). A typical free energy F^* exists such that for all cavity complexity functions

$$\frac{\partial}{\partial F}\Sigma_a(F)\bigg|_{F^*} \simeq \frac{\partial}{\partial F}\Sigma_i(F)\bigg|_{F^*} \simeq y\,, \tag{2.30}$$

  up to corrections of order $1/N$.

- We neglect the F dependence of the quantities Q and P, in the neighbourhood of F^*

$$\frac{\partial}{\partial F}\ln Q_{a}(\{h_{j\to a}\}|F)\Big|_{F^*} = \mathcal{O}\left(\tfrac{1}{N}\right)\,, \qquad \frac{\partial}{\partial F}\ln P_{i}(\{u_{b\to i}\}|F)\Big|_{F^*} = \mathcal{O}\left(\tfrac{1}{N}\right)\,. \tag{2.31}$$

  We thus define the densities

$$Q_{a}(\{h_{j\to a}\}) := Q_{a}(\{h_{j\to a}\}|F^*)\,; \qquad P_{i}(\{u_{b\to i}\}) := P_{i}(\{u_{b\to i}\}|F^*)\,. \tag{2.32}$$

- We consider the functions (2.32) almost factorized

$$Q_{a}(\{h_{j\to a}\}) \simeq \prod_{j\in(\!(a)\!)} Q_{j\to a}(h_{j\to a})\,, \qquad P_{i}(\{u_{b\to i}\}) \simeq \prod_{b\in(\!(i)\!)} P_{b\to i}(u_{b\to i})\,, \tag{2.33}$$

  with each $Q_{j\to a}(h_{j\to a})$ and $P_{b\to i}(u_{b\to i})$ normalized on $\mathbf{h}$.

Under these hypotheses equation (2.29) considerably simplifies. Integrating both sides over the variable $u_{a\to i}$, on the left side the delta function cancels out, while on the right side, using the factorization of P_{i} and the normalization of $P_{a\to i}(u_{a\to i})$, it just produces a factor 1. Then, on the left side we can integrate over the variables $\{h_{j\to a}\}_{j\in(\!(a)\!)_i}$. This gives a prefactor depending only on y

$$\int\Big(\prod_{j\in(\!(a)\!)_i} \mathrm{d}h_{j\to a}\, Q_{j\to a}(h_{j\to a})\Big) e^{-y\psi_{a\to i}(\{h_{j\to a}\})} = C_{ai}(y)\,. \tag{2.34}$$

Finally, also integration over free-energy variables F_{a} and F_{i} produces two prefactors depending only on y

$$\int \mathrm{d}F_{a}\, e^{-yF_{a}+N\Sigma_{a}(F_{a}/N)} = C_a(y)\,, \qquad \int \mathrm{d}F_{i} e^{-yF_{i}+N\Sigma_{i}(F_{i}/N)} = C_i(y)\,. \tag{2.35}$$

Since $Q_{i\to a}$ and $P_{a\to i}$ are normalized, these prefactors do not play a significant role. Eventually, we are left with the so-called *1-RSB Cavity Equations* (or *Survey Propagation Equations*)

$$Q_{i\to a}(h_{i\to a}) \propto \int \prod_{b\in(\!(i)\!)_a} \Big(\mathrm{d}u_{b\to i}\, P_{b\to i}(u_{b\to i})\Big)\; e^{-y\varphi_{i\to a}(\{u_{b\to i}\})}\; \delta\Big(h_{i\to a}, \Phi_{i\to a}(\{u_{b\to i}\})\Big)\,, \tag{2.36a}$$

$$P_{a\to i}(u_{a\to i}) \propto \int \prod_{j\in(\!(a)\!)_i} \Big(\mathrm{d}h_{j\to a}\, Q_{j\to a}(h_{j\to a})\Big)\; e^{-y\psi_{a\to i}(\{h_{j\to a}\})}\; \delta\Big(u_{a\to i}, \Psi_{a\to i}(\{h_{j\to a}\})\Big)\,. \tag{2.36b}$$

These equations have many points of contact with the Cavity Equations (2.11). The main differencies are essentially two:

- the cavity fields $h_{i\to a}$ and $u_{a\to i}$ are replaced by *surveys* of cavity fields, related to the process of handling the various pure phases simultaneously;
- a new factor $e^{-y\Delta F}$ accounts for the reweighting of pure phases due to free-energy shifts. In a numerical implementation, the reweighting parameter y must be tuned on the value of $\frac{\mathrm{d}}{\mathrm{d}F}\Sigma(F)$ in correspondence of the expected free-energy window explored by the algorithm in that moment.

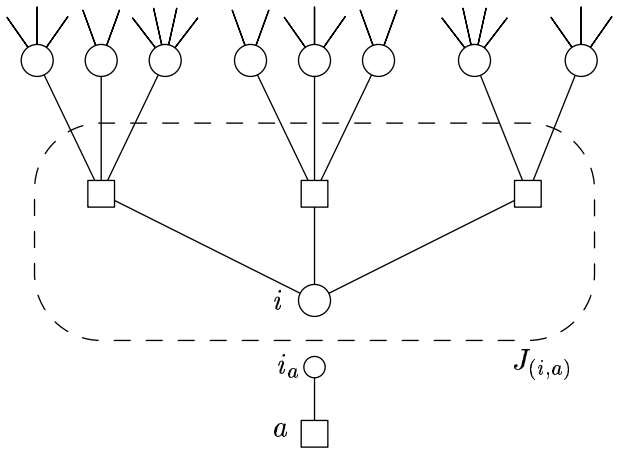


Figure 2.3 Subset $J_{(i,a)}$ of nodes of the factor graph associated to the pair (i, a).

2.6 Cavity Equations as Distributional Equations

So far, we have discussed Cavity Equations in systems with a phase space of the RS or 1-RSB kind, valid for a given instance of the disorder. We now want to perform an average over the ensemble of the disorder. From now on, we will suppose the validity of RS equations (2.11), for simplicity of discussion, and since in the rest of the work, this will be the case.

Call $J = \{E_a, W_i\}$ the set of parameters describing the Hamiltonian, sampled with a certain measure $\mu(J)$. Given a pair (i, a), consider the subset $J_{(i,a)} \subseteq J$

$$J_{(i,a)} = (W_i, \{E_b\}_{b \in (\!(i)\!)_a}) \,. \tag{2.37}$$

The set $J_{(i,a)}$ corresponds to the set of interactions involving only the vertex i and its first neighbours on the cavity system $\mathcal{H}_{a \nleftrightarrow i}$ (see figure 2.6). We make some assumptions on the measure induced on this set by the measure on the whole instance, $\mu(J)$:

- The probability distribution on $J_{(i,a)}$ does not depend on the rest of the system

$$\text{prob}\left(J_{(i,a)} | \, J \setminus J_{(i,a)}\right) = p_{(i,a)}\left(J_{(i,a)}\right) \,. \tag{2.38}$$

- There is a macroscopic number NM of pairs (i, a), so a priori we have to deal with a macroscopic number of probability distributions $p_{(i,a)}\left(J_{(i,a)}\right)$. We assume that the measure on the instances $\mu(J)$ has a sufficiently large group of invariances, so that a small number of pair classes exists, such that $p_{(i,a)}\left(J_{(i,a)}\right) = p_{(i',a')}\left(J_{(i',a')}\right)$ if (i, a) and (i', a') are in the same class. When averaging over the instances, eventually we will have a distributional equation per class of pairs.

- This probability is factorized over each interaction and one-body term appearing in the system:

$$p_{(i,a)}\left(J_{(i,a)}\right) = p_1\left(W_i, c(i)\right) \prod_{b=1}^{c(i)-1} p_2\left(E_b | W_i, c(i)\right) . \tag{2.39}$$

The first assumption is essentially an hypothesis of independence. It is easily verified in systems defined over random graphs *à la* Erdös-Renyi. Furthermore, it is verified also in some systems for which the measure over the factor graph has both a deterministic and a stochastic part, for example, in the Assignment Problem, where the factor graph does not change from instance to instance, while the values of one-body terms are random, i.i.d. quantities.

The second assumption is an homogeneity hypothesis for the measure over the instances. It is typically true for all random systems, with only one class of variable-interaction pairs. An example in which there are two classes is given by the variant of Assignment Problem on rectangular grids, with N_1 rows and N_2 columns: it is necessary to make a distinction between links connecting variables to row constraints, and links connecting variables to column constraints, as the factor graph of the problem, which does not change from instance to instance, contains "row" interaction-nodes with coordination N_2 and "column" interaction-nodes with coordination N_1. To avoid notation complicancies, we will deal with the case in which all edges are equivalent.

The third assumption is delicate. The expression itself is slightly ill-defined, as on the right side we have a product over all interactions *but one*. Suppose to have a one-parameter family of Hamiltonians of the form

$$\mathcal{H}(\boldsymbol{\sigma}) = \sum_{a=1}^{M_1} E_a(\{\sigma_{i_1}, \ldots, \sigma_{i_{k_a}}\}) + t \sum_{a'=M_1+1}^{M_2} E_{a'}(\{\sigma_{i_1}, \ldots, \sigma_{i_{k_{a'}}}\}) + \sum_{i=1}^{N} W_i(\sigma_i) . \tag{2.40}$$

We have a system in our general form, but when $t \neq 0$ there are M_2 interactions, while when $t = 0$ there are only M_1 interactions. We expect the distribution of the cavity fields to vary in a continuous way with t, while equation (2.39) suggests a qualitative difference at $t = 0$. The proper way to proceed is to recognize that there are two classes of pairs, the ones involved in an interaction not scaled with t, and the ones involved in an interaction scaled with t. We would have a probability distribution per each class

$$p_{(i,a)}\left(J_{(i,a)}\right) = p_1\left(W_i, c(i), c'(i)\right) \prod_{b=1}^{c(i)-1} p_2\left(E_a | W_i, c(i)\right) \prod_{b'=1}^{c'(i)} p'_2\left(E_{a'} | W_i, c(i)\right) ; \tag{2.41a}$$

$$p'_{(i,a')}\left(J_{(i,a')}\right) = p_1\left(W_i, c(i), c'(i)\right) \prod_{b=1}^{c(i)} p_2\left(E_a | W_i, c(i)\right) \prod_{b'=1}^{c'(i)-1} p'_2\left(E_{a'} | W_i, c(i)\right) . \tag{2.41b}$$

Now the apparent paradox above is solved: the distribution of cavity biases outcoming from interactions not scaled with t varies in a continuous way, while the distribution of cavity biases outcoming from interactions scaled with t concentrates on zero.

Accounting for this fact is not necessary in many cases. Indeed, in the case in which the average coordination is macroscopically large, the product on all interactions *but one* in (2.39) can be replaced by a product on all interactions, with a negligible error. This is the case of Assignment Problem, of SK Model and of p-spin Models.

Besides, for finite-connectivity models like Viana-Bray Model, K-SAT, $(2+p)$-SAT, in which all families γ of interaction functions are distributed independently, the conditional probability distributions for the coordinations of function nodes in each family, of the form

$$p_1\left(W_i, \{c_{\gamma'}(i) - \delta(\gamma', \gamma)\} \,|\, a \text{ is of kind } \gamma\right) ,$$

do *not* depend on index γ. Moreover, if the coordinations are Poissonianly distributed, for each family γ

$$p_1\left(W_i, \{c_{\gamma'}(i)\}\right) = p_1\left(W_i, \{c_{\gamma'}(i) - \delta(\gamma', \gamma)\} \,|\, a \text{ is of kind } \gamma\right) . \tag{2.42}$$

So, in all these cases, the distribution of the shapes of the neighbourhood of i in the cavity system $\mathcal{H}_{a \nleftrightarrow i}$, conditioned to having a given interaction in node a involving variable i, coincides with the distribution of the shapes of the neighbourhood of a variable i in the whole system.

Equations (2.11) can be summarized as

$$h_{i\to a} = \Phi_{i\to a}\left(\left\{\Psi_{b\to i}(\{h_{j\to b}\}_{j\in(\!(b)\!)_i})\right\}_{b\in(\!(i)\!)_a}\right) , \tag{2.43}$$

for any choice of $J_{(i,a)}$. When averaging over disorder, this equation turns into a self-consistency equation for the probability distribution $f(x)$ of the cavity fields, assuming that they are distributed indipendently, consistently with Cavity Ansatz 2.3.1. The functionals $\Phi_{i\to a}$ and $\Psi_{b\to i}$ depend respectively on the elements W_i and E_b of $J_{(i,a)}$, which are distributed according to (2.39), so, with abuse of notation, we could say that $\Phi_{i\to a}$ is distributed with $p_1(\Phi)$, and $\Psi_{b\to i}$ is distributed with $p_2(\Psi|\Phi_{i\to a})$. Equation (2.43) thus is promoted to a *distributional* equation, determining the distribution of the field:

$$x \stackrel{d}{=} \Phi\left(\{\Psi_b(\{x_b\})\}\right) , \tag{2.44}$$

with x and x_b i.i.d. with $f(x)$, and Φ and Ψ_b distributed as described above.

3. A short introduction to Optimization Theory

The theory of disordered system has recently received great attentions from the scientific community because of its connections with optimization problems in information theory. The theoretical and algorithmic treatment of these problems has a practical counterpart in the design of performing algorithms for applications in everyday life. This chapter briefly presents the basic definitions of the theory of algorithmic complexity. In particular, we focus our attention on the class of Polynomial and Non-deterministic Polynomial problems. In the last section, we explain how an optimization problem can be restated, and possibly solved, in terms of a physical disordered system, and we sketch some of the possible applications of physical concepts to algorithm design.

3.1 The theory of Algorithmic Complexity

The theory of Algorithms concerns the solution of complex problems via a prescription on a sequence of elementary steps (called an *algorithm*). The most varied and heterogeneous problems can be treated. In order to allow for an unitary mathematical formalization, it is useful to define classes of problems sharing the same definition frame. Algorithmic Complexity [21], in particular, studies time employed by an algorithm to solve a given problem. Of course we are not interested in a detailed quantification of the amount of time: we would rather know the *asymptotic behaviour* of algorithm time, with respect to the *size of the problem.* So Algorithmic Complexity focuses on problems for which a size can be reasonably defined, and arbitrarily large problem instances can be given, in order to study asymptotics.

In our treatment we mainly deal with problems such that, at given size N, a set X of *feasible solutions* exists, of exponentially large size, $|X| \sim \exp(P(N))$ with $P(N)$ a power-law at leading order, and the problem can be stated as a search problem for a solution in X. We define two classes in this framework: *optimization problems* and *satisfiability problems.* For optimization problems, there is a cost function $C : X \to \mathbb{R}$, and the problem is solved when a solution $x_s \in X$ such that $C(x_s) = \min_{x \in X} C(x)$ is found. For satisfiability problems, there is both a

cost function C and a threshold a, and the problem is solved either when a solution $x_s \in X$ such that $C(x_s) < a$ is found, or a proof that such a solution does not exists (called a *certificate*) is given.

For example, one of the most famous optimization problems is the *Travelling Salesman Problem*: given a set of n points (*towns*, in a pictorial description), a cost is associated to each pair of points (the *length of the road*). The set X is the set of all possible connected tours passing only once from each town and connecting all of them, and the cost function is the total length of the tour. In this example, the problem is fully described by the set of $N = n(n-1)/2$ pairwise lengths, and the number of possible tours is $\frac{1}{2}(n-1)!$, so $|X| \sim \exp(\sqrt{N} \ln N)$.

Optimization problems and satisfiability problems are intimately related. Is is obvious that the solution of an optimization problem contains the solutions for all the related satisfiability problems, i.e. with the same cost function, in the whole range of thresholds a. A less obvious fact is that a "fast" algorithm for a satisfiability problem in the whole range of thresholds typically can be trivially adapted to a "reasonably fast" algorithm for the associated optimization problem. If reasonable bounds on the range of values of $\min_X C(x)$ can be given, then the optimization problem can be solved iteratively applying the algorithm for the satisfiability problem, bisecating over the range of thresholds. This would be the case, for example, of the Travelling Salesman Problem, with integer distances: as the cost function is the sum of n terms, the interval between a trivial lower-bound and a trivial upper-bound, say the sum of the n shortest lengths and the sum of the n longest lengths, is at most of order nL, with L the difference between the largest distance and the shortest distance. Bisecating over a range of nL values involves $\log_2(nL)$ iterations in the worst case, which is certainly a very small complexity, w.r.t. the intrinsic complexity of the satisfiability version of the Travelling Salesman Problem.

We have proved that, for the Travelling Salesman Problem with integer lengths and bounded largest-distance – shortest-distance interval, the complexity of *the best* algorithm for the optimization problem is not smaller than the complexity of *any* algorithm for the satisfiability problem, and the complexity of *the best* algorithm for the satisfiability problem is not smaller than the complexity of *any* algorithm for the optimization problem, up to a factor $\log_2(nL)$. The proof has been performed rephrasing a problem of the first kind as a problem of the second kind, and vice versa.

In general, studying the complexity of one algorithm for a given problem is at its worst technically complicated. On the other side, the algorithmic complexity *of the problem itself*, defined via a minimization over the whole set of possible algorithms for the problem, is typically a very delicate question.

3.2 Polynomial and Non-deterministic Polynomial Algorithmic Classes

The knowledge of the algorithmic complexity of a family of problems induces a fine hyerarchy of classes, depending from the asymptotic resolution time for each problem: we would thus deal with *linear-time* problems, *quadratic-time* problems, and so on. It is intrisecally difficult to perform such a fine classification, but for particularly simple problems for which a direct analysis is possible.

Besides, the precise mathematical definition of the algorithmic time, involved in the determination of a problem's complexity, depends on the mathematical formalization of the concept of algorithm. Several possibilities exist [22], but it can be proven that all consistent definitions are equivalent up to polynomial factors in the resulting complexity.

In the previous section, we have shown how it is possible to solve a satisfiability problem, if we dispose of an algorithm for solving the corresponding optimization problem. We can reduce each instance of the saisfiability problem to an instance of the optimization one: we thus say that the satisfiability Travelling Salesman Problem is *reducible* to the optimization TSP. We have also seen that the converse reduction holds up to a logarithmic factor. If, in the spirit of the previous discussion, we are interested in the complexity up to polynomials in the problem size, it is natural to extend the concept of reducibility to problem pairs such that the complexity of the firt bounds the complexity of the second up to a polynomial factor. Naturally, reducibility is transitive. Besides, if problem A is reducible to problem B, and B is reducible to A, A and B are equivalent under their complexity aspect.

It is thus relatively easy, using reducibility arguments, to determine if the complexity of a given problem is polynomial, or to state that the complexities of two given problems must be either both polynomial, or both non-polynomial.

So, one is naturally led to dealing with a large class of *Polynomial-time* problems (P problems), insensitive of the degree of the polynomial describing the complexity: any problem belonging to this class can be solved by an algorithm working in polynomial time. A relevant class of problems is defined in an operative way: define the class of *Non-deterministic Polynomial-time* problems (NP problems) as the class of search problems such that, given a feasible solution x_s, the subproblem of checking whether x_s is a solution or not is a polynomial problem. Clearly, a polynomial problem must also be in the NP class. A main question is whether the inclusion of class P in class NP is tight or if there are NP problems which are intrinsically non-polynomial.

This question is of great relevance, and is one of the main open questions of Computer Science. Indeed, the class NP contains a large number of problems of practical relevance: among them, for instace, is the Travelling Salesman Problem discussed above, and the Boolean K-

Satisfiabilily problem (K-SAT problem) for $K \geq 3$, which is stated in the form of a logic boolean expression, and thus is naturally related to computer implementation.

Reducibility arguments induce also the introduction of the concept of *completeness* for problems in a given class: given a complexity class of problems, we define the subclass of *complete* problems as the set of problems to which *all* problems in the class can be reduced, up to polynomial factors in the complexity. For a given class, a priori, this subclass can be empty. But, if there is at least one representative, reducibility arguments can enlarge it via the obvious transitivity property.

The K-Satisfiabilily problem, discussed before, is of great importance also for a reason transcending its practical applications: it is NP-complete. This fact has been proved by Cook [23], starting from the formal definition of problems in NP, and of non-deterministic Turing Machines. The proof of Cook's Theorem establishes that K-SAT is NP-complete by showing reduction of general SAT to K-SAT, and of the formal class of NP-complete problems to SAT. This last step relies on the complete generality of formulation of boolean satisfiability instances, which mimic any abstract non-deterministic Turing Machine. Using reducibility arguments, also the TSP can be shown to be NP-complete.

Although the question whether P$\equiv$NP or P$\subset$NP is open, it is widely believed that NP problems are *not* polynomial problems. Heuristic arguments give a hint that their algorithmic hardness is due to some intrinsic structure of the problem. Nonetheless, formalizing these arguments seems a hard task, and one of the main difficulties lies of the fact that P and NP problems can often be stated in a very similar way. As an example, consider the following problems. For a given graph $\mathcal{G}$, with set of edges $E(\mathcal{G})$, and a length function $l : E \to \mathbb{N}$, we can ask if there is a self-avoiding path connecting two given vertices whose total length is

- smaller than a given threshold $L_{\max}$;
- larger than a given threshold $L_{\min}$.

Despite of the formal similarity of the two problems, it can be shown that the first one is in P, while the second one is NP-complete.

3.3 Average case analysis of algorithmic complexity

So far, we have concentrated on defining the algorithmic complexity of a problem. We have focused on algorithms capable of solving a problems for any instance, since the definition of algorithmic complexity is based on a worst-case analysis.

Nevertheless, in many NP problems, a deep exploration of the possible instances of given size show that, in certain "regions" of this space the problem is computationally "easy", while

the "hard" instances concentrate on a region where performing the search of a solution becomes structurally difficult. For example, in the Travelling Salesman Problem, if we force the towns to be on a planar map, triangular inequality between distances must be satisfied, and the notion of nearby towns allows for high-performance deterministic algorithms. In small words, if we grow the path connecting the endpoint to one of the nearest unvisited towns, and trying to leave at the end only towns in the same region, typically we achieve a good final cost. On the other side, if distances are simply randomly chosen, partial paths created connecting the endpoint to one of the nearest unvisited town easily get stuck at the end, because of the lack of a global strategy based on the geometric structure.

It thus becomes relevant to dispone of an *average case analysis* of the difficulty of a problem, depending on the *ensemble* of problem instances. Indeed, for many practical purposes, beside the formal optimization problem, a "reasonably good" solution will do nearly as well (say, a solution of the satisfiability problem, with a threshold depending from the application), with a cost "close" to the optimal one. In order to apply this strategy, it is advisable to know *a priori* the expected value of the optimal cost, and how much it fluctuates from instance to instance: this information is contained in an average-case treatment of the problem. Another situation in which average-case analysis becomes an important tool is the case in which we search for the most-performing algorithm meeting certain requirements on the time employed (say, for the Travelling Salesman Problem, the algorithm, in the class of quadratic-time algorithms, which on average finds the shortest path).

In this scenario, heuristic incomplete algorithms become relevant. Such an algorithm is typically highly performant for certain ensemble of instances, and poorly performant for other ones, and it is frequent that an average-case analysis highlights the mechanisms for which the algorithm gives the wrong answer, or in case gets stuck.

Once we have motivated the study of average-case complexity, we want to show how this study is strictly related to a statistical-mechanics rephrasing of the Computer-Science problem.

3.4 Combinatorial Optimization and Statistical Mechanics

It is easy to see that an optimization problem can be stated as a physical problem. Indeed, the set of possible solutions can be interpreted as a configuration space, and its cost can be chosen as the Hamiltonian of the system: cost minimization turns into finding the ground state of the physical system, when frozen at zero temperature.

The cost function, as we said before, depends on a large set of parameters, and one is interested in the average case w.r.t. a measure on the parameter space. Thus, the physical system is a disordered system, in the sense described in section 2.6, where the probability

measure aver the disorder corresponds to this measure on the parameter space, i.e. on the space of possible instances.

The "hardness" of a hard optimization problem is related to the fact that the corresponding disordered system is frustrated, and, as the ground state configuration is not the one which simply minimizes all local interactions (as usually happens in ordered Ising-like systems), it is possible that a local algorithm for finding the ground-state does not exist.

Note that the probability measure over the disorder has a crucial role: changing the measure can change the statistical properties of the system, analogously to the fact that "hard" problems can be "easy" for certain regions of the space of instances.

Also in this case, as happens in physical problems on disordered systems, it is interesting to study the problem both at fixed instance, and averaging over disorder. In the first case, a statistical mechanics approach can be fruitful in order to design good heuristic algorithms based on some physically meaningful approximations. In the second, one can answer to the important questions of average-case complexity described in the previous section.

For what concerns algorithm design, two major applications have been achieved. One is the use of the so-called *simulated annealing* algorithms. Since these kind of problems usually are frustrated, algorithms trying to minimize locally the energy often get trapped in a local minimum of the cost function, and the resulting configuration not only has a cost higher than the optimal one, but also is very different from the ground state. A solution to this problem comes from the physical formulation: if we introduce a temperature, and thus allow for thermal fluctuations, it is possible to avoid being stuck in local minima, exploring a larger part of the configuration space. When the temperature decreases, first the system gets trapped in the valley of the energy landscape containing the minimum, then, inside the valley, a further cooling allows to find the ground state. In case this procedure should fail, a "reheating" can be performed, in order to select a new valley. How this annealing is to be done is in general a difficult question that can be answered with a careful statistical analysis of the whole phase space. It is thus clear that the physical formulation of optimization problems becomes interesting even at non-zero temperature, also from an informatics point of view.

A second class of algorithms based on a statistical mechanics approach is very recent, and is inspired by the cavity method of chapter 2. The cavity method approach discussed therein allows to analyse also single instances of a given disordered system, and leads to self-consistency relations for the set of cavity fields (equations (2.11)), when the phase space is described by a RS-pattern, or for the surveys of cavity fields over various pure phases (equations (2.36)), when the phase space is described by a 1-RSB-pattern. It is possible to write an heuristic algorithm which uses cavity equations (2.11) or (2.36), at zero temperature, to find the minimum cost solution [18, 24, 20].

In fact, programmers have been using for a long time an heuristic message-passing algorithms, called *Belief Propagation*, which have been recognize to correspond to cavity equations in the RS approximation [25, 26]. Belief Propagation is highly performing on certain problems, and low performing on other problems. The physical interpretation of this fact is that Belief Propagation fails when the phase-space landscape contains many phases, and that Survey Propagation Algorithm is the proper generalization to 1-RSB landscapes.

The first application of Survey Propagation has been to the 3-SAT problem. We shortly describe here this exemplar case. An admissible configuration is a set of N boolean variables $x_i \in \{0,1\}^N$, and it must satisfy a set of M logical constraints, or *clauses*, involving three literals (say, a clause could be $x_5 \vee \overline{x}_8 \vee x_{18}$). Each clause can be viewed as an interaction term of a Hamiltonian, so the constraint that all clauses are satisfied translates into the requirement that the corresponding Hamiltonian is not frustrated. In an average-case analysis, a natural control parameter is the clause-to-variable ratio $\alpha = M/N$. If the value of α is sufficiently low, the problem is satisfied with probability going to 1 in the thermodynamic limit, while if it is sufficiently high, the problem is unsatisfied with probability going to 1. It turns out that, in the thermodynamic limit, the satisfiability probability undergoes a sharp *SAT-UNSAT phase transition*: the probability $p_N(\alpha)$ that a random 3-SAT instance of size N and ratio α is satisfied, in the large N limit, is given by

$$\lim_{N\to\infty} p_N(\alpha) = \theta(\alpha_c - \alpha)\,, \tag{3.1}$$

with $\alpha_c = 4.267\ldots$ In the UNSAT region, the encoding itself of a certificate of unsatisfiability requires an exponential time [27], whose rate is larger at lower values of the parameter α. The SAT region presents a second transition point: for values of α below $\alpha_d = 3.921\ldots$, the system contains an exponentially large number of solutions, and the phase space is depicted by a RS-pattern, while for values of α between α_d and α_c there is an exponential number both of solutions and of pure phases, whose rates vanish at α_c, and the pure-phase clusterization pattern is well-described by a 1-RSB ansatz.

This statistical mechanics analysis has a practical counterpart. Heuristic algorithms which only use "local" informations undergo a dynamical transition at the ratio $\alpha_d^{(\mathrm{alg})}$: beyond this value, they fail to retrieve a solution. When there are many pure phases, these algorithms always get stuck on metastable valleys, so α_d is an upper-bound for all the $\alpha_d^{(\mathrm{alg})}$, found in any algorithm of this class. On the other hand, when using Survey-Propagation-inspired algorithms, the non-local information propagated by the survey over pure phases, and by the parameter y which accounts for the reshuffling of free-energies, allows to retrieve a solution also for large sizes, and values of α near to α_c (for example, a random instance with $N = 10^6$ and $\alpha = 4.22$ has been solved in 3338.6 seconds on a 2.4 GHz Pc, cfr. [24]

4. The Random Assignment problem

In this chapter we present the classical solution of the Random Assignment Problem (or Random Matching Problem) [9, 28, 29], at zero temperature, using the Cavity Method approach in the Replica Symmetric ansatz. Some interesting thermodynamical quantities are derived, too.

Then, we solve the problem also at finite temperature, using a similar technique, and a property of Poisson Point Processes, discussed in Appendix A.

4.1 The problem

Given a $N \times N$ matrix ϵ, with positive entries, consider the cost function

$$\mathcal{H}_\epsilon(n) = \sum_{i,j} \epsilon_{ij} n_{ij} \,; \tag{4.1}$$

defined on the ensemble of feasible configurations $n \in \{0,1\}^{N^2}$, with the property that there is one entry $n_{ij} = 1$ per row and per column.

There exists a natural one-to-one mapping of the ensemble of feasible configurations into the set of permutations $\mathfrak{S}_N$. The mapping is given by

$$n \leftrightarrow \pi \in \mathfrak{S}_N : \qquad n_{ij} = 1 \iff j = \pi(i) \,.$$

A combinatorial optimization problem is finding, for a given instance ϵ, the configuration $n^{(1)}(\epsilon)$ which minimizes the cost function (4.1), and the relative cost $E_{\min}(\epsilon)$. The related satisfiability problem would be, for a given cost threshold E, finding either a configuration n such that $\mathcal{H}_\epsilon(n) \leq E$, or a certificate that such a configuration does not exist. We will discuss the complexity of these problems and some algorithmic solutions in chapter 6.

The statistical mechanic problem, in which the cost function (4.1) is interpreted as the Hamiltonian of a physical system, can be studied also at finite temperature. The partition function and the free energy for a given system ϵ are given by

$$Z(\epsilon) = \sum_n e^{-\beta \mathcal{H}_\epsilon(n)} \,; \qquad F(\epsilon) = -\frac{1}{\beta} \log Z(\epsilon) \,; \tag{4.2}$$

and the minimum cost is related to the zero-temperature limit of the free energy,

$$E_{\min}(\epsilon) = \lim_{\beta\to\infty} F_\beta(\epsilon)\,. \tag{4.3}$$

We can consider the problem both for a given instance, and on average over a certain ensemble of instances ϵ, weighted with a certain measure given *a priori*. In the following we will always use factorized measures, in which the entries ϵ_{ij} are i.i.d. quantities distributed with some normalized function $\mu(\epsilon)$ over $\mathbb{R}^+$. The situations most frequently studied in literature are

- Flat measure over a compact interval, $\mu(\epsilon) = \theta(1-\epsilon)\theta(\epsilon)$.. This measure exploits the symmetry $F_\beta(\epsilon) + F_{-\beta}(1-\epsilon) = N$, as it is symmetric under the involution $\epsilon_{ij} \to 1 - \epsilon_{ij}$.

- Exponential measure $\mu(\epsilon) = e^{-\epsilon}$. This measure is in some sense "natural" in the context of Poisson Point Processes (the spacings between i.i.d. variables with exponential distribution are still distributed exponentially). This leads to an elegant combinatorial proof of the statistical properties of Random Assignment [29], as summarized at the end of the section.

- Exponential measure with a cutoff. A sampled value $\epsilon_{ij} = \epsilon \in [0, \frac{c}{N}]$ with probability $e^{-\epsilon}\mathrm{d}\epsilon$, and $+\infty$ with probability $e^{-c/N}$. This measure can be seen also as a deformation of the exponential measure above, in which entries larger than c/N are removed. In the large N limit, this process does not change the choice of the optimal cost configuration for values of $c \gg \mathcal{O}(\log_2 N)$. Working at finite c allows us to deal with graphs of finite connectivity, thus providing a stronger mathematical control of Cavity Theory. We can perform the limit $c \to \infty$ at the end of calculations in order to obtain the results of the exponential measure problem.

- Double-delta measure $\mu(\epsilon) = \frac{c}{N}\delta(\epsilon) + \left(1 - \frac{c}{N}\right)\delta(\epsilon - 1)$. This measure leads to the combinatorial problem of dimer packing on random graphs *à la* Erdös-Renyi, of average connectivity c. This problem is discussed in [30].

The precise choice of the measure $\mu(\epsilon)$ is not important as long as we are interested to the thermodynamics of the system for sufficiently large inverse temperatures β, and in particular for the zero-temperature limit involved in the calculations of average properties of optimal cost.

Indeed, consider the extraction of N i.i.d. variables $\{x_i\}$ with a given probability distribution $\mu(x)$ as above. Then this set of variables, sorted and rescaled by a factor N, are distributed according to a Poisson Process of rate $\mu(x/N)$. So, if $\mu(\epsilon)$ is finite and regular in a sufficiently large right neighbourhood of the origin, with $\mu(0) = \mu_0$, the first items of the process are asymptotically distributed according to a Poisson Point Process of rate equal to the function $\mu_0\theta(x)$. For this reason, the first two examples above, (and *a posteriori* also the third one for $c \gg \mathcal{O}(\log N)$), are all thermodynamically equivalent.

In the context of replica symmetric approximation, it has been proven in [9] that the average value, over the disorder measure $\mu(\epsilon)$, of the minimal cost in the infinite N limit is

$$\langle E_{\min}\rangle = \frac{\pi^2}{6} = \zeta(2) = \sum_{n=1}^{\infty} \frac{1}{n^2}\,. \tag{4.4}$$

This fact has been recently proven with exact methods on the exponential measure [29], together with the finite-size conjectures

$$\langle E_{\min}\rangle_N = \sum_{n=1}^{N} \frac{1}{n^2}\,, \tag{4.5}$$

(Parisi Conjecture), and the more general Coppersmith and Sorkin Conjecture [31], for the case of rectangular cost matrices ϵ.

As we will see in the derivation, the $\zeta(2)$ limit of random assignment [28] is related to the moments of the logistic distribution

$$f(x) = \frac{1}{\left(e^{x/2} + e^{-x/2}\right)^2}\,,$$

which is indeed the distribution of the cavity magnetic fields in the zero-temperature limit.

4.2 Cavity Equations for the Assignment Problem

The Hamiltonian (4.1) is of the general form (2.1) of section 2.1, with variable indices $\{(i,j)\} \in \{1,\dots,N\}^2$, and interaction indices $\{i\}^{(\text{rows})} \cup \{j\}^{(\text{columns})} \in \{1^{(\text{row})},\dots,N^{(\text{row})}, 1^{(\text{col.})},\dots,N^{(\text{col.})}\}$. One-body terms are given by

$$e^{-\beta W_{ij}(n_{ij})} = e^{-\beta \epsilon_{ij} n_{ij}}\,, \tag{4.6}$$

while interactions correspond to

$$e^{-\beta E_i^{(\text{row})}(n_{ij})} = \delta\Big(\sum_j n_{ij}, 1\Big)\,, \qquad e^{-\beta E_j^{(\text{col.})}(n_{ij})} = \delta\Big(\sum_i n_{ij}, 1\Big)\,. \tag{4.7}$$

Specializing equation (2.11) to this problem, we obtain the Cavity Equations

$$\begin{cases} h_{(ij)\to i} = \epsilon_{ij} + u_{j\to(ij)}\,; \\ u_{j\to(ij)} = -\dfrac{1}{\beta} \ln \dfrac{\sum_{n_{ij'}} \delta\Big(1 + \sum n_{ij'}, 1\Big) e^{-\beta \sum h_{(ij')\to i} n_{ij'}}}{\sum_{n_{ij'}} \delta\Big(\sum n_{ij'}, 1\Big) e^{-\beta \sum h_{(ij')\to i} n_{ij'}}}\,; \end{cases} \tag{4.8}$$

which, solving w.r.t. cavity fields, can be written for short

$$u_{i\to(ij)} = -\frac{1}{\beta} \ln \frac{\displaystyle\sum_{n_{ij'}} \delta\Big(\sum_{j'\neq j} n_{ij'}, 0\Big) \exp\Big(-\beta \sum_{j'\neq j} (\epsilon_{ij'} + u_{j'\to(ij')}) n_{ij'}\Big)}{\displaystyle\sum_{n_{ij'}} \delta\Big(\sum_{j'\neq j} n_{ij'}, 1\Big) \exp\Big(-\beta \sum_{j'\neq j} (\epsilon_{ij'} + u_{j'\to(ij')}) n_{ij'}\Big)}\,. \tag{4.9}$$

Because of the delta constraint, the numerator reduces to a single term, since all $n_{ij'}$ must be zero, while the denominator is the sum of $N-1$ terms since only one $n_{ij'}$ must be non-zero. From now on we call for short

$$u_{i\to(ij)} \equiv g_{i\to j}\,, \qquad\qquad u_{j\to(ij)} \equiv h_{j\to i}\,; \tag{4.10}$$

equation (4.9) becomes

$$g_{i\to j} = \frac{1}{\beta}\ln\sum_{j'\neq j} e^{-\beta(\epsilon_{ij'}+h_{j'\to i})}\,, \tag{4.11a}$$

$$h_{j\to i} = \frac{1}{\beta}\ln\sum_{i'\neq i} e^{-\beta(\epsilon_{i'j}+g_{i'\to j})}\,. \tag{4.11b}$$

4.3 Exact solution of the distributional equation

In this section we search for a distributional identity involving the average cavity field distribution, for a measure over instances factorized on the entries ϵ_{ij}, which are i.i.d. with $\mu(\epsilon) = \theta(\epsilon)e^{-\epsilon}$. As we will see, the entries of an instance on a certain row in the large N limit are distributed as a Poisson Point Process, and this will allow us to solve the equation using general properties of Independent Point Processes discussed in A.

Consider equation (4.11a) for a given pair (i,j) averaged over the instances. The cavity fields $g_{i\to j}$ and $\{h_{j'\to i}\}_{j'\neq j}$ are assumed to be identically distributed with some normalized $f(x)$, which we want to find self-consistently. We will come back to this point in section 6.4. We can reorder the $N-1$ summands w.r.t. the values of $\epsilon_{ij'}$, and rescale the entries and the fields of a factor $\frac{1}{N}$. Call $\{\xi_i\}_{i=1,\dots,N-1}$ the rescaled entries, and x, $\{x_i\}_{i=1,\dots,N-1}$ respectively the rescaled fields g and h. We can write the distributional equation

$$x \stackrel{d}{=} \frac{1}{\beta}\ln\sum_{i=1}^{N-1} e^{-\beta(\xi_i+x_i)}\,, \tag{4.12}$$

with $\{\xi_i\}$ a Poisson Point Process of rate $\theta(\xi)$, as motivated in page 4.1 and x, $\{x_i\}_{i=1,\dots,N-1}$ i.i.d. with $f(x)$. We will find in the end that the differentail equation we find allows for a solution, $f(x)$, with a finite limit for $N\to\infty$, and is a fast-decreasing function. So, for values of β of order 1, or larger, the equation above allows for a simple infinite N limit, we can write

$$x \stackrel{d}{=} \frac{1}{\beta}\ln\sum_{i=1}^{\infty} e^{-\beta(\xi_i+x_i)}\,, \tag{4.13}$$

up to corrections of order $\exp(-\beta N)$, as the contribution of terms with index larger than N in the sum is exponentially small.

The zero-temperature limit of this equation provides us the relevant results for the optimization problem. For $\beta \to +\infty$, the equation becomes

$$x \stackrel{d}{=} \max_i(-\xi_i - x_i)\,. \tag{4.14}$$

The quantity on the right-hand side is of the same kind discussed in Appendix A. The points $\xi + x$ are the result of a reshuffling procedure, and are consequently distributed with $(\theta * f)(x) = F(x)$, where $F(x)$ is the primitive of $f(x)$:

$$F(x) = \int_{-\infty}^{x} \mathrm{d}x'\, f(x')\,. \tag{4.15}$$

We ask for the probability distribution of the maximum value in the set $\{-\xi_i - x_i\}$. As we deal with an independent Poisson Process, the probability that this value is *smaller* than a given value y is expressed as the probability that, for a given free gas of particles distributed with density $F(-x)$, there are exactly zero particles in the region $[y, +\infty)$, and thus is a poissonian of rate $R = \int_y^\infty \mathrm{d}y'\, F(-y')$, calculated in $n = 0$:

$$\mathrm{prob}(\max(-\xi_i - x_i) \le y) = \mathrm{Poiss}_R(0) = e^{-\Phi(-y)}\,, \tag{4.16}$$

where

$$\Phi(x) = \int_{-\infty}^{x} \mathrm{d}x'\, F(x')\,. \tag{4.17}$$

The distributional equation (4.14) implies in particular

$$\mathrm{prob}(x \le y) = \mathrm{prob}(\max_i(-\xi_i - x_i) \le y)\,,$$

and, by definition (4.15), we have

$$\mathrm{prob}(x \le y) = F(y)\,,$$

from which we obtain

$$F(x) = e^{-\Phi(-x)}\,. \tag{4.18}$$

Furthermore, differentiating equation (4.16) we find

$$\mathrm{prob}(\max(-\xi_i - x_i) = y) = -\frac{\mathrm{d}}{\mathrm{d}y}\mathrm{prob}(\max(-\xi_i - x_i) \le y) = F(-y)e^{-\Phi(-y)}\,, \tag{4.19}$$

which by the distributional equation (4.14) implies

$$f(x) = F(-x)e^{-\Phi(-x)}\,, \tag{4.20}$$

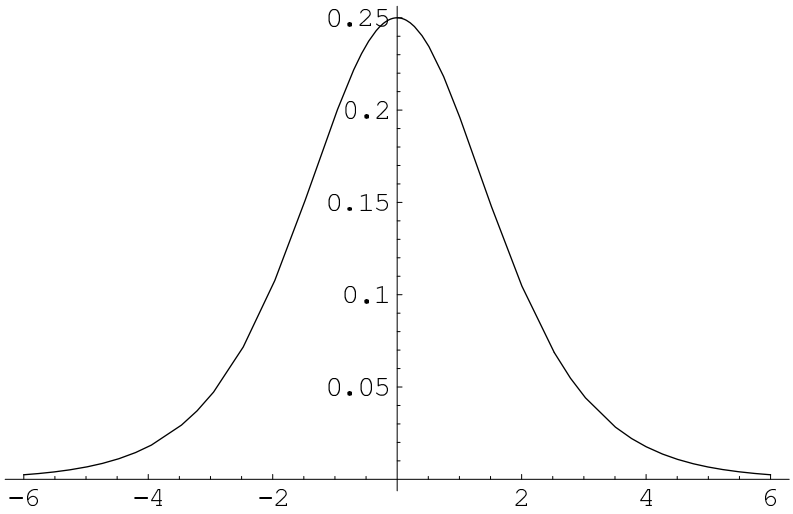


Figure 4.1 Plot of the logistic distribution $f(x) = 1/(4\cosh^2(x/2))$.

which can be also obtained directly by differentiation of (4.18). Combining (4.18) and (4.20) we get

$$f(x) = F(x)F(-x)\,, \tag{4.21}$$

from which we deduce that $f(x)$ is an even function. So, we find

$$f(x) = F(x)(1 - F(x))\,; \tag{4.22}$$

this differential equation jointly with the fact that $f(x)$ is positive and is normalized, allows us to identify the distribution of the cavity fields with the logistic distribution

$$f(x) = \frac{1}{4\cosh^2(x/2)} = \frac{e^x}{(1+e^x)^2}\,. \tag{4.23}$$

The explicit forms of its first and second primitives, with boundary condition fixed by symmetry, are

$$F(x) = \frac{e^x}{1+e^x}\,, \tag{4.24}$$

$$\Phi(x) = \ln(1+e^x)\,. \tag{4.25}$$

4.4 Statistical properties of the minimum energy configuration

From distributional equation (4.14) we can deduce more than the mere distribution of the fields (4.23). Indeed, the distribution $p(\xi)$, for the entry ξ on which the maximum of $-\xi_i - x_i$ is

realized, corresponds to the distribution of instance entries ϵ_{ij} chosen in the solution, rescaled by a factor N. From this distribution we can extract the average value of the minimum energy, $\langle E_{\min} \rangle$.

The probability $p(\xi)$ can be found by independence of Poisson Processes. Say we ask for the probability $p(\overline{\xi}, \overline{x})$ that the maximum value is realized for a given pair $(\overline{\xi}, \overline{x})$. The remaining entries (ξ_i, x_i), conditioned to the presence of the entry $(\overline{\xi}, \overline{x})$, are still an independent point process with rate $\rho(\xi, x) = \theta(\xi) f(x)$, thus the maximum of $-\xi_i - x_i$ over these remaining entries is still distributed with $f(x)$. So $p(\overline{\xi}, \overline{x}) = \text{prob}(-\overline{\xi} - \overline{x} > x)$, with x and $\overline{x}$ i.i.d. with measure $f(x)$. After integration we find

$$p(\xi) = \theta(\xi) \int \mathrm{d}\overline{x} \int \mathrm{d}x \, f(\overline{x}) f(x) \theta(-\overline{\xi} - \overline{x} - x) = \theta(\xi) \frac{e^{-\xi} - 1 + \xi}{4 \cosh^2(\xi/2)} \,. \tag{4.26}$$

The average value of the minimum energy corresponds to the first moment of this distribution (the factor $1/N$ for the rescaling of the entries cancels out with the factor N of the sum):

$$\langle E_{\min} \rangle = \int \mathrm{d}\xi \, \xi \, p(\xi) = \frac{\pi^2}{6} \,. \tag{4.27}$$

Starting from the distribution of the fields we can find another interesting quantity connected to the statistical properties of the minimum energy configuration. We want to calculate the probability that the k-th smallest entry in a row is chosen in the minimum cost configuration, in the limit where the number of entries goes to infinity. Remember that, for a fixed row, we have relabeled the entries $\{\xi_i\}$, so that they are sorted in increasing order. We seek for the probability p_k that the index realizing the maximum of $\{-\xi_i - x_i\}_i$ is equal to k, where, as usual, the entries $\{\xi_i\}$ are a Poisson Point Process of rate $\theta(\xi)$, and the set of fields x_i are distributed according to a logistic distribution (4.23).

Say that the maximum of $-\xi_i - x_i$ is realized with the values $\overline{\xi}$ and $\overline{x}$; the probability we are looking for is given by

$$\begin{aligned} p_k = \int \mathrm{d}\overline{\xi} \, \mathrm{d}\overline{x} \, p(k, \overline{\xi}, \overline{x}) \,, \qquad p(k, \overline{\xi}, \overline{x}) &= \text{prob}(\xi_k = \overline{\xi}, x_k = \overline{x}, \max_{i \neq k}(-\xi_i - x_i) < -\overline{\xi} - \overline{x}) \\ &= \theta(\overline{\xi}) f(\overline{x}) \, \text{Poiss}_A(0) \, \text{Poiss}_B(k-1) \\ &= \theta(\overline{\xi}) f(\overline{x}) \frac{B^{k-1}}{(k-1)!} e^{-A-B} \,, \end{aligned} \tag{4.28}$$

where

- $\text{Poiss}_A(0)$ implements the constraint that the point $-\overline{\xi} - \overline{x}$ is the maximum; A is the integral of the density of the two-dimensional Poisson Point Process with rate $\theta(\xi) f(x)$, over the region (ξ, x) such that $(-\xi - x > -\overline{\xi} - \overline{x})$.

- $\mathrm{Poiss}_B(k-1)$ implements the constraint that the number of entries ξ_i smaller than $\overline{\xi}$ is $k-1$; B is the integral of the density of the two-dimensional Poisson Point Process with rate $\theta(\xi)f(x)$, over the region (ξ, x) such that $(-\xi - x < -\overline{\xi} - \overline{x})$ and $\xi < \overline{\xi}$.

So, using definitions (4.15), the quantities A and B are

$$\begin{aligned} A &= \int \mathrm{d}\xi \int \mathrm{d}x\, \theta(\xi) f(x) \theta(\overline{\xi} + \overline{x} - \xi - x) \\ &= \int \mathrm{d}\xi\, \theta(\xi) F(\overline{\xi} + \overline{x} - \xi) = \Phi(\overline{\xi} + \overline{x})\,, \end{aligned} \tag{4.29a}$$

$$\begin{aligned} B &= \int \mathrm{d}\xi \int \mathrm{d}x\, \theta(\xi) f(x) \theta(\overline{\xi} - \xi) \theta(-\overline{\xi} - \overline{x} + \xi + x) \\ &= \int_0^{\xi} \mathrm{d}\xi\, F(-\overline{\xi} - \overline{x} + \xi) = \Phi(-\overline{x}) - \Phi(-\overline{\xi} - \overline{x})\,. \end{aligned} \tag{4.29b}$$

Now substitute expression (4.25). The generating function $p(\omega)$ is given by

$$p(\omega) = \sum_{k=1}^{\infty} \omega^{k-1} p_k = \int_0^{\xi} \mathrm{d}\xi \int_{-\infty}^{+\infty} \mathrm{d}x\, \frac{e^{-x}}{(1+e^{-x})^3} e^{-x-\xi} \left(\frac{1+e^{-x}}{1+e^{-x-\xi}} \right)^{\omega}\,, \tag{4.30}$$

performing the change of variables

$$u = e^{-x-\xi}\,, \qquad t = e^{-x}\,,$$

we finally find

$$p(\omega) = \int_0^{\infty} \mathrm{d}t\, \frac{1}{(1+t)^3} \int_0^{t} \mathrm{d}u\, \frac{1}{(1+u)^{\omega}} = \frac{1}{2-\omega}\,. \tag{4.31}$$

So the probability of chosing the kth smallest entry in a row for the minimum cost configuration is simply given by

$$p_k = \frac{1}{2^k}\,. \tag{4.32}$$

This result is found with different techniques in [28], while it is derived in a similar way in [32].

4.5 Finite-temperature distributional equations

We now turn our attention to equation (4.13). In a fashion similar to what we did for its zero-temperature limit (4.14), we want to translate that distributional equation into a functional identity for $f_\beta(x)$, the distribution of the quantities x and $\{x_i\}$ at inverse temperature β: this identity, together with the normalization and the boundary conditions $\lim_{x\to\pm\infty} f_\beta(x) = 0$, univocally identifies the distribution of cavity fields.

We rewrite equation (4.13) as

$$e^{\beta x} \stackrel{d}{=} \sum_i e^{-\beta(\xi_i + x_i)}\,. \tag{4.33}$$

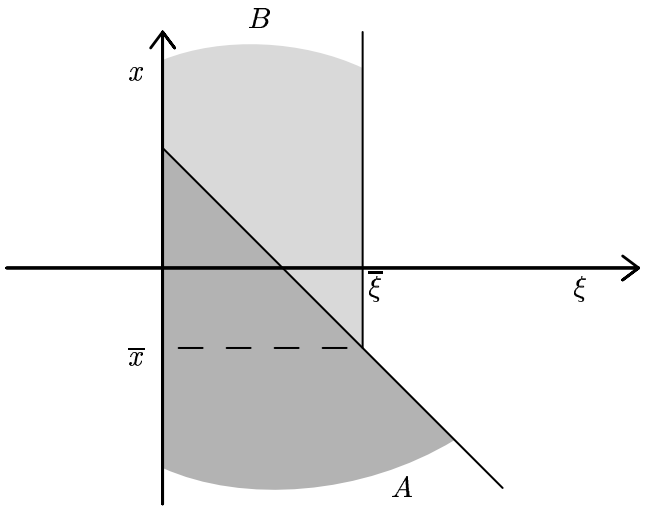


Figure 4.2 Regions of integration for the quantities A and B in (4.29).

Let's focus on the left-hand side first. If x is distributed with $f_\beta(x)$, the quantity $y = e^{\beta x}$ is distributed with

$$g(y) = f_\beta(x(y)) \left| \frac{\mathrm{d}x\,(y)}{\mathrm{d}y} \right| = \frac{1}{\beta y} f_\beta(x(y))\,. \tag{4.34}$$

Since $y > 0$, we can introduce the Laplace transform $\widetilde{g}(\eta)$ of this distribution on the domain $\mathfrak{Re}(\eta) > 0$:

$$\widetilde{g}(\eta) := \int_0^{+\infty} \mathrm{d}y\, e^{-\eta y} g(y) = \int_{-\infty}^{+\infty} \mathrm{d}x'\, f_\beta(x') \exp(-\eta e^{\beta x'})\,. \tag{4.35}$$

For the right-hand side, we define

$$y = \sum_i y_i\,, \qquad y_i = e^{-\beta(\xi_i + x_i)}\,. \tag{4.36}$$

Each y_i is sampled according to a distribution $g_\xi(y)$:

$$g_\xi(y) = \frac{1}{\beta y} f_\beta\Big(-\xi - \frac{\ln y}{\beta} \Big)\,. \tag{4.37}$$

We can find the distribution $g(y)$ of the sum $\sum_i y_i$ using the results of Appendix (A). We again consider the Laplace Transform of the distribution $g_\xi(y)$

$$\widetilde{g}_\xi(\eta) := \int_0^{+\infty} \mathrm{d}y\, \frac{e^{-\eta y}}{\beta y} f_\beta\Big(-\xi - \frac{\ln y}{\beta} \Big) = \int_{-\infty}^{+\infty} \mathrm{d}x'\, \exp(-\eta e^{\beta(\xi + x')}) f_\beta(-x')\,, \tag{4.38}$$

where we have made the change of variable

$$x' = \xi + \frac{\ln y}{\beta}\,.$$

We now apply equation (A.8), and find

$$\widetilde{g}(\eta) = \exp\left[- \int_{-\infty}^{+\infty} \mathrm{d}x' \, f_\beta(-x') \int_0^{+\infty} \mathrm{d}\xi \, \left(1 - \exp\left(- \eta e^{\beta(\xi + x')} \right) \right) \right] , \tag{4.39}$$

where we have used also the normalization of f. Equating (4.39) with (4.35) we get a differential equation allowing to determine $f_\beta(x)$. Expressions of the kind $\exp(-e^{-\beta x})$ are the only ones in which the temperature parameter explicitly appears. In the $\beta \to \infty$ limit they just produce $\theta(x)$. At finite β we recognize the primitive of a universal normalized function, the Gumbel function $t(x)$, rescaled of a factor β:

$$t(x) = \exp\left(-x - e^{-x} \right) ; \qquad t_\beta(x) = \beta \, t(\beta x) ; \tag{4.40}$$

$$\exp(-e^{-\beta x}) = T_\beta(x) := \int_{-\infty}^{x} \mathrm{d}x' \, t_\beta(x') . \tag{4.41}$$

Set

$$\eta = e^{-\beta x} . \tag{4.42}$$

Equation (4.35) becomes

$$\int_{-\infty}^{\infty} \mathrm{d}x' \, f_\beta(x') T_\beta(x - x') = (F_\beta * t_\beta)(x) , \tag{4.43}$$

while the argument in the exponent of equation (4.39) becomes

$$\begin{aligned} \int \mathrm{d}x' \int \mathrm{d}\xi \, f_\beta(-x') \, \theta(\xi) \left(1 - T_\beta(x - \xi - x') \right) &= \int \mathrm{d}y \int \mathrm{d}\xi \, f_\beta(\xi - y) \, \theta(\xi) \left(1 - T_\beta(x - y) \right) \\ &= \int \mathrm{d}y \, F_\beta(y) \left(1 - T_\beta(x - y) \right) = (\Phi_\beta * t_\beta^{(s)})(-x) , \end{aligned} \tag{4.44}$$

with $t_\beta^{(s)}(x) = t_\beta(-x)$. We finally get

$$(F_\beta * t_\beta)(x) = \exp\left[-(\Phi_\beta * t_\beta^{(s)})(-x) \right] , \tag{4.45}$$

or, deriving w.r.t. x the previous equation:

$$(f_\beta * t_\beta)(x) = (F_\beta * t_\beta)(x)(F_\beta * t_\beta^{(s)})(-x) . \tag{4.46}$$

The $\beta \to \infty$ limit of this equation is again equation (4.21): indeed

$$\lim_{\beta \to \infty} t_\beta(x) = \delta(x) ; \qquad \delta^{(s)}(x) \equiv \delta(x) ; \qquad (f * \delta)(x) = \delta(x) \quad \forall \, f(x) . \tag{4.47}$$

4.6 Low-temperature expansion of cavity field distribution

Equation (4.46) allows us to find finite-temperature corrections to the distribution of the fields, at any order in $\frac{1}{\beta}$.

First note that equation is invariant under the transformation

$$t(x) \longrightarrow t_a(x) = t(x-a)\,, \tag{4.48}$$

as deducible from the explicit form

$$\int \mathrm{d}x'\, f_\beta(x') t_\beta(x-x') = \int \mathrm{d}x'\, F_\beta(x') t_\beta(x-x') \int \mathrm{d}x'\, F_\beta(x') t_\beta(x+x')\,. \tag{4.49}$$

So we can replace $t(x)$ with its central part $t(x+\gamma_{EM})$, such that $\int \mathrm{d}x\, x\, t(x+\gamma_{EM}) = 0$.

Given an analytic function f and a fast-decreasing function t, their convolution can be expressed as

$$(f * t)(x) = \sum_k \frac{M_k}{k!}(-1)^k \frac{\mathrm{d}^k}{\mathrm{d}x^k} f(x)\,, \tag{4.50}$$

where M_k is the k-th moment of t

$$M_k = \int \mathrm{d}x\, x^k t(x)\,. \tag{4.51}$$

Indeed, from the definition of convolution

$$(f * t)(x) = \int \mathrm{d}x'\, f(x') t(x-x') = \int \mathrm{d}x' \left(\sum_k \frac{\mathrm{d}^k}{\mathrm{d}x^k} f(x) \frac{(x'-x)^k}{k!} \right) t(x-x')\,, \tag{4.52}$$

where we have expanded the function f around x.

In equation (4.46), where translation of Gumbel function in the barycenter is understood, M_k are the central moments, which can be obtained from the generating function

$$\sum_k \frac{M_k}{k!}\xi^k = \Gamma(1-\xi) e^{-\gamma_{EM}\xi} = \exp\left(\sum_{k=2}^{\infty} \frac{\zeta(k)}{k} \xi^k \right)\,. \tag{4.53}$$

Considering also the scaling (4.40) in β of $t_\beta(x)$, we get

$$(F_\beta * t_\beta)(x) = \sum_k \frac{M_k}{\beta^k k!}(-1)^k \frac{\mathrm{d}^k}{\mathrm{d}x^k} F_\beta(x)\,. \tag{4.54}$$

Besides, the (central) moments of the function $t_\beta^{(s)}(x)$ are identical to those of the Gumbel distribution, but an extra factor $(-1)^k$ appearing in the k-th moment, so that

$$(F_\beta * t_\beta^{(s)})(-x) = \sum_k \frac{M_k}{\beta^k k!} \frac{\mathrm{d}^k}{\mathrm{d}x^k} F_\beta(-x)\,. \tag{4.55}$$

We can now plunge these expansions in equation (4.46); in order to find corrections in β, we make an ansatz on the temperature dependence of the unknown field distribution primitive, $F_\beta(x)$

$$F_\beta(x) = \sum_k \frac{1}{\beta^k k!}[G_k(x) + H_k(x)]\,, \tag{4.56}$$

$$G_k(x) = G_k(-x)\,, \qquad H_k(x) = -H_k(-x)\,, \tag{4.57}$$

where we have highlighted the symmetric and antisymmetric part. Substituting in equation (4.46), and using the parity of the two parts, we get

$$\begin{aligned}
&\sum_{k=0}^{\infty} \frac{1}{\beta^k k!} \sum_{l=0}^{k} (-1)^l \binom{k}{l} M_l D^{(l+1)}[G_{k-l}(x) + H_{k-l}(x)] \\
&\qquad = \sum_{k=0}^{\infty} \frac{1}{\beta^k k!} \sum_{k_1=0}^{k} \sum_{l_1=0}^{k_1} \sum_{l_2=0}^{k_2 \equiv k-k_1} \binom{k}{k_1} \binom{k_1}{l_1} \binom{k_2}{l_2} (-1)^{l_1+l_2} M_{l_1} M_{l_2} \\
&\qquad\qquad \left[D^{(l_1)} G_{k_1-l_1}(x) D^{(l_2)} G_{k_2-l_2}(x) - D^{(l_1)} H_{k_1-l_1}(x) D^{(l_2)} H_{k_2-l_2}(x) \right]\,,
\end{aligned} \tag{4.58}$$

where we have denoted $\frac{\mathrm{d}^n}{\mathrm{d}x^n} \equiv D^{(n)}$.

To find the corrections of the k-th order in β, we must match the k-th summand of the left side with the k-th summand of the right side; in this way we find equations where β no longer appears. This system of equations is hyerarchical, since the k-th order involves only $G_0, \ldots, G_k$ and $H_0, \ldots, H_k$. That is, the k-th order correction can be found solving the equation:

$$\begin{aligned}
&\sum_{l=0}^{k} (-1)^l \binom{k}{l} M_l D^{(l+1)}[G_{k,l}(x) + H_{k,l}(x)] = \\
&\qquad \sum_{k_1=0}^{k} \binom{k}{k_1} \left(\sum_{l_1=0}^{k_1} \binom{k_1}{l_1} (-1)^{l_1} M_{l_1} [D^{(l_1)}(G_{k_1-l_1}(x) + H_{k_1-l_1}(x))] \right) \\
&\qquad\qquad \left(\sum_{l_2=0}^{k_2 \equiv k-k_1} \binom{k_2}{l_2} (-1)^{l_2} M_{l_2} [D^{(l_2)}(G_{k_2-l_2}(x) + H_{k_2-l_2}(x))] \right)
\end{aligned} \tag{4.59}$$

The first equation, for $k = 0$, is the only non-linear one and coincides with equation (4.22); it is solved by

$$G_0(x) = \frac{1}{2}\,, \qquad H_0(x) = \frac{1}{2} \tanh\left(\frac{x}{2}\right)\,. \tag{4.60}$$

For $k \geq 1$, G_k and H_k are the solutions of a linear system, where the only operators acting on the functions involved are derivatives (of order k at most) and products. Note that the closure

of 1 and $\tanh\left(\frac{x}{2}\right)$ w.r.t. these operators is contained in the linear space generated by the set of monomials

$$\left\{ \frac{1}{\left(\cosh\left(\frac{x}{2}\right)\right)^{2k}} ; \frac{\tanh\left(\frac{x}{2}\right)}{\left(\cosh\left(\frac{x}{2}\right)\right)^{2k}} \right\}_k$$

Set

$$t = \tanh\left(\frac{x}{2}\right) , \qquad y = \frac{1}{\cosh\left(\frac{x}{2}\right)^2} . \tag{4.61}$$

The following formal rules are valid for these variables:

$$\begin{cases} D(y^k) = -kty^k , \\ D(ty^k) = -ky^k + (k+\frac{1}{2})y^{k+1} , \\ t^2 \equiv 1 - y . \end{cases} \tag{4.62}$$

With this substitution, obeying the set of formal rules, no longer do we have to cope with a set of differential equations; instead, we have a linear system of polynomials in y and t, of degree at most 1 in t. Since this set is closed w.r.t. products and derivatives, as we noticed before, we

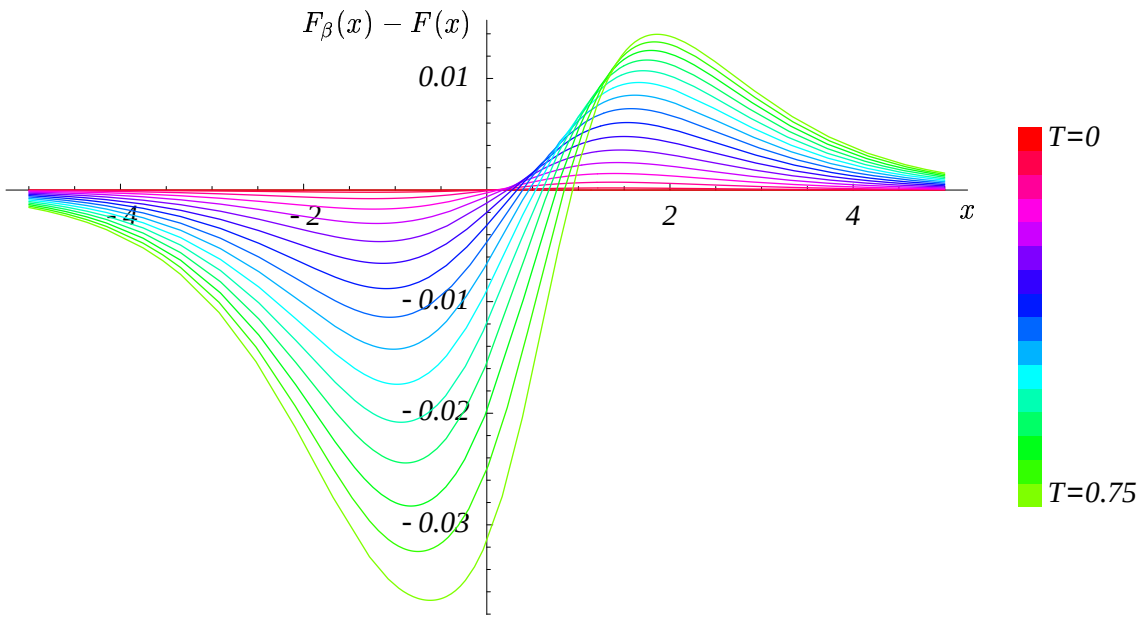


Figure 4.3 Plot of the finite-temperature corrections $F_\beta(x) - F(x)$, at temperatures $T = \frac{1}{\beta} = 0, 0.05, 0.10, \ldots, 0.75$ (from red to green). Included corrections are up to order β^{-8}.

are guaranteed that the unknown functions G_k and H_k are of the following form:

$$G_k(x) = \sum_h A_{k,h} y^h \,, \tag{4.63}$$

$$H_k(x) = \sum_h B_{k,h} t y^h \,. \tag{4.64}$$

From the zero order equation, we recognize:

$$A[0,0] = B[0,0] = \frac{1}{2}\,; \qquad A[0,h] = B[0,h] = 0 \qquad \forall\ h \geq 1\,. \tag{4.65}$$

The calculation of coefficients $A[k,h]$, $B[k,h]$ for $k \geq 1$ can be implemented using a computer program[1]. It turns out that, at all orders k, when non-null, $G_k(x) \sim \cosh^{-4}\left(\frac{x}{2}\right)$ and $H_k(x) \sim \tanh\left(\frac{x}{2}\right)\cosh^{-2}\left(\frac{x}{2}\right)$ at large x.

The explicit analytic form of the solution, up to order β^{-4}, is

$$\begin{aligned} F_\beta(x) = & \frac{1}{1+e^{-x}} \\ & + \frac{1}{\beta^2}\frac{\pi^2}{48}\cosh^{-3}\left(\frac{x}{2}\right)\sinh\left(\frac{x}{2}\right) \\ & - \frac{1}{\beta^3}\frac{\zeta(3)}{16}\cosh^{-4}\left(\frac{x}{2}\right)(1+\sinh(x)) \\ & + \frac{1}{\beta^4}\frac{\pi^4}{32\cdot 6!}\cosh^{-5}\left(\frac{x}{2}\right)\left(11\,\sinh\left(\frac{x}{2}\right) - \sinh\left(\frac{3x}{2}\right)\right) \\ & + \mathcal{O}(\beta^{-5})\,. \end{aligned} \tag{4.66}$$

[1] We used a Mathematica™ 4 program.

5. The k-assignment problem

The k-assignment problem is a generalization of the assignment problem to the case in which each person can execute k jobs simultaneously, and each job can be shared among k persons.

The statistical analysis can be performed in a fashion similar to the one of chapter 4, using the Cavity Equations in RS approximation. As we will see, the original case $k = 1$ shows some peculiarities. This is a first motivation for the present analysis.

Furthermore, the case $k = 2$, in a delicate limit of "number of colours" going to zero, can be related to the famous *Travelling Salesman Problem*, which is known to be a NP-complete problem, and, under many aspects, is considered as a prototype problem for Complexity Theory. The consequences of this fact are briefly discussed in the Conclusions.

In this chapter, instead, we focus on the thermodynamics of this problem. Following the same procedure of the previous chapter, we start specializing the general Cavity Equations (2.11), derived in chapter 2, to this case. An average over the ensemble of disorder leads to their promotion to distributional equations: a local differential equation, solved with numerical methods, is derived. We also explain how to calculate the average value of the minimum cost, and we find analytically the large k behaviour of the fields, solving the distributional equation.

5.1 The problem

Consider the ensemble of instances $\{\epsilon\}$, where each instance is a $N \times N$ grid of i.i.d. random variables ϵ_{ij}, $i = 1, ..., N$, sampled according to some normalized distribution $\rho(\epsilon)$. To fix the ideas, let us suppose that $\rho(\epsilon)$ has support on $\mathbb{R}_+$, and is finite and continuous in a right-neighbouring of the origin.

We will show that the statistical properties of the minimization problem (corresponding to a zero-temperature limit of the related statistical-mechanics problem), in the thermodynamic limit only depend on the value of the distribution at the origin, $\rho_0 \equiv \rho(0)$.

The set of allowed configurations is the set of N^2 variables $n_{ij} \in \{0, 1\}$, such that for each $i \leq N$ we have $\sum_j n_{ij} = k$, and for each $j \leq N$ we have $\sum_i n_{ij} = k$.

The cost function (Hamiltonian) is

$$\mathcal{H}_\epsilon(n) = \sum_{i,j} \epsilon_{ij} n_{ij} \,. \tag{5.1}$$

5.2 Cavity equations for the k-assignment

The argument exposed in section 4.2 to find cavity equations for the assignment problem can be easily generalized to the k-assignment problem. Variables and interactions are labeled in the very same way as the Assignment problem: variable indices are $\{(i,j)\} \in \{1,\dots,N\}^2$ and interaction indices are $\{i\}^{(\text{rows})} \cup \{j\}^{(\text{columns})} \in \{1^{(\text{row})},\dots,N^{(\text{row})},1^{(\text{col.})},\dots,N^{(\text{col.})}\}$. One-body terms are again given by

$$e^{-\beta W_{ij}(n_{ij})} = e^{-\beta \epsilon_{ij} n_{ij}} \,, \tag{5.2}$$

the only difference being in the constraints, which correspond to the interactions:

$$e^{-\beta E_i^{(\text{row})}(\{n_{ij}\})} = \delta\Big(\sum_j n_{ij}, k\Big) \,, \qquad e^{-\beta E_j^{(\text{col.})}(\{n_{ij}\})} = \delta\Big(\sum_i n_{ij}, k\Big) \,. \tag{5.3}$$

Specializing equations (2.11) to this problem, we obtain the Cavity Equations

$$\begin{cases} h_{(ij)\to j} = \epsilon_{ij} + u_{i\to(ij)} \,; \\ u_{i\to(ij)} = -\dfrac{1}{\beta} \ln \dfrac{\sum_{n_{ij'}} \delta\Big(1 + \sum_{j'} n_{ij'}, k\Big) e^{-\beta \sum_{j'} h_{(ij')\to i} n_{ij'}}}{\sum_{n_{ij'}} \delta\Big(\sum_{j'} n_{ij'}, k\Big) e^{-\beta \sum_{j'} h_{(ij')\to i} n_{ij'}}} \,; \end{cases} \tag{5.4}$$

It is again convenient to solve w.r.t. cavity fields, getting

$$u_{i\to(ij)} = -\frac{1}{\beta} \ln \frac{\sum_{n_{ij'}} \delta\Big(\sum_{j'\neq j} n_{ij'}, k-1\Big) \exp\Big(-\beta \sum_{j'\neq j} (\epsilon_{ij'} + u_{j'\to(ij')}) n_{ij'}\Big)}{\sum_{n_{ij'}} \delta\Big(\sum_{j'\neq j} n_{ij'}, k\Big) \exp\Big(-\beta \sum_{j'\neq j} (\epsilon_{ij'} + u_{j'\to(ij')}) n_{ij'}\Big)} \,. \tag{5.5}$$

In this case, the numerator involves $\binom{N-1}{k-1}$ terms, while the denominator includes $\binom{N-1}{k}$ terms. In analogy with the Assignment Problem, we turn to the notation:

$$u_{i\to(ij)} \equiv g_{i\to j} \,, \qquad u_{j\to(ij)} \equiv h_{j\to i} \,; \tag{5.6}$$

equation (5.5) becomes

$$g_{i\to j} = \frac{1}{\beta} \ln \frac{\sum_{n_{ij'}} \delta\Big(\sum_{j'} n_{ij'}, k-1\Big) e^{-\beta \sum_{j'} (\epsilon_{ij'} + h_{j'\to i}) n_{ij'}}}{\sum_{n_{ij'}} \delta\Big(\sum_{j'} n_{ij'}, k\Big) e^{-\beta \sum_{j'} (\epsilon_{ij'} + h_{j'\to i}) n_{ij'}}} \,, \tag{5.7a}$$

$$h_{j\to i} = \frac{1}{\beta} \ln \frac{\sum_{n_{ij'}} \delta\Big(\sum_{i'} n_{i'j}, k-1\Big) e^{-\beta \sum_{i'} (\epsilon_{i'j} + g_{i'\to j}) n_{i'j})}}{\sum_{n_{ij'}} \delta\Big(\sum_{i'} n_{i'j}, k\Big) e^{-\beta \sum_{i'} (\epsilon_{i'j} + g_{i'\to j}) n_{i'j}}} \,. \tag{5.7b}$$

In the zero temperature limit, they have a much simpler form. On equation (5.7a), the leading term in the denominator is the product of the k largest values $\exp(-\beta(\epsilon_{ij'} + h_{j'\to i}))$, with $j' \neq j$, while the leading term in the numerator is the product of the $k-1$ largest values in the same set. Corrections due to subleading terms are of order $\mathcal{O}(e^{-\beta\Delta x_k}, e^{-\beta\Delta x_{k+1}})$, with Δx_k equal to the difference between the k-th and the $(k-1)$-th smallest $x_{j'} = \epsilon_{ij'} + h_{j'\to i}$. If we neglect them, the first $k-1$ factors $\exp(-\beta(\epsilon_{ij'} + h_{j'\to i}))$ simplify, and we are left with the k-th in the denominator. If we denote with $k\text{-th}_i(x_i)$ the k-th largest element in a set $\{x_i\}$, we obtain the zero-temperature equations

$$g_{i\to j} = \underset{j'\neq j}{k\text{-th}}(-\epsilon_{ij'} - h_{j'\to i})\,, \tag{5.8a}$$

$$h_{j\to i} = \underset{i'\neq i}{k\text{-th}}(-\epsilon_{i'j} - g_{i'\to j})\,. \tag{5.8b}$$

5.3 Distribution of the fields

Following the procedure described in section 4.3, we require that the fields h_j and g_i are distributed with the same (normalized) distribution $f_k(x)$, satisfying the distributional equation

$$x \overset{d}{=} \underset{i\geq 1}{k\text{-th}}(-\xi_i - x_i)\,, \tag{5.9}$$

where the sequence $\{0 = \xi_0, \xi_1, \xi_2, \ldots\}$ is a rate 1 Poisson point process, i.e. the quantities $\xi_i - \xi_{i-1}$ for $i \geq 1$ are independent identically distributed as $\theta(\xi)\mathrm{e}^{-\xi}$, and the quantities x, x_i are i.i.d. as $f_k(x)$. The pairs (ξ_i, x_i) form a two-dimensional point process with density $\rho_k(x,\xi) = f_k(x)\theta(\xi)$, and we search for the k-th item with respect to the gradient function $-\xi - x$. The probability that $k\text{-th}_i(-\xi_i - x_i) \in [y, y+dy]$ is the product of two factors:

1. the differential probability that there is a point in the slice $[y, y+dy]$, given by

$$\int_{-\xi-x\geq y} \mathrm{d}\xi\,\mathrm{d}x\,\rho_k(x,\xi) - \int_{-\xi-x\geq y+dy} \mathrm{d}\xi\,\mathrm{d}x\,\rho_k(x,\xi)\,;$$

2. the probability that the number of points such that $(-\xi - x) \geq y$ is exactly $k-1$. As we have an independent point process, the number of points inside a certain region is Poissonianly distributed, with average given by the integral of the density inside the region. In our specific case, we have

$$\langle n\rangle = \int_{-\xi-x\geq y} \mathrm{d}\xi\,\mathrm{d}x\,\rho_k(x,\xi)\,.$$

Define the integral:

$$\mathcal{I}_k(y) := \int_{-\xi-x\geq y} \mathrm{d}\xi\,\mathrm{d}x\,f_k(x)\theta(\xi)\,, \tag{5.10}$$

then the distributional equation (5.9) reads

$$f_k(y)\mathrm{d}y \;=\; -\mathrm{d}y\,\left(\frac{\mathrm{d}}{\mathrm{d}y}\mathcal{I}_k(y)\right)\,\frac{\mathcal{I}_k(y)^{k-1}}{(k-1)!}\,\exp(-\mathcal{I}_k(y))\,. \tag{5.11}$$

Note that, by definition, $\lim_{y\to-\infty}\mathcal{I}_k(y) = 0^+$ as the support is empty in the limit.

We expect $f_k(x)$ to be a fast-decreasing function. We introduce its primitive

$$F_k(y) := \int_{-\infty}^{y} dy' f_k(y') \tag{5.12}$$

for which we must have

$$\lim_{y\to-\infty} F_k(y) = 0^+\,; \qquad \lim_{y\to+\infty} F_k(y) = 1^-\,;$$

Apply the operator $\int_y^{+\infty}$ on both sides of equation (5.11). On the right side, integrating by parts and using the asymptotic limit of $\mathcal{I}_k(y)$:

$$\begin{aligned} 1 - F_k(y) &= \int_y^{+\infty} \mathrm{d}y'\, f_k(y') = -\int_y^{+\infty} \mathrm{d}y'\,\left(\frac{\mathrm{d}}{\mathrm{d}y'}\mathcal{I}_k(y')\right)\,\frac{\mathcal{I}_k(y')^{k-1}}{(k-1)!}\,\exp(-\mathcal{I}_k(y')) \\ &= \left(\exp(-\mathcal{I}_k(y))\sum_{h=0}^{k-1}\frac{\mathcal{I}_k(y)^h}{h!}\right)\Bigg|_{y_0}^{\infty} = 1 - \left(\sum_{h=0}^{k-1}\frac{\mathcal{I}_k(y_0)^h}{h!}\right)\exp(-\mathcal{I}_k(y_0)) \end{aligned} \tag{5.13}$$

Simplifying, we have:

$$F_k(y) = \exp(-\mathcal{I}_k(y))\sum_{h=0}^{k-1}\frac{\mathcal{I}_k(y)^h}{h!} \tag{5.14}$$

The function $\mathcal{I}_k(y)$ itself is related to $F_k(y)$. Indeed, from the definition (5.10) we have:

$$\begin{aligned} \mathcal{I}_k(y) &= \int_{x\le-\xi-y} \mathrm{d}\xi\,\mathrm{d}x\, f_k(x)\theta(\xi) = \int_{-\infty}^{+\infty}\mathrm{d}\xi\,\theta(\xi)F_k(-\xi-y) \\ &= \int_0^{+\infty}\mathrm{d}\xi\,F_k(-\xi-y) = \int_{-\infty}^{-y}\mathrm{d}\xi\,F_k(\xi)\,. \end{aligned} \tag{5.15}$$

We define also the primitive of $F_k(x)$:

$$\Phi_k(x) = \int_{-\infty}^{x}\mathrm{d}y\,F_k(y) \equiv \mathcal{I}_k(-x)\,; \tag{5.16}$$

with the asymptotic behaviour

$$\lim_{x\to-\infty}\Phi_k(x) = 0^+\,, \qquad \lim_{x\to+\infty}\frac{\Phi_k(x)}{x} = 1\,.$$

We recall the definition of Euler Gamma function $\Gamma(k,z)$ and of Regularized Gamma function $\mathcal{Q}(k,z)$

$$\Gamma(k,z) = \int_z^\infty \mathrm{d}t\, t^{k-1}e^{-t} = \Gamma(k)e^{-z}\sum_{h=0}^{k-1}\frac{z^h}{h!} \qquad k \in \mathbb{N}^+ ; \tag{5.17}$$

$$\mathcal{Q}(k,z) = \frac{\Gamma(k,z)}{\Gamma(k)} . \tag{5.18}$$

Using definition (5.16), equation (5.14) can be written as

$$\Phi_k'(-x) = \frac{\Gamma(k,\Phi_k(x))}{\Gamma(k)} = \mathcal{Q}(k,\Phi_k(x)) . \tag{5.19}$$

In the $k=1$ case, an important role is played by the symmetric combination

$$F(-x)F(x) = f(x) .$$

Also for this general case we can write a symmetric combination:

$$\psi_k(x) = \frac{\Phi_k'(-x)\Phi_k'(x)}{k} = \frac{1}{k}\Phi_k'(x)\mathcal{Q}(k,\Phi_k(x)) \tag{5.20}$$

The function $\psi_k(x)$ turns out to be normalized. Indeed, we have for its right primitive:

$$\begin{aligned}\Psi_k(x) &:= \int_x^\infty \mathrm{d}x'\,\psi_k(x') = \int_x^\infty \mathrm{d}x'\,\Phi_k'(x')\frac{\mathcal{Q}(k,\Phi_k(x'))}{k} \\ &= \mathcal{Q}(k+1,\Phi_k(x)) - \frac{\Phi_k(x)}{k}\mathcal{Q}(k,\Phi_k(x)) ;\end{aligned} \tag{5.21}$$

and, since $\Phi_k(x) \to 0$ as $x \to -\infty$,

$$\lim_{x\to-\infty}\Psi_k(x) = 1 . \tag{5.22}$$

Note that for this function the following identity holds

$$\Psi_k(x) + \Psi_k(-x) = 1 . \tag{5.23}$$

From the symmetry of the function $\psi_k(x)$ we derive a relation for the numerical value of $\Phi_k(0)$:

$$\Psi_k(0) = \mathcal{Q}(k+1,\Phi_k(0)) - \frac{\Phi_k(0)}{k}\mathcal{Q}(k,\Phi_k(0)) = \frac{1}{2} . \tag{5.24}$$

This initial condition, together with the linear differential equation (5.19), allows us to find numerically the distribution function for the cavity fields. A plot of $f_k(x)$ for the first values of k is shown in figure 5.3.

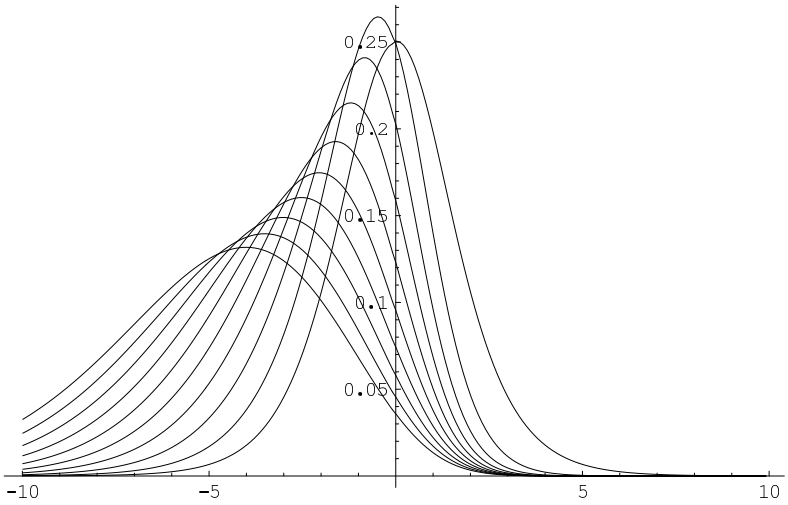


Figure 5.1 Plot of the field distributions $f_k(x)$, for $k = 1, \ldots, 10$. The central symmetric function is the logistic distribution $f_1(x)$ of equation (4.23), already shown in figure 4.3.

5.4 Calculation of the expectation value for the minimum energy

As shown in Aldous [28] for the case of Random Assignment, the average value of the minimum energy can be calculated from the distribution of the cavity fields, $f(x)$. Following the reasoning which led us to the distributional equation (5.9), we understand that the entry chosen for the minimum energy configuration is the one realizing the maximum. So, if in row i we have that $\max_j(-\epsilon_{ij} - h_j)$ is realized on the index $\bar{\jmath}$, the energy contribution of row i is given by $\epsilon_{i\bar{\jmath}}$.

The probability $\mu(\xi)$ that, for a certain row, we have a summand ξ as energy contribution, is the product of two events:

- the probability that a point $\xi_{\bar{\imath}} \in [\xi, \xi + \mathrm{d}\xi]$ is sampled in our process, that is

$$\theta(\xi)\mathrm{d}\xi\,;$$

- the probability, conditioned to the event above, that all the other points in the process have gradient function $-\xi_i - x_i$ smaller than $-\xi_{\bar{\imath}} - x_{\bar{\imath}}$, where $x_{\bar{\imath}}$ is sampled according to $f(x)$:[1]

$$\mathrm{prob}(-\xi - x) \leq \max_i(-\xi_i - x_i))\,.$$

[1]Note that in an independent point process, the process marginalized to the presence of some points, in which these points are removed, coincides with the original process.

Since the right hand side of the inequality is distributionally identical to $f(x)$, we can replace it with a single random variable x', distributed with $f(x')$. So we have

$$\mu(\xi) = \theta(\xi)\text{prob}(-\xi \leq x + x') \,. \tag{5.25}$$

In the general k-assignment problem, the distributional equation for the fields (5.9) is

$$f_k(x) \stackrel{d}{=} \underset{i\geq 1}{k\text{-th}}(-\xi_i - x_i) \,,$$

and the k entries chosen for the minimum configuration are the ones corresponding to the k largest values of the gradient function $-\xi_i - x_i$. So, if in row i the first k values of $-\epsilon_{ij} - h_j$ are obtained for $\{j_1^{(i)}, \ldots, j_k^{(i)}\} \subseteq \{1, \ldots, N\}$, the energy contribution for row i is given by $\epsilon_{ij_1^{(i)}} + \ldots, + \epsilon_{ij_k^{(i)}}$.

Following a reasoning similar to the one above, for a given row, one of the k energy contributions is inside the interval $[\xi, \xi + \mathrm{d}\xi]$ with probability $\mu_k(\xi)$ given by

$$\mu_k(\xi) = \theta(\xi)\,\text{prob}(-\xi - x \leq \underset{i}{k\text{-th}}(-\xi_i - x_i)) \,. \tag{5.26}$$

As we have k contributions, the measure $\mu_k(\xi)$ is normalized to k. Since the right hand side of the inequality is distributionally identical to $f_k(x)$, we can replace it with a single random variable x', distributed with $f_k(x')$. So we have

$$\mu_k(\xi) = \theta(\xi)\,\text{prob}(-\xi \leq x + x') \,. \tag{5.27}$$

The average energy contribution per row is given by the expectation value of ξ over the measure $\mu_k(\xi)$, that is

$$\begin{aligned}
\langle E_{\min} \rangle_k &= \int_0^{+\infty} \mathrm{d}\xi\, \xi\, \text{prob}(-\xi \leq x + x') = \int_0^{+\infty} \mathrm{d}\xi\, \xi \int_{-\infty}^{+\infty} \mathrm{d}x \int_{-\infty}^{-\xi - x} \mathrm{d}x'\, f_k(x) f_k(x') \\
&= \int_0^{+\infty} \mathrm{d}\xi\, \xi \int_{-\infty}^{+\infty} \mathrm{d}x\, f_k(x) F_k(-\xi - x) = \int_{-\infty}^{+\infty} \mathrm{d}x\, F_k(x) \int_0^{+\infty} \mathrm{d}\xi\, \xi f_k(-\xi - x)
\end{aligned}$$

and, integrating by parts in ξ,

$$\begin{aligned}
&\int_{-\infty}^{+\infty} \mathrm{d}x\, F_k(x) \int_0^{+\infty} \mathrm{d}\xi\, F_k(-\xi - x) = \int_{-\infty}^{+\infty} \mathrm{d}x\, \Phi_k'(x) \Phi_k(-x) \\
&= \int_{-\infty}^{+\infty} \mathrm{d}x\, \Phi_k(x) \mathcal{Q}(k, \Phi_k(x)) \,.
\end{aligned}$$

The expectation value of the free energy is thus given by

$$\langle E_{\min} \rangle_k = \int_{-\infty}^{+\infty} \mathrm{d}x\, \Phi_k(x) \mathcal{Q}(k, \Phi_k(x)) \,. \tag{5.28}$$

5.5 Asymptotic behaviour for large k

The distributional equation (5.19) can be easily solved in the limit $k \to \infty$ using the asymptotic properties of the Poisson distribution. The density of the points $y_i = \xi_i + x_i$, given by the reshuffling argument of A, is $\rho(y) = F_k(y)$. Suppose that the width of $f_k(x)$ is $o(k)$, then in the relevant region where $\int_{-\infty}^{y} \mathrm{d}\rho(y') \sim k$, the density is approximatively 1. So, the k-th point is distributed approximatively as a Gaussian of variance k, centered in the point y such that $\int_{-\infty}^{y} \mathrm{d}\rho(y') = k$, that is $k + \langle x \rangle$ in our hypothesis. From the distributional equation we deduce that $f(x)$ itself should be a Gaussian of variance k, centered on some point $\langle x \rangle$. The width of $f(x)$, being of order $\sqrt{k}$, is consistent with the previous ansatz. The average value is determined by

$$\langle x \rangle = \left\langle \underset{i}{k\text{-th}}(-y_i) \right\rangle = -k - \langle x \rangle \; ; \qquad \langle x \rangle = -\frac{k}{2} \, .$$

The first non-trivial correction can be calculated by a careful analysis of equation (5.19). First we note that this equation admits a scaling limit for $k \to \infty$. At order zero we find

$$\lim_{k \to \infty} \mathcal{Q}(k, kx) = \theta(1 - x) \, , \tag{5.29}$$

by saddle-point reasonings[2], from which we deduce that the proper rescaling of $\Phi_k(x)$ is of a factor k. This factor is balanced on the left side of (5.19) if we also scale the argument x of a factor k, because of the Jacobian in the derivative. Thus, defining a rescaled function

$$\widetilde{\Phi}_k(x) = \frac{1}{k} \Phi_k(kx) \, ; \qquad \widetilde{\Phi}(x) = \lim_{k \to \infty} \widetilde{\Phi}_k(x) \, ,$$

at zero order equation (5.19) becomes

$$\widetilde{\Phi}'(-x) = \theta(1 - \widetilde{\Phi}(x)) \, . \tag{5.30}$$

Using the boundary condition $\lim_{x \to -\infty} \widetilde{\Phi}(x) = 0$ we find the solution

$$\widetilde{\Phi}(x) = \begin{cases} 0 & x \leq -1/2 \, ; \\ x + 1/2 & x > -1/2 \, . \end{cases}$$

From $\widetilde{\Phi}''(x) = \delta(x + 1/2)$ we recover the fact that $\langle x \rangle = -k/2$. We expect corrections to this solution in a neighbourhood of the singularity of size $\mathcal{O}(1/\sqrt{k})$. Indeed, the leading correction to (5.29) is given by[3]

[2] Going back to definition (5.17),

$$\mathcal{Q}(k, kx) = \int_{kx}^{+\infty} \mathrm{d}t \, t^{k-1} e^{-t} \propto \int_{x}^{+\infty} \frac{\mathrm{d}t}{t} e^{-k(t - \ln t)} \, ,$$

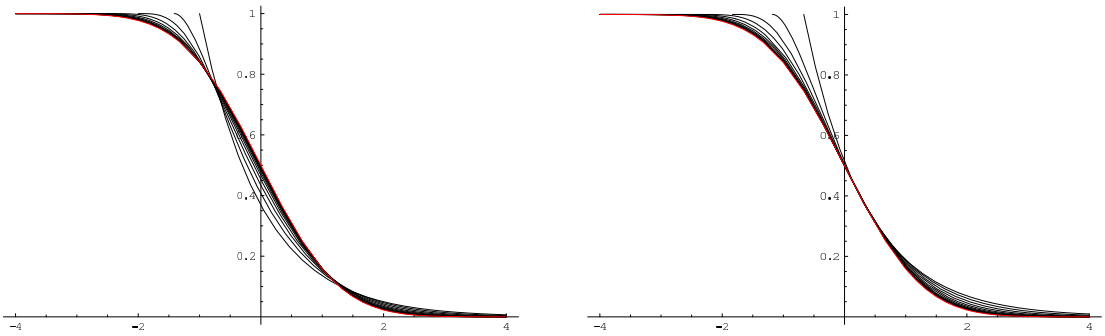


Figure 5.2 On the left, plot of the rescaled regularized gamma functions $\mathcal{Q}(k, k+\sqrt{k}x)$ in (5.31), for $k = 2^n$, and $n = 0, \ldots, 10$. On the right, plot of $\mathcal{Q}(k, k+\sqrt{k}x+c)$, with the translation coefficient $c = -1/3$ which enhances the asymmetry suppression, as described in equation (5.33). Comparison with the asymptotic curve, plotted in red (gray for B&W prints) is shown.

$$\lim_{k\to\infty} \mathcal{Q}(k, k+\sqrt{k}x) = \frac{1}{\sqrt{2\pi}} \int_x^{+\infty} \mathrm{d}t\, e^{-\frac{t^2}{2}} + \mathcal{O}(k^{-1/2})\,, \tag{5.31}$$

Defining

$$\widetilde{\widetilde{\Phi}}_k(x) = \frac{1}{\sqrt{k}} \widetilde{\Phi}_k(\sqrt{k}x - 1/2)\,; \qquad \widetilde{\widetilde{\Phi}}(x) = \lim_{k\to+\infty} \widetilde{\widetilde{\Phi}}_k(x)\,,$$

and using the fact that $\widetilde{\Phi}(x) = x + 1/2 + \mathcal{O}(e^{-k})$ in a neighbourhood of $x = 1/2$, we can write

$$\widetilde{\widetilde{\Phi}}'(x) = \frac{1}{\sqrt{2\pi}} \int_{-x}^{+\infty} \mathrm{d}t\, e^{-\frac{t^2}{2}} + \mathcal{O}(1/\sqrt{k})\,. \tag{5.32}$$

From the original definition $f(x) = F'(x) = \Phi''(x)$, which applies also to $\widetilde{\widetilde{\Phi}}(x)$ up to a scale factor, we obtain that $f(x)$ at leading order is a Gaussian, as expected from the previous reasoning.

Keeping the next correction in the expansion of $\mathcal{Q}(k, x)$, which describes the asymmetry of this function, we find

$$\lim_{k\to\infty} \frac{1}{\sqrt{k}} \Big(1 - \sum_{\pm} \mathcal{Q}(k, k \pm \sqrt{k}x + c)\Big) = \sqrt{\frac{2}{\pi}} \left(c + \frac{1-x^2}{3}\right) e^{-\frac{x^2}{2}}\,. \tag{5.33}$$

---

and the saddle-point action $S(t) = t - \ln t$ has a minimum in $t = 1$.

[3] Expanding around the saddle point

$$\int_{x\sqrt{k}}^{+\infty} \frac{\mathrm{d}t}{t} e^{-k(t-\ln t)} \propto \int_x^{+\infty} \mathrm{d}t\, e^{-\frac{t^2}{2} + \mathcal{O}(k^{-1})}\,.$$

So the next order in equation (5.32) is given by

$$\tilde{\tilde{\Phi}}'(x) = \frac{1}{\sqrt{2\pi}} \int_{-x}^{+\infty} \mathrm{d}t\, e^{-\frac{t^2}{2}} + \sqrt{\frac{2}{k\pi}} \left(\frac{1-x^2}{3} \right) e^{-\frac{x^2}{2}} ; \tag{5.34}$$

$$\tilde{\tilde{f}}(x) = \tilde{\tilde{\Phi}}''(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \left(1 + \frac{2}{\sqrt{k}} \left(-x + \frac{x^3}{3} \right) + \mathcal{O}\left(\frac{1}{k} \right) \right) . \tag{5.35}$$

6. Survey of Algorithms for Random Assignment Problems

In this chapter we describe some algorithms to solve Assignment Problems. We start with a short description of a general method for solving Linear Programming problems, the Simplex method; the linear programming formulation of the Assignment and the k-assignment Problems are discussed subsequently. Another exact algorithm for the Assignment Problem, the Hungarian Method, is presented. Also an algorithmic implementation of Cavity Equations is described, as well as several problems we have faced in this task, which have highlighted some peculiarities of the Assignment, and stressed the validity extent of the approximations done in the theoretical treatment. The final section deals with the satisfiability version of the Assignment Problem.

6.1 The Simplex method of Linear Programming

A Linear Programming problem [33] is any problem that can be stated so that the desired solution is a set of values $\{x_1, \ldots, x_n\}$ such that:

- all $x_j \geq 0$ (*non-negativity condition*);
- they are a solution of the system in n unknowns, composed of m *constraints*

$$a_{11}x_1 + \ldots + a_{1n}x_n = b_1\,, \tag{6.1a}$$
$$a_{21}x_1 + \ldots + a_{2n}x_n = b_2\,, \tag{6.1b}$$
$$\vdots$$
$$a_{m1}x_1 + \ldots + a_{mn}x_n = b_m\,; \tag{6.1c}$$

- they minimize the *objective function*

$$\mathcal{H}_c(x) = c_1x_1 + \ldots + c_nx_n = z\,. \tag{6.2}$$

Any configuration satisfying the first two properties but not the third, is called a *feasible solution*. A *basic solution* is a feasible solution in which $n - m$ unknowns are set to zero. The set of all

feasible solutions is a convex $(n-m)$-dimensional polygon in $\mathbb{R}^n$, and its vertices are the basic solutions[1]. As a consequence, the optimal solution, if it exists and is unique, is a vertex of the convex. If it is not unique, it is the simplex convex combination of a subset of vertices, all optimal solutions.

Linear programming problems admits a certain number of algorithms of solution, but almost all of them use the basic principles already present in the simplest algorithmic approach, named *simplex method*, which we shortly describe in this section.

The simplex method is an iterative procedure which consists of moving from an extreme point of the convex to an adjacent one[2], following the steepest gradient of the cost function.

Assume that all the equations (6.1) are linearly independent (or restrict to an independent subset). Suppose we have found a feasible basic solution (this is potentially a delicate point), and $x_1, \ldots, x_m$ are the basic variables. With a set of linear transformations we can rearrange (6.1), in the so-called *canonical form*

$$\begin{array}{rcccccl} x_1 & & & + a'_{1,m+1}x_{m+1} & + \ldots & + a'_{1,n}x_n & = b'_1\,, \\ & x_2 & & + a'_{2,m+1}x_{m+1} & + \ldots & + a'_{2,n}x_n & = b'_2\,, \\ & & \ddots & & & & \\ & & x_m & + a'_{m,m+1}x_{m+1} & + \ldots & + a'_{mn}x_n & = b'_m\,; \\ & & & & & & \\ & & & c'_{m+1}x_{m+1} & + \ldots & + c'_n x_n & = z - z_0\,. \end{array} \tag{6.3}$$

In the last line z_0 is the value of the cost function in the basis under consideration.

If the smallest coefficient c'_s is positive or equal to zero, the basic solution we are considering is the optimal solution, because all extreme adjacent points correspond to higher values of the cost function. Indeed, suppose we move to another vertex; there will be some x_l, whose value was set to zero in the old basic solution, now assuming a positive value, and the cost funtion value in the new basis will be equal to $z = z_0 + c'_l x_l$. So, say that the minimum c'_s is negative; in the column s look for the positive coefficients a'_{is} (if they are all negative or null, this means that the cost function is unbounded -and this is advised not to happen). It is possible to see from the canonical form that the non-negativity constraint implies that for all i such that $a_{is} > 0$, x_s must be smaller than b_i/a_{is}. Choose the row r realizing the smallest value of these ratios. Using element a'_{rs} as a pivot, we move with another set of linear tranformations to the adjacent basis where x_r takes the place of x_s. This new basis has been chosen in order to decrease

[1] It is convex as, given two feasible solutions $\vec{x}^{(0)}$ and $\vec{x}^{(1)}$, the convex combination $\vec{x}^{(t)} = t\vec{x}^{(0)} + (1-t)\vec{x}^{(1)}$ still satisfy both non-negativity condition and the linear system (6.1); it is a polygon since it is the intersection of a $(n-m)$-dimensional vector space and n semi-spaces in $\mathbb{R}^n$, and the basis solutions correspond to the points satisfying tightly at least $n-m$ disequalities, so they are the vertices of the polygon.

[2] Two extreme points are adiacents if their basis differ of only a pair of variables.

the cost function as much as possible (we have taken the smallest coefficient c'_s) while satisfying non-negativity constraint. Performing another linear transformation, we can write the canonical form in the new basic feasible solution and restart the whole procedure, eventually halting when all coefficients in the cost function are negative.

6.2 Linear Programming formulation for Assignment and k-Assignment Problem

It is known that the assignment problem has a formulation in terms of an integer linear programming instance. There are:

- N^2 variables
$$n_{ij} \geq 0\,;$$
- $2N-1$ independent equations
$$\begin{cases} \sum_i n_{ij} = 1 & \forall j \leq N-1\,; \\ \sum_j n_{ij} = 1 & \forall i \leq N\,; \end{cases}$$
- the cost function $\mathcal{H}_\epsilon(n)$.

The constraints $n_{ij} \leq 1$ are not needed, as they are implicit in the conditions above.

This fact doesn't hold any longer in the generic k case, for which a linear programming formulation requires a larger instance. Introduce the N^2 slack variables $\overline{n}_{ij}$, complementary to n_{ij}. The desired instance consists of:

- $2N^2$ variables
$$n_{ij} \geq 0\,; \qquad \overline{n}_{ij} \geq 0\,;$$
- $N^2 + 2N - 1$ independent equations
$$\begin{cases} \sum_i n_{ij} = k & \forall j \leq N-1\,; \\ \sum_j n_{ij} = k & \forall i \leq N\,; \\ n_{ij} + \overline{n}_{ij} = 1 & \forall i,j \leq N\,; \end{cases}$$
- the cost function $\mathcal{H}_\epsilon(n)$.

This is a first hint of a different behaviour of the $k = 1$ case with respect to the $k > 1$ case.

6.3 The Hungarian Algorithm

In this section we describe in detail a performing algorithm for solving in polynomial time the Assignment Problem on each instance (up to a zero measure set) of the ensemble of random matrices (4.1). The algorithm described is a modification of the so-called *Hungarian algorithm* [34], which we have chosen because of its conceptual simplicity, its interesting algebraic aspects, and its resemblance with some features of the cavity method approach.

First we note the existence of a *gauge invariance* in the definition of the cost function. Generalize the family of cost functions (4.1) introducing an arbitrary cost shift

$$\mathcal{H}_{\epsilon,h_0}(n) = \sum_{ij} \epsilon_{ij} n_{ij} + h_0 \,. \tag{6.4}$$

If all $\epsilon_{ij} \geq 0$, the quantity h_0 is a lower bound for the minimum cost $E_{\min}(\epsilon, h_0)$. Furthermore, the cost function (6.4) is invariant under the linear transformation

$$\epsilon_{ij} \longrightarrow \epsilon_{ij} - \lambda_i - \mu_j \,, \qquad\qquad h_0 \longrightarrow h_0 + \sum_i \lambda_i + \sum_j \mu_j \,, \tag{6.5}$$

which are thus a sort of gauge transformations for the system. The space of gauge choices is isomorphic to $\mathbb{R}^{2N}$; the subset for which $\epsilon_{ij} \geq 0$ is a closed convex. Since now on we restrict ourselves to this subset.

The Hungarian algorithm consists in finding a sequence of gauge transformations, increasing at each step the lower bound h_0, so that eventually it will reach the minimum cost $E_{\min}$. At the same time the algorithm will provide the permutation $\pi(i)$ such that $E_{\min} = \sum_i \epsilon_{i\pi(i)}$; the entries $\epsilon_{i\pi(i)}$ chosen in the solution will all be equal to zero in the final gauge.

We will assume in the following that the costs ϵ_{ij} are sampled with a non-singular real valued function. A consequence of this assumption is that the subset of instances where accidental degeneracies are present has a null measure (in the Lebesgue sense), that is, any linear combination of the entries with non-null coefficients c_{ij}, is zero only on a subset of zero measure of the instances. In particular, for gauge-invariant combinations (i.e. such that $\sum_j c_{ij} = 0$ for each i, and $\sum_j c_{ij} = 0$ for each j), this is true under any arbitrary gauge transformation.

In particular, for each sequence

$$(i_1, j_1, i_2, j_2, \ldots, i_l, j_l) \,,$$

the choice $c_{i_k j_k} = 1$, $c_{i_k j_{k+1}} = c_{i_l j_1} = -1$ and $c_{ij} = 0$ otherwise corresponds to a gauge-invariant combination, so

$$\mathrm{prob}\Big(\sum_{ij} c_{ij}\epsilon_{ij} = 0\Big) = 0 \tag{6.6}$$

after each gauge transformation, i.e. at each step of Hungarian Algorithm.

For a subset of gauge choices, which we call *S-gauges*, there is at least one null entry for each row and for each column. Given a gauge not fulfilling these conditions, just apply the gauge transformation which shifts all rows of their minimum entry, and then does the same for the columns, to have a S-gauge. The Hungarian Algorithm works applying gauge transformation in the set of S-gauges.

Consider the complete bipartite graph $\mathcal{K}_{N,N}$, where each vertex of the first kind is associated with a row of the original matrix, and each vertex of the second kind to a column. The subgraph $\mathcal{G}_0 \subseteq \mathcal{K}_{N,N}$ has vertices $V(\mathcal{G}_0) \equiv V(\mathcal{K}_{N,N})$, and an edge (i,j) is in $E(\mathcal{G}_0)$ if and only if $\epsilon_{ij} = 0$. As a consequence of the property (6.6), no loops are present, so $\mathcal{G}_0$ is a forest (that is, a collection of trees). Furthermore, in a S-gauge, $\mathcal{G}_0$ is a *spanning* forest, i.e. all the vertices have coordination larger than one. Note that the lower-bound is tight if and only if $\mathcal{G}_0$ allows for a dimer-covering, as the configuration in which $n_{ij} = 1$ on pairs (i,j) of the covering has cost $\mathcal{H}_{\epsilon,h_0}(n) = h_0 \equiv E_{\min}$.

This fact is crucial in the way Hungarian Algorithm works: the desired optimal configuration is the dimer-covering of $\mathcal{K}_{N,N}$ which minimizes the cost function – but searching for an optimal dimer-covering on a generic graph is in principle a difficult task, because of "information feedback" due to the presence of loops. On the contrary, the Hungarian Algorithm involves as a first step the search for a dimer-covering on the forest $\mathcal{G}_0$, then, as long as contraddictions appear, a second step involves a global move on the entries defining the instance. It is a peculiarity of the Assignment problem that, as long as $h_0 < E_{\min}$, it is possible to find a gauge transfomation which improves the bound h_0. The same approach can be applied to variants of Assignment which are NP-complete [35], and of course it fails to provide a complete deterministic algorithm for these problems. Their failure is due to the lack of this last property.

So the main property of Assignment Problem used by the Hungarian Algorithm is the following

Theorem 3 *As long as $\mathcal{G}_0$ does not allow for a dimer-covering, a gauge transformation exists such that the lower bound h_0 is improved in the new gauge.*

We prove this fact as a sequence of lemmas. First note that the search for a dimer-covering on a forest can be performed separately in each tree. Then, in a given S-gauge, either all trees allow for a dimer covering (in this case the algorithm stops for the reasons described above), or a certain tree does not allow for a dimer-covering. Define a *contraddiction subtree* a rooted subtree of a tree (that is, a subtree with a privileged vertex) such that

- the root is the only vertex connected to the rest of the tree (if the subtree does not coincide with the tree);
- the root has coordination two in the subtree;

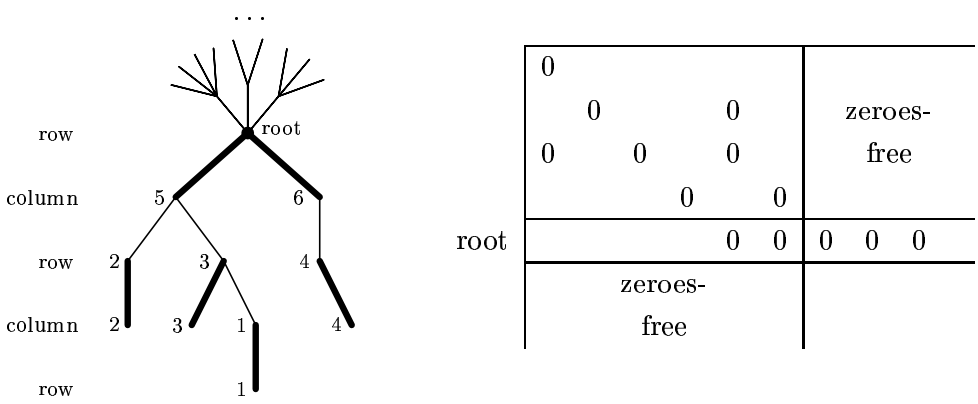


Figure 6.1 An example of a contraddiction subtree, and the corresponding matrix ϵ_{ij}.

- each of the two branches connected to the root allows for a dimer-covering which contains the edge incident on the root.

We want to prove that

Lemma 1 *If a tree does not allow for a dimer-covering, it contains a contraddiction subtree (which can be found in short time).*

Start imposing that leaf edges are in the covering, then that edge neighbouring leaf-edges are not in the covering and so on. This procedure is fast and deterministic on a tree (the time of steps is poportional to the number of edges), and either produces a dimer-covering or identifies a contraddiction subtree (cfr. fig. 6.3 for an example).

Lemma 2 *If in a given S-gauge there is a contraddiction subtree, a gauge transformation exists (and can be found in short time) improving the lower bound h_0.*

The contraddiction subtree inherits from $\mathcal{K}_{N,N}$ the property of being bipartite. In fact the number of vertices at odd distance from the root, minus the number of vertices at even distance, is exactly two. This is true for the smallest contraddiction subtree with two edges, •═(root)═•, and inductively true on larger contraddiction subtrees, as they can be reconducted to the smallest one by recursively removing subgraphs of the kind (Rest)—•═•. So, up to a relabeling of rows and columns, and a matrix transposition, a contraddiction subtree corresponds to a $n \times (n+1)$ block such that

- the n-th row corresponds to the root vertex;
- rows from 1 to $n-1$ correspond to vertices at even distance;

- columns from 1 to $n+1$ correspond to vertices at odd distance;
- the top right $(n-1)\times(N-n-1)$ block and the bottom left $(N-n)\times(n+1)$ block are free of zeroes.

Because of the last property, for δ sufficiently small, an allowed gauge transformation is

$$\lambda_i = -\delta \qquad i \le n\,; \qquad\qquad \mu_j = \delta \qquad j \le n+1\,; \tag{6.7}$$

$$\lambda_i = 0 \qquad i > n\,; \qquad\qquad \mu_j = 0 \qquad j > n+1\,. \tag{6.8}$$

The largest allowed value of δ is the minimum entry in the bottom left $(N-n)\times(n+1)$ block. In the new gauge, $\mathcal{G}_0$ has slightly changed: we have one more edge connecting one of the first $n+1$ column-vertices to one of the last $N-n$ row-vertices, and all the edges (possibly none) connecting the root vertex to the rest of the tree are removed. The value of h_0 has increased of δ, which is expected to be of order $\frac{1}{(N-n)(n+1)}$. In a first part of the algorithm, the forest is composed of a macroscopic number of trees, whose size is order 1, so the gain is order $1/N$. At the very end, we are left with large trees, so, in the worst case $n \simeq N/2$, the gain is of order $1/(N^2)$ and the complexity of a single move is order N. Thus, a rough estimate of the algorithm complexity is N^3. Note that the peak of complexity is on the last moves, so we expect a crossover in the growth rate of the lower bound. We will turn back to this point in section 6.5.

6.4 Algorithmic Implementation of Cavity Equations

Since Cavity Equations are valid for a given instance of the disorder, it is possible to use them also in the special case of the Assignment Problem as a starting point for an algorithm. If we want to solve the optimization problem, we must use the zero-temperature limit of equations (4.11a):

$$g_{i\to j} = \max_{j'\neq j}(-\epsilon_{ij'} - h_{j'\to i})\,, \tag{6.9a}$$

$$h_{j\to i} = \max_{i'\neq i}(-\epsilon_{i'j} - g_{i'\to j})\,. \tag{6.9b}$$

Cavity equations at finite temperature can be used, in case, for an annealing to improve performances. We remark that the gauge invariance pointed out in section 6.3 has a counterpart also in cavity equations implementation: if we change

$$\epsilon_{ij} \to \epsilon_{ij} - \lambda_i - \mu_j\,; \tag{6.10a}$$

$$g_{i\to j} \to g_{i\to j} + \lambda_i\,; \qquad\qquad h_{j\to i} \to h_{j\to i} + \mu_j\,; \tag{6.10b}$$

equations (6.9) do not change.

In this section, we outline the simplest implementation, without paying attention to several possible smart ways of reducing algorithmic complexity: our main goal is to point out some general difficulties we have noticed.

A possible implementation is an algorithm iterating at each time step t the following maps:

$$g_{i\to j}^{(t+1)} = \max_{j'\neq j}(-\epsilon_{ij'} - h_{j'\to i}^{(t)})\,, \tag{6.11a}$$

$$h_{j\to i}^{(t+1)} = \max_{i'\neq i}(-\epsilon_{i'j} - g_{i'\to j}^{(t)})\,. \tag{6.11b}$$

The cavity fields h and g are $N\times N$ matrices, that can be initialized to zero, or to a random real value. Consistently with the discussion of section 4.1 a good choice for the cost matrix entries ϵ_{ij} is to draw them randomly with, say, flat measure on the interval $[0, N]$.

From the first step of the iteration on, fields $g_{i\to j}$ with the same row-index i will all assume the same value, but at most one, assuming a smaller value. The same would happen for fields $h_{j\to i}$ with the same column-index j. According to the general Belief- and Survey-Propagation prescriptions, the iteration should stop when all fields are stationary, then a certain number of variables should be fixed, in order to reduce the instance size, finally the procedure should restart from scratch.

A criterium for fixing the variables could be the following. As all the fields $g_{i\to j}$ in a row but at most one are equal, and the same holds for fields $h_{j\to i}$ in a column, and as we have the gauge invariance (6.10), we can shift the fields as

$$\widetilde{g}_{i\to j}^{(t)} = g_{i\to j}^{(t)} - \max_{j'} g_{i\to j'}^{(t)}\,; \qquad\qquad \widetilde{h}_{j\to i}^{(t)} = h_{j\to i}^{(t)} - \max_{i'} h_{j\to i'}^{(t)}\,. \tag{6.12}$$

Now all the fields are zero up to at most one $\widetilde{g}_{i\to j}$ per row and one $\widetilde{h}_{j\to i}$ per column, which are negative. Large negative values for a field $\widetilde{g}_{i\to j}$ (or $\widetilde{h}_{j\to i}$) means that the row-interaction node i (or respectively the column-interaction j) strongly suggests to take the corresponding entry n_{ij}. We could fix $n_{ij} = 1$ on the variables for which the quantity

$$\min(-\widetilde{g}_{i\to j}, -\widetilde{h}_{j\to i})$$

is larger, then restrict the instance matrix ϵ to the minor which complements the fixed entries.

Some comments are to be done. The main problem with this procedure is that *the fields do not converge to a stationary solution.* This is a consequence of several facts. First, two zero modes of the iteration are present. The first one has eigenvalue 1, and is due to the fact that the graph is bipartite. Indeed, say $\{h^*, g^*\}$ is a fixed point of the map (6.11). Then, another fixed point is given by

$$\begin{cases} h = h^* + \delta\,, \\ g = g^* - \delta\,. \end{cases} \tag{6.13}$$

This is no major surprise. Going back to section (4.3), also in the theoretic treatment it was necessary to impose that the fields had the same probability distribution. This problem can thus be fixed in the same way: we impose that the average value of fields of both species is (almost) the same in the following way. We start with the same distribution for both sets of fields; at each time step we iterate the fields using maps (6.11), and getting $h^{(t+1)}$ and $g^{(t+1)}$; we calculate their average value over all pairs[3], $\langle h \rangle$ and $\langle g \rangle$; the new value of the fields is given by

$$\begin{cases} h^{\text{new}} = h^{(t+1)} - \frac{1}{2}(\langle h \rangle - \langle g \rangle)\,, \\ g^{\text{new}} = g^{(t+1)} + \frac{1}{2}(\langle h \rangle - \langle g \rangle)\,. \end{cases} \tag{6.14}$$

The second zero mode of the iterative map corresponds to an eigenvalue of -1: if $\{h^*, g^*\}$ are fixed points of the map (6.11), another point, given by

$$\begin{cases} h = h^* + \delta(-1)^t\,, \\ g = g^* + \delta(-1)^t\,, \end{cases} \tag{6.15}$$

gives a cycle of period 2. This mode is related to the discrete dynamics we are using. One way to fix this problem is to change the dynamics. Indeed, say evolution at each time step is given by

$$g_{i\to j}^{(t+1)} = \tau\Big(\max_{j'\neq j}(-\epsilon_{ij'} - h_{j'\to i}^{(t)})\Big) + (1-\tau)\; g_{i\to j}^{(t)}\,, \tag{6.16a}$$

$$h_{j\to i}^{(t+1)} = \tau\Big(\max_{i'\neq i}(-\epsilon_{i'j} - g_{i'\to j}^{(t)})\Big) + (1-\tau)\; h_{j\to i}^{(t)}\,. \tag{6.16b}$$

For $\tau \in (0,1)$ this map shares the same fixed points of (6.11), and the eigenvalue related to the mode (6.15) becomes $-1+2\tau$: the cycle is no longer stable for any value of $\tau \in (0,1)$. In particular the mode is suppressed in just one iteration if we choose $\tau = \frac{1}{2}$.

The third reason why the fields do not converge using equations (6.11) is more subtle and is not easy to fix, since it is a problem intimately related to the cavity method and the approximations it relies on. Consider the following example. Given the map (6.11), we want to find the optimum assignment for the matrix of size $N \times N$

$$\epsilon_{ij} = \epsilon(1-\delta_{ij})\,, \tag{6.17}$$

initializing all fields to zero. It is clear that the optimal configuration is $n_{ij}^{(\text{opt})} = \delta_{ij}$. Writing the evolution equations for the cavity fields $g_{i\to j}$ and $h_{j\to i}$,

If we initialize all cavity fields to the same value, the evolution reduces to only two equations, one for the diagonal fields $h_{1\to 1} = g_{1\to 1} = \ldots = g_{N\to N} \equiv h_{\text{yes}}$ and one for the out-of-diagonal

[3]This time-consuming procedure can be helped, as for all practical purposes it is sufficient to average only on a few pairs.

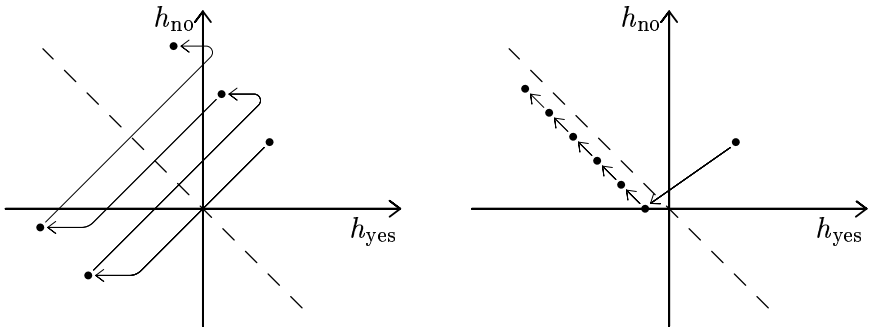


Figure 6.2 On the left, dynamics of the map ϕ of equation (6.18). Iteration can be view as a reflection w.r.t. the dash-drawed diagonal, plus a traslation of $(-\epsilon, 0)$. On the right, dynamics of the enhanced map $\phi_{1/2} = \frac{1}{2}(\phi + I)$ of equation (6.19). Iteration can be view as a projection on the dash-drawed diagonal, plus a traslation of $(-\epsilon/2, 0)$.

ones $h_{j\to i} = g_{i\to j} \equiv h_{\rm no}$ for all $i \neq j$:

$$\begin{pmatrix} h_{\rm yes} \\ h_{\rm no} \end{pmatrix}^{(t+1)} = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} h_{\rm yes} \\ h_{\rm no} \end{pmatrix}^{(t)} + \begin{pmatrix} -\epsilon \\ 0 \end{pmatrix} . \tag{6.18}$$

The behaviour of the solution is drawn in figure 6.4 (left). It shows the period-two oscillations discussed previously. We observe that the fields do not converge to a finite fixed-point: but rather the fields corresponding to the solution are drifting towards $-\infty$ while the others are drifting towards $+\infty$, with velocity linear in ϵ. If we use the strategy of equations (6.16), with $\tau = \frac{1}{2}$, the iteration map becomes

$$\begin{pmatrix} h_{\rm yes} \\ h_{\rm no} \end{pmatrix}^{(t+1)} = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} h_{\rm yes} \\ h_{\rm no} \end{pmatrix}^{(t)} + \frac{1}{2} \begin{pmatrix} -\epsilon \\ 0 \end{pmatrix} , \tag{6.19}$$

for which the period-2 mode is suppressed in one step, but the drift of the fields is still present (cfr. right part of 6.4).

This behaviour is not a peculiarity of this particular example. In fact, this seems to be the typical case. In random finite size systems, qualitatively what happens is that there is a feedback of information through loops, as the constraints are stiff and inference propagation is not damped. This fact shows that the general theory we have developed in the previous chapters has two weak points. On one hand, the assumption we made in section 2.3 when deriving Cavity Equations, that the inference is in some sense small in cavity systems, so that the effect of loops can be neglected, turns out to be uncorrect in this case. On the other hand, in section 4.3 we

have assumed the field distributions to correspond to stationary points of Cavity Equations, and this assumption, too, turns out to be wrong. Nevertheless, Cavity Equations, albeit non valid for cavity iteration at finite size, do find the right thermodynamic quantities, like the average minimum energy, which are intrinsically static; moreover they predict the correct distribution of fields, but only for metastable times, scaling with the size. This has been observed numerically in our simulations (see figure 6.4).

Thus, even in the simple case of the Assignment Problem, we find many interesting phenomena which require a deep investigation, as we note in the concluding chapter 7.

6.5 Average-case complexity for the Satisfiability Assignment Problem

We have seen on chapter 3 that any optimization problem can be related to a satisfiability problem in a natural way. For the case of Assignment, this leads to consider the following problem: given a $N \times N$ matrix ϵ_{ij} and a threshold value α, we search either for a solution n_{ij} satisfying the traditional constraints $\sum_i n_{ij} = 1$, $\sum_j n_{ij} = 1$, and such that the cost function $\mathcal{H}_\epsilon(n) = \sum_{ij} n_{ij}\epsilon_{ij}$ is smaller than α, or for a certificate that such a solution does not exist: it is known that, in the large N limit, the minimum cost $E_{\min}$ concentrates on the value $\pi^2/6$. So we deduce that, in the parameter α, the satisfiability variant of assignment undergoes a SAT-UNSAT transition at the critical value $\alpha_c = \pi^2/6$.

In chapter 3 we discuss a simple reduction argument which typically applies to these pairs of optimization- and related satisfiability-problem. This argument suggests that the worst-case complexity of the optimization problem is strictly related to the worst-case complexity of the satisfiability problem. What typically happens is that the average-case complexity of the satisfiability problem at fixed threshold value α has a peak in correspondence of the SAT-UNSAT transition.

This can be understood via a general procedure, useful for hard problems, called *branch and bound* method. Organize the search process in a tree structure: feasible solutions are located on the leaves of the tree, while solutions having in common the choice of the first k variables have a common ancestor at level k. For example, a branch-and-bound tree for the set of feasible solutions of Assignment Problem, in the case $N = 4$, is shown in figure 6.5. The search process explore the tree following the branching structure. At a given node, all the leaves below it have in common a certain fraction of variable choices. One can evaluate a (computationally simple) lower-bound on the minimum cost for the remaining subproblem, and, at each level, for a certain fraction of branches, determine a priori that no solutions are present, thus skipping the search inside the branch.

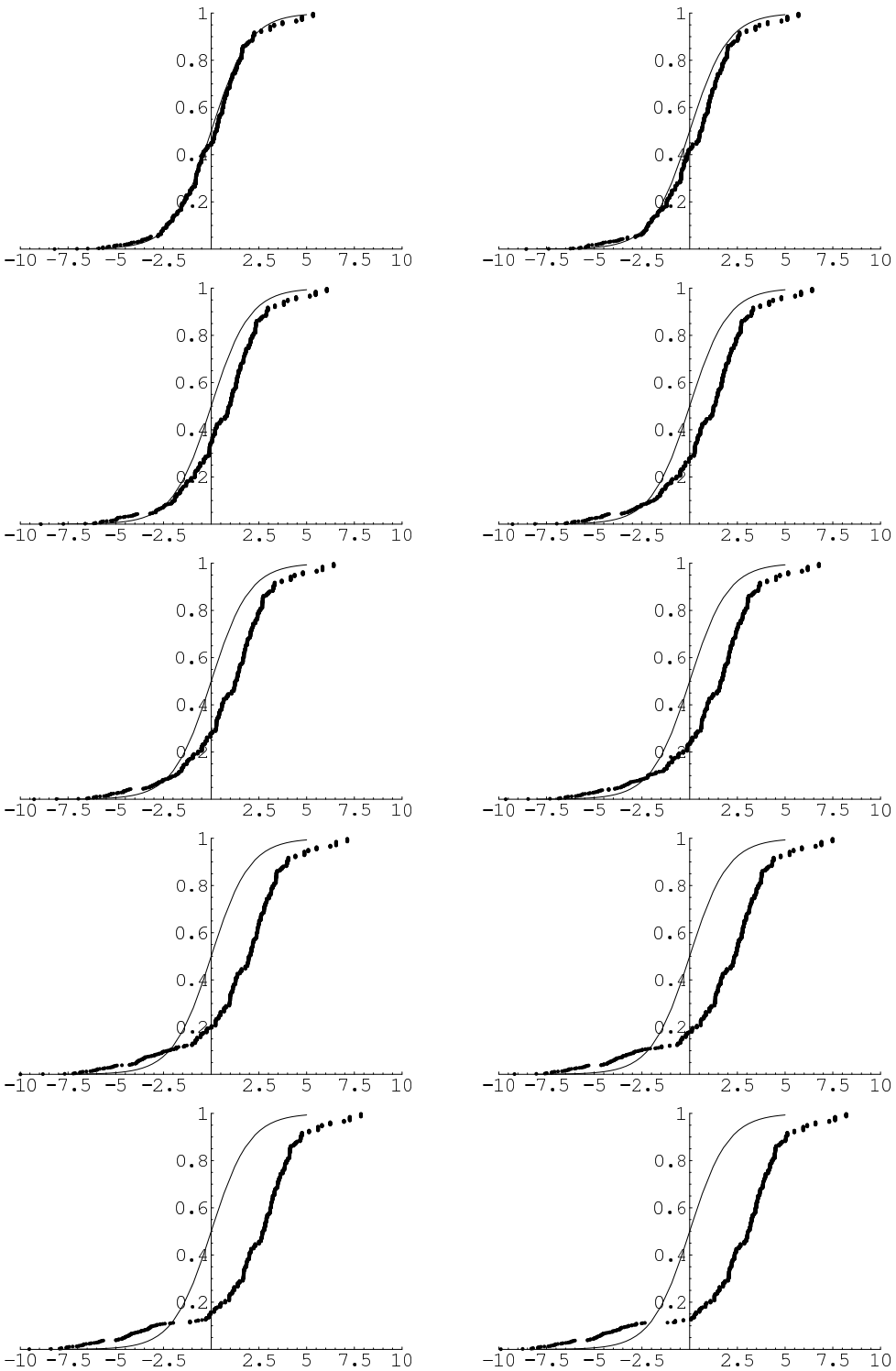


Figure 6.3 Time shots of numerical integrated distributions of fields $F(x)$, for a system of size $N = 100$, each 100 iterations of cavity equations. In solid line, the theoretical prevision $F(x) = e^x/(1 + e^x)$.

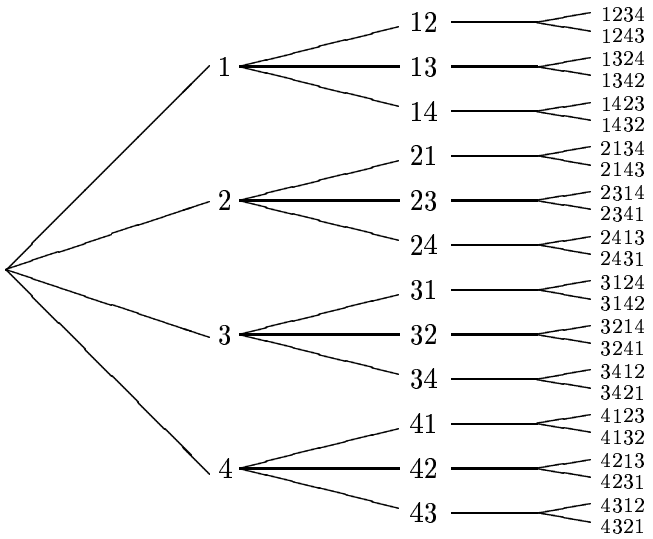


Figure 6.4 A branch-and-bound tree for the set of feasible solutions of Assignment Problem, in the case $N = 4$. Each string of $k \leq N$ entries corresponds to the sequence $\pi(1), \ldots, \pi(k)$, where $\pi(i) = j$ if and only if $n_{ij} = 1$.

When one is deep inside the SAT phase, a heuristic algorithm which follows the more "reasonable" branches first can fastly find a solution with a small fraction of backtracking moves. On the other side, deep inside the UNSAT phase, in most cases the bound procedure succeeds at the very first nodes, thus skipping large regions of the tree.

In the case of the Assignment Problem, it is straightforward to recognize all these general features, using the diverse algorithmic implementations discussed so far. In the region where α is larger than the expected average value, $\pi^2/6$, we are in the SAT phase. Many algorithms, suitable also for the optimization problem, e.g. Simplex method or Cavity method, work suggesting a sequence of possible solutions of decreasing cost. The satisfiability variant, in the region $\alpha > \alpha_c$, can in case stop before the sequence reaches for the optimal solution, if a configuration is reached, whose cost is lower than the threshold. On the other hand, below the critical value, it is necessary to reach for the ground state before getting a certificate of unsatisfiability. Thus, the average case complexity of these algorithms is a monotonic decreasing function of the threshold for $\alpha > \alpha_c$, while below the critical value it is constant on the maximum value, reached in α_c.

The Hungarian Algorithm, on the other hand, works differently: instead of having a sequence of decreasing costs, associated to each possible solution, there is a sequence of lower bounds of $E_{\min}$, which increase until this value is eventually reached. In the satisfiability variant, when looking for a UNSAT certificate, the Hungarian Algorithm will stop, as soon as a lower bound higher than the threshold is reached. In the SAT region, instead, the Hungarian Algorithm must

always find the minimum energy configuration in order to have the largest possible lower bound.

Furthermore, we have already noticed in the end of section 6.3 that the rate of growth of the lower bound h_0 has a size-dependence crossover approaching the end of the procedure. The complete optimization Hungarian Algorithm has a power-law complexity. From the previous observation, we expect that the satisfiability Hungarian Algorithm in the large N limit reaches for this complexity only for values of α approaching α_c.

Thus, a mixed algorithm, working in the best possible way both in SAT and UNSAT region, will have a peak of complexity only in proximity of the critical value, which becomes more and more narrow in the large N limit.

6.6 Application of Assignment to Scheduling

The Assignment Problem has several practical applications. One of the most relevant is in scheduling packet traffic in routers [36]. Internet, for instance, is basically a large packed switched network: travelling informations are divided into packets, and the most important switching system in Internet are the routers. Their main task is to receive packets from input ports, find their local destination port on the basis of the routing table, and then deliver them to the output ports. Routers of all throughtputs (i.e. the traffic volume supported) are present, but in particular there is a relatively small number of basic routers, named backbone routers, switching almost all information traffic, and they must meet strict requirements in terms of speed and reliability.

The problems routers face is basically an Assignment problem. A good model to describe how a router works is the following: there are N input ports and N output ports; at each discretized time slot one information packet, or none, arrives at each input port, carrying a label with its destination. The router can support at most one outcoming packet per port, so, if at a given time slot two incoming packets must go out of the same port, one of them must wait.

The router thus must schedule the crossing times of packets, and manage queues. Queue scheduling is a complex and delicate topic, and we will not treat here. However, whatever the scheduling policy, the result is an "priority" parameter for each pair of input-output ports, related to the number of packets waiting, reflecting how long they have been waiting, and possibly other informations, like their urgence, level of priority, and so on.

At this point the router must find an assignment, in order to decide which packets should cross the switch: in our notation of chapter 4, the cost matrix ϵ_{ij} is related to the priority parameters, and if $n_{ij} = 1$, a packet arrived at port i and directed to port j is routed.

The algorithm chosen to find an Assignment must meet a certain number of requirements. Speed is the very first requisit: as the algorithm must produce a solution at each trigger time

of the switcher, the complexity of the algorithm must not scale with the number of ports. On the contrary, optimality is secondary, provided that, for the largest possible set of incoming traffic patterns, the gap between the cost of the optimal solution and the one of the solution identified by the algorithm remains of order 1. Another requisit for a performing algorithm is parallelization. It is not possible to solve a problem of macroscopic size in a time of order 1, unless one makes use of a macroscopic scalable circuitry.

For example, the Hungarian Algorithm, which develops a particularly smart combinatorial idea, is a good complete algorithm for software implementation, but is *not* a good algorithm for a hardware implementation in a specific router circuitry.

On the contrary, the Cavity Method approach seems to meet a large number of the requirements exposed above: the iterative equations are extremely simple and scalable, and processable in parallel. So an heuristic message-passing algorithm could be feasible for applications in this field. Nonetheless, a wider numeric study of performance for various algorithmic implementations and traffic patterns are necessary.

7. Conclusions and perspectives

In this thesis, two main research topics have been treated. On one hand, a revisitation of the general theory of the cavity method has led to the understanding of some potentially delicate points; a numeric implementation of Cavity Equations for a simple model, the Assignment Problem, has confirmed that these points were indeed delicate, since the effect of inference loops at finite size, neglected in Cavity Ansatz, does play a relevant role. Recipes on the solution of these problems in the design of a Cavity algorithm for Assignment have been proposed.

On the other hand, we have formalized the averaging procedure of Cavity Equations in the mathematical framework of Independent Poisson Processes. This led us to deal with distributional equations, which we have solved also in some new cases, obtaining original results (the finite-temperature analysis of the Assignment Problem; the k-assignment Problem for generic value of k).

Both these research areas suggest some directions for further investigation, which we outline in the following.

7.1 General derivation of Cavity Theory

In chapter 2 we have given heuristic arguments for the fact that, for a certain family of models (*weak-inference* models), yet to be outlined precisely, in the infinite-size limit, Cavity Approximations 2.3.1 hold. Moreover, for systems showing symmetry breaking of 1-RSB kind, a further set of hypotheses concerning the complexity of the system (cfr. page 2.5) are expected to be valid in the infinite-size limit.

Here we argument that these sets of approximations can both be handled at finite-size in a systematic way, inside the theoretical framework of a system with many pure phases. Indeed, finite-size corrections to this picture are expected to be asymptotically irrelevant w.r.t. corrections related to assuming factorization hypotheses and neglecting higher-order dependence in the free energy of cavity-fields distributions, as the latter are expected to be of power-law order in N, while the first are expected to scale exponentially with N. For instance, as free-energy barriers are expected to scale with a power-law in N, $\Delta F_{\mathrm{barr}}^{(\alpha,\alpha')} \sim N^{\gamma}$, metastability times are

expected to scale as

$$\tau_{(\alpha,\alpha')} \sim e^{-c\Delta F_{\rm barr}^{(\alpha,\alpha')}} \sim e^{-c' N^{\gamma}} .$$

Corrections to Cavity Ansatz 2.3.1, inside a pure phase, can be derived from a perturbative evaluation of correlation functions, calculated self-consistently using the cavity fields which solve the set of infinite-size Cavity Equations (2.11). We expect these corrections to be relevant w.r.t. corrections to the pure-phase picture, at least for models defined over finite-connectivity random graphs, as, although correlation functions are expected to decrease exponentially with distance

$$G_{x,x'}(\sigma,\sigma') \sim \exp -|x-x'|/\xi \,,$$

the average loop lengths, which determines average distances of variables in cavity systems, are expected to be of logarithmic order in size.

A systematic expansion on correlation functions, inside a pure phase, analogous to the Dobrushin-Lanford-Ruelle expansion for ordered systems [1], is eventually possible. Furthermore, the study of the first-order correction should lead to a more precise definition of a weak-inference system, a concept treated only heuristically in literature so far.

For what concerns 1-RSB hypotheses of page 2.5, also when restricting to systems whose complexity has the form outlined in section 2.4 (which are the most interesting ones, for many purposes), a certain number of approximations play a role. Assume that approximation of $\Sigma(f)$ with a continuous concave function is justified, thus the reshuffling parameter y select one typical value F^* of free energy only up to statistical fluctuations of order $\mathcal{O}(\sqrt{N})$, as

$$e^{-yF+N\Sigma(F/N)} = C(y) \exp\left(-\frac{1}{2N} \left|\Sigma''(\tfrac{F^*}{N})\right| (F-F^*)^2 + \mathcal{O}\left(\frac{1}{N^2}\right)\right) .$$

So we expect that the error due to the approximation

$$Q_{\partial\!\!\!/}(\{h_{j\to a}\}|N(f^*+\Delta f)) \simeq Q_{\partial\!\!\!/}(\{h_{j\to a}\}|Nf^*) \equiv Q_{\partial\!\!\!/}(\{h_{j\to a}\})$$

is proportional to the function

$$\frac{1}{\sqrt{N}} \frac{\partial}{\partial f} Q_{\partial\!\!\!/}(\{h_{j\to a}\}|Nf)\Big|_{f^*} .$$

Nonetheless, from a more careful analysis of equation (2.29) we derive that this function is averaged with a symmetric gaussian at first-order correction, so the first non-vanishing corrections are of order $\mathcal{O}(\frac{1}{N})$.

A further source of error is the factorization hypothesis of equations (2.33). The analysis of this contribution is connected to the quantification of inference, in analogy with the previous discussion on estimating corrections to Cavity Ansatz, although with some subtle difference. Anyway, corrections should be of the same order of magnitude.

A last (and also least) correction derives from neglecting discrepancies between complexity functions of different cavity systems. As free-energy differencies of pure phases between different cavity systems are expected to be of order 1, one could naïvely argue that this correction is of order $1/N$. Some care is required in this case: complexity functions are involved with their derivaties, and, if some regularity hypothesis are missing, discrepancies could be potentially large. Nonetheless, these regularity hypotheses can typically be related to physically reasonable hypotheses on decorrelation between free-energy shifts in different pure phases, and are believed to hold for most interesting systems.

7.2 Applications of distributional-equation formalism

The systematic discussion of a method for transposing averages of cavity equations into distributional equations, outlined in section 2.6, is original in its generality. It has allowed us to derive in a relatively straighforward way some new results on the statistical properties of the Assignment and k-assignment Problems, and seems promising for further future applications.

In particular, it is natural to extend it to two models intimately related to the Assignment Problems, treated in this work. The first is the (D, D')-*assignment Problem*, which is a multidimensional variant of Random Assignment: the grid ϵ is generalized to dimension $D + D' > 2$, and the constraints apply to subgrids of a given dimension D, with $1 \leq D < D + D'$. This problem reduces to traditional Assignment for $(D, D') = (1, 1)$, while it is NP-complete for each pair $(D, D') \neq (1, 1)$ (cfr. [35] for a proof in the case $(D, D') = (1, 2)$). The derivation of Cavity Equations for this problem is a straightforward generalization of section 4.2 using the recipe given in chapter 2. An algorithmic study of this problem with a Survey Propagation approach seems interesting, since the NP-completeness of the problem.

A second problem is the famous *Travelling Salesman Problem* (TSP), shortly introduced also in chapter 3. In the 2-assignment problem, for each feasible solution n_{ij}, the graph $\mathcal{G}(n) \subset \mathcal{K}_{N,N}$ such that $E(\mathcal{G}) = \{(i, j), n_{ij} = 1\}$ consists of a set of self-avoiding closed paths. The TSP is a variant of the 2-assignment where a non-local constraint is present: $\mathcal{G}(n)$ must be composed of a single loop. This makes TSP a NP-complete problem, but, since this constraint decimates the phase space only for a subexponental factor (bounded below by the factor $1/N$), it is expected that the RS ansatz, valid for the 2-assignment, is valid also for this problem [37]. This non-local constraint can be reconducted to a local one introducing a (colour) auxiliary variable on occupied bonds, and letting the number of colour go to zero [38]. The corresponding cavity equations, even though involving some technical difficulties, seem suitable both for an algorithmic implementation and for a theoretical investigation, extending the formalism of distributional equations, and complementing the results of Parisi and Mézard, obtained via the

replica approach [10].

7.3 Numerical investigations on Assignment Problem

At the beginning of our work, the Assignment Problem was considered as a toy model for applying cavity-based algorithms, eventually with direct applications as an algorithm for switching problem in network routers. Instead it has shown unexpected striking features: the algorithmic implementation has presented, and theoretical treatment has partially explained, the need for modifying *ad hoc* the general recipe [24] for this model, in order to make the cavity fields converge in cavity iteration.

Indeed, using a traditional iterative dynamics we find essentially three problems

- presence of zero-modes in the dynamics;
- occurrence of cycles of period 2;
- metastability of the traditional solution, and asymptotic drift of the fields toward $\pm\infty$.

The peculiarities of the model which we believe to account for these problems are mainly

- gauge invariance of Cavity Equations;
- bipartition of the interaction-node part of the factor graph (all second-neighbours of row-interaction nodes on the factor graph are column-interaction nodes, and vice versa);
- presence of stiff constraints with just one allowed configuration in certain marginals of variables in cavity systems.

In our advice, a further theoretical and numerical investigation of these aspects should be done. Particularly, if we correct the wrong stability assumption in the derivation of field distribution, into an assumption of asymptotically stationary drift, it should be possible to extract the asymptotic form of field distributions in the drift regime. In perspective, when the solution of these highlighted subtleties has been achieved, the primary applicative idea to routers could be reasonably tried.

As a last remark, the connection between the Hungarian Algorithm and a certain prescription on implementing Cavity Iteration, seems interesting. On the other hand, the Hungarian Algorithm seems to be related also to the Simplex Method, which is a paradigm in approaching problems that can be stated as Linear Programming Problems (note that a large number of famous problems, as e.g. the K-SAT, belong to this class). Working out these connections might lead to a better understanding of Cavity-Equations based algorithms.

A. Mathematical properties of Independent Point Processes

A.1 Poisson Point Distributions

We define a *Poisson Point Distribution* (PPD) as a measure over the space Ξ of (possibly infinite) configurations of points on the line $\mathbb{R}$ such that any two disjoint intervals of $\mathbb{R}$ behave independently. We refer the reader to [39] for a short clear physical description of the field, and [40], for a textbook.

Although defined on the huge space Ξ, a PPD is fully identified by a single functional parameter, $\rho(\xi)$, describing the density of points. With an eye to physical application, and in order to avoid tedious technicalities, in the following we will consider only finite densities, integrable on every compact interval of $\mathbb{R}$.

A definition involving a measure over the space of infinite configurations, although intuitive, is mathematically imprecise. Nonetheless, the independence property allows to use as a definition the natural property that, for each pair of measurable sets $\Omega_1, \Omega_2 \in \mathbb{R}$ such that

$$\int_{\Omega_1} \mathrm{d}\rho(\xi) = \alpha_1 \,, \qquad \int_{\Omega_2} \mathrm{d}\rho(\xi) = \alpha_2 \,, \qquad \int_{\Omega_1 \cap \Omega_2} \mathrm{d}\rho(\xi) = 0 \,,$$

the probability distribution of the number of points n_1 inside Ω_1 and n_2 inside Ω_2, averaged over the PPD, is factorized:

$$p(n_1, n_2) = p_{\Omega_1}(n_1) p_{\Omega_2}(n_2) \,. \tag{A.1}$$

Applying independence reasoning on infinitesimal intervals, it is simple to show that for any set $\Omega \in \mathbb{R}$ such that $\int_\Omega \mathrm{d}\rho(\xi) = \alpha$ is finite, the probability distribution of the number n of points inside Ω is a Poissonian of rate α:

$$\mathrm{Poiss}_\alpha(n) = e^{-\alpha} \frac{\alpha^n}{n!} \,. \tag{A.2}$$

A physical intuition of these distributions is given by a one-dimensional free gas, in the grand-canonical ensemble. The mathematical independence property is the correspective of the absence

of interaction in the gas, and the density is simply related to the chemical potential μ, by equation $\rho(\xi) = \exp[-\beta\mu(\xi)]$.

When $\int \mathrm{d}\rho(\xi) = \mathcal{N}$ is finite, we deal with a *finite PPD*. We can write a configuration $\vec{\xi} \in \Xi$ as $(\xi_1, \ldots, \xi_n)$, with $\xi_i < \xi_{i+1}$ and n distributed Poissonianly with rate $\mathcal{N}$.

For a finite PPD, a clear algorithmic recipe exists for a process of construction of a configuration $\vec{\xi} \in \Xi$ with the proper distribution: first extract the total number of points, n, with probability $\mathrm{Poiss}_{\mathcal{N}}(n)$, then extract n i.i.d. random variables ξ_i on $\mathbb{R}$, with (normalized) measure $\frac{1}{\mathcal{N}}\mathrm{d}\rho\,(\xi)$.

When $\int \mathrm{d}\rho(\xi) = \infty$ each configuration $\vec{\xi} \in \Xi$ contains an infinite number of points, accumulating only at $+\infty$. We call such a distribution a *semi-infinite PPD*. We can write the configuration as an infinite sequence, $\vec{\xi} = (\xi_1, \xi_2, \xi_3, \ldots)$, with $\xi_i < \xi_{i+1}$ for each $i \in \mathbb{N}$.

Semi-infinite PPDs allow for a finite-size limit procedure of construction. Define

$$\mathrm{d}\rho_{\,L}(\xi) = \mathrm{d}\rho\,(\xi)\theta(L - \xi)\,. \tag{A.3}$$

For each finite L we deal with a finite PPD, and thus the construction above is possible. If we enlarge the domain to $L + \Delta L$, from independence property A.1, it follows that the enlarged distribution coincides with the previous one on the restriction to $(-\infty, L]$.

A certain number of transformations preserve the independence property:

Merging. Given two PPDs with densities $\rho_1(\xi)$ and $\rho_2(\xi)$, for each pair of configurations $(\vec{\xi_1}, \vec{\xi_2})$, extracted with the two measures, consider the configuration $\vec{\xi} = \vec{\xi_1} \cup \vec{\xi_2}$. The distribution obtained by this procedure is still a PPD, with density

$$\rho'(\xi) = \rho_1(\xi) + \rho_2(\xi)\,.$$

Filtration. Given a PPD with density $\rho(\xi)$, and a function $p(\xi) : \mathbb{R} \to [0, 1]$, for each configuration $\vec{\xi}$ reject each point $\xi_i \in \vec{\xi}$ independently with probability $1 - p(\xi_i)$. The distribution obtained by this procedure is still a PPD, with density

$$\rho'(\xi) = \rho(\xi)p(\xi)\,.$$

Reshuffling. Given a PPD with density $\rho(\xi)$, and a normalized distribution $f(x)$, we shift each point ξ_i of a configuration $\vec{xi}$ by an independent random quantity x_i sampled with $f(x)$. The distribution obtained by this procedure is still a PPD with density

$$\rho'(\xi) = \int \mathrm{d}x\, f(x)\rho(\xi - x) = (\rho * f)\,(\xi)\,.$$

A.2 Independent Poisson Processes

We stress here a fact which has been anticipated in the previous section. As, by hypothesis, dealing with finite or semi-infinite PPDs, the density is integrable from the left, i.e. for each allowed density $\rho(x)$ and each real value x, the function

$$R(x) = \int_{-\infty}^{x} \mathrm{d}\rho(x')$$

is finite, for any configuration $\vec{\xi}$ a natural ordering of the points $(\xi_1 < \xi_2 < \ldots)$ is defined.

A non-trivial consequence of independence is that one can define an iterative process of construction of a configuration, such that at the k-th step the k smallest points $\{\xi_i\}_{i\leq k}$ of configuration $\vec{\xi}$ are sampled. Moreover, the probability distribution of the $(k+1)$-th point, conditioned to the first k points already extracted, only depends on the k-th point. For the simplest case of the semi-infinite PPD the sorted sequences $(\xi_1, \xi_2, \ldots)$ should be extracted with the memory-1 Markov Process [1]

$$p(\xi_n|\xi_{n-1}, \cdots, \xi_1) = p(\xi_n|\xi_{n-1}) = \rho(\xi_n) e^{-\int_{\xi_{n-1}}^{\xi_n} \mathrm{d}\rho(\xi)} , \tag{A.4}$$

where conventionally $\xi_{-1} \equiv -\infty$. For finite PPDs, the conventional symbol BREAK should be added: when a BREAK is extracted, the sequence $(\xi_1, \xi_2, \ldots, \xi_k)$ terminates. The memory-1 Markov Process associated to finite PPDs is

$$p(\xi_n|\xi_{n-1}, \cdots, \xi_1) = p(\xi_n|\xi_{n-1}) = \begin{cases} \xi_n & \mathrm{d}\rho(\xi_n)\, e^{-\int_{\xi_{n-1}}^{\xi_n} \mathrm{d}\rho(\xi)} \\ \text{BREAK} & 1 - e^{-\int_{\xi_{n-1}}^{+\infty} \mathrm{d}\rho(\xi)} \end{cases} \tag{A.5}$$

This formulation as a point process accounts for the name of *Poisson Point Process* (PPP) devoted to Markov Chains related to PPDs in the way described above. The simple facts about merging, filtration and reshuffling holds also for PPPs. An interesting fact is that the PPP-construction of a semi-infinite process allows to deal directly with an infinite number of points, while, on the other hand, the construction used for i.i.d. extraction on semi-infinite PPDs was based on the introduction of a finite-size cut-off.

Now consider some simple applications. Given a PPP with density $\rho(\xi)$, and a normalized distribution $f(x)$, we consider a reshuffling process as described in the previous section. For a given threshold a, we ask for the probability that all the points $\xi_i + x_i$ are larger than a. We can

[1] Note that probabilities are automatically normalized, as

$$\int_{\xi_{n-1}}^{+\infty} \mathrm{d}\xi\, \rho(\xi) e^{-\int_{\xi_{n-1}}^{\xi} \mathrm{d}\rho(\xi')} = -e^{-\int_{\xi_{n-1}}^{\xi} \mathrm{d}\rho(\xi')} \Big|_{\xi_{n-1}}^{+\infty} = 1 .$$

understand this problem either as a Reshuffling procedure, or as a Filtration procedure. Indeed in the first context we have

$$\text{prob}(\xi_i - x_i > a \quad \forall\, i) = e^{-\int_{-\infty}^{a} \mathrm{d}\xi\, (\rho * f)(\xi)} = e^{-(\rho * F)(a)}\,, \tag{A.6}$$

while in the second

$$\text{prob}(\xi_i - x_i > a \quad \forall\, i) = e^{-\int \mathrm{d}\xi\, \rho(\xi)\text{prob}(\xi - x < a)} = e^{-\int \mathrm{d}\xi\, \rho(\xi) F(a-\xi)} = e^{-(\rho * F)(a)}\,. \tag{A.7}$$

A less trivial calculation is the following. Consider a PPP with density $\rho(\xi)$, and a one-parameter family of distributions $g_\xi(y)$ on an abelian additive group Y (say, $Y = \mathbb{R}$). Given a configuration $\vec{\xi} \in \Xi$, for each point ξ_i sample a value y_i with probability $g_{\xi_i}(y)$. The total value $y = \sum y_i$ is distributed as $g(y)$. Consider an integral transform on Y such that a convolution theorem holds

$$\mathcal{T}(f_1 * f_2) = \mathcal{T}(f_1)\mathcal{T}(f_2)\,,$$

and for which $\mathcal{T}(\delta(y)) = 1$, with $\delta(y)$ the delta function in the group Y. A simple example is, for $Y = \mathbb{R}$, the Fourier Transform or the Laplace Transform, while, for a generic abelian group Y, the Peter-Weyl Transform has the desired properties. Our statement is that the transform $\widetilde{g}(\eta)$ is given by

$$\widetilde{g}(\eta) = \exp\left[-\int \mathrm{d}\rho(\xi)\,(1 - \widetilde{g}_\xi(\eta))\right]\,. \tag{A.8}$$

To prove this equation, introduce the distribution $g_{(\xi)}(y)$ for the truncated sum $y^{(\xi)} = \sum_{i:\xi_i<\xi} y_i$. We have the differential equation

$$g_{(\xi+\mathrm{d}\xi)}(y) = g_{(\xi)}(y) * [\delta(y)(1 - \mathrm{d}\rho(\xi)\,) + \mathrm{d}\rho(\xi)\, g_\xi(y)]\,, \tag{A.9}$$

and, transforming

$$\widetilde{g}_{(\xi+\mathrm{d}\xi)}(\eta) = \widetilde{g}_{(\xi)}(\eta)\,[1 - \mathrm{d}\rho(\xi)\,(1 - \widetilde{g}_\xi(\eta))]\,, \tag{A.10}$$

$$\ln \widetilde{g}_{(\xi+\mathrm{d}\xi)}(\eta) = \ln \widetilde{g}_{(\xi)}(\eta) - \mathrm{d}\rho(\xi)\,(1 - \widetilde{g}_\xi(\eta))\,, \tag{A.11}$$

from which, jointly with the boundary condition $\lim_{\xi\to-\infty} g_{(\xi)}(y) = \delta(y)$, we finally obtain

$$\ln \widetilde{g}(\eta) = \lim_{\xi\to+\infty} \ln \widetilde{g}_{(\xi)}(\eta) = -\int \mathrm{d}\rho(\xi)\,(1 - \widetilde{g}_\xi(\eta))\,. \tag{A.12}$$

As a first specialization of this general formula, we can rederive equation (A.6) for the probability that all points of the reshuffled process are larger than a threshold a.

Say that $y_i = 1$ if $\xi_i + x_i < a$ and $y_i = 0$ otherwise. The quantity $y = \sum_i y_i$ counts the number of points below the threshold, and the desired probability is $g(y = 0)$. The distribution $g_\xi(y)$ is

$$g_\xi(y) = \delta(y)(1 - F(a - \xi)) + \delta(y - 1)F(a - \xi)\,.$$

We use Fourier Transform. Since we have

$$\int \mathrm{d}\rho(\xi)\,(1-\widetilde{g}_\xi(\eta)) = \int \mathrm{d}\rho(\xi)\,F(a-\xi)(1-e^{i\eta}) = (1-e^{i\eta})\,(\rho * F)(a)$$

and since, by Taylor expansion,

$$\int \mathrm{d}\eta\, e^{-i\eta y-\alpha(1-e^{i\eta})} = \mathrm{Poiss}_\alpha(y)$$

we get equation (A.6).

B. Mathematics of Distributional Equations

In this work we have made large use of mathematical expressions of the kind

$$(\text{something}) \stackrel{d}{=} (\text{something else})\,, \tag{B.1}$$

adding somewhere nearby in the text sentences like "... where 'something' is distributed as ... and 'something else' is distributed as ...". We have called this kind of expressions *distributional equations*. In this appendix we want to clarify the mathematical meaning of this framework, and illustrate some elementary algebraic manipulations which concern this kind of expressions.

As clear from the nomenclature, we deal with expressions which are true "on average", given a probability distribution over the variables involved. More precisely, say that an (in case infinite) set of variables $\vec{x} = \{x_i\}$ is distributed with $\mu(\vec{x})$, and that an (in case infinite) set of variables $\vec{y} = \{y_j\}$ is distributed with $\lambda(\vec{y})$. Then, given two functionals $\Phi(\vec{x})$ and $\Psi(\vec{y})$, the equation

$$\Phi(\vec{x}) \stackrel{d}{=} \Psi(\vec{y}) \tag{B.2}$$

means that the marginal distribution of $\Phi(\vec{x})$ w.r.t. the probability distribution $\mu(\vec{x})$ coincides with the marginal distribution of $\Psi(\vec{y})$ w.r.t. the probability distribution $\lambda(\vec{y})$

$$\forall\, z \in \mathbb{R} \qquad \int \mathrm{d}\mu(\vec{x})\, \delta(z, \Phi(\vec{x})) = \int \mathrm{d}\lambda(\vec{y})\, \delta(z, \Psi(\vec{y}))\,. \tag{B.3}$$

As a simple example of the resolution of a distributional equation, consider the following equation

$$\max(x_1, \cdots, x_k) \stackrel{d}{=} y\,, \tag{B.4}$$

where the k variables are i.i.d. with $f(x)$, and y is distributed with $g(y)$. The probability that the maximum value of k identically distributed variables $\{x_i\}$ is inside the interval $[y, y + \mathrm{d}y]$ is given by

- a factor k for the choices of the index realizing the maximum-value;
- a factor $f(y)\mathrm{d}y$ for the probability that a variable is in $[y, y + \mathrm{d}y]$;

- a factor $F(y) = \int_{-\infty}^{y} \mathrm{d}y' \, f(y')$ per each of the other variables, which gives the probability that all the other variables are smaller;

putting together all these factors, we get

$$g(y) = k f(y) F^{k-1}(y) = \frac{\mathrm{d}}{\mathrm{d}y} F^k(y) \,. \tag{B.5}$$

Note that $g(y)$ is automatically normalized.

We remark the importance of having *different* variables on the two sides of the distributional equation: two independent averaging procedures on the two sides of the equation must be done. A simple example is explicative of a certain number of useful warnings: consider the distributional equation

$$x \stackrel{d}{=} y_1 + y_2 \,, \tag{B.6}$$

where x is distributed with $f(x)$ and $y_{1,2}$ are distributed with $g_{1,2}(y)$. The translation in the form of equation (B.3) of this example is

$$f(x) = \int \mathrm{d}y \, g_1(y) g_2(x-y) = (g_1 * g_2)(x) \,. \tag{B.7}$$

The traditional algebraic manipulations of equations are *not* admitted anymore. For example, the distributional equation

$$x - y_2 \stackrel{d}{=} y_1 \tag{B.8}$$

is *not* verified by the same triplets of functions, [1]

$$f(x) = \int \mathrm{d}y \, g_1(y) g_2(x-y) \qquad \not\Longleftrightarrow \qquad \int \mathrm{d}x \, f(x) g_2(x-y) = g_1(y) \,. \tag{B.9}$$

Nonetheless a few manipulations are still possible. Consider the case in which the functionals Φ and Ψ take value in $\mathbb{R}^n$, and the distributional equation is intended in n dimensions. One can write a system of real equations, one per component

$$\begin{cases} \Phi^1(\vec{x}) \stackrel{d}{=} \Psi^1(\vec{y}) \\ \vdots \\ \Phi^n(\vec{x}) \stackrel{d}{=} \Psi^n(\vec{y}) \end{cases} \tag{B.10}$$

Since the distributional identity does not depend on the choice of basis in $\mathbb{R}^n$, a linear combination mixing the n components in the same way on the two sides of the equations is allowed.

[1] In particular, using Fourier Transform, one easily deduce that the two equations (B.9) are satisfied if and only if $|\widetilde{g}_2(\xi)|^2 = 1$ on the whole region in which $\widetilde{f}(\xi)$ and $\widetilde{g}_1(\xi)$ are different from zero, i.e., if y_2 is a real variable, essentially only if $g_2(y_2) = \delta(y_2, \overline{y})$.

For example, if the system of distributional equations

$$\begin{cases} x_1 \stackrel{d}{=} y + y_1 \\ x_2 \stackrel{d}{=} y + y_2 \end{cases} \tag{B.11}$$

holds in the sense above, i.e., given the probability distributions $\mu(x_1, x_2)$ and $\lambda(y, y_1, y_2)$, the distributional equation system translates into

$$\forall\ (x_1, x_2) \in \mathbb{R}^2 \qquad \mu(x_1, x_2) = \int \mathrm{d}y\, \lambda(y, x_1 - y, x_2 - y)\,, \tag{B.12}$$

then the distributional equation

$$x_1 - x_2 \stackrel{d}{=} y_1 - y_2 \tag{B.13}$$

is valid.

Acknowledgments

This thesis is the fruit of a collaboration with Prof. Sergio Caracciolo and Dr. Andrea Sportiello. I wish to thank them for their essential contributions, for all the things they have taught me, for their patience and for their support.

I also wish to thank Dr. Andrea Montanari for many useful suggestions and discussions.

Bibliography

[1] G. Parisi, *Statistical Field Theory,* Perseus books (1998)

[2] K. Wilson, J. Kogut *The Renormalization Group and the ϵ-expansion,* Physics Reports **12C**, 75-200 (1974).

[3] L. Onsager, *Crystal Statistics. I. A Two-Dimensional Model with a Order-Disorder Transition,* Phys. Rev. **65**, 117-149 (1944).

[4] N.N. Bogolubov, A.A. Logunov, I.T. Todorov, *Introduction to axiomatic quantum field theory,* Mathematical physics monograph series Vol.18, Benjamin Press, New York (1975) part VI

[5] F. Strocchi, *Selected Topics on the General Properties of Quantum Field Theory,* World Scientific Publishing Co., Singapore (1993)

[6] R.F. Streater and A.S. Wightman, *PCT, Spin and Statistics and all That,* 5th edition, Princeton University Press Series: Princeton Landmarks in Mathematics and Physics (2000)

[7] M. Mézard, G. Parisi and M.A. Virasoro, *Spin Glass Theory and Beyond,* Word Scientific, Singapore, (1987)

[8] G. Parisi, *Glasses, replicas and all that*, `cond-mat/0301157`, (2003)

[9] M. Mézard, G. Parisi, *Optimization and replicas,* J. Physique Lett. **46**, L771-L778 (1985); M. Mézard, G. Parisi, *On the solution of the random link matching problem,* J. Physique **48**, 1451-1459 (1987)

[10] M. Mézard, G. Parisi, *Replica Analysis of the Travelling Salesman Problem,* J. Physique **47**, 1285-1296 (1986)

[11] R. Monasson, R. Zecchina, *Entropy of the K-Satisfiability Problem,* Phys. Rev. Lett. **76**, 3881-3885 (1996)

[12] R. Monasson, R. Zecchina, *Statistical mechanics of the random K-satisfiability model* Phys. Rev. **E 56**, 1357-1370 (1997)

[13] L. Viana, A.J. Bray, *Phase diagrams for dilute spin-glasses,* J. Phys. **C 18**, 3037 (1985)

[14] E. Gardner, *Spin glasses with p-spin interactions,* Nucl. Phys. **B 257**, 747-765 (1985)

[15] F. Guerra, F.L. Toninelli, *The Thermodynamic Limit in Mean Field Spin Glass Models,* Commun. Math. Phys. **230**, 1, 71-79 (2002)

[16] S. Franz, M. Leone, *Replica bounds for optimization problems and diluted spin systems,* `cond-mat/0208280` (2002)

[17] M. Mézard, G. Parisi *Bethe lattice spin glass revisited*, Eur. Phys. J. **B 20**, 217 (2001) `cond-mat//0009418`

[18] M. Mézard, R. Zecchina, Phis. Rev. **E 66**, 056126 (2002)

[19] T. Castellani, V. Napolitano, F. Ricci-Tersenghi, R. Zecchina, *Bicolouring Random Hypergraphs,* J: Phys. **A 36** 11037 (2003)

[20] A. Braunstein, R. Mulet, A. Pagnani, M. Weigt, R. Zecchina *Polynomial iterative algorithms for coloring and analyzing random graphs,* Phys. Rev. **E 68**, 036702 (2003)

[21] M. Garey, D.S. Johnson, *Computers and Intractability. A Guide to the Theory of NP-Completeness* W.H. Freeman and Co., San Francisco (1979).

[22] P. Bürgisser, *Completeness and reduction in algebraic complexity theory*, Berlin Springer-Verlag, (2000)

[23] S.A. Cook, *The complexity of theorem-proving procedures,* in Proc. 3rd Ann. ACM Symp. on Theory of Computing, 151-158. ACM, New York (1971)

[24] A. Braunstein, M. Mézard, R. Zecchina, *Survey propagation: an algorithm for satisfiability* `cs.CC/0212002` (2002)

[25] R.G. Gallager, *Low Density Parity Check Codes,* Monograph, M.I.T. Press (1963)

[26] A. Montanari, *The glassy phase of Gallager codes,* `cond-mat/0104079`

[27] V. Chvatal, E. Szemeredi, *Many hard examples for resolution,* JACM, **35**, 759 (1988)

[28] D.J. Aldous, *The $\zeta(2)$ limit of Random Assignment Problem,* Random Structures Algorithms **18**, no 4, 381-418, (2001). `math.PR/0010063`

[29] C. Nair, B. Prabhakar, M. Sharma, *Proof of the Parisi and Coppersmith-Sorkin Conjectures for the Finite Random Assignment Problem,* FOCS 2003, on web at `http://www.stanford.edu/~mchandra/`

[30] Haijun Zhou, Zhong-can Ou-Yang *Maximum matching on random graphs,* `cond-mat/0309348` (2003)

[31] D. Coppersmith, G. B. Sorkin, *On the Expected Incremental Cost of a Minimum Assignment,* Bela Bollobas ed., Contemporary Combinatorics, Bolyai Society Mathematical Studies **10** chapter 8, Springer (2002)

[32] G. Parisi, M. Ratieville, *On the finite size corrections to some random matching problems,* `cond-mat/0204595`

[33] W.W. Garvin, *Introduction to Linear programming,* McGrow-Hill Book Company, (1960)

[34] H.W. Kuhn, *The hungarian method for solving the assignment problem,* Naval Research Logistics Quarterly, **36**, 2:83-97 (1955)

[35] R.M. Karp, *Reducibility among combinatorial problems,* Complexity of Computer Computations, Miller and Thatcher, eds., Plenum Press, New York, 85-104 (1972)

[36] D. Shah, P. Giaccone, B. Prabhakar, *An Efficient Randomized Algorithm for Input-Queued Switch Scheduling,* IEEE Hot Interconnects 9, Stanford, CA, USA, (2001)
P. Giaccone, *Queueing and scheduling algorithms for high performance routers,* Ph.D. dissertation, available at `http://www.tlc-networks.polito.it/giaccone/all.shtml` (2002)

[37] A.G. Percus, O.C. Martin, *The stochastic traveling salesman problem: Finite size scaling and the cavity prediction,* Journal of Statistical Physics **94**, 739-758 (1999)

[38] P.G. DeGennes, *Exponents for the excluded volume problem as derived by the Wilson method,* Phys. Lett., **38A**, 339-340 (1972)

[39] D. Ruelle, *A mathematical reformulation of Derrida's REM and GREM,* Math. Phys. **108**, pp. 225-239 (1987)

[40] Yu.A. Kutoyants, *Statistical inference for spatial poisson processes*, Lecture notes in statistics – 0134, Springer-Verlag, New York (1998)