Aspects of Statistical Physics in Computational Complexity

Stefano Gogioso
Quantum Group, Department of Computer Science
University of Oxford, UK
stefano.gogioso@cs.ox.ac.uk
24 June 2013

Abstract
The aim of this review paper is to give a panoramic of the impact of spin glass theory and statistical physics in the study of the K-sat problem, as summarised by the words of Amin Coja-Oghlan ([2], Warwick 2010)

“Random K-sat is a spin glass problem (with a combinatorial flavor)”.

The introduction of spin glass theory in the study of the random K-sat problem has had profound effects on the field, leading to some groundbreaking descriptions of the geometry of its solution space and helping to shed light on why it seems to be so hard to solve. Most of the geometrical intuitions have their roots in the Sherrington-Kirkpatrick model of spin glass: its simple formulation and complex free-energy landscape make it the ideal place to start our exploration of the statistical physics of random K-sat.

We’ll start Chapter 2 by introducing the SK model from a mathematical point of view, presenting some rigorous results on free-entropy density and factorisation of the Gibbs measure and giving a first intuition about the cavity method. We’ll then switch to a physical perspective and start exploring concepts like pure states, hierarchical clustering and replica symmetry breaking in the sandbox provided by the SK model.

Chapter 3 will be devoted to the spin glass formulation of K-sat: we’ll introduce factor graphs, draw the connection between pure states and clusters of solutions, and define the complexity. The most important phase transitions of K-sat (clustering, condensation, freezing and SAT/UNSAT) will be extensively discussed in Chapter 4 with respect their complexity, free-entropy density and the so-called Parisi 1RSB parameter: rigorous results and physically-inspired analysis will blend together to give as much intuition as possible about the geometry of the phase space in the various regimes, with a special focus on clustering and condensation.

The so-called algorithmic barrier will be presented in Chapter 5 and exemplified in detail on the Belief Propagation (BP) algorithm. The BP algorithm will be introduced and motivated, and numerical analysis of a BP-guided decimation algorithm will be used to show the role of the clustering, condensation and freezing phase transitions in creating an algorithmic barrier for BP.

Taking from the failure of BP in the clustered and condensed phases, Chapter 6 will finally introduce the Cavity Method to deal with the shattering of the solution space, and present its application to the development of the Survey Propagation algorithm.
## Contents

1 Introduction
   1.1 The K-sat problem ................................................. 4
   1.2 Statistical mechanics ............................................ 5
   1.3 The Hamiltonian for random K-sat ............................... 6

2 Spin glass fundamentals
   2.1 The SK model, for mathematicians ............................... 8
      2.1.1 Formulation ............................................... 8
      2.1.2 The free-entropy density: replica symmetric (RS) case .... 9
      2.1.3 A first look at the cavity method ......................... 11
   2.2 The SK model, for physicists ................................... 12
      2.2.1 Pure state ............................................... 12
      2.2.2 Pure state overlaps ...................................... 12
      2.2.3 Ultrametricity and hierarchical clustering ............... 13
      2.2.4 The replica method ...................................... 16
      2.2.5 Replica symmetric (RS) solution ......................... 17
      2.2.6 Replica symmetry breaking (RSB) ......................... 18
      2.2.7 The $n \to 0$ limit in RSB (a.k.a. the messy part) ....... 19

3 From SK to K-sat .................................................. 20
   3.1 Diluted SK and the K-sat Hamiltonian ......................... 20
   3.2 Factor graphs ................................................. 21
   3.3 Pure states and clusters of solutions ......................... 23
   3.4 The complexity .............................................. 24

4 The phases of K-sat ................................................ 25
   4.1 The Parisi 1RSB parameter ..................................... 25
   4.2 Replica symmetry breaking ..................................... 26
   4.3 Phase transitions in K-sat .................................... 27
   4.4 Clustering: the Dynamical phase transition RS $\to$ d1RSB .... 28
   4.5 Condensation phase transition d1RSB $\to$ 1RSB ............... 30
   4.6 Freezing phase transition .................................... 30
   4.7 SAT/UNSAT phase transition .................................. 33

5 Algorithms for K-sat ............................................. 35
   5.1 The algorithmic barrier ....................................... 35
   5.2 Belief propagation ............................................ 36
      5.2.1 An typical example of BP ................................ 36
      5.2.2 The BP algorithm for tree factor graphs ............... 41
      5.2.3 The BP algorithm for general factor graphs .......... 40
      5.2.4 The BP algorithm for K-sat ............................. 41
   5.3 BP guided decimation ........................................ 42
      5.3.1 BP guided decimation algorithm ......................... 42
      5.3.2 Frozen variables and success probability for BP ....... 43
| Section | Title                                                                 | Page |
|---------|----------------------------------------------------------------------|------|
| 5.3.3   | A condensation phase transition for the residual free-entropy density | 44   |
| 5.3.4   | The algorithmic barrier for BP                                       | 49   |
| 5.4     | From BP to SP                                                        | 50   |
| 6       | The Cavity Method and Survey Propagation                             | 51   |
| 6.1     | The Cavity Method                                                    | 51   |
| 6.1.1   | From messages to surveys                                             | 52   |
| 6.1.2   | The cavity method: free-entropy density, overlaps and marginals      | 53   |
| 6.2     | Survey Propagation                                                   | 54   |
| 6.2.1   | Encoding the surveys                                                 | 54   |
| 6.2.2   | The SP algorithm                                                     | 54   |
| 6.2.3   | Convergence of SP                                                    | 55   |
| 7       | Conclusion                                                           | 55   |
1 Introduction

1.1 The K-sat problem

The K-satisfiability problem, better known as K-sat, asks whether one can satisfy $M$ given constraints over $N$ boolean variables, i.e. whether there is an assignment $b$ of boolean values to the $N$ variables that satisfies a given K-CNF $I$ (the assignment is then called a satisfying assignment, or a solution for $I$). A CNF is a boolean formula involving only connectives $\lor$, $\land$ and $\lnot$ and written in conjunctive normal form, i.e. as the conjunction of many clauses, each clause being the disjunction of many literals (a literal is either a variable $x_i$ or the negation $\lnot x_i$ of a variable). A K-CNF is a CNF where each clause has exactly K literals.

Despite the simplicity of its formulation, there is no known (deterministic) algorithm which can solve the K-sat problem in polynomial time. There are, on the other hand, efficient solvers that can find solutions with high probability for reasonably low constraint density $\alpha \triangleq M/N$: an understanding of the properties of these algorithms goes through an understanding of statistical properties of the K-sat solution space.

We call random K-sat the (conceptual) variant of the K-sat problem where the instance $I$ is allowed to be a random variable $\xi$ (and thus so is its solution space), governed by some probability measure over the space of instances of K-sat. Two commonly used measures are

(a) the uniform model: the instance $I$ is chosen with the uniform probability in the space of instances of K-sat (and the uniform probability measure is introduced on its solution space).

(b) the planted model: the issue with the uniform model is that the instance could be unsatisfiable, so one first choses a random assignment $b$ of boolean values to the variables, and then chooses $M$ clauses satisfied by it, uniformly over all such clauses.

The two models induce different measures on the solution space $\{(I, b) \text{ s.t. } b \text{ satisfies } I\}$: the former yields immediate results on algorithms but is very hard to work with, while the second is easy to work with but hard to connect to algorithms.

The idea behind the key paper [6] (which will prove the existence of a clustering phase transition in the solution space of $I$) is to work in the planted model (where the slightly higher abundance of solution-rich instances allows a successful use of the so-called 2nd moment method), and then transfer the results to the uniform model via the following

**Theorem 1.1.** (Transfer theorem)

There is a sequence $\xi_K \rightarrow 0$ s.t. if a property holds with probability $1 - \exp[-\xi_K N]$ in the planted model then it holds with high probability (see below) in the uniform model.

When talking about an event $E(I)$ depending on a random instance $I \equiv I(N, M)$ of K-sat, we’ll say that $E(I)$ happens with high probability (w.h.p.) iff

$$\lim_{N \rightarrow \infty, M/N \rightarrow \alpha} \mathbb{P}(E(I)) = 1.$$

\footnote{From now on we’ll write r.v. for random variable.}
1.2 Statistical mechanics

Unless otherwise stated, this section is based on [13].

In statistical physics one considers systems of many identical particles: the focus is on the ensemble properties of the system, i.e. in the statistical properties of the system as a whole, ignoring the details of the single microstates. As the number $N$ of particles is big, we’ll usually be working in the so-called thermodynamic limit $N \to \infty$, and we’ll often write $f(N) \approx g(N)$ for $\lim_{N \to \infty} f(N) = \lim_{N \to \infty} g(N)$.

Our starting point is the definition of a temperature $T > 0$ for our system, and a much more useful inverse temperature $\beta \overset{\text{def}}{=} \frac{1}{K_B T}$: high temperature corresponds to $\beta \ll 1$ while low temperature corresponds to $\beta \gg 1$, the zero-temperature limit being $\beta \to \infty$.

The system will have a Hamiltonian $H$, an operator on the space of states describing the energy $E_\sigma$ of each state $\sigma$ of the system: $H(\sigma) = E_\sigma$. We define the following probability measure on the space of states, called the Gibbs measure:

$$
\mu(\sigma) \overset{\text{def}}{=} \frac{1}{Z} \exp[-\beta H(\sigma)]
$$

where $Z$ is called the partition function for the system, and $\sum_\sigma$ stands for the sum with $\sigma$ ranging over the full space of states.

The quantity $\log Z$, called the log-partition function is fundamental, as it allows to reconstruct all the moments of the energy of the system by taking derivatives w.r.t. $\beta$. For example mean and variance of $E$ are reconstructed as

$$
\langle E \rangle = \sum_\sigma \mu(\sigma) H(\sigma) = -\frac{\partial}{\partial \beta} \log Z
$$

$$
\text{Var}[E] = -\frac{\partial}{\partial \beta} \langle E \rangle = \frac{\partial^2}{\partial \beta^2} \log Z
$$

The log-partition function is also connected to the free-energy $^4$ by $F = -\frac{1}{\beta} \log Z$.

An ubiquitous example is provided by two state systems (also called spin-$\frac{1}{2}$) systems, and we’ll refer to such system as spins. The state space of a spin is given by its two possible configurations $\{\pm\}$; its Hamiltonian is $H(\pm) = \mp \epsilon$, and the Gibbs measure becomes

$$
\mu(\pm) = \frac{1}{\exp[+\beta \epsilon] + \exp[-\beta \epsilon]} \exp[\pm \beta \epsilon] = \frac{\exp[\pm \beta \epsilon]}{2 \cosh[\beta \epsilon]}
$$

$^2$We’re usually talking about $N \approx 10^{23}$ particles, plus or minus a handful of orders of magnitude.

$^3$Where $K_B$ is the Boltzmann constant, making $\frac{1}{\beta}$ a measure of energy.

$^4$Representing the amount of energy in a system that can be used to do physical work.
Indeed the first, humble step in the application of statistical physics to the K-sat problem will be realising that, at the end of the day, a boolean variable is nothing but a spin.

In what follows we’ll be interested in systems of many interacting\footnote{As opposed to many free spins, where the Gibbs measure is given by the normalised product of the individual spin measures.} spins $\sigma_1, \ldots, \sigma_N$. One iconic system of interacting spins is the Ising model:

$$H(\sigma) = -\frac{1}{\sqrt{N}} \sum_{i<j} g_{ij} \sigma_i \sigma_j - \sum_i h_i \sigma_i$$  \hspace{1cm} (1.4)

where the spins lie on a $\mathbb{Z}^d$ lattice and the spin couplings $g_{ij}$ are non-zero only for $i, j$ nearest neighbours on the lattice, and have the same value over all the nearest neighbour pairs. The Ising model is the simplest model for magnetic systems which exhibits non-trivial behaviour:

1. the spins correspond to magnetic orientations along some fixed axis;
2. the couplings correspond to the ferromagnetic ($g_{ij} > 0$, i.e. tendency to align) or antiferromagnetic ($g_{ij} < 0$, i.e. tendency to antialign) interaction of nearby spins;
3. the external fields $h_i$ correspond to external magnetic fields that influence the single spins (in the classical model $h_i = h$ has the same value for all spins, i.e. there is a uniform external field).

We’ll be interested in the much more complex case of spin glasses, where we’ll allow couplings and fields to become random variables.

1.3 The Hamiltonian for random K-sat

Given an instance of K-sat $I = \bigwedge_{a=1}^M C_a$ and an assignment $x_1 = b_1, \ldots, x_n = b_N$ of boolean values to the variables, a natural way to define the energy of an assignment $b$ is

$$H_N(b) = \# \text{ of clauses of } I \text{ violated by } b$$  \hspace{1cm} (1.5)

Thus minimising the energy is equivalent to minimising the number of violated clauses: if the instance is satisfiable, the satisfying assignments will be exactly those with zero energy.

The first step towards spin glasses is to go from boolean assignments $b = (b_1, \ldots, b_N) \in \{0, 1\}^N$ to spin configurations $\sigma = (\sigma_1, \ldots, \sigma_N) \in \{\pm 1\}^N$: the correspondence is a matter of convention, and we’ll take it to be $b_i = 0, 1 \leftrightarrow \sigma_i = -1, +1$. From now on we’ll treat the K-sat problem as if it was formulated in terms of spins: when talking of variable $i$ or $x_i$ we’ll usually refer to the index of the spin, when talking of spin $\sigma_i$ we’ll be talking about the r.v. encoding the boolean assignment of a variable.

Given clause $C_a = z_{i(a, 1)}, \ldots, z_{i(a, K)}$, where $z_i \in \{x_i, \neg x_i\}$, we’re interested in having an indicator function $W_a(\sigma)$ for its violation (dependence on $N$ is kept implicit):

$$W_a(\sigma) \overset{\text{def}}{=} \begin{cases} 1, & \text{if } \sigma \text{ violates clause } C_a \\ 0, & \text{if } \sigma \text{ satisfies clause } C_a \end{cases}$$  \hspace{1cm} (1.6)
Define the spins $J_a = (J^1_a, \ldots, J^K_a)$ to be

$$J^r_a \overset{\text{def}}{=} \begin{cases} +1, & \text{if } z_i(a,r) = \neg x_i(a,r) \\ -1, & \text{if } z_i(a,r) = x_i(a,r) \end{cases}$$ (1.7)

i.e. $J^r_a$ is the spin that makes $z_i(a,r)$ false. Then $W_a(\sigma) = 1 \iff$ each spin $\sigma_i(a,r)$ of $\sigma$ involved in clause $C_a$ falsifies its the corresponding literal $z_i(a,r) \iff$ each spin $\sigma_i(a,r)$ is aligned with $J^r_a$:

$$W_a(\sigma) = \prod_{r=1}^{K} \frac{1 + J^r_a \sigma_i(a,r)}{2}$$ (1.8)

The Hamiltonian for instance $\mathcal{I}$ is then

$$H_N(\sigma) = \sum_{a=1}^{M} W_a(\sigma)$$ (1.9)

which is a random function of $\sigma$ since $\mathcal{I}$ is a random instance of K-sat.

The partition function is

$$Z_N = \sum_{\sigma} \exp \left[ -\beta \left( \text{# of clauses of } \mathcal{I} \text{ violated by } \sigma \right) \right]$$ (1.10)

Thus in the zero temperature limit $\beta \to \infty$ only the satisfying assignments contribute to the sum and we have the following neat way of counting them

$$Z_N|_{\beta=\infty} = \# \text{ of satisfying assignments for } \mathcal{I}$$ (1.11)

Indeed at zero temperature the Gibbs measure is concentrated on the satisfying assignments only.
2 Spin glass fundamentals

2.1 The SK model, for mathematicians

Except where otherwise stated, this section is based on [3][4].

2.1.1 Formulation

Consider the spin space of an $N$-spin system

\[ \Sigma_N \overset{\text{def}}{=} \{-1, +1 \}^N \]  

equipped with the Hamming distance

\[ d(\sigma, \tau) \overset{\text{def}}{=} \#\{1 \leq i \leq N \text{ s.t. } \sigma_i \neq \tau_i\} \]  

Consider also a family of iid standard gaussian r.v.s (the spin couplings, or the disorder)

\[ \{g_{ij}\}_{1 \leq i < j \leq N} \text{ with } \mathbb{E} g_{ij} = 0 \] and \[ \text{Cov}[g_{ij}, g_{i'j'}] = \mathbb{E} g_{ij} g_{i'j'} = \delta_{ii'}\delta_{jj'} \]  

and a vector $h \in \mathbb{R}^N$ (the components are the external fields for the spins).

The Hamiltonian $H_N$ of the SK model is the following random function over $\Sigma_N$

\[ H_N(\sigma) \overset{\text{def}}{=} -\frac{1}{\sqrt{N}} \sum_{i<j} g_{ij} \sigma_i \sigma_j - \sum_i h_i \sigma_i \]  

Eq'n 2.4 shows how the model assigns energy to spin configurations:

- An individual spin contributes to a lower system energy when aligned with its external field
- A pair of spins $\sigma_i, \sigma_j$ with positive coupling (i.e. $g_{ij} > 0$) contribute to a lower system energy when aligned (i.e. $\text{sgn}[\sigma_i] = \text{sgn}[\sigma_j]$)
- A pair of spins $\sigma_i, \sigma_j$ with negative coupling (i.e. $g_{ij} < 0$) contribute to a lower system energy when anti-aligned (i.e. $\text{sgn}[\sigma_i] \neq \text{sgn}[\sigma_j]$)
- A pair of spins with null coupling (i.e. $g_{ij} = 0$) does not contribute to the system energy

The SK model introduces not one but two sources of randomness on the spin system:

(a) the space $\Sigma_N$ is endowed with a Gibbs measure

\[ \mu_N(\sigma) \overset{\text{def}}{=} \frac{1}{Z_N} \exp[-\beta H_N(\sigma)] \]  

where $Z_N \overset{\text{def}}{=} \sum_{\sigma} \exp[-\beta H_N(\sigma)]$ is the partition function as usual.

(b) the measure $\mu_N$ is itself random, as it depends on the $\frac{N(N-1)}{2}$ r.v.s $\{g_{ij}\}_{ij}$. 

8
Given a function $f$ on $\Sigma_N$, we’ll use the following notation (the brackets) for the expectation taken over the Gibbs measure
\[
\langle f \rangle \overset{\text{def}}{=} \frac{1}{Z_N} \sum_{\sigma} f(\sigma) \exp[-\beta H_N(\sigma)]
\] (2.6)

Then $\langle f \rangle$ is still a random variable, depending on the random couplings. On the other hand we’ll denote the expectation taken over the couplings by the usual $E$ symbol, and we’ll end up writing $E\langle f \rangle$ for the full expectation over all sources of randomness. Please note that $E f$ is also a random variable, this time depending on the random spin configuration (governed by the Gibbs measure and usually denoted by $\sigma$).

When dealing with several spin configurations at a time we’ll assume to have at our disposal a sequence of iid r.v.s $(\sigma_k)_{k \geq 1}$, the replicas, all governed by the Gibbs measure of eq’n 2.5.

Replicas are primarily used to linearise products of brackets, as in
\[
\langle g \rangle^n = \langle g(\sigma^1) \cdot \ldots \cdot g(\sigma^n) \rangle
\] (2.7)

where we’ve extended brackets in the obvious way
\[
\langle f(\sigma^1, \ldots, \sigma^n) \rangle \overset{\text{def}}{=} \frac{1}{(Z_N)^n} \sum_{\sigma^1, \ldots, \sigma^n} f(\sigma^1, \ldots, \sigma^n) \exp[-\beta \sum_{1 \leq k \leq n} H_N(\sigma^k)]
\] (2.8)

Despite this linearisation of brackets, the energies of different spin configurations are not independent: their covariance is given by
\[
\text{Cov}[H_N(\sigma^1), H_N(\sigma^2)] = \frac{1}{N} \sum_{i < j} \sigma^1_i \sigma^1_j \sigma^2_i \sigma^2_j = \frac{N}{2} R^2(\sigma^1, \sigma^2) - \frac{1}{2}
\] (2.9)

where $R(\sigma^1, \sigma^2)$ is the overlap of the two spin configurations
\[
R(\sigma^1, \sigma^2) \overset{\text{def}}{=} \frac{1}{N} \sum_i \sigma^1_i \sigma^2_i = 1 - 2d(\sigma^1, \sigma^2)
\] (2.10)

Because all $R(\sigma_i, \sigma_j)$ for $i \neq j$ have the same distribution, when under expectation (and/or not concerned with which pair of spins we choose) we’ll usually just write $R$ for the overlap.

### 2.1.2 The free-entropy density: replica symmetric (RS) case

As we’ve seen in section 1.2, the log-partition function $\log Z_N$ gives us a lot of information about the system (e.g. the energy distribution or, in the case of K-sat, the distribution of the number of violated constraints). In the SK model it is a random variable, and we also consider its quenched average $\overset{6}{7}$ (also known, in the thermodynamic limit, as the free-entropy density)
\[
\rho_N \overset{\text{def}}{=} \frac{1}{N} E \log Z_N
\] (2.11)

---

6 As opposed to the annealed average $\overset{4}{5} \log E Z_N$, which is much easier to compute.
7 To get some physical intuition, notice that the quantity $\frac{1}{2} \rho_N$ is the expected free energy per spin.
Physical considerations (similar to *equipartition of energy*) suggest that the r.v. \( \frac{1}{N} \log Z_N \) should be self-averaging in the thermodynamic limit \(^7\) and we expect the quenched average to capture most of the information the log-partition function gave us in the classical non-random cases.

Since \( p_N \) captures so much information about the system, we expect its exact computation to be far from trivial (especially in the low temperature regime where spin glasses are characterised by a complex free-energy landscape): the quest for its value has given birth to some of the most interesting tools of spin glass theory, like the replica method and the cavity method. The rest of this section 2.1 is dedicated to a concise survey of the main results on the free-entropy density of the SK model.

**Theorem 2.1. (Guerra’s replica-symmetric bound)**

For all \( \beta, q > 0, \) and any choice of law for the external fields \( h \) (assume them iid), we have

\[
p_N(\beta, h) \leq \log 2 + \mathbb{E} \log \cosh[\beta(\sqrt{q} + h)] + \frac{\beta^2}{4} (1 - q)^2
\]

(2.12)

The RHS is usually denoted \( SK(\beta, q, h) \), and \( g \) is a r.v. distributed like the spin couplings.

Optimising over \( q \) one obtains the tightest possible bound for

\[
q = \mathbb{E} \tanh^2[\beta(\sqrt{q} + h)]
\]

(2.13)

A theorem by Latala and Guerra then guarantees the existence of a unique such \( q \), and from now on we’ll consider that specific value (and we’ll write \( SK(\beta, h) \)).

The overlaps can be shown, in the high-temperature region \( \beta < \frac{1}{2} \), to converge in distribution to the \( q \) of 2.13, and this implies the convergence of \( p_N(\beta, h) \) to \( SK(\beta, h) \) in the same high-temperature region \(^9\).

**Theorem 2.2. (Guerra’s replica-symmetric solution)**

Assume \( \beta < \frac{1}{2} \) and any choice of law for the external fields \( h \), then

\[
|p_N(\beta, h) - SK(\beta, h)| \leq O(N^{-1})
\]

(2.14)

The constants hidden by the big-O notation are allowed to depend on \( \beta \) and the law of \( h \). In fact the authors prove convergence in distribution, not just in first moment.

Control of the overlaps (the variance of which bounds the correlation of finite sets of spins) also allows to show that the Gibbs measure fully factorises in the thermodynamic limit:

**Theorem 2.3. (Factorisation of the Gibbs measure)**

Assume \( \beta < \frac{1}{2} \), and let \( \mu_{N,p} \) be the marginal distribution of \((\sigma_1, ..., \sigma_p)\) under \( \mu_N \). Then we have

\[
\mathbb{E} ||\mu_{N,p} - \nu_p||^2 \leq O(N^{-1})
\]

(2.15)

where \( ||\mu - \nu|| \overset{\text{def}}{=} \sum_{s \in \Sigma_p} |\mu(s) - \nu(s)| \) is the total variational distance, and \( \nu_p \) is the product probability measure on \( \Sigma_p \):

\[
\nu_p(s_1, ..., s_p) \overset{\text{def}}{=} \frac{1}{2^p} \prod_{i \leq p} (1 + \langle \sigma_i \rangle s_i)
\]

(2.16)

---

\(^8\)i.e. its fluctuations around its expected value \( p_N \) should become small, with variance falling as \( \frac{1}{N} \): this is the case for the SK model, as shown by Theorem 2.2.

\(^9\)A different solution is needed for the low temperature region.
2.1.3 A first look at the cavity method

The cavity method is arguably the most important tool in spin glass theory: its applications encompass all the results we’ve seen and a good share of the ones to come. Here we introduce its basic formulation in the SK model, and in section 6 we’ll develop its full potential within the factor graph formalism.

In short, the cavity method is induction over $N$ for Gibbs averages: it reduces the computation of brackets for an $N$ spin system to that for an $N - 1$ spin system by

1. fixing the value of a spin $\sigma_N$ to +1
2. removing the spin from the system, thus creating a cavity and raising the temperature
3. measuring the Gibbs average in the smaller system
4. taking the mean of the two Gibbs averages obtained in (1a)-(3b)

Mathematically this is the observation that Hamiltonian $H_N$ can be written as that of a smaller $N - 1$ spin system plus a term coupling the smaller system with the last spin $\sigma_N$:

$$H_N(\sigma) = \frac{\sqrt{N-1}}{\sqrt{N}} H_{N-1}(\sigma_1, \ldots, \sigma_{N-1}) + \alpha_N \sigma_N$$  \hspace{1cm} (2.17)

where the coupling $\alpha_N$ is given by

$$\alpha_N \triangleq \frac{1}{\sqrt{N}} \sum_{i<N} g_{iN} \sigma_i + h_N$$  \hspace{1cm} (2.18)

Absorbing the $\frac{\sqrt{N-1}}{\sqrt{N}}$ factor into the temperature of the $N - 1$ spin system, denoting by $Av$ the average over $\sigma_N = \pm 1$, and expanding the brackets $\langle \cdot \rangle$ of the original system in terms of the brackets $\langle \cdot \rangle_-$ of the smaller system, we get the desired result

$$\langle f \rangle = \frac{\langle Av \left( f(\sigma) \exp[\beta \alpha_N \sigma_N] \right) \rangle_-}{\langle Av \left( \exp[\beta \alpha_N \sigma_N] \right) \rangle_-}$$  \hspace{1cm} (2.19)

---

10 Raising is good, as our results tend to hold from a given temperature up, rather than down.
2.2 The SK model, for physicists

Except where otherwise stated, this section is based on [20][5].

In order to investigate the low temperature region (where the replica-symmetric solution breaks), we’ll continue our survey of the SK model with a shift in perspective: we’ll put mathematical rigour aside, and try to gain some intuition about replica symmetry (and replica symmetry breaking) through the eyes of a physicist.

From now on we’ll assume no external field is present in our model, and throughout this section we’ll use the following simplified Hamiltonian

\[ H_{N}(\sigma) \overset{\text{def}}{=} -\frac{1}{\sqrt{N}} \sum_{i \neq j} g_{ij} \sigma_{i} \sigma_{j} \tag{2.20} \]

2.2.1 Pure state

In analogy with quantum mechanics, physicists refer to ergodic components of the Gibbs measure as pure states. Suppose the Gibbs measure can be decomposed as

\[ \mu_{N}(\sigma) = \sum_{\psi} \omega_{\psi} \mu_{N}^{(\psi)}(\sigma) \tag{2.21} \]

where \( \mu_{N}^{(\psi)} \) are probability measures on \( \Sigma_{N} \). Then the \( \mu_{N}^{(\psi)} \) are pure states iff the following holds:

(a) the \( \mu_{N}^{(\psi)} \) satisfy cluster decomposition: for any set of local functions \( A_{1}, \ldots, A_{p} \) we have that \( \langle A_{1} \cdot \ldots \cdot A_{p} \rangle_{\psi} \approx \langle A_{1} \rangle_{\psi} \cdot \ldots \cdot \langle A_{p} \rangle_{\psi} \) in the large \( N \) limit \(^{12}\), where \( \langle \cdot \rangle_{\psi} \) is the expectation taken w.r.t. measure \( \mu_{N}^{(\psi)} \). This property is also called correlation-decay property, and it can be rephrased as large distance implies vanishing correlation.

(b) the Gibbs probabilities \( \omega_{\psi} \) are all positive.

(c) the \( \mu_{N}^{(\psi)} \) themselves cannot be further decomposed in measures satisfying condition (a)-(b).

Sets of positive measure for a pure state have measure zero for all the other pure states. Also, due to theorem 2.3, the Gibbs measure for the SK model in the high temperature region \( \beta < \frac{1}{2} \) only has one pure state.

2.2.2 Pure state overlaps

The notion of overlap is now transferred to pure states, where the overlap \( q_{\phi\psi} \) of two pure states \( \phi \) and \( \psi \) is defined as:

\[ q_{\phi\psi} \overset{\text{def}}{=} \frac{1}{N} \sum_{i} \langle \sigma_{i} \rangle_{\phi} \langle \sigma_{i} \rangle_{\psi} \tag{2.22} \]

\(^{11}\)i.e. only involving a finite number of spins.

\(^{12}\)Technically a remainder \( r(\sigma_{i_{1}}, \ldots, \sigma_{i_{r}}) \) (which will be a function of the spins \( \sigma_{i_{1}}, \ldots, \sigma_{i_{r}} \) involved in the expression) is allowed on the RHS of the approx equation, as long as its magnitude vanishes in the large \( N \) limit when averaged over all the spins, i.e. as long as \( \lim_{N \to \infty} \frac{1}{N} \sum_{i_{1}, \ldots, i_{r}} |r(\sigma_{i_{1}}, \ldots, \sigma_{i_{r}})| = 0 \). We’ll not mention this \( r \) anywhere else.
Then in particular $q_{\psi \phi} = \frac{1}{N} \sum_i \langle \sigma_i \rangle^2_\phi$ is the *average magnetisation* in pure state $\psi$. It turns out \cite{32} that $q_{\psi \phi}$ is independent of both $\psi$ and the spin couplings: it is thus an invariant of the system depending only on $\beta$, and is sometimes called the Edwards-Andersen order parameter $q_{EA}$.

One can study the distribution of the pure state overlaps by considering the following **pure state overlap probability density function (PDF)**

$$P(q') \overset{df}{=} \sum_{\phi, \psi} \omega_{\psi} \omega_{\phi} \delta(q' - q_{\phi \psi}) = \langle \delta(q' - \frac{1}{N} \sum_i \sigma_i^1 \sigma_i^2) \rangle$$

(2.23)

Analytical computation of $P(q)$ at fixed disorder is in general not possible, but the cavity method allows computation of the disorder-averaged overlap PDF $\mathbb{E} P(q)$.

### 2.2.3 Ultrametricity and hierarchical clustering

A key result comes from studying the following **overlap triangle PDF**

$$P(q_1, q_2, q_3) \overset{df}{=} \sum_{\phi, \psi, \varphi} \omega_{\psi} \omega_{\phi} \omega_{\varphi} \delta(q_1 - q_{\phi \psi}) \delta(q_2 - q_{\phi \varphi}) \delta(q_3 - q_{\psi \varphi})$$

(2.24)

Using replicas and restricting to positive pure state overlaps this can \cite{33} be written as

$$P(q_1, q_2, q_3) = \frac{1}{2} \int_0^{q_1} P(q') dq' P(q_1) \delta(q_1 - q_2) \delta(q_2 - q_3) + \frac{1}{2} \left[ P(q_1) P(q_2) \theta(q_1 - q_2) \delta(q_2 - q_3) + \text{two more permutations} \right]$$

(2.25)

where $\theta(x)$ is the Heaviside step function. The first line of eq’n \ref{2.25} makes $(q_1, q_2, q_3)$ into the sides of an equilateral triangle, while the three terms on the second line cover the case of non-equilateral, isosceles triangles. If we now define a metric on pure states by

$$d(\phi, \psi) \overset{df}{=} \frac{1}{N} \sum_i (\langle \sigma_i \rangle_\phi - \langle \sigma_i \rangle_\psi)^2 = 2(q_{EA} - q_{\phi \psi})$$

(2.26)

then eq’n \ref{2.25} implies that under $d$ all triangles are either isosceles or equilateral, i.e. that the space of pure states exhibits **ultrametricity**.

Ultrametricity is ubiquitous in biology, where it appears in all metrics based on some concept of *most recent common ancestor* \cite{13}, and in fact ultrametricity is equivalent to the possibility of clustering objects into a hierarchical tree. The ultrametric nature of the space of pure states is most evident in figure 1 (p. 15), coming from the yet-to-appear \cite{35} and obtained as follows.

1. 100 spin configurations (of a SK model with $N = 800$ spins) are sampled 100 times each with a Monte Carlo method and then averaged: the Monte Carlo method starts the system at a high temperature $T = 1.2 T_c$ and cools it down to low temperature $T = 0.1 T_c$.

\(^{13}\text{e.g. distance between species, difference of DNA or representative sets of proteins, evolutionary trees, etc.}\)

\(^{14}\text{I.e. above the critical temperature } T_c \text{ corresponding to the so called spin glass phase transition, where the replica-symmetric solution breaks.}\)
2. each configuration of spin averages is taken to be representative of the $(\langle \sigma_i \rangle_{\psi})_{i=1,...,N}$ for some pure state $\psi$ of the system (so to effectively sample 100 pure states from the system). The overlaps of the pure states are then computed and clustered. The 100x100 heatmaps show the overlaps, with red indicating maximum positive overlap $q_{\phi \psi} = +1$, blue indicating maximum negative overlap $q_{\phi \psi} = -1$ and white indicating no overlap $q_{\phi \psi} = 0$.

3. the dendrogram plots on top of the heatmaps show the clustering of the samples, with vertical height being proportional to temperature. Each pair of pure states $\phi, \psi$ then defines a specific temperature (or time) in the cooling (or evolution) of the system, the temperature of their most-recent common ancestor in the dendrogram. The difference $\tau(\phi, \psi)$ between that temperature and the final temperature $T_f = 0.1 \cdot T_c$ is then a metric on the space of pure states, itself showing ultrametricity 

Figure 1 (p.15) shows snapshots at $T/T_c = 1.2, 0.86, 0.55, 0.12$ of the evolution of the system:

1. At high temperature $T = 1.2 \cdot T_c$ the pure states show little or no overlap, and all lie in a single big cluster.

2. Having just crossed the critical temperature, at $T = 0.86 \cdot T_c$ the pure states start showing visible overlaps and shatter in two recognisable clusters (the bigger of which features two barely-recognisable subclusters).

3. At about half the critical temperature ($T = 0.55 \cdot T_c$) the overlaps continue to increase and the pure states are clustered into two sharply separated clusters, each of which features two clearly recognisable subclusters.

4. At very low temperature $T = 0.12 \cdot T_c$ the pure states are organised into 8 sharp clusters, hierarchically organised in what is to good approximation a binary tree.

For two states $\phi, \psi$ in different clusters, $\tau(\phi, \psi)$ is the temperature below which the two clusters shatter and rapidly become separated by impassably high energy barriers.

---

15 It is closely related to the original metric $d$, but the details of the relation depend on the clustering method used. Unless differently stated we’ll assume the metric on the space of pure states to be $d$.

16 With the exception of 9 isolated states.

17 A rough way to obtain a hierarchical tree from the dendrogram tree is by fixing a temperature difference $\Delta T$, discarding all the isolated states and progressively merging all nodes of the dendrogram tree that are nearer in temperature than $\Delta T$, in order of increasing difference of temperature. This procedure applied to the $T = 0.12 \cdot T_c$ dendrogram tree with $\Delta T \lesssim 0.06 \cdot T_c$ yields a binary hierarchy of clusters.
Figure 1: Dendrograms and overlap matrices for 100 randomly sampled pure states, at $T/T_c = 1.2, 0.86, 0.55, 0.12$. Red is maximally positive overlap $q_{\phi\psi} = +1$, blue is maximally negative overlap $q_{\phi\psi} = -1$, white is no overlap $q_{\phi\psi} = 0$. Figure from [20].
2.2.4 The replica method

The **Replica Method** has found, for both its simplicity and effectiveness, widespread use in physical literature as a tool for computing quenched averages in the theory of spin glasses. It is based on the observation that

(a) the problem of computing the quenched average $E \log Z$ could be reduced (with lots of care in interchanging $E$ and $\lim_{n \to 0}$) to the problem of computing $E Z^n$ via

$$E \log Z = \lim_{n \to 0} \frac{E Z^n - 1}{n}$$  \hspace{1cm} (2.27)

(b) for $n \in \mathbb{N}$, $E Z^n_N$ can be computed as an annealed average by using $n$ replicas

The procedure implementing the replica method is never fully justified in the literature, but has served the field well; it is certainly worth presenting here, as it encompasses all the core ideas behind replica symmetry breaking (RSB):

1. evaluate $E Z^n_N$ in the thermodynamic limit as in point (b) above, obtaining an expression in $n$ that holds for $n \in \mathbb{N}$ and can be extended analytically to the positive reals (or at least a neighbourhood of $n = 0$)

2. postulate that the expression for $\lim_{N \to \infty} Z^n_N$ continues to hold when analytically extended, postulate/prove that the limits $N \to \infty$ and $n \to 0$ can be exchanged and obtain the quenched average as suggested in point (a) above:

$$\lim_{N \to \infty} E Z^n_N = \lim_{n \to 0} \lim_{N \to \infty} \frac{E Z^n_N - 1}{n}$$

Employing $n$ replicas $\{\sigma^a\}_a$ (we’ll use indices from the first letters of the alphabet $1 \leq a, b, ... \leq n$) and integrating away the couplings one gets

$$E Z^n_N = \int dg_{ij} e^{-\frac{1}{2} \sum_{i \neq j} g_{ij}^2} \sum_{\sigma \in \{\pm 1\}^{nN}} \exp \left[ -\beta \frac{1}{\sqrt{N}} \sum_a \sum_{i \neq j} g_{ij} \sigma_i^a \sigma_j^a \right]$$

$$\sum_{\sigma^a, \ldots, \sigma^n} \exp \left[ \frac{\beta^2}{2N} \sum_{a,b} \sum_{i \neq j} \sigma_i^a \sigma_j^b \sigma_i^a \sigma_j^b \right]$$  \hspace{1cm} (2.28)

Something interesting happened here: the $n$ replicas, originally independent copies of the system, got coupled because they shared the same spin couplings.

The second line of eq’n (2.28) shows formal symmetry between replica indices $a, b$ and spin indices $i, j$: the process of eq’n (2.28) is then inverted\(^{19}\) this time considering **replica couplings**\(^{20}\) $\{Q_{ab}\}_{1 \leq a, b \leq n}$, to decouple the spins:

$$E Z^n_N = \int dQ_{ab} e^{-\frac{1}{\sqrt{N}} \sum_{a \neq b} Q_{ab}^2} \sum_{\sigma \in \{\pm 1\}^{nN}} \exp \left[ -\beta^2 \sum_i \sum_{a \neq b} Q_{ab} \sigma_i^a \sigma_i^b \right]$$  \hspace{1cm} (2.29)

\(^{18}\)For reason of clarity, we have absorbed the distribution normalisation factors into the differentials.

\(^{19}\)Again we have absorbed the distribution normalisation factors into the differentials.

\(^{20}\)The $N\beta^2$ is there because the $Q_{ab}$ couplings correspond to the $\frac{1}{\sqrt{N}} g_{ij}$ couplings, which had variance $\frac{1}{N}$. 

16
The spin indices finally disappear by rewriting eq’n 2.29 as
\[
\mathbb{E} Z_N^n = \int dQ_{ab} e^{-\frac{1}{2} N \beta^2 \sum_{a \neq b} Q_{ab}^2 \left( \sum_{S \in \{\pm 1\}^n} \exp[-\beta^2 \sum_{a \neq b} Q_{ab} S^a S^b] \right)^N} \tag{2.30}
\]
In the thermodynamic limit \( N \to \infty \) the integral in eq’n 2.30 is evaluated by saddle point method, finding the dominant critical point \( Q^* \equiv \{Q_{ab}^*\}_{ab} \) of eq’n 2.31 for all \( n \). The expression for \( \mathbb{E} Z_N^n \) valid at large \( N \) is then obtained from eq’n 2.33:
\[
\mathcal{F}(Q) = \frac{1}{2} \beta \sum_{a,b} Q_{ab}^2 - \frac{1}{\beta} \log \mathcal{Z}(Q) \tag{2.31}
\]
\[
\mathcal{Z}(Q) = \sum_{S \in \{\pm 1\}^n} \exp[\beta^2 \sum_{a,b} Q_{ab} S^a S^b] \tag{2.32}
\]
\[
\mathbb{E} Z_N^n = \exp[-\beta N \mathcal{F}(Q^*(n))] \tag{2.33}
\]
Finally the trick from eq’n 2.27 gives 21 the quenched average in terms of the saddle-point free energy functional \( \mathcal{F}(Q) \):
\[
\lim_{N \to \infty} \frac{1}{N} \mathbb{E} \log Z_N = -\frac{1}{\beta} \frac{\partial}{\partial n} \mathcal{F}(Q^*(n)) \bigg|_{n=0} \tag{2.34}
\]

2.2.5 Replica symmetric (RS) solution

The first step on the way to getting an ansatz for the matrix \( Q_{ab}^* \) is to understand its physical meaning, which is hidden in the saddle point equations \( \mathcal{F}'(Q^*) = 0 \)
\[
Q_{ab}^* = \frac{1}{\mathcal{Z}(Q^*)} \sum_{S \in \{\pm 1\}^n} S^a S^b \exp[\beta^2 \sum_{c,d} Q_{cd} S^c S^d] \tag{2.35}
\]
The RHS of eq’n 2.35 is the Gibbs average \( \langle S^a S^b \rangle_{Q^*} \) in an \( n \)-spin system, and the \( Q_{ab}^* \) will thus be called the overlaps of the replicas.

Sherrington and Kirkpatrick give in [36] their ansatz for \( Q_{ab}^* \), the only ansatz that leaves replica symmetry 22 unbroken:
\[
Q_{ab}^* = u \delta_{ab} + v (1 - \delta_{ab}) \tag{2.36}
\]
A visualisation of the resulting matrix is given on the left in figure 2 (p. 19).

Constrained extremisation over \( u, v \) (yielding \( u = 1 \) and \( v = q \)) then reproduces Guerra’s replica-symmetric solution from theorem 2.2 (p. 10):
\[
\lim_{N \to \infty} p_N = -\frac{1}{\beta} \frac{\partial}{\partial n} \mathcal{F}(Q^*(n)) \bigg|_{n=0} = \log 2 + \mathbb{E} \log[\cosh(\beta g \sqrt{v})] + \frac{\beta^2}{4} (u - v)^2 \tag{2.37}
\]

21 Along with the observation that consistency with \( Z_0 = 1 \) implies the requirement \( \mathcal{F}(Q^*)|_{n=0} = 0 \).

22 I.e. permutational symmetry of the replicas.
2.2.6 Replica symmetry breaking (RSB)

Unless otherwise stated, this section is based on [21][22][23].

The ansatz of eq’n 2.36 corresponds to full $S_n$ permutational symmetry of the $n$ replicas: each fixed replica $a$ has overlap $Q^*_{aa} = 1$ with itself and $Q^*_{ab} = q$ with any other replica $b$. But in the low temperature region this ansatz stops holding, and we have to break the $S_n$ symmetry.

The minimal symmetry breaking scheme considered is

$$S_n \to S_{m_1} \times S_{n/m_1} \quad (2.38)$$

corresponding to the $n$ replicas splitting into $n/m_1$ identical clusters of size $m_1$, with full permutation symmetry within each cluster. Each fixed replica $a$ has the following overlaps:

- $Q^*_{aa} = 1$ with itself
- $Q^*_{a0} = q_0$ with any of the $m_1 - 1$ replicas $b_0$ in its cluster
- $Q^*_{a1} = q_1$ with any of the other $n - m_1$ replicas $b_1$

This scheme is called 1-step replica symmetry breaking or 1RSB, and is the one we’ll work with when talking of K-sat later on.

The replica breaking scheme can be generalised, introducing an ultrametric hierarchy of clusters: the general case is called K-step replica symmetry breaking scheme, or K-RSB, and consists of a K-level hierarchy of clusters of sizes $1 < m_1 < m_2 < ... < m_K < n$, where each cluster of size $m_{i+1}$ contains $m_{i+1}/m_i$ clusters of size $m_i$. It corresponds to the symmetry breaking scheme

$$S_n \to S_{m_1/m_0} \times S_{m_2/m_1} \times S_{m_3/m_2} \times ... \times S_{m_{K+1}/m_K} \quad (2.39)$$

where by convention we define $m_0 = 1$ (the trivial cluster containing a single replica) and $m_{K+1} = n$ (the trivial cluster containing all the $n$ replicas). The overlaps $1, q_0, q_1, ..., q_K$ generalise in the obvious way.

The K-RSB scheme is encoded by the non-increasing cluster size function (we set $0 = q_{K+1}$ and $1 = q_{-1}$ by convention):

$$m(q) \overset{\text{def}}{=} m_i \text{ when } q_i \leq q < q_{i-1}, \text{ for all } 0 \leq i \leq K + 1$$

$$= \max \text{ size of clusters with overlaps all of absolute value } > q \quad (2.40)$$

In the context of RSB the $n \times n$ matrix $Q^*_{ab}$ is called the Parisi matrix, and the terminology is extended to the RS case. Figure 2 (p.19) shows an example of RS Parisi matrix (on the left) and 3RSB Parisi matrix (on the right).

Similarly to the RS solution, also the 1RSB solution for $\lim_{N \to \infty} p_N$ stops holding as $\beta$ grows bigger, but the full family of K-RSB ansatzes for $K \geq 1$ is enough to cover the entire low-temperature region. We’ll talk about K-RSB phase\(^{24}\) when referring to regions of the phase space where the K-RSB ansatz holds.

---

\(^{23}\)This K has nothing to do with the K of K-sat, it just follows the convention in the literature.

\(^{24}\)For a specific value of $K$, and we’ll talk about the RS phase similarly.
Figure 2: An example of RS Parisi matrix (on the left) and 3RSB Parisi matrix (on the right), both with \( n = 90 \). The RS case has only the trivial cluster sizes \( m_0 = 1, m_1 = n \); its only overlap is \( q_0 = q = 0.4 \). The 3RSB case, on the other hand, has the trivial cluster sizes \( m_0 = 1 \) (not visualised) and \( m_4 = n \), and the non-trivial cluster sizes \( m_1, m_2, m_3 = 3, 15, 45 \); its overlaps are \( q_0 = q = 0.9 \) and \( q_1, q_2, q_3 = 0.9, 0.6, 0.4, 0.1 \). Original figure from [20], modified by the author.

2.2.7 The \( n \to 0 \) limit in RSB (a.k.a. the messy part)

RSB concerns clustering of replicas, but what does it say about clustering of pure states? A hint to the answer is given by the pure state overlap PDF, which in the K-RSB phase is

\[
\mathbb{E} P(q') = \lim_{n \to 0} \sum_{i=0}^{K} (m_i - m_{i+1}) \delta(q' + q_i) \tag{2.41}
\]

The Parisi matrix construction has to be modified if we want some sort of \( n \to 0 \) limit to exist: we need \( 1 = m_0 > m_1 > \ldots > m_K > 0 = m_{K+1} \), and the \( m_i \) will assume a probabilistic interpretation along with the function \( m(q) \) which encodes them. Formally

\[
\mathbb{E} P(q') = \lim_{n \to 0} \frac{\partial}{\partial q} m(q) \tag{2.42}
\]

and the function \( x(q) \overset{\text{def}}{=} \lim_{n \to 0} m(q) \) is then akin to a pure state overlap cumulative distribution function. This way the hierarchical K-level clustering of replicas will transfer to pure state space, as we wanted. But to do this properly one has [3] to understand what a space of \( 0 \times 0 \) matrices would look like, and that’s a story for another time.
3 From SK to K-sat

Unless otherwise stated, this section is based on [3][6][8][10][12][15].

3.1 Diluted SK and the K-sat Hamiltonian

The connection between SK and K-sat goes through the so-called diluted SK model: the Hamiltonian is obtained by multiplying the spin couplings in the SK model by iid Bernoulli r.v.s $\gamma_{ij}$ taking values in $\{0, 1\}$:

$$H_N(\sigma) \overset{\text{def}}{=} -\sum_{i<j} g_{ij} \gamma_{ij} \sigma_i \sigma_j$$  \hspace{1cm} (3.1)

Taking $E \gamma_{ij} = P(\gamma_{ij} = 1) = \gamma/N$ for some fixed constant $\gamma > 0$, we get that each spin interacts on average with $\gamma$ other spins (independently of $N$).

Luckily it turns out [3] that a number of results from the SK model can be transferred to the diluted SK model, and more in general to a broader class of Hamiltonians in the form

$$H_N(\sigma) \overset{\text{def}}{=} \sum_{a \leq M} W_a(\sigma_{i(a,1)}, \ldots, \sigma_{i(a,p)})$$  \hspace{1cm} (3.2)

where

1. the $W_a : \{\pm 1\}^p \rightarrow \mathbb{R}$ are iid random functions
2. for some fixed constant $\alpha$ we have $M = \alpha N$ or, alternatively, $M$ is Poisson with mean $\alpha N$
3. the sets $\{1 \leq i(a,1) < \ldots < i(a,p) \leq N\}$ are iid and uniformly distributed
4. the three sources of randomness above are independent of each other

The diluted SK model is then given by $p = 2$ and

$$W_a(\sigma_{i(a,1)}, \sigma_{i(a,2)}) = -g_{i(a,1)i(a,2)} \gamma_{i(a,1)i(a,2)} \sigma_{i(a,1)} \sigma_{i(a,2)}$$  \hspace{1cm} (3.3)

K-sat belongs to said class, since its Hamiltonian from section 1.3 (p.6) is given by $p = K$ and

$$W_a(\sigma_{i(a,1)}, \ldots, \sigma_{i(a,K)}) = \prod_{r=1}^{K} \left(1 + J_{a}^r \sigma_{i(a,r)}\right)$$  \hspace{1cm} (3.4)

In the case of K-sat, the $W_a$ functions are called the indicator functions of the clauses.

Therefore the hope is that good part of the behaviour of the SK model will transfer, at least qualitatively, to random K-sat: this observation has been used throughout the years to gain intuition on what to expect, and what to look for, in the statistical treatment of the problem.

---

25 The $\frac{1}{\sqrt{N}}$ factor in front of the SK Hamiltonian, which gave the spin couplings a variance of $\frac{1}{N}$, has been absorbed into the expectation of the Bernoulli r.v.s
3.2 Factor graphs

A natural way to study spin systems with sparse interactions (like the Ising model, diluted SK or K-sat) is to focus on the geometry of the interactions themselves: this is done by studying the so-called factor graph of the Gibbs measure.

Consider a collection of $N$ variables $\sigma = (\sigma_1, ..., \sigma_N)$ and a function $\mu(\sigma)$ that factors as

$$\mu(\sigma) = \frac{1}{Z} \prod_{a \in F} \mu_a(\sigma|_{\partial_a})$$

(3.5)

where

1. $Z$ is the normalisation constant
2. the $\mu_a$ functions are called the factors
3. $\partial_a \overset{\text{def}}{=} \{1 \leq i \leq N \text{ s.t. } \mu_a \text{ depends on } \sigma_i\}$ is the neighbourhood of factor $\mu_a$
4. $\sigma|_{\partial_a} = (\sigma_{i(a,1)}, ..., \sigma_{i(a,p_a)})$ where we wrote $\partial_a = \{i(a,1), ..., i(a,p_a)\}$

We'll use indices $a, b, c, d, ...$ from the beginning of the alphabet for the factors, and indices $i, j, k, l, ...$ from the middle of the alphabet for the variables. $F$ is the set of factor indices and we'll let $V$ denote the set of variable indices.

The factor graph for eq'n 3.5 is then defined to be the bipartite graph with node classes

1. variable nodes $\sigma_1, ..., \sigma_N$
2. factor nodes $\mu_1, ..., \mu_M$

and edge $\sigma_i \leftrightarrow \mu_a$ if and only if $i \in \partial_a$. Then $\{\sigma_i \text{ s.t. } i \in \partial_a\}$ is the graph neighbourhood of factor node $W_a$. Also we can define $\partial_i \overset{\text{def}}{=} \{1 \leq a \leq M \text{ s.t. } \mu_a \text{ depends on } \sigma_i\}$, the neighbourhood of variable $\sigma_i$, so that $\{\mu_a \text{ s.t. } a \in \partial_i\}$ is the graph neighbourhood of variable node $\sigma_i$.

The Hammersley-Clifford theorem [15][11] guarantees that all positive Markov fields and Gibbs ensembles can be represented by factor graphs, and the Gibbs measure for the K-sat problem takes indeed the form of eq'n 3.5

$$\mu_N(\sigma) = \frac{1}{Z_N} \exp \left[ -\beta \sum_{a \in F} W_a(\sigma|_{\partial_a}) \right] = \frac{1}{Z_N} \prod_{a \in F} \exp \left[ -\beta W_a(\sigma|_{\partial_a}) \right]$$

(3.6)

The factors nodes correspond to clauses, and the neighbourhood of a factor node is composed of the variables involved in the clauses; the variable node correspond to variables/spins.

An example of factor graph for a 3-sat instance can be found in figure 3 (p.22). No distinction is made in the graph topology between positive and negated occurrence of variables in clauses: that information is encoded in the indicator function associated to each factor node.

\[26\] And we'll confuse indices with the corresponding variables and factors
Figure 3: A small region of a factor graph for 3-sat. Factor nodes are depicted as squares, spin nodes are depicted as circles. Inspired by [9].

The connectivity of factor nodes in K-sat is always K; the connectivities of the variable nodes, on the other hand, depend on the generative model chosen, but become iid poisson r.v.s with mean $K\alpha$ in the thermodynamic limit

$$M, N \to \infty \text{ with } M/N \to \alpha \text{ w.h.p.}$$

We expect, from theorem 2.3 (p.10), that the Gibbs measure will factor in the thermodynamic limit. This results in the following behaviour of the factor graph at loop level

$$\text{typical size of a loop} = O(\log(N))$$

and has two key consequences:

1. the factor graph is almost tree-like, i.e. breadth first exploration of the graph from any node will typically produce trees of diameter $O(\log(N))$ before encountering a loop. \footnote{This is part of the reason why Belief Propagation and Survey Propagation work so well on random K-sat.}

2. consider a factor node and any two of the K spins connected to it: the spins have initially distance 2, but upon removal of the factor node their distance will typically jump up to $O(\log(N))$. \footnote{This is part of the reason why the Cavity method is so effective: removal of a single factor node, i.e. creation of a cavity, will make its neighbouring spins to good approximation independent of each other.}
3.3 Pure states and clusters of solutions

Now that we have factor graphs in our hands we are able to give a better definition of pure states, and it will become evident how pure states, rather than spin configurations, are the correct way to study the random K-sat problem.

A probability measure
\[ \mu_N^{(\psi)} \] on \( \Sigma_N \) is a pure state iff its correlation function

\[
C_N(r) = \sup_{I,J \subseteq V \text{ s.t. } \inf_{i \in I, j \in J} d(\sigma_i, \sigma_j) \geq r} \left| \sum_{i \in I, j \in J} \mu_N^{(\psi)}(\{\sigma_i, \sigma_j\}) - \mu_N^{(\psi)}(\sigma_i) \mu_N^{(\psi)}(\sigma_j) \right|
\]

decays at large \( r \) in the thermodynamic limit, i.e. iff we have

\[
\limsup_{N \to \infty} C_N(r) \to 0 \text{ as } r \to \infty
\]

This definition, although conceptually similar to the one given in section 2.2.1, is rather cumbersome, and we’ll not use it directly: instead we’ll restrict our attention to the zero temperature limit of the so-called satisfiable phase.

In the zero temperature limit, the measure concentrates on the solutions of \( \mathcal{I} \): it can be shown that, at zero temperature, a pure state \( \psi \) is composed of a set of spin configurations that are

1. all of same energy \( E^{(\psi)} \)
2. connected by 1-spin flips, i.e. connected in \( \Sigma_N \) with graph structure given by the Hamming distance
3. locally stable, in the sense that the energy cannot be decreased by any 1-spin flip

In [6] these are called clusters of solutions, as it is noted that, in the satisfiable phase, they are nothing but the connected components of the space \( \mathcal{S}(\mathcal{I}) \subset \Sigma_N \) of solutions (i.e. satisfying assignments) to our instance \( \mathcal{I} \).

From now on when talking of a pure state \( \psi \) we’ll use notation \( \psi \) to denote the cluster of solutions as well as labelling the state, and we’ll adopt the zero temperature definition above as our working definition.

---

[29] To be precise, a family of probability measures \( (\mu_N^{(\psi)})_N \).

[30] I.e. the region where at least a solution exists

[31] I.e. the measure is supported by.

[32] I.e. an edge will connect a pair of spin configurations \( (\sigma, \tau) \) if and only if \( d(\sigma, \tau) = 1 \) as per eq’n 2.2
3.4 The complexity

Physical experience with spin glasses and other frustrated systems (see e.g. [5]) suggests that, in the RSB phases, pure states should grow exponentially in \( N \), at least to leading order.

If we denote by \( \mathcal{N}(\omega;\alpha) \) the expected number of pure states \( \psi \) with free-entropy density \( \omega \)

\[
\omega \equiv \frac{1}{N} \log Z_N^{(\psi)}
\]  \hspace{1cm} (3.9)

then \( \mathcal{N}(\omega;\alpha) \) allows to define the complexity \( \Sigma(\omega;\alpha) \) as the unique function satisfying

\[
\mathcal{N}(\omega) \approx \exp \left[ N \Sigma(\omega;\alpha) \right]
\]  \hspace{1cm} (3.10)

What is \( \omega \) exactly? Eqn 3.9 tells us that at zero temperature and in the satisfiable phase

\[
\omega = \frac{1}{N} \log \left[ \text{# of solutions in } \psi \right]
\]  \hspace{1cm} (3.11)

and thus \( \omega \) coincides with the entropy density \( s \) (by definition the RHS of eqn 3.11).

A plot of the complexity \( \Sigma(\phi;\alpha) \) for 4-sat is given in figure 4: negative complexity for a value \( \omega \) means that w.h.p. there is no cluster with that free-entropy density, while vanishing of the complexity just implies a sub-exponential number of clusters. Notice also that the complexity curve only covers a reduced range of \( \omega \) for \( \alpha \) small enough.

![Figure 4: Plot of the complexity \( \Sigma(\omega;\alpha) \) of 4-sat, as a function of the free-entropy density \( \omega \) (here denoted \( \phi \)) and for various values of \( \alpha \). Figure from [8].](image)
4 The phases of K-sat

4.1 The Parisi 1RSB parameter

This section is based on [8].

The definition of pure state given in section 2.2.1 (p.12) is recovered for K-sat if we write

\[
Z_N^{(\psi)} \overset{\text{def}}{=} \sum_{\sigma \in \psi} \prod_{a \in F} \exp \left[ -\beta W_a(\sigma|\partial_a) \right]
\]

\[
\mu_N^{(\psi)}(\sigma) \overset{\text{def}}{=} \frac{1}{Z_N^{(\psi)}} \prod_{a \in F} \exp \left[ -\beta W_a(\sigma|\partial_a) \right]
\]

\[
\mu_N(\sigma) = \sum_\psi w_\psi \mu_N^{(\psi)}(\sigma), \quad \text{with} \quad w_\psi \overset{\text{def}}{=} \frac{Z_N^{(\psi)}}{Z_N} = \text{# of solutions in } \psi \text{ over total # of solutions}
\]

The presentation of the phases of K-sat will cover the RS and 1RSB phases only, as the higher RSB phases are poorly understood in terms of sparse factor graphs. The relevant parameters from section 2.2.6 will be the intra-state overlap \(q_0\), the inter-state overlap \(q_1\) and the Parisi 1RSB parameter \(m_1 \in [0, 1]\), which from now on we'll denote \(m\).

The main quantity to study in the context of RSB is the replicated free-entropy density

\[
\Phi(m) \overset{\text{def}}{=} \lim_{N \to \infty} \frac{1}{N} \mathbb{E} \log \sum_\psi \left( Z_N^{(\psi)} \right)^m
\]

The replicated free-entropy is directly related to the complexity by the equations

\[
\Phi(m) = \sup_{\omega \in [\omega_-, \omega_+]} (\Sigma(\omega) + m \cdot \omega)
\]

\[
\Sigma(\omega(m)) = \Phi(m) - m \Phi'(m) = \Phi(m) - m \cdot \omega(m)
\]

where \(\Sigma(\omega)\) is defined and positive on \([\omega_-, \omega_+]\). Eqn 4.5 is a Legendre inversion and requires \(\Sigma(\omega)\) to be concave and \(m\) to be in a range \([m_-, m_+]\) s.t. the supremum of eqn 4.5 is found in the interior of \([\omega_-, \omega_+]\). The total free-entropy density is, in the 1RSB approximation, the minimum of \(\Phi(m)/m\).

We'll overload the notation and write \(\Sigma(m = m_0) \equiv \Sigma(\omega(m_0))\). Note that the slope of curve \(\Sigma(m = m_0)\) at any particular value of \(m_0\) of \(m\) is \(\Sigma(m = m_0)' = -m_0\): the point \(m = 0\) marks the maximum of the complexity curve, while the point \(m = 1\) is, when \(\Sigma\) is defined there, the point where the complexity attains slope \(\Sigma(m = 1)' = -1\).

Figure 4 (p.24) shows the complexity for 4-sat at different values of \(\alpha\): for \(\alpha = 9.3\) the complexity is defined only on \([\omega_-, \omega_+] \approx [0.02, 0.04]\), and has no point of slope \(-1\); for \(\alpha = 9.45\) the complexity attains slope \(-1\) at a value of \(m\) where it is still positive (the point \((\Sigma(m = 1), \omega(m = 1))\) is marked with the black circle); for higher values of alpha the complexity vanishes before reaching slope \(-1\), at some \(m = m_s < 1\) (the points \((\Sigma(m = m_s), \omega(m = m_s))\) are marked with black squares). More in general we'll denote by \(m_s\) the Parisi 1RSB parameter describing the thermodynamically relevant clusters in RSB phases (this will become clear in the coming section).

\[33\] The complexity is always concave for K-sat.
4.2 Replica symmetry breaking

This section is based on [8]. Replica symmetry breaking manifests in K-sat in the following phases:

RS In the **replica symmetric phase** the measure is concentrated \(^{34}\) in a single thermodynamically relevant cluster, that is to say

\[
\max_{\psi} w_\psi \to 1 \text{ as } N \to \infty \text{ w.h.p.} \tag{4.7}
\]

The replicated free-entropy is given by \(\Phi_{RS}(m) = m \cdot \omega_*\), where \(\omega_*\) is the contribution of the single dominant cluster. We also have \(Z_N \approx \exp[N \omega_*]\).

d1RSB In the **dynamical 1RSB phase** the measure is concentrated into

\[
\mathcal{N}(\omega_*) \approx \exp[N \Sigma_*] \text{ clusters} \tag{4.8}
\]

all with the same weight \(^{35}\)

\[
w_\psi \approx \exp[-N \Sigma_*] \text{ and free-entropy density } \omega_*.\]

\(\Phi(m)/m\) is minimised at \(m = 1\), with \(\Sigma_* \equiv \Sigma(\omega_*) = \Phi(1) - \Phi'(1) > 0\) and \(\omega_* = \Phi'(1)\). The thermodynamically relevant clusters are thus described by a 1RSB solution with Parisi parameter \(m = m_s = 1\), which means they indeed all have the same size.

1RSB In the **1RSB phase** the measure is concentrated into a sub-exponential number of clusters \(\psi_1, \psi_2, ...\) (w.l.o.g., consider them in order of decreasing weight). The sequence of weights \(w_\psi\) converges to a Poisson-Dirichlet Point process\(^{36}\) of parameter \(m_s \in (0, 1)\)

\[
1 \geq w_{\psi_1} \geq w_{\psi_2} \geq ... \geq 0, \text{ satisfying } \sum_{i=1}^{\infty} w_{\psi_i} = 1
\]

\[
\text{and } w_{\psi_i} = z_i \cdot \prod_{1 \leq j < i} (1 - z_j) \tag{4.9}
\]

for \((z_i)\) i.i.d. with density \(m_s (1 - z)^{m_s - 1}\).

For more details about Poisson-Dirichlet Point processes see [3][26][46].

The thermodynamically relevant clusters are thus described by a 1RSB solution with a Parisi parameter \(m = m_s < 1\) minimising \(\Phi(m)/m\). The free-entropy density of these states is \(\omega_* = \Phi'(m_s)\), and the complexity \(\Sigma_*\) vanishes as expected from the sub-exponential number of states.

Note that, in the d1RSB and 1RSB phases, the value of \(m\) describing the thermodynamically relevant clusters is always the one minimizing \(\Phi(m)/m\).

\(^{34}\)From now on by the measure is concentrated we’ll mean that the states we’re ignoring are thermodynamically irrelevant, i.e. their collective measure vanishes in the thermodynamic limit.

\(^{35}\)This is 1RSB, so we expect a single layer of clusters, all statistically identical in the thermodynamic limit.

\(^{36}\)Unsurprisingly, the same process that governs allele frequencies in infinite coalescent trees [47].
4.3 Phase transitions in K-sat

This and the following sections are based on [8][7][6][2], with some insights from [1][10][26] (the figure on top is from [37]). As $\alpha$ grows, K-sat undergoes the following structural phase transitions, some of which will be presented in detail in the next sections:

$\alpha_{d,+}$ A positive complexity appears for some values of $\omega$, but K-sat is still in the RS phase: there is a single dominant cluster which at this phase transition starts shedding an exponential number of thermodynamically irrelevant clusters described by a 1RSB solution with $m = 1$.

$\alpha_d$ K-sat undergoes the clustering phase transition\(^{37}\) from the RS phase to the d1RSB phase (from now on also called the clustered phase): the main cluster shatters into an exponential number of exponentially small, almost identical, thermodynamically relevant clusters (described by a 1RSB solution with $m = m_s = 1$).

$\alpha_c$ K-sat undergoes the condensation phase transition from the d1RSB phase to the 1RSB phase (from now on also called the condensed phase): the solutions tend to condensate into a sub-exponential number of thermodynamically relevant clusters (described by a 1RSB solution with $m = m_s < 1$). In 3-sat this coincides with $\alpha_d$ (i.e. no clustered phase in 3-sat).

$\alpha_f$ Frozen variables appear in thermodynamically dominating clusters.

$\alpha_s$ K-sat undergoes the SAT/UNSAT phase transition, with a random instance $I$ going from satisfiable w.h.p. to unsatisfiable w.h.p.

\(^{37}\)Also known as dynamic phase transition.
4.4 Clustering: the Dynamical phase transition RS $\rightarrow$ d1RSB

Take $D$ to be the ball of radius $l$ around some random variable node in the factor graph, and $\tau \equiv \sigma|_{\mathcal{V}\setminus D}$ to be its complement. As $\alpha$ crosses $\alpha_d$, the spins become [7] globally correlated under the Gibbs measure, i.e. the following point-set correlation stops vanishing in the $l \to \infty$ limit

$$C_l \overset{\text{def}}{=} \mathbb{E} \sum_{\tau} \mu_N(\tau) \| \mu_N(\cdot|\tau) - \mu_N(\cdot) \|$$

[2] mentions a 2010+ result from Coja-Oghlan and Gerke by which $\lim_{l \to \infty} C_l = 1/2$ once crossed $\alpha_d$. Figure 5 (p.31) shows this for 4-sat.

This phase transition is best characterised by the clustering of solution space, which the following groundbreaking results from [6] presents at its finest.

**Theorem 4.1. (Shattering)**

We define a region to be any non-empty union of clusters of solutions, and we also define the height of a path $\sigma(1), ..., \sigma(T)$ to be

$$\text{height}(\sigma(1), ..., \sigma(T)) \overset{\text{def}}{=} \max_{1 \leq t \leq T} H_N(\sigma(t))$$

Then there is a sequence $\epsilon_K \to 0$ s.t. for all $\alpha$ in the region

$$\alpha_d = (1 + \epsilon_K) \frac{2K}{K} \log K \leq \alpha \leq (1 - \epsilon_K) \frac{2K}{K} \log 2$$

the solution space of a random $K$-sat instance $\mathcal{I}$ shatters, i.e. w.h.p. there exists at least $\exp[O(N)]$ regions with the following properties:

1. each region contains at least an $\exp[-O(N)]$ fraction of all solutions
2. the distance between any two vertices in distinct regions is at least $O(N)$
3. every path between any two vertices in distinct regions has height at least $O(N)$

The picture is the following: for low $\alpha$ the solution space is a single big cluster, but as it crosses $\alpha_d$ the big cluster shatters into an exponential amount of exponentially small regions, very far from each other and separated by very high energy barriers.

The proof found in [6] is probabilistic and very rigorous, based on the establishment of a connection between the planted model and the uniform model that enables the authors to work with typical solutions but still leave space of manoeuvre by allowing a (small but exponential) number of atypical solutions. This level of rigour is relevant because, up until then, similar results had been based on the physics-inspired cavity method, which has not yet received a proper mathematical formalisation.

---

38 For $K$ small enough the region might be empty. [6] mentions quick calculations suggesting the result to hold at least for $K \geq 8$, and the evidence presented in the rest of this chapter would suggest the result to hold for $K \geq 4$.

39 The farthest they can be is indeed $O(N)$, more precisely $N$.

40 The highest they can be is $O(M) = O(N)$, more precisely $M$. 

28
A rigorous value for the clustering phase transition is given in [6] to be

$$\alpha_d = (1 + o(1))\frac{2^K}{K} \log K$$

(4.13)

while [7] and [8] use the cavity method to obtain the more refined (but less rigorous)

$$\alpha_d = \frac{2^K}{K} \left( \log K + \log \log K + 1 + O\left( \frac{\log \log K}{\log K} \right) \right)$$

(4.14)

Now it’s time to give a look at the statistical mechanics of the clustering phase transition.

Figure 7 (p.33) shows the clustering phase transition $\alpha_d$ to be the point at which a positive complexity $\Sigma(m = 1) > 0$ starts existing: this makes $m = 1$ the new minimum for $\Phi(m)/m$, and corresponds to the appearance of an exponential number of thermodynamically relevant clusters with hierarchical structure described by a 1RSB solution $m = m_s = 1$.

Figure 7 also shows that the RS estimate $\omega_{RS}$ for the free-entropy density can be expressed as

$$\omega_{RS} = \Sigma(m = 1) + \omega(m = 1)$$

(4.15)

which confirms the picture of the clustering phase transition as shattering of a single solution cluster of entropy density $\omega_{RS}$ into $\exp[N \Sigma(m = 1)]$ clusters of entropy density $\omega(m = 1)$, without any discontinuous loss of global entropy (or, equivalently, of number of solutions) in the process. The free-entropy density of the typical solution, though, has a discontinuous jump down.

It should be noted that the maximum $\Sigma(m = 0)$ of the complexity curve is defined way before the clustering phase transition: this corresponds to the existence of an exponential number of clusters described by a 1RSB solution with parameter $m = 0$. But in the RS phase these solutions are thermodynamically irrelevant: this is indeed the $\alpha_{d,+}$ phase transition. Figure 7 shows this for 4-sat, at $\alpha_{d,+} \approx 8.297$.

Figure 4 (p.24) shows more in detail this evolution of the complexity curve: for $\alpha < \alpha_d$ only a small part of it is defined, around $\omega(m = 0)$, while for $\alpha > \alpha_d$ the curve is defined at least up to $\omega(m = 1)$ (and much further): it is indeed the appearance of a point $\Sigma'(m = 1)$ of slope $-1$ that marks the transition (it’s what sets the minimum for $\Phi(m)/m$ at $m = 1$).

Figure 6 (p.32) confirms that $m_s = 1$ constantly in the region $\alpha_d \leq \alpha \leq \alpha_c$, i.e. that in the d1RSB phase the thermodynamically relevant clusters are indeed described by a 1RSB solution with parameter $m = 1$ (and are thus all approximately equal in size).

Finally the overlaps in the d1RSB phase are given by $q_0(m = 1)$ and $q_1(m = 1)$, and are shown in red in figure 8 (p.34).
4.5 Condensation phase transition d1RSB → 1RSB

Similarly to the clustering phase transition, the condensation phase transition can be formulated in terms of some notion of correlation decay. Specifically the factorisation result of theorem 2.3 (p.10), which holds for K-sat in the region $\alpha < \alpha_c$, fails in the condensed phase: for $\alpha > \alpha_c$ the following quantity stops vanishing [7] in the thermodynamic limit

$$E \sum_{\sigma_i^{(1)}(\cdot)} \left| \mu_N(\sigma_i^{(1)}, \ldots, \sigma_i^{(n)}) - \mu_N(\sigma_i^{(1)}) \cdot \ldots \cdot \mu_N(\sigma_i^{(n)}) \right|$$

(4.16)

The condensation phase transition is shown in [7] and [8] to happen at

$$\alpha_c = 2^K \log 2 - \frac{3}{2} \log 2 + O(2^{-K})$$

(4.17)

The best way to understand what’s happening this time is to go straight to the statistical mechanics.

Figure 7 (p.33) shows the condensation phase transition to be the point at which $\Sigma(m = 1)$ vanishes. Above $\alpha_c$ the quantity $\Phi(m)/m$ is maximised by the value $m = m_s < 1$ s.t. $\Sigma(m = m_s) = 0$, i.e. the highest value at which the complexity is non-negative: this means that the measure is concentrated into a sub-exponential (since $\Sigma = 0$) number of clusters described by a 1RSB solution of parameter $m = m_s < 1$.

Figure 7 also confirms that at $\alpha_c$ we get $\omega_{RS} = \omega(m = 1)$, and that the decrease in free-entropy density for the typical solution is continuous at $\alpha_c$. It also shows, though, that the free-entropy density has discontinuous derivative at condensation, a phenomenon that we’ll encounter again when talking about the residual free-entropy density for BP-guided decimation in section 5.3.3.

Figure 6 (p.32) shows the value of $m_s$ decreasing continuously from $m_s(\alpha_c) = 1$ to $m_s(\alpha_s) = 0$, where it vanishes as $m_s(\alpha) \approx \sqrt{\alpha_s - \alpha}$: the cluster weights oscillate wildly up until the point where all clusters vanish at $\alpha_s$.

Figure 8 (p.34) finally shows, in blue, the overlaps $q_0(m = m_s)$ and $q_1(m = m_s)$ for the clusters in the 1RSB phase.

4.6 Freezing phase transition

A question of relevance for K-sat solvers is: how free am I to set an arbitrary value for a spin? We’ll see in section 5 that the general case is related to the so-called residual free-entropy density and will require us to compute the marginal distribution of the single spins (using Belief Propagation). The extreme case, though, is that where spins get frozen, i.e. they take only one value within a cluster, and can be understood in terms of the freezing phase transition.

Given a cluster $\psi$, we’ll define the projection on the $i^{th}$ spin by

$$\pi_{\psi}(i) = \{s \in \{\pm 1\} \text{ s.t. } \exists \sigma \in \psi \text{ s.t. } \sigma_i = s\}$$

(4.18)

and we’ll say that the variable $x_i$ is frozen in cluster $\psi$ if $\pi_{\psi}(i) \neq \{\pm 1\}$, i.e. if its value in the cluster is fixed.
Figure 5: The point-set correlation $C_l$ for 4-sat: left to right $\alpha = 9.30, 9.33, 9.35 < \alpha_d \approx 9.38$ (where it vanishes in the long-range limit $l \to \infty$) and $\alpha = 9.40 > \alpha_d$ (where it stops vanishing). Ignore the inset. Figure from [7].

Then a theorem from [16] guarantees that a freezing phase transition exists, i.e. there is a value $\alpha_f$ such that for $\alpha > \alpha_f$ every cluster will w.h.p. contain a majority of frozen variables:

$$\alpha_f = \left(\frac{4}{5} + o(1)\right) 2^K \log 2$$

(4.19)

In fact the authors prove that for any $\varepsilon \in (0, 1)$ there is an $\alpha_f(\varepsilon)$ s.t. every cluster will w.h.p. contain at least $\varepsilon N$ frozen variables (and $\alpha_f \equiv \alpha_f(1/2)$).

From the point of view of the Parisi 1RSB parameter, we can ask which value $m_f$ describes the biggest clusters that w.h.p. will have frozen variables: the freezing phase transition is then the point $\alpha_f$ s.t. $m_f(\alpha_f) = m_\sigma(\alpha_f)$, which figure 6 (p.32) shows for 4-sat to be at $\alpha_f \approx 9.88$. The situation presented by figure 6 is not generic though: for $K = 4$ we have $\alpha_f > \alpha_c$ while for $K \geq 6$ it can be shown that $\alpha_d < \alpha_f < \alpha_c$.

But we can do better: given a frozen variable $x_i$ and a satisfying assignment $\sigma$, we can ask how far we have to go in the solution space to find an assignment $\tau$ where $x_i$ takes a different value.

\footnote{For $K$ big enough: the authors of [16] prove it for $K \geq 9$ and report evidence suggesting that no freezing exists for $K = 3$; we'll see that experimental results from [8] suggest that freezing takes place for $K \geq 4$.}
We start by defining the following notions of rigidity:

(a) \( x_i \) is \( f(N) \)-rigid if for every satisfying assignment \( \tau \) we have \( \tau_i \neq \sigma_i \Rightarrow d(\sigma, \tau) > f(N) \). We'll (slightly) change our notion of frozen variable to: \( x_i \) is \textbf{frozen} if it is \( \log(N) \)-rigid.

(b) \( x_i \) is \( f(N) \)-loose if there is a satisfying assignment \( \tau \) s.t. \( \tau_i \neq \sigma_i \land d(\sigma, \tau) \leq f(N) \). We'll say that \( x_i \) is \textbf{fluid} if it is \( \log(N) \)-loose.

Then the authors of [6] and [2] proved the following result:

\textbf{Theorem 4.2. (Frozen and fluid variables)}

Let \((\mathcal{I}, \sigma)\) a random instance-solution pair \footnote{First choose a uniformly random satisfiable instance \( \mathcal{I} \), then choose a uniformly random satisfying assignment \( \sigma \) for \( \mathcal{I} \). The distribution of \((\mathcal{I}, \sigma)\) is not uniform over all instance-solution pairs.} s.t. \( \alpha \) is in the region of eq’n \ref{eq:4.12} (p.28). Then w.h.p. the number of \( \Omega(N) \)-rigid (and thus frozen) variables in \( \sigma \) will be at least \( \gamma_K N \), for a sequence \( \gamma_K \to 1 \). Furthermore the \( \Omega(N) \) bound is tight, as w.h.p. \( \sigma \) will have \( \Omega(N) \) variables that are not bound by any constraint (and thus cannot be rigid).

Finally in the region \( \alpha < \alpha_d \) every variable of \( \sigma \) will w.h.p. be \( o(n) \)-loose, and in fact fluid.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure6.png}
\caption{In black/red the Parisi 1RSB parameter \( m_s \) for the thermodynamically relevant clusters, going from the \( m_s = 1 \) of the clustered phase to the \( m_s = 0 \) of the UNSAT phase. In blue the parameter \( m_f \) describing the biggest clusters containing frozen variables. Figure from [8].}
\end{figure}
Figure 7: The complexity $\Sigma$ and free-entropy density $\omega$ (here denoted $\phi$) of 4-sat. The dynamical phase transition is at $\alpha_d \approx 9.38$, the condensation phase transition is at $\alpha_c \approx 9.547$ and the SAT/UNSAT phase transition is at $\alpha_s \approx 9.931$. Quantities are colour-coded based on the phase for which they are of interest: green are of interest for the RS phase, red are of interest for the clustering phase, and blue are of interest for the condensation phase. Figure from [8].

4.7 SAT/UNSAT phase transition

Finally the SAT/UNSAT phase transition marks the point where $I$ goes from being w.h.p. satisfiable to being w.h.p. unsatisfiable. Figure 7 (p.33) shows this to coincide with the point at which the maximum $\Sigma(m = 0)$ of the complexity vanishes.

Also we see that $\omega(m = m_s)$ gets to coincide with $\omega(m = 0)$: the thermodynamically dominant, condensed clusters (sub-exponential in number) decrease in free-entropy density and size up to those of the thermodynamically irrelevant (but exponential in number) clusters, and then the exponential family of clusters composing the whole solution space suddenly disappears at $\alpha_s$. 
Figure 8: The RS overlap is shown before the clustering phase transition. The 1RSB overlaps are shown in the clustered phase for \( m = 1 \) and in the condensed phase for \( m = m_s \). Figure from [8].

Figure 6 (p.32) shows this to coincide with the point at which \( m_s \) vanishes, suggesting that the thermodynamically dominant clusters become described, just before \( \alpha_s \), by the same 1RSB solution with \( m = 0 \) that until then described the thermodynamically irrelevant clusters generated at \( \alpha_{d,+} \).

Figure 8 provides further confirmation of this by showing that the overlaps \( q_0(m = m_s), q_1(m = m_s) \) of the 1RSB solution describing the condensed clusters go to coincide with the overlaps \( q_0(m = 0), q_1(m = 0) \) of the 1RSB solution describing the clusters shedded at \( \alpha_{d,+} \).

Cavity method calculations from [8] [7] set

\[
\alpha_s = 2^K \log 2 - \frac{1 + \log 2}{2} + O(2^{-K}) \tag{4.20}
\]

The following rigorous result is proven in [44] via the second moment method:

\[
2^K \log 2 - (K + 1) \frac{\log 2}{2} - 1 - o(1) \leq \alpha_s \leq 2^K \log 2 - \frac{1 + \log 2}{2} + o(1) \tag{4.21}
\]
5 Algorithms for K-sat

5.1 The algorithmic barrier

Unless otherwise stated, this section is based on [6][7][2][1].

The search for a clustering phase transition was initiated by the following empirical observation: all known efficient (i.e. poly-time w.h.p.) algorithms for K-sat stopped finding solutions at densities $\approx \frac{2^k}{K}$, much lower than the well known SAT/UNSAT threshold; in fact, no efficient algorithm performed, asymptotically in K, better than the naive **Unit Clause Propagation**\(^{43}\). In 2010, just 2 years after the proof of existence of the clustering phase transition at $\alpha_d \approx \frac{2^k}{K \log K}$ appeared in [6], the algorithm Fix, which succeeds w.h.p. up to $\alpha_d$ was introduced and rigorously analysed in [30]. The following table [30][2] gives **algorithmic barriers**\(^{44}\) for the best known efficient K-sat solvers.

| Algorithm               | Density $m/n < \cdots$               | Success probability |
|-------------------------|--------------------------------------|---------------------|
| Pure Literal            | $o(1)$ as $k \to \infty$             | w.h.p.              |
| Walksat, rigorous       | $\frac{1}{6} \cdot 2^k/k^2$         | w.h.p.              |
| Walksat, nonrigorous    | $2^k/k$                              | w.h.p.              |
| Unit Clause             | $\frac{1}{2} \left( \frac{k-1}{k-2} \right) \cdot \frac{k-2}{k} \cdot \frac{2^k}{k}$ | $\Omega(1)$        |
| Shortest Clause         | $\frac{1}{8} \left( \frac{k-1}{k-3} \right) \cdot \frac{k-3}{k-2} \cdot \frac{2^k}{k}$ | w.h.p.              |
| SC + backtracking       | $\sim 1.817 \cdot \frac{2^k}{k}$   | w.h.p.              |
| BP + decimation         | $e \cdot \frac{2^k}{k}$             | w.h.p.              |

So what is the intuition behind these barriers? If a problem has non-zero complexity, local algorithms will easily get stuck in the exponential multitude of local minima. This is the common problem of Walksat, Unit Clause, Shortest clause, all of which start failing at $\approx \frac{2^k}{K}$, just before the clustering phase transition. This clustering, with its exponential number of small, far-away dominant clusters, high energy barriers and long-range correlations, is widely believed to be the ultimate barrier for local algorithms. Indeed no efficient algorithm is rigorously proven to succeed past it.

We’ll now move onto the most important non-local algorithms, Belief Propagation and Survey Propagation: neither of them is rigorously known to succeed on densities higher than $\alpha_d$ (in fact we’ll see that Belief Propagation is guaranteed to fail w.h.p. at densities higher than $O\left(\frac{2^k}{K}\right)$), but the ideas involved in their formulation and analysis are invaluable for a thorough understanding of the connection between the phase transitions of K-sat and the efficiency of algorithms to solve it.

\(^{43}\)If there is a unit clause, satisfy it, otherwise assign a random value to a random variable

\(^{44}\)I.e. densities above which the algorithms stop succeeding w.h.p. (in polytime, when relevant).
5.2 Belief propagation

Unless otherwise stated, this section is based on [15][10][24], with some insights from [8][1].

The following DPLL algorithm\(^{45}\) forms the basis of many K-sat solvers, and succeeds w.h.p. in linear time on constraint densities up to \(\alpha = O(2^k/k)\):

\[
\text{DPLL}(I): \\
1. \text{Apply the pure literal rule} \(^{46}\) and satisfy any unit clause \(^{47}\) until no pure literal and/or unit clauses remain. Call the result \(I'\).
   (a) Exit returning SATISFIABLE if \(I'\) is empty.
   (b) Exit if a contradiction is generated.
2. Select a variable \(\sigma_i\) appearing in \(I'\) and a random value \(s_i \in \{\pm 1\}\)
3. DPLL(\(I'|_{\sigma_i = +s_i}\))
4. DPLL(\(I'|_{\sigma_i = -s_i}\))

A way to improve step 2 of DPLL is to compute the marginal distribution

\[
\mu_N(\sigma_i) = \sum_{\sigma_j \text{ s.t. } j \neq i} \mu_N(\sigma_1, ..., \sigma_N) \tag{5.1}
\]

and then setting \(s_i\) to its most likely value under the marginal \(\mu_N(\sigma_i)\). The purpose of the Belief Propagation (BP) algorithm is that of computing that marginal efficiently by exploiting the factor graph. BP returns the exact marginal if the factor graph is a tree, and a (hopefully converging) series of approximations if the factor graph has loops.

5.2.1 An typical example of BP

First we see a typical example of computation of marginals, taken from [15]. Suppose that \(N = 5\) and that the factor graph is given by figure \(9\) (p.37): then the measure factors as

\[
Z_5 \cdot \mu_5(\sigma_1, ..., \sigma_5) = f_A(\sigma_1)f_B(\sigma_2)f_C(\sigma_1, \sigma_2, \sigma_3)f_D(\sigma_3, \sigma_4)f_E(\sigma_3, \sigma_5) \tag{5.2}
\]

Using distributivity of sum and product (which is what BP is all about) we write the marginal as

\[
Z_5 \cdot \mu_5(\sigma_1) = f_A(\sigma_1)\left(\sum_{\sigma_2} \sum_{\sigma_3} f_B(\sigma_2)f_C(\sigma_1, \sigma_2, \sigma_3)\left(\sum_{\sigma_4} f_D(\sigma_3, \sigma_4)\left(\sum_{\sigma_5} f_E(\sigma_3, \sigma_5)\right)\right)\right) \tag{5.3}
\]

\(^{45}\)And in fact a much wider family of DPLL algorithms obtained by replacing step 1 with better heuristics.  
\(^{46}\)Satisfy all pure literals, i.e. literals that appear always with the same polarisation, i.e. literals the complement of which doesn’t appear in the formula.  
\(^{47}\)I.e. a clause with just one literal.
We see that to compute the marginal $\mu_5(\sigma_1)$ we need $f_A(\sigma_1)$ and $f_{BCDE}(\sigma_1, \sigma_2, \sigma_3) \downarrow \sigma_1$, where we defined the summary operator $\downarrow$ by

$$
f(\sigma|\partial a) \downarrow \sigma_i \overset{\text{def}}{=} \sum_{\sigma_j \text{ s.t. } j \in \partial a \setminus \{i\}} f(\ldots) \quad (5.4)
$$

We have also defined the shorthand

$$
f_{BCDE}(\sigma_1, \sigma_2, \sigma_3) \equiv f_B(\sigma_2)f_C(\sigma_1, \sigma_2, \sigma_3) \left( \left( \sum_{\sigma_4} f_D(\sigma_3, \sigma_4) \right) \left( \sum_{\sigma_5} f_E(\sigma_3, \sigma_5) \right) \right) \quad (5.5)
$$

In turn to compute $f_{BCDE}(\sigma_1, \sigma_2, \sigma_3)$ we need $f_B(\sigma_2)$, $f_C(\sigma_1, \sigma_2, \sigma_3)$ and $f_{DE}(\sigma_3, \sigma_4, \sigma_5) \downarrow \sigma_3$, where we have defined another shorthand

$$
f_{DE}(\sigma_3) \equiv \left( \sum_{\sigma_4} f_D(\sigma_3, \sigma_4) \right) \left( \sum_{\sigma_5} f_E(\sigma_3, \sigma_5) \right) \quad (5.6)
$$

Finally to compute $f_{DE}(\sigma_3, \sigma_4, \sigma_5)$ we need $f_D(\sigma_3, \sigma_4) \downarrow \sigma_3$ and $f_E(\sigma_3, \sigma_5) \downarrow \sigma_3$.

Figure 9: Factor graph for the example of BP done in this section.

Figure 10: Dependencies for the computation of the marginal $\mu_5(\sigma_1)$.
The previous decomposition of the dependences suggest a computation for $\mu_{5}(\sigma_{1})$ based on the idea of passing messages $\eta_{a \rightarrow i}$ and $\eta_{i \rightarrow a}$ on directed edges $f_{a} \rightarrow \sigma_{i}$ and $\sigma_{i} \rightarrow f_{a}$: this is illustrated in figure 11 (p.38) and reported in detail below.

1. Start with messages from the leaves: $\eta_{4 \rightarrow D}(\sigma_{4}) = 1$, $\eta_{5 \rightarrow E}(\sigma_{5}) = 1$

\[
\eta_{B \rightarrow 2}(\sigma_{2}) = f_{B}(\sigma_{2}) \downarrow \sigma_{2} = \sum_{\sigma_{j} \text{ s.t. } j \in \partial D \setminus \{3\}} f_{B}(\sigma_{2}) \cdot \eta_{j \rightarrow D} = \eta_{B \rightarrow 2}(\sigma_{2})
\]

\[
\eta_{A \rightarrow 1}(\sigma_{1}) = f_{A}(\sigma_{1}) \downarrow \sigma_{1} = \sum_{\sigma_{j} \text{ s.t. } j \in \partial A \setminus \{2\}} f_{A}(\sigma_{1}) \cdot \eta_{j \rightarrow A} = \eta_{A \rightarrow 1}(\sigma_{1})
\]

Notice that we can write $1 = \prod_{\emptyset} \prod_{j \in \partial B \setminus \{2\}} \eta_{j \rightarrow B} = \prod_{j \in \partial A \setminus \{1\}} \eta_{j \rightarrow A}$.

2a. Then proceed to compute messages

\[
\eta_{D \rightarrow 3}(\sigma_{3}) = f_{D}(\sigma_{3}, \sigma_{4}) \downarrow \sigma_{3} = \sum_{\sigma_{j} \text{ s.t. } j \in \partial D \setminus \{3\}} f_{D}(\sigma_{3}, \sigma_{4}) \cdot \eta_{4 \rightarrow D}(\sigma_{4})
\]

\[
\eta_{E \rightarrow 3}(\sigma_{3}) = f_{E}(\sigma_{3}, \sigma_{5}) \downarrow \sigma_{3} = \sum_{\sigma_{j} \text{ s.t. } j \in \partial E \setminus \{3\}} f_{E}(\sigma_{3}, \sigma_{5}) \cdot \eta_{5 \rightarrow E}(\sigma_{5})
\]

Notice that we can write $\eta_{1 \rightarrow D}(\sigma_{4}) = \prod_{j \in \partial D \setminus \{3\}} \eta_{j \rightarrow D}$ and similarly for $\eta_{5 \rightarrow E}$.
2b. Also compute message

\[ \eta_{2 \rightarrow C}(\sigma_2) = f_B(\sigma_2) \downarrow \sigma_2 = \eta_{B \rightarrow 2}(\sigma_2) \]  

(5.13)

Notice that we can write \( \eta_{B \rightarrow 2}(\sigma_2) = \prod_{b \in \partial_2 \setminus \{C\}} \eta_b(\sigma_2) \).

3. Continue by computing the message (see eq’n 5.6)

\[ \eta_{3 \rightarrow C}(\sigma_3) = f_{DE}(\sigma_3) \downarrow \sigma_3 = \prod_{d \in \partial_3 \setminus \{C\}} \eta_d(\sigma_3) \]  

(5.14)

4. Finally compute the message

\[ \eta_{C \rightarrow 1}(\sigma_1) = f_{BCDE}(\sigma_1, \sigma_2, \sigma_3) \downarrow \sigma_1 = \sum_{\sigma_j \text{ s.t.}} f_C(\sigma_1, \sigma_2, \sigma_3) \cdot \prod_{j \in \partial C \setminus \{1\}} \eta_j(\sigma_j) \]

\[ = \sum_{\sigma_2 = \pm} \sum_{\sigma_3 = \pm} f_C(\sigma_1, \sigma_2, \sigma_3) \cdot \eta_{2 \rightarrow C}(\sigma_2) \cdot \eta_{3 \rightarrow C}(\sigma_3) \]  

(5.15)

(5.16)

And now we’re done: all we need to do to compute the marginal \( \mu_5(\sigma_1) \) is observe that

\[ Z_5 \cdot \mu_5(\sigma_1) = \prod_{a \in \partial 1} \eta_{a \rightarrow 1}(\sigma_1) \]

\[ Z_5 = \sum_{\sigma_1 = \pm} Z_5 \cdot \mu_5(\sigma_1) \]  

(5.17)

One thing worth noting is that the messages passed on directed edges are functions, which might seem a little too abstract from an implementative point of view. What has to be kept in mind, though, is that the messages are functions \( \eta_{a \rightarrow 1}(\sigma_i) \) or \( \eta_{i \rightarrow a}(\sigma_i) \) of a binary spin variable \( \sigma_i \), and thus what we’re passing on are effectively pairs of real numbers.
5.2.2 The BP algorithm for tree factor graphs

The previous example clearly spells out what the BP algorithm for tree factor graphs should be. We want to compute $Z_N \cdot \mu_N(\sigma_t)$:

1. **Building the schedule:** the factor graph is explored breadth-first starting from variable node $\sigma_t$: whenever a node $v$ is explored as a child of some node $u$, we set a direction $u \leftarrow v$ to the edge $uv$. The directed graph (a tree) obtained at the end will have exactly one outgoing edge for all nodes, except for node $\sigma_t$ (the root), which has only incoming edges.

2. **Computing the messages:** starting from leaves (no incoming edges) of the tree, the outgoing messages of all nodes are computed, the outgoing message of node $u$ being computed only when all incoming messages have been already computed.

3. **Computing the marginal:** the exact marginal for $\sigma_t$ is computed as $Z_N \cdot \mu_N(\sigma_t) = \prod_{a \in \partial t} \eta_{a \rightarrow t}$.

The rules for computing the messages (also known as message-passing fixed point eq’ns) are:

$$
\eta_{a \rightarrow i}(\sigma_i) = \sum_{\sigma_j \text{ s.t. } j \in \partial a \setminus \{i\}} f_a(\sigma_{\partial a}) \cdot \prod_{j \in \partial a \setminus \{i\}} \eta_{j \rightarrow a}(\sigma_j)
$$

$$
\eta_{i \rightarrow a}(\sigma_i) = \prod_{b \in \partial i \setminus \{a\}} \eta_{b \rightarrow i}(\sigma_i)
$$

(5.18)

5.2.3 The BP algorithm for general factor graphs

The situation becomes more complicated for factor graphs with loops: there isn’t a natural way to build the schedule, and the computation will not be exact. On the other hand there are a number of well-studied scheduling algorithms for message passing, and a number of results on convergence of the computation to marginals: we’ll only cover the general principles here, and refer the reader to [15] for detailed descriptions and results.

1. **Building the schedule:** each edge of the factor graph is replaced with two directed edges (instead of one), and each directed edge is initialised with some random message\footnote{A random function of the spin involved in the edge, i.e. random pair of real numbers.} The exact details depend on the scheduling algorithm.

2. **Computing the messages:** the messages are updated with the rules of eq’n 5.18: all incoming edges have values on them at any time and are updated according to some scheduling algorithm\footnote{The simplest algorithm updates all messages in subsequent generations, but much more efficient schedules exists.}. This time the node $\sigma_t$ is not treated differently from the others.

3. **Computing the marginal:** an approximation to the marginal at $\sigma_t$ can be computed at any sweep of the algorithm as $Z_N \cdot \mu_N(\sigma_t) \approx \prod_{a \in \partial t} \eta_{a \rightarrow t}(\sigma_t)$: the approximation is expected to be accurate and convergent in the RS phase, but a better algorithm (Survey Propagation) will be required in the 1RSB and RSB phases.
5.2.4 The BP algorithm for K-sat

The functions on the factor nodes of K-sat all have the exponential form

\[ \mu_a(\sigma|\partial a) = \exp[-\beta W_a(\sigma|\partial a)] \] (5.19)

so it is convenient to express the messages in exponential form as well: all the multiplicative operations involved in BP message processing become additive operations over the exponents.

Concretely we write the messages as

\[ \eta_{i\rightarrow a}(\sigma_i) = \exp[\beta (h_{i\rightarrow a}\sigma_i + \omega_{i\rightarrow a})] \]
\[ \eta_{a\rightarrow i}(\sigma_i) = \exp[\beta (u_{a\rightarrow i}\sigma_i + \omega_{a\rightarrow i})] \] (5.20)

and the BP rules for updating the messages become

\[ h_{i\rightarrow a} = \sum_{b\in\partial i\setminus\{a\}} u_{b\rightarrow i} \]
\[ \exp[\beta u_{a\rightarrow i}\sigma_i] = \sum_{\sigma_j \text{ s.t. } j\in\partial a\setminus\{i\}} \exp \left[ \beta \left( -W_a(\sigma|\partial a) + \sum_{j\in\partial a\setminus\{i\}} h_{j\rightarrow a}\sigma_j \right) \right] \] (5.21)

Under eq’ns we consider the message passed on edge \( i \rightarrow a \) to be the cavity-field \( h_{i\rightarrow a} \), and the message passed on edge \( a \rightarrow i \) to be the cavity-bias \( u_{a\rightarrow i} \).

An interpretation for the BP messages is given by

\[ \frac{1}{z_{a\rightarrow i}} \eta_{a\rightarrow i}(\sigma_i) \approx \text{marginal law of } \sigma_i \text{ when all factor nodes in } \partial i\setminus\{a\} \text{ are removed.} \]
\[ \frac{1}{z_{i\rightarrow a}} \eta_{i\rightarrow a}(\sigma_i) \approx \text{marginal law of } \sigma_i \text{ when factor node } a \text{ is removed.} \] (5.22)

where \( z_{u\rightarrow v} = \eta_{u\rightarrow v}(+1) + \eta_{u\rightarrow v}(-1) \) is the normalisation constant. Thus only the parameters \( h_{i\rightarrow a} \) and \( u_{a\rightarrow i} \) are of relevance, as the \( \omega_{u\rightarrow v} \) can be eliminated through normalisation (but the message passing equations for normalised messages are more complicated).

The computation of the marginal for \( \sigma_i \) can then be written as

\[ \mu_N(\sigma_i) \approx \frac{\exp(\beta h_i\sigma_i)}{2 \cosh(\beta h_i)} \] (5.23)

where we have defined the local field at \( \sigma_i \) by

\[ h_i = \sum_{a\in\partial i} u_{a\rightarrow i} \] (5.24)

\[ ^{50}\text{Approximations are exact in tree factor graphs, by Markov property of } \mu_N \text{ (see Hammersley–Clifford theorem).} \]
\[ ^{51}\text{We’ll often write } \eta_{u\rightarrow v} \text{ when we don’t need to distinguish the direction of the message.} \]
5.3 BP guided decimation

Unless otherwise stated, this section is based on [28][27][29].

5.3.1 BP guided decimation algorithm

To understand the limits of BP, we will now see the BP-guided decimation algorithm. Starting with an instance $I_0$ of K-sat we proceed by decimation, i.e. we assign values to one spin at a time by running BP on a sequence of progressively simplified CNF formulae $I_0, I_1, ..., I_t, ..., I_N$. We’ll denote by $U_t$ to be the set of indices of spins that have been fixed after step $t$, and $\tau|_{U_t}$ the family, indexed by $U_t$, of values that we’ve assigned to those spins. The spins $\sigma|_{U_t}$ will be called the fixed spins/variables at time $t$.

Set $U_0 = \emptyset$, $\tau|_{U_0} = ()$ and for $t = 1, ..., N$ do:

1. choose a random $i_t \in V \setminus U_{t-1}$ and set $U_t := U_{t-1} \cup \{i_t\}$
2. run BP on the factor graph of $I_{t-1}$ to approximate the marginal of $\sigma_{i_t}$ in it:
   \[ \mu_{N-t+1}(\sigma_{i_t}) \equiv \mu_N(\sigma_{i_t}|\tau_{U_{t-1}}) \] (5.25)
3. fix the spin value $\tau_{i_t} := \pm 1$ with probability $\mu_{N-t+1}(\pm 1)$
4. simplify the CNF formula $I_t := I_{t-1}|_{\sigma_{i_t}=\tau_{i_t}}$

The algorithm can stop for only two reasons:

(a) at some point the partial assignment $\tau|_{U_{t-1}}$ is not compatible with any solution: in this case BP fails to compute the marginal $\mu_{N-t+1}(\sigma_{i_t})$ because $Z_{N-t+1}\mu_{N-t+1}(\sigma_{i_t}) = 0$ and cannot be normalised;

(b) a value has been assigned to all spins: this means that BP managed to compute all marginals, and thus the assignment $\sigma = \tau$ is a satisfying assignment for $I$.

In fact the mechanism that makes BP fail is equivalent to the Unit Clause Propagation (UCP) algorithm, which sequentially progressively simplifies all unit clauses in a CNF $I_t$:

1. pick a unit clause $C_a = \zeta_i \in \{x_i, \neg x_i\}$
2. set $x_i = 0, 1$ to satisfy clause $C_a$, i.e. to get $z_i = 1$
3. for all other clauses $C_b$
   (a) if $z_i$ appears in a clause $C_b$, remove clause $C_b$ from the CNF (as it is now satisfied)
   (b) if $\neg z_i$ appears in a clause $C_b$, remove $\neg z_i$ from $C_b$. If $\neg z_i$ was the only literal of $C_b$, return instance $I_t$ to be unsatisfiable (i.e. a contradiction has been discovered)

Running UCP until no unit clauses remain provides a family of spins whose values are directly implied at time $t$ by the fixed spins $\sigma|_{U_t}$. We’ll say that a spin $\sigma_i$ is frozen at time $t$ if either $i \in U_t$ or $\sigma_i$ is directly implied at time $t$, and denote by $W_t$ the set of indices of all frozen spins at time $t$ (thus $W_t \setminus U_t$ is the set of indices of all spins which are directly implied at time $t$).

.. footref{52}{Just for this specific instance we’ll go back to the boolean formulation.}
.. footref{53}{Slightly different concept of frozen from the one of section 4.6}
5.3.2 Frozen variables and success probability for BP

Unless specified otherwise, this section is based on [29].

Start by defining the set of **newly frozen spins/variables** at time $t$ by $\sigma|_{Z_t}$, where

$$Z_t \overset{def}{=} W_t \backslash W_{t-1} \quad \text{(5.26)}$$

Then $|Z_{\theta N}|$ is the amount of newly frozen variables when the fraction of fixed spins is $\theta$.

Now consider the subgraph $G_{\theta N}$ of the factor graph of $I_{\theta N}$ obtained by only considering variables in $Z_{\theta N}$ and their adjacent factor nodes: if $|Z_{\theta N}|$ remains bounded in the thermodynamic limit then $G_{\theta N}$ will w.h.p. be a tree (see [29]); on the other hand if $|Z_{\theta N}|$ diverges then $G_{\theta N}$ will w.h.p. contain loops (see [29][28]). The following theorem then connects the asymptotic behaviour of $Z_{\theta N}$ to the success probability of BP:

**Theorem 5.1.** If $G_t$ is a tree then no contradiction will arise at time $t$, i.e. UCP will find no contradiction. As a consequence if $|Z_{\theta N}|$ remains bounded in the thermodynamic limit then BP will succeed w.h.p. at fraction $\theta$ of fixed variables.

Finally define the **fraction of frozen spins/variables**

$$\phi(\theta) \overset{def}{=} \frac{1}{N} E |Z_{\theta N}| \quad \text{(5.27)}$$

to get, under reasonable regularity conditions, that

$$\lim_{N \to \infty} Z_{\theta N} = \frac{d\phi}{d\theta} \quad \text{(5.28)}$$

Thus BP succeeds w.h.p. in the region where $\phi(\theta)$ has bounded derivative, i.e. where the amount of newly frozen variables is bounded in the thermodynamic limit.

The *cavity method* provides a fixed point equation for $\phi(\theta)$ which can be approximately solved via iterative methods: for $\alpha$ low enough, any initial condition will converge to the same fixed point $\phi(\theta)$; for $\alpha$ high enough, on the other hand, the fixed point equation has 3 distinct solutions $\phi_-(\theta)$, $\phi_0(\theta)$ and $\phi_+(\theta)$, satisfying $\phi_+(\theta) \leq \phi_0(\theta) \leq \phi_-(\theta)$ at all $\theta$.

We’ll denote the value of $\alpha$ after which multiple fixed points appear by $\alpha_*$, and we will be interested in $\phi_{\pm}(\theta)$ only: careful analysis of the initial conditions for population dynamics that lead to $\phi_{\pm}(\theta)$ shows that $\phi_{+}(\theta)$ is the solution for the RS regime, while $\phi_{-}(\theta)$ is the solution for the 1RSB regime. Both $\phi_{\pm}(\theta)$ have vertical slope and are discontinuous at some fraction of fixed variables $\theta'_{\pm}$, with $\theta'_- \leq \theta'_+$. Figure 13 (p.46) shows $\phi(\theta)$ for a value $\alpha < \alpha_*$ and $\phi_{\pm}(\theta)$ for a value $\alpha > \alpha_*$. The bottom plot in figure 15 (p.48) shows, amongst other things, the lines $\theta'_{\pm}(\alpha)$ for 4-sat, and allows to estimate the critical point where they both originate to be $(\alpha_*, \theta_*) \approx (8.05, 0.35)$.

---

54 The expectation is taken over all sources of randomness, including the set of fixed variables.
55 We’ll present it in section 6, but we’ll not go over the computation of $\phi(\theta)$.
56 Or multivalued, depending on the point of view.
5.3.3 A condensation phase transition for the residual free-entropy density

Unless specified otherwise, this section is based on [29].

It would be neat if $\alpha^{\star}$ marked the point where BP stops being effective, but this is not the case: as we’ll see in section 5.3.4, the algorithm behaves well up to a higher constraint density $\alpha_{a}$, very near to another phase transition $\alpha^{(res)}_{c}$ which we’ll now explore.

Firstly we need to generalised the quenched average $p_{N}$ of the log-partition function to the so-called decimated regime, i.e. to ensembles where a fraction $\theta$ of the spins is allowed to be fixed: in this context it is called the residual free-entropy density (or simply residual entropy) and it is denoted by $\omega(\theta)$:

$$Z_{N}[\tau_{U}] \overset{def}{=} \sum_{\sigma} \prod_{a \in F} \exp[-\beta W_{a}(\sigma|\partial a)] \cdot 1_{\sigma|U=\tau_{U}}$$

$$\omega(\theta) \overset{def}{=} \lim_{N \to \infty} \frac{1}{N} \mathbb{E} \log Z_{N}[\tau_{U}]$$

(5.29)

where we have put $U \equiv U_{\theta N}$. Please note that $\omega(0) = \lim_{N \to \infty} p_{N}$ is the usual quenched average.

Computation of the residual entropy proceeds through the cavity method, and it turns out that it can be expressed entirely in terms of the fraction of frozen variables $\phi(\theta)$, i.e. as $\omega(\theta) = \hat{\omega}(\phi(\theta))$:

(a) in the region $\alpha < \alpha^{\star}$ the fixed point $\phi(\theta)$ is unique and there is no ambiguity for $\omega(\theta)$

(b) in the region $\alpha^{\star} < \alpha < \alpha^{(res)}_{c}$ there are multiple solutions for $\phi(\theta)$, but again there is no ambiguity for $\omega(\theta)$ because the expressions $\hat{\omega}(\phi_{\pm}(\theta))$ coincide

(c) in the region $\alpha > \alpha^{(res)}_{c}$ the two distinct expressions $\hat{\omega}(\phi_{\pm}(\theta))$ concur in determining $\omega(\theta)$ as

$$\omega(\theta) = \max[\hat{\omega}(\phi_{-}(\theta)), \hat{\omega}(\phi_{+}(\theta))].$$

(5.30)

Both $\hat{\omega}(\phi_{\pm}(\theta))$ are decreasing, and for $\alpha > \alpha^{(res)}_{c}$ they intersect at a unique point $\theta_{c}(\alpha)$, where the derivative of $\omega(\theta)$ develops a discontinuity: the point $\alpha^{(res)}_{c}$ marking the appearance of this discontinuity is called the condensation point of the residual entropy. Figure 14 (p.47) shows plots of the residual entropy for 4-sat below and above its condensation point $\alpha^{(res)}_{c} \approx 9.05$.

So what happens after $\alpha^{(res)}_{c}$ to $\omega(\theta)$, and why is it called a condensation point? Careful analysis of $\hat{\omega}(\phi_{\pm}(\theta))$ shows that for $\alpha > \alpha^{(res)}_{c}$ there exists a region $[\theta_{-}(\alpha), \theta_{+}(\alpha)]$ where both solutions coexists, and that

(a) $\omega(\theta) = \hat{\omega}(\phi_{+}(\theta))$ for $\theta \in [\theta_{-}(\alpha), \theta_{c}(\alpha)]$ (i.e. the RS solution dominates)

(b) $\omega(\theta) = \hat{\omega}(\phi_{-}(\theta))$ for $\theta \in [\theta_{c}(\alpha), \theta_{+}(\alpha)]$ (i.e. the 1RSB solution dominates)

57The expectation is taken over all sources of randomness, including the set of fixed variables.
Figure 15 (p.48) shows the three curves $\theta_{\pm,c}(\alpha)$ of 4-sat stemming from the critical point $(\alpha_{c}^{(\text{res})}, \theta_{c}^{*}) \approx (9.05, 0.045)$. The **condensation curve** $\theta_{c}(\alpha)$ touches $\theta = 0$ at $\alpha_{c} \approx 9.547$, asserting the status of $\theta_{c}(\alpha)$ as the condensation phase transition in the $(\alpha, \theta)$ plane. The **clustering curve** $\theta_{-}(\alpha)$ touches $\theta = 0$ at $\alpha_{d} \approx 9.38$, asserting the status of $\theta_{-}(\alpha)$ as the clustering phase transition in the $(\alpha, \theta)$ plane. These interpretations are confirmed by the following considerations.

In the region $[\theta_{-}(\alpha), \theta_{c}(\alpha)]$ the residual entropy is dominated by the RS solution $\phi_{+}(\theta)$, but the 1RSB solution exists and the complexity (at Parisi 1RSB parameter $m = 1$) can be computed as:

$$
\hat{\omega}(\phi_{+}(\theta)) - \hat{\omega}(\phi_{-}(\theta)) = \lim_{N \to \infty} \mathbb{E} \log \frac{Z_{+}}{Z_{-}} = \lim_{N \to \infty} \mathbb{E} \log \mathcal{N}(\omega(\theta)) = \Sigma(\omega(\theta); \alpha) \bigg|_{m=1}
$$

We thus have a well-defined, positive complexity, and recalling section 4.4 we conclude the region between the curves $\theta_{-}(\alpha)$ and $\theta_{c}(\alpha)$ to be the clustering phase of K-sat in the $(\alpha, \theta)$ plane.

In the region $[\theta_{c}(\alpha), \theta_{+}(\alpha)]$ the residual entropy is dominated by the 1RSB solution $\phi_{-}(\theta)$, and the complexity (at Parisi 1RSB parameter $m = 1$), still defined, becomes negative: recalling section 4.5 we conclude the region between the curves $\theta_{c}(\alpha)$ and $\theta_{+}(\alpha)$ to be the condensation phase of K-sat in the $(\alpha, \theta)$ plane.

---

**Figure 12: Success probability for BP-guided decimation as a function of $\alpha$ for random 4-sat.** The vertical line marks $\alpha_{c}^{(\text{res})}$, value beyond which a condensation phase transition appears for the residual entropy (see e.g. figure 15 (p.48)). Figure from [29].

---

58 I.e. it generalises the condensation phase transition from the condensation point on the $\alpha$ line to the condensation curve in the $(\alpha, \theta)$ plane.
Figure 13: Fraction of frozen variables $\phi(\theta)$ for random 4-sat, as a function of the fraction of fixed variables $\theta$. In the figure above we have $\alpha = 7.0 < \alpha_d^{(\text{res})}$ and there is only one curve for $\phi(\theta)$ (the solid line). In the figure below we have $\alpha = 8.4 > \alpha_d^{(\text{res})}$ and two curves appear for $\phi(\theta)$, superimposed on the range $\theta > \theta'_+$. The curve below (solid line) is the RS solution $\phi_+(\theta)$ and is discontinuous at $\theta'_+$, while the curve above (dashed line) is the 1RSB solution $\phi_-(\theta)$ and is discontinuous at $\theta'_-$ (where it starts being defined). For the dependence of $\theta'_\pm$ on $\alpha$ see figure 15 (p. 48). At small $\theta$, $\phi(\theta)$ is close to $\theta$ and its slope is close to 1: almost all frozen variables are just fixed variables, and the only newly frozen variables to be expected are the ones just fixed. At some intermediate value of $\theta$ the slope of $\phi(\theta)$ reaches its maximum (diverging in the bottom figure): enough variables have been fixed to induce sizeable cascades of unit clauses. At high $\theta$ the slope of $\phi(\theta)$ tends to zero: the expected number of newly frozen variables tends to zero, as most of the variables fixed are already directly implied. Figure from [29].
Figure 14: Plots of the residual entropy for random 4-sat. In the figure above we have $\alpha = 8.8 < \alpha_{c}^{\text{(res)}}$ and the residual entropy is a smooth function of $\theta$. In the figure below we have $\alpha = 9.3 > \alpha_{c}^{\text{(res)}}$ and a singular point appears (in detail in the inset). The dashed lines in the inset correspond to the continuation of the two solutions in their region of coexistence. Figure from [29].
Figure 15: Some phase transitions in the $(\alpha, \theta)$ plane of random 4-sat. In the figure above we have a plot of the three curves $\theta_-(\alpha)$, $\theta_c(\alpha)$ and $\theta_+(\alpha)$: notice the critical point $(\alpha^{(res)}_c, \theta^*_c) \approx (9.05, 0.045)$ in the top left corner. Also notice, in the bottom right corner, $\theta_-(\alpha)$ and $\theta_c(\alpha)$ touching $\theta = 0$ at $\alpha_d \approx 9.38$ and $\alpha_c \approx 9.547$ respectively. In the figure below we have a larger region of the $(\alpha, \theta)$ plane, showing the three $\theta_{\pm,c}(\alpha)$ curves in the bottom right corner and the $\theta_{\pm}'(\alpha)$ curves in the top half: notice the critical point $(\alpha^*_c, \theta^*_c) \approx (8.05, 0.35)$ in the top left corner. Figure from [29].
The algorithmic barrier for BP

Now that we’ve seen the clustering and condensation phases in the \((\alpha, \theta)\) plane\(^{59}\), we’re ready to explore the algorithmic barrier \(\alpha_a\) for BP, i.e. the set of results (both rigorous and experimental) connecting the success probability of BP to the geometry of its phase space. Firstly it’s interesting to consider the main hypothesis \(^{27}\) backing the introduction of BP in the context of decimation algorithms.

**Hypothesis 5.2.** Define \(\mu_{N-t+1}(\sigma_{it}; r)\) to be the marginal computed by BP when the factor graph of \(I_{t-1}\) is restricted to the ball of radius \(2r\) around node \(\sigma_{it}\). Then we expect BP to be effective in computing the marginal \(\mu_{N-t+1}(\sigma_{it})\) if the following conditions of locality hold:

1. \(\forall \epsilon > 0 \exists r_{\epsilon} \text{ s.t. } \forall t \left| \mu_{N-t+1}(\sigma_{it}) - \mu_{N-t+1}(\sigma_{it}; r_{\epsilon}) \right| \leq \epsilon\)
2. there is a function \(r(N)\) s.t. \(\mu_{N-t+1}(\sigma_{it}) \approx \mu_{N-t+1}(\sigma_{it}; r(N))\) w.h.p.

The validity of hypothesis 5.2 for K-sat is established in \(^{43}\) up to \(\alpha_u = \frac{2\log K}{K} + o(1)\) (called the Gibbs uniqueness phase transition for K-sat) by considering correlation decay in tree factor graphs, and using the fact that for \(1 < r \ll \log N\) the ball around any variable node is w.h.p. a tree.

Figures 12 (p.45) and 16 (p.50) show a common scenario in numerical experiments for BP-guided decimation over small \(K\): there is an \(\alpha_a\) near to the condensation point \(\alpha_c^{(res)}\) s.t.

1. the success probability is positively bounded below in the thermodynamic limit for all \(\alpha < \alpha_a\)
2. the success probability vanishes in the thermodynamic limit for all \(\alpha > \alpha_a\)

This has prompted the further hypothesis that, in the undecimated regime \(\theta = 0\), BP-guided decimation should perform well up to the condensation phase transition \(\alpha_c = 2^K \log 2 - \frac{3}{2} \log 2 + O(2^{-K})\). Unfortunately the following rigorous result from \(^{27}\) shows that this is just an artefact of the small \(K\) analysis\(^{60}\).

**Theorem 5.3.** *(Success probability of BP)*

There is a constant \(\rho > 0\) s.t. for all \(\rho \frac{2^K}{K} \geq \alpha \geq \rho \geq \frac{2^K \log 2}{2} \) we have

\[
\text{Success probability for } \text{BP-guided decimation over } I \leq \exp\left[-\Omega(N)\right] \text{ w.h.p.} \quad (5.32)
\]

Furthermore this is not an artefact of the specific implementation of the algorithm: hypothesis 5.2 fails as well for \(\rho \frac{2^K}{K} \geq \alpha \geq 2^K \log 2\), regardless of the choice of function \(r(N)\).

The proof of the theorem begins with the observation that, upon success, BP-guided decimation doesn’t just find a solution, but it finds a uniformly distributed one. The proof then pushes this observation further to an analysis of the BP operator on a random decimated instances of K-sat: these are shown to obey some specific quasi-randomness properties that turn out to be sufficient for a probabilistic analysis of the computation.

---

\(^{59}\)The \((\alpha, \theta)\) plane is the natural phase space for BP-guided decimation, and all other applications of BP to compute marginals of partially reduced instances of K-sat.

\(^{60}\)The only ones around: it’s really hard to do any numerical simulation even for \(K \geq 10\)
In conclusion an algorithmic barrier $\alpha_a = O(2^K/K)$ exists for BP on K-sat, exactly the same as for all other solvers (except for Fix). It is related to the condensation phase transition only for small $K$: for large $K$ the barrier not only doesn’t reach the condensation point, but in fact appears before the clustering phase transition $\alpha_d$.

Figure 16: Probabilities of non-convergence within 1000 iterations (panel above) and mean convergence time (panel below) for BP-guided decimation in random 4-sat, at various values of $\alpha$ approaching the (conjectured) algorithmic phase transition $\alpha_a \approx 9.05$. Both the probability of non-convergence and the mean convergence time peak at a value $\theta_{\text{max}}$, with sharper and taller peaks as $\alpha$ approaches $\alpha_a$. Figure from [29].

5.4 From BP to SP

The solution to computing marginals in the clustering and condensation phases is the Survey Propagation algorithm, which we’ll see in section 6.2: the clever idea behind it is to overcome the shattering of the solution space by considering the uniform measure over clusters rather than the uniform measure over spin configurations. Where BP computes global marginals, SP computes distributions of marginals restricted to clusters: the tool used to compute those distribution is the cavity method, to which we’ll now turn.

---

61 Which makes messages all the more complicated (and interesting), as now we have to pass distributions around.
6 The Cavity Method and Survey Propagation

6.1 The Cavity Method

Unless otherwise stated this section is based on [9][10][7][8]. The cavity method was developed in the study of spin glasses and is, at least in principle, equivalent to the replica method. It does, however, possess two key advantages over the latter:

(a) it is a standard probabilistic analysis of the factor graph, which makes explicit hypotheses on the correlation of the variable nodes. This makes it easier to turn into a rigorous method (or at least distinctly more rigorous than the replica method).

(b) the replica method starts by averaging over the disorder, while the cavity method works at fixed disorder (and only takes the average in the end, when statistical properties have to be computed with it). This makes it suitable for actual algorithms.

Consider a spin $\sigma_N$ connected to some number $k$ of factor nodes $f_{a_1},...,f_{a_k}$, each $f_a$ being connected to $K$ other spins $\sigma_{i(a,1)},...,\sigma_{i(a,K-1)}$. The spins $\sigma_{i(a,1)}$ are then very near to each other, but (and this is the key point) upon removal of $\sigma_N$ (a.k.a. creation of a cavity) their distance typically jumps up to $\log(N)$.

In the RS phase, where only one cluster exists, the correlation-decay property of pure states allows us to claim that large distance implies vanishing correlation, and to conclude that the spins $\sigma_{i(a,l)}$ are, upon removal of $\sigma_N$, uncorrelated in the thermodynamic limit: recalling the interpretation of eq’n 5.22 (p.41), one can then justify writing the marginal for $\sigma_N$ as the product of the $\eta_{\sigma_{i(a,l)}\rightarrow N}$, themselves defined in terms of the $\eta_{\sigma_{i(a,l)}\rightarrow i}$. This is the cavity method derivation of the BP message-passing eq’ns 5.18 (p.40): in fact the BP messages embody the cavity method itself, the induction process behind the latter being encoded in the fixed-point equations governing the formers. From now on we’ll phrase all our results in terms of messages (cavity-fields and cavity-biases).

In the d1RSB and 1RSB phases, on the other hand, one cannot claim a connection between large distance and vanishing correlation, as shown in section 4.3. The issue is the presence of multiple pure states, and the solution is to apply the cavity method while keeping separate messages for each pure state: this is equivalent to running BP on each pure state via the following procedure (which is given in more detail in [8]):

1. restrict the factor graph to a tree-like neighborhood $T(\sigma_N)$ of $\sigma_N$, and let $B$ be its boundary.

2. impose boundary conditions to the messages from boundary nodes in $B$ that isolate a single pure state $\psi$.

3. run BP on $T(\sigma_N)$ (constrained to the boundary conditions) and obtain a family of cavity-fields $(h_{i\rightarrow a}^{(\psi)})_{i\psi}$ and of cavity-biases $(u_{a\rightarrow i}^{(\psi)})_{i\psi}$ for the pure state $\psi$.

Each edge $i \rightarrow a$ of the factor graph will then have a family of cavity-fields $(h_{i\rightarrow a}^{(\psi)})_{i\psi}$ and cavity-biases $(u_{a\rightarrow i}^{(\psi)})_{i\psi}$ indexed by the pure states, and eq’n 5.24 (p.41) will compute $\mu_{N}^{(\psi)}$.

---

62 A poisson r.v. with mean $\alpha K$ (in the thermodynamic limit).

63 And thus strongly correlated.

64 In the RS phase. Notice that this derivation starts failing at $\alpha_d$, way before the clustering phase transition.
6.1.1 From messages to surveys

The first step in dealing with the multi-cluster scenarios of d1RSB and 1RSB is to take the BP message passing eq’s 5.21 (p.41) and reinterpret them in a probabilistic way, by considering them as the deterministic case of more general message-passing eq’ns involving probability distributions. The idea is that the BP messages are accurate descriptions of the marginals when restricted to a single pure state (which has the right correlation decay properties by definition), and that the case of many pure states can be treated by passing around PDFs of messages over states.

One then changes the messages into:

\[ P_{i \rightarrow a}(h) \overset{def}{=} \text{probability density for a cavity-field } h^{(\psi)}_{i \rightarrow a} \text{ to take value } h \]
\[ Q_{a \rightarrow i}(u) \overset{def}{=} \text{probability density for a cavity-bias } u^{(\psi)}_{a \rightarrow i} \text{ to take value } u \]  

(6.1)

and recovers BP as the 1-state, deterministic case

\[ P_{i \rightarrow a}(h) = \delta(h - h^{(\psi)}_{i \rightarrow a}) \]
\[ Q_{a \rightarrow i}(u) = \delta(u - u^{(\psi)}_{a \rightarrow i}) \]  

(6.2)

The \( P_{i \rightarrow a}(h) \) messages are called h-surveys, while the \( Q_{a \rightarrow i}(u) \) messages are called u-surveys.

The BP message-passing fixed-point eq’ns 5.21 (p.41) (or rather their normalised version) then become the 1-state, deterministic limit of the following 1RSB distributional fixed-point equations for surveys:

\[
P_{i \rightarrow a}(h) = \frac{1}{Z_i[(Q_{b \rightarrow i})_b]} \int \prod_b dQ_{b \rightarrow i}(u_{b \rightarrow i}) \delta \left( h - \sum_b u_{b \rightarrow i} \right) z_i[(u_{b \rightarrow i})_b]^m
\]
where we have taken \( b \) to range over \( \partial i \setminus \{ a \} \)

\[
Q_{a \rightarrow i}(u) = \frac{1}{Z_a[(P_{j \rightarrow a})_j]} \int \prod_j dP_{j \rightarrow a}(h_{j \rightarrow a}) \delta \left( u - f[(h_{j \rightarrow a})_j] \right) z_a'[(h_{j \rightarrow a})_j]^m
\]
where we have taken \( j \) to range over \( \partial a \setminus \{ i \} \)

(6.3)

where \( (u_{a \rightarrow i} - f[(h_{j \rightarrow a})_j]) \) is obtained from eq’n 5.21 (p.41). The effect of the Parisi 1RSB parameter \( m \) is the same as in the replicated free-entropy density of section 4.1 (p.25): it favours clusters described by a specific 1RSB solution, concentrating the measure over them.

\( Z_i \) and \( Z_a \) are normalisation factors and we have defined the following

\[
z_i[(u_{b \rightarrow i})_b] \overset{def}{=} \exp \left[ + \sum_b u_{b \rightarrow i} + \exp \left[ - \sum_b u_{b \rightarrow i} \right] \right] \prod_b \cosh u_{b \rightarrow i}
\]
\[
z_a'[(h_{j \rightarrow a})_j] \overset{def}{=} 1 + \exp \left[ -2f[(h_{j \rightarrow a})_j] \right]
\]

(6.4)

Note that \( P_{i \rightarrow a}(h) \) and \( Q_{a \rightarrow i}(u) \) are themselves r.v.s (depending on the r.v. \( I \)): we define \( P(P) \) and \( Q(Q) \) to be their respective PDFs, and will abuse notation by writing \( P(h) \) and \( Q(u) \) for the distributions of cavity-fields and cavity-biases after having taken the randomness of \( I \) into account.
6.1.2 The cavity method: free-entropy density, overlaps and marginals

The quantities $z_i, Z_i$ and $Z_a$ are called **free-entropy shifts**, and contribute to the cavity method formulation of the RS and 1RSB free-entropy densities, which we’ll now see.

The cavity method estimate for the **RS free-entropy density** is

$$\omega = -\alpha K \mathbb{E} \log z_{ia}(h, u) + \alpha \mathbb{E} \log z_a(h_1, ..., h_K) + \mathbb{E} \log z_i(u_1, ..., u_k)$$  \hspace{1cm} (6.5)

where $k$ is poisson of mean $\alpha K$, and $u, u_1, u_2, ...$ are iid distributed according to $Q(u)$, and $h, h_1, h_2, ...$ are iid distributed according to $P(h)$. The **RS free-entropy shifts** $z_{ia}, z_a$ and $z_i$ are defined by eq’n 6.4 and the following

$$z_{ia}(h, u) \overset{\text{def}}{=} 1 + \tanh h \tanh u$$

$$z_a(h_1, ..., h_r) \overset{\text{def}}{=} 1 - \prod_{i=1}^{r} \frac{\exp[-h_i]}{2 \cosh[h_i]}$$  \hspace{1cm} (6.6)

The cavity method estimate for the **1RSB (replicated) free-entropy density** is

$$\Phi(m) = -\alpha K \mathbb{E} \log Z_{ia}(P, Q) + \alpha \mathbb{E} \log Z_a(P_1, ..., P_K) + \mathbb{E} \log Z_i(Q_1, ..., Q_k)$$  \hspace{1cm} (6.7)

where $k$ is poisson of mean $\alpha K$, and $Q, Q_1, Q_2, ...$ are iid distributed according to $Q(Q)$, and $P, P_1, P_2, ...$ are iid distributed according to $P(P)$. The **1RSB free-entropy shifts** $Z_{ia}, Z_a$ and $Z_i$ are defined by eq’n 6.4 and the following

$$Z_{ia}(P, Q) \overset{\text{def}}{=} \int dP(h) dQ(u) \ z_{ia}(h, u)^m$$

$$Z_a(P_1, ..., P_r) \overset{\text{def}}{=} \int \prod_{i=1}^{r} dP_i(h_i) \ z_a(h_1, ..., h_r)^m$$

$$Z_i(Q_1, ..., Q_l) \overset{\text{def}}{=} \int \prod_{a=1}^{l} dQ_a(u_a) \ z_i(u_1, ..., h_l)^m$$  \hspace{1cm} (6.8)

Furthermore the surveys allow computation of the overlaps: the intra-state overlap of the dominant RS cluster is given by

$$q_{RS} = \mathbb{E} \ \tanh^2 h$$  \hspace{1cm} (6.9)

while the overlaps for the d1RSB and 1RSB phases are given by

$$q_0(m) = \mathbb{E} \left( \int \tanh h \ dP(h) \right)^2$$

$$q_1(m) = \mathbb{E} \int \tanh^2 h \ dP(h)$$  \hspace{1cm} (6.10)

Finally we can obtain the **surveys of local fields** as

$$P_i(h) = \frac{1}{Z_i[Q_{b \rightarrow i}, h]} \int dQ_{b \rightarrow i}(u_{b \rightarrow i}) \ \delta \left( h - \sum_b u_{b \rightarrow i} \right) \ z_i[(u_{b \rightarrow i})^m]$$  \hspace{1cm} (6.11)

where we have taken $b$ to range over $\partial i$ this time.
6.2 Survey Propagation

Unless otherwise stated, this section is based on the seminal papers [9][10].

6.2.1 Encoding the surveys

We want to use the surveys in an algorithm, but how do we practically carry PDFs around? The trick is to work in the zero temperature limit $\beta \to \infty$, where the second rule of 5.20 (p.41) simplifies to a minimisation operation. The cavity-biases are then only allowed to take a finite number of standard values \(^65\) (e.g. in 3-sat at zero temperature we have $u_{a\to i} \in \{0, \pm 1\}$) and the u-surveys can be encoded as histograms.

6.2.2 The SP algorithm

The SP algorithm is an evolution of BP, adapted to the multi-cluster scenario of the d1RSB and 1RSB phases: the messages passed around are now the u-surveys $Q_{a\to i}(u)$, and instead of computing a local field $h_i$ (which gave us a marginal via eq’n 5.24 (p.41)) we compute a distribution $P_i(h)$ of local fields (which gives us a distribution of marginals). The algorithm works at a fixed value of the Parisi 1RSB parameter $m$, i.e. focusing the measure on a specific family of clusters.

0. Initialise all the u-surveys at random
1. Select a random factor node $a$
2. For each $i \in \partial a$, compute the h-survey $P_{i\to a}(h)$ by using eq’n 6.3 (p.52)
3. For each $i \in \partial a$, update the u-survey $Q_{a\to i}(u)$ by using eq’n 6.3 (p.52)
4. Test convergence. If convergence is not reached, go to step 1. Otherwise:
   (a) compute the surveys of local fields $(P_i(h))_i$ by using eq’n 6.11 (p.53) and the u-surveys which were just computed
   (b) compute the free-entropy density $\Phi(m)$ by using eq’n 6.7 (p.53), the u-surveys and the h-surveys which were just computed
   (c) return $(P_i(h))_i$ and $\Phi(m)$

The SP algorithm can be used in a decimation procedure, called SID, similar to BP-guided decimation: steps 2 and 3 of the procedure from section 5.3.1 are changed into

2. Run SP, obtain the surveys $(P_i(h))_i$ and for each survey compute $w_i^+, w_i^0$:

$$w_i^+ = \int_{0^+}^{+\infty} dP_i(h), \quad w_i^- = \int_{-\infty}^{-} dP_i(h), \quad w_i^0 = 1 - w_i^+ - w_i^- \quad (6.12)$$

3a. Check if the system is in the paramagnetic phase (i.e. for every spin $\sigma_i$ we have $w_i^0 = 1$).
   If it is, try running a some fast local search algorithm (like simulated annealing or Walksat), and if the algorithm finds a solution return it. Otherwise proceed to step 3b

3b. Select and fix the \(^66\) most biased spin (i.e. largest $|w_i^+ - w_i^-|$)

\(^65\)If the random initialisation of BP is done within those standard values, that is.

\(^66\)Or a random one if more than one exists with same bias.
6.2.3 Convergence of SP

SP converges to different solutions depending on the value of \( m \) used, i.e. depending on which clusters we are concentrating our measure onto. It will usually be necessary to run SP at different values of \( m \) to find the best solution: the most efficient way to do this is to start at a high value of \( m \) (like \( m = 1 \) or a bit above \( m_s \), if known) and progressively lower the value of SP, using the u-surveys computed at higher-\( m \) runs as initial condition for lower-\( m \) runs, to speed up convergence. The value of \( m \) describing the thermodynamically relevant clusters is then, as usual, the one minimising \( \Phi(m)/m \) (where we use the estimate for \( \Phi(m) \) returned by SP). The following experimental results about convergence are reported in [10].

For \( \alpha < \alpha_{d, +} \) SP always converges to the trivial paramagnetic solution \((Q_{a \rightarrow i}(u) = \delta(u) \text{ and } P_i(h) = \delta(h) \text{ for all } i)\), irrespective of the value of \( m \): this is to be expected, as there is only one pure state. Fast local search algorithms can be used effectively in this phase.

For \( \alpha_d < \alpha < \alpha_c \) the behaviour depends on \( m \):

(a) for \( m \) low enough the algorithm converges to the trivial paramagnetic solution.

(b) for \( m \) near enough to 1 the algorithm converges to a unique non-trivial solution, describing the surveys amongst the thermodynamically dominating clusters.

For \( \alpha_c < \alpha < \alpha_s \) the behaviour depends even more strongly on \( m \):

(a) for \( m \) low enough the algorithm converges to the trivial paramagnetic solution.

(b) for intermediate values of \( m \) the algorithm converges to a unique non-trivial solution, describing the surveys amongst the thermodynamically dominating clusters.

(c) For larger values of \( m \) the algorithm stops converging. The range of values for which the algorithm converges to a non-trivial solution is reported to be sufficient, in the numerical experiments of [10], for the free-entropy density to converge to the value expected for the thermodynamically relevant clusters.

7 Conclusion

Starting from the Sherrington-Kirkpatrick model of spin glass, we have presented the main contributions given, between the years 2002 and 2010, to the application of spin glass theory to understanding the ensemble properties of the solution space of K-sat. From the Hamiltonian formulation to the condensation phase transition, we have reviewed, amalgamated and consolidated a decade of work to provide a global understanding of the achievements and potential of this field.

An earlier version of this work was presented as the author’s 2013 Part III Essay at DAMTP, University of Cambridge.

---

67 There they are reported as a function of the parameter \( y = \frac{\partial C}{\partial \epsilon} \), where \( \epsilon \) is the density of violated clauses, and a (fairly straightforward) rephrasing in terms of the parameter \( m \) has been done.

68 It’s reasonable to assume this means for values of \( m \approx m_s \), but the authors don’t make this observation explicitly.

69 Not counting the spin glass works it’s based on, which span almost 30 years.
References

[-] **Topic survey references:**

[1] D Achlioptas, *Random satisfiability*, Handbook of Satisfiability 185 (2009).

[2] A Coja-Oghlan, *A statistical mechanics perspective on hard computational problems*, Slides online for the Warwick Statistical Mechanics Seminar of 11 Feb 2010.

[3] M Talagrand, *Mean field models for spin glasses*, Springer (2010).

[4] M Talagrand, *Spin glasses: A challenge for mathematicians*, Springer (2003).

[5] M Mézard, G Parisi, M A Virasoro, *Spin glass theory and beyond*, (1987).

[-] **Core references:**

[6] D Achlioptas, A Coja-Oghlan, *Algorithmic barriers from phase transitions*, Foundations of Computer Science, IEEE 49th Annual IEEE Symposium on (2008).

[7] F Krzakala, A Montanari, F Ricci-Tersenghi, G Semerjian, L Zdeborova *Gibbs states and the set of solutions of random constraint satisfaction problems* P of the Nat. Acad. of Sc. (2007).

[8] A Montanari, F Ricci-Tersenghi, G Semerjian *Clusters of solutions and replica symmetry breaking in random k-satisfiability*, Journal of Statistical Mechanics: Theory and Experiment (2008)[version used: http://arxiv.org/abs/0802.3627v2].

[9] M Mezard, G Parisi, R Zecchina, *Analytic and algorithmic solution of random satisfiability problems*, Science Vol 297 (2002).

[10] M Mezard, R Zecchina, *The random K-satisfiability problem: from an analytic solution to an efficient algorithm*, Physical Review E (2002) [version used: http://arxiv.org/abs/cond-mat/0207194v3].

[-] **References on Statistical Physics and Gibbs measures:**

[11] G R Grimmett, *A theorem about random fields*, B of the London Mathematical Society (1973).

[12] R Monasson, R Zecchina. *Statistical mechanics of the random K-satisfiability model* (1997).

[13] D Tong, *Lectures on Statistical Physics*, Notes online from the author’s DAMTP teaching page.

[14] H O Georgii, *Gibbs measures and phase transitions*, Walter de Gruyter (2011).

[15] F R Kschischang, B J Frey, H A Loeliger, *Factor graphs and the sum-product algorithm*, IEEE Transactions on Information Theory (2001)

[-] **Additional references on Solution Space Geometry:**

[16] D Achlioptas , F Ricci-Tersenghi *On the solution-space geometry of random constraint satisfaction problems*, P of the 38thannual ACM symposium on Theory of computing (2006).

[17] D Achlioptas, *Solution Clustering in Random Satisfiability*, B of the Europ. Phys. J (2008).
D Achlioptas, F Ricci-Tersenghi, Random formulas have frozen variables, SIAM Journal on Computing (2009).

**Additional references on Replica Symmetry Breaking:**

- F Denef, String glasses, Slides available online
- F Denef, TASI lectures on complex structures, http://arxiv.org/pdf/1104.0254.pdf (2011).
- G Parisi, Order Parameter for Spin-Glasses, Physical Review Letters (1983)
- G Parisi, Infinite Number of Order Parameters for Spin-Glasses, Physical Review L (1979)
- G Parisi, The order parameter for spin glasses - A function on the interval 0-1, (1979)

**Additional references on Belief Propagation, Cavity Method, Survey Propagation:**

- E Maneva, E Mossel, M J Wainwright, A New Look at Survey Propagation and its Generalizations, (2005) [version used: http://arxiv.org/abs/cs/0409012v3].
- A Braunstein, R Zecchina, Survey propagation as local equilibrium equations, Journal of Statistical Mechanics: Theory and Experiment, IOP Publishing (2004) [version used: http://arxiv.org/abs/cond-mat/0312483].
- M Mezard, A Montanari, Information, Physics, and Computation
- A Coja-Oghlan, On Belief Propagation Guided Decimation for Random k-SAT, http://arxiv.org/abs/1007.1328v1 (2010)
- A Montanari, F Ricci-Tersenghi, G Semerjian, Solving constraint satisfaction problems through belief propagation-guided decimation, http://arxiv.org/abs/0709.1667v2 (2007)
- F Ricci-Tersenghi, G Semerjian, On the cavity method for decimated random constraint satisfaction problems and the analysis of belief propagation guided decimation algorithms, J. Stat. Mech. (2009)
- A Coja-Oghlan, A Better Algorithm for Random k-SAT, SIAM J Comput (2010)
- M Mézard, G Parisi, The cavity method at zero temperature, Journ of Statistical Physics (2003).

**Other references:**

- M. Mezard, M Virasoro, The microstructure of ultrametricity, Journal de Physique (1985)
- M Mezard, G Parisi, N Sourlas, G Toulouse, M Virasoro, Nature of the spin-glass phase, Phisical Review Letters (1984)
- A Crisanti, T Rizzo, Analysis of the infinity-replica symmetry breaking solution of the Sherrington-Kirkpatrick model, Physical Review E, APS (2002)
- D Anninos, T Anous, J Barandes, F Denef, String Glasses (to appear in the future)
- D Sherrington, S Kirkpatrick, Solvable Model of a Spin-Glass, Physical Review Letters (1975)
[37] L Zdebrodova, F Krzakala, *Phase transitions in the coloring of random graphs*, P Rev. E (2007)

[38] L Zdeborova, F Krzakala, *Hiding Quiet Solutions in Random Constraint Satisfaction Problems*,
http://arxiv.org/abs/0901.2130v2

[39] M Mezard, M Palassini, O Rivoire *Landscape of solutions in constraint satisfaction problems*,
http://arxiv.org/abs/cond-mat/0507451v2

[40] C H Papadimitriou, *On selecting a satisfying truth assignment*, FOCS (1991)

[41] M Alekhnovich, E Ben-Sasson, *Linear upper bounds for random walk on small density random 3-CNFs* SIAM J. Comput. (2007)

[42] J S Yedidia, W T Freeman, Y Weiss, *Understanding Belief Propagation and its generalizations*, IJCAI (2001)

[43] A Montanari, D Shah, *Counting good truth assignments of random k-SAT formulae*, Proc. 18th SODA, (2007)

[44] D Achlioptas, Y Peres, *The threshold for random K-SAT is $2^K \log 2 - O(k)$*, J AMS (2004).

[45] R C Griffiths, *On the distribution of points in a Poisson Dirichlet process*, J Appl Probab (1988).

[46] R C Griffiths, *On the distribution of points in a Poisson Dirichlet process*, J Appl Probab (1988).

[47] Durrett, *Probability models for DNA sequence evolution*, Springer (2008)