Replica and dynamics, similarities and differences.

Andrea Mazzei

October 26, 2015

Abstract

Aim of this short project is to give an overview of the most important analytical techniques used to deal with mean field spin glasses. The final sections also will present some numerical simulations.

1 Replica approach

1.1 Introduction to replicas

First of all we need to introduce the replica method, which is a tool used in the framework of disordered systems to integrate out over the disorder distribution function. Consider a generic spin system described by a Hamiltonian that's disorder dependent, and a partition function given by:

$$Z_J = \sum_{\{S\}} e^{-\beta \mathcal{H}_J} \tag{1}$$

where the pedix J describes the disorder dependance. If we want to evaluate the thermodynamic properties of the system we first have to evaluate the free energy functional, $-\beta F_J = \log Z_J$, and then perform the average over the disorder distribution (the so called quenched average). If the evaluation of $\langle F \rangle_J$ is unaccessible but we can evaluate every moment of the disorder-integrated partition function $\{Z, Z^2, \ldots, Z^n\}$ we can obtain $\langle F \rangle_J$ using a known limit.

$$-\beta \langle F \rangle_J = \langle \log Z_J \rangle = \lim_{n \to 0} \frac{\langle Z_J^n \rangle_J - 1}{n} \tag{2}$$

which is the essence of the replica method. When we evaluate $\langle Z^n \rangle_J$ we are generating n non interacting copies of the system, which become correlated after averaging over the disorder distribution.

Replica method (sometimes called replica trick) has been successfully applied to a variety of disordered systems, the most notable examples are the Sherrington-Kirkpatrick model [Parisi] and to the spherical p-spin model [Crisanti Sommers, Leuzzi].

1

In this first part we will concentrate on the spherical p-spin model, which is a RS-1RSB, depending on p, system. Consider the Hamiltonian

$$\mathcal{H} = \sum_{(i_1 \ldots i_p)} J_{i_1 \ldots i_p} S_{i_1} \ldots S_{i_p} - \sum_i h_i S_i \tag{3}$$

where each J has been taken by a random distribution and the summation on $(i_1 \ldots i_p)$ is a sum over each p-uplet of the system. The dynamical variables S_i are constrained by $\sum S_i^2 = N$. We will start by describing shortly the case p = 2.

1.2   RS 2-spin system

The Hamiltonial for the 2 spin system is:

$$\mathcal{H} = \sum_{(ij)} J_{ij} S_i S_j - \sum_i h_i S_i \tag{4}$$

The partition function for replicated system will be the product of each single system partition function.

$$\langle Z^n \rangle = \sum_{\{S\}^n} e^{-\beta \sum_a^n H[S^a]} \tag{5}$$

If we suppose Gaussian distributed disorder we can integrate over its distribution using an Hubbard Stratonovich integration:

$$\langle Z^n \rangle = \int \prod_{(ij)} dJ_{ij} \sum_{\{S\}^n} e^{-\beta \sum_a^n H[S^a] - \frac{J_{ij}^2}{2N}} \tag{6}$$

$$= \sum_{\{S\}^n} e^{(\beta \sum_a \sum_i h_i S_i^a)} e^{\frac{\beta^2}{4N^{p-1}} \sum_{i_1 \ldots i_p} \sum_{ab} S_{i_1}^a \ldots S_{i_p}^a \ldots S_{i_1}^b \ldots S_{i_p}^b} \tag{7}$$

The trace over spin states can be evaluated transforming the spherical constraint into its integral representation

$$\sum_{\{S\}^n} [\ldots] = \int \prod_{(ia)} dS_a \delta(\sum_i S_i^2 - N)[\ldots] = \int \frac{d\lambda_a}{\sqrt{2\pi}} e^{i\lambda_a(\sum_i S_i^2 - N)} \tag{8}$$

Now we may evaluate the integrals by saddle point methods and obtain the values of q_{ab} and λ_a. If we adopt the RS ansatz for this calculation, i.e. $q_{ab} = q*$ and $\lambda_a =?$ we get the RS solution of this model.

By breaking this RS ansatz one should observe that RSB is not present. This could be either achieved by evaluating the eigenvalues of the free energy's Hessian matrix or noticing that a one-step RSB ansatz would predict the same results of the RS one.

2

1.3 RSB p-spin systems

If we try the same evaluation scheme for the generic $p > 2$ spin model we have to notice that now the quadratic form in S is missing and is replaced by a p-power term. So we start by rewriting

DA FINIRE

$$\langle Z^n \rangle = \int \prod_{(ij)} dJ_{ij} \sum_{\{S\}^n} e^{-\beta \sum_a^n H[S^a] - \frac{J_{ij}^2}{2N}} = \dots \tag{9}$$

When we write down the free energy function we see that RS solution is now unstable: this can done both by evaluating the Hessian matrix eigenvalues and going into a one-step RSB ansatz and see if it returns a different solution from the RS one.

1.4 Few words on Ising p-spin

The framework is very different for the Ising p-spin. It has been shown in literature [MISSING] that this model in zero field presents three different phases. One at high temperatures, where the RS solution is correct and essentially returns the paramagnetic solution. One which appears to be 1 RSB, at intermediate temperatures, and one which is full RSB, just like the Sherrington Kirkpatrick model. The 1RSB solution is typical of p-spin systems (with p greater than 2) and seems to be connected with the nonlinearity of associated Langevin equation, that we will describe below.

2 Dynamic approach

The dynamic approach is the application of Langevin dynamics to the disordered spin models.

2.1 Few words on Langevin dynamics

Consider the Hamiltonian discussed before. If we want to implement a dynamic for the observables S_i we may use the Langevin equation

$$\frac{dS_i}{dt} = -\frac{\partial \mathcal{H}}{\partial S_i} + \eta_i \tag{10}$$

where η is a random noise. At this point we have not to confuse η with the intrinsic disorder of the system (which is inside the definition of Hamiltonian). Noise is now delta correlated, i.e. $\bar{\eta} = 0$, $\overline{\eta_i(t)\eta_j(t')} = 2T\delta(t - t')\delta_{ij}$. It is also necessary to distinguish the average over the disorder from the noise average: the latter will always be performed with an overline, just like the previous formula.

The noise distribution can be formally written as:

$$P[\eta] = \mathcal{N} e^{-\frac{1}{2} \int dt \int dt' \eta(t) D(t,t') \eta(t')} \tag{11}$$

where D is the inverse of the variance of the process.

We can evaluate the average of a generic operator using the integral representation for the Langevin operator

$$
\langle O(S) \rangle = \int D[\eta]P[\eta] \int D[S]\delta(S - S_\eta)O(S) \tag{12}
$$

$$
= \int D[S]\delta(S - S_\eta)O(S) \int D[\eta]P[\eta]\delta(\mathcal{L}[\mathcal{S}] - \eta) \tag{13}
$$

$$
= \int D[S] \int D[\hat{S}]O(S) \int D[\eta]P[\eta]e^{i \int dt\hat{S}[\mathcal{L}[\mathcal{S}] - \eta]} \tag{14}
$$

It is important to remark that the Langevin operator is still disorder- dependent. When averaging over the disorder distribution the integrals in the previous formula become connected and it becomes possible to write down a new Langevin dynamic with noise ξ, where now ξ is not δ correlated. A pratical example, with most of calculations omitted, will be presented in the next section.

2.2 Long time dynamics for the p-spin spherical spin glass

If we evaluate Langevin dynamic for the spherical p-spin model first of all we can integrate over the spherical constraint, similarly to what we've done in the replica framework. In these lines we will follow the calculations presented in [1]. We will skip passages and go straight to the point.

$$
\delta(\sum_i S_i^2 - N) = \int d\mu(t)e^{i\mu(\sum_i S_i^2 - N)} \tag{15}
$$

At this point μ has to be computed in a self-consistent way.

The dynamic equation now reads

$$
\frac{dS_i}{dt} = -\frac{\partial \mathcal{H}}{\partial S_i} - \mu(t)S_i + \eta_i \tag{16}
$$

After a bit of calculation we can get the time evolution equation for the correlation function

$$
EQTNMISSING
$$

LOTS OF EQTNS MISSING (17)

When combining this equation with the criterion that C is time-decreasing at any given time, we get that the inequality cannot hold if the temperature is lower than T_d, with

$$T_d = \sqrt{\frac{p(p-2)^{p-2}}{2(p-1)^{p-1}}} \qquad (18)$$

which is always higher than the static temperature mentioned in the replica approach. The meaning of this picture is that at a temperature in the T_d neighborhood, metastable states (thus not weighing in the static measure) dominate the dynamic of the system. In this sense, the T_d transition is purely dynamical.

At this point an observation is important. Both the dynamical temperature and the static one are, in the 2-spin spherical mean field spin glass, equal to zero.

3   Common Ground

So far we have described two methods used to approach to the study of a disordered system. One is static, its aim is to find the thermodynamic of the systems by explicitly evaluating the average over the boltzmann measure and the disorder distribution. The information on eventual transitions and other critical phenomena are encoded in the usual thermodynamic functions, such as entropy and free energy. The other is dynamical, meaning that we are searching for transitions and other critical phenomena observing the time evolution of specific correlation functions.

In the high temperature phases the two methods coincide. In the low temperature phase both present interesting issues and phenomena, which essentially decode some aspects of the same complex behaviour, though very important differences begin to arise. We saw that, for example, the replica approach fails to determine the presence of the dynamical transition.

MANCA UN COMMENTO SULLA DINAMICA

4   RSB and Ergodicity breaking in simulations

In the next section I will present some numerical simulations on the Ising p-spin system. The choice of a different spin system is purely for sake of simplicity in the simulation. If we'd want to simulate a system with the spherical constraint (which is of course a continous system) we'd have to take into account some additional problematics:

- implement a dynamic which preserves the constraint

  Our Montecarlo simulation will have to satisfy the constraint for each move. This can be done in two ways: either by satisfying the constraint at each move or by implement a soft spin version which has our original constraint as a limit. The first method must of course update at least two spins (or more) simultaneously, and I was not sure about nor the details of convergence of the algorithm nor the satisfiability of detailed balancing, even though it seems that this method was already been investigated in

5

literature. The second method is indeed easier to implement. It could be a choice but it still soffers two issues described below these lines.

- Mean field systems require more memory and much more operations

  A connectivity tensor of a system with N sites requires an $NN \dots N$ matrix with inside the information on the quenched disorder. If we insted consider a random graph with Z neighbors per site, we could store the same information in two matrices $N(Z-1)P$, name it C and J, where C remembers the label of each neighbor present in the p-uplet and each p-uplet, and J has inside the numerical value o of the disorder.

  Furthermore, any update in mean field requires the evaluation of, and some operations on, a large number of spins. This would severely slow down the code.

- Continous dynamic require a double arithmetic

  A bool arithmetic is much more efficient than a double one.

4.1   Few words on multispin coding

Multispin coding is a coding technique implemented to run fast simulations on Ising spins. Since an ising spin is essentially a boolean variable, we are allowed to imagine that a long integer variable (32 bits), when expressed in bit representation, contains informations on 32 spins. If we have N long int, we can imagine that we have informations on 32 systems of N sites each.

The interesting thing is that if we can write the Hamiltonian using bitwise operations, instead of using integer (or worse, floating point) algebra, we will have two great advantages: First is that bitwise operations are pretty fast by itselves, so we will gain some computation time in this operation. Second, and more important, is that we can evolve 32 spins using a single operation! Of course, we may have to multiply the result of a bitwise operation with a floating point number, so that the algebra is essentially floating point. But the crucial operations of the simulations are still boosted.

As example, the definition of Hamiltonian is given for a 3-spin Ising model:

$$\mathcal{H}(S_1, S2, S_3) = A(S_1 \oplus S_2 \oplus S_3) + B \tag{19}$$

where the $\oplus$ sign represents the bitwise XOR operation.

4.2   Dynamic arrest in simulations

In this section i will present some simulation on Ising p-spin and show that dynamic arrest occurs, similarly to what happens in the Dynamical framework presented before. If we save from simulations the time evolution of the overlap matrix, we can see that, in the high temperature phase all the overlaps fastly decrease to zero, but if we are approaching low temperatures those seem to stuck to nonzero values.

$$S_1 \qquad 10100101001011101$$
$$S_2 \qquad 10010111101011100$$
$$S_3 \qquad 11101101100011010$$

$$\Delta E \qquad 11011111000111011$$

Figure 1: Multispin Coding: 3 sites have R replicas, all of them are updated using a single bitwise operation

This is indeed a method to roughly determine via-simulations the eventual presence of dynamical arrest, even though I found some difficulties in the evaluation of the exact T_d.

4.3 Tempering

By looking the relaxation curves for the replica overlaps of the system, we can see that below the dynamical temperature dynamic arrest occurs. That is, the overlap values do not go to zero, while the equilibrium properties at this temperature suggest us that the overlap distribution should be $P(Q) = \delta(Q-0)$.

However, there is a possible workaround to estimate the correct overlap distribution at equilibrium, and is given by parallel tempering.

Parallel tempering is a technique for simulation boosting invented to deal with problems involving multi-basin sampling. The idea is to evolve a certain number of 'replicas' (technically these are not replicas of the system, since they are at higher temperatures) at temperatures higher that the sampling ones. HIgher temperatures allows those probes to sample the energy landscape without being freezed in a single basin. When a high temperature probe finds a basin that has similar energy to the real system, the simulation allows the exchange of temperatures, letting the cold probe to sample the new basin, and turning the old sampler into an high temperature probe. By means of this implementation, parallel tempering is usually also referred as replica exchange method.

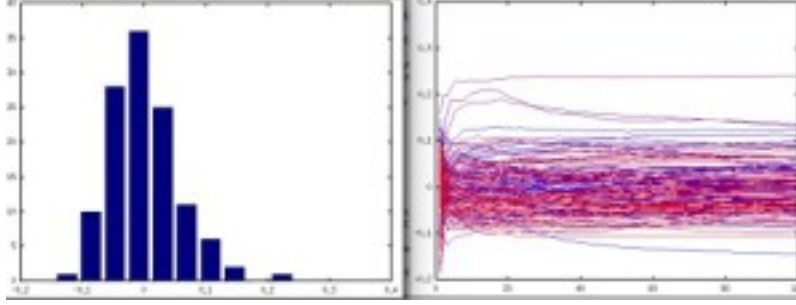The detail of the algorithm is very similar to the one used in the Metropolis

Figure 2: Comparison in slowing down on the time series: A metropolis simulation showing dynamical arrest in the region $\beta = 1$. Time in the x axis is scaled on sampling time, $N^4/100$

algorithm. When a certain number of Montecarlo sweeps have been made on the system, replica exchange picks two configurations at random, $\{S_a, \beta_a\}, \{S_b, \beta_b\}$, and evaluates the probability of the move, according to Metropolis Hastings algorithm.

$$p = \min(1, \frac{\exp(-\beta_a E[S_b] - \beta_b E[S_a])}{\exp(-\beta_a E[S_a] - \beta_b E[S_b])}) \tag{20}$$

The ratio of the two exponential may be re-evaluated with some simple algebra to obtain

$$p = \min(1, \exp(\Delta\beta\Delta E) \tag{21}$$

It is easy to see that this condition satisfies detailed balance.

In this way we should be able to recover the static properties of the system under the condition of dynamical arrest. As we can see from pictures, at low but not extreme temperatures, the system is now able to relax around the correct overlap distribution, which is $P(Q) = \delta(Q)$.
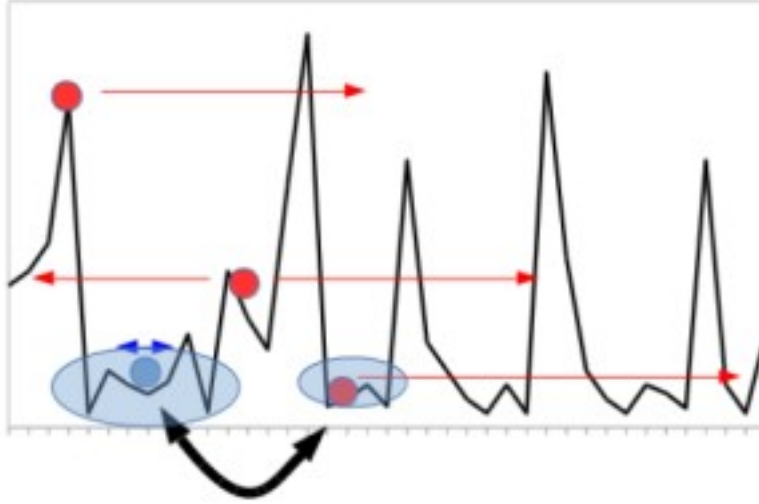
Figure 3: Parallel tempering allows sampling in different basins, thus preventing the effects of critical slowing down

References

[1] A. Crisanti, H.J. Sommers, The spherical p-spin interaction spin-glass model, Dynamics , Z. Phys. B 92, 25

[2] A. Crisanti, L. Leuzzi, The spherical p-spin interaction spin-glass model, Static, Z.Phys.B Condensed Matter 87, 341-354 (1992))

[3] G. Parisi, Title, Journal/Editor, (year)

[4] Kosterlitz, Title, Journal/Editor, (year)

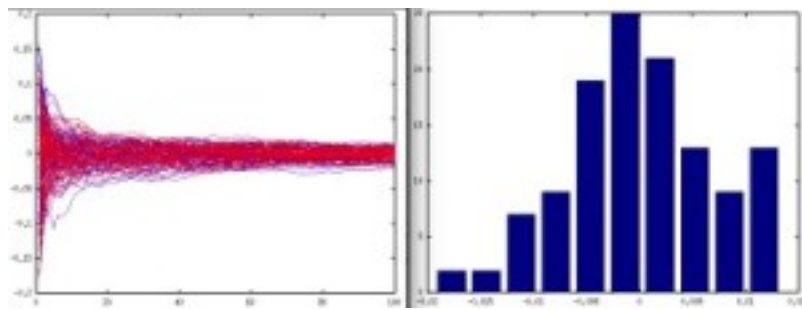[5] C. Castellani, A. Cavagna, Spin glass theory for pedestrians

Figure 4: Comparison in slowing down on the time series: A parallel tempering in the region $\beta = 1$. We can see a very different picture if we compare this with the preceding one Time in the x axis is scaled on sampling time, $N^4/100$