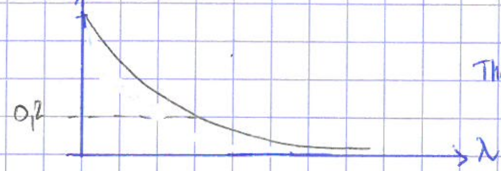
11/07/2017

iii/ Assume we have a binary signal: $P_x(x) = \pm 1$ with proba $1/2$ $\Rightarrow \begin{cases} m = \int \mathcal{D}u \, \tanh(\lambda + \sqrt{\lambda} u) \\ q_0 = 1 \end{cases}$ $\quad \mathcal{D}u = du \frac{e^{-u^2/2}}{\sqrt{2\pi}}$

$MMSE = q_0 - m = 1 - m$



The best one can do according to the SNR.

AVERAGE FREE ENERGY: $\boxed{F = -\mathbb{E}_{x^*, z}(\log Z) = -\int dx^* P_x(x^*) \int \mathcal{D}z \log \int dx \, e^{-\frac{\lambda}{2}x^2 + x\lambda x^* + x\sqrt{\lambda} z} P_x(x)}$

↳ we'll need $\varphi_{\text{growing}}(\lambda^{(1)}, \lambda^{(2)}) = -\int dx^* P_x(x^*) \int \mathcal{D}z \log \int dx \, e^{-\frac{\lambda^{(1)}}{2}x^2 + x\lambda^{(2)} x^* + x\sqrt{\lambda^{(1)}} z} P_x(x)$.

↳ One can show that $\boxed{I(x;Y) = F + \frac{\lambda \mathbb{E}(x^2)}{2}}$ Ex: $\rightarrow \frac{\partial I}{\partial \lambda} = \frac{1}{2}\underbrace{\left(\langle x \rangle_{MMSE} - x^*\right)^2}_{= MMSE}$ I-MMSE theorem.

LECTURE II : SPIKE MATRIX MODEL

II a introduction

Our signal will have this time many component $\vec{x}^* = \begin{cases} x_1^* \\ \vdots \\ x_N^* \end{cases}$ with $x_i^* \sim P_x(x_i)$ i.i.d.

Our noisy measurement is a $N \times N$ matrix: $Y = \underbrace{\sqrt{\frac{\lambda}{N}} \vec{x}^* \vec{x}^{*T}}$ $+ W$ with $w_{ij} = w_{ji} \sim \mathcal{N}(0,1)$.

perturbation to the random matrix W (wigner) corresponding to changing 1 e.V only → this is why thinking in terms of spectrum we talk about "spike".

<u>QUESTION:</u> Can we recover X^* from Y? Is there an efficient algorithm to do so? N^2 measures for N components ↳ EASY! if λ large enough (spike out of bulk)

↳ <u>Rk</u>: this definition can be extended to tensors: $y_{ijk} = \frac{\sqrt{\lambda} x_i^* x_j^* x_k^*}{N^{3-1/2}} + w_{ijk}$ ORDER 3, RANK 1 TENSOR

↳ much harder because there is no such thing as eigen decomposition for tensors

Application, binary variables + noise = random flips → SBM

let's compute the posterior: $\begin{cases} P(\vec{X}|Y) \propto P_x(\vec{x}) \prod_{i \le j} e^{-\frac{1}{2}\left(y_{ij} - \sqrt{\frac{\lambda}{n}} x_i x_j\right)^2} \propto \prod P_x(x_i) \prod_{i \le j} e^{-\frac{\lambda}{2N} x_i^2 x_j^2 + y_{ij} x_i x_j \sqrt{\lambda/N}} \\ Z(Y) = \int d\vec{x} \prod_i P_x(x_i) \prod_{i \le j} e^{-\lambda/_{2N} x_i^2 x_j^2 + y_{ij} x_i x_j \sqrt{\lambda/N}} \end{cases}$

Restricting ourselves to $x_i = \pm 1$: $(x_i^2 = 1) \Rightarrow P(\vec{X}|Y) \propto e^{\sqrt{\lambda} \sum_{i \le j} \frac{y_{ij}}{\sqrt{N}} x_i x_j}$ → Disordered Ising model with correlated interactions.

unlike the SK model, interactions are not iid: $y_{ij} = w_{ij} + \sqrt{\frac{\lambda^*}{n}} x_i^* x_j^*$ ← let's assume there is a ground truth X^* that we might not know.

$P(\vec{x}|Y) \propto e^{\sqrt{\lambda} \sum_{i \le j} \left(\frac{w_{ij}}{\sqrt{N}} + x_i^* x_j^* \sqrt{\frac{\lambda^*}{N}}\right) x_i x_j}$

$(x_i^*)^2 = 1$

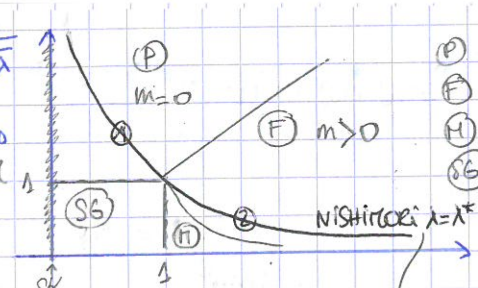
$\propto \exp\left[\sqrt{\lambda} \sum_{i \le j} \left(\frac{1}{\sqrt{N}} w_{ij} x_i^* x_j^* + \sqrt{\frac{\lambda^*}{N}}\right) x_i x_i^* x_j x_j^*\right]$

→ by the gauge transform we got rid of correlations at the price of a non zero mean.

non-zero mean

diagram of model parameters

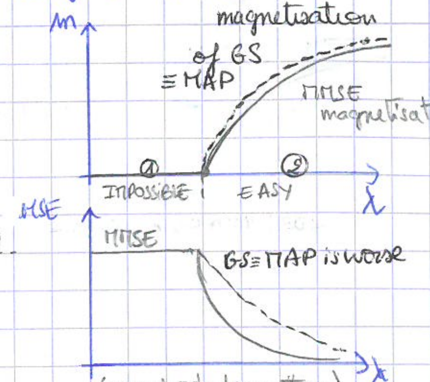
PHASE DIAGRAM



(P): paramagnetic
(F): ferromagnetic
(M): mixed: m>0, full RSB
(SG): spin glass; full RSB

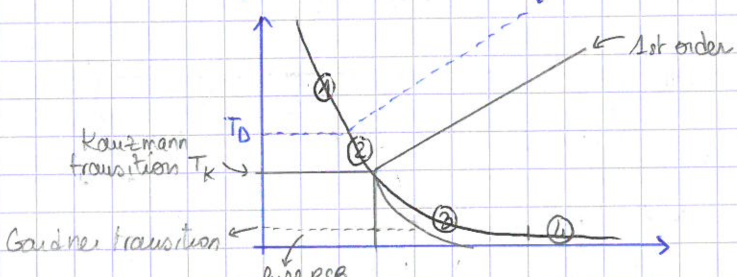
where the Nishimori relations hold.

along the Nishimori line

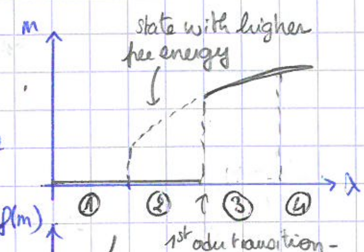


(second order transition)

Now in the case of a tensor of order 3:

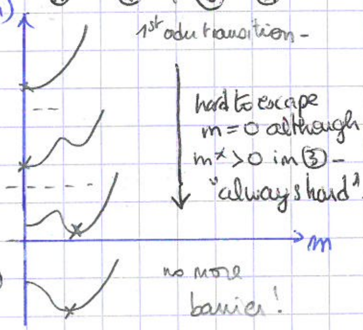


now 3 phase along Nishimori line



COMPUTATIONAL COMPLEXITY:

- ① IMPOSSIBLE (indistinguishable) ②
- POSSIBLE but HARD ③
- EASY ④



↳ This two cases are representative of the different scenario: → 1st order with impossible → easy; ↘ 2^{nd} order with impossible + hard → easy.

We'll see how to make the necessary computations on the Nishimori line.

↳ REPLICA COMPUTATION ON THE EXERCISE SHEET.

$$\lim_{N\to+\infty} \frac{F}{N} = \min_{m} \left[\mathcal{G}_{den}(\lambda m, \lambda m) + \frac{\lambda m^2}{4} \right]$$

$m \to m_{eq}$.

→ replicas lengthy (crazy as well...)
↘ rigorous proof

Lecture III Free energies and algorithms

III.1 Replica free energy

Recall the hamiltonian of the spike matrix model:

$$\mathcal{H}_{SM} = -\sum_{i \leq j} -\frac{\lambda}{2N} x_i^2 x_j^2 + \frac{\lambda}{N} x_i x_j x_i^* x_j^* + \sqrt{\frac{\lambda}{N}} x_i x_j W_{ij}$$

↳ exercise shows:

$$\begin{cases} f = \lim_{N\to+\infty} \dfrac{\mathbb{E}_{W,x^*}(-\log Z)}{N} = \min_m \underbrace{\left[\varphi_{den}(\lambda m, \lambda m) + \dfrac{\lambda m^2}{4} \right]}_{f_{RS}(m)} \\ m_{eq} = \operatorname{argmin}_m f_{RS}(m) \end{cases} \quad (1)$$

(1) for computation later.

In some cases, the crazy replica method can be actually rigourously proven. The main technique to do so is the Guerra interpolation.

In our problem we have no replica symmetry breaking → pretty simple setting.

↳ idea: interpolate between our problem and the simple model of 1 spin in a field.

simple → $\mathcal{H}_{DEN} = -\sum_{i=1}^{N} -\frac{\lambda x_i^2}{2} + x_i x_i^* \lambda + x_i \overset{\text{noise}}{\xi} \sqrt{\lambda}$ → independent spins $y_i = \sqrt{\lambda} x_i + \xi_i$.