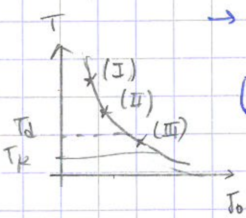
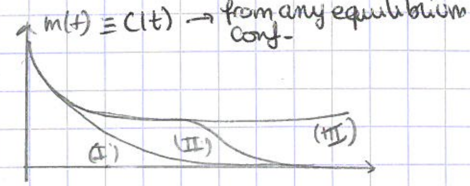
looking back at the phase diagram = spin-tensor factorization.

- For $T > T_K$, on the Nishimori line, paramagnetic phase → indistinguishable from random problem, same as if $J_0 = 0$ = unplanted

→ but if you planted the problem, you know one equilibrium configuration, which is of great help in many circumstances.





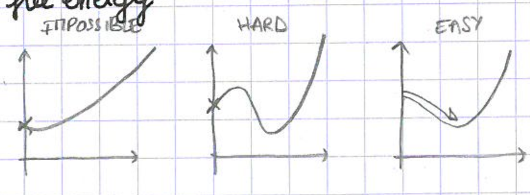
Monte Carlo simulation from large magnetization on the Nishimori line

= simulation from any equilibrium conf, which would be very costly to obtain otherwise at low Temperature, especially $T < T_d$ → GREAT TRICK!

14/07/2017

III 2 Belief propagation, TAP free energy

We have this situation



to find the solution we can run monte carlo but alternative, faster algorithms → Belief Propagation

Belief propagation was rediscovered many times, but originally Bethe-Peierls in the 30s.

The idea is the following (see exercise sheet):

hamiltonian: $H = -\sum_{i,j} J_{ij} S_i S_j$

1. Write recursive BP "tree with high connectivity":
$m_{i\to j} = f(\{m_{k\to i}\}) = \tanh\left(\sum_{k\in\partial i, j} \operatorname{atanh}(\tanh(\beta J_{ik}) m_{k\to i})\right)$

2. Expand in J (small J)

large connectivity assumption, marginals ← 3. Close the equations on the marginals = magnetizations
$m_i = \tanh\left(\beta \sum_j J_{ij} m_j - \beta^2 \sum_j J_{ij}^2 (1 - m_j^2) m_i^{t-1}\right)$
↳ time indexing is crucial for convergence!!

This algorithm can be analyzed: it is basically doing a gradient descent along the replica symmetric potential → TAP magnetization get to (local) minima of the RS free energy.
↳ still stuck for the hard phase.
↳ finds optimal solution in easy phase.

Rk: Belief in stat mech community that message passing algorithm are somewhat optimal: there are not polynomial algorithm that can succeed in the hard phase.

Lecture IV: GENERALISED LINEAR MODELS

The situation is still signal recovery from noisy observations.

signal $X^* \in \mathbb{R}^N$ distributed $P_X(x)$ → observation in $\mathbb{R}^M$ $Y = f[\Phi X^*]$ (KNOWN; measurement matrix $\mathbb{R}^{M\times N}$ KNOWN!)

component-wise! $y_\mu = f[\Phi_\mu . X^*]$

$\begin{cases} \alpha = M/N \\ \rho = \text{non zeros in } X^*/N \\ \Phi_{i\mu} \sim \mathcal{N}(0, 1/N) \text{ iid.} \end{cases}$

different situations for different f: GENERALISED LINEAR MODELS!

1) $f(x) = x$ → linear model → statphys community Tanaka 2000 something

2) stochastic version: $f(x) = x + \xi$, $\xi \sim \mathcal{N}(0, \Delta)$

3) non-linearities: $f(x) = \operatorname{sign}(x + \xi)$ – perceptron – classification problem

hamiltonian depending on the noise → $p(x|y) \propto p_{out}(y|x)\, p_x(x)$ where P_{out} encodes the $f(\,)$.

$z_\mu = \Phi_\mu \cdot x \rightarrow P_{out}(y_\mu | z_\mu)$ –

↳ one can compute the associated free energy:

$$\lim_{n\to\infty} \frac{\log Z}{N} = \underset{m,\hat{m}}{\mathrm{Extr}} \Big[\alpha \int dy\, \mathcal{D}\xi\, \mathcal{D}s\, P_{out}\big(y \,|\, s\sqrt{\rho - m} + \xi\sqrt{m}\big) \log \int \mathcal{D}x\, P_{out}\big(y \,|\, x\sqrt{\rho - m} + \xi\sqrt{m}\big)$$
$$- \frac{m\hat{m}}{2} + \int \mathcal{D}\xi\, ds\, dx\, P_x(s)\, e^{-\hat{m}s^2/2 + \xi s\sqrt{\hat{m}}} \log\Big(\int dx\, P_x(x)\, e^{-\frac{1}{2}\hat{m}x^2 + \xi x\sqrt{\hat{m}}}\Big)\Big]$$

example of P_{out}: 1) $f(x) = x \rightarrow P_{out}(y|z) = \delta(y - z)$

2) $f(x) = x + \varepsilon \rightarrow P_{out}(y|z) = \mathcal{N}(z, \Delta)$.

Historically: - Mezard first write TAP for generalized linear model 89'

- Kabashima 06'

dynamic TAP equations !

- Renaissance AMP approximate message passing Donoho, Maleki, Montanari 09'

Bayati - Montanari

↳ proof, understand the t-1 index.

- GAMP Sundeep Rangan.

IV 1 Compressed sensing

Assume signal is sparse $X^* = \begin{pmatrix} 0 \\ 0 \\ 0.7 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1.2 \\ -1.7 \\ 0 \\ 0 \end{pmatrix} \in \mathbb{R}^N$ $\{ K = \#$ non zeros, $\rho = K/N$ sparsity $)$

and you have access to a series of M random projections

$y_1 = \Phi_1^T X^*$
$y_2 = \Phi_2^T X^*$
$\vdots$
$y_M = \Phi_M^T X^*$

We denote $\alpha = \frac{M}{N}$ the number of measurements / N, controls the difficulty of the problem:

→ $\alpha \geq 1$ then you can trivially invert the linear system.

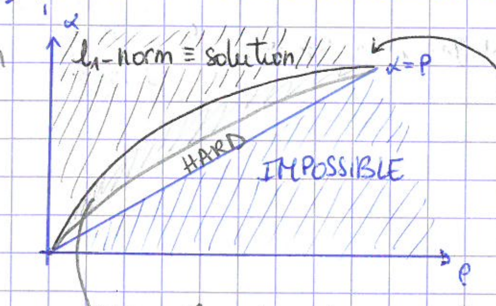
→ but we don't even need that many measurement, as we only have K non-zero in X.
we only need: $\alpha > \rho$ and take $\hat{X}$ = solution of $Y = \Phi \hat{X}$ with largest number of $0 = X^*$

↳ NP complete strategy!

can be proven

↳ instead of minimizing ℓ_0 norm (= number of non-zero), minimizing the ℓ_1-norm $|x|_1 = \sum_i |x_i|$ is much easier and working! $\hat{X}$ = solution of $Y = \Phi \hat{X}$ with smallest $|x|_1$

→ phase diagram



Donoho - - -

line → AMP will give you the solution.

↳ AMP with Nishimori line = knowledge of the prior
e.g. $P(x) = (1-\rho)\,\delta(x) + \rho\, \mathcal{N}(0,1)$

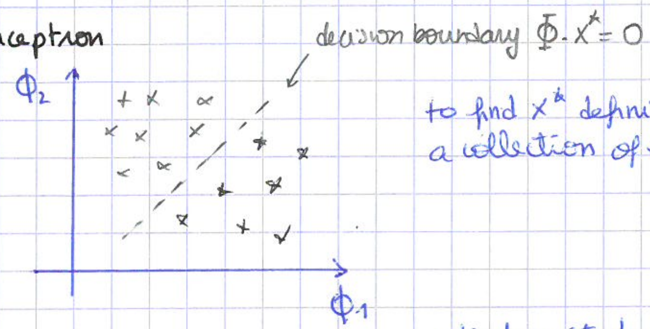
IV 2 Superposition codes

Divide signal in B element sections with only one 1 in each section: $X^* = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \in \mathbb{R}^{(L \times B)}$ } B elements, L sections

$M \times N$, $M < N$.

Compress signal: $W = \Phi X^*$, send signal threw AWGN → $Y = \Phi X^* + Z$ → { sparsest ... X solution

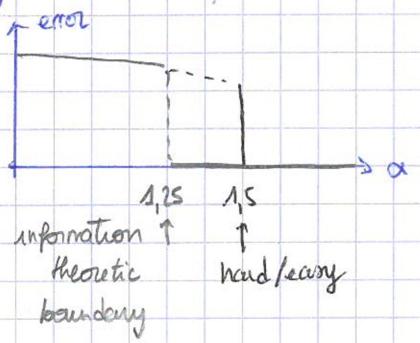
IV 3 Perceptron



to find x^* defining the decision boundary, one is given a collection of examples: $\begin{cases} y^{(1)} = \text{sign}(\Phi^{(1)} \cdot x^*) ; \Phi^{(1)} \\ y^{(2)}, \Phi^{(2)} \\ \vdots \\ y^{(M)}, \Phi^{(M)} \end{cases}$

→ teacher student perceptron.
baby machine learning problem.

how many examples one needs?



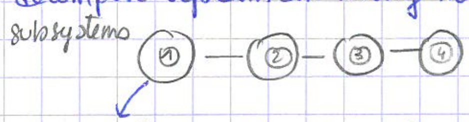
→ again the same picture.

↳ Now, not everything is linearly separable.
⇒ need to go to multi-layer perceptron / deeplearning to solve this problem.

CONCLUSION: How to get out of the hard phase? SPATIAL COUPLING.

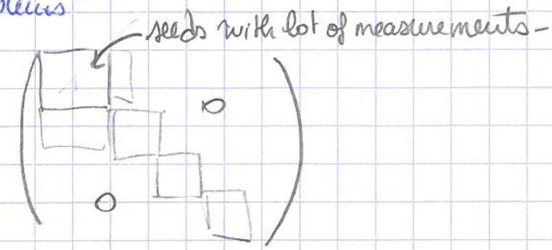
→ The problem is how to get out of the metastable state?
{ In real world, it is easy to nucleate supercooled liquid, but limited.
{ In our MF settings energy barriers are $O(N)$!

↳ "de mean field the problem", capture some
↳ decompose system in many mean field subsystems coupled in 1d.



→ with large α here to make a seed, something easy to solve.
⟶ succinctly solve the different subsystems

↳ in practice compressed sensing with Φ



→ resolution wave that can be understood by the physics of nucleation!