Belief Propagation, TAP free energy

We have this situation

\[ \text{HARD} \quad \text{EASY} \]

to find the solution we can run Monte Carlo but alternative, faster algorithms → Belief Propagation

The idea is the following, (see exercise sheet):

- Write recursive BP "tree with high connectivity":
  \[ m_{i;j} = \text{tanh} \left( \sum_j \text{tanh} \left( \beta J_{i;j} m_j \right) \right) \]

- Expand in \( T \) (small \( T \))

Large connectivity assumption, marginals → 3. Close the equations on the marginals of magnetizations

\[ m_i = \text{tanh} \left( \sum_j J_{i;j} m_j \right) \]

\( \beta \) is fixed, time indexing is omitted for convergence!!

This algorithm can be analyzed: it is basically doing a gradient descent along the replica symmetric potential → TAP magnetization gets to (local) minima of the RS free energy.

This still works for the hard phase and finds optimal solution in every phase.

Bethe/Stat mech community has message passing algorithms are somewhat optimal: there are no polynomial algorithms that can succeed in the hard phase.

Lecture IV: GENERALISED LINEAR MODELS

The situation is still signal recovery from noisy observations:

\[ x^* \in \mathbb{R}^N \text{ distributed } P_0(x) \rightarrow \text{observation } y \in \mathbb{R}^H \]

\[ y = f(\Phi x^* + \epsilon) \]

\[ \Phi \text{ a known measurement matrix} \]

Different situations for different \( f \): GENERALISED LINEAR MODELS:

1) \( f(x) = x \) → linear model
   \[ \Phi \text{ II} \Phi \text{ community Katrina 2000 reconstructing} \]

2) stochastic version: \( f(x) = x + \epsilon \)
   \[ \Phi \text{ non-linear} \]

3) non-linearities: \( f(x) = \text{sign}(x + \epsilon) \) → perceptron classification problem
Hamiltonian depending on the noise \( \mathcal{H}(x,y) = p(y|x)\log p(y|x)p(x) \), when four encodes the \( f(x) = \phi_x \).\n
You can compute the associated free energy:
\[
\begin{aligned}
\lim_{m \to \infty} \frac{1}{N} \sum_{\text{Pa}(m)} \mathbb{E}_{p(x)} \left[ \mathcal{H}(x,y) \right] &= \sum_{m} \mathbb{E}_{p(x)} \left[ \mathcal{H}(x,y) \right] \\
&= \min_{m} \left[ \frac{1}{2} \sum_{x} \mathcal{H}_{m}(x) \right] \log \mathbb{P}(x) \end{aligned}
\]

Example of Part 1: \( f(x) = x \rightarrow \mathbb{P}(y|x) = \delta(y-x) \)
\( f(x) = x + \epsilon \rightarrow \mathbb{P}(y|x) = N(2, \epsilon) \).

Historically: Vand and Fieffe TAP for generalized linear model 89

Renaissance AMP approximates message passing, Koshic, Kalb, Tackman 90

Bayati, Montanari

Coproof, understand Var-1 index

AHT: send deep Rangan.

IV. 1 Compressed sensing

Assume signal is sparse \( x = \sum_{i} \Phi_{i} \).

\( y = \mathcal{P}_{\epsilon} x \), and you have access to a series of \( H \) random projections.

We define \( \epsilon = \frac{\ell}{N} \), the number of measurements, \( N \), controls the difficulty of the problem:

- \( \epsilon > 1 \) then, you can trivially invert the linear system
- But, we don't even need that many measurements, as we only have \( \epsilon \) nonzero in \( x \).

We only need: \( \epsilon > \epsilon_{c} \) and take \( \hat{x} = \Phi_{\epsilon} y \), with largest number of \( 0 = x^{*} \).

AMP complete strategy! can be proven

\( \epsilon \) instead of minimizing \( \ell_{1} \)-norm \( \| x \|_{1} \), minimizing \( \ell_{1} \)-norm instead of minimizing \( \ell_{1} \)-norm \( \| x \|_{1} \).

HARD IMPOSSIBLE

Phase diagram: \( \ell_{1} \)-norm = solution

AHT with Nicholoni line = knowledge of the prior

IV.2 Superposition codes

Divide signal in 8 element sections with only one 1 in each section: \( x^{*} \).

Complex signal: \( w = \Phi x^{*} \), send signal through AWGN \( \rightarrow y = \Phi x^{*} + z \).
IV 3. Perception

to find \( x^* \) defining the decision boundary, one is given
a collection of examples: \( y^{(i)} = \text{sign}(\Phi_i x^*) \); \( \Phi^{(i)} \)

\[ \begin{align*}
\Phi^{(2)} \quad y^{(2)} \\
\Phi^{(3)} \quad y^{(3)} \\
\Phi^{(m)} \quad y^{(m)}
\end{align*} \]

- teachable, student, perception
- baby machine learning problem

\[ \rightarrow \text{again the same picture} \]

\[ \begin{align*}
\Phi_2 \quad y^{(1)} \\
\Phi_3 \quad y^{(2)} \\
\Phi_m \quad y^{(m)}
\end{align*} \]

\[ \rightarrow \text{now, not everything is linearly separable} \]
\[ \Rightarrow \text{need to go to multi-layer perceptron/deep learning to solve this problem.} \]

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

\[ \rightarrow \text{The problem is how to get out of the metastable state?} \]
\[ \text{In real world, it is easy to nucleate supercooled liquid, but limited.} \]
\[ \text{In our HF setting, energy barrier are } O(N). \]

"de-mean field the problem", capture some
\[ \text{1. decompose system in many mean-field subsystems coupled in 1d} \]
\[ \text{subsystems} \]
\[ \text{with huge x. how to make a seed, something easy to solve} \]
\[ \text{1. success: solve the different subsystems} \]
\[ \text{in practice: compared sensing with} \ \Phi \]
\[ \text{made with } O \text{ of measurements.} \]

\[ \rightarrow \text{resolution wave that can be understood by the physics of nucleation!} \]