FEDERICO ③

19/07/2017

The fact that Bethe free entropy is bounded ⇒ at least one extremum.
⟺ BP has at least one fixed point.

Computing long range correlations with the linear response:
↳ Suppose $\Psi_i(x_i) \leftarrow \Psi_i(x_i)e^{H_i x_i} \Rightarrow \chi_{ij} = \langle x_i x_j\rangle - \langle x_i x_j\rangle = \frac{\partial\langle x_i\rangle}{\partial H_j}$

why better than $C_{ij}$? Argument in Parisi's book (non trivial)

ISING PAIRWISE MODELS

$$H(\underline{x}) = \sum_{(ij)\in\tilde{E}} J_{ij}x_i x_j - \sum_i H_i x_i \longrightarrow p(\underline{x}) = \frac{e^{-\beta H(x)}}{Z}$$ on any graph $\tilde{E}$. ⚠ ≠ edges between factors and nodes.

In this special case $a \equiv (i,j)$, $\Psi_a(\underline{x}_{\partial a}) = \Psi_{ij}(x_i,x_j) = e^{\beta J_{ij}x_i x_j}$

For binary spins, marginals only depend on one number. Consequently we introduce the corresponding parameter: $\nu_{i\to(ij)} \propto e^{\beta h_{i\to j}x_i}$ $\quad \hat{\nu}_{(ij)\to i} \propto e^{\beta u_{i\to j}x_j}$

TRICK: $x\in\{0,1\}$, $\exp(Ax) = ch(A)(1+x\, th(A))$

The BP equations can be rewritten: $h_{i\to j} = H_i + \sum_{k\in\partial i\setminus j} u_{k\to i}$

and $1 + x_j\, th(\beta u_{i\to j}) \propto \sum_{x_i}\left[1 + th(\beta J_{ij})x_i x_j\right]\left[1 + th(\beta h_{i\to j})x_i\right]$
$\propto 1 + th(\beta J_{ij})\, th(\beta h_{i\to j})x_j$

$$\Rightarrow \boxed{\begin{cases} h_{i\to j} = H_i + \sum_{k\in\partial i\setminus j} u_{k\to i} \\ u_{i\to j} = \hat{u}(J_{ij}, h_{i\to j}) \text{ with } \hat{u}(J,h) = 1/\beta\ \mathrm{arcth}(th(\beta J)\,th(\beta h)) \end{cases}}$$

BP EQUATIONS!

EXERCISE: Solve homogeneous Ising model $J_{ij} = J = 1$, $H_i = H$ on a graph of fix degree $|\partial i| = d$. ($d=4$, square lattice)
Compute $M, E, T_c$ (→ more than one solution) → compare to MF
→ compare to exact result 2D ($d=4$).

LINEAR RESPONSE: $\frac{\partial\langle x_i\rangle}{\partial H_j}$ : $\delta H_j \to$ modification of messages propagating $\to \delta\langle x_i\rangle$.
With BP, after convergence we can estimate: $\langle x_i\rangle = th(\beta H_i + \beta\sum_{j\in\partial i} u_{j\to i})$
Perturbation at the origin: $\frac{\partial\langle x_i\rangle}{\partial H_0} = (1-\langle x_i\rangle^2)\,\beta\sum_{j\in\partial i}\frac{\partial u_{j\to i}}{\partial H_0}$

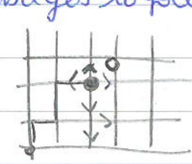
attenuation factor

depends on the strength of the couplings.

and
$$\begin{cases} \frac{\partial u_{j\to i}}{\partial H_0} = \left.\frac{\partial\hat{u}}{\partial h}\right|_{h^*_{ij}} \frac{\partial h_{i\to j}}{\partial H_0} \\ \frac{\partial h_{i\to j}}{\partial H_0} = \delta_{i0} + \sum_{k\in\partial i\setminus j}\frac{\partial u_{k\to i}}{\partial H_0} \end{cases}$$

source term

SUSCEPTIBILITY PROPAGATION
Linear equations in terms of the derivates of the cavity messages to follow the linear perturbation.



$i$

→ sum over all the possible paths through the recursive procedure

Rk: ferromagnetic susceptibility: $\chi_F = \frac{1}{N}\sum_{ij}\chi_{ij} \overset{\text{some homogeneity}}{=} \sum_i \chi_{i0} = 0$ if random signs for $J_{ij} \sim \mathcal{N}(0,J)$

need for spin glass susceptibility: $\chi_{SG} = \frac{1}{N}\sum_{ij}\chi_{ij}^2 = \sum_i \chi_{i0}^2$

Rk: We are connecting physics/computation → physical susceptibility
↳ stability of the BP fixed point

↳ no convergence? Might not be a numerical problem and instead a true property of the model which susceptibility is diverging!

THOULESS-ANDERSON-PALMER EQUATIONS: SIMPLIFICATION FOR DENSE MODELS.

Fully connected model $H = -\sum_{ij} J_{ij} x_i x_j$ ferromagnetic $J \sim 1/N$
spin glass $J \sim \pm 1/\sqrt{N}$.

↳ using BP ⟹ $N^2$ equations → let's try to simplify.

⟹ expansion in small $J$: $\begin{cases} h_{i\to j}^{(t+1)} = H_i + \sum_{k \in \partial i \setminus j} u_{k\to i}^{(t)} \\ u_{k\to i}^{(t)} = J_{ik}\,\mathrm{th}(\beta h_{k\to i}^{(t)}) \end{cases}$

considering the magnetization (not the cavity messages):

$$m_i^{(t+1)} = \mathrm{th}\Big[\beta H_i + \beta \sum_{j\in\partial i} u_{j\to i}^{(t)}\Big] = \mathrm{th}\Big[\beta H_i + \beta \sum_j J_{ij} \underbrace{\mathrm{th}(\beta h_{j\to i}^{(t)})}_{\simeq m_j^{(t)} - \dots}\Big]$$

$$\begin{cases} m_j^{(t)} = \mathrm{th}\big[\beta h_{j\to i}^{(t)} + \beta u_{i\to j}^{(t-1)}\big] \\ \mathrm{th}(\beta h_{j\to i}^{(t)}) = m_j^{(t)} - (1-m_j^{(t)})^2 J_{ij}\, \underbrace{\mathrm{th}(\beta h_{i\to j}^{(t-1)})}_{\simeq m_i^{(t-1)}} \end{cases}$$

⟹ TAP $\boxed{m_i^{(t+1)} = \mathrm{th}\Big[\beta H_i + \beta \sum_i J_{ij} m_j^{(t)} - \beta \sum_j J_{ij}^2 (1-m_j^{(t)2})\, m_i^{(t-1)}\Big]}$

↳ BP for dense model
with only N mean field parameters

Rk: the TAP equations are only valid for N large, can run into numerical problems, where the fixed point are not attractive → yet are these states interesting? saddle point of $F_{Bethe}(b)$? Metastable state? → SK N finite, no solution... $\min V(m) = \sum_i \left(\frac{\partial F_{TAP}}{\partial m_i}\right)^2 = 0$? flat?

Rk: In the ferromagnetic case → $\sum_j J_{ij}^2 \sim O(1/N) \ll \sum_i J_{ij} \sim O(1)$. Indeed MF is correct!

CSPs ZERO TEMPERATURE LIMIT.

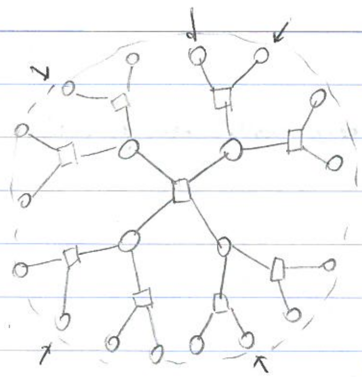
EXERCISE: Consider zero temperature limit of BP for CSP. Where $\psi_a \sim e^{-E_a}$ $E_a = \begin{cases} 1 & \text{UNSAT} \\ 0 & \text{SAT} \end{cases}$ ; $H = \#\text{UNSAT}$

⟹ warning propagation algorithm

$\begin{cases} \tilde{u}_{a\to i}(x_i) \in \{0,1\} : \text{"you can/cannot take value } x_i \text{ according to variables } \underline{x}_{\partial a \setminus i}\text{"} \\ h_{i\to a}(x_i) \in \{0,1\} : \text{"I can/cannot take value } x_i \text{ according to clauses } b \in \partial i / a\text{"} \end{cases}$

↳ parallelized algorithm handy for optimization!!

$\{ J = 1 \to (d-1)\,\mathrm{th}(\beta_c) = 1$



locall neighborhood $R$:

propagate boundary messages $\rightarrow$ measure $\mu_R(\underline{x}_R)$

↳ assumption if looking at neighborhood of distance $R-1$, entering boundary computed from $\mu_R$ through BP. $\equiv$ messages at boundaries indep.

↳ assumption, internal measure well approximated by tree.

<u>WRONG ASSUMPTION</u> when: → loopy graphs (Yedidia) - GRAPHICAL MODEL APPROX - CLUSTER VARIATIONAL METHODS (Kikuchi)

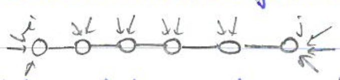
→ long range correlations (long loops)

SPARSE RANDOM GRAPHS

When getting long <u>range correlations</u>, what should we do with BP?

First, what correlations to measure?

→ connected correlations from distant variables: $\langle x_i x_j\rangle_c \equiv \langle x_i x_j\rangle - \langle x_i\rangle\langle x_j\rangle$



in 1d correlations always decay: $C(r) = \langle x_i x_j\rangle_c\big|_{|i-j|=r} \sim e^{-r/\xi}$

But the number of neighbors increases with distance: $\chi_F = \sum_i \chi_{i0} = \sum_r C(r)\, d(d-1)^{r-1}$ ← $d$ random regular graph

$= \frac{d}{d-1}\sum_r \exp\left(-\frac{r}{\xi} + r\ln(d-1)\right)$

↳ criticality for finite correlation length: $\xi_c = \frac{1}{\ln d-1}$ - !

20/07/2017

Define measure of how $i$ and $j$ are correlated:

$\chi^{(2)} = \frac{1}{N}\sum_{ij} \| \mu_{ij}(\cdot,\cdot) - \mu_i(\cdot)\mu_j(\cdot)\|$ → norm over probability measures

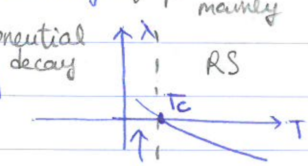
$\Rightarrow \chi^{(n)} = \frac{1}{N^{n-1}}\sum_{(i_1 \dots i_n)\in V} \| \mu_{i_1\dots i_n}(\,) - \mu_{i_1}\cdots\mu_{i_n}\|$   eg: $\ell_1$- "Frobenius" $\|A(x_i,x_j)\| = \frac{1}{2}\sum_{x_i x_j}|A(x_i,x_j)|$.

The model (BP) is <u>stable to small perturbation</u> if for all finite $n$, $\frac{\chi^{(n)}}{N} \rightarrow 0$ as $N\rightarrow\infty$

→ divergence of susceptibility ↔ instability of BP fixed points

In practice, should look at all $\chi^{(n)}$, but instead typically look at $\chi^{(2)}$ only (instability featured in pairwise models mainly)

↳ perturbating the fixed point: $\overset{*}{\nu} + \delta\nu$ → rate of decay? $\lambda = \lim_{t\to\infty} \frac{1}{t}\|\delta\nu\|$ ← exponential decay




BP stops converging

can be measured by population dynamics

Instability making the RS ansatz incorrect.

↳ POINT-TO-SET CORRELATION FUNCTION

Even if BP fixed points are stable, the system can be creating some kind of order in