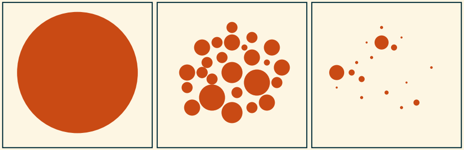
Diluted spin glass models

Amin Coja-Oghlan
TU Dortmund

Overview

- replica symmetry and Belief Propagation
- construction of Bethe states
- application: random matrices and the k-XORSAT threshold

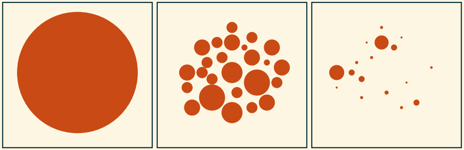
Reminder: predictions from the cavity method



Replica symmetry breaking [MP00,KMRTSZ07]

- replica symmetry
- dynamic replica symmetry breaking
- (static) replica symmetry breaking

Reminder: predictions from the cavity method

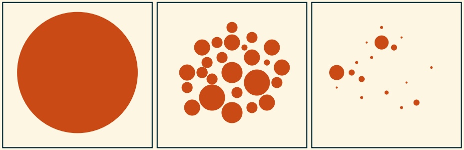


Replica symmetry

[MP00,KMRTSZ07]

- $\mu_{\mathbb{G},\beta}(\{\boldsymbol{\sigma}_{x_1} = s, \boldsymbol{\sigma}_{x_2} = t\}) \sim \mu_{\mathbb{G},\beta}(\{\boldsymbol{\sigma}_{x_1} = s\})\mu_{\mathbb{G},\beta}(\{\boldsymbol{\sigma}_{x_2} = t\})$
- in other words, $\mu_{\mathbb{G},\beta}$ is $o(1)$-extremal

Reminder: predictions from the cavity method



Bethe states [MPRTRLZ99,MP00,KMRTSZ07]

- the phase space decomposes into pure states
- each of them induces a BP fixed point (but not vice versa)
- replica symmetry iff there is just one Bethe state

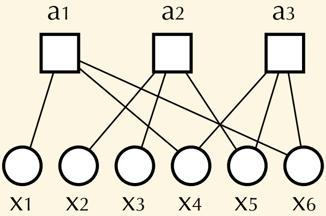
The replica symetric case

An Erdős-Rényi factor graph model

- a random factor graph model $\mathbb{G} = \mathbb{G}(n, \boldsymbol{m})$ with variables $x_1, \ldots, x_n$
- the variable range over Ω
- k-ary factor nodes $a_1, \ldots, a_{\boldsymbol{m}}$ with $\boldsymbol{m} \sim \mathrm{Po}(dn/k)$
- the factor nodes are *independent*
- suppose they all carry the same weight function $\psi(\cdot) > 0$
- the model induces a Boltzmann distribution

$$\mu_{\mathbb{G}}(\sigma) = \frac{1}{Z(\mathbb{G})} \prod_{i=1}^{\boldsymbol{m}} \psi(\sigma_{\partial a_i})$$

The replica symetric case



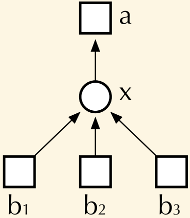
Replica symmetry assumption

- we assume that

$$\sum_{s,t\in\Omega} \mathrm{E}\left|\mu_{\mathbb{G}}(\{\boldsymbol{\sigma}_{x_1} = s, \boldsymbol{\sigma}_{x_2} = t\}) - \mu_{\mathbb{G}}(\{\boldsymbol{\sigma}_{x_1} = s)\mu_{\mathbb{G}}(\{\boldsymbol{\sigma}_{x_2} = t)\right| = o(1)$$

- in other words, $\mu_{\mathbb{G}}$ is $o(1)$-extremal with high probability

The replica symetric case



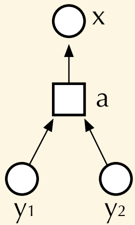
The standard messages

- obtain $\mathbb{G} - a$ by removing a from $\mathbb{G}$ and let

$$\mu_{\mathbb{G},x\to a}(s) = \mu_{\mathbb{G}-a,x}(s) \qquad\qquad (s \in \Omega)$$

- *the marginal of x in $\mathbb{G} - a$*

The replica symetric case



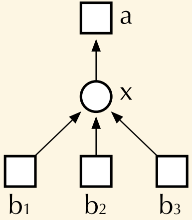
The standard messages

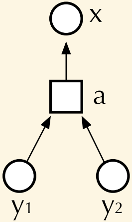
- obtain $\mathbb{G} - (\partial x \setminus a)$ by removing all $b \in \partial x \setminus a$

$$\mu_{\mathbb{G}, a \to x}(s) = \mu_{\mathbb{G} - (\partial x \setminus a), x}(s) \qquad\qquad (s \in \Omega)$$

- *the marginal of* x *in* $\mathbb{G} - (\partial x \setminus a)$

The replica symetric case





Reminder: Belief Propagation

We expect that

$$\mu_{\mathbb{G},x\to a}(s) \propto \prod_{b\in\partial x\setminus a} \mu_{\mathbb{G},b\to x}(s)$$

$$\mu_{\mathbb{G},a\to x}(s) \propto \sum_{\sigma\in\Omega^{\partial a}} \mathbf{1}\{\sigma_x = s\}\psi_a(\sigma) \prod_{y\in\partial a\setminus x} \mu_{\mathbb{G},y\to a}(\sigma_y)$$

The replica symetric case

Theorem [COP18]

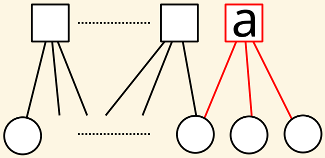
If replica symmetry holds we have for all $s \in \Omega$,

$$\lim_{n\to\infty} \sum_{a \in \partial x_1} \left| \mu_{\mathbb{G}, x \to a}(s) - \frac{\prod_{b \in \partial x \setminus a} \mu_{\mathbb{G}, b \to x}(s)}{\sum_{t \in \Omega} \prod_{b \in \partial x \setminus a} \mu_{\mathbb{G}, b \to x}(t)} \right| = 0$$

$$\lim_{n\to\infty} \sum_{a \in \partial x_1} \left| \mu_{\mathbb{G}, a \to x}(s) - \frac{\sum_{\sigma \in \Omega^{\partial a}} \mathbf{1}\{\sigma_x = s\} \psi_a(\sigma) \prod_{y \in \partial a \setminus x} \mu_{\mathbb{G}, y \to a}(\sigma_y)}{\sum_{\sigma \in \Omega^{\partial a}} \psi_a(\sigma) \prod_{y \in \partial a \setminus x} \mu_{\mathbb{G}, y \to a}(\sigma_y)} \right| = 0$$

“the standard messages satisfy the BP equations”

The replica symetric case

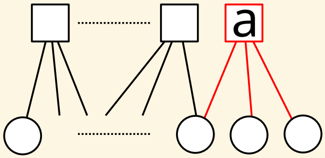


Lemma

Obtain $\mathbb{G} + a$ by adding a single factor node a arbitrarily. Then w.h.p. $\mu_{\mathbb{G}+a}$ is $o(1)$-extremal and

$$\frac{1}{n}\sum_{i=1}^{n} d_{\mathrm{TV}}(\mu_{\mathbb{G}+a,x_i}, \mu_{\mathbb{G},x_i}) = o(1)$$

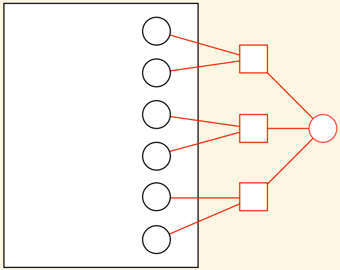
The replica symetric case



Reminder: cut metic

$$d_{\square}(\mu,\nu)=\frac{1}{n}\min_{\gamma\in\Gamma(\mu,\nu)}\max_{\substack{I\subset\{1,\ldots,n\}\\ B\subset\Omega^n\times\Omega^n\\ \omega\in\Omega}}\left|\sum_{i\in I}\sum_{(\sigma,\tau)\in B}\gamma(\sigma,\tau)\left(\mathbf{1}\{\sigma_i=\omega\}-\mathbf{1}\{\tau_i=\omega\}\right)\right|$$

The replica symetric case



Proof of the theorem

- add a new random variable node along with adjacent factors
- the attachment points are random
- due to extremalilty, their joint distribution factorises
- we can therefore verify the BP equations

The replica symetric case

Corollary [COP18]

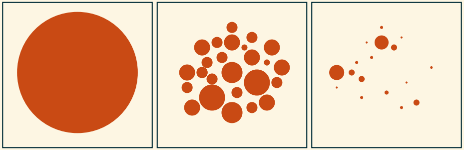
Assume replica symmetry. Then and that the Bethe free entropy $\mathscr{B}(\mathbb{G})$ of the standard messages satisfies

$$\lim_{n\to\infty} \frac{1}{n}\mathscr{B}(\mathbb{G}) = B \in \mathbb{R} \qquad \text{in probability.}$$

Then

$$\lim_{n\to\infty} \frac{1}{n}\left|\log Z(\mathbb{G}) - B\right| = 0 \qquad \text{in probability.}$$

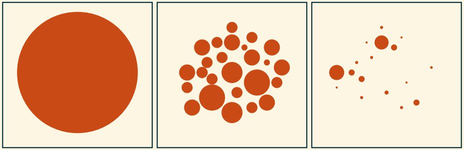
Bethe states



Absence of replica symmetry

- let us drop the assumption of replica symmetry
- we expect any number of Bethe states

Bethe states



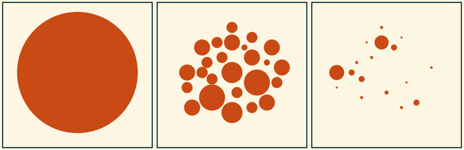
The conditional standard messages

For $S \subset \Omega^n$ let

$$\mu_{\mathbb{G},x\to a}(s \mid S) = \mu_{\mathbb{G}-a,x}(s \mid S) \qquad (s \in \Omega)$$

$$\mu_{\mathbb{G},a\to x}(s \mid S) = \mu_{\mathbb{G}-(\partial x \setminus a),x}(s \mid S) \qquad (s \in \Omega)$$

Bethe states



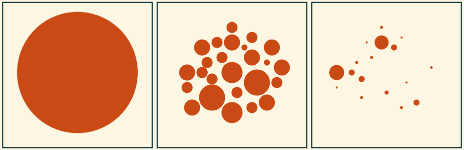
Definition

A set $S \subset \Omega^n$ is an *ε-Bethe state* if

$$\frac{1}{n}\sum_{i=1}^{n}\sum_{a\in\partial x_i}\left|\mu_{\mathbb{G},x_i\to a}(s\mid S)-\frac{\prod_{b\in\partial x_i\setminus a}\mu_{\mathbb{G},b\to x_i}(s\mid S)}{\sum_{t\in\Omega}\prod_{b\in\partial x_i\setminus a}\mu_{\mathbb{G},b\to x_i}(t\mid S)}\right|<\varepsilon$$

$$\frac{1}{n}\sum_{i=1}^{n}\sum_{a\in\partial x_i}\left|\mu_{\mathbb{G},a\to x_i}(s\mid S)-\frac{\sum_{\sigma\in\Omega^{\partial a}}\mathbf{1}\{\sigma_{x_i}=s\}\psi_a(\sigma)\prod_y\mu_{\mathbb{G},y\to a}(\sigma_y\mid S)}{\sum_{\sigma\in\Omega^{\partial a}}\psi_a(\sigma)\prod_{y\in\partial a\setminus x_i}\mu_{\mathbb{G},y\to a}(\sigma_y\mid S)}\right|<\varepsilon$$

Bethe states



Theorem [COP19]

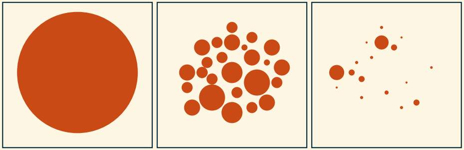
For any $\varepsilon > 0$ there exist $L \geq 1$ and $n_0 > 0$ such that for all $n > n_0$ with high probability there exist $S_1, \ldots, S_\ell \subset \Omega^n$, $\ell \leq L$, such that

(i) $S_1, \ldots, S_\ell$ are ε-Bethe states

(ii) $\sum_{i=1}^{\ell} \mu_{\mathbb{G}}(S_i) > 1 - \varepsilon$

"any random factor graph model has a Bethe state decomposition"

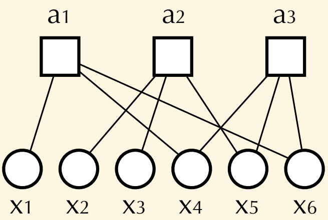
Bethe states



Proof

- apply pinning repeatedly like in the decomposition theorem
- to each sub-cube apply a coupling argument as in the RS case
- delicate point: the “edits” shift the relative weights of the Bethe states

Random k-XORSAT



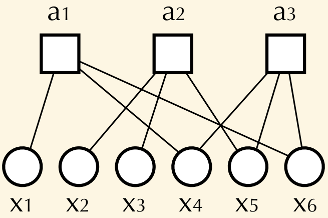
The random k-XORSAT problem

- variables $x_1, \ldots, x_n$ ranging over $\Omega = \mathbb{F}_2$
- check nodes $a_1, \ldots, a_{\boldsymbol{m}}$ represent k-XOR constraints:

$$x_{i_1} + x_{i_2} + \cdots + x_{i_k} = y_i$$

- with $\boldsymbol{m} \sim \mathrm{Po}(dn/k)$, for what d is it possible to satisfy all constraints?

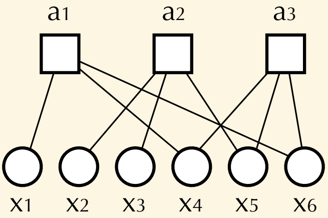
Random k-XORSAT



Equivalent formulation

- let $\boldsymbol{A}$ be the random (bi)adjacency matrix
- for what d, k does $\boldsymbol{A}$ have full row rank?
- equivalently, determine the dimension of the kernel
- $Z = \#\{\text{solutions to } \boldsymbol{A}x = 0\}$
- $\mu_{\boldsymbol{A}}(\sigma) = \mathbf{1}\{\boldsymbol{A}\sigma = 0\}/Z$

Random k-XORSAT



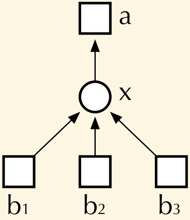
Trivial BP solution

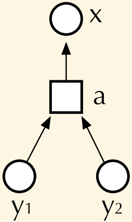
- set all messages to 1/2:

$$\mu_{x\to a}(0) = \mu_{x\to a}(1) = \frac{1}{2}$$

$$\mu_{a\to x}(0) = \mu_{a\to x}(1) = \frac{1}{2}$$

Random k-XORSAT





Trivial BP solution

- set all messages to 1/2:

$$\mu_{x\to a}(0) = \mu_{x\to a}(1) = \frac{1}{2} \qquad \mu_{a\to x}(0) = \mu_{a\to x}(1) = \frac{1}{2}$$

- Bethe free entropy $= n(1 - d/k)\log 2$
- *Are there other fixed points* (there'd better be)?

Random k-XORSAT

Lemma

For any $\boldsymbol{m} \times n$-matrix A for a random $\boldsymbol{\sigma} \in \ker A$ we have

$$\mathrm{P}\left[\boldsymbol{\sigma}_i = 0\right] \in \{1/2, 1\}$$

If $\mathrm{P}\left[\boldsymbol{\sigma}_i = 0\right] = 1$, *call coordinate i frozen*.

Proof

- consider a basis $\xi_1, \ldots, \xi_\ell$ of the kernel
- $\boldsymbol{\sigma} = \boldsymbol{\omega}_1 \xi_1 + \cdots + \boldsymbol{\omega}_\ell \xi_\ell$

Random k-XORSAT

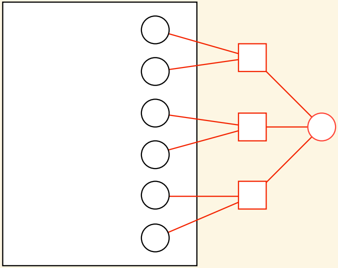
Bethe states of k-XORSAT

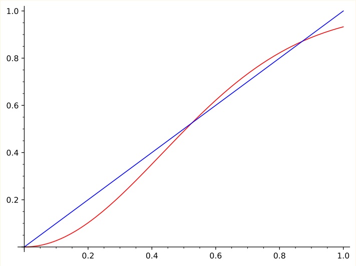
- consider the Bethe states $S_1, \ldots, S_L$
- for each of them let $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_L$ be the fraction of frozen variables:

$$\boldsymbol{\alpha}_j = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}\left\{\mathrm{P}\left[\boldsymbol{\sigma}_i = 0 \mid S_j\right] = 1\right\}$$

- *what values can $\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_L$ take?*

Random k-XORSAT





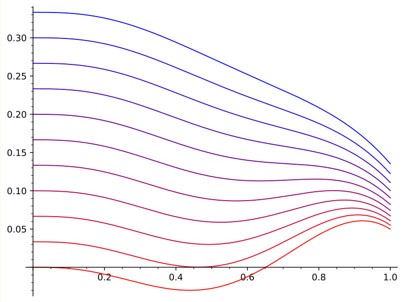
The fixed point property

- The $\boldsymbol{\alpha}_i$ satisfy the fixed point equation

$$\alpha = 1 - \exp(-d\alpha^{k-1})$$

- this equation has at most three solutions
- at most two of them are stable

Random k-XORSAT



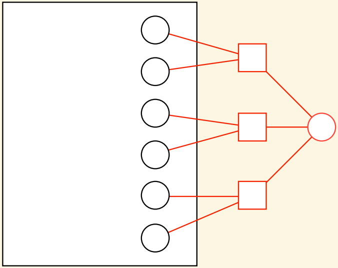
The ensuing Bethe free entropy

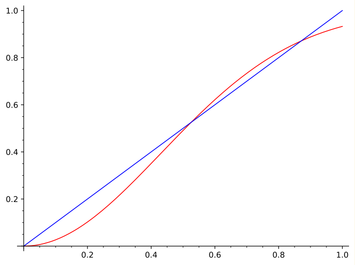
- the fixed points translate into stationary points of the BFE

$$\mathscr{B}(\alpha) = \exp(-d\alpha^{k-1}) - \frac{d}{k}\left(1 - k\alpha^{k-1} + (k-1)\alpha^k\right)$$

- the stable ones are local maxima
- from a certain threshold $d^* = d^*(k)$ the positive one dominates

Random k-XORSAT

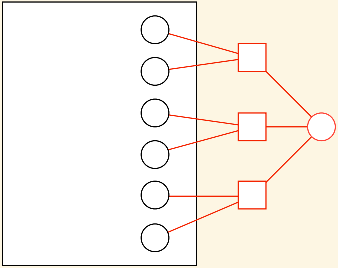


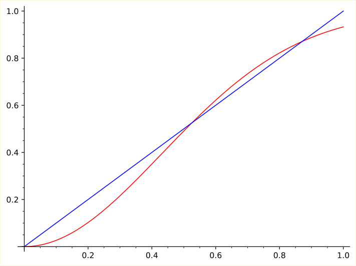


The method of moments

- we can finally calculate the *expected* number of solutions to $\boldsymbol{A}x = 0$ with a certain fraction of frozen coordinates
- if that fraction is a fixed point, the answers (with suitable conditioning) boils down to $\mathscr{B}(\alpha)$

Random k-XORSAT





Theorem [DM02,MRTZ02,DGMMPR10,PS16]

The random k-XORSAT satisfiability threshold equals $d^*(k)$.

Random k-XORSAT

Summary

- random factor graph models possess Bethe states
- they can be constructed obliviously via pinning
- we can harness Bethe decompositions to derive combinatorial results
- *Example:* the k-XORSAT threshold