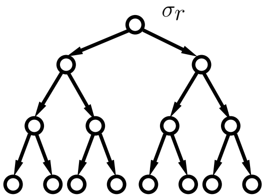
Decay of Correlations for the Hardcore Model in Random Regular Graphs

Nayantara Bhatnagar (University of Delaware)

joint work with

Allan Sly (UC Berkeley) and Prasad Tetali (Georgia Tech)

Information Flow on Trees



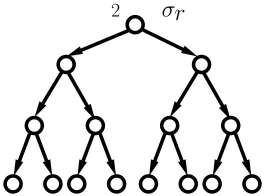
Broadcast Model

- Infinite d-ary tree T_d.
- Spins $[q]$, $\sigma_v \in [q]$ is spin at v.
- Choose $\sigma_r \sim \pi$.
- Independent Markov chain M on each edge of T.

Motivation

- Evolution of DNA, Noisy computation.
- Broadcast model gives rise to the **free Gibbs measure**.
- Random constraint satisfaction.

Information Flow on Trees



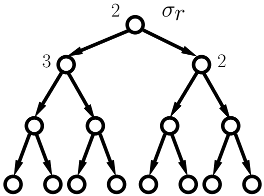
Broadcast Model

- Infinite d-ary tree T_d.
- Spins $[q]$, $\sigma_v \in [q]$ is spin at v.
- Choose $\sigma_r \sim \pi$.
- Independent Markov chain M on each edge of T.

Motivation

- Evolution of DNA, Noisy computation.
- Broadcast model gives rise to the **free Gibbs measure**.
- Random constraint satisfaction.

Information Flow on Trees



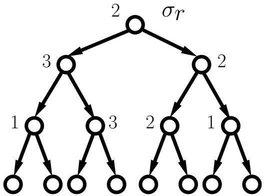
Broadcast Model

- Infinite d-ary tree T_d.
- Spins $[q]$, $\sigma_v \in [q]$ is spin at v.
- Choose $\sigma_r \sim \pi$.
- Independent Markov chain M on each edge of T.

Motivation

- Evolution of DNA, Noisy computation.
- Broadcast model gives rise to the **free Gibbs measure**.
- Random constraint satisfaction.

Information Flow on Trees



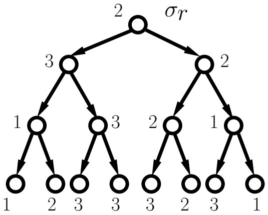
Broadcast Model

- Infinite d-ary tree T_d.
- Spins $[q]$, $\sigma_v \in [q]$ is spin at v.
- Choose $\sigma_r \sim \pi$.
- Independent Markov chain M on each edge of T.

Motivation

- Evolution of DNA, Noisy computation.
- Broadcast model gives rise to the **free Gibbs measure**.
- Random constraint satisfaction.

Information Flow on Trees



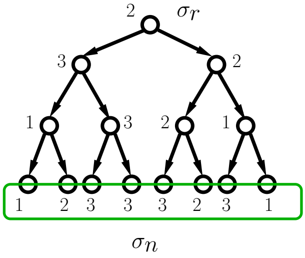
Broadcast Model

- Infinite d-ary tree T_d.
- Spins $[q]$, $\sigma_v \in [q]$ is spin at v.
- Choose $\sigma_r \sim \pi$.
- Independent Markov chain M on each edge of T.

Motivation

- Evolution of DNA, Noisy computation.
- Broadcast model gives rise to the **free Gibbs measure**.
- Random constraint satisfaction.

Information Flow on Trees



Broadcast Model

- Infinite d-ary tree T_d.
- Spins $[q]$, $\sigma_v \in [q]$ is spin at v.
- Choose $\sigma_r \sim \pi$.
- Independent Markov chain M on each edge of T.

Motivation

- Evolution of DNA, Noisy computation.
- Broadcast model gives rise to the **free Gibbs measure**.
- Random constraint satisfaction.

What is reconstruction?

- Reconstruction is **solvable** if for some i, j

$$\limsup_n d_{tv}(\sigma_n^i, \sigma_n^j) > 0$$

- Mutual information between values at level n and root does not go to 0.

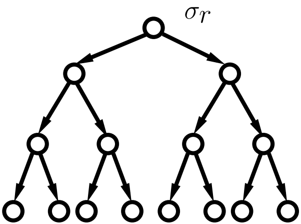
$$\mathbb{P}(\sigma_r = i | \sigma_n) \not\to \pi_i \quad \text{a.s.}$$

- Otherwise, **non-solvable** or **non-reconstruction**.

Related to...

- Can the phylogenetic tree be constructed efficiently from DNA?
- Extremality of the free Gibbs measure.
- Phase transitions in the solution space of random constraint satisfaction problems.

Symmetric channel

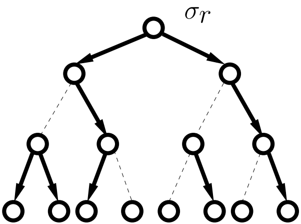


- Keep state w.p. λ, randomize w.p., $1-\lambda$.
- Second eigenvalue $\lambda(M) = \lambda$.
- $$M_{i,j} = \lambda \mathbb{1}_{i=j} + \frac{1-\lambda}{q}.$$
- Critical λ_c s.t.
  - Reconstruction solvable for $\lambda > \lambda_c$.
  - Non-solvable for $0 < \lambda < \lambda_c$.

Percolation bound

Non-solvable when $\lambda d \leq 1$. All components finite a.s.

Symmetric channel

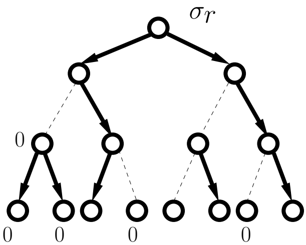


- Keep state w.p. λ, randomize w.p., $1-\lambda$.
- Second eigenvalue $\lambda(M) = \lambda$.
- $$M_{i,j} = \lambda \mathbb{1}_{i=j} + \frac{1-\lambda}{q}.$$
- Critical λ_c s.t.
  - Reconstruction solvable for $\lambda > \lambda_c$.
  - Non-solvable for $0 < \lambda < \lambda_c$.

Percolation bound

Non-solvable when $\lambda d \leq 1$. All components finite a.s.

Symmetric channel

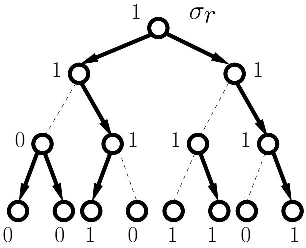


- Keep state w.p. λ, randomize w.p., $1-\lambda$.
- Second eigenvalue $\lambda(M) = \lambda$.
- $$M_{i,j} = \lambda \mathbb{1}_{i=j} + \frac{1-\lambda}{q}.$$
- Critical λ_c s.t.
  - Reconstruction solvable for $\lambda > \lambda_c$.
  - Non-solvable for $0 < \lambda < \lambda_c$.

Percolation bound

Non-solvable when $\lambda d \leq 1$. All components finite a.s.

Symmetric channel



- Keep state w.p. λ, randomize w.p., $1-\lambda$.
- Second eigenvalue $\lambda(M)=\lambda$.
- $$M_{i,j}=\lambda\mathbb{1}_{i=j}+\frac{1-\lambda}{q}.$$
- Critical λ_c s.t.
  - Reconstruction solvable for $\lambda>\lambda_c$.
  - Non-solvable for $0<\lambda<\lambda_c$.

Percolation bound

Non-solvable when $\lambda d\le 1$. All components finite a.s.

Kesten-Stigum bound

Theorem (Kesten-Stigum '68)

For $\lambda^2 d > 1$, reconstruction is (census) solvable.

Theorem (Mossel-Peres '03)

The threshold for census reconstruction is $\lambda^2 d = 1$.

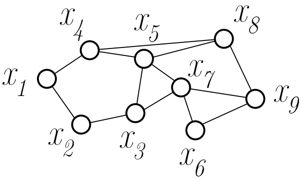
- *Kesten-Stigum CLT's imply a.s. the frequencies are π.*

Tightness of Kesten-Stigum

- [Bleher-Ruiz-Zagrebnov '95], [Ioffe '96], [Evans-Kenyon-Peres-Schulman '00] K-S bound tight for binary symmetric channel ($q = 2$, Ising model).
- [Mossel '01] When $q \geq 18$, K-S bound is **not** tight. Reconstruction beyond census.
- [Mezard-Montanari '06] Statistical physics predictions: for $\theta > 0$ (ferromagnetic)
  - K-S bound tight when $q \leq 4$.
  - K-S bound **not** tight for $q \geq 5$.

  [Sly '08] rigorously verified these predictions, for large enough d.
- [Brightwell-Winkler '04] Reconstruction beyond K-S in the "hardcore model" of lattice gases.

Graphical models



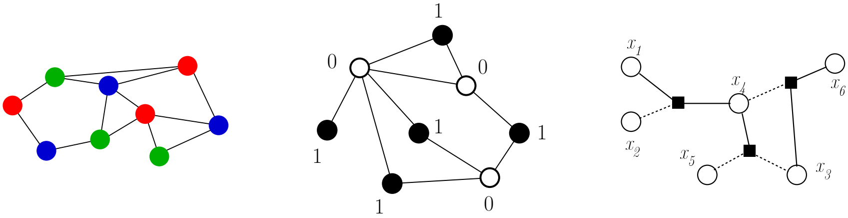
$$G = (V, E), \quad |V| = n, \quad \underline{x} = (x_1, \dots, x_n), \quad x_i \in [q], \quad \psi : [q] \times [q] \to \mathbb{R}.$$

$$\mu(\underline{x}) = \frac{1}{Z} \prod_{(x_i \sim x_j)} \psi(x_i, x_j)$$

- G has bounded degree.
- G is random.
- G has girth $g(n) \to \infty$, except for $o(n)$ vertices.

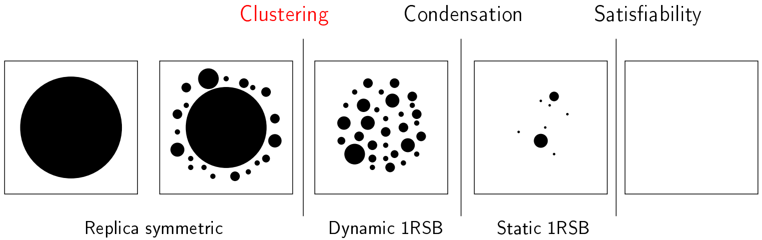
Random constraint satisfaction problems

A CSP is a set of variables subject to some constraints: decide whether there exists some variable assignment satisfying all constraints.



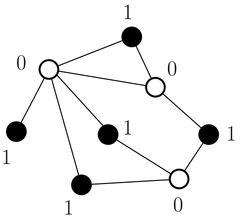
- Proper k-colorings of random graph (of average degree d).
- Independent sets (hardcore model) of a random graph.
- Random k-satisfiability.
- Using theory of spin glasses (random instances of satisfiability)...

Clustering of the solution space in random CSPs



- Replica symmetric (decay of correlations) vs. replica symmetry breaking (long range correlations).
- [MM '06] conjectured that 1 step replica symmetry breaking occurs exactly when the K-S bound is not tight.
- For $d < d_c(k)$ almost all solutions are contained in a giant cluster.
- For $d > d_c(k)$, there are exponentially many clusters. [Mezard-Mora-Zecchina '05], [Achlioptas-Coja-Oghlan '08]
- [Mezard-Montanari '06] conjectured that clustering threshold coincides with tree reconstruction. Shown asymptotically in [Montanari-Restrepo-Tetali '09].

Hardcore Model



- Probability distribution over independent sets (ISs), fugacity $\lambda > 0$,

$$\mathbb{P}(\sigma) = \frac{1}{Z}\lambda^{|\sigma|}\mathbb{1}(\sigma \text{ is an independent set})$$

- Partition function Z

$$Z = \sum_{\sigma \in I(G)} \lambda^{|\sigma|}$$

Reconstruction for hardcore measure on the tree T_d

- Broadcast process for the hardcore measure on T_d

$$M = \begin{pmatrix} p_{11} & p_{10} \\ p_{01} & p_{00} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ \frac{\alpha}{1-\alpha} & \frac{1-2\alpha}{1-\alpha} \end{pmatrix}$$

- Density of independent sets $\alpha = \alpha(\lambda, d)$ where

$$\lambda = \frac{\alpha}{1-\alpha}\left(\frac{1-\alpha}{1-2\alpha}\right)^d .$$

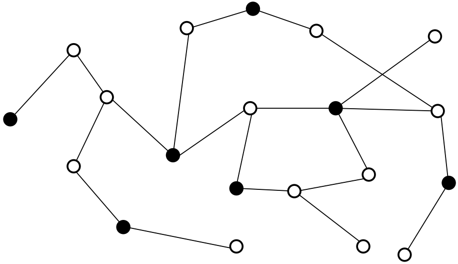
- Reconstruction solvable for λ if $\mathbb{P}(\sigma_r = 1|\sigma_n) \not\to \alpha(\lambda, d)$.
- [Brightwell-Winkler '04] Reconstruction solvability for

$$\alpha \geq \frac{\ln d + \ln\ln d + 1 + o(1)}{d}, \quad \lambda \geq (e + o(1))(\ln d)^2.$$

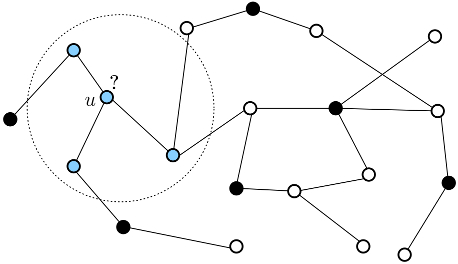
- [Martin '03] Non-solvable for $\lambda \leq e - 1$.
- [B-Sly-Tetali '10] Non-solvable for $\lambda \leq \frac{(\ln 2 - o_d(1))\ln(d)}{2\ln\ln d}$,

$$\alpha \leq \frac{\ln d + \ln\ln d - \ln\ln\ln d - \ln 2 + \ln\ln 2 - o(1)}{d}, .$$

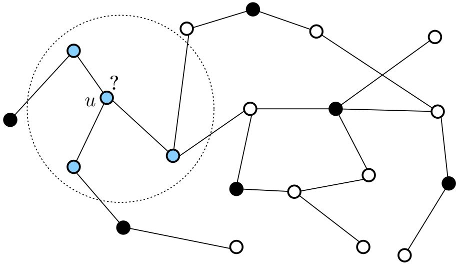
Reconstruction in graphs



Reconstruction in graphs



Reconstruction in graphs



- $\{G_n\}$ a family of random graphs, $|V(G_n)| \to \infty$.
- $\{G_n\}$ has non-reconstruction if for uniform $u \in V(G_n)$.

$$\lim_{L\to\infty} \limsup_{n} \mathbb{E}\Big|\mathbb{P}\left(\sigma_u = 1 | \sigma(\partial B_u(L)), u\right) - \alpha(\lambda, d)\Big| = 0$$

Reconstruction in graphs through local weak convergence

Theorem (B-Sly-Tetali '16)

For $\lambda < \lambda_c$, and large enough d, the d-regular random graph has non-reconstruction iff the d-regular tree has non-reconstruction.

$$\mathbb{P}_{T_d}(\sigma_x = 1 | \sigma_{L_\ell}) \xrightarrow{\mathbb{P}} \alpha(\lambda, d) \text{ as } \ell \to \infty.$$

$$\Leftrightarrow$$

$$\lim_{\ell \to \infty} \limsup_n \mathbb{E} \left| \mathbb{P}_{G_n} \left(\sigma_u = 1 | \sigma(\partial B_u(\ell)) \right) - \alpha(\lambda, d) \right| = 0$$

where λ_c is the fugacity corresponding to

$$\alpha_c = \frac{2 \ln d - (3 + o(1) \ln \ln d)}{d}.$$

Convergence in probability locally

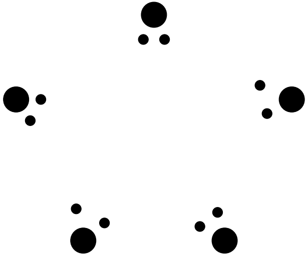
[Montanari-Mossel-Sly '12] Convergence in probability locally:

- For all r, $\varepsilon > 0$

$$\lim_{n\to\infty} \mathbb{P}_{U_n}\left(d_{tv}(\mathbb{P}_{G_n}(\sigma(B_r(u) \in \cdot)), \mathbb{P}_{T_d}(\sigma(B_r(x)) \in \cdot))) > \varepsilon\right) = 0$$

- Idea: use a type of second moment argument to show convergence in probability locally of hardcore measure to tree measure.

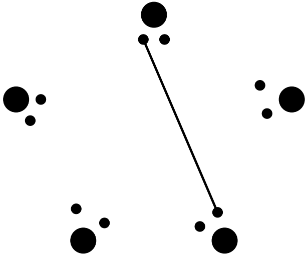
The configuration model



- Analyze the configuration model - d-regular multigraphs on n-vertices $\mathcal{C}(n, d)$.

$$\mathbb{P}(\text{Simple}) = (1 + o(1)) \exp\left(\frac{1 - d^2}{4}\right)$$

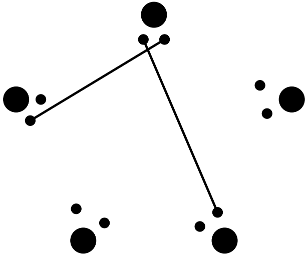
The configuration model



- Analyze the configuration model - d-regular multigraphs on n-vertices $\mathcal{C}(n, d)$.

$$\mathbb{P}(\text{Simple}) = (1 + o(1)) \exp\left(\frac{1 - d^2}{4}\right)$$

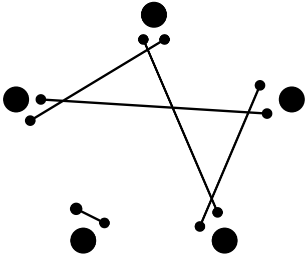
The configuration model



- Analyze the configuration model - d-regular multigraphs on n-vertices $\mathcal{C}(n,d)$.

$$\mathbb{P}(\text{Simple}) = (1+o(1))\exp\left(\frac{1-d^2}{4}\right)$$

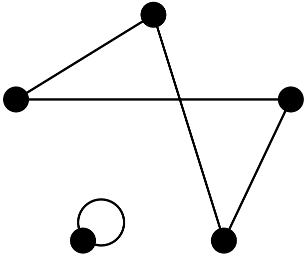
The configuration model



- Analyze the configuration model - d-regular multigraphs on n-vertices $\mathcal{C}(n, d)$.

$$\mathbb{P}(\text{Simple}) = (1 + o(1)) \exp\left(\frac{1 - d^2}{4}\right)$$

The configuration model



- Analyze the configuration model - d-regular multigraphs on n-vertices $\mathcal{C}(n, d)$.

$$\mathbb{P}(\text{Simple}) = (1 + o(1)) \exp\left(\frac{1 - d^2}{4}\right)$$

The partition function of the hardcore model on $\mathcal{C}(n,d)$

$$Z_{G,\alpha} := \sum_{\sigma:|\sigma|=\alpha n} \lambda^{\alpha n}, \qquad Z_G := \sum_{\alpha} Z_{G,\alpha}.$$

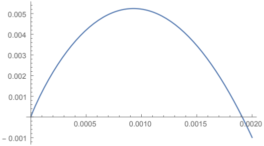


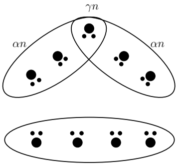
Figure: $\Phi(\alpha)$ vs. α.

$$\mathbb{E}\left(Z_{G,\alpha}\right) = \binom{n}{\alpha n}\lambda^{\alpha n}\prod_{i=0}^{\alpha n d-1}\frac{(1-\alpha)nd-i}{nd-1-2i} = \widetilde{\Theta}(1)\exp\left(n\Phi(\alpha)\right).$$

$\Phi(\alpha)$ maximized at α^* which solves $\lambda = \frac{\alpha}{1-\alpha}\left(\frac{1-\alpha}{1-2\alpha}\right)^d$.

$$\alpha^*(\lambda,d) = (1+o_d(1))\frac{\ln(\lambda d/\ln(\lambda d))}{d}.$$

The second moment of $Z_{G,\alpha}$



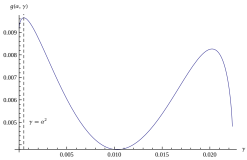
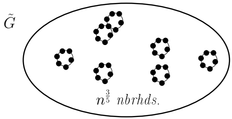


Figure: Projection $g(\alpha, \gamma)$ of f onto the maximizer $\overline{\varepsilon}(\alpha, \gamma)$.

Main facts/results: $G \sim \mathcal{C}(n, d)$

- $\mathbb{E}((Z_{G,\alpha})^2) = \sum_{\gamma,\varepsilon} \mathbb{E}(Z^{(2)}_{G,\alpha,\gamma,\varepsilon}) = \exp(nf(\alpha,\gamma,\varepsilon) + O(\ln(n)))$
- For any $0 \le \alpha \le \frac{1}{2}$, $2\Phi(\alpha) = f(\alpha, \alpha^2, \alpha(1-2\alpha))$.
- Let $\lambda < \lambda_c$. The function f is maximized at $(\alpha^*, (\alpha^*)^2, \alpha^*(1-2\alpha^*))$ and decays quadratically about this point.
- $\mathbb{E}((Z_G)^2) = \tilde{\Theta}(1)(\mathbb{E}(Z_G))^2$.

Relating graph and tree measures: punctured graph $\tilde{G}$



$\tilde{G}$: Choose $n^{3/5}$ vertices at random and delete their local neighborhood of radius r.

Study Z_G conditioned on the boundaries of local neighborhoods B.

Main results:

- Whp $O(n^{3/5})$ vertices of B have degree $d-1$ and $O(n^{1/5})$ have degree $d-2$.
- $\mathbb{E}(Z_{\tilde{G},\sigma})$ is proportional to a product measure on B.
- Second moment calculation for $\tilde{G}$. For each $\sigma \in \{0,1\}^B$

$$\mathbb{E}((Z_{G,\sigma})^2) = \exp\left(O(n^{1/5})\right)\mathbb{E}(Z_{G,\sigma})^2 \tag{1}$$

A set of bad boundaries $\mathcal{B}$

Local weak convergence: Let $G_n \sim \mathcal{C}(n,d)$. Then,

$$\frac{|\{i : \|\mathbb{P}_{G_n}(\partial B_r(s_i) \in \cdot), \mathbb{P}_{T_d}(\partial B_r(x) \in \cdot)\|_{tv} > \varepsilon\}|}{n^{3/5}} \xrightarrow{\mathbb{P}} 0$$

If not, there is a set $\mathcal{B}$ of boundary configurations on which graph and tree measure differ.

- By (1) and Azuma's ineq, $\mathcal{B}$ has large stationary prob. :

$$\mathbb{P}\left(\sum_{\sigma\in\mathcal{B}} Z_{G,\sigma} \geq \exp\left(-O(n^{\frac{1}{2}+\delta(\varepsilon)})\right)\mathbb{E}\left(Z_G\right)\right) \geq 1 - 2\exp\left(-n^{2\delta(\varepsilon)}\right).$$

- $\mathbb{E}(Z_G)$ and $\mathbb{E}\left(\sum_{\sigma\in\mathcal{B}} Z_{G,\sigma}\right)$ related by tree measure $\mathbb{P}_{T_d}(\mathcal{B})$.
- On these events, the part. fn. is much larger than expected

$$\sum_{\sigma\in\mathcal{B}} Z_{G,\sigma} \geq \exp\left(cn^{\frac{3}{5}}\right)\mathbb{E}\left(\sum_{\sigma\in\mathcal{B}} Z_{G,\sigma}\right).$$

Second moment implies $\mathcal{B}$ is unlikely

$$
\begin{aligned}
&\mathbb{P}\left(\sum_{\sigma\in\mathcal{B}} Z_{G,\sigma} \geq \exp\left(-O(n^{\frac{1}{2}+\delta})\right)\mathbb{E}\left(Z_G\right)\right) \\
&\leq \frac{\mathbb{E}\left(\sum_{\sigma\in\mathcal{B}} Z_{G,\sigma}\right)}{\exp\left(-O(n^{\frac{1}{2}+\delta})\right)\mathbb{E}\left(Z_G\right)} \qquad \text{(Markov's ineq.)} \\
&\leq \frac{\exp\left(-cn^{\frac{3}{5}}\right)}{\exp\left(-O(n^{\frac{1}{2}+\delta})\right)} \leq \exp\left(-cn^{\frac{3}{5}}\right)
\end{aligned}
$$

by choosing δ such that $\frac{3}{5} > \frac{1}{2} + \delta$.

Phase transitions for the hardcore measure

- [Barbier-Krzakala-Zdeborová-Zhang '13] Several earlier works showed the hardcore model undergoes a continuous "full replica symmetry breaking" transition for graphs of small degree.
- [Coja-Oghlan-Efthymiou '11] Clustering for hardcore measure in Erdös-Rényi graphs beyond $(1 + o_d(1)) \log(d)/d$. [Rahman-Virag '14, Gamarnik-Sudan '14] random regular graphs.
- [BKZZ13] use cavity methods to show that for degree $d > 16$, there is a discontinuous phase transition while for $d < 16$, it is continuous.
- [BKZZ13] "1 RSB approach": condensation transition exists and is at

$$\alpha_c \approx \frac{2(\log(d) + \log\log(d) + 1 - \log(2))}{d} + o(1/d).$$

Thank you for your attention!