Disordered Systems and Random Graphs 1

Amin Coja-Oghlan
Goethe University

based on joint work with Dimitris Achlioptas, Oliver Gebhard, Max Hahn-Klimroth, Joon Lee, Philipp Loick, Noela Müller, Manuel Penschuck, Guangyan Zhou

Overview

Lecture 1: introduction

- random graphs and phase transitionss
- the cavity method
- first/second moment method
- Belief Propagation and density evolution

Overview

Lecture 2: random 2-SAT

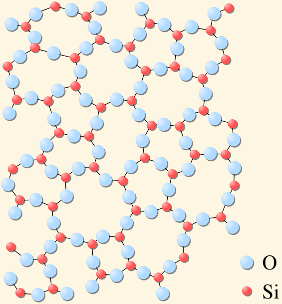
- the contraction method
- spatial mixing
- the Aizenman-Sims-Starr scheme
- the interpolation method

Overview

Lecture 3: group testing

- basics of Bayesian inference
- analysis of combinatorial algorithms
- spatial coupling
- information-theoretic lower bounds

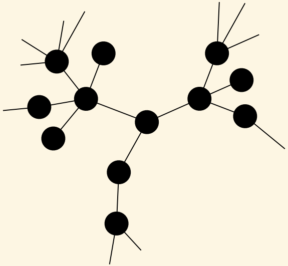
Disordered systems



From glasses to random graphs [MP00]

- (spin) glasses are disordered materials rather than crystals
- lattice models are difficult to grasp even non-rigorously
- classical mean-field models: complete interaction
- *diluted mean-field models:* sparse random graph topology

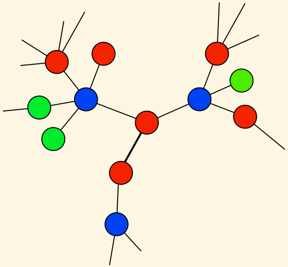
Disordered systems



The binomial random graph $\mathbb{G} = \mathbb{G}(n, p)$ [ER60]

- vertex set $x_1, \ldots, x_n$
- connect any two vertices w/ probability $p = \frac{d}{n}$ independently
- local structure converges to Po(d) Galton-Watson tree

The Potts antiferromagnet



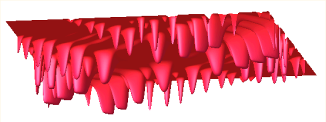
Definition

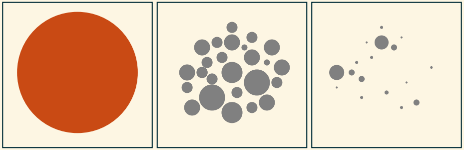
- fix $d > 0$, $q \geq 2$ and $\beta > 0$
- the Boltzmann distribution reads

$$\mu_{\mathbb{G},\beta}(\sigma) = \frac{1}{Z(\mathbb{G},\beta)} \prod_{vw \in E(\mathbb{G})} \exp(-\beta \mathbf{1}\{\sigma_v = \sigma_w\}) \qquad (\sigma \in \{1,\ldots,q\}^n)$$

$$Z(\mathbb{G},\beta) = \sum_{\tau \in \{1,\ldots,q\}^n} \prod_{vw \in E(\mathbb{G})} \exp(-\beta \mathbf{1}\{\tau_v = \tau_w\})$$

The physics story: replica symmetry breaking





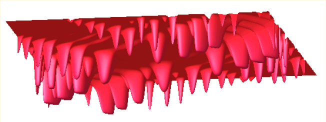
Replica symmetry [KMRTSZ07]

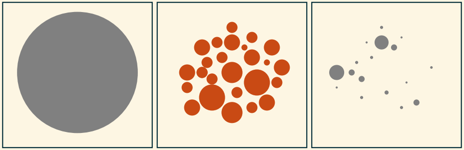
- fix a large d and increase β
- for small β there are no extensive long-range correlations

$$\mu_{\mathbb{G},\beta}(\{\boldsymbol{\sigma}_{x_1}=\tau_1,\boldsymbol{\sigma}_{x_2}=\tau_2\})\sim q^{-2} \qquad (\tau_1,\tau_2\in\{1,\ldots,q\})$$

- in fact, there is non-reconstruction and rapid mixing

The physics story: replica symmetry breaking





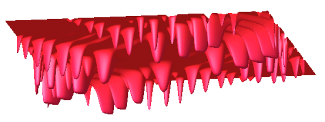
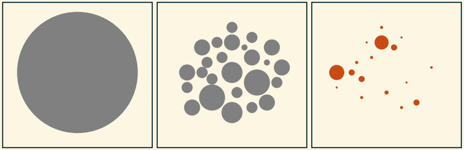
Dynamic replica symmetry breaking [KMRTSZ07]

- still no extensive long-range correlations for moderate β

$$\mu_{\mathbb{G},\beta}(\{\boldsymbol{\sigma}_{x_1} = \tau_1, \boldsymbol{\sigma}_{x_2} = \tau_2\}) \sim q^{-2} \qquad (\tau_1, \tau_2 \in \{1, \ldots, q\})$$

- but there is reconstruction and torpid mixing

The physics story: replica symmetry breaking

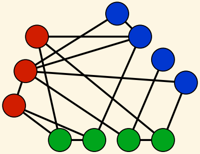
Static replica symmetry breaking [KMRTSZ07]

- for large β long-range correlations emerge

$$\mu_{\mathbb{G},\beta}(\{\boldsymbol{\sigma}_{x_1}=\tau_1,\boldsymbol{\sigma}_{x_2}=\tau_2\})\not\sim q^{-2}\qquad(\tau_1,\tau_2\in\{1,\ldots,q\})$$

- a few pure states dominate

The stochastic block model



The Potts model as an inference problem [DKMZ11]

- choose a random colouring $\boldsymbol{\sigma}^* \in \{1, \ldots, q\}^n$
- then choose a random graph $\mathbb{G}^*$ with

$$\mathrm{P}\left[\mathbb{G}^* = G \mid |E(\mathbb{G}^*)| = |E(G)|\right] \propto \mu_{G,\beta}(\boldsymbol{\sigma}^*)$$

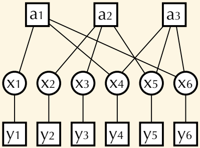
- given $\mathbb{G}^*$ can we (partly) infer $\boldsymbol{\sigma}^*$?

Rigorous work

Techniques

- Classical random graphs techniques
  - method of moments
  - branching processes
  - large deviations
- Mathematial physics techniques
  - coupling arguments
  - exchangeable arrays and the cut metric
  - Belief Propagation and the contraction method
  - the interpolation method

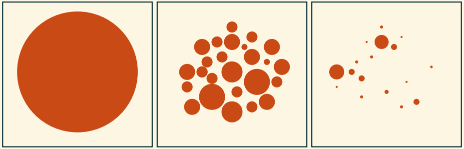
Rigorous work



Success stories

- solution space geometry [ACO08,M12]
- random k-SAT [AM02,AP03,COP16,DSS15]
- low-density parity check codes [G63,KRU13]
- stochastic block model [AS15,M14,MNS13,MNS14,COKPZ16]
- group testing [MTT08,COGHKL20]
- …

Rigorous work



Theorem [COKPZ17]

Let $\Lambda(x) = x \log x$ and

$$\mathscr{B}^*_{q,\beta}(d) = \sup_\pi \mathscr{B}_{q,\beta,d}(\pi) \qquad \text{where}$$

$$\mathscr{B}_{q,\beta,d}(\pi) = \mathrm{E}\left[\frac{\Lambda(\sum_{\sigma=1}^{q} \prod_{i=1}^{\gamma} 1-(1-\mathrm{e}^{-\beta})\boldsymbol{\mu}_i^{(\pi)}(\sigma))}{q(1-(1-\mathrm{e}^{-\beta})/q)^\gamma} - \frac{d}{2}\frac{\Lambda(1-(1-\mathrm{e}^{-\beta})\sum_{\sigma=1}^{q}\boldsymbol{\mu}_1^{(\pi)}(\sigma)\boldsymbol{\mu}_2^{(\pi)}(\sigma))}{1-(1-\mathrm{e}^{-\beta})/q}\right].$$

Then

$$d_{\mathrm{cond}}(q,\beta) = \inf\left\{d > 0 : \mathscr{B}^*_{q,\beta}(d) = \ln q + \frac{d}{2}\ln(1-(1-e^{-\beta})/q)\right\}.$$

Random 2-SAT

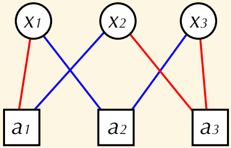
The 2-SAT problem

- Boolean variables $x_1, \ldots, x_n$
- truth values $+1$ and -1
- four types of clauses:

$$x_i \vee x_j \qquad x_i \vee \neg x_j \qquad \neg x_i \vee x_j \qquad \neg x_i \vee \neg x_j$$

- a 2-SAT formula is a conjunction $\Phi = \bigwedge_{i=1}^{m} a_i$ of clauses
- $S(\Phi) =$ set of satisfying assignments
- $Z(\Phi) = |S(\Phi)|$

Random 2-SAT



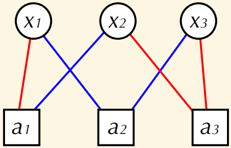
Example

- $\Phi = (\neg x_1 \vee x_2) \wedge (x_1 \vee x_3) \wedge (\neg x_2 \vee \neg x_3)$
- $Z(\Phi) = 2$ and $S(\Phi)$ consists of the two assignments

$$\sigma_{x_1} = +1 \qquad \sigma_{x_2} = +1 \qquad \sigma_{x_3} = -1$$
$$\sigma_{x_1} = -1 \qquad \sigma_{x_2} = -1 \qquad \sigma_{x_3} = +1$$

- glassy because variables may appear with opposing signs

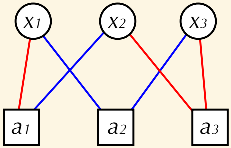
Random 2-SAT



Computational complexity

- 2-SAT admits an efficient decision algorithm [K67]
- in fact, `WalkSAT` solves the problem efficiently [P91]
- the problem is `NL`-complete [IS87,P94]
- however, computing $\log Z(\Phi)$ is #`P`-hard [V79]

Random 2-SAT



Random 2-SAT

- for a fixed $0 < d < \infty$ let $\boldsymbol{m} = \text{Po}(dn/2)$
- $\boldsymbol{\Phi}$ = conjunction of $\boldsymbol{m}$ independent random clauses
- variable degrees have distribution $\text{Po}(d)$
- *Key questions:* is $Z(\boldsymbol{\Phi}) > 0$ and if so, what is

$$\lim_{n\to\infty} \frac{1}{n} \log Z(\boldsymbol{\Phi}) \quad ?$$

Random 2-SAT

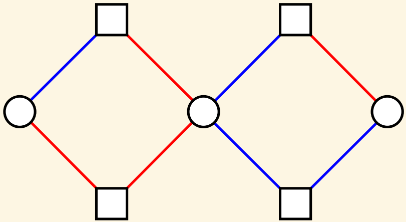
Prior work

- the threshold for $S(\mathbf{\Phi}) = \emptyset$ occurs at $d = 2$ [CR92,G96]
- computation of $\log Z(\mathbf{\Phi})$ via replica/cavity method [MZ96]
- the scaling window [BBCKW01]
- partial results on 'soft' version [T01,MS07,P14]
- existence of a function $\phi(d)$ such that [AM14]

$$\lim_{n\to\infty} \frac{\log Z(\mathbf{\Phi})}{n} = \phi(d)$$

for almost all $d \in (0,2)$

The satisfiability threshold



Bicycles

- the clause $l \vee l'$ is logically equivalent to the two implications

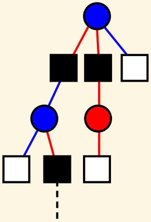
$$l \vee l' \equiv (\neg l \to l') \wedge (\neg l' \to l)$$

- $\mathbf{\Phi}$ is satisfiable unless there is an implication chain

$$x_i \to \cdots \to \neg x_i \to \cdots \to x_i$$

- such chains are called *bicycles*

The satisfiability threshold



Theorem [CR92,G96]

- If $d < 2$ then Φ does not contain a bicycle w.h.p.
- If $d > 2$ then Φ contains a bicycle w.h.p.

The second moment method

A naive attempt

- we aim to compute $\log Z(\mathbf{\Phi})$ for a typical $\mathbf{\Phi}$
- Jensen's inequality shows that

$$\log Z(\mathbf{\Phi}) \le \log \mathrm{E}[Z(\mathbf{\Phi}) \mid \boldsymbol{m}] + o(n) \qquad \text{w.h.p.}$$

The second moment method

The first moment

- computing $\mathrm{E}[Z(\mathbf{\Phi}) \mid \boldsymbol{m}]$ is a cinch:

$$\mathrm{E}[Z(\mathbf{\Phi}) \mid \boldsymbol{m}] = 2^n \cdot \left(\frac{3}{4}\right)^{\boldsymbol{m}}$$

- hence,

$$\frac{1}{n} \log Z(\mathbf{\Phi}) \leq (1-d)\log 2 + \frac{d}{2}\log 3 \qquad \text{w.h.p.}$$

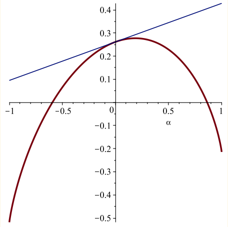
The second moment method

The second moment

- this bound is tight if $\mathrm{E}[Z(\mathbf{\Phi})^2] = O(\mathrm{E}[Z(\mathbf{\Phi})]^2)$
- we calculate

$$\begin{aligned}
\mathrm{E}[Z(\mathbf{\Phi})^2 \mid \boldsymbol{m}] &= \sum_{\sigma,\tau\in\{\pm1\}^n} \mathrm{P}\left[\mathbf{\Phi} \models \sigma, \mathbf{\Phi} \models \tau \mid \boldsymbol{m}\right] \\
&= \sum_{\ell=-n}^{n} \sum_{\sigma,\tau:\sigma\cdot\tau=\ell} \left(\frac{1}{2} + \frac{(1+\ell/n)^2}{16}\right)^{\boldsymbol{m}} \\
&= \sum_{\ell=-n}^{n} \binom{n}{(n+\ell)/2} \left(\frac{1}{2} + \frac{(1+\ell/n)^2}{16}\right)^{\boldsymbol{m}}
\end{aligned}$$

The second moment method



The second moment

- hence,

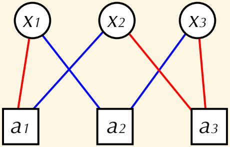
$$\frac{1}{n}\log \mathrm{E}[Z(\mathbf{\Phi})^2 \mid \boldsymbol{m}] \sim \max_{-1\leq\alpha\leq 1} H((1+\alpha)/2) + \frac{d}{2}\log\left(\frac{1}{2}+\frac{(1+\alpha)^2}{16}\right)$$

- at $\alpha = 0$ the above function evaluates to

$$\log 2 + d\log\frac{3}{4} \sim \frac{2}{n}\log \mathrm{E}[Z(\mathbf{\Phi}) \mid \boldsymbol{m}]$$

- therefore, we succeed iff the max is attained at $\alpha = 0$:(

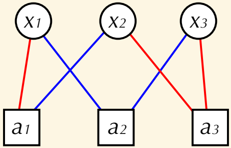
The cavity method



The factor graph

- vertices $x_1, \ldots, x_n$ represent variables
- vertices $a_1, \ldots, a_{\boldsymbol{m}}$ represent clauses
- the graph $G(\boldsymbol{\Phi})$ contains few short cycles
- locally $G(\boldsymbol{\Phi})$ resembles a Galton-Watson branching process

The cavity method



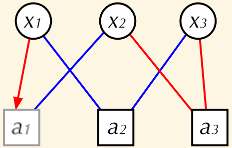
The Boltzmann distribution

- assuming $S(\mathbf{\Phi}) \neq \emptyset$ define

$$\mu_{\mathbf{\Phi}}(\sigma) = \frac{\mathbf{1}\{\sigma \in S(\mathbf{\Phi})\}}{Z(\mathbf{\Phi})} \qquad (\sigma \in \{\pm 1\}^{\{x_1,\ldots,x_n\}})$$

- let $\boldsymbol{\sigma} = \boldsymbol{\sigma}_{\mathbf{\Phi}}$ be a sample from $\mu_{\mathbf{\Phi}}$

The cavity method



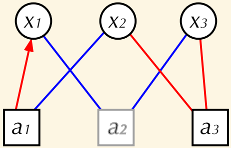
Belief Propagation

- define the variable–to–clause messages by

$$\mu_{\Phi,x\to a}(\sigma) = \mu_{\Phi-a}(\boldsymbol{\sigma}_x = \sigma) \qquad (\sigma = \pm 1)$$

- "marginal of x upon removal of a"

The cavity method



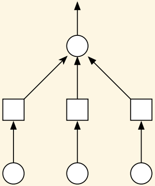
Belief Propagation

- define the clause–to–variable messages by

$$\mu_{\mathbf{\Phi},a\to x}(\sigma) = \mu_{\mathbf{\Phi}-(\partial x\setminus a)}(\boldsymbol{\sigma}_x = \sigma) \qquad\qquad (\sigma = \pm 1)$$

- “marginal of x upon removal of all neighbours $b \in \partial x$, $b \neq a$”

The cavity method



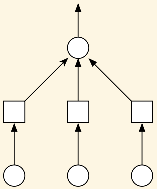
The replica symmetric ansatz

The messages (approximately) satisfy

$$\mu_{\mathbf{\Phi},x\to a}(\sigma) \propto \prod_{b\in\partial x\setminus a} \mu_{\mathbf{\Phi},b\to x}(\sigma)$$

$$\mu_{\mathbf{\Phi},a\to x}(\sigma) \propto 1 - \mathbf{1}\left\{\sigma \neq \mathrm{sign}(x,a)\right\} \mu_{\mathbf{\Phi},\partial a\setminus x}(-\mathrm{sign}(\partial a \setminus x))$$

The cavity method



The Bethe free entropy

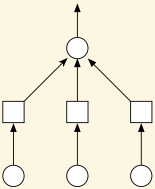
- we expect that

$$\log Z(\mathbf{\Phi}) \sim \sum_{i=1}^{n} \log \sum_{\sigma=\pm 1} \prod_{a \in \partial x_i} \mu_{\mathbf{\Phi},a\to x}(\sigma)$$

$$+ \sum_{i=1}^{\boldsymbol{m}} \log\left(1 - \prod_{x \in \partial a_i} \boldsymbol{\mu}_{\mathbf{\Phi},x\to a_i}(-\mathrm{sign}(x,a_i))\right)$$

$$- \sum_{i=1}^{n} \sum_{a \in \partial x_i} \log \sum_{\sigma=\pm 1} \mu_{\mathbf{\Phi},x\to a_i}(\sigma)\mu_{\mathbf{\Phi},a_i\to x}(\sigma)$$

The cavity method



Density evolution

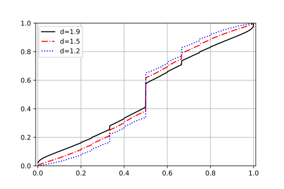
- consider the empirical distribution of the messages:

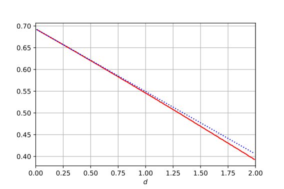
$$\pi_{\Phi} = \frac{1}{2\boldsymbol{m}} \sum_{i=1}^{n} \sum_{a \in \partial x_i} \delta_{\mu_{\Phi, x \to a}(+1)}$$

- $\boldsymbol{d}^+, \boldsymbol{d}^- \sim \mathrm{Po}(d/2)$, $\boldsymbol{\mu}_0, \boldsymbol{\mu}_1, \boldsymbol{\mu}_2, \ldots$ samples from π_{Φ}

$$\boldsymbol{\mu}_0 \overset{\mathrm{d}}{=} \frac{\prod_{i=1}^{\boldsymbol{d}_+} \boldsymbol{\mu}_i}{\prod_{i=1}^{\boldsymbol{d}_+} \boldsymbol{\mu}_i + \prod_{i=1}^{\boldsymbol{d}_-} \boldsymbol{\mu}_{i+\boldsymbol{d}^+}}$$

The cavity method





Summary: the replica symmetric prediction [MZ96]

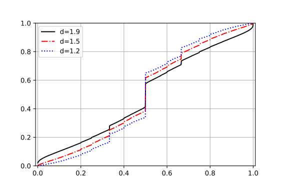
For $d < 2$ there is a unique distribution π_d on $(0,1)$ s.t.

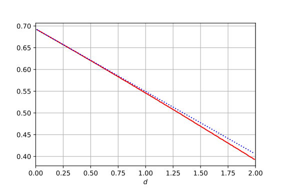
$$\boldsymbol{\mu}_0 \stackrel{\mathrm{d}}{=} \frac{\prod_{i=1}^{\boldsymbol{d}_+} \boldsymbol{\mu}_i}{\prod_{i=1}^{\boldsymbol{d}_+} \boldsymbol{\mu}_i + \prod_{i=1}^{\boldsymbol{d}_-} \boldsymbol{\mu}_{i+\boldsymbol{d}^+}}$$

and $\lim_{n\to\infty} n^{-1} \log Z(\boldsymbol{\Phi}) = \mathscr{B}_d$ where

$$\mathscr{B}_d = \mathrm{E}\left[\log\left(\prod_{i=1}^{\boldsymbol{d}_+} \boldsymbol{\mu}_i + \prod_{i=1}^{\boldsymbol{d}_-} \boldsymbol{\mu}_{i+\boldsymbol{d}_+}\right) - \frac{d}{2}\log\left(1 - \boldsymbol{\mu}_1 \boldsymbol{\mu}_2\right)\right]$$

The cavity method





Theorem [ACOHKLMPZ20]

For $d < 2$ there is a unique distribution π_d on $(0,1)$ s.t.

$$\boldsymbol{\mu}_0 \stackrel{\mathrm{d}}{=} \frac{\prod_{i=1}^{\boldsymbol{d}_+} \boldsymbol{\mu}_i}{\prod_{i=1}^{\boldsymbol{d}_+} \boldsymbol{\mu}_i + \prod_{i=1}^{\boldsymbol{d}_-} \boldsymbol{\mu}_{i+\boldsymbol{d}^+}}$$

and $\lim_{n\to\infty} n^{-1} \log Z(\boldsymbol{\Phi}) = \mathscr{B}_d$ where

$$\mathscr{B}_d = \mathrm{E}\left[\log\left(\prod_{i=1}^{\boldsymbol{d}_+} \boldsymbol{\mu}_i + \prod_{i=1}^{\boldsymbol{d}_-} \boldsymbol{\mu}_{i+\boldsymbol{d}_+}\right) - \frac{d}{2}\log\left(1 - \boldsymbol{\mu}_1 \boldsymbol{\mu}_2\right)\right]$$

References

- M. Mézard, G. Parisi: The Bethe lattice spin glass revisited. Eur Phys J B **20** (2001)
- P. Erdős, A. Rényi: On the evolution of random graphs. Publ Math Inst Hung Acad Sci (1960)
- F. Krzakala, A. Montanari, F. Ricci-Tersenghi, G. Semerjian, L. Zdeborová: Gibbs states and the set of solutions of random constraint satisfaction problems. PNAS **104** (2007)
- A. Decelle, F. Krzakala, C. Moore, L. Zdeborová: Asymptotic analysis of the stochastic block model for modular networks and its algorithmic applications. Phys Rev E **84** (2011)
- A. Coja-Oghlan, F. Krzakala, W. Perkins, L. Zdeborová: Information-theoretic thresholds from the cavity method. Advances in Mathematics **333** (2018).
- E. Abbé, A. Montanari: On the concentration of the number of solutions of random satisfiability formulas. RS&A **45** (2014)