Biased landscape in random constraint satisfaction problems

Louise Budzynski

LPENS, PhD with Guilhem Semerjian

June 30, 2020

Table of contents

Introduction

Random constraint satisfaction problems

Constraint satisfaction problems (CSPs): N discrete variables subjected to M constraints. A solution is an assignement that satisfies all the constraints.

An example of CSP: k-hypergraph bicoloring problem

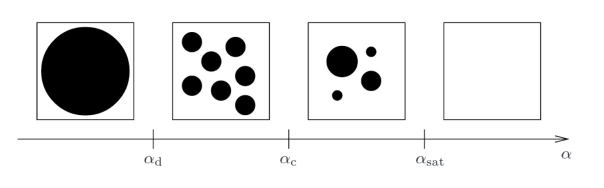
- hypergraph $G = (V, E)$: N vertices and M hyperedges (between k vertices)
- $\sigma_i \in \{+1, -1\}$ on the vertices
- $\underline{\sigma} = \{\sigma_1, \dots, \sigma_N\}$ solution $\iff$ for all hyperedges $\langle i_1, \dots, i_k \rangle \in E$, $\sigma_i \neq \sigma_j$, $\sigma_{i_1}, \dots, \sigma_{i_k}$ not all equal (at least one $+1$ and one -1)

Random graph ensembles: Thermodynamic limit $N, M \to \infty$ at $M/N = \alpha$ finite

regular ensemble: degree $l = \alpha k$ fixed, or Erdös Rényi: $\langle l \rangle = \alpha k$

Phase transitions in random CSPs

[Monasson, Zecchina 97], [Biroli, Monasson, Weigt 00], [Mézard, Parisi, Zecchina 02], [Krzakala, Montanari, Ricci-Tersenghi, Semerjian, Zdeborova 07], [Achlioptas, Coja-Oghlan 08], [Ding, Sly, Sun 14]



Focus on the clustering transition:

- clustering of the solution set
- exponential relaxation time of Monte Carlo Markov Chain [Montanari, Semerjian, 06]
- reconstruction on tree, apparition of long-range point-to-set correlations

Algorithmic performances

Open questions:

- estimate the putative algorithmic barrier $\alpha_{alg}(k)$ above which no algorithm can find solution in polynomial time, on large typical instances
- Can we relate α_{alg} to one of the phase transitions ?

Small values of k algos almost reach α_{sat} [Marino, Parisi, Ricci-Tersenghi, 15]

Large values of k:

- $\alpha_{sat}(k) \sim 2^{k-1} \ln 2$ and $\alpha_d(k) \sim 2^{k-1} \ln k/k$
- Best algorithm reaches α_d at the leading order (on k-SAT) [Coja-Oghlan, 10]

Biased measure over the set of solutions

Biased measure over the set of solutions

Phase transitions (clustering) are obtained for uniform measure:

$$\mu(\underline{\sigma}) = \frac{1}{Z} \begin{cases} 1 \text{ if } \underline{\sigma} \text{ is a solution} \\ 0 \text{ if } \underline{\sigma} \text{ is not a solution} \end{cases} \tag{1}$$

Introduce a non-uniform measure

$$\mu(\underline{\sigma}) = \frac{1}{Z} \begin{cases} b(\underline{\sigma}) & \text{if } \underline{\sigma} \text{ is a solution} \\ 0 & \text{if } \underline{\sigma} \text{ is not a solution} \end{cases} \tag{2}$$

α_d, α_c modified, but not α_{sat}.

Goal: Moving the clustering threshold α_d

Related works

- On hard spheres, with additional soft interactions [Sellitto, Zamponi 13], [Maimbourg, Sellito, Semerjian, Zamponi, 18]
- On the bicoloring problem, according to the number of frozen variables [Braunstein, Dall'Asta, Semerjian, Zdeborova 16]
- Local entropy [Baldassi, Ingrosso, Lucibello, Saglietti, Zecchina 16]

Biased measure

Specific bias: intra-clauses interactions [Budzynski, Ricci-Tersenghi, Semerjian, 19]

$$\mu(\underline{\sigma}) = \frac{1}{Z} \prod_{a \in E} \omega(\underline{\sigma}_{\partial a}) \tag{3}$$

with

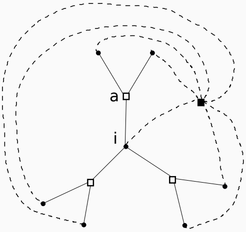
$$\omega(\sigma_1, \dots \sigma_k) = \begin{cases} 0 \text{ if } \sum_i \sigma_i = \pm k (\text{all equal}) \\ 1 - \epsilon \text{ if } \sum_i \sigma_i = \pm (k-2) \\ 1 \text{ otherwise} \end{cases} \tag{4}$$

Biased measure

Specific bias: interactions at distance 1 [Budzynski, Semerjian, in preparation]

$$\mu(\underline{\sigma}) = \frac{1}{Z} \prod_{a \in E} \mathbb{I}[\underline{\sigma}_{\partial a} \ n.a.e] \prod_{i \in V} \varphi(\sigma_i, \{\underline{\sigma}_{\partial a \setminus i}\}_{a \in \partial i})$$

Bias φ counts the number of forcing clauses: clause a is forcing i when $\underline{\sigma}_{\partial a \setminus i}$ all equal



$$\varphi(\sigma_i, \{\underline{\sigma}_{\partial a \setminus i}\}_{a \in \partial i}) = \psi(p_i), \quad p_i = \sum_{a \in \partial i} \mathbb{I}[\underline{\sigma}_{\partial a \setminus i} \text{ all equal}]$$

Finite k results

Clustering threshold l_d on random regular ensemble ($l = \alpha k$)

k	uniform	intra-clause	distance 1	l_{sat}
5	47	48	49	52
6	108	113	115	129

Table 1: Clustering threshold l_d optimized over the biases

Intra-clause: $\psi(p) = (1-\epsilon)^p$

distance 1: $\psi(0) = 1$, $\psi(1) = b_1$, $\psi(p \geq 2) = b_2(1-\epsilon)^p$

Large k asymptotics of the clustering transition

Asymptotics for the uniform measure

Scaling for the clustering transition:

$$\alpha_d(k) = \frac{2^{k-1}}{k}(\ln k + \ln\ln k + \gamma_d + o(1)) \tag{5}$$

What is known rigorously:

- dominant term (for a large class of models including bicoloring on k-hypergraphs) [Montanari, Restrepo, Tetali 11]
- for q-coloring: $1 - \ln 2 \leq \gamma_d \leq 1$ [Sly 09]
- for q-coloring: $\gamma_d < 1$ [Sly, Zhang 16]

Claim: [Budzynski, Semerjian 19]

$\gamma_d \simeq 0.871$ (for bicoloring on k-hypergraphs and q-coloring)

Asymptotics for the biased measure

Scalings for the bias parameters:

1. intra-clause bias: take $\epsilon = \frac{\widetilde{\epsilon}\sqrt{2}}{\sqrt{k \ln k}}$, $\widetilde{\epsilon}$ constant
2. interactions at distance 1: $\psi(0) = 1$, $\psi(p \geq 1) = b$, take b constant

Claim: [Budzynski, Semerjian, in preparation]

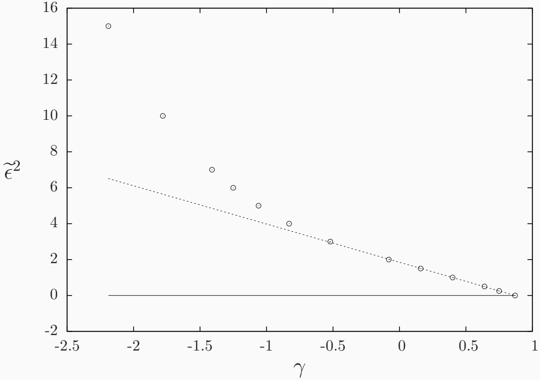
With this choice the clustering transition occurs at the same scale

$$\alpha_d = \frac{2^{k-1}}{k}(\ln k + \ln \ln k + \gamma_d + o(1)) \tag{6}$$

where γ_d depends on $\widetilde{\epsilon}, b$

Asymptotics for the biased measure

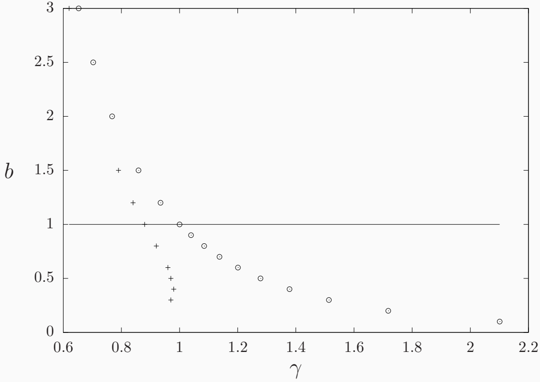
Results for the intra-clause bias:



$\gamma_d(\widetilde{\epsilon})$ with $\widetilde{\epsilon} \neq 0$ smaller than $\gamma_d(\widetilde{\epsilon} = 0) \simeq 0.87$ (uniform case)

Asymptotics for the biased measure

Results for the bias with interactions at distance 1:



Optimal value at $b_{opt} = 0.4$: $\gamma_d(b_{opt}) \simeq 0.98$ larger than $\gamma_d(b = 1) \simeq 0.87$ (uniform case)

Conclusion and perspectives

Using the biased measure we could:

- Increase the clustering threshold at small k, improvement of the performances of Simulated Annealing
- At large k, improve on the third term of the asymptotic expansion: $\gamma_d \simeq 0.98$ (compare to uniform case $\gamma_d \simeq 0.87$)

Perspectives:

- more generic biases, larger range of interactions ?
- not only information on the number of forced clauses ?
- Is it possible to improve on the more dominant terms in the asymptotic expansion at large k ?

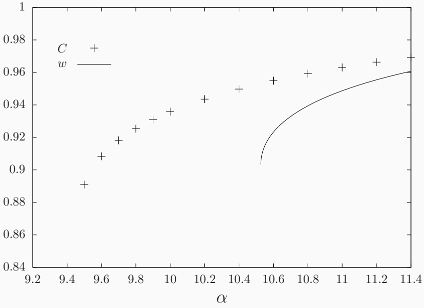
Thank you !

Asymptotics of the clustering transition

Order parameter: point-to-set correlation function C:

- Draw a solution of the CSP according to the measure μ
- Observe the spins at large distance from a root vertex
- C quantifies the amount of information on the value of the root (reconstruction threshold: $C > 0$ for $\alpha > \alpha_d$)
- Simpler lower-bound: $C \geq w$, with w the probability to be sure of the value at the root (naive reconstruction threshold: $w > 0$ for $\alpha > \alpha_r$, rigidity transition)

Asymptotics of the clustering transition



Assumption: $C(\gamma) \simeq 1 - \frac{\widehat{C}(\gamma)}{k \ln k}$ for $\gamma \geq \gamma_d$

$C > 0$ for $\alpha > \alpha_d$

$w > 0$ for $\alpha > \alpha_r$

Scaling of w, α_r: [Semerjian, 08]

- $\alpha_r = \frac{2^{k-1}}{k}(\ln k + \ln \ln k + 1 + o(1)$
- $w(\gamma) \simeq 1 - \frac{\widehat{w}(\gamma)}{k \ln k}$ for $\gamma \geq \gamma_r$

Asymptotics of the clustering transition

- Assumption: $C(\gamma) \simeq 1 - \frac{\widehat{C}(\gamma)}{k \ln k}$ for $\gamma \geq \gamma_d$
- Rescaled order parameter $\widehat{C}(\gamma)$ (diverges below γ_d)
- For $\gamma \in (\gamma_d, \gamma_r)$ one has $C \simeq 1$: quasi-hard fields that does not contribute to $\widehat{C}(\gamma)$. Need a reweighting of the soft field distribution
- For the biased measures with the appropriate scaling: obtain similar scaling for the overlap $\widehat{C}(\gamma, \widetilde{\epsilon})$ and $\widehat{C}(\gamma, b)$