Diluted spin glass models

Amin Coja-Oghlan
TU Dortmund

Overview

- discrete probability measures
- extremality and pinning
- the cut metric and pure state decompositions
- limits and compactness

Discrete probability measures

Notation

- let $\Omega \neq \emptyset$ be a finite set
- $\mathscr{P}(\Omega)$ comprises all probability distributions on Ω
- recall the total variation distance:

$$d_{\mathrm{TV}}(\mu, \nu) = \frac{1}{2} \sum_{\omega \in \Omega} |\mu(\omega) - \nu(\omega)| \qquad (\mu, \nu \in \mathscr{P}(\Omega))$$

Discrete probability measures

Probability measures on cubes

- consider a distribution $\mu \in \mathscr{P}(\Omega^n)$
- μ_i is the marginal of the i-th entry:

$$\mu_i(s) = \mu(\{\boldsymbol{\sigma}_i = s\}) \qquad\qquad (s \in \Omega)$$

- more generally, for a set $I \subset [n]$ let

$$\mu_I(\boldsymbol{s}) = \mu(\{\forall i \in I : \boldsymbol{\sigma}_i = s_i\}) \qquad\qquad (\boldsymbol{s} \in \Omega^I)$$

Discrete probability measures

Example 1

- $\Omega = \{\pm 1\}$ and $n = 3$ with

$$\mu(\boldsymbol{\sigma}) = \frac{\mathbf{1}\{\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2\boldsymbol{\sigma}_3 = 1\}}{4} \qquad (\boldsymbol{\sigma} \in \Omega^n)$$

- $\mu_1(\pm 1) = \mu_2(\pm 1) = \mu_3(\pm 1) = 1/2$
- $\mu_{i,j}(\pm 1, \pm 1) = 1/4$ for all $1 \leq i < j \leq 3$
- however, $\boldsymbol{\sigma}_1, \boldsymbol{\sigma}_2, \boldsymbol{\sigma}_3$ are not independent

Discrete probability measures

Example 2

- $\Omega = \{\pm 1\}^n$ for $n \gg 1$
- $\mu = \mu_{K_n,\beta}$ =Curie-Weiss Boltzmann distribution with $\beta = b/n$
- if $b < 1$ then for any set I of size $|I| = O(1)$

$$\mu_I(\sigma) = \prod_{i \in I} \mu_i(\sigma_i) = 2^{-|I|} \qquad (\sigma \in \{\pm 1\}^I)$$

- if $b > 1$ then

$$\mu_I(\sigma) \not\sim \prod_{i \in I} \mu_i(\sigma_i) = 2^{-|I|} \qquad (\sigma \in \{\pm 1\}^I)$$

- however,

$$\mu_I\left(\sigma \mid \sum_{i \le n} \boldsymbol{\sigma}_i \ge 0\right) \sim \prod_{i \in I} \mu_i\left(\sigma_i \mid \sum_{i \le n} \boldsymbol{\sigma}_i \ge 0\right)$$

Extremality and pinning

Extremality [BCO16]

- A distribution $\mu \in \mathscr{P}(\Omega^n)$ is *ε-extremal* if

$$\frac{1}{n^2}\sum_{i,j=1}^{n}\left|\mu_{i,j}(s,t)-\mu_i(s)\mu_j(t)\right| \le \varepsilon \qquad (s,t\in\Omega)$$

- *not the similarity with (static) "replica symmetry" [KMRTSZ07]*

Extremality and pinning

Extremality [BCO16]

- A distribution $\mu \in \mathscr{P}(\Omega^n)$ is (ε,ℓ)-*extremal* if

$$\sum_{i_1,\ldots,i_\ell=1}^{n} \left| \mu_{i_1,\ldots,i_\ell}(s_1,\ldots,s_\ell) - \prod_{j=1}^{\ell} \mu_{i_j}(s_j) \right| \le \varepsilon n^\ell \qquad (s_i \in \Omega)$$

Extremality and pinning

Lemma [BCO16]

For any Ω, $\varepsilon > 0, \ell > 2$ there exist $\delta > 0$, $n_0 > 0$ such that any δ-extremal $\mu \in \mathscr{P}(\Omega^n)$ is (ε, ℓ)-extremal.

"approximate pairwise independence implies approximate ℓ-wise independence"

Extremality and pinning

The pinning operation [T08,M08,RT12,CKPZ17]

- consider $\mu \in \mathscr{P}(\Omega^n)$
- for a set $I \subset [n]$ and $\chi \in \Omega^n$ let

$$\mathscr{S}^{I,\chi} = \left\{\sigma \in \Omega^n : \forall i \in I : \sigma_i = \chi_i\right\}$$

- define

$$\mu^{I,\chi}(\sigma) = \mu\left(\sigma \mid \mathscr{S}^{I,\chi}\right)$$

- *pin the coordinates I according to* σ

Extremality and pinning

Randomised pinning [T08,M08,RT12,CKPZ17]

- let $1 \leq \Theta \leq n$ and choose $1 \leq \ell \leq \Theta$ randomly
- consider a random set $I \subset [n]$ of size ℓ
- draw χ from μ
- then define

$$\hat{\mu}(\cdot) = \mu^{[\Theta]}(\cdot) = \mu^{I,\chi}(\cdot)$$

- observe that $\hat{\mu}(\cdot)$ is *random*!
- *pin a random set of coordinates according to a sample from* μ

Extremality and pinning

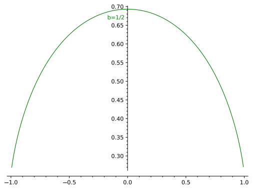
Theorem [T08,M08,RT12,CKPZ17]

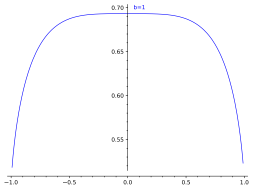
For any $\Omega, \varepsilon > 0$ there exist $\Theta > 0$, $n_0 > \Theta$ such that for all $n > n_0$ and all $\mu \in \mathscr{P}(\Omega^n)$,

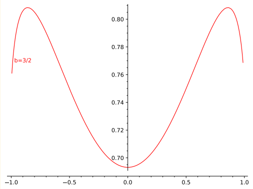
$$\mathrm{P}\left[\hat{\mu} \text{ is } \varepsilon\text{-extremal}\right] > 1 - \varepsilon.$$

crucially, Θ *depends on* Ω *and* ε *only, but not on* μ *or even* n

Extremality and pinning







Example: the Curie-Weiss model with $b > 1$

- fix a small $\varepsilon > 1$ and let $n \gg 1$
- $\mu = \mu_{K_n,\beta}$ fails to be ε-extremal
- however, $\mu^{[\ell]}$ is ε-extremal for large enough ℓ

Extremality and pinning

Proof [RT12]

- the conditional mutual information

$$I(X,Y \mid Z) = \sum_{x,y,z\in\mathcal{X}} \mathrm{P}\left[X=x,\ Y=y,\ Z=z\right] \cdot \log \frac{\mathrm{P}\left[X=x,\ Y=y \mid Z=z\right]}{\mathrm{P}\left[X=x \mid Z=z\right]\mathrm{P}\left[Y=y \mid Z=z\right]}$$

- the conditional entropy

$$H(X \mid Y) = \sum_{x,y\in\mathcal{X}} \mathrm{P}\left[X=x,\ Y=y\right] \log \mathrm{P}\left[X=x \mid Y=y\right].$$

- they satisfy

$$I(X,Y \mid Z) = H(X \mid Z) - H(X \mid Y,Z).$$

Extremality and pinning

Proof [RT12]

- let $\boldsymbol{i}, \boldsymbol{i}', \boldsymbol{i}_1, \boldsymbol{i}_2, \ldots \in [n]$ be independent and uniform
- for every $\theta \geq 0$,

$$
\begin{aligned}
&I(\boldsymbol{\sigma}_{\boldsymbol{i}}, \boldsymbol{\sigma}_{\boldsymbol{i}'} \mid \boldsymbol{\sigma}_{\boldsymbol{i}_1}, \ldots, \boldsymbol{\sigma}_{\boldsymbol{i}_\theta}) \\
&\quad = H(\boldsymbol{\sigma}_{\boldsymbol{i}} \mid \boldsymbol{\sigma}_{\boldsymbol{i}_1}, \ldots, \boldsymbol{\sigma}_{\boldsymbol{i}_\theta}) - H(\boldsymbol{\sigma}_{\boldsymbol{i}} \mid \boldsymbol{\sigma}_{\boldsymbol{i}_1}, \ldots, \boldsymbol{\sigma}_{\boldsymbol{i}_\theta}, \boldsymbol{\sigma}_{\boldsymbol{i}'}) \\
&\quad \stackrel{\mathrm{d}}{=} H(\boldsymbol{\sigma}_{\boldsymbol{i}} \mid \boldsymbol{\sigma}_{\boldsymbol{i}_1}, \ldots, \boldsymbol{\sigma}_{\boldsymbol{i}_\theta}) - H(\boldsymbol{\sigma}_{\boldsymbol{i}} \mid \boldsymbol{\sigma}_{\boldsymbol{i}_1}, \ldots, \boldsymbol{\sigma}_{\boldsymbol{i}_\theta}, \boldsymbol{\sigma}_{\boldsymbol{i}_{\theta+1}}).
\end{aligned}
$$

- summing on $\theta = 1, \ldots, T$ (with $\mathrm{E}[\cdot]$ referring to the choice of $\boldsymbol{i}, \boldsymbol{i}', \ldots$), we thus obtain

$$
\sum_{\theta=0}^{T} \mathrm{E}\left[I(\boldsymbol{\sigma}_{\boldsymbol{i}}, \boldsymbol{\sigma}_{\boldsymbol{i}'} \mid \boldsymbol{\sigma}_{\boldsymbol{i}_1}, \ldots, \boldsymbol{\sigma}_{\boldsymbol{i}_\theta})\right] = \mathrm{E}\left[H(\boldsymbol{\sigma}_{i})\right] - \mathrm{E}\left[H(\boldsymbol{\sigma}_{i} \mid \boldsymbol{\sigma}_{\boldsymbol{i}_1}, \ldots, \boldsymbol{\sigma}_{\boldsymbol{i}_{T+1}})\right]
$$

Extremality and pinning

Proof [RT12]

- since $H(\boldsymbol{\sigma}_i) \leq \log|\Omega|$ and $H(\boldsymbol{\sigma}_i \mid \boldsymbol{\sigma}_{\boldsymbol{i}_1}, \ldots, \boldsymbol{\sigma}_{\boldsymbol{i}_{T+1}}) \geq 0$, we obtain

$$\sum_{\theta=0}^{T} \mathrm{E}\left[I(\boldsymbol{\sigma}_{\boldsymbol{i}}, \boldsymbol{\sigma}_{\boldsymbol{i}'} \mid \boldsymbol{\sigma}_{\boldsymbol{i}_1}, \ldots, \boldsymbol{\sigma}_{\boldsymbol{i}_\theta})\right] \leq \log|\Omega|$$

- recalling the definition of the mutual information, we conclude that

$$\mathrm{E}\left[D_{\mathrm{KL}}\left(\hat{\mu}_{\boldsymbol{i},\boldsymbol{i}'} \| \hat{\mu}_{\boldsymbol{i}} \otimes \hat{\mu}_{\boldsymbol{i}'}\right)\right] \leq \frac{\log|\Omega|}{T}$$

Extremality and pinning

Proof [RT12]

- finally, let us recall Pinsker's inequality:

$$d_{\mathrm{TV}}(\mu, \nu) \le \sqrt{D_{\mathrm{KL}}\left(\mu \| \nu\right)/2}.$$

- applying Pinsker's inequality and Jensen's inequality, we see

$$\begin{aligned}
\mathrm{E}\left[\left\| \hat{\mu}_{\boldsymbol{i},\boldsymbol{i}'} - \hat{\mu}_{\boldsymbol{i}} \otimes \mu_{\boldsymbol{i}'} \right\|_{\mathrm{TV}}\right] &\le \mathrm{E}\left[\sqrt{D_{\mathrm{KL}}\left(\hat{\mu}_{\boldsymbol{i},\boldsymbol{i}'} \| \hat{\mu}_{\boldsymbol{i}} \otimes \hat{\mu}_{\boldsymbol{i}'}\right)/2}\right] \\
&\le \sqrt{\mathrm{E}\left[D_{\mathrm{KL}}\left(\hat{\mu}_{\boldsymbol{i},\boldsymbol{i}'} \| \hat{\mu}_{\boldsymbol{i}} \otimes \hat{\mu}_{\boldsymbol{i}'}\right)\right]/2} \\
&\le \sqrt{\frac{\log|\Omega|}{2T}}
\end{aligned}$$

thereby completing the proof

The cut metric and pure states

Couplings

- suppose that $\mu, \nu \in \mathscr{P}(\Omega)$
- a *coupling* of μ, ν is a probability γ on $\Omega \times \Omega$ such that

$$\sum_{y \in \Omega} \gamma(x, y) = \mu(x) \qquad (x \in \Omega),$$
$$\sum_{y \in \Omega} \gamma(y, x) = \nu(x) \qquad (x \in \Omega).$$

- let $\Gamma(\mu, \nu)$ comprise all couplings of μ, ν
- *Example:* the product measure $\mu \otimes \nu$

The cut metric and pure states

Coupling lemma

For any $\mu, \nu \in \mathscr{P}(\Omega)$ we have

$$\max\left\{\sum_{x \in \Omega} \gamma(x,x) : \gamma \in \Gamma(\mu,\nu)\right\} + d_{\mathrm{TV}}(\mu,\nu) = 1.$$

The cut metric and pure states

The cut metric [FK99,BCO16,COHK21]

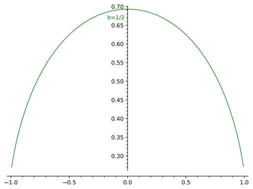
For two probability distributions $\mu, \nu \in \mathscr{P}(\Omega^n)$ define

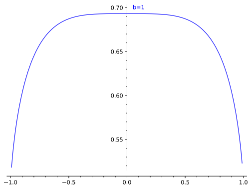
$$d_\square(\mu,\nu) = \frac{1}{n} \min_{\gamma \in \Gamma(\mu,\nu)} \max_{\substack{I \subset \{1,\ldots,n\} \\ B \subset \Omega^n \times \Omega^n \\ \omega \in \Omega}} \left| \sum_{i \in I} \sum_{(\sigma,\tau) \in B} \gamma(\sigma,\tau) (\mathbf{1}\{\sigma_i = \omega\} - \mathbf{1}\{\tau_i = \omega\})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

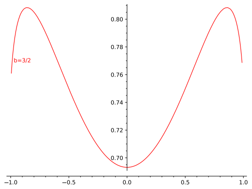
Explanation

- first couple μ, ν as best as possible $\rightsquigarrow \gamma$
- then identify the largest discrepancy $\rightsquigarrow I, B, \omega$
- (the cut metric satisfies the triangle inequality)

The cut metric and pure states





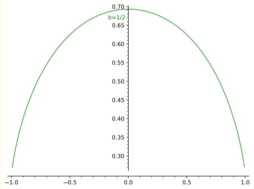


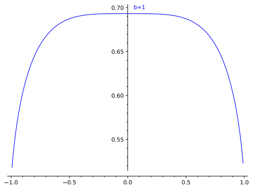
Example: Curie–Weiss

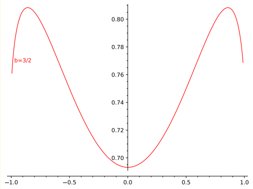
- let ν be the uniform distribution on $\{\pm 1\}$
- then for any $b < 1$,

$$\lim_{n\to\infty} d_{\square}(\mu_{K_n,\beta}(\cdot), \nu^{\otimes n}) = 0$$

The cut metric and pure states







Example: Curie–Weiss

- suppose $b > 1$
- let $\lambda_\pm$ be the positive/negative maximiser
- let $\nu_\pm \in \mathscr{P}(\{\pm 1\})$ have mean $\lambda_\pm$

$$\lim_{n\to\infty} d_\square\left(\mu_{K_n,\beta}(\cdot), \frac{1}{2}\left(\nu_+^{\otimes n} + \nu_-^{\otimes n}\right)\right) = 0$$

The cut metric and pure states

Product measures

- suppose that $\mu \in \mathscr{P}(\Omega^n)$
- then we let

$$\bar{\mu} = \bigotimes_{i=1}^{n} \mu_i$$

be the product measure with the same marginals as μ

The cut metric and pure states

Lemma

For any Ω, $\varepsilon > 0$ there exist $\delta > 0$, $n_0 > 0$ such that for all $n > n_0$ and all $\mu \in \mathscr{P}(\Omega^n)$:

- if μ is δ-extremal, then $d_{\square}(\mu, \bar{\mu}) < \varepsilon$
- if $d_{\square}(\mu, \bar{\mu}) < \delta$, then μ is ε-extremal

The cut metric and pure states

Decomposition theorem [BCO16,COHK21]

For any Ω, $\varepsilon > 0$ there exist $n_0 > 0$, $\Theta > 0$ such that for a random $1 \leq \boldsymbol{\theta} < \Theta$ for all $n > n_0$ and $\mu \in \mathscr{P}(\Omega^n)$:

- let $I \subset [n]$ be a random subset of size $\boldsymbol{\theta}$
- let
$$\bar{\mu}^I = \sum_{\chi \in \Omega^I} \mu_I(\chi) \bar{\mu}^{I,\chi}$$
- then
$$\mathrm{E}[d_{\square}(\mu, \bar{\mu}^I)] < \varepsilon$$

The cut metric and pure states

Explanation

- $\boldsymbol{\theta}$ is bounded independently of n, μ
- actually $\Theta \leq (\varepsilon^{-1} \log|\Omega|)^c$
- proof: pinning lemma and triangle inequality
- related to the "Szemerédi regularity lemma"
- *"any distribution can be approximated by a mixture of a small number of product measures"*

The cut metric and pure states

Pure states

- we obtain a decomposition of the phase space Ω^n into

$$\mathscr{S}^{I,\chi}\{\sigma \in \Omega^n : \forall i \in I : \sigma_i = \chi_i\} \qquad\qquad (\chi \in \Omega^I)$$

- for all but an ε-measure of χ, the conditional

$$\mu(\cdot \mid \mathscr{S}^{I,\chi})$$

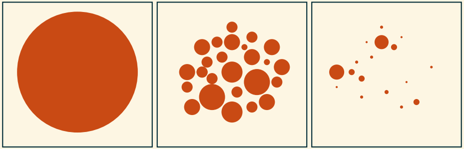
  is ε-extremal
- think of the $\mathscr{S}^{I,\chi}$ as “pure states”

The cut metric and pure states

Example: Curie–Weiss

- suppose $b > 1$
- actually the model has two pure states
- the decomposition theorem renders a partition that is "too fine"
- ...but not by "too much"
- ...and under the cut metric, the sub-states amalgamate

The cut metric and pure states



Example: sparse random factor graphs

- in the RS/dRSB phase we obtain an approximation by a single product measure $\rightsquigarrow$ one pure state
- in the static RSB phase, we obtain several pure states
- their number diverges (slowly) as $\varepsilon \to 0$ and $n \to \infty$

Limits and Aldous–Hoover

Limit objects

- let's stick to $\Omega = \{\pm 1\}$
- let $\mathscr{S}$ contain all measurable $s : [0,1] \to [-1,1]$
- equip $\mathscr{S}$ with the L^1 metric:

$$d_1(s,t) = \int_0^1 |s(x) - t(x)| \mathrm{d}x$$

- let $\mathbb{S}$ contain all measurable bijections $[0,1] \to [0,1]$

Limits and Aldous–Hoover

Limit objects

- for probability measures μ, ν on $\mathscr{S}$ define

$$d_{\square}(\mu,\nu) = \min_{\substack{\gamma\in\Gamma(\mu,\nu)\\ \varphi\in\mathbb{S}}} \max_{\substack{I\subset[0,1]\\ B\subset\mathscr{S}\times\mathscr{S}}} \left|\int_I \int_B s(x) - t(x)\mathrm{d}\gamma(s,t)\mathrm{d}x\right|$$

- obtain a space $L_{\square}(\{\pm 1\})$ by identifying μ, ν with $d_{\square}(\mu,\nu) = 0$

Limits and Aldous–Hoover

Proposition

Endowed with $d_\square(\cdot,\cdot)$, the space $L_\square(\{\pm 1\})$ is compact and separable.

Limits and Aldous–Hoover

Embedding discrete mesures

- any configuration $\sigma \in \{\pm 1\}^n$ yields a step function

$$s : [0,1] \to [-1,1], \qquad x \mapsto s(\lceil nx \rceil)$$

- this turns a distribution $\mu \in \mathscr{P}(\{\pm 1\}^n)$ into an element of $L_\square(\{\pm 1\})$

Limits and Aldous–Hoover

Random probability measures

- the Boltzmann distribution $\mu_{\mathbb{G},\beta}$ of, say, an Ising model on the random regular graph $\mathbb{G} = \mathbb{G}(n,d)$ is random itself
- hence, $\mu_{\mathbb{G},\beta}$ induces a distribution on $L_\square(\{\pm 1\})$
- thus, disordered systems map to the space

$$\mathscr{P}(L_\square(\{\pm 1\}))$$

- this is a compact separable space

Limits and Aldous–Hoover

Aldous–Hoover [P13]

- this embedding is equivalent to the Aldous–Hoover representation of exchangeable arrays
- specifically, given $p \in \mathscr{P}(L_\square(\{\pm 1\}))$, we ultimately represent the Boltzmann disordered as follows:
  - choose a random measure $\mu \in L_\square(\{\pm 1\})$ from p
  - choose $s \in \mathscr{S}$ from μ
  - draw $x \in [0,1]$ randomly
  - draw $\sigma \in \{\pm 1\}$ from a Rademacher with mean $s(x)$

Limits and Aldous–Hoover

Example: Curie–Weiss

- the Boltzmann distributions $\mu_{K_n,\beta}$ coverges to a limit $\mu^* \in L_\square(\{\pm 1\})$
- in the case $b \leq 1$ we obtain

$$\mu^* = \delta_0$$

- in the case $b > 1$ we obtain

$$\mu^* = \frac{1}{2}\left(\delta_{\lambda_-} + \delta_{\lambda_+}\right)$$

Summary

- the pinning operation
- decomposition into pure states
- embedding into a compact space
- Aldous–Hoover