Approximate Ultrametricity with Applications to Spin Glasses

Aukosh S. Jagannath[1]

[1]Courant Institute of Mathematical Sciences
New York University

July 21, 2014

Introduction

- Know ultrametricity for limiting Dovbysh-Sudakov measure (Asymptotic Gibbs Measure) [Panchenko '13]
- Gives pure state decomposition and Ruelle Probability Cascade picture

Question

Is there a sense in which one can see a 'pure state' decomposition arising at large finite N?

Approximate Ultrametricity occurs

Theorem

Let $\{\mu_N\}$ be a sequence of random probability measures supported on the unit ball of ℓ_2. Suppose that this sequence satisfies the Approximate Ghirlanda-Guerra Identities and that for some ζ,

$$\zeta_N = \mathbb{E}\mu_N^{\otimes 2}((\sigma^1, \sigma^2)_{\ell_2} \in \cdot) \to \zeta.$$

Then this sequence is **regularly approximately ultrametric with respect to** ζ.

Translation

Include $\{-1,+1\}^N \hookrightarrow \ell_2$ by

$$(\sigma_1,\ldots,\sigma_N) \overset{i}{\mapsto} (\frac{1}{\sqrt{N}}\sigma_1,\ldots,\frac{1}{\sqrt{N}}\sigma_N,0,\ldots)$$

Key:

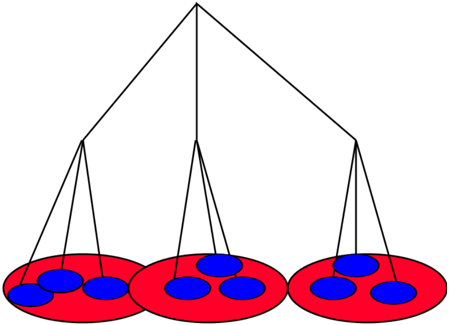
$$\begin{aligned} \text{Gibbs measure } G_N &\iff \mu_N \\ \text{overlap } R_{12} &\iff (\sigma^1,\sigma^2)_{\ell_2} \\ \text{overlap distribution} &\iff \zeta_N \\ \text{limiting overlap distirbution} &\iff \zeta \end{aligned}$$

What do we mean by Approximately Ultrametric?

- Has a "pure state"-like decomposition $\implies$ there is a collection of hierarchically arranged "balls" that almost exhaust μ
- Need to abstract properties of balls in ultrametric measure spaces
- Focus two properties:
  1. measure theoretic: Hierarchical Exhaustion
  2. geometric: Hierarchical Clustering

Hierarchical Exhaustion

- In ultrametric measure spaces balls form hierarchical partition
- In “almost” ultrametric measure spaces form hierarchical “almost” partition

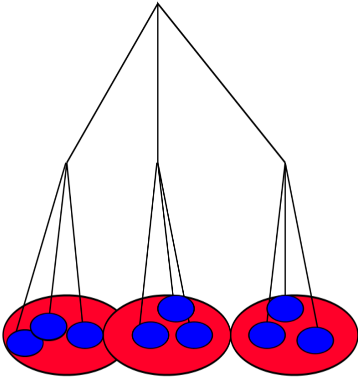


Hierarchical Exhaustion

Definition

$\{C_\alpha\}_{\alpha \in \tau_r}$ is an (ϵ, δ)-**hierarchical exhaustion** of μ if:

1. $$C_\alpha \cap C_\beta = C_\alpha \text{ if } \beta \precsim \alpha.$$
2. $$\sum_{\alpha,\beta \in \tau, \alpha \nsim \beta} \mu(C_\alpha \cap C_\beta) \leq \delta.$$
3. $$\sum_{|\alpha|=k} \mu(C_\alpha) \geq 1 - \epsilon$$
4. $$\mu(C_\alpha) - \sum_{\beta \in child(\alpha)} \mu(C_\beta) \in [0, \epsilon).$$



Hierarchical Exhaustion

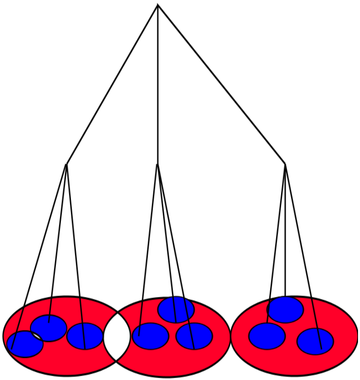
Definition

$\{C_\alpha\}_{\alpha\in\mathcal{T}_r}$ is an $(\epsilon, 0)$-**hierarchical exhaustion** of μ if:

1. $$C_\alpha \cap C_\beta = \begin{cases} C_\alpha & \text{if } \beta \precsim \alpha \\ \emptyset & \text{if } \beta \not\sim \alpha. \end{cases}$$

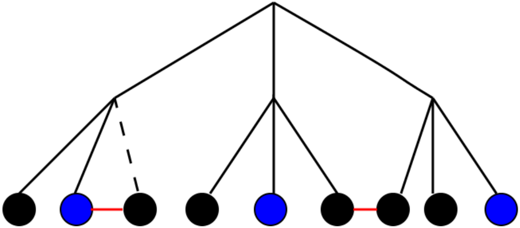
2. $$\sum_{|\alpha|=k} \mu(C_\alpha) \geq 1-\epsilon$$

3. $$\mu(C_\alpha) - \sum_{\beta\in child(\alpha)} \mu(C_\beta) \in [0,\epsilon).$$



Hierarchical Clustering

- In ultrametric measure space: points in the same R-ball of should be R-close, in different R-balls should be R-far
- In "almost" ultrametric measure space: probability of points in same R-ball being $(R+\epsilon)$-far or points in different R-balls being $(R-\epsilon)$-close is small



Hierarchical Clustering

Definition

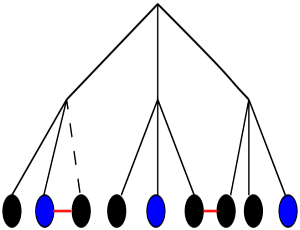
$\{C_\alpha\}_{\alpha\in\tau_r}$ is an (ϵ,δ)-**hierarchical clustering** for μ with respect to $\{q_k\}_{k=1}^r$ if:

1. same cluster $\implies$ close:

$$\mu\left(\sigma^1,\sigma^2\in C_\alpha : (\sigma^1,\sigma^2)_{\ell_2}\leq q_{|\alpha|}-\epsilon\right)\leq\delta$$

2. cousins $\implies$ uniformly far apart: $\alpha\nsim\beta$ with same parent

$$\mu\left(\sigma^1\in C_\alpha,\sigma^2\in C_\beta : (\sigma^1,\sigma^2)_{\ell_2}\geq q_{|\alpha|}+\epsilon\right)\leq\delta$$



What radii are allowed?

Definition

Let ζ be a probability measure on the interval $[0,1]$. A finite increasing sequence of points $\{q_k\}_{k=1}^r$ on the interval $(0,1)$ is said to be ζ-**admissible** if:

1. They are all continuity points:
$$\zeta(\{q_k\}) = 0.$$
2. There is mass between them:
$$\zeta[q_k, q_{k+1}] > 0.$$
3. There is some but not all of the ζ-mass lies between 0 and q_1, and q_r and 1:
$$0 < \zeta[0, q_1] < 1 \text{ and } 0 < \zeta[q_r, 1] < 1$$

Recap

Three properties:

1. A measure theoretic property: hierarchical exhaustion
2. A geometric property: hierarchical clustering
3. allowed radii

Together will make Approximate Ultrametricity.

Approximate Ultrametricity

Definition (Approximately ultrametric)

$\{\mu_N\}_{N=1}^{\infty}$ is **approximately ultrametric** with respect to ζ if for every r and every ζ-admissible sequence, $\{q_k\}_{k=1}^{r}$, there is a sequence of finite rooted trees of depth r, $\{\tau_{N,r}\}$, and sequences a_N, b_N, and ϵ_N all tending to 0 such that with probability tending to one, there are sets $\{C_{\alpha,N}\}_{\alpha\in\tau_{N,r}}$ that are $(\epsilon_N, 0)$-hierarchically exhausting and (a_N, b_N)- clustering.

It is **regularly** approximately ultrametric if there is a sequence $m_N \to \infty$ such that $\tau_{N,r}$ can be taken to be the m_N-regular tree of depth r.

Approximate Ultrametricity occurs

Theorem

Let $\{\mu_N\}$ be a sequence of random probability measures supported on the unit ball of ℓ_2. Suppose that this sequence satisfies the Approximate Ghirlanda-Guerra Identities and that

$$\zeta_N = \mathbb{E}\mu_N^{\otimes 2}((\sigma^1, \sigma^2)_{\ell_2} \in \cdot) \to \zeta$$

for some ζ. Then this sequence is regularly approximately ultrametric with respect to ζ.

Laws of cluster weights

Fix $\{q_k\}_{k=1}^r$ ζ-admissible, $C_{\alpha,N}$ corresponding clusters, set

$$\tilde{Y}_{\alpha,N} = \begin{cases} C_{\alpha,N} & \text{for } \alpha \in \tau_{N,r} \\ \emptyset & \mathcal{A}_r \setminus \tau_{N,r} \end{cases}$$

Let $(Y_{\alpha,N})_{\alpha\in\mathcal{A}_r}$ be $(\mu_N(\tilde{Y}_{\alpha,N}))_{\alpha\in\mathcal{A}_r}$ arranged in standard order.

Theorem

Let (Y_α) be distributed like the weights of a Ruelle Probability Cascade with parameters $\zeta_k = \zeta[q_k, q_{k+1}]$. Then

$$(Y_{\alpha,N}) \overset{(d)}{\to} (Y_\alpha)$$

Applications

- Provides "pure state" decomposition of hypercube at low temperature and **finite** N for:
  - Generic mixed p-spin models [Auffinger-Chen '13],
  - REM, GREM [Bovier-Kurkova '04]
- Provides new sense of convergence of REM and GREM to Ruelle Probability Cascade
- Resolves Orthogonal Structures conjecture

Orthogonal Structures

Theorem (Orthogonal Structures Conjecture)

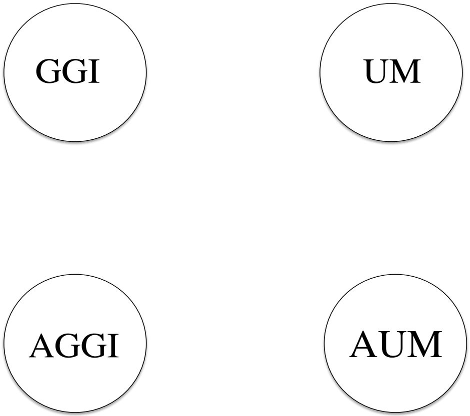
For a generic mixed p-spin glass model, if $\zeta(\{0\}) > 0$ then the Gibbs measures G_N admit an **Orthogonal Structure**.

Definition

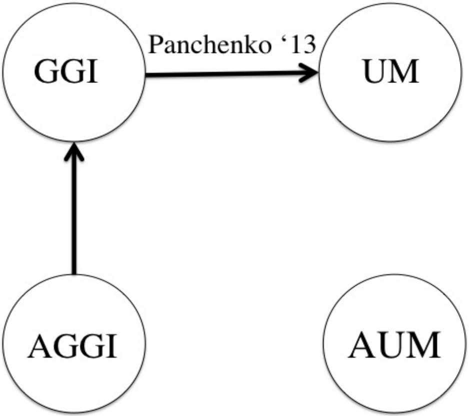
A spin glass model on the hypercube admits an **Orthogonal Structure** if there is a sequence $(a_k)_{k\geq 0}$ such that for any $k_0 \in \mathbb{N}$ and ϵ positive, there is an N_0 such that for $N \geq N_0$, with probability at least $3/4$, there is a random collection of sets $\{A_{k,N}\}_{k\leq k_0} \subset \Sigma_N$ such that

- $G_N(A_{k,N}) \geq a_k$
- $\forall k \neq l : \left\langle |R_{12}| \mathbb{1}_{\sigma^1 \in A_{k,N}, \sigma^2 \in A_{l,N}} \right\rangle < \epsilon$

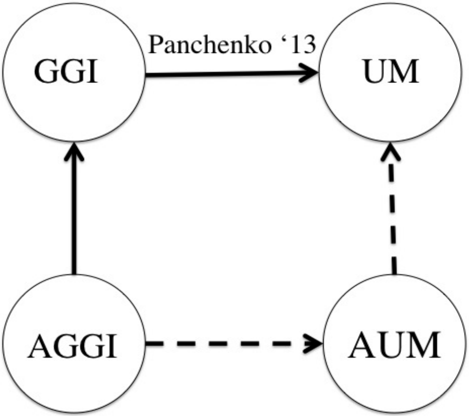
Relation to what is known



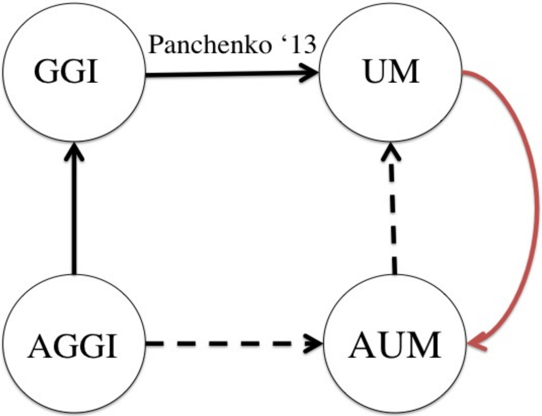
Relation to what is known



Relation to what is known



Relation to what is known



Consequences of Ghirlanda-Guerra Identities

Key consequences of Ghirlanda-Guerra Identities we use:

1. Determines structure of limiting Dovbysh-Sudakov measure:
    - Support is ultrametric
    - Weights of pure states are "Ruelle Probability Cascade"-like
2. Convergence of $\zeta_N \implies$ convergence of overlap array.

Strategy for Theorem 1

- Picture should arise for heaviest balls
- Potential difficulties with looking at heaviest balls
- Instead work with randomly chosen balls
- Low temperature $\implies$ will not sample "dust"
- Sample enough balls to get all information needed

Ball sampling procedure

- Draw $(\sigma^\alpha)_{\alpha \in \mathcal{A}_r}$ iid from μ_N
- Pick $\{B_\alpha\}_{\alpha \in \mathcal{A}_r}$ by

$$B_\alpha(\boldsymbol{\sigma}) = \bigcap_{\beta \prec \alpha} B(\sigma^\beta, q_{|\beta|}).$$

- look at

$$\mathbf{V}^N = (V_\alpha^N) = (\mu_N B_\alpha)$$

Lemma

Let $\mathbf{V}$ *be generated the same way from* μ*, the limiting Dovbysh-Sudakov Measure*

$$\mathbf{V}^N \overset{(d)}{\to} \mathbf{V}$$

Hierarchical Exhaustion

Lemma

For every ϵ positive, r and ζ-admissible sequence $\{q_k\}_{k=1}^r$,

$$\lim_{N\to\infty} \mathbb{P}(\exists \tau_{N,r} : \exists(\epsilon, 0) - \textit{exhaustion of } \mu_N) = 1.$$

Proof.

- μ is ultrametric $\implies$ μ has $(\epsilon, 0)$-hierarchical exhaustions for some τ.
- Weak convergence to conclude τ works for large *but finite N*
- Key Point: $\exists \tau$ for $\mathbf{V}^N$ is an open condition.

□

Hierarchical Clustering and Approximate Ultrametricity

Theorem

For $\{\mu_N\}$ as above is regularly approximately ultrametric w.r.t. ζ

Proof.

- Show (a, b)-hierarchical clustering
  - Encode distance control in functions of replica f and g
  - Use ultrametricity of limit to show f and g are uniformly small
- Choose errors in exhaustion and clustering to vanish in N and degree $m_N \to \infty$

□

Convergence of Leaves to Poisson-Dirichlet

Let (v_n^N) be the weights of the leaves rearranged in decreasing order.

Theorem

The weights (v_n^N) converge in distribution to $PD(\zeta[q_r, 1])$ as elements of the space of mass-partitions.

- Talagrand's identities for PD: Recursion relation for moments via Ghirlanda-Guerra

Convergence of weights to Ruelle Probability Cascade

- Hierarchical analogue of space of mass-partitions: space of cascades of depth r

$$\mathcal{C}_r = \{(w_\alpha)_{\alpha\in\mathcal{A}_r} \in [0,1]^{\mathcal{A}_r} : w_{\alpha 1} \geq w_{\alpha 2} \geq \ldots \geq 0; \sum_{|\alpha|=k} w_\alpha \leq 1, \forall k \in [r]; w_\alpha \geq \sum_{\beta\in child(\alpha)} w_\beta\}$$

- Problem: obvious analogues of Talagrand's Identities does not seem to come from GGIs
- Work around: **ROSt encoding**

ROSt Encoding argument

$$\mathbf{v}^N \overset{?}{\dashrightarrow} \mathbf{v}$$

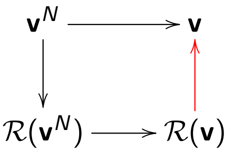
ROSt Encoding argument

- Encode $\mathbf{v}_N$ in ROSt
- convergence of $\mathbf{v}^N \iff$ convergence of ROSt

$$\begin{array}{ccc} \mathbf{v}^N & \overset{?}{\dashrightarrow} & \mathbf{v} \\ \downarrow & & \downarrow \\ \mathcal{R}(\mathbf{v}^N) & \longrightarrow & \mathcal{R}(\mathbf{v}) \end{array}$$

ROSt Encoding argument

- Encode $\mathbf{v}_N$ in ROSt
- convergence of $\mathbf{v}^N \iff$ convergence of ROSt
- Uniqueness of Dovbysh-Sudakov measure gives correspondence $\mathbf{v} \to$ROSt



Correspondence is not continuous

Convergence of ROSt

Theorem

The sequence $\mathbf{v}^N$ converges to the weights of a Ruelle Probability Cascade.

Proof.

- Key Observation: $\mathbf{v}^{N_k} \stackrel{(d)}{\to} \mathbf{v}$ gives $\mathbb{E}\left\langle f(R^n)\right\rangle_{\mathcal{R}(\mathbf{v}^{N_k})} \to \mathbb{E}\left\langle f(R^n)\right\rangle_{\mathcal{R}(\mathbf{v})}$
- Careful application of convergence of leaves circumvents issues caused by compactness
- $\mathcal{R}(\mathbf{v}^N) \stackrel{(d)}{\to} RPC$
- Uniqueness of Dovbysh-Sudakov + Subsequence principle

□

Recap

- Introduced approximate ultrametricity to capture existence of "pure state"-like decomposition
- Showed approximate ultrametricity occurs in large class of models
- Characterized the laws of the weights of the "pure states"

Thanks

Thanks for listening!

Thanks to: D. Panchenko, A. Auffinger, G. Ben Arous.

This material is based on work supported by the NSF Graduate Research Fellowship DGE-0813964 and NSF Grant DMS 1209165.

Hierarchical Exhaustion

Lemma

For every ϵ positive, r and ζ-admissible sequence $\{q_k\}_{k=1}^r$,

$$\lim_{N\to\infty} \mathbb{P}(\exists \tau_{N,r} : \exists(\epsilon, 0) - \textit{exhaustion of } \mu_N) = 1.$$

Proof.

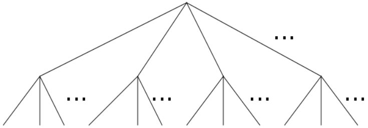
- μ is UM $\implies$ μ has hierarchical exhaustions with $\delta = 0$ for some τ regular.
- Property test: do not work with largest, just find collection that has our property
- Prune $\mathcal{A}_r$ into shape of τ: $(\sigma^\alpha)_{\alpha\in\mathcal{A}_r} \mapsto (\sigma^{\tau^i})_{i=1}^N$
- $(\mu B_\alpha)_{\alpha\in\tau}$ has correct properties is an open condition $\implies$ $\exists\tau$ is an open condition
- Use LLN and Weak convergence to conclude τ works for large *but finite* N w.h.p.
- Clean up B_α's: remove intersection with cousins. These are the clusters.

□

Lemma

For every ϵ positive, r and ζ-admissible sequence $\{q_k\}_{k=1}^r$,

$$\lim_{N\to\infty} \mathbb{P}(\exists \tau_{N,r} : \exists(\epsilon, 0) - \textit{exhaustion of } \mu_N) = 1.$$

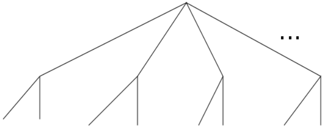


Hierarchical Exhaustion

Lemma

For every ϵ positive, r and ζ-admissible sequence $\{q_k\}_{k=1}^r$,

$$\lim_{N\to\infty} \mathbb{P}(\exists \tau_{N,r} : \exists(\epsilon, 0) - \textit{exhaustion of } \mu_N) = 1.$$



Hierarchical Exhaustion

Lemma

For every ϵ positive, r and ζ-admissible sequence $\{q_k\}_{k=1}^r$,

$$\lim_{N\to\infty} \mathbb{P}(\exists \tau_{N,r} : \exists(\epsilon, 0) - \textit{exhaustion of } \mu_N) = 1.$$

Proof.

- μ is UM $\implies$ μ has hierarchical exhaustions with $\delta = 0$ for some τ regular.
- Property test: do not work with largest, just find collection that has our property
- Prune $\mathcal{A}_r$ into shape of τ: $(\sigma^\alpha)_{\alpha\in\mathcal{A}_r} \mapsto (\sigma^{\tau^i})_{i=1}^N$
- $(\mu B_\alpha)_{\alpha\in\tau}$ has correct properties is an open condition $\implies$ $\exists\tau$ is an open condition
- Use LLN and Weak convergence to conclude τ works for large *but finite* N w.h.p.
- Clean up B_α's: remove intersection with cousins. These are the clusters.

□

Hierarchical Clustering and Approximate Ultrametricity

Theorem

For $\{\mu_N\}$ as above is approximately ultrametric w.r.t. ζ

Proof.

- Given errors, find τ as above m-regular and before time M
- Show hierarchical clustering
  - Encode uniform distance control:

$$f^N_{|\alpha|,\epsilon}(\sigma^\alpha) = \mu_N^{\otimes 2}\left(\sigma^1,\sigma^2 \in C_{\alpha,N} : R_{12} \leq q_{|\alpha|} - \epsilon\right)$$
$$g^N_{\alpha,\beta,\epsilon}(\boldsymbol{\sigma}) = \mu_N^{\otimes 2}\left(\sigma^1 \in C_{\alpha,N}, \sigma^2 \in C_{\beta,N} : R_{12} \geq q_{|\alpha\wedge\beta|+1} + \epsilon\right)$$

  - show f and g are uniformly small w.h.p.
- Chose all of the errors in the above to vanish in N and $m_N \to \infty$
- Play with indices to make everything happen simultaneously

□

ROSt Encoding

- Fix $\mathbf{v} \in \mathcal{C}_r$ and $\{q_k\}_{k=1}^r$
- Place dust in dustbin ∂:

$$v_\partial = 1 - \sum_{|\alpha|=r} v_\alpha.$$

- As with Ruelle Probability Cascades

$$h_\alpha = \begin{cases} \sum_{\beta \precsim \alpha} (q_{|\alpha|} - q_{|\alpha|-1})^{\frac{1}{2}} e_\beta & \alpha \in \partial \mathcal{A}_r \\ h_\partial = q_r^{\frac{1}{2}} e_\partial & \alpha = \partial \end{cases}$$

- ROSt encoding

$$\mathcal{R}(\mathbf{v}, \{q_k\}; \{h_\alpha\}) = v_\alpha$$

Convergence of ROSt

Lemma

Fix $\{q_k\}$. Let $\mathbf{v}$ be a limit point of the sequence $\mathbf{v}^N$ augmented by v_∂ and let $\mathbf{v}^{N_k} \to \mathbf{v}$. It follows that

$$\mathcal{R}(\mathbf{v}_{N_k}) \overset{(samp)}{\to} \mathcal{R}(\mathbf{v})$$

Proof.

- use PD convergence to conclude the dustbins converge
- use Skorokhod Representation +Scheffé's lemma to conclude $\mathbb{E}\langle f\rangle_{\mathcal{R}(\mathbf{v}^N)} \to \mathbb{E}\langle f\rangle_{\mathcal{R}(\mathbf{v}))}$

□

$$\mathcal{R}(\mathbf{v}^N) \longrightarrow \mathcal{R}(\mathbf{v})$$

Convergence of ROSt

Lemma

Fix $\{q_k\}$. Let $\mathbf{v}$ be a limit point of the sequence $\mathbf{v}^N$ augmented by v_∂ and let $\mathbf{v}^{N_k} \to \mathbf{v}$. It follows that

$$\mathcal{R}(\mathbf{v}_{N_k}) \stackrel{(samp)}{\rightarrow} \mathcal{R}(\mathbf{v})$$

Lemma

The ROSt corresponding to $\mathcal{R}(\mathbf{v}^N)$ converges in law to an RPC.

Proof.

- Convergence of $\mathcal{R}$ is equivalent to convergence of R^n for each n
- R^n only takes on finitely many values $\implies$ suffices to compute all of these probabilities
- To calculate $\langle R^n = Q^n \rangle$, encode Q^n in a tree $\tau(Q)$ of depth $r+1$
- Key Idea: Draw replica indexed by the leaves of $\tau(Q)$, indicator function that they land in clusters which obey metric induced by τ is well-approximated by $R_{ij} \in [q_{k(ij)}, q_{k(ij)+1}]$

□

Convergence of ROSt

Lemma

Fix $\{q_k\}$. Let $\mathbf{v}$ be a limit point of the sequence $\mathbf{v}^N$ augmented by v_∂ and let $\mathbf{v}^{N_k} \to \mathbf{v}$. It follows that

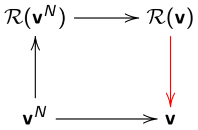
$$\mathcal{R}(\mathbf{v}_{N_k}) \overset{(samp)}{\to} \mathcal{R}(\mathbf{v})$$

Lemma

The ROSt corresponding to $\mathcal{R}(\mathbf{v}^N)$ converges in law to an RPC.

Theorem

The sequence $\mathbf{v}^N$ converges to the weights of an RPC.



Dotsenko-Franz-Mézard Conjecture

Conjecture (Dotsenko-Franz-Mézard)

Suppose Z_N is the partition function for mixed p-spin glass Hamiltonian at $h = 0$, then for all a negative

$$\lim_{N\to\infty} \frac{1}{aN} \log \mathbb{E} Z_N^a = \lim_{N\to\infty} \frac{1}{N} \mathbb{E} \log Z_N$$

Theorem (Dotsenko-Franz-Mézard for generic models)

This holds for generic mixed p-spin glass Hamiltonian such that $\zeta\{0\} > 0$, at low temperature

Proof.

- Existence of Orthogonal Structure $\implies$ Dotsenko-Franz-Mézard [Talagrand '07]

□