

# Average Gromov hyperbolicity and the Parisi ansatz

Sourav Chatterjee

(Joint work with Leila Sloman)

# Mathematical models of spin glasses

- ▶ A **spin glass** is a kind of magnetic object with properties that are quite different than ferromagnets.
- ▶ A mathematical model of a spin glass assigns a *random* probability measure  $\mu_n$  on a set  $\Sigma_n$ .
- ▶ Usually,  $\Sigma_n$  is either the hypercube  $\{-1, 1\}^n$  or the sphere of radius  $\sqrt{n}$  centered at the origin in  $\mathbb{R}^n$ . For the rest of this talk,  $\Sigma_n$  will be one of these two.
- ▶ One of the most famous examples is the **Sherrington–Kirkpatrick (SK) model**.
- ▶  $\mu_n$  is called the Gibbs measure, and the set  $\Sigma_n$  is called the configuration space.

# Replicas and their overlaps

- ▶ An important quantity in spin glass theory is the **overlap** between two configurations  $\sigma^1, \sigma^2 \in \Sigma_n$ , defined as

$$R_{1,2} := \frac{1}{n} \sum_{i=1}^n \sigma_i^1 \sigma_i^2.$$

- ▶ Note that  $R_{1,2} \in [-1, 1]$ .
- ▶ An i.i.d. sequence  $\sigma^1, \sigma^2, \dots$  of configurations drawn from the Gibbs measure  $\mu_n$  is called a sequence of **replicas**.
- ▶ The overlap between the replicas  $\sigma^i$  and  $\sigma^j$  is usually denoted by  $R_{i,j}$ .

- ▶ Parisi (1979) famously conjectured that certain spin glass models have the property that in the “ $n \rightarrow \infty$  limit”,

$$R_{1,2} \geq \min\{R_{1,3}, R_{2,3}\}.$$

- ▶ This is known as the **Parisi ultrametricity ansatz**.
- ▶ It was the basis of the Mézard–Parisi **broken replica symmetry** method for solving mean-field spin glass models.

# Panchenko's theorem

- ▶ Following a long line of deep contributions by various authors (Aizenman, Arguin, Contucci, Ghirlanda, Guerra, Talagrand, ...), the Parisi conjecture was finally proved by Panchenko (2013) for spin glass models that satisfy a certain set of equations known as the **generalized Ghirlanda–Guerra (GG) identities**.
- ▶ For such models, Panchenko's theorem says that for all  $\epsilon > 0$ ,

$$\lim_{n \rightarrow \infty} \mathbb{E} \langle \mathbb{1}_{\{R_{1,2} \geq \min\{R_{1,3}, R_{2,3}\} - \epsilon\}} \rangle = 1,$$

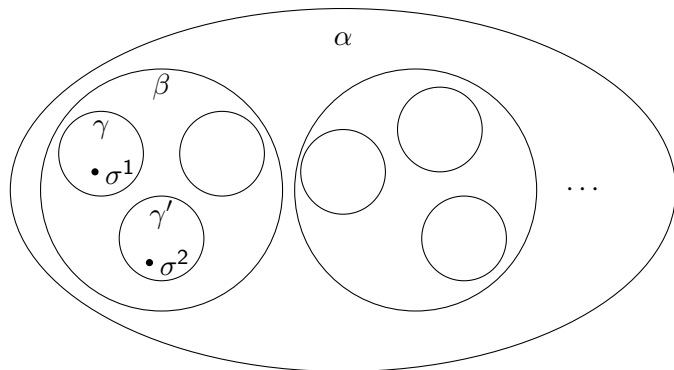
where  $\langle \cdot \rangle$  denotes expectation with respect to the three-fold product of the Gibbs measure  $\mu_n$ ,  $\mathbb{E}$  denotes expectation with respect to the randomness in  $\mu_n$ , and  $\mathbb{1}_A$  denotes the function that is 1 on the set  $A$  and 0 elsewhere.

- ▶ The generalized GG identities have been proved in a large class of models.

# Hierarchical organization of pure states

- ▶ It was conjectured (e.g. in the book of Mézard, Parisi and Virasoro) that ultrametricity happens because the infinite volume limit of the Gibbs measure can be decomposed into “hierarchically organized pure states”.
- ▶ Roughly speaking, this means that the configuration space admits a hierarchical clustering, with a number  $q_\alpha \in [-1, 1]$  attached to each cluster  $\alpha$ , so that if  $\sigma^1$  and  $\sigma^2$  are drawn independently from the Gibbs measure, then with high probability,  $R_{1,2} \approx q_\alpha$ , where  $\alpha$  is the smallest cluster containing both  $\sigma^1$  and  $\sigma^2$ .

# Illustration



**Figure:** Hierarchical organization of pure states. Here  $\alpha$ ,  $\beta$ ,  $\gamma$  and  $\gamma'$  are hierarchically nested clusters representing various pure states, and  $\sigma^1 \in \gamma$ ,  $\sigma^2 \in \gamma'$ . But  $R_{1,2} \approx q_\beta$ , since  $\beta$  is the smallest cluster that contains both  $\sigma^1$  and  $\sigma^2$ .

# From ultrametricity to hierarchically organized pure states

- ▶ It is not difficult to prove that ultrametricity implies the hierarchical organization of pure states if  $R_{1,2}$  can take only finitely many values in the infinite volume limit.
- ▶ This is the basis of the heuristic sketched in physics papers.
- ▶ However, if this condition does not hold — in which case the system is said to exhibit “full replica symmetry breaking” — then it is not clear how to construct hierarchically organized pure states even if the Parisi ansatz is known to be valid.



# Rigorous construction of pure states

- ▶ There are two kinds of systems where the pure state picture has been rigorously established.
- ▶ The first is the class of **pure  $p$ -spin spherical models**, where the pure state construction was given recently by Subag (2017), building on the earlier contributions of Auffinger, Ben Arous, Černý, Chen, Subag, Zeitouni, ...
- ▶ The second is the class of models that satisfy the generalized GG identities.
- ▶ For these models, the construction of pure states was given by Panchenko (2013) in the infinite volume limit, and recently by Jagannath (2017) in the setting of large but finite  $n$ .
- ▶ However, it is natural to expect that hierarchically organized pure states exist in any system that satisfies the ultrametricity ansatz. This will be the first main result of this talk.

# The generalized Parisi ansatz

- ▶ It is believed that the ultrametricity ansatz and the generalized GG identities are equivalent, but this has not been proved.
- ▶ Besides, there are important models such as the SK model, where it is known that the generalized GG identities **do not hold**.
- ▶ In the SK model, it is believed that  $|R_{1,2}|$ , rather than  $R_{1,2}$ , satisfies the ultrametric property.
- ▶ To account for such cases, we formulate a generalized version of the ansatz.
- ▶ We will say that a sequence of spin glass models satisfy the **generalized Parisi ansatz** if for some bounded measurable  $f : [-1, 1] \rightarrow \mathbb{R}$  and any  $\epsilon > 0$ ,

$$\lim_{n \rightarrow \infty} \mathbb{E} \langle \mathbb{1}_{\{f(R_{1,2}) \geq \min\{f(R_{1,3}), f(R_{2,3})\} - \epsilon\}} \rangle = 1.$$

# Ultrametricity implies hierarchically organized pure states

## Theorem (C. & Sloman, 2019)

*Consider any sequence of spin glass models that satisfy the generalized Parisi ultrametricity ansatz for some bounded measurable function  $f$ . Then there are sequences  $\epsilon_n$  and  $\delta_n$  tending to zero, such that with probability at least  $1 - \epsilon_n$ , the following happens. There is a hierarchical clustering of the configuration space  $\Sigma_n$ , such that the number of clusters is finite (but may depend on  $n$ ), each cluster is measurable, and for each cluster  $\alpha$  there is a number  $q_\alpha$  that is a function of its the depth in the hierarchy, with the property that*

$$\langle |f(R_{1,2}) - q_\alpha| \rangle \leq \delta_n,$$

*where  $\alpha = \alpha(\sigma^1, \sigma^2)$  is the smallest cluster containing two configurations  $\sigma^1$  and  $\sigma^2$  drawn independently from the Gibbs measure and  $R_{1,2}$  is their overlap.*

- ▶ Note that the theorem works for finite RSB as well as full RSB.
- ▶ It is a corollary of a more general result about “almost hyperbolic” metric spaces, which I will now present.
- ▶ The general result has nothing to do with spin glasses.

# Gromov hyperbolicity of a metric space

- ▶ Let  $(S, d)$  be a metric space.
- ▶ The **Gromov product** of  $x, y \in S$  with respect to  $z \in S$  is defined as  $(x, y)_z := \frac{1}{2}(d(x, z) + d(y, z) - d(x, y))$ .
- ▶  $(S, d)$  is called  $\delta$ -hyperbolic if for any  $x, y, z, w \in S$ ,

$$(x, y)_w \geq \min\{(x, z)_w, (y, z)_w\} - \delta. \quad (\star)$$

- ▶ The smallest  $\delta$  for which this is satisfied is known as the Gromov hyperbolicity of  $(S, d)$ .
- ▶ The condition  $(\star)$  is known as Gromov's four point condition.
- ▶ If  $(\star)$  is satisfied for all  $x, y, z$  for a given  $w_0$ , then it can be shown that it is satisfied for all  $w$  with  $2\delta$  in place of  $\delta$ .
- ▶ Thus, we may equivalently define hyperbolicity using a three point condition, by fixing  $w$ .
- ▶ If  $(\star)$  is satisfied for all  $x, y, z$  for some fixed  $w$ , then we say that the space is  **$\delta$ -hyperbolic with base point  $w$** .

- ▶ Gromov hyperbolicity measures how “tree-like” the space is. To make this connection precise, we need to recall the notion of a **real tree**.
- ▶ If  $(T, \rho)$  is a metric space and  $x, y \in T$ , an **arc** from  $x$  to  $y$  is the image of a topological embedding  $\gamma : [a, b] \rightarrow T$  with  $\gamma(a) = x$  and  $\gamma(b) = y$ , where  $[a, b]$  is a closed interval in  $\mathbb{R}$  (allowing the possibility that  $a = b$ ).
- ▶ A **geodesic segment** from  $x$  to  $y$  is the image of an isometric embedding  $\gamma : [a, b] \rightarrow T$  with  $\gamma(a) = x$  and  $\gamma(b) = y$ .
- ▶ A metric space  $(T, \rho)$  is called a **real tree** if for any  $x, y \in T$ , there exist a **unique** arc from  $x$  to  $y$ , and this arc is a geodesic segment.
- ▶ A real tree with a distinguished point  $r \in T$  is called a rooted real tree with root  $r$ .

# Embedding of Gromov hyperbolic spaces in real trees

- ▶ It is easy to show that a metric space  $(S, d)$  is 0-hyperbolic if and only if it is isometric to a subset of a real tree.
- ▶ Now suppose that  $(S, d)$  is  $\delta$ -hyperbolic for some small but nonzero  $\delta$ . Is it approximately isometric to a subset of a real tree in some sense?
- ▶ It is known that this is true when  $S$  has finite cardinality, with an error proportional to  $\delta \log |S|$  (which cannot be improved). The proof uses induction on the size of the space.

## Theorem (Ghys & de la Harpe, 1990)

Let  $(S, d)$  be a  $\delta$ -hyperbolic metric space with base point  $w$  and finite cardinality. Let  $k$  be a positive integer such that  $|S| \leq 2^k + 2$ . Then there exists a real tree  $(T, \rho)$  with root  $r$  and a map  $\Phi : S \rightarrow T$  such that for all  $x \in S$ ,  $d(x, w) = \rho(\Phi(x), r)$ , and for all  $x, y \in S$ ,  $d(x, y) - 2k\delta \leq \rho(\Phi(x), \Phi(y)) \leq d(x, y)$ .

# Average hyperbolicity

- ▶ The notion of Gromov hyperbolicity has found great success in many areas of mathematics and science.
- ▶ However, it has one weakness. Gromov's four point condition is a worst-case condition: The space is not  $\delta$ -hyperbolic if there is even a single four-tuple  $(x, y, z, w)$  for which it fails.
- ▶ For the construction of pure states in spin glasses, we need an **average** notion of hyperbolicity, which stipulates that the four point condition holds in **most** but not all cases, where "most" is with respect to the Gibbs measure.
- ▶ Average hyperbolicity has been considered in the literature on networks and phylogenetic trees, but its mathematical properties have not been analyzed.
- ▶ In particular, what we need is an analog of the embedding theorem.
- ▶ I will present such a result. The proof, surprisingly, involves **Szemerédi's regularity lemma** from graph theory.



# Our definition of average hyperbolicity

- ▶ We will go beyond metric spaces.
- ▶ Let  $S$  be a set equipped with a countably generated  $\sigma$ -algebra  $\mathcal{F}$  and a probability measure  $\mathbb{P}$  defined on  $\mathcal{F}$ .
- ▶ Let  $s : S \times S \rightarrow [0, 1]$  be a measurable function satisfying  $s(x, y) = s(y, x)$  for all  $x, y \in S$ . We will say that  $s$  is a **similarity function**.
- ▶ Similarity functions generalize the notion of Gromov product: If  $S$  has diameter 1 with respect to a separable metric and is endowed with the Borel  $\sigma$ -algebra generated by this metric, the Gromov product  $(x, y)_w$  is a similarity function for any base point  $w \in S$ .
- ▶ **Definition:** We will say that  $(S, \mathcal{F}, \mathbb{P}, s)$  is  $\delta$ -hyperbolic if

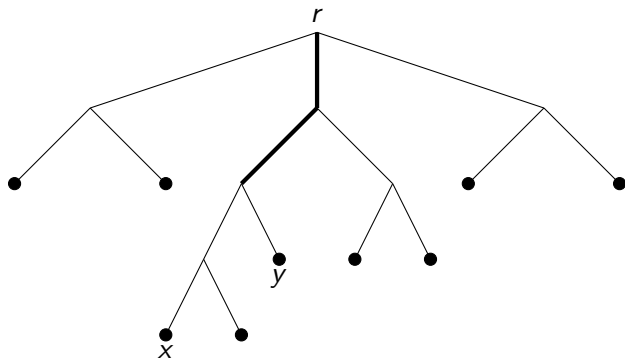
$$\text{Hyp}(S, \mathcal{F}, \mathbb{P}, s) := \mathbb{E}(\min\{s(X, Z), s(Y, Z)\} - s(X, Y))_+ \leq \delta,$$

where  $x_+$  denotes the positive part of a real number  $x$ , and  $X, Y, Z$  are i.i.d.  $S$ -valued random variables with law  $\mathbb{P}$ .

# Hierarchical clustering

- ▶ Recall that a (graph-theoretic) tree is a connected undirected graph without self-loops or closed paths.
- ▶ A rooted tree is a tree where one distinguished node is called the root. A node of a rooted tree is called a leaf if it is not the root and it has degree one.
- ▶ **Definition:** We will say that a tree  $T$  with root  $r$  is **compatible with  $(S, \mathcal{F})$**  if the following three conditions are satisfied:
  - ▶  $S$  is the set of leaves of  $T$ ,
  - ▶  $T \setminus S$  is a finite set, and
  - ▶ for any node  $v \in T \setminus S$ , the set of leaves that are the descendants of  $v$  is a measurable subset of  $S$ .
- ▶  $T$  gives a hierarchical clustering of  $S$  into a finite number of measurable clusters. Conversely, any such clustering defines a compatible tree.
- ▶ For  $x, y \in S$ , let  $(x, y)_r$  be the Gromov product of  $x$  and  $y$  under the graph distance on  $T$  with respect to the base point  $r$ .

# Example



**Figure:** A tree  $T$  compatible with  $S$ , with root  $r$ . The leaves of  $T$ , shown using dots, are the elements of  $S$ . The number of edges in the thickened path equals the Gromov product  $(x, y)_r$  for the graph distance on  $T$ . (Recall:  $(x, y)_r = \frac{1}{2}(d(x, r) + d(y, r) - d(x, y))$ .)

# How well can the space be embedded in a tree?

- ▶ **Definition:** We will say that  $(S, \mathcal{F}, \mathbb{P}, s)$  is  $\delta$ -tree-like if

$$\text{Tree}(S, \mathcal{F}, \mathbb{P}, s) := \inf_{T, \alpha} \mathbb{E} |s(X, Y) - \alpha(X, Y)_r| \leq \delta,$$

where  $X$  and  $Y$  are independent  $S$ -valued random variables with law  $\mathbb{P}$ , and the infimum is taken over over all  $\alpha \geq 0$  and all rooted trees  $T$  that are compatible with  $(S, \mathcal{F})$ . Here  $r$  is the root of  $T$  and  $(X, Y)_r$  is the Gromov product of  $X$  and  $Y$  under the graph distance on  $T$ , with respect to the base point  $r$ .

- ▶ If  $S$  is a metric space and  $s$  is the Gromov product with respect to some base point, then  $S$  is  $\delta$ -tree-like for some small  $\delta$  if and only if it is “approximately isometric to a real tree in an average sense”.

## Theorem (C. & Sloman, 2019)

*Given any  $\epsilon > 0$ , there is some  $\delta > 0$  depending only on  $\epsilon$ , such that if  $\text{Hyp}(S, \mathcal{F}, \mathbb{P}, s) < \delta$ , then  $\text{Tree}(S, \mathcal{F}, \mathbb{P}, s) < \epsilon$ . Conversely, given any  $\epsilon > 0$  there is some  $\delta > 0$  depending only on  $\epsilon$ , such that if  $\text{Tree}(S, \mathcal{F}, \mathbb{P}, s) < \delta$ , then  $\text{Hyp}(S, \mathcal{F}, \mathbb{P}, s) < \epsilon$ .*

- ▶ The dependence of  $\delta$  on  $\epsilon$  does not involve the space or the similarity function.
- ▶ The approximate embedding into a tree given by the above theorem does not have the unpleasant factor of  $\log |S|$  in it, unlike the result about Gromov hyperbolic spaces. In particular, our result works for infinite  $S$  too.
- ▶ To get the result about spin glasses, just take  $s$  to be some appropriate function of the overlap.
- ▶ Since any Gromov hyperbolic space is also hyperbolic in the average sense, the above theorem implies a new embedding result for Gromov hyperbolic spaces.

# Proof idea

- ▶ Take any  $t \in [0, 1]$ , and let  $G_t$  be the graph on  $S$  that puts an edge between  $x$  and  $y$  if and only if  $s(x, y) \geq t$ .
- ▶ If the space is 0-hyperbolic, then  $G_t$  is exactly a disjoint union of cliques. So we may expect that if the space is  $\delta$ -hyperbolic (in the average sense) for some small  $\delta$ , then  $G_t$  is **approximately** a disjoint union of cliques (call it  $\tilde{G}_t$ ).
- ▶ Such claims are usually not easy to prove. For example, if a graph on  $n$  vertices has  $o(n^3)$  triangles, one may expect that it can be made triangle-free by removing  $o(n^2)$  edges. But this is in fact a difficult result of Ruzsa and Szemerédi, known as the **triangle removal lemma**.
- ▶ Our claim can also be proved along similar lines, using **Szemerédi's regularity lemma** (but with much more work).
- ▶ The cliques can be constructed in such a way  $\tilde{G}_s$  is a subgraph of  $\tilde{G}_t$  whenever  $s \geq t$ . This yields the desired hierarchical clustering.

# Summary and future directions

- ▶ We introduced an average version of Gromov's measure of hyperbolicity of metric spaces.
- ▶ A space that is hyperbolic in the average sense is shown to be approximately isometric to a subset of a tree, also in an average sense. The proof is based on Szemerédi's regularity lemma.
- ▶ This result is used to construct hierarchically organized pure states in any spin glass model that satisfies the (generalized) Parisi ultrametricity ansatz.
- ▶ Gromov hyperbolicity has been very useful in a wide array of applications, such as social networks and phylogenetic trees. It is possible that average hyperbolicity will also find uses in similar contexts. (It may, in fact, be better suited for applications.)