Optimization of full-RSB spherical spin glasses

Eliran Subag

June, 2020

מכון ויצמן למדע
WEIZMANN INSTITUTE OF SCIENCE

Spherical spin glasses

Random functions (polynomials) on the unit sphere in $\mathbb{R}^N$,

$$\mathbb{S}^N := \left\{ \mathbf{x} \in \mathbb{R}^N : \ \|\mathbf{x}\| = 1 \right\}.$$

Spherical spin glasses

Random functions (polynomials) on the unit sphere in $\mathbb{R}^N$,

$$\mathbb{S}^N := \Big\{ \mathbf{x} \in \mathbb{R}^N : \ \|\mathbf{x}\| = 1 \Big\}.$$

⋆ **Pure *p*-spin:** for $p \geq 2$,

$$H_{N,p}(\mathbf{x}) = \sqrt{N} \sum_{i_1,\dots,i_p=1}^{N} J_{i_1,\dots,i_p} x_{i_1} x_{i_2} \cdots x_{i_p},$$

where $J_{i_1,\dots,i_p} \sim \mathrm{N}(0,1)$ i.i.d., $\mathbf{x} = (x_1, \dots, x_N)$.

Spherical spin glasses

Random functions (polynomials) on the unit sphere in $\mathbb{R}^N$,

$$\mathbb{S}^N := \Big\{ \mathbf{x} \in \mathbb{R}^N : \ \|\mathbf{x}\| = 1 \Big\}.$$

$\star$ **Pure *p*-spin:** for $p \geq 2$,

$$H_{N,p}(\mathbf{x}) = \sqrt{N} \sum_{i_1,\dots,i_p=1}^{N} J_{i_1,\dots,i_p} x_{i_1} x_{i_2} \cdots x_{i_p},$$

where $J_{i_1,\dots,i_p} \sim \mathrm{N}(0,1)$ i.i.d., $\mathbf{x} = (x_1, \dots, x_N)$.

$\star$ **Mixed *p*-spin:** fix some coefficients $\gamma_p \geq 0$, $p \geq 2$,

$$H_N(\mathbf{x}) = \sum_{p=2}^{\infty} \gamma_p H_{N,p}(\mathbf{x}).$$

Optimization

Given all the coefficients $J_{i_1,\ldots,i_p}$, find a ground-state configuration,

$$\frac{1}{N} H_N(\mathbf{x}_\star) \approx E_\star := \lim_{N\to\infty} \frac{1}{N} \max_{\mathbf{x}\in\mathbb{S}^N} H_N(\mathbf{x}),$$

with minimal time complexity as $N \to \infty$.

Optimization

Given all the coefficients $J_{i_1,\ldots,i_p}$, find a ground-state configuration,

$$\frac{1}{N} H_N(\mathbf{x}_\star) \approx E_\star := \lim_{N\to\infty} \frac{1}{N} \max_{\mathbf{x}\in\mathbb{S}^N} H_N(\mathbf{x}),$$

with minimal time complexity as $N \to \infty$.

Denote $p_{\max} = \max\{p : \gamma_p \neq 0\}$. (recall: $H_N(\mathbf{x}) = \sum_p \gamma_p H_{N,p}(\mathbf{x})$)

Optimization

Given all the coefficients $J_{i_1,\ldots,i_p}$, find a ground-state configuration,

$$\frac{1}{N}H_N(\mathbf{x}_\star) \approx E_\star := \lim_{N\to\infty}\frac{1}{N}\max_{\mathbf{x}\in\mathbb{S}^N} H_N(\mathbf{x}),$$

with minimal time complexity as $N \to \infty$.

Denote $p_{\max} = \max\{p : \gamma_p \neq 0\}$. (recall: $H_N(\mathbf{x}) = \sum_p \gamma_p H_{N,p}(\mathbf{x})$)

Theorem (S. '18)

If at zero temperature the support of the **Parisi measure** is $[0,1]$ and $p_{\max} < \infty$, then there is an algorithm that finds $\mathbf{x}_\star$ with

$$\left|\frac{1}{N}H_N(\mathbf{x}_\star) - E_\star\right| < \epsilon,$$

and time complexity $O(C(\epsilon)N^{p_{\max}})$.

Optimization

Theorem (S. '18)

If at zero temperature the support of the **Parisi measure** is $[0,1]$ and $p_{\max} < \infty$, then there is an algorithm that finds $\mathbf{x}_\star$ with

$$\left|\frac{1}{N}H_N(\mathbf{x}_\star) - E_\star\right| < \epsilon,$$

and time complexity $O(C(\epsilon)N^{p_{\max}})$.

Define the mixture polynomial

$$\nu(t) = \sum_{p=2}^{\infty} \gamma_p^2 t^p.$$

Support $= [0,1] \iff \nu''(t)^{-\frac{1}{2}}$ **is concave** on $(0,1]$.

Optimization

Theorem (S. '18)

If at zero temperature the support of the **Parisi measure** is $[0,1]$ and $p_{\max} < \infty$, then there is an algorithm that finds $\mathbf{x}_\star$ with

$$\left|\frac{1}{N}H_N(\mathbf{x}_\star) - E_\star\right| < \epsilon,$$

and time complexity $O(C(\epsilon)N^{p_{\max}})$.

Theorem (Montanari '18, El Alaoui-Montanari-Sellke '20)

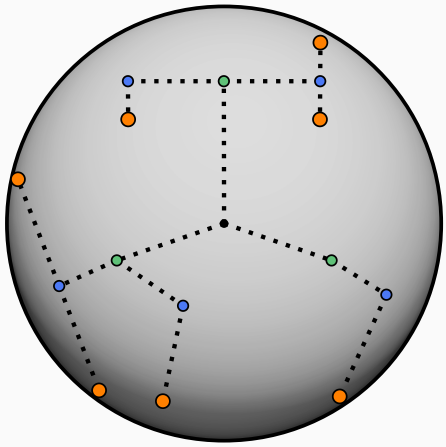
Similar result for **Ising spins**, i.e. optimization over $\boldsymbol{\sigma} \in \{\pm 1\}^N$, using massage passing algorithms.

The ultrametric tree at zero temp.

W.p.$\to 1$, there exists a tree with vertex set

$$T \subset \mathbb{B}^N = \{\mathbf{x} \in \mathbb{R}^N : \|\mathbf{x}\| \leq 1\}$$

with root $= 0$, leaves $\in \mathbb{S}^N$ such that:



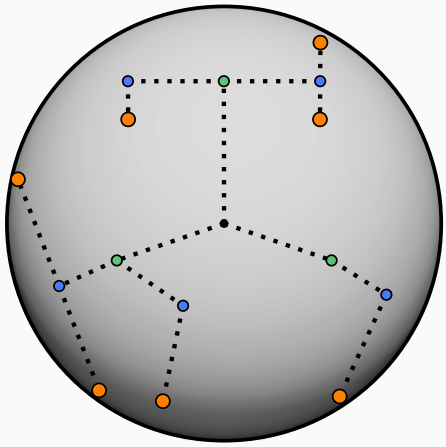
The ultrametric tree at zero temp.

W.p.$\to 1$, there exists a tree with vertex set

$$T \subset \mathbb{B}^N = \{\mathbf{x} \in \mathbb{R}^N : \|\mathbf{x}\| \leq 1\}$$

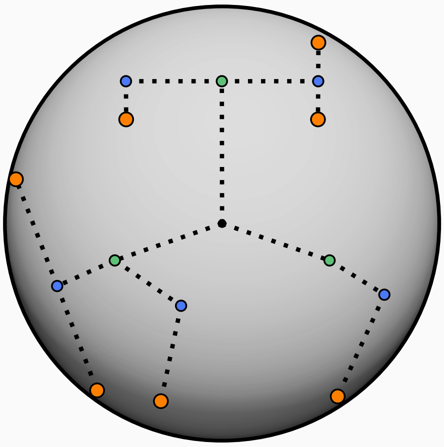
with root $= 0$, leaves $\in \mathbb{S}^N$ such that:

1. Any two **edges are orthogonal.**



The ultrametric tree at zero temp.

W.p.$\to 1$, there exists a tree with vertex set



$$T \subset \mathbb{B}^N = \{\mathbf{x} \in \mathbb{R}^N : \|\mathbf{x}\| \le 1\}$$

with root $= 0$, leaves $\in \mathbb{S}^N$ such that:

1. Any two **edges are orthogonal.**
2. **Maximality:** for any $\mathbf{x} \in T$,

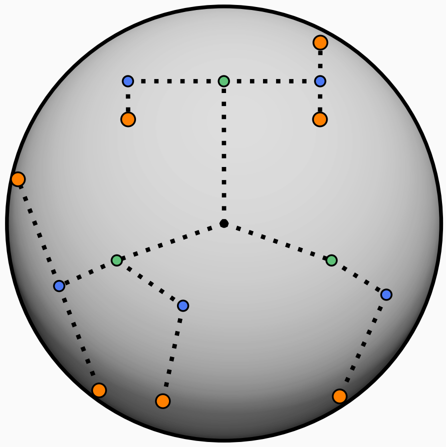
$$\frac{1}{N} H_N(\mathbf{x}) = \frac{1}{N} \max_{\|\mathbf{y}\|=\|\mathbf{x}\|} H_N(\mathbf{y}) + o(1).$$

Remark: $H_N(\mathbf{x})$ is defined on $\mathbb{B}^N$ as on the sphere $\mathbb{S}^N$.

The ultrametric tree at zero temp.

W.p.$\to 1$, there exists a tree with vertex set

$$T \subset \mathbb{B}^N = \{\mathbf{x} \in \mathbb{R}^N : \|\mathbf{x}\| \leq 1\}$$



with root $= 0$, leaves $\in \mathbb{S}^N$ such that:

1. Any two **edges are orthogonal.**
2. **Maximality:** for any $\mathbf{x} \in T$,

$$\frac{1}{N} H_N(\mathbf{x}) = \frac{1}{N} \max_{\|\mathbf{y}\|=\|\mathbf{x}\|} H_N(\mathbf{y}) + o(1).$$

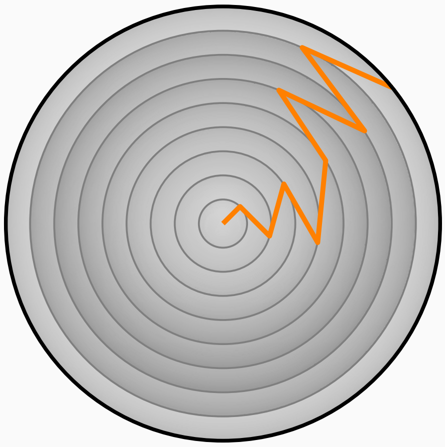
3. $\#$ levels depends on the size of the support of the Parisi measure. If support $= [0, 1]$: $\#$ levels$\to \infty$ as $N \to \infty$, vertices at all radii.

One path from the tree

What do we have for a single path from the tree?

$$\mathbf{x}_q : [0,1] \to \mathbb{R}^N, \quad \|\mathbf{x}_q\| = \sqrt{q}$$

(linearly interpolated between nodes of the tree...)



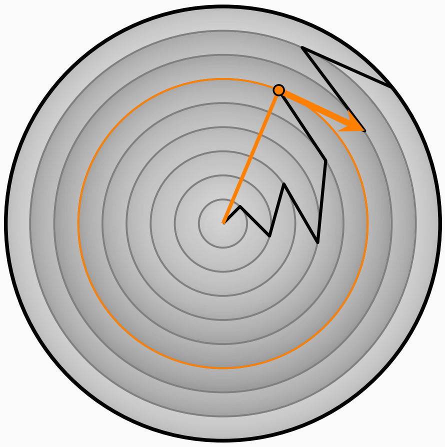
One path from the tree

What do we have for a single path from the tree?

$$\mathbf{x}_q : [0,1] \to \mathbb{R}^N, \quad \|\mathbf{x}_q\| = \sqrt{q}$$

(linearly interpolated between nodes of the tree...)

1. Each segment is orthogonal to the current position.



One path from the tree

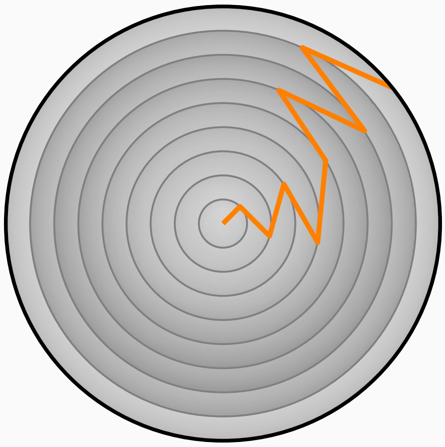
What do we have for a single path from the tree?

$$\mathbf{x}_q : [0, 1] \to \mathbb{R}^N, \quad \|\mathbf{x}_q\| = \sqrt{q}$$

(linearly interpolated between nodes of the tree...)

1. Each segment is orthogonal to the current position.
2. The energy is consistently maximal,

$$\frac{1}{N} H_N(\mathbf{x}_q) \approx \frac{1}{N} \max_{\|\mathbf{x}\|=\sqrt{q}} H_N(\mathbf{x}).$$

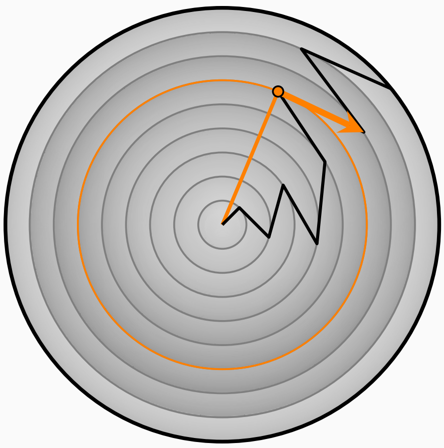


Constructing a good path

To 'imitate' the path on the tree, we set

$$\mathbf{x}_i = \sqrt{\frac{1}{k}} \sum_{j=0}^{i-1} \mathbf{v}_j, \quad i = 0, 1, \ldots, k,$$

where k is large and $\mathbf{v}_i$, $\|\mathbf{v}_i\| = 1$ are directions with $\mathbf{v}_i \perp \mathbf{x}_i$.



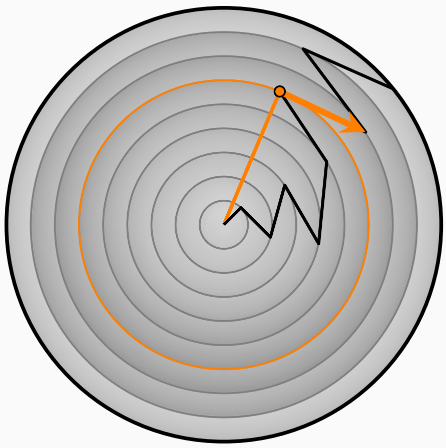
Constructing a good path

To 'imitate' the path on the tree, we set

$$\mathbf{x}_i = \sqrt{\frac{1}{k}} \sum_{j=0}^{i-1} \mathbf{v}_j, \quad i = 0, 1, \dots, k,$$

where k is large and $\mathbf{v}_i$, $\|\mathbf{v}_i\| = 1$ are directions with $\mathbf{v}_i \perp \mathbf{x}_i$.

How do we choose $\mathbf{v}_i$ to maximize the energy?



Constructing a good path

To ‘imitate’ the path on the tree, we set

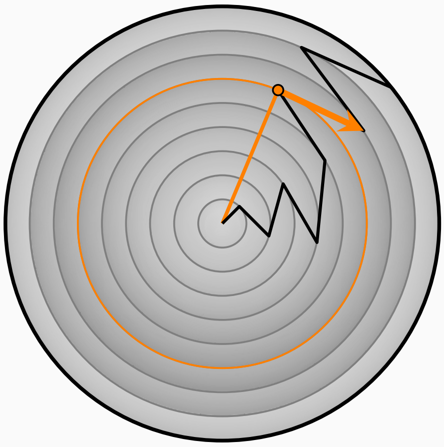
$$\mathbf{x}_i = \sqrt{\frac{1}{k}} \sum_{j=0}^{i-1} \mathbf{v}_j, \quad i = 0, 1, \ldots, k,$$

where k is large and $\mathbf{v}_i$, $\|\mathbf{v}_i\| = 1$ are directions with $\mathbf{v}_i \perp \mathbf{x}_i$.

How do we choose $\mathbf{v}_i$ to maximize the energy?

- We should have

$$H_N(\mathbf{x}_i) \approx \max_{\|\mathbf{y}\|=\|\mathbf{x}_i\|} H_N(\mathbf{y}) \implies \nabla_\perp H_N(\mathbf{x}_i) \approx 0.$$



Constructing a good path

To 'imitate' the path on the tree, we set

$$\mathbf{x}_i = \sqrt{\frac{1}{k}} \sum_{j=0}^{i-1} \mathbf{v}_j, \quad i = 0, 1, \ldots, k,$$

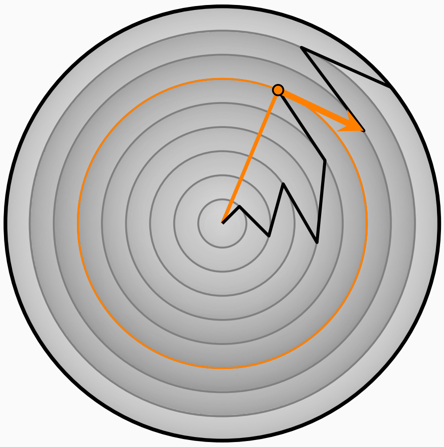
where k is large and $\mathbf{v}_i$, $\|\mathbf{v}_i\| = 1$ are directions with $\mathbf{v}_i \perp \mathbf{x}_i$.

How do we choose $\mathbf{v}_i$ to maximize the energy?

- We should have

$$H_N(\mathbf{x}_i) \approx \max_{\|\mathbf{y}\|=\|\mathbf{x}_i\|} H_N(\mathbf{y}) \implies \nabla_\perp H_N(\mathbf{x}_i) \approx 0.$$

- Go in a direction of large e.v. of $\nabla^2_\perp H_N(\mathbf{x}_i)$.



Algorithm

- ⋆ Fix some large k and small δ.

Algorithm

- ⋆ Fix some large k and small δ.
- ⋆ **Initialize:** $\mathbf{x}_1$ arbitrary with $\|\mathbf{x}_1\| = \frac{1}{\sqrt{k}}$.

Algorithm

- ⋆ Fix some large k and small δ.
- ⋆ **Initialize:** $\mathbf{x}_1$ arbitrary with $\|\mathbf{x}_1\| = \frac{1}{\sqrt{k}}$.
- ⋆ **Iteration:** find $\mathbf{v}$ s.t. $\|\mathbf{v}\| = 1$, $\mathbf{v} \perp \mathbf{x}_i$, $\mathbf{v} \cdot \nabla H_N(\mathbf{x}_i) \geq 0$,

$$\frac{1}{N}\mathbf{v}^T \nabla^2 H_N(\mathbf{x}_i)\mathbf{v} \geq \max_{\mathbf{u}} \frac{1}{N}\mathbf{u}^T \nabla^2 H_N(\mathbf{x}_i)\mathbf{u} - \delta,$$

and update

$$\mathbf{x}_{i+1} = \mathbf{x}_i + \frac{1}{\sqrt{k}}\mathbf{v}.$$

Algorithm

- ⋆ Fix some large k and small δ.
- ⋆ **Initialize:** $\mathbf{x}_1$ arbitrary with $\|\mathbf{x}_1\| = \frac{1}{\sqrt{k}}$.
- ⋆ **Iteration:** find $\mathbf{v}$ s.t. $\|\mathbf{v}\| = 1,\ \mathbf{v} \perp \mathbf{x}_i,\ \mathbf{v} \cdot \nabla H_N(\mathbf{x}_i) \geq 0,$

$$\tfrac{1}{N}\mathbf{v}^T \nabla^2 H_N(\mathbf{x}_i)\mathbf{v} \geq \max_{\mathbf{u}} \tfrac{1}{N}\mathbf{u}^T \nabla^2 H_N(\mathbf{x}_i)\mathbf{u} - \delta,$$

and update

$$\mathbf{x}_{i+1} = \mathbf{x}_i + \tfrac{1}{\sqrt{k}}\mathbf{v}.$$

Energy on path: one can show that

$$\tfrac{1}{N} H_N(\mathbf{x}_i) \geq \max_{\|\mathbf{y}\|=\|\mathbf{x}_i\|} \tfrac{1}{N} H_N(\mathbf{y}) - \epsilon,$$

where $\epsilon \to 0$ as $k \to \infty$ and $\delta \to 0$.

Algorithm

- ⋆ Fix some large k and small δ.
- ⋆ **Initialize:** $\mathbf{x}_1$ arbitrary with $\|\mathbf{x}_1\| = \frac{1}{\sqrt{k}}$.
- ⋆ **Iteration:** find $\mathbf{v}$ s.t. $\|\mathbf{v}\| = 1$, $\mathbf{v} \perp \mathbf{x}_i$, $\mathbf{v} \cdot \nabla H_N(\mathbf{x}_i) \geq 0$,

$$\frac{1}{N}\mathbf{v}^T \nabla^2 H_N(\mathbf{x}_i)\mathbf{v} \geq \max_{\mathbf{u}} \frac{1}{N}\mathbf{u}^T \nabla^2 H_N(\mathbf{x}_i)\mathbf{u} - \delta,$$

  and update

$$\mathbf{x}_{i+1} = \mathbf{x}_i + \frac{1}{\sqrt{k}}\mathbf{v}.$$

Energy on path: one can show that

$$\frac{1}{N} H_N(\mathbf{x}_i) \geq \max_{\|\mathbf{y}\| = \|\mathbf{x}_i\|} \frac{1}{N} H_N(\mathbf{y}) - \epsilon,$$

where $\epsilon \to 0$ as $k \to \infty$ and $\delta \to 0$.

Complexity: Each step takes time $O(C(\delta)N^{\deg \nu})$.

Proof sketch

Proposition (S. '18)

W.h.p., denoting $q = \|\mathbf{x}\|^2$, uniformly in $\|\mathbf{x}\| \leq 1$:

$$\text{'edge' of } \tfrac{1}{N}\nabla_\perp^2 H_N(\mathbf{x}) \approx 2\nu''(q)^{\frac{1}{2}}.$$

For $\mathbf{x}_i$, $\mathbf{v}$ as before, by a Taylor expansion,

$$\frac{1}{N}H_N(\mathbf{x}_{i+1} + \tfrac{1}{\sqrt{k}}\mathbf{v}) \geq \frac{1}{N}H_N(\mathbf{x}_i) + \tfrac{1}{k}\nu''(q)^{\frac{1}{2}} - O(\tfrac{1}{k^{3/2}}).$$

If we concatenate many small intervals like above, we get a path

$$\frac{1}{N}H_N(\mathbf{x}_q) \gtrsim \int_0^q \nu''(q)^{\frac{1}{2}} \qquad (\|\mathbf{x}_q\| = \sqrt{q}).$$

From (e.g.) Parisi formula, assuming the condition on the support,

$$\lim_{N\to\infty} \frac{1}{N} \max_{\|\mathbf{x}\|=\sqrt{q}} H_N(\mathbf{x}) = \int_0^q \nu''(q)^{\frac{1}{2}}.$$

Thank You!