## Sequential Dynamics in Ising Spin Glasses

David Gamarnik (MIT)

Joint work with Yatin Dandi, Francisco Pernice and Lenka Zdeborová

August, 2025

- Given a random Hamiltonian  $H(\sigma, \mathcal{R}), \sigma \in \Sigma$ ,  $\mathcal{R}$  encodes randomness of the model.
- Suppose an iterative algorithm  $\mathcal{A}$  is implemented sequentially  $\sigma^t = \mathcal{A}(\mathcal{R}, \sigma_1, \dots, \sigma^{t-1}), t = 1, 2, \dots$ , with fixed (random) start  $\sigma^0$ .
- Examples: Markov Chain Monte Carlo/Glauber dynamics, Greedy, Simulated Annealing, etc.
- Goal: Characterise  $H(\sigma_t, \mathcal{R})$  for "large" t.
- Verify that for bounded t, say  $t = N^{O(1)}$ , where N is model dimension.

$$H(\sigma_t, \mathcal{R}) < \max_{\sigma} H(\sigma, \mathcal{R})$$

- Given a random Hamiltonian  $H(\sigma, \mathcal{R}), \sigma \in \Sigma$ ,  $\mathcal{R}$  encodes randomness of the model.
- Suppose an iterative algorithm  $\mathcal{A}$  is implemented sequentially  $\sigma^t = \mathcal{A}(\mathcal{R}, \sigma_1, \dots, \sigma^{t-1}), t = 1, 2, \dots$ , with fixed (random) start  $\sigma^0$ .
- Examples: Markov Chain Monte Carlo/Glauber dynamics, Greedy, Simulated Annealing, etc.
- **Goal:** Characterise  $H(\sigma_t, \mathcal{R})$  for "large" t.
- Verify that for bounded t, say  $t = N^{O(1)}$ , where N is model dimension.

$$H(\sigma_t, \mathcal{R}) < \max_{\sigma} H(\sigma, \mathcal{R})$$

- Given a random Hamiltonian  $H(\sigma, \mathcal{R}), \sigma \in \Sigma$ ,  $\mathcal{R}$  encodes randomness of the model.
- Suppose an iterative algorithm  $\mathcal{A}$  is implemented sequentially  $\sigma^t = \mathcal{A}(\mathcal{R}, \sigma_1, \dots, \sigma^{t-1}), t = 1, 2, \dots$ , with fixed (random) start  $\sigma^0$ .
- Examples: Markov Chain Monte Carlo/Glauber dynamics, Greedy, Simulated Annealing, etc.
- **Goal:** Characterise  $H(\sigma_t, \mathcal{R})$  for "large" t.
- Verify that for bounded t, say  $t = N^{O(1)}$ , where N is model dimension.

$$H(\sigma_t, \mathcal{R}) < \max_{\sigma} H(\sigma, \mathcal{R})$$

- Given a random Hamiltonian  $H(\sigma, \mathcal{R}), \sigma \in \Sigma$ ,  $\mathcal{R}$  encodes randomness of the model.
- Suppose an iterative algorithm  $\mathcal{A}$  is implemented sequentially  $\sigma^t = \mathcal{A}(\mathcal{R}, \sigma_1, \dots, \sigma^{t-1}), t = 1, 2, \dots$ , with fixed (random) start  $\sigma^0$ .
- Examples: Markov Chain Monte Carlo/Glauber dynamics, Greedy, Simulated Annealing, etc.
- **Goal:** Characterise  $H(\sigma_t, \mathcal{R})$  for "large" t.
- Verify that for bounded t, say  $t = N^{O(1)}$ , where N is model dimension.

$$H(\sigma_t, \mathcal{R}) < \max_{\sigma} H(\sigma, \mathcal{R})$$

- Given a random Hamiltonian  $H(\sigma, \mathcal{R}), \sigma \in \Sigma$ ,  $\mathcal{R}$  encodes randomness of the model.
- Suppose an iterative algorithm  $\mathcal{A}$  is implemented sequentially  $\sigma^t = \mathcal{A}(\mathcal{R}, \sigma_1, \dots, \sigma^{t-1}), t = 1, 2, \dots$ , with fixed (random) start  $\sigma^0$ .
- Examples: Markov Chain Monte Carlo/Glauber dynamics, Greedy, Simulated Annealing, etc.
- **Goal:** Characterise  $H(\sigma_t, \mathcal{R})$  for "large" t.
- Verify that for bounded t, say  $t = N^{O(1)}$ , where N is model dimension.

$$H(\sigma_t, \mathcal{R}) < \max_{\sigma} H(\sigma, \mathcal{R})$$

- Given a random Hamiltonian  $H(\sigma, \mathcal{R}), \sigma \in \Sigma$ ,  $\mathcal{R}$  encodes randomness of the model.
- Suppose an iterative algorithm  $\mathcal{A}$  is implemented sequentially  $\sigma^t = \mathcal{A}(\mathcal{R}, \sigma_1, \dots, \sigma^{t-1}), t = 1, 2, \dots$ , with fixed (random) start  $\sigma^0$ .
- Examples: Markov Chain Monte Carlo/Glauber dynamics, Greedy, Simulated Annealing, etc.
- **Goal:** Characterise  $H(\sigma_t, \mathcal{R})$  for "large" t.
- Verify that for bounded t, say  $t = N^{O(1)}$ , where N is model dimension.

$$H(\sigma_t, \mathcal{R}) < \max_{\sigma} H(\sigma, \mathcal{R})$$

- This work: computing the dynamics  $H(\sigma_t, \mathcal{R})$  for spin glasses, specifically Sherrington-Kirkpatrick (2-body) N-spin model for linear time scale T = O(N).
- The remainder of this and next talk:
  - Model background
  - Main results
  - Prior literature
  - Algorithmic implications
  - Part II by Fran Pernice. Proof approaches: (i) Gaussian conditioning principle (ii) Cavity argument

- This work: computing the dynamics  $H(\sigma_t, \mathcal{R})$  for spin glasses, specifically Sherrington-Kirkpatrick (2-body) N-spin model for linear time scale T = O(N).
- The remainder of this and next talk:
  - Model background
  - Main results
  - Prior literature
  - Algorithmic implications
  - Part II by Fran Pernice. Proof approaches: (i) Gaussian conditioning principle (ii) Cavity argument

- This work: computing the dynamics  $H(\sigma_t, \mathcal{R})$  for spin glasses, specifically Sherrington-Kirkpatrick (2-body) N-spin model for linear time scale T = O(N).
- The remainder of this and next talk:
  - Model background
  - Main results
  - Prior literature
  - Algorithmic implications
  - Part II by Fran Pernice. Proof approaches: (i) Gaussian conditioning principle (ii) Cavity argument

- This work: computing the dynamics  $H(\sigma_t, \mathcal{R})$  for spin glasses, specifically Sherrington-Kirkpatrick (2-body) N-spin model for linear time scale T = O(N).
- The remainder of this and next talk:
  - Model background
  - Main results
  - Prior literature
  - Algorithmic implications
  - Part II by Fran Pernice. Proof approaches: (i) Gaussian conditioning principle (ii) Cavity argument

- This work: computing the dynamics  $H(\sigma_t, \mathcal{R})$  for spin glasses, specifically Sherrington-Kirkpatrick (2-body) N-spin model for linear time scale T = O(N).
- The remainder of this and next talk:
  - Model background
  - Main results
  - Prior literature
  - Algorithmic implications
  - Part II by Fran Pernice. Proof approaches: (i) Gaussian conditioning principle (ii) Cavity argument

- This work: computing the dynamics  $H(\sigma_t, \mathcal{R})$  for spin glasses, specifically Sherrington-Kirkpatrick (2-body) N-spin model for linear time scale T = O(N).
- The remainder of this and next talk:
  - Model background
  - Main results
  - Prior literature
  - Algorithmic implications
  - Part II by Fran Pernice. Proof approaches: (i) Gaussian conditioning principle (ii) Cavity argument

- This work: computing the dynamics  $H(\sigma_t, \mathcal{R})$  for spin glasses, specifically Sherrington-Kirkpatrick (2-body) N-spin model for linear time scale T = O(N).
- The remainder of this and next talk:
  - Model background
  - Main results
  - Prior literature
  - Algorithmic implications
  - Part II by Fran Pernice. Proof approaches: (i) Gaussian conditioning principle (ii) Cavity argument

- This work: computing the dynamics  $H(\sigma_t, \mathcal{R})$  for spin glasses, specifically Sherrington-Kirkpatrick (2-body) N-spin model for linear time scale T = O(N).
- The remainder of this and next talk:
  - Model background
  - Main results
  - Prior literature
  - Algorithmic implications
  - Part II by Fran Pernice. Proof approaches: (i) Gaussian conditioning principle (ii) Cavity argument

- Given  $J = (J_{ij}, 1 \le i < j \le N)$  i.i.d.  $\mathcal{N}(0, 1/N)$  Gaussian symmetric matrix.
- Sherrington-Kirkpatrick model (2-spin model).

$$H(\sigma, J) = \langle \sigma, J\sigma \rangle = \sum_{ij} J_{ij}\sigma_i\sigma_j, \qquad \sigma \in \{\pm 1\}^N.$$

• *p*-spin model:  $J = (J_{i_1,...,i_p}) \in \mathbb{R}^{N \otimes p}$  i.i.d.  $\mathcal{N}\left(0, \frac{1}{N^{\frac{p-1}{2}}}\right)$ .

$$H(\sigma, J) = \langle J, \sigma^{\otimes p} \rangle = \sum_{i_1, \dots, i_p} J_{i_1, \dots, i_p} \sigma_{i_1} \sigma_{i_2} \cdots \sigma_{i_p}, \qquad \sigma \in \{\pm 1\}^N.$$

- Given  $J = (J_{ij}, 1 \le i < j \le N)$  i.i.d.  $\mathcal{N}(0, 1/N)$  Gaussian symmetric matrix.
- Sherrington-Kirkpatrick model (2-spin model).

$$H(\sigma, J) = \langle \sigma, J\sigma \rangle = \sum_{ij} J_{ij}\sigma_i\sigma_j, \qquad \sigma \in \{\pm 1\}^N.$$

• *p*-spin model: 
$$J=(J_{i_1,...,i_p})\in\mathbb{R}^{N\otimes p}$$
 i.i.d.  $\mathcal{N}\left(0,\frac{1}{N^{\frac{p-1}{2}}}\right)$ .

$$H(\sigma,J)=\langle J,\sigma^{\otimes p}\rangle=\sum_{i_1,\ldots,i_p}J_{i_1,\ldots,i_p}\sigma_{i_1}\sigma_{i_2}\cdots\sigma_{i_p},\qquad \sigma\in\{\pm 1\}^N.$$

- Given  $J = (J_{ij}, 1 \le i < j \le N)$  i.i.d.  $\mathcal{N}(0, 1/N)$  Gaussian symmetric matrix.
- Sherrington-Kirkpatrick model (2-spin model).

$$H(\sigma, J) = \langle \sigma, J\sigma \rangle = \sum_{ij} J_{ij}\sigma_i\sigma_j, \qquad \sigma \in \{\pm 1\}^N.$$

• 
$$p$$
-spin model:  $J=(J_{i_1,...,i_p})\in\mathbb{R}^{N\otimes p}$  i.i.d.  $\mathcal{N}\left(0,\frac{1}{N^{\frac{p-1}{2}}}\right)$ .

$$H(\sigma, J) = \langle J, \sigma^{\otimes p} \rangle = \sum_{i_1, i_2, \dots, i_p} J_{i_1, \dots, i_p} \sigma_{i_1} \sigma_{i_2} \cdots \sigma_{i_p}, \qquad \sigma \in \{\pm 1\}^N.$$

- Given  $J = (J_{ij}, 1 \le i < j \le N)$  i.i.d.  $\mathcal{N}(0, 1/N)$  Gaussian symmetric matrix.
- Sherrington-Kirkpatrick model (2-spin model).

$$H(\sigma, J) = \langle \sigma, J\sigma \rangle = \sum_{ij} J_{ij}\sigma_i\sigma_j, \qquad \sigma \in \{\pm 1\}^N.$$

• p-spin model:  $J=(J_{i_1,...,i_p})\in\mathbb{R}^{N\otimes p}$  i.i.d.  $\mathcal{N}\left(0,\frac{1}{N^{\frac{p-1}{2}}}\right)$ .

$$H(\sigma, J) = \langle J, \sigma^{\otimes p} \rangle = \sum_{i_1, \dots, i_p} J_{i_1, \dots, i_p} \sigma_{i_1} \sigma_{i_2} \cdots \sigma_{i_p}, \qquad \sigma \in \{\pm 1\}^N.$$

- Given  $J = (J_{ij}, 1 \le i < j \le N)$  i.i.d.  $\mathcal{N}(0, 1/N)$  Gaussian symmetric matrix.
- Sherrington-Kirkpatrick model (2-spin model).

$$H(\sigma, J) = \langle \sigma, J\sigma \rangle = \sum_{ij} J_{ij} \sigma_i \sigma_j, \qquad \sigma \in \{\pm 1\}^N.$$

• p-spin model:  $J=(J_{i_1,...,i_p})\in\mathbb{R}^{N\otimes p}$  i.i.d.  $\mathcal{N}\left(0,\frac{1}{N^{\frac{p-1}{2}}}\right)$ .

$$H(\sigma, J) = \langle J, \sigma^{\otimes p} \rangle = \sum_{i_1, \dots, i_p} J_{i_1, \dots, i_p} \sigma_{i_1} \sigma_{i_2} \cdots \sigma_{i_p}, \qquad \sigma \in \{\pm 1\}^N.$$

## Ground states and free energy

### Parisi [1980], Talagrand [2006]

• The limit ground state (Parisi constant)

$$\lim_{N} \frac{1}{N} \max_{\sigma \in \{\pm\}^{N}} H(\sigma, J) \triangleq \eta_{\rho, \text{Parisi}}$$

exists and can be computed via PDE.

The limit free energy

$$\lim_{N} \frac{1}{N} \log \left( \sum_{\sigma \in \{\pm\}^{N}} \exp\left(-\beta H(\sigma, J)\right) \right) \triangleq f(p, \beta)$$

exists and can be computed via (similar) PDE.

## Ground states and free energy

### Parisi [1980], Talagrand [2006]

• The limit ground state (Parisi constant)

$$\lim_{N} \frac{1}{N} \max_{\sigma \in \{\pm\}^{N}} H(\sigma, J) \triangleq \eta_{\rho, \text{Parisi}}$$

#### exists and can be computed via PDE.

The limit free energy

$$\lim_{N} \frac{1}{N} \log \left( \sum_{\sigma \in \{\pm\}^{N}} \exp\left(-\beta H(\sigma, J)\right) \right) \triangleq f(p, \beta)$$

exists and can be computed via (similar) PDE

## Ground states and free energy

#### Parisi [1980], Talagrand [2006]

The limit ground state (Parisi constant)

$$\lim_{N} \frac{1}{N} \max_{\sigma \in \{\pm\}^{N}} H(\sigma, J) \triangleq \eta_{p, \text{Parisi}}$$

exists and can be computed via PDE.

The limit free energy

$$\lim_{N} \frac{1}{N} \log \left( \sum_{\sigma \in \{\pm\}^{N}} \exp \left( -\beta H(\sigma, J) \right) \right) \triangleq f(p, \beta)$$

exists and can be computed via (similar) PDE.

## **Fields**

• Given  $\sigma \in \{\pm 1\}^N$  define the field at spin *i* as

$$h_i = h_i(\sigma) \triangleq \sum_{i \neq i} J_{ij}\sigma_j.$$

Note:

$$H(\sigma) = \sum_{i} \sigma_{i} h_{i}.$$

## **Fields**

• Given  $\sigma \in \{\pm 1\}^N$  define the field at spin *i* as

$$h_i = h_i(\sigma) \triangleq \sum_{j \neq i} J_{ij}\sigma_j.$$

Note:

$$H(\sigma) = \sum_{i} \sigma_{i} h_{i}.$$

## **Fields**

• Given  $\sigma \in \{\pm 1\}^N$  define the field at spin *i* as

$$h_i = h_i(\sigma) \triangleq \sum_{j \neq i} J_{ij}\sigma_j.$$

Note:

$$H(\sigma) = \sum_{i} \sigma_{i} h_{i}.$$

Order of updated spins i(t):

- (a) i.i.d. uniform random  $i(t) \in [N]$ .
- (b) Scanning dynamics: i(1) = 1, i(2) = 2, ..., i(N) = N, i(N + 1) = 1, i(N + 2) = 2, ... etc. T scans = TN updates.

## Greedy

- Initialize  $\sigma^{(0)} \in \{\pm 1\}^N$ .
- ② For t = 1, 2, ... update one coordinate i(t) greedily, keep others the same

$$\begin{split} \sigma_{i(t)}^{(t)} &= \arg\max_{\sigma = \pm 1} \sigma h_i^t = \mathrm{Sign} \left( h_i^t \right) \\ \sigma_j^{(t)} &= \sigma_j^{(t-1)}, \qquad j \neq i(t). \end{split}$$

## Order of updated spins i(t):

- (a) i.i.d. uniform random  $i(t) \in [N]$ .
- (b) Scanning dynamics: i(1) = 1, i(2) = 2, ..., i(N) = N, i(N+1) = 1, i(N+2) = 2, ... etc. T scans = TN updates.

## Greedy.

- Initialize  $\sigma^{(0)} \in \{\pm 1\}^N$ .
- ② For t = 1, 2, ... update one coordinate i(t) greedily, keep others the same

$$\sigma_{i(t)}^{(t)} = \arg\max_{\sigma = \pm 1} \sigma h_i^t = \operatorname{Sign} \left( h_i^t \right)$$
$$\sigma_j^{(t)} = \sigma_j^{(t-1)}, \qquad j \neq i(t).$$

## Order of updated spins i(t):

- (a) i.i.d. uniform random  $i(t) \in [N]$ .
- (b) Scanning dynamics: i(1) = 1, i(2) = 2, ..., i(N) = N, i(N+1) = 1, i(N+2) = 2, ... etc. T scans = TN updates.

## Greedy.

- Initialize  $\sigma^{(0)} \in \{\pm 1\}^N$ .
- ② For t = 1, 2, ... update one coordinate i(t) greedily, keep others the same

$$\begin{split} \sigma_{i(t)}^{(t)} &= \arg\max_{\sigma = \pm 1} \sigma h_i^t = \mathrm{Sign} \left( h_i^t \right) \\ \sigma_j^{(t)} &= \sigma_j^{(t-1)}, \qquad j \neq i(t). \end{split}$$

## Order of updated spins i(t):

- (a) i.i.d. uniform random  $i(t) \in [N]$ .
- (b) Scanning dynamics: i(1) = 1, i(2) = 2, ..., i(N) = N, i(N+1) = 1, i(N+2) = 2, ... etc. T scans = TN updates.

### Greedy.

- Initialize  $\sigma^{(0)} \in \{\pm 1\}^N$ .
- For t = 1, 2, ... update one coordinate i(t) greedily, keep others the same

$$\sigma_{i(t)}^{(t)} = \arg\max_{\sigma = \pm 1} \sigma h_i^t = \operatorname{Sign}(h_i^t)$$
  
$$\sigma_j^{(t)} = \sigma_j^{(t-1)}, \qquad j \neq i(t).$$

# Dynamics. Glauber

### Glauber dynamics.

- 1 Initialize  $\sigma^{(0)} \in \{\pm 1\}^N$ .
- ② For t = 1, 2, ... update one coordinate i(t) randomly, keep others the same

$$\mathbb{P}\left(\sigma_{i(t)}^{(t)} = \pm 1\right) = \frac{\exp(\pm\beta h_i)}{\exp(\beta h_i) + \exp(-\beta h_i)}$$
$$\sigma_{j}^{(t)} = \sigma_{j}^{(t-1)}, \quad j \neq i(t).$$

**General dynamics:** any randomized rule of the form  $\mathbb{P}(\sigma_i = 1) = c(h_i^t)$  for some fixed function  $c(\cdot)$ .  $c(x) = (1/2)(\tanh(\beta x) + 1)$  for the Glauber dynamics

# Dynamics. Glauber

### Glauber dynamics.

- **1** Initialize  $\sigma^{(0)} \in \{\pm 1\}^N$ .
- **②** For t = 1, 2, ... update one coordinate i(t) randomly, keep others the same

$$\mathbb{P}\left(\sigma_{i(t)}^{(t)} = \pm 1\right) = \frac{\exp(\pm \beta h_i)}{\exp(\beta h_i) + \exp(-\beta h_i)}$$
$$\sigma_j^{(t)} = \sigma_j^{(t-1)}, \quad j \neq i(t).$$

**General dynamics:** any randomized rule of the form  $\mathbb{P}(\sigma_i = 1) = c(h_i^t)$  for some fixed function  $c(\cdot)$ .  $c(x) = (1/2)(\tanh(\beta x) + 1)$  for the Glauber dynamics

# Dynamics. Glauber

### Glauber dynamics.

- Initialize  $\sigma^{(0)} \in \{\pm 1\}^N$ .
- ② For t = 1, 2, ... update one coordinate i(t) randomly, keep others the same

$$\mathbb{P}\left(\sigma_{i(t)}^{(t)} = \pm 1\right) = \frac{\exp(\pm \beta h_i)}{\exp(\beta h_i) + \exp(-\beta h_i)}$$
$$\sigma_j^{(t)} = \sigma_j^{(t-1)}, \qquad j \neq i(t).$$

**General dynamics:** any randomized rule of the form  $\mathbb{P}(\sigma_i = 1) = c(h_i^t)$  for some fixed function  $c(\cdot)$ .  $c(x) = (1/2)(\tanh(\beta x) + 1)$  for the Glauber dynamics.

## Main result. Scanning dynamics, informally

- $\sigma^t \in \{\pm 1\}^N$  spin configuration at the end of scan  $t = 1, 2, \dots, T$ .
- Fix  $x \in [0, 1]$ . We give a recursive characterization of the limiting distribution of  $\left(\left(\sigma_{xN}^t, h_{xN}^t\right), 1 \le t \le T\right)$  for every constant T

$$((\sigma_{xN}^t, h_{xN}^t), 1 \le t \le T) \xrightarrow{\text{means}} ((\sigma_x^t, h_x^t), 1 \le t \le T),$$
a certain r.v. 
$$((\sigma_x^t, h_x^t), 1 \le t \le T) \in \{\pm 1\}^T \times \mathbb{R}^T \text{ which}$$

Asymptotic expected energy achieved at time TN (after T scans) satisfies

$$\lim_{N} \frac{1}{N} \sum_{i} \sigma_{i}^{T} h_{i}^{T} = \sum_{t < T} \mathbb{E} \int_{0}^{1} (\sigma_{x}^{t} - \sigma_{x}^{t-1}) h_{x}^{t} dx.$$

# Main result. Scanning dynamics, informally

- $\sigma^t \in \{\pm 1\}^N$  spin configuration at the end of scan  $t = 1, 2, \dots, T$ .
- Fix  $x \in [0, 1]$ . We give a recursive characterization of the limiting distribution of  $(\sigma_{xN}^t, h_{xN}^t)$ ,  $1 \le t \le T$  for every constant T

$$\left(\left(\sigma_{xN}^{t}, h_{xN}^{t}\right), 1 \leq t \leq T\right) \stackrel{\text{weakly}}{\longrightarrow} \left(\left(\sigma_{x}^{t}, h_{x}^{t}\right), 1 \leq t \leq T\right),$$

for a certain r.v.  $((\sigma_X^t, h_X^t), 1 \le t \le T) \in \{\pm 1\}^T \times \mathbb{R}^T$  which we characterize next.

Asymptotic expected energy achieved at time TN (after T scans) satisfies

$$\lim_{N} \frac{1}{N} \sum_{i} \sigma_{i}^{T} h_{i}^{T} = \sum_{t \leq T} \mathbb{E} \int_{0}^{1} (\sigma_{x}^{t} - \sigma_{x}^{t-1}) h_{x}^{t} dx.$$

# Main result. Scanning dynamics, informally

- $\sigma^t \in \{\pm 1\}^N$  spin configuration at the end of scan t = 1, 2, ..., T.
- Fix  $x \in [0, 1]$ . We give a recursive characterization of the limiting distribution of  $((\sigma_{xN}^t, h_{xN}^t), 1 \le t \le T)$  for every constant T

$$\left(\left(\sigma_{xN}^{t}, h_{xN}^{t}\right), 1 \leq t \leq T\right) \stackrel{\text{weakly}}{\longrightarrow} \left(\left(\sigma_{x}^{t}, h_{x}^{t}\right), 1 \leq t \leq T\right),$$
 for a certain r.v.  $\left(\left(\sigma_{x}^{t}, h_{x}^{t}\right), 1 \leq t \leq T\right) \in \{\pm 1\}^{T} \times \mathbb{R}^{T}$  which we characterize next.

Asymptotic expected energy achieved at time TN (after T scans) satisfies

$$\lim_{N} \frac{1}{N} \sum_{i} \sigma_{i}^{T} h_{i}^{T} = \sum_{t \leq T} \mathbb{E} \int_{0}^{1} (\sigma_{x}^{t} - \sigma_{x}^{t-1}) h_{x}^{t} dx.$$

#### Main result. Scanning dynamics, informally

- $\sigma^t \in \{\pm 1\}^N$  spin configuration at the end of scan  $t = 1, 2, \dots, T$ .
- Fix  $x \in [0, 1]$ . We give a recursive characterization of the limiting distribution of  $((\sigma_{xN}^t, h_{xN}^t), 1 \le t \le T)$  for every constant T

$$\left(\left(\sigma_{xN}^{t},h_{xN}^{t}\right),1\leq t\leq T\right)\overset{\text{weakly}}{\longrightarrow}\left(\left(\sigma_{x}^{t},h_{x}^{t}\right),1\leq t\leq T\right),$$
 for a certain r.v.  $\left(\left(\sigma_{x}^{t},h_{x}^{t}\right),1\leq t\leq T\right)\in\{\pm 1\}^{T}\times\mathbb{R}^{T}$  which we characterize next.

Asymptotic expected energy achieved at time TN (after T scans) satisfies

$$\lim_{N} \frac{1}{N} \sum_{i} \sigma_{i}^{T} h_{i}^{T} = \sum_{t \leq T} \mathbb{E} \int_{0}^{1} (\sigma_{x}^{t} - \sigma_{x}^{t-1}) h_{x}^{t} dx.$$

- Fix a psd  $K \in \mathbb{R}^{T \times T}$  and a sequence of vectors  $v^t \in \mathbb{R}^t$ , 1 < t < T.
- Let  $G = (G^1, ..., G^T) \stackrel{d}{=} \mathcal{N}(0, K)$ .
- Define  $((\sigma^1, h^1), \dots, (\sigma^T, h^T)) \in \{\pm 1\}^T \times \mathbb{R}^T$  recursively via  $\sigma_0 = \pm 1$  u.a.r. and

$$\begin{cases} h^t = G^t + \langle v^t, (\sigma_0, \dots, \sigma_{t-1}) \rangle \\ \sigma^t = c(h^t, U^t). \end{cases}$$

- Fix a psd  $K \in \mathbb{R}^{T \times T}$  and a sequence of vectors  $v^t \in \mathbb{R}^t$ ,  $1 \le t \le T$ .
- Let  $G = (G^1, ..., G^T) \stackrel{d}{=} \mathcal{N}(0, K)$ .
- Define  $((\sigma^1, h^1), \dots, (\sigma^T, h^T)) \in \{\pm 1\}^T \times \mathbb{R}^T$  recursively via  $\sigma_0 = \pm 1$  u.a.r. and

$$\begin{cases} h^t = G^t + \langle v^t, (\sigma_0, \dots, \sigma_{t-1}) \rangle, \\ \sigma^t = c(h^t, U^t). \end{cases}$$

- Fix a psd  $K \in \mathbb{R}^{T \times T}$  and a sequence of vectors  $v^t \in \mathbb{R}^t$ , 1 < t < T.
- Let  $G = (G^1, ..., G^T) \stackrel{d}{=} \mathcal{N}(0, K)$ .
- Define  $((\sigma^1, h^1), \dots, (\sigma^T, h^T)) \in \{\pm 1\}^T \times \mathbb{R}^T$  recursively via  $\sigma_0 = \pm 1$  u.a.r. and

$$\left\{ egin{array}{ll} h^t &= G^t + \langle v^t, (\sigma_0, \ldots, \sigma_{t-1}) 
angle, \ \sigma^t &= c(h^t, U^t). \end{array} 
ight.$$

- Fix a psd  $K \in \mathbb{R}^{T \times T}$  and a sequence of vectors  $v^t \in \mathbb{R}^t$ ,  $1 \le t \le T$ .
- Let  $G = (G^1, ..., G^T) \stackrel{d}{=} \mathcal{N}(0, K)$ .
- Define  $((\sigma^1, h^1), \dots, (\sigma^T, h^T)) \in \{\pm 1\}^T \times \mathbb{R}^T$  recursively via  $\sigma_0 = \pm 1$  u.a.r. and

$$\begin{cases} h^t &= G^t + \langle v^t, (\sigma_0, \dots, \sigma_{t-1}) \rangle, \\ \sigma^t &= c(h^t, U^t). \end{cases}$$

- Fix a psd  $K \in \mathbb{R}^{T \times T}$  and a sequence of vectors  $v^t \in \mathbb{R}^t$ ,  $1 \le t \le T$ .
- Let  $G = (G^1, ..., G^T) \stackrel{d}{=} \mathcal{N}(0, K)$ .
- Define  $((\sigma^1, h^1), \dots, (\sigma^T, h^T)) \in \{\pm 1\}^T \times \mathbb{R}^T$  recursively via  $\sigma_0 = \pm 1$  u.a.r. and

$$\begin{cases} h^t = G^t + \langle v^t, (\sigma_0, \dots, \sigma_{t-1}) \rangle, \\ \sigma^t = c(h^t, U^t). \end{cases}$$

Introduce

$$\mathcal{C}(K, v) = \mathbb{E}\sigma^{s}\sigma^{t},$$
 Correlation function,  $0 \leq s, t \leq T$   
 $\mathcal{R}(K, v) = \mathbb{E}G^{s}\sigma^{t},$  Response function,  $0 \leq s, t \leq T$ 

• Consider an ODE uniquely defining  $K(x) \in \mathbb{R}^{T \times T}, v(x) \in \mathbb{R} \times \mathbb{R}^2 \times \cdots \mathbb{R}^T, x \in [0, 1]$ :

$$\frac{dK(x)}{dx} = C_{1:T,1:T}(K(x), v(x)) - C_{0:T-1,1:T-1}(K(x), v(x)),$$

$$\frac{dv(x)}{dx} = \mathcal{R}_{1:T,T}(K(x), v(x)) - C_{0:T-1,T-1}(K(x), v(x)).$$

#### Introduce

$$\mathcal{C}(K, v) = \mathbb{E}\sigma^{s}\sigma^{t},$$
 Correlation function,  $0 \leq s, t \leq T$ ,  $\mathcal{R}(K, v) = \mathbb{E}G^{s}\sigma^{t},$  Response function,  $0 \leq s, t \leq T$ .

• Consider an ODE uniquely defining  $K(x) \in \mathbb{R}^{T \times T}, v(x) \in \mathbb{R} \times \mathbb{R}^2 \times \cdots \mathbb{R}^T, x \in [0, 1]$ :

$$\frac{dK(x)}{dx} = C_{1:T,1:T}(K(x), v(x)) - C_{0:T-1,1:T-1}(K(x), v(x)),$$

$$\frac{dv(x)}{dx} = \mathcal{R}_{1:T,T}(K(x), v(x)) - C_{0:T-1,T-1}(K(x), v(x)).$$

Introduce

$$\mathcal{C}(K, v) = \mathbb{E}\sigma^{s}\sigma^{t},$$
 Correlation function,  $0 \leq s, t \leq T$ ,  $\mathcal{R}(K, v) = \mathbb{E}G^{s}\sigma^{t},$  Response function,  $0 \leq s, t \leq T$ .

• Consider an ODE uniquely defining  $K(x) \in \mathbb{R}^{T \times T}, v(x) \in \mathbb{R} \times \mathbb{R}^2 \times \cdots \mathbb{R}^T, x \in [0, 1]$ :

$$\frac{dK(x)}{dx} = C_{1:T,1:T}(K(x), v(x)) - C_{0:T-1,1:T-1}(K(x), v(x)),$$

$$\frac{dv(x)}{dx} = \mathcal{R}_{1:T,T}(K(x), v(x)) - C_{0:T-1,T-1}(K(x), v(x)).$$

## Boundary condition

$$K(0) = \int_0^1 C_{0:T-1,0:T-1}(K(x), v(x)) dx,$$

$$v(0) = \int_0^1 R_{0:T-1,T-1}(K(x), v(x)) dx.$$

#### Theorem

Fix  $x_1, \ldots, x_R \in [0, 1]$ . Then the distribution of

$$(\sigma_{x_rN}^t, h_{x_rN}^t, 1 \leq t \leq T, 1 \leq r \leq R) \stackrel{\text{weakly}}{\longrightarrow} (\sigma_{x_r}^t, h_{x_r}^t, 1 \leq t \leq T).$$

converges to the independent product of  $(\sigma_{x_r}^t, h_{x_r}^t, 1 \le t \le T), 1 \le r \le R.$ 

#### Corollary

Let  $(\sigma_x^t, h_x^t, 1 \le t \le T) \in \{\pm 1\}^T \times \mathbb{R}^T$  be distributed according the stochastic process driven by (K(x), v(x)) solving this ODE. For every "nice" update function c and (pseudo-Lipschitz) test function  $\phi : \{\pm 1\}^T \times \mathbb{R}^T \to \mathbb{R}$ 

$$\lim_N \frac{1}{N} \sum_{1 \leq i \leq N} \phi(\sigma_i^t, h_i^t, 1 \leq t \leq T) = \int_0^1 \mathbb{E} \phi(\sigma_x^t, h_x^t, 1 \leq t \leq T) dx.$$

- Physics derivation of equations for Langevin dynamics for spherical spin glasses. Discovery of "aging" Cugliandolo and Kurchan [93]
- Rigorous verification of CK equations Ben Arous, Dembo and Guionnet, [95],[97],[01]

$$C(s,t) \sim \frac{1}{\left(\frac{t}{s}\right)^{\frac{3}{4}}}, \qquad 0 << s << t.$$

- for large s you need  $t \geq s(1/\epsilon)^{4/3}$  to get  $C < \epsilon$ .
- Energy achieved by Langevin dynamics in spherical models in linear time scale. Sellke [23]

- Physics derivation of equations for Langevin dynamics for spherical spin glasses. Discovery of "aging" Cugliandolo and Kurchan [93]
- Rigorous verification of CK equations Ben Arous, Dembo and Guionnet, [95],[97],[01]

$$C(s,t) \sim \frac{1}{\left(\frac{t}{s}\right)^{\frac{3}{4}}}, \qquad 0 << s << t.$$

- for large s you need  $t \ge s(1/\epsilon)^{4/3}$  to get  $C < \epsilon$ .
- Energy achieved by Langevin dynamics in spherical models in linear time scale. Sellke [23]

- Physics derivation of equations for Langevin dynamics for spherical spin glasses. Discovery of "aging" Cugliandolo and Kurchan [93]
- Rigorous verification of CK equations Ben Arous, Dembo and Guionnet, [95],[97],[01]

$$C(s,t) \sim \frac{1}{\left(\frac{t}{s}\right)^{\frac{3}{4}}}, \qquad 0 << s << t.$$

- for large s you need  $t \ge s(1/\epsilon)^{4/3}$  to get  $C < \epsilon$ .
- Energy achieved by Langevin dynamics in spherical models in linear time scale. Sellke [23]

- Physics derivation of equations for Langevin dynamics for spherical spin glasses. Discovery of "aging" Cugliandolo and Kurchan [93]
- Rigorous verification of CK equations Ben Arous, Dembo and Guionnet, [95],[97],[01]

$$C(s,t) \sim \frac{1}{\left(\frac{t}{s}\right)^{\frac{3}{4}}}, \qquad 0 << s << t.$$

- for large s you need  $t \ge s(1/\epsilon)^{4/3}$  to get  $C < \epsilon$ .
- Energy achieved by Langevin dynamics in spherical models in linear time scale. Sellke [23]

- Dynamics using TAP approach Biroli [99]
- Simulations for versions of Greedy (zero temperature) algs Parisi [03]
- MCMC for cliques in random graphs (Erdös-Rényi and Hidden Clique Model) with empty starting point Chen, Mossel, Zadik [22]
- Gradient descent from complete set for Hidden Clique Model Gheissari, Jagannath, Yu [23]
- Dynamics in  $\mathbb{Z}^d$  from warm start Alaoui, Eldan, Gheissari, Piana [25]

- Dynamics using TAP approach Biroli [99]
- Simulations for versions of Greedy (zero temperature) algs Parisi [03]
- MCMC for cliques in random graphs (Erdös-Rényi and Hidden Clique Model) with empty starting point Chen, Mossel, Zadik [22]
- Gradient descent from complete set for Hidden Clique Model Gheissari, Jagannath, Yu [23]
- Dynamics in  $\mathbb{Z}^d$  from warm start Alaoui, Eldan, Gheissari, Piana [25]

- Dynamics using TAP approach Biroli [99]
- Simulations for versions of Greedy (zero temperature) algs Parisi [03]
- MCMC for cliques in random graphs (Erdös-Rényi and Hidden Clique Model) with empty starting point Chen, Mossel, Zadik [22]
- Gradient descent from complete set for Hidden Clique Model Gheissari, Jagannath, Yu [23]
- Dynamics in  $\mathbb{Z}^d$  from warm start Alaoui, Eldan, Gheissari, Piana [25]

- Dynamics using TAP approach Biroli [99]
- Simulations for versions of Greedy (zero temperature) algs Parisi [03]
- MCMC for cliques in random graphs (Erdös-Rényi and Hidden Clique Model) with empty starting point Chen, Mossel, Zadik [22]
- Gradient descent from complete set for Hidden Clique Model Gheissari, Jagannath, Yu [23]
- Dynamics in  $\mathbb{Z}^d$  from warm start Alaoui, Eldan, Gheissari, Piana [25]

- Dynamics using TAP approach Biroli [99]
- Simulations for versions of Greedy (zero temperature) algs Parisi [03]
- MCMC for cliques in random graphs (Erdös-Rényi and Hidden Clique Model) with empty starting point Chen, Mossel, Zadik [22]
- Gradient descent from complete set for Hidden Clique Model Gheissari, Jagannath, Yu [23]
- Dynamics in  $\mathbb{Z}^d$  from warm start Alaoui, Eldan, Gheissari Piana [25]

- Dynamics using TAP approach Biroli [99]
- Simulations for versions of Greedy (zero temperature) algs Parisi [03]
- MCMC for cliques in random graphs (Erdös-Rényi and Hidden Clique Model) with empty starting point Chen, Mossel, Zadik [22]
- Gradient descent from complete set for Hidden Clique Model Gheissari, Jagannath, Yu [23]
- Dynamics in  $\mathbb{Z}^d$  from warm start Alaoui, Eldan, Gheissari, Piana [25]

- Does the Greedy algorithm reach near optimality (ground state) η<sub>p,Parisi</sub>? No for p ≥ 3 in linear time because of the Overlap Gap Property (OGP).
- Near optimal algorithm for p = 2 (Sherrington-Kirkpatrick) exists, Montanari [19].
- No known als for  $p \ge 3$ . OGP, which exists for p = 4, 6, 8 and all  $p \ge 10$  implies no near optimal *stable* algs.
- Tight value achievable by stable algs via branching (ultrametric) OGP, Alaoui, Montanari, Sellke [21], Huang, Sellke [22].
- But Greedy/Glauber are not stable...

- Does the Greedy algorithm reach near optimality (ground state) η<sub>p,Parisi</sub>? No for p ≥ 3 in linear time because of the Overlap Gap Property (OGP).
- Near optimal algorithm for p = 2 (Sherrington-Kirkpatrick) exists, Montanari [19].
- No known als for  $p \ge 3$ . OGP, which exists for p = 4, 6, 8 and all  $p \ge 10$  implies no near optimal *stable* algs.
- Tight value achievable by stable algs via branching (ultrametric) OGP, Alaoui, Montanari, Sellke [21], Huang, Sellke [22].
- But Greedy/Glauber are not stable...

- Does the Greedy algorithm reach near optimality (ground state)  $\eta_{p,\text{Parisi}}$ ? No for  $p \geq 3$  in linear time because of the Overlap Gap Property (OGP).
- Near optimal algorithm for p = 2 (Sherrington-Kirkpatrick) exists, Montanari [19].
- No known als for  $p \ge 3$ . OGP, which exists for p = 4, 6, 8 and all  $p \ge 10$  implies no near optimal *stable* algs.
- Tight value achievable by stable algs via branching (ultrametric) OGP, Alaoui, Montanari, Sellke [21], Huang, Sellke [22].
- But Greedy/Glauber are not stable...

- Does the Greedy algorithm reach near optimality (ground state)  $\eta_{p,\text{Parisi}}$ ? No for  $p \geq 3$  in linear time because of the Overlap Gap Property (OGP).
- Near optimal algorithm for p = 2 (Sherrington-Kirkpatrick) exists, Montanari [19].
- No known als for  $p \ge 3$ . OGP, which exists for p = 4, 6, 8 and all  $p \ge 10$  implies no near optimal *stable* algs.
- Tight value achievable by stable algs via branching (ultrametric) OGP, Alaoui, Montanari, Sellke [21], Huang, Sellke [22].
- But Greedy/Glauber are not stable...

- Does the Greedy algorithm reach near optimality (ground state)  $\eta_{p,\text{Parisi}}$ ? No for  $p \geq 3$  in linear time because of the Overlap Gap Property (OGP).
- Near optimal algorithm for p = 2 (Sherrington-Kirkpatrick) exists, Montanari [19].
- No known als for  $p \ge 3$ . OGP, which exists for p = 4, 6, 8 and all  $p \ge 10$  implies no near optimal *stable* algs.
- Tight value achievable by stable algs via branching (ultrametric) OGP, Alaoui, Montanari, Sellke [21], Huang, Sellke [22].
- But Greedy/Glauber are not stable...

- Does the Greedy algorithm reach near optimality (ground state) η<sub>p,Parisi</sub>? No for p ≥ 3 in linear time because of the Overlap Gap Property (OGP).
- Near optimal algorithm for p = 2 (Sherrington-Kirkpatrick) exists, Montanari [19].
- No known als for  $p \ge 3$ . OGP, which exists for p = 4, 6, 8 and all  $p \ge 10$  implies no near optimal *stable* algs.
- Tight value achievable by stable algs via branching (ultrametric) OGP, Alaoui, Montanari, Sellke [21], Huang, Sellke [22].
- But Greedy/Glauber are not stable...

- Does the Greedy algorithm reach near optimality (ground state) η<sub>p,Parisi</sub>? No for p ≥ 3 in linear time because of the Overlap Gap Property (OGP).
- Near optimal algorithm for p = 2 (Sherrington-Kirkpatrick) exists, Montanari [19].
- No known als for  $p \ge 3$ . OGP, which exists for p = 4, 6, 8 and all  $p \ge 10$  implies no near optimal *stable* algs.
- Tight value achievable by stable algs via branching (ultrametric) OGP, Alaoui, Montanari, Sellke [21], Huang, Sellke [22].
- But Greedy/Glauber are not stable...

#### Overlap Gap Property (OGP)

## Theorem (Chen, G, Panchenko, Rahman [19], G, Jagannath, Kizildag [23])

p-spin model exhibits OGP for  $p=4,6,8,\geq 10$ . Specifically, there exists  $\eta_{p,\text{OGP}}<\eta_{p,\text{Parisi}}$  s.t. for every  $\sigma_1,\sigma_2\in\{\pm\}^N$  satisfying  $H(\sigma_1),H(\sigma_2)\geq\eta_{p,\text{OGP}}$  it holds

$$\frac{1}{N}\langle \sigma_1, \sigma_2 \rangle \notin (\nu_1, \nu_2)$$

for some  $0 < \nu_1 < \nu_2 < 1$ .

OGP is an obstruction to stable algs.

# OGP is a barrier to sequential dynamics at linear time scale

• Consider smoothed dynamics where  $\tau_i \in [-1, 1]$  using smooth approximation  $\phi$  of the  $\operatorname{sign}(x)$  function. Consider the implied dynamics  $\tau^t \in [-1, 1]^N$ ,  $h^t \in \mathbb{R}^N$ .

#### Theorem (This work

- (a) Suppose  $J \approx \hat{J}$ . Then for every T = O(1) (linear time scale),  $\tau^T(J) \approx \tau^T(\hat{J})$ . Namely  $\tau^T$  is stable as a function of disorder J. As a result  $H(\tau^T) \leq \eta_{p,OGP}$  whp.
- (b) Suppose  $\phi \approx \text{sign}(x)$ . Then  $(\sigma^T, h^T) \stackrel{\circ}{\approx} (\tau^T, h^T)$ . Then  $H(\sigma^T) \approx H(\tau^T) \leq \eta_{p,\text{OGP}}$  OGP is a barrier to sequential dynamics at linear time scale.

# OGP is a barrier to sequential dynamics at linear time scale

• Consider smoothed dynamics where  $\tau_i \in [-1, 1]$  using smooth approximation  $\phi$  of the  $\operatorname{sign}(x)$  function. Consider the implied dynamics  $\tau^t \in [-1, 1]^N$ ,  $h^t \in \mathbb{R}^N$ .

#### Theorem (This work)

- (a) Suppose  $J \approx \hat{J}$ . Then for every T = O(1) (linear time scale),  $\tau^T(J) \approx \tau^T(\hat{J})$ . Namely  $\tau^T$  is stable as a function of disorder J. As a result  $H(\tau^T) \leq \eta_{p, OGP}$  whp.
- (b) Suppose  $\phi \approx \text{sign}(x)$ . Then  $(\sigma^T, h^T) \stackrel{d}{\approx} (\tau^T, h^T)$ . Then  $H(\sigma^T) \approx H(\tau^T) \leq \eta_{p, \text{OGP}}$  OGP is a barrier to sequential dynamics at linear time scale.

- Analysis beyond linear time. Conjecture: Greedy alg gets stuck in O(N log N) steps and does not reach ground state energy
- Sparse graphs using the second (cavity) argument which does not require Gaussian conditioning principle
- Glauber dynamics for Number Partitioning Problem. (Greedy fails Kizildag [25] and gives  $\sqrt{\pi/8}$  vs  $\exp(-\Theta(\log^2 n))$  for the best poly-time alg).

- Analysis beyond linear time. Conjecture: Greedy alg gets stuck in O(N log N) steps and does not reach ground state energy
- Sparse graphs using the second (cavity) argument which does not require Gaussian conditioning principle
- Glauber dynamics for Number Partitioning Problem. (Greedy fails Kizildag [25] and gives  $\sqrt{\pi/8}$  vs  $\exp(-\Theta(\log^2 n))$  for the best poly-time alg).

- Analysis beyond linear time. Conjecture: Greedy alg gets stuck in O(N log N) steps and does not reach ground state energy
- Sparse graphs using the second (cavity) argument which does not require Gaussian conditioning principle
- Glauber dynamics for Number Partitioning Problem. (Greedy fails Kizildag [25] and gives  $\sqrt{\pi/8}$  vs  $\exp(-\Theta(\log^2 n))$  for the best poly-time alg).

- Analysis beyond linear time. Conjecture: Greedy alg gets stuck in O(N log N) steps and does not reach ground state energy
- Sparse graphs using the second (cavity) argument which does not require Gaussian conditioning principle
- Glauber dynamics for Number Partitioning Problem. (Greedy fails Kizildag [25] and gives  $\sqrt{\pi/8}$  vs  $\exp(-\Theta(\log^2 n))$  for the best poly-time alg).