The Threshold Energy of Low Temperature Langevin Dynamics for Pure Spherical Spin Glasses

Mark Sellke

Definition of a Spherical Spin Glass

Pure p-spin Hamiltonian: random function $H_N : \mathbb{R}^N \to \mathbb{R}$ given by

$$H_N(\sigma) = N^{-(p-1)/2} \sum_{1 \le i_1, i_2, \dots, i_p \le N} J_{i_1,\dots,i_p} \sigma_{i_1} \dots \sigma_{i_p}$$

with i.i.d. Gaussian coefficients $J_{i,j} \sim \mathcal{N}(0,1)$.

Inputs σ will be on the sphere: $\mathcal{S}_N = \{\sigma \in \mathbb{R}^N \ : \ \sum_{i=1}^N \sigma_i^2 = N\}$.

Definition of a Spherical Spin Glass

Pure p-spin Hamiltonian: random function $H_N : \mathbb{R}^N \to \mathbb{R}$ given by

$$H_N(\sigma) = N^{-(p-1)/2} \sum_{1 \le i_1, i_2, \ldots, i_p \le N} J_{i_1,\ldots,i_p} \sigma_{i_1} \ldots \sigma_{i_p}$$

with i.i.d. Gaussian coefficients $J_{i,j} \sim \mathcal{N}(0,1)$.

Inputs σ will be on the sphere: $\mathcal{S}_N = \{\sigma \in \mathbb{R}^N \ : \ \sum_{i=1}^N \sigma_i^2 = N\}$.

Quick facts:

1. Rotationally invariant Gaussian process: $\mathbb{E} H_N(\sigma) H_N(\rho) = N \left(\frac{\langle \sigma, \rho \rangle}{N}\right)^p$.
2. Scaling: $\|H_N\|_{L^\infty(\mathcal{S}_N)} \asymp N$, $\quad \|\nabla H_N(\sigma)\| \asymp \sqrt{N}$, $\quad \|\nabla^2 H_N(\sigma)\|_{\mathrm{op}} \asymp 1$.

Langevin dynamics on $\mathcal{S}_N$:

$$\mathrm{d}\boldsymbol{x}_t = \left(\beta \nabla_{\mathrm{sp}} H_N(\boldsymbol{x}_t) - \frac{(N-1)\boldsymbol{x}_t}{2N}\right) \mathrm{d}t + P^{\perp}_{\boldsymbol{x}_t} \, \mathrm{d}\boldsymbol{B}_t.$$

Invariant for Gibbs measure $\mu_\beta(\mathrm{d}\sigma) = e^{\beta H_N(\sigma)} \mathrm{d}\sigma / Z_N(\beta)$. Much is known even at low temperature:

- Free energy is 1-RSB [Talagrand 06]
- Geometric description: supported on deep wells, PD statistics [Subag 17].

Spherical Langevin Dynamics

Langevin dynamics on $\mathcal{S}_N$:

$$\mathrm{d}\boldsymbol{x}_t = \left(\beta\nabla_{\mathrm{sp}} H_N(\boldsymbol{x}_t) - \frac{(N-1)\boldsymbol{x}_t}{2N}\right)\mathrm{d}t + P^{\perp}_{\boldsymbol{x}_t}\,\mathrm{d}\boldsymbol{B}_t.$$

Invariant for Gibbs measure $\mu_\beta(\mathrm{d}\sigma) = e^{\beta H_N(\sigma)}\mathrm{d}\sigma / Z_N(\beta)$. Much is known even at low temperature:

- Free energy is 1-RSB [Talagrand 06]
- Geometric description: supported on deep wells, PD statistics [Subag 17].

However $t_{mix}(\beta) \geq e^{\Omega(N)}$ for large β [Ben Arous-Jagannath 18]. $\mu_\beta(\mathrm{d}\sigma)$ is inaccessible.

This belief motivated the study of dynamics on $O(1)$ time-scales independent of N.

1. Exact description via Cugliandolo-Kurchan equations [Crisanti-Horner-Sommers 93].
   - [Ben Arous-Dembo-Guionnet 06]: **Yes** (for soft spherical spins)

Physics Predictions and Rigorous Results

1. Exact description via Cugliandolo-Kurchan equations [Crisanti-Horner-Sommers 93].
   - [Ben Arous-Dembo-Guionnet 06]: **Yes** (for soft spherical spins)

2. FDT and exponential decay of correlations at high temperature.
   - [Dembo-Guionnet-Mazza 07]: **Yes**, using 1.

Physics Predictions and Rigorous Results

1. Exact description via Cugliandolo-Kurchan equations [Crisanti-Horner-Sommers 93].
   - [Ben Arous-Dembo-Guionnet 06]: **Yes** (for soft spherical spins)
2. FDT and exponential decay of correlations at high temperature.
   - [Dembo-Guionnet-Mazza 07]: **Yes**, using 1.
3. Aging at low temperatures [Cugliandolo-Kurchan 93].
   - [Ben Arous-Dembo-Guionnet 01]: **Yes**, but so far only for $p = 2$.
   - Results for related dynamics on longer time-scales, e.g. [Ben Arous-Bovier-Černỳ 08]

Physics Predictions and Rigorous Results

1. Exact description via Cugliandolo-Kurchan equations [Crisanti-Horner-Sommers 93].
   - [Ben Arous-Dembo-Guionnet 06]: **Yes** (for soft spherical spins)
2. FDT and exponential decay of correlations at high temperature.
   - [Dembo-Guionnet-Mazza 07]: **Yes**, using 1.
3. Aging at low temperatures [Cugliandolo-Kurchan 93].
   - [Ben Arous-Dembo-Guionnet 01]: **Yes**, but so far only for $p = 2$.
   - Results for related dynamics on longer time-scales, e.g. [Ben Arous-Bovier-Černỳ 08]
4. Threshold energy equals $E_\infty(p) \equiv 2\sqrt{\frac{p-1}{p}}$ as $\beta \to \infty$ [Biroli 99].
   - [Ben Arous-Gheissari-Jagannath 18]: **non-sharp bounds**, without 1.
   - [**S** 23]: **Yes**, without 1.

Cugliandolo-Kurchan Equations

Closed system of equations as $N \to \infty$ for:

$$C(s,t) \equiv \langle \boldsymbol{x}_s, \boldsymbol{x}_t \rangle / N,$$
$$R(s,t) \equiv \langle \boldsymbol{x}_s, \boldsymbol{B}_t \rangle / N.$$

Cugliandolo-Kurchan Equations

Closed system of equations as $N \to \infty$ for:

$$C(s,t) \equiv \langle \boldsymbol{x}_s, \boldsymbol{x}_t \rangle / N,$$
$$R(s,t) \equiv \langle \boldsymbol{x}_s, \boldsymbol{B}_t \rangle / N.$$

Tells you everything in principle, but hard to work with! For $s \geq t \geq 0$:

$$\partial_s R(s,t) = -\mu(s)R(s,t) + \beta^2 p(p-1)\int_t^s R(u,t)R(s,u)C(s,u)^{p-2}\,\mathrm{d}u,$$

$$\partial_s C(s,t) = -\mu(s)C(s,t) + \beta^2 p(p-1)\int_0^s C(u,t)R(s,u)C(s,u)^{p-2}\,\mathrm{d}u$$
$$+\beta^2 p\int_0^t C(s,u)^{p-1}R(t,u)\,\mathrm{d}u;$$

$$\mu(s) \equiv \frac{1}{2} + \beta^2 p^2 \int_0^s C(s,u)^{p-1}R(s,u)\,\mathrm{d}u.$$

Intuition for E_∞

[Biroli 99]: $E_\infty = 2\sqrt{\frac{p-1}{p}}$ is the threshold energy in the limit of low temperatures.

Intuition for E_∞

[Biroli 99]: $E_\infty = 2\sqrt{\frac{p-1}{p}}$ is the threshold energy in the limit of low temperatures.

Explanation:

- For $\boldsymbol{x} \in \mathcal{S}_N$, the spherical Hessian $\nabla^2_{\mathrm{sp}} H_N(\boldsymbol{x})$ is a **shifted GOE**:

$$\nabla^2_{\mathrm{sp}} H_N(\boldsymbol{x}) \stackrel{d}{=} \sqrt{p(p-1)}\, GOE(N-1) - p \cdot \frac{H_N(\boldsymbol{x})}{N}.$$

Intuition for E_∞

[Biroli 99]: $E_\infty = 2\sqrt{\frac{p-1}{p}}$ is the threshold energy in the limit of low temperatures.

Explanation:

- For $\boldsymbol{x} \in \mathcal{S}_N$, the spherical Hessian $\nabla^2_{\mathrm{sp}} H_N(\boldsymbol{x})$ is a **shifted GOE**:

$$\nabla^2_{\mathrm{sp}} H_N(\boldsymbol{x}) \stackrel{d}{=} \sqrt{p(p-1)}\, GOE(N-1) - p \cdot \frac{H_N(\boldsymbol{x})}{N}.$$

- The real prediction was that $\lambda_{\max}\left(\nabla^2_{\mathrm{sp}} H_N(\boldsymbol{x}_T)\right) \approx 0$.
- These are consistent since $\lambda_{\max}(GOE(N-1)) \approx 2$.

Intuition for E_∞

[Biroli 99]: $E_\infty = 2\sqrt{\frac{p-1}{p}}$ is the threshold energy in the limit of low temperatures.

Explanation:

- For $\boldsymbol{x} \in \mathcal{S}_N$, the spherical Hessian $\nabla^2_{\mathrm{sp}} H_N(\boldsymbol{x})$ is a **shifted GOE**:

$$\nabla^2_{\mathrm{sp}} H_N(\boldsymbol{x}) \overset{d}{=} \sqrt{p(p-1)}\, GOE(N-1) - p \cdot \frac{H_N(\boldsymbol{x})}{N}.$$

- The real prediction was that $\lambda_{\max}\left(\nabla^2_{\mathrm{sp}} H_N(\boldsymbol{x}_T)\right) \approx 0$.
- These are consistent since $\lambda_{\max}(GOE(N-1)) \approx 2$.

As $\beta \to \infty$, we expect to be **on the border of being at a local maximum**:

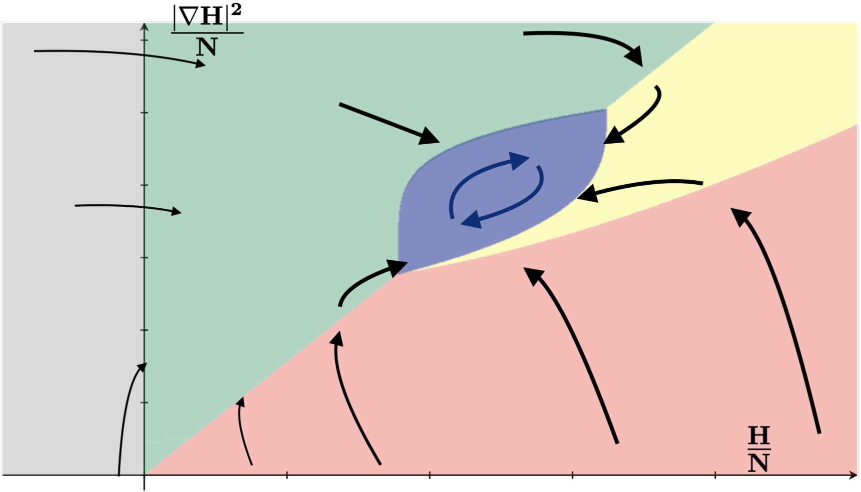
The Hessian's maximum eigenvalue "equals zero at zero temperature, as is expected for a dynamics in a rugged energy landscape".

For mixed models, suggests threshold energy is in $[E^-_\infty, E^+_\infty]$ by [Auffinger-Ben Arous 13].

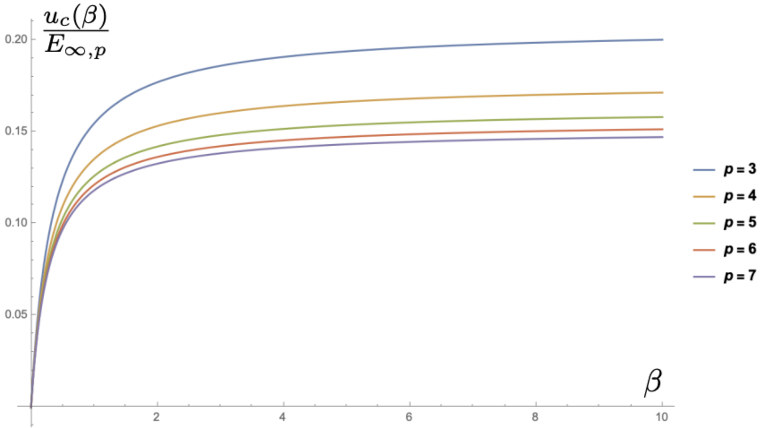
Bounding Flows Approach

Rigorously understanding the Cugliandolo-Kurchan equations is difficult at low temperature.

[Ben Arous-Gheissari-Jagannath 18]: **bounding flows** method of differential inequalities.



Yields quantitative lower bounds on the energy achieved:



This method is inexact but works for disorder dependent $x_0 \in \mathcal{S}_N$.

New Result: E_∞ is the Threshold Energy as $\beta \to \infty$

Theorem (S 23, Upper Bound)

For any β there is $\delta > 0$ such that for any T, if $\boldsymbol{x}_0 \in \mathcal{S}_N$ is independent of H_N:

$$\mathbb{P}\left[\sup_{t\in[0,T]} H_N(\boldsymbol{x}_t)/N \le E_\infty - \delta\right] \ge 1 - e^{-cN}.$$

New Result: E_∞ is the Threshold Energy as $\beta \to \infty$

Theorem (S 23, Upper Bound)

For any β there is $\delta > 0$ such that for any T, if $\boldsymbol{x}_0 \in \mathcal{S}_N$ is independent of H_N:

$$\mathbb{P}\left[\sup_{t\in[0,T]} H_N(\boldsymbol{x}_t)/N \le E_\infty - \delta\right] \ge 1 - e^{-cN}.$$

Theorem (S 23, Lower Bound)

*For any $\eta > 0$, with $T_0 \ge T_0(\eta)$ and $\beta \ge \beta_0(\eta)$, **even if $\boldsymbol{x}_0$ is disorder dependent**:*

$$\mathbb{P}\left[\inf_{t\in[T_0, T_0+e^{cN}]} H_N(\boldsymbol{x}_t)/N \ge E_\infty - \eta\right] \ge 1 - e^{-cN}.$$

New Result: E_∞ is the Threshold Energy as $\beta \to \infty$

Theorem (S 23, Upper Bound)

For any β there is $\delta > 0$ such that for any T, if $\mathbf{x}_0 \in \mathcal{S}_N$ is independent of H_N:

$$\mathbb{P}\left[\sup_{t\in[0,T]} H_N(\mathbf{x}_t)/N \leq E_\infty - \delta\right] \geq 1 - e^{-cN}.$$

Theorem (S 23, Lower Bound)

For any $\eta > 0$, with $T_0 \geq T_0(\eta)$ and $\beta \geq \beta_0(\eta)$, ***even if $\mathbf{x}_0$ is disorder dependent****:*

$$\mathbb{P}\left[\inf_{t\in[T_0,T_0+e^{cN}]} H_N(\mathbf{x}_t)/N \geq E_\infty - \eta\right] \geq 1 - e^{-cN}.$$

In particular for large constant times $t \in [T_0, T]$, the energy stays slightly below E_∞:

$$H_N(\mathbf{x}_t)/N \in [E_\infty - \delta_1(\beta), E_\infty - \delta_2(\beta)].$$

Ideas for the Upper Bound

The upper bound uses prior work with Brice Huang on stable optimization algorithms.

Definition

An L-**Lipschitz optimization algorithm** is an L-Lipschitz function $\mathcal{A}_N : \mathbb{R}^{N^p} \to \mathcal{B}_N$.

Ideas for the Upper Bound

The upper bound uses prior work with Brice Huang on stable optimization algorithms.

Definition

An L-**Lipschitz optimization algorithm** is an L-Lipschitz function $\mathcal{A}_N : \mathbb{R}^{N^p} \to \mathcal{B}_N$.

Theorem (Huang-S 21 & 23)

Fix any $L, \eta > 0$. If $\mathcal{A}_N$ is an L-Lipschitz algorithm, then for N large enough,

$$\mathbb{P}[H_N(\mathcal{A}_N(H_N))/N \leq E_\infty + \eta] \geq 1 - e^{-cN}.$$

(Informally: Lipschitz algorithms cannot access energies above E_∞.)

Ideas for the Upper Bound

The upper bound uses prior work with Brice Huang on stable optimization algorithms.

Definition

An L-**Lipschitz optimization algorithm** is an L-Lipschitz function $\mathcal{A}_N : \mathbb{R}^{N^p} \to \mathcal{B}_N$.

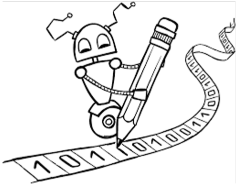
Theorem (Huang-S 21 & 23)

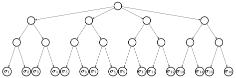
Fix any $L, \eta > 0$. If $\mathcal{A}_N$ is an L-Lipschitz algorithm, then for N large enough,

$$\mathbb{P}[H_N(\mathcal{A}_N(H_N))/N \leq E_\infty + \eta] \geq 1 - e^{-cN}.$$

(Informally: Lipschitz algorithms cannot access energies above E_∞.)

Proof: branching overlap gap property.
Extends OGP from [Gamarnik-Sudan 14,...].





Ideas for the Upper Bound

For the upper bound, we approximate $\boldsymbol{x}_T$ by a $L(\beta, T)$-Lipschitz algorithm for each $\boldsymbol{B}_{[0,T]}$.

Previously known for modified versions of spherical Langevin [Ben Arous-Dembo-Guionnet 06, Huang-**S** 21].

Ideas for the Upper Bound

For the upper bound, we approximate $\boldsymbol{x}_T$ by a $L(\beta, T)$-Lipschitz algorithm for each $\boldsymbol{B}_{[0,T]}$.

Previously known for modified versions of spherical Langevin [Ben Arous-Dembo-Guionnet 06, Huang-**S** 21].

Improving from $E_\infty + \eta$ to $E_\infty - \delta(\beta)$:

- [Ben Arous-Gheissari-Jagannath 18]: $\|\nabla_{\text{sp}} H_N(\boldsymbol{x}_t)\| \geq \delta'(\beta)\sqrt{N}$.
- Quenching the dynamics to zero temperature for time $[t, t+\varepsilon]$ gives slightly more energy.

Ideas for the Upper Bound

For the upper bound, we approximate $\boldsymbol{x}_T$ by a $L(\beta, T)$-Lipschitz algorithm for each $\boldsymbol{B}_{[0,T]}$.

Previously known for modified versions of spherical Langevin [Ben Arous-Dembo-Guionnet 06, Huang-**S** 21].

Improving from $E_\infty + \eta$ to $E_\infty - \delta(\beta)$:

- [Ben Arous-Gheissari-Jagannath 18]: $\|\nabla_{\mathrm{sp}} H_N(\boldsymbol{x}_t)\| \geq \delta'(\beta)\sqrt{N}$.
- Quenching the dynamics to zero temperature for time $[t, t+\varepsilon]$ gives slightly more energy.

The Lipschitz approximation goes via *soft* spherical Langevin dynamics. As a byproduct, we extend [Ben Arous-Dembo-Guionnet 06] to the hard spherical case:

Corollary (S 23)

Langevin dynamics on the sphere obeys the Cugliandolo-Kurchan equations.

Lower Bound: Reaching Approximate Local Maxima

Definition

$x \in \mathcal{S}_N$ is an ε-**approximate local maximum** if both:

1. $\|\nabla_{\mathrm{sp}} H_N(x)\| \le \varepsilon\sqrt{N}$.
2. $\lambda_{\varepsilon N}\big(\nabla^2_{\mathrm{sp}} H_N(x)\big) \le \varepsilon$.

If 1 holds but 2 doesn't, then x is an ε-**approximate saddle**.

Lower Bound: Reaching Approximate Local Maxima

Definition

$x \in \mathcal{S}_N$ is an **ε-approximate local maximum** if both:

1. $\|\nabla_{\mathrm{sp}} H_N(x)\| \leq \varepsilon\sqrt{N}$.
2. $\lambda_{\varepsilon N}\big(\nabla^2_{\mathrm{sp}} H_N(x)\big) \leq \varepsilon$.

If 1 holds but 2 doesn't, then x is an **ε-approximate saddle**.

Proposition (Specific to Pure p-Spin Models)

With probability $1 - e^{-cN}$, all ε-approximate local maxima satisfy $H_N(x)/N \geq E_\infty - \delta(\varepsilon)$.

Lower Bound: Reaching Approximate Local Maxima

Definition

$x \in \mathcal{S}_N$ is an ε-**approximate local maximum** if both:

1. $\|\nabla_{\mathrm{sp}} H_N(x)\| \leq \varepsilon\sqrt{N}$.
2. $\lambda_{\varepsilon N}\big(\nabla^2_{\mathrm{sp}} H_N(x)\big) \leq \varepsilon$.

If 1 holds but 2 doesn't, then x is an ε-**approximate saddle**.

Proposition (Specific to Pure p-Spin Models)

With probability $1 - e^{-cN}$, all ε-approximate local maxima satisfy $H_N(x)/N \geq E_\infty - \delta(\varepsilon)$.

Theorem (General; Only Uses 3rd-Order Smoothness of H_N)

Suppose all ε-approximate local maxima satisfy $H_N(x)/N \geq E_*(\varepsilon)$.
Then for large T_0, β and $x_0 \in \mathcal{S}_N$ possibly depending on H_N:

$$\mathbb{P}\left[\inf_{t\in[T_0, T_0+e^{cN}]} H_N(x_t)/N \geq E_*(\varepsilon) - \delta(\varepsilon)\right] \geq 1 - e^{-cN}.$$

Energy Gain While Below $E_*(\varepsilon)$

Lemma

Fix $\beta \geq \beta_0(\varepsilon)$. *For any stopping time* τ, *with probability* $1 - e^{-cN}$ *conditioned on* $\mathcal{F}_\tau$*:*

1. *If* $\|\nabla_{\mathrm{sp}} H_N(\boldsymbol{x}_\tau)\| \geq C\sqrt{N}/\beta$*:*
$$\frac{H_N(\boldsymbol{x}_{\tau+\beta^{-3}}) - H_N(\boldsymbol{x}_\tau)}{N} \geq \beta^{-3}/C.$$
2. *If* $\|\nabla_{\mathrm{sp}} H_N(\boldsymbol{x}_\tau)\| \leq C\sqrt{N}/\beta$ ***and*** $\lambda_{\varepsilon N}\big(\nabla^2_{\mathrm{sp}} H_N(\boldsymbol{x}_\tau)\big) \geq \varepsilon$*:*
$$\frac{H_N(\boldsymbol{x}_{\tau+\overline{C}(\varepsilon)\beta^{-1}}) - H_N(\boldsymbol{x}_\tau)}{N} \geq \beta^{-1}.$$

Energy Gain While Below $E_*(\varepsilon)$

Lemma

Fix $\beta \geq \beta_0(\varepsilon)$. For any stopping time τ, with probability $1 - e^{-cN}$ conditioned on $\mathcal{F}_\tau$:

1. *If $\|\nabla_{\mathrm{sp}} H_N(\boldsymbol{x}_\tau)\| \geq C\sqrt{N}/\beta$:*

$$\frac{H_N(\boldsymbol{x}_{\tau+\beta^{-3}}) - H_N(\boldsymbol{x}_\tau)}{N} \geq \beta^{-3}/C.$$

2. *If $\|\nabla_{\mathrm{sp}} H_N(\boldsymbol{x}_\tau)\| \leq C\sqrt{N}/\beta$ **and** $\lambda_{\varepsilon N}(\nabla^2_{\mathrm{sp}} H_N(\boldsymbol{x}_\tau)) \geq \varepsilon$:*

$$\frac{H_N(\boldsymbol{x}_{\tau+\overline{C}(\varepsilon)\beta^{-1}}) - H_N(\boldsymbol{x}_\tau)}{N} \geq \beta^{-1}.$$

Lemma says $H_N(\boldsymbol{x}_t)$ increases whenever $H_N(\boldsymbol{x}_t)/N \leq E_*(\varepsilon)$. This yields the lower bound.

Energy Gain While Below $E_*(\varepsilon)$

Lemma

Fix $\beta \geq \beta_0(\varepsilon)$. For any stopping time τ, with probability $1 - e^{-cN}$ conditioned on $\mathcal{F}_\tau$:

1. *If $\|\nabla_{\mathrm{sp}} H_N(\boldsymbol{x}_\tau)\| \geq C\sqrt{N}/\beta$:*
$$\frac{H_N(\boldsymbol{x}_{\tau+\beta^{-3}}) - H_N(\boldsymbol{x}_\tau)}{N} \geq \beta^{-3}/C.$$
2. *If $\|\nabla_{\mathrm{sp}} H_N(\boldsymbol{x}_\tau)\| \leq C\sqrt{N}/\beta$ **and** $\lambda_{\varepsilon N}(\nabla^2_{\mathrm{sp}} H_N(\boldsymbol{x}_\tau)) \geq \varepsilon$:*
$$\frac{H_N(\boldsymbol{x}_{\tau+\overline{C}(\varepsilon)\beta^{-1}}) - H_N(\boldsymbol{x}_\tau)}{N} \geq \beta^{-1}.$$

Lemma says $H_N(\boldsymbol{x}_t)$ increases whenever $H_N(\boldsymbol{x}_t)/N \leq E_*(\varepsilon)$. This yields the lower bound.

1 is easy: energy gain from the gradient overwhelms Ito term.
2 is the key. **Langevin gains energy from approximate saddles.**

Gaining Energy From Approximate Saddles

Suppose $\|\nabla_{sp} H_N(\boldsymbol{x}_\tau)\| \le C\sqrt{N}/\beta$ and $\lambda_{\varepsilon N}\left(\nabla^2_{sp} H_N(\boldsymbol{x}_\tau)\right) \ge \varepsilon$, where $\beta \gg 1/\varepsilon$.

Wishful thinking: suppose H_N was **quadratic** and $\mathcal{S}_N$ was $\mathbb{R}^{N-1}$.

Gaining Energy From Approximate Saddles

Suppose $\|\nabla_{\mathrm{sp}} H_N(\boldsymbol{x}_\tau)\| \le C\sqrt{N}/\beta$ and $\lambda_{\varepsilon N}\big(\nabla^2_{\mathrm{sp}} H_N(\boldsymbol{x}_\tau)\big) \ge \varepsilon$, where $\beta \gg 1/\varepsilon$.

Wishful thinking: suppose H_N was **quadratic** and $\mathcal{S}_N$ was $\mathbb{R}^{N-1}$.

Then $\boldsymbol{x}_{\tau+t}$ would be a multi-dimensional OU process. Easy to analyze!

- Exponentially fast energy gain from positive eigenvalues.
- Rapid equilibration for negative eigenvalues.
- Altogether, energy gain of $N\beta^{-1}$ after time $\overline{C}(\varepsilon)\beta^{-1}$.
- (But, energy can initially drop. This is a problem for differential inequalities.)

Gaining Energy From Approximate Saddles

Suppose $\|\nabla_{\mathrm{sp}} H_N(\boldsymbol{x}_\tau)\| \le C\sqrt{N}/\beta$ and $\lambda_{\varepsilon N}\big(\nabla^2_{\mathrm{sp}} H_N(\boldsymbol{x}_\tau)\big) \ge \varepsilon$, where $\beta \gg 1/\varepsilon$.

Wishful thinking: suppose H_N was **quadratic** and $\mathcal{S}_N$ was $\mathbb{R}^{N-1}$.

Then $\boldsymbol{x}_{\tau+t}$ would be a multi-dimensional OU process. Easy to analyze!

- Exponentially fast energy gain from positive eigenvalues.
- Rapid equilibration for negative eigenvalues.
- Altogether, energy gain of $N\beta^{-1}$ after time $\overline{C}(\varepsilon)\beta^{-1}$.
- (But, energy can initially drop. This is a problem for differential inequalities.)

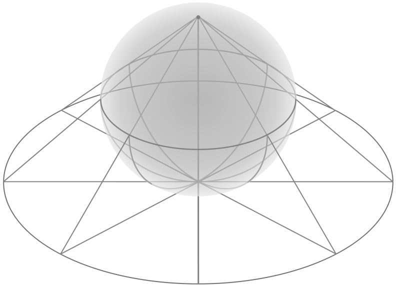
Key Idea: map $\mathcal{S}_N$ to flat space and **Taylor expand the dynamics to an OU process**.

- Approximation error is $O(\sqrt{N}/\beta)$ in distance.
- Leads to energy error $O(N\beta^{-3/2}) \ll N\beta^{-1}$ because $\|\nabla_{\mathrm{sp}} H_N(\boldsymbol{x}_\tau)\|$ is small.

Taylor Expansion of Langevin Dynamics

Use stereographic projection centered at $-\boldsymbol{x}_\tau$:

$$\Gamma_{\boldsymbol{x}_\tau} : \mathcal{S}_N \backslash \{-\boldsymbol{x}_\tau\} \to \mathbb{R}^{N-1},$$
$$\Gamma_{\boldsymbol{x}_\tau}(\boldsymbol{x}_\tau) = \vec{0},$$
$$\Gamma_{\boldsymbol{x}_\tau}(\boldsymbol{x}_t) = \widetilde{\boldsymbol{x}}_t.$$



Taylor Expansion of Langevin Dynamics

Use stereographic projection centered at $-\boldsymbol{x}_\tau$:

$$\Gamma_{\boldsymbol{x}_\tau} : \mathcal{S}_N \backslash \{-\boldsymbol{x}_\tau\} \to \mathbb{R}^{N-1},$$
$$\Gamma_{\boldsymbol{x}_\tau}(\boldsymbol{x}_\tau) = \vec{0},$$
$$\Gamma_{\boldsymbol{x}_\tau}(\boldsymbol{x}_t) = \widetilde{\boldsymbol{x}}_t.$$

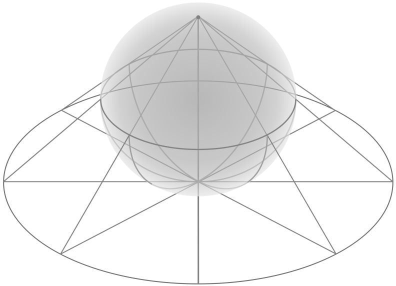


Image dynamics in $\mathbb{R}^{N-1}$:

$$\mathrm{d}\widetilde{\boldsymbol{x}}_t = \vec{b}_t(\widetilde{\boldsymbol{x}}_t)\,\mathrm{d}t + \sigma_t \mathrm{d}\boldsymbol{W}_t \approx \beta\,\Gamma'_{\boldsymbol{x}_\tau}(\boldsymbol{x}_t)\nabla_{\mathrm{sp}} H_N(\boldsymbol{x}_t)\,\mathrm{d}t + \mathrm{d}\boldsymbol{W}_t.$$

Bound OU process error via usual coupling. Need to be precise with powers of β.

A simplification is possible due to conformal flatness of the sphere. Means the matrix σ_t is scalar, so can reset to identity via time-change.

Quick Summary of Relevant Estimates

Modulo $O(1/N)$ Brownian terms, OU approximation is analyzed via differential inequalities.

- Movement is small on $O(1/\beta)$ time-scales due to small gradient:
$$\|\boldsymbol{x}_{t+\overline{C}\beta^{-1}} - \boldsymbol{x}_t\| \leq O_{\overline{C}}(\beta^{-1/2}\sqrt{N}).$$
- Since $\|\nabla H_N(\boldsymbol{x}_t)\| \leq C\beta^{-1}\sqrt{N}$ and H_N is smooth, get
$$\|\nabla H_N(\boldsymbol{x}_{t+\overline{C}\beta^{-1}})\| \leq O_{\overline{C}}(\beta^{-1/2}\sqrt{N}).$$
- Careful Grönwall implies the OU process $\widetilde{\boldsymbol{x}}_t$ stays close to $\boldsymbol{x}_t$:
$$\|\boldsymbol{x}_{t+\overline{C}\beta^{-1}} - \widetilde{\boldsymbol{x}}_{t+\overline{C}\beta^{-1}}\| \leq O_{\overline{C}}(\beta^{-1}\sqrt{N}).$$
- Combining the previous two,
$$\|H_N(\boldsymbol{x}_{t+\overline{C}\beta^{-1}}) - H_N(\widetilde{\boldsymbol{x}}_{t+\overline{C}\beta^{-1}})\| \leq O_{\overline{C}}(\beta^{-3/2}N).$$
- Since $\widetilde{\boldsymbol{x}}_t$ gains energy $\beta^{-1}N$, we conclude:
$$H_N(\boldsymbol{x}_{t+\overline{C}\beta^{-1}}) - H_N(\boldsymbol{x}_t) \approx H_N(\widetilde{\boldsymbol{x}}_{t+\overline{C}\beta^{-1}}) - H_N(\widetilde{\boldsymbol{x}}_t) \geq \beta^{-1}N.$$

Thank You!

Summary:

- Pure p-spin Hamiltonian:

$$H_N(\sigma) = N^{-(p-1)/2} \sum_{1 \le i_1, i_2, \ldots, i_p \le N} J_{i_1,\ldots,i_p} \sigma_{i_1} \ldots \sigma_{i_p}$$

- Main result: for spherical Langevin dynamics as $T, \beta \to \infty$:

$$H_N(x_T)/N \to E_\infty(p) \equiv 2\sqrt{\frac{p-1}{p}}.$$

- Upper bound holds for Lipschitz algorithms via branching overlap gap property.
- Lower bound: dynamics reach approximate local maxima.
  - Works even for worst-case $\boldsymbol{x}_0 = \boldsymbol{x}_0(H_N)$ and for $t \in [T_0, T_0 + e^{cN}]$.