Stochastic Localization with Non-Gaussian Tilts and Applications to Tensor Ising Models

Arianna Piana
joint work with Dan Mikulincer (University Washington)

February 25, 2025

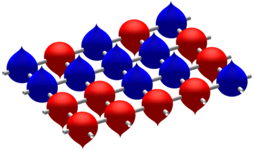
מכון ויצמן למדע
WEIZMANN INSTITUTE OF SCIENCE

Contents

Introduction

The Ising model

We consider the *Ising model* defined on the n-dim. hypercube $\mathcal{C}_n := \{-1, +1\}^n$.



Energy is given by the Hamiltonian

$$-\mathcal{H}(x) = \langle x, Jx \rangle + \langle h, x \rangle,$$

where J is the *interaction matrix* and h the *external field*.
The Gibbs measure is

$$\mu(x) = \frac{1}{Z_\beta} \exp(-\beta \mathcal{H}(x)).$$

Glauber Dynamics (GLD)

Markov chain $(X_t)_{t \geq 0}$ with $Law(X_t) \xrightarrow{t \to \infty} \mu$.

- Start from a configuration $x = (x_1, \ldots, x_n)$.
- Repeat the following at each step
  1. pick a coordinate $i \in [n]$ uniformly at random,
  2. update $(X_t)_i$ according to $\mu | (X_t)_{j \neq i}$.

GLD has an associated reversible transition kernel P_{GLD}.
The mixing time is defined as

$$t_{\mathrm{mix}}(\varepsilon) := \min \left\{ t \geq 0 : \; \forall x \in \mathcal{C}_n \; \| P^t_{\mathrm{GLD}}(x, \cdot) - \mu \|_{\mathrm{TV}} \leq \varepsilon \right\}.$$

Spectral gap and Poincaré inequality

Consider the eigenvalues of P_{GLD}: $1 = \lambda_1 > \lambda_2 \geq \cdots \geq \lambda_n > -1$. The *spectral gap* is defined as

$$\mathsf{gap} := \lambda_1 - \lambda_2 = 1 - \lambda_2.$$

It holds:

$$t_{\mathrm{mix}}(\varepsilon) \leq \left\lceil \frac{n}{\mathsf{gap}} \left(\log\left(\frac{1}{\min_{x \in \mathcal{C}_n} \mu(x)} \right) + \log\left(\frac{1}{2\varepsilon} \right) \right) \right\rceil .$$

Fix $\varphi : \mathcal{C}_n \to \mathbb{R}$. For GLD, having gap is equivalent to satisfy a Poincaré inequality:

$$\mathrm{Var}_\mu(\varphi) \leq C_{\mathrm{P}}(\mu)\mathcal{E}_\mu(\varphi),$$

where

$$\mathcal{E}_\mu(\varphi) := \frac{1}{2} \sum_{x,y \in \mathcal{C}_n} \big(\varphi(x) - \varphi(y)\big)^2 \mu(x) P_{\mathrm{GLD}}(x,y),$$

and

$$\mathrm{Var}_\mu(\varphi) = \int \varphi^2 \mathrm{d}\mu - \left(\int \varphi \mathrm{d}\mu \right)^2 .$$

Stochastic Localization and two questions

Idea of Stochastic Localization (SL)

SL is a measure-valued process constructed by Eldan in 2013.

Start from a measure on $\mathcal{C}_n$ (or $\mathbb{R}^n$) of the form $\frac{\mathrm{d}m(x)}{\mathrm{d}x} = \exp(-w(x))$.
Perturb it by a Gaussian tilt

$$\exp\left(-w(x)\right)\exp\left(\langle x, \mathsf{d}\rangle - \mathsf{s}\|x\|_2^2\right).$$

w must be "nice", for instance convex or quadratic.

SL definition

Fix μ on $\mathcal{C}_n$. Let $(B_t)_{t\geq 0}$ be a standard Brownian motion in $\mathbb{R}^n$.
SL is a process of the form $\mu_t = \mathsf{F}_t\mu$, where F_t is the solution of

$$\mathrm{d}\mathsf{F}_t(x) = \langle x - \mathsf{a}_t, \mathsf{C}_t \mathrm{d}B_t \rangle, \quad \mathsf{F}_0(x) = 1,$$

- a_t is the *barycenter*:

$$\mathsf{a}_t := \int x \mathrm{d}\mu_t(x),$$

- C_t is the *driving matrix*: encodes constraints on the system.

Role of a_t

a_t allows μ_t to be a probability measure. In fact:

$$\mathrm{d}\int \mu_t(x)\mathrm{d}x = \mathrm{d}\int \mathsf{F}_t(x)\mu(x)\mathrm{d}x = \left\langle \int x\mathsf{F}_t(x)\mu(x)\mathrm{d}x - \mathsf{a}_t, \mathsf{C}_t\mathrm{d}B_t \right\rangle = 0.$$

SL is a Gaussian tilt

Apply Itô's formula:

$$\mathrm{d}\log \mathsf{F}_t = \langle x - \mathsf{a}_t, \mathsf{C}_t \mathrm{d}B_t \rangle - \frac{1}{2}\|C_t(x - \mathsf{a}_t)\|_2^2 \mathrm{d}t.$$

We get:

$$\begin{aligned} \mu_t(x) &= \exp\left(\int_0^t \langle x - \mathsf{a}_s, \mathsf{C}_s \mathrm{d}B_s \rangle - \frac{1}{2}\int_0^t \|\mathsf{C}_s(x - \mathsf{a}_s)\|_2^2 \mathrm{d}s \right) \\ &\propto \exp\left(-\langle x, \mathsf{Q}_t x \rangle + \langle \mathsf{L}_t, x \rangle \right) \mu(x), \end{aligned}$$

where

- L_t is some adapted process in $\mathbb{R}^n$,
- $\mathsf{Q}_t = \frac{1}{2}\int_0^t \mathsf{C}_s^2 \mathrm{d}s$ is a matrix-valued process.

Can we localize μ with a non-Gaussian tilt?

Poincaré inequality via SL

Let τ be a suitable stopping time. Suppose we are able to prove

$$\mathsf{Var}_{\mu_\tau}(\varphi) \le C_{\mathsf{P}}(\mu_\tau)\mathcal{E}_{\mu_\tau}(\varphi)\,,$$

Let $\mathsf{M}_t := \int_{\mathcal{C}_n} \varphi \, \mathrm{d}\mu_t$. Then

$$\mathsf{Var}_{\mu}(\varphi) = \mathbb{E}\big[[\mathsf{M}]_t\big] + \mathbb{E}\big[\mathsf{Var}_{\mu_t}(\varphi)\big].$$

Suppose $[\mathsf{M}]_t = 0$ a.s. Then

$$\mathsf{Var}_{\mu}(\varphi) = \mathbb{E}[\mathsf{Var}_{\mu_\tau}(\varphi)] \le C_{\mathsf{P}}(\mu_\tau)\mathbb{E}[\mathcal{E}_{\mu_\tau}(\varphi)] \le C_{\mathsf{P}}(\mu_\tau)\mathbb{E}[\mathcal{E}_{\mu}(\varphi)]\,.$$

Eldan-Koehler-Zeitouni work: role of C_t

Consider $\mu(x) \propto \exp(\langle x, Jx \rangle)$. C_t encodes two constraints

1. $\mathbb{E}_{\mu_t}[\varphi]$ constant in time $\Rightarrow$ $\mathrm{Var}_{\mu_t}(\varphi)$ is a martingale.
2. $J_t = J - \int_0^t \mathsf{C}_s^2 ds$ decreasing, if $\mathsf{rank}(J_t) \geq 2$.

Define $\tau = \min\{t : \mathsf{rank}(J_t) = 1\}$. SL decomposes μ into a mixture of measures of the form

$$w_{u,v}(x) \propto \exp\left(\langle u, x \rangle^2 + \langle v, x \rangle\right).$$

It holds:

1. $\|u\|^2 = \|J_t\|_{\mathrm{op}} \leq \|J\|_{\mathrm{op}}$
2. $C_{\mathrm{P}}(w_{u,v}) \leq (1 - 2\|u\|_2^2)^{-1}$

If $\|J\|_{\mathrm{op}} < \frac{1}{2} \Rightarrow C_{\mathrm{P}}(\mu) \leq (1 - 2\|J\|_{\mathrm{op}})^{-1}$.

What happens if J is a fourth-order tensor?

Naïve answers

About fourth-order tensors

- $T \in (\mathbb{R}^n)^{\otimes 4}$.
- $T : \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^n) \to \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^n)$. "$n^2 \times n^2$ matrix"
- T symmetric: $T_{i_1,i_2,i_3,i_4} = T_{i_{\sigma(1)},i_{\sigma(2)},i_{\sigma(3)},i_{\sigma(4)}}$ for any permutation σ.
- Injective norm: $\|T\|_{\text{inj}} = \sup_{x\in\mathbb{S}^{n-1}} |T(x,x,x,x)|$.

Naïve approach

Let $T \in (\mathbb{R}^n)^{\otimes 4}$. Consider $\nu(x) \propto \exp(T(x))$ on $\mathcal{C}_n$.
Define the following version of SL: $\nu_t = \mathtt{F}_t \nu$, where $\mathtt{F}_t$ solves

$$\mathrm{d}\mathtt{F}_t(x) = \langle x^{\otimes 2} - \mathtt{a}_t, \mathtt{M}_t \mathrm{d}\mathtt{W}_t \rangle_{\mathrm{HS}}, \quad \mathtt{F}_0(x) = 1,$$

with

$$\mathtt{a}_t = \int x^{\otimes 2} \mathrm{d}\nu_t(x),$$

and where $\mathtt{W}_t$ is a Dyson Brownian motion, and $\mathtt{M}_t$ is a fourth-order tensor operating on matrices.

Problems

Apply Itô's formula:

$$\nu_t(x) = \exp\left(\int_0^t \langle x^{\otimes 2} - \mathtt{a}_s, \mathtt{M}_s \mathrm{d}\mathtt{W}_s\rangle_{\mathrm{HS}} - \frac{1}{2}\int_0^t \|\mathtt{M}_s(x^{\otimes 2} - \mathtt{a}_s)\|_{\mathrm{HS}}^2 \mathrm{d}s\right)\mu(x)$$
$$\propto \exp\left(\left\langle x^{\otimes 2}, \left(T - \frac{1}{2}\int_0^t \mathtt{M}_s^2 \mathrm{d}s\right) x^{\otimes 2}\right\rangle_{\mathrm{HS}} + \left\langle \mathtt{L}_t, x^{\otimes 2}\right\rangle_{\mathrm{HS}}\right),$$

where

$$\mathtt{L}_t = \int_0^t \left(\mathtt{M}_s \mathrm{d}\mathtt{W}_s - \mathtt{M}_s^2 \mathtt{a}_s \mathrm{d}s\right).$$

Problems:

1. $\left\langle \mathtt{L}_t, x^{\otimes 2}\right\rangle_{\mathrm{HS}}$ is not linear in x. Spectral gap can deteriorate.
2. We treat T as a *matrix* $\Rightarrow$ $T - \frac{1}{2}\int_0^t \mathtt{M}_s^2 \mathrm{d}s$ collapses to a rank 1 matrix.

Tensor Stochastic Localization

Solution to problem 1

Problem: $\mathtt{L}_t = \int_0^t \left(\mathtt{M}_s \mathrm{d}\mathtt{W}_s - \mathtt{M}_s^2 \mathtt{a}_s \mathrm{d}s\right)$ not linear in x.
Solution: eliminate it. How?

1. Remove barycenter. Choose v_t such that $\mathtt{L}_t$ is arbitrarily small.
2. Encode in $\mathtt{M}_t$ a new constraint to fix $\int 1 \mathrm{d}\nu_t$

$\Rightarrow$ $\mathtt{M}_t$ encodes mass preservation and variance martingale.

$$\mathrm{d}F_t(x) = \langle x^{\otimes 2} - v_t, C_t \mathrm{d}W_t \rangle_{\mathrm{HS}}, \quad F_0(x) = 1,$$

with $\mu_t = F_t \mu$ and where

- C_t is a 4-tensor operating on matrices such that $C_t \mathrm{d}W_t$ is a matrix.
- v_t satisfies:
  Proposition $\exists$ adapted drift process $v_t^\delta \in \mathrm{Mat}(\mathbb{R}^n, \mathbb{R}^n)$, st if X_t^δ satisfies

$$\mathrm{d}X_t^\delta = C_t \mathrm{d}W_t - C_t^2 v_t^\delta \mathrm{d}t, \quad X_0^\delta = 0,$$

then $\sup_{t \geq 0} \|X_t^\delta\|_{\mathrm{HS}} \leq \delta$ a.s.

Solution to problem 2

Define

$$T_t := T - \frac{1}{2}\int_0^t C_s^2 \mathrm{d}s \quad \text{and} \quad \tau = \inf\{t \geq 0 : \mathrm{rank}(T_t) \leq 2\}.$$

We can decompose the measure

$$\begin{aligned}\mu(x) &= \mathbb{E}[\mu_\tau] \\ &\propto \mathbb{E}[\exp(\langle x^{\otimes 2} \otimes x^{\otimes 2}, M_1 \otimes M_1 + M_2 \otimes M_2\rangle_{\mathrm{HS}})] \\ &= \mathbb{E}[\exp(\langle x, M_1 x\rangle^2 + \langle x, M_2 x\rangle^2)]\end{aligned}$$

with $\|M_1\|_{\mathrm{OP}}^2, \|M_2\|_{\mathrm{OP}}^2 \leq \|T\|_{\mathrm{inj}}$.

Decomposition of the measure $\mu(x) \propto \exp(T(x))$

$$\exp\left(T(x)\right)$$

$$\exp\left(\langle x, M_1 x\rangle^2\right) \quad \exp\left(\langle x, M_2 x\rangle^2\right)$$

$$\exp\Big(\left(\langle v_1, x\rangle^2 + \langle v_2, x\rangle^2\right) \quad \langle x, M_1 x\rangle\Big) \quad \exp\Big(\langle x, M_2 x\rangle \quad \left(\langle v_3, x\rangle^2 + \langle v_4, x\rangle^2\right)\Big)$$

$$\exp\Big(\left(\langle v_1, x\rangle^2 + \langle v_2, x\rangle^2\right) \quad \left(\langle u_1, x\rangle^2 + \langle u_2, x\rangle^2\right)\Big) \quad \exp\Big(\left(\langle u_3, x\rangle^2 + \langle u_4, x\rangle^2\right) \quad \left(\langle v_3, x\rangle^2 + \langle v_4, x\rangle^2\right)\Big)$$

Decomposition theorem

Let $\varphi : \mathcal{C}_n \to \mathbb{R}$ test function, let $\mu \propto \exp(T(x))$. Then, $\exists$ decomposition

$$\mu = \int \mu_{\bar{u},\bar{v},\bar{w},\ell}\,\eta(\mathrm{d}\bar{u}, \mathrm{d}\bar{v}, \mathrm{d}\bar{w}, \mathrm{d}\ell),$$

where $\bar{u} = \{u_1, u_2, u_3, u_4\}$, $\bar{v} = \{v_1, v_2, v_3, v_4\}$, and $\bar{w} = \{w_1, w_2\}$, with $u_i, v_i, w_j, \ell \in \mathbb{R}^n$ for $i \in [4]$ and $j = 1, 2$, and $\mu_{\bar{u},\bar{v},\bar{w},\ell}$ are probability measures

$$\mu_{\bar{u},\bar{v},\bar{w},\ell}(x) \propto \exp\left(\sum_{i,j=1}^{2} \langle u_i, x\rangle^2 \langle v_j, x\rangle^2 + \sum_{i,j=3}^{4} \langle u_i, x\rangle^2 \langle v_j, x\rangle^2 + \sum_{i=1}^{2} \langle x, w_i\rangle^2 + \langle \ell, x\rangle\right).$$

Properties:

- For $i \in [4]$, $\|u_i\|_2^2, \|v_i\|_2^2 \leq 2\sqrt{\|T\|_{\mathrm{inj}}}$, $\|w_i\|_2^2 \leq 4n\|T\|_{\mathrm{inj}}$.
- Variance decomposition: $\mathrm{Var}_\mu(\varphi) = \int \mathrm{Var}_{\mu_{\bar{u},\bar{v},\bar{w},\ell}}(\varphi)\eta(\mathrm{d}\bar{u}, \mathrm{d}\bar{v}, \mathrm{d}\bar{w}, \mathrm{d}\ell)$.

Spectral gap for tensor Ising model

Theorem Let μ be a measure on $\mathcal{C}_n$ given by $\mu(x) \propto \exp(T(x))$. Assume $\|T\|_{\text{inj}} \leq \frac{1}{336n}$. Then

$$C_{\mathrm{P}}(\mu) \leq \frac{1}{1 - 336n\|T\|_{\text{inj}}}.$$

Corollary T has independent $\mathcal{N}(0, n^{-3})$ off-diagonal entries and 0 everywhere else. Let μ be a measure on $\mathcal{C}_n$ given by $\mu(x) \propto \exp(\beta T(x))$. If $\beta \lesssim \frac{1}{1205.568}$, then

$$C_{\mathrm{P}}(\mu) \leq \frac{1}{1 - 1205.568\beta}.$$

Another approach

Trick: Normalized variance

Let ν be an unormalized measure. Consider normalized variance

$$\overline{\mathsf{Var}}_\nu(\varphi) = \mathbb{E}_\nu\left[\varphi^2\right] - \frac{\mathbb{E}_\nu\left[\varphi\right]^2}{\mathbb{E}_\nu\left[1\right]}.$$

Define $\tilde{\nu} = \frac{\nu}{\mathbb{E}_\nu[1]}$. Then

- $\overline{\mathsf{Var}}_\nu(\varphi) = \mathbb{E}_\nu[1]\mathsf{Var}_{\tilde{\nu}}(\varphi)$.
- If $C_{\mathsf{P}}(\tilde{\nu}) < \infty$. Then, $\forall \varphi : \mathcal{C}_n \to \mathbb{R}$

$$\overline{\mathsf{Var}}_\nu(\varphi) \leq C_{\mathsf{P}}(\tilde{\nu})\mathbb{E}_\nu[1]\mathcal{E}_{\tilde{\nu}}(\varphi) = C_{\mathsf{P}}(\tilde{\nu})\mathcal{E}_\nu(\varphi).$$

Adjustments to TSL

We allow $\mathbb{E}_{\mu_t}[1]$ to vary.
$\Rightarrow$ C_t has one constraint: $\frac{\mathbb{E}_{\nu_t}[\varphi]}{\mathbb{E}_{\nu_t}[1]}$ is constant. This allows to preserve $\overline{\mathrm{Var}}_{\nu_t}$.

$$\exp\left(T(x)\right)$$

$$\exp\left(\langle x, Mx\rangle \; \langle x, Mx\rangle + \langle x, Rx\rangle\right)$$

$$\exp\left(\langle x, Mx\rangle \; \langle v, x\rangle^2 + \langle x, Rx\rangle\right)$$

$$\exp\left(\langle u, x\rangle^2 \; \langle v, x\rangle^2 + \langle x, Rx\rangle\right)$$

$$\exp\left(\langle u, x\rangle^2 \; \langle v, x\rangle^2 + \langle x, w\rangle^2 + \langle \ell, x\rangle\right)$$

Decomposition theorem bis

Let $\varphi : \mathcal{C}_n \to \mathbb{R}$ and $\mu \propto \exp(T(x))$. Then, $\forall \delta > 0$, $\exists$ decomposition

$$\mu = \int \mu_{u,v,w,\ell,\psi} \eta(\mathrm{d}u, \mathrm{d}v, \mathrm{d}w, \mathrm{d}\ell, \mathrm{d}\psi),$$

where with $u, v, w, \ell \in \mathbb{R}^n$ and $\psi : \mathcal{C}_n \to \mathbb{R}$, and $\mu_{u,v,w,\ell,\psi}$ is a non-negative measure of the form

$$\mu_{u,v,w,\ell,\psi}(x) \propto \exp\left(\langle u, x\rangle^2 \langle v, x\rangle^2 + \langle x, w\rangle^2 + \langle \ell, x\rangle + \psi(x)\right).$$

Properties:

- $\|u\|_2^2, \|v\|_2^2 \leq 2\sqrt{\|T\|_{\mathrm{inj}}}$, $\|w\|_2^2 \leq 4n\|T\|_{\mathrm{inj}}$.
- $\max_{x \in \mathcal{C}_n} |\psi(x)| \leq \delta$.
- Variance decomposition:
  $\mathrm{Var}_\mu(\varphi) \leq \int \overline{\mathrm{Var}}_{\mu_{u,v,w,\ell,\psi}}(\varphi) \eta(\mathrm{d}u, \mathrm{d}v, \mathrm{d}w, \mathrm{d}\ell, \mathrm{d}\psi) + \delta.$

Spectral gap for tensor Ising model bis

Theorem Let μ be a measure on $\mathcal{C}_n$ given by $\mu(x) \propto \exp(T(x))$. Assume $\|T\|_{\text{inj}} \leq \frac{1}{56n}$. Then

$$C_{\mathrm{P}}(\mu) \leq \frac{1}{1 - 56n\|T\|_{\text{inj}}}.$$

Corollary T has independent $\mathcal{N}(0, n^{-3})$ off-diagonal entries and 0 everywhere else. Let μ be a measure on $\mathcal{C}_n$ given by $\mu(x) \propto \exp(\beta T(x))$. If $\beta \lesssim \frac{1}{200.928}$, then

$$C_{\mathrm{P}}(\mu) \leq \frac{1}{1 - 200.928\beta}.$$

Thank you!