Poincaré and logarithmic Sobolev inequalities by decomposition of the energy landscape

André Schlichting (University Bonn), Georg Menz (MPI Leipzig)

Introduction

Hamiltonian $H : \mathbb{R}^n \to \mathbb{R}$, Morse, growth at ∞

Glauber dynamic $dX_t = -\nabla H(X_t)\,dt + \sqrt{2\varepsilon}\,dB_t$ at temperature $\varepsilon \ll 1$

Gibbs measure $\mu(dx) = \frac{1}{Z_\mu} e^{-\frac{H}{\varepsilon}}\,dx$ with $Z_\mu = \int e^{-\frac{H}{\varepsilon}}\,dx$.

Generator If law $X_0 = f_0\mu$, then law $X_t = f_t\mu$, where f_t solves

Fokker-Planck equation $\partial_t f_t = L f_t := \varepsilon\Delta f_t - \nabla H \cdot \nabla f_t$.

Definition *PI and LSI*

The measure μ satisfies the Poincaré inequality PI(ϱ) with constant $\varrho > 0$ if

$$\forall f : \mathbb{R}^n \to \mathbb{R} \qquad \mathrm{var}_\mu(f) := \int \left(f - \int f d\mu\right)^2 d\mu \le \frac{1}{\varrho}\int |\nabla f|^2\,d\mu. \qquad \mathrm{PI}(\varrho)$$

and the logarithmic Sobolev inequality LSI(α) with constant $\alpha > 0$ if

$$\forall f : \mathbb{R}^n \to \mathbb{R} \qquad \mathrm{Ent}_\mu(f^2) := \int f^2 \log \frac{f^2}{\int f^2 d\mu} d\mu \le \frac{2}{\alpha}\int |\nabla f|^2\,d\mu. \qquad \mathrm{LSI}(\alpha)$$

Basic properties: LSI(α) implies PI(ϱ) for μ with $\varrho \ge \alpha$.

PI(ϱ) and LSI(α) imply exponential convergence to invariant measure μ:

$\mathrm{PI}(\varrho) \Rightarrow \mathrm{var}_\mu(f_t) \le \mathrm{var}_\mu(f_0)e^{-2\varrho\varepsilon t}$ and $\mathrm{LSI}(\alpha) \Rightarrow \mathrm{Ent}_\mu(f_t) \le \mathrm{Ent}_\mu(f_0)e^{-2\alpha\varepsilon t}$.

Partitions and split of variance

- Define for every local minima m_i of H its domain of attraction by

$$\Omega_i := \left\{ y \in \mathbb{R}^n : \lim_{t\to\infty} y_t = x_i, \dot{y}_t = -\nabla H(y_t), y_0 = y \right\}.$$

- $\{\Omega_0, \Omega_1\}$ is partition of $\mathbb{R}^n = \overline{\Omega}_0 \cup \overline{\Omega}_1$ and $\Omega_0 \cap \Omega_1 = \emptyset$.
- Introduce the restricted measures μ_0 and μ_1

$$\mu_i(dx) := \mu(dx)\llcorner\Omega_i = \frac{1}{Z_i Z_\mu} e^{-\frac{H}{\varepsilon}}\Big|_{\Omega_i} dx, \qquad Z_i := \mu(\Omega_i) \qquad i = 0, 1.$$

$\Rightarrow$ *mixture representation:* $\mu = Z_0\mu_0 + Z_1\mu_1$.

Motivated from [CM10] (discrete setting) and extended for LSI:

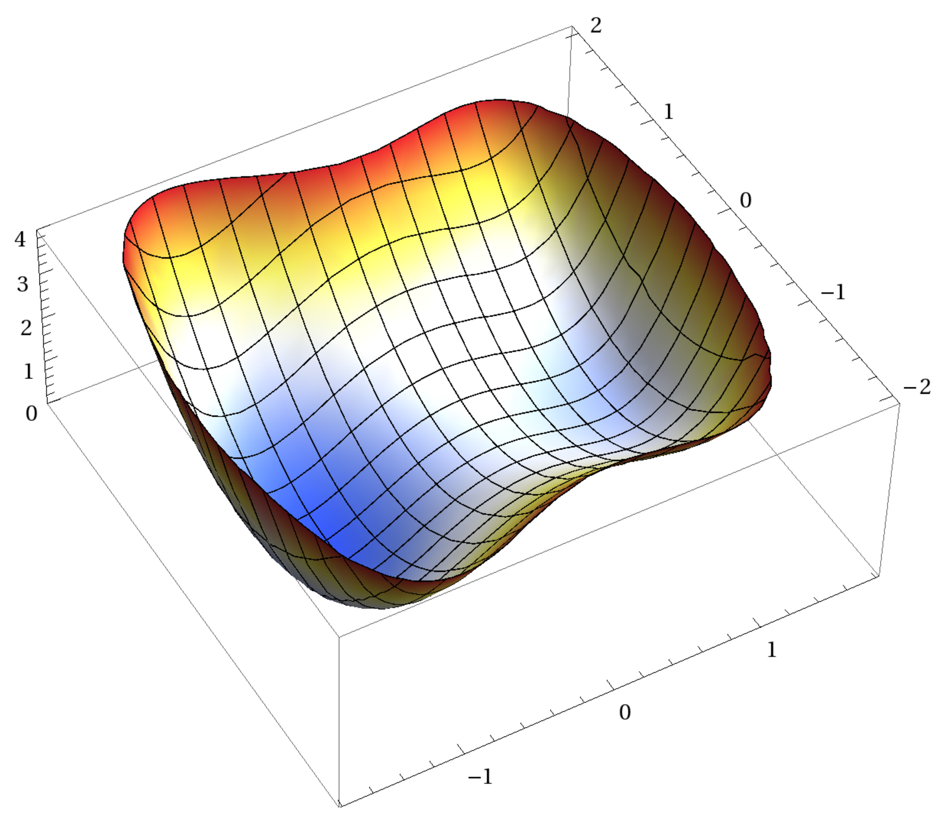
Lemma *Splitting of the variance and entropy*

$$\mathrm{var}_\mu(f) = \underbrace{Z_0\,\mathrm{var}_{\mu_0}(f) + Z_1\,\mathrm{var}_{\mu_1}(f)}_{\text{local variances}} + Z_0 Z_1 \underbrace{\left(\mathbb{E}_{\mu_0}(f) - \mathbb{E}_{\mu_1}(f)\right)^2}_{\text{mean difference}}$$

$$\mathrm{Ent}_\mu(f) \le Z_0\,\mathrm{Ent}_{\mu_0}(f) + Z_1\,\mathrm{Ent}_{\mu_1}(f) + \frac{Z_0 Z_1}{\Lambda(Z_0, Z_1)}\left(\mathrm{var}_{\mu_0}(f) + \mathrm{var}_{\mu_1}(f) + \left(\mathbb{E}_{\mu_0}(f) - \mathbb{E}_{\mu_1}(f)\right)^2\right),$$

where $\Lambda(Z_0, Z_1) := \frac{Z_0 - Z_1}{\log Z_0 - \log Z_1}$ is the logarithmic mean.

Example of an energy landscape



A *generic* double well

Main results

Theorem 1 *Local* PI *and* LSI

The measures μ_0 and μ_1 satisfy PI(ϱ_{loc}) and LSI(α_{loc}) with

$$\varrho_{\mathrm{loc}}^{-1} = O(\varepsilon) \quad \text{and} \quad \alpha_{\mathrm{loc}}^{-1} = O(1).$$

Proof: Use [CGW10] and explicitly construct a Lyapunov function.

Theorem 2 *Mean difference estimate*

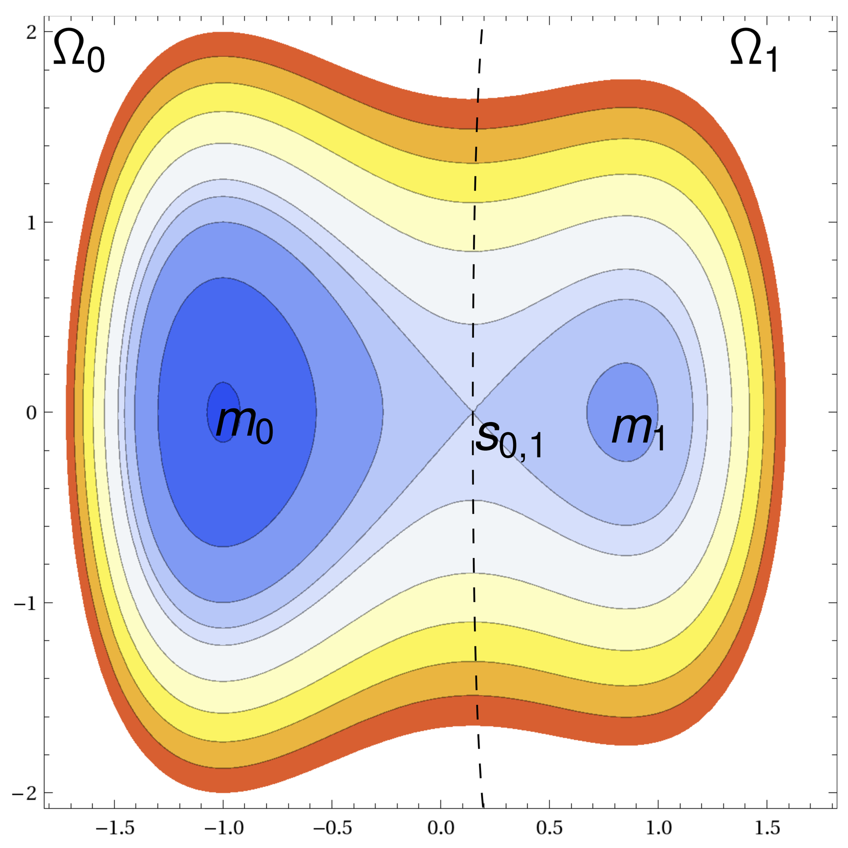
Notation: $\lesssim$ is $\le$ up to factor $1 + o(1)$ as $\varepsilon \to 0$

$$\left(\mathbb{E}_{\mu_0}(f) - \mathbb{E}_{\mu_1}(f)\right)^2 \lesssim \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}} \frac{2\pi\varepsilon\sqrt{|\det \nabla^2 H(s_{0,1})|}}{|\lambda^-(\nabla^2 H(s_{0,1}))|} e^{\varepsilon^{-1}H(s_{0,1})} \int |\nabla f|^2\,d\mu,$$

where $\lambda^-(\nabla^2 H(s_{0,1}))$ is the negative eigenvalue of $\nabla^2 H(s_{0,1})$.

Proof: Coupling and transport argument

Example of the according contour



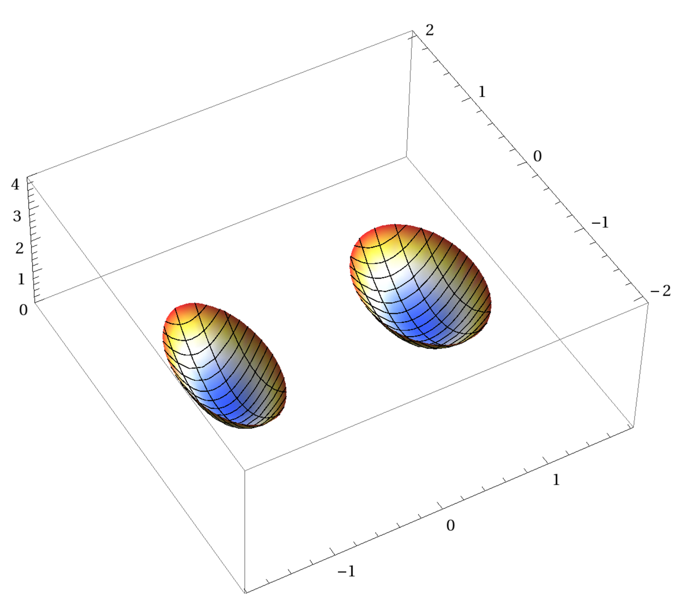
Ad Theorem 2: Coupling measures

Truncated Gaussian measures ν_0 and ν_1:

$\nu_i \sim \mathcal{N}(m_i, \Sigma_i)\llcorner B_{\sqrt{\varepsilon}}$ where $\Sigma_i^{-1} = \nabla^2 H(x_i)$.

Use ν_0 and ν_i as coupling between μ_0 and μ_1:

$$\left(\mathbb{E}_{\mu_0}(f) - \mathbb{E}_{\mu_1}(f)\right)^2 \lesssim \underbrace{\left(\mathbb{E}_{\nu_0}(f) - \mathbb{E}_{\nu_1}(f)\right)^2}_{\text{weighted transport cost}} + \sum_{i=\{0,1\}} \underbrace{\left(\mathbb{E}_{\mu_i}(f) - \mathbb{E}_{\nu_i}(f)\right)^2}_{\text{approximation bound}}.$$



Approximation bound follows from Theorem 1.

Ad Theorem 2: Weighted transport cost

Definition *weighted transport cost*

For $\nu_0, \nu_1 \ll \mu$ the weighted transport cost $\mathcal{T}_\mu(\nu_0, \nu_1)$ is defined by

$$\mathcal{T}_\mu(\nu_0, \nu_1) = \inf_{\{\Phi_s\}_{s\in[0,1]}} \int \left| \int_0^1 \dot{\Phi}_s \circ \Phi_s^{-1} \frac{d\nu_s}{d\mu} ds \right|^2 d\mu,$$

where $\{\Phi_s\}_{s\in[0,1]}$ is a transport interpolation between ν_0 and ν_1 and $\nu_s = (\Phi_s)_\sharp \nu_0$ the pushforward of ν_0 under Φ_s.

Mean difference can be estimated by $\mathcal{T}_\mu(\nu_0, \nu_1)$:

$$\left(\mathbb{E}_{\nu_0}(f) - \mathbb{E}_{\nu_1}(f)\right)^2 = \left(\int\int_0^1 \frac{df(\Phi_s(x))}{ds} ds\,d\nu_0\right)^2 = \left(\int_0^1 \int \left\langle \nabla f \circ \Phi_s, \dot{\Phi}_s \right\rangle d\nu_0\,ds\right)^2$$

$$= \left(\int \left\langle \nabla f, \int_0^1 \dot{\Phi}_s \circ \Phi_s^{-1} \frac{d\nu_s}{d\mu} ds \right\rangle d\mu\right)^2 \le \mathcal{T}_\mu(\nu_0, \nu_1) \int |\nabla f|^2\,d\mu.$$

Ansatz: explicit affine transport interpolation $\{\Phi_s\}_{s\in[0,1]}$ proofs Theorem 2.

Relation to $H^{-1}(\mu)$ norm: Note $\mathbb{E}_{\nu_0}(f) - \mathbb{E}_{\nu_1}(f) = (f, \nu_0 - \nu_1)_{L^2(\mu)}$

Interpolation estimate $(f, \nu_0 - \nu_1)^2_{L^2(\mu)} \le \|\nu_0 - \nu_1\|^2_{H^{-1}(\mu)}\,\|f\|^2_{H^1(\mu)}$.

Indeed, it holds $\mathcal{T}_\mu(\nu_0, \nu_1) = \|\nu_0 - \nu_1\|_{H^{-1}(\mu)}$.

$\Rightarrow$ *Dynamic representation of $H^{-1}(\mu)$ norm allows sharp estimates.*

This representation generalizes to Wasserstein type distances [DNG09].

Discussion of main results

Theorem 1 + 2 reproofs the lower bound of [BGK05] for PI and extends to LSI

Corollary *Eyring-Kramers formula for* PI *and* LSI

The measure μ satisfies PI(ϱ) and LSI(α) with

$$\varrho \approx \frac{(2\pi\varepsilon)^{\frac{n}{2}}}{Z_0 Z_1 Z_\mu} \frac{|\lambda^-(\nabla^2 H(s_{0,1}))|}{2\pi\varepsilon\sqrt{|\det \nabla^2 H(s_{0,1})|}} e^{-\frac{H(s_{0,1}) - H(x_1)}{\varepsilon}} \quad \text{and} \quad \alpha \gtrsim \Lambda(Z_0, Z_1)\,\varrho.$$

In special cases, an asymptotic evaluation of the factor $\frac{(2\pi\varepsilon)^{\frac{n}{2}}}{Z_0 Z_1 Z_\mu}$ leads to:

$H(m_0) < H(m_1)$: $\quad \frac{\varrho}{\alpha} \lesssim O(\varepsilon^{-1})$

$H(m_0) = H(m_1)$: $\quad \frac{\varrho}{\alpha} \lesssim \frac{\frac{\kappa_0 + \kappa_1}{2}}{\Lambda(\kappa_0, \kappa_1)} = \frac{\text{arithmetic mean}}{\text{logarithmic mean}}, \quad \kappa_i := \sqrt{\det \nabla^2 H(m_i)}$.

Especially, for $\kappa_0 = \kappa_1$ holds $\varrho \approx \alpha \quad \Rightarrow$ *LSI likes symmetry in the system.*

Question: Relation to cut-off phenomenon of Markov chains?

Outlook and open questions

- discrete setting (Curie-Weiss model with random field)
- infinite dimension (stochastic PDEs like 1d stochastic Allen-Cahn equation)
- non-reversible processes (∇H replaced by $b \in C^2(\mathbb{R}^n, \mathbb{R}^n)$)
- other functional inequalities (interpolation between PI and LSI)

References

[BGK05] A. Bovier, V. Gayrard, and M. Klein. Metastability in reversible diffusion processes II. *J. Eur. Math. Soc. (JEMS)*, 7(1):69–99, 2005.

[CGW10] P. Cattiaux, A. Guillin, and L. Wu. A note on Talagrand's transportation inequality and LSI. *Probab. Theory Related Fields*, 148(1-2):285–304, 2010.

[CM10] D. Chafaï and F. Malrieu. On fine properties of mixtures with respect to concentration of measure and Sobolev type inequalities. *Ann. Inst. Henri Poincaré*, 2010.

[MS12] G. Menz and A. Schlichting. PI and LSI by decomposition of the energy landscape. `http://arxiv.org/abs/1202.1510`, 2012.

[DNG09] J. Dolbeault, B. Nazaret, and G. Savaré. A new class of transport distances between measures. *Calc. Var. Partial Differential Equations*, 34(2):193–231, 2009.

Schlichting@uni-bonn.de