Two-scale decomposition for metastable dynamics in continuous and discrete setting.

André Schlichting

joint work with Georg Menz (Stanford) and Martin Slowik (TU Berlin)

ERC Workshop on
Energy/Entropy-Driven Systems and Applications

October 9, 2013

Introduction

Metastability

A system possesses metastable states if:

- (most) initial states converge quickly to some metastable state
- the lifetime of the metastable state is very long
- the probability of return is very small

flatness of the state space

- Allen-Cahn equation (Otto-Reznikoff)
- Becker-Döring systems (Penrose)

barriers in the state space

- chemical reactions (Eyring-Kramers)
- Nucleation in spin systems (Bovier, den Hollander, Spitoni)

Introduction

Metastability

A system possesses metastable states if:

- (most) initial states converge quickly to some metastable state
- the lifetime of the metastable state is very long
- the probability of return is very small

flatness of the state space

- Allen-Cahn equation (Otto-Reznikoff)
- Becker-Döring systems (Penrose)

barriers in the state space

- chemical reactions (Eyring-Kramers)
- Nucleation in spin systems (Bovier, den Hollander, Spitoni)

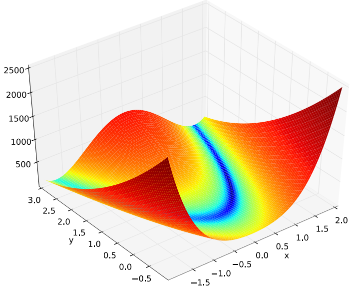
Introduction

Metastability

A system possesses metastable states if:

- (most) initial states converge quickly to some metastable state
- the lifetime of the metastable state is very long
- the probability of return is very small

flatness of the state space



barriers in the state space

- chemical reactions (Eyring-Kramers)
- Nucleation in spin systems (Bovier, den Hollander, Spitoni)

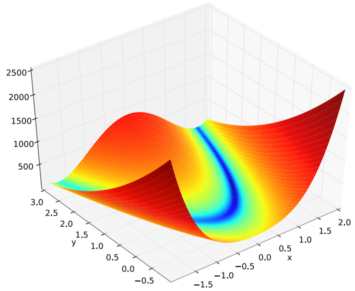
Introduction

Metastability

A system possesses metastable states if:

- (most) initial states converge quickly to some metastable state
- the lifetime of the metastable state is very long
- the probability of return is very small

flatness of the state space



barriers in the state space

- chemical reactions (Eyring-Kramers)
- Nucleation in spin systems (Bovier, den Hollander, Spitoni)

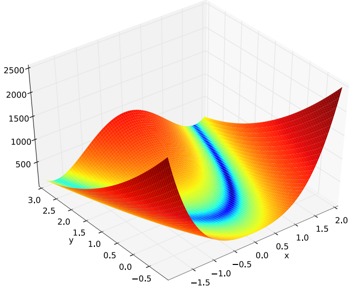
Introduction

Metastability

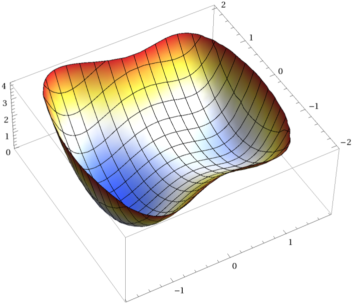
A system possesses metastable states if:

- (most) initial states converge quickly to some metastable state
- the lifetime of the metastable state is very long
- the probability of return is very small

flatness of the state space



barriers in the state space



Introduction

Quantification for Metastability in Fokker-Planck equation

Spectral characterization vs. energetic/entropic characterization

Fokker-Planck equation with potential $H : \mathbb{R}^n \to \mathbb{R}$:

$$\partial_t u_t = \nabla \cdot (\varepsilon \nabla u_t + u_t \nabla H) \qquad \text{equilibrium: } \mu = \exp\left(-\varepsilon^{-1} H\right)$$

- Spectral gap λ: Exponential convergence of $\int |u - \mu|^2 \, dx$
- [JKO98]: gradient flow of energy

$$\mathcal{E}(u) = \int u \log u \, dx + \varepsilon^{-1} \int H u \, dx = \int \frac{u}{\mu} \log \frac{u}{\mu} \, d\mu$$

Logarithmic Sobolev inequality $\Rightarrow$ exponential convergence

Goal:	Metastability in "entropic" setting for Foker-Planck
Necessary:	Technique providing sharp asymptotics for $\varepsilon \ll 1$
Outlook:	gradient flow structure for Markov chains (Maas, Mielke): Metastability in "entropic" setting for Markov chains

Introduction

Quantification for Metastability in Fokker-Planck equation

Spectral characterization vs. energetic/entropic characterization

Fokker-Planck equation with potential $H : \mathbb{R}^n \to \mathbb{R}$:

$$\partial_t u_t = \nabla \cdot (\varepsilon \nabla u_t + u_t \nabla H) \qquad \text{equilibrium: } \mu = \exp\left(-\varepsilon^{-1} H\right)$$

- Spectral gap λ: Exponential convergence of $\int |u - \mu|^2 \, \mathrm{d}x$
- [JKO98]: gradient flow of energy

$$\mathcal{E}(u) = \int u \log u \, \mathrm{d}x + \varepsilon^{-1} \int H u \, \mathrm{d}x = \int \frac{u}{\mu} \log \frac{u}{\mu} \, \mathrm{d}\mu$$

Logarithmic Sobolev inequality $\Rightarrow$ exponential convergence

Goal: Metastability in “entropic” setting for Foker-Planck

Necessary: Technique providing sharp asymptotics for $\varepsilon \ll 1$

Outlook: gradient flow structure for Markov chains (Maas, Mielke): Metastability in “entropic” setting for Markov chains

Spectral characterization vs. energetic/entropic characterization

Fokker-Planck equation with potential $H : \mathbb{R}^n \to \mathbb{R}$:

$$\partial_t u_t = \nabla \cdot (\varepsilon \nabla u_t + u_t \nabla H) \qquad \text{equilibrium: } \mu = \exp\left(-\varepsilon^{-1} H\right)$$

- Spectral gap λ: Exponential convergence of $\int |u - \mu|^2 \, \mathrm{d}x$
- [JKO98]: gradient flow of energy

$$\mathcal{E}(u) = \int u \log u \, \mathrm{d}x + \varepsilon^{-1} \int H u \, \mathrm{d}x = \int \frac{u}{\mu} \log \frac{u}{\mu} \, \mathrm{d}\mu$$

Logarithmic Sobolev inequality $\Rightarrow$ exponential convergence

Goal: Metastability in "entropic" setting for Foker-Planck

Necessary: Technique providing sharp asymptotics for $\varepsilon \ll 1$

Outlook: gradient flow structure for Markov chains (Maas, Mielke): Metastability in "entropic" setting for Markov chains

Spectral characterization vs. energetic/entropic characterization

Fokker-Planck equation with potential $H : \mathbb{R}^n \to \mathbb{R}$:

$$\partial_t u_t = \nabla \cdot (\varepsilon \nabla u_t + u_t \nabla H) \qquad \text{equilibrium: } \mu = \exp\left(-\varepsilon^{-1} H\right)$$

- Spectral gap λ: Exponential convergence of $\int |u - \mu|^2 \, \mathrm{d}x$
- [JKO98]: gradient flow of energy

$$\mathcal{E}(u) = \int u \log u \, \mathrm{d}x + \varepsilon^{-1} \int H u \, \mathrm{d}x = \int \frac{u}{\mu} \log \frac{u}{\mu} \, \mathrm{d}\mu$$

Logarithmic Sobolev inequality $\Rightarrow$ exponential convergence

Goal: Metastability in "entropic" setting for Foker-Planck

Necessary: Technique providing sharp asymptotics for $\varepsilon \ll 1$

Outlook: gradient flow structure for Markov chains (Maas, Mielke): Metastability in "entropic" setting for Markov chains

Spectral characterization vs. energetic/entropic characterization

Fokker-Planck equation with potential $H : \mathbb{R}^n \to \mathbb{R}$:

$$\partial_t u_t = \nabla \cdot (\varepsilon \nabla u_t + u_t \nabla H) \qquad \text{equilibrium: } \mu = \exp\left(-\varepsilon^{-1} H\right)$$

- Spectral gap λ: Exponential convergence of $\int |u - \mu|^2 \,\mathrm{d}x$
- [JKO98]: gradient flow of energy

$$\mathcal{E}(u) = \int u \log u \,\mathrm{d}x + \varepsilon^{-1} \int H u \,\mathrm{d}x = \int \frac{u}{\mu} \log \frac{u}{\mu} \,\mathrm{d}\mu$$

Logarithmic Sobolev inequality $\Rightarrow$ exponential convergence

Goal: Metastability in “entropic” setting for Foker-Planck

Necessary: Technique providing sharp asymptotics for $\varepsilon \ll 1$

Outlook: gradient flow structure for Markov chains (Maas, Mielke): Metastability in “entropic” setting for Markov chains

Spectral characterization vs. energetic/entropic characterization

Fokker-Planck equation with potential $H : \mathbb{R}^n \to \mathbb{R}$:

$$\partial_t u_t = \nabla \cdot (\varepsilon \nabla u_t + u_t \nabla H) \qquad \text{equilibrium: } \mu = \exp\left(-\varepsilon^{-1} H\right)$$

- Spectral gap λ: Exponential convergence of $\int |u - \mu|^2 \, \mathrm{d}x$
- [JKO98]: gradient flow of energy

$$\mathcal{E}(u) = \int u \log u \, \mathrm{d}x + \varepsilon^{-1} \int H u \, \mathrm{d}x = \int \frac{u}{\mu} \log \frac{u}{\mu} \, \mathrm{d}\mu$$

Logarithmic Sobolev inequality $\Rightarrow$ exponential convergence

Goal: Metastability in “entropic” setting for Foker-Planck

Necessary: Technique providing sharp asymptotics for $\varepsilon \ll 1$

Outlook: gradient flow structure for Markov chains (Maas, Mielke): Metastability in “entropic” setting for Markov chains

Poincaré and logarithmic Sobolev inequality

Generator L gives evolution of $f_t = u_t/\mu$:

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

Definition

μ satisfies the Poincaré inequality $\mathrm{PI}(\lambda)$ if $\forall f : \mathbb{R}^n \to \mathbb{R}$

$$\operatorname{var}_\mu(f) := \int \left(f - \int f \,\mathrm{d}\mu \right)^2 \mathrm{d}\mu \le \frac{1}{\lambda} \int |\nabla f|^2 \,\mathrm{d}\mu. \qquad \mathrm{PI}(\lambda)$$

and the logarithmic Sobolev inequality $\mathrm{LSI}(\alpha)$ if $\forall f : \mathbb{R}^n \to \mathbb{R}$

$$\operatorname{Ent}_\mu(f) := \int f \log \frac{f}{\int f \,\mathrm{d}\mu} \mathrm{d}\mu \le \frac{1}{\alpha} \int \frac{|\nabla f|^2}{2f} \mathrm{d}\mu. \qquad \mathrm{LSI}(\alpha)$$

$\mathrm{PI}(\lambda)$ and $\mathrm{LSI}(\alpha)$ imply exponential convergence to μ:

$$\mathrm{PI}(\lambda) \;\Rightarrow\; \operatorname{var}_\mu(f_t) \le \operatorname{var}_\mu(f_0) e^{-2\lambda\varepsilon t}$$

$$\mathrm{LSI}(\alpha) \;\Rightarrow\; \operatorname{Ent}_\mu(f_t) \le \operatorname{Ent}_\mu(f_0) e^{-2\alpha\varepsilon t}.$$

Poincaré and logarithmic Sobolev inequality

Generator L gives evolution of $f_t = u_t/\mu$:

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

Definition

μ satisfies the Poincaré inequality PI(λ) if $\forall f : \mathbb{R}^n \to \mathbb{R}$

$$\operatorname{var}_\mu(f) := \int \left(f - \int f \, d\mu \right)^2 d\mu \le \frac{1}{\lambda} \int |\nabla f|^2 \, d\mu. \qquad \text{PI}(\lambda)$$

and the logarithmic Sobolev inequality LSI(α) if $\forall f : \mathbb{R}^n \to \mathbb{R}$

$$\operatorname{Ent}_\mu(f) := \int f \log \frac{f}{\int f \, d\mu} d\mu \le \frac{1}{\alpha} \int \frac{|\nabla f|^2}{2f} d\mu. \qquad \text{LSI}(\alpha)$$

PI(λ) and LSI(α) imply exponential convergence to μ:

$$\text{PI}(\lambda) \ \Rightarrow \ \operatorname{var}_\mu(f_t) \le \operatorname{var}_\mu(f_0) e^{-2\lambda\varepsilon t}$$

$$\text{LSI}(\alpha) \ \Rightarrow \ \operatorname{Ent}_\mu(f_t) \le \operatorname{Ent}_\mu(f_0) e^{-2\alpha\varepsilon t}.$$

Poincaré and logarithmic Sobolev inequality

Generator L gives evolution of $f_t = u_t/\mu$:

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

Definition

μ satisfies the Poincaré inequality $\mathrm{PI}(\lambda)$ if $\forall f : \mathbb{R}^n \to \mathbb{R}$

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

and the logarithmic Sobolev inequality $\mathrm{LSI}(\alpha)$ if $\forall f : \mathbb{R}^n \to \mathbb{R}$

$$\mathrm{Ent}_\mu(f) := \int f \log \frac{f}{\int f \,\mathrm{d}\mu} \mathrm{d}\mu \le \frac{1}{\alpha} \int \frac{|\nabla f|^2}{2f} \mathrm{d}\mu. \qquad \mathrm{LSI}(\alpha)$$

$\mathrm{PI}(\lambda)$ and $\mathrm{LSI}(\alpha)$ imply exponential convergence to μ:

$$\mathrm{PI}(\lambda) \;\Rightarrow\; \mathrm{var}_\mu(f_t) \le \mathrm{var}_\mu(f_0) e^{-2\lambda\varepsilon t}$$

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

Poincaré and logarithmic Sobolev inequality

Generator L gives evolution of $f_t = u_t/\mu$:

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

Definition

μ satisfies the Poincaré inequality PI(λ) if $\forall f : \mathbb{R}^n \to \mathbb{R}$

$$\operatorname{var}_\mu(f) := \int \left(f - \int f \, d\mu \right)^2 d\mu \le \frac{1}{\lambda} \int |\nabla f|^2 \, d\mu. \qquad \text{PI}(\lambda)$$

and the logarithmic Sobolev inequality LSI(α) if $\forall f : \mathbb{R}^n \to \mathbb{R}$

$$\operatorname{Ent}_\mu(f^2) := \int f^2 \log \frac{f^2}{\int f^2 d\mu} d\mu \le \frac{2}{\alpha} \int |\nabla f|^2 \, d\mu. \qquad \text{LSI}(\alpha)$$

PI(λ) and LSI(α) imply exponential convergence to μ:

$$\text{PI}(\lambda) \;\Rightarrow\; \operatorname{var}_\mu(f_t) \le \operatorname{var}_\mu(f_0) e^{-2\lambda\varepsilon t}$$

$$\text{LSI}(\alpha) \;\Rightarrow\; \operatorname{Ent}_\mu(f_t) \le \operatorname{Ent}_\mu(f_0) e^{-2\alpha\varepsilon t}.$$

Poincaré and logarithmic Sobolev inequality

Generator L gives evolution of $f_t = u_t/\mu$:

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

Definition

μ satisfies the Poincaré inequality PI(λ) if $\forall f : \mathbb{R}^n \to \mathbb{R}$

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

and the logarithmic Sobolev inequality LSI(α) if $\forall f : \mathbb{R}^n \to \mathbb{R}$

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

PI(λ) and LSI(α) imply exponential convergence to μ:

$$\mathrm{PI}(\lambda) \;\Rightarrow\; \mathrm{var}_\mu(f_t) \le \mathrm{var}_\mu(f_0) e^{-2\lambda\varepsilon t}$$

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

Goal: Optimal constants in PI and LSI

Accurate estimates of λ and α in the regime $\varepsilon \ll 1$:

$$\lambda = C_\lambda(\varepsilon) e^{-\frac{\Delta H}{\varepsilon}} (1 + o(1)) \qquad \text{and} \qquad \alpha = C_\alpha(\varepsilon) e^{-\frac{\Delta H}{\varepsilon}} (1 + o(1)).$$

Goal: Optimal constants in PI and LSI

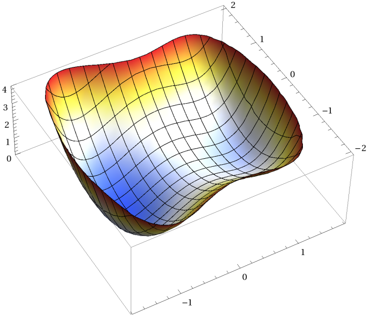
Accurate estimates of λ and α in the regime $\varepsilon \ll 1$:

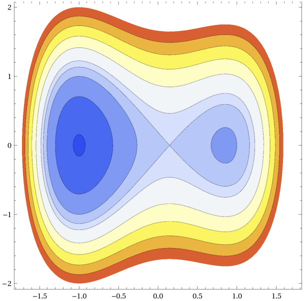
$$\lambda = C_\lambda(\varepsilon) e^{-\frac{\Delta H}{\varepsilon}}(1 + o(1)) \qquad \text{and} \qquad \alpha = C_\alpha(\varepsilon) e^{-\frac{\Delta H}{\varepsilon}}(1 + o(1)).$$

Heuristics

$$\mathrm{d}X_t = -\nabla H(X_t)\,\mathrm{d}t + \sqrt{2\varepsilon}\,\mathrm{d}W_t$$

- particle follows $-\nabla H$ as long as $|\nabla H| \sim 1$
- noise is dominant, if $|\nabla H| \lesssim \sqrt{\varepsilon}$





Heuristics

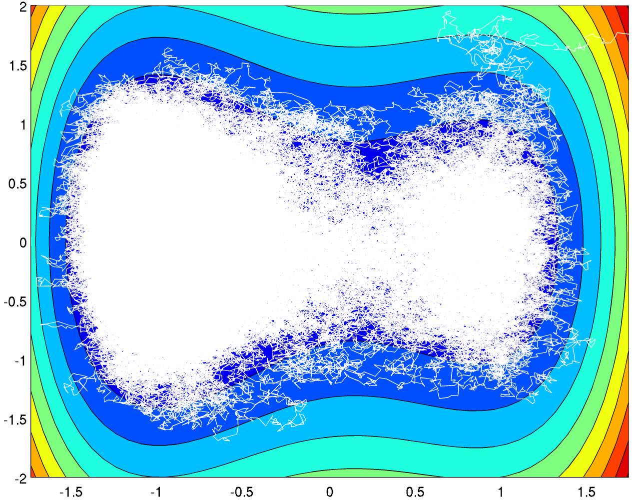


Figure : Trajectory for $\varepsilon = 0.4$

Heuristics

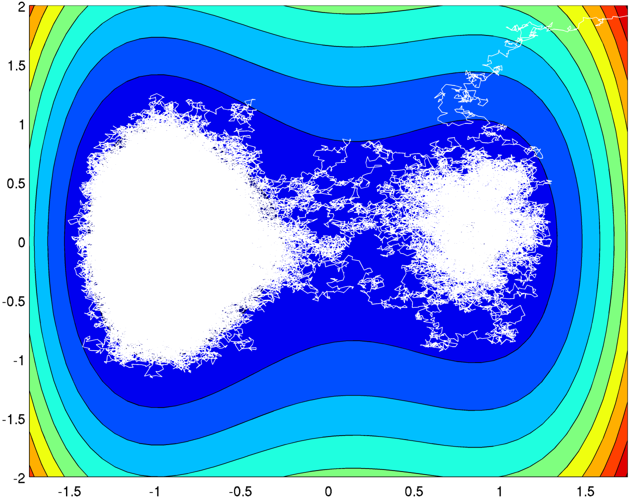


Figure : Trajectory for $\varepsilon = 0.2$

Heuristics

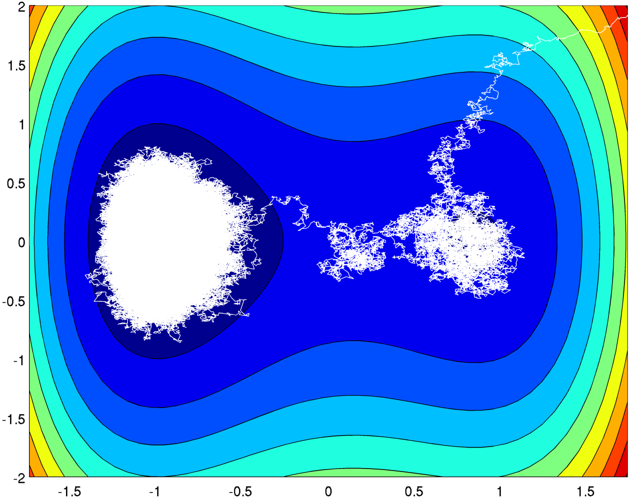


Figure : Trajectory for $\varepsilon = 0.1$

Heuristics

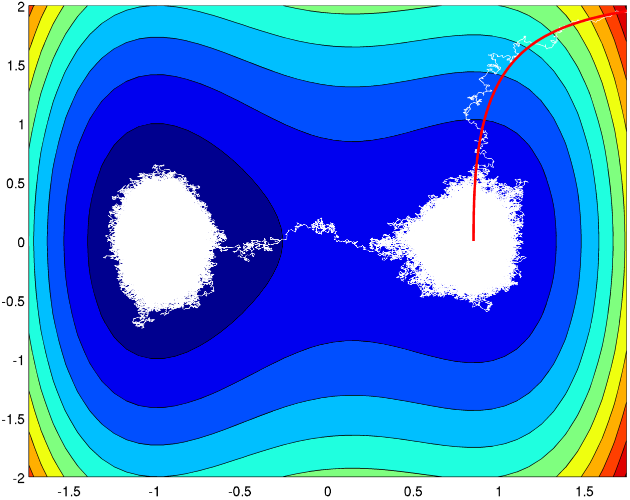


Figure : Trajectory for $\varepsilon = 0.05$ (red $\varepsilon = 0$)

Two scales by decomposition à la [GOVW09][1]

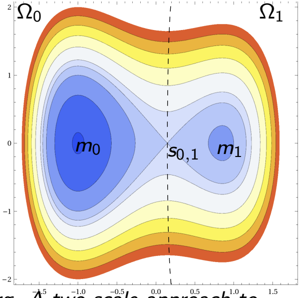
The partition $\biguplus_i \Omega_i = \mathbb{R}^n$ is called admissible for μ if:

(i) For each local minimum $m_i \in \mathcal{M}$ exists Ω_i with $m_i \in \Omega_i$

(ii) The partition sum of each Ω_i is approximately Gaussian

$$\mu(\Omega_i) Z_\mu = \frac{(2\pi\varepsilon)^{\frac{n}{2}}}{\sqrt{\det \nabla^2 H(m_i)}} \exp\left(-\frac{H(m_i)}{\varepsilon}\right)(1 + o(1)) .$$

Restricted measures: $\mu_i := \mu \llcorner \Omega_i, \quad i = 0, 1.$



Macroscopic measures $\bar{\mu}$ on $\{0, 1\}$:

$$\bar{\mu} := Z_0 \delta_0 + Z_1 \delta_1 .$$

Mixture representation:

$$\mu = Z_0 \mu_0 + Z_1 \mu_1 \text{ with } Z_i := \mu(\Omega_i).$$

[1] N. Grunewald, F. Otto, C. Villani, and M. G. Westdickenberg, *A two-scale approach to logarithmic Sobolev inequalities and the hydrodynamic limit,* Annales de l'Institut Henri Poincaré, Probabilités et Statistiques, 45:**2**, 2009.

Splitting

Ideas motivated from [CM10][2]

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

$$\mathrm{Ent}_\mu(f^2) = \overbrace{Z_0\,\mathrm{Ent}_{\mu_0}(f^2) + Z_1\,\mathrm{Ent}_{\mu_1}(f^2)}^{\text{local entropies}} + \overbrace{\mathrm{Ent}_{\bar{\mu}}\left(\mathbb{E}_{\mu_\bullet}(f^2)\right)}^{\text{macroscopic entropy}}$$

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

Expect from heuristics:

- *good* estimate for local variances/entropies
- *exponential* estimate for mean-difference

[2] D. Chafaï and F. Malrieu, *On fine properties of mixtures with respect to concentration of measure and Sobolev type inequalities*, Annales de l'Institut Henri Poincaré, Probabilités et Statistiques, 46:**1**, 2010.

Splitting

Ideas motivated from [CM10][2]

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

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

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

Expect from heuristics:

- *good* estimate for local variances/entropies
- *exponential* estimate for mean-difference

[2] D. Chafaï and F. Malrieu, *On fine properties of mixtures with respect to concentration of measure and Sobolev type inequalities*, Annales de l'Institut Henri Poincaré, Probabilités et Statistiques, 46:**1**, 2010.

Splitting

Ideas motivated from [CM10][2]

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

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

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

Expect from heuristics:

- *good* estimate for local variances/entropies
- *exponential* estimate for mean-difference

[2] D. Chafaï and F. Malrieu, *On fine properties of mixtures with respect to concentration of measure and Sobolev type inequalities*, Annales de l'Institut Henri Poincaré, Probabilités et Statistiques, 46:**1**, 2010.

Main results

Theorem (Local PI and LSI [Menz, S. 2012])

There exists an admissible partition $\biguplus_i \Omega_i = \mathbb{R}^n$ *such that each local measure* $\mu_i = \mu \llcorner \Omega_i$ *satisfies* $\mathrm{PI}(\lambda_{loc})$ *and* $\mathrm{LSI}(\alpha_{loc})$ *with*

$$\lambda_{loc}^{-1} = O(\varepsilon) \quad \textit{and} \quad \alpha_{loc}^{-1} = O(1).$$

- PI is as good as for convex potential
- Non-convexity of potential worsens LSI
- Both results scale optimal in one dimension

Theorem (Mean-difference estimate [Menz, S. 2012])

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

"$\lesssim$": *up to multiplicative error* $1 + o(1)$ *as* $\varepsilon \to 0$.

Main results

Theorem (Local PI and LSI [Menz, S. 2012])

There exists an admissible partition $\biguplus_i \Omega_i = \mathbb{R}^n$ such that each local measure $\mu_i = \mu \llcorner \Omega_i$ satisfies $\mathrm{PI}(\lambda_{loc})$ *and* $\mathrm{LSI}(\alpha_{loc})$ *with*

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

- PI is as good as for convex potential
- Non-convexity of potential worsens LSI
- Both results scale optimal in one dimension

Theorem (Mean-difference estimate [Menz, S. 2012])

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

"$\lesssim$": *up to multiplicative error* $1 + o(1)$ *as* $\varepsilon \to 0$.

Main results

Theorem (Local PI and LSI [Menz, S. 2012])

There exists an admissible partition $\biguplus_i \Omega_i = \mathbb{R}^n$ *such that each local measure* $\mu_i = \mu \llcorner \Omega_i$ *satisfies* $\mathrm{PI}(\lambda_{loc})$ *and* $\mathrm{LSI}(\alpha_{loc})$ *with*

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

- PI is as good as for convex potential
- Non-convexity of potential worsens LSI
- Both results scale optimal in one dimension

Theorem (Mean-difference estimate [Menz, S. 2012])

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

"$\lesssim$": *up to multiplicative error* $1 + o(1)$ *as* $\varepsilon \to 0$.

Main results

Theorem (Local PI and LSI [Menz, S. 2012])

There exists an admissible partition $\biguplus_i \Omega_i = \mathbb{R}^n$ such that each local measure $\mu_i = \mu \llcorner \Omega_i$ satisfies $\mathrm{PI}(\lambda_{loc})$ *and* $\mathrm{LSI}(\alpha_{loc})$ *with*

$$\lambda_{loc}^{-1} = O(\varepsilon) \quad \textit{and} \quad \alpha_{loc}^{-1} = O(1).$$

- PI is as good as for convex potential
- Non-convexity of potential worsens LSI
- Both results scale optimal in one dimension

Theorem (Mean-difference estimate [Menz, S. 2012])

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

"$\lesssim$": up to multiplicative error $1 + o(1)$ as $\varepsilon \to 0$.

Main results

Theorem (Local PI and LSI [Menz, S. 2012])

There exists an admissible partition $\biguplus_i \Omega_i = \mathbb{R}^n$ *such that each local measure* $\mu_i = \mu \llcorner \Omega_i$ *satisfies* $\mathrm{PI}(\lambda_{loc})$ *and* $\mathrm{LSI}(\alpha_{loc})$ *with*

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

- PI is as good as for convex potential
- Non-convexity of potential worsens LSI
- Both results scale optimal in one dimension

Theorem (Mean-difference estimate [Menz, S. 2012])

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

"$\lesssim$": *up to multiplicative error* $1 + o(1)$ *as* $\varepsilon \to 0$.

Eyring-Kramers formula

Corollary

The measure μ *satisfies* $\mathrm{PI}(\lambda)$ *and* $\mathrm{LSI}(\alpha)$ *with*

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

Asymptotic evaluation of the factor $\Lambda(Z_0, Z_1)$ for two special cases:

$$H(m_0) < H(m_1): \quad \frac{\lambda}{\alpha} \approx \frac{1}{2}\left(\frac{H(m_1) - H(m_0)}{\varepsilon} + \log\left(\frac{\kappa_0}{\kappa_1}\right)\right) = O(\varepsilon^{-1})$$

$$H(m_0) = H(m_1): \quad \frac{\lambda}{\alpha} \approx \frac{\frac{\kappa_0+\kappa_1}{2}}{\Lambda(\kappa_0, \kappa_1)} = O(1),$$

where $\kappa_i := \sqrt{\det \nabla^2 H(m_i)}$.

Eyring-Kramers formula

Corollary

The measure μ *satisfies* $\mathrm{PI}(\lambda)$ *and* $\mathrm{LSI}(\alpha)$ *with*

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

Asymptotic evaluation of the factor $\Lambda(Z_0, Z_1)$ for two special cases:

$$H(m_0) < H(m_1): \quad \frac{\lambda}{\alpha} \approx \frac{1}{2}\left(\frac{H(m_1) - H(m_0)}{\varepsilon} + \log\left(\frac{\kappa_0}{\kappa_1}\right)\right) = O(\varepsilon^{-1})$$

$$H(m_0) = H(m_1): \quad \frac{\lambda}{\alpha} \approx \frac{\frac{\kappa_0+\kappa_1}{2}}{\Lambda(\kappa_0, \kappa_1)} = O(1),$$

where $\kappa_i := \sqrt{\det \nabla^2 H(m_i)}$.

Proof: Local PI and LSI

Theorem (Local PI and LSI [Menz, S. 2012])

There exists an admissible partition $\biguplus_i \Omega_i = \mathbb{R}^n$ such that each local measure $\mu_i = \mu \llcorner \Omega_i$ satisfies PI(λ_{loc}) *and* LSI(α_{loc}) *with*

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

- lack of convexity of H on Ω
  $\Rightarrow$ rules out Bakry-Émery criterion
- non-exponential behavior of constants
  $\Rightarrow$ rules out Holley-Stroock perturbation principle
- optimality available in one dimension
  $\Rightarrow$ Muckenhoupt and Bobkov/Götze functional

Proof: Local PI and LSI

Theorem (Local PI and LSI [Menz, S. 2012])

There exists an admissible partition $\biguplus_i \Omega_i = \mathbb{R}^n$ such that each local measure $\mu_i = \mu \llcorner \Omega_i$ satisfies $\mathrm{PI}(\lambda_{loc})$ and $\mathrm{LSI}(\alpha_{loc})$ with

$$\lambda_{loc}^{-1} = O(\varepsilon) \quad \textit{and} \quad \alpha_{loc}^{-1} = O(1).$$

- lack of convexity of H on Ω
  $\Rightarrow$ rules out Bakry-Émery criterion
- non-exponential behavior of constants
  $\Rightarrow$ rules out Holley-Stroock perturbation principle
- optimality available in one dimension
  $\Rightarrow$ Muckenhoupt and Bobkov/Götze functional

Proof: Local PI and LSI

Theorem (Local PI and LSI [Menz, S. 2012])

There exists an admissible partition $\biguplus_i \Omega_i = \mathbb{R}^n$ *such that each local measure* $\mu_i = \mu \llcorner \Omega_i$ *satisfies* PI(λ_{loc}) *and* LSI(α_{loc}) *with*

$$\lambda_{loc}^{-1} = O(\varepsilon) \quad \textit{and} \quad \alpha_{loc}^{-1} = O(1).$$

- lack of convexity of H on Ω
  $\Rightarrow$ rules out Bakry-Émery criterion
- non-exponential behavior of constants
  $\Rightarrow$ rules out Holley-Stroock perturbation principle
- optimality available in one dimension
  $\Rightarrow$ Muckenhoupt and Bobkov/Götze functional

Proof: Local PI and LSI

Theorem (Local PI and LSI [Menz, S. 2012])

There exists an admissible partition $\biguplus_i \Omega_i = \mathbb{R}^n$ such that each local measure $\mu_i = \mu \llcorner \Omega_i$ satisfies PI(λ_{loc}) *and* LSI(α_{loc}) *with*

$$\lambda_{loc}^{-1} = O(\varepsilon) \quad \textit{and} \quad \alpha_{loc}^{-1} = O(1).$$

- lack of convexity of H on Ω
  $\Rightarrow$ rules out Bakry-Émery criterion
- non-exponential behavior of constants
  $\Rightarrow$ rules out Holley-Stroock perturbation principle
- optimality available in one dimension
  $\Rightarrow$ Muckenhoupt and Bobkov/Götze functional

Proof: Local PI and LSI via Lyapunov condition

Technique developed by Bakry, Barthe, Cattiaux, Guillin, Wang and Wu 2008–

Definition (Lyapunov condition on domains)

L satisfies a Lyapunov condition with constants $\lambda, b > 0$ and some $U \subset \Omega$, if there exists a Lyapunov function $W : \Omega \to [1, \infty)$ satisfying

$$\frac{LW}{\varepsilon W} \leq -\lambda + b\, \mathbb{1}_U.$$

and Neumann boundary condition on Ω, such that integration by parts holds

$$\int_\Omega f(-LW)\, \mathrm{d}\mu = \varepsilon \int_\Omega \langle \nabla f, \nabla W \rangle\, \mathrm{d}\mu.$$

Theorem ([BBCG08], [Menz, S. '12 for domains])

Suppose L satisfies a Lyapunov condition on Ω and $\mu \llcorner U$ satisfies $\mathrm{PI}(\lambda_U)$, *then $\mu \llcorner \Omega$ satisfies* $\mathrm{PI}(\lambda_\Omega)$ *with*

$$\lambda_\Omega \geq \frac{\lambda}{b + \lambda_U} \lambda_U$$

Proof: Local PI and LSI via Lyapunov condition

Technique developed by Bakry, Barthe, Cattiaux, Guillin, Wang and Wu 2008–

Definition (Lyapunov condition on domains)

L satisfies a Lyapunov condition with constants $\lambda, b > 0$ and some $U \subset \Omega$, if there exists a Lyapunov function $W : \Omega \to [1, \infty)$ satisfying

$$\frac{LW}{\varepsilon W} \leq -\lambda + b\, \mathbb{1}_U.$$

and Neumann boundary condition on Ω, such that integration by parts holds

$$\int_\Omega f(-LW)\,\mathrm{d}\mu = \varepsilon \int_\Omega \langle \nabla f, \nabla W \rangle\,\mathrm{d}\mu.$$

Theorem ([BBCG08], [Menz, S. '12 for domains])

Suppose L satisfies a Lyapunov condition on Ω and $\mu \llcorner U$ satisfies $\mathrm{PI}(\lambda_U)$, then $\mu \llcorner \Omega$ satisfies $\mathrm{PI}(\lambda_\Omega)$ with

$$\lambda_\Omega \geq \frac{\lambda}{b + \lambda_U} \lambda_U$$

Proof: Local PI and LSI

Lyapunov function

- Task: Find a function $W : \Omega \to [1, \infty)$ such that

$$\frac{LW}{W} \le -\lambda + b\, \mathbb{1}_{B_{a\sqrt{\varepsilon}}(m)}.$$

- Ansatz $W = \exp\left(\frac{\tilde{H}}{2\varepsilon}\right)$, where $\tilde{H}$ is an ε-perturbation of H

$$\frac{\tilde{L}W}{W} = \frac{1}{2}\Delta \tilde{H} - \frac{1}{4\varepsilon}|\nabla \tilde{H}|^2 \overset{!}{\le} -\lambda.$$

  - if x is $\sqrt{\varepsilon}$-away from critical points: $\varepsilon^{-1}|\nabla \tilde{H}(x)|^2 \ge 4\lambda$
  - if x is $\sqrt{\varepsilon}$-nearby a critical point of index $k \ge 1$

$$\Delta \tilde{H}(x) = \underbrace{\tilde{\lambda}_1^- + \cdots + \tilde{\lambda}_k^-}_{<0} + \underbrace{\tilde{\lambda}_{k+1}^+ + \cdots + \tilde{\lambda}_n^+}_{>0} + O(\sqrt{\varepsilon})$$

Can negative eigenvalues be enforced such that $\Delta \tilde{H}(x) \le -2\lambda$?

Proof: Local PI and LSI

Lyapunov function

- Task: Find a function $W : \Omega \to [1, \infty)$ such that

$$\frac{LW}{W} \leq -\lambda + b\, \mathbb{1}_{B_{a\sqrt{\varepsilon}}(m)}.$$

- Ansatz $W = \exp\left(\frac{\tilde{H}}{2\varepsilon}\right)$, where $\tilde{H}$ is an ε-perturbation of H

$$\frac{\tilde{L}W}{W} = \frac{1}{2}\Delta\tilde{H} - \frac{1}{4\varepsilon}|\nabla\tilde{H}|^2 \stackrel{!}{\leq} -\lambda.$$

  - if x is $\sqrt{\varepsilon}$-away from critical points: $\varepsilon^{-1}|\nabla\tilde{H}(x)|^2 \geq 4\lambda$
  - if x is $\sqrt{\varepsilon}$-nearby a critical point of index $k \geq 1$

$$\Delta\tilde{H}(x) = \underbrace{\tilde{\lambda}_1^- + \cdots + \tilde{\lambda}_k^-}_{<0} + \underbrace{\tilde{\lambda}_{k+1}^+ + \cdots + \tilde{\lambda}_n^+}_{>0} + O(\sqrt{\varepsilon})$$

Can negative eigenvalues be enforced such that $\Delta\tilde{H}(x) \leq -2\lambda$?

Proof: Local PI and LSI

Lyapunov function

- Task: Find a function $W : \Omega \to [1, \infty)$ such that

$$\frac{LW}{W} \leq -\lambda + b\,\mathbb{1}_{B_{a\sqrt{\varepsilon}}(m)}.$$

- Ansatz $W = \exp\left(\frac{\tilde{H}}{2\varepsilon}\right)$, where $\tilde{H}$ is an ε-perturbation of H

$$\frac{\tilde{L}W}{W} = \frac{1}{2}\Delta\tilde{H} - \frac{1}{4\varepsilon}|\nabla\tilde{H}|^2 \overset{!}{\leq} -\lambda.$$

  - if x is $\sqrt{\varepsilon}$-away from critical points: $\varepsilon^{-1}|\nabla\tilde{H}(x)|^2 \geq 4\lambda$
  - if x is $\sqrt{\varepsilon}$-nearby a critical point of index $k \geq 1$

$$\Delta\tilde{H}(x) = \underbrace{\tilde{\lambda}_1^- + \cdots + \tilde{\lambda}_k^-}_{<0} + \underbrace{\tilde{\lambda}_{k+1}^+ + \cdots + \tilde{\lambda}_n^+}_{>0} + O(\sqrt{\varepsilon})$$

Can negative eigenvalues be enforced such that $\Delta\tilde{H}(x) \leq -2\lambda$?

Proof: Local PI and LSI

Lyapunov function

- Task: Find a function $W : \Omega \to [1, \infty)$ such that

$$\frac{LW}{W} \leq -\lambda + b\, \mathbb{1}_{B_{a\sqrt{\varepsilon}}(m)}.$$

- Ansatz $W = \exp\left(\frac{\tilde{H}}{2\varepsilon}\right)$, where $\tilde{H}$ is an ε-perturbation of H

$$\frac{\tilde{L}W}{W} = \frac{1}{2}\Delta\tilde{H} - \frac{1}{4\varepsilon}|\nabla\tilde{H}|^2 \overset{!}{\leq} -\lambda.$$

  - if x is $\sqrt{\varepsilon}$-away from critical points: $\varepsilon^{-1}|\nabla\tilde{H}(x)|^2 \geq 4\lambda$
  - if x is $\sqrt{\varepsilon}$-nearby a critical point of index $k \geq 1$

$$\Delta\tilde{H}(x) = \underbrace{\tilde{\lambda}_1^- + \cdots + \tilde{\lambda}_k^-}_{<0} + \underbrace{\tilde{\lambda}_{k+1}^+ + \cdots + \tilde{\lambda}_n^+}_{>0} + O(\sqrt{\varepsilon})$$

Can negative eigenvalues be enforced such that $\Delta\tilde{H}(x) \leq -2\lambda$?

Proof: Local PI and LSI

Lyapunov function

- Task: Find a function $W : \Omega \to [1, \infty)$ such that

$$\frac{LW}{W} \leq -\lambda + b\,\mathbb{1}_{B_{a\sqrt{\varepsilon}}(m)}.$$

- Ansatz $W = \exp\left(\frac{\tilde{H}}{2\varepsilon}\right)$, where $\tilde{H}$ is an ε-perturbation of H

$$\frac{\tilde{L}W}{W} = \frac{1}{2}\Delta\tilde{H} - \frac{1}{4\varepsilon}|\nabla\tilde{H}|^2 \overset{!}{\leq} -\lambda.$$

  - if x is $\sqrt{\varepsilon}$-away from critical points: $\varepsilon^{-1}|\nabla\tilde{H}(x)|^2 \geq 4\lambda$
  - if x is $\sqrt{\varepsilon}$-nearby a critical point of index $k \geq 1$

$$\Delta\tilde{H}(x) = \underbrace{\tilde{\lambda}_1^- + \cdots + \tilde{\lambda}_k^-}_{<0} + \underbrace{\tilde{\lambda}_{k+1}^+ + \cdots + \tilde{\lambda}_n^+}_{>0} + O(\sqrt{\varepsilon})$$

Can negative eigenvalues be enforced such that $\Delta\tilde{H}(x) \leq -2\lambda$?

Proof: Local PI and LSI

Lyapunov function

- Task: Find a function $W : \Omega \to [1,\infty)$ such that

$$\frac{LW}{W} \leq -\lambda + b\,\mathbb{1}_{B_{a\sqrt{\varepsilon}}(m)}.$$

- Ansatz $W = \exp\left(\frac{\tilde{H}}{2\varepsilon}\right)$, where $\tilde{H}$ is an ε-perturbation of H

$$\frac{\tilde{L}W}{W} = \frac{1}{2}\Delta\tilde{H} - \frac{1}{4\varepsilon}|\nabla\tilde{H}|^2 \overset{!}{\leq} -\lambda.$$

  - if x is $\sqrt{\varepsilon}$-away from critical points: $\varepsilon^{-1}|\nabla\tilde{H}(x)|^2 \geq 4\lambda$
  - if x is $\sqrt{\varepsilon}$-nearby a critical point of index $k \geq 1$

$$\Delta\tilde{H}(x) = \underbrace{\tilde{\lambda}_1^- + \cdots + \tilde{\lambda}_k^-}_{<0} + \underbrace{\tilde{\lambda}_{k+1}^+ + \cdots + \tilde{\lambda}_n^+}_{>0} + O(\sqrt{\varepsilon})$$

Can negative eigenvalues be enforced such that $\Delta\tilde{H}(x) \leq -2\lambda$?

Proof: Local PI and LSI

Lyapunov function

- Task: Find a function $W : \Omega \to [1, \infty)$ such that

$$\frac{LW}{W} \leq -\lambda + b\, \mathbb{1}_{B_{a\sqrt{\varepsilon}}(m)}.$$

- Ansatz $W = \exp\left(\frac{\tilde{H}}{2\varepsilon}\right)$, where $\tilde{H}$ is an ε-perturbation of H

$$\frac{\tilde{L}W}{W} = \frac{1}{2}\Delta\tilde{H} - \frac{1}{4\varepsilon}|\nabla\tilde{H}|^2 \overset{!}{\leq} -\lambda.$$

  - if x is $\sqrt{\varepsilon}$-away from critical points: $\varepsilon^{-1}|\nabla\tilde{H}(x)|^2 \geq 4\lambda$
  - if x is $\sqrt{\varepsilon}$-nearby a critical point of index $k \geq 1$

$$\Delta\tilde{H}(x) = \underbrace{\tilde{\lambda}_1^- + \cdots + \tilde{\lambda}_k^-}_{<0} + \underbrace{\tilde{\lambda}_{k+1}^+ + \cdots + \tilde{\lambda}_n^+}_{>0} + O(\sqrt{\varepsilon})$$

Can negative eigenvalues be enforced such that $\Delta\tilde{H}(x) \leq -2\lambda$? YES!

Proof: Local PI and LSI

Construction of Lyapunov function

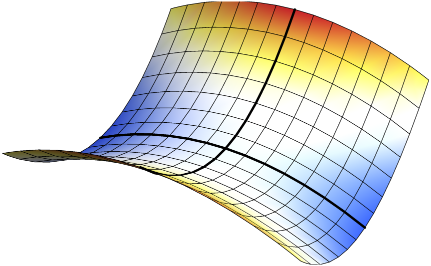


Figure : H around a saddle point

$\tilde{H}$ is quadratic perturbation of H in $\sqrt{\varepsilon}$-neighborhoods of critical points:

$$\sup_x \left| H(x) - \tilde{H}(x) \right| = O(\varepsilon).$$

Proof: Local PI and LSI

Construction of Lyapunov function

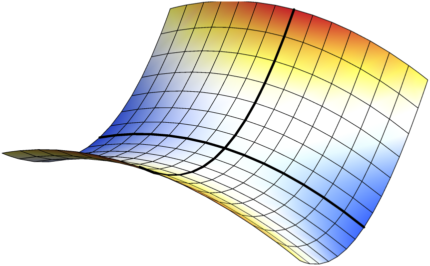


Figure : H around a saddle point

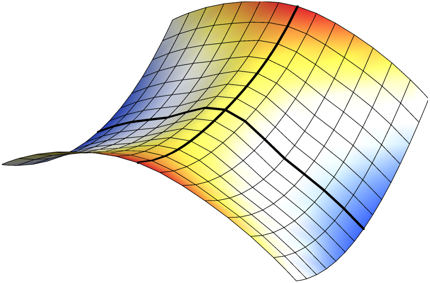


Figure : $\tilde{H}$ around a saddle point

$\tilde{H}$ is quadratic perturbation of H in $\sqrt{\varepsilon}$-neighborhoods of critical points:

$$\sup_x \left| H(x) - \tilde{H}(x) \right| = O(\varepsilon).$$

Proof: Mean-difference estimate

Goal: Find a good estimate for C in

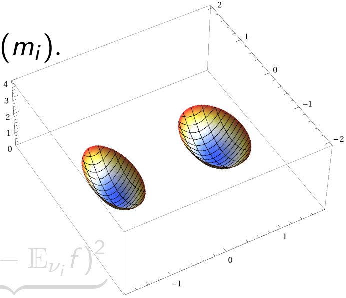
$$\left(\mathbb{E}_{\mu_0}(f) - \mathbb{E}_{\mu_1}(f)\right)^2 \leq C \int |\nabla f|^2 \, \mathrm{d}\mu.$$

Goal: Find a good estimate for C in

$$\left(\mathbb{E}_{\mu_0}(f) - \mathbb{E}_{\mu_1}(f)\right)^2 \leq C \int |\nabla f|^2 \, \mathrm{d}\mu.$$

Step 1: Approximate μ_0 and μ_1 by truncated Gaussians ν_0 and ν_1:

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



Introduce ν_0 and ν_1 as coupling:

$$\left(\mathbb{E}_{\mu_0} f - \mathbb{E}_{\mu_1} f\right)^2 \leq (1+\tau) \underbrace{\left(\mathbb{E}_{\nu_0} f - \mathbb{E}_{\nu_1} f\right)^2}_{\text{transport argument}}$$

$$+ 2(1+\tau^{-1}) \sum_{i=\{0,1\}} \underbrace{\left(\mathbb{E}_{\mu_i} f - \mathbb{E}_{\nu_i} f\right)^2}_{\text{approximation bound}}$$

$\Rightarrow$ Approximation bound follows from local PI and local LSI.

Proof: Mean-difference estimate

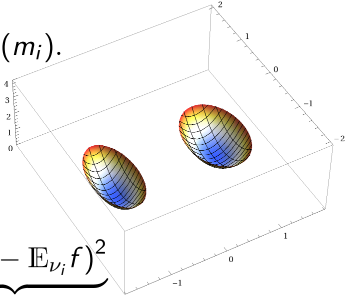
Approximation step

Goal: Find a good estimate for C in

$$\left(\mathbb{E}_{\mu_0}(f) - \mathbb{E}_{\mu_1}(f)\right)^2 \leq C \int |\nabla f|^2 \,\mathrm{d}\mu.$$

Step 1: Approximate μ_0 and μ_1 by truncated Gaussians ν_0 and ν_1:

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



Introduce ν_0 and ν_1 as coupling:

$$\left(\mathbb{E}_{\mu_0} f - \mathbb{E}_{\mu_1} f\right)^2 \leq (1+\tau) \underbrace{\left(\mathbb{E}_{\nu_0} f - \mathbb{E}_{\nu_1} f\right)^2}_{\text{transport argument}}$$

$$+\, 2(1+\tau^{-1}) \sum_{i=\{0,1\}} \underbrace{\left(\mathbb{E}_{\mu_i} f - \mathbb{E}_{\nu_i} f\right)^2}_{\text{approximation bound}}$$

$\Rightarrow$ Approximation bound follows from local PI and local LSI.

Proof: Mean-difference estimate

Transport interpolation

Goal: Find a good estimate for C in

$$(\mathbb{E}_{\nu_0}(f) - \mathbb{E}_{\nu_1}(f))^2 \leq C \int |\nabla f|^2 \, \mathrm{d}\mu.$$

Step 2: Transport $(\Phi_s \in \mathrm{Diff}(\mathbb{R}^n, \mathbb{R}^n))_{s\in[0,1]}$ interpolating $(\Phi_s)_\sharp \nu_0 = \nu_s$

Proof: Mean-difference estimate

Transport interpolation

Goal: Find a good estimate for C in

$$\left(\mathbb{E}_{\nu_0}(f) - \mathbb{E}_{\nu_1}(f)\right)^2 \le C \int |\nabla f|^2 \, \mathrm{d}\mu.$$

Step 2: Transport $(\Phi_s \in \mathrm{Diff}(\mathbb{R}^n, \mathbb{R}^n))_{s \in [0,1]}$ interpolating $(\Phi_s)_\sharp \nu_0 = \nu_s$

$$\left(\int f \, \mathrm{d}\nu_0 - \int f \, \mathrm{d}\nu_1\right)^2 = \left(\int \int_0^1 \frac{\mathrm{d}}{\mathrm{d}s}(f \circ \Phi_s) \, \mathrm{d}s \, \mathrm{d}\nu_0\right)^2$$

Proof: Mean-difference estimate

Transport interpolation

Goal: Find a good estimate for C in

$$\left(\mathbb{E}_{\nu_0}(f) - \mathbb{E}_{\nu_1}(f)\right)^2 \leq C \int |\nabla f|^2 \, \mathrm{d}\mu.$$

Step 2: Transport $(\Phi_s \in \mathrm{Diff}(\mathbb{R}^n, \mathbb{R}^n))_{s\in[0,1]}$ interpolating $(\Phi_s)_\sharp \nu_0 = \nu_s$

$$\begin{aligned}\left(\int f \, \mathrm{d}\nu_0 - \int f \, \mathrm{d}\nu_1\right)^2 &= \left(\int \int_0^1 \frac{\mathrm{d}}{\mathrm{d}s}(f \circ \Phi_s) \, \mathrm{d}s \, \mathrm{d}\nu_0\right)^2 \\ &= \left(\int \int_0^1 \left\langle \dot{\Phi}_s, \nabla f \circ \Phi_s \right\rangle \mathrm{d}s \, \mathrm{d}\nu_0\right)^2\end{aligned}$$

Proof: Mean-difference estimate

Transport interpolation

Goal: Find a good estimate for C in

$$\left(\mathbb{E}_{\nu_0}(f) - \mathbb{E}_{\nu_1}(f)\right)^2 \leq C \int |\nabla f|^2 \, \mathrm{d}\mu.$$

Step 2: Transport $(\Phi_s \in \mathrm{Diff}(\mathbb{R}^n, \mathbb{R}^n))_{s \in [0,1]}$ interpolating $(\Phi_s)_\sharp \nu_0 = \nu_s$

$$\begin{aligned}\left(\int f \, \mathrm{d}\nu_0 - \int f \, \mathrm{d}\nu_1\right)^2 &= \left(\int \int_0^1 \frac{\mathrm{d}}{\mathrm{d}s}(f \circ \Phi_s) \, \mathrm{d}s \, \mathrm{d}\nu_0\right)^2 \\ &= \left(\int_0^1 \int \left\langle \dot{\Phi}_s, \nabla f \circ \Phi_s \right\rangle \mathrm{d}\nu_0 \, \mathrm{d}s\right)^2\end{aligned}$$

Proof: Mean-difference estimate

Transport interpolation

Goal: Find a good estimate for C in

$$\left(\mathbb{E}_{\nu_0}(f) - \mathbb{E}_{\nu_1}(f)\right)^2 \leq C \int |\nabla f|^2 \, \mathrm{d}\mu.$$

Step 2: Transport $(\Phi_s \in \mathrm{Diff}(\mathbb{R}^n, \mathbb{R}^n))_{s\in[0,1]}$ interpolating $(\Phi_s)_\sharp \nu_0 = \nu_s$

$$\begin{aligned}
\left(\int f \, \mathrm{d}\nu_0 - \int f \, \mathrm{d}\nu_1\right)^2 &= \left(\int \int_0^1 \frac{\mathrm{d}}{\mathrm{d}s}(f \circ \Phi_s) \, \mathrm{d}s \, \mathrm{d}\nu_0\right)^2 \\
&= \left(\int_0^1 \int \left\langle \dot{\Phi}_s, \nabla f \circ \Phi_s \right\rangle \mathrm{d}\nu_0 \, \mathrm{d}s\right)^2 \\
&= \left(\int_0^1 \int \left\langle \dot{\Phi}_s \circ \Phi_s^{-1}, \nabla f \right\rangle \mathrm{d}\nu_s \, \mathrm{d}s\right)^2
\end{aligned}$$

Proof: Mean-difference estimate

Transport interpolation

Goal: Find a good estimate for C in

$$\left(\mathbb{E}_{\nu_0}(f) - \mathbb{E}_{\nu_1}(f)\right)^2 \leq C \int |\nabla f|^2 \, \mathrm{d}\mu.$$

Step 2: Transport $(\Phi_s \in \mathrm{Diff}(\mathbb{R}^n, \mathbb{R}^n))_{s\in[0,1]}$ interpolating $(\Phi_s)_\sharp \nu_0 = \nu_s$

$$\begin{aligned}
\left(\int f \, \mathrm{d}\nu_0 - \int f \, \mathrm{d}\nu_1\right)^2 &= \left(\int \int_0^1 \frac{\mathrm{d}}{\mathrm{d}s}(f \circ \Phi_s) \, \mathrm{d}s \, \mathrm{d}\nu_0\right)^2 \\
&= \left(\int_0^1 \int \left\langle \dot{\Phi}_s, \nabla f \circ \Phi_s \right\rangle \mathrm{d}\nu_0 \, \mathrm{d}s\right)^2 \\
&= \left(\int_0^1 \int \left\langle \dot{\Phi}_s \circ \Phi_s^{-1}, \nabla f \right\rangle \frac{\mathrm{d}\nu_s}{\mathrm{d}\mu} \, \mathrm{d}\mu \, \mathrm{d}s\right)^2
\end{aligned}$$

Proof: Mean-difference estimate

Transport interpolation

Goal: Find a good estimate for C in

$$\left(\mathbb{E}_{\nu_0}(f) - \mathbb{E}_{\nu_1}(f)\right)^2 \le C \int |\nabla f|^2 \,\mathrm{d}\mu.$$

Step 2: Transport $(\Phi_s \in \mathrm{Diff}(\mathbb{R}^n, \mathbb{R}^n))_{s\in[0,1]}$ interpolating $(\Phi_s)_\sharp \nu_0 = \nu_s$

$$\begin{aligned}\left(\int f \,\mathrm{d}\nu_0 - \int f \,\mathrm{d}\nu_1\right)^2 &= \left(\int \int_0^1 \frac{\mathrm{d}}{\mathrm{d}s}(f \circ \Phi_s) \,\mathrm{d}s \,\mathrm{d}\nu_0\right)^2 \\ &= \left(\int_0^1 \int \left\langle \dot{\Phi}_s, \nabla f \circ \Phi_s \right\rangle \mathrm{d}\nu_0 \,\mathrm{d}s\right)^2 \\ &= \left(\int \int_0^1 \left\langle \dot{\Phi}_s \circ \Phi_s^{-1}, \nabla f \right\rangle \frac{\mathrm{d}\nu_s}{\mathrm{d}\mu} \,\mathrm{d}s \,\mathrm{d}\mu\right)^2\end{aligned}$$

Proof: Mean-difference estimate

Transport interpolation

Goal: Find a good estimate for C in

$$(\mathbb{E}_{\nu_0}(f) - \mathbb{E}_{\nu_1}(f))^2 \leq C \int |\nabla f|^2 \, \mathrm{d}\mu.$$

Step 2: Transport $(\Phi_s \in \mathrm{Diff}(\mathbb{R}^n, \mathbb{R}^n))_{s\in[0,1]}$ interpolating $(\Phi_s)_\sharp \nu_0 = \nu_s$

$$\begin{aligned}
\left(\int f \, \mathrm{d}\nu_0 - \int f \, \mathrm{d}\nu_1\right)^2 &= \left(\int \int_0^1 \frac{\mathrm{d}}{\mathrm{d}s}(f \circ \Phi_s) \, \mathrm{d}s \, \mathrm{d}\nu_0\right)^2 \\
&= \left(\int_0^1 \int \left\langle \dot{\Phi}_s, \nabla f \circ \Phi_s \right\rangle \mathrm{d}\nu_0 \, \mathrm{d}s\right)^2 \\
&= \left(\int \left\langle \int_0^1 \dot{\Phi}_s \circ \Phi_s^{-1} \, \frac{\mathrm{d}\nu_s}{\mathrm{d}\mu} \, \mathrm{d}s, \nabla f \right\rangle \mathrm{d}\mu\right)^2
\end{aligned}$$

Goal: Find a good estimate for C in

$$\left(\mathbb{E}_{\nu_0}(f) - \mathbb{E}_{\nu_1}(f)\right)^2 \leq C \int |\nabla f|^2 \, \mathrm{d}\mu.$$

Step 2: Transport $(\Phi_s \in \mathrm{Diff}(\mathbb{R}^n, \mathbb{R}^n))_{s\in[0,1]}$ interpolating $(\Phi_s)_\sharp \nu_0 = \nu_s$

$$\begin{aligned}
\left(\int f \, \mathrm{d}\nu_0 - \int f \, \mathrm{d}\nu_1\right)^2 &= \left(\int \int_0^1 \frac{\mathrm{d}}{\mathrm{d}s}(f \circ \Phi_s) \, \mathrm{d}s \, \mathrm{d}\nu_0\right)^2 \\
&= \left(\int_0^1 \int \left\langle \dot{\Phi}_s, \nabla f \circ \Phi_s \right\rangle \mathrm{d}\nu_0 \, \mathrm{d}s\right)^2 \\
&= \left(\int \left\langle \int_0^1 \dot{\Phi}_s \circ \Phi_s^{-1} \, \frac{\mathrm{d}\nu_s}{\mathrm{d}\mu} \, \mathrm{d}s, \nabla f \right\rangle \mathrm{d}\mu\right)^2 \\
&\leq \int \left| \int_0^1 \dot{\Phi}_s \circ \Phi_s^{-1} \, \frac{\mathrm{d}\nu_s}{\mathrm{d}\mu} \, \mathrm{d}s \right|^2 \mathrm{d}\mu \quad \int |\nabla f|^2 \, \mathrm{d}\mu
\end{aligned}$$

Proof: Mean-difference estimate

Sideremark: Weighted transport distance

Definition

For $\nu_0, \nu_1 \ll \mu$ define the weighted transport distance by

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

$(\Phi_s)_{s\in[0,1]}$ is absolutely continuous in s: $(\Phi_s)_\sharp \nu_0 = \nu_s$.

Mean-difference revisited: Identify $\int |\nabla f|^2 \, d\mu = \|f\|^2_{\dot{H}^1(\mu)}$, then

$$\left(\int f \, d\nu_0 - \int f \, d\nu_1 \right)^2 = \left({}_{\dot{H}^{-1}(\mu)}\langle \nu_0 - \nu_1, f \rangle_{\dot{H}^1(\mu)} \right)^2$$
$$\leq \mathcal{T}_\mu^2(\nu_0, \nu_1) \, \|f\|^2_{\dot{H}^1(\mu)}.$$

Indeed, it holds: $\mathcal{T}_\mu^2(\nu_0, \nu_1) = \|\nu_0 - \nu_1\|^2_{\dot{H}^{-1}(\mu)}$.

Proof: Mean-difference estimate

Sideremark: Weighted transport distance

Definition

For $\nu_0, \nu_1 \ll \mu$ define the weighted transport distance by

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

$(\Phi_s)_{s\in[0,1]}$ is absolutely continuous in s: $(\Phi_s)_\sharp \nu_0 = \nu_s$.

Mean-difference revisited: Identify $\int |\nabla f|^2 \, \mathrm{d}\mu = \|f\|^2_{\dot{H}^1(\mu)}$, then

$$\left(\int f \, \mathrm{d}\nu_0 - \int f \, \mathrm{d}\nu_1 \right)^2 = \left({}_{\dot{H}^{-1}(\mu)}\langle \nu_0 - \nu_1, f \rangle_{\dot{H}^1(\mu)} \right)^2$$
$$\leq \mathcal{T}_\mu^2(\nu_0, \nu_1) \, \|f\|^2_{\dot{H}^1(\mu)} .$$

Indeed, it holds: $\mathcal{T}_\mu^2(\nu_0, \nu_1) = \|\nu_0 - \nu_1\|^2_{\dot{H}^{-1}(\mu)}$.

Proof: Mean-difference estimate

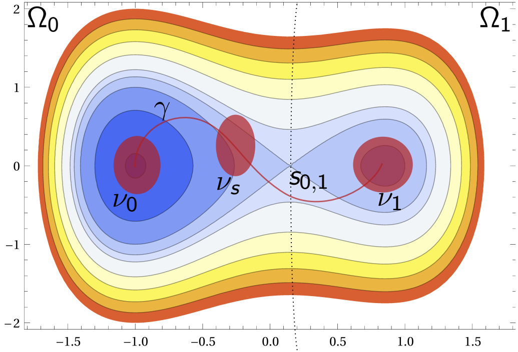
Construction of transport interpolation

Step 3: Ansatz Φ_s such that $\nu_s = (\Phi_s)_\sharp \nu_0 = \mathcal{N}(\gamma_s, \Sigma_s) \llcorner B_{\sqrt{\varepsilon}}(\gamma_s)$

(1) optimize $\gamma \Rightarrow$ passage of saddle $\gamma_{\tau^*} = s_{0,1}$

(2) optimize $\dot\gamma_{\tau^*} \Rightarrow$ direction of eigenvector to $\lambda^-(\nabla^2 H(s_{0,1}))$

(3) optimize $\Sigma_{\tau^*} \Rightarrow \Sigma_{\tau^*}^{-1} = \nabla^2 H(s_{0,1})$ on stable manifold of $s_{0,1}$



$$(\mathbb{E}_{\nu_0} f - \mathbb{E}_{\nu_1} f)^2 \lesssim \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}} \frac{2\pi\varepsilon\sqrt{|\det \nabla^2 H(s_{0,1})|}}{|\lambda^-(\nabla^2 H(s_{0,1}))|} e^{\varepsilon^{-1} H(s_{0,1})} \int |\nabla f|^2 \,\mathrm{d}\mu.$$

Proof: Mean-difference estimate

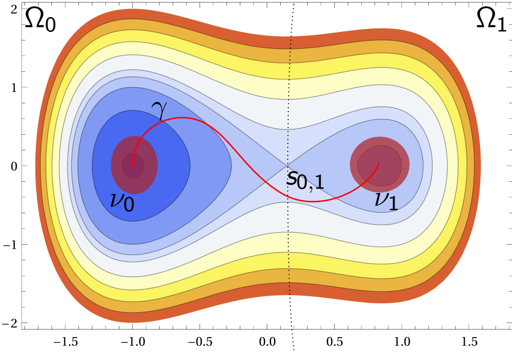
Construction of transport interpolation

Step 3: Ansatz Φ_s such that $\nu_s = (\Phi_s)_\sharp \nu_0 = \mathcal{N}(\gamma_s, \Sigma_s) \llcorner B_{\sqrt{\varepsilon}}(\gamma_s)$

(1) optimize $\gamma \Rightarrow$ passage of saddle $\gamma_{\tau^*} = s_{0,1}$

(2) optimize $\dot{\gamma}_{\tau^*} \Rightarrow$ direction of eigenvector to $\lambda^-(\nabla^2 H(s_{0,1}))$

(3) optimize $\Sigma_{\tau^*} \Rightarrow \Sigma_{\tau^*}^{-1} = \nabla^2 H(s_{0,1})$ on stable manifold of $s_{0,1}$



$$\left(\mathbb{E}_{\nu_0} f - \mathbb{E}_{\nu_1} f\right)^2 \lesssim \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}} \frac{2\pi\varepsilon\sqrt{|\det \nabla^2 H(s_{0,1})|}}{|\lambda^-(\nabla^2 H(s_{0,1}))|} e^{\varepsilon^{-1} H(s_{0,1})} \int |\nabla f|^2 \, \mathrm{d}\mu.$$

Proof: Mean-difference estimate

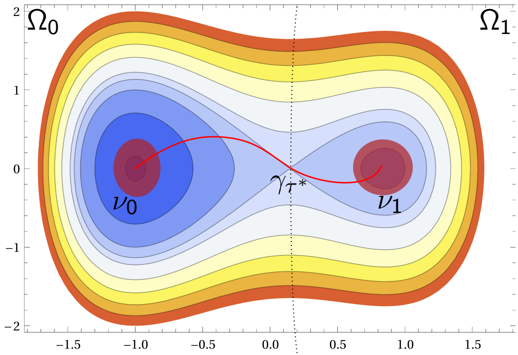
Construction of transport interpolation

Step 3: Ansatz Φ_s such that $\nu_s = (\Phi_s)_\sharp \nu_0 = \mathcal{N}(\gamma_s, \Sigma_s) \llcorner B_{\sqrt{\varepsilon}}(\gamma_s)$

(1) optimize $\gamma \Rightarrow$ passage of saddle $\gamma_{\tau^*} = s_{0,1}$

(2) optimize $\dot\gamma_{\tau^*} \Rightarrow$ direction of eigenvector to $\lambda^-(\nabla^2 H(s_{0,1}))$

(3) optimize $\Sigma_{\tau^*} \Rightarrow \Sigma_{\tau^*}^{-1} = \nabla^2 H(s_{0,1})$ on stable manifold of $s_{0,1}$



$$(\mathbb{E}_{\nu_0} f - \mathbb{E}_{\nu_1} f)^2 \lesssim \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}} \frac{2\pi\varepsilon\sqrt{|\det \nabla^2 H(s_{0,1})|}}{|\lambda^-(\nabla^2 H(s_{0,1}))|} e^{\varepsilon^{-1} H(s_{0,1})} \int |\nabla f|^2 \, d\mu.$$

Proof: Mean-difference estimate

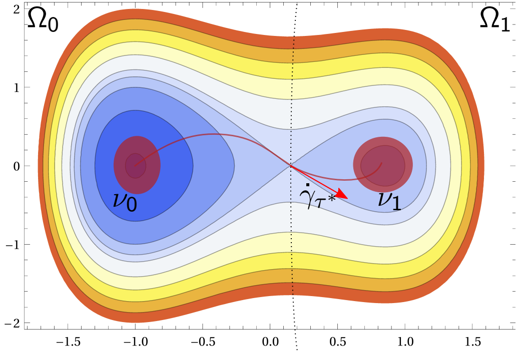
Construction of transport interpolation

Step 3: Ansatz Φ_s such that $\nu_s = (\Phi_s)_\sharp \nu_0 = \mathcal{N}(\gamma_s, \Sigma_s) \llcorner B_{\sqrt{\varepsilon}}(\gamma_s)$

(1) optimize $\gamma \Rightarrow$ passage of saddle $\gamma_{\tau^*} = s_{0,1}$

(2) optimize $\dot{\gamma}_{\tau^*} \Rightarrow$ direction of eigenvector to $\lambda^-(\nabla^2 H(s_{0,1}))$

(3) optimize $\Sigma_{\tau^*} \Rightarrow \Sigma_{\tau^*}^{-1} = \nabla^2 H(s_{0,1})$ on stable manifold of $s_{0,1}$



$$(\mathbb{E}_{\nu_0} f - \mathbb{E}_{\nu_1} f)^2 \lesssim \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}} \frac{2\pi\varepsilon\sqrt{|\det \nabla^2 H(s_{0,1})|}}{|\lambda^-(\nabla^2 H(s_{0,1}))|} e^{\varepsilon^{-1} H(s_{0,1})} \int |\nabla f|^2 \, d\mu.$$

Proof: Mean-difference estimate

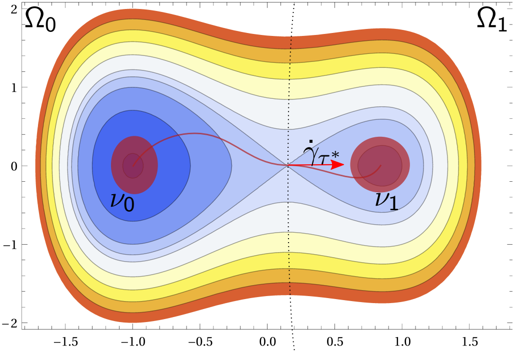
Construction of transport interpolation

Step 3: Ansatz Φ_s such that $\nu_s = (\Phi_s)_\sharp \nu_0 = \mathcal{N}(\gamma_s, \Sigma_s) \llcorner B_{\sqrt{\varepsilon}}(\gamma_s)$

(1) optimize $\gamma \Rightarrow$ passage of saddle $\gamma_{\tau^*} = s_{0,1}$

(2) optimize $\dot\gamma_{\tau^*} \Rightarrow$ direction of eigenvector to $\lambda^-(\nabla^2 H(s_{0,1}))$

(3) optimize $\Sigma_{\tau^*} \Rightarrow \Sigma_{\tau^*}^{-1} = \nabla^2 H(s_{0,1})$ on stable manifold of $s_{0,1}$



$$\left(\mathbb{E}_{\nu_0} f - \mathbb{E}_{\nu_1} f\right)^2 \lesssim \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}} \, \frac{2\pi\varepsilon\sqrt{|\det \nabla^2 H(s_{0,1})|}}{|\lambda^-(\nabla^2 H(s_{0,1}))|} \, e^{\varepsilon^{-1} H(s_{0,1})} \int |\nabla f|^2 \, d\mu.$$

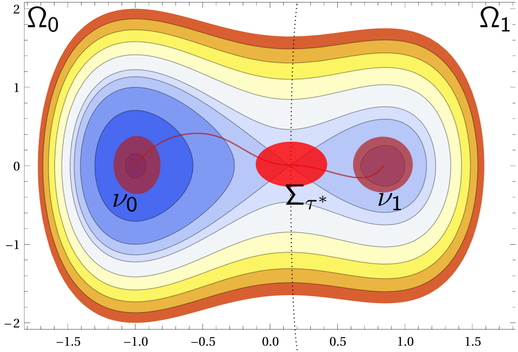
Construction of transport interpolation

Step 3: Ansatz Φ_s such that $\nu_s = (\Phi_s)_\sharp \nu_0 = \mathcal{N}(\gamma_s, \Sigma_s) \llcorner B_{\sqrt{\varepsilon}}(\gamma_s)$

(1) optimize $\gamma \Rightarrow$ passage of saddle $\gamma_{\tau^*} = s_{0,1}$

(2) optimize $\dot{\gamma}_{\tau^*} \Rightarrow$ direction of eigenvector to $\lambda^-(\nabla^2 H(s_{0,1}))$

(3) optimize $\Sigma_{\tau^*} \Rightarrow \Sigma_{\tau^*}^{-1} = \nabla^2 H(s_{0,1})$ on stable manifold of $s_{0,1}$



$$(\mathbb{E}_{\nu_0} f - \mathbb{E}_{\nu_1} f)^2 \lesssim \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}} \frac{2\pi\varepsilon\sqrt{|\det \nabla^2 H(s_{0,1})|}}{|\lambda^-(\nabla^2 H(s_{0,1}))|} e^{\varepsilon^{-1} H(s_{0,1})} \int |\nabla f|^2 \, d\mu.$$

Proof: Mean-difference estimate

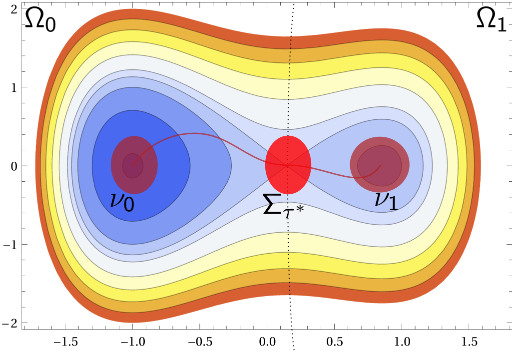
Construction of transport interpolation

Step 3: Ansatz Φ_s such that $\nu_s = (\Phi_s)_\sharp \nu_0 = \mathcal{N}(\gamma_s, \Sigma_s) \llcorner B_{\sqrt{\varepsilon}}(\gamma_s)$

(1) optimize $\gamma \Rightarrow$ passage of saddle $\gamma_{\tau^*} = s_{0,1}$

(2) optimize $\dot{\gamma}_{\tau^*} \Rightarrow$ direction of eigenvector to $\lambda^-(\nabla^2 H(s_{0,1}))$

(3) optimize $\Sigma_{\tau^*} \Rightarrow \Sigma_{\tau^*}^{-1} = \nabla^2 H(s_{0,1})$ on stable manifold of $s_{0,1}$



$$(\mathbb{E}_{\nu_0} f - \mathbb{E}_{\nu_1} f)^2 \lesssim \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}} \frac{2\pi\varepsilon\sqrt{|\det \nabla^2 H(s_{0,1})|}}{|\lambda^-(\nabla^2 H(s_{0,1}))|} e^{\varepsilon^{-1} H(s_{0,1})} \int |\nabla f|^2 \, d\mu.$$

Proof: Mean-difference estimate

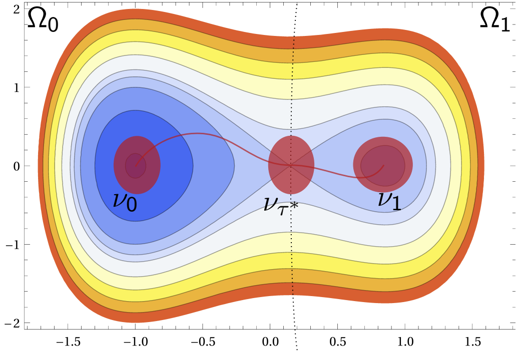
Construction of transport interpolation

Step 3: Ansatz Φ_s such that $\nu_s = (\Phi_s)_\sharp \nu_0 = \mathcal{N}(\gamma_s, \Sigma_s) \llcorner B_{\sqrt{\varepsilon}}(\gamma_s)$

(1) optimize $\gamma \Rightarrow$ passage of saddle $\gamma_{\tau^*} = s_{0,1}$

(2) optimize $\dot{\gamma}_{\tau^*} \Rightarrow$ direction of eigenvector to $\lambda^-(\nabla^2 H(s_{0,1}))$

(3) optimize $\Sigma_{\tau^*} \Rightarrow \Sigma_{\tau^*}^{-1} = \nabla^2 H(s_{0,1})$ on stable manifold of $s_{0,1}$



$$\left(\mathbb{E}_{\nu_0} f - \mathbb{E}_{\nu_1} f\right)^2 \lesssim \frac{Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}} \, \frac{2\pi\varepsilon\sqrt{|\det \nabla^2 H(s_{0,1})|}}{|\lambda^-(\nabla^2 H(s_{0,1}))|} \, e^{\varepsilon^{-1} H(s_{0,1})} \int |\nabla f|^2 \, \mathrm{d}\mu.$$

Discrete state space

reversible Markov chain

Markov chain $\{X(t)\}_{t\geq 0}$ on a finite state space $\mathcal{S}$ with generator

$$(Lf)(x) = \sum_{y\in\mathcal{S}} p(x,y)\,(f(x) - f(y))$$

and with detailed balance

$$\mu(x)p(x,y) = \mu(y)p(y,x) \qquad \text{forall } x,y \in \mathcal{S}.$$

For any $A \subset \mathcal{S}$: first hitting time of A

$$\tau_A = \inf\left\{t > 0 \,\middle|\, X(t) \in A\right\}.$$

Discrete state space

reversible Markov chain

Markov chain $\{X(t)\}_{t\geq 0}$ on a finite state space $\mathcal{S}$ with generator

$$(Lf)(x) = \sum_{y\in\mathcal{S}} p(x,y)\,(f(x) - f(y))$$

and with detailed balance

$$\mu(x)p(x,y) = \mu(y)p(y,x) \qquad \text{forall } x,y \in \mathcal{S}.$$

For any $A \subset \mathcal{S}$: first hitting time of A

$$\tau_A = \inf\left\{t > 0 \,\middle|\, X(t) \in A\right\}.$$

Metastable points and valleys

Definition (Metastable points [Bovier])

Let $\mathcal{M} = \{a_1, \ldots, a_k\} \subset \mathcal{S}$. The set $\mathcal{M}$ is called a *set of metastable points*, if there exists a $\varrho > 0$ such that

$$\frac{\max_{a \in \mathcal{M}} \mathbb{P}_a\left[\tau_{\mathcal{M}\setminus a} < \tau_a\right]}{\min_{x \in \mathcal{S}\setminus\mathcal{M}} \mathbb{P}_x\left[\tau_{\mathcal{M}} < \tau_x\right]} \leq \varrho \ll \frac{1}{|\mathcal{S}|} \qquad \varepsilon := \varrho|\mathcal{S}| \ll 1 \quad \text{as } |\mathcal{S}| \to \infty.$$

Definition (Metastable partition)

Metastable points $\mathcal{M} \subset \mathcal{S}$ give rise to a *metastable partition* $\{\Omega_a\}_{a \in \mathcal{M}}$ if:

i) $\forall a \in \mathcal{M} : \Omega_a \subseteq \{x \in \mathcal{S} \mid \mathbb{P}_x[\tau_a \leq \tau_x] \geq \mathbb{P}_x[\tau_{\mathcal{M}\setminus a} \leq \tau_x]\}$;

ii) $\biguplus_{a \in \mathcal{M}} \Omega_a = \mathcal{S}$;

iii) $\forall a \in \mathcal{M}\ \forall x \in \Omega_a : \mu[x] \leq \mu[a]$.

Metastable points and valleys

Definition (Metastable points [Bovier])

Let $\mathcal{M} = \{a_1, \ldots, a_k\} \subset \mathcal{S}$. The set $\mathcal{M}$ is called a *set of metastable points*, if there exists a $\varrho > 0$ such that

$$\frac{\max_{a\in\mathcal{M}} \mathbb{P}_a\left[\tau_{\mathcal{M}\setminus a} < \tau_a\right]}{\min_{x\in\mathcal{S}\setminus\mathcal{M}} \mathbb{P}_x\left[\tau_{\mathcal{M}} < \tau_x\right]} \leq \varrho \ll \frac{1}{|\mathcal{S}|} \qquad \varepsilon := \varrho|\mathcal{S}| \ll 1 \quad \text{as } |\mathcal{S}| \to \infty.$$

Definition (Metastable partition)

Metastable points $\mathcal{M} \subset \mathcal{S}$ give rise to a *metastable partition* $\{\Omega_a\}_{a\in\mathcal{M}}$ if:

i) $\forall a \in \mathcal{M} : \Omega_a \subseteq \{x \in \mathcal{S} \mid \mathbb{P}_x[\tau_a \leq \tau_x] \geq \mathbb{P}_x[\tau_{\mathcal{M}\setminus a} \leq \tau_x]\}$;

ii) $\biguplus_{a\in\mathcal{M}} \Omega_a = \mathcal{S}$;

iii) $\forall a \in \mathcal{M}\ \forall x \in \Omega_a : \mu[x] \leq \mu[a]$.

Metastable points and valleys

Definition (Metastable points [Bovier])

Let $\mathcal{M} = \{a_1, \ldots, a_k\} \subset \mathcal{S}$. The set $\mathcal{M}$ is called a *set of metastable points*, if there exists a $\varrho > 0$ such that

$$\frac{\max_{a\in\mathcal{M}} \mathbb{P}_a\left[\tau_{\mathcal{M}\setminus a} < \tau_a\right]}{\min_{x\in\mathcal{S}\setminus\mathcal{M}} \mathbb{P}_x\left[\tau_{\mathcal{M}} < \tau_x\right]} \leq \varrho \ll \frac{1}{|\mathcal{S}|} \qquad \varepsilon := \varrho|\mathcal{S}| \ll 1 \quad \text{as } |\mathcal{S}| \to \infty.$$

Definition (Metastable partition)

Metastable points $\mathcal{M} \subset \mathcal{S}$ give rise to a *metastable partition* $\{\Omega_a\}_{a\in\mathcal{M}}$ if:

i) $\forall a \in \mathcal{M} : \Omega_a \subseteq \{x \in \mathcal{S} \mid \mathbb{P}_x[\tau_a \leq \tau_x] \geq \mathbb{P}_x[\tau_{\mathcal{M}\setminus a} \leq \tau_x]\}$;

ii) $\biguplus_{a\in\mathcal{M}} \Omega_a = \mathcal{S}$;

iii) $\forall a \in \mathcal{M}\ \forall x \in \Omega_a : \mu[x] \leq \mu[a]$.

Splitting

induced by metastable partition

Conditional measures for $a \in \mathcal{M}$

$$\mu_a(x) := \frac{\mathbb{1}_{\Omega_a}(x)}{\mu(\Omega_a)}\mu(x)$$

Marginal measure

$$\bar{\mu} := \sum_{a \in \mathcal{M}} Z_a \delta_a \qquad \text{with} \qquad Z_a := \mu(\Omega_a).$$

$$\begin{aligned}
\operatorname{var}_\mu(f) &= Z_a \operatorname{var}_{\mu_a}(f) + Z_b \operatorname{var}_{\mu_b}(f) + Z_a Z_b (\mathbb{E}_{\mu_a}(f) - \mathbb{E}_{\mu_b}(f))^2 \\
\operatorname{Ent}_\mu(f^2) &\le Z_a \operatorname{Ent}_{\mu_0}(f^2) + Z_1 \operatorname{Ent}_{\mu_b}(f^2) \\
&\quad + \frac{Z_a Z_b}{\Lambda(Z_a, Z_b)} \left(\operatorname{var}_{\mu_a}(f) + \operatorname{var}_{\mu_b}(f) + (\mathbb{E}_{\mu_a}(f) - \mathbb{E}_{\mu_b}(f))^2 \right),
\end{aligned}$$

where $\Lambda(Z_a, Z_b) = \frac{Z_a - Z_b}{\log Z_a - \log Z_b}$ is the logarithmic mean.

Splitting

induced by metastable partition

Conditional measures for $a \in \mathcal{M}$

$$\mu_a(x) := \frac{\mathbb{1}_{\Omega_a}(x)}{\mu(\Omega_a)}\mu(x)$$

Marginal measure

$$\bar{\mu} := \sum_{a\in\mathcal{M}} Z_a \delta_a \qquad \text{with} \qquad Z_a := \mu(\Omega_a).$$

$$\begin{aligned}
\operatorname{var}_\mu(f) &= Z_a \operatorname{var}_{\mu_a}(f) + Z_b \operatorname{var}_{\mu_b}(f) \;+\; Z_a Z_b \big(\mathbb{E}_{\mu_a}(f) - \mathbb{E}_{\mu_b}(f)\big)^2 \\
\operatorname{Ent}_\mu(f^2) &\le Z_a \operatorname{Ent}_{\mu_0}(f^2) + Z_1 \operatorname{Ent}_{\mu_b}(f^2) \\
&\quad + \frac{Z_a Z_b}{\Lambda(Z_a, Z_b)} \Big(\operatorname{var}_{\mu_a}(f) + \operatorname{var}_{\mu_b}(f) + \big(\mathbb{E}_{\mu_a}(f) - \mathbb{E}_{\mu_b}(f)\big)^2\Big),
\end{aligned}$$

where $\Lambda(Z_a, Z_b) = \frac{Z_a - Z_b}{\log Z_a - \log Z_b}$ is the logarithmic mean.

Capacities

Definition (Harmonic functions and capacities)

Let $A, B \subset \mathcal{S}$ be disjoint, then the harmonic function between A to B is defined by

$$\begin{cases} (Lh_{A,B}) = 0 & , x \in \mathcal{S} \setminus (A \cup B) \\ h_{A,B}(x) = \mathbb{1}_A(x) & , x \in A \cup B. \end{cases}$$

The capacity of a capacitor (A, B) is defined as

$$\mathrm{cap}(A, B) = (h_{A,B}, -Lh_{A,B})_\mu .$$

Probabilistic interpretation

$$A, B \subset \mathcal{S} \qquad h_{A,B}(x) = \mathbb{P}_x(\tau_A < \tau_B), \qquad x \in \mathcal{S} \setminus (A \cup B)$$

$$a, b \in \mathcal{S} \qquad \mathrm{cap}(a, b) = \mu(a)\, \mathbb{P}_a(\tau_b < \tau_a) = \mu(b)\, \mathbb{P}_b(\tau_a < \tau_b).$$

Capacities

Definition (Harmonic functions and capacities)

Let $A, B \subset \mathcal{S}$ be disjoint, then the harmonic function between A to B is defined by

$$\begin{cases} (Lh_{A,B}) = 0 & , x \in \mathcal{S} \setminus (A \cup B) \\ h_{A,B}(x) = \mathbb{1}_A(x) & , x \in A \cup B. \end{cases}$$

The capacity of a capacitor (A, B) is defined as

$$\operatorname{cap}(A, B) = \left(h_{A,B}, -Lh_{A,B}\right)_\mu .$$

Probabilistic interpretation

$$A, B \subset \mathcal{S} \qquad h_{A,B}(x) = \mathbb{P}_x(\tau_A < \tau_B), \qquad x \in \mathcal{S} \setminus (A \cup B)$$

$$a, b \in \mathcal{S} \qquad \operatorname{cap}(a, b) = \mu(a)\,\mathbb{P}_a(\tau_b < \tau_a) = \mu(b)\,\mathbb{P}_b(\tau_a < \tau_b).$$

PI and LSI for Markov chains

Theorem ([S., Slowik '13])

For a metastable Markov chain with $\mathcal{M} = \{a, b\}$ *holds*

$$\frac{1}{\lambda} = \frac{1}{2}\frac{Z_a Z_b}{\operatorname{cap}(a,b)}\left(1 + o(\varrho|\mathcal{S}|)\right).$$

Moreover, under assuming a good local LSI, i.e. $\alpha_{\mathrm{loc}}^{-1} = o(\varrho|\mathcal{S}|)$, *holds*

$$\frac{1}{\alpha} = \frac{Z_a Z_b}{\Lambda(Z_a, Z_b)}\frac{1}{\operatorname{cap}(a,b)}\left(1 + o(\varrho|\mathcal{S}|)\right) = \frac{1}{\Lambda(Z_a, Z_b)}\frac{1}{\varrho}\left(1 + o(\varrho|\mathcal{S}|)\right)$$

Ingredients:

- local PI: Donsker-Varadhan variational characterization
  $\Rightarrow$ precursor of Lyapunov technique of continuous case
  Needs refinement for local LSI!
- mean-difference estimate via H^{-1}-norms in discrete setting

PI and LSI for Markov chains

Theorem ([S., Slowik '13])

For a metastable Markov chain with $\mathcal{M} = \{a, b\}$ *holds*

$$\frac{1}{\lambda} = \frac{1}{2} \frac{Z_a Z_b}{\mathrm{cap}(a,b)} \left(1 + o(\varrho|\mathcal{S}|)\right).$$

Moreover, under assuming a good local LSI, i.e. $\alpha_{\mathrm{loc}}^{-1} = o(\varrho|\mathcal{S}|)$, *holds*

$$\frac{1}{\alpha} = \frac{Z_a Z_b}{\Lambda(Z_a, Z_b)} \frac{1}{\mathrm{cap}(a,b)} \left(1 + o(\varrho|\mathcal{S}|)\right) = \frac{1}{\Lambda(Z_a, Z_b)} \frac{1}{\varrho} \left(1 + o(\varrho|\mathcal{S}|)\right)$$

Ingredients:

- local PI: Donsker-Varadhan variational characterization
  $\Rightarrow$ precursor of Lyapunov technique of continuous case
  Needs refinement for local LSI!
- mean-difference estimate via H^{-1}-norms in discrete setting

PI and LSI for Markov chains

Theorem ([S., Slowik '13])

For a metastable Markov chain with $\mathcal{M} = \{a, b\}$ holds

$$\frac{1}{\lambda} = \frac{1}{2} \frac{Z_a Z_b}{\operatorname{cap}(a,b)} \left(1 + o(\varrho|\mathcal{S}|)\right).$$

Moreover, under assuming a good local LSI, i.e. $\alpha_{\text{loc}}^{-1} = o(\varrho|\mathcal{S}|)$, *holds*

$$\frac{1}{\alpha} = \frac{Z_a Z_b}{\Lambda(Z_a, Z_b)} \frac{1}{\operatorname{cap}(a,b)} \left(1 + o(\varrho|\mathcal{S}|)\right) = \frac{1}{\Lambda(Z_a, Z_b)} \frac{1}{\varrho} \left(1 + o(\varrho|\mathcal{S}|)\right)$$

Ingredients:

- local PI: Donsker-Varadhan variational characterization
  $\Rightarrow$ precursor of Lyapunov technique of continuous case
  Needs refinement for local LSI!
- mean-difference estimate via H^{-1}-norms in discrete setting

Mean-Difference

Negative Sobolev norms in discrete setting

Identify Dirichlet form with H^1-norm

$$\|f\|^2_{H^1(\mu)} := (f, -Lf)_\mu = \frac{1}{2} \sum_{x,y \in \mathcal{S}} \mu(x) p(x,y) \left(f(x) - f(y) \right)^2 .$$

Negative Sobolev norm as dual norm

$$\|f\|^2_{H^{-1}(\mu)} := \sup_{g \in H^1(\mu)} \left(2(f,g)_\mu - \|g\|^2_{H^1(\mu)} \right) .$$

Mean-difference becomes for two probability measure ν_A, ν_B

$$\left| \mathbb{E}_{\nu_A}(f) - \mathbb{E}_{\nu_B}(f) \right| = \left| \left(f, \frac{\nu_A}{\mu} - \frac{\nu_B}{\mu} \right)_\mu \right| \leq \|f\|_{H^1(\mu)} \left\| \frac{\nu_A}{\mu} - \frac{\nu_B}{\mu} \right\|_{H^{-1}(\mu)} .$$

Mean-Difference

Negative Sobolev norms in discrete setting

Identify Dirichlet form with H^1-norm

$$\|f\|^2_{H^1(\mu)} := (f, -Lf)_\mu = \frac{1}{2} \sum_{x,y\in\mathcal{S}} \mu(x)p(x,y)\left(f(x) - f(y)\right)^2 .$$

Negative Sobolev norm as dual norm

$$\|f\|^2_{H^{-1}(\mu)} := \sup_{g\in H^1(\mu)} \left(2(f,g)_\mu - \|g\|^2_{H^1(\mu)}\right) .$$

Mean-difference becomes for two probability measure ν_A, ν_B

$$\left|\mathbb{E}_{\nu_A}(f) - \mathbb{E}_{\nu_B}(f)\right| = \left|\left(f, \frac{\nu_A}{\mu} - \frac{\nu_B}{\mu}\right)_\mu\right| \leq \|f\|_{H^1(\mu)} \left\|\frac{\nu_A}{\mu} - \frac{\nu_B}{\mu}\right\|_{H^{-1}(\mu)} .$$

Mean-Difference

Identify Dirichlet form with H^1-norm

$$\|f\|_{H^1(\mu)}^2 := (f, -Lf)_\mu = \frac{1}{2} \sum_{x,y\in\mathcal{S}} \mu(x)p(x,y)\left(f(x) - f(y)\right)^2 .$$

Negative Sobolev norm as dual norm

$$\|f\|_{H^{-1}(\mu)}^2 := \sup_{g\in H^1(\mu)} \left(2(f,g)_\mu - \|g\|_{H^1(\mu)}^2\right) .$$

Mean-difference becomes for two probability measure ν_A, ν_B

$$\left|\mathbb{E}_{\nu_A}(f) - \mathbb{E}_{\nu_B}(f)\right| = \left|\left(f, \frac{\nu_A}{\mu} - \frac{\nu_B}{\mu}\right)_\mu\right| \leq \|f\|_{H^1(\mu)} \left\|\frac{\nu_A}{\mu} - \frac{\nu_B}{\mu}\right\|_{H^{-1}(\mu)} .$$

Discrete flows

Definition (Flow)

flow: an antisymmetric map $\varphi : \mathcal{S} \times \mathcal{S} \to \mathbb{R}$ with $\sum_{y \in \mathcal{S}} \varphi(x, y) = 0$ for all $x \in \mathcal{S}$

divergence: $\operatorname{div} \varphi(x) := \sum_y (\varphi(x, y) - \varphi(y, x))$

form: $\mathcal{E}(\varphi) := \frac{1}{2} \sum_{x,y \in \mathcal{S}} \frac{1}{\mu(x)\, p(x,y)} |\varphi(x, y)|^2.$

Analog of weighted transport distance

Lemma (Representation of H^{-1}-norm)

For two probability measure ν_a and ν_b holds

$$\left\| \frac{\nu_a}{\mu} - \frac{\nu_b}{\mu} \right\|^2_{H^{-1}(\mu)} = \min \left\{ \mathcal{E}(\varphi) : \operatorname{div} \varphi(x) = \nu_a(x) - \nu_b(x), \forall x \in \mathcal{S} \right\}.$$

Goal: Construct flow providing a good upper bound for mean-difference.

Discrete flows

Definition (Flow)

flow: an antisymmetric map $\varphi : \mathcal{S} \times \mathcal{S} \to \mathbb{R}$ with $\sum_{y \in \mathcal{S}} \varphi(x,y) = 0$ for all $x \in \mathcal{S}$

divergence: $\operatorname{div} \varphi(x) := \sum_y (\varphi(x,y) - \varphi(y,x))$

form: $\mathcal{E}(\varphi) := \frac{1}{2} \sum_{x,y \in \mathcal{S}} \frac{1}{\mu(x)\, p(x,y)} |\varphi(x,y)|^2.$

Analog of weighted transport distance

Lemma (Representation of H^{-1}-norm)

For two probability measure ν_a and ν_b holds

$$\left\| \frac{\nu_a}{\mu} - \frac{\nu_b}{\mu} \right\|^2_{H^{-1}(\mu)} = \min \left\{ \mathcal{E}(\varphi) : \operatorname{div} \varphi(x) = \nu_a(x) - \nu_b(x), \forall x \in \mathcal{S} \right\}.$$

Goal: Construct flow providing a good upper bound for mean-difference.

Discrete flows

Definition (Flow)

flow: an antisymmetric map $\varphi : \mathcal{S} \times \mathcal{S} \to \mathbb{R}$ with $\sum_{y \in \mathcal{S}} \varphi(x,y) = 0$ for all $x \in \mathcal{S}$

divergence: $\operatorname{div} \varphi(x) := \sum_y (\varphi(x,y) - \varphi(y,x))$

form: $\mathcal{E}(\varphi) := \frac{1}{2} \sum_{x,y \in \mathcal{S}} \frac{1}{\mu(x)\, p(x,y)} |\varphi(x,y)|^2$.

Analog of weighted transport distance

Lemma (Representation of H^{-1}-norm)

For two probability measure ν_a and ν_b holds

$$\left\| \frac{\nu_a}{\mu} - \frac{\nu_b}{\mu} \right\|^2_{H^{-1}(\mu)} = \min \left\{ \mathcal{E}(\varphi) : \operatorname{div} \varphi(x) = \nu_a(x) - \nu_b(x), \forall x \in \mathcal{S} \right\}.$$

Goal: Construct flow providing a good upper bound for mean-difference.

Discrete flows

Definition (Flow)

flow: an antisymmetric map $\varphi : \mathcal{S} \times \mathcal{S} \to \mathbb{R}$ with $\sum_{y \in \mathcal{S}} \varphi(x,y) = 0$ for all $x \in \mathcal{S}$

divergence: $\operatorname{div} \varphi(x) := \sum_y (\varphi(x,y) - \varphi(y,x))$

form: $\mathcal{E}(\varphi) := \frac{1}{2} \sum_{x,y \in \mathcal{S}} \frac{1}{\mu(x)\, p(x,y)} |\varphi(x,y)|^2$.

Analog of weighted transport distance

Lemma (Representation of H^{-1}-norm)

For two probability measure ν_a and ν_b holds

$$\left\| \frac{\nu_a}{\mu} - \frac{\nu_b}{\mu} \right\|^2_{H^{-1}(\mu)} = \min \left\{ \mathcal{E}(\varphi) : \operatorname{div} \varphi(x) = \nu_a(x) - \nu_b(x), \forall x \in \mathcal{S} \right\}.$$

Goal: Construct flow providing a good upper bound for mean-difference.

Canonical flow

Harmonic flows between points $z, z' \in \mathcal{S}$:

$$\varphi_{z,z'}(x,y) := \frac{\mu(x)\,p(x,y)}{\operatorname{cap}(z,z')}\left(h_{z,z'}(x) - h_{z,z'}(y)\right) \qquad \begin{cases} Lh_{z,z'} = 0, \text{ on } \mathcal{S} \setminus \{z,z'\} \\ h_{z,z'}(z) = 1,\, h_{z,z'}(z') = 0. \end{cases}$$

Then

$$\mathcal{E}(\varphi_{z,z'}) = \frac{1}{2}\sum_{x,y} \frac{\mu(x)\,p(x,y)}{\operatorname{cap}(z,z')^2}\left(h_{z,z'}(x) - h_{z,z'}(y)\right)^2 = \frac{1}{\operatorname{cap}(z,z')}.$$

Lemma

For two probability measures ν_A, ν_B let the canonical flow be defined by

$$\varphi_{\nu_A,\nu_B}(x,y) := \sum_{z\in A, z'\in B} \nu_A(z)\nu_B(z')\frac{\mu(x)p(x,y)}{\operatorname{cap}(z,z')}\left(h_{z,z'}(x) - h_{z,z'}(y)\right).$$

Then φ_{ν_A,ν_B} is a flow and it holds

$$\mathcal{E}(\varphi_{\nu_A,\nu_B}) \le \sum_{x\in A, y\in B} \frac{\nu_A(x)\nu_B(y)}{\operatorname{cap}(x,y)}.$$

Canonical flow

Harmonic flows between points $z, z' \in \mathcal{S}$:

$$\varphi_{z,z'}(x,y) := \frac{\mu(x)\,p(x,y)}{\operatorname{cap}(z,z')}\left(h_{z,z'}(x) - h_{z,z'}(y)\right) \qquad \begin{cases} Lh_{z,z'} = 0, \text{ on } \mathcal{S} \setminus \{z,z'\} \\ h_{z,z'}(z) = 1, h_{z,z'}(z') = 0. \end{cases}$$

Then

$$\mathcal{E}(\varphi_{z,z'}) = \frac{1}{2}\sum_{x,y} \frac{\mu(x)\,p(x,y)}{\operatorname{cap}(z,z')^2}\left(h_{z,z'}(x) - h_{z,z'}(y)\right)^2 = \frac{1}{\operatorname{cap}(z,z')}.$$

Lemma

For two probability measures ν_A, ν_B let the canonical flow be defined by

$$\varphi_{\nu_A,\nu_B}(x,y) := \sum_{z\in A, z'\in B} \nu_A(z)\nu_B(z')\frac{\mu(x)p(x,y)}{\operatorname{cap}(z,z')}\left(h_{z,z'}(x) - h_{z,z'}(y)\right).$$

Then φ_{ν_A,ν_B} is a flow and it holds

$$\mathcal{E}(\varphi_{\nu_A,\nu_B}) \leq \sum_{x\in A, y\in B} \frac{\nu_A(x)\nu_B(y)}{\operatorname{cap}(x,y)}.$$

Canonical flow

Harmonic flows between points $z, z' \in \mathcal{S}$:

$$\varphi_{z,z'}(x,y) := \frac{\mu(x)\, p(x,y)}{\operatorname{cap}(z,z')} \left(h_{z,z'}(x) - h_{z,z'}(y)\right) \qquad \begin{cases} Lh_{z,z'} = 0, & \text{on } \mathcal{S} \setminus \{z,z'\} \\ h_{z,z'}(z) = 1, \, h_{z,z'}(z') = 0. \end{cases}$$

Then

$$\mathcal{E}(\varphi_{z,z'}) = \frac{1}{2} \sum_{x,y} \frac{\mu(x)\, p(x,y)}{\operatorname{cap}(z,z')^2} \left(h_{z,z'}(x) - h_{z,z'}(y)\right)^2 = \frac{1}{\operatorname{cap}(z,z')}.$$

Lemma

For two probability measures ν_A, ν_B let the canonical flow be defined by

$$\varphi_{\nu_A,\nu_B}(x,y) := \sum_{z \in A, z' \in B} \nu_A(z) \nu_B(z') \frac{\mu(x) p(x,y)}{\operatorname{cap}(z,z')} \left(h_{z,z'}(x) - h_{z,z'}(y)\right).$$

Then φ_{ν_A,ν_B} is a flow and it holds

$$\mathcal{E}(\varphi_{\nu_A,\nu_B}) \leq \sum_{x \in A, y \in B} \frac{\nu_A(x) \nu_B(y)}{\operatorname{cap}(x,y)}.$$

Truncate by capacity

Tradeoff between approximation and comparison of capacities

Introduce coupling measures ν_a^δ and ν_b^δ between μ_a, μ_b:

$$
\begin{aligned}
&\Big(\mathbb{E}_{\mu_a}[f] - \mathbb{E}_{\mu_b}[f]\Big)^2 \\
&\quad \leq (1+\tau)\Big(\mathbb{E}_{\nu_a^\delta}[f] - \mathbb{E}_{\nu_b^\delta}[f]\Big)^2 + 2\Big(1+\frac{1}{\tau}\Big)\sum_{c\in\{a,b\}}\Big(\mathbb{E}_{\mu_c}[f] - \mathbb{E}_{\nu_c^\delta}[f]\Big)^2 \\
&\quad \leq (1+\tau)\left\|\frac{\nu_a^\delta}{\mu} - \frac{\nu_b^\delta}{\mu}\right\|_{H^{-1}(\mu)}^2 \mathcal{D}(f) + 2\Big(1+\frac{1}{\tau}\Big)\sum_{c\in\{a,b\}} \operatorname{cov}_{\mu_c}^2\left(f, \frac{\nu_c^\delta}{\mu_c}\right) \\
&\quad \leq (1+\tau)\,\mathcal{E}(\varphi_{\nu_a^\delta,\nu_b^\delta})\,\mathcal{D}(f) + 2\Big(1+\frac{1}{\tau}\Big)\sum_{c\in\{a,b\}} \operatorname{var}_{\mu_c}\left(\frac{\nu_c^\delta}{\mu_c}\right)\operatorname{var}_{\mu_c}(f) \\
&\quad \leq (1+\tau)\sum_{x,y}\frac{\nu_a^\delta(x)\nu_b^\delta(y)}{\operatorname{cap}(x,y)}\,\mathcal{D}(f) + 2\Big(1+\frac{1}{\tau}\Big)\sum_{c\in\{a,b\}}\frac{\operatorname{var}_{\mu_c}\left(\frac{\nu_c^\delta}{\mu_c}\right)}{\lambda_{\text{loc}}}\,\mathcal{D}(f)
\end{aligned}
$$

Truncate by capacity

Tradeoff between approximation and comparison of capacities

Interpolation between of H^{-1} and H^1norm + recognize covariance:

$$\begin{aligned}
&\Big(\mathbb{E}_{\mu_a}[f] - \mathbb{E}_{\mu_b}[f]\Big)^2 \\
&\quad\le (1+\tau)\Big(\mathbb{E}_{\nu_a^\delta}[f] - \mathbb{E}_{\nu_b^\delta}[f]\Big)^2 + 2\Big(1+\frac{1}{\tau}\Big)\sum_{c\in\{a,b\}}\Big(\mathbb{E}_{\mu_c}[f] - \mathbb{E}_{\nu_c^\delta}[f]\Big)^2 \\
&\quad\le (1+\tau)\left\|\frac{\nu_a^\delta}{\mu} - \frac{\nu_b^\delta}{\mu}\right\|_{H^{-1}(\mu)}^2 \mathcal{D}(f) + 2\Big(1+\frac{1}{\tau}\Big)\sum_{c\in\{a,b\}} \operatorname{cov}_{\mu_c}^2\left(f, \frac{\nu_c^\delta}{\mu_c}\right) \\
&\quad\le (1+\tau)\,\mathcal{E}(\varphi_{\nu_a^\delta,\nu_b^\delta})\,\mathcal{D}(f) + 2\Big(1+\frac{1}{\tau}\Big)\sum_{c\in\{a,b\}} \operatorname{var}_{\mu_c}\left(\frac{\nu_c^\delta}{\mu_c}\right)\operatorname{var}_{\mu_c}(f) \\
&\quad\le (1+\tau)\sum_{x,y}\frac{\nu_a^\delta(x)\nu_b^\delta(y)}{\operatorname{cap}(x,y)}\,\mathcal{D}(f) + 2\Big(1+\frac{1}{\tau}\Big)\sum_{c\in\{a,b\}} \frac{\operatorname{var}_{\mu_c}\left(\frac{\nu_c^\delta}{\mu_c}\right)}{\lambda_{\mathrm{loc}}}\,\mathcal{D}(f)
\end{aligned}$$

Truncate by capacity

Tradeoff between approximation and comparison of capacities

Use canonical flow for H^{-1}-norm + Cauchy-Schwarz on covariance:

$$
\begin{aligned}
&\Big(\mathbb{E}_{\mu_a}[f] - \mathbb{E}_{\mu_b}[f]\Big)^2 \\
&\quad\leq (1+\tau)\Big(\mathbb{E}_{\nu_a^\delta}[f] - \mathbb{E}_{\nu_b^\delta}[f]\Big)^2 + 2\Big(1+\frac{1}{\tau}\Big)\sum_{c\in\{a,b\}}\Big(\mathbb{E}_{\mu_c}[f] - \mathbb{E}_{\nu_c^\delta}[f]\Big)^2 \\
&\quad\leq (1+\tau)\left\|\frac{\nu_a^\delta}{\mu} - \frac{\nu_b^\delta}{\mu}\right\|^2_{H^{-1}(\mu)}\mathcal{D}(f) + 2\Big(1+\frac{1}{\tau}\Big)\sum_{c\in\{a,b\}}\operatorname{cov}^2_{\mu_c}\left(f, \frac{\nu_c^\delta}{\mu_c}\right) \\
&\quad\leq (1+\tau)\,\mathcal{E}(\varphi_{\nu_a^\delta,\nu_b^\delta})\,\mathcal{D}(f) + 2\Big(1+\frac{1}{\tau}\Big)\sum_{c\in\{a,b\}}\operatorname{var}_{\mu_c}\left(\frac{\nu_c^\delta}{\mu_c}\right)\operatorname{var}_{\mu_c}(f) \\
&\quad\leq (1+\tau)\sum_{x,y}\frac{\nu_a^\delta(x)\nu_b^\delta(y)}{\operatorname{cap}(x,y)}\,\mathcal{D}(f) + 2\Big(1+\frac{1}{\tau}\Big)\sum_{c\in\{a,b\}}\frac{\operatorname{var}_{\mu_c}\left(\frac{\nu_c^\delta}{\mu_c}\right)}{\lambda_{\mathrm{loc}}}\,\mathcal{D}(f)
\end{aligned}
$$

Tradeoff between approximation and comparison of capacities

Estimate canonical flow + Local Poincaré inequality:

$$
\begin{aligned}
&\Big(\mathbb{E}_{\mu_a}[f] - \mathbb{E}_{\mu_b}[f]\Big)^2 \\
&\quad \leq (1+\tau)\Big(\mathbb{E}_{\nu_a^\delta}[f] - \mathbb{E}_{\nu_b^\delta}[f]\Big)^2 + 2\Big(1+\frac{1}{\tau}\Big)\sum_{c\in\{a,b\}}\Big(\mathbb{E}_{\mu_c}[f] - \mathbb{E}_{\nu_c^\delta}[f]\Big)^2 \\
&\quad \leq (1+\tau)\left\|\frac{\nu_a^\delta}{\mu} - \frac{\nu_b^\delta}{\mu}\right\|_{H^{-1}(\mu)}^2 \mathcal{D}(f) + 2\Big(1+\frac{1}{\tau}\Big)\sum_{c\in\{a,b\}}\operatorname{cov}_{\mu_c}^2\left(f, \frac{\nu_c^\delta}{\mu_c}\right) \\
&\quad \leq (1+\tau)\,\mathcal{E}(\varphi_{\nu_a^\delta,\nu_b^\delta})\,\mathcal{D}(f) + 2\Big(1+\frac{1}{\tau}\Big)\sum_{c\in\{a,b\}}\operatorname{var}_{\mu_c}\left(\frac{\nu_c^\delta}{\mu_c}\right)\operatorname{var}_{\mu_c}(f) \\
&\quad \leq (1+\tau)\sum_{x,y}\frac{\nu_a^\delta(x)\nu_b^\delta(y)}{\operatorname{cap}(x,y)}\,\mathcal{D}(f) + 2\Big(1+\frac{1}{\tau}\Big)\sum_{c\in\{a,b\}}\frac{\operatorname{var}_{\mu_c}\left(\frac{\nu_c^\delta}{\mu_c}\right)}{\lambda_{\mathrm{loc}}}\,\mathcal{D}(f)
\end{aligned}
$$

Truncate by capacity

Tradeoff between approximation and comparison of capacities

$$\left(\mathbb{E}_{\mu_a}[f] - \mathbb{E}_{\mu_b}[f]\right)^2$$

$$\leq (1+\tau)\sum_{x,y} \frac{\nu_a^\delta(x)\nu_b^\delta(y)}{\operatorname{cap}(x,y)}\,\mathcal{D}(f) + 2\left(1+\frac{1}{\tau}\right)\sum_{c\in\{a,b\}} \frac{\operatorname{var}_{\mu_c}\left(\frac{\nu_c^\delta}{\mu_c}\right)}{\lambda_{\mathrm{loc}}}\,\mathcal{D}(f)$$

Choice of ν_c^δ, $c \in \{a,b\}$, has to provide:

a) Comparison of capacities: $\operatorname{cap}(x,y) \approx \operatorname{cap}(a,b)$ for $x \in \operatorname{supp}\nu_a^\delta$ and $y \in \operatorname{supp}\nu_b^\delta$.

b) Approximation in variance: $\operatorname{var}_{\mu_c}\left(\nu_c^\delta/\mu_c\right) \lesssim o(\varrho\mathcal{S})$

For $\delta > 0$ and $c \in \{a,b\}$ define:

$$\Omega_c^\delta := \{x \in \mathcal{S} : \operatorname{cap}(a,b) \leq \delta \operatorname{cap}(c,x)\} \quad \text{and} \quad \nu_c^\delta[x] := \frac{\mathbb{1}_{\Omega_c^\delta}(x)}{\mu\left(\Omega_c^\delta\right)}\,\mu[x].$$

- Provides a) and b) for appropriate choice of δ.
- Optimization of $\tau \Rightarrow \tau = o(\varrho|\mathcal{S}|)$ proves mean-difference Theorem.

Truncate by capacity

Tradeoff between approximation and comparison of capacities

$$\Big(\mathbb{E}_{\mu_a}[f] - \mathbb{E}_{\mu_b}[f]\Big)^2$$

$$\leq (1+\tau)\sum_{x,y} \frac{\nu_a^\delta(x)\nu_b^\delta(y)}{\operatorname{cap}(x,y)}\,\mathcal{D}(f) + 2\Big(1+\frac{1}{\tau}\Big)\sum_{c\in\{a,b\}} \frac{\operatorname{var}_{\mu_c}\Big(\frac{\nu_c^\delta}{\mu_c}\Big)}{\lambda_{\mathrm{loc}}}\,\mathcal{D}(f)$$

Choice of ν_c^δ, $c \in \{a,b\}$, has to provide:

a) Comparison of capacities: $\operatorname{cap}(x,y) \approx \operatorname{cap}(a,b)$
for $x \in \operatorname{supp}\nu_a^\delta$ and $y \in \operatorname{supp}\nu_b^\delta$.

b) Approximation in variance: $\operatorname{var}_{\mu_c}\left(\nu_c^\delta/\mu_c\right) \lesssim o(\varrho\mathcal{S})$

For $\delta > 0$ and $c \in \{a,b\}$ define:

$$\Omega_c^\delta := \{x \in \mathcal{S} : \operatorname{cap}(a,b) \leq \delta \operatorname{cap}(c,x)\} \quad \text{and} \quad \nu_c^\delta[x] := \frac{\mathbb{1}_{\Omega_c^\delta}(x)}{\mu\left(\Omega_c^\delta\right)}\,\mu[x].$$

- Provides a) and b) for appropriate choice of δ.
- Optimization of $\tau \Rightarrow \tau = o(\varrho|\mathcal{S}|)$ proves mean-difference Theorem.

Truncate by capacity

Tradeoff between approximation and comparison of capacities

$$\Big(\mathbb{E}_{\mu_a}[f] - \mathbb{E}_{\mu_b}[f]\Big)^2$$

$$\leq (1+\tau)\sum_{x,y} \frac{\nu_a^\delta(x)\nu_b^\delta(y)}{\operatorname{cap}(x,y)}\,\mathcal{D}(f) + 2\Big(1+\frac{1}{\tau}\Big)\sum_{c\in\{a,b\}} \frac{\operatorname{var}_{\mu_c}\Big(\frac{\nu_c^\delta}{\mu_c}\Big)}{\lambda_{\mathrm{loc}}}\,\mathcal{D}(f)$$

Choice of ν_c^δ, $c \in \{a,b\}$, has to provide:

a) Comparison of capacities: $\operatorname{cap}(x,y) \approx \operatorname{cap}(a,b)$

for $x \in \operatorname{supp} \nu_a^\delta$ and $y \in \operatorname{supp} \nu_b^\delta$.

b) Approximation in variance: $\operatorname{var}_{\mu_c}\left(\nu_c^\delta/\mu_c\right) \lesssim o(\varrho \mathcal{S})$

For $\delta > 0$ and $c \in \{a,b\}$ define:

$$\Omega_c^\delta := \{x \in \mathcal{S} : \operatorname{cap}(a,b) \leq \delta \operatorname{cap}(c,x)\} \quad \text{and} \quad \nu_c^\delta[x] := \frac{\mathbb{1}_{\Omega_c^\delta}(x)}{\mu\left(\Omega_c^\delta\right)}\,\mu[x].$$

- Provides a) and b) for appropriate choice of δ.
- Optimization of $\tau \Rightarrow \tau = o(\varrho|\mathcal{S}|)$ proves mean-difference Theorem.

Truncate by capacity

Tradeoff between approximation and comparison of capacities

$$\Big(\mathbb{E}_{\mu_a}[f] - \mathbb{E}_{\mu_b}[f]\Big)^2$$
$$\leq (1+\tau) \sum_{x,y} \frac{\nu_a^\delta(x)\nu_b^\delta(y)}{\operatorname{cap}(x,y)}\, \mathcal{D}(f) + 2\Big(1+\frac{1}{\tau}\Big) \sum_{c\in\{a,b\}} \frac{\operatorname{var}_{\mu_c}\Big(\frac{\nu_c^\delta}{\mu_c}\Big)}{\lambda_{\mathrm{loc}}}\, \mathcal{D}(f)$$

Choice of ν_c^δ, $c \in \{a,b\}$, has to provide:

a) Comparison of capacities: $\operatorname{cap}(x,y) \approx \operatorname{cap}(a,b)$

for $x \in \operatorname{supp} \nu_a^\delta$ and $y \in \operatorname{supp} \nu_b^\delta$.

b) Approximation in variance: $\operatorname{var}_{\mu_c}\left(\nu_c^\delta/\mu_c\right) \lesssim o(\varrho\mathcal{S})$

For $\delta > 0$ and $c \in \{a,b\}$ define:

$$\Omega_c^\delta := \{x \in \mathcal{S} : \operatorname{cap}(a,b) \leq \delta \operatorname{cap}(c,x)\} \quad \text{and} \quad \nu_c^\delta[x] := \frac{\mathbb{1}_{\Omega_c^\delta}(x)}{\mu\left(\Omega_c^\delta\right)}\,\mu[x].$$

- Provides $a)$ and $b)$ for appropriate choice of δ.
- Optimization of $\tau \Rightarrow \tau = o(\varrho|\mathcal{S}|)$ proves mean-difference Theorem.

Truncate by capacity

Tradeoff between approximation and comparison of capacities

$$\Big(\mathbb{E}_{\mu_a}[f] - \mathbb{E}_{\mu_b}[f]\Big)^2$$
$$\leq (1+\tau)\sum_{x,y} \frac{\nu_a^\delta(x)\nu_b^\delta(y)}{\mathrm{cap}(x,y)}\,\mathcal{D}(f) + 2\Big(1+\frac{1}{\tau}\Big)\sum_{c\in\{a,b\}} \frac{\mathrm{var}_{\mu_c}\left(\frac{\nu_c^\delta}{\mu_c}\right)}{\lambda_{\mathrm{loc}}}\,\mathcal{D}(f)$$

Choice of ν_c^δ, $c \in \{a,b\}$, has to provide:

a) Comparison of capacities: $\mathrm{cap}(x,y) \approx \mathrm{cap}(a,b)$
for $x \in \operatorname{supp}\nu_a^\delta$ and $y \in \operatorname{supp}\nu_b^\delta$.

b) Approximation in variance: $\mathrm{var}_{\mu_c}\left(\nu_c^\delta/\mu_c\right) \lesssim o(\varrho\mathcal{S})$

For $\delta > 0$ and $c \in \{a,b\}$ define:

$$\Omega_c^\delta := \{x \in \mathcal{S} : \mathrm{cap}(a,b) \leq \delta\,\mathrm{cap}(c,x)\} \quad \text{and} \quad \nu_c^\delta[x] := \frac{\mathbb{1}_{\Omega_c^\delta}(x)}{\mu\left(\Omega_c^\delta\right)}\,\mu[x].$$

- Provides a) and b) for appropriate choice of δ.
- Optimization of $\tau \Rightarrow \tau = o(\varrho|\mathcal{S}|)$ proves mean-difference Theorem.

Summary

- Examples of entropic gradient flows showing metastability:
  - Fokker-Planck equation at low temperature
  - Metastable reversible Markov chains for growing system
- Partitions and splitting induced from dynamic (two scales)
- Optimal constants in PI and LSI follow from two ingredients:
  - *good* local mixing
    $\Rightarrow$ Donsker-Varadhan/Lyapunov technique
  - *sharp* estimates of mean-difference
    $\Rightarrow$ transport/flow representation of H^{-1}-norm and optimization

Summary

- Examples of entropic gradient flows showing metastability:
  - Fokker-Planck equation at low temperature
  - Metastable reversible Markov chains for growing system
- Partitions and splitting induced from dynamic (two scales)
- Optimal constants in PI and LSI follow from two ingredients:
  - *good* local mixing
    $\Rightarrow$ Donsker-Varadhan/Lyapunov technique
  - *sharp* estimates of mean-difference
    $\Rightarrow$ transport/flow representation of H^{-1}-norm

Summary

- Examples of entropic gradient flows showing metastability:
  - Fokker-Planck equation at low temperature
  - Metastable reversible Markov chains for growing system
- Partitions and splitting induced from dynamic (two scales)
- Optimal constants in PI and LSI follow from two ingredients:
  - *good* local mixing
    $\Rightarrow$ Donsker-Varadhan/Lyapunov technique
  - *sharp* estimates of mean-difference
    $\Rightarrow$ transport/flow representation of H^{-1}-norm and optimization

Summary

- Examples of entropic gradient flows showing metastability:
  - Fokker-Planck equation at low temperature
  - Metastable reversible Markov chains for growing system
- Partitions and splitting induced from dynamic (two scales)
- Optimal constants in PI and LSI follow from two ingredients:
  - *good* local mixing
    $\Rightarrow$ Donsker-Varadhan/Lyapunov technique
  - *sharp* estimates of mean-difference
    $\Rightarrow$ transport/flow representation of H^{-1}-norm and optimization

Summary

- Examples of entropic gradient flows showing metastability:
  - Fokker-Planck equation at low temperature
  - Metastable reversible Markov chains for growing system
- Partitions and splitting induced from dynamic (two scales)
- Optimal constants in PI and LSI follow from two ingredients:
  - *good* local mixing
    $\Rightarrow$ Donsker-Varadhan/Lyapunov technique
  - *sharp* estimates of mean-difference
    $\Rightarrow$ transport/flow representation of H^{-1}-norm and optimization

Summary

- Examples of entropic gradient flows showing metastability:
  - Fokker-Planck equation at low temperature
  - Metastable reversible Markov chains for growing system
- Partitions and splitting induced from dynamic (two scales)
- Optimal constants in PI and LSI follow from two ingredients:
  - *good* local mixing
    $\Rightarrow$ Donsker-Varadhan/Lyapunov technique
  - *sharp* estimates of mean-difference
    $\Rightarrow$ transport/flow representation of H^{-1}-norm and optimization

Summary

- Examples of entropic gradient flows showing metastability:
  - Fokker-Planck equation at low temperature
  - Metastable reversible Markov chains for growing system
- Partitions and splitting induced from dynamic (two scales)
- Optimal constants in PI and LSI follow from two ingredients:
  - *good* local mixing
    $\Rightarrow$ Donsker-Varadhan/Lyapunov technique
  - *sharp* estimates of mean-difference
    $\Rightarrow$ transport/flow representation of H^{-1}-norm and optimization