Eyring-Kramers formula
for
Poincaré
and
logarithmic Sobolev inequalities

André Schlichting

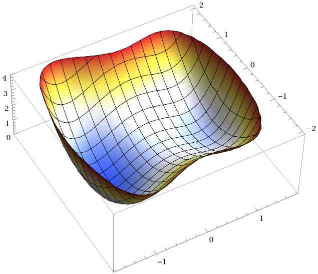
Institute for Applied Mathematics, University of Bonn

5th Workshop on Random Dynamical Systems, Bielefeld.

October 5, 2012

Introduction
Overdamped Langevin dynamics

Hamiltonian $H : \mathbb{R}^n \to \mathbb{R}$ *energy landscape*



Dynamic at temperature $\varepsilon \ll 1$

$$dX_t = -\nabla H(X_t)dt + \sqrt{2\varepsilon}\, dW_t$$

Gibbs measure $\mu(dx) = \frac{1}{Z_\mu} \exp\left(-\frac{H}{\varepsilon}\right) dx,$

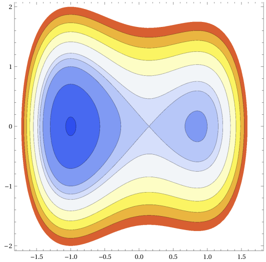
where $Z_\mu = \int e^{-\frac{H}{\varepsilon}} dx$

Generator law $X_t = f_t \mu$ evolves

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

Dirichlet form $\mathcal{E}(f) := \int (-Lf) f \, d\mu$

$$= \varepsilon \int |\nabla f|^2 \, d\mu.$$



Introduction

Overdamped Langevin dynamics

Hamiltonian $H : \mathbb{R}^n \to \mathbb{R}$ *energy landscape*

Dynamic at temperature $\varepsilon \ll 1$

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

Gibbs measure $\mu(\mathrm{d}x) = \frac{1}{Z_\mu} \exp\left(-\frac{H}{\varepsilon}\right) \mathrm{d}x,$

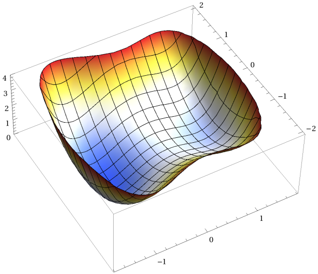
where $Z_\mu = \int e^{-\frac{H}{\varepsilon}} \mathrm{d}x$

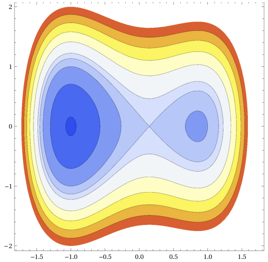
Generator law $X_t = f_t \mu$ evolves

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

Dirichlet form $\mathcal{E}(f) := \int (-Lf) f \,\mathrm{d}\mu$

$$= \varepsilon \int |\nabla f|^2 \,\mathrm{d}\mu.$$





Introduction

Overdamped Langevin dynamics

Hamiltonian $H : \mathbb{R}^n \to \mathbb{R}$ *energy landscape*

Dynamic at temperature $\varepsilon \ll 1$

$$dX_t = -\nabla H(X_t)dt + \sqrt{2\varepsilon}\, dW_t$$

Gibbs measure $\mu(dx) = \frac{1}{Z_\mu} \exp\left(-\frac{H}{\varepsilon}\right) dx,$

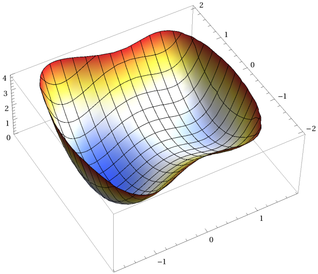
where $Z_\mu = \int e^{-\frac{H}{\varepsilon}} dx$

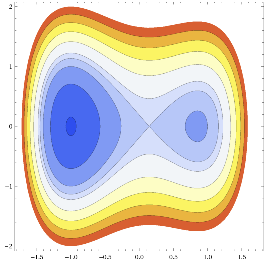
Generator law $X_t = f_t \mu$ evolves

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

Dirichlet form $\mathcal{E}(f) := \int (-Lf) f \, d\mu$

$$= \varepsilon \int |\nabla f|^2 \, d\mu.$$





Introduction

Overdamped Langevin dynamics

Hamiltonian $H : \mathbb{R}^n \to \mathbb{R}$ *energy landscape*

Dynamic at temperature $\varepsilon \ll 1$

$$dX_t = -\nabla H(X_t)dt + \sqrt{2\varepsilon}\, dW_t$$

Gibbs measure $\mu(dx) = \frac{1}{Z_\mu} \exp\left(-\frac{H}{\varepsilon}\right) dx,$

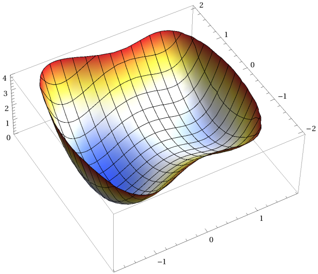
where $Z_\mu = \int e^{-\frac{H}{\varepsilon}} dx$

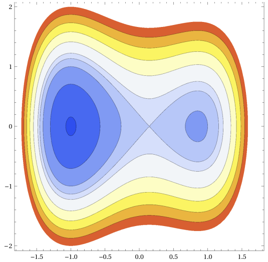
Generator law $X_t = f_t \mu$ evolves

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

Dirichlet form $\mathcal{E}(f) := \int (-Lf) f \, d\mu$

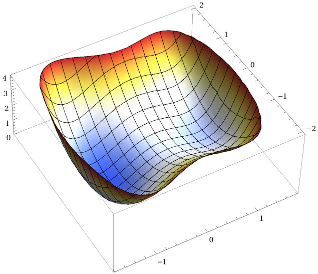
$$= \varepsilon \int |\nabla f|^2 \, d\mu .$$





Introduction

Overdamped Langevin dynamics



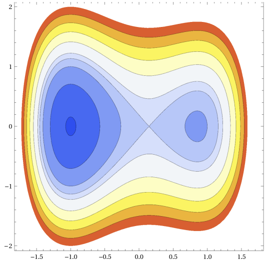
Hamiltonian $H : \mathbb{R}^n \to \mathbb{R}$ *energy landscape*

Dynamic at temperature $\varepsilon \ll 1$

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

Gibbs measure $\mu(\mathrm{d}x) = \frac{1}{Z_\mu} \exp\left(-\frac{H}{\varepsilon}\right) \mathrm{d}x,$

where $Z_\mu = \int e^{-\frac{H}{\varepsilon}} \mathrm{d}x$



Generator law $X_t = f_t \mu$ evolves

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

Dirichlet form $\mathcal{E}(f) := \int (-Lf) f \,\mathrm{d}\mu$

$$= \varepsilon \int |\nabla f|^2 \,\mathrm{d}\mu.$$

Poincaré and logarithmic Sobolev inequality

Definition

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

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

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

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

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

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

Poincaré and logarithmic Sobolev inequality

Definition

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

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

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

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

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

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

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

Poincaré and logarithmic Sobolev inequality

Definition

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

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

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

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

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

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

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

Poincaré and logarithmic Sobolev inequality

Definition

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

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

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

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

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

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

Partitions

Basins of attraction $\Omega_0 \uplus \Omega_1 = \mathbb{R}^n$ of local minima m_0, m_1:

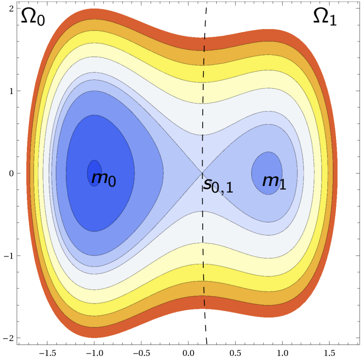
$$\Omega_i := \{y_0 \in \mathbb{R}^n : \dot{y}_t = -\nabla H(y_t), y_t \to m_i\}.$$

Restricted measures μ_0, μ_1:

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

Mixture representation

$$\mu = Z_0 \mu_0 + Z_1 \mu_1, \quad Z_i := \mu(\Omega_i).$$



Splitting

Lemma

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

$$\begin{aligned}\mathrm{Ent}_\mu(f^2) \le\ & \overbrace{Z_0 \,\mathrm{Ent}_{\mu_0}(f^2) + Z_1 \,\mathrm{Ent}_{\mu_1}(f^2)}^{\textit{local entropies}} \\ & + \frac{Z_0 Z_1}{\Lambda(Z_0, Z_1)} \left(\mathrm{var}_{\mu_0}(f) + \mathrm{var}_{\mu_1}(f) + \left(\mathbb{E}_{\mu_0}(f) - \mathbb{E}_{\mu_1}(f)\right)^2\right),\end{aligned}$$

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

Splitting

Lemma

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

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

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

Splitting

Lemma

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

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

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

Main results

Theorem (Local PI and LSI)

The measures μ_0 *and* μ_1 *satisfy* $\mathrm{PI}(\varrho_{loc})$ *and* $\mathrm{LSI}(\alpha_{loc})$ *with*

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

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

Theorem (Mean-difference estimate)

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

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

Main results

Theorem (Local PI and LSI)

The measures μ_0 and μ_1 satisfy $\mathrm{PI}(\varrho_{loc})$ *and* $\mathrm{LSI}(\alpha_{loc})$ *with*

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

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

Theorem (Mean-difference estimate)

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

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

Main results

Theorem (Local PI and LSI)

The measures μ_0 *and* μ_1 *satisfy* $\mathrm{PI}(\varrho_{loc})$ *and* $\mathrm{LSI}(\alpha_{loc})$ *with*

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

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

Theorem (Mean-difference estimate)

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

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

Main results

Theorem (Local PI and LSI)

The measures μ_0 *and* μ_1 *satisfy* $\mathrm{PI}(\varrho_{loc})$ *and* $\mathrm{LSI}(\alpha_{loc})$ *with*

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

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

Theorem (Mean-difference estimate)

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

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

Main results

Theorem (Local PI and LSI)

The measures μ_0 and μ_1 satisfy $\mathrm{PI}(\varrho_{loc})$ *and* $\mathrm{LSI}(\alpha_{loc})$ *with*

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

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

Theorem (Mean-difference estimate)

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

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

Eyring-Kramers formula

Corollary

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

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

Asymptotic evaluation of the factor $\frac{Z_0 Z_1 Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}}$ for two special cases:

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

$$H(m_0) = H(m_1): \quad 1 \leq \frac{\varrho}{\alpha} \lesssim \frac{\frac{\kappa_0+\kappa_1}{2}}{\Lambda(\kappa_0, \kappa_1)}, \quad \text{where } \kappa_i := \sqrt{\det \nabla^2 H(m_i)}.$$

Eyring-Kramers formula

Corollary

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

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

Asymptotic evaluation of the factor $\frac{Z_0 Z_1 Z_\mu}{(2\pi\varepsilon)^{\frac{n}{2}}}$ for two special cases:

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

$$H(m_0) = H(m_1): \quad 1 \le \frac{\varrho}{\alpha} \lesssim \frac{\frac{\kappa_0 + \kappa_1}{2}}{\Lambda(\kappa_0, \kappa_1)}, \quad \text{where } \kappa_i := \sqrt{\det \nabla^2 H(m_i)}.$$

Proof: Mean-difference estimate

Goal: Find a good estimate for C in

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

Proof: Mean-difference estimate

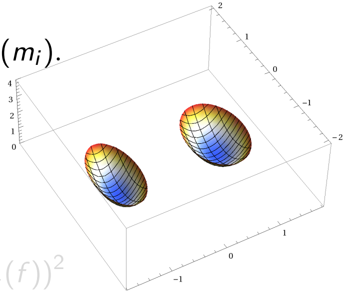
Approximation step

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:

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



Introduce ν_0 and ν_1 as coupling:

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

Approximation bound follows from local PI and local LSI.

Proof: Mean-difference estimate

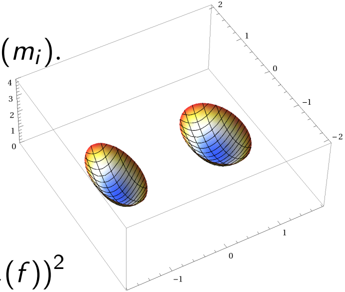
Approximation step

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:

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



Introduce ν_0 and ν_1 as coupling:

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

Approximation bound follows from local PI and local LSI.

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

Proof: Mean-difference estimate

Transport interpolation

Goal: Find a good estimate for C in

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

Step 2: Transport $(\Phi_s : \mathbb{R}^n \to \mathbb{R}^n)_{s \in [0,1]}$ 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 $(\Phi_s : \mathbb{R}^n \to \mathbb{R}^n)_{s\in[0,1]}$ 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 $(\Phi_s : \mathbb{R}^n \to \mathbb{R}^n)_{s\in[0,1]}$ 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

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

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

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

Proof: Mean-difference estimate

Transport interpolation

Goal: Find a good estimate for C in

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

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

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

Proof: Mean-difference estimate

Transport interpolation

Goal: Find a good estimate for C in

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

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

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

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

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

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 $(\Phi_s : \mathbb{R}^n \to \mathbb{R}^n)_{s\in[0,1]}$ 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

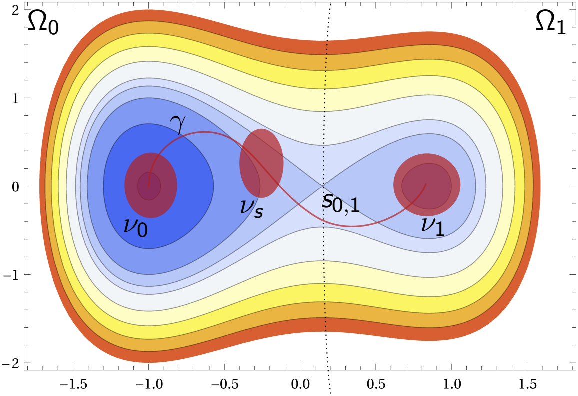
Construction of transport interpolation

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 $s_{0,1}$



Proof: Mean-difference estimate

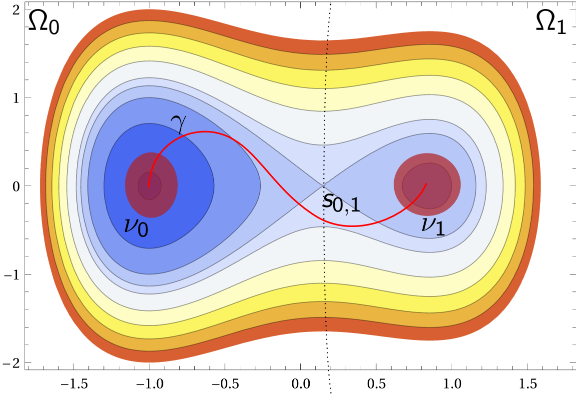
Construction of transport interpolation

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 $s_{0,1}$



Proof: Mean-difference estimate

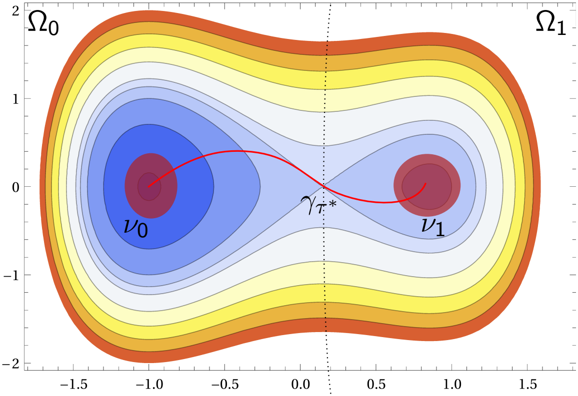
Construction of transport interpolation

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



Proof: Mean-difference estimate

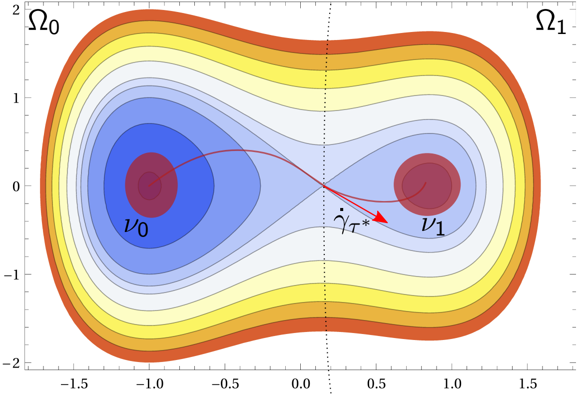
Construction of transport interpolation

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 $s_{0,1}$



Proof: Mean-difference estimate

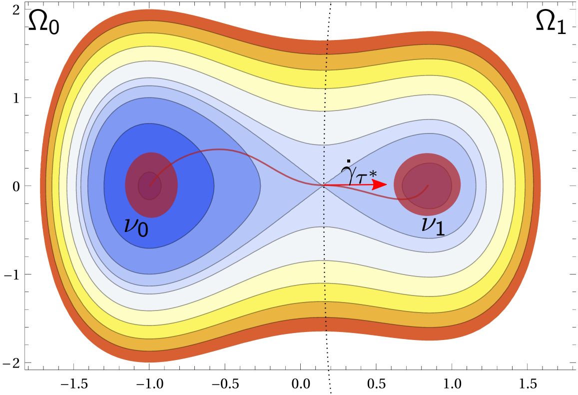
Construction of transport interpolation

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 $s_{0,1}$



Proof: Mean-difference estimate

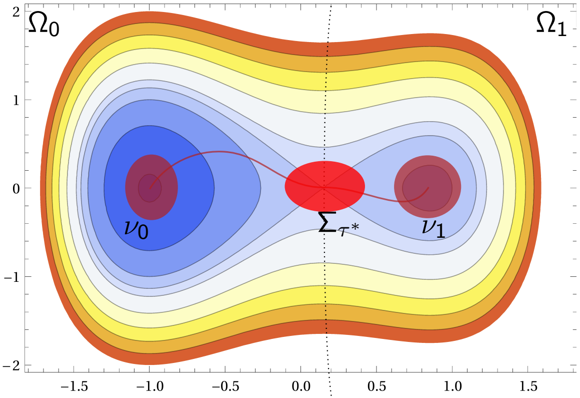
Construction of transport interpolation

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



Proof: Mean-difference estimate

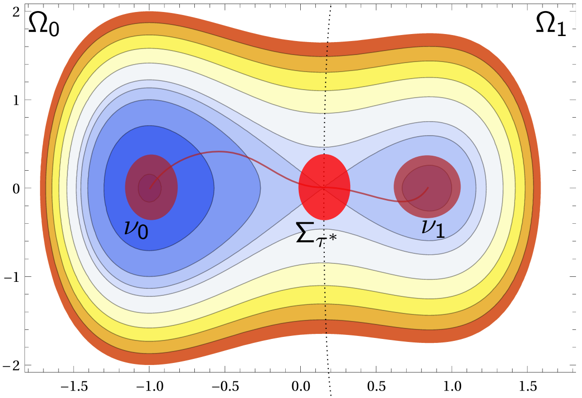
Construction of transport interpolation

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 $s_{0,1}$



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 $s_{0,1}$