Transport Inequalities For Markov Processes

C. Léonard

A joint work with Arnaud Guillin, Liming Wu and Nian Yao
Université Paris 10 & École Polytechnique

Beijing, July 14-18, 2007

Outline

Outline

Outline

1. Setting And Definitions
   - The Markov Process
   - Optimal Transport
   - Inequalities
2. A Connection With Large deviations
   - Large Deviations Of The Occupation Measure
   - The Basic Result
   - Tensorization
3. Criteria
   - Poincaré and log-Sobolev
   - Poincaré implies Pinsker
   - Spectral gap in C^0_{Lip}
   - Lyapunov function condition

Outline

The Markov process

Consider

- State space: $\mathcal{X}$ is Polish
- Invariant probability measure: $\mu \in \mathcal{P}_\mathcal{X}$
- Markov generator: $\mathcal{L}$
  - $\mathcal{L}$ self-adjoint in $L^2(\mu)$
  - the semigroup $(P_t = e^{t\mathcal{L}})_{t \geq 0}$ is μ-ergodic :

$$P_t f = f, \mu\text{-a.e.}, \forall t \Rightarrow f = c, \mu\text{-a.e.}$$

- Dirichlet form:

$$\mathcal{E}(g) := \langle -\mathcal{L}g, g \rangle_{L^2(\mu)}, \quad g \in \mathbb{D}_2(\mathcal{L}) \subset L^2(\mu).$$

  $\mathcal{E}$ is closable in $L^2(\mu)$ and $\mathbb{D}(\mathcal{E}) = \mathbb{D}_2(\sqrt{-\mathcal{L}}) \subset L^2(\mu)$.

- Fisher-Donsker-Varadhan information

$$I(\nu|\mu) = \begin{cases} \mathcal{E}\left(\sqrt{\frac{d\nu}{d\mu}}\right), & \text{if } \sqrt{\frac{d\nu}{d\mu}} \in \mathbb{D}(\mathcal{E}); \\ +\infty, & \text{otherwise.} \end{cases}$$

The Markov process

Consider

- State space: $\mathcal{X}$ is Polish
- Invariant probability measure: $\mu \in \mathcal{P}_{\mathcal{X}}$
- Markov generator: $\mathcal{L}$
  - $\mathcal{L}$ self-adjoint in $L^2(\mu)$
  - the semigroup $(P_t = e^{t\mathcal{L}})_{t\geq 0}$ is μ-ergodic :

$$P_t f = f, \mu\text{-a.e.}, \forall t \Rightarrow f = c, \mu\text{-a.e.}$$

- Dirichlet form:

$$\mathcal{E}(g) := \langle -\mathcal{L}g, g\rangle_{L^2(\mu)}, \quad g \in \mathbb{D}_2(\mathcal{L}) \subset L^2(\mu).$$

  $\mathcal{E}$ is closable in $L^2(\mu)$ and $\mathbb{D}(\mathcal{E}) = \mathbb{D}_2(\sqrt{-\mathcal{L}}) \subset L^2(\mu)$.

- Fisher-Donsker-Varadhan information

$$I(\nu|\mu) = \begin{cases} \mathcal{E}\left(\sqrt{\frac{d\nu}{d\mu}}\right), & \text{if } \sqrt{\frac{d\nu}{d\mu}} \in \mathbb{D}(\mathcal{E}); \\ +\infty, & \text{otherwise.} \end{cases}$$

The Markov process

Consider

- State space: $\mathcal{X}$ is Polish
- Invariant probability measure: $\mu \in \mathcal{P}_{\mathcal{X}}$
- Markov generator: $\mathcal{L}$
  - $\mathcal{L}$ self-adjoint in $L^2(\mu)$
  - the semigroup $(P_t = e^{t\mathcal{L}})_{t \geq 0}$ is μ-ergodic :

$$P_t f = f, \mu\text{-a.e.}, \forall t \Rightarrow f = c, \mu\text{-a.e.}$$

- Dirichlet form:

$$\mathcal{E}(g) := \langle -\mathcal{L}g, g \rangle_{L^2(\mu)}, \quad g \in \mathbb{D}_2(\mathcal{L}) \subset L^2(\mu).$$

  $\mathcal{E}$ is closable in $L^2(\mu)$ and $\mathbb{D}(\mathcal{E}) = \mathbb{D}_2(\sqrt{-\mathcal{L}}) \subset L^2(\mu)$.
- Fisher-Donsker-Varadhan information

$$I(\nu|\mu) = \begin{cases} \mathcal{E}\left(\sqrt{\frac{d\nu}{d\mu}}\right), & \text{if } \sqrt{\frac{d\nu}{d\mu}} \in \mathbb{D}(\mathcal{E}); \\ +\infty, & \text{otherwise.} \end{cases}$$

The Markov process

Consider

- State space: $\mathcal{X}$ is Polish
- Invariant probability measure: $\mu \in \mathcal{P}_{\mathcal{X}}$
- Markov generator: $\mathcal{L}$
  - $\mathcal{L}$ self-adjoint in $L^2(\mu)$
  - the semigroup $(P_t = e^{t\mathcal{L}})_{t\geq 0}$ is μ-ergodic :

$$P_t f = f, \mu\text{-a.e.}, \forall t \Rightarrow f = c, \mu\text{-a.e.}$$

- Dirichlet form:

$$\mathcal{E}(g) := \langle -\mathcal{L}g, g\rangle_{L^2(\mu)}, \quad g \in \mathbb{D}_2(\mathcal{L}) \subset L^2(\mu).$$

  $\mathcal{E}$ is closable in $L^2(\mu)$ and $\mathbb{D}(\mathcal{E}) = \mathbb{D}_2(\sqrt{-\mathcal{L}}) \subset L^2(\mu)$.

- Fisher-Donsker-Varadhan information

$$I(\nu|\mu) = \begin{cases} \mathcal{E}\left(\sqrt{\frac{d\nu}{d\mu}}\right), & \text{if } \sqrt{\frac{d\nu}{d\mu}} \in \mathbb{D}(\mathcal{E}); \\ +\infty, & \text{otherwise.} \end{cases}$$

The Markov process

Consider

- State space: $\mathcal{X}$ is Polish
- Invariant probability measure: $\mu \in \mathcal{P}_\mathcal{X}$
- Markov generator: $\mathcal{L}$
  - $\mathcal{L}$ self-adjoint in $L^2(\mu)$
  - the semigroup $(P_t = e^{t\mathcal{L}})_{t\geq 0}$ is μ-ergodic :

$$P_t f = f, \mu\text{-a.e.}, \forall t \Rightarrow f = c, \mu\text{-a.e.}$$

- Dirichlet form:

$$\mathcal{E}(g) := \langle -\mathcal{L}g, g\rangle_{L^2(\mu)}, \quad g \in \mathbb{D}_2(\mathcal{L}) \subset L^2(\mu).$$

  $\mathcal{E}$ is closable in $L^2(\mu)$ and $\mathbb{D}(\mathcal{E}) = \mathbb{D}_2(\sqrt{-\mathcal{L}}) \subset L^2(\mu)$.
- Fisher-Donsker-Varadhan information

$$I(\nu|\mu) = \begin{cases} \mathcal{E}\left(\sqrt{\frac{d\nu}{d\mu}}\right), & \text{if } \sqrt{\frac{d\nu}{d\mu}} \in \mathbb{D}(\mathcal{E}); \\ +\infty, & \text{otherwise.} \end{cases}$$

Outline

Optimal transport cost

Let $\nu, \mu \in \mathcal{P}_\mathcal{X}$ be probability measures on $\mathcal{X}$.

Definition (Optimal transport cost)

$$\mathcal{T}_c(\nu, \mu) = \inf_{P(\nu,\mu)} \iint_{\mathcal{X}^2} c(x, y)\, \pi(dxdy)$$

- $P(\nu, \mu) = \{\pi \in \mathcal{P}_{\mathcal{X}^2}; \pi_1 = \nu, \pi_2 = \mu\}$
- $c(x, y)$ is a $[0, \infty)$-valued lower semicontinuous function

Theorem (Kantorovich duality)

$$\mathcal{T}_c(\nu, \mu) = \sup_{(u,v)\in\Phi_c} \{\nu(u) - \mu(v)\}$$

- $\Phi_c := \{(u, v) \text{ functions}; u(x) - v(y) \leq c(x, y), \forall x, y \in \mathcal{X}\}$.

Optimal transport cost

Let $\nu, \mu \in \mathcal{P}_{\mathcal{X}}$ be probability measures on $\mathcal{X}$.

Definition (Optimal transport cost)

$$\mathcal{T}_c(\nu, \mu) = \inf_{P(\nu,\mu)} \iint_{\mathcal{X}^2} c(x, y)\, \pi(dxdy)$$

- $P(\nu, \mu) = \{\pi \in \mathcal{P}_{\mathcal{X}^2}; \pi_1 = \nu, \pi_2 = \mu\}$
- $c(x, y)$ is a $[0, \infty)$-valued lower semicontinuous function

Theorem (Kantorovich duality)

$$\mathcal{T}_c(\nu, \mu) = \sup_{(u,v)\in\Phi_c} \{\nu(u) - \mu(v)\}$$

- $\Phi_c := \{(u, v) \text{ functions}; u(x) - v(y) \le c(x, y), \forall x, y \in \mathcal{X}\}$.

Optimal transport cost

Let $\nu, \mu \in \mathcal{P}_\mathcal{X}$ be probability measures on $\mathcal{X}$.

Definition (Optimal transport cost)

$$\mathcal{T}_c(\nu,\mu) = \inf_{P(\nu,\mu)} \iint_{\mathcal{X}^2} c(x,y)\,\pi(dxdy)$$

- $P(\nu,\mu) = \{\pi \in \mathcal{P}_{\mathcal{X}^2}; \pi_1 = \nu, \pi_2 = \mu\}$
- $c(x,y)$ is a $[0,\infty)$-valued lower semicontinuous function

Theorem (Kantorovich duality)

$$\mathcal{T}_c(\nu,\mu) = \sup_{(u,v)\in\Phi_c} \{\nu(u) - \mu(v)\}$$

- $\Phi_c := \{(u,v) \text{ functions}; u(x) - v(y) \le c(x,y), \forall x,y \in \mathcal{X}\}$.

Optimal transport cost

Let $\nu, \mu \in \mathcal{P}_{\mathcal{X}}$ be probability measures on $\mathcal{X}$.

Definition (Optimal transport cost)

$$\mathcal{T}_c(\nu, \mu) = \inf_{P(\nu,\mu)} \iint_{\mathcal{X}^2} c(x, y)\, \pi(dxdy)$$

- $P(\nu, \mu) = \{\pi \in \mathcal{P}_{\mathcal{X}^2}; \pi_1 = \nu, \pi_2 = \mu\}$
- $c(x, y)$ is a $[0, \infty)$-valued lower semicontinuous function

Theorem (Kantorovich duality)

$$\mathcal{T}_c(\nu, \mu) = \sup_{(u,v)\in\Phi_c} \{\nu(u) - \mu(v)\}$$

- $\Phi_c := \{(u, v) \text{ functions}; u(x) - v(y) \le c(x, y), \forall x, y \in \mathcal{X}\}.$

Optimal transport cost

Let $\nu, \mu \in \mathcal{P}_{\mathcal{X}}$ be probability measures on $\mathcal{X}$.

Definition (Optimal transport cost)

$$\mathcal{T}_c(\nu, \mu) = \inf_{P(\nu,\mu)} \iint_{\mathcal{X}^2} c(x, y)\, \pi(dxdy)$$

- $P(\nu, \mu) = \{\pi \in \mathcal{P}_{\mathcal{X}^2}; \pi_1 = \nu, \pi_2 = \mu\}$
- $c(x, y)$ is a $[0, \infty)$-valued lower semicontinuous function

Theorem (Kantorovich duality)

$$\mathcal{T}_c(\nu, \mu) = \sup_{(u,v)\in\Phi_c} \{\nu(u) - \mu(v)\}$$

- $\Phi_c := \{(u, v) \text{ functions}; u(x) - v(y) \le c(x, y), \forall x, y \in \mathcal{X}\}.$

Optimal transport cost

Definition (Wasserstein's metric)

Let d be a lower semicontinuous metric on $\mathcal{X}$, the cost $c(x, y) = d^p(x, y)$ with $1 \leq p < \infty$, gives

$$W_p(\mu, \nu) := \mathcal{T}_{d^p}(\mu, \nu)^{1/p}.$$

Result: It is a metric on (a subset of) $\mathcal{P}_{\mathcal{X}}$.

With $c = d$ we have $\mathcal{T}_c = W_1$.

Theorem (Kantorovich-Rubinstein)

$$W_1(\nu, \mu) = \sup\{\nu(u) - \mu(u); \|u\|_{\mathrm{Lip}} \leq 1\} := \|\nu - \mu\|^*_{\mathrm{Lip}}.$$

where $\|u\|_{\mathrm{Lip}} := \sup_{x \neq y} \frac{|u(x) - u(y)|}{d(x,y)}$.

Optimal transport cost

Definition (Wasserstein's metric)

Let d be a lower semicontinuous metric on $\mathcal{X}$, the cost $c(x,y) = d^p(x,y)$ with $1 \le p < \infty$, gives

$$W_p(\mu,\nu) := \mathcal{T}_{d^p}(\mu,\nu)^{1/p}.$$

Result: It is a metric on (a subset of) $\mathcal{P}_\mathcal{X}$.

With $c = d$ we have $\mathcal{T}_c = W_1$.

Theorem (Kantorovich-Rubinstein)

$$W_1(\nu,\mu) = \sup\{\nu(u) - \mu(u); \|u\|_{\mathrm{Lip}} \le 1\} := \|\nu - \mu\|^*_{\mathrm{Lip}}.$$

where $\|u\|_{\mathrm{Lip}} := \sup_{x \neq y} \frac{|u(x)-u(y)|}{d(x,y)}$.

Optimal transport cost

Definition (Wasserstein's metric)

Let d be a lower semicontinuous metric on $\mathcal{X}$, the cost $c(x,y) = d^p(x,y)$ with $1 \le p < \infty$, gives

$$W_p(\mu,\nu) := \mathcal{T}_{d^p}(\mu,\nu)^{1/p}.$$

Result: It is a metric on (a subset of) $\mathcal{P}_{\mathcal{X}}$.

With $c = d$ we have $\mathcal{T}_c = W_1$.

Theorem (Kantorovich-Rubinstein)

$$W_1(\nu,\mu) = \sup\{\nu(u) - \mu(u); \|u\|_{\mathrm{Lip}} \le 1\} := \|\nu - \mu\|^*_{\mathrm{Lip}}.$$

where $\|u\|_{\mathrm{Lip}} := \sup_{x \ne y} \frac{|u(x)-u(y)|}{d(x,y)}$.

Optimal transport cost

Definition (Wasserstein's metric)

Let d be a lower semicontinuous metric on $\mathcal{X}$, the cost $c(x, y) = d^p(x, y)$ with $1 \leq p < \infty$, gives

$$W_p(\mu, \nu) := \mathcal{T}_{d^p}(\mu, \nu)^{1/p}.$$

Result: It is a metric on (a subset of) $\mathcal{P}_{\mathcal{X}}$.

With $c = d$ we have $\mathcal{T}_c = W_1$.

Theorem (Kantorovich-Rubinstein)

$$W_1(\nu, \mu) = \sup \{\nu(u) - \mu(u); \|u\|_{\mathrm{Lip}} \leq 1\} := \|\nu - \mu\|^*_{\mathrm{Lip}}.$$

where $\|u\|_{\mathrm{Lip}} := \sup_{x \neq y} \frac{|u(x)-u(y)|}{d(x,y)}$.

Outline

The generalized transport cost

Definition (The generalized transport cost)

$$\mathcal{T}_\Phi(\nu,\mu) := \sup_{(u,v)\in\Phi} \{\nu(u) - \mu(v)\}, \quad \nu,\mu \in \mathcal{P}_\mathcal{X}.$$

- $\Phi = \Phi_c = \{(u,v) \text{ functions}; u(x) - v(y) \le c(x,y), \forall x,y \in \mathcal{X}\}$, gives $\mathcal{T}_\Phi = \mathcal{T}_c$.
- $\Phi = \Phi_{\|\cdot\|} := \{(u,u); u : \mathcal{X} \to \mathbb{R}, \|u\| \le 1\}$, gives $\mathcal{T}_\Phi(\nu,\mu) = \|\nu-\mu\|^*$.

Example

With $\Phi = \Phi_d$ or $\Phi = \Phi_{\|\cdot\|_{\mathrm{Lip}}}$,

$$\mathcal{T}_\Phi(\nu,\mu) = W_1(\nu,\mu) = \|\nu-\mu\|^*_{\mathrm{Lip}}.$$

The generalized transport cost

Definition (The generalized transport cost)

$$\mathcal{T}_\Phi(\nu,\mu) := \sup_{(u,v)\in\Phi} \{\nu(u) - \mu(v)\}, \quad \nu,\mu \in \mathcal{P}_\mathcal{X}.$$

- $\Phi = \Phi_c = \{(u,v) \text{ functions}; u(x) - v(y) \le c(x,y), \forall x,y \in \mathcal{X}\}$, gives $\mathcal{T}_\Phi = \mathcal{T}_c$.
- $\Phi = \Phi_{\|\cdot\|} := \{(u,u); u : \mathcal{X} \to \mathbb{R}, \|u\| \le 1\}$, gives $\mathcal{T}_\Phi(\nu,\mu) = \|\nu - \mu\|^*$.

Example

With $\Phi = \Phi_d$ or $\Phi = \Phi_{\|\cdot\|_{\mathrm{Lip}}}$,

$$\mathcal{T}_\Phi(\nu,\mu) = W_1(\nu,\mu) = \|\nu - \mu\|^*_{\mathrm{Lip}}.$$

The generalized transport cost

Definition (The generalized transport cost)

$$\mathcal{T}_\Phi(\nu,\mu) := \sup_{(u,v)\in\Phi} \{\nu(u) - \mu(v)\}, \quad \nu,\mu \in \mathcal{P}_\mathcal{X}.$$

- $\Phi = \Phi_c = \{(u,v) \text{ functions}; u(x) - v(y) \le c(x,y), \forall x,y \in \mathcal{X}\}$, gives $\mathcal{T}_\Phi = \mathcal{T}_c$.
- $\Phi = \Phi_{\|\cdot\|} := \{(u,u); u : \mathcal{X} \to \mathbb{R}, \|u\| \le 1\}$, gives $\mathcal{T}_\Phi(\nu,\mu) = \|\nu - \mu\|^*$.

Example

With $\Phi = \Phi_d$ or $\Phi = \Phi_{\|\cdot\|_{\mathrm{Lip}}}$,

$$\mathcal{T}_\Phi(\nu,\mu) = W_1(\nu,\mu) = \|\nu - \mu\|^*_{\mathrm{Lip}}.$$

The generalized transport cost

Definition (The generalized transport cost)

$$\mathcal{T}_\Phi(\nu,\mu) := \sup_{(u,v)\in\Phi} \{\nu(u) - \mu(v)\}, \quad \nu,\mu \in \mathcal{P}_\mathcal{X}.$$

- $\Phi = \Phi_c = \{(u,v) \text{ functions}; u(x) - v(y) \le c(x,y), \forall x,y \in \mathcal{X}\}$, gives $\mathcal{T}_\Phi = \mathcal{T}_c$.
- $\Phi = \Phi_{\|\cdot\|} := \{(u,u); u : \mathcal{X} \to \mathbb{R}, \|u\| \le 1\}$, gives $\mathcal{T}_\Phi(\nu,\mu) = \|\nu - \mu\|^*$.

Example

With $\Phi = \Phi_d$ or $\Phi = \Phi_{\|\cdot\|_{\mathrm{Lip}}}$,

$$\mathcal{T}_\Phi(\nu,\mu) = W_1(\nu,\mu) = \|\nu - \mu\|^*_{\mathrm{Lip}}.$$

Transport-Information Inequalities

Definition (The class $\mathcal{C}$)

The function $\alpha : [0, \infty] \mapsto [0, \infty]$ is in the class $\mathcal{C}$ if it is convex increasing left continuous and $\alpha(0) = 0$.

Definition (Transport-Information Inequality)

The probability measure μ satisfies $T_\Phi I(\alpha)$ with $\alpha \in \mathcal{C}$ if

$$\alpha(\mathcal{T}_\Phi(\nu, \mu)) \leq I(\nu|\mu), \text{ for all } \nu \in \mathcal{P}_\mathcal{X} \qquad T_\Phi I(\alpha)$$

Example ($W_1 I$, $W_2 I$)

The probability measure μ satisfies $W_1 I(c)$ or $W_2 I(c)$ if

$$\begin{aligned} \mathcal{T}_d(\nu, \mu)^2 &:= W_1^2(\nu, \mu) \leq 4c^2 I(\nu|\mu), \ \forall \nu \qquad W_1 I(c) \\ \mathcal{T}_{d^2}(\nu, \mu) &:= W_2^2(\nu, \mu) \leq 4c^2 I(\nu|\mu), \ \forall \nu \qquad W_2 I(c) \end{aligned}$$

Corresponds to $\alpha_1(r) = r^2/(4c^2)$ and $\alpha_2(r) = r/(4c^2)$, $r \geq 0$.

Transport-Information Inequalities

Definition (The class $\mathcal{C}$)

The function $\alpha : [0, \infty] \mapsto [0, \infty]$ is in the class $\mathcal{C}$ if it is convex increasing left continuous and $\alpha(0) = 0$.

Definition (Transport-Information Inequality)

The probability measure μ satisfies $T_\Phi I(\alpha)$ with $\alpha \in \mathcal{C}$ if

$$\alpha(\mathcal{T}_\Phi(\nu, \mu)) \leq I(\nu|\mu), \text{ for all } \nu \in \mathcal{P}_\mathcal{X} \qquad T_\Phi I(\alpha)$$

Example ($W_1 I$, $W_2 I$)

The probability measure μ satisfies $W_1 I(c)$ or $W_2 I(c)$ if

$$\mathcal{T}_d(\nu, \mu)^2 := W_1^2(\nu, \mu) \leq 4c^2 I(\nu|\mu), \forall \nu \qquad W_1 I(c)$$
$$\mathcal{T}_{d^2}(\nu, \mu) := W_2^2(\nu, \mu) \leq 4c^2 I(\nu|\mu), \forall \nu \qquad W_2 I(c)$$

Corresponds to $\alpha_1(r) = r^2/(4c^2)$ and $\alpha_2(r) = r/(4c^2)$, $r \geq 0$.

Transport-Information Inequalities

Definition (The class $\mathcal{C}$)

The function $\alpha : [0,\infty] \mapsto [0,\infty]$ is in the class $\mathcal{C}$ if it is convex increasing left continuous and $\alpha(0) = 0$.

Definition (Transport-Information Inequality)

The probability measure μ satisfies $T_\Phi I(\alpha)$ with $\alpha \in \mathcal{C}$ if

$$\alpha(\mathcal{T}_\Phi(\nu,\mu)) \le I(\nu|\mu), \text{ for all } \nu \in \mathcal{P}_\mathcal{X} \qquad T_\Phi I(\alpha)$$

Example ($W_1 I$, $W_2 I$)

The probability measure μ satisfies $W_1 I(c)$ or $W_2 I(c)$ if

$$\mathcal{T}_d(\nu,\mu)^2 := W_1^2(\nu,\mu) \le 4c^2 I(\nu|\mu), \ \forall \nu \qquad W_1 I(c)$$
$$\mathcal{T}_{d^2}(\nu,\mu) := W_2^2(\nu,\mu) \le 4c^2 I(\nu|\mu), \ \forall \nu \qquad W_2 I(c)$$

Corresponds to $\alpha_1(r) = r^2/(4c^2)$ and $\alpha_2(r) = r/(4c^2)$, $r \ge 0$.

Transport-Information Inequalities

Definition (The class $\mathcal{C}$)

The function $\alpha : [0, \infty] \mapsto [0, \infty]$ is in the class $\mathcal{C}$ if it is convex increasing left continuous and $\alpha(0) = 0$.

Definition (Transport-Information Inequality)

The probability measure μ satisfies $T_\Phi I(\alpha)$ with $\alpha \in \mathcal{C}$ if

$$\alpha(\mathcal{T}_\Phi(\nu, \mu)) \leq I(\nu|\mu), \text{ for all } \nu \in \mathcal{P}_\mathcal{X} \qquad T_\Phi I(\alpha)$$

Example ($W_1 I$, $W_2 I$)

The probability measure μ satisfies $W_1 I(c)$ or $W_2 I(c)$ if

$$\mathcal{T}_d(\nu, \mu)^2 := W_1^2(\nu, \mu) \leq 4c^2 I(\nu|\mu), \ \forall \nu \qquad W_1 I(c)$$

$$\mathcal{T}_{d^2}(\nu, \mu) := W_2^2(\nu, \mu) \leq 4c^2 I(\nu|\mu), \ \forall \nu \qquad W_2 I(c)$$

Corresponds to $\alpha_1(r) = r^2/(4c^2)$ and $\alpha_2(r) = r/(4c^2)$, $r \geq 0$.

Transport-Entropy Inequalities

- Relative entropy: $H(\nu|\mu) = \int_{\mathcal{X}} \log\left(\frac{d\nu}{d\mu}\right) \, d\nu.$

Definition (Transport-Entropy Inequality)

The probability measure μ satisfies $T_\Phi H(\alpha)$ if

$$\alpha\big(\mathcal{T}_\Phi(\nu, \mu)\big) \le H(\nu|\mu), \text{ for all } \nu \in \mathcal{P}_{\mathcal{X}} \qquad T_\Phi H(\alpha)$$

Example ([$W_1 H$, $W_2 H$])

$$\begin{array}{llllll} \mathcal{T}_d(\nu,\mu)^2 & := & W_1^2(\nu,\mu) & \le & cH(\nu|\mu), \ \forall \nu & W_1 H \\ \mathcal{T}_{d^2}(\nu,\mu) & := & W_2^2(\nu,\mu) & \le & cH(\nu|\mu), \ \forall \nu & W_2 H \end{array}$$

- $W_2 H$ is Talagrand's T_2 inequality.
- $T_\Phi H(\alpha)$ is investigated by N. Gozlan and CL (PTRF online).

Transport-Entropy Inequalities

- Relative entropy: $H(\nu|\mu) = \int_{\mathcal{X}} \log\left(\frac{d\nu}{d\mu}\right)\, d\nu.$

Definition (Transport-Entropy Inequality)

The probability measure μ satisfies $T_\Phi H(\alpha)$ if

$$\alpha\big(\mathcal{T}_\Phi(\nu,\mu)\big) \le H(\nu|\mu), \text{ for all } \nu \in \mathcal{P}_{\mathcal{X}} \qquad T_\Phi H(\alpha)$$

Example ([$W_1 H$, $W_2 H$)

$$\mathcal{T}_d(\nu,\mu)^2 := W_1^2(\nu,\mu) \le cH(\nu|\mu),\ \forall\nu \qquad W_1 H$$
$$\mathcal{T}_{d^2}(\nu,\mu) := W_2^2(\nu,\mu) \le cH(\nu|\mu),\ \forall\nu \qquad W_2 H$$

- $W_2 H$ is Talagrand's T_2 inequality.
- $T_\Phi H(\alpha)$ is investigated by N. Gozlan and CL (PTRF online).

Transport-Entropy Inequalities

- Relative entropy: $H(\nu|\mu) = \int_{\mathcal{X}} \log\left(\frac{d\nu}{d\mu}\right)\, d\nu.$

Definition (Transport-Entropy Inequality)

The probability measure μ satisfies $T_\Phi H(\alpha)$ if

$$\alpha\big(\mathcal{T}_\Phi(\nu,\mu)\big) \le H(\nu|\mu), \text{ for all } \nu \in \mathcal{P}_{\mathcal{X}} \qquad T_\Phi H(\alpha)$$

Example ([$W_1 H$, $W_2 H$)

$$\mathcal{T}_d(\nu,\mu)^2 := W_1^2(\nu,\mu) \le cH(\nu|\mu),\ \forall\nu \qquad W_1 H$$
$$\mathcal{T}_{d^2}(\nu,\mu) := W_2^2(\nu,\mu) \le cH(\nu|\mu),\ \forall\nu \qquad W_2 H$$

- $W_2 H$ is Talagrand's T_2 inequality.
- $T_\Phi H(\alpha)$ is investigated by N. Gozlan and CL (PTRF online).

Transport-Entropy Inequalities

- Relative entropy: $H(\nu|\mu) = \int_{\mathcal{X}} \log\left(\frac{d\nu}{d\mu}\right)\, d\nu.$

Definition (Transport-Entropy Inequality)

The probability measure μ satisfies $T_\Phi H(\alpha)$ if

$$\alpha\big(\mathcal{T}_\Phi(\nu,\mu)\big) \le H(\nu|\mu), \text{ for all } \nu \in \mathcal{P}_{\mathcal{X}} \qquad T_\Phi H(\alpha)$$

Example ([$W_1 H$, $W_2 H$)

$$\begin{array}{llllr} \mathcal{T}_d(\nu,\mu)^2 & := & W_1^2(\nu,\mu) & \le & cH(\nu|\mu),\ \forall\nu & W_1 H \\ \mathcal{T}_{d^2}(\nu,\mu) & := & W_2^2(\nu,\mu) & \le & cH(\nu|\mu),\ \forall\nu & W_2 H \end{array}$$

- $W_2 H$ is Talagrand's T_2 inequality.
- $T_\Phi H(\alpha)$ is investigated by N. Gozlan and CL (PTRF online).

Transport-Entropy Inequalities

- Relative entropy: $H(\nu|\mu) = \int_{\mathcal{X}} \log\left(\frac{d\nu}{d\mu}\right) d\nu.$

Definition (Transport-Entropy Inequality)

The probability measure μ satisfies $T_\Phi H(\alpha)$ if

$$\alpha\big(\mathcal{T}_\Phi(\nu,\mu)\big) \le H(\nu|\mu), \text{ for all } \nu \in \mathcal{P}_{\mathcal{X}} \qquad T_\Phi H(\alpha)$$

Example ([W_1H, W_2H)

$$\mathcal{T}_d(\nu,\mu)^2 := W_1^2(\nu,\mu) \le cH(\nu|\mu),\ \forall\nu \qquad W_1H$$
$$\mathcal{T}_{d^2}(\nu,\mu) := W_2^2(\nu,\mu) \le cH(\nu|\mu),\ \forall\nu \qquad W_2H$$

- W_2H is Talagrand's T_2 inequality.
- $T_\Phi H(\alpha)$ is investigated by N. Gozlan and CL (PTRF online).

Outline

The occupation measure

The main actor is the random probability measure

Definition (Occupation measure of the process X)

$$L_t := \frac{1}{t}\int_0^t \delta_{X_s}\,ds, \quad t \geq 0$$

$$L_t(A) = \text{ratio of time spent by } X \text{ in } A \text{ during } [0,t], \quad A \subset \mathcal{X}.$$

Let

- $\beta = \mathrm{Law}(X_0)$ is the initial law and
- $\mathbb{P}_\beta$ is the corresponding law of $(X_t)_{t\geq 0}$.

Theorem (Ergodic theorem)

$$\lim_{t\to\infty} L_t = \mu, \ \mathbb{P}_\beta\text{-a.s. with respect to } \sigma(\mathcal{P}_\mathcal{X}, B_\mathcal{X}).$$

The occupation measure

The main actor is the random probability measure

Definition (Occupation measure of the process X)

$$L_t := \frac{1}{t}\int_0^t \delta_{X_s}\, ds, \quad t \geq 0$$

$$L_t(A) = \text{ratio of time spent by } X \text{ in } A \text{ during } [0,t], \quad A \subset \mathcal{X}.$$

Let

- $\beta = \mathrm{Law}(X_0)$ is the initial law and
- $\mathbb{P}_\beta$ is the corresponding law of $(X_t)_{t\geq 0}$.

Theorem (Ergodic theorem)

$$\lim_{t\to\infty} L_t = \mu, \ \mathbb{P}_\beta\text{-a.s. } \textit{with respect to } \sigma(\mathcal{P}_\mathcal{X}, B_\mathcal{X}).$$

The occupation measure

The main actor is the random probability measure

Definition (Occupation measure of the process X)

$$L_t := \frac{1}{t}\int_0^t \delta_{X_s}\, ds, \quad t \geq 0$$

$$L_t(A) = \text{ratio of time spent by } X \text{ in } A \text{ during } [0,t], \quad A \subset \mathcal{X}.$$

Let

- $\beta = \mathrm{Law}(X_0)$ is the initial law and
- $\mathbb{P}_\beta$ is the corresponding law of $(X_t)_{t\geq 0}$.

Theorem (Ergodic theorem)

$$\lim_{t\to\infty} L_t = \mu, \ \mathbb{P}_\beta\text{-a.s. } \textit{with respect to } \sigma(\mathcal{P}_{\mathcal{X}}, B_{\mathcal{X}}).$$

Large deviations of $\{L_t\}_{t\geq 0}$

Theorem (Ergodic theorem)

$$\lim_{t\to\infty} L_t = \mu,\ \mathbb{P}_\beta\text{-a.s. } \textit{with respect to } \sigma(\mathcal{P}_\mathcal{X}, B_\mathcal{X}).$$

Theorem (Donsker-Varadhan, Deuschel-Stroock)

Under mixing assumptions,
$\{L_t\}$ obeys the Large Deviation Principle as $t\to\infty$:

$$\mathbb{P}_\beta(L_t\in A) \underset{t\to\infty}{\asymp} \exp\Big(-t\inf_{\nu\in A} I(\nu|\mu)\Big), \quad A\subset\mathcal{P}_\mathcal{X}$$

"uniformly" in the initial law β.

Remark: $I(\nu|\mu)=0$ iff $\nu=\mu$.

Large deviations of $\{L_t\}_{t\geq 0}$

Theorem (Ergodic theorem)

$$\lim_{t\to\infty} L_t = \mu,\ \mathbb{P}_\beta\text{-a.s. with respect to } \sigma(\mathcal{P}_\mathcal{X}, B_\mathcal{X}).$$

Theorem (Donsker-Varadhan, Deuschel-Stroock)

Under mixing assumptions,
$\{L_t\}$ *obeys the Large Deviation Principle as* $t \to \infty$*:*

$$\mathbb{P}_\beta(L_t \in A) \underset{t\to\infty}{\asymp} \exp\Big(-t \inf_{\nu\in A} I(\nu|\mu)\Big), \quad A \subset \mathcal{P}_\mathcal{X}$$

"uniformly" in the initial law β.

Remark: $I(\nu|\mu) = 0$ iff $\nu = \mu$.

Large deviations of $\{L_t\}_{t\geq 0}$

Theorem (Wu Liming)

Assume that $\beta \prec \mu$ with $d\beta/d\mu \in L^2(\mu)$.
$\{L_t\}_{t\geq 0}$ obeys the LDP in $\mathcal{P}_\mathcal{X}$ with rate function $\nu \mapsto I(\nu|\mu)$: For all measurable subset A of $\mathcal{P}_\mathcal{X}$,

$$
\begin{aligned}
-\inf_{\nu\in\mathrm{int}\ A} I(\nu|\mu) \leq \quad & \liminf_{t\to\infty} \frac{1}{t}\log \mathbb{P}_\beta(L_t\in A) \\
\leq \quad & \limsup_{t\to\infty} \frac{1}{t}\log \mathbb{P}_\beta(L_t\in A) \leq \quad -\inf_{\nu\in\mathrm{cl}\ A} I(\nu|\mu)
\end{aligned}
$$

where $\mathrm{int}\ A$ and $\mathrm{cl}\ A$ are the interior and closure of A with respect to the topology $\sigma(\mathcal{P}_\mathcal{X}, B_\mathcal{X})$.

Shortly and approximately: $\mathbb{P}_\beta(L_t\in A) \underset{t\to\infty}{\asymp} \exp[-t\inf_{\nu\in A} I(\nu|\mu)]$

- Remark: $\mathbb{P}_\beta = \frac{d\beta}{d\mu}(X_0)\mathbb{P}_\mu$.
- No mixing assumptions.

Large deviations of $\{L_t\}_{t\geq 0}$

Theorem (Wu Liming)

Assume that $\beta \prec \mu$ *with* $d\beta/d\mu \in L^2(\mu)$.
$\{L_t\}_{t\geq 0}$ *obeys the LDP in* $\mathcal{P}_\mathcal{X}$ *with rate function* $\nu \mapsto I(\nu|\mu)$: *For all measurable subset A of* $\mathcal{P}_\mathcal{X}$,

$$
\begin{aligned}
-\inf_{\nu\in \mathrm{int}\ A} I(\nu|\mu) \leq &\liminf_{t\to\infty} \frac{1}{t}\log \mathbb{P}_\beta(L_t\in A)\\
\leq &\limsup_{t\to\infty} \frac{1}{t}\log \mathbb{P}_\beta(L_t\in A) \leq -\inf_{\nu\in \mathrm{cl}\ A} I(\nu|\mu)
\end{aligned}
$$

where $\mathrm{int}\ A$ *and* $\mathrm{cl}\ A$ *are the interior and closure of A with respect to the topology* $\sigma(\mathcal{P}_\mathcal{X}, B_\mathcal{X})$.

Shortly and approximately: $\mathbb{P}_\beta(L_t\in A) \underset{t\to\infty}{\asymp} \exp[-t\inf_{\nu\in A} I(\nu|\mu)]$

- Remark: $\mathbb{P}_\beta = \frac{d\beta}{d\mu}(X_0)\mathbb{P}_\mu$.
- No mixing assumptions.

Large deviations of $\{L_t\}_{t\geq 0}$

Theorem (Wu Liming)

Assume that $\beta \prec \mu$ with $d\beta/d\mu \in L^2(\mu)$.
$\{L_t\}_{t\geq 0}$ obeys the LDP in $\mathcal{P}_\mathcal{X}$ with rate function $\nu \mapsto I(\nu|\mu)$: For all measurable subset A of $\mathcal{P}_\mathcal{X}$,

$$\begin{aligned} -\inf_{\nu\in \mathrm{int}\ A} I(\nu|\mu) \leq & \liminf_{t\to\infty} \frac{1}{t}\log \mathbb{P}_\beta(L_t\in A) \\ \leq & \limsup_{t\to\infty} \frac{1}{t}\log \mathbb{P}_\beta(L_t\in A) \leq -\inf_{\nu\in \mathrm{cl}\ A} I(\nu|\mu) \end{aligned}$$

where $\mathrm{int}\ A$ and $\mathrm{cl}\ A$ are the interior and closure of A with respect to the topology $\sigma(\mathcal{P}_\mathcal{X}, B_\mathcal{X})$.

Shortly and approximately: $\mathbb{P}_\beta(L_t\in A) \underset{t\to\infty}{\asymp} \exp[-t\inf_{\nu\in A} I(\nu|\mu)]$

- Remark: $\mathbb{P}_\beta = \frac{d\beta}{d\mu}(X_0)\mathbb{P}_\mu$.
- No mixing assumptions.

Large deviations of $\{L_t\}_{t\geq 0}$

Theorem (Wu Liming)

Assume that $\beta \prec \mu$ *with* $d\beta/d\mu \in L^2(\mu)$.
$\{L_t\}_{t\geq 0}$ *obeys the LDP in* $\mathcal{P}_\mathcal{X}$ *with rate function* $\nu \mapsto I(\nu|\mu)$: *For all measurable subset* A *of* $\mathcal{P}_\mathcal{X}$,

$$\begin{aligned} -\inf_{\nu\in \mathrm{int}\, A} I(\nu|\mu) \leq & \liminf_{t\to\infty} \frac{1}{t} \log \mathbb{P}_\beta(L_t \in A) \\ \leq & \limsup_{t\to\infty} \frac{1}{t} \log \mathbb{P}_\beta(L_t \in A) \leq -\inf_{\nu\in \mathrm{cl}\, A} I(\nu|\mu) \end{aligned}$$

where $\mathrm{int}\, A$ *and* $\mathrm{cl}\, A$ *are the interior and closure of* A *with respect to the topology* $\sigma(\mathcal{P}_\mathcal{X}, B_\mathcal{X})$.

Shortly and approximately: $\mathbb{P}_\beta(L_t \in A) \underset{t\to\infty}{\asymp} \exp[-t \inf_{\nu\in A} I(\nu|\mu)]$

- Remark: $\mathbb{P}_\beta = \frac{d\beta}{d\mu}(X_0)\mathbb{P}_\mu$.
- No mixing assumptions.

Large deviations of $\{L_t\}_{t\geq 0}$

Theorem (Wu Liming)

Assume that $\beta \prec \mu$ with $d\beta/d\mu \in L^2(\mu)$.
$\{L_t\}_{t\geq 0}$ obeys the LDP in $\mathcal{P}_\mathcal{X}$ with rate function $\nu \mapsto I(\nu|\mu)$: For all measurable subset A of $\mathcal{P}_\mathcal{X}$,

$$
\begin{aligned}
-\inf_{\nu\in \mathrm{int}\, A} I(\nu|\mu) \leq \quad & \liminf_{t\to\infty} \frac{1}{t} \log \mathbb{P}_\beta(L_t \in A) \\
\leq \quad & \limsup_{t\to\infty} \frac{1}{t} \log \mathbb{P}_\beta(L_t \in A) \leq \quad -\inf_{\nu\in \mathrm{cl}\, A} I(\nu|\mu)
\end{aligned}
$$

where $\mathrm{int}\, A$ and $\mathrm{cl}\, A$ are the interior and closure of A with respect to the topology $\sigma(\mathcal{P}_\mathcal{X}, B_\mathcal{X})$.

Shortly and approximately: $\mathbb{P}_\beta(L_t \in A) \underset{t\to\infty}{\asymp} \exp[-t \inf_{\nu\in A} I(\nu|\mu)]$

- Remark: $\mathbb{P}_\beta = \frac{d\beta}{d\mu}(X_0)\mathbb{P}_\mu$.
- No mixing assumptions.

Large deviations of $\{L_t\}_{t\geq 0}$

Theorem (Wu Liming)

Assume that $\beta \prec \mu$ *with* $d\beta/d\mu \in L^2(\mu)$.
$\{L_t\}_{t\geq 0}$ *obeys the LDP in* $\mathcal{P}_\mathcal{X}$ *with rate function* $\nu \mapsto I(\nu|\mu)$: *For all measurable subset A of* $\mathcal{P}_\mathcal{X}$,

$$\begin{aligned} -\inf_{\nu\in \mathrm{int}\ A} I(\nu|\mu) \leq & \liminf_{t\to\infty} \frac{1}{t} \log \mathbb{P}_\beta(L_t \in A) \\ \leq & \limsup_{t\to\infty} \frac{1}{t} \log \mathbb{P}_\beta(L_t \in A) \leq -\inf_{\nu\in \mathrm{cl}\ A} I(\nu|\mu) \end{aligned}$$

where $\mathrm{int}\ A$ *and* $\mathrm{cl}\ A$ *are the interior and closure of A with respect to the topology* $\sigma(\mathcal{P}_\mathcal{X}, B_\mathcal{X})$.

Shortly and approximately: $\mathbb{P}_\beta(L_t \in A) \underset{t\to\infty}{\asymp} \exp[-t \inf_{\nu\in A} I(\nu|\mu)]$

- Remark: $\mathbb{P}_\beta = \frac{d\beta}{d\mu}(X_0)\mathbb{P}_\mu$.
- No mixing assumptions.

Outline

Statement of the basic result

For $\alpha \in \mathcal{C}$, define: $\alpha^{\circledast}(\lambda) := \sup_{r\geq 0}\{r\lambda - \alpha(r)\}$, $\lambda \geq 0$.

Theorem

The following statements are equivalent.

(a) $\alpha(\mathcal{T}_\Phi(\nu,\mu)) \leq I(\nu|\mu), \quad \forall \nu \in \mathcal{P}_\mathcal{X}.$

(b) $\forall \lambda \geq 0, \forall (u,v) \in \Phi,$

$$\limsup_{t\to\infty} \frac{1}{t} \log \mathbb{E}_\mu \exp\left(\lambda \int_0^t u(X_s)\,ds\right) \leq \lambda\mu + \alpha^{\circledast}(\lambda)$$

(c) $\forall r,t > 0, \forall (u,v) \in \Phi, \forall \beta \in \mathcal{P}_\mathcal{X}$ *such that* $d\beta/d\mu \in L^2(\mu)$,

$$\mathbb{P}_\beta\left(\frac{1}{t}\int_0^t u(X_s)\,ds \geq \mu(v) + r\right) \leq \left\|\frac{d\beta}{d\mu}\right\|_{L^2(\mu)} e^{-t\alpha(r)}$$

- Remark: $\frac{1}{t}\int_0^t u(X_s)\,ds = \int_\mathcal{X} u\,dL_t = L_t(u)$.

Statement of the basic result

For $\alpha \in \mathcal{C}$, define: $\alpha^{\circledast}(\lambda) := \sup_{r \geq 0}\{r\lambda - \alpha(r)\}, \quad \lambda \geq 0.$

Theorem

The following statements are equivalent.

(a) $\alpha(\mathcal{T}_\Phi(\nu, \mu)) \leq I(\nu|\mu), \quad \forall \nu \in \mathcal{P}_\mathcal{X}.$

(b) $\forall \lambda \geq 0, \forall (u, v) \in \Phi,$

$$\limsup_{t\to\infty} \frac{1}{t} \log \mathbb{E}_\mu \exp\left(\lambda \int_0^t u(X_s)\, ds\right) \leq \lambda\mu + \alpha^{\circledast}(\lambda)$$

(c) $\forall r, t > 0, \forall (u, v) \in \Phi, \forall \beta \in \mathcal{P}_\mathcal{X}$ *such that* $d\beta/d\mu \in L^2(\mu)$,

$$\mathbb{P}_\beta\left(\frac{1}{t}\int_0^t u(X_s)\, ds \geq \mu(v) + r\right) \leq \left\|\frac{d\beta}{d\mu}\right\|_{L^2(\mu)} e^{-t\alpha(r)}$$

- Remark: $\frac{1}{t}\int_0^t u(X_s)\, ds = \int_\mathcal{X} u\, dL_t = L_t(u).$

Statement of the basic result

For $\alpha \in \mathcal{C}$, define: $\quad \alpha^{\circledast}(\lambda) := \sup_{r \geq 0}\{r\lambda - \alpha(r)\}, \quad \lambda \geq 0.$

Theorem

The following statements are equivalent.

(a) $\alpha(\mathcal{T}_\Phi(\nu, \mu)) \leq I(\nu|\mu), \quad \forall \nu \in \mathcal{P}_\mathcal{X}.$

(b) $\forall \lambda \geq 0, \forall (u, v) \in \Phi,$

$$\limsup_{t\to\infty} \frac{1}{t} \log \mathbb{E}_\mu \exp\left(\lambda \int_0^t u(X_s)\, ds\right) \leq \lambda\mu + \alpha^{\circledast}(\lambda)$$

(c) $\forall r, t > 0, \forall (u, v) \in \Phi, \forall \beta \in \mathcal{P}_\mathcal{X}$ *such that* $d\beta/d\mu \in L^2(\mu)$,

$$\mathbb{P}_\beta\left(\frac{1}{t}\int_0^t u(X_s)\, ds \geq \mu(v) + r\right) \leq \left\|\frac{d\beta}{d\mu}\right\|_{L^2(\mu)} e^{-t\alpha(r)}$$

- Remark: $\frac{1}{t}\int_0^t u(X_s)\, ds = \int_\mathcal{X} u\, dL_t = L_t(u).$

Statement of the basic result

For $\alpha \in \mathcal{C}$, define: $\alpha^{\circledast}(\lambda) := \sup_{r \geq 0}\{r\lambda - \alpha(r)\}, \quad \lambda \geq 0.$

Theorem

The following statements are equivalent.

(a) $\alpha(\mathcal{T}_\Phi(\nu, \mu)) \leq I(\nu|\mu), \quad \forall \nu \in \mathcal{P}_{\mathcal{X}}.$

(b) $\forall \lambda \geq 0, \forall (u, v) \in \Phi,$

$$\limsup_{t \to \infty} \frac{1}{t} \log \mathbb{E}_\mu \exp\left(\lambda \int_0^t u(X_s)\, ds\right) \leq \lambda\mu + \alpha^{\circledast}(\lambda)$$

(c) $\forall r, t > 0, \forall (u, v) \in \Phi, \forall \beta \in \mathcal{P}_{\mathcal{X}}$ *such that* $d\beta/d\mu \in L^2(\mu)$,

$$\mathbb{P}_\beta\left(\frac{1}{t}\int_0^t u(X_s)\, ds \geq \mu(v) + r\right) \leq \left\|\frac{d\beta}{d\mu}\right\|_{L^2(\mu)} e^{-t\alpha(r)}$$

- Remark: $\frac{1}{t}\int_0^t u(X_s)\, ds = \int_{\mathcal{X}} u\, dL_t = L_t(u).$

An immediate corollary

Specializing with W_1 one obtains the

Corollary

The following statements are equivalent.

(a) $W_1^2(\nu, \mu) \le 4c^2 I(\nu|\mu), \quad \forall \nu;$

(b) $\forall \lambda \ge 0, \forall u \in C_{\mathrm{Lip}} : \|u\|_{\mathrm{Lip}} \le 1, \mu(u) = 0,$

$$\limsup_{t\to\infty} \frac{1}{t} \log \mathbb{E}_\mu \exp\left(\lambda \int_0^t u(X_s)\, ds\right) \le c^2 \lambda^2$$

(c) $\forall r, t > 0, \forall u \in C_{\mathrm{Lip}}, \forall \beta \in \mathcal{P}_{\mathcal{X}}$ *such that* $d\beta/d\mu \in L^2(\mu),$

$$\mathbb{P}_\beta\left(\frac{1}{t}\int_0^t u(X_s)\, ds \ge \mu(u) + r\right) \le \left\|\frac{d\beta}{d\mu}\right\|_{L^2(\mu)} \exp\left(-\frac{tr^2}{4c^2\|u\|_{\mathrm{Lip}}^2}\right)$$

- Interest: Estimation of $\mu(u)$ by $L_t(u)$.
- Analogous result for $W_2 I$.

An immediate corollary

Specializing with W_1 one obtains the

Corollary

The following statements are equivalent.

(a) $W_1^2(\nu, \mu) \le 4c^2 I(\nu|\mu), \quad \forall \nu;$

(b) $\forall \lambda \ge 0, \forall u \in C_{\mathrm{Lip}} : \|u\|_{\mathrm{Lip}} \le 1, \mu(u) = 0,$

$$\limsup_{t\to\infty} \frac{1}{t} \log \mathbb{E}_\mu \exp\left(\lambda \int_0^t u(X_s)\, ds\right) \le c^2 \lambda^2$$

(c) $\forall r, t > 0, \forall u \in C_{\mathrm{Lip}}, \forall \beta \in \mathcal{P}_{\mathcal{X}}$ *such that* $d\beta/d\mu \in L^2(\mu)$,

$$\mathbb{P}_\beta\left(\frac{1}{t}\int_0^t u(X_s)\, ds \ge \mu(u) + r\right) \le \left\|\frac{d\beta}{d\mu}\right\|_{L^2(\mu)} \exp\left(-\frac{tr^2}{4c^2\|u\|_{\mathrm{Lip}}^2}\right)$$

- Interest: Estimation of $\mu(u)$ by $L_t(u)$.
- Analogous result for $W_2 I$.

An immediate corollary

Specializing with W_1 one obtains the

Corollary

The following statements are equivalent.

(a) $W_1^2(\nu,\mu) \le 4c^2 I(\nu|\mu), \quad \forall\nu;$

(b) $\forall\lambda \ge 0, \forall u \in C_{\mathrm{Lip}} : \|u\|_{\mathrm{Lip}} \le 1, \mu(u) = 0,$

$$\limsup_{t\to\infty} \frac{1}{t} \log \mathbb{E}_\mu \exp\left(\lambda \int_0^t u(X_s)\, ds\right) \le c^2\lambda^2$$

(c) $\forall r, t > 0, \forall u \in C_{\mathrm{Lip}}, \forall\beta \in \mathcal{P}_{\mathcal{X}}$ *such that* $d\beta/d\mu \in L^2(\mu)$,

$$\mathbb{P}_\beta\left(\frac{1}{t}\int_0^t u(X_s)\, ds \ge \mu(u) + r\right) \le \left\|\frac{d\beta}{d\mu}\right\|_{L^2(\mu)} \exp\left(-\frac{tr^2}{4c^2\|u\|_{\mathrm{Lip}}^2}\right)$$

- Interest: Estimation of $\mu(u)$ by $L_t(u)$.
- Analogous result for $W_2 I$.

Deviation functions, Transport functions

Definition (Transport function)

A function α in $\mathcal{C}$ is a transport function if μ satisfies $T_\Phi I(\alpha)$

$$\alpha\big(\mathcal{T}_\Phi(\mu,\nu)\big) \le I(\nu|\mu), \text{ for all } \nu$$

Denote

$$\mathcal{T}(L_t) = \mathcal{T}_\Phi(L_t,\mu)$$

Definition (Deviation function)

A function α in $\mathcal{C}$ is a deviation function if for all $r \geq 0$

$$\limsup_{t\to\infty} \frac{1}{t} \log \mathbb{P}_\beta(\mathcal{T}(L_t) \geq r) \leq -\alpha(r)$$

Deviation functions, Transport functions

Definition (Transport function)

A function α in $\mathcal{C}$ is a transport function if μ satisfies $T_\Phi I(\alpha)$

$$\alpha\big(\mathcal{T}_\Phi(\mu,\nu)\big) \le I(\nu|\mu), \text{ for all } \nu$$

Denote

$$\mathcal{T}(L_t) = \mathcal{T}_\Phi(L_t,\mu)$$

Definition (Deviation function)

A function α in $\mathcal{C}$ is a deviation function if for all $r \ge 0$

$$\limsup_{t\to\infty} \frac{1}{t} \log \mathbb{P}_\beta(\mathcal{T}(L_t) \ge r) \le -\alpha(r)$$

A deviation function is a transport function

Borrowed from [Gozlan-Léonard].

Recipe

Any deviation function α is a transport function.

Idea of proof.

- Suppose $\mathcal{T}$ is regular enough for the sets $A_r = \{\nu \in \mathcal{P}_\mathcal{X}; \mathcal{T}(\nu) \geq r\}$ to be good sets. For all $r \geq 0$,

$$\mathbb{P}_\beta(\mathcal{T}(L_t) \geq r) = \mathbb{P}_\beta(L_t \in A_r) \asymp \exp[-ti(r)]$$

  with $i(r) = \inf\{I(\nu|\mu); \nu \in \mathcal{P}_\mathcal{X} : \mathcal{T}(\nu) \geq r\}$.
- Let α be a deviation function. Then, $\exp[-ti(r)] \lessapprox \exp[-t\alpha(r)], \forall r \geq 0$ i.e. $\alpha(r) \leq i(r), \forall r \geq 0$.
- Taking $r = \mathcal{T}(\nu)$ leads to: $\alpha(\mathcal{T}(\nu)) \leq i(\mathcal{T}(\nu)) \leq I(\nu|\mu), \forall \nu \in \mathcal{P}_\mathcal{X}$. □

Mostly based on the LD lower bound.

A deviation function is a transport function

Borrowed from [Gozlan-Léonard].

Recipe

Any deviation function α is a transport function.

Idea of proof.

- Suppose $\mathcal{T}$ is regular enough for the sets $A_r = \{\nu \in \mathcal{P}_{\mathcal{X}}; \mathcal{T}(\nu) \geq r\}$ to be good sets. For all $r \geq 0$,

$$\mathbb{P}_\beta(\mathcal{T}(L_t) \geq r) = \mathbb{P}_\beta(L_t \in A_r) \asymp \exp[-ti(r)]$$

  with $i(r) = \inf\{I(\nu|\mu); \nu \in \mathcal{P}_{\mathcal{X}} : \mathcal{T}(\nu) \geq r\}$.
- Let α be a deviation function. Then, $\exp[-ti(r)] \lessapprox \exp[-t\alpha(r)], \forall r \geq 0$ i.e. $\alpha(r) \leq i(r), \forall r \geq 0$.
- Taking $r = \mathcal{T}(\nu)$ leads to: $\alpha(\mathcal{T}(\nu)) \leq i(\mathcal{T}(\nu)) \leq I(\nu|\mu), \forall \nu \in \mathcal{P}_{\mathcal{X}}$. □

Mostly based on the LD lower bound.

A deviation function is a transport function

Borrowed from [Gozlan-Léonard].

Recipe

Any deviation function α is a transport function.

Idea of proof.

- Suppose $\mathcal{T}$ is regular enough for the sets $A_r = \{\nu \in \mathcal{P}_\mathcal{X}; \mathcal{T}(\nu) \geq r\}$ to be good sets. For all $r \geq 0$,

$$\mathbb{P}_\beta(\mathcal{T}(L_t) \geq r) = \mathbb{P}_\beta(L_t \in A_r) \asymp \exp[-ti(r)]$$

with $i(r) = \inf\{I(\nu|\mu); \nu \in \mathcal{P}_\mathcal{X} : \mathcal{T}(\nu) \geq r\}$.

- Let α be a deviation function. Then, $\exp[-ti(r)] \lessapprox \exp[-t\alpha(r)], \forall r \geq 0$ i.e. $\alpha(r) \leq i(r), \forall r \geq 0$.
- Taking $r = \mathcal{T}(\nu)$ leads to: $\alpha(\mathcal{T}(\nu)) \leq i(\mathcal{T}(\nu)) \leq I(\nu|\mu), \forall \nu \in \mathcal{P}_\mathcal{X}$. □

Mostly based on the LD lower bound.

A deviation function is a transport function

Borrowed from [Gozlan-Léonard].

Recipe

Any deviation function α is a transport function.

Idea of proof.

- Suppose $\mathcal{T}$ is regular enough for the sets $A_r = \{\nu \in \mathcal{P}_\mathcal{X}; \mathcal{T}(\nu) \geq r\}$ to be good sets. For all $r \geq 0$,

$$\mathbb{P}_\beta(\mathcal{T}(L_t) \geq r) = \mathbb{P}_\beta(L_t \in A_r) \asymp \exp[-ti(r)]$$

 with $i(r) = \inf\{I(\nu|\mu); \nu \in \mathcal{P}_\mathcal{X} : \mathcal{T}(\nu) \geq r\}$.
- Let α be a deviation function. Then, $\exp[-ti(r)] \lessapprox \exp[-t\alpha(r)], \forall r \geq 0$ i.e. $\alpha(r) \leq i(r), \forall r \geq 0$.
- Taking $r = \mathcal{T}(\nu)$ leads to: $\alpha(\mathcal{T}(\nu)) \leq i(\mathcal{T}(\nu)) \leq I(\nu|\mu), \forall \nu \in \mathcal{P}_\mathcal{X}$. □

Mostly based on the LD lower bound.

A deviation function is a transport function

Borrowed from [Gozlan-Léonard].

Recipe

Any deviation function α is a transport function.

Idea of proof.

- Suppose $\mathcal{T}$ is regular enough for the sets $A_r = \{\nu \in \mathcal{P}_\mathcal{X}; \mathcal{T}(\nu) \geq r\}$ to be good sets. For all $r \geq 0$,

$$\mathbb{P}_\beta(\mathcal{T}(L_t) \geq r) = \mathbb{P}_\beta(L_t \in A_r) \asymp \exp[-ti(r)]$$

with $i(r) = \inf\{I(\nu|\mu); \nu \in \mathcal{P}_\mathcal{X} : \mathcal{T}(\nu) \geq r\}$.

- Let α be a deviation function. Then, $\exp[-ti(r)] \lesssim \exp[-t\alpha(r)], \forall r \geq 0$ i.e. $\alpha(r) \leq i(r), \forall r \geq 0$.
- Taking $r = \mathcal{T}(\nu)$ leads to: $\alpha(\mathcal{T}(\nu)) \leq i(\mathcal{T}(\nu)) \leq I(\nu|\mu), \forall \nu \in \mathcal{P}_\mathcal{X}$. □

Mostly based on the LD lower bound.

A deviation function is a transport function

Borrowed from [Gozlan-Léonard].

Recipe

Any deviation function α is a transport function.

Idea of proof.

- Suppose $\mathcal{T}$ is regular enough for the sets $A_r = \{\nu \in \mathcal{P}_\mathcal{X}; \mathcal{T}(\nu) \geq r\}$ to be good sets. For all $r \geq 0$,

$$\mathbb{P}_\beta(\mathcal{T}(L_t) \geq r) = \mathbb{P}_\beta(L_t \in A_r) \asymp \exp[-ti(r)]$$

  with $i(r) = \inf\{I(\nu|\mu); \nu \in \mathcal{P}_\mathcal{X} : \mathcal{T}(\nu) \geq r\}$.
- Let α be a deviation function. Then, $\exp[-ti(r)] \lessapprox \exp[-t\alpha(r)], \forall r \geq 0$ i.e. $\alpha(r) \leq i(r), \forall r \geq 0$.
- Taking $r = \mathcal{T}(\nu)$ leads to: $\alpha(\mathcal{T}(\nu)) \leq i(\mathcal{T}(\nu)) \leq I(\nu|\mu), \forall \nu \in \mathcal{P}_\mathcal{X}$. □

Mostly based on the LD lower bound.

A deviation function is a transport function

Borrowed from [Gozlan-Léonard].

Recipe

Any deviation function α is a transport function.

Idea of proof.

- Suppose $\mathcal{T}$ is regular enough for the sets $A_r = \{\nu \in \mathcal{P}_\mathcal{X}; \mathcal{T}(\nu) \geq r\}$ to be good sets. For all $r \geq 0$,

$$\mathbb{P}_\beta(\mathcal{T}(L_t) \geq r) = \mathbb{P}_\beta(L_t \in A_r) \asymp \exp[-ti(r)]$$

  with $i(r) = \inf\{I(\nu|\mu); \nu \in \mathcal{P}_\mathcal{X} : \mathcal{T}(\nu) \geq r\}$.
- Let α be a deviation function. Then, $\exp[-ti(r)] \lessapprox \exp[-t\alpha(r)], \forall r \geq 0$ i.e. $\alpha(r) \leq i(r), \forall r \geq 0$.
- Taking $r = \mathcal{T}(\nu)$ leads to: $\alpha(\mathcal{T}(\nu)) \leq i(\mathcal{T}(\nu)) \leq I(\nu|\mu), \forall \nu \in \mathcal{P}_\mathcal{X}$. □

Mostly based on the LD lower bound.

Outline

$T_cI(\alpha)$ for product measures

The ingredients

- $(\mathcal{X}_1, \mu_1), (\mathcal{X}_2, \mu_2), (\mathcal{X}_1 \times \mathcal{X}_2, \mu_1 \otimes \mu_2)$
- $c_1, c_2, c_1 \oplus c_2\Big((x_1, x_2), (y_1, y_2)\Big) = c_1(x_1, y_1) + c_2(x_2, y_2)$

Theorem (Tensorization)

Suppose that the following T_cIs hold:

$$\alpha_1(\mathcal{T}_{c_1}(\nu_1, \mu_1)) \leq I(\nu_1 \mid \mu_1), \ \forall \nu_1 \in \mathcal{P}(\mathcal{X}_1)$$
$$\alpha_2(\mathcal{T}_{c_2}(\nu_2, \mu_2)) \leq I(\nu_2 \mid \mu_2), \ \forall \nu_2 \in \mathcal{P}(\mathcal{X}_2)$$

Then, on $\mathcal{X}_1 \times \mathcal{X}_2$, we have:

$$\alpha_1 \square \alpha_2(\mathcal{T}_{c_1 \oplus c_2}(\nu, \mu_1 \otimes \mu_2)) \leq I(\nu \mid \mu_1 \otimes \mu_2), \quad \forall \nu \in \mathcal{P}(\mathcal{X}_1 \times \mathcal{X}_2)$$

- $\alpha_1 \square \alpha_2(r) = \inf\{\alpha_1(r_1) + \alpha_2(r_2); r_1, r_2 \geq 0 : r_1 + r_2 = r\}$
  is the inf-convolution of α_1 and α_2.

$T_cI(\alpha)$ for product measures

The ingredients

- $(\mathcal{X}_1, \mu_1), (\mathcal{X}_2, \mu_2), (\mathcal{X}_1 \times \mathcal{X}_2, \mu_1 \otimes \mu_2)$
- $c_1, c_2, c_1 \oplus c_2\Big((x_1, x_2), (y_1, y_2)\Big) = c_1(x_1, y_1) + c_2(x_2, y_2)$

Theorem (Tensorization)

Suppose that the following T_cIs hold:

$$\alpha_1(\mathcal{T}_{c_1}(\nu_1, \mu_1)) \le I(\nu_1 \mid \mu_1), \ \forall \nu_1 \in \mathcal{P}(\mathcal{X}_1)$$
$$\alpha_2(\mathcal{T}_{c_2}(\nu_2, \mu_2)) \le I(\nu_2 \mid \mu_2), \ \forall \nu_2 \in \mathcal{P}(\mathcal{X}_2)$$

Then, on $\mathcal{X}_1 \times \mathcal{X}_2$, we have:

$$\alpha_1 \Box \alpha_2(\mathcal{T}_{c_1 \oplus c_2}(\nu, \mu_1 \otimes \mu_2)) \le I(\nu \mid \mu_1 \otimes \mu_2), \quad \forall \nu \in \mathcal{P}(\mathcal{X}_1 \times \mathcal{X}_2)$$

- $\alpha_1 \Box \alpha_2(r) = \inf\{\alpha_1(r_1) + \alpha_2(r_2); r_1, r_2 \ge 0 : r_1 + r_2 = r\}$
  is the inf-convolution of α_1 and α_2.

$T_cI(\alpha)$ for product measures

The ingredients

- $(\mathcal{X}_1, \mu_1), (\mathcal{X}_2, \mu_2), (\mathcal{X}_1 \times \mathcal{X}_2, \mu_1 \otimes \mu_2)$
- $c_1, c_2, c_1 \oplus c_2\Big((x_1, x_2), (y_1, y_2)\Big) = c_1(x_1, y_1) + c_2(x_2, y_2)$

Theorem (Tensorization)

Suppose that the following T_cIs hold:

$$\alpha_1(\mathcal{T}_{c_1}(\nu_1, \mu_1)) \leq I(\nu_1 \mid \mu_1), \ \forall \nu_1 \in \mathcal{P}(\mathcal{X}_1)$$
$$\alpha_2(\mathcal{T}_{c_2}(\nu_2, \mu_2)) \leq I(\nu_2 \mid \mu_2), \ \forall \nu_2 \in \mathcal{P}(\mathcal{X}_2)$$

Then, on $\mathcal{X}_1 \times \mathcal{X}_2$, we have:

$$\alpha_1 \square \alpha_2\big(\mathcal{T}_{c_1 \oplus c_2}(\nu, \mu_1 \otimes \mu_2)\big) \leq I(\nu \mid \mu_1 \otimes \mu_2), \quad \forall \nu \in \mathcal{P}(\mathcal{X}_1 \times \mathcal{X}_2)$$

- $\alpha_1 \square \alpha_2(r) = \inf\{\alpha_1(r_1) + \alpha_2(r_2); r_1, r_2 \geq 0 : r_1 + r_2 = r\}$
is the inf-convolution of α_1 and α_2.

$T_cI(\alpha)$ for product measures

The ingredients

- $(\mathcal{X}_1, \mu_1), (\mathcal{X}_2, \mu_2), (\mathcal{X}_1 \times \mathcal{X}_2, \mu_1 \otimes \mu_2)$
- $c_1, c_2, c_1 \oplus c_2\Big((x_1, x_2), (y_1, y_2)\Big) = c_1(x_1, y_1) + c_2(x_2, y_2)$

Theorem (Tensorization)

Suppose that the following T_cIs hold:

$$\alpha_1(\mathcal{T}_{c_1}(\nu_1, \mu_1)) \le I(\nu_1 \mid \mu_1), \ \forall \nu_1 \in \mathcal{P}(\mathcal{X}_1)$$
$$\alpha_2(\mathcal{T}_{c_2}(\nu_2, \mu_2)) \le I(\nu_2 \mid \mu_2), \ \forall \nu_2 \in \mathcal{P}(\mathcal{X}_2)$$

Then, on $\mathcal{X}_1 \times \mathcal{X}_2$, we have:

$$\alpha_1 \square \alpha_2 (\mathcal{T}_{c_1 \oplus c_2}(\nu, \mu_1 \otimes \mu_2)) \le I(\nu \mid \mu_1 \otimes \mu_2), \quad \forall \nu \in \mathcal{P}(\mathcal{X}_1 \times \mathcal{X}_2)$$

- $\alpha_1 \square \alpha_2(r) = \inf\{\alpha_1(r_1) + \alpha_2(r_2); r_1, r_2 \ge 0 : r_1 + r_2 = r\}$ is the inf-convolution of α_1 and α_2.

$T_cI(\alpha)$ for product measures

Typical LD result

Let $\{Z_t^1\}_t$ and $\{Z_t^2\}_t$ be two *independent* LD systems on $\mathbb{R}$ with respective rate functions α_1 and α_2. Then, $\{Z_t := Z_t^1 + Z_t^2\}_t$ obeys the LDP with rate function $\alpha = \alpha_1 \Box \alpha_2$.

n-Tensorization

On $\mathcal{X}^n$ define the sum-cost

- $\bigoplus c(x_1,\dots,x_n;y_1,\dots,y_n) := \sum_{1\le i\le n} c(x_i;y_i), \quad x_1,\cdots,y_n \in \mathcal{X}$

and the sum-Dirichlet form

- $\bigoplus \mathcal{E}(g) := \int_{\mathcal{X}^n} \sum_{1\le i\le n} \mathcal{E}(g_i^{\hat{x}_i}) \prod_{1\le i\le n} \mu(dx_i)$

where $g_i^{\hat{x}_i}(x_i) = g(x_1,\dots,x_n)$ with $\hat{x}_i = (x_1,\cdots,x_{i-1},x_{i+1},\cdots,x_n)$.

It is the Dirichlet form of $(X_t^1,\cdots,X_t^n)_{t\ge 0}$ with $X^1,\dots,X^n$ independent copies of X.

n-Tensorization

On $\mathcal{X}^n$ define the sum-cost

- $\bigoplus c(x_1, \ldots, x_n; y_1, \ldots, y_n) := \sum_{1 \le i \le n} c(x_i; y_i), \quad x_1, \cdots, y_n \in \mathcal{X}$

and the sum-Dirichlet form

- $\bigoplus \mathcal{E}(g) := \int_{\mathcal{X}^n} \sum_{1 \le i \le n} \mathcal{E}(g_i^{\hat{x}_i}) \prod_{1 \le i \le n} \mu(dx_i)$

where $g_i^{\hat{x}_i}(x_i) = g(x_1, \ldots, x_n)$ with $\hat{x}_i = (x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_n)$.

It is the Dirichlet form of $(X_t^1, \cdots, X_t^n)_{t \ge 0}$ with $X^1, \ldots, X^n$ independent copies of X.

n-Tensorization

On $\mathcal{X}^n$ define the sum-cost

- $\bigoplus c(x_1,\dots,x_n;y_1,\dots,y_n) := \sum_{1\le i\le n} c(x_i;y_i), \quad x_1,\cdots,y_n \in \mathcal{X}$

and the sum-Dirichlet form

- $\bigoplus \mathcal{E}(g) := \int_{\mathcal{X}^n} \sum_{1\le i\le n} \mathcal{E}(g_i^{\hat{x}_i}) \prod_{1\le i\le n} \mu(dx_i)$

where $g_i^{\hat{x}_i}(x_i) = g(x_1,\dots,x_n)$ with $\hat{x}_i = (x_1,\cdots,x_{i-1},x_{i+1},\cdots,x_n)$.

It is the Dirichlet form of $(X_t^1,\cdots,X_t^n)_{t\ge 0}$ with $X^1,\dots,X^n$ independent copies of X.

n-Tensorization

Theorem

If μ satisfies T_cI, then $\mu^{\otimes n}$ satisfies

$$n\alpha\left(\frac{T_{\oplus c}(\nu,\mu^{\otimes n})}{n}\right) \le I_{\oplus \mathcal{E}}(\nu|\mu^{\otimes n}), \quad \forall \nu \in \mathcal{P}(\mathcal{X}^n)$$

In particular, for all $(u,v) \in \Phi_c$, $\beta \in \mathcal{P}(\mathcal{X}^n)$, $t, r > 0$,

$$\mathbb{P}_\beta\left(\frac{1}{n}\sum_{i=1}^n \frac{1}{t}\int_0^t u(X_s^i)\,ds \ge \mu(v) + r\right) \le \left\|\frac{d\beta}{d\mu^{\otimes n}}\right\|_{L^2(\mu^{\otimes n})} e^{-n t \alpha(r)}.$$

Proof.

$\alpha_n(r) = \Box^n \alpha(r) = n\alpha(r/n)$, $r \ge 0$. □

n-Tensorization

Theorem

If μ satisfies $T_c I$, then $\mu^{\otimes n}$ satisfies

$$n\alpha\left(\frac{T_{\oplus c}(\nu,\mu^{\otimes n})}{n}\right) \le I_{\oplus\mathcal{E}}(\nu|\mu^{\otimes n}), \quad \forall \nu \in \mathcal{P}(\mathcal{X}^n)$$

In particular, for all $(u,v) \in \Phi_c$, $\beta \in \mathcal{P}(\mathcal{X}^n)$, $t, r > 0$,

$$\mathbb{P}_\beta\left(\frac{1}{n}\sum_{i=1}^n \frac{1}{t}\int_0^t u(X_s^i)\,ds \ge \mu(v) + r\right) \le \left\| \frac{d\beta}{d\mu^{\otimes n}} \right\|_{L^2(\mu^{\otimes n})} e^{-nt\alpha(r)}.$$

Proof.

$\alpha_n(r) = \square^n \alpha(r) = n\alpha(r/n)$, $r \ge 0$. □

n-Tensorization

Theorem

If μ satisfies $T_c I$, then $\mu^{\otimes n}$ satisfies

$$n\alpha\left(\frac{T_{\oplus c}(\nu, \mu^{\otimes n})}{n}\right) \le I_{\oplus\mathcal{E}}(\nu|\mu^{\otimes n}), \quad \forall \nu \in \mathcal{P}(\mathcal{X}^n)$$

In particular, for all $(u, v) \in \Phi_c$, $\beta \in \mathcal{P}(\mathcal{X}^n)$, $t, r > 0$,

$$\mathbb{P}_\beta\left(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{t}\int_0^t u(X_s^i)\,ds \ge \mu(v) + r\right) \le \left\|\frac{d\beta}{d\mu^{\otimes n}}\right\|_{L^2(\mu^{\otimes n})} e^{-nt\alpha(r)}.$$

Proof.

$\alpha_n(r) = \Box^n\alpha(r) = n\alpha(r/n)$, $r \ge 0$. □

n-Tensorization

Example ($W_1 I$)

As $W_1 I(c)$: $W_1^2(\nu, \mu) \leq 4c^2 I(\nu|\mu)$ corresponds to $\alpha_1(r) = r^2/(4c^2)$ we have

$$W_1^2(\nu, \mu^{\otimes n}) \leq 4c^2 n I(\nu|\mu^{\otimes n})$$

Example ($W_2 I$)

As $W_2 I(c)$: $W_2(\nu, \mu) \leq 4c^2 I(\nu|\mu)$ corresponds to $\alpha_2(r) = r/(4c^2)$ we have

$$W_2(\nu, \mu^{\otimes n}) \leq 4c^2 I(\nu|\mu^{\otimes n})$$

- The constant $4c^2 n$ in $W_1 I$ explodes as n.
- The constant is *dimension-free* in $W_2 I$.
- Beautiful applications of $W_2 I$ for Gibbs measures in situations where log-Sobolev is unknown (Gao and Wu, 2007).

n-Tensorization

Example ($W_1 I$)

As $W_1 I(c)$: $W_1^2(\nu,\mu) \le 4c^2 I(\nu|\mu)$ corresponds to $\alpha_1(r) = r^2/(4c^2)$ we have

$$W_1^2(\nu,\mu^{\otimes n}) \le 4c^2 n I(\nu|\mu^{\otimes n})$$

Example ($W_2 I$)

As $W_2 I(c)$: $W_2(\nu,\mu) \le 4c^2 I(\nu|\mu)$ corresponds to $\alpha_2(r) = r/(4c^2)$ we have

$$W_2(\nu,\mu^{\otimes n}) \le 4c^2 I(\nu|\mu^{\otimes n})$$

- The constant $4c^2 n$ in $W_1 I$ explodes as n.
- The constant is *dimension-free* in $W_2 I$.
- Beautiful applications of $W_2 I$ for Gibbs measures in situations where log-Sobolev is unknown (Gao and Wu, 2007).

n-Tensorization

Example ($W_1 I$)

As $W_1 I(c)$: $W_1^2(\nu, \mu) \leq 4c^2 I(\nu|\mu)$ corresponds to $\alpha_1(r) = r^2/(4c^2)$ we have

$$W_1^2(\nu, \mu^{\otimes n}) \leq 4c^2 n I(\nu|\mu^{\otimes n})$$

Example ($W_2 I$)

As $W_2 I(c)$: $W_2(\nu, \mu) \leq 4c^2 I(\nu|\mu)$ corresponds to $\alpha_2(r) = r/(4c^2)$ we have

$$W_2(\nu, \mu^{\otimes n}) \leq 4c^2 I(\nu|\mu^{\otimes n})$$

- The constant $4c^2 n$ in $W_1 I$ explodes as n.
- The constant is *dimension-free* in $W_2 I$.
- Beautiful applications of $W_2 I$ for Gibbs measures in situations where log-Sobolev is unknown (Gao and Wu, 2007).

n-Tensorization

Example ($W_1 I$)

As $W_1 I(c)$: $W_1^2(\nu,\mu) \le 4c^2 I(\nu|\mu)$ corresponds to $\alpha_1(r) = r^2/(4c^2)$ we have

$$W_1^2(\nu,\mu^{\otimes n}) \le 4c^2 n I(\nu|\mu^{\otimes n})$$

Example ($W_2 I$)

As $W_2 I(c)$: $W_2(\nu,\mu) \le 4c^2 I(\nu|\mu)$ corresponds to $\alpha_2(r) = r/(4c^2)$ we have

$$W_2(\nu,\mu^{\otimes n}) \le 4c^2 I(\nu|\mu^{\otimes n})$$

- The constant $4c^2 n$ in $W_1 I$ explodes as n.
- The constant is *dimension-free* in $W_2 I$.
- Beautiful applications of $W_2 I$ for Gibbs measures in situations where log-Sobolev is unknown (Gao and Wu, 2007).

n-Tensorization

Example ($W_1 I$)

As $W_1 I(c)$: $W_1^2(\nu,\mu) \le 4c^2 I(\nu|\mu)$ corresponds to $\alpha_1(r) = r^2/(4c^2)$ we have

$$W_1^2(\nu,\mu^{\otimes n}) \le 4c^2 n I(\nu|\mu^{\otimes n})$$

Example ($W_2 I$)

As $W_2 I(c)$: $W_2(\nu,\mu) \le 4c^2 I(\nu|\mu)$ corresponds to $\alpha_2(r) = r/(4c^2)$ we have

$$W_2(\nu,\mu^{\otimes n}) \le 4c^2 I(\nu|\mu^{\otimes n})$$

- The constant $4c^2 n$ in $W_1 I$ explodes as n.
- The constant is *dimension-free* in $W_2 I$.
- Beautiful applications of $W_2 I$ for Gibbs measures in situations where log-Sobolev is unknown (Gao and Wu, 2007).

Outline

1. Setting And Definitions
   - The Markov Process
   - Optimal Transport
   - Inequalities
2. A Connection With Large deviations
   - Large Deviations Of The Occupation Measure
   - The Basic Result
   - Tensorization
3. Criteria
   - **Poincaré and log-Sobolev**
   - Poincaré implies Pinsker
   - Spectral gap in C^0_{Lip}
   - Lyapunov function condition

Poincaré and log-Sobolev

Definition (Poincaré inequality)

$$\mathrm{Var}_\mu\left(\sqrt{\frac{d\nu}{d\mu}}\right) \le cI(\nu|\mu), \quad \forall\nu \qquad P(c)$$

Definition (log-Sobolev inequality)

$$H(\nu|\mu) \le 2cI(\nu|\mu), \quad \forall\nu \qquad LS(c)$$

Theorem

$$c_P(\mu) \le c_{W_2I}(\mu) \le c_{LS}(\mu).$$

Proof.

Follow Otto and Villani (JFA, 2000). □

Poincaré and log-Sobolev

Definition (Poincaré inequality)

$$\mathrm{Var}_\mu\left(\sqrt{\frac{d\nu}{d\mu}}\right) \le cI(\nu|\mu), \quad \forall \nu \qquad P(c)$$

Definition (log-Sobolev inequality)

$$H(\nu|\mu) \le 2cI(\nu|\mu), \quad \forall \nu \qquad LS(c)$$

Theorem

$$c_P(\mu) \le c_{W_2 I}(\mu) \le c_{LS}(\mu).$$

Proof.

Follow Otto and Villani (JFA, 2000). □

Poincaré and log-Sobolev

Definition (Poincaré inequality)

$$\mathrm{Var}_\mu\left(\sqrt{\frac{d\nu}{d\mu}}\right) \le cI(\nu|\mu), \quad \forall \nu \qquad P(c)$$

Definition (log-Sobolev inequality)

$$H(\nu|\mu) \le 2cI(\nu|\mu), \quad \forall \nu \qquad LS(c)$$

Theorem

$$c_P(\mu) \le c_{W_2 I}(\mu) \le c_{LS}(\mu).$$

Proof.

Follow Otto and Villani (JFA, 2000). □

Poincaré and log-Sobolev

Definition (Poincaré inequality)

$$\mathrm{Var}_\mu\left(\sqrt{\frac{d\nu}{d\mu}}\right) \le cI(\nu|\mu), \quad \forall \nu \qquad P(c)$$

Definition (log-Sobolev inequality)

$$H(\nu|\mu) \le 2cI(\nu|\mu), \quad \forall \nu \qquad LS(c)$$

Theorem

$$c_P(\mu) \le c_{W_2 I}(\mu) \le c_{LS}(\mu).$$

Proof.

Follow Otto and Villani (JFA, 2000). □

Outline

Poincaré implies a Pinsker type inequality

Theorem

If μ satisfies $P(c)$, then

$$\|\nu - \mu\|^2_{\mathrm{TV}} \le 4cI(\nu|\mu), \ \forall \nu.$$

In particular, $\forall t, r, \beta, u,$

$$\mathbb{P}_\beta \left(\frac{1}{t} \int_0^t u(X_s)\, ds \ge \mu(u) + r \right) \le \left\| \frac{d\beta}{d\mu} \right\|_{L^2(\mu)} \exp\left(-\frac{tr^2}{c\delta(u)^2} \right)$$

where $\delta(u) := \sup(u) - \inf(u)$.

Idea of proof.

- $\|\nu - \mu\|_{\mathrm{TV}} = \|\nu - \mu\|^*_{\mathrm{Lip}}$ with $d(x, y) = 2_{x \neq y}$.
- $\|\nu - \mu\|^2_{\mathrm{TV}} \le 4\mathrm{Var}_\mu \left(\sqrt{\frac{d\nu}{d\mu}} \right) \le 4cI(\nu|\mu)$.

Poincaré implies a Pinsker type inequality

Theorem

If μ satisfies $P(c)$, then

$$\|\nu-\mu\|^2_{\mathrm{TV}} \le 4cI(\nu|\mu),\ \forall \nu.$$

In particular, $\forall t, r, \beta, u$,

$$\mathbb{P}_\beta\left(\frac{1}{t}\int_0^t u(X_s)\,ds \ge \mu(u)+r\right) \le \left\|\frac{d\beta}{d\mu}\right\|_{L^2(\mu)} \exp\left(-\frac{tr^2}{c\delta(u)^2}\right)$$

where $\delta(u) := \sup(u) - \inf(u)$.

Idea of proof.

- $\|\nu-\mu\|_{\mathrm{TV}} = \|\nu-\mu\|^*_{\mathrm{Lip}}$ with $d(x,y) = \mathbf{2}_{x\neq y}$.
- $\|\nu-\mu\|^2_{\mathrm{TV}} \le 4\mathrm{Var}_\mu\left(\sqrt{\frac{d\nu}{d\mu}}\right) \le 4cI(\nu|\mu)$.

Poincaré implies a Pinsker type inequality

Theorem

If μ satisfies $P(c)$, then

$$\|\nu-\mu\|^2_{\mathrm{TV}} \le 4cI(\nu|\mu),\ \forall \nu.$$

In particular, $\forall t, r, \beta, u,$

$$\mathbb{P}_\beta\left(\frac{1}{t}\int_0^t u(X_s)\,ds \ge \mu(u)+r\right) \le \left\|\frac{d\beta}{d\mu}\right\|_{L^2(\mu)} \exp\left(-\frac{tr^2}{c\delta(u)^2}\right)$$

where $\delta(u) := \sup(u) - \inf(u)$.

Idea of proof.

- $\|\nu-\mu\|_{\mathrm{TV}} = \|\nu-\mu\|^*_{\mathrm{Lip}}$ with $d(x,y) = \mathbf{2}_{x\neq y}$.
- $\|\nu-\mu\|^2_{\mathrm{TV}} \le 4\mathrm{Var}_\mu\left(\sqrt{\frac{d\nu}{d\mu}}\right) \le 4cI(\nu|\mu)$.

□

Poincaré implies a Pinsker type inequality

Theorem

If μ satisfies $P(c)$, then

$$\|\nu-\mu\|^2_{\mathrm{TV}} \le 4cI(\nu|\mu), \ \forall \nu.$$

In particular, $\forall t, r, \beta, u$,

$$\mathbb{P}_\beta\left(\frac{1}{t}\int_0^t u(X_s)\,ds \ge \mu(u)+r\right) \le \left\|\frac{d\beta}{d\mu}\right\|_{L^2(\mu)} \exp\left(-\frac{tr^2}{c\delta(u)^2}\right)$$

where $\delta(u) := \sup(u) - \inf(u)$.

Idea of proof.

- $\|\nu-\mu\|_{\mathrm{TV}} = \|\nu-\mu\|^*_{\mathrm{Lip}}$ with $d(x,y) = \mathbf{2}_{x\neq y}$.
- $\|\nu-\mu\|^2_{\mathrm{TV}} \le 4\mathrm{Var}_\mu\left(\sqrt{\frac{d\nu}{d\mu}}\right) \le 4cI(\nu|\mu)$.

□

Poincaré implies a Pinsker type inequality

Theorem

$$\|\nu-\mu\|^2_{\mathrm{TV}} \le 4\mathrm{Var}_\mu\left(\sqrt{\frac{d\nu}{d\mu}}\right), \quad \forall \mu,\nu.$$

Idea of proof.

This is W_1I with

- $d(x,y) = \mathbf{2}_{x\neq y}$ so that: $\quad W_1(\nu,\mu) = \|\nu-\mu\|_{\mathrm{TV}}$;
- the process X defined as follows. Take
  - $(Y_n)_{n\ge 0}$ independent with common law μ;
  - $(N_t)_{t\ge 0}$ a Poisson(1) process;
  - $(Y_n)_{n\ge 0}$ and N independent;
  - $X_t = Y_{N_t}$, $t \ge 0$

so that: $\quad I(\nu|\mu) = \mathrm{Var}_\mu\left(\sqrt{\frac{d\nu}{d\mu}}\right), \ \forall \nu.$

Poincaré implies a Pinsker type inequality

Theorem

$$\|\nu-\mu\|^2_{\mathrm{TV}} \le 4\mathrm{Var}_\mu\left(\sqrt{\frac{d\nu}{d\mu}}\right), \quad \forall \mu,\nu.$$

Idea of proof.

This is $W_1 I$ with

- $d(x,y)=\mathbf{2}_{x\neq y}$ so that: $\quad W_1(\nu,\mu)=\|\nu-\mu\|_{\mathrm{TV}}$;
- the process X defined as follows. Take
  - $(Y_n)_{n\ge 0}$ independent with common law μ;
  - $(N_t)_{t\ge 0}$ a Poisson(1) process;
  - $(Y_n)_{n\ge 0}$ and N independent;
  - $X_t := Y_{N_t},\ t\ge 0$

so that: $\quad I(\nu|\mu)=\mathrm{Var}_\mu\left(\sqrt{\frac{d\nu}{d\mu}}\right),\ \forall \nu.$

Poincaré implies a Pinsker type inequality

Theorem

$$\|\nu-\mu\|^2_{\mathrm{TV}} \le 4\mathrm{Var}_\mu\left(\sqrt{\frac{d\nu}{d\mu}}\right), \quad \forall \mu,\nu.$$

Idea of proof.

This is $W_1 I$ with

- $d(x,y) = \mathbf{2}_{x\neq y}$ so that: $W_1(\nu,\mu) = \|\nu-\mu\|_{\mathrm{TV}}$;
- the process X defined as follows. Take
  - $(Y_n)_{n\ge 0}$ independent with common law μ;
  - $(N_t)_{t\ge 0}$ a Poisson(1) process;
  - $(Y_n)_{n\ge 0}$ and N independent;
  - $X_t := Y_{N_t},\ t\ge 0$

so that: $I(\nu|\mu) = \mathrm{Var}_\mu\left(\sqrt{\frac{d\nu}{d\mu}}\right),\ \forall\nu.$

Poincaré implies a Pinsker type inequality

Theorem

$$\|\nu-\mu\|^2_{\mathrm{TV}} \le 4\mathrm{Var}_\mu\left(\sqrt{\frac{d\nu}{d\mu}}\right), \quad \forall \mu,\nu.$$

Idea of proof.

This is $W_1 I$ with

- $d(x,y)=\mathbf{2}_{x\neq y}$ so that: $W_1(\nu,\mu)=\|\nu-\mu\|_{\mathrm{TV}}$;
- the process X defined as follows. Take
  - $(Y_n)_{n\ge 0}$ independent with common law μ;
  - $(N_t)_{t\ge 0}$ a Poisson(1) process;
  - $(Y_n)_{n\ge 0}$ and N independent;
  - $X_t := Y_{N_t}$, $t\ge 0$

so that: $I(\nu|\mu)=\mathrm{Var}_\mu\left(\sqrt{\frac{d\nu}{d\mu}}\right)$, $\forall \nu$.

Poincaré implies a Pinsker type inequality

Theorem

$$\|\nu - \mu\|^2_{\mathrm{TV}} \le 4\mathrm{Var}_\mu\left(\sqrt{\frac{d\nu}{d\mu}}\right), \quad \forall \mu, \nu.$$

Idea of proof.

This is $W_1 I$ with

- $d(x, y) = \mathbf{2}_{x \neq y}$ so that: $W_1(\nu, \mu) = \|\nu - \mu\|_{\mathrm{TV}}$;
- the process X defined as follows. Take
  - $(Y_n)_{n \ge 0}$ independent with common law μ;
  - $(N_t)_{t \ge 0}$ a Poisson(1) process;
  - $(Y_n)_{n \ge 0}$ and N independent;
  - $X_t := Y_{N_t}$, $t \ge 0$

so that: $I(\nu|\mu) = \mathrm{Var}_\mu\left(\sqrt{\frac{d\nu}{d\mu}}\right)$, $\forall \nu$.

Poincaré implies a Pinsker type inequality

Theorem

$$\|\nu-\mu\|^2_{\mathrm{TV}} \le 4\mathrm{Var}_\mu\left(\sqrt{\frac{d\nu}{d\mu}}\right), \quad \forall \mu,\nu.$$

Idea of proof.

This is $W_1 I$ with

- $d(x,y) = \mathbf{2}_{x\neq y}$ so that: $W_1(\nu,\mu) = \|\nu-\mu\|_{\mathrm{TV}}$;
- the process X defined as follows. Take
  - $(Y_n)_{n\ge 0}$ independent with common law μ;
  - $(N_t)_{t\ge 0}$ a Poisson(1) process;
  - $(Y_n)_{n\ge 0}$ and N independent;
  - $X_t := Y_{N_t},\ t\ge 0$

so that: $I(\nu|\mu) = \mathrm{Var}_\mu\left(\sqrt{\frac{d\nu}{d\mu}}\right), \ \forall \nu.$

Poincaré implies a Pinsker type inequality

Theorem

$$\|\nu - \mu\|_{\mathrm{TV}}^2 \le 4\mathrm{Var}_\mu\left(\sqrt{\frac{d\nu}{d\mu}}\right), \quad \forall \mu, \nu.$$

Idea of proof.

This is $W_1 I$ with

- $d(x, y) = \mathbf{2}_{x \neq y}$ so that: $\quad W_1(\nu, \mu) = \|\nu - \mu\|_{\mathrm{TV}}$;
- the process X defined as follows. Take
  - $(Y_n)_{n \ge 0}$ independent with common law μ;
  - $(N_t)_{t \ge 0}$ a Poisson(1) process;
  - $(Y_n)_{n \ge 0}$ and N independent;
  - $X_t := Y_{N_t}, \ t \ge 0$

so that: $\quad I(\nu|\mu) = \mathrm{Var}_\mu\left(\sqrt{\frac{d\nu}{d\mu}}\right), \ \forall \nu.$

Outline

Spectral gap in C^0_{Lip}

Definition (Spectral gap in C^0_{Lip})

There exists $c < \infty$ such that for all $g \in C_{\mathrm{Lip}} \cap L^2_0(\mu)$, there exists $f \in L^2_0(\mu)$ such that

- $-\mathcal{L}f = g$;
- $\|\tilde{f}\|_{\mathrm{Lip}} \leq c\|g\|_{\mathrm{Lip}}, \quad \tilde{f}$: version of f.

- $\mathrm{gap}(-\mathcal{L}) = 1/c$.

Theorem

With $\mathcal{L} = \Delta - \nabla V \cdot \nabla$, the spectral gap in C^0_{Lip} implies $W_1 I(c)$.

- Remark: $\mu = e^{-V}\,dx$.

Spectral gap in C^0_{Lip}

Definition (Spectral gap in C^0_{Lip})

There exists $c < \infty$ such that for all $g \in C_{\mathrm{Lip}} \cap L^2_0(\mu)$, there exists $f \in L^2_0(\mu)$ such that

- $-\mathcal{L}f = g$;
- $\|\tilde{f}\|_{\mathrm{Lip}} \le c\|g\|_{\mathrm{Lip}}$, $\quad \tilde{f}$: version of f.

- $\mathrm{gap}(-\mathcal{L}) = 1/c$.

Theorem

With $\mathcal{L} = \Delta - \nabla V \cdot \nabla$, the spectral gap in C^0_{Lip} implies $W_1 I(c)$.

- Remark: $\mu = e^{-V}\,dx$.

Spectral gap in C^0_{Lip}

Definition (Spectral gap in C^0_{Lip})

There exists $c < \infty$ such that for all $g \in C_{\mathrm{Lip}} \cap L^2_0(\mu)$, there exists $f \in L^2_0(\mu)$ such that

- $-\mathcal{L}f = g$;
- $\|\tilde{f}\|_{\mathrm{Lip}} \le c\|g\|_{\mathrm{Lip}}, \quad \tilde{f}$: version of f.

- $\mathrm{gap}(-\mathcal{L}) = 1/c$.

Theorem

With $\mathcal{L} = \Delta - \nabla V \cdot \nabla$, the spectral gap in C^0_{Lip} implies $W_1 I(c)$.

- Remark: $\mu = e^{-V}\,dx$.

Spectral gap in C^0_{Lip}

Definition (Spectral gap in C^0_{Lip})

There exists $c < \infty$ such that for all $g \in C_{\mathrm{Lip}} \cap L^2_0(\mu)$, there exists $f \in L^2_0(\mu)$ such that

- $-\mathcal{L}f = g$;
- $\|\tilde{f}\|_{\mathrm{Lip}} \le c\|g\|_{\mathrm{Lip}}, \quad \tilde{f}$: version of f.

- $\mathrm{gap}(-\mathcal{L}) = 1/c$.

Theorem

With $\mathcal{L} = \Delta - \nabla V \cdot \nabla$, the spectral gap in C^0_{Lip} implies $W_1 I(c)$.

- Remark: $\mu = e^{-V}\, dx$.

Spectral gap in C^0_{Lip}

Definition (Spectral gap in C^0_{Lip})

There exists $c < \infty$ such that for all $g \in C_{\mathrm{Lip}} \cap L^2_0(\mu)$, there exists $f \in L^2_0(\mu)$ such that

- $-\mathcal{L}f = g$;
- $\|\tilde{f}\|_{\mathrm{Lip}} \le c\|g\|_{\mathrm{Lip}}, \quad \tilde{f}$: version of f.

- $\mathrm{gap}(-\mathcal{L}) = 1/c$.

Theorem

With $\mathcal{L} = \Delta - \nabla V \cdot \nabla$, the spectral gap in C^0_{Lip} implies $W_1 I(c)$.

- Remark: $\mu = e^{-V}\,dx$.

Spectral gap in C^0_{Lip}

Application

With $\mathcal{L} = \Delta - \nabla V \cdot \nabla$, we have: $\quad \mathrm{Ric} + \nabla^2 V \geq K > 0 \Rightarrow W_1 I(K^{-1})$.

Remark: By Bakry-Emery criterion log-Sob holds.
Proof without log-Sob.

Application

With $dX_t = \sqrt{2}\sigma(X_t)\,dB_t + b(X_t)\,dt$ and

$$\begin{cases} \mathrm{trace}[(\sigma(y)-\sigma(x))(\sigma(y)-\sigma(x))^T] + \langle y-x, b(y)-b(x)\rangle \leq -\delta|y-x|^2, \forall x, y \\ (P_t) \text{ is symmetric in } L^2(\mu) \end{cases}$$

then $W_1 I(\|\sigma\|_\infty/\delta)$ holds true.

- unique invariant measure μ, unknown to be estimated;
- log-Sob unknown.

Spectral gap in C^0_{Lip}

Application

With $\mathcal{L} = \Delta - \nabla V \cdot \nabla$, we have: $\mathrm{Ric} + \nabla^2 V \geq K > 0 \Rightarrow W_1 I(K^{-1})$.

Remark: By Bakry-Emery criterion log-Sob holds.
Proof without log-Sob.

Application

With $dX_t = \sqrt{2}\sigma(X_t)\, dB_t + b(X_t)\, dt$ and

$$\begin{cases} \mathrm{trace}[(\sigma(y) - \sigma(x))(\sigma(y) - \sigma(x))^T] + \langle y - x, b(y) - b(x) \rangle \leq -\delta |y - x|^2, \forall x, y \\ (P_t) \text{ is symmetric in } L^2(\mu) \end{cases}$$

then $W_1 I(\|\sigma\|_\infty / \delta)$ holds true.

- unique invariant measure μ, unknown to be estimated;
- log-Sob unknown.

Spectral gap in C^0_{Lip}

Application

With $\mathcal{L} = \Delta - \nabla V \cdot \nabla$, we have: $\mathrm{Ric} + \nabla^2 V \geq K > 0 \Rightarrow W_1 I(K^{-1})$.

Remark: By Bakry-Emery criterion log-Sob holds.
Proof without log-Sob.

Application

With $dX_t = \sqrt{2}\sigma(X_t)\, dB_t + b(X_t)\, dt$ and

$$\begin{cases} \mathrm{trace}[(\sigma(y) - \sigma(x))(\sigma(y) - \sigma(x))^T] + \langle y - x, b(y) - b(x) \rangle \leq -\delta |y - x|^2, \forall x, y \\ (P_t) \text{ is symmetric in } L^2(\mu) \end{cases}$$

then $W_1 I(\|\sigma\|_\infty / \delta)$ holds true.

- unique invariant measure μ, unknown to be estimated;
- log-Sob unknown.

Spectral gap in C^0_{Lip}

Application

With $\mathcal{L} = \Delta - \nabla V \cdot \nabla$, we have: $\mathrm{Ric} + \nabla^2 V \geq K > 0 \Rightarrow W_1 I(K^{-1})$.

Remark: By Bakry-Emery criterion log-Sob holds.
Proof without log-Sob.

Application

With $dX_t = \sqrt{2}\sigma(X_t)\, dB_t + b(X_t)\, dt$ and

$$\begin{cases} \mathrm{trace}[(\sigma(y) - \sigma(x))(\sigma(y) - \sigma(x))^T] + \langle y - x, b(y) - b(x)\rangle \leq -\delta|y - x|^2, \forall x, y \\ (P_t) \text{ is symmetric in } L^2(\mu) \end{cases}$$

then $W_1 I(\|\sigma\|_\infty/\delta)$ holds true.

- unique invariant measure μ, unknown to be estimated;
- log-Sob unknown.

Spectral gap in C^0_{Lip}

Application

With $\mathcal{L} = \Delta - \nabla V \cdot \nabla$, we have: $\mathrm{Ric} + \nabla^2 V \geq K > 0 \Rightarrow W_1 I(K^{-1})$.

Remark: By Bakry-Emery criterion log-Sob holds.
Proof without log-Sob.

Application

With $dX_t = \sqrt{2}\sigma(X_t)\, dB_t + b(X_t)\, dt$ and

$$\begin{cases} \mathrm{trace}[(\sigma(y) - \sigma(x))(\sigma(y) - \sigma(x))^T] + \langle y - x, b(y) - b(x)\rangle \leq -\delta |y - x|^2, \forall x, y \\ (P_t) \text{ is symmetric in } L^2(\mu) \end{cases}$$

then $W_1 I(\|\sigma\|_\infty / \delta)$ holds true.

- unique invariant measure μ, unknown to be estimated;
- log-Sob unknown.

Outline

Lyapunov function condition

Definition (Lyapunov function condition)

With $U : \mathcal{X} \to [1, \infty)$ continuous; $\phi : \mathcal{X} \to [0, \infty)$; $b > 0$,

$$-\frac{\mathcal{L}U}{U} \geq \phi - b, \quad \mu\text{-a.e.}$$

Theorem

Under this Lyapunov function condition, if μ satisfies Poincaré inequality, then

$$\|\phi.(\nu - \mu)\|_{\mathrm{TV}} \leq C \left(\sqrt{I(\nu|\mu)} + I(\nu|\mu) \right), \quad \forall \nu.$$

- $d(x, y) = [\phi(x) + \phi(y)]\mathbf{1}_{x \neq y}$;
- $\alpha(r) = O_{r \simeq 0}(r^2)$ and $\alpha(r) = O_{r \simeq +\infty}(r)$

Lyapunov function condition

Definition (Lyapunov function condition)

With $U : \mathcal{X} \to [1, \infty)$ continuous; $\phi : \mathcal{X} \to [0, \infty)$; $b > 0$,

$$-\frac{\mathcal{L}U}{U} \geq \phi - b, \quad \mu\text{-a.e.}$$

Theorem

Under this Lyapunov function condition, if μ satisfies Poincaré inequality, then

$$\|\phi.(\nu - \mu)\|_{\mathrm{TV}} \leq C\left(\sqrt{I(\nu|\mu)} + I(\nu|\mu)\right), \quad \forall \nu.$$

- $d(x, y) = [\phi(x) + \phi(y)]\mathbf{1}_{x \neq y}$;
- $\alpha(r) = O_{r \simeq 0}(r^2)$ and $\alpha(r) = O_{r \simeq +\infty}(r)$

Lyapunov function condition

Definition (Lyapunov function condition)

With $U : \mathcal{X} \to [1, \infty)$ continuous; $\phi : \mathcal{X} \to [0, \infty)$; $b > 0$,

$$-\frac{\mathcal{L}U}{U} \geq \phi - b, \quad \mu\text{-a.e.}$$

Theorem

Under this Lyapunov function condition, if μ satisfies Poincaré inequality, then

$$\|\phi.(\nu - \mu)\|_{\mathrm{TV}} \leq C\left(\sqrt{I(\nu|\mu)} + I(\nu|\mu)\right), \quad \forall \nu.$$

- $d(x, y) = [\phi(x) + \phi(y)]\mathbf{1}_{x \neq y}$;
- $\alpha(r) = O_{r \simeq 0}(r^2)$ and $\alpha(r) = O_{r \simeq +\infty}(r)$

Lyapunov function condition

Definition (Lyapunov function condition)

With $U : \mathcal{X} \to [1, \infty)$ continuous; $\phi : \mathcal{X} \to [0, \infty)$; $b > 0$,

$$-\frac{\mathcal{L}U}{U} \geq \phi - b, \quad \mu\text{-a.e.}$$

Theorem

Under this Lyapunov function condition, if μ satisfies Poincaré inequality, then

$$\|\phi.(\nu - \mu)\|_{\mathrm{TV}} \leq C \left(\sqrt{I(\nu|\mu)} + I(\nu|\mu) \right), \quad \forall \nu.$$

- $d(x, y) = [\phi(x) + \phi(y)]\mathbf{1}_{x \neq y}$;
- $\alpha(r) = O_{r \simeq 0}(r^2)$ and $\alpha(r) = O_{r \simeq +\infty}(r)$

Lyapunov function condition

Definition (Lyapunov function condition)

With $U : \mathcal{X} \to [1, \infty)$ continuous; $\phi : \mathcal{X} \to [0, \infty)$; $b > 0$,

$$-\frac{\mathcal{L}U}{U} \geq \phi - b, \quad \mu\text{-a.e.}$$

Theorem

Under this Lyapunov function condition, if μ satisfies Poincaré inequality, then

$$\|\phi.(\nu - \mu)\|_{\mathrm{TV}} \leq C \left(\sqrt{I(\nu|\mu)} + I(\nu|\mu) \right), \quad \forall \nu.$$

- $d(x, y) = [\phi(x) + \phi(y)]\mathbf{1}_{x \neq y}$;
- $\alpha(r) = O_{r \simeq 0}(r^2)$ and $\alpha(r) = O_{r \simeq +\infty}(r)$

Lyapunov function condition

Corollary

With $\mathcal{L} = \Delta - \nabla V \cdot \nabla : \mu = e^{-V}\,dx$. *If*

- $|x - x_o|^2 \le c(1 + |\nabla V|^2(x)), \forall x;$
- $\limsup_{|x|\to\infty} \Delta V(x)/|\nabla V|^2(x) < 1,$

then $W_1 I(1/(4c))$ *holds and for all* $t, r > 0$

$$\mathbb{P}_\beta\left(\frac{1}{t}\int_0^t u(X_s)\,ds \ge \mu(u) + r\right) \le \left\|\frac{d\beta}{d\mu}\right\|_{L^2(\mu)} e^{-tr^2/c}.$$