Modified Log-Sobolev inequalities and transportation cost inequalities in Euclidean space

JINGHAI SHAO

Beijing Normal University

July 14, 2007

Consider $\mathrm{d}\mu = \frac{1}{Z} e^{-V} \mathrm{d}x$ on $\mathbb{R}^n$ and give out a sufficient condition on V such that μ satisfies modified logarithmic Sobolev inequality and transportation cost inequality.

As an example, we establish one kind of modified Log-Sobolev inequality and transportation cost inequality for $\mathrm{d}\mu_p = \frac{1}{Z_p} e^{-|x|^p} \mathrm{d}x$ with $p > 1$, and give out the explicit estimates of their constants.

- First, review some known results in this aspect;
- then introduce our main results;
- at last, introduce the idea of proof and key points in the proof.

Consider $\mathrm{d}\mu = \frac{1}{Z}e^{-V}\mathrm{d}x$ on $\mathbb{R}^n$ and give out a sufficient condition on V such that μ satisfies modified logarithmic Sobolev inequality and transportation cost inequality.

As an example, we establish one kind of modified Log-Sobolev inequality and transportation cost inequality for $\mathrm{d}\mu_p = \frac{1}{Z_p}e^{-|x|^p}\mathrm{d}x$ with $p > 1$, and give out the explicit estimates of their constants.

- First, review some known results in this aspect;
- then introduce our main results;
- at last, introduce the idea of proof and key points in the proof.

Given a probability measure μ on $\mathbb{R}^n$, μ is said to satisfy Log-Sobolev inequality if $\exists\, c > 0$ such that

$$\mathrm{Ent}_\mu(f^2) \leq c \int |\nabla f|^2 \, \mathrm{d}\mu. \qquad (1)$$

for $f \in C^1(\mathbb{R}^n)$, where $\mathrm{Ent}_\mu(f^2) = \int f^2 \log f^2 \, \mathrm{d}\mu - \int f^2 \mathrm{d}\mu \log \int f^2 \mathrm{d}\mu$. Now consider $\mu_p(\mathrm{d}x) = \dfrac{1}{Z_p} e^{-|x|^p} \mathrm{d}x$ for $p \geq 1$ on $\mathbb{R}^n$. Z_p is normalizing constant.

μ_p satisfies Log-Sobolev inequality (1) if and only if $p \geq 2$.

How about μ_p for $1 < p < 2$?

- In this case, μ_p satisfies Poincaré inequality (PI.)

$$\mathrm{Var}_{\mu_p}(f) \le C \int |\nabla f|^2 \,\mathrm{d}\mu_p.$$

Latała -Oleszkiewicz(2000):

$$\int f^2 \,\mathrm{d}\mu_p - \Big(\int |f|^\alpha \,\mathrm{d}\mu_p \Big)^{2/\alpha} \le C(2-\alpha)^\beta \int |\nabla f|^2 \,\mathrm{d}\mu_p, \quad \forall\, \alpha \in [1,2),$$

where $\beta = 2(p-1)/p$. This is an interpolation between Poincaré inequality and Log-Sobolev inequality. $\beta = 0$ is PI. and $\beta = 1$ is Log-Sobolev inequality.

Wang(2005) explored the relationship between previous inequality with F-Sobolev and super-Poincaré inequalities.

- In $\mathbb{R}$, I. Gentil, Guillin and Miclo(2005,2007) proved that for $1 < p < 2$,

$$\mathrm{Ent}_{\mu_p}(f^2) \leq C \int_{\mathbb{R}} H_p\Big(\frac{f'}{f}\Big) f^2 \,\mathrm{d}\mu_p, \tag{2}$$

where C is positive constant, and

$$H_p(x) = \max(|x|^2, |x|^q), \quad q = p/(p-1).$$

This result can be extended to more general log-concave measure on $\mathbb{R}$ between the exponential and Gaussian measure and to product measure on $\mathbb{R}^n$ using the product property of entropy.
Barthe and Roberto(2007) gave out a sufficient condition for modified Log-Sobolev inequality to hold. All these results rely on Hardy inequality on the line.

Wang(2007): M a connected complete Reimannian manifold, μ probability measure. $\log^\delta$-Sobolev inequality with power $\delta > 0$

$$\mu(f^2 \log^\delta(1+f^2)) \le C_1\mu(|\nabla f|^2) + C_2, \quad \mu(f^2) = 1 \tag{3}$$

implies modified log-Sobolev inequality

$$\mu(f^2 \log f^2) \le C\mu(|\nabla f|^2 \vee \left(\frac{|\nabla f|}{f}\right)^{2/\delta} f^2), \quad \mu(f^2) = 1. \qquad (4).$$

Therefore, all known sufficient conditions for $\log^\delta$-Sobolev inequality are also sufficient for (4). Especially, on $\mathbb{R}$, Wang gave out a sufficient condition for (4).

From modified log-Sobolev, under the help of Hamilton-Jacobi semigroup, transportation cost inequality can be deduced.

$$W_{\rho_\delta}(f\mu,\mu) \le C\mu(f\log f), \quad f \ge 0,\ \mu(f) = 1,$$

for some constant $C > 0$, where ρ is Riemannian distance, $\rho_\delta = \rho \wedge \rho^{1/(2-\delta)}$,

$$W_{\rho_\delta}(f\mu,\mu) = \inf_{\pi \in \mathscr{C}(f\mu,\mu)} \int_{M\times M} \rho_\delta(x,y)^2 \pi(\mathrm{d}x,\mathrm{d}y).$$

Here $\mathscr{C}(f\mu,\mu)$ denotes the set of all couplings of $f\mu$ and μ.

Main results

Consider the absolutely continuous measure $e^{-V(x)}\mathrm{d}x$ on $\mathbb{R}^n$.

Theorem

Let V be a C^1 function on $\mathbb{R}^n$. Assume that there exists some continuous function $\Psi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ and constant $c > 0$, such that

$$V(z+h) - V(z) - \langle \nabla V(z), h\rangle \geq c\Psi(z,h), \quad z,\ h \in \mathbb{R}^n. \tag{5}$$

Set $\mu(\mathrm{d}x) = e^{-V}\mathrm{d}x$. Then it holds

$$\mathrm{Ent}_\mu(f^2) \leq c\int_{\mathbb{R}^n} \Psi_*\Big(\frac{2\nabla f}{cf}, z\Big) f^2\,\mathrm{d}\mu(z), \quad \text{for positive } f \in C^1(\mathbb{R}^n), \tag{6}$$

where $\Psi_*(x,z) := \sup_{h\in\mathbb{R}^n}\{\langle x,h\rangle - \Psi(z,h)\}$ for $x,\ z \in \mathbb{R}^n$.

Remark:

- If $\mathrm{Hess}_x V(\xi,\xi) \geq c|\xi|^2$ for some constant $c > 0$, then condition (5) is satisfied by taking $\Psi(z,h) = |h|^2/2$. Hence, $\Psi_*(x,z) = |x|^2/2$. (6) turns out to be
$$\mathrm{Ent}_\mu(f^2) \leq \frac{1}{C} \int |\nabla f|^2 \,\mathrm{d}\mu.$$
This is just the Bakry-Emery criterion.
- If $\mathrm{Hess}_x V$ is not uniformly bounded below, our result can also help to deduce some kind of modified log-Sobolev inequality.

As an application, consider $\mu_p(\mathrm{d}x) = \frac{1}{Z_p} e^{-|x|^p} \mathrm{d}x$ on $\mathbb{R}^n$. We proved that

$$\text{for } p \geq 2, \quad |z+h|^p - |z|^p - p|z|^{p-2}\langle z, h\rangle \geq c|h|^p,$$
$$\text{for } 1 < p < 2, \quad |z+h|^p - |z|^p - p|z|^{p-2}\langle z, h\rangle \geq \tilde{c}\min\{|z|^{p-2}|h|^2, |h|^p\},$$

then it follows that

Theorem

For all positive C^1 function f, it holds that

$$p \geq 2, \mathrm{Ent}_{\mu_p}(f^2) \leq \frac{2^q}{q(cp)^{q-1}} \int_{\mathbb{R}^n} \Big(\frac{|\nabla f|}{f}\Big)^q f^2 \,\mathrm{d}\mu_p,$$
$$1 < p < 2, \mathrm{Ent}(f^2) \leq \int \max\Big\{\frac{|x|^{2-p}}{4\tilde{c}}\Big(\frac{|\nabla f|}{f}\Big)^2, \frac{2^q}{q(\tilde{c}p)^{q-1}}\Big(\frac{|\nabla f|}{f}\Big)^q\Big\} f^2 \,\mathrm{d}\mu_p(x$$

where $q = \frac{p}{p-1}$, $c = \min\left(\frac{p}{2^{p-1}}, \frac{(p-1)^{\frac{p}{p-2}} + p(p-1)^{\frac{1}{p-2}} + 1}{\left((p-1)^{\frac{1}{p-2}}+1\right)^p} \mathbf{1}_{p>2} + \mathbf{1}_{p=2}\right)$ and $\tilde{c} = 2^p - 1 - p$.

As an application, consider $\mu_p(\mathrm{d}x) = \frac{1}{Z_p} e^{-|x|^p} \mathrm{d}x$ on $\mathbb{R}^n$. We proved that

$$\text{for } p \geq 2, \quad |z+h|^p - |z|^p - p|z|^{p-2}\langle z, h\rangle \geq c|h|^p,$$

$$\text{for } 1 < p < 2, \quad |z+h|^p - |z|^p - p|z|^{p-2}\langle z, h\rangle \geq \tilde{c} \min\{|z|^{p-2}|h|^2, |h|^p\},$$

then it follows that

Theorem

For all positive C^1 function f, it holds that

$$p \geq 2, \mathrm{Ent}_{\mu_p}(f^2) \leq \frac{2^q}{q(cp)^{q-1}} \int_{\mathbb{R}^n} \left(\frac{|\nabla f|}{f}\right)^q f^2 \, \mathrm{d}\mu_p,$$

$$1 < p < 2, \mathrm{Ent}(f^2) \leq \int \max\left\{\frac{|x|^{2-p}}{4\tilde{c}}\left(\frac{|\nabla f|}{f}\right)^2, \frac{2^q}{q(\tilde{c}p)^{q-1}}\left(\frac{|\nabla f|}{f}\right)^q\right\} f^2 \, \mathrm{d}\mu_p(x$$

where $q = \frac{p}{p-1}$, $c = \min\left(\frac{p}{2^{p-1}}, \frac{(p-1)^{\frac{p}{p-2}} + p(p-1)^{\frac{1}{p-2}} + 1}{\left((p-1)^{\frac{1}{p-2}} + 1\right)^p} \mathbf{1}_{p>2} + \mathbf{1}_{p=2}\right)$ *and* $\tilde{c} = 2^p - 1 - p$.

Transportation cost inequality

Talagrand(1996) established

$$W_2^2(f\gamma,\gamma) \le 2\int_{\mathbb{R}^n} |\nabla f|^2 \,\mathrm{d}\gamma, \quad f \ge 0,\ \int_{\mathbb{R}^n} f = 1,$$

where $\gamma(\mathrm{d}x)$ is Gaussian measure on $\mathbb{R}^n$, and

$$W_2^2(f\gamma,\gamma) = \inf_{\pi\in\mathscr{C}(f\gamma,\gamma)} \Big(\int_{\mathbb{R}^n\times\mathbb{R}^n} |x-y|^2 \,\mathrm{d}\pi(x,y)\Big).$$

Recall that $\mathscr{C}(\mu,\nu)$ denotes the set of all couplings of μ and ν. In general, given a lower semi-continuous function $C(\cdot,\cdot) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$, define

$$W_c(\mu,\nu) = \inf_{\pi\in\mathscr{C}(\mu,\nu)} \Big(\int_{\mathbb{R}^n\times\mathbb{R}^n} C(x,y)\,\mathrm{d}\pi(x,y)\Big).$$

Theorem

Let $\mu(\mathrm{d}x) = e^{-V(x)}\mathrm{d}x$, $V \in C^1(\mathbb{R}^n)$. *Assume*

$$V(x) - V(y) - \langle \nabla V(y), x - y \rangle \geq \Psi(x, y), \quad x, y \in \mathbb{R}^n,$$

where $\Psi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$ *is lower semi-continuous function. Then for any* $f \geq 0,\ \mu(f) = 1,$

$$W_\Psi(\mu, f\mu) \leq \mathrm{Ent}_\mu(f). \tag{7}$$

N. Gozlan(2006) gave a sufficient condition for a transportation cost inequality to hold on $\mathbb{R}$ for the measure $\mathrm{d}\mu = e^{-V(x)}\mathrm{d}x$, where V is "good potential" .

Indeed, V is a good potential if $V \in C^2(\mathbb{R})$ and

- $\exists\ x_0 > 0$ such that $V' > 0$ on $(-\infty, -x_0] \cup [x_0, \infty)$,
- $\frac{V''(x)}{V'^2(x)} \to 0$ as $x \to \pm\infty$.

Then μ satisfies the transportation cost inequality with the cost function $C(x, y) = \alpha(|x - y|)$ where α is a even "good potential" and satisfies some conditions. One of them is

$$\exists\ \lambda > 0, \text{ such that } \lim_{x \to \pm\infty} \frac{\alpha'(\lambda x)}{V'(x + m)} < +\infty.$$

One application of our theorem is that when $\mathrm{Hess} V$ is not uniformly bounded below, we get

Proposition

If $V \in C^2(\mathbb{R}^n)$ *and* $\exists$ *measurable function* $k(x)$ *such that* $\mathrm{Hess}_x V(\xi, \xi) \geq k(x)|\xi|^2$, *then*

$$\inf_{\pi \in \mathscr{C}(\mu, f\mu)} \left(\int K(x,y)|x-y|^2 \, \pi(\mathrm{d}x, \mathrm{d}y) \right) \leq \mathrm{Ent}_\mu(f),$$

where $K(x,y) = \int_0^1 s k(sx + (1-s)y) \, \mathrm{d}s$.

Another application is to deduce the transportation cost inequality for μ_p on $\mathbb{R}^n$.

Proposition

When $p \geq 2$, *for each* $f \geq 0$, $\int f\,\mu_p = 1$,

$$W_p^p(f\mu_p, \mu_p) := \inf_{\pi \in \mathscr{C}(f\mu_p, \mu_p)} \left(\int_{\mathbb{R}^n \times \mathbb{R}^n} |x-y|^p \,\mathrm{d}\pi(x,y) \right) \leq \frac{1}{c} \mathrm{Ent}_{\mu_p}(f),$$

When $1 < p < 2$, *for each* $f \geq 0$, $\int f\,\mu_p = 1$,

$$\begin{aligned} &W_p^p(f\mu_p, \mu_p) \\ &:= \inf_{\pi \in \mathscr{C}(f\mu, \mu)} \left(\int_{\mathbb{R}^n \times \mathbb{R}^n} \min\left(|y|^{p-2}|x-y|^2, |x-y|^p \right) \mathrm{d}\pi(x,y) \right) \\ &\leq \frac{1}{\tilde{c}} \mathrm{Ent}_{\mu_p}(f). \end{aligned}$$

Here, c *and* $\tilde{c}$ *are the same constants as in previous theorem.*

Idea of the proof   Using the idea of Bobkov and Ledoux(2000), the Prékopa-Leindler inequality implies modified Log-Sobolev inequality and transportation cost inequality. Prékopa-Leindler inequality is:
For some $s,\ t > 0$, $s + t = 1$, u, v, w are non-negative measurable functions on $\mathbb{R}^n$ such that

$$w(sx + ty) \geq su(x) + tv(y), \quad x,\ y \in \mathbb{R}^n,$$

then

$$\int_{\mathbb{R}^n} w \geq \Big(\int_{\mathbb{R}^n} u\Big)^s \Big(\int_{\mathbb{R}^n} v\Big)^t.$$

Conditions in Bobkov and Ledoux(2000) are: V is a convex function such that for some $c > 0$, for all $s,\ t > 0,\ s + t = 1$,

$$tV(x) + sV(y) - V(tx + sy) \geq \frac{c}{p}(s + o(s))|x - y|^p,$$

where $p \geq 2$, $o(s)/s \to 0$, as $s \to 0^+$.

Lemma: Let B be an compact set in $\mathbb{R}^n$, $V \in C^1(\mathbb{R}^n)$. Then

$$\begin{aligned} &\lim_{s\to 0^+} \sup_{h\in B} \frac{1}{s}[sV(z + th) + tV(z - th) - sV(z)] \\ &= \sup_{h\in B} \lim_{s\to 0^+} \frac{1}{s}[sV(z + th) + tV(z - th) - sV(z)] \\ &= \sup_{h\in B} V(z + h) - V(z) - \langle \nabla V(z), h\rangle. \end{aligned}$$

Proof of the Lemma: Set $\Phi_s(x,y) := sV(x) + tV(y) - V(sx+ty)$ and $\Psi(z,h) = V(z+h) - V(z) - \langle \nabla V(z), h\rangle$. For each fixed $z \in \mathbb{R}^n$, there exists a constant M such that $|\nabla V(x)| \le M$ for any $x \in \{x : \ z+\lambda B, \ \lambda \in [0,1]\}$ due to $V \in C^1$. Since

$$\lim_{s\to 0^+} \frac{\Phi_s(z+th, z-sh)}{s} = \Psi(z,h),$$

then for any $\varepsilon > 0$, for each $h \in B$, $\exists\ \delta_h > 0$ such that $\forall\ 0 < s < \delta_h$,

$$\left|\frac{\Phi_s(z+th, z-sh)}{s} - \Psi(z,h)\right| < \varepsilon.$$

Let

$$A_n = \left\{h \in B : \ \max_{0\le s\le 1/n}\left|\frac{\Phi_s(z+th, z-sh)}{s} - \Psi(z,h)\right| < \varepsilon\right\}.$$

Then $\bigcup_{n=1}^{\infty} A_n = B$. Each A_n is an open set.

Indeed, if $h_0 \in A_n$, then

$$a := \max_{0 \le t \le 1/n} \left| \frac{\Phi_s(z + th_0, z - sh_0)}{s} - \Psi(z, h_0) \right| < \varepsilon.$$

$h \in B$ such that $|h - h_0| < \frac{\varepsilon - a}{4M}$ belong to A_n. So A_n is an open set. By the compactness of B, there exists a N such that $\bigcup_{n=1}^{N} A_n = B$. Then for any $0 < s \le 1/N$,

$$\sup_{h \in B} \left\{ \Psi(z, h) \right\} - \varepsilon \le \sup_{h \in B} \left\{ \frac{\Phi_s(z + th, z - sh)}{s} \right\} \le \sup_{h \in B} \left\{ \Psi(z, h) \right\} + \varepsilon.$$

Hence,

$$\sup_{h \in B} \left\{ \Psi(z, h) \right\} - \varepsilon \le \lim_{s \to 0^+} \sup_{h \in B} \left\{ \frac{\Phi_s(z + th, z - sh)}{s} \right\} \le \sup_{h \in B} \left\{ \Psi(z, h) \right\} + \varepsilon.$$

By the arbitrariness of ε, the proof is complete.

Proof of transportation cost inequality: Let $\mu_N(\mathrm{d}x) = \mathbf{1}_{B_N}(x)\mu(\mathrm{d}x)/Z_N$, where $Z_N = \mu(B_N)$, B_N is the open ball in $\mathbb{R}^n$ with radius N centered at 0. Set $f_N = \mathbf{1}_{B_N} f/\mu_N(f)$. Given a Lipschitz continuous function g, for $s,\ t > 0$, $s + t = 1$, set $w \equiv 1$, $u(x) = e^{tg_t(x)}$, $v(y) = e^{-sg(y)}$, where

$$g_t(x) := \inf_{y \in B_N} \left\{ g(y) + \frac{\Phi_s(x,y)}{st} \right\}, \quad x \in \mathbb{R}^n.$$

Recall $\Phi_s(x,y) = sV(x) + tV(y) - V(sx+ty)$. It is easy to check that

$$w(sx+ty) \geq u(x)^s v(y)^t e^{-\Phi_s(x,,y)}, \quad x,\ y \in B_N.$$

By the Prékopa-Leindler inequality, we get

$$1 \geq \left(\int_{B_N} e^{tg_t} \,\mathrm{d}\mu \right)^s \left(\int_{B_N} e^{-sg} \,\mathrm{d}\mu \right)^t$$

$$1 \geq \left(\int_{B_N} e^{tg_t} \,\mathrm{d}\mu \right)^{\frac{1}{t}} \left(\int_{B_N} e^{-sg} \,\mathrm{d}\mu \right)^{\frac{1}{s}}$$

Letting $t \to 1 (s \to 0)$, and noting that $g_t(x)$ converges to

$$\widetilde{Q}g(x) := \inf_{y \in B_N} \Big\{ g(y) + V(x) - V(y) - \langle \nabla V(y), x - y \rangle \Big\},$$

it yields that

$$\int_{B_N} e^{\widetilde{Q}g} \, \mathrm{d}\mu_N \cdot e^{-\int_{B_N} g \, \mathrm{d}\mu_N} \leq 1. \tag{8}$$

By the variation formula of entropy $\mathrm{Ent}_\mu(f) = \sup\{\int fg \, \mathrm{d}\mu \big| \int e^g \, \mathrm{d}\mu \leq 1\}$, we have

$$\int_{B_N} f_N \widetilde{Q}g \, \mathrm{d}\mu_N - \int_{B_N} g \, \mathrm{d}\mu_N \leq \mathrm{Ent}_{\mu_N}(f_N).$$

By the assumption on V, we have

$$Qg(x) := \inf_{y \in B_N} \Big\{ g(y) + \Psi(x, y) \Big\} \leq \widetilde{Q}g(x).$$

The Kantorovich dual representation for the Wasserstein distance yields that

$$W_\Psi(f_N \mu_N, \mu_N) \leq \mathrm{Ent}_{\mu_N}(f_N).$$

Since $\Psi(x,y)$ is lower semi-continuous on $\mathbb{R}^n \times \mathbb{R}^n$, and μ_N, $f_N\mu_N$ weakly converge to μ, $f\mu$ respectively, we obtain

$$W_\Psi(f\mu,\mu) \leq \lim_{N\to+\infty} W_\Psi(f_N\mu_N,\mu_N) \leq \lim_{N\to+\infty} \mathrm{Ent}_{\mu_N}(f_N) = \mathrm{Ent}_\mu(f).$$

The proof of this theorem is completed now.

Thank You!