



Transportation-information inequalities for Markov processes (II) : relations with other functional inequalities

Arnaud Guillin, Christian Léonard, Feng-Yu Wang, Liming Wu

► To cite this version:

Arnaud Guillin, Christian Léonard, Feng-Yu Wang, Liming Wu. Transportation-information inequalities for Markov processes (II) : relations with other functional inequalities. 2009. <hal-00360854>

HAL Id: hal-00360854

<https://hal.archives-ouvertes.fr/hal-00360854>

Submitted on 12 Feb 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

TRANSPORTATION-INFORMATION INEQUALITIES FOR MARKOV PROCESSES (II) : RELATIONS WITH OTHER FUNCTIONAL INEQUALITIES

ARNAUD GUILLIN, CHRISTIAN LÉONARD, FENG-YU WANG, AND LIMING WU

ABSTRACT. We continue our investigation on the transportation-information inequalities $W_p I$ for a symmetric markov process, introduced and studied in [13]. We prove that $W_p I$ implies the usual transportation inequalities $W_p H$, then the corresponding concentration inequalities for the invariant measure μ . We give also a direct proof that the spectral gap in the space of Lipschitz functions for a diffusion process implies $W_1 I$ (a result due to [13]) and a Cheeger type's isoperimetric inequality. Finally we exhibit relations between transportation-information inequalities and a family of functional inequalities (such as Φ -log Sobolev or Φ -Sobolev).

keywords: Wasserstein distance; entropy; Fisher information; transport-information inequality; deviation inequality.

MSC 2000: 60E15, 60K35; 60G60.

1. INTRODUCTION

Let (\mathcal{X}, d) be a complete and separable metric space (say Polish) and μ a given probability measure on $(\mathcal{X}, \mathcal{B})$ where \mathcal{B} is the Borel σ -field. Let $(X_t)_{t \geq 0}$ be a μ -symmetric ergodic conservative Markov process valued in \mathcal{X} , with transition semigroup (P_t) (which is symmetric on $L^2(\mu)$), and Dirichlet form $(\mathcal{E}(\cdot, \cdot), \mathbb{D}(\mathcal{E}))$ where $\mathbb{D}(\mathcal{E})$ is the domain of \mathcal{E} in $L^2(\mu) := L^2(\mathcal{X}, \mathcal{B}, \mu)$. Here the ergodicity means simply : for $g \in \mathbb{D}(\mathcal{E})$, $\mathcal{E}(g, g) = 0$ iff $g = c$.

For $1 \leq p < +\infty$ fixed and for any probability measure ν on \mathcal{X} (written as $\nu \in \mathcal{M}_1(\mathcal{X})$), consider

(i): L^p -Wasserstein distance between ν and μ :

$$W_p(\nu, \mu) := \inf_{\pi \in \mathcal{C}(\nu, \mu)} \iint_{E^2} d^p(x, y) \pi(dx, dy) \quad (1.1)$$

where $\mathcal{C}(\nu, \mu)$ are the set of all couplings of (ν, μ) , i.e., probability measures π on $E \times E$ such that $\pi(A \times E) = \nu(A)$ and $\pi(\mathcal{X} \times A) = \mu(A)$ for all $A \in \mathcal{B}$.

(ii): Relative entropy or Kullback's information of ν w.r.t. μ

$$H(\nu|\mu) := \begin{cases} \int \frac{d\nu}{d\mu} \log \frac{d\nu}{d\mu} d\mu, & \text{if } \nu \ll \mu; \\ +\infty, & \text{otherwise.} \end{cases} \quad (1.2)$$

def-entropy-a

(iii): The Fisher information of ν w.r.t. μ :

Date: First Version: March 2008.

$$I(\nu|\mu) := \begin{cases} \mathcal{E}(\sqrt{f}, \sqrt{f}). & \text{if } \nu = f\mu, \sqrt{f} \in \mathbb{D}(\mathcal{E}), \\ +\infty & \text{otherwise.} \end{cases} \quad (1.3) \quad \boxed{\text{Fisher}}$$

The usual transport inequalities $W_p H$, introduced and studied by K. Marton [16] and M. Talagrand [18] mean that

$$W_p(\nu, \mu)^2 \leq 2CH(\nu|\mu), \quad \forall \nu \in M_1(\mathcal{X}). \quad (W_p H(C))$$

Its study is very active: see Bobkov-Götze [4], Otto-Villani [17], Bobkov-Gentil-Ledoux [3], Djellout-Guillin-Wu [10] and references therein. Furthermore Gozlan-Léonard [12] consider the following generalized transportation cost from ν to μ :

$$T_{\mathcal{V}}(\nu, \mu) := \sup\{\nu(u) - \mu(v); (u, v) \in \mathcal{V}\}$$

($\mu(u) := \int_E u d\mu$) where \mathcal{V} is some given family of $(u, v) \in (b\mathcal{B})^2$ so that

- (A1) $u \leq v$ for all $(u, v) \in \mathcal{V}$ (or equivalently $T_{\mathcal{V}}(\nu, \nu) \leq 0$ for all $\nu \in M_1(\mathcal{X})$);
- (A2) For all $\nu_1, \nu_2 \in M_1(\mathcal{X})$, there exists $(u, v) \in \mathcal{V}$ such that $\int u d\nu_1 - \int v d\nu_2 \geq 0$ (or equivalently $T_{\mathcal{V}}(\nu_1, \nu_2) \geq 0$ for all $\nu_1, \nu_2 \in M_1(\mathcal{X})$).

And they introduced the following generalization of $W_p H$: for some convex, non-decreasing and left continuous function α on \mathbb{R}^+ ,

$$\alpha(T_{\mathcal{V}}(\nu, \mu)) \leq 2CH(\nu|\mu), \quad \forall \nu \in M_1(\mathcal{X}) \quad (\alpha - T_{\mathcal{V}} H(C))$$

and they established its equivalence with some concentration inequality of the underlying measure μ and of the i.i.d. sequences of common law μ .

Recall that $T_{\mathcal{V}}(\nu, \mu) = W_p(\nu, \mu)^2$ iff $\mathcal{V} = \mathcal{V}(p, d)$, the family of all couples (u, v) of real bounded measurable functions on \mathcal{X} such that

$$u(x) - v(y) \leq d^p(x, y), \quad \forall x, y \in E. \quad (1.4) \quad \boxed{\text{A31}}$$

Guillin-Leonard-Wu-Yao [13] propose a new transport-information inequality, adapted to Markov processes (and in particular to consider deviation inequalities for integral functionals of Markov processes)

$$W_p(\nu, \mu)^2 \leq 4C^2 I(\nu|\mu), \quad \forall \nu \in M_1(\mathcal{X}) \quad (W_p I(C))$$

or the more general

$$\alpha(T_{\mathcal{V}}(\nu, \mu)) \leq I(\nu|\mu), \quad \forall \nu \in M_1(\mathcal{X}). \quad (\alpha - T_{\mathcal{V}} I)$$

Using large deviations techniques they prove the following characterization:

thm-GLWY

Theorem 1.1. ([13]) *Let $((X_t)_{t \geq 0}, \mathbb{P}_\mu)$ be the μ -symmetric and ergodic Markov associated with the Dirichlet form $(\mathcal{E}, \mathbb{D}(\mathcal{E}))$, $\alpha : \mathbb{R}^+ \rightarrow [0, +\infty]$ a left-continuous non-decreasing convex function with $\alpha(0) = 0$, and \mathcal{V} as above.*

The following properties are equivalent:

- (a): μ satisfies the transport-information inequality $(\alpha - T_{\mathcal{V}} I)$.
- (b): For all $(u, v) \in \mathcal{V}$ and all $\lambda \geq 0$

$$\lambda_{\max}(\mathcal{L} + \lambda u) := \sup_{g \in \mathbb{D}(\mathcal{E}) : \mu(g^2) = 1} \left[\lambda \int u g^2 d\mu - \mathcal{E}(g, g) \right] \leq \lambda \mu(v) + \alpha^*(\lambda) \quad (1.5) \quad \boxed{\text{L2bound}}$$

where \mathcal{L} is the generator of (P_t) on $L^2(\mathcal{X}, \mu)$ and

$$\alpha^*(\lambda) = \sup_{r \geq 0} \{\lambda r - \alpha(r)\}, \quad \forall \lambda \geq 0$$

is the semi-Legendre transformation of α .

(c): For any initial measure $\nu = f\mu$ with $f \in L^2(\mu)$ and for all $(u, v) \in \mathcal{V}$

$$\mathbb{P}_\nu \left(\frac{1}{t} \int_0^t u(X_s) ds \geq \mu(v) + r \right) \leq \|f\|_2 e^{-t\alpha(r)}, \quad \forall t, r > 0. \quad (1.6)$$

thmA11a

Remarks 1.2. The meaning of the deviation inequality characterization (1.6) of α - $W_p I$ is clear in the ergodic behavior of the Markov process (X_t) , as well as (1.5) in the study of the Schrödinger operator $\mathcal{L} + u$. That is one more reason why α - $T_\nu I$ inequality is useful.

Remarks 1.3. If \mathcal{V} is some family of $(u, u) \in (b\mathcal{B})^2$, (1.6) becomes a deviation inequality of the empirical (time) mean from its space mean $\mu(u)$ for the observable u so that $(u, u) \in \mathcal{V}$. Notice that if $\mathcal{V} = \{(u, u); u \in b\mathcal{B}, \|u\|_{Lip} \leq 1\}$ then $T_\nu(\nu, \mu) = W_1(\nu, \mu)$, and $W_1 I(C)$ is equivalent to the Gaussian deviation inequality (1.6) with $\alpha(r) = r^2/(4C^2)$ for the Lipschitzian observable u with Lipschitzian coefficient $\|u\|_{Lip} \leq 1$, which generalizes the well known Hoeffding's inequality in the i.i.d. case.

Three criteria for $W_1 I(C)$ are established in [13]: spectral gap in $L^2(\mu)$; spectral gap in the space of Lipschitz functions and a very general Lyapunov function criterion if $\mathcal{V} = \{(u, u); |u| \leq \phi\}$ where $\phi > 0$ is some fixed weight function. And it is also shown that on a Riemannian manifold \mathcal{X} equipped with the Riemannian metric d , the log-Sobolev inequality

$$H(\nu|\mu) \leq 2CI(\nu|\mu), \quad \forall \nu \in M_1(\mathcal{X}). \quad (HI(C))$$

implies $W_2 I(C)$, which in turn implies the Poincaré inequality

$$Var_\mu(g) \leq c_P \mathcal{E}(g, g), \quad \forall g \in L^2(\mu) \cap \mathbb{D}(\mathcal{E}) \quad (P(C))$$

where $Var_\mu(g) = \mu(g^2) - \mu(g)^2$ is the variance. Furthermore $W_2 I(C) \implies HI(C')$ once if the Ricci-Bakry-Emery curvature of μ is bounded from below.

We organize this paper around the four questions below:

- (i): Investigate the relations between $W_p I$ with $W_p H$. That is the objective of §2.
- (ii): Prove that the spectral gap in the space of Lipschitz functions implies a Cheeger type's isoperimetric inequality, which is stronger than $W_1 I$. That is the purpose of §3. We will also establish deviation inequalities under natural quantities such as the variance of the test function, refining [13].
- (iii): In §4 we study relations between $(\alpha - W_2 I)$ and the β -log-Sobolev inequality:

$$\beta \circ \mu(g^2 \log g^2) \leq \mathcal{E}(g, g), \quad \mu(g^2) = 1, g \in \mathbb{D}(\mathcal{E}), \quad (1.7)$$

Phi

where β is a positive increasing function. This inequality was connected in [24] to the well developed F -Sobolev inequality introduced in [21], so that known criteria for the later can be applied directly to (1.7).

- (iv): Finally we present in §5 applications of Φ -Sobolev inequality

$$\|g^2\|_\Phi \leq C_1 \mathcal{E}(g, g) + C_2 \mu(g^2)$$

in transportation-information inequalities $\alpha - T_\nu I$ and then in the concentration phenomena of $\frac{1}{t} \int_0^t u(X_s) ds$ under integrability conditions on u .

2. $W_p I$ IMPLIES $W_p H$ ON A RIEMANNIAN MANIFOLD

Recall (cf. Villani [19]) the well known Kantorovitch's dual characterization :

$$W_p^p(\nu, \mu) = \sup_{(u,v) \in \mathcal{V}(p,d)} \int u d\nu - \int v d\mu \quad (2.1) \quad \boxed{\text{A32}}$$

where $\mathcal{V}(p, d)$ is given in (1.4), and Kantorovitch-Robinstein's identity

$$W_1(\nu, \mu) = \sup_{\|u\|_{Lip} \leq 1} \int u d(\nu - \mu). \quad (2.2) \quad \boxed{\text{A33}}$$

Throughout this section \mathcal{X} is a connected complete Riemannian manifold equipped with the Riemannian metric d , and $\mu = e^{-V} dx/Z$ (Z being the normalization constant assumed to be finite) with $V \in C^1(\mathcal{X})$, and $(\mathcal{E}, \mathbb{D}(\mathcal{E}))$ is the closure of

$$\mathcal{E}(g, g) := \int_E |\nabla g|^2 d\mu(x), \quad g \in C_0^\infty(\mathcal{X})$$

where ∇ is the gradient on \mathcal{X} , and $C_0^\infty(\mathcal{X})$ is the space of infinitely differentiable functions on \mathcal{X} with compact support. In such case our Fisher information of $f\mu$ with $0 < f \in C^1(\mathcal{X})$ w.r.t. μ becomes

$$I(f\mu|\mu) = \frac{1}{4} \int \frac{|\nabla f|^2}{f} d\mu = \frac{1}{4} \int |\nabla \log f|^2 d\mu.$$

2.1. $W_1 I(C) \implies W_1 H(C)$.

thm21

Theorem 2.1. *Assume that μ satisfies $W_1 I(C)$. Then*

$$W_1(\nu, \mu)^2 \leq 2CH(\nu|\mu), \quad \forall \nu \in \mathcal{M}_1(\mathcal{X})$$

i.e., μ satisfies $W_1 H(C)$.

Proof. By Bobkov-Götze's criterion [4] for $W_1 H(C)$, it is enough to show that for any bounded $g \in C^1(\mathcal{X})$ with $|\nabla g| \leq 1$ and $\lambda \geq 0$,

$$\int e^{\lambda(g - \mu(g))} d\mu \leq e^{\lambda^2 C^2 / 2}. \quad (2.3) \quad \boxed{\text{thm21b}}$$

To this end we may assume that $\mu(g) = 0$. Consider

$$Z(\lambda) = \int e^{\lambda g} d\mu, \quad \mu_\lambda := \frac{e^{\lambda g}}{Z(\lambda)} \mu.$$

We have by Kantorovitch's identity (2.2)

$$\frac{d}{d\lambda} \log Z(\lambda) = \mu_\lambda(g) \leq W_1(\mu_\lambda, \mu)$$

but by $W_1 I(C)$,

$$W_1(\mu_\lambda, \mu) \leq 2C \sqrt{I(\mu_\lambda|\mu)} = C\lambda \sqrt{\int |\nabla g|^2 d\mu_\lambda} \leq C\lambda.$$

Thus

$$\log Z(\lambda) \leq \int_0^\lambda C t dt = \frac{C\lambda^2}{2}$$

the desired control (2.3). \square

The implication " $W_1 I(C) \implies W_1 H(C)$ " is strict, as shown by the following simple counter-example ([10]).

Example 2.2. Let $\mathcal{X} = [-2, -1] \cup [1, 2]$ and $\mu(dx) = (1_{[-2, -1]} + 1_{[1, 2]})dx/2$. The Dirichlet form $(\mathcal{E}, \mathbb{D}(\mathcal{E}))$ is given by

$$\mathcal{E}(f, f) = \int f'^2 d\mu(x), \quad \forall f \in \mathbb{D}(\mathcal{E}) = H^1(\mathcal{X})$$

where $H^1(\mathcal{X})$ is the space of those functions $f \in L^2(\mu)$ so that $f' \in L^2(\mu)$ (in the distribution sense). It corresponds to the reflecting Brownian Motion in \mathcal{X} , which is not ergodic. But $W_1I(C)$ implies always the ergodicity. Thus μ does not satisfy $W_1I(C)$. However μ satisfies $W_1H(C)$ by the Gaussian integrability criterion in [10].

The argument above can be extended to more general transportation information inequality $\alpha - W_1I$:

prop21

Proposition 2.3. Let $\alpha : \mathbb{R}^+ \rightarrow [0, +\infty]$ be a left-continuous non-decreasing convex function with $\alpha(0) = 0$. Assume that μ satisfies $\alpha - W_1I$. Then μ satisfies

$$\tilde{\alpha}(W_1(\nu, \mu)) \leq H(\nu|\mu), \quad \forall \nu \in \mathcal{M}_1(\mathcal{X}) \quad (\tilde{\alpha} - W_1H) \quad \text{prop21a}$$

where $\tilde{\alpha}(r) = 2 \int_0^r \sqrt{\alpha(s)} ds$. In particular for any Lipschitzian function g with $\|g\|_{Lip} \leq 1$,

$$\mu(g > \mu(g) + r) \leq e^{-\tilde{\alpha}(r)}, \quad \forall r > 0.$$

Proof. By Gozlan-Léonard's criterion [12] for $\tilde{\alpha} - W_1H$, it is enough to show that for any bounded $g \in C^1(\mathcal{X})$ with $|\nabla g| \leq 1$ and $\lambda \geq 0$,

$$\int e^{\lambda(g - \mu(g))} d\mu \leq e^{\tilde{\alpha}^*(\lambda)}, \quad \forall \lambda \geq 0 \quad (2.4) \quad \text{prop21c}$$

which implies the last concentration inequality in this Proposition by Chebychev's inequality. To show (2.4) we may assume that $\mu(g) = 0$. Let $Z(\lambda)$ and μ_λ be as in the previous proof of Theorem 2.1, we have

$$\frac{d}{d\lambda} \log Z(\lambda) = \mu_\lambda(g) \leq W_1(\mu_\lambda, \mu).$$

But by the assumed $\alpha - W_1I$,

$$W_1(\mu_\lambda, \mu) \leq \alpha^{-1}(I(\mu_\lambda|\mu)) = \alpha^{-1}\left(\frac{\lambda^2}{4} \int |\nabla g|^2 d\mu_\lambda\right) \leq \alpha^{-1}(\lambda^2/4)$$

where $\alpha^{-1}(t) := \inf\{t \geq 0; \alpha(r) > t\}$, $t \geq 0$. Thus

$$\log Z(\lambda) \leq \int_0^\lambda \alpha^{-1}(t^2/4) dt =: h(\lambda).$$

Now by Fenchel-Legendre theorem, $h = (h^*)^*$, but

$$h^*(r) = \sup_{\lambda \geq 0} (\lambda r - h(\lambda)) = 2 \int_0^r \sqrt{\alpha(s)} ds,$$

which completes the proof of the desired control (2.4). \square

2.2. $W_2I(C) \implies W_2H(C)$.

thm22 **Theorem 2.4.** *Assume that μ satisfies $W_2I(C)$. Then*

$$W_2(\nu, \mu)^2 \leq 2CH(\nu|\mu), \quad \forall \nu \in \mathcal{M}_1(\mathcal{X})$$

i.e., μ satisfies $W_2H(C)$.

Proof. We shall use the method of Hamilton-Jacobi equation due to Bobkov-Gentil-Ledoux [3]. Consider the inf-convolution

$$Q_t g(x) := \inf_{y \in E} (g(y) + \frac{1}{2t} d^2(x, y))$$

which is viscosity solution of the Hamilton-Jacobi equation

$$\partial_t Q_t g + \frac{1}{2} |\nabla Q_t g|^2 = 0. \quad (2.5) \quad \text{HJequ}$$

By Bobkov-Götze's criterion [4] for $W_2H(C)$, it is enough to show that for any $g \in C_b^1(\mathcal{X})$,

$$\int e^{Q_{1g}/C} d\mu \leq e^{\mu(g)/C}. \quad (2.6) \quad \text{thm22b}$$

To this end we may and will assume that $\mu(g) = 0$. Let $\lambda = \lambda(t) = \kappa t$ where $\kappa > 0$ will be determined later and consider

$$Z(t) = \int e^{\lambda Q_t g} d\mu, \quad \mu_t := \frac{e^{\lambda Q_t g}}{Z(t)} \mu.$$

We have

$$\begin{aligned} \frac{d}{dt} \log Z(t) &= \frac{1}{Z(t)} \int [\lambda'(t) Q_t g + \lambda(t) \partial_t Q_t g] e^{\lambda Q_t g} d\mu \\ &= \kappa \int Q_t g d\mu_t - \frac{\lambda}{2} \int |\nabla Q_t g|^2 d\mu_t \\ &= \kappa \int Q_t g d\mu_t - \frac{2}{\lambda} I(\mu_t|\mu). \end{aligned}$$

But by Kantorovitch's identity (2.1),

$$\int Q_t g d\mu_t \leq \frac{1}{2t} W_2^2(\mu_t, \mu) \quad (2.7) \quad \text{thm22c}$$

and the assumed $W_2I(C)$ gives $W_2^2(\mu_t, \mu) \leq 4C^2 I(\mu_t|\mu)$. Thus for every $t > 0$,

$$\frac{d}{dt} \log Z(t) \leq \left(\frac{2\kappa C^2}{t} - \frac{2}{\kappa t} \right) I(\mu_t|\mu)$$

Putting $\kappa = 1/C$, we obtain $\frac{d}{dt} \log Z(t) \leq 0$ for all $t > 0$, which implies by the continuity of $\log Z(t)$ on \mathbb{R}^+ that

$$\int e^{Q_{1g}/C} d\mu = Z(1) \leq Z(0) = 1$$

the desired (2.6). □

Remarks 2.5. The proof above is adapted from that of Bobkov-Gentil-Ledoux [3] for the implication $HI(C) \implies W_2H(C)$, originally established by Otto-Villani [17].

Remarks 2.6. We have thus established in this section

$$HI(C) \Rightarrow W_2I(C) \Rightarrow W_2H(C).$$

It was also established in [13] that under a lower bound of the Ricci-Bakry-Emery curvature of μ that W_2I implies back to HI , and with additional conditions on this lower bound that W_2H implies back HI . It is then a natural question to know if the condition on the lower bound of the Ricci-Bakry-Emery curvature is also necessary to get the reverse implication. A partial answer was provided in [7] where an example of a real probability measure, with infinite lower bounded curvature, was shown to verify W_2H but not HI . Inspired by this example, we furnish here an example where W_1I holds (using Lyapunov conditions of [13, Section 5]) but not HI . Let then consider $d\mu(x) = e^{-V(x)}dx$, where V is symmetric C^2 (at least) and given for large x by

$$V(x) = x^4 + 4x^3 \sin^2(x) + x^\beta.$$

Consider also the natural reversible process associated to this measure given by generator $Lf = f'' - V'f'$. Using $W(x) = e^{ax^4}$, by easy calculus, one sees that $LW \leq -cx^4W + b$ (for some positive b and c) if $\beta > 2$. This Lyapunov condition also implies a Poincaré inequality (see [1] for example), so that using a slight modification of [13, Lem. 5.7], we get that W_1I holds and also W_2H by [7]. Remark now that if $\beta < 3$ then V/V'^2 is not bounded, which is a known necessary condition for HI to hold (see [7]). Unfortunately, we are not up to now able to prove that W_2I holds.

3. W_1I AND THE ISOPERIMETRIC INEQUALITY OF CHEEGER'S TYPE BY MEANS OF THE SPECTRAL GAP IN C_{Lip}

In this section we return to the general Polish space case (\mathcal{X}, d) . We assume that μ charges all non-empty open subsets of \mathcal{X} .

Let C_{Lip} be the space of all real functions g on \mathcal{X} which are Lipschitz-continuous, i.e., $\|g\|_{Lip} := \sup_{x \neq y} \frac{|f(x) - f(y)|}{d(x, y)} < +\infty$. We assume that there is an algebra $\mathcal{A} \subset C_{Lip} \cap \mathbb{D}_2(\mathcal{L})$ (here $\mathbb{D}_2(\mathcal{L})$ is the domain of the generator \mathcal{L} in $L^2(\mu)$ associated with $(\mathcal{E}, \mathbb{D}(\mathcal{E}))$), which is a form core for $(\mathcal{E}, \mathbb{D}(\mathcal{E}))$. Hence the carré-du-champs operator

$$\Gamma(f, g) := \frac{1}{2} (\mathcal{L}(fg) - f\mathcal{L}g - g\mathcal{L}f), \quad \forall f, g \in \mathcal{A}$$

admits a unique continuous extension $\Gamma : \mathbb{D}(\mathcal{E}) \times \mathbb{D}(\mathcal{E}) \rightarrow L^1(\mathcal{X}, \mu)$. Throughout this section we assume that Γ is a differentiation, that is, for all $(h_k)_{1 \leq k \leq n} \subset \mathcal{A}$, $g \in \mathcal{A}$ and $F \in C_b^1(\mathbb{R}^n)$,

$$\Gamma(F(h_1, \dots, h_n), g) = \sum_{i=1}^n \partial_i F(h_1, \dots, h_n) \Gamma(h_i, g)$$

(this can be extended to $\mathbb{D}(\mathcal{E})$).

thm31

Theorem 3.1. *Assume that $\int d^2(x, x_0)d\mu(x) < +\infty$ for some (or all) $x_0 \in E$ and Γ is a differentiation. Suppose that there is a form core $\mathcal{D} \subset C_{Lip} \cap \mathbb{D}_2(\mathcal{L})$ of $(\mathcal{E}, \mathbb{D}(\mathcal{E}))$ such that $1 \in \mathcal{D}$ and*

$$W_1(\nu, \mu) = \sup_{g \in \mathcal{D}: \|g\|_{Lip} \leq 1} \left\{ \int gd(\nu - \mu) \right\}, \quad \forall \nu \text{ with } I(\nu|\mu) < +\infty \quad (3.1) \quad \text{thm31a}$$

and

$$\sqrt{\Gamma(g, g)} \leq \sigma \|g\|_{Lip}, \quad \mu - a.s., \quad \forall g \in C_{Lip} \cap \mathbb{D}_2(\mathcal{L}) \quad (3.2) \quad \text{thm31b}$$

and for some constant $C > 0$ and for any $g \in \mathcal{D}$ with $\mu(g) = 0$, there is $G \in C_{Lip} \cap \mathbb{D}_2(\mathcal{L})$ so that

$$-\mathcal{L}G = g, \quad \|G\|_{Lip} \leq C\|g\|_{Lip}. \quad (3.3) \quad \boxed{\text{thm31c}}$$

Then the Poincaré inequality holds with $c_P \leq C$, and the following isoperimetric inequality of Cheeger's type

$$W_1(f\mu, \mu) \leq \sigma C \int \sqrt{\Gamma(f, f)} d\mu, \quad 0 \leq f \in \mathbb{D}(\mathcal{E}), \mu(f) = 1 \quad (3.4) \quad \boxed{\text{thm31d}}$$

holds true. In particular,

$$W_1(\nu, \mu)^2 \leq 4(\sigma C)^2 I(\nu|\mu), \quad \forall \nu \in \mathcal{M}_1(\mathcal{X}). \quad (3.5) \quad \boxed{\text{thm31e}}$$

Furthermore for any observable g with $\|g\|_{Lip} = 1$,

$$\int g d(\nu - \mu) \leq 2\sqrt{I(\nu|\mu) \left[\frac{V(g)}{2} + 2(\sigma C)^2 \sqrt{c_P I(\nu|\mu)} \right]} \quad (3.6) \quad \boxed{\text{thm31f}}$$

and for any $t, r, \delta > 0$,

$$\begin{aligned} & \mathbb{P}_\beta \left(\frac{1}{t} \int_0^t g(X_s) ds > \mu(g) + r \right) \\ & \leq \left\| \frac{d\beta}{d\mu} \right\|_2 \exp \left(-t \frac{r^2}{(1+\delta)V(g) + \sqrt{[(1+\delta)V(g)]^2 + \frac{8c_P(\sigma C)^4}{\delta V(g)} r^2}} \right) \end{aligned} \quad (3.7) \quad \boxed{\text{thm31g}}$$

where $V(g) := \lim_{t \rightarrow \infty} \frac{1}{t} \text{Var}_{\mathbb{P}_\mu} \left(\int_0^t g(X_s) ds \right) = 2 \int_0^\infty \langle g - \mu(g), P_t g \rangle_\mu dt$ is the asymptotic variance of g .

Proof. Under the Lipschitzian spectral gap condition (3.3), it is noted in [13] that the Poincaré inequality holds with $c_P \leq C$.

For both (3.4) and (3.5) we may assume that $\nu = f\mu$ with $f \in \mathbb{D}(\mathcal{E})$, $f \geq \varepsilon > 0$. For any $g \in \mathcal{D}$ with $\|g\|_{Lip} \leq 1$ and $\mu(g) = 0$, letting $G := (-\mathcal{L})^{-1}g$ be the unique solution of the Poisson equation with $\mu(G) = 0$, we have

$$\int g d\nu - \int g d\mu = \langle g, f \rangle_\mu = \mathcal{E}(G, f) = \int \Gamma(G, f) d\mu \leq \|\sqrt{\Gamma(G, G)}\|_\infty \int \sqrt{\Gamma(f, f)} d\mu.$$

Taking the supremum over all such g and observing $\|\sqrt{\Gamma(G, G)}\|_\infty \leq \sigma \|G\|_{Lip} \leq \sigma C$ we obtain (3.4). Furthermore by Cauchy-Schwarz and the fact that Γ is a differentiation, we have

$$\int \sqrt{\Gamma(f, f)} d\mu \leq \sqrt{\int \frac{\Gamma(f, f)}{f} d\mu} \sqrt{\int f d\mu} = 2\sqrt{I(\nu|\mu)}$$

where (3.5) follows from (3.4).

For (3.6) writing $f = h^2$, we have

$$\begin{aligned} \int g d\nu - \int g d\mu &= \int \Gamma(G, f) d\mu = 2 \int h \Gamma(G, h) d\mu \\ &\leq 2\sqrt{\int \Gamma(h, h) d\mu} \cdot \int \Gamma(G, G) h^2 d\mu. \end{aligned}$$

Using the inequality in [13, Theorem 3.1]

$$\int \Gamma(G, G)h^2 d\mu - \int \Gamma(G, G)d\mu \leq \|\Gamma(G, G)\|_\infty \|h^2\mu - \mu\|_{TV} \leq (\sigma C)^2 \sqrt{4c_P I(\nu|\mu)}$$

and noting that $V(g) = 2\langle (-\mathcal{L})^{-1}g, g \rangle_\mu = 2\mathcal{E}(G, G) = 2 \int \Gamma(G, G)d\mu$, we obtain

$$\int g d\nu - \int g d\mu \leq 2\sqrt{I(\nu|\mu) \left[\frac{V(g)}{2} + 2(\sigma C)^2 \sqrt{c_P I(\nu|\mu)} \right]}$$

which is (3.6). Using $2I^{3/2} \leq \varepsilon I + I^2/\varepsilon$ in (3.6), we obtain (3.7) by Theorem 1.1. \square

Remarks 3.2. The $W_1 I(\sigma C)$ inequality (3.5) is due to Guillin and al. [13], but the method therein is based on the Lyons-Meyer-Zheng forward-backward martingale decomposition. The argument here is simpler and direct, and yields the stronger Cheeger type's isoperimetric inequality (3.4).

Remarks 3.3. Letting δ be close to 0, we see that (3.7) is sharp for small r by the central limit theorem.

Set $C_{Lip,0} = \{g \in C_{Lip,0}; \mu(g) = 0\}$. Under the Lipschitzian spectral gap condition (3.3), the Poisson operator $(-\mathcal{L})^{-1} : C_{Lip,0} \rightarrow C_{Lip,0}$ is a well defined bounded linear operator w.r.t. the Lipschitzian norm, and the best constant C in (3.3) is the Lipschitzian norm $\|(-\mathcal{L})^{-1}\|_{Lip}$ and will be denoted by $c_{Lip,P}$ (the index P is referred to Poincaré).

We now present four examples for illustrating usefulness of Theorem 3.1.

Example 3.4. (Ornstein-Uhlenbeck process) Consider the Ornstein-Uhlenbeck process $dX_t = \sqrt{2}dB_t - \sigma^{-2}X_t dt$ on $\mathcal{X} = \mathbb{R}$ where $\sigma > 0$ and B_t is the standard Brownian motion on \mathbb{R} . Its unique invariant measure is $\mu = \mathcal{N}(0, \sigma^2)$. For $f \in C_b^\infty(\mathbb{R})$, from the explicit solution $X_t = e^{-\sigma^{-2}t} \left(X_0 + \int_0^t e^{\sigma^{-2}s} \sqrt{2}dB_s \right)$, we see that $(P_t f)' = e^{-\sigma^{-2}t} P_t f'$. Hence $c_{Lip,P} = \|(-\mathcal{L})^{-1}\|_{Lip} = \sigma^2$. Therefore μ satisfies $W_1 I(C)$ with $C = c_{Lip} = \sigma^2$ by Theorem 3.1.

Furthermore $C = c_{Lip} = \sigma^2$ is also the best constant in $W_1 I(C)$. Indeed $W_1 I(C) \implies W_1 H(C)$ and the best constant in $W_1 H(C)$ of μ is $C = \sigma^2$. In other words Theorem 3.1 produces the exact best constant C in $W_1 I(C)$ for this example.

Example 3.5. (Reflected Brownian Motion) Consider the reflected Brownian Motion X_t^D on the interval $\mathcal{X} = [0, D]$ ($D > 0$) equipped with the usual Euclidean metric, whose generator is given by $\mathcal{L}f = f''$ with Neumann boundary condition at $0, D$. The unique invariant measure μ is the uniform law on $[0, D]$. For every $g \in C_b^2([0, D])$ with $\int_0^D g(x)dx = 0$, the solution G of the Poisson equation $-\mathcal{L}G = g$ satisfies

$$G'(x) = - \int_0^x g(t)dt, \quad x \in [0, D].$$

It is now easy to see that $c_{Lip,P} = \sup_{\|g\|_{Lip}=1} \|G'\|_\infty$ is attained with $g(x) = x - D/2$ and then $c_{Lip,P} = D^2/8$. Thus by Theorem 3.1, the optimal constant $C_{W_1 I}$ for this process satisfies $C_{W_1 I} \leq c_{Lip} = D^2/8$. In comparison recall that the best Poincaré constant $c_P = D^2/\pi^2$.

Since $W_1 I(C) \implies W_1 H(C)$ and the best constant of $W_1 H(C)$ for the uniform law μ on $[0, D]$ is $D^2/12$, so we obtain

$$\frac{D^2}{12} \leq c_{W_1 I} \leq \frac{D^2}{8}.$$

We do not know the exact value of $c_{W_1 I}$ for this simple example.

Example 3.6. Let \mathcal{X} be a compact connected Riemannian manifold of dimension n with empty or convex boundary. Assume that the Ricci curvature is nonnegative and its diameter is D . Consider the Brownian Motion (with reflection in the presence of the boundary) generated by the Laplacian operator Δ .

In [30] it is shown that $c_{Lip,P} = \|(-\Delta)^{-1}\|_{Lip} \leq D^2/8$ (the latest quantity is exactly $c_{Lip,P}$ for the reflected Brownian Motion on the interval $[0, D]$). Thus by Theorem 3.1, $W_1I(C)$ holds with $C = D^2/8$.

See [30] for more examples for which $c_{Lip,P}$ is estimated.

Example 3.7. (One-dimensional diffusions) Now let us consider the one-dimensional diffusion with values in the interval (x_0, y_0) generated by

$$\mathcal{L}f = a(x)f'' + b(x)f', f \in C_0^\infty(x_0, y_0)$$

where a, b are continuous such that $a(x) > 0$ for all $x \in (x_0, y_0)$. Let $((X_t)_{0 \leq t < \tau}, \mathbb{P}_x)$ be the martingale solution associated with \mathcal{L} and initial position x , where τ is the explosion time. With a fixed $c \in (x_0, y_0)$,

$$s'(x) := \exp\left(-\int_c^x \frac{b(t)}{a(t)} dt\right), \quad m'(x) := \frac{1}{a(x)} \exp\left(\int_c^x \frac{b(t)}{a(t)} dt\right)$$

are respectively the derivatives of Feller's scale and speed functions. Assume that

$$Z := \int_{x_0}^{y_0} m'(x) dx < +\infty \tag{3.8} \quad \boxed{D1}$$

and let $\mu(dx) = m'(x)dx/Z$. It is well known that $(\mathcal{L}, C_0^\infty(x_0, y_0))$ is symmetric on $L^2(\mu)$. Assume also that

$$\int_c^{y_0} s'(x) dx \int_c^x m'(x) dx = \int_{x_0}^c s'(x) \int_x^c m'(x) dx = +\infty \tag{3.9} \quad \boxed{D2}$$

which, in the Feller's classification, means that x_0, y_0 are no accessible or equivalently $\tau = \infty$, \mathbb{P}_x -a.s. In this case by the L^1 -uniqueness in [25], the Dirichlet form

$$\begin{aligned} \mathbb{D}(\mathcal{E}) &= \left\{ f \in \mathcal{AC}(x_0, y_0) \cap L^2(\mu); \int_{x_0}^{y_0} (f')^2 d\mu < +\infty \right\}, \\ \mathcal{E}(f, f) &= \int_{x_0}^{y_0} (f')^2 d\mu, \quad f \in \mathbb{D}(\mathcal{E}) \end{aligned}$$

is associated with (X_t) , where $\mathcal{AC}(x_0, y_0)$ is the space of absolutely continuous functions on (x_0, y_0) .

Fix some $\rho \in C^1(x_0, y_0)$ such that $\rho \in L^2(\mu)$ and $\rho'(x) > 0$ everywhere, consider the metric $d_\rho(x, y) = |\rho(x) - \rho(y)|$. A function f on (x_0, y_0) is Lipschitz with respect to d_ρ (written as $f \in C_{Lip(\rho)}$) if and only if $f \in \mathcal{AC}(x_0, y_0)$ and

$$\|f\|_{Lip(\rho)} = \sup_{x_0 < x < y < y_0} \frac{|f(y) - f(x)|}{\rho(y) - \rho(x)} = \left\| \frac{f'}{\rho'} \right\|_\infty.$$

The argument below is borrowed from [11]. Assume that

$$C(\rho) := \sup_{x \in (x_0, y_0)} \frac{1}{\rho'(x)} \int_x^{y_0} [\rho(t) - \mu(\rho)] m'(t) dt < +\infty. \tag{3.10} \quad \boxed{C-rho}$$

For every $g \in C_{Lip(\rho)}$ with $\mu(g) = 0$, then $f(x) = \int_c^x dy \int_y^{y_0} g(t) m'(t) dt - A$ (in C^2) solves

$$-(af'' + bf') = g. \tag{3.11} \quad \boxed{cor44a}$$

It is obvious that

$$\|f\|_{\text{Lip}(\rho)} = \sup_{x \in (x_0, y_0)} \frac{1}{\rho'(x)} \int_x^{y_0} g(t) m'(t) dt.$$

An elementary exercise (as done in [11]) shows that the last quantity is always not greater than $C(\rho)\|g\|_{\text{Lip}(\rho)}$. Thus $f \in L^2(\mu)$ (for $\rho \in L^2(\mu)$). By Ito's formula, $f \in \mathbb{D}_2(\mathcal{L})$. With the constant A so that $\mu(f) = 0$, f given above is the unique solution in $L^2(\mu)$ with zero mean of (3.11) by the ergodicity of (X_t) . We see also that $C(\rho)$ is the best constant by taking $g = \rho - \mu(\rho)$. In other words condition (3.3) is verified with the best constant $C = c_{\text{Lip}, P} = C(\rho)$. Hence from Theorem 3.1, we get

cor44

Corollary 3.8. *Let $a, b : (x_0, y_0) \rightarrow \mathbb{R}$ be continuous such that $a(x) > 0$ for all x and conditions (3.8) (3.9) be satisfied. Assume (3.10) and $\sigma := \sup_{x \in (x_0, y_0)} \sqrt{a(x)}\rho'(x) < +\infty$. Then μ satisfies $W_1I(\sigma C(\rho))$ on $((x_0, y_0), d_\rho)$. In particular for*

$$\rho_a(x) = \int_c^x \frac{1}{\sqrt{a(t)}} dt$$

(d_{ρ_a} is the metric associated with the carré-du-champs operator of the diffusion), if $C(\rho_a) < +\infty$, then μ satisfies $W_1I(C(\rho_a))$ on $((x_0, y_0), d_{\rho_a})$.

Remarks 3.9. The quantity $C(\rho)$ in (3.10) is not innocent: Chen-Wang's variational formula for the spectral gap λ_1 says that ([8, 23]): $\lambda_1 = \sup_{\rho} \frac{1}{C(\rho)}$.

4. FUNCTIONAL INEQUALITIES AND W_2I INEQUALITIES

Throughout this section we consider the framework of Section 2, i.e. \mathcal{X} is a connected complete Riemannian manifold M with $\mu(dx) := e^{-V(x)}dx/Z$ for some $V \in C(M)$ with $Z := \int_M e^{-V(x)}dx < \infty$. Recall that in [17, 13] was proven the fact that a logarithmic Sobolev inequality implies W_2I , and that (using HWI inequalities) under a lower bounded curvature, the converse was also true. We extend here this assertion for $\alpha - W_2I$ inequalities.

T4.1

Theorem 4.1. (1) *Let $\beta \in C([0, \infty))$ be increasing with $\beta(0) = 0$ such that*

$$\gamma(r) := \frac{1}{2} \int_0^r \frac{ds}{\sqrt{\beta(s)}} < \infty, \quad r > 0.$$

Then the following β -log-Sobolev inequality

$$\beta \circ \mu(g^2 \log g^2) \leq \mu(|\nabla g|^2), \quad g \in C_b^1(M), \mu(g^2) = 1, \quad (4.1) \quad \boxed{\text{P1}}$$

implies

$$\alpha(W_2(\nu, \mu)) \leq I(\nu|\mu), \quad \nu \in M_1(\mathcal{X}) \quad (4.2) \quad \boxed{\text{AA}}$$

for $\alpha(s) := \beta \circ \gamma^{-1}(s)$, $s \geq 0$.

(2) *Assume that $\text{Ric} + \text{Hess}_V \geq -K$ for some $K \geq 0$. Then (4.2) implies (4.1) for*

$$\beta(r) := \inf \{s > 0 : 2\sqrt{2s}\alpha^{-1}(s) + K(\alpha^{-1}(s))^2 \geq r\}, \quad r \geq 0.$$

Proof. (1) According to [24, Theorem 2.2], (4.1) implies

$$W_2(f\mu, \mu) \leq \gamma \circ \mu(f \log f), \quad f \geq 0, \mu(f) = 1. \quad (4.3) \quad \boxed{\text{TP}}$$

Then (4.2) follows from (4.3) and (4.1). For readers' convenience, we include below a brief proof of (4.3), inspired by the seminal work [17] pushed further in [22].

Since a continuous function can be uniformly approximated by smooth ones, we may and do assume that V is smooth. Let P_t be the diffusion semigroup generated by $\mathcal{L} := \Delta - \nabla V \cdot \nabla$. Then P_t is symmetric in $L^2(\mu)$. For fixed $f > 0$ with $\mu(f) = 1$, let $\mu_t = (P_t f)\mu$, $t > 0$. According to [22, page 176] for $p = 2$ (see also [17] under a curvature condition), we have

$$\frac{d^+}{dt} \{-W_2(\mu, \mu_t)\} := \limsup_{s \downarrow 0} \frac{W_2(\mu, \mu_t) - W_2(\mu, \mu_{t+s})}{s} \leq 2\mu(|\nabla \sqrt{P_t f}|^2)^{1/2}. \quad (4.4) \quad \boxed{\text{W2.1}}$$

Let

$$\gamma(r) = \frac{1}{2} \int_0^r \frac{ds}{\sqrt{\beta(s)}}, \quad r > 0.$$

It suffices to prove for the case that $\gamma(r) < \infty$ for $r > 0$. By (4.1) we have

$$\begin{aligned} \frac{d}{dt} \gamma \circ \mu(P_t f \log P_t f) &= 4\gamma' \circ \mu(P_t f \log P_t f) \mu(|\nabla \sqrt{P_t f}|^2) \\ &= \frac{2\mu(|\nabla \sqrt{P_t f}|^2)}{\sqrt{\beta \circ \mu(P_t f \log P_t f)}} \geq 2\mu(|\nabla \sqrt{P_t f}|^2)^{1/2}. \end{aligned}$$

Combining this with (4.4) we obtain

$$\frac{d^+}{dt} \{-W_2(\mu, \mu_t)\} \leq \frac{d}{dt} \gamma \circ \mu(P_t f \log P_t f),$$

which implies (4.3) by noting that $P_t f \rightarrow \mu(f) = 1$ as $t \rightarrow \infty$.

(2) By the HWI inequality (see [17, 3]), we have

$$\mu(g^2 \log g^2) \leq \frac{2(e^{2Kt} - 1)}{K} \mu(|\nabla g|^2) + \frac{K e^{2Kt}}{e^{2Kt} - 1} W_2(g^2 \mu, \mu)^2, \quad \mu(g^2) = 1, t > 0.$$

Combining this with (4.2) we obtain

$$\mu(g^2 \log g^2) \leq \inf_{t>0} \left\{ \frac{2(e^{2Kt} - 1)}{K} \mu(|\nabla g|^2) + \frac{K e^{2Kt}}{e^{2Kt} - 1} [\alpha^{-1}(\mu(|\nabla g|^2))]^2 \right\}.$$

Taking $t > 0$ such that

$$e^{2Kt} = 1 + K \frac{\alpha^{-1}(\mu(|\nabla g|^2))}{\sqrt{2\mu(|\nabla g|^2)}},$$

we obtain

$$\mu(g^2 \log g^2) \leq 2\sqrt{2\mu(|\nabla g|^2)} \alpha^{-1}(\mu(|\nabla g|^2)) + K[\alpha^{-1}(\mu(|\nabla g|^2))]^2.$$

This completes the proof. \square

Let us give a natural family of examples, namely when β is a power function.

C4.2 **Corollary 4.2.** *For any $\delta \in [1, 2)$,*

$$\mu(g^2 \log g^2)^\delta \leq C \mu(|\nabla g|^2), \quad g \in C_b^1(\mathcal{X}), \mu(g^2) = 1$$

implies

$$W_2(\nu, \mu)^2 \leq \frac{C^{2/\delta}}{(2-\delta)^2} I(\nu|\mu)^{(2-\delta)/\delta}.$$

Inversely if $\text{Ric} + \text{Hess}_V$ is bounded below, then

$$W_2(\nu, \mu)^2 \leq CI(\nu|\mu)^{(2-\delta)/\delta}$$

implies

$$\mu(g^2 \log g^2)^\delta \leq C' \mu(|\nabla g|^2), \quad g \in C_b^1(\mathcal{X}), \mu(g^2) = 1$$

for some $C' > 0$.

Proof. For $\beta(r) := r^\delta/C$ we have $\gamma(r) = \frac{\sqrt{C}}{2-\delta} r^{(2-\delta)/2}$ so that

$$\beta \circ \gamma^{-1}(s) = \frac{1}{C} \left(\frac{2-\delta}{\sqrt{C}} s \right)^{2\delta/(2-\delta)} = \frac{(2-\delta)^{2\delta/(2-\delta)}}{C^{2/(2-\delta)}} s^{2\delta/(2-\delta)}.$$

Then the first assertion follows from Theorem 4.1(1).

Next, for $\alpha(r) = r^{2\delta/(2-\delta)} C^{-\delta/(2-\delta)}$, we have $\alpha^{-1}(s) = \sqrt{C} s^{(2-\delta)/2\delta}$. Since $2-\delta \leq 1$, Theorem 4.1(2) implies

$$\mu(g^2 \log g^2) \leq 2\sqrt{2C} \mu(|\nabla g|^2)^{1/\delta} + KC \mu(|\nabla g|^2)^{(2-\delta)/\delta}, \quad \mu(g^2) = 1. \quad (4.5) \quad \boxed{W*}$$

Since $2-\delta \leq 1$, this implies

$$\mu(g^2 \log g^2) \leq C' \mu(|\nabla g|^2)^{1/\delta}, \quad \mu(g^2) = 1, \mu(|\nabla g|^2) \geq 1 \quad (4.6) \quad \boxed{W*2}$$

for some $C' > 0$. Moreover, since $\delta \geq 1$, (4.5) implies the defective log-Sobolev inequality

$$\mu(g^2 \log g^2) \leq C_1 \mu(|\nabla g|^2) + C_2, \quad \mu(g^2) = 1$$

for some $C_1, C_2 > 0$, which in particular implies that the spectrum of $L := \Delta + \nabla V$ is discrete (see e.g. [21, 27]), and hence the Poincaré inequality holds since $\lambda_0 = 0$ is the simple eigenvalue due to the connection of the manifold. Thus, the strict log-Sobolev inequality

$$\mu(g^2 \log g^2) \leq C' \mu(|\nabla g|^2), \quad \mu(g^2) = 1$$

for some constant $C' > 0$. The proof is then completed by combining this with (4.6). \square

Example 4.3. Let Ric be bounded below, and ρ_o the Riemannian distance function to a fixed point $o \in E$. Let $V \in C(\mathcal{X})$ such that $V - a\rho_o^\theta$ is bounded for some $a > 0$ and $\theta \geq 2$. Then (4.2) holds for $\alpha(r) = Cr^{2(\theta-1)}$ for some $C > 0$, i.e.

$$CW_2(\nu, \mu)^{2(\theta-1)} \leq I(\nu|\mu), \quad \nu \in M_1(\mathcal{X}). \quad (4.7) \quad \boxed{W*3}$$

The power $2(\theta-1)$ is sharp, i.e. the above inequality does not hold if this power is replaced by any larger number, as seen from Proposition 2.3.

Indeed, by [21, Corollaries 2.5 and 3.3], we have

$$\mu(g^2 \log^{2(\theta-1)/\theta}(g^2 + 1)) \leq C_1 \mu(|\nabla g|^2) + C_2, \quad \mu(g^2) = 1$$

holds for some $C_1, C_2 > 0$. By Jensen's inequality we obtain

$$\mu(g^2 \log g^2)^{2(\theta-1)/\theta} \leq \mu(g^2 \log(g^2 + 1))^{2(\theta-1)/\theta} \leq C_1 \mu(|\nabla g|^2) + C_2, \quad \mu(g^2) = 1.$$

Combining this with the log-Sobolev inequality as in the proof of Corollary 4.2, we obtain

$$\mu(g^2 \log g^2)^{2(\theta-1)/\theta} \leq \mu(g^2 \log(g^2 + 1))^{2(\theta-1)/\theta} \leq C' \mu(|\nabla g|^2), \quad \mu(g^2) = 1$$

for some constant $C' > 0$. According to Corollary 4.2, this implies (4.7).

5. Φ -SOBOLEV INEQUALITY AND CONCENTRATION INEQUALITY
FOR UNBOUNDED OBSERVABLES UNDER INTEGRABILITY CONDITION

Let $\Phi : \mathbb{R}^+ \rightarrow [0, +\infty]$ be a convex, increasing and left continuous function with $\Phi(0) = 0$, such that

$$\lim_{r \rightarrow +\infty} \frac{\Phi(r)}{r} = +\infty. \quad (5.1) \quad \boxed{51}$$

Consider the Orlicz space $L^\Phi(\mu)$ of those measurable functions g on \mathcal{X} so that its gauge norm

$$N_\Phi(g) := \inf\{c > 0; \int \Phi(|g|/c) d\mu \leq 1\}$$

is finite, where the convention $\inf \emptyset := +\infty$ is used. The Orlicz norm of g is defined by

$$\|g\|_\Phi := \sup\left\{\int gu d\mu; N_\Psi(u) \leq 1\right\}$$

where

$$\Psi(r) := \sup_{\lambda \geq 0} (\lambda r - \Phi(\lambda)), \quad r \geq 0 \quad (5.2)$$

is the convex conjugation of Φ . The so called (defective) Φ -Sobolev inequality says that for some two nonnegative constants $C_1, C_2 \geq 0$

$$\|g^2\|_\Phi \leq C_1 \mathcal{E}(g, g) + C_2 \mu(g^2), \quad \forall g \in \mathbb{D}(\mathcal{E}), \mu(g^2) = 1. \quad (5.3) \quad \boxed{52}$$

Under the assumption of the Poincaré inequality with the best constant C_P , (5.3) can be transformed into the following tight version

$$\|(g - \mu(g))^2\|_\Phi \leq (C_1 + C_2 C_P) \mathcal{E}(g, g), \quad \forall g \in \mathbb{D}(\mathcal{E}) \quad (5.4) \quad \boxed{53}$$

called sometimes Orlicz-Poincaré inequality.

thm51

Theorem 5.1. *Assume the Φ -Sobolev inequality (5.3) and the Poincaré inequality with constant C_P . Then*

(a): *for any μ -probability density f ,*

$$\|f - 1\|_\Phi \leq \sqrt{C'_1 I(f\mu|\mu)^2 + C'_2 I(f\mu|\mu)} \quad (5.5) \quad \boxed{\text{thm51a}}$$

where $C'_1 = (C_1 + 2C_2 C_P)C_1$, $C'_2 = (C_1 + 2C_2 C_P) \cdot 4C_2$; or equivalently for any observable $u \in L^\Psi(\mu)$ (Ψ being the convex conjugation of Φ given above) so that $N_\Psi(u) \leq 1$ and for all $t, r > 0$,

$$\mathbb{P}_\nu \left(\frac{1}{t} \int_0^t u(X_s) ds > \mu(u) + r \right) \leq \left\| \frac{d\nu}{d\mu} \right\|_2 \exp \left(-t \cdot \frac{\sqrt{4C'_1 r^2 + (C'_2)^2} - C'_2}{2C'_1} \right). \quad (5.6) \quad \boxed{\text{thm51b}}$$

(b): *for any μ -probability density f ,*

$$\sup_{u \in b\mathcal{B}: N_\Psi(u^2) \leq 1} \int (f - 1) u d\mu \leq \sqrt{2(C_1 + 4C_2 C_P) I(f\mu|\mu)} \quad (5.7) \quad \boxed{\text{thm51c}}$$

or equivalently for any $u \in L^1(\mu)$ such that $u^2 \in L^\Psi(\mu)$,

$$\mathbb{P}_\nu \left(\frac{1}{t} \int_0^t u(X_s) ds > \mu(u) + r \right) \leq \left\| \frac{d\nu}{d\mu} \right\|_2 \exp \left(-t \frac{r^2}{2(C_1 + 4C_2 C_P) \|u^2\|_\Psi} \right), \quad \forall t, r > 0. \quad (5.8) \quad \boxed{\text{thm51d}}$$

(c): More generally for any $p \in [1, +\infty)$, there is a constant $\kappa > 0$ depending only of p, C_1, C_2, C_P such that for any μ -probability density f ,

$$\alpha_p \left(\sup_{u \in b\mathcal{B}: N_\Psi(u^2) \leq 1} \int (f - 1) u d\mu \right) \leq I(f\mu|\mu) \quad (5.9) \quad \boxed{\text{thm51e}}$$

where $\alpha_p(r) = (1 + r^2/\kappa)^{p/2} - 1$; or equivalently for any $u \in L^1(\mu)$ such that $N_\Psi(|u|^p) \leq 1$,

$$\mathbb{P}_\nu \left(\frac{1}{t} \int_0^t u(X_s) ds > \mu(u) + r \right) \leq \left\| \frac{d\nu}{d\mu} \right\|_2 \exp \left(-t [1 + r^2/\kappa]^{p/2} - 1 \right), \quad \forall t, r > 0. \quad (5.10) \quad \boxed{\text{thm51f}}$$

As there are numerous practical criteria for the Φ -Sobolev inequality (see e.g. [8, 14, 23]), this theorem is very useful and gives different concentration behaviors for $\frac{1}{t} \int_0^t u(X_s) ds$, according to the integrability condition $|u|^p \in L^\Psi(\mu)$ where $p \in [1, +\infty)$.

This result generalizes the sharp concentration inequality under the log-Sobolev inequality in Wu [26]. For applications of Φ -Sobolev inequalities in large deviations see Wu and Yao [29].

Remarks 5.2. As the l.h.s. of (5.5), (5.7) and (5.9) are the transportation cost $T_{\mathcal{V}}(f\mu, \mu)$, with $\mathcal{V} = \{(u, u); u \in b\mathcal{B}, N_\Psi(|u|^p) \leq 1\}$, $p = 1, 2, p \geq 1$ respectively, so they are the transportation-information inequality. In this point of view, the equivalence between (5.5) and (5.6) in part (a), that between (5.7) and (5.8) in part (b) and that between (5.9) and (5.10) in part (c) are all immediate from Theorem 1.1 (the passage from bounded u to general u in the concentration inequalities (5.6), (5.8) and (5.10) can be realized easily by dominated convergence).

Remarks 5.3. The concentration inequalities (5.6), (5.8) and (5.10) are all sharp in order. Indeed consider the Ornstein-Uhlenbeck process on \mathbb{R} generated by $\mathcal{L}f = f'' - xf'$: the Φ -Sobolev inequality (X_t) holds with $\Phi(r) = (1+r) \log(1+r)$ and $\mu = \mathcal{N}(0, 1)$. Consider $u(x) := |x|^{2/p}$ where $p \geq 1$. Then $u^p \in L^\Psi(\mu)$, and $\frac{1}{t} \int_0^t u(X_s) ds = \frac{1}{t} \int_0^t |X_s|^p ds$ possess exactly the concentration behaviors exhibited by the r.h.s. of (5.10) for large deviation value r , and for small deviation value r of order $1/\sqrt{t}$ if t is large enough (by the central limit theorem).

Proof of Theorem 5.1. As explained in the previous remarks, (5.6) (resp. (5.8); (5.10)) is equivalent to (5.5) (resp. (5.7), (5.9)), all by Theorem 1.1.

It is not surprising that the proof relies on the ideas first used in [5], establishing criterions for W_1H under integrability criteria. Note also that the reader may easily adapt the proof to use conditions on F -Sobolev inequalities (equivalent to some Orlicz-Poincaré inequality) and integrability on u (rather than Orlicz norm of u).

(a) For (5.5) we may assume that $I(f\mu|\mu)$ is finite, i.e., $\sqrt{f} \in \mathbb{D}(\mathcal{E})$ (and then $I(f\mu|\mu) = \mathcal{E}(\sqrt{f}, \sqrt{f})$). For any $u \in L^\Psi(\mu)$ with $N_\Psi(u) \leq 1$, we have by Cauchy-Schwartz

$$\begin{aligned} \int |(f-1)u|d\mu &= \int |\sqrt{f}-1|(\sqrt{f}+1)|u|d\mu \\ &\leq \sqrt{\int (\sqrt{f}-1)^2|u|d\mu} \sqrt{\int (\sqrt{f}+1)^2|u|d\mu} \\ &\leq \sqrt{\|(\sqrt{f}-1)^2\|_\Phi \|(\sqrt{f}+1)^2\|_\Phi} \end{aligned}$$

But by the assumed Φ -Sobolev inequality (5.3),

$$\|(\sqrt{f}-1)^2\|_\Phi \leq C_1\mathcal{E}(\sqrt{f}, \sqrt{f}) + C_2 \int (\sqrt{f}-1)^2d\mu$$

and $\int (\sqrt{f}-1)^2d\mu = 2(1-\mu(\sqrt{f})) \leq 2\text{Var}_\mu(\sqrt{f}) \leq 2C_P\mathcal{E}(\sqrt{f}, \sqrt{f})$; moreover

$$\|(\sqrt{f}+1)^2\|_\Phi \leq C_1\mathcal{E}(\sqrt{f}, \sqrt{f}) + C_2 \int (\sqrt{f}+1)^2d\mu$$

and $\int (\sqrt{f}+1)^2d\mu \leq 4$. Thus we get

$$\int |(f-1)u|d\mu \leq \sqrt{(C_1+2C_2C_P)\mathcal{E}(\sqrt{f}, \sqrt{f})(C_1\mathcal{E}(\sqrt{f}, \sqrt{f})+4C_2)}$$

where (5.5) follows by recalling $I(f\mu|\mu) = \mathcal{E}(\sqrt{f}, \sqrt{f})$.

(b) For any u so that $N_\Psi(u^2) \leq 1$ we use now differently Cauchy-Schwartz inequality to get:

$$\int |(f-1)u|d\mu \leq \sqrt{\int (\sqrt{f}-1)^2d\mu} \sqrt{\int (\sqrt{f}+1)^2u^2d\mu}$$

But as noticed in the proof of (a),

$$\int (\sqrt{f}-1)^2d\mu \leq 2\text{Var}_\mu(\sqrt{f}) \leq 2\min\{C_P I(f\mu|\mu), 1\}$$

and

$$\int (\sqrt{f}+1)^2u^2d\mu \leq \|(\sqrt{f}+1)^2\|_\Phi \leq C_1 I(f\mu|\mu) + 4C_2.$$

Plugging those two estimates into the previous inequality we get (5.7).

(c). Letting $q := p/(p-1)$ we have by Hölder's inequality,

$$\int |(f-1)u|d\mu \leq (\mu(|f-1|))^{1/q} \left(\int |f-1||u|^pd\mu \right)^{1/p}$$

Note that $\mu(|f-1|) \leq 2$ and by [13, Theorem 3.3],

$$(\mu(|f-1|))^2 \leq 4\text{Var}_\mu(\sqrt{f}) \leq 4C_P I$$

where $I := I(f\mu|\mu)$. On the other hand by part (a),

$$\int |f-1||u|^pd\mu \leq \|f-1\|_\Phi \leq \sqrt{C'_1 I^2 + C'_2 I}.$$

Substituting those estimates into the first inequality we get

$$\begin{aligned} \left(\int |(f-1)u|d\mu \right)^2 &\leq (\max\{4, 4C_P I\})^{1/q} (C'_1 I^2 + C'_2 I)^{1/p} \\ &\leq \begin{cases} 4^{1/q} (C'_1 + C'_2 C_P)^{1/p} \cdot I^{2/p}, & \text{if } C_P I \geq 1; \\ 4^{1/q} (C'_1 + C'_2 C_P)^{1/p} C_P^{(p-2)/p} \cdot I, & \text{otherwise.} \end{cases} \end{aligned}$$

The last term is less than $\kappa[(1 + I)^{2/p} - 1]$ for some constant $\kappa > 0$. That yields to (5.9). \square

Let us finally relate previous inequalities to usual $\alpha - WI$ inequalities.

cor51

Corollary 5.4. *Assume the Φ -Sobolev inequality (5.3) and the Poincaré inequality. Assume that $d^p(x, x_0) \in L^\Psi(\mu)$ for some $p \geq 1$ where Ψ is the convex conjugation of Φ . Then there are positive constants C'_1, C'_2 and κ such that for all $\nu \in \mathcal{M}_1(\mathcal{X})$,*

$$W_p^p(\nu, \mu) \leq \sqrt{C'_1 I(\nu, \mu)^2 + C'_2 I(\nu|\mu)},$$

and

$$\kappa \left([1 + W_1(\nu, \mu)^2]^{p/2} - 1 \right) \leq I(\nu|\mu)$$

and when $p \geq 2$,

$$\kappa \left([1 + W_2(\nu, \mu)^4]^{p/4} - 1 \right) \leq I(\nu|\mu).$$

Proof. Recall the following fact ([19, Proposition 7.10]),

$$W_p^p(\nu, \mu) \leq 2^{p-1} \|d(\cdot, x_0)^p(\nu - \mu)\|_{TV}.$$

Then this corollary follows directly from Theorem 5.1. \square

REFERENCES

- [1] F. Barthe, and D. Bakry and P. Cattiaux and A. Guillin. A simple proof of the Poincaré inequality for a large class of probability measure including the logconcave case. *Electronic Communications in Probability*, 13:60–66, 2008.
- [2] D. Bakry and M. Emery. Diffusions hypercontractives. In *Séminaire de Probabilités*, number 1123 in Lecture Notes in Mathematics. Springer-Verlag, 1985.
- [3] S.G. Bobkov, I. Gentil, and M. Ledoux. Hypercontractivity of Hamilton-Jacobi equations. *Journal de Mathématiques Pures et Appliquées*, 80(7):669–696, 2001.
- [4] S.G. Bobkov and F. Götze. Exponential integrability and transportation cost related to logarithmic Sobolev inequalities. *Journal of Functional Analysis*, 163:1–28, 1999.
- [5] F. Bolley and C. Villani. Weighted Csiszar-Kullback-Pinsker inequalities and applications to transportation inequalities *Ann. Math. Fac. Toulouse* 14(6):331–352, 2005,
- [6] P. Cattiaux and A. Guillin. Deviation bounds for additive functionals of Markov process. *ESAIM P&S*, 12:12–29, 2008.
- [7] P. Cattiaux and A. Guillin. On quadratic transportation cost inequalities. *Journal de Mathématiques Pures et Appliquées*, 86:342–361, 2006.
- [8] M.F. Chen. *Eigenvalues, inequalities, and ergodic theory*. Probability and its Applications. Springer-Verlag, 2005.
- [9] J-D. Deuschel and D.W. Stroock. *Large Deviations*, volume 137 of *Pure and Applied Mathematics*. Academic Press, 1989.
- [10] H. Djellout, A. Guillin, and L. Wu. Transportation cost-information inequalities for random dynamical systems and diffusions. *Annals of Probability*, 32(3B):2702–2732, 2004.
- [11] H. Djellout and L. Wu. Spectral gap of one dimensional diffusions in Lipschitzian norm and application to log-Sobolev inequalities for Gibbs measures. In preparation.
- [12] N. Gozlan and C. Léonard. A large deviation approach to some transportation cost inequalities. *Probability Theory and Related Fields*, 1-2:235–283, 2007.
- [13] A. Guillin, Ch. Léonard, L. Wu and N. Yao, Transport-information inequalities for Markov processes (I), Preprint 2007.

- [14] M. Ledoux. *The Concentration of Measure Phenomenon*. Mathematical Surveys and Monographs 89. American Mathematical Society, Providence RI, 2001.
- [15] W. Liu and Y. Ma, Spectral gap and deviation inequalities for birth-death processes, *Preprint 06, contained in the Ph.D thesis of Y. Ma at Université La Rochelle 2007*.
- [16] K. Marton. Bounding \bar{d} -distance by informational divergence: a way to prove measure concentration. *Annals of Probability*, 24:857–866, 1996.
- [17] F. Otto and C. Villani. Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. *Journal of Functional Analysis*, 173:361–400, 2000.
- [18] M. Talagrand. Transportation cost for gaussian and other product measures. *Geometric and Functional Analysis*, 6:587–600, 1996.
- [19] C. Villani. *Topics in Optimal Transportation*. Graduate Studies in Mathematics 58. American Mathematical Society, Providence RI, 2003.
- [20] C. Villani. Saint-Flour Lecture Notes, 2005. Optimal transport, old and new. Available online via <http://www.umpa.ens-lyon.fr/~cvillani/>.
- [21] F.-Y. Wang, *Functional inequalities for empty essential spectrum*, *J. Funct. Anal.* 170(2000), 219–245.
- [22] F.-Y. Wang, *Probability distance inequalities on Riemannian manifolds and path spaces*, *J. Funct. Anal.* 206(2004), 167–190.
- [23] F.Y. Wang, *Functional inequalities, Markov Semigroup and Spectral Theory*, Chinese Sciences Press, Beijing/New York, 2005.
- [24] F.Y. Wang, *Transportation cost inequalities for the log-Sobolev inequality with powers*, 2007 preprint.
- [25] L. Wu, Uniqueness of Nelson’s diffusions, *Probability Theory Rel. Fields* **114**, 549-585 (1999)
- [26] L. Wu. A deviation inequality for non-reversible Markov processes. *Ann. Inst. Henri Poincaré (série Probab. Stat.)*, 36:435–445, 2000.
- [27] L. Wu. Uniformly integrable operators and large deviations for Markov processes. *J. Funct. Anal.*, 172:301–376, 2000.
- [28] L. Wu, Essential spectral radius for Markov semigroups (I) : discrete time case, *Probab. Th. Rel. Fields* **128**, 255-321 (2004)
- [29] L. Wu and N. Yao, Large deviations for Markov processes via Φ -Sobolev inequalities, *Elect. Comm. Probab.*, 13:10–23, 2008.
- [30] L. Wu, Lipschitzian norm of Poisson operators on Riemannian manifolds. *preprint 2009*

ARNAUD GUILLIN. ECOLE CENTRALE DE MARSEILLE ET LATP, CENTRE DE MATHÉMATIQUES ET INFORMATIQUE. TECHNOPÔLE DE CHÂTEAU-GOMBERT. 13453 MARSEILLE, FRANCE
E-mail address: guillin@cmi.univ-mrs.fr

CHRISTIAN LÉONARD. MODAL-X, UNIVERSITÉ PARIS 10. BÂTIMENT G, 200 AVENUE DE LA RÉPUBLIQUE.92001 NANTERRE, FRANCE AND CMAP, ÉCOLE POLYTECHNIQUE. 91128 PALAISEAU, FRANCE
E-mail address: christian.leonard@polytechnique.fr

FENG YU WANG. DEPARTMENT OF MATHEMATICS, SWANSEA UNIVERSITY, SINGLETON PARK, SA2 8PP, SWANSEA, UK
E-mail address: f.y.wang@swansea.ac.uk

LIMING WU. LABORATOIRE DE MATHÉMATIQUES APPLIQUÉES, CNRS-UMR 6620, UNIVERSITÉ BLAISE PASCAL, 63177 AUBIÈRE, FRANCE. AND DEPARTMENT OF MATHEMATICS, WUHAN UNIVERSITY, 430072 HUBEI, CHINA
E-mail address: Li-Ming.Wu@math.univ-bpclermont.fr