Gong, F.-Z., Liu, Y. and Wen, Z.-Y.
Osaka J. Math.
50 (2013), 491–502

RICCI CURVATURE OF MARKOV CHAINS ON POLISH SPACES REVISITED

FU-ZHOU GONG, YUAN LIU and ZHI-YING WEN

(Received December 1, 2010, revised September 26, 2011)

Abstract

Recently, Y. Ollivier defined the Ricci curvature of Markov chains on Polish spaces via the contractivity of transition kernels under the L^1 Wasserstein metric. In this paper, we will discuss further the spectral gap, entropy decay, and logarithmic Sobolev inequality for the λ-range gradient operator. As an application, given resistance forms (i.e. symmetric Dirichlet forms with finite effective resistance) on fractals, we can construct Markov chains with positive Ricci curvature, which yields the Gaussian-then-exponential concentration of invariant distributions for Lipschitz test functions.

1. Introduction

Recently, Ollivier [9] introduced the *Ricci curvature* of Markov chains on Polish spaces, which was characterized by the contractivity of transition probabilities under the L^1 Wasserstein metric. In this paper, we will discuss and improve some results about the spectral gap, entropy decay, logarithmic Sobolev inequality (LSI in short) for the λ-*range gradient operator*, and related topics.

Let (X, d) be a Polish space endowed with a *random walk* $m = \{m_x\}_{x\in X}$, i.e. a family of Borel probability measures on X. Denote by Lip_1 the set of 1-Lipschitz functions, $\mathcal{M}_1$ the set of Borel probability measures, and $\mathcal{C}(\mu, \nu)$ the set of *couplings* of μ and ν (i.e. all joint distributions on $X \times X$ with marginals μ and ν). Suppose m_x depends on x measurably, and has a finite first moment (i.e. $\int d(o, y)\,\mathrm{d}m_x(y) < \infty$ for some $o \in X$). Define the L^1 *Wasserstein metric* (or *transportation distance*) between m_x and m_y as

$$W_1(m_x, m_y) := \inf_{\pi\in\mathcal{C}(m_x,m_y)} \int_{X\times X} d(\xi, \eta)\,\mathrm{d}\pi(\xi, \eta).$$

Then, $(\mathcal{M}_1, W_1)$ is a complete metric space (for example, see Villani [12]). Equivalently, the Kantorovich dual theorem reads (see also [12])

$$W_1(m_x, m_y) = \sup_{f\in\mathrm{Lip}_1} \left| \int f\,\mathrm{d}m_x - \int f\,\mathrm{d}m_y \right|. \tag{1.1}$$

2000 Mathematics Subject Classification. Primary 60J10, 47D07; Secondary 28A80, 31C25.

According to [9] Definition 3, define the Ricci curvature of (X, d, m) as

$$\kappa(x, y) := 1 - \frac{W_1(m_x, m_y)}{d(x, y)}, \quad \forall x \neq y. \tag{1.2}$$

Suppose $\kappa(x, y) \geqslant \kappa \geqslant 0$ throughout this paper. Denote by δ_x the Dirac measure at x and $\mu * m(\cdot) = \int m_x(\cdot)\, \mathrm{d}\mu(x)$, when $\kappa > 0$, (1.2) implies

$$W_1(\delta_x * m, \delta_y * m) = W_1(m_x, m_y) \leqslant (1 - \kappa)\, d(x, y) = (1 - \kappa) W_1(\delta_x, \delta_y),$$

which means the transition $*m$ is a strictly contractive map onto $(\mathcal{M}_1, W_1)$ with the factor $1 - \kappa$. So there exists a unique probability measure ν such that $\nu * m = \nu$, which is called an *invariant distribution*. By convention, denote by $(\cdot, \cdot)$ the inner product of $L^2(X, \nu)$, $\|\cdot\|_2$ the L^2-norm, and νf the expectation of f.

Let's give an overview of main results in this paper as follows.

Invariant distribution for nonnegative curvature. In Section 2, we will discuss the existence and uniqueness of the invariant distribution for the critical case $\kappa = 0$, which means $*m$ is non-expansive on $(\mathcal{M}_1, W_1)$. Define the *averaging operator*

$$\mathrm{M} f(x) = \int f \,\mathrm{d}m_x,$$

transition kernel $\mathrm{M}(x, A) := \mathrm{M}\mathbf{1}_A(x) = m_x(A)$, and denote by M^n the n-step transition kernel. According to Szarek [11], we have

Proposition 1.1. *Suppose $\kappa = 0$. Let $\overline{\mathrm{M}}_N = (1/N) \sum_{1 \leqslant n \leqslant N} \mathrm{M}^n$, suppose also*
($\mathfrak{C}$) $\exists z, \forall \delta > 0, \exists x,$ *s.t.* $\limsup_{N\to\infty} \overline{\mathrm{M}}_N(x, B(z, \delta)) > 0$.
Then $\{\mathrm{M}^n(z, \cdot)\}_{n\geqslant 1}$ is tight and there exists an invariant distribution. Moreover, if $\kappa(x, y) > 0$ everywhere, the invariant distribution is unique.

REMARK 1.1. ($\mathfrak{C}$) is necessary to the existence of invariant distributions.

Spectral gap. Spectral gap will be revisited in Section 3. Define the *diffusion constant* at x as $\sigma(x) = \big((1/2)\iint d(y,z)^2 \,\mathrm{d}m_x(y)\,\mathrm{d}m_x(z)\big)^{1/2}$, and *average diffusion constant* $\sigma = \|\sigma(x)\|_2$. According to [9] Proposition 30, assume $\sigma < \infty$ and either M is self-adjoint or X is finite, then the spectral radius of M on $L^2(X, \nu)/\{\text{const}\}$ is less than $1 - \kappa$. So the spectral gap of $\mathrm{Id} - \mathrm{M}$ in $L^2(X, \nu)$ is at least κ.

Fu-Zhou Gong and Li-Ming Wu give a counterexample (which will be mentioned later) to say that, if M is non self-adjoint, there may be no spectral gap in $L^2(X, \nu)$. If M is self-adjoint, we can remove the original assumption $\sigma < \infty$ in [9] to show that the spectral radius of M is strictly less than 1 yet.

Theorem 1.2. *Suppose* $\kappa > 0$ *and* M *is self-adjoint. Then the spectral radius of* M *on* $L^2(X, \nu)/\{const\}$ *is at most* $\sqrt{1-\kappa}$.

In [9] Section 3.3.6, there defines a reversible random walk on $\mathbb{N}$ with two parameters a and b, which admits positive Ricci curvature and a unique invariant distribution ν. Letting $a = b$ gives $\sigma = \infty$. We point out, from the above theorem, the spectral gap of $\mathrm{Id} - \mathrm{M}$ in $L^2(X, \nu)$ does exist, and then the Poincaré inequality holds. According to Aida and Stroock [1] or Ledoux [7], one might look forward to proving the exponential concentration of ν for Lipschitz test functions. However, for this moment, ν has a heavy tail, which means there is no exponential concentration at all. Hence, this example shows us that $\sigma < \infty$ is essentially necessary to derive the exponential concentration for Lipschitz test functions from the Poincaré inequality.

Entropy decay. Section 4 will give an entropy-variance inequality, which implies $\mathrm{Ent}_\nu(\mathrm{M}^n f)$ has an exponential decay by using Proposition 1.2.

Proposition 1.3. *There exists a constant* C_1 *not greater than* $2(2 + \log 2)$ *such that for positive* $f \in L^2(X, \nu)$

$$\mathrm{Ent}_\nu\, f \leqslant C_1 \sqrt{\mathrm{Var}_\nu\, f}.$$

Suppose $\kappa > 0$ *and* M *is self-adjoint, then for any* $t \in \mathbb{N}$

$$\mathrm{Ent}_\nu\, \mathrm{M}^t f \leqslant C_1 \sqrt{\mathrm{Var}_\nu\, f}\,(1-\kappa)^{t/2}.$$

REMARK 1.2. From the inequality $a \log a \leqslant a^2 - a$, it follows $\mathrm{Ent}_\nu\, f \leqslant \mathrm{Var}_\nu\, f/(\nu f)$. The right-hand fraction can be much bigger than $\sqrt{\mathrm{Var}_\nu\, f}$.

Define the Dirichlet form $\mathcal{E}^*(f, g) = (f, (\mathrm{Id} - \mathrm{M}^2)g)$, which satisfies a modified LSI if m_x is absolutely continuous to ν.

Proposition 1.4. *Suppose* $\kappa > 0$ *and* M *is self-adjoint. Suppose* $\mathrm{d}m_x(y) = p(x, y)\,\mathrm{d}\nu(y)$ *with* $\|p(x, \cdot)\|_2 \leqslant C_2$ *for all* x. *Then for positive* $f \in L^2(X, \nu)$

$$\mathrm{Ent}_\nu\, f \leqslant \mathcal{E}^*(f, \log f) + \frac{2C_1C_2}{1-\sqrt{1-\kappa}}\nu f.$$

In fact, it is hopeless to obtain the standard LSI $\mathrm{Ent}_\nu(f^2) \leqslant C\mathcal{E}^*(f, f)$, since $\mathcal{E}^*(f, f) \leqslant \mathrm{Var}_\nu\, f$ by the Hölder inequality.

LSI for λ-range gradient operator. Section 5 will be devoted to the standard LSI for the λ-range gradient operator D. According to [9] Section 4, define

$$\mathrm{D} f(x) := \sup_{y,y'\in X} \frac{|f(y)-f(y')|}{d(y,y')} e^{-\lambda\, d(x,y)-\lambda\, d(y,y')}. \tag{1.3}$$

Assume $\sigma_\infty := \sup_x (1/2)\,\mathrm{diam}(\mathrm{Supp}\, m_x) < \infty$. Then a modified LSI holds, i.e. there exists some $\lambda > 0$ such that for positive f

$$\mathrm{Ent}_\nu\, f := \nu\left(f \cdot \log \frac{f}{\nu f}\right) \leqslant \left(\sup_x \frac{4\sigma(x)^2}{\kappa n_x}\right) \int \frac{(\mathrm{D} f)^2}{f}\, \mathrm{d}\nu, \tag{1.4}$$

where define the *local dimension* $n_x = \sigma(x)^2/\sup\{\mathrm{Var}_{m_x} f : f \in \mathrm{Lip}_1\}$ and the *variance* $\mathrm{Var}_{m_x} f = (1/2)\iint |f(y)-f(z)|^2\, \mathrm{d}m_x(y)\, \mathrm{d}m_x(z)$. In fact, $0 < \lambda \leqslant 1/(20\sigma_\infty(1+U))$, where U takes the supremum of *unstability*, which will be mentioned later.

Theorem 1.5. *Suppose $\kappa > 0$ and $\sigma_\infty < \infty$. Then for $0 < \lambda \leqslant 1/(80\sigma_\infty(1+U))$*

$$\mathrm{Ent}_\nu(f^2) \leqslant \frac{32}{\kappa\lambda} \sup_x \sqrt{\frac{\sigma(x)^2}{n_x}} \int (\mathrm{D} f)^2\, \mathrm{d}\nu.$$

In general, there is no way to define a gradient-type symmetric Dirichlet form $\mathcal{E}$ such that $\mathcal{E}(f,f) \simeq \int (\mathrm{D} f)^2\, \mathrm{d}\nu$. Otherwise, the above standard LSI yields the Gaussian concentration for Lipschitz test functions (see [1]), which contradicts the case that ν is allowed to be Poisson-like.

Ricci curvature for resistance forms. In Section 6, let's consider a probability space (X,μ) equipped with a resistance form $(\mathcal{E},\mathcal{F})$, namely *symmetric Dirichlet form* with finite *effective resistance* $R(x,y)$, see Kigami [5]. Moreover, suppose $(\mathcal{E},\mathcal{F})$ is conservative.

Proposition 1.6. *Let $\{G_\alpha\}_{\alpha>0}$ be the resolvent operator family associated to $(\mathcal{E},\mathcal{F})$, and m a random walk with its average operator $\mathrm{M} = \alpha G_\alpha$. Then, μ is invariant to m, and $(X,\sqrt{R},m)$ has a positive Ricci curvature at least $\kappa > 0$ provided that $2\alpha \int R(o,x)\, \mathrm{d}\mu(x) \leqslant (1-\kappa)^2$ for some $o \in X$. Moreover, if $R(o,x)$ is uniformly bounded, μ satisfies the Gaussian-then-exponential concentration for Lipschitz test functions by* [9] Theorem 33.

REMARK 1.3. Equip the real line $\mathbb{R}$ with a probability measure $\mathrm{d}\mu(x) = C\exp(-c|x|^\alpha)\, \mathrm{d}x$ for $\alpha > 2$. Define $\mathcal{E}(f,f) = (1/2)\int |f'|^2\, \mathrm{d}\mu$. Then $R(0,x)$ is not uniformly bounded, but satisfies $\int R(0,x)\, \mathrm{d}\mu(x) < \infty$.

For a basic theory of Dirichlet forms, we refer to Fukushima, Oshima and Takeda [4], or Ma and Röckner [8]. Note that for resistance forms on fractal sets (for example, the Sirpiński gasket), writing $\mathcal{G}$ to be the collection of Lipschitz functions under the metric $\sqrt{R(x, y)}$, there usually occurs $\mathcal{F} \subsetneq \mathcal{G}$.

Finally, we would like to mention that, all above functional inequalities use a unified approach in the viewpoint of Ricci curvature from [9], but might be not sharp for concrete models of Markov chains.

2. Invariant distribution for nonnegative curvature

Suppose $\kappa \geqslant 0$, (1.1) and (1.2) imply that for any $n \geqslant 1$ and $f \in \mathrm{Lip}_1$

$$|\mathrm{M}^n f(x) - \mathrm{M}^n f(y)| \leqslant d(x, y), \quad \forall x, y \in X.$$

So $\{\mathrm{M}^n f\}_{n\geqslant 1}$ is a (uniformly) equicontinuous family on X.

Now, let's prove Proposition 1.1.

Proof of Proposition 1.1. If $(\mathfrak{C})$ holds, the tightness of $\{\mathrm{M}^n(z, \cdot)\}_{n\geqslant 1}$ and existence of invariant distributions directly follow from [11] Proposition 2.1 due to equicontinuity.

Moreover, suppose $\kappa(x, y) > 0$ everywhere. Let Ξ be an optimal coupling of tow distinct $\mu_1, \mu_2 \in \mathcal{M}_1$. Similar to [9] Proposition 20, let $\xi_{x,y}$ be an optimal coupling of m_x and m_y, depending on (x, y) measurably. Then $\int \xi_{x,y} \mathrm{d}\Xi(x, y)$ is a coupling of $\mu_1 * m$ and $\mu_2 * m$, which yields

$$\begin{aligned} W_1(\mu_1 * m, \mu_2 * m) &\leqslant \int d(x, y)\, \mathrm{d}\Big\{\int \xi_{x',y'}\, \mathrm{d}\Xi(x', y')\Big\}(x, y) \\ &= \iint d(x, y)\, \mathrm{d}\xi_{x',y'}(x, y)\, \mathrm{d}\Xi(x', y') \\ &= \int d(x', y')(1 - \kappa(x', y'))\, \mathrm{d}\Xi(x', y') < W_1(\mu_1, \mu_2). \end{aligned}$$

It follows $\mu_1 = \mu_2$ if both μ_1 and μ_2 are invariant distributions. □

3. Spectral gap

Let's point out, if M is non self-adjoint, there may be no spectral gap in $L^2(X, \nu)$. Fu-Zhou Gong and Li-Ming Wu give a counterexample as follows.

Let $X = \{0, 1\}^{\mathbb{N}}$ be the symbol space of one-sided infinite words, equipped with a metric $d(x, y) := 2^{-\inf\{n:\, x_n \neq y_n\}}$ for $x = x_0x_1\cdots$ and $y = y_0y_1\cdots$. Let ξ be a $\{0, 1\}$-valued Bernoulli random variable with the law $\mathbb{P}(\xi = 0) = \mathbb{P}(\xi = 1) = 1/2$. Define the average operator as $\mathrm{M}f(x) = \mathbb{E}[f(\xi x)]$, which determines a random walk m. Then, (X, d, m) becomes a compact Polish space and has a Ricci curvature at least $(1/2)$. The unique invariant distribution ν is the infinite product measure $\mathbb{P}^{\mathbb{N}}$. However, the spectral radius of M on $L^2(X, \nu)/\{\mathrm{const}\}$ equals 1, since $\mathrm{M}^*\mathrm{M} = \mathrm{Id}$.

However, the spectral gap in the set of Lipschitz functions always exists, since $\|\mathrm{M} f - \nu f\|_{\mathrm{Lip}} \leqslant (1-\kappa)\|f\|_{\mathrm{Lip}}$.

Now, we prove Theorem 1.2 without the original condition $\sigma < \infty$ in [9].

Proof of Theorem 1.2. Define a new metric $\hat{d}$ as

$$\hat{d}(x, y) = \begin{cases} d(x, y), & \text{if} \quad d(x, y) \leqslant 1, \\ \sqrt{d(x, y)}, & \text{otherwise.} \end{cases}$$

Denote by $\mathrm{Lip}_1(\hat{d})$ the set of 1-Lipschitz functions with respect to $\hat{d}$. When $\hat{d}(x,y) \leqslant 1$, we have due to $\mathrm{Lip}_1(\hat{d}) \subset \mathrm{Lip}_1$ that

$$W_1^{\hat{d}}(m_x, m_y) := \sup_{f \in \mathrm{Lip}_1(\hat{d})} |\mathrm{M} f(x) - \mathrm{M} f(y)| \leqslant (1-\kappa)\, \hat{d}(x, y).$$

When $\hat{d}(x, y) > 1$, we have for any coupling π of m_x and m_y that

$$\left(\int_{X \times X} \hat{d}(\xi, \eta)\, \mathrm{d}\pi(\xi, \eta) \right)^2 \leqslant \int_{X \times X} d(\xi, \eta)\, \mathrm{d}\pi(\xi, \eta),$$

which implies

$$W_1^{\hat{d}}(m_x, m_y) \leqslant \sqrt{W_1(m_x, m_y)} \leqslant \sqrt{1-\kappa} \cdot \hat{d}(x, y).$$

Hence, $(X, \hat{d}, m)$ has a positive curvature at least $\hat{\kappa} = 1 - \sqrt{1-\kappa} > 0$. The associated diffusion constant $\hat{\sigma}(x)$ is L^2-integrable since ν has a finite first moment by [9] Corollary 21. Then, we can apply the spectral gap result in [9] Proposition 30 to $(X, \hat{d}, m)$. □

4. Entropy decay

In this section, we can prove an entropy-variance inequality which implies the exponential decay of entropy. Then we will obtain a modified LSI if the density function $\mathrm{d}m_x/\mathrm{d}\nu$ belongs to $L^2(X, \nu)$.

According to the proof of Theorem 45 in [9], it follows

Lemma 4.1. *Let* $f > 0$ *and* $t \in \mathbb{N}$. *Then* $\mathrm{Ent}_\nu f = \sum_{t \geqslant 0} \int \mathrm{Ent}_{m_x}(\mathrm{M}^t f)\, \mathrm{d}\nu(x)$.

Proof. This summation formula can be verified straightforward. □

Referring to Barthe and Roberto [3], we give a preliminary inequality, where the control constant is a bit different.

Lemma 4.2. *Let* $\Psi_r(s) = s\log(s/r) - (s-r)$ *for any* $s, r > 0$. *Then*

$$\Psi_{r^2}(s^2) \leqslant 2(1+\log\rho)(s-r)^2, \quad \forall s \in [0, \rho r].$$

Proof. Let $\Phi(s) = 2(1+\log\rho)(s-r)^2 - \Psi_{r^2}(s^2)$, which satisfies $\mathrm{d}\Phi/\mathrm{d}s_{|s=r} = 0$ and $\mathrm{d}^2\Phi/\mathrm{d}s^2 \geqslant 0$ on $[0, \rho r]$. □

Now, let's prove Proposition 1.3. Some ideas come from [3].

Proof of Proposition 1.3. By the definition of $\Psi_r(s)$, we have $\mathrm{Ent}_\nu f = \int \Psi_{\nu f}(f)\,\mathrm{d}\nu$. Denote $E = \{x\colon f(x) \geqslant 2\nu f\}$, which implies for all $y \in E$

$$f(x) \leqslant 2(f(x) - \nu f), \quad \log\frac{f(x)}{\nu f} \leqslant \sqrt{\frac{f(x)}{\nu f}}. \tag{4.1}$$

Using (4.1) and the Hölder inequality gives

$$(\mathbb{I}) := \int_E \Psi_{\nu f}(f(x))\,\mathrm{d}\nu(x) \leqslant \int_E 2(f(x) - \nu f)\sqrt{\frac{f(x)}{\nu f}}\,\mathrm{d}\nu(x) \leqslant 2\sqrt{\mathrm{Var}_\nu f}.$$

Moreover, putting $\rho = 2$ and $C = 2(1+\log\rho)$, we have by Lemma 4.2 that

$$\begin{aligned}(\mathbb{II}) &:= \int_{E^c} \Psi_{\nu f}(f(x))\,\mathrm{d}\nu(x)\\ &\leqslant C\int_{E^c}(\sqrt{f(x)} - \sqrt{\nu f})^2\,\mathrm{d}\nu(x) \leqslant C\sqrt{\mathrm{Var}_\nu f}.\end{aligned}$$

Combining the above estimates and Theorem 1.2 yields

$$\mathrm{Ent}_\nu f \leqslant (\mathbb{I}) + (\mathbb{II}) \leqslant (2+C)\sqrt{\mathrm{Var}_\nu f}.$$

Hence, we obtain the exponential decay of $\mathrm{Ent}_\nu \mathrm{M}^t f$ by Proposition 1.2. □

Now we prove Proposition 1.4, a modified LSI.

Proof of Proposition 1.4. Put $\mathrm{Q}(f) = \int \mathrm{Ent}_{m_x} f\,\mathrm{d}\nu(x)$. Lemma 4.1 and Proposition 1.3 give

$$\mathrm{Ent}_\nu f = \sum_{t\geqslant 0}\mathrm{Q}(\mathrm{M}^t f) \leqslant \mathrm{Q}(f) + \frac{C_1}{1-\sqrt{1-\kappa}}\sqrt{\mathrm{Var}_\nu(\mathrm{M}f)}.$$

Then, we have by the concavity of logarithm and self-adjointness of M that

$$\mathrm{Q}(f) = \int f\cdot\log f - \mathrm{M}f\cdot\log\mathrm{M}f\,\mathrm{d}\nu \leqslant \mathcal{E}^*(f, \log f).$$

Recall the notation $p(x, y) = \mathrm{d}m_x(y)/\mathrm{d}\nu(y)$, we also have by the Minkowski inequality that

$$\sqrt{\mathrm{Var}_\nu(\mathrm{M}f)} = \left(\int \left|\int f(y) - \nu f \,\mathrm{d}m_x(y)\right|^2 \mathrm{d}\nu(x)\right)^{1/2}$$
$$\leqslant \int \left(\int |f(y) - \nu f|^2 p(x, y)^2 \,\mathrm{d}\nu(x)\right)^{1/2} \mathrm{d}\nu(y) \leqslant 2C_2 \nu f.$$

Combining above estimates, we complete the proof. □

5. LSI for λ-range gradient operator

Let $\xi_{x,y}$ be an optimal coupling of m_x and m_y for $x \neq y$, and

$$\kappa_+(x, y) := \frac{1}{d(x, y)} \int (d(x, y) - d(x', y'))_+ \,\mathrm{d}\xi_{x,y}(x', y'),$$
$$\kappa_-(x, y) := \frac{1}{d(x, y)} \int (d(x, y) - d(x', y'))_- \,\mathrm{d}\xi_{x,y}(x', y'),$$

satisfying $\kappa(x, y) = \kappa_+(x, y) - \kappa_-(x, y)$. According to [9] Definition 42, the *unstability* $U(x, y)$ is defined as $U(x, y) = \kappa_-(x, y)/\kappa(x, y)$ and $U = \sup_{x \neq y} U(x, y)$.

Recall the definition of λ-range gradient operator (1.3), let's address two facts from [9] Theorem 44 and Lemma 48 respectively.

Lemma 5.1 (Gradient Contraction)**.** *Suppose* $\kappa > 0$ *and* $\sigma_\infty < \infty$. *If* $0 < \lambda \leqslant 1/(20\sigma_\infty(1 + U))$, *then for any* f *with* $\mathrm{D}f < \infty$ *and any* x

$$(\mathrm{DM}f)(x) \leqslant \left(1 - \frac{\kappa}{2}\right)\mathrm{M}(\mathrm{D}f)(x).$$

Lemma 5.2. *Let* f *satisfying* $\mathrm{D}f < \infty$. *Then for any* $y, z \in \mathrm{Supp}\, m_x$

$$|f(y) - f(z)| \leqslant e^{4\lambda\sigma_\infty} d(y, z)\mathrm{M}(\mathrm{D}f)(x).$$

Denote by D_4 the 4λ-range gradient operator.

Lemma 5.3. *Let* f *satisfying* $\mathrm{D}f < \infty$. *Then for all* x

$$\mathrm{D}_4(f^2)(x) \leqslant 2|f(x)|\mathrm{D}f(x) + \frac{2}{e\lambda}(\mathrm{D}f)^2(x).$$

Proof. We have by using $a \leqslant e^{a-1}$ that

$$|f(y) - f(x)| \leqslant \left(\frac{|f(y) - f(x)|}{d(y, x)} e^{-2\lambda\, d(y,x)}\right)\frac{e^{3\lambda d(y,x)}}{e\lambda} \leqslant \mathrm{D}f(x)\frac{e^{3\lambda d(y,x)}}{e\lambda},$$

and for any $y, y' \in X$

$$|f^2(y) - f^2(y')| \leqslant |f(y) - f(y')|(2|f(x)| + |f(y) - f(x)| + |f(y') - f(x)|).$$

Combining the above estimates yields an upper bound of $\mathrm{D}_4(f^2)(x)$. □

Now, let's prove Theorem 1.5, the standard LSI.

Proof. Take $4\lambda \leqslant 1/(20\sigma_\infty(1+U))$ such that Lemma 5.1 holds for $\mathrm{D}_4 f$. For simplicity, denote $\rho(x) = \sigma(x)^2/n_x$ (see (1.4)), $\alpha = e^{16\lambda\sigma_\infty} \leqslant e^{1/5}$ and $\beta = 1/(e\lambda)$.

Given f with $f^2 > 0$ everywhere. Denote $h = \mathrm{M}^t(f^2)$ for $t \in \mathbb{N}$. For any $x \in X$, due to $a \log a \leqslant a^2 - a$, we have $\mathrm{Ent}_{m_x}(h) \leqslant \mathrm{Var}_{m_x} h/\mathrm{M}h(x)$. Applying Lemma 5.2 to $\mathrm{D}_4 h$ yields

$$\frac{\mathrm{Var}_{m_x} h}{\mathrm{M}h(x)} \leqslant \alpha^2\rho(x)\frac{(\mathrm{M}(\mathrm{D}_4 h)(x))^2}{\mathrm{M}h(x)}.$$

Moreover, it follows from Lemma 5.1 and 5.3 that

$$\mathrm{M}(\mathrm{D}_4 h) \leqslant \Big(1 - \frac{\kappa}{2}\Big)^t \mathrm{M}^{t+1}\big(\mathrm{D}_4(f^2)\big) \leqslant 2\Big(1 - \frac{\kappa}{2}\Big)^t \mathrm{M}^{t+1}(|f|\mathrm{D}f + \beta(\mathrm{D}f)^2).$$

Abbreviate $A = \mathrm{M}^{t+1}((\mathrm{D}f)^2)(x)$ and $B = \mathrm{M}h(x) = \mathrm{M}^{t+1}(f^2)(x)$, then combining the above estimates and Hölder inequality yields

$$\mathrm{Ent}_{m_x}(h) \leqslant 4\alpha^2\rho(x)\Big(1 - \frac{\kappa}{2}\Big)^{2t}\Big(\sqrt{A} + \beta\frac{A}{\sqrt{B}}\Big)^2. \tag{5.1}$$

On the other hand, recall the proof of Proposition 1.3 in Section 4 (replace ν by m_x), denote $\Psi_r(s) = s\log(s/r) - (s - r)$ for any $r, s > 0$, then

$$\mathrm{Ent}_{m_x} h = \int \Psi_{\mathrm{M}h(x)}(h)\,\mathrm{d}m_x \leqslant 2(2 + \log 2)\sqrt{\mathrm{Var}_{m_x} h}.$$

Hence, by the similar argument to derive (5.1), there is another bound

$$\mathrm{Ent}_{m_x} h \leqslant 4(2 + \log 2)\alpha\sqrt{\rho(x)}\Big(1 - \frac{\kappa}{2}\Big)^t(\sqrt{AB} + \beta A). \tag{5.2}$$

Now, let $\gamma > 0$ be a parameter which will be determined later. Using either (5.1) if $\beta^2 A \leqslant \gamma^2 B$ or (5.2) otherwise, we have

$$\begin{aligned}\mathrm{Ent}_{m_x} h \leqslant{}& \max\{4\alpha^2(1+\gamma)^2\sqrt{\rho(x)}, 4(2+\log 2)\alpha\beta(1+\gamma^{-1})\}\\ &\cdot\sqrt{\rho(x)}\cdot\Big(1 - \frac{\kappa}{2}\Big)^t \cdot A.\end{aligned} \tag{5.3}$$

Note that $\sqrt{\rho(x)} \leqslant \sqrt{2}\sigma_\infty < 1/\lambda$, the above maximum is not greater than $16/\lambda$ by taking $\gamma = (\sqrt{5} - 1)/2$. Finally, we apply Lemma 4.1 to (5.3). □

6. Ricci curvature for resistance forms

Let μ be a Borel probability measure on the Polish space X. Let $(\mathcal{E}, \mathcal{F})$ be a symmetric Dirichlet form in $L^2(X,\mu)$ associated to a conservative Markov process with a transition kernel $p_t(x, \mathrm{d}y)$. According to [5], define

$$R(x, y) := \sup_{\mathcal{E}(f,f)\neq 0} \frac{|f(x) - f(y)|^2}{\mathcal{E}(f, f)}, \quad \forall x, y \in X, \tag{6.1}$$

which is called effective resistance if it is finite everywhere. In fact, R is a metric on X and $(\mathcal{E}, \mathcal{F})$ called a resistance form. For resistance forms on fractals, we refer to [2, 5, 6] and references therein. Sturm [10] gave another kind of gradient-type construction via the Γ-*convergence* argument.

Define a random walk $m = \{m_x\}$ (depending on α) by

$$\mathrm{d}m_x(y) := \int_0^\infty \alpha e^{-\alpha t} p_t(x, \mathrm{d}y)\, \mathrm{d}t.$$

Then $m_x \in \mathcal{M}_1$ due to the conservativeness, μ is invariant to m, and the averaging operator M is self-adjoint since $(\mathcal{E}, \mathcal{F})$ is symmetric. Moreover, recall the definition of resolvent operators family, we have $\mathrm{M}f = \alpha G_\alpha f$.

Now, let's prove Proposition 1.6 that $(X, \sqrt{R}, m)$ has a positive curvature.

Proof of Proposition 1.6. Denote by Lip_1 the set of 1-Lipschitz functions under $\sqrt{R}$. Let $f \in \mathrm{Lip}_1$ and $f(o) = 0$ for $o \in X$, then $f \in L^2(X,\mu)$, $\mathrm{M}f = \alpha G_\alpha f \in \mathcal{F}$ and

$$|\mathrm{M}f(x) - \mathrm{M}f(y)| \leqslant \sqrt{R(x, y)} \cdot \sqrt{\mathcal{E}(\mathrm{M}f, \mathrm{M}f)}.$$

Applying the basic properties of resolvent operators, we have

$$\mathcal{E}(\mathrm{M}f, \mathrm{M}f) = \alpha(f - \alpha G_\alpha f, \alpha G_\alpha f) = \alpha(f - \mathrm{M}f, \mathrm{M}f),$$

and estimate respectively

$$|f(x) - \mathrm{M}f(x)| \leqslant \int \sqrt{R(x, y)}\, \mathrm{d}m_x(y), \ |\mathrm{M}f(x)| \leqslant \int \sqrt{R(o, y)}\, \mathrm{d}m_x(y).$$

Denote $g(x) = \int \sqrt{R(o, y)}\, \mathrm{d}m_x(y)$, we obtain by the Hölder inequality that

$$\mathcal{E}(\mathrm{M}f, \mathrm{M}f) \leqslant \alpha \int \sqrt{R(o, x)} g(x) + g^2(x)\, \mathrm{d}\mu(x) \leqslant 2\alpha \int R(o, x)\, \mathrm{d}\mu(x).$$

Hence, the desired result follows from (1.2). □

In particular, a standard LSI (6.2) holds if $R(x, y)$ is uniformly bounded.

Proposition 6.1. *Suppose $R(x, y) \leqslant C$ for all x, y. Then for any $f \in \mathcal{F}$*

$$\mathrm{Ent}_\mu(f^2) \leqslant 2C\mathcal{E}(f, f). \tag{6.2}$$

Proof. Due to $a \log a \leqslant a^2 - a$, we have

$$\mathrm{Ent}_\mu(f^2) \leqslant \frac{\mathrm{Var}_\mu(f^2)}{\mu(f^2)} = \frac{1}{2\mu(f^2)} \iint |f^2(x) - f^2(y)|^2 \,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y).$$

Since $|f(x) - f(y)|^2 \leqslant R(x, y)\mathcal{E}(f, f)$ by (6.1), we complete the proof. □

REMARK 6.1. $R(x, y) \leqslant C$ means $\sigma_\infty < \infty$, so we can use Theorem 1.5 and Proposition 1.6 to show the LSI but with a constant much bigger than $2C$.

ACKNOWLEDGEMENTS. The authors gratefully thank M. Takeda and L.-M. Wu for helpful discussions and comments. The first and second named authors thank the financial support from Key Laboratory of Random Complex Structures and Data Science of CAS 2008DP173182. Moreover, respectively, Fu-Zhou Gong is supported by NSFC-11021161 and 973 program-2011CB808000, Yuan Liu supported by research grant-Y129161ZZ1, and Zhi-Ying Wen supported by NSFC-61071066. The main results of this paper were presented in "2nd International Conference on Stochastic Analysis and its Applications" held at Seoul National university, Seoul, Korea on May 28–31, 2008, and in "IMS-China International Conference on Statistics and Probability" held at Zhejiang university, Hangzhou, China on June 11–13, 2008.

References

[1] S. Aida and D. Stroock: *Moment estimates derived from Poincaré and logarithmic Sobolev inequalities*, Math. Res. Lett. **1** (1994), 75–86.

[2] M.T. Barlow and R.F. Bass: *Brownian motion and harmonic analysis on Sierpinski carpets*, Canad. J. Math. **51** (1999), 673–744.

[3] F. Barthe and C. Roberto: *Modified logarithmic Sobolev inequalities on* $\mathbb{R}$, Potential Anal. **29** (2008), 167–193.

[4] M. Fukushima, Y. Oshima and M. Takeda: Dirichlet Forms and Symmetric Markov Processes, de Gruyter, Berlin, 1994.

[5] J. Kigami: Analysis on Fractals, Cambridge Univ. Press, Cambridge, 2001.

[6] S. Kusuoka and X.Y. Zhou: *Dirichlet forms on fractals*: *Poincaré constant and resistance*, Probab. Theory Related Fields **93** (1992), 169–196.

[7] M. Ledoux: The Concentration of Measure Phenomenon, Mathematical Surveys and Monographs **89**, Amer. Math. Soc., Providence, RI, 2001.

[8] Z.-M. Ma and M. Röckner: Introduction to the Theory of (Non-Symmetric) Dirchlet Forms, Springer-Verlag, Berlin, 1992.

[9] Y. Ollivier: *Ricci curvature of Markov chains on metric spaces*, J. Funct. Anal. **256** (2009), 810–864.
[10] K.T. Sturm: *Diffusion processes and heat kernels on metric spaces*, Ann. Probab. **26** (1998), 1–55.
[11] T. Szarek: *Feller processes on nonlocally compact spaces*, Ann. Probab. **34** (2006), 1849–1863.
[12] C. Villani: Topics in Optimal Transportation, Graduate Studies in Mathematics **58**, Amer. Math. Soc., Providence, RI, 2003.

Fu-Zhou Gong
Institute of Applied Mathematics
Academy of Mathematics and Systems Science
Chinese Academy of Sciences
Beijing 100190
China
e-mail: fzgong@amt.ac.cn

Yuan Liu
Institute of Applied Mathematics
Academy of Mathematics and Systems Science
Chinese Academy of Sciences
Beijing 100190
China
e-mail: liuyuan@amss.ac.cn

Zhi-Ying Wen
Department of Mathematical Sciences
University of Tsinghua
Beijing 100084
China
e-mail: wenzy@tsinghua.edu.cn