Complete logarithmic Sobolev inequalities via Ricci curvature bounded below

Michael Brannan [a], Li Gao [a,*], Marius Junge [b]

[a] *Department of Mathematics, Texas A&M University, College Station, TX 77840, USA*
[b] *Department of Mathematics, University of Illinois, Urbana, IL 61801, USA*

A R T I C L E   I N F O

A B S T R A C T

We prove that for a symmetric Markov semigroup, Ricci curvature bounded from below by a non-positive constant combined with a finite L_∞-mixing time implies the modified log-Sobolev inequality. Such L_∞-mixing time estimates always hold for Markov semigroups that have spectral gap and finite Varopoulos dimension. Our results apply to non-ergodic quantum Markov semigroups with noncommutative Ricci curvature bounds recently introduced by Carlen and Maas. As an application, we prove that the heat semigroup on a compact Riemannian manifold admits a uniform modified log-Sobolev inequality for all its matrix-valued extensions.

Published by Elsevier Inc.

1. Introduction

In differential geometry, Ricci curvature lower bounds have many applications in topology, geometry and analysis. One pioneering work that connects Ricci curvature with analysis of heat semigroups is the Bakry-Emery theorem [2]. It implies that if the Ricci curvature of a compact Riemannian manifold (M, g) is bounded from below by a positive

constant, then the heat semigroup satisfies a logarithmic Sobolev inequality. In this paper, motivated by quantum information theory, we present a uniform approach to obtain logarithmic Sobolev inequalities from a non-positive Ricci curvature lower bound for both classical and quantum Markov semigroups. Indeed, we show that a non-positive Ricci curvature lower bound plus a L_∞-mixing time implies logarithmic Sobolev inequality in the noncommutative non-ergodic setting.

In the past decades, the notion of Ricci curvature lower bound has been largely extended beyond Riemannain manifolds using ideas from optimal transport. Motivated by Gromov's precompactness theorem [27], Lott-Villani [42] and Strum [55] independently introduced a notion of Ricci curvature lower bound for metric measure spaces. Such a space has Ricci curvature bounded below by a constant λ if the entropy, as a functional on the state space (space of all probability measures), is λ-convex along geodesics of the L_2-Wasserstein distance. Later, similar ideas were extended to Markov semigroups on discrete spaces and noncommutative spaces. The key ingredient is to construct an analog of the Wasserstein distance W on the state space such that the semigroup is the gradient flow of the entropy functional with respect to W. Such gradient flow constructions were obtained independently in [43,44,13] for Markov process on finite state spaces, and [9,10,45,47] for finite dimensional quantum systems. More recently, the noncommutative Wasserstein metric has been further studied on finite von Neumann algebras [60,31]. Based on these, the notions of Ricci curvature lower bound via λ-convexity of entropy have been studied by Erbar and Maas [21] for discrete spaces and by Carlen-Maas [11], Datta-Rouzé [15] and Wirth [60] for noncommutative spaces. Thanks to the gradient flow structure, the connection between Ricci curvature and functional inequalities, including the extensions of the Bakry-Emery theorem, has been obtained in all the above settings.

The logarithmic Sobolev inequalities were first introduced by Gross [29,28] as a reformulation of hypercontractivity, and have been intensively studied since then (see [30] for an overview). The focus of this paper is the L_1-version of the log-Sobolev inequality, also called the modified log-Sobolev inequality. Indeed, let $T_t = e^{-At} : L_\infty(\Omega, \mu) \to L_\infty(\Omega, \mu)$ be a Markov semigroup with Dirichlet form $\mathcal{E}(f) = (f, Af)$. We say T_t satisfies a λ-modified log-Sobolev inequality (in short, λ-MLSI) if for any probability density function f,

$$2\lambda \int f \log f \, d\mu \leq \mathcal{E}(f, \log f) \, , \, \forall \, f \geq 0, \int f \, d\mu = 1$$

The integral on the left hand side of the above inequality is the entropy $H(f) = \int f \log f \, d\mu$ and the right hand side is called the Fisher information $I(f) = \int (Af) \log f \, d\mu$, which describes the decreasing rate of entropy: $I(T_t f) = -\frac{d}{dt} H(T_t(f))$. Intuitively, MLSI characterizes the exponential decay of entropy along the time evolution of the semigroup. In the smooth setting, MLSI is equivalent to the more common L_2-log-Sobolev inequality

$$\lambda \int g^2 \log g^2 \, d\mu \leq 2\mathcal{E}(g, g) \, , \, \forall \, \int g^2 \, d\mu = 1 \, . \tag{1}$$

However, it is weaker than (1) in discrete and noncommutative cases. See [38] for a review article on the interplay between spectral gap, log-Sobolev inequalities and Ricci curvature. Otto-Villani [49] proved that MLSI also implies Talagrand's transport cost inequality, which further bounds spectral gap and derives concentration of measure phenomena. More recently these applications of MLSI have also been extended to (finite dimensional) quantum Markov semigroups [10,52], which suggest a uniform picture of functional inequalities in both classical and noncommutative settings.

Quantum Markov semigroups are noncommutative generalization of classical Markov semigroups, where the underlying function spaces are replaced by matrix algebras or operator algebras. A quantum Markov semigroup on a von Neumann algebra $\mathcal{M}$ is an ultra-weakly continuous family $(T_t)_{t\geq 0} : \mathcal{M} \to \mathcal{M}$ of normal unital completely positive maps. When $\mathcal{M} = B(H)$ the bounded operators on a Hilbert space H, quantum Markov semigroups model the time evolution of dissipative open quantum system. In operator algebras, quantum Markov semigroups have been widely studied in the context of approximation properties, structure theory, and noncommutative harmonic analysis (see e.g. [12,34]). In this paper, we will focus on *symmetric quantum Markov semigroup* on finite von Neumann algebras. That is, $\mathcal{M}$ is a von Neumann algebra equipped with a normal faithful tracial state τ, and the semigroup $T_t : \mathcal{M} \to \mathcal{M}$ is given by self-adjoint maps with respect to the τ-inner product. This setting avoids the technicalities of Tomita–Takesaki theory, but is still broad enough to cover many examples of wide interests, such as classical Markov semigroups on probability spaces, finite dimensional dissipative systems in quantum information theory and quantum many-body system, and also various infinite dimensional examples in operator algebras.

One of the main motivations for this work is to prove a MLSI for quantum Markov semigroups that is stable under tensor products. For classical Markov semigroups, it is known that if a pair of semigroups S_t, T_t satisfy λ-MLSI, then $S_t \otimes T_t$ satisfies λ-MLSI. Tensorization is a useful property that allows us to obtain MLSI for composite systems by studying smaller, more tractable subsystems. In the noncommutative setting, tensor stability of MLSI generally requires not only MLSI but a "completely bounded" version of MLSI: T_t is said to satisfy a λ-complete log-Sobolev inequality (λ-CLSI) if all of its matrix-valued extensions $T_t \otimes \mathrm{id}_{M_n}$ satisfy λ-MLSI. For quantum Markov semigroups, CLSI has the tensor-stability that S_t and T_t satisfy λ-CLSI $\Rightarrow S_t \otimes T_t$ satisfies λ-CLSI [22]. For classical Markov semigroups, CLSI simply means an uniform MLSI constant for all matrix-valued functions, and for quantum Markov semigroups, CLSI has applications in estimating decay rates of entanglement. The study of CLSI naturally leads us to consider non-ergodic semigroups, because $T_t \otimes \mathrm{id}$ always has non-trivial fixed-point space.

We now describe the content of paper and state our main results. Section 2 reviews the basic definitions and proves some preliminary lemmas.

The main theorem of this paper is discussed in Section 3, which we illustrate here using the example of the heat semigroup. Let $T_t = e^{-\Delta t}$ be the heat semigroup on a compact Riemannnian manifold (M, g). There are two key ingredients in our proof. The

first one is (displaced) monotonicity of Fisher information. The idea goes back to the Bakry-Emery theorem, in the proof of which they actually showed the implications

$$\left\{ \text{Ricci curvature lower bound } \lambda \right\} \overset{\lambda \in \mathbb{R}}{\Longrightarrow} \left\{ I(T_t f) \le e^{-2\lambda t} I(f) \ \forall t \ge 0 \right\} \overset{\lambda > 0}{\Longrightarrow} \lambda\text{-MLSI.} \tag{2}$$

We call the middle inequality "λ-Fisher monotonicity", as for $\lambda = 0$, it asserts that $I(T_t f)$ is non-increasing in t. For $\lambda > 0$, this immediately implies λ-MLSI, which in classical setting have been studied under the name "convex entropy decay" [8]. For $\lambda \le 0$, we will need a second ingredient, which is the finiteness of the following L_∞-mixing time

$$t_{cb} = \inf\{t > 0 | \ \|T_t - E : L_1(M, d\mu) \to L_\infty(M, d\mu)\| \le 1/2\} < \infty$$

Here $E(f) = (\int f d\mu)1$ is the averaging map. We prove that this L_∞-mixing time is the half-decay time for entropy $H(T_t f)$, and t_{cb} is always finite by the spectral gap of Δ and standard heat kernel estimates. All the notions mentioned above including the implication (2) are fully adapted to the noncommutative non-ergodic setting, which leads to the statement of our main theorem.

Theorem 1.1 *(cf. Theorem 3.6). Let $T_t : \mathcal{M} \to \mathcal{M}$ be a symmetric quantum Markov semigroup and $E : \mathcal{M} \to \mathcal{N}$ be the conditional expectation onto its fixed point algebra $\mathcal{N}$. Suppose*

i) *T_t satisfies λ-Fisher monotonicity for some $\lambda \in \mathbb{R}$: for all densities ρ,*

$$I(T_t(\rho)) \le e^{-\lambda t} I(\rho) \ , \forall t \ge 0$$

ii) *T_t has finite completely bounded return time:*

$$t_{cb} = \inf\{t > 0 \mid \|T_t - E : L_\infty^1(\mathcal{N} \subset \mathcal{M}) \to L_\infty(\mathcal{M})\|_{cb} \le 1/2\} < \infty \ .$$

Then T_t-satisfies $\kappa(\lambda, t_{cb})$-MSLI for $\kappa(\lambda, t) = \begin{cases} \frac{\lambda}{2(1 - e^{-2t\lambda})}, & \text{if } \lambda \ne 0 \\ \frac{1}{4t}, & \text{if } \lambda = 0. \end{cases}$

For classical Markov semigroups, it is well-known that the L_∞-mixing time itself implies the log-Sobolev inequality (see [18]). Nevertheless, this standard approach via hypercontractivity does not apply to the matrix-valued setting because the famous Rothaus Lemma as a crucial step is no longer valid (see [22, Section 7.5]). We emphasis that our main theorem, using ideas from quantum information theory, applies to fully non-ergodic noncommutative setting. It allows one to derive MSLI for matrix-valued functions or endomorphism maps on vector bundle, and also the tensor-stable CLSI for quantum Markov semigroups.

In Section 4 we apply the main theorem to various examples in both the classical and quantum contexts. Section 4.1 discusses the connection to Bakry-Emery's curvature dimension condition for Markov diffusion semigroups. An important class of such semigroups are heat semigroups on (weighted) Riemannian manifolds, for which we have the following result

Theorem 1.2 *(cf. Theorem 4.4). Every heat semigroup on a connected compact (weighted) Riemannian manifold satisfies CLSI.*

In Section 4.3, we show that any "central" semigroup on a compact group has entropy curvature bound zero, and based on that, we estimate the optimal CLSI constant for the heat semigroup on d-torus $\mathbb{T}^d$. For noncommutative examples, Section 4.4 studies entropy Ricci curvature bounds and MLSI constants for depolarizing semigroups. We also consider Schur multiplier semigroups and semigroups of random unitary channels in Section 4.5 & 4.6. We end our paper discussion with an appendix on approximations of relative entropy. This paper is the first part of a series of two. In the second part [7], we continue our study to infinite dimensional noncommutative examples in operators algebras, including group algebra, quantum groups and q-Gaussian algebras.

Acknowledgments

2. Preliminaries

2.1. Entropy and relative entropy

Throughout the paper, we let $\mathcal{M}$ be a finite von Neumann algebra equipped with a normal faithful finite tracial state τ. For $0 < p < \infty$, the L_p-space $L_p(\mathcal{M})$ is defined as the completion of $\mathcal{M}$ with respect to the norm

$$\|a\|_p = \tau(|a|^p)^{1/p} .$$

We identify $L_\infty(\mathcal{M}) := \mathcal{M}$ and the predual space $\mathcal{M}_* \cong L_1(\mathcal{M})$ via the duality

$$a \in L_1(\mathcal{M}) \longleftrightarrow \phi_a \in \mathcal{M}_*, \ \phi_a(x) = \tau(ax) .$$

We say $\rho \in L_1(\mathcal{M})$ is a density operator (or simply density) if $\rho \geq 0$ and $\tau(\rho) = 1$. The set of all densities correspond to the normal states of $\mathcal{M}$, which we denote by $S(\mathcal{M})$. In

all the discussion, states always mean normal states and are identified with their density operators.

Recall that for two normal positive linear functionals ρ and σ, the Umegaki-Araki relative entropy is

$$D(\rho||\sigma) = \begin{cases} \langle \rho^{1/2} | \log \Delta(\rho, \sigma) | \rho^{1/2} \rangle, & \text{if } \operatorname{supp}(\rho) \leq \operatorname{supp}(\sigma) \\ +\infty, & \text{otherwise,} \end{cases}$$

where $\Delta(\rho, \sigma)(x) = \rho x \sigma^{-1}$ is the relative modular operator and $|\rho^{1/2}\rangle$ is the vector of $\rho^{1/2}$ in $L_2(\mathcal{M})$. In the tracial setting

$$D(\rho||\sigma) = \tau(\rho \log \rho - \rho \log \sigma) ,$$

provided $\rho \log \rho, \rho \log \sigma \in L_1(\mathcal{M})$. The entropy of ρ is then given by $H(\rho) = D(\rho||1)$. (Note that H is actually the Boltzmann H-functional, which differs with the usual von Neumann entropy in quantum information theory by a negative sign.) We say a linear map $\Phi : L_1(\mathcal{M}) \to L_1(\mathcal{M})$ is completely positive trace preserving (CPTP) if its adjoint $\Phi^\dagger : \mathcal{M} \to \mathcal{M}$ is normal, unital, and completely positive (UCP). The monotonicity of the relative entropy under CPTP maps (also called the *data processing inequality*) states that for any CPTP Φ and any two states ρ, σ,

$$D(\rho||\sigma) \geq D(\Phi(\rho)||\Phi(\sigma)) .$$

In particular, we have $D(\rho||\sigma) \geq 0$ for any ρ and σ, and the equality $D(\rho||\sigma) = 0$ holds if and only if $\rho = \sigma$.

Let $\mathcal{N} \subset \mathcal{M}$ be a von Neumann subalgebra. The trace preserving conditional expectation $E : \mathcal{M} \to \mathcal{N}$ on to $\mathcal{N}$ is the (unique) completely positive unital map determined by

$$\tau(xy) = \tau(xE(y)), \forall x \in \mathcal{N}, y \in \mathcal{M} .$$

The conditional expectation E is normal and its pre-adjoint map gives an embedding $L_1(\mathcal{N}) \subset L_1(\mathcal{M})$. For a state ρ, the relative entropy with respect to $\mathcal{N}$ is defined as follows

$$D(\rho||\mathcal{N}) := \inf_{\sigma \in S(\mathcal{N})} D(\rho||\sigma) = D(\rho||E(\rho)) ,$$

where the infimum is always attained by $E(\rho)$. Indeed, we have the chain rule [33, Lemma 3.4] that for $\sigma \in \mathcal{S}(\mathcal{N})$

$$D(\rho||\sigma) = D(\rho||E(\rho)) + D(E(\rho)||\sigma) ,$$

and the infimum is attained if and only if $D(E(\rho)||\sigma)$ is zero. If $H(\rho) = D(\rho||1) < \infty$ is finite, so does

$$H(E(\rho)) = D(E(\rho)||1) \leq D(\rho||1) = H(\rho) < \infty$$

and

$$D(\rho||\mathcal{N}) = \tau(\rho \log \rho - \rho \log E(\rho)) = \tau(\rho \log \rho) - \tau(E(\rho) \log E(\rho)) = H(\rho) - H(E(\rho)) .$$

If Φ is CPTP and $\Phi(L_1(\mathcal{N})) \subset L_1(\mathcal{N})$ (or equivalently $\Phi^\dagger(\mathcal{N}) \subset \mathcal{N}$), we have the data processing inequality for $D(\rho||\mathcal{N})$,

$$D(\Phi(\rho)||\mathcal{N}) \leq D(\Phi(\rho)||\Phi \circ E(\rho)) \leq D(\rho||E(\rho)) = D(\rho||\mathcal{N}) .$$

Here the second inequality follows from $\Phi \circ E(\rho) \in S(\mathcal{N})$. As already seen in [4,22], the relative entropy $D(\rho||\mathcal{N})$ is crucial in functional inequalities for non-ergodic Markov semigroups.

2.2. Quantum Markov semigroups

A quantum Markov semigroup is a family of linear maps $(T_t)_{t \geq 0} : \mathcal{M} \to \mathcal{M}$ with the following properties

i) T_t is a normal UCP map for all $t \geq 0$.
ii) $T_t \circ T_s = T_{s+t}$ for any $t, s \geq 0$ and $T_0 = \mathrm{id}$.
iii) for each $x \in \mathcal{M}$, $t \mapsto T_t(x)$ is continuous in ultra-weak topology.

The generator of the semigroup is defined as

$$Ax = w^* - \lim_{t \to 0} \frac{x - T_t(x)}{t} \ , \ T_t = e^{-At} \ ,$$

where A is a closable densely defined operator on $L_2(\mathcal{M})$. We say a quantum Markov semigroup (T_t) is *symmetric* if for any t, T_t is a self-adjoint map for the τ-inner product,

$$\tau(x^* T_t(y)) = \tau(T_t(x)^* y) \ , \ x, y \in \mathcal{M}.$$

We refer to [16] for the basic properties of symmetric quantum Markov semigroups. A symmetric quantum Markov semigroup is determined by its *Dirichlet form*

$$\mathcal{E} : L_2(\mathcal{M}) \to [0, \infty] \ , \ \mathcal{E}(x, x) = \tau(x^* Ax) \ .$$

We write $\mathrm{dom}(A)$ for the domain of A and $\mathrm{dom}(A^{1/2})$ for the domain of $\mathcal{E}$. The Dirichlet subalgebra $\mathcal{A}_\mathcal{E} := \mathrm{dom}(A^{1/2}) \cap \mathcal{M}$ is a dense $*$-subalgebra of $\mathcal{M}$ and a core of $A^{1/2}$ [16].

For symmetric semigroups, $T_t = T_t^{\dagger}$ are unital completely positive and trace preserving (in short, UCPTP), and the generator A is self-adjoint and positive on $L_2(\mathcal{M})$. Let $\mathcal{N}$ be the common multiplicative domain for (T_t), defined as follows

$$\mathcal{N} = \{a \in \mathcal{M} \mid T_t(a^*)T_t(a) = T_t(a^*a) \text{ and } T_t(a)T_t(a^*) = T_t(aa^*) , \forall \, t \geq 0\} \qquad (3)$$

Let E be the conditional expectation onto $\mathcal{N}$. For symmetric (T_t), we have

$$T_t \circ E = E \circ T_t = E .$$

Then $\mathcal{N} = \{x \in \mathcal{M} \mid T_t(x) = x, \forall t\}$ is the fixed-point subalgebra, and each T_t is an $\mathcal{N}$-bimodule map,

$$T_t(axb) = aT_t(x)b , \ \forall \, a, b \in \mathcal{N}, x \in \mathcal{M}$$

In particular, we have $A(\mathcal{N}) = 0$ and $\mathcal{N} \subset \mathcal{A}_{\mathcal{E}}$.

We say (T_t) is *ergodic* (also called *primitive*) if $\mathcal{N} = \mathbb{C}1$ is trivial. This means the semigroup admits a unique invariant state. We specify the conditional expectation onto the scalars $\mathbb{C}1$ as $E_\tau(\rho) = \tau(\rho)1$. Throughout the paper, we will focus on symmetric quantum Markov semigroups that are not necessarily *ergodic*. Recall that the *gradient form* (or *carré du champ*) of the generator A is the operator given by

$$\Gamma(x, y) = \frac{1}{2}\Big((Ax^*)y + x^*Ay - A(x^*y)\Big) . \qquad (4)$$

Γ is a (completely) positive sesquilinear form because

$$\Gamma(x, x) = \lim_{t \to 0} \frac{1}{t}(T_t(x^*x) - T_t(x^*)T_t(x)) ,$$

where the right hand side is always positive by the Kadison-Schwarz inequality for unital completely positive maps. We recall the following fundamental Markov dilation result from the preprint [36].

Theorem 2.1 *([36]). Let $T_t = e^{-At} : \mathcal{M} \to \mathcal{M}$ be a symmetric quantum Markov semigroup. Suppose $\Gamma(x, x) \in L_1(\mathcal{M})$ for all $x \in \mathrm{dom}(A^{1/2})$. Then there exists a trace-preserving embedding $\mathcal{M} \subseteq (\hat{\mathcal{M}}, \tau)$ into a finite von Neumann algebra $\hat{\mathcal{M}}$, and a closed symmetric derivation $\delta : \mathrm{dom}(A^{1/2}) \to L_2(\hat{\mathcal{M}})$, meaning that*

i) *$\delta : \mathrm{dom}(A^{1/2}) \to L_2(\hat{\mathcal{M}})$ is a closed linear map such that $\delta(x^*) = \delta(x)^*$.*
ii) *δ satisfies the Leibniz rule: for any $a, b \in \mathrm{dom}(A^{1/2}) \cap \mathcal{M}$,*

$$\delta(ab) = \delta(a)b + a\delta(b) .$$

Moreover, the gradient form Γ and the derivation δ are related through

iii) *for all $z \in \mathcal{M}$,*

$$\tau(\Gamma(x,y)z) \;=\; \hat{\tau}(\delta(x)^*\delta(y)z) \;. \tag{5}$$

Equivalently, $E_{\mathcal{M}}(\delta(x)^*\delta(y)) = \Gamma(x,y)$ where $E_{\mathcal{M}} : \hat{\mathcal{M}} \to \mathcal{M}$ is the conditional expectation. As a consequence, $A = \delta^*\delta$ as an operator on $L_2(\mathcal{M})$.

The above Theorem in 2.1 gives a more specific construction of derivation than the representation theorem for completely Dirichlet forms by Cipriani and Sauvageot [14]. Instead of having a larger von Neumann algebra $\hat{\mathcal{M}}$, [14, Theorems 8.2 & 8.3] ensures the existence of a closed derivation $\partial : \mathrm{dom}(A^{1/2}) \to H$ into a Hilbert $\mathcal{M}$-bimodule, and the derivation ∂ satisfies the Leibniz rule with respect to the bimodule action and

$$\tau(\Gamma(x,y)z) = \langle z\partial(x), \partial(y)\rangle_H \,, \forall, z \in \mathcal{M}, x, y \in \mathrm{dom}(A^{1/2})$$

which is analogous to the property (5). The derivation construction in this setting is used in [60] and [31] to construct the noncommutative Wasserstein distance. Throughout the paper, we will focus on symmetric quantum Markov semigroups in order to ensure the existence of the derivation δ in Theorem 2.1, making heavy use of (5) and also the von Neumann algebra structure of $\hat{\mathcal{M}}$. These ideas are close to the works [10,11] by Carlen and Maas (and also [15]). Nevertheless, the derivation in Theorem 2.1 is clearly a special case of [14, Theorem 8.2 & 8.3] by choosing $\mathcal{H} = L_2(\mathcal{M})$ with natural left and right multiplication. This enables us to apply the results from [14] and [60] in our discussion. We recall the following definition from [39].

Definition 2.2. We say $(\mathcal{A}, \hat{\mathcal{M}}, \delta)$ is a derivation triple for $T_t : \mathcal{M} \to \mathcal{M}$ if

 i) $(\delta, \hat{\mathcal{M}})$ satisfies properties i)-iii) in the Theorem 2.1
 ii) $\mathcal{A} \subset \mathcal{M}$ is a w^*-dense subalgebra such that $\mathcal{A} \subset \mathrm{dom}(A^{1/2})$, $T_t(\mathcal{A}) \subset \mathcal{A}$.

Note that Dirichlet subalgebra $\mathcal{A}_{\mathcal{E}} = \mathrm{dom}(A^{1/2}) \cap \mathcal{M}$ always satisfies ii). Then it is guaranteed by Theorem 2.1 that derivation triples always exist for symmetric semigroups. It was proved in [14, Lemma 7.2] that $\mathcal{A}_{\mathcal{E}}$ is closed under C^1-functional calculus. Indeed, let $x \in \mathcal{M}$ be a self-adjoint element with spectrum $\mathrm{spec}(x) \subset (a,b)$ and let $f : (a,b) \to \mathbb{R}$ be a function with continuous bounded derivative. We have $f(x) \in \mathcal{A}_{\mathcal{E}}$ and its gradient is given by the double operator integral,

$$\delta(f(x)) = J_F^x(\delta(x)) := \int\limits_{\mathbb{R}} \int\limits_{\mathbb{R}} F(x,y)dE_s\delta(x)dE_t$$

where E_s is spectral projection of x and F is the bi-variable function

$$F : \mathbb{R} \times \mathbb{R} \to \mathbb{R} \ , \ F(s,t) = \begin{cases} \frac{f(s) - f(t)}{s - t}, & \text{if } s \neq t \\ f'(s), & \text{if } s = t. \end{cases}$$

For concrete examples, it maybe more convenient to work with some smaller algebra $\mathcal{A} \subset \mathcal{A}_{\mathcal{E}}$ usually with strong regularity. Indeed, for most of examples in our discussions, the derivation triple $(\mathcal{A}, \hat{\mathcal{M}}, \delta)$ will be concretely described, and by assumption $\mathcal{A} \subset \mathcal{A}_{\mathcal{E}}$ always holds. Thus the C^1-functional calculus is also applicable for $\mathcal{A}$ (with $f(x)$ in $\mathcal{A}_{\mathcal{E}}$). It follows from Kaplansky density theorem (cf. [56, Theorem II.4.8]) that $\mathcal{A}$ is norm dense in $L_1(\mathcal{M})$ and $L_2(\mathcal{M})$. Moreover, we denote $\mathcal{A}_0 = \cup_{t>0} T_t(\mathcal{A})$. Then $\mathcal{A}_0 \subset \mathrm{dom}(A)$ is w*-dense in $\mathcal{M}$ and norm-dense in $L_p(\mathcal{M})$ for all $1 \leq p < \infty$ (see [16, Proposition 2.14 & 3.1]).

2.3. Modified logarithmic Sobolev inequalities

Let $T_t = e^{-At} : \mathcal{M} \to \mathcal{M}$ be a symmetric quantum Markov semigroup and let $(\mathcal{A}, \hat{\mathcal{M}}, \delta)$ be a derivation triple of T_t. We first specify some subsets of state space.

$$S_H(\mathcal{M}) = \{ \rho \in S(\mathcal{M}) \mid H(\rho) < \infty \} ,$$
$$S_B(\mathcal{M}) = \{ \rho \in S(\mathcal{M}) \mid \lambda 1 \leq \rho \leq \mu 1 \ , \ \text{for some } \lambda, \mu > 0 \}$$
$$S_B(\mathcal{A}_0) = S_B(\mathcal{M}) \cap \mathcal{A}_0 .$$

Here $S_H(\mathcal{M})$ are states with finite entropy, $S_B(\mathcal{M})$ are states with bounded invertible density and $S_B(\mathcal{A}_0)$ are bounded invertible densities in $\mathcal{A}_0 = \bigcup_{t>0} T_t(\mathcal{A})$. All the three are norm-dense subset of the state space $S(\mathcal{M})$. Recall that the *Fisher information* for $\rho \in S_B(\mathcal{A}_0)$ is defined as

$$I(\rho) := \tau\big((A\rho) \log \rho\big)$$

Definition 2.3. We say a quantum Markov semigroup $T_t = e^{-At}$ satisfies the λ-modified logarithmic Sobolev inequality (in short, λ-MLSI) for $\lambda > 0$ if

$$2\lambda D(\rho \| \mathcal{N}) \leq I(\rho) \ , \ \forall \rho \in S_B(\mathcal{A}_0)$$

Note that we have the constant 2 in the definition to match with curvature constant introduced later. The definition of Fisher information and the derivative relation (6) can be further extended to $\rho \in \mathrm{dom}(A^{1/2})$ as

$$I(\rho) := \lim_{n \to \infty} \mathcal{E}(\rho, \log_{(n)} \rho)$$

where $\log_{(n)}$ is the function $\log_{(n)}(x) = \log(x + e^{-n}) \wedge n$. See [60, Definition 5.17 & Proposition 5.23]. Nevertheless, it suffices (is more convenient) to consider $\rho \in S_B(\mathcal{A}_0)$ for MLSI.

Proposition 2.4. *A semigroup T_t satisfies λ-MLSI if and only if*

$$D(T_t(\rho)||\mathcal{N}) \leq e^{-2\lambda t} D(\rho||\mathcal{N}) \,, \ \forall\, \rho \in S(\mathcal{M}).$$

The proof of the above proposition is a standard density argument included in Appendix and here we illustrate the heuristic. The Fisher information is the negative derivative of (relative) entropy along the semigroup flow

$$I(\rho) = -\frac{d}{dt} D(T_t(\rho)||\mathcal{N})|_{t=0} = -\frac{d}{dt} H(T_t(\rho))|_{t=0} \,, \tag{6}$$

where the second equality follows from

$$D(T_t(\rho)||\mathcal{N}) = D(T_t(\rho)||E(\rho)) = H(T_t(\rho)) - H(E(\rho)) \,.$$

In particular, we have $I(\rho) \geq 0$ by the data processing inequality $D(T_t(\rho)||\mathcal{N}) \leq D(\rho||\mathcal{N})$. Then by $\lim_{t\to\infty} D(T_t(\rho)||E(\rho)) = 0$ and Gronwall's Lemma, MLSI is equivalent to exponential decay of relative entropy (see [22,4])

$$D(T_t(\rho)||\mathcal{N}) \leq e^{-\lambda t} D(\rho||\mathcal{N}) \,, \forall \rho \in S_B(\mathcal{A}). \tag{7}$$

The intuition here is that for non-ergodic semigroups, the semigroup flow $T_t(\rho)$ for an initial state ρ does not converge to one unique equilibrium state, but to its conditional expectation $E(\rho)$. Thus only the relative entropy $D(T_t(\rho)||\mathcal{N}) = D(T_t(\rho)||E(\rho))$ decay to 0, and the entropy $H(T_t(\rho)) = D(T_t(\rho)||1)$ does not converge to 0. Based on the non-ergodic MLSI, we introduce the complete bounded version of MLSI.

Definition 2.5. We say $(T_t)_{t\geq 0}$ satisfies λ-complete logarithmic Sobolev inequality (λ-CLSI) if $\mathrm{id}_{\mathcal{R}} \otimes T_t$ satisfy λ-MLSI for any finite von Neumann algebra $\mathcal{R}$.

Note that CLSI was studied in [22] under the definition that $\mathrm{id}_{M_n} \otimes T_t$ satisfy λ-MLSI for every matrix algebra M_n. Here in this paper, we will work with the potentially stronger definition that $\mathcal{R}$ can be any finite von Neumann algebra. (It remains open that whether these two definitions are equivalent.) The MLSI is a L_1-version of the Gross' logarithmic Sobolev inequality that is often stated for L_2-elements. For an ergodic symmetric Markov semigroup T_t, T_t is said to satisfy λ-logarithmic Sobolev inequality (λ-LSI) if for any $x \in \mathrm{dom}(A^{1/2})$ with $\|x\|_2 = 1$,

$$\lambda H(|x|^2) \leq 2\mathcal{E}(x,x) \,.$$

It was proved in [37, Section III.A.1] that all (finite dimensional) symmetric quantum Markov semigroup satisfies strong L_1-regularity: $4\mathcal{E}(\rho^{1/2}, \rho^{1/2}) \leq I(\rho)$. Thus we have λ-LSI $\Longrightarrow \lambda$-MLSI for ergodic symmetric Markov semigroups. On the other hand, it was pointed out in [22, Section 7.4] and [5, Theorem 5.1] that for non-ergodic cases, LSI does

not hold for the basic example such as $A = I - E$. This suggests that LSI does not hold for many non-ergodic cases neither the complete version. This is in contrast to MLSI and its complete version CLSI (see [22, Section 5] for a density result and [25,24] more recent progress in finite dimensions).

2.4. Noncommutative Wassersetin distance

Let $T_t : \mathcal{M} \to \mathcal{M}$ be a symmetric quantum Markov semigroup and $(\mathcal{A}, \hat{\mathcal{M}}, \delta)$ be a derivation triple for T_t. For simplicity of notation, we write τ for the trace on both $\mathcal{M}$ and $\hat{\mathcal{M}}$. For a state $\rho \in S(\mathcal{M})$, define the operator

$$[\rho]x := \int_0^1 \rho^s x \rho^{1-s} ds = R_\rho \circ f(\Delta_\rho)(x) .$$

Here R_ρ (resp. L_ρ) is the right (resp. left) multiplication operator and $\Delta_\rho = L_\rho R_\rho^{-1}$ is the modular operator of ρ. $f(\Delta_\rho)$ is the functional calculus of Δ_ρ for the function $f(w) = \int_0^1 \omega^s ds = (w-1)/\log w$. The inverse operator (on the support of ρ) is

$$[\rho]^{-1}x = R_\rho^{-1} \circ \frac{1}{f}(\Delta_\rho)x = J_{\log}^\rho(x) = \int_0^\infty (\rho+s)^{-1} x(\rho+s)^{-1} ds,$$

where $J_{\log}^\rho$ is the double operator integral for the function $f(t) = \log t$ and operator ρ. The last equality follows from $\frac{\ln x - \ln y}{x-y} = \int_0^\infty (x+s)^{-1}(y+s)^{-1} ds$. We define the *weighted L_2-(semi)norm* on $\hat{\mathcal{M}}$ by

$$\langle \xi, \eta \rangle_\rho := \langle \xi, [\rho]\eta \rangle_{L_2(\hat{\mathcal{M}},\tau)} = \int_0^1 \tau(\xi^* \rho^{1-s} \eta \rho^s) ds .$$

Denote $\hat{\mathcal{H}}_\rho \subset L_2(\hat{\mathcal{M}}, \rho)$ as the closure of $\delta(\mathcal{A}_\mathcal{E})$. Let I be an interval. Following [60], we say a curve $\gamma : (a,b) \to S(\mathcal{M})$ is *admissible* if

i) for any $a \in \mathcal{A}$, $s \mapsto \tau(a\gamma(s))$ is locally absolutely continuous.
ii) there exists $\xi \in L_{loc}^2((a,b), \hat{\mathcal{H}}_{\gamma(t)})$ such that

$$\frac{d}{ds}\tau(a\gamma(s)) = \langle \delta a, \xi(s) \rangle_\rho , \ a.e. \ s \in (a,b) \tag{8}$$

Such ξ is unique since $\delta(\mathcal{A})$ is dense in $\hat{\mathcal{H}}_\rho$ and we write this as $\xi(s) = D\gamma(s)$.

Definition 2.6. For $\rho, \sigma \in S(\mathcal{M})$, the noncommutative Wasserstein distance is defined as

$$W(\rho, \sigma) = \inf_{\gamma} \int_0^1 \|D\gamma(s)\|_{\gamma(s)} \, ds$$

where the infimum is taken over all admissible curves $\gamma : [0,1] \to S(\mathcal{M})$ such that $\gamma(0) = \rho, \gamma(1) = \sigma$.

We say an admissible curve $\gamma : [0,1] \to (S(\mathcal{M}), W)$ is a *geodesic* if γ attains the infimum of $W(\gamma(0), \gamma(1))$. We say that γ is a *geodesic with constant speed* if $W(\gamma(s), \gamma(t)) = |s - t| W(\gamma(0), \gamma(1))$. It was proved in [60, Lemma 4.19] that under the assumption that the smooth subalgebra $\mathcal{A}$ is dense and $L_1(\mathcal{M})$ is separable, then the infimum above can be taken to be over smooth curves.

For simplicity, we now illustrate the Riemannian metric for smooth curves on $S_B(\mathcal{M})$ as in [10]. The Wasserstein distance induces a pseudo-metric on $S_B(\mathcal{M})$: for $z \in \mathcal{M}$,

$$\|z\|_{g,\rho} := \inf\{ \|\xi\|_\rho \mid \delta^*([\rho]\xi) = z \},$$

where δ^* is the adjoint of $\delta : L_2(\mathcal{M}, \tau) \to L_2(\hat{\mathcal{M}}, \tau)$. The infimum is taken over all $\xi \in \hat{\mathcal{M}}$ satisfying the continuity equation $z = \delta^*([\rho]\xi)$. Here the L_2-closure of $\delta^*(\mathcal{A}\delta(\mathcal{A}))$ is exactly $(I - E)L_2(\mathcal{M}) = L_2(\mathcal{N})^\perp$, the orthogonal complement of $L_2(\mathcal{N})$. So for $z \notin L_2(\mathcal{N})^\perp$, $\|z\|_{g,\rho} = +\infty$. Thus we only need to consider the metric $\|\cdot\|_{g,\rho}$ restricted to

$$\mathcal{H} = \{a - E(a) \mid a = a^* \in \mathcal{M}\}$$

which is the horizontal direction on $S_B(\mathcal{M})$. Indeed, for any $z \in \mathcal{H}$ there exists a unique self-adjoint element $\xi \in \overline{\mathrm{ran}(\delta)} = \ker(\delta^*)^\perp \in L_2(\hat{\mathcal{M}})$ such that

$$z = \delta^*([\rho]\xi) , \quad \|z\|_{g,\rho} = \|\xi\|_\rho \tag{9}$$

(see [10, Theorem 7.3] and [22, Lemma 6.2]). Thus for an admissible smooth curve $\gamma : (a,b) \to S_B(\mathcal{M})$, we have

$$\gamma'(s) = \delta^*([\gamma(s)]D\gamma(s)) , \quad \|\gamma'(s)\|_{g,\gamma(s)} = \|D\gamma(s)\|_{\gamma(s)} .$$

The Wasserstein distance is then the (sub-)Riemannian distance induced by the metric $\langle \cdot, \cdot \rangle_{g,\rho}$,

$$W(\rho, \sigma) = \inf_{\gamma} \int_0^1 \|\gamma'(s)\|_{g,\gamma(s)} \, ds$$

where the infimum is taken over admissible smooth curve $\gamma \in C^1([0,1], S_B(\mathcal{M}))$. In the following we denote by $\mathcal{H}_\rho$ the closure of $\mathcal{H}$ with respect to the $\|\cdot\|_{g,\rho}$ norm. $\mathcal{H}_\rho$ should be thought of as the horizontal tangent space at the point $\rho \in S_B(\mathcal{M})$, equipped with

sub-Riemannian metric $\|\cdot\|_{g,\rho}$. The element $z \in \mathcal{H}_\rho$ is in one to one correspondence with $\xi \in \hat{\mathcal{H}}_\rho$ by the relation (9).

Let $F : S_B(\mathcal{M}) \to \mathbb{C}$ be a function. We say F admits a *(horizontal) gradient* at ρ if there exists a vector $\xi \in \hat{\mathcal{H}}_\rho$ such that for every smooth path $\rho : (-\varepsilon, \varepsilon) \to S_B(\mathcal{M})$ with $\rho(0) = \rho$,

$$\rho'(0) = \delta^*([\rho]\xi_0) \quad \Longrightarrow \quad \frac{d}{dt} F(\rho(t))|_{t=0} = \langle \xi, \xi_0 \rangle_\rho ,$$

and we write $\xi = \mathrm{grad}_\rho F$. By the relation (9), this is equivalent to the gradient for the metric $\|\cdot\|_{g,\rho}$ in the usual Riemannian sense,

$$\frac{d}{dt} F(\rho(t))|_{t=0} = \langle \rho'(0), \delta^*([\rho]\mathrm{grad}_\rho F) \rangle_{g,\rho} .$$

An admissible smooth curve $\gamma : I \to S_B(\mathcal{M})$ in the bounded density space is said to follow the path of *steepest descent* or gradient flow with respect to F if for any $a \in \mathcal{A}$ and $s \in (a, b)$

$$\frac{d}{ds} \tau(a\gamma(s)) = -\langle \delta(a), \mathrm{grad}_{\gamma(s)} F \rangle_{\gamma(s)} ,$$

or equivalently, $\gamma'(s) = -\delta^*([\gamma(s)]\mathrm{grad}_\gamma(s)F)$ weakly. One immediate consequence is that along a gradient flow γ,

$$\frac{dF(\gamma(s))}{ds} = -\|\delta^*([\gamma(s)]\mathrm{grad}_\gamma(s)F)\|_{g,\gamma(s)}^2 = -\|\mathrm{grad}_{\gamma(s)} F\|_{\gamma(s)}^2 . \tag{10}$$

Now we take $F(\rho) = H(\rho)$ as the entropy functional. Note that it is equivalent to take the relative entropy $D(\rho\|\mathcal{N})$ because an admissible curve $E(\gamma(s))$ is independent of s and $D(\gamma(s)\|\mathcal{N}) = H(\gamma(s)) - H(E(\gamma(s)))$. The next lemma shows that for $\rho \in S_B(\mathcal{A}_0)$, $\rho_t = T_t(\rho)$ is the gradient flow of H as well as other convenient properties of ρ_t. The key point is that it suffices to consider $\rho \in S_B(\mathcal{A}_0)$ for functional inequalities and we do not need assume curvature condition comparing to [60].

Lemma 2.7. *Let $\rho \in S_B(\mathcal{A}_0)$ and denote $\rho_t = T_t(\rho)$. Then*

i) *(ρ_t) is an admissible curve with $D(\rho_t) = \delta(\log \rho_t)$ and $\|D(\rho_t)\|_{\rho_t} = I(\rho_t)$.*
ii) *$t \mapsto I(\rho_t)$ is continuous and (ρ_t) is the gradient flow with respect to entropy H.*
iii) *For any t, $W(\rho_t, \rho) < \infty$ and $\lim_{t\to 0} W(\rho_t, \rho) = 0$.*
iv) *$\lim_{t\to\infty} \|\rho_t - E(\rho)\|_2 = 0$ and $\lim_{t\to\infty} D(\rho_t\|\mathcal{N}) = 0$.*

Proof. By assumption on $\mathcal{A}$, we have $T_t(\rho) \subset \mathcal{A} \cap \mathrm{dom}(A)$ and $\log \rho \in \mathrm{dom}(A^{1/2})$. Then we have the derivative

$$\frac{d}{dt}\rho_t = A\rho_t = \delta^*\delta(\rho_t) = \delta^*([\rho_t]\delta(\log\rho_t)) \ .$$

By definition (8), this implies $D(\rho_t) = \delta(\log\rho_t)$.

$$
\begin{aligned}
\|\delta(\log\rho_t)\|_{\rho_t}^2 &= \langle [\rho_t]\delta(\log\rho_t), \delta(\log\rho_t)\rangle \\
&= \langle [\rho_t][\rho_t]^{-1}\delta(\rho_t), \delta(\log\rho_t)\rangle \\
&= \tau(\delta(\rho_t)^*\delta(\log\rho_t)) = \mathcal{E}(\rho_t, \log\rho_t) = I(\rho_t)
\end{aligned}
$$

where we have used the derivation relation $\delta(\log\rho) = J_{\log}^\rho(\delta(\rho)) = [\rho]^{-1}\delta(\rho)$. The admissibility of (ρ_t) follows from the continuity of $t \mapsto I(\rho_t)$. Indeed, by assumption $\mu_1 1 \le \rho \le \mu_2 1$ and $A\rho \in L_2(\mathcal{M})$. Using the continuity of semigroup [16, Proposition 3.1], we have $\rho_t \mapsto \rho$ and $A\rho_t = T_t(A\rho) \mapsto A\rho$ in L_2. Since $f(x) = \log x$ is a Lipschitz continuous on $[\mu_1, \mu_2]$, $\lim_{t\to 0} \|\log\rho_t - \log\rho\|_2 = 0$ by [17, Corollary 7.5]. Then for the Fisher information,

$$
\begin{aligned}
\lim_{t\to 0} I(\rho_t) - I(\rho) &= \lim_{t\to 0} \tau(A\rho_t \log\rho_t) - \tau(A\rho\log\rho) \\
&\le \lim_{t\to 0} \tau(A\rho_t(\log\rho_t - \log\rho)) + \tau((A\rho_t - A\rho)\log\rho) \\
&= \lim_{t\to 0} \tau(T_t(A\rho)(\log\rho_t - \log\rho)) + \tau((T_t(A\rho) - A\rho)\log\rho) \\
&\le \lim_{t\to 0} \|T_t(A\rho)\|_2\|\log\rho_t - \log\rho\|_2 + \|T_t(A\rho) - A\rho\|_2\|\log\rho\|_2 = 0 \ .
\end{aligned}
$$

Applying semigroup property, we have $t \mapsto I(\rho_t)$ is continuous. For the gradient flow, given a self-adjoint $\beta = \delta^*([\rho]\xi_0)$,

$$
\begin{aligned}
\frac{d}{dt}H(\rho + t\beta)|_{t=0} &= \tau(\beta\log\rho) = \langle\delta^*([\rho]\xi_0), \log\rho\rangle_\tau \\
&= \langle[\rho]\xi_0, \delta(\log\rho)\rangle_\tau = \langle\xi_0, \delta(\log\rho)\rangle_\rho \ .
\end{aligned}
$$

Thus $\mathrm{grad}_\rho H = \delta(\log\rho)$ and the gradient flow for $H(\cdot)$ is given by the equation

$$
\begin{aligned}
\rho'(t) &= -\delta^*([\rho(t)]\mathrm{grad}_{\rho(t)}H) = -\delta^*([\rho(t)]\delta(\log\rho(t))) \\
&= -\delta^*([\rho(t)][\rho(t)]^{-1}\delta(\rho(t))) = -A(\rho(t)) \ ,
\end{aligned}
$$

whose solution is the semigroup flow $\rho(t) = T_t(\rho(0))$. For iii), since $s \mapsto \rho_s$ is admissible

$$\lim_{t\to 0} W(\rho_t, \rho) \le \lim_{t\to 0}\int_0^t \|D\rho_s\|_{\rho_s} \, ds = \lim_{t\to 0}\int_0^t I(\rho_s)^{1/2}ds = 0 \ .$$

For iv), we first show the L_2-convergence. Consider A as a positive self-adjoint operator on $L_2(\mathcal{M})$ and denote e_s (resp. e_0) as the spectral projection for the spectrum $[0, s)$ (resp. $\{0\}$). Clearly, $e_0(L_2(\mathcal{M})) = L_2(\mathcal{N})$. Write $\overset{\circ}{\rho} = \rho - E(\rho)$. We have

$$\overset{\circ}{\rho} \in e_0^{\perp} \ , \ \| \overset{\circ}{\rho} \|_2 \leq \| \rho \|_2 \ , \ T_t(\rho) - E(\rho) = T_t(\overset{\circ}{\rho}) \ .$$

Then $\lim_{s \to 0} \| e_s(\overset{\circ}{\rho}) \|_2 = 0$. For any $\epsilon > 0$, we can find $s > 0$ and then large enough t such $\| e_s(\overset{\circ}{\rho}) \|_2 < \epsilon$ and $e^{-st} \| \rho \|_2 < \epsilon$. Thus

$$\| T_t(\rho) - E(\rho) \|_2 = \| T_t(\overset{\circ}{\rho}) \|_2 \leq \| T_t(e_s(\overset{\circ}{\rho})) \|_2 + \| T_t(\overset{\circ}{\rho} - e_s(\overset{\circ}{\rho})) \|_2$$

$$\leq \epsilon + e^{-st} \| \rho \|_2 \leq 2\epsilon \ .$$

Therefore $\lim_{t \to \infty} \| T_t(\rho) - E(\rho) \|_2 = 0$. This further implies $\lim_{t \to \infty} \| T_t(\rho) - E(\rho) \|_1 = 0$ and by Lemma A.2,

$$\lim_{t \to \infty} D(T_t(\rho) || \mathcal{N}) = D(E(\rho) || \mathcal{N}) = 0 \ . \quad \square$$

3. Fisher monotonicity and CB-return time

3.1. Monotonicity of Fisher information

Our first ingredient is the monotonicity of Fisher information, which can be equivalently characterized by the following conditions.

Proposition 3.1. *Let $\lambda \in \mathbb{R}$. For a state $\rho \in S(\mathcal{M})$, denote $T_t(\rho) = \rho_t$. The following conditions are equivalent*

i) *for any $\rho \in S_B(\mathcal{A}_0)$ and $t \geq 0$,*

$$I(\rho_t) \leq e^{-2\lambda t} I(\rho).$$

ii) *for any $\rho \in S_H(\mathcal{M})$ and $s, t \geq 0$,*

$$D(\rho_t || \mathcal{N}) - D(\rho_{s+t} || \mathcal{N}) \leq e^{-2\lambda t} (D(\rho || \mathcal{N}) - D(\rho_s || \mathcal{N})) \ .$$

iii) *for any $\rho \in S_H(\mathcal{M})$ and $s, t \geq 0$,*

$$H(\rho_t) - H(\rho_{s+t}) \leq e^{-2\lambda t} (H(\rho) - H(\rho_s)) \ .$$

Proof. Let $\rho \in S_B(\mathcal{A}_0)$. Combined Lemma (2.7) with [60, Proposition 5.23]), we have for $\rho \in S_B(\mathcal{A}_0)$

$$D(\rho || \mathcal{N}) - D(\rho_t || \mathcal{N}) = H(\rho) - H(\rho_t) = \int\limits_0^t I(\rho_u) du,$$

$$I(\rho) = \lim_{t \to 0} \frac{D(\rho || \mathcal{N}) - D(\rho_t || \mathcal{N})}{t}. \tag{11}$$

Then ii) follows from i) since for $\rho \in S_B(\mathcal{A}_0)$,

$$D(\rho||\mathcal{N}) - D(\rho_s||\mathcal{N}) = \int_0^s I(\rho_u)du \geq \int_0^s e^{2\lambda t}I(\rho_{t+u})du$$

$$= e^{2\lambda t}\Big(\int_t^{s+t} I(\rho_u)du\Big) = e^{2\lambda t}\big(D(\rho_t||\mathcal{N}) - D(\rho_{s+t}||\mathcal{N})\big) \, .$$

For general $\rho \in S_H(\mathcal{M})$, we use the approximation in Lemma A.3. On the other hand, i) follows from ii) since for $\rho \in S_B(\mathcal{A}_0)$,

$$I(\rho) = \lim_{s\to 0} \frac{D(\rho||\mathcal{N}) - D(\rho_s||\mathcal{N})}{s}$$

$$\geq \lim_{s\to 0} e^{2\lambda t} \cdot \frac{D(\rho_t||\mathcal{N}) - D(\rho_{s+t}||\mathcal{N})}{s} \geq e^{2\lambda t}I(\rho_t) \, .$$

The equivalence to iii) follows from the fact that $D(\rho||\mathcal{N}) = H(\rho) - H(E(\rho))$ for $\rho \in S_H(\mathcal{M})$ and $E(\rho) = E(T_t(\rho))$. $\square$

Definition 3.2. We say a semigroup T_t is λ-Fisher monotone for $\lambda \in \mathbb{R}$ (in short, λ-FM) if T_t satisfies one of the above conditions in Proposition (3.1). We say T_t is λ-complete Fisher monotone (λ-CFM) if for any finite von Neumann algebra $\mathcal{R}$, $id_\mathcal{R} \otimes T_t$ is λ-FM. For $\lambda = 0$, we simply say T_t is (complete) Fisher monotone.

The idea of following proposition goes back to the Γ-calculus in [2].

Proposition 3.3. *For $\lambda > 0$, λ-FM implies λ-MLSI.*

Proof. For $\rho \in S_B(\mathcal{A}_0)$, denote $f(t) = D(\rho||\mathcal{N}) - D(\rho_t||\mathcal{N})$ and hence $I(\rho_t) = f'(t)$. Then λ-FM means that

$$f'(t) \leq e^{-2\lambda t}f'(0)$$

Integrating both sides from 0 to t,

$$D(\rho||\mathcal{N}) - D(\rho_t||\mathcal{N}) \leq \frac{e^{-2\lambda t} - 1}{-2\lambda}I(\rho)$$

Taking $t \to \infty$,

$$2\lambda D(\rho||\mathcal{N}) = \lim_{t\to\infty} 2\lambda(D(\rho||\mathcal{N}) - D(\rho_t||\mathcal{N})) \leq \lim_{t\to\infty}(1 - e^{-2\lambda t})I(\rho) = I(\rho) \, ,$$

which is λ-MLSI. Here we used the assumption $\lambda > 0$ and the property $\lim_{t\to\infty} D(\rho_t||\mathcal{N}) = 0$ from Lemma 2.7. $\square$

3.2. Complete bounded return time

Let $\mathcal{M}$ be a finite von Neumann algebra and $\mathcal{N} \subset \mathcal{M}$ be a subalgebra. The conditional L_1 space $L_\infty^1(\mathcal{N} \subset \mathcal{M})$ is defined as the completion of $\mathcal{M}$ with respect to the norm

$$\|x\|_{L_\infty^1(\mathcal{N} \subset \mathcal{M})} = \sup_{a,b \in L_2(\mathcal{N}), \|a\|_2 = \|b\|_2 = 1} \|axb\|_1 ,$$

where the supremum takes over all $a, b \in L_2(\mathcal{N})$ with $\|a\|_2 = \|b\|_2 = 1$. The operator space structure of $L_\infty^1(\mathcal{N} \subset \mathcal{M})$ is given by

$$M_n(L_\infty^1(\mathcal{N} \subset \mathcal{M})) = L_\infty^1(M_n(\mathcal{N}) \subset M_n(\mathcal{M}))$$

(see [35] and [23, Appendix]). We consider again $T_t : \mathcal{M} \to \mathcal{M}$ be a symmetric quantum Markov semigroup and $\mathcal{N}$ be the fixed point subalgebra with conditional expectation E. We define the complete bounded (CB) return time of T_t as follows

$$t_{cb} = \inf\{ t \geq 0 \mid \|T_t - E : L_\infty^1(\mathcal{N} \subset \mathcal{M}) \to L_\infty(\mathcal{M})\|_{cb} \leq 1/2\}$$

If such t does not exist, we write $t_{cb} = +\infty$. Recall the following lemma from [22].

Lemma 3.4 (*Lemma 3.15 of [22]*). *Let $T : \mathcal{M} \to \mathcal{M}$ be a unital completely positive $\mathcal{N}$-bimodule map such that*

$$\|T - E : L_\infty^1(\mathcal{N} \subset \mathcal{M}) \to \mathcal{M}\|_{cb} \leq \frac{1}{2} .$$

Then $T \geq_{cp} \frac{1}{2}E$, i.e. $T - \frac{1}{2}E$ is completely positive.

We refer [22] for the complete proof and illustrate here the argument for the ergodic case. Namely, we consider $\mathcal{N} = \mathbb{C}1$ and $L_\infty^1(\mathcal{N} \subset \mathcal{M}) = L_1(\mathcal{M})$. The CB return time becomes

$$t_{cb} = \inf\{ t \geq 0 \mid \|T_t - E : L_1(\mathcal{M}) \to L_\infty(\mathcal{M})\|_{cb} \leq 1/2\} .$$

This completely bounded norm is by no means abstract. Indeed, by Effros-Ruan Theorem (see [19] and also [6]),

$$\|T : L_1(\mathcal{M}) \to L_\infty(\mathcal{M})\|_{cb} = \|C_T\|_{\mathcal{M}^{op} \overline{\otimes} \mathcal{M}}$$

where C_T is the kernel of T (also called Choi matrix, in finite dimensions) given by the relation

$$T(a) = \tau \otimes \mathrm{id}(C_T(a \otimes 1)) , a \in L_1(\mathcal{M}) \cong (\mathcal{M}^{op})_* .$$

Here $\mathcal{M}^{op}$ is the opposite algebra of $\mathcal{M}$. Moreover, the correspondence $T \leftrightarrow C_T$ is also order preserving: T is completely positive if and only if C_T as a operator is positive in $\mathcal{M}^{op} \overline{\otimes} \mathcal{M}$. In particular, the conditional expectation onto scalars $E_\tau(a) = \tau(a)1$ has kernel as the identity $1 \otimes 1 \in \mathcal{M}^{op} \overline{\otimes} \mathcal{M}$. For this special case,

$$\| T - E_\tau : L_1(\mathcal{M}) \to \mathcal{M} \|_{cb} \leq 1/2 \Longleftrightarrow \| C_T - 1 \otimes 1 \|_{\mathcal{M}^{op} \overline{\otimes} \mathcal{M}} \leq 1/2$$

$$\Longrightarrow C_T \leq \frac{1}{2} 1 \otimes 1 \iff C_T \geq_{cp} \frac{1}{2} C_{E_\tau} ,$$

where the implication "$\Rightarrow$" is evident from spectrum calculus for a self-adjoint operator C_T. This proves the above lemma for the special case $\mathcal{N} = \mathbb{C}1$. The general case for non-trivial $\mathcal{N}$ is an extension for bimodule maps.

The next lemma shows t_{cb} is the half-life for the decay of relative entropy.

Lemma 3.5. *Let $\mathcal{N} \subset \mathcal{M}$ be a subalgebra and E be the condition expectation onto $\mathcal{N}$. Suppose for $\alpha \in (0,1)$, $\Phi - \alpha E$ is a positive map and $\Phi(L_1(\mathcal{N})) \subset L_1(\mathcal{N})$. Then for any $\rho \in S(\mathcal{M})$,*

$$D(\Phi(\rho)||\mathcal{N}) \leq (1-\alpha)D(\rho||\mathcal{N}) . \tag{12}$$

If in additional, $\Phi - \alpha E$ is a completely positive map, the same assertion holds for $\Phi \otimes \mathrm{id}_{\mathcal{R}}$.

Proof. Define $\Psi := \frac{1}{1-\alpha}(\Phi - \alpha E)$. By assumption that $\Phi - \alpha E$ is positive, Ψ is a positive trace preserving map such that $\Psi(L_1(\mathcal{N})) \subset L_1(\mathcal{N})$. Thus $\Phi = (1-\alpha)\Psi + \alpha E$. Note that the data processing inequality holds for positive trace preserving maps [32]. Then by the convexity of relative entropy and the data processing inequality of $D(\cdot||\mathcal{N})$ give

$$D(\Phi(\rho)||\mathcal{N}) = D((1-\alpha)\Psi(\rho) + \alpha E(\rho)||\mathcal{N}) \leq (1-\alpha)D(\Psi(\rho)||\mathcal{N}) + \alpha D(E(\rho)||\mathcal{N})$$

$$= (1-\alpha)D(\Psi(\rho)||\mathcal{N}) \leq (1-\alpha)D(\rho||\mathcal{N}) .$$

The same argument applies to $\Phi \otimes \mathrm{id}_{\mathcal{R}}$. $\square$

We now prove our main technical theorem that (complete) Fisher monotonicity plus CB-return time implies MLSI (resp. CLSI). Define the function

$$\kappa(\lambda, t) = \begin{cases} \frac{1}{4t}, & \text{if } \lambda = 0 \\ \frac{\lambda}{2(1-e^{-2\lambda t})}, & \text{if } \lambda \neq 0. \end{cases}$$

For each t, $\lambda \mapsto \kappa(\lambda, t)$ is continuous at 0.

Theorem 3.6. *Let $T_t : \mathcal{M} \to \mathcal{M}$ be a symmetric quantum Markov semigroup. Suppose*

i) *T_t satisfies λ-FM for some $\lambda \in \mathbb{R}$*

ii) T_t has finite CB-return time $t_{cb} < \infty$.

Then T_t-satisfies $\kappa(\lambda, t_{cb})$-MSLI. The same assertions hold replacing "FM" with "CFM" and "MLSI" with "CLSI".

Proof. Write $t_{cb} = t_0$. As a consequence of Lemma 3.5, we have

$$D(T_{t_0}(\rho)||\mathcal{N}) \leq \frac{1}{2}D(\rho||\mathcal{N})$$

Write $T_t(\rho) = \rho_t$. For $\lambda < 0$, we have

$$\frac{1}{2}D(\rho||\mathcal{N}) \leq D(\rho||\mathcal{N}) - D(\rho_{t_0}||\mathcal{N}) = \int\limits_0^{t_0} I(\rho_s)ds$$

$$\leq \Big(\int\limits_0^{t_0} e^{-2\lambda s}ds\Big)I(\rho) = \frac{e^{-2\lambda t_0}-1}{-2\lambda}I(\rho)$$

where we used λ-FM in the second inequality. The argument above remains valid for $\lambda = 0$ and $T_t \otimes \mathrm{id}_{\mathcal{R}}$. This completes the proof. $\square$

Remark 3.7. For the ergodic classical Markov semigroups, it was proved by Diaconis and Saloff-Coste in [18, Theorem 3.10] that the L_∞-return time in the ergodic setting (the complete boundedness is automatic here)

$$t_\infty := \{t \geq 0 \mid \ \|T_t - E : L_1(\Omega) \to L_\infty(\Omega)\| \leq 1/e^2\}$$

itself implies $\frac{1}{t_\infty}$-LSI, which further implies MLSI. Nevertheless, their argument went through hypercontractive estimate that does not apply to non-ergodic or non-commutative setting.

The CB-return time can be estimated by standard argument.

Proposition 3.8. Let $T_t = e^{-At} : \mathcal{M} \to \mathcal{M}$ be a symmetric quantum Markov semigroup and $\mathcal{N}$ be its fixed-point subalgebra. Suppose

i) for some $t_0 \geq 0$, $\|T_{t_0} : L_\infty^1(\mathcal{N} \subset \mathcal{M}) \to L_\infty(\mathcal{M})\|_{cb} \leq C$.
ii) the generator A has spectral gap $\sigma > 0$, that is

$$\|A^{-1}(\mathrm{id} - E) : L_2(\mathcal{M}) \to L_2(\mathcal{M})\| \leq \sigma^{-1}.$$

Then for $t \geq t_0$,

$$\|T_t - E : L_\infty^1(\mathcal{N} \subset \mathcal{M}) \to L_\infty(\mathcal{M})\|_{cb} \leq Ce^{-\sigma(t-t_0)}.$$

As a consequence, $t_{cb} \leq \sigma^{-1}(\log 2C) + t_0$.

Proof. Note that $T_t - E$ is an $\mathcal{N}$-bimodule map. We have for $t \geq t_0$,

$$
\begin{aligned}
&\| T_t - E : L^1_\infty(\mathcal{N} \subset \mathcal{M}) \to L_\infty(\mathcal{M}) \|_{cb} \\
={}& \| T_{\frac{t}{2}} - E : L^1_\infty(\mathcal{N} \subset \mathcal{M}) \to L^2_\infty(\mathcal{N} \subset \mathcal{M}) \|_{cb}^2 \\
\leq{}& \| T_{\frac{t_0}{2}} : L^1_\infty(\mathcal{N} \subset \mathcal{M}) \to L^2_\infty(\mathcal{N} \subset \mathcal{M}) \|_{cb}^2 \| T_{\frac{t-t_0}{2}} \\
&- E : L^2_\infty(\mathcal{N} \subset \mathcal{M}) \to L^2_\infty(\mathcal{N} \subset \mathcal{M}) \|_{cb}^2 \\
={}& \| T_{t_0} : L^1_\infty(\mathcal{N} \subset \mathcal{M}) \to L_\infty(\mathcal{M}) \|_{cb} \| T_{\frac{t-t_0}{2}} - E : L_2(\mathcal{M}) \to L_2(\mathcal{M}) \|_{cb}^2 \\
={}& Ce^{-\sigma(t-t_0)} \, ,
\end{aligned}
$$

Here the first equality uses [22, Lemma 3.13] and the third equality uses [22, Lemma 3.12]. □

The above estimates have the following two corollaries. The first one is the non-ergodic version of [22, Proposition 3.2]. It basically says that the spectral gap plus a non-ergodic Varopoulos dimension condition implies finite CB-return time.

Lemma 3.9. *Let $T_t : \mathcal{M} \to \mathcal{M}$ be a symmetric quantum Markov semigroup and $\mathcal{N} \subset \mathcal{M}$ be the fixed-point subalgebra. Suppose*

i) $\| T_t : L^1_\infty(\mathcal{N} \subset \mathcal{M}) \to L_\infty(\mathcal{M}) \|_{cb} \leq ct^{-d/2}$ *for some $c, d > 0$ and all $0 < t < 1$;*
ii) *the generator A has spectral gap $\sigma > 0$*

Then the CB-return time satisfies

$$
t_{cb} \leq \frac{1}{2} + \frac{d-1}{2} \log 2 + \frac{1}{\sigma} \log c
$$

Proof. Choose $t_0 = 1/2$ in Lemma 3.8. □

The second case is related to von Neumann subalgebra index introduced by Pimsner and Popa [50]. Recall that for two states ρ, ω, the maximal relative entropy is

$$
D_\infty(\rho \| \omega) = \log \inf \{ \, \alpha > 0 \mid \rho \leq \alpha \omega \, \} \, .
$$

For an inclusion $\mathcal{N} \subset \mathcal{M}$ of finite von Neumann algebras, the maximal relative entropy D_∞ of $\mathcal{M}$ to $\mathcal{N}$ and its CB-version $D_{\infty,cb}$ is defined as

$$
D_\infty(\mathcal{M} \| \mathcal{N}) = \sup_{\rho \in S(\mathcal{M})} D_\infty(\rho \| \mathcal{N}) \, , \quad D_{\infty,cb}(\mathcal{M} \| \mathcal{N}) = \sup_m D_\infty(M_m(\mathcal{M}) \| M_m(\mathcal{N}))
$$

It was proved in [23, Theorem 3.9] that

$$D_{\infty,cb}(\mathcal{M}||\mathcal{N}) = \log \| \mathrm{id} : L^1_\infty(\mathcal{N} \subset \mathcal{M}) \to L_\infty(\mathcal{M}) \|_{cb} \ .$$

The next proposition gives the estimate of t_{cb} given that $D_{\infty,cb}(\mathcal{M}||\mathcal{N})$ is finite and spectral gap is positive.

Proposition 3.10. *Let $T_t : \mathcal{M} \to \mathcal{M}$ be a symmetric quantum Markov semigroup and $\mathcal{N}$ be its fixed-point subalgebra. Suppose $D_{cb,\infty}(\mathcal{M}||\mathcal{N}) < \infty$ is finite and T_t has spectral gap $\sigma > 0$. Then*

$$\| T_t - E : L^1_\infty(\mathcal{N} \subset \mathcal{M}) \to L_\infty(\mathcal{M}) \|_{cb} \le e^{D_{cb,\infty}(\mathcal{M}||\mathcal{N})} e^{-\sigma t} \ .$$

As a consequence, $t_{cb} \le \sigma^{-1}(D_{cb,\infty}(\mathcal{M}||\mathcal{N}) + \log 2)$.

Proof. Choose $t_0 = 0$ in Lemma 3.8. □

The maximal relative entropy $D_{cb,\infty}(\mathcal{M}||\mathcal{N})$ connects to the von Neumann algebra subalgebra index and is explicit for many examples. It was proved in [23, Theorem 3.1] that $D_\infty(\mathcal{M}||\mathcal{N}) = \log \lambda(\mathcal{M} : \mathcal{N})^{-1}$ for $\mathcal{M}, \mathcal{N}$ being II_1 factors or finite dimensional, where $\lambda(\mathcal{M} : \mathcal{N})$ is the Pimsner-Popa index in [50]. In particular, for II_1 factors, $D_{cb,\infty}(\mathcal{M}||\mathcal{N}) = \log[\mathcal{M} : \mathcal{N}]$ where $[\mathcal{M} : \mathcal{N}]$ is the Jones subfactor index; for $\mathcal{M}, \mathcal{N}$ finite dimensional, the explicit formula of $D_\infty(\mathcal{M}||\mathcal{N})$ is calculated in [50, Theorem 6.1], from which $D_{cb,\infty}(\mathcal{M}||\mathcal{N})$ are also known. For example,

$$D_{cb,\infty}(M_n||\mathbb{C}) = 2\log n \ , \ \ D_{cb,\infty}(M_n||l^n_\infty) = \log n \ , \ \ D_{cb,\infty}(l^n_\infty||\mathbb{C}) = \log n.$$

For any $\mathcal{N} \subset M_n$, $D_{cb,\infty}(M_n||\mathcal{N}) \le D_{cb,\infty}(M_n||\mathbb{C}1) = 2\log n$.

3.3. Entropy Ricci curvature bound

We shall now discuss the connection between Fisher monotonicity and Ricci curvature lower bound and give a non-egordic version of Bakry-Emery theorem. Following [20], we call Ricci curvature bound defined through geodesic convexity of entropy H as *entropy Ricci curvature bound*. We first review the different formulations of entropy Ricci curvature bound discussed in [60,11,15]. For a function $f : [0, a) \to \infty$, we introduce the notation

$$\frac{d^+}{dt} f = \limsup_{t \to 0} \frac{1}{t}(f(t) - f(0)) \ .$$

Recall that $S_H(\rho) = \{\rho \in S(\mathcal{M}) \mid H(\rho) < \infty\}$ is the state space with finite entropy and we write $\rho_t = T_t(\rho)$.

Definition 3.11. Let $T_t = e^{-At}$ be a symmetric quantum Markov semigroup and let $(\mathcal{A}, \delta, \hat{\mathcal{M}})$ be a derivation triple of T_t. For $\lambda \in \mathbb{R}$, define the following conditions

i) Gradient Estimate: we say T_t satisfies a λ-gradient estimate (λ-GE) if for any $\rho \in S(\mathcal{M})$ and $x \in \text{dom}(A^{1/2})$ with $E(x) = 0$,

$$\| \delta(T_t(x)) \|_\rho^2 \leq e^{-2\lambda t} \| \delta(x) \|_{\rho_t}^2 , \forall t \geq 0 .$$

ii) Evolution Variational Inequality: we say T_t satisfies a λ-evolution variational inequality (λ-EVI) if for all $\rho, \sigma \in S_H(\mathcal{M})$ with $W(\rho, \sigma) < \infty$ and $t \geq 0$

$$\frac{1}{2} \frac{d^+}{dt} W(\rho_t, \sigma)^2 + \frac{\lambda}{2} W(\rho_t, \sigma)^2 + H(\rho_t) \leq H(\sigma) .$$

iii) Displacement Convexity: we say the entropy functional H is geodesically λ-convex if for any constant speed geodesic $\gamma : [0, 1] \to (S_H(\mathcal{M}), W)$,

$$H(\gamma(s)) \leq (1 - s)H(\gamma(0)) + sH(\gamma(1)) - \frac{\lambda(1 - s)s}{2} W(\gamma(0), \gamma(1))^2 .$$

When $\mathcal{M}$ is a finite dimensional C^*-algebra and T_t being a primitive semigroup (including non-symmetric cases with detailed balance condition), all three of the above conditions are proved to be equivalent and are referred to as a λ-*Ricci lower bound* in [15,11]. For finite von Neumann algebras $\mathcal{M}$, it has been proved in [60, Theorem 7.12] that

(i) $\Rightarrow$ {W is non-degenerate and (ii)} $\Rightarrow$ {$(S_H(\mathcal{M}), W)$ is a geodesic space and (iii)}

For this reason, we take the gradient estimate condition λ-GE as our working definition of entropy Ricci curvature bound.

Remark 3.12. For EVI and displacement convexity above, it is equivalent to replace the entropy $H(\rho)$ by the relative entropy $D(\rho||\mathcal{N})$. This is because for $\rho \in S_H(\mathcal{M})$, $D(\rho||E(\rho)) = H(\rho) - H(E(\rho)) < \infty$. For λ-EVI, $W(\rho, \sigma) < \infty$ implies $E(\rho) = E(\sigma)$ and hence

$$\frac{1}{2} \frac{d^+}{dt} W(\rho_t, \sigma)^2 + \frac{\lambda}{2} W(\rho_t, \sigma)^2 + D(\rho_t||\mathcal{N}) \leq D(\sigma||\mathcal{N}) . \tag{13}$$

For λ-displacement convexity, $E(\gamma(s)) = E(\gamma(t))$ for any admissible curve γ and hence

$$D(\gamma(s)||\mathcal{N}) \leq (1 - s)D(\gamma(0)||\mathcal{N}) + sD(\gamma(1)||\mathcal{N}) - \frac{\lambda(1 - s)s}{2} W(\gamma(0), \gamma(1))^2 .$$

Remark 3.13. A semigroup T_t can admit distinct derivation triples $(\mathcal{A}, \delta, \hat{\mathcal{M}})$. For example, let M_2 be 2×2 matrix algebra and consider the depolarizing semigroup

$$D_t : M_2 \to M_2, D_t(\rho) = e^{-t}\rho + (1 - e^{-t})\tau(\rho)1 \ ,$$

where τ is the normalized trace $\tau(\rho) = \frac{1}{2}\text{Tr}(\rho)$. It was discussed in [11, Section 5.6] that D_t admits a derivation

$$\delta : M_2 \to \oplus_{j=1}^{3} M_2 \ , \delta(a) = \frac{1}{2\sqrt{2}}(i[X,a], i[Y,a], i[Z,a]) \ ,$$

where X, Y, Z are Pauli matrices. This follows from that the depolarizing map E is an average of unitary conjugation by Pauli matrices,

$$E(\rho) = \frac{tr(\rho)}{2}1 = \frac{1}{4}(\rho + X\rho X + Y\rho Y + Z\rho Z) \ .$$

On the other hand, the depolarizing map E can also be seen as the following average of unitary conjugations over the unitary group $U(2) \subset M_2$,

$$E(\rho) = \int_{U(2)} u^*\rho u \, d\mu(u)$$

where μ is the Haar measure on $U(2)$. Then one can construct an alternative derivation

$$\tilde{\delta} : M_2 \to L_\infty(U(2), M_2) \ , \tilde{\delta}(a)(u) = i[u, a] \ ,$$

where $L_\infty(U(2), M_2)$ is the M_2-valued function on the Lie group $U(2)$. For more examples of distinct derivation triples, see Example 4.14.

The next proposition shows that the gradient estimate is independent of the choice of derivation triple $(\mathcal{A}, \hat{\mathcal{M}}, \delta)$.

Proposition 3.14. *The definition of the gradient estimate is independent of the choice of derivation.*

Proof. We show that the norm

$$\| \delta(x) \|^2_{L_2(\hat{\mathcal{M}}, \rho)} = \int_0^1 \tau(\delta(x)^*\rho^s \delta(x)\rho^{1-s}) ds$$

is independent of δ. Recall that the Dirichlet algebra $\mathcal{A}_{\mathcal{E}} = \text{dom}(A^{1/2}) \cap \mathcal{M}$ is a core for δ and closed under C^1-functional calculus. For $x, \rho \in \mathcal{A}_{\mathcal{E}}$, we have $\rho^s \in \mathcal{A}_{\mathcal{E}}$ and by Leibniz rule

$$\rho^s \delta(x) = \delta(x\rho^s) - \delta(\rho^s)x \ , x \in \text{dom}(A^{1/2}) \ .$$

Then for each $s \in [0,1]$,

$$
\begin{aligned}
\tau(\delta(x)^* \rho^s \delta(x)\rho^{1-s}) =& \tau(\delta(x)^* \delta(\rho^s x)\rho^{1-s}) - \tau(\delta(x)^* \delta(\rho^s)x\rho^{1-s}) \\
=& \tau\Big(E_M(\delta(x)^* \delta(\rho^s x)\rho^{1-s})\Big) - \tau\Big(E_M(\delta(x)^* \delta(\rho^s)x\rho^{1-s})\Big) \\
=& \tau\Big(E_M(\delta(x)^* \delta(\rho^s x))\rho^{1-s}\Big) - \tau\Big(E_M(\delta(x)^* \delta(\rho^s))x\rho^{1-s}\Big) \\
=& \tau\Big(\Gamma(x, \rho^s x)\rho^{1-s}\Big) - \tau\Big(\Gamma(x, \rho^s)x\rho^{1-s}\Big),
\end{aligned}
$$

which is completely determined by gradient form Γ. We now show for general ρ, x, the norm $\{\|\delta(x)\|_{L_2(\hat{\mathcal{M}},\rho)}\}$ can be approximated by $\rho, x \in \mathcal{A}_{\mathcal{E}}$. For $x \in \mathrm{dom}(A^{1/2})$, we chose a sequence $x_n \to x$ in the graph norm of δ. In particular, $\delta(x_n) \to \delta(x)$ in L_2. Then for $\rho \in \mathcal{A}_{\mathcal{E}}$,

$$
\begin{aligned}
\lim_{n\to\infty} \tau(\delta(x_n)^* \rho^s \delta(x_n)\rho^{1-s}) &= \lim_{n\to\infty} \|\rho^{s/2}\delta(x_n)\rho^{(1-s)/2}\|_2^2 \\
&= \|\rho^{s/2}\delta(x)\rho^{(1-s)/2}\|_2^2 = \tau(\delta(x)^* \rho^s \delta(x)\rho^{1-s})
\end{aligned}
$$

For any $\rho \in S(\mathcal{M})$, we take sequence $\rho_n = \rho \wedge n \in \mathcal{A}_{\mathcal{E}}$ and $\rho_n \nearrow \rho$ in L_1. Then for any $x \in \mathrm{dom}(A^{1/2})$, we apply the Fatou lemma

$$
\begin{aligned}
\limsup_{n\to\infty} \tau(\delta(x)^* \rho_n^s \delta(x)\rho_n^{1-s}) \leq& \limsup_{n\to\infty} \tau(\delta(x)^* \rho_n^s \delta(x)\rho^{1-s}) \\
\leq& \tau(\delta(x)^* \rho^s \delta(x)\rho^{1-s}) \leq \liminf_{n\to\infty} \tau(\delta(x)^* \rho_n^s \delta(x)\rho_n^{1-s})
\end{aligned}
$$

which implies $\tau(\delta(x)^* \rho_n^s \delta(x)\rho_n^{1-s}) \nearrow \tau(\delta(x)^* \rho^s \delta(x)\rho^{1-s})$. Then by monotone convergence theorem, $\lim_n \|\delta(x)\|_{L_2(\hat{\mathcal{M}},\rho_n)}^2 = \|\delta(x)\|_{L_2(\hat{\mathcal{M}},\rho)}^2$. That completes the proof. $\quad\square$

The next proposition shows that entropy Ricci curvature bound implies Fisher monotonicity.

Proposition 3.15. *For any symmetric quantum Markov semigroup $T_t : \mathcal{M} \to \mathcal{M}$ and $\lambda \in \mathbb{R}$, λ-GE implies λ-FM.*

Proof. Let $\rho \in S_B(\mathcal{A}_0)$ and $\rho_t = T_t(\rho)$ be the semigroup path. By Lemma 2.7, (ρ_t) is an admissible curve with

$$
\|D\rho_t\|^2 = \|\delta(\log \rho_t)\|_{\rho_t}^2 = I(\rho_t) ,
$$

and $t \mapsto I(\rho_t)$ is continuous. Then it follows from [60, Theorem 6.9] that for any $s > 0$

$$
I(T_{s+t}(\rho)) = I(T_s(\rho_t)) = \|DT_s(\rho_t)\|_{T_s(\rho_t)}^2 \leq e^{-2\lambda s} \|D\rho_t\|_{\rho_t}^2 = e^{-2\lambda s} I(\rho_t) . \quad\square
$$

For $\lambda > 0$, the above Proposition and Proposition 3.3 combined gives λ-GE $\Rightarrow \lambda$-FM $\Rightarrow \lambda$-MLSI, which is a noncommutative non-ergodic version of Bakry-Emery theorem. In the following, we take another approach using Otto-Villani's HWI inequality introduced in [49]. The quantum HWI inequality is obtained in [15, Corollary 2] for finite dimensional ergodic case (see also [11]). For finite von Neumann algebra, this idea is also used in [60, Proposition 7.9]. Here the major difference to [60] is that we do not need to assume λ-GE for some $\lambda > 0$. The proof of the following two propositions follows [49].

Proposition 3.16. *Let T_t be a semigroup satisfying λ-EVI for $\lambda \in \mathbb{R}$: for any $\rho, \sigma \in S_H(\mathcal{M})$ with $W(\rho, \sigma) < \infty$,*

$$\frac{1}{2}\frac{d^+}{dt}W(\rho_t, \sigma)^2 + \frac{\lambda}{2}W(\rho_t, \sigma)^2 + H(\rho_t) \leq H(\sigma) .$$

Then T_t satisfies the following λ-HWI inequality: for any $\rho \in S_B(\mathcal{A}_0), \sigma \in S_H(\mathcal{M})$ with $W(\rho, \sigma) < \infty$,

$$H(\rho) - H(\sigma) \leq W(\rho, \sigma)\sqrt{I(\rho)} - \frac{\lambda}{2}W(\rho, \sigma)^2 .$$

Proof. By Lemma 2.7, we know that for $\rho \in S_B(\mathcal{A}_0)$, $t \mapsto I(\rho_t)$ is continuous and $t \mapsto \rho_t$ is an admissible curve with $\| D\rho_t \|_{\rho_t}^2 = I(\rho_t)$. By triangle inequality,

$$\frac{d}{dt}^+ W(\rho_{t+s}, \sigma) \leq \limsup_{t \to 0} \frac{1}{t}W(\rho_{t+s}, \sigma) - W(\rho_s, \sigma)$$

$$\leq \limsup_{t \to 0} \frac{1}{t}W(\rho_{t+s}, \rho_s) \leq \limsup_{t \to 0} \frac{1}{t}\int_0^t \| D\rho_{t+s} \|_{\rho_{t+s}} \, ds = \sqrt{I(\rho_s)} .$$

Therefore,

$$-\frac{1}{2}\frac{d^+}{dt}W(\rho_t, \sigma)^2 = \liminf_{t \to 0} \frac{1}{2t}(W(\rho, \sigma)^2 - W(\rho_t, \sigma)^2)$$

$$\leq \limsup_{t \to 0} \frac{1}{2t}(W(\rho_t, \rho)^2 + 2W(\rho_t, \rho)W(\rho_t, \sigma))$$

$$\leq \limsup_{t \to 0} \frac{1}{2t}W(\rho_t, \rho)^2 + \frac{1}{t}W(\rho_t, \rho)W(\rho_t, \sigma)$$

$$\leq W(\rho, \sigma)\sqrt{I(\rho_t)},$$

where in the last inequality we used Lemma 2.7 iii),

$$\lim_{t \to 0} W(\rho_t, \rho) = 0 \ , \ \lim_{t \to 0} W(\rho_t, \sigma) \leq \lim_{t \to 0} W(\rho_t, \rho) + W(\rho, \sigma) = W(\rho, \sigma) . \quad \square$$

Proposition 3.17. *For $\lambda > 0$, λ-HWI implies λ-MLSI.*

Proof. Since $W(\rho, T_s(\rho)) \leq \int_0^s I(\rho_t)^{1/2} dt < \infty$, we can choose $\sigma = T_s(\rho)$ in HWI inequality for any $s > 0$. By Lemma 2.7 (iv),

$$\lim_{s \to \infty} H(T_s(\rho)) - H(E(\rho)) = \lim_{s \to \infty} D(T_s(\rho)||E(\rho)) = 0 \ .$$

Then for any $\rho \in S_B(\mathcal{A}_0)$, we apply HWI inequality for $\sigma = T_s(\rho)$

$$\begin{aligned} D(\rho||\mathcal{N}) =& H(\rho) - H(E(\rho)) = H(\rho) - \lim_{s \to \infty} H(T_s(\rho)) \\ \leq& \lim_{s \to \infty} W(\rho, T_s(\rho)) \sqrt{I(\rho)} - \frac{\lambda}{2} W(\rho, T_s(\rho))^2 \\ \leq& \frac{1}{2\lambda} I(\rho) \end{aligned}$$

Here, in the last step we used the elementary inequality

$$xy \leq cx^2 + \frac{y^2}{c} \ , \ x, y, c > 0 \ ,$$

for $x = W(\rho, T_s(\rho)), y = I(\rho), c = \lambda/2$. $\square$

Remark 3.18. Here we can not choose $\sigma = E(\rho)$ because in general we do not know $W(\rho, E(\rho)) < \infty$ for $\rho \in S_B(\mathcal{A}_0)$. In particular, the finite distance for $\rho \in S_H(\mathcal{M})$ and $E(\rho)$ is a consequence of MLSI via the transport cost inequality (see [22, Section 6]) as follows,

$$W(\rho, E(\rho)) \leq \sqrt{\frac{2D(\rho||E(\rho))}{\lambda}} \ . \tag{14}$$

We call the above inequality (14) λ-transport cost inequality or in short λ-TC.

Now we have two ways to reach Bakry-Emery Theorem.

Corollary 3.19 *(Non-ergodic Bakry-Emery theorem). For $\lambda > 0$, λ-GE implies λ-MLSI.*

Proof. We can either use λ-GE $\Rightarrow$ λ-FM $\Rightarrow$ λ-MSLI or λ-GE $\Rightarrow$ λ-HWI $\Rightarrow$ λ-MSLI. $\square$

Beyond positive curvature lower bound, we also have two ways for MLSI. The first one is to apply our Theorem 3.6 with the above discussion. Recall that the function $\kappa(\lambda, t) = \lambda(2 - 2e^{-2\lambda t})^{-1}$.

Corollary 3.20. *Let $T_t : \mathcal{M} \to \mathcal{M}$ be a symmetric quantum Markov semigroup. Suppose*

 i) T_t *satisfies λ-GE for some $\lambda \in \mathbb{R}$;*
 ii) T_t *has finite CB-return time $t_{cb} < \infty$.*

Then T_t-satisfies $\kappa(\lambda, t_{cb})$-MLSI.

Remark 3.21. Note that for $\lambda t > \ln \sqrt{2}$, $\kappa(\lambda, t) > \lambda$. This means when the CB-return time $t_{cb} < \lambda^{-1} \ln \sqrt{2}$, Corollary 3.20 gives stronger MLSI-constant than Bakry-Emery Theorem. Also for $\lambda > 0$, $\kappa(\lambda, t) \to \lambda/2$ when $t_{cb} \to \infty$.

One can compare the above corollary to the approach in [49, Corollary 3.1] using the transport inequality.

Corollary 3.22. *Let $T_t : \mathcal{M} \to \mathcal{M}$ be a symmetric quantum Markov semigroup. Suppose*

i) *T_t satisfies λ-GE for some $\lambda \in \mathbb{R}$;*
ii) *T_t satisfies γ-transport cost inequality in (14) for $\gamma \geq \max\{-\lambda, 0\}$*

Then T_t-satisfies α-MLSI for $\alpha = \max\{\lambda, \frac{\gamma}{4}(1 + \frac{\lambda}{\gamma})^2\}$.

The proof is similar to [49, Corollary 3.1]. One could also replace "TC" in condition ii) by the so called "MLSI+TC" inequality

$$W(\rho, E(\rho)) \leq \sqrt{\frac{I(\rho)}{\gamma}} \, , \tag{15}$$

to obtain a similar estimate as in [49, Corollary 3.2].

3.4. Bochner's inequality

We shall now discuss the curvature lower bound condition introduced in [39]. Let $T_t : \mathcal{M} \to \mathcal{M}$ be a symmetric quantum Markov semigroup and $(\mathcal{A}, \hat{\mathcal{M}}, \delta)$ be a derivation triple for T_t. Denote Ω_δ as the closure of $\mathcal{A}\delta(\mathcal{A})$ in $L_2(\hat{\mathcal{M}})$. It follows from Leibniz rule that Ω_δ is a $\mathcal{A}$-bimodule. To distinguish with the entropy Ricci curvature lower bound, we refer the following notion from [39] as geometric Ricci curvature lower bound.

Definition 3.23. We say $(\mathcal{A}, \hat{\mathcal{M}}, \delta)$ satisfies a geometric Ricci curvature lower bound λ for $\lambda \in \mathbb{R}$ (in short GRic $\geq \lambda$) if there exists a symmetric quantum Markov semigroup $\hat{T}_t = e^{-\hat{A}t} : \hat{\mathcal{M}} \to \hat{\mathcal{M}}$ with generator $\hat{A}$ such that

i) $\hat{T}_t|_{\mathcal{M}} = T_t$ for any $t \geq 0$.
ii) $\delta(\mathcal{A}_0) \subset \mathrm{dom}(\hat{A})$ and there exists a $\mathcal{A}$-bimodule operator $\mathrm{Ric} : \Omega_\delta \to L_2(\hat{\mathcal{M}})$ such that for $x \in \mathcal{A}_0$,

$$\mathrm{Ric}(\delta(x)) = \delta A(x) - \hat{A}\delta(x). \tag{16}$$

iii) for any $y \in \Omega_\delta$,

$$\langle y, \mathrm{Ric}(y) \rangle \geq \lambda \langle y, y \rangle \ , \tag{17}$$

where $\langle \cdot, \cdot \rangle$ is the trace inner product of $(\hat{\mathcal{M}}, \tau)$.

We call the bimodule map Ric "Ricci operator" as an analog of Ricci tensor in geometry. The above definition is of course an imitation of Bochner–Weitzenböck–Lichnerowicz formula (cf. p. 374 [58])

$$\Delta - \nabla\nabla^* + \mathrm{Ric} = 0 \ , \tag{18}$$

where $\Delta = \nabla^*\nabla$ is the Laplace-Beltrami operator on a Riemannian manifold and ∇ is the gradient operator. When acting on a gradient ∇f, (18) becomes

$$\Delta(\nabla f) - \nabla(\Delta f) + \mathrm{Ric}(\nabla f) = 0 \ ,$$

which is the motivation for (16). Note that the above Definition (3.23) adds a little flexibility that $\hat{A}$ can be any generator extending A on $\mathcal{M}$. We discuss more on the connection to classic Ricci curvature in Section 4.2.

On the other hand, we emphasize that Definition 3.23 is different from the entropy Ricci lower bound in Definition 3.11. One major difference is that Definition 3.23 is automatically "complete" in the sense that if T_t has GRic $\geq \lambda$ (in our sense), then $T_t \otimes \mathrm{id}_{\mathcal{R}}$ has GRic $\geq \lambda$ for any finite von Neumann algebra $\mathcal{R}$. Indeed, both the algebraic equation (16) and the L_2 inequality (17) naturally extends to $T_t \otimes \mathrm{id}_{\mathcal{R}}$. The complete bounded version of gradient estimate is studied in [61].

We recall the following results from [39].

Theorem 3.24 (Theorem 3.6 of [39]). For $\lambda \in \mathbb{R}$, T_t has GRic $\geq \lambda$ implies that $T_t \otimes \mathrm{id}_{\mathcal{R}}$ has λ-GE for any finite von Neumann algebra $\mathcal{R}$.

The next theorem is inspired by the discussion in [10, Section 8.3] (see also [11, Theorem 10.8] and [15, Proposition 5]).

Theorem 3.25. Let $T_t : \mathcal{M} \to \mathcal{M}$ be a symmetric quantum Markov semigroup and let $(\mathcal{A}, \hat{\mathcal{M}}, \delta)$ be a derivation triple of T_t. Suppose that there exists a symmetric quantum Markov semigroup $\hat{T}_t : \hat{\mathcal{M}} \to \hat{\mathcal{M}}$ such that for any $t \geq 0$,

$$\tilde{T}_t|_{\mathcal{M}} = T_t \ , \ \text{and} \ \delta \circ T_t = e^{-\lambda t} \hat{T}_t \circ \delta \tag{19}$$

for some $\lambda \in \mathbb{R}$. Then T_t satisfies GRic $\geq \lambda$. Moreover, the Ricci operator GRic can be taken to a constant multiple of the identity operator.

Proof. Let $\hat{A}$ be the generator of $\hat{T}_t$. For $x \in \mathcal{A}_0$,

$$\lim_{t \to 0} \frac{1}{t}(e^{-\lambda t}\hat{T}_t(\delta(x)) - \delta(x)) = \lim_{t \to 0} \frac{e^{-\lambda t}}{t}(\hat{T}_t(\delta(x)) - \delta(x)) + \frac{1}{t}(e^{-\lambda t}\delta(x) - \delta(x))$$

$$= -\hat{A}\delta(x) - \lambda\delta(x) ,$$

which converges in w^*-topology because $\delta(\mathcal{A}_0) \subset \text{dom}(\hat{A})$. On the other hand, for $y \in \delta(\mathcal{A}_0)$ and $\delta^*(y) \in L_2(\mathcal{M})$,

$$\lim_{t \to 0} \frac{1}{t}\big(\tau(y\delta(x)) - \tau(y\delta(T_t(x)))\big) = \lim_{t \to 0} \frac{1}{t}\big(\tau(\delta^*(y)x) - \tau(\delta^*(y)T_t(x))\big)$$

$$= \tau(\delta^*(y)A(x)) ,$$

which implies $\lim_{t \to 0} \frac{1}{t}(\delta(T_t(x)) - \delta(x)) = -\delta(A(x))$ weakly. Thus we have for $x \in \mathcal{A}_0$,

$$-\delta(A(x)) = -\hat{A}\delta(x) - \lambda\delta(x) ,$$

which means the Ricci operator is constant $\text{Ric}(\delta(x)) = \lambda\delta(x)$. $\square$

As we see in the above proof, the relation (19) is equivalent to the Ricci operator in (16) equaling to a multiple of the identity. We emphasize this special case by giving the following definition.

Definition 3.26. We say a semigroup T_t satisfies constant λ-Ricci curvature condition (λ-GRic) if T_t admits a derivation triple satisfying (19).

We remark that the λ-GRic relation does not give the meaning that Ricci curvature is constant λ but just a lower bound by λ. We revisit the Orstein-Unlenbeck semigroup discussed in [10].

Example 3.27. Let $\mathbb{R}^n$ be the n-dimensional real Euclidean space and μ the standard Gaussian distribution. The Orstein-Unlenbeck (OU) semigroup $T_t = e^{-At}$: $L_\infty(\mathbb{R}^n, \mu) \to L_\infty(\mathbb{R}^n, d\mu)$ is given by

$$T_t f(x) = \int_{\mathbb{R}^n} f(e^{-t}x + \sqrt{1 - e^{-2t}}y)d\mu(y) .$$

Denote $\partial_j := \frac{\partial}{\partial x_j}$ as the partial derivative. The generator of the OU semigroup is given by

$$A = \Delta + x \cdot \nabla = \nabla^*\nabla + x \cdot \nabla = \sum_{j=1}^{n}(-\partial_j^2 + x_j\partial_j).$$

Consider the derivation

$$\delta : C^\infty(\mathbb{R}^n) \to \oplus_{j=1}^n C^\infty(\mathbb{R}^n) \ , \ \delta(f) = (\partial_j f)_{j=1}^n \ .$$

As observed in [10, Section 8.1], we have the relation $[\partial_j, -\Delta + x \cdot \nabla] = \partial_j$ for $j = 1, \ldots, n$. This translates to the equality

$$(\hat{A} \otimes \mathrm{id}) \circ \delta - A \circ \delta = \delta \ ,$$

where $\hat{A} = A \otimes \mathrm{id}$ is the extension of A to $\oplus_{j=1}^n C^\infty(\mathbb{R}^n) \cong C^\infty(\mathbb{R}^n) \otimes l_\infty^n$, which is clearly the generator of the semigroup $\hat{T}_t = T_t \otimes \mathrm{id}$ on $L_\infty(\mathbb{R}^n) \otimes l_\infty^n$. In particular, this gives a derivation triple for the OU semigroup that satisfies 1-GRic. Moreover since T_t has spectral gap 1, we can therefore conclude the sharp complete version result that $T_t \otimes \mathrm{id}_{\mathcal{R}}$ satisfies 1-GE for any finite von Neumann algebra $\mathcal{R}$, and T_t satisfies 1-CFM and 1-CLSI.

We have a complete version of Corollary 3.20

Corollary 3.28. *Let $T_t : \mathcal{M} \to \mathcal{M}$ be a symmetric quantum Markov semigroup. Suppose*

  i) *T_t satisfies $\mathrm{GRic} \geq \lambda$ for some $\lambda \in \mathbb{R}$;*
 ii) *T_t has finite CB-return time $t_{cb} < \infty$.*

Then T_t-satisfies $\kappa(\lambda, t_{cb})$-CLSI.

4. Examples

In this section, we discuss applications to classical Markov semigroups and finite dimensional quantum Markov semigroups.

4.1. Diffusion semigroups

Our motivation for Fisher monotonicity was from Bakry-Emery's curvature dimension condition for diffusion Markov semigroup. We refer to [3] for more information on classical diffusion Markov semigroup.

Let (Ω, μ) be a Borel space equipped with a Borel probability measure μ. Let $T_t : L_\infty(\Omega, \mu) \to L_\infty(\Omega, \mu)$ be an ergodic Markov semigroup and A be its generator. We say T_t satisfy *diffusion property* if its gradient form Γ satisfies the following product rule,

$$\Gamma(fh, g) = f\Gamma(h, g) + h\Gamma(f, g) \ . \tag{20}$$

Denote $\Gamma(f) := \Gamma(f, f)$. It then follows from polynomial approximation that for a smooth function $\psi : \mathbb{R} \to \mathbb{R}$,

$$\Gamma(\psi(f), g) = \psi'(f)\Gamma(f, g) \,, \Gamma(\psi(f)) = \psi'(f)^2\Gamma(f, g)$$

For a density function $f \in L_\infty(\Omega, \mu)$, the entropy $H(f)$ (also called Boltzman H-functional) and the Fisher information $I(f)$ are given by

$$H(\rho) = D(\rho||1) = \int_\Omega \rho \log \rho \, d\mu$$

$$I(f) = -\int (Af) \log f \, d\mu = \int \Gamma(f, \log f) d\mu = \int f\Gamma(\log f) d\mu$$

Recall that the Γ_2 operator is defined as

$$\Gamma_2(f, g) = \frac{1}{2}\Big(\Gamma(Af, g) + \Gamma(f, Ag) - A\Gamma(f, g)\Big) \,.$$

Denote $\Gamma_2(f) := \Gamma_2(f, f)$. Γ_2 can be realized as

$$\Gamma_2(f) = \lim_{t \to 0} \frac{T_t(\Gamma(f)) - \Gamma(T_t(f))}{t} \,.$$

The derivative of Fisher information is

$$\frac{dI(T_t f)}{dt} = -2 \int T_t f \Gamma_2(\log T_t f) d\mu \,. \tag{21}$$

Recall that T_t satisfies (λ, ∞)-curvature dimension condition for $\lambda \in \mathbb{R}$ (in short, CD(λ, ∞)) if for any $f \in \text{dom}(A)$

$$\Gamma_2(f) \geq \lambda \Gamma(f) \,.$$

It follows immediately CD(λ, ∞) implies λ-FM. For $\lambda > 0$, it is the Barky-Emery theorem that CD(λ, ∞) $\Rightarrow$ λ-FM $\Rightarrow$ λ-MLSI. For general $\lambda \in \mathbb{R}$, we have the following theorem for diffusion Markov semigroups.

Theorem 4.1. *Let $T_t : L_\infty(\Omega, \mu) \to L_\infty(\Omega, \mu)$ be an ergodic symmetric diffusion Markov semigroup. Suppose T_t satisfies curvature-dimension condition $CD(\lambda, \infty)$. If in addition, we assume*

i) $\|T_t : L_1(\Omega) \to L_\infty(\Omega)\| \leq ct^{-d/2}$ *for some $c, d > 0$ and all $0 < t < 1$;*
ii) *the generator A satisfies spectral gap $\sigma > 0$.*

Then T_t-satisfies $m(\lambda)$-MLSI for

$$m(\lambda) = \begin{cases} \Big(2 + 2(d-1)\log 2 + \frac{4}{\sigma}\log c\Big)^{-1}, & \text{if } \lambda = 0 \\ \lambda\Big(2 - 2^{1-(d-1)\lambda}c^{-\frac{2\lambda}{\sigma}}\Big)^{-1}, & \text{if } \lambda \neq 0. \end{cases}$$

Proof. The condition i) is the Varopoulos' dimension condition. Here the CB-norm estimate is automatic:

$$\|T_t - E_\tau : L_1(\Omega) \to L_\infty(\Omega)\| = \|T_t - E_\tau : L_1(\Omega) \to L_\infty(\Omega)\|_{cb}.$$

This is because $L_\infty(\Omega)$ is a commutative space (see [51, Proposition 1.10]). The assertions follow from Theorem 3.6 and the return time estimates in Lemma 3.9. □

Remark 4.2. It is well known that if $T_t f(x) = \int_\Omega k_t(x,y) f(y) d\mu$ is given by the kernel function $k_t(x, y)$. Then

$$\|T_t - E_\tau : L_1(\Omega) \to L_\infty(\Omega)\| = \| k_t - 1 \|_{L_\infty(\Omega \times \Omega)}$$

is a kernel estimate.

4.2. Heat semigroups

We shall now discuss the heat semigroups. We refer to [58] for more information on analysis of heat semigroups on manifolds. Let (M, g) be a complete compact Riemannian manifold equipped with Riemannian metric g. Let Δ be the Laplace-Beltrami operator given by

$$\Delta f = \nabla^* \nabla f \,,$$

where ∇ is the gradient operator and $\nabla^* = \text{div}$ is the divergence. The heat semigroup $T_t = e^{-\Delta t} : L_\infty(M, d\text{vol}) \to L_\infty(M, d\text{vol})$ is a Markov semigroup with respect to the volume form "$d\text{vol}$" induced by g. Recall the Bochner–Weitzenböck–Lichnerowicz formula that for the vector field $\nabla \phi$,

$$-\frac{1}{2}\Delta|\nabla\phi|^2 + \nabla\phi \cdot \nabla(\Delta\phi) + \| \nabla\phi \|_2^2 + \text{Ric}(\nabla\phi, \nabla\phi) = 0,$$

which translates to

$$-\Delta + \nabla\nabla^* + \text{Ric} = 0 \,. \tag{22}$$

The $C^\infty(M)$-bimodule property of Ric is exactly the fact that the Ricci curvature is a smooth tensor over M.

The same argument applies to weighted Riemannian manifolds $(M, g, e^{-W} d\text{vol})$ where e^{-W} is a smooth density function with respect to $d\text{vol}$. The weighted Laplacian is

$$\Delta_W = \nabla^*\nabla = \Delta - \nabla W \cdot \nabla \,,$$

where ∇^* is adjoint of ∇ with respect to $L_2(M, e^{-W} d\mu)$ and Δ_W is a self-adjoint operator on $L_2(M, e^{-W} d\mu)$. Then the weighted heat semigroup $T_t = e^{-\Delta_W t}$ is an ergodic symmetric Markov semigroup with the unique invariant measure $e^{-W} d\mu$. In this case,

$$\Delta_W - \nabla \nabla^* = \mathrm{Ric}_W \ ,$$

where $\mathrm{Ric}_W = \mathrm{Ric}_g + \nabla \nabla W$ is the sum of Ricci curvature tensor of the metric g and the Hessen of the function W. The weighted Ricci curvature bound $\mathrm{Ric}_W \geq \lambda$ is that $\mathrm{Ric}_W(\xi, \xi) \geq \lambda g(\xi, \xi)$ for any vector field $\xi \in TM$. When $\lambda > 0$, $\mathrm{Ric}_W \geq \lambda$ implies $T_t = e^{-\Delta_W t}$ satisfies λ-MLSI by the Bakry-Emery Theorem.

It is proved in [39, Section 4] that $\mathrm{Ric}_W \geq \lambda$ actually implies $\mathrm{GRic} \geq \lambda$, which implies a complete version of Bakry-Emery theorem.

Theorem 4.3 *([39]). If $\mathrm{Ric}_W(\xi, \xi) \geq \lambda g(\xi, \xi)$ for any $\xi \in TM$, then the weighted heat semigroup $T_t = e^{-\Delta_W t}$ satisfies $\mathrm{GRic} \geq \lambda$. In particular, if $\mathrm{Ric}_W \geq \lambda > 0$, $T_t = e^{-\Delta t}$ satisfies λ-CLSI.*

The proof uses the Clifford bundle $Cl(M)$ as the quantization of tangent bundle TM. Then the $\mathrm{GRic} \geq \lambda$ is a realization of the Bochner identity on $Cl(M)$. We refer to [39] for details.

Now we apply our method for general compact weighted manifolds. It follows from compactness and continuity that $\mathrm{Ric}_W \geq \lambda$ always holds for some real $\lambda \in \mathbb{R}$. Indeed, for each $x \in M$, Ric_W at x is a real symmetric matrix with respect to an orthonormal basis of g. Hence

$$(\mathrm{Ric}_W)_x \geq \lambda_{min}(x) g \geq \min_{x \in M} \lambda_{min}(x) g$$

Here $\lambda_{min}(x)$ is the smallest eigenvalue of $(\mathrm{Ric}_W)_x$ with respect to metric g, which is continuous depending on $x \in M$. Define that $\mathrm{Ric}(\Delta_W) = \min_{x \in M} \lambda_{min}(x)$ as the global minimum of $\lambda_{min}(x)$. Thus the heat semigroup $T_t = e^{-\Delta_W t}$ always satisfies $\mathrm{GRic} \geq \lambda$ for some real $\lambda = \mathrm{Ric}(\Delta_W)$. The following is an application of Theorem 3.28.

Theorem 4.4. *Let $(M, g, e^{-W} dvol)$ be a compact connected weighted Riemannian manifold. Then the weighted heat semigroup $T_t = e^{-\Delta_W t}$ satisfies λ-CLSI for some $\lambda > 0$.*

Proof. We know from Theorem 4.3 that $T_t = e^{-\Delta_W t}$ always satisfies $\mathrm{GRic} \geq \mathrm{Ric}(\Delta_W) \in \mathbb{R}$. On the other hand, both spectral gap and finite Varopoulos dimension of Δ_W are well-known for compact weighted manifolds. See [26, Theorem 10.23] for spectral gap and [26, Theorem 14.19 & Exercise 15.2] for Varopoulos dimension. Indeed, the $T_t = e^{-\Delta_W t}$ satisfies the ultra-contractive estimates of dimension $n = \dim(M)$,

$$\| T_t : L_1(M, dvol) \to L_\infty(M, dvol) \| \leq c t^{-n/2} \ , \ 0 < t \leq 1.$$

Then it follows from Lemma 3.9 and Corollary 3.28 that $T_t = e^{-\Delta_W t}$ satisfies λ-CLSI where λ is determined by $\mathrm{Ric}(\Delta_W)$, spectral gap of Δ_W and the ultra-contractive estimate of $e^{-\Delta_W t}$. $\square$

The above theorem has the following refined form.

Theorem 4.5. *Let (M, g) be a connected compact Riemannian manifold and let Δ be the Laplace-Beltrami operator. Suppose the Ricci curvature of M is bounded below by K for some $K \in \mathbb{R}$.*

(i) *the heat semigroup $T_t = e^{-\Delta t}$ satisfies λ-CLSI for*

$$\lambda = \begin{cases} K, & \text{if } K > 0 \\ \left(4 + \frac{4}{\sigma} \log(2C_1)\right)^{-1}, & \text{if } K = 0 \\ K\left(2 - e^{-2K}\left(\frac{2c(K,n)}{V}\right)^{-\frac{2K}{\sigma}}\right)^{-1}, & \text{if } K < 0, \end{cases} \tag{23}$$

where σ is the spectral gap of Δ, V is the minimum volume of radius 1 ball in M, C_1 is a universal constant and $C_2(K, n)$ only depends on K and the dimension $n = \dim(M)$.

(ii) *Let W be a smooth function on M such that e^{-W} is a probability density function for the volume form $d\mathrm{vol}$. Then the weighted heat semigroup $T_t = e^{-\Delta_W t}$ satisfies $c\lambda$-CLSI where λ is given in (23) and $c = e^{\min W - \max W}$.*

Proof. The case $K > 0$ is in Theorem 4.3. We argue for the case $K \leq 0$. Denote $k : M \times M \times \mathbb{R}_+ \to \mathbb{R}$ as the heat kernel. Recall the famous Li-Yau estimate that for a complete Riemannian manifold with Ricci curvature bounded below by $\mathrm{Ric}(M) \geq -K$ for some $K \geq 0$, the heat kernel satisfies

$$k(x, y, t) \leq \frac{C_1}{\sqrt{V(x, \sqrt{t})V(y, \sqrt{t})}} \exp\left(C_2 Kt - \frac{d(x, y)^2}{5t}\right),$$

where $d(x, y)$ is the Riemannian distance, $V(x, \sqrt{t})$ is the volume of geodesic ball center at x with radius $\sqrt{t}$, C_1 is some universal constant and C_2 only depends on the dimension $\dim(M) = n$. (We choose the parameter $\epsilon = 1$ in statement of [40, Corollary 3.1].) On diagonal $x = y$, we have

$$k(x, x, t) \leq \frac{C_1}{V(x, \sqrt{t})} \exp\left(C_2 Kt\right).$$

Take $V = \min_{x \in M} V(x, 1)$ as the minimum volume of radius 1 ball in M. Then for $t = 1$,

$$k(x, x, 1) \leq C_1 V(x, 1)^{-1} \exp\left(C_2 Kt\right) \leq c(K, n)V^{-1},$$

where $c(K, n) = C_1 \exp(C_2 K)$ is a constant only depending on $\dim(M) = n$ and curvature bound K (for $K = 0$, $C(0, n)$ is also independent of n). The ultra-contractive estimate is given by heat kernel on the diagonal,

$$\| T_1 : L_1(M, d\text{vol}) \to L_\infty(M, d\text{vol}) \| = \sup_x k(x, x, 1) \leq c(K, n) V^{-1} .$$

Let σ be the spectral gap of Δ. By Lemma 3.8, we have

$$t_{cb} \leq 1 + \frac{1}{\sigma} \log(2c(K, n) V^{-1})$$

The assertion follows from Corollary 3.28. This proves i). ii) follows from the change measure [39, Lemma 2.11]. Indeed, for smooth (operator-valued) function f

$$I_{\Delta_W}(f) = \int \langle \nabla f, \nabla \log f \rangle e^{-W} d\text{vol} \geq e^{-\max W} \int \langle \nabla f, \nabla \log f \rangle d\text{vol} = I_\Delta(f) ,$$

where I_Δ is the Fisher information for the standard Laplacian and I_{Δ_W} for the weighted Laplacian Δ_W. The comparison for relative entropy follows from [39, Lemma 2.8]. $\quad\square$

4.3. Central semigroups on compact groups

In this subsection, we consider Markov semigroups on compact groups. Let G be a compact group. We denote by $C(G)$ (resp. $C^\infty(G)$) the space of continuous (resp. smooth) functions on G and denote by $L_\infty(G) = L_\infty(G, m)$ the L_∞-space with respect to the Haar probability measure m. Let $L_g : L_\infty(G) \to L_\infty(G)$ (resp. R_g) be the left (resp. right) translation operator.

$$(L_g f)(h) = f(gh) , \; (R_g f)(h) = f(hg) .$$

We say a Markov semigroup $T_t : L_\infty(G) \to L_\infty(G)$ is *left (resp. right) invariant* if $L_g \circ T_t = T_t \circ L_g$ (resp. $R_g \circ T_t = T_t \circ R_g$) for all $g \in G$. We say T_t is *central* if it is both left and right invariant. Recall that a function $k \in L_1(G)$ is *central* if $k(sgs^{-1}) = k(g)$ for a.e. $g, s \in G$. This is equivalent to the condition $f \star k = k \star f$ for all $f \in L_1(G)$, where $\star$ denotes the convolution product on $L_1(G)$. We denote the subalgebra of central functions in $L_1(G)$ by $ZL_1(G)$. It is well known that a Markov semigroup T_t on $L_\infty(G)$ is central if and only if there exists a *convolution semigroup* of central probability densities $(k_t)_{t \geq 0} \subset ZL_1(G)$

$$T_t f(g) = (f \star k_t)(g) = \int_G f(y) k_t(y^{-1}g) dm(y), \; f \in L_\infty(G).$$

Now consider the co-multiplication map $\alpha : L_\infty(G, m) \to L_\infty(G \times G, m \times m)$,

$$\alpha(f)(g, h) = f(gh) , \; \alpha(f)(g, \cdot) = L_g f , \; \alpha(f)(\cdot, h) = R_h f$$

It is clear that α is a m to $m \times m$ measure preserving $*$-monomorphism. Moreover, if T_t is a left invariant semigroup we have the commutation relation $\alpha \circ T_t = (\mathrm{id} \otimes T_t) \circ \alpha$. Indeed,

$$\alpha(T_t f)(g, \cdot) = L_g(T_t f) = T_t(L_g f) = \mathrm{id} \otimes T_t(\alpha(f))(g, \cdot)$$

Similarly, if T_t is right invariant, we have $\alpha \circ T_t = (T_t \otimes \mathrm{id}) \circ \alpha$. Thus for a central semigroup T_t, we have the following commutative diagram

$$
\begin{array}{ccc}
L_\infty(G \times G) & \xrightarrow{\mathrm{id}_G \otimes T_t \text{ or } T_t \otimes \mathrm{id}_G} & L_\infty(G \times G) \\
\uparrow \alpha & & \uparrow \alpha \\
L_\infty(G) & \xrightarrow{T_t} & L_\infty(G)
\end{array}
\qquad (24)
$$

This is a crucial point in the following lemma.

Lemma 4.6. *Let G be a compact group and $T_t : L_\infty(G) \to L_\infty(G)$ be a central Markov semigroup. Then T_t satisfies $\mathrm{GRic} \geq 0$ and hence complete Fisher monotonicity.*

Proof. Let A be the generator of T_t and $(\mathcal{A}_\mathcal{E}, \mathcal{M}, \delta)$ be a derivation triple for T_t. That is, $\delta : \mathcal{A}_\mathcal{E} \to L_2(\mathcal{M})$ is a $*$-preserving derivation such that

$$E(\delta(x)^* \delta(y)) = \Gamma_A(x, y) ,$$

where E is the conditional expectation on to $L_\infty(G) \subseteq \mathcal{M}$, and $\mathcal{A}_\mathcal{E} = L_\infty(G) \cap \mathrm{dom}(A^{1/2})$ is the Dirichlet subalgebra. We show that

$$\partial = (\delta \otimes \mathrm{id}) \circ \alpha : L_\infty(G) \to L_\infty(G, \mathcal{M}) \cong \mathcal{M} \bar{\otimes} L_\infty(G)$$

is also a derivation for T_t. Let $E_\alpha : L_\infty(G \times G) \to L_\infty(G)$ be the conditional expectation obtained as the adjoint of α. Using the commutative diagram (24), we have $E_\alpha(A \otimes \mathrm{id})\alpha = A$, which follows by differentiating $\alpha \circ T_t = (T_t \otimes \mathrm{id}) \circ \alpha$. Then for the gradient forms associated to A and $A \otimes \mathrm{id}$ (the latter which acts on $\alpha(\mathcal{A}_\mathcal{E}) \subset \alpha(L_\infty(G)))$, we have

$$
\begin{aligned}
\Gamma_A(x, y) &= x^* A y + (Ax)^* y - A(x^* y) \\
&= x^* E_\alpha(A \otimes \mathrm{id})\alpha(y) + (E_\alpha(A \otimes \mathrm{id})\alpha(x))^* y - E_\alpha(A \otimes \mathrm{id})\alpha(x^* y) \\
&= E_\alpha(\alpha(x)^*(A \otimes \mathrm{id})\alpha(y) + (A \otimes \mathrm{id})\alpha(x)^* \alpha(y) - (A \otimes \mathrm{id})\alpha(x^* y)) \\
&= E_\alpha(\Gamma_{A \otimes \mathrm{id}}(\alpha(x), \alpha(y))) \\
&= E_\alpha \circ (E \otimes \mathrm{id})((\delta \otimes id)\alpha(x)^*(\delta \otimes id)\alpha(y)) \\
&= E_\alpha \circ (E \otimes \mathrm{id})(\partial(x)^* \partial(y))
\end{aligned}
$$

where we have used the fact $(\delta \otimes \mathrm{id})$ is a derivation for $T_t \otimes \mathrm{id}$. Here $E_\alpha \circ (E \otimes \mathrm{id})$ is exactly the conditional expectation onto $\alpha(L_\infty(G)) \subset L_\infty(G, \mathcal{M})$. Thus we have shown

that $(\mathcal{A}_{\mathcal{E}}, L_{\infty}(G, \mathcal{M}), \partial)$ is a new derivation triple for T_t. Now for this derivation, we have

$$
\begin{aligned}
\partial \circ T_t &= (\delta \otimes \mathrm{id}_G) \circ \alpha \circ T_t = (\delta \otimes \mathrm{id}_G)(\mathrm{id}_G \otimes T_t) \circ \alpha \\
&= (\mathrm{id}_{\mathcal{M}} \otimes T_t)(\delta \otimes \mathrm{id}) \circ \alpha = (\mathrm{id}_{\mathcal{M}} \otimes T_t)\partial \,,
\end{aligned}
$$

where $\mathrm{id}_G \otimes T_t$ (resp. $\mathrm{id}_{\mathcal{M}} \otimes T_t$) is the extension semigroup of T_t on $L_{\infty}(G \times G)$ (resp. $\mathcal{M} \overline{\otimes} L_{\infty}(G)$). Note that here we used the other part of (24) $\alpha \circ T_t = (\mathrm{id}_G \otimes T_t) \circ \alpha$ by the right invariance of T_t. This verifies the algebraic relation in Theorem 3.25 for $\lambda = 0$, which implies the assertions. $\square$

Example 4.7 *(Heat semigroups).* Let G be a compact Lie group and $\mathfrak{g}$ be its Lie algebra of left invariant vector fields. Let $X = \{X_1, ..., X_r\}$ be an orthonormal basis of $\mathfrak{g}$ with respect to its Killing form. We consider the heat semigroup $T_t = e^{-\Delta t}$ generated by the Casimir operator $\Delta = \sum_j X_r^2$. The natural derivation for Δ is the gradient

$$
\nabla : C^{\infty}(G) \to \oplus_{j=1}^r C^{\infty}(G) \,, \nabla(f) = (X_j f)_{j=1}^r
$$

It is known from representation theory that $\Delta = \sum_j X_j^2$ as a generator is central. Indeed, recall that for an irreducible continuous representation $\pi : G \to B(H_{\pi})$ on the Hilbert space H_{π}, the coefficient function space associated to π is the finite-dimensional subspace

$$
\mathcal{E}_{\pi}(G) = \{\, f \in C(G) : f(g) = \langle h_1, \pi(g)h_2 \rangle_{H_{\pi}} \mid h_1, h_2 \in H_{\pi} \,\} \subset L_2(G) \,.
$$

Denote E_{π} as the Hilbert projection from $L_2(G)$ to the closure of $\mathcal{E}_{\pi}(G)$. The Casimir operator Δ then admits a spectral decomposition of the form

$$
\Delta = \sum_{\pi \in Irr(G)} \lambda_{\pi} E_{\pi}
$$

where the summation is over all irreducible representation π and λ_{π} is the common eigenvalue for all coefficient functions of π. Since the E_{π} is invariant for both left translation and right translation, this implies Δ and the semigroup $e^{-\Delta t}$ are central. By the construction in Theorem 4.6, the algebraic relation curvature relation 0-GRic is satisfied with the following alternative derivation

$$
\partial : C^{\infty}(G) \to \oplus_{j=1}^r C^{\infty}(G \times G) \,, \partial f = (\nabla \otimes \mathrm{id})\alpha(f)(g, h) = (X_j f(gh))_{j=1}^r \,.
$$

Combined with the heat kernel estimate and spectral gap (see e.g. [57]), we have the following corollary.

Theorem 4.8. *Let G be a compact Lie group and let Δ be the Casimir operator. For $r \in (0, 1]$, denote $T_t^r = e^{-\Delta^r t} : L_{\infty}(G) \to L_{\infty}(G)$ as the heat semigroup ($r = 1$) and its*

subordinated semigroup $(0 < r < 1)$. Then for each $r \in (0, 1]$, T_t^r satisfies GRic ≥ 0, complete Fisher monotonicity, and $\lambda(r)$-CLSI for

$$\lambda(r) = \left(4 + 4\sigma^{-r} \log(2c(r, n) + \frac{C}{V_1})\right)^{-1} ,$$

where C is an absolute constant, $c(r, n)$ is a constant only depending on $0 < r \leq 1$ and $n = \dim(G)$, σ is the spectral gap of Δ and V_1 is the volume of unit geodesic ball.

Proof. For all $r \in (0, 1]$, $\Delta^r = \sum_\pi \lambda_\pi^r E_\pi$ is a central generator. Thus T_t^r are central semigroup hence has 0-GRic. It is well-known (see e.g. [57]) that the heat semigroup $T_t^1 = e^{-\Delta t}$ has ultra-contractive estimate

$$\| T_t : L_1(G, m) \to L_\infty(G, m) \| = Ct^{-\frac{n}{2}} , 0 < t \leq 1$$

where $n = \dim(G)$. By the discussion in [57, Section II.3], the subordinated semigroup T_t^r has spectral gap σ^r and Varopoulos dimension $\frac{1}{r} \dim(G)$. Then assertions follow from Theorem 3.28.

We now give the concrete ultra-contractive estimates of T_t^r for each r. Let V_t be the volume of geodesic ball of radius t. Since G has non-negative Ricci curvature, by Bishop-Gromov volume comparison theorem (cf. [53, Theorem 5.6.4]), for $0 < t \leq 1$, $V(t) \geq V(1)t^n$. Then for $r = 1$ and $T_t := T_t^r$, using the Li-Yau estimate [40, Corollary 3.1] again,

$$\| T_t : L_1(G) \to L_\infty(G) \| = k(x, x, t) \leq \frac{C}{V_{\sqrt{t}}} \leq \frac{C}{V_1} t^{-\frac{n}{2}} ,$$

where C is some absolute constant, $k(x, y, t)$ is the heat kernel of T_t, x is some point in G, and $V_{\sqrt{t}}$ (resp. V_1) is the volume of geodesic ball in G with radius $\sqrt{t}$ (resp. 1). Denote $C_1 = C/V_1$. For the subordinated semigroup, we then use the argument from [57, Section II.3],

$$T_t^\alpha = e^{-\Delta^\alpha t} = \int\limits_0^\infty f_\alpha(v) T_{vt^{1/\alpha}} dv ,$$

where f_α is the function whose Laplace transform is $s \mapsto e^{-s^\alpha}$. In particular, $f_\alpha \geq 0$ and $\int_0^\infty f_\alpha(v) dv = 1$. Then for $t = 1$,

$$\| T_1^r : L_1(G) \to L_\infty(G) \|$$

$$\leq \int\limits_0^\infty f_\alpha(v) \| T_v : L_1(G) \to L_\infty(G) \| \, dv$$

$$\leq \int_0^1 f_\alpha(v) \, \|T_v : L_1(G) \to L_\infty(G)\| \; dv + \int_1^\infty f_\alpha(v) \, \|T_v : L_1(G) \to L_\infty(G)\| \; dv$$

$$\leq \int_0^1 f_\alpha(v) v^{-\frac{n}{2}} \, dv + \int_1^\infty f_\alpha(v) \, \|T_1 : L_1(G) \to L_\infty(G)\| \; dv$$

$$\leq \int_0^1 f_\alpha(v) v^{-n/2} dv + C_1 \int_1^\infty f_\alpha(v) dv$$

$$\leq c(\alpha, n) + C_1$$

where $c(\alpha, n) = \int_0^1 f_\alpha(v) v^{-n/2} dv \leq \int_0^\infty f_\alpha(v) v^{-n/2} dv < \infty$. By Lemma 3.8, we have

$$t_{cb} \leq 1 + \sigma^{-\alpha} \log \left(2c(\alpha, n) + \frac{C}{V_1} \right).$$

The assertion follows from Corollary 3.28. $\quad\square$

Remark 4.9. a) In [46, Section 7] Milnor proved that for any bi-invariant metric on G, the Ricci curvature is non-negative. Theorem 4.6 recovers the non-negativity of Ricci curvature for all heat semigroups with bi-invariant metric. Furthermore, it also applies to subordinated semigroup beyond the Laplacian case.

b) Based on the derivation of heat semigroup $T_t = e^{-\Delta t}$, derivation triple for subordinated group can be constructed as in [14, Section 10.4]. Note that the CLSI of subordinate semigroup was obtained in [22] using a completely different method.

c) By Theorem 4.5, the constant for the heat semigroup T_t^1 has the following explicit form

$$\lambda(1) = \left(2 + 2(n-1) \log 2 + \frac{4}{\sigma} \log(\frac{C}{V}) \right)^{-1}$$

where σ is the spectral gap, V is the volume of unit ball and C is some absolute constant.

It was also pointed out in [46] that Ricci curvature of a left invariant metric is strictly positive if the fundamental group of G is finite. It means for semi-simple Lie groups Theorem 4.3 usually gives better CLSI constant than Theorem 4.6. Nevertheless, for non semi-simple Lie group with zero curvature lower bound, Theorem 4.6 gives us an effective way to obtain lower bounds of CLSI constant.

Example 4.10 *(Circle).* Let $\mathbb{T} = \{z \in \mathbb{C} \mid |z| = 1\}$ be the unit circle. Then $\{z^n | n \in \mathbb{Z}\}$ is a orthonormal basis of $L_2(\mathbb{T})$. The heat semigroup is given by

$$T_t(z^m) = e^{-m^2 t} z^m,$$

and the associated heat kernel is given by $k_t(z) = \sum_{m \in \mathbb{Z}} e^{-m^2 t} z^m$. Now we estimate the cb-return time of T_t:

$$
\begin{aligned}
\|T_t - E_\tau : L_1(\mathbb{T}) \to L_\infty(\mathbb{T})\| &= \| \sum_{m \in \mathbb{Z} \setminus \{0\}} e^{-m^2 t} z^m w^{-m} \|_{L_\infty(\mathbb{T}^2)} \\
&= \| \sum_{m \in \mathbb{Z} \setminus \{0\}} e^{-m^2 t} z^m \|_{L_\infty(\mathbb{T})} \\
&= \| k_t - 1 \|_{L_\infty(\mathbb{T})} \\
&= k_t(e) - 1 \\
&= 2 \sum_{m=1}^{\infty} e^{-m^2 t}.
\end{aligned}
$$

In the above, the first equality follows from the isometric identification

$$
L_\infty(\mathbb{T}^2) \cong L_\infty(\mathbb{T}) \bar{\otimes} L_\infty(\mathbb{T}) \cong B(L_1(\mathbb{T}), L_\infty(\mathbb{T})); \quad (\varphi \otimes \psi)(f) = \left(\int_{\mathbb{T}} \psi(w) f(w) dw \right) \varphi.
$$

The third equality follows from the fact that k_t is a positive definite function on $\mathbb{T}$. Denote $f(t) = 2 \sum_{m=1}^{\infty} e^{-m^2 t}$, so that

$$
t_{cb} = \inf\{t | f(t) \le 1/2\} .
$$

Using standard estimates, we have

$$
2e^{-t} \le f(t) = k_t(0) - 1 \le \frac{2e^{-t}}{1 - e^{-t}} \qquad (t > 1).
$$

These estimates yield concrete bounds of the form

$$
1.38629 \sim \ln 4 \le t_{cb} \le \ln 5 \sim 1.60944.
$$

Numerical calculation shows that $t_{cb} \le 1.41 < 1.5$, and therefore the heat semigroup on $\mathbb{T}$ has $\frac{1}{6}$-CLSI.

Example 4.11 (*d-Torus*). Let $\mathbb{T}^d = \{z = (z_1, z_2, \cdots, z_d) \in \mathbb{C}^d \mid |z_i| = 1, i = 1, \cdots, d\}$ be the d-Torus. For a multi-index $m = (m_1, \cdots, m_d) \in \mathbb{Z}^d$, write $|m|^2 = m_1^2 + m_2^2 + \cdots + m_d^2$ and define the polynomials $z^m := z_1^{m_1} z_2^{m_2} \cdots z_d^{m_d}$. The set $\{z^m | m \in \mathbb{Z}^d\}$ is an orthonormal basis of $L_2(\mathbb{T})$. The heat kernel $k_t^{(d)}$ and heat semigroup T_t on $\mathbb{T}^d$ are given by

$$
k_t^{(d)}(z) = \sum_{m \in \mathbb{Z}^d} e^{-|m|^2 t} z^m, \quad T_t(z^m) = e^{-|m|^2 t} z^m .
$$

We then proceed as in the previous example to compute the CB-return time:

$$\| T_t - E_\tau : L_1(\mathbb{T}^d) \to L_\infty(\mathbb{T}^d) \| = \| \sum_{m \in \mathbb{Z}^d, m \neq 0} e^{-m^2 t} z^m w^{-m} \|_{L_\infty(\mathbb{T}^d \times \mathbb{T}^d)}$$

$$= \| \sum_{m \in \mathbb{Z}^d, m \neq 0} e^{-m^2 t} z^m \|_{L_\infty(\mathbb{T}^d)}$$

$$= \| k_t^{(d)} - 1 \|_{L_\infty(\mathbb{T}^d)}$$

$$= k_t^{(d)}(e) - 1$$

$$= (2 \sum_{m=1} e^{-m^2 t})^d = f(t)^d,$$

where $f(t) := 2 \sum_{m=1} e^{-m^2 t}$ is as in the previous example. Thus we have a CB-return time estimate depending on the dimension d

$$t_{cb}(d) = \inf\{t \mid f(t) \leq 2^{-\frac{1}{d}}\}.$$

Using the same heat kernel estimates as in the previous example, we then conclude that

$$(1 + \frac{1}{d}) \ln 2 \leq t_{cb}(d) \leq \ln(2^{(1+\frac{1}{d})} + 1).$$

For example, $t_{cb}(2) \leq 1.35$ and $t_{cb}(3) \leq 1.26$. (Numerical suggests $t_{cb}(2) \leq 1.08$ and $t_{cb}(3) \leq 0.98$.)

Denote $T_{\mathbb{T}^d, t}$ as the semigroup on $\mathbb{T}^d$, and write $\mathrm{CLSI}(T_{\mathbb{T}^d, t})$ as the optimal CLSI constant of $T_{\mathbb{T}^d, t}$. Note that the CLSI constant $(4t_{cb}(d))^{-1}$ obtained above is monotone increasing for d. This is better than the constant obtained from tenzorization $T_{\mathbb{T}^d, t} = T_{\mathbb{T}, t}^{\otimes d}$ and

$$\mathrm{CLSI}(T_{\mathbb{T}^d, t}) \geq \mathrm{CLSI}(T_{\mathbb{T}, t}) \geq \frac{1}{4t_{cb}(1)}.$$

This leads to the following dimension free estimates.

Theorem 4.12. *Let $d \geq 1$ and $\mathbb{T}^d$ be the unit d-torus. The heat semigroup on $\mathbb{T}^d$ (in the above normalization) satisfies complete Fisher monotonicity and λ-CLSI for $\lambda = (4 \ln 3)^{-1}$.*

Proof. By Example 4.11, we have

$$\mathrm{CLSI}(T_{\mathbb{T}^d, t}) \geq \left(4 \inf\{ t \mid (2 \sum_{m=1} e^{-m^2 t})^d \leq 1/2 \}\right)^{-1} \geq \left(4 \ln(2^{(1+\frac{1}{d})} + 1)\right)^{-1}$$

For $m \leq d$, $\mathbb{T}^d = \mathbb{T}^m \times \mathbb{T}^{d-m}$. Consider the embedding $\pi_{m,d} : C(\mathbb{T}^m) \to C(\mathbb{T}^d)$

$$\pi_{m,d}(f) = f \otimes 1_{d-m} \ , f \in C(\mathbb{T}^m) \ ,$$

where 1_{d-m} is the identity function on $\mathbb{T}^{d-m}$. Namely, $\pi_{m,d}(f)(z_1, \cdots, z_d) = f(z_1, \cdots, z_m)$. It is clear that

$$\pi_{m,d} \circ T_{\mathbb{T}^m,t} = T_{\mathbb{T}^d,t} \circ \pi_{m,d} \ .$$

Hence the heat semigroup $T_{\mathbb{T}^m,t}$ on m-torus is a sub-semigroup for $T_{\mathbb{T}^d,t}$ on d-torus. We have for any $d \geq m$,

$$\mathrm{CLSI}(T_{\mathbb{T}^m,t}) \geq \mathrm{CLSI}(T_{\mathbb{T}^d,t}) \geq \left(4 \ln(2^{(1+\frac{1}{d})} + 1) \right)^{-1} \ .$$

Taking $d \to \infty$, we have $\mathrm{CLSI}(T_{\mathbb{T}^m,t}) \geq (4 \ln 3)^{-1}$ for any m. That completes the proof. $\quad\square$

Remark 4.13. It was proved by Weissler [59] that on the circle $\mathbb{T}$, both the heat semigroup $T_t(z^m) = e^{-m^2} z^m$ and the Possion semigroup $P_t(z^m) = e^{-|m|t} z^m$ satisfies sharp 1-LSI hence sharp 1-MLSI (because the spectral gap is 1). We show in the second part [7] of this series that the Possion semigroup P_t on $\mathbb{T}$ satisfies sharp GRic ≥ 1 and hence sharp 1-CLSI. The sharp CLSI constant for the heat semigroup on $\mathbb{T}$ remains open.

Example 4.14 *(Finite groups)*. Let G be a finite group and $l_\infty(G)$ be the function space on G equipped with counting probability measure. Let

$$T_t : l_\infty(G) \to l_\infty(G) \ , (T_t f)(g) = \sum_{g \in G} k_t(g^{-1}h) f(h)$$

be a symmetric central Markov semigroup with kernel function $k_t \in Zl_1(G)$. Let A be generator of T_t, which acts on the $l_2(G)$:

$$A : l_2(G) \to l_2(G), A(e_h) = \sum_{g \in G} A_{g,h} e_g$$

The entries of A are given by

$$A_{g,h} = \begin{cases} \sum_{h \neq g} w_{g,h}, & \text{if } h = g \\ -w_{g,h}, & \text{otherwise,} \end{cases}$$

where $w_{g,h} > 0$ are the transition rates. If T_t is symmetric and central,

$$w_{g,h} = w_{h,g} = w_{sg,sh} = w_{gs,hs} \ , \forall \ s, g, h \in G \ .$$

Here we use the derivation of finite Markov chain from [21]. Denote $B = \sum_{g \neq h} \sqrt{w_{g,h}} e_{g,h}$, where $e_{g,h}$ are matrix units in $B(l_2(G))$. Consider the standard embedding $\pi : l_\infty(G) \hookrightarrow B(l_2(G))$ as diagonal matrices $\pi(f) = \sum_g f(g) e_{g,g}$. We have the following derivation.

$$\delta : l_\infty(G) \to B(l_2(G)) \ , \ \delta(f) = \sum_{g,h} b_{g,h}(f(h) - f(g))e_{g,h} = i[B, \pi(f)]$$

For the gradient form,

$$2\Gamma(e_g, e_h) = \left(e_g^*(Ae_h) + (Ae_g)^* e_h - A(e_g e_h) \right)$$

$$= \begin{cases} \sum_{s \neq g} w_{s,g}(e_s + e_g), & \text{if } g = h \\ -w_{g,h}(e_g + e_h), & \text{otherwise.} \end{cases}$$

Note that $[B, \pi(e_g)] = \sum_{s \neq g} \sqrt{w_{s,g}}(e_{s,g} - e_{g,s})$. Then for $g \neq h$

$$E([B, \pi(e_g)]^*[B, \pi(e_h)]) = E\Big(\big(\sum_{s \neq g} \sqrt{w_{s,g}}(e_{g,s} - e_{s,g})\big)\big(\sum_{r \neq h} \sqrt{w_{r,h}}(e_{r,h} - e_{h,r})\big)\Big)$$

$$= E\Big(\sum_r \sqrt{w_{r,h}}\sqrt{w_{r,g}}e_{g,h} - \sum_s \sqrt{w_{s,g}}\sqrt{w_{g,h}}e_{s,h}$$

$$- \sum_r \sqrt{w_{h,g}}\sqrt{w_{r,h}}e_{g,r} + \delta_{g,h}\sum_{s,r} \sqrt{w_{s,g}}\sqrt{w_{r,g}}e_{s,r} \Big)$$

$$= -w_{g,h}e_h - w_{h,g}e_g$$

For $g = h$,

$$E([B, \pi(e_g)]^*[B, \pi(e_h)]) = E\Big(\big(\sum_{s \neq g} \sqrt{w_{s,g}}(e_{g,s} - e_{s,g})\big)\big(\sum_{r \neq h} \sqrt{w_{r,h}}(e_{r,h} - e_{h,r})\big)\Big)$$

$$= \sum_{r \neq g} w_{r,g}e_g + \sum_{s \neq g} w_{s,g}e_s$$

$$= \sum_s w_{s,g}(e_s + e_g)$$

Thus we have verified that

$$\Gamma(e_g, e_h) = E(\delta(e_g)^*\delta(e_h)) \ ,$$

which extends bi-linearly to $l_\infty(G) \times l_\infty(G)$. Now we have

$$\delta \circ T_t(e_g) = \delta(\sum_r k_t(r)e_{gr^{-1}})$$

$$= \sum_r k_t(r) \sum_{s \neq gr^{-1}} \sqrt{w_{s,gr^{-1}}}(e_{s,gr^{-1}} - e_{gr^{-1},s})$$

$$= \sum_r k_t(r)U_r^* \Big(\sum_{sr \neq g} \sqrt{w_{sr,g}}(e_{sr,g} - e_{g,sr}) \Big)U_r$$

$$= \sum_r k_t(r) U_r^* \Big(\delta(e_g) \Big) U_r$$

$$= \hat{T}_t \circ \delta(e_g)$$

In the third equality above we used the central property $w_{s,gr^{-1}} = w_{sr,g}$. The extension semigroup on $B(l_2(G))$ is

$$\hat{T}_t(\rho) = \sum_r k_t(r) U_r \rho U_r^* ,$$

where $U_r e_g = U_r e_{gr}$ is the right shifting unitary. $\hat{T}_t$ is a extension of $T_t : l_\infty(G) \to l_\infty(G)$ on $B(l_2(G))$. Indeed,

$$T_t(e_{g,g}) = \sum_r k_t(r) e_{gr^{-1},gr^{-1}} = \sum_r k_t(r) U_{r^{-1}} e_{g,g} U_{r^{-1}}^* = \sum_r k_t(r) U_r^* e_{g,g} U_r .$$

This verifies that T_t satisfies 0-GRic via a construction different from Lemma 4.6.

Corollary 4.15. *Let $T_t : l_\infty(G) \to l_\infty(G)$ be a central Markov semigroup with spectral gap σ. Then T_t satisfies $\mathrm{GRic} \geq 0$, complete Fisher monotonicity and λ-CLSI for*

$$\lambda = \frac{\sigma}{4(\log 2|G|)} .$$

Proof. This follows from Theorem 3.28, Proposition 3.10 and $D_{cb}(l_\infty(G)||\mathbb{C}) = |G|$. $\quad\square$

4.4. Generalized depolarizing semigroups

Let $\mathcal{N} \subset \mathcal{M}$ be a subalgebra and let $E : \mathcal{M} \to \mathcal{N}$ be the conditional expectation. We now discuss curvature bounds and MLSI (resp. CLSI) constants for the generalized depolarizing semigroup

$$T_t(\rho) = e^{-\lambda t}\rho + (1 - e^{-\lambda t})E(\rho) .$$

The generator is $A = \lambda(I - E)$ whose spectral gap is clearly λ (here I is the identity operator on $L_2(\mathcal{M})$). In the following we show that T_t has $\lambda/2$-GE. This result is independently obtained by Melchior Wirth and Haonan Zhang and the case for ergodic depolarizing semigroup on matrix algebras was obtained in [15, Section 3.4].

Theorem 4.16. *The generalizing depolarizing semigroup*

$$T_t(\rho) = e^{-\lambda t}\rho + (1 - e^{-\lambda t})E(\rho) ,$$

satisfies $(\lambda/2)$-GE.

Proof. Let $(\mathcal{A}, \hat{\mathcal{M}}, \delta)$ be a derivation triple of T_t. Since $\delta(x) = 0$ for $x \in \mathcal{N}$, we have for $x \in \mathcal{A}$,

$$\delta(T_t(x)) = \delta\big(e^{-\lambda t}(x - E(x)) + E(x)\big) = e^{-\lambda t}\delta(x) .$$

Then we have

$$\| \delta(T_t(x)) \|_\rho^2 = \| e^{-\lambda t}\delta(x) \|_\rho^2 = e^{-2\lambda t} \| \delta(x) \|_\rho^2,$$

where

$$\| \delta(x) \|_\rho^2 = \int\limits_0^1 \tau(\delta(x)^* \rho^s \delta(x) \rho^{1-s}) ds.$$

It follows from Lieb's concavity theorem [41] that for each $s \in [0,1]$,

$$(\rho, \sigma) \to \tau(\delta(x)^* \rho^s \delta(x) \sigma^{1-s})$$

is jointly concave for (ρ, σ). For $\rho_t := T_t(\rho) = e^{-\lambda t} + (1 - e^{\lambda t})E(\rho)$,

$$\tau(\delta(x)^* \rho_t^s \delta(x) \rho_t^{1-s}) \geq e^{-\lambda t}\tau(\delta(x)^* \rho^s \delta(x) \rho^{1-s}) + (1 - e^{\lambda t})\tau(\delta(x)^* E(\rho)^s \delta(x) E(\rho)^{1-s})$$

Integrating over s,

$$\| \delta(x) \|_{T_t(\rho)}^2 \geq e^{-\lambda t} \| \delta(x) \|_\rho^2 + (1 - e^{\lambda t}) \| \delta(x) \|_{E(\rho)}^2 \geq e^{-\lambda t} \| \delta(x) \|_\rho^2 .$$

Then

$$\| \delta(T_t(x)) \|_\rho^2 = e^{-2\lambda t} \| \delta(x) \|_\rho^2 \leq e^{-2\lambda t} e^{\lambda t} \| \delta(x) \|_{T_t(\rho)}^2 = e^{-\lambda t} \| \delta(x) \|_{T_t(\rho)}^2$$

which proves the gradient estimates. $\square$

Remark 4.17. In second part of the series [7], we will prove a stronger result that $T_t = e^{-\lambda(I-E)t}$ satisfies GRic $\geq \lambda/2$ based the free product property discussed there.

Note that the above theorem implies the generator $A = (I - E)$ has 1/2-CLSI. This can be verified directly via its Fisher information

$$\begin{aligned}
I(\rho) &= \tau((I - E)(\rho)\log\rho) = \tau(\rho\log\rho) - \tau(E(\rho)\log\rho) \\
&= \tau(\rho\log\rho - \rho\log E(\rho)) + \tau(\rho\log E(\rho) - E(\rho)\log\rho) \\
&= \tau(\rho\log\rho - \rho\log E(\rho)) + \tau(E(\rho)\log E(\rho) - E(\rho)\log\rho) \\
&= D(\rho||E(\rho)) + D(E(\rho)||\rho) \geq D(\rho||E(\rho)) \tag{25}
\end{aligned}$$

where in the third equality we used the definition of the conditional expectation. It follows from $D(E(\rho)||\rho) \geq 0$ that $A = (I - E)$ has 1/2-MLSI and also 1/2-CLSI by the same argument for $(I - E) \otimes \mathrm{id}$. In the following discussion, we denote $\mathrm{MSLI}(A)$ (resp. $\mathrm{CLSI}(A)$ and $\mathrm{GE}(A)$) as the optimal constant λ of MSLI (resp. CLSI and GE) for the generator A.

Example 4.18 *(Depolarizing semigroup).* Let M_d be the algebra of $d \times d$ matrix. Consider the depolarizing semigroup

$$D_t : M_d \to M_d \ , \ D_t(\rho) = e^{-t}\rho + (1 - e^{-t})\tau_d(\rho)1,$$

where $\tau_d(x) = \frac{1}{d}\mathrm{Tr}(x)$ is the normalized matrix trace on M_d. It is proved in [37] that the optimal LSI constant is

$$\mathrm{LSI}(I - \tau_d) = \frac{2 - 4/d}{\log(d-1)} \ , \ \mathrm{LSI}(I - \tau_2) = 1$$

This implies

$$\mathrm{MLSI}(I - \tau_d) \geq \frac{2 - 4/d}{\log(d-1)} \ .$$

Here we show that $\mathrm{GE}(I - \tau_3) \leq \mathrm{MLSI}(I - \tau_3) < 1$. In M_3, we choose the normalized density $\rho = \frac{3}{2}e_1 + \frac{3}{4}e_2 + \frac{3}{4}e_3$ where e_1, e_2, e_3 are orthogonal rank one projections. Then

$$D(\rho||1) = \frac{1}{2}\log(3/2) + \frac{1}{4}\log(3/4) + \frac{1}{4}\log(3/4) = \frac{1}{2}\log(9/8) = \log(3/2\sqrt{2})$$

$$D(1||\rho) = \frac{1}{3}\log(2/3) + \frac{1}{3}\log(4/3) + \frac{1}{3}\log(4/3) = \frac{1}{3}\log(32/27)$$

$$= \log(2^{5/3}/3) < \log(3/2\sqrt{2})$$

This means $D(\rho||1) > D(1||\rho)$ and

$$I(\rho) = D(\rho||1) + D(1||\rho) < 2D(\rho||1)$$

This implies on M_3, the depolarizing semigroup $A = I - \tau_3$ does not have 1-MLSI nor 1-GE. Similar examples can be found for dimension $d \geq 3$.

Remark 4.19. By (25), the optimal MLSI constant is

$$\mathrm{MLSI}(I - \tau_d) = \frac{1}{2}(1 + \inf_{\rho \in S(M_d)} \frac{D(1||\rho)}{D(\rho||1)}) \ .$$

It is clear that $(I - \tau_d)$ has the same MLSI constant for the classical depolarizing semigroup

$$S_t : l_\infty^d \to l_\infty^d \ , \ S_t(f) = e^{-t} f + (1 - e^{-t}) \frac{(\sum_i f(i))}{d} 1 \ .$$

Maas and Erbar showed in [21] that $GE(S_t) \geq \frac{1}{2} + \frac{1}{2d}$.

We show that the above GE constant also holds for M_n. We are indebt to Melchior Wirth for pointing out our earlier mistake on the following proposition.

Proposition 4.20. *The d-dimensional depolarizing semigroup*

$$D_t : M_d \to M_d \ , \ D_t(\rho) = e^{-t} \rho + (1 - e^{-t}) \tau_d(\rho) 1$$

satisfies $\frac{1}{2} + \frac{1}{2d}$-GE.

Proof. Note that for any derivation δ of T_t,

$$\delta(a) = \delta(a - \tau_d(a)) \ , \delta(T_t(a)) = e^{-t}(a - \tau_d(a)) \ .$$

Then

$$\| \delta(T_t(a)) \|_{L_2(\hat{\mathcal{M}}, \rho)}^2 = e^{-2t} \| \delta(a) \|_{L_2(\hat{\mathcal{M}}, \rho)}^2$$

Let $\alpha > 0$. The D_t satisfies $\frac{1}{2} + \frac{1}{2\alpha}$-GE means that for any $a \in M_d$,

$$e^{-2t} \| \delta(a) \|_{L_2(\hat{\mathcal{M}}, \rho)}^2 = \| \delta(T_t(a)) \|_{L_2(\hat{\mathcal{M}}, \rho)}^2 \leq e^{-2(\frac{1}{2} + \frac{1}{2\alpha})t} \| \delta(a) \|_{L_2(\hat{\mathcal{M}}, T_t(\rho))}^2$$

This is equivalent to the function

$$h(t) := e^{(1 - \frac{1}{\alpha})t} \| \delta(a) \|_{L_2(\hat{\mathcal{M}}, T_t(\rho))}^2$$

is increasing. Denote the function $f(t) := \| \delta(a) \|_{L_2(\hat{\mathcal{M}}, T_t(\rho))}^2$. Write $\rho_t = T_t(\rho)$ for a density ρ. We have the derivative,

$$\frac{d}{dt} \rho_t = -(I - \tau)(\rho) = (1 - \rho)$$

and ($\hat{\tau}$ is the trace on the derivation triple)

$$\frac{df(t)}{dt} \Big|_{t=0} = \frac{d}{dt} \Big|_{t=0} \Big(\int_0^1 \hat{\tau}(\delta(a)^* \rho_t^{1-s} \delta(a) \rho_t^s) ds \Big)$$

$$= \int_0^1 \Big(\frac{d}{dt} \Big|_{t=0} \hat{\tau}(\delta(a)^* \rho_t^{1-s} \delta(a) \rho_t^s) \Big) ds$$

Let $\rho = \sum_j p_j e_j$ be the orthogonal decomposition of ρ. By double operator integral,

$$\frac{d}{dt}\rho_t^s|_{t=0} = \sum_{j,k} \frac{p_j^s - p_k^s}{p_j - p_k} e_j(1-\rho)e_k$$

$$= (1-\rho) \sum_{j,k} \frac{p_j^s - p_k^s}{p_j - p_k} e_j e_k$$

$$= \lambda s(1-\rho)\rho^{s-1}$$

and similarly $\frac{d}{dt}\rho_t^{1-s}|_{t=0} = \lambda(1-s)(1-\rho)\rho^{-s}$. For a bi-variable function $F : (0,\infty) \times (0,\infty) \to (0,\infty)$, we introduce the notation

$$I_{F,\rho}(X) = \sum_{j,k} F(p_j, p_k)e_j X e_k \ .$$

Then

$$\frac{d}{dt}(\rho_t^{1-s}\delta(a)\rho_t^s)|_{t=0} = (1-\rho)I_{F_s,\rho}(\delta(a)) + I_{G_s,\rho}(\delta(a))(1-\rho)$$

where $F_s(x,y) = (1-s)(1-x)x^{-s}y^s$ and $G_s = sx^{1-s}y^{s-1}(1-y)$. Integrating over s on $[0,1]$, we have

$$\frac{df(t)}{dt}|_{t=0} = \frac{d}{dt}|_{t=0}\Big(\int\limits_0^1 \hat{\tau}(\delta(x)^*\rho_t^{1-s}\delta(x)\rho_t^s)ds\Big)$$

$$= \int\limits_0^1 \tau(\delta(x)^*I_{F_s,\rho}(\delta(x))) + \tau(\delta(x)^*I_{G_s,\rho}(\delta(x)))ds$$

$$= \tau(\delta(x)^*I_{H,\rho}(\delta(x)))$$

where H is the function given by

$$H(x,y) = \int\limits_0^1 F_s(x,y) + G_s(x,y)ds = \frac{(x-y)(x-y-xy(\log(x)-\log(y)))}{xy(\log(x)-\log(y))^2}$$

$$= \frac{(x-y)}{\log(x)-\log(y)}\frac{(x-y-xy(\log(x)-\log(y)))}{xy(\log(x)-\log(y))}$$

$$= \frac{(x-y)}{\log(x)-\log(y)}\Big(\frac{(x-y)}{xy(\log(x)-\log(y))}-1\Big)$$

here log is natural log. On the other hand,

$$f(0) = \int\limits_0^1 \tau(\delta(a)^* \rho^{1-s} \delta(a) \rho^s) \,,$$

$$= \tau(\delta(a)^* I_{J,\rho} \delta(a))$$

where $J(x,y) = \dfrac{x-y}{\log x - \log y}$. Then the derivative of $h(t) = e^{(1-\frac{1}{\alpha})t} f(t)$ is

$$h'(0) = (1 - \frac{1}{\alpha}) f(0) + f'(0)$$

$$= (1 - \frac{1}{\alpha}) \tau(\delta(x)^* I_{J,\rho}(\delta(x))) + \tau(\delta(x)^* I_{H,\rho}(\delta(x)))$$

Thus it suffices to require $(1 - \frac{1}{\alpha})J + H$ be a positive function on the spectrum of ρ. Indeed,

$$(1 - \frac{1}{\alpha}) J(x,y) + H(x,y) = (1 - \frac{1}{\alpha}) \frac{(x-y)}{\log(x) - \log(y)}$$

$$+ \frac{(x-y)}{\log x - \log y} \Big(\frac{(x-y)}{xy(\log x - \log y)} - 1 \Big)$$

$$= \frac{(x-y)}{\log x - \log y} \Big(\frac{(x-y)}{xy(\log x - \log y)} - \frac{1}{\alpha} \Big)$$

Because $\frac{(x-y)}{\log x - \log y} \geq 0$, it suffices to require

$$\frac{(x-y)}{xy(\log x - \log y)} - \frac{1}{\alpha} > 0$$

or equivalently

$$\frac{xy(\log x - \log y)}{x - y} \leq \alpha \,.$$

Here for M_d, the domain of (x,y) is contained in $S_d := \{0 \leq x, y \leq d\}$ since x, y are eigenvalues of a normalized density $\rho \in M_d$. By elementary calculus, one can show

$$\max_{(x,y) \in S_d} \frac{xy(\log x - \log y)}{x - y} = d \,.$$

Thus α can be d and we finish the proof. $\square$

We now use a similar idea to consider the MLSI constant of $D_t \otimes \mathrm{id}_2 : M_2 \otimes M_2 \to M_2 \otimes M_2$ where D_t is the depolarizing on M_2. Let $E : M_2 \otimes M_2 \to M_2 \otimes M_2$, $E(\rho) = (\tau \otimes \mathrm{id}(\rho)) \otimes 1$ be the partial trace map. Consider the basis of Bell states

$$|\phi_1\rangle = \frac{1}{\sqrt{2}}(|0\rangle|0\rangle + |1\rangle|1\rangle) \,, \ |\phi_2\rangle = \frac{1}{\sqrt{2}}(|0\rangle|0\rangle - |1\rangle|1\rangle)$$

$$|\phi_3\rangle = \frac{1}{\sqrt{2}}(|0\rangle|1\rangle + |0\rangle|1\rangle) \,, \ |\phi_4\rangle = \frac{1}{\sqrt{2}}(|0\rangle|1\rangle - |0\rangle|1\rangle)$$

Using the identification

$$|0\rangle|0\rangle \to |1\rangle \,, |0\rangle|1\rangle \to |2\rangle \,, |1\rangle|0\rangle \to |3\rangle \,, |1\rangle|1\rangle \to |4\rangle$$

we have the densities in $M_4 \cong M_2 \otimes M_2$ represented as

$$\phi_1 = \begin{bmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{bmatrix} \qquad \phi_2 = \begin{bmatrix} \frac{1}{2} & 0 & 0 & -\frac{1}{2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -\frac{1}{2} & 0 & 0 & \frac{1}{2} \end{bmatrix}$$

$$\phi_3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \phi_4 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Now we choose the state $\rho = \frac{5}{8}\phi_1 + \frac{1}{8}(\phi_2 + \phi_3 + \phi_4)$. The reduced density is

$$E(\rho) = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \otimes \frac{1}{2} = \frac{1}{4} 1 \otimes 1 \,.$$

Thus

$$D(\rho||E(\rho)) = D(\rho||\frac{1}{4}) = \frac{5}{8}\log(5/2) + \frac{3}{8}\log(1/2) \simeq 0.313$$

$$D(E(\rho)||\rho) = D(\frac{1}{4}||\rho) = \frac{1}{4}\log(2/5) + \frac{3}{4}\log 2 \simeq 0.291$$

Then we have $D(\rho||E(\rho)) > D(E(\rho)||\rho)$, which implies that $\mathrm{GE}((I - \tau_2) \otimes \mathrm{id}_{M_2}) \leq \mathrm{MLSI}((I - \tau_2) \otimes \mathrm{id}_{M_2}) < 1$. Note that $\mathrm{MLSI}(I - \tau_2) = 1$. We have the following corollary.

Proposition 4.21. *Let D_t be the depolarizing semigroup on M_2 and $(I - \tau_2)$ be its generator. Then*

$$\mathrm{MLSI}((I - \tau_2) \otimes \mathrm{id}_{M_2}) < 1 = \mathrm{MLSI}(I - \tau_2)$$

In particular, $\mathrm{CLSI}(I - \tau_2) < \mathrm{MLSI}(I - \tau_2)$.

For classical Markov semigroups, the MLSI is stable under tensorization [3]. The above example shows that tensorization of MLSI does not hold for quantum cases if we allow non-ergodic semigroup.

4.5. Schur multipliers

Let M_m be the $m \times m$ matrix algebra and $a = (a_{ij})_{i,j=1}^m \in M_m$. The Schur multiplier of a is

$$T_a : M_n \to M_n \ , \ T_a(x_{ij}) = (a_{ij}x_{ij})$$

Consider a semigroup of Schur multiplier $T_t : M_n \to M_n \ , \ T_t((x_{ij})) = (e^{-b_{ij}t}x_{ij})$. The generator is the Schur multiplier of $b = (b_{ij})$,

$$A((x_{ij})) = (b_{ij}x_{ij}) \ .$$

By Schoenberg's theorem [54], T_t is a symmetric quantum Markov semigroup (unital completely positive and self-adjoint) if and only if $b_{ii} = 0, b_{ij} = b_{ji} \geq 0$ and *conditionally negative definite*, i.e. for any real sequence $(c_1, \cdots, c_m)$ with $\sum_{i=1}^m c_i = 0$,

$$\sum_{i,j=1}^m c_i c_j b_{ij} \leq 0 \ .$$

Moreover, there exists a real Hilbert space H and a family of vector $b(1), \cdots, b(n) \in H$ such that

$$b_{ij} = \| b(i) - b(j) \|^2 \ .$$

For T_t, the fixed point subalgebra $\mathcal{N}$ is

$$\mathcal{N} = \{(x_{ij}) \in M_n \mid x_{ij} = 0 \text{ for all } (i,j) \text{ that } b_{ij} \neq 0\} \ ,$$

where $e_{ij} \in M_m$ are the matrix units. It is clear that the diagonal matrices $l_\infty^m \subset \mathcal{N}$. Thus T_t are always non-ergodic. Because e_{ij} are eigenvectors of the generator A with eigenvalue b_{ij}, the spectral gap is

$$\sigma = \min\{ b_{ij} \mid b_{ij} \neq 0 \} \ .$$

The gradient form is given by

$$\Gamma(e_{ij}, e_{lk}) = \frac{1}{2}\delta_{il}(b_{ij} + b_{lk} - b_{jk})e_{jk} \ .$$

Here δ_{il} is the Kroenecker-delta notation. For $i = l$, we have

$$\Gamma(e_{ij}, e_{ik}) = \frac{1}{2}(\| b(i) - b(j) \|^2 + \| b(i) - b(k) \|^2 - \| b(j) - b(k) \|^2)e_{jk}$$

$$= \frac{1}{2}(\| b(i) - b(j) \|^2 + \| b(i) - b(k) \|^2 - \| b(j) - b(k) \|^2)e_{jk}$$

$$= \langle b(i) - b(j), b(i) - b(k)\rangle e_{jk}$$

We recall the derivation structure from [1]. Recall that for a real Hilbert space H, an H-isonormal process on a standard probability space (Ω, m) is a linear mapping $W : H \to L_0(\Omega)$ satisfying the following properties:

i) for any $v \in H$, the random variable $W(v)$ is a centered real Gaussian.
ii) for any $v_1, v_2 \in H$, we have $E_\Omega(W(v_1)W(v_2)) = \langle v_1, v_2 \rangle_H$
iii) The linear span of the products $\{W(v_1)W(v_2)\cdots W(v_n) \mid v_1, \cdots, v_n \in H\}$ is dense in the real Hilbert space $L_2(\Omega)$

Here $L_0(\Omega)$ denote the space of measurable functions on Ω. Now we define the derivation

$$\delta : M_m \to M_m \otimes L_2(\Omega) \, , \delta(e_{ij}) = e_{ij} \otimes \sqrt{-1}(W(b(i)) - W(b(j))) \, .$$

We verify that δ is a derivation,

$$\delta(e_{ij})e_{jk} + e_{ij}\delta(e_{jk}) = e_{ik} \otimes \sqrt{-1}(W(b(i)) - W(b(j))) + e_{ik} \otimes \sqrt{-1}(W(b(j)) - W(b(k)))$$
$$= e_{ik} \otimes \sqrt{-1}(W(b(i)) - W(b(k))) = \delta(e_{ik}) = \delta(e_{ij}e_{jk}).$$

Moreover for the gradient form

$$E(\delta(e_{ij})^*\delta(e_{lk})) = E\Big(\big(e_{ji} \otimes (W(b(i)) - W(b(j)))\big)\big(e_{lk} \otimes (W(b(l)) - W(b(k)))\big)\Big)$$
$$= \delta_{il}e_{jk} \otimes E\Big((W(b(i)) - W(b(j)))(W(b(i)) - W(b(k)))\Big)$$
$$= \delta_{il}\langle b(i) - b(j), b(i) - b(k)\rangle e_{jk} \, .$$

Then it is readily seen that

$$\delta \circ T_t = (T_t \otimes id_\Omega) \circ \delta \, ,$$

where $T_t \otimes id_\Omega$ is the extension of T_t on $M_m \otimes L_\infty(\Omega)$. By Theorem 3.25, this implies T_t satisfies 0-GRic. Combined with CB-return time estimates in Proposition 3.10, we have

Theorem 4.22. *Let* $T_t : M_m \to M_m, T_t((x_{ij})) = (e^{-b_{ij}t}x_{ij})$ *be a symmetric quantum Markov semigroup of Schur multipliers. Then* T_t *satisfies* 0-GRic *and complete Fisher monotonicity. Denote* $\sigma = \min\{b_{ij} \mid b_{ij} \neq 0\}$ *as the spectral gap of* T_t. *Then* T_t *satisfies* λ-CLSI *with constant*

$$\lambda = \frac{\sigma}{4(D_{cb}(M_m||\mathcal{N}) + \log 2)}$$

In particular, $D_{cb}(M_m||\mathcal{N}) \leq D_{cb}(M_m||l_\infty^m) = \log m$.

4.6. Random unitary channels

A CPTP map $T : M_m \to M_m$ is called a *random unitary channel* if it is a convex combination of unitary conjugations,

$$T(\rho) = \sum_{j=1}^n p_j U_j \rho U_j^* \quad (p_i \geq 0, \ \sum_i p_i = 1) \ .$$

In this subsection, we discuss semigroups of random unitary channels arising from group representations. Let G be a finite group. Recall that a projective unitary representation $U : G \to U(M_m)$ satisfies

$$U_g U_h = \sigma(g, h) U_{gh} \ , \ \forall \ g, h \in G$$

where $\sigma : G \times G \to \mathbb{C}$ is a group 2-cocycle with $|\sigma(g, h)| = 1$. Let $T_t : M_m \to M_m$ be the quantum Markov semigroup given by

$$T_t(\rho) = \frac{1}{|G|} \sum_g k_t(g) U_g \rho U_g^* \ ,$$

where $k_t(g)$ is the weight function that satisfies $k_t(g) \geq 0$, $\sum_g k_t(g) = 1$ and

$$k_{t+s}(g) = \frac{1}{|G|} \sum_h k_t(gh^{-1}) k_s(h) = (k_t \star k_s)(g) \ .$$

Thus k_t forms the right invariant kernel on G. Let

$$S_t : l_\infty(G) \to l_\infty(G), S_t(f)(g) = \sum_h k_t(gh^{-1}) f(h)$$

be the right invariant Markov semigroup on $l_\infty(G)$. We have the transference

$$
\begin{array}{ccc}
l_\infty(G, M_m) & \overset{S_t \otimes \mathrm{id}_{M_m}}{\longrightarrow} & l_\infty(G, M_m) \\
\uparrow \alpha & & \uparrow \alpha \\
M_m & \overset{T_t}{\longrightarrow} & M_m
\end{array}
\quad , \tag{26}
$$

where $\alpha : M_m \to l_\infty(G, M_m), \alpha(x)(g) = U_g x U_g^*$ is a trace preserving $*$-monomorphism. Thus $T_t = (S_t \otimes \mathrm{id})|_{\alpha(M_m)}$ is a subsystem of the semigroup $(S_t \otimes \mathrm{id}_{M_m})$.

Theorem 4.23. *Let G be a finite group and let $U : G \to M_m$ be a projective unitary representation. Let $T_t : M_m \to M_m$ be a quantum Markov semigroup given by*

$$T_t(\rho) = \frac{1}{|G|} \sum_g k_t(g) U_g \rho U_g^* \ .$$

Suppose k_t is central and T_t has spectral gap σ. Then T_t satisfies complete Fisher monotonicity and λ-CLSI with constant

$$\lambda = \frac{\sigma}{4(\log 2m^2)}$$

Proof. If k_t are central, it follows from Theorem 4.6 that the classical semigroup S_t satisfies complete Fisher monotonicity, which pass to T_t as a subsystem. The CLSI constant follows from Proposition 3.10 and $D_{cb}(M_m||\mathcal{N}) \leq D_{cb}(M_m||\mathbb{C}) = m^2$. $\quad\square$

Example 4.24. Recall the m-dimensional generalized Pauli matrices are

$$X|j\rangle = |j+1\rangle \,, Z|j\rangle = e^{\frac{2\pi ij}{m}}|j\rangle \,.$$

It is clear that $\{X^k Z^l\}$ forms a projective representation of $\mathbb{Z}_m^2$. Since $\mathbb{Z}_m^2$ is abelian, so every function on $\mathbb{Z}^2$ is a central. Thus the above theorem applies to every semigroup of random Pauli unitaries

$$T_t(\rho) = \frac{1}{m^2}\sum_{j,l} k_t(j,l)X^j Z^l \rho (X^j Z^l)^* \,.$$

Appendix A

In this appendix we provide the approximation lemmas in terms of entropy. We start with a standard density argument.

Lemma A.1. *Suppose $\mathcal{A} \subset \mathcal{M}$ is a w^*-dense unital $*$-subalgebra $\mathcal{A} \subset \mathcal{M}$. Denote B as the unit ball of $\mathcal{M}$. Then $\mathcal{A}$ is norm dense in $L_2(\mathcal{M})$ and $L_1(\mathcal{M})$. Moreover, the positive part $\mathcal{A}_+$ (resp. $\mathcal{A}_+ \cap B$) is dense in $L_1(\mathcal{M})_+$ (resp. $L_1(\mathcal{M})_+ \cap B$).*

Proof. By Kaplansky density theorem (cf. [56, Theorem 4.8]), $\mathcal{A} \cap B$ is also strong operator topology (SOT) dense in $\mathcal{M} \cap B$. Then for any $\xi \in L_2(\mathcal{M})$, we have a net $(x_\alpha) \subset \mathcal{A}$ such that $x_\alpha \to \xi$ in SOT topology and hence norm dense in $L_2(\mathcal{M})$. For L_1, it suffices to show that $\mathcal{A}$ is L_1-norm dense in $L_1(\mathcal{M}) \cap \mathcal{M}$. Indeed, for any positive $\rho \in L_1(\mathcal{M}) \cap B$, we take $x_\alpha \to \rho^{1/2}$ in SOT topology and in $L_2(\mathcal{M}) \cap B$. Then for any subsequence $(x_n) \subset (x_\alpha)$,

$$\lim_{n\to\infty} \|x_n^* x_n - \rho\|_1 \leq \|x_n^* x_n - \rho^{1/2} x_n\|_1 + \|\rho^{1/2} x_n - \rho\|_1$$

$$\leq \lim_{n\to\infty} \|x_n^* - \rho^{1/2}\|_2 \|x_n\|_2 + \|\rho^{1/2}\|_2 \|\rho^{1/2} - x_n\|_2 = 0$$

Then $x_n^* x_n \to \rho \in L_1(\mathcal{M})$ and $x_n^* x_n \in \mathcal{A} \cap B$ since $\mathcal{A}$ is a $*$-subalgebra. $\quad\square$

The next lemma shows that the relative entropy is continuous in L_1-norm for bounded invertible densities.

Lemma A.2. *Let $\rho \in S_B(\mathcal{M})$ and ρ_n be a sequence in $L_1(\mathcal{M})_+$ such that $\|\rho_n - \rho\|_1 = 0$. Suppose there exist $m, M > 0$ such that $m1 \leq \rho_n \leq M1$ for any n. Then $\lim\limits_{n \to \infty} H(\rho_n) = H(\rho)$ and $\lim\limits_{n \to \infty} D(\rho_n || \mathcal{N}) = D(\rho || \mathcal{N})$.*

Proof. We assume that $m1 \leq \rho \leq M1$. The lower semi-continuity inherited from relative entropy,

$$H(\rho) = D(\rho||1) \leq \liminf_{n \to \infty} D(\rho_n||1) = \liminf_{n \to \infty} H(\rho)$$

For the upper continuity, we use Klein's inequality [60, Theorem 5.9] for $h(s) = s \log s$

$$H(\rho_n) - H(\rho) = \tau(h(\rho_n) - h(\rho)) \leq \tau(h'(\rho_n)(\rho_n - \rho)),$$

where $h'(s) = 1 + \log s$ is the derivative of h. Because $m1 \leq \rho_n \leq M1$, we have $\|h'(\rho_n)\|_\infty \leq \max\{\log M, -\log m\} + 1$ is uniform bounded for n. Thus

$$\limsup_{n \to \infty} H(\rho_n) - H(\rho) \leq \limsup_{n \to \infty} \tau(h'(\rho_n)(\rho_n - \rho))$$

$$\leq \limsup_{n \to \infty} (\max\{\log M, -\log m\} + 1) \|\rho_n - \rho\|_1 = 0 \,,$$

which implies $\limsup_n H(\rho_n) \leq \limsup_n H(\rho_n) = H(\rho)$. For $D(\rho||\mathcal{N})$ we use the decomposition $D(\rho||\mathcal{N}) = H(\rho) - H(E(\rho))$. Note that $m1 = mE(1) \leq E(\rho) \leq ME(1) = M1$ and

$$\lim_n \|E(\rho_n) - E(\rho)\|_1 \leq \lim_n \|\rho_n - \rho\|_1 = 0 \,.$$

By the same argument, we obtain $H(E(\rho)) = \lim_n H(E(\rho_n))$. $\square$

Now we can show that λ-MLSI inequality for density in $S_B(\mathcal{A}_0)$ is equivalent to entropy decay property for all density in $S(\mathcal{M})$. Recall that $S_B(\mathcal{A}_0) = S_B(\mathcal{M}) \cap \mathcal{A}_0$ where $\mathcal{A}_0 = \bigcup_{t>0} T_t(\mathcal{A}) \subset \mathrm{dom}(A)$. Note that by the continuity of T_t on L_1 (see [16, Proposition 2.14]), the positive part $(\mathcal{A}_0)_+$ is norm dense in $\mathcal{A}_+$ hence by Lemma A.1 also dense in $L_1(\mathcal{M})_+$. Moreover, since $\mathcal{A}_0$ is a linear subspace containing unit, $S_B(\mathcal{A}_0)$ is norm dense in $S(\mathcal{M})$.

Proposition A.3. *A semigroup T_t satisfies λ-MLSI if and only if*

$$D(T_t(\rho)||\mathcal{N}) \leq e^{-2\lambda t} D(\rho||\mathcal{N}) \,, \; \forall \, \rho \in S(\mathcal{M}).$$

Proof. By the heuristic discussion and the equation (11), we know that our Definition 2.3 of λ-MLSI is equivalent to

$$D(\rho_t||\mathcal{N}) \leq e^{-2\lambda t} D(\rho||\mathcal{N}) \,, \forall \rho \in S_B(\mathcal{A}_0) \,.$$

To extend the exponential decay to all of $S(\mathcal{M})$, it suffices to show that for any $\rho \in S(\mathcal{M})$, there exists a sequence of $\rho_n \in S_B(\mathcal{A}_0)$ such that

$$\rho_n \to \rho \text{ in weakly } L_1 \ , \ \lim_{n \to \infty} D(\rho_n||\mathcal{N}) = D(\rho||\mathcal{N}) \ . \tag{27}$$

This is because by the lower semicontinuity of relative entropy (cf. [48, Corollary 5.12]) w.r.t. L_1-norm,

$$D(T_t(\rho)||\mathcal{N}) \le \liminf_n D(T_t(\rho_n)||\mathcal{N}) \le \liminf_n e^{-2\lambda t} D(\rho_n||\mathcal{N}) = e^{-2\lambda t} D(\rho||\mathcal{N}) \ ,$$

which implies the assertion. We verify the claim by two steps: (1) for any $\rho \in S(\mathcal{M})$, there exists a sequence $\rho_n \in S_B(\mathcal{M})$ satisfying (27); (2) for any $\rho \in S_B(\mathcal{M})$, there exists a sequence $\rho_n \in S_B(\mathcal{A}_0)$ satisfying (27). We first prove (2). By Lemma A.1, for $\rho \in S_B(\mathcal{M})$ with $\rho \le M1$, there exists a sequence $\rho_n \in S(\mathcal{A})$ such that $\rho_n \to \rho$ in L_1 and $\rho_n \le M1$. Since $T_{t_n}(\rho_n) \to \rho_n$ in L_1 as $t_n \to 0$, we can assume $\rho_n \in S(\mathcal{A}_0)$ by replacing ρ_n by $T_{t_n}(\rho_n)$ for some small t_n. For any $0 < \epsilon < 1$, we define

$$\rho_{n,\epsilon} = (1-\epsilon)\rho_n + \epsilon 1 \ , \ \rho_\epsilon = (1-\epsilon)\rho + \epsilon 1 \ .$$

Then for each ϵ, we have $\rho_{n,\epsilon} \to \rho_\epsilon$ in L_1 and by Lemma A.2, $\lim_n D(\rho_{n,\epsilon}||\mathcal{N}) = D(\rho_\epsilon||\mathcal{N})$ because $\epsilon 1 \le \rho_{n,\epsilon} \le M1$. Moreover, by convexity and lower semi-continuity

$$\limsup_{\epsilon \to 0} D(\rho_\epsilon||\mathcal{N}) \le \limsup_{\epsilon \to 0}(1-\epsilon)D(\rho||\mathcal{N}) = D(\rho||\mathcal{N}) \le \liminf_{\epsilon \to 0} D(\rho_\epsilon||\mathcal{N}) \ .$$

Thus $D(\rho||\mathcal{N}) = \lim_{\epsilon \to 0} D(\rho_\epsilon||\mathcal{N})$ and this proves (2). For (1), we denote e_n as the spectral projection of $E(\rho)$ for the spectrum $[1/n, n]$ and $e_n^\perp = 1 - e_n$. Without losing generosity, we assume ρ is faithful otherwise we restrict the discussion on its support. Note that $\|e_n^\perp\|_1 = \tau(e_n^\perp) \to 0$. For each n, we define CPTP map

$$P_n : L_1(\mathcal{M}) \to L_1(\mathcal{M}) \ , \ P_n(x) = e_n x e_n + \tau(x e_n^\perp)1$$

We have $P_n(L_1(\mathcal{N})) \subset L_1(\mathcal{N})$ and hence by data processing

$$D(P_n(\rho)||\mathcal{N}) \le D(\rho||\mathcal{N}) \ , \ \forall n \ . \tag{28}$$

On the other hand, $E(P_n(\rho)) = e_n E(\rho) e_n + \tau(E(\rho)e_n^\perp)1$ converges to $E(\rho)$ in L_1-norm and $P_n(\rho) \to \rho$ in weakly. Indeed, for any $y \in \mathcal{M}$,

$$\lim_n |\tau(\rho y) - \tau(e_n \rho e_n y)| \le \lim_n |\tau(e_n^\perp \rho y)| + |\tau(e_n \rho e_n^\perp y)|$$

$$\le \lim_n \|e_n^\perp\|_1 \|\rho\|_1 \|y\|_\infty + \|e_n^\perp\|_1 \|\rho\|_1 \|y\|_\infty = 0 \ .$$

Thus by the lower semicontinuity again

$$D(\rho||\mathcal{N}) = D(\rho||E(\rho)) \leq \liminf_n D(P_n(\rho)||E(P_n(\rho))) = D(P_n(\rho)||\mathcal{N}) \ .$$

Combined with (28), we have $\lim_n D(P_n(\rho)||\mathcal{N}) = D(\rho||\mathcal{N})$. That completes the proof. $\square$

References

[1] C. Arhancet, C. Kriegler, Riesz transforms, Hodge-Dirac operators and functional calculus for multipliers i, arXiv preprint, arXiv:1903.10151, 2019.

[2] D. Bakry, M. Émery, Diffusions hypercontractives, in: Séminaire de Probabilités XIX 1983/84, Springer, 1985, pp. 177–206.

[3] D. Bakry, I. Gentil, M. Ledoux, Analysis and Geometry of Markov Diffusion Operators, vol. 348, Springer Science & Business Media, 2013.

[4] I. Bardet, Estimating the decoherence time using non-commutative functional inequalities, arXiv preprint, arXiv:1710.01039, 2017.

[5] I. Bardet, C. Rouzé, Hypercontractivity and logarithmic Sobolev inequality for non-primitive quantum Markov semigroups and estimation of decoherence rates, arXiv preprint, arXiv:1803.05379, 2018.

[6] D.P. Blecher, V.I. Paulsen, Tensor products of operator spaces, J. Funct. Anal. 99 (2) (1991) 262–292.

[7] M. Brannan, L. Gao, M. Junge, Complete logarithmic Sobolev inequality via Ricci curvature bounded below ii, J. Topol. Anal. (2021) 1–54.

[8] P. Caputo, P. Dai Pra, G. Posta, Convex entropy decay via the Bochner-Bakry-emery approach, Ann. Inst. Henri Poincaré Probab. Stat. 45 (2009) 734–753.

[9] E.A. Carlen, J. Maas, An analog of the 2-Wasserstein metric in non-commutative probability under which the fermionic Fokker–Planck equation is gradient flow for the entropy, Commun. Math. Phys. 331 (3) (2014) 887–926.

[10] E.A. Carlen, J. Maas, Gradient flow and entropy inequalities for quantum Markov semigroups with detailed balance, J. Funct. Anal. 273 (5) (2017) 1810–1869.

[11] E.A. Carlen, J. Maas, Non-commutative calculus, optimal transport and functional inequalities in dissipative quantum systems, J. Stat. Phys. 178 (2) (2020) 319–378.

[12] M. Caspers, Gradient forms and strong solidity of free quantum groups, arXiv preprint, arXiv:1802.01968, 2018.

[13] S.-N. Chow, W. Huang, Y. Li, H. Zhou, Fokker–Planck equations for a free energy functional or Markov process on a graph, Arch. Ration. Mech. Anal. 203 (3) (2012) 969–1008.

[14] F. Cipriani, J.-L. Sauvageot, Derivations as square roots of Dirichlet forms, J. Funct. Anal. 201 (1) (2003) 78–120.

[15] N. Datta, C. Rouzé, Relating relative entropy, optimal transport and Fisher information: a quantum hwi inequality, in: Annales Henri Poincaré, Springer, 2020, pp. 1–36.

[16] E.B. Davies, J.M. Lindsay, Non-commutative symmetric Markov semigroups, Math. Z. 210 (1) (1992) 379–411.

[17] B. De Pagter, H. Witvliet, F.A. Sukochev, Double operator integrals, J. Funct. Anal. 192 (1) (2002) 52–111.

[18] P. Diaconis, L. Saloff-Coste, et al., Logarithmic Sobolev inequalities for finite Markov chains, Ann. Appl. Probab. 6 (3) (1996) 695–750.

[19] E.G. Effros, R.Z. Jin, Operator Spaces, Clarendon Press, 2000.

[20] M. Erbar, M.F. Poincaré, Modified logarithmic Sobolev and isoperimetric inequalities for Markov chains with non-negative Ricci curvature, J. Funct. Anal. 274 (11) (2018) 3056–3089.

[21] M. Erbar, J. Maas, Ricci curvature of finite Markov chains via convexity of the entropy, Arch. Ration. Mech. Anal. 206 (3) (2012) 997–1038.

[22] L. Gao, M. Junge, N. LaRacuente, Fisher Information and Logarithmic Sobolev Inequality for Matrix-Valued Functions, Annales Henri Poincaré, vol. 21, Springer, 2020, pp. 3409–3478.

[23] L. Gao, M. Junge, N. LaRacuente, Relative entropy for von Neumann subalgebras, Int. J. Math. 31 (06) (2020) 2050046.

[24] L. Gao, M. Junge, H. Li, Geometric approach towards complete logarithmic Sobolev inequalities, arXiv preprint, arXiv:2102.04434, 2021.

[25] L. Gao, C. Rouzé, Spectral methods for entropy contraction coefficients, arXiv preprint, arXiv: 2102.04146, 2021.
[26] A. Grigoryan, Heat Kernel and Analysis on Manifolds, vol. 47, American Mathematical Soc., 2009.
[27] M. Gromov, Structures métriques pour les variétés riemanniennes, in: J. Lafontaine, P. Pansu (Eds.), Cedic/Fernand Nathan, Paris, 1981.
[28] L. Gross, Hypercontractivity and logarithmic Sobolev inequalities for the Clifford-Dirichlet form, Duke Math. J. 42 (3) (1975) 383–396.
[29] L. Gross, Logarithmic Sobolev inequalities, Am. J. Math. 97 (4) (1975) 1061–1083.
[30] L. Gross, Hypercontractivity, logarithmic Sobolev inequalities, and applications: a survey of surveys, in: Diffusion, Quantum Theory, and Radically Elementary Mathematics, vol. 47, 2014, pp. 45–73.
[31] D.F. Hornshaw, l^2-Wasserstein distances of tracial w^*-algebras and their disintegration problem, arXiv preprint, arXiv:1806.01073, 2018.
[32] A. Jenčová, Rényi Relative Entropies and Noncommutative l_p-Spaces, Annales Henri Poincaré, vol. 19, Springer, 2018, pp. 2513–2542.
[33] M. Junge, N. LaRacuente, C. Rouzé, Stability of logarithmic Sobolev inequalities under a noncommutative change of measure, arXiv preprint, arXiv:1911.08533, 2019.
[34] M. Junge, T. Mei, J. Parcet, An invitation to harmonic analysis associated with semigroups of operators, in: Harmonic Analysis and Partial Differential Equations, vol. 612, 2013, pp. 107–122.
[35] M. Junge, J. Parcet, Mixed-Norm Inequalities and Operator Space L_p Embedding Theory, American Mathematical Soc., 2010.
[36] M. Junge, E. Ricard, D. Shlyahktenko, Noncommutative Diffusion Semigroups and Free Probability, In preparation.
[37] M.J. Kastoryano, K. Temme, Quantum logarithmic Sobolev inequalities and rapid mixing, J. Math. Phys. 54 (5) (2013) 052202.
[38] M. Ledoux, Spectral gap, logarithmic Sobolev constant, and geometric bounds, Surv. Differ. Geom. 9 (1) (2004) 219–240.
[39] H. Li, M. Junge, N. LaRacuente, Graph Hörmander systems, arXiv preprint, arXiv:2006.14578, 2020.
[40] P. Li, S.T. Yau, et al., On the parabolic kernel of the Schrödinger operator, Acta Math. 156 (1986) 153–201.
[41] E.H. Lieb, Convex trace functions and the Wigner-Yanase-Dyson conjecture, Adv. Math. 11 (3) (1973) 267–288.
[42] J. Lott, C. Villani, Ricci curvature for metric-measure spaces via optimal transport, Ann. Math. (2009) 903–991.
[43] J. Maas, Gradient flows of the entropy for finite Markov chains, J. Funct. Anal. 261 (8) (2011) 2250–2292.
[44] A. Mielke, A gradient structure for reaction–diffusion systems and for energy-drift-diffusion systems, Nonlinearity 24 (4) (2011) 1329.
[45] A. Mielke, Dissipative quantum mechanics using generic, in: Recent Trends in Dynamical Systems, Springer, 2013, pp. 555–585.
[46] J. Milnor, Curvatures of left invariant metrics on Lie groups, 1976.
[47] M. Mittnenzweig, A. Mielke, An entropic gradient structure for lindblad equations and couplings of quantum systems to macroscopic models, J. Stat. Phys. 167 (2) (2017) 205–233.
[48] M. Ohya, D. Petz, Quantum Entropy and Its Use, Springer Science & Business Media, 2004.
[49] F. Otto, C. Villani, Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality, J. Funct. Anal. 173 (2) (2000) 361–400.
[50] M. Pimsner, S. Popa, Entropy and index for subfactors, Ann. Sci. Éc. Norm. Supér. 19 (1986) 57–106.
[51] G. Pisier, Introduction to Operator Space Theory, vol. 294, Cambridge University Press, 2003.
[52] C. Rouzé, N. Datta, Concentration of quantum states from quantum functional and transportation cost inequalities, J. Math. Phys. 60 (1) (2019) 012202.
[53] L. Saloff-Coste, Aspects of Sobolev-Type Inequalities, vol. 289, Cambridge University Press, 2002.
[54] I.J. Schoenberg, Metric spaces and positive definite functions, Trans. Am. Math. Soc. 44 (3) (1938) 522–536.
[55] K.-T. Sturm, et al., On the geometry of metric measure spaces, Acta Math. 196 (1) (2006) 65–131.
[56] M. Takesaki, Theory of operator algebra i, 1979.
[57] N.T. Varopoulos, L. Saloff-Coste, T. Coulhon, Analysis and Geometry on Groups, vol. 100, Cambridge University Press, 2008.
[58] C. Villani, Optimal Transport: Old and New, vol. 338, Springer Science & Business Media, 2008.

[59] F.B. Weissler, Logarithmic Sobolev inequalities and hypercontractive estimates on the circle, J. Funct. Anal. 37 (2) (1980) 218–234.
[60] M. Wirth, A noncommutative transport metric and symmetric quantum Markov semigroups as gradient flows of the entropy, arXiv preprint, arXiv:1808.05419, 2018.
[61] M. Wirth, H. Zhang, Complete gradient estimates of quantum Markov semigroups, Commun. Math. Phys. (2021) 1–31.