Ann. Henri Poincaré 21 (2020), 2115–2150

1424-0637/20/072115-36
published online February 5, 2020
https://doi.org/10.1007/s00023-020-00891-8

Annales Henri Poincaré

Relating Relative Entropy, Optimal Transport and Fisher Information: A Quantum HWI Inequality

Nilanjana Datta and Cambyse Rouzé

Abstract. Quantum Markov semigroups characterize the time evolution of an important class of open quantum systems. Studying convergence properties of such a semigroup and determining concentration properties of its invariant state have been the focus of much research. Quantum versions of functional inequalities (like the modified logarithmic Sobolev and Poincaré inequalities) and the so-called transportation cost inequalities have proved to be essential for this purpose. Classical functional and transportation cost inequalities are seen to arise from a single geometric inequality, called the Ricci lower bound, via an inequality which interpolates between them. The latter is called the HWI inequality, where the letters I, W and H are, respectively, acronyms for the Fisher information (arising in the modified logarithmic Sobolev inequality), the so-called Wasserstein distance (arising in the transportation cost inequality) and the relative entropy (or Boltzmann H function) arising in both. Hence, classically, the above inequalities and the implications between them form a remarkable picture which relates elements from diverse mathematical fields, such as Riemannian geometry, information theory, optimal transport theory, Markov processes, concentration of measure and convexity theory. Here, we consider a quantum version of the Ricci lower bound introduced by Carlen and Maas and prove that it implies a quantum HWI inequality from which the quantum functional and transportation cost inequalities follow. Our results hence establish that the unifying picture of the classical setting carries over to the quantum one.

1. Introduction

Realistic physical systems that are relevant for quantum information processing are inherently open. They undergo unwanted but unavoidable interactions

with the surrounding environment and are hence subject to noise and decoherence. Under the Markovian approximation, which is valid when the system is only weakly coupled to its environment, the resulting dissipative dynamics of the system is described by a quantum Markov semigroup (QMS), whose generator we denote by $\mathcal{L}$. The analysis of quantum Markov semigroups is hence a key component of the theory of open quantum systems and quantum information. An important problem in the study of a QMS is the analysis of its convergence properties, in particular its *mixing time*, which is the time taken by any state evolving under the action of the QMS to come close to its invariant state.[1]

1.1. Functional and Transportation Cost Inequalities

Classically, given a measure μ, functional inequalities, e.g. the *Poincaré inequality* (usually denoted as PI(λ)) [29] and the *(modified) logarithmic Sobolev inequality* (or log-Sobolev in short), denoted as MLSI(α) [22], constitute a powerful tool for deriving mixing times of a Markov semigroup with invariant measure μ and determining concentration properties of μ. They are also related to the so-called *transportation cost* inequalities denoted by $\mathrm{TC}_1(c_1)$ and $\mathrm{TC}_2(c_2)$. Here α, c_1 and c_2 denote constants appearing in the respective inequalities. Consider a compact manifold $\mathcal{M}$, and let $\mathcal{P}(\mathcal{M})$ be the set of probability measures on $\mathcal{M}$. Given a measure $\mu \in \mathcal{P}(\mathcal{M})$, the inequality $\mathrm{TC}_1(c_1)$ (resp. $\mathrm{TC}_2(c_2)$) provides an upper bound on the so-called Wasserstein distance W_1 (resp. W_2), between any probability measure $\nu \in \mathcal{P}(\mathcal{M})$ and the given measure μ, in terms of the square root of the relative entropy of ν with respect to μ. Since μ is fixed, this relative entropy is simply a functional of ν and, due to its close links with the Boltzmann H-functional, is often denoted by the letter H in the literature. The notion of Wasserstein distances first appeared in the theory of optimal transport, which was initiated by Monge [21] and later analysed by Kantorovich [15]. In its original formulation by Monge, the problem of optimal transport concerns finding the optimal way, in the sense of minimal transportation cost, of moving a sand pile between two locations (see also [30]). In 1986, Marton [20] showed that transportation cost inequalities are also useful for deriving *concentration of measure* properties of the given measure μ.[2]

The classical inequalities discussed above can be shown to be obtainable from a single geometric inequality, involving a quantity called the *Ricci curvature* of the manifold $\mathcal{M}$, and referred to as the *Ricci lower bound*. In fact, there is an inherent relation between the geometry of the manifold and a diffusion process (whose associated Markov semigroup has generator $\mathcal{L}$, say) defined on

[1] Here we assume that the QMS is primitive, i.e. it has a unique invariant state.

[2] Given a metric space $(\mathcal{X}, d)$, a probability measure μ is said to satisfy Gaussian (resp. exponential) concentration on it if there exist positive constants a, b such that for any $A \subseteq \mathcal{X}$, and $r > 0$,

$$\mu(A) \geq 1/2 \implies \mu(A_r) \geq 1 - ae^{-bf(r)}.$$

where $A_r := \{x \in \mathcal{X} : d(x, A) < r\}$ and $f(r) = r^2$ (resp. r).

it: the diffusion process can be used to explore the geometry of $\mathcal{M}$, and conversely, the latter determines the mixing time of the diffusion process. Finding a quantum analogue of this appealing geometric inequality is hence a problem of fundamental interest and is considered in this paper. Before we present our results on this problem, we first need to explain the statement of the Ricci lower bound in the classical setting. In fact, it is instructive to start from the very definition of curvature which generalizes to the Ricci curvature for the case of a Riemannian manifold.

1.2. Ricci Curvature and Ricci Lower Bound (Classical Setting)

Given a surface $\mathcal{S}$ embedded in the Euclidean space $\mathbb{R}^3$, the *Gauss curvature* κ of $\mathcal{S}$ is a measure of its local boundedness. More precisely, given a point $x \in \mathcal{S}$ and any two mutually orthogonal unit tangent vectors u, v at x, the distance between two geodesics γ_u and γ_v, starting at x, with respective directions u and v, obeys the following Taylor expansions:

$$d_{\mathrm{g}}(\gamma_{\mathbf{u}}(t), \gamma_{\mathbf{v}}(t)) = \sqrt{2}t \left(1 - \frac{\kappa(x)}{12} t^2 + \mathcal{O}_{t\to 0}(t^3)\right), \quad t \geq 0, \tag{1.1}$$

where d_{g} is the geodesic distance defined with respect to the metric g induced on $\mathcal{S}$ by the Euclidean metric. In the case when $\kappa = 0$ uniformly on the surface, the latter is flat, and we recover the Pythagoras theorem from Eq. (1.1). More generally, let x be a point on a d-dimensional compact Riemannian manifold $\mathcal{M}$, let u belong to the tangent space $T_x\mathcal{M}$ at the point x of $\mathcal{M}$, and complete the vector u into an orthonormal basis $(u, v_2, \ldots, v_d)$ of $T_x\mathcal{M}$. Then, the *Ricci curvature* of $\mathcal{M}$, evaluated at u, is the averaged Gauss curvature over orthogonal surfaces defined by all the geodesics starting at x with direction given by the unit vectors belonging to the vector subspace spanned by u and any other vector v_i, $i = 2, \ldots, d$. The expression for the Ricci curvature [30] is given in terms of the Laplace–Beltrami operator (denoted simply as Δ), and hence, the curvature is usually denoted as $\mathrm{Ric}(\Delta)$. Since Δ is the generator of the heat semigroup, the curvature provides a bridge between the geometry of the manifold and the evolution on it induced by heat diffusion. There is an important inequality, known as the *Ricci lower bound*, which is denoted by $\mathrm{Ric}(\Delta) \geq \kappa$ [3], and is the property that the Ricci curvature is uniformly bounded below by a real parameter $\kappa \geq 0$. Intuitively, the inequality is related to concentration of the uniform measure on $\mathcal{M}$, which is known to be the unique invariant measure of heat diffusion. For example, in the case of the sphere, which has constant Ricci curvature given in terms of its radius, the Haar measure can be shown to concentrate around any great circle. One can relax the condition of uniformity of the measure in order to allow for the study of concentration of measure phenomena for different measures μ, invariant for other diffusions processes on $\mathcal{M}$. In this more general framework, the Ricci lower bound is denoted by $\mathrm{Ric}(\mathcal{L}) \geq \kappa$, where $\mathcal{L}$ denotes the generator of the diffusion semigroup associated with μ (Fig. 1).

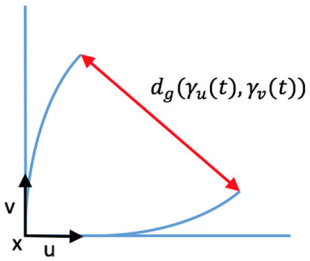


FIGURE 1. The Gauss curvature

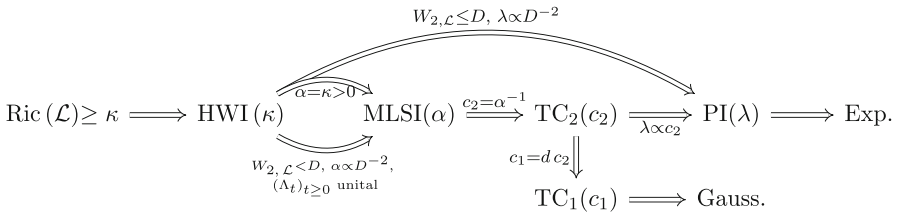
More recently, Sturm [27,28] and Lott–Villani [18] showed that $\mathrm{Ric}(\mathcal{L}) \geq \kappa$ can be viewed as a (refined) convexity property (called the κ-*displacement convexity*) of H along geodesics on the Riemannian manifold obtained by endowing the set of probability measures $\mathcal{P}(\mathcal{M})$ on $\mathcal{M}$ with the Wasserstein distance W_2 [31]. This discovery led to a more robust notion of a Ricci lower bound, which does not explicitly depend on the expression of the Ricci curvature, and hence can be extended to more general metric spaces. Starting from this convexity property, one can then construct a diffusion semigroup for which H decreases the most along the direction of evolution induced by the semigroup. In this case, the path on the Riemannian manifold $(\mathcal{P}(\mathcal{M}), W_2)$, which corresponds to the actual evolution under the diffusion, is said to be *gradient flow* for H. It is a striking fact that this diffusion coincides with the one whose generator appears in the Bakry–Émery condition (see [10,14]).

In [23], the authors introduced the so-called HWI(κ)-*interpolation inequality*, using which they reproved the so-called Bakry–Émery theorem, which states that for $\kappa > 0$, $\mathrm{Ric}(\mathcal{L}) \geq \kappa$ implies MLSI(α) (for diffusions on $\mathbb{R}^n$ with associated generator $\mathcal{L}$). The letters W, I and H are, respectively, acronyms for the Wasserstein distance W_2 (appearing in $\mathrm{TC}_2(c_2)$), the Fisher information (which arises in MLSI(α)) and the relative entropy (also called the Boltzmann H-functional, as mentioned above) which appears in both these inequalities. They also showed that MLSI(α) implies $\mathrm{TC}_2(c_2)$. The term interpolation here comes from the fact that in the case $\kappa = 0$ and $c > 0$, $\mathrm{TC}_2(c)$ together with HWI(0) gives back MLSI(α).

In [11,12,19], a modified version of the Ricci lower bound was defined for Markov processes on finite sets, which led to the unification of the previously discussed functional and concentration inequalities in this discrete framework. In particular, it was proved in [11] that one can recover the Poincaré and modified log-Sobolev inequalities from the Ricci lower bound, provided the diameter of $\mathcal{P}(\mathcal{M})$, with respect to the Wasserstein distance, W_2, is bounded.

1.3. Ricci Lower Bound (Quantum Setting)

In the case of a quantum system with a finite-dimensional Hilbert space $\mathcal{H}$, the set $\mathcal{P}(\mathcal{M})$ is replaced by the set $\mathcal{D}(\mathcal{H})$ of quantum states (i.e. density matrices) on $\mathcal{H}$. Then, in analogy with the classical case, starting with a primitive QMS



$$\mathrm{TC}_2(c_2) + \mathrm{HWI}(\kappa) \Rightarrow \mathrm{MLSI}(\alpha), \quad \alpha = \max\left[\frac{1}{4c_2}(1 + c_2\kappa)^2, \kappa\right].$$

FIGURE 2. Chain of quantum functional- and Talagrand inequalities and related concentrations for a primitive semigroup $(\Lambda_t)_{t\geq 0}$ with generator $\mathcal{L}$ defined on a Hilbert space of dimension d. The implication $\mathrm{MLSI}(\alpha) \Rightarrow \mathrm{PI}(\lambda)$ was proved in [16]. Here, "Exp." refers to the notion of exponential concentration, whereas "Gauss." refers to the stronger notion of Gaussian concentration. The implications $\mathrm{MLSI}(\alpha) \Rightarrow \mathrm{TC}_2(c_2) \Rightarrow \mathrm{PI}(\lambda) \Rightarrow$ Exp., as well as $\mathrm{TC}_2(c_2) \Rightarrow \mathrm{TC}_1(c_1) \Rightarrow$ Gauss. were proved in [24]

with generator $\mathcal{L}$, Carlen and Maas [6,7] defined a quantum Wasserstein distance $W_{2,\mathcal{L}}$ which renders $\mathcal{D}(\mathcal{H})$ with a Riemannian structure, and for which the master equation associated with the QMS is gradient flow for the quantum relative entropy.

In [7], the authors proved that a quantum $\mathrm{MLSI}(\alpha)$, first introduced in [16], holds provided the quantum relative entropy (between a state on a geodesic on this manifold and the invariant state of the QMS) satisfies a quantum analogue of the κ-displacement convexity property along geodesics, for $\alpha = \kappa > 0$. This is denoted below by $\mathrm{Ric}(\mathcal{L}) \geq \kappa$ in analogy with the classical case, with $\mathcal{L}$ being the generator of the QMS.

The quantum versions of the Ricci lower bound, the HWI inequality, and the functional and transportation cost inequalities, all fit into a unifying picture which is analogous to the classical setting. It is given in Fig. 2.

Our Contribution: In this paper, we analyse the quantum version of the Ricci lower bound introduced by Carlen and Maas [7] and derive various implications of it in Theorem 3. Moreover, we show that $\mathrm{Ric}(\mathcal{L}) \geq \kappa$ implies a quantum version of the celebrated $\mathrm{HWI}(\kappa)$ inequality which interpolates between the modified logarithmic Sobolev inequality and the transportation cost inequality (Theorem 5). We show that, in the case of $\kappa > 0$, $\mathrm{HWI}(\kappa) \Rightarrow \mathrm{MLSI}(\kappa)$ (Corollary 2), recovering the result of [7]. On the other hand, in Corollary 3, we establish that in the case when $\kappa \in \mathbb{R}$, $\mathrm{Ric}(\mathcal{L} \geq \kappa)$ together with $\mathrm{TC}_2(c_2)$ implies $\mathrm{MLSI}(\alpha)$. Moreover, in the case when $\kappa = 0$, we show that, under the assumption of boundedness of the diameter D of the set of states with respect to the quantum Wasserstein distance $W_{2,\mathcal{L}}$, $\mathrm{Ric}(\mathcal{L}) \geq 0$ implies $\mathrm{PI}(c_1 D^{-2})$ for some universal constant c_1 (Theorem 6). Moreover, in the case of a unital QMS

(i.e. one which has the completely mixed state as its unique invariant state), we show that it also implies MLSI($c_2 D^{-2}$) for some universal positive constant c_2 (Theorem 7). We hence extend the results of [11] to the quantum regime.

1.4. Layout of the Paper

In Sect. 2, we introduce the necessary notations and definitions, including quantum Markov semigroups, the quantum Wasserstein distance and quantum functional inequalities. The quantum version of κ-displacement convexity is studied in Sect. 3. In Sect. 4, we prove the quantum HWI(κ) inequality, show that it implies MLSI(κ) in the case when $\kappa > 0$ and derive interpolation results between MLSI(κ) and TC$_2$(c_2) from it. In Sect. 5, we show that in the case in which $\kappa = 0$, PI(λ) holds with a constant λ proportional to D^{-2}, where D stands for the diameter of the set of states. In Sect. 6, we show that under the further assumption of the QMS being unital, MLSI(α_1) holds with a constant α_1 also proportional to D^{-2}.

2. Notations and Preliminaries

2.1. Operators, States and Entropic Quantities

In this paper, we denote by $(\mathcal{H}, \langle .|. \rangle)$ a finite-dimensional Hilbert space of dimension d with associated inner product $\langle .|. \rangle$, by $\mathcal{B}(\mathcal{H})$ the algebra of linear operators acting on $\mathcal{H}$, and by $\mathcal{B}_{sa}(\mathcal{H}) \subset \mathcal{B}(\mathcal{H})$ the subspace of self-adjoint operators. Moreover, the Hilbert–Schmidt inner product $\langle ., . \rangle$, where $\langle A, B \rangle = \mathrm{Tr}(A^* B)\ \forall A, B \in \mathcal{B}(\mathcal{H})$, provides $\mathcal{B}(\mathcal{H})$ with a Hilbert space structure. Here, the trace Tr is unnormalized, adopting the uses of the community of quantum information theory, so that $\mathrm{Tr}(\mathbb{I}) = d$. Let $\mathcal{P}(\mathcal{H})$ be the cone of positive semi-definite operators on $\mathcal{H}$ and $\mathcal{P}_+(\mathcal{H}) \subset \mathcal{P}(\mathcal{H})$ the set of (strictly) positive operators. Further, let $\mathcal{D}(\mathcal{H}) := \{\rho \in \mathcal{P}(\mathcal{H}) \mid \mathrm{Tr}\,\rho = 1\}$ denote the set of density operators (or states) on $\mathcal{H}$, and $\mathcal{D}_+(\mathcal{H}) := \mathcal{D}(\mathcal{H}) \cap \mathcal{P}_+(\mathcal{H})$ denote the subset of faithful states. We denote the support of an operator A by $\mathrm{supp}(A)$. Let $\mathbb{I} \in \mathcal{P}(\mathcal{H})$ be the identity operator on $\mathcal{H}$, and $\mathrm{id} : \mathcal{B}(\mathcal{H}) \mapsto \mathcal{B}(\mathcal{H})$ the identity map on operators on $\mathcal{H}$. For $p, q \geq 1$, the p-Schatten norm of an operator $A \in \mathcal{B}(\mathcal{H})$ is denoted by $\|A\|_p := (\mathrm{Tr}\,|A|^p)^{1/p}$, and the $p \to q$-norm of a superoperator $\Lambda : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ by $\|\Lambda\|_{p\to q}$. Such a linear map is said to be unital if $\Lambda(\mathbb{I}) = \mathbb{I}$. Given two states $\rho, \sigma \in \mathcal{D}(\mathcal{H})$, the quantum relative entropy between ρ and σ is defined as:

$$D(\rho\|\sigma) := \begin{cases} \mathrm{Tr}(\rho(\log\rho - \log\sigma)) & \text{if } \mathrm{supp}(\rho) \subseteq \mathrm{supp}(\sigma), \\ +\infty & \text{else.} \end{cases}$$

2.2. Quantum Markov Semigroups and the Detailed Balance Condition

In the Heisenberg picture, a quantum Markov semigroup (QMS) on a finite-dimensional Hilbert space $\mathcal{H}$ is given by a one-parameter family $(\Lambda_t)_{t\geq 0}$ of linear, completely positive, unital maps on $\mathcal{B}(\mathcal{H})$ satisfying the following properties

- $\Lambda_0 = \mathrm{id}$;

- $\Lambda_t \circ \Lambda_s = \Lambda_{t+s}$—semigroup property;
- $\forall X \in \mathcal{B}(\mathcal{H}), \lim_{t\to 0} ||\Lambda_t(X) - X||_\infty = 0$—strong continuity.

The parameter t plays the role of time. For each quantum Markov semigroup, there exists an operator $\mathcal{L}$ called the *generator*, or *Lindbladian*, of the semigroup, such that

$$\frac{\mathrm{d}}{\mathrm{d}t}\Lambda_t = \Lambda_t \circ \mathcal{L} = \mathcal{L} \circ \Lambda_t. \tag{2.1}$$

In the Schrödinger picture, the dual of Λ_t is written Λ_{*t}, for any $t \geq 0$. Similarly, we denote by $\mathcal{L}_*$ the dual of $\mathcal{L}$. The QMS is said to be *primitive* (or *ergodic*) if there exists a unique invariant state σ, i.e. such that $\Lambda_{*t}(\sigma) = \sigma$. Such a QMS is said to satisfy the *detailed balance condition* if the following holds:

$$\mathrm{Tr}(\sigma \mathcal{L}(X)^* Y) = \mathrm{Tr}(\sigma X^* \mathcal{L}(Y)), \quad X, Y \in \mathcal{B}(\mathcal{H}). \tag{2.2}$$

In the context of quantum logarithmic Sobolev inequalities (introduced later), the *quantum Fisher information* of ρ with respect to the state σ, first defined in [25], is particularly useful:

$$\mathrm{I}_\sigma(\rho) := \begin{cases} -\mathrm{Tr}(\mathcal{L}_*(\rho)(\log\rho - \log\sigma)), & \rho \in \mathcal{D}_+(\mathcal{H}) \\ +\infty, & \text{otherwise.} \end{cases}$$

This quantity is also referred to as *entropy production* and denoted by EP_σ in the literature. We will use both notations in what follows. The following theorem provides a structure for the generators of primitive QMS satisfying the detailed balance condition:

Theorem 1 ([1,7]). *Let $\sigma \in \mathcal{D}_+(\mathcal{H})$, and let $(\Lambda_t)_{t\geq 0}$ be a quantum Markov semigroup on $\mathcal{B}(\mathcal{H})$. Suppose that the generator $\mathcal{L}$ of $(\Lambda_t)_{t\geq 0}$ satisfies the detailed balance condition with respect to a full-rank invariant state σ. Then there exists an index set $\mathcal{J}$ of cardinality $|\mathcal{J}| \leq d^2 - 1$, where $d = \dim(\mathcal{H})$, such that $\mathcal{L}$ takes the following form for any $f \in \mathcal{B}(\mathcal{H})$:*

$$\mathcal{L}(f) = \sum_{j\in\mathcal{J}} c_j \left(\mathrm{e}^{-\omega_j/2} \tilde{L}_j^* [f, \tilde{L}_j] + \mathrm{e}^{\omega_j/2} [\tilde{L}_j, f] \tilde{L}_j^* \right) \tag{2.3}$$

where $\omega_j \in \mathbb{R}$ and $c_j > 0$ for all $j \in \mathcal{J}$, and $\{\tilde{L}_j\}_{j\in\mathcal{J}}$ is a set of operators in $\mathcal{B}(\mathcal{H})$ with the properties:

1. $\frac{1}{\dim(\mathcal{H})}\mathrm{Tr}(\tilde{L}_j^* \tilde{L}_k) = \delta_{k,j}$ *for all* $j, k \in \mathcal{J}$
2. $\mathrm{Tr}(\tilde{L}_j) = 0$ *for all* $j \in \mathcal{J}$
3. $\{\tilde{L}_j\}_{j\in\mathcal{J}} = \{\tilde{L}_j^*\}_{j\in\mathcal{J}}$
4. $\{\tilde{L}_j\}_{j\in\mathcal{J}}$ *consists of eigenvectors of the modular operator* $\Delta_\rho : f \mapsto \rho f \rho^{-1}$ *with*

$$\Delta_\sigma(\tilde{L}_j) = \mathrm{e}^{-\omega_j} \tilde{L}_j.$$

Finally for each $j \in \mathcal{J}$

$$c_j = c_{j'}, \text{ and } \omega_j = -\omega_{j'} \text{ when } \tilde{L}_j^* = \tilde{L}_{j'}. \tag{2.4}$$

Conversely, given any faithful (full-rank) state σ, any set $\{\tilde{L}_j\}_{j\in\mathcal{J}}$ satisfying the above four conditions for some $\{\omega_j\}_{j\in\mathcal{J}} \subset \mathbb{R}$ and any set $\{c_j\}_{j\in\mathcal{J}}$ of positive numbers satisfying the symmetry condition (2.4), the operator $\mathcal{L}$ given by Eq. (2.3) is the generator of a quantum Markov semigroup $(\Lambda_t)_{t\geq 0}$ which satisfies the detailed balance condition.

2.3. The Wasserstein Distance $W_{2,\mathcal{L}}$

In this section, we recall the construction of the Wasserstein metric $W_{2,\mathcal{L}}$ first defined in [7]. Assume given a generator $\mathcal{L}$ of a primitive QMS, with invariant state σ, of the form of (2.3). Given an operator $X \in \mathcal{B}(\mathcal{H})$, its non-commutative *gradient* is defined as:

$$\nabla X := (\partial_1 X, \ldots, \partial_{\mathcal{J}} X), \quad X \in \mathcal{B}(\mathcal{H}),$$

where $\partial_j X = [\tilde{L}_j, X]$ for all $j \in \mathcal{J}$. Similarly, given a vector $\mathbf{A} \equiv (A_1, \ldots, A_{|\mathcal{J}|}) \in \bigoplus_{j\in\mathcal{J}} \mathcal{B}(\mathcal{H})$, the *divergence* of $\mathbf{A}$ is defined as

$$\mathrm{div}(\mathbf{A}) := \sum_{j\in\mathcal{J}} c_j [A_j, \tilde{L}_j^*] \equiv -\sum_{j\in\mathcal{J}} c_j \partial_j^* A_j,$$

where $\partial_j^* X := [\tilde{L}_j^*, X]$. For $\vec{\omega} := (\omega_1, \ldots, \omega_{|\mathcal{J}|})$, define the linear operator $[\rho]_{\vec{\omega}}$ on $\bigoplus_{j\in\mathcal{J}} \mathcal{B}(\mathcal{H})$ through

$$[\rho]_{\vec{\omega}} \mathbf{A} := ([\rho]_{\omega_1} A_1, \ldots, [\rho]_{\omega_{|\mathcal{J}|}} A_{|\mathcal{J}|}), \ \mathbf{A} \equiv (A_1, \ldots, A_{|\mathcal{J}|}),$$

where for any $\omega \in \mathbb{R}$,

$$[\rho]_\omega := R_\rho \circ f_\omega(\Delta_\rho), \ f_\omega(t) := \mathrm{e}^{\omega/2} \frac{t - \mathrm{e}^{-\omega}}{\log t + \omega}, \ t \in \mathbb{R}, \tag{2.5}$$

where $R_\rho : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ denotes the operator of left multiplication by ρ. Intuitively, $[\rho]_\omega$ can be understood as a non-commutative way of multiplying by ρ:

Lemma 1 (see Lemma 5.8 of [7]). *For any $\omega \in \mathbb{R}$, and $\rho \in \mathcal{D}_+(\mathcal{H})$,*

$$[\rho]_\omega(A) = \int_0^1 \mathrm{e}^{\omega(1/2-s)} \rho^s A \rho^{1-s} \mathrm{d}s.$$

Let $(\gamma(s))_{s\in(-\varepsilon,\varepsilon)}$ be a differential path in $\mathcal{D}_+(\mathcal{H})$ for some $\varepsilon > 0$ and denote $\rho := \gamma(0)$. Then $\mathrm{Tr}(\dot{\gamma}(0)) = \frac{\mathrm{d}}{\mathrm{d}s}\big|_{s=0} \mathrm{Tr}(\gamma(s)) = 0$. Carlen and Maas proved that there is a unique vector field $\mathbf{V} \in \bigoplus_{j\in\mathcal{J}} \mathcal{B}(\mathcal{H})$ of the form $\mathbf{V} = \nabla U$, where $U \in \mathcal{B}(\mathcal{H})$ is traceless and self-adjoint, for which the following non-commutative continuity equation holds:

$$\dot{\gamma}(0) = -\mathrm{div}([\rho]_{\vec{\omega}} \nabla U). \tag{2.6}$$

Define the inner product $\langle ., . \rangle_{\mathcal{L},\rho}$ on $\bigoplus_{j\in\mathcal{J}} \mathcal{B}(\mathcal{H})$ through:

$$\langle \mathbf{W}, \mathbf{V} \rangle_{\mathcal{L},\rho} := \sum_{j\in\mathcal{J}} c_j \langle W_j, [\rho]_{\omega_j} V_j \rangle, \tag{2.7}$$

where $\langle A, B \rangle := \mathrm{Tr}(A^* B)$ denotes the usual Hilbert–Schmidt inner product on $\mathcal{B}(\mathcal{H})$. Hence, looking upon $\mathcal{D}_+(\mathcal{H})$ as a manifold, for each $\rho \in \mathcal{D}_+(\mathcal{H})$, we

can identify the tangent space T_ρ at ρ with the set of gradient vector fields $\{\nabla U : U \in \mathcal{B}(\mathcal{H}),\ U = U^*\}$ through the correspondence provided by the continuity equation (2.6). Defining the metric $g_{\mathcal{L}}$ through the relation

$$\|\dot\gamma(0)\|^2_{g_{\mathcal{L},\rho}} := \|\mathbf{V}(0)\|^2_{\mathcal{L},\rho}, \tag{2.8}$$

this endows the manifold $\mathcal{D}_+(\mathcal{H})$ with a smooth Riemannian structure. In this framework, Carlen and Maas then defined the *modified non-commutative Wasserstein distance* $W_{2,\mathcal{L}}$ to be the energy associated with the metric $g_{\mathcal{L}}$, i.e.:

$$W_{2,\mathcal{L}}(\rho,\sigma) := \inf_\gamma \left\{ \left(\int_0^1 \|\mathbf{V}(s)\|^2_{\mathcal{L},\gamma(s)} \mathrm{d}s \right)^{1/2} : \gamma(0) = \rho, \gamma(1) = \sigma \right\}, \tag{2.9}$$

where the infimum is taken over smooth paths $\gamma : [0,1] \to \mathcal{D}_+(\mathcal{H})$, and $\mathbf{V} : [0,1] \to \bigoplus_{i\in\mathcal{J}} \mathcal{B}(\mathcal{H})$ is related to γ through the continuity equation (2.6). The paths achieving the infimum, if they exist, are the minimizing geodesics with respect to the metric $g_{\mathcal{L}}$. The following lemma, proved in [24], follows from a standard argument:

Lemma 2. *With the above notations, the Wasserstein distance between two faithful states ρ, σ is equal to the minimal length over the smooth paths joining ρ and σ:*

$$W_{2,\mathcal{L}}(\rho,\sigma) = \inf_{\gamma(s)\ \mathit{const.\ speed}} \left\{ \int_0^1 \|\dot\gamma(s)\|_{g_{\mathcal{L},\gamma(s)}} \mathrm{d}s :\ \gamma(0) = \rho,\ \gamma(1) = \sigma \right\}, \tag{2.10}$$

where the infimum is taken over curves γ of constant speed, i.e. such that $s \mapsto \|\dot\gamma(s)\|_{g_{\mathcal{L},\gamma(s)}}$ is constant on $[0,1]$.

This definition for the quantum Wasserstein distance, $W_{2,\mathcal{L}}$, is natural in the sense that the master equation

$$\dot\rho_t = \mathcal{L}_* \rho_t,$$

is *gradient flow* for $D(.\|\sigma)$, where σ is the invariant state associated with $\mathcal{L}$. This means that $\mathcal{L}_*\rho = -\operatorname{grad}_{\mathcal{L}} D(\rho\|\sigma)$, where the gradient $\operatorname{grad}_{\mathcal{L}}$ of a differentiable functional $\mathcal{F} : \mathcal{D}_+(\mathcal{H}) \to \mathbb{R}$ is defined as the unique element in the tangent space at ρ so that

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{F}(\gamma(t))\bigg|_{t=0} = g_{\mathcal{L},\rho}(\dot\gamma, \operatorname{grad}_{g_{\mathcal{L},\rho}} \mathcal{F}(\rho)) \tag{2.11}$$

for all smooth paths $\gamma(t)$ defined on $(-\varepsilon,\varepsilon)$ for some $\varepsilon > 0$ with $\gamma(0) = \rho$. In particular, for $\gamma(t) = \rho_t \equiv \Lambda_{*t}(\rho)$,

$$\frac{\mathrm{d}}{\mathrm{d}t} D(\rho_t\|\sigma)\bigg|_{t=0} = -g_{\mathcal{L},\rho}(\mathcal{L}_*(\rho), \mathcal{L}_*(\rho)) = -\|\mathcal{L}_*(\rho)\|^2_{g_{\mathcal{L},\rho}}. \tag{2.12}$$

The following lemma is going to play a crucial role in the rest of this paper:

Lemma 3. *For any $\rho \in \mathcal{D}_+(\mathcal{H})$, the map $D_{\vec{\omega}}(\rho) : U \mapsto -\operatorname{div}([\rho]_{\vec{\omega}}\nabla U)$ is invertible and positive in the sense of Loewner order on the Hilbert space of self-adjoint, traceless operators. Moreover, if $\rho \geq \varepsilon \mathbb{I}$ for some $\varepsilon > 0$, then:*

$$D_{\vec{\omega}}(\rho)^{-1} \leq K_{\mathcal{L}}\ \varepsilon^{-1}\ \mathrm{id},$$

where $K_{\mathcal{L}} := \sup_{j\in\mathcal{J}} \frac{\omega_j}{\mathrm{e}^{\omega_j/2}-\mathrm{e}^{-\omega_j/2}} \|(-\operatorname{div}\circ\nabla(.))^{-1}\|_{2\to 2} > 0$.

Proof. Let $\mathcal{W}$ be the space of self-adjoint, traceless operators on $\mathcal{H}$. From Theorem 7.3 of [7], for any C^1 path $(\gamma(t))_{t\in(-\varepsilon,\varepsilon)}$, with $\gamma(0) = \rho$, there exists a unique vector field of the form ∇U for which the continuity equation $\dot{\gamma}(0) = -\operatorname{div}([\rho]_{\vec{\omega}}(\nabla U))$ holds. Moreover, by ergodicity of $(\Lambda_t)_{t\geq 0}$, $\ker(\nabla)$ consists of multiples of the identity. Therefore, there exists a unique $U \in \mathcal{W}$ such that $\dot{\gamma}(0) = -\operatorname{div}([\rho]_{\vec{\omega}}(\nabla U))$. Now, for any $U, V \in \mathcal{W}$:

$$\begin{aligned}
\langle U, D_{\vec{\omega}}(\rho)[V]\rangle &= \sum_{j\in\mathcal{J}} c_j \langle \partial_j U, [\rho]_{\omega_j}\partial_j V\rangle \\
&= \sum_{j\in\mathcal{J}} c_j \langle [\rho]_{\omega_j}\partial_j U, \partial_j V\rangle \\
&= \sum_{j\in\mathcal{J}} c_j \langle \partial_j^*([\rho]_{\omega_j}\partial_j U), V\rangle \\
&= -\langle \operatorname{div}([\rho]_{\vec{\omega}}\nabla U), V\rangle \\
&= \langle D_{\vec{\omega}}(\rho)[U], V\rangle,
\end{aligned}$$

which means that $D_{\vec{\omega}}(\rho)$ is indeed self-adjoint. By the same argument, we can show that the superoperators $\nabla : \mathcal{B}(\mathcal{H}) \to \oplus_{j\in\mathcal{J}}\mathcal{B}(\mathcal{H})$ and $-\operatorname{div} : \oplus_{j\in\mathcal{J}}\mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ are adjoint to each other, where $\mathcal{B}(\mathcal{H})$ and $\oplus_{j\in\mathcal{J}}\mathcal{B}(\mathcal{H})$ are provided with the inner products $\langle .,.\rangle$ and $\sum_{j\in\mathcal{J}} c_j \langle .,.\rangle$, respectively. Indeed for any $U \in \mathcal{B}(\mathcal{H})$ and $\mathbf{V} = (V_1, \ldots, V_{|\mathcal{J}|}) \in \oplus_{j\in\mathcal{J}}\mathcal{B}(\mathcal{H})$:

$$\begin{aligned}
\sum_{j\in\mathcal{J}} c_j \langle V_j, \partial_j U\rangle &= \sum_{j\in\mathcal{J}} c_j \operatorname{Tr}((V_j)^*[\tilde{L}_j, U]) \\
&= \sum_{j\in\mathcal{J}} c_j \operatorname{Tr}([V_j^*, \tilde{L}_j]U) \\
&= -\sum_{j\in\mathcal{J}} c_j \operatorname{Tr}([V_j, \tilde{L}_j^*]^* U) \\
&= \langle -\operatorname{div}(\mathbf{V}), U\rangle.
\end{aligned}$$

Assume now that $\rho \geq \varepsilon\mathbb{I}$ for some $\varepsilon > 0$, so that for any $j \in \mathcal{J}$:

$$[\rho]_{\omega_j} = \int_0^1 \mathrm{e}^{\omega_j(1/2-\alpha)} L_\rho^\alpha R_\rho^{1-\alpha} \mathrm{d}\alpha \geq \varepsilon \frac{\mathrm{e}^{\omega_j/2} - \mathrm{e}^{-\omega_j/2}}{\omega_j}\mathrm{id} > 0.$$

Hence, $-\operatorname{div}\circ[\rho]_{\vec{\omega}}\circ\nabla(.)$ is positive and

$$(-\operatorname{div}\circ[\rho]_{\vec{\omega}}\circ\nabla(.))^{-1} \leq \varepsilon^{-1} \sup_{j\in\mathcal{J}} \frac{\omega_j}{\mathrm{e}^{\omega_j/2} - \mathrm{e}^{-\omega_j/2}} (-\operatorname{div}\circ\nabla(.))^{-1},$$

and the result follows. □

The above lemma allows us to extend the definition of the Wasserstein distance to non-faithful states:

Proposition 1 (Extension of the metric to $\mathcal{D}(\mathcal{H})$). *Let $\rho, \omega \in \mathcal{D}(\mathcal{H})$ and let $\{\rho_n\}_{n\in\mathbb{N}}$ and $\{\omega_n\}_{n\in\mathbb{N}}$ be sequences of faithful states satisfying*

$$\mathrm{Tr}[(\rho-\rho_n)^2] \to 0, \ \ \mathrm{Tr}[(\omega-\omega_n)^2] \to 0, \tag{2.13}$$

as $n \to \infty$. Then the sequence $\{W_{2,\mathcal{L}}(\rho_n, \omega_n)\}_{n\in\mathbb{N}}$ converges. Moreover, the limit does not depend on the choice of the approximating sequences $\{\rho_n\}_{n\in\mathbb{N}}$ and $\{\omega_n\}_{n\in\mathbb{N}}$.

Proof. The proof is similar to the one given in Proposition 4.5 of [6]. It is enough to show that $\{W_{2,\mathcal{L}}(\rho_n, \omega_n)\}_{n\in\mathbb{N}}$ is Cauchy. By the triangle inequality, it is even enough to prove that $W_{2,\mathcal{L}}(\rho_n, \rho_m) \to 0$ as $m, n \to \infty$. Let $\varepsilon \in (0,1)$ and set $\bar{\rho} := (1-\varepsilon)\rho + \varepsilon \frac{\mathbb{I}}{d}$. Let $N \in \mathbb{N}$ be such that for any $n \geq N$, $\mathrm{Tr}[(\rho-\rho_n)^2] \leq \varepsilon^2$. For $n \geq N$, consider the convex interpolation $\gamma(s) := (1-s)\rho_n + s\bar{\rho}$. Since $\gamma(s) \geq \varepsilon s \frac{\mathbb{I}}{d}$ for $s \in [0,1]$, we find from Eq. 2.10 that

$$\begin{aligned}
W_{2\mathcal{L}}(\rho_n, \bar{\rho}) &\leq \int_0^1 \|\dot{\gamma}(s)\|_{g_{\mathcal{L},\gamma(s)}} \mathrm{d}s \\
&= \int_0^1 \Bigg[\sum_{j\in\mathcal{J}} c_j \langle \partial_j(-\operatorname{div} \circ [\gamma(s)]_{\vec{\omega}} \circ \nabla)^{-1}(\dot{\gamma}(s)), \\
&\qquad [\gamma(s)]_{\omega_j} \partial_j(-\operatorname{div} \circ [\gamma(s)]_{\vec{\omega}} \circ \nabla)^{-1}(\dot{\gamma}(s)) \rangle \Bigg]^{\frac{1}{2}} \mathrm{d}s \\
&= \int_0^1 \sqrt{\langle \dot{\gamma}(s), (-\operatorname{div}([\gamma(s)]_{\vec{\omega}} \nabla(.)))^{-1} \ \dot{\gamma}(s) \rangle} \mathrm{d}s \\
&\leq \sqrt{\frac{d\, K_{\mathcal{L}}}{\varepsilon}} \int_0^1 s^{-1/2} \sqrt{\mathrm{Tr}[(\dot{\gamma}(s))^2]} \mathrm{d}s,
\end{aligned}$$

where we used Lemma 3 in the second, third and fourth lines above. Now

$$\begin{aligned}
\mathrm{Tr}[(\dot{\gamma}(s))^2] &= \mathrm{Tr}[(\rho - \rho_n + \varepsilon(\mathbb{I}/d - \rho))^2] \\
&\leq 2\,\mathrm{Tr}[(\rho-\rho_n)^2] + 2\varepsilon^2\,\mathrm{Tr}[(\mathbb{I}/d-\rho)^2] \\
&\leq 2\left(1 + \mathrm{Tr}[(\mathbb{I}/d-\rho)^2]\right)\varepsilon^2.
\end{aligned}$$

Hence, $W_{2,\mathcal{L}}(\rho_n, \bar{\rho}) \leq \sqrt{K(\mathcal{L},\rho)\varepsilon}$, for some constant $K(\mathcal{L},\rho)$ depending on ρ and $\mathcal{L}$. Since ε is arbitrary, we conclude by triangle inequality that $W_{2,\mathcal{L}}(\rho_m, \rho_n) \leq W_{2,\mathcal{L}}(\rho_m, \bar{\rho}) + W_{2,\mathcal{L}}(\bar{\rho}, \rho_n) \to 0$. □

The above proposition justifies the following definition: The modified Wasserstein distance $W_{2,\mathcal{L}}$ between two states $\rho, \omega \in \mathcal{D}(\mathcal{H})$ is defined as

$$W_{2,\mathcal{L}}(\rho,\omega) := \lim_{n\to\infty} W_{2,\mathcal{L}}(\rho_n, \omega_n),$$

where $\{\rho_n\}_{n\in\mathbb{N}}$ and $\{\omega_n\}_{n\in\mathbb{N}}$ are arbitrary sequences in $\mathcal{D}_+(\mathcal{H})$ satisfying (2.13). It can be shown that $(\mathcal{D}(\mathcal{H}), W_{2,\mathcal{L}})$ forms a complete metric space as follows:

Lemma 4. *For any $\rho, \omega \in \mathcal{D}(\mathcal{H})$,*

$$\|\rho-\omega\|_1 \leq 2\left(\sum_{j\in\mathcal{J}} c_j(\mathrm{e}^{-\omega_j/2}+\mathrm{e}^{\omega_j/2})\|\tilde{L}_j\|_\infty^2\right)^{1/2} W_{2,\mathcal{L}}(\rho,\omega).$$

Proof. The proof follows from a direct application of inequality (2.39) of Lemma 6 of [24]: for any $X \in \mathcal{B}_{sa}(\mathcal{H})$:

$$|\operatorname{Tr}(X(\rho-\omega))| \leq \sqrt{d}\ \|X\|_{\mathrm{Lip}} W_{2,\mathcal{L}}(\rho,\omega),$$

where

$$\begin{aligned}\|X\|_{\mathrm{Lip}} &:= \left(\frac{1}{d}\sum_{j\in\mathcal{J}} c_j(\mathrm{e}^{-\omega_j/2}+\mathrm{e}^{\omega_j/2})\|\partial_j X\|_\infty^2\right)^{1/2}\\ &\leq \frac{2}{\sqrt{d}}\left(\sum_{j\in\mathcal{J}} c_j(\mathrm{e}^{-\omega_j/2}+\mathrm{e}^{\omega_j/2})\|\tilde{L}_j\|_\infty^2\right)^{1/2}\|X\|_\infty.\end{aligned}$$

The result follows from the duality relation between the norms $\|.\|_\infty$ and $\|.\|_1$. □

Proposition 2. *The metric space $(\mathcal{D}(\mathcal{H}), W_{2,\mathcal{L}})$ is complete.*

Proof. This directly follows from Lemma 4 and Proposition 1: assume that $\{\rho_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence in $(\mathcal{D}(\mathcal{H}), W_{2,\mathcal{L}})$, that is $W_{2,\mathcal{L}}(\rho_n,\rho_m)\to 0$ as $m,n\to\infty$. Then, by Lemma 4, $\{\rho_n\}_{n\in\mathbb{N}}$ is also Cauchy with respect to the trace norm $\|.\|_1$. By completeness of the normed vector space $(\mathcal{B}(\mathcal{H}), \|.\|_1)$, this implies existence of $\rho_\infty \in \mathcal{B}(\mathcal{H})$ such that $\|\rho_n-\rho_\infty\|_1\to 0$ as $n\to\infty$. Moreover, $\rho_\infty\in\mathcal{D}(\mathcal{H})$: indeed, for any $\psi\in(\mathcal{H},\langle.|.\rangle)$,

$$\langle\psi|\rho_\infty\psi\rangle = \langle\psi|(\rho_\infty-\rho_n)\psi\rangle + \langle\psi|\rho_n\,\psi\rangle,$$

which implies the positivity of ρ_∞, since $|\langle\psi|(\rho_\infty-\rho_n)\psi\rangle| \leq \|\rho_\infty-\rho_n\|_1\langle\psi|\psi\rangle \to 0$ as $n\to\infty$, and $\langle\psi|\rho_n\psi\rangle\geq 0$ for all n. Moreover

$$|\operatorname{Tr}(\rho_n-\rho_\infty)| \leq \|\rho_n-\rho_\infty\|_1 \to 0,\ n\to\infty,$$

which implies $\operatorname{Tr}\rho_\infty = 1$. We conclude that $W_{2,\mathcal{L}}(\rho_n,\rho_\infty)\to W_{2,\mathcal{L}}(\rho_\infty,\rho_\infty) = 0$ by Proposition 1. □

2.4. Quantum Functional and Transportation Cost Inequalities

A primitive QMS $(\Lambda_t)_{t\geq 0}$ with unique invariant state σ is said to satisfy:

1. a Poincaré inequality with constant $\lambda>0$, if for all $f\in\mathcal{B}_{sa}(\mathcal{H})$ with $\operatorname{Tr}(\sigma f)=0$,

$$\lambda \operatorname{Var}_\sigma(f)^2 \leq -\operatorname{Tr}(\sigma f\mathcal{L}(f)), \tag{PI(λ)}$$

where $\operatorname{Var}_\sigma(f) := \operatorname{Tr}(\sigma f^2) - \operatorname{Tr}(\sigma f)^2$.

2. a modified logarithmic Sobolev inequality with constant $\alpha_1 > 0$ if for all $\rho \in \mathcal{D}_+(\mathcal{H})$,
$$2\alpha_1 D(\rho\|\sigma) \leq \mathrm{EP}_\sigma(\rho) = \mathrm{I}_\sigma(\rho). \tag{MLSI(α_1)}$$
3. a transportation cost inequality of order 2 with constant $c_2 > 0$ if for all $\rho \in \mathcal{D}_+(\mathcal{H})$,
$$W_{2,\mathcal{L}}(\rho,\sigma) \leq \sqrt{2c_2 D(\rho\|\sigma)}. \tag{TC$_2$(c_2)}$$
4. MLSI+TC$_2$(c) inequality with constant $c > 0$ if for all $\rho \in \mathcal{D}_+(\mathcal{H})$,
$$W_{2,\mathcal{L}}(\rho,\sigma) \leq c\sqrt{\mathrm{I}_\sigma(\rho)}. \tag{MLSI+TC$_2$(c)}$$

That (MLSI(α_1)) implies (TC$_2$(c_2)) for $c_2 = \alpha_1^{-1}$ was proved in [24]. Hence, the following corollary easily follows:

Corollary 1. *Assume that* $(\Lambda_t)_{t\geq 0}$ *satisfies* (MLSI(α_1)) *for some* $\alpha_1 > 0$. *Then it also satisfies* (MLSI+TC$_2$(c)) *with* $c = \alpha_1^{-1}$.

3. Quantum Ricci Lower Bound and κ-Displacement Convexity

In their celebrated paper [3] (see also [2]), Bakry and Emery found an elegant criterion which implies the logarithmic Sobolev inequality in the setting of diffusions. In this case of Markov semigroups defined on a Riemannian manifold $\mathcal{M}$, this criterion, called the *Ricci lower bound*, which is a special case of the Bakry–Emery condition, was shown later on to be equivalent to the so-called κ-displacement convexity of the relative entropy along geodesics in the Wasserstein space of probability measures on $\mathcal{M}$ in [26]. This notion of κ-displacement convexity was extended to the framework of (necessarily non-diffusive) finite Markov chains by Maas in [19]. Carlen and Maas generalized this notion to the quantum regime in [7] and proved that it implies the modified logarithmic Sobolev inequality as well as the contractivity of the Wasserstein metric under the flow associated with the underlying quantum semigroup $(\Lambda_t)_{t\geq 0}$. In their previous article [6], the same authors had already studied this quantum extension of the notion of κ-displacement convexity in the particular case of the fermionic Fokker–Planck equation. In this section, we provide a systematic analysis of the κ-displacement convexity, including a study of the geodesic equations on the Riemannian manifold $(\mathcal{D}_+(\mathcal{H}), g_{\mathcal{L}})$.

3.1. Geodesic Equations

Similarly to Theorem 2.4 of [12], Carlen and Maas provided in [7] the set of faithful states $\mathcal{D}_+(\mathcal{H})$ with a Riemannian structure with associated Riemannian distance given by $W_{2,\mathcal{L}}$. Therefore, the local existence and uniqueness of constant speed geodesics is guaranteed by standard Riemannian geometry. We first recall that a constant speed geodesic $(\gamma(s), U(s))_{s\in[0,1]}$, where U is related to γ through Eq. (2.6), satisfies a Euler–Lagrange equation that we derive in Theorem 2. This result is a direct generalization of Theorem 5.3 in [6]. We start by recalling the abstract framework. Let $(\mathcal{V}, \langle .,. \rangle)$ be a finite-dimensional real Hilbert space. Let $\mathcal{W} \subset \mathcal{V}$ be a subspace of $\mathcal{V}$, and $z \in \mathcal{V}\backslash\mathcal{W}$. Consider

the affine subspace $\mathcal{W}_z := z + \mathcal{W}$, and let $\mathcal{M} \subset \mathcal{W}_z$ be a relatively open subset. Let $D : \mathcal{M} \to \mathcal{B}(\mathcal{W})$ be a smooth function such that $D(x)$ is self-adjoint and invertible for all $x \in \mathcal{M}$. We shall write $C(x) := D(x)^{-1}$. Consider the Lagrangian $L : \mathcal{W} \times \mathcal{M} \to \mathbb{R}$ defined by $L(p, x) = \langle C(x)p, p\rangle$ and the associated minimization problem:

$$\inf_{u(.)\in C^1([0,1],\mathcal{M})} \left(\int_0^1 L(u'(t), u(t))dt : \ u(0) = u_0, \ u(1) = u_1 \right),$$

where $u_0, u_1 \in \mathcal{M}$ are given boundary values. Then, the Euler–Lagrange equations are equivalent to the following system of equations:

$$\begin{cases} u'(t) - D(u(t))v(t) = 0, \\ v'(t) + \dfrac{1}{2}\langle \partial_x D(u(t))v(t), v(t)\rangle = 0. \end{cases} \tag{3.1}$$

Here, we apply this abstract result to the case where $\mathcal{V} = \mathcal{B}_{sa}(\mathcal{H})$, with inner product $\langle ., .\rangle$ the usual Hilbert–Schmidt inner product, $\mathcal{W} = \{A \in \mathcal{V} : \ \mathrm{Tr}(A) = 0\}$, $z := \mathbb{I}/\dim(\mathcal{H})$, and $\mathcal{M} = \mathcal{D}_+(\mathcal{H})$. Indeed, any density operator ρ can be written as $\rho = \mathbb{I}/\dim \mathcal{H} + K$, for some self-adjoint and traceless operator K. For any $\rho \in \mathcal{D}_+(\mathcal{H})$, we already proved in Lemma 3 that $D_{\vec{\omega}}(\rho) : U \mapsto -\operatorname{div}([\rho]_{\vec{\omega}}\nabla U)$ is invertible and self-adjoint. Now we use the following identity (see [6] p. 21):

$$\frac{\mathrm{d}}{\mathrm{d}t}(\rho + tA)^\alpha\Big|_{t=0} = \int_0^1 \int_0^\alpha \frac{\rho^{\alpha-\beta}}{(1-s)\mathbb{I} + s\rho} A \frac{\rho^\beta}{(1-s)\mathbb{I} + s\rho} \mathrm{d}\beta \mathrm{d}s \tag{3.2}$$

for any $0 < \alpha < 1$, $\rho \in \mathcal{D}_+(\mathcal{H})$ and $A \in \mathcal{W}$. Hence for all $A, U \in \mathcal{W}$,

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \langle D_{\vec{\omega}}(\rho + tA)[U], U\rangle &= \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \sum_{j\in\mathcal{J}} c_j \langle \partial_j U, [\rho + tA]_{\omega_j} \partial_j U\rangle \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} \sum_{j\in\mathcal{J}} c_j \langle \partial_j U, \int_0^1 \mathrm{e}^{\omega_j(1/2-\alpha)} (\rho + tA)^\alpha \\ &\quad \partial_j U (\rho + tA)^{1-\alpha}\rangle \mathrm{d}\alpha \\ &= \langle A, \nabla U ._\rho \nabla U\rangle, \end{aligned} \tag{3.3}$$

where for two vectors $\vec{V}_1, \vec{V}_2$ in $\bigoplus_j \mathcal{B}(\mathcal{H})$,

$$\begin{aligned} \vec{V}_1 ._\rho \vec{V}_2 := \sum_{j\in\mathcal{J}} c_j \int_0^1 \int_0^1 \mathrm{e}^{\omega_j(1/2-\alpha)} (\chi_j(\vec{V}_1, \vec{V}_2^*, \rho, \alpha, s) \\ + \chi_j(\vec{V}_1^*, \vec{V}_2, \rho, 1-\alpha, s)) \mathrm{d}\alpha \mathrm{d}s, \end{aligned} \tag{3.4}$$

where

$$\chi_j(\vec{V}_1, \vec{V}_2, \rho, \alpha, s) := \int_0^\alpha \frac{\rho^\beta}{(1-s)\mathbb{I} + s\rho} (V_1)_j \ \rho^{1-\alpha} \ (V_2)_j \frac{\rho^{\alpha-\beta}}{(1-s)\mathbb{I} + s\rho} \mathrm{d}\beta.$$

Therefore, in our context the Euler–Lagrange equations (3.1) reduce to the following:

Theorem 2. *The geodesic equations in the Riemannian manifold* $(\mathcal{D}_+(\mathcal{H}), W_{2,\mathcal{L}})$ *are given by*

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}s}\gamma(s) + \mathrm{div}([\gamma(s)]_{\vec{\omega}}\nabla U(s)) = 0, \\ \dfrac{\mathrm{d}U(s)}{\mathrm{d}s} + \dfrac{1}{2}\nabla U(s).{}_{\gamma(s)}\nabla U(s) = 0. \end{cases} \tag{3.5}$$

3.2. Different Formulations of Quantum κ-Displacement Convexity

In analogy with [12], we say that a primitive quantum Markov semigroup $(\Lambda_t)_{t\geq 0}$ with associated invariant state σ and generator $\mathcal{L}$ of the form of Eq. (2.3) has *Ricci curvature bounded from below* by a constant $\kappa \in \mathbb{R}$ if the following inequality holds:

$$\frac{\mathrm{d}^2}{\mathrm{d}s^2}\Big|_{s=0} D(\gamma(s)\|\sigma) \geq \kappa \|\dot{\gamma}(0)\|^2_{g_{\mathcal{L},\rho}}, \qquad (\mathrm{Ric}(\mathcal{L}) \geq \kappa)$$

where $(\gamma(s), U(s))_{s\in(-\varepsilon,\varepsilon)}$ is the unique solution to the geodesic equation (3.5) such that $\mathcal{D}_+(\mathcal{H}) \ni \rho := \gamma(0)$ and $U(0) = U$. We also refer to the above inequality as the *quantum Ricci lower bound*. Theorem 2 is useful to derive an expression for the second derivative of the relative entropy $D(\gamma(s)\|\sigma)$ with respect to s, where $(\gamma(s))_{s\in(-\varepsilon,\varepsilon)}$ is a constant speed geodesic with associated tangent vector $\nabla U(s)$ at each s. We already know from the gradient flow equation (2.12) that

$$\begin{aligned} &\frac{\mathrm{d}}{\mathrm{d}s} D(\gamma(s)\|\sigma) \\ &\quad = -g_{\mathcal{L},\gamma(s)}(\dot{\gamma}(s), \mathcal{L}_*(\gamma(s))) \\ &\quad = \sum_{j\in\mathcal{J}} c_j \langle \partial_j U(s), [\gamma(s)]_{\omega_j} \partial_j (\log\gamma(s) - \log\sigma)\rangle \\ &\quad = \sum_{j\in\mathcal{J}} c_j \langle \partial_j U(s), [\gamma(s)]_{\omega_j} (\tilde{L}_j \log(\mathrm{e}^{-\omega_j/2}\gamma(s)) - \log(\mathrm{e}^{\omega_j/2}\gamma(s))\tilde{L}_j)\rangle, \end{aligned}$$

where the second line comes from Theorem 5.10 in [7], and the last identity comes from Lemma 5.9 of [7]. Now by identity (5.6) of the same paper,

$$[\gamma(s)]_{\omega_j}(\tilde{L}_j \log(\mathrm{e}^{-\omega_j/2}\gamma(s)) - \log(\mathrm{e}^{\omega_j/2}\gamma(s))\tilde{L}_j) = \mathrm{e}^{-\omega_j/2}\tilde{L}_j\gamma(s) - \mathrm{e}^{\omega_j/2}\gamma(s)\tilde{L}_j,$$

so that we finally get

$$\frac{\mathrm{d}}{\mathrm{d}s} D(\gamma(s)\|\sigma) = \sum_{j\in\mathcal{J}} c_j \langle \partial_j U(s), \mathrm{e}^{-\omega_j/2}\tilde{L}_j\gamma(s) - \mathrm{e}^{\omega_j/2}\gamma(s)\tilde{L}_j\rangle.$$

Differentiating once more, we get:

$$\begin{aligned} \frac{\mathrm{d}^2}{\mathrm{d}s^2} D(\gamma(s)\|\sigma)\Big|_{s=0} = \sum_{j\in\mathcal{J}} c_j \Big\{ &\langle \partial_j \frac{\mathrm{d}}{\mathrm{d}s} U(s)\Big|_{s=0}, \mathrm{e}^{-\omega_j/2}\tilde{L}_j\rho - \mathrm{e}^{\omega_j/2}\rho\tilde{L}_j\rangle \\ &+ \langle \partial_j U, \mathrm{e}^{-\omega_j/2}\tilde{L}_j\dot{\gamma}(0) - \mathrm{e}^{\omega_j/2}\dot{\gamma}(0)\tilde{L}_j\rangle \Big\}. \end{aligned} \tag{3.6}$$

We first take care of the second line of Eq. (3.6). Using Theorem 2 as well as Eq. (2.3), we find

$$\begin{aligned}
&\langle \partial_j U, \mathrm{e}^{-\omega_j/2}\tilde{L}_j\dot{\gamma}(0) - \mathrm{e}^{\omega_j/2}\dot{\gamma}(0)\tilde{L}_j\rangle \\
&\quad = -\langle \partial_j U, \mathrm{e}^{-\omega_j/2}\tilde{L}_j \operatorname{div}([\rho]_{\vec{\omega}}\nabla U) - \mathrm{e}^{\omega_j/2}\operatorname{div}([\rho]_{\vec{\omega}}\nabla U)\tilde{L}_j\rangle \\
&\quad = -\langle \partial_j U, \mathrm{e}^{-\omega_j/2}\tilde{L}_j \sum_{k\in\mathcal{J}} c_k[[\rho]_{\omega_k}\partial_k U, \tilde{L}_k^*] - \mathrm{e}^{\omega_j/2}\sum_{k\in\mathcal{J}} c_k[[\rho]_{\omega_k}\partial_k U, \tilde{L}_k^*]\tilde{L}_j\rangle \\
&\quad = \sum_{k\in\mathcal{J}} c_k\left(\mathrm{e}^{-\omega_j/2}\langle \partial_k(\tilde{L}_j^*\partial_j U), [\rho]_{\omega_k}\partial_k U\rangle - \mathrm{e}^{\omega_j/2}\langle \partial_k(\partial_j U\tilde{L}_j^*), [\rho]_{\omega_k}\partial_k U\rangle\right) \\
&\quad = \sum_{k\in\mathcal{J}} c_k\langle \partial_k\left(\mathrm{e}^{-\omega_j/2}\tilde{L}_j^*\partial_j U - \mathrm{e}^{\omega_j/2}\partial_j U\tilde{L}_j^*\right), [\rho]_{\omega_k}\partial_k U\rangle.
\end{aligned}$$

Hence by (2.3),

$$\begin{aligned}
\sum_{j\in\mathcal{J}} c_j\langle \partial_j U, \mathrm{e}^{-\omega_j/2}\tilde{L}_j\dot{\gamma}(0) - \mathrm{e}^{\omega_j/2}\dot{\gamma}(0)\tilde{L}_j\rangle &= -\sum_k c_k\langle \partial_k\mathcal{L}(U), [\rho]_{\omega_k}\partial_k U\rangle \\
&= -\langle \nabla\mathcal{L}(U), \nabla U\rangle_{\mathcal{L},\rho}. \qquad (3.7)
\end{aligned}$$

By (3.5), the first line of (3.6) is equal to

$$\begin{aligned}
&\frac{1}{2}\sum_{j\in\mathcal{J}} c_j\langle \partial_j(\nabla U\cdot_\rho \nabla U), \mathrm{e}^{\omega_j/2}\rho\tilde{L}_j - \mathrm{e}^{-\omega_j/2}\tilde{L}_j\rho\rangle \\
&\quad = \frac{1}{2}\sum_{j\in\mathcal{J}} c_j\langle \nabla U\cdot_\rho\nabla U, [\tilde{L}_j^*, \rho\tilde{L}_j]\mathrm{e}^{\omega_j/2} - \mathrm{e}^{-\omega_j/2}[\tilde{L}_j^*, \tilde{L}_j\rho]\rangle \\
&\quad = \frac{1}{2}\langle \nabla U\cdot_\rho\nabla U, \mathcal{L}_*(\rho)\rangle, \qquad (3.8)
\end{aligned}$$

where we used that, replacing $\tilde{L}_j$ by $\tilde{L}_j^*$ so that $\omega_j \to -\omega_j$ and $c_j \to c_j$,

$$\begin{aligned}
\mathcal{L}_*(\rho) &= \sum_{j\in\mathcal{J}} c_j\left(\mathrm{e}^{\omega_j/2}[\tilde{L}_j^*\rho, \tilde{L}_j] + \mathrm{e}^{-\omega_j/2}[\tilde{L}_j, \rho\tilde{L}_j^*]\right) \\
&= \sum_{j\in\mathcal{J}} c_j\left(\mathrm{e}^{-\omega_j/2}[\tilde{L}_j\rho, \tilde{L}_j^*] + \mathrm{e}^{\omega_j/2}[\tilde{L}_j^*, \rho\tilde{L}_j]\right),
\end{aligned}$$

Hence, using (3.7) and (3.8), (3.6) reduces to

$$\frac{\mathrm{d}^2}{\mathrm{d}s^2}D(\gamma(s)\|\sigma)\Big|_{s=0} = \frac{1}{2}\langle \nabla U\cdot_\rho\nabla U, \mathcal{L}_*(\rho)\rangle - \langle \nabla\mathcal{L}(U), \nabla U\rangle_{\mathcal{L},\rho}. \qquad (3.9)$$

One can compare this expression with the one derived in Proposition 4.3 of [12]. To make this analogy more clear, we denote the quantity on the right-hand side of Eq. (3.9) by $B(\rho, U)$ so that

$$\frac{\mathrm{d}^2}{\mathrm{d}s^2}D(\gamma(s)\|\sigma)\Big|_{s=0} = B(\rho, U). \qquad (3.10)$$

The following lemma extends Lemma 4.6 of [12] to the quantum regime, as well as part of the proof of Proposition 5.11 of [6], and is proven to be useful in what follows:

Lemma 5. *Let $(\gamma(s))_{s\in[0,1]}$ be a smooth curve in $\mathcal{D}_+(\mathcal{H})$. For each $t \geq 0$, set $\gamma(s,t) := \Lambda_{*st}(\gamma(s))$, and let $(U(s,t))_{s\in[0,1]}$ be a smooth curve satisfying the continuity equation*

$$\partial_s(\gamma(s,t)) + \operatorname{div}([\gamma(s,t)]_{\vec{\omega}}\nabla U(s,t)) = 0, \ \ s \in [0,1]. \tag{3.11}$$

Therefore,

$$\frac{1}{2}\partial_t \|\partial_s\gamma(s,t)\|^2_{g_{\mathcal{L},\gamma(s,t)}} + \partial_s D(\gamma(s,t)\|\sigma) = -sB(\gamma(s,t),U(s,t)).$$

Proof. Start by noticing that

$$\begin{aligned}
\partial_s D(\gamma(s,t)\|\sigma) &= \partial_s \operatorname{Tr}(\gamma(s,t)(\log\gamma(s,t) - \log\sigma)) \\
&= \operatorname{Tr}(\partial_s\gamma(s,t)(\log\gamma(s,t) - \log\sigma)) \\
&= -\operatorname{Tr}((\log\gamma(s,t) - \log\sigma)\operatorname{div}([\gamma(s,t)]_{\vec{\omega}}\nabla U(s,t))) \\
&= -\langle \log\gamma(s,t) - \log\sigma, \operatorname{div}([\gamma(s,t)]_{\vec{\omega}}\nabla U(s,t))\rangle \\
&= -\sum_{j\in\mathcal{J}} c_j \langle \log\gamma(s,t) - \log\sigma, [[\gamma(s,t)]_{\omega_j}(\partial_j U(s,t)), \tilde{L}_j^*]\rangle \\
&= \sum_{j\in\mathcal{J}} c_j \langle \partial_j(\log\gamma(s,t) - \log\sigma), [\gamma(s,t)]_{\omega_j}(\partial_j U(s,t))\rangle \\
&= \sum_{j\in\mathcal{J}} c_j \langle [\gamma(s,t)]_{\omega_j}(\partial_j(\log\gamma(s,t) - \log\sigma)), \partial_j U(s,t)\rangle \\
&= \sum_{j\in\mathcal{J}} c_j \langle \partial_j^*[\gamma(s,t)]_{\omega_j}(\partial_j(\log\gamma(s,t) - \log\sigma)), U(s,t)\rangle \\
&= -\langle \mathcal{L}_*(\gamma(s,t)), U(s,t)\rangle,
\end{aligned}$$

where in the third line we used (3.11), in the last line we used Theorem 5.10 of [7], and in the second line we used that

$$\begin{aligned}
\operatorname{Tr}(\gamma(s,t)\partial_s\log\gamma(s,t)) &= \operatorname{Tr}\left(\gamma(s,t)\partial_s\int_0^\infty \frac{1}{(1+u)\mathbb{I}} - \frac{1}{\gamma(s,t)+u\mathbb{I}}\,\mathrm{d}u\right) \\
&= \operatorname{Tr}\gamma(s,t)\int_0^\infty \frac{1}{\gamma(s,t)+u\mathbb{I}}\partial_s\gamma(s,t)\frac{1}{\gamma(s,t)+u\mathbb{I}}\,\mathrm{d}u \\
&= \operatorname{Tr}\partial_s\gamma(s,t) \\
&= 0.
\end{aligned}$$

Moreover, by definition of the metric $g_{\mathcal{L}}$ through Eqs. (2.8) and (2.7),

$$\begin{aligned}
&\frac{1}{2}\partial_t\|\partial_s\gamma(s,t)\|^2_{g_{\mathcal{L},\gamma(s,t)}} \\
&\quad= \frac{1}{2}\partial_t\sum_{j\in\mathcal{J}} c_j\langle\partial_j U(s,t), [\gamma(s,t)]_{\omega_j}\partial_j U(s,t)\rangle \\
&\quad= \sum_{j\in\mathcal{J}} c_j\Big(\langle\partial_t(\partial_j U(s,t)), [\gamma(s,t)]_{\omega_j}\partial_j U(s,t)\rangle \\
&\qquad+ \frac{1}{2}\langle\partial_j U(s,t), \partial_t([\gamma(s,t)]_{\omega_j})\partial_j U(s,t)\rangle\Big).
\end{aligned} \tag{3.12}$$

From (3.3),

$$\begin{aligned}\sum_{j\in\mathcal{J}} c_j\langle\partial_j U(s,t),\partial_t([\gamma(s,t)]_{\omega_j})\partial_j U(s,t)\rangle\\ &= \langle\partial_t\gamma(s,t),\nabla U(s,t).\nabla U(s,t)\rangle\\ &= s\langle\mathcal{L}_*(\gamma(s,t)),\nabla U(s,t).\nabla U(s,t)\rangle. \end{aligned}\tag{3.13}$$

Moreover,

$$\begin{aligned}&\sum_{j\in\mathcal{J}} c_j\langle\partial_t(\partial_j U(s,t)),[\gamma(s,t)]_{\omega_j}\partial_j U(s,t)\rangle\\ &\quad= -\sum_{j\in\mathcal{J}} c_j\langle\partial_t U(s,t),[[\gamma(s,t)]_{\omega_j}(\partial_j U(s,t)),\tilde{L}_j^*]\rangle\\ &\quad= -\langle\partial_t U(s,t),\mathrm{div}([\gamma(s,t)]_{\vec{\omega}}(\nabla U(s,t)))\rangle\\ &\quad= \langle\partial_t U(s,t),\partial_s\gamma(s,t)\rangle\\ &\quad= \partial_t(\langle U(s,t),\partial_s\gamma(s,t)\rangle) - \langle U(s,t),\partial_s\partial_t\gamma(s,t)\rangle\\ &\quad= \partial_t\|\partial_s\gamma(s,t)\|^2_{g_{\mathcal{L},\gamma(s,t)}} - \langle U(s,t),\partial_s\partial_t\gamma(s,t)\rangle\\ &\quad= \partial_t\|\partial_s\gamma(s,t)\|^2_{g_{\mathcal{L},\gamma(s,t)}} - \langle U(s,t),\partial_s(s\mathcal{L}_*(\gamma(s,t)))\rangle, \end{aligned}\tag{3.14}$$

where we used once again (3.11) in the third and fifth lines above. Therefore, using (3.13) and (3.14), the right-hand side of (3.12) reduces to

$$\begin{aligned}\frac{1}{2}\partial_t\|\partial_s\gamma(s,t)\|^2_{g_{\mathcal{L},\gamma(s,t)}} &= \langle U(s,t),\partial_s(s\mathcal{L}_*(\gamma(s,t)))\rangle\\ &\quad - \frac{1}{2}s\langle\mathcal{L}_*(\gamma(s,t)),\nabla U(s,t).\nabla U(s,t)\rangle.\end{aligned}$$

Hence,

$$\begin{aligned}&\frac{1}{2}\partial_t\|\partial_s\gamma(s,t)\|^2_{g_{\mathcal{L},\gamma(s,t)}} + \partial_s D(\gamma(s,t)\|\sigma)\\ &\quad= s\langle U(s,t),\mathcal{L}_*\partial_s(\gamma(s,t))\rangle - \frac{1}{2}s\langle\mathcal{L}_*\gamma(s,t),\ \nabla U(s,t).\nabla U(s,t)\rangle\\ &\quad= -s\langle\mathcal{L}(U(s,t)),\mathrm{div}([\gamma(s,t)]_{\vec{\omega}}\nabla U(s,t))\rangle\\ &\qquad - \frac{1}{2}s\langle\mathcal{L}_*\gamma(s,t),\nabla U(s,t).\nabla U(s,t)\rangle\\ &\quad= s\langle\nabla\mathcal{L}(U(s,t)),\nabla U(s,t)\rangle_{\mathcal{L},\gamma(s,t)}\\ &\qquad - \frac{1}{2}s\langle\mathcal{L}_*(\gamma(s,t)),\nabla U(s,t).\nabla U(s,t)\rangle\\ &\quad= -sB(\gamma(s,t),U(s,t)),\end{aligned}$$

which is what needed to be proved. □

Theorem 3. *Let $\mathcal{L}$ be the generator of an ergodic QMS $(\Lambda_t)_{t\geq 0}$, with unique invariant state σ, of the form of Eq. 2.3). Then, for $\kappa\in\mathbb{R}$, the following are equivalent:*

(i) $\mathrm{Ric}(\mathcal{L})\geq\kappa$

(ii) *For all $\rho \in \mathcal{D}_+(\mathcal{H})$, and $U \in \mathcal{W}$, the space of self-adjoint, traceless operators on $\mathcal{H}$,*

$$B(\rho, U) \geq \kappa \|\nabla U\|^2_{\mathcal{L},\rho}\,.$$

(iii) *For all $\rho, \omega \in \mathcal{D}_+(\mathcal{H})$ and all $t \geq 0$, writing $\rho_t := \Lambda_{*t}(\rho)$:*

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\bigg|_{t^+} (W_{2,\mathcal{L}}(\rho_t, \omega))^2 + \frac{\kappa}{2} W_{2,\mathcal{L}}(\rho_t, \omega)^2 \leq D(\omega\|\sigma) - D(\rho_t\|\sigma). \tag{3.15}$$

(iv) *Equation (3.15) holds for any $\rho, \omega \in \mathcal{D}(\mathcal{H})$.*

(v) *κ-displacement convexity of the relative entropy: given any constant speed geodesic $(\gamma(s))_{s\in[0,1]}$ in $\mathcal{D}(\mathcal{H})$,*

$$D(\gamma(s)\|\sigma) \leq (1-s)D(\gamma(0)\|\sigma) + sD(\gamma(1)\|\sigma) - \frac{\kappa}{2}s(1-s)W_{2,\mathcal{L}}(\gamma(0),\gamma(1))^2. \tag{3.16}$$

Proof. The proof is inspired by the one of Theorem 4.5 of [12]. That $(i) \Leftrightarrow (ii)$ follows from Eq. (3.10). We use Lemma 5 to show that $(ii) \Rightarrow (iii)$: Take a smooth path $(\gamma(s), U(s))_{s\in[0,1]}$ such that $\gamma(0) = \omega$, $\gamma(1) = \rho$ and

$$\int_0^1 \|\dot\gamma(s)\|^2_{g_{\mathcal{L},\gamma(s)}} \mathrm{d}s \leq W_{2,\mathcal{L}}(\rho,\omega)^2 + \varepsilon. \tag{3.17}$$

With the notations of Lemma 5,

$$\frac{1}{2}\partial_t\left(\mathrm{e}^{2\kappa st}\|\partial_s\gamma(s,t)\|^2_{g_{\mathcal{L},\gamma(s,t)}}\right) + \partial_s\left(\mathrm{e}^{2\kappa st}D(\gamma(s,t)\|\sigma)\right) \leq 2\kappa t\mathrm{e}^{2\kappa st}D(\gamma(s,t)\|\sigma).$$

Integrating with respect to $t \in [0,h]$, for some $h > 0$, and $s \in [0,1]$,

$$\begin{aligned}\frac{1}{2}\int_0^1 \left(\mathrm{e}^{2\kappa sh}\|\partial_s\gamma(s,h)\|^2_{g_{\mathcal{L},\gamma(s,h)}} - \|\partial_s\gamma(s,0)\|^2_{g_{\mathcal{L},\gamma(s,0)}}\right)\mathrm{d}s \qquad (3.18)\\ + \int_0^h \left(\mathrm{e}^{2\kappa t}D(\gamma(1,t)\|\sigma) - D(\gamma(0,t)\|\sigma)\right)\mathrm{d}t\\ \leq 2\kappa\int_0^1 \mathrm{d}s\int_0^h \mathrm{d}t\; t\,\mathrm{e}^{2\kappa st}D(\gamma(s,t)\|\sigma). \qquad (3.19)\end{aligned}$$

The following inequality, for which a classical equivalent is given in the proof of Theorem 4.5 of [12], can be derived similarly to Lemma 5.1 of [9]:

$$m(\kappa h)W_{2,\mathcal{L}}(\rho_h,\omega)^2 \leq \int_0^1 \mathrm{e}^{2\kappa sh}\|\partial_s\gamma(s,h)\|^2_{g_{\mathcal{L},\gamma(s,h)}}\mathrm{d}s, \tag{3.20}$$

where $m(x) := x\mathrm{e}^x/\sinh(x)$. Indeed, define $f : s \mapsto \mathrm{e}^{2\kappa sh}$ and denote $L_f := \int_0^1 \frac{1}{f(s)}\mathrm{d}s$. Then, let $g : [0,1] \mapsto [0,1]$ be the smooth increasing map defined as $g(s) = L_f^{-1}\int_0^s \frac{1}{f(u)}\mathrm{d}u$ and denote its inverse k such that $k'(g(s)) = L_f f(s)$. Then define the reparametrized curve $(\gamma(k(r),h),\, k'(r)\,U(k(r),h))_{r\in[0,1]}$ which satisfies the continuity equation:

$$\begin{aligned}\partial_r\gamma(k(r),h) &= k'(r)\partial_1\gamma(k(r),h)\\ &= -k'(r)\,\mathrm{div}([\gamma(k(r),h)]_{\vec\omega}\nabla U(k(r),h)),\end{aligned}$$

where we used Eq. (3.11) in order to establish the second line. This curve satisfies $\gamma(k(0),h)=\omega$ and $\gamma(k(1),h)=\rho_h$, so that

$$\begin{aligned} W_{2,\mathcal{L}}(\rho_h,\omega)^2 &\leq \int_0^1 \|\partial_r\gamma(k(r),h)\|^2_{g_{\mathcal{L},\gamma(k(r),h)}}\mathrm{d}r \\ &= \int_0^1 k'(r)^2\|\nabla U(k(r),h)\|^2_{\mathcal{L},\gamma(k(r),h)}\mathrm{d}r \\ &= \int_0^1 k'(g(s))\|\nabla U(s,h)\|^2_{\mathcal{L},\gamma(s,h)}\mathrm{d}s \\ &= L_f\int_0^1 f(s)\|\partial_s\gamma(s,h)\|^2_{g_{\mathcal{L},\gamma(s,h)}}\mathrm{d}s, \end{aligned}$$

which directly leads to (3.20). This inequality, together with (3.17), implies

$$\begin{aligned} &\frac{m(h\kappa)}{2}W_{2,\mathcal{L}}(\rho_h,\omega)^2-\frac{1}{2}W_{2,\mathcal{L}}(\rho,\omega)^2-\varepsilon+\int_0^h \mathrm{e}^{2\kappa t}\mathrm{d}t\ D(\rho_h\|\sigma)-hD(\omega\|\sigma) \\ &\leq \frac{1}{2}\int_0^1 \mathrm{e}^{2\kappa sh}\|\partial_s\gamma(s,t)\|^2_{g_{\mathcal{L},\gamma(s,t)}}\mathrm{d}s-\frac{1}{2}\int_0^1\|\dot\gamma(s)\|^2_{g_{\mathcal{L},\gamma(s)}}\mathrm{d}s \\ &\quad+\int_0^h \mathrm{e}^{2\kappa t}D(\rho_t\|\sigma)\ \mathrm{d}t-hD(\omega\|\sigma) \\ &\leq 2\kappa\int_0^1\int_0^h t\mathrm{e}^{2\kappa st}D(\gamma(s,t)\|\sigma)\,\mathrm{d}t\,\mathrm{d}s. \end{aligned}$$

where, in the first inequality, we also used the monotonicity of the relative entropy so that $D(\rho_h\|\sigma)=D(\rho_h\|\Lambda_{*h}\sigma)\leq D(\rho_t\|\Lambda_{*t}\sigma)=D(\rho_t\|\sigma)$, and in the second one that for all $t>0$, $\gamma(1,t)=\rho_t$, $\gamma(0,t)=\omega$, as well as (3.19). Since for all $s\in[0,1]$, $t\mapsto D(\gamma(s,t)\|\sigma)$ is bounded,

$$\lim_{h\to 0}\frac{1}{h}\int_0^1\int_0^h t\mathrm{e}^{2\kappa st}D(\gamma(s,t)\|\sigma)\mathrm{d}t\ \mathrm{d}s=0.$$

Moreover,

$$\lim_{h\to 0}\frac{1}{h}\left(\int_0^h \mathrm{e}^{2\kappa t}\mathrm{d}t\ D(\rho_h\|\sigma)-hD(\omega\|\sigma)\right)=D(\rho\|\sigma)-D(\omega\|\sigma)$$

Since $\varepsilon>0$ is arbitrary, we arrive at

$$\frac{\mathrm{d}}{\mathrm{d}h}\Big|_{h=0^+}\left(\frac{m(\kappa h)}{2}W_{2,\mathcal{L}}(\rho_h,\omega)^2\right)+D(\rho\|\sigma)-D(\omega\|\sigma)\leq 0.$$

The result for $t=0$ follows from the fact that the first term in the left-hand side above is equal to $\frac{\kappa}{2}W_{2,\mathcal{L}}(\rho_h,\omega)^2+\frac{1}{2}\ \frac{\mathrm{d}}{\mathrm{d}h}\big|_{h=0^+}W_{2,\mathcal{L}}(\rho_h,\omega)^2$. The case $t\geq 0$ directly follows from the case $t=0$.

$(iii)\Rightarrow(iv)$ follows from Theorem 3.3 of [9] together with the fact that $(\mathcal{D}(\mathcal{H}),W_{2,\mathcal{L}})$ is complete (cf. Proposition 2).

$(iv)\Rightarrow(v)$ follows directly from Theorem 3.2 of [9].

$(v)\Rightarrow(i)$ can easily be proved as follows: let $0<\varepsilon<\varepsilon'$, and without loss of generality, let $\gamma:(-\varepsilon',\varepsilon')\to\mathcal{D}_+(\mathcal{H})$ be speed 1 geodesic, and that $\gamma(0)=\rho$. Then,

construct the following constant speed geodesic $\tilde{\gamma} : [0,1] \to \mathcal{D}_+(\mathcal{H})$ as follows: for any $s \in [0,1]$, $\tilde{\gamma}(s) := \gamma(2\varepsilon s - \varepsilon)$. It then follows that $W_{2,\mathcal{L}}(\tilde{\gamma}(0), \tilde{\gamma}(1)) = 2\varepsilon$. Moreover, by applying (3.16) to $\tilde{\gamma}$, we find, after a suitable rearrangement of the terms:

$$\frac{D(\gamma(\varepsilon)\|\sigma) - 2D(\rho\|\sigma) + D(\gamma(-\varepsilon)\|\sigma)}{\varepsilon^2} \geq \kappa.$$

The result follows after taking the limit $\varepsilon \to 0$. □

3.3. Other Equivalent Formulations of Displacement Convexity

Here, we provide other characterizations of the Ricci curvature lower bound in terms of some contraction properties of the Wasserstein metric along the semigroup $(\mathcal{P}_t)_{t\geq 0}$. In the next theorem, the characterization of displacement convexity in terms of gradient estimates can be interpreted as a non-commutative version of Bakry–Émery's original gradient bound (see Theorem 4.7.2 of [4]): in particular they showed that the Ricci curvature lower bound is equivalent to the following pointwise inequality for smooth enough functions:

$$\Gamma(P_t(f), P_t(f)) \leq \mathrm{e}^{-2\kappa t} P_t(\Gamma(f,f)), \tag{3.21}$$

where Γ stands for the carré du champ operator:

$$\Gamma(f,g) := \nabla f . \nabla g .$$

Proposition 3 (Gradient estimate). *$\mathrm{Ric}(\mathcal{L}) \geq \kappa$ is equivalent to the following gradient estimate: for any $\rho \in \mathcal{D}_+(\mathcal{H})$, any $U \in \mathcal{B}_{sa}(\mathcal{H})$ with $\mathrm{Tr}(U) = 0$, and all $t > 0$:*

$$\|\nabla(\Lambda_t(U))\|^2_{\mathcal{L},\rho} \leq \mathrm{e}^{-2\kappa t} \|\nabla U\|^2_{\mathcal{L},\Lambda_{*t}(\rho)}. \tag{3.22}$$

Proof. Define for $u \in [0,t]$ $\rho_u \equiv \Lambda_{*u}(\rho)$ and $U_u \equiv \Lambda_u(U)$. Then,

$$\Phi(s) := \mathrm{e}^{-2\kappa s} \|\nabla U_{t-s}\|^2_{\mathcal{L},\rho_s} \equiv \mathrm{e}^{-2\kappa s} \sum_{j\in\mathcal{J}} c_j \langle \partial_j(U_{t-s}), [\rho_s]_{\omega_j} \partial_j(U_{t-s}) \rangle.$$

Then, $\Phi(0) = \|\nabla U_t\|^2_{\mathcal{L},\rho}$ and $\Phi(t) = \mathrm{e}^{-2\kappa t} \|\nabla U\|^2_{\mathcal{L},\rho_t}$. It is then enough to prove that Φ has non-negative derivative to prove the claim. But:

$$\begin{aligned}
\Phi'(s) &= 2\mathrm{e}^{-2\kappa s}\Big[-\kappa \|\nabla U_{t-s}\|^2_{\mathcal{L},\rho_s} + \frac{1}{2}\frac{\partial}{\partial s} \|\nabla U_{t-s}\|^2_{\mathcal{L},\rho_s} \Big] \\
&= 2\mathrm{e}^{-2\kappa s}\Big[-\kappa \|\nabla U_{t-s}\|^2_{\mathcal{L},\rho_s} + \sum_{j\in\mathcal{J}} c_j \langle \partial_j \partial_s U_{t-s}, [\rho_s]_{\omega_j} \partial_j U_{t-s} \rangle \\
&\quad + \frac{1}{2} \sum_{j\in\mathcal{J}} c_j \langle \partial_j U_{t-s}, \partial_s([\rho_s]_{\omega_j}) \partial_j U_{t-s} \rangle \Big] \\
&= 2\mathrm{e}^{-2\kappa s}\Big[-\kappa \|\nabla U_{t-s}\|^2_{\mathcal{L},\rho_s} - \langle \nabla \mathcal{L}(U_{t-s}), \nabla U_{t-s} \rangle_{\mathcal{L},\rho_s} \\
&\quad + \frac{1}{2} \langle \mathcal{L}_*(\rho_s), \nabla U_{t-s} . \rho_s \nabla U_{t-s} \rangle \Big] \\
&= 2\mathrm{e}^{-2\kappa s}\Big[-\kappa \|\nabla U_{t-s}\|^2_{\mathcal{L},\rho_s} + B(\rho_s, U_{t-s}) \Big]
\end{aligned}$$

where we used Eq. (3.3) in the second line. We conclude by a use of (ii) of Theorem 3. For the reverse implication, assume that Eq. (3.22) holds. Then,

$$\begin{aligned} 0 &\leq \mathrm{e}^{-2\kappa t}\|\nabla U\|^2_{\mathcal{L},\rho_t} - \|\nabla\left(\Lambda_t U\right)\|^2_{\mathcal{L},\rho} \\ &= (\mathrm{e}^{-2\kappa t}-1)\|\nabla U\|^2_{\mathcal{L},\rho} + \mathrm{e}^{-2\kappa t}(\|\nabla U\|^2_{\mathcal{L},\rho_t} - \|\nabla U\|^2_{\mathcal{L},\rho}) \\ &\quad - \|\nabla(\Lambda_t(U))\|^2_{\mathcal{L},\rho} + \|\nabla U\|^2_{\mathcal{L},\rho}\,. \end{aligned}$$

By dividing by t and letting $t \to 0$, we once again obtain that $-\kappa\|\nabla U\|^2_{\mathcal{L},\rho} + B(\rho, U) \geq 0$. □

In the commutative diffusive setting, the Ricci curvature lower bound is also known to be equivalent to the contraction of the Wasserstein distance along the semigroup $(P_t)_{t\geq 0}$ (see Theorem 9.7.2 of [4]):

$$W_2(P_{t*}(\nu),\, P_{t*}(\nu')) \leq \mathrm{e}^{-\kappa t} W_2(\nu, \nu')\,.$$

This still holds true in the non-commutative, finite-dimensional setting:

Proposition 4. *For any $\kappa \in \mathbb{R}$, $\mathrm{Ric}(\mathcal{L}) \geq \kappa$ is equivalent to the contraction of the Wasserstein distance along the flow generated by $(\Lambda_t)_{t\geq 0}$: for any $\rho, \omega \in \mathcal{D}_+(\mathcal{H})$*

$$W_{2,\mathcal{L}}(\Lambda_{t*}(\rho), \Lambda_{t*}(\omega)) \leq \mathrm{e}^{-\kappa t} W_{2,\mathcal{L}}(\rho, \omega)\,. \tag{3.23}$$

Proof. The direct implication follows from Proposition 3.1 of [9] and Theorem 3(iii). The reverse implication is proved as in inequality (2.12) of [9], using the smooth Riemannian structure provided by $(\mathcal{D}_+(\mathcal{H}), W_{2,\mathcal{L}})$ in the finite-dimensional case. □

In the commutative diffusive setting, the contraction of (3.21) is actually known to be equivalent to its "square root" version, usually referred to as the *strong gradient bound*:

$$\sqrt{\Gamma(P_t(f), P_t(f))} \leq \mathrm{e}^{-\kappa t} P_t(\sqrt{\Gamma(f,f)})\,. \tag{3.24}$$

The proof of (3.24)⇒(3.21) follows by a simple use of Jensen's inequality, the converse being the content of Theorem 3.3.18 of [4]. The advantage of this formulation arises from the fact that some canonical semigroups (e.g. the quantum Ornstein–Uhlenbeck semigroup on $\mathbb{R}^n$) saturate the inequality, or equivalently:

$$[L,\, \nabla] = \kappa\nabla\,.$$

Therefore, the Ricci lower bound is equivalent to comparing the commutation of a semigroup with the gradient to the one of a canonical semigroup. A similar reasoning recently lead [13] to formulate a Bakry–Émery condition for birth and death processes on $\mathbb{N}$ in terms of a comparison to the Poisson process. Going back to our non-commutative setting, [7] showed that the quantum Ornstein–Uhlenbeck semigroup, as well as its fermionic version on the Clifford algebra, does satisfy such a commutation relation. They used this fact to derive the modified logarithmic Sobolev constant for these QMS via the contraction (3.23). In the next proposition, we recall their argument:

Proposition 5. *Assume that the following equalities hold: there exists* $\kappa \in \mathbb{R}$ *such that, for any* $j \in \mathcal{J}$ *and any* $t \geq 0$,

$$\partial_j \circ \Lambda_t = \mathrm{e}^{-\kappa t} \Lambda_t \circ \partial_j \,. \tag{3.25}$$

Then, $\mathrm{Ric}(\mathcal{L}) \geq \kappa$ *holds.*

Proof. From Proposition 4, it is enough to prove that (3.23) holds. Assume that $(\gamma(s))_{s\in[0,1]}$ is a minimal geodesic relating ρ to σ and denote by $(\mathbf{A}(s))_{s\in[0,1]}$ the unique solution of the continuity equation

$$\dot{\gamma}(s) = \mathrm{div}\, \mathbf{A}(s) \,.$$

By duality, Eq. (3.25) implies that $\Lambda_{t*}\dot{\gamma}(s) = \mathrm{e}^{-\kappa t}\, \mathrm{div}\, \vec{\Lambda}_{t*}\mathbf{A}(s)$, where $\vec{\Lambda}_{t*}\mathbf{A} := (\Lambda_{t*}A_j)_{j\in\mathcal{J}}$. Then, denoting $\gamma(s,t) := \Lambda_{t*}(\gamma(s))$,

$$\begin{aligned} g_{\mathcal{L},\gamma(s,t)}(\dot{\gamma}(s,t),\dot{\gamma}(s,t)) &= \mathrm{e}^{-2\kappa t} \sum_{j\in\mathcal{J}} \langle \mathcal{P}_{*t}(A_j(s)),\, [\gamma(s,t)]^{-1}_{\omega_j}(\mathcal{P}_{t*}(A_j(s)))\rangle \\ &\leq \mathrm{e}^{-2\kappa t} \sum_{j\in\mathcal{J}} \langle A_j(s),\, [\gamma(s)]^{-1}_{\omega_j} A_j(s))\rangle \\ &= \mathrm{e}^{-2\kappa t} g_{\mathcal{L},\gamma(s)}(\dot{\gamma}(s),\dot{\gamma}(s)) \,, \end{aligned}$$

where the inequality arises from the property of monotonicity of Fisher information metrics (cf. [17]). The result follows after taking the integral over the geodesic path. □

3.4. Example: The Quantum Depolarizing Semigroup

In this section, we derive a Ricci curvature lower bound on perhaps the simplest possible QMS: the depolarizing semigroup: define $(\Lambda_t^{\mathrm{dep}})_{t\geq 0}$ on $\mathcal{B}(\mathbb{C}^d)$ as follows,

$$\Lambda_t^{\mathrm{dep}}(X) = \mathrm{e}^{-t} X \, + (1 - \mathrm{e}^{-t}) \frac{1}{d} \mathrm{Tr}(X)\, \mathbb{I} \,.$$

Theorem 4. *The quantum depolarizing semigroup* $(\Lambda_t^{\mathrm{dep}})_{t\geq 0}$ *satisfies* $\mathrm{Ric}(\mathcal{L}^{\mathrm{dep}}) \geq \frac{1}{2}$.

Proof. In the Schrödinger picture, the generator $\mathcal{L}_*^{\mathrm{dep}}$ can be written as

$$\mathcal{L}_*^{\mathrm{dep}}(\rho) = \frac{\mathbb{I}}{d} \mathrm{Tr}(\rho) - \rho = \frac{1}{d^2} \sum_{j=1}^{d^2} (U_j\, \rho\, U_j - \rho), \;\; \rho \in \mathcal{D}(\mathbb{C}^d),$$

where the operators U_j can be chosen to be self-adjoint (e.g. generalized Pauli matrices [32]). In this case, $c_j = \frac{1}{2d^2}$, $\tilde{L}_j = U_j$ and $\omega_j = 0$, $j = 1, \ldots, d^2$. Now,

given a vector $\vec{V} = (V_1, \dots, V_{d^2}) \in \bigoplus_j \mathcal{B}(\mathbb{C}^d)$ and $\rho \in \mathcal{D}_+(\mathbb{C}^d)$,

$$\begin{aligned}\langle \vec{V}, \vec{V} \rangle_{\mathcal{L}^{\mathrm{dep}}, \rho} &= \frac{1}{2d^2} \sum_{j=1}^{d^2} \langle V_j, [\rho]_0(V_j) \rangle \\ &= \frac{1}{2d^2} \sum_{j=1}^{d^2} \int_0^1 \langle V_j, \rho^s V_j \rho^{1-s} \rangle \, \mathrm{d}s \\ &= \frac{1}{2d^2} \sum_{j=1}^{d^2} \int_0^1 \mathrm{Tr}(V_j^* \rho^s V_j \rho^{1-s}) \, \mathrm{d}s\end{aligned}$$

Now, given $U \in \mathcal{B}_{sa}(\mathbb{C}^d)$ and $\rho \in \mathcal{D}_+(\mathbb{C}^d)$, the following holds:

$$\begin{aligned}&\nabla U._\rho \nabla U \\ &= -\frac{1}{2d^2} \sum_{j=1}^{d^2} \int_0^1 \int_0^1 \int_0^\alpha \frac{\rho^\beta}{(1-s)\mathbb{I} + s\rho} \, \partial_j U \, \rho^{1-\alpha} \partial_j U \frac{\rho^{\alpha-\beta}}{(1-s)\mathbb{I} + s\rho} \mathrm{d}\beta \mathrm{d}s \mathrm{d}\alpha \\ &\quad + \int_0^1 \int_0^1 \int_0^{1-\alpha} \frac{\rho^\beta}{(1-s)\mathbb{I} + s\rho} \, \partial_j U \, \rho^{\alpha} \partial_j U \frac{\rho^{1-\alpha-\beta}}{(1-s)\mathbb{I} + s\rho} \mathrm{d}\beta \mathrm{d}s \mathrm{d}\alpha,\end{aligned}$$

where we used that, since $U_j = U_j^*$, $(\nabla U)^* = -\nabla U$. Then:

$$\begin{aligned}&\frac{1}{2} \langle \nabla U._\rho \nabla U, \, \mathcal{L}_*^{\mathrm{dep}}(\rho) \rangle \\ &\quad = -\frac{1}{4d^2} \sum_{j=1}^{d^2} \int_0^1 \int_0^1 \mathrm{d}s \mathrm{d}\alpha \left\{ \int_0^\alpha \mathrm{Tr} \left(\frac{\mathbb{I}}{d} - \rho \right) \right. \\ &\qquad \frac{\rho^\beta}{(1-s)\mathbb{I} + s\rho} \, \partial_j U \, \rho^{1-\alpha} \partial_j U \frac{\rho^{\alpha-\beta}}{(1-s)\mathbb{I} + s\rho} \mathrm{d}\beta \\ &\qquad \left. + \int_0^{1-\alpha} \mathrm{Tr} \left(\frac{\mathbb{I}}{d} - \rho \right) \frac{\rho^\beta}{(1-s)\mathbb{I} + s\rho} \, \partial_j U \, \rho^{\alpha} \partial_j U \frac{\rho^{1-\alpha-\beta}}{(1-s)\mathbb{I} + s\rho} \mathrm{d}\beta \right\}.\end{aligned}$$

By cyclicity of the trace, and since $\int_0^1 \frac{1}{((1-s)\mathbb{I} + s\rho)^2} \mathrm{d}s = \rho^{-1}$, forgetting about the positive contributions coming from the terms in $d^{-1}\mathbb{I}$, the above expression can be lower bounded as follows:

$$\begin{aligned}&\frac{1}{2} \langle \nabla U._\rho \nabla U, \, \mathcal{L}_*^{\mathrm{dep}}(\rho) \rangle \\ &\quad \geq \frac{1}{4d^2} \sum_{j=1}^{d^2} \int_0^1 \left\{ \int_0^\alpha \mathrm{Tr}(\rho^\beta \partial_j U \rho^{1-\alpha} \partial_j U \rho^{\alpha-\beta}) \mathrm{d}\beta \right. \\ &\qquad \left. + \int_0^{1-\alpha} \mathrm{Tr}(\rho^\beta \partial_j U \rho^{\alpha} \partial_j U \rho^{1-\alpha-\beta}) \mathrm{d}\beta \right\} \mathrm{d}\alpha\end{aligned}$$

$$= \frac{1}{4d^2}\sum_{j=1}^{d^2}\int_0^1 \left\{\alpha\,\mathrm{Tr}(\partial_j U\rho^{1-\alpha}\partial_j U\rho^{\alpha}) + (1-\alpha)\,\mathrm{Tr}(\partial_j U\,\rho^{\alpha}\partial_j U\rho^{1-\alpha})\right\}\,\mathrm{d}\alpha$$

$$= \frac{1}{4d^2}\sum_{j=1}^{d^2}\int_0^1 \mathrm{Tr}(\partial_j U\,\rho^{\alpha}\partial_j U\rho^{1-\alpha})\,\mathrm{d}\alpha.$$

On the other hand,

$$\langle \nabla\mathcal{L}^{\mathrm{dep}}(U),\,\nabla U\rangle_{\mathcal{L}^{\mathrm{dep}},\rho} = \frac{1}{2d^2}\sum_{j=1}^{d^2}\int_0^1 \mathrm{Tr}\left[\partial_j\left(\frac{\mathbb{I}}{d}\mathrm{Tr}(U)-U\right)^{*}\rho^s\,\partial_j U\,\rho^{1-s}\right]\mathrm{d}s$$

$$= \frac{1}{2d^2}\sum_{j=1}^{d^2}\int_0^1 \mathrm{Tr}[\,\partial_j U\,\rho^s\partial_j U\,\rho^{1-s}\,]\,\mathrm{d}s$$

Therefore,

$$B(\rho,U) \geq -\frac{1}{4d^2}\sum_{j=1}^{d^2}\int_0^1 \mathrm{Tr}[\,\partial_j U\,\rho^s\partial_j U\,\rho^{1-s}\,]\,\mathrm{d}s = \frac{1}{2}\|\nabla U\|^2_{\mathcal{L}^{\mathrm{dep}},\rho},$$

and the result follows. □

Remark 1. This result was independently found in [8].

4. A Quantum HWI Inequality

In [7] it was proved that, in the case when $\kappa > 0$, $\mathrm{Ric}(\mathcal{L}) \geq \kappa$ implies $\mathrm{MLSI}(\alpha_1)$ for $\kappa = \alpha_1$. This is, for example, the case of the classical and quantum Ornstein–Uhlenbeck processes. Here, we study the case of $\kappa \in \mathbb{R}$. In [12], the authors proved that, in the classical discrete framework, $\mathrm{Ric}(\mathcal{L}) \geq \kappa$ for $\kappa \in \mathbb{R}$ implies an HWI-like inequality (see Theorem 7.3). Here, we provide a quantum generalization of their result.

Theorem 5. *Assume that* $\mathrm{Ric}(\mathcal{L}) \geq \kappa$, *for some* $\kappa \in \mathbb{R}$. *Then* $\mathcal{L}$ *satisfies the following inequality*

$$\forall \rho \in \mathcal{D}_+(\mathcal{H}),\;\; D(\rho\|\sigma) \leq W_{2,\mathcal{L}}(\rho,\sigma)\sqrt{\mathrm{I}_\sigma(\rho)} - \frac{\kappa}{2}W_{2,\mathcal{L}}(\rho,\sigma)^2. \qquad (\mathrm{HWI}(\kappa))$$

Proof. By Theorem 3, for any $\rho,\omega \in \mathcal{D}_+(\mathcal{H})$

$$\frac{1}{2}\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0^+} W_{2,\mathcal{L}}(\rho_t,\omega)^2 + \frac{\kappa}{2}W_{2,\mathcal{L}}(\rho,\omega)^2 \leq D(\omega\|\sigma) - D(\rho\|\sigma).$$

Taking $\omega := \sigma$, this implies that

$$D(\rho\|\sigma) \leq -\frac{1}{2}\left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0^+} W_{2,\mathcal{L}}(\rho_t\|\sigma)^2 - \frac{\kappa}{2}W_{2,\mathcal{L}}(\rho,\sigma)^2. \qquad (4.1)$$

Then,

$$\begin{aligned}
-\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0^+} W_{2,\mathcal{L}}(\rho_t,\sigma)^2 &= \liminf_{s\to 0^+}\frac{1}{2s}(W_{2,\mathcal{L}}(\rho,\sigma)^2 - W_{2,\mathcal{L}}(\rho_s,\sigma)^2)\\
&\leq \limsup_{s\to 0^+}\frac{1}{2s}(W_{2,\mathcal{L}}(\rho,\rho_s)^2 + 2W_{2,\mathcal{L}}(\rho,\rho_s)W_{2,\mathcal{L}}(\rho_s,\sigma))\\
&\leq \limsup_{s\to 0^+}\frac{1}{2s}W_{2,\mathcal{L}}(\rho,\rho_s)^2 + W_{2,\mathcal{L}}(\rho,\sigma)\sqrt{\mathrm{I}_\sigma(\rho)}\\
&= W_{2,\mathcal{L}}(\rho,\sigma)\sqrt{\mathrm{I}_\sigma(\rho)}.
\end{aligned}$$

where the second inequality follows from Lemma 7 of [24]. The result follows from inserting this back into (4.1). □

In the case when $\kappa > 0$, we recover the result of [7]:

Corollary 2 (Quantum Bakry–Émery theorem). *Assume that* $\mathrm{Ric}(\mathcal{L}) \geq \kappa$, *for some* $\kappa > 0$. *Then* $\mathcal{L}$ *satisfies* $\mathrm{MLSI}(\kappa)$.

Proof. By Theorem 5, $\mathcal{L}$ satisfies HWI(κ). MLSI(κ) follows from an application of Young's inequality:

$$xy \leq cx^2 + \frac{1}{4c}y^2,\ \forall x, y \in \mathbb{R}, c > 0, \tag{4.2}$$

in which we set $x = W_{2,\mathcal{L}}(\rho,\sigma)$, $y = \sqrt{\mathrm{I}_\sigma(\rho)}$, and $c = \frac{\kappa}{2}$. □

In the case $\mathrm{Ric}(\mathcal{L}) \geq \kappa$ for $\kappa \in \mathbb{R}$, HWI(κ) still implies a modified log-Sobolev inequality under the further condition that a transportation cost inequality holds. This is a direct quantum generalization of Theorem 7.8 of [12] (see also Corollary 3.1 of [23])

Corollary 3. *Assume that* $\mathrm{Ric}(\mathcal{L}) \geq \kappa$, $\kappa \in \mathbb{R}$, *and that* $\mathrm{TC}_2(c_2)$ *holds with* $c_2^{-1} \geq \max(0, -\kappa)$. *Then* $\mathrm{MLSI}(\alpha_1)$ *holds for*

$$\alpha_1 = \max\left[\frac{1}{4c_2}\left(1 + c_2\kappa\right)^2, \kappa\right]$$

Proof. The proof is identical to the one of Corollary 3.1 of [23]. □

Similarly, we can show that $\mathrm{Ric}(\mathcal{L}) \geq \kappa$ for $\kappa \in \mathbb{R}$ implies MLSI as long as MLSI+TC_2 holds.

Corollary 4. *Assume that* $\mathrm{Ric}(\mathcal{L}) \geq \kappa$, $\kappa \in \mathbb{R}$, *and that the inequality* MLSI $+\,\mathrm{TC}_2(c)$ *(defined in Sect. 2.4) holds with* $c^{-1} \geq \max(\kappa, 0)$, *then* $\mathrm{MLSI}(\alpha_1)$ *holds, with*

$$\alpha_1 = \frac{1}{c\,(2 - \kappa c)}.$$

Proof. See Corollary 3.2 of [23]. □

The *diameter* of $\mathcal{D}(\mathcal{H})$ in the Wasserstein distance $W_{2,\mathcal{L}}$ is defined as follows:

$$\mathrm{Diam}_{\mathcal{L}}(\mathcal{D}(\mathcal{H})) = \sup_{\rho,\omega\in\mathcal{D}(\mathcal{H})} W_{2,\mathcal{L}}(\rho,\omega).$$

Another straightforward consequence of the κ-displacement convexity of the quantum relative entropy for $\kappa > 0$ is the following estimate on the diameter $\mathrm{Diam}_{\mathcal{L}}(\mathcal{D}(\mathcal{H}))$, which is a quantum analogue of the Bonnet–Myers theorem (see Proposition 7.3 of [11]).

Proposition 6. *Assume that* $\mathrm{Ric}(\mathcal{L}) \geq \kappa$ *holds for* $\kappa > 0$. *Then for any two states* $\rho, \omega \in \mathcal{D}(\mathcal{H})$,

$$W_{2,\mathcal{L}}(\rho,\omega)^2 \leq \frac{4}{\kappa}(D(\rho\|\sigma) + D(\omega\|\sigma)).$$

Therefore,

$$\mathrm{Diam}_{\mathcal{L}}(\mathcal{D}(\mathcal{H})) \leq \sup_{\rho\in\mathcal{D}(\mathcal{H})} \sqrt{\frac{8}{\kappa}D(\rho\|\sigma)}.$$

Proof. The result follows directly from the convexity of the quantum relative entropy (cf. (v) of Theorem 3):

$$0 \leq D(\gamma(1/2)\|\sigma) \leq \frac{1}{2}D(\rho\|\sigma) + \frac{1}{2}D(\omega\|\sigma) - \frac{\kappa}{8}W_{2,\mathcal{L}}(\rho,\omega)^2.$$

for a given constant speed geodesic $(\gamma(s))_{s\in[0,1]}$ relating ρ and ω. □

5. From Ricci Lower Bound to the Poincaré Inequality

In this section, we show that $\mathrm{Ric}(\mathcal{L}) \geq 0$ together with a condition of finiteness of the diameter of $\mathcal{D}(\mathcal{H})$ with respect to the distance $W_{2,\mathcal{L}}$ implies the Poincaré inequality, hence extending Proposition 5.9 of [11] to our non-commutative setting. Throughout this section, we fix $(\mathcal{P}_t)_{t\geq 0}$ to be a primitive QMS on $\mathcal{B}(\mathcal{H})$, $\mathcal{H}$ finite dimensional, with unique invariant state σ and associated generator $\mathcal{L}$, satisfying σ-DBC. The next result is a non-commutative extension of the fourth equivalent statement in Theorem 4.7.2 of [4]:

Proposition 7. (Reverse quantum Poincaré inequality) *Assume that* $\mathrm{Ric}(\mathcal{L}) \geq \kappa$ *holds. Then for any* $\rho \in \mathcal{D}_+(\mathcal{H})$, *any* $U \in \mathcal{B}_{sa}(\mathcal{H})$ *and all* $t > 0$*:*

$$\mathrm{Tr}(\Lambda_{*t}(\rho)U^2) - \mathrm{Tr}(\rho(\Lambda_t(U))^2) \geq \frac{\mathrm{e}^{2\kappa t} - 1}{2\kappa}\|\nabla\Lambda_t(U)\|^2_{\mathcal{L},\rho} \tag{5.1}$$

Proof. The proof is similar to the one of Theorem 3.5 of [11]. For $u \geq 0$, let $\rho_u \equiv \Lambda_{*u}(\rho)$ and $U_u \equiv \Lambda_u(U)$. Then, from Proposition 3,

$$
\begin{aligned}
&2\mathrm{e}^{2\kappa s}\|\nabla U_t\|^2_{\mathcal{L},\rho} \\
&\quad = 2\mathrm{e}^{2\kappa s}\|\nabla(\Lambda_s U_{t-s})\|^2_{\mathcal{L},\rho} \\
&\quad \leq 2\|\nabla U_{t-s}\|^2_{\mathcal{L},\rho_s} \\
&\quad = 2\sum_{j\in\mathcal{J}} c_j \langle \partial_j U_{t-s}, [\rho_s]_{\omega_j}\partial_j U_{t-s}\rangle \\
&\quad \leq \sum_{j\in\mathcal{J}} c_j \langle \partial_j U_{t-s}, (\mathrm{e}^{-\omega_j/2}R_{\rho_s} + \mathrm{e}^{\omega_j/2}L_{\rho_s})\partial_j U_{t-s}\rangle \\
&\quad = \sum_{j\in\mathcal{J}} c_j \left(\mathrm{e}^{\omega_j/2}\,\mathrm{Tr}[\rho_s \partial_j U_{t-s}(\partial_j U_{t-s})^*] + \mathrm{e}^{-\omega_j/2}\,\mathrm{Tr}[\rho_s(\partial_j U_{t-s})^*\partial_j U_{t-s}]\right) \\
&\quad = \sum_{j\in\mathcal{J}} c_j \left(\mathrm{e}^{\omega_j/2}\,\mathrm{Tr}(\rho_s[\tilde{L}_j, U_{t-s}][U_{t-s}, \tilde{L}_j^*]) + \mathrm{e}^{-\omega_j/2}\,\mathrm{Tr}(\rho_s[U_{t-s}, \tilde{L}_j^*][\tilde{L}_j, U_{t-s}])\right) \\
&\quad = \sum_{j\in\mathcal{J}} c_j \Big(\mathrm{Tr}(\rho_s(\tilde{L}_j U_{t-s}^2 \tilde{L}_j^* \mathrm{e}^{\omega_j/2} + \mathrm{e}^{-\omega_j/2}\tilde{L}_j^* U_{t-s}^2 \tilde{L}_j)) \\
&\qquad +\mathrm{e}^{\omega_j/2}\,\mathrm{Tr}(\rho_s(-\tilde{L}_j U_{t-s}\tilde{L}_j^* U_{t-s} - U_{t-s}\tilde{L}_j U_{t-s}\tilde{L}_j^* + U_{t-s}\tilde{L}_j\tilde{L}_j^* U_{t-s})) \\
&\qquad +\mathrm{e}^{-\omega_j/2}\,\mathrm{Tr}(\rho_s(U_{t-s}\tilde{L}_j^*\tilde{L}_j U_{t-s} - \tilde{L}_j^* U_{t-s}\tilde{L}_j U_{t-s} - U_{t-s}\tilde{L}_j^* U_{t-s}\tilde{L}_j))\Big) \\
&\quad = \mathrm{Tr}(\rho_s \mathcal{L}(U_{t-s}^2)) \\
&\qquad + \sum_{j\in\mathcal{J}} c_j \Big(\mathrm{e}^{\omega_j/2}\,\mathrm{Tr}(\rho_s(-\tilde{L}_j U_{t-s}\tilde{L}_j^* U_{t-s} - U_{t-s}\tilde{L}_j U_{t-s}\tilde{L}_j^* \\
&\qquad +U_{t-s}\tilde{L}_j\tilde{L}_j^* U_{t-s} + U_{t-s}^2\tilde{L}_j\tilde{L}_j^*)) \\
&\qquad + \mathrm{e}^{-\omega_j/2}\,\mathrm{Tr}(\rho_s(U_{t-s}\tilde{L}_j^*\tilde{L}_j U_{t-s} - \tilde{L}_j^* U_{t-s}\tilde{L}_j U_{t-s} \\
&\qquad -U_{t-s}\tilde{L}_j^* U_{t-s}\tilde{L}_j + \tilde{L}_j^*\tilde{L}_j U_{t-s}^2))\Big) \\
&\quad = \mathrm{Tr}(\rho_s \mathcal{L}(U_{t-s}^2)) - \mathrm{Tr}(\rho_s U_{t-s}\mathcal{L}(U_{t-s})) - \mathrm{Tr}(\rho_s \mathcal{L}(U_{t-s})U_{t-s}) \\
&\quad = \frac{\partial}{\partial s}\,\mathrm{Tr}(\rho_s U_{t-s}^2)
\end{aligned}
$$

where we used (2.38) of [24], with $R_\rho(A) \equiv A\rho$ and $L_\rho(A) \equiv \rho A$, in the fourth line. The claim follows after integrating the above inequality from 0 to t. □

Theorem 6. $\mathrm{Ric}(\mathcal{L}) \geq 0 \ + \ \mathrm{Diam}_{\mathcal{L}}(\mathcal{D}(\mathcal{H})) \leq D \Rightarrow \mathrm{PI}(\frac{1}{\mathrm{e}D^2})$.

Proof. Let $f \in \mathcal{B}_{sa}(\mathcal{H})$ be an eigenvector of $\mathcal{L}$ with associated eigenvalue opposite to the spectral gap λ of $\mathcal{L}$. Without loss of generality, $\|f\|_\infty = 1$, and by primitivity of $(\Lambda_t)_{t\geq 0}$, $\mathrm{Tr}(\sigma f) = 0$. Now, note that $\Lambda_t(f) = \mathrm{e}^{-\lambda t} f$. Therefore, the reverse Poincaré inequality (5.1) in the case when $\kappa = 0$ implies that for any $\rho \in \mathcal{D}_+(\mathcal{H})$,

$$\|\nabla f\|^2_{\mathcal{L},\rho} \leq \frac{\mathrm{e}^{2\lambda t}}{2t}\|f\|^2_\infty.$$

Optimizing over t and using $\|f\|_\infty = 1$, we find

$$\|\nabla f\|^2_{\mathcal{L},\rho} \leq \mathrm{e}\lambda \|f\|^2_\infty.$$

Given the following spectral decomposition of $f = \sum_\mu \mu P_\mu$, since $\mathrm{Tr}(\sigma f) = 0$, the minimum and maximum eigenvalues of f, respectively, denoted by $\mu_{\min}$ and $\mu_{\max}$, obey $\mu_{\min} < 0 < \mu_{\max}$. Since we assumed $\|f\|_\infty = 1$, this implies that given a path $(\gamma(s), U(s))_{s\in[0,1]}$ in $\mathcal{D}(\mathcal{H})$ joining the states $\gamma(0) = \frac{P_{\mu_{\max}}}{\mathrm{Tr}(P_{\mu_{\max}})}$ and $\gamma(1) = \frac{P_{\mu_{\min}}}{\mathrm{Tr}(P_{\mu_{\min}})}$ such that $\int_0^1 \|\dot\gamma(s)\|^2_{g_{\mathcal{L},\gamma(s)}} \mathrm{d}s \leq W_{2,\mathcal{L}}(\gamma(0),\gamma(1))^2 + \varepsilon$,

$$\begin{aligned}
1 \leq |\mu_{\max} - \mu_{\min}| &= \left| \mathrm{Tr}\, f \left(\frac{P_{\mu_{\min}}}{\mathrm{Tr}(P_{\mu_{\min}})} - \frac{P_{\mu_{\max}}}{\mathrm{Tr}(P_{\mu_{\max}})} \right) \right| \\
&= \left| \mathrm{Tr} \left(f \int_0^1 \dot\gamma(s) \mathrm{d}s \right) \right| \\
&= \left| \int_0^1 \sum_{j\in\mathcal{J}} c_j \langle \partial_j f, [\gamma(s)]_{\omega_j} \partial_j U(s) \rangle \mathrm{d}s \right| \\
&\leq \sqrt{(D^2+\varepsilon)} \left(\int_0^1 \|\nabla f\|^2_{\mathcal{L},\gamma(s)} \mathrm{d}s \right)^{1/2} \leq \sqrt{(D^2+\varepsilon)\lambda \mathrm{e}},
\end{aligned}$$

where in the last line we used the Cauchy–Schwarz inequality with respect to the inner product $\sum_{j\in\mathcal{J}} c_j \langle .\ \ , \int_0^1 [\gamma(s)]_{\omega_j} \mathrm{d}s\ \ . \rangle$, and the result directly follows. □

6. From Ricci Lower Bound to Modified Log-Sobolev Inequality

In [11], a modified logarithmic Sobolev inequality was proved to hold under the conditions that $\mathrm{Ric}(\mathcal{L}) \geq 0$ and that the diameter of the underlying space, in terms of the modified Wasserstein distance, is bounded. Here, we extend their results to the quantum regime under the further assumption that the semigroup $(\Lambda_t)_{t\geq 0}$ is unital, leaving the study of the general case to later. The idea of the proof is to get a non-tight logarithmic Sobolev inequality from HWI(0) and then to tighten it using ideas borrowed from [5]. In what follows, we denote by d the dimension of $\mathcal{H}$.

Given two states $\rho, \omega \in \mathcal{D}(\mathcal{H})$, with associated spectral decompositions $\rho = \sum_{i\in\mathcal{A}} \lambda_i P_i$, $\omega = \sum_{j\in\mathcal{B}} \mu_j Q_j$, where $\mathcal{A}$ and $\mathcal{B}$ are two finite index sets, a *coupling* of ρ and ω is a probability distribution q on $\mathcal{A}\times\mathcal{B}$ such that

$$\begin{aligned}
\sum_{i\in\mathcal{A}} q(i,j) &= \mu_j \,\mathrm{Tr}(Q_j) \\
\sum_{j\in\mathcal{B}} q(i,j) &= \lambda_i \,\mathrm{Tr}(P_i).
\end{aligned}$$

The set of couplings between ρ and ω is denoted by $\Pi(\rho,\omega)$. In analogy with the classical literature (see, for example, [12]), given an ergodic semigroup $(\Lambda_t)_{t\geq 0}$

with associated generator $\mathcal{L}$, the *coupling Wasserstein distance of order two* between ρ and ω is defined as follows:

$$W_{2,\mathcal{L},c}(\rho,\omega)^2 := \inf_{q\in\Pi(\rho,\sigma)} \sum_{i\in\mathcal{A},\ j\in\mathcal{B}} q(i,j) W_{2,\mathcal{L}}(\rho_i,\omega_j)^2,$$

where

$$\rho_i := \frac{P_i}{\operatorname{Tr} P_i},\ \omega_j := \frac{Q_j}{\operatorname{Tr}(Q_j)},\ i\in\mathcal{A},\ j\in\mathcal{B}.$$

The following result is a quantum generalization of Proposition 2.14 of [12]:

Proposition 8. *Let $(\Lambda_t)_{t\geq 0}$ be a primitive QMS, with unique invariant state σ and associated generator $\mathcal{L}$, satisfying the detailed balance condition. Then, for any $\rho,\omega\in\mathcal{D}_+(\mathcal{H})$,*

$$W_{2,\mathcal{L}}(\rho,\omega)\leq W_{2,\mathcal{L},c}(\rho,\omega).$$

Proof. Let $\rho=\sum_{i\in\mathcal{A}}\lambda_i P_i$, $\omega=\sum_{j\in\mathcal{B}}\mu_j Q_j$ the spectral decompositions of the states ρ and ω. For $(i,j)\in\mathcal{A}\times\mathcal{B}$, define $\rho_i := \frac{P_i}{\operatorname{Tr}(P_i)}$, $\omega_j := \frac{Q_j}{\operatorname{Tr}(Q_j)}$, and let $\varepsilon>0$. By definition of the Wasserstein distance $W_{2,\mathcal{L}}$, there exists a curve $\gamma_{ij}:[0,1]\mapsto\mathcal{D}(\mathcal{H})$ from ρ_i to ω_j such that

$$\int_0^1 \|\dot\gamma_{ij}(s)\|^2_{g_{\mathcal{L},\gamma_{ij}(s)}}\mathrm{d}s \leq W_{2,\mathcal{L}}(\rho_i,\omega_j)^2+\varepsilon.$$

For any coupling $q:\mathcal{A}\times\mathcal{B}\to\mathbb{R}_+$ of the states ρ and ω, define the path $(\gamma(s))_{s\in[0,1]}$ on $\mathcal{D}(\mathcal{H})$ as

$$\gamma(s)=\sum_{i\in\mathcal{A},\ j\in\mathcal{B}} q(i,j)\,\gamma_{ij}(s).$$

Therefore, $\gamma(0)=\rho$ and $\gamma(1)=\omega$. Now,

$$\begin{aligned} W_{2,\mathcal{L}}(\rho,\omega)^2 &\leq \int_0^1 \|\dot\gamma(s)\|^2_{g_{\mathcal{L},\gamma(s)}}\mathrm{d}s \\ &\leq \sum_{i\in\mathcal{A},\ j\in\mathcal{B}} q(i,j)\int_0^1 \|\dot\gamma_{ij}(s)\|^2_{g_{\mathcal{L},\gamma(s)}}\mathrm{d}s \\ &\leq \sum_{i\in\mathcal{A},\ j\in\mathcal{B}} q(i,j)\ W_{2,\mathcal{L}}(\rho_i,\omega_j)^2+\varepsilon. \end{aligned}$$

where we used the convexity of $g_{\mathcal{L}}$ in the second line (see equation (8.15) of [7]). As ε is arbitrary, the result follows after optimizing over the couplings q. □

In what follows, we restrict our analysis to the case of a primitive QMS $(\Lambda_t)_{t\geq 0}$ with unique invariant state $\mathbb{I}/d$ that satisfies the detailed balance condition. In order to prove the main result of this section, we need the following two lemmas that are extensions of Lemmas 6.2 and 6.3 of [11]:

Lemma 6. *Assume that* $\mathrm{Ric}(\mathcal{L}) \geq 0$ *and* $\mathrm{Diam}_{\mathcal{L}}(\mathcal{D}(\mathcal{H})) \leq D$. *Then for any* $\delta > 0$ *and* $f \in \mathcal{B}_{sa}(\mathcal{H})$ *such that* $\mathrm{Tr}(f^2) = d$:

$$D\left(f^2/d \| \mathbb{I}/d\right) \leq \delta D^2 \, \mathrm{I}_{\mathbb{I}/d}(f^2/d) + \frac{1}{4d\,\delta} \mathrm{Tr}(f^2 \mathbf{1}_{[1,\infty)}(f^2))$$

Proof. The case when f is not of full support is trivial, as then $\mathrm{I}_\sigma(f^2/d) = \infty$. Without loss of generality, we assume that f has full support, so that $f^2/d \in \mathcal{D}_+(\mathcal{H})$. Write $f = \sum_{i\in\mathcal{A}} \varphi(i) P_i$ the spectral decomposition of f, for some index set $\mathcal{A}$. From HWI(0), and Young's inequality (4.2) with $c = \delta D^2$, $x = \sqrt{\mathrm{I}_\sigma(f^2/d)}$, and $y = W_{2,\mathcal{L}}(f^2/d, \mathbb{I}/d)$:

$$D(f^2/d \| \mathbb{I}/d) \leq \delta D^2 \, \mathrm{I}_{\mathbb{I}/d}(f^2/d) + \frac{1}{4\delta D^2} W_{2,\mathcal{L}}(f^2/d, \mathbb{I}/d)^2.$$

From Proposition 8, for any coupling $q : \mathcal{A} \times \mathcal{B} \to \mathbb{R}_+$ between f^2/d and $\mathbb{I}/d$ such that $q(i,j) = 0$ for all $j \neq i$ whenever $\varphi(i)^2 \leq 1$,

$$\begin{aligned} D(f^2/d \| \mathbb{I}/d) &\leq \delta D^2 \, \mathrm{I}_{\mathbb{I}/d}(f^2/d) + \frac{1}{4\delta D^2} \\ &\qquad \sum_{i,j:\ \varphi(i)^2 > 1} q(i,j) W_{2,\mathcal{L}} \left(\frac{P_i}{\mathrm{Tr}(P_i)}, \frac{P_j}{\mathrm{Tr}(P_j)} \right)^2, \\ &\leq \delta D^2 \, \mathrm{I}_{\mathbb{I}/d}(f^2/d) + \frac{1}{4d\,\delta} \mathrm{Tr}(f^2 \mathbf{1}_{(1,\infty)}(f^2)). \end{aligned}$$

We recall that such a coupling q exists due to Strassen's theorem. □

Lemma 7. *For any* $A > 1$, *there exists* $\gamma > 0$ *such that for any* $f \in \mathcal{B}_{sa}(\mathcal{H})$ *with* $\mathrm{Tr}(f^2) = d$,

$$\frac{1}{d} \mathrm{Tr}(f^2 \mathbf{1}_{[A^2,\infty)}(f^2)) \leq \left(\frac{A}{A-1} \right)^2 \mathrm{Var}_{\mathbb{I}/d}(f), \tag{6.1}$$

$$D(f^2/d \| \mathbb{I}/d) \leq \gamma \, \mathrm{Var}_{\mathbb{I}/d}(f) + \frac{1}{d} \mathrm{Tr}(f^2 \log f^2 \mathbf{1}_{[A^2,\infty)}(f^2)). \tag{6.2}$$

Proof. This is a direct rewriting of Lemma 2.5 of [5]. □

Theorem 7. *Let* $(\Lambda_t)_{t\geq 0}$ *be a primitive semigroup with unique invariant state* $\mathbb{I}/d$ *and associated generator* $\mathcal{L}$. *Assume that* $\mathrm{Ric}(\mathcal{L}) \geq 0$ *and that* $\mathrm{Diam}_{\mathcal{L}}(\mathcal{D}(\mathcal{H})) \leq D$. *Then* $\mathrm{MLSI}(cD^{-2})$ *holds, for some universal constant* c.

Proof. Let $A > 1$ and $f \in \mathcal{B}_{sa}(\mathcal{H})$ of spectral decomposition $f = \sum_{i\in\mathcal{A}} \varphi(i) P_i$, with $\mathrm{Tr}(f^2) = d$. Without loss of generality, we can assume f positive definite. Then, set $f_A := f \vee A \equiv \sum_{i:\ \varphi(i)\geq A} \varphi(i) P_i + A \, \mathbf{1}_{(-\infty,A)}(f)$. Define the state $\rho_A = f_A^2 / \mathrm{Tr}(f_A^2)$. By (6.2),

$$D(f^2/d \| \mathbb{I}/d) \leq \gamma \, \mathrm{Var}_{\mathbb{I}/d}(f) + \frac{1}{d} \mathrm{Tr}(f^2 \log f^2 \mathbf{1}_{[A^2,\infty)}(f^2)). \tag{6.3}$$

By Theorem 6,

$$\gamma \operatorname{Var}_{\mathbb{I}/d}(f) \leq -2\mathrm{e}D^2\gamma\frac{1}{d}\langle f, \mathcal{L}(f)\rangle \leq \frac{\gamma \mathrm{e}D^2}{2}\operatorname{I}_{\mathbb{I}/d}(f^2/d), \tag{6.4}$$

where in the last inequality, we used the strong regularity of Dirichlet forms of unital semigroups (see [16]). Moreover,

$$\begin{aligned}\frac{1}{d}\operatorname{Tr}(f^2\log f^2\mathbf{1}_{[A^2,\infty)}(f^2)) &= \frac{1}{d}\operatorname{Tr}(f_A^2\log f_A^2) - \frac{1}{d}A^2\log A^2\operatorname{Tr}(\mathbf{1}_{(-\infty,A)}(f))\\ &\leq (1+A^2)D(\rho_A\|\mathbb{I}/d) + \frac{\operatorname{Tr}(f_A^2)}{d}\log\left(\frac{\operatorname{Tr} f_A^2}{d}\right) - \frac{A^2\log A^2}{d}\operatorname{Tr}(\mathbf{1}_{(-\infty,A)}(f)),\end{aligned} \tag{6.5}$$

where in the last line we used that $\frac{1}{d}\operatorname{Tr}(f_A^2) \leq 1 + A^2$. However, from Lemma 6 applied to f_A, since $\operatorname{I}_{\mathbb{I}/d}(\rho_A) \leq \frac{d}{\operatorname{Tr}(f_A^2)}\operatorname{I}_{\mathbb{I}/d}(f^2/d)$ by convexity of monotone Riemannian metrics (see, for example, equation (8.16) of [7]),

$$\begin{aligned}D(\rho_A\|\mathbb{I}/d) &\leq \frac{d\,\delta\,D^2}{\operatorname{Tr}(f_A^2)}\operatorname{I}_{\mathbb{I}/d}(f^2/d) + \frac{1}{4\operatorname{Tr}(f_A^2)\delta}\operatorname{Tr}(f_A^2\mathbf{1}_{[\frac{1}{d}\operatorname{Tr}(f_A^2),\infty)}(f_A^2))\\ &\leq \frac{\delta D^2}{A^2}\operatorname{I}_{\mathbb{I}/d}(f^2/d) + \frac{1}{4d\,\delta A^2}\operatorname{Tr}(f^2\mathbf{1}_{[A^2,\infty)}(f^2)),\end{aligned} \tag{6.6}$$

where in the last line we used that $A^2 \leq \frac{1}{d}\operatorname{Tr}(f_A^2)$. Using (6.1) and (6.4) together with Theorem 6,

$$\begin{aligned}D(\rho_A\|\mathbb{I}/d) &\leq \frac{\delta D^2}{A^2}\operatorname{I}_{\mathbb{I}/d}(f^2/d) + \frac{1}{4\delta(A-1)^2}\operatorname{Var}_{\mathbb{I}/d}(f)\\ &\leq \frac{\delta D^2}{A^2}\operatorname{I}_{\mathbb{I}/d}(f^2/d) - \frac{\mathrm{e}D^2}{2\delta(A-1)^2}\frac{1}{d}\langle f, \mathcal{L}(f)\rangle\\ &\leq \left(\frac{\delta D^2}{A^2} + \frac{\mathrm{e}D^2}{8\delta(A-1)^2}\right)\operatorname{I}_{\mathbb{I}/d}(f^2/d)\end{aligned}$$

Now,

$$\begin{aligned}&\frac{\operatorname{Tr}(f_A^2)}{d}\log\left(\frac{\operatorname{Tr}(f_A^2)}{d}\right) - \frac{A^2\log(A^2)}{d}\operatorname{Tr}(\mathbf{1}_{(-\infty,A)}(f))\\ &\leq \frac{1}{d}\Big(A^2\operatorname{Tr}(\mathbf{1}_{(-\infty,A)}(f)) + \operatorname{Tr}(f^2\mathbf{1}_{[A^2,\infty)}(f^2))\Big)\\ &\quad\log\left[\frac{1}{d}\Big(A^2\operatorname{Tr}(\mathbf{1}_{(-\infty,A)}(f)) + \operatorname{Tr}(f^2\mathbf{1}_{[A^2,\infty)}(f^2))\Big)\right]\\ &\quad - \frac{A^2\log(A^2)}{d}\operatorname{Tr}(\mathbf{1}_{(-\infty,A)}(f))\\ &= A^2\frac{1}{d}\operatorname{Tr}(\mathbf{1}_{(-\infty,A)}(f))\log\left(\frac{1}{d}\operatorname{Tr}(\mathbf{1}_{(-\infty,A)}(f)) + \frac{\operatorname{Tr}(f^2\mathbf{1}_{[A^2,\infty)}(f^2))}{dA^2}\right)\\ &\quad + \frac{1}{d}\operatorname{Tr}(f^2\mathbf{1}_{[A^2,\infty)}(f^2))\log\left(A^2\frac{1}{d}\operatorname{Tr}(\mathbf{1}_{(-\infty,A]}(f)) + \frac{1}{d}\operatorname{Tr}(f^2\mathbf{1}_{[A^2,\infty)}(f^2))\right)\end{aligned}$$

$$\leq A^2 \log\left(1 + \frac{\mathrm{Tr}(f^2 \mathbf{1}_{[A^2,\infty)}(f^2))}{dA^2}\right) + \frac{1}{d}\,\mathrm{Tr}(f^2 \mathbf{1}_{[A^2,\infty)}(f^2)) \log(1 + A^2)$$
$$\leq (1 + \log(1 + A^2))\frac{1}{d}\,\mathrm{Tr}(f^2 \mathbf{1}_{[A^2,\infty)}(f^2)),$$

where in the fourth line we used that $\frac{1}{d}\,\mathrm{Tr}(f^2) = 1$. Using once more (6.1) and (6.4) together with Theorem 6, we find

$$\frac{\mathrm{Tr}(f_A^2)}{d} \log\left(\frac{\mathrm{Tr}(f_A^2)}{d}\right) - \frac{A^2 \log(A^2)}{d}\,\mathrm{Tr}(\mathbf{1}_{(-\infty,A]}(f))$$
$$\leq \frac{\mathrm{e}D^2 A^2 (1 + \log(1 + A^2))}{2(1 - A)^2}\,\mathrm{I}_{\mathbb{I}/d}(f^2/d) \tag{6.7}$$

The result follows after combining (6.7), (6.3), (6.4), (6.5) and (6.6). □

7. Conclusion

In this paper, we prove that a classical picture, relating various inequalities which are useful in the analysis of Markov semigroups, carries over to the quantum setting. Classically, a key element of this picture is a geometric inequality called the Ricci lower bound. Functional and transportation cost inequalities, which play an important role in the study of mixing times of a primitive Markov semigroup and concentration properties of its invariant measure, can be obtained from this geometric inequality. The connection between them is provided by an interpolating inequality called the HWI inequality. In this paper, we analyse a quantum version of the Ricci lower bound (due to Carlen and Maas [7]) and show that it implies a quantum HWI inequality, from which quantum versions of the functional and transportation cost inequalities (which are relevant for the analysis of quantum Markov semigroups) follow.

Acknowledgements

Open Access funding provided by Projekt DEAL. CR acknowledges support by the DFG cluster of excellence 2111 (Munich Center for Quantum Science and Technology). The authors would like to thank Ivan Bardet for helpful discussions.

References

[1] Alicki, R.: On the detailed balance condition for non-Hamiltonian systems. Rep. Math. Phys. **10**(2), 249–258 (1976)

[2] Bakry, D.: L'hypercontractivité et son utilisation en théorie des semigroupes. In: Bernard, P. (ed.) Lectures on Probability Theory, pp. 1–114. Springer, Berlin (1992)

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

[4] Bakry, D., Gentil, I., Ledoux, M.: Analysis and Geometry of Markov Diffusion Operators. Springer, Berlin (2014)

[5] Barthe, F., Kolesnikov, A.V.: Mass transport and variants of the logarithmic sobolev inequality. J. Geom. Anal. **18**(4), 921–979 (2008)

[6] Carlen, E.A., Maas, J.: 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), 887–926 (2014)

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

[8] Carlen, E.A., Maas, J.: Non-commutative calculus, optimal transport and functional inequalities in dissipative quantum systems (2018). arXiv preprint arXiv:1811.04572

[9] Daneri, S., Savaré, G.: Eulerian calculus for the displacement convexity in the Wasserstein distance. SIAM J. Math. Anal. **40**(3), 1104–1122 (2008)

[10] Erbar, M.: The heat equation on manifolds as a gradient flow in the Wasserstein space. Ann. Inst. H. Poincaré Probab. Stat. **46**(1), 1–23 (2010). 02

[11] Erbar, M., Fathi, M.: Poincaré, modified logarithmic Sobolev and isoperimetric inequalities for Markov chains with non-negative Ricci curvature (2016). arXiv preprint arXiv:1612.00514

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

[13] Johnson, O.: A discrete log-Sobolev inequality under a Bakry–Émery type condition. Ann. l'Inst. Henri Poincaré Probab. Stat. **53**(4), 1952–1970 (2017). 11

[14] Jordan, R., Kinderlehrer, D., Otto, F.: The variational formulation of the Fokker–Planck equation. SIAM J. Math. Anal. **29**(1), 1–17 (1998)

[15] Kantorovich, L.: On the translocation of masses. Dokl. Akad. Nauk. USSR **133**, 199–201 (1942)

[16] Kastoryano, M.J., Temme, K.: Quantum logarithmic Sobolev inequalities and rapid mixing. J. Math. Phys. **54**(5), 052202 (2013)

[17] Lesniewski, A., Ruskai, M.B.: Monotone Riemannian metrics and relative entropy on noncommutative probability spaces. J. Math. Phys. **40**(11), 5702–5724 (1999)

[18] Lott, J., Villani, C.: Ricci curvature for metric-measure spaces via optimal transport. Ann. Math. **169**, 903–991 (2009)

[19] Maas, J.: Gradient flows of the entropy for finite Markov chains. J. Funct. Anal. **261**(8), 2250–2292 (2011)

[20] Marton, K.: A simple proof of the blowing-up lemma (corresp.). IEEE Trans. Inf. Theory **32**(3), 445–446 (1986)

[21] Monge, G.: Mémoire sur la théorie des déblais et des remblais. Hist. l'Acad. R. Sci. Paris, pp. 666 – 704 (1781)

[22] Olkiewicz, R., Zegarlinski, B.: Hypercontractivity in noncommutative Lp spaces. J. Funct. Anal. **161**(1), 246–285 (1999)

[23] Otto, F., Villani, C.: Generalization of an inequality by Talagrand and links with the logarithmic Sobolev inequality. J. Funct. Anal. **173**(2), 361–400 (2000)

[24] Rouzé, C., Datta, N.: Concentration of Quantum States from Quantum Functional and Talagrand Inequalities (2017). arXiv preprint arXiv:1704.02400

[25] Spohn, H.: Entropy production for quantum dynamical semigroups. J. Math. Phys. **19**(5), 1227–1230 (1978)

[26] Sturm, K.-T.: Transport inequalities, gradient estimates, entropy, and Ricci curvature. Commun. Pure Appl. Math. **LVIII**, 923–940 (2005)

[27] Sturm, K.-T.: On the geometry of metric measure spaces. Acta Math. **196**(1), 65–131 (2006)

[28] Sturm, K.-T.: On the geometry of metric measure spaces. II. Acta Math. **196**(1), 133–177 (2006)

[29] Temme, K., Kastoryano, M.J., Ruskai, M., Wolf, M.M., Verstraete, F.: The χ2-divergence and mixing times of quantum Markov processes. J. Math. Phys. **51**(12), 122201 (2010)

[30] Villani, C.: Optimal Transport: Old and New, vol. 338. Springer, Berlin (2008)

[31] von Renesse, M.-K., Sturm, K.-T.: Transport inequalities, gradient estimates, entropy and Ricci curvature. Commun. Pure Appl. Math. **58**(7), 923–940 (2005)

[32] Wolf, M.M.: Quantum Channels and Operations: Guided tour. Lecture notes, vol. 5 (2012). http://www-m5.ma.tum.de/foswiki/pubM. Accessed 13 Mar 2019

Nilanjana Datta
Statistical Laboratory, Centre for Mathematical Sciences
University of Cambridge
Cambridge CB 30WB
UK

and

DAMTP, Centre for Mathematical Sciences
University of Cambridge
Cambridge CB 30WA
UK

Cambyse Rouzé
Department of Mathematics
Technische Universität München
85748 Garching
Germany
e-mail: rouzecambyse@gmail.com

Communicated by David Pérez-García.
Received: July 9, 2019.
Accepted: January 21, 2020.