



# Non-commutative Calculus, Optimal Transport and Functional Inequalities in Dissipative Quantum Systems

Eric A. Carlen<sup>1</sup> · Jan Maas<sup>2</sup>

Received: 20 February 2019 / Accepted: 12 November 2019 / Published online: 27 November 2019  
© The Author(s) 2019

## Abstract

We study dynamical optimal transport metrics between density matrices associated to symmetric Dirichlet forms on finite-dimensional  $C^*$ -algebras. Our setting covers arbitrary skew-derivations and it provides a unified framework that simultaneously generalizes recently constructed transport metrics for Markov chains, Lindblad equations, and the Fermi Ornstein–Uhlenbeck semigroup. We develop a non-nommutative differential calculus that allows us to obtain non-commutative Ricci curvature bounds, logarithmic Sobolev inequalities, transport-entropy inequalities, and spectral gap estimates.

**Keywords** Non-commutative optimal transport · Functional inequalities · Lindblad equation · Gradient flow

## 1 Introduction

In the context of diffusion semigroups, a great deal of recent progress has been made based on two different gradient flow interpretations of the heat flow, namely as

- (1) The gradient flow of the Dirichlet energy in  $L^2$ ;
- (2) The gradient flow of the Boltzmann entropy in the space of probability measures endowed with the 2-Kantorovich metric.

In this paper we study the analogs of (1) and (2) for non-commutative probability, in the setting of Neumann algebras, and we establish the equivalence of (1) and (2) in this setting. This naturally involves the construction of non-commutative analogs of the 2-Kantorovich metric, a topic that was investigated in our earlier papers [8, 10] and in the independent work

---

Communicated by Alessandro Giuliani.

---

✉ Jan Maas  
jan.maas@ist.ac.at

Eric A. Carlen  
carlen@math.rutgers.edu

<sup>1</sup> Department of Mathematics, Hill Center, Rutgers University, 110 Frelinghuysen Road, Piscataway, NJ 08854-8019, USA

<sup>2</sup> Institute of Science and Technology Austria (IST Austria), Am Campus 1, 3400 Klosterneuburg, Austria

[34,36]. Recently the subject received the attention of a number of authors; see [11,12] for noncommutative transport metrics, [22,45,46] for functional inequalities, and [28,48] for results in infinite dimensions. We refer to [6,23] for different non-commutative variants of the 2-Kantorovich metric in other contexts.

Our focus in this paper is on developing the relations between (1) and (2) in the non-commutative setting with the aim of proving functional inequalities relevant to the study of the rate of approach to equilibrium for quantum Markov semigroups, in close analogy with what has been accomplished along these lines in the classical setting in recent years.

In order not to obscure the main ideas we shall work in a finite-dimensional setting and postpone the infinite-dimensional extension to a future work. The finite-dimensional case is of direct interest in quantum information theory, and the essential aspects of our new results are interesting even in this setting where they can be explained to a wider audience that is not thoroughly familiar with the Tomita–Takesaki theory. We now briefly describe the content of the paper. Any unfamiliar terminology is explained in the next subsection, but hopefully many readers will not need to look ahead.

The central object of study in this paper is a quantum Markov semigroup (QMS)  $(\mathcal{P}_t)_{t>0}$  on  $\mathcal{A}$ , a finite-dimensional  $C^*$ -algebra containing the identity  $\mathbf{1}$ . That is, for each  $t$ ,  $\mathcal{P}_t\mathbf{1} = \mathbf{1}$  and  $\mathcal{P}_t$  is completely positive. The generators  $\mathcal{L}$  of such semigroups have been characterized in [24,31].

We are concerned with the case in which there is a unique faithful invariant state  $\sigma$  for the dual semigroup; i.e.,  $\mathcal{P}_t^\dagger\sigma = \sigma$  for all  $t$ . The paper [47] is an excellent source for the physical context and makes it clear that assuming that the invariant state  $\sigma$  is tracial, which we do not do, would preclude a great many physical applications. Let  $\mathfrak{F}_+$  denote the space of faithful states. We would like to know, for instance, when there is a Riemannian metric  $g$  on  $\mathfrak{F}_+$  such that the flow on  $\mathfrak{F}_+$  given by the dual semigroup  $(\mathcal{P}_t^\dagger)_{t>0}$  is the gradient flow driven by the relative entropy functional  $\text{Ent}_\sigma(\rho) = \text{Tr}[\rho(\log \rho - \log \sigma)]$  with respect to the Riemannian metric. In [10,36], it is shown that when each  $\mathcal{P}_t$  is self-adjoint with respect to the Gelfand–Naimark–Segal (GNS) inner product induced on  $\mathcal{A}$  by  $\sigma$ , this is the case. We constructed the metric using ideas from optimal mass transport, and showed that, as in the classical case, the framework provided an efficient means for proving functional inequalities. This has been taken up and further developed by other authors, in particular Rouzé and Datta [45,46]. As in the classical case, Ricci curvature bounds are essential for the framework to be used to obtain functional inequalities. As shown in [10,46], once one has Ricci curvature bounds, a host of functional inequalities follow. A central problem then is to prove such bounds. A main contribution of the present paper is a flexible framework for doing this. It turns out that there are many ways to write a given QMS generator  $\mathcal{L}$  (that is self-adjoint in the GNS sense) in “divergence form” for non-commutative derivatives. Each of the different ways of doing this can be associated to a Riemannian metric on  $\mathfrak{F}_+$ . Different ways of writing  $\mathcal{L}$  in divergence form may have advantages over others, for example in proving Ricci curvature bounds. Hence it is important to have as much flexibility here as possible. We shall use this flexibility to give new examples in which we can obtain sharp Ricci curvature bounds. The machinery is useful for other functionals and other flows; the methods of this paper are not by any means restricted to gradient flow for relative entropy, despite our focus on this example here in the introduction.

An interesting problem remains: For each way of writing  $\mathcal{L}$  in divergence form, we have a Riemannian metric. The formulas are different, but in principle, all of the metrics might be the same. That is, they might all be determined by  $\mathcal{L}$ , and not the particular way of writing in divergence form, even though doing this one way or another may facilitate certain computations.

The problem of writing QMS as gradient flow for the relative entropy was also taken up independently by Mittnenzweig and Mielke [36], and although their framework is somewhat different, their approach also works in the case that each  $\mathcal{P}_t$  is self-adjoint with respect to the GNS inner product induced on  $\mathcal{A}$  by  $\sigma$ . Here, we shall show that if  $(\mathcal{P}_t)_{t \geq 0}$  can be written as gradient flow for  $\text{Ent}_\sigma$  with respect to some continuously differentiable Riemannian metric, then each  $\mathcal{P}_t$  is necessarily self-adjoint with respect to another inner product associated to  $\sigma$ , the Bogoliobov-Kubo-Mori (BKM) inner product. As we show, the class of QMS with this self-adjointness property is strictly larger than the class of QMS with the GNS self-adjointness property. Thus, there is at present an interesting gap between the known necessary condition for the construction of the Riemannian metric, and the known sufficient condition. Of course, in the classical setting, the two notions of self-adjointness coincide, and one has a pleasing characterization of reversible Markov chains in terms of gradient flow [15].

### 1.1 Notation

Let  $\mathcal{A}$  be finite-dimensional  $C^*$ -algebra containing the identity  $\mathbf{1}$ . In the finite-dimensional setting, all topologies one might impose on  $\mathcal{A}$  are equivalent, and  $\mathcal{A}$  is also a von Neumann algebra. In particular, it is generated by the projections it contains. We may regard any such algebra as a  $*$ -subalgebra of  $\mathbb{M}_n(\mathbb{C})$ , the set of all complex  $n \times n$  matrices. Let  $\mathcal{A}_h$  be the subset of hermitian elements in  $\mathcal{A}$ , and let  $\mathcal{A}_+ \subseteq \mathcal{A}$  denote the class of elements that are positive definite (i.e.,  $\text{sp}(A) \subseteq (0, \infty)$  for  $A \in \mathcal{A}_+$ ). For  $\mathcal{A} = \mathbb{M}_n(\mathbb{C})$  we write  $\mathcal{A}_+ = \mathbb{M}_n^+(\mathbb{C})$ .

Throughout this section we fix a positive linear functional  $\tau$  on  $\mathcal{A}$  that is *tracial* (i.e.,  $\tau[AB] = \tau[BA]$  for all  $A, B \in \mathcal{A}$ ) and *faithful* (i.e.,  $A = 0$  whenever  $\tau[A^*A] = 0$ ). Under these assumptions,  $\tau$  induces a scalar product on  $\mathcal{A}$  given by  $\langle A, B \rangle_{L^2(\mathcal{A}, \tau)} = \tau[A^*B]$  for  $A, B \in \mathcal{A}$ . In our applications,  $\tau$  will often be the usual trace  $\text{Tr}$  on  $\mathbb{M}_n(\mathbb{C})$  in which case the scalar product is the Hilbert–Schmidt scalar product, but it will be useful to include different situations, e.g., the trace induced by a non-uniform probability measure on a finite set.

A *state* on  $\mathcal{A}$  is a positive linear functional  $\varphi$  on  $\mathcal{A}$  such that  $\varphi(\mathbf{1}) = 1$ . If  $\varphi$  is a state, there is a uniquely determined  $\sigma \in \mathcal{A}$  such that  $\varphi(A) = \tau[\sigma A]$  for all  $A \in \mathcal{A}$ . Note that  $\sigma$  is a *density matrix*; i.e., it is positive semidefinite and  $\tau[\sigma] = 1$ . Let  $\mathfrak{P}(\mathcal{A})$  denote the set of density matrices. We write  $\mathfrak{P}_+(\mathcal{A}) = \{\rho \in \mathfrak{P}(\mathcal{A}) : \rho \text{ is positive definite}\}$ . We will simply write  $\mathfrak{P} = \mathfrak{P}(\mathcal{A})$  and  $\mathfrak{P}_+ = \mathfrak{P}_+(\mathcal{A})$  if the algebra  $\mathcal{A}$  is clear from the context.

We always use  $\dagger$  to denote the adjoint of a linear transformation on  $\mathcal{A}$  with respect to the scalar product  $\langle \cdot, \cdot \rangle_{L^2(\mathcal{A}, \tau)}$ . If  $\mathcal{K}$  is such a linear transformation,

$$\langle A, \mathcal{K}B \rangle_{L^2(\mathcal{A}, \tau)} = \langle \mathcal{K}^\dagger A, B \rangle_{L^2(\mathcal{A}, \tau)}. \tag{1.1}$$

Though we suppose no familiarity with the Tomita–Takesaki Theory of standard forms of von Neumann algebras, we will make use of the so-called modular and relative modular operators that arise there. In our setting, these operators have a simple direct definition:

**Definition 1.1** (*The relative modular operator*) Let  $\sigma, \rho \in \mathfrak{P}_+$ . The corresponding *relative modular operator*  $\Delta_{\sigma, \rho}$  is the linear transformation on  $\mathcal{A}$  defined by

$$\Delta_{\sigma, \rho}(A) = \sigma A \rho^{-1}. \tag{1.2}$$

The *modular operator* corresponding to  $\sigma$ ,  $\Delta_\sigma$ , is defined by  $\Delta_\sigma := \Delta_{\sigma, \sigma}$ .

Since  $\langle B, \Delta_{\sigma, \rho}A \rangle_{L^2(\mathcal{A}, \tau)} = \tau[(\sigma^{1/2}B\rho^{-1/2})^*(\sigma^{1/2}A\rho^{-1/2})]$  for all  $A, B \in \mathcal{A}$ , the operator  $\Delta_{\sigma, \rho}$  is positive definite on  $L^2(\mathcal{A}, \tau)$ . In case that  $\tau$  is the restriction of the usual trace

Tr to  $\mathcal{A} \subseteq \mathbb{M}_n(\mathbb{C})$ , the operators  $\sigma$  and  $\rho$  are also positive density matrices in  $\mathbb{M}_n(\mathbb{C})$ , and the same computations are valid for all  $A, B \in \mathbb{M}_n(\mathbb{C})$ . We may regard  $\Delta_\sigma$  as an operator on  $\mathbb{M}_n(\mathbb{C})$ , equipped with the Hilbert–Schmidt inner product, and then, so extended, it is still positive definite.

We are interested in evolution equations on  $\mathfrak{P}_+(\mathcal{A})$  that correspond to forward Kolmogorov equations for ergodic Markov processes satisfying a detailed balance condition, or in other words a reversibility condition, with respect to their unique invariant probability measure. Before presenting our results, we introduce the class of quantum Markov semigroups satisfying a detailed balance condition that are the focus of our investigation.

## 2 Quantum Markov Semigroups with Detailed Balance

Let  $\mathcal{A} \subseteq B(\mathcal{H})$  be a  $C^*$ -algebra of operators acting on a finite-dimensional Hilbert space  $\mathcal{H}$ . Let  $\tau$  be a tracial and faithful positive linear functional on  $\mathcal{A}$ . A *quantum Markov semigroup* on  $\mathcal{A}$  is a  $C_0$ -semigroup of operators  $(\mathcal{P}_t)_{t \geq 0}$  acting on  $\mathcal{A}$ , satisfying

- (1)  $\mathcal{P}_t \mathbf{1} = \mathbf{1}$ ;
- (2)  $\mathcal{P}_t$  is *completely positive*, i.e.,  $\mathcal{P}_t \otimes I_{\mathbb{M}_n(\mathbb{C})}$  is a positivity preserving operator on  $\mathcal{A} \otimes \mathbb{M}_n(\mathbb{C})$  for all  $n \in \mathbb{N}$ .

Note that (2) implies that  $\mathcal{P}_t$  is *real*, i.e.,  $(\mathcal{P}_t A)^* = \mathcal{P}_t A^*$  for all  $A \in \mathcal{A}$ . Let  $\mathcal{P}_t^\dagger$  be the Hilbert–Schmidt adjoint of  $\mathcal{P}_t$  satisfying  $\tau[A^* \mathcal{P}_t^\dagger B] = \tau[(\mathcal{P}_t A)^* B]$  for all  $A, B \in \mathcal{A}$ . It follows that  $\mathcal{P}_t^\dagger$  is trace-preserving and completely positive.

It is well known [24,31] that the generator  $\mathcal{L}$  of the semigroup  $\mathcal{P}_t = e^{t\mathcal{L}}$  can be written in *Lindblad form*

$$\mathcal{L}A = i[\tilde{H}, A] + \sum_{j \in \mathcal{J}} V_j^*[A, V_j] + [V_j^*, A]V_j, \tag{2.1}$$

$$\mathcal{L}^\dagger \rho = -i[\tilde{H}, \rho] + \sum_{j \in \mathcal{J}} [V_j, \rho V_j^*] + [V_j \rho, V_j^*], \tag{2.2}$$

where  $\mathcal{J}$  is a finite index set,  $V_j \in B(\mathcal{H})$  (not necessarily belonging to  $\mathcal{A}$ ) for all  $j \in \mathcal{J}$ , and the Hamiltonian  $\tilde{H} \in B(\mathcal{H})$  is self-adjoint.

### 2.1 Detailed Balance

The starting point of our investigations is the assumption that  $(\mathcal{P}_t)_{t \geq 0}$  satisfies the condition of *detailed balance*.

In the commutative setting, if  $P = (P_{ij})$  is the transition matrix of a Markov chain on  $\{1, \dots, n\}$  with invariant probability vector  $\sigma$ , we say that detailed balance holds if  $\sigma_i P_{ij} = \sigma_j P_{ji}$  for all  $i, j$ . An analytic way to formulate this condition is that  $P$  is self-adjoint with respect to the weighted inner product on  $\mathbb{C}^n$  given by  $\langle f, g \rangle_\sigma = \sum_{j=1}^n \sigma_j f_j g_j$ .

In the quantum setting, with a reference density matrix  $\sigma$  that is not a multiple of the identity, there are many candidates for such a weighted inner product. E.g., given  $\sigma \in \mathfrak{P}_+$ , and  $s \in [0, 1]$  one can define an inner product on  $\mathcal{A}$  by

$$\langle X, Y \rangle_s = \tau[X^* \sigma^s Y \sigma^{1-s}]. \tag{2.3}$$

Note that by cyclicity of the trace,  $\langle X, X \rangle_s = \tau[|\sigma^{s/2} X \sigma^{(1-s)/2}|^2] \geq 0$ , so that  $\langle \cdot, \cdot \rangle_s$  is indeed a positive definite sesquilinear form. The inner products for  $s = 0$  and  $s = \frac{1}{2}$  will come up

frequently in what follows, and they have their own names:  $\langle \cdot, \cdot \rangle_0$  is the Gelfand–Naimark–Segal inner product, denoted  $\langle \cdot, \cdot \rangle_{L^2_{\text{GNS}}(\sigma)}$ , and  $\langle \cdot, \cdot \rangle_{1/2}$  is the Kubo–Martin–Schwinger inner product, denoted  $\langle \cdot, \cdot \rangle_{L^2_{\text{KMS}}(\sigma)}$ . We shall write  $\mathcal{A} = L^2_{\text{GNS}}(\mathcal{A}, \sigma)$  (resp.  $\mathcal{A} = L^2_{\text{KMS}}(\mathcal{A}, \sigma)$ ) if we want to stress this Hilbert space structure.

Suppose, for some  $s \in [0, 1]$ , that  $\mathcal{P}_t$  is self-adjoint with respect to the  $\langle \cdot, \cdot \rangle_s$  inner product. Then, for all  $A \in \mathcal{A}$ ,

$$\begin{aligned} \tau[(\mathcal{P}_t^\dagger \sigma)A] &= \tau[\sigma \mathcal{P}_t A] = \tau[\sigma^{1-s} \mathbf{1} \sigma^s \mathcal{P}_t A] = \langle \mathbf{1}, \mathcal{P}_t A \rangle_s \\ &= \langle \mathcal{P}_t \mathbf{1}, A \rangle_s = \langle \mathbf{1}, A \rangle_s = \tau[\sigma A] . \end{aligned}$$

Hence for each of these inner products, self-adjointness of  $\mathcal{P}_t$  implies that  $\sigma$  is invariant under  $\mathcal{P}_t^\dagger$ .

The following lemma of Alicki [1] relates some of the possible definitions of detailed balance; a proof may be found in [10].

**Lemma 2.1** *Let  $\mathcal{K}$  be a real linear transformation on  $\mathcal{A}$ . If  $\mathcal{K}$  is self-adjoint with respect to the  $\langle \cdot, \cdot \rangle_s$  inner product for some  $s \in [0, 1/2) \cup (1/2, 1]$ , then  $\mathcal{K}$  commutes with  $\Delta_\sigma$ , and  $\mathcal{K}$  is self-adjoint with respect to  $\langle \cdot, \cdot \rangle_s$  for all  $s \in [0, 1]$ , including  $s = 1/2$ .*

As we have remarked, for a QMS  $(\mathcal{P}_t)_{t \geq 0}$ , each  $\mathcal{P}_t$  is real, and so  $\mathcal{P}_t$  is self-adjoint with respect to the GNS inner product if and only if it is self-adjoint with respect to the  $\langle \cdot, \cdot \rangle_s$  inner product for all  $s \in [0, 1]$ . However, if each  $\mathcal{P}_t$  is self-adjoint with respect to the KMS inner product, then it need not be self-adjoint with respect to the GNS inner product: There exist QMS for which each  $\mathcal{P}_t$  is self-adjoint with respect to the KMS inner product, but for which  $\mathcal{P}_t$  does not commute with  $\Delta_\sigma$ , and therefore cannot be self-adjoint with respect to the GNS inner product. A simple example is provided in appendix B of [10]. The generators of QMS such that  $\mathcal{P}_t$  is self-adjoint with respect to the KMS inner product have been investigated by Fagnola and Umanita [20]. However, there is a third notion of detailed balance that is natural in the present context, namely the requirement that each  $\mathcal{P}_t$  be self-adjoint with respect to the Bogoliubov–Kubo–Mori inner product:

**Definition 2.2** (*BKM inner product*) The *BKM inner product* is defined by

$$\langle A, B \rangle_{L^2_{\text{BKM}}(\sigma)} = \int_0^1 \langle A, B \rangle_s \, ds . \tag{2.4}$$

By what we have remarked above, if each  $\mathcal{P}_t$  is self-adjoint with respect to the GNS inner product, then each  $\mathcal{P}_t$  is self-adjoint with respect to the BKM inner product. However, as will be discussed at the end of this section, the converse is not in general true. The relevance of the BKM version of detailed balance is due to the following result that we show in Theorem 2.9: If the forward Kolmogorov equation for an ergodic QMS  $(\mathcal{P}_t)_{t \geq 0}$  with invariant state  $\sigma \in \mathfrak{P}_+$  is gradient flow for the quantum relative entropy  $\text{Ent}_\sigma(\rho) := \tau[\rho(\log \rho - \log \sigma)]$  with respect to some continuously differentiable Riemannian metric on  $\mathfrak{P}_+$ , then each  $\mathcal{P}_t$  is self-adjoint with respect to the BKM inner product. The BKM inner product is closely connected to the relative entropy functional, and for this reason it appears in some of the functional inequalities that we consider in Sect. 11.

On the other hand, only when each  $\mathcal{P}_t$  is self-adjoint with respect to the GNS inner product do we have a construction of such a Riemannian metric. The same is true for other constructions of Riemannian metrics on  $\mathfrak{P}_+$  for which QMS become gradient flow for  $\text{Ent}_\sigma(\rho)$ , in particular see [36]. Since most of this paper is concerned with our construction and its consequences, we make the following definition:

**Definition 2.3** (*Detailed balance*) Let  $\sigma \in \mathcal{A}$  be non-negative. We say that a quantum Markov semigroup  $(\mathcal{P}_t)_{t \geq 0}$  satisfies the *detailed balance condition* with respect to  $\sigma$  if for each  $t > 0$ ,  $\mathcal{P}_t$  is self-adjoint with respect to the GNS inner product on  $\mathcal{A}$  induced by  $\sigma$ , i.e.,

$$\tau[\sigma A^* \mathcal{P}_t B] = \tau[\sigma (\mathcal{P}_t A)^* B] \quad \text{for all } A, B \in \mathcal{A} .$$

We shall write that  $(\mathcal{P}_t)_t$  satisfies  $\sigma$ -DBC for brevity.

The following result gives the general form of the generator of quantum Markov semigroups on  $B(\mathcal{H})$  satisfying detailed balance. This result is due to Alicki [1, Theorem 3]; see [10] for a detailed proof.

**Theorem 2.4** (Structure of Lindblad operators with detailed balance) *Let  $\mathcal{P}_t = e^{t\mathcal{L}}$  be a quantum Markov semigroup on  $B(\mathcal{H})$  satisfying detailed balance with respect to  $\sigma \in \mathfrak{P}_+$ . Then the generator  $\mathcal{L}$  and its adjoint  $\mathcal{L}^\dagger$  have the form*

$$\mathcal{L} = \sum_{j \in \mathcal{J}} e^{-\omega_j/2} \mathcal{L}_j , \quad \mathcal{L}_j(A) = V_j^*[A, V_j] + [V_j^*, A]V_j , \tag{2.5}$$

$$\mathcal{L}^\dagger = \sum_{j \in \mathcal{J}} e^{-\omega_j/2} \mathcal{L}_j^\dagger , \quad \mathcal{L}_j^\dagger(\rho) = [V_j, \rho V_j^*] + [V_j \rho, V_j^*] , \tag{2.6}$$

where  $\mathcal{J}$  is a finite index set, the operators  $V_j \in B(\mathcal{H})$  satisfy  $\{V_j\}_{j \in \mathcal{J}} = \{V_j^*\}_{j \in \mathcal{J}}$ , and  $\omega_j \in \mathbb{R}$  satisfies

$$\Delta_\sigma V_j = e^{-\omega_j} V_j \quad \text{for all } j \in \mathcal{J} . \tag{2.7}$$

For  $j \in \mathcal{J}$ , let  $j^* \in \mathcal{J}$  be an index such that  $V_{j^*} = V_j^*$ . It follows from (2.7) that

$$\omega_{j^*} = -\omega_j .$$

Moreover, if we define  $H = -\log \sigma$ , (2.7) is equivalent to the commutator identity  $[V_j, H] = -\omega_j V_j$ . Furthermore, in our finite-dimensional context, the identity

$$\Delta_\sigma^t V_j = e^{-\omega_j t} V_j \tag{2.8}$$

is valid for some  $t \neq 0$  in  $\mathbb{R}$  if and only if it is valid for all  $t \in \mathbb{C}$ .

### 2.2 Gradient Flow Structure for the Non-commutative Dirichlet Energy

Let  $(\mathcal{P}_t)_{t \geq 0}$  be a quantum Markov semigroup satisfying detailed balance with respect to  $\sigma \in \mathfrak{P}_+(A)$ . Let  $\mathcal{L}$  be the generator, so that for each  $t > 0$ ,  $\mathcal{P}_t = e^{t\mathcal{L}}$ . As explained in the discussion leading up to Definition 2.3, for each  $t$ ,  $\mathcal{P}_t$  is self-adjoint with respect to both the GNS and the KMS inner products induced by  $\sigma$ . Therefore, we may define a *Dirichlet form*  $\mathcal{E}$  on  $A$  by

$$\mathcal{E}(A, A) = \lim_{t \downarrow 0} \frac{1}{t} \langle A, (I - \mathcal{P}_t)A \rangle \tag{2.9}$$

where the inner product is either the GNS or the KMS inner product. Then, either way, the *Kolmogorov backward equation*  $\partial_t A = \mathcal{L}A$  is a gradient flow equation for the energy  $\mathcal{E}(A, A)$  with respect to the chosen  $L^2$  metric.

The class of bilinear forms  $\mathcal{E}$  defined in terms of a self-adjoint QMS  $(\mathcal{P}_t)_{t \geq 0}$  through (2.9) is, by definition, the class of *conservative completely Dirichlet forms* on  $A$  in the specified

inner product. The abstract Beurling–Deny Theorem, discussed in the next section, provides an intrinsic characterization of such bilinear forms.

Although Definition 2.3 might seem to suggest that the natural choice of the  $L^2$  metric is the one given by the GNS inner product, we shall show that in some sense it is the KMS inner product that is more natural: The Dirichlet form defined by (2.9) using the KMS inner product induced by  $\sigma$  can be expressed in terms of a “squared gradient”, and the associated non-commutative differential calculus will turn out to be very useful for investigating properties of the flow specified by  $\partial_t A = \mathcal{L}A$ . A somewhat different construction leading to the representation of Dirichlet forms with respect to the KMS metric in terms of derivations has been given by Cipriani and Sauvageot [13]. Our “derivatives” are not always derivations, and this more general structure is suited to applications. Indeed, one of the first non-commutative Dirichlet forms to be investigated in mathematical physics, the *Clifford Dirichlet form* of Gross, is most naturally expressed in terms of a sum of squares of *skew derivations*. The flexibility of our framework will be essential to our later applications. In this part of the introduction, we present only some of the key computations in a simple setting involving derivations to explain the roles of the KMS inner product. Our more general framework will be presented in Sect. 4.

Consider a Lindblad generator  $\mathcal{L}$  given as in Theorem 2.4. To bring out the analogy with classical Kolmogorov backward diffusion equations of the form

$$\frac{\partial}{\partial t} f(x, t) = \Delta f(x, t) + (\nabla \log \sigma(x)) \cdot \nabla f(x, t), \tag{2.10}$$

where  $\sigma$  is a smooth, strictly positive probability density on  $\mathbb{R}^n$ , we define the following *partial derivative operators on  $\mathcal{A}$* :

$$\partial_j A = [V_j, A], \tag{2.11}$$

$j \in \mathcal{J}$ . Note that  $\partial_j^\dagger = \partial_{j^*}$ , where we recall that  $j^*$  denotes an index such that  $V_{j^*} = V_j^*$ . An easy computation shows that the adjoint of  $\partial_j$  with respect to  $\langle \cdot, \cdot \rangle_{L^2_{\text{KMS}}(\sigma)}$  is given by

$$\partial_{j,\sigma}^\dagger A = \sigma^{-1/2} \partial_j^\dagger (\sigma^{1/2} A \sigma^{1/2}) \sigma^{-1/2}. \tag{2.12}$$

**Proposition 2.5** (Divergence form representation of  $\mathcal{L}$ ) *For all  $A \in \mathcal{A}$  we have*

$$\mathcal{L}A = - \sum_{j \in \mathcal{J}} \partial_{j,\sigma}^\dagger \partial_j A.$$

**Proof** Using (2.12) and (2.8) we obtain

$$\begin{aligned} \sum_{j \in \mathcal{J}} \partial_{j,\sigma}^\dagger \partial_j A &= \sum_{j \in \mathcal{J}} \partial_{j,\sigma}^\dagger (V_j A - A V_j) \\ &= \sum_{j \in \mathcal{J}} \sigma^{-1/2} \partial_j^\dagger (\sigma^{1/2} (V_j A - A V_j) \sigma^{1/2}) \sigma^{-1/2} \\ &= \sum_{j \in \mathcal{J}} \sigma^{-1/2} \left( V_j^* \sigma^{1/2} (V_j A - A V_j) \sigma^{1/2} - \sigma^{1/2} (V_j A - A V_j) \sigma^{1/2} V_j^* \right) \sigma^{-1/2} \\ &= \sum_{j \in \mathcal{J}} \left( e^{-\omega_j/2} V_j^* (V_j A - A V_j) - e^{\omega_j/2} (V_j A - A V_j) V_j^* \right) \end{aligned}$$

$$\begin{aligned}
 &= \sum_{j \in \mathcal{J}} \left( e^{-\omega_j/2} V_j^* (V_j A - A V_j) - e^{-\omega_j/2} (V_j^* A - A V_j^*) V_j \right) \\
 &= - \sum_{j \in \mathcal{J}} e^{-\omega_j/2} \mathcal{L}_j(A) = -\mathcal{L}A,
 \end{aligned}$$

as desired. □

Proposition 2.5 can be stated equivalently as an integration by parts identity

$$\sum_{j \in \mathcal{J}} \langle \partial_j A, \partial_j B \rangle_{L^2_{\text{KMS}}(\sigma)} = -\langle A, \mathcal{L}B \rangle_{L^2_{\text{KMS}}(\sigma)} \quad \text{for } A, B \in \mathcal{A}. \tag{2.13}$$

It is now immediate that the backward equation  $\partial_t A = \mathcal{L}A$  with  $\mathcal{L}$  given by (2.1), is the gradient flow equation for the energy  $\mathcal{E}(A, A)$  with respect to the KMS inner product induced by  $\sigma$ . What makes this particular gradient flow representation especially useful is that the Dirichlet form  $\mathcal{E}$  is written, in (2.13), as the expectation of a squared gradient. That is, the gradient flow structure given here is analogous to the gradient flow formulation for the Kolmogorov backward equation (2.10) for the Dirichlet energy  $\mathcal{D}_{\text{class}}(f) = \frac{1}{2} \int_{\mathbb{R}^n} |\nabla f(x)|^2 \sigma(x) dx$ . This would not be the case if we had considered the Dirichlet form based on the GNS inner product: We would have a gradient flow structure, but the Dirichlet form would not be the expectation of a squared gradient in any meaningful sense; see however, Proposition 4.12 below for a related representation.

In the next section we show how the non-commutative differential calculus associated to the Dirichlet form  $\mathcal{E}$  allows us to write the corresponding *forward equation* as gradient flow for the relative entropy with respect to a Riemannian metric constructed in terms of this differential calculus.

### 2.3 A Gradient Flow Structure for the Quantum Relative Entropy

Consider the *quantum relative entropy functionals*  $\text{Ent}_\sigma : \mathfrak{P}_+ \rightarrow \mathbb{R}$  given by

$$\text{Ent}_\sigma(\rho) := \tau[\rho(\log \rho - \log \sigma)].$$

Our goal is to sketch a proof of one of the results of [10,36], namely that the quantum master equation  $\partial_t \rho = \mathcal{L}^\dagger \rho$ , which is a Kolmogorov *forward equation*, can be formulated as the gradient flow equation for  $\text{Ent}_\sigma$  with respect to a suitable Riemannian metric on  $\mathfrak{P}_+$ . The construction of the Riemannian metric will make use of the “quantum directional derivatives”  $\partial_j$  introduced in the last subsection.

Since  $\mathfrak{P}_+$  is a relatively open subset of the  $\mathbb{R}$ -affine subspace  $\{A \in \mathcal{A}_h : \tau[A] = 1\}$ , we may identify, at each point in  $\rho \in \mathfrak{P}_+$ , its tangent space  $T_\rho \mathfrak{P}_+$  with  $\mathcal{A}_0 := \{A \in \mathcal{A}_h : \tau[A] = 0\}$ . The cotangent space  $T_\rho^\dagger \mathfrak{P}_+$  may also be identified with  $\mathcal{A}_0$  through the duality pairing  $\langle A, B \rangle = \tau[AB]$  for  $A, B \in \mathcal{A}_0$ .

Let  $(g_\rho)_{\rho \in \mathfrak{P}_+}$  be a Riemannian metric on  $\mathfrak{P}_+$ , i.e., a collection of positive definite bilinear forms  $g_\rho : T_\rho \mathfrak{P}_+ \times T_\rho \mathfrak{P}_+ \rightarrow \mathbb{R}$  depending smoothly on  $\rho \in \mathfrak{P}_+$ . Consider the associated operator  $\mathcal{G}_\rho : T_\rho \mathfrak{P}_+ \rightarrow T_\rho^\dagger \mathfrak{P}_+$  defined by  $\langle A, \mathcal{G}_\rho B \rangle = g_\rho(A, B)$  for  $A, B \in T_\rho \mathfrak{P}_+$ . Clearly,  $\mathcal{G}_\rho$  is invertible and self-adjoint with respect to the Hilbert–Schmidt inner product on  $\mathcal{A}_0$ . Define  $\mathcal{K}_\rho : T_\rho^\dagger \mathfrak{P}_+ \rightarrow T_\rho \mathfrak{P}_+$  by  $\mathcal{K}_\rho = (\mathcal{G}_\rho)^{-1}$ , so that

$$g_\rho(A, B) = \langle A, \mathcal{K}_\rho^{-1} B \rangle. \tag{2.14}$$

In many situations of interest it is convenient to define the metric  $g_\rho$  by specifying the operator  $\mathcal{K}_\rho$ . In such cases, there is often no explicit formula available for  $\mathcal{G}_\rho$  and  $g_\rho$ .



For a smooth functional  $\mathcal{F} : \mathfrak{P}_+ \rightarrow \mathbb{R}$  and  $\rho \in \mathfrak{P}_+$ , its differential  $D\mathcal{F}(\rho) \in T_\rho^\dagger \mathfrak{P}_+$  is defined by  $\lim_{\varepsilon \rightarrow 0} \varepsilon^{-1}(\mathcal{F}(\rho + \varepsilon A) - \mathcal{F}(\rho)) = \langle A, D\mathcal{F}(\rho) \rangle$  for  $A \in T_\rho \mathfrak{P}_+$  (independently of the Riemannian metric  $g_\rho$ ). Its gradient  $\nabla_g \mathcal{F}(\rho) \in T_\rho \mathfrak{P}_+$  depends on the Riemannian metric through the duality formula  $g_\rho(A, \nabla_g \mathcal{F}(\rho)) = \langle A, D\mathcal{F}(\rho) \rangle$  for  $A \in T_\rho \mathfrak{P}_+$ . It follows that  $\mathcal{G}_\rho \nabla_g \mathcal{F}(\rho) = D\mathcal{F}(\rho)$ , or equivalently

$$\nabla_g \mathcal{F}(\rho) = \mathcal{K}_\rho D\mathcal{F}(\rho) .$$

The gradient flow equation  $\partial_t \rho = -\nabla_g \mathcal{F}(\rho)$  takes the form

$$\partial_t \rho = -\mathcal{K}_\rho D\mathcal{F}(\rho) .$$

Let us now focus on the relative entropy functional  $\text{Ent}_\sigma$  for some  $\sigma \in \mathfrak{P}_+$ , and note that its differential is given by

$$D\text{Ent}_\sigma(\rho) = \log \rho - \log \sigma . \tag{2.15}$$

Consider a generator  $\mathcal{L}^\dagger$  written in the form (2.6), i.e.,

$$\mathcal{L}^\dagger = \sum_{j \in \mathcal{J}} e^{-\omega_j/2} \mathcal{L}_j^\dagger , \quad \mathcal{L}_j^\dagger(\rho) = [V_j, \rho V_j^*] + [V_j \rho, V_j^*] ,$$

where  $\{V_j\}_{j \in \mathcal{J}}$  is a finite set of eigenvectors of  $\Delta_\sigma$  such that  $\{V_j^*\}_{j \in \mathcal{J}} = \{V_j\}_{j \in \mathcal{J}}$ , and where  $\Delta_\sigma V_j = e^{-\omega_j} V_j$  for some  $\omega_j \in \mathbb{R}$ . As before, we use the notation  $\partial_j A := [V_j, A]$ .

For  $\rho \in \mathfrak{P}$  we define  $\widehat{\rho}_j \in \mathcal{A} \otimes \mathcal{A}$  by

$$\widehat{\rho}_j = \int_0^1 (e^{\omega_j/2} \rho)^{1-s} \otimes (e^{-\omega_j/2} \rho)^s \, ds .$$

We shall frequently make use of the contraction operator  $\# : (\mathcal{A} \otimes \mathcal{A}) \times \mathcal{A} \rightarrow \mathcal{A}$  defined by

$$(A \otimes B)\#C := ACB \tag{2.16}$$

and linear extension. A crucial step towards obtaining the gradient flow structure is the following chain rule for the commutators  $\partial_j$ , which involves the differential of the entropy.

**Lemma 2.6** (Chain rule for the logarithm) *For all  $\rho \in \mathfrak{P}_+$  and  $j \in \mathcal{J}$  we have*

$$e^{-\omega_j/2} V_j \rho - e^{\omega_j/2} \rho V_j = \widehat{\rho}_j \# \partial_j (\log \rho - \log \sigma) . \tag{2.17}$$

**Proof** Using (2.7) we infer that

$$\partial_j (\log \rho - \log \sigma) = V_j \log(e^{-\omega_j/2} \rho) - \log(e^{\omega_j/2} \rho) V_j .$$

Consider the spectral decomposition  $\rho = \sum_\ell \lambda_\ell E_\ell$ , where  $\lambda_\ell > 0$  for all  $i$ , and  $\{E_\ell\}_\ell$  are the spectral projections, so that  $E_\ell E_m = \delta_{\ell m} E_\ell$  and  $\sum_\ell E_\ell = \mathbf{1}$ . Observe that

$$\widehat{\rho}_j = \sum_{\ell, m} \Lambda(e^{\omega_j/2} \lambda_\ell, e^{-\omega_j/2} \lambda_m) E_\ell \otimes E_m ,$$

where  $\Lambda(\xi, \eta) = \int_0^1 \xi^{1-s} \eta^s \, ds = \frac{\xi - \eta}{\log \xi - \log \eta}$  denotes the logarithmic mean of  $\xi$  and  $\eta$ . Thus,

$$\begin{aligned} & \widehat{\rho}_j \# (\partial_j (\log \rho - \log \sigma)) \\ &= \sum_{\ell, m, p} \Lambda(e^{\omega_j/2} \lambda_\ell, e^{-\omega_j/2} \lambda_m) E_\ell \left( \log(e^{-\omega_j/2} \lambda_p) V_j E_p - \log(e^{\omega_j/2} \lambda_p) E_p V_j \right) E_m \end{aligned}$$

$$\begin{aligned}
 &= \sum_{\ell,m} \Lambda(e^{\omega_j/2}\lambda_\ell, e^{-\omega_j/2}\lambda_m) \left( \log(e^{-\omega_j/2}\lambda_m) - \log(e^{\omega_j/2}\lambda_\ell) \right) E_\ell V_j E_m \\
 &= \sum_{\ell,m} (e^{-\omega_j/2}\lambda_m - e^{\omega_j/2}\lambda_\ell) E_\ell V_j E_m \\
 &= e^{-\omega_j/2} V_j \rho - e^{\omega_j/2} \rho V_j,
 \end{aligned}$$

which proves (2.17). □

For  $\rho \in \mathfrak{P}_+$  we define the operator  $\mathcal{K}_\rho : \mathcal{A} \rightarrow \mathcal{A}$  by

$$\mathcal{K}_\rho A := \sum_{j \in \mathcal{J}} \partial_j^\dagger (\widehat{\rho}_j \# \partial_j A). \tag{2.18}$$

Since  $\text{Tr}(A^* \mathcal{K}_\rho B) = \overline{\text{Tr}(B^* \mathcal{K}_\rho A)}$  for  $A, B \in \mathcal{A}$ , it follows that  $\mathcal{K}_\rho$  is a non-negative self-adjoint operator on  $L^2(\mathcal{A}, \tau)$  for each  $\rho \in \mathfrak{P}_+$ . Assuming that  $\mathcal{P}_t$  is ergodic, the operator  $\mathcal{K}_\rho : \mathcal{A}_0 \rightarrow \mathcal{A}_0$  is invertible for each  $\rho \in \mathfrak{P}_+$  (see Corollary 7.4 below for a proof of this statement). Since  $\mathcal{K}_\rho$  depends smoothly on  $\rho$ , it follows that  $\mathcal{K}_\rho$  induces a Riemannian metric on  $\mathfrak{P}_+$  defined by (2.14).

The following result shows that the Kolmogorov forward equation  $\partial_t \rho = \mathcal{L}^\dagger \rho$  can be formulated as the gradient flow equation for  $\text{Ent}_\sigma$ .

**Proposition 2.7** *For  $\rho \in \mathfrak{P}_+$  we have the identity*

$$\mathcal{L}^\dagger \rho = -\mathcal{K}_\rho \text{D Ent}_\sigma(\rho),$$

hence the gradient flow equation of  $\text{Ent}_\sigma$  with respect to the Riemannian metric induced by  $(\mathcal{K}_\rho)_\rho$  is the master equation  $\partial_t \rho = \mathcal{L}^\dagger \rho$ .

**Proof** Using the identity (2.15), the chain rule from Lemma 2.6, and the fact that  $\{V_j\} = \{V_j^*\}$  and  $\omega_{j^*} = -\omega_j$ , we obtain

$$\begin{aligned}
 \mathcal{K}_\rho \text{D Ent}_\sigma(\rho) &= \sum_{j \in \mathcal{J}} \partial_j^\dagger (\widehat{\rho}_j \# \partial_j (\log \rho - \log \sigma)) \\
 &= \sum_{j \in \mathcal{J}} \partial_j^\dagger (e^{-\omega_j/2} V_j \rho - e^{\omega_j/2} \rho V_j) \\
 &= \frac{1}{2} \sum_{j \in \mathcal{J}} \left( \partial_j^\dagger (e^{-\omega_j/2} V_j \rho - e^{\omega_j/2} \rho V_j) + \partial_j (e^{\omega_j/2} V_j^* \rho - e^{-\omega_j/2} \rho V_j^*) \right) \\
 &= -\frac{1}{2} \sum_{j \in \mathcal{J}} e^{-\omega_j/2} \left( [V_j, \rho V_j^*] + [V_j \rho, V_j^*] \right) + e^{\omega_j/2} \left( [V_j^*, \rho V_j] + [V_j^* \rho, V_j] \right) \\
 &= -\sum_{j \in \mathcal{J}} e^{-\omega_j/2} \left( [V_j, \rho V_j^*] + [V_j \rho, V_j^*] \right) \\
 &= -\mathcal{L}^\dagger \rho,
 \end{aligned}$$

which is the desired identity. □

In this paper we extend this result into various directions: we consider more general entropy functionals, more general Riemannian metrics, and nonlinear evolution equations.

**Remark 2.8** The gradient flow structure given in Proposition 2.7 can be viewed as a non-commutative analogue of the Kantorovich gradient flow structure obtained by Jordan, Kinderlehrer and Otto [29] for the *Kolmogorov backward equation*

$$\frac{\partial}{\partial t} \rho(x, t) = \Delta \rho(x, t) - \nabla \cdot (\rho(x, t) \nabla \log \sigma(x)) .$$

This structure is formally given in terms of the operator  $K_\rho$  defined by

$$K_\rho \psi = -\nabla \cdot (\rho \nabla \psi) ,$$

for probability densities  $\rho$  on  $\mathbb{R}^n$  and suitable functions  $\psi : \mathbb{R}^n \rightarrow \mathbb{R}$  in analogy with (2.18). As the differential of the relative entropy  $\text{Ent}_\sigma(\rho) = \int_{\mathbb{R}^n} \rho(x) \log \frac{\rho(x)}{\sigma(x)} dx$  is given by  $D \text{Ent}_\sigma(\rho) = 1 + \log \frac{\rho}{\sigma}$ , we have

$$K_\rho D \text{Ent}_\sigma(\rho) = -\Delta \rho + \nabla \cdot (\rho \nabla \log \sigma) ,$$

which is the commutative counterpart of Proposition 2.7.

### 2.4 The Necessity of BKM-Detailed Balance

In the classical setting of irreducible finite Markov chain, Dietert [15] has proven that if the Kolmogorov forward equation for a Markov semigroup can be written as gradient flow for the relative entropy with respect to the unique invariant measure for some continuously differentiable Riemannian metric, then the Markov chain is necessarily reversible. That is, it satisfies the classical detailed balance condition.

**Theorem 2.9** *Let  $(\mathcal{P}_t)_{t \geq 0}$  be an ergodic QMS with generator  $\mathcal{L}$  and invariant state  $\sigma \in \mathfrak{P}_+$ . If there exists a continuously differentiable Riemannian metric  $(g_\rho)$  on  $\mathfrak{P}_+$  such that the quantum master equation  $\partial_t \rho = \mathcal{L}^\dagger \rho$  is the gradient flow equation for  $\text{Ent}_\sigma$  with respect to  $(g_\rho)$ , then each  $\mathcal{P}_t$  is self-adjoint with respect to the BKM inner product associated to  $\sigma$ .*

Before beginning the proof, we recall some relevant facts, and introduce some notation. Regarding  $\sigma$  as an element of  $\mathbb{M}_n(\mathbb{C})$ , we define the operator  $\mathcal{M}$  on  $\mathbb{M}_n(\mathbb{C})$  by

$$\mathcal{M} A = \int_0^1 \sigma^{1-s} A \sigma^s ds .$$

A simple calculation shows that  $\mathcal{M}$  is the derivative of the matrix exponential function. Its inverse is the derivative of the matrix logarithm function:

$$\mathcal{M}^{-1} A = \int_0^\infty \frac{1}{t + \sigma} A \frac{1}{t + \sigma} dt ,$$

(see Example 6.5 below for more details). While the matrix logarithm function is monotone, the matrix exponential is not. Thus  $\mathcal{M}^{-1}$  preserves positivity, but  $\mathcal{M}$  does not. In fact  $A \mapsto \mathcal{M}^{-1} A$  is evidently completely positive. The BKM inner product can now be written as

$$\langle A, B \rangle_{L^2_{\text{BKM}}(\sigma)} = \tau[A^* \mathcal{M} B] = \tau[\mathcal{M}(A^*) B] .$$

**Proof of Theorem 2.9** As before, it will be convenient to consider the operators  $(\mathcal{K}_\rho)$  defined by (2.14). Since  $D \text{Ent}_\sigma(\rho) = \log \rho - \log \sigma$ , the gradient flow equation  $\partial_t \rho = -\mathcal{K}_\rho D \text{Ent}_\sigma(\rho)$  becomes

$$\mathcal{L}^\dagger \rho = -\mathcal{K}_\rho(\log \rho - \log \sigma) . \tag{2.19}$$

Applying this identity to  $\rho_\varepsilon = \sigma + \varepsilon A$  for  $A \in \mathcal{A}_0$ , and differentiating at  $\varepsilon = 0$ , we obtain using the identity  $\partial_\varepsilon|_{\varepsilon=0} \log \rho_\varepsilon = \mathcal{M}^{-1}A$  that

$$\mathcal{L}^\dagger A = -\mathcal{H}_\sigma \mathcal{M}^{-1}A, \tag{2.20}$$

Consequently, for  $A, B \in \mathcal{A}$ ,

$$\langle \mathcal{L}A, B \rangle_{L^2_{\text{BKM}}(\sigma)} = \tau[(\mathcal{L}A)^* \mathcal{M}B] = \tau[A^* \mathcal{L}^\dagger \mathcal{M}B] = -\tau[A^* \mathcal{H}_\sigma B].$$

As  $g_\sigma$  is a symmetric bilinear form, the operator  $\mathcal{H}_\sigma$  is self-adjoint with respect to the Hilbert-Schmidt scalar product. This implies the result.  $\square$

We are unaware of any investigation of the nature of the class of QMS generators that are self-adjoint for the BKM inner product associated to their invariant state  $\sigma$ . Therefore we briefly demonstrate that this class *strictly* includes the class of QMS generators that are self-adjoint for the GNS inner product associated to their invariant state  $\sigma$ .

Let  $\mathcal{P}$  be a unital completely positive map such that  $\mathcal{P}^\dagger \sigma = \sigma$ , and define

$$\tilde{\mathcal{P}}(A) = \mathcal{M}^{-1}(\sigma^{1/2} \mathcal{P}(A) \sigma^{1/2}).$$

Note that

$$\mathcal{M}^{-1}(\sigma^{1/2} A \sigma^{1/2}) = \int_0^\infty \frac{\sigma^{1/2}}{t + \sigma} A \frac{\sigma^{1/2}}{t + \sigma} dt$$

defines a completely positive and unital operator, and hence  $\tilde{\mathcal{P}}$  is completely positive and unital. Moreover,

$$\tilde{\mathcal{P}}^\dagger(A) = \mathcal{P}^\dagger(\mathcal{M}^{-1}(\sigma^{1/2} A \sigma^{1/2})),$$

and hence  $\tilde{\mathcal{P}}^\dagger \sigma = \sigma$ . Now observe that  $\tilde{\mathcal{P}}$  is self-adjoint with respect to the BKM inner product if and only if  $\mathcal{P}$  is self-adjoint for the KMS inner product. In fact, for all  $A, B \in \mathcal{A}$ ,

$$\langle \tilde{\mathcal{P}}A, B \rangle_{L^2_{\text{BKM}}(\sigma)} = \tau[\mathcal{M}(\mathcal{M}^{-1}(\sigma^{1/2} \mathcal{P}(A^*) \sigma^{1/2}))B] = \langle \mathcal{P}A, B \rangle_{L^2_{\text{KMS}}(\sigma)}.$$

Next, it is clear that  $\tilde{\mathcal{P}}$  commutes with  $\Delta_\sigma$  if and only if  $\mathcal{P}$  commutes with  $\Delta_\sigma$ . Since there exist completely positive unital maps  $\mathcal{P}$  satisfying  $\mathcal{P}^\dagger \sigma = \sigma$  that are KMS symmetric but do not commute with  $\Delta_\sigma$ , there exists completely positive unital maps  $\tilde{\mathcal{P}}$  satisfying  $\tilde{\mathcal{P}}^\dagger \sigma = \sigma$  that are BKM symmetric but do not commute with  $\Delta_\sigma$ .

Moreover, the class of completely positive unital maps  $\tilde{\mathcal{P}}$  satisfying  $\tilde{\mathcal{P}}^\dagger \sigma = \sigma$  that are BKM symmetric is in some sense larger than the class of completely positive unital maps  $\mathcal{P}$  satisfying  $\mathcal{P}^\dagger \sigma = \sigma$  that are KMS symmetric: The map  $\mathcal{P} \mapsto \tilde{\mathcal{P}}$  is invertible, but  $\mathcal{M}$  is not even positivity preserving, let alone completely positive, so that

$$\mathcal{P}(A) = \sigma^{-1/2} \mathcal{M}(\tilde{\mathcal{P}}(A)) \sigma^{-1/2}$$

need not be completely positive. It is therefore an interesting problem to characterize the QMS generators that are self-adjoint with respect to the BKM inner product.

### 3 Beurling–Deny Theory in Finite-Dimensional von Neumann Algebras

In this section we recall some key results of Beurling–Deny theory that will be used in our construction of Dirichlet forms in Sect. 4. We present some proofs of known results for the reader’s convenience, especially when available references suppose a familiarity with the

Tomita–Takesaki theory. However, Theorem 3.8, which singles out the KMS inner product, is new.

### 3.1 Abstract Beurling–Deny Theory

In this subsection,  $\mathcal{H}$  always denotes a *real* Hilbert space with inner product  $\langle \cdot, \cdot \rangle$ . Let  $\mathcal{P}$  be a cone in  $\mathcal{H}$ . That is,  $\mathcal{P}$  is a convex subset of  $\mathcal{H}$  such that if  $\varphi \in \mathcal{P}$ , then  $\lambda\varphi \in \mathcal{P}$  for all  $\lambda > 0$ . The cone  $\mathcal{P}$  is *pointed* in case  $\varphi \in \mathcal{P}$  and  $-\varphi \in \mathcal{P}$  together imply that  $\varphi = 0$ . In particular, a subspace of  $\mathcal{H}$  is a cone, but it is not a pointed cone.

**Definition 3.1** (*Dual cone*) The *dual cone*  $\mathcal{P}^\circ$  of a cone  $\mathcal{P}$  is the set

$$\mathcal{P}^\circ := \{ \psi \in \mathcal{H} : \langle \psi, \varphi \rangle \geq 0 \text{ for all } \varphi \in \mathcal{P} \}. \tag{3.1}$$

A cone  $\mathcal{P}$  is *self-dual* in case  $\mathcal{P}^\circ = \mathcal{P}$ .

Let  $\mathcal{P}$  be a non-empty self-dual cone in  $\mathcal{H}$ , and take  $\varphi \in \mathcal{H}$ . Since  $\mathcal{P}$  is a non-empty closed, convex set, the Projection Lemma ensures the existence of  $P_{\mathcal{P}}(\varphi) \in \mathcal{P}$  such that

$$\| \varphi - P_{\mathcal{P}}(\varphi) \| < \| \varphi - \psi \| \text{ for all } \psi \in \mathcal{P}, \psi \neq P_{\mathcal{P}}(\varphi). \tag{3.2}$$

**Theorem 3.2** (*Decomposition Theorem*) *Let  $\mathcal{P}$  be a non-empty self-dual cone in  $\mathcal{H}$ . Then for each  $\varphi \in \mathcal{H}$ , there exists a unique pair  $\varphi_+, \varphi_- \in \mathcal{P}$  such that*

$$\varphi = \varphi_+ - \varphi_- \text{ and } \langle \varphi_+, \varphi_- \rangle = 0. \tag{3.3}$$

*In fact,  $\varphi_+ = P_{\mathcal{P}}(\varphi)$  and  $\varphi_- = P_{\mathcal{P}}(-\varphi)$ , where  $P_{\mathcal{P}}$  denotes projection onto (the closed convex set)  $\mathcal{P}$ .*

**Proof** Define  $\varphi_+ := P_{\mathcal{P}}(\varphi)$ . Then define  $-\varphi_- := \varphi - \varphi_+$ . We claim that  $\varphi_- \in \mathcal{P}$ . Indeed, for any  $\psi \in \mathcal{P}$  and any  $\epsilon > 0$ ,  $\varphi_+ + \epsilon\psi \in \mathcal{P}$ , and hence,

$$\| \varphi_- \|^2 = \| \varphi - \varphi_+ \|^2 < \| \varphi - (\varphi_+ + \epsilon\psi) \|^2 = \| \varphi_- \|^2 + 2\epsilon \langle \varphi_-, \psi \rangle + \epsilon^2 \| \psi \|^2.$$

Therefore,  $\langle \varphi_-, \psi \rangle \geq 0$  for all  $\psi \in \mathcal{P}$ . Since  $\mathcal{P}$  is self-dual, the claim follows.

To see that  $\varphi_+$  and  $\varphi_-$  are orthogonal, let  $\epsilon \in (-1, 1)$ , so that  $(1 + \epsilon)\varphi_+ \in \mathcal{P}$ . It follows that  $\| \varphi_- \|^2 = \| \varphi - \varphi_+ \|^2 \leq \| \varphi - (1 + \epsilon)\varphi_+ \|^2 = \| \varphi_- \|^2 + 2\epsilon \langle \varphi_-, \varphi_+ \rangle + \epsilon^2 \| \varphi_+ \|^2$  which yields a contradiction for negative  $\epsilon$  sufficiently close to zero, unless  $\langle \varphi_-, \varphi_+ \rangle = 0$ . This proves existence of the decomposition. Now the fact that  $\varphi_- = P_{\mathcal{P}}(-\varphi)$  follows from a theorem of Moreau [37], as does the uniqueness of the decomposition, though both points can be proved directly by variations on the arguments just provided.  $\square$

**Definition 3.3** Let  $\mathcal{H}$  be a real Hilbert space with a non-empty self-dual cone  $\mathcal{P}$ . For  $\varphi$  in  $\mathcal{H}$ , define  $\varphi_+$  and  $\varphi_-$  as in Theorem 3.2. Then  $\varphi_+$  is the *positive part* of  $\varphi$ ,  $\varphi_-$  is the *negative part* of  $\varphi$ , and  $|\varphi| := \varphi_+ + \varphi_-$  is the *absolute value* of  $\varphi$ . If  $\varphi_- = 0$ , we write  $\varphi \geq 0$ .

We next recall some elements of the abstract theory of symmetric Dirichlet forms. A *bilinear form* on a real Hilbert space  $\mathcal{H}$  is a bilinear mapping  $\mathcal{E} : \mathcal{D} \times \mathcal{D} \rightarrow \mathbb{R}$  where  $\mathcal{D} \subseteq \mathcal{H}$  is a linear subspace (called the *domain* of  $\mathcal{E}$ ). We say that  $\mathcal{E}$  is *non-negative* if  $\mathcal{E}(\varphi, \varphi) \geq 0$  for all  $\varphi \in \mathcal{D}$ ; *symmetric* if  $\mathcal{E}(\varphi, \psi) = \mathcal{E}(\psi, \varphi)$  for all  $\psi, \varphi \in \mathcal{D}$ ; *closed* if  $\mathcal{D}$  is complete when endowed with the norm  $\| \varphi \|_{\mathcal{E}} = (\| \varphi \|^2 + \mathcal{E}(\varphi, \varphi))^{1/2}$ ; and *densely defined* if  $\mathcal{D}$  is dense in  $\mathcal{H}$ .

**Definition 3.4** (*Dirichlet form*) Let  $\mathcal{H}$  be a real Hilbert space with a non-empty self-dual cone  $\mathcal{P}$ . A non-negative, symmetric, closed bilinear form  $\mathcal{E}$  on  $\mathcal{H}$  with dense domain  $\mathcal{D}$  is a *Dirichlet form* in case  $|\varphi| \in \mathcal{D}$  for all  $\varphi \in \mathcal{D}$ , and

$$\mathcal{E}(|\varphi|, |\varphi|) \leq \mathcal{E}(\varphi, \varphi), \tag{3.4}$$

or equivalently, if for all  $\varphi \in \mathcal{D}$ ,

$$\mathcal{E}(\varphi_+, \varphi_-) \leq 0. \tag{3.5}$$

To see the equivalence of (3.4) and (3.5), note that

$$\mathcal{E}(|\varphi|, |\varphi|) - \mathcal{E}(\varphi, \varphi) = 4\mathcal{E}(\varphi_+, \varphi_-).$$

Given a non-negative, symmetric, closed bilinear form  $\mathcal{E}$ , the operator  $\mathcal{L} : \mathcal{D}_{\mathcal{L}} \subseteq \mathcal{H} \rightarrow \mathcal{H}$  associated to  $\mathcal{E}$  is defined by

$$\mathcal{D}_{\mathcal{L}} := \{\psi \in \mathcal{D} \mid \exists \xi \in \mathcal{H} : \mathcal{E}(\varphi, \psi) = -\langle \varphi, \xi \rangle \ \forall \varphi \in \mathcal{D}\}, \quad \mathcal{L}\psi := \xi.$$

This operator is well-defined since  $\mathcal{D}_{\mathcal{L}}$  is dense. Moreover,  $\mathcal{L}$  is non-positive and self-adjoint.

The following abstract result by Ouhabaz [40] characterizes the invariance of closed convex sets under the associated semigroup (in a more general setting that includes nonsymmetric Dirichlet forms).

**Theorem 3.5** (Ouhabaz’ Theorem) *Let  $\mathcal{H}$  be a real Hilbert space, and let  $\mathcal{E}$  be a non-negative, symmetric, closed bilinear form with domain  $\mathcal{D}$  and associated operator  $\mathcal{L}$ . Let  $\mathcal{C} \subseteq \mathcal{H}$  be closed and convex. Then, the following assertions are equivalent:*

- (1)  $e^{t\mathcal{L}}\varphi \in \mathcal{C}$  for all  $\varphi \in \mathcal{C}$  and all  $t \geq 0$ ;
- (2)  $P_{\mathcal{C}}\varphi \in \mathcal{D}$  and  $\mathcal{E}(P_{\mathcal{C}}\varphi, \varphi - P_{\mathcal{C}}\varphi) \leq 0$  for all  $\varphi \in \mathcal{D}$ .

Combining Theorems 3.2 and 3.5 we obtain the following result.

**Corollary 3.6** (Abstract Beurling–Deny Theorem) *Let  $\mathcal{H}$  be a real Hilbert space with a non-empty self-dual cone  $\mathcal{P}$ . Let  $\mathcal{E}$  be a non-negative, symmetric, closed bilinear form with domain  $\mathcal{D}$ . Then,  $\mathcal{E}$  is a Dirichlet form if and only if  $e^{t\mathcal{L}}\varphi \geq 0$  for all  $t \geq 0$  and all  $\varphi \geq 0$ .*

### 3.2 Completely Dirichlet Forms

Let  $\mathcal{E}$  be a Dirichlet form on  $(\mathcal{A}, \langle \cdot, \cdot \rangle_{L^2_{\text{KMS}}(\sigma)})$  with the KMS inner product specified by a faithful state  $\sigma$ . Here, the notion of Dirichlet form is understood with respect to the self-dual cone consisting of all positive semidefinite matrices belonging to  $\mathcal{A}$ ; see Lemma 3.10 below. Let  $\mathcal{P}_t = e^{t\mathcal{L}}$  where  $\mathcal{L}$  is the semigroup generator associated to  $\mathcal{E}$ . Recall that the Dirichlet form  $\mathcal{E}$  is said to be *completely Dirichlet* in case for each  $t$ ,  $\mathcal{P}_t$  is completely positive.

The condition that  $\mathcal{E}$  be completely Dirichlet may be expressed in terms of  $\mathcal{E}$  itself, permitting one to check the property directly from a specification of  $\mathcal{E}$ .

For  $m \in \mathbb{N}$ , let  $E_{ij}$  denote the matrix whose  $(i, j)$ -entry is 1, with all other entries being 0. Alternatively,  $E_{ij}$  represents the linear transformation taking  $\mathbf{e}_j$  to  $\mathbf{e}_i$ , while annihilating  $\mathbf{e}_k$  for  $k \neq j$ . (Here  $\{\mathbf{e}_1, \dots, \mathbf{e}_m\}$  is the standard orthonormal basis of  $\mathbb{C}^m$ .) It follows that  $E_{ij}E_{k\ell} = \delta_{jk}E_{i\ell}$ . The general element of  $\mathcal{A} \otimes \mathbb{M}_m(\mathbb{C})$  can be written as

$$\mathbf{A} = \sum_{i,j=1}^m A_{ij} \otimes E_{ij} \tag{3.6}$$

where each  $A_{ij} \in \mathcal{A}$ . With  $\tau_m$  denoting the normalized trace on  $\mathbb{M}_m(\mathbb{C})$ , the state  $\sigma \otimes \tau_m$  on  $\mathcal{A} \otimes \mathbb{M}_m(\mathbb{C})$  is defined by

$$\sigma \otimes \tau_m(\mathbf{A}) := \frac{1}{m} \sum_{j=1}^m \sigma(A_{jj}),$$

where  $\mathbf{A}$  is given by (3.6). The corresponding KMS inner product on  $\mathcal{A} \otimes \mathbb{M}_m(\mathbb{C})$  is denoted  $\langle \cdot, \cdot \rangle_{L^2_{\text{KMS}}(\sigma \otimes \tau_m)}$ . One readily checks that for  $\mathbf{A}, \mathbf{B} \in \mathcal{A} \otimes \mathbb{M}_m(\mathbb{C})$ ,

$$\langle \mathbf{B}, \mathbf{A} \rangle_{L^2_{\text{KMS}}(\sigma \otimes \tau_m)} = \frac{1}{m} \sum_{i,j=1}^m \langle B_{ij}, A_{ij} \rangle_{L^2_{\text{KMS}}(\sigma)}.$$

Define  $\mathcal{P}_t^{(m)}$  on  $\mathcal{A} \otimes \mathbb{M}_m(\mathbb{C})$  by

$$\mathcal{P}_t^{(m)} \mathbf{A} = \sum_{i,j=1}^m \mathcal{P}_t A_{ij} \otimes E_{ij} \tag{3.7}$$

where  $\mathbf{A}$  is given by (3.6). One then computes

$$-\frac{d}{dt} \langle \mathbf{A}, \mathcal{P}_t^{(m)} \mathbf{A} \rangle_{L^2_{\text{KMS}}(\sigma \otimes \tau_m)} \Big|_{t=0} = -\frac{1}{m} \sum_{i,j=1}^m \langle A_{ij}, \mathcal{L} A_{ij} \rangle_{L^2_{\text{KMS}}(\sigma)} = \frac{1}{m} \sum_{i,j=1}^m \mathcal{E}(A_{ij}, A_{ij}).$$

Thus, we define  $\mathcal{E}^{(m)}$  on  $(\mathcal{A} \otimes \mathbb{M}_m(\mathbb{C}), \langle \cdot, \cdot \rangle_{L^2_{\text{KMS}}(\sigma \otimes \tau_m)})$  by

$$\mathcal{E}^{(m)}(\mathbf{A}, \mathbf{A}) = \frac{1}{m} \sum_{i,j=1}^m \mathcal{E}(A_{ij}, A_{ij}) \tag{3.8}$$

where  $\mathbf{A}$  is given by (3.6). In view of Corollary 3.6,  $\mathcal{E}$  is completely Dirichlet if and only if for each  $m \in \mathbb{N}$ ,  $\mathcal{E}^{(m)}$  is Dirichlet.

A QMS  $(\mathcal{P}_t)_t$  is not only completely positive; it also satisfies  $\mathcal{P}_t \mathbf{1} = \mathbf{1}$  for all  $t$ . This too may be expressed in terms of the Dirichlet form  $\mathcal{E}$ : A Dirichlet form  $\mathcal{E}$  is *conservative* in case  $\mathcal{E}(A, \mathbf{1}) = 0$  for all  $A \in \mathcal{A}$ , and one readily sees that this is equivalent to the condition that  $\mathcal{P}_t \mathbf{1} = \mathbf{1}$  for all  $t$ .

### 3.3 Moreau Decomposition with Respect to the Cone of Positive Matrices

Let  $\mathbb{H}_n(\mathbb{C})$  denote the set of self-adjoint  $n \times n$  matrices, which contains a distinguished pointed cone  $\mathcal{P}$ , namely the cone of positive semidefinite matrices  $A$ . If we equip  $\mathbb{H}_n(\mathbb{C})$  with the Hilbert–Schmidt inner product  $\langle X, Y \rangle = \text{Tr}[XY]$ , then  $\mathcal{P}$  is self-dual: for  $X \in \mathbb{H}_n(\mathbb{C})$ ,  $\langle X, A \rangle \geq 0$  for all  $A \in \mathcal{P}$  if and only if  $\langle v, Xv \rangle \geq 0$  for all  $v \in \mathbb{C}^n$ , as one sees by considering rank one projections and using the spectral theorem.

The next result characterizes the Moreau decomposition in  $(\mathbb{H}_n(\mathbb{C}), \langle \cdot, \cdot \rangle)$  in spectral terms. For  $X \in \mathbb{H}_n(\mathbb{C})$ , there is the *spectral decomposition*  $X = X_{(+)} - X_{(-)}$  where

$$X_{(+)} = X \mathbf{1}_{(0, \infty)}(X) \quad \text{and} \quad X_{(-)} = -X \mathbf{1}_{(-\infty, 0)}(X). \tag{3.9}$$

**Theorem 3.7** (Moreau decomposition for Hilbert–Schmidt) *Let  $\mathcal{H}$  be  $\mathbb{H}_n(\mathbb{C})$  equipped with the Hilbert–Schmidt inner product, and let  $\mathcal{P}$  be the cone of positive semidefinite matrices. Then the spectral decomposition of  $X \in \mathcal{H}$  coincides with the decomposition of  $X$  into its positive and negative parts with respect to  $\mathcal{P}$ .*

**Proof** Let  $X \in \mathbb{H}_n(\mathbb{C})$ , and let  $X = X_+ - X_-$  be the decomposition determined by  $\mathcal{P}$ . Then, for  $v$  in the range of  $X_+$ , we have  $X_+ - \epsilon|v\rangle\langle v| \in \mathcal{P}$  for all sufficiently small  $\epsilon > 0$ . Therefore,

$$\|X_-\|^2 = \|X - X_+\|^2 \leq \|X - (X_+ - \epsilon|v\rangle\langle v|)\|^2 = \|X_-\|^2 - 2\epsilon\langle v, X_-v \rangle + \epsilon^2\|v\|^2.$$

It follows that  $\langle v, X_-v \rangle \leq 0$ , but since  $X_- \in \mathcal{P}$ , this yields  $\langle v, X_-v \rangle = 0$ . Hence the range of  $X_+$  lies in the null-space of  $X_-$ , so that  $X_-X_+ = 0$ . Taking the adjoint, we find that  $X_+X_- = 0$ . Therefore,  $X_-$  and  $X_+$  commute with each other, and hence with  $X$ . Thus, the projectors onto the ranges of  $X_+$  and  $X_-$  are both spectral projectors of  $X$ . Since  $X = X_+ - X_-$  it follows that  $X_+ = X_{(+)}$  and  $X_- = X_{(-)}$ .  $\square$

The situation is more interesting for other inner products on  $\mathbb{H}_n(\mathbb{C})$ . Let  $\sigma$  be an invertible density matrix. For  $s \in [0, 1]$ , let  $\langle \cdot, \cdot \rangle_s$  be the inner product on  $\mathbb{M}_n(\mathbb{C})$  given by  $\langle A, B \rangle_s = \text{Tr}[A^* \sigma^s B \sigma^{1-s}]$ .

**Theorem 3.8** *Let  $\sigma$  be an invertible  $n \times n$  density matrix that is not a multiple of the identity. Then the cone  $\mathcal{P}$  of positive matrices in  $\mathbb{H}_n(\mathbb{C})$  is self-dual with respect to the inner product  $\langle \cdot, \cdot \rangle_s$  determined by  $\sigma$  if and only if  $s = \frac{1}{2}$ .*

**Proof** Let  $X \in \mathbb{H}_n(\mathbb{C})$  and  $A \in \mathcal{P}$ . Then  $\langle X, A \rangle_s = \text{Tr}[X \sigma^s A \sigma^{1-s}] = \text{Tr}[(\sigma^{1-s} X \sigma^s) A]$ . Therefore,  $\langle X, A \rangle_s \geq 0$  for all  $A \in \mathcal{P}$  if and only if  $\sigma^{1-s} X \sigma^s \in \mathcal{P}$ . If  $\sigma^{1-s} X \sigma^s \in \mathcal{P}$ , then  $\sigma^{1-s} X \sigma^s$  is self-adjoint, and hence  $\sigma^{1-s} X \sigma^s = \sigma^s X \sigma^{1-s}$ , or, what is the same,  $[\sigma^{1-2s}, X] = 0$ . Let  $X := |v\rangle\langle v|$  with  $v$  chosen *not* to be an eigenvector of  $\sigma$ . Then for  $s \neq \frac{1}{2}$ ,  $[\sigma^{1-2s}, X] \neq 0$ . Therefore,  $X \in \mathcal{P}$ , but  $X \notin \mathcal{P}^\circ$ . Hence,  $\mathcal{P}$  is not self-dual when  $\mathcal{H}$  is equipped with the inner product  $\langle \cdot, \cdot \rangle_s$  for  $s \neq \frac{1}{2}$ .

One the other hand,

$$\langle X, A \rangle_{1/2} = \text{Tr}[X \sigma^{1/2} A \sigma^{1/2}] = \text{Tr}[(\sigma^{1/4} X \sigma^{1/4})(\sigma^{1/4} A \sigma^{1/4})].$$

Since  $\sigma$  is invertible, as  $A$  ranges over  $\mathcal{P}$ ,  $\sigma^{1/4} A \sigma^{1/4}$  ranges over  $\mathcal{P}$ , and so  $\langle X, A \rangle_{1/2} \geq 0$  for all  $A \in \mathcal{P}$  if and only if  $\sigma^{1/4} X \sigma^{1/4} \in \mathcal{P}$ . Again, since  $\sigma$  is invertible, this is the case if and only if  $X \in \mathcal{P}$ . Hence,  $\mathcal{P}$  is self-dual for  $\langle \cdot, \cdot \rangle_{1/2}$ , the KMS inner product.  $\square$

The Moreau decomposition for the KMS scalar product can easily be obtained from Theorem 3.7 by a unitary transformation.

**Theorem 3.9** (Moreau decomposition for KMS) *Let  $\sigma$  be an invertible  $n \times n$  density matrix and let  $X \in \mathbb{H}_n(\mathbb{C})$ . Then, with respect to the KMS norm on  $\mathbb{H}_n(\mathbb{C})$ ,*

$$\|X - \sigma^{-1/4}(\sigma^{1/4} X \sigma^{1/4})_{(+)} \sigma^{-1/4}\|_{L^2_{\text{KMS}}(\sigma)} \leq \|X - A\|_{L^2_{\text{KMS}}(\sigma)} \tag{3.10}$$

for all  $A \in \mathcal{P}$ . Consequently, the positive part of  $X$  in the decomposition according to  $\mathcal{P}$ ,  $X_+$ , is given by

$$X_+ = \sigma^{-1/4}(\sigma^{1/4} X \sigma^{1/4})_{(+)} \sigma^{-1/4}. \tag{3.11}$$

**Proof** The map  $Y \mapsto \sigma^{1/4} Y \sigma^{1/4}$  is unitary from  $\mathbb{H}_n(\mathbb{C})$  equipped with the KMS inner product to  $\mathbb{H}_n(\mathbb{C})$  equipped with the Hilbert–Schmidt inner product. That is,

$$\|X - A\|_{L^2_{\text{KMS}}(\sigma)}^2 = \text{Tr}[\sigma^{1/4} X \sigma^{1/4} - \sigma^{1/4} A \sigma^{1/4}]^2$$

for  $X, A \in \mathbb{H}_n(\mathbb{C})$ . By Theorem 3.7,  $\min\{\text{Tr}[\sigma^{1/4} X \sigma^{1/4} - B]^2 : B \in \mathcal{P}\}$  is achieved at  $B = (\sigma^{1/4} X \sigma^{1/4})_{(+)}$ .  $\square$



We conclude the section by extending the results above to an arbitrary  $*$ -subalgebra  $\mathcal{A}$  of  $\mathbb{M}_n(\mathbb{C})$ . Let  $\sigma$  be an invertible  $n \times n$  density matrix belonging to  $\mathcal{A}$ .

**Lemma 3.10** *Let  $\mathcal{H}$  be  $\mathcal{A}_h$  equipped with the KMS inner product induced by  $\sigma$ , and let  $\mathcal{P}$  be the positive matrices in  $\mathbb{M}_n(\mathbb{C})$ , and let  $\mathcal{P}_{\mathcal{A}} := \mathcal{P} \cap \mathcal{A}$ . Then  $\mathcal{P}_{\mathcal{A}}$  is self-dual in  $\mathcal{H}$ .*

**Proof** Let  $X \in \mathcal{P}_{\mathcal{A}}$ . For any  $A \in \mathcal{P}_{\mathcal{A}}$  we have  $\sigma^{1/2}A\sigma^{1/2} \geq 0$ , hence  $\langle X, A \rangle_{L^2_{\text{KMS}}(\sigma)} = \text{Tr}[X\sigma^{1/2}A\sigma^{1/2}] \geq 0$ , which shows that  $X \in \mathcal{P}_{\mathcal{A}}^{\circ}$ .

Conversely, suppose that  $X \in \mathcal{A}_h$  belongs to  $\mathcal{P}_{\mathcal{A}}^{\circ}$ . For every  $A \in \mathcal{P}_{\mathcal{A}}$  we then have  $\text{Tr}[X\sigma^{1/2}A\sigma^{1/2}] = \langle X, A \rangle_{L^2_{\text{KMS}}(\sigma)} \geq 0$ . Since  $\sigma$  is invertible, it follows that  $\text{Tr}[XB] \geq 0$  for every  $B \in \mathcal{P}_{\mathcal{A}}$ . Therefore, the spectrum of  $X$  is non-negative, which implies that  $X$  belongs to  $\mathcal{P}$  and hence to  $\mathcal{P}_{\mathcal{A}}$ .  $\square$

**Lemma 3.11** *Let  $X$  be a self-adjoint element of  $\mathcal{A}$ . Then the decomposition of  $X$  with respect to  $\mathcal{P}_{\mathcal{A}}$  is given by  $X = X_+ - X_-$  where*

$$X_+ := \sigma^{-1/4}(\sigma^{1/4}X\sigma^{1/4})_{(+)}\sigma^{-1/4} \quad \text{and} \quad X_- := \sigma^{-1/4}(\sigma^{1/4}X\sigma^{1/4})_{(-)}\sigma^{-1/4} .$$

**Proof** Let  $X$  be a self-adjoint element of  $\mathcal{A}$ . Then by Theorem 3.9,  $\min\{\|X - A\|_{L^2_{\text{KMS}}(\sigma)} : A \in \mathcal{P}\}$  is achieved at  $A = \sigma^{-1/4}(\sigma^{1/4}X\sigma^{1/4})_{(+)}\sigma^{-1/4}$ , and since this belongs to  $\mathcal{A}$ , this same choice of  $A$  also achieves the minimum in  $\min\{\|X - A\|_{L^2_{\text{KMS}}(\sigma)} : A \in \mathcal{P}_{\mathcal{A}}\}$ .  $\square$

### 4 Construction of Dirichlet Forms on a Finite-Dimensional von Neumann Algebra

Motivated by the results in Sects. 2 and 3 we introduce a general framework in which various gradient flow structures can be studied naturally. This setting unifies and extends several previous approaches to gradient flows, in particular for reversible Markov chains on finite spaces [32,35], the fermionic Fokker-Planck equation [8], and Lindblad equations with detailed balance [10,36]

While the results in Sect. 2 show that the general QMS satisfying the  $\sigma$ -DBC can be represented in terms of a Dirichlet form specified in terms of derivations, our applications require us to work with representations for the generator  $\mathcal{L}$  in terms of “partial derivative operators”  $\partial_j$  that are not simply derivations. The reason is that, to obtain functional inequalities and sharp rates of convergence to equilibrium, it will be important to obtain commutation relations of the form  $[\partial_j, \mathcal{L}] = -a\partial_j$  for  $a \in \mathbb{R}$ . We shall demonstrate that such commutation relations may hold for the general class of representations introduced in this section, but not for the simpler representation in terms of derivations discussed in Sect. 2.

Our starting point is a finite-dimensional von Neumann algebra  $\mathcal{A}$  which we may regard as a subalgebra of  $\mathbb{M}_n(\mathbb{C})$  for some  $n \in \mathbb{N}$ . On account of the finite-dimensionality of  $\mathcal{A}$ , there is always a tracial positive linear functional  $\tau$  on  $\mathcal{A}$ : One choice is the normalized trace  $\tau[A] = n^{-1} \text{Tr}[A]$ . However, if  $\mathcal{A}$  is commutative (hence isomorphic to  $\ell_n^{\infty}$ ), there will be many other tracial positive linear functionals — any positive measure on  $\{1, \dots, n\}$  specifies such a positive linear functional. In what follows,  $\tau$  will denote any faithful positive linear functional on  $\mathcal{A}$  that is tracial; i.e., such that  $\tau[AB] = \tau[BA]$  for all  $A, B \in \mathcal{A}$ . Since  $\tau$  is faithful, every state  $\sigma$  on  $\mathcal{A}$  can be represented as  $\sigma(A) = \tau[\sigma A]$ , where on the right side  $\sigma \in \mathcal{A} \subseteq \mathbb{M}_n(\mathbb{C})$  is the  $n \times n$  density matrix belonging to  $\mathcal{A}$  determined by the state  $\sigma$ .

The basic operation in terms of which we shall construct completely Dirichlet forms on  $\mathcal{A}$  has several components.

Let  $\mathcal{B}$  be another finite-dimensional von Neumann algebra with tracial state  $\tau_{\mathcal{B}}$ . A unital  $*$ -homomorphism  $\ell$  from  $(\mathcal{A}, \tau)$  to  $(\mathcal{B}, \tau_{\mathcal{B}})$  is  $(\tau, \tau_{\mathcal{B}})$ -compatible in case for all  $A \in \mathcal{A}$ ,

$$\tau_{\mathcal{B}}[\ell(A)] = \tau[A]. \tag{4.1}$$

Equivalently,  $\ell$  is  $(\tau, \tau_{\mathcal{B}})$ -compatible in case its adjoint  $\ell^\dagger : L^2(\mathcal{B}, \tau_{\mathcal{B}}) \rightarrow L^2(\mathcal{A}, \tau)$  satisfies  $\ell^\dagger(\mathbf{1}_{\mathcal{B}}) = \mathbf{1}_{\mathcal{A}}$ .

Let  $0 \neq V \in \mathcal{B}$ , and let  $\ell$  and  $r$  be a pair of  $(\tau, \tau_{\mathcal{B}})$ -compatible unital  $*$ -homomorphisms from  $\mathcal{A}$  into  $\mathcal{B}$ . Then define the operator  $\partial_V : \mathcal{A} \rightarrow \mathcal{B}$  by

$$\partial_V A := Vr(A) - \ell(A)V. \tag{4.2}$$

If  $\mathcal{B} = \mathcal{A}$  and both  $\ell$  and  $r$  are the identity, this reduces to (2.11). The following Leibniz rule shows that  $\partial_V$  is an  $(\ell, r)$ -skew derivation.

**Lemma 4.1** (Leibniz rule for  $\partial_V$ ) *For  $A, B \in \mathcal{A}$  we have*

$$\partial_V(AB) = (\partial_V A)r(B) + \ell(A)\partial_V B. \tag{4.3}$$

**Proof** Since  $\ell$  and  $r$  are algebra homomorphisms,

$$\begin{aligned} \partial_V(AB) &= Vr(AB) - \ell(AB)V \\ &= (Vr(A) - \ell(A)V)r(B) + \ell(A)(Vr(B) - \ell(B)V) = (\partial_V A)r(B) + \ell(A)\partial_V B, \end{aligned}$$

which is the desired identity. □

**Remark 4.2** Since  $\ell$  and  $r$  are algebra  $*$ -homomorphisms, it follows that

$$\ell^\dagger(\ell(A_1)B\ell(A_2)) = A_1\ell^\dagger(B)A_2 \quad \text{and} \quad r^\dagger(r(A_1)Br(A_2)) = A_1r^\dagger(B)A_2 \tag{4.4}$$

for all  $A_1, A_2 \in \mathcal{A}$  and  $B \in \mathcal{B}$ . Moreover,  $\ell^\dagger(B)^* = \ell^\dagger(B^*)$  and  $r^\dagger(B)^* = r^\dagger(B^*)$  for all  $B \in \mathcal{B}$ .

Let  $\sigma \in \mathcal{A}$  be the density matrix (with respect to  $\tau$ ) of a faithful state on  $\mathcal{A}$ . Since  $\ell$  and  $r$  are  $(\tau, \tau_{\mathcal{B}})$ -compatible,  $\ell(\sigma)$  and  $r(\sigma)$  are density matrices (with respect to  $\tau_{\mathcal{B}}$  on  $\mathcal{B}$ ). The inner product that we use on  $\mathcal{B}$  is a KMS inner product based on both  $\ell(\sigma)$  and  $r(\sigma)$  defined in terms of the relative modular operator  $\Delta_{\ell(\sigma), r(\sigma)}$ :

$$\Delta_{\ell(\sigma), r(\sigma)}(B) := \ell(\sigma)Br(\sigma)^{-1}. \tag{4.5}$$

It is easily verified that  $\Delta_{\ell(\sigma), r(\sigma)}$  is a positive operator on  $L^2(\mathcal{B}, \tau_{\mathcal{B}})$ , and hence we may define an inner product on  $\mathcal{B}$  through

$$\begin{aligned} \langle B_1, B_2 \rangle_{L^2_{\text{KMS}}(\mathcal{B}, \ell(\sigma), r(\sigma))} &:= \langle B_1r(\sigma)^{1/2}, \Delta_{\ell(\sigma), r(\sigma)}^{1/2}(B_2r(\sigma)^{1/2}) \rangle_{L^2(\mathcal{B}, \tau_{\mathcal{B}})} \\ &= \tau_{\mathcal{B}}[B_1^*\ell(\sigma^{1/2})B_2r(\sigma^{1/2})]. \end{aligned}$$

Given a faithful state  $\sigma$  on  $\mathcal{A}$ ,  $V \in \mathcal{B}$ , and two pairs  $(\ell, r)$  and  $(\ell_*, r_*)$  of  $(\tau, \tau_{\mathcal{B}})$ -compatible  $*$ -homomorphisms of  $\mathcal{A}$  into  $\mathcal{B}$ , define  $\partial_V$  by (4.2), and define

$$\partial_{V^*} = V^*r_*(A) - \ell_*(A)V^*$$

in accordance with (4.2), but using  $V^*$ ,  $\ell_*$  and  $r_*$ . Then define a sesquilinear form  $\mathcal{E}$  on  $\mathcal{A}$  by

$$\mathcal{E}(A_1, A_2) = \langle \partial_V A_1, \partial_V A_2 \rangle_{L^2_{\text{KMS}}(\mathcal{B}, \ell(\sigma), r(\sigma))} + \langle \partial_{V^*} A_1, \partial_{V^*} A_2 \rangle_{L^2_{\text{KMS}}(\mathcal{B}, \ell_*(\sigma), r_*(\sigma))}. \tag{4.6}$$

Our immediate goal in this section is to determine conditions on  $V$ ,  $(\ell, r)$  and  $(\ell_*, r_*)$  under which  $\mathcal{E}$  is a conservative completely Dirichlet form on  $\mathcal{A}$  equipped with the KMS inner product induced by  $\sigma$ .

It is first of all necessary that the operator  $\mathcal{L}$  determined by  $\mathcal{E}$  through  $\mathcal{E}(A_1, A_2) = -\langle B, \mathcal{L}A \rangle_{L^2_{\text{KMS}}(\sigma)}$  be real; i.e.,  $(\mathcal{L}(A))^* = \mathcal{L}A^*$ . Since  $\langle A_1, A_2 \rangle_{L^2_{\text{KMS}}(\sigma)} = \langle A_2^*, A_1^* \rangle_{L^2_{\text{KMS}}(\sigma)}$  for all  $A_1, A_2 \in \mathcal{A}$ , it is easily seen that  $\mathcal{L}$  is real if and only if  $\mathcal{E}(A_1, A_2) = \mathcal{E}(A_2^*, A_1^*)$  for all  $A_1, A_2 \in \mathcal{A}$ .

**Lemma 4.3** *Under the condition that for all  $A_1, A_2 \in \mathcal{A}$ ,*

$$\tau_{\mathcal{B}}[V^* \ell(A_1) V r(A_2)] = \tau_{\mathcal{B}}[V^* r_*(A_1) V \ell_*(A_2)], \tag{4.7}$$

*we have  $\mathcal{E}(A_1, A_2) = \mathcal{E}(A_2^*, A_1^*)$  for all  $A_1, A_2 \in \mathcal{A}$ .*

**Remark 4.4** One can satisfy (4.7) in a trivial way by taking  $\ell, r, \ell_*$  and  $r_*$  each to be the identity. Almost as trivially, one may take  $\ell_* = r$  and  $r_* = \ell$ . However, we shall see that one can also satisfy (4.7) with  $\ell_* = \ell$  and  $r_* = r = I_{\mathcal{B}}$  with a non-trivial  $*$ -homomorphism  $\ell$ ; see the discussion in the next section on the Clifford Dirichlet form. Other non-trivial realizations of (4.7) arise in practice.

**Proof of Lemma 4.3** We compute

$$\langle \partial_V A_1, \partial_V A_2 \rangle_{L^2_{\text{KMS}}(\mathcal{B}, \ell(\sigma), r(\sigma))} = \tau_{\mathcal{B}}[r(A_1^*) V^* \ell(\sigma^{1/2}) V r(A_2) r(\sigma^{1/2})] \tag{4.8}$$

$$\begin{aligned} &+ \tau_{\mathcal{B}}[V^* \ell(A_1^*) \ell(\sigma^{1/2}) \ell(A_2) V r(\sigma^{1/2})] \\ &- \tau_{\mathcal{B}}[r(A_1^*) V^* \ell(\sigma^{1/2}) \ell(A_2) V r(\sigma^{1/2})] \\ &- \tau_{\mathcal{B}}[V^* \ell(A_1^*) \ell(\sigma^{1/2}) V r(A_2) r(\sigma^{1/2})]. \end{aligned} \tag{4.9}$$

By cyclicity of the trace  $\tau_{\mathcal{B}}$ , the homomorphism property of  $\ell$  and  $r$ , and (4.7),

$$\begin{aligned} \tau_{\mathcal{B}}[r(A_1^*) V^* \ell(\sigma^{1/2}) V r(A_2) r(\sigma^{1/2})] &= \tau_{\mathcal{B}}[r(A_2 \sigma^{1/2} A_1^*) V^* \ell(\sigma^{1/2}) V] \\ &= \tau_{\mathcal{B}}[\ell_*(A_2 \sigma^{1/2} A_1^*) V^* r_*(\sigma^{1/2}) V] \\ &= \tau_{\mathcal{B}}[V \ell_*(A_2) \ell_*(\sigma^{1/2}) \ell_*(A_1^*) V^* r_*(\sigma^{1/2})]. \end{aligned}$$

This shows that the quantity in (4.8) is what we obtain from the quantity in (4.9) if we replace  $\ell$  by  $\ell_*$ ,  $r$  by  $r_*$ ,  $V$  by  $V^*$ ,  $A_1$  by  $A_2^*$ , and  $A_2$  by  $A_1^*$ . Similar computations then yield the identity

$$\langle \partial_V A_1, \partial_V A_2 \rangle_{L^2_{\text{KMS}}(\mathcal{B}, \ell(\sigma), r(\sigma))} = \langle \partial_{V^*} A_2^*, \partial_{V^*} A_1^* \rangle_{L^2_{\text{KMS}}(\mathcal{B}, \ell_*(\sigma), r_*(\sigma))},$$

and this implies  $\mathcal{E}(A_1, A_2) = \mathcal{E}(A_2^*, A_1^*)$ . □

Thus, the condition (4.7) suffices to ensure that the sesquilinear form  $\mathcal{E}$  defined in (4.6) is real. In the rest of this section, we suppose that this condition is satisfied, and then since  $\mathcal{E}$  is real, it suffices to consider its bilinear restriction to  $\mathcal{A}_h$ .

One further condition is required to ensure that  $\mathcal{E}$  be a Dirichlet form on  $\mathcal{A}_h$ , and we shall see that under this same condition  $\mathcal{E}$  is actually a completely Dirichlet form. The assumption is that  $V$  (resp.  $V^*$ ) is an eigenvector of the relative modular operator  $\Delta_{\ell(\sigma), r(\sigma)}$  (resp.  $\Delta_{\ell_*(\sigma), r_*(\sigma)}$ ). Since the relative modular operator is positive, there exist  $\omega, \omega_* \in \mathbb{R}$  such that

$$\Delta_{\ell(\sigma), r(\sigma)} V = e^{-\omega} V \quad \text{and} \quad \Delta_{\ell_*(\sigma), r_*(\sigma)} V^* = e^{-\omega_*} V^*. \tag{4.10}$$

There are several equivalent formulations of this condition that will be useful.

**Lemma 4.5** *The first condition in (4.10) is equivalent to the condition*

$$\partial_V \log \sigma = \omega V, \tag{4.11}$$

and to the condition that for all  $t \in \mathbb{R}$ ,

$$\Delta_{\ell(\sigma), r(\sigma)}^t V = e^{-t\omega} V. \tag{4.12}$$

Moreover, (4.10) implies that

$$\omega_* = -\omega. \tag{4.13}$$

**Proof** Note that  $(\Delta_{\ell(\sigma), r(\sigma)}^t)_{t \in \mathbb{R}}$  is a group of linear operators on  $\mathcal{B}$ , and the generator  $\mathcal{G}$  of this group is given by  $\mathcal{G}B = \ell(\log \sigma)B - Br(\log \sigma)$ , thus  $\mathcal{G}V = -\partial_V \log \sigma$ . The equivalences thus follow from basic spectral theory.

Using (4.7) with  $A_1 = \sigma$  and  $A_2 = \sigma^{-1}$ , and two applications of (4.12), we obtain

$$e^{-\omega} \tau_{\mathcal{B}}[V^*V] = \tau_{\mathcal{B}}[V^* \ell(\sigma) V r(\sigma^{-1})] = \tau_{\mathcal{B}}[V \ell_*(\sigma^{-1}) V^* r_*(\sigma)] = e^{\omega_*} \tau_{\mathcal{B}}[V V^*].$$

Since  $V \neq 0$ , this yields (4.13). □

We are now ready to state the main result of this section.

**Theorem 4.6** *Let  $\sigma$  be a faithful state on  $\mathcal{A}$ . Let  $V \in \mathcal{B}$  and two pairs  $(\ell, r)$  and  $(\ell_*, r_*)$  of  $(\tau, \tau_{\mathcal{B}})$ -compatible  $*$ -homomorphisms be given. Suppose also that (4.7) is satisfied, and suppose that  $V$  (resp.  $V^*$ ) is an eigenvector of the relative modular operator  $\Delta_{\ell(\sigma), r(\sigma)}$  (resp.  $\Delta_{\ell_*(\sigma), r_*(\sigma)}$ ) satisfying (4.10). Then the sesquilinear form  $\mathcal{E} : \mathcal{A} \times \mathcal{A} \rightarrow \mathbb{C}$  given by (4.6) defines a conservative completely Dirichlet form on  $L^2_{\text{KMS}}(\mathcal{A}_h, \sigma)$ .*

**Proof** To explain the crucial role of the assumption that  $V$  is an eigenvector of the relative modular operator, so that (4.10) is satisfied, we fix  $V, W \in \mathcal{B}$  and (temporarily) define the operators  $\partial, \partial_* : \mathcal{A} \rightarrow \mathcal{B}$  by  $\partial A := Vr(A) - \ell(A)W$  and  $\partial_* A := V^*r_*(A) - \ell_*(A)W^*$ , and set

$$\mathcal{E}(A_1, A_2) = \langle \partial A_1, \partial A_2 \rangle_{L^2_{\text{KMS}}(\mathcal{B}, \ell(\sigma), r(\sigma))} + \langle \partial_* A_1, \partial_* A_2 \rangle_{L^2_{\text{KMS}}(\mathcal{B}, \ell_*(\sigma), r_*(\sigma))}.$$

We will show:

- (1) If  $W = e^{\omega/2} \Delta_{\ell(\sigma), r(\sigma)}^{1/2} V$  and  $W^* = e^{\omega_*/2} \Delta_{\ell_*(\sigma), r_*(\sigma)}^{1/2} V^*$  for some  $\omega, \omega_* \in \mathbb{R}$ , then  $\mathcal{E}$  defines a Dirichlet form on  $L^2_{\text{KMS}}(\mathcal{A}_h, \sigma)$ .
- (2) If, in addition, (4.10) holds, then  $\mathcal{E}(\mathbf{1}, A) = 0$  for all  $A \in \mathcal{A}_h$ , hence  $\mathcal{E}$  is conservative.

Consider the unitary transformation  $\mathcal{U} : L^2_{\text{KMS}}(\mathcal{A}, \sigma) \rightarrow L^2(\mathcal{A}, \tau)$  given by  $\mathcal{U}A := \sigma^{1/4} A \sigma^{1/4}$ . For brevity we write  $\mathcal{T}B := \Delta_{\ell(\sigma), r(\sigma)}^{1/4} B = \ell(\sigma^{1/4}) B r(\sigma^{-1/4})$ , and likewise,  $\mathcal{T}_*B := \Delta_{\ell_*(\sigma), r_*(\sigma)}^{1/4} B = \ell_*(\sigma^{1/4}) B r_*(\sigma^{-1/4})$ .

For  $A \in \mathcal{A}_h$  we need to show that  $\mathcal{E}(A_+, A_-) \leq 0$ . For  $A_1, A_2 \in \mathcal{A}$  we have

$$\begin{aligned} & \langle \partial A_1, \partial A_2 \rangle_{L^2_{\text{KMS}}(\mathcal{B}, \ell(\sigma), r(\sigma))} \\ &= \tau_{\mathcal{B}} \left[ r(\sigma^{1/4}) \left( r(A_1^*) V^* - W^* \ell(A_1^*) \right) \ell(\sigma^{1/2}) \left( Vr(A_2) - \ell(A_2)W \right) r(\sigma^{1/4}) \right] \\ &= \tau_{\mathcal{B}} \left[ \left( r(\mathcal{U}A_1^*) (\mathcal{T}V)^* - (\mathcal{T}^{-1}W)^* \ell(\mathcal{U}A_1^*) \right) \left( (\mathcal{T}V)r(\mathcal{U}A_2) - \ell(\mathcal{U}A_2)\mathcal{T}^{-1}W \right) \right] \end{aligned}$$

$$\begin{aligned}
 &= \tau_{\mathcal{B}} \left[ (TV)r(\mathcal{U}A_2\mathcal{U}A_1^*)(TV)^* - (TV)^*\ell(\mathcal{U}A_2)T^{-1}Wr(\mathcal{U}A_1^*) \right. \\
 &\quad \left. - r(\mathcal{U}A_2)(T^{-1}W)^*\ell(\mathcal{U}A_1^*)(TV) + (T^{-1}W)^*\ell(\mathcal{U}A_1^*\mathcal{U}A_2)T^{-1}W \right]. \tag{4.14}
 \end{aligned}$$

For  $A \in \mathcal{A}_h$  we have  $A_{\pm} = \mathcal{U}^{-1}(\mathcal{U}A)_{(\pm)}$  by Lemma 3.11, thus

$$\mathcal{U}A_+\mathcal{U}A_- = (\mathcal{U}A)_{(+)}(\mathcal{U}A)_{(-)} = 0 \quad \text{and} \quad \mathcal{U}A_-\mathcal{U}A_+ = (\mathcal{U}A)_{(-)}(\mathcal{U}A)_{(+)} = 0.$$

We obtain

$$\begin{aligned}
 \langle \partial A_+, \partial A_- \rangle_{L^2_{\text{KMS}}(\mathcal{B}, \ell(\sigma), r(\sigma))} &= -\tau_{\mathcal{B}} \left[ (TV)^*\ell(\mathcal{U}A_-)T^{-1}Wr(\mathcal{U}A_+) \right. \\
 &\quad \left. + (T^{-1}W)^*\ell(\mathcal{U}A_+)(TV)r(\mathcal{U}A_-) \right]. \tag{4.15}
 \end{aligned}$$

Since  $r(\mathcal{U}A_{\pm}) \geq 0$ , it follows that  $\langle \partial A_+, \partial A_- \rangle_{L^2_{\text{KMS}}(\mathcal{B}, \ell(\sigma), r(\sigma))} \leq 0$  if we can show that

$$(TV)^*\ell(\mathcal{U}A_-)T^{-1}W \geq 0 \quad \text{and} \quad (T^{-1}W)^*\ell(\mathcal{U}A_+)(TV) \geq 0. \tag{4.16}$$

To show this, we make the assumption that  $W = e^{\omega/2} \Delta_{\ell(\sigma), r(\sigma)}^{1/2} V$  for some  $\omega \in \mathbb{R}$ . Equivalently, this means that  $T^{-1}W = e^{\omega/2} TV$ , and since  $\ell(\mathcal{U}A_{\pm}) \geq 0$ , we obtain (4.16). This proves that  $\langle \partial A_+, \partial A_- \rangle_{L^2_{\text{KMS}}(\mathcal{B}, \ell(\sigma), r(\sigma))} \leq 0$ .

An entirely analogous argument shows that  $\langle \partial_* A_+, \partial_* A_- \rangle_{L^2_{\text{KMS}}(\mathcal{B}, \ell_*(\sigma), r_*(\sigma))} \leq 0$ , and this proves that  $\mathcal{E}(A_+, A_-)$  is a Dirichlet form.

Observe now that  $\partial \mathbf{1} = V - W$  and  $\partial_* \mathbf{1} = V^* - W^*$ . Thus, to conclude that  $\partial \mathbf{1} = \partial_* \mathbf{1} = 0$ , we need to assume that  $V$  is an eigenvector of  $\Delta_{\ell(\sigma), r(\sigma)}$  with eigenvalue  $e^{-\omega}$ , and that  $V^*$  is an eigenvector of  $\Delta_{\ell_*(\sigma), r_*(\sigma)}$  with eigenvalue  $e^{-\omega^*}$ . It immediately follows that  $\mathcal{E}(\mathbf{1}, A) = 0$  for all  $A \in \mathcal{A}_h$ , hence  $\mathcal{E}$  is conservative.

It remains to prove that under the given conditions,  $\mathcal{E}$  is completely Dirichlet. Let  $\text{Tr}$  be the standard trace on  $\mathbb{M}_m(\mathbb{C})$ . Let  $\mathbf{H}$  be a self-adjoint element of  $\mathcal{A} \otimes \mathbb{M}_m(\mathbb{C})$ , and let  $\mathbf{H}_+$  and  $\mathbf{H}_-$  be the elements of its decomposition  $\mathbf{H} = \mathbf{H}_+ - \mathbf{H}_-$  in  $L^2_{\text{KMS}}(\sigma \otimes \text{Tr})$ , where  $\mathbf{H}_+$  and  $\mathbf{H}_-$  are positive and such that  $\langle \mathbf{H}_+, \mathbf{H}_- \rangle_{L^2_{\text{KMS}}(\sigma \otimes \text{Tr})} = 0$ .

Let  $\sigma = \sum_{j=1}^m \sigma \otimes E_{jj}$  and write  $\tilde{\mathbf{H}} = \sigma^{1/4} \mathbf{H} \sigma^{1/4}$  for brevity. By Theorem 3.9,  $\mathbf{H}_+ = \sigma^{-1/4} \tilde{\mathbf{H}}_{(+)} \sigma^{-1/4}$ , hence  $[\mathbf{H}_+]_{ij} = \sigma^{-1/4} [\tilde{\mathbf{H}}_{(+)}]_{ij} \sigma^{-1/4}$ . It follows that

$$\sum_{i,j=1}^m \tau \left[ \mathcal{U}([\mathbf{H}_+]_{ij}) \mathcal{U}([\mathbf{H}_-]_{ij}) \right] = \langle \tilde{\mathbf{H}}_{(+)}, \tilde{\mathbf{H}}_{(-)} \rangle_{L^2(\tau \otimes \text{Tr})} = 0.$$

Using this identity, (4.15) with  $V = W$  yields

$$\begin{aligned}
 &\sum_{i,j=1}^m \langle \partial[\mathbf{H}_+]_{ij}, \partial[\mathbf{H}_-]_{ij} \rangle_{L^2_{\text{KMS}}(\mathcal{B}, \ell(\sigma), r(\sigma))} \\
 &= - \sum_{i,j=1}^m \tau_{\mathcal{B}} \left[ V^* \ell(\mathcal{U}[\mathbf{H}_-]_{ij}) Vr(\mathcal{U}[\mathbf{H}_+]_{ij}) + V^* \ell(\mathcal{U}[\mathbf{H}_+]_{ij}) Vr(\mathcal{U}[\mathbf{H}_-]_{ij}) \right] \\
 &= - \sum_{i,j=1}^m \tau_{\mathcal{B}} \left[ V^* \ell([\tilde{\mathbf{H}}_{(-)}]_{ij}) Vr([\tilde{\mathbf{H}}_{(+)}]_{ij}) + V^* \ell([\tilde{\mathbf{H}}_{(+)}]_{ij}) Vr([\tilde{\mathbf{H}}_{(-)}]_{ij}) \right] \\
 &= -\tau_{\mathcal{B}} \otimes \text{Tr} \left[ (V \otimes \mathbf{1}_m)^* \ell(\tilde{\mathbf{H}}_{(-)}) (V \otimes \mathbf{1}_m) r(\tilde{\mathbf{H}}_{(+)}) + (V \otimes \mathbf{1}_m)^* \ell(\tilde{\mathbf{H}}_{(+)}) (V \otimes \mathbf{1}_m) r(\tilde{\mathbf{H}}_{(-)}) \right],
 \end{aligned}$$

where  $\mathbf{1}_m$  denotes the identity matrix in  $\mathbb{M}_m(\mathbb{C})$ , and in the last line, we simply write  $\ell$  and  $r$  to denote their canonical extensions  $\ell \otimes I$  and  $r \otimes I$ . Since  $r(\tilde{\mathbf{H}}_{(\pm)}) \geq 0$  and  $(V \otimes$

$I)^*\ell(\tilde{\mathbf{H}}_{(\mp)}) (V \otimes \mathbf{1}_m) \geq 0$ , it is now evident that the right-hand side is non-positive. An analogous argument applies if we replace  $\partial$  by  $\partial_*$ , and therefore,

$$\mathcal{E}^{(m)}(\mathbf{H}_+, \mathbf{H}_-) = \sum_{i,j=1}^m \mathcal{E}([\mathbf{H}_+]_{ij}, [\mathbf{H}_-]_{ij}) \leq 0 .$$

In summary, this proves that  $\mathcal{E}^{(m)}$  is a Dirichlet form for all  $m \in \mathbb{N}$ , and hence that  $\mathcal{E}$  is completely Dirichlet. □

Evidently, the sum of a finite set of conservative completely Dirichlet forms on  $\mathcal{A}$  is a conservative completely Dirichlet form. Thus, we may construct a large class of conservative completely Dirichlet forms by taking sums of forms of the type considered in Theorem 4.6. In the remainder of this section, we consider such a conservative, completely Dirichlet form and the associated QMS  $\mathcal{P}_t$ .

It will be convenient going forward to streamline our notation. In the rest of this section we are working in the framework specified as follows:

**Definition 4.7** Let  $\mathcal{A}$  be a finite-dimensional von Neumann algebra  $\mathcal{A}$  endowed with a faithful tracial positive linear functional  $\tau$ . A *differential structure* on  $\mathcal{A}$  consists of the following:

- (1) A finite index set  $\mathcal{J}$ , and for each  $j \in \mathcal{J}$ , a finite dimensional von Neumann algebra  $\mathcal{B}_j$  endowed with a faithful tracial positive linear functional  $\tau_j$ .
- (2) For each  $j \in \mathcal{J}$ , a pair  $(\ell_j, r_j)$  of unital  $*$ -homomorphisms from  $\mathcal{A}$  to  $\mathcal{B}_j$  such that for each  $A \in \mathcal{A}$  and each  $j \in \mathcal{J}$ ,  $\tau_j(\ell_j(A)) = \tau_j(r_j(A)) = \tau(A)$ , and a non-zero  $V_j \in \mathcal{B}_j$ .
- (3) It is further required that for each  $j \in \mathcal{J}$ , there is a unique  $j^*$  such that  $V_j^* = V_{j^*}$ , hence  $\{V_j\}_{j \in \mathcal{J}} = \{V_{j^*}\}_{j \in \mathcal{J}}$  and  $\mathcal{B}_{j^*} = \mathcal{B}_j$ . Moreover, for  $j \in \mathcal{J}$  and  $A_1, A_2 \in \mathcal{A}$ ,

$$\tau_j[V_j^* \ell_j(A_1) V_j r_j(A_2)] = \tau_j[V_{j^*}^* r_{j^*}(A_1) V_j \ell_{j^*}(A_2)] . \tag{4.17}$$

- (4) An invertible density matrix  $\sigma \in \mathfrak{P}_+$ , such that, for each  $j \in \mathcal{J}$ ,  $V_j$  is an eigenvector of the relative modular operator  $\Delta_{\ell_j(\sigma), r_j(\sigma)}$  on  $\mathcal{B}_j$  with

$$\Delta_{\ell_j(\sigma), r_j(\sigma)}(V_j) = e^{-\omega_j} V_j \tag{4.18}$$

for some  $\omega_j \in \mathbb{R}$ .

Then for each  $j \in \mathcal{J}$ , we define the linear operator  $\partial_j : \mathcal{A} \rightarrow \mathcal{B}_j$  by

$$\partial_j A := V_j r_j(A) - \ell_j(A) V_j \tag{4.19}$$

for  $A \in \mathcal{A}$ , and set

$$\nabla A := (\partial_j A)_{j \in \mathcal{J}} \in \mathcal{B} , \quad \mathcal{B} = \prod_{j \in \mathcal{J}} \mathcal{B}_j .$$

We refer to  $\nabla A$  as the gradient of  $A$ , or derivative of  $A$ , with respect to the differential structure on  $\mathcal{A}$  defined above. We will denote the differential structure by the triple  $(\mathcal{A}, \nabla, \sigma)$ .

For  $s \in [0, 1]$  we endow  $\mathcal{B}_j$  with the inner product

$$\langle B_1, B_2 \rangle_{s,j} := \tau_j[B_1^* \ell_j(\sigma^s) B_2 r_j(\sigma^{1-s})] .$$

The most relevant case for our purposes is  $s = \frac{1}{2}$ , in which case we write

$$\langle B_1, B_2 \rangle_{L^2_{\text{KMS},j}(\sigma)} := \langle B_1, B_2 \rangle_{1/2,j} .$$

It follows immediately from Theorem 4.6 that

$$\mathcal{E}(A_1, A_2) := \sum_{j \in \mathcal{J}} \langle \partial_j A_1, \partial_j A_2 \rangle_{L^2_{\text{KMS},j}(\sigma)} \tag{4.20}$$

is a conservative completely Dirichlet form on  $L^2_{\text{KMS}}(\mathcal{A}_h, \sigma)$ .

**Remark 4.8** As we have seen earlier in this section, (3) ensures that the sesquilinear form  $\mathcal{E}$  defined by (4.20) is real and leads to the symmetry condition (4.13), and then (4) ensures that  $\mathcal{E}$  is completely Dirichlet.

Having the gradient  $\nabla$  at our disposal, we can define a corresponding divergence operator by trace duality. For  $\mathbf{B} = (B_j)_{j \in \mathcal{J}} \in \mathcal{B}$  we shall use the notation

$$\text{div } \mathbf{B} = - \sum_{j \in \mathcal{J}} \partial_j^\dagger B_j . \tag{4.21}$$

**Proposition 4.9** *Let  $s \in [0, 1]$ . The adjoint of the differential operator  $\partial_j : (\mathcal{A}, \langle \cdot, \cdot \rangle_s) \rightarrow (\mathcal{B}_j, \langle \cdot, \cdot \rangle_{s,j})$  is given by*

$$\partial_{j,\sigma}^{\dagger,(s)} B = e^{-s\omega_j} r_j^\dagger (V_j^* B) - e^{(1-s)\omega_j} \ell_j^\dagger (B V_j^*) . \tag{4.22}$$

*In particular, the adjoint of the operator  $\partial_j : L^2_{\text{KMS}}(\mathcal{A}, \sigma) \rightarrow L^2_{\text{KMS},j}(\mathcal{B}_j, \sigma)$  is given by*

$$\partial_{j,\sigma}^\dagger B = e^{-\omega_j/2} r_j^\dagger (V_j^* B) - e^{\omega_j/2} \ell_j^\dagger (B V_j^*) \tag{4.23}$$

for  $B \in \mathcal{B}_j$ .

**Proof** For  $A \in \mathcal{A}$  we obtain using (4.4) and (4.12),

$$\begin{aligned} & \langle \partial_j A, B \rangle_{s,j} \\ &= \tau_j [(V_j r_j(A) - \ell_j(A) V_j)^* \ell_j(\sigma^s) B r_j(\sigma^{1-s})] \\ &= \tau_j [r_j(A)^* V_j^* \ell_j(\sigma^s) B r_j(\sigma^{1-s}) - \ell_j(A)^* \ell_j(\sigma^s) B r_j(\sigma^{1-s}) V_j^*] \\ &= \tau [A^* r_j^\dagger (V_j^* \ell_j(\sigma^s) B r_j(\sigma^{1-s})) - A^* \ell_j^\dagger (\ell_j(\sigma^s) B r_j(\sigma^{1-s}) V_j^*)] \\ &= \tau [A^* \sigma^s (r_j^\dagger (r_j(\sigma^{-s}) V_j^* \ell_j(\sigma^s) B) - \ell_j^\dagger (B r_j(\sigma^{1-s}) V_j^* \ell_j(\sigma^{s-1}))) \sigma^{1-s}] \\ &= \tau [A^* \sigma^s (e^{-s\omega_j} r_j^\dagger (V_j^* B) - e^{(1-s)\omega_j} \ell_j^\dagger (B V_j^*)) \sigma^{1-s}] \\ &= \langle A, e^{-s\omega_j} r_j^\dagger (V_j^* B) - e^{(1-s)\omega_j} \ell_j^\dagger (B V_j^*) \rangle_s , \end{aligned}$$

which proves (4.22). □

The following result provides an explicit expression for  $\mathcal{L}$ .

**Proposition 4.10** *The operator  $\mathcal{L}$  associated to the Dirichlet form (4.20) is given by*

$$\begin{aligned} \mathcal{L}A &= \sum_{j \in \mathcal{J}} e^{-\omega_j/2} r_j^\dagger \left( -r_j(A) V_j^* V_j + 2V_j^* \ell_j(A) V_j - V_j^* V_j r_j(A) \right) \\ &= \sum_{j \in \mathcal{J}} e^{\omega_j/2} \ell_j^\dagger \left( V_j r_j(A) V_j^* - \ell_j(A) V_j V_j^* \right) - e^{-\omega_j/2} r_j^\dagger \left( V_j^* V_j r_j(A) - V_j^* \ell_j(A) V_j \right) \end{aligned}$$

for  $A \in \mathcal{A}$ . Its Hilbert space adjoint with respect to  $L^2(\mathcal{A}, \tau)$  is given by

$$\begin{aligned} \mathcal{L}^\dagger \rho &= \sum_{j \in \mathcal{J}} e^{-\omega_j/2} \left( -r_j^\dagger(r_j(\rho) V_j^* V_j) + 2\ell_j^\dagger(V_j r_j(\rho) V_j^*) - r_j^\dagger(V_j^* V_j r_j(\rho)) \right) \\ &= \sum_{j \in \mathcal{J}} e^{\omega_j/2} \left( r_j^\dagger(V_j^* \ell_j(\rho) V_j) - \ell_j^\dagger(\ell_j(\rho) V_j V_j^*) \right) \\ &\quad - e^{-\omega_j/2} \left( r_j^\dagger(V_j^* V_j r_j(\rho)) - \ell_j^\dagger(V_j r_j(\rho) V_j^*) \right) \end{aligned}$$

for  $\rho \in \mathcal{A}$ .

**Proof** Using Proposition 4.9 we obtain

$$\begin{aligned} \mathcal{L}A &= - \sum_{j \in \mathcal{J}} \partial_{j,\sigma}^\dagger \partial_j A \\ &= - \sum_{j \in \mathcal{J}} e^{-\omega_j/2} r_j^\dagger(V_j^* \partial_j A) - e^{\omega_j/2} \ell_j^\dagger((\partial_j A) V_j^*) \\ &= - \sum_{j \in \mathcal{J}} e^{-\omega_j/2} r_j^\dagger(V_j^* V_j r_j(A) - V_j^* \ell_j(A) V_j) - e^{\omega_j/2} \ell_j^\dagger(V_j r_j(A) V_j^* - \ell_j(A) V_j V_j^*), \end{aligned}$$

which yields the second expression for  $\mathcal{L}$ . The first expression is obtained using (4.17) and the fact that  $\omega_j^* = -\omega_j$ . The formulas for  $\mathcal{L}^\dagger$  follow by direct computation.  $\square$

The following result is an immediate consequence.

**Proposition 4.11** *We have*

$$\text{Ker}(\mathcal{L}) = \text{Ker}(\nabla) \quad \text{and} \quad \text{Ran}(\mathcal{L}^\dagger) = \text{Ran}(\text{div}).$$

**Proof** The identity  $\mathcal{L}A = - \sum_{j \in \mathcal{J}} \partial_{j,\sigma}^\dagger \partial_j A$  implies that  $\text{Ker}(\nabla) \subseteq \text{Ker}(\mathcal{L})$ . The reverse inclusion follows from the identity  $-\langle \mathcal{L}A, A \rangle_{L^2_{\text{KMS}}(\sigma)} = \sum_{j \in \mathcal{J}} \langle \partial_j A, \partial_j A \rangle_{L^2_{\text{KMS},j}(\sigma)}$ . The identification of the ranges is a consequence of duality.  $\square$

**Proposition 4.12** *For  $s \in [0, 1]$  and  $A_1, A_2 \in \mathcal{A}$  we have the identity*

$$-\langle \mathcal{L}A_1, A_2 \rangle_s = \sum_{j \in \mathcal{J}} e^{(s-\frac{1}{2})\omega_j} \langle \partial_j A_1, \partial_j A_2 \rangle_{s,j}.$$

Consequently, the operator  $\mathcal{L}$  is self-adjoint with respect to  $\langle \cdot, \cdot \rangle_s$  for all  $s \in [0, 1]$ , and in particular, the detailed balance condition holds in the sense of Definition 2.3.

**Proof** This follows from a direct computation using (4.22).  $\square$



### 5 Examples

We provide a number of examples of conservative completely Dirichlet forms defined in the context of a differential structure on a finite-dimensional von Neumann algebra  $\mathcal{A}$  equipped with a faithful state  $\sigma$ .

#### 5.1 Generators of Quantum Markov Semigroups in Lindblad Form

We have seen in Sect. 2 that generators of quantum Markov semigroups satisfying detailed balance (see Theorem 2.4) naturally fit into the framework of Sect. 4 by taking  $\mathcal{A} = \mathcal{B}_j = B(\mathcal{H})$  and  $\ell_j = r_j = I_{\mathcal{A}}$ .

The framework also includes quantum Markov semigroups on subalgebras  $\mathcal{A}$  of  $B(\mathcal{H})$ . In this case we set  $\mathcal{B}_j = B(\mathcal{H})$ , so that the situation in which  $V_j \notin \mathcal{A}$  is covered. Such a situation also arises naturally in the following example.

#### 5.2 Classical Reversible Markov Chains in the Lindblad Framework

For  $n \geq 2$ , Let  $\{e_1, \dots, e_n\}$  be an orthonormal basis of  $\mathbb{R}^n$  and set  $E_{kp} = |e_k\rangle\langle e_p|$ . Note that  $E_{kp}E_{rs} = \delta_{pr}E_{ks}$  and  $E_{kp}^* = E_{pk}$ . We consider the algebra  $\mathcal{A} \subseteq \mathbb{M}_n(\mathbb{C})$  consisting of all operators that are diagonal in the basis given by the  $e_i$ 's:

$$\mathcal{A} = \left\{ \sum_{i=1}^n \psi_i E_{ii} : \psi_1, \dots, \psi_n \in \mathbb{C} \right\}.$$

Furthermore, for each  $k, p$ , we set  $\mathcal{B}_{kp} = \mathbb{M}_n(\mathbb{C})$ , and we endow  $\mathcal{A}$  and  $\mathcal{B}_{kp}$  with the usual normalized trace given by  $\tau(B) = \frac{1}{n} \sum_i \langle B e_i, e_i \rangle$ . Let  $\ell_{kp} = r_{kp}$  be the canonical embedding from  $\mathcal{A}$  into  $\mathcal{B}_{kp}$ . It then follows that  $\ell_{kp}^\dagger(B) = r_{kp}^\dagger(B) = \sum_i \langle B e_i, e_i \rangle E_{ii}$ .

For  $k \neq p$ , let  $q_{kp} \geq 0$  be the transition rate of a continuous-time Markov chain on  $\{1, \dots, n\}$ . We set  $V_{kp} = 2^{-1/2}(q_{kp}q_{pk})^{1/4} E_{kp}$  so that  $V_{kp}^* = V_{pk}$ . Moreover, it is immediate to see that the identity in (4.17) holds. Fix positive weights  $\pi_1, \dots, \pi_n$ . It then follows that  $\sigma = \sum_i \pi_i E_{ii}$  satisfies (4.18) with  $\omega_{kp} = \log(\pi_p/\pi_k)$ .

By Proposition 4.10, the operator  $\mathcal{L}$  associated to the Dirichlet form (4.20) is given by

$$\mathcal{L}A = \frac{1}{2} \sum_{k \neq p} \sqrt{\frac{q_{kp}q_{pk}\pi_k}{\pi_p}} \left( E_{kp}^*[A, E_{kp}] + [E_{kp}^*, A]E_{kp} \right)$$

for  $A \in \mathcal{A}$ . Assume now that  $\pi_1, \dots, \pi_n$  satisfy the classical detailed balance condition, i.e.,  $\pi_k q_{kp} = \pi_p q_{pk}$  for all  $k, p$ . Then we have

$$\mathcal{L}A = \frac{1}{2} \sum_{k \neq p} q_{pk} \left( E_{kp}^*[A, E_{kp}] + [E_{kp}^*, A]E_{kp} \right).$$

More explicitly,

$$\mathcal{L} \left( \sum_i \psi_i E_{ii} \right) = \sum_{k,p} q_{kp} (\psi_p - \psi_k) E_{kk}.$$

Hence, under the identification  $(\psi_1, \dots, \psi_n) \leftrightarrow \sum_{i=1}^n \psi_i E_{ii}$ , the operator  $\mathcal{L}$  corresponds to the operator  $\mathcal{L}_M$  given by  $(\mathcal{L}_M \psi)_k = \sum_p q_{kp} (\psi_p - \psi_k)$ , which is the generator of the continuous-time Markov chain on  $\{1, \dots, n\}$  with transition rates from  $k$  to  $p$  given by  $q_{kp}$ .

### 5.3 Another Approach to Reversible Markov Chains

Let us now give an alternative way to put reversible Markov chains in the framework of this paper, which corresponds to the construction in [32,33]. As above, let  $q_{kp} \geq 0$  be the transition rate of a continuous-time Markov chain on  $\{1, \dots, n\}$ , and assume that the positive weights  $\pi_1, \dots, \pi_n$  satisfy the detailed balance condition  $\pi_k q_{kp} = \pi_p q_{pk}$ . Let  $\mathcal{J} := \{(k, p) : q_{kp} > 0\}$  be the edge set of the associated graph. We consider the (non-)commutative probability spaces  $(\mathcal{A}, \tau)$  and  $(\mathcal{B}_{kp}, \tau_{kp})$  given by

$$\mathcal{A} := \ell_n^\infty, \quad \tau(A) := \sum_{i=1}^n A_i \pi_i, \quad \mathcal{B}_{kp} = \mathbb{C}, \quad \tau_{kp}(B) := \frac{B}{2} \pi_k q_{kp}.$$

The operators  $\partial_{kp}$  are determined by  $V_{kp} = 1$ ,  $\ell_{kp}(A) = A_k$ , and  $r_{kp}(A) = A_p$  for  $A \in \ell_n^\infty$ . It follows that  $\ell_{kp}^\dagger(B) = \frac{B}{2} q_{kp} e_k$  and  $r_{kp}^\dagger(B) = \frac{B}{2} q_{pk} e_p$ , where  $e_k$  is the  $k$ 'th unit vector in  $\ell_n^\infty$ . Therefore,

$$\partial_{kp} A = A_p - A_k \quad \text{and} \quad \partial_{kp}^\dagger B = \frac{B}{2} (q_{pk} e_p - q_{kp} e_k).$$

Moreover, as  $\sigma = \mathbf{1}$  satisfies (4.18) with  $\omega_{kp} = 0$ , it is readily checked that this defines a differentiable structure in the sense of Definition 4.7. Using Proposition 4.10, we infer that the operator  $\mathcal{L}$  is given by

$$(\mathcal{L}A)_k = \sum_p q_{kp} (A_p - A_k),$$

so that  $\mathcal{L}$  is indeed the generator of the continuous time Markov chain with transition rates  $q_{kp}$ .

### 5.4 The Discrete Hypercube

For a given Markov chain generator, there are different ways to write the generator in the framework of this paper, and it is often useful to represent  $\mathcal{L}$  using set  $\mathcal{J}$  that is smaller than in Example 5.3; see also [21]. We illustrate this for the simple random walk on the discrete hypercube  $\mathcal{Q}^n = \{-1, 1\}^n$ . Set  $\mathcal{J} = \{1, \dots, n\}$ , and let  $s_j : \mathcal{Q}^n \rightarrow \mathcal{Q}^n$  define the  $j$ -th coordinate swap defined by  $s_j(x_1, \dots, x_n) = (x_1, \dots, -x_j, \dots, x_n)$ .

Consider the (non-)commutative probability spaces  $(\mathcal{A}, \tau)$  and  $(\mathcal{B}_j, \tau_j)$  determined by

$$\mathcal{A} := \ell^\infty(\mathcal{Q}^n), \quad \tau(A) := 2^{-n} \sum_{x \in \mathcal{Q}^n} A(x), \quad \mathcal{B}_j = \mathcal{A}, \quad \tau_j := \tau.$$

Furthermore, set  $\sigma = \mathbf{1}$  and  $\omega_j = 0$ . We define  $V_j = 1$ ,  $\ell_j = I$ , and  $r_j A(x) = A(s_j x)$ , so that  $\partial_j A(x) = A(s_j x) - A(x)$ . This defines a differential structure with  $\sigma = \mathbf{1}$ . It follows that  $r_j^\dagger = r_j$  and

$$\partial_j A(x) = \partial_j^\dagger A(x) = A(s_j x) - A(x).$$

It follows that

$$\mathcal{L}A(x) = 2 \sum_{j=1}^n (A(s_j x) - A(x)),$$

which is the discrete Laplacian on  $\mathcal{Q}^n$  that generates the simple random walk.

### 5.5 The Fermionic Ornstein–Uhlenbeck Equation

A non-commutative example in which it is advantageous to work with  $\ell_j$  not equal to the identity, is the Fermionic Ornstein–Uhlenbeck operator, for which a non-commutative transport metric was constructed in [8]. Let  $(Q_1, \dots, Q_n)$  be self-adjoint operators on a finite-dimensional Hilbert space satisfying the *canonical anti-commutation relations* (CAR):

$$Q_i Q_j + Q_j Q_i = 2\delta_{ij} .$$

The *Clifford algebra*  $\mathfrak{C}^n$  is the  $2^n$ -dimensional algebra generated by  $\{Q_j\}_{j=1}^n$ . Let  $\Gamma : \mathfrak{C}^n \rightarrow \mathfrak{C}^n$  be the principle automorphism on  $\mathfrak{C}^n$ , i.e., the unique algebra homomorphism satisfying  $\Gamma(Q_j) = -Q_j$  for all  $j$ . Let  $\tau$  be the canonical trace on  $\mathfrak{C}^n$ , determined by  $\tau(Q_1^{\alpha_1} \dots Q_n^{\alpha_n}) := \delta_{0,|\alpha|}$  for all  $\alpha = (\alpha_j)_j \in \{0, 1\}^n$ , where  $|\alpha| := \sum_j \alpha_j$ . We then set  $\mathcal{J} = \{1, \dots, n\}$ ,  $\mathcal{A} := \mathcal{B}_j := \mathfrak{C}^n$ , and  $\tau_j := \tau$ . Furthermore we set  $V_j = Q_j$ ,  $\ell_j = \Gamma$ , and  $r_j = I$ . Then  $\ell_j^\dagger = \Gamma$ , and the operators  $\partial_j$  and  $\partial_j^\dagger$  are skew-derivations given by

$$\partial_j A = Q_j A - \Gamma(A) Q_j , \quad \partial_j^\dagger A = Q_j A + \Gamma(A) Q_j .$$

Taking  $\sigma = \mathbf{1}$  and  $\omega_j = 0$  we obtain

$$\mathcal{L}A = 2 \sum_{j=1}^n (Q_j A Q_j - A) ,$$

which implies that  $\mathcal{L} = -4\mathcal{N}$ , where  $\mathcal{N}$  is the *fermionic number operator* (see [8,9] for more details).

### 5.6 The Depolarizing Channel

This is one of the simplest non-commutative examples. Given a non-commutative probability space  $(\mathcal{A}, \tau)$  and  $\gamma > 0$ , the generator is defined by

$$\mathcal{L}A = \gamma(\tau[A]\mathbf{1} - A) . \tag{5.1}$$

In the case where  $\mathcal{A} = \mathcal{B}_j = \mathbb{M}_2(\mathbb{C})$  and  $\tau$  is the usual trace, this operator can be written in Lindblad form using the Pauli matrices

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} , \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} , \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} .$$

We set  $V_j = \sqrt{\gamma}\sigma_j$  and  $\ell_j = r_j = I_{\mathcal{A}}$ , so that the differential operators  $\partial_x, \partial_y$  and  $\partial_z$  are the commutators

$$\partial_j A = \sqrt{\gamma}[\sigma_j, A]$$

for  $j \in \{x, y, z\}$ . This yields a differentiable structure with  $\sigma = \mathbf{1}$  and  $\omega_j = 0$ , and a direct computation shows that  $\mathcal{L}$  is indeed given by (5.1).

### 6 Non-commutative Functional Calculus

Let  $\mathcal{A}$  be a finite-dimensional  $C^*$ -algebra. Let  $A, B \in \mathcal{A}$  be self-adjoint with spectral decompositions

$$A = \sum_i \lambda_i A_i \quad \text{and} \quad B = \sum_k \mu_k B_k \tag{6.1}$$

for some eigenvalues  $\lambda_i, \mu_k \in \mathbb{R}$  and spectral projections  $A_i, B_k \in \mathcal{A}$  satisfying  $A_i A_k = \delta_{ik} A_i, B_i B_k = \delta_{ik} B_i$ , and  $\sum_i A_i = \sum_k B_k = \mathbf{1}_{\mathcal{A}}$ . For a function  $\theta : \text{sp}(A) \times \text{sp}(B) \rightarrow \mathbb{R}$  we define  $\theta(A, B) \in \mathcal{A} \times \mathcal{A}$  to be the *double operator sum*

$$\theta(A, B) = \sum_{i,k} \theta(\lambda_i, \mu_k) A_i \otimes B_k . \tag{6.2}$$

**Remark 6.1** A systematic theory of infinite-dimensional generalizations of  $\theta(A, B)$  has been developed under the name of *double operator integrals*, see, e.g., [5,43].

Double operator sums are compatible with the usual functional calculus, in the sense that

$$\theta(f(A), g(B)) = (\theta \circ (f, g))(A, B) \tag{6.3}$$

for all  $f : \text{sp}(A) \rightarrow \mathbb{R}, g : \text{sp}(B) \rightarrow \mathbb{R}$  and  $\theta : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ . Moreover, recalling that the contraction operator has been defined in (2.16), we have

$$\theta_2(A, B) \# (\theta_1(A, B) \# C) = (\theta_2 \cdot \theta_1)(A, B) \# C \tag{6.4}$$

The straightforward proof of these identities is left to the reader.

Let  $\mathcal{I} \subseteq \mathbb{R}$  be an interval. Of particular relevance for our purposes is the special case where  $\theta = \delta f : \mathcal{I} \times \mathcal{I} \rightarrow \mathbb{R}$  is the *discrete derivative* of a differentiable function  $f : \mathcal{I} \rightarrow \mathbb{R}$ , defined by

$$\delta f(\lambda, \mu) := \begin{cases} \frac{f(\lambda) - f(\mu)}{\lambda - \mu}, & \lambda \neq \mu, \\ f'(\lambda), & \lambda = \mu. \end{cases} \tag{6.5}$$

Using the contraction operator we can write the following useful chain rule:

$$f(A) - f(B) = \delta f(A, B) \# (A - B) . \tag{6.6}$$

We can also formulate a chain rule for the operator  $\partial_V$  defined in (4.2), which plays a crucial role in the sequel.

**Proposition 6.2** (Chain rule for  $\partial_V$ ) *Let  $A \in \mathcal{A}_h$ . For any function  $f : \text{sp}(A) \rightarrow \mathbb{R}$  we have*

$$\partial_V f(A) = \delta f(\ell(A), r(A)) \# \partial_V A . \tag{6.7}$$

**Proof** Let  $A = \sum_i \lambda_i A_i$  be the spectral decomposition with eigenvalues  $\lambda_i \in \mathbb{R}$  and spectral projections  $A_i \in \mathcal{A}$  satisfying  $A_i A_k = \delta_{ik} A_i$  and  $\sum_i A_i = \mathbf{1}_{\mathcal{A}}$ . Since  $\ell(\mathbf{1}_{\mathcal{A}}) = r(\mathbf{1}_{\mathcal{A}}) = \mathbf{1}_{\mathcal{A}}$  by assumption, it follows that  $\sum_i \ell(A_i) = \sum_i r(A_i) = \mathbf{1}_{\mathcal{B}}$  for all  $j$ . Therefore,

$$\begin{aligned} \partial_V A &= \sum_i \lambda_i \left( Vr(A_i) - \ell(A_i) V \right) \\ &= \sum_{i,k} (\lambda_k - \lambda_i) \ell(A_i) Vr(A_k) . \end{aligned}$$

Consequently, since  $\ell(A_p)\ell(A_i) = \ell(A_p A_i) = \delta_{pi}\ell(A_i)$  and  $r(A_k)r(A_m) = \delta_{km}r(A_k)$ ,

$$\begin{aligned} \delta f(\ell(A), r(A))\# \partial A &= \sum_{i,k,p,m} \delta f(\lambda_p, \lambda_m)(\lambda_k - \lambda_i)\ell(A_p A_i) V r(A_k A_m) \\ &= \sum_{i,k} \delta f(\lambda_i, \lambda_k)(\lambda_k - \lambda_i)\ell(A_i) V r(A_k) \\ &= \sum_{i,k} (f(\lambda_k) - f(\lambda_i))\ell(A_i) V r(A_k) \\ &= \partial_V f(A) . \end{aligned} \tag{6.8}$$

□

**Remark 6.3** Note that the function  $f$  is not required to be differentiable in Proposition 6.2. In this case,  $\delta f$  is not defined on the diagonal, but the second line in (6.8) shows that its diagonal value is irrelevant.

The following well-known chain rule can also be formulated in terms of  $\delta f$ .

**Proposition 6.4** Let  $A : \mathcal{I} \rightarrow \mathcal{A}_h$  be differentiable on an interval  $\mathcal{I} \subseteq \mathbb{R}$  and let  $f$  be a real-valued function on an interval containing  $\text{sp}(A(t))$  for all  $t \in \mathcal{I}$ . Then:

$$\frac{d}{dt} f(A(t)) = \delta f(A(t), A(t))\# A'(t) , \tag{6.9}$$

$$\frac{d}{dt} \tau[f(A(t))] = \tau[f'(A(t))A'(t)] . \tag{6.10}$$

**Proof** The first assertion follows by passing to the limit in (6.6). The second identity follows easily using the definition of  $\delta f$  and the cyclicity of the trace. □

**Example 6.5** We illustrate the proposition above with a well-known computation that will be useful below. For  $\rho, \sigma \in \mathfrak{P}_+(A)$  and  $v \in \mathcal{A}_h$  with  $\tau[v] = 0$ , set  $\rho_t := \rho + tv$ . It follows from (6.10) that

$$\partial_t \text{Ent}_\sigma(\rho_t) = \tau[v(\log \rho_t - \log \sigma)] . \tag{6.11}$$

Since  $\delta \log(r, s) = \frac{\log r - \log s}{r - s} = \int_0^\infty (x + r)^{-1} (x + s)^{-1} dx$ , we have  $\delta \log(R, S) = \int_0^\infty (x + R)^{-1} \otimes (x + S)^{-1} dx$ . Thus, (6.9) yields

$$\partial_t^2 \text{Ent}_\sigma(\rho_t) = \int_0^\infty \tau \left[ v \frac{1}{x + \rho_t} v \frac{1}{x + \rho_t} \right] dx . \tag{6.12}$$

We finish this subsection with some useful properties of the sesquilinear form  $(A, B) \mapsto \langle A, \varphi(R, S)\#B \rangle_{L^2(\tau)}$  on  $\mathcal{A}$ .

**Lemma 6.6** Let  $R, S \in \mathcal{A}$  be self-adjoint and let  $\varphi : \text{sp}(R) \times \text{sp}(S) \rightarrow \mathbb{R}_+$  be given. Then, for all  $A \in \mathcal{A}$ ,

$$\langle A, \varphi(R, S)\#A \rangle_{L^2(\tau)} \geq 0 .$$

**Proof** Using the spectral decompositions  $R = \sum_i \lambda_i R_i$  and  $S = \sum_k \mu_k S_k$  we may write

$$\langle A, \varphi(R, S)\#A \rangle_{L^2(\tau)} = \sum_{i,k} \varphi(\lambda_i, \mu_k) \tau[A^* R_i A S_k] .$$

Since  $\tau[A^* R_i A S_k] = \tau[(R_i A S_k)^*(R_i A S_k)] \geq 0$  the result follows. □

**Proposition 6.7** *Let  $R, S \in \mathcal{A}$  be self-adjoint and suppose that  $\varphi : \text{sp}(R) \times \text{sp}(S) \rightarrow \mathbb{R}$  is strictly positive. Then the sesquilinear form*

$$(A, B) \mapsto \langle A, \varphi(R, S)\#B \rangle_{L^2(\tau)}$$

defines a scalar product on  $\mathcal{A}$ .

**Proof** Consider the spectral decompositions  $R = \sum_i \lambda_i R_i$  and  $S = \sum_k \mu_k S_k$ . Using basic properties of the trace, we obtain

$$\overline{\tau[A^* R_i B S_k]} = \tau[(A^* R_i B S_k)^*] = \tau[S_k B^* R_i A] = \tau[B^* R_i A S_k],$$

and therefore, since  $\varphi$  is real-valued,

$$\begin{aligned} \overline{\langle A, \varphi(R, S)\#B \rangle_{L^2(\tau)}} &= \sum_{i,k} \varphi(\lambda_i, \mu_k) \overline{\tau[A^* R_i B S_k]} \\ &= \sum_{i,k} \varphi(\lambda_i, \mu_k) \tau[B^* R_i A S_k] = \langle B, \varphi(R, S)\#A \rangle_{L^2(\tau)}. \end{aligned}$$

Moreover, since  $\varphi$  is strictly positive on the finite set  $\text{sp}(R) \times \text{sp}(S)$ , we have  $\varphi \geq \varepsilon$  for some  $\varepsilon > 0$ . Thus Lemma 6.6 implies that  $\langle A, \varphi(R, S)\#A \rangle \geq \varepsilon \|A\|_{L^2(\tau)}^2$ . It follows that  $\langle A, \varphi(R, S)\#A \rangle \geq 0$ , with equality if and only if  $A = 0$ . □

### Higher Order Expressions

In the sequel we will use versions of Propositions 6.2 and 6.4 for higher order derivatives, for which we need to introduce more notation. For  $x = (x_1, \dots, x_n) \in \mathbb{R}^n$  and  $1 \leq i \leq m \leq n$  we will use the shorthand notation  $x_i^m = (x_i, x_{i+1}, \dots, x_{m-1}, x_m)$ . For a function  $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}$  and  $j = 1, \dots, n$  we consider the discrete derivative  $\delta_j \varphi : \mathbb{R}^{n+1} \rightarrow \mathbb{R}$  defined by

$$\delta_j \varphi(x_1^{j-1}, (x_j, \tilde{x}_j), x_{j+1}^n) := \delta \varphi(x_1^{j-1}, \cdot, x_{j+1}^n)(x_j, \tilde{x}_j), \tag{6.13}$$

where  $\delta$  denotes the discrete derivative given by (6.5). Iterating this procedure, one arrives at expressions that can be naturally encoded using rooted planar binary trees. Indeed, for a given function  $\theta : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  and  $x, y \in \mathbb{R}$ , we write

$$\theta(x, y) = \theta(x \searrow \bullet \swarrow y).$$

The left and right child in this tree correspond to the variables  $x$  and  $y$  in  $\theta(x, y)$  respectively. More complicated trees are then constructed by iteratively replacing one of the children  $\bullet$  by  $\vee$ . This will correspond to discrete differentiation with respect to the respective variables, e.g.,

$$\delta \theta((x, y), z) = \delta \theta \left( x \searrow \bullet \swarrow z \right) := \frac{\theta(x, z) - \theta(y, z)}{x - y}, \tag{6.14}$$

$$\delta \theta(x, (y, z)) = \delta \theta \left( x \searrow \bullet \swarrow z \right) := \frac{\theta(x, y) - \theta(x, z)}{y - z}, \tag{6.15}$$

$$\delta \theta(((x, y), z), w) = \delta \theta \left( x \searrow \bullet \swarrow w \right) := \frac{\frac{\theta(x, w) - \theta(z, w)}{x - z} - \frac{\theta(y, w) - \theta(z, w)}{y - z}}{x - y}. \tag{6.16}$$

The middle expressions are valid whenever the variables are distinct. If some of the variables are equal, finite differences are to be interpreted as derivatives. For instance, if  $x = y \neq z$  in (6.16), we have

$$\delta\theta\left(\begin{array}{c} x \quad \cdot \quad x \\ \diagdown \quad \diagup \\ \cdot \\ \diagdown \quad \diagup \\ z \quad \cdot \quad w \end{array}\right) = \partial_x\left(\delta\theta\left(\begin{array}{c} x \quad \cdot \quad z \\ \diagdown \quad \diagup \\ \cdot \\ \diagdown \quad \diagup \\ z \quad \cdot \quad w \end{array}\right)\right) = \frac{(x-z)\partial_1\theta(x,w) - (\theta(x,w) - \theta(z,w))}{(x-z)^2}.$$

If  $x = y = z$  in (6.16), then the formula above becomes

$$\delta\theta\left(\begin{array}{c} x \quad \cdot \quad x \\ \diagdown \quad \diagup \\ \cdot \\ \diagdown \quad \diagup \\ x \quad \cdot \quad w \end{array}\right) = \frac{1}{2}\partial_1^2\theta(x,w).$$

The functional calculus (6.2) generalizes naturally to functions of several variables. Let  $A^{(1)}, \dots, A^{(n)}$  be self-adjoint elements in  $\mathcal{A}$  with spectral decompositions

$$A^{(k)} = \sum_i \lambda_i^{(k)} A_i^{(k)}$$

for some eigenvalues  $\lambda_i^{(k)} \in \mathbb{R}$  and spectral projections  $A_i^{(k)} \in \mathcal{A}$  with  $\sum_i A_i^{(k)} = \mathbf{1}_{\mathcal{A}}$ . For a function  $\theta : \text{sp}(A^{(1)}) \times \dots \times \text{sp}(A^{(n)}) \rightarrow \mathbb{R}$  we define  $\theta(A_1, \dots, A_n) \in \mathcal{A}^{\otimes n}$  to be the multiple operator sum

$$\theta(A_1, \dots, A_n) = \sum_{i_1, \dots, i_n} \theta(\lambda_{i_1}^{(1)}, \dots, \lambda_{i_n}^{(n)}) A_{i_1}^{(1)} \otimes \dots \otimes A_{i_n}^{(n)}. \tag{6.17}$$

In the sequel we shall apply this definition to  $\delta\theta$  in order to define expressions such as  $\delta\theta(A, B)$ . The tree notation is useful when considering generalizations of the contraction operation (2.16) to higher order tensor products. Each of the nodes that is a parent can be used to indicate the position at which an operator for contraction is inserted: e.g., we write

$$\begin{aligned} \theta\left(\begin{array}{c} A^{(1)} \quad \cdot \quad A^{(2)} \\ \diagdown \quad \diagup \\ B \end{array}\right) &= \theta(A^{(1)}, A^{(2)})\#B = \sum_{i,j} \theta(\lambda_i^{(1)}, \lambda_j^{(2)}) A_i^{(1)} B A_j^{(2)}, \\ \delta\theta\left(\begin{array}{c} A^{(1)} \quad \cdot \quad A^{(2)} \\ \diagdown \quad \diagup \\ \cdot \\ \diagdown \quad \diagup \\ A^{(3)} \end{array}\right) &= \sum_{i,j,k} \frac{\theta(\lambda_i^{(1)}, \lambda_k^{(3)}) - \theta(\lambda_j^{(2)}, \lambda_k^{(3)})}{\lambda_i^{(1)} - \lambda_j^{(2)}} A_i^{(1)} B A_j^{(2)} \otimes A_k^{(3)}, \\ \delta\theta\left(\begin{array}{c} A^{(2)} \quad \cdot \quad A^{(3)} \\ \diagdown \quad \diagup \\ \cdot \\ \diagdown \quad \diagup \\ A^{(1)} \end{array}\right) &= \sum_{i,j,k} \frac{\theta(\lambda_i^{(1)}, \lambda_j^{(2)}) - \theta(\lambda_i^{(1)}, \lambda_k^{(3)})}{\lambda_j^{(2)} - \lambda_k^{(3)}} A_i^{(1)} \otimes A_j^{(2)} B A_k^{(3)}, \end{aligned}$$

where the fractions at the right-hand side are to be understood in the sense of limits if the denominator vanishes. These expressions appear naturally in the following chain rule that will be useful in Sect. 7.

**Proposition 6.8** *Let  $A, B : \mathcal{I} \rightarrow \mathcal{A}_h$  be differentiable on an interval  $\mathcal{I} \subseteq \mathbb{R}$ , and let  $\theta : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  be differentiable. Then:*

$$\partial_t\theta(A_t, B_t) = \delta\theta\left(\begin{array}{c} A_t \quad \cdot \quad A_t \\ \diagdown \quad \diagup \\ \partial_t A_t \quad \cdot \quad B_t \end{array}\right) + \delta\theta\left(\begin{array}{c} B_t \quad \cdot \quad B_t \\ \diagdown \quad \diagup \\ A_t \quad \cdot \quad \partial_t B_t \end{array}\right).$$

**Proof** We have  $\partial_t\theta(A_t, B_t) = \partial_s|_{s=t}\theta(A_s, B_t) + \partial_s|_{s=t}\theta(A_t, B_s)$ . Since we can write  $\theta(A_t, B_s) = \sum_k \theta(A_t, \mu_{s,k}) \otimes F_{s,k}$ , where  $B_s = \sum_k \mu_{s,k} F_{s,k}$  denotes the spectral decomposition of  $B_s$ , the result follows by applying (6.9) from Proposition 6.4 twice.  $\square$

Higher order derivatives can also be naturally expressed in terms of trees, but since this will not be needed in the sequel, we will not go into details here.

## 7 Riemannian Structures on the Space of Density Matrices

In this section we shall analyze a large class of Riemannian metrics on the space of density matrices. Throughout the section we fix a differentiable structure  $(\mathcal{A}, \nabla, \sigma)$  in the sense of Definition 4.7. The generator of the associated quantum Markov semigroup  $(\mathcal{P}_t)_t$  will be denoted by  $\mathcal{L}$ .

### 7.1 Riemannian Structures on Density Matrices

Consider the  $\mathbb{R}$ -linear subspace

$$\mathcal{A}_0 := \text{Ran}(\mathcal{L}^\dagger) \cap \mathcal{A}_h .$$

We shall study Riemannian structures on relatively open subsets of  $\mathfrak{P}_+$ , the set of all strictly positive elements in  $\mathfrak{P}$ . These subsets are of the form

$$\mathcal{M}_\rho := (\rho + \mathcal{A}_0) \cap \mathfrak{P}_+ ,$$

where  $\rho \in \mathfrak{P}_+$ . At each point of  $\mathcal{M}_\rho$ , the tangent space of  $\mathcal{M}_\rho$  is thus naturally given by  $\mathcal{A}_0$ .

**Remark 7.1** Of special interest is the *ergodic* case, i.e., the case where  $\text{Ker}(\mathcal{L}) = \text{lin}\{\mathbf{1}\}$ . In this case we have  $\mathcal{A}_0 = \{A \in \mathcal{A}_h : \tau[A] = 0\}$ , and therefore  $\mathcal{M}_\rho = \mathfrak{P}_+$  for all  $\rho \in \mathfrak{P}_+$ .

In order to define a Riemannian structure, we shall fix for each  $j \in \mathcal{J}$  a function  $\theta_j : [0, \infty) \times [0, \infty) \rightarrow \mathbb{R}$  satisfying the following properties:

**Assumption 7.2** For  $j \in \mathcal{J}$  the functions  $\theta_j : [0, \infty) \times [0, \infty) \rightarrow \mathbb{R}$  are continuous. Moreover, on  $(0, \infty) \times (0, \infty)$ , the function  $\theta_j$  is  $C^\infty$  and strictly positive, and we have the symmetry condition

$$\theta_j(r, s) = \theta_{j^*}(s, r) . \tag{7.1}$$

Recalling the definition of the double operator sum in (6.2), we will use the shorthand notation

$$\widehat{\rho}_j = \theta_j(\ell_j(\rho), r_j(\rho)) \in \mathcal{B}_j \otimes \mathcal{B}_j , \quad \widehat{\rho} = (\widehat{\rho}_j)_{j \in \mathcal{J}} \quad \text{for } \rho \in \mathfrak{P} , \tag{7.2}$$

$$\check{\rho}_j = \frac{1}{\theta_j}(\ell_j(\rho), r_j(\rho)) \in \mathcal{B}_j \otimes \mathcal{B}_j , \quad \check{\rho} = (\check{\rho}_j)_{j \in \mathcal{J}} \quad \text{for } \rho \in \mathfrak{P}_+ . \tag{7.3}$$

Let us now define the class of quantum transport metrics that we are interested in. For  $\rho \in \mathfrak{P}$ , we define the operator  $\mathcal{K}_\rho : \mathcal{A} \rightarrow \mathcal{A}$  by

$$\mathcal{K}_\rho A := -\text{div}(\widehat{\rho} \# \nabla A) = \sum_{j \in \mathcal{J}} \partial_j^\dagger (\widehat{\rho}_j \# \partial_j A) , \tag{7.4}$$

where we use the vector notation  $\widehat{\rho} \# \nabla A = (\widehat{\rho}_j \# \partial_j A)_{j \in \mathcal{J}}$  and we recall that the divergence operator has been defined in (4.21). To define the Riemannian metric we need a lemma concerning the unique solvability of the continuity equation in the class of “gradient vector fields”. Therefore we need to identify the kernel and the range of the linear operator  $\mathcal{K}_\rho$ .



**Lemma 7.3** (Mapping properties of  $\mathcal{K}_\rho$ ) For  $\rho \in \mathfrak{P}_+$  the operator  $\mathcal{K}_\rho$  is non-negative and self-adjoint on  $L^2(\mathcal{A}, \tau)$ . Moreover, we have

$$\text{Ran}(\mathcal{K}_\rho) = \text{Ran}(\mathcal{L}^\dagger) = \text{Ran}(\text{div}) , \quad \text{Ker}(\mathcal{K}_\rho) = \text{Ker}(\mathcal{L}) = \text{Ker}(\nabla) . \quad (7.5)$$

Furthermore,  $\mathcal{K}_\rho$  is real, i.e., for  $A \in \mathcal{A}$  we have  $(\mathcal{K}_\rho A)^* = \mathcal{K}_\rho A^*$ .

**Proof** For  $A, B \in \mathcal{A}$ , Lemma 6.7 yields

$$\langle \mathcal{K}_\rho A, B \rangle_{L^2(\tau)} = \sum_{j \in \mathcal{J}} \langle \widehat{\rho}_j \# \partial_j A, \partial_j B \rangle_{L^2(\tau_j)} = \sum_{j \in \mathcal{J}} \langle \partial_j A, \widehat{\rho}_j \# \partial_j B \rangle_{L^2(\tau_j)} = \langle A, \mathcal{K}_\rho B \rangle_{L^2(\tau)} ,$$

hence  $\mathcal{K}_\rho$  is self-adjoint on  $L^2(\mathcal{A}, \tau)$ .

The identities  $\text{Ker}(\mathcal{L}) = \text{Ker}(\nabla)$  and  $\text{Ran}(\mathcal{L}^\dagger) = \text{Ran}(\text{div})$  have already been proved in Proposition 4.11. Clearly,  $\text{Ker}(\nabla) \subseteq \text{Ker}(\mathcal{K}_\rho)$ . To prove the opposite inclusion, we note that since  $\rho \in \mathfrak{P}_+$ , there exists  $c > 0$  with  $\theta_j|_{\text{sp}(\rho)} \geq c > 0$  for all  $j \in \mathcal{J}$ . Lemma 6.6 implies that

$$\langle A, \mathcal{K}_\rho A \rangle_{L^2(\mathcal{A}, \tau)} = \sum_{j \in \mathcal{J}} \langle \partial_j A, \widehat{\rho}_j \# \partial_j A \rangle_{L^2(\mathcal{B}_j, \tau_j)} \geq c \sum_{j \in \mathcal{J}} \|\partial_j A\|_{L^2(\mathcal{B}_j, \tau_j)}^2 ,$$

from which we infer that  $\text{Ker}(\mathcal{K}_\rho) \subseteq \text{Ker}(\nabla)$ . This proves the second identity in (7.5), and the nonnegativity of  $\mathcal{K}_\rho$  follows as well. The first identity in (7.5) follows using elementary linear algebra, since the self-adjointness of  $\mathcal{K}_\rho$  in  $L^2(\mathcal{A}, \tau)$  yields

$$\text{Ran}(\mathcal{K}_\rho) = (\text{Ker}(\mathcal{K}_\rho))^\perp = (\text{Ker}(\nabla))^\perp = \text{Ran}(\text{div}) .$$

To prove that  $\mathcal{K}_\rho$  preserves self-adjointness, we consider the spectral decomposition  $\rho = \sum_k \lambda_k E_k$ , and write  $\theta_j^{km} := \theta_j(\lambda_k, \lambda_m)$  for brevity. We have

$$\begin{aligned} \langle B, \mathcal{K}_\rho A^* \rangle_{L^2(\tau)} &= \sum_{j \in \mathcal{J}} \tau_j [(\partial_j B)^* \widehat{\rho}_j \# \partial_j A^*] \\ &= \sum_{j \in \mathcal{J}} \sum_{k, m} \theta_j^{km} \tau_j \left[ \left( V_j r_j(B) - \ell_j(B) V_j \right)^* \ell_j(E_k) \left( V_j r_j(A^*) - \ell_j(A^*) V_j \right) r_j(E_m) \right] \\ &= \sum_{j \in \mathcal{J}} \sum_{k, m} \theta_j^{km} \tau_j \left[ r_j(B^*) V_j^* \ell_j(E_k) V_j r_j(A^*) r_j(E_m) - V_j^* \ell_j(B^*) \ell_j(E_k) V_j r_j(A^*) r_j(E_m) \right. \\ &\quad \left. - r_j(B^*) V_j^* \ell_j(E_k) \ell_j(A^*) V_j r_j(E_m) + V_j^* \ell_j(B^*) \ell_j(E_k) \ell_j(A^*) V_j r_j(E_m) \right] \\ &= \sum_{j \in \mathcal{J}} \sum_{k, m} \theta_j^{km} \tau_j \left[ V_j^* \ell_j(E_k) V_j r_j(A^* E_m B^*) - V_j^* \ell_j(B^* E_k) V_j r_j(A^* E_m) \right. \\ &\quad \left. - V_j^* \ell_j(E_k A^*) V_j r_j(E_m B^*) + V_j^* \ell_j(B^* E_k A^*) V_j r_j(E_m) \right] . \end{aligned} \quad (7.6)$$

On the other hand,

$$\begin{aligned} \langle B, (\mathcal{K}_\rho A)^* \rangle_{L^2(\tau)} &= \sum_{j \in \mathcal{J}} \tau_j [(\partial_j B^*) (\widehat{\rho}_j \# \partial_j A)^*] \\ &= \sum_{j \in \mathcal{J}} \sum_{k, m} \theta_j^{km} \tau_j \left[ \left( V_j r_j(B^*) - \ell_j(B^*) V_j \right) r_j(E_m) \left( V_j r_j(A) - \ell_j(A) V_j \right)^* \ell_j(E_k) \right] \end{aligned}$$

$$\begin{aligned}
 &= \sum_{j \in \mathcal{J}} \sum_{k,m} \theta_j^{km} \tau_j \left[ V_j r_j(B^*) r_j(E_m) r_j(A^*) V_j^* \ell_j(E_k) - V_j r_j(B^*) r_j(E_m) V_j^* \ell_j(A^*) \ell_j(E_k) \right. \\
 &\quad \left. - \ell_j(B^*) V_j r_j(E_m) r_j(A^*) V_j^* \ell_j(E_k) + \ell_j(B^*) V_j r_j(E_m) V_j^* \ell_j(A^*) \ell_j(E_k) \right] \\
 &= \sum_{j \in \mathcal{J}} \sum_{k,m} \theta_j^{km} \tau_j \left[ V_j^* \ell_j(E_k) V_j r_j(B^* E_m A^*) - V_j^* \ell_j(A^* E_k) V_j r_j(B^* E_m) \right. \\
 &\quad \left. - V_j^* \ell_j(E_k B^*) V_j r_j(E_m A^*) + V_j^* \ell_j(A^* E_k B^*) V_j r_j(E_m) \right].
 \end{aligned}$$

Thus, using (4.17), and then changing  $j$  by  $j^*$  and using that  $\theta_j^{km} = \theta_{j^*}^{mk}$  by Assumption 7.2, we obtain

$$\begin{aligned}
 &\langle B, (\mathcal{K}_\rho A)^* \rangle_{L^2(\tau)} \\
 &= \sum_{j \in \mathcal{J}} \sum_{k,m} \theta_j^{km} \tau_j \left[ V_j \ell_{j^*}(B^* E_m A^*) V_j^* r_{j^*}(E_k) - V_j \ell_{j^*}(B^* E_m) V_j^* r_{j^*}(A^* E_k) \right. \\
 &\quad \left. - V_j \ell_{j^*}(E_m A^*) V_j^* r_{j^*}(E_k B^*) + V_j \ell_{j^*}(E_m) V_j^* r_{j^*}(A^* E_k B^*) \right] \\
 &= \sum_{j \in \mathcal{J}} \sum_{k,m} \theta_j^{mk} \tau_j \left[ V_j^* \ell_j(B^* E_m A^*) V_j r_j(E_k) - V_j^* \ell_j(B^* E_m) V_j r_j(A^* E_k) \right. \\
 &\quad \left. - V_j^* \ell_j(E_m A^*) V_j r_j(E_k B^*) + V_j^* \ell_j(E_m) V_j r_j(A^* E_k B^*) \right]
 \end{aligned}$$

which coincides with (7.6) after interchanging  $m$  and  $k$ . □

The following result expressing the unique solvability of the continuity equation is now an immediate consequence.

**Corollary 7.4** *For  $\rho \in \mathfrak{P}_+$ , the linear mapping  $\mathcal{K}_\rho$  is a bijection on  $\mathcal{A}_0$  that depends smoothly ( $C^\infty$ ) on  $\rho$ .*

**Proof** It follows from Lemma 7.3 that  $\mathcal{K}_\rho$  maps  $\mathcal{A}_0$  into itself. Since the restriction of a self-adjoint operator to its range is injective, the result follows. Smooth dependence on  $\rho$  follows from the smoothness of  $\theta$ . □

The following elementary variational characterization is of interest.

**Proposition 7.5** *Fix  $\rho \in \mathfrak{P}_+$  and  $v \in \mathcal{A}_0$ . Among all vector fields  $\mathbf{B} \in \mathcal{B}$  satisfying the continuity equation*

$$v + \operatorname{div}(\widehat{\rho} \# \mathbf{B}) = 0 \tag{7.7}$$

*there is a unique one that is a gradient. Moreover, among all vector fields  $\mathbf{B}$  solving (7.7), this vector field is the unique minimizer of the “kinetic energy functional”  $\mathcal{E}_\rho$  given by*

$$\mathcal{E}_\rho(\mathbf{B}) = \sum_{j \in \mathcal{J}} \langle \widehat{\rho}_j \# B_j, B_j \rangle_{L^2(\tau_j)}.$$

**Proof** Existence of a gradient vector  $\mathbf{B}$  field solving (7.7) follows from Corollary 7.4. To prove uniqueness, suppose that  $\operatorname{div}(\widehat{\rho} \# \nabla A) = -v = \operatorname{div}(\widehat{\rho} \# \nabla \widetilde{A})$  for some  $A, \widetilde{A} \in \mathcal{A}$ . This means that  $\mathcal{K}_\rho A = \mathcal{K}_\rho \widetilde{A}$ , hence Lemma 7.3 yields  $\nabla A = \nabla \widetilde{A}$ . The remaining part follows along the lines of the proof of [8, Theorem 3.17]. □

We are now ready to define a class of Riemannian metrics that are the main object of study in this paper.

**Definition 7.6** (*Quantum transport metric*) Fix  $\rho \in \mathfrak{P}_+$  and let  $\theta = (\theta_j)_j$  satisfy Assumption 7.2. The associated quantum transport metric is the Riemannian metric on  $\mathcal{M}_\rho$  induced by the operator  $\mathcal{K}_\rho$ , i.e., for  $\dot{\rho}_1, \dot{\rho}_2 \in \mathcal{A}_0$ ,

$$\langle \dot{\rho}_1, \dot{\rho}_2 \rangle_\rho = \langle \mathcal{K}_\rho^{-1} \dot{\rho}_1, \dot{\rho}_2 \rangle_{L^2(\tau)},$$

or, more explicitly,

$$\begin{aligned} \langle \dot{\rho}_1, \dot{\rho}_2 \rangle &= \langle \nabla A_1, \widehat{\rho} \# \nabla A_2 \rangle_{L^2(\tau)} \\ &= \sum_j \langle \partial_j A_1, \theta_j(\ell_j(\rho), r_j(\rho)) \# \partial_j A_2 \rangle_{L^2(\tau_j)} \quad \text{for } \rho \in \mathfrak{P}_+, \end{aligned} \tag{7.8}$$

where, for  $i = 1, 2$ ,  $A_i$  is the unique solution in  $\mathcal{A}_0$  to the continuity equation

$$\dot{\rho}_i + \operatorname{div}(\widehat{\rho} \# \nabla A_i) = 0.$$

It follows from Lemma 7.3 and Corollary 7.4 that  $\mathcal{K}_\rho$  indeed induces a Riemannian metric on  $\mathcal{M}_\rho$ .

### 7.2 Gradient Flows of Entropy Functionals

In this section we shall show that various evolution equations of interest can be interpreted as gradient flow equations with respect to suitable quantum transport metrics introduced in Sect. 7.1.

We consider the operator  $\mathcal{K}_\rho : \mathcal{A}_0 \rightarrow \mathcal{A}_0$  given by

$$\mathcal{K}_\rho A := \sum_{j \in \mathcal{J}} \partial_j^\dagger (\widehat{\rho}_j \# \partial_j A),$$

where  $\widehat{\rho}_j = \theta_j(\ell_j(\rho), r_j(\rho))$  is defined in terms of a well-chosen function  $\theta_j$  that depends on the context and will be determined below.

**Theorem 7.7** (Gradient flow structure for the relative entropy) *Consider the operator  $\mathcal{K}_\rho$  defined using the functions  $\theta_j$  given by  $\theta_j(r, s) := \Lambda(e^{\omega_j/2}r, e^{-\omega_j/2}s)$ , where  $\Lambda(r, s) = \frac{r-s}{\log r - \log s}$  is the logarithmic mean. Then we have the identity*

$$\mathcal{L}^\dagger \rho = -\mathcal{K}_\rho \operatorname{DEnt}_\sigma(\rho)$$

for all  $\rho \in \mathfrak{P}_+$ , thus the gradient flow equation for the relative von Neumann entropy functional  $\operatorname{Ent}_\sigma$  with respect to the Riemannian metric on  $\mathcal{M}_\sigma$  induced by  $(\mathcal{K}_\rho)_{\rho \in \mathcal{M}_\sigma}$  is the Kolmogorov forward equation  $\partial_t \rho = \mathcal{L}^\dagger \rho$ .

This result generalises the gradient flow structure from [10,36] as described in Sect. 2. The proof relies on the following version of the chain rule.

**Lemma 7.8** (Chain rule for the logarithm) *Define  $\theta_j(r, s) := \Lambda(e^{\omega_j/2}r, e^{-\omega_j/2}s)$ , where  $\Lambda(r, s) = \frac{r-s}{\log r - \log s}$  is the logarithmic mean. Then, for all  $\rho \in \mathfrak{P}_+$  we have*

$$e^{-\omega_j/2} V_j r_j(\rho) - e^{\omega_j/2} \ell_j(\rho) V_j = \widehat{\rho}_j \# \partial_j (\log \rho - \log \sigma). \tag{7.9}$$

**Proof** Using (4.18) we infer that

$$\partial_j(\log \rho - \log \sigma) = V_j \log(e^{-\omega_j/2} r_j(\rho)) - \log(e^{\omega_j/2} \ell_j(\rho)) V_j .$$

We consider the spectral decomposition  $\rho = \sum_k \lambda_k E_k$  as before, and observe that

$$\widehat{\rho}_j = \theta_j(\ell_j(\rho), r_j(\rho)) = \sum_{k,m} \Lambda(e^{\omega_j/2} \lambda_k, e^{-\omega_j/2} \lambda_m) \ell_j(E_k) \otimes r_j(E_m) .$$

Using this identity, we obtain

$$\begin{aligned} & \widehat{\rho}_j \# (\partial_j(\log \rho - \log \sigma)) \\ &= \sum_{k,m,p} \Lambda(e^{\omega_j/2} \lambda_k, e^{-\omega_j/2} \lambda_m) \ell_j(E_k) \\ & \quad \times \left( \log(e^{-\omega_j/2} \lambda_p) V_j r_j(E_p) - \log(e^{\omega_j/2} \lambda_p) \ell_j(E_p) V_j \right) r_j(E_m) \\ &= \sum_{k,m} \Lambda(e^{\omega_j/2} \lambda_k, e^{-\omega_j/2} \lambda_m) \left( \log(e^{-\omega_j/2} \lambda_m) - \log(e^{\omega_j/2} \lambda_k) \right) \ell_j(E_k) V_j r_j(E_m) \\ &= \sum_{k,m} (e^{-\omega_j/2} \lambda_m - e^{\omega_j/2} \lambda_k) \ell_j(E_k) V_j r_j(E_m) \\ &= e^{-\omega_j/2} V_j r_j(\rho) - e^{\omega_j/2} \ell_j(\rho) V_j , \end{aligned}$$

which yields (7.9). □

**Proof of Theorem 7.7** Since  $D \text{Ent}_\sigma(\rho) = \log \rho - \log \sigma$ , the chain rule from Lemma 7.8 yields, using Proposition 4.9,

$$\begin{aligned} \mathcal{K}_\rho D \text{Ent}_\sigma(\rho) &= \sum_{j \in \mathcal{J}} \partial_j^\dagger (\widehat{\rho}_j \# \partial_j(\log \rho - \log \sigma)) \\ &= \sum_{j \in \mathcal{J}} \partial_j^\dagger (e^{-\omega_j/2} V_j r_j(\rho) - e^{\omega_j/2} \ell_j(\rho) V_j) \\ &= \sum_{j \in \mathcal{J}} e^{-\omega_j/2} \left\{ r_j^\dagger (V_j^* V_j r_j(\rho)) - \ell_j^\dagger (V_j r_j(\rho) V_j^*) \right\} \\ & \quad - e^{\omega_j/2} \left\{ r_j^\dagger (V_j^* \ell_j(\rho) V_j) - \ell_j^\dagger (\ell_j(\rho) V_j V_j^*) \right\} , \end{aligned}$$

which equals the expression for  $-\mathcal{L}^\dagger \rho$  given in Proposition 4.10. □

Let us now consider the special case where  $\sigma = \mathbf{1}$ . Then (4.10) reduces to  $\omega_j = 0$  for all  $j \in \mathcal{J}$ , and we will be able to formulate a natural nonlinear generalization of Theorem 7.7. Let  $f \in C^2((0, \infty); \mathbb{R})$  be strictly convex, and consider the functional  $\mathcal{F} : \mathfrak{P}_+ \rightarrow \mathbb{R}$  given by

$$\mathcal{F}(\rho) = \tau[f(\rho)] ,$$

where  $f(\rho)$  is interpreted in the sense of functional calculus. Let  $\varphi \in C^1((0, \infty); \mathbb{R})$  be strictly increasing, and consider the operator  $\mathcal{K}_\rho$  as defined before, with  $\theta_j = \theta$  given by

$$\theta(\lambda, \mu) = \begin{cases} \frac{\varphi(\lambda) - \varphi(\mu)}{f'(\lambda) - f'(\mu)}, & \text{if } \lambda \neq \mu, \\ \frac{\varphi'(\lambda)}{f''(\lambda)}, & \text{otherwise.} \end{cases} \tag{7.10}$$

The following result is a non-commutative analogue of a seminal result by Otto [38], which states that the porous medium equation is the gradient flow equation for the Rényi entropy in with respect to the 2-Kantorovich metric.

**Theorem 7.9** (Gradient flow structures with general entropy functionals) *Consider a differentiable structure with  $\sigma = \mathbf{1}$ , and let  $\theta_j$  be given by (7.10). Then we have the identity*

$$\mathcal{L}\varphi(\rho) = \mathcal{L}^\dagger\varphi(\rho) = -\mathcal{K}_\rho D\mathcal{F}(\rho) \tag{7.11}$$

for  $\rho \in \mathfrak{P}_+$ , thus the gradient flow equation for  $\mathcal{F}$  with respect to the Riemannian metric on  $\mathcal{M}_1$  induced by  $(\mathcal{K}_\rho)_{\rho \in \mathcal{M}_1}$  is given by

$$\partial_t \rho = \mathcal{L}\varphi(\rho) .$$

**Proof** The first identity in (7.11) follows immediately from the construction of  $\mathcal{L}$  since  $\sigma = \mathbf{1}$ . The chain rule (6.10) implies that the derivative of  $\mathcal{F}$  is given by

$$D\mathcal{F}(\rho) = f'(\rho) .$$

Recalling (6.5), we note that  $\theta_j$  is defined to satisfy the identity  $\theta \cdot \delta f' = \delta\varphi$ . Using (6.4), (7.10), and the chain rule from Proposition 6.2 we infer that

$$\begin{aligned} \widehat{\rho}_j \# \partial_j D\mathcal{F}(\rho) &= \theta(\ell_j(\rho), r_j(\rho)) \# (\delta f'(\ell_j(\rho), r_j(\rho)) \# \partial_j \rho) \\ &= \delta\varphi(\ell_j(\rho), r_j(\rho)) \# \partial_j \rho \\ &= \partial_j \varphi(\rho) . \end{aligned}$$

We obtain

$$\mathcal{K}_\rho D\mathcal{F}(\rho) = \sum_{j \in \mathcal{J}} \partial_j^\dagger (\widehat{\rho}_j \# \partial_j D\mathcal{F}(\rho)) = \sum_{j \in \mathcal{J}} \partial_j^\dagger \partial_j \varphi(\rho) = -\mathcal{L}\varphi(\rho) ,$$

which is the desired identity. □

**Remark 7.10** The result remains true if  $f$  is required to be strictly concave and  $\varphi$  is required to be strictly decreasing. Note that  $\theta$  is positive in this case, so that  $(\mathcal{K}_\rho)_\rho$  induces a Riemannian metric.

**Remark 7.11** This result contains various known results as special cases. Take  $f(\lambda) = \lambda \log \lambda$  and  $\varphi(r) = r$ . Then the functional  $\mathcal{F}$  is the von Neumann entropy  $\mathcal{F}(\rho) = \tau[\rho \log \rho]$ , and we recover the special case of Theorem 7.7 with  $\sigma = \mathbf{1}$ . It also contains the gradient flow structure for the fermionic Fokker-Planck equation from [8]. In the special case where  $\mathcal{L}$  is the generator of a reversible Markov chain, we recover the gradient flow structure for discrete porous medium equations obtained in [19].

**Remark 7.12** In some situations the expression for  $\widehat{\rho}_j = \theta(\ell_j(\rho), r_j(\rho))$  can be simplified. If  $f(\lambda) = \lambda \log \lambda$  and  $\varphi(\lambda) = \lambda$ , it follows that  $\theta(\lambda, \mu) = \frac{\lambda - \mu}{\log \lambda - \log \mu}$  is the logarithmic mean. The integral representation  $\theta(\lambda, \mu) = \int_0^1 \lambda^{1-s} \mu^s ds$  allows one to express  $\widehat{\rho}_j$  in terms of the functional calculus for  $\ell_j(\rho)$  and  $r_j(\rho)$ :

$$\widehat{\rho}_j = \theta(\ell_j(\rho), r_j(\rho)) = \int_0^1 \ell_j(\rho)^{1-s} \otimes r_j(\rho)^s ds .$$

More generally, take  $m \in \mathbb{R} \setminus \{0, 1\}$ , and set  $\varphi(\lambda) = \lambda^m$  and  $f(\lambda) = \frac{1}{m-1}\lambda^m$ . We shall consider the *power difference means* defined by

$$\theta_m(\lambda, \mu) := \frac{m-1}{m} \frac{\lambda^m - \mu^m}{\lambda^{m-1} - \mu^{m-1}},$$

with the convention that  $\theta_m(\lambda, \lambda) = \lambda$ . A systematic study of the operator means associated to these functions has been carried out in [25]. Various classical means are contained as special cases:

$$\theta_m(\lambda, \mu) = \begin{cases} \frac{2\lambda\mu}{\lambda + \mu}, & m = -1 \text{ (harmonic mean)}, \\ \frac{\lambda\mu(\log \lambda - \log \mu)}{\lambda - \mu}, & m \rightarrow 0, \\ \sqrt{\lambda\mu}, & m = \frac{1}{2} \text{ (geometric mean)}, \\ \frac{\lambda - \mu}{\log \lambda - \log \mu}, & m \rightarrow 1 \text{ (logarithmic mean)}, \\ \frac{\lambda + \mu}{2}, & m = 2 \text{ (arithmetic mean)}. \end{cases}$$

The following integral representation holds:

$$\theta_m(\lambda, \mu) = \int_0^1 \left( (1-\alpha)\lambda^{m-1} + \alpha\mu^{m-1} \right)^{\frac{1}{m-1}} d\alpha. \tag{7.12}$$

If  $m = 2$  and  $\ell_j = r_j = I_{\mathcal{A}}$ , one has  $\widehat{\rho}_j \# A = \frac{1}{2}(\rho A + A\rho)$ , which corresponds to the anti-commutator case studied in [12].

Another special case is obtained by taking  $\varphi(\lambda) = \lambda$  and  $f(\lambda) = \lambda^2/2$ , which yields  $\theta(\lambda, \mu) \equiv 1$ , so that  $\mathcal{K}_\rho = -\mathcal{L}$  for all  $\rho$ , and  $\mathcal{F}(\rho) = \frac{1}{2}\tau[\rho^2] = \frac{1}{2}\|\rho\|_{L^2(\tau)}^2$ . In this case, the distance associated to  $\mathcal{K}_\rho$  may be regarded as a non-commutative analogue of the Sobolev  $H^{-1}$ -metric.

### 7.3 Geodesics

As before we consider the operator  $\mathcal{K}_\rho : \mathcal{A}_0 \rightarrow \mathcal{A}_0$  given by

$$\mathcal{K}_\rho A := \partial_j^\dagger (\widehat{\rho}_j \# \partial_j A).$$

For fixed  $\bar{\rho} \in \mathfrak{P}_+$  we will compute the geodesic equations associated to the Riemannian structure on  $\mathcal{M}_{\bar{\rho}}$  induced by the operator  $(\mathcal{K}_\rho)_\rho$ . The Riemannian distance  $d_{\mathcal{K}}$  is given by

$$\begin{aligned} d_{\mathcal{K}}(\tilde{\rho}_0, \tilde{\rho}_1)^2 &= \inf_{\rho, A} \left\{ \int_0^1 \langle \mathcal{K}_{\rho_t} A_t, A_t \rangle_{L^2(\mathcal{A}, \tau)} dt : \partial_t \rho_t = \mathcal{K}_{\rho_t} A_t, \rho_0 = \tilde{\rho}_0, \rho_1 = \tilde{\rho}_1 \right\} \\ &= \inf_{\rho, A} \left\{ \int_0^1 \langle \widehat{\rho}_t \# \nabla A_t, \nabla A_t \rangle_{L^2(\mathcal{A}, \tau)} dt : \right. \\ &\quad \left. \partial_t \rho_t + \operatorname{div}(\widehat{\rho}_t \# \nabla A_t) = 0, \rho_0 = \tilde{\rho}_0, \rho_1 = \tilde{\rho}_1 \right\}, \end{aligned}$$

where the infimum runs over smooth curves  $\{\rho_t\}_{t \in [0,1]}$  in  $\mathcal{M}_{\bar{\rho}}$  and  $\{A_t\}_{t \in [0,1]}$  in  $\mathcal{A}_0$  satisfying the stated conditions.

The geodesic equations are the Euler–Lagrange equations associated to this constrained minimization problem, given by

$$\begin{aligned} \partial_s \rho_s &= \mathcal{K}_{\rho_s} A_s \\ \partial_s A_s &= -\frac{1}{2} \langle D \mathcal{K}_{\rho_s} A_s, A_s \rangle_{L^2(\mathcal{A}, \tau)}. \end{aligned} \tag{7.13}$$

Note that the latter equation is equivalent to

$$\partial_s \tau[A_s B] = -\frac{1}{2} \langle D_B \mathcal{K}_{\rho_s} A_s, A_s \rangle_{L^2(\mathcal{A}, \tau)}$$

for  $B \in \mathcal{A}_0$ , where  $D_B \mathcal{K}_{\rho} = \lim_{\varepsilon \rightarrow 0} \varepsilon^{-1} (\mathcal{K}_{\rho + \varepsilon B} - \mathcal{K}_{\rho})$  denotes the directional derivative.

**Proposition 7.13** (Geodesic equations) *The geodesic equations for  $(\rho_s, A_s)_s$  are given by*

$$\partial_s \rho_s + \operatorname{div}(\widehat{\rho}_s \# \nabla A_s) = 0, \tag{7.14}$$

$$\partial_s A_s + \Phi(\rho_s, A_s) = 0, \tag{7.15}$$

where

$$\begin{aligned} \Phi(\rho, A) &= \sum_{j \in \mathcal{J}} \sum_{k, m, p} \delta \theta_j \left( \begin{array}{c} \lambda_k \quad \cdot \quad \lambda_m \\ \cdot \quad \cdot \quad \cdot \\ \lambda_p \end{array} \right) E_m \ell_j^\dagger \left( (\partial_j A) r_j(E_p) (\partial_j A)^* \right) E_k \\ &= \sum_{j \in \mathcal{J}} \sum_{k, m, p} \delta \theta_j \left( \begin{array}{c} \lambda_p \quad \lambda_m \\ \cdot \quad \cdot \\ \lambda_k \end{array} \right) E_k r_j^\dagger \left( (\partial_j A)^* \ell_j(E_p) \partial_j A \right) E_m. \end{aligned}$$

Here,  $\rho = \sum_k \lambda_k E_k$  denotes the spectral decomposition of  $\rho$ .

**Remark 7.14** In the sequel we will use (7.15) in the weak formulation:

$$\partial_s \tau[A_s B] = - \sum_{j \in \mathcal{J}} \tau_j \left[ (\partial_j A_s)^* \mathcal{N}_{\rho, B}^{(\eta), j} \# (\partial_j A_s) \right], \tag{7.16}$$

for all  $B \in \mathcal{A}$  and  $\eta = 1, 2$ , where

$$\mathcal{N}_{\rho, B}^{(1), j} = \delta \theta_j \left( \begin{array}{c} \ell_j(\rho) \quad \cdot \quad \ell_j(B) \\ \cdot \quad \cdot \quad \cdot \\ r_j(\rho) \end{array} \right), \quad \mathcal{N}_{\rho, B}^{(2), j} = \delta \theta_j \left( \begin{array}{c} r_j(\rho) \quad \cdot \quad r_j(B) \\ \cdot \quad \cdot \quad \cdot \\ \ell_j(\rho) \end{array} \right).$$

**Remark 7.15** If  $\theta_j(r, s) := \Lambda(e^{\omega_j/2} r, e^{-\omega_j/2} s)$  where  $\Lambda$  is the logarithmic mean, the expression above can be simplified. In this case we have the integral representation

$$\delta \Lambda \left( \begin{array}{c} y \quad \cdot \quad z \\ \cdot \quad \cdot \quad \cdot \\ x \end{array} \right) = \int_0^1 \int_0^1 \int_0^\alpha \left[ x^{1-\alpha} \cdot \frac{y^{\alpha-\beta}}{(1-s) + sy} \cdot \frac{z^\beta}{(1-s) + sz} \right] d\beta \, d\alpha \, ds,$$

so that

$$\begin{aligned} \delta \theta_j \left( \begin{array}{c} y \quad \cdot \quad z \\ \cdot \quad \cdot \quad \cdot \\ x \end{array} \right) &= \int_0^1 \int_0^1 \int_0^\alpha e^{-\omega_j \alpha} \left[ x^{1-\alpha} \cdot \frac{y^{\alpha-\beta}}{(1-s) + se^{-\omega_j/2} y} \cdot \frac{z^\beta}{(1-s) + se^{-\omega_j/2} z} \right] d\beta \, d\alpha \, ds, \end{aligned}$$

which implies that

$$\begin{aligned} \Phi(\rho, A) &= \sum_{j \in \mathcal{J}} \int_0^1 \int_0^1 \int_0^\alpha e^{\omega_j \alpha} \left[ \frac{\rho^{\alpha-\beta}}{(1-s)I + se^{\omega_j/2} \rho} \ell_j^\dagger \left( (\partial_j A) r_j (\rho^{1-\alpha}) (\partial_j A)^* \right) \right. \\ &\quad \left. \times \frac{\rho^\beta}{(1-s)I + se^{\omega_j/2} \rho} \right] d\beta \, d\alpha \, ds \\ &= \sum_{j \in \mathcal{J}} \int_0^1 \int_0^1 \int_0^\alpha e^{-\omega_j \alpha} \left[ \frac{\rho^{\alpha-\beta}}{(1-s)I + se^{-\omega_j/2} \rho} r_j^\dagger \left( (\partial_j A)^* \ell_j (\rho^{1-\alpha}) \partial_j A \right) \right. \\ &\quad \left. \times \frac{\rho^\beta}{(1-s)I + se^{-\omega_j/2} \rho} \right] d\beta \, d\alpha \, ds . \end{aligned}$$

**Proof of Proposition 7.13** Proposition 6.8 yields

$$\begin{aligned} \partial_\varepsilon|_{\varepsilon=0} \theta_j \left( \ell_j(\rho + \varepsilon B), r_j(\rho + \varepsilon B) \right) \\ = \delta \theta_j \left( \ell_j^{(\rho)} \underset{\cdot}{\leftarrow} \ell_j^{(B)} \underset{\cdot}{\leftarrow} r_j^{(\rho)} \right) + \delta \theta_j \left( \ell_j^{(\rho)} \underset{\cdot}{\leftarrow} r_j^{(\rho)} \underset{\cdot}{\leftarrow} r_j^{(B)} \underset{\cdot}{\leftarrow} r_j^{(\rho)} \right) , \end{aligned}$$

and therefore

$$\begin{aligned} \partial_\varepsilon|_{\varepsilon=0} \langle \mathcal{K}_{\rho+\varepsilon B} A, A \rangle_{L^2(\mathcal{A}, \tau)} \\ = \sum_{j \in \mathcal{J}} \tau_j \left[ (\partial_j A)^* \left\{ \delta \theta_j \left( \ell_j^{(\rho)} \underset{\cdot}{\leftarrow} \ell_j^{(B)} \underset{\cdot}{\leftarrow} r_j^{(\rho)} \right) + \delta \theta_j \left( \ell_j^{(\rho)} \underset{\cdot}{\leftarrow} r_j^{(\rho)} \underset{\cdot}{\leftarrow} r_j^{(B)} \underset{\cdot}{\leftarrow} r_j^{(\rho)} \right) \right\} \right] . \end{aligned}$$

Since  $A$  is self-adjoint, it follows using (7.1) and (4.17) that

$$\begin{aligned} \tau_j \left[ (\partial_j^* A)^* \delta \theta_j^* \left( \ell_{j^*}^{(\rho)} \underset{\cdot}{\leftarrow} \ell_{j^*}^{(B)} \underset{\cdot}{\leftarrow} r_{j^*}^{(\rho)} \underset{\cdot}{\leftarrow} r_{j^*}^{(B)} \underset{\cdot}{\leftarrow} r_{j^*}^{(\rho)} \right) \right] \\ = \sum_{k,m,p} \delta \theta_{j^*} \left( \lambda_p \underset{\cdot}{\leftarrow} \lambda_k \underset{\cdot}{\leftarrow} \lambda_m \right) \tau_j \left[ (\partial_j^* A)^* \ell_{j^*} (E_p) (\partial_j^* A) r_{j^*} (E_k B E_m) \right] \\ = \sum_{k,m,p} \delta \theta_j \left( \lambda_k \underset{\cdot}{\leftarrow} \lambda_m \underset{\cdot}{\leftarrow} \lambda_p \right) \\ \quad \times \tau_j \left[ \left( r_{j^*} (A) V_j - V_j \ell_{j^*} (A) \right) \ell_{j^*} (E_p) \left( V_j^* r_{j^*} (A) - \ell_{j^*} (A) V_j^* \right) r_{j^*} (E_k B E_m) \right] \\ = \sum_{k,m,p} \delta \theta_j \left( \lambda_k \underset{\cdot}{\leftarrow} \lambda_m \underset{\cdot}{\leftarrow} \lambda_p \right) \\ \quad \times \tau_j \left[ \left( \ell_j (A) V_j - V_j r_j (A) \right) r_j (E_p) \left( V_j^* \ell_j (A) - r_j (A) V_j^* \right) \ell_j (E_k B E_m) \right] \\ = \sum_{k,m,p} \delta \theta_j \left( \lambda_k \underset{\cdot}{\leftarrow} \lambda_m \underset{\cdot}{\leftarrow} \lambda_p \right) \tau_j \left[ (\partial_j A)^* \ell_j (E_k B E_m) (\partial_j A) r_j (E_p) \right] \\ = \tau_j \left[ (\partial_j A)^* \delta \theta_j \left( \ell_j^{(\rho)} \underset{\cdot}{\leftarrow} \ell_j^{(B)} \underset{\cdot}{\leftarrow} r_j^{(\rho)} \right) \right] . \end{aligned}$$

This implies the equality of the two sums in (7.16), and it also follows that

$$\partial_\varepsilon|_{\varepsilon=0} \langle \mathcal{K}_{\rho+\varepsilon B} A, A \rangle_{L^2(\mathcal{A}, \tau)} = 2 \sum_{j \in \mathcal{J}} \tau_j \left[ (\partial_j A)^* \delta \theta_j \left( \ell_j^{(\rho)} \underset{\cdot}{\leftarrow} \ell_j^{(B)} \underset{\cdot}{\leftarrow} r_j^{(\rho)} \right) \right] , \tag{7.17}$$



which yields the weak formulation (7.16) in view of (7.13). To obtain (7.15), we compute using (4.4),

$$\begin{aligned} & \tau_j \left[ (\partial_j A)^* \delta \theta_j \left( \begin{matrix} \ell_j^{(\rho)} & - & \ell_j^{(\rho)} \\ \partial_j A & - & r_j(\rho) \end{matrix} \right) \right] \\ &= \sum_{k,m,p} \delta \theta_j \left( \begin{matrix} \lambda_k & \leftarrow & \lambda_m \\ \cdot & \leftarrow & \cdot \\ \lambda_p & & \lambda_p \end{matrix} \right) \tau_j \left[ \ell_j(BE_m)(\partial_j A)r_j(E_p)(\partial_j A)^* \ell_j(E_k) \right] \\ &= \sum_{k,m,p} \delta \theta_j \left( \begin{matrix} \lambda_k & \leftarrow & \lambda_m \\ \cdot & \leftarrow & \cdot \\ \lambda_p & & \lambda_p \end{matrix} \right) \tau_j \left[ B \ell_j^\dagger \left( \ell_j(E_m)(\partial_j A)r_j(E_p)(\partial_j A)^* \ell_j(E_k) \right) \right] \\ &= \sum_{k,m,p} \delta \theta_j \left( \begin{matrix} \lambda_k & \leftarrow & \lambda_m \\ \cdot & \leftarrow & \cdot \\ \lambda_p & & \lambda_p \end{matrix} \right) \tau_j \left[ BE_m \ell_j^\dagger \left( (\partial_j A)r_j(E_p)(\partial_j A)^* \right) E_k \right] \\ &= \tau_j \left[ B \Phi_j(\rho, A) \right], \end{aligned}$$

where

$$\Phi_j(\rho, A) = \sum_{k,m,p} \delta \theta_j \left( \begin{matrix} \lambda_k & \leftarrow & \lambda_m \\ \cdot & \leftarrow & \cdot \\ \lambda_p & & \lambda_p \end{matrix} \right) E_m \ell_j^\dagger \left( (\partial_j A)r_j(E_p)(\partial_j A)^* \right) E_k .$$

An analogous computation shows that

$$\Phi_j(\rho, A) = \sum_{k,m,p} \delta \theta_{j^*} \left( \begin{matrix} \lambda_p & \leftarrow & \lambda_m \\ \cdot & \leftarrow & \cdot \\ \lambda_k & & \lambda_k \end{matrix} \right) E_k r_{j^*}^\dagger \left( (\partial_{j^*} A)^* \ell_{j^*}(E_p) \partial_{j^*} A \right) E_m .$$

We thus obtain

$$\langle D_{\mathcal{K}_\rho} A, A \rangle_{L^2(\mathcal{A}, \tau)} = 2\Phi(\rho, A) ,$$

hence the result follows from the Euler–Lagrange equations (7.13). □

We will use the geodesic equations to compute the Hessian of some interesting functionals on  $\mathcal{M}_\rho$ . Note that the Hessian is obtained from the formula

$$\text{Hess}_{\mathcal{K}} \mathcal{E}(\rho_0)[A_0, A_0] := \partial_s^2|_{s=0} \mathcal{E}(\rho_s)$$

for  $A_0 \in \mathcal{A}_0$ , where  $(\rho_s, A_s)_s$  evolves according to the geodesic equations (7.13) with initial conditions  $\rho|_{s=0} = \rho_0$  and  $\partial_s|_{s=0} \rho_s = \mathcal{K}_{\rho_0} A_0$ .

**Proposition 7.16** *For  $\bar{\rho} \in \mathfrak{P}_+$ , let  $\mathcal{E} : \mathcal{M}_{\bar{\rho}} \rightarrow \mathbb{R}$  be a smooth functional, and let us write  $\mathcal{M}(\rho) := \mathcal{K}_\rho D\mathcal{F}(\rho)$  for the Riemannian gradient of  $\mathcal{F}$  induced by  $(\mathcal{K}_\rho)_\rho$ . Then, the Hessian of  $\mathcal{E}$  is given by*

$$\text{Hess}_{\mathcal{K}} \mathcal{E}(\rho)[A, A] = \tau[AD_{\mathcal{K}_\rho A} \mathcal{M}(\rho)] - \tau \left[ (\nabla A)^* \mathcal{N}_{\rho, \mathcal{M}(\rho)}^{(\eta)} \# (\nabla A) \right] \tag{7.18}$$

for  $A \in \mathcal{A}_0$  and  $\eta = 1, 2$ , where  $D_B \mathcal{M}(\rho) = \lim_{\varepsilon \rightarrow 0} \varepsilon^{-1} (\mathcal{M}(\rho + \varepsilon B) - \mathcal{M}(\rho))$  denotes the directional derivative. In particular, if  $\mathcal{M}(\rho) = -\mathcal{L}^\dagger \rho$  (as is the case in setting of Theorem 7.7, where  $\mathcal{E}(\rho) = \text{Ent}_\sigma(\rho)$ ), we have

$$\text{Hess}_{\mathcal{K}} \mathcal{E}(\rho)[A, A] = -\tau[(\nabla \mathcal{L} A)^* \widehat{\rho} \# \nabla A] + \tau \left[ (\nabla A)^* \mathcal{N}_{\rho, \mathcal{L}^\dagger \rho}^{(\eta)} \# (\nabla A) \right]. \tag{7.19}$$

**Proof** Let  $(\rho_s, A_s)_s$  satisfy the geodesic equations (7.14)–(7.15). Then:

$$\partial_s \mathcal{E}(\rho_s) = \tau[\mathbf{D}\mathcal{E}(\rho_s)\partial_s \rho_s] = \tau[\mathbf{D}\mathcal{E}(\rho_s)\mathcal{K}_{\rho_s} A_s] = \tau[A_s \mathcal{K}_{\rho_s} \mathbf{D}\mathcal{E}(\rho_s)] = \tau[A_s \mathcal{M}(\rho_s)].$$

Thus, by (7.16),

$$\begin{aligned} \partial_s^2 \mathcal{E}(\rho_s) &= \tau[A_s \partial_s \mathcal{M}(\rho_s)] + \tau[(\partial_s A_s) \cdot \mathcal{M}(\rho_s)] \\ &= \tau[A_s \mathbf{D}_{\mathcal{K}_{\rho_s} A_s} \mathcal{M}(\rho_s)] - \sum_{j \in \mathcal{J}} \tau_j \left[ (\partial_j A_s)^* \mathcal{N}_{\rho_s, \mathcal{L}^\dagger \rho_s}^{(\eta), j} \# (\partial_j A_s) \right], \end{aligned}$$

for  $\eta = 1, 2$ , which proves (7.18).

If  $\mathcal{M}(\rho) = -\mathcal{L}^\dagger \rho$  we have  $\mathbf{D}_B \mathcal{M}(\rho) = -\mathcal{L}^\dagger B$ , hence the expression above simplifies to

$$\begin{aligned} \partial_s^2 \mathcal{E}(\rho_s) &= -\tau[A_s \mathcal{L}^\dagger \mathcal{K}_{\rho_s} A_s] + \sum_{j \in \mathcal{J}} \tau_j \left[ (\partial_j A_s)^* \mathcal{N}_{\rho_s, \mathcal{L}^\dagger \rho_s}^{(\eta), j} \# (\partial_j A_s) \right] \\ &= -\tau[(\nabla \mathcal{L} A_s)^* \widehat{\rho}_s \# \nabla A_s] + \sum_{j \in \mathcal{J}} \tau_j \left[ (\partial_j A_s)^* \mathcal{N}_{\rho_s, \mathcal{L}^\dagger \rho_s}^{(\eta), j} \# (\partial_j A_s) \right]. \end{aligned}$$

□

**Remark 7.17** In the setting of the theorem above, we remark that the following equivalent expression holds as well:

$$\text{Hess}_{\mathcal{K}} \mathcal{E}(\rho)[A, A] = \tau[\mathbf{A} \mathbf{D}_{\mathcal{K}_{\rho} A} \mathcal{M}(\rho)] - \tau[\Phi(\rho, A) \cdot \mathcal{M}(\rho)].$$

### 8 Preliminaries on Quasi-entropies

In this section we collect some known results on trace functionals that will be useful in the study of quantum transport metrics. Special cases of the results in this section already played a key role in the proof of functional inequalities in [10].

Let  $\mathcal{A}$  be a finite-dimensional von Neumann algebra endowed with a positive tracial linear functional  $\tau$ . We consider the mapping  $\mathcal{J}_{\theta, p} : \mathcal{A}_+ \times \mathcal{A}_+ \times \mathcal{A} \rightarrow \mathbb{R}$  given by

$$\mathcal{J}_{\theta, p}(R, S; A) := \langle A, \theta^{-p}(R, S) \# A \rangle = \sum_{k, \ell} \theta^{-p}(\lambda_k, \mu_\ell) \tau[A^* E_k^R A E_\ell^S],$$

where  $\theta : (0, \infty) \times (0, \infty) \rightarrow (0, \infty)$  and  $p \in \mathbb{R}$ , and  $R = \sum_k \lambda_k E_k^R$  and  $S = \sum_\ell \mu_\ell E_\ell^S$  denotes the spectral decomposition. The main cases of interest to us are  $p = \pm 1$ .

In this section we shall assume that the function  $\theta$  is 1-homogeneous, i.e.,  $\theta(\lambda r, \lambda s) = \lambda \theta(r, s)$  for all  $\lambda, r, s > 0$ . Clearly, this assumption is satisfied if and only if there exists a function  $f : (0, \infty) \rightarrow (0, \infty)$  such that  $\theta(r, s) = sf(r/s)$  for all  $r, s > 0$ , in which case we have  $f(r) = \theta(r, 1)$ . To simplify notation, we write  $k(r) = 1/f(r)$ .

**Remark 8.1** (Relation to the relative modular operator) It is instructive to see how the definition of  $\theta(R, S)$  can be formulated in terms of the relative modular operator, if  $\theta$  is 1-homogeneous. Given  $S \in \mathcal{A}_+$ , let  $\mathbf{L}_S$  and  $\mathbf{R}_S$  denote the left- and right-multiplication operators defined by  $\mathbf{L}_S(A) = SA$  and  $\mathbf{R}_S(A) = AS$ . Then the relative modular operator  $\Delta_{R, S} : \mathcal{A} \rightarrow \mathcal{A}$  defined by  $\Delta_{R, S} A = RAS^{-1}$  can be expressed as  $\Delta_{R, S} = \mathbf{L}_R \circ \mathbf{R}_{S^{-1}} = \mathbf{R}_{S^{-1}} \circ \mathbf{L}_R$ . Let  $\{\xi_k\}$  (resp.  $\{\eta_\ell\}$ ) be an orthonormal basis of  $\mathbb{C}^n$  consisting of eigenvectors of  $R$  (resp.  $S$ ), let  $\{\lambda_k\}$  (resp.  $\{\mu_\ell\}$ ) be the corresponding eigenvalues, and set  $E_{k\ell} := |\xi_k\rangle\langle \eta_\ell|$ . It follows

that  $\Delta_{R,S}(E_{k\ell}) = \frac{\lambda_k}{\mu_\ell} E_{k\ell}$ , hence the  $E_{k\ell}$ 's form a complete basis of eigenvectors of  $\Delta_{R,S}$ . Moreover, the  $E_{k\ell}$ 's are orthonormal with respect to the Hilbert–Schmidt inner product  $\langle A, B \rangle_{L^2(\text{Tr})} = \text{Tr}[A^*B]$  on  $\mathbb{M}_n(\mathbb{C})$ . Consequently, the spectral decomposition of  $\Delta_{R,S}$  is given by

$$\Delta_{R,S} = \sum_{k,\ell} \frac{\lambda_k}{\mu_\ell} |E_{k\ell}\rangle\langle E_{k\ell}|,$$

and for functions  $f : (0, \infty) \rightarrow \mathbb{R}$  we find  $f(\Delta_{R,S})(A) = \sum_{k,\ell} f(\lambda_k/\mu_\ell) \langle E_{k\ell}, A \rangle_{L^2(\text{Tr})} E_{k\ell}$ . Note that

$$\langle E_{k\ell}, A \rangle_{L^2(\text{Tr})} E_{k\ell} = \sum_m \langle E_{k\ell} \eta_m, A \eta_m \rangle E_{k\ell} = \langle \xi_k, A \eta_\ell \rangle E_{k\ell} = E_k^R A E_\ell^S,$$

where  $E_k^R = |\xi_k\rangle\langle \xi_k|$  and  $E_\ell^S = |\eta_\ell\rangle\langle \eta_\ell|$ . It follows that

$$f(\Delta_{R,S})(A) = \sum_{k,\ell} f(\lambda_k/\mu_\ell) E_k^R A E_\ell^S,$$

and therefore, since  $f(r/s)s = \theta(r, s)$ ,

$$\theta(R, S) \# A = (R_S \circ f(\Delta_{R,S}))(A).$$

**Example 8.2** Let us recall our main examples of interest. A central role is played by the *tilted logarithmic mean*  $\theta_{1,\beta}$  given by

$$\begin{aligned} \theta_{1,\beta}(r, s) &= \int_0^1 (e^{-\beta/2r})^{1-\alpha} (e^{\beta/2s})^\alpha \, d\alpha = \frac{e^{-\beta/2r} - e^{\beta/2s}}{-\beta + \log r - \log s}, \\ f_{1,\beta}(r) &= \frac{e^{-\beta/2r} - e^{\beta/2}}{-\beta + \log r}, \end{aligned}$$

for  $\beta \in \mathbb{R}$ . More generally, in view of Remark 7.12 we are interested in the class of power difference quotients  $\theta_m$  given by  $f_{m,\beta}(r) = \theta_{m,\beta}(r, 1)$ , where

$$\begin{aligned} \theta_{m,\beta}(r, s) &= \int_0^1 \left( (1-\alpha)(e^{-\beta/2r})^{m-1} + \alpha(e^{\beta/2s})^{m-1} \right)^{\frac{1}{m-1}} \, d\alpha \\ &= \frac{m-1}{m} \frac{(e^{-\beta/2r})^m - (e^{\beta/2s})^m}{(e^{-\beta/2r})^{m-1} - (e^{\beta/2s})^{m-1}}. \end{aligned}$$

Consider the mapping  $\Upsilon_{f,p} : \mathcal{A}_+ \times \mathcal{A}_+ \times \mathcal{A} \rightarrow \mathbb{R}$  given by

$$\Upsilon_{f,p}(R, S; A) := \mathcal{J}_{\theta,p}(R, S; A) = \langle A, \theta^{-p}(R, S) \# A \rangle$$

Our goal is to characterize its convexity and contractivity properties in terms of  $f$  and  $m$ . For this purpose we recall that a function  $f : (0, \infty) \rightarrow (0, \infty)$  is said to be *operator monotone*, whenever  $f(A) \leq f(B)$  for all positive matrices  $A \leq B$  in all dimensions. Each operator monotone function is continuous, non-decreasing and concave. We set  $f(0) := \inf_{t>0} f(t)$ .

The following result has been obtained in [27, Theorem 2.1]. The implication “(2)  $\Rightarrow$  (1)”, as well as the reverse implication for fixed  $p = 1$  had already been proved in [26].

**Theorem 8.3** (Characterization of convexity of  $\Upsilon_{f,p}$ ) *Let  $f : (0, \infty) \rightarrow (0, \infty)$  be a function and let  $p \in \mathbb{R} \setminus \{0\}$ . The following assertions are equivalent.*

- (1) *The function  $\Upsilon_{f,p}$  is jointly convex in its three variables;*

(2) The function  $f$  is operator monotone and  $p \in (0, 1]$ .

Applying this result to the functions  $f = f_{m,\beta}$ , we obtain the following result.

**Corollary 8.4** (Characterization of convexity of  $\Upsilon_{f,p}$  for power difference quotients) *For  $m \in \mathbb{R} \setminus \{0\}$  and  $\beta \in \mathbb{R}$ , let  $f = f_{m,\beta}$  and  $\theta = \theta_{m,\beta}$  be as in Example 8.2. Then, the associated mapping  $\Upsilon_{f,p}$  is jointly convex if and only if  $m \in [-1, 2]$ ,  $p \in (0, 1]$ , and  $\beta \in \mathbb{R}$ . In particular, the mapping*

$$(R, S, A) \mapsto \langle A^*, \theta_{1,\beta}^{-1}(R, S) \# A \rangle = \tau \left[ \int_0^\infty A^* \frac{1}{x + e^{-\beta/2}R} A \frac{1}{x + e^{\beta/2}S} dx \right]$$

is jointly convex for all  $\beta \in \mathbb{R}$ .

**Proof** Since  $f_{m,\beta}(s) = e^{\beta/2} f_{m,0}(e^{-\beta} s)$ , the operator monotonicity of  $f_{m,\beta}$  does not depend on  $\beta$ . It has been proved in [25, Proposition 4.2], that  $f_{m,0}$  is operator monotone if and only if  $m \in [-1, 2]$ . Hence, the first assertion follows from Theorem 8.3. The second assertion is the special case  $m = p = 1$ , noting that

$$\frac{1}{\theta_{1,\beta}(r, s)} = \int_0^\infty \frac{1}{x + e^{-\beta/2}r} \frac{1}{x + e^{\beta/2}s} dx .$$

□

**Remark 8.5** In the case where  $\theta = \theta_{1,\beta}$ , the operator monotonicity of  $f_{1,\beta}$  can be checked elementarily, by writing  $f_{1,\beta}(r) = \int_0^1 e^{-\beta(1/2-\alpha)r^\alpha} d\alpha$ , and applying the Löwner-Heinz Theorem (e.g., [7, Theorem 2.6]), which asserts that the function  $r \mapsto r^\alpha$  is operator monotone for  $\alpha \in [0, 1]$ .

The following result is proved in [26, Theorem 5].

**Theorem 8.6** (Contractivity of  $\Theta_{f,p}$  under CPTC maps) *Suppose that  $f : (0, \infty) \rightarrow (0, \infty)$  is operator monotone. Then, for any  $R, S \in \mathcal{A}_+$  and  $A \in \mathcal{A}$ , and for any completely positive and trace preserving map  $\mathcal{T} : \mathcal{A} \rightarrow \mathcal{A}$ , we have*

$$\Upsilon_{f,1}(\mathcal{T}(R), \mathcal{T}(S); \mathcal{T}(A)) \leq \Upsilon_{f,1}(R, S; A) . \tag{8.1}$$

In the case where  $f = f_{m,\beta}$  as in Example 8.2, we obtain the following result.

**Corollary 8.7** (Contractivity of  $\Theta_{f,p}$  for power difference quotients) *Let  $m \in [-1, 2]$  and  $\beta \in \mathbb{R}$ , and let  $f = f_{m,\beta}$  and  $\theta = \theta_{m,\beta}$  be as in Example 8.2. Then, for any  $R, S \in \mathcal{A}_+$  and  $A \in \mathcal{A}$ , and for any completely positive and trace preserving map  $\mathcal{P} : \mathcal{A} \rightarrow \mathcal{A}$ , (8.1) holds. In particular, for  $m = 1$  we obtain*

$$\begin{aligned} & \tau \left[ \int_0^\infty \mathcal{T}(A)^* \frac{1}{x + e^{-\beta/2}\mathcal{T}(R)} \mathcal{T}(A) \frac{1}{x + e^{\beta/2}\mathcal{T}(S)} dx \right] \\ & \leq \tau \left[ \int_0^\infty A^* \frac{1}{x + e^{-\beta/2}R} A \frac{1}{x + e^{\beta/2}S} dx \right] . \end{aligned}$$

**Proof** This follows from Theorem 8.6, as the operator monotonicity of  $f_{m,p}$  had already been noted in Corollary 8.4. □

### 9 The Riemannian Distance

Fix a differentiable structure  $(\mathcal{A}, \nabla, \sigma)$  in the sense of Definition 4.7 and a collection of functions  $(\theta_j)_j$  satisfying Assumption 7.2. For simplicity we restrict ourselves to the ergodic case, so that  $\mathcal{M}_\rho = \mathfrak{P}_+$  for all  $\rho \in \mathfrak{P}_+$ .

In this section we study basic properties of the Riemannian distance  $\mathcal{W}$  associated to the operators  $(\mathcal{K}_\rho)_\rho$  defined in (7.4). For  $\rho_0, \rho_1 \in \mathfrak{P}_+$  this distance is given by

$$\begin{aligned} \mathcal{W}(\rho_0, \rho_1)^2 &= \inf \left\{ \int_0^1 \langle \mathcal{K}_{\rho_t} A_t, A_t \rangle dt : \partial_t \rho_t = \mathcal{K}_{\rho_t} A_t, \rho_t|_{t=0} = \rho_0, \rho_t|_{t=1} = \rho_1 \right\} \\ &= \inf \left\{ \int_0^1 \tau [(\nabla A_t)^* \widehat{\rho}_t \# \nabla A_t] dt : \right. \\ &\quad \left. \partial_t \rho_t + \operatorname{div}(\widehat{\rho}_t \# \nabla A_t) = 0, \rho_t|_{t=0,1} = \rho_{0,1} \right\}, \end{aligned} \tag{9.1}$$

where the infimum runs over smooth curves  $(\rho_t)_{t \in [0,1]}$  in  $\mathfrak{P}_+$  and  $(A_t)_{t \in [0,1]}$  in  $\mathcal{A}_0$  satisfying the stated conditions.

In the classical theory of optimal transport, it is a useful fact that the following equivalent formulations hold for the 2-Kantorovich distance on  $\mathbb{R}^n$ :

$$\begin{aligned} W_2(\rho_0, \rho_1)^2 &= \inf \left\{ \int_0^1 |\nabla \psi_t(x)|^2 d\rho_t(x) dt : \partial_t \rho_t + \operatorname{div}(\rho_t \nabla \psi_t) = 0, \rho_t|_{t=0,1} = \rho_{0,1} \right\} \\ &= \inf \left\{ \int_0^1 \frac{|P_t(x)|^2}{\rho_t(x)} dx dt : \partial_t \rho_t + \operatorname{div} P_t = 0, \rho_t|_{t=0,1} = \rho_{0,1} \right\}. \end{aligned} \tag{9.2}$$

The latter formulation has the advantage that the minimisation problem is convex, due to the convexity of the function  $(p, r) \mapsto \frac{|p|^2}{r}$  on  $\mathbb{R}^n \times (0, \infty)$ .

Using the convexity results presented in Sect. 8 we will show that an analogous result holds in the non-commutative setting. We use the shorthand notation

$$\langle \mathbf{B}, \mathbf{C} \rangle_\rho = \sum_j \tau_j [B_j^* (\widehat{\rho}_j \# C_j)], \quad \langle \mathbf{B}, \mathbf{C} \rangle_{-1, \rho} = \sum_j \tau_j [B_j^* (\check{\rho}_j \# C_j)],$$

to denote the scalar products that will frequently appear below. The corresponding norms are given by  $\|\mathbf{B}\|_\rho = \sqrt{\langle \mathbf{B}, \mathbf{B} \rangle_\rho}$  and  $\|\mathbf{B}\|_{-1, \rho} = \sqrt{\langle \mathbf{B}, \mathbf{B} \rangle_{-1, \rho}}$ . It will occasionally be convenient to write

$$\mathcal{A}(\rho; \mathbf{B}, \mathbf{C}) = \langle \mathbf{B}, \mathbf{C} \rangle_{-1, \rho} \quad \text{and} \quad \mathcal{A}(\rho, \mathbf{B}) = \|\mathbf{B}\|_{-1, \rho}^2.$$

We start with a non-commutative analogue of (9.2).

**Lemma 9.1** *For  $\rho_0, \rho_1 \in \mathfrak{P}_+$  we have*

$$\mathcal{W}(\rho_0, \rho_1)^2 = \inf \left\{ \int_0^1 \|\mathbf{B}_t\|_{-1, \rho_t}^2 dt : \partial_t \rho_t + \operatorname{div} \mathbf{B}_t = 0, \rho_t|_{t=0,1} = \rho_{0,1} \right\}, \tag{9.3}$$

where the infimum runs over all smooth curves  $(\rho_t)_{t \in [0,1]}$  in  $\mathfrak{P}_+$  and  $(\mathbf{B}_t)_{t \in [0,1]}$  in  $\mathcal{B}$ .

**Proof** Any admissible curve  $(A_t)$  in (9.1) yields an admissible curve  $(\mathbf{B}_t)$  in (9.3) given by  $\mathbf{B}_t = \widehat{\rho}_t \nabla A_t$ , that satisfies  $\|\nabla A_t\|_{\rho_t} = \|\mathbf{B}_t\|_{-1, \rho_t}$ . This implies the inequality “ $\geq$ ” in (9.3).

To prove the reverse inequality, we take an admissible curve  $(\rho_t, \mathbf{B}_t)_t$  in (9.3). We consider the linear space of gradient vector fields  $\mathcal{G} = \{\nabla A : A \in \mathcal{A}_0\}$ , and let  $\mathcal{D}_t \subseteq \mathcal{G}$  denote its

orthogonal complement in  $\mathcal{B}$  with respect to the scalar product product  $\langle \cdot, \cdot \rangle_{\rho_t}$ . Consider the orthogonal decomposition

$$\check{\rho}_j \# \mathbf{B}_t = \nabla A_t + D_t \in \mathcal{G} \oplus \mathcal{D}_t .$$

Since  $\langle \nabla \tilde{A}, D_t \rangle_{\rho_t} = 0$  for all  $\tilde{A} \in \mathcal{A}_0$ , it follows that  $\operatorname{div}(\widehat{\rho} \# D_t) = 0$ . Therefore,  $\partial_t \rho_t + \operatorname{div}(\widehat{\rho}_t \# \nabla A_t) = 0$ . Moreover,

$$\tau [(\nabla A_t)^* \widehat{\rho}_t \# \nabla A_t] = \|\nabla A_t\|_{\rho_t}^2 \leq \|\check{\rho}_j \# \mathbf{B}_t\|_{\rho_t}^2 = \|\mathbf{B}_t\|_{-1, \rho_t}^2 ,$$

which yields the inequality “ $\leq$ ” in (9.3). □

**Proposition 9.2** (Extension of the distance to the boundary) *Suppose that  $\theta_j(a, b) \geq C \min\{a, b\}^p$  for some  $C > 0$  and  $p < 2$ . Then the distance function  $\mathcal{W} : \mathfrak{P}_+ \times \mathfrak{P}_+ \rightarrow \mathbb{R}$  extends continuously to a metric on  $\mathfrak{P}$ .*

**Proof** Let  $\rho_0, \rho_1 \in \mathfrak{P}$  and let  $\{\rho_0^n\}_n, \{\rho_1^n\}_n$  be sequences in  $\mathfrak{P}_+$  satisfying  $\tau [|\rho_i^n - \rho_i|^2] \rightarrow 0$  as  $n \rightarrow \infty$  for  $i = 0, 1$ . We claim that the sequence  $\{\mathcal{W}(\rho_0^n, \rho_1^n)\}_n$  is Cauchy.

To prove this, it suffices to show that  $\mathcal{W}(\rho_i^n, \rho_i^m) \rightarrow 0$  as  $n, m \rightarrow \infty$  for  $i = 0, 1$ , since

$$|\mathcal{W}(\rho_0^n, \rho_1^n) - \mathcal{W}(\rho_0^m, \rho_1^m)| \leq \mathcal{W}(\rho_0^n, \rho_0^m) + \mathcal{W}(\rho_1^n, \rho_1^m) .$$

Fix  $\varepsilon \in (0, 1)$ , and set  $\tilde{\rho} := (1 - \varepsilon)\rho_0 + \varepsilon \mathbf{1}$ . Take  $N \geq 1$  so large that  $\tau [|\rho_0^n - \rho_0|^2] \leq \varepsilon^2$  whenever  $n \geq N$ . For  $n \geq N$  we consider the linear interpolation  $\rho_t^n = (1 - t)\rho_0^n + t\tilde{\rho}$ . Then  $\dot{\rho}_t^n = \tilde{\rho} - \rho_0^n$  for all  $t \in (0, 1)$ . Since  $\mathcal{X}_1$  is invertible on  $\mathcal{A}_0$  by Lemma 7.3 and ergodicity, we may define  $A := \mathcal{X}_1^{-1}(\tilde{\rho} - \rho_0^n) \in \mathcal{A}_0$ , and we have  $\dot{\rho}_t^n = \operatorname{div}(\nabla A)$ . Since  $\rho_t^n \geq t\varepsilon \mathbf{1}$  for  $t \in [0, 1]$ , we have  $\frac{1}{\theta_j} \Big|_{\operatorname{sp}(\rho_t^n)} \leq C(t\varepsilon)^{-p}$ , and thus  $\tau[(\nabla A)^* \check{\rho}_t^n \# \nabla A] \leq C(t\varepsilon)^{-p} \tau[|\nabla A|^2]$  by Lemma 6.6. It follows that

$$\mathcal{W}(\rho_0^n, \tilde{\rho}) \leq \int_0^1 \sqrt{\tau[(\nabla A)^* \check{\rho}_t^n \# \nabla A]} \, dt \leq C\varepsilon^{-p/2} \|\nabla A\|_{L^2(\tau)} ,$$

since  $p < 2$ . Using the boundedness of  $\nabla \circ \mathcal{X}_1^{-1}$  we obtain

$$\begin{aligned} C^{-1} \|\nabla A\|_{L^2(\tau)} &\leq \|\tilde{\rho} - \rho_0^n\|_{L^2(\tau)} \\ &= \|\rho_0 - \rho_0^n\|_{L^2(\tau)} + \varepsilon \|\mathbf{1} - \rho_0\|_{L^2(\tau)} \\ &\leq \varepsilon(1 + \|\mathbf{1} - \rho_0\|_{L^2(\tau)}) . \end{aligned}$$

We infer that  $\mathcal{W}(\rho_0^n, \tilde{\rho}) \leq C\varepsilon^{1-p/2}$  for some  $C < \infty$  depending on  $\rho_0$ . It follows that  $\mathcal{W}(\rho_0^n, \rho_0^m) \leq C\varepsilon^{1-p/2}$  for  $n, m \geq N$ . Since  $p < 2$ , this proves the claim.

We can thus extend  $\mathcal{W}$  to  $\mathfrak{P}$  by setting  $\mathcal{W}(\rho_0, \rho_1) = \lim_{n \rightarrow \infty} \mathcal{W}(\rho_0^n, \rho_1^n)$ . It immediately follows that  $\mathcal{W}$  is symmetric and the triangle inequality extends to  $\mathfrak{P}$ . The fact that  $\mathcal{W}(\rho_0, \rho_1) \neq 0$  whenever  $\rho_0$  and  $\rho_1$  are distinct, follows from Proposition 9.4 below. □

Our next aim is to prove Proposition 9.4 below, which yields a lower bound on the distance  $\mathcal{W}$  in terms of a non-commutative analogue of the 1-Kantorovich metric. To formulate the result, we use the notation

$$\|\mathbf{B}\|_{\mathcal{B}, 2} := \sqrt{\frac{1}{2} \sum_j \|\ell_j^\dagger(B_j B_j^*) + r_j^\dagger(B_j^* B_j)\|_{\mathcal{A}}}$$

for  $\mathbf{B} = (B_j)_{j \in \mathcal{J}} \in \mathcal{B}$ .

**Lemma 9.3** *There exists  $M < \infty$  such that  $\|\mathbf{B}\|_\rho \leq M\|\mathbf{B}\|_{\mathcal{B},2}$  for all  $\rho \in \mathfrak{P}_+$  and  $\mathbf{B} \in \mathcal{B}$ . If  $\theta_j(r, s) \leq \frac{1}{2}(r + s)$  for all  $r, s > 0$ , then this estimate holds with  $M = 1$ :*

$$\|\mathbf{B}\|_\rho \leq \|\mathbf{B}\|_{\mathcal{B},2} .$$

**Proof** Recalling that  $\|\cdot\|_{\mathcal{B}_j}$  denotes the norm on  $\mathcal{B}_j$ , we define

$$M_j := \sup \left\{ \tau_j[|\widehat{\rho}_j \# B|] : \rho \in \mathfrak{P}_+, \|B\|_{\mathcal{B}_j} \leq 1 \right\}, \quad \text{and} \quad \widetilde{M} := \sup_{j \in \mathcal{J}} M_j .$$

Since our setting is finite-dimensional,  $\widetilde{M}$  is finite and all norms on  $\mathcal{B}$  are equivalent. Thus, for a suitable constant  $M < \infty$ , it follows that

$$\|\mathbf{B}\|_\rho^2 = \sum_j \tau_j[B_j^* \widehat{\rho}_j \# B_j] \leq \sum_j \|B_j^*\|_{\mathcal{B}_j} \tau_j[|\widehat{\rho}_j \# B_j|] \leq \widetilde{M} \sum_j \|B_j\|_{\mathcal{B}_j}^2 \leq M\|\mathbf{B}\|_{\mathcal{B},2}^2 ,$$

which proves the first statement.

Suppose now that  $\theta_j(r, s) \leq \frac{1}{2}(r + s)$ . Since  $\rho$  is positive and the operators  $\ell_j$  and  $r_j$  preserve positivity, we obtain using Lemma 6.6,

$$\begin{aligned} \|\mathbf{B}\|_\rho^2 &= \sum_j \tau_j[B_j^* \widehat{\rho}_j \# B_j] \\ &\leq \frac{1}{2} \sum_j \tau_j[\ell_j(\rho) B_j B_j^* + r_j(\rho) B_j^* B_j] \\ &= \frac{1}{2} \sum_j \tau[\rho(\ell_j^\dagger(B_j B_j^*) + r_j^\dagger(B_j^* B_j))] \\ &\leq \frac{1}{2} \sum_j \|\ell_j^\dagger(B_j B_j^*) + r_j^\dagger(B_j^* B_j)\|_{\mathcal{A}} , \end{aligned}$$

which yields the result. □

For  $\rho_0, \rho_1 \in \mathfrak{P}$  we set

$$W_1(\rho_0, \rho_1) := \sup \left\{ \tau[(\rho_1 - \rho_0)A] : A \in \mathcal{A}, \|\nabla A\|_{\mathcal{B},2} \leq 1 \right\} . \tag{9.4}$$

By analogy with the dual Kantorovich formulation of the commutative 1-Kantorovich metric  $W_1$  in terms of Lipschitz functions, this metric can be seen as a non-commutative analogue of  $W_1$ . The following result generalizes a result from [18] from the discrete to the non-commutative setting; see also [46] for non-commutative results of this type.

**Proposition 9.4** *Let  $M$  be as in Lemma 9.3 and set  $N := \sup\{\|\nabla A\|_{\mathcal{B},2} : \|A\|_{\mathcal{A}} \leq 1\}$ . Then, for  $\rho_0, \rho_1 \in \mathfrak{P}$  we have*

$$N^{-1} \tau[|\rho_0 - \rho_1|] \leq W_1(\rho_0, \rho_1) \leq M \mathscr{W}(\rho_0, \rho_1) .$$

**Proof** The first inequality follows from the definitions, since  $\tau[|B|] = \sup_{\|A\|_{\mathcal{A}} \leq 1} \tau[AB]$  for  $B \in \mathcal{A}$ .

Fix  $\varepsilon > 0$ , take  $\bar{\rho}_0, \bar{\rho}_1 \in \mathfrak{P}$ , and let  $(\rho_t, B_t)_t$  be a solution to the continuity equation with approximately optimal action, i.e.,

$$\partial_t \rho_t + \operatorname{div}(\widehat{\rho}_t \# \nabla B_t) = 0 \quad \text{and} \quad \left( \int_0^1 \|\nabla B_t\|_{\rho_t}^2 \, dt \right)^{\frac{1}{2}} \leq \mathscr{W}(\bar{\rho}_0, \bar{\rho}_1) + \varepsilon .$$

For any  $A \in \mathcal{A}_h$  we obtain using Lemma 9.3

$$\begin{aligned} |\tau[A(\bar{\rho}_0 - \bar{\rho}_1)]| &= \left| \int_0^1 \tau[A\dot{\rho}_t] dt \right| \\ &= \left| \int_0^1 \tau[A \operatorname{div}(\widehat{\rho}_t \# \nabla B_t)] dt \right| \\ &= \left| \int_0^1 \langle \nabla A, \nabla B_t \rangle_{\rho_t} dt \right| \\ &\leq \left( \int_0^1 \|\nabla A\|_{\rho_t}^2 dt \right)^{1/2} \left( \int_0^1 \|\nabla B_t\|_{\rho_t}^2 dt \right)^{1/2} \\ &\leq M \|\nabla A\|_{\mathcal{B},2} (\mathcal{W}(\bar{\rho}_0, \bar{\rho}_1) + \varepsilon). \end{aligned}$$

Since  $\varepsilon > 0$  is arbitrary, the result follows by definition of  $W_1$ . □

In the remainder of this section we impose the following natural additional conditions in addition to Assumption 7.2.

**Assumption 9.5** The functions  $\theta_j : [0, \infty) \times [0, \infty) \rightarrow [0, \infty)$  are 1-homogeneous (which implies that  $\theta_j(r, s) = sf_j(r/s)$  for some function  $f_j$ ). The functions  $f_j$  are assumed to be operator monotone.

Under this assumption, we will prove some crucial convexity properties for the action functional and the squared distance.

**Proposition 9.6** (Convexity of the action) *Let  $\rho^i \in \mathfrak{P}$  and  $\mathbf{B}^i \in \mathcal{B}$  for  $i = 0, 1$ . For  $s \in [0, 1]$  set  $\rho^s := (1 - s)\rho^0 + s\rho^1$  and  $\mathbf{B}^s := (1 - s)\mathbf{B}^0 + s\mathbf{B}^1$ . Then we have*

$$\mathcal{A}(\rho^s, \mathbf{B}^s) \leq (1 - s)\mathcal{A}(\rho^0, \mathbf{B}^0) + s\mathcal{A}(\rho^1, \mathbf{B}^1).$$

**Proof** This follows immediately from Theorem 8.3 in view of Assumption 9.5. □

**Theorem 9.7** (Convexity of the squared distance) *For  $i = 0, 1$ , let  $\rho_0^i, \rho_1^i \in \mathfrak{P}$ , and for  $s \in [0, 1]$  set  $\rho_0^s := (1 - s)\rho_0^0 + s\rho_0^1$  and  $\rho_1^s := (1 - s)\rho_1^0 + s\rho_1^1$ . Then:*

$$\mathcal{W}(\rho_0^s, \rho_1^s)^2 \leq (1 - s)\mathcal{W}(\rho_0^0, \rho_1^0)^2 + s\mathcal{W}(\rho_0^1, \rho_1^1)^2.$$

**Proof** Fix  $\varepsilon > 0$ . By continuity, it suffices to prove the inequality for  $\rho_0^i, \rho_1^i \in \mathfrak{P}_+$  and  $i = 0, 1$ . Let  $(\rho_t^i, \mathbf{B}_t^i)_t$  be such that  $\partial_t \rho_t^i + \operatorname{div} \mathbf{B}_t^i = 0$  and  $\int_0^1 \mathcal{A}(\rho_t^i, \mathbf{B}_t^i) dt \leq \mathcal{W}(\rho_0^i, \rho_1^i)^2 + \varepsilon$ . For  $s \in [0, 1]$  we define

$$\rho_t^s := (1 - s)\rho_t^0 + s\rho_t^1 \quad \text{and} \quad \mathbf{B}_t^s := (1 - s)\mathbf{B}_t^0 + s\mathbf{B}_t^1.$$

It follows that  $\partial_t \rho_t^s + \operatorname{div} \mathbf{B}_t^s = 0$ , and by Lemma 9.1 and Proposition 9.6 we obtain

$$\begin{aligned} \mathcal{W}(\rho_0^s, \rho_1^s)^2 &\leq \int_0^1 \mathcal{A}(\rho_t^s, \mathbf{B}_t^s) dt \\ &\leq (1 - s) \int_0^1 \mathcal{A}(\rho_t^0, \mathbf{B}_t^0) dt + s \int_0^1 \mathcal{A}(\rho_t^1, \mathbf{B}_t^1) dt \\ &\leq (1 - s)\mathcal{W}(\rho_0^0, \rho_1^0)^2 + s\mathcal{W}(\rho_0^1, \rho_1^1)^2 + 2\varepsilon. \end{aligned}$$

Since  $\varepsilon > 0$  is arbitrary, the desired inequality follows. □



Using these convexity properties, the existence of constant speed geodesics for the metric  $\mathscr{W}$  follows by standard arguments; cf. [18, Theorem 3.2]) for a proof in the commutative setting and [46] for a proof in a non-commutative context.

**Theorem 9.8** (Existence of  $\mathscr{W}$ -geodesics) *For any  $\bar{\rho}_0, \bar{\rho}_1 \in \mathfrak{P}$  there exists a curve  $\rho : [0, 1] \rightarrow \mathfrak{P}$  satisfying  $\rho_0 = \bar{\rho}_0, \rho_1 = \bar{\rho}_1$ , and  $\mathscr{W}(\rho_s, \rho_t) = |s - t|\mathscr{W}(\rho_0, \rho_1)$  for all  $s, t \in [0, 1]$ .*

### 10 Geodesic Convexity of the Entropy

In this section we will analyse geodesic convexity of the relative entropy functional  $\text{Ent}_\sigma$ . Throughout this section we fix a differential structure  $(\mathcal{A}, \nabla, \sigma)$  and assume that the associated quantum Markov semigroup  $(\mathcal{P}_t)$  is ergodic. We consider the transport metric  $\mathscr{W}$  defined in Theorem 7.7 using the functions  $\theta_j$  given by  $\theta_j(r, s) := \Lambda(e^{\omega_j/2r}, e^{-\omega_j/2s})$ , so that the Kolmogorov forward equation  $\partial_t \rho = \mathcal{L}^\dagger \rho$  is the gradient flow of the relative von Neumann entropy  $\text{Ent}_\sigma$  with respect to the Riemannian metric induced by  $(\mathcal{K}_\rho)_\rho$ .

The following terminology will be useful.

**Definition 10.1** Let  $(\mathcal{X}, d)$  be a metric space. A functional  $\mathcal{F} : \mathcal{X} \rightarrow \mathbb{R} \cup \{+\infty\}$  is said to be

- *weakly geodesically  $\lambda$ -convex* if any pair  $x_0, x_1 \in \mathcal{X}$  can be connected by a geodesic  $(\gamma_t)_{t \in [0,1]}$  in  $(\mathcal{X}, d)$  along which  $\mathcal{F}$  satisfies the  $\lambda$ -convexity inequality

$$\mathcal{F}(\gamma_t) \leq (1 - t)\mathcal{F}(\gamma_0) + t\mathcal{F}(\gamma_1) - \frac{\kappa}{2}t(1 - t)d(x_0, x_1)^2. \tag{10.1}$$

- *strongly geodesically  $\lambda$ -convex* if (10.1) holds for any geodesic  $(\gamma_t)_{t \in [0,1]}$  in  $(\mathcal{X}, d)$ .

The following result, shows in particular that these concepts are equivalent in our setting and provides several equivalent characterizations of geodesic  $\lambda$ -convexity. We shall use the notation

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

We refer to [18] for a version of this result in the discrete setting, and to [46] for the Lindblad setting.

**Theorem 10.2** (Characterizations of geodesic  $\lambda$ -convexity) *Let  $\lambda \in \mathbb{R}$ . For a differential structure  $(\mathcal{A}, \nabla, \sigma)$  the following assertions are equivalent:*

- (1)  $\text{Ent}_\sigma$  is weakly geodesically  $\lambda$ -convex on  $(\mathfrak{P}, \mathscr{W})$ ;
- (2)  $\text{Ent}_\sigma$  is strongly geodesically  $\lambda$ -convex on  $(\mathfrak{P}, \mathscr{W})$ ;
- (3) For all  $\rho, \nu \in \mathfrak{P}$ , the following ‘evolution variational inequality’ holds for all  $t \geq 0$ :

$$\frac{1}{2} \frac{d^+}{dt} \mathscr{W}^2(\mathcal{P}_t^\dagger \rho, \nu) + \frac{\lambda}{2} \mathscr{W}^2(\mathcal{P}_t^\dagger \rho, \nu) \leq \text{Ent}_\sigma(\nu) - \mathcal{H}(\mathcal{P}_t^\dagger \rho); \tag{10.2}$$

- (4) For all  $\rho \in \mathfrak{P}_+$  and  $A \in \mathcal{A}_0$  we have

$$\text{Hess}_{\mathcal{K}} \text{Ent}_\sigma(\rho)[A, A] \geq \lambda \tau[A \mathcal{K}_\rho A].$$

**Proof** “(4)  $\Rightarrow$  (3)” This can be proved by an argument from [14]; see [18, Theorem 4.5] for a proof in a similar setting.

“(3)  $\Rightarrow$  (2)”: This follows from an application of [14, Theorem 3.2] to the metric space  $(\mathfrak{P}, \mathscr{W})$ .

“(2)  $\Rightarrow$  (1)”: Since  $(\mathfrak{P}, \mathscr{W})$  is a geodesic space, this implication is immediate.

“(1)  $\Rightarrow$  (4)”: Obvious. □

In the classical setting, the Ricci curvature on a Riemannian manifold  $\mathcal{M}$  is bounded from below by  $\lambda \in \mathbb{R}$  if and only if the entropy (with respect to the volume measure) is geodesically  $\lambda$ -convex in the space of probability measures  $\mathcal{P}(\mathcal{M})$  endowed with the Kantorovich metric  $W_2$ . This characterisation is the starting point for the synthetic theory of metric measure spaces with lower Ricci curvature bounds, which has been pioneered by Lott, Sturm and Villani.

By analogy, we make the following definition in the non-commutative setting, which extends the corresponding definition in the discrete setting [18].

**Definition 10.3** (*Ricci curvature*) Let  $\lambda \in \mathbb{R}$ . We say that a differential structure  $(\mathcal{A}, \nabla, \sigma)$  has Ricci curvature bounded from below by  $\lambda$  if the equivalent conditions of Theorem 10.2 hold. In this case, we write  $\text{Ric}(\mathcal{A}, \nabla, \sigma) \geq \lambda$ .

It is possible to characterize Ricci curvature in terms of a gradient estimate in the spirit of Bakry–Émery; see [17] for the corresponding statement in the setting of finite Markov chains and [46] for an implementation in the Lindblad setting.

**Theorem 10.4** (*Gradient estimate*) Let  $\lambda \in \mathbb{R}$ . A differential structure  $(\mathcal{A}, \nabla, \sigma)$  satisfies  $\text{Ric}(\mathcal{A}, \nabla, \sigma) \geq \lambda$  if and only if the following gradient estimate holds for all  $\rho \in \mathfrak{P}$ ,  $A \in \mathcal{A}_0$  and  $t \geq 0$ :

$$\|\nabla \mathcal{P}_t A\|_{\rho}^2 \leq e^{-2\lambda t} \|\nabla A\|_{\mathcal{P}_t^{\dagger} \rho}^2. \tag{10.3}$$

**Proof** We follow a standard semigroup interpolation argument. Clearly, (10.3) holds for any  $\rho \in \mathfrak{P}$  if and only if it holds for any  $\rho \in \mathfrak{P}_+$ .

Fix  $t > 0$ ,  $\rho \in \mathfrak{P}_+$  and  $A \in \mathcal{A}_0$ , and define  $f : [0, t] \rightarrow \mathbb{R}$  by

$$f(s) := e^{-2\lambda s} \langle \mathcal{H}_{\mathcal{P}_s^{\dagger} \rho} \mathcal{P}_{t-s} A, \mathcal{P}_{t-s} A \rangle_{L^2(\mathcal{A}, \tau)} = e^{-2\lambda s} \|\nabla \mathcal{P}_{t-s} A\|_{\mathcal{P}_s^{\dagger} \rho}^2.$$

Writing  $\rho_s = \mathcal{P}_s^{\dagger} \rho$  and  $A_s = \mathcal{P}_s A$ , it follows by (7.17) and Proposition 7.16 that

$$\begin{aligned} f'(s) &= e^{-2\lambda s} \tau \left[ (\nabla A_{t-s})^* (\mathcal{N}_{\rho_s, \mathcal{L}^{\dagger} \rho_s}^{(1)} + \mathcal{N}_{\rho_s, \mathcal{L}^{\dagger} \rho_s}^{(2)}) \# \nabla A_{t-s} \right. \\ &\quad \left. - 2(\nabla \mathcal{L} A_{t-s})^* \widehat{\rho}_s \# \nabla A_{t-s} - 2\lambda (\nabla A_{t-s})^* \widehat{\rho}_s \# \nabla A_{t-s} \right] \\ &= 2e^{-2\lambda s} \left( \text{Hess}_{\mathcal{H}} \text{Ent}_{\sigma}(\rho_s)[A_{t-s}, A_{t-s}] - \lambda \tau[A_{t-s} \mathcal{H}_{\rho_s} A_{t-s}] \right). \end{aligned}$$

Assume now that  $\text{Ric}(\mathcal{A}, \nabla, \sigma) \geq \lambda$ . Applying (4) from Theorem 10.2, we obtain  $f'(s) \geq 0$  for all  $s$ . This implies that  $f(t) \geq f(0)$ , which is (10.3).

To prove the converse, set  $g(t) = e^{2\lambda t} \|\nabla \mathcal{P}_t A\|_{\rho}^2$  and  $h(t) = \|\nabla A\|_{\mathcal{P}_t^{\dagger} \rho}^2$ . Then (10.3) implies that  $g(t) \leq h(t)$  for all  $t \geq 0$ . Since  $g(0) = h(0)$ , we infer that  $g'(0) \leq h'(0)$ . Since

$$\begin{aligned} g'(0) &= 2\tau \left[ (\nabla \mathcal{L} A)^* \widehat{\rho} \# \nabla A \right] + 2\lambda \|A\|_{\rho}^2, \\ h'(0) &= \tau \left[ (\nabla A)^* (\mathcal{N}_{\rho, \mathcal{L}^{\dagger} \rho}^{(1)} + \mathcal{N}_{\rho, \mathcal{L}^{\dagger} \rho}^{(2)}) \# \nabla A \right], \end{aligned}$$

we obtain  $\text{Hess}_{\mathcal{K}} \text{Ent}_{\sigma}(\rho)[A, A] \geq \lambda \tau[A \mathcal{K}_{\rho} A]$  in view of the expression for the Hessian in Proposition 7.16. □

An immediate consequence of a Ricci curvature bound is the following contractivity estimate for the associated semigroup, which was independently proved by Rouzé in [44].

**Proposition 10.5** ( $\lambda$ -Contractivity) *If  $\text{Ric}(\mathcal{A}, \nabla, \sigma) \geq \lambda$ , then the  $\lambda$ -contractivity bound*

$$\mathcal{W}(\mathcal{P}_t^{\dagger} \rho_0, \mathcal{P}_t^{\dagger} \rho_1) \leq e^{-\lambda t} \mathcal{W}(\rho_0, \rho_1)$$

holds for all  $\rho_0, \rho_1 \in \mathfrak{P}$  and  $t \geq 0$ .

**Proof** This is a well-known consequence of the evolution variational inequality (10.2); see [14, Proposition 3.1]. □

Using the techniques developed in this paper, we can explicitly compute the Ricci curvature for the depolarizing channel defined in Sect. 5.6. The result has been obtained independently by Rouzé in [44].

**Theorem 10.6** (Ricci bound for the depolarizing channel) *Let  $\gamma > 0$ , and let  $(\mathcal{A}, \nabla, \tau)$  be a differential structure for the generator of the depolarizing channel given by  $\mathcal{L}A = \gamma(\tau[A]\mathbf{1} - A)$ . Then  $\text{Ric}(\mathcal{A}, \nabla, \tau) \geq \gamma$ .*

**Proof** Since  $\mathcal{L}A = \gamma(\tau[A]\mathbf{1} - A)$  and  $\partial_j \mathbf{1} = 0$ , we have  $\partial_j \mathcal{L}A = -\gamma \partial_j A$ , independently of the choice of the operators  $\partial_j$ . We will show that the result follows from this identity.

First we note that

$$-\tau[(\nabla \mathcal{L}A)^* \widehat{\rho} \# \nabla A] = \gamma \tau[(\nabla A)^* \widehat{\rho} \# \nabla A]. \tag{10.4}$$

Moreover, since  $\partial_1 \Lambda(a, b) = \int_0^1 (1-s)a^{-s}b^s ds$  we obtain (using the notation from (7.16)),

$$\begin{aligned} \mathcal{N}_{\rho, \mathcal{L}\rho}^{(1),j} &= \gamma \delta \Lambda \left( \begin{matrix} \rho & & \rho \\ & (1-\rho) & \\ & & \rho \end{matrix} \right) \\ &= \gamma \sum_{k,m,p} \frac{\Lambda(\lambda_k, \lambda_p) - \Lambda(\lambda_m, \lambda_p)}{\lambda_k - \lambda_m} E_k (\mathbf{1} - \rho) E_m \otimes E_p \\ &= \gamma (\mathbf{1} - \rho) \sum_{k,p} \partial_1 \Lambda(\lambda_k, \lambda_p) E_k \otimes E_p \\ &= \gamma (\mathbf{1} - \rho \otimes \mathbf{1}) \partial_1 \Lambda(\rho, \rho) \end{aligned}$$

Similarly, we have  $\mathcal{N}_{\rho, \mathcal{L}\rho}^{(2),j} = \gamma (\mathbf{1} \otimes (\mathbf{1} - \rho)) \partial_2 \Lambda(\rho, \rho)$ . Using the scalar identity  $a \partial_1 \Lambda(a, b) + b \partial_2 \Lambda(a, b) = \Lambda(a, b)$ , it follows that

$$\mathcal{N}_{\rho, \mathcal{L}\rho}^{(1),j} + \mathcal{N}_{\rho, \mathcal{L}\rho}^{(2),j} = \gamma (\partial_1 \Lambda + \partial_2 \Lambda - \Lambda)(\rho, \rho).$$

Moreover, we note that  $\partial_1 \Lambda(a, b) + \partial_2 \Lambda(a, b) \geq 1 \geq \Lambda(a, b)$  for  $a, b \in [0, 1]$  (and hence for  $a, b \in \text{sp}(\rho)$ ). Therefore, for  $\eta = 1, 2$ , we obtain using Lemma 6.6,

$$\begin{aligned} \tau[(\nabla A)^* \mathcal{N}_{\rho, \mathcal{L}^{\dagger} \rho}^{(\eta)} \# (\nabla A)] &= \frac{1}{2} \tau[(\nabla A)^* (\mathcal{N}_{\rho, \mathcal{L}^{\dagger} \rho}^{(1)} + \mathcal{N}_{\rho, \mathcal{L}^{\dagger} \rho}^{(2)}) \# (\nabla A)] \\ &= \frac{\gamma}{2} \tau[(\nabla A)^* (\partial_1 \Lambda + \partial_2 \Lambda - \Lambda)(\rho, \rho) \# (\nabla A)] \tag{10.5} \\ &\geq 0. \end{aligned}$$

Combining (10.4) and (10.5), it follows from (7.19) that

$$\text{Hess}_{\mathcal{K}} \text{Ent}(\rho)[A, A] \geq \gamma \tau[(\nabla A)^* \widehat{\rho} \# \nabla A] = \gamma \langle \mathcal{K}_\rho A, A \rangle_{L^2(\tau)},$$

which proves the result. □

Since the spectral gap of  $\mathcal{L}$  equals  $\gamma$ , it follows from the results in Sect. 11 that the obtained constant is optimal.

### 10.1 Geodesic Convexity Via Intertwining

In this subsection we provide a useful technique for proving Ricci curvature bounds, which has the advantage that it does not require an explicit computation of the Hessian of the entropy. Instead, it relies on the following intertwining property between the gradient and the quantum Markov semigroup.

**Definition 10.7** (*Intertwining property*) For  $\lambda \in \mathbb{R}$ , we say that a collection of linear operators  $(\mathcal{P}_t)_{t \geq 0}$  on  $\mathcal{B}$  is  $\lambda$ -intertwining for the quantum Markov semigroup  $(\mathcal{P}_t)_{t \geq 0}$ , if the following conditions hold:

- (1) For all  $A \in \mathcal{A}$  and  $t \geq 0$ , we have  $\nabla \mathcal{P}_t A = \mathcal{P}_t \nabla A$ ;
- (2) For all  $\rho \in \mathfrak{P}_+$ ,  $\mathbf{B} = (B_j) \in \mathcal{B}$  and  $t \geq 0$ , we have

$$\mathcal{A}(\rho, \mathcal{P}_t^\dagger \mathbf{B}) \leq e^{-2\lambda t} \mathcal{A}(\rho, (\mathcal{P}_t^\dagger B_j)_j). \tag{10.6}$$

By duality, the intertwining relation (1) implies the identity

$$\mathcal{P}_t \text{div}(\mathbf{A}) = \text{div}(\mathcal{P}_t^\dagger \mathbf{A}), \quad \text{for } \mathbf{A} \in \mathcal{B}. \tag{10.7}$$

The following lemma allows us to check the  $\lambda$ -intertwining property in several examples of interest.

**Lemma 10.8** *Let  $\lambda \in \mathbb{R}$ , and suppose that  $\partial_j \mathcal{L} A = (\mathcal{L} - \lambda) \partial_j A$  for all  $A \in \mathcal{A}$ . Then the semigroup  $(\mathcal{P}_t)_t$  defined by  $(\mathcal{P}_t \mathbf{B})_j = e^{-\lambda t} \mathcal{P}_t B_j$  is  $\lambda$ -intertwining for the quantum Markov semigroup  $(\mathcal{P}_t)_{t \geq 0}$ .*

**Proof** By spectral theory, the stated condition on the generator is equivalent to the semigroup property  $\partial_j \mathcal{P}_t A = e^{-\lambda t} \mathcal{P}_t \partial_j A$  for all  $t \geq 0$ . Thus, the semigroup  $(\mathcal{P}_t)_t$  satisfies (1) in Definition 10.7. Since  $(\mathcal{P}_t^\dagger \mathbf{B})_j = e^{-\lambda t} \mathcal{P}_t^\dagger B_j$ , condition (2) follows as well. □

**Theorem 10.9** (Lower Ricci bound via intertwining) *Let  $(\mathcal{A}, \nabla, \sigma)$  be a differential structure, and let  $\lambda \in \mathbb{R}$ . If there exists a collection of linear operators  $(\mathcal{P}_t)_{t \geq 0}$  on  $\mathcal{B}$  that is  $\lambda$ -intertwining for the associated QMS  $(\mathcal{P}_t)_{t \geq 0}$ , then  $\text{Ric}(\mathcal{A}, \nabla, \sigma) \geq \lambda$ .*

**Proof of Theorem 10.9** The proof is a variation on an argument by Dolbeault, Nazaret and Savaré [16].

Fix  $\bar{\rho}, v \in \mathfrak{P}$ , and let  $(\rho_s, \mathbf{B}_s)_{s \in [0,1]}$  be a solution to the continuity equation

$$\partial_s \rho_s + \text{div} \mathbf{B}_s = 0, \quad \rho_0 = v, \quad \rho_1 = \bar{\rho},$$

that minimizes the action functional (9.3). This implies that  $(\rho_s)_s$  is a constant speed geodesic, and

$$\mathcal{A}(\rho_s, \mathbf{B}_s) = \mathcal{W}(v, \bar{\rho})^2 \tag{10.8}$$

for all  $s \in [0, 1]$ . We define  $\rho_s^t := \mathcal{P}_{st}^\dagger \rho_s$ , so that  $\partial_s \rho_s^t = \mathcal{P}_{st}^\dagger (\partial_s \rho_s) - t \mathcal{L}^\dagger \mathcal{P}_{st}^\dagger \rho_s$ . Using this identity, we obtain

$$\begin{aligned} \partial_s \rho_s^t &= \mathcal{P}_{st}^\dagger (\partial_s \rho_s) - t \mathcal{L}^\dagger \mathcal{P}_{st}^\dagger \rho_s \\ &= -\mathcal{P}_{st}^\dagger (\operatorname{div} \mathbf{B}_s) - t \mathcal{L}^\dagger \rho_s^t \\ &= -\operatorname{div}(\mathcal{P}_{st}^\dagger \mathbf{B}_s) - t \mathcal{L}^\dagger \rho_s^t. \end{aligned}$$

Write  $\tilde{\nabla} = (\tilde{\partial}_j)_j$ , where  $\tilde{\partial}_j = e^{-\omega_j/2} V_j r_j(\rho) - e^{\omega_j/2} \ell_j(\rho) V_j$ . It then follows from Lemma 7.8 and Theorem 7.7 that  $\mathcal{L}^\dagger = \operatorname{div} \tilde{\nabla}$ . Hence, we infer that the curve  $(\rho_s^t)_{s \in [0,1]}$  satisfies the continuity equation  $\partial_s \rho_s^t + \operatorname{div} \mathbf{B}_s^t = 0$ , where

$$\mathbf{B}_s^t = \mathcal{P}_{st}^\dagger \mathbf{B}_s - t \tilde{\nabla} \rho_s^t.$$

Using the bilinearity of  $\mathcal{A}(\rho_s^t, \cdot, \cdot)$ , we obtain

$$\begin{aligned} \mathcal{W}(v, \mathcal{P}_t^\dagger \bar{\rho})^2 &\leq \int_0^1 \mathcal{A}(\rho_s^t, \mathbf{B}_s^t) \, ds \\ &= \int_0^1 \mathcal{A}(\rho_s^t, \mathcal{P}_{st}^\dagger \mathbf{B}_s) - 2t \mathcal{A}(\rho_s^t, \mathbf{B}_s^t, \tilde{\nabla} \rho_s^t) - t^2 \mathcal{A}(\rho_s^t, \tilde{\nabla} \rho_s^t) \, ds. \end{aligned} \tag{10.9}$$

Using (10.6) and Corollary 8.7 we infer that

$$\mathcal{A}(\rho_s^t, \mathcal{P}_{st}^\dagger \mathbf{B}_s) \leq e^{-2\lambda st} \mathcal{A}(\mathcal{P}_{st}^\dagger \rho_s, (\mathcal{P}_{st}^\dagger B_{j,s})_j) \leq e^{-2\lambda st} \mathcal{A}(\rho_s, \mathbf{B}_s)$$

hence (10.8) yields

$$\int_0^1 \mathcal{A}(\rho_s^t, \mathcal{P}_{st}^\dagger \mathbf{B}_s) \, ds \leq \frac{1 - e^{-2\lambda t}}{2\lambda t} \mathcal{W}(v, \bar{\rho})^2,$$

A direct computation using Lemma 7.8 shows that

$$\begin{aligned} \partial_s \operatorname{Ent}_\sigma(\rho_s^t) &= \tau[(\log \rho_s^t - \log \sigma) \partial_s \rho_s^t] = -\tau[(\log \rho_s^t - \log \sigma) \operatorname{div} \mathbf{B}_s^t] \\ &= \tau[(\nabla(\log \rho_s^t - \log \sigma))^* \mathbf{B}_s^t] = \tau[(\rho_s^t \# \tilde{\nabla} \rho_s^t)^* \mathbf{B}_s^t] = \mathcal{A}(\rho_s^t, \mathbf{B}_s^t, \tilde{\nabla} \rho_s^t). \end{aligned}$$

Estimating the final term in (10.9) by 0, we infer that

$$\frac{1}{2t} \left( \mathcal{W}(v, \mathcal{P}_t^\dagger \bar{\rho})^2 - \mathcal{W}(v, \bar{\rho})^2 \right) \leq \frac{1}{2t} \left( \frac{1 - e^{-2\lambda t}}{2\lambda t} - 1 \right) \mathcal{W}(v, \bar{\rho})^2 - \int_0^1 \partial_s \operatorname{Ent}_\sigma(\rho_s^t) \, ds.$$

Since  $t \mapsto \operatorname{Ent}_\sigma(\rho_s^t)$  is continuous, we observe that the right-hand side converges as  $t \downarrow 0$ . Letting  $t \downarrow 0$  we infer that

$$\frac{1}{2} \frac{d^+}{dt} \Big|_{t=0} \mathcal{W}(v, \mathcal{P}_t \bar{\rho})^2 \leq \frac{\lambda}{2} \mathcal{W}(v, \bar{\rho})^2 + \operatorname{Ent}_\sigma(\bar{\rho}) - \operatorname{Ent}_\sigma(v),$$

which proves the evolutionary variational inequality from Theorem 10.2 for  $t = 0$ . By the semigroup property, the inequality holds for all  $t \geq 0$ , hence the result follows.  $\square$

**Remark 10.10** As pointed out by an anonymous referee, the condition from Lemma 10.8 is preserved under taking tensor products of quantum Markov semigroups. Therefore, Theorem 10.9 yields a lower Ricci curvature bound for tensor product semigroups of this type. It is an interesting open question whether such a tensorisation property holds for arbitrary quantum Markov semigroups, as is known to be true in the Markov chain setting [18].

We finish the section with the example of the Fermionic Ornstein–Uhlenbeck equation from Sect. 5.5, which was already discussed in [10]. For the convenience of the reader we provide the details.

**Proposition 10.11** (Intertwining for fermions) *In the fermionic setting, we have the commutation relations  $[\partial_j, \mathcal{L}] = -\partial_j$  for  $j = 1, \dots, n$ . Consequently, the intertwining property holds with  $\lambda = 1$ .*

**Proof** We use the well-known fact that the differential operator  $\partial_j$  is the *annihilation operator*: it maps the  $k$ -particle space  $\mathcal{H}^k$  into the  $(k - 1)$ -particle space  $\mathcal{H}^{k-1}$  for any  $0 \leq k \leq n$  (with the convention that  $\mathcal{H}^{-1} = \{0\}$ ). On the other hand,  $-\mathcal{L}$  is the *number operator*, which satisfies  $\mathcal{L}A = -kA$  for all  $A \in \mathcal{H}^k$ . Hence, for  $A \in \mathcal{H}^k$ , we have  $\partial_j \mathcal{L}A = -k\partial_j A$ , whereas  $\mathcal{L}\partial_j A = -(k - 1)\partial_j A$ . This yields the desired commutation relation  $[\partial_j, \mathcal{L}] = -\partial_j$  on  $\mathcal{H}^k$ , which extends to  $\mathfrak{C}^n$  by linearity. The result thus follows from Lemma 10.8.  $\square$

We immediately obtain the following result.

**Corollary 10.12** *The differential structure for the fermionic Ornstein–Uhlenbeck equation in Sect. 5.5 satisfies  $\text{Ric}(\mathfrak{C}^n, \nabla, \tau) \geq 1$  in any dimension  $n \geq 1$ .*

It follows from the results in the following section that the constant 1 is optimal.

### 11 Functional Inequalities

One of the advantages of the framework of this paper is that it allows one to prove a sequence of implications between several useful functional inequalities. Throughout this section we assume that  $(\mathcal{P}_t)_t$  is ergodic.

Recall that

$$\text{Ent}_\sigma(\rho) := \text{Tr}[\rho(\log \rho - \log \sigma)], \quad \mathcal{I}_\sigma(\rho) := -\text{Tr}[(\log \rho - \log \sigma)\mathcal{L}^\dagger \rho],$$

and note that  $\frac{d}{dt} \text{Ent}_\sigma(\mathcal{P}_t^\dagger \rho) = -\mathcal{I}_\sigma(\mathcal{P}_t^\dagger \rho)$  for  $\rho \in \mathfrak{P}_+$ . The quantity  $\mathcal{I}_\sigma$  is a quantum version of the Fisher information (or entropy production) relative to  $\sigma$ ; we refer to [42] for an introduction to several notions of Fisher information in the quantum setting.

The gradient flow structure from Theorem 7.7 implies that  $\mathcal{L}^\dagger \rho = \text{div}(\widehat{\rho} \# \nabla(\log \rho - \log \sigma))$ , which yields  $\mathcal{I}_\sigma(\rho) = \|\nabla(\log \rho - \log \sigma)\|_\rho^2$ . Recall that for  $\rho \in \mathfrak{P}$  and  $A \in \mathcal{A}$  we denote the associated Bogolioubov–Kubo–Mori scalar product and norm by

$$\langle A, B \rangle_{L^2_{\text{BKM}}(\rho)} = \int_0^1 \tau [A^* \rho^{1-s} B \rho^s] ds, \quad \|A\|_{L^2_{\text{BKM}}(\rho)} = \sqrt{\langle A, A \rangle_{\text{BKM}}}$$

The results presented in this section have been obtained in the classical discrete setting of finite Markov chains in [18], and in the setting of Lindblad operators in [46]. Here we state and prove the results in the more general framework that includes arbitrary differential structures  $(\mathcal{A}, \nabla, \sigma)$ . The proofs closely follow the original arguments by Otto and Villani [39], which were adapted in [18,46]. In our finite-dimensional setting, most of the results follow directly from Riemannian considerations, though some additional care is needed due to the degeneracy of the metric at the boundary  $\mathfrak{P} \setminus \mathfrak{P}_+$ .

**Definition 11.1** A differential structure  $(\mathcal{A}, \nabla, \sigma)$  satisfies

(1) a *modified logarithmic Sobolev inequality* with constant  $\lambda > 0$  if for all  $\rho \in \mathcal{P}(\mathcal{X})$ ,

$$\text{Ent}_\sigma(\rho) \leq \frac{1}{2\lambda} \mathcal{I}_\sigma(\rho). \tag{MLSI(\lambda)}$$

(2) an *H $\mathcal{W}$ I inequality* with constant  $\kappa \in \mathbb{R}$  if for all  $\rho \in \mathcal{P}(\mathcal{X})$ ,

$$\text{Ent}_\sigma(\rho) \leq \mathcal{W}(\rho, \sigma) \sqrt{\mathcal{I}_\sigma(\rho)} - \frac{\kappa}{2} \mathcal{W}(\rho, \sigma)^2. \tag{H\mathcal{W}I(\kappa)}$$

(3) a *modified Talagrand inequality* with constant  $\lambda > 0$  if for all  $\rho \in \mathfrak{P}$ ,

$$\mathcal{W}(\rho, \sigma) \leq \sqrt{\frac{2}{\lambda} \text{Ent}_\sigma(\rho)}. \tag{T_{\mathcal{W}}(\lambda)}$$

(4) a *T<sub>1</sub>-transport inequality* with constant  $\lambda > 0$  if for all  $\rho \in \mathfrak{P}$ ,

$$W_1(\rho, \sigma) \leq \sqrt{\frac{2}{\lambda} \text{Ent}_\sigma(\rho)}. \tag{T_1(\lambda)}$$

(5) a *Poincaré inequality* (or *spectral gap inequality*) with constant  $\lambda > 0$  if for all  $A \in \mathcal{A}_h$  with  $\tau[\int_0^1 \sigma^{1-s} A \sigma^s ds] = 0$ ,

$$\|A\|_{L^2_{\text{BKM}}(\sigma)}^2 \leq \frac{1}{\lambda} \|\nabla A\|_\sigma^2. \tag{P(\lambda)}$$

It is well known and an easy consequence of Gronwall’s inequality, that **MLSI(\lambda)** is equivalent to the exponential decay of the entropy with rate  $2\lambda$ :

$$\text{Ent}_\sigma(\mathcal{P}_t^\dagger \rho) \leq e^{-2\lambda t} \text{Ent}_\sigma(\rho). \tag{11.1}$$

There are other approaches to some of these inequalities and variants of them; see, e.g., [3,4,9,30,41].

Recall that for an absolutely continuous curve  $(\rho_t)_t \in (\mathfrak{P}, \mathcal{W})$ , its *metric derivative*

$$|\rho'_t| := \lim_{h \rightarrow 0} \frac{\mathcal{W}(\rho_{t+h}, \rho_t)}{|h|}$$

exists for a.e.  $t \in [0, T]$ ; see [2, Theorem 1.1.2].

**Proposition 11.2** *Let  $\rho, \nu \in \mathfrak{P}_+$ . For all  $t \geq 0$  we have*

$$\frac{d^+}{dt} \mathcal{W}(\mathcal{P}_t^\dagger \rho, \nu) \leq \sqrt{\mathcal{I}_\sigma(\mathcal{P}_t^\dagger \rho)}. \tag{11.2}$$

*In particular, the metric derivative of the heat flow with respect to  $\mathcal{W}$  satisfies  $|(\mathcal{P}_t^\dagger \rho)'| \leq \sqrt{\mathcal{I}_\sigma(\mathcal{P}_t^\dagger \rho)}$ .*

**Proof** Set  $\rho_t := \mathcal{P}_t^\dagger \rho$ . Using the triangle inequality for  $\mathcal{W}$  we obtain

$$\begin{aligned} \frac{d^+}{dt} \mathcal{W}(\rho_t, \nu) &= \limsup_{s \downarrow 0} \frac{1}{s} (\mathcal{W}(\rho_{t+s}, \nu) - \mathcal{W}(\rho_t, \nu)) \\ &\leq \limsup_{s \downarrow 0} \frac{1}{s} \mathcal{W}(\rho_t, \rho_{t+s}). \end{aligned}$$

In view of the gradient flow identity  $\partial_t \rho = \operatorname{div}(\widehat{\rho} \# \nabla(\log \rho - \log \sigma))$ , the definition of  $\mathscr{W}$  yields

$$\begin{aligned} \limsup_{s \downarrow 0} \frac{1}{s} \mathscr{W}(\rho_t, \rho_{t+s}) &\leq \limsup_{s \downarrow 0} \frac{1}{s} \int_t^{t+s} \|\nabla(\log \rho_r - \log \sigma)\|_{\rho_r} \, dr \\ &= \limsup_{s \downarrow 0} \frac{1}{s} \int_t^{t+s} \sqrt{\mathcal{I}_\sigma(\rho_r)} \, dr \\ &= \sqrt{\mathcal{I}_\sigma(\rho_t)}. \end{aligned}$$

The last equality follows from the continuity of  $r \mapsto \sqrt{\mathcal{I}_\sigma(\rho_r)}$ . □

The following result is a non-commutative analogue of a well-known result by Otto and Villani [39].

**Theorem 11.3** *Assume that  $\operatorname{Ric}(\mathcal{A}, \nabla, \sigma) \geq \kappa$  for some  $\kappa \in \mathbb{R}$ . Then  $\operatorname{H}\mathscr{W}\operatorname{I}(\kappa)$  holds as well.*

**Proof** Fix  $\rho \in \mathfrak{P}$ . If  $\mathcal{I}_\sigma(\rho) = +\infty$  there is nothing to prove, so we will assume without loss of generality that  $\rho \in \mathfrak{P}_+$ . Set  $\rho_t := \mathcal{P}_t^+ \rho$ . From Theorem 10.2 and the lower bound on the Ricci curvature we know that the curve  $(\rho_t)$  satisfies  $\operatorname{EVI}(\kappa)$ , i.e., equation (10.2). Choosing  $\nu = \sigma$  and  $t = 0$  in the  $\operatorname{EVI}(\kappa)$  yields

$$\operatorname{Ent}_\sigma(\rho) \leq -\frac{1}{2} \frac{d^+}{dt} \Big|_{t=0} \mathscr{W}(\rho_t, \sigma)^2 - \frac{\kappa}{2} \mathscr{W}(\rho, \sigma)^2.$$

It remains to show that

$$-\frac{1}{2} \frac{d^+}{dt} \Big|_{t=0} \mathscr{W}(\rho_t, \sigma)^2 \leq \mathscr{W}(\rho, \sigma) \sqrt{\mathcal{I}_\sigma(\rho)}.$$

To see this, we use the triangle inequality to estimate

$$\begin{aligned} -\frac{1}{2} \frac{d^+}{dt} \Big|_{t=0} \mathscr{W}(\rho_t, \sigma)^2 &= \liminf_{t \downarrow 0} \frac{1}{2t} (\mathscr{W}(\rho, \sigma)^2 - \mathscr{W}(\rho_t, \sigma)^2) \\ &\leq \limsup_{t \downarrow 0} \frac{1}{2t} (\mathscr{W}(\rho, \rho_t)^2 + 2\mathscr{W}(\rho, \rho_t) \cdot \mathscr{W}(\rho, \sigma)), \end{aligned}$$

Using Proposition 11.2 with  $\nu = \rho$  and  $t = 0$  we see that the second term on the right-hand side is bounded by  $\mathscr{W}(\rho, \sigma) \sqrt{\mathcal{I}_\sigma(\rho)}$ , while the first term vanishes. □

The following result is now a simple consequence.

**Theorem 11.4** (Quantum Bakry–Émery Theorem) *Suppose that  $\operatorname{Ric}(\mathcal{A}, \sigma, \nabla) \geq \lambda$  for some  $\lambda > 0$ . Then the modified logarithmic Sobolev inequality  $\operatorname{MLSI}(\lambda)$  holds.*

**Proof** Take  $\rho \in \mathfrak{P}_+$ . It follows from Theorem 11.3 that  $(\mathcal{A}, \sigma, \nabla)$  satisfies  $\operatorname{H}\mathscr{W}\operatorname{I}(\lambda)$ . Using this inequality followed by Young’s inequality we obtain

$$\operatorname{Ent}_\sigma(\rho) \leq \mathscr{W}(\rho, \sigma) \sqrt{\mathcal{I}_\sigma(\rho)} - \frac{\lambda}{2} \mathscr{W}(\rho, \sigma)^2 \leq \frac{1}{2\lambda} \mathcal{I}_\sigma(\rho),$$

which is  $\operatorname{MLSI}(\lambda)$ . □



**Theorem 11.5** (Quantum Otto–Villani Theorem) *Suppose that the differential structure  $(\mathcal{A}, \nabla, \sigma)$  satisfies  $\text{MLSI}(\lambda)$  for some  $\lambda > 0$ . Then the Talagrand inequality  $T_{\mathcal{W}}(\lambda)$  holds as well.*

**Proof** It suffices to prove  $T_{\mathcal{W}}(\lambda)$  for  $\rho \in \mathfrak{P}_+$ , since the inequality for general  $\rho \in \mathfrak{P}$  can then be obtained by approximation.

Fix  $\rho \in \mathfrak{P}_+$  and set  $\rho_t = \mathcal{P}_t^\dagger \rho$ . As  $t \rightarrow \infty$ , we use (11.1) to infer that

$$\text{Ent}_\sigma(\rho_t) \rightarrow 0 \quad \text{and} \quad \mathcal{W}(\rho, \rho_t) \rightarrow \mathcal{W}(\rho, \sigma). \tag{11.3}$$

Define  $F : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  by

$$F(t) := \mathcal{W}(\rho, \rho_t) + \sqrt{\frac{2}{\lambda} \text{Ent}_\sigma(\rho_t)}.$$

We have  $F(0) = \sqrt{\frac{2}{\lambda} \text{Ent}_\sigma(\rho)}$  and  $F(t) \rightarrow \mathcal{W}(\rho, \sigma)$  as  $t \rightarrow \infty$  by (11.3). Hence it is sufficient to show that  $\frac{d^+}{dt} F(t) \leq 0$  for all  $t \geq 0$ . If  $\rho_t \neq \sigma$ , we use Proposition 11.2 and the identity  $\frac{d}{dt} \text{Ent}_\sigma(\rho_t) = -\mathcal{I}_\sigma(\rho_t)$  to obtain

$$\frac{d^+}{dt} F(t) \leq \sqrt{\mathcal{I}_\sigma(\rho_t)} - \frac{\mathcal{I}_\sigma(\rho_t)}{\sqrt{2\lambda \text{Ent}_\sigma(\rho_t)}} \leq 0,$$

where the last inequality follows from  $\text{MLSI}(\lambda)$ . If  $\rho_t = \sigma$ , then the same inequality holds, since this implies that  $\rho_r = \sigma$  for all  $r \geq t$ . □

It is known that the modified logarithmic Sobolev inequality implies a Poincaré inequality by a linearization argument. The following result shows that Poincaré inequality is in fact implied by the Talagrand inequality, which is weaker than the  $\text{MLSI}$  in view of the previous theorem. The  $\text{BKM}$  metric in the left-hand side of  $\text{P}(\lambda)$  appears since it also appears in the second order expansion of the relative entropy of  $\text{Ent}_\sigma(\rho)$  around  $\rho = \sigma$ ; see (6.12).

**Proposition 11.6** *Assume that the triple  $(\mathcal{A}, \sigma, \nabla)$  satisfies  $T_{\mathcal{W}}(\lambda)$  for some  $\lambda > 0$ . Then the Poincaré inequality  $\text{P}(\lambda)$  and the  $T_1$ -transport inequality  $T_1(\lambda)$  hold as well. Moreover,  $\text{Ric}(\mathcal{A}, \sigma, \nabla) \geq \lambda$  implies  $\text{P}(\lambda)$ .*

**Proof** The fact that  $T_{\mathcal{W}}(\lambda)$  implies the  $T_1$ -inequality is an immediate consequence of Proposition 9.4.

Suppose that  $T_{\mathcal{W}}(\lambda)$  holds and let us show  $\text{P}(\lambda)$ . Fix  $v \in \mathcal{A}_0$  and set  $\rho^\varepsilon := \sigma + \varepsilon v$ . Then  $\rho^\varepsilon \in \mathfrak{P}_+$  for sufficiently small  $\varepsilon > 0$ . For such  $\varepsilon > 0$ , let  $(\rho_t^\varepsilon, \mathbf{B}_t^\varepsilon)_t$  be an action minimizing curve connecting  $\rho_0^\varepsilon = \rho^\varepsilon$  and  $\rho_1^\varepsilon = \sigma$ . Thus we have  $\partial_t \rho_t^\varepsilon + \text{div}(\widehat{\rho}_t^\varepsilon \# \mathbf{B}_t^\varepsilon) = 0$  and  $\int_0^1 \tau [(\mathbf{B}_t^\varepsilon)^* \widehat{\rho}_t^\varepsilon \# \mathbf{B}_t^\varepsilon] dt = \mathcal{W}(\rho^\varepsilon, \sigma)^2$ .

Write  $A = \int_0^\infty (x + \sigma)^{-1} v (x + \sigma)^{-1} dx$  so that  $v = \int_0^1 \sigma^{1-s} A \sigma^s ds$ . Using the continuity equation we obtain

$$\|A\|_{L_{\text{BKM}}^2(\sigma)}^2 = \frac{1}{\varepsilon} \tau[A^*(\rho^\varepsilon - \sigma)] = \frac{1}{\varepsilon} \tau[A^* \text{div}(\widehat{\rho}_t^\varepsilon \# \mathbf{B}_t^\varepsilon)] = -\frac{1}{\varepsilon} \int_0^1 \tau [(\nabla A)^* \widehat{\rho}_t^\varepsilon \# \mathbf{B}_t^\varepsilon] dt .$$

The Cauchy-Schwarz inequality yields

$$\begin{aligned} \|A\|_{L^2_{\text{BKM}}(\sigma)}^2 &\leq \frac{1}{\varepsilon} \left( \int_0^1 \|\nabla A\|_{\rho_t^\varepsilon}^2 dt \right)^{1/2} \left( \int_0^1 \|\mathbf{B}_t^\varepsilon\|_{\rho_t^\varepsilon}^2 dt \right)^{1/2} \\ &= \frac{1}{\varepsilon} \left( \int_0^1 \|\nabla A\|_{\rho_t^\varepsilon}^2 dt \right)^{1/2} \mathcal{W}(\rho^\varepsilon, \sigma), \end{aligned}$$

since  $(\rho_t^\varepsilon)_t$  is a  $\mathcal{W}$ -geodesic. Using  $T_{\mathcal{W}}(\lambda)$  we obtain

$$\limsup_{\varepsilon \rightarrow 0} \frac{\mathcal{W}(\rho^\varepsilon, \sigma)}{\varepsilon} \leq \limsup_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon} \sqrt{\frac{2}{\lambda} \text{Ent}_\sigma(\rho^\varepsilon)} \leq \frac{1}{\sqrt{\lambda}} \|A\|_{L^2_{\text{BKM}}(\sigma)},$$

since  $\text{Ent}_\sigma(\rho^\varepsilon) = \frac{1}{2} \varepsilon^2 \|A\|_{L^2_{\text{BKM}}(\sigma)}^2 + o(\varepsilon^2)$  by (6.11) and (6.12). It remains to show that, as  $\varepsilon \rightarrow 0$ ,

$$\int_0^1 \|\nabla A\|_{\rho_t^\varepsilon}^2 dt \rightarrow \|\nabla A\|_\sigma^2.$$

To see this, note that  $\tau[|\rho^\varepsilon - \sigma|] \rightarrow 0$ , hence  $\mathcal{W}(\rho^\varepsilon, \sigma) \rightarrow 0$ . Since  $\mathcal{W}(\rho_t^\varepsilon, \sigma) = (1 - t)\mathcal{W}(\rho^\varepsilon, \sigma)$ , it follows that  $\mathcal{W}(\rho_t^\varepsilon, \sigma) \rightarrow 0$  as  $\varepsilon \rightarrow 0$  for all  $t \in [0, 1]$ , which implies that  $\|\nabla A\|_{\rho_t^\varepsilon}^2 \rightarrow \|\nabla A\|_\sigma^2$  for all  $t \in [0, 1]$ . The result now follows using dominated convergence, since  $\|\nabla A\|_{\rho_t^\varepsilon}^2 \leq \|\nabla A\|_B$  by Lemma 9.3.

The final assertion of the proposition follows by combining this result with Theorem 11.4 and Theorem 11.5. □

**Acknowledgements** Open access funding provided by Institute of Science and Technology (IST Austria). Eric A. Carlen gratefully acknowledges support through NSF grant DMS-174625. Jan Maas gratefully acknowledges support by the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (Grant agreement No 716117), and by the Austrian Science Fund (FWF), Project SFB F65. We are grateful to the anonymous referees for carefully reading the original manuscript and making useful comments.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

1. Alicki, R.: On the detailed balance condition for non-Hamiltonian systems. *Rep. Math. Phys.* **10**(2), 249–258 (1976)
2. Ambrosio, L., Gigli, N., Savaré, G.: *Gradient Flows in Metric Spaces and in the Space of Probability Measures*. Lectures in Mathematics ETH Zürich, 2nd edn. Birkhäuser Verlag, Basel (2008)
3. Bardet, I.: Estimating the decoherence time using non-commutative functional inequalities. [arXiv:1710.01039](https://arxiv.org/abs/1710.01039) (2017)
4. Bardet, I., Rouzé, C.: Hypercontractivity and logarithmic Sobolev inequality for non-primitive quantum Markov semigroups and estimation of decoherence rates. [arXiv:1803.05379](https://arxiv.org/abs/1803.05379) (2018)
5. Birman, M., Solomyak, M.: Double operator integrals in a Hilbert space. *Integr. Equ. Oper. Theory* **47**(2), 131–168 (2003)
6. Brenier, Y., Vorotnikov, D.: On optimal transport of matrix-valued measures. [arXiv:1808.05064](https://arxiv.org/abs/1808.05064) (2018)
7. Carlen, E.: Trace inequalities and quantum entropy: an introductory course. In: *Entropy and the Quantum*. *Contemp. Math.*, vol. 529, pp. 73–140. Amer. Math. Soc., Providence, RI (2010)

8. Carlen, E., 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)
9. Carlen, E.A., Lieb, E.H.: Optimal hypercontractivity for Fermi fields and related noncommutative integration inequalities. *Commun. Math. Phys.* **155**(1), 27–46 (1993)
10. 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)
11. Chen, Y., Gangbo, W., Georgiou, T., Tannenbaum, A.: On the matrix Monge–Kantorovich problem. *Eur. J. Appl. Math.* (2017). <https://doi.org/10.1017/S0956792519000172>
12. Chen, Y., Georgiou, T., Tannenbaum, A.: Matrix optimal mass transport: a quantum mechanical approach. *IEEE Trans. Autom. Control* **63**(8), 2612–2619 (2018)
13. Cipriani, F., Sauvageot, J.-L.: Derivations as square roots of Dirichlet forms. *J. Funct. Anal.* **201**(1), 78–120 (2003)
14. Daneri, S., Savaré, G.: Eulerian calculus for the displacement convexity in the Wasserstein distance. *SIAM J. Math. Anal.* **40**(3), 1104–1122 (2008)
15. Dietert, H.: Characterisation of gradient flows on finite state Markov chains. *Electron. Commun. Probab.* **20**(29), 8 (2015)
16. Dolbeault, J., Nazaret, B., Savaré, G.: A new class of transport distances between measures. *Calc. Var. Partial Differ. Equ.* **34**(2), 193–231 (2009)
17. Erbar, M., Fathi, M.: Poincaré, modified logarithmic Sobolev and isoperimetric inequalities for Markov chains with non-negative Ricci curvature. *J. Funct. Anal.* **274**(11), 3056–3089 (2018)
18. Erbar, M., Maas, J.: Ricci curvature of finite Markov chains via convexity of the entropy. *Arch. Ration. Mech. Anal.* **206**(3), 997–1038 (2012)
19. Erbar, M., Maas, J.: Gradient flow structures for discrete porous medium equations. *Discrete Contin. Dyn. Syst.* **34**(4), 1355–1374 (2014)
20. Fagnola, F., Umanità, V.: Generators of detailed balance quantum Markov semigroups. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* **10**(3), 335–363 (2007)
21. Fathi, M., Maas, J.: Entropic Ricci curvature bounds for discrete interacting systems. *Ann. Appl. Probab.* **26**(3), 1774–1806 (2016)
22. Gao, L., Junge, M., LaRacuente, N.: Fisher information and logarithmic Sobolev inequality for matrix valued functions. [arXiv:1807.08838](https://arxiv.org/abs/1807.08838) (2018)
23. Golse, F., Mouhot, C., Paul, T.: On the mean field and classical limits of quantum mechanics. *Commun. Math. Phys.* **343**(1), 165–205 (2016)
24. Gorini, V., Kossakowski, A., Sudarshan, E.C.G.: Completely positive dynamical semigroups of  $N$ -level systems. *J. Math. Phys.* **17**(5), 821–825 (1976)
25. Hiai, F., Kosaki, H.: Means for matrices and comparison of their norms. *Indiana Univ. Math. J.* **48**(3), 899–936 (1999)
26. Hiai, F., Petz, D.: From quasi-entropy to various quantum information quantities. *Publ. Res. Inst. Math. Sci.* **48**(3), 525–542 (2012)
27. Hiai, F., Petz, D.: Convexity of quasi-entropy type functions: Lieb’s and Ando’s convexity theorems revisited. *J. Math. Phys.* **54**(6), 062201, 21 (2013)
28. Hornshaw, D.F.:  $L^2$ -Wasserstein distances of tracial  $W^*$ -algebras and their disintegration problem. [arXiv:1806.01073](https://arxiv.org/abs/1806.01073) (2018)
29. Jordan, R., Kinderlehrer, D., Otto, F.: The variational formulation of the Fokker–Planck equation. *SIAM J. Math. Anal.* **29**(1), 1–17 (1998)
30. Kastoryano, Michael J., Temme, Kristan: Quantum logarithmic Sobolev inequalities and rapid mixing. *J. Math. Phys.* **54**(5), 052202, 30 (2013)
31. Lindblad, G.: On the generators of quantum dynamical semigroups. *Commun. Math. Phys.* **48**(2), 119–130 (1976)
32. Maas, J.: Gradient flows of the entropy for finite Markov chains. *J. Funct. Anal.* **261**(8), 2250–2292 (2011)
33. Mielke, A.: A gradient structure for reaction-diffusion systems and for energy-drift-diffusion systems. *Nonlinearity* **24**(4), 1329–1346 (2011)
34. Mielke, A.: Dissipative quantum mechanics using GENERIC. In: *Recent Trends in Dynamical Systems*. Springer Proc. Math. Stat., vol. 35, pp. 555–585. Springer, Basel (2013)
35. Mielke, A.: Geodesic convexity of the relative entropy in reversible Markov chains. *Calc. Var. Partial Differ. Equ.* **48**(1–2), 1–31 (2013)
36. Mittenzweig, M., Mielke, A.: An entropic gradient structure for Lindblad equations and couplings of quantum systems to macroscopic models. *J. Stat. Phys.* **167**(2), 205–233 (2017)
37. Moreau, J.-J.: Décomposition orthogonale d’un espace hilbertien selon deux cônes mutuellement polaires. *C. R. Acad. Sci. Paris* **255**, 238–240 (1962)

38. Otto, F.: The geometry of dissipative evolution equations: the porous medium equation. *Commun. Partial Differ. Equ.* **26**(1–2), 101–174 (2001)
39. 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)
40. Ouhabaz, E.-M.: Invariance of closed convex sets and domination criteria for semigroups. *Potential Anal.* **5**(6), 611–625 (1996)
41. Pérez-García, D., Wolf, M., Petz, D., Ruskai, M.B.: Contractivity of positive and trace-preserving maps under  $L_p$  norms. *J. Math. Phys.* **47**(8), 083506, 5 (2006)
42. Petz, D., Ghinea, C.: Introduction to quantum Fisher information. In: *Quantum Probability and Related Topics. QP–PQ: Quantum Probab. White Noise Anal.*, vol. 27, pp. 261–281. World Sci. Publ., Hackensack, NJ (2011)
43. Potapov, D., Sukochev, F.: Operator–Lipschitz functions in Schatten–von Neumann classes. *Acta Math.* **207**(2), 375–389 (2011)
44. Rouzé, C.: Functional inequalities in quantum information theory. Ph.D. thesis, Cambridge University
45. Rouzé, C., Datta, N.: Concentration of quantum states from quantum functional and transportation cost inequalities. *J. Math. Phys.* **60**, 012202 (2019). <https://doi.org/10.1063/1.5023210>
46. Rouzé, C., Datta, N.: Relating relative entropy, optimal transport and Fisher information: a quantum HWI inequality. [arXiv:1709.07437](https://arxiv.org/abs/1709.07437) (2017)
47. Spohn, H., Lebowitz, J.L.: Stationary non-equilibrium states of infinite harmonic systems. *Commun. Math. Phys.* **54**(2), 97–120 (1977)
48. Wirth, M.: A noncommutative transport metric and symmetric quantum Markov semigroups as gradient flows of the entropy. [arXiv:1808.05419](https://arxiv.org/abs/1808.05419) (2018)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.