# Chapter 10
# Wasserstein Geometry of Quantum States and Optimal Transport of Matrix-Valued Measures

**Yongxin Chen, Tryphon T. Georgiou and Allen Tannenbaum**

**Abstract** We overview recent results on generalizations of the Wasserstein 2-metric, originally defined on the space of scalar probability densities, to the space of Hermitian matrices and of matrix-valued distributions, as well as some extensions of the theory to vector-valued distributions and discrete spaces (weighted graphs).

## 10.1 Introduction

Optimal mass transport (OMT) is currently a very active area of research with applications to areas both applied and theoretical including control, transportation, econometrics, fluid dynamics, probability theory, statistical physics, shape optimization, expert systems, and meteorology; see [25, 30] for extensive lists of references. The original problem was first formulated by the civil engineer Gaspar Monge in 1781, and concerned finding the optimal way, in the sense of minimal transportation cost, of moving a pile of soil from one site to another. Much later the problem was extensively analyzed by Kantorovich [18], and is now known as the Monge–Kantorovich (MK) or optimal mass transport (OMT) problem.

Y. Chen
Department of Electrical and Computer Engineering, Iowa State University,
Ames, IA 50011, USA
e-mail: yongchen@iastate.edu

T. T. Georgiou (✉)
Department of Mechanical Engineering, University of California, Irvine, CA, USA
e-mail: tryphon@uci.edu

A. Tannenbaum
Departments Computer Science and Applied Mathematics and Statistics,
Stony Brook University, New York City, USA
e-mail: allen.tannenbaum@stonybrook.edu

In this paper, we present certain generalizations of OMT to matrix and vector-valued transportation. Our original motivation for this rather nontraditional viewpoint was provided by problems in Signal Analysis, more specifically, the need of a weakly continuous metric to compare (matrix-valued) power spectra of multivariate time series (see [24]). Soon afterward it became apparent Quantum Mechanics was another field that would stand to benefit from such an unusual extension of OMT. In fact, it was this latter subject that provided some of the clues of how to properly set up noncommutative OMT.

The basis of the new theory is a suitable extension of the ***Liouville (continuity) equation*** that allows flows in matrix or other spaces. To this end, in [8], we first proposed such a continuity equation and a noncommutative counterpart of OMT where probability distributions are replaced by density matrices (i.e., Hermitian positive-definite matrices with unit trace). The appropriate ***Wasserstein metric*** now corresponds to the minimal value of an action integral evaluated on flows connecting end-point density matrices. The key insight, to use such a dynamic formulation in seeking the needed generality was provided by the seminal approach of Benamou and Brenier [3]. Indeed, the Benamou–Brenier formulation recasts OMT as a stochastic control problem. The work we are reporting herein takes this idea along several different directions, and in particular to OMT between matrices, matricial distribution and vectorial distributions [8, 10]. Extensions of these results to distributions that have end-point distributions of unequal overall mass (unbalanced) are reported in [11] (and not included in the current survey).

We note that at about the same time as [8] was originally reported, closely related approaches were formulated independently and simultaneously in [6, 20]. In fact, in our work, we greatly benefited from earlier work by Carlen and Maas [7] on a fermionic Fokker–Planck equation.

## 10.2   Quantum Continuity Equation

The three papers [6, 8, 20] all begin with the *Lindblad equation* that describes the evolution of open quantum systems. Open quantum systems are thought of as coupled to a larger system (referred to as the environment or the ancilla) and, thereby, cannot, in general, be described by the Schrödinger equation [17]. In this case, the evolution of density operators $\rho$ [17] is given by the Lindblad equation

$$\dot{\rho} = -i[H, \rho] \tag{10.1}$$
$$+ \sum_{k=1}^{N} (L_k \rho L_k^* - \frac{1}{2} \rho L_k^* L_k - \frac{1}{2} L_k^* L_k \rho),$$

where $^*$ denotes conjugate transpose, and throughout, we assume that $\hbar = 1$. The first term on the right-hand side describes the evolution of the state under the effect of the Hamiltonian $H$ (Schrödinger unitary evolution). The other terms on the right-hand

side model diffusion and, thereby, capture the dissipation of energy–they constitute the quantum analogue of Laplace's operator $\Delta$ and are referred to as the Lindblad terms.

Our approach [8, 9] relies on a suitable continuity equation in the space of Hermitian matrices $\mathscr{H}$ (of a given dimension). To this end, we invoke suitable definitions for the gradient $\nabla_L$ and divergence $\nabla_L^*$ operators on spaces of matrices that are explained below and express the continuity equation in the familiar form

$$\dot{\rho} = \nabla_L^* J, \tag{10.2}$$

where $J$ is a *matricial flux*, in complete analogy with the continuity equation on scalar densities.

Throughout, $\rho(t) \in \mathscr{H}$ is a positive-semidefinite matrix of trace one, i.e., a *density matrix* of quantum mechanics. Regarding notation, we let $\mathscr{H}_+$ and $\mathscr{H}_{++}$ denote the cones of nonnegative and positive-definite matrices, respectively,

$$\mathscr{D}_+ := \{\rho \in \mathscr{H}_{++} \mid \text{tr}(\rho) = 1\}$$

the space of density matrices, and $\mathscr{S}$ the space of skew-Hermitian matrices (of the same dimension as $\mathscr{H}$). The flux $J$ is taken in $\mathscr{S}^N$, i.e., a vector with matrix entries. Flux typically arises in the form

$$J = \rho \circ v \ \text{ or in the form } \ J = \nabla_L \rho.$$

The symbol $\rho \circ v$ denotes one of several possible choices of *noncommutative multiplication*. We have considered specifically the following two choices, referred to as the *anticommutator* multiplication (i) and *Kubo-Mori* product (ii), respectively:

$$\text{(i)} \ \ \rho \circ v = \frac{1}{2}(\rho v + v\rho) \text{ and (ii)} \ \ \rho \circ v = \int_0^1 \rho^s v \rho^{1-s} ds,$$

where, for $\rho \in \mathscr{H}$ and $v \in \mathscr{S}^N$,

$$v\rho := \begin{bmatrix} v_1 \rho \\ \vdots \\ v_N \rho \end{bmatrix}, \text{ and } \rho v := \begin{bmatrix} \rho v_1 \\ \vdots \\ \rho v_N \end{bmatrix}.$$

On the other hand, we define the gradient operator with respect to $L \in \mathscr{H}^N$ to be

$$\nabla_L : \mathscr{H} \to \mathscr{S}^N, \ \ X \mapsto \begin{bmatrix} L_1 X - X L_1 \\ \vdots \\ L_N X - X L_N \end{bmatrix}.$$

With respect to the standard Hilbert–Schmidt inner product $\langle X, Y \rangle = \text{tr}(X^*Y)$ (and, for the case when $X, Y$ are in $\mathscr{H}^N$ or $\mathscr{S}^N$, the inner product $\langle X, Y \rangle = \sum_{k=1}^{N} \text{tr}(X_k^* Y_k)$), the divergence operator turns out to be

$$\nabla_L^* : \mathscr{S}^N \to \mathscr{H}, \ Y = \begin{bmatrix} Y_1 \\ \vdots \\ Y_N \end{bmatrix} \mapsto \sum_{k}^{N} L_k Y_k - Y_k L_k,$$

and this is what is used in (10.2). We note that for technical reasons, the definition of gradient and divergence require that $L_k = L_k^*$, i.e., $L \in \mathscr{H}^N$, as above. Also, one can easily verify that $\nabla_L$ is a derivation, in that,[1]

$$\nabla_L(XY + YX) = (\nabla_L X)Y + X(\nabla_L Y)$$
$$+ (\nabla_L Y)X + Y(\nabla_L X), \quad \forall X, Y \in \mathscr{H}.$$

With these definitions in place, we define the (matricial) *Laplacian* as

$$\Delta_L X := -\nabla_L^* \nabla_L X$$
$$= \sum_{k=1}^{N} (2 L_k X L_k - X L_k L_k - L_k L_k X), \ X \in \mathscr{H},$$

which is exactly (after scaling by $1/2$) the diffusion term in the Lindblad equation[2] (10.1). Hence, Lindblad's equation can be rewritten as

$$\dot{\rho} = -\nabla_H^*(\rho i) + \frac{1}{2} \nabla_L^*(\nabla_L \rho)$$
$$= -\nabla_H(\rho i) + \frac{1}{2} \Delta_L \rho.$$

## 10.3 Matricial Wasserstein 2-Metric

From here on we consider the *continuity equation*,

$$\dot{\rho} = \nabla_L^*(\rho \circ v), \tag{10.3}$$

without the diffusion term, but for a general velocity field $v \in \mathscr{S}^N$. A tacit assumption throughout is that the identity matrix $I$ spans the null space or $\nabla_L$; this can be ensured

---

[1] The domain of $\nabla_L$ is $\mathscr{H}$, hence the identity requires $XY + YX$, instead of simply $XY$.

[2] The Lindblad term is in the so-called *symmetric form* since the coefficients are Hermitian.

if one chooses $L_1, \ldots, L_N$ to form a basis of $\mathscr{H}$ (which is a sufficient, but not a necessary condition).

A Wasserstein distance between two density matrices can now be defined as the *least action* (minimum control problem) to steer one density matrix to another,

$$W_{2,a}(\rho_0, \rho_1)^2 := \min_{\rho, v} \int_0^1 \langle v, \rho \circ v \rangle dt, \tag{10.4a}$$

$$\dot{\rho} = \frac{1}{2} \nabla_L^*(\rho \circ v), \tag{10.4b}$$

$$\rho(0) = \rho_0, \quad \rho(1) = \rho_1. \tag{10.4c}$$

In this, $\rho_0$ and $\rho_1$ in $\mathscr{D}_+$ and the optimization is over $\rho(\cdot) \in \mathscr{D}_+$ and $v \in \mathscr{S}^N$. In fact, for $v \in \mathscr{S}^N$, (10.3) already preserves positive definiteness and trace of $\rho(\cdot)$.

The choice of the anticommutator product $\rho \circ v = \frac{1}{2}(\rho v + v\rho)$ is especially appealing since, in this case, the matricial Wasserstein metric (10.4) is readily computable. Indeed, (10.4) can be cast as a convex optimization problem in a manner analogous to that in the scalar case [3]. To see this, let $u := \rho v = [u_1^*, \ldots, u_N^*]^*$ and $u_* := [u_1, \ldots, u_N]^*$, and observe that

$$\mathrm{tr}(\rho v^* v) = \sum_{k=1}^N \mathrm{tr}(\rho v_k^* v_k)$$

$$= \sum_{k=1}^N \mathrm{tr}((\rho v_k)^* \rho^{-1} \rho v_k) = \mathrm{tr}(u^* \rho^{-1} u).$$

Thus, (10.4) can be equivalently expressed as

$$W_2(\rho_0, \rho_1)^2 = \min_{\rho, u} \int_0^1 \mathrm{tr}(u^* \rho^{-1} u) dt, \tag{10.5a}$$

$$\dot{\rho} = \frac{1}{2} \nabla_L^*(u - u_*), \tag{10.5b}$$

$$\rho(0) = \rho_0, \quad \rho(1) = \rho_1. \tag{10.5c}$$

In this, it turns out that although we do not require any structural constraint on $u$, the optimal $u$ satisfies $u = \rho v$ for some $v \in \mathscr{S}^N$.

The choice of the Kubo-Mori product, on the other hand, provides a matricial version of the Wasserstein metric for which the gradient flow of the von Neuman entropy $\mathrm{tr}(\rho \log(\rho))$ is the Lindblad equation [6, 8, 20]. Thus, it generalizes to the noncommutative setting, the famous result by Jordan, Kinderlehrer, and Otto [16] for the ordinary Wasserstein-2 metric on probability densities where the heat equation is the gradient flow of the entropy. However, it is interesting to note that computation of the Wasserstein metric for the Kubo-Mori product appears challenging as compared to the one based on the anticommutator product above.

To characterize the form of minimizer one can proceed to consider the dual problem, which for the case of the anticommutator product goes as follows. With $\lambda(\cdot) \in \mathcal{H}$ a smooth Lagrangian multiplier for the constraints we construct the Lagrangian

$$\mathcal{L}(\rho, v, \lambda) = \int_0^1 \left\{ \frac{1}{2} \text{tr}(\rho v^* v) - \text{tr}(\lambda(\dot{\rho} - \frac{1}{2} \nabla_L^*(\rho v + v\rho))) \right\} dt$$

$$= \int_0^1 \left\{ \frac{1}{2} \text{tr}(\rho v^* v) + \frac{1}{2} \text{tr}((\nabla_L \lambda)^*(\rho v + v\rho)) + \text{tr}(\dot{\lambda}\rho) \right\} dt - \text{tr}(\lambda(1)\rho_1) + \text{tr}(\lambda(0)\rho_0).$$

Point-wise minimization over $v$ yields

$$v_{opt}(t) = -\nabla_L \lambda(t)$$

and the expression for the corresponding minimum

$$\int_0^1 \left\{ -\frac{1}{2} \text{tr}(\rho(\nabla_L \lambda)^*(\nabla_L \lambda)) + \text{tr}(\dot{\lambda}\rho) \right\} dt$$

$$-\text{tr}(\lambda(1)\rho_1) + \text{tr}(\lambda(0)\rho_0),$$

from which we conclude the following sufficient condition for optimality: Suppose there exists $\lambda(\cdot) \in \mathcal{H}$ satisfying

$$\dot{\lambda} = \frac{1}{2}(\nabla_L \lambda)^*(\nabla_L \lambda) = \frac{1}{2} \sum_{k=1}^N (\nabla_L \lambda)_k^*(\nabla_L \lambda)_k \tag{10.6a}$$

such that the solution of

$$\dot{\rho} = -\frac{1}{2} \nabla_L^*(\rho \nabla_L \lambda + \nabla_L \lambda \rho) \tag{10.6b}$$

matches the marginals $\rho(0) = \rho_0$, $\rho(1) = \rho_1$. Then the pair $(\rho, v)$ with $v = -\nabla_L \lambda$ solves (10.4).

The Wasserstein metric induces a Riemannian structure

$$\langle \delta_1, \delta_2 \rangle_\rho = \frac{1}{2} \text{tr}(\rho \nabla \lambda_1^* \nabla \lambda_2 + \rho \nabla \lambda_2^* \nabla \lambda_1)$$

on the tangent space of Hermitian matrices with a specified trace,

$$T_\rho = \{\delta \in \mathcal{H} \mid \text{tr}(\delta) = 0\}.$$

Here $\lambda_j$, $j = 1, 2$ is the solution to the *Poisson equation*

$$\delta_j = -\frac{1}{2}\nabla_L^*(\rho\nabla_L\lambda_j + \nabla_L\lambda_j\rho). \tag{10.7}$$

The proof of existence and uniqueness of the solution of (10.7) follows exactly along the same lines as in [7]; details are given in [9]. In fact, given a *tangent direction* $\delta$, $-\nabla_L\lambda$ is the unique minimizer of $\mathrm{tr}(\rho v^*v)$ over all $v \in \mathscr{S}^N$ satisfying

$$\delta = \frac{1}{2}\nabla_L^*(\rho v + v\rho).$$

With the above definition of inner product, $W_2(\cdot, \cdot)$ indeed defines a metric on $\mathscr{D}_+$ for which $\mathscr{D}_+$ is a geodesic space, i.e., the distance between two given $\rho_0, \rho_1 \in \mathscr{D}_+$ can be rewritten as

$$W_{2,a}(\rho_0, \rho_1) = \min_{\rho} \int_0^1 \sqrt{\langle\dot{\rho}(t), \dot{\rho}(t)\rangle_{\rho(t)}}dt,$$

where the minimum is taken over all piecewise smooth paths on the manifold $\mathscr{D}_+$.

We finally note that, more generally, OMT can be formulated on the space of matrix-valued distributions. In this case, the mass constraint becomes $\int \mathrm{tr}\rho(x)dx = 1$, where $x$ represents a vector of spatial coordinates and $dx$ the volume element. Transport along spatial coordinates, e.g., with $x \in \mathbb{R}^m$, is effected by a term $\nabla_x \cdot (\rho \circ w)$ in the continuity equation, with $w \in \mathscr{H}^m$, i.e.,

$$\dot{\rho} = \nabla_L^*(\rho \circ v) - \nabla_x \cdot (\rho \circ w).$$

Likewise, the cost of transport is duly penalized in a corresponding problem to minimize a suitable action integral; see [8] for details.

## 10.4  Vector-Valued Transport on $\mathbb{R}^N$

A vector-valued density $\rho = [\rho_1, \rho_2, \cdots, \rho_\ell]^T$, on $\mathbb{R}^N$ or on a discrete space, may represent power reflected off a surface at different frequencies/colors. The "mass" of these components may transfer between different entries of the density vector (e.g., due to different angles of reflection) along time flows of the vectorial density. Thus, while the total power may be invariant (under some lighting conditions), the proportion of power at different frequencies or polarization may smoothly vary with viewing angle. As another example consider the case where the entries of $\rho$ represent densities of different species, or particles, and allow for the possibility that mass transfers from one species to another ("mutate"), i.e., between entries of $\rho$. Thus, in

general, we postulate that transport of vector-valued quantities captures flow across space as well as between entries of the density vector.

In this context, an OMT-inspired geometry is aimed to express a suitable continuity and to quantify transport cost for such vectorial distributions. We highlight some of the key elements in [10] for such a theory. It follows a line which is analogous to development of quantum transport that was discussed above.

We begin by considering a vector-valued density $\rho$ on $\mathbb{R}^N$, i.e., a map from $\mathbb{R}^N$ to $\mathbb{R}_+^\ell$ such that

$$\sum_{i=1}^{\ell} \int_{\mathbb{R}^N} \rho_i(x) dx = 1,$$

and consider the entries of $\rho$ as representing density or mass of species/particles that can mutate between one another while maintaining total mass. We denote the set of all vector-valued densities and its interior by $\mathscr{D}$ and $\mathscr{D}_+$, respectively. The dynamics are described by the following ***continuity equation***:

$$\frac{\partial \rho_i}{\partial t} + \nabla_x \cdot (\rho_i v_i) - \sum_{j \neq i} (\rho_j w_{ji} - \rho_i w_{ij}) = 0, \quad \forall i = 1, \ldots, \ell. \tag{10.8}$$

Here $v_i$ is the velocity field of particles $i$ and $w_{ij} \geq 0$ is the transfer rate from $i$ to $j$. Equation (10.8) allows for the possibility to mutate between each pair of entries. More generally, mass transfer may only be permissible between specific types of particles and can be modeled by a graph $\mathscr{G} = (\mathscr{V}, \mathscr{E})$ (where the entries denote nodes and edges, respectively), in which case, for a subset of indices, the transfer rates $w_{ji}$ may be restricted to be zero.

Given $\mu, \nu \in \mathscr{D}_+$, we formulate the optimal mass transport

$$W_2(\mu, \nu)^2 := \inf_{\rho, v, w} \int_0^1 \int_{\mathbb{R}^N} \left\{ \sum_{i=1}^{\ell} \rho_i(t, x) \|v_i(t, x)\|^2 + \gamma \sum_{i,j=1}^{\ell} \rho_i w_{ij}^2(t, x) \right\} dx dt \tag{10.9}$$

$$\frac{\partial \rho_i}{\partial t} + \nabla_x \cdot (\rho_i v_i) - \sum_{j \neq i} (\rho_j w_{ji} - \rho_i w_{ij}) = 0, \quad \forall i = 1, \ldots, \ell$$

$$w_{ij}(t, x) \geq 0, \quad \forall i, j, t, x$$

$$\rho(0, \cdot) = \mu(\cdot), \ \rho(1, \cdot) = \nu(\cdot).$$

The coefficient $\gamma > 0$ specifies the relative cost between transporting mass in space and trading mass between different types of particles. When $\gamma$ is large, the solution reduces to independent OMT problems for the different entries to the degree possible. In general, $W_2$ is a quasi-metric in that it satisfies the triangle inequality and positivity, but may not be symmetric.

Setting $p_{ij} = \rho_i w_{ij} \geq 0$ and $u_i = \rho_i v_i$, we have $\rho_i w_{ij}^2 = \rho_i^{-1} p_{ij}^2$, and $\rho_i \|v_i\|^2 = \rho_i^{-1} \|u_i\|^2$. It follows that

$$W_2(\mu, \nu)^2 = \inf_{\rho, u, p} \int_0^1 \int_{\mathbb{R}^N} \left\{ \sum_{i=1}^{\ell} \rho_i(t, x)^{-1} \|u_i(t, x)\|^2 + \gamma \sum_{i,j=1}^{\ell} \rho_i^{-1} p_{ij}^2(t, x) \right\} dx dt$$

$$\frac{\partial \rho_i}{\partial t} + \nabla_x \cdot u_i - \sum_{j \neq i} (p_{ji} - p_{ij}) = 0, \quad \forall i = 1, \ldots, \ell$$

$$p_{ij}(t, x) \geq 0, \quad \forall i, j, t, x$$

$$\rho(0, \cdot) = \mu(\cdot), \quad \rho(1, \cdot) = \nu(\cdot) \text{ which is a convex problem.}$$

Finally, a Riemannian-like metric on $\mathscr{D}_+$ can be obtained by symmetrizing the above expression [10]. This is,

$$W_{2,\text{sym}}(\mu, \nu)^2 = \inf_{\rho, u, p} \int_0^1 \int_{\mathbb{R}^N} \left\{ \sum_{i=1}^{\ell} \rho_i(t, x)^{-1} \|u_i(t, x)\|^2 \right. \tag{10.10}$$

$$\left. + \frac{\gamma}{2} \sum_{i,j=1}^{\ell} (\rho_i^{-1} + \rho_j^{-1}) p_{ij}^2(t, x) \right\} dx dt$$

under the same constraints. This vector-valued OMT structure is further explored and developed in [10].

## 10.5   Vector-Valued Transport on Graphs

We conclude by highlighting elements of an OMT theory solely on graphs, cast in the setting of vector-valued densities [10]. As explained earlier, such densities may represent the distribution of multiple species/resources that are allowed to mutate between each other as they transition from node to node. The theory is aimed to capture cost of transport in such a setting.

A vector-valued mass distribution on the graph $\mathscr{G} = (\mathscr{V}, \mathscr{E})$ (with $n$ nodes and $m$ edges) is a $\ell$-tuple $\rho = (\rho_1, \cdots, \rho_\ell)$ with each $\rho_i = (\rho_{i,1}, \cdots, \rho_{i,n})^T$ being a vector in $\mathbb{R}_+^n$ such that

$$\sum_{i=1}^{\ell} \sum_{k=1}^{n} \rho_{i,k} = 1.$$

That is, each entry $\rho_i$, for $i \in \{1, \ldots, \ell\}$, is a vector with nonnegative $n$-entries representing, e.g., color intensity for the $i$-th color, at the node corresponding to the respective entry. We denote the set of all nonnegative vector-valued mass distributions

with $\mathscr{D}$ and its interior with $\mathscr{D}_+$. Equation (10.8) is now replaced by the following continuity equation:

$$\dot{\rho}_i - \nabla_{\mathscr{G}}^* \left( (D_2^T \rho_i) \circ v_i - (D_1^T \rho_i) \circ \bar{v}_i \right) - \sum_{j \neq i} (\rho_j \circ w_{ji} - \rho_i \circ w_{ij}) = 0, \quad \forall i = 1, \ldots, \ell,$$

(10.11)

since the spatial domain is now also discrete (i.e., it is $\mathscr{G}$ instead of $\mathbb{R}^N$). Here, $D = D_1 - D_2$ is the incident matrix of the graph, with $D_1, D_2$ are matrices with positive entries reflecting the position of sources $(D_1)$ and sinks $(D_2)$ by a entry equal to 1 in the corresponding place. Thus, the vector $D_1^T \rho$ represents density at the sources of an edge and, likewise, $D_2^T \rho$ represents density at the sinks. Then, also, $\nabla_{\mathscr{G}}$ represents differencing between neighboring nodes and $\nabla_{\mathscr{G}}^*$ represents its dual (i.e., negative divergence) [10]. Finally, $\circ$ represents entry-wise multiplication (Shur) between two vectors.

Now following the Benamou–Brenier [3] philosophy once again, given two marginal densities $\mu, \nu \in \mathscr{D}_+$, we define their Wasserstein distance as

$$W_2(\mu, \nu)^2 := \inf_{\rho, v, w} \int_0^1 \left\{ \sum_{i=1}^{\ell} [v_i^T ((D_2^T \rho) \circ v_i) + \bar{v}_i^T ((D_1^T \rho) \circ \bar{v}_i)] + \gamma \sum_{i,j=1}^{\ell} \sum_{k=1}^{n} \rho_{i,k} w_{ij,k}^2 \right\} dt$$

subject to (10.11) as well as $w_{ij} \geq 0$, $v_i \geq 0$, $\bar{v}_i \geq 0$ for all (or a subset) of $(i, j)$'s and $\rho(0) = \mu$, $\rho(1) = \nu$.

We should note that the problem of transporting vector-valued mass on a graph is quite simpler than in the case where the underlined space is continuous, since it reduces essentially to a scalar mass situation on a suitably larger graph. Indeed, we can view the vector-valued mass as a scalar mass distribution on $\ell$ identical layers of the graph $\mathscr{G}$ where the same nodes at different layers are connected through a complete graph. The two velocity fields $v$, $w$ represent mass transfer within the same layer and between different layers, respectively.

Once again, the computation of the metric has a ***convex*** formulation by changing optimization variable from $(\rho, v_i, \bar{v}_i, w_i j, k)$ to momenta "mass" $\rho$ and momenta $u_i = \rho_i v_i$ and $p_{ij} = \rho_i w_{ij}$, instead.

## 10.6   Conclusion

The basic fluid dynamical formulation of OMT can be generalized to flows on the space of matrices or vectors, that belong to a simplex of a suitable positive cone. A Wasserstein metric in these spaces can then be defined as a minimal quadratic cost for transferring between two end points. Such metrics appear natural as, in particular, for the space of quantum density matrices, render the Lindblad equation as the gradient flow of the von Neumann entropy. Our interest stems from problems in signal analysis and, more specifically, spectral and image analysis. In both of these

application areas, the relevance of weakly continuous metric that can be used to quantify distances between, e.g., matrix-valued power spectra or multicolor images, is self-evident. In particular, geodesics in such spaces naturally model flows and allow morphing between spectra and images, respectively.

## References

1. Ambrosio, L.: Euro Summer School Mathematical Aspects of Evolving Interfaces. Lecture Notes on Optimal Transport Theory. CIME Series of Springer Lecture Notes. Madeira, Portugal, Springer-Verlag, New York (2000)
2. Angenent, S., Haker, S., Tannenbaum, A.: Minimizing flows for the Monge-Kantorovich problem. SIAM J. Math. Anal. **35**, 61–97 (2003)
3. Benamou, J.-D., Brenier, Y.: A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. Numerische Mathematik **84**, 375–393 (2000)
4. Benamou, J.-D.: Numerical resolution of an unbalanced mass transport problem. ESAIM. Math. Model. Numer. Anal. **37**(5), 851–868 (2010)
5. Carlier, G., Salomon, J.: A monotonic algorithm for the optimal control of the Fokker-Planck equation. In: IEEE Conference on Decision and Control (2008)
6. Carlen, E., Maas, J.: Gradient flow and entropy inequalities for quantum Markov semigroups with detailed balance (2016). https://arxiv.org/abs/1609.01254
7. 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**, 887–926 (2014)
8. Chen, Y., Georgiou, T.T., Tannenbaum, A.: Matrix optimal mass transport: a quantum mechanical approach (2016). https://arxiv.org/abs/1610.03041
9. Chen, Y., Gangbo, W., Georgiou, T.T., Tannenbaum, A.: On the matrix Monge-Kantorovich problem. https://arxiv.org/abs/1701.02826
10. Chen, Y., Georgiou, T.T., Tannenbaum, A.: Transport distance on graphs and vector-valued optimal mass transport (2016). https://arxiv.org/pdf/1611.09946v1.pdf
11. Chen, Y., Georgiou, T.T., Tannenbaum, A.: Interpolation of density matrices and matrix-valued measures: the unbalanced case (2017). https://arxiv.org/abs/1612.05914
12. Chen, Y., Georgiou, T.T., Pavon, M.: On the relation between optimal transport and Schrödinger bridges: a stochastic control viewpoint. J. Optim. Theory Appl. **169**(2), 671–691 (2016)
13. Chen, Y., Georgiou, T.T., Pavon, M., Tannenbaum, A.: Robust transport over networks. IEEE Trans. Autom. Control (2016). https://doi.org/10.1109/TAC.2016.2626796
14. Chen, Y., Georgiou, T.T., Pavon, M.: Entropic and displacement interpolation: a computational approach using the Hilbert metric. SIAM J. Appl. Math. **76**(6), 2375–2396 (2016)
15. Evans, L.C.: Partial Differential Equations and Monge-Kantorovich Mass Transfer, in Current Developments in Mathematics, pp. 65–126. International Press, Boston, MA (1999)
16. Jordan, R., Kinderlehrer, D., Otto, F.: The variational formulation of the Fokker-Planck equation. SIAM J. Math. Anal. **29**, 1–17 (1998)
17. Gustafson, S., Sigal, I.M.: Mathematical Concepts of Quantum Mechanics. Springer, New York (2011)
18. Kantorovich, L.V.: On a problem of Monge. Uspekhi Mat. Nauk. **3**, 225–226 (1948)
19. Kumar, A., Tannenbaum, A., Balas, G.: Optical flow: a curve evolution approach. IEEE Trans. Image Process. **5**, 598–611 (1996)
20. Mittnenzweig, M., Mielke, A.: An entropic gradient structure for Lindblad equations and GENERIC for quantum systems coupled to macroscopic models (2016). https://arxiv.org/abs/1609.05765
21. Léonard, C.: From the Schrödinger problem to the Monge-Kantorovich problem. J. Funct. Anal. **262**, 1879–1920 (2012)

22. Maas, J.: Gradient flows of the entropy for finite Markov chains. J. Funct. Anal. **261**(8), 2250–2292 (2011)
23. McCann, R.: Existence and uniqueness of monotone measure-preserving maps. Duke Math. J. **80**, 309–323 (1995)
24. Ning, L., Georgiou, T., Tannenbaum, A.: On matrix-valued Monge-Kantorovich optimal mass transport. IEEE Trans. Autom. Control **60**(2), 373–382 (2015)
25. Rachev, S., Rüschendorf, L.: Mass Transportation Problems, vol. I. Springer-Verlag, New York (1998) (Probab. Appl.)
26. Sandhu, R., Georgiou, T., Reznik, E., Zhu, L., Kolesov, I., Senbabaoglu, Y., Tannenbaum, A.: Graph curvature for differentiating cancer networks. Sci. Rep. (Nat.), **5**, 12323 (2015). https://doi.org/10.1038/srep12323
27. Sandhu, R., Georgiou, T., Tannenbaum, A.: Ricci curvature: an economic indicator for market fragility and systemic risk. Sci. Adv. **2** (2016). https://doi.org/10.1126/sciadv.1501495
28. Tannenbaum, E., Georgiou, T., Tannenbaum, A.: Optimal mass transport for problems in control, statistical estimation, and image processing. In: Dym, H., de Oliveira, M.C., Putinar, M. (eds.) Mathematical Methods in Systems, Optimization, and Control. Birkhauser, Basel (2012)
29. Tannenbaum, E., Georgiou, T., Tannenbaum, A.: Signals and control aspects of optimal mass transport and the Boltzmann entropy. In: 49th IEEE Conference on Decision and Control, pp. 1885–1890 (2010)
30. Villani, C.: Topics in Optimal Transportation, Graduate Studies in Mathematics, vol. 58. AMS, Providence, RI (2003)
31. Villani, C.: Trend to equilibrum for dissipative equations, functional inequalities and mass transportation. In: de Carvalho, M., Rodrigues, J-F. (eds.) Contemporary Mathematics: Recent Advances in the Theory and Applications of Mass Transport. American Mathematical Society Publications (2004)
32. Villani, C.: Optimal Transport, Old and New. Springer, New York (2008)
33. Wang, C., Jonckheere, E., Banirazi, R.: Wireless network capacity versus Ollivier-Ricci curvature under Heat Diffusion (HD) protocol. In: Proceedings of ACC (2013)
34. Yamamoto, K., Chen, Y., Ning, L., Georgiou, T., Tannenbaum, A.: Regularization and interpolation of positive matrices (2016). https://arxiv.org/abs/1611.07945