Matrix-valued optimal mass transportation and its applications

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Lipeng Ning

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
Doctor of Philosophy

Professor Tryphon T. Georgiou

November, 2013

© Lipeng Ning 2013
ALL RIGHTS RESERVED

Acknowledgements

First of all, I want to express my gratitude and appreciation to my advisor Professor Tryphon T. Georgiou. It has been such a great honor to work with such a deep thinker in the past five years. His profound knowledge, passion for research and great patience make him an idea mentor. His advice made me strong both intellectually and physically. His encouragement and confidence in me will always be a source of energy. In addition to his intellectual guidance, I am also thankful for his financial support via the grants from NSF and AFOSR.

I wish to express my sincere appreciation to my final exam committee members, Professor Mihailo Jovanovic, Professor Zhi-Quan (Tom) Luo and Professor Peter Seiler, for reviewing my thesis and offering constructive comments. The collaboration with Professor Allen Tannenbaum, Professor Stephen Boyd and my advisor Professor Tryphon Georgiou was very educational and had great impact on my research throughout the last years of my PhD studies. I am thankful to Professor Murti Salapaka and Professor Guillermo Sapiro for being my preliminary oral exam committee members. I am also grateful to Professor William Peria for his help in improving my language.

I would like to acknowledge my former colleague Xianhua Jiang. It was such an exciting time working with her on the problem of multivariate power spectra and covariance distances. I want to thank Donnatello Materassi, Francesca P. Carli, Hamza Farooq, Hartmut Durchschlag and Eric Dahlberg for sharing the office with me. My thanks also goes to Yongxin Chen for the time we spent together in gym and on discussing research problems. I appreciate my friendship with all the Mihailo's group members: Fu Lin, Rashad Moarref, Binh Lieu, Armin Zare, Neil Dingra, Xiaofan Wu. I also want to thank my friend Johan Karlsson for inviting and hosting me for my visit to Florida.

Finally, I owe my deepest gratitude to my family. My parents are always supportive for me to pursue my PhD. The journey of life becomes more joyful with my wife's accompaniment.

Dedication

To my parents and my wife

Abstract

The subject of this thesis is the geometry of matrix-valued density functions. The main motivation is the need for quantitative measures to compare power spectral densities of multivariate time-series. Distance measures between statistical objects provide fundamental tools for estimation, tracking and classification. In particular, for power spectra, such notions of distance are especially relevant in modeling slowly varying time-series. To this end, power spectra estimated from short observation records are considered as data points on a statistical manifold and can be connected by a regression geodesic induced by a suitable distance measure. Moreover, metrics to compare power spectra are key in quantifying resolution in spectral analysis and in various problems in statistical estimation and smoothing.

We study classical notions of distance, such as Fisher information metric, Kullback-Leibler and Itakura-Saito distance, and their multivariable generalizations. We explore the Riemannian geometric structure and derive geodesics on the corresponding statistical manifolds, we draw connections with analogous notions of distance in Quantum mechanics to compare density matrices.

We introduce two formulations of matrix-valued Monge-Kantorovich optimal mass transport (OMT) problem. In the first formulation, we use a notion of non-positive transportation plan and we show that the induced Wasserstein metric is weakly continuous. The second formulation leads to a rotation-aware distance measure between the end-point power spectra that takes into account the transference of power over frequencies as well as the rotation of the principle directions. In this, we show that the optimal transportation plan is no longer supported on a monotonically increasing thin set. Applications to spectral tracking and spectral morphing highlight the relevance of the proposed distance.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

A basic problem in signal processing and control is to quantify dissimilarities between signals and between statistical models [1, 2]. Indeed, distance measures play a central role in detection [3], tracking [4], pattern recognition [5], and clustering [6]. A variety of choices are readily available for comparing deterministic signals and systems. These included various L_p norms on signal space and induced norms in spaces of systems. On the other hand, statistical models, such as covariances, histograms, probability distributions, and power spectra, are not elements in of a vector space. Their geometry is dictated by positivity constraints and hence, they lie on suitable cones or simplices. Various theories have been developed for comparing statistical models such as those in the classical work of C.R. Rao and R.A Fisher on "information geometry" [7, 8, 9, 10], the generalization of information theory to quantum mechanics [11] and the optimal transportation theory [12]. The present work focuses on the problem of comparing matrix-valued density functions. Our interest originates in spectral analysis of multivariate stochastic process and stems from an apparent need for weakly-continuous "rotation-aware" geometric tools that could be used to quantify, interpolate and approximate matrix-valued data.

Spectral analysis is central in antenna arrays, speech signal analysis, stock analysis and many other areas in signal processing. Typically, a power spectral density (PSD) is estimated from a second-order stationary time-series to represent the distribution of energy over frequencies. This can be used to estimate the locations of sources, formants and identifying other periodic patterns. Much effort has been devoted for PSD estimation from a finite record of observations. We refer to the book of Stoica and Moses [13] for a detailed exposition of spectral estimation methods and a series of works in [14, 15, 16, 17] on a generalized framework for high

resolution analysis. Non-stationary time-series analysis techniques often rely on the assumption that the spectral content of the signal does not change too fast. Thus, various parametric methods assume that the parameters of the model, such as an ARMA or AR model, change slowly with time [18, 19, 20]. Non-parametric methods are similarly based on spectral estimation over sufficient short time intervals [21]. A novel method was proposed recently in [22] where a geodesic path on the manifold of power spectra is considered as a non-parametric model for time-varying spectra. In this, short windowed spectral density estimates are considered as points on the spectral manifold tracing a (geodesic) path corresponding to the time-variability. Construction of such a (geodesic) path relies heavily on the availability of an appropriate metric between spectra. In fact, because the Wasserstein metric possesses the desirable weak-continuity property, it was used in [22] as being a natural metric.

In signal analysis, it is often that the distance measure or divergence is used to describe a nonnegative function of the two variables

$$\mathrm{D}(f_0, f_1)$$

where typically f_0, f_1 are power spectra. Examples include the Kullback-Leibler divergence, the Itakura-Saito distance and the distance devised in [23] based on the theory of optimal prediction. However, they are often not symmetric and they do not satisfy the triangular inequality.

Distances between multivariate spectra have only recently received any attention. In this direction, there are generalizations of Hellinger and Itakura-Saito distances [24, 25, 26, 27] and the Umegaki-von Neumann relative entropy [28]. A distance based on optimal prediction theory was devised in [23] and has been recently generalized to a multivariate version in [29]. However, all these distance measures fail to be weakly continuous. This motivates the generalization of the Monge-Kantorovich geometry to matrix-valued density functions and this is the subject of the present work.

Throughout the thesis, we use normal fonts such as f, m, μ for scalar or scalar-valued functions. We use bold symbols such as $\boldsymbol{f}$, $\boldsymbol{m}$, $\boldsymbol{\mu}$ for matrices (or vectors) and matrix-valued functions. In the following section we review preliminaries on matrix-valued power spectra and introduce notations.

1.1 Matrix-valued power spectra

Consider a multivariate discrete-time, zero-mean, weakly stationary stochastic process $\{\boldsymbol{u}(k), k \in \mathbb{Z}\}$ with $\boldsymbol{u}(k)$ taking values in $\mathbb{C}^{\ell \times 1}$ with $\ell \in \mathbb{N}$. We denote by

$$R_k = \mathscr{E}(\boldsymbol{u}(i)\boldsymbol{u}^*(i-k)) \text{ for } i,k \in \mathbb{Z}$$

the sequence of matrix covariances, $\mathscr{E}$ the expectation, and $*$ the complex conjugate transpose. Let $d\boldsymbol{\mu}(x)$ be the corresponding matrix-valued spectral measure, then

$$R_k = \int_{-\pi}^{\pi} e^{-\mathrm{j}k\theta} \frac{d\boldsymbol{\mu}(\theta)}{2\pi}$$

for $k \in \mathbb{Z}$, see, e.g. [30] where $\mathrm{j} = \sqrt{-1}$.

It is well-known that a covariance sequence $\{R_k : \ell \in \mathbb{Z} \text{ and } R_{-k} = R_k^*\}$ is completely characterized by the non-negativity of the block-Toeplitz matrices

$$\begin{bmatrix} R_0 & R_1 & \dots & R_k \\ R_{-1} & R_0 & \dots & R_{k-1} \\ \vdots & \vdots & \ddots & \vdots \\ R_{-k} & R_{-k+1} & \dots & R_0 \end{bmatrix} \tag{1.1}$$

for all k. Such an infinite sequence of positive semi-definite block-Toeplitz matrices qualifies as a covariance sequence of a stochastic process and vice versa. At the same time, the infinite sequence of $\{R_k : k \in \mathbb{Z}\}$ defines the spectral measure $d\boldsymbol{\mu}$ and also vice versa. For the most part, unless we specifically indicate otherwise, we will be concerned with the case of a non-deterministic process of full rank with an absolutely continuous power spectrum. Hence,

$$d\boldsymbol{\mu}(\theta) = \boldsymbol{f}(\theta)d\theta$$

with $\boldsymbol{f}(\theta)$ being a matrix-valued power spectral density function. For any fixed θ, $\boldsymbol{f}(\theta)$ is a positive semi-definite Hermitian matrix. If the dimension $\ell = 1$, then $\boldsymbol{f}$ is a scalar-valued PSD function and it is denoted as f (without bold).

The computation of $\boldsymbol{f}$ requires the infinite sequence of covariances $\{R_k : k \in \mathbb{Z}\}$. Usually, when a finite observation record is available, we can only obtain a finite sequence of estimated covariances. Spectral estimation is about the problem of determining a power spectral density function that is consistent with the finite record of observations. Estimation usually involves

parametric or nonparametric method for extending the finite covariance sequence to an infinite one corresponding to a positive semi-definite infinite block-Toeplitz matrix as in (1.1). In this, we also mention that distance measures between power spectra are important in spectral estimation and in reconciling data with prior information [31, 28, 25, 32, 33, 27].

1.2 Outline and contributions

In this thesis, we study the geometry for matrix-valued power spectra, and propose generalizations of the Monge-Kantorovich optimal mass transportation distance for comparing matrix-valued density functions. These lead to natural distance measures that are suitable for comparing and approximating power spectra, for modeling slow time-varying stochastic processes, for morphing and deforming the spectral content of time-series and for other spectral analysis tasks.

In Chapter 2, we review various distances that have already been used for comparing matrix-valued power spectra. These include the generalized Kullback-Leibler relative entropy, generalized Itakura-Saito distance, generalized Hellinger distance and the distances that are devised based on the optimal prediction theory. The corresponding Riemannian structures induced by these distances are also discussed. These distances lack a desired weak-continuity property and this motivates our work of exploring generalizations of the Monge-Kantorovich geometry.

In Chapter 3, we present some background materials of OMT on comparing scalar-valued probability density functions. We discuss a generalization of OMT for density functions with unbalanced masses, and also study in detail about transportation between Gaussian distributions. In fact, the Wasserstein distance between Gaussian distributions induces a metric between corresponding covariance matrices as well. Computation of such distances can be cast as a solution of a linear matrix inequality (LMI) problem. Because of this fact, the corresponding metric is natural and easy to use for covariance approximation problems. We finally study an extension of the optimal transportation problem between distributions on the circle and discuss the properties of the geometry of the corresponding optimal transference plan.

In Chapter 4, we propose a generalized framework of the Monge-Kantorovich problem for matrix-valued density functions that have equal integral. Our formulation generalizes the so-called Kantorovich-Rubinstein transshipment problem. In this, we introduce a notion of non-positive matrix-valued transference plan. The nuclear norm of the matrix-valued density

is penalized. We also study an extension of this formulation to the case where the densities have non-equal integral. We show that the weak continuity property of the Wasserstein metric is nicely inherited by the metrics we devised in this context.

In Chapter 5, an alternative formulation of the Monge-Kantorovich problem for matrix-valued density functions is proposed which is derived under the more general assumption that the integrals of density functions have equal trace. In this, the transportation cost includes two part: a cost of transference of "mass" across frequencies and a cost of rotating the "direction" of matrix densities as a functions of frequency. The two endpoint matrix-valued densities can be thought of as marginals of a joint density on a tensor product space. We cast the problem as one of convex optimization. Contrary to the classical setting of the Monge-Kantorovich OMT, the optimal transport plan for matrix-valued densities is no longer supported on a monotonically increasing thin set.

In Chapter 6, we review certain generalizations of the Wasserstein metric to noncommutative geometry. The relation between the distance we introduced in Chapter 4 and the spectral distance in noncommutative geometry is discussed and a family of metrics is suggested that relies on suitable set of test functions.

In Chapter 7, we first compare the various distances for matrix-valued spectra. Naturally, OMT is more suitable to quantify shift of content across frequencies. Then we present numerical examples of the proposed distances in multivariate spectral analysis. In this, we apply the proposed OMT-based distance to the problems of spectral morphing and spectral tracking and we demonstrate the ability of the distance in tracking the shift of power across frequencies as well as the rotation of power content between channels.

Chapter 2

Geometry of matrix-valued power spectra

Various methods have been devised to compare scalar-valued PSD functions. Key among those are the Itakura-Saito distance and the logarithmic spectral deviation see e.g., [2, page 370]. A very related alternative to the Itakura-Saito distance is devised in [23] to quantify the dissimilarity of models in the context of optimal prediction theory for second-order stationary stochastic processes. Recently, the optimal mass transportation theory has been adopted in [22] to compare PSD functions and to model slowly time-varying process. It has demonstrated remarkable performance in capturing the drift of power across frequencies over time.

Matrix-valued power spectra are utilized in analyzing multivariate time-series. Distances between matrix-valued power spectra have only recently received attention. In this direction, we note the generalized Hellinger distance [24, 25], the generalized Itakura-Saito distances [26, 27], and the Umegaki-von Neumann relative entropy [28, 25]. A metric based on the optimal prediction theory was devised in [23] and has been recently generalized to a multivariate version in [29].

2.1 Distance measures

The problem of comparing dissimilarities between matrix-valued PSD functions is timely [24, 25, 32]. The evident goal is to provide a means to quantify deviation and uncertainties in the

spectral domain in a way that is suitable with particular applications. Appropriate distance measures can be used in averaging of multiple spectra [22] and to regularize spectral approximation problems by suitably weighing second-order statistics [28, 25]. We overview certain distance measures that have been devised to compare multivariate power spectra.

Generalized Kullback-Leibler relative entropy

Herein we use p to denote probability density functions on a support set (e.g. $p(x), x \in \mathbb{R}$) and likewise $\boldsymbol{\rho}$ for density matrices. These are non-commutative probability vectors. They are Hermitian positive semi-definite matrices with trace equal to one.

The Kullback-Leibler (K-L) relative entropy

$$\mathrm{D}_{\mathrm{KL}}(p_0||p_1) = \int p_0 \log p_0 - p_0 \log p_1 dx$$

has played a central role in information theory and provides a non-symmetric measure of the difference between two probability distributions p_0 and p_1 [34]. It is non-negative and it is zero if and only if $p_0 = p_1$. Since both probability densities and spectral densities are positive, the K-L relative entropy can be adapted to power spectra by normalizing these to have the same integral [27].

In quantum mechanics, the Umegaki-von Neumann relative entropy [35]

$$\mathrm{d}_{\mathrm{Umegaki}}(\boldsymbol{\rho}_0||\boldsymbol{\rho}_1) = \mathrm{tr}(\boldsymbol{\rho}_0 \log \boldsymbol{\rho}_0 - \boldsymbol{\rho}_0 \log \boldsymbol{\rho}_1) \tag{2.1}$$

generalizes the K-L relative entropy to density matrices. Likewise, it can be adopted to compare matrix-valued power spectra. In fact, for $\boldsymbol{f}_0, \boldsymbol{f}_1 > 0$, the expression

$$\mathrm{D}_{\mathrm{KL}}(\boldsymbol{f}_0||\boldsymbol{f}_1) = \int_{-\pi}^{\pi} \mathrm{tr}\,(\boldsymbol{f}_0 \log \boldsymbol{f}_0 - \boldsymbol{f}_0 \log \boldsymbol{f}_1)\, d\theta \tag{2.2}$$

was used in [28] in spectral approximation problems. Apparently, D_{KL} may take negative values for some pairs of $\boldsymbol{f}_0$ and $\boldsymbol{f}_1$. For example, if $\boldsymbol{f}_1 > \boldsymbol{f}_0$ then $\log(\boldsymbol{f}_1) > \log(f_0)$, thus $\mathrm{D}_{\mathrm{KL}}(\boldsymbol{f}_0||\boldsymbol{f}_1) < 0$. We note the following inequality

$$\mathrm{tr}\,(\boldsymbol{f}_0 \log \boldsymbol{f}_0 - \boldsymbol{f}_0 \log \boldsymbol{f}_1) \geq \mathrm{tr}(\boldsymbol{f}_0 - \boldsymbol{f}_1)$$

see, [11, page 174]. Thus $\mathrm{D}_{\mathrm{KL}}(\boldsymbol{f}_0||\boldsymbol{f}_1)$ is nonnegative whenever

$$\mathrm{tr}(\int_{-\pi}^{\pi} \boldsymbol{f}_0 d\theta) = \mathrm{tr}(\int_{-\pi}^{\pi} \boldsymbol{f}_1 d\theta)$$

and this can be ensured by suitable normalization.

Generalized Itakura-Saito distance

The Itakura-Saito (I-S) distance

$$\mathrm{D}_{\mathrm{IS}}(f_0, f_1) = \int_{-\pi}^{\pi} \left(\frac{f_0(\theta)}{f_1(\theta)} - \log \frac{f_0(\theta)}{f_1(\theta)} - 1 \right) \frac{d\theta}{2\pi} \tag{2.3}$$

between power spectral density functions has deep connection with the K-L relative entropy. To explain the connection, consider n-dimensional Gaussian distributions functions p_0 and p_1 with covariance matrices T_0 and T_1. Then

$$\mathrm{D}_{\mathrm{KL}}(p_0||p_1) = \frac{1}{2}(\mathrm{tr}(T_0 T_1^{-1}) - \log|T_0 T_1^{-1}| - n).$$

The relation between K-L relative entropy and I-S distance is now apparent. In fact, if p_0 and p_1 represent the law of Gaussian stochastic processes the two coincides.

Formally, let $\{\ldots,\ r_{i,-1},\ r_{i,0},\ r_{i,1}, \ldots\}$ be covariance sequences of two stationary Gaussian processes with spectral density functions f_i for $i = 0, 1$, and let

$$T_{i,n} = \begin{bmatrix} r_{i,0} & r_{i,1} & \ldots & r_{i,n-1} \\ r_{i,-1} & r_{i,0} & \ldots & r_{i,n-2} \\ \vdots & \vdots & \ddots & \vdots \\ r_{i,1-n} & r_{i,2-n} & \ldots & r_{i,0} \end{bmatrix}$$

be the corresponding $n \times n$ Toeplitz structured covariance matrices. If $p_{0,n}$ and $p_{1,n}$ are Gaussian densities with covariance $T_{0,n}$ and $T_{1,n}$, then

$$\frac{2}{n}\mathrm{D}_{\mathrm{KL}}(p_{0,n}||p_{1,n}) = \frac{1}{n}(\mathrm{tr}(T_0 T_1^{-1}) - \log|T_0 T_1^{-1}|) - 1$$

and in the limit we obtain

$$\lim_{n\to\infty} \frac{2}{n}\mathrm{D}_{\mathrm{KL}}(p_{0,(n)}||p_{1,(n)}) = \mathrm{D}_{\mathrm{IS}}(f_0, f_1)$$

follow [36, 24]. The I-S distance has been used in speech signal discrimination [2] and spectral approximation [27]. The non-negativity of $\mathrm{D}_{\mathrm{IS}}(f_0, f_1)$ is guaranteed from the inequality

$$x \geq 1 + \log x, \text{ for } x > 0.$$

For matrix-valued power spectra, the I-S distance is generalized as

$$\mathrm{D}_{\mathrm{IS}}(\boldsymbol{f}_0, \boldsymbol{f}_1) = \int_{-\pi}^{\pi} \mathrm{tr}\left(\boldsymbol{f}_0 \boldsymbol{f}_1^{-1} - \log(\boldsymbol{f}_0 \boldsymbol{f}_1^{-1}) - I\right) \frac{d\theta}{2\pi} \tag{2.4}$$

which is studied in [24] in the context of spectral approximation problem.

Distance based on optimal prediction error variance

The Itakura-Saito distance is closely related to the optimal prediction theory [2] and a distance devised in [23, 29] which quantify the dissimilarity of PSD's based on prediction error variance. We now develop this alternative distance.

Consider a multivariate, discrete-time, zero-mean, weakly stationary stochastic process $\{\boldsymbol{u}(k) : k \in \mathbb{Z}\}$ with $\boldsymbol{u}(k)$ taking values in $\mathbb{C}^{\ell\times 1}$. In least-variance linear prediction, we consider evaluating

$$\min_{P_k}\left\{\mathrm{tr}\mathscr{E}(\boldsymbol{e}\boldsymbol{e}^*)\ :\ \boldsymbol{e}=\boldsymbol{u}(0)-\sum_{k>0}P_k\boldsymbol{u}(-k),\ P_k\in\mathbb{C}^{\ell\times\ell}\right\}. \tag{2.5}$$

Let Ω denote the optimal prediction error variance $\mathscr{E}(\boldsymbol{e}\boldsymbol{e}^*)$ and let $\boldsymbol{f}(\theta)$ be the PSD of the stochastic process. If $\log\det\boldsymbol{f}(\theta)\in L_1[-\pi,\pi]$, the variance is finite and the well-known Szegö-Kolmogorov formula [30, pg. 369]

$$\det\Omega=\exp\{\int_{-\pi}^{\pi}\log\det\boldsymbol{f}(\theta)\frac{d\theta}{2\pi}\} \tag{2.6}$$

relates the error variance with the corresponding PSD. Let $\hat{P}_k$'s be the optimal solution of (2.5). Then,

$$\boldsymbol{p}(z)=I-\sum_{k>0}\hat{P}_kz^k$$

is the optimal prediction filter and relates to a spectral factorization of $\boldsymbol{f}(\theta)$ as follows. The power spectral density function $\boldsymbol{f}(\theta)$ admits a unique factorization

$$\boldsymbol{f}(\theta)=\boldsymbol{f}_+(e^{\mathrm{j}\theta})\boldsymbol{f}_+(e^{\mathrm{j}\theta})^* \tag{2.7}$$

with $\boldsymbol{f}_+(z)$ being analytic inside the disc $\{z:|z|<1\}$ and $\boldsymbol{f}_+(0)=\Omega^{\frac{1}{2}}$. The matrix-function $\boldsymbol{f}_+(z)$ is the so-called outer factor and the optimal prediction filter $\boldsymbol{p}(z)$ is explicitly expressed as

$$\boldsymbol{p}(z)=\boldsymbol{f}_+(0)\boldsymbol{f}_+^{-1}(z).$$

Consider now two time-series with PSD's $\boldsymbol{f}_0$ and $\boldsymbol{f}_1$, respectively. Let $\boldsymbol{p}_0(z)$ and $\boldsymbol{p}_1(z)$ be the optimal prediction filter for $\boldsymbol{f}_0$ and $\boldsymbol{f}_1$, respectively, and let Ω_0 and Ω_1 be the corresponding optimal prediction error variance. By using $\boldsymbol{p}_1(z)$ as a prediction filter for the time-series corresponds to $\boldsymbol{f}_0$, we obtain a prediction error variance $\Omega_{0,1}$ which satisfies

$$\Omega_{0,1}\geq\Omega_0$$

since $\boldsymbol{p}_1$ may not be optimal for $\boldsymbol{f}_0$. The matrix $\Omega_{0,1}$ represents a degraded error-variance since a suboptimal prediction filter may have been used. By normalizing into

$$\Omega_0^{-\frac{1}{2}}\Omega_{0,1}\Omega_0^{-\frac{1}{2}}$$

the variance of the prediction error, we can quantify mismatch between $\boldsymbol{f}_0$ and $\boldsymbol{f}_1$. Indeed if $\boldsymbol{f}_0$ and $\boldsymbol{f}_1$ coincide (or if $\boldsymbol{p}_0(z) = \boldsymbol{p}_1(z)$) then

$$\Omega_0^{-\frac{1}{2}}\Omega_{0,1}\Omega_0^{-\frac{1}{2}} = I.$$

Thus we quantify the dissimilarity by

$$\mathrm{D}_{\mathrm{pred}}(\boldsymbol{f}_0, \boldsymbol{f}_1) := \log\det(\Omega_0^{-\frac{1}{2}}\Omega_{0,1}\Omega_0^{-\frac{1}{2}}).$$

This distance can be expressed in term of the spectral densities as

$$\mathrm{D}_{\mathrm{pred}}(\boldsymbol{f}_0, \boldsymbol{f}_1) = \mathrm{tr}\left(\log\int_{-\pi}^{\pi}\boldsymbol{f}_{0+}^{-1}\boldsymbol{f}_1\boldsymbol{f}_{0+}^{-*}\frac{d\theta}{2\pi} - \int_{-\pi}^{\pi}\log(\boldsymbol{f}_{0+}^{-1}\boldsymbol{f}_1\boldsymbol{f}_{0+}^{-*})\frac{d\theta}{2\pi}\right) \tag{2.8}$$

which generalizes the distance devised in [23] for scalar-valued case.

Distance based on flatness of innovation power spectra

An alternative way of comparing spectral densities using optimal prediction theory is to measure the flatness of the innovation power spectra [29]. Consider a stochastic process with power spectral density $\boldsymbol{f}$, and assume it admits a spectral factorization as in (2.7). If we apply the scaled optimal prediction filter $\boldsymbol{f}_+^{-1}(z)$ to the stochastic process, the innovation process has a constant power spectrum $\boldsymbol{f}_+^{-1}\boldsymbol{f}\boldsymbol{f}_+^{-*} = I$ which indicates that the innovation process has a "flat" spectrum.

Consider two stochastic processes with PSD's $\boldsymbol{f}_0$ and $\boldsymbol{f}_1$, respectively. If $\boldsymbol{f}_{1,+}^{-1}$ is applied to the process with PSD $\boldsymbol{f}_0$, the power spectrum of the output process is

$$\boldsymbol{f}_{1,+}^{-1}\boldsymbol{f}_0\boldsymbol{f}_{1,+}^{-*} \neq I$$

which is not "flat". By quantifying the difference between the non-flat spectrum from being the constant I in a symmetric way, we are lead to the distance introduced in [29]

$$\mathrm{D}_{\mathrm{flat}}(\boldsymbol{f}_0, \boldsymbol{f}_1) = \int_{-\pi}^{\pi}\mathrm{tr}(\boldsymbol{f}_0^{-\frac{1}{2}}\boldsymbol{f}_1\boldsymbol{f}_0^{-\frac{1}{2}} + \boldsymbol{f}_0^{\frac{1}{2}}\boldsymbol{f}_1^{-1}\boldsymbol{f}_0^{\frac{1}{2}} - 2I)\frac{d\theta}{2\pi}. \tag{2.9}$$

The non-negativity of (2.9) is guarantied by the inequality

$$x + \frac{1}{x} \geq 2, \text{ for } x > 0.$$

Generalized Hellinger/Bures distance

The Hellinger distance between two (scalar-valued) probability density functions f_0 and f_1 is given by

$$\frac{1}{\sqrt{2}}\left(\int(\sqrt{f_0}-\sqrt{f_1})^2dx\right)^{\frac{1}{2}}$$

see e.g. [37]. It turns out that there is a very similar expression in quantum mechanics that has been used to compare density matrices as we pointed out in [38]. More specifically, the Bures distance is defined as [39]

$$\mathrm{d}_{\mathrm{Bures}}(\boldsymbol{\rho}_0,\boldsymbol{\rho}_1)=\min_U\left\{\|\boldsymbol{\rho}_0^{\frac{1}{2}}-\boldsymbol{\rho}_1^{\frac{1}{2}}U\|_{\mathrm{F}}\mid UU^*=I\right\} \tag{2.10}$$

where U is a unitary matrix, $\boldsymbol{\rho}_0^{\frac{1}{2}}$ and $\boldsymbol{\rho}_1^{\frac{1}{2}}$ are respectively the unique positive semi-definite square root of $\boldsymbol{\rho}_0$ and $\boldsymbol{\rho}_1$, respectively, and $\|\cdot\|_{\mathrm{F}}$ denotes the Frobenius norm. This distance has been extensively used in quantum mechanics to measure distance between quantum states. The optimization over U in (2.10) is natural since U represents a gauge transformation. The optimal rotation $\hat{U}$ in (2.10) "lining up" the two factors is

$$\hat{U}=\boldsymbol{\rho}_1^{-\frac{1}{2}}\boldsymbol{\rho}_0^{-\frac{1}{2}}(\boldsymbol{\rho}_0^{\frac{1}{2}}\boldsymbol{\rho}_1\boldsymbol{\rho}_0^{\frac{1}{2}})^{\frac{1}{2}} \tag{2.11}$$

and the Bures distance has a closed form as

$$\mathrm{d}_{\mathrm{Bures}}(\boldsymbol{\rho}_0,\boldsymbol{\rho}_1)=\mathrm{tr}(\boldsymbol{\rho}_0+\boldsymbol{\rho}_1-2(\boldsymbol{\rho}_0^{\frac{1}{2}}\boldsymbol{\rho}_1\boldsymbol{\rho}_0^{\frac{1}{2}}))^{\frac{1}{2}} \tag{2.12}$$

As we pointed out in [38], the Bures distance $\mathrm{d}_{\mathrm{Bures}}$ is equivalent to the 2-Wasserstein distance between two zero-mean Gaussian distributions with covariance matrices $\boldsymbol{\rho}_0$ and $\boldsymbol{\rho}_1$, respectively, which will be explained in detail in Section 3.2.1 Chapter 3.

The interest in this distance measure in the context of comparing PSD's was initiated in [25] where a matrix generalization of the Hellinger distance was first introduced as

$$\mathrm{D}_{\mathrm{H}}(\boldsymbol{f}_0,\boldsymbol{f}_1)=\left(\int_{-\pi}^{\pi}\mathrm{d}_{\mathrm{Bures}}(\boldsymbol{f}_0,\boldsymbol{f}_1)^2\frac{d\theta}{2\pi}\right)^{\frac{1}{2}} \tag{2.13}$$

and used in the context of approximation and spectral estimation problems.

2.2 Riemannian structures

Consider a "small" perturbation $\boldsymbol{f} + \Delta$ away from a nominal power spectral density $\boldsymbol{f}$. All the distances/divergences we have seen so far are continuous in their arguments and in the small, can be approximated by a quadratic form in Δ which depends on $\boldsymbol{f}$. This is what it referred to as a Riemannian metric[1] . The availability of a metric dictates how perturbations in various directions compare to each other. It also provides the concepts of geodesics and geodesic distances. Geodesics are paths of shortest length connecting the start and the finish; the length is the geodesic distance.

Geodesics in the space of power spectral densities represent deformations from a "starting" power spectral density $\boldsymbol{f}_0$ to an "end-point" $\boldsymbol{f}_1$. These are natural structures for modeling changes and deformations. A key motivation behind the present work is to model time-varying spectra via a geodesic path in a suitable metric space as in [22]. This viewpoint provides a non-parametric model for nonstationary spectra, analogous to a spectrogram, but one which takes into account the inherent geometry of power spectral densities.

We consider the distance measures shown in the previous section for infinitesimal perturbations about a given power spectral density function. We will show the Riemannian metric that are derived by each of the distances. For simplicity, we assume that the power spectral densities are strictly positive definite and differentiable. We also assume that the perturbations are well-behaved so that the perturbed density functions still have the desired properties.

We note that all the distances, except $\mathrm{D}_{\mathrm{pred}}$, are based on evaluating an integral of the form

$$\int_{-\pi}^{\pi} \mathrm{d}(\boldsymbol{f}_0, \boldsymbol{f}_1) d\theta$$

with $\mathrm{d}(\cdot,\cdot)$ being a distance between positive semi-definite matrices. The Riemannian metrics induced by the distance $\mathrm{d}(\cdot,\cdot)$ on the manifold of positive semi-definite matrices induce the Riemannian structures for power spectral density functions.

Riemannian metric induced by the generalized K-L divergence

We consider the Umegaki-von Neumann relative entropy for density matrices. Using variational analysis we derive an expression for the corresponding Riemannian metric. Let X and Y be

[1] A positive-definite quadratic form $\mathrm{g}_{\boldsymbol{f}}(\Delta)$ in Δ which continuously depends on the point $\boldsymbol{f}$ on the manifold.

perturbations of a positive semidefinite matrix M. Starting from the Hessian

$$-\frac{\partial^2}{\partial t \partial s} \mathrm{d}_{\mathrm{Umegaki}}(M+tX||M+sY) = \int_0^\infty \mathrm{tr}\left(X(M+uI)^{-1}Y(M+uI)^{-1}\right) du, \tag{2.14}$$

we obtain the Kubo-Mori metric [40]

$$\mathrm{g}_{1,M}(\Delta) := \int_0^\infty \mathrm{tr}\left(\Delta(M+uI)^{-1}\Delta(M+uI)^{-1}\right) du. \tag{2.15}$$

This is a Riemannian metric on the manifold of positive definite matrices. The expression (2.14) is based on the expansion of the logarithm

$$\log(M+\Delta) = \log(M) + \int_0^\infty (M+uI)^{-1}\Delta(M+uI)^{-1} du + o(\|\Delta\|)$$

where

$$\int_0^\infty (M+uI)^{-1}\Delta(M+uI)^{-1} du$$

can be thought of as a non-commutative division of Δ by M.

Equation (2.15) gives rise to

$$\mathrm{g}_{1,\boldsymbol{f}}(\Delta) = \int_{-\pi}^{\pi}\int_0^\infty \mathrm{tr}\left(\Delta(\boldsymbol{f}+uI)^{-1}\Delta(\boldsymbol{f}+uI)^{-1}\right) du d\theta \tag{2.16}$$

as a metric between PSD's. We point out that the geodesics induced by the Kubo-Mori metric (2.15) on the manifold of positive definite matrices are still unknown.

Riemannian metric induced by the I-S distance and "flatness" of spectra

The generalized Itakura-Saito distance (2.4) and the distance based on the "flatness" of spectra of innovation process (2.9) are respectively related to the following distances between positive definite matrices M_0 and M_1

$$\begin{aligned} \mathrm{d}_{\mathrm{IS}}(M_0,M_1) &:= \mathrm{tr}(M_0M_1^{-1} - \log(M_0M_1^{-1}) - I) \\ \mathrm{d}_{\mathrm{flat}}(M_0,M_1) &:= \mathrm{tr}(M_0^{-1}M_1 + M_1^{-1}M_0 - 2I)\,. \end{aligned}$$

Consider Δ as a perturbation of M. By expanding $\mathrm{d}_{\mathrm{IS}}(M+\Delta,M)$ and $\mathrm{d}_{\mathrm{flaw}}(M+\Delta,M)$ in terms of Δ, we find that they have the same quadratic term, see [29], that is expressed as

$$\mathrm{g}_{2,M}(\Delta) := \mathrm{tr}(M^{-1}\Delta M^{-1}\Delta). \tag{2.17}$$

This is a Riemannian metric on the manifold of positive definite matrices and it coincides to the Fisher-Rao metric for zero-mean Gaussian distributions [41].

The Fisher-Rao metric has been studied extensively in recent years [41, 42]. The geodesic and geodesic length induced by this metric has been derived explicitly. Given two positive definite matrices M_0 and M_1, and let $M(\tau)$ with $\tau \in [0,1]$ be a smooth path on the manifold of positive definite matrices. Then the geodesic connecting M_0 and M_1 induced by $\mathrm{g}_{2,M}$ is the solution of

$$\inf_{M(\tau)} \left\{ \int_0^1 \sqrt{\mathrm{g}_{2,M}(\dot{M})} dt \mid M(0) = M_0,\ M(1) = M_1 \right\} \tag{2.18}$$

which is explicitly given as

$$M(\tau) = M_0^{\frac{1}{2}} (M_0^{-\frac{1}{2}} M_1 M_0^{-\frac{1}{2}})^{\tau} M_0^{\frac{1}{2}}.$$

The the geodesic distance is given as

$$\| \log(M_0^{-\frac{1}{2}} M_1 M_0^{-\frac{1}{2}}) \|_{\mathrm{F}}.$$

The Riemannian metric for matrix-valued PSD's induced by D_{IS} and $\mathrm{D}_{\mathrm{flat}}$ is

$$\mathrm{g}_{2,\boldsymbol{f}}(\Delta) = \int_{-\pi}^{\pi} \mathrm{tr}\left(\boldsymbol{f}^{-1}\Delta\boldsymbol{f}^{-1}\Delta\right) \frac{d\theta}{2\pi}.$$

The corresponding geodesic connecting $\boldsymbol{f}_0$ and $\boldsymbol{f}_1$ is

$$\boldsymbol{f}(\tau) = \boldsymbol{f}_0^{\frac{1}{2}} (\boldsymbol{f}_0^{-\frac{1}{2}} \boldsymbol{f}_1 \boldsymbol{f}_0^{-\frac{1}{2}})^{\tau} \boldsymbol{f}_0^{\frac{1}{2}} \tag{2.19}$$

which has been studied in [29] for analyzing the deformation of PSD's.

Riemannian metric induced by sub-optimality of prediction

Consider two matrix-valued PSD's $\boldsymbol{f}$ and $\boldsymbol{f}+\Delta$, the quadratic form approximation of

$$\mathrm{D}_{\mathrm{pred}}(\boldsymbol{f}, \boldsymbol{f}+\Delta)$$

in term of Δ leads to the Riemannian metric [29]

$$\mathrm{g}_{3,\boldsymbol{f}}(\Delta) := \mathrm{tr} \int_{-\pi}^{\pi} (\boldsymbol{f}_+^{-1}\Delta\boldsymbol{f}_+^{-*})^2 \frac{d\theta}{2\pi} - \mathrm{tr}\left(\int_{-\pi}^{\pi} \boldsymbol{f}_+^{-1}\Delta\boldsymbol{f}_+^{-*} \frac{d\theta}{2\pi} \right)^2.$$

In the case when $\boldsymbol{f}$ is a scalar-valued PSD, this reduces to the metric derived in [23]. The geodesics induced by $\mathrm{g}_{3,\boldsymbol{f}}$ are unknown in general. But in the scalar case, a geodesic given in [23] is expressed as

$$f(\tau) = f_0 \left(\frac{f_1}{f_0}\right)^{\tau} \text{ for } \tau \in [0,\ 1]$$

which coincides with (2.19). The geodesics induced by $\mathrm{g}_{3,\boldsymbol{f}}$ are not unique [23].

Riemmannian metric induced by the generalized Hellinger/Bures distance

The distance $\mathrm{d}_{\mathrm{Bures}}$ in (2.12) induces the Bure's metric [39, 43]

$$\mathrm{g}_{4,M}(\Delta) := \mathrm{tr}(\Delta X),\ \text{where } \frac{1}{2}(MX + XM) = \Delta \tag{2.20}$$

where X is thought of as a non-commutative division of Δ by M. Given two positive definite matrices M_0 and M_1, let

$$W_0 = M_0^{\frac{1}{2}} \text{ and } W_1 = M_0^{-\frac{1}{2}}(M_0^{\frac{1}{2}} M_1 M_0^{\frac{1}{2}})^{\frac{1}{2}}.$$

Then $M_i = W_i W_i^*$ for $i = 0, 1$. The distance $\mathrm{d}_{\mathrm{Bures}}(M_0, M_1) = \|W_0 - W_1\|_{\mathrm{F}}$. The geodesic connecting M_0 and M_1 is given by the straight line connecting W_0 and W_1, and is expressed as

$$M(\tau) = ((1-\tau)W_0 + \tau W_1)\left((1-\tau)W_0 + \tau W_1)^{\frac{1}{2}}\right)^*. \tag{2.21}$$

By applying (2.20) frequency-wise to $\boldsymbol{f}$, the Riemannian metric is given as

$$\mathrm{g}_{4,\boldsymbol{f}}(\Delta) := \int_{-\pi}^{\pi} \mathrm{tr}(\Delta X) d\theta,\ \text{where } \frac{1}{2}(\boldsymbol{f}X + X\boldsymbol{f}) = \Delta. \tag{2.22}$$

The geodesic induced by $\mathrm{g}_{4,\boldsymbol{f}}$ is

$$\boldsymbol{f}(\tau) = \left((1-\tau)\boldsymbol{f}_0^{\frac{1}{2}} + \tau \boldsymbol{f}_0^{-\frac{1}{2}}(\boldsymbol{f}_0^{\frac{1}{2}} \boldsymbol{f}_1 \boldsymbol{f}_0^{\frac{1}{2}})^{\frac{1}{2}}\right)\left((1-\tau)\boldsymbol{f}_0^{\frac{1}{2}} + \tau \boldsymbol{f}_0^{-\frac{1}{2}}(\boldsymbol{f}_0^{\frac{1}{2}} \boldsymbol{f}_1 \boldsymbol{f}_0^{\frac{1}{2}})^{\frac{1}{2}}\right)^* \tag{2.23}$$

which has been used in [38] in the context of spectral deformation.

2.3 Remarks

In spectral analysis, a main interest is to identify the frequencies that have the dominant power. In antennas array, these frequencies reflect the direction of sources with respect to the receiver

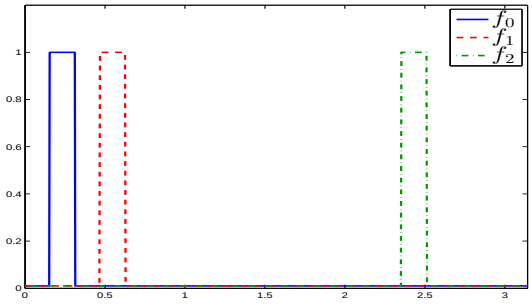

Figure 2.1: Power spectra $f_0(\theta)$, $f_1(\theta)$ and $f_2(\theta)$ for $\theta \in [0, \pi]$.

[13]. In this, there is a natural property that is desirable for the distance measures between power spectra. A natural distance should reflect the closeness of the locations of power.

Consider the three power spectra shown in Figure 2.1. We expect the distance between f_0 and f_1 to be much smaller than the distance between f_0 and f_2. The distance measures we have seen so far as well as many other distances, including the total variation distance

$$\|f_0 - f_1\|_{\mathrm{TV}} = \int |f_0 - f_1| dx,$$

have a common property that they compare the PSD's frequency-wise. For this densities shown in the figure, $\|f_0 - f_1\|_{\mathrm{TV}}$ is the same as $\|f_0 - f_2\|_{\mathrm{TV}}$. So these distances don't have the expected property.

A natural distance has the desired property is given by the Wasserstein metric that originates from the theory of OMT. It has already been used in the context of spectral analysis in [22, 44] . In the next chapter, we introduce some background materials on the theory of OMT and the Wasserstein metric.

Chapter 3

Optimal mass transport

We introduce some background materials on the Monge-Kantorovich optimal mass transport problem and on the Wasserstein metric for comparing scalar-valued probability density functions. The 2-Wasserstein distance between Gaussian distributions is discussed in detail. We also present a survey the generalizations of the OMT distance for density functions with unbalanced masses. Finally, we discuss the structure of the optimal transference plan for densities on a circle.

3.1 The Monge-Kantorovich problem

Consider a pile of sand on the ground. Assume that we are given a task to move the sand to fill up a hole on the ground, and assume that the hole has equal volume as the sand. Since it costs energy to move mass around, it is natural to seek for the optimal strategy that costs the minimum amount of energy. The optimal solution for this mass allocation problem is considered in the optimal mass transportation (OMT) theory.

We assume that the total mass is normalized to one, and model the pile and the hole by two density functions f_0 and f_1 on some space X and Y, respectively. Let the amount of energy that is needed to transport one ton of sand from location $x \in X$ to $y \in Y$ be $c(x,y)$. The original formulation of mass transportation problem is to find a mass-preserving map $t(x) : X \to Y$ that minimizes

$$\inf_t \int_X c(x,t(x)) f_0(x) dx \tag{3.1}$$

and satisfies that

$$f_0(x) = f_1(t(x))|\det \nabla t(x)|.$$

This problem was formulated by Monge in 1781 [45]. In this, there is a strong requirement that no mass splits, i.e. all the mass at x is transferred to $t(x)$. However, there may not be admissible solutions, and examples can be easily found when f_0 and f_1 are discrete distributions.

A more general formulation is proposed by Kantorovich in 1941 [46] which allows the split of mass. Let $m(x,y)$ denote the mass that is transferred from x to y. Since all the mass moved out from the point x equals to $f_0(x)$ and all the mass that is transported to y equals to $f_1(y)$, then $m(x,y)$ satisfies that

$$\int_Y m(x,y)dy = f_0(x), \ \int_X m(x,y)dx = f_1(y). \tag{3.2}$$

Thus $m(x,y)$ is thought as a density function on the product space $X \times Y$ with marginals on X and Y given by f_0 and f_1, respectively. The joint density function $m(x,y)$ is usually referred to as the transference plan. Let

$$M(f_0,f_1) := \{m \mid m \geq 0, \text{ (3.2) holds}\}$$

denote the set of all admissible transference plans. The optimal transportation cost is computed as

$$T_c(f_0,f_1) = \min_{m \in M(f_0,f_1)} \int_{X\times Y} c(x,y)m(x,y)dxdy. \tag{3.3}$$

The optimization problem in (3.3) is usually referred to as the Monge-Kantorovich problem see e.g. [12, page 19]. The optimal solution in (3.3) is called the optimal transference plan.

The Monge-Kantorovich problem also has a probabilistic interpretation. Consider two random variables $\mathbf{x} \in X$ and $\mathbf{y} \in Y$ which are distributed according to f_0 and f_1, respectively. The minimum $\mathscr{E}(c(\mathbf{x},\mathbf{y}))$ over all joint density functions of $(\mathbf{x},\mathbf{y})$ is exactly the optimal transportation cost $T_c(f_0,f_1)$.

The Monge-Kantorovich problem (3.3) admits a dual formulation

$$T_c(f_0,f_1) = \sup_{\phi,\psi} \left\{ \int_X \phi(x)f_0(x)dx + \int_Y \psi(y)f_1(y)dy \mid \phi(x)+\psi(y) \leq c(x,y) \right\} \tag{3.4}$$

see e.g. [12, page 19]. An interesting informal interpretation of the dual problem is provided in [12] and is explained next. Consider the problem of transporting sand given in the first paragraph. If we transport the mass by ourselves, then it costs us $c(x,y)$ for moving one ton of sand from x to y. Suppose there is a moving company that offers trucks to delivery the mass

and we don't need to worry about the transference plan. The price that we need to pay for the delivery is $\phi(x)$ per ton of mass that is loaded at x and $\psi(y)$ per ton of mass that is unloaded at y. For the price to be reasonable, the sum $\phi(x)+\psi(y)$ should be not larger than $c(x,y)$. Otherwise, we will handle the shipping by ourselves. The Kantorovich duality (3.4) tells that if the moving company decides the prices cleverly, we will pay them as much as the amount we would have spent if we handle the transportation by ourselves.

3.2 The Wasserstein metric

Consider probability density functions have support on a compact subset $X \subset \mathbb{R}^n$. Let the cost function $c(x,y)$ in (3.3) be of the form $c(x,y) = \|x-y\|^p$ for $p > 0$. We denote the corresponding optimal transportation cost as $T_p(f_0, f_1)$ with a subscript p for the exponent in the cost function. Then

$$W_p := T_p^{\min(1,\frac{1}{p})}$$

defines a metric on the set of probability density functions [12]. The metric is usually referred to as the *p-Wasserstein metric*.

An interesting property of the Wasserstein metric is that it metrizes weak convergence of probability measures, see e.g. [12]. We use the same notation $W_p(\mu_0, \mu_1)$ when we compare probability measures. Consider a probability measure μ and a sequence of measures $\{\mu_k : k \in \mathbb{N}\}$ on X. The weak-continuouity property of W_p implies that $W_p(\mu_k, \mu)$ converges to zero if and only if for any given continuous and bounded function ϕ

$$\int_X \phi d\mu_k \underset{k\to\infty}{\longrightarrow} \int_X \phi d\mu.$$

The weak convergence of probability measures is truly a weak* convergence [47, page 68]. However, we use the terminology "weak convergence" throughout as is common in probability theory.

A basic problem in optimal mass transport is to identify the geometry of the optimal transference plan. In the case of a quadratic cost function, i.e. $c(x,y) = \|x-y\|^2$, the transference plan can be characterized very neatly. A fundamental result shows that a transference plan is optimal if and only if it is concentrated on the sub-differential of a convex function. A transference plan m is the unique optimal solution in (3.3) with a quadratic cost if and only if it only has support on a thin subset $\{(x, \nabla\phi(x)) : x \in X\}$ with $\nabla\phi$ the unique gradient of a convex function.

This result is due to Knott and Smith [48] and Brenier [49]. In this case, $\nabla\phi(x)$ is the unique solution $t(x)$ to the Monge's problem (3.1).

The 2-Wasserstein distance has a simple expression when f_0 and f_1 only have support on the real line. Let F_0 and F_1 be the cumulative distribution functions of f_0 and f_1, respectively. Then

$$W_2(f_0,f_1) = \left(\int_0^1 |F_0^{-1}(t) - F_1^{-1}(t)|^2 dt\right)^{\frac{1}{2}}.$$

The optimal map in Monge's problem is given by $t(x) = F_1^{-1}(F_0(x))$ which satisfies that

$$\int_{-\infty}^{x} f_0(x)dx = \int_{-\infty}^{t(x)} f_1(y)dy. \tag{3.5}$$

The optimal transference plan in situations when $c(x,y) = \|x-y\|^p$ with $0 < p < 1$ behaves very differently from the solution when $p > 1$. First, the common mass between f_0 and f_1 has to stay in place. Secondly, the optimal mapping tends to have reverse orientation from the map induced by a convex cost, see e.g. [50]. If $0 < p \leq 1$, then the cost function $c(x,y) = \|x-y\|^p$ is a metric, i.e. it satisfies the triangular inequality. In this case, the Monge-Kantorovich problem (3.3) is equivalent to

$$W_1(f_0,f_1) = \min_m \left\{ \int c(x,y)m(x,y)dxdy \mid m(x,y) \geq 0, \int m(x,y) - m(y,x)dy = f_0(x) - f_1(x) \right\} \tag{3.6}$$

which is referred to as the Kantorovich-Rubinstein transshipment problem. Its dual expression is

$$W_p(f_0,f_1) = \sup_\phi \left\{ \int_X \phi(x)(f_0(x) - f_1(x))dx \;\mid\; \phi(x) - \phi(y) \leq c(x,y) \right\}. \tag{3.7}$$

We notice that the distance only depends on the difference $f_0 - f_1$.

3.2.1 Distance between Gaussian distributions

Consider two random variables $\mathbf{x}$ and $\mathbf{y}$ taking values in $\mathbb{R}^d$. Let f_0 and f_1 denote the probability density functions. Then

$$W_2(f_0,f_1)^2 = \inf_m \left\{ \mathscr{E}(|\mathbf{x}-\mathbf{y}|^2) \;\mid\; m(x,y) \geq 0, \int m(x,y)dy = f_0(x), \int m(x,y)dx = f_1(y) \right\} \tag{3.8}$$

where the expectation $\mathscr{E}$ is with respect to the joint probability density function m. Assume $\mathbf{x}$ and $\mathbf{y}$ are zero-mean Gaussian random variables with covariance R_0 and R_1, respectively. Let

$$C := \mathscr{E}(\mathbf{x}\mathbf{y}'),$$

then

$$\mathscr{E}(\|\mathbf{x}-\mathbf{y}\|^2) = \mathrm{tr}(R_0+R_1-C-C')$$

and

$$W_2(f_0,f_1)^2 = \min_C \left\{ \mathrm{tr}(R_0+R_1-C-C') \;\middle|\; \begin{bmatrix} R_0 & C \\ C' & R_1 \end{bmatrix} \geq 0 \right\}. \tag{3.9}$$

A closed form solution is given as

$$C = R_1^{-\frac{1}{2}}(R_1^{\frac{1}{2}}R_0R_1^{\frac{1}{2}})^{\frac{1}{2}}R_1^{\frac{1}{2}} \tag{3.10}$$

and $W_2(f_0,f_1)$ is given by

$$W_2(f_0,f_1) = \mathrm{tr}(R_0+R_1-2(R_1^{\frac{1}{2}}R_0R_1^{\frac{1}{2}})^{\frac{1}{2}})^{\frac{1}{2}} \tag{3.11}$$

see e.g. [51, 38]. By comparing this formula with the Bures distance $\mathrm{d}_{\mathrm{Bures}}$ in (2.12), we readily have the following proposition.

Proposition 1. *For f_0 and f_1 Gaussian zero mean distributions with covariances R_0 and R_1, respectively,*

$$W_2(f_0,f_1) = \mathrm{d}_{Bures}(R_0,R_1).$$

Proof. We now show that given $R_0, R_1 > 0$, the solution of (3.9) has an explicit closed-form expression given in (3.10).

Consider the Schur complement

$$P := R_0 - CR_1^{-1}C'.$$

This is clearly nonnegative definite. Then,

$$CR_1^{-\frac{1}{2}} = (R_0-P)^{\frac{1}{2}}U, \tag{3.12}$$

where $UU' = I$, and therefore

$$C = (R_0-P)^{\frac{1}{2}}UR_1^{\frac{1}{2}}. \tag{3.13}$$

Moreover,

$$\begin{aligned} \mathrm{tr}(C) &= \mathrm{tr}((R_0-P)^{\frac{1}{2}}UR_1^{\frac{1}{2}}) \\ &= \mathrm{tr}(R_1^{\frac{1}{2}}(R_0-P)^{\frac{1}{2}}U). \end{aligned} \tag{3.14}$$

Since R_0 and R_1 are given, minimizing $\mathrm{tr}(R_0+R_1-C-C')$ is equivalent to maximizing $\mathrm{tr}(C)$. Let $U_C\Lambda_C V_C'$ be the singular value decomposition of $R_1^{\frac{1}{2}}(R_0-P)^{\frac{1}{2}}$, and

$$U_0 := \arg\max_U \{\mathrm{tr}(R_1^{\frac{1}{2}}(R_0-P)^{\frac{1}{2}}U) \mid UU' = I\}.$$

Then, U_0 must satisfy $V_C'U_0 = U_C'$ and

$$R_1^{\frac{1}{2}}(R_0-P)^{\frac{1}{2}}U_0 = (R_1^{\frac{1}{2}}(R_0-P)R_1^{\frac{1}{2}})^{\frac{1}{2}}. \tag{3.15}$$

From (3.14) we have $\mathrm{tr}(C) = \mathrm{tr}((R_1^{\frac{1}{2}}(R_0-P)R_1^{\frac{1}{2}})^{\frac{1}{2}})$. Since $P \geq 0$, the $\mathrm{tr}(C)$ is maximal when $P = 0$.

Thus, setting $P = 0$ into (3.15), we have

$$U_0 = R_0^{-\frac{1}{2}}R_1^{-\frac{1}{2}}(R_1^{\frac{1}{2}}R_0R_1^{\frac{1}{2}})^{\frac{1}{2}}.$$

By substituting the expression of U_0 and $P = 0$ into (3.13), we obtain the optimal C given in (3.10). We note that the optimal C can also be written as

$$C = R_0^{\frac{1}{2}}(R_0^{\frac{1}{2}}R_1R_0)^{\frac{1}{2}}R_0^{-\frac{1}{2}}.$$

□

The distance $\mathrm{d}_{\mathrm{Bures}}$ brings to us a natural tool to quantify statistical errors in covariance matrices since it reflects the 2-Wasserstein metric between the corresponding Gaussian distributions. We note that for a given $\varepsilon \geq$ and a covariance matrix R_0, the set of covariances

$$\{R \mid \mathrm{d}_{\mathrm{Bures}}(R_0,R) \leq \varepsilon\}$$

is convex and it is equivalent to

$$\left\{R \mid \exists\, C \text{ such that } \mathrm{tr}(R_0+R-C-C') \leq \varepsilon^2 \text{ and } \begin{bmatrix} R_0 & C \\ C' & R \end{bmatrix} \geq 0\right\}.$$

This expression has been used in [52] to quantify errors in covariance matrices. The distance $\mathrm{d}_{\mathrm{Bures}}$ also brings a computational efficient tool for covariance approximation problem. Suppose that we are looking for a covariance matrix R in a set of covariance matrices $\mathfrak{S}$ that is close to a sample covariance R_0. It is natural to consider a covariance approximation problem of the form

$$\min_{R\in\mathfrak{S}} \mathrm{d}_{\mathrm{Bures}}(R_0,R).$$

From the LMI problem (3.9), this is equivalent to

$$\min_{R\in\mathfrak{S},C}\left\{\mathrm{tr}(R_0+R-C-C') \;\mid\; \begin{bmatrix} R_0 & C \\ C' & R \end{bmatrix}\geq 0\right\}.$$

This has been used in [38, 53] in the context of Toeplitz structured covariance approximation problem.

3.3 A Riemannian structure

The 2-Wasserstein metric leads to a Riemannian metric between probability density functions which was first introduced by Benamou and Brenier in [54]. Let $f(t,x)$ be a time-varying density function with $t\in[0,1]$ and $x\in\mathbb{R}^d$. Consider the continuity equations

$$\partial_t f+\mathrm{div}(f\boldsymbol{v})=0 \tag{3.16a}$$

where $\boldsymbol{v}$ denotes a velocity field and $\partial_t f$ corresponds to a tangent vector at f. If the initial and final conditions are given as

$$f(0,x)=f_0(x) \;\text{ and }\; f(1,x)=f_1(x) \tag{3.16b}$$

then it was shown in [54] that the optimal transportation cost can also be expressed as

$$T_2(f_0,f_1)=\inf_{\boldsymbol{v}}\left\{\int_0^1\int_{\mathbb{R}^d} f(t,x)|\boldsymbol{v}(t,x)|^2dxdt \mid f(t,x) \text{ satisfies (3.16a) and (3.16b)}\right\}. \tag{3.17}$$

The optimal velocity field satisfies that

$$\boldsymbol{v}(t,x)=\nabla\phi(t,x)$$

where ϕ is a potential function. For a tangent vector at $\partial_t f$, there is a unique $\nabla\phi$ such that

$$\partial_t f+\mathrm{div}(f\nabla\phi)=0. \tag{3.18a}$$

The optimal potential function $\phi(t,x)$ that makes $f(t,x)$ be a minimizer (3.17) satisfies the Hamilton-Jacobi equation

$$\partial_t\phi+\frac{1}{2}|\nabla\phi|^2=0. \tag{3.18b}$$

The equations (3.18a) and (3.18b) determine the geodesics induced by the 2-Wasserstein metric.

The Riemannian structure induced by the 2-Wasserstein metric is based on equations (3.17) and (3.18a). Let $\dot{f}$ be a tangent vector at f and let $\nabla\phi$ satisfy

$$\dot{f} = -\operatorname{div}(f\nabla\phi).$$

Then the Riemannian metric is defined as

$$g_f(\dot{f},\dot{f}) := \int f|\nabla\phi|^2 dx. \tag{3.19}$$

If f only has support on $\mathbb{R}$, the Riemannian metric is explicitly given as

$$g_f(\dot{f},\dot{f}) = \int_{\mathbb{R}} \frac{\dot{F}^2}{f} dx \tag{3.20}$$

where

$$\dot{F}(x) = \int_{-\infty}^{x} \dot{f}(y)dy$$

is a tangent vector at the cumulative density function F, see e.g. [22].

The geodesic f_τ for $\tau \in [0,1]$ between the two end-points f_0 and f_0 is determined by the gradient flow of the potential function [12, page 252]. For the special case when f_0 and f_1 are densities on the real line, the geodesic is computed via

$$F_\tau((1-\tau)x + \tau t(x)) = F_0(x) \tag{3.21}$$

where $t(x)$ is computed via (3.5) and F_τ is the cumulative function of f_τ.

The geodesics for Gaussian distributions induced by the 2-Wasserstein metric can also be computed explicitly. Consider two zero-mean Gaussian distribution functions f_0 and f_1 with covariances R_0 and R_1, respectively. The probability density functions $f(\tau)$ on the geodesic between f_0 and f_1 induced by the 2-Wasserstein metric are also zero-mean Gaussian distributions. The corresponding covariances are given by

$$R(\tau) = \left((1-\tau)R_0^{\frac{1}{2}} + \tau R_0^{-\frac{1}{2}}(R_0^{\frac{1}{2}}R_1R_0^{\frac{1}{2}})^{\frac{1}{2}}\right)\left((1-\tau)R_0^{\frac{1}{2}} + \tau R_0^{-\frac{1}{2}}(R_0^{\frac{1}{2}}R_1R_0^{\frac{1}{2}})^{\frac{1}{2}}\right)' \tag{3.22}$$

for $\tau \in [0,1]$ which is exactly the geodesic between the covariance matrices induced by $\mathrm{d}_{\mathrm{Bures}}$ in (2.21). In the situation when f_0 and f_1 are not zero-mean Gaussian distributions, the corresponding $f(\tau)$ is still a Gaussian distribution function with the covariance matrix given in (3.22) and the mean is given by $(1-\tau)\mathscr{E}(\mathbf{x}) + \tau\mathscr{E}(\mathbf{y})$, see e.g. [55].

3.4 The gradient flows with respect to the 2-Wasserstein metric

A remarkable result in optimal mass transportation theory is that many differential equations for the evolution of probability density functions can be viewed as the gradient flow with respect to the 2-Wasserstein metric [56, 57]. This was pointed out by Jordan, Kinderlehrer and Otto in [56]. The derivation of this result is based on the Riemannian metric introduced in the previous section.

Consider a probability density function f on $\mathbb{R}^d$. Let $\dot{f}$ be a tangent vector at f. Then the gradient direction of a functional $G(f)$ induced by the 2-Wasserstein metric, denoted by $\text{grad}_W G(f)$, is defined by the equation

$$\frac{dG(f+t\dot{f})}{dt}\,|_{t=0} = g_f(\text{grad}_W G(f), \dot{f})$$

where $g_f(\cdot,\cdot)$ denotes the Riemannian metric induced by the 2-Wasserstein metric in (3.19). Both $\text{grad}_W G(f)$ and $\dot{f}$ are tangent vectors at f. There are two potential functions ϕ and ψ that satisfy

$$\begin{aligned} \text{grad}_W G(f) &= -\text{div}(f\nabla\phi)\,, \\ \dot{f} &= -\text{div}(f\nabla\psi)\,. \end{aligned}$$

Following (3.19), we have

$$\begin{aligned} g_f(\text{grad}_W G(f), \dot{f}) &= \int (\nabla\phi\cdot\nabla\psi) f dx \\ &= \int -\phi\,\text{div}(\nabla\psi) dx \\ &= \int \phi\dot{f} dx \\ &= \int \frac{\delta G}{\delta f}\dot{f} dx \end{aligned}$$

where we have used integration by parts. Then we have $\phi = \frac{\delta G}{\delta f}$ and

$$\text{grad}_W G(f) = -\text{div}(f\nabla\frac{\delta G}{\delta f}) \tag{3.23}$$

see also [12, page 251]. If $\dot{f}$ satisfies that

$$\dot{f} = \text{grad}_W G(f) = -\text{div}(f\nabla\frac{\delta G}{\delta f}),$$

then it is a gradient direction of G with respect to the 2-Wasserstein metric. In particular, the heat equation

$$\dot{f} = f_{xx}$$

is obtain with $G(f) = -\int f \log f$.

We note that a noncommutative analogy the 2-Wasserstein metric has been recently introduced by Carlen and Maas in [58]. The probability density functions are generalized to positive operators in a Clifford algebra with trace one. The gradient flows induced by the generalized metric are also derived.

3.5 Generalizations of the Wasserstein metric

We study the generalizations of the Wasserstein metric in two different aspects. In the first part, we provide a survey of the methods that have been used to compare density functions that don't have equal integrals. Some the methods will be used later to introduce transportation distance between matrix-valued density functions. In the second part, we focus on the OMT problem between density functions on a circle and we discuss the structure of the optimal transference plan.

3.5.1 Densities with non-equal masses

Consider two density functions f_0 and f_1 that have non-equal mass, i.e. $\int_X f_0 dx \neq \int_X f_1 dx$. The Wasserstein metric is no longer suitable to compare f_0 and f_1. A distance suggested in [44] is based on a mixture of optimal transportation cost and total variation distance. For $\kappa > 0$, the proposed distance is

$$T_{c,\kappa}(f_0, f_1) := \inf_{\tilde{f}_0, \tilde{f}_1} T_c(\tilde{f}_0, \tilde{f}_1) + \kappa \sum_{k=0}^{1} \|f_k - \tilde{f}_k\|_{TV} \tag{3.24}$$

where $\tilde{f}_k$'s for $k = 0, 1$ are considered as the noise-free density functions that have equal mass while the difference between f_k's and $\tilde{f}_k$'s are caused by additive noise. For $c(x,y) = \|x-y\|^p$ with $p > 0$, it was shown in [44] that

$$T_{c,\kappa}^{\min(1,\frac{1}{p})}$$

also leads to a metric that has the weak-continuity property. The optimization problem (3.24) can be interpreted as a transportation problem on the set $X \cup \infty$ and a mass is added at ∞ as

needed so that the measures have equal mass. The corresponding cost function is modified so that

$$\hat{c}(x,y) = \begin{cases} \min(c(x,y), 2\kappa) & \text{for } x,y \in X, \\ \kappa & \text{for } x \in X,\ y = \infty, \\ \kappa & \text{for } y \in X,\ x = \infty, \\ 0 & \text{for } x = \infty,\ y = \infty. \end{cases}$$

In the case when $p = 1$, the dual of (3.24) has a neat formulation given as follows

$$\max_{\substack{\|\phi\| \le \kappa \\ \|\phi\|_{\text{Lip}} \le 1}} \int \phi(x)(f_0(x) - f_1(x))dx \tag{3.25}$$

where

$$\|\phi\|_{\text{Lip}} := \sup_{x,y} \frac{|f(x) - f(y)|}{\|x - y\|}$$

is the Lipschitz semi-norm.

Several other methods have also been proposed to generalize the Wasserstein metric for unbalanced masses. In [59], Benamou proposed a distance which is a mixtures of the transportation cost and the quadratic cost as

$$\inf_{\tilde{f}_1} \left\{ T_2(f_0, \tilde{f}_1) + \frac{\gamma}{2} \|f_1 - \tilde{f}_1\|^2 \right\}$$

where γ is a positive parameter which is used to adjust the relative significance of the two parts. This distance was recently used in the context of variational data assimilation to quantify the error in states in [60]. Figalli [61] studies the transportation between a portion of two distributions with non-equal mass. Recently, Piccoli and Rossi [62] proposes a very similar formulation as (3.24) and they consider the problem

$$\inf_{\tilde{f}_0, \tilde{f}_1} aW_p(\tilde{f}_0, \tilde{f}_1) + b \sum_{k=0}^{1} \|f_k - \tilde{f}_k\|_{TV}$$

with $a, b > 0$ and $p > 1$.

3.5.2 Mass transport on the circle

We consider a mass transportation problem between density functions on the circle and study the property of the optimal transference plan. For simplicity, the density functions f_0, f_1 are

assumed to be continuous and positive. In the classical OMT problem, if f_0 and f_1 have support on the interval $[0,1] \subset \mathbb{R}$ and if the transportation cost is given by the usual quadratic cost function, then the optimal map is monotonically increasing which is determined by the equation (3.5). The points 0 and 1 in the interval $[0,1]$ are two fixed points of this optimal map, i.e. the mass at 0 and 1 stay in place. This is no longer true for the optimal transference plan between densities on the circle.

We parameterize the circle as the segment $[0,1]$ with identified end-points. The transportation cost function is considered as

$$c(x,y) = \min\{|x-y|^2, (1-|x-y|)^2\}, \text{ for } x,y \in [0,1]. \tag{3.26}$$

The corresponding optimal transport cost is denoted as

$$\begin{aligned} T_{\text{circle}}(f_0,f_1) := \min_m \Big\{ \int_0^1 \int_0^1 c(x,y)m(x,y)dxdy \;\; | \;\; & m(x,y) \geq 0, \\ & \int_0^1 m(x,y)dy = f_0(x), \\ & \int_0^1 m(x,y)dx = f_1(y) \Big\}. \end{aligned} \tag{3.27}$$

We characterize the structure of the optimal transference plan and show that there are at least two fixed points in the optimal map. The proof of these results are based on the following lemmas.

For any $x \in \mathbb{R}$, let $[x]$ denote the element in the equivalent class $x+\mathbb{Z}$ that belongs to $[0,1)$. Let $\hat{m}(x,y)$ be the optimal solution of (3.27). For any given θ, let

$$m_\theta(x,y) := \hat{m}([x],[y]), \; \forall x,y \in [\theta, \theta + \frac{1}{2}].$$

Lemma 2. *If $m_\theta(x,y)$ has support on (x_1,y_1) and (x_2,y_2), then*

$$(x_1-x_2)(y_1-y_2) > 0.$$

Proof. Since $\hat{m}$ is the optimal solution of (3.27), m_θ is the optimal transference plan for the masses that is transported form the half circle $[\theta, \theta+\frac{1}{2}]$ to itself. However, the transportation cost function on this half circle is the same as the cost on a straight line which is $|x-y|^2$. Thus the transference plan corresponds to a monotonically increasing function on each half circle. The lemma follows directly from Theorem 2.12 in [12, page 66]. □

Corollary 3. *Let $\hat{m}$ be the optimal solution of* (3.27) *for two continuous and positive functions f_0 and f_1. Then for any $x \in [0,1)$, $\hat{m}(x,[x+\frac{1}{2}]) = 0$.*

Proof. Assume that there exist $x \in [0,1)$ such that $\hat{m}(x,[x+\frac{1}{2}]) \neq 0$, i.e. the mass at x is transferred to $[x+\frac{1}{2}]$. So the mass at $[x+\frac{1}{2}]$ is not transferred to the half circle $[x,x+\frac{1}{2}]$ and it must be transferred to the half circle $[x+\frac{1}{2},x+1)$. But on the half circle $[x+\frac{1}{2},x+1]$, the mass on $x+1$ is transferred to $x+\frac{1}{2}$ while the mass on $x+\frac{1}{2}$ is transferred to the same interval. This implies that all the masses in f_1 are concentrated at $[x+\frac{1}{2}]$. This contradicts to the assumption that f_1 is continuous and positive. Thus the corollary is proved. □

Based on the above lemma and corollary, we are ready to show the following proposition.

Proposition 4. *Given two continuous and positive probability densities f_0, f_1 on the circle, let $\hat{m}(x,y)$ be the optimal solution of* (3.27)*. If $\hat{m}(0,\theta) \neq 0$, then $\hat{m}$ only has support on the set*

$$\{(x,[t(x)]) \mid x \in [0,1]\}$$

with $t : [0,1] \to [\theta,\theta+1]$ being a continuous and monotonic function.

Proof. Without loss of generality, we consider the case when the mass at $x = 0$ is transported to $t(0) = \theta \in [0,\frac{1}{2})$. The case when $t(0) \in (\frac{1}{2},1)$ can be cast to the former one by switching the two density functions f_0 and f_1. The situation when $t(0) = \frac{1}{2}$ is ruled out by the above corollary. From Lemma 2, the optimal map at every half circle is monotonically increasing. We need to show that this map is also continuous. It suffices to show that it is continuous at 0, i.e. for $x \in [0,\frac{1}{2})$, as $x \to 0$, $t(x) \to \theta$. The continuity at other points can be shown similarly by a changing of coordinate. We only prove that the map is upper semi-continuous at $x = 0$, the case of lower semiconscious is proved analogously by changing the coordinate from x to $1-x$.

Assume that the optimal maps is not continuous at $x = 0$, and for $x \in [0,\frac{1}{2})$, $\lim_{x\to 0} t(x) = \theta + c$ for some $c > 0$. Then from Lemma 2, $\theta + c \in (\theta,\theta+\frac{1}{2}]$. Moreover,

$$\text{for } x \in [\theta+c-\frac{1}{2},\theta+c] \cup [0,\frac{1}{2}],\ t(x) \notin (\theta,\theta+c),$$

and

$$\text{for } x \in (\theta,\theta+c),\ t(x) \notin [\theta+c-\frac{1}{2},\theta+c].$$

Otherwise, there is a contradiction to Lemma 2. So there exists at least a point $x \in [\frac{1}{2},\theta+c+\frac{1}{2}]$ such that $y := t(x) \in (\theta,\theta+c)$. If $t(y) \in [y-\frac{1}{2},y]$, then we find that the transference plan has

support on $(0,\theta)$ and $(y,t(y))$ which satisfy

$$(y-0)(t(y)-\theta)<0.$$

This contradicts to Lemma 2. So $t(y)$ has to be in the interval $(y,y+\frac{1}{2})$. Since $x\in(y,y+\frac{1}{2})$ and $t(x)=y$, $t(y)$ cannot be in $(y,y+\frac{1}{2})$ otherwise there is also a contradiction to Lemma 2. Thus the assumption that the optimal map t is not continuous at $x=0$ fails and the proof is completed. □

From Proposition 4 the optimal transference plan is represented by a shifted monotonically increasing function. We note that this result is also pointed out in [63] but with a different approach. For the analysis of the structure of optimal transference plan, it is convenient to lift the densities f_0 and f_1 from the circle to the real line and consider them as periodic functions, i.e. consider f_0 and f_1 as periodic functions on $\mathbb{R}$ with period 1. For $x\in[0,1]$, the values of f_0 and f_1 are the same as the values on the circle. From Proposition 4, the transference plan monotonically maps the mass of f_0 on $[0,1]$ to f_1 on $[\theta,\theta+1]$. Let F_0 and F_1 be the cumulative functions of f_0 and f_1, respectively, with $F_0(0)=F_1(0)=0$. Then, from the monotonicity of t, it follows that $F_0(x)=F_1(t(x))-F_1(\theta)$. Thus

$$t(x)=F_1^{-1}(F_0(x)+F_1(\theta)) \tag{3.28}$$

which is a strictly increasing function of x. To simplify the notations, we denote by $\tilde{F}_0$ and $\tilde{F}_1$ the inverse functions F_0^{-1} and F_1^{-1}, respectively. Then the optimal transport cost

$$\begin{aligned} T_{\text{circle}}(f_0,f_1) &= \int_0^1 f_0(x)|t(x)-x|^2dx && \text{(3.29a)}\\ &= \int_0^1 |\tilde{F}_0(v)-\tilde{F}_1(v+F_1(\theta))|^2dv && \text{(3.29b)} \end{aligned}$$

where (3.29b) is obtained via a changing of variable with $v=F_0(x)$.

Proposition 5. *For the optimal map $t:[0,1]\to[\theta,\theta+1]$ in Proposition 4, there are at least two points $x_1,x_2\in[0,\ 1]$ such that $t(x_i)=x_i$, for $i=1,2$.*

Proof. For $\omega\in[0,1]$, let

$$t_\omega(x)=\tilde{F}_0(F_0(x)+F_1(\omega)). \tag{3.30}$$

Then the optimal map $t(x)$ is equal to $t_\omega(x)$ when $\omega = \theta$. Then for any fixed x, $t_\omega(x)$ is strictly increasing with respect to ω. Let

$$T_{\omega,\text{circle}}(f_0, f_1) := \int_0^1 f_0(x)|t_\omega(x) - x|^2 dx$$

and we consider $T_{\omega,\text{circle}}(f_0, f_1)$ as a function of ω. Then from the optimality of the $t(x)$ with $t(0) = \theta$ it holds that

$$\theta = \arg\min_\omega T_{\omega,\text{circle}}(f_0, f_1). \tag{3.31}$$

The first order derivative of $T_{\omega,\text{circle}}(f_0, f_1)$ satisfies that

$$\frac{dT_{\omega,\text{circle}}}{d\omega} |_{\omega=\theta} = 2\int_0^1 f_0(x)(t(x) - x)\frac{dt_\omega(x)}{d\omega} |_{\omega=\theta}\, dx = 0 \tag{3.32}$$

where $\frac{dt_\omega(x)}{d\omega} > 0$. If there is no x such that $t(x) = x$, and from the continuity and monotonicity of $t(x)$ then either $t(x) - x > 0$ for all $x \in [0,1]$ or $t(x) - x < 0$ for all $x \in [0,1]$. In both cases, (3.32) is not equal to zero. For (3.32) holds, from the continuity and monotonicity of $t(x)$, we need at least two fixed points x_i such that $t(x_i) = x_i$. □

We illustrate the result in Proposition 4 and Proposition 5 based on the following two probability density functions f_0 f_1 that are supported on $[0,1]$ shown in Figure 3.1. The optimal

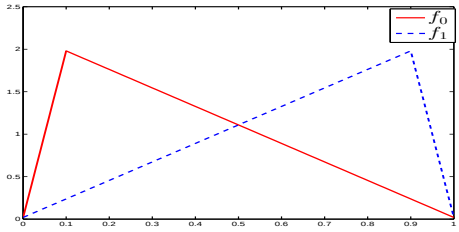

Figure 3.1: Two probability density functions f_0, f_1 supported on $[0,1]$.

transference plan between f_0 and f_1 with the circular cost function (3.26) is shown in Figure 3.2. The transference plan is a two dimensional density function having support on a thin set which is shown as the gray curve. The value of the density function is illustrated in the gray scale colors. The two points which are the intersection between the transference plan and the line $y = x$ are the two fixed points in the transference plan. For comparison, we consider the

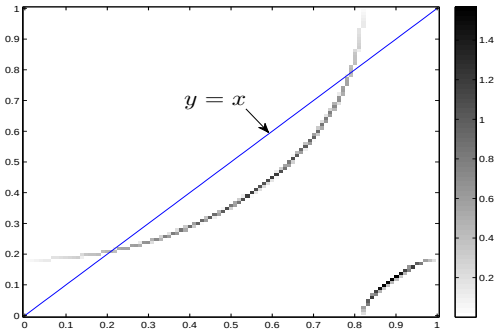

Figure 3.2: The optimal transference plan between f_0, f_1 with the circular cost function (3.26). The intersections of the line $y = x$ and the transference plan corresponds to two fixed points.

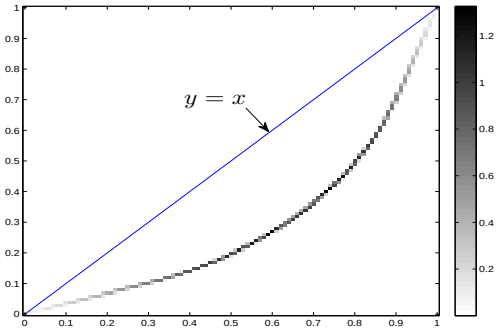

Figure 3.3: The optimal transference plan between f_0, f_1 with the quadratic cost function $c(x,y) = |x-y|^2$. The two fixed points of the transference plan are $x = 0, 1$.

optimal transference plan associated with the quadratic cost function $c(x,y) = |x-y|^2$ which is shown in Figure 3.3. We note that the transference plan shown in Figure 3.3 corresponds to a monotonic increasing function and the transference plan in Figure 3.2 corresponds to a shifted and circular function.

3.6 Remarks

This chapter presents certain preliminary materials on the Monge-Kantorovich OMT problem and the Wasserstein metric. The 2-Wasserstein metric between zero-mean Gaussian distributions is studied in detail. We point out the 2-Wasserstein metric between two zero-mean

Gaussian distributions coincides with the Bures distance between the corresponding covariance matrices. Several methods on generalizing the Monge-Kantorovich problem for non-balanced masses are reviewed. Some of these methods will be used later in the next chapter to formulate a distance between matrix-valued density functions. As an extension, the optimal mass transportation problem between densities on the circle is studied. It is shown that the optimal transference plan always has at least two fixed points.

Chapter 4

OMT for matrix-valued densities with equal integral

In this chapter, we introduce a generalization of the Monge-Kantorovich problem to compare matrix-valued densities. We start by considering matrix-valued density functions in the set

$$\mathscr{S} := \left\{ \boldsymbol{f} \;\middle|\; \boldsymbol{f} \text{ has support on } [0,1], \boldsymbol{f}(x) \geq 0, \text{ and } \int_0^1 \boldsymbol{f}(x)dx = I_{\ell\times\ell} \right\}$$

The elements in $\mathscr{S}$ are positive semi-definite matrix valued functions that have the same integral which is the identity matrix. We introduce a formulation of optimal mass transportation for comparing $\boldsymbol{f}_0, \boldsymbol{f}_1 \in \mathscr{S}$ which leads to a generalization of the Kantorovich-Rubinstein problem. Obviously, the assumption that the matrix-valued density functions have equal integral is too strong. We then relax this assumption and present an alternative distance that is suitable for density functions with not necessarily balanced masses and is weakly continuous.

4.1 Matrix-valued Kantorovich-Rubinstein problem

We begin with an observation on the Kantorovich-Rubinstein problem for scalar-valued density functions. Consider two scalar-valued density functions f_0 and f_1 having support on $[0,1]$, the Wasserstein metric $W_p(f_0, f_1)$ corresponding to a cost function

$$c(x,y) = \|x-y\|^p \text{ with } 0 < p \leq 1$$

is defined as

$$W_p(f_0,f_1) \quad = \quad \min_m \Bigg\{ \int_0^1 \int_0^1 c(x,y)m(x,y)dxdy \;\mid\; m(x,y) \geq 0, \int_0^1 m(x,y)dy = f_0(x), \int_0^1 m(x,y)dx = f_1(y) \Bigg\}. \tag{4.1a}$$

Because the exponent $0 < p \leq 1$, the cost is a metric and

$$W_p(f_0,f_1) = \sup_\phi \left\{ \int_0^1 \phi(x)\,(f_0(x) - f_1(x))\,dx \;\mid\; \phi(x) - \phi(y) \leq c(x,y), \forall x,y \right\} \tag{4.1b}$$

where (4.1b) is referred to as the Kantorovich-Rubinstein mass transshipment problem. We observe that the positive constraint on $m(x,y)$ can be removed by using an absolute value of penalty of $m(x,y)$ in the objective function and the optimal value does not change. This is elaborated next.

Proposition 6. *Given two probability density functions f_0 and f_1 having support on $[0,1]$, let $W_p(f_0,f_1)$ denote the p-Wasserstein metric with $0 < p \leq 1$, then*

$$W_p(f_0,f_1) = \inf_m \Bigg\{ \int_0^1 \int_0^1 c(x,y)|m(x,y)|dxdy \quad \mid \quad \int_0^1 m(x,y)dy = f_0(x), \int_0^1 m(x,y)dx = f_1(x) \Bigg\} \tag{4.2}$$

Proof. Let $\hat{m}(x,y)$ denote the optimal solution of (4.2). Since the set

$$\left\{ m \mid m \geq 0, \int_0^1 m(x,y)dy = f_0(x), \int_0^1 m(x,y)dx = f_1(y) \right\}$$

is only a subset of the feasible set in (4.2), then clearly

$$W_p(f_0,f_1) \geq \int_0^1 \int_0^1 c(x,y)|\hat{m}(x,y)|dxdy.$$

We will show that the inequality also holds in the other direction. We decompose $\hat{m}$ as

$$\hat{m}(x,y) = \hat{m}_+(x,y) - \hat{m}_-(x,y)$$

with $\hat{m}_+(x,y), \hat{m}_-(x,y) \geq 0$ and $|\hat{m}(x,y)| = \hat{m}_+(x,y) + \hat{m}_-(x,y)$. Let

$$f_{0,+}(x) := \int \hat{m}_+(x,y)dy \;\text{ and }\; f_{1,+}(y) := \int \hat{m}_+(x,y)dx$$

$$f_{0,-}(x) := \int \hat{m}_-(x,y)dy \text{ and } f_{1,-}(y) := \int \hat{m}_-(x,y)dx.$$

Then $f_{0,+}(x) - f_{0,-}(x) = f_0(x)$ and $f_{1,+}(x) - f_{1,-}(x) = f_1(x)$. Thus $\hat{m}_+$ and $\hat{m}_-$ are transportation plans for $f_{0,+}$, $f_{1,+}$ and $f_{0,-}$, $f_{1,-}$, respectively. From the optimality of $\hat{m}(x,y)$,

$$\int_0^1 \int_0^1 c(x,y)|\hat{m}(x,y)|dxdy = W_p(f_{0,+}, f_{1,+}) + W_p(f_{0,-}, f_{1,-}).$$

From (3.7),

$$\begin{aligned}
& W_p(f_{0,+}, f_{1,+}) + W_p(f_{0,-}, f_{1,-}) \\
= \; & \sup_{\phi_+} \left\{ \int_0^1 \phi_+(f_{0,+} - f_{1,+})dx \;\middle|\; \phi_+(x) - \phi_+(y) \le c(x,y) \right\} \\
& + \sup_{\phi_-} \left\{ \int_0^1 \phi_-(f_{1,-} - f_{0,-})dx \;\middle|\; \phi_-(x) - \phi_-(y) \le c(x,y) \right\} \\
\ge \; & \sup_{\phi} \left\{ \int_0^1 \phi\left((f_{0,+} - f_{0,-}) - (f_{1,+} - f_{1,-})\right) \mid \phi(x) - \phi(y) \le c(x,y) \right\} \\
= \; & \sup_{\phi} \left\{ \int_0^1 \phi\left(f_0 - f_1\right) \mid \phi(x) - \phi(y) \le c(x,y) \right\} \\
= \; & W_p(f_0, f_1).
\end{aligned}$$

So the proposition is proved. □

We note that the condition that $c(x,y) = |x-y|^p$ with $0 < p \le 1$ is crucial in the lemma. If $p > 1$, the Kantorovich-Rubinstein transhipment cost vanishes [12, page 37]. We note that the main difference between (4.2) and the Kantorovich problem (3.3) is that the variable m in the above is not constrained to be nonnegative and its absolute value is penalized in the objective function. The penalty of absolute values of m introduces the possibility of considering "negative mass".

We propose an optimal mass transportation problem that generalizes (4.2) to matrix-valued density functions. A natural generalization of the absolute value for scalar values to matrices is the *nuclear norm*. For $n \times n$ matrix $\boldsymbol{m}$ with singular values $\sigma_0, \ldots, \sigma_n \ge 0$, its nuclear norm is

$$\|\boldsymbol{m}\|_* := \sum_{i=1}^{n} \sigma_i.$$

Given two matrix-valued density functions $\boldsymbol{f}_0, \boldsymbol{f}_1 \in \mathscr{S}$, we consider a formulation of matrix-valued mass transportation problem as follows

$$\mathrm{D}_W(\boldsymbol{f}_0, \boldsymbol{f}_1) = \inf_{\boldsymbol{m}} \left\{ \int_0^1 \int_0^1 c(x,y) \|\boldsymbol{m}(x,y)\|_* \;\; | \;\; \int_0^1 \boldsymbol{m}(x,y)dy = \boldsymbol{f}_0(x), \right. \\ \left. \int_0^1 \boldsymbol{m}(x,y)dx = \boldsymbol{f}_1(y) \right\} \tag{4.3}$$

where $\|\boldsymbol{m}\|_*$ denotes the nuclear norm of $\boldsymbol{m}$. The optimization problem (4.3) reduces to (4.2) when $\boldsymbol{f}_0$ and $\boldsymbol{f}_1$ are scalar-valued densities.

The constraint in the optimization problem (4.3) is always feasible. A trivial feasible solution is given as

$$\boldsymbol{m}(x,y) = \boldsymbol{f}_0(x) + \boldsymbol{f}_1(y) - I.$$

Analogous to the scalar-valued case, it is natural to ask if the optimization problem (4.3) still make sense if it is constrained in the subset

$$\left\{ \boldsymbol{m} \;\; | \;\; \boldsymbol{m}(x,y) \geq 0, \int_0^1 \boldsymbol{m}(x,y)dy = \boldsymbol{f}_0(x), \int_0^1 \boldsymbol{m}(x,y)dx = \boldsymbol{f}_1(y) \right\} \tag{4.4}$$

where all the "mass" $\boldsymbol{m}(x,y)$ is constrained to be non-negative definite. However, this set is not always feasible. To see this, consider the following two discrete matrix-valued density functions

$$\boldsymbol{f}_0(x) = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 0 \end{bmatrix} \delta(x - x_1) + \begin{bmatrix} 0 & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \delta(x - x_2), \\ \boldsymbol{f}_1(x) = \begin{bmatrix} \frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{1}{4} \end{bmatrix} \delta(x - x_1) + \begin{bmatrix} \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} \end{bmatrix} \delta(x - x_2).$$

It is easy to see that there is no feasible element in (4.4) that satisfies the given marginals since we can not split the singular matrix-valued masses.

The nuclear norm is a convex functional, thus the optimization problem in (4.3) is convex. Its dual problem has a formulation very similar to (4.1b).

Proposition 7. *Let $\boldsymbol{f}_0, \boldsymbol{f}_1 \in \mathscr{S}$, then*

$$\mathrm{D}_W(\boldsymbol{f}_0, \boldsymbol{f}_1) = \sup_{\boldsymbol{\phi}} \left\{ \int_0^1 \mathrm{tr}\Big(\boldsymbol{\phi}(\boldsymbol{f}_0 - \boldsymbol{f}_1) \Big) dx \;\; | \;\; \boldsymbol{\phi}(x) - \boldsymbol{\phi}(y) \leq c(x,y) I, \; \forall x, y \right\}. \tag{4.5}$$

Proof. For a Hermitian matrix $\boldsymbol{m}$, its nuclear norm

$$\|\boldsymbol{m}\|_* = \max_{\|\boldsymbol{w}\|\le 1} \operatorname{tr}(\boldsymbol{w}\boldsymbol{m}).$$

Let $\boldsymbol{\phi}_0$ and $\boldsymbol{\phi}_1$ be the multipliers of the constraints in (4.3), we have

$$\begin{aligned} \mathrm{D_W}(\boldsymbol{f}_0,\boldsymbol{f}_1) &= \inf_{\boldsymbol{m}} \sup_{\substack{\|\boldsymbol{w}\|\le 1\\ \boldsymbol{\phi}_0,\boldsymbol{\phi}_1}} \int_0^1\int_0^1 \operatorname{tr}\Big((\boldsymbol{w}(x,y)c(x,y)-\boldsymbol{\phi}_0(x)-\boldsymbol{\phi}_1(y))\boldsymbol{m}(x,y)\Big)dxdy \\ &\qquad + \int_0^1 \operatorname{tr}\left(\boldsymbol{\phi}_0(x)\boldsymbol{f}_0(x)\right)dx + \int_0^1 \operatorname{tr}\left(\boldsymbol{\phi}_1(y)\boldsymbol{f}_1(y)\right)dy \qquad (4.6a) \\ &= \sup_{\substack{\|\boldsymbol{w}\|\le 1\\ \boldsymbol{\phi}_0,\boldsymbol{\phi}_1}} \inf_{\boldsymbol{m}} \int_0^1\int_0^1 \operatorname{tr}\Big((\boldsymbol{w}(x,y)c(x,y)-\boldsymbol{\phi}_0(x)-\boldsymbol{\phi}_1(y))\boldsymbol{m}(x,y)\Big)dxdy \\ &\qquad + \int_0^1 \operatorname{tr}\left(\boldsymbol{\phi}_0(x)\boldsymbol{f}_0(x)\right)dx + \int_0^1 \operatorname{tr}\left(\boldsymbol{\phi}_1(y)\boldsymbol{f}_1(y)\right)dy. \qquad (4.6b) \end{aligned}$$

Then the optimal $\boldsymbol{w}$, $\boldsymbol{\phi}_0$ and $\boldsymbol{\phi}_1$ must satisfy

$$\boldsymbol{w}(x,y)c(x,y)-\boldsymbol{\phi}_0(x)-\boldsymbol{\phi}_1(y)=0$$

otherwise the first term of (4.6b) goes to $-\infty$. Thus

$$\boldsymbol{\phi}_0(x)+\boldsymbol{\phi}_1(y)=\boldsymbol{w}(x,y)c(x,y) \text{ for all } x,y\in[0,1].$$

Since $c(x,x)=0$, then we have $\boldsymbol{\phi}_1(x)=-\boldsymbol{\phi}_0(x)$. Thus

$$\boldsymbol{\phi}_0(x)-\boldsymbol{\phi}_0(y)=\boldsymbol{w}(x,y)c(x,y).$$

From the constraint $\|\boldsymbol{w}\|\le 1$, we have $\|\boldsymbol{\phi}_0(x)-\boldsymbol{\phi}_0(y)\|\le c(x,y),\forall x,y$ which is equivalent to

$$\boldsymbol{\phi}_0(x)-\boldsymbol{\phi}_0(y)\le c(x,y)I$$

where the inequality is in the sense of positive semi-definiteness. Then the dual problem (4.5) is obtained from (4.6b) with $\boldsymbol{\phi}=\boldsymbol{\phi}_0$. □

Proposition 8. $\mathrm{D}_W(\cdot,\cdot)$ *defines a metric for matrix-valued density functions in* $\mathscr{S}$.

Proof. For any $\boldsymbol{f}_0,\boldsymbol{f}_1\in\mathscr{S}$, it is obviously that $\mathrm{D_W}(\boldsymbol{f}_0,\boldsymbol{f}_1)=\mathrm{D_W}(\boldsymbol{f}_1,\boldsymbol{f}_0)\ge 0$ and $\mathrm{D_W}(\boldsymbol{f}_0,\boldsymbol{f}_1)=0$ if and only if $\boldsymbol{f}_0=\boldsymbol{f}_1$. We will show that the triangular inequality is also satisfied.

For $\boldsymbol{f}_0, \boldsymbol{f}_1, \boldsymbol{f}_2 \in \mathscr{S}$,

$$\begin{aligned}
&\mathrm{D_W}(\boldsymbol{f}_0, \boldsymbol{f}_1) + \mathrm{D_W}(\boldsymbol{f}_1, \boldsymbol{f}_2) \\
&= \sup_{\boldsymbol{\phi}} \left\{ \int_0^1 \mathrm{tr}\,(\boldsymbol{\phi}(\boldsymbol{f}_0 - \boldsymbol{f}_1))\, dx \;\mid\; \boldsymbol{\phi}(x) - \boldsymbol{\phi}(y) \le c(x,y) I \right\} \\
&\quad + \sup_{\boldsymbol{\psi}} \left\{ \int_0^1 \mathrm{tr}\,(\boldsymbol{\psi}(x)(\boldsymbol{f}_1 - \boldsymbol{f}_2))\, dx \;\mid\; \boldsymbol{\psi}(x) - \boldsymbol{\psi} \le c(x,y) I \right\} \\
&\ge \sup_{\boldsymbol{\phi}} \left\{ \int_0^1 \mathrm{tr}\,(\boldsymbol{\phi}(\boldsymbol{f}_0 - \boldsymbol{f}_1 + \boldsymbol{f}_1 - \boldsymbol{f}_2))\, dx \;\mid\; \boldsymbol{\phi}(x) - \boldsymbol{\phi}(y) \le c(x,y) I \right\} \\
&= \mathrm{D_W}(\boldsymbol{f}_0, \boldsymbol{f}_2).
\end{aligned}$$

□

Clearly, if $\boldsymbol{f}_0, \boldsymbol{f}_1$ do not have equal integral, the value in (4.5) is infinity. In the next section, we discuss a generalization of the distance $\mathrm{D_W}$ to matrix-valued densities with unbalanced masses. The idea of the generalization is adopted from [44] where we use an additional bound for $\boldsymbol{\phi}$.

4.2 Generalization for densities with non-equal "masses"

In the formulation of the distance $\mathrm{D_W}$ in (4.3), we have a very restrictive assumption that the functions having the same integral which is normalized to the identity I. If $\boldsymbol{f}_0$ and $\boldsymbol{f}_1$ are power spectra of two multivariate time-series, this assumption implies that the two time-series have the same covariance $R_0 = I$. This is unrealistic since there may be correlations between channels, and it is very likely that the covariances can not be normalized to the identity by the same congruence transform. We now extend the distance to compare density functions in the set

$$\hat{\mathscr{S}} := \{\boldsymbol{f} \;\mid\; \boldsymbol{f} \text{ has support on } [0,1], \boldsymbol{f}(x) \ge 0, \text{ and } \int_0^1 \boldsymbol{f}(x)dx \text{ is bounded}\}.$$

If $\boldsymbol{f}_0, \boldsymbol{f}_1 \in \hat{\mathscr{S}}$ do not have equal integral, the primal problem (4.3) is infeasible and the dual problem (8) will be unbounded. So we need to develop a generalized formulation of the distance $\mathrm{D_W}$. The generalization below is based on the metric introduced in [44] for scalar-valued densities which was also explained in Section 3.5.1 of Chapter 3.

For $\boldsymbol{f}_0, \boldsymbol{f}_1 \in \hat{\mathscr{S}}$ and $\kappa > 0$, we define

$$\mathrm{D_{W,\kappa}}(\boldsymbol{f}_0, \boldsymbol{f}_1) := \inf_{\tilde{\boldsymbol{f}}_0, \tilde{\boldsymbol{f}}_1} \left\{ \mathrm{D_W}(\tilde{\boldsymbol{f}}_0, \tilde{\boldsymbol{f}}_1) + \kappa \sum_{k=0}^{1} \|\boldsymbol{f}_k(x) - \tilde{\boldsymbol{f}}_k(x)\|_{\mathrm{TV}} \right\} \tag{4.7}$$

where

$$\|\boldsymbol{f}-\tilde{\boldsymbol{f}}\|_{\mathrm{TV}} := \int_0^1 \|\boldsymbol{f}(x)-\tilde{\boldsymbol{f}}(x)\|_* dx.$$

The expression $\|\boldsymbol{f}-\tilde{\boldsymbol{f}}\|_{\mathrm{TV}}$ represents a generalized total variation distance for matrix-valued density functions. In (4.7), $\tilde{\boldsymbol{f}}_0$ and $\tilde{\boldsymbol{f}}_1$ can be thought as the noise-free components and the difference $\boldsymbol{f}_k-\tilde{\boldsymbol{f}}_k$, for $k=0,1$, can be thought as the perturbation or contribution of noise. The parameter κ is adjusted for the relative significance of the "noise-free" and noise components.

The optimization problem in (4.7) is convex. Its dual problem has a similar formulation as (3.25) is given below.

Proposition 9. *The dual of* (4.7) *is*

$$\mathrm{D}_{W,\kappa}(\boldsymbol{f}_0,\boldsymbol{f}_1) = \sup_{\boldsymbol{\phi}} \left\{ \mathrm{tr}\left(\int \boldsymbol{\phi}(\boldsymbol{f}_0-\boldsymbol{f}_1)dx \right) \;\middle|\; \boldsymbol{\phi}(x)-\boldsymbol{\phi}(y) \le c(x,y)I,\ \|\boldsymbol{\phi}(x)\| \le \kappa \right\}. \quad (4.8)$$

Proof. Since $\mathrm{D}_{\mathrm{W}}(\tilde{\boldsymbol{f}}_0,\tilde{\boldsymbol{f}}_1)$ is also the solution of an optimization problem, $\mathrm{D}_{\mathrm{W},\kappa}(\boldsymbol{f}_0,\boldsymbol{f}_1)$ can be written as

$$\begin{aligned}\mathrm{D}_{\mathrm{W},\kappa}(\boldsymbol{f}_0,\boldsymbol{f}_1) = \inf_{\boldsymbol{m},\tilde{\boldsymbol{f}}_0,\tilde{\boldsymbol{f}}_1} \Bigg\{ & \int_0^1\int_0^1 c(x,y)\|\boldsymbol{m}(x,y)\|_* + \kappa \sum_{k=0}^{1} \|\boldsymbol{f}_k(x)-\tilde{\boldsymbol{f}}_k(x)\|_* dxdy \;\Big|\\ & \int_0^1 \boldsymbol{m}(x,y)dy = \tilde{f}_0(x),\ \int_0^1 \boldsymbol{m}(x,y)dx = \tilde{f}_1(y) \Bigg\}. \end{aligned} \quad (4.9)$$

We note the following dual formulation of the nuclear norm

$$\begin{aligned}\|\tilde{\boldsymbol{m}}\|_* &= \max_{\|\boldsymbol{w}\|\le 1} \mathrm{tr}(\boldsymbol{w}\boldsymbol{m})\\ \|\boldsymbol{f}_k-\tilde{\boldsymbol{f}}_k\|_* &= \max_{\|\boldsymbol{\lambda}_k\|\le 1} \mathrm{tr}(\boldsymbol{\lambda}_k(\boldsymbol{f}_k-\tilde{\boldsymbol{f}}_k)).\end{aligned}$$

Let $\boldsymbol{\phi}_0, \boldsymbol{\phi}_1$ be the multipliers for the two constraints. Then,

$$\begin{aligned}
&\mathrm{D}_{\mathrm{W},\kappa}(\boldsymbol{f}_0, \boldsymbol{f}_1) \\
&= \inf_{\substack{\boldsymbol{m},\tilde{\boldsymbol{f}}_0,\tilde{\boldsymbol{f}}_1 \\ \|\boldsymbol{w}\|,\|\boldsymbol{\lambda}_0\|,\|\boldsymbol{\lambda}_1\|\le 1}} \sup_{\boldsymbol{\phi}_0,\boldsymbol{\phi}_1} \Big\{ \mathrm{tr} \int \Big((c(x,y)\boldsymbol{w}(x,y) - \boldsymbol{\phi}_0(x) - \boldsymbol{\phi}_1(y))\, \boldsymbol{m}(x,y) + (\boldsymbol{\phi}_0(x) - \kappa\boldsymbol{\lambda}_0(x))\tilde{\boldsymbol{f}}_0(x) \\
&\qquad + (\boldsymbol{\phi}_1(y) - \kappa\boldsymbol{\lambda}_1(y))\tilde{\boldsymbol{f}}_1(y) + \kappa\boldsymbol{\lambda}_0(x)\boldsymbol{f}_0(x) + \kappa\boldsymbol{\lambda}_1(y)\boldsymbol{f}_1(y) \Big) dxdy \Big\} \qquad (4.10a) \\
&= \sup_{\substack{\boldsymbol{\phi}_0,\boldsymbol{\phi}_1 \\ \|\boldsymbol{w}\|,\|\boldsymbol{\lambda}_0\|,\|\boldsymbol{\lambda}_1\|\le 1}} \inf_{\boldsymbol{m},\tilde{\boldsymbol{f}}_0,\tilde{\boldsymbol{f}}_1} \Big\{ \mathrm{tr} \int \Big((c(x,y)\boldsymbol{w}(x,y) - \boldsymbol{\phi}_0(x) - \boldsymbol{\phi}_1(y))\, \boldsymbol{m}(x,y) + (\boldsymbol{\phi}_0(x) - \kappa\boldsymbol{\lambda}_0(x))\tilde{\boldsymbol{f}}_0(x) \\
&\qquad + (\boldsymbol{\phi}_1(y) - \kappa\boldsymbol{\lambda}_1(y))\tilde{\boldsymbol{f}}_1(y) + \kappa\boldsymbol{\lambda}_0(x)\boldsymbol{f}_0(x) + \kappa\boldsymbol{\lambda}_1(y)\boldsymbol{f}_1(y) \Big) dxdy \Big\}. \qquad (4.10b)
\end{aligned}$$

The optimal $\boldsymbol{w}, \boldsymbol{\phi}_k, \boldsymbol{\lambda}_k$ must satisfy that

$$\begin{aligned}
c(x,y)\boldsymbol{w}(x,y) - \boldsymbol{\phi}_0(x) - \boldsymbol{\phi}_1(y) &= 0 \\
\boldsymbol{\phi}_0(x) - \kappa\boldsymbol{\lambda}_0(x) &= 0 \\
\boldsymbol{\phi}_1(y) - \kappa\boldsymbol{\lambda}_1(y) &= 0
\end{aligned}$$

which imply that $\boldsymbol{\phi}_0 = -\boldsymbol{\phi}_1 =: \boldsymbol{\phi}$ and

$$\boldsymbol{\phi}(x) - \boldsymbol{\phi}(y) \le c(x,y)I, \text{ and } \|\boldsymbol{\phi}\| \le \kappa.$$

By substituting these conditions to (4.10b), we obtain (4.8). □

The transportation cost $\mathrm{D}_{\mathrm{W},\kappa}$ actually defines a metric between matrix-valued density functions on the set $\hat{\mathscr{S}}$. The proof can be carried out in a similar manner as Proposition 8. Moreover, the metric $\mathrm{D}_{\mathrm{W},\kappa}$ also inherits the weak-continuity property of its scalar-valued counter part which is explained next.

4.3 On weak-continuity of the new metric

An interesting property of the Wasserstein metric is that it metrizes weak convergence of probability measures. Thus, small changes in the Wasserstein metric reflect in small changes of any

statistics of the corresponding measures. This property is inherited by the metric introduced in [44] for comparing distributions with non-equal mass (c.f. Chapter 3). In the present section, we show that a similar property also holds for $\mathrm{D}_{\mathrm{W},\kappa}$. To explain this property, we will use the notation $\mathrm{D}_{\mathrm{W},\kappa}(\boldsymbol{\mu}_0,\boldsymbol{\mu})$ when we compare matrix-valued measures.

Let us denote $C_b(\mathbb{H}^{\ell\times\ell},[0,1])$ as the set of continuous and bounded Hermitian matrix-valued functions that have support on $[0,1]$. We show that convergence in $\mathrm{D}_{\mathrm{W},\kappa}$ is equivalent to weak convergence.

Proposition 10. *Let $\boldsymbol{\mu}$ be a matrix-valued measures on $[0,1]$ and let $\{\boldsymbol{\mu}_k : k\in\mathbb{N}\}$ be a sequence of matrix-valued measures with the same dimension as $\boldsymbol{\mu}$. Then*

$$\mathrm{D}_{W,\kappa}(\boldsymbol{\mu}_k,\boldsymbol{\mu}) \underset{k\to\infty}{\longrightarrow} 0 \tag{4.11a}$$

if and only if

$$\forall \boldsymbol{\phi}\in C_b(\mathbb{H}^{\ell\times\ell},[0,1]),\ \mathrm{tr}\left(\int_0^1 \boldsymbol{\phi}\, d\boldsymbol{\mu}_k\right) \underset{k\to\infty}{\longrightarrow} \mathrm{tr}\left(\int_0^1 \boldsymbol{\phi}\, d\boldsymbol{\mu}\right). \tag{4.11b}$$

Proof. Since the set

$$\{\boldsymbol{\phi} \mid \boldsymbol{\phi}(x)-\boldsymbol{\phi}(y)\le c(x,y)I, \|\boldsymbol{\phi}\|\le\kappa\}$$

is a subset of $C_b(\mathbb{H}^{\ell\times\ell},[0,1])$, then clearly (4.11b) implies (4.11a). We will show that (4.11a) also implies (4.11b). The proof below follows the proof in [12, page 216] for the scalar-valued Wasserstein metric.

Since all metrics on the real line are equivalent, we only consider $c(x,y)=|x-y|$. We denote

$$\|\boldsymbol{\phi}\|_{\mathrm{Lip}} := \sup \frac{\|\boldsymbol{\phi}(x)-\boldsymbol{\phi}(y)\|}{|x-y|}$$

the Lipschitz semi-norm. By replacing $\boldsymbol{\phi}$ with $\boldsymbol{\phi}/\max\{\|\boldsymbol{\phi}\|_{\mathrm{Lip}}, \frac{\max\|\boldsymbol{\phi}\|}{\kappa}\}$, we see that (4.11b) holds for all Lipschitz and bounded functions. The proof requires showing that any function in $C_b(\mathbb{H}^{\ell\times\ell},[0,1])$ can be approached by sequences of Lipschitz functions from above and below respectively.

We will need the following expression inf and sup of a $\boldsymbol{\phi}\in C_b(\mathbb{H}^{\ell\times\ell},[0,1])$:

$$\inf_{x\in[0,1]} \boldsymbol{\phi}(x) := \arg\sup_{\boldsymbol{m}} \{\mathrm{tr}(\boldsymbol{m}) \mid \boldsymbol{\phi}(x)-\boldsymbol{m}\ge 0\} \tag{4.12a}$$

$$\sup_{x\in[0,1]} \boldsymbol{\phi}(x) := \arg\inf_{\boldsymbol{m}} \{\mathrm{tr}(\boldsymbol{m}) \mid \boldsymbol{m}-\boldsymbol{\phi}(x)\ge 0\} \tag{4.12b}$$

which are the lower and upper bounds of the matrix-valued function $\boldsymbol{\phi}(x)$ in term of positive semi-definiteness. For $n \in \mathbb{N}$, we denote

$$\begin{aligned}\boldsymbol{\phi}_{\mathrm{low},n}(x) &:= \inf_{y\in[0,1]} \{\boldsymbol{\phi}(y) + nc(x,y)I\} \\ \boldsymbol{\phi}_{\mathrm{up},n}(x) &:= \sup_{y\in[0,1]} \{\boldsymbol{\phi}(y) - nc(x,y)I\}.\end{aligned}$$

Then, $\boldsymbol{\phi}_{\mathrm{low},n}(x) \le \boldsymbol{\phi}(x) \le \boldsymbol{\phi}_{\mathrm{up},n}(x)$. For any fixed x, $\{\boldsymbol{\phi}_{\mathrm{low},n}(x) : n \in \mathbb{N}\}$ and $\{\boldsymbol{\phi}_{\mathrm{up},n}(x) : n \in \mathbb{N}\}$ are increasing sequence and decreasing sequence, respectively, and they both converge to $\boldsymbol{\phi}(x)$. It is also important to note that $\|\boldsymbol{\phi}_{\mathrm{low},n}\|_{\mathrm{Lip}} \le n$ and $\|\boldsymbol{\phi}_{\mathrm{up},n}\|_{\mathrm{Lip}} \le n$. To see this, for any $x, y \in [0,1]$, we have

$$\begin{aligned}\boldsymbol{\phi}_{\mathrm{low},n}(x) - \boldsymbol{\phi}_{\mathrm{low},n}(y) &= \inf_{z_1}\left\{\boldsymbol{\phi}(z_1) + nc(x,z_1)I\right\} - \inf_{z_2}\left\{\boldsymbol{\phi}(z_2) + nc(y,z_2)I\right\} \\ &\le \sup_{z}\left\{(\boldsymbol{\phi}(z) + nc(x,z)I) - (\boldsymbol{\phi}(z) + nc(y,z)I)\right\} \\ &= \sup_{z} n(c(x,z) - c(y,z))I \\ &\le nc(x,y)I.\end{aligned}$$

Thus $\|\boldsymbol{\phi}_{\mathrm{low},n}\|_{\mathrm{Lip}} \le n$, and $\|\boldsymbol{\phi}_{\mathrm{up},n}\|_{\mathrm{Lip}} \le n$ can be proved in a similar manner. Then we have

$$\begin{aligned}\limsup_{k\to\infty} \mathrm{tr}\left(\int_0^1 \boldsymbol{\phi} d\boldsymbol{\mu}_k\right) &\le \liminf_{n\to\infty} \limsup_{k\to\infty} \mathrm{tr}\left(\int_0^1 \boldsymbol{\phi}_{\mathrm{up},n} d\boldsymbol{\mu}_k\right) \\ &= \liminf_{n\to\infty} \mathrm{tr}\left(\int_0^1 \boldsymbol{\phi}_{\mathrm{up},n} d\boldsymbol{\mu}\right) \\ &= \mathrm{tr}\left(\int_0^1 \boldsymbol{\phi} d\boldsymbol{\mu}\right).\end{aligned}$$

Similarly, using the sequence $\{\boldsymbol{\phi}_{\mathrm{low},n}(x) : n \in \mathbb{N}\}$ we can also show that

$$\liminf_{k\to\infty} \mathrm{tr}\left(\int_0^1 \boldsymbol{\phi} d\boldsymbol{\mu}_k\right) \ge \mathrm{tr}\left(\int_0^1 \boldsymbol{\phi} d\boldsymbol{\mu}\right).$$

Then the proof is completed. □

4.4 Generalization with convex cost functions

The distance $\mathrm{D_W}$ relies on the cost being concave. Indeed, when the cost function is convex, e.g. $c(x,y) = |x-y|^2$, $\mathrm{D_W}$ is zero. In this section, we explore a possibility of defining a meaningful notion of distance that is also suitable for convex cost.

For $\boldsymbol{f}_0, \boldsymbol{f}_1 \in \hat{\mathscr{S}}$ and $\kappa \geq 0$, we define

$$\hat{\mathrm{D}}_{\mathrm{W},\kappa}(\boldsymbol{f}_0, \boldsymbol{f}_1) := \min_{\boldsymbol{m}, \tilde{\boldsymbol{f}}_0, \tilde{\boldsymbol{f}}_1} \left\{ \int_0^1 \int_0^1 \mathrm{tr}(\boldsymbol{m}(x)c(x,y))dxdy + \kappa \sum_{k=0}^{1} \|\tilde{\boldsymbol{f}}_k - \boldsymbol{f}_k\|_{\mathrm{TV}} \;\; | \right.$$
$$\left. \boldsymbol{m}(x,y) \geq 0, \int_0^1 \boldsymbol{m}(x,y)dy = \tilde{\boldsymbol{f}}_0(x), \int_0^1 \boldsymbol{m}(x,y)dx = \tilde{\boldsymbol{f}}_1(y) \right\}. \quad (4.13)$$

We note that the above $\boldsymbol{m}(x,y)$ is constrained to be positive semi-definite whereas the $\boldsymbol{m}(x,y)$ in (4.9) for defining $\mathrm{D}_{\mathrm{W},\kappa}$ is not. Apparently,

$$\hat{\mathrm{D}}_{\mathrm{W},\kappa}(\boldsymbol{f}_0, \boldsymbol{f}_1) \geq \mathrm{D}_{\mathrm{W},\kappa}(\boldsymbol{f}_0, \boldsymbol{f}_1).$$

The distance $\hat{\mathrm{D}}_{\mathrm{W},\kappa}$ relates to (3.24) and they coincide when $\boldsymbol{f}_0$ and $\boldsymbol{f}_1$ are both scalar-valued densities. Using Lagrangian multiplier method, we obtain the dual of (4.13) as

$$\hat{\mathrm{D}}_{\mathrm{W},\kappa}(\boldsymbol{f}_0, \boldsymbol{f}_1) = \sup_{\boldsymbol{\phi}, \boldsymbol{\psi}} \left\{ \int_0^1 \mathrm{tr}\Big(\boldsymbol{\phi}(x)\boldsymbol{f}_0(x) + \boldsymbol{\psi}(x)\boldsymbol{f}_1(x) \Big) dx \;\; | \;\; \boldsymbol{\phi}(x) + \boldsymbol{\psi}(y) \leq c(x,y)I, \right.$$
$$\left. \|\boldsymbol{\phi}\| \leq \kappa, \; \|\boldsymbol{\psi}\| \leq \kappa \right\}. \quad (4.14)$$

Since the two functions $\boldsymbol{\phi}$ and $\boldsymbol{\psi}$ are bounded, $\hat{\mathrm{D}}_{\mathrm{W},\kappa}$ is finite. However, $\hat{\mathrm{D}}_{\mathrm{W},\kappa}$ is not a metric since the triangular inequality is not satisfied.

4.5 Remarks

We extend the formulation of the Monge-Kantorovich problem to matrix-valued density functions in (4.3). This generalization is based on an observation that in the Kantorovich-Rubinstein problem, we can use an absolute value penalty of the transportation plan and remove the non-negative constraint and the optimal value does not change. This observation introduces the possibility of using "negative mass" in the mass transportation problem. In the matrix-valued formulation, the negative mass is generalized to joint Hermitian matrix-valued densities that are not necessary positive semi-definite. The nuclear norm of the matrix-valued density is penalized in the objective function. This formulation is extended to compare densities with unbalanced masses. It is shown that this is a natural generalization of the scalar-valued case in the sense that the weak-continuity property of the Wasserstein metric is nicely inherited.

4.5.1 On a "rotation-aware" metric

The distances $\mathrm{D_W}$ and $\mathrm{D_{W,\kappa}}$ are mathematically natural and straightforward generalizations of the scalar-valued counterparts. They also inherit the weak-continuity property of the Wasserstein metric. However, the distances relate to unrealistic assumptions of the density functions that they have equal integral. Though the metric $\mathrm{D_{W,\kappa}}$ is suitable for comparing density functions with non-equal integral, its formulation implicitly requires that there are some "noise-free" components $\tilde{\boldsymbol{f}}_k$'s that have equal integral. In spectral analysis of multivariate time-series, the equal integral assumption implies that the time-series have the same R_0 which is too strong to be realistic.

Our motivation for studying distances between matrix-valued densities originates in the analysis of slowly time-varying process. In this case, the content of matrix-valued spectra may be transported over frequencies and may also rotate between the channels. We are interested in comparing the power spectra by quantifying the proximity of their spectral content across frequencies and directions is weighted similarly. A possible formulation that captures rotation across "channels" is to consider a metric of the form

$$\min_{U,V} \{\mathrm{D_{W,\kappa}}(\boldsymbol{f}_0, U\boldsymbol{f}_1U^*) + \mathrm{D_{W,\kappa}}(V\boldsymbol{f}_0V^*, \boldsymbol{f}_1) + \|U-I\| + \|V-I\| \;\mid\; UU^* = VV^* = I\}$$

where U and V are both unitary matrices that are used to rotate the masses. However, the computation of the above is challenging and we know of no general method.

Chapter 5

OMT for matrix-valued densities with equal integral trace

In the previous chapter, we extended the Monge-Kantorovich problem to matrix-valued density functions based on the assumptions that the density functions have equal integral. In spectral analysis of multivariate time-series, this assumption implies that the time-series have the same covariance R_0 which is unrealistic. In this chapter, we explore a different formulation of the OMT problem for matrix-valued density functions in the set

$$\mathscr{F} := \left\{ \boldsymbol{f} \;\middle|\; \boldsymbol{f} \text{ has support on } [0,1],\ \boldsymbol{f}(x) \geq 0, \text{ and } \operatorname{tr}\left(\int_0^1 \boldsymbol{f}(x)dx \right) = 1 \right\}.$$

In spectral analysis, this equal trace assumption implies that the total power over all the channels of the time-series are fixed. Any non-zero matrix-valued density functions can be scaled to have unit trace by multiplying a scalar.

5.1 Matrix-valued Monge-Kantorovich problem

Our formulation of the OMT problem of matrix-valued densities is based on a generalization of the transportation cost and the joint density function. The new transportation cost consists of a cost of transference of the "mass" across frequencies and a cost of rotating the "direction" of matrix densities. We find a natural formulation of the joint matrix-valued density function is in the tensor product space.

5.1.1 Tensor product and partial trace

Consider two ℓ-dimensional real or complex (Hilbert) spaces $\mathscr{H}_0$ and $\mathscr{H}_1$. Let $\mathscr{L}(\mathscr{H}_0)$ and $\mathscr{L}(\mathscr{H}_1)$ denote the space of linear operators on $\mathscr{H}_0$ and $\mathscr{H}_1$, respectively, and let $\boldsymbol{f}_0 \in \mathscr{L}(\mathscr{H}_0)$ and $\boldsymbol{f}_1 \in \mathscr{L}(\mathscr{H}_1)$. Thus, in the present section, $\boldsymbol{f}_i$ ($i \in \{0,1\}$) are constant. We denote their tensor product by $\boldsymbol{f}_0 \otimes \boldsymbol{f}_1 \in \mathscr{L}(\mathscr{H}_0 \otimes \mathscr{H}_1)$ which is formally defined via

$$\boldsymbol{f}_0 \otimes \boldsymbol{f}_1 \ : \ u \otimes v \mapsto \boldsymbol{f}_0 u \otimes \boldsymbol{f}_1 v.$$

Since our spaces are finite-dimensional, this can be identified with the Kronecker product of the corresponding matrix representation of the two operators. The space $\mathscr{L}(\mathscr{H}_0 \otimes \mathscr{H}_1)$ is the span of all products $\boldsymbol{f}_0 \otimes \boldsymbol{f}_1$ with $\boldsymbol{f}_i \in \mathscr{L}(\mathscr{H}_i)$ for $i \in \{0,1\}$.

Consider $\boldsymbol{f} \in \mathscr{L}(\mathscr{H}_0 \otimes \mathscr{H}_1)$. The partial traces $\mathrm{tr}_{\mathscr{H}_0}$ and $\mathrm{tr}_{\mathscr{H}_1}$, or tr_0 and tr_1 for brevity, are linear maps

$$\begin{aligned} \mathrm{tr}_1 \ : \ \mathscr{L}(\mathscr{H}_0 \otimes \mathscr{H}_1) \ &\to \ \mathscr{L}(\mathscr{H}_0) \ : \ \boldsymbol{f} \mapsto \mathrm{tr}_1(\boldsymbol{f}) \\ \mathrm{tr}_0 \ : \ \mathscr{L}(\mathscr{H}_0 \otimes \mathscr{H}_1) \ &\to \ \mathscr{L}(\mathscr{H}_1) \ : \ \boldsymbol{f} \mapsto \mathrm{tr}_0(\boldsymbol{f}) \end{aligned}$$

defined uniquely by the property that on simple products it acts as follows:

$$\mathrm{tr}_1(\boldsymbol{f}_0 \otimes \boldsymbol{f}_1) = \mathrm{tr}(\boldsymbol{f}_1)\boldsymbol{f}_0 \text{ and } \mathrm{tr}_0(\boldsymbol{f}_0 \otimes \boldsymbol{f}_1) = \mathrm{tr}(\boldsymbol{f}_0)\boldsymbol{f}_1$$

for any $\boldsymbol{f}_0 \in \mathscr{L}(\mathscr{H}_0)$ and $\boldsymbol{f}_1 \in \mathscr{L}(\mathscr{H}_1)$. Alternatively, $\boldsymbol{f} \in \mathscr{L}(\mathscr{H}_0 \otimes \mathscr{H}_1)$ can be represented by a matrix $\left[\boldsymbol{f}_{im,jn}\right]$ of size $\ell^2 \times \ell^2$ as it maps a basis element $u_i \otimes v_m \in \mathscr{H}_0 \otimes \mathscr{H}_1$ to $\sum_{j,n} \boldsymbol{f}_{im,jn} u_j \otimes v_n$. Then, the partial trace e.g., $\mathrm{tr}_1(\boldsymbol{f})$ is represented by the $\ell \times \ell$ matrix with (i,j)-th entry

$$[\mathrm{tr}_1(\boldsymbol{f})]_{i,j} = \sum_k \boldsymbol{f}_{ik,jk}, \text{ for } 1 \le i,j \le \ell.$$

Likewise the (j,n)-th entry of $\mathrm{tr}_0(\boldsymbol{f})$ is

$$[\mathrm{tr}_0(\boldsymbol{f})]_{m,n} = \sum_k \boldsymbol{f}_{km,kn}, \text{ for } 1 \le m,n \le \ell.$$

See [11] for its significance in the context of quantum mechanics. For example, the partial traces of

$$\boldsymbol{f} = \begin{bmatrix} a & b & c & d \\ b & e & f & g \\ c & f & h & i \\ d & g & i & j \end{bmatrix}$$

are given as

$$\mathrm{tr}_1(\boldsymbol{f}) = \begin{bmatrix} a+e & c+g \\ c+g & h+j \end{bmatrix} \text{ and } \mathrm{tr}_0(\boldsymbol{f}) = \begin{bmatrix} a+h & b+i \\ b+i & e+j \end{bmatrix}.$$

5.1.2 Joint matrix-valued density

We now return to considering matrix-valued density functions $\boldsymbol{f}_0, \boldsymbol{f}_1 \in \mathscr{F}$. A naive attempt for a joint matrix-valued density $\boldsymbol{m}$ is to consider $\boldsymbol{m}(x,y)$ in the same way as in Chapter 4 where we consider $\boldsymbol{m}$ as $\ell \times \ell$ symmetric matrices and its integration over y gives $\boldsymbol{f}_0(x)$ and the integration over x gives $\boldsymbol{f}_1(y)$. However, this formulation requires that $\boldsymbol{f}_0$ and $\boldsymbol{f}_1$ have equal integral, otherwise there is no feasible solution. We may also need to introduce non-positive semi-definite mass in order to find a transference plan. An example is given by the following discrete matrix-valued density functions

$$\boldsymbol{f}_0(x) = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 0 \end{bmatrix} \delta(x) + \begin{bmatrix} 0 & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \delta(x-1), \tag{5.1a}$$

$$\boldsymbol{f}_1(x) = \begin{bmatrix} \frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & \frac{1}{4} \end{bmatrix} \delta(x) + \begin{bmatrix} \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} \end{bmatrix} \delta(x-1). \tag{5.1b}$$

There is no positive semi-definite matrix-valued joint density functions of size 2×2 that have the two marginals. We introduce a different and more natural formulation of transference plan based on the tensor product and partial trace.

For any $(x,y) \in [0,1] \times [0,1]$ and

$$\boldsymbol{m}(x,y) \text{ is } \ell^2 \times \ell^2 \text{ positive semi-definite Hermitian matrix,} \tag{5.2a}$$

let

$$\boldsymbol{m}_0(x,y) := \mathrm{tr}_1(\boldsymbol{m}(x,y)),\ \boldsymbol{m}_1(x,y) := \mathrm{tr}_0(\boldsymbol{m}(x,y)), \tag{5.2b}$$

$$\int_0^1 \boldsymbol{m}_0(x,y)dy = \boldsymbol{f}_0(x),\ \int_0^1 \boldsymbol{m}_1(x,y)dx = \boldsymbol{f}_1(y). \tag{5.2c}$$

Denote

$$\boldsymbol{M}(\boldsymbol{f}_0, \boldsymbol{f}_1) := \Big\{ \boldsymbol{m} \mid (5.2a) - (5.2c) \text{ are satisfied} \Big\}$$

which is the set of transference plan we will consider in. Given the two marginal densities, there is always an admissible joint distribution in $\boldsymbol{M}(\boldsymbol{f}_0, \boldsymbol{f}_1)$ since, clearly,

$$\boldsymbol{f}_0 \otimes \boldsymbol{f}_1 \in \boldsymbol{M}(\boldsymbol{f}_0, \boldsymbol{f}_1).$$

5.1.3 Transportation cost

We interpret $\mathrm{tr}(\boldsymbol{m}(x,y))$ as the amount of "mass" that is being transferred from x to y. This value is always non-negative. Thus, for a scalar cost function $c(x,y)$ as before, one may simply consider a "mass transference" cost

$$\min_{\boldsymbol{m}\in\boldsymbol{M}(\boldsymbol{f}_0,\boldsymbol{f}_1)} \int_0^1 \int_0^1 c(x,y)\mathrm{tr}(\boldsymbol{m}(x,y))dxdy. \tag{5.3}$$

However, if $\mathrm{tr}(\boldsymbol{f}_0(x)) = \mathrm{tr}(\boldsymbol{f}_1(x)),\ \forall x \in \mathbb{R}$, then the optimal value of (5.3) is zero. Thus (5.3) fails to quantify the mismatch of directionality in the present matricial setting. We need to use a different formulation of transportation cost.

For simplicity, throughout, we only consider marginals $\boldsymbol{f}$ which point-wise satisfy $\mathrm{tr}(\boldsymbol{f}) > 0$; $\mathrm{tr}(\boldsymbol{f}(x))$ is a scalar-valued density representing mass at location x while $\frac{\boldsymbol{f}(x)}{\mathrm{tr}(\boldsymbol{f}(x))}$ has trace 1 and contains directional information. In a similar manner, for the joint density $\boldsymbol{m}(x,y)$, assuming that $\boldsymbol{m}(x,y) \neq 0$, we define the normalized partial traces

$$\underline{\mathrm{tr}}_0(\boldsymbol{m}(x,y)) := \mathrm{tr}_0(\boldsymbol{m}(x,y))/\mathrm{tr}(\boldsymbol{m}(x,y))$$
$$\underline{\mathrm{tr}}_1(\boldsymbol{m}(x,y)) := \mathrm{tr}_1(\boldsymbol{m}(x,y))/\mathrm{tr}(\boldsymbol{m}(x,y)).$$

Since $\underline{\mathrm{tr}}_0(\boldsymbol{m}(x,y))$ and $\underline{\mathrm{tr}}_1(\boldsymbol{m}(x,y))$ are normalized to have unit trace, their difference captures the directional mismatch between the two partial traces. We take the "rotational cost" of the form

$$\mathrm{d}^2\left(\underline{\mathrm{tr}}_0(\boldsymbol{m}), \underline{\mathrm{tr}}_1(\boldsymbol{m})\right)\mathrm{tr}(\boldsymbol{m}(x,y)) \tag{5.4}$$

where $\mathrm{d}(\cdot,\cdot)$ denotes a distance measure between positive semi-definite matrices. Note that in the scalar-valued case, $\underline{\mathrm{tr}}_0(m) = \underline{\mathrm{tr}}_1(m) = 1$, thus (5.4) reduces to zero. We discuss two examples of the rotational cost next.

Using the Frobenius distance for $\mathrm{d}(\cdot,\cdot)$, we take

$$\mathrm{tr}(\|(\underline{\mathrm{tr}}_0 - \underline{\mathrm{tr}}_1)\boldsymbol{m}(x,y)\|_{\mathrm{F}}^2\boldsymbol{m}(x,y)) \tag{5.5}$$

for the rotational mismatch (which may be thought as "rotational kinetic energy"). The above motivates the following cost functional that includes both terms, rotational and linear:

$$\mathrm{tr}\left((c(x,y) + \lambda\|(\underline{\mathrm{tr}}_0 - \underline{\mathrm{tr}}_1)\boldsymbol{m}(x,y)\|_{\mathrm{F}}^2)\boldsymbol{m}(x,y)\right) \tag{5.6}$$

with $\lambda > 0$ to weigh in the relative significance of the two terms. A more interesting option for $\mathrm{d}(\cdot,\cdot)$ is the Bures distance $\mathrm{d}_{\mathrm{Bures}}$ given in (2.12). The corresponding cost function is given as

$$\mathrm{tr}\bigg((c(x,y)+\lambda\,\mathrm{d}^2_{\mathrm{Bures}}(\underline{\mathrm{tr}}_0(\boldsymbol{m})-\underline{\mathrm{tr}}_1(\boldsymbol{m})))\boldsymbol{m}(x,y)\bigg). \tag{5.7}$$

As we pointed out in Chapter 3, the Bures distance $\mathrm{d}_{\mathrm{Bures}}(\underline{\mathrm{tr}}_0(\boldsymbol{m})-\underline{\mathrm{tr}}_1(\boldsymbol{m}))$ coincides with the 2-Wasserstein metric between two zero-mean Gaussian distributions with covariance matrices $\underline{\mathrm{tr}}_0(\boldsymbol{m})$ and $\underline{\mathrm{tr}}_1(\boldsymbol{m})$, respectively.

5.1.4 Matrix-valued OMT

We use the cost functionals in (5.6) and (5.7), respectively, to formulate the corresponding OMT problem for matrix-valued densities.

Matrix-valued OMT based on quadratic cost

In view of (5.5), we now arrive at the following formulation of a matrix-valued version of the OMT, namely the determination of

$$\mathscr{T}_\lambda(\boldsymbol{f}_0,\boldsymbol{f}_1) := \min_{\boldsymbol{m}\in\boldsymbol{M}(\boldsymbol{f}_0,\boldsymbol{f}_1)}\int_0^1\int_0^1 \mathrm{tr}\bigg((c+\lambda\|(\underline{\mathrm{tr}}_0-\underline{\mathrm{tr}}_1)\boldsymbol{m}\|_{\mathrm{F}}^2)\boldsymbol{m}\bigg)\,dxdy. \tag{5.8}$$

Interestingly, (5.8) can be cast as a convex optimization problem as shown next.

By definition,

$$\begin{aligned}\underline{\mathrm{tr}}_0(\boldsymbol{m})\mathrm{tr}(\boldsymbol{m}) &= \mathrm{tr}_0(\boldsymbol{m}),\\ \underline{\mathrm{tr}}_1(\boldsymbol{m})\mathrm{tr}(\boldsymbol{m}) &= \mathrm{tr}_1(\boldsymbol{m}).\end{aligned}$$

Hence,

$$\begin{aligned}\|(\underline{\mathrm{tr}}_0-\underline{\mathrm{tr}}_1)\boldsymbol{m}\|_{\mathrm{F}}^2\mathrm{tr}(\boldsymbol{m}) &= \frac{\|(\underline{\mathrm{tr}}_0-\underline{\mathrm{tr}}_1)\boldsymbol{m}\|_{\mathrm{F}}^2\mathrm{tr}(\boldsymbol{m})^2}{\mathrm{tr}(\boldsymbol{m})}\\ &= \frac{\|(\mathrm{tr}_0-\mathrm{tr}_1)\boldsymbol{m}\|_{\mathrm{F}}^2}{\mathrm{tr}(\boldsymbol{m})}.\end{aligned}$$

Now let $m(x,y)=\mathrm{tr}(\boldsymbol{m}(x,y))$ and $\boldsymbol{m}_0(x,y)$ and $\boldsymbol{m}_1(x,y)$ be as in (5.2). The expression for the

optimal cost in (5.8) can be equivalently written as

$$\min_{\boldsymbol{m}_0,\boldsymbol{m}_1,m}\Big\{\int\ \left(c(x,y)m(x,y)+\lambda\,\frac{\|\boldsymbol{m}_0-\boldsymbol{m}_1\|_{\mathrm{F}}^2}{m}\right)dxdy\ \mid\ \boldsymbol{m}_0(x,y),\ \boldsymbol{m}_1(x,y)\geq 0,$$
$$\mathrm{tr}(\boldsymbol{m}_0(x,y))=\mathrm{tr}(\boldsymbol{m}_1(x,y))=m(x,y),$$
$$\int \boldsymbol{m}_0(x,y)dy=\boldsymbol{f}_0(x),$$
$$\int \boldsymbol{m}_1(x,y)dx=\boldsymbol{f}_1(y)\Big\}. \tag{5.9}$$

Since, the function $\frac{(y-z)^2}{x}$ is a jointly convex function when $x>0$ and $y,z\in\mathbb{R}$. It readily follows that the integral in (5.9) is a convex functional of its arguments. All additional constraints in (5.9) are convex as well and therefore, so is the optimization problem.

Matrix-valued OMT based on the Bures distance

Using (5.7), we are lead to the following problem

$$\min_{\boldsymbol{m}\in\boldsymbol{M}(\boldsymbol{f}_0,\boldsymbol{f}_1)}\int_0^1\int_0^1 \mathrm{tr}\left((c+\lambda\,\mathrm{d}^2_{\mathrm{Bures}}(\underline{\mathrm{tr}}_0(\boldsymbol{m}),\underline{\mathrm{tr}}_1(\boldsymbol{m}))\boldsymbol{m}\right)dxdy. \tag{5.10}$$

This can also be cast as a convex optimization problem as shown next.

From (2.10), we have

$$\mathrm{d}^2_{\mathrm{Bures}}(\underline{\mathrm{tr}}_0(\boldsymbol{m}),\underline{\mathrm{tr}}_1(\boldsymbol{m}))\mathrm{tr}(\boldsymbol{m})=\mathrm{d}^2_{\mathrm{Bures}}(\mathrm{tr}_0(\boldsymbol{m}),\mathrm{tr}_1(\boldsymbol{m})).$$

As pointed out in Section 3.2.1, $\mathrm{d}^2_{\mathrm{Bures}}(\mathrm{tr}_0(\boldsymbol{m}),\mathrm{tr}_1(\boldsymbol{m}))$ is the solution of an LMI as follows

$$\mathrm{d}^2_{\mathrm{Bures}}(\mathrm{tr}_0(\boldsymbol{m}),\mathrm{tr}_1(\boldsymbol{m}))=\min_{\boldsymbol{c}}\left\{\mathrm{tr}\left(\begin{bmatrix}I & -I\\ -I & I\end{bmatrix}\begin{bmatrix}\mathrm{tr}_0(\boldsymbol{m}) & \boldsymbol{c}\\ \boldsymbol{c}^* & \mathrm{tr}_1(\boldsymbol{m})\end{bmatrix}\right)\ \mid\ \begin{bmatrix}\mathrm{tr}_0(\boldsymbol{m}) & \boldsymbol{c}\\ \boldsymbol{c}^* & \mathrm{tr}_1(\boldsymbol{m})\end{bmatrix}\geq 0\right\}.$$

Let $\boldsymbol{m}_0(x,y)$ and $\boldsymbol{m}_1(x,y)$ be as in (5.2). Then $\mathrm{tr}(\boldsymbol{m}_0)=\mathrm{tr}(\boldsymbol{m}_1)=\mathrm{tr}(\boldsymbol{m})$. We obtain the following expression of (5.10)

$$\min_{\boldsymbol{c},\boldsymbol{m}_0,\boldsymbol{m}_1}\Big\{\int_0^1\int_0^1 \mathrm{tr}\left(\begin{bmatrix}(\lambda+\frac{c}{2})I & -\lambda I\\ -\lambda I & (\lambda+\frac{c}{2})I\end{bmatrix}\begin{bmatrix}\boldsymbol{m}_0 & \boldsymbol{c}\\ \boldsymbol{c}^* & \boldsymbol{m}_1\end{bmatrix}\right)dxdy\ \mid\ \begin{bmatrix}\boldsymbol{m}_0 & \boldsymbol{c}\\ \boldsymbol{c}^* & \boldsymbol{m}_1\end{bmatrix}\geq 0,$$
$$\mathrm{tr}(\boldsymbol{m}_0)=\mathrm{tr}(\boldsymbol{m}_1),$$
$$\int \boldsymbol{m}_0(x,y)dy=\boldsymbol{f}_0(x),$$
$$\int \boldsymbol{m}_1(x,y)dx=\boldsymbol{f}_1(y)\Big\}.$$

The objective function is linear and the constraints are convex, so the optimization problem is convex as well.

For simplicity, in the following, we only consider (5.8) for the OMT problem for matrix-valued densities. Next, we discuss the structure of the corresponding optimal transference plan.

5.2 On the optimal transference plan

A standard result in the (scalar) OMT theory is that the transference plan is the sub-differential of a convex function. As a consequence the transference plan has support on a monotonic zero-measure set. This is no longer true for the optimal transference plan for matrix-valued density functions as we discuss next. This is explained in the following subsections.

5.2.1 The supporting set of transference plan

In optimal transport of scalar-valued distributions, the optimal transference plan has a certain cyclically monotonic property [12]. More specifically, if (x_1, y_1), (x_2, y_2) are two points where the transference plan has support (i.e., $m(x_k, y_k) \neq 0$ for k=1, 2), then $x_2 > x_1$ implies $y_2 \geq y_1$. The interpretation is that optimal transportation paths of mass elements do not cross. For the case of matrix-valued distributions as in (13), this property may not hold in the same way. However, interestingly, a weaker monotonicity property holds for the supporting set of the optimal matrix transference plan. The property is defined next and the precise statement is given in Proposition 12 below.

Definition 11. *A set $\mathscr{S} \subset \mathbb{R}^2$ is called a* λ-monotonically non-decreasing*, for $\lambda > 0$, if for any two points $(x_1, y_1), (x_2, y_2) \in \mathscr{S}$, it holds that*

$$(x_2 - x_1)(y_1 - y_2) \leq \lambda.$$

A geometric interpretation for a λ-monotonically non-decreasing set is that if (x_1, y_1), $(x_2, y_2) \in \mathscr{S}$ and $x_2 > x_1$, $y_1 > y_2$, then the area of the rectangle with vertices (x_i, y_j) $(i, j \in \{1, 2\})$ is not larger than λ. The transference plan of the scalar-valued optimal transportation problem with a quadratic cost has support on a 0-monotonically non-decreasing set.

Proposition 12. *Given $\boldsymbol{f}_0, \boldsymbol{f}_1 \in \mathscr{F}$, let $\boldsymbol{m}$ be the optimal transference plan in* (5.8) *with $\lambda > 0$. Then $\boldsymbol{m}$ has support on at most a $(4 \cdot \lambda)$-monotonically non-decreasing set.*

Proof. We need to show that if $\boldsymbol{m}(x_1,y_1) \neq 0$ and $\boldsymbol{m}(x_2,y_2) \neq 0$, then $x_2 > x_1$, $y_1 > y_2$ implies

$$(x_2 - x_1)(y_1 - y_2) \leq 4\lambda. \tag{5.11}$$

Without loss of generality let

$$\boldsymbol{m}(x_i,y_j) \quad = \quad m_{ij} \cdot A_{ij} \otimes B_{ij} \tag{5.12}$$

with $A_{ij}, B_{ij} \geq 0$, $\mathrm{tr}(A_{ij}) = \mathrm{tr}(B_{ij}) = 1$ and $i,j \in \{1,2\}$. Note that m_{12} and m_{21} could be zero if $\boldsymbol{m}$ does not have support on the particular point. We assume that the condition in the proposition fails and that

$$(x_2 - x_1)(y_1 - y_2) > 4\lambda. \tag{5.13}$$

We then show that by rearranging the mass, the cost can be reduced.

We first consider the situation when $m_{22} \geq m_{11}$. By rearranging the value of $\boldsymbol{m}$ at the four points (x_i,y_j) with $i,j \in \{1,2\}$, we construct a new transference plan $\tilde{\boldsymbol{m}}$ at these four locations as follows

$$\begin{aligned}
\tilde{\boldsymbol{m}}(x_1,y_1) &= 0 && \text{(5.14a)}\\
\tilde{\boldsymbol{m}}(x_1,y_2) &= (m_{11}+m_{12}) \cdot \tilde{A}_{12} \otimes \tilde{B}_{12} && \text{(5.14b)}\\
\tilde{\boldsymbol{m}}(x_2,y_1) &= (m_{11}+m_{21}) \cdot \tilde{A}_{21} \otimes \tilde{B}_{21} && \text{(5.14c)}\\
\tilde{\boldsymbol{m}}(x_2,y_2) &= (m_{22}-m_{11}) \cdot A_{22} \otimes B_{22} && \text{(5.14d)}
\end{aligned}$$

where

$$\begin{aligned}
\tilde{A}_{12} &= \frac{m_{11}A_{11}+m_{12}A_{12}}{m_{11}+m_{12}}, \tilde{B}_{12} = \frac{m_{11}B_{22}+m_{12}B_{12}}{m_{11}+m_{12}}\\
\tilde{A}_{21} &= \frac{m_{11}A_{22}+m_{21}A_{21}}{m_{11}+m_{21}}, \tilde{B}_{21} = \frac{m_{11}B_{11}+m_{21}B_{21}}{m_{11}+m_{21}}.
\end{aligned}$$

This new transference plan $\tilde{\boldsymbol{m}}$ has the same marginals as $\boldsymbol{m}$ at x_1, x_2 and y_1, y_2. The original cost incurred by $\boldsymbol{m}$ at these four locations is

$$\sum_{i=1}^{2}\sum_{j=1}^{2} m_{ij}\left((x_i - y_j)^2 + \lambda \|A_{ij} - B_{ij}\|_{\mathrm{F}}^2\right) \tag{5.15}$$

while the cost incurred by $\tilde{\boldsymbol{m}}$ is

$$\begin{aligned}
&(m_{11}+m_{12})\left((x_1-y_2)^2 + \lambda\|\tilde{A}_{12} - \tilde{B}_{12}\|_{\mathrm{F}}^2\right)\\
+&(m_{11}+m_{21})\left((x_2-y_1)^2 + \lambda\|\tilde{A}_{21} - \tilde{B}_{21}\|_{\mathrm{F}}^2\right)\\
+&(m_{22}-m_{11})\left((x_2-y_2)^2 + \lambda\|A_{22} - B_{22}\|_{\mathrm{F}}^2\right).
\end{aligned} \tag{5.16}$$

After simplification, to show that (5.15) is larger than (5.16), it suffices to show that

$$2m_{11}(x_2 - x_1)(y_1 - y_2) \tag{5.17}$$

is larger than

$$\lambda m_{11} \left(\sum_{i=1}^{2} \sum_{j \neq i} \|\tilde{A}_{ij} - \tilde{B}_{ij}\|_{\mathrm{F}}^2 - \sum_{i=1}^{2} \|A_{ii} - B_{ii}\|_{\mathrm{F}}^2 \right) \tag{5.18a}$$

$$+ \lambda m_{12} \left(\|\tilde{A}_{12} - \tilde{B}_{12}\|_{\mathrm{F}}^2 - \|A_{12} - B_{12}\|_{\mathrm{F}}^2 \right) \tag{5.18b}$$

$$+ \lambda m_{21} \left(\|\tilde{A}_{21} - \tilde{B}_{21}\|_{\mathrm{F}}^2 - \|A_{21} - B_{21}\|_{\mathrm{F}}^2 \right). \tag{5.18c}$$

From (5.13), it follows that the value in (5.17) is greater than $8\lambda m_{11}$. We derive upper bounds for each term in (5.18). First,

$$(5.18a) \leq \lambda m_{11} \left(\|\tilde{A}_{12} - \tilde{B}_{12}\|_{\mathrm{F}}^2 + \|\tilde{A}_{21} - \tilde{B}_{21}\|_{\mathrm{F}}^2 \right) \leq 4\lambda m_{11}$$

where the last inequality follows from the fact that

$$\|A - B\|_{\mathrm{F}}^2 = \mathrm{tr}(A^2 - 2AB + B^2) \leq \mathrm{tr}(A^2 + B^2) \leq 2$$

for any $A, B \geq 0$ with $\mathrm{tr}(A) = \mathrm{tr}(B) = 1$. Now consider

$$\begin{aligned}
&\|\tilde{A}_{12} - \tilde{B}_{12}\|_{\mathrm{F}}^2 - \|A_{12} - B_{12}\|_{\mathrm{F}}^2 \\
=&\mathrm{tr}\left((\tilde{A}_{12} - \tilde{B}_{12} + A_{12} - B_{12})(\tilde{A}_{12} - \tilde{B}_{12} - A_{12} + B_{12}) \right) \\
=&\frac{m_{11}}{m_{11} + m_{12}} \left(\|A_{11} - B_{22}\|_{\mathrm{F}}^2 - \|A_{12} - B_{12}\|_{\mathrm{F}}^2 - \frac{m_{12}}{m_{11} + m_{12}} \|A_{11} - B_{22} - A_{12} + B_{12}\|_{\mathrm{F}}^2 \right) \\
\leq&\frac{m_{11}}{m_{11} + m_{12}} \|A_{11} - B_{22}\|_{\mathrm{F}}^2 \\
\leq&2\frac{m_{11}}{m_{11} + m_{12}}
\end{aligned}$$

where the second equality follows from the definitions of $\tilde{A}_{12}$ and $\tilde{B}_{12}$ while the last inequality is obtained by bounding the terms in the trace. Thus,

$$(5.18b) \leq 2\lambda m_{12} \frac{m_{11}}{m_{11} + m_{12}} \leq 2\lambda m_{11}.$$

In a similar manner, $(5.18c) \leq 2\lambda m_{11}$. Therefore,

$$(5.18) \leq 8\lambda m_{11} < (5.17)$$

which implies that the cost incurred by $\tilde{\boldsymbol{m}}$ is smaller than the cost incurred by $\boldsymbol{m}$.

For the case where $m_{11} > m_{22}$, we can prove the claim by constructing a new transference plan $\hat{\boldsymbol{m}}$ with values

$$\begin{aligned}
\hat{\boldsymbol{m}}(x_1, y_1) &= (m_{11} - m_{22}) \cdot A_{11} \otimes B_{11} \\
\hat{\boldsymbol{m}}(x_1, y_2) &= (m_{12} + m_{22}) \cdot \hat{A}_{12} \otimes \hat{B}_{12} \\
\hat{\boldsymbol{m}}(x_2, y_1) &= (m_{21} + m_{22}) \cdot \hat{A}_{21} \otimes \hat{B}_{21} \\
\hat{\boldsymbol{m}}(x_2, y_2) &= 0
\end{aligned}$$

with

$$\begin{aligned}
\hat{A}_{12} &= \frac{m_{12}A_{12} + m_{22}A_{11}}{m_{12} + m_{22}}, \hat{B}_{12} = \frac{m_{12}B_{12} + m_{22}B_{22}}{m_{12} + m_{22}} \\
\hat{A}_{21} &= \frac{m_{21}A_{21} + m_{22}A_{22}}{m_{21} + m_{22}}, \hat{B}_{21} = \frac{m_{21}B_{21} + m_{22}B_{11}}{m_{21} + m_{22}}.
\end{aligned}$$

The rest of the proof is carried out in a similar manner. □

We illustrate the proposition using a synthetic example as follows. Let

$$A = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \text{ and } B = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

For f_0 and f_1 shown in Figure 3.1, let

$$\boldsymbol{f}_0(x) = f_0 \cdot (A(1-x) + Bx) \text{ and } \boldsymbol{f}_1(x) = f_1 \cdot (Ax + B(1-x))$$

for $x \in [0,1]$. Then the trace of the optimal transference plan with $\lambda = 0.2$, which is the optimal m in (5.9) is shown in Figure 5.1. The points (x_1, y_1) and (x_2, y_2) are two points where the transference plan has support on. The area of the rectangle shown in Figure 5.1 is upper bounded by 4λ.

Further, the optimal transportation cost $\mathscr{T}_\lambda(\boldsymbol{f}_0, \boldsymbol{f}_1)$ satisfies:

1. $\mathscr{T}_\lambda(\boldsymbol{f}_0, \boldsymbol{f}_1) = \mathscr{T}_\lambda(\boldsymbol{f}_1, \boldsymbol{f}_0)$,
2. $\mathscr{T}_\lambda(\boldsymbol{f}_0, \boldsymbol{f}_1) \geq 0$,
3. $\mathscr{T}_\lambda(\boldsymbol{f}_0, \boldsymbol{f}_1) = 0$ if and only if $\boldsymbol{f}_0 = \boldsymbol{f}_1$.

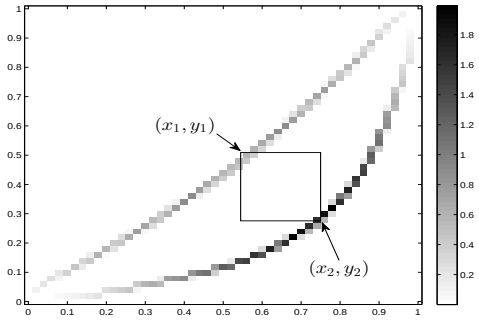

Figure 5.1: The trace of optimal transference plan between $\boldsymbol{f}_0$, $\boldsymbol{f}_1$ with $\lambda = 0.2$.

Thus, although $\mathscr{T}_\lambda(\boldsymbol{f}_0, \boldsymbol{f}_1)$ can be used to compare matrix-valued densities, it is not a metric and neither is

$$\mathrm{D}_{\mathrm{T},\lambda} := \mathscr{T}_\lambda^{\frac{1}{2}} \tag{5.19}$$

since the triangular inequality does not hold in general. This is explained in the next.

Let $\boldsymbol{f}_0$ and $\boldsymbol{f}_1$ be given as in (5.1) and let

$$\boldsymbol{f}_2 = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \delta(x).$$

Since $\boldsymbol{f}_2$ can be split as

$$\boldsymbol{f}_2 = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 0 \end{bmatrix} \delta(x) + \begin{bmatrix} 0 & 0 \\ 0 & \frac{1}{2} \end{bmatrix} \delta(x),$$

where the two matrices are compatible with those in $\boldsymbol{f}_0$, then $\mathscr{T}_\lambda(\boldsymbol{f}_2, \boldsymbol{f}_0)$ only consists of the mass transference cost. Thus

$$\mathscr{T}_\lambda(\boldsymbol{f}_2, \boldsymbol{f}_0) = \frac{1}{2}$$

which is independent of λ. In a similar manner, we also have

$$\mathscr{T}_\lambda(\boldsymbol{f}_2, \boldsymbol{f}_1) = \frac{1}{2}.$$

On the other hand, since the matrices in $\boldsymbol{f}_0$ and $\boldsymbol{f}_1$ do not have compatible directions then $\mathscr{T}_\lambda(\boldsymbol{f}_1, \boldsymbol{f}_0)$ always consists of a part due to the "rotational" cost. Thus the value $\mathscr{T}_\lambda(\boldsymbol{f}_1, \boldsymbol{f}_0)$ depends on λ. We can show that

$$\mathscr{T}_\lambda(\boldsymbol{f}_1, \boldsymbol{f}_0) \geq \lambda$$

which is larger than $\mathscr{T}_\lambda(\boldsymbol{f}_2,\boldsymbol{f}_0)+\mathscr{T}_\lambda(\boldsymbol{f}_2,\boldsymbol{f}_0)$ when $\lambda>1$. Thus, $\mathscr{T}_\lambda$ fails to have the triangular inequality property.

In the following, we introduce a slightly different formulation of a transportation problem which does give rise to a metric.

5.2.2 On a constrained transference plan

We restrict attention to a certain subset of transport plans of $\boldsymbol{M}(\boldsymbol{f}_0,\boldsymbol{f}_1)$ and show that the corresponding optimal transportation cost induces a metric. More specifically, let

$$\boldsymbol{M}_0(\boldsymbol{f}_0,\boldsymbol{f}_1):=\left\{\boldsymbol{m}\mid\ \boldsymbol{m}(x,y)=\boldsymbol{f}_0(x)\otimes\boldsymbol{f}_1(y)a(x,y),\ \boldsymbol{m}\in\boldsymbol{M}\right\}.$$

For $\boldsymbol{m}(x,y)\in\boldsymbol{M}_0(\boldsymbol{f}_0,\boldsymbol{f}_1)$,

$$\underline{\mathrm{tr}}_0(\boldsymbol{m}(x,y)):=\boldsymbol{f}_1(x)/\mathrm{tr}(\boldsymbol{f}_1(x))$$
$$\underline{\mathrm{tr}}_1(\boldsymbol{m}(x,y)):=\boldsymbol{f}_0(y)/\mathrm{tr}(\boldsymbol{f}_0(y)).$$

Given $\boldsymbol{f}_0$ and $\boldsymbol{f}_1$, the "orientation" of the mass of $\boldsymbol{m}(x,y)$ is fixed. Thus, in this case, the optimal transportation cost is

$$\mathscr{T}_{\mathrm{sub},\lambda}(\boldsymbol{f}_0,\boldsymbol{f}_1):=\min_{\boldsymbol{m}\in\boldsymbol{M}_0(\boldsymbol{f}_0,\boldsymbol{f}_1)}\int\mathrm{tr}\left((c+\lambda\|(\underline{\mathrm{tr}}_0-\underline{\mathrm{tr}}_1)\boldsymbol{m}(x,y)\|_{\mathrm{F}}^2)\boldsymbol{m}\right)dxdy. \tag{5.20}$$

In this problem, given the marginals $\boldsymbol{f}_0,\boldsymbol{f}_1$, we can compute the transportation cost function $(c+\lambda\|(\underline{\mathrm{tr}}_0-\underline{\mathrm{tr}}_1)\boldsymbol{m}(x,y)\|_{\mathrm{F}}^2)$ explicitly. So that (5.20) can be viewed as a linear programming problem.

Proposition 13. *For $\mathscr{T}_{sub,\lambda}(\boldsymbol{f}_0,\boldsymbol{f}_1)$ as in* (5.20)*,*

$$\hat{\mathrm{D}}_{T,\lambda}:=\mathscr{T}_{sub,\lambda}(\boldsymbol{f}_0,\boldsymbol{f}_1)^{\frac{1}{2}} \tag{5.21}$$

defines a metric on $\mathscr{F}$.

Proof. It is straightforward to see that

$$\hat{\mathrm{D}}_{\mathrm{T},\lambda}(\boldsymbol{f}_0,\boldsymbol{f}_1)=\hat{\mathrm{D}}_{\mathrm{T},\lambda}(\boldsymbol{f}_1,\boldsymbol{f}_0)\geq 0$$

and that $\hat{\mathrm{D}}_{\mathrm{T},\lambda}(\boldsymbol{f}_0,\boldsymbol{f}_1)=0$ if and only if $\boldsymbol{f}_0=\boldsymbol{f}_1$. We now show that the triangle inequality holds as well. For $\boldsymbol{f}_0,\boldsymbol{f}_1,\boldsymbol{f}_2\in\mathscr{F}$, let

$$\boldsymbol{m}_{01}(x,y)=\frac{\boldsymbol{f}_0(x)}{\mathrm{tr}(\boldsymbol{f}_0(x))}\otimes\frac{\boldsymbol{f}_1(y)}{\mathrm{tr}(\boldsymbol{f}_1(y))}m_{01}(x,y)$$
$$\boldsymbol{m}_{12}(y,z)=\frac{\boldsymbol{f}_1(y)}{\mathrm{tr}(\boldsymbol{f}_1(y))}\otimes\frac{\boldsymbol{f}_2(z)}{\mathrm{tr}(\boldsymbol{f}_2(z))}m_{12}(y,z)$$

denote the optimal transference plan for the pairs $(\boldsymbol{f}_0,\boldsymbol{f}_1)$ and $(\boldsymbol{f}_1,\boldsymbol{f}_2)$, respectively, where m_{01} and m_{12} are two (scalar-valued) joint densities on $\mathbb{R}^2$ with marginals $\mathrm{tr}(\boldsymbol{f}_0)$, $\mathrm{tr}(\boldsymbol{f}_1)$ and $\mathrm{tr}(\boldsymbol{f}_1)$, $\mathrm{tr}(\boldsymbol{f}_2)$, respectively. Given $m_{01}(x,y)$ and $m_{12}(y,z)$ there is a joint density function $m(x,y,z)$ on $\mathbb{R}^3$ with m_{01} and m_{12} as the marginals on the corresponding subspaces [12, page 208]. We set

$$\boldsymbol{m}(x,y,z)=\frac{\boldsymbol{f}_0(x)}{\mathrm{tr}(\boldsymbol{f}_0(x))}\otimes\frac{\boldsymbol{f}_1(y)}{\mathrm{tr}(\boldsymbol{f}_1(y))}\otimes\frac{\boldsymbol{f}_2(z)}{\mathrm{tr}(\boldsymbol{f}_2(z))}m(x,y,z)$$

and note that it has $\boldsymbol{m}_{01}$ and $\boldsymbol{m}_{12}$ as matrix-valued marginal distributions.

Now, let $\boldsymbol{m}_{02}(x,z)=\frac{\boldsymbol{f}_0(x)}{\mathrm{tr}\boldsymbol{f}_0(x)}\otimes\frac{\boldsymbol{f}_2(z)}{\mathrm{tr}\boldsymbol{f}_2(z)}m_{02}(x,z)$ be the marginal of $\boldsymbol{m}(x,y,z)$ when tracing out the y-component. This $\boldsymbol{m}_{02}(x,z)$ is a possible transference plan between $\boldsymbol{f}_0$ and $\boldsymbol{f}_2$. Hence,

$$\begin{aligned}
&\hat{\mathrm{D}}_{\mathrm{T},\lambda}(\boldsymbol{f}_0,\boldsymbol{f}_2)\\
&\leq\left(\int_{\mathbb{R}^2}\left((x-z)^2+\lambda\|\frac{\boldsymbol{f}_0(x)}{\mathrm{tr}\boldsymbol{f}_0(x)}-\frac{\boldsymbol{f}_2(z)}{\mathrm{tr}\boldsymbol{f}_2(z)}\|_{\mathrm{F}}^2\right)m_{02}dxdz\right)^{\frac{1}{2}}\\
&=\left(\int_{\mathbb{R}^3}\left((x-z)^2+\lambda\|\frac{\boldsymbol{f}_0(x)}{\mathrm{tr}\boldsymbol{f}_0(x)}-\frac{\boldsymbol{f}_2(z)}{\mathrm{tr}\boldsymbol{f}_2(z)}\|_{\mathrm{F}}^2\right)mdxdydz\right)^{\frac{1}{2}}\\
&=\left(\int_{\mathbb{R}^3}\left((x-y+y-z)^2+\lambda\|\frac{\boldsymbol{f}_0(x)}{\mathrm{tr}\boldsymbol{f}_0(x)}-\frac{\boldsymbol{f}_1(y)}{\mathrm{tr}\boldsymbol{f}_1(y)}+\frac{\boldsymbol{f}_1(y)}{\mathrm{tr}\boldsymbol{f}_1(y)}-\frac{\boldsymbol{f}_2(z)}{\mathrm{tr}\boldsymbol{f}_2(z)}\|_{\mathrm{F}}^2\right)mdxdydz\right)^{\frac{1}{2}}\\
&\leq\left(\int_{\mathbb{R}^2}\left((x-y)^2+\lambda\|\frac{\boldsymbol{f}_0(x)}{\mathrm{tr}\boldsymbol{f}_0(x)}-\frac{\boldsymbol{f}_1(y)}{\mathrm{tr}\boldsymbol{f}_1(y)}\|_{\mathrm{F}}^2\right)m_{01}dxdy\right)^{\frac{1}{2}}\\
&+\left(\int_{\mathbb{R}^2}\left((y-z)^2+\lambda\|\frac{\boldsymbol{f}_1(y)}{\mathrm{tr}\boldsymbol{f}_1(y)}-\frac{\boldsymbol{f}_2(z)}{\mathrm{tr}\boldsymbol{f}_2(z)}\|_{\mathrm{F}}^2\right)m_{12}dydz\right)^{\frac{1}{2}}\\
&=\hat{\mathrm{D}}_{\mathrm{T},\lambda}(\boldsymbol{f}_0,\boldsymbol{f}_1)+\hat{\mathrm{D}}_{\mathrm{T},\lambda}(\boldsymbol{f}_1,\boldsymbol{f}_2)
\end{aligned}$$

where the last inequality is due to the metric property of the L_2 distance. □

We also derive the geometric property of the optimal transference plan induced by $\hat{\mathrm{D}}_{\mathrm{T},\lambda}$ as follows.

Proposition 14. *Given $\boldsymbol{f}_0, \boldsymbol{f}_1 \in \mathscr{F}$, let $\boldsymbol{m}$ be the optimal transference plan in* (5.21)*, then $\boldsymbol{m}$ has support on at most a $(2 \cdot \lambda)$-monotonically non-decreasing set.*

Proof. We need to prove that if $\boldsymbol{m}(x_1, y_1) \neq 0$ and $\boldsymbol{m}(x_2, y_2) \neq 0$, then $x_2 > x_1$, $y_1 > y_2$ implies

$$(y_1 - y_2)(x_2 - x_1) \leq 2\lambda. \tag{5.22}$$

Assume that $\boldsymbol{m}$ evaluated at the four points (x_i, y_j) with $i, j \in \{1, 2\}$, is as follows

$$\boldsymbol{m}(x_i, y_j) = m_{ij} \cdot A_i \otimes B_j$$

with

$$A_i = \frac{\boldsymbol{f}_0(x_i)}{\mathrm{tr}(\boldsymbol{f}_0(x_i))},\ B_i = \frac{\boldsymbol{f}_1(y_i)}{\mathrm{tr}(\boldsymbol{f}_1(y_i))},$$

and $m_{11}, m_{22} > 0$. The steps of the proof are similar to those of Proposition 12 detailed in the appendix: first, we assume that Proposition 14 fails and that

$$(y_1 - y_2)(x_2 - x_1) > 2\lambda.$$

Then we show that a smaller cost can be incurred by rearranging the "mass". Consider the situation when $m_{22} \geq m_{11}$. Let $\hat{\boldsymbol{m}}$ be a new transference plan with

$$\begin{aligned}
\hat{\boldsymbol{m}}(x_1, y_1) &= 0 \\
\hat{\boldsymbol{m}}(x_1, y_2) &= (m_{11} + m_{12}) \cdot A_1 \otimes B_2 \\
\hat{\boldsymbol{m}}(x_2, y_1) &= (m_{11} + m_{21}) \cdot A_2 \otimes B_1 \\
\hat{\boldsymbol{m}}(x_2, y_2) &= (m_{22} - m_{11}) \cdot A_2 \otimes B_2.
\end{aligned}$$

Then, $\hat{\boldsymbol{m}}, \boldsymbol{m}$ have the same marginals at the four points, the cost incurred by $\boldsymbol{m}$ is

$$\sum_{i=1}^{2} \sum_{j=1}^{2} m_{ij} \left((x_i - y_j)^2 + \lambda \left\| A_i - B_j \right\|_{\mathrm{F}}^2 \right), \tag{5.23}$$

and the cost incurred by $\hat{\boldsymbol{m}}$ is

$$\begin{aligned}
&(m_{11} + m_{12}) \left((x_1 - y_2)^2 + \lambda \left\| A_1 - B_2 \right\|_{\mathrm{F}}^2 \right) \\
+&(m_{11} + m_{21}) \left((x_2 - y_1)^2 + \lambda \left\| A_2 - B_1 \right\|_{\mathrm{F}}^2 \right) \\
+&(m_{22} - m_{11}) \left((x_2 - y_2)^2 + \lambda \left\| A_2 - B_2 \right\|_{\mathrm{F}}^2 \right).
\end{aligned} \tag{5.24}$$

To show that (5.23) is larger than (5.24), after canceling common terms, it suffices to show that

$$\begin{aligned}&(y_1-x_1)^2+(y_2-x_2)^2+\lambda\|A_1-B_1\|_{\mathrm{F}}^2+\lambda\|A_2-B_2\|_{\mathrm{F}}^2\\ \geq\ &(y_2-x_1)^2+(y_1-x_2)^2+\lambda\|A_1-B_2\|_{\mathrm{F}}^2+\lambda\|A_2-B_1\|_{\mathrm{F}}^2.\end{aligned}$$

However, this holds true since

$$\begin{aligned}&(y_1-x_1)^2+(y_2-x_2)^2+\lambda\|A_1-B_1\|_{\mathrm{F}}^2+\lambda\|A_2-B_2\|_{\mathrm{F}}^2\\ \geq&(y_1-x_1)^2+(y_2-x_2)^2\\ =&(y_1-x_2)^2+(y_2-x_1)^2+2(x_2-x_1)(y_1-y_2)\\ >&(y_1-x_2)^2+(y_2-x_1)^2+4\lambda\\ \geq&(y_1-x_2)^2+(y_1-x_2)^2+\lambda(\|A_1-B_2\|_{\mathrm{F}}^2+\|A_2-B_1\|_{\mathrm{F}}^2).\end{aligned}$$

The case $m_{11}>m_{22}$ proceeds similarly. □

We explain that the metric $\hat{\mathrm{D}}_{\mathrm{T},\lambda}$ fails to have the desired weak-continuity property in the following example. Consider

$$\boldsymbol{f}=\begin{bmatrix}\frac{1}{2}&0\\0&\frac{1}{2}\end{bmatrix}\delta(x)\ \text{ and }\ \boldsymbol{f}_k=\begin{bmatrix}\frac{1}{2}&0\\0&0\end{bmatrix}\delta(x)+\begin{bmatrix}0&0\\0&\frac{1}{2}\end{bmatrix}\delta(x-x_k).$$

Obviously, if $x_k=0$ then $\hat{\mathrm{D}}_{\mathrm{T},\lambda}(\boldsymbol{f},\boldsymbol{f}_k)=0$. If $x_k\neq 0$ and $x_k\to 0$ as $k\to\infty$, we expect $\hat{\mathrm{D}}_{\mathrm{T},\lambda}(\boldsymbol{f},\boldsymbol{f}_k)$ converges to 0 as k increases. But this is not true. Since the matrices $\begin{bmatrix}\frac{1}{2}&0\\0&\frac{1}{2}\end{bmatrix}$ and $\begin{bmatrix}0&0\\0&\frac{1}{2}\end{bmatrix}$ indicate different directions, then there is always a "rotational" cost in $\hat{\mathrm{D}}_{\mathrm{T},\lambda}(\boldsymbol{f},\boldsymbol{f}_k)$. It can be derived that

$$\hat{\mathrm{D}}_{\mathrm{T},\lambda}(\boldsymbol{f},\boldsymbol{f}_k)\geq\frac{\lambda}{2}$$

for any given $\lambda>0$. If $\lambda\neq 0$, then $\hat{\mathrm{D}}_{\mathrm{T},\lambda}(\boldsymbol{f},\boldsymbol{f}_k)$ does not converge to 0 as $k\to\infty$.

5.3 Remarks

We introduce a formulation of the Monge-Kantorovich problem for matrix-valued density functions that have unit trace. In this, we consider the joint matrix-valued density on a tensor product space. The transportation cost consists of a cost of transference of "mass" and a cost of rotating the "direction" of the mass densities. We show that the matrix-valued OMT can be cast as

a convex optimization problem. In a special case when the transference plan has constrained directions, the optimal solution gives rise to a metric. The structure of the optimal transference plan is studied. We show that the optimal transference plan is no longer supported on a monotonically increasing thin set. Though the distance is able to quantify the transference of mass and rotation of directions, it lacks the desired weak-continuity property. In the following, we discuss an idea on the Riemannian structure on the manifold of matrix-valued power spectra.

On a length space structure

Our interest for studying the Monge-Kantorovich geometry of matrix-valued density functions originates in the analysis of slowly varying time-series. The goal is to devise a computational efficient tool that is suitable for comparing matrix-valued spectra, for spectral morphing and tracking. One such tool is provided by the geodesics on the manifold of density functions. The geodesics induced by the 2-Wasserstein metric for scalar-valued densities have already been used for modeling of slowly time-varying processes. However, the Rienmannian structure induced by OMT for matrix-valued densities is not available. For this, we discuss an idea on a Riemannian structure for matrix-valued spectra.

The Riemannian structure induced by the 2-Wasserstein metric was introduced by Benamou and Brenier in [54] (c.f. Chapter 3). For scalar-valued probability densities having support on the real line, the Riemannian metric is explicitly given in (3.20) as

$$\int_{\mathbb{R}} \frac{\dot{F}^2}{f} dx \tag{5.25}$$

where

$$\dot{F}(x) = \int_{-\infty}^{x} \dot{f}(y) dy.$$

Then $\dot{F}(x)$ is a tangent vector of the cumulative function. We consider a generalization of (5.25) to matrix-valued density of the form

$$\int_{\mathbb{R}} \mathrm{tr}(\dot{F} D_{\boldsymbol{f}}(\dot{F})) dx$$

where $D_{\boldsymbol{f}}(\dot{F})$ denotes a noncommutative division of $\dot{F}$ by $\boldsymbol{f}$. Some examples of the noncommutative division are

$$
\begin{aligned}
D_{\boldsymbol{f},1}(\dot{F}) &:= \boldsymbol{f}^{-1}\dot{F} \\
D_{\boldsymbol{f},2}(\dot{F}) &:= \int_0^\infty (\boldsymbol{f}+uI)^{-1}\dot{F}(\boldsymbol{f}+uI)^{-1}du \\
D_{\boldsymbol{f},3}(\dot{F}) &:= X, \text{ where } \frac{1}{2}(\boldsymbol{f}X + X\boldsymbol{f}) = \dot{F}.
\end{aligned}
$$

In the scalar-valued case, these all lead to $\frac{\dot{F}}{\boldsymbol{f}}$. The shortest paths induced by these metrics are natural candidates for the connection of matrix-valued density functions. The explicit expressions for the paths are still need to be explored.

Chapter 6

Generalization of Wasserstein metric in noncommutative geometry

We review some generalizations of the Wasserstein metric in noncommutative geometry. The main tool in noncommutative geometry is Alain Connes's theory of spectral triple [64]. It has been pointed out by Rieffel in [65] that the Wasserstein distance is exactly a special case of Connes's spectral distance. Recently, alternative viewpoints have been proposed as generalizations of the Wasserstein metric in noncommutative geometry [66, 67, 58]. In the following, we present preliminaries of Conne's spectral distance and study the alternative viewpoints by D'Andrea and Martinetti [66] and Martinetti [67]. Their relations with our proposed distances are discussed. We also suggest a family of metrics that relies on suitable set of test functions.

6.1 Lipschitz semi-norm and the Wasserstein metric

Consider a function ϕ on a complete m dimensional Riemannian manifold $\mathscr{M}$. The Lipschitz semi-norm of ϕ is defined as

$$\|\phi\|_{\text{Lip}} := \sup_{x,y} \frac{|\phi(x) - \phi(y)|}{d(x,y)} \tag{6.1}$$

where $d(x,y)$ denotes the geodesic distance between $x, y \in \mathscr{M}$. It is a called a semi-norm since for any constant function ϕ we have $\|\phi\|_{\text{Lip}} = 0$. If ϕ is smooth, then

$$\|\phi\|_{\text{Lip}} = \|\frac{d\phi}{ds}\|_{\infty}$$

with

$$ds = \sqrt{g_{i,j} dx^i dx^j}$$

and g being the Riemannian metric tensor where $g_{i,j} dx^i dx^j$ is simplified for $\sum_{i,j=1}^m g_{i,j} dx^i dx^j$ according to the usual convention. If ϕ is smooth then

$$\phi(x+\delta) = \phi(x) + \partial_x \phi^T \delta + o(\|\delta\|^2)$$

where $\partial_x \phi$ denote the derivative of ϕ. Note that

$$\max_{\delta} \left\{ \partial_x \phi^T \delta \mid g_{i,j} \delta^i \delta^j \leq 1 \right\} = \sqrt{\|g^{i,j} \partial_i \phi \partial_j \phi\|}$$

where $g^{i,j}$ denote the (i,j)th entry of the inverse of Riemannian metric g and ∂_i is short for ∂_{x_i}. Thus the Lipschitz semi-norm of a smooth function equals

$$\|\phi\|_{\text{Lip}}^2 = \sup_x \|g^{i,j} \partial_i \phi \partial_j \phi\| = \|g^{i,j} \partial_i \phi \partial_j \phi\|_\infty .$$

From (6.1), the distance $d(\cdot,\cdot)$ can be recovered from $\|\cdot\|_{\text{Lip}}$ as

$$d(x,y) = \sup_{\phi} \left\{ |\phi(x) - \phi(y)| \ \mid \ \|\phi\|_{\text{Lip}} \leq 1 \right\}. \tag{6.2}$$

Let δ_x and δ_y denote the point masses on x and y, respectively. We denote

$$\delta_x(\phi) := \int \phi(y) \delta_x(y) dy = \phi(x).$$

Then we can also view (6.2) as a distance on between δ_x and δ_y as

$$d(\delta_x, \delta_y) = \sup_{\phi} \left\{ |\delta_x(\phi) - \delta_y(\phi)| \ \mid \ \|\phi\|_{\text{Lip}} \leq 1 \right\}.$$

By generalizing δ_x and δ_y to probability measures, we obtain the Wasserstein metric (3.25). In particular, consider two probability measures μ_0 and μ_1. Let

$$\mu_i(\phi) := \int \phi(x) d\mu_i(x), \text{ for } i = 0, 1.$$

Then

$$d(\mu_0, \mu_1) = \sup_{\phi} \left\{ |\mu_0(\phi) - \mu_1(\phi)| \ \mid \ \|\phi\|_{\text{Lip}} \leq 1 \right\} \tag{6.3}$$

is exactly the Wasserstein metric (3.7).

Alain Connes discovered that the Lipschitz semi-norm can also be defined in algebraic terms [64]. Let D be the Dirac operator

$$D = -i\boldsymbol{\gamma}^k \partial_k \tag{6.4}$$

where $\boldsymbol{\gamma}^k$'s for $k = 1, \ldots, m$, are the self-adjoint Dirac gamma matrices. The matrices $\boldsymbol{\gamma}$'s satisfy

$$\boldsymbol{\gamma}^k \boldsymbol{\gamma}^\ell + \boldsymbol{\gamma}^\ell \boldsymbol{\gamma}^k = 2g^{k\ell} I$$

with I being the identity matrix of dimension 2^n if $m = 2n$ or $2n+2$. Then we have

$$[D, \phi] f = D\phi f - \phi D f = -i\boldsymbol{\gamma}^k \partial_k (\phi f) + i\phi \boldsymbol{\gamma}^k \partial_k f = -i\boldsymbol{\gamma}^k \partial_k \phi f.$$

Thus $[D, \phi]$ acts on f as multiplying by $-i\boldsymbol{\gamma}^k \partial_k \phi$. The operator norm of $[D, \phi]$ equals to the operator norm of $\|\boldsymbol{\gamma}^k \partial_k \phi\|$.

Moreover, we derive the Lipschitz semi-norm as

$$\begin{aligned} \|[D, \phi]\|^2 &= \|\boldsymbol{\gamma}^k \partial_k \phi\|^2 = \|(\boldsymbol{\gamma}^k \partial_k \phi)(\boldsymbol{\gamma}^\ell \partial_\ell \phi)\| \\ &= \|\frac{1}{2}(\boldsymbol{\gamma}^\ell \boldsymbol{\gamma}^\ell + \boldsymbol{\gamma}^\ell \boldsymbol{\gamma}^k) \partial_k \phi \partial_\ell \phi\| = \|g^{k\ell} \partial_k \phi \partial_\ell \phi I\| \\ &= \|g^{k\ell} \partial_k \phi \partial_\ell \phi\|_\infty = \|\phi\|_{\mathrm{Lip}}^2, \end{aligned}$$

see e.g. [66] which implies that

$$\|\phi\|_{\mathrm{Lip}} = \|[D, \phi]\|. \tag{6.5}$$

It is also pointed out in [64, 66] that (6.5) holds for all continuous functions. The operator D that induces Lipschitz semi-norm is not unique.

In noncommutative geometry, Alain Connes introduced the theory of spectral triples and spectral distance which relies on a generalized Dirac operator. It turns out that the Wasserstein metric is a special case of the spectral distance.

6.2 Connes' spectral distance

Let $C(\mathscr{M})$ denote the set of continuous functions on $\mathscr{M}$. For any $\phi, \psi \in C(\mathscr{M})$, $\phi + \psi, \phi\psi \in C(\mathscr{M})$ and any scalar $a \in \mathbb{C}$, we have

$$\phi + \psi,\ \phi\psi,\ a\phi,\ a\psi \in C(\mathscr{M})$$

moreover $\phi\psi = \psi\phi$. Thus, $C(\mathscr{M})$ forms a commutative algebra.

In a noncommutative algebra, the multiplication may not be commutative. For example, the set of $n \times n$ complex-valued matrices $M_n(\mathbb{C})$ forms a noncommutative algebra. This is because for $A, B \in M_n(\mathbb{C})$, AB may not equal to BA. Thus commutative algebra is a special case of noncommutative algebra.

A state $\boldsymbol{\rho}$ of a noncommutative algebra of operators $\mathscr{A}$ is a positive linear functional that maps $\phi \in \mathscr{A}$ to $\mathbb{C}$. We denote $S(\mathscr{A})$ as the set of states of $\mathscr{A}$. For example, for the commutative algebra $C(\mathscr{M})$, a state $\boldsymbol{\rho} \in S(C(\mathscr{M}))$ uniquely corresponds to a probability measure μ such that

$$\boldsymbol{\rho}(\phi) = \int_{\mathscr{M}} \phi d\mu$$

for any $\phi \in C(\mathscr{M})$. In quantum mechanics, for the noncommutative algebra $\mathscr{A}$ of observable of a physical system, a state $\boldsymbol{\rho} \in S(\mathscr{A})$ is a density matrix which is positive semi-definite with trace equals to one. For $\boldsymbol{\phi} \in \mathscr{A}$,

$$\boldsymbol{\rho}(\boldsymbol{\phi}) = \mathrm{tr}(\boldsymbol{\rho}\boldsymbol{\phi}).$$

The set of states $S(\mathscr{A})$ is convex. The extremal points are called the purse states. The pure states in $S(C(\mathscr{M}))$ are the δ_x's, and an extremal point of the density matrices is a rank one matrix. We denote $P(\mathscr{A})$ as the subset of pure states in $S(\mathscr{A})$.

In noncommutative geometry, the distance between states is studied via Connes' theory of spectral triple $(\mathscr{A}, \mathscr{H}, D)$ which includes a involutive algebra of operators $\mathscr{A}$ on a Hilbert space $\mathscr{H}$ and an self-adjoint densely defined operator D such that the commutator $[D, \phi]$ is bounded for all $\phi \in \mathscr{A}$. The operator D is a generalized Dirac operator. The spectral distance between two states $\boldsymbol{\rho}_0, \boldsymbol{\rho}_1 \in S(\mathscr{A})$ is defined as

$$d_D(\boldsymbol{\rho}_0, \boldsymbol{\rho}_1) = \sup_{\boldsymbol{\phi} \in \mathscr{A}} \left\{ |\boldsymbol{\rho}_0(\boldsymbol{\phi}) - \boldsymbol{\rho}_1(\boldsymbol{\phi})| \mid \|[D, \boldsymbol{\phi}]\| \leq 1 \right\} \tag{6.6}$$

where $\|\cdot\|$ denotes the operator norm of the operators in $\mathscr{A}$ on $\mathscr{H}$.

Consider $\mathscr{A}$ as the algebra of continuous functions $C(\mathscr{M})$. Let two states $\boldsymbol{\rho}_0, \boldsymbol{\rho}_1$ be given as probability measures μ_0 and μ_1, respectively. Let D be the Dirac operator as in (6.4). Then

$$d_D(\boldsymbol{\rho}_0, \boldsymbol{\rho}_1) = \sup \left\{ \int_{\mathscr{M}} \phi (d\mu_0 - d\mu_1) \mid \|[D, \phi]\| \leq 1 \right\}$$

is exactly the Wasserstein metric between μ_0 and μ_1 as in (6.3) (see [66] for a rigorous proof). The relation of spectral distance and Wasserstein metric was pointed out by Rieffel in [65].

In the noncommutative setting, to define a spectral distance, an operator D should be specified which is thought as a generalized Dirac operator. Usually, it is still referred to as the Dirac

operator to indicate how it is employed. Then, $\|[D,\boldsymbol{\phi}]\|$ is a generalization of the Lipschitz semi-norm. In the following example, we illustrate that the spectral distance may be infinite.

Example 15. *Consider an noncommutative algebra $\mathscr{A}$ as the 2×2 complex-valued matrices $M_2(\mathbb{C})$. Let the Dirac operator D be given as*

$$D=\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix}.$$

We consider $S(\mathscr{A})$ as the set of positive semi-definite matrices with trace equals to one. A state $\boldsymbol{\rho}$ acts on a $\boldsymbol{\phi}\in\mathscr{A}$ via trace. For simplicity, we only restrict $\boldsymbol{\phi}\in\mathscr{A}$ to be real-valued and denote it as

$$\boldsymbol{\phi}=\begin{bmatrix}a & b\\ c & d\end{bmatrix}.$$

Then

$$[D,\boldsymbol{\phi}]=\begin{bmatrix}c-b & d-a\\ a-d & b-c\end{bmatrix}.$$

Clearly, the Lipschitz semi-norm $\|[D,\boldsymbol{\phi}]\|$ does not change when we add or subtract any matrix of the form $\begin{bmatrix}x & y\\ y & x\end{bmatrix}$ to $\boldsymbol{\phi}$. Consider two states

$$\boldsymbol{\rho}_0=\begin{bmatrix}p_0 & q_0\\ q_0 & 1-p_0\end{bmatrix} \text{ and } \boldsymbol{\rho}_1=\begin{bmatrix}p_1 & q_1\\ q_1 & 1-p_1\end{bmatrix}.$$

Their spectral distance is

$$\begin{aligned} d_D(\boldsymbol{\rho}_0,\boldsymbol{\rho}_1) &= \sup_{\boldsymbol{\phi}\in\mathscr{A}}\{|\mathrm{tr}(\boldsymbol{\rho}_0\boldsymbol{\phi})-\mathrm{tr}(\boldsymbol{\rho}_1\boldsymbol{\phi})| \;\mid\; \|[D,\boldsymbol{\phi}]\|\le 1\}\\ &= \sup\left\{|(p_0-p_1)(a-d)+(q_0-q_1)(b+c)| \;\mid\; \|\begin{bmatrix}c-b & d-a\\ a-d & b-c\end{bmatrix}\|\le 1\right\}. \end{aligned}$$

If $q_0=q_1$, then

$$d_D(\boldsymbol{\rho}_0,\boldsymbol{\rho}_1)=\sup\left\{|(p_0-p_1)(a-d)| \mid \|\begin{bmatrix}c-b & d-a\\ a-d & b-c\end{bmatrix}\|\le 1\right\}=|p_0-p_1|$$

is bounded. If $q_0\neq q_1$, then $d_D(\boldsymbol{\rho}_0,\boldsymbol{\rho}_1)$ is unbounded since we can make $b=c$ to be an arbitrary large number.

To introduce meaningful and bounded distances, we propose some methods that relate the OMT distance we have introduced for densities with unbalanced masses. The basic idea is to introduce a suitable way to bound the set of test functions. Following (3.25) and (4.8), one method is to introduce an additional constraint on the test function by bounding its norm. For example, we introduce the constraint $\|\boldsymbol{\phi}\| \leq \kappa$ with a given $\kappa > 0$ in (6.6) and we consider the problem

$$\sup_{\boldsymbol{\phi}\in\mathscr{A}} \big\{ |\boldsymbol{\rho}_0(\boldsymbol{\phi}) - \boldsymbol{\rho}_1(\boldsymbol{\phi})| \mid \|[D,\boldsymbol{\phi}]\| \leq 1,\ \|\boldsymbol{\phi}\| \leq \kappa \big\}. \tag{6.7}$$

Clearly, this always leads to a bounded distance.

We introduce an alternative idea on a bounded distance that is based on additional Lipschitz semi-norms. Consider the above example. Let

$$\hat{D} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

denote an alternative Dirac operator. Then

$$[\hat{D},\boldsymbol{\phi}] = 2\begin{bmatrix} 0 & b \\ -c & 0 \end{bmatrix}.$$

By adding a constraint $\|[\hat{D},\boldsymbol{\phi}]\| \leq 1$ in (6.6), we have

$$\sup_{\boldsymbol{\phi}\in\mathscr{A}} \big\{ |\mathrm{tr}(\boldsymbol{\rho}_0\boldsymbol{\phi}) - \mathrm{tr}(\boldsymbol{\rho}_1\boldsymbol{\phi})| \ \mid\ \|[D,\boldsymbol{\phi}]\| \leq 1, \|[\hat{D},\boldsymbol{\phi}]\| \leq 1 \big\} \tag{6.8}$$

$$= \sup \left\{ |(p_0-p_1)(a-d)+(q_0-q_1)(b+c)| \ \mid\ \| \begin{bmatrix} c-b & d-a \\ a-d & b-c \end{bmatrix} \| \leq 1, \| \begin{bmatrix} 0 & b \\ -c & 0 \end{bmatrix} \| \leq \frac{1}{2} \right\}. \tag{6.9}$$

This optimal value gives a bounded distance between the states. We note that each Dirac operator gives the freedom of $\boldsymbol{\phi}$ in one direction. Since the directions are not compatible with each other, the values of $a-d$ and $b+c$ are always bounded for any feasible $\boldsymbol{\phi}$.

Apparently, the OMT distance for matrix-valued densities $\mathrm{D_W}$ introduced in Section 6.1 is also closely related to the spectral distance. It is natural to consider the set of matrix-valued spectral measures with unit trace as the set of states. The distance $\mathrm{D_W}$ is unbounded if the states do not have equal integral. The Lipschitz condition

$$\boldsymbol{\phi}(x) - \boldsymbol{\phi}(y) \leq |x-y| I, \forall x, y$$

corresponds to $\|[D,\boldsymbol{\phi}]\| \leq 1$ with D being the Dirac operator. The distance $\mathrm{D_W}$ can be written as

$$\sup_{\boldsymbol{\phi}} \left\{ \int_0^1 \mathrm{tr}\left(\boldsymbol{\phi}(d\boldsymbol{\mu}_0 - d\boldsymbol{\mu}_1)\right) \;\middle|\; \|[D,\boldsymbol{\phi}]\| \leq 1 \right\}. \tag{6.10}$$

Following (6.8), it is interesting to explore the idea of multiple Lipschitz semi-norms in the OMT distance. For this, we explain a preliminary idea based on scalar-valued measures.

We introduce a generalized Lipschitz semi-norm for scalar-valued functions. Given $g \in C(\mathscr{M})$, for $\phi \in C(\mathscr{M})$ we define

$$\|\phi\|_{\mathrm{Lip,g}} := \|\phi g\|_{\mathrm{Lip}}.$$

Clearly, $\|\phi\|_{\mathrm{Lip,g}} = 0$ implies that ϕg is constant. It is considered as a Lipschitz semi-norm with respect to g. Given $g_1, g_2 \in C(\mathscr{M})$, we consider

$$\sup_{\boldsymbol{\phi}} \left\{ \int_0^1 \phi(d\mu_0 - d\mu_1) \;\middle|\; \|\phi\|_{\mathrm{Lip,g_1}} \leq 1, \|\phi\|_{\mathrm{Lip,g_2}} \leq 1 \right\} \tag{6.11}$$

as a generalized Wasserstein metric between μ_0 and μ_1. If $g_1 \neq g_2$, then the feasible set of ϕ is bounded. Thus the optimal value is finite even if μ_0 and μ_1 have different masses which can be considered as a generalized Wasserstein metric.

6.3 D'Andrea-Martinetti's viewpoint

We review the generalization of the Wasserstein metric in the noncommutative geometry presented in [66, 67]. The basic idea in these works is to consider the pure states as the analogy of data points on a manifold. The distance between purrs states are used to define transportation cost and to define Lipschitz semi-norm.

Optimal transport on pure state space

The set of pure states $P(\mathscr{A})$ of the algebra $\mathscr{A}$ contains all the extremal points in $S(\mathscr{A})$. Any state in $S(\mathscr{A})$ can be computed as a convex combination of the pure states. This convex combination is represented as a probability measure on the pure state space. In [66, 67], a Monge-Kantrovich OMT problem is introduced to compare the probability measures. This OMT cost provides a

distance measure between the states. In this, the spectral distance between the pure states is considered as the transportation cost. This idea is elaborated bellow.

Consider $\boldsymbol{\rho}_0 = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} \end{bmatrix}$. It can be decomposed as

$$\boldsymbol{\rho}_0 = \frac{1}{4}\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} + \frac{3}{4}\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} =: \frac{1}{4}\boldsymbol{\rho}_{0,1} + \frac{3}{4}\boldsymbol{\rho}_{0,2}.$$

Thus, a probability measure induced by this decomposition on the set of pure states is given as

$$\mu_0 = \frac{1}{4}\delta(\boldsymbol{\rho} - \boldsymbol{\rho}_{0,1}) + \frac{3}{4}\delta(\boldsymbol{\rho} - \boldsymbol{\rho}_{0,2}).$$

Consider another state $\boldsymbol{\rho}_1 = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$. We decompose it as

$$\boldsymbol{\rho}_1 = \frac{1}{2}\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{bmatrix} + \frac{1}{2}\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} =: \frac{1}{2}\boldsymbol{\rho}_{1,1} + \frac{1}{2}\boldsymbol{\rho}_{1,2}.$$

The corresponding probability measure is

$$\mu_1 = \frac{1}{2}\delta(\boldsymbol{\rho} - \boldsymbol{\rho}_{1,1}) + \frac{1}{2}\delta(\boldsymbol{\rho} - \boldsymbol{\rho}_{1,2}).$$

Then we compare $\boldsymbol{\rho}_0$ and $\boldsymbol{\rho}_1$ by computing the Monge-Kantorovich optimal transportation cost between μ_0 and μ_1. The transportation cost is set as the spectral distance between the corresponding pure states $d_D(\boldsymbol{\rho}_{0,i}, \boldsymbol{\rho}_{1,j})$. Then the OMT cost $T(\mu_0, \mu_1)$ can be used to compare the states. Apparently, a problem in this idea is that the decomposition of a state is not unique. For example, the $\boldsymbol{\rho}_0$ in above can also be decomposed as

$$\boldsymbol{\rho}_0 = \frac{3}{8}\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \frac{5}{8}\begin{bmatrix} \frac{1}{5} & \frac{2}{5} \\ \frac{2}{5} & \frac{4}{5} \end{bmatrix} =: \frac{3}{8}\hat{\boldsymbol{\rho}}_{0,(1)} + \frac{5}{8}\hat{\boldsymbol{\rho}}_{0,(2)}$$

which corresponds to a probability measure

$$\hat{\mu}_0 = \frac{3}{8}\delta(\boldsymbol{\rho} - \hat{\boldsymbol{\rho}}_{0,(1)}) + \frac{5}{8}\delta(\boldsymbol{\rho} - \hat{\boldsymbol{\rho}}_{0,(2)}).$$

Clearly, the OMT cost $T(\mu_0, \hat{\mu}_0)$ is not zero since they corresponds to different pure states. But μ_0 and $\hat{\mu}_0$ corresponds to the same state. Their distance is expected to be zero. In [67], Martinetti suggests to compare $\boldsymbol{\rho}_0$ and $\boldsymbol{\rho}_1$ by minimizing

$$\inf_{\mu_0, \mu_1} W(\mu_0, \mu_1)$$

over all possible probability measures that represent the two states, respectively. It is not clear that whether this defines a metric or not.

Lipschitz semi-norm based on pure state space

An alternative idea in [67] for extending the Wasserstein metric in noncommutative geometry is based on a generalization of the Lipschitz semi-norm. For scalar-valued functions, the set of Lipschitz function on a Riemannian manifold $\mathscr{M}$ is defined as

$$\{\phi \mid \phi \in C(\mathscr{M}), |\phi(x)-\phi(y)| \leq d(x,y), \forall x,y \in \mathscr{M}\}.$$

Consider the noncommutative algebra $\mathscr{A}$ as a generalization of $C(\mathscr{M})$. It is natural to consider the set of pure states $P(\mathscr{A})$ as the noncommutative analogy of the points on a manifold. It suggested in [67] to consider

$$\mathrm{Lip}_D(\mathscr{A}) = \{\boldsymbol{\phi} \mid \boldsymbol{\phi} \in \mathscr{A},\ |\boldsymbol{\rho}_0(\boldsymbol{\phi}) - \boldsymbol{\rho}_1(\boldsymbol{\phi})| \leq d_D(\boldsymbol{\rho}_0, \boldsymbol{\rho}_1),\ \forall \boldsymbol{\rho}_0, \boldsymbol{\rho}_1 \in P(\mathscr{A})\}$$

as the set of Lipschitz operators in $\mathscr{A}$. The generalized Wasserstein metric is then defined by

$$W_D(\boldsymbol{\rho}_0, \boldsymbol{\rho}_1) = \sup_{\phi \in \mathrm{Lip}_D(\mathscr{A})} |\boldsymbol{\rho}_0(\boldsymbol{\phi}) - \boldsymbol{\rho}_1(\boldsymbol{\phi})|.$$

Since

$$\{\boldsymbol{\phi} \mid \boldsymbol{\phi} \in \mathscr{A},\ \|[D, \boldsymbol{\phi}]\| \leq 1\} \subset \mathrm{Lip}_D(\mathscr{A}),$$

then

$$d_D \leq W_D$$

where d_D is the spectral distance in (6.6). It was also shown in [67] that $d_D(\boldsymbol{\rho}_0, \boldsymbol{\rho}_1) = W_D(\boldsymbol{\rho}_0, \boldsymbol{\rho}_1)$ if $\boldsymbol{\rho}_0$ and $\boldsymbol{\rho}_1$ are the convex combination of any given two pure states.

We conclude this chapter with an example discussed in [67]. Consider a spectral triple $(\mathscr{A}, \mathscr{H}, D)$ with spectral distance denoted by d_D. Let $C(\mathscr{M})$ denote the set of continuous functions on a manifold $\mathscr{M}$. Let $d(x,y)$ denote the geodesic distance between $x, y \in \mathscr{M}$. Then the pure states in the product algebra $\mathscr{A}' := C(\mathscr{M}) \otimes \mathscr{A}$ are represented as

$$x_i := (\delta_x, \boldsymbol{\rho}_i)$$

with $\boldsymbol{\rho}_i \in P(\mathscr{A})$. It was suggested in [67] to consider a cost function between two pure states x_0 and y_1 as

$$c(x_0, y_1) := \sqrt{d(x,y)^2 + d_D(\boldsymbol{\rho}_0, \boldsymbol{\rho}_1)^2}.$$

It is interesting to compare this cost function with the one we defined in (5.6). They all consist of the cost in two different spaces.

6.4 Remarks

This chapter presents a survey of different viewpoints on generalization of the Wasserstein metric in noncommutative geometry. The relation between our proposed OMT distance for matrix-valued densities and Conne's spectral distance is discussed. We suggest a family of distances that is based on suitable set of test functions. An alternative viewpoints suggested by D'Andrea and Martinetti are also reviewed. We mention that the Wasserstein metric is also studied in the context of free probability [68]. In [58] an noncommutative generalization of the Wasserstein metric is also introduced to compare positive operators in a Clifford algebra with trace one. The relations between these viewpoints need to be explored in future work.

Chapter 7

Comparison and examples

First, we compare the various distance measures that we introduced in the previous chapters for matrix-valued spectra. Then, we present applications of the proposed distances to multivariate spectral analysis, in particular, OMT distance to the problems of spectral morphing and spectral tracking.

7.1 Comparison of distances of power spectra

Consider three spectral density functions constructed as follows

$$\boldsymbol{f}_0(\theta) = \begin{bmatrix} 1 & 0.4 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0.01 & 0 \\ 0 & \frac{1}{|a_0(e^{\mathrm{j}\theta})|^2} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0.4 & 1 \end{bmatrix}$$

$$\boldsymbol{f}_1(\theta) = \begin{bmatrix} 1 & 0.5 \\ 0.5e^{\mathrm{j}\theta} & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{|a_1(e^{\mathrm{j}\theta})|^2} & 0 \\ 0 & \frac{1}{|a_1(e^{\mathrm{j}\theta})|^2} \end{bmatrix} \begin{bmatrix} 1 & 0.5e^{-\mathrm{j}\theta} \\ 0.5 & 1 \end{bmatrix}$$

$$\boldsymbol{f}_2(\theta) = \begin{bmatrix} 1 & 0 \\ 0.4e^{\mathrm{j}\theta} & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{|a_2(e^{\mathrm{j}\theta})|^2} & 0 \\ 0 & 0.01 \end{bmatrix} \begin{bmatrix} 1 & 0.4e^{-\mathrm{j}\theta} \\ 0 & 1 \end{bmatrix}$$

where

$$a_0(z) = (1 - 1.9\cos(\frac{\pi}{6})z + 0.95^2 z^2)(1 - 1.5\cos(\frac{\pi}{3})z + 0.75^2 z^2)$$
$$a_1(z) = (1 - 1.9\cos(\frac{5\pi}{12})z + 0.95^2 z^2)(1 - 1.5\cos(\frac{\pi}{2})z + 0.75^2 z^2)$$
$$a_2(z) = (1 - 1.9\cos(\frac{2\pi}{3})z + 0.95^2 z^2)(1 - 1.5\cos(\frac{5\pi}{8})z + 0.75^2 z^2)$$

with $\theta \in [0, \pi]$. We scale all the functions by multiplying a scalar so that their integral has unit trace. Typically, the power spectra reflect the dynamic relationship between two time series; in turn these may represent noise input/output of dynamical systems or measurements across independent array of sensors, etc.

The functions $\boldsymbol{f}_0, \boldsymbol{f}_1$ and $\boldsymbol{f}_2$ are shown in Figure 7.1 red dashed line, blue solid line and green dash-dot line, respectively. Since the value of a power spectral density $\boldsymbol{f}$ at each point in frequency is a Hermitian positive semi-definite matrix, our convention is to show in the $(1,1)$, $(1,2)$ and $(2,2)$ subplots the magnitude of the entries $\boldsymbol{f}_{(1,1)}$, $\boldsymbol{f}_{(1,2)}$ (which is the same as $\boldsymbol{f}_{(2,1)}$) and $\boldsymbol{f}_{(2,2)}$, respectively. Since $\boldsymbol{f}_{(1,2)}$ is complex and is the complex conjugate of $\boldsymbol{f}_{(2,1)}$, we plot its phase in the $(2,1)$ subplot.

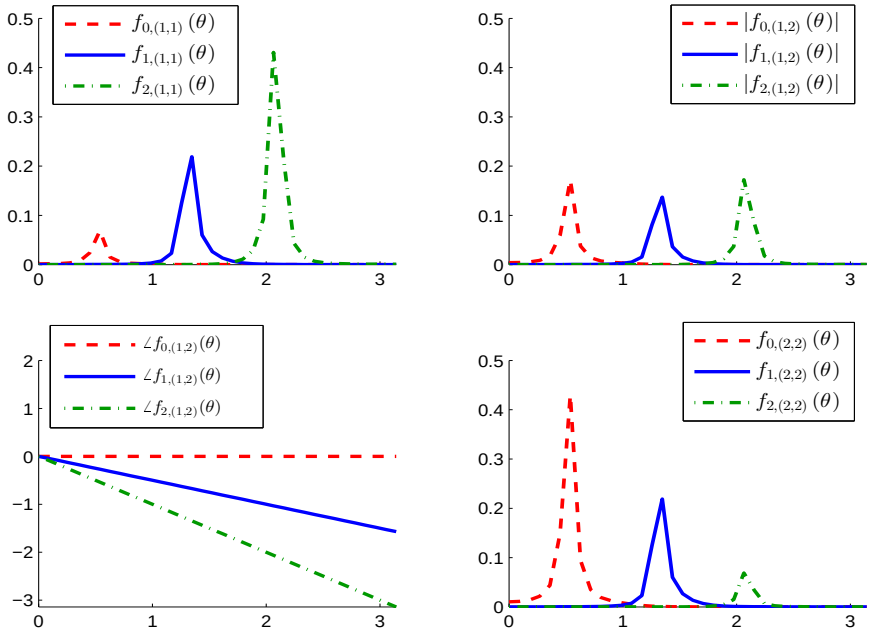

Figure 7.1: Matrix-valued power spectra $\boldsymbol{f}_0$, $\boldsymbol{f}_1$ and $\boldsymbol{f}_2$ are shown in red dashed line, blue solid line and green dashdot line, respectively. Subplots (1,1), (1,2) and (2,2) show $\boldsymbol{f}_{i,(1,1)}, |\boldsymbol{f}_{i,(1,2)}|$ (same as $|\boldsymbol{f}_{i,(2,1)}|$) and $\boldsymbol{f}_{i,(2,2)}$. Subplot (2,1) shows $\angle(\boldsymbol{f}_{i,(1,2)})$ for $i \in \{0,1,2\}$.

We observe that the dominant power in $\boldsymbol{f}_0$ locates at around $\theta = \frac{\pi}{6}$ in the frequency domain

and the majority of the power is in the second channel, i.e. in $\boldsymbol{f}_{0,(2,2)}$. The power of $\boldsymbol{f}_1$ is spread equality in the two channels, and the location of the dominant power is at around $\theta = \frac{5\pi}{12}$. The first channel in $\boldsymbol{f}_2$ contains most of the power. The dominant peak is at around $\theta = \frac{2\pi}{3}$. The cross spectra, $\boldsymbol{f}_{k,(1,2)}$ for $k = 0,1,2$, have similar magnitude. The phase angle of $\boldsymbol{f}_{1,(1,2)}$ is about the average of the phase angles of $\boldsymbol{f}_{1,(1,2)}$ and $\boldsymbol{f}_{2,(1,2)}$.

We use the distances introduced in previous chapters to compare each pair of the three functions. The distances used for this comparison include D_{KL} in (2.2), D_{IS} in (2.4), $\mathrm{D}_{\mathrm{pred}}$ in (2.8), $\mathrm{D}_{\mathrm{flat}}$ in (2.9), D_{H} in (2.13) and the OMT distance $\mathrm{D}_{\mathrm{W},\kappa}$ in (4.7), $\mathrm{D}_{\mathrm{T},\lambda}$ in (5.19) and $\hat{\mathrm{D}}_{\mathrm{T},\lambda}$ in (5.21). We use the parameter $\kappa = \lambda = 1$ in this example. For $\mathrm{D}_{\mathrm{W},\kappa}$, the transportation cost function $c(x,y) = |x-y|$. For completeness, we also denote the total variation distance by

$$\mathrm{D}_{\mathrm{TV}}(\boldsymbol{f}_0, \boldsymbol{f}_1) := \|\boldsymbol{f}_0 - \boldsymbol{f}_1\|_{\mathrm{TV}}.$$

The distances between the spectra are tabulated in Table 7.1.

	Distance between power spectra		
	$\boldsymbol{f}_0, \boldsymbol{f}_1$	$\boldsymbol{f}_1, \boldsymbol{f}_2$	$\boldsymbol{f}_0, \boldsymbol{f}_2$
D_{KL}	5.86	6.92	9.36
D_{IS}	3.44×10^3	5.36×10^4	9.27×10^4
$\mathrm{D}_{\mathrm{pred}}$	3.90×10^2	3.80×10^2	4.40×10^2
$\mathrm{D}_{\mathrm{flat}}$	8.86×10^4	5.56×10^4	3.22×10^5
D_{H}	1.31	1.32	1.40
D_{TV}	1.95	1.96	2.00
$\mathrm{D}_{\mathrm{W},1}$	1.37	1.65	2.29
$\mathrm{D}_{\mathrm{T},1}$	1.01	1.09	2.05
$\hat{\mathrm{D}}_{\mathrm{T},1}$	0.93	0.97	1.86

Table 7.1: Distance measures between power spectra.

We observe that for $\mathrm{D}_{\mathrm{pred}}$, D_{H}, D_{TV} the distances among the three pairs are close to each other. The distance D_{IS} gives rather distorted view of the geometry since we expect similar distances for the two pairs $\boldsymbol{f}_0, \boldsymbol{f}_1$ and $\boldsymbol{f}_1, \boldsymbol{f}_2$. For the distance $\mathrm{D}_{\mathrm{flat}}$, $\mathrm{D}_{\mathrm{flat}}(\boldsymbol{f}_0, \boldsymbol{f}_2)$ is much larger than the distance between the other two pairs. For the generalized relative entropy, $\mathrm{D}_{\mathrm{KL}}(\boldsymbol{f}_0, \boldsymbol{f}_1)$ is close to $\mathrm{D}_{\mathrm{KL}}(\boldsymbol{f}_1, \boldsymbol{f}_2)$ and they both are smaller then $\mathrm{D}_{\mathrm{KL}}(\boldsymbol{f}_0, \boldsymbol{f}_2)$. Similar behavior is also

found for the OMT distance $\mathrm{D}_{\mathrm{W},1}$. For the other two OMT distances $\mathrm{D}_{\mathrm{T},1}$ and $\hat{\mathrm{D}}_{\mathrm{T},1}$, the two pairs $\boldsymbol{f}_0, \boldsymbol{f}_1$ and $\boldsymbol{f}_1, \boldsymbol{f}_2$ have similar distances and they are approximated one half of the distance between $\boldsymbol{f}_0$ and $\boldsymbol{f}_2$. This suggests that $\boldsymbol{f}_1$ is close to be a "mean" value between $\boldsymbol{f}_0$ and $\boldsymbol{f}_2$.

From this comparison, it is shown that the OMT distances we proposed are suitable to quantify the shift of spectral content over frequency and the rotation over channels.

7.2 Spectral morphing

We highlight the relevance of the matrix-valued OMT to spectral analysis by presenting a numerical example of spectral morphing. The idea is to model slowly time-varying changes in the spectral domain by geodesics in a suitable geometry (see e.g., [22, 29]). As noted earlier, the importance of OMT stems from the fact that it induces a weakly continuous metric. Thereby, geodesics smoothly shift spectral power across frequencies lessening the possibility of a fade-in fade-out phenomenon. The classical theory of OMT allows constructing such geodesics for scalar-valued distributions. The example below demonstrates analogous construction of geodesics of matrix-valued power spectra.

Starting with $\boldsymbol{f}_0, \boldsymbol{f}_1 \in \mathscr{F}$ we approximate the geodesic between them by identifying $N-1$ points between the two. More specifically, we set $\boldsymbol{f}_{\tau_0} = \boldsymbol{f}_0$ and $\boldsymbol{f}_{\tau_N} = \boldsymbol{f}_1$, and determine $\boldsymbol{f}_{\tau_k} \in \mathscr{F}$ for $k = 1, \ldots, N-1$ by solving

$$\min_{\boldsymbol{f}_{\tau_k}, 0<k<N} \sum_{k=0}^{N-1} \mathscr{T}_\lambda(\boldsymbol{f}_{\tau_{k+1}}, \boldsymbol{f}_{\tau_k}) \tag{7.1}$$

where $\mathscr{T}_\lambda$ is the optimal transportation cost defined in (5.8). As noted in Chapter 5, numerically this can be solved via convex programming. For computational efficiency, we will solve (7.1) approximately in a hierarchical manner where we compute $\boldsymbol{f}_{\frac{1}{2}}$ using $\boldsymbol{f}_0$ and $\boldsymbol{f}_1$ then compute $\boldsymbol{f}_{\frac{1}{4}}$ using $\boldsymbol{f}_0$ and $\boldsymbol{f}_{\frac{1}{2}}$ and so on. The numerical example is based on two matrix-valued power spectral densities:

$$\boldsymbol{f}_0 = \begin{bmatrix} 1 & 0 \\ 0.2e^{-\mathrm{j}\theta} & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{|a_0(e^{\mathrm{j}\theta})|^2} & 0 \\ 0 & 0.01 \end{bmatrix} \begin{bmatrix} 1 & 0.2e^{\mathrm{j}\theta} \\ 0 & 1 \end{bmatrix}$$

$$\boldsymbol{f}_1 = \begin{bmatrix} 1 & 0.2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0.01 & 0 \\ 0 & \frac{1}{|a_1(e^{\mathrm{j}\theta})|^2} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0.2 & 1 \end{bmatrix}$$

with

$$a_0(z) = (z^2 - 1.8\cos(\frac{\pi}{4})z + 0.9^2)(z^2 - 1.4\cos(\frac{\pi}{3})z + 0.7^2)$$
$$a_1(z) = (z^2 - 1.8\cos(\frac{\pi}{6})z + 0.9^2)(z^2 - 1.5\cos(\frac{2\pi}{15})z + 0.75^2),$$

shown in Figure 7.2. Since the value of a power spectral density at each point in frequency is a 2×2 Hermitian matrix, we have used the $(1,1)$, $(1,2)$, and $(2,2)$ subplots to display the magnitude of the corresponding entries, i.e., $|\boldsymbol{f}_{(1,1)}|$, $|\boldsymbol{f}_{(1,2)}|$ $(= |\boldsymbol{f}_{(2,1)}|)$ and $|\boldsymbol{f}_{(2,2)}|$, respectively, and the $(2,1)$ subplot to display the phase $\angle \boldsymbol{f}_{(1,2)}$ $(= -\angle \boldsymbol{f}_{(2,1)})$.

We denote the solution of (7.1) as $\boldsymbol{f}_{\text{Transp}}$ which is an approximation of a geodesic. The three dimensional plots in Figure 7.3 show a solution $\boldsymbol{f}_{\text{Transp}}$ with $\lambda = 0.1$. The two boundary plots represent the power spectra $\boldsymbol{f}_0$ and $\boldsymbol{f}_1$ shown in blue and red, respectively, using the same convention about magnitudes and phases. There are in total 7 power spectra $\boldsymbol{f}_{\tau_k}$, $k = 1, \ldots, 7$ shown along the geodesic between $\boldsymbol{f}_0$ and $\boldsymbol{f}_1$, and the time-indices correspond to $\tau_k = \frac{k}{8}$. It is interesting to observe the smooth shift of the energy over the geodesic path from the one "channel" to the other while, at the same time, the corresponding peak shifts from one frequency to another. One should bear in mind that the so-constructed geodesic is a non-parametric path interpolating/linking the given spectra.

For comparison, we also use the geodesics that were derived in Chapter 2 to deform the spectra from $\boldsymbol{f}_0$ to $\boldsymbol{f}_1$. Using (2.19), we compute the geodesic induced by the Fisher-Rao metric as

$$\boldsymbol{f}_{\text{Fisher},\tau_k} = \boldsymbol{f}_0^{\frac{1}{2}} (\boldsymbol{f}_0^{-\frac{1}{2}} \boldsymbol{f}_1 \boldsymbol{f}_0^{-\frac{1}{2}})^{\tau_k} \boldsymbol{f}_0^{\frac{1}{2}}. \tag{7.2}$$

Using (2.23), the geodesic induced by the generalized Hellinger distance is computed as

$$\boldsymbol{f}_{\text{H},\tau_k} = \left((1-\tau_k)\boldsymbol{f}_0^{\frac{1}{2}} + \tau_k \boldsymbol{f}_0^{-\frac{1}{2}} (\boldsymbol{f}_0^{\frac{1}{2}} \boldsymbol{f}_1 \boldsymbol{f}_0^{\frac{1}{2}})^{\frac{1}{2}} \right) \left((1-\tau_k)\boldsymbol{f}_0^{\frac{1}{2}} + \tau_k \boldsymbol{f}_0^{-\frac{1}{2}} (\boldsymbol{f}_0^{\frac{1}{2}} \boldsymbol{f}_1 \boldsymbol{f}_0^{\frac{1}{2}})^{\frac{1}{2}} \right)^*. \tag{7.3}$$

The interpolated spectra using (7.2) and (7.3) are shown in Figure 7.4 and Figure 7.5, respectively. We note that the main difference from the results shown in Figure 7.3 is that the spectra $\boldsymbol{f}_{\text{FR},\tau_k}$ and $\boldsymbol{f}_{\text{H},\tau_k}$ have the fade-in and fade-out effect on the geodesics. For $\boldsymbol{f}_{\text{FR},\tau_k}$ and $\boldsymbol{f}_{\text{H},\tau_k}$, the power disappears in one frequency then it appears in a different frequency. The decreasing

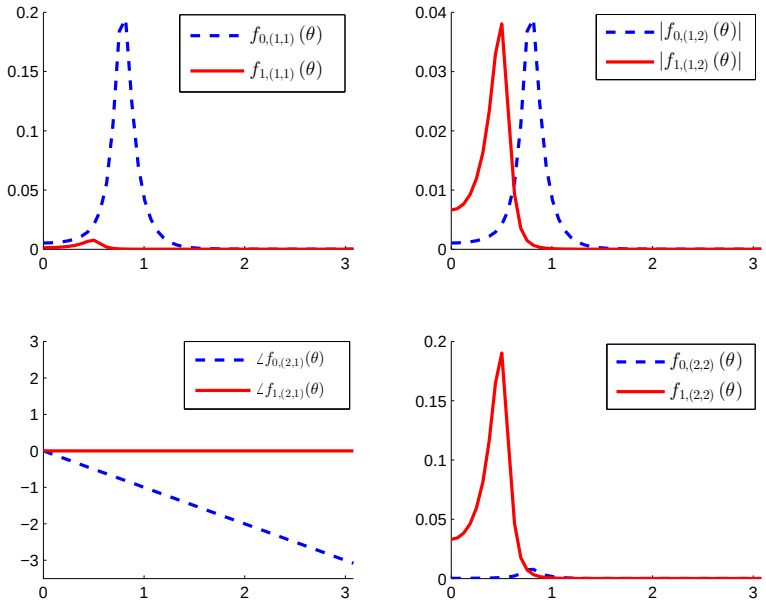

Figure 7.2: Two power spectra $\boldsymbol{f}_0$ and $\boldsymbol{f}_1$ are shown by blue solid line and red dashed line, respectively. Subplots (1,1), (1,2) and (2,2) show $\boldsymbol{f}_{i,(1,1)}, |\boldsymbol{f}_{i,(1,2)}|$ (same as $|\boldsymbol{f}_{i,(2,1)}|$) and $\boldsymbol{f}_{i,(2,2)}$. Subplot (2,1) shows $\angle(\boldsymbol{f}_{i,(1,2))}$ for $i \in \{0,1\}$.

rate of $\boldsymbol{f}_{\mathrm{FR},\tau_k}$ is faster than the rate of $\boldsymbol{f}_{\mathrm{H},\tau_k}$. There is no smooth shifting of the contents over frequency and over channels.

7.3 Regularization with geodesics

Consider two time-series

$$x_1(t) = a_1(t)\cos(\theta_1(t)t + \phi_{1a}) + a_2(t)\cos(\theta_2(t)t + \phi_{1b}) + w_1(t)$$
$$x_2(t) = a_2(t)\cos(\theta_1(t)t + \phi_{2a}) + a_1(t)\cos(\theta_2(t)t + \phi_{2b}) + w_2(t)$$

for $t = 0, 1, \ldots, 2000$, both consisting of sinusoidal signals with time-varying amplitude and frequency together with additive white (discrete-time) noise $w_1(t)$ and $w_2(t)$. Amplitude $a_1(t)$ decreases from 1.2 to 0.1 while $a_2(t)$ increases from 0.1 to 1.2. Frequency $\theta_1(t)$ decreases from $\frac{\pi}{4}$ to $\frac{\pi}{4} - \frac{\pi}{30}$ while $\theta_2(t)$ increases from $\frac{\pi}{3}$ to $\frac{\pi}{3} + \frac{\pi}{30}$. Then $[w_1(t), w_2(t)]'$ is white (independent)

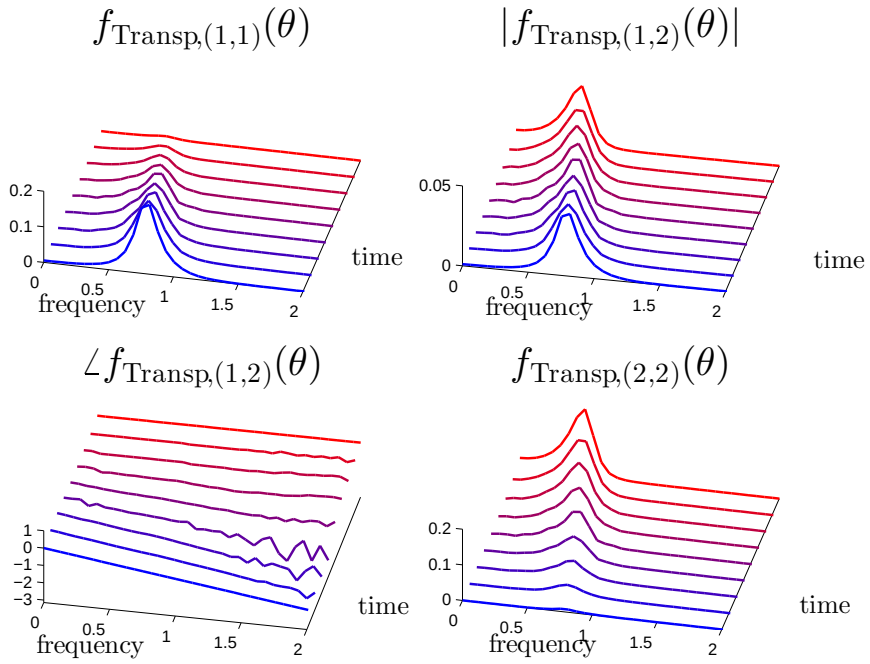

Figure 7.3: The interpolated results $\boldsymbol{f}_{\tau_k}$ for $k = 0,\ldots,8$ computed from (7.1) with $\boldsymbol{f}_0$ and $\boldsymbol{f}_1$ as the two boundary points: subplots (1,1), (1,2) and (2,2) show $\boldsymbol{f}_{\tau_k,(1,1)}, |\boldsymbol{f}_{\tau_k,(1,2)}|$ (same as $|\boldsymbol{f}_{\tau_k,(2,1)}|$) and $\boldsymbol{f}_{\tau_k,(2,2)}$, subplot (2,1) shows $\angle(\boldsymbol{f}_{\tau_k,(2,1)})$.

and sampled from a zero-mean Gaussian distribution with covariance $\begin{bmatrix} 3 & 1.5 \\ 1.5 & 3 \end{bmatrix}$. The initial phases of the sinusoids are randomly selected in $[0, 2\pi]$.

Since we are dealing with a non-stationary situation, we truncate the observed time-series to segments of length equal to 200. For each of these segments, we estimate the 2×2 matrix-valued PSD for the 2-vector-valued time-series using an autoregressive model of order 20. Thus, we have obtained 10 PSD's from the complete observation record which are denoted as $\hat{\boldsymbol{f}}_{\tau_k}, k = 1,\ldots,10$ with $\tau_0 = 0$ and $\tau_{10} = 1$. The spectrogram of the whole observation record which are shown in Figure 7.6.

We use an OMT-geodesic to regularize the estimated PSD's. This idea was proposed and carried out in [22] for scalar time-series and PSD's. For the present matrix-valued setting, the

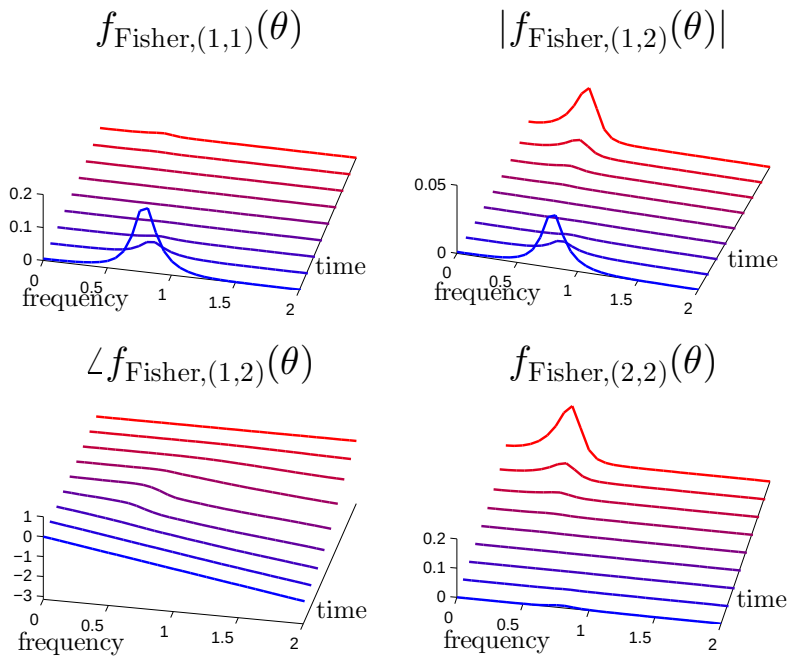

Figure 7.4: The interpolated results $\boldsymbol{f}_{\text{Fisher},\tau_k}$ for $k = 0,\ldots,8$ computed from (7.2): subplots (1,1), (1,2) and (2,2) show $\boldsymbol{f}_{\text{Fisher},(1,1)}, |\boldsymbol{f}_{\text{Fisher},(1,2)}|$ (same as $|\boldsymbol{f}_{\text{Fisher},(2,1)}|$) and $\boldsymbol{f}_{\text{Fisher},(2,2)}$, subplot (2,1) shows $\angle(\boldsymbol{f}_{\text{Fisher},(2,1)})$.

geodesic is obtained by solving

$$\min_{\boldsymbol{f}_{\tau_k}} \left\{ \sum_{k=1}^{10} \mathscr{T}_\lambda(\boldsymbol{f}_{\tau_k}, \hat{\boldsymbol{f}}_{\tau_k}) \mid \boldsymbol{f}_{\tau_k} \text{ are on an OMT geodesic} \right\}. \tag{7.4}$$

An explicit formula of the OMT geodesic is not available. However, in view of Proposition 12, (7.4) can be approximated for small λ as follows.

Let $\hat{f}_{\tau_k} = \text{tr}(\hat{\boldsymbol{f}}_{\tau_k})$ for $k = 1,\ldots,10$. These are scalar-valued PSD's. Let $\hat{F}_{\tau_k}$ denote the corresponding cumulative distribution functions. For λ small, following [22], we compute $f_0 := \text{tr}(\boldsymbol{f}_0)$ and $f_1 := \text{tr}(\boldsymbol{f}_1)$ via solving

$$\min_{\mu_0,\mu_1} \sum_{k=1}^{10} \int_0^1 ((1-\tau_k)F_0^{-1}(v) + \tau_k F_1^{-1}(v) - \hat{F}_{\tau_k}^{-1}(v))^2 dv$$

with F_0 and F_1 representing the cumulative distribution function of f_0 and f_1, respectively. Then,

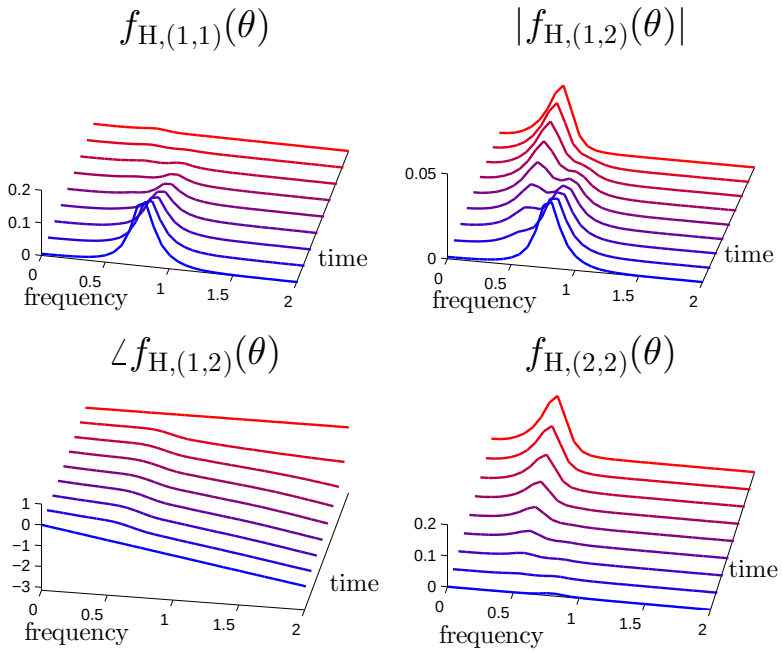

Figure 7.5: The interpolated results $\boldsymbol{f}_{\mathrm{H},\tau_k}$ for $k = 0,\ldots,8$ computed from (7.3): subplots (1,1), (1,2) and (2,2) show $\boldsymbol{f}_{\tau_k,(1,1)}, |\boldsymbol{f}_{\tau_k,(1,2)}|$ (same as $|\boldsymbol{f}_{\tau_k,(2,1)}|$) and $\boldsymbol{f}_{\tau_k,(2,2)}$, subplot (2,1) shows $\angle(\boldsymbol{f}_{\tau_k,(2,1)})$.

as was shown in [22], the f_{τ_k}'s for $1 < k < 10$ can be computed via

$$F_{\tau_k}^{-1}(v) = (1-\tau_k)F_0^{-1}(v) + \tau_k F_1^{-1}(v).$$

The matrix-valued PSD's $\boldsymbol{f}_0$ and $\boldsymbol{f}_1$ are obtained by solving

$$\min_{\boldsymbol{f}_0,\boldsymbol{f}_1} \sum_{k=1}^{10} \int_0^1 \|(1-\tau_k)\frac{\boldsymbol{f}_0(F_0^{-1}(v))}{f_0(F_0^{-1}(v))} + \tau_k \frac{\boldsymbol{f}_1(F_1^{-1}(v))}{f_1(F_1^{-1}(v))} - \frac{\hat{\boldsymbol{f}}_{\tau_k}(\hat{F}_{\tau_k}^{-1}(v))}{\hat{f}_{\tau_k}(\hat{F}_{\tau_k}^{-1}(v))}\|_{\mathrm{F}}^2 dv$$

and the $\boldsymbol{f}_{\tau_k}$'s for $1 < k < 10$ are computed via

$$\frac{\boldsymbol{f}_{\tau_k}(F_{\tau_k}^{-1}(v))}{f_{\tau_k}(F_{\tau_k}^{-1}(v))} = (1-\tau_k)\frac{\boldsymbol{f}_0(F_0^{-1}(v))}{f_0(F_0^{-1}(v))} + \tau_k \frac{\boldsymbol{f}_0(F_1^{-1}(v))}{f_0(F_1^{-1}(v))}.$$

We display this geodesic-fitted spectrogram in Figure 7.7. It can be seen that the shift of energy from one channel to another and between resonant frequencies is smoother than that shown in

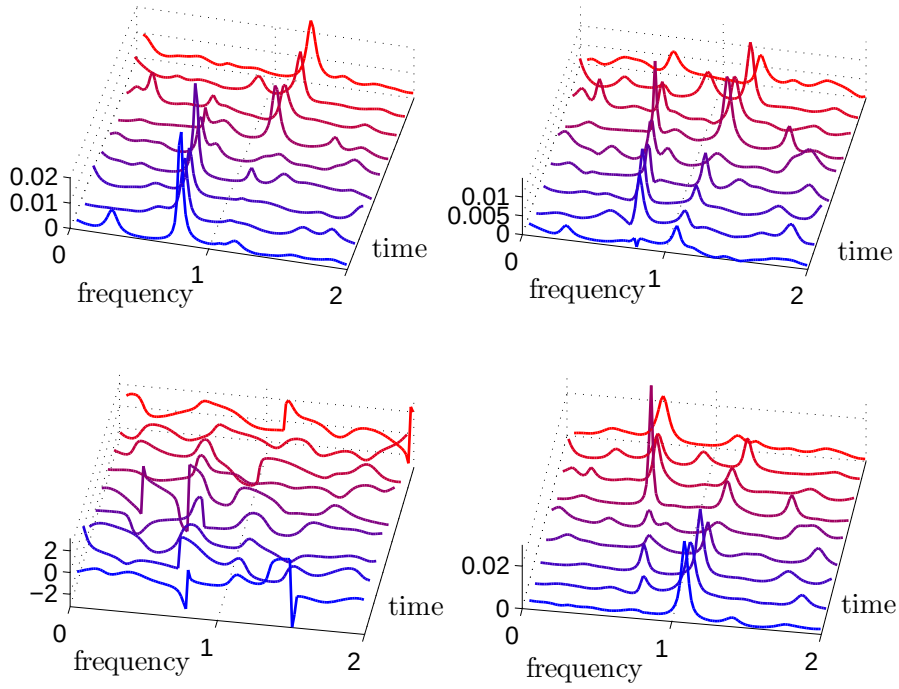

Figure 7.6: The estimated spectrogram of the observed time-series: subplots (1,1), (1,2) and (2,2) show $\boldsymbol{f}_{\tau_k,(1,1)}$, $|\boldsymbol{f}_{\tau_k,(1,2)}|$ (same as $|\boldsymbol{f}_{\tau_k,(2,1)}|$) and $\boldsymbol{f}_{\tau_k,(2,2)}$, subplot (2,1) shows $\angle(\boldsymbol{f}_{\tau_k,(2,1)})$.

Figure 7.6 (which is a spectrogram based on auto-regressive models). The comparison of the directionality of the power between the two (spectrogram vs. geodesic) discussed next is even more revealing.

For each $\hat{\boldsymbol{f}}_{\tau_k}(\theta)$, we find two frequencies θ_1 and θ_2 where the power spectral densities (PSD's) have locally maximal power, i.e. the two frequencies where $\mathrm{tr}(\hat{\boldsymbol{f}}_{\tau_k}(\theta))$ has the largest peaks. Then we compute the (normalized) eigenvectors corresponding to the dominant eigenvalues of $\boldsymbol{f}_{\tau_k}(\theta_1)$ and $\boldsymbol{f}_{\tau_k}(\theta_2)$, respectively. These eigenvectors are shown in Figure 7.8 using black dashed lines. The red and green plots in Figure 7.8 represent the path of the two eigenvectors as τ_k increases from 0 to 1. The axes in Figure 7.8 correspond to the two channels/components of the time-series and τ_k. (Should all the power be present in one of the two channels, the eigenvector would line up, accordingly.) The values of the eigenvector when projected onto

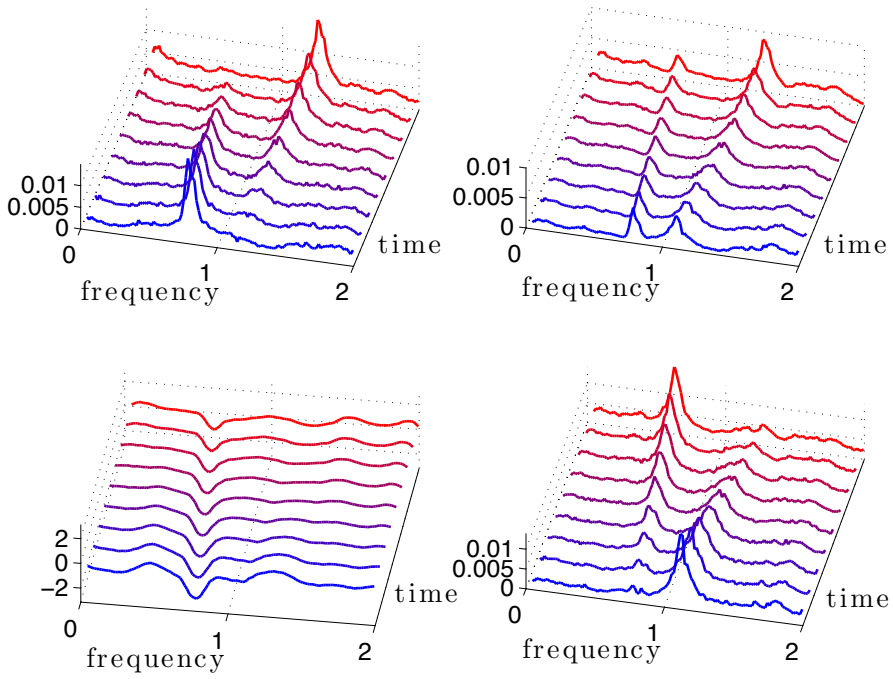

Figure 7.7: The the geodesic-fitted spectrogram: subplots (1,1), (1,2) and (2,2) show $\boldsymbol{f}_{\tau_k,(1,1)}$, $|\boldsymbol{f}_{\tau_k,(1,2)}|$ (same as $|\boldsymbol{f}_{\tau_k,(2,1)}|$) and $\boldsymbol{f}_{\tau_k,(2,2)}$, subplot (2,1) shows $\angle(\boldsymbol{f}_{\tau_k,(2,1)})$.

the two channels/axes, reflect the energy of the signals in the corresponding channels. Thus, in antenna-array applications, the direction of eigenvectors correspond to the direction of a scatterer relative to the array. Statistical errors are reflected in the jagged nature of the paths when these are based on a spectrogram as in Figure 7.8. However, when comparing with the eigenvectors of the regularized spectrogram $\boldsymbol{f}_{\tau_k}$, the corresponding paths shown in Figure 7.9 are smooth. Direct comparison between Figure 7.9 and Figure 7.8 highlights the potential advantages of using geodesic fitting to regularize power distribution in non-stationary time-series.

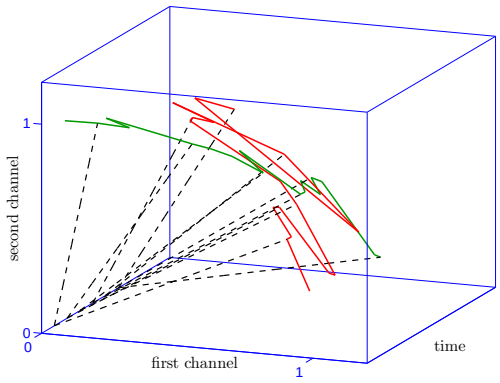

Figure 7.8: The trajectories of the dominant eigenvector of $\hat{\boldsymbol{f}}_{\tau_k}(\theta_1)$ and $\hat{\boldsymbol{f}}_{\tau_k}(\theta_2)$ are shown in red and green, respectively.

7.4 Remarks

Several applications of using the proposed OMT distances in the context of multivariate spectral analysis are discussed. First, we compare the our proposed distances with several other alternatives that are devised from statistics, information theory and quantum mechanics for matrix-valued power spectra. Naturally, our proposed OMT distance are more suitable to quantify the shift of content over frequencies. We apply the proposed distance to problems of spectral morphing and spectral tracking and demonstrate the its ability of tracking the shift of power across frequencies as well as the rotation of power contact between channels.

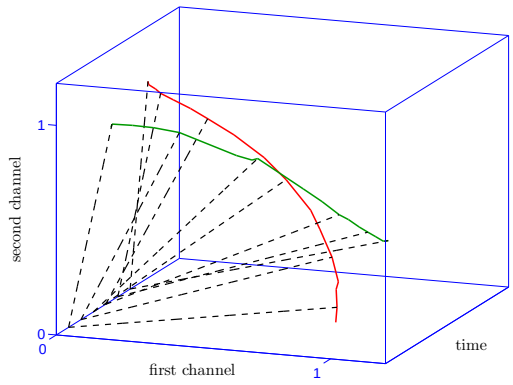

Figure 7.9: The trajectories of the dominant eigenvector of the geodesic-fitted spectrogram $\boldsymbol{f}_{\tau_k}(\theta_1)$ and $\boldsymbol{f}_{\tau_k}(\theta_2)$ are shown in red and green, respectively.

Chapter 8

Conclusion and Discussion

The subject of this thesis is the geometry of matrix-valued density functions. Our interest originates in spectral analysis of multivariate stochastic process and stems from the need for weakly-continuous geometric tools that could be used to quantify, interpolate and approximate matrix-valued power spectra. We first study several distance measures that are motivated from information theory, quantum mechanics and optimal prediction theory. These include a generalized Kullback-Leiber relative entropy, Itakura-Saito distance, Hellinger distance and the distance we have devised based on the optimal prediction theory. Then we present a generalization of the optimal mass transportation distance for matrix-valued density functions which is the main contribution of this thesis.

We propose two frameworks of the OMT problem for matrix-valued density functions. The fist one is based on the assumption that the densities have equal integral. We introduce a notion of non-positive semi-definite matrix-valued transportation plan and use the nuclear norm as representative "mass". This leads to a generalization of the Kantorovich-Rubinstein problem. We then compare densities with unbalanced masses and show that the distance has the weak-continuity property.

Our second formulation of matrix-valued OMT distance is derived under the assumption that their integral have equal trace. In this, we find that it is natural to consider the joint matrix-valued densities in a tensor product space. We introduce a transportation cost that includes a cost of transference of spectral content over frequencies together with a cost of rotating the principle directions. The transportation distance is obtained by solving a convex optimization problem. We show that the corresponding optimal transportation plan is no longer supported on

a monotonic increasing thin set.

The applications of the proposed distances are explored in multivariate spectral analysis. By comparing the various distances for matrix-valued spectra, we show that OMT-based distance is more suitable to quantify shift of content over frequencies. We then demonstrate the ability of the proposed distance in tracking the shift of power across frequencies as well as the rotation of power content between channels.

We note that the applications of OMT distance is not limited to spectral analysis. Many problems in signal processing, fluid dynamics, meteorology are closely related to optimal mass transport. For example, the OMT distance has been used in variational data assimilation to quantify errors between states [60] and in image processing to study image deformation [69]. The proposed OMT-based distance for matrix-valued densities is also a natural tool in the area of diffusion tensor imaging (DTI) in diffusion magnetic resonance imaging. In DTI, a positive semi-definite tensor of size 3×3 is used to model the orientation of the fiber population in each voxel of the tissues image (see e.g. [70]). Since the diffusion tensor images can be thought of as matrix-valued functions, the OMT-based distances we proposed in this thesis is suitable to compare images. Moreover, we can also compare each pair of tensors using the OMT distance between zero-mean Gaussian distributions. These applications will be explored in future.

References

[1] M. Basseville. Distance measures for signal processing and pattern recognition. *Signal processing*, 18(4):349–369, 1989.

[2] R. Gray, A. Buzo, A. Jr. Gray, and Y. Matsuyama. Distortion measures for speech processing. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 28(4):367–376, 1980.

[3] T. Kailath. The divergence and Bhattacharyya distance measures in signal selection. *IEEE Transactions on Communication Technology*, 15(1):52–60, 1967.

[4] Y. Rathi, N. Vaswani, A. Tannenbaum, and A. Yezzi. Tracking deforming objects using particle filtering for geometric active contours. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 29(8):1470–1475, 2007.

[5] A. Sanfeliu and K.-S. Fu. A distance measure between attributed relational graphs for pattern recognition. *IEEE Transactions on Systems, Man and Cybernetics*, SMC-13(3):353–362, 1983.

[6] A. Banerjee, S. Merugu, I.S. Dhillon, and J. Ghosh. Clustering with Bregman divergences. *The Journal of Machine Learning Research*, 6:1705–1749, 2005.

[7] C.R. Rao. Information and the accuracy attainable in the estimation of statistical parameters. *Bull. Calcutta Math. Soc.*, 37:81–91, 1945.

[8] S-I. Amari and H. Nagaoka. *Methods of information geometry*. Amer. Math. Soc., 2000.

[9] N.N. Cencov. *Statistical decision rules and optimal inference*. Amer. Math. Soc., 1982.

[10] R.E. Kass and P.W. Vos. *Geometrical foundations of asymptotic inference*. Wiley New York, 1997.

[11] D. Petz. *Quantum Information Theory and Quantum Statistics (Theoretical and Mathematical Physics)*. Berlin: Springer, 2008.

[12] C. Villani. *Topics in optimal transportation*, volume 58. American Mathematical Society, 2003.

[13] P. Stoica and R.L. Moses. *Introduction to spectral analysis*, volume 1. Prentice hall New Jersey:, 1997.

[14] T.T. Georgiou. Spectral analysis based on the state covariance: the maximum entropy spectrum and linear fractional parametrization. *IEEE Transactions on Automatic Control*, 47(11):1811–1823, 2002.

[15] T.T. Georgiou. The structure of state covariances and its relation to the power spectrum of the input. *IEEE Transactions on Automatic Control*, 47(7):1056–1066, 2002.

[16] A.N. Amini and T.T. Georgiou. Tunable line spectral estimators based on state-covariance subspace analysis. *IEEE Transactions on Signal Processing*, 54(7):2662–2671, 2006.

[17] T.T. Georgiou. The Carathéodory–Fejér–Pisarenko decomposition and its multivariable counterpart. *IEEE Transactions on Automatic Control*, 52(2):212–228, 2007.

[18] T.S. Rao. The fitting of non-stationary time-series models with time-dependent parameters. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 312–322, 1970.

[19] A. Kaderli and A.S. Kayhan. Spectral estimation of nonstationary ARMA processes using the evolutionary cepstrum. *IEEE Signal Processing Letters*, 9(4):130–132, 2002.

[20] D. Rudoy and T.T. Georgiou. Regularized parametric models of nonstationary processes. In *Proc. the 19th International Symposium on Mathematical Theory of Networks and Systems (MTNS2010)*, pages 5–9, 2010.

[21] L. Cohen. *Time-frequency analysis*. Englewood Cliffs, NJ: Prentice Hall, 1995.

[22] X. Jiang, ZQ. Luo, and T.T. Georgiou. Geometric methods for spectral analysis. *IEEE Transactions on Signal Processing*, 60(3):1064–1074, 2012.

[23] T.T. Georgiou. Distance and Riemannian metrics for spectral density functions. *IEEE Transactions on Signal Processing*, 55(8):3995–4003, 2007.

[24] A. Ferrante, C. Masiero, and M. Pavon. Time and spectral domain relative entropy: A new approach to multivariate spectral estimation. *IEEE Transactions on Automatic Control*, 57(10):2561–2575, 2012.

[25] A. Ferrante, M. Pavon, and F. Ramponi. Hellinger versus Kullback–Leibler multivariable spectrum approximation. *IEEE Transactions on Automatic Control*, 53(4):954–967, 2008.

[26] B. Musicus and R. Johnson. Multichannel relative-entropy spectrum analysis. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 34(3):554–564, 1986.

[27] P. Enqvist and J. Karlsson. Minimal Itakura-Saito distance and covariance interpolation. In *47th IEEE Conference on Decision and Control*, pages 137–142, 2008.

[28] T.T. Georgiou. Relative entropy and the multivariable multidimensional moment problem. *IEEE Transactions on Information Theory*, 52(3):1052–1066, 2006.

[29] X. Jiang, L. Ning, and T.T. Georgiou. Distances and Riemannian metrics for multivariate spectral densities. *IEEE Transactions on Automatic Control*, 57(7):1723–1735, 2012.

[30] P. Masani. *Recent trends in multivariable prediction theory. (Krishnaiah, P.R., Editor), Multivariate Analysis, pp. 351-382*. Academic Press, 1966.

[31] T.T. Georgiou and A. Lindquist. Kullback-Leibler approximation of spectral density functions. *IEEE Transactions on Information Theory*, 49(11):2910–2917, 2003.

[32] A. Ferrante, M. Pavon, and M. Zorzi. A maximum entropy enhancement for a family of high-resolution spectral estimators. *IEEE Transactions on Automatic Control*, 57(2):318–329, 2012.

[33] M. Pavon and A. Ferrante. On the Georgiou-Lindquist approach to constrained Kullback-Leibler approximation of spectral densities. *IEEE Transactions on Automatic Control*, 51(4):639–644, 2006.

[34] S. Kullback and R.A. Leibler. On information and sufficiency. *The Annals of Mathematical Statistics*, 22(1):79–86, 1951.

[35] H. Umegaki. Conditional expectation in an operator algebra, IV (entropy and information). *Kodai Mathematical Journal*, 14(2):59–85, 1962.

[36] R.M. Gray. *Toeplitz and circulant matrices: A review*. Now Pub, 2006.

[37] I. Csiszár and P.C. Shields. *Information theory and statistics: A tutorial*. Now Publishers Inc, 2004.

[38] L. Ning, X. Jiang, and T.T. Georgiou. On the geometry of covariance matrices. *IEEE Signal Processing Letters*, 20(8):787, 2013.

[39] A. Uhlmann. Geometric phases and related structures. *Reports on Mathematical Physics*, 36(2):461–481, 1995.

[40] D. Petz. Geometry of canonical correlation on the state space of a quantum system. *Journal of Mathematical Physics*, 35:780, 1994.

[41] S-I. Amari. Differential-geometrical methods in statistics. *Lecture Notes in Statistics, Berlin, Springer Verlag*, 28, 1985.

[42] R. Bhatia. *Positive definite matrices*. Princeton University Press, 2007.

[43] M. Hübner. Explicit computation of the Bures distance for density matrices. *Physics Letters A*, 163(4):239 – 242, 1992.

[44] T.T. Georgiou, J. Karlsson, and M.S. Takyar. Metrics for power spectra: an axiomatic approach. *IEEE Transactions on Signal Processing*, 57(3):859–867, 2009.

[45] G. Monge. *Mémoire sur la théorie des déblais et des remblais*. De l'Imprimerie Royale, 1781.

[46] L.V. Kantorovich. On the transfer of masses. In *Dokl. Akad. Nauk. SSSR*, volume 37, pages 227–229, 1942.

[47] W. Rudin. *Functional analysis*. McGraw-Hill, Inc., New York, 1991.

[48] C.S. Smith and M. Knott. Note on the optimal transportation of distributions. *Journal of Optimization Theory and Applications*, 52(2):323–329, 1987.

[49] Y. Brenier. Polar factorization and monotone rearrangement of vector-valued functions. *Communications on pure and applied mathematics*, 44(4):375–417, 1991.

[50] R.J. McCann. Exact solutions to the transportation problem on the line. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 455(1984):1341–1380, 1999.

[51] M. Knott and C. S. Smith. On the optimal mapping of distributions. *Journal of Optimization Theory and Applications*, 43(1):39–49, 1984.

[52] L. Ning, T.T. Georgiou, A. Tannenbaum, and S.P. Boyd. Linear models based on noisy data and the Frisch scheme. *arXiv:1304.3877*, 2013.

[53] L. Ning, X. Jiang, and T.T. Georgiou. Geometric methods for estimation of structured covariances. *arXiv:1110.3695*, 2011.

[54] J.D. Benamou and Y. Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numerische Mathematik*, 84(3):375–393, 2000.

[55] R.J. McCann. A convexity principle for interacting gases. *Advances in Mathematics*, 128(1):153–179, 1997.

[56] R. Jordan, D. Kinderlehrer, and F. Otto. The variational formulation of the Fokker-Planck equation. *SIAM Journal on Mathematical Analysis*, 29(1):1–17, 1998.

[57] E. Tannenbaum, T.T. Georgiou, and A. Tannenbaum. Signals and control aspects of optimal mass transport and the Boltzmann entropy. In *49th IEEE Conference on Decision and Control*, pages 1885–1890, 2010.

[58] E.A. Carlen and J. Maas. An analog of the 2-Wasserstein metric in non-commutative probability under which the Fermionic Fokker-Planck equation is gradient flow for the entropy. *arXiv:1203.5377*, 2012.

[59] J.-D. Benamou. Numerical resolution of an unbalanced mass transport problem. *ESAIM: Mathematical Modelling and Numerical Analysis*, 37(05):851–868, 2003.

[60] L. Ning, F.P. Carli, M. Ebtenhaj, E. Foufoula-Georgiou, and T.T. Georgiou. Coping with model uncertainty in data assimilation using optimal mass transport. *submitted*, 2013.

[61] A. Figalli. The optimal partial transport problem. *Archive for rational mechanics and analysis*, 195(2):533–560, 2010.

[62] B. Piccoli and F. Rossi. Generalized Wasserstein distance and its application to transport equations with source. *arXiv:1206.3219*, 2012.

[63] J. Delon, J. Salomon, and A. Sobolevski. Fast transport optimization for Monge costs on the circle. *SIAM Journal on Applied Mathematics*, 70(7):2239–2258, 2010.

[64] A. Connes. *Noncommutative geometry*. Academic press, 1995.

[65] M.A. Rieffel. Metrics on state spaces. *Doc. Math., J. DMV*, 4:559–600, 1999.

[66] F. D'Andrea and P. Martinetti. A view on optimal transport from noncommutative geometry. *SIGMA*, 6(057):24, 2010.

[67] P. Martinetti. Towards a Monge-Kantorovich metric in noncommutative geometry. *Zap. Nauch. Semin. POMI (arXiv:1210.6573)*, page 411, 2013.

[68] P. Biane and D. Voiculescu. A free probability analogue of the Wasserstein metric on the trace-state space. *Geometric and Functional Analysis*, 11(6):1125–1138, 2001.

[69] S. Haker, L. Zhu, A. Tannenbaum, and S. Angenent. Optimal mass transport for registration and warping. *International Journal of Computer Vision*, 60(3):225–240, 2004.

[70] P.J. Basser, J. Mattiello, and D. LeBihan. Estimation of the effective self-diffusion tensor from the NMR spin echo. *Journal of Magnetic Resonance*, 103(3):247–254, 1994.