# Triangle inequalities of quantum Wasserstein distances on noncommutative tori

by

Jinshi Fu

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Thursday June 6, 2024 at 10:00 AM.

| | |
|---|---|
| Student number: | 5720672 |
| Master's Programme: | Applied Mathematics |
| Specialisation: | Mathematics of Quantum Technology and Computation |
| Thesis committee: | Dr. M.P.T. Caspers, TU Delft, Chair, Supervisor |
| | Dr. D. de Laat, TU Delft |
| | Dr. S.E. Zegers, TU Delft |

An electronic version of this thesis is available at http://repository.tudelft.nl/.

# Abstract

In 2022, Golse and Paul defined a pseudometric for quantum optimal transport that extends the classical Wasserstein distance. They proved that the pseudometric satisfies the triangle inequality in certain cases. This thesis reviews their proof in the case where the middle point is a classical density. Motivated by that proof, we formulate the optimal transport problem and propose the quantum Wasserstein distance on the noncommutative 2-torus. This thesis also proves that the proposed quantum Wasserstein distance satisfies the triangle inequality in the case where the middle point is a classical density on the 2-torus.

# Acknowledgements

On the occasion of completing this thesis, I would like to express my deepest gratitude to all those who have helped and supported me during my master's studies.

First, I would like to sincerely thank my supervisor, Martijn Caspers, for his support throughout this thesis project. At the initial stage, the inspiring discussions with him during the weekly meetings are indispensable for the establishment of the objective and methodology of this thesis. In the midterm, his valuable suggestions helped me solve some of the problems I encountered in proving the main result. During the final writing period, with his detailed feedback, I made this thesis more and more complete.

I would also like to thank other staff at the Delft Institute of Applied Mathematics. Special thanks to David de Laat and Sophie Zegers for agreeing to be on the thesis committee, and to Bas Janssens for his teaching and providing helpful information about my course selection.

Finally, I must acknowledge my family. I am grateful to my parents for their support all the time and to my uncles for helping me adapt to life in a foreign country. I would also like to thank my girlfriend for her company.

*Jinshi Fu*
*Delft, May 2024*

# Contents

# Chapter 1

# Introduction

Optimal transport enjoys widespread popularity in various fields such as probability theory, optimization, economics, and machine learning. Central to this problem is the Wasserstein distance, which defines a metric on the set of probability measures. Recent advancements in quantum technologies have motivated the study of quantum versions of optimal transport. Among the literature, Golse and Paul [21] proposed an extended Wasserstein distance that can be used to compare classical probability measures and quantum density operators. It is formally easy and satisfies triangle inequalities in certain cases. Prompted by the development of noncommutative geometry, the notion of noncommutative tori [34] was introduced. We will formulate the optimal transport problem and quantum Wasserstein distances on the noncommutative 2-torus. Moreover, we will investigate the triangle inequality property of the proposed quantum Wasserstein distance.

## 1.1 Optimal transport and Wasserstein distances

The optimal transport problem concerns how to efficiently transport mass from one distribution to another. Intuitively, assuming there is a unit amount of sand piled on a construction site, the sand now needs to be piled in a different way. Moving a certain amount of sand costs certain work, which depends on the starting and ending locations where the sand is taken from and placed. The aim is to use minimal effort to pile the sand in the intended way. This problem is also known as the Monge-Kantorovich problem, which is named after the French mathematician Gaspard Monge and the Soviet mathematician and economist Leonid Kantorovich. In 1781, Gaspard Monge first formalized this problem in [30]. Since then, it has become a classical subject in probability theory, economics, and optimization [38]. In the 20th century, significant progress was made in the study of this problem. In 1942, Leonid Kantorovich [25] proposed a more generalized formulation than Monge's formulation using the notions from measure theory. Unlike Monge's approach, which seeks a deterministic map, Kantorovich's formulation allows for "splitting" mass, thereby considering probabilistic mixtures of destination points. For instance, the Dirac measure can only be transported to another Dirac measure in Monge's formulation, while it can be transported to an arbitrary probability measure in Kantorovich's formulation. With his formulation, the Kantorovich duality [38, Theorem 1.3] was proposed, which is a powerful tool to solve the optimal transport problem. In 1987, another important advancement of the optimal transport problem was made by Yann Brenier [5]. His work connected optimal transport with partial differential equations, fluid mechanics, geometry, probability theory, and functional analysis. Currently, optimal transport is widely applied in various fields such as image retrieval [32], signal processing, and machine learning [27].

To make the optimal transport problem more concrete, suppose the distributions are probability measures $\mu$ and $\nu$ which are defined on an appropriate metric space $(M, d)$, where $d : M \times M \to \mathbb{R}$ is the metric on $M$ (Polish spaces, i.e., separable completely metrizable topological spaces, are appropriate spaces for the optimal transport problem [38, Section 1.1.1]). Let $p \geq 0$ be a nonnegative real number. If the two probability measures have finite $p$-th moments, that is, for some (and thus any) $x_0 \in M$,

$$\int_M d(x, x_0)^p \, d\mu(x) < \infty \quad \text{and} \quad \int_M d(x, x_0)^p \, d\nu(x) < \infty,$$

then it is natural to define the transport cost as a function $c : M \times M \to \mathbb{R}$ which is given by $c : (x, y) \mapsto d(x, y)^p$. A transport plan $\pi$ is a joint probability measure on $M \times M$ whose marginals are $\mu$ and $\nu$ respectively, i.e., for all measurable subsets of $M$,

$$\pi(A \times M) = \mu(A), \quad \pi(M \times A) = \nu(A).$$

Denote by $\mathcal{C}(\mu, \nu)$ the set of transport plans between $\mu$ and $\nu$. Then the Wasserstein distance between $\mu$ and $\nu$ is defined as

$$W_p(\mu, \nu) = \begin{cases} \left( \inf\limits_{\pi \in \mathcal{C}(\mu,\nu)} \displaystyle\int_{M \times M} d(x, y)^p \, d\pi(x, y) \right)^{1/p}, & p \in [1, \infty), \\ \inf\limits_{\pi \in \mathcal{C}(\mu,\nu)} \displaystyle\int_{M \times M} d(x, y)^p \, d\pi(x, y), & p \in [0, 1). \end{cases} \tag{1.1}$$

Denote by $\mathcal{P}_p(M)$ the set of probability measures on $M$ with finite $p$-th moments. The Wasserstein distance $W_p$ defines a metric on $\mathcal{P}_p(M)$ [38, Theorem 7.3]. $W_p$ shall be called the Monge-Kantorovich distance with exponent $p$. In particular, $W_1$ is called the Kantorovich-Rubinstein distance, and $W_2$ is called the quadratic Wasserstein distance. The Wasserstein distance is named after Leonid Vaserstein [37] (Wasserstein is the German spelling of Vaserstein) by Roland Dobrushin in 1970. However, this metric was first introduced by Leonid Kantorovich [26] in 1939. It is also known as the earth mover's distance in computer science.

## 1.2 Quantum optimal transport

Beyond the classical optimal transport problem and Wasserstein distances, their quantum analogues have been proposed and studied in the literature (see, e.g., [6,7,14,15,18,19,21,23,42]). It is a fact that quantum mechanics can be well approximated by classical mechanics [28], and the classical limit of a quantum state can be expressed in terms of the convergence of the Wigner, or the Husimi functions on the phase space associated to the orthogonal projections on the line spanned by the wave functions [21]. In addition, the Wasserstein distances metrize the weak convergence of probability measures on Euclidean spaces [38, Theorem 7.12]. Based on these facts, Zyczkowski and Slomczynski [42] proposed the idea of comparing a quantum state and its classical limit by measuring the Wasserstein distance between its Husimi function and its weak limit, and comparing two quantum states by measuring the Wasserstein distance between their Husimi functions.

However, the evolution of the Husimi function described by the Schrödinger equation is quite complicated. To avoid the difficulties, Golse and Paul [21] proposed to extend the classical Wasserstein distances using the well known formal analogy between quantum density operators on $L^2(\mathbb{R}^d)$ and Borel probability measure on $\mathbb{R}^{2d}$. Then the extended Wasserstein distances can be applied to directly compare a quantum density operator with a probability measures of classical mechanics (though they are different objects) and compare two quantum density operators. Denote by $\mathcal{P}_2^{ac}(\mathbb{R}^{2d})$ the set of absolutely continuous probability measures on $\mathbb{R}^{2d}$ with finite second moments. Denote by $\mathcal{D}(L^2(\mathbb{R}^d))$ the set of density operators on $L^2(\mathbb{R}^d)$, that is,

$$\mathcal{D}(L^2(\mathbb{R}^d)) = \{T \in B(L^2(\mathbb{R}^d)) : T = T^* \geq 0 \text{ and } \operatorname{tr}(T) = 1\}.$$

Let $x$ be the position operator on $L^2(\mathbb{R}^d)$, and let $\nabla_x$ and $\Delta_x = \nabla_x^2$ be the gradient operator and the Laplace operator on $L^2(\mathbb{R}^d)$ respectively. Then the Hamiltonian $\hat{H}$ of the harmonic oscillator is given by

$$\hat{H} = \hat{H}(x, \hbar\nabla_x) = -\frac{1}{2}\hbar^2\Delta_x + \frac{1}{2}|x|^2,$$

where the first term is the kinetic energy operator and the second term is the potential energy operator. Denote by $\mathcal{D}_2(L^2(\mathbb{R}^d))$ the set of density operators on $L^2(\mathbb{R}^d)$ with finite energy for the harmonic oscillator, that is,

$$\mathcal{D}_2(L^2(\mathbb{R}^d)) \coloneqq \{T \in \mathcal{D}(L^2(\mathbb{R}^d)) : \operatorname{tr}(T^{1/2}\hat{H}T^{1/2}) < \infty\}.$$

They intended to extend the Wasserstein distance to the set $\mathfrak{D}_2 \coloneqq \mathcal{D}_2(L^2(\mathbb{R}^d)) \cup \mathcal{P}_2(\mathbb{R}^{2d})$. The couplings between a quantum density operator and an absolutely continuous probability measure, and the couplings between two density operators, are defined analogously to the classical case. Let $x$ and $\xi$ be the $d$-dimensional position variable and the momentum variable in classical mechanics, respectively. The pair $(x, \xi)$ is called the phase space variable. Let $\mu, \nu \in \mathcal{P}_2(\mathbb{R}^{2d})$ and $f dx d\xi \in \mathcal{P}_2^{ac}(\mathbb{R}^{2d})$ be probability measures on $\mathbb{R}^{2d}$, and let $R, S \in \mathcal{D}_2(L^2(\mathbb{R}^d))$. A coupling between $f$ and $R$ is the $B(L^2(\mathbb{R}^d))$-valued measurable functions $Q(x, \xi)$ such that

$$Q(x, \xi) = Q(x, \xi)^* \geq 0 \quad \text{and} \quad \operatorname{tr}(Q(x, \xi)) = f(x, \xi) \quad \text{for a.e. } (x, \xi) \in \mathbb{R}^{2d},$$

while

$$\int_{\mathbb{R}^{2d}} Q(x, \xi)\, dx d\xi = R.$$

Denote by $\mathcal{C}(f, R)$ the set of couplings between $f$ and $R$. The set of couplings between the operators $R$ and $S$ is given by

$$\mathcal{C}(R, S) = \{T \in \mathcal{D}(L^2(\mathbb{R}^d) \otimes L^2(\mathbb{R}^d)) : \operatorname{tr}(T(A \otimes I + I \otimes B)) = \operatorname{tr}(RA + SB)\}.$$

The transport cost functions are also defined analogously to the classical case, but there are extra terms about momentum operators. The transport cost function between a classical probability with the phase space variable $(x, \xi)$ and a quantum density operator with the position operator $y$ and the momentum operator $-i\hbar\nabla_y$ is given by

$$c_\hbar(x, \xi) = c(x, \xi, y, -i\hbar\nabla_y) \coloneqq \sum_{j=1}^{d}((x_j - y_j)^2 + (\xi_j + i\hbar\partial_{y_j})^2). \tag{1.2}$$

The transport cost function between two quantum density operators with the position operators $x, y$ and the momentum operators $-i\hbar\nabla_x, -i\hbar\nabla_y$ is given by

$$C_\hbar = C(x, -i\hbar\nabla_x, y, -i\hbar\nabla_y) \coloneqq \sum_{j=1}^{d}((x_j - y_j)^2 - \hbar^2(\partial_{x_j} - \partial_{y_j})^2). \tag{1.3}$$

Then the extended Wasserstein distance $\mathfrak{d}$ is defined as

$$\begin{cases} \mathfrak{d}(\mu,\nu) := W_2(\mu,\nu), \\ \mathfrak{d}(f,R) := \left( \inf_{Q\in\mathcal{C}(f,R)} \int_{\mathbb{R}^{2d}} \operatorname{tr}(Q(x,\xi)^{1/2} c_\hbar(x,\xi) Q(x,\xi)^{1/2})\, dx d\xi \right)^{1/2}, \\ \mathfrak{d}(R,S) := \left( \inf_{T\in\mathcal{C}(R,S)} \operatorname{tr}(T^{1/2} C_\hbar T^{1/2}) \right)^{1/2}. \end{cases} \tag{1.4}$$

Note that, in the classical-to-quantum case, the distance $\mathfrak{d}$ is only defined for probability measures in $\mathcal{P}_2^{ac}(\mathbb{R}^{2d})$ at present but not for all probability measures in $\mathcal{P}_2(\mathbb{R}^{2d})$. Hence, regarding this $\mathfrak{d}$, they first proved that it admits a unique extension on $\mathfrak{D}_2 \times \mathfrak{D}_2$. Unlike $W_2$ which defines a metric on $\mathcal{P}_2(\mathbb{R}^{2d})$, the extended Wasserstein distance $\mathfrak{d}$ is not a metric on $\mathfrak{D}_2$ since $\mathfrak{d}(R,R) \geq \sqrt{2d\hbar}$ for any $R \in \mathcal{D}_2(L^2(\mathbb{R}^d))$ [18, Theorem 2.3]. Though $\mathfrak{d}$ failed to be a metric, it is still meaningful to check whether it satisfies other properties of a metric. The symmetry and positivity are easy to see. Concerning the triangle inequality, they proved that $\mathfrak{d}$ satisfies the triangle inequality if the middle point is a probability measure, or one of the points is a rank-one density operator [21, Theorem 3.1]. The original statement is formulated in the following theorem.

**Theorem 1.1.** *There is a unique extension of* $\mathfrak{d}$ *defining a map* $\mathfrak{D}_2 \times \mathfrak{D}_2 \to [0,+\infty)$ *still denoted* $\mathfrak{d}$*, satisfying the triangle inequality for each* $\rho_1, \rho_2, \rho_3 \in \mathfrak{D}_2$*:*

$$\mathfrak{d}(\rho_1,\rho_3) \leq \mathfrak{d}(\rho_1,\rho_2) + \mathfrak{d}(\rho_2,\rho_3)$$

*if* $\rho_2 \in \mathcal{P}_2(\mathbb{R}^{2d})$*, or if* $\rho_j \in \{R \in \mathcal{D}_2(L^2(\mathbb{R}^d)) : R^2 = R\}$ *for some* $j \in \{1,2,3\}$*.*

## 1.3 Optimal transport on noncommutative tori

The correspondence between commutative algebras and geometric spaces is a basic idea of algebraic geometry. For example, Gelfand duality states the correspondence between a locally compact Hausdorff space and the C*-algebra of continuous functions on it. Analogously, the concept of noncommutative geometry arises from corresponding the would-be spaces with noncommutative algebras [9].

The interest in studying noncommutative spaces originated from attempts at problems in theoretical physics, in particular quantum mechanics and quantum field theory [12]. For instance, in classical mechanics, the observables are formulated as continuous functions on the phase space, where the functions form a commutative algebra with pointwise multiplication. By contrast, in quantum mechanics, the observables are formulated as the self-adjoint operators on a Hilbert space which do not commute in general. It can be viewed as a noncommutative analogue of functions, and hence an investigation of the underlying geometric space of the algebra of observables is desired.

The field of noncommutative geometry was largely developed by Alain Connes [8, 9] in the 1980s. Many classical tools like measure theory, topology, differential calculus, and Riemannian geometry are extended to the noncommutative case [9]. In particular, the topological structures of the underlying noncommutative spaces are captured by the corresponding C*-algebras. The noncommutative counterpart of the measure theory of classical spaces, which is essential for classical optimal transport, is described by von Neumann algebras. We refer to [29] for a historical description and background of noncommutative geometry.

As a fundamental example of noncommutative spaces, the noncommutative torus is well studied. In [10], Alain Connes introduced basic notions of noncommutative differential geometry and applied them to noncommutative tori. The noncommutative torus is not a single algebra but a family of algebras that are parametrized by antisymmetric matrices $\theta$. There are different ways to see the noncommutative torus. In particular, the noncommutative 2-torus $A_\theta$ with $\theta \in [0,1]$ can be defined as follows:

- **Universal C*-algebras:** The noncommutative 2-torus is the universal C*-algebra generated by two unitaries $u, v$ subject to the relation [1, Section 6 and 7]

$$uv = e^{i\theta} vu.$$

- **Crossed product of C*-algebras:** Let $C(\mathbb{T})$ be the C*-algebra of continuous functions on $\mathbb{T}$, and let $\alpha$ be an action of $\mathbb{Z}$ on $C(\mathbb{T})$ given by

$$\alpha_k(f)(z) = f(e^{ik\theta} z), \quad \text{for } k \in \mathbb{Z} \text{ and } f \in C(\mathbb{T}).$$

  The noncommutative 2-torus is the C*-algebra crossed product $C(\mathbb{T}) \rtimes_\alpha \mathbb{Z}$ [40]. From this point of view, if $\theta$ is irrational, the von Neumann algebra generated by the noncommutative 2-torus is a hyperfinite $\mathrm{II}_1$ factor [39, Section 6.6].

- **Foliations:** For fixed irrational $\theta$, the noncommutative 2-torus can be considered as a foliation C*-algebra for a Kronecker foliation on a torus [11, Section 6].

Among these three constructions, we will introduce the universal C*-algebra construction in detail in Section 2.2 and 2.3. The noncommutative 2-torus is a noncommutative generalization of continuous functions on the 2-torus. In this thesis, we will formulate the optimal transport problem and propose the quantum Wasserstein distances on the noncommutative 2-torus. In particular, we will investigate the triangle inequalities of the proposed quantum Wasserstein distances when the middle point is a classical probability density.

To formulate the optimal transport problem on the noncommutative torus, the notions of densities, couplings between densities, and transport cost are needed. The densities are formulated by the positive and trace-one operators in the noncommutative $L^1$-space associated with the von Neumann algebra generated by $A_\theta$. The couplings between two density operators $A$ and $B$ are defined analogously to the classical setting. They are density operators on the tensor product space with marginals equal to $A$ and $B$ respectively. The quantum Wasserstein distance is defined by formal analogy with the classical Wasserstein distance on the probability measures on the 2-torus. Then, as in Theorem 1.1 [21, Theorem 3.1], we consider the triangle inequality problems of the proposed quantum Wasserstein distances. Since the von Neumann algebra generated by $A_\theta$ is a $\mathrm{II}_1$ factor for irrational $\theta$, it does not contain minimal projections and neither contains rank-one operators. We only considered the case when the middle point is a classical probability density. Following the ideas of [21], we first proved an inequality for the cost functions and some assertions about the spectral measures. Finally, we applied the results above and proved that the triangle inequality of the proposed quantum Wasserstein holds when the middle point is a classical probability density.

## 1.4 Structure of this thesis

This thesis is organized as follows: In Chapter 2, we first review the notion of noncommutative $L^p$-spaces which is necessary for the definition of density operators on the noncommutative torus. Based on universal C*-algebras, the noncommutative 2-torus is defined. The spectral theorem, which will be extensively used in the proof, is reviewed. Then this chapter ends with comparisons between classical and quantum optimal transport problems. In Chapter 3, we give a formal statement of the optimal transport problem on the noncommutative 2-torus by defining the couplings between two density operators and the quantum Wasserstein distances. Then the triangle inequality problem of the proposed quantum Wasserstein distances, which is the main concern of this thesis, is introduced. In Chapter 4, we first review the original proof of Theorem 1.1 [21, Theorem 3.1] and discuss its inspiration for the proof of the triangle inequality problem concerned in this thesis. Then detailed proof of the problem is given. In Chapter 5, we will formulate the main results and discuss the open problems.

# Chapter 2

# Preliminaries

In this chapter, we will introduce the basic terminologies that are needed for the formulation of optimal transport on the noncommutative torus, and present some theorems that will be applied in the proof. To be specific, the notion of noncommutative $L^p$-spaces associated with a semifinite von Neumann algebra (see, e.g., [16,33]) is introduced; Based on the notion of universal C*-algebras, the definition of the noncommutative 2-torus is given (see, e.g., [1]); The spectral theorem for normal operators on a Hilbert space (see, e.g., [13]) is introduced. Finally, this chapter ends with comparisons between key notions (densities and transport plans) of optimal transport problems in different settings.

## 2.1 Noncommutative $L^p$-spaces

To discuss density operators associated with a von Neumann algebra, we need the terminology of noncommutative $L^p$-spaces, specifically, the noncommutative $L^1$-space. This section will follow the formulation in [33,41]. At first, we give the definition of the trace on a von Neumann algebra, which is a noncommutative analogue of the classical integral.

**Definition 2.1** (Trace on a von Neumann algebra)**.** Let $\mathcal{M}$ be a von Neumann algebra, and let $\mathcal{M}_+$ denote the set of positive elements of $\mathcal{M}$. A map $\tau : \mathcal{M}_+ \to [0, \infty]$ is a *trace* on $\mathcal{M}$ if

(i) $\tau(x + y) = \tau(x) + \tau(y)$, for all $x, y \in \mathcal{M}_+$;

(ii) $\tau(\lambda x) = \lambda\tau(x)$, for all $\lambda \geq 0$, $x \in \mathcal{M}_+$;

(iii) $\tau(u^*u) = \tau(uu^*)$, for all $u \in \mathcal{M}$.

$\tau$ is said to be *normal* if $\sup_\alpha(x_\alpha) = \tau(\sup_\alpha x_\alpha)$ for any bounded increasing net $(x_\alpha)$ in $\mathcal{M}_+$, *semifinite* if for any non-zero $x \in \mathcal{M}_+$ there is a non-zero $y \in \mathcal{M}_+$ such that $y \leq x$ and $\tau(y) < \infty$, and *faithful* if $\tau(x) = 0$ implies $x = 0$. If $\tau(1) < \infty$, $\tau$ is said to be finite.

*Remark.* A von Neumann algebra is called semifinite if it admits a normal semifinite faithful (*n.s.f.*) trace $\tau$, which is assumed in this section. Let $\mathcal{S}_+(\mathcal{M})$ be the set that all $x \in \mathcal{M}_+$ such that $\tau(\operatorname{supp} x) < \infty$, where $\operatorname{supp} x$ is the support projection of $x$ which is the smallest projection $p \in \mathcal{M}$ such that $xp = x$, or equivalently for positive elements, $px = x$. The linear span of $\mathcal{S}_+(\mathcal{M})$ is denoted by $\mathcal{S}(\mathcal{M})$. Consider the set of families of pairwise orthogonal finite projections with inclusion order. There exist a maximal element $\{p_i\}_{i\in I}$ by Zorn's lemma. It can be verified that $\sum_{i\in I} p_i = 1$ under the assumption that $\tau$ is semifinite. For any finite subset

$J \subset I$, let $e_J = \sum_{i \in J} p_i$. Then $e_J$ is finite and strong operator topology (SOT) converges to 1 as $J$ tends to $I$. For any element $x \in \mathcal{M}_+$, $e_J x e_J \in \mathcal{S}_+(\mathcal{M})$ and SOT converges to $x$. Thus, $\mathcal{S}(\mathcal{M})$ is SOT dense in $\mathcal{M}$. It can also be verified that $\mathcal{S}(\mathcal{M})$ is a self-adjoint ideal in $\mathcal{M}$ [41, Theorem 5.1.3.(3)]. If there is no ambiguity referring the von Neumann algebra, we abbreviate $\mathcal{S}_+(\mathcal{M})$ and $\mathcal{S}(\mathcal{M})$ as $\mathcal{S}_+$ and $\mathcal{S}$ respectively.

**Definition 2.2** (Noncommutative $L^p$-spaces)**.** Let $\mathcal{M}, \tau$ and $\mathcal{S}$ be as stated above. Define

$$\|x\|_p := [\tau(|x|^p)]^{1/p}, \quad x \in \mathcal{S} \text{ and } 0 < p < \infty.$$

Then $\|\cdot\|_p$ is a norm on $\mathcal{S}$ if $1 \leq p < \infty$ and a quasi-norm if $0 < p < 1$. The noncommutative $L^p$-space associated with $(\mathcal{M}, \tau)$ is the completion of $\mathcal{S}$ with respect to $\|\cdot\|_p$

$$L^p(\mathcal{M}, \tau) = \overline{\mathcal{S}}^{\|\cdot\|_p}.$$

Set $L^\infty(\mathcal{M}, \tau) = \mathcal{M}$ equipped with the operator norm. The pair $(\mathcal{M}, \tau)$ is sometimes called *noncommutative measure space.*

Instead of defining the noncommutative $L^p$-spaces as the completion of $\mathcal{S}$, there is another equivalent definition describing its elements as the closed densely defined operators on $H$ ($H$ is the underlying Hilbert space on which $\mathcal{M}$ acts). To begin with, the definitions of affiliated operators and $\tau$-measurable operators are introduced.

**Definition 2.3** (Affiliated operators)**.** Suppose $\mathcal{M}$ is a von Neumann algebra that acts on the Hilbert space $H$. A closed densely defined operator $x$ on $H$ is said to be *affiliated* with $\mathcal{M}$, denoted as $x\eta\mathcal{M}$, if $xu = ux$ for any unitary $u$ in the commutant $\mathcal{M}'$ of $\mathcal{M}$.

**Definition 2.4** ($\tau$-measurable operators)**.** An affiliated operator $x$ is said to be *$\tau$-measurable* if $\tau(E_{|x|}(\lambda, \infty)) < \infty$ for some $\lambda > 0$, where $E_{|x|}$ denotes the spectral measure (see Section 2.4) for $|x|$ and

$$E_{|x|}(\lambda, \infty) = \int \chi_{(\lambda,\infty)} \, dE_{|x|},$$

where $\chi$ is the indicator function. The space of all $\tau$-measurable operators associated with $\mathcal{M}$ is denoted by $L^0(\mathcal{M}, \tau)$.

*Remark.* For $\tau$-measurable operators $x$ and $y$, the closures of $x + y$ and $xy$ are $\tau$-measurable [16, Theorem 2.3.8]. For convenience, we will still denote their closures by $x + y$ and $xy$ respectively.

*Remark.* Clearly, $\mathcal{M} \subset L^0(\mathcal{M})$. Besides, if $\tau$ is finite, then any affiliated operator is $\tau$-measurable.

Note that $\tau$ is only defined for operators in $\mathcal{M}_+$ at this moment, it will then be defined on the positive part of $\tau$-measurable operators using the notion of generalized singular numbers.

**Definition 2.5** (Generalized singular numbers)**.** For any measurable operator $x$, the *generalized singular numbers* are defined as

$$\mu_t(x) := \inf\{\lambda > 0 : \tau(E_{|x|}(\lambda, \infty)) \leq t\}, \quad t > 0.$$

*Remark.* Let

$$V(\varepsilon, \delta) = \{x \in L^0(\mathcal{M}, \tau) : \mu_\varepsilon(x) \leq \delta\}.$$

Then $\{V(\varepsilon, \delta) : \varepsilon, \delta > 0\}$ forms a system of neighbourhoods at 0 for which $L^0(\mathcal{M}, \tau)$ becomes a metrizable topological $*$-algebra. The convergence with respect to this topology is called the *convergence in measure*. Then $\mathcal{M}$ is dense in $L^0(\mathcal{M}, \tau)$ and $\tau$ can be extended to a positive tracial functional on the positive part $L^0_+(\mathcal{M}, \tau)$ of $L^0(\mathcal{M}, \tau)$, still denoted by $\tau$, satisfying

$$\tau(x) = \int_0^\infty \mu_t(x) \, dt, \quad x \in L^0_+(\mathcal{M}, \tau).$$

**Definition 2.6** (Noncommutative $L^p$-spaces)**.** For $0 < p < \infty$, the noncommutative $L^p$-space associated with $(\mathcal{M}, \tau)$ is

$$L^p(\mathcal{M}, \tau) = \{x \in L^0(\mathcal{M}, \tau) : \tau(|x|^p) < \infty\} \quad \text{and} \quad \|x\|_p = (\tau(|x|^p))^{1/p}.$$

*Remark.* Definition 2.2 and Definition 2.6 are equivalent.

Now, with the noncommutative $L^1$-space defined, we are ready to define the density operators associated with a semifinite von Neumann algebra.

**Definition 2.7** (Densities associated with a von Neumann algebra)**.** The density operators associated with $(\mathcal{M}, \tau)$ are positive elements $\rho \in L^1(\mathcal{M})$ such that $\tau(\rho) = 1$. The set of density operators associated with $\mathcal{M}$ is denoted by $\mathcal{D}(\mathcal{M})$.

To discuss the joint density operators associated with two von Neumann algebras, we need the notions of tensor product of von Neumann algebras and tensor product of traces [41, Theorem 5.5.1].

**Definition 2.8** (Tensor product of von Neumann algebras)**.** Let $\mathcal{M}$ and $\mathcal{N}$ be two von Neumann algebras on Hilbert spaces $H_1$ and $H_2$ respectively. The tensor product von Neumann algebra $\mathcal{M}\bar{\otimes}\mathcal{N}$ is the SOT closure in $B(H_1 \otimes H_2)$ of the algebraic tensor product of $\mathcal{M}$ and $\mathcal{N}$, that is,

$$\mathcal{M}\bar{\otimes}\mathcal{N} = \overline{\mathcal{M} \underset{\text{alg}}{\otimes} \mathcal{N}}^{\text{SOT}}.$$

**Theorem 2.1** (Tensor product of traces)**.** *Let $\mathcal{M}_1, \mathcal{M}_2$ be two von Neumann algebras equipped with n.s.f. trace $\tau_1, \tau_2$ respectively. Then there exists a unique n.s.f. trace $\tau$ on $\mathcal{M} = \mathcal{M}_1\bar{\otimes}\mathcal{M}_2$ such that*

$$\tau(x_1 \otimes x_2) = \tau_1(x_1)\tau_2(x_2), \quad x_1 \in \mathcal{S}(\mathcal{M}_1), x_2 \in \mathcal{S}(\mathcal{M}_2).$$

*$\tau$ is called the tensor product of $\tau_1$ and $\tau_2$, denoted by $\tau_1 \otimes \tau_2$.*

## 2.2 Universal C*-algebras

To define the noncommutative 2-torus, we need the notion of universal C*-algebras. It is a C*-algebra described in terms of generators and relations. The construction of the universal C*-algebra is similar to it of the universal algebra which is just the free algebra quotient by some relations. However, due to the norm structure of C*-algebras, the universal C*-algebra may not exist since it is possible that there is no appropriate C*-seminorm for the universal $*$-algebra to become a C*-algebra. This section will follow [1, Section 6] to formulate the construction of universal C*-algebras, discuss the existence of universal C*-algebras, and introduce the universal property of universal C*-algebras.

**Definition 2.9** (Free algebra)**.** Let elements $E = \{x_i \mid i \in I\}$ be given, where $I$ is some index set.

(i) A *noncommutative monomial* in $E$ is a word $x_{i_1} \cdots x_{i_m}$ with $i_1, ..., i_m \in I$ and $m \in \mathbb{N} \setminus \{0\}$.

(ii) A *noncommutative polynomial* in $E$ is a formal complex linear combination of noncommutative monomials: $\sum_{k=1}^{N} \alpha_k y_k$ with $N \in \mathbb{N}, \alpha_k \in \mathbb{C}$ and $y_1, ..., y_n$ being noncommutative monomials in $E$.

(iii) On noncommutative monomials, we consider the *concatenation of words*,

$$(x_{i_1} \cdots x_{i_m}) \cdot (x_{j_1} \cdots x_{j_n}) \coloneqq x_{i_1} \cdots x_{i_m} x_{j_1} \cdots x_{j_n}$$

where $x_{i_1} \cdots x_{i_m}$ and $x_{j_1} \cdots x_{j_n}$ are two monomials.

(iv) The *free (complex) algebra* on the *generator* set $E$ is given as the set of noncommutative polynomials in $E$ together with the canonical addition and scalar multiplication, and the multiplication of elements given by the concatenation. The elements in $E$ are understood as being distinct.

Given $E = \{x_i \mid i \in I\}$, we add another set (disjoint with $E$) of generators $E^* \coloneqq \{x_i^* \mid i \in I\}$, and we define an involution on the free algebra on $E \cup E^*$ by extending

$$(\alpha x_{i_1}^{\varepsilon_1} \cdots x_{i_m}^{\varepsilon_m}) \coloneqq \bar{\alpha} x_{i_m}^{\bar{\varepsilon}_m} \cdots x_{i_1}^{\bar{\varepsilon}_1}$$

to linear combinations, where $\alpha \in \mathbb{C}, \varepsilon_k \in \{1, *\}$ and

$$\bar{\varepsilon}_k \coloneqq \begin{cases} 1, & \text{if } \varepsilon_k = *, \\ *, & \text{if } \varepsilon_k = 1. \end{cases}$$

In this way, we obtain the *free* $*$*-algebra* $P(E)$ on the generator set $E$. Note that any polynomial $p \in P(E)$ can be viewed as an algebraic relation when considering the equation $p = 0$.

**Definition 2.10** (Universal $*$-algebra)**.** Let $E = \{x_i \mid i \in I\}$ be a set of elements with $I$ some index set, and let $R \subset P(E)$ be a set of polynomials. Let $J(R) \subset P(E)$ be the two-sided self-adjoint ideal generated by $R$. The *universal* $*$*-algebra* with *generators* $E$ and *relations* $R$ is defined as the quotient $A(E \mid R) \coloneqq P(E)/J(R)$.

*Remark.* The equivalence class of $x_i \in E$ in $A(E \mid R)$ will be denoted as $x_i$ by abuse of notation.

**Definition 2.11** (C*-seminorms)**.** Let $A$ be a $*$-algebra. A C*-seminorm on $A$ is a map $p : A \to [0, \infty)$ such that

(i) $p(\lambda x) = |\lambda| p(x)$ and $p(x + y) \leq p(x) + p(y)$ for all $x, y \in A$ and $\lambda \in \mathbb{C}$;

(ii) $p(xy) \leq p(x)p(y)$ for all $x, y \in A$;

(iii) $p(x^*x) = p(x)^2$ for all $x \in A$.

**Definition 2.12** (Universal C*-algebra)**.** Let $E$ be a set of generators and $R \subset P(E)$ be relations. Put

$$\|x\| \coloneqq \sup\{p(x) \mid p \text{ is a C*-seminorm on } A(E \mid R)\}.$$

If $\|x\| < \infty$ for all $x \in A(E \mid R)$, we can see that $\|\cdot\|$ is a C*-seminorm and $\{x \in A(E \mid R) \mid \|x\| = 0\}$ is a two-sided self-adjoint ideal. In that case, we define the *universal C*-algebra* $C^*(E \mid R)$ as the completion with respect to $\|\cdot\|$:

$$C^*(E \mid R) \coloneqq \overline{A(E \mid R)/\{x \in A(E \mid R) \mid \|x\| = 0\}}^{\|\cdot\|}.$$

**Lemma 2.2** (Existence of universal C*-algebras)**.** *Let* $E = \{x_i \mid i \in I\}$ *be generators and* $R \subset P(E)$ *be relations.*

(i) *If* $\|x\| < \infty$ *for all* $x \in A(E \mid R)$*, then* $C^*(E \mid R)$ *is a C*-algebra and we say that the universal C*-algebra of* $E$ *and* $R$ *exists.*

(ii) *If there is a constant $C > 0$ such that $p(x_i) < C$ for all $i \in I$ and all C*-seminorms $p$ on $A(E \mid R)$, then $\|x\| < \infty$ for all $x \in A(E \mid R)$, that is, $C^*(E \mid R)$ exists.*

*Remark.* The existence of universal C*-algebra is assured by the above lemma. However, it is still possible that the constructed universal C*-algebra is trivial, namely, $C^*(E \mid R) = 0$. If there are non-trivial $*$-homomorphisms from $C^*(E \mid R)$ to another non-trivial C*-algebra, we can conclude that $C^*(E \mid R)$ is non-trivial. Concerning this issue, we introduce the universal property of universal C*-algebras in the next proposition.

**Proposition 2.3** (Universal property). *Let $E = \{x_i \mid i \in I\}$ be generators and $R \subset P(E)$ be relations such that the universal C*-algebra $C^*(E \mid R)$ exists. Let $B$ be a C*-algebra containing a subset $E' = \{y_i \mid i \in I\}$. If the elements in $E'$ satisfy the relations $R$, then there is a unique $*$-homomorphism $\varphi : C^*(E \mid R) \to B$ sending $x_i$ to $y_i$, for all $i \in I$.*

*Remark.* Note that every $*$-homomorphism $\varphi : C^*(E \mid R) \to B$ between C*-algebras gives rise to a C*-seminorm $p_\varphi(\cdot) = \|\varphi(\cdot)\|_B$ on $C^*(A \mid R)$. Also note that every $*$-homomorphism between C*-algebras is contractive. It is now clear that why the norm on the universal C*-algebra is defined as the supremum of all its C*-seminorms.

## 2.3 Noncommutative 2-torus

The Gelfand duality (see, e.g., Section IV.4 in [24]) asserts the equivalence of categories between the 2-torus $\mathbb{T}^2$ and the C*-algebra of continuous functions on it $C(\mathbb{T}^2)$. So we can investigate $\mathbb{T}^2$ by looking at $C(\mathbb{T}^2)$ indirectly. By Stone-Weierstrass theorem, we know that $C(\mathbb{T}^2)$ can be uniformly approximated by polynomial functions $\mathrm{Pol}(\mathbb{T}^2)$. Utilizing the concept of universal C*-algebras, the C*-algebra $C(\mathbb{T}^2)$ can be viewed as the universal C*-algebra generated by two commuting unitaries. This can be justified as follows [1, Lemma 7.4]:

**Proposition 2.4.** *Let $E = \{u, v\}$ be generators, and let $R = \{u, v \text{ are unitaries}, uv = vu\}$ be relations. Then $C^*(E \mid R)$ exists and is isomorphic to $C(\mathbb{T}^2)$.*

*Proof.* Suppose $p$ is a C*-seminorm on $A(E \mid R)$, then $p(1) = p(1^*1) = p(1)^2$ implies that $p(1) \in \{0, 1\}$. Also note that $p(u)^2 = p(u^*u) = p(1)$ and similarly $p(v)^2 = p(1)$. So $p(u), p(v) \le 1$ and then $C^*(E \mid R)$ exists by Lemma 2.2.

For convenience, denote $C^*(E \mid R)$ by $A_0$. Also note that $A(E \mid R)$ is just the usual polynomials with two indeterminates.

By the Gelfand representation: $A_0 \cong C_0(\Omega(A_0))$ and we claim that $\Omega(A_0)$ is homeomorphic to $\mathbb{T}^2$. Indeed, for any character $\pi \in \Omega(A_0)$, it is determined by its values at $u$ and $v$ due to the construction of universal C*-algebra. This gives a correspondence $f : \pi \mapsto (\pi(u), \pi(v)) \in S^1 \times S^1 = \mathbb{T}^2$ since

$$|\pi(u)|^2 = \pi(u)^*\pi(u) = \pi(u^*u) = \pi(1) = 1 = \pi(1) = \pi(v^*v) = \pi(v)^*\pi(v) = |\pi(v)|^2.$$

Suppose $(\pi_n)_{n\ge1}$ is a sequence in $\Omega(A_0)$ such that it converges to $\pi$ in weak* topology, then

$$\lim_{n\to\infty} \|f(\pi) - f(\pi_n)\| = \lim_{n\to\infty} \sqrt{|\pi(u) - \pi_n(u)|^2 + |\pi(v) - \pi_n(v)|^2} = 0.$$

This implies that $f$ is continuous. Conversely, for $(z_1, z_2) \in \mathbb{T}^2$, $z_1, z_2$ are unitaries and satisfy $z_1 z_2 = z_2 z_1$. By the universal property of $A_0$ (Proposition 2.3), there is a unique $*$-homomorphism $\phi : A_0 \to \mathbb{C}$ maps $u, v$ to $z_1, z_2$. So it is a character in $\Omega(A_0)$. Denote this correspondence by $g : \mathbb{T}^2 \to \Omega(A_0)$ and note that it is the inverse of $f$. Suppose $(z_{n1}, z_{n2})_{n\ge1}$ is a sequence in $\mathbb{T}^2$ such that it converges to $(z_1, z_2)$. For any $a \in A_0$ and $\varepsilon > 0$, there is $b \in A(E \mid R)$ such that $\|a - b\| < \varepsilon$. Then

$$\begin{aligned}
|g(z_1, z_2)(a) - g(z_{n1}, z_{n2})(a)| =& |g(z_1, z_2)(a) - g(z_1, z_2)(b) + g(z_1, z_2)(b) - g(z_{n1}, z_{n2})(b) \\
&+ g(z_{n1}, z_{n2})(b) - g(z_{n1}, z_{n2})(a)| \\
\leq& |g(z_1, z_2)(a) - g(z_1, z_2)(b)| + |g(z_1, z_2)(b) - g(z_{n1}, z_{n2})(b)| \\
&+ |g(z_{n1}, z_{n2})(b) - g(z_{n1}, z_{n2})(a)| \\
\leq& \|g(z_1, z_2)\| \cdot \|a - b\| + |g(z_1, z_2)(b) - g(z_{n1}, z_{n2})(b)| + \\
& \|g(z_{n1}, z_{n2})\| \cdot \|a - b\| \\
<& \varepsilon + |g(z_1, z_2)(b) - g(z_{n1}, z_{n2})(b)| + \varepsilon.
\end{aligned}$$

Since polynomials are continuous, $|g(z_1, z_2)(b) - g(z_{n1}, z_{n2})(b)| \to 0$ as $n \to \infty$. In addition, $\varepsilon$ is arbitrary, we obtain

$$\lim_{n\to\infty} |g(z_1, z_2)(a) - g(z_{n1}, z_{n2})(a)| = 0.$$

This implies that $(g(z_{n1}, z_{n2}))_{n\geq 1}$ converges to $g(z_1, z_2)$ in weak* topology and thus $g$ is continuous. This concludes the claim and we obtain $A_0 \cong C_0(\Omega(A_0)) \cong C_0(\mathbb{T}^2) = C(\mathbb{T}^2)$. □

Further, we can consider the case where the two unitary generators do not commute but satisfy certain commuting relation. Next we introduce the definition of *noncommutative 2-torus* which is also known as *rotation algebra*.

**Definition 2.13** (Noncommutative 2-torus)**.** The noncommutative 2-torus $A_\theta$ is defined as the universal C*-algebra generated by unitaries $u, v$ subject to the relation $uv = e^{i\theta} vu$ with $\theta \in \mathbb{R}$,

$$A_\theta = C^*(u, v \mid u, v \text{ are unitaries, } uv = e^{i\theta} vu).$$

*Remark.* The existence of $A_\theta$ can be derived as in Proposition 2.4. For the non-triviality, consider the following example. Let $H$ be a separable Hilbert space with orthonormal basis $(e_n)_{n\in\mathbb{Z}}$. The bilateral shift operator $\tilde{S} \in B(H)$ is given by $\tilde{S}e_n = e_{n+1}, n \in \mathbb{Z}$. For $e^{i\theta} \in S^1 = \mathbb{T}$, the diagonal operator $d(e^{i\theta}) \in B(H)$ is given by $d(e^{i\theta})e_n = e^{in\theta}e_n$. It can be verified that $\tilde{S}$ and $d(e^{i\theta})$ are unitaries and satisfy $d(e^{i\theta})\tilde{S} = e^{i\theta}\tilde{S}d(e^{i\theta})$ (Lemma 7.2 in [1]). By the universal property of $A_\theta$, there is a $*$-homomorphism from $A_\theta$ to $B(H)$ sending $u$ and $v$ to $d(e^{i\theta})$ and $\tilde{S}$ respectively, and thus $A_\theta$ is non-trivial.

*Remark.* Analogous to commutative case, $A_\theta$ can also be denoted by $C(\mathbb{T}^2_\theta)$. For convenience, denote by $\mathrm{Pol}(\mathbb{T}^2_\theta)$ the set of "polynomials" on the noncommutative 2-torus given by

$$\mathrm{Pol}(\mathbb{T}^2_\theta) := \left\{ x \in A_\theta \,\middle|\, x = \sum_{k,\ell\in\mathbb{Z}} \alpha_{k\ell} u^k v^\ell, \text{ for finitely many nonzero } \alpha_{k\ell} \in \mathbb{C} \right\}.$$

By the construction of universal C*-algebras, we can see that $\mathrm{Pol}(\mathbb{T}^2_\theta)$ is dense in $A_\theta$.

In the next proposition, we consider a faithful tracial state on $A_\theta$ [1, Proposition 7.10].

**Proposition 2.5.** *Let $\sum_{k,\ell\in\mathbb{Z}} \alpha_{k\ell} u^k v^\ell$ be an arbitrary element in $\mathrm{Pol}(\mathbb{T}^2_\theta)$, and let $\tau$ be the linear map given by*

$$\tau\Big(\sum_{k,\ell\in\mathbb{Z}} \alpha_{k\ell} u^k v^\ell\Big) = \alpha_{00}.$$

*Then $\tau$ can be extended to $A_\theta$ as a faithful tracial state, that is, $\tau$ satisfies*

(i) $\tau(x^*x) \geq 0$, *for all* $x \in A_\theta$ *and the equality holds if and only if* $x = 0$;

(ii) $\tau(xy) = \tau(yx)$, *for all* $x, y \in A_\theta$;

(iii) $\tau(\mathbf{1}) = 1$.

For the proof of this proposition, we will follow Lemma 7.6, 7.7 and Proposition 7.10 in [1] to construct a faithful tracial state on $A_\theta$ that extends $\tau$.

*Proof.* For $\zeta, \mu \in \mathbb{T}$, consider the liner map $\rho_{\zeta,\mu} : A_\theta \to A_\theta$ given by

$$\rho_{\zeta,\mu}(u) = \zeta u, \quad \rho_{\zeta,\mu}(v) = \mu v.$$

This map exists by the universal property (Proposition 2.3) since $\zeta u$ and $\mu v$ satisfy the commuting relation $(\zeta u)(\mu v) = \zeta\mu uv = \zeta\mu e^{i\theta} vu = e^{i\theta}(\mu v)(\zeta u)$. Note that $\rho_{\bar{\zeta},\bar{\mu}}$ is the inverse of $\rho_{\zeta,\mu}$, so $\rho_{\zeta,\mu}$ is a $*$-isomorphism.

For fixed $x = \sum_{k,\ell\in\mathbb{Z}} \alpha_{k\ell} u^k v^\ell \in \mathrm{Pol}(\mathbb{T}^2_\theta)$, the map $\rho_{1,e^{2\pi it}}(x) : [0,1] \to A_\theta$ is norm continuous. Indeed, for $t, t' \in [0,1]$, observe that

$$\begin{aligned}
\|\rho_{1,e^{2\pi it}}(x) - \rho_{1,e^{2\pi it'}}(x)\| &= \left\| \sum_{k,\ell\in\mathbb{Z}} \alpha_{k\ell} u^k (e^{2\pi it} v)^\ell - \sum_{k,\ell\in\mathbb{Z}} \alpha_{k\ell} u^k (e^{2\pi it'} v)^\ell \right\| \\
&= \left\| \sum_{k,\ell\in\mathbb{Z}} \alpha_{k\ell} u^k (e^{2\pi it\ell} - e^{2\pi it'\ell}) v^\ell \right\| \\
&\leq \sum_{k,\ell\in\mathbb{Z}} |\alpha_{k,\ell}| |e^{2\pi it\ell} - e^{2\pi it'\ell}| \|u^k\| \|v^\ell\| \\
&= \sum_{k,\ell\in\mathbb{Z}} |\alpha_{k,\ell}| |e^{2\pi i\ell(t-t')} - 1|,
\end{aligned}$$

and it tends to 0 as $|t - t'|$ tends to 0. Since $\mathrm{Pol}(\mathbb{T}^2_\theta)$ is dense in $A_\theta$, for any $x \in A_\theta$ and $\varepsilon > 0$, there exists $y \in \mathrm{Pol}(\mathbb{T}^2_\theta)$ such that $\|x - y\| < \varepsilon$. Note that

$$\begin{aligned}
&\|\rho_{1,e^{2\pi it}}(x) - \rho_{1,e^{2\pi it'}}(x)\| \\
=&\|\rho_{1,e^{2\pi it}}(x) - \rho_{1,e^{2\pi it}}(y) + \rho_{1,e^{2\pi it}}(y) - \rho_{1,e^{2\pi it'}}(y) + \rho_{1,e^{2\pi it'}}(y) - \rho_{1,e^{2\pi it'}}(x)\| \\
\leq&\|\rho_{1,e^{2\pi it}}(x) - \rho_{1,e^{2\pi it}}(y)\| + \|\rho_{1,e^{2\pi it}}(y) - \rho_{1,e^{2\pi it'}}(y)\| + \|\rho_{1,e^{2\pi it'}}(y) - \rho_{1,e^{2\pi it'}}(x)\| \\
=&\|\rho_{1,e^{2\pi it}}(x - y)\| + \|\rho_{1,e^{2\pi it}}(y) - \rho_{1,e^{2\pi it'}}(y)\| + \|\rho_{1,e^{2\pi it'}}(y - x)\| \\
=&\|x - y\| + \|\rho_{1,e^{2\pi it}}(y) - \rho_{1,e^{2\pi it'}}(y)\| + \|y - x\| \\
<&\|\rho_{1,e^{2\pi it}}(y) - \rho_{1,e^{2\pi it'}}(y)\| + 2\varepsilon,
\end{aligned}$$

where the last equality holds by the fact that any $*$-isomorphism between C*-algebras is isometric. This implies that $\rho_{1,e^{2\pi it}}(x)$ is a norm continuous map for any fixed $x \in A_\theta$.

Further, for fixed $x \in A_\theta$, the map $\rho_{1,e^{2\pi it}}(x)$ is Bochner integrable [17, Chapter 64] in the following sense. For norm continuous function $f : [0,1] \to A_\theta$ and positive integer $n$, there exists a partition $\{\Delta^n_k\}_{k=1}^{m_n}$ of $[0,1]$ such that each $\Delta^n_k$ is measurable and

$$\|f(t) - f(t')\| < \frac{1}{n}, \quad t, t' \in \Delta^n_k.$$

For arbitrary $t_k^n \in \Delta_k^n$, consider the following simple function

$$s_n(t) \coloneqq \sum_{k=1}^{m_n} f(t_k^n)\chi_{\Delta_k^n}(t).$$

Then $\|f(t) - s_n(t)\| < 1/n$ for all $t \in [0,1]$ and

$$\int_0^1 \|f(t) - s_n(t)\|\, dt < \int_0^1 \frac{1}{n}\, dt = \frac{1}{n},$$

which implies $f$ is Bochner integrable [17, Definition 64.9]. Denote by $|\Delta_k^n|$ the measure of $\Delta_k^n$. It can be proven that $\lim_{n\to\infty} \sum_{k=1}^{m_n} f(t_k^n)|\Delta_k^n|$ exists and independent of the choice of the sequence [17, Lemma 64.10]. Then the Bochner integral of $f$ is defined as

$$\int_0^1 f(t)\, dt = \lim_{n\to\infty} \sum_{k=1}^{m_n} f(t_k^n)|\Delta_k^n|.$$

Moreover, we consider $\varphi_1 : A_\theta \to A_\theta$ given by

$$\varphi_1(x) \coloneqq \int_0^1 \rho_{1,e^{2\pi it}}(x)\, dt.$$

By the definition of Bochner integral and the fact that $\rho_{1,e^{2\pi it}}$ is a $*$-isomorphism, we have

$$\|\varphi_1(x)\| \le \int_0^1 \|\rho_{1,e^{2\pi it}}(x)\|\, dt = \int_0^1 \|x\|\, dt = \|x\|,$$

and $\varphi_1$ is linear, unital and positive. For the faithfulness of $\varphi_1$ (a positive linear map is faithful if the images of nonzero positive elements are nonzero), let $x$ be a nonzero positive element in $A_\theta$, then $\rho_{1,e^{2\pi it}}(x)$ is nonzero and positive for any $t \in [0,1]$. For some fixed $t_0 \in [0,1]$, there exists a state $\psi$ on $A_\theta$ such that $\psi(\rho_{1,e^{2\pi it_0}}(x)) = \|\rho_{1,e^{2\pi it_0}}(x)\| \neq 0$ [31, Theorem 3.3.6]. By linearity and continuity of $\psi$ [17, Corollary 64.14], we have

$$\psi(\varphi_1(x)) = \psi\left(\int_0^1 \rho_{1,e^{2\pi it}}(x)\, dt\right) = \int_0^1 \psi(\rho_{1,e^{2\pi it}}(x))\, dt > 0.$$

This implies $\varphi_1$ is faithful. Let $k, \ell \in \mathbb{Z}$, observe that

$$\varphi_1(u^k v^\ell) = \int_0^1 \rho_{1,e^{2\pi it}}(u^k v^\ell)\, dt = \int_0^1 u^k (e^{2\pi it} v)^\ell\, dt = u^k v^\ell \int_0^1 e^{2\pi it\ell}\, dt = u^k \delta_{\ell 0},$$

where the second to last equality holds by Corollary 64.14 in [17].

Similarly, the map $\varphi_2 : A_\theta \to A_\theta$ given by

$$\varphi_2(x) \coloneqq \int_0^1 \rho_{e^{2\pi it},1}(x)\, dt,$$

is linear, bounded, unital, positive, faithful and satisfies $\varphi_2(u^k v^\ell) = \delta_{k0} v^\ell$.

Hence, the composition map $\varphi_2 \circ \varphi_1$ is linear, bounded, unital, positive and faithful. For any $\sum_{k,\ell\in\mathbb{Z}} \alpha_{k\ell} u^k v^\ell \in \mathrm{Pol}(\mathbb{T}_\theta^2)$, it satisfies

$$\varphi_2 \circ \varphi_1 \left(\sum_{k,\ell\in\mathbb{Z}} \alpha_{k\ell} u^k v^\ell\right) = \varphi_2\left(\sum_{k\in\mathbb{Z}} \alpha_{k0} u^k\right) = \alpha_{00},$$

which implies $\varphi_2 \circ \varphi_1$ extends $\tau$. We may just denote $\varphi_2 \circ \varphi_1$ by $\tau$.

For the tracial property, let $x = u^k v^\ell$ and $y = u^m v^n$, note that

$$\begin{aligned}\tau(xy) =&\tau(u^k v^\ell u^m v^n) = e^{-i\ell m\theta}\tau(u^{k+m}v^{\ell+n}) = e^{-i\ell m\theta}\delta_{k,-m}\delta_{\ell,-n} = e^{i\ell k\theta}\delta_{k,-m}\delta_{\ell,-n},\\ \tau(yx) =&\tau(u^m v^n u^k v^\ell) = e^{-ink\theta}\tau(u^{k+m}v^{\ell+n}) = e^{-ink\theta}\delta_{k,-m}\delta_{\ell,-n} = e^{i\ell k\theta}\delta_{k,-m}\delta_{\ell,-n}.\end{aligned}$$

By linearity, we have $\tau(xy) = \tau(yx)$ for all $x, y \in \mathrm{Pol}(\mathbb{T}^2_\theta)$. Since $\tau$ is bounded and $\mathrm{Pol}(\mathbb{T}^2_\theta)$ is dense in $A_\theta$, we obtain $\tau(xy) = \tau(yx)$ for all $x, y \in A_\theta$.

Therefore, we conclude $\tau = \varphi_2 \circ \varphi_1$ is a faithful tracial state on $A_\theta$. □

*Remark.* If $\theta$ is irrational, $\tau = \varphi_2 \circ \varphi_1$ is the unique faithful tracial state on $A_\theta$ [1, Proposition 7.10].

Observe that $\langle x, y\rangle := \tau(y^*x)$ is a nondegenerate and positive definite sesquilinear form (inner product) on $A_\theta$, the completion of $A_\theta$ with respect to the norm associated to this sesquilinear form is a Hilbert space which will be denoted by $L^2(\mathbb{T}^2_\theta)$. Then $A_\theta$ can be viewed as a C*-subalgebra of $B(L^2(\mathbb{T}^2_\theta))$ by treating the elements of $A_\theta$ as the left multiplication operators on $L^2(\mathbb{T}^2_\theta)$. The von Neumann algebra generated by $A_\theta$ will be denoted by $L^\infty(\mathbb{T}^2_\theta)$.

*Remark.* The above statement of viewing $A_\theta$ as a C*-subalgebra of $B(L^2(\mathbb{T}^2_\theta))$ is just the GNS representation (see, e.g., Section 3.4 in [31]) of $A_\theta$ associated to $\tau$. The von Neumann algebra generated by $A_\theta$ can be described in the following ways [31, Section 4.1 and 4.2]:

- The SOT closure of $A_\theta$ in $B(L^2(\mathbb{T}^2_\theta))$;
- The weak operator topology (WOT) closure in $B(L^2(\mathbb{T}^2_\theta))$;
- $A_\theta''$ the double commutant of $A_\theta$ by von Neumann's double commutant theorem.

If $L^\infty(\mathbb{T}^2_\theta)$ is a semifinite von Neumann algebra, we can consider the noncommutative $L^p$-spaces associated to it. The next proposition states that $L^\infty(\mathbb{T}^2_\theta)$ is indeed finite.

**Proposition 2.6.** *Let $\tau$ be the faithful tracial state on $A_\theta$ as stated in Proposition 2.5, then $\tau$ can be extended to a normal finite faithful trace on $L^\infty(\mathbb{T}^2_\theta)$.*

*Proof.* Let $(\pi_\tau, H_\tau)$ be the GNS representation of $A_\theta$ associated with $\tau$, and let $\Omega$ be the cyclic vector such that $\tau(x) = \langle \pi_\tau(x)\Omega, \Omega\rangle$ for all $x \in A_\theta$. Note that $A_\theta$ is isomorphic to $\pi_\tau(A_\theta)$ since $\tau$ is faithful. By abuse of notation, we identify $A_\theta$ with $\pi_\tau(A_\theta)$, and also identify $\tau$ on $A_\theta$ with $\langle(\cdot)\Omega, \Omega\rangle$ on $\pi_\tau(A_\theta)$. Observe that $\tau(\cdot) = \langle(\cdot)\Omega, \Omega\rangle$ is WOT continuous by definition, it extends to the WOT closure of $A_\theta$ which is just the von Neumann algebra $L^\infty(\mathbb{T}^2_\theta)$. The extension will still be denoted by $\tau$. It is finite since $\tau(\mathbf{1}) = 1$.

For any $x, y \in L^\infty(\mathbb{T}^2_\theta)$, there exist sequences $\{x_n\}_{n=1}^\infty$ and $\{y_m\}_{m=1}^\infty$ in $A_\theta$ such that they SOT converge to $x$ and $y$ respectively. Observe that the left multiplication and the right multiplication are SOT continuous. Namely, for any fixed $b \in B(H_\tau)$, the maps $B(H_\tau) \to B(H_\tau)$ given by $a \mapsto ba$ and $a \mapsto ab$ are SOT continuous. Also note that $\tau$ is tracial on $A_\theta$. So

$$\tau(xy) = \lim_{m\to\infty} \tau(xy_m) = \lim_{m\to\infty}\lim_{n\to\infty} \tau(x_n y_m) = \lim_{m\to\infty}\lim_{n\to\infty} \tau(y_m x_n) = \lim_{m\to\infty} \tau(y_m x) = \tau(yx),$$

which implies $\tau$ is tracial on $L^\infty(\mathbb{T}^2_\theta)$.

For the faithfulness, assume $x \in L^\infty(\mathbb{T}^2_\theta)_+$ such that $\tau(x) = 0$. For any $a \in A_\theta$, note that

$$\|x^{1/2}a\Omega\|^2_{H_\tau} := \langle x^{1/2}a\Omega, x^{1/2}a\Omega\rangle = \langle a^*xa\Omega, \Omega\rangle = \tau(a^*xa) = \tau(x^{1/2}aa^*x^{1/2}) \le \|a\|^2\tau(x) = 0.$$

This implies $x^{1/2}a\Omega = 0$. Since $A_\theta\Omega$ is dense in $H_\tau$, we obtain $x = 0$.

To see $\tau$ is normal, consider a bounded increasing net $(x_\alpha)$ in $L^\infty(\mathbb{T}_\theta^2)_+$. By Theorem 4.1.1 in [31], the net $(x_\alpha)$ is strongly convergent, namely, the operator $\sup_\alpha(x_\alpha)$ exists and $(x_\alpha)$ SOT converges to $\sup_\alpha(x_\alpha)$. So

$$\sup_\alpha \tau(x_\alpha) = \sup_\alpha \langle x_\alpha \Omega, \Omega\rangle = \langle \sup_\alpha x_\alpha \Omega, \Omega\rangle = \tau(\sup_\alpha x_\alpha).$$

Therefore, we conclude $\tau$ extends to a normal finite faithful trace on $L^\infty(\mathbb{T}_\theta^2)$. □

*Remark.* For $L^\infty(\mathbb{T}^2)$, the trace $\tau$ is just the normalized Lebesgue integral on $\mathbb{T}^2$. To see this, let $u : z \mapsto z_1$ and $v : z \mapsto z_2$ with $z = (z_1, z_2) \in \mathbb{T}^2$. Note that $u, v$ are commuting unitaries. Let $dz$ be the normalized Lebesgue measure on $\mathbb{T}^2$. Suppose $\sum_{k,\ell\in\mathbb{Z}} \alpha_{k\ell} u^k v^\ell \in \mathrm{Pol}(\mathbb{T}^2)$, then

$$\int_{\mathbb{T}^2} \left( \sum_{k,\ell\in\mathbb{Z}} \alpha_{k\ell} u^k v^\ell \right) dz = \sum_{k,\ell\in\mathbb{Z}} \left( \alpha_{k\ell} \int_{\mathbb{T}^2} z_1^k z_2^\ell \, dz \right) = \sum_{k,\ell\in\mathbb{Z}} \left( \alpha_{k\ell} \int_{\mathbb{T}} z_1^k \, dz_1 \int_{\mathbb{T}} z_2^\ell \, dz_2 \right)$$

Note that $\int_{\mathbb{T}} z_i^n \, dz_i = \delta_{n0}$, $i = 1, 2$, we obtain

$$\sum_{k,\ell\in\mathbb{Z}} \left( \alpha_{k\ell} \int_{\mathbb{T}} z_1^k \, dz_1 \int_{\mathbb{T}} z_2^\ell \, dz_2 \right) = \sum_{k,\ell\in\mathbb{Z}} (\alpha_{k\ell} \delta_{k0} \delta_{\ell 0}) = \alpha_{00} = \tau \left( \sum_{k,\ell\in\mathbb{Z}} \alpha_{k\ell} u^k v^\ell \right).$$

So $\tau$ coincides with the normalized Lebesgue integral on $\mathrm{Pol}(\mathbb{T}^2)$ and thus on $C(\mathbb{T}^2)$ and $L^\infty(\mathbb{T}^2)$. In this case, the trace is denoted by $\int$ instead of $\tau$, namely,

$$\int (f) := \int_{\mathbb{T}^2} f(z) \, dz, \quad \text{for all } f \in L^\infty(\mathbb{T}^2).$$

Also note that for $f \in L^\infty(\mathbb{T}^2)$, we have

$$\int (|f|^p) = \int_{\mathbb{T}^2} |f(z)|^p \, dz.$$

So, by Definition 2.2, the noncommutative $L^p$-spaces associated with $L^\infty(\mathbb{T}^2)$ coincide with the usual $L^p$-spaces $L^p(\mathbb{T}^2)$, that is, $L^p(L^\infty(\mathbb{T}^2)) \cong L^p(\mathbb{T}^2)$. For convenience, we denote the noncommutative $L^p$-spaces associated with $L^\infty(\mathbb{T}_\theta^2)$ by $L^p(\mathbb{T}_\theta^2) := L^p(L^\infty(\mathbb{T}_\theta^2))$. In particular, the Hilbert space $L^2(\mathbb{T}_\theta^2)$ which was mentioned in the previous remark is just the noncommutative $L^2$-space associated with $L^\infty(\mathbb{T}_\theta^2)$.

An alternative approach to identify $L^p(L^\infty(\mathbb{T}^2))$ and $L^p(\mathbb{T}^2)$ from the point of view of $\tau$-measurable operators (Definition 2.6) can be found in Example 2.1.5 in [16].

## 2.4 Spectral theorem

Spectral decomposition is a powerful tool when considering normal operators. For a finite-dimensional Hilbert space, the normal operators (or just normal matrices) on it can be diagonalized. For an arbitrary Hilbert space, the compact normal operators on it also admit eigenvalue decompositions. Unlike the cases above, some normal operators may not have eigenvalues, but they admit spectral decompositions which are generalized analogues of eigenvalue decompositions. This section will follow [13, Section IX.1, IX.2 and X.4] to formulate the notion of spectral measures, explain the integration with respect to a spectral measure and introduce the spectral decomposition of a normal operator.

**Definition 2.14** (Spectral measures)**.** If $X$ is a set, $\Omega$ is a $\sigma$-algebra of subsets of $X$ and $H$ is a Hilbert space. A *spectral measure* for $(X, \Omega, H)$ is a function $E : \Omega \to B(H)$ such that:

(i) For each $\Delta$ in $\Omega$, $E(\Delta)$ is a projection;

(ii) $E(\varnothing) = 0$ and $E(X) = 1$;

(iii) $E(\Delta_1 \cap \Delta_2) = E(\Delta_1)E(\Delta_2)$ for $\Delta_1$ and $\Delta_2$ in $\Omega$;

(iv) (Countably additive) If $\{\Delta_n\}_{n=1}^{\infty}$ are pairwise disjoint sets from $\Omega$, then

$$E\left(\bigcup_{n=1}^{\infty} \Delta_n\right) = \sum_{n=1}^{\infty} E(\Delta_n).$$

In the next proposition [13, Proposition IX.1.10], the integration with respect to a spectral measure is explained.

**Proposition 2.7.** *If $E$ is a spectral measure for $(X, \Omega, H)$ and $\phi : X \to \mathbb{C}$ is a bounded $\Omega$-measurable function, then there is a unique operator $A$ in $B(H)$ such that if $\varepsilon > 0$ and $\{\Delta_1, ..., \Delta_n\}$ is an $\Omega$-partition of $X$ with $\sup\{|\phi(x) - \phi(x')| : x, x' \in \Delta_k\} < \varepsilon$ for $1 \leq k \leq n$, then for any $x_k$ in $\Delta_k$,*

$$\left\| A - \sum_{k=1}^{n} \phi(x_k)E(\Delta_k) \right\| < \varepsilon.$$

*$A$ is called the integral of $\phi$ with respect to $E$ and is denoted by*

$$A = \int \phi \, dE.$$

*Remark.* Let $E$ be a spectral measure for $(X, \Omega, H)$ and $\phi : X \to \mathbb{C}$ be an $\Omega$-measurable function (not necessarily bounded) and for each $n$ let $\Delta_n = \{x \in X : n - 1 \leq |\phi(x)| < n\}$. So $\chi_{\Delta_n}\phi$ is a bounded $\Omega$-measurable function. Put $H_n = E(\Delta_n)H$. Since $\bigcup_{n=1}^{\infty} \Delta_n = X$ and the sets $\{\Delta_n\}$ are pairwise disjoint, $\bigoplus_{n=1}^{\infty} H_n = H$. If $E_n(\Delta) = E(\Delta \cap \Delta_n)$, then $E_n$ is a spectral measure for $(X, \Omega, H_n)$. Also, $\int \phi \, dE_n$ is a normal operator on $H_n$. Define

$$D_\phi = \left\{ h \in H : \sum_{n=1}^{\infty} \left\| \left( \int \phi \, dE_n \right) E(\Delta_n)h \right\|^2 < \infty \right\}$$

By Lemma X.4.4 in [13], $N_\phi : H \to H$ given by

$$N_\phi h = \sum_{i=1}^{\infty} \left( \int \phi \, dE_n \right) E(\Delta_n)h, \quad h \in D_\phi$$

is a normal operator. The operator $N_\phi$ is also denoted by

$$N_\phi = \int \phi \, dE.$$

After understanding the integration of bounded and unbounded measurable functions with respect to a spectral measure, we introduce the spectral theorem.

**Theorem 2.8** (Spectral theorem)**.** *If $N$ is a (unbounded) normal operator on a Hilbert space $H$, then there is a unique spectral measure $E$ on the Borel subsets of $\mathbb{C}$ such that*

$$N = \int z \, dE(z).$$

*In addition, if $\Delta$ is a Borel subset of $\mathbb{C}$, then $E(\Delta) = 0$ if and only if $\Delta \cap \sigma(N) = \varnothing$.*

## 2.5 Comparisons between classical and quantum optimal transport

In this section, we will discuss the differences and similarities between the key notions of optimal transport problems in different settings. In particular, the notions of classical probability measures, quantum density operators, and density operators associated with a von Neumann algebra will be compared. The similarity between the classical and quantum transport plans will be explained. In the end, an equivalent definition of quantum transport plans that uses partial traces is introduced.

### 2.5.1 Classical and quantum densities

First we recall briefly the definition of classical probability measures. A probability measure is a set function defined on a $\sigma$-algebra which ranges in the interval $[0, 1]$, and maps the empty set and the universal set to 0 and 1 respectively, and satisfies the countable additivity. Suppose $\mu$ and $\nu$ are two probability measures on a $\sigma$-algebra, then their pointwise products $\mu\nu$ and $\nu\mu$ are equal. This fact indicates the commutativity of probability measures, which is a main difference from their noncommutative counterparts, the quantum density operators and the density operators associated with a noncommutative von Neumann algebra.

In quantum mechanics, the quantum states of a physical system are described by density operators. They are the bounded positive trace-one operators on a Hilbert space. In general, the density operators do not commute. For instance, consider the Hilbert space $\mathbb{C}^2$ and density operators

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad B = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

Then

$$AB = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \neq \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} = BA.$$

Nonetheless, there are some density operators that commute with each other. Suppose the Hilbert space $H$ is separable, then there exists a countable orthonormal basis $\{x_n\}_{n=1}^{\infty}$ of $H$. Let $\{p_n\}_{n=1}^{\infty}$be a probability distribution on $\mathbb{N}$, namely, the elements are positive and sum to one. Then $\sum_{n=1}^{\infty} p_n|x_n\rangle\langle x_n|$ is a density operator (Dirac notation). Moreover, density operators of this form commute with each other and can be related to the corresponding classical probability distributions.

The density operators associated with a semifinite von Neumann algebra are positive trace-one operators in the noncommutative $L^1$-space associated with the von Neumann algebra. In particular, the set of all bounded operators on a Hilbert space is a von Neumann algebra, and the usual trace on it is normal, semifinite and faithful. Let $H$ be a Hilbert space, and let $B(H)$ be the set of all bounded operators on $H$. The noncommutative $L^p$-spaces associated with $B(H)$ are also known as the $p$-th Schatten classes. Specifically, the first Schatten class is the space of trace class operators. Indeed, if a von Neumann algebra is a type I factor, then it is isomorphic to the von Neumann algebra of all bounded operators on some Hilbert space. So the density operators associated with a von Neumann algebra can be considered as a generalization of quantum density operators. Moreover, for commutative von Neumann algebras, the noncommutative $L^1$-space associated to it can be identified as the space of absolute integrable functions on a measure space. So the density operators are just the positive absolute integrable functions whose integrals are one. These functions can be interpreted as probability measures by calculating their integrals over some measurable sets. Indeed, these functions are just the probability density functions of some random variables. Though there are some probability measures that are not of this form (e.g., Dirac measures), the notion of absolutely continuous probability measures are generalized and perfectly captured by the notion of density operators associated with a von Neumann algebra.

### 2.5.2 Classical and quantum transport plans

Then we consider the transport plans. Recall the definition of transport plans in classical optimal transport problems. Suppose $\mu$ and $\nu$ are two probability measures on Polish spaces $X$ and $Y$ respectively, a transport plan between $\mu$ and $\nu$ is the probability measure $\pi$ on the product space $X \times Y$ which satisfies

$$\pi(A \times Y) = \mu(A), \quad \pi(X \times B) = \nu(B)$$

for all measurable subsets $A$ of $X$ and $B$ of $Y$. Or, equivalently, for all functions $\varphi, \psi$ in a suitable class of test functions (e.g., $(\varphi, \psi) \in L^\infty(d\mu) \times L^\infty(d\nu)$), it satisfies

$$\int_{X\times Y} (\varphi(x) + \psi(y))\, d\pi(x, y) = \int_X \varphi(x)\, d\mu(x) + \int_Y \psi(y)\, d\nu(y).$$

It is worth noting that the trace is to density operators as the integral is to probability measures. Suppose $R$ and $S$ are two density operators on a Hilbert space $H$. We can analogously define the transport plans between $R$ and $S$ as the density operators $T$ on the tensor product Hilbert space $H \otimes H$ such that

$$\operatorname{tr}((A \otimes I)T) = \operatorname{tr}(AR), \quad \operatorname{tr}((I \otimes A)T) = \operatorname{tr}(AS) \tag{2.1}$$

for each $A \in B(H)$. This is just the definition of transport plans between two density operators in [21] and [18, Definition 2.1].

### 2.5.3 Partial traces

The condition (2.1) has an equivalent formulation in terms of partial traces. Denote by $L^1(H)$ the trace class operators on $H$. Recall the partial traces are the maps $\operatorname{id} \otimes \operatorname{tr}, \operatorname{tr} \otimes \operatorname{id} : L^1(H \otimes H) \to L^1(H)$ given by

$$\operatorname{id} \otimes \operatorname{tr} : x \otimes y \mapsto x \operatorname{tr}(y), \quad \operatorname{tr} \otimes \operatorname{id} : x \otimes y \mapsto \operatorname{tr}(x) y.$$

By Theorem 2.28 in [2], the condition (2.1) is equivalent to

$$(\operatorname{id} \otimes \operatorname{tr})(T) = R, \quad (\operatorname{tr} \otimes \operatorname{id})(T) = S. \tag{2.2}$$

Observe that the condition (2.2) is formally simpler and more intuitive than condition (2.1). So we will define the transport plans between two density operators associated with some von Neumann algebra analogously to (2.2). Indeed, this is equivalent to defining the transport plans analogously to (2.1). To see this, let $(\mathcal{M}, \tau)$ be a noncommutative measure space. The partial traces $\operatorname{id} \otimes \tau$ and $\tau \otimes \operatorname{id}$ are defined similarly to the usual partial traces above. For any $x \in \mathcal{M}$

and $T=\sum_{k=1}^{n} c_k T_{1k} \otimes T_{2k} \in L^1(\mathcal{M}) \otimes L^1(\mathcal{M})$, we have

$$\begin{aligned}(\tau \otimes \tau)((x \otimes \mathbf{1})T) = \sum_{k=1}^{n} c_k (\tau \otimes \tau)(x T_{1k} \otimes T_{2k}) &= \sum_{k=1}^{n} c_k \tau(x T_{1k}) \tau(T_{2k}) \\ &= \tau(x \sum_{k=1}^{n} c_k T_{1k} \tau(T_{2k})) = \tau(x(\mathrm{id} \otimes \tau)(T)).\end{aligned}$$

Note that if $(\tau \otimes \tau)((x \otimes \mathbf{1})T) = \tau(xy)$ for some $y \in L^1(\mathcal{M})$, then $(\mathrm{id} \otimes \tau)(T) = y$ since $\mathcal{M}$ is the dual of $L^1(\mathcal{M})$ [16, Theorem 3.4.24]. Conversely, if $(\mathrm{id} \otimes \tau)(T) = y$, then $(\tau \otimes \tau)((x \otimes \mathbf{1})T) = \tau(xy)$. Also note the fact that $L^1(\mathcal{M}) \otimes L^1(\mathcal{M})$ is dense in $L^1(\mathcal{M} \bar{\otimes} \mathcal{M})$ with respect to $\|\cdot\|_1$, and $\mathrm{id} \otimes \tau$ is continuous with respect to $\|\cdot\|_1$ [22, Lemma 5.3]. Thus, for any $T \in L^1(\mathcal{M} \bar{\otimes} \mathcal{M})$ and $x \in \mathcal{M}$, we have $(\tau \otimes \tau)((x \otimes I)T) = \tau(x(\mathrm{id} \otimes \tau)(T))$. Similarly, we also have $(\tau \otimes \tau)((I \otimes x)T) = \tau(x(\tau \otimes \mathrm{id})(T))$.

# Chapter 3

# Problem statement

In this chapter, the optimal transport problem on the noncommutative 2-torus is formulated. First, we will define the transport plans between two density operators associated with $L^\infty(\mathbb{T}^2_\theta)$, the transport cost function, and the quantum Wasserstein distance on the noncommutative 2-torus. Then we will state the main concern of this thesis, the triangle inequality problem of the proposed distance.

Let $\theta_1$ and $\theta_2$ be real numbers. Consider the von Neumann algebras $L^\infty(\mathbb{T}^2_{\theta_1})$ and $L^\infty(\mathbb{T}^2_{\theta_2})$, and corresponding normal finite faithful traces $\tau_1$ and $\tau_2$. Let $\rho_1 \in \mathcal{D}(\mathbb{T}^2_{\theta_1})$ and $\rho_2 \in \mathcal{D}(\mathbb{T}^2_{\theta_2})$ be two density operators. The set of transport plans, or couplings, between $\rho_1$ and $\rho_2$ is

$$\mathcal{C}(\rho_1, \rho_2) \coloneqq \{T \in \mathcal{D}(\mathbb{T}^2_{\theta_1} \bar{\otimes} \mathbb{T}^2_{\theta_2}) : (\mathrm{id} \otimes \tau_2)(T) = \rho_1 \text{ and } (\tau_1 \otimes \mathrm{id})(T) = \rho_2\}.$$

Suppose $u_i$ and $v_i$ are the unitary generators of $L^\infty(\mathbb{T}^2_\theta)$, and $\mathbf{1}_i$ is the unit of $L^\infty(\mathbb{T}^2_{\theta_i})$ $(i = 1, 2)$. The cost function for transport plans in $\mathcal{C}(\rho_1, \rho_2)$ is defined as

$$C(u_1, v_1, u_2, v_2) = |u_1 \otimes \mathbf{1}_2 - \mathbf{1}_1 \otimes u_2|^2 + |v_1 \otimes \mathbf{1}_2 - \mathbf{1}_1 \otimes v_2|^2. \tag{3.1}$$

Then the quantum Wasserstein distance between $\rho_1$ and $\rho_2$ is defined as

$$d(\rho_1, \rho_2) = \left( \inf_{T \in \mathcal{C}(\rho_1, \rho_2)} (\tau_1 \otimes \tau_2)(CT) \right)^{\frac{1}{2}}. \tag{3.2}$$

*Remark.* If one of the $\theta_i$ is equal to 0, the cost function is an operator-valued function on $\mathbb{T}^2$. To see this, assume $\theta_2 = 0$ without loss of generality, the unitary generators of $L^\infty(\mathbb{T}^2)$ can be chosen as the functions $z_k : z \mapsto z_k$ with $k = 1, 2$ and $z = (z_1, z_2) \in \mathbb{T}^2$ (Here $z_k$ represents the $k$-th coordinate of $z$ and also represents the function that maps $z$ to its $k$-th coordinate by abuse of notation). Then for fixed $z \in \mathbb{T}^2$, the cost function can be written as

$$C(u_1, v_1, z_1, z_2) = |u_1 \otimes 1 - \mathbf{1}_1 \otimes z_1|^2 + |v_1 \otimes 1 - \mathbf{1}_1 \otimes z_2|^2 = |u_1 - z_1 \cdot \mathbf{1}_1|^2 + |v_1 - z_2 \cdot \mathbf{1}_1|^2.$$

In this case, we will abbreviate $C(u_1, v_1, z_1, z_2)$ to $C(u_1, v_1, z)$, namely,

$$C(u_1, v_1, z) = C(u_1, v_1, z_1, z_2) = |u_1 - z_1 \cdot \mathbf{1}_1|^2 + |u_2 - z_2 \cdot \mathbf{1}_1|^2.$$

*Remark.* If both of the $\theta_i$ are equal to 0, the cost function is just $|z - \zeta|^2$ with $z = (z_1, z_2), \zeta = (\zeta_1, \zeta_2) \in \mathbb{T}^2$. To see this, choose the unitary generators of $L^\infty(\mathbb{T}^2)$ similarly to the previous remark. Then the cost function can be written as

$$C(z_1, z_2, \zeta_1, \zeta_2) = |z_1 \otimes 1 - 1 \otimes \zeta_1|^2 + |z_2 \otimes 1 - 1 \otimes \zeta_2|^2 = |z_1 - \zeta_1|^2 + |z_2 - \zeta_2|^2 = |z - \zeta|^2.$$

So the distance between classical densities $f, g \in \mathcal{D}(\mathbb{T}^2)$ is

$$d(f,g)^2 = \inf_{T\in\mathcal{C}(f,g)} (\int \otimes \int)(CT) = \inf_{T\in\mathcal{C}(f,g)} \iint_{\mathbb{T}^2\times\mathbb{T}^2} |z-\zeta|^2 T(z,\zeta)\, dz d\zeta.$$

In this classical case, the quantum Wasserstein distance between $f$ and $g$ reduces to the quadratic Wasserstein distance between $f$ and $g$, considering only absolutely continuous transport plans.

With the distance defined (3.2), it is natural to check whether it satisfies the properties of a metric. The symmetry is clear from the construction. The positivity remains to be studied, but the non-negativity is clear. Indeed, observe that the cost function $C = C(u_1, v_1, u_2, v_2)$ is positive and bounded. The transport plan $T \in \mathcal{C}(\rho_1, \rho_2)$ is a positive element in $L^1(\mathbb{T}^2_{\theta_1} \bar{\otimes} \mathbb{T}^2_{\theta_2})$. Then $C^{\frac{1}{2}}TC^{\frac{1}{2}}$ is positive, and $C^{\frac{1}{2}}T$ and $C^{\frac{1}{2}}TC^{\frac{1}{2}}$ are in $L^1(\mathbb{T}^2_{\theta_1} \bar{\otimes} \mathbb{T}^2_{\theta_2})$ [16, Proposition 3.4.1]. The non-negativity follows from

$$(\tau_1 \otimes \tau_2)(CT) = (\tau_1 \otimes \tau_2)(C^{\frac{1}{2}}C^{\frac{1}{2}}T) = (\tau_1 \otimes \tau_2)(C^{\frac{1}{2}}TC^{\frac{1}{2}}) \geq 0,$$

where the second equality is obtained from Proposition 3.4.2 in [16].

The property of the triangle inequality is the main concern of this thesis. As in Theorem 1.1, we will investigate the triangle inequality of (3.2) in some special case. Since $L^\infty(\mathbb{T}^2_\theta)$ does not contain rank-one projection, we will only investigate the case where the middle point is a classical density. Namely, suppose $\rho_1, \rho_3 \in \mathcal{D}(\mathbb{T}^2_\theta)$ and $g \in \mathcal{D}(\mathbb{T}^2)$, we will prove the following inequality

$$d(\rho_1, \rho_3) \leq d(\rho_1, g) + d(g, \rho_3). \tag{3.3}$$

# Chapter 4

# Triangle inequality when the middle point is a classical density

In this chapter, we first review the original proof of Theorem 1.1 [21, Theorem 3.1] and discuss how it motivated the proof of the problem stated in Chapter 3. Then we prove an inequality for the cost function and some assertions about spectral measures. In the end, we will consider the case where the middle point is a classical density, and prove the triangle inequality of the quantum Wasserstein distance (3.2) by applying the obtained inequality and assertions.

## 4.1 Review of Golse and Paul's proof

In this section, we will first interpret the original proof of Theorem 1.1 given in Section 3 of [21] by Golse and Paul. Then we will discuss the main idea and inspiration of their proof, and outline the essential steps of the proof for the problem considered in Chapter 3.

### 4.1.1 Original proof

In this subsection, we will only review the proof to Theorem 1.1 for the case where the middle point is a classical probability measure. The arguments in this subsection closely follow the original proof [21, Section 3].

For convenience, denote by $\mathfrak{H}$ the Hilbert space $L^2(\mathbb{R}^d)$. Recall the definition of $\mathfrak{d}$ by (1.4), the distance $\mathfrak{d}$ is only defined for absolutely continuous probability measures with finite second moments $\mathcal{P}_2^{ac}(\mathbb{R}^{2d})$ in the classical-to-quantum case. In order to extend $\mathfrak{d}$ to all probability measures with finite second moments $\mathcal{P}_2(\mathbb{R}^{2d})$ in this case, some density arguments are used. They proved that $\mathcal{P}_2^{ac}(\mathbb{R}^{2d})$ is dense in $\mathcal{P}_2(\mathbb{R}^{2d})$ with respect to the metric $W_2$ [21, Lemma 3.2]. Then for any fixed finite energy density operator $\rho \in \mathcal{D}_2(\mathfrak{H})$, if $\mathfrak{d}(\cdot, \rho)$ is continuous on $\mathcal{P}_2^{ac}(\mathbb{R}^{2d})$ with respect to $W_2$, then we can conclude $\mathfrak{d}(\cdot, \rho)$ admits a unique extension to $\mathcal{P}_2(\mathbb{R}^{2d})$. Indeed, it is a consequence of Theorem 3.5 in [20] that $\mathfrak{d}(\cdot, \rho)$ is Lipschitz continuous. Namely, for all $\rho_1, \rho_2 \in \mathcal{P}_2^{ac}(\mathbb{R}^{2d})$, we have

$$|\mathfrak{d}(\rho_1, \rho) - \mathfrak{d}(\rho_2, \rho)| \leq \mathfrak{d}(\rho_1, \rho_2) = W_2(\rho_1, \rho_2).$$

Therefore, $\mathfrak{d}$ can be uniquely extended to $\mathfrak{D}_2 \times \mathfrak{D}_2$, where $\mathfrak{D}_2 := \mathcal{D}_2(\mathfrak{H}) \cup \mathcal{P}_2(\mathbb{R}^{2d})$.

Next we turn to the proof of the triangle inequality when the middle point $\rho_2$ is in $\mathcal{P}_2(\mathbb{R}^{2d})$. For $\rho_1, \rho_3 \in \mathfrak{D}_2(\mathfrak{H})$, there are three possibilities in total. Namely,

- $\rho_1, \rho_3 \in \mathcal{P}_2(\mathbb{R}^{2d})$. This is just the triangle inequality of classical quadratic Wasserstein distance [38, Theorem 7.3].
- $\rho_1 \in \mathcal{P}_2(\mathbb{R}^{2d})$ and $\rho_3 \in \mathcal{D}_2(\mathfrak{H})$, or $\rho_1 \in \mathcal{D}_2(\mathfrak{H})$ and $\rho_3 \in \mathcal{P}_2(\mathbb{R}^{2d})$. The density arguments above and Theorem 3.5 in [20] justify the triangle inequality.
- $\rho_1, \rho_3 \in \mathcal{D}_2(\mathfrak{H})$. This is the case that needs to be studied.

Based on the previous density arguments, we can just assume $\rho_2 \in \mathcal{P}_2^{ac}(\mathbb{R}^{2d})$. In the remainder of this subsection, we will consider the triangle inequality in the case where $\rho_1, \rho_3 \in \mathcal{D}_2(\mathfrak{H})$ and $\rho_2 = \rho_2(x,\xi)dxd\xi \in \mathcal{P}_2^{ac}(\mathbb{R}^{2d})$.

Observe that the distance between two densities are determined by the set of transport plans between them. To establish the connection between the three distances $\mathfrak{d}(\rho_1,\rho_2), \mathfrak{d}(\rho_1,\rho_3)$ and $\mathfrak{d}(\rho_2,\rho_3)$, we can think of establishing connections between transport plans. Given two transport plans $Q^1(x,\xi) \in \mathcal{C}(\rho_1,\rho_2)$ and $Q^3(x,\xi) \in \mathcal{C}(\rho_2,\rho_3)$, we hope to construct a "joint transport plan" $Q_{13}(x,\xi)$ such that

- $Q^{13}$ becomes $Q^1$ after tracing out its third "coordinate";
- $Q^{13}$ becomes $Q^3$ after tracing out its first "coordinate";
- $Q^{13}$ becomes a transport plan between $\rho_1$ and $\rho_3$ after integrating over the second "coordinate".

To this end, we consider the following disintegration result [20, Lemma A.4].

**Lemma 4.1.** *Let $f \in \mathcal{P}^{ac}(\mathbb{R}^{2d})$ and $R \in \mathcal{D}_2(\mathfrak{H})$, and let $Q \in \mathcal{C}(f,R)$. There exists a weakly measurable $B(\mathfrak{H})$-valued function $(x,\xi) \mapsto Q_f(x,\xi)$ defined a.e. on $\mathbb{R}^{2d}$ such that*

$$Q_f(x,\xi) = Q_f(x,\xi)^* \geq 0, \quad \operatorname{tr}(Q_f(x,\xi)) = 1, \quad \textit{and} \quad Q(x,\xi) = f(x,\xi)Q_f(x,\xi)$$

*for a.e. $(x,\xi) \in \mathbb{R}^{2d}$.*

With this lemma, we construct the "joint transport plan" $Q^{13}$ for $Q^1$ and $Q^3$ as

$$Q^{13}(x,\xi) := Q^1(x,\xi) \otimes Q^3_{\rho_2}(x,\xi).$$

Then we examine the properties of $Q^{13}$. First observe that

$$Q^{13}(x,\xi) = Q^{13}(x,\xi)^* \geq 0.$$

Define

$$T_{13} := \int_{\mathbb{R}^{2d}} Q^{13}(x,\xi)\,dxd\xi.$$

For any $\psi \in \mathfrak{H} \otimes \mathfrak{H}$, since $Q^{13}(x,\xi)$ is positive a.e. $(x,\xi) \in \mathbb{R}^{2d}$, we have

$$\langle T_{13}\psi, \psi\rangle = \int_{\mathbb{R}^{2d}} \langle Q^{13}(x,\xi)\psi,\psi\rangle\,dxd\xi \geq 0.$$

So $T_{13}$ is positive. Also note that

$$\begin{aligned}
\operatorname{tr}(T_{13}) &= \int_{\mathbb{R}^{2d}} \operatorname{tr}(Q^{13}(x,\xi))\,dxd\xi \\
&= \int_{\mathbb{R}^{2d}} \operatorname{tr}(Q^1(x,\xi))\operatorname{tr}(Q^2_{\rho_2}(x,\xi))\,dxd\xi \\
&= \int_{\mathbb{R}^{2d}} \rho_2(x,\xi)\,dxd\xi \\
&=1,
\end{aligned}$$

so $T_{13}$ is a density operator. Moreover, it is a transport plan between $\rho_1$ and $\rho_3$. Indeed, for all $A \in B(\mathfrak{H})$, we have

$$
\begin{aligned}
\operatorname{tr}(T_{13}(A \otimes I)) &= \int_{\mathbb{R}^{2d}} \operatorname{tr}(Q^{13}(x,\xi)(A \otimes I))\, dxd\xi \\
&= \int_{\mathbb{R}^{2d}} \operatorname{tr}(Q^{1}(x,\xi)A)\operatorname{tr}(Q^{3}_{\rho_2}(x,\xi))\, dxd\xi \\
&= \int_{\mathbb{R}^{2d}} \operatorname{tr}(Q^{1}(x,\xi)A)\, dxd\xi \\
&= \operatorname{tr}\left(\int_{\mathbb{R}^{2d}} Q^{1}(x,\xi)\, dxd\xi\, A\right) \\
&= \operatorname{tr}(\rho_1 A),
\end{aligned}
$$

and

$$
\begin{aligned}
\operatorname{tr}(T_{13}(I \otimes A)) &= \int_{\mathbb{R}^{2d}} \operatorname{tr}(Q^{13}(x,\xi)(I \otimes A))\, dxd\xi \\
&= \int_{\mathbb{R}^{2d}} \operatorname{tr}(Q^{1}(x,\xi))\operatorname{tr}(Q^{3}_{\rho_2}(x,\xi)A) \\
&= \int_{\mathbb{R}^{2d}} \rho_2(x,\xi)\operatorname{tr}(Q^{3}_{\rho_2}(x,\xi)A)\, dxd\xi \\
&= \int_{\mathbb{R}^{2d}} \operatorname{tr}(Q^{3}(x,\xi)A)\, dxd\xi \\
&= \operatorname{tr}\left(\int_{\mathbb{R}^{2d}} Q^{3}(x,\xi)\, dxd\xi\, A\right) \\
&= \operatorname{tr}(\rho_3 A).
\end{aligned}
$$

Therefore, $T_{13} \in \mathcal{C}(\rho_1, \rho_3)$ and $Q^{13}(x,\xi)$ is the desired "joint transport plan".

Next, we will prove an inequality involving quantum-to-quantum, quantum-to-classical and classical-to-quantum cost operators, which is the essence of the desired triangle inequality. To achieve this, recall the elementary Peter-Paul inequality. For $a, b \in \mathbb{R}$ and $\varepsilon > 0$, we have $2ab \le \varepsilon a^2 + b^2/\varepsilon$. The next lemma is an operator version of Peter-Paul inequality [20, Lemma A.1].

**Lemma 4.2.** *Let $L_1, L_2$ be (unbounded) self-adjoint operators on $L^2(\mathbb{R}^d)$. For each $\alpha > 0$*

$$\langle\psi|L_1L_2 + L_2L_1|\psi\rangle \le \alpha\langle\psi|L_1^2|\psi\rangle + \frac{1}{\alpha}\langle\psi|L_2^2|\psi\rangle, \quad \psi \in \operatorname{dom}(L_1) \cap \operatorname{dom}(L_2). \tag{4.1}$$

*Remark.* The inequality (4.1) used Dirac notation. It is heuristic since it is formally similar to the elementary Peter-Paul inequality. However, this formulation may be ambiguous. The expression $\langle\psi|L_1L_2|\psi\rangle$ could mean $\langle L_1L_2\psi, \psi\rangle$ or $\langle L_2\psi, L_1\psi\rangle$, while it is possible that $L_2\psi \notin \operatorname{dom}(L_1)$. It is more rigorous to formulate (4.1) in terms of inner product as

$$\langle L_2\psi, L_1\psi\rangle + \langle L_1\psi, L_2\psi\rangle \le \alpha\langle L_1\psi, L_1\psi\rangle + \frac{1}{\alpha}\langle L_2\psi, L_2\psi\rangle.$$

Then we consider the quantum-to-quantum cost operator. For fixed $(y,\eta) \in \mathbb{R}^{2d}$ $(y, \eta \in \mathbb{R}^d)$, we can rewrite $C(x, \hbar\nabla_x, z, \hbar\nabla_z)$ as

$$C(x, \hbar\nabla_x, z, \hbar\nabla_z) = \sum_{j=1}^{d}(x_j - y_j + y_j - z_j)^2 - (-i\hbar\partial_{x_j} - \eta_j + \eta_j + i\hbar\partial_{z_j})^2.$$

By expanding the square, we have

$$C(x,\hbar\nabla_x,z,\hbar\nabla_z)=c(y,\eta,x,-i\hbar\nabla_x)+c(y,\eta,z,-i\hbar\nabla_z)\\+2\sum_{j=1}^{d}\left((x_j-y_j)(y_j-z_j)+(-i\hbar\partial_{x_j}-\eta_j)(n_j+i\hbar\partial_{z_j})\right).$$

Note that the vectors in $\mathfrak{H}\otimes\mathfrak{H}$ are functions of $(x,z)$ $(x,z\in\mathbb{R}^d)$. They can also be viewed as $\mathfrak{H}$-valued functions on $\mathbb{R}^d$. To specify the variables, denote by $\mathbb{R}^d_x$ the $d$-dimensional Euclidean space for $x$, and denote by $\mathbb{R}^d_z$ the $d$-dimensional Euclidean space for $z$. Then $\mathfrak{H}\otimes\mathfrak{H}$ can be interpreted as

$$\mathfrak{H}\otimes\mathfrak{H}\cong L^2(\mathbb{R}^d_x;L^2(\mathbb{R}^d_z))\cong L^2(\mathbb{R}^d_z;L^2(\mathbb{R}^d_x)).$$

Observe that $(x_j-y_j)$ is the multiplication operator on $L^2(\mathbb{R}^d_z;L^2(\mathbb{R}^d_x))$, and its domain is given by

$$\mathrm{dom}(x_j-y_j)=L^2(\mathbb{R}^d_z;L^2(\mathbb{R}^d_x,|x|^2dx)).$$

Similarly,

$$\mathrm{dom}(y_j-z_j)=L^2(\mathbb{R}^d_x;L^2(\mathbb{R}^d_z,|z|^2dz)),\\\mathrm{dom}(-i\hbar\partial_{x_j}-\eta_j)=L^2(\mathbb{R}^d_z;H^1(\mathbb{R}^d_x)),\\\mathrm{dom}(n_j+i\hbar\partial_{z_j})=L^2(\mathbb{R}^d_x;H^1(\mathbb{R}^d_z)),$$

where $H^1$ denotes the Sobolev space. By applying the previous lemma to the cross terms, we obtain the following inequality for cost operators.

**Lemma 4.3.** *For each $\alpha>0$ and each $(y,\eta)\in\mathbb{R}^{2d}$, one has*

$$C(x,\hbar\nabla_x,z,\hbar\nabla_z)\le(1+\alpha)c(y,\eta,x,-i\hbar\nabla_x)\otimes I_{\mathfrak{H}}+(1+\frac{1}{\alpha})I_{\mathfrak{H}}\otimes c(y,\eta,z,-i\hbar\nabla_z).$$

*In other words, for all $\alpha>0$, if*

$$\phi\equiv\phi(x,z)\in L^2\left(\mathbb{R}^d_x;L^2(\mathbb{R}^d_z,|z|^2dz)\cap H^1(\mathbb{R}^d_z)\right)\cap L^2\left(\mathbb{R}^d_z;L^2(\mathbb{R}^d_x,|x|^2dx)\cap H^1(\mathbb{R}^d_x)\right),\tag{4.2}$$

*then*

$$\langle\phi|C(x,\hbar\nabla_x,z,\hbar\nabla_z)|\phi\rangle\le(1+\alpha)\int_{\mathbb{R}^d}\langle\phi(\cdot,z)|c_\hbar(y,\eta)|\phi(\cdot,z)\rangle\,dz\\+(1+\frac{1}{\alpha})\int_{\mathbb{R}^d}\langle\phi(x,\cdot)|c_\hbar(y,\eta)|\phi(x,\cdot)\rangle\,dx.$$

Next we will derive the desired triangle inequality from previous preparations. To show the condition (4.2) of Lemma 4.3 can be satisfied, the spectral decompositions of transport plans are considered. For a.e. $(y,\eta)\in\mathbb{R}^{2d}$, the operators $Q^1(y,\eta),Q^3(y,\eta)$ are positive and trace-class (and thus compact), so they admit eigenvalue decompositions

$$Q^1(y,\eta)=\sum_{k\ge1}\lambda^1_k(y,\eta)|e^1_k(y,\eta)\rangle\langle e^1_k(y,\eta)|,\\Q^3(y,\eta)=\sum_{k\ge1}\lambda^3_k(y,\eta)|e^3_k(y,\eta)\rangle\langle e^3_k(y,\eta)|,$$

where $e_k^1(y,\eta), e_k^3(y,\eta)$ are the unit eigenvectors of $Q^1(y,\eta), Q^3(y,\eta)$, and $\lambda_k^1(y,\eta), \lambda_k^3(y,\eta)$ are the corresponding eigenvalues. Moreover, $(e_k^1(y,\eta))_{k\geq 1}$ and $(e_k^3(y,\eta))_{k\geq 1}$ are two orthonormal bases for $\mathfrak{H}$. Recall that $Q^1 \in \mathcal{C}(\rho_1,\rho_2)$ and $Q^3 \in \mathcal{C}(\rho_2,\rho_3)$, we have $\lambda_k^1, \lambda_k^3$ are measurable real-valued functions and nonnegative a.e., while $e_k^1, e_k^3$ are weakly measurable $\mathfrak{H}$-valued functions. In addition,

$$\rho_2(y,\eta) = \mathrm{tr}(Q^1(y,\eta)) = \sum_{k\geq 1} \langle Q^1(y,\eta)e_k^1(y,\eta), e_k^1(y,\eta)\rangle = \sum_{k\geq 1} \lambda_k^1(y,\eta).$$

Similarly,

$$\rho_2(y,\eta) = \sum_{k\geq 1} \lambda_k^3(y,\eta).$$

Recall that $Q^3(y,\eta) = \rho_2(y,\eta)Q^3_{\rho_2}(y,\eta)$, so $Q^3_{\rho_2}(y,\eta)$ and $Q^3(y,\eta)$ commute. Then $Q^3_{\rho_2}(y,\eta)$ can also be decomposed under the orthonormal basis $(e_k^3(y,\eta))_{k\geq 1}$ as

$$Q^3_{\rho_2}(y,\eta) = \sum_{k\geq 1} \mu_k(y,\eta)|e_k^3(y,\eta)\rangle\langle e_k^3(y,\eta)|,$$

where $\mu_k(y,\eta)$ are eigenvalues of $Q^3_{\rho_2}(y,\eta)$ and satisfy

$$\mu_k(y,\eta) \geq 0, \quad \sum_{k\geq 1} \mu_k(y,\eta) = 1, \quad \rho_2(y,\eta)\mu_k(y,\eta) = \lambda_k^3(y,\eta).$$

Hence,

$$\begin{aligned} Q^{13}(y,\eta) =& Q^1(y,\eta) \otimes Q^3_{\rho_2}(y,\eta) \\ =& \sum_{k,l\geq 1} \lambda_k^1(y,\eta)\mu_l(y,\eta)|e_k^1(y,\eta)\otimes e_l^3(y,\eta)\rangle\langle e_k^1(y,\eta)\otimes e_l^3(y,\eta)| \\ =& \sum_{k,l\geq 1} \lambda_{kl}^{13}(y,\eta)|e_k^1(y,\eta)\otimes e_l^3(y,\eta)\rangle\langle e_k^1(y,\eta)\otimes e_l^3(y,\eta)|, \end{aligned}$$

where $\lambda_{kl}^{13}(y,\eta) = \lambda_k^1(y,\eta)\mu_l(y,\eta)$. Then $\lambda_{kl}^{13}$ are measurable real-valued functions on $\mathbb{R}^{2d}$ and positive a.e., and satisfy

$$\sum_{k\geq 1} \lambda_{kl}^{13}(y,\eta) = \mu_l(y,\eta)\sum_{k\geq 1} \lambda_k^1(y,\eta) = \rho_2(y,\eta)\mu_l(y,\eta) = \lambda_l^3(y,\eta),$$

and

$$\sum_{l\geq 1} \lambda_{kl}^{13}(y,\eta) = \lambda_k^1(y,\eta)\sum_{l\geq 1} \mu_l(y,\eta) = \lambda_k^1(y,\eta).$$

In particular,

$$\lambda_{kl}^{13}(y,\eta) \leq \min(\lambda_k^1(y,\eta), \lambda_l^3(y,\eta)).$$

So, if $\lambda_{kl}^{13}(y,\eta) > 0$, then $\lambda_k^1(y,\eta) > 0$ and $\lambda_l^3(y,\eta) > 0$. It has been proven in Lemma 2.1 of [21] that

$$\begin{aligned} &\sum_{k\geq 1}\int_{\mathbb{R}^{2d}} \lambda_k^1(y,\eta)\langle e_k^1(y,\eta)|\hat{H}(x,\hbar\nabla_x)|e_k^1(y,\eta)\rangle\, dyd\eta \\ =& \int_{\mathbb{R}^{2d}} \mathrm{tr}(Q^1(y,\eta)^{\frac{1}{2}}\hat{H}(x,\hbar\nabla_x)Q^1(y,\eta)^{\frac{1}{2}})\, dyd\eta \\ =& \,\mathrm{tr}(\rho_1^{\frac{1}{2}}\hat{H}(x,\hbar\nabla_x)\rho_1^{\frac{1}{2}}) < \infty. \end{aligned}$$

Observe that if $\lambda_k^1(y,\eta)>0$, then

$$e_k^1(y,\eta)\in L^2(\mathbb{R}_x^d,|x|^2dx)\cap H^1(\mathbb{R}_x^d).$$

Similarly, if $\lambda_k^3(y,\eta)>0$, then

$$e_l^3(y,\eta)\in L^2(\mathbb{R}_z^d,|z|^2dz)\cap H^1(\mathbb{R}_z^d).$$

Therefore, if $\lambda_{kl}^{13}(y,\eta)>0$, then

$$e_k^1(y,\eta)\otimes e_l^3(y,\eta)\in L^2\left(\mathbb{R}_x^d;L^2(\mathbb{R}_z^d,|z|^2dz)\cap H^1(\mathbb{R}_z^d)\right)\cap L^2\left(\mathbb{R}_z^d;L^2(\mathbb{R}_x^d,|x|^2dx)\cap H^1(\mathbb{R}_x^d)\right).$$

By Lemma 4.3, we have

$$\begin{aligned}
&\langle e_k^1(y,\eta)\otimes e_l^3(y,\eta)|C(x,\hbar\nabla_x,z,\hbar\nabla_z)|e_k^1(y,\eta)\otimes e_l^3(y,\eta)\rangle\\
\le&(1+\alpha)\langle e_k^1(y,\eta)\otimes e_l^3(y,\eta)|c_\hbar(y,\eta)\otimes I_{\mathfrak{H}}|e_k^1(y,\eta)\otimes e_l^3(y,\eta)\rangle\\
&+(1+\frac{1}{\alpha})\langle e_k^1(y,\eta)\otimes e_l^3(y,\eta)|I_{\mathfrak{H}}\otimes c_\hbar(y,\eta)|e_k^1(y,\eta)\otimes e_l^3(y,\eta)\rangle\\
=&(1+\alpha)\langle e_k^1(y,\eta)|c_\hbar(y,\eta)|e_k^1(y,\eta)\rangle+(1+\frac{1}{\alpha})\langle e_l^3(y,\eta)|c_\hbar(y,\eta)|e_l^3(y,\eta)\rangle.
\end{aligned}$$

Then

$$\begin{aligned}
&\operatorname{tr}\left(Q^{13}(y,\eta)^{\frac12}C_\hbar Q^{13}(y,\eta)^{\frac12}\right)\\
=&\sum_{k,l\ge1}\lambda_{kl}^{13}(y,\eta)\langle e_k^1(y,\eta)\otimes e_l^3(y,\eta)|C_\hbar|e_k^1(y,\eta)\otimes e_l^3(y,\eta)\rangle\\
\le&(1+\alpha)\sum_{k,l\ge1}\lambda_{kl}^{13}(y,\eta)\langle e_k^1(y,\eta)|c_\hbar(y,\eta)|e_k^1(y,\eta)\rangle\\
&+(1+\frac{1}{\alpha})\sum_{k,l\ge1}\lambda_{kl}^{13}(y,\eta)\langle e_l^3(y,\eta)|c_\hbar(y,\eta)|e_l^3(y,\eta)\rangle\\
=&(1+\alpha)\sum_{k\ge1}\lambda_k^1(y,\eta)\langle e_k^1(y,\eta)|c_\hbar(y,\eta)|e_k^1(y,\eta)\rangle\\
&+(1+\frac{1}{\alpha})\sum_{l\ge1}\lambda_l^3(y,\eta)\langle e_l^3(y,\eta)|c_\hbar(y,\eta)|e_l^3(y,\eta)\rangle\\
=&(1+\alpha)\operatorname{tr}(Q^1(y,\eta)^{\frac12}c_\hbar(y,\eta)Q^1(y,\eta)^{\frac12})\\
&+(1+\frac{1}{\alpha})\operatorname{tr}(Q^3(y,\eta)^{\frac12}c_\hbar(y,\eta)Q^3(y,\eta)^{\frac12}).
\end{aligned}$$

For unbounded cost operator $C_\hbar$, the operator $(I+\frac{1}{n}C_\hbar)^{-1}C_\hbar$ is bounded. Therefore,

$$\begin{aligned}
\operatorname{tr}\left((I+\frac{1}{n}C_\hbar)^{-1}C_\hbar T_{13}\right)&=\operatorname{tr}\left((I+\frac{1}{n}C_\hbar)^{-1}C_\hbar\int_{\mathbb{R}^{2d}}Q^{13}(y,\eta)\,dyd\eta\right)\\
&=\int_{\mathbb{R}^{2d}}\operatorname{tr}\left((I+\frac{1}{n}C_\hbar)^{-1}C_\hbar Q^{13}(y,\eta)\right)\,dyd\eta.
\end{aligned}$$

Note that $(I+\frac{1}{n}C_\hbar)^{-1}C_\hbar$ is increasing and $\lim_{n\to\infty}(I+\frac{1}{n}C_\hbar)^{-1}C_\hbar=C_\hbar$, by monotone convergence [20, Proposition A.3], we have

$$\lim_{n\to\infty}\operatorname{tr}\left((I+\frac{1}{n}C_\hbar)^{-1}C_\hbar T_{13}\right)=\lim_{n\to\infty}\operatorname{tr}\left(T_{13}^{\frac12}(I+\frac{1}{n}C_\hbar)^{-1}C_\hbar T_{13}^{\frac12}\right)=\operatorname{tr}\left(T_{13}^{\frac12}C_\hbar T_{13}^{\frac12}\right),$$

and

$$\lim_{n\to\infty} \operatorname{tr}\left((I+\frac{1}{n}C_\hbar)^{-1}C_\hbar Q^{13}(y,\eta)\right) = \operatorname{tr}\left(Q^{13}(y,\eta)^{\frac{1}{2}}C_\hbar Q^{13}(y,\eta)^{\frac{1}{2}}\right).$$

By monotone convergence, we obtain

$$\begin{aligned}
\operatorname{tr}\left(T_{13}^{\frac{1}{2}}C_\hbar T_{13}^{\frac{1}{2}}\right) &= \lim_{n\to\infty}\int_{\mathbb{R}^{2d}} \operatorname{tr}\left((I+\frac{1}{n}C_\hbar)^{-1}C_\hbar Q^{13}(y,\eta)\right) dyd\eta \\
&= \int_{\mathbb{R}^{2d}} \lim_{n\to\infty} \operatorname{tr}\left((I+\frac{1}{n}C_\hbar)^{-1}C_\hbar Q^{13}(y,\eta)\right) dyd\eta \\
&= \int_{\mathbb{R}^{2d}} \operatorname{tr}\left(Q^{13}(y,\eta)^{\frac{1}{2}}C_\hbar Q^{13}(y,\eta)^{\frac{1}{2}}\right) dyd\eta.
\end{aligned}$$

Since $T_{13}\in\mathcal{C}(\rho_1,\rho_3)$, we obtain

$$\begin{aligned}
\mathfrak{d}(\rho_1,\rho_3)^2 \le& \operatorname{tr}\left(T_{13}^{\frac{1}{2}}C_\hbar T_{13}^{\frac{1}{2}}\right) \\
=& \int_{\mathbb{R}^{2d}} \operatorname{tr}\left(Q^{13}(y,\eta)^{\frac{1}{2}}C_\hbar Q^{13}(y,\eta)^{\frac{1}{2}}\right) dyd\eta \\
\le& (1+\alpha)\int_{\mathbb{R}^{2d}} \operatorname{tr}\left(Q^{1}(y,\eta)^{\frac{1}{2}}c_\hbar(y,\eta) Q^{1}(y,\eta)^{\frac{1}{2}}\right) dyd\eta \\
&+ (1+\frac{1}{\alpha})\int_{\mathbb{R}^{2d}} \operatorname{tr}\left(Q^{3}(y,\eta)^{\frac{1}{2}}c_\hbar(y,\eta) Q^{3}(y,\eta)^{\frac{1}{2}}\right) dyd\eta.
\end{aligned}$$

Minimizing the last right hand side in $Q^1\in\mathcal{C}(\rho_1,\rho_2)$ and in $Q^3\in\mathcal{C}(\rho_2,\rho_3)$ shows that

$$\mathfrak{d}(\rho_1,\rho_3)^2 \le (1+\alpha)\mathfrak{d}(\rho_1,\rho_2)^2 + (1+\frac{1}{\alpha})\mathfrak{d}(\rho_2,\rho_3)^2.$$

Let $\alpha = \frac{\mathfrak{d}(\rho_2,\rho_3)}{\mathfrak{d}(\rho_1,\rho_2)}$. Then

$$\begin{aligned}
\mathfrak{d}(\rho_1,\rho_3)^2 \le& (1+\alpha)\mathfrak{d}(\rho_1,\rho_2)^2 + (1+\frac{1}{\alpha})\mathfrak{d}(\rho_2,\rho_3)^2 \\
=& \mathfrak{d}(\rho_1,\rho_2)^2 + \mathfrak{d}(\rho_2,\rho_3)^2 + 2\mathfrak{d}(\rho_1,\rho_2)\mathfrak{d}(\rho_2,\rho_3) \\
=& (\mathfrak{d}(\rho_1,\rho_2) + \mathfrak{d}(\rho_2,\rho_3))^2
\end{aligned}$$

Thus

$$\mathfrak{d}(\rho_1,\rho_3) \le \mathfrak{d}(\rho_1,\rho_2) + \mathfrak{d}(\rho_2,\rho_3).$$

### 4.1.2 Observations and motivations

In Golse and Paul's original proof, they first constructed a "joint transport plan" $Q^{13}$ among $\rho_1, \rho_2$, and $\rho_3$ from two given transport plans $Q^1\in\mathcal{C}(\rho_1,\rho_2)$ and $Q^3\in\mathcal{C}(\rho_2,\rho_3)$. "Tracing out" one of the coordinates of the "joint transport plan" $Q^{13}$, it becomes a transport plan between the remaining two coordinates. Then, $Q^{13}$ "associates" the three distances $\mathfrak{d}(\rho_1,\rho_2), \mathfrak{d}(\rho_2,\rho_3)$ and $\mathfrak{d}(\rho_1,\rho_3)$. Concerning the quantum-to-quantum cost operator $C_\hbar$ and the quantum-to-classical cost operator $c_\hbar$, they proved an operator inequality that is fundamental to the triangle inequality of $\mathfrak{d}$. Since the cost operators are unbounded, the operator inequality only holds for appropriate vectors. To specify this, they applied the spectral decomposition to the transport plans and showed that the vectors which lie in the nonzero eigenspaces of $Q^{13}$ are indeed in the appropriate domain. Finally, by tracing out all the coordinates, they concluded the triangle inequality.

The idea of Golse and Paul's proof greatly inspired the proof presented in this thesis of the triangle inequality problem considered in Chapter 3. We will construct the joint transport plan from two given transport plans in a slightly different manner, and prove the triangle inequality of the cost functions by using a common triangle inequality trick. Deriving the triangle inequality from the previous preparations is a crucial step. Note that the cost functions considered in Chapter 3 are bounded operators, while the density operators on the noncommutative torus can be unbounded. Let $T \in \mathcal{D}(\mathbb{T}^2_\theta)$, and let $C_1 \le C_2$ be positive operators in $L^\infty(\mathbb{T}^2_\theta)$. It turns out that we can prove the triangle inequality if we can show

$$\tau(C_1 T) \le \tau(C_2 T).$$

This is certainly true for bounded $T$, since

$$\tau(C_1 T) = \tau(T^{\frac{1}{2}} C_1 T^{\frac{1}{2}}) \le \tau(T^{\frac{1}{2}} C_2 T^{\frac{1}{2}}) = \tau(C_2 T).$$

Motivated by Golse and Paul's proof, we can apply spectral decomposition to unbounded operators. Here we will need the notion of spectral measures since the eigenvalue decomposition is not applicable for operators in $\mathcal{D}(\mathbb{T}^2_\theta)$. Suppose $T = \int \lambda \, dE$ is the spectral decomposition of $T$ and $\Delta$ is a Borel subset of $\mathbb{C}$. Then $E(\Delta)$ is bounded since it is a projection by definition. So we have the following inequality

$$\tau(C_1 E(\Delta)) \le \tau(C_2 E(\Delta)).$$

If we can prove $\tau(C_1 E), \tau(C_2 E)$ are Borel measures on $\mathbb{C}$ and

$$\tau(C_1 T) = \tau(C_1 \int \lambda \, dE) = \int \lambda \, d\tau(C_1 E),$$
$$\tau(C_2 T) = \tau(C_2 \int \lambda \, dE) = \int \lambda \, d\tau(C_2 E).$$

Then by the inequality above, we may obtain the desired inequality

$$\tau(C_1 T) = \tau(C_1 \int \lambda \, dE) = \int \lambda \, d\tau(C_1 E) \le \int \lambda \, d\tau(C_2 E) = \tau(C_2 \int \lambda \, dE) = \tau(C_2 T).$$

Another way is to investigate the properties of noncommutative $L^2$-spaces. If we have the following assertions

- $T^{\frac{1}{2}} C_1 T^{\frac{1}{2}}, T^{\frac{1}{2}} C_2 T^{\frac{1}{2}} \in L^1(\mathbb{T}^2_\theta)$;
- $\tau(C_1 T) = \tau(T^{\frac{1}{2}} C_1 T^{\frac{1}{2}})$ and $\tau(C_2 T) = \tau(T^{\frac{1}{2}} C_2 T^{\frac{1}{2}})$;
- $T^{\frac{1}{2}} C_1 T^{\frac{1}{2}} \le T^{\frac{1}{2}} C_2 T^{\frac{1}{2}}$,

then it follows that

$$\tau(C_1 T) = \tau(T^{\frac{1}{2}} C_1 T^{\frac{1}{2}}) \le \tau(T^{\frac{1}{2}} C_2 T^{\frac{1}{2}}) = \tau(C_2 T).$$

## 4.2 The inequality for cost functions

In this section, we will prove the inequality for cost functions which is fundamental to the proof of the triangle inequality. By analogy with the proof of Lemma 4.3 [21, Lemma 3.5], we will first prove a variation of "Peter-Paul" inequality for operators (Lemma A.1 in [20]). To be specific, the bounded operators on $L^2(\mathbb{T}^2_\theta)$ are considered. Then the inequality for cost operators are derived from the auxiliary "Peter-Paul" inequality.

**Lemma 4.4** (Peter-Paul inequality)**.** *Let $T, S$ be bounded operators on $L^2(\mathbb{T}^2_\theta)$. Then, for all $\alpha > 0$, one has*

$$T^*S + S^*T \leq \alpha T^*T + \frac{1}{\alpha}S^*S.$$

*Proof.* Indeed, for each $\alpha > 0$, one has

$$\alpha T^*T + \frac{1}{\alpha}S^*S - (T^*S + S^*T) = (\sqrt{\alpha}T - \frac{1}{\sqrt{\alpha}}S)^*(\sqrt{\alpha}T - \frac{1}{\sqrt{\alpha}}S) \geq 0.$$

□

Next, we apply the preceding lemma to get the following inequality about the cost function.

**Lemma 4.5.** *Let $(u_1, v_1)$ and $(u_3, v_3)$ be two pairs of unitary generators of the noncommutative torus $L^\infty(\mathbb{T}^2_\theta)$ and let $\mathbf{1}$ denote the unit. Let $z = (z_1, z_2) \in \mathbb{T}^2$ and then $z_1, z_2$ are the unitary generators of $L^\infty(\mathbb{T}^2)$ (viewing $z_1, z_2$ as functions). For each $\alpha > 0$ and $z \in \mathbb{T}^2$, we have*

$$C(u_1, v_1, u_3, v_3) \leq (1 + \alpha)C(u_1, v_1, z) \otimes \mathbf{1} + (1 + \frac{1}{\alpha})\mathbf{1} \otimes C(z, u_3, v_3).$$

*Proof.* Observe that for any fixed $z = (z_1, z_2) \in \mathbb{T}^2$ and $\mathbf{1} \otimes \mathbf{1} \in B(L^2(\mathbb{T}^2_\theta) \otimes L^2(\mathbb{T}^2_\theta))$

$$z_i \cdot (\mathbf{1} \otimes \mathbf{1}) = z_i\mathbf{1} \otimes \mathbf{1} = \mathbf{1} \otimes z_i\mathbf{1}, \quad i = 1, 2,$$

is also an operator on $L^2(\mathbb{T}^2_\theta) \otimes L^2(\mathbb{T}^2_\theta)$, so we can rewrite $C(u_1, v_1, u_3, v_3)$ as

$$\begin{aligned} C(u_1, v_1, u_3, v_3) =& |u_1 \otimes \mathbf{1} - \mathbf{1} \otimes u_3|^2 + |v_1 \otimes \mathbf{1} - \mathbf{1} \otimes v_3|^2 \\ =& |u_1 \otimes \mathbf{1} - \mathbf{1} \otimes z_1\mathbf{1} + z_1\mathbf{1} \otimes \mathbf{1} - \mathbf{1} \otimes u_3|^2 \\ &+ |v_1 \otimes \mathbf{1} - \mathbf{1} \otimes z_2\mathbf{1} + z_2\mathbf{1} \otimes \mathbf{1} - \mathbf{1} \otimes v_3|^2. \end{aligned}$$

Substitute the absolute value of an operator by definition

$$\begin{aligned} &|u_1 \otimes \mathbf{1} - \mathbf{1} \otimes z_1\mathbf{1} + z_1\mathbf{1} \otimes \mathbf{1} - \mathbf{1} \otimes u_3|^2 \\ =&(u_1 \otimes \mathbf{1} - \mathbf{1} \otimes z_1\mathbf{1} + z_1\mathbf{1} \otimes \mathbf{1} - \mathbf{1} \otimes u_3)^*(u_1 \otimes \mathbf{1} - \mathbf{1} \otimes z_1\mathbf{1} + z_1\mathbf{1} \otimes \mathbf{1} - \mathbf{1} \otimes u_3) \\ =&(u_1 \otimes \mathbf{1} - \mathbf{1} \otimes z_1\mathbf{1})^*(u_1 \otimes \mathbf{1} - \mathbf{1} \otimes z_1\mathbf{1}) + (z_1\mathbf{1} \otimes \mathbf{1} - \mathbf{1} \otimes u_3)^*(z_1\mathbf{1} \otimes \mathbf{1} - \mathbf{1} \otimes u_3) \\ &+ (u_1 \otimes \mathbf{1} - \mathbf{1} \otimes z_1\mathbf{1})^*(z_1\mathbf{1} \otimes \mathbf{1} - \mathbf{1} \otimes u_3) + (z_1\mathbf{1} \otimes \mathbf{1} - \mathbf{1} \otimes u_3)^*(u_1 \otimes \mathbf{1} - \mathbf{1} \otimes z_1\mathbf{1}) \\ =&|u_1 \otimes \mathbf{1} - \mathbf{1} \otimes z_1\mathbf{1}|^2 + |z_1\mathbf{1} \otimes \mathbf{1} - \mathbf{1} \otimes u_3|^2 \\ &+ (u_1 \otimes \mathbf{1} - \mathbf{1} \otimes z_1\mathbf{1})^*(z_1\mathbf{1} \otimes \mathbf{1} - \mathbf{1} \otimes u_3) + (z_1\mathbf{1} \otimes \mathbf{1} - \mathbf{1} \otimes u_3)^*(u_1 \otimes \mathbf{1} - \mathbf{1} \otimes z_1\mathbf{1}) \end{aligned}$$

By letting $T, S$ in Lemma 4.4 be $u_1 \otimes \mathbf{1} - \mathbf{1} \otimes z_1\mathbf{1}$ and $z_1\mathbf{1} \otimes \mathbf{1} - \mathbf{1} \otimes u_3$ respectively, we obtain

$$\begin{aligned} &(u_1 \otimes \mathbf{1} - \mathbf{1} \otimes z_1\mathbf{1})^*(z_1\mathbf{1} \otimes \mathbf{1} - \mathbf{1} \otimes u_3) + (z_1\mathbf{1} \otimes \mathbf{1} - \mathbf{1} \otimes u_3)^*(u_1 \otimes \mathbf{1} - \mathbf{1} \otimes z_1\mathbf{1}) \\ \leq&\alpha|u_1 \otimes \mathbf{1} - \mathbf{1} \otimes z_1\mathbf{1}|^2 + \frac{1}{\alpha}|z_1\mathbf{1} \otimes \mathbf{1} - \mathbf{1} \otimes u_3|^2 \\ =&\alpha|u_1 \otimes \mathbf{1} - z_1\mathbf{1} \otimes \mathbf{1}|^2 + \frac{1}{\alpha}|\mathbf{1} \otimes z_1\mathbf{1} - \mathbf{1} \otimes u_3|^2 \\ =&\alpha|u_1 - z_1\mathbf{1}|^2 \otimes \mathbf{1} + \frac{1}{\alpha}\mathbf{1} \otimes |z_1\mathbf{1} - u_3|^2. \end{aligned}$$

So we have

$$|u_1 \otimes \mathbf{1} - \mathbf{1} \otimes z_1\mathbf{1} + z_1\mathbf{1} \otimes \mathbf{1} - \mathbf{1} \otimes u_3|^2 \leq (1+\alpha)|u_1 - z_1\mathbf{1}|^2 \otimes \mathbf{1} + (1+\frac{1}{\alpha})\mathbf{1} \otimes |z_1\mathbf{1} - u_3|^2.$$

Similarly,

$$|v_1 \otimes \mathbf{1} - \mathbf{1} \otimes z_2\mathbf{1} + z_2\mathbf{1} \otimes \mathbf{1} - \mathbf{1} \otimes v_3|^2 \leq (1+\alpha)|v_1 - z_2\mathbf{1}|^2 \otimes \mathbf{1} + (1+\frac{1}{\alpha})\mathbf{1} \otimes |z_2\mathbf{1} - v_3|^2.$$

Therefore, we obtain the desired inequality

$$\begin{aligned} C(u_1, v_1, u_3, v_3) \leq & (1+\alpha)|u_1 - z_1\mathbf{1}|^2 \otimes \mathbf{1} + (1+\frac{1}{\alpha})\mathbf{1} \otimes |z_1\mathbf{1} - u_3|^2 \\ & + (1+\alpha)|v_1 - z_2\mathbf{1}|^2 \otimes \mathbf{1} + (1+\frac{1}{\alpha})\mathbf{1} \otimes |z_2\mathbf{1} - v_3|^2 \\ = & (1+\alpha)(|u_1 - z_1\mathbf{1}|^2 + |v_1 - z_2\mathbf{1}|^2) \otimes \mathbf{1} \\ & + (1+\frac{1}{\alpha})\mathbf{1} \otimes (|z_1\mathbf{1} - u_3|^2 + |z_2\mathbf{1} - v_3|^2) \\ = & (1+\alpha)C(u_1, v_1, z) \otimes \mathbf{1} + (1+\frac{1}{\alpha})\mathbf{1} \otimes C(z, u_3, v_3). \end{aligned}$$

□

## 4.3 Assertions about spectral measures

In this section, we first prove lemmas concerning how the trace is applied to the spectral decompositions of normal operators. We will also discuss the tensor product of spectral measures. Then the trace and "partial traces" of the operators on the tensor product space will also be investigated. In the following discussion, we assume implicitly the Hilbert spaces are from the GNS constructions of the von Neumann algebras.

**Lemma 4.6.** *Let $(\mathcal{M}, \tau)$ be a noncommutative measure space. Let $x$ be a normal element in $L^1(\mathcal{M})$ and let $E$ be the spectral measure for $x$. Then*

(i) *The map $\tau(E) : \Delta \mapsto \tau(E(\Delta))$ is a measure on the Borel subsets of $\sigma(x)$;*

(ii) *If $\tau$ is finite and $C$ is a positive element in $\mathcal{M}$, then $\tau(CE) : \Delta \mapsto \tau(CE(\Delta))$ is a measure on the Borel subsets of $\sigma(x)$.*

*Proof.* For assertion (i), note that if $\Delta$ is a Borel subset of $\sigma(x)$, then $E(\Delta) \in \mathcal{M}_+$ by Proposition 2.1.4.(iv) in [16] and thus $\tau$ and $E$ can be composed.

$\tau(E)$ is non-negative since $E(\Delta)$ is a projection and $\tau : \mathcal{M}_+ \to [0, \infty]$. For empty set $\varnothing$, $\tau(E(\varnothing)) = \tau(0) = 0$. Let $\{\Delta_n\}_{n\geq 1}$ be a family of mutually disjoint sets from the Borel subsets of $\sigma(x)$. The countable additivity follows from

$$\tau(E(\bigcup_{n=1}^{\infty} \Delta_n)) = \tau(\sum_{n=1}^{\infty} E(\Delta_n)) = \sum_{n=1}^{\infty} \tau(E(\Delta_n)).$$

The last equality holds since $\tau$ is normal. Therefore, $\tau(E)$ is a measure on the Borel subsets of $\sigma(x)$.

For assertion (ii), $\tau$ extends to $\mathcal{M}$ since it is finite, so $\tau(CE)$ is properly defined.

$\tau(CE)$ is non-negative since $C^{1/2}E(\Delta)C^{1/2} \geq 0$ and $\tau(CE(\Delta)) = \tau(C^{1/2}E(\Delta)C^{1/2}) \geq 0$. $\tau(CE(\varnothing)) = \tau(0) = 0$. For the countable additivity, observe that

$$\tau(CE(\bigcup_{n=1}^{\infty} \Delta_n)) = \tau(C^{\frac{1}{2}}E(\bigcup_{n=1}^{\infty} \Delta_n)C^{\frac{1}{2}}) = \tau(C^{\frac{1}{2}} \sum_{n=1}^{\infty} E(\Delta_n)C^{\frac{1}{2}}).$$

Also note that

$$C^{\frac{1}{2}} \sum_{n=1}^{\infty} E(\Delta_n)C^{\frac{1}{2}} = \sup_{n\in\mathbb{N}} C^{\frac{1}{2}} \sum_{k=1}^{n} E(\Delta_k)C^{\frac{1}{2}} = \sup_{n\in\mathbb{N}} C^{\frac{1}{2}} E(\sum_{k=1}^{n} \Delta_k)C^{\frac{1}{2}}.$$

By the normality of $\tau$, we obtain

$$\begin{aligned}\tau(CE(\bigcup_{n=1}^{\infty} \Delta_n)) &= \sup_{n\in\mathbb{N}} \tau(C^{\frac{1}{2}}E(\sum_{k=1}^{n} \Delta_k)C^{\frac{1}{2}}) \\ &= \sup_{n\in\mathbb{N}} \tau(CE(\sum_{k=1}^{n} \Delta_k)) \\ &= \sup_{n\in\mathbb{N}} \sum_{k=1}^{n} \tau(CE(\Delta_k)) \\ &= \sum_{n=1}^{\infty} \tau(CE(\Delta_n)).\end{aligned}$$

Therefore, $\tau(CE)$ is a measure on the Borel subsets of $\sigma(x)$. □

**Lemma 4.7.** *Let $(\mathcal{M}, \tau)$ be a noncommutative measure space with finite $\tau$. Let $A$ be a positive operator in $L^1(\mathcal{M})$ with spectral decomposition $A = \int \lambda\, dE(\lambda)$. If $C$ is a positive element in $\mathcal{M}$, then*

$$\tau(CA) = \tau(C \int \lambda\, dE) = \int \lambda\, d\tau(CE).$$

*Proof.* We first prove this holds for bounded $A$. Let $\varepsilon > 0$ and $\{\Delta_1, ..., \Delta_n\}$ be a Borel partition of $\sigma(A)$ such that $|x - x'| \leq \varepsilon$ with $x, x' \in \Delta_k, 1 \leq k \leq n$. Then Proposition 2.7 claims that

$$\left\| A - \sum_{k=1}^{n} x_k E(\Delta_k) \right\| \leq \varepsilon, \quad x_k \in \Delta_k.$$

Apply $\tau$ to $CA - C\sum_{k=1}^{n} x_k E(\Delta_k)$

$$\left| \tau(CA) - \sum_{k=1}^{n} x_k \tau(CE(\Delta_k)) \right| = \left| \tau(CA - C\sum_{k=1}^{n} x_k E(\Delta_k)) \right| \leq \|\tau\| \cdot \|C\| \cdot \varepsilon.$$

Note that $\tau(CE)$ is a measure on the Borel subsets of $\sigma(A)$ by Lemma 4.6. It follows from the definition of Lebesgue integral that

$$\tau(CA) = \int \lambda\, d\tau(CE(\lambda)).$$

Then we consider an arbitrary positive operator $A$ in $L^1(\mathcal{M})$. Denote

$$\Delta_n = \{x \in \mathbb{C} : 0 \leq |x| < n\}, \quad E_n(\Delta) = \begin{cases} E(\Delta \cap \Delta_n) + E(\mathbb{C} \setminus \Delta_n), & 0 \in \Delta, \\ E(\Delta \cap \Delta_n), & 0 \notin \Delta, \end{cases} \quad n = 1, 2, 3, \cdots$$

Then $\int \lambda \, dE_n$ is bounded and $A = \int \lambda \, dE = \sup_{n \in \mathbb{N}} \int \lambda \, dE_n$. By the arguments about bounded operators, we obtain

$$\tau(C \int \lambda \, dE_n) = \int \lambda \, d\tau(CE_n)$$

By Proposition 3.4.1 in [16], we have $CA, C^{\frac{1}{2}}AC^{\frac{1}{2}} \in L^1(\mathcal{M})$. Together with Proposition 3.4.2 in [16], we have

$$\tau(CA) = \tau(C^{\frac{1}{2}}AC^{\frac{1}{2}}) < \infty.$$

Also note that

$$C^{\frac{1}{2}}AC^{\frac{1}{2}} \geq 0, \quad C^{\frac{1}{2}} \int \lambda \, dE_n C^{\frac{1}{2}} \geq 0$$

and $\sup_{n \in \mathbb{N}} C^{\frac{1}{2}} \int \lambda \, dE_n C^{\frac{1}{2}} = C^{\frac{1}{2}}AC^{\frac{1}{2}}$ (Proposition 2.2.25 in [16]). So, by Proposition 3.3.3 in [16], we have

$$\begin{aligned} \tau(C^{\frac{1}{2}}AC^{\frac{1}{2}}) &= \sup_{n \in \mathbb{N}} \tau(C^{\frac{1}{2}} \int \lambda \, dE_n C^{\frac{1}{2}}) = \sup_{n \in \mathbb{N}} \tau(C \int \lambda \, dE_n) \\ &= \sup_{n \in \mathbb{N}} \int \lambda \, d\tau(CE_n) = \lim_{n \to \infty} \int \lambda \, d\tau(CE_n) = \lim_{n \to \infty} \int \lambda \cdot \chi_{|\lambda| < n} d\tau(CE), \end{aligned}$$

where $\chi$ is the characteristic function. Note that $\lambda$ is the pointwise limit of $\lambda \cdot \chi_{|\lambda|<n}$, by Fatou's lemma

$$\int \lambda \, d\tau(CE) \leq \lim_{n \to \infty} \int \lambda \cdot \chi_{|\lambda|<n} \, d\tau(CE) = \tau(CA)$$

Since $0 < \lambda \cdot \chi_{|\lambda|<n} \leq \lambda$, we also have

$$\tau(CA) = \lim_{n \to \infty} \int \lambda \cdot \chi_{|\lambda|<n} \, d\tau(CE) \leq \int \lambda \, d\tau(CE).$$

Thus,

$$\tau(CA) = \tau(C \int \lambda \, dE) = \int \lambda \, d\tau(CE).$$

□

**Corollary 4.8.** *Let $(\mathcal{M}, \tau)$ be a noncommutative measure space with finite $\tau$. Let $A$ be a normal operator in $L^1(\mathcal{M})$ with spectral decomposition $A = \int \lambda \, dE(\lambda)$. Then*

$$\tau(A) = \int \lambda \, d\tau(E).$$

*Proof.* Immediate from Lemma 4.7 by letting $C$ be the unit element of $\mathcal{M}$. □

Given two density operators $A$ and $B$, suppose the $E_A$ and $E_B$ are the corresponding spectral measures. We shall consider the "joint" density operator $A \otimes B$ by taking the tensor product. Moreover, $A \otimes B$ can be viewed as the integral on the "joint" spectrum $\sigma(A) \times \sigma(B)$ with respect to the "joint" spectral measure $E_A \otimes E_B$ [35, Theorem 8.2]. According to Theorem 2′ in [4], the tensor product of a finite number of spectral measures is a spectral measure, and so does $E_A \otimes E_B$. Its proof is only briefly stated since it is similar to the proof of Theorem 1 in [4]. Hereby the result about the tensor product of two spectral measures and its proof are formulated below:

**Theorem 4.9** (Tensor product of spectral measures). *Let $H_i$ be separable Hilbert spaces and let $E_i$ be spectral measures for $(X_i, \Omega_i, H_i)$, $i = 1, 2$. Put $X = X_1 \times X_2$, $\Omega$ the minimal $\sigma$-algebra generated by $\Omega_1 \times \Omega_2$ and $H = H_1 \otimes H_2$. Then there exists a spectral measure $E$ for $(X, \Omega, H)$ such that $E|_{\Omega_1 \times \Omega_2} = E_1 \otimes E_2$.*

*Proof.* First we define

$$E(\Delta) \coloneqq E_1(\delta_1) \otimes E_2(\delta_2), \quad \Delta = \delta_1 \times \delta_2 \in \Omega_1 \times \Omega_2.$$

By the property of tensor product, it is easy to check that

- $E(\Delta)$ is a projection;
- $E(\varnothing) = 0, E(X) = \mathrm{id}_X$;
- $E(\Delta_1 \cap \Delta_2) = E(\Delta_1)E(\Delta_2)$.

For the countable additivity, let $\{\Delta_n\}_{n=1}^{\infty}$ be pairwise disjoint sets from $\Omega_1 \times \Omega_2$ such that $\Delta = \bigcup_{n=1}^{\infty} \Delta_n \in \Omega_1 \times \Omega_2$. We need to show that

$$E(\Delta)\xi = \sum_{n=1}^{\infty} E(\Delta_n)\xi, \quad \text{for all } \xi \in H. \tag{4.3}$$

Note that for $N \geq 1$

$$(E(\Delta) - \sum_{n=1}^{N} E(\Delta_n))^* = E(\Delta)^* - \sum_{n=1}^{N} E(\Delta_n)^* = E(\Delta) - \sum_{n=1}^{N} E(\Delta_n)$$

and

$$\begin{aligned}
(E(\Delta) - \sum_{n=1}^{N} E(\Delta_n))^2 =& E(\Delta)^2 - E(\Delta)\sum_{n=1}^{N} E(\Delta_n) - \sum_{n=1}^{N} E(\Delta_n)E(\Delta) + (\sum_{n=1}^{N} E(\Delta_n))^2 \\
=& E(\Delta) - \sum_{n=1}^{N} E(\Delta \cap \Delta_n) - \sum_{n=1}^{N} E(\Delta_n \cap \Delta) + \sum_{1 \leq m,n \leq N} E(\Delta_m)E(\Delta_n) \\
=& E(\Delta) - 2\sum_{n=1}^{N} E(\Delta_n) + \sum_{1 \leq m,n \leq N} E(\Delta_m \cap \Delta_n) \\
=& E(\Delta) - 2\sum_{n=1}^{N} E(\Delta_n) + \sum_{n=1}^{N} E(\Delta_n) \\
=& E(\Delta) - \sum_{n=1}^{N} E(\Delta_n).
\end{aligned}$$

This implies that $E(\Delta) - \sum_{n=1}^N E(\Delta_n)$ is an orthogonal projection. So

$$\begin{aligned}\|E(\Delta)\xi - \sum_n^N E(\Delta_n)\xi\|^2 =& \|(E(\Delta) - \sum_{n=1}^N E(\Delta_n))\xi\|^2 \\ =& \langle (E(\Delta) - \sum_{n=1}^N E(\Delta_n))\xi, (E(\Delta) - \sum_{n=1}^N E(\Delta_n))\xi\rangle \\ =& \langle (E(\Delta) - \sum_{n=1}^N E(\Delta_n))\xi, \xi\rangle.\end{aligned}$$

So (4.3) is equivalent to

$$\langle E(\Delta)\xi, \xi\rangle = \sum_{n=1}^\infty \langle E(\Delta_n)\xi, \xi\rangle, \quad \text{for all } \xi \in H.$$

In fact, it is sufficient to check on "monomials" $\xi = \xi_1 \otimes \xi_2 \in H$ by linearity and density. Observe that for $\delta_i \in \Omega_i$, we have

$$\begin{aligned}\langle E(\delta_1 \times \delta_2)\xi, \xi\rangle =& \langle E_1(\delta_1) \otimes E_2(\delta_2)(\xi_1 \otimes \xi_2), \xi_1 \otimes \xi_2\rangle \\ =& \langle E_1(\delta_1)\xi_1, \xi_1\rangle \langle E_2(\delta_2)\xi_2, \xi_2\rangle.\end{aligned}$$

This is the product of 2 complex measures by Lemma IX.1.9 in [13], and thus is countably additive.

The countable additivity-preserving extension from $\Omega_1 \times \Omega_2$ to $\Omega$ exists and is unique by Theorem 5.2.6 in [3]. □

*Remark.* The Hilbert space $L^2(\mathbb{T}^2_\theta)$ is separable since the linear span of $\{u^k v^\ell\}_{k,\ell \in \mathbb{Z}}$ has dense linear span in it.

Let $N_1$ and $N_2$ be two positive elements in $L^1(\mathbb{T}^2_\theta)$. They admit spectral decompositions

$$N_1 = \int \lambda \, dE_{N_1} \quad \text{and} \quad N_2 = \int \mu \, dE_{N_2}.$$

Let $E$ denote the tensor product of spectral measures $E_{N_1}$ and $E_{N_2}$. In the remaining part of this section, we will consider the operator

$$\int \lambda\mu \, dE = \int \lambda\mu \, dE_{N_1} \otimes E_{N_2}.$$

By Theorem 2.1, $\tau \otimes \tau$ is the normal finite faithful trace on $L^\infty(\mathbb{T}^2_\theta)\bar{\otimes}L^\infty(\mathbb{T}^2_\theta)$ satisfying $(\tau \otimes \tau)(x \otimes y) = \tau(x)\tau(y)$ for $x, y \in L^\infty(\mathbb{T}^2_\theta)$. In the next lemma, we will investigate how $\tau \otimes \tau$ is applied to $\int \lambda\mu \, dE_{N_1} \otimes E_{N_2}$.

**Lemma 4.10.** *Let $N_1, N_2, E_{N_1}, E_{N_2}$ and $E$ be as stated above. Let $C$ be a positive operator in $L^\infty(\mathbb{T}^2_\theta)\bar{\otimes}L^\infty(\mathbb{T}^2_\theta)$. Then*

$$(\tau \otimes \tau)(C \int \lambda\mu \, dE) = (\int \lambda\mu \, d(\tau \otimes \tau)(CE))$$

*Proof.* First we check that $(\tau \otimes \tau)(CE)$ is properly defined. Let $\{\delta_{in}\}_{n=1}^{\infty}, i = 1, 2$ be Borel subsets of $\mathbb{C}$. Note that $E_{N_1}(\delta_{1n})$ and $E_{N_2}(\delta_{2n})$ are in $L^\infty(\mathbb{T}_\theta^2)$, so

$$E(\delta_{1k} \times \delta_{2k}) = E_{N_1}(\delta_{1k}) \otimes E_{N_2}(\delta_{2k}) \in L^\infty(\mathbb{T}_\theta^2) \otimes L^\infty(\mathbb{T}_\theta^2),$$

and thus $\sum_{k=1}^{n} E(\delta_{1k} \times \delta_{2k})$ belongs to $L^\infty(\mathbb{T}_\theta^2) \otimes L^\infty(\mathbb{T}_\theta^2)$. So the range projection $P_n$ of $\sum_{k=1}^{n} E(\delta_{1k} \times \delta_{2k})$ is contained in $L^\infty(\mathbb{T}_\theta^2) \bar{\otimes} L^\infty(\mathbb{T}_\theta^2)$ [31, Theorem 4.1.9]. In fact,

$$P_n = E(\bigcup_{k=1}^{n} (\delta_{1k} \times \delta_{2k})).$$

Observe that $\{P_n\}_{n=1}^{\infty}$ is increasing and

$$\sup_{n \in \mathbb{N}} P_n = E(\bigcup_{n=1}^{\infty} (\delta_{1k} \times \delta_{2k})).$$

Since $L^\infty(\mathbb{T}_\theta^2) \bar{\otimes} L^\infty(\mathbb{T}_\theta^2)$ is SOT closed, we obtain $E(\bigcup_{n=1}^{\infty} (\delta_{1k} \times \delta_{2k})) \in L^\infty(\mathbb{T}_\theta^2) \bar{\otimes} L^\infty(\mathbb{T}_\theta^2)$. Furthermore, $E$ of the complement of $\bigcup_{k=1}^{n} (\delta_{1k} \times \delta_{2k})$ is

$$E(\mathbb{C}^2 \setminus \bigcup_{k=1}^{n} (\delta_{1k} \times \delta_{2k})) = \mathbf{1} - E(\bigcup_{k=1}^{n} (\delta_{1k} \times \delta_{2k})) \in L^\infty(\mathbb{T}_\theta^2) \bar{\otimes} L^\infty(\mathbb{T}_\theta^2).$$

Then note that $E$ is defined on the $\sigma$-algebra generated by the product of two Borel algebras of subsets of $\mathbb{C}$ which is just the Borel algebra of the subsets of $\mathbb{C}^2$, so for any Borel subset $\Delta$ of $\mathbb{C}^2$, we have $E(\Delta) \in L^\infty(\mathbb{T}_\theta^2) \bar{\otimes} L^\infty(\mathbb{T}_\theta^2)$. This shows that $(\tau \otimes \tau)(CE)$ is properly defined. Repeat the arguments in Lemma 4.6, we obtain $(\tau \otimes \tau)(CE)$ is a measure on Borel subsets of $\mathbb{C}^2$.

Similar to the proof of Lemma 4.7, we first prove this assertion for $E$ to be compactly supported. Let $\varepsilon > 0$ and $\{\Delta_1, ..., \Delta_n\}$ be a Borel partition of $\sigma(N_1) \times \sigma(N_2)$ such that $|\lambda\mu - \lambda'\mu'| \le \varepsilon$ with $(\lambda, \mu), (\lambda', \mu') \in \Delta_k, 1 \le k \le n$. This partition exists since $\sigma(N_1) \times \sigma(N_2)$ is compact. Then Proposition 2.7 claims that

$$\left\| \int \lambda\mu \, dE - \sum_{k=1}^{n} \lambda_k \mu_k E(\Delta_k) \right\| \le \varepsilon, \quad (\lambda_k, \mu_k) \in \Delta_k.$$

Apply $\tau \otimes \tau$ to $C \int \lambda\mu \, dE - C \sum_{k=1}^{n} \lambda_k \mu_k E(\Delta_k)$,

$$\begin{aligned}
&\left| (\tau \otimes \tau) \left( C \int \lambda\mu \, dE \right) - \sum_{k=1}^{n} \lambda_k \mu_k (\tau \otimes \tau)(CE(\Delta_k)) \right| \\
=&\left| (\tau \otimes \tau) \left( C \int \lambda\mu \, dE - C \sum_{k=1}^{n} \lambda_k \mu_k E(\Delta_k) \right) \right| \\
\le&\|\tau \otimes \tau\| \|C\| \varepsilon.
\end{aligned}$$

Note that $(\tau \otimes \tau)(CE)$ is a measure on the Borel subsets of $\mathbb{C}^2$. It follows from the definition of Lebesgue integral that

$$(\tau \otimes \tau) \left( C \int \lambda\mu \, dE \right) = \int \lambda\mu \, d(\tau \otimes \tau)(CE(\lambda, \mu)). \tag{4.4}$$

Next we consider $N_1$ and $N_2$ to be unbounded. Denote

$$\Delta_n = \{(\lambda,\mu) \in \mathbb{C}^2 : 0 \le |\lambda|, |\mu| < n\},$$
$$E_n(\Delta) = \begin{cases} E(\Delta \cap \Delta_n) + E(\mathbb{C}^2 \setminus \Delta_n), & (0,0) \in \Delta, \\ E(\Delta \cap \Delta_n), & (0,0) \notin \Delta, \end{cases} \quad n = 1,2,3,\cdots$$

Then $E_n$ is compactly supported and $\int \lambda\mu \, dE = \sup_{n\in\mathbb{N}} \int \lambda\mu \, dE_n$. By (4.4) for compactly supported spectral measure, we obtain

$$(\tau \otimes \tau)(C \int \lambda\mu \, dE_n) = \int \lambda\mu \, d(\tau \otimes \tau)(CE_n).$$

We claim that $\int \lambda\mu \, dE \in L^1(\mathbb{T}^2_\theta \bar{\otimes} \mathbb{T}^2_\theta)$. Let $\psi : \mathbb{C}^2 \to \mathbb{C}$ be the map given by $(\lambda,\mu) \mapsto \lambda\mu$. By Theorem 1.5.8 in [16], we can rewrite the integral as

$$\int \lambda\mu \, dE(\lambda,\mu) = \int \nu \, d\psi(E)(\nu) := \int \nu \, dE(\psi^{-1}(\nu)),$$

which is the usual spectral decomposition of $\int \lambda\mu \, dE$. For arbitrary Borel subset $\delta$ of $\mathbb{C}$, the set $\Delta = \{(\lambda,\mu) \in \mathbb{C}^2 : \lambda\mu \in \delta\}$ is a Borel subset of $\mathbb{C}^2$, and thus $E(\Delta) \in L^\infty(\mathbb{T}^2_\theta) \bar{\otimes} L^\infty(\mathbb{T}^2_\theta)$ by previous arguments. It follows from Proposition 2.1.4 in [16] that $\int \lambda\mu \, dE$ is affiliated with $L^\infty(\mathbb{T}^2_\theta) \bar{\otimes} L^\infty(\mathbb{T}^2_\theta)$, and thus is $\tau \otimes \tau$-measurable. Then by Proposition 3.3.3 in [16], we have

$$\begin{aligned}
(\tau \otimes \tau)(\int \lambda\mu \, dE) &= \sup_{n\in\mathbb{N}} (\tau \otimes \tau)(\int \lambda\mu \, dE_n) \\
&= \sup_{n\in\mathbb{N}} \int_{|\lambda|,|\mu|<n} \lambda\mu \, d(\tau \otimes \tau)(E_n) \\
&= \sup_{n\in\mathbb{N}} \int_{|\lambda|,|\mu|<n} \lambda\mu \, d\tau(E_{N_1}) \, d\tau(E_{N_2}) \\
&= \sup_{n\in\mathbb{N}} \int_{|\lambda|<n} \lambda \, d\tau(E_{N_1}) \int_{|\mu|<n} \mu \, d\tau(E_{N_2}) \\
&= \int \lambda \, d\tau(E_{N_1}) \int \mu \, d\tau(E_{N_2}) \\
&= \tau(N_1)\tau(N_2) < \infty,
\end{aligned}$$

where the last equality is assured by Corollary 4.8. So $\int \lambda\mu \, dE$ is in $L^1(\mathbb{T}^2_\theta \bar{\otimes} \mathbb{T}^2_\theta)$.

Then repeat the arguments in Lemma 4.7, we will obtain

$$(\tau \otimes \tau)(C \int \lambda\mu \, dE) = \int \lambda\mu \, d(\tau \otimes \tau)(CE).$$

□

Next we consider the "partial traces" $\mathrm{id} \otimes \tau$ and $\tau \otimes \mathrm{id}$ of $\int \lambda\mu \, dE$, where id is the identity map on $L^\infty(\mathbb{T}^2_\theta)$. Recall that we have shown $\tau$ is WOT continuous in Proposition 2.6. Since WOT is weaker that $\sigma$-weak topology, it follows that $\tau$ is $\sigma$-weakly continuous. Moreover, $\tau$ is completely positive. Then by Proposition IV.5.13 in [36], the $\sigma$-weakly continuous and completely positive extension of $\mathrm{id} \otimes \tau$ on $L^\infty(\mathbb{T}^2_\theta) \bar{\otimes} L^\infty(\mathbb{T}^2_\theta)$ into $L^\infty(\mathbb{T}^2_\theta)$ exists, and we still denote it by $\mathrm{id} \otimes \tau$. Furthermore, it can also be extended to a positive and bounded map on $L^1(\mathbb{T}^2_\theta \bar{\otimes} \mathbb{T}^2_\theta)$ into $L^1(\mathbb{T}^2_\theta)$ by Lemma 5.3 in [22].

**Lemma 4.11.** *Let $N_1, N_2$ be two positive elements in $L^1(\mathbb{T}^2_\theta)$ with spectral decompositions of the form*

$$N_1 = \int \lambda \, dE_{N_1}, \quad N_2 = \int \mu \, dE_{N_2},$$

*where $E_{N_1}, E_{N_2}$ are spectral measures for $N_1, N_2$ respectively. Let $E$ be the tensor product spectral measure $E_{N_1} \otimes E_{N_2}$. Then*

$$(\mathrm{id} \otimes \tau)(\int \lambda\mu \, dE) = \iint \lambda\mu \, d\tau(E_{N_2}) \, dE_{N_1},$$
$$(\tau \otimes \mathrm{id})(\int \lambda\mu \, dE) = \iint \lambda\mu \, d\tau(E_{N_1}) \, dE_{N_2}.$$

*Proof.* We first prove this assertion for compactly supported $E_{N_1}$ and $E_{N_2}$, that is, $N_1$ and $N_2$ are bounded operators. For arbitrary $\varepsilon > 0$, let $\{\delta_{ik}\}_{k=1}^n, i = 1, 2$ be Borel partitions of $\sigma(N_i)$ and then $\{\Delta_{k\ell} | \Delta_{k\ell} = \delta_{1k} \times \delta_{2\ell}, 1 \leq k, \ell \leq n\}$ is a Borel partition of $\sigma(N_1) \times \sigma(N_2)$ such that $|\lambda\mu - \lambda'\mu'| < \varepsilon$, with $(\lambda, \mu), (\lambda', \mu') \in \Delta_{k\ell}, 1 \leq k, \ell \leq n$. This partition exists since $\lambda\mu$ is uniformly continuous on a compact subset of $\mathbb{C}^2$. Then by Proposition 2.7, we have

$$\left\| \int \lambda\mu \, dE - \sum_{k,\ell=1}^{n} \lambda_k \mu_\ell E(\Delta_{k\ell}) \right\| \leq \varepsilon, \quad \text{for } \lambda_k \in \delta_{1k} \text{ and } \mu_\ell \in \delta_{2\ell}.$$

Note that $\mathrm{id} \otimes \tau$ is $\sigma$-weakly continuous and $\sigma$-weak topology is weaker that norm topology, it follows that $\mathrm{id} \otimes \tau$ is bounded. So

$$\begin{aligned} &\left\| (\mathrm{id} \otimes \tau) \left( \int \lambda\mu \, dE \right) - (\mathrm{id} \otimes \tau) \left( \sum_{k,\ell=1}^{n} \lambda_k \mu_\ell E(\Delta_{k\ell}) \right) \right\| \\ =& \left\| (\mathrm{id} \otimes \tau) \left( \int \lambda\mu \, dE - \sum_{k,\ell=1}^{n} \lambda_k \mu_\ell E(\Delta_{k\ell}) \right) \right\| \\ \leq& \|\mathrm{id} \otimes \tau\| \varepsilon. \end{aligned} \tag{4.5}$$

Observe that

$$\begin{aligned} (\mathrm{id} \otimes \tau) \left( \sum_{k,\ell=1}^{n} \lambda_k \mu_\ell E(\Delta_{k\ell}) \right) =& (\mathrm{id} \otimes \tau) \left( \sum_{k,\ell=1}^{n} \lambda_k \mu_\ell E_{N_1}(\delta_{1k}) \otimes E_{N_2}(\delta_{2\ell}) \right) \\ =& \sum_{k,\ell=1}^{n} \lambda_k \mu_\ell E_{N_1}(\delta_{1k}) \otimes \tau(E_{N_2}(\delta_{2\ell})) \\ =& \sum_{k,\ell=1}^{n} \lambda_k \mu_\ell \tau(E_{N_2}(\delta_{2\ell})) E_{N_1}(\delta_{1k}). \end{aligned}$$

By the assumption of the partition $\Delta_{k\ell}$, we have

$$\left| \sum_{\ell=1}^{n} \lambda_k \mu_\ell \tau(E_{N_2}(\delta_{2\ell})) - \int \lambda_k \mu \, d\tau(E_{N_2}) \right| \leq \max_{\mu \in \delta_{2\ell}, 1 \leq \ell \leq n} |\lambda_k \mu_\ell - \lambda_k \mu| \cdot \tau(\mathbf{1}) < \varepsilon \cdot \tau(\mathbf{1}) = \varepsilon.$$

For arbitrary $\lambda$ in $\delta_{1k}$, we also have

$$\left| \int \lambda_k \mu \, d\tau(E_{N_2}) - \int \lambda\mu \, d\tau(E_{N_2}) \right| \leq \max_{\lambda \in \delta_{1k}} |\lambda_k \mu - \lambda\mu| \cdot \tau(\mathbf{1}) < \varepsilon \cdot \tau(\mathbf{1}) = \varepsilon.$$

Thus,

$$\begin{aligned}
&\left| \sum_{\ell=1}^{n} \lambda_k \mu_\ell \tau(E_{N_2}(\delta_{2\ell})) - \int \lambda\mu \, d\tau(E_{N_2}) \right| \\
=&\left| \sum_{\ell=1}^{n} \lambda_k \mu_\ell \tau(E_{N_2}(\delta_{2\ell})) - \int \lambda_k \mu \, d\tau(E_{N_2}) + \int \lambda_k \mu \, d\tau(E_{N_2}) - \int \lambda\mu \, d\tau(E_{N_2}) \right| \\
\leq&\left| \sum_{\ell=1}^{n} \lambda_k \mu_\ell \tau(E_{N_2}(\delta_{2\ell})) - \int \lambda_k \mu \, d\tau(E_{N_2}) \right| + \left| \int \lambda_k \mu \, d\tau(E_{N_2}) - \int \lambda\mu \, d\tau(E_{N_2}) \right| \\
<&2\varepsilon \cdot \tau(\mathbf{1}) = 2\varepsilon.
\end{aligned}$$

So, by Proposition 2.7, we obtain

$$\left\| \sum_{k,\ell=1}^{n} \lambda_k \mu_\ell \tau(E_{N_2}(\delta_{2\ell})) E_{N_1}(\delta_{1k}) - \iint \lambda\mu \, d\tau(E_{N_2}) \, dE_{N_1} \right\| < 2\varepsilon \cdot \tau(\mathbf{1}) < 2\varepsilon.$$

Therefore, together with (4.5), we obtain

$$(\mathrm{id} \otimes \tau) \left( \int \lambda\mu \, dE \right) = \iint \lambda\mu \, d\tau(E_{N_2}) \, dE_{N_1}.$$

Then we consider $N_1$ and $N_2$ to be unbounded. Denote

$$\begin{aligned}
\Delta_n =& \{x \in X : 0 \leq |\lambda|, |\mu| < n\}, \\
E_n(\Delta) =& \begin{cases} E(\Delta \cap \Delta_n) + E(\mathbb{C}^2 \setminus \Delta_n), & (0,0) \in \Delta, \\ E(\Delta \cap \Delta_n), & (0,0) \notin \Delta, \end{cases} \qquad n = 1, 2, 3, \cdots
\end{aligned}$$

Note that

$$\int \lambda\mu \, dE = \sup_{n \in \mathbb{N}} \int \lambda\mu \, dE_n = \lim_{n \to \infty} \int \lambda\mu \, dE_n.$$

Then

$$\begin{aligned}
\lim_{n\to\infty} \left\| \int \lambda\mu \, dE - \int \lambda\mu \, dE_n \right\|_1 =& \lim_{n\to\infty} (\tau \otimes \tau) \left( \int \lambda\mu \, dE - \int \lambda\mu \, dE_n \right) \\
=&(\tau \otimes \tau)(\int \lambda\mu \, dE) - \lim_{n\to\infty} (\tau \otimes \tau)(\int \lambda\mu \, dE_n) \\
=&(\tau \otimes \tau)(\int \lambda\mu \, dE) - \sup_{n \in \mathbb{N}} (\tau \otimes \tau)(\int \lambda\mu \, dE_n) \\
=&0.
\end{aligned}$$

So Lemma 5.3 in [22] implies that

$$\lim_{n\to\infty} \left\| (\mathrm{id} \otimes \tau)(\int \lambda\mu \, dE - \int \lambda\mu \, dE_n) \right\|_1 = 0.$$

Note that $\mathrm{id} \otimes \tau$ is positive and $\int \lambda\mu \, dE \geq \int \lambda\mu \, dE_n$, hence

$$\begin{aligned}
&\lim_{n\to\infty} \left\| (\mathrm{id} \otimes \tau)(\int \lambda\mu \, dE - \int \lambda\mu \, dE_n) \right\|_1 \\
=& \lim_{n\to\infty} \tau \left( (\mathrm{id} \otimes \tau)(\int \lambda\mu \, dE - \int \lambda\mu \, dE_n) \right) \\
=& \tau \left( (\mathrm{id} \otimes \tau)(\int \lambda\mu \, dE) \right) - \lim_{n\to\infty} \tau \left( (\mathrm{id} \otimes \tau)(\int \lambda\mu \, dE_n) \right).
\end{aligned}$$

For compactly supported $E_n$, we have proved that

$$(\mathrm{id} \otimes \tau)(\iint \lambda\mu \, dE_n) = \iint_{\Delta_n} \lambda\mu \, d\tau(E_{N_2}) \, dE_{N_1}.$$

Also observe that

$$\sup_{n\in\mathbb{N}} \iint_{\Delta_n} \lambda\mu \, d\tau(E_{N_2}) \, dE_{N_1} = \iint \lambda\mu \, d\tau(E_{N_2}) \, dE_{N_1}.$$

Thus,

$$\begin{aligned}
\lim_{n\to\infty} \tau \left( (\mathrm{id} \otimes \tau)(\int \lambda\mu \, dE_n) \right) &= \lim_{n\to\infty} \tau \left( \iint_{\Delta_n} \lambda\mu \, d\tau(E_{N_2}) \, dE_{N_1} \right) \\
&= \sup_{n\in\mathbb{N}} \tau \left( \iint_{\Delta_n} \lambda\mu \, d\tau(E_{N_2}) \, dE_{N_1} \right) \\
&= \tau \left( \iint \lambda\mu \, d\tau(E_{N_2}) \, dE_{N_1} \right).
\end{aligned}$$

So we have

$$\begin{aligned}
0 =& \tau \left( (\mathrm{id} \otimes \tau)(\int \lambda\mu \, dE) \right) - \tau \left( \iint \lambda\mu \, d\tau(E_{N_2}) \, dE_{N_1} \right) \\
=& \tau \left( (\mathrm{id} \otimes \tau)(\int \lambda\mu \, dE) - \iint \lambda\mu \, d\tau(E_{N_2}) \, dE_{N_1} \right).
\end{aligned}$$

Note that

$$(\mathrm{id} \otimes \tau) \left( \int \lambda\mu \, dE \right) \geq \sup_{n\in\mathbb{N}} (\mathrm{id} \otimes \tau) \left( \int \lambda\mu \, dE_n \right) = \iint \lambda\mu \, d\tau(E_{N_2}) \, dE_{N_1}$$

and $\tau$ is faithful by Proposition 3.3.3 in [16]. Therefore,

$$(\mathrm{id} \otimes \tau) \left( \int \lambda\mu \, dE \right) = \int \lambda\mu \, d\tau(E_{N_2}) \, dE_{N_1}.$$

□

## 4.4 Proof of the triangle inequality

In this section, the triangle inequality (3.3) is proved. We will first construct the "joint transport plan" for $\rho_1, \rho_3$ and $g$. Then we will apply the proved inequality for cost functions and assertions for spectral measures to show the triangle inequality (3.3) holds.

Recall the notations in Chapter 3, suppose $\rho_1, \rho_3 \in \mathcal{D}(\mathbb{T}^2_\theta)$ and $\rho_2 = g \in \mathcal{D}(\mathbb{T}^2)$. Let $\tau$ and $\int$ be the normal finite faithful traces on $L^\infty(\mathbb{T}^2_\theta)$ and $L^\infty(\mathbb{T}^2)$ respectively. Let $T_{12} \in \mathcal{C}(\rho_1, g)$ and $T_{23} \in \mathcal{C}(g, \rho_3)$ be arbitrary couplings, then they satisfy the marginal conditions

$$\begin{aligned} (\mathrm{id} \otimes \int) T_{12} = \rho_1, \quad (\tau \otimes \mathrm{id}) T_{12} = g, \\ (\mathrm{id} \otimes \tau) T_{23} = g, \quad (\int \otimes \mathrm{id}) T_{23} = \rho_3. \end{aligned} \tag{4.6}$$

Note that $T_{12}$ and $T_{23}$ belong to $L^1(\mathbb{T}^2_\theta \bar{\otimes} \mathbb{T}^2)$ and $L^1(\mathbb{T}^2 \bar{\otimes} \mathbb{T}^2_\theta)$. In the next lemma, we will see that they can be viewed as operator-valued functions on $\mathbb{T}^2$.

**Lemma 4.12.** *Let $\mathbb{T}^2$ be the 2-torus equipped with the normalized Lebesgue measure and let $(\mathcal{M}, \tau)$ be a noncommutative measure space with finite trace $\tau$. For $0 < p < \infty$,*

$$L^p(\mathcal{M} \bar{\otimes} L^\infty(\mathbb{T}^2)) \cong L^p(\mathbb{T}^2; L^p(\mathcal{M})).$$

*Proof.* Denote by $\hat{\tau}$ the tensor product trace on $L^\infty(\mathbb{T}^2) \bar{\otimes} \mathcal{M}$ such that

$$\hat{\tau}(x \otimes f) = \tau(x) \cdot \int_{\mathbb{T}^2} f(z)\, dz = \int_{\mathbb{T}^2} \tau(f(z)x)\, dz, \quad f \in L^\infty(\mathbb{T}^2) \text{ and } x \in \mathcal{M}.$$

Denote by $L^\infty_{\mathrm{wo}}(\mathbb{T}^2; \mathcal{M})$ the space of all WOT measurable $\mathcal{M}$-valued functions on $\mathbb{T}^2$. Consider a map $\pi : \mathcal{M} \bar{\otimes} L^\infty(\mathbb{T}^2) \to L^\infty_{\mathrm{wo}}(\mathbb{T}^2; \mathcal{M})$ given by

$$\pi : x \otimes f \mapsto f \cdot x, \quad x \in \mathcal{M}, f \in L^\infty(\mathbb{T}^2).$$

The map $\pi$ is a bijective $*$-homomorphism and trace-preserving [16, Remark 3.9.7], that is,

$$\hat{\tau}(T) = \int_{\mathbb{T}^2} \tau(\pi(T)(z))\, dz, \quad T \in \mathcal{M} \bar{\otimes} L^\infty(\mathbb{T}^2).$$

For $T \in \mathcal{M} \bar{\otimes} L^\infty(\mathbb{T}^2)$, Note that

$$\|\pi(T)\|_p^p = \int_{\mathbb{T}^2} \tau(|\pi(T)|^p(z))\, dz = \int_{\mathbb{T}^2} \tau(\pi(|T|^p)(z))\, dz = \hat{\tau}(|T|^p) = \|T\|_p^p,$$

where $|\pi(T)|^p = \pi(|T|^p)$ follows from the fact that $\pi$ is a (contractive) homomorphism and continuous functional calculus. Therefore, we obtain

$$L^p(\mathcal{M} \bar{\otimes} L^\infty(\mathbb{T}^2)) \cong L^p(\mathbb{T}^2; L^p(\mathcal{M})).$$

□

Then we can view the couplings $T_{12}$ and $T_{23}$ as $L^1(\mathbb{T}^2_\theta)$-valued functions on $\mathbb{T}^2$, that is,

$$\begin{aligned} T_{12}(z) &\coloneqq \pi(T_{12})(z) \in L^1(\mathbb{T}^2; L^1(\mathbb{T}^2_\theta)), \\ T_{23}(z) &\coloneqq \pi(T_{23})(z) \in L^1(\mathbb{T}^2; L^1(\mathbb{T}^2_\theta)). \end{aligned}$$

For fixed $z \in \mathbb{T}^2$, observe that $T_{12}(z)$ and $T_{23}(z)$ are positive elements in $L^1(\mathbb{T}^2_\theta)$. Suppose that $T_{12}(z) = A_z$ and $T_{23}(z) = B_z$ and consider the spectral decomposition of $A_z$ and $B_z$

$$\begin{aligned} A_z &= \int \lambda\, dE_{A_z}(\lambda), \\ B_z &= \int \mu\, dE_{B_z}(\mu), \end{aligned}$$

where $E_{A_z}$ and $E_{B_z}$ are spectral measures for $A_z$ and $B_z$ respectively.

We may rewrite (4.6) as

$$\begin{aligned}
\rho_1 &= (\mathrm{id} \otimes \int)(T_{12}) = \int_{\mathbb{T}^2} T_{12}(z)\, dz = \int_{\mathbb{T}^2} A_z\, dz, \\
g(z) &= (\tau \otimes \mathrm{id})(T_{12})(z) = \tau(T_{12}(z)) = \tau(A_z) = \int \lambda\, d\tau(E_{A_z}(\lambda)), \\
g(z) &= (\mathrm{id} \otimes \tau)(T_{23})(z) = \tau(T_{23}(z)) = \tau(B_z) = \int \mu\, d\tau(E_{B_z}(\mu)), \\
\rho_3 &= (\int \otimes \mathrm{id})(T_{23}) = \int_{\mathbb{T}^2} T_{23}(z)\, dz = \int_{\mathbb{T}^2} B_z\, dz,
\end{aligned}$$

where the integrals of operators are Bochner integrals [17, Definition 64.11] and $\tau(A_z) = g(z) = \tau(B_z)$ holds for a.e. $z \in \mathbb{T}^2$ and we will just assume it holds for all $z \in \mathbb{T}^2$.

Now we start to construct the "joint coupling". Let $E_{A_z} \otimes E_{B_z}$ still denote the extension of $E_{A_z} \otimes E_{B_z}$ assured by Theorem 4.9. Then define $T_{123} \in L^1(\mathbb{T}^2_\theta \bar{\otimes} \mathbb{T}^2 \bar{\otimes} \mathbb{T}^2_\theta) \cong L^1(\mathbb{T}^2; L^1(\mathbb{T}^2_\theta \bar{\otimes} \mathbb{T}^2_\theta))$ as

$$T_{123}(z) = g(z)^{-1} \iint \lambda\mu\, dE_{A_z}(\lambda) \otimes E_{B_z}(\mu),$$

where

$$g^{-1}(z) = \begin{cases} 1/g(z), & g(z) \neq 0, \\ 0, & g(z) = 0. \end{cases} \tag{4.7}$$

By Theorem 1.5.3 in [16],

$$\begin{aligned}
T_{123}(z)^* =& \overline{g(z)^{-1}} \int_{\sigma(A_z)\times\sigma(B_z)} \overline{\lambda\mu}\, dE_{A_z}(\lambda) \otimes E_{B_z}(\mu) \\
=& g(z)^{-1} \int_{\sigma(A_z)\times\sigma(B_z)} \lambda\mu\, dE_{A_z}(\lambda) \otimes E_{B_z}(\mu) \\
=& T_{123}(z).
\end{aligned}$$

This implies that $T_{123}$ is self-adjoint. Note that, for any $\xi \in L^2(\mathbb{T}^2_\theta) \otimes L^2(\mathbb{T}^2_\theta)$,

$$\langle T_{123}(z)\xi, \xi\rangle = \int_{\sigma(A_z)\times\sigma(B_z)} \lambda\mu g(z)^{-1}\, d\langle E_{A_z} \otimes E_{B_z}\xi, \xi\rangle \geq 0.$$

So $T_{123}(z)$ is positive. Also note that

$$\begin{aligned}
\|T_{123}\|_1 =& \int_{\mathbb{T}^2} g^{-1}(z)(\tau \otimes \tau)\left(\int \lambda\mu\, dE_{A_z} \otimes E_{B_z}\right) dz \\
=& \int_{\mathbb{T}^2} g^{-1}(z) \iint \lambda\mu\, d(\tau \otimes \tau)(E_{A_z} \otimes E_{B_z})\, dz \\
=& \int_{\mathbb{T}^2} g^{-1}(z) \int \lambda\, d\tau(E_{A_z}) \int \mu\, d\tau(E_{B_z})\, dz \\
=& \int_{\mathbb{T}^2} g^{-1}(z)\tau(A_z)\tau(B_z)\, dz \\
=& \int_{\mathbb{T}^2} g(z)\, dz \\
=& 1 < \infty.
\end{aligned}$$

So $T_{123}$ belongs to $L^1(\mathbb{T}^2_\theta \bar{\otimes} \mathbb{T}^2 \bar{\otimes} \mathbb{T}^2_\theta)$.

Observe that

$$(\mathrm{id} \otimes \tau)(\mathrm{id} \otimes \int \otimes\, \mathrm{id}) = (\mathrm{id} \otimes \int \otimes \tau) = (\mathrm{id} \otimes \int)(\mathrm{id} \otimes \mathrm{id} \otimes \tau).$$

Thus,

$$\begin{aligned}(\mathrm{id} \otimes \tau)(\int_{\mathbb{T}^2} T_{123}(z)\, dz) &= \int_{\mathbb{T}^2} (\mathrm{id} \otimes \tau)(T_{123}(z))\, dz \\ &= \int_{\mathbb{T}^2} (\mathrm{id} \otimes \tau) \left( g^{-1}(z) \int \lambda\mu\, dE_{A_z} \otimes E_{B_z} \right) dz \\ &= \int_{\mathbb{T}^2} g^{-1}(z)(\mathrm{id} \otimes \tau) \left( \int \lambda\mu\, dE_{A_z} \otimes E_{B_z} \right) dz.\end{aligned}$$

By Lemma 4.11, we have

$$\begin{aligned} g^{-1}(z)(\mathrm{id} \otimes \tau) \left( \int \lambda\mu\, dE_{A_z} \otimes E_{B_z} \right) =& g^{-1}(z) \iint \lambda\mu\, d\tau(E_{B_z})\, dE_{A_z} \\ =& g^{-1}(z) \int \lambda \int \mu\, d\tau(E_{B_z})\, dE_{A_z} \\ =& g^{-1}(z) \int \lambda\tau(B_z)\, dE_{A_z} \\ =& g^{-1}(z)\tau(B_z)A_z \\ =& g^{-1}(z)g(z)A_z \\ =& A_z. \end{aligned}$$

The last equality holds since $0 = g(z) = \tau(B_z) = \tau(A_z)$ implies $A_z = B_z = 0$. Therefore,

$$(\mathrm{id} \otimes \tau)(\int_{\mathbb{T}^2} T_{123}(z)\, dz) = \int_{\mathbb{T}^2} A_z\, dz = (\mathrm{id} \otimes \int)(T_{12}) = \rho_1.$$

Similarly,

$$(\tau \otimes \mathrm{id})(\int_{\mathbb{T}^2} T_{123}(z)\, dz) = \rho_3.$$

This indicates $T_{13} \coloneqq \int_{\mathbb{T}^2} T_{123}(z)\, dz$ is a coupling between $\rho_1$ and $\rho_3$. Then

$$\begin{aligned} d(\rho_1, \rho_3)^2 \leq& (\tau \otimes \tau)(C(u_1, v_1, u_3, v_3)T_{13}) \\ =& (\tau \otimes \tau) \left( C(u_1, v_1, u_3, v_3) \int_{\mathbb{T}^2} T_{123}(z)\, dz \right) \\ =& \int_{\mathbb{T}^2} (\tau \otimes \tau) \left( C(u_1, v_1, u_3, v_3)T_{123}(z) \right) dz \\ =& \int_{\mathbb{T}^2} (\tau \otimes \tau) \left( C(u_1, v_1, u_3, v_3)g^{-1}(z) \int \lambda\mu\, dE_{A_z} \otimes E_{B_z} \right) dz \\ =& \int_{\mathbb{T}^2} g^{-1}(z)(\tau \otimes \tau) \left( C(u_1, v_1, u_3, v_3) \int \lambda\mu\, dE_{A_z} \otimes E_{B_z} \right) dz. \end{aligned}$$

By Lemma 4.10, we have

$$(\tau \otimes \tau)\left(C(u_1, v_1, u_3, v_3) \int \lambda\mu \, dE_{A_z} \otimes E_{B_z}\right) = \int \lambda\mu \, d(\tau \otimes \tau)\left(C(u_1, v_1, u_3, v_3)(E_{A_z} \otimes E_{B_z})\right).$$

For any $\alpha > 0$, Lemma 4.5 implies

$$\begin{aligned}
&(\tau \otimes \tau)\left(C(u_1, v_1, u_3, v_3)(E_{A_z} \otimes E_{B_z})\right) \\
=&(\tau \otimes \tau)\left((E_{A_z} \otimes E_{B_z})C(u_1, v_1, u_3, v_3)(E_{A_z} \otimes E_{B_z})\right) \\
\leq&(\tau \otimes \tau)\left((E_{A_z} \otimes E_{B_z})\left((1+\alpha)C(u_1, v_1, z) \otimes \mathbf{1} + (1+\frac{1}{\alpha})\mathbf{1} \otimes C(z, u_3, v_3)\right)(E_{A_z} \otimes E_{B_z})\right) \\
=&(\tau \otimes \tau)\left(\left((1+\alpha)C(u_1, v_1, z) \otimes \mathbf{1} + (1+\frac{1}{\alpha})\mathbf{1} \otimes C(z, u_3, v_3)\right)(E_{A_z} \otimes E_{B_z})\right) \\
=&(1+\alpha)(\tau \otimes \tau)\left((C(u_1, v_1, z) \otimes \mathbf{1})(E_{A_z} \otimes E_{B_z})\right) \\
&+ (1+\frac{1}{\alpha})(\tau \otimes \tau)\left((\mathbf{1} \otimes C(z, u_3, v_3))(E_{A_z} \otimes E_{B_z})\right) \\
=&(1+\alpha)(\tau \otimes \tau)((C(u_1, v_1, z)E_{A_z}) \otimes E_{B_z}) + (1+\frac{1}{\alpha})(\tau \otimes \tau)\left(E_{A_z} \otimes (C(z, u_3, v_3)E_{B_z})\right).
\end{aligned}$$

Thus,

$$\begin{aligned}
&\int \lambda\mu \, d(\tau \otimes \tau)(C(u_1, v_1, u_3, v_3)(E_{A_z} \otimes E_{B_z})) \\
\leq&(1+\alpha) \int \lambda\mu \, d(\tau \otimes \tau)((C(u_1, v_1, z)E_{A_z}) \otimes E_{B_z}) \\
&+ (1+\frac{1}{\alpha}) \int \lambda\mu \, d(\tau \otimes \tau)(E_{A_z} \otimes (C(z, u_3, v_3)E_{B_z})) \\
=&(1+\alpha) \iint \lambda\mu \, d\tau(C(u_1, v_1, z)E_{A_z}) \, d\tau(E_{B_z}) \\
&+ (1+\frac{1}{\alpha}) \iint \lambda\mu \, d\tau(E_{A_z}) \, d\tau(C(z, u_3, v_3)E_{B_z}) \\
=&(1+\alpha) \int \lambda \, d\tau(C(u_1, v_1, z)E_{A_z}) \int \mu \, d\tau(E_{B_z}) \\
&+ (1+\frac{1}{\alpha}) \int \lambda \, d\tau(E_{A_z}) \int \mu \, d\tau(C(z, u_3, v_3)E_{B_z}).
\end{aligned}$$

By Lemma 4.7 and Corollary 4.8, we have

$$\begin{aligned}
\int \lambda \, d\tau(C(u_1, v_1, z)E_{A_z}) \int \mu \, d\tau(E_{B_z}) =&\tau\left(C(u_1, v_1, z) \int \lambda \, dE_{A_z}\right) \tau\left(\int \mu \, dE_{B_z}\right) \\
=&\tau(C(u_1, v_1, z)A_z)\tau(B_z),
\end{aligned}$$

and similarly,

$$\int \lambda \, d\tau(E_{A_z}) \int \mu \, d\tau(C(z, u_3, v_3)E_{B_z}) = \tau(A_z)\tau(C(z, u_3, v_3)B_z).$$

Therefore,

$$
\begin{aligned}
d(\rho_1,\rho_3)^2 \leq &\int_{\mathbb{T}^2} g^{-1}(z)\left((1+\alpha)\tau(C(u_1,v_1,z)A_z)\tau(B_z)+(1+\frac{1}{\alpha})\tau(A_z)\tau(C(z,u_3,v_3)B_z)\right)dz\\
=&(1+\alpha)\int_{\mathbb{T}^2}\tau(C(u_1,v_1,z)T_{12}(z))\,dz+(1+\frac{1}{\alpha})\int_{\mathbb{T}^2}\tau(C(z,u_3,v_3)T_{23}(z))\,dz\\
=&(1+\alpha)(\tau\otimes\int)(C(u_1,v_1,z)T_{12})+(1+\frac{1}{\alpha})(\int\otimes\tau)(C(z,u_3,v_3)T_{23}).
\end{aligned}
\tag{4.8}
$$

Minimizing the right hand side in $T_{12}\in\mathcal{C}(\rho_1,g)$ and in $T_{23}\in\mathcal{C}(g,\rho_3)$ shows that

$$
d(\rho_1,\rho_3)^2\leq(1+\alpha)d(\rho_1,g)^2+(1+\frac{1}{\alpha})d(g,\rho_3)^2.
\tag{4.9}
$$

Minimizing right hand side

$$
\begin{aligned}
(1+\alpha)d(\rho_1,g)^2+(1+\frac{1}{\alpha})d(g,\rho_3)^2=&d(\rho_1,g)^2+d(g,\rho_3)^2+\alpha d(\rho_1,g)^2+\frac{1}{\alpha}d(g,\rho_3)^2\\
\geq&d(\rho_1,g)^2+d(g,\rho_3)^2+2d(\rho_1,g)d(g,\rho_3)\\
=&(d(\rho_1,g)+d(g,\rho_3))^2,
\end{aligned}
\tag{4.10}
$$

equality holds if and only if $\alpha d(\rho_1,g)^2=\frac{1}{\alpha}d(g,\rho_3)^2$, that is, $\alpha=\frac{d(g,\rho_3)}{d(\rho_1,g)}$. Thus,

$$
d(\rho_1,\rho_3)\leq d(\rho_1,g)+d(g,\rho_3).
$$

## 4.5 An alternative proof

In this section, we provide an alternative proof for the triangle inequality (3.3). The assertions mentioned in the last paragraph of Section 4.1.2 are indeed true. This proof is based on those assertions, and is shorter and more direct than the previous one.

Let $T_{12}\in\mathcal{C}(\rho_1,g)$ and let $T_{23}\in\mathcal{C}(g,\rho_3)$. By Lemma 4.12, we identify $T_{12},T_{23}$ as $L^1(\mathbb{T}^2_\theta)$-valued functions on $\mathbb{T}^2$. Define

$$
T_{123}\coloneqq g^{-1}(z)T_{12}(z)\otimes T_{23}(z),
$$

where $g^{-1}(z)$ is given by (4.7). Observe that

$$
\|T\|_1=\int_{\mathbb{T}^2}(\tau\otimes\tau)(T_{123}(z))\,dz=\int_{\mathbb{T}^2}g(z)\,dz=1,
$$

so $T\in L^1(\mathbb{T}^2_\theta\bar{\otimes}\mathbb{T}^2\bar{\otimes}\mathbb{T}^2_\theta)$. Also note that

$$
(\mathrm{id}\otimes\tau)\,(T_{123}(z))=g^{-1}(z)T_{12}(z)\tau(T_{23}(z))=T_{12}(z).
$$

It follows that

$$
(\mathrm{id}\otimes\tau)\left(\int_{\mathbb{T}^2}T_{123}(z)\,dz\right)=\int_{\mathbb{T}^2}(\mathrm{id}\otimes\tau)(T_{123}(z))\,dz=\int_{\mathbb{T}^2}T_{12}(z)\,dz=\rho_1.
$$

Similarly,

$$
(\tau\otimes\mathrm{id})(T_{123}(z))=T_{23}(z),\quad(\tau\otimes\mathrm{id})\left(\int_{\mathbb{T}^2}T_{123}(z)\,dz\right)=\rho_3.
$$

This implies that $\int_{\mathbb{T}^2} T_{123}(z)\,dz \in \mathcal{C}(\rho_1,\rho_3)$.

We have $T_{123}(z)^{\frac{1}{2}} \in L^2(\mathbb{T}^2_\theta \bar{\otimes} \mathbb{T}^2_\theta)$ since $\tau(|T(z)^{\frac{1}{2}}|^2) = \tau(T_{123}(z)) < \infty$. It follows from Proposition 3.3.4 in [16] that

$$\begin{aligned}(\tau\otimes\tau)(|C(u_1,v_1,u_3,v_3)T_{123}(z)^{\frac{1}{2}}|^2) =&(\tau\otimes\tau)(T_{123}(z)^{\frac{1}{2}}C(u_1,v_1,u_3,v_3)^2T_{123}(z)^{\frac{1}{2}})\\ =&(\tau\otimes\tau)(C(u_1,v_1,u_3,v_3)T_{123}C(u_1,v_1,u_3,v_3))\\ \le&\|C(u_1,v_1,u_3,v_3)\|^2(\tau\otimes\tau)(T_{123}).\end{aligned}$$

This implies $C(u_1,v_1,u_3,v_3)T_{123}^{\frac{1}{2}}(z) \in L^2(\mathbb{T}^2_\theta \bar{\otimes} \mathbb{T}^2_\theta)$. Then by Proposition 3.4.3 in [16], we obtain

$$\begin{aligned}(\tau\otimes\tau)(C(u_1,v_1,u_3,v_3)T_{123}(z)) =&(\tau\otimes\tau)(C(u_1,v_1,u_3,v_3)T_{123}(z)^{\frac{1}{2}}T_{123}(z)^{\frac{1}{2}})\\ =&(\tau\otimes\tau)(T_{123}(z)^{\frac{1}{2}}C(u_1,v_1,u_3,v_3)T_{123}(z)^{\frac{1}{2}}).\end{aligned}$$

By Lemma 4.5 and Proposition 2.2.24 in [16], we obtain

$$\begin{aligned}T_{123}(z)^{\frac{1}{2}}C(u_1,v_1,u_3,v_3)T_{123}(z)^{\frac{1}{2}} \le&(1+\alpha)T_{123}(z)^{\frac{1}{2}}\left(C(u_1,v_1,z)\otimes\mathbf{1}\right)T_{123}(z)^{\frac{1}{2}}\\ &+(1+\frac{1}{\alpha})T_{123}(z)^{\frac{1}{2}}\left(\mathbf{1}\otimes C(z,u_3,v_3)\right)T_{123}(z)^{\frac{1}{2}}.\end{aligned}$$

Then it follows that

$$\begin{aligned}&(\tau\otimes\tau)(T_{123}^{\frac{1}{2}}(z)C(u_1,v_1,u_3,v_3)T_{123}^{\frac{1}{2}}(z))\\ \le&(1+\alpha)(\tau\otimes\tau)(T_{123}^{\frac{1}{2}}(z)(C(u_1,v_1,z)\otimes\mathbf{1})T_{123}^{\frac{1}{2}}(z))\\ &+(1+\frac{1}{\alpha})(\tau\otimes\tau)(T_{123}^{\frac{1}{2}}(z)(\mathbf{1}\otimes C(z,u_3,v_3))T_{123}^{\frac{1}{2}}(z)).\end{aligned}$$

Therefore,

$$\begin{aligned}d(\rho_1,\rho_3)^2 \le&(\tau\otimes\tau)(C(u_1,v_1,u_3,v_3)T_{13})\\ =&(\tau\otimes\tau)\left(C(u_1,v_1,u_3,v_3)\int_{\mathbb{T}^2}T_{123}(z)\,dz\right)\\ =&\int_{\mathbb{T}^2}(\tau\otimes\tau)(C(u_1,v_1,u_3,v_3)T_{123}(z))\,dz\\ \le&(1+\alpha)\int_{\mathbb{T}^2}(\tau\otimes\tau)((C(u_1,v_1,z)\otimes\mathbf{1})T_{123}(z))\,dz\\ &+(1+\frac{1}{\alpha})\int_{\mathbb{T}^2}(\tau\otimes\tau)((\mathbf{1}\otimes C(z,u_3,v_3))T_{123}(z))\,dz\\ =&(1+\alpha)\int_{\mathbb{T}^2}\tau(C(u_1,v_1,z)T_{12}(z))\tau(T_{23}(z))g^{-1}(z)\,dz\\ &+(1+\frac{1}{\alpha})\int_{\mathbb{T}^2}\tau(T_{12}(z))g^{-1}(z)\tau(C(z,u_3,v_3)T_{23}(z))\,dz\\ =&(1+\alpha)\int_{\mathbb{T}^2}\tau(C(u_1,v_1,z)T_{12}(z))\,dz+(1+\frac{1}{\alpha})\int_{\mathbb{T}^2}\tau(C(z,u_3,v_3)T_{23}(z))\,dz\\ =&(1+\alpha)(\tau\otimes\int)(C(u_1,v_1,z)T_{12})+(1+\frac{1}{\alpha})(\int\otimes\tau)(C(z,u_3,v_3)T_{23}).\end{aligned}$$

This is exactly inequality (4.8). Repeat the minimizing steps as in (4.9) and (4.10), we obtain

$$d(\rho_1,\rho_3) \le d(\rho_1,g)+d(g,\rho_3).$$

# Chapter 5

# Conclusion and discussion

In this concluding chapter, we summarize the main findings of this thesis and discuss some open problems of interest. First, the key notions of the optimal transport problem on the noncommutative 2-torus are reviewed. Then the metric properties of the classical Wasserstein distance $W_p$ (1.1), the extended Wasserstein distance $\mathfrak{d}$ (1.4), and the quantum Wasserstein distance $d$ (3.2) are discussed. In particular, the triangle inequality of $d$ in the case where the middle point is a classical density is formulated. The main method used in the proof, namely spectral decomposition, is discussed. In the end, we propose a possible improvement to the cost function $C$ (3.1) that may render it more physically meaningful and point out the unverified metric properties of the quantum Wasserstein distance $d$.

## 5.1 Review of problem formulation

To formulate the optimal transport problem on noncommutative 2-tori, we introduced a series of notions. First, the noncommutative 2-torus and densities on it are defined by universal C*-algebras (Definition 2.12) and noncommutative $L^p$-spaces (Definition 2.2, 2.6). Specifically, the noncommutative 2-torus $A_\theta$ is defined as the universal C*-algebra generated by two unitaries $u, v$ subject to the commuting relation $uv = e^{i\theta} vu$. In Proposition 2.5, we introduced a faithful tracial state $\tau$ on $A_\theta$. To define the densities on the noncommutative 2-torus, the von Neumann algebra $L^\infty(\mathbb{T}^2_\theta)$ generated by $A_\theta$ is considered. Meanwhile, we also see in Proposition 2.6 that $\tau$ extends to a normal finite faithful trace on $L^\infty(\mathbb{T}^2_\theta)$. Then the pair $(L^\infty(\mathbb{T}^2_\theta), \tau)$ is a noncommutative measure space, and the density operator is defined as the positive trace-one operators in the noncommutative $L^1$-space associated with $L^\infty(\mathbb{T}^2_\theta)$.

To properly define the transport plans for two density operators, in Section 2.5.2, we reviewed the definition of transport plans for probability measures in classical optimal transport and quantum transport plans for quantum and classical densities proposed by Golse and Paul [21]. Concerning the transport plans for density operators defined in [21], we introduced an equivalent definition that uses partial traces $\mathrm{id} \otimes \mathrm{tr}$ and $\mathrm{tr} \otimes \mathrm{id}$ which make the marginal conditions (2.2) more concise. Then the transport plans for density operators $\rho_1, \rho_2$ associated with $L^\infty(\mathbb{T}^2_{\theta_1}), L^\infty(\mathbb{T}^2_{\theta_2})$ are analogously defined as the density operators $T$ associated with $L^\infty(\mathbb{T}^2_{\theta_1} \bar{\otimes} \mathbb{T}^2_{\theta_2})$ such that their partial traces satisfy

$$(\mathrm{id} \otimes \tau)(T) = \rho_1, \quad (\tau \otimes \mathrm{id})(T) = \rho_2.$$

To find an appropriate cost function, we reviewed the cost function for classical Wasserstein distances (1.1) and the cost functions $c_\hbar, C_\hbar$ defined in (1.2), (1.3) respectively. Consequently, we defined the cost function $C(u_1, v_1, u_2, v_2)$ for transport plans between $\rho_1, \rho_2$ as

$$C(u_1, v_1, u_2, v_2) = |u_1 \otimes \mathbf{1} - \mathbf{1} \otimes u_2|^2 + |v_1 \otimes \mathbf{1} - 1 \otimes v_2|^2. \tag{5.1}$$

In the classical case, namely $\theta_1 = \theta_2 = 0$, this cost function is just the cost function for quadratic Wasserstein distance (as remarked in Chapter 3). Then the cost function $C(u_1, v_1, u_2, v_2)$ can be considered as the square of the difference between two "position operators".

With the density operators, transport plans, and cost function defined, the quantum Wasserstein distance on $\bigcup_{\theta \in \mathbb{R}} \mathcal{D}(\mathbb{T}_\theta^2)$ is defined as

$$d(\rho_1, \rho_2) = \left( \inf_{T \in \mathcal{C}(\rho_1, \rho_2)} (\tau \otimes \tau) \left( C(u_1, v_1, u_2, v_2) T \right) \right)^{\frac{1}{2}}. \tag{5.2}$$

This concludes the formulation of the optimal transport problem on the noncommutative 2-torus.

## 5.2 Triangle inequalities and other metric properties

Moreover, we introduced and investigated the triangle inequality and other metric properties for the classical Wasserstein distance $W_p$ (1.1), the extended Wasserstein distance $\mathfrak{d}$ (1.4), and the quantum Wasserstein distance $d$ (3.2).

For the classical Wasserstein distance $W_p$, Theorem 7.8 in [38] states that it defines a metric on the space of probability measure with finite $p$-th moments on some Polish space. It is consistent with the intuition that moving one distribution to another identical distribution should not cost work; moving one distribution to a different distribution should cost some amount of work; moving one distribution to another distribution should cost the same work as the opposite direction; and moving one distribution to another distribution directly should cost less work than first moving it to some intermediate distribution and then moving it to the intended destination distribution. Meanwhile, since $W_p$ is a metric, it is possible to discuss the convergence of probability measures with respect to $W_p$. Therefore, the classical Wasserstein distance provides a meaningful and rigorous way to measure the difference between probability distributions, and it has a wide range of applications.

For the distance $\mathfrak{d}$ defined by Golse and Paul, it is proved in Lemma 2.1 of [21] that $\mathfrak{d}$ has a nonzero lower bound if one of the densities is a quantum density operator, and thus $\mathfrak{d}$ fails to be metric. They also checked the other properties of metrics for $\mathfrak{d}$. The symmetry of it can be observed from its definition. The triangle inequality property of $\mathfrak{d}$ was investigated using the method of spectral decomposition. They proved that $\mathfrak{d}$ satisfies triangle inequality in the case where the middle point is a classical density or one of the densities is a rank-one operator (Theorem 1.1).

Motivated by their proof for Theorem 1.1, we sought to obtain a similar triangle inequality result for the quantum Wasserstein distance $d$. It is observed from their proof that when the middle point is a classical density, the "joint transport plan" can be constructed by considering the tensor product of two transport plans; when one of the densities is a rank-one operator, the only transport plan between it and another density is their tensor product [21, Lemma 3.6] which makes it easy to construct the "joint transport plan" and prove the triangle inequality. However, if it is not either of these two cases, finding the "joint transport plan" is not easy. Since $L^\infty(\mathbb{T}^2)$ is a $\mathrm{II}_1$ factor, there is no minimal projection, and thus the structure of the transport plan between a noncommutative density operator and another density is not simple. Consequently, if the middle point is not classical, it is difficult to find the "joint transport plan". Thus, we only considered the case where the middle point is a classical density and proved that the triangle inequality of the quantum Wasserstein distance $d$ holds. The statement of this result is formulated in the following theorem:

**Theorem 5.1.** *The map* $d : \bigcup_{\theta\in\mathbb{R}} \mathcal{D}(\mathbb{T}^2_\theta) \times \bigcup_{\theta\in\mathbb{R}} \mathcal{D}(\mathbb{T}^2_\theta) \to \mathbb{R}_{\geq 0}$ *satisfies the triangle inequality*

$$d(\rho_1, \rho_3) \leq d(\rho_1, g) + d(g, \rho_3),$$

*if* $\rho_1, \rho_3 \in \mathcal{D}(\mathbb{T}^2_\theta)$ *and* $g \in \mathcal{D}(\mathbb{T}^2)$ *for some* $\theta \in \mathbb{R}$.

## 5.3 Spectral decompositions

In the proof for Theorem 1.1, the spectral decompositions of density operators are considered. Since the density operators on the Hilbert space are compact, their spectral decompositions can be written as infinite sums of rank-one projections with eigenvalues as coefficients. Moreover, the unit vectors in the images of the projections form an orthonormal basis for the Hilbert space, which played an important role in computing the trace of $Q^{13}(y, \eta)^{\frac{1}{2}} C_\hbar Q^{13}(y, \eta)^{\frac{1}{2}}$ and deriving the triangle inequality.

This method motivated our proof for Theorem 5.1. Since $L^\infty(\mathbb{T}^2)$ does not contain rank-one projections, we considered the spectral decompositions based on spectral measures, which can be viewed as the continuous generalization of the spectral decompositions for compact operators. Regarding the generalized spectral decompositions, we proved a series of assertions that are analogous to the existing results about compact operator spectral decompositions. For example, the trace of a positive trace-class operator is the sum of its eigenvalues. Similarly, we showed in Corollary 4.8 that the trace of a positive operator in the noncommutative $L^1$-space associated with $L^\infty(\mathbb{T}^2_\theta)$ is equal to the integral of its spectrum with respect to the composition of the trace and its spectral measure. Thus, the proof for Theorem 5.1 in Section 4.4 can be viewed as a continuous analogue of the proof for Theorem 1.1.

## 5.4 Open problems

There are still some issues that need to be addressed. Both of the cost functions $c_\hbar$ (1.2) and $C_\hbar$ (1.3) are the sums of the square of the difference in "position variables" and the square of the difference in "momentum operators". Such cost functions can be interpreted as the sum of the difference in potential energy and the difference in kinetic energy. From this point of view, the quantum Wasserstein distance (5.2) may lack a physical interpretation. It only captured the difference in potential energy, not the difference in kinetic energy. Note that the "momentum operator" in $C_\hbar$ is given by $-i\hbar\nabla$. To complete the physical meaning, further studies can add the square of the difference in "momentum operators" to the cost function (5.1) by introducing the differential operators on the noncommutative 2-torus.

As discussed previously about the importance of metric properties, we hope that the quantum Wasserstein distance (5.2) also satisfies metric properties. In Chapter 3, we have shown that $d(\cdot, \cdot)$ is symmetric and non-negative. However, it is not yet clear whether $d(\rho, \rho) = 0$ for any $\rho \in L^1(\mathbb{T}^2_\theta)$, or whether $d(\rho_1, \rho_2) > 0$ for any $\rho_1 \neq \rho_2$ where $\rho_1 \in L^1(\mathbb{T}^2_{\theta_1}), \rho_2 \in L^1(\mathbb{T}^2_{\theta_1})$. Moreover, the triangle inequality property in general cases also requires further study.

# Bibliography


[1] C*-algebras and dynamics. ISem24 lecture notes, 2021. https://www.math.uni-sb.de/ag/speicher/weber/ISem24/ISem24LectureNotes.pdf. Accessed 19 May 2024.

[2] S. Attal. Tensor products and partial traces. Lectures in Quantum Noise Theory. http://math.univ-lyon1.fr/~attal/Partial_traces.pdf. Accessed 19 May 2024.

[3] M. Birman and M. Solomyak. *Spectral Theory of Self-Adjoint Operators in Hilbert Space*. Springer Dordrecht, 1987. (Originally published in Russian).

[4] M. Birman and M. Solomyak. Tensor product of a finite number of spectral measures is always a spectral measure. *Integral Equations and Operator Theory*, 24(2):179–187, 1996.

[5] Y. Brenier. Décomposition polaire et réarrangement monotone des champs de vecteurs. *C. R. Acad. Sci. Série I Math.*, 305:805–808, 1987.

[6] 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. *Communications in Mathematical Physics*, 331(3):887–926, 2014.

[7] E. A. Carlen and J. Maas. Non-commutative calculus, optimal transport and functional inequalities in dissipative quantum systems. *Journal of Statistical Physics*, 178(2):319–378, 2020.

[8] A. Connes. Non-commutative differential geometry. *Publications Mathématiques de l'IHÉS*, 62:41–144, 1985.

[9] A. Connes. *Noncommutative Geometry*. Academic Press, 1994.

[10] A. Connes. C*algebras and differential geometry. arXiv: hep-th/0101093, 2001. (Originally published in French).

[11] A. Connes and M. Marcolli. A Walk in the Noncommutative Garden. In M. Khalkhali and M. Matilde, editors, *An Invitation to Noncommutative Geometry*, pages 1–128. World Scientific Publishing Company, Singapore, 2008.

[12] A. Connes and M. Marcolli. *Noncommutative Geometry, Quantum Fields and Motives*, volume 55 of *Colloquium Publications*. American Mathematical Society, 2008.

[13] J. B. Conway. *A Course in Functional Analysis*. Springer New York, 2007.

[14] N. Datta and C. Rouzé. Relating relative entropy, optimal transport and Fisher information: a quantum HWI inequality. *Annales Henri Poincaré*, 21(7):2115–2150, 2020.

[15] G. De Palma and D. Trevisan. Quantum optimal transport with quantum channels. *Annales Henri Poincaré*, 22(10):3199–3234, 2021.

[16] P. Dodds, B. de Pagter, and F. Sukochev. *Noncommutative Integration and Operator Theory.* Birkhäuser, 2024.

[17] A. Ern and J. Guermond. *Bochner integration*, chapter 64, pages 111–122. Texts in Applied Mathematics. Springer, Cham, 2021.

[18] F. Golse, C. Mouhot, and T. Paul. On the mean field and classical limits of quantum mechanics. *Communications in Mathematical Physics*, 343:165–205, 2016.

[19] F. Golse and T. Paul. The Schrödinger equation in the mean-field and semiclassical regime. *Archive for Rational Mechanics and Analysis*, 223:57–94, 2017.

[20] F. Golse and T. Paul. Semiclassical evolution with low regularity. *Journal de Mathématiques Pures et Appliquées*, 151:257–311, 2021.

[21] F. Golse and T. Paul. Optimal transport pseudometrics for quantum and classical densities. *Journal of Functional Analysis*, 282(9):109417, 2022.

[22] U. Haagerup, M. Junge, and Q. Xu. A reduction method for noncommutative $L_p$-spaces and applications. *Transactions of the American Mathematical Society*, 362(4):2125–2165, 2010.

[23] K. Ikeda. Foundation of quantum optimal transport and applications. *Quantum Information Processing*, 19(1):25, 2019.

[24] P. T. Johnstone. *Stone Spaces.* Cambridge University Press, 1986.

[25] L. V. Kantorovich. On the translocation of masses. *Dokl. Akad. Nauk SSSR*, 37(7–8):227–229, 1942.

[26] L. V. Kantorovich. Mathematical methods of organizing and planning production. *Management science*, 6(4):366–422, 1960.

[27] S. Kolouri, S. R. Park, M. Thorpe, D. Slepcev, and G. K. Rohde. Optimal mass transport: Signal processing and machine-learning applications. *IEEE Signal Processing Magazine*, 34(4):43–59, 2017.

[28] L. D. Landau and E. M. Lifshitz. *Quantum Mechanics: Non-Relativistic Theory.* Pergamon Press, 3rd edition, 1977.

[29] E. Lupercio. Non-commutative geometry indomitable. *Notices of the American Mathematical Society*, 68(1), 2021.

[30] G. Monge. Mémoire sur la théorie des déblais et des remblais. In *Histoire de l'Académie Royale des Sciences de Paris, avec les Mémoires de Mathématique et de Physique pour la même année*, pages 666–704. De l'Imprimerie Royale, 1781.

[31] G. J. Murphy. *C*-Algebras and Operator Theory.* Academic Press, 1990.

[32] O. Pele and M. Werman. A linear time histogram metric for improved sift matching. In D. Forsyth, P. Torr, and A. Zisserman, editors, *Computer Vision – ECCV 2008*, pages 495–508, Berlin, Heidelberg, 2008. Springer Berlin Heidelberg.

[33] G. Pisier and Q. Xu. Non-Commutative $L^p$-Spaces. In W. Johnson and J. Lindenstrauss, editors, *Handbook of the Geometry of Banach Spaces*, volume 2, chapter 34, pages 1459–1517. Elsevier Science B.V., 2003.

[34] M. Rieffel. C*-algebras associated with irrational rotations. *Pacific Journal of Mathematics*, 93(2):415–429, 1981.

[35] W. F. Stinespring. Integration theorems for gages and duality for unimodular groups. *Transactions of the American Mathematical Society*, 90(1):15–56, 1959.

[36] M. Takesaki. *Theory of Operator Algebras I.* Springer Berlin Heidelberg, 2001.

[37] L. N. Vaserstein. Markov processes over denumerable products of spaces, describing large systems of automata. *Problemy Peredachi Informatsii*, 5(3):64–72, 1969.

[38] C. Villani. *Topics in Optimal Transportation.* Graduate studies in mathematics. American Mathematical Society, 2003.

[39] N. Weaver. *Mathematical Quantization.* Studies in Advanced Mathematics. CRC Press, 2001.

[40] D. Williams. *Crossed Products of C*-Algebras.* American Mathematical Society, 2007.

[41] Q. Xu, T. N. Bekjan, and Z. Chen. *Introduction to Operator Algebras and Noncommutative* $L_p$*-Spaces.* Science Press, Beijing, 2010. (in Chinese).

[42] K. Zyczkowski and W. Slomczynski. The Monge distance between quantum states. *Journal of Physics A: Mathematical and General*, 31(45):9095, 1998.