On Quantum Optimal Transport

Sam Cole1 · Michał Eckstein2 · Shmuel Friedland3 · Karol Życzkowski2,4

Received: 15 December 2022 / Accepted: 11 May 2023 / Published online: 6 June 2023
© The Author(s), under exclusive licence to Springer Nature B.V. 2023

Abstract
We analyze a quantum version of the Monge–Kantorovich optimal transport problem. The quantum transport cost related to a Hermitian cost matrix $C$ is minimized over the set of all bipartite coupling states $\rho^{AB}$ with fixed reduced density matrices $\rho^A$ and $\rho^B$ of size $m$ and $n$. The minimum quantum optimal transport cost $T^Q_C(\rho^A, \rho^B)$ can be efficiently computed using semidefinite programming. In the case $m = n$ the cost $T^Q_C$ gives a semidistance if and only if $C$ is positive semidefinite and vanishes exactly on the subspace of symmetric matrices. Furthermore, if $C$ satisfies the above conditions, then $\sqrt{T^Q_C}$ induces a quantum analogue of the Wasserstein-2 distance. Taking the quantum cost matrix $C^Q$ to be the projector on the antisymmetric subspace, we provide a semi-analytic expression for $T^Q_{C^Q}$ for any pair of single-qubit states and show that its square root yields a transport distance on the Bloch ball. Numerical simulations suggest that this property holds also in higher dimensions. Assuming that the cost matrix suffers decoherence and that the density matrices become diagonal, we study the quantum-to-classical transition of the Monge–Kantorovich distance, propose a continuous family of interpolating distances, and demonstrate that the quantum transport is cheaper than the classical one. Furthermore, we introduce a related quantity—the SWAP-fidelity—
and compare its properties with the standard Uhlmann–Jozsa fidelity. We also discuss the quantum optimal transport for general $d$-partite systems.

**Keywords** Quantum optimal transport · Classical optimal transport · Coupling of density matrices · Semidefinite programming · Wasserstein-2 distance

**Mathematics Subject Classification** 81P40 · 90C22 · 15A69

1 Introduction

1.1 Classical Optimal Transport for Discrete Variables

Let us recall the discrete optimal transport problem [61], as stated in Hitchcock [34] and Kantorovich [39] (prepared in 1939), which is a variation of the classical transport problem initiated by Monge [46]. Suppose we have $m$ factories producing an amount $G$ of the same product that has to be dispatched to $n$ consumers. Assume that $x_{i}^{AB}$ is the proportion of the goods sent from the factory $i$ to consumer $j$. Then $x_{i}^{A}$ and $x_{j}^{B}$ are the proportions of the goods produced by factory $i$ and received by consumer $j$ respectively:

$$x_{i}^{A} = \sum_{j=1}^{n} x_{i}^{AB}, \quad i \in [m], \quad x_{j}^{B} = \sum_{i=1}^{m} x_{i}^{AB}, \quad j \in [n], \quad (1.1)$$

where $[m] = \{1, 2, \ldots, m\}$. It is convenient to introduce the random variables $X^{A}, X^{B}$ such that

$$x_{i}^{A} = P(X^{A} = i), \quad i \in [m], \quad x_{j}^{B} = P(X^{B} = j), \quad j \in [n].$$

Then the nonnegative matrix $X^{AB} = [x_{ij}^{AB}] \in \mathbb{R}_{+}^{m \times n}$ satisfying the above equalities is the joint distribution of the random variable $X^{AB}: x_{ij}^{AB} = P(X^{AB} = (i, j))$. The random variable $X^{AB}$, or the matrix $X^{AB}$, is called a coupling of $X^{A}$ and $X^{B}$. Let $x^{A} = (x_{1}^{A}, \ldots, x_{m}^{A})^{\top}, x^{B} = (x_{1}^{B}, \ldots, x_{n}^{B})^{\top}$ be the probability vectors corresponding to $X^{A}$ and $X^{B}$ respectively. The set of all coupling matrices $X^{AB}$ corresponding to $x^{A}, x^{B}$ is denoted by $\Gamma^{cl}(x^{A}, x^{B})$. Note that $X = x^{A}(x^{B})^{\top}$, corresponding to the independent coupling of $X^{A}$ and $X^{B}$, is in $\Gamma^{cl}(x^{A}, x^{B})$. Let $C = [c_{ij}] \in \mathbb{R}_{+}^{m \times n}$ be a nonnegative matrix where $c_{ij}$ is the transport cost of a unit of goods from the factory $i$ to the consumer $j$. The classical optimal transport problem, abbreviated as OT, is

$$\text{Tr}_{c}^{cl}(x^{A}, x^{B}) = \min_{X \in \Gamma^{cl}(x^{A}, x^{B})} \text{Tr} \ C \ X^{\top}, \quad (1.2)$$

where Tr denotes the trace of a square matrix, and $X^{\top}$ the transpose of $X$. The optimal transport problem is a linear programming problem (LP) which can be solved in polynomial time in the size of the inputs $x^{A}, x^{B}$ and the matrix $C$ [17].
Assume now that \( m = n \). Let \( C = [c_{ij}] \in \mathbb{R}^{n \times n}_+ \) be a symmetric nonnegative matrix with zero diagonal and positive off-diagonal entries such that \( c_{ij} \) induces a distance on \([n]\): \( \text{dist}(i, j) = c_{ij} \). That is, in addition to the above conditions one has the triangle inequality \( c_{ij} \leq c_{ik} + c_{kj} \) for \( i, j, k \in [n] \). For \( p > 0 \) denote \((C^o)^p = [c_{ij}^p] \in \mathbb{R}^{n \times n}_+ \). Then the quantity

\[
W_{C, p}^{cl}(x^A, x^B) = \left( T_{(C^o)^p}(x^A, x^B) \right)^{1/p}, \quad p \geq 1
\]  

(1.3)

is the Wasserstein-\( p \) distance on the simplex of probability vectors, \( \Pi_n \subset \mathbb{R}^n_+ \). This follows from the continuous version of the Wasserstein-\( p \) distance, as in \([60]\). See \([18]\) for \( p = 1 \). It turns out that \( T_{C}^{cl}(x^A, x^B) \) has many recent applications in machine learning \([2, 3, 43, 47, 55]\), statistics \([8, 24, 49, 57]\) and computer vision \([10, 53, 54]\).

### 1.2 Optimal Transport for Density Matrices and Related Distances

Several attempts to generalize the notion of the Monge–Kantorovich distance in quantum information theory are known. An early contribution defines the distance between any two quantum states by the Monge distance between the corresponding Husimi functions \([66, 67]\). As this approach depends on the choice of the set of coherent states, other efforts were undertaken \([1, 15, 21, 32, 33]\) to introduce the transport distance between quantum states by applying the Kantorovich–Wasserstein optimization over the set of bipartite quantum states with fixed marginals. However, none of the proposed ‘distances’ satisfy all of the properties of a genuine distance. Even though the matrix transport problem has been often investigated in the recent literature \([5, 7, 16, 22, 27–29]\), this aim has not been fully achieved until now \([37, 52, 65]\). Meanwhile, the quantum optimal transport has found a number of applications in quantum physics: the measure of proximity of quantum states \([13, 14, 19, 20]\), quantum metrology \([12, 42, 56]\), and quantum machine learning \([6, 15, 41, 44]\).

Denote by \( \Omega_m \) the convex set of density matrices, i.e., the set of \( m \times m \) Hermitian positive semidefinite matrices of trace one. Let \( \rho^A \in \Omega_m \) and \( \rho^B \in \Omega_n \). A quantum coupling of \( \rho^A, \rho^B \) is a density matrix \( \rho^{AB} \in \Omega_{mn} \), whose partial traces give \( \rho^A, \rho^B \) respectively: \( \text{Tr}_B \rho^{AB} = \rho^A \) and \( \text{Tr}_A \rho^{AB} = \rho^B \). The set of all quantum couplings of \( \rho^{AB} \) is denoted by \( \Gamma^Q(\rho^A, \rho^B) \). Observe that we always have \( \rho^A \otimes \rho^B \in \Gamma^Q(\rho^A, \rho^B) \).

Let \( C \) be a given positive Hermitian matrix of order \( mn \). The quantum optimal transport problem, abbreviated as QOT, is defined as follows:

\[
T^Q_C(\rho^A, \rho^B) = \min_{\rho^{AB} \in \Gamma^Q(\rho^A, \rho^B)} \text{Tr} C \rho^{AB}.
\]  

(1.4)

The matrix \( C \) can be viewed as a “cost matrix” in certain instances that will be explained later. Such a formulation of the quantum optimal transport was, to our best knowledge, first proposed in \([32]\) in the context of density operators on the infinite dimensional Hilbert space \( L^2(\mathbb{R}^d) \) (cf. also \([63]\)). The finite dimensional version (1.4), which is a natural generalization of (1.2), appeared first in \([15]\) and also in \([52]\).

The quantum optimal transport has a simple operational interpretation. Suppose that Alice and Bob represent two parties who share a bipartite state \( \rho^{AB} \). Their local
concerning i.a. the dual problem (Theorem 3.2), the notion of the weak distance, which induces a quantum distance (Theorem 6.1) and QOT for correspond the process of decoherence caused by the interaction of the system with 123 is also a semidistance for any and positive definite on \( \rho \).

The goal of this work is to analyze the properties of \( T^Q_C(\rho^A, \rho^B) \) and determine the class of quantum cost matrices, which induce a quantum analog of the Wasserstein distance (1.3). Some of the presented results were discussed, in the context of quantum information processing, in the companion paper [28]. Here we give the complete account of our mathematical study related to quantum optimal transport. It contains the full technical proofs of the results announced in [28], along with some new material concerning i.a. the dual problem (Theorem 3.2), the notion of the weak distance, which induces a quantum distance (Theorem 6.1) and QOT for \( d \)-partite systems (Sect. 7). The physical interpretation of our results on QOT and their potential applications in quantum information processing are discussed in the companion paper [28].

It is easy to show (see Sect. 3 and also [15]) that finding the value of \( T^Q_C(\rho^A, \rho^B) \) is a semidefinite programming problem (SDP). Using standard complexity results for SDP, as in [59, Theorem 5.1], we show that the complexity of finding the value of \( T^Q_C(\rho^A, \rho^B) \) within a given precision \( \varepsilon > 0 \) is polynomial in the size of the given data and \( \log \frac{1}{\varepsilon} \). Notably, there are quantum algorithms that offer a speedup for SDP [11].

It is useful to compare \( T^Q_C \) with its classical counterpart \( T^cl_{Ccl} \) defined as follows. Observe that the diagonal entries of \( \rho^A \) and \( \rho^B \) form two probability vectors \( p^A \) and \( p^B \). Physically, the suppression of the off-diagonal terms in a density matrix of a system corresponds the process of decoherence caused by the interaction of the system with some environment. For \( x \in \mathbb{R}^n, X \in \mathbb{R}^{n \times n} \) denote by \( \operatorname{diag}(x), \operatorname{diag}(X) \in \mathbb{R}^{n \times n} \) the diagonal matrices induced by the entries of \( x \) and the diagonal entries of \( X \), respectively. For \( p^A \in \Pi_m, p^B \in \Pi_n \) denote by \( \Gamma^Q_{de}(\operatorname{diag}(p^A), \operatorname{diag}(p^B)) \) the convex subset of diagonal matrices in \( \Gamma^Q(\operatorname{diag}(p^A), \operatorname{diag}(p^B)) \). We show that \( \Gamma^Q_{de}(\operatorname{diag}(p^A), \operatorname{diag}(p^B)) \) is isomorphic to the set \( \Gamma^cl_{p^A, p^B} \) of classical coupling matrices. Let now \( C_{cl} \in \mathbb{R}^{m \times n} \) be the matrix induced by the diagonal entries of \( C \) (see Sect. 4). Then,

\[
\Gamma^Q_C(\operatorname{diag}(p^A), \operatorname{diag}(p^B)) \leq T^cl_{C_{cl}}(p^A, p^B) \quad \text{for} \quad p^A \in \Pi_m, p^B \in \Pi_n. \tag{1.5}
\]

We give examples where strict inequality holds. Specific cases of this inequality were studied in [13].

Let us now concentrate on the most important case \( m = n \). We would like to find an analog of the Wasserstein-\( p \) distance on \( \Omega_n \). A symmetric function \( \operatorname{sdist}: \Omega_n \times \Omega_n \to [0, \infty) \) is called a semidistance when \( \operatorname{sdist}(\rho^A, \rho^B) = 0 \) if and only if \( \rho^A = \rho^B \). We show that \( T^Q_C \) is a semidistance if and only if \( C \) is zero on \( H_S \) and positive definite on \( H_A \), where \( H_S \) and \( H_A \) are the subspaces of symmetric and skew-symmetric \( n \times n \) matrices viewed as subspaces of \( \mathbb{C}^n \otimes \mathbb{C}^n = \mathbb{C}^{n \times n} = H_S \oplus H_A \). If \( C \) is zero on \( H_S \) and positive definite on \( H_A \), then so is \( C^p \) for any \( p > 0 \). Consequently, \( (T^Q_{C^p})^{1/p} \) is also a semidistance for any \( p > 0 \). We further show that \( \sqrt{T^Q_C} \) is a weak distance, i.e. there exists a distance \( D' \) on \( \Omega_n \) such that \( \sqrt{T^Q_C(\rho^A, \rho^B)} \geq D'(\rho^A, \rho^B) \) for all \( \rho^A, \rho^B \in \Omega_n \)—see Theorem 6.1). Then, we prove that for such \( C \) there exists a unique
maximum distance $D'$ on $\Omega_n$, which we shall call the induced quantum Wasserstein-2 distance, given by the formula:

$$W^Q_{C}(\rho^A, \rho^B) = \lim_{N \to \infty} \min_{\rho^{A_0}=\rho^A, \rho^{A_{N+1}}=\rho^B} \sum_{i=1}^{N+1} \sqrt{T^Q_C(\rho^{A_{i-1}}, \rho^{A_i})}.$$  \hfill (1.6)

Similar construction can be done for any $p \geq 2$. The distance (1.6) does not seem to be easily computable for a general cost matrix, however for some choices of $C$ formula (1.6) simplifies significantly.

A simple example of the quantum cost matrix is provided by $C^Q$—the orthogonal projection of $\mathbb{C}^{n \times n}$ on $\mathcal{H}_A$, as advocated also in [15, 22, 65] and [52]. It is straightforward to show that $C^Q = \frac{1}{2}(I - S)$, where $S$ is the SWAP operator, $S(x \otimes y) = y \otimes x$, while $I$ is the identity operator on $\mathbb{C}^n \otimes \mathbb{C}^n$. Note that $C^Q$ is a projection, hence $(C^Q)^p = C^Q$ for any $p > 0$. We show that $(T^Q_C)^{1/p}$ does not satisfy the triangle inequality for $p \in [1, 2)$. On the other hand, for the single-qubit case, $n = 2$, the square root of the optimal quantum transport cost, $\sqrt{T^Q_C}$, does form a distance. In fact, we show that $W^Q_{C^Q} = \sqrt{T^Q_C}$ for qubits. Furthermore, $\sqrt{T^Q_C}$ is a distance on pure states for any $n$ and numerical simulations strongly suggest that $\sqrt{T^Q_C}$ satisfies the triangle inequality for all density matrices in $\Omega_n$ for $n = 3, \ldots, 8$ (see [28] for the details). It is also remarkable, that the optimal quantum transport cost $T^Q_C$ is monotonous under all single-qubit quantum channels [9], though this property fails in higher dimensions [48].

For the specific cost matrix $C^Q$ one can study a related quantity, the SWAP-fidelity between any two states, defined as [28],

$$F_S(\rho^A, \rho^B) = \max_{\rho^{AB} \in \Gamma^Q(\rho^A, \rho^B)} \text{Tr} S \rho^{AB} = 1 - 2T^Q_C(\rho^A, \rho^B).$$ \hfill (1.7)

It is a symmetric function with the following properties: it is continuous, jointly concave, unitarily invariant, and super-multiplicative with respect to the tensor product. Moreover, $F_S$ is bounded from below by the standard Uhlmann–Jozsa fidelity [38, 58], $F$, and from above by its square root, $\sqrt{F}$.

A simple generalization of $C^Q$ is the following operator that vanishes on $\mathcal{H}_S$ and is positive definite on $\mathcal{H}_A$:

$$C^Q_E = \sum_{1 \leq i < j \leq n} e_{ij} \frac{1}{\sqrt{2}} (|i\rangle\langle j| - |j\rangle\langle i|)(|i\rangle\langle j| - |j\rangle\langle i|),$$ \hfill (1.8)

with $e_{ij} > 0$ for $1 \leq i < j \leq n$.

Here $|1\rangle, \ldots, |n\rangle$ is any orthonormal basis in $\mathcal{H}_n$, while the entries of a fixed symmetric matrix $e_{ij}$ can be interpreted as classical distances between the sites $i$ and $j$. We show
that decoherence of the marginal states, $\rho \rightarrow \text{diag}(\rho)$, decreases the cost of QOT for $C_{E}^{Q}$:

$$T_{C_{E}^{Q}}^{Q}(\text{diag}(\rho^{A}), \text{diag}(\rho^{B})) \leq T_{C_{E}^{Q}}^{Q}(\rho^{A}, \rho^{B}) \text{ for } \rho^{A}, \rho^{B} \in \Omega_{n}. \quad (1.9)$$

Similarly to the results in [27, 30] for the product distribution of finite measures, we define the quantum transport problem for $d$-partite states in Sect. 7. The analog of $C_{Q}^{Q}$ for multipartite systems is the projection onto the orthogonal complement of the bosonic subspace—the subspace of symmetric tensors in $\otimes^{d}C^{n}$.

1.3 Outline of the Article

To make the paper accessible to a wide mathematical audience we use a fusion of standard mathematical notation with the notation of Dirac common in the physics community (see Sect. 2.1). Furthermore, at the beginning of each section we summarize the results included therein. Therefore, we only briefly state here the main topics that each section deals with. In Sect. 2 we present some preliminary results that are used in the rest of the paper. Section 3 discusses the connection between QOT and semidefinite programming. In Sect. 4 we compare the classical and quantum optimal transport problems for diagonal density matrices. Section 5 gives lower bounds on the cost of the QOT with the cost matrix $C_{Q}^{Q}$ for any pair of density matrices of any dimension, with equality holding for qubits. This result (Theorem 5.1) is one of the main results of the paper, and will allow us to derive a semi-analytic formula for $T_{C_{Q}^{Q}}^{Q}$ for qubits, as discussed in Appendix B. In Sect. 6 we show that any positive semidefinite cost matrix $C \in S(\mathcal{H}_{n} \otimes \mathcal{H}_{n})$ that vanishes exactly on the subspace of symmetric matrices yields the induced Wasserstein-2 distance (1.6). For the qubit cost matrix $C_{Q}^{Q}$ this Wasserstein-2 distance is simply $\sqrt{T_{C_{Q}^{Q}}^{Q}(\rho^{A}, \rho^{B})}$. Section 7 discusses the quantum optimal transport for $d$-partite systems for $d \geq 3$, denoted as $T_{C}^{Q}(\rho^{A_{1}}, \ldots, \rho^{A_{d}})$. In Appendix A we briefly review the basic properties of partial traces. In Appendix B we discuss additional properties of the QOT for qubits. Appendix C gives a closed formula for the QOT for some pairs of diagonal qutrits, i.e., classical states of a three-level quantum system.

2 Preliminary Results

The aim of this section is fivefold. First, we discuss briefly our notation. Second, Proposition 2.1 shows that the coupling set $\Gamma_{Q}^{Q}(\rho^{A}, \rho^{B})$ contains a rank one matrix if and only if $\rho^{A}$ and $\rho^{B}$ are isospectral. Third, we discuss some basic properties of $T_{C}^{Q}(\rho^{A}, \rho^{B})$. Fourth, we introduce the SWAP operator, and the corresponding cost matrices $C_{Q}^{Q}, C_{E}^{Q}$, which are positive semidefinite and vanish on the set of symmetric matrices, the two-qubit bosons. Fifth, we discuss SWAP fidelity and related quantities.
2.1 Notations

In what follows we combine the standard mathematical notation with the Dirac notation used in quantum theory. We view \( \mathbb{C}^n \), the vector space of column vectors over the complex field \( \mathbb{C} \), as a Hilbert space \( \mathcal{H}_n \) with the inner product

\[
\langle y, x \rangle = y^\dagger x = \langle y|x \rangle.
\]

Then \( |i\rangle \in \mathcal{H}_n \) is identified with the unit vector \( e_i = (\delta_{ij}, \ldots, \delta_{nj})^T \) for \( i \in [n] \). Let \( B(\mathcal{H}_n) \supset S(\mathcal{H}_n) \supset S_+(\mathcal{H}_n) \supset \Omega_n \) be the space of linear operators, the real subspace of selfadjoint operators, the cone of positive semidefinite operators, and the convex set of density operators, respectively. For \( \rho \in B(\mathcal{H}_n) \) we denote \( |\rho| = \sqrt{\rho \rho^\dagger} \in S_+(\mathcal{H}_n) \). Then \( \|\rho\|_1 = \text{Tr} |\rho| \). For \( \rho, \sigma \in S(\mathcal{H}_n) \) we write \( \rho \geq \sigma \) and \( \rho > \sigma \) if if the eigenvalues of \( \rho - \sigma \) are all nonnegative or positive respectively.

The space of \( n \times n \) complex valued matrices, denoted as \( \mathbb{C}^{n \times n} \), is a representation of \( B(\mathcal{H}_n) \), where the matrix \( \rho = [\rho_{ij}] \in \mathbb{C}^{n \times n} \) represents the operator \( \rho \in B(\mathcal{H}_n) \). The set of density operators in \( B(\mathcal{H}_n) \) are viewed as \( \Omega_n \): the convex set of \( n \times n \) Hermitian positive semidefinite trace-one matrices. The tensor product \( \mathcal{H}_m \otimes \mathcal{H}_n \) is represented by \( \mathbb{C}^{m \times n} \). An element of \( \mathbb{C}^{m \times n} \) is a matrix \( X = [x_{ip}] = \sum_{i=1}^m \sum_{p=1}^n x_{ip}|i \rangle \langle p| \), which correspond to a bipartite state. Observe that \( x \otimes y = |x\rangle |y\rangle \) is represented by the rank-one matrix \( xy^\dagger \). We denote by \( X^\dagger = \{X| \) the complex conjugate of the transpose of \( X \in \mathbb{C}^{m \times n} \). The inner product of bipartite states \( X, Y \in \mathbb{C}^{m \times n} \) is \( \langle X, Y \rangle = \langle X|Y\rangle = \text{Tr} X^\dagger Y \). We identify \( B(\mathcal{H}_m \otimes \mathcal{H}_n) \) with \( \mathbb{C}^{(mn) \times (mn)} \) as follows.

An operator \( \rho^{AB} \in B(\mathcal{H}_m \otimes \mathcal{H}_n) \) is represented by a matrix \( R \in \mathbb{C}^{(mn) \times (mn)} \), whose entries are indexed with two pairs of indices \( r(i,p)(j,q) \) where \( i, j \in [m], \ p, q \in [n] \). Then the partial traces of \( R \) are defined as follows:

\[
\text{Tr}_A R = \left[ \sum_{i=1}^m r(i,p)(i,q) \right] = \rho^B \in \mathbb{C}^{m \times n},
\]

\[
\text{Tr}_B R = \left[ \sum_{p=1}^n r(i,p)(j,p) \right] = \rho^A \in \mathbb{C}^{m \times m}.
\]  

(2.1)

Recall that \( \text{Tr} R = \text{Tr}(\text{Tr}_A R) = \text{Tr}(\text{Tr}_B R) \). Some more known facts about partial traces that we use in this paper are discussed in Appendix A.

Let \( M : B(\mathcal{H}_m \otimes \mathcal{H}_n) \rightarrow B(\mathcal{H}_m) \oplus B(\mathcal{H}_n) \) be the partial trace map: \( \rho^{AB} \mapsto (\rho^A, \rho^B) \). We identify \( M \) with the map \( M : \mathbb{C}^{(mn) \times (mn)} \rightarrow \mathbb{C}^{m \times m} \oplus \mathbb{C}^{n \times n} \). For \( \rho^A \in \Omega_m, \rho^B \in \Omega_n \) we denote by \( \Gamma^Q(\rho^A, \rho^B) \) the set of all quantum coupling matrices—bipartite density matrices \( \rho^{AB} \) whose partial traces are \( \rho^A \) and \( \rho^B \) respectively.

\[
\Gamma^Q(\rho^A, \rho^B) = \{\rho^{AB} \in \Omega_{mn}, \text{Tr}_B \rho^{AB} = \rho^A, \text{Tr}_A \rho^{AB} = \rho^B\}.
\]

Then \( \Omega_{mn} \) fibers over \( \Omega_m \times \Omega_n \), that is, \( \Omega_{mn} = \bigcup_{(\rho^A, \rho^B) \in \Omega_m \times \Omega_n} \Gamma^Q(\rho^A, \rho^B) \). The Hausdorff distance between \( \Gamma^Q(\rho^A, \rho^B) \) and \( \Gamma^Q(\rho^C, \rho^D) \) is a complete distance on the fibers \([29]\).
2.2 Isospectral Density Matrices

We identify $\mathcal{H}_n \otimes \mathcal{H}_n$ as the space of $n \times n$ complex valued matrices $\mathbb{C}^{n \times n}$ as follows. Let $e_i = (\delta_{i1}, \ldots, \delta_{in})^\top \equiv |i\rangle$, $i \in [n]$ be the standard basis in $\mathbb{C}^n \equiv \mathcal{H}_n$. Then a state $|\psi\rangle \in \mathcal{H}_n \otimes \mathcal{H}_n$ is given by $|\psi\rangle = \sum_{i,j=1}^n x_{ij} |i\rangle |j\rangle$. Thus we associate with $|\psi\rangle$ the matrix $X = [x_{ij}] \in \mathbb{C}^{n \times n}$. Then $|\psi\rangle$ is a normalized state if and only if $\|X\|_2^2 = \text{Tr} XX^\dagger = 1$. Suppose we change the orthonormal basis $e_1, \ldots, e_n$ to an orthonormal basis $f_1, \ldots, f_n$, where $e_i = \sum_{p=1}^n u_{pi} f_p$. Here $U = [u_{ip}] \in \mathbb{C}^{n \times n}$ is a unitary matrix. Then $|\psi\rangle = \sum_{p,q=1}^n y_{pq} \langle f_p | f_q \rangle$, where $Y = UXU^\top$.

We now consider a pure state density operator $|\psi\rangle \langle \psi| = (\sum_{i,j=1}^n x_{ij} |i\rangle |j\rangle)(\sum_{p,q=1}^n \bar{x}_{pq} \langle p | \langle q|) = \sum_{i,j,p,q=1}^n x_{ij} \bar{x}_{pq} |i\rangle \langle j| \langle p | \langle q|$.

We identify the coefficient matrix with the Kronecker product $X \otimes \bar{X}$. Then

$$
\rho^A = \text{Tr}_B |\psi\rangle \langle \psi| = \sum_{i,p=1}^n (XX^\dagger)_{ip} |i\rangle \langle p|,
$$

$$
\rho^B = \text{Tr}_A |\psi\rangle \langle \psi| = \sum_{j,q=1}^n (X^\dagger \bar{X})_{jq} |j\rangle \langle q|.
$$

Thus in the standard basis of $\mathcal{H}_n$ we can identify $\rho^A$ and $\rho^B$ with the density matrices

$$
\rho^A = XX^\dagger, \quad \rho^B = X^\dagger \bar{X}. \quad (2.2)
$$

If we change from the standard basis to the basis $f_1, \ldots, f_n$, using a unitary matrix $U$, the both partial traces read,

$$
\tilde{\rho}^A = \bar{X} \bar{X}^\dagger = U(XX^\dagger)U^\dagger = U \rho^A U^\dagger,
$$

$$
\tilde{\rho}^B = X^\dagger \bar{X} = U(X^\dagger \bar{X})U^\dagger = U \rho^B U^\dagger. \quad (2.3)
$$

Note that if $\nu_1 \geq \cdots \geq \nu_n \geq 0$ are the singular values of the matrix $X$ then $\lambda_1 = \nu_1^2 \geq \cdots \geq \lambda_n = \nu_n^2 \geq 0$ are the eigenvalues of $\rho^A$ and $\rho^B$. That is $\rho^A$ and $\rho^B$ are isospectral. Vice versa:

**Proposition 2.1** Let $\rho^A, \rho^B \in \Omega_n$. Then $\Gamma^Q(\rho^A, \rho^B)$ contains a matrix $R$ of rank one if and only if $\rho^A$ and $\rho^B$ are isospectral.
Proof Suppose first that $\rho^A$ and $\rho^B$ are isospectral, i.e., have the same eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$. Assume that $\rho^A$ and $\rho^B$ have the following spectral decompositions:

\[
\rho^A = \sum_{i=1}^{n} \lambda_i |x_i\rangle \langle x_i|, \quad \langle x_i, x_j \rangle = \delta_{ij},
\]

\[
\rho^B = \sum_{j=1}^{n} \lambda_j |y_j\rangle \langle y_j|, \quad \langle y_i, y_j \rangle = \delta_{ij}.
\]

Then the set $\Gamma^Q(\rho^A, \rho^B)$ of coupling matrices contains the rank-one matrix

\[
R = \left( \sum_{i=1}^{n} \sqrt{\lambda_i} |x_i\rangle \right) \left( \sum_{j=1}^{n} \sqrt{\lambda_j} |y_j\rangle \right).
\]

Vice versa, if $R$ represents a projector onto a pure bipartite state $|\psi\rangle$ in $S_+(H_n \otimes H_n)$, then its Schmidt decomposition, related to the Singular Value Decomposition (SVD) of the above matrix $X = [x_{ij}]$, takes the above form [25]. Hence $\text{Tr}_A R$ and $\text{Tr}_B R$ are isospectral density matrices.

2.3 Some Properties of $T^Q_C(\rho^A, \rho^B)$

We first recall the definition of the Hausdorff distance between non-empty sets in a space $X$ equipped with a distance $D : X \times X \to [0, \infty)$. Assume that $Y, Z \subset X$. Then the Hausdorff distance between $Y$ and $Z$ is:

\[
D_H(Y, Z) = \max \left( \sup_{y \in Y} \inf_{z \in Z} D(y, z), \sup_{z \in Z} \inf_{y \in Y} D(y, z) \right).
\]

We will need the Hausdorff distance in the proof of the following result.

Proposition 2.2 For $C \in S(H_m \otimes H_n)$ the function $T^Q_C(\cdot, \cdot)$ is a continuous convex function on $\Omega_m \times \Omega_n$: for any $0 < a < 1$,

\[
T^Q_C(a\rho^A + (1-a)\sigma^A, a\rho^B + (1-a)\sigma^B) \leq aT^Q_C(\rho^A, \rho^B) + (1-a)T^Q_C(\sigma^A, \sigma^B).
\]

Furthermore, if $C \geq 0$ then $T^Q_C(\cdot, \cdot)$ is nonnegative.

Proof Assume that

\[
T^Q_C(\rho^A, \rho^B) = \text{Tr} C \rho^{AB}, \quad \rho^{AB} \in \Gamma^Q(\rho^A, \rho^B),
\]

\[
T^Q_C(\sigma^A, \sigma^B) = \text{Tr} C \sigma^{AB}, \quad \sigma^{AB} \in \Gamma^Q(\sigma^A, \sigma^B).
\]

Let $\tau^{AB} = a\rho^{AB} + (1-a)\sigma^{AB}$. Then $\tau^{AB} \in \Gamma^Q(a\rho^A + (1-a)\sigma^A, a\rho^B + (1-a)\sigma^B)$. Clearly $\text{Tr} C \tau^{AB} = aT^Q_C(\rho^A, \rho^B) + (1-a)T^Q_C(\sigma^A, \sigma^B)$. The minimal characteriza-
tion (1.4) of $T$ yields the first inequality of the lemma. Clearly if $C \geq 0$ then $T^Q_C (\cdot, \cdot)$ is nonnegative. This yields the second inequality of the lemma.

The continuity of $T^Q_C (\cdot, \cdot)$ follows from the following arguments. Assume that

$$
\rho^A = \lim_{k \to \infty} \rho^A_k, \quad \rho^B = \lim_{k \to \infty} \rho^B_k,
$$

$$
T^Q_C (\rho^A_k, \rho^B_k) = \text{Tr} \ C \rho^A_k \rho^B_k, \quad \rho^A_k, \rho^B_k \in \Gamma^Q (\rho^A, \rho^B), \quad k \in \mathbb{N}.
$$

As the set of density matrices of a fixed dimension is a compact set, there exists a subsequence $\{k_l, k \in \mathbb{N}\}$ such that $\rho^A_{k_l}$ converges to a density matrix $\rho$. Hence

$$
\text{Tr}_A \rho = \lim_{k \to \infty} \text{Tr}_A \rho^A_{k_l} = \lim_{k \to \infty} \rho^B_{k_l} = \rho^B,
$$

$$
\text{Tr}_B \rho = \lim_{k \to \infty} \text{Tr}_B \rho^A_{k_l} = \lim_{k \to \infty} \rho^A_{k_l} = \rho^A.
$$

Hence, $\rho \in \Gamma^Q (\rho^A, \rho^B)$. Therefore,

$$
\lim_{k \to \infty} T^Q_C (\rho^A_{k_l}, \rho^B_{k_l}) = \lim_{k \to \infty} \text{Tr} \ C \rho^A_{k_l} = \text{Tr} \ C \rho \geq T^Q_C (\rho^A, \rho^B) = \text{Tr} \ C \rho^A.
$$

Observe that for each $\rho^A \in \Omega_m, \rho^B \in \Omega_n$, the set $\Gamma^Q (\rho^A, \rho^B)$, viewed as a fiber over $(\rho^A, \rho^B)$, is a compact convex set. Hence, one can define the Hausdorff distance (distance) on the fibers. It is shown in [29, Theorem 5.2] that the Hausdorff distance is a complete distance. Furthermore, the sequence $\Gamma^Q (\rho^A_k, \rho^B_k), k \in \mathbb{N}$ converges to $\Gamma^Q (\rho^A, \rho^B)$ in the Hausdorff distance if and only if $\lim_{k \to \infty} (\rho^A_k, \rho^B_k) = (\rho^A, \rho^B)$. Hence, there exists a sequence $\omega_k \in \Gamma^Q (\rho^A_{k_l}, \rho^B_{k_l})$ such that $\lim_{k \to \infty} \|\omega_k - \rho^A\| = 0$.

(Here we let $\|\eta\| = \sqrt{\text{Tr} \eta^2}$ for any Hermitian operator on $\mathcal{H}_N$.) Clearly,

$$
T^Q_C (\rho^A_{k_l}, \rho^B_{k_l}) \leq \text{Tr} \ C \omega_k = \text{Tr} \ C \rho^A + \text{Tr} \ C (\omega_k - \rho^A) = T^Q_C (\rho^A, \rho^B) + \text{Tr} \ C (\omega_k - \rho^A).
$$

Let $k \to \infty$ to deduce the inequality $\lim_{k \to \infty} T^Q_C (\rho^A_{k_l}, \rho^B_{k_l}) \leq T^Q_C (\rho^A, \rho^B)$. Combine that with (2.6) to deduce the equality $\lim_{k \to \infty} T^Q_C (\rho^A_{k_l}, \rho^B_{k_l}) = T^Q_C (\rho^A, \rho^B)$. Clearly, the sequence $\{T^Q_C (\rho^A_{k_l}, \rho^B_{k_l}), k \in \mathbb{N}\}$ is a bounded sequence. We showed that from this sequence we can always extract a subsequence which converges to $T^Q_C (\rho^A, \rho^B)$. Hence, this sequence converges to $T^Q_C (\rho^A, \rho^B)$.

The following Proposition shows that to compute $T^Q_C (\rho^A, \rho^B)$ one can assume that the eigenvalues of $C$ are in the interval $[0, 1]$:

**Proposition 2.3** Assume that $C \in \mathcal{S}(\mathcal{H}_m \otimes \mathcal{H}_n)$ is not a scalar operator, $C \not\in \mathcal{C}$. Let

$$
\tilde{C} = \frac{1}{\lambda_{\text{max}} (C) - \lambda_{\text{min}} (C)} (C - \lambda_{\text{min}} (C) \mathbb{I}).
$$
Then $0 \leq \tilde{C} \leq I$. Furthermore for $\rho^A \in \Omega_m$, $\rho^B \in \Omega_n$ the following equality holds:

$$T_C^Q(\rho^A, \rho^B) = (\lambda_{\max}(C) - \lambda_{\min}(C))T_C^Q(\rho^A, \rho^B) + \lambda_{\min}(C). \quad (2.7)$$

**Proof** Clearly $C = (\lambda_{\max}(C) - \lambda_{\min}(C))\tilde{C} + \lambda_{\min}(C)I$. Furthermore

$$\text{Tr} C \rho^{AB} = (\lambda_{\max}(C) - \lambda_{\min}(C)) \text{Tr} \tilde{C} \rho^{AB} + \lambda_{\min}(C), \quad \rho^{AB} \in \Gamma^Q(\rho^A, \rho^B).$$

As $\lambda_{\max}(C) - \lambda_{\min}(C) > 0$ we deduce (2.7). \hfill $\square$

We next observe that one can reduce the computation of $T_C^Q(\rho^A, \rho^B)$ to a smaller dimension problem if either $\rho^A$ or $\rho^B$ are not positive definite:

**Proposition 2.4** Assume that $\rho^A \in \Omega_m$, $\rho^B \in \Omega_n$. Let $m'$ and $n'$ be the dimensions of range $\rho^A = H_m$ and range $\rho^B = H_n$ respectively. Denote by $\rho^{A'} \in \Omega_{m'}$, and $\rho^{B'} \in \Omega_{n'}$ the restrictions of $\rho^A$ and $\rho^B$ to $H_m$ and $H_n$ respectively. Assume that $C \in S(H_m \otimes H_n)$, and denote by $C' \in S(H_{m'} \otimes H_{n'})$ the restriction of $C$ to $H_m \otimes H_n$.

Then

$$T_C^Q(\rho^A, \rho^B) = T_{C'}^Q(\rho^{A'}, \rho^{B'}).$$

**Proof** Without loss of generality we can assume that we chose orthonormal bases in $H_m$ and $H_n$ to be the eigenvectors of $\rho^A$ and $\rho^B$ respectively. Thus to prove the lemma it is enough to consider the following case: $\rho^A = \rho^C \oplus 0_{m-l}$ where $\rho^C \in \Omega_{l}$, $l < m$ and $0_l$ is an $l \times l$ zero matrix. Let $\tilde{C} \in S(H_{l} \otimes H_{n})$ be the restriction of $C$ to $H_l \otimes H_n$.

We claim that

$$T_C^Q(\rho^A, \rho^B) = T_{\tilde{C}}^Q(\rho^C, \rho^B). \quad (2.8)$$

Let $R = [R_{ij}(j,q)] \in \Gamma^Q(\rho^A, \rho^B)$. As $R \geq 0$ it follows that the submatrix $R_{ii} = [R_{ij}(j,q)]$, $p, q \in [n]$ is positive semidefinite for each $i \in [m]$. Since $\text{Tr} R = \rho^A$ we deduce that $\rho^A_{ii} = \sum_{p,q \in [n]} R_{ij}(j,q) = \text{Tr} R_{ii} = 0$ for $i > l$. Therefore $R_{ii} = 0$, that is, $R_{ij}(j,q) = 0$ for $p, q \in [n]$ and $i > l$. Let $R'$ be the following submatrix of $R$: $[R_{ij}(j,q)]$, $i, j \in [l]$, $p, q \in [n]$. Then $R' \in \Gamma^Q(\rho^C, \rho^B)$. Vice versa, given $R' \in \Gamma^Q(\rho^C, \rho^B)$, one can enlarge trivially $R'$ to $R$ in $\Gamma^Q(\rho^C, \rho^B)$. Clearly $\text{Tr} CR = \text{Tr} \tilde{C} R'$. Repeating the same process with $\rho^B$ establishes (2.8). \hfill $\square$

As we point out in the next subsection, it is natural to consider the case $m = n$. However, if either $\rho^A$ or $\rho^B$ are singular density matrices then we can reduce the computation of $T_C^Q(\rho^A, \rho^B)$ to a lower-dimensional problem, and after this reduction it may happen that the dimensions are no longer equal.

**2.4 Quantum Transport Problem Induced by the SWAP Operator**

When describing any two distinguishable physical objects one can introduce an operation $S$ which exchanges them. On the composite space $H_n \otimes H_n$ it corresponds to
a natural isometry induced by swapping the two factors $x \otimes y \mapsto y \otimes x$. On the space of square matrices the SWAP operator is the map $S : X \mapsto X^\top$. This map is of fundamental importance in quantum information theory. It allows to observe some interesting properties of bipartite system and is useful in the criterion for separability by Peres and Horodecki [36, 50]. It is shown below that the SWAP operator $S$ induces a cost matrix

$$C^Q = \frac{1}{2}(I - S),$$

(2.9)

for the quantum transport problem, which enjoys several nice properties.

For $\mathcal{H}_n \otimes \mathcal{H}_n$ the SWAP operation $S \in B(\mathcal{H}_n \otimes \mathcal{H}_n)$ acts on the product states as follows: $S(|x\rangle|u\rangle) = |u\rangle|x\rangle$. So $S$ is both unitary and an involution operator: $S^\dagger S = I$ and $S^2 = I$. Hence the eigenvalues of $S$ are $\pm 1$ and $S$ is selfadjoint, $S^\dagger = S$. The invariant subspaces of $S$ corresponding to the eigenvalues $1$ and $-1$ are the symmetric and skew-symmetric tensors respectively, which can be identified with the symmetric $\mathcal{H}_S = \mathcal{S}^n$ and skew-symmetric $\mathcal{H}_A = \mathcal{A}^n$ matrices in $\mathbb{C}^{n \times n}$, respectively. Note that the decomposition of a matrix $X$ into a sum of symmetric and skew-symmetric matrices $X = (1/2)(X + X^\top) + (1/2)(X - X^\top)$ is an orthogonal decomposition. That is

$$\mathcal{H}_n \otimes \mathcal{H}_n = \mathcal{H}_S \oplus \mathcal{H}_A = \mathbb{C}^{n \times n} = \mathcal{S}^n \oplus \mathcal{A}^n$$

is an orthogonal decomposition. Observe that $S(X) = X^\top$. Hence the action of $S$ on a rank-one operator $|X\rangle\langle Y|$ in $B(\mathcal{H}_n \otimes \mathcal{H}_n)$ is $S(|X\rangle\langle Y|) = |X^\top\rangle\langle Y|$. Therefore the action of $S$ on rank one product operator in $B(\mathcal{H}_n \otimes \mathcal{H}_n)$ is given by

$$S(|x\rangle|u\rangle\langle y|\langle v|) = S(|x\rangle|u\rangle)\langle y|\langle v| = |u\rangle|x\rangle\langle y|\langle v|.$$  

Hence

$$\text{Tr} \ S(|x\rangle|u\rangle\langle y|\langle v|) = (\langle y|\langle v|)(|u\rangle|x\rangle)) = \langle y|\langle v|$$

Similarly

$S(|x\rangle|u\rangle\langle y|\langle v|)S^\dagger = |u\rangle|x\rangle\langle y|\langle v|.$

Use the identity (A.1) and the above results to deduce that

$$\text{Tr} \ S(|x\rangle|u\rangle\langle y|\langle v|) = \langle y|\langle v| = \text{Tr} \ S(|u\rangle|x\rangle\langle y|\langle v|),$$

$$\text{Tr}_A \ S(|x\rangle|u\rangle\langle y|\langle v|)S^\dagger = \langle y|\langle v| = \text{Tr}_B \ S(|x\rangle|u\rangle\langle y|\langle v|),$$

$$\text{Tr}_B \ S(|x\rangle|u\rangle\langle y|\langle v|)S^\dagger = \langle y|\langle v| = \text{Tr}_A \ S(|x\rangle|u\rangle\langle y|\langle v|).$$

Use (A.1) to deduce

$$S(|x\rangle \otimes |u\rangle)\langle y|\langle v| = |u\rangle|x\rangle\langle y|\langle v| = (|u\rangle\langle y|) \otimes (|x\rangle\langle v|).$$
Combine the above equalities to obtain the following identities:

\[\text{Tr } S(\rho^A \otimes \rho^B) = \text{Tr } \rho^A \rho^B, \quad \rho^A, \rho^B \in B(H_n),\]

\[\text{Tr}_A S \rho^{AB} S^\dagger = \text{Tr}_B \rho^{AB}, \quad \text{Tr}_B S \rho^{AB} S^\dagger = \text{Tr}_A \rho^{AB}, \quad \rho^{AB} \in B(H_n \otimes H_n).\] (2.10)

The first identity is due to Werner [62], see also [45].

Denote by \(\ker C\) the kernel of a linear operator \(C : H_n \otimes H_n \to H_n \otimes H_n\). An operator \(C\) is said to vanish exactly on symmetric matrices if \(\ker C = H_S\). Thus a positive semidefinite \(C\) vanishes exactly on \(H_S\) if and only if it has \(n(n - 1)/2\) positive eigenvalues (counting with multiplicities) with the corresponding skew-symmetric eigenvectors.

Let \(|\psi_{ij}\rangle\) be an orthonormal basis in \(H_n\). Define (as in [28]) the maximally entangled singlet states spanned on two dimensional subspaces:

\[|\psi_{ij}\rangle = \frac{1}{\sqrt{2}}(|i\rangle|j\rangle - |j\rangle|i\rangle) \text{ for } 1 \leq i < j \leq n.\] (2.11)

Given a classical distance matrix \(E = [e_{ij}]_{i,j=1}^{n}\) with \(e_{ij} > 0\) for all \(1 \leq i < j \leq n\), the following operator is positive semidefinite and vanishes exactly on the symmetric subspace, \(S^2 \mathbb{C}^n\) [28, (11)]:

\[C_Q^E = \sum_{1 \leq i < j \leq n} e_{ij} |\psi_{ij}\rangle \langle \psi_{ij}|.\] (2.12)

Consider the operator \(C_Q^E\) given by (2.9). Then \(C_Q^E\) is an orthogonal projection of \(\mathbb{C}^{n \times n}\) onto antisymmetric subspace, \(A^2 \mathbb{C}^n\). Hence \(C_Q^E\) is of the form (2.12), where \(e_{ij} = 1\) for all \(i < j\), so such a distance matrix \(E\) represents the simplex configuration. Denote by \(U(n) \subset \mathbb{C}^{n \times n}\) the group of unitary matrices. The following lemma shows that \(T_Q C_Q^E\) is invariant under conjugation by a unitary matrix:

**Proposition 2.5** Assume that \(\rho^A, \rho^B \in \Omega_n\) and \(\rho^{AB} \in \Gamma_Q(\rho^A, \rho^B)\). Then for \(U \in U(n)\) the following equalities hold:

\[\text{Tr}_B((U \otimes U)\rho^{AB}(U^\dagger \otimes U^\dagger)) = U \rho^A U^\dagger,\]
\[\text{Tr}_A((U \otimes U)\rho^{AB}(U^\dagger \otimes U^\dagger)) = U \rho^B U^\dagger,\]
\[(U \otimes U)\Gamma_Q(\rho^A, \rho^B)(U^\dagger \otimes U^\dagger) = \Gamma_Q(U \rho^A U^\dagger, U \rho^B U^\dagger),\]
\[T_Q^C(\rho^A, \rho^B) = T_Q^C((U \otimes U)\rho^A U^\dagger, U \rho^B U^\dagger).\] (2.13)

In particular, the optimal transport cost for \(C\) given by (2.9) is unitarily invariant,

\[T_Q^C(\rho^A, \rho^B) = T_Q^C(U \rho^A U^\dagger, U \rho^B U^\dagger).\] (2.14)
Proof Assume that $R$ is a pure state $R = |\psi\rangle\langle\psi|$. The state $|\psi\rangle$ corresponds to a matrix $X \in \mathbb{C}^{n \times n}$ with $\text{Tr} XX^\dagger = 1$. Then $\text{Tr}_B R = XX^\dagger$ and $\text{Tr}_A R = X^\dagger \hat{X}$. Recall that $(U \otimes U)|\psi\rangle$ is represented by $\hat{X} = UXU^\dagger$. Now use (2.3) to deduce the first two equalities in (2.13) if $R \in \Gamma^Q(\rho^A, \rho^B)$. Recall that any $\rho^{AB} \in \Gamma^Q(\rho^A, \rho^B)$ is a convex combination of pure states $R_i = |\psi_i\rangle\langle\psi_i|$, $i \in [k]$. That is $R = \sum_{i=1}^k a_i R_i$, where $a_i > 0$ and $\sum_{i=1}^k a_i = 1$. Then $\text{Tr}_B R_i = \rho^A_i$, $\text{Tr}_A R_i = \rho^B_i$. Now use the above results for $R_i$ to deduce the first two equalities in (2.13), which imply two following equalities. Equality (2.14) is deduced from the fact

$$\textbf{(U} \otimes \textbf{U}) \textbf{C}_Q(\textbf{U}^\dagger \otimes \textbf{U}^\dagger) = \textbf{C}_Q.$$  \hfill (2.15)

\hfill $\Box$

2.5 SWAP Fidelity and Related Quantities

Recall the definitions of fidelity, Bures distance and root infidelity [4]. For $\rho^A, \rho^B \in \Omega_n$ these are defined, respectively, as

$$F(\rho^A, \rho^B) = \left(\text{Tr} \sqrt{\sqrt{\rho^A} \sqrt{\rho^B}}\right)^2,$$  \hfill (2.16)

$$B(\rho^A, \rho^B) = \sqrt{2} \sqrt{1 - F(\rho^A, \rho^B)},$$  \hfill (2.17)

$$I(\rho^A, \rho^B) = \sqrt{1 - F(\rho^A, \rho^B)}.$$  \hfill (2.18)

It is known that the root infidelity [31] and Bures distance lead to metrics on $\Omega_n$. As $0 \leq F(\rho^A, \rho^B) \leq 1$ it follows that

$$B(\rho^A, \rho^B) \leq \sqrt{2} I(\rho^A, \rho^B).$$

We now describe basic properties of the SWAP-fidelity $F_S(\rho^A, \rho^B)$ given by (1.7):

Proposition 2.6 Assume that $\rho^A, \rho^B \in \Omega_n$ and $\rho^{AB} \in \Gamma^Q(\rho^A, \rho^B)$. Then the SWAP-fidelity function has the following properties:

(a) One has the equality

$$F_S(\rho^A, \rho^B) = 1 - 2 \text{T}_C^Q(\rho^A, \rho^B).$$  \hfill (2.19)

(b) The function $F_S$ is a symmetric concave function on $\Omega_n \times \Omega_n$ with values in the interval $[0, 1]$.

(c) $F_S(\rho^A, \rho^B) = 1$ if and only if $\rho^A = \rho^B$.

(d) $F_S(\rho^A, \rho^B) = 0$ if and only if $\text{Tr} \rho^A \rho^B = 0$. 

Springer
(e) The following equalities and inequalities hold
\[
F_S(\rho^A, \rho^B) = F_S(U \rho^A U^\dagger, U \rho^B U^\dagger), \quad \text{for } U \in U(n),
\]
\[
F(\rho^A, \rho^B) \leq F_S(\rho^A, \rho^B) \leq \sqrt{F(\rho^A, \rho^B)},
\]
\[
F_S(\rho^A \otimes \sigma^C, \rho^B \otimes \sigma^D) \geq F_S(\rho^A, \rho^B) F_S(\sigma^C, \sigma^D), \quad \text{for } \sigma^C, \sigma^D \in \Omega_m.
\]

(f) \(F_S(\rho^A, \rho^B) = F(\rho^A, \rho^B)\) if either \(\rho^A\) or \(\rho^B\) is a pure state.

**Proof**

(a) Observe that \(\text{Tr } S\rho^{AB} = 1 - 2 \text{Tr } ((1/2)(I - S)\rho^{AB})\). Then the maximal characterization of \(F_S(\rho^A, \rho^B)\) in (1.7) and the minimum characterization of \(T_{\Omega}^Q(\rho^A, \rho^B)\) yield equality (2.19).

(b) Proposition 2.2 and equality (2.19) yield that \(F_S\) is a concave function on \(\Omega_n \times \Omega_n\) whose value is at most 1. Part (d) of Theorem 6.3 yields that \(T_{\Omega}^Q(\rho^A, \rho^B) \leq 1/2\). Equality (2.19) yields that \(F_S \geq 0\).

(c) Clearly, \(F_S(\rho^A, \rho^B) = 1 \iff T_{\Omega}^Q(\rho^A, \rho^B) = 0\). Part (c) of Theorem 6.3 yields that this happens if and only if \(\rho^A = \rho^B\).

(e) Equation (2.20) follows from the equalities (2.19) and (2.14). Inequality (2.21) is stated and proved in Theorem 10 in [65]. Finally, inequality (2.22) follows from Lemma A.4 and the following observations. Let \(S_n\) and \(S_{m,n}\) be the SWAP operators on \(\mathcal{H}_n \otimes \mathcal{H}_n\) and \(\mathcal{H}_{nm} := (\mathcal{H}_n \otimes \mathcal{H}_m) \otimes (\mathcal{H}_n \otimes \mathcal{H}_m)\), respectively. Denote by \(R_{n,m} : \mathcal{H}_n \otimes \mathcal{H}_n \otimes \mathcal{H}_m \otimes \mathcal{H}_m \rightarrow \mathcal{H}_{nm}\) the SWAP of the two middle factors. Assume that
\[
F_S(\rho^A, \rho^B) = \text{Tr } S_n \rho^{AB}, \quad F_S(\sigma^C, \sigma^D) = \text{Tr } S_m \sigma^{CD}.
\]
Let \(\tau^{ACBD} := R_{n,m}(\rho^{AB} \otimes \sigma^{CD})R_{n,m} \in \Omega_{nm}\). The first line of (A.2) yields \(\tau^{ACBD} \in \Gamma_Q(\rho^A \otimes \sigma^C, \rho^B \otimes \sigma^D)\). The second line of (A.2) yields
\[
\text{Tr } S_{n,m} \tau^{ACBD} = (\text{Tr } S_n \rho^{AB})(\text{Tr } S_m \sigma^{CD}) = F_S(\rho^A, \rho^B) F_S(\sigma^A, \sigma^B).
\]

The maximum characterization of \(F_S(\rho^A \otimes \sigma^C, \rho^B \otimes \sigma^D)\) yields inequality (2.22).

(d) Recall that \(\text{Tr } \rho^A \rho^B = 0\) if and only if the eigenvectors of \(\rho^A\) and \(\rho^B\) corresponding to positive eigenvalues are orthogonal. This is equivalent to \(F(\rho^A, \rho^B) = 0\). Use inequality (2.21) to deduce part (d).

(f) Assume that \(\rho^A\) is a pure state. Then \(\sqrt{\rho^A} = \rho^A\). Hence,
\[
|\sqrt{\rho^A} \sqrt{\rho^B}|^2 = \sqrt{\rho^A \rho^B} \sqrt{\rho^A} = \rho^A \rho^B \rho^A = (\text{Tr } \rho^A \rho^B) \rho^A
\]
\[
\Rightarrow |\sqrt{\rho^A} \sqrt{\rho^B}| = \sqrt{\text{Tr } \rho^A \rho^B} \rho^A \Rightarrow F_S(\rho^A, \rho^B) = \text{Tr } \rho^A \rho^B.
\]

Use (2.19) and (6.2) to deduce \(F_S(\rho^A, \rho^B) = F(\rho^A, \rho^B)\). \(\square\)

Recall the definition of \(\mathcal{W}_{\Omega}^Q\) (1.6).
Corollary 2.7 Let $\rho^A, \rho^B \in \Omega_n$. Then the following inequalities hold:

\[
\frac{1}{2} \left( 1 - \sqrt{F(\rho^A, \rho^B)} \right) \leq T_{C_0}^Q(\rho^A, \rho^B) \leq \frac{1}{2} \left( 1 - \frac{F(\rho^A, \rho^B)}{2} \right),
\]

\[
\frac{1}{2} B(\rho^A, \rho^B) \leq \mathcal{W}_{C_0}^Q(\rho^A, \rho^B) \leq \sqrt{T_{C_0}^Q(\rho^A, \rho^B)} \leq \frac{1}{\sqrt{2}} I(\rho^A, \rho^B).
\]

Moreover, if either of the states is pure, then $T_{C_0}^Q(\rho^A, \rho^B) = (1 - F(\rho^A, \rho^B))/2$.

Proof The bound (2.23) follows directly from inequalities (2.21). The third inequality in (2.24) is an immediate consequence of (2.23), the first and second one follow from Theorem 6.1. The last statement is a consequence of equality (6.2), which holds if either of the states is pure, and point (f) of Proposition 2.6.

Proposition 2.6 shows that the introduced SWAP-fidelity shares many features with the standard quantum fidelity introduced by Uhlmann [58] and Jozsa [38]. Notably, $F_S$ is jointly concave, as is the root fidelity, $\sqrt{F}$, but not the fidelity itself [4]. It is also remarkable that the SWAP-fidelity is supermultiplicative with respect to the tensor product, as is the superfidelity introduced in [45]. This feature might prove relevant for applications in quantum machine learning—see [28] and references therein.

3 Quantum Optimal Transport as a Semidefinite Programming Problem

One of the main results of this paper is the observation that the computation of the quantum transport is carried out efficiently using semidefinite programming (SDP) [59]. The main results of this section are the statement of QOT as the direct and the dual semidefinite programs:

Theorem 3.1 Assume that $C \in S(\mathcal{H}_m \otimes \mathcal{H}_n)$, $\rho^A \in \Omega_m$, $\rho^B \in \Omega_n$. Then the computation of $T_{C}^Q(\rho^A, \rho^B)$ is a semidefinite programming problem. The value of $T_{C}^Q(\rho^A, \rho^B)$ can be approximated within precision $\varepsilon > 0$ in polynomial time in the size of the data and log $1/\varepsilon$.

Theorem 3.2 Assume that $\rho^A \in \Omega_m$, $\rho^B \in \Omega_n$ and $C \in S(\mathcal{H}_m \otimes \mathcal{H}_n)$. Then the dual problem to (1.4) is

\[
\sup \{ \operatorname{Tr} \sigma^A \rho^A + \operatorname{Tr} \sigma^B \rho^B, \sigma^A \in S(\mathcal{H}_m), \sigma^B \in S(\mathcal{H}_n), C - \sigma^A \otimes I_n - I_m \otimes \sigma^B \geq 0 \}.
\]

Furthermore, the above supremum is equal to $T_{C}^Q(\rho^A, \rho^B)$. Moreover, for a coupling matrix $\rho^{AB} \in \Gamma(\rho^A, \rho^B)$ and $F = C - \sigma^A \otimes I_n - I_m \otimes \sigma^B \geq 0$ the following complementary implication holds:

\[
\operatorname{Tr} F \rho^{AB} = 0 \iff \operatorname{Tr} C \rho^{AB} = \operatorname{Tr} \sigma^A \rho^A + \operatorname{Tr} \sigma^B \rho^B = T_{C}^Q(\rho^A, \rho^B).
\]
In particular, if $\text{Tr} \ F \rho^{AB} = 0$ then rank $F \leq mn - \text{rank} \ \rho^{AB}$.

Assume that $\rho^A, \rho^B > 0$. Then the above supremum is achieved: There exist $\sigma^A \in S(\mathcal{H}_m), \sigma^B \in S(\mathcal{H}_n)$ such that

$$
T^Q_C(\rho^A, \rho^B) = \text{Tr}(\sigma^A \rho^A + \sigma^B \rho^B), \quad C = \sigma^A \otimes \mathbb{I}_n - \mathbb{I}_m \otimes \sigma^B \geq 0.
$$

(3.3)

We remark that the equality (3.1) is stated in [15, (4.2)].

3.1 Proofs of Theorems 3.1 and 3.2

**Proof of Theorem 3.1** Assume that $\rho^A = [a_{ij}] \in \Omega_m, \rho^B = [b_{pq}] \in \Omega_n$. Denote the entries of the Hermitian matrix $C$ by $c(i, p)(j, q)$, i.e., $c(i, p)(j, q) = \overline{c(j, q)(i, p)}$. Let $i = \sqrt{-1}$, and

$$
E^A_{ij} = |i\rangle\langle j|, \quad G^A_{ij} = \frac{1}{2}(E^A_{ij} + E^A_{ji}), \quad H^A_{ij} = \frac{1}{2}i(E^A_{ij} - E^A_{ji}), \quad i, j \in [m],
$$

$$
E^B_{pq} = |p\rangle\langle q|, \quad G^B_{pq} = \frac{1}{2}(E^B_{pq} + E^B_{qp}), \quad H^B_{pq} = \frac{1}{2}i(E^B_{pq} - E^B_{qp}), \quad p, q \in [n].
$$

Thus $|i\rangle, i \in [m], E^A_{ij}, i, j \in [m], G^A_{ij}, 1 \leq i \leq j \leq m, H^A_{ij}, 1 \leq i < j \leq m$ are the standard bases in $\mathbb{C}^m$, $\mathbb{C}^{m \times m}$, and in the subspace of $m \times m$ Hermitian matrices respectively. A similar observation applies when we replace $A$ and $m$ by $B$ and $n$. The conditions $\text{Tr}_B \rho^{AB} = \rho^A, \text{Tr}_A \rho^{AB} = \rho^B$ are stated as the following linear conditions:

$$
\text{Tr} \ \rho^{AB}(G_{ij} \otimes \mathbb{I}_n) = \Re a_{ij}, \quad i \leq j,
\text{Tr} \ \rho^{AB}(H_{ij} \otimes \mathbb{I}_n) = \Im a_{ij}, \quad i < j,
\text{Tr} \ \rho^{AB}(\mathbb{I}_m \otimes G_{pq}) = \Re b_{pq}, \quad p \leq q,
\text{Tr} \ \rho^{AB}(\mathbb{I}_m \otimes H_{pq}) = \Im b_{pq}, \quad p < q.
$$

(3.4)

Here $\Re z, \Im z$ are the real and the imaginary part of the complex number $z \in \mathbb{C}$. We assume that $\rho^{AB} \succeq 0$. Hence $T^Q_C(\rho^A, \rho^B)$ is a semidefinite problem for $\rho^{AB}$.

Assume first that $\rho^A, \rho^B$ are positive definite. Then $\rho^A \otimes \rho^B$, viewed as a Kronecker tensor product, is positive definite. Thus $\Gamma^Q(\rho^A, \rho^B)$ contains a positive definite operator $\rho^A \otimes \rho^B$. The standard SDP theory [59, Theorem 5.1] yields that $T^Q_C(\rho^A, \rho^B)$ can be computed in polynomial time with precision $\varepsilon > 0$.

(Note that the standard SDP is stated for real symmetric positive semidefinite matrices. It is well known that Hermitian positive semidefinite matrices can be encoded as special real symmetric matrices of double dimension. See the proof of Theorem 3.2 for details.)

Assume that $\rho^A, \rho^B \succeq 0$. Then the restrictions $\rho^A = \rho^A|_{\text{range} \ \rho^A}$ and $\rho^B = \rho^B|_{\text{range} \ \rho^B}$ are positive definite. Use Proposition 2.4 to deduce that $T^Q_C(\rho^A, \rho^B)$ can be computed in polynomial time in precision $\varepsilon > 0$.

**Proof of Theorem 3.2** Let us first consider the simplified case where $\rho^A, \rho^B, C$ are real symmetric. Let $S_k \supset S_{k,+} \supset S_{k,+1}$ be the space of $k \times k$ real symmetric matrices,
the cone of positive semidefinite matrices and the convex set of real density matrices. Define

\[ \Gamma^Q(\rho^A, \rho^B, \mathbb{R}) = S_{mn,+} \cap \Gamma^Q(\rho^A, \rho^B), \]

\[ T^Q_C(\rho^A, \rho^B, \mathbb{R}) = \min_{\rho^{AB} \in \Gamma^Q(\rho^A, \rho^B, \mathbb{R})} \text{Tr} \ C \rho^{AB}. \]

We claim that the dual problem to \( T^Q_C(\rho^A, \rho^B, \mathbb{R}) \) is given by

\[
\sup \{ \text{Tr} \sigma^A \rho^A + \text{Tr} \sigma^B \rho^B, \sigma^A \in S_m, \sigma^B \in S_n, C - \sigma^A \otimes I_n - I_m \otimes \sigma^B \geq 0 \}.
\]

(3.5)

Indeed, the conditions \( \text{Tr}_B \rho^{AB} = \rho^A, \text{Tr}_A \rho^{AB} = \rho^B \) for \( \rho^{AB} \in S_{mn,+} \) are stated as the linear conditions given by the first part of (3.4). Assume that \( \rho^A = [a_{ij}] \in \Omega_m, \rho^B = [b_{ij}] \in \Omega_n \). Recall the definition of the matrices \( G_{ij,m} \) introduced in the beginning of the proof of Theorem 3.1. Then the standard dual characterization of the above semidefinite problem over \( \Gamma^Q(\rho^A, \rho^B, \mathbb{R}) \) has the following form (see [59, Theorem 3.1] or [26, (2.4)]):

\[
\max \left\{ \sum_{1 \leq i \leq j \leq m} a_{ij} \tilde{u}_{ij} + \sum_{1 \leq p \leq q \leq n} b_{pq} \tilde{v}_{pq}, \tilde{u}_{ij}, \tilde{v}_{pq} \in \mathbb{R}, \left( \sum_{1 \leq i \leq j \leq m} \tilde{u}_{ij} (G_{ij,m} \otimes I_n) + \sum_{1 \leq p \leq q \leq n} \tilde{v}_{pq} (I_m \otimes G_{pq,n}) \right) \leq C \right\}.
\]

Let

\[
\sigma^A = \sum_{1 \leq i \leq j \leq m} \tilde{u}_{ij} G_{ij,m}, \quad \sigma^B = \sum_{1 \leq p \leq q \leq n} \tilde{v}_{pq} G_{pq,n}.
\]

Then the last condition of the above maximum is \( \sigma^A \otimes I_n + I_m \otimes \sigma^B \leq C \). Next observe that

\[
\text{Tr} \sigma^A \rho^A + \text{Tr} \sigma^B \rho^B = \left( \sum_{1 \leq i \leq j \leq m} a_{ij} \tilde{u}_{ij} \right) + \left( \sum_{1 \leq p \leq q \leq n} b_{pq} \tilde{v}_{pq} \right).
\]

Hence the dual to \( T^Q_C(\rho^A, \rho^B, \mathbb{R}) \) is given by (3.5). Observe that we can choose \( \sigma^A = -a I_m, \sigma^B = 0 \), where \( a \) is a positive big number such that

\[
C - \sigma^A \otimes I_n - I_m \otimes \sigma^B = C + a I_{mn} > 0.
\]

Hence the duality theorem [59, Theorem 3.1] yields that the supremum (3.5) is equal to \( T^Q_C(\rho^A, \rho^B, \mathbb{R}) \). Assume that \( \rho^A, \rho^B > 0 \). Then \( 0 < \rho^A \otimes \rho^B \in \Gamma^Q(\rho^A, \rho^B, \mathbb{R}) \). Theorem 3.1 in [59] yields that the supremum (3.5) is achieved.
We now discuss the Hermitian case. Let $i = \sqrt{-1}$. There is a standard injective map $L : S(H_m) \to S_{2^m}$:

$$L(X + iY) = \begin{bmatrix} X & Y \\ -Y & X \end{bmatrix}, \quad X, Y \in \mathbb{R}^{m \times m}, X^\top = X, Y^\top = -Y.$$ 

Note that $L(X + iY) \geq 0 \iff X + iY \geq 0$ and $L(X + iY) > 0 \iff X + iY > 0$. Hence it is possible to translate an SDP problem over Hermitian matrices to an SDP problem over reals. This yields the proof that the supremum in (3.1) is equal to $T^Q_{\mathcal{C}}(\rho^A, \rho^B)$.

Assume that $\rho^{AB} \in \Omega_1^{Q}(\rho^A, \rho^B)$ and $F = C - \sigma^A \otimes \mathbb{I}_n - \mathbb{I}_m \otimes \sigma^B \geq 0$. As $\rho^{AB}$ and $F$ are positive semidefinite we obtain

$$0 \leq \text{Tr} F \rho^{AB} = \text{Tr} C \rho^{AB} - \text{Tr} \sigma^A \rho^A - \text{Tr} \sigma^B \rho^B$$

The characterization (3.1) yields the implication (3.2). As $F$ and $\rho^{AB}$ are positive semidefinite the condition $\text{Tr} F \rho^{AB} = 0$ yields that rank $F$ + rank $\rho^{AB} \leq mn$.

Assume that $\rho^A, \rho^B > 0$. Then the above arguments show that the supremum in (3.1) is achieved. \qed

In Sect. B.3 we give an example of $\rho^A, \rho^B \in \Omega_2$, where $\rho^A$ is a pure state, for which the supremum (3.1) is not achieved. Note that the dual problem has an advantage over the original problem, as we are not constrained by linear conditions (3.4). Also the number of variables is smaller, as the supremum is restricted to $S(H_m) \times S(H_n)$. However we have to deal with the condition $\sigma^A \otimes \mathbb{I}_n + \mathbb{I}_m \otimes \sigma^B \leq C$.

### 4 Comparison of Classical and Quantum Optimal Transports for Diagonal Density Matrices

The main result of this section is a technical elaboration of the statement from [13] that the cost of quantum optimal transport is cheaper than that of the classical optimal transport. This result is stated in Theorem 4.3. In Sect. 4.1 we give a simple example in which the cost of the classical optimal transport is seven times higher than the cost of the quantum optimal transport. In Sect. 4.2 we construct a family of matrices $C^Q_{\alpha} = \alpha C^Q + (1-\alpha) \text{diag}(C^Q)$ for $\alpha \in [0, 1]$, which interpolates between the quantum cost matrix $\alpha = 1$ and its classical counterpart $\alpha = 0$. We show that in general the optimal transport cost $T^Q_{C^Q_{\alpha}}$ strictly decreases as a function of the coherence parameter $\alpha$.

**Lemma 4.1** Assume that $\rho^A, \rho^B \in \Omega_n$ and $C^Q_E$ is defined by (2.12). Then

$$T^Q_{C^Q_E}(\text{diag}(\rho^A), \text{diag}(\rho^B)) \leq T^Q_{C^Q_E}(\rho^A, \rho^B).$$

**Proof** Without loss of generality we can assume that the basis $|1\rangle, \ldots, |n\rangle$ used in (2.12) is the standard orthonormal basis in $\mathcal{H}_n = \mathbb{C}^n$. Denote by $\mathcal{D}_n \subset \mathbb{C}^{n \times n}$ the
subgroup of diagonal matrices whose diagonal entries are \( \pm 1 \). Note that \(|D_n| = 2^n\) and \( D_n \) is a subgroup of unitary matrices. Observe next that, for \( D \in D_n \),

\[
(D \otimes D)\langle \psi_{ij}^- | \psi_{ij}^- \rangle (D \otimes D) = \langle \psi_{ij}^- | \psi_{ij}^- \rangle \Rightarrow (D \otimes D)C_E^Q (D \otimes D) = C_E^Q.
\]

Hence \( T_{C_E^Q}^Q(\rho^A, \rho^B) = T_{C_E^Q}^Q(D\rho^A D, D\rho^B D) \) for each \( D \in D_n \). Clearly,

\[
\text{diag}(\rho^A) = 2^{-n} \sum_{D \in D_n} D\rho^A D, \quad \text{diag}(\rho^B) = 2^{-n} \sum_{D \in D_n} D\rho^B D.
\]

Use the convexity of \( T_{C_E^Q}^Q(\rho^A, \rho^B) \) to obtain

\[
T_{C_E^Q}^Q(\text{diag}(\rho^A), \text{diag}(\rho^B)) \leq 2^{-n} \sum_{D \in D_n} T_{C_E^Q}^Q(D\rho^A D, D\rho^B D) = T_{C_E^Q}^Q(\rho^A, \rho^B).
\]

\[\square\]

Assume that \( p_A \in \Pi_m, p_B \in \Pi_n \). The following lemma gives the isomorphism of \( \Gamma_{cl}(p_A, p_B) \) and \( \Gamma_{de}^Q(\text{diag}(p_A), \text{diag}(p_A)) \) mentioned in the Introduction. Furthermore, it describes a special \( \rho^{AB} \in \Gamma^Q(\text{diag}(p_A), \text{diag}(p_B)) \) induced by \( p^{AB} \in \Gamma_{cl}(p_A, p_B) \).

**Lemma 4.2** Let \( \rho^A \in \Omega_m, \rho^B \in \Omega_n \) and assume that \( p_A \in \Pi^m, p_B \in \Pi_n \) are induced by the diagonal entries of \( \rho^A, \rho^B \) respectively. Then

(a) Each matrix \( X = [x_{ip}]_{i \in [m], p \in [n]} \in \Gamma_{cl}(p_A, p_B) \) induces the following two matrices

\[
R = [r_{(i,p)\langle j,q \rangle}], \quad \tilde{R} = [\tilde{r}_{(i,p)\langle j,q \rangle}] \in \Gamma^Q(\text{diag}(p_A), \text{diag}(p_B)), \quad i, j \in [m], p, q \in [n].
\]

The matrix \( R \) is diagonal with \( r_{(i,p)\langle i,p \rangle} = x_{ip} \) for \( i \in [m], p \in [n] \), and \( \tilde{R} - R \) is a matrix whose only possible nonzero entries are the entries \( (i, p)(p, i) \) for \( i, p \in [\min(m, n)] \) and \( i \neq p \) which are equal to \( \sqrt{x_{ip}x_{pi}} \). Furthermore, rank \( R \leq mn - \min(m, n)(\min(m, n) - 1)/2 \).

(b) Each matrix \( R = [r_{(i,p)\langle j,q \rangle}] \in \Gamma^Q(\rho^A, \rho^B) \) induces the following two matrices: first, \( X = [x_{ip}] \in \Gamma_{cl}(p_A, p_B) \), where \( x_{ip} = r_{(i,p)\langle i,p \rangle} \) for \( i \in [m], p \in [n] \). Second, \( \tilde{R} \in \Gamma^Q(\text{diag}(\rho^A), \text{diag}(\rho^B)) \), which is obtained by replacing the entries of \( R \) at places \( (i, p)(j, q) \) by zero unless either \( (i, p)(j, q) = (i, p)(i, p) \) for \( i \in [m], p \in [n] \) or \( (i, p)(j, q) = (i, p)(p, i) \) for \( i, p \in [\min(m, n)], i \neq p \).

**Proof** (a) As \( X \in \Gamma_{cl}(p_A, p_B) \) we deduce

\[
\sum_{j=1}^{n} x_{ij} = p_i^A, \quad i \in [m], \quad \sum_{i=1}^{m} x_{ij} = p_j^B, \quad j \in [n].
\]
Assume that $R$ is a diagonal matrix with $r_{(i,p)(i,p)} = x_{ip}$. Use (2.1) to deduce that $R \in \Gamma^Q(\text{diag}(p^A), \text{diag}(p^B))$.

Consider now the matrix $\hat{R}$. In view of (2.1) we deduce that $\text{Tr}_B \hat{R} = \text{diag}(p^A)$ and $\text{Tr}_A \hat{R} = \text{diag}(p^B)$. It is left to show that $\hat{R}$ is positive semidefinite. Observe that $\hat{R}$ is a direct sum of $\left( mn - \min(m,n)(\min(m,n) - 1) \right)$ blocks of size one and $\left( \min(m,n)\min(m,n) - 1 \right)/2$ blocks of size two: $[x_{ii}]$ for $i \in [\min(m,n)]$, $[x_{ip}]$ for $i \in [m]$, $p \in [n]$, max$(i, p) > \min(m,n)$, and

$$X_{ip} = \left[ \begin{array}{c} \frac{x_{ip}}{\sqrt{x_{ip}\chi_{pi}}} \\
\sqrt{x_{ip}\chi_{pi}} \\
x_{pi} \end{array} \right], \quad \text{for } 1 \leq i < p \leq \min(m,n). \quad (4.2)$$

As $X \succeq 0$ each block is positive semidefinite and has rank at most 1. Hence rank $\hat{R} \leq mn - \min(n,n)(\min(m,n) - 1)/2$.

(b) Assume that $R \in \Gamma^Q(\rho^A, \rho^B)$. As $R$ is positive semidefinite we deduce that $r_{(i,p)(i,p)} \geq 0$. The above arguments yield that the matrix $X = [r_{(i,p)(i,p)}] \in \Gamma^cl(\rho^A, \rho^B)$. Observe next that $\hat{R}$ is a direct sum of blocks of size one and two: $[r_{(i,i)(i,i)}]$ for $i \in [\min(m,n)]$, $[r_{(i,p)(i,p)}]$ for max$(i, p) > \min(m,n)$, and

$$R_{ip} = \left[ \begin{array}{c} r_{(i,p)(i,p)} \\
\frac{r_{(p,i)(i,p)}}{r_{(p,i)(p,i)}} \end{array} \right], \quad \text{for } 1 \leq i < p \leq \min(m,n). \quad (4.3)$$

Clearly all these blocks of size one and two are principal submatrices of $R$. As $R$ is positive semidefinite, each such submatrix is positive semidefinite. Hence $\hat{R}$ is positive semidefinite. Use (2.1) to deduce that $\text{Tr}_B \hat{R} = \text{diag}(p^A)$ and $\text{Tr}_A \hat{R} = \text{diag}(p^B)$. □

**Theorem 4.3** Assume that $p^A \in \Pi_m, p^B \in \Pi_n$ are induced by the diagonal entries of $\rho^A \in \Omega_m, \rho^B \in \Omega_n$ respectively. Let $C = [C_{(i,p)(j,q)}]$ for $i, j \in [m], p, q \in [n]$ be a Hermitian matrix. Define $C^cl = [C^cl_{ip}]$ by $C^cl_{ip} = C_{(i,p)(i,p)}$ for $i \in [m], p \in [n]$.

Let $\Gamma^Q_{de}(\text{diag}(p^A), \text{diag}(p^B)) \subset \Gamma^Q(\text{diag}(p^A), \text{diag}(p^B))$ be the subset of diagonal matrices. Define $T^Q_{c,de}(\text{diag}(p^A), \text{diag}(p^B)) = \min_{R \in \Gamma^Q_{de}(p^A,p^B)} \text{Tr} CR$.

Then

(a) $T^cl_{c,de}(p^A, p^B) = T^Q_{c,de}(\text{diag}(p^A), \text{diag}(p^B)) = T^Q_{\text{diag}(C)}(\rho^A, \rho^B) \geq T^Q_{c}(\text{diag}(p^A), \text{diag}(p^B)). \quad (4.4)$

(b) Assume that $m \leq n$, and $C^Q = [C^Q_{(i,p)(j,q)}] \in S_+(\mathcal{H}_m \otimes \mathcal{H}_n)$. Denote by $C^Q_{m,n} \in S_+(\mathcal{H}_m \otimes \mathcal{H}_n)$ the submatrix of $C^Q$ whose entries are $C^Q_{(i,p)(j,q)}$ for $i, j \in [m], p, q \in [n]$. Let $C^cl_{m,n}$ be the $m \times n$ nonnegative matrix induced by the
diagonal entries of $C_{m,n}^Q$. Then

$$T_{C_{m,n}}^{cl}(p^A, p^B) = \frac{1}{2} \min_{X \in \Gamma^{cl}(p^A, p^B)} \left( \sum_{1 \leq i < p \leq m} (x_{ip} + x_{pi}) + \sum_{1 \leq i \leq m, m+1 \leq p \leq n} x_{ip} \right).$$

$$T_{C_{m,n}}^Q \left( \text{diag}(p^A), \text{diag}(p^B) \right) = \frac{1}{2} \min_{X \in \Gamma^{cl}(p^A, p^B)} \left( \sum_{1 \leq i < p \leq m} \left( x_{ip} + x_{pi} - 2 \sqrt{x_{ip} x_{pi}} \right) \right).$$

(4.5)

(c) Suppose that $m = n = 2$. Assume that $s = (s_1, s_2)^\top, t = (t_1, t_2)^\top$ are two probability vectors. Then

$$T_{C_Q}^Q(\text{diag}(s), \text{diag}(t)) = \begin{cases} \left( \sqrt{s_1} - \sqrt{t_1} \right)^2, & \text{if } s_2 \geq t_1, \\ \left( \sqrt{s_2} - \sqrt{t_2} \right)^2, & \text{if } s_2 < t_1. \end{cases}$$

(4.6)

Furthermore

$$T_{C_Q}^Q(\text{diag}(s), \text{diag}(t)) = \frac{1}{2} \max \left( \left( \sqrt{s_1} - \sqrt{t_1} \right)^2, \left( \sqrt{s_2} - \sqrt{t_2} \right)^2 \right).$$

(4.7)

Proof (a) Let $X = [x_{ij}] \in \Gamma^{cl}(p^A, p^B)$ correspond to a diagonal matrix

$R \in \Gamma_Q^Q(\text{diag}(p^A), \text{diag}(p^B))$ as in Lemma 4.2. Then $\text{Tr } C^{cl} X^\top = \text{Tr } C R$. This shows the first equality in (4.4). To show the second equality in (4.4) observe that for $R \in \Gamma_Q^Q(\rho^A, \rho^B)$ we have $\text{Tr } \text{diag}(C) R = \text{Tr } \text{diag}(C) \text{diag}(R)$. Next observe that $\text{diag}(R) \in \Gamma_Q^Q(\text{diag}(p^A), \text{diag}(p^B))$. As

$$\Gamma_Q^Q(\text{diag}(p^A), \text{diag}(p^B)) \supset \Gamma_Q(\text{diag}(p^A), \text{diag}(p^B))$$

we deduce the inequality

$$T_{C,Q}^Q(\text{diag}(p^A), \text{diag}(p^B)) \geq T_C^Q(\text{diag}(p^A), \text{diag}(p^B)).$$

The proof of (4.4) is complete.

(b) Let $R \in \Gamma_Q^Q(\text{diag}(p^A), \text{diag}(p^B))$. Define $X \in \Gamma^{cl}(p^A, p^B)$ and

$\tilde{R} \in \Gamma_Q^Q(\text{diag}(p^A), \text{diag}(p^B))$ as in part (b) of Lemma 4.2. Furthermore, let $\tilde{R} \in \Gamma_Q^Q(\text{diag}(p^A), \text{diag}(p^B))$ be defined as in part (a) of Lemma 4.2. It is straightforward to show that

$$\text{Tr } \text{diag}(C_{m,n}^Q) R = \text{Tr } C_{m,n}^{cl} X^\top, \quad \text{Tr } C_{m,n}^Q R = \text{Tr } C_{m,n}^Q \tilde{R}.$$
We now show the second equality in (4.5). As each $R_{ip}$ in (4.3) is positive semidefinite we deduce the relations:

$$\text{Tr} C_{m,n}^{Q} \hat{R} = \frac{1}{2} \left( \sum_{1 \leq i < p \leq m} (r(i,p)(i,p) + r(p,i)(p,i) - 2\sqrt{r(i,p)(p,i)}) + \sum_{1 \leq i \leq m, m+1 \leq p \leq n} r(i,p)(i,p) \right)$$

$$\geq \frac{1}{2} \left( \sum_{1 \leq i < p \leq m} (r(i,p)(i,p) + r(p,i)(p,i) - 2\sqrt{r(i,p)(p,i)}) + \sum_{1 \leq i \leq m, m+1 \leq p \leq n} r(i,p)(i,p) \right)$$

$$\geq \frac{1}{2} \left( \sum_{1 \leq i < p \leq m} (x_{ij} + x_{ji} - 2\sqrt{x_{ij}x_{ji}}) + \sum_{1 \leq i \leq m, m+1 \leq p \leq n} x_{ip} \right) = \text{Tr} C_{m,n}^{Q} \hat{R}.$$ 

This establishes the second equality in (4.5).

(c) Assume that $s_2 \geq t_1$. Then $A = \begin{bmatrix} 0 & s_1 \\ t_1 & s_2 - t_1 \end{bmatrix} \in \Gamma^{cl}(s, t)$. Therefore

$$T_{C_0}^{Q} (\text{diag}(s), \text{diag}(t)) \leq \frac{1}{2} (t_1 + s_1 - 2\sqrt{t_1s_1}) = \frac{1}{2} (\sqrt{s_1} - \sqrt{t_1})^2.$$ 

If $s_1t_1 = 0$ then $\Gamma^{cl}(s, t) = \{A\}$, and $T_{C_0}^{Q}(s, t) = \frac{1}{2} (\sqrt{s_1} - \sqrt{t_1})^2$.

Assume that $s_1t_1 > 0$. Then $\Gamma^{cl}(s, t)$ is an interval $[A, B]$. Indeed, let $C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. So $A + tC \in \Gamma^{cl}(s, t)$ for $t$ small and positive, and $B = A + t_0C$ for some $t_0 > 0$. Let $g(t) = f(A + tC)$ for $t \in [0, t_0]$. Recall that $g(t)$ is a convex function on $[0, t_0]$. Observe next that

$$g'(0+) = \frac{1}{2} (-2 + s_1^{-1/2}t_1^{1/2} + s_1^{1/2}t_1^{-1/2}) = \frac{1}{2} s_1^{-1/2}t_1^{-1/2} (\sqrt{s_1} - \sqrt{t_1})^2 \geq 0.$$ 

Hence $g(t) \geq g(0)$ for $t \in [0, t_0]$. This proves (4.6) for $s_2 \geq t_1$.

To show the equality (4.7) for $s_2 \geq t_1$ we need to show that $(\sqrt{s_1} - \sqrt{t_1})^2 \geq (\sqrt{s_2} - \sqrt{t_2})^2$. Let $x \in [0, 1/2]$. Observe that the function $\sqrt{1/2 + x} + \sqrt{1/2 - x}$ is strictly decreasing on $[0, 1/2]$. Hence

$$\sqrt{s_1} + \sqrt{s_2} \leq \sqrt{t_1} + \sqrt{t_2} \iff \max(s_1, s_2) \geq \max(t_1, t_2),$$

$$\sqrt{s_1} + \sqrt{s_2} \geq \sqrt{t_1} + \sqrt{t_2} \iff \max(s_1, s_2) \leq \max(t_1, t_2).$$

Suppose first that $s_2 \geq t_2$. Hence $s_2 \geq \max(t_1, t_2)$, and $s_1 = 1 - s_2 \leq 1 - t_2 = t_1$. Thus

$$|\sqrt{s_1} - \sqrt{t_1}| = \sqrt{t_1} - \sqrt{s_1} \geq \sqrt{s_2} - \sqrt{t_2} = |\sqrt{s_2} - \sqrt{t_2}|.$$
Suppose second that \( s_2 < t_2 \). Hence \( t_2 \geq s_1 > t_1 \). Thus \( \max(t_1, t_2) \geq \max(s_1, t_1) \).

Hence

\[
|\sqrt{s_1} - \sqrt{t_1}| = \sqrt{s_1} - \sqrt{t_1} \geq \sqrt{t_2} - \sqrt{s_2} = |\sqrt{s_2} - \sqrt{t_2}|.
\]

This proves (4.7) in the case \( s_2 \geq t_1 \). Similar arguments proves (4.6) and (4.7) in the case \( s_2 < t_1 \).

On the set of rectangular matrices \( \mathbb{R}^{m \times n} \), where \( m \leq n \), define

\[
f(X) = \frac{1}{2} \left( \sum_{1 \leq i < p \leq m} (x_{ip} + x_{pi} - 2\sqrt{x_{ip}x_{pi}}) + \sum_{1 \leq i \leq m, \ m+1 \leq p \leq n} x_{ip} \right),
\]

\[
X = [x_{ip}] \in \mathbb{R}^{m \times n}.
\]

(4.8)

Note that the second sum is zero if \( m = n \). As the function \( \sqrt{xy} \) is a concave function on \( \mathbb{R}_+^2 \) it follows that \( f(X) \) is a convex function on \( \mathbb{R}_+^{m \times n} \). Hence \( T_{C_0}^{Q}(\text{diag}(p^A), \text{diag}(p^B)) \) is the minimum of the convex function \( f(X) \) on \( \Gamma_{\text{cl}}^c(p^A, p^B) \). Therefore this minimum can be computed in polynomial time within precision \( \varepsilon > 0 \).

**Remark 4.4** Note that the second equality in (4.5) can be extended for a broad class of cost matrices \( C_0 \) introduced in (2.12).

Lemma 11 in [65] shows that

\[
C_{0}^{Q} \geq \frac{1}{2} \left( \sum_{i=1}^{n} (\sqrt{s_i} - \sqrt{t_i})^2 - \min_{j[n]} (\sqrt{s_j} - \sqrt{t_j})^2 \right),
\]

\[
\text{(4.9)}
\]

for \( s, t \in \Pi_n \). Moreover, Algorithm 1 in [65] gives \( X \in \Gamma_{\text{cl}}(s, t) \) such that \( f(X) \) is bounded from above by the right hand side of (4.9). Note (4.7) yields that for \( n = 2 \) the inequality (4.9) is sharp.

### 4.1 Quantum Optimal Transport is Less Than Classical Optimal Transport

In this subsection we give a simple example on \( \mathcal{H}_2 \otimes \mathcal{H}_2 \) that shows that the quantum optimal transport is less than the classical optimal transport by a factor of seven. That is, strict inequality holds in (4.4). A different example is given in [13]. It is straightforward to show that \( C_0^{Q} \) on \( \mathcal{H}_2 \otimes \mathcal{H}_2 \), given by (2.9), is equal to

\[
\frac{1}{2} \begin{bmatrix}
0 & 0 & 0 & 0 \\
0 & 1 & -1 & 0 \\
0 & -1 & 1 & 0 \\
0 & 0 & 0 & 0
\end{bmatrix}.
\]

(4.10)
Let \( s = (16/25, 9/25)\top, t = (9/25, 16/25)\top \). The equality (4.7) yields

\[
T^Q_C (\text{diag}(s), \text{diag}(t)) = \frac{1}{2} \left( \sqrt{\frac{16}{25}} - \sqrt{\frac{9}{25}} \right)^2 = \frac{1}{50}.
\]

Then the set of all coupling matrices is

\[
\Gamma^{cl}(s, t) = \left\{ \begin{bmatrix} x & 16/25 - x \\ 9/25 - x & x \end{bmatrix}, 0 \leq x \leq 9/25 \right\}.
\]

The classical cost matrix induced by the diagonal entries of \( C^Q \) is

\[
C^{cl} = \frac{1}{2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.
\]

(4.11)

Then, Formula (1.2) yields

\[
T^{cl}_{C^Q} (s, t) = \min_{X \in \Gamma^{cl}(s, t)} \text{Tr} C^{cl} X = \frac{1}{2} \min_{x \in [0, 9/25]} (1 - 2x) = \frac{7}{50}.
\]

4.2 Decoherence of the Quantum Cost Matrix

Let us denote

\[
C^{Q}_\alpha = \frac{1}{2} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -\alpha & 0 \\ 0 & -\alpha & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \alpha C^Q + (1 - \alpha) \text{diag}(C^Q), \quad \alpha \in [0, 1]. \quad (4.12)
\]

Since \( C^Q_1 \) reduces to \( C^Q \) defined in Eq. (4.10) the number \( \alpha \) can be called the coherence parameter. The case \( \alpha = 0 \) corresponds to the full decoherence, as the matrix \( C^Q_0 \) is diagonal, so that \( \text{Tr} C^Q_0 \rho^{AB} = \text{Tr} C^Q_0 \text{diag}(\rho^{AB}) \). The entries of \( \text{diag}(\rho^{AB}) \) correspond to the diagonal elements of \( \Gamma^{cl}(\mathbf{x}^A, \mathbf{x}^B) \), where \( \mathbf{x}^A, \mathbf{x}^B \) are the probability vectors induced by the diagonal entries of \( \rho^A \) and \( \rho^B \) respectively. Note that the cost matrix \( C^{cl} \) induced by \( C^Q_0 \) is given by (4.11). Then the quantity

\[
T^Q_\alpha (\rho^A, \rho^B) = \min_{\rho^{AB} \in \Gamma^Q(\rho^A, \rho^B)} \text{Tr} C^Q_\alpha \rho^{AB}.
\]

(4.13)

describes a continuous interpolation between the quantum and classical optimal transports, related to the gradual decoherence of the quantum state \( |\psi^-\rangle\langle\psi^-| \), which plays the role of the quantum cost matrix \( C^Q \). We will show that, for two diagonal states, \( T^Q_\alpha \) is a decreasing function of the coherence parameter \( \alpha \) on \([0, 1]\) and provide an exact expression for it.
Lemma 4.5 Let \( s, t \) be two probability vectors in \( \mathbb{R}^2 \). Assume that \( 0 \leq \alpha \leq 1 \) and denote

\[
T^Q(s, t, \alpha) = T^Q(\text{diag}(s), \text{diag}(t)), \\
f_\alpha(X) = \frac{1}{2}(x_{12} + x_{21} - 2\alpha \sqrt{x_{12}x_{21}}), \quad X = [x_{ij}] \in \Gamma^c(s, t).
\]

Then

\[
T^Q(s, t, \alpha) = \min_{X \in \Gamma^c(s, t)} f_\alpha(X). \tag{4.14}
\]

Let \( T^Q(s, t, 1) = T^Q_C(\text{diag}(s), \text{diag}(t)) \) be given by (4.7). Assume that \( T^Q(s, t, 1) = (\sqrt{s_i} - \sqrt{t_i})^2 \). If either \( \min(s_i, t_i) = 0 \) or \( s = t \) then

\[
T^Q(s, t, \alpha) = T^Q(s, t, 1) \quad \text{for all } \alpha \in [0, 1].
\]

Otherwise \( T^Q(s, t, \alpha) \) is a strictly decreasing function for \( \alpha \in [0, 1] \) given by the formula

\[
T^Q(s, t, \alpha) = \frac{1}{2} \begin{cases} 
\sqrt{1 - \alpha^2}|s_i - t_i|, & \text{for } 0 \leq \alpha < \frac{2\sqrt{s_i t_i}}{s_i + t_i}, \\
2T^Q(s, t, 1) + 2(1 - \alpha)\sqrt{s_i t_i}, & \text{for } \frac{2\sqrt{s_i t_i}}{s_i + t_i} \leq \alpha \leq 1.
\end{cases} \tag{4.15}
\]

**Proof** The equality (4.14) is deduced as the second equality in (4.5). Observe next that \( C^Q_\alpha \) is positive semidefinite. Hence \( T^Q(s, t, \alpha) \geq 0 \). Therefore for \( s = t \) we choose \( X = I \in \Gamma^c(s, t) \) to deduce from (4.14) that \( T^Q(s, t, \alpha) = 0 \). Assume that \( \min(s_i, t_i) = 0 \). Then \( \Gamma^c(s, t) = \{B\} \), where \( B \) has one zero off-diagonal element, and \( T^Q(s, t, \alpha) = T^Q(s, t, 1) \).

Assume that \( \min(s_i, t_i) > 0 \) and \( s \neq t \). Suppose first that \( s_2 \geq t_1 \). Then for \( \alpha = 1 \) (4.6) yields that \( T^Q(s, t, 1) = \frac{1}{2}(\sqrt{s_1} - \sqrt{t_1})^2 \), i.e., \( i = 1 \). Thus \( \min(s_1, t_1) > 0 \). The proof of part (c) of Theorem 4.3 implies that the minimum of \( f_1(X) \) is achieved at the matrix \( A = \begin{bmatrix} 0 & s_1 \\
t_1 & s_2 - t_1 \end{bmatrix} \), which is an extreme point of \( \Gamma^c(s, t) \). As \( s_1, t_1 > 0 \) it follows that \( \Gamma^c(s, t) \) is an interval, where the second extreme matrix is \( C = \begin{bmatrix} 1 \quad s_1 - \min(s_1, t_1) \\
t_1 - \min(s_1, t_1) & s_2 - t_1 + \min(s_1, t_1) \end{bmatrix} \). Thus we can move from \( A \) to the relative interior of \( \Gamma^c(s, t) \) by considering \( A(x) = A + xB \), where \( B = \begin{bmatrix} 1 & -1 \\
-1 & 1 \end{bmatrix} \) and \( x > 0 \).

Denoting

\[
g_\alpha(x) = f_\alpha(A(x)) = \frac{1}{2}(s_1 + t_1 - 2x - 2\alpha \sqrt{s_1 - x} \sqrt{t_1 - x}),
\]

\( \square \) Springer
one obtains
\[ g'_\alpha(0^+) = \frac{1}{2} \left[ -2 + \alpha \left( \frac{\sqrt{t_1}}{s_1} + \frac{\sqrt{s_1}}{\sqrt{t_1}} \right) \right]. \]

(Here \( h'(x^-) \) and \( h'(x^+) \) are the one-sided derivatives of a function \( h \) as \( x \) is approached from the left and right, respectively.) Hence this derivative is nonnegative for \( \alpha \geq \frac{2\sqrt{s_1 t_1}}{s_1 + t_1} \) and negative for \( 0 \leq \alpha < \frac{2\sqrt{s_1 t_1}}{s_1 + t_1} \). As \( g_\alpha(x) \) is convex on the interval \( [0, \min(s_1, t_1)] \) we obtain that for \( \frac{2\sqrt{s_1 t_1}}{s_1 + t_1} \leq \alpha \leq 1 \) the minimum of \( g_\alpha \) for \( \frac{2\sqrt{s_1 t_1}}{s_1 + t_1} \) is achieved at \( x = 0 \). This proves the second part of (4.15). So assume that \( 0 \leq \alpha < \frac{2\sqrt{s_1 t_1}}{s_1 + t_1} \). Clearly the minimum of \( f_0(X) \) on \( \Gamma^{cl}(s, t) \) is achieved at \( A(\min(s_1, t_1)) \). For \( \alpha > 0 \) we have \( g'_\alpha(\min(s_1, t_1)^-) = \infty \).

Hence for \( 0 < \alpha < \frac{2\sqrt{s_1 t_1}}{s_1 + t_1} \) the minimum \( g_\alpha(x) \) is achieved at a critical point \( x \in (0, \min(s_1, t_1)) \). This critical point is unique, as \( g_\alpha(x) \) is strictly convex on \( (0, \min(s_1, t_1)) \) and satisfies the quadratic equation
\[ 4(s_1 - x)(t_1 - x) - \alpha^2(s_1 + t_1 - 2x)^2 = 0, \quad 0 \leq \alpha < \frac{2\sqrt{s_1 t_1}}{s_1 + t_1}. \] (4.16)

We claim that the critical point is given by
\[ x(\alpha) = \frac{1}{2} \left( s_1 + t_1 - \frac{|s_1 - t_1|}{\sqrt{1 - \alpha^2}} \right), \quad 0 \leq \alpha < \frac{2\sqrt{s_1 t_1}}{s_1 + t_1}. \]

A direct computation shows that \( x(\alpha) \) satisfies (4.16). Next observe that as \( s_1 \neq t_1 \) the function \( x(\alpha) \) is a strictly decreasing function on \( [0, 1) \). Clearly
\[ x(0) = \min(s_1, t_1), \quad x\left( \frac{2\sqrt{s_1 t_1}}{s_1 + t_1} \right) = 0. \]

Hence \( x(\alpha) \in (0, \min(s_1, t_1)) \). Note that for \( x(\alpha) \) we have equality
\[ 2\sqrt{s_1} - x(\alpha)\sqrt{t_1} - x(\alpha) = \alpha(s_1 + t_1 - 2x(\alpha)). \]

This proves the first part of (4.15) in the case for \( i = 1 \). Similar arguments show the first part of (4.15) in the case for \( i = 2 \). Clearly for \( s_i \neq t_i \) and \( \min(s_i, t_i) > 0 \) the function \( T_Q^C(s, t, \alpha) \) is strictly decreasing on the interval \( [0, 1] \). \( \square \)

5 A Lower Bound on the Quantum Transport Cost \( T_Q^C(\rho^A, \rho^B) \)

The main result of this section is Theorem 5.1. Inequality (5.1) will show that \( \sqrt{T_Q^C(\rho^A, \rho^B)} \) is a weak distance, and (5.2) will allow us to obtain an explicit formula for \( T_Q^C \) for qubits.
Theorem 5.1 Let $\rho^A, \rho^B \in \Omega_n$. Then the following statements hold:

(a) For any $n$ we have

$$T_{CQ}^Q(\rho^A, \rho^B) \geq \frac{1}{2} \max_{U \in U(n), i \in [n]} \left( \sqrt{(U^\dagger \rho^A U)_{ii}} - \sqrt{(U^\dagger \rho^B U)_{ii}} \right)^2 \tag{5.1}$$

(v) For $n = 2$ equality holds in (5.1):

$$T_{CQ}^Q(\rho^A, \rho^B) = \frac{1}{2} \max_{U \in U(2), i \in [2]} \left( \sqrt{(U^\dagger \rho^A U)_{ii}} - \sqrt{(U^\dagger \rho^B U)_{ii}} \right)^2 \tag{5.2}$$

We first start with the diagonal density matrices.

5.1 A Lower Bound on $T_{CQ}^Q(\text{diag}(s), \text{diag}(t))$

The following lemma is of independent interest and is used in the proof of Theorem 5.3.

Lemma 5.2 Assume that $s, t \in \mathbb{R}^n$ are nonnegative probability vectors and $\rho^A = \text{diag}(s), \rho^B = \text{diag}(t)$. Then the dual supremum problem (3.1) can be restricted to diagonal matrices $\sigma^A = -\text{diag}(a), \sigma^B = -\text{diag}(b)$ for $a, b \in \mathbb{R}^n$ which satisfy the condition that $F = C^Q + \text{diag}(a) \otimes \mathbb{I}_n + \mathbb{I}_n \otimes \text{diag}(b)$ is positive semidefinite.

Let $X^* = [x_{ij}^*] \in \Gamma_{cl}^Q(s, t)$ be a solution to the second minimum problem in (4.5), where $p^A = s, p^B = t$ and $m = n$. Assume that the maximum in the dual supremum problem (3.1) is achieved by a matrix of the form $F^* = C^Q + \text{diag}(a^*) \otimes \mathbb{I}_n + \mathbb{I}_n \otimes \text{diag}(b^*)$, where $\rho^A = \text{diag}(s), \rho^B = \text{diag}(t), \sigma^A = -\text{diag}(a), \sigma^B = -\text{diag}(b)$. Then the following equalities hold:

$$x_{ii}^* (a_i^* + b_i^*) = 0, \text{ for } i \in [n],$$
$$x_{ij}^* (a_i^* + b_j^* + 1/2) + x_{ji}^* (a_j^* + b_i^* + 1/2) - \sqrt{x_{ij}^* x_{ji}^*} = 0, \text{ for } 1 \leq i < j \leq n. \tag{5.3}$$

Furthermore the following conditions are satisfied

(a) For $i \neq j$ either $x_{ij}^* x_{ji}^* > 0$ or $x_{ij}^* = x_{ji}^* = 0$.

(b) Assume that $x_{ii}^* x_{jj}^* > 0$. Then $x_{ij}^* = x_{ji}^*$. Let $X(t)$ be obtained from $X^*$ by replacing the entries $x_{ii}^*, x_{ij}^*, x_{ji}^*, x_{jj}^*$ with $x_{ii}^* - t, x_{ij}^* + t, x_{ji}^* + t, x_{jj}^* - t$. Then $X(t)$ is also a solution to the second minimum problem in (4.5) for $t \in [-x_{ii}^*, \min(x_{ij}^*, x_{jj}^*)]$. Furthermore, $a_i^* = a_j^* = -b_i^* = -b_j^*$. 

Springer
Suppose that $x^{*}_{ip}, x^{*}_{iq}, x^{*}_{jp}, x^{*}_{jq}$ are positive for $i \neq j, p \neq q$, where $i, j, p, q \in [n]$. Then

$$\frac{\sqrt{x^{*}_{pi}} + \sqrt{x^{*}_{qj}}}{\sqrt{x^{*}_{ip}} - \sqrt{x^{*}_{j}} - \sqrt{x^{*}_{jp}}} = 0, \quad \text{if } i \neq p, i \neq q, j \neq i, j \neq q,$$

$$1 + \frac{\sqrt{x^{*}_{pj}}}{\sqrt{x^{*}_{jq}} - \sqrt{x^{*}_{ij}}} = 0, \quad \text{if } i = p, i \neq q, i \neq j, j \neq q. \quad (5.4)$$

Furthermore, there exists a minimizing matrix $X^*$, satisfying the above conditions, such that it has at most one nonzero diagonal entry even if a maximizing $F^*$ does not exist.

**Proof** Let $a = (a_1, \ldots, a_n)^T, b = (b_1, \ldots, b_n)^T \in \mathbb{R}^n$, and consider the matrix $F = C^Q + \text{diag}(a) \otimes I_n + I_n \otimes \text{diag}(b)$. Then $F$ is a direct sum of $n$ blocks of size one of the form $a_i + b_i$ corresponding to the diagonal entries $((i, i), (i, i))$ and $n(n-1)/2$ blocks of size two corresponding to the entries $((i, j)(i, j)), ((i, j)(j, i)), ((j, i)(j, i))$: $X^*$ induces a solution to the original SDP $X^* \in \Gamma^Q(\text{diag}(s), \text{diag}(t))$ of the form described in part (a) of Lemma 4.2. That is, the diagonal entries of $X^*$ are $R^{*}_{(i, j)(i, j)} = x^{*}_{i j}$ with additional nonnegative entries: $R^{*}_{(i, j)(j, i)} = \sqrt{x^{*}_{i j}x^{*}_{j i}}$ for $i \neq j$. 

\[ \text{Tr } \sigma^A \text{ diag}(s) = -\text{Tr } \text{diag}(a) \text{ diag}(s), \quad \text{Tr } \sigma^B \text{ diag}(t) = -\text{Tr } \text{diag}(b) \text{ diag}(t). \]

Hence the dual supremum problem (3.1) can be restricted to diagonal matrices $\sigma^A = -\text{diag}(a), \sigma^B = -\text{diag}(b)$ for $a, b \in \mathbb{R}^n$ that satisfy the condition that $F$ is positive semidefinite.
Clearly, \( R^* \) is a direct sum of \( n \) submatrices of order 1 and \( n(n - 1)/2 \) of order 2 as above. The implication (3.2) yields that \( \text{Tr} \ F^* R^* = 0 \).

As \( F^* \) is positive semidefinite we deduce the conditions (5.6) for \( \mathbf{a}^* \) and \( \mathbf{b}^* \). The blocks \([x_{ii}^*] \) and \([a_i^* + b_i^*] \) contribute 1 to the ranks of \( R^* \) and \( F^* \) if and only if \( x_{ii}^* > 0 \) and \( a_i^* + b_i^* > 0 \). Each \( 2 \times 2 \) block of \( R^* \) is of the form

\[
\begin{bmatrix}
\frac{x_{ij}^*}{\sqrt{x_{ij}^* x_{ji}^*}} & \sqrt{x_{ij}^* x_{ji}^*} \\
\sqrt{x_{ij}^* x_{ji}^*} & x_{ji}^*
\end{bmatrix}
\]

for \( 1 \leq i < j \leq n \). Note that the rank of this block is either zero or one. Each corresponding \( 2 \times 2 \) submatrix of \( F^* \) is of the form \( M_{ij}^* \) given by (5.5). Thus \( M_{ij}^* \) is positive semidefinite with rank at least one. This matrix has rank one if and only if the following quadratic condition holds:

\[
(a_i^* + b_j^* + 1/2)(a_j^* + b_i^* + 1/2) - 1/4 = 0, \text{ for } 1 \leq i < j \leq n.
\]

(5.7)

Recall the complementary condition

\[
0 = \text{Tr} \ F^* F^* = \sum_{i=1}^{n} x_{ii}^*(a_i^* + b_i^*)
\]

\[
+ \sum_{1 \leq i < j \leq n} (x_{ij}^*(a_j^* + b_j^* + 1/2) + x_{ji}^*(a_j^* + b_i^* + 1/2) - \sqrt{x_{ij}^* x_{ji}^*}).
\]

As all three \( 1 \times 1 \) and \( 2 \times 2 \) corresponding blocks of \( R^* \) and \( F^* \) are positive semidefinite, it follows that we have the complementary conditions (5.3).

We now show the second part of the lemma.

(a) Assume that \( x_{ij}^* = 0 \) for \( i \neq j \). Then the second part of (5.3) yields \( x_{ij}^*(a_j^* + b_i^* + 1/2) = 0 \). The second condition in (5.5) yield that \( x_{ii}^* = 0 \).

(b) Observe that \( X(t) \in \Gamma^{cl}(s, \mathbf{t}) \) for \( t \in [-\min(x_{ij}^*, x_{ji}^*), \min(x_{ij}^*, x_{ji}^*)] \). Assume first that \( x_{ij}^* x_{ji}^* > 0 \). Let \( f(X) \) be defined as in (4.8). As \( t = 0 \) is an interior point of this interval, and \( X(0) = X^* \) we have the critical condition \( \frac{d}{dt} f(X(t)) \big|_{t=0} \), with \( f \) given by (4.8). This yields the equality \( 2 - \frac{x_{ij}^*}{\sqrt{x_{ij}^*}} - \frac{x_{ji}^*}{\sqrt{x_{ji}^*}} = 0 \). Hence \( x_{ij}^* = x_{ji}^* \) and thus \( f(X(t)) = f(X(0)) \) for \( t \in [-x_{ij}^*, \min(x_{ij}^*, x_{ji}^*)] \).

Assume now that \( x_{ij}^* \neq x_{ji}^* = 0 \). Then \( f(X(t)) = f(X(0)) \) for \( t \in [0, \min(x_{ij}^*, x_{ji}^*)] \).

It is left to show that \( a_i^* = a_j^* = -b_i^* = -b_j^* \). First observe that the first set of conditions (5.3) yield that \( a_i^* + b_j^* = a_j^* + b_i^* = 0 \). By replacing \( \mathbf{a}^* \) and \( \mathbf{b}^* \) by \( \mathbf{a}^* - \mathbf{c} \mathbf{1} \) and \( \mathbf{b}^* + \mathbf{c} \mathbf{1} \) we do not change \( F^* \). Hence we can assume that \( a_i^* = b_j^* = 0 \). Set \( b_i^* = -a_i^* \). Then the assumption that the diagonal entries of \( M_{ij}^* \) are nonnegative yields that \( |a_i^*| \leq 1/2 \). Use the assumption that \( \det M_{ij}^* \geq 0 \) to deduce that \( 0 = a_i^* = -b_i^* \).

(c) Let \( X(t) \) be the matrix obtained from \( X^* \) by replacing \( x_{ip}^*, x_{iq}^*, x_{jp}^*, x_{jq}^* \) with \( x_{ip}^* - t, x_{iq}^* + t, x_{jp}^* + t, x_{jq}^* - t \). Then for \( t \in [-\min(x_{iq}^*, x_{jp}^*), \min(x_{ip}^*, x_{jq}^*)] \) we have \( X(t) \in \Gamma^{cl}(s, \mathbf{t}) \). As \( t = 0 \) is an interior point of this interval we deduce that \( \frac{d}{dt} f(X(t)) \big|_{t=0} \).
Suppose first that \( i \neq p, i \neq q, j \neq i, j \neq q \). Then Eq. (4.8) yields

\[
f(X(t)) = -\left(\sqrt{(x_{ip}^* - t)x_{pi}^*} + \sqrt{(x_{iq}^* + t)x_{qi}^*} + \sqrt{(x_{jp}^* + t)x_{pj}^*} + \sqrt{(x_{jq}^* - t)x_{qj}^*}\right) + C,
\]

where \( C \) is a term that does not depend on \( t \). The condition \( \frac{d}{dt} f(X(t)) \big|_{t=0} \) yields the first condition (5.4).

Assume now that \( i = p \) and \( i \neq q, j \neq i, j \neq q \). Then we have

\[
f(X(t)) = t/2 - \left(\sqrt{(x_{iq}^* + t)x_{qi}^*} + \sqrt{(x_{jp}^* + t)x_{pj}^*} + \sqrt{(x_{jq}^* - t)x_{qj}^*}\right) + C,
\]

where \( C \) does not depend on \( t \). Now, the condition \( \frac{d}{dt} f(X(t)) \big|_{t=0} \) yields the second condition in (5.4).

Finally, we need to prove the existence of an \( X^* \) with at most one nonzero entry that satisfies the conditions of the lemma. Assume first that \( s, t > 0 \). Then Theorem 3.2 yields that there exists a maximizing matrix \( F^* \) to the dual supremum problem. As we showed above we can assume that \( F^* = CQ + \text{diag}(a^*) \otimes I_n + I_n \otimes \text{diag}(b^*) \).

Let \( X^* \) be a minimizing matrix with at most \( k \) zeros on the diagonal. Assume to the contrary that \( x_{ii}^*x_{jj}^* > 0 \) for \( 1 \leq i < j \leq n \). Part (b) yields that for \( t \in [-\min(x_{ii}^*, x_{jj}^*), \min(x_{ii}^*, x_{jj}^*)] \) the matrix \( X(t) \) minimizes \( f \). Choose \( t^* = \min(x_{ii}^*, x_{jj}^*) \). Then \( X(t^*) \) is a minimizing matrix with at least \( k + 1 \) zeros on the diagonal, which contradicts our choice of \( X^* \).

Assume now that \( s, t \) are nonnegative. Let \( s_k, t_k > 0, k \in \mathbb{N} \) be two sequences that converge to \( s, t \) respectively. Let \( X_k^* \) be a minimizing matrix of \( f(X) \) corresponding to \( s_k, t_k \) that has at most one nonzero diagonal element. Clearly, there exists a subsequence \( X_{k_l}^* \) which has either all zero diagonal elements or exactly one positive diagonal element in a fixed diagonal entry. Choose a subsequence \( [\tilde{x}_{ij,l}^*], l \in \mathbb{N} \) of this subsequence which converges to \( X^* \). Clearly \( X^* \) is a minimizing matrix of \( f(X) \) corresponding to \( s, t \). If \( x_{ij}^* > 0 \) then \( \tilde{x}_{ij,l}^* > 0 \) for \( l \gg 1 \). Hence \( X^* \) satisfies the conditions of the lemma. \( \square \)

**Theorem 5.3** Assume that \( s = (s_1, \ldots, s_n)^\top, t = (t_1, \ldots, t_n)^\top \in \mathbb{R}^n_+ \) are probability vectors and \( U \in U(n) \). Then

\[
T_{CQ}^Q(U^\dagger \text{diag}(s)U, U^\dagger \text{diag}(t)U) \geq \frac{1}{2} \max_{i \in [n]} \left(\sqrt{s_i} - \sqrt{t_i}\right)^2 \quad (5.8)
\]

Equality holds if and only if there exists \( i \in [n] \) such that

\[
\begin{align*}
either s_j & \geq t_j and t_i t_j \geq s_i s_j for all j \neq i, 
\text{or} \quad t_j & \geq s_j and s_i s_j \geq t_i t_j for all j \neq i. \quad (5.9)
\end{align*}
\]
In particular, for \( n = 2 \) equality in (5.8) holds:

\[
T_{Q}^{Q} \left( U^{\dagger} \text{diag}(s_{1}, s_{2})^{\top} U, \ U^{\dagger} \text{diag}(t_{1}, t_{2})^{\top} U \right) \\
= \frac{1}{2} \max((\sqrt{s_{1}} - \sqrt{t_{1}})^{2}, (\sqrt{s_{2}} - \sqrt{t_{2}})^{2}).
\]

(5.10)

**Proof** Without loss of generality we can assume that \( U = I_{n} \). Suppose first that \( s, t > 0 \). Lemma 5.2 yields that \( T_{Q}^{Q} \) is the maximum of the dual problem where \( F = C^{Q} + \text{diag}(a) \otimes I_{n} + I_{n} \otimes \text{diag}(b) \) is positive semidefinite. Choose \( i \in [n] \).

Assume that the coordinates of \( a, b \) are given as follows:

\[
a_{i} = \frac{1}{2} \left( \frac{\sqrt{s_{i}}}{\sqrt{s_{i}}} - 1 \right), \quad b_{i} = \frac{1}{2} \left( \frac{\sqrt{s_{i}}}{\sqrt{t_{i}}} - 1 \right), \quad a_{j} = b_{j} = 0 \quad \text{for} \quad j \neq i.
\]

Clearly

\[
a_{i} + b_{i} = \frac{(\sqrt{s_{i}} - \sqrt{t_{i}})^{2}}{2\sqrt{s_{i}t_{i}}} \geq 0, \quad a_{j} + b_{j} = 0, \quad \text{for} \quad j \neq i.
\]

\[
1/2 + a_{i} > 0, \quad 1/2 + b_{i} > 0, \quad 1/2 + a_{j} = 1/2 + b_{j} = 1/2, \quad \text{for} \quad j \neq i.
\]

\[
(a_{i} + b_{j} + 1/2)(a_{j} + b_{i} + 1/2) = (a_{i} + 1/2)(b_{i} + 1/2) = 1/4, \quad \text{for} \quad j \neq i.
\]

\[
(a_{j} + b_{p} + 1/2)(a_{p} + b_{j} + 1/2) = 1/2 \times 1/2 = 1/4, \quad \text{for} \quad p \neq j \in [n] \setminus \{i\}.
\]

Thus \( F \geq 0 \). Therefore

\[
T_{Q}^{Q}(\text{diag}(s), \text{diag}(t)) \geq -\text{Tr} \left( \text{diag}(a) \text{diag}(s) + \text{diag}(b) \text{diag}(t) \right) \\
= \frac{1}{2} \left[ \left( 1 - \frac{\sqrt{s_{i}}}{\sqrt{s_{i}}} \right) s_{i} + \left( 1 - \frac{\sqrt{s_{i}}}{\sqrt{t_{i}}} \right) t_{i} \right] = \frac{1}{2} (\sqrt{s_{i}} - \sqrt{t_{i}})^{2}.
\]

As we let \( i \in [n] \) we deduce the inequality (5.8). Since \( T_{Q}^{Q}(\text{diag}(s), \text{diag}(t)) \) is continuous on \( \Pi_{n} \times \Pi_{n} \) we deduce the inequality (5.8) for all \( (s, t) \in \Pi_{n} \times \Pi_{n} \).

We now discuss the equality case in (5.8). Clearly \( \max_{i \in [n]}(\sqrt{s_{i}} - \sqrt{t_{i}})^{2} = 0 \) if and only if \( s = t \), in which case \( T_{Q}^{Q}(\text{diag}(s), \text{diag}(t)) = 0 \). Assume that \( T_{Q}^{Q}(\text{diag}(s), \text{diag}(t)) > 0 \). Suppose first that equality holds in (5.8). Then there exists an index \( i \in [n] \) such that \( T_{Q}^{Q}(\text{diag}(s), \text{diag}(t)) = \frac{1}{2} (\sqrt{s_{i}} - \sqrt{t_{i}})^{2} > 0 \). By renaming indices and interchanging \( s \) and \( t \) if needed we can assume that \( t_{1} > s_{1} \) and \( T_{Q}^{Q}(\text{diag}(s), \text{diag}(t)) = \frac{1}{2} (\sqrt{s_{1}} - \sqrt{t_{1}})^{2} \). Let \( X = X^{*} \) be a solution to the second minimum problem in (4.5). Recall that \( f(X^{*}) = \frac{1}{2} (\sqrt{t_{1}} - \sqrt{s_{1}})^{2} \). Suppose first that \( s_{1} = 0 \). Then the first row of each \( X \in \Gamma^{cl}(s, t) \) is zero. Hence

\[
2f(X) = \sum_{j=2}^{n} x_{j1} + \sum_{2 \leq j < k \leq n} (\sqrt{x_{jk}} - \sqrt{x_{kj}})^{2} = t_{1} + \sum_{2 \leq j < k \leq n} (\sqrt{x_{jk}} - \sqrt{x_{kj}})^{2}.
\]
for $X \in \Gamma^{cl}(s, t)$. As $f(X^*) = t_1$ we deduce that the submatrix $Y = [x^*_{jk}]_{j,k \geq 2}$ is a nonnegative symmetric matrix. Thus for $j \geq 2$

$$s_j = \sum_{k=1}^{n} x^*_{jk} = x^*_{j1} + \sum_{k=2}^{n} x^*_{jk} = x^*_{j1} + \sum_{k=2}^{n} x^*_{kj} = x^*_{j1} + t_j.$$ 

Therefore $s_j \geq t_j$ and $t_1t_j \geq 0 = s_1s_j$ for $j \geq 2$. Hence the conditions (5.9) hold.

Assume now that $s_1 > 0$. Let $F$ be defined as above for $i = 1$. Our assumption is that $F = F^*$ is a solution to the maximum dual problem. Lemma 5.2 yields the equalities (5.3). Hence $x^*_{11} = 0$. Next consider the second part of the equalities (5.3) for $i = 1$ and $j \geq 2$:

$$\frac{\sqrt{t_1}}{\sqrt{s_1}} x^*_{1j} = \frac{\sqrt{t_1}}{\sqrt{s_1}} x^*_{j1} = c_j \geq 0 \text{ for } j \geq 2.$$ 

Observe next that

$$s_1 = \sum_{j=2}^{n} x^*_{j1} = \frac{\sqrt{s_1}}{\sqrt{t_1}} \sum_{j=2}^{n} c_j \Rightarrow \sum_{j=2}^{n} c_j = \sqrt{s_1t_1}.$$ 

Therefore

$$\sum_{j=2}^{n} (x^*_{1j} + x^*_{j1} - 2\sqrt{x^*_{1j}x^*_{j1}}) = s_1 + t_1 - 2\sum_{j=2}^{n} c_j = s_1 + t_1 - 2\sqrt{s_1t_1} = (\sqrt{s_1} - \sqrt{t_1})^2.$$ 

Hence

$$2f(X^*) = (\sqrt{s_1} - \sqrt{t_1})^2 + \sum_{2 \leq j < k \leq n} (\sqrt{x^*_{jk}} - \sqrt{x^*_{kj}})^2 = (\sqrt{s_1} - \sqrt{t_1})^2.$$ 

Therefore the submatrix $Y = [x^*_{jk}]_{j,k \geq 2}$ is a nonnegative symmetric matrix. Observe next that

$$s_j = x^*_{j1} + \sum_{k=2}^{n} x^*_{jk} = \frac{\sqrt{t_1}}{\sqrt{s_1}} c_j + \sum_{k=2}^{n} x^*_{jk},$$

$$t_j = x^*_{1j} + \sum_{k=2}^{n} x^*_{kj} = \frac{\sqrt{s_1}}{\sqrt{t_1}} c_j + \sum_{k=2}^{n} x^*_{kj}, \text{ for } j \geq 2.$$
As \( Y \) is symmetric we obtain that
\[
 s_j - t_j = \frac{(t_1 - s_1)c_j}{\sqrt{s_1 t_1}} \geq 0 \quad \Rightarrow \quad c_j = \frac{(s_j - t_j)\sqrt{s_1 t_1}}{t_1 - s_1}.
\]

As
\[
 s_j \geq x_j^* = \sqrt{\frac{t_1}{s_1}}c_j = \frac{(s_j - t_j)t_1}{t_1 - s_1}
\]
we deduce that \( t_1 t_j \geq s_1 s_j \). Hence conditions (5.9) hold.

Assume now that the conditions (5.9) hold. To be specific we assume that \( t_1 \geq s_1 \) and \( s_j \geq t_j \) for \( j \geq 2 \). If \( s_j = t_j \) for \( j \geq 2 \) then \( s = t \) and equality holds in (5.8). Hence we assume that \( t_1 > s_1 \). Define \( X = [x_{ij}] \) as follows:
\[
 x_{11} = 0, \quad x_{1j} = \frac{s_1(s_j - t_j)}{t_1 - s_1}, \quad x_{j1} = \frac{t_1(s_j - t_j)}{t_1 - s_1}, \quad x_{jk} = \frac{t_1 t_j - s_1 s_j}{t_1 - s_1} \delta_{jk} \quad \text{for} \quad j, k \geq 2.
\]
Then \( X \in \Gamma_{c^1}(s, t) \). Furthermore \( 2f(X) = s_1 + t_1 - 2\sqrt{s_1 t_1} = (\sqrt{s_1} - \sqrt{t_1})^2 \). Therefore \( 2T_{CQ}^Q(s, t) \leq (\sqrt{s_1} - \sqrt{t_1})^2 \). On the other hand, inequality (5.8) yields that \( 2T_{CQ}^Q(\text{diag}(s), \text{diag}(t)) \geq (\sqrt{s_1} - \sqrt{t_1})^2 \). Consequently, we conclude that \( T_{CQ}^Q(\text{diag}(s), \text{diag}(t)) = \frac{1}{2}(\sqrt{s_1} - \sqrt{t_1})^2 \).

Assume that \( n = 2 \). Then \( s_1 + s_2 = t_1 + t_2 = 1 \). Assume for simplicity of exposition that \( t_1 t_2 \geq s_1 s_2 \). Then \( s_i = \min(s_1, s_2) \leq \min(t_1, t_2) \). Hence \( s_j = 1 - s_i \geq t_j = 1 - t_i \) and the first condition of (5.9) holds. Hence, (5.10) is holds.

\[\Box\]

### 5.2 Proof of Theorem 5.1

(a) Recall inequality (4.1) and the fact that \( T_{CQ}^Q \) is unitarily invariant, \( T_{CQ}^Q(\rho^A, \rho^B) = T_{CQ}^Q(U^\dagger \rho^A U, U^\dagger \rho^B U) \), for \( U \in U(n) \). Use the inequality (5.8) with \( U = I_n \) to deduce
\[
 T_{CQ}^Q(\rho^A, \rho^B) = T_{CQ}^Q(U^\dagger \rho^A U, U^\dagger \rho^B U) \geq T_{CQ}^Q(\text{diag}(U^\dagger \rho^A U), \text{diag}(U^\dagger \rho^B U)) \geq \frac{1}{2} \max_{i \in [n]} \left( \sqrt{(U^\dagger \rho^A U)_{ii}} - \sqrt{(U^\dagger \rho^B U)_{ii}} \right)^2.
\]

Take the maximum over \( U \in U(n) \) to deduce (5.1).
(b) First observe that $F$ that is given in (3.2) is of the form:

$$\sigma^A = -\begin{bmatrix} a & b \\ b & c \end{bmatrix}, \quad \sigma^B = -\begin{bmatrix} e & f \\ f & g \end{bmatrix}, \quad a, c, e, g \in \mathbb{R}, \ b, f \in \mathbb{C},$$

$$F = \begin{bmatrix} a + e & f & b & 0 \\ f & a + g + 1/2 & -1/2 & b \\ b & -1/2 & c + e + 1/2 & f \\ 0 & b & \bar{f} & c + g \end{bmatrix}.$$  \hfill (5.12)

We now assume that $\rho^A, \rho^B$ are positive definite and non-isospectral. Proposition 2.1 yields that $\Gamma^Q(\rho^A, \rho^B)$ does not contain a matrix of rank one. Let $\rho^{AB}$ and $F$ be the matrices for which (3.2) holds. Our assumptions yield that rank $\rho^{AB} \geq 2$. Proposition 3.2 yields that $\text{Tr} \, \rho^{AB} = 0$. Hence rank $F \leq 4 - 2 = 2$. Note that the second and the third columns of $F$ are nonzero. Hence rank $F \geq 1$.

For $U \in U(2)$ we have the equalities

$$T^Q_{CQ}(\rho^A, \rho^B) = T^Q_{CQ}(U^\dagger \rho^A U, U^\dagger \rho^B U) = \text{Tr} \big( \sigma^A \rho^B + \sigma^B \rho^B \big)$$

$$= \text{Tr} \big( (U^\dagger \sigma^A U)(U^\dagger \rho^B U) + (U^\dagger \sigma^B U)(U^\dagger \rho^B U) \big)$$

$$F = (U^\dagger \otimes U^\dagger) F (U \otimes U^\dagger) = C^Q - (U^\dagger \sigma^A U) \otimes I_2 - I_2 \otimes (U^\dagger \sigma^B U) \geq 0.$$

We now choose $V \in U(2)$ so that $V^\dagger \sigma^A V$ is a diagonal matrix. Let

$$\rho^A = V^\dagger \rho^A V, \quad \rho^B = V^\dagger \rho^B V,$$

$$\sigma^A = V^\dagger \sigma^A V = -\begin{bmatrix} a & 0 \\ 0 & c \end{bmatrix}, \quad \sigma^B = V^\dagger \sigma^B V = -\begin{bmatrix} e & f \\ f & g \end{bmatrix}, \quad a, c, e, g \in \mathbb{R}, \ f \in \mathbb{C},$$

$$F = \begin{bmatrix} a + e & f & 0 & 0 \\ f & a + g + 1/2 & -1/2 & 0 \\ 0 & -1/2 & c + e + 1/2 & f \\ 0 & 0 & \bar{f} & c + g \end{bmatrix}.$$  \hfill (5.12)

Clearly rank $F = \text{rank} \, F \leq 2$. We claim that rank $F = 2$. Assume to the contrary that rank $F = 1$. As the third column is nonzero we deduce that the fourth column is a multiple of the third column. Hence the fourth column is zero. That is, $f = c + g = 0$. Similarly $a + e = 0$. Next observe that we can replace $\sigma^A, \sigma^B$ by $\sigma^A - a I_2, \sigma^B + a I_2$ without affecting the supremum in (3.1). This is equivalent to the assumption that $a = 0$. Hence $e = 0$ and $g = -c$. As $F$ is Hermitian and rank $F = 1$ we have the condition

$$0 = (-c + 1/2)(c + 1/2) - 1/4 = -c^2.$$

Hence $c = -g = 0$. Thus we can assume that $\sigma^A = \sigma^B = 0$. Equality (3.3) yields that $T^Q_{CQ}(\rho^A, \rho^B) = 0$, which implies that $\rho^A = \rho^B$. This contradicts our assumption that $\rho^A$ and $\rho^B$ are not similar. Hence rank $F = \text{rank} \, F = 2$. 

 Springer
We claim that either $x = a + e$ or $z = c + g$ are zero. Assume to the contrary that $\bar{x}, \bar{z} > 0$. (Recall that $\bar{F} > 0$.) Let $c_1, c_2, c_3, c_4$ be the four columns of $F$. Clearly $c_1, c_4$ are linearly independent. Hence $c_2 = uc_1 + vc_4$. As the fourth coordinate of $c_2$ is zero we deduce that $v = 0$. Hence $c_2 = uc_1$. This is impossible since the third coordinate of $c_1$ is 0 and the third coordinate of $c_2$ is $-1/2$. Hence either $\bar{x} = a + e$ or $\bar{z} = c + g$ are zero. Suppose that $\bar{x} = 0$. As $F$ is positive semidefinite we deduce that the first row and column of $F$ is zero. Hence $f = 0$. Similarly, if $\bar{z} = 0$ we deduce that $f = 0$. Thus $\sigma^A$ and $\sigma^B$ are diagonal matrices. Therefore

$$T^Q_C(\rho^A, \rho^B) = \text{Tr}(\sigma^A \rho^A + \sigma^B \rho^B) = \text{Tr}(\sigma^A \text{diag}(\rho^A) + \sigma^B \text{diag}(\rho^B)).$$

As $F \geq 0$, the maximum dual characterization yields

$$\text{Tr}(\sigma^A \text{diag}(\rho^A) + \sigma^B \text{diag}(\rho^B)) \leq T^Q_C(\text{diag}(\rho^A), \text{diag}(\rho^B)).$$

Hence $T^Q_C(\rho^A, \rho^B) \leq T^Q_C(\text{diag}(\rho^A), \text{diag}(\rho^B))$. Compare that with (4.1) to deduce the equalities

$$T^Q_C(\rho^A, \rho^B) = T^Q_C(\text{diag}(\rho^A), \text{diag}(\rho^B)).$$

Use the last part of Theorem 5.3 to deduce

$$T^Q_C(\rho^A, \rho^B) = T^Q_C(\text{diag}(\rho^A), \text{diag}(\rho^B)) = \frac{1}{2} \max \left[ (\sqrt{\rho^A_{11}} - \sqrt{\rho^B_{11}})^2, (\sqrt{\rho^A_{22}} - \sqrt{\rho^B_{22}})^2 \right].$$

The inequality (5.1) yields (5.2) for $\rho^A$ and $\rho^B$ positive definite and non-isospectral. Clearly every pair $\rho^A, \rho^B \in \Omega_2$ can be approximated by $\hat{\rho}^A, \hat{\rho}^B \in \Omega_2$ which are positive definite and non-isospectral. Use the continuity of $T^Q_C(\rho^A, \rho^B)$ on $\Omega_2 \times \Omega_2$ (Proposition 2.2) to deduce (5.2) in the general case.

### 6 The Induced Quantum Wasserstein-2 Distance

The main theorem of this section is:

**Theorem 6.1** Assume that $C \in S(\mathcal{H}_n \otimes \mathcal{H}_n)$ is positive semidefinite and vanishes exactly on $\mathcal{H}_S$, the symmetric subspace of $\mathcal{H}_n \otimes \mathcal{H}_n$. Then $\sqrt{T^Q_C}$ is a weak distance, and $\mathcal{W}^Q_C$ defined by (1.6) is the maximum distance majorized by $\sqrt{T^Q_C}$. For $n = 2$ the equality $\mathcal{W}^Q_C(\rho^A, \rho^B) = \sqrt{T^Q_C(\rho^A, \rho^B)}$ holds.

Let $X$ be a set of points. Assume that $D : X \times X \to \mathbb{R}_+([0, \infty))$. Then $D(\cdot, \cdot)$ is called a distance on $X$ if it satisfies the following three properties:
(a) **Symmetry:** \( D(x, y) = D(y, x) \);
(b) **Positivity:** \( D(x, y) \geq 0 \), and equality holds if and only if \( x = y \).
(c) **Triangle inequality:** \( D(x, y) + D(y, z) \geq D(x, z) \).

The function \( D(\cdot, \cdot) \) is called a semidistance if it satisfies the above first two conditions. A semidistance is called a weak distance if there exists a distance \( D'(\cdot, \cdot) \) such that

\[
D'(x, y) \leq D(x, y) \quad \text{for all } x, y \in X. 
\]

**Proposition 6.2** Assume that \( D \) is a weak distance on the space \( X \) satisfying (6.1), where \( D' \) is a distance on \( X \). For each positive integer \( N \) define the following function:

\[
D_N(x, y) = \inf_{z_1, \ldots, z_N \in X, z_0 = x, z_{N+1} = y} \sum_{i=0}^{N} D(z_i, z_{i+1}) \quad \text{for } x, y \in X.
\]

Then

(a) For each \( N \) the function \( D_N(\cdot, \cdot) \) is a weak distance that satisfies the inequality (6.1).
(b) For each \( x, y \in X \) and \( N \) we have the inequalities \( 0 \leq D_{N+1}(x, y) \leq D_N(x, y) \leq D(x, y) \).
(c) For each \( M, N \geq 1 \) we have the inequality

\[
D_M(x, u) + D_N(u, y) \geq D_{M+N+1}(x, y) \quad \text{for } x, u, y \in X.
\]
(d) Denote by \( D_\infty(x, y) = \lim_{N \to \infty} D_N(x, y) \). Then \( D_\infty(x, y) \) is a distance, called the induced distance of \( D \). Furthermore, \( D_\infty \) is the maximum distance \( D' \) that satisfies (6.1).

**Proof** (a) Clearly \( D_N(x, y) \geq 0 \). As \( D(x, y) = D(y, x) \) it follows that

\[
D(z_0, z_1) + \cdots + D(z_N, z_{N+1}) = D(z_N+1, z_N) + \cdots + D(z_1, z_0).
\]

Hence \( D_N(x, y) = D_N(y, x) \). Assume that \( y = x \). Choose \( z_1 = \cdots = z_N = x \). As \( D(x, x) = 0 \) we deduce that \( \sum_{i=0}^{N} D(z_i, z_{i+1}) = 0 \). Hence \( D_N(x, x) = 0 \). As \( D' \) is a distance we deduce

\[
\sum_{i=0}^{N} D'(z_i, z_{i+1}) \geq D'(z_0, z_{N+1}) = D'(x, y).
\]

Use (6.1) to deduce that

\[
\sum_{i=0}^{N} D(z_i, z_{i+1}) \geq \sum_{i=0}^{N} D'(z_i, z_{i+1}) \geq D'(x, y).
\]
Hence $D_N$ satisfies the inequality (6.1). In particular, if $x \neq y$ then $D_N(x, y) \geq D'(x, y) > 0$. Therefore $D_N$ is a weak distance.

(b) Assume that $z_1 = \ldots = z_N = x$, $z_{N+1} = y$. Then $\sum_{i=0}^N D(z_i, z_{i+1}) = D(x, y)$. Hence $D_N(x, y) \leq D(x, y)$. Now let $z_{N+1} = z_{N+2} = y$. Then

$$\sum_{i=0}^N D(z_i, z_{i+1}) = \sum_{i=0}^{N+1} D(z_i, z_{i+1}).$$

Hence $D_{N+1}(x, y) \leq D_N(x, y)$.

(c) Choose $z_0 = x$, $z_{M+1} = u$, $z_{M+N+2} = y$, and $z_1, \ldots, z_{M+N+1}$ arbitrarily. Then $\sum_{i=0}^{M+N+1} D(z_i, z_{i+1}) \geq D_{M+N+1}(x, y)$. Compare that with the definitions of $D_M(x, u)$ and $D_N(u, y)$ to deduce the inequality $D_M(x, u) + D_N(u, y) \geq D_{M+N+1}(x, y)$.

(d) As $\{D_N(x, y)\}$ is a non-increasing sequence such that $D_N(x, y) \geq D'(x, y)$ we deduce that the limit $D_\infty(x, y)$ exists and $D(x, y) \geq D_\infty(x, y) \geq D'(x, y)$. Since $D_N(x, y) = D_N(y, x)$ it follows that $D_\infty(x, y) = D_\infty(y, x)$. Hence $D_\infty(x, y) \geq 0$ and equality holds if and only if $x = y$. In the inequality $D_M(x, u) + D_N(u, x) \geq D_{M+N+1}(x, y)$ let $M = N \to \infty$ to deduce that $D_\infty$ satisfies the triangle inequality. Hence $D_\infty$ is a distance. The inequality $D(x, y) \geq D_\infty(x, y) \geq D'(x, y)$ yields that $D_\infty$ is a maximum distance $D'$ that satisfies (6.1).

\[\square\]

**Theorem 6.3** Let $C \in \mathcal{S}(\mathcal{H}_n \otimes \mathcal{H}_n)$. Then $T^Q_C$ is a semidistance on $\Omega_n \times \Omega_n$ if and only if $C$ is positive semidefinite and $\ker(C) = \mathcal{H}_S$. Furthermore, for $\rho^A, \rho^B \in \Omega_n$ the following statements hold:

(a) $T^Q_C(\rho^A, \rho^B) = T^Q_C(\rho^B, \rho^A)$.

(b) $T^Q_C(\rho^A, \rho^B) \geq 0$.

(c) $T^Q_C(\rho^A, \rho^B) = 0$ if and only if $\rho^A = \rho^B$.

(d) For $C = C^Q$ the inequality $T^Q_C(\rho^A, \rho^B) \leq \frac{1}{2} (1 - \text{Tr} \rho^A \rho^B)$ holds. Furthermore

$$T^Q_C(\rho^A, \rho^B) = \frac{1}{2} (1 - \text{Tr} \rho^A \rho^B) \text{ if either } \rho^A \text{ or } \rho^B \text{ is a pure state.} \quad (6.2)$$

(e) $\sqrt{T^Q_C(\rho^A, \rho^B)}$ is a distance on pure states.

**Proof** We first show the second part of the theorem. Assume that $C$ is positive semidefinite and vanishes exactly on symmetric matrices.

(a) As $S$ is an involution with the eigenspaces $S^2 \mathbb{C}^n$ and $A^2 \mathbb{C}^n$ corresponding to the eigenvalues 1 and $-1$ respectively, and $CS^2 \mathbb{C}^n = 0$, it follows that $SC = CS = -C$. Hence $SCS^\dagger = C$. The second equality in (2.10) yields that $\Sigma^Q(\rho^A, \rho^B)S^\dagger = \Gamma^Q(\rho^B, \rho^A)$. As $Tr C \rho^{AB} = Tr CS \rho^{AB} S^\dagger$ we deduce (a).

(b) Since $C \geq 0$, for any $\rho^{AB} \in \Omega_n^2$ we get that $\text{Tr} C \rho^{AB} \geq 0$. This proves (b).

(c) Suppose that $\rho^A = \rho^B = \rho$. Consider the spectral decomposition of $\rho$ given by (2.4). Then a purification of $\rho$ is...
Clearly \( R \in \Gamma^Q(\rho, \rho) \). As \( X = \sum_{i=1}^{n} \sqrt{\lambda_i} |x_i\rangle |x_i\rangle \) is a symmetric matrix it follows that \( CX = 0 \). Hence \( \mathrm{Tr} CR = 0 \) and \( T^Q_C(\rho, \rho) = 0 \).

Assume now that \( T^Q_C(\rho^A, \rho^B) = 0 \). Hence \( \mathrm{Tr} C \rho^{AB} = 0 \) for some \( \rho^{AB} \in \Gamma^Q(\rho^A, \rho^B) \). That is, the eigenvectors of \( \rho^{AB} \) are symmetric matrices. Therefore \( \rho^{AB} = \sum_{j=1}^{k} p_j |\psi_j\rangle \langle \psi_j| \) where each \( |\psi_j\rangle \) is a symmetric matrix and \( p_j > 0 \). We claim that each \( |\psi_j\rangle \langle \psi_j| \) is of the form (6.3). This is equivalent to the Autonne–Takagi factorization theorem [35, Corollary 4.4.4, part (c)] that any symmetric \( X \in \mathbb{C}^{n \times n} \) is of the form

\[
X = \sum_{i=1}^{n} d_i |x_i\rangle \langle x_i| = U DU^\top, \quad D = \text{diag}(d), \quad U \in U(n),
\]

where the columns of \( U \) represent vectors, \( x_1, \ldots, x_n \). Clearly \( \mathrm{Tr} A |\psi_j\rangle \langle \psi_j| = \mathrm{Tr} B |\psi_j\rangle \langle \psi_j| \). Hence \( \rho^B = \mathrm{Tr} A \rho^{AB} = \mathrm{Tr} B \rho^{AB} = \rho^A \).

(d) As \( \rho^A \otimes \rho^B \in \Gamma^Q(\rho^A, \rho^B) \) it follows that \( T^Q_C(\rho^A, \rho^B) \leq \mathrm{Tr} C \rho^{AB} = \mathrm{Tr} (\rho^A \otimes \rho^B) \). Clearly \( \mathrm{Tr} (\rho^A \otimes \rho^B) = 1 \). The first part of (2.10) yields that \( \mathrm{Tr} S(\rho^A \otimes \rho^B) = \mathrm{Tr} (\rho^A \otimes \rho^B) \). Hence \( \mathrm{Tr} C \rho^{AB} = \frac{1}{2} (1 - \mathrm{Tr} \rho^A \rho^B) \), and \( T^Q_C(\rho^A, \rho^B) \leq \frac{1}{2} (1 - \mathrm{Tr} \rho^A \rho^B) \). Assume that either \( \rho^A \) or \( \rho^B \) is a pure state. Lemma A.3 yields that \( \Gamma^Q(\rho^A \rho^B) = \{ \rho^A \otimes \rho^B \} \). Hence (6.2) holds.

(e) It is known that if \( \rho^A, \rho^B \) are pure state then [51]

\[
\sqrt{1 - \mathrm{Tr} \rho^A \rho^B} = \frac{1}{2} \| \rho^A - \rho^B \|_1,
\]

\[
\rho^A = |x\rangle \langle x|, \quad \rho^B = |y\rangle \langle y|, \quad \langle x|x\rangle = \langle y|y\rangle = 1.
\]

Note that if one of the states is pure then \( \sqrt{1 - \mathrm{Tr} \rho^A \rho^B} \) reduces to the root infidelity [31, 45]—see also Eq. (2.18). We give a short proof for completeness. By changing the orthonormal basis in \( \mathcal{H}_n \) we can assume that \( n = 2 \) and

\[
\rho^A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \rho^B = \begin{bmatrix} b & c \\ c & 1 - b \end{bmatrix}, \quad 0 \leq b \leq 1, \quad 0 \leq c, \quad c^2 = b(1 - b).
\]

As \( \mathrm{Tr}(\rho^A - \rho^B) = 0 \) it follows that the two eigenvalues of \( \rho^A - \rho^B \) are

\[
\pm \sqrt{-\det(\rho^A - \rho^B)} = \pm \sqrt{(1 - b)^2 + c^2} = \pm \sqrt{1 - b} = \pm \sqrt{1 - \mathrm{Tr} \rho^A \rho^B}.
\]

This proves (6.4). Hence \( \frac{1}{2} \| \rho^A - \rho^B \|_1 + \frac{1}{2} \| \rho^B - \rho^C \|_1 \geq \frac{1}{2} \| \rho^A - \rho^C \|_1 \). Combine that with (d) to deduce (e).
We now show the first part of the theorem. Suppose that $C$ is positive semidefinite and vanishes exactly on symmetric matrices. Then parts (a)–(c) of the theorem show that $T_C^Q$ is a semidistance.

Assume now that $C \in S(H_n \otimes H_n)$ and $T_C^Q$ is a semidistance. For $n = 1$ it is straightforward to see that $C = 0$. Assume that $n > 1$. As $T_C^Q(\rho^A, \rho^B) > 0$ for $\rho^A \neq \rho^B \in \Omega_n$ it follows that $C \neq 0$. Let $R \in S(H_n \otimes H_n)$ be nonzero and positive semidefinite. We claim that $\text{Tr} \ CR \geq 0$. It is enough to assume that $\text{Tr} \ R = 1$. Set $\rho^A = \text{Tr}_B \ R$, $\rho^B = \text{Tr}_A \ R$. Then $R \in \Gamma^Q(\rho^A, \rho^B)$. Thus $0 \leq T_C^Q(\rho^A, \rho^B) \leq \text{Tr} \ CR$.

Suppose that $C = \sum_{k=1}^{n^2} \mu_k |\psi_k\rangle\langle\psi_k|$, where $|\psi_1\rangle, \ldots, |\psi_{n^2}\rangle$ is an orthonormal basis for $H_n \otimes H_n$. Choose rank-one $R_k = |\psi_k\rangle\langle\psi_k|$ $\geq 0$. Thus $\mu_k = \text{Tr} \ CR_k \geq 0$ for $k \in \{n^2\}$. Hence $C \geq 0$. Let $\rho = |x\rangle\langle x|$ be a pure state. Lemma A.3 yields that $\Gamma^Q(\rho, \rho) = \{\rho \otimes \rho\}$. Hence $0 = T_C^Q(\rho, \rho) = \text{Tr} \ C(\rho \otimes \rho)$. Noting that $\rho \otimes \rho = (|x\rangle\langle x|)(|x\rangle\langle x|)$, as $C$ is positive semidefinite we deduce that $C(|x\rangle\langle x|) = 0$. So $C$ vanishes on all rank one symmetric matrices, hence $C \mathcal{H}_S = 0$.

It is left to show that $C|Y\rangle \neq 0$ if $Y$ is a nonzero skew-symmetric matrix. Assume to the contrary that $C|Y\rangle = 0$ for some nonzero skew-symmetric matrix $Y$. Let $Z \in S^2 \mathbb{C}^n$ be the unique symmetric matrix with zero diagonal such that $X = Z + Y$ is a nonzero lower triangular matrix with zero diagonal. Note that $C|X\rangle = 0$. Normalize $X$ such that $\text{Tr} \ XX^\dagger = 1$. Let $R = |X\rangle\langle X|$, $\rho^A = \text{Tr}_B \ R$, $\rho^B = \text{Tr}_A \ R \in \Omega_n$. Clearly $\text{Tr} \ CR = 0$. Hence $0 \leq T_C^Q(\rho^A, \rho^B) \leq \text{Tr} \ CR = 0$. As $T_C^Q$ is a semidistance we deduce that $\rho^A = \rho^B$. We now contradict this equality. Indeed, consider the equality (2.2). As $X$ is lower triangular with zero diagonal its first row is zero. Hence $\rho^A_{11} = 0$. Hence $\rho^B_{11} = 0$. Note that $\rho^B_{11}$ is the norm squared of the first column of $X$. Hence the first column of $X$ is zero. Therefore the second row of $X$ is zero. Thus $\rho^A_{22} = 0$, which yields that $\rho^B_{22} = 0$. Therefore the second column of $X$ is zero. Repeat this argument to deduce that $X = 0$, which contradicts our assumption that $\text{Tr} \ XX^\dagger = 1$. \hfill \square

We now give a very general distance on positive semidefinite matrices, inspired by our lower bound (5.1) on $T_C^Q(\rho^A, \rho^B)$, which is exact on qubit density matrices.

**Proposition 6.4** Let $v : \mathbb{R}^n \to [0, \infty)$ be a norm. Assume that $f : [0, \infty) \to [0, \infty)$ is a continuous, strictly increasing function. For $\rho^A, \rho^B$ positive semidefinite define

$$D(\rho^A, \rho^B) = \max_{U \in U(n)} v\left((f((U^\dagger \rho^A U)_{11}), \ldots, f((U^\dagger \rho^A U)_{nn}))^T - (f((U^\dagger \rho^B U)_{11}), \ldots, f((U^\dagger \rho^B U)_{nn}))^T\right).$$

(6.5)

Then $D(\rho^A, \rho^B)$ is a distance on positive semidefinite matrices. In particular,

$$D_0(\rho^A, \rho^B) = \max_{U \in U(n), i \in [n]} \left| f((U^\dagger \rho^A U)_{ii}) - f((U^\dagger \rho^B U)_{ii}) \right|$$

$$= \max_{U \in U(n)} \left| f((U^\dagger \rho^A U)_{11}) - f((U^\dagger \rho^B U)_{11}) \right|$$

(6.6)

is a distance on positive semidefinite matrices.
Let us note that the distance \((6.6)\) admits an operational interpretation: It quantifies the maximal distinguishability between two given states, \(\rho^A, \rho^B \in \Omega_n\), achievable through local unitary rotations followed by a projective measurement.

**Proof** By definition \(D(\rho^A, \rho^B) = D(\rho^B, \rho^A) \geq 0\). Assume that \(D(\rho^A, \rho^B) = 0\). Then \(f((U^\dagger \rho^A U)_{ii}) = f((U^\dagger \rho^B U)_{ii})\) for each \(i \in [n]\) and \(U \in U(n)\). As \(f\) is strictly increasing we deduce that \((U^\dagger \rho^A U)_{ii} = (U^\dagger \rho^B U)_{ii}\) for each \(i \in [n]\). That is for each \(U \in U(n)\) the diagonal entries of \(U^\dagger (\rho^A - \rho^B) U\) are 0. Choose a unitary \(V\) so that \(V^\dagger (\rho^A - \rho^B) V = 0\). Hence \(\rho^A = \rho^B\). It is left to show the triangle inequality.

Denote by \(f(\rho)\) the vector \(\left(f(\rho_{11}), \ldots, f(\rho_{nn})\right)^\top\). Since \(f\) is continuous there exists \(V \in U(n)\) such that \(D(\rho^A, \rho^B) = v(\left(f(V^\dagger \rho^A V) - f(V^\dagger \rho^B V)\right))\). Hence

\[
D(\rho^A, \rho^B) = v\left(f(V^\dagger \rho^A V) - f(V^\dagger \rho^B V)\right) \\
\leq v\left(f(V^\dagger \rho^A V) - f(V^\dagger \rho^C V)\right) + v\left(f(V^\dagger \rho^C V) - f(V^\dagger \rho^B V)\right) \\
\leq D(\rho^A, \rho^C) + D(\rho^C, \rho^B).
\]

To show that \(D_0(\cdot, \cdot)\) is a distance we observe that \(D_0(\rho^A, \rho^B) = D(\rho^A, \rho^B)\) where \(v\left((x_1, \ldots, x_n)^\top\right) = \max_{i \in [n]} |x_i|\). To show the last equality of \((6.6)\) let \(P_\permut\subset U(n)\) denote the group of permutation matrices. Then

\[
\max_{i \in [n]} |f((U^\dagger \rho^A U)_{ii}) - f((U^\dagger \rho^B U)_{ii})| = \max_{P \in P_\permut} |f\left(((UP)^\dagger \rho^A (UP))_{11}\right) - f\left(((UP)^\dagger \rho^B (UP))_{11}\right)|.
\]

\(\square\)

**Proof of Theorem 6.1** As \(C\) is positive semidefinite that vanishes exactly on symmetric matrices, Theorem 6.3 yields that \(T^C_Q\) is a semidistance on \(\Omega_n\). Next observe that \(C \geq a C_Q\) for some \(a > 0\). Hence \(T^C_Q(\rho^A, \rho^B) \geq a T^Q_C(\rho^A, \rho^B)\). Let \(D_0(\rho^A, \rho^B)\) be the distance defined in \((6.6)\), with \(f(x) = \sqrt{x/2}\) for \(x \geq 0\). The inequality \((5.1)\) yields that \(\sqrt{T^C_Q(\rho^A, \rho^B)} \geq \sqrt{\bar{a}} D_0(\rho^A, \rho^B)\). Hence \(\sqrt{T^C_Q(\rho^A, \rho^B)}\) is a weak distance. Proposition 6.2 shows that \(\sqrt{T^C_Q}\) yields the induced Wasserstein-2 distance given by \((1.6)\), which is the maximum distance majorized by \(\sqrt{T^C_Q}\).

Assume that \(n = 2\). Then \((5.2)\) yields that \(\sqrt{T^C_Q}(\rho^A, \rho^B) = D_0(\rho^A, \rho^B)\). Hence,

\[
T^C_Q(\rho^A, \rho^B) = T^Q_C(\rho^A, \rho^B).
\]

\(\square\)

Observe that if \(\sqrt{T^C_Q}\) is a weak distance, then so is \((T^C_Q)^{1/p}\) for any \(p \geq 2\), which dominates the distance \(D_0^{2/p}\). Then, one can define \(W^C_{p} Q\) in a similar way as in formula \((1.6)\).
7 Quantum Optimal Transport for $d$-Partite Systems

We now explain briefly how to state the quantum optimal transport problem for a $d$-partite system, where $d \geq 3$, similarly to what was done in [27, 30]. The main result of this section is Theorem 7.1.

Let $\mathcal{H}_{n_j}$ be a Hilbert space of dimension $n_j$ for $j \in [d]$. We consider the $d$-partite tensor product space $\otimes_{j=1}^{d} \mathcal{H}_{n_j}$. A product state in Dirac’s notation is $\otimes_{i=1}^{d} |x_i\rangle$. Then

$$\langle \otimes_{i=1}^{d} x_i, \otimes_{j=1}^{d} y_j \rangle = \langle \otimes_{i=1}^{d} (x_i) | \otimes_{j=1}^{d} (y_j) \rangle = \prod_{j=1}^{d} \langle x_j | y_j \rangle.$$ 

Consider the space $\mathcal{B}(\otimes_{j=1}^{d} \mathcal{H}_{n_j})$ of linear operators from $\otimes_{j=1}^{d} \mathcal{H}_{n_j}$ to itself. A rank-one product operator is of the form $\langle \otimes_{j=1}^{d} (x_j) | \otimes_{j=1}^{d} (y_j) \rangle$ and acts on a product state as follows:

$$\langle \otimes_{i=1}^{d} (x_i) | \otimes_{j=1}^{d} (y_j) | \otimes_{k=1}^{d} (z_k) \rangle = \left( \prod_{j=1}^{d} \langle y_j | z_j \rangle \right) \langle \otimes_{i=1}^{d} (x_i) \rangle.$$ 

Given $\rho^{A_1,\ldots,A_d} \in \mathcal{B}(\otimes_{j=1}^{d} \mathcal{H}_{n_j})$ one can define a $k$-partial trace on $k \in [d]$:

$$\text{Tr}_k : \mathcal{B}(\otimes_{j=1}^{d} \mathcal{H}_{n_j}) \to \mathcal{B}\left(\otimes_{j \in [d] \setminus \{k\}} \mathcal{H}_{n_j}\right),$$

$$\text{Tr}_k(\otimes_{i=1}^{d} |x_i\rangle)(\otimes_{j=1}^{d} |y_j\rangle) = \langle y_k | (\otimes_{i \in [d] \setminus \{k\}} |x_i\rangle)(\otimes_{j \in [d] \setminus \{k\}} |y_j\rangle |).$$

We will denote $\text{Tr}_k \rho^{A_1,\ldots,A_d}$ by $\rho^{A_1,\ldots,A_{k-1},A_{k+1},\ldots,A_d}$. Let $\rho^{A_k} \in \mathcal{B}(\mathcal{H}_{n_k})$ be the operator obtained from $\rho^{A_1,\ldots,A_d}$ by tracing out all but the $k$-th component. Thus we have the map

$$\tilde{\text{Tr}} : \mathcal{B}(\otimes_{j=1}^{d} \mathcal{H}_{n_j}) \to \otimes_{j=1}^{d} \mathcal{B}(\mathcal{H}_{n_j}),$$

$$\tilde{\text{Tr}}(\rho^{A_1,\ldots,A_d}) = (\rho^{A_1}, \ldots, \rho^{A_d}).$$

Let $N = \prod_{j=1}^{d} n_j$ and view the set of density matrices $\Omega_N$ as a subset of selfadjoint operators on $\mathcal{H}_N = \otimes_{j=1}^{d} \mathcal{H}_{n_j}$. For $\rho^{A_i} \in \Omega_{n_i}, i \in [d]$ denote

$$\Gamma^Q(\rho^{A_1}, \ldots, \rho^{A_d}) = \{\rho^{A_1,\ldots,A_d} \in \Omega_N, \tilde{\text{Tr}}(\rho^{A_1,\ldots,A_d}) = (\rho^{A_1}, \ldots, \rho^{A_d})\}.$$ 

Assume that $C$ is a selfadjoint operator on $\mathcal{H}_N$. We define the quantum optimal transport as

$$\text{Tr}_C^Q(\rho^{A_1}, \ldots, \rho^{A_d}) = \min_{\rho^{A_1,\ldots,A_d} \in \Gamma^Q(\rho^{A_1}, \ldots, \rho^{A_d})} \text{Tr} C \rho^{A_1,\ldots,A_d}. \quad (7.1)$$
We now give an analog of a result in [27]. Assume that \( d = 2\ell \geq 4 \), and \( n_1 = \cdots = n_d = n \). Then \( \mathcal{H}_n^{\otimes d} = \bigotimes^d \mathcal{H}_n \). We want to give a semidistance between two ordered \( \ell \)-tuples of density matrices \( (\rho^{A_1}, \ldots, \rho^{A_{2\ell}}), (\rho^{A_{\ell+1}}, \ldots, \rho^{A_{2\ell}}) \in \Omega_n^\ell \). We view \( \mathcal{H}_n^{\otimes (2\ell)} \) as bipartite states \( \mathcal{H}_n^{\otimes \ell} \otimes \mathcal{H}_n^{\otimes \ell} \). Let \( S \in \mathcal{B}(\mathcal{H}_n^{\otimes (2\ell)}) \) be the SWAP operator:

\[
S(\otimes^\ell_{j=1} |x_j\rangle) = (\otimes^\ell_{j=1} |x_{j+\ell}\rangle) \otimes (\otimes^\ell_{j=1} |x_j\rangle).
\]

Denote by \( C^Q = \frac{1}{2}(I - S) \). Then \( T^Q_{C^Q}(\rho^{A_1}, \ldots, \rho^{A_{2\ell}}) \geq 0 \). Equality holds if and only if \( (\rho^{A_1}, \ldots, \rho^{A_{\ell}}) = (\rho^{A_{\ell+1}}, \ldots, \rho^{A_{2\ell}}) \). Also

\[
T^Q_{C^Q}(\rho^{A_1}, \ldots, \rho^{A_{2\ell}}) = T^Q_{C^Q}(\rho^{A_{1+\ell}}, \ldots, \rho^{A_{2\ell}}, \rho^{A_1}, \ldots, \rho^{A_{\ell}}).
\]

Hence \( T^Q_{C^Q}(\rho^{A_1}, \ldots, \rho^{A_{2\ell}}) \) is a semidistance on \( \Omega_n^\ell \). As in the case of \( \ell = 1 \) we can show that \( \sqrt{T^Q_{C^Q}(\rho^{A_1}, \ldots, \rho^{A_{2\ell}})} \) is a weak distance. Denote by \( W^Q_{C^Q}((\rho^{A_1}, \ldots, \rho^{A_{\ell}}), (\rho^{A_{\ell+1}}, \ldots, \rho^{A_{2\ell}})) \) the Wasserstein-2 distance on \( \Omega_n^\ell \) induced by the weak distance \( \sqrt{T^Q_{C^Q}(\rho^{A_1}, \ldots, \rho^{A_{2\ell}})} \).

Let \( \Sigma_\ell \) be the group of bijections \( \pi : [\ell] \to [\ell] \). Then

\[
\min_{\pi \in \Sigma_\ell} W^Q_{C^Q}((\rho^{A_{\pi(1)}}, \ldots, \rho^{A_{\pi(\ell)}}), (\rho^{1+\ell}, \ldots, \rho^{2\ell}))
\]
gives a distance on unordered \( \ell \)-tuples of density matrices. We call this distance the quantum Wasserstein-2 distance on the set of unordered \( \ell \)-tuples \( (\rho^{A_1}, \ldots, \rho^{A_{\ell}}) \).

On \( \mathcal{H}_n^{\otimes d} \) we define for two integers \( 1 \leq p < q \leq d \) the SWAP operator \( S_{pq} \in \mathcal{B}(\mathcal{H}_n^{\otimes d}) \), which swaps \( x_p \) with \( x_q \) in the tensor product \( |x_1\rangle \otimes \cdots \otimes |x_d\rangle \). Note that \( S_{pq} \) is unitary and involutive. Hence \( S_{pq} \) is selfadjoint with eigenvalues \( \pm 1 \). The common invariant subspace of \( \mathcal{H}_n^{\otimes d} \) for all \( S_{pq} \) is the subspace of symmetric tensors —―“bosons”―—, denoted as \( S^d \mathcal{H}_n \). Let \( C^B \in S_+(\mathcal{H}_n^{\otimes d}) \) be the projection on the orthogonal complement of \( S^d \mathcal{H}_n \). Note that \( C^B = C^Q \) for \( d = 2 \). We now have a partial analog of Theorem 6.3:

**Theorem 7.1** Let \( \rho^{A_1}, \ldots, \rho^{A_d} \in \Omega_n \). Then

(a) \( T^Q_{C^B}(\rho^{A_1}, \ldots, \rho^{A_d}) \geq 0 \).

(b) \( T^Q_{C^B}(\rho^{A_1}, \ldots, \rho^{A_d}) = 0 \) if and only if \( \rho^{A_1} = \cdots = \rho^{A_d} \).

(c) Assume that at least \( d - 1 \) out of \( \rho^{A_1}, \ldots, \rho^{A_d} \) are pure states. Then

\[
T^Q_{C^B}(\rho^{A_1}, \ldots, \rho^{A_d}) = \text{Tr} C^B (\otimes^d_{j=1} \rho^{A_j}).
\]

**Proof** (a) This follows from the fact that \( \text{Tr} C^B \rho^{A_1,\ldots,A_d} \geq 0 \).

(b) Assume that \( T^Q_{C^B}(\rho^{A_1}, \ldots, \rho^{A_d}) = \text{Tr} C^B \rho^{A_1,\ldots,A_d} = 0 \). Hence all the eigenvectors of \( \rho^{A_1,\ldots,A_d} \) corresponding to positive eigenvalues are symmetric tensors. So \( S_{pq} \rho^{A_1,\ldots,A_d} S_{pq} = \rho^{A_1,\ldots,A_d} \). Therefore \( \text{Tr}(\rho^{A_1,\ldots,A_d}) = (\rho, \ldots, \rho) \). Thus
\( \rho^{A_1} = \ldots = \rho^{A_d} = \rho \). We now show that \( T^Q_{CB}(\rho, \ldots, \rho) = 0 \). Suppose that \( \rho \) has the spectral decomposition (2.4). Let us take a \( d \)-purification of \( \rho \)

\[
\rho_{\text{pur}, d} = \left( \sum_{i=1}^{n} \sqrt{\lambda_i} \otimes^d |x_i\rangle \right) \left( \sum_{j=1}^{n} \sqrt{\lambda_j} \otimes^d \langle x_j| \right).
\]

Clearly we have \( \rho_{\text{pur}, d} \in \Gamma^Q(\rho, \ldots, \rho) \). As \( \rho_{\text{pur}, d} \) is a pure state whose eigenvector corresponding to its positive eigenvalue is a symmetric tensor we deduce that \( \text{Tr} C_B \rho_{\text{pur}, d} = 0 \).

(c) Assume for simplicity of the exposition that \( \rho^{A_2}, \ldots, \rho^{A_d} \) are pure states. Then \( \rho^B = \otimes_{j=2}^d \rho^{A_j} \) is a pure state. Lemma A.3 yields that \( \Gamma^Q(\rho^{A_1}, \rho^B) = \{ \rho^{A_1} \otimes \rho^B \} \). Hence \( \Gamma^Q(\rho^{A_1}, \ldots, \rho^{A_d}) = \{ \otimes_{j=1}^d \rho^{A_j} \} \). This proves part (c) of the theorem. \( \square \)

The next question concerns the optimal technique to compute \( \text{Tr} C_B (\otimes_{j=1}^d \rho^{A_j}) \). This problem is related to the permanent function. Assume first that each \( \rho^{A_j} \) is a pure state \( |x_j\rangle\langle x_j| \), where \( \langle x_j| x_j \rangle = 1 \). Then \( \otimes_{j=1}^d \rho^{A_j} \) is a pure product state with the positive eigenvector \( \otimes_{j=1}^d |x_j\rangle \). A symmetrization of \( \otimes_{j=1}^d |x_j\rangle \) is the orthogonal projection on the subspace of symmetric tensors, given by

\[
(\mathbb{I} - C_B) (\otimes_{j=1}^d |x_j\rangle) = \frac{1}{d!} \sum_{\pi \in \Sigma_d} \otimes_{j=1}^d |x_{\pi(j)}\rangle.
\]

Hence

\[
\left\| (\mathbb{I} - C_B) (\otimes_{j=1}^d |x_j\rangle) \right\|^2 = \frac{1}{d!} \sum_{\pi \in \Pi_d} \prod_{j=1}^d \langle x_j| x_{\pi(j)} \rangle.
\]

Let \( X = [x_1 \ldots x_d] \in \mathbb{C}^{n \times d} \) be the matrix whose columns are the vectors \([x_1, \ldots, x_d]\). The \( G(x_1, \ldots, x_d) = X^\dagger X \) is the Gramian matrix \([\langle x_i| x_j \rangle] \in H_{d, +} \). Note that since \( \|x_1\| = \cdots = \|x_d\| = 1 \) the diagonal entries of \( G(x_1, \ldots, x_d) \) are all 1, and \( G(x_1, \ldots, x_d) \) is called a complex covariance matrix. It now follows that \( \| (\mathbb{I} - C_B) \otimes_{j=1}^d |x_j\rangle \|^2 \) is \( \frac{1}{d!} \) times the permanent of \( G(x_1, \ldots, x_d) \), denoted as \( \text{per} G(x_1, \ldots, x_d) \).

\[
\text{Tr} C_B (\otimes_{i=1}^d |x_i\rangle) (\otimes_{j=1}^d |x_j\rangle) = 1 - \frac{1}{d!} \text{per} G(x_1, \ldots, x_d), \quad \|x_1\| = \cdots = \|x_d\| = 1.
\]

**Lemma 7.2** Assume that \( \rho^{A_1}, \ldots, \rho^{A_d} \in \Omega_n \) have the following spectral decomposition:

\[
\rho^{A_j} = \sum_{i=1}^{n} \lambda_{i, j} |x_{i, j}\rangle \langle x_{i, j}|, \quad j \in [d].
\]
Then

$$\text{Tr} \ C^B(\otimes_{j=1}^d \rho^{A_j}) = 1 - \frac{1}{d!} \sum_{i_1,...,i_d \in [n]} \prod_{j=1}^d \lambda_{i_j,j} \text{ per } G(x_{i_1,1}, \ldots, x_{i_d,d}). \quad (7.2)$$

The proof of this lemma follows straightforwardly from the multilinearity of $\otimes_{j=1}^d \rho^{A_j}$.

We now state the analog to part (d) of Theorem 6.3, which is a corollary to the above lemma:

**Corollary 7.3** Let $\rho^{A_1}, \ldots, \rho^{A_d}$ be density matrices with the spectral decomposition given by Lemma 7.2. Then

$$T^Q_{CB}(\rho^{A_1}, \ldots, \rho^{A_d}) \leq 1 - \frac{1}{d!} \sum_{i_1,...,i_d \in [n]} \prod_{j=1}^d \lambda_{i_j,j} \text{ per } G(x_{i_1,1}, \ldots, x_{i_d,d}).$$

If at least $d - 1$ density matrices are pure states then equality holds.

**Acknowledgements** It is a pleasure to thank Rafal Bistroń, John Calsamiglia, Matt Hoogsteder, Tomasz Miller, Wojciech Słomczyński and Andreas Winter for numerous inspiring discussions and helpful remarks. Financial support by Simons collaboration Grant for mathematicians, Narodowe Centrum Nauki under the Maestro Grant number DEC-2015/18/A/ST2/00274 and by the Foundation for Polish Science under the Team-Net Project no. POIR.04.04.00-17C1/18-00 is gratefully acknowledged.

**Appendix A: Basic Properties of Partial Traces**

In order to understand the partial traces on $B(H_m \otimes H_n)$ it is convenient to view this space as a 4-mode tensor space [29] and use Dirac notation. Denote by $H_m'$ the space of linear functionals on $H_m$, i.e., the dual space. Then $y^\vee = \langle y | \in H_m'$ acts on $z \in H_m$ as follows: $y^\vee(z) = \langle y, z \rangle = \langle y | z \rangle$. Hence a rank-one operator in $B(H_m)$ is of the form $x \otimes y^\vee = |x\rangle \langle y|$, where $\langle (|x\rangle \langle y|)(z) = \langle y | z \rangle |x\rangle$. So $|x\rangle \langle y|$ can be viewed a matrix $\rho = xy^\dagger \in \mathbb{C}^{m \times m}$. Assume that $V_1, V_2$ are linear transformations from $H_m$ to itself. Then $V_1 \otimes V_2$ is a sesquilinear transformation from $H_m \otimes H_m'$ to itself, which acts on rank one operators as follows:

$$(V_1 \otimes V_2)(|x\rangle \langle y|) = |V_1 x\rangle \langle V_2 y| = V_1(|x\rangle \langle y|) V_2^\dagger, \quad x, y \in H_m.$$  

Assume now that $W_1, W_2$ are linear transformations from $H_n$ to itself. Then

$$(V_1 \otimes W_1)|x\rangle |v\rangle = |V_1 x\rangle |W_1 v\rangle, \quad x \in H_m, y \in H_n.$$  

A tensor product of two rank-one operators is identified a 4-tensor:

$$|x\rangle \langle y| \otimes |u\rangle \langle v| = |x\rangle |u\rangle \langle y| |v\rangle, \quad x, y \in H_m, u, v \in H_n. \quad (A.1)$$
Thus
\[
(|x⟩|u⟩⟨y|⟨v|)(|z⟩|w⟩) = (y|z⟩⟨v|w⟩|x⟩|u⟩), \quad x, y, z ∈ H_m, u, v, w ∈ H_n.
\]

Observe next that \(V_1 ⋄ W_1 ⋄ V_2 ⋄ W_2\) is a multi-sesquilinear transformation of \(B(H_m ⋄ H_n)\) to itself, which acts on a rank-one product operator as follows:
\[
(V_1 ⋄ W_1 ⋄ V_2 ⋄ W_2)(|x⟩|u⟩⟨y|⟨v|) = |V_1x⟩|W_1u⟩⟨V_2y|⟨W_2v|
\]
\[
= (V_1 ⋄ W_1)(|x⟩|u⟩⟨v|)(V_2^† ⋄ W_2^†).
\]
(In the last equality we view \(|x⟩|u⟩⟨v|\) as an \((mn) × (mn)\) matrix.) As \(\text{Tr} |x⟩⟨y| = ⟨y|x⟩\) we deduce the following lemma:

**Lemma A.1** Let
\[
x, y ∈ H_m, u, v ∈ H_n, \quad V_1, V_2 ∈ B(H_m), W_1, W_2 ∈ B(H_n).
\]

Then
\[
\text{Tr}_A |x⟩|u⟩⟨y|⟨v| = ⟨y|x⟩|u⟩⟨v|,
\]
\[
\text{Tr}_B |x⟩|u⟩⟨y|⟨v| = ⟨v|u⟩⟨x|y⟩,
\]
\[
\text{Tr}_A(V_1 ⋄ W_1 ⋄ V_2 ⋄ W_2)(|x⟩|u⟩⟨y|⟨v|) = ⟨V_2y|V_1x⟩|W_1u⟩⟨W_2v|,
\]
\[
\text{Tr}_B(V_1 ⋄ W_1 ⋄ V_2 ⋄ W_2)(|x⟩|u⟩⟨y|⟨v|) = ⟨W_2v|W_1u⟩|V_1x⟩⟨V_2y|.
\]

In particular, if \(V_1 = V_2 = V\) and \(W_1 = W_2 = W\) are unitary then
\[
\text{Tr}_A(V ⋄ W ⋄ V ⋄ W)(|x⟩|u⟩⟨y|⟨v|) = ⟨y|x⟩|Wu⟩⟨Wv|,
\]
\[
\text{Tr}_B(V ⋄ W ⋄ V ⋄ W)(|x⟩|u⟩⟨y|⟨v|) = ⟨v|u⟩|Vx⟩⟨Vy|.
\]

**Corollary A.2** Let \(ρ^A ∈ Ω_m, ρ^B ∈ Ω_n, V ∈ B(H_m), W ∈ B(H_n)\) be unitary and \(C ∈ S(H_m ⋄ H_n)\). Then
\[
Γ^Q(V ρ^AV^†, W ρ^BW^†) = (V ⋄ W)Γ^Q(ρ^A, ρ^B)(V^† ⋄ W^†),
\]
\[
T^Q_C(ρ^A, ρ^B) = T_{(V ⋄ W)C(V^† ⋄ W^†)}(V ρ^AV^†, W ρ^BW^†).
\]

**Proof** View \(ρ^A ∈ Ω_m\) as an element in \(H_m ⋄ H_m^∗\) to deduce \(V ρ^AV^† = (V ⋄ V)ρ^A\).
Suppose that
\[
ρ^{AB} = \sum_{i,j ∈ [m], p,q ∈ [n]} r_{(i,p)(j,q)}|i⟩⟨p| |j⟩⟨q| ∈ Γ^Q(ρ^A, ρ^B).
\]
Let $\tilde{\rho}^{AB} = (V \otimes W \otimes V \otimes W)\rho^{AB}$. Observe that

\[
\text{Tr}_A \tilde{\rho}^{AB} = \sum_{p,q \in [n]} \left( \sum_{i \in [m]} r(i,p)(i,q) \right) |p\rangle\langle q| = \rho^B,
\]

\[
\text{Tr}_A \tilde{\rho}^{AB} = \sum_{p,q \in [n]} \left( \sum_{i \in [m]} r(i,p)(i,q) \right) \langle q|W^\dagger(W|p) = W\rho_B W^\dagger.
\]

Similarly $\text{Tr}_B \tilde{\rho}^{AB} = V\rho^A V^\dagger$. Hence

\[
(V \otimes W \otimes V \otimes W)\Gamma^Q(\rho^A, \rho^B) \subseteq \Gamma^Q(V\rho^A V^\dagger, W\rho_B W^\dagger).
\]

and

\[
(V^\dagger \otimes W^\dagger \otimes V^\dagger \otimes W^\dagger)\Gamma^Q(V\rho^A V^\dagger, W\rho_B W^\dagger) \subseteq \Gamma^Q(\rho^A, \rho^B).
\]

Hence we deduce the first part of the corollary. The second part of the corollary follows from the identity

\[
\text{Tr} \ C \rho^{AB} = \text{Tr}(V \otimes W)C(V^\dagger \otimes W^\dagger)(V \otimes W)\rho^{AB}(V^\dagger \otimes W^\dagger).
\]

\[\square\]

The following result appeared in the literature [29] and we state it here for completeness. For $\rho^A \in \mathcal{B}(\mathcal{H}_m)$ denote by $\text{range } \rho^A \subseteq \mathcal{H}_m$ the range of $\rho^A$.

**Lemma A.3** Let $\rho^A \in \Omega_m$, $\rho^B \in \Omega_n$. Then

\[
\Gamma^Q(\rho^A, \rho^B) \subseteq \mathcal{B}(\text{range } \rho^A) \otimes \mathcal{B}(\text{range } \rho^B).
\]

In particular if either $\rho^A$ or $\rho^B$ is a pure state then $\Gamma^Q(\rho^A, \rho^B) = \{\rho^A \otimes \rho^B\}$.

**Proof** It is enough to show that $\Gamma^Q(\rho^A, \rho^B) \subseteq \mathcal{B}(\text{range } \rho^A) \otimes \mathcal{B}(\mathcal{H}_n)$. To show this condition we can assume that range $\rho^A$ is a nonzero strict subspace of $\mathcal{H}_m$. By choosing a corresponding orthonormal basis consisting of eigenvectors of $\rho^A$ we can assume that $\rho^A$ is a diagonal matrix whose first $1 \leq \ell < m$ diagonal entries are positive, and whose last $n - \ell$ diagonal entries are zero. Write down $\rho^{AB}$ as a block matrix $[R_{pq}] \in \mathbb{C}^{(mn) \times (mn)}$, were $R_{pq} \in \mathbb{C}^{m \times m}$, $p, q \in [n]$. Then $\text{Tr}_B \rho^{AB} = \sum_{p=1}^n R_{pp} = \rho^A$. As $R_{pp} \geq 0$ we deduce that $\rho^A = [a_{ij}] \geq R_{pp} \geq 0$. As $a_{ii} = 0$ for $i > \ell$ it follows that the $(i, i)$ entry of each $R_{pp}$ is zero. As $\rho^{AB}$ positive semidefinite it follows that the $((p-1)n + i)$th row and column of $\rho^{AB}$ are zero. This proves $\Gamma^Q(\rho^A, \rho^B) \subseteq \mathcal{B}(\text{range } \rho^A) \otimes \mathcal{B}(\mathcal{H}_n)$. Apply the same argument for $\rho^B$ to deduce $\Gamma^Q(\rho^A, \rho^B) \subseteq \mathcal{B}(\text{range } \rho^A) \otimes \mathcal{B}(\text{range } \rho^B)$.

Assume that $\rho^A = |1\rangle\langle 1|$ and $\rho^{AB} \in \Gamma^Q(\rho^A, \rho^B)$. Then $\rho^{AB} = \rho^A \otimes \rho^B$. \[\square\]

More information concerning the partial trace and its properties can be found in a recent work [23].

The following results are used in the proof of Proposition 2.6.
Lemma A.4 Denote by $S_N$ the SWAP operator on $\mathcal{H}_{N^2} := \mathcal{H}_N \otimes \mathcal{H}_N$, and by $S_{n,m}$ and $R_{n,m}$ the following SWAP operators on $\mathcal{H}_{(nm)^2} := \mathcal{H}_n \otimes \mathcal{H}_m \otimes \mathcal{H}_n \otimes \mathcal{H}_m$:

$$S_{n,m}(|x| |u| |y| |v\rangle) = |y| |v\rangle |x| |u\rangle, \quad R_{n,m}(|x| |u| |y| |v\rangle) = |x| |y\rangle |u| |v\rangle.$$ 

(a) Assume that $|i\rangle$, with $i \in [N]$, is an orthonormal basis in $\mathcal{H}_N$. Suppose that

$$\rho = \sum_{i,j,p,q \in [N]} \rho_{i(p,j,q)i}|p\rangle\langle j| q\rangle \in B(\mathcal{H}_N \otimes \mathcal{H}_N).$$

Then $\text{Tr} \ S_N \rho = \sum_{i,p \in [N]} \rho_{p(i,i)}$. 

(b) Assume that

$$\rho^{AB} \in B(\mathcal{H}_n \otimes \mathcal{H}_n), \quad \text{Tr}_B \rho^{AB} = \rho^A \in B(\mathcal{H}_n), \quad \text{Tr}_A \rho^{AB} = \rho^B \in B(\mathcal{H}_n),$$

$$\sigma^{CD} \in B(\mathcal{H}_m \otimes \mathcal{H}_m), \quad \text{Tr}_D \sigma^{CD} = \sigma^C \in B(\mathcal{H}_m), \quad \text{Tr}_C \sigma^{CD} = \sigma^D \in B(\mathcal{H}_m).$$

Then $\tau^{ACBD} := R_{n,m}(\rho^{AB} \otimes \sigma^{CD}) R_{n,m}$ is in $B(\mathcal{H}_{(nm)^2})$. Furthermore

$$\text{Tr}_B \text{Tr} \tau^{ACBD} = \rho^A \otimes \sigma^C, \quad \text{Tr}_A \text{Tr} \tau^{ACBD} = \rho^B \otimes \sigma^D,$$

$$\text{Tr} \ S_{n,m} \tau^{ACBD} = \left( \text{Tr} \ S_n \rho^{AB} \right) \left( \text{Tr} \ S_m \sigma^{CD} \right). \quad (A.2)$$

Proof (a) View $S_N$ and $\rho$ as $N^2 \times N^2$ matrices with entries indexed by the row $(i, p)$ and the column $(j, q)$. Observe that $S_N$ is a symmetric permutation matrix. Then $(S_N \rho)_{(i,p),(j,q)} = \rho_{(p,i),(j,q)}$. The trace of $S_N \rho$ is obtained by summation on the entries $p = q$ and $i = j$.

Clearly, $\tau^{ACBD} \in B(\mathcal{H}_{(nm)^2})$. Assume that

$$\rho^{AB} = \sum_{i_A,i_B,j_A,j_B \in [n]} \rho_{i_A,i_B}(j_A,j_B)|i_B\rangle\langle j_A| \langle j_B|,$$

$$\sigma^{CD} = \sum_{p_C,p_D,q_C,q_D \in [m]} \sigma_{p_C,p_D}(q_C,q_D)|p_C\rangle\langle p_D| \langle q_C| \langle q_D|.$$ 

Then

$$\tau^{ACBD} = \sum_{i_A,i_B,j_A,j_B \in [n]} \rho_{i_A,i_B}(j_A,j_B) \sigma_{p_C,p_D}(q_C,q_D)|i_A\rangle\langle i_B| |p_C\rangle \langle p_D| |q_C\rangle \langle q_D|. $$
Observe next that $\text{Tr}_{BD} \tau^{ACBD}$ is obtained when we sum on $i_B = j_B$ and $p_D = q_D$. Hence

$$\text{Tr}_{BD} \tau^{ACBD} = \sum_{i_A, j_A \in [n]} (\sum_{i_B = 1}^{n} \rho(i_A,i_B)(j_A,i_B))(\sum_{p_D = 1}^{m} \sigma(p_C,p_D)(q_C,p_D))|i_A\rangle|p_C\rangle\langle j_A|\langle q_C|$$

$$= \sum_{i_A, j_A \in [n]} \rho^A_{i_A j_A} \sigma^C_{p_C q_C} |i_A\rangle|p_C\rangle\langle j_A|\langle q_C| = \rho^A \otimes \sigma^C.$$ 

Similarly $\text{Tr}_{AC} \tau^{ACBD} = \rho^B \otimes \sigma^D$. This proves the first line in (A.2).

We now use (a) to compute $\text{Tr}_{S_n,m} \tau^{ACBD}$:

$$\text{Tr}_{S_n,m} \tau^{ACBD} = \sum_{i_A, i_B \in [n]} \rho(i_B,i_A)(i_A,i_B)\sigma(p_D,p_C)(p_C,p_D) = (\text{Tr}_{S_n} \rho^A) (\text{Tr}_{S_m} \sigma^D).$$

This proves the second line in (A.2).

\[\square\]

**Appendix B: Quantum States of a Single Qubit System**

In this Appendix we discuss additional properties of the quantum optimal transport for qubits. Section B.1 provides (Theorem B.1) a closed formula for $T^Q_{CQ}(\rho^A, \rho^B)$ in terms of solutions of the trigonometric equation (B.1). Lemma B.2 shows that this trigonometric equation is equivalent to a polynomial equation of degree at most 6. Section B.2 gives a nice closed formula for the value of QOT for two isospectral qubit density matrices. In Section B.3 we present a simple example where the supremum of the dual SDP problem to QOT is not achieved.

**B.1: A Semi-analytic Formula for the Single-Qubit Optimal Transport**

We begin by introducing a convenient notation for qubits in the $y = 0$ section of the Bloch ball $\Omega_2$—see [4, Sect. 5.2]. Let $O$ denote the orthogonal rotation matrix,

$$O(\theta) = \begin{bmatrix} \cos(\theta/2) & -\sin(\theta/2) \\ \sin(\theta/2) & \cos(\theta/2) \end{bmatrix}, \quad \text{for } \theta \in [0, 2\pi),$$

and define, for $r \in [0, 1]$,

$$\rho(r, \theta) = O(\theta) \begin{bmatrix} r & 0 \\ 0 & 1-r \end{bmatrix} O(\theta)^\top.$$ 

Because of unitary invariance (2.14), the quantum transport problem between two arbitrary qubits $\rho^A, \rho^B \in \Omega_2$ can be reduced to the case $\rho^A = \rho(s, 0)$ and $\rho^B = \rho(r, \theta)$, with three parameters, $s, r \in [0, 1]$ and $\theta \in [0, 2\pi)$. The parameter $\theta$ is the angle
between the Bloch vectors associated with $\rho^A$ and $\rho^B$. With such a parametrization we can further simplify the single-qubit transport problem.

Observe first that if $s \in [0, 1]$ then $\rho^A$ is pure, and if $r \in [0, 1]$ then $\rho^B$ is pure. In any such case an explicit solution of the qubit transport problem is given (6.2).

**Theorem B.1** Let $\rho^A = \rho(s, 0)$, $\rho^B = \rho(r, \theta)$ and assume that $0 < r, s < 1$. Then

$$T_{CQ}^Q(\rho^A, \rho^B) = \max_{\phi \in \Phi(s, r, \theta)} \frac{1}{4} \left( \sqrt{1 + (2s - 1) \cos \phi} - \sqrt{1 + (2r - 1) \cos(\theta + \phi)} \right)^2,$$

where $\Phi(s, r, \theta)$ is the set of all $\phi \in [0, 2\pi)$ satisfying the equation

$$\frac{(2s - 1)^2 \sin^2 \phi}{1 + (2s - 1) \cos \phi} = \frac{(2r - 1)^2 \sin^2(\theta + \phi)}{1 + (2r - 1) \cos(\theta + \phi)}.$$  

(B.1)

**Proof** A unitary $2 \times 2$ matrix $U$ can be parametrized, up to a global phase, with three angles $\alpha, \beta, \phi \in [0, 2\pi)$,

$$U = \begin{bmatrix} e^{i\alpha} & 0 \\ 0 & e^{-i\alpha} \end{bmatrix} O(\phi) \begin{bmatrix} e^{i\beta} & 0 \\ 0 & e^{-i\beta} \end{bmatrix}.$$  

Thus, setting $f(r, \theta; \alpha, \phi) = (U^\dagger \rho(r, \theta) U)_{11}$, we have

$$f(r, \theta; \alpha, \phi) = \frac{1}{2} \left( 1 + (2r - 1)(\cos(\theta) \cos(\phi) + \cos(2\alpha) \sin(\theta) \sin(\phi)) \right).$$

This quantity does not depend on the parameter $\beta$, so we can set $\beta = 0$. Note also that $f(s, 0; \alpha, \phi)$ does not depend on $\alpha$. With $\rho^A = \rho(s, 0)$, $\rho^B = \rho(r, \theta)$, Theorem 5.1 yields

$$T_{CQ}^Q(\rho^A, \rho^B) = \frac{1}{2} \max_{\alpha, \phi \in [0, 2\pi)} \left( \sqrt{f(s, 0; 0, \phi)} - \sqrt{f(r, \theta; \alpha, \phi)} \right)^2.$$

Now, note that the equation $\partial_\alpha f(r, \theta; \alpha, \phi) = 0$ yields the extreme points $\alpha_0 = k\pi/2$, with $k \in \mathbb{Z}$. Since $f(r, \theta; \alpha + \pi, \phi) = f(r, \theta; \alpha, \phi)$ we can take just $\alpha_0 \in [0, \pi/2]$. Consequently,

$$T_{CQ}^Q(\rho^A, \rho^B) = \max_{\phi \in [0, 2\pi)} \{ g_-(s, r, \theta; \phi), g_+(s, r, \theta; \phi) \},$$

where we introduce the auxilliary functions

$$g_\pm(s, r, \theta; \phi) = \frac{1}{4} \left( \sqrt{1 + (2s - 1) \cos \phi} - \sqrt{1 + (2r - 1) \cos(\theta \pm \phi)} \right)^2. \quad \text{(B.2)}$$

But since $g_-(s, r, \theta; 2\pi - \phi) = g_+(s, r, \theta; \phi)$ we can actually drop the $\pm$ index in the above formula. In conclusion, we have shown that it is sufficient to take $U = O(\phi)$ for $\phi \in [0, 2\pi)$ in Formula (5.2).
Finally, it is straightforward to show that the equation \( \partial_\phi g(s, r, \theta; \phi) = 0 \) is equivalent to (B.1). Hence, \( \Phi(s, r, \theta) \) is the set of extreme points, and (B.1) follows.

**Lemma B.2** The Eq. (B.1) has at most six solutions \( \phi \in [0, 2\pi) \) for given \( r, s \in (0, 1), \theta \in [0, 2\pi) \). Moreover there is an open set of \( s, r \in (0, 1), \theta \in [0, 2\pi) \) where there are exactly six distinct solutions.

**Proof** Write \( z = e^{i\phi}, \zeta = e^{i\theta} \). Then

\[
2 \cos \phi = z + \frac{1}{z}, \quad 2i \sin \phi = z - \frac{1}{z},
\]

\[
2 \cos(\theta + \phi) = \zeta z + \frac{1}{\zeta z}, \quad 2i \sin(\theta + \phi) = \zeta z - \frac{1}{\zeta z}.
\]

Thus (B.1) is equivalent to

\[
(1 - 2r)^2 \left[ (2s - 1) \left( z^2 + 1 \right) + 2z \right] \left( \zeta^2 z^2 - 1 \right)^2 \]

\[- \zeta (1 - 2s)^2 \left( z^2 - 1 \right)^2 \left[ (2r - 1) \left( \zeta^2 z^2 + 1 \right) + 2\zeta z \right] = 0. \quad \text{(B.3)}
\]

This a 6th order polynomial equation in the variable \( z \), so it has at most 6 real solutions. Since we must have \( |z| = 1 \), not every complex root of (B.3) will yield a real solution to the original (B.1). Nevertheless, it can be shown that there exist open sets in the parameter space \( s, r \in (0, 1), \theta \in [0, 2\pi) \) on which (B.1) does have 6 distinct solutions.

Observe that if \( \theta = 0 \) and \( s, r \in (0, 1) \) and \( s \neq r \) then two solutions to the equality (B.1) are \( \phi \in \{0, \pi\} \), which means that \( z = \pm 1 \). In this case the equality (B.1) is

\[
\sin^2 \phi \left( \frac{(2s - 1)^2}{1 + (2s - 1) \cos \phi} - \frac{(2r - 1)^2}{1 + (2r - 1) \cos(\phi)} \right) = 0.
\]

As \( \sin^2 \phi = -(1/4)z^{-2}(z^2 - 1)^2 \) we see that \( z = \pm 1 \) is a double root.

Another solution \( \phi \notin \{0, \pi\} \) is given by

\[
\cos \phi = \frac{(2s - 1)^2 - (2r - 1)^2}{(2r - 1)(2s - 1)} = \frac{2(1 - r - s)}{(2r - 1)(2s - 1)}.
\]

Assume that \( r + s = 1 \). Then \( \cos \phi = 0 \), so \( \phi \in \{\pi/2, 3\pi/2\} \). Thus if \( r + s \) is close to 1 we have that \( \phi \) has two values close to \( \pi/2 \) and \( 3\pi/2 \) respectively. Hence in this case we have 6 solutions counting with multiplicities.

We now take a small \( |\theta| > 0 \). The two simple solutions \( \phi \) are close to \( \pi/2 \) and \( 3\pi/2 \). We now need to show that the double roots \( \pm 1 \) split to two pairs of solutions on the unit disc: one pair close to 1 and the other pair close to \( -1 \). Let us consider the pair close to 1, i.e., \( \phi \) close to zero. Then the equation (B.1) can be written in the form

\[
(2s - 1)^2(1 + (2r - 1) \cos(\theta + \phi)) \sin^2 \phi - (2r - 1)^2(1 + (2s - 1) \cos \phi) \sin^2(\theta + \phi) = 0.
\]
Replacing $\sin \phi$, $\sin(\theta + \phi)$ by $\phi$, $\theta + \phi$ respectively we see that the first term gives the equation: $(2s - 1)^2(2r)\phi^2 - (2r - 1)^2 2s(\theta + \phi)^2 = 0$. Then we obtain two possible Taylor series of $\phi$ in terms of $\theta$:

\[
\begin{align*}
\phi_1(\theta) &= \frac{(2r - 1) \sqrt{s} \theta}{(2s - 1) \sqrt{r} - (2r - 1) \sqrt{s}} + \theta^2 E_1(\theta), \\
\phi_2(\theta) &= -\frac{(2r - 1) \sqrt{s} \theta}{(2s - 1) \sqrt{r} + (2r - 1) \sqrt{s}} + \theta^2 E_2(\theta).
\end{align*}
\]

Use the implicit function theorem to show that $E_1(\theta)$ and $E_2(\theta)$ are analytic in $\theta$ in the neighborhood of 0. Hence in this case we have 6 different solutions. \qed

We have thus shown that the general solution of the quantum transport problem of a single qubit with cost matrix $C^Q = \frac{1}{2}(I_4 - S)$ is equivalent to solving a 6th degree polynomial equation with certain parameters. For some specific values of these parameters an explicit analytic solution can be given. This is discussed in the next subsection.

**B.2: Two Isospectral Density Matrices of a Single Qubit**

In view of unitary invariance (2.14) and the results of the previous section we can assume that two isospectral qubits have the following form: $\rho^A = \rho(s, 0)$ and $\rho^B = \rho(s, \theta)$ for some $s \in [0, 1]$ and $\theta \in [0, 2\pi)$.

**Theorem B.3** For any $s \in [0, 1]$ and $\theta \in [0, 2\pi)$ we have

\[
T^Q_{C^Q}(\rho(s, 0), \rho(s, \theta)) = \left(\frac{1}{2} - \sqrt{s(1 - s)}\right) \sin^2(\theta/2).
\] (B.4)

**Proof** Note first that if the states $\rho^A, \rho^B$ are pure, i.e. $s = 0$ or $s = 1$, formula (B.4) gives $T^Q_{C^Q}(\rho(s, 0), \rho(s, \theta)) = \frac{1}{2} \sin^2(\theta/2)$, which agrees with (6.2).

From now on we assume that that $\rho^A, \rho^B$ are not pure. When $r = s$, (B.3) simplifies to the following:

\[
(\xi - 1)(1 - 2s)^2 \left(\xi z^2 - 1\right) \times \\
\times \left[4s(\xi + 1) \left(\xi z^2 + 1\right) z + (2s - 1)(z - 1)^2(\xi z - 1)^2\right] = 0. \tag{B.5}
\]

Equation (B.5) is satisfied when $z = \pm \xi^{-1/2}$. This corresponds to $\phi_0 = -\theta/2$ or $\phi_0' = \pi - \theta/2$. Observe, however, that we have $g(s, s, \theta; \phi_0) = g(s, s, \theta; \phi_0') = 0$, so we can safely ignore $\phi_0, \phi_0' \in \Phi(s, s, \theta)$ in the maximum in (B.1).

Hence, we are left with a 4th order equation

\[
4s(\xi + 1) \left(\xi z^2 + 1\right) z + (2s - 1)(z - 1)^2(\xi z - 1)^2 = 0, \tag{B.6}
\]
which reads

\[(2s - 1)[2 + \cos(\theta + 2\phi) + \cos(\theta)] + 2[\cos(\theta + \phi) + \cos(\phi)] = 0. \quad (B.7)\]

Now, observe that if \(\phi\) satisfies (B.7), then so does \(\phi' = -\phi - \theta\). This translates to the fact that if \(z\) satisfies (B.6), then so does \((z\zeta)^{-1}\). Furthermore, \(g(s, s, \theta; \phi) = g(s, s, \theta; \phi')\). Hence, in the isospectral case we are effectively taking the maximum over just two values of \(\phi\).

Let us now seek an angle \(\phi_1 \in [0, 2\pi)\) such that \(g(s, s, \theta; \phi_1) = \text{righthand side of (B.4)}\). The latter equation reads

\[
\begin{align*}
\left\{(2s - 1)[\cos(\theta + \phi_1) + \cos(\phi_1)] - (2\sqrt{s(1-s)} - 1)(\cos(\theta) - 1) + 2\right\}^2 \\
= 4[(2s - 1)\cos(\phi_1) + 1][(2s - 1)\cos(\theta + \phi_1) + 1].
\end{align*}
\]

In terms of \(z\) and \(\zeta\), the above is equivalent to a 4th order polynomial equation in \(z\), which can be recast in the following form:

\[
\begin{align*}
\left[\zeta(1 - 2s)z^2 + (\zeta + 1)(2\sqrt{s(1-s)} - 1)z - 2s + 1\right]^2 = 0. \quad (B.8)
\end{align*}
\]

Hence, (B.8) has two double roots:

\[
z_{1,2}^\pm = \left[2\zeta(1 - 2s)\right]^{-1}\left\{(\zeta + 1)(1 - 2\sqrt{s(1-s)}) \pm \sqrt{(\zeta + 1)^2(1 - 2\sqrt{s(1-s)})^2 - 4\zeta(1 - 2s)^2}\right\}.
\]

Furthermore, one can check that \(z_{1,2}^- = (z_{1,2}^+)^{-1}\).

Now, it turns out that \(z_{1,2}^\pm\) are also solutions to (B.6), as one can quickly verify using MATHEMATICA [64]. We thus conclude that \(\phi_1, \phi_1' \in \Phi(s, s, \theta)\).

We now divide the polynomial in (B.6) by \((z - z_1^+)(z - z_1^-)\). We are left with the following quadratic equation

\[
\zeta\left[(2s - 1)\left(\zeta z^2 + 1\right) + (\zeta + 1)(2\sqrt{(1-s)s} + 1)z\right] = 0.
\]

Its solutions are

\[
z_{2}^\pm = \left[2\zeta(1 - 2s)\right]^{-1}\left\{(\zeta + 1)(1 + 2\sqrt{s(1-s)}) \pm \sqrt{(\zeta + 1)^2(1 + 2\sqrt{s(1-s)})^2 - 4\zeta(1 - 2s)^2}\right\}.
\]
Again, we have $z_2^- = (\zeta z_2^+)^{-1}$, in agreement with the symmetry argument. Setting $z_2^+ =: e^{i\phi_2}$ and $z_2^- =: e^{i\phi'_2}$ we have $\phi_2, \phi'_2 \in \Phi(s, s, \theta)$. Then we deduce that

$$g(s, s, \theta; \phi_2) = g(s, s, \theta; \phi'_2) = \frac{1}{4} \left[ (1 - 6\sqrt{(1-s)s} - (1 + 2\sqrt{(1-s)s}) \cos(\theta) \right].$$

Finally, we observe that

$$g(s, s, \theta; \phi_1) - g(s, s, \theta; \phi_2) = \sqrt{(1-s)s} \left( 1 + \cos(\theta) \right) \geq 0.$$ 

This shows that, for any $s \in (0, 1), \theta \in [0, 2\pi)$,

$$T^Q_{C_0}(\rho(s, 0), \rho(s, \theta)) = g(s, s, \theta; \phi_1),$$

and (B.4) follows.

Note that $g(s, s, \theta; \phi_2)$ can become negative for certain values of $s$ and $\theta$. This means that for such values the set $\Phi$ of phases defined in Theorem B.1 reads, $\Phi(s, s, \theta) = \{\phi_0, \phi'_0, \phi_1, \phi'_1\}$.

### B.3: An Example Where the Supremum (3.1) is not Achieved

Assume that $m = n = 2, C = C^Q, \rho^A = |0\rangle\langle 0| = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $\rho^B = \mathbb{I}_2/2$. Recall that in such a case, $\Gamma^Q(\rho^A, \rho^B) = \{\rho^A \otimes \rho^B\}$ and

$$\rho^A \otimes \rho^B = \begin{bmatrix} \frac{1}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$ 

Hence $T^Q(\rho^A, \rho^B) = 1/4$. We can easily see that the supremum in (3.1) is not attained in this case. Let $F$ be of the form (5.12). Suppose that there exists $\sigma^A, \sigma^B \in \mathcal{S}(\mathcal{H}_2)$ such that $F \geq 0$ and $T^Q_{C_0}(\rho^A, \rho^B) = \text{Tr}(\sigma^A \rho^A + \sigma^B \rho^B)$. As in the proof of Theorem 3.2 we deduce that $\text{Tr} F(\rho^A \otimes \rho^B) = 0$. Hence the $(1, 1)$ and $(2, 2)$ entries of $F$ are zero. Since $F \geq 0$ it follows that the first and the second row and column of $F$ are zero. Observe next that the $(2, 3)$ and $(3, 2)$ entries of $F$ are $-1/2$. Hence such $\sigma^A, \sigma^B$ do not exist.

Since $\rho^A$ is not positive definite and $\rho^B$ is positive definite, as pointed out in the proof of Proposition 2.4, one can replace $\rho^A$ by $\rho'^A = [1] \in \Omega_1$. Then the dual problem for $\rho'^A, \rho^B$ boils down to

$$\sigma'^A = -a', \quad \sigma^B = \text{diag}(-e, -g), \quad F = \text{diag}(a' + e, a' + g + 1/2) \geq 0,$$

$$\max_{a'+e \geq 0, a'+g+1/2 \geq 0} \left( -a' - \frac{e + g}{2} \right).$$
Then the above maximum is 1/4, achieved for \( a' = -1/2 + t, \ e = 1/2 - t, \ g = -t \) for each \( t \in \mathbb{R} \).

To summarize: the supremum of the dual problem to \( T_{CQ}^{Q}(\rho^A, \rho^B) \) is achieved at \( \text{infinity} \), while the supremum of the dual problem to \( T_{CQ}^{Q}(\rho^{a'}, \rho^B) \) is achieved on an unbounded set.

### Appendix C: Diagonal States of a Qutrit

In this section we provide a closed formula for \( T_{CQ}^{Q}(\rho^A, \rho^B) \) for a large class of classical qutrits, i.e. diagonal matrices \( \rho^A, \rho^B \in \Omega_3 \).

**Theorem C.1** Let \( \mathbf{s} = (s_1, s_2, s_3)^\top, \ \mathbf{t} = (t_1, t_2, t_3)^\top \in \mathbb{R}^3 \) be probability vectors. Then the quantum optimal transport problem for diagonal qutrits is determined by the given formulas in the following cases:

(a)

\[
T_{CQ}^{Q}(\text{diag}(\mathbf{s}), \text{diag}(\mathbf{t})) = \frac{1}{2} \max_{p \in [3]} (\sqrt{s_p} - \sqrt{t_p})^2
\]

if and only if the conditions (5.9) hold for \( n = 3 \).

(b) Suppose that there exists a renaming of 1, 2, 3 by \( p, q, r \) such that

\[
t_r \geq s_p + s_q \text{ and } \quad \text{either } s_p \geq t_p > 0, \ s_q \geq t_q > 0 \text{ or } t_p \geq s_p > 0, \ t_q \geq s_q > 0.
\]

Then

\[
T_{CQ}^{Q}(\text{diag}(\mathbf{s}), \text{diag}(\mathbf{t})) = \frac{1}{2} \left( (\sqrt{s_p} - \sqrt{t_p})^2 + (\sqrt{s_q} - \sqrt{t_q})^2 \right).
\]

(c) Suppose that there exists \( \{p, q, r\} = \{1, 2, 3\} \) such that

\[
s_p > t_q > 0, \quad t_p > s_q > 0, \quad s_q + s_r \geq t_p.
\]

and

\[
1 + \frac{\sqrt{t_q}}{\sqrt{s_q}} - \frac{s_p - t_q}{t_p - s_q} \geq 0, \quad 1 + \frac{\sqrt{s_q}}{\sqrt{t_q}} - \frac{t_p - s_q}{s_p - t_q} \geq 0,
\]

\[
\left( 1 + \frac{\sqrt{t_q}}{\sqrt{s_q}} - \frac{s_p - t_q}{t_p - s_q} \right) \left( 1 + \frac{\sqrt{s_q}}{\sqrt{t_q}} - \frac{t_p - s_q}{s_p - t_q} \right) \geq 1,
\]

\[
\max \left( \frac{s_q}{t_q}, \frac{t_q}{s_q} \right) \geq \max \left( \frac{s_p - t_q}{t_p - s_q}, \frac{t_p - s_q}{s_p - t_q} \right).
\]
Then
\[
T_{CQ}^Q (\text{diag}(s), \text{diag}(t)) = \frac{1}{2} \left( (\sqrt{s_q} - \sqrt{t_q})^2 + (\sqrt{s_p} - t_q - \sqrt{s_p} - s_q)^2 \right). \quad (C.5)
\]

(d) Assume that \( s = (s_1, s_2, 0)^\top, t = (t_1, t_2, t_3)^\top \) are probability vectors. Then
\[
T_{CQ}^Q (\text{diag}(s), \text{diag}(t)) = \begin{cases}
\frac{1}{2} ( (\sqrt{t_1} - \sqrt{t_2})^2 + t_3), & \text{if } s_1 \geq t_2 \text{ and } s_2 \geq t_1, \\
\frac{1}{2} ( (\sqrt{t_1} - \sqrt{s_1})^2 + t_3), & \text{if } s_1 < t_2, \\
\frac{1}{2} ( (\sqrt{t_2} - \sqrt{s_2})^2 + t_3), & \text{if } s_2 < t_1.
\end{cases}
\]
\[(C.6)\]

If \( s = (s_1, s_2, s_3)^\top, t = (t_1, t_2, 0)^\top, \) then formula (C.6) holds after the swapping \( s_i \leftrightarrow t_i. \)

**Proof**

(a) This follows from Theorem 5.3.

(b) Suppose that the condition (C.1) holds. By relabeling the coordinates and interchanging \( s \) and \( t \) if needed we can assume the conditions (C.1) are satisfied with \( p = 1, q = 2, r = 3: \)

\[
s_1 \geq t_1 > 0, \quad s_2 \geq t_2 > 0, \quad t_3 \geq s_1 + s_2.
\]

Hence
\[
X^* = \begin{bmatrix}
0 & 0 & s_1 \\
0 & 0 & s_2 \\
t_1 & t_2 & t_3 -(s_1 + s_2)
\end{bmatrix} \in \Gamma^{cl}(s, t). \quad (C.7)
\]

We claim that the conditions (C.1) yield that \( X^* \) is a minimizing matrix for \( T_{CQ}^Q (\text{diag}(s), \text{diag}(t)) \) as given in (4.5). To show that we use the complementary conditions in Lemma 5.2. Let \( R^* \in \Gamma^Q (\text{diag}(s), \text{diag}(t)) \) be the matrix induced by \( X^* \) of the form described in part (a) of Lemma 4.2. That is, the diagonal entries of \( R^* \) are \( R^*_{(i,j)(i,j)} = x_{ij}^* \) with additional nonnegative entries: \( R^*_{(i,j)(j,i)} = \sqrt{x_{ij}^* x_{ji}^*} \) for \( i \neq j \).

Clearly, \( R^* \) is a direct sums of 3 submatrices of order 1 and 3 of order 2 as above. Let \( F^* \) be defined as in Lemma 5.2 with the following parameters:

\[
a_1^* = \frac{1}{2} \left( \frac{\sqrt{t_1}}{\sqrt{s_1}} - 1 \right), \quad b_1^* = \frac{1}{2} \left( \frac{\sqrt{s_1}}{\sqrt{t_1}} - 1 \right), \\
a_2^* = \frac{1}{2} \left( \frac{\sqrt{t_2}}{\sqrt{s_2}} - 1 \right), \quad b_2^* = \frac{1}{2} \left( \frac{\sqrt{s_2}}{\sqrt{t_2}} - 1 \right), \\
a_3^* = b_3^* = 0.
\]
\[(C.8)\]

We claim that the conditions (C.1) yield that \( F^* \) is positive semidefinite. We verify that the three blocks of size one and the three blocks of size two of \( F^* \) are positive.
semidefinite. The condition \( a_i^* + b_i^* \geq 0 \) for \( i \in [3] \) is straightforward. The conditions for \( M_{12}^* \) and \( M_{13}^* \) are straightforward. We now show that \( M_{12}^* \) is positive semidefinite. First note that as \( s_1 \geq t_1 \) and \( s_2 \geq t_2 \) we get that \( b_1^* \geq 0 \) and \( b_2^* \geq 0 \). Clearly \( a_1^* > -1/2 \) and \( a_2^* > -1/2 \). Hence the diagonal entries of \( M_{12}^* \) are positive. It is left to show that \( \det M_{12}^* \geq 0 \). Set \( u = \sqrt{t_1}/\sqrt{s_1} \leq 1 \) and \( v = \sqrt{s_2}/\sqrt{t_2} \geq 1 \). Then

\[
2(a_1^* + b_2^* + 1/2) = u + v - 1, \quad 2(a_2^* + b_1^* + 1/2) = 1/u + 1/v - 1,
\]

\[
4 \det M_{12}^* = (u + v - 1)(1/u + 1/v - 1) - 1
\]

\[
= (1/(uv))(u + v - 1)(u + v - 1 - uv)
\]

\[
= (1/(uv))(u + v - 1 - u)(v - 1) \geq 0.
\]

We next observe that equalities (5.3) hold. The first three equalities hold as \( x_{11}^* = x_{22}^* = (a_1^* + b_2^*) = 0 \). The equality of \( i = 1, j = 2 \) holds as \( x_{12}^* = x_{21}^* = 0 \). The equalities for \( i = 1, j = 3 \) and \( i = 2, j = 3 \) follow from the following equalities:

\[
x_{13}^*(a_1^* + b_3^* + 1/2) + x_{31}^*(a_3^* + b_1^* + 1/2) = \frac{1}{2}(s_1 \sqrt{t_1} + t_1 \sqrt{s_1}) = \sqrt{s_1 t_1} = \sqrt{x_{13}^* x_{31}^*},
\]

\[
x_{23}^*(a_2^* + b_3^* + 1/2) + x_{32}^*(a_3^* + b_2^* + 1/2) = \frac{1}{2}(s_2 \sqrt{t_2} + t_2 \sqrt{s_2}) = \sqrt{s_2 t_2} = \sqrt{x_{23}^* x_{32}^*}.
\]

Hence \( \text{Tr } R^* F^* = 0 \) and \( X^* \) is a minimizing matrix. Therefore (C.2) holds for \( p = 1, q = 2 \).

(c) Suppose that the condition (C.3) holds. By relabeling the coordinates we can assume the conditions (C.3) are satisfied with \( p = 1, q = 2, r = 3 \):

\[
s_1 > t_2, \quad t_1 > s_2, \quad s_2 + s_3 - t_1 \geq 0.
\]

Hence

\[
X^* = \begin{bmatrix}
0 & t_2 & s_1 - t_2 \\
t_1 - s_2 & 0 & 0 \\
0 & s_2 + s_3 - t_1
\end{bmatrix} \in \Gamma^{cl}(s, t). \tag{C.9}
\]

We claim that the conditions (C.4) yield that \( X^* \) is a minimizing matrix for \( T_{C,Q}(\text{diag}(s), \text{diag}(t)) \) as given in (4.5). To show this we use the complementary conditions in Lemma 5.2. Let \( R^* \in \Gamma^{Q}(\text{diag}(s), \text{diag}(t)) \) be the matrix induced by \( X^* \) of the form described in part (a) of Lemma 4.2. Recall that \( R^* \) is a direct sum of 3 submatrices of order 1 and 3 of order 2 as above. Let \( F^* \) correspond to

\[
a_1^* = \frac{1}{2} \left( \frac{\sqrt{t_1} - s_2}{\sqrt{s_1} - t_2} - 1 \right), \quad a_2^* = \frac{1}{2} \left( \frac{\sqrt{t_2} - s_1}{\sqrt{s_2} - t_1} - \frac{s_1 - t_2}{t_2 - s_1} \right), \quad a_3^* = 0, \tag{C.10}
\]

\[
b_1^* = \frac{1}{2} \left( \frac{\sqrt{t_2} - s_1}{\sqrt{s_1} - s_2} - 1 \right), \quad b_2^* = \frac{1}{2} \left( \frac{\sqrt{t_2} - s_1}{\sqrt{s_1} - t_2} - \frac{t_1 - s_2}{s_1 - t_2} \right), \quad b_3^* = 0.
\]
We claim that (C.4) yield that $F^*$ is positive semidefinite. We verify that the three blocks of size one and the three blocks of size two matrices of $F^*$ are positive semidefinite. The condition $a_1^* + b_1^* \geq 0$ is straightforward. To show the condition $a_2^* + b_2^*_2 \geq 0$ we argue as follows. Let

$$u = \frac{\sqrt{t_1}}{\sqrt{s_1}}, \quad v = \sqrt{\frac{s_1 - t_2}{t_2 - s_1}}.$$

Then $2(a_2^* + b_2^*_2) = u + 1/u - (v + 1/v).$ The fourth condition of (C.4) is max $(u, 1/u) \geq$ max $(v, 1/v).$ As $w + 1/w$ increases on $[1, \infty)$ we deduce that $a_2^* + b_2^*_2 \geq 0.$ Clearly $a_3^* + b_3^* = 0.$ We now show that the matrices $(5.5)$ are positive semidefinite, where the last three inequalities follow from the first three inequalities of (C.4):

$$2(a_1^* + b_2^* + 1/2) = \frac{\sqrt{s_2}}{\sqrt{t_2}} > 0, \quad 2(a_2^* + b_1^* + 1/2) = \frac{\sqrt{t_2}}{\sqrt{s_2}} > 0,$$

$$(a_1^* + b_2^* + 1/2)(a_2^* + b_1^* + 1/2) - 1/4 = 0,$$

$$2(a_1^* + b_3^* + 1/2) = \frac{\sqrt{t_1} - s_2}{\sqrt{s_1} - t_2} > 0, \quad 2(a_3^* + b_1^* + 1/2) = \frac{\sqrt{s_1} - t_2}{\sqrt{t_1} - s_2} > 0,$$

$$(a_1^* + b_3^* + 1/2)(a_3^* + b_1^* + 1/2) - 1/4 = 0,$$

$$2(a_2^* + b_3^* + 1/2) = \frac{s_2}{\sqrt{t_2}} - \sqrt{\frac{t_1 - s_2}{s_1 - t_2}} + 1 \geq 0,$$

$$2(a_3^* + b_2^* + 1/2) = \frac{\sqrt{t_2}}{\sqrt{s_2}} - \sqrt{\frac{s_1 - t_2}{t_1 - s_2}} + 1 \geq 0,$$

$$(a_2^* + b_3^* + 1/2)(a_3^* + b_2^* + 1/2) - 1/4 \geq 0.$$

Moreover, the conditions (5.3) hold: as $x_{11}^* = x_{22}^* = a_3^* + b_3^* = 0$ the first three conditions of (5.3) hold, and as $x_{23}^* = x_{32}^* = 0$ the second conditions of (5.3) for $p = 2, q = 3$ trivially hold. The other two conditions follow from the following equalities:

$$x_{12}(a_1^* + b_2^* + 1/2) + x_{21}(a_2^* + b_1^* + 1/2) - \sqrt{x_{12}x_{21}^*}$$

$$= t_2 \frac{\sqrt{s_2}}{2\sqrt{t_2}} + s_2 \frac{\sqrt{t_2}}{2\sqrt{s_2}} - \sqrt{t_2}s_2 = 0,$$

$$x_{13}(a_1^* + b_3^* + 1/2) + x_{31}(a_3^* + b_1^* + 1/2) - \sqrt{x_{13}x_{31}^*}$$

$$= (s_1 - t_2) \frac{\sqrt{t_1 - s_2}}{2\sqrt{s_1 - t_2}} + s_2 \frac{\sqrt{t_2}}{2\sqrt{s_2}} - \sqrt{(s_1 - t_2)(t_1 - s_2)} = 0.$$
The probability vectors $\hat{F}$ are the solutions to the following optimization problem:

$$
\min_{W \in \Gamma} \langle X, W \rangle
$$

where $\Gamma$ is the set of all matrices with non-negative entries and row sums equal to 1. The optimal transport map $\hat{F}$ is obtained by solving this problem.

Using continuity arguments it is enough to consider the case where $Q$ is a rank-one matrix. Using the continuity argument we may assume that $s$ is a minimizing matrix in $\Gamma_{1,3}^c$. Clearly, $\gamma_{1,3}^c(Y)$ is obtained from $\gamma_{1,3}^c(s, t)$ by deleting the third row in each matrix in $\Gamma_{1,3}^c(s, t)$. Proposition 2.4 yields that

$$
\gamma_{1,3}^c(s, t) = \gamma_{1,3}^c(s', t).
$$

(See Lemma 4.3 for the definition of $\gamma_{1,3}^c$.) We use now the minimum characterization of $\gamma_{1,3}^c(s', t)$ given in (4.5). Assume that the minimum is achieved for $X^* = [x_{i,l}^*] \in \Gamma_{1,3}^c(s', t)$.

Let $Y = [x_{i,l}^*]$, $i, l \in [2]$. Suppose first that $Y = 0$. Then $t_1 = t_2 = 0$ and $t_3 = 1$. So $\gamma(t)$ is a rank-one matrix and $\gamma(\gamma(s), \gamma(t)) = 0$. The equality (6.2) yields that $\gamma_{1,3}^c(s, t) = 1$. Clearly, $s_1 \geq t_2 = 0, s_2 \geq t_1 = 0$. Hence (C.6) holds.

Suppose second that $Y \neq 0$. Then $t_1 + t_2$, the sum of the entries of $Y$, is positive. Using continuity arguments it is enough to consider the case $t_1, t_2, t_3 > 0$. Denote by $\Gamma'$ the set of all matrices $X = [x_{i,l}] \in \Gamma_{1,3}^c(s', t)$ such that $x_{i,3} = x_{i,3}^*$ for $i = 1, 2$. Let $f$ be defined by (4.8). Clearly, $\min_{A \in \Gamma'} f(A) = f(Y)$. We now translate this minimum to the minimum problem we studied above.

Let $Z = \frac{1}{t_1 + t_2} Y$. The vectors corresponding to the row sums and the column sums $Z$ are the probability vectors $\hat{s} = (\hat{s}_1, \hat{s}_2)^T$ and $\hat{t} = \frac{1}{t_1 + t_2} (t_1, t_2)^T$ respectively. Consider the minimum problem $\min_{W \in \Gamma_{1,3}^c(s, t)} f(W)$. The proof of Lemma 4.5 yields that this minimum is achieved at $W^*$ which has at least one zero diagonal element. Hence $Y$ has at least one zero diagonal element.

Assume first that $Y$ has two zero diagonal elements. Then $X^* = \begin{bmatrix} 0 & t_2 & s_1 - t_2 \\ t_1 & 0 & s_2 - t_1 \end{bmatrix}$. This corresponds to the first case of (C.6). It is left to show that $X^*$ is a minimizing matrix. Using the continuity argument we may assume that $s_1 > t_2, s_2 > t_1$. Let $B \in \mathbb{R}^{2 \times 3}$ be a nonzero matrix such that $X^* + cB \in \Gamma_{1,3}^c(s', t)$ for $c \in [0, \varepsilon]$ for some small positive $\varepsilon$. Then $B = \begin{bmatrix} a & -b & -a + b \\ -a & b & a - b \end{bmatrix}$, where $a, b \geq 0$ and $a^2 + b^2 > 0$. It is clear that $f(X^*) < f(X + cB)$ for each $c \in [0, \varepsilon]$. This proves the first case of (C.6).

Assume second that $x_{i,1}^* = 0$ and $x_{i,2}^* > 0$. Observe that $x_{i,1}^* = t_1 > 0$. We claim that $x_{i,3}^* = 0$. Indeed, suppose that it is not the case. Let $B = \begin{bmatrix} 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}$. Then $X^* + cB \in \Gamma_{1,3}^c(s', t)$ for $c \in [0, \varepsilon]$ for some positive $\varepsilon$. Clearly $f(X^* + cB) < f(X^*)$
for $c \in (0, \varepsilon]$. Thus contradicts the minimality of $X^*$. Hence $x^*_{13} = 0$. Therefore $X^* = \begin{bmatrix} 0 & s_1 & 0 \\ t_1 & t_2 - s_1 & t_3 \end{bmatrix}$. This corresponds to the second case of (C.6).

The third case is when $x^*_{11} > 0$ and $x^*_{22} = 0$. We show, as in the second case, that $x^*_{23} = 0$. Then $X^* = \begin{bmatrix} t_1 - s_2 & t_2 & t_3 \\ s_2 & 0 & 0 \end{bmatrix}$. This corresponds to the third case of (C.6).

The case $s = (s_1, s_2, s_3)^T$, $t = (t_1, t_2, 0)^T$ is completely analogous, hence the proof is complete. □

Basing on the numerical studies we conjecture that the cases (a)–(d) exhaust the parameter space $\Pi_3 \times \Pi_3$. Nevertheless, we include for completeness an analysis of the quantum optimal transport $T^Q_{CQ}(\text{diag}(s), \text{diag}(t))$ under the assumption that this is not the case. The employed techniques might prove useful when studying more general qutrit states or diagonal ququarts.

**Proposition C.2** Let $O \subseteq \Pi_3 \times \Pi_3$ be the set of pairs $s, t$, which do not meet neither of conditions (a)–(d) from Theorem C.1. Suppose that $O$ is nonempty. Then each minimizing $X^*$ in the characterization (4.5) of $T^Q_{CQ}(\text{diag}(s), \text{diag}(t))$ has zero diagonal. Let $O' \subseteq O$ be an open dense subset of $O$ such that for each $(s, t) \in O'$ and each triple $\{i, j, k\} = [3]$ the inequalities $s_p \neq t_p$ and $s_p + s_q \neq t_r$ hold. Assume that $(s, t) \in O'$. The set of matrices in $\Gamma^{cl}(s, t)$ with zero diagonal is an interval spanned by two distinct extreme points $E_1, E_2$, which have exactly five positive off-diagonal elements. Let $Z(u) = uE_1 + (1-u)E_2$ for $u \in [0, 1]$. Then the minimum of the function $f(Z(u))$, $u \in [0, 1]$, where $f$ is defined by (4.8), is attained at a unique point $u^* \in (0, 1)$. The point $u^*$ is the unique solution in the interval $(0, 1)$ to a polynomial equation of degree at most 12. The matrix $X^* = Z(u^*)$ is the minimizing matrix for the second minimum problem in (4.5), and $T^Q_{CQ}(\text{diag}(s), \text{diag}(t)) = f(X^*)$.

**Proof** Assume first that the set $O \subseteq \Pi_3 \times \Pi_3$ is nonempty and satisfies the conditions (i)-(iv). Combine Theorem 5.3 with part (a) of the theorem to deduce that if the conditions (5.9) do hold for $n = 3$ then

$$T^Q_{CQ}(\text{diag}(s), \text{diag}(t)) > \max_{p \in [3]} \frac{1}{2} \left( \sqrt{s_p} - \sqrt{t_p} \right)^2. \quad (C.11)$$

In view of our assumption the above inequality holds. We first observe that $s_p \neq t_p$ for each $p \in [3]$. Assume to the contrary that $s_p = t_p$. Without loss of generality we can assume that $s_3 = t_3$. Assume that in addition $s_q = t_q$ for some $q \in [2]$. Then $s = t$ and

$$T^Q_{CQ}(\text{diag}(s), \text{diag}(t)) = \frac{1}{2} \max_{p \in [3]} \left( \sqrt{s_p} - \sqrt{t_p} \right)^2 = 0$$

This contradicts (C.11). Hence there exists $q \in [2]$ such that $s_q > t_q$ for $q \in [2]$. Without loss of generality we can assume that $s_2 > t_2$, therefore $s_1 < t_1$, as $s_1 + s_2 = t_1 + t_2 = 1 - s_3 = 1 - t_3$. Hence for $Y = \begin{bmatrix} s_1 & 0 \\ t_1 - s_1 & t_2 \end{bmatrix}$ we have $X = Y \oplus [s_3] \in \Gamma^{cl}(s, t)$. □
Recall that $s, t > 0$. We replace $Y$ by $Y^* = Y + u^* \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ such that $u^* > 0$, $Y^* \geq 0$ and one of the diagonal elements of $Y^*$ is zero. By relabeling $\{1, 2\}$ if necessary we can assume that $Y^* = \begin{bmatrix} 0 & s_1 \\ t_1 & t_2 - s_1 \end{bmatrix}$ so $t_2 \geq s_1$ and $X^* = Y^* \oplus [s_3] \in \Gamma^{cl}(s, t)$. The minimal characterization (4.5) of $T_{C_0}^Q(\text{diag}(s), \text{diag}(t))$ yields

$$T_{C_0}^Q(\text{diag}(s), \text{diag}(t)) \leq f(X^*) = \frac{1}{2}(\sqrt{s_1} - \sqrt{t_1})^2.$$ 

This contradicts (C.11).

As $s, t > 0$ there exists a maximizing matrix $F^*$ to the dual problem of the form given by Lemma 5.2. Let $X^*$ be the corresponding minimizing matrix. We claim that $X^*$ has zero diagonal. Assume first that $X^*$ has a positive diagonal. Then the arguments in part (b) of Lemma 5.2 yield that $X^*$ is a symmetric matrix. Thus $s = t$, and this contradicts (C.11).

Assume second that $X^*$ has two positive diagonal entries. By renaming the indices we can assume that $x^*_{11} = 0, x^*_{22}, x^*_{33} > 0$. Part (b) of Lemma 5.2 and the arguments of its proof yield that we can assume that $a^*_3 = b^*_3 = 0$. Let $u^* = a^*_1 + 1/2, v^* = b^*_1 + 1/2$. As $M^*_{12}$ is positive semidefinite we have the inequalities: $u^* \geq 0, v^* \geq 0, u^*v^* \geq 1/4$. Hence $x^* > 0, y^* > 0$. Recall that $F^*$ is a maximizing matrix for the dual problem (3.1). Hence

$$T_{C_0}^Q\left(\text{diag}(s), \text{diag}(t)\right) = -(u^* - 1/2)s_1 - (v^* - 1/2)t_1$$

$$= -u^*s_1 - v^*t_1 + (s_1 + t_1)/2$$

$$\leq -u^*s_1 - t_1/(4u^*) + (s_1 + t_1)/2$$

$$\leq -\sqrt{s_1t_1} + (s_1 + t_1)/2 = (\sqrt{s_1} - \sqrt{t_1})^2/2.$$ 

This contradicts (C.11).

We now assume that $X^*$ has one positive diagonal entry. Be renaming the indices 1, 2, 3 we can assume that $x^*_{11} = x^*_{22} = 0, x^*_{33} > 0$. The conditions (5.3) yield that $a^*_3 + b^*_3 = 0$. Since we can choose $b^*_3 = 0$ we assume that $a^*_3 = b^*_3 = 0$.

Let us assume, case (A1), that $X^*$ has six positive off-diagonal entries. We first claim that either $x^*_{13} = x^*_{31}$ or $x^*_{23} = x^*_{32}$. (Those are equivalent conditions if we interchange the indices 1 and 2.) We deduce these conditions and an extra condition using the second conditions of (5.4). First we consider $x^*_{12}, x^*_{13}, x^*_{22}, x^*_{33}$, that is $i = p = 3, j = 1, q = 2$. By replacing these entries by $x^*_{12} - v, x^*_{13} + v, x^*_{22} + v, x^*_{33} - v$ we obtain the equalities

$$1 + x = y + z, \quad x = \sqrt{x^*_{22}} \quad y = \sqrt{x^*_{31}} \quad z = \sqrt{x^*_{32}}.$$ 

\[ \square \]
Second we consider \( x_{21}^*, x_{23}^*, x_{31}^*, x_{33}^* \). By replacing these entries by \( x_{21}^* - v, x_{23}^* + v, x_{31}^* + v, x_{33}^* - v \) we obtain the equality:

\[
1 + \frac{1}{x} = \frac{1}{z} + \frac{1}{y}.
\]

Multiply the first and the second equality to deduce

\[
x + \frac{1}{x} = u + \frac{1}{u}, \quad u = \frac{y}{x} \Rightarrow \text{either } x = u \text{ or } x = \frac{1}{u}.
\]

Assume first that \( x = u = y/z \). Substitute that into the first equality to deduce that \( z = 1 \), which implies that \( x_{23}^* = x_{32}^* \). Similarly, if \( x = 1/u \) we deduce that \( y = 1 \), which implies that \( x_{13}^* = x_{31}^* \). Let us assume for simplicity of exposition that \( x_{23}^* = x_{32}^* \). Let \( X(w) \) be obtained from \( X \) by replacing \( x_{22}^* = 0, x_{23}^*, x_{32}^*, x_{33}^* \) with \( x_{22}^* + w, x_{23}^* - w, x_{32}^* - w, x_{33}^* + w \) for \( 0 < w < x_{23}^* \). Then \( X(w) \) is a minimizing matrix and has two positive diagonal entries. This contradicts our assumption that \( X^* \) has only one positive diagonal entry.

We now consider the case (A2) that \( x_{ii}^* = 0 \) for some \( i \neq j \). Part (a) of Lemma 5.2 yields that \( x_{ii}^* = 0 \). We claim that all four off-diagonal entries are positive. Assume to the contrary that \( x_{pq}^* = 0 \) for some \( p \neq q \) and \( \{ p, q \} \neq \{i, j \} \). Then \( x_{ij}^* = 0 \). As \( s, t > 0 \) we must have that \( x_{12}^*, x_{21}^* > 0 \) and all four other off-diagonal entries are zero. But then \( s_1 = t_2, t_1 = s_2, s_3 = t_3 \). This is impossible since we showed that \( s_3 \neq t_3 \). Hence \( X^* \) has exactly four positive off-diagonal entries.

Let us assume first that \( x_{12}^* = x_{21}^* = 0 \). Then \( X^* \) is of the form given by (C.7), where \( t_3 > s_1 + s_2 \). We now recall again the conditions (5.3). As we already showed, we can assume that \( a_3^* = b_3^* = 0 \). As \( x_{11}^* = x_{22}^* = 0 \) all of the first three conditions of (5.3) hold. As \( x_{12}^* = x_{21}^* = 0 \) the second condition of (5.3) holds trivially for \( i = 1, j = 2 \). The conditions for \( i = 1, j = 3 \) and \( i = 2, j = 3 \) are

\[
\begin{align*}
s_1(a_1^* + 1/2) + t_1(b_1^* + 1/2) &= \sqrt{s_1 t_1}, \\
s_2(a_2^* + 1/2) + t_2(b_2^* + 1/2) &= \sqrt{s_2 t_2}.
\end{align*}
\]

We claim that (C.8) holds. Using the assumption that \( \text{det } M_{13}^* \geq 1/4 \) and the inequality of arithmetic and geometric means we deduce that \( \text{det } M_{13}^* = 1/4 \). Hence

\[
\begin{align*}
a_1^* + 1/2 &= u, & b_1^* + 1/2 &= 1/(4u), \quad \text{for some } u > 0, \\
s_1 u + t_1/(4u)t_1 &\geq \sqrt{s_1 t_1}.
\end{align*}
\]

Equality holds if and only if \( u = \sqrt{t_1}/(2\sqrt{s_1}) \). This shows the first equality in (C.8). The second equality in (C.8) is deduced similarly. We now show that the conditions (C.1) hold for \( i = 1, j = 2, k = 3 \). As \( t_3 > s_1 + s_2 \) the first condition of (C.8) holds. We use the conditions that \( M_{12}^* \) is positive semidefinite. Let \( u = \sqrt{t_1}/\sqrt{s_1}, v = \sqrt{s_2}/\sqrt{t_2} \).
Then the arguments of the proof of part (b) yield

\[ 2(a_1^* + b_2^* + 1) = u + v - 1 > 0, \quad 2(a_2^* + b_1^* + 1) = (1/u + 1/v - 1) > 0, \]
\[ 4 \det M_{12}^* = (1/(uv))(1-u)(v-1). \]

So either \( u \geq 1 \) and \( v \leq 1 \), or \( u \leq 1 \) and \( v \geq 1 \). Hence (C.1) holds for \( i = 1, j = 2, k = 3 \). This contradicts our assumption that (C.1) does not hold.

Let us assume second that \( x_1^* > 0, x_2^* > 0 \). Then either \( x_1^* = x_2^* = 0 \) or \( x_1^* = x_3^* = 0 \). By relabeling \( 1, 2 \) we can assume that \( x_1^* = x_3^* = 0 \). Hence \( X^* \) is of the form (C.9), where \( s_1 > t_2 > 0, t_1 > s_2 > 0, s_2 + s_3 > t_1 \). Hence the conditions (C.3) are satisfied with \( i = 1, j = 2, k = 3 \). We now obtain a contradiction by showing that the conditions (C.4) are satisfied. This is done using the same arguments as in the previous case as follows. First observe that the second nontrivial conditions of (5.3) are:

\[ t_2(a_1^* + b_2^* + 1/2) + s_2(a_2^* + b_1^* + 1/2) = \sqrt{s_2 t_2}, \]
\[ (s_1 - t_2)(a_1^* + 1/2) + t_1(b_1^* + 1/2) = \sqrt{(s_1 - t_2)(t_1 - s_2)}. \]

As in the previous case we deduce that

\[ a_1^* + b_2^* + 1/2 = \sqrt{s_2/(2\sqrt{t_2})}, \]
\[ a_1^* + 1/2 = \sqrt{t_1 - s_2/(2\sqrt{t_1 - t_2})}, \]
\[ b_1^* + 1/2 = \sqrt{s_1 - t_2/(2\sqrt{t_1 - s_2})}. \]

Hence (C.10) holds. We now recall the proof of part (c) of the theorem. We have thus shown that the minimizing matrix \( X^* \) has zero diagonal.

We now show that \( O \) is an open set. Clearly, the set of all pairs of probability vectors \( O_1 \subset \Pi_3 \times \Pi_3 \) such that at least one of them has a zero coordinate is a closed set. Let \( O_2, O_3, O_4 \subset \Pi_3 \times \Pi_3 \) be the sets which satisfy the conditions (a), (b),(c) of the theorem respectively. It it straightforward to show: \( O_2 \) is a closed set, and Closure(\( O_3 \)) \( \subset \) (\( O_3 \cup O_1 \)). We now show that Closure(\( O_4 \)) \( \subset \) \( O_2 \cup O_4 \cup O_1 \). Indeed, assume that we have a sequence \( (s_l, t_l) \in O_4, l \in \mathbb{N} \) that converges to \((s, t)\). It is enough to consider the case where \( s, t > 0 \). Again we can assume for simplicity that each \((s_l, t_l)\) satisfies the conditions (C.3) and (C.4) for \( i = 1, j = 2, k = 3 \). Then we deduce that the limit of the minimizing matrices \( X_l^* \) is of the form (C.9). Hence \( \lim_{l \to \infty} X_l^* = X^* \), where \( X^* \) is of the form (C.9). Also \( X^* \) is a minimizing matrix for \( T_{C_0}^Q(\text{diag}(s_l), \text{diag}(t_l)) \). Recall that \( s_2, t_2 > 0 \). If \( s_1 - t_2 > 0, t_1 - s_2 > 0 \) then \((s_l, t_l) \in O_4 \). So assume that \( (s_1 - t_2)(t_1 - s_2) = 0 \). As \( X^* \) is minimizing \( T_{C_0}^Q(\text{diag}(s_l), \text{diag}(t_l)) \) and \( s, t > 0 \), part (a) of Lemma 5.2 yields that \( s_1 = t_2, t_1 = s_2 \). Hence \( s_3 = t_3 \). As \( X^* \) is minimizing \( T_{C_0}^Q(\text{diag}(s_l), \text{diag}(t_l)) \) we get that \( T_{C_0}^Q = \frac{1}{2}(\sqrt{s_2} - \sqrt{t_2})^2 \). Hence \((s, t) \in O_2 \). This shows that \( O_1 \cup O_2 \cup O_3 \cup O_4 \) is a closed set. Therefore \( O = \Pi_3 \times \Pi_3 \backslash (O_1 \cup O_2 \cup O_3 \cup O_4) \) is an open set. If \( O \) is an empty set then proof of the theorem is concluded.
Assume that \( O \) is a nonempty set. Let \( O' \subset O \) be an open dense subset of \( O \) such that for each \((s, t) \in O'\) and each triple \(\{p, q, r\} = [3]\) the inequality \(s_p \neq t_q\) and \(s_p + s_q \neq t_r\) hold.

Assume that \((s, t) \in O'\). Let \(\Gamma^c_0(s, t)\) be the convex subset of \(\Gamma^c(s, t)\) of matrices with zero diagonal. We claim that any \(X \in \Gamma^c_0(s, t)\) has at least 5 nonzero entries. Indeed, suppose that \(X \in \Gamma^c_0(s, t)\) has two zero off-diagonal entries. As \(s, t > 0\) they cannot be in the same row or column. By relabeling the rows we can assume that the two zero elements are in the first and the second row. Suppose first that \(x_{12}^* = x_{23}^* = 0\). Then \(X = \begin{bmatrix} 0 & s_1 & 0 \\ s_2 & 0 & 0 \\ t_1 - s_2 & t_2 & 0 \end{bmatrix}\). Thus \(s_1 = t_3\) which is impossible. Assume now that \(x_{12}^* = x_{21}^* = 0\). Then \(s_1 + s_2 = t_3\) which is impossible. All other choices also are impossible.

We claim that \(\Gamma^c_0(s, t)\) is spanned by two distinct extreme points \(E_1, E_2\), which have exactly five positive off-diagonal elements. Suppose first that there exists \(X \in \Gamma^c_0(s, t)\) which has six positive off-diagonal elements. Let

\[
B = \begin{bmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix}.
\]

Then all matrices in \(\Gamma^c_0(s, t)\) are of the form \(X^* + uB, u \in [u_1, u_2]\) for some \(u_1 < u_2\). Consider the matrix \(E_1 = X^* + u_1B\). It has at least one zero off-diagonal entry hence we conclude that \(E_1\) has exactly five off-diagonal positive elements. Similarly \(E_2 = X + u_2B\) has five positive off-diagonal elements. Assume now that \(E \in \Gamma^c_0(s, t)\) has five positive off-diagonal elements. Hence there exits a small \(u > 0\) such that either \(E + uB\) or \(E - uB\) has six positive off-diagonal elements. Hence \(\Gamma^c_0(s, t)\) contains a matrix with six positive diagonal elements. Therefore \(\Gamma^c_0(s, t)\) is an interval spanned by \(E_1 \neq E_2 \in \Gamma^c_0(s, t)\), where \(E_1\) and \(E_2\) have five positive off-diagonal elements. Part (a) of Lemma 5.2 yields that \(X^*\) has six positive off-diagonal elements. Consider \(E_1\) and assume that the \((1, 2)\) entry of \(E_1\) is zero. Then

\[
E_1 = \begin{bmatrix} 0 & 0 & s_1 \\ s_1 + s_2 - t_3 & 0 & t_3 - s_1 \\ s_3 - t_2 & t_2 & 0 \end{bmatrix}.
\]

As \(f(E_1 + uB)\) is strictly convex on \([0, u_3]\), there exists a unique \(u^* \in (0, u_3)\) which satisfies the equation

\[
\sqrt{\frac{s_1 + s_2 - t_3 - u}{u}} - \frac{\sqrt{s_1 - u}}{\sqrt{u}} + \frac{\sqrt{u}}{\sqrt{s_1 + s_2 - t_3 - u}} - \frac{\sqrt{s_1 - u}}{\sqrt{s_3 - t_2 + u}} + \frac{\sqrt{s_3 - t_2 + u}}{\sqrt{s_1 - u}} - \frac{\sqrt{u}}{\sqrt{s_3 - t_2 + u}} + \frac{\sqrt{s_3 - s_1 + u}}{\sqrt{t_3 - s_1 + u}} + \frac{\sqrt{t_2 - u}}{\sqrt{t_2 - u}} = 0.
\]
It is not difficult to show that the above equation is equivalent to a polynomial equation of degree at most 12 in \( u \). Indeed, group the six terms into three groups, multiply by the common denominator, and pass the last group to the other side of the equality to obtain the equality:

\[
\sqrt{(s_1 - u)(s_3 - t_2 + u)(t_3 - s_1 + u)(t_2 - u)}(2u + t_3 - s_1 - s_2)
+ \sqrt{u(s_1 + s_2 - t_3 - u)(t_3 - s_1 + u)(t_2 - u)(2u + t_3 - s_1 - s_2)(2u + s_3 - s_1 - t_2)}
= \sqrt{u(s_1 + s_2 - t_3 - u)(s_3 - t_2 + u)(t_3 - s_1 + u)(-2u + s_1 + t_2 - t_3)}.
\]

Raise this equality to the second power. Put all polynomial terms of degree 6 on the left hand side, and the one term with a square radical on the other side. Raise to the second power to obtain a polynomial equation in \( u \) of degree at most 12. Hence \( X^* = E_1 + u^*B \). This completes the proof of (e). \( \square \)

References

1. Agredo, J., Fagnola, F.: On quantum versions of the classical Wasserstein distance. Stochastics 89, 910 (2017)
2. Altschuler, J., Weed, J., Rigollet, P.: Near-linear time approximation algorithms for optimal transport via Sinkhorn iteration. In: NIPS’17: Proceedings of the 31st International Conference on Neural Information Processing Systems, pp. 1961–1971 (2017)
3. Arjovsky, M., Chintala, S., Bottou, L.: Wasserstein generative adversarial networks. In: Proceedings of the 34th International Conference on Machine Learning, PMLR, vol. 70, p. 214 (2017)
4. Bengtsson, I., Życzkowski, K.: Geometry of Quantum States, 2nd edn. Cambridge University Press, Cambridge (2017)
5. Bhatia, R., Gaubert, S., Jain, T.: Matrix versions of the Hellinger distance. Lett. Math. Phys. 109, 1777–1804 (2019)
6. Biamonte, J., Wittek, P., Pancotti, N., Rebentrost, P., Wiebe, N., Lloyd, S.: Quantum machine learning. Nature 549, 195 (2017)
7. Biane, P., Voiculescu, D.: A free probability analogue of the Wasserstein distance on the trace-state space. Geom. Funct. Anal. 11, 1125 (2001)
8. Bigot, J., Gouet, R., Klein, T., López, A.: Geodesic PCA in the Wasserstein space by convex PCA. Ann. Inst. H. Poincaré Probab. Stat. 53, 1–26 (2017)
9. Bistroń, R., Eckstein, M., Życzkowski, K.: Monotonicity of the quantum 2-Wasserstein distance. J. Phys. A 56, 095301 (2023)
10. Bonneel, N., van de Panne, M., Paris, S., Heidrich, W.: Displacement interpolation using Lagrangian mass transport. ACM Trans. Graph. 30, 158 (2011)
11. Brandão, F.G.S.L., Svore, K.: Quantum speed-ups for solving semidefinite programs. In: 2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS), pp. 415–426
12. Braunstein, S.L., Caves, C.M.: Statistical distance and the geometry of quantum states. Phys. Rev. Lett. 72, 3439 (1994)
13. Caglioti, E., Golse, F., Paul, T.: Quantum optimal transport is cheaper. J. Stat. Phys. 181, 149 (2020)
14. Carlen, E.A., Maas, J.: Non-commutative calculus, optimal transport and functional inequalities in dissipative quantum systems. J. Stat. Phys. 178, 319 (2020)
15. Chakrabarti, S., Huang, Y., Li, T., Feizi, S., Wu, X.: Quantum Wasserstein generative adversarial networks. In: 33rd Conference on Neural Information Processing Systems (NeurIPS 2019), Vancouver, Canada, arXiv:1911.00111
16. Chen, Y., Gangbo, W., Georgiou, T.T., Tannenbaum, A.: On the matrix Monge-Kantorovich problem. Eur. J. Appl. Math. 31, 574 (2020)
17. Cook, W.J., Cunningham, W.H., Pulleyblank, W.R., Schrijver, A.: Combinatorial Optimization. Wiley, New York (1998)
18. Cuturi, M.: Sinkhorn distances: lightspeed computation of optimal transport. In: Burges, C.J.C., Bottou, L., Welling, M., Ghahramani, Z., Weinberger, K.Q. (eds.) Advances in Neural Information Processing Systems, vol. 26, pp. 2292–2300. Curran Associates Inc., New York (2013)

19. Datta, N., Rouzé, C.: Relating relative entropy, optimal transport and Fisher information: a quantum HWI inequality. Ann. H. Poincaré 21, 2115 (2020)

20. De Palma, G., Marvian, M., Trevisan, D., Lloyd, S.: The quantum Wasserstein distance of order 1. IEEE Trans. Inf. Theory 67, 6627–6643 (2021). https://doi.org/10.1109/TIT.2021.3076442

21. De Palma, G., Trevisan, D.: Quantum optimal transport with quantum channels. Ann. Henri Poincaré 22, 3199–3234 (2021)

22. Duvenhage, R.: Quadratic Wasserstein metrics for von Neumann algebras via transport plans. J. Operator Theory 88, 289–308 (2022)

23. Filipiak, K., Klein, D., Vojtková, E.: The properties of partial trace and block trace operators of partitioned matrices. Electron. J. Linear Algebra 33, 3–15 (2018)

24. Flamary, R., Cuturi, M., Courty, N., Rakotomamonjy, A.: Wasserstein discriminant analysis. Mach. Learn. 107, 1923–1945 (2018)

25. Friedland, S.: Matrices: Algebra, Analysis and Applications, p. 596. World Scientific, Singapore (2016)

26. Friedland, S.: Notes on semidefinite programming, Fall 2017, http://homepages.math.uic.edu/~friedlan/SDPNov17.pdf

27. Friedland, S., Eckstein, M., Cole, S., Życzkowski, K.: Quantum Monge-Kantorovich problem and transport distance between density matrices. Phys. Rev. Lett. 129, 110402 (2022)

28. Friedland, S., Ge, J., Zhi, L.: Quantum Strassen’s theorem. Infin. Dimens. Anal. Quantum Probab. Relat. Top. 23, 2050020 (2020)

29. Friedland, S., Ge, J., Zhi, L.: Quantum Fisher information matrix and multiparameter estimation. J. Phys. A 53, 045001 (2020)

30. Friesecke, G., Vögler, D.: Breaking the curse of dimension in multi-marginal Kantorovich optimal transport on finite state spaces. SIAM J. Math. Anal. 50(4), 3996–4019 (2018)

31. Gilchrist, A., Langford, N.K., Nielsen, M.A.: Distance measures to compare real and ideal quantum processes. Phys. Rev. A 71, 062310 (2005)

32. Golse, F., Mouhot, C., Paul, T.: On the mean field and classical limits of quantum mechanics. Commun. Math. Phys. 343, 165–205 (2016)

33. Golse, F., Paul, T.: Wave packets and the quadratic Monge-Kantorovich distance in quantum mechanics. Comptes Rendus Math. 356, 177–197 (2018)

34. Hitchcock, F.L.: The distribution of a product from several sources to numerous localities. J. Math. Phys. Mass. Inst. Tech. 20, 224–230 (1941)

35. Horn, R.A., Johnson, C.R.: Matrix Analysis, 2nd edn. Cambridge University Press, Cambridge (2013)

36. Horodecki, M., Horodecki, P., Horodecki, R.: Separability of mixed states: necessary and sufficient conditions. Phys. Lett. A 223, 1–8 (1996)

37. Ikeda, K.: Foundation of quantum optimal transport and applications. Quantum Inform. Process. 19, 25 (2020)

38. Jozsa, R.: Fidelity for mixed quantum states. J. Mod. Opt. 41, 2315–23 (1994)

39. Kantorovich, L.V.: Mathematical methods of organizing and planning production. Manag. Sci. 6, 366–422 (1959/60)

40. Keyl, M.: Fundamentals of quantum information theory. Phys. Rep. 369, 431–548 (2002)

41. Kiani, B.T., De Palma, G., Marvian, M., Liu, Z.-W., Lloyd, S.: Learning quantum data with the quantum earth mover’s distance. Quantum Sci. Technol. 7, 045002 (2022)

42. Liu, J., Yuan, H., Lu, X.-M., Wang, X.: Quantum Fisher information matrix and multiparameter estimation. J. Phys. A 53, 023001 (2020)

43. Lloyd, J.R., Ghahramani, Z.: Statistical model criticism using kernel two sample tests. In: Proceedings of the 28th International Conference on Neural Information Processing Systems, NIPS’15, pp. 829–837. MIT Press, Cambridge (2015)

44. Lloyd, S., Weedbrook, C.: Quantum generative adversarial learning. Phys. Rev. Lett. 121, 040502 (2018)

45. Miszczak, J.A., Puchala, Z., Horodecki, P., Uhlmann, A., Życzkowski, K.: Sub- and super-fidelity as bounds for quantum fidelity. Quantum Inf. Comp. 9, 0103–0130 (2009)

46. Monge, G.: Mémoire sur la théorie des déblais et des remblais. Histoire de l’Académie Royale des Sciences de Paris, avec les Mémoires de Mathématique et de Physique pour la même année, pp. 666–704 (1781)
47. Mueller, J., Jaakkola, T.: Principal differences analysis: interpretable characterization of differences between distributions, In: Proceedings of the 28th International Conference on Neural Information Processing Systems, NIPS’15, pp. 1702–1710. MIT Press, Cambridge (2015)
48. Müller-Hermes, A.: On the monotonicity of a quantum optimal transport cost, preprint arXiv:2211.11713 (2022)
49. Panaretos, V.M., Zemel, Y.: Amplitude and phase variation of point processes. Ann. Stat. 44, 771–812 (2016)
50. Peres, A.: Separability criterion for density matrices. Phys. Rev. Lett. 77, 1413–1415 (1996)
51. Renner, R.: Quantum Information Theory, Exercise Sheet 9, http://edu.itp.phys.ethz.ch/hs15/QIT/ex09.pdf
52. Riera, M.H.: A transport approach to distances in quantum systems, Bachelor’s thesis for the degree in Physics, Universitat Autònoma de Barcelona (2018)
53. Rubner, Y., Tomasi, C., Guibas, L.J.: The earth mover’s distance as a distance for image retrieval. Int. J. Comput. Vis. 40, 99–121 (2000)
54. Solomon, J., de Goes, F., Peyré, G., Cuturi, M., Butscher, A., Nguyen, A., Du, T., Guibas, L.: Convolutional Wasserstein distances: efficient optimal transportation on geometric domains. ACM Trans. Graph. 34, 66 (2015)
55. Sandler, R., Lindenbaum, M.: Nonnegative matrix factorization with earth mover’s distance distance for image analysis. IEEE Trans. Pattern Anal. Mach. Intell. 33, 1590–1602 (2011)
56. Safránek, D.: Discontinuities of the quantum Fisher information and the Bures distance. Phys. Rev. A 95, 052320 (2017)
57. Székely, G.J., Rizzo, M.L.: Testing for equal distributions in high dimension. Inter-Stat. 11, 1–16 (2004)
58. Uhlig, A.: The ‘transition probability’ in the state space of a *-algebra. Rep. Math. Phys. 9, 273 (1976)
59. Vandenberghe, L., Boyd, S.: Semidefinite programming. SIAM Rev. 38, 49–95 (1996)
60. Vasershtein, L.N.: Markov processes over denumerable products of spaces describing large system of automata. Probl. Inf. Transmission 5, 47–52 (1969)
61. Villani, C.: Optimal Transport, Old and New, Grundlehren der Mathematischen Wissenschaften, 338. Springer, Berlin (2009)
62. Werner, R.F.: Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable mode. Phys. Rev. A 40, 4277 (1989)
63. Winter, A.: Tight uniform continuity bounds for quantum entropies: conditional entropy, relative entropy distance and energy constraints. Commun. Math. Phys. 347, 291–313 (2016)
64. Wolfram Research, Inc., Mathematica, Version 12.2, Champaign, IL, USA (2020), https://www.wolfram.com/mathematica
65. Zhou, L., Yu, N., Ying, S., Ying, M.: Quantum earth mover’s distance, no-go quantum Kantorovich-Rubinstein theorem, and quantum marginal problem. J. Math. Phys. 63, 102201 (2022)
66. Życzkowski, K., Słomczyński, W.: Monge distance between quantum states. J. Phys. A 31, 9095–9104 (1998)
67. Życzkowski, K., Słomczyński, W.: The Monge distance on the sphere and geometry of quantum states. J. Phys. A 34, 6689 (2001)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.