Monotonicity of a quantum 2-Wasserstein distance

R Bistroń\textsuperscript{1,*}, M Eckstein\textsuperscript{1} and K Życzkowski\textsuperscript{1,2}

\textsuperscript{1} Institute of Theoretical Physics, Jagiellonian University, ul. Łojasiewicza 11, 30–348 Kraków, Poland
\textsuperscript{2} Center for Theoretical Physics, Polish Academy of Sciences, Al. Lotników 32/46, 02-668 Warszawa, Poland

E-mail: rafal.bistron@uj.edu.pl

Received 13 September 2022; revised 28 November 2022
Accepted for publication 7 February 2023
Published 20 February 2023

Abstract
We study a quantum analogue of the 2-Wasserstein distance as a measure of proximity on the set $\Omega_N$ of density matrices of dimension $N$. We show that such (semi-)distances do not induce Riemannian metrics on the tangent bundle of $\Omega_N$ and are typically not unitarily invariant. Nevertheless, we prove that for $N = 2$ dimensional Hilbert space the quantum 2-Wasserstein distance (unique up to rescaling) is monotonous with respect to any single-qubit quantum operation and the solution of the quantum transport problem is essentially unique. Furthermore, for any $N \geq 3$ and the quantum cost matrix proportional to a projector we demonstrate the monotonicity under arbitrary mixed unitary channels. Finally, we provide numerical evidence which allows us to conjecture that the unitary invariant quantum 2-Wasserstein semi-distance is monotonous with respect to all CPTP maps for dimension $N = 3$ and 4.

Keywords: quantum transport problem, Monge–Kantorovich distance, monotonicity with respect to quantum channels

(Some figures may appear in colour only in the online journal)
1. Introduction

The distances on the space of probability vectors are a primary tool to compare and analyse various statistical distributions. On the mathematical side, they lie at the heart of information geometry [1, 2], which unveils a powerful interplay between Riemannian geometry and statistics. In particular, with some distances on the space of probability vectors one can associate a Riemannian metric on the suitable tangent bundle. In the set of all such distances the one generated by the Fischer–Rao metric is distinguished as the unique continuous distance monotone under classical stochastic maps (Cencov theorem [3]).

In the modern field of quantum information [4, 5] classical probability vectors are replaced by density matrices. The quantum analogue of a stochastic map is a completely positive trace preserving (CPTP) linear map, \( \Phi : \Omega_N \to \Omega_M \), where \( \Omega_N := \{ \rho \in \mathbb{C}^{N \times N} \mid \rho = \rho^*, \rho \geq 0, \text{Tr}\rho = 1 \} \) is the set of density matrices of order \( N \). A semi-distance \( D_{\text{mon}} \) (i.e. positive, symmetric, non-degenerate function) on the space of quantum states is called monotone if the inequality

\[
D_{\text{mon}}(\Phi(\rho^A), \Phi(\rho^B)) \leq D_{\text{mon}}(\rho^A, \rho^B)
\]

holds for all states \( \rho^A, \rho^B \in \Omega_N \), for any dimension \( N \), and all CPTP maps \( \Phi \).

Some of the commonly used distances on the set of quantum states are based on the Uhlmann–Jozsa fidelity \( F \) [6, 7]. The latter is reverse monotone [7], i.e. \( F(\Phi(\rho^A), \Phi(\rho^B)) \geq F(\rho^A, \rho^B) \), for all CPTP maps \( \Phi \). Consequently, any distance on the space of quantum state determined by a strictly decreasing function \( h : [0, 1] \to \mathbb{R}^+ \), \( D_h, \rho^A(\rho^A, \rho^B) = h(F(\rho^A, \rho^B)) \), will automatically be monotone. The prominent examples include the root infidelity, \( I := \sqrt{1-F} \), the Bures distance, \( B := \sqrt{2(1-\sqrt{F})} \) and the Bures angle, \( A := \frac{1}{2} \arccos \sqrt{F} \).

All these three distances are Riemannian, i.e. they generate Riemannian metrics on the tangent bundle of \( \Omega_N \). More generally, the set of all monotone Riemannian metrics on the tangent bundle of \( \Omega_N \), can be characterised, via the Morozova–Cencov–Petz theorem [8, 9], through a single-parameter operator-monotone function. On the other hand, there exist important monotone distances, which do not induce a Riemannian metric—most notably, the trace distance \( D_{\text{tr}}(\rho^A, \rho^B) = \text{Tr}(\rho^A - \rho^B) \), with \( |X| = \sqrt{XX^*} \).

The property of monotonicity is crucial for applications in quantum information processing. It implies that the distinguishability, quantified by a given distance function, cannot increase under an application of any quantum map.

First attempts to generalise the classical Monge transport distance to the quantum setting with the use of Husimi function were done in [10, 11]. More recently, a more general approach of Kantorovich [12] and Wasserstein [13] was extended to the quantum case an analysed from the mathematical [14–18], physical [19–23] and information-theoretical [24–27] perspective.

Among these, a recent proposal [28, 29] is based on the optimization over the set of bipartite quantum states with fixed marginals of the expectation value of a certain ‘cost observable’ (see also [30–32]). Concretely, for any classical distance \( d \) on the set of \( N \) points one can define [29] an associated quantum cost operator

\[
C^Q_E = \sum_{j>i=1}^N E_{ij} |\psi_{ij}^\perp \rangle \langle \psi_{ij}^\perp|,
\]

where \( E_{ij} = d(x_i, x_j) \) is the distance between points \( x_i, x_j \), while \( |\psi_{ij}^\perp \rangle \langle \psi_{ij}^\perp| \) is the projector on the antisymmetric subspace spanned by two base vectors \( |i \rangle, |j \rangle \), i.e. \( |\psi_{ij}^\perp \rangle = \frac{1}{\sqrt{2}} (|i \rangle |j \rangle - |j \rangle |i \rangle) \). For any two density matrices \( \rho^A, \rho^B \in \Omega_N \) one introduces the set \( \Gamma^Q(\rho^A, \rho^B) \) of bi-partite coupling
matrices $\rho^{AB}$, such that both partial traces are fixed, $\text{Tr}_A \rho^{AB} = \rho^A$ and $\text{Tr}_B \rho^{AB} = \rho^B$. Minimising the Hilbert–Schmidt scalar product of the cost matrix and $\rho^{AB}$ over all possible coupling matrices we arrive at the optimal quantum transport cost,

$$T^Q_E (\rho^A, \rho^B) := \min_{\rho^{AB} \in \Gamma^Q (\rho^A, \rho^B)} \text{Tr} \left( C^Q \rho^{AB} \right).$$

(1.3)

It was shown in [29] that $T^Q_E$ is a weak metric (i.e. a semi-metric bounded from below by a genuine metric) on $\Omega_N$ for any $N$ and any quantum cost matrix $C^Q$. Following the classical analogy, one can thus define [28] the corresponding quantum 2-Wasserstein semi-distance

$$W_E (\rho^A, \rho^B) := \sqrt{\min_{\rho^{AB} \in \Gamma^Q (\rho^A, \rho^B)} \text{Tr} \left( (C^Q)^2 \rho^{AB} \right)}.$$

(1.4)

The optimal quantum transport cost $T^Q_E$ is continuous on $\Omega_N \times \Omega_N$, and hence so is the quantum 2-Wasserstein semi-distance $W_E$ [29].

On the 2-point set all classical distance matrices $E$ are equivalent up to rescaling. In higher dimensions a distinguished role is played by the simplex geometry, for which $E_{ij} = 1 - \delta_{ij}$. In this case, the corresponding quantum cost matrix [1.2] forms a projector onto the antisymmetric subspace of the Hilbert space $\mathcal{C}^N \otimes \mathcal{C}^N$. For sake of brevity we shall denote such a quantum cost matrix by

$$C^Q := \frac{1}{2} (I_N - S), \quad \text{where} \quad S(|x\rangle\langle y|) = |y\rangle\langle y| |x\rangle$$

(1.5)

and, since $(C^Q)^2 = C^Q$, write

$$W (\rho^A, \rho^B) := \sqrt{C^Q (\rho^A, \rho^B)} \quad \text{with} \quad T^Q (\rho^A, \rho^B) := \min_{\rho^{AB} \in \Gamma^Q (\rho^A, \rho^B)} \text{Tr} (C^Q \rho^{AB}).$$

(1.6)

The quantum optimal transport with the specific cost matrix [1.5] was studied in [30–33]. With the transport cost $T^Q$ one can associate the SWAP-fidelity [28],

$$F_S (\rho^A, \rho^B) := \max_{\rho^{AB} \in \Gamma^Q (\rho^A, \rho^B)} \left( \text{Tr} \rho^{AB} \right) = 1 - 2 T^Q (\rho^A, \rho^B).$$

(1.7)

It shares many properties with the standard Uhlmann–Jozsa quantum fidelity $F$ and equals to the latter if either of the states $\rho^A, \rho^B$ is pure. Hence, one could conjecture that $F_S$ is also reverse monotone. Clearly, the monotonicity of $T^Q$ under CPTP maps is equivalent to the reverse monotonicity of the SWAP-fidelity.

For single-qubit states, $N = 2$, it was shown [28, 29] that $W$ enjoys the triangle inequality and hence is a genuine distance on the Bloch ball $\Omega_2$. Furthermore, the numerical simulations strongly suggest that the triangle inequality actually holds for the quantum 2-Wasserstein semi-distance [1.6] with the cost matrix [1.5] in any dimension $N$. However, in the problem of monotonicity the triangle inequality does not play any role, hence we shall—for simplicity—focus on the general semi-distances $T^Q_E$ and the specific case of $T^Q$ determined by the projection cost matrix [1.5].

The quantum optimal transport problem [1.3] admits a dual formulation [29, 30]. The optimization takes place over the set of pairs $\sigma^A, \sigma^B \in H_N$ of $N \times N$ Hermitian matrices, satisfying a certain algebraic constraint determined by the cost matrix $C_E$. Concretely, let

$$\Sigma_N := \{ \sigma^A, \sigma^B \in H_N \mid F := C_E - \sigma^A \otimes I_N - I_N \otimes \sigma^B \succeq 0 \},$$

(1.8)

then

$$T^Q_E (\rho^A, \rho^B) = \sup_{(\sigma^A, \sigma^B) \in \Sigma_N} \text{Tr} (\sigma^A \rho^A + \sigma^B \rho^B).$$

(1.9)
If neither of the states $\rho_A, \rho_B$ is pure, then the supremum is achieved [29].

The classical Wasserstein distances have multifarious applications in physics and computer science. Recently, the field of machine learning greatly benefited from the introduction of the Wasserstein Generative Adversarial Networks [34]. A natural use of quantum Wasserstein (semi-)distances is to quantify the proximity between quantum states [28]. As it is the case of their classical counterparts, they found applications in the emerging field of quantum machine learning [27, 30, 35, 36]. In particular, the quantum 2-Wasserstein semi-distance (1.6) offers advantages in the quantum Generative Adversarial Network scheme [28, 30], which can be employed to efficiently approximate complex quantum circuits [30]. It is also worth noting that the quantum optimal transport cost (1.3) can be efficiently computed using semidefinite programming [29].

From the viewpoint of applications, the monotonicity of a distance is a desirable property, which ensures robustness against the noise. On the other hand, in the context of quantum machine learning it might be desirable to use distances, which are not unitary invariant [25, 27].

The primary purpose of this work is to study the monotonicity with respect to the quantum channels of the optimal quantum transport cost $T_Q^E$ defined above. We start in section 2 by considering some general properties of (1.3) and show that for $N \geq 3$ and a general classical cost matrix $E$ the corresponding quantum transport cost is not unitarily invariant. Then we demonstrate, in section 3, that $T_Q^E$ does not induce a Riemannian metric, hence the standard Morozova–Cencov–Petz theorem does not apply. Section 4 includes a proof of monotonicity of $T^Q$ under general mixed unitary channels in any dimension $N$. Then, in section 5 we focus on $N = 2$ and provide a complete proof of monotonicity of $T^Q$ for arbitrary single-qubits channels. Finally, in section 6, we show that for any two mixed non-isospectral qubits there exists a unique optimal coupling $\rho^{AB}$ yielding the minimum in (1.3). Furthermore, we provide explicit formulae for the optimal coupling, and the optimal dual observables $\sigma^A$, $\sigma^B$, for isospectral or commuting qubits.

The article is supplemented by two appendices. Appendix A contains the proofs of technical results, while in appendix B we provide substantial numerical evidence for the monotonicity of $T^Q$ under all qutrit and ququart channels. Based on these observations, we are tempted to conjecture that the unitary invariant quantum 2-Wasserstein semi-distance $W$ is actually monotone for all CPTP maps, at least in dimensions $N \leq 4$.

2. No unitary invariance for general quantum 2-Wasserstein semi-distances

In the single-qubit case, $N = 2$, the only (up to a trivial multiplicative factor) quantum cost matrix (1.2) is a projector in the antisymmetric subspace (1.5). But already in the qutrit case, $N = 3$, there exists different cost matrices $C_Q^E$, for instance the one induced by the classical Euclidean distance on the line, $E_{12} = E_{23} = 1, E_{13} = 2$ (see the Supplemental Material in [28]).

For the projection matrix $C_Q^O$ the optimal quantum transport cost (1.3) is invariant under unitary channels:

$$T^Q(\rho A^\dagger, \rho B^\dagger) = T^Q(\rho^A, \rho^B), \quad \text{for any} \quad U \in U(N).$$

This stems from the fact that $(U \otimes U)C_Q^O(U^\dagger \otimes U^\dagger) = C_Q^O$, for any unitary matrix $U \in U(N)$. 
The latter property, however, does not hold for a general quantum cost matrix (1.2). This implies that one cannot expect the monotonicity to hold in full generality for $T_E^Q$, even under unitary channels. Indeed, let us take two qutrit states

$$\rho^A = \frac{1}{2} (|1\rangle + |2\rangle) (|1\rangle + \langle 2|), \quad \rho^B = \frac{1}{3} (|1\rangle + 2|2\rangle) (|1\rangle + 2\langle 2|)$$

and the quantum cost matrix induced by the line geometry,

$$C_E^Q = |\psi_{13}^+\rangle\langle\psi_{13}^+| + |\psi_{23}^+\rangle\langle\psi_{23}^+| + 2|\psi_{13}^-\rangle\langle\psi_{13}^-|.$$ 

Consider now a unitary channel, which interchanges the states $|2\rangle$ and $|3\rangle$,

$$U = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$$

and yields

$$\eta^A = U\rho^A U^\dagger = \frac{1}{2} (|1\rangle + |3\rangle) (|1\rangle + \langle 3|), \quad \eta^B = U\rho^B U^\dagger = \frac{1}{3} (|1\rangle + 2|3\rangle) (|1\rangle + 2\langle 3|).$$

Since all of the involved states are pure, there is only one coupling matrix for both pairs [29, Lemma A.3], $\rho^{AB} = \rho^A \otimes \rho^B$ and $\eta^{AB} = \eta^A \otimes \eta^B$. We thus have

$$T_E^Q(U\rho^A U^\dagger, U\rho^B U^\dagger) = \text{Tr}C_E^Q\eta^{AB} = \frac{1}{10} > T_E^Q(\rho^A, \rho^B) = \text{Tr}C_E^Q\rho^{AB} = \frac{1}{20}.$$

Let us note that the lack of unitary invariance of a distance on the space of quantum states can be a desirable property in certain applications. In particular, the quantum Wasserstein distance of order 1 proposed in [25], which is not unitarily invariant, offers improved efficiency in quantum learning algorithms [27].

3. The lack of Riemannian structure

A semi-distance $d : \Omega_N \times \Omega_N \rightarrow \mathbb{R}^+$ generates a Riemannian metric $g$ on the tangent bundle $T_p\Omega_N$ if the expansion (see e.g. [37] and [11])

$$d(\rho, \rho + tv)^2 = g_p(v, v) t^2 + o(t^2), \text{as } t \rightarrow 0^+,$$ (3.1)

holds for any $\rho \in \Omega_N$ and any $v \in T_p\Omega_N = \{ v \in H_N, \text{with Tr} v = 0 \}$.

All monotone Riemannian metrics on the set of quantum states, are characterised via the Morozova–Cencov–Petz theorem [8, 9]. The latter gives an explicit formula for the metric $g$ in terms of a single operator monotone function. However, there are important distances, which do not generate a Riemannian metric, but are nevertheless monotone. A classical example is the $l_1$-distance (the ‘taxicab distance’), which corresponds to the trace distance in the quantum case. It turns out that the quantum 2-Wasserstein distance (1.6) shares this feature with the trace distance.

Let us focus on the case $N = 2$ and show the failure of (3.1) for the quantum 2-Wasserstein distance (1.6). Because of the unitary invariance of $W$ we can restrict ourselves to the states in the real slice of the Bloch ball, $\Omega_2^\mathbb{R} \subset \Omega_2$,

$$\rho(r, \theta) = \frac{1}{2} [1 + (2r - 1)(\sigma_1 \sin \theta + \sigma_3 \cos \theta)], \quad \text{with } r \in [0, 1], \theta \in [0, \pi).$$ (3.2)
Figure 1. The plot of $h(r, v_2) := G(r, (1, v_2)) - g_{11}(r) - g_{22}(r)v_2^2$ as a function of $v_2$ for $r = 1/3$, $r = 1/4$ and $r = 1/5$ — colours orange, red and purple, respectively. If Formula (3.4) would be true then, for any $r$, $h(r, v_2)$ would be a linear function of $v_2$, which clearly is not the case.

where $\sigma_i$ denote the Pauli matrices. Consequently, $T_{\rho(r, 0)}^{\mathbb{R}^2}\Omega^2$ is a real vector space diffeomorphic to $\mathbb{R}^2$. Furthermore, by unitary invariance, it is sufficient to consider the tangent space at the point $\rho(r, 0)$. Basing on the results of [29] we can compute the square of the metric derivative (see e.g. [38] for a precise definition) of the quantum 2-Wasserstein distance $W$ on $\Omega_2$ in the direction of a vector $v \in T_{\rho(r, 0)}^{\mathbb{R}^2}\Omega^2$.

**Proposition 3.1.** For any $r \in (0, 1)$ and any tangent vector $v = (v_1, v_2) \in T_{\rho(r, 0)}^{\mathbb{R}^2}\Omega^2$ we have

$$G(r, v) := \lim_{t \rightarrow 0^+} \frac{T^0(\rho(r, 0), \rho(r + tv_1, tv_2))}{t^2} = \max_{\phi \in [0, 2\pi]} \frac{(2v_1 \cos(\phi) - (2r - 1)v_2 \sin(\phi))^2}{16(1 + (2r - 1)\cos(\phi))}.$$  

(3.3)

**Proof.** The proof, which bases on a semi-analytical formula derived in [29], can be found in appendix A. \qed

If $W$ would generate a Riemannian metric $g$ on $T\Omega_2$, then formula (3.1) would imply that $g$ can be recovered from the equality

$$G(r, v) = g_{\rho(r, 0)}(v, v) = g_{11}(r)v_1^2 + 2g_{12}(r)v_1v_2 + g_{22}(r)v_2^2,$$  

(3.4)

which should be valid for all $r \in (0, 1)$ and $v \in \mathbb{R}^2$. The functions $g_{11}$ and $g_{22}$ can be computed explicitly from the analytic formulae for the 2-Wasserstein distance between commuting and isospectral qubits, (A.2) and (A.3), respectively. On the other hand, one can convince oneself with the help of formula (3.3) that $g_{12}$ depends not only on $r$, but also on the tangent vector $v$—see figure 1. Hence, formula (3.4) fails and the square of the metric derivative (3.3) of the 2-Wasserstein distance does not induce a Riemannian metric on $T\Omega_2$.

Finally, let us note that the failure of formula (3.4) extends to the general case of $\Omega_N$ and any quantum cost matrix $C^Q_{\rho}$. This is because $\Omega_2$ is embedded isometrically in $\Omega_N$ for any $N \geq 3$ and the restriction of any $C^Q_{\rho}$ to a suitable four-dimensional subspace of $\mathbb{C}^{N \times N}$ is proportional to $C^Q$ — see [29, proposition 2.4].
4. Monotonicity under general mixed unitary channels

Recall that the CPTP maps can be conveniently characterised with the help of Kraus operators. For any channel \( \Phi : \Omega_N \to \Omega_M \) there exists a set of \( N \times M \) matrices \( \{ K_i \}_{i=1}^R \), with \( R \leq NM \), called Kraus operators, such that

\[
\sum_i K_i^\dagger K_i = I_N
\]  

and for any \( \rho \in \Omega_N \)

\[
\Phi(\rho) = \sum_i K_i \rho K_i^\dagger. \tag{4.2}
\]

Conversely, quantum channels can be constructed through Kraus operators, since every set of \( N \times M \) matrices satisfying (4.1) defines a CPTP map via (4.2).

A particular class of quantum channels arises from the statistical mixtures of unitary maps. In such a case, we have \( K_i = \sqrt{p_i} U_i \), where \( U_i \in U(N) \) and \( p_i \geq 0 \), \( \sum p_i = 1 \). The general properties of the projective quantum transport cost imply the following result:

**Proposition 4.1.** Let \( \Psi \) be a mixed unitary channel on \( \Omega_N \), then for any \( \rho^A, \rho^B \in \Omega_N \) we have

\[
T_O^Q(\Psi(\rho^A), \Psi(\rho^B)) \leq T_O^Q(\rho^A, \rho^B).
\]

**Proof.** This follows directly from the joint convexity of the transport cost [29, proposition 2.2] and the unitary invariance of \( T_O^Q \).

The convexity of the optimal quantum transport cost \( T_O^Q \) implies a more general result, which holds for any quantum cost matrix (1.2).

**Proposition 4.2.** Let \( \Phi_1, \ldots, \Phi_R \) be quantum channels under which the optimal quantum transport cost is monotonous:

\[
T_O^Q(\Phi_i(\rho^A), \Phi_i(\rho^B)) \leq T_O^Q(\rho^A, \rho^B), \quad \text{for all } \rho^A, \rho^B \in \Omega_N \text{ and } i \in \{1, \ldots, R\}.
\]

Then, \( T_O^Q \) is monotonous under any convex combination of the channels \( \Phi_1, \ldots, \Phi_R \).

\[
T_O^Q(\Psi(\rho^A), \Psi(\rho^B)) \leq T_O^Q(\rho^A, \rho^B), \quad \text{for all } \Psi = \sum_i p_i \Phi_i, \text{ with } p_i \geq 0, \sum p_i = 1.
\]

The last result implies that it is sufficient to study the monotonicity of the optimal quantum transport under the extremal channels, i.e. the ones which cannot be decomposed as a convex combination of CPTP maps.

We can also invoke the dual formulation (1.9) to derive the following useful result:

**Proposition 4.3.** Let \( \Phi : \Omega_N \to \Omega_M \) be a CPTP map characterised by Kraus operators \( \{ K_i \} \) and denote by \( \Phi^* : H_M \to H_N \) the map dual to \( \Phi \), i.e. \( \Phi^*(\sigma) = \sum_i K_i^\dagger \sigma K_i \). If for any pair of matrices \( (\sigma^A, \sigma^B) \in \Sigma_M \), with \( \Sigma_M \) defined in (1.8), we have \( (\Phi^*(\sigma^A), \Phi^*(\sigma^B)) \in \Sigma_N \), then \( T_E^Q \) is monotone with respect to the map \( \Phi \), \( T_E^Q(\Phi(\rho^A), \Phi(\rho^B)) \leq T_E^Q(\rho^A, \rho^B) \).

**Proof.** Clearly, if \( \sigma \) is a Hermitian matrix, then so is \( \Phi^*(\sigma) \). From the dual formulation of the optimal quantum transport problem (1.9), we deduce
The optimal quantum transport cost is monotonous under all CPTP maps. This fact is shown in the course of the proof of theorem 5.1 in \[ \text{[Equation]} \]

Thanks to the unitary invariance of the transport cost \( \Phi \) channels \( \Omega \) and let \( \sigma \) such that \( \| \Phi(\sigma^A) \| \leq 1 \). Furthermore, the matrix \( M \) can be diagonalised with the help of two unitary rotations \( U, V \in U(2) \) as follows,

\[
\Phi(\rho) = U[\Phi_d(V\rho V^\dagger)] U^\dagger, \tag{5.1}
\]

where \( \Phi_d \) is a quantum channel with a diagonal matrix \( M \).

In \[ \text{[Equation]} \] it was shown that any such single-qubit map \( \Phi_d \), which belongs to the closure of the set of extremal quantum channels, can be realised with two Kraus operators

\[
K_1 = \begin{bmatrix} \cos \left( \frac{1}{2} v \right) \cos \left( \frac{1}{2} u \right) \\ \sin \left( \frac{1}{2} v \right) \sin \left( \frac{1}{2} u \right) \end{bmatrix} I + \begin{bmatrix} \sin \left( \frac{1}{2} v \right) \sin \left( \frac{1}{2} u \right) \\ \cos \left( \frac{1}{2} v \right) \cos \left( \frac{1}{2} u \right) \end{bmatrix} \sigma_3, \tag{5.2}
K_2 = \begin{bmatrix} \sin \left( \frac{1}{2} v \right) \cos \left( \frac{1}{2} u \right) \\ \cos \left( \frac{1}{2} v \right) \sin \left( \frac{1}{2} u \right) \end{bmatrix} \sigma_1 + i \begin{bmatrix} \cos \left( \frac{1}{2} v \right) \sin \left( \frac{1}{2} u \right) \\ \sin \left( \frac{1}{2} v \right) \cos \left( \frac{1}{2} u \right) \end{bmatrix} \sigma_2,
\]

with \( u \in [0, 2\pi) \) and \( v \in [0, \pi) \). For such quantum channels we have the following technical result proven in appendix A.

**Lemma 5.1.** Let \( \Phi : \Omega_2 \to \Omega_2 \) be a CPTP map determined by Kraus operators (5.2) and assume that \( \sigma^A, \sigma^B \in \mathbb{H}_2 \) are diagonal. If \( F = C^0 - \sigma^A \otimes I_N - I_N \otimes \sigma^B \geq 0 \) then also \( F^B = C^0 - \Phi^*(\sigma^A) \otimes I_N - I_N \otimes \Phi^*(\sigma^B) \geq 0 \).

It turns out that the matrices \( \sigma^A, \sigma^B \) yielding the maximum in (1.9) actually commute and thus can be simultaneously diagonalised.

**Lemma 5.2.** Let \( \rho^A, \rho^B \in \Omega_2 \) be mixed quantum states of full rank and with different spectra, and let \( \sigma^A, \sigma^B \) be two Hermitian matrices saturating the dual problem (1.9), i.e.

\[
T^Q(\rho^A, \rho^B) = \text{Tr}(\sigma^A \rho^A + \sigma^B \rho^B).
\]

Then, \( \sigma^A \) and \( \sigma^B \) commute.

**Proof.** This fact is shown in the course of the proof of theorem 5.1 in [29].

We are now in a position to present the main result of this paper.

**Theorem 5.3.** The optimal quantum transport cost is monotonous under all CPTP maps \( \Phi : \Omega_2 \to \Omega_2 \).

**Proof.** Thanks to the unitary invariance of the transport cost \( T^Q \) we can restrict ourselves to channels \( \Phi_d \) defined by equation (5.1). Furthermore, by proposition 4.2, it is sufficient to check
the monotoncity for extremal single-qubit channels. Such channels are completely characterised by two Kraus operators (5.2), with \( u \in [0, 2\pi) \), \( v \in [0, \pi) \) such that either \( \sin u \cdot \sin v \neq 0 \) or \( |\cos u \cdot \cos v| < 1 \) (see theorem 1.4 in [40]). We can thus use the characterisation (5.2) and apply lemma 5.2.

Now, assume that the states \( \rho^A, \rho^B \in \Omega_2 \) are non-isospectral and neither of them is pure. By lemma 5.2 we know that the Hermitian matrices saturating the dual problem, \( \sigma^A, \sigma^B \), commute. Let then \( V \in U(2) \) be the matrix diagonalizing simultaneously \( \sigma^A \) and \( \sigma^B \). Using the unitary invariance of \( T^Q \) we obtain

\[
T^Q(\Phi(\rho^A), \Phi(\rho^B)) = T^Q\left(V^\dagger \Phi(\rho^A)V, V^\dagger \Phi(\rho^B)V\right)
\]

\[
= \max_{(\sigma^A, \sigma^B) \in \Sigma_2} \text{Tr}(\sigma^A V^\dagger \Phi(\rho^A)V + \sigma^B V^\dagger \Phi(\rho^B)V)
\]

\[
= \text{Tr}(\sigma^A V^\dagger \Phi(\rho^A)V + \sigma^B V^\dagger \Phi(\rho^B)V)
\]

\[
= \text{Tr}(\Phi^*(V\sigma^AV^\dagger)\rho^A + \Phi^*(V\sigma^BV^\dagger)\rho^B).
\]

Now, lemma 5.1 implies that \((\Phi^*(V\sigma^AV^\dagger), \Phi^*(V\sigma^BV^\dagger)) \in \Sigma_2\), hence we conclude that

\[
T^Q(\Phi(\rho^A), \Phi(\rho^B)) = \text{Tr}(\Phi^*(V\sigma^AV^\dagger)\rho^A + \Phi^*(V\sigma^BV^\dagger)\rho^B)
\]

\[
\leq \sup_{(\rho^A, \rho^B) \in \Sigma_2} \text{Tr}(\sigma^A \rho^A + \sigma^B \rho^B) = T^Q(\rho^A, \rho^B).
\]

This inequality extends by the continuity of \( T^Q \) and \( \Phi \) to the limiting cases of isospectral and pure states \( \rho^A, \rho^B \).

**6. Solutions to the single-qubit transport problem**

In this section, we study the solutions to original (1.3) and dual (1.9) optimal quantum transport problems for single-qubit states. We start with the general discussion of the uniqueness of solutions, and then present explicit solutions to the quantum optimal transport in the special cases of commuting and isospectral qubits.

**6.1. The uniqueness of solutions**

Let us start with some general remarks, which apply in any dimension \( N \). While the original problem always has at least one solution yielding the minimum in (1.3), the supremum in the dual problem (1.9) is attained only if both of the states of the states \( \rho^A, \rho^B \) are positive definite (see [29, theorem 3.2]). On the other hand, if at least one of the states is pure, then there exists only one coupling, \( \rho^{AB} = \rho^A \otimes \rho^B \) — see [29, lemma A.3]. Note also that the solution of the dual problem (1.9) is always given up to a shift: \((\sigma^A, \sigma^B) \rightarrow (\sigma^A - cI_{2^2}, \sigma^B + cI_{2^2})\), where \( c \in \mathbb{R} \) is an arbitrary constant. Such a shift changes neither the value of the transport cost nor the matrix \( F := C_F - \sigma^A \otimes I_N - I_N \otimes \sigma^B \) defined in (1.8). Consequently, by a unique solution to the dual problem we shall always mean the uniqueness of the matrix \( F \).

We now show the uniqueness of the solution to the optimal quantum transport for nonisospectral mixed qubits 3. To this end, we first need the following result.

**Lemma 6.1.** Let \( \rho^A, \rho^B \in \Omega_2 \) be not isospectral. For such states, let \( \rho^{AB} \in \Gamma(\rho^A, \rho^B) \) be an optimal coupling yielding the minimum in (1.3) and let \( F \) be the optimal matrix solving the dual problem (1.9). Then, \( \text{rank } \rho^{AB} = \text{rank } F = 2 \).

3 We are indebted to Shmuel Friedland for providing us arguments leading to a shorter proof of this result.
This fact is shown in the course of the proof of theorem 5.1 in [29].

Theorem 6.2. Let $\rho^{A}, \rho^{B} \in \Omega_{2}$ be two non-isospectral mixed density matrices. Then, there exists a unique solution to the original and the dual quantum transport problem.

Proof. Assume that both $\rho^{A}, \rho^{B}$ are of full rank. Lemma 6.1 shows that any optimal coupling $\rho_{1}^{AB}$ has rank two. Suppose that $\rho_{1}^{AB}$ and $\rho_{2}^{AB}$ are two different optimal solutions to the original problem. Suppose first that $\rho_{1}^{AB}$ and $\rho_{2}^{AB}$ have different supports. Then, for any $a \in (0, 1)$, the state $\rho_{a}^{AB} := a\rho_{1}^{AB} + (1 - a)\rho_{2}^{AB}$ is an optimal coupling for $\rho^{A}, \rho^{B}$. But rank $\rho_{a}^{AB} \geq 3$, which contradicts lemma 6.1. Suppose then that $\rho_{1}^{AB}$ and $\rho_{2}^{AB}$ are both supported by the same two-dimensional subspace. Now, set $\rho_{t}^{AB} := -t\rho_{2}^{AB} + (1 + t)\rho_{1}^{AB}$ for any $t > 0$. Any such $\rho_{t}^{AB}$ is an optimal coupling for $\rho^{A}, \rho^{B}$. But since the set of density matrices is bounded, there exists a maximal $t$ such that for each $t' > t$ the coupling $\rho_{t'}^{AB}$ has one negative eigenvalue. Hence $\rho_{t}^{AB}$ has only one nonzero eigenvalue, which contradicts lemma 6.1.

Along the same lines one shows the uniqueness of the optimal matrix $E$ solving the dual problem. The only difference appears in the last part, because the set $\Sigma_{N}$ is not bounded, but it is a subset of a positive cone as $F \geq 0$. Suppose that $F_{1}, F_{2}$ are two solutions of the dual problem supported by the same two dimensional subspace. Then, there always exists a maximal $t'$ such that either $F_{1} = tF_{2} + (1 - t)F_{2}'$ or $F_{1} = -tF_{1} + (1 + t)F_{2}'$ is outside of the positive cone for each $t > t'$. Hence, either $F_{1}'$ or $F_{2}'$ is a solution of the dual problem with rank smaller than 2, which contradicts lemma 6.1.

6.2. Commuting states

We shall now present an explicit formula for the optimal coupling of two commuting qubits. Using the Bloch vectors,

$$\rho^{A} = \frac{1}{2}(1 + \vec{a} \cdot \vec{\sigma}), \quad \rho^{B} = \frac{1}{2}(1 + \vec{b} \cdot \vec{\sigma}),$$

it is convenient to write an element of $\Gamma(\rho^{A}, \rho^{B})$ in the Fano form [41]

$$\rho^{AB} = \frac{1}{4} \left( I_{4} + \sum_{i=1}^{3} a_{i} \sigma_{i} \otimes I_{2} + \sum_{i=1}^{3} b_{i} I_{2} \otimes \sigma_{i} + \sum_{i,j=1}^{3} R_{ij} \sigma_{i} \otimes \sigma_{j} \right).$$

It is straightforward to show that the quantum transport cost depends only on the trace of the correlation matrix,

$$\text{Tr} \left[ C^{O} \rho^{AB} \right] = \frac{1}{4} \left( 1 - \text{Tr} R \right).$$

Note however that the requirement of positive semi-definiteness of $\rho^{AB}$ induces constraints on the matrix $R$ — see appendix A.3. Consequently, the minimisation of (6.3) is performed over a complicated convex subset of the 9-dimensional parameter space of $R$.

If $\rho^{A}$ and $\rho^{B}$ commute then their Bloch vectors can be aligned via a unitary transformation and hence we can assume that only $a_{1}$ and $b_{3}$ are non-zero. We claim that the correlation matrix $R$ of the optimal coupling $\rho^{AB}$ is diagonal with entries

$$R_{11} = R_{22} = \frac{1}{2} \sqrt{(a_{3} + b_{3})^{2} - (a_{3} - b_{3})^{2}}, \quad R_{33} = |a_{3} + b_{3}| - 1.$$ 

Indeed, by inserting (6.4) to formula (6.3), for $a_{3} + b_{3} > 0$ one gets

$$\text{Tr} \left[ C^{O} \rho^{AB} \right] = \frac{1}{4} \left( \sqrt{1 - a_{3}} - \sqrt{1 - b_{3}} \right)^{2}.$$
and in the case of \( a_3 + b_3 < 0 \) one gets
\[
\text{Tr} \left[ C_{AB}^0 \right] = \frac{1}{4} \left( \sqrt{1 + a_3} - \sqrt{1 + b_3} \right)^2.
\]
Equivalently, we can write
\[
\text{Tr} \left[ C_{AB}^0 \right] = \frac{1}{4} \max \left( \left( \sqrt{1 - a_3} - \sqrt{1 - b_3} \right)^2, \left( \sqrt{1 + a_3} - \sqrt{1 + b_3} \right)^2 \right),
\]
which agrees with equation (A.2) derived in [28]. The optimal coupling for commuting qubit states is of rank 2, as follows from lemma 6.1, and its nonzero eigenvalues are \( \frac{1}{2} |a_3 + b_3| \) and \( 1 - \frac{1}{2} |a_3 + b_3| \).

Given the optimal coupling, \( \rho_{AB} \) one can compute the optimal observables, \( \sigma^A \) and \( \sigma^B \), saturating the dual problem (1.9). To do so one may use theorem 3.2 from [29] which states that for any optimal coupling \( \rho_{AB} \) and any optimal \( F \) the following equality holds:
\[
\text{Tr} \left[ F \rho_{AB} \right] = 0. \tag{6.5}
\]

Thanks to the non-negativity of \( \rho_{AB} \) and \( F \) this is equivalent to the statement that \( \rho_{AB} \) and \( F \) are supported on the orthogonal subspaces of the Hilbert space \( \mathbb{C}^d \). By inserting the solution of the original problem (6.4) and a general form of \( F (1.8) \) into (6.5) we obtained a system of linear equations with the unique solution:

if \( a_3 + b_3 > 0 \),
\[
\sigma^A = \frac{1}{8} \begin{pmatrix}
\frac{|a_3 - b_3|}{\sqrt{1 - a_3} \sqrt{1 + a_3}} & 0 & 4 - \frac{2(2 - a_3 - b_3) + |a_3 - b_3|}{\sqrt{1 - a_3} \sqrt{1 + b_3}} \\
0 & \sqrt{1 - a_3} & 4 - \frac{2(2 - a_3 - b_3) - |a_3 - b_3|}{\sqrt{1 - a_3} \sqrt{1 + b_3}} \\
0 & 4 - \frac{2(2 - a_3 - b_3) - |a_3 - b_3|}{\sqrt{1 - a_3} \sqrt{1 + b_3}} & \sqrt{1 + a_3}
\end{pmatrix}, \tag{6.6}
\]

if \( a_3 + b_3 < 0 \),
\[
\sigma^A = \frac{1}{8} \begin{pmatrix}
4 - \frac{2(2 + a_3 + b_3) - |a_3 - b_3|}{\sqrt{1 + a_3} \sqrt{1 + b_3}} & 0 & \frac{-|a_3 - b_3|}{\sqrt{1 + a_3} \sqrt{1 + b_3}} \\
0 & \sqrt{1 + a_3} & \frac{-|a_3 - b_3|}{\sqrt{1 + a_3} \sqrt{1 + b_3}} \\
0 & \frac{-|a_3 - b_3|}{\sqrt{1 + a_3} \sqrt{1 + b_3}} & \sqrt{1 + a_3}
\end{pmatrix}, \tag{6.7}
\]

6.3. Isospectral states

Now, let us discuss the case of isospectral states. We start by finding an exemplary solution to the quantum transport problem between the isospectral states \( \rho^A, \rho^B = U \rho^A U^\dagger \). By unitary invariance of transport cost we may assume that state \( \rho^A \) is diagonal, and state \( \rho^B \) is real, so \( U = \exp(i \sigma y / 2) \).

We use the following anzatz for the coupling matrix:
\[
\rho_{AB} = (I_2 \otimes U) \rho^A (I_2 \otimes U^\dagger), \tag{6.8}
\]
where \( \rho^A = \rho^A \otimes I_2 + I_2 \otimes \rho^A + \sum_{i=1}^{3} R_i \rho_{ij} \otimes \sigma_i \) is some coupling between \( \rho^A \) and itself with a diagonal correlation matrix \( R \). Therefore the matrix of correlation coefficients \( R \) for \( \rho_{AB} \) has
the form $R = R' O'$, where $O$ is the rotation matrix changing the Bloch vector $\vec{a}$ into the Bloch vector $\vec{b}$. Moreover, since $\rho_{AB}$ and $\rho_{AA}$ are connected by unitary map $\rho_{AB} \geq 0 \iff \rho_{AA} \geq 0$.

The transport cost is obtained by minimizing the formula:

$$\text{Tr} C Q \rho_{AB} = \frac{1}{8} \left( 1 + \sqrt{2} R'_{11} + R'_{33} \right) (1 - \cos(\phi)),$$

with $R'_{ij} = (R'_{11} - R'_{22}) / \sqrt{2}$. The minimization of (6.9) over $R'_{33}$ and $R'_{ij}$ under the constraints imposed by non-negativity of $\rho_{AB}$ (see equation (A.16) in the appendix) leads to:

$$T_{CQ} = \frac{1}{2} \left( 1 - \sqrt{1 - a_3^2} \right) \sin(\phi/2)^2,$$

which agrees with equation (A.3) obtained in [28]. The correlation matrix $R$ determining the optimal coupling $\rho_{AB}$ has the following form:

$$R = \begin{pmatrix}
-\sqrt{1 - a_3^2} \cos(\phi) & 0 & \sqrt{1 - a_3^2} \sin(\phi) \\
0 & \sqrt{1 - a_3^2} & 0 \\
\sin(\phi) & 0 & \cos(\phi)
\end{pmatrix}.$$

(6.10)

The optimal coupling $\rho_{AB}$ has rank one and thus can be written as $\rho_{AB} = |\psi\rangle \langle \psi|$ with

$$|\psi\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix}
\sqrt{1 + a_3 \cos(\phi/2)} \\
\sqrt{1 + a_3 \sin(\phi/2)} \\
\sqrt{1 - a_3 \sin(\phi/2)} \\
-\sqrt{1 - a_3 \cos(\phi/2)}
\end{pmatrix}.$$

Since the construction of solutions for the dual problem, based on orthogonality of $F$ and $\rho_{AB}$ supports, is straightforward but tedious, we present it in appendix A.4. We found a one-parameter family of solutions for the dual problem, such that for each $\phi$ there exist at least one optimal $F$ with rank 3. Equation (6.5) and the fact that both $F$ and $\rho_{AB}$ are positive semi-definite, together imply that for each pair of isospectral states $\rho^A$ and $\rho^B$, there exist only one optimal $\rho_{AB} = |\psi\rangle \langle \psi|$, with $|\psi\rangle$ being the unique eigenvector to 0 eigenvalue of above mentioned optimal $F$. Hence, the original quantum optimal transport problem possesses only one solution, determined by the correlation matrix (6.10).

Note added in the proof: In a recent preprint [42] Müller-Hermes provided an example of a quantum channel in dimension $N = 8$, for which the monotonicity of the quantum 2-Wasserstein semi-distance (1.6) fails.

Data availability statement

No new data were created or analysed in this study.

Acknowledgment

It is a pleasure to thank Shmuel Friedland and Tomasz Miller for their helpful remarks. We also thank the anonymous Referees for their pertinent comments. Financial support by Narodowe Centrum Nauki under the Maestro grant number DEC- 2015/18/A/ST2/00274 and by Foundation for Polish Science under the Teams-Net Project No. POIR.04.00.00-17C1/18-00 is gratefully acknowledged.
Appendix A. Proofs of technical results

A.1. Semi-analytic formulae for $T^Q$ for $N = 2$

In this appendix we present the proof of proposition 3.1, along with the semi-analytic and analytic formulae for the optimal quantum transport cost for qubits, which we invoked in section 3.

We first recall a semi-analytic formula for $T^Q$ between two single-qubit states in the Bloch parametrisation (3.2) derived in [29]:

$$T^Q(\rho(s,0), \rho(r,\theta)) = \max_{\phi \in [0, 2\pi]} \frac{1}{4} \left( \sqrt{1 + (2s - 1) \cos \phi} - \sqrt{1 + (2r - 1) \cos (\theta + \phi)} \right)^2. \quad (A.1)$$

For the special case of commuting or isospectral states one can derive [28] explicit analytic formulæ:

$$T^Q(\rho(s,0), \rho(r,0)) = \frac{1}{2} \max \left\{ \left( \sqrt{r} - \sqrt{s} \right)^2, \left( \sqrt{1-r} - \sqrt{1-s} \right)^2 \right\}, \quad (A.2)$$

$$T^Q(\rho(r,0), \rho(r,\theta)) = \left( \frac{1}{2} - \sqrt{r(1-r)} \right) \sin^2(\theta/2). \quad (A.3)$$

We shall now prove proposition (3.1) via the following lemma:

**Lemma A.1.** For any $r \in (0, 1)$ and any tangent vector $v = (v_1, v_2) \in T_{\rho(r,0)} \Omega_2^R$ we have

$$T^Q(\rho(r,0), \rho(r + v_1t, v_2t)) = G(r,v)t^2 + O(t^3), \quad \text{as } t \downarrow 0, \quad (A.4)$$

where the function $G$ is defined by formula (3.3),

$$G(r,v) = \max_{\phi \in [0, 2\pi]} \frac{2v_1 \cos(\phi) - 2r(2r - 1)v_2 \sin(\phi)}{16(1 + (2r - 1) \cos(\phi))}.$$

**Proof.** In order to slightly simplify the notation we set $\xi := 2r - 1 \in (-1, 1)$ and assume that $\xi \neq 0$, i.e., $r \neq 1/2$. Note that the maximally mixed state $\rho(1/2, 0)$ commutes with any other state $\rho^\delta$ in the Bloch ball, hence one can use Formula (A.2) to compute the value of $T^Q(\rho(1/2, 0), \rho^\delta)$ (cf equation (12) in [28]).

Let us start with writing, for sufficiently small $t > 0$,

$$W(\rho(r,0), \rho(r + v_1t, v_2t)) = \frac{1}{2} \max_{\phi \in [0, 2\pi]} \left| \sqrt{1 + \xi \cos \phi} - \sqrt{1 + (\xi + 2v_1 t) \cos(\phi + v_2 t)} \right|$$

$$\leq \sqrt{G(r,v)} t + \frac{1}{2} \max_{\phi \in [0, 2\pi]} R(\xi, v, \phi, t), \quad (A.5)$$

where

$$R(\xi, v, \phi, t): = \left| \sqrt{1 + (\xi + 2v_1 t) \cos(\phi + v_2 t)} - \sqrt{1 + \xi \cos \phi} \right.$$  
$$- \frac{2v_1 t \cos \phi - \xi v_2 t \sin \phi}{2\sqrt{1 + \xi \cos \phi}}. \quad (A.6)$$

Moreover, the 2-Wasserstein distance can be also bounded from below, for sufficiently small $t > 0$, by

$$W(\rho(r,0), \rho(r + v_1t, v_2t)) \geq \sqrt{G(r,v)} t - \frac{1}{2} \max_{\phi \in [0, 2\pi]} R(\xi, v, \phi, t). \quad (A.7)$$

Our goal is to establish a uniform bound on $R(\xi, v, \phi, t)$, which is at least quadratic in $t$ and does not depend on $\phi$. 

13
Let us identify the three terms in Formula (A.6),

\[ A := 1 + (\xi + 2v_t)\cos(\phi + v_2 t), \quad B := 1 + \xi \cos \phi, \]

\[ C := 2v_t \cos \phi - \xi v_2 \sin \phi. \]

Then, we have

\[
R = \left| \sqrt{A} - \sqrt{B} - \frac{C}{2\sqrt{B}} \right| = \frac{|A - B - C|}{\sqrt{A + \sqrt{B}}} + \frac{C(B - A)}{2\sqrt{B}(\sqrt{A} + \sqrt{B})^2}
\]

\[
\leq \frac{|A - B - C|}{\sqrt{B}} + \frac{|C| \cdot |A - B|}{2B^{3/2}}.
\]

The most obvious estimates are

\[ B \geq 1 - |\xi|, \quad |C| \leq (2|v_1| + |\xi v_2|) t. \]

In order to establish suitable bounds on the remaining terms we will use the sum-to-product trigonometric identities [43]. We have

\[
|A - B - C| = \left| (\xi + 2v_t)\left[ \cos(\phi + v_2 t) - \cos \phi \right] + \xi v_2 \sin \phi \right|
\]

\[
= \left| 2(\xi + 2v_t) \sin \left( \phi + \frac{v_2 t}{2} \right) \sin \left( \frac{v_2 t}{2} \right) - \xi v_2 \sin \phi \right|
\]

\[
\leq 2|\xi| \cdot \left| \sin \left( \phi + \frac{v_2 t}{2} \right) \sin \left( \frac{v_2 t}{2} \right) - \frac{1}{2} \xi v_2 \sin \phi \right| + 4t|v_1| \cdot \left| \sin \left( \frac{v_2 t}{2} \right) \sin \phi \right|
\]

\[
\leq 2|\xi| \cdot \left| \left| \sin \left( \phi + \frac{v_2 t}{2} \right) \sin \left( \frac{v_2 t}{2} \right) - \frac{1}{2} \xi v_2 \sin \phi \right| + 2|v_1| v_2 |r^2
\]

\[
\leq 2|\xi| \cdot \left| \sin \left( \frac{v_2 t}{2} \right) \right| \cdot \left| \cos \left( \phi + \frac{v_2 t}{4} \right) \right| \cdot \left| v_2 |r + \frac{|v_2|^3 r^3}{48} \right| + 2|v_1| v_2 |r^2
\]

\[
\leq \frac{1}{2} \left( |\xi v_2|^2 + 4|v_1 v_2| \right) t^2 + \frac{1}{24} |v_2|^3 t^3.
\]

Similarly, we obtain

\[
|A - B| = \left| \xi \left[ \cos(\phi + v_2 t) - \cos \phi \right] + 2v_1 t \cos(\phi + v_2 t) \right|
\]

\[
\leq 2|\xi| \cdot \left| \sin \left( \phi + \frac{v_2 t}{2} \right) \right| \cdot \left| \sin \left( \frac{v_2 t}{2} \right) \right| + 2t|v_1| v_2 |r^2
\]

\[
\leq \left( |\xi v_2| + 2|v_1| \right) t.
\]

Summa summarum, we arrive at the following estimate

\[
R(\xi, v, \phi, t) \leq \left( \frac{|\xi v_2|^2 + 4|v_1 v_2|}{2 \sqrt{1 - |\xi|}} + \frac{(|\xi v_2| + 2|v_1|)^2}{(1 - |\xi|)^{3/2}} \right) t^2 + \frac{|v_2|^3}{24 \sqrt{1 - |\xi|}} t^3.
\]

This provides the desired uniform bound, which we can insert into Formula (A.5) and conclude that

\[
W(\rho(r, 0), \rho(r + v_1 t, v_2 t)) \leq \sqrt{G(r, \nu) t} + c_1 t^2 + c_2 t^3.
\]

The square of the above equality implies equation (A.4).
One can show (cf [29, appendix B]) that the angle \( \phi \) yielding the maximum in equation (3.3) defining the function \( G \) can be obtained from the solution to a quartic equation. For the specific cases of two commuting or isospectrual qubits we have
\[
G((r, 1.0)) = \frac{1}{8} \max \left\{ \frac{1}{1-r}, \frac{1}{r} \right\}, G((r, 0, 1)) = \frac{1}{8} \left( 1 - 2 \sqrt{r(1-r)} \right)
\]  
(A.8)

Equipped with these formulae one can reproduce figure 1 and convince oneself that equation (3.4) fails.

### A.2. Proof of lemma 5.1

In this appendix we present the proof of lemma 5.1, which is the key to theorem 5.3. Let us fix a single-qubit channel \( \Phi \) determined by two Kraus operators (5.2). The channel, as well as its dual, \( \Phi^*(\sigma) = \sum_{i=1,2} K_i^\dagger \sigma K_i \), is parametrised by two angles \( u \in [0, 2\pi) \) and \( v \in [0, \pi) \).

Lemma 5.1 claims that if \( \sigma^A \) and \( \sigma^B \) are two diagonal Hermitian \( 2 \times 2 \) matrices, then the positivity of the operator \( F = C^0 - \sigma^A \otimes I - I \otimes \sigma^B \), appearing in the dual optimisation problem (1.9), implies the positivity of the operator \( F^A = C^0 - \Phi^* (\sigma^A) \otimes I - I \otimes \Phi^* (\sigma^A) \), for the extremal single-qubit channel \( \Phi \) determined by equation (5.2).

Since both matrices \( \sigma^A \) and \( \sigma^B \) are diagonal, we have four independent real parameters. Note, however, that a gauge transformation \( \sigma^A \mapsto \sigma^A + \alpha I, \sigma^B \mapsto \sigma^B - \alpha I \), with any \( \alpha \in \mathbb{R} \), does not affect neither the operator \( F \), nor the value of the optimal quantum transport cost (1.9).

Consequently, we can eliminate one of these four parameters. It is convenient to use the following parametrisation:
\[
c := \text{Tr}(\sigma^A \sigma^3), \quad d := \text{Tr}(\sigma^B \sigma^3), \quad x := \text{Tr}(\sigma^A + \sigma^B),
\]
where \( \sigma_3 \) denotes the third Pauli matrix. Then, both hermitian matrices \( F \) and \( F^\Phi \) take the following form
\[
F = \frac{1}{2} \begin{pmatrix}
F_{11} & 0 & 0 & 0 \\
0 & F_{22} & F_{23} & 0 \\
0 & F_{23} & F_{33} & 0 \\
0 & 0 & 0 & F_{44}
\end{pmatrix}, \quad F^\Phi = \frac{1}{2} \begin{pmatrix}
F^\Phi_{11} & 0 & 0 & 0 \\
0 & F^\Phi_{22} & F^\Phi_{23} & 0 \\
0 & F^\Phi_{23} & F^\Phi_{33} & 0 \\
0 & 0 & 0 & F^\Phi_{44}
\end{pmatrix}, \quad (A.9)
\]
with entries:
\[
F_{11} = -(c + d + x), \quad F_{22} = 1 - c + d - x, \quad F_{23} = F_{32} = -1, \quad (A.10)
\]
\[
F_{33} = 1 + c - d - x, \quad F_{44} = c + d - x,
\]
\[
F^\Phi_{11} = -(c + d) \cos(u - v) - x, \quad F^\Phi_{22} = 1 - c \cos(u - v) + d \cos(u + v) - x, \quad F^\Phi_{23} = F^\Phi_{32} = -1, \quad (A.11)
\]
\[
F^\Phi_{33} = 1 + c \cos(u + v) - d \cos(u - v) - x,
\]
\[
F^\Phi_{44} = (c + d) \cos(u + v) - x.
\]

The demand \( F \geq 0 \) implies, in particular, the non-negativity of the diagonal elements of \( F \), which yields the following constraints on the parameters,
\[
x \leq 0, \quad x \leq c + d \leq -x, \quad x - 1 \leq c - d \leq 1 - x. \quad (A.12)
\]

The non-negativity of the central \( 2 \times 2 \) minor of \( F \), which is the only non-trivial one, gives an additional constraint
\[
(c - d)^2 \leq x(x - 2). \quad (A.13)
\]
To simplify notation let us set
\[ \alpha := \cos(u - v) \quad \text{and} \quad \beta := \cos(u + v). \]

We start with showing that the diagonal elements of \( F^\Phi \) are non-negative. If \( \alpha \in [0, 1] \) then (A.12) yields
\[ F^\Phi_{11} = -(c + d) \alpha - x \geq x(\alpha - 1) \geq 0, \]
while if \( \alpha \in [-1, 0] \) then we obtain
\[ F^\Phi_{11} \geq -x(\alpha + 1) \geq 0. \]
Analogously, we deduce that \( F^\Phi_{44} \geq 0. \)

Let us now rewrite
\[ F^\Phi_{22} = 1 - x - \frac{1}{2}(c + d)(\alpha - \beta) + \frac{1}{2}(d - c)(\alpha + \beta) \]
and assume that \( (\alpha - \beta) \geq 0 \) and \( (\alpha + \beta) \geq 0 \). Then, inequalities (A.12) yield,
\[ F^\Phi_{22} \geq 1 - x + \frac{1}{2}x(\alpha - \beta) + \frac{1}{2}(x - 1)(\alpha + \beta) = \left[ 1 - \frac{1}{2}(\alpha + \beta) \right] + \frac{1}{2}x(\alpha - 1) \geq 0, \]
because \( \alpha \in [0, 1] \) and \( (\alpha + \beta) \in [0, 2] \). Analogously, one shows that \( F^\Phi_{33} \geq 0 \) under three other possible assumptions about the signs of \( \alpha - \beta \) and \( \alpha + \beta \). Along the same lines, one can prove that \( F^\Phi_{33} \geq 0 \).

Finally, let us consider the central \( 2 \times 2 \) minor of \( F^\Phi \). With \( \pm \) denoting the sign of \( \alpha + \beta \) and using the previous estimates on \( F^\Phi_{22} \) and \( F^\Phi_{33} \), along with constraint (A.13), we can write the determinant of this minor as
\[
F^\Phi_{22} F^\Phi_{33} - 1 = \left[ 1 - x - \frac{1}{2}(c + d)(\alpha - \beta) - \frac{1}{2}(c - d)(\alpha + \beta) \right] \times \\
\times \left[ 1 - x - \frac{1}{2}(c + d)(\alpha - \beta) + \frac{1}{2}(c - d)(\alpha + \beta) \right] - 1 \geq 0.
\]

The last inequality holds because \( \alpha, \beta \in [-1, 1] \), while \( x \leq 0 \) from constraints (A.12).

Since the central minor of \( F^\Phi \) is the only non-trivial one we conclude that, indeed, if \( F \) is positive semidefinite, then \( F^\Phi \) is so, for any quantum channel determined by Kraus operators (5.2).

### A.3. Conditions for positive semi-definiteness in Fano form

In this appendix, we derive a useful form of the constraint for the coupling matrix between two qubits to be positive semidefinite, using its Fano form [41]. The obtained results are based on general formulas from [44].
The density matrix representing any quantum state must be Hermitian, positive semidefinite and of trace 1. The first and the last property can be expressed easily using parametrisation via generators of the $su(n)$ Lie algebras, i.e. Hermitian traceless matrices. This leads to the Fano representation of a bipartite quantum state $\rho^{AB}$, which for two qubits takes the following form:

$$\rho^{AB} = \frac{1}{4} \left( I_4 + \sqrt{6} \vec{n} \cdot \vec{\lambda} \right),$$

where we temporarily adapt the normalization form [44], and use the basis of Hermitian operators on $\mathbb{C}^{2 \times 2}$ defined by the tensor product of Pauli operators:

$$\lambda_i, \ i = 1, 2, 3 \quad \leftrightarrow \quad \frac{1}{\sqrt{2}} \sigma_j \otimes I_2,$$

$$\lambda_i, \ i = 4, 5, 6 \quad \leftrightarrow \quad \frac{1}{\sqrt{2}} I_2 \otimes \sigma_j,$$

$$\lambda_i, \ i = 7, 8, 9 \quad \leftrightarrow \quad \frac{1}{\sqrt{2}} \sigma_1 \otimes \sigma_j,$$

$$\lambda_i, \ i = 10, 11, 12 \quad \leftrightarrow \quad \frac{1}{\sqrt{2}} \sigma_5 \otimes \sigma_j,$$

$$\lambda_i, \ i = 13, 14, 15 \quad \leftrightarrow \quad \frac{1}{\sqrt{2}} \sigma_3 \otimes \sigma_j,$$

with $j = 1, 2, 3$ and the vector $\vec{n}$ has the form:

$$\vec{n} = (a_1, a_2, a_3, b_1, b_2, b_3, R_{11}, R_{12}, R_{13}, R_{21}, R_{22}, R_{23}, R_{31}, R_{32}, R_{33}) .$$

Here $\vec{a}, \vec{b}$ are proportional to the Bloch vectors for consecutive subsystems and the matrix $R$ is defined in equation (6.2). The density matrix $\rho^{AB}$ has non-negative eigenvalues if and only if all coefficients in its characteristic polynomial are non-negative [44],

$$\det[\rho^{AB} - \lambda I] = \lambda^4 - S_1 \lambda^3 + S_2 \lambda^2 - S_3 \lambda + S_4 = 0 .$$

In [44] it was shown that these coefficients can be written as:

$$S_1 = 1 ,$$

$$S_2 = \frac{3}{8} (1 - \vec{n} \cdot \vec{n}) ,$$

$$S_3 = \frac{1}{16} (1 - 3 \vec{n} \cdot \vec{n} + 2 \vec{n} \cdot (\vec{n} \star \vec{n})) ,$$

$$S_4 = \frac{1}{64} (1 - 6 \vec{n} \cdot \vec{n} + 8 \vec{n} \cdot (\vec{n} \star \vec{n}) + 9 (\vec{n} \star \vec{n})^2 - 12 (\vec{n} \star \vec{n}) \cdot (\vec{n} \star \vec{n}) ) ,$$

where $\star$ denotes the product defined [44] via symmetric structure constants $d_{ijk}$ of the Lie algebra $su(4)$,

$$(\vec{a} \star \vec{b})_k = \frac{\sqrt{6}}{2} d_{ijk} a_i b_j .$$

The expression $\vec{n} \cdot (\vec{n} \star \vec{n})$ has a compact form

$$\vec{n} \cdot (\vec{n} \star \vec{n}) = 3 \sqrt{3} \left( \vec{a}^T R \vec{b} - \det(R) \right) ,$$

while the expression $(\vec{n} \star \vec{n}) \cdot (\vec{n} \star \vec{n})$ can be simplified to
\[(\vec{n} \times \vec{n}) \cdot (\vec{n} \times \vec{n}) = 3 \left[ b_1 R^T R b_1 + a_1^T R R^T a_1 + (\det R)^2 \sum_{i,j} (R^{-1})_{ij}^2 - 2(\det R) b_1 R^{-1} a_1 + a_1^T b_1^2 \right],\]

under the assumption that the matrix \( R \) is invertible. Hence the coefficients \( S_1, S_2, S_3, S_4 \) entering (A.14) can be rewritten as:

\[
S_1 = 1,
S_2 = \frac{3}{8} \left( 1 - \sum_{i,j} (R_{ij})^2 - a_i^2 - b_j^2 \right),
S_3 = \frac{1}{16} \left[ 1 - 3 \left( \sum_{i,j} (R_{ij})^2 + a_i^2 + b_j^2 \right) + 6 \sqrt{3} \left( a_i^T R b_j - (\det R) \right) \right],
S_4 = \frac{1}{64} \left[ 1 - 6 \left( \sum_{i,j} (R_{ij})^2 + a_i^2 + b_j^2 \right) + 24 \sqrt{3} \left( a_i^T R b_j - (\det R) \right) + 9 \left( \sum_{i,j} (R_{ij})^2 + a_i^2 + b_j^2 \right)^2 - 36 \left( b_1 R^T R b_1 + a_1^T R R^T a_1 + (\det R)^2 \sum_{i,j} (R^{-1})_{ij}^2 - 2(\det R) b_1 R^{-1} a_1 + a_1^T b_1^2 \right) \right].
\]

The necessary and sufficient conditions for \( \rho^{AB} \) to be a density matrix come down to

\[
0 \leq S_i \quad \text{for} \quad i = 1, 2, 3, 4.
\]

We end this section by discussing the special case of \( \rho^A = \rho^B \), i.e. \( \vec{a} = \vec{b} = (0, 0, a_3) \), and a diagonal correlation matrix \( R = \text{diag}(R_{11}, R_{22}, R_{33}) \). Adopting the normalization from (6.2), the expression form (A.15) for \( S_3 \) and \( S_4 \) simplifies to the product of two planes and a hyperbola:

\[
S_3 = (R_{33} - 1) \left( 2a_3^2 - R_{33} + R_{-}^2 \right) - (R_{33} + 1)R_{+}^2 \geq 0,
S_4 = (1 - R_{33} + \sqrt{2}R_+)(1 - R_{33} - \sqrt{2}R_+) \left( 4a_3^2 + 2R_-^2 - (1 + R_{33})^2 \right) \geq 0,
\]

where \( R_{\pm} = (R_{11} \pm R_{22})/\sqrt{2} \).

A.4. Solutions of dual quantum transport problem for isospectral qubit states

In this appendix we present the construction of a one parametric family of solutions to dual quantum optimal transport problem for isospectral qubit states. Afterwards, we use those solutions to show that the solution of the original transport problem (6.10) is unique.

For isospectral states \( \rho^A, \rho^B \) the optimal coupling \( \rho^{AB} \), given by (6.10), has only one nonzero eigenvalue, which demonstrates that theorem 6.2 cannot hold and we can expect multiple solutions both to the original and the dual problem.

The construction of the solutions to the dual problem for the isospectral states is based on the fact that any optimal coupling \( \rho^{AB} \) has disjoint support with any optimal \( \rho \), as follows from
equation (6.5). Hence, by the eigendecomposition of $\rho^{AB}$ we know that the optimal $F$ must have the form

$$F = \sum_{i,j} v_i^j s_{ij},$$

(A.17)

where $v_i$ are eigenvectors to zero eigenvalues of $\rho^{AB}$ and $s_{ij} = s_{ji}^*$ are some a priori unknown coefficients. On the other hand, the general form of $F$ is given by

$$F = \left( \frac{1}{4} - x_0 \right) I_4 - \sum_{i=1}^{3} c_i \sigma_i \otimes I_2 - \sum_{i=1}^{3} d_i I_2 \otimes \sigma_i - \frac{1}{4} \sum_{i=1}^{3} \sigma_i \otimes \sigma_i,$$

(A.18)

where $c_i = \frac{1}{4} \text{Tr}[\sigma_i^4], d_i = \frac{1}{2} \text{Tr}[\sigma_i^4], x_0 = \text{Tr}[\sigma_i^4] + \text{Tr}[\sigma_i^3]$, and $\sigma_i$ are Pauli matrices.

Now, we have to compare (A.17) with (A.18) to eliminate as many coefficients as possible and then check when the resulting matrix $F$ is positive semidefinite. Elementwise comparison gives the following form of optimal couplings $\bar{\sigma}^A, \bar{\sigma}^B$

$$\begin{align*}
(\bar{\sigma}^A)_{11} &= \frac{a_3 (2w + 2a_3 - 1) - 2w (4a_3 d_3 + 1 - w) \sec \phi - 2(a_3 - 1)(w - 1) \cos \phi}{8a_3 w}, \\
(\bar{\sigma}^A)_{12} &= \frac{\sqrt{1 - a_3 (4a_3 d_3 + 1 - w) \tan \phi}}{4a_3 \sqrt{1 + a_3}} + \\
&\quad - \frac{(\cos \phi + 1) \tan(\phi/2) (4wa_3 d_3 - (a_3 + 1)(w - 1) \cos \phi + w - w^2)}{4a_3 (a_3 + 1) \cos \phi}, \\
(\bar{\sigma}^A)_{21} &= (\bar{\sigma}^A)_{12}, \\
(\bar{\sigma}^A)_{22} &= - (\bar{\sigma}^A)_{11} + \frac{(w - 1)(1 - \cos \phi)}{2w}, \\
\bar{\sigma}^B &= \frac{1}{4d_3} \begin{pmatrix} 4a_3 d_3 & (4a_3 d_3 - w + 1) \tan \phi \\ (4a_3 d_3 - w + 1) \tan \phi & -4a_3 d_3 \end{pmatrix},
\end{align*}$$

(A.19)

where $w = \sqrt{1 - a_3^2}$ and $d_3$ is the only parameter that has not been eliminated. For generic values of $d_3$ parameter, those solutions correspond to optimal $F$ which has a rank equal to three.

To ensure the positivity of $F$, we check the values of coefficients $S_1, S_2, S_3, S_4$ of the characteristic polynomial of $F$, similarly as in appendix A.3.

$$\det[F - \lambda I] = \lambda^4 - S_1 \lambda^3 + S_2 \lambda^2 - S_3 \lambda + S_4.$$  

Firstly, let us note that $S_1 = \text{Tr}[F] = (1 - 4x_0) \geq 0$ if and only if $(1 - w) \cos(\phi) \leq 1$, which is true for any values of $w = \sqrt{1 - a_3^2}$ and $\phi$. Since $F$ has at most 3 nonzero eigenvalues, the coefficient $S_4 = \det F$ must be equal to 0. The $S_2$ coefficient imposes a quadratic condition in $d_3$:

$$-32d_3^2 a_3^3 w^3 \sec^2(\phi) - 8d_3 a_3 w^2 (\sec(\phi - 1)(2w(1 - w) \sec(\phi - 1) - (1 - w)^2) \cos \phi + \sec^2(\phi)w(1 - w)^2 (2(1 - w)w) - a_3^2 \cos 3\phi + 2(1 + w)^2 \cos 2\phi + (w(4 + 3w) - 3) \cos \phi) \geq 0,$$

(A.20)

and $S_3$ imposes a quadratic condition in $d_3$ as well

$$8d_3 a_3^3 \sin^2(\phi/2) \sec^2(\phi) + \sin^4(\phi/2) \sec^2(\phi)w(1 - w)^2 (w - \cos(\phi)) + 4d_3 a_3 \sin^4(\phi/2) \sec(\phi)w(1 - w)(w \sec(\phi) + 1) - 1 \leq 0.$$  

(A.21)
For both these quadratic inequalities, the discriminants are greater than 0. Therefore, there exist intervals of $d_3$ satisfying each of them separately. Moreover, the roots of (A.20) lie between the roots of (A.21). Hence, introducing the auxiliary variables

\[ d_{3,1} = \frac{1}{16a_3w} \left[ (w - 1)(1 + 3w) - (1 - w)^2 \cos 2\phi + 2a_3 \cos \phi \left( a_3 - \sqrt{\frac{2}{a_3}} \sin (\phi/2) \right) \right] \]
\[ d_{3,2} = \frac{1}{16a_3w} \left[ (w - 1)(1 + 3w) - (1 - w)^2 \cos 2\phi + 2a_3 \cos \phi \left( a_3 + \sqrt{\frac{2}{a_3}} \sin (\phi/2) \right) \right] \]

we obtain the following bounds for the value of $d_3$,

\[ \text{if } \phi \leq \pi/2 \quad \text{then} \quad d_{3,1} \leq d_3 \leq d_{3,2}, \]
\[ \text{if } \phi \geq \pi/2 \quad \text{then} \quad d_{3,1} \geq d_3 \geq d_{3,2}. \]

In the singular case, $\phi = \pi/2$, the entire interval of allowed values of $d_3$ shrinks to a single point:

\[ d_3 = d_{3,1} = d_{3,2} = \frac{w - 1}{4a_3} \]

resulting in only a single solution of the dual problem, with rank($F$) = 3 for any value of $a_3 \in (0,1)$.

We constructed a family of the dual problem solutions in which for any values of $\phi$ and $a_3$ there exists an optimal $F$ with rank 3. Therefore, the corresponding optimal coupling $\rho_{AB}$ is uniquely defined by the eigenvector to zero eigenvalue of the abovementioned $F$, hence the original problem has only one solution (6.10).

**Appendix B. Monotonicity: numerical results**

In this appendix the results of numerical calculations are presented. We performed Monte Carlo simulations in which we drew pairs of quantum states and a quantum channel and checked how the optimal quantum transport cost between the selected states changes after the application the channel.

We examined the monotonicity of the optimal quantum transport for qutrits ($N = 3$) and ququarts ($N = 4$) using two different types of tests. In the first one we picked random quantum channels corresponding to dynamical Choi matrices of a fixed rank. In the second one we employed random extremal quantum channels, defined up to unitary pre- and post-processing of the quantum states. Because of the unitary invariance of $T^Q$ and its convexity (recall proposition 4.2), the monotonicity of $T^Q$ under such channels implies the monotonicity with respect to all CPTP maps.

To generate a random quantum state of order $N$ we first drew a rectangular complex random matrix $X$ of size $N \times k$, with $k \leq N$. Setting $\rho = XX^\dagger / \text{Tr}[XX^\dagger]$ we ensure that a random matrix $\rho$ of size $N$ and rank $k$ is non-negative and $\text{Tr} \rho = 1$. Our calculations show that the rank of states does not affect significantly the results. Therefore we focus on full rank states to maximize the explored space.
Random quantum channels \( \Phi : \Omega_N \to \Omega_N \) of a given rank \( k \) were generated via the Choi–Jamiołkowski isomorphism. We generated a random \( N^2 \times k \) complex matrix \( X \) and created an auxiliary \( N^2 \times N^2 \) positive matrix \( Y = XX^\dagger \geq 0 \). By the Choi–Jamiołkowski isomorphism a quantum state corresponds to a quantum channel if the second of its partial traces is a totally mixed state. To assure this property we defined \( Y = \text{Tr}_1 Y \), where \( \text{Tr}_1 \) is the partial trace on the first subsystem. Then, we set \( \rho_k \propto (I_N \otimes Y^{-1/2}) (I_N \otimes Y^{-1/2}) \) with the proportionality coefficient adjusted by the condition \( \text{Tr}[\rho_k] = 1 \). For further details on random quantum channel generation we encourage the Reader to consult [45, 46].

Having a normalized density matrix \( \rho_k \), we can calculate the action of channel \( \Phi \) on any state \( \rho \) using the definition of Choi–Jamiołkowski isomorphism:

\[
N \text{Tr}_2[\rho_k (I_N \otimes \rho^T)] = N \text{Tr}_2 \left[ (\Phi \otimes I_N)((\psi^+) (\psi^+)) (I_N \otimes \rho^T) \right]
\]

\[
= \text{Tr}_2 \left[ (\Phi \otimes I_N) \left( \sum_{i=1}^{N} |ii\rangle \langle ii| \right) \left( \sum_{j=1}^{N} \langle jj| \right) \right] (I_N \otimes \rho^T)
\]

\[
= \sum_{q} \Phi(|i\rangle \langle j|) \text{Tr}[|i\rangle \langle j| \rho^T] = \Phi(\rho),
\]

where \( |\psi^+\rangle = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} |ii\rangle \) is a maximally entangled state.

In the test of the second type we focused on extremal quantum channels. Using the methods presented in [47] we described all such channels by their Kraus operators, \( K_i = U_i D_i \) with \( i = 1, \ldots, N \), where \( U_i \) and \( D_i \) denote, respectively, the unitary matrices and diagonal matrices of order \( N \). An explicit form of \( U_i \) and \( D_i \) for a qutrit reads,

\[
D_1 = \text{diag}(a, b, c), \quad D_2 = \text{diag}(d, e, f), \quad D_3 = \sqrt{1 - D_1^2 - D_2^2},
\]

\[
U_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad U_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad U_3 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.
\]

An analogous parametrisation works for \( N = 4 \) (see [47]). All of the parameters in the above formulae are non-negative and chosen so that the matrices \( D_3 \) (for the qutrit case) and \( D_4 \) (for the ququart case) are positive semi-definite.

In each Monte Carlo simulation we generated a random pair of initial states \( \rho^A, \rho^B \) of order \( N \) and a quantum channel \( \Phi \) using the methods presented above. Then, we calculated the quantum transport cost before and after the application of the channel, \( T^Q(\rho^A, \rho^B) \) and \( T^Q(\Phi(\rho^A), \Phi(\rho^B)) \), respectively and compared the results.

The simulations were programmed in the Python language, using Numpy library for algebraic calculations, Cvxpy library for the transport costs calculations and a solver supplied by the MOSEK Optimizer for solving semidefinite programming problems. The plots were generated with help of the Matplotlib library. The numerical precision of calculations was set to \( 10^{-12} \). We checked the accuracy of the optimisation algorithm by generating random unitary channels, which correspond to \( k = 1 \) in the procedure described above. The deviation between the obtained values of \( T^Q \) before and after the application of a unitary channel was not greater than \( 10^{-9} \), hence we can take this value as the numerical accuracy of the algorithm.

The results of the simulations are presented in table B1 and figure B1. In all of the considered cases the difference \( T^Q(\rho^A, \rho^B) - T^Q(\Phi(\rho^A), \Phi(\rho^B)) \) was positive and a few orders of magnitude larger than the numerical accuracy. These results allow us to conjecture that optimal quantum transport cost \( T^Q \) for dimensions 3 and 4 is monotonous under all CPTP maps.
Table B1. The results of the Monte Carlo simulations discussed in the text. ‘Samples’ refers to number of drawn triples \((\rho_A, \rho_B, \Phi)\), whereas ‘\(\min(T^Q - T^Q\circ \Phi)\)’ refers to the smallest difference \(T^Q(\rho_A, \rho_B) - T^Q(\Phi(\rho_A), \Phi(\rho_B))\) found in the samples.

| \(N = 3\), qutrit | \(N = 4\), ququart |
|-------------------|-------------------|
| Random channels   |                   |                   |
| \(\text{rank } k\) | \(\text{samples}\) | \(\text{rank } k\) | \(\text{samples}\) |
| 2                 | 300 000           | 2                 | 200 000           |
| 3                 | 300 000           | 3                 | 200 000           |
| 9                 | 600 000           | 4                 | 200 000           |
|                   |                   | 16                | 400 000           |
| \(\min(T^Q - T^Q\circ \Phi)\) | 0.001 46 | \(\min(T^Q - T^Q\circ \Phi)\) | 0.006 44 |
| Extremal channels |                   |                   |
| \(\text{samples}\) | 7 500 000         | \(\text{samples}\) | 7 500 000         |
| \(\min(T^Q - T^Q\circ \Phi)\) | 0.000 234 85 | \(\min(T^Q - T^Q\circ \Phi)\) | 0.0061 979 01 |

Figure B1. The results of the Monte Carlo simulations of random density matrices of order \(N = 3\) (top) and \(N = 4\) (bottom) discussed in the text. The optimal quantum transport cost, \(T^Q\circ \Phi\), after the action of a quantum channel \(\Phi\) is plotted against the initial value of \(T^Q\). Results obtained for channels of a fixed rank \(k\) are presented on the left and the extremal channels on the right. The dashed line marks the equality \(T^Q(\rho_A, \rho_B) = T^Q(\Phi(\rho_A), \Phi(\rho_B))\), which is saturated for unitary channels \((k = 1)\).
ORCID iDs

R Bistroń ☝️ https://orcid.org/0000-0002-0837-8644
M Eckstein ☝️ https://orcid.org/0000-0003-4343-5031
K Życzkowski ☝️ https://orcid.org/0000-0002-0653-3639

References

[1] Amari S and Nagaoka H 2000 Methods of Information Geometry (Philadelphia, PA: American Mathematical Society)
[2] Amari S 2016 Information Geometry and its Applications (Berlin: Springer)
[3] Čencov N N 1982 Statistical Decision Rules and Optimal Inference vol 53 English edn (Providence, RI: American Mathematical Society)
[4] Nielsen M and Chuang I 2010 Quantum Computation and Quantum Information (Cambridge: Cambridge University Press)
[5] Bengtsson I and Życzkowski K 2017 Geometry of Quantum States 2nd edn (Cambridge: Cambridge University Press)
[6] Uhlmann A 1976 The ‘transition probability’ in the state space of a *-algebra Rep. Math. Phys. 9 273
[7] Jozsa R 1994 Fidelity for mixed quantum states J. Mod. Opt. 41 2315–23
[8] Morozova E A and Čentsov N N 1991 Markov invariant geometry on manifolds of states J. Sov. Math. 56 2648–69
[9] Petz D 1996 Monotone metrics on matrix spaces Linear Algebr. Appl. 244 81–96
[10] Życzkowski K and Slomczyński W 1998 Monge distance between quantum states J. Phys. A 31 9095–104
[11] Życzkowski K and Slomczyński W 2001 The Monge metric on the sphere and geometry of quantum states J. Phys. A 34 6689
[12] Kantorovich L V 1948 On a problem of Monge Usp. Mat. Nauk. 3 225
[13] Villani C 2009 Optimal Transport, Old and New (Grundlehren der Mathematischen Wissenschaften vol 338) (Berlin: Springer)
[14] Agredo J and Fagnola F 2017 On quantum versions of the classical Wasserstein distance Stochastics 89 6–7
[15] Chen Y, Gangbo W, Georgiou T T and Tannenbaum A 2020 On the matrix Monge-Kantorovich problem Eur. J. Appl. Math. 31 574
[16] Bhatia R, Gaubert S and Jain T 2019 Matrix versions of the Hellinger distance Lett. Math. Phys. 109 777–804
[17] Friedland S 2020 Tensor optimal transport, distance between sets of measures and tensor scaling (arXiv:2005.00945)
[18] Duvenhage R 2021 Quadratic Wasserstein metrics for von Neumann algebras via transport plans J. Oper. Theory 88 289–308
[19] Golse F, Mouhot C and Paul T 2016 On the mean field and classical limits of quantum mechanics Commun. Math. Phys. 343 165–205
[20] Golse F and Paul T 2018 Wave packets and the quadratic Monge-Kantorovich distance in quantum mechanics C. R. Math. 356 177–97
[21] Caglioti E, Golse F and Paul T 2020 Quantum optimal transport is cheaper J. Stat. Phys. 181 149
[22] Carlen E A and Maas J 2020 Non-commutative calculus, optimal transport and functional inequalities in dissipative quantum systems J. Stat. Phys. 178 319
[23] Ikeda K 2020 Foundation of quantum optimal transport and applications Quantum Inf. Process. 19 25
[24] Chen Y, Georgiou T T, Ning L and Tannenbaum A 2017 Matricial Wasserstein-1 distance* IEEE Control Syst. Lett. 1 14
[25] De Palma G, Marvian M, Trevisan D and Lloyd S 2021 The quantum Wasserstein distance of order 1 IEEE Trans. Inf. Theory 67 6627–43
[26] De Palma G and Trevisan D 2021 Quantum optimal transport with quantum channels Ann. Henri Poincaré 22 3199–234
Kiani B T, De Palma G, Marvian M, Liu Z-W and Lloyd S 2022 Quantum earth mover’s distance: a new approach to learning quantum data Quantum Sci. Technol. 7 045002

Friedland S, Eckstein M, Cole S and Życzkowski K 2022 Quantum monge-kantorovich problem and transport distance between density matrices Phys. Rev. Lett. 129 110403

Cole S, Eckstein M, Friedland S and Życzkowski K 2021 Quantum optimal transport (arXiv:2105.06922)

Chakrabarti S, Huang Y, Li T, Feizi S and Wu X 2019 Quantum Wasserstein generative adversarial networks Advances in Neural Information Processing Systems p 3674 (https://doi.org/10.48550/arXiv.1911.00111)

Yu N, Zhou L, Ying S and Ying M 2022 Quantum earth mover’s distance, no-go quantum Kantorovich-Rubinstein theorem and quantum marginal problem J. Math. Phys. 63 102201

Reira M H 2018 A transport approach to distances in quantum systems Bachelor’s Thesis Universitat Autònoma de Barcelona

Duvenhage R 2022 Quadratic Wasserstein metrics for von Neumann algebras via transport plans J. Oper. Theory 88 289–308

Arjovsky M, Chintala S and Bottou L 2017 Wasserstein Generative Adversarial Networks Proc. Mach. Learn. Res. 70 214

Biamonte J, Witte P, Pancotti N, Rebentrost P, Wiebe N and Lloyd S 2017 Quantum machine learning Nature 549 195

Lloyd S and Weedbrook C 2018 Quantum generative adversarial learning Phys. Rev. Lett. 121 040502

Spohn D, Illuminati F, Orszag M and Roga W 2017 Geometric measures of quantum correlations with bures and hellinger distances Lectures on General Quantum Correlations and Their Applications, ed F Fanchini, D Soares Pinto and G Adesso (Cham: Springer)

Ambrosio L, Gigli N and Savaré G 2005 Gradient Flows in Metric Spaces and in the Space of Probability Measures (Birkhäuser Verlag: ETH Zürich)

King C and Ruskai M B 2001 Minimal entropy of states emerging from noisy quantum channels IEEE Trans. Inf. Theory 47 192–209

Ruskai M B, Szarek S and Werner E 2001 An analysis of completely positive trace-preserving maps on $M_2$, Linear Algebra Appl. 347 159

Fano U 1957 Description of States in Quantum Mechanics by Density Matrix and Operator Techniques Rev. Mod. Phys. 29 74

Müller-Hermes A 2022 On the monotonicity of a quantum optimal transport cost (arXiv:2211.11713v1)

Abramowitz M and Stegun I A (eds) 1972 Handbook of Mathematical Functions With Formulas, Graphs and Mathematical Tables (New York: Dover Publications)

Byrd M and Khaneja N 2003 Characterization of the positivity of the density matrix in terms of the coherence vector representation Phys. Rev. A 68 062322

Kukulski R, Nechoita I, Pawela L, Puchala Z and Życzkowski K 2021 Generating random quantum channels J. Math. Phys. 62 062201

Życzkowski K, Penson K, Nechoita I and Collins B 2011 Generating random density matrices J. Math. Phys. 52 062201

Dixit K and Sudarshan E 2009 Extremal channels for a general quantum system (arXiv:0909.3744v1)